MONASH University

The role of Hfq and sRNAs in regulation of *Pasteurella multocida* gene expression

Marianne Nelly Fleurette Mégroz

BSc (Hons)

A thesis submitted for the degree of Doctor of Philosophy at

Monash University

Department of Microbiology

Monash Biomedicine Discovery Institute

2020

Copyright notice

Table of Contents

List of Tables

Appendix B

Appendix C

List of Figures

Appendix B

Figure S1: Growth rate of *P. multocida* wild-type VP161 and the *hfq* mutant in BHI broth at 37°C with constant shaking at 200 rpm.

Appendix D

Figure S2: Electrophoretic separation of PCR products generated from analysis of the *hfq::hfq*FLAG strain AL2842.

Figure S3: Southern hybridisation analysis of *hyaD/hfq* strain AL3089.

List of Abbreviations and Symbols

%	Percent
~	Approximately
α	Alpha
β	Beta
BHI	Brain Heart Infusion
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
cDNA	Complementary DNA
CI	Competitive index
CFU	Colony Forming Unit(s)
°C	Degrees Celsius
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
ddH_2O	Double-distilled water
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
FDR	False Discovery Rate
g	Gram(s)
g	Gravitational force
h	Hour(s)
HA	Hyaluronc acid
HI	Heart Infusion
HCl	Hydrochloric acid
HRP	Horseradish peroxidase
ID	Infective dose
Kan^R	Kanamycin resistance
kb	Kilobase pairs
kDa	Kilodaltons
kPa	Kilopascal(s)
L	Litre(s)

LB	Lysogeny (Luria) broth
LPS	Lipopolysaccharide
µF	Microfarad(s)
µg	Microgram(s)
µL	Microlitre(s)
µm	Micrometre(s)
µM	Micromolar
mg	Milligram(s)
mL	Millilitre(s)
mM	Millimolar
min	Minute(s)
M	Molar
MCS	Multiple cloning site
mRNA	Messenger ribonucleic acid
Mw	Molecular weight
NA	Not applicable
ND	Not detected
ng	Nanogram(s)
nm	Nanometre(s)
Ω	Ohm
OD_{600}	Optical Density 600nm wavelength
ORF	Open reading frame
P	Probability that null hypothesis is correct
PCR	Polymerase chain reaction
pH	Potential of hydrogen
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
rpm	Revolutions per minute
RNA	Ribonucleic acid
RNA-Seq	RNA Sequencing
RT	Room temperature
SD	Standard deviation

SDS	Sodium dodecyl sulphate
$Spec^R$	Spectinomycin resistance
Tet^R	Tetracycline resistance
Tris	Tris (hydroxymethyl) aminomethane
U	Units
UV	Ultraviolet
V	Volt
v/v	Volume/volume
w/v	Weight/volume
WT	Wild-type

Summary

The Gram-negative bacterium *Pasteurella multocida* is the causative agent of a number of economically important animal diseases, including avian fowl cholera. A small number of *P. multocida* virulence factors have been identified, including capsule, lipopolysaccharide (LPS) and filamentous hemagglutinin, but little is known about how the expression of these virulence factors is regulated. Small non-coding RNA molecules (sRNAs) are important regulators of bacterial gene expression and protein production, with essential roles in controlling diverse bacterial functions including virulence, carbon, amino acid and iron metabolism, outer membrane protein production and quorum sensing. Hfq is an RNA-binding protein that facilitates riboregulation via interaction with small noncoding RNA molecules (sRNAs) and their mRNA targets.

In order to assess the role of sRNAs in *P. multocida* gene regulation, a *hfq* mutant was constructed in *P. multocida* strain VP161. The *hfq* mutant produced significantly less hyaluronic acid capsule during all growth phases and displayed reduced *in vivo* fitness and virulence compared with the wild-type strain. The *hfq* mutant also displayed global changes in gene expression and protein production, including altered expression of the key *P. multocida* virulence factors, capsule, filamentous hemagglutinin and LPS. These data indicate that Hfq and associated sRNAs are likely to be critical regulators of *P. multocida* virulence.

To further examine sRNA-mediated riboregulation in *P. multocida*, putative sRNAs were identified using bioinformatic analysis of whole genome RNA-Sequencing, using RNA isolated from cells grown in a range of growth conditions to maximise sRNA discovery. These analyses identified 38 putative sRNAs including the GcvB sRNA, known to regulate amino acid metabolism and transport, and the HrrF sRNA, known to be involved in iron homeostasis in *Haemophilus influenzae*. Most predicted sRNAs were encoded in intergenic regions, highly expressed, and contained putative Rho-independent terminators. A number of the putative sRNAs were differentially expressed under specific conditions, such as during growth in iron-limited medium or under reduced oxygen conditions, suggesting that they may have roles in regulating global responses to these changing conditions.

Hfq co-immunoprecipitation experiments, including UV-crosslinking, ligation and sequencing of hybrids (Hfq-CLASH), were then performed to identify Hfq-associated sRNAs and mRNAs. A total of 90 transcripts were identified as Hfq-associated using native Hfq Co-IP, including the sRNA GcvB that is Hfq-dependent in other species, and two sRNAs we termed Prrc10 and Prrc19.

Other transcripts that were identified as Hfq-associated from the analysis included the transcript representing outer membrane lipoprotein *plpB* and the phosphocholine (PCho) transferase *pcgD*, which both displayed increased transcription in the *hfq* mutant, suggesting a potential role for Hfq in the destabilization/degradation of these transcripts. Seventeen Hfq-associated RNA:RNA hybrids were identified using Hfq-CLASH, including one predicted sRNA:tRNA association.

Lastly, a number of putative sRNAs were selected for experimental validation using fluorescent primer extension and northern blotting. Specifically, sRNAs were chosen that were predicted to be involved in the regulation of *P. multocida* virulence phenotypes. However, although most were confirmed to be sRNA species, no clear phenotypes were identified for any of these putative sRNAs.

Publications during enrolment

Mégroz, M., Kleifeld, O., Wright, A., Powell, D., Harrison, P., Adler, B., Harper, M., Boyce, JD. (2016). The RNA-binding chaperone Hfq is an important global regulator of gene expression in *Pasteurella multocida* and plays a crucial role in production of a number of virulence factors including hyaluronic acid capsule. *Infection and Immunity* **84**: 1361-70. (Chapter 2).

Gulliver, E., Wright, A., Deveson Lucas, D., Mégroz, M., Kleifeld, O., B Schittenhelm, R., Powell, D., Seemann, T., B Bulitta, J., Harper, M., Boyce, J. (2018). Determination of the small RNA GcvB regulon in the Gram-negative bacterial pathogen *Pasteurella multocida* and identification of the GcvB seed binding region. *RNA*. **24**(5):704-720.

Conference proceedings

2018 (Oral presentation): The Australian Society for Microbiology (ASM 2018), Brisbane, Australia: "Identification of *Pasteurella multocida* small RNAs: unravelling the regulatory network"

2018 (Poster presentation): 5th Meeting of Regulation with RNA in Bacteria and Archaea, Seville, Spain: "Identification of *Pasteurella multocida* small RNAs: unravelling the regulatory network"

2018 (Poster presentation): Biomedicine Discovery Institute Graduate Student Symposium (Monash University), Melbourne, Australia: "Identification of *Pasteurella multocida* small RNAs: unravelling the regulatory network"

2018 (Poster presentation): meRNA Student Symposium (Monash University), Melbourne, Australia: "Identification of *Pasteurella multocida* small RNAs: unravelling the regulatory network"

2017 (Poster presentation): BacPath meeting, Adelaide, Australia: "Identification of *Pasteurella multocida* small RNAs: unravelling the regulatory network"

2016 (Poster presentation): Walter and Eliza Hall Institute VIIN Young Investigator Symposium, Melbourne, Australia: "Identification of *Pasteurella multocida* small RNAs: unravelling the regulatory network"

2016 (Poster presentation): Monash University Microbiology Postgraduate Society (MUMPS) Student Conference, Melbourne, Australia: "The role of Hfq and sRNAs in *Pasteurella multocida* gene expression"

2016 (Poster presentation): Walter and Eliza Hall Institute VIIN Young Investigator Symposium, Melbourne, Australia: "Inactivation of the RNA-binding chaperone Hfq in *Pasteurella multocida* results in reduced capsule expression and reduced *in vivo* fitness"

2015 (Oral presentation): BacPath meeting, Victoria, Australia: "Inactivation of the RNA-binding chaperone Hfq in *Pasteurella multocida* results in reduced capsule expression and reduced *in vivo* fitness"

Thesis including published works declaration

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes one original paper published in a peer reviewed journal. The core theme of the thesis is riboreulgation in *Pasteurella multocida*. The ideas, development and preparation of all manuscripts in the thesis were the principal responsibility of myself, the student, working within the Department of Microbiology under the supervision of Associate Professor John Boyce and Dr Marina Harper.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of Chapter 2 and Appendices A and B, my contribution to the work involved the following:

Thesis Chapter	Publication Title	Status	Nature and % of student contribution	Co-author name(s), nature and % of Co-author's contribution	Co-authors, Monash student Y/N
2	The RNA-binding chaperone Hfq is an important global regulator of gene expression in *Pasteurella multocida* and plays a crucial role in production of a number of virulence factors, including hyaluronic acid capsule	Published	60%. Laboratory work, data analysis, concept and manuscript preparation.	1) Oded Kleifeld 2.5%, proteomic experimental platform assistance and data analysis. 2) Amy Wright 5%, proteomic laboratory work and assistance, manuscript preparation. 3) David Powell 5%, bioinformatic data analysis. 4) Paul Harrison 2.5%, bioinformatic data analysis. 5) Ben Adler 5%, laboratory assistance (mouse work) manuscript preparation and experimental design. 6) Marina Harper 10%, laboratory assistance (chicken trials), manuscript preparation and experimental design. 7) John D. Boyce 10%, (chicken trials), manuscript preparation and experimental design.	1) N 2) N 3) N 4) N 5) N 6) N 7) N

In Chapter 2, proteomic sample preparation was performed by Amy Wright and mass spectrometry and initial data analysis was performed by Dr Oded Kleifeld (Monash Biomedical Proteomics Facility). Final analyses of the proteomics data were performed by this candidate, Dr David Powell and Dr Paul Harrison (Monash University Bioinformatics Platform), whom also provided assistance with analysis of RNA-Seq data. Virulence trials in chickens were performed under the guidance of Assoc. Prof. John Boyce and Dr Marina Harper. Competitive growth assays in mice were performed under the guidance of Prof. Ben Adler. All other laboratory work, including all other bioinformatic analyses was performed by this candidate. The PhD thesis and publication were prepared by this candidate, with major editing and concept input performed by the candidate's PhD supervisors, Dr Marina Harper and Assoc. Prof. John Boyce. Minor edits of the Hfq manuscript (Chapter 2) were performed by Prof. Ben Adler and Amy Wright.

I have renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student name: Marianne Mégroz

Student signature:  **Date: 20.02.2020**

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work. In instances where I am not the responsible author I have consulted with the responsible author to agree on the respective contributions of the authors.

Main Supervisor name: John Boyce

Main Supervisor signature:  **Date: 20.02.2020**

Acknowledgements

Undertaking my PhD has been one of the most challenging yet rewarding experiences of my life. It has pushed me out of my comfort zone and has taught me so much scientifically, professionally and personally. I have been enormously privileged to have had a strong support network to encourage and uplift me during my project, especially when things did not seem ideal nor on track. I am very grateful for their support and would not have been able to complete this PhD without them.

Firstly, I would like to thank my wonderful PhD supervisors, A/Prof John Boyce and Dr Marina Harper for your unending support, guidance and knowledge throughout my PhD. You have both been constantly willing to help and advise me, and to answer my endless questions; your knowledge and experience is admirable. I sincerely appreciate the time and effort that you have each put towards guiding me through my project and I am very thankful for the opportunity to have worked alongside you and as a part of your team.

Thank you also to the Australian government for providing me with a Research Training Program Scholarship, allowing me to focus my efforts on my research. Thank you also to the Micromon genomics platform for providing me with travel grants to attend conferences both nationally and internationally.

I would like to thank our collaborator, Dr Jai Tree, as well as his team for their contribution to the Hfq UV-CRAC/CLASH experiments in my project. Thank you for all of your assistance and time, and also to Dr Emily Gulliver for your efforts with these experiments during your time at the University of New South Wales.

I would like to thank all members of the Boyce Lab, both past and present, for their insight, support and friendship throughout my project. These include (in no particular order) Amy, Caitlin, Dee, Jess, Jordan, Michael, Tom, Tim, Vali, Vicki (and previous honorary lab member, Hasini). I have been immensely lucky to be part of such an enthusiastic and cooperative team who are always there for advice as well as a good chat (with a side of banter). I have particularly enjoyed our regular Cinque café morning teas and the daily quiz sessions run by Amy. I especially thank Jess and Amy for your friendship and for always being there for me throughout my project, both professionally and personally. Thank you also to the Monash Microbiology Department and the

PhD student cohort for welcoming me and being a huge part of my life over the past years. I particularly thank Sherrie for helping everything run smoothly throughout my PhD process.

To my parents, thank you for all of your help and support throughout my studies, and for always believing in me. I would not be where I am today without your encouragement and sacrifices. I am very thankful to be able to come home to your company and delicious home cooked meals. To my sisters Yvonne and Emilia, thank you for all of your assistance and motivation during my studies, and for enduring my 'study-inflicted' stress and odd behaviours. Thanks for being there for me through the good and bad, I am proud to have you both as my sisters.

To my family dogs Toby (dearly missed) and Shelby, thank you for your therapeutic pats and cuddles whenever I needed a furry hug.

To Robbie, thank you for your unending support, encouragement and patience throughout my project, especially through the final write-up period. You have always pushed me to go for every opportunity, to not let things stand in my way, and to see things with a more positive perspective. I really appreciate the late nights you stayed up to support and check in with me while working, especially when my sanity was questionable. I am very lucky to have had you by my side through this rollercoaster and in life.

I would also like to thank my extended family for all of their support and positivity throughout my studies. To my grandparents, both with and without me now, thank you for believing in me and encouraging me throughout my education. To my friends, especially Nikki and Danielle, thank you for your constant encouragement, positivity and support. I also thank my thesis writing buddies, Nat and Diana, for your moral support during our write-up stages, as well as our Cinque café breakfasts which always provided some morning motivation.

Thank you all, and to all those not mentioned and whom I have met along the way. I feel extremely fortunate to be surrounded by such supportive people in my life.

Chapter One: Introduction

1.1 Introduction to *Pasteurella multocida*

Louis Pasteur first isolated *Pasteurella multocida* from chickens in 1880 and determined that the organism was the causative agent of the disease fowl cholera (Harper et al., 2006). By inoculating chickens with a laboratory attenuated strain of *P. multocida,* Pasteur also developed the first live bacterial vaccine that could stimulate a protective immune response against a specific bacterial disease. *P. multocida* is the most widely studied member of the *Pasteurella* genus and is the primary causative agent of many other important animal diseases that result in significant global economic burden (Boyce et al., 2010). In addition to fowl cholera, these diseases include haemorrhagic septicaemia in ungulates, atrophic rhinitis in swine, respiratory disease in rabbits, and bite wound-associated skin or soft tissue infections in humans (Boyce et al., 2010).

P. multocida is a Gram-negative bacterium that is encapsulated, coccobacillus in shape, non-motile and a facultative anaerobe (Boyce et al., 2010). It is a heterogeneous species with strains classified up until recently into five serogroups (A, B, D, E and F), based on capsular antigens, and into 16 Heddleston serovars (1 to 16), based on lipopolysaccharide (LPS) antigens (Carter, 1952; Heddleston et al., 1972). Each isolate is then designated by a capsular serogroup number followed by a Heddleston LPS serovar letter, for example A:1 indicates a strain that is capsular group A and Heddleston serovar 1. Recently, two new multiplex polymerase chain reaction (PCR) systems have been developed to replace serological identification methods. The 16 Heddleston LPS serovar types can be classified into eight unique LPS genotypes (L1 through to L8) based on the LPS outer core biosynthesis genes (Harper et al., 2015) and capsule type based on the capsule locus (still genes designated A, B, D, E and F) (Townsend et al., 2001). Though *P. multocida* can cause a range of disease syndromes in a variety of hosts, particular diseases are often associated with specific capsular serogroups (Boyce et al., 2010).

1.1.1 Diseases caused by *P. multocida*

P. multocida is a component of the normal flora of the oropharynx of many species, but is also the primary causative agent of a range of serious animal diseases. Haemorrhagic septicaemia (HS) is a systemic disease of ungulates that has high a very mortality rate (Wilson and Ho, 2013) and is endemic to tropical regions of Africa (strain E:2) and Asia (strain B:2). Cattle and buffalo are particularly susceptible (Benkirane and De Alwis, 2002; Shivachandra et al., 2011). The site of initial infection is likely the tonsils; *P. multocida* then spreads to local lymph nodes, predicted to occur via internalisation into macrophages, and into the surrounding tissues causing local oedema and tissue necrosis (Wilkie et al., 2012). Infected animals succumb to the disease unless effective

antibiotic treatment occurs very rapidly after appearance of the initial signs of HS (Wilkie et al., 2012).

Atrophic rhinitis (AR) is a disease caused by toxigenic *P. multocida* strains belonging to capsular type D and, less commonly, type A (Wilson and Ho, 2013). This chronic mucosal nasal disease results in nasal turbinate destruction and maxillary dysplasia (Orth, 2009). It is most commonly observed in swine but has also been reported in housed rabbits, goats and cattle (DiGiacomo et al., 1989; Dutt and Kameswaran, 2005; Roumen and van de Goor, 1988). AR results in large economic losses in swine industries globally due to the reduced growth rate of infected swine (Wilkie et al., 2012). The effects of the disease are due almost entirely to the action of a single toxin, the *Pasteurella multocida* toxin (PMT). PMT is a dermo-necrotoxin produced by the organism during growth within the upper respiratory tract (Wilkie et al., 2012). PMT disrupts osteoblast function, affecting the activity of osteoclasts and resulting in the promotion of bone resorption and loss of nasal turbinate structure (Mullan and Lax, 1998).

P. multocida infection can also result in respiratory diseases such as 'snuffles' (rhinitis) and pneumonia in rabbits and ungulates (Jarvinen et al., 1998). 'Snuffles' of rabbits is caused by colonisation of the upper respiratory tract, usually by *P. multocida* capsular serogroup A, or sometimes D, strains (Wilkie et al., 2012). Snuffles is characterised by an acute chronic exudative rhinitis, and may lead to bronchopneumonia or other localised infections and eventually death (Deeb et al., 1990; Jarvinen et al., 1998).

P. multocida can also play a secondary role in disease as an opportunistic pathogen, causing various soft tissue infections (Baillot et al., 2011; Talan et al., 1999). Any species, including humans, can develop skin infections such as cellulitis or subcutaneous abscesses as a result of bite wounds from species infected with, or colonised by, *P. multocida* (Boyce et al., 2010; Talan et al., 1999). Cases of other infections affecting a range of body sites including endocarditis, urinary tract infections and bacteremic meningitis have also been reported in humans as a result of *P. multocida* infection via contact with pets (Costanzo et al., 2017; Graf et al., 2007; Katechakis et al., 2019; Singh and Harrington, 2017; Wilson and Ho, 2013).

1.2 *P. multocida* as the causative agent of fowl cholera

1.2.1 Fowl cholera

Fowl cholera is an important disease caused by *P. multocida*, that can affect most avian species including chickens, turkeys, ducks and wild waterfowl (Blackall and Miflin, 2000; Boyce et al., 2010). Infections result in a high mortality rate and cause major economic burden to poultry industries worldwide (Blackall and Miflin, 2000; Carpenter et al., 1988; Chung et al., 2005). Strains belonging to capsular serogroup A are most commonly associated with fowl cholera; however, strains belonging to other capsular types, such as serogroups F and D, have also been associated with this disease (Wilson and Ho, 2013).

Fowl cholera is transmitted via contaminated water and feed, or through direct contact with diseased or dead birds (Wilkie et al., 2012). *P. multocida* can be readily isolated from pond water during outbreaks in wild waterfowl, while infection has been observed in birds of prey following consumption of infected water birds (Backstrand and Botzler, 1986; Williams et al., 1987). Although the exact route of infection of fowl cholera is unclear, the respiratory tract is thought to be the main site of entry. Virulent *P. multocida* can colonise the mucosa of the upper respiratory tract, and further spread to air sacs, lungs and into the circulation system (Matsumoto et al., 1991; Wilkie et al., 2012). The organisms can then rapidly multiply in other tissues, particularly the liver and spleen (Rhoades and Rimler, 1990). Experiments have shown that *P. multocida* can adhere to turkey air sac macrophages, and it is predicted that this may facilitate the spread to other tissues (Pruimboom et al., 1999). However, entry of *P. multocida* through leg abrasions, which was observed during a fowl cholera outbreak in ducklings, suggests that other routes of entry are possible (Wilkie et al., 2012).

Fowl cholera may present in peracute, acute or chronic forms, with clinical course of disease varying from a few hours to several days. While sudden death may be the first evidence of disease in peracute and acute cases, a more prolonged clinical course may occur in other cases, involving depression, lethargy, inappetence, fever, serous exudation from nostrils, swelling of wattles and combs and respiratory complications (Boyce et al., 2010; Woo and Kim, 2006). Morbidity and mortality are generally high unless a low virulence isolate is involved, in which case birds may slowly recover. Recovery may be complete, or in some cases exudative arthritis in leg or wing joints may develop (Wilkie et al., 2012).

Effective biosecurity and vaccination are the primary methods of prevention of fowl cholera (Blackall and Miflin, 2000). Killed whole cell vaccines (bacterins) and live-attenuated vaccines

are available to prevent disease, however these are limited in efficacy especially against heterologous strains (Christiansen et al., 1992; Chung et al., 2005; Heddleston et al., 1970). A more recent study demonstrated that *P. multocida* LPS mutants presented as killed whole cell vaccines provide protection only against other strains with identical or nearly identical surface LPS structures, while live-attenuated vaccines provided broader protection, largely independent of LPS structure (Harper et al., 2016).

1.2.2 Virulence factors associated with fowl cholera

Toxins, such as PMT are not usually associated with fowl cholera however a number of other *P. multocida* virulence factors have been identified which contribute to fowl cholera (Harper et al., 2006; Peng et al., 2019). These include polysaccharide capsule, lipopolysaccharide (LPS), iron sequestering systems, adhesins and sialometabolism (the uptake of sialic acid as a nutrient or for decoration of bacterial surface components) (Harper et al., 2006; Wilson and Ho, 2013).

1.2.2.1 Capsule

The polysaccharide capsule is a critical virulence factor of *P. multocida,* allowing the bacteria to evade host immune defence mechanisms (Chung et al., 2001; DeAngelis et al., 2002). Acapsular mutants are highly attenuated for virulence in mice and chickens and show increased sensitivity to both the bactericidal action of complement in chicken serum and phagocytosis (Boyce and Adler, 2000; Chung et al., 2001). Resistance to complement-mediated lysis is important for virulence and correlates with possession of capsule in *P. multocida* type A strains (Hansen and Hirsh, 1989; Snipes and Hirsh, 1986).

Serogroup A *P. multocida* strains, which are most commonly associated with fowl cholera, express a capsule primarily composed of hyaluronic acid (Carter, 1958). The genes for the biosynthesis, transport and surface localisation of this capsule are found in a single locus comprising ten genes; namely, *phyBA-hyaEDCB-hexDCBA*. The genes *phyBA* are predicted to encode proteins with roles in polysaccharide lipidation and surface attachment, *hyaEDCB* encode proteins involved in synthesis of the capsular polysaccharide, and *hexDCBA* encode proteins required for transport of the polysaccharide to the bacterial surface (Chung et al., 1998).

Expression of the *P. multocida* capsule is, in part, regulated by Fis, a global transcriptional regulator (Steen et al., 2010). Fis mutants express highly reduced levels of extracellular capsule and are completely attenuated for virulence in chickens and mice (Steen et al., 2010). Whole genome transcriptional analysis of the *P. multocida fis* mutants showed that Fis may also directly

regulate the production of other virulence factors, such as *pfhB2*, encoding a filamentous hemagglutinin, and the cross-protective surface antigen PlpE (Steen et al., 2010).

1.2.2.2 Lipopolysaccharide (LPS)

LPS forms the outer leaflet of the outer membrane of Gram-negative bacteria. *P. multocida* expresses LPS that lacks a repeating O-antigen, but strains belonging to different serovars/LPS genotypes each express LPS with a different outer core structure (Harper et al., 2011). LPS plays a critical role in the pathogenesis of fowl cholera; defined *P. multocida* mutants with truncated LPS are attenuated for virulence in chickens (Fuller et al., 2000; Harper et al., 2007b; Harper et al., 2004). In the *P. multocida* strain VP161, reduced virulence in fowl cholera disease models has been observed for mutant strains expressing truncated LPS, and these mutants showed increased susceptibility to the chicken antimicrobial peptide, fowlicidin (Boyce et al., 2009; Harper et al., 2007a; Harper et al., 2004). LPS is also a protective antigen, and immunisation of animals with killed bacterins results in protection against strains expressing identical or very similar LPS (homologous strains) (Harper et al., 2011).

1.2.2.3 Iron sequestering systems

Iron sequestering systems have been shown to contribute to *P. multocida* virulence in fowl cholera. These systems act to allow sufficient acquisition of host iron for the survival and growth of the bacteria during growth *in vivo*, where levels of free iron are usually very low (Boyce and Adler, 2006). *P. multocida* mutants with disruption in the functionally related iron uptake genes *tonB, exbD,* and *exbB,* or the haemoglobin-binding protein gene *hgbA* were attenuated for virulence in mice (Bosch et al., 2002a; Bosch et al., 2002b).

1.2.2.4 Fimbriae and adhesins

Type IV fimbriae (pili) are known to be involved in virulence in many bacterial species through their role in attachment to host cell surfaces (Ruffolo et al., 1997). Type IV fimbriae have been identified in *P. multocida* strains A, B and D, and fimbriae have been identified on strains with the ability to adhere to the mucosal epithelium, but not on strains unable to adhere (Glorioso et al., 1982). However, a type IV fimbrial mutant (*ptfA*) was still able to cause disease in chickens although time to death was slightly extended (Boyce and Harper, pers. comm.). *P. multocida* expresses two large proteins PfhB1 and PfhB2 with significant similarity to the filamentous hemagglutinins of *Bordetella pertussis* (Tatum et al., 2005). In *B. pertussis*, these hemagglutinin proteins play an important role in the colonisation of the upper respiratory tract (Kimura et al., 1990). Mutagenesis of *pfhB2* in *P. multocida* resulted in a significant reduction of virulence in turkeys (Tatum et al., 2005); while birds infected with the wild-type parent strain all died of acute

fowl cholera within 28 h, birds that received the same low intravenous dose of the *pfhB2* mutant showed no clinical signs (Tatum et al., 2005).

1.2.2.5 Sialometabolism

Host sialic acid may be utilised by bacteria as a source of carbon, nitrogen or amino sugars (Vimr et al., 2004). Incorporation of sialic acid into cell surface polysaccharides may promote biofilm formation, and more importantly for virulence of *P. multocida*, may aid in the evasion of the host immune system (Steenbergen et al., 2005; Vimr et al., 2004). Sialic acid has been identified in LPS expressed by *P. multocida* strains Pm70 and P1059 (Steenbergen et al., 2005; Tatum et al., 2009). Although *P. multocida* cannot synthesise sialic acid *de novo*, the organism utilises the sialidases NanH and NanB to obtain sialic acid from the host (Mizan et al., 2000). Two sialyltransferases, Pm0508 and Pm0188, involved in the transfer of active sialic acid to cell surface components, have been identified in *P. multocida*. While the role of the sialyltransferase Pm0508 in virulence remains unclear, inactivation of the *P. multocida* Pm0188 sialyltransferase resulted in significant attenuation of growth of *P. multocida* in chickens, indicating its importance in virulence (Boyce et al., 2010).

1.3 Riboregulation and small RNAs

Riboregulation involves the regulation of gene expression through the action of regulatory RNA molecules (Papenfort and Vogel, 2010). Small noncoding RNA molecules (sRNAs) have increasingly been recognised as important regulators of bacterial gene expression, with essential roles in controlling diverse bacterial functions including virulence, carbon, amino acid and iron metabolism, outer membrane protein production and quorum sensing (Guillier et al., 2006; Lenz et al., 2004; Li et al., 2012b; Mandin and Guillier, 2013; Massé and Gottesman, 2002; Romby et al., 2006; Sharma et al., 2011; Sonnleitner et al., 2009). Small RNAs are short transcripts, usually 50-400 nucleotides in length, that facilitate regulatory effects, either directly or indirectly through interaction with other RNA molecules. These molecules primarily facilitate gene regulation via interaction with one or more mRNAs via short imperfect base pairing, and can alter protein production via multiple post-transcriptional mechanisms (Gottesman and Storz, 2011; Vogel and Luisi, 2011). Binding of sRNAs at or near the ribosome binding site (RBS) of an mRNA can result in inhibition of translational initiation (Gottesman and Storz, 2011). Conversely, as has been shown for *Escherichia coli rpoS*, some sRNA molecules bind to an mRNA secondary structure that naturally occludes the RBS, thus favouring translational initiation and protein production (Majdalani et al., 1998). Binding of an sRNA can also induce RNase E-mediated degradation of the target RNA (Massé et al., 2003). The seed region of an sRNA is a short region of sequence

complementarity, usually 13-24 bp in length, where an sRNA interacts with its target mRNA molecule via base pairing (Peer and Margalit, 2011; Storz et al., 2011). Although sRNA-encoding regions typically lack nucleotide motifs such as ribosome binding sites, they may still contain 'orphan' transcription promoter and terminator sequences (Waters and Storz, 2009). However, some sRNAs lack identifiable promoter/terminator sequences and some may be generated via endonuclease cleavage of longer transcripts (De Lay and Garsin, 2016).

In general, sRNAs may be classified as either *cis*- or *trans*-encoded. *Cis*-encoded sRNAs, also referred to as antisense sRNAs, are encoded on the opposite strand of the gene that they target to facilitate riboregulatory functions (Brantl, 2007; Wagner et al., 2002). *Cis*-encoded sRNAs therefore share long regions of complementarity to their targets (Brantl, 2007). Although *cis*-encoded sRNAs were primarily identified in plasmids, transposons and phages (Wagner et al., 2002), these sRNAs are now increasingly being found as chromosomally encoded transcripts (Waters and Storz, 2009). In contrast to *cis*-encoded sRNAs, *trans*-encoded sRNAs are encoded in intergenic regions and only share limited complementarity (usually ~10-25 bp) with their target RNAs (Waters and Storz, 2009). *Trans*-encoded sRNAs can therefore interact with target mRNAs encoded anywhere on the genome, and they typically target multiple mRNAs (Gottesman and Storz, 2011; Waters and Storz, 2009). As most sRNA molecules have multiple mRNA targets, and many mRNA targets are regulated by multiple sRNA molecules, a complex inter-linked sRNA-mRNA regulatory network is formed that allows for precisely coordinated responses to many different environments (Papenfort and Vogel, 2009). Other regulatory RNA species are also known to play important roles in bacterial riboregulation. These include *cis*-acting riboswitches, thermosensors and other *cis*-acting regulatory elements of mRNAs, as well as tRNA and rRNA cleavage products (De Lay and Garsin, 2016; Gottesman and Storz, 2011; Kavita et al., 2018).

The expression of sRNAs is known to vary considerably under different bacterial growth conditions, including anaerobic growth, iron-limited growth and when under oxidative stress (Gottesman and Storz, 2011). Therefore, a given sRNA may only be expressed during growth under a specific condition, or during a particular growth phase. There is also considerable variation in the number and sequence of sRNAs expressed by different bacterial species. Model organisms used in the initial sRNA studies, *E. coli* and *Salmonella enterica* serovar *Typhimurium* (*S.* Typhimurium), are now known to each express more than 200 sRNAs (Barquist and Vogel, 2015). Importantly, sRNA riboregulatory functions often crucially rely on interactions with one of the RNA chaperone proteins, namely Hfq, ProQ and CsrA (Holmqvist and Vogel, 2018).

1.4 The Hfq protein

1.4.1 Properties of Hfq

The action of each sRNA species is often critically dependent on the activity of the RNA-binding chaperone protein, Hfq. First discovered as a host factor required for the *in vitro* replication of the RNA phage Qβ in *E. coli,* Hfq regulates gene expression by modulating the binding of sRNA molecules to their cognate target mRNAs via the formation of an sRNA-mRNA complex. These interactions can result in both positive and negative regulation of gene expression (Papenfort and Vogel, 2010). Hfq homologues have been identified in many bacterial species (Chao and Vogel, 2010). The protein is a member of the Sm and Sm-like (LSm) RNA-binding protein family and forms a doughnut shaped homo-hexameric ring, approximately 65 Å in diameter, with a central pore (Figure 1.1) (Franze De Fernandez et al., 1968; Moller et al., 2002; Sobrero and Valverde, 2012; Valentin-Hansen et al., 2004). Within each Hfq monomer there is an Sm fold, consisting of five antiparallel β-strands with an N-terminal α-helix (Brennan and Link, 2007; Schumacher et al., 2002).

The Hfq homo-hexamer has four surfaces for binding RNA molecules, including two asymmetrical faces, a proximal face, with an N-terminal α-helix and a distal face. The other binding surfaces include a C-terminal tail which extends laterally from the distal face, and an outer ring which is referred to as the rim or lateral face (Figure 1.1) (Beich-Frandsen et al., 2011; Sauer, 2013; Updegrove et al., 2016). Each of these four surfaces provides a binding site for single-stranded RNA molecules and has a unique architecture (Holmqvist and Vogel, 2018; Vogel and Luisi, 2011). The asymmetrical surfaces of Hfq present a net positive electrostatic potential, which allows for nucleic acid binding and interactions with Hfq targets (Sobrero and Valverde, 2012).

A range of *in vivo* and *in vitro* studies have been performed to establish a general model for Hfq interaction with RNAs at the different binding surfaces; these binding characteristics have primarily been determined from studies performed in *E. coli* and *Salmonella* (Holmqvist and Vogel, 2018). Crystallisation and binding studies have shown that the proximal face of Hfq preferentially binds to uridine-rich sequences present at the 3′ end of sRNA molecules. (Dimastrogiovanni et al., 2014; Sauer and Weichenrieder, 2011). Uridine-rich sequences at the 3′ end of sRNA molecules are characteristic of Rho-independent terminator tail sequences, and are likely to significantly contribute to sRNA recognition by Hfq (Sauer and Weichenrieder, 2011). The rim region of the Hfq protein may also be involved with sRNA positioning that allows the sRNA seed region to be exposed for interaction with a cognate RNA molecule (Dimastrogiovanni

et al., 2014). The distal face of Hfq has been shown to preferentially interact with single-stranded adenine-rich sequences typically present in the 5′ untranslated regions (UTRs) of mRNAs, facilitating close contact between the mRNAs and sRNAs bound to the proximal face of Hfq (Link et al., 2009; Mikulecky et al., 2004; Peng et al., 2014). The UA-rich sequences of sRNA and mRNA pairs simultaneously associating with Hfq are then able to interact with each other via the rim surface region of Hfq (Panja et al., 2013).

The C-terminal region of Hfq is intrinsically disordered and variable across different species, both in length and sequence, and its role in riboregulation is controversial. In *E. coli*, the C-terminus of Hfq was shown to have specificity for mRNAs and was necessary for RNA interactions (Sobrero and Valverde, 2012; Večerek et al., 2008). Conversely, other studies have shown that the *E. coli* Hfq C-terminus plays only a minor role, if any, in riboregulation (Olsen et al., 2010). Recent studies suggest that the C-terminal region acts to displace Hfq bound sRNA-mRNA complexes and allow Hfq to rapidly cycle between competing RNA targets (Santiago-Frangos et al., 2016; Santiago-Frangos and Woodson, 2018).

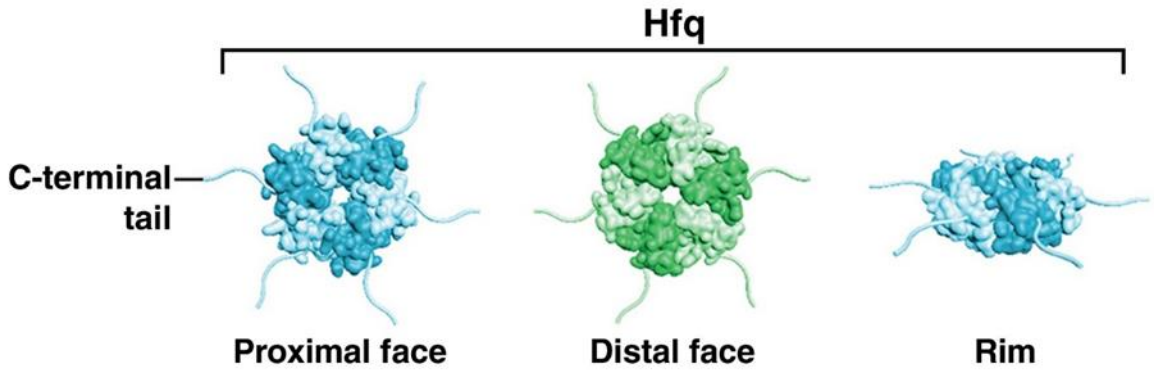


Figure 1.1. *E. coli* Hfq (teal for the proximal face and rim views and green for the distal face view) employs four solvent exposed surfaces to interact with RNA (Figure and legend taken from Updegrove et al., 2016.

1.4.2 Hfq and its riboregulatory functions

By preferentially binding to A/U-rich sequences on RNA molecules, Hfq plays an important role in mediating the various post-transcriptional riboregulatory mechanisms of sRNAs at the level of translational efficiency and RNA stability (Schumacher et al., 2002). Several general regulatory mechanisms have been identified that allow Hfq to facilitate both positive and negative regulation of gene expression (Brennan and Link, 2007; Papenfort and Vogel, 2010). Most commonly, Hfq regulates gene expression by modulating the binding of sRNA molecules to their cognate target mRNA (Figure 1.2.a) (Brennan and Link, 2007; Sobrero and Valverde, 2012). These antisense pairing interactions between the mRNA and sRNA usually occur in close proximity to the ribosome binding site, as Hfq can promote the binding of an sRNA to the 5′ region of its target mRNA (Sobrero and Valverde, 2012; Vogel and Luisi, 2011). Therefore, Hfq may act to effectively block ribosomal access to an mRNA and inhibit translation of that message into a polypeptide (Figure 1.2.a) (Vogel and Luisi, 2011). Other riboregulatory functions of Hfq include the activation of mRNA translation through the formation of a Hfq/sRNA complex where the translation initiation region on the mRNA becomes more exposed, allowing for 30S ribosomal binding, and thereby positively regulating mRNA translation (Figure 1.2.b) (Soper et al., 2010; Vogel and Luisi, 2011). Hfq may also act to stabilise or protect an sRNA from cleavage by ribonuclease E (Figure 1.2.c) leading to increased message lifetime (Vogel and Luisi, 2011). Additionally, Hfq has been shown to facilitate coupled sRNA and target mRNA degradation (Figure 1.2.d).

The Hfq protein can also act independently of sRNAs to regulate mRNA activity (Figure 1.2.e) (Sobrero and Valverde, 2012; Vogel and Luisi, 2011). In *E. coli*, Hfq has been shown to directly interact with poly(A) polymerase to induce mRNA polyadenylation and subsequent exoribonuclease degradation of some mRNAs (Mohanty et al., 2004). Furthermore, direct binding of Hfq to the *mutS* mRNA facilitates translational repression of the gene during mismatch repair processes in *E. coli* (Chen and Gottesman, 2017). In addition, it has been shown that Hfq represses mRNA translation in *Pseudomonas aeruginosa* when stimulated by the catabolite repression control protein Crc; upon binding of RNA to the distal side of Hfq, Crc and Hfq form a complex, enhancing the stability of the Hfq–RNA interactions and promoting translational repression (Sonnleitner and Bläsi, 2014; Sonnleitner et al., 2017).

Although Hfq associates with other nucleic acid species, such as tRNAs, rRNAs and in some cases DNA (Geinguenaud et al., 2011; Lee and Feig, 2008; Takada et al., 1997), co-immunoprecipitation

analyses with *E. coli* and *Salmonella* Hfq have demonstrated a preferential association with mRNA and sRNA molecules (Sittka et al., 2008; Sittka et al., 2009; Zhang et al., 2003).

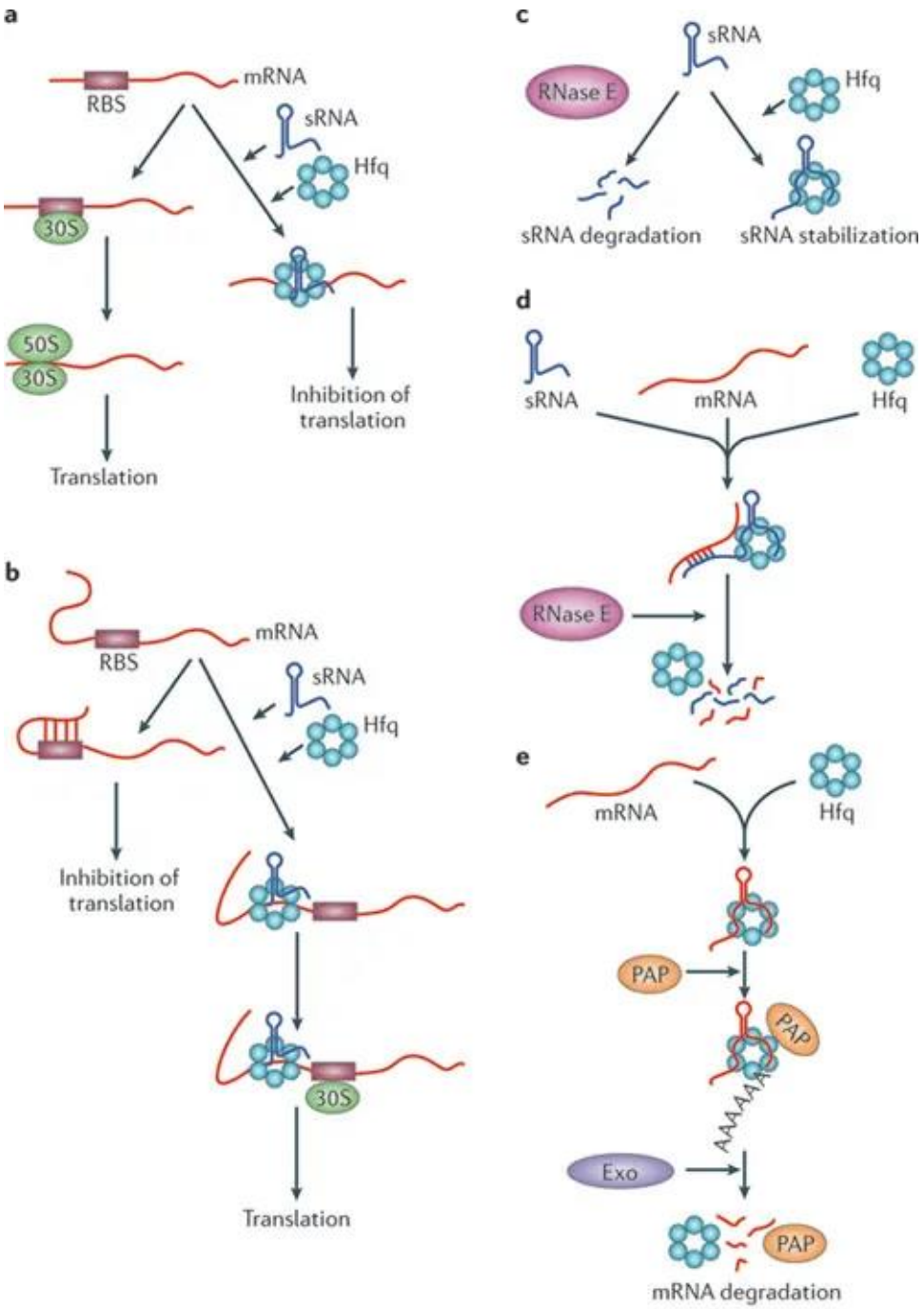


Figure 1.2. Widely accepted modes of Hfq activity. a. Hfq in association with a small RNA (sRNA) may sequester the ribosome-binding site (RBS) of a target mRNA, thus blocking binding of the 30S and 50S ribosomal subunits and repressing translation. **b.** In some mRNAs, a secondary structure in the 5′ untranslated region (UTR) can mask the RBS and inhibit translation. A complex formed by Hfq and a specific sRNA may activate the translation of one of these mRNAs by exposing the translation initiation region for 30S binding. **c.** Hfq may protect some sRNAs from ribonuclease cleavage, which is carried out by ribonuclease E (RNaseE) in many cases. **d.** Hfq may induce the cleavage (often by RNaseE) of some sRNAs and their target mRNAs. **e**. Hfq may stimulate the polyadenylation of an mRNA by poly(A) polymerase (PAP), which in turn triggers 3′-to-5′ degradation by an exoribonuclease (Exo). In *Escherichia coli*, the exoribonuclease can be polynucleotide phosphorylase, RNase R or RNase II. Figure and legend taken from Vogel and Luisi, 2011.

1.4.3 *hfq* mutagenesis studies

Hfq homologues have been identified in a range of Gram-positive and Gram-negative bacterial species, including many pathogenic species (Chao and Vogel, 2010; Koo et al., 2011). Most *hfq* mutant strains display a broad range of pleiotropic phenotypes, suggesting a global role for Hfq in bacterial physiology (Sobrero and Valverde, 2012; Vogel and Luisi, 2011). These pleiotropic effects have been observed more commonly in Gram-negative species including the pathogens *S.* Typhimurium, *Yersinia pestis, Moraxella catarrhalis, Neisseria meningitidis, Brucella abortus,* and *Vibrio cholera,* but also in some Gram-positive species such as *Listeria monocytogenes* (Attia et al., 2008; Christiansen et al., 2004; Ding et al., 2004; Fantappiè et al., 2009; Figueroa-Bossi et al., 2006; Geng et al., 2009; Robertson and Roop, 1999). Indeed, Hfq was required for maximum fitness of many species in a range of *in vitro* and *in vivo* niches (Dietrich et al., 2009; Sonnleitner et al., 2003; Tsui et al., 1994). Many *hfq* mutants show increased sensitivity to host defence mechanisms, reduced growth rate and virulence, and increased susceptibility to a range of environmental stresses (Dietrich et al., 2009; Sobrero and Valverde, 2012; Sonnleitner et al., 2003; Tsui et al., 1994). Pleiotropic changes at the mRNA and protein levels have also resulted from inactivation of *hfq* in mutagenesis studies (Ansong et al., 2009; Dietrich et al., 2009). Transcriptomic analysis of *hfq* mutants has indicated that Hfq is involved in the regulation of at least 20% of *Salmonella* genes, and approximately 15% of *N. gonorrhoeae* and *P. aeruginosa* genes (Chao and Vogel, 2010; Dietrich et al., 2009; Sittka et al., 2008; Sonnleitner et al., 2006). Importantly, an overrepresentation of differentially expressed genes with virulence and stress-associated functions has been shown for *hfq* mutants in a number of bacterial species using transcriptomic analyses (Ansong et al., 2009; Geng et al., 2009).

As *hfq* mutants in a range of bacterial pathogens have reduced virulence in the relevant disease models, these attenuated strains have been considered for use as live vaccines (Sobrero and Valverde, 2012). Indeed, a single oral immunisation with a *S.* Typhimurium *hfq* mutant conferred protection against an oral challenge in mice with the virulent wild-type strain (Allam et al., 2011; Sittka et al., 2007). Similarly, vaccination of zebrafish with a live *Vibrio alginolyticus hfq* mutant elicited significant protection against wild-type challenge (Liu et al., 2011). These studies indicate that *hfq* mutants may be valuable candidates for novel live oral vaccines against many bacterial infectious diseases (Sobrero and Valverde, 2012).

1.4.4 Regulation of Hfq

It is clear that Hfq plays a major role in global regulation of gene expression; however, studies exploring the regulation of Hfq synthesis are limited. It has been estimated that there are

approximately 30,000-60,000 individual Hfq molecules per *E. coli* cell (up to 10,000 Hfq hexamers) during exponential growth (Ali Azam et al., 1999; Kajitani et al., 1994). Autoregulation of Hfq at the translational level has also been demonstrated in *E. coli* via the Hfq hexamer binding to the 5′ UTR of the *hfq* mRNA (Večerek et al., 2005). Recent studies have supported these findings and showed that this binding occurred primarily via the distal face of Hfq (Morita and Aiba, 2019).

1.4.5 *P. multocida* Hfq

Bioinformatic analysis shows that a Hfq orthologue is present in all available *P. multocida* genomes. The *P. multocida* Hfq is 96 amino acids in length and shares 93% amino acid identity with the *E. coli* Hfq protein over the first 71 amino acids. As has been observed for other Hfq proteins, the C-terminal amino acid sequence shares little identity with other Hfq proteins (Sobrero and Valverde, 2012). The protein contains Hfq and Sm-like superfamily domains and conserved Sm1 and Sm2 sequence motifs, the latter being common to members of the Sm/LSm family (Figure 1.3) (Moller et al., 2002).

In most bacterial genomes, *hfq* is located directly upstream of *hflX*, which encodes a GTPase that interacts with the 50S ribosomal subunit (Sobrero and Valverde, 2012; Wu et al., 2010). Moreover, *hflX* is typically located only a few nucleotides downstream from the stop codon of *hfq*, and the two genes are likely co-transcribed as well as translationally coupled (Sobrero and Valverde, 2012). However, co-transcription has only been demonstrated in a few species including *E. coli* and *Francisella tularensis* (Meibom et al., 2009; Tsui et al., 1994). This genetic organisation of *hfq* and *hflX* was also identified in *P. multocida* genomes, but there is no known association between the two encoded products in this species; in previous transcriptomic studies on *P. multocida* strain VP161, the *hfq* RNA transcript levels were approximately five times higher than those of *hflX*, suggesting that they are not co-transcribed (Boyce and Harper, pers. comm.).

Although gene regulation in *P. multocida* has not been extensively studied, Fis has been shown to be involved in the regulation of more than forty genes in *P. multocida* (Steen et al., 2010). Fis was shown to be essential for capsule production and for the expression of other important virulence factors including the filamentous hemagglutinin *pfhB2* (Steen et al., 2010). Interestingly, the expression of *hfq* during early-exponential growth phase was 2.1-fold (FDR = 0.03) higher in a *fis* mutant compared to expression in the wild-type strain, suggesting a potential regulatory link between Hfq and Fis (Boyce and Harper, pers. comm.).

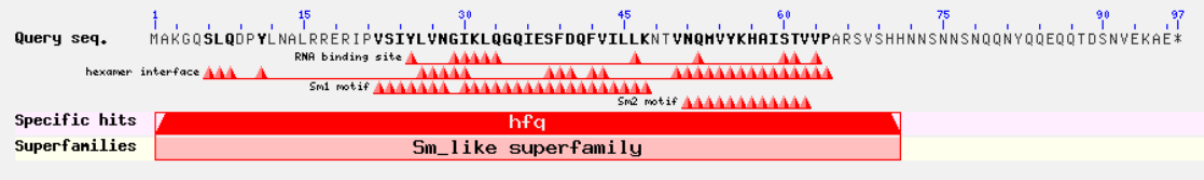


Figure 1.3. Conserved domains and motifs present in the *P. multocida* VP161 Hfq as determined by NCBI Conserved Domain Search. The red triangles on the conserved features indicate residues that have mapped to the Hfq query sequence, also indicated by the bolded amino acids in the Hfq query sequence.

1.4.6 Other riboregulatory proteins: ProQ and CsrA

ProQ has recently emerged as an important RNA chaperone involved in global bacterial post-transcriptional regulation via sRNAs (Gonzalez et al., 2017; Smirnov et al., 2016). The ProQ protein contains a FinO domain and in *E. coli* and *Salmonella* has been shown to associate with hundreds of RNA transcripts including many sRNAs (Holmqvist et al., 2018; Smirnov et al., 2016; Smirnov et al., 2017). Although there is currently limited information on the role of ProQ in bacterial pathogenesis, transcriptomic analyses have shown ProQ as an important regulator of virulence gene expression in *Salmonella* (Westermann et al., 2019). Initial ProQ analyses in *P. multocida,* performed in parallel to this study, identified a range of mRNAs, tRNAs and sRNAs showed an association with ProQ (Gulliver, Boyce Laboratory, manuscript in preparation).

The carbon storage regulator A protein, CsrA, is another well-characterised and highly conserved RNA-binding protein that facilitates post-transcriptional regulation in many bacterial species (Romeo and Babitzke, 2018; Vakulskas et al., 2015). CsrA has been shown to play an important role in the regulation of genes involved in bacterial virulence, carbon metabolism, motility, quorum sensing and biofilm formation (Jackson et al., 2002; Romeo et al., 1993; Sabnis et al., 1995; Wei et al., 2001; Yakhnin et al., 2013; Yakhnin et al., 2011). The CsrA protein functions by binding to target mRNA molecules via various mechanisms to alter RNA stability, secondary structure and/or Rho-mediated transcription termination (Potts et al., 2017; Romeo and Babitzke, 2018). Multiple sRNA antagonists have been identified that regulate the activity of CsrA by competing with mRNAs for CsrA binding (Romeo and Babitzke, 2018).

1.5 Methods of bacterial sRNA and sRNA-target identification

Regulatory sRNAs have been most widely studied in model organisms such as *E. coli* and *S.* Typhimurium (Barquist and Vogel, 2015). A range of techniques have been utilised to identify sRNAs and their targets across many bacterial species, including bioinformatic and experimental-based approaches. Experimental validation methods for sRNAs have also been developed to confirm the authenticity of predicted sRNAs.

1.5.1 Bioinformatic strategies for identification of bacterial sRNAs in whole genome sequences

With increased availability of sequenced bacterial genomes, developments in computational methods have allowed for the bioinformatic prediction of many sRNAs across diverse bacterial species (Sridhar and Gunasekaran, 2013; Vogel and Sharma, 2005). Databases such as Rfam, BSRD and sRNAdb (Kalvari et al., 2017; Li et al., 2012a; Pischimarov et al., 2012) now contain sRNAs from both Gram-negative and Gram-positive species, but most were identified using genomic screens of the model organisms *E. coli* and *S.* Typhimurium. However, sRNAs are short sequences and lack clearly identifiable nucleotide motifs, so they are often difficult to predict bioinformatically (Pichon and Felden, 2008; Sridhar and Gunasekaran, 2013). In addition, while sRNAs are often conserved among closely related bacterial species, they are highly diverse and largely unidentifiable in distantly related genomes if only sequence-based approaches are used.

The computational bioinformatic prediction tools often utilise multiple lines of information to predict sRNA sequences (Li et al., 2012b; Pichon and Felden, 2008; Sridhar and Gunasekaran, 2013). These may include properties of RNA primary sequences and secondary structures, as predicted from structural similarity information and free energy models. Additionally, base composition statistical analysis of genomic sequences or comparative genomics of phylogenetically related bacteria can be utilised. Computational sRNA identification tools can incorporate into the algorithm searches for 'typical' sRNA properties, identified in other known and well conserved sRNA sequences. These include 'orphan' promotor and Rho-independent terminator sequences. Using combinations of the above properties and analyses, a number of sRNA identification programs have been developed and successfully applied to discover sRNA species, including sRNAPredict, RNAz, QRNA, SIPHT and PSoL (Livny et al., 2005; Livny et al., 2008; Rivas and Eddy, 2001; Wang et al., 2006; Washietl et al., 2005). However, despite the advances in bioinformatic sRNA prediction algorithms and tools, they cannot be wholly relied upon and experimental *in vivo* validation is an important process in sRNA identification.

1.5.2 RNA-Sequencing: analysis of whole genome sequencing data for sRNA identification

Advances in RNA-Sequencing (RNA-Seq) technologies have allowed for the large-scale study of bacterial gene expression under a range of conditions and across a range of bacterial species, revolutionising the discovery of bacterial sRNAs. Early studies used microarray analysis for the identification of sRNAs from bacterial transcriptomes (Sharma and Vogel, 2009). Since then, many novel sRNAs have been discovered using RNA-Seq analyses combined with an algorithm, such as APERO and sRNA-Detect, that have been specifically designed for sRNA identification (Leonard et al., 2019; Peña-Castillo et al., 2016). However, visual analysis of RNA-Seq reads mapped to bacterial genomes is also an effective method for sRNA identification; regions encoding putative sRNAs are typically short, highly expressed sequence regions. *Trans*-encoded sRNAs are encoded in intergenic regions and are often followed by a Rho-independent terminator comprising a stable hairpin followed by a string of uridines, making these sRNAs readily identifiable by visual analysis of RNA-Seq data (Rath et al., 2017).

To maximise the discovery of sRNAs using RNA-Seq, cDNA libraries can be specifically enriched for sRNA by including size-selection in the method. Over 500 novel sRNAs were identified in *P. aeruginosa* following RNA-Seq analysis of three different size-selected libraries, generated from cells harvested at different growth phases (Gómez-Lozano et al., 2012). Importantly, unique sRNAs were identified in each different size-selected library.

1.5.3 Experimental identification of sRNAs using co-immunoprecipitation analysis

Many sRNA species require interaction with RNA chaperone proteins, such as Hfq, to facilitate their riboregulatory functions (Sobrero and Valverde, 2012). This knowledge has been used to develop protein-RNA co-immunoprecipitation (Co-IP) assays that have become important experimental strategies for sRNA discovery, most commonly using the Hfq protein as ‘bait’ (Faner and Feig, 2013; Sharma and Vogel, 2009). Hfq Co-IP experiments typically involve the addition of a sequence encoding a triple FLAG-tag epitope to the chromosomally encoded *hfq* gene. This tagging enables purification of Hfq and any bound/associated RNAs from cellular lysates using a specific anti-FLAG antibody (Sittka et al., 2008; Uzzau et al., 2001) (Figure 1.4). The Hfq-associated RNAs can then be purified and sequenced to identify each RNA species; untagged Hfq in the wild-type strain is usually used as a negative control. Hfq-associated RNA species identified by these analyses may include sRNAs as well as their target mRNAs. In a Hfq Co-IP study in

E. coli, 297 mRNAs, sRNAs, and tRNAs were identified in the Hfq-enriched RNA samples (Bilusic et al., 2014). However, limitations to Hfq Co-IP experiments include non-specific binding of RNA species to tagged-Hfq, as well as consideration of growth conditions, as particular transcripts may only be expressed under very specific conditions (Faner and Feig, 2013). It is also important to consider that some Hfq-RNA interactions may be transient and any RNA species that undergo short-lived interactions with chaperone proteins such as Hfq may not be purified by this Co-IP method.

More recently, crosslinking and immunoprecipitation (CLIP) techniques have been developed to identify Hfq-associated RNA species, such as UV-crosslinking and analysis of DNA by high-throughput sequencing (UV-CRAC) and UV-crosslinking, ligation and sequencing of hybrids (UV-CLASH) (Iosub et al., 2018; Kudla et al., 2011; Waters et al., 2017). Hfq UV-CRAC/CLASH involves an initial *in vivo* live cell UV step that crosslinks tagged-Hfq with interacting RNAs. This method directly reflects *in vivo* protein-RNA interactions and improves purification of RNAs that naturally have transient interactions with the RNA chaperone (Figure 1.5). Hfq-RNA complexes are then purified by Co-IP and Hfq-bound RNA molecules (sRNA/mRNA/tRNA) are identified using high-throughput sequencing (UV-CRAC component). In addition, a ligation step can be performed to covalently link RNA molecules that are simultaneously interacting with Hfq, referred to as hybrids (UV-CLASH component). These hybrids may include potential sRNA:mRNA regulatory pairs. In summary, Hfq UV-CRAC/CLASH is a useful tool for the identification of Hfq-associated RNA, and of interacting RNA molecules bound to Hfq. Interestingly, hundreds of novel sRNA-target interactions in *E. coli* that occurred during adaptation to changes in nutrient availability were identified using Hfq UV-CRAC/CLASH (Iosub et al., 2018). Importantly, these methods can be particularly useful in the identification of *cis*-encoded sRNAs, which may be more difficult to identify using techniques such as bioinformatic analysis of RNA-Seq data.

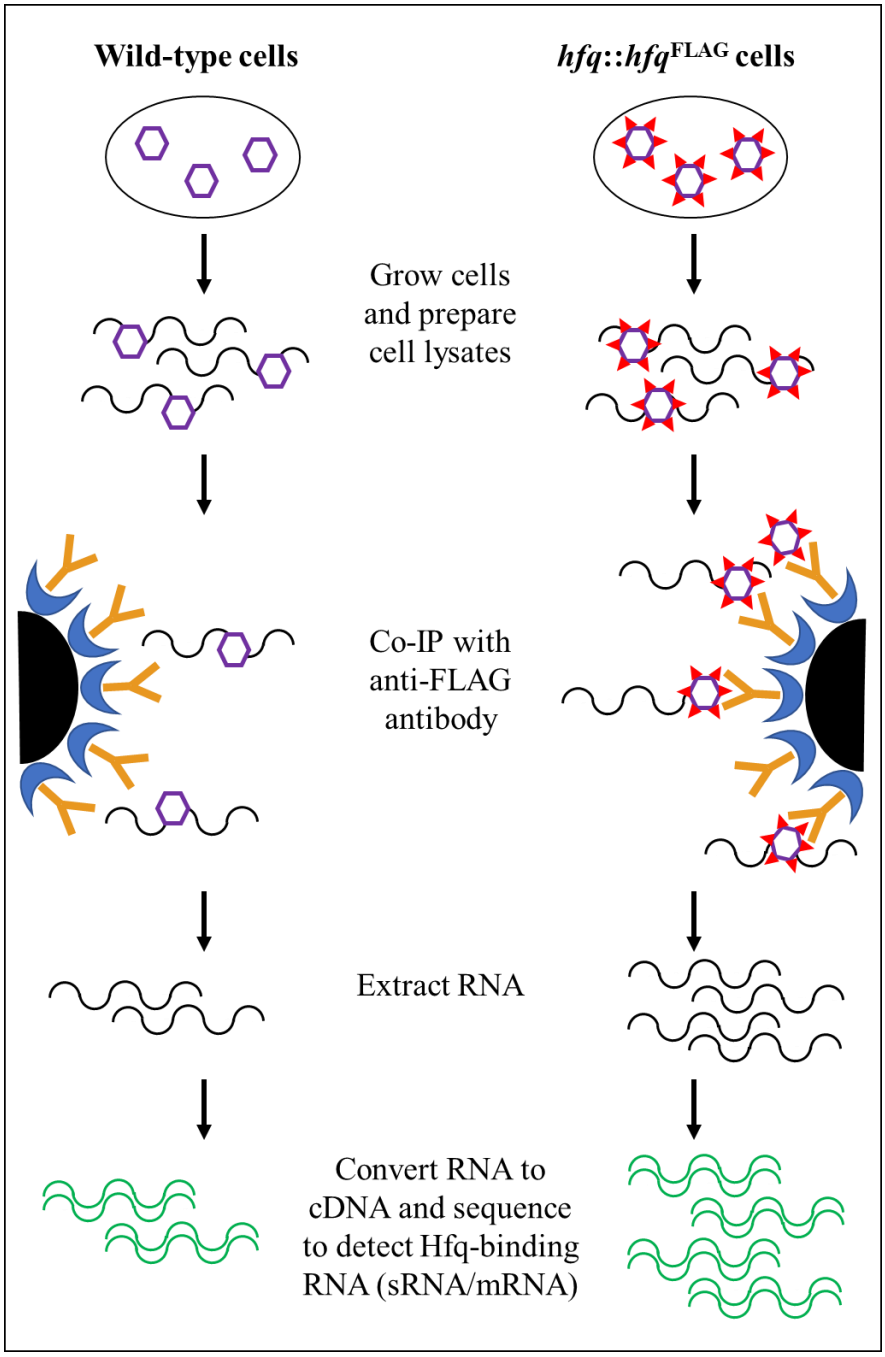


Figure 1.4. Method of identification of Hfq targets using Co-IP. In this example, one strain is expressing FLAG-tagged Hfq (*hfq*^{FLAG}) and the other is the wild-type strain, used as a negative control.

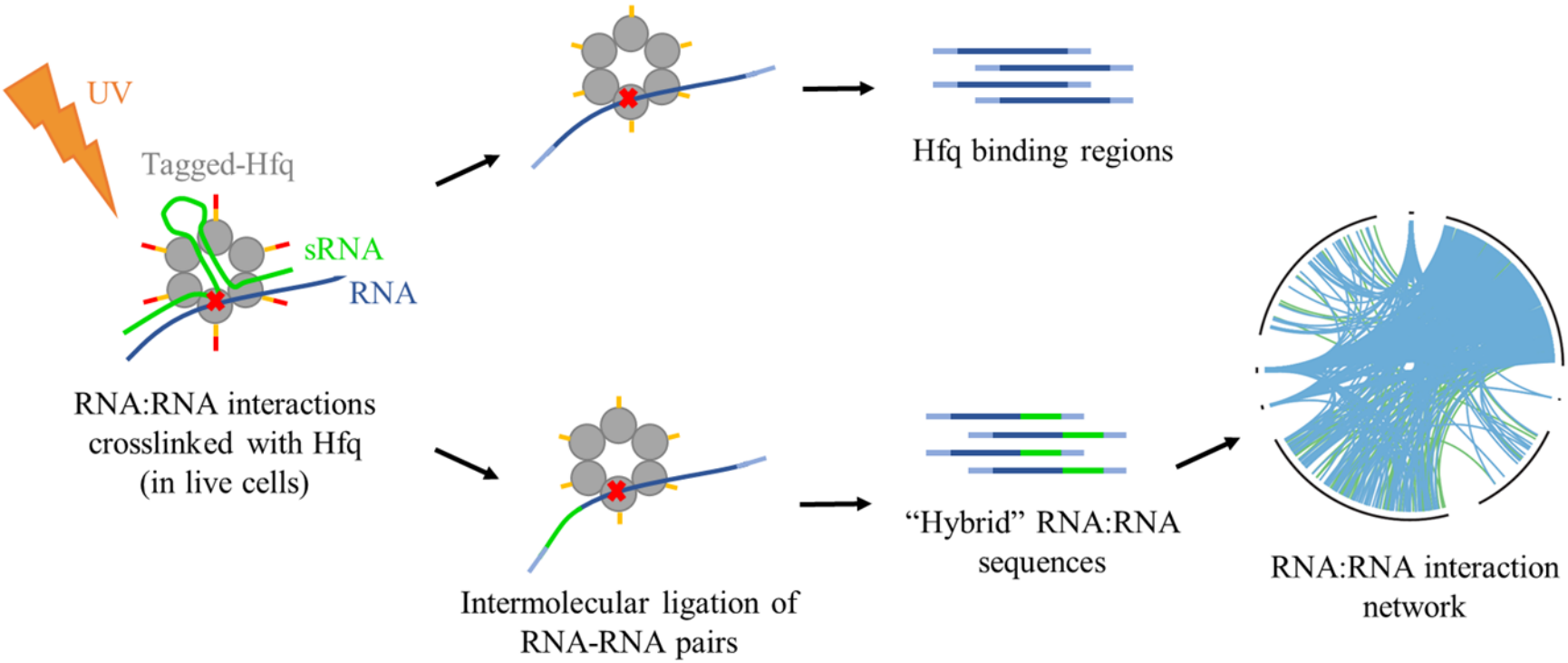


Figure 1.5. Hfq UV-CRAC/CLASH method, used for identification of RNA species interacting with Hfq. Both Hfq-tagged and untagged (control) strains are used in the analysis. The tagged Hfq hexamer (grey, with tags indicated by the attached yellow and red lines) facilitates the interaction between one or more RNA molecules; an sRNA molecule is indicated by the green line and another interacting RNA molecule is indicated by the dark blue line. The interacting RNA molecules are then UV-crosslinked (orange thunderbolt symbol) to the Hfq protein. The Hfq-RNA complexes undergo co-immunoprecipitation steps, before the Hfq-associated RNA molecules are isolated (dark blue lines). Following adapter ligation (light blue lines) to allow for sequencing, the RNA molecules are sequenced and the Hfq-associated RNA identified (UV-CRAC; top panel). To identify directly interacting RNA molecules associated with Hfq, an intermolecular ligation step is performed following UV-crosslinking, resulting in isolation of hybrid RNA sequences (UV-CLASH; lower panel). These hybrid RNA sequences may include mRNA-sRNA pairs (dark blue and green lines), and can be sequenced to generate a Hfq mediated sRNA-RNA interaction network. Modified from Waters et al., 2017.

1.5.4 Methods of identifying the RNA targets of sRNAs

In order to understand bacterial riboregulatory networks and determine the precise function of bacterial sRNAs, it is important to identify the specific RNA target molecules for each sRNA species. A range of computational based sRNA target prediction tools have been developed, including TargetRNA2, IntaRNA, CopraRNA, and RNAPredator (Busch et al., 2008; Eggenhofer et al., 2011; Kery et al., 2014; Wright et al., 2014). These tools generally consider many factors, such as RNA sequence and structure, in order to predict specific mRNA targets and can also consider binding energy between the RNA pair (Pain et al., 2015). Recently an sRNA Target Prediction Organising Tool (SPOT), has been developed to combine a number of available target prediction tools with experimental data to streamline the prediction of the RNA molecules targeted by an sRNA (King et al., 2019).

Although computational sRNA prediction tools can be helpful, they generally have low accuracy with high rates of false positives, likely due to the complexity of factors involved in sRNA and target mRNA binding (Pain et al., 2015). Due to this, experimental approaches for target prediction are more favourable. These include analysis of RNA expression data from an sRNA mutant and its corresponding overexpression strain, both compared to expression data obtained from the wild-type strain, and pulse expression of sRNAs in combination with transcriptome analysis. Global small non-coding RNA target identification by ligation and sequencing (GRIL-Seq) is another experimental method for identification of sRNA targets; this method involves co-expression of a specific sRNA with a T4 RNA ligase protein followed by *in vivo* ligation of sRNAs to their targets (Han et al., 2016).

Transcriptomic analyses can reveal genes with altered expression in sRNA mutants and in sRNA overexpression strains. However, these analyses also detect pleiotropic effects on gene regulation and therefore interpretation is limited, because genes can be affected both directly and indirectly by the presence or overabundance of a particular sRNA. To overcome some of these limitations, pulse expression analyses can be used to overexpress an sRNA of interest for only a short period of time (usually 10-15 minutes), so that only rapid changes in RNA levels are detected (Massé et al., 2005; Papenfort et al., 2006). Thus, pulse expression experiments are more likely to show gene expression changes that are a direct response to increasing sRNA expression (Hör et al., 2018).

1.5.5 Experimental validation of sRNAs and sRNA-target interactions

It is important to experimentally validate predicted sRNAs and their interactions with other RNA. Northern blotting is a well-established, experimental method for the separation and visualisation of RNA transcripts using probes that are complementary to part of, or the entire, sequence of interest. The specific size and the abundance of RNA transcripts under different conditions can be analysed using this method (López-Gomollón, 2011).

In order to determine the precise 5′ transcriptional start site of sRNAs, techniques such as fluorescent primer extension and rapid amplification of cDNA ends (5′ RACE) are commonly used (Fekete et al., 2003; Liu et al., 2018; Lloyd et al., 2005). These techniques can also be used to estimate the size of sRNAs, in conjunction with knowledge of the predicted transcriptional terminator site. Knowledge of precise sRNA 5′ end, and to some extent the 3′ end, is important for experiments involving recombinant expression of specific sRNAs.

Experimental evidence of interactions between sRNAs and mRNA targets is critical for predicting sRNA function. Binding analyses such as biolayer interferometry (BLItz), electrophoretic mobility shift assay (EMSA) or surface plasmon resonance (SPR) can be applied to determine specific RNA interactions, and have been used to determine RNA interactions with the Hfq protein and the *hfq* transcript (Dimastrogiovanni et al., 2014; Morita et al., 2012; Salim and Feig, 2010).

1.6 Aims of this study

While a number of *P. multocida* virulence factors have been identified, little is known about how the expression of these virulence factors is regulated. In this study, we aimed to unravel the role of the Hfq chaperone protein and sRNAs in the *P. multocida* regulatory network, specifically focusing on riboregulatory functions that impact *P. multocida* pathogenesis. The role of *P. multocida* Hfq was first investigated using a *hfq* TargeTron mutant constructed in the virulent A:1 strain VP161 (Chapter 2). The effect of *hfq* inactivation on *P. multocida* and its virulence was assessed using phenotypic and *in vivo* analyses. To identify genes and proteins whose expression were controlled by Hfq and sRNAs, the RNA expression and protein production levels in wild-type and *hfq* mutant strains were analysed by transcriptomic and proteomic methods.

There is currently very limited data on *P. multocida* sRNAs. Our experiments focused on identification of the range of sRNA molecules produced by *P. multocida* (Chapter 3). Bioinformatic analysis of whole genome RNA-Sequencing data was utilised to maximise the range of sRNAs identified and a range of different growth conditions were used. To augment these findings, Hfq co-immunoprecipitation analyses were employed to identify the range of Hfq-

associated RNA species, including sRNA and mRNA sequences (Chapter 4). These Co-IP analyses included Hfq UV-crosslinking, ligation and sequencing of hybrids (CLASH), which allowed for the identification of specific RNA:RNA pairs in simultaneous association with Hfq.

Finally, a number of identified putative sRNAs were selected for experimental validation and characterisation (Chapter 5). Specifically, sRNAs predicted to be involved in the regulation of *P. multocida* virulence phenotypes were selected for investigation. Together, these results contribute novel insights to the understanding of *P. multocida* virulence via riboregulatory mechanisms.

Chapter Two: The RNA-binding chaperone Hfq is an important global regulator of gene expression in *Pasteurella multocida* and plays a crucial role in production of a number of virulence factors including hyaluronic acid capsule

Based on:
Mégroz, M., Kleifeld, O., Wright, A., Powell, D., Harrison, P., Adler, B., Harper, M., Boyce, J. D. The RNA-binding chaperone Hfq is an important global regulator of gene expression in *Pasteurella multocida* and plays a crucial role in production of a number of virulence factors including hyaluronic acid capsule. Infect Immun. **84(5):**1361-70.

Please Note: since this publication, the *P. multocida* VP161 genome has been re-sequenced and re-annotated and genes have been assigned new locus tags (Smallman, Boyce Laboratory, manuscript in preparation). All other chapters will use the new VP161 annotation.

2.1. Introduction

Small noncoding RNA (sRNA) molecules play essential roles in regulating the production of a wide range of proteins, including those involved in virulence; quorum sensing; and the metabolism of carbon, amino acids, and iron (Lenz et al., 2004; Massé and Gottesman, 2002; Romby et al., 2006; Sharma et al., 2011; Sonnleitner et al., 2009). Typically, sRNA molecules are between 50 and 400 nucleotides in length and act primarily by interacting with one or more target mRNAs via short imperfect base pairing (Sobrero and Valverde, 2012; Vogel and Luisi, 2011). Protein production can be controlled by sRNAs via multiple mechanisms, including inhibition of translation, activation of translation, and alteration of transcript degradation rates (Bobrovskyy and Vanderpool, 2013; Gottesman and Storz, 2011). The regulation of bacterial protein expression via sRNAs often requires the activity of Hfq, an RNA-binding chaperone protein that belongs to the Sm-like RNA-binding protein family and displays a highly conserved core sequence (Moller et al., 2002; Sobrero and Valverde, 2012). Hfq monomers form a homohexameric ring-shaped structure that preferentially binds to A/U-rich sequences on target RNA molecules and mediates the various posttranscriptional regulatory mechanisms of sRNAs (Brennan and Link, 2007; De Lay et al., 2013; Schumacher et al., 2002). Hfq can also directly protect sRNA molecules from RNase E-mediated degradation (Folichon et al., 2003) and can have an sRNA-independent role in the regulation of mRNA decay, via direct interaction with poly(A) polymerase (Mohanty et al., 2004).

Hfq homologues have been identified in a number of Gram-positive and Gram-negative bacterial species, including many pathogenic species (Chao and Vogel, 2010). Most *hfq* mutants display a broad range of pleiotropic phenotypes, suggesting a global role for Hfq in bacterial physiology (Robertson and Roop, 1999; Sobrero and Valverde, 2012; Tsui et al., 1994). These pleiotropic effects have been observed more commonly in Gram-negative species, including the pathogens *Brucella abortus*, *S.* Typhimurium, *Yersinia pestis*, *Moraxella catarrhalis*, *Neisseria meningitidis*, and *Vibrio cholerae*, but also in some Gram-positive species such as *Listeria monocytogenes* (Attia et al., 2008; Christiansen et al., 2004; Ding et al., 2004; Fantappiè et al., 2009; Figueroa-Bossi et al., 2006; Geng et al., 2009; Robertson and Roop, 1999). Many *hfq* mutants show reduced growth rates, reduced virulence, and increased susceptibility to host defence mechanisms and environmental stresses (Chao and Vogel, 2010; Sobrero and Valverde, 2012).

Pasteurella multocida is an encapsulated, Gram-negative, facultative anaerobe that is the causative agent of a number of important animal diseases, including fowl cholera (Boyce et al., 2010). Fowl cholera can affect most avian species, including chickens, turkeys, ducks, and wild waterfowl,

with infections resulting in high mortality rates and major economic losses to poultry industries worldwide (Blackall and Miflin, 2000; Wilson and Ho, 2013). *P. multocida* is a heterogeneous species, with strains being classified into five serogroups (A, B, D, E, and F) based on capsular composition (Carter, 1952) and into eight lipopolysaccharide (LPS) genotypes (L1 through to L8) based on the LPS outer core biosynthesis genes (Harper et al., 2015). A number of *P. multocida* virulence factors have been defined, including the polysaccharide capsule, LPS, filamentous hemagglutinin adhesins, iron-sequestering systems, and sialic acid uptake (Bosch et al., 2002a; Harper et al., 2012; Tatum et al., 2005; Vimr, 2013). The polysaccharide capsule is a critical virulence factor of *P. multocida* that allows the bacteria to evade host immune defence mechanisms (Boyce and Adler, 2000; Chung et al., 2001). Very little is known about how the expression of these virulence factors is regulated, as only one *P. multocida* regulatory protein, Fis, has been identified and characterized. Fis has been shown to be essential for the production of capsule, and a *fis* mutant displayed highly reduced expression levels of many genes, including those within the capsule biosynthesis locus and *pfhB2*, encoding filamentous hemagglutinin (Steen et al., 2010).

In this study, the role of Hfq in *P. multocida* gene expression and protein production was characterized by using a *hfq* mutant constructed in the *P. multocida* fowl cholera strain VP161. As Hfq is a crucial modulator of sRNA action, an understanding of the role of Hfq provides an initial indication of the significance of sRNAs in the regulation of *P. multocida* genes and will underpin future work on specific sRNAs in this bacterium. The VP161 *hfq* mutant displayed highly reduced hyaluronic acid (HA) capsule production as well as reduced *in vivo* fitness. RNA sequencing (RNA-Seq) and high-throughput quantitative liquid proteomics analyses indicated that Hfq plays an important role in controlling the expression of >100 genes/proteins. In particular, genes/proteins belonging to the capsule biosynthetic locus were expressed at reduced levels in the *hfq* mutant, consistent with the reduced levels of HA capsule production. These analyses also revealed that Hfq plays a role in the production of filamentous hemagglutinin and proteins involved in LPS biosynthesis. Taken together, these data indicate that Hfq plays an important role in the regulation of key *P. multocida* virulence factors and is essential for full *in vivo* fitness.

2.2. Materials and Methods

2.2.1. Bacterial strains, plasmids, and culture conditions

The bacterial strains and plasmids used in this study are listed in Table 2.1. *Escherichia coli* DH5α was grown in lysogeny broth (LB), and *P. multocida* strains were grown in either heart infusion (HI) or brain heart infusion (BHI) broth (Oxoid). Liquid cultures were incubated at 37°C with shaking (200 rpm). Solid media were obtained by the addition of 1.5% agar. When required, the media were supplemented with antibiotics at the following concentrations: 10 µg/ml (*E. coli*) or 2.5 µg/ml (*P. multocida*) tetracycline, and 50 µg/ml kanamycin (*P. multocida*).

2.2.2. Measurement of bacterial growth rates in rich medium in flasks

Cultures of *P. multocida* strains grown overnight were subcultured at a 1:50 dilution in BHI or HI broth and grown until cultures reached an optical density at 600 nm (OD_{600}) of 0.2 to 0.3. A 1-ml aliquot (standardized to an OD_{600} of 0.2) of each culture was transferred to 50 ml of sterile BHI broth and incubated at 37°C with shaking. To determine the growth rate, optical density measurements of the liquid cultures (at 600 nm) were taken at regular intervals by using the WPA CO8000 Biowave cell density meter, until the late-exponential growth phase was reached.

2.2.3. Ethics statement

All animal experiments were carried out in accordance with the provisions of the Prevention of Cruelty to Animal Act, 1986; the *Australian Code of Practice for the Care and Use of Animals for Scientific Purposes* (NHMRC, 2004); and Monash University Animal Welfare Committee guidelines and policies. The protocols were approved by the Monash Animal Research Platform 2 (MARP-2) Animal Ethics Committee (AEC) of Monash University (AEC number MARP/2011/066, Understanding How *Pasteurella multocida* Causes Disease in Animals).

2.2.4. DNA manipulations

Restriction digests, ligations, and PCR amplifications were performed using enzymes obtained from NEB or Roche, according to the manufacturer's instructions. Plasmid DNA was prepared by using the NucleoSpin plasmid kit (Macherey-Nagel), while genomic DNA was prepared by using the HiYield genomic DNA minikit (Real Genomics). PCR amplification of DNA was performed by using *Taq*DNA polymerase (Roche) or Phusion high-fidelity DNA polymerase (NEB) on an Eppendorf Mastercycler. Amplified DNA was purified by using the NucleoSpin gel or PCR cleanup kit (Macherey-Nagel). The primers used in this study are listed in Table 2.2. Column-purified DNA samples were quantified by using the NanoDrop 1000 spectrophotometer (Thermo

Scientific) and/or by gel electrophoresis. Sequencing reactions were performed by using Applied Biosystems Prism BigDye terminator mix, version 3.1. Electropherograms were generated on a capillary platform Applied Biosystems 3730 genetic analyzer and analyzed by using Vector NTI Advance, version 11.5 (Life Technologies).

Table 2.1. Bacterial strains and plasmids used in this study.

Strain or plasmid	Relevant description	Source or reference
Strains		
E. coli		
DH5α	*deoR endA1 gyrA96 hsdR17*(r_K^- m_K^+) *recA1 relA1 supE44 thi-1* (*lacZYA-argFV169*) φ80*lacZ*ΔM15, F^-	Bethseda Research Laboratories
AL1995	DH5α harbouring pAL953; Kan^R $Spec^R$	(Harper et al., 2013)
AL2224	DH5α harbouring pAL99T; Tet^R	(Harper et al., 2013)
AL2227	DH5α harbouring pAL1069; Kan^R $Spec^R$	This study
AL2523	DH5α harbouring pAL1108; Tet^R	This study
AL2760	DH5α harbouring pAL1104; Kan^R $Spec^R$	This study
P. multocida		
VP161	Serotype A:1: Vietnamese isolate from chickens	(Wilkie et al., 2000)
AL2234	VP161 *hyaD* TargeTron mutant; Kan^R	This study
AL2521	VP161 *hfq* TargeTron mutant; Kan^R	This study
AL2526	AL2521 containing pAL1108; Kan^R, Tet^R	This study
AL2527	AL2521 containing pAL99T; Kan^R, Tet^R	This study
Plasmids		
pAL99T	*P. multocida-E. coli* expression plasmid; contains the constitutive *tpiA* promoter upstream of the cloning site; derivative of pAL99; Tet^R	(Harper et al., 2013)
pAL953	*P. multocida*-specific TargeTron plasmid; intron targeted to the *gatA* gene; Kan^R $Spec^R$	(Harper et al., 2013)
pAL1069	*P. multocida* TargeTron plasmid for inactivation of *hyaD*; Kan^R $Spec^R$	This study
pAL1104	*P. multocida* TargeTron plasmid for inactivation of *hfq*; Kan^R $Spec^R$	This study
pAL1108	Wild-type *hfq* cloned into pAL99T; Tet^R	This study

Table 2.2. Primers used in this study.

Primer	Sequence (5′ – 3′)	Description
BAP2067	GGAAGGAACAGTTTCTCTGGATTG	Forward primer located within VP161 *hyaD*
BAP2679	TTGTGTGGAATTGTGAGCGGA	Primer located upstream of the *tpi* promoter and MCS[a] on pAL99T
BAP7153	AAAAAAGCTTATAATTATCCTTACTGAACCAGACTGTGCGC CCAGATAGGGTG	TargeTron IBS primer specific for inactivation of *hyaD*
BAP7154	CAGATTGTACAAATGTGGTGATAACAGATAAGTCCAGACTG TTAACTTACCTTTCTTTGT	TargeTron EBS1d primer specific for inactivation of *hyaD*
BAP7155	TGAACGCAAGTTTCTAATTTCGGTTTTCAGTCGATAGAGGA AAGTGTCT	TargeTron EBS2 primer specific for inactivation of *hyaD*
BAP7389	TGAACGCAAGTTTCTAATTTCGATTCTTGTTCGATAGAGGA AAGTGTCT	TargeTron EBS2 primer specific for inactivation of *hfq*
BAP7390	AAAAAAGCTTATAATTATCCTTAACAAGCTCAGATGTGCGC CCAGATAGGGTG	TargeTron IBS primer specific for inactivation of *hfq*
BAP7391	CAGATTGTACAAATGTGGTGATAACAGATAAGTCTCAGATT GTAACTTACCTTTCTTTGT	TargeTron EBS1d primer specific for inactivation of *hfq*
BAP7400	CAGAGGATCCTGTGTTTAGCTAGTTGTATC	Forward primer located upstream of VP161 *hfq*; contains BamHI site
BAP7401	AGATGTCGACCACTTGCAATAAATGTGTTC	Reverse primer located downstream of VP161 *hfq*; contains SalI site
EBS Universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	Universal TargeTron primer, located ~250 bp from the 5′ end of the intron

2.2.5. Construction of *P. multocida hfq* and *hyaD* mutants

The TargeTron (Sigma-Aldrich) mutagenesis system was used to separately inactivate *hfq* and, as a control for HA assays, the capsular biosynthesis gene *hyaD* in *P. multocida* A:1 strain VP161, as described previously (Harper et al., 2013). The group II intron within *P. multocida* TargeTron plasmid pAL953 (Table 2.1) was retargeted to *hfq* or *hyaD* according to the TargeTron user manual and with specific primers designed by using the TargeTron design site (Sigma-Aldrich) (Table 2.2). The retargeted TargeTron plasmids pAL1104 (*hfq* targeted) and pAL1069 (*hyaD* targeted) were then used to separately transform *P. multocida* VP161 by electroporation, and transformants were selected on solid agar containing kanamycin. Following transformation, strains were cured of free-replicating plasmid as described previously (Harper et al., 2013), and putative *hfq* and *hyaD* mutants were identified by patching for $Kan^R/Spec^S$ colonies. Colony PCRs using the TargeTron-specific primer EBS universal, together with a *hfq*- or *hyaD*-specific primer (BAP7400 and BAP2067, respectively) (Table 2.2), were used to initially identify TargeTron mutants (data not shown). To confirm that *P. multocida hfq* or *hyaD* mutants contained only a single TargeTron insertion, Southern hybridization was performed (data not shown), using a TargeTron intron-specific digoxigenin (DIG)-labelled probe. In addition, the exact position of the intron within the genome was determined by direct sequencing using genomic DNA isolated from each TargeTron mutant as the template with the intron-specific EBS universal primer (Table 2.2). Two mutants that generated unambiguous sequencing data identical to data for the correct target gene sequence (immediately adjacent to the intron insertion site) were selected for further study and designated AL2521 (VP161 *hfq* mutant) and AL2234 (VP161 *hyaD* mutant).

2.2.6. Complementation of the *hfq* mutant

For complementation of the *hfq* mutant, an intact copy of the *hfq* gene was amplified from VP161 genomic DNA by using primers BAP7400 and BAP7401 that contained BamHI and SalI restriction sites, respectively (Table 2.2). After digestion with BamHI and SalI, the amplified fragment was cloned into BamHI and SalI-digested pAL99T (Table 2.1), such that the *P. multocida* constitutive promoter *tpiA* would drive the transcription of *hfq*. The ligated products were then used to transform *E. coli* strain DH5α. *E. coli* transformants containing the correct plasmid were identified by colony PCR using pAL99T-specific and *hfq*-specific primers BAP2679 and BAP7401, respectively. DNA sequencing was used to confirm the fidelity of the *hfq* sequence, and one correct recombinant plasmid was selected and designated pAL1108. The complementation plasmid pAL1108 and the vector pAL99T were then used to separately

transform the VP161 *hfq* mutant (AL2521) via electroporation, generating strains AL2526 and AL2527, respectively (Table 2.1).

2.2.7. Quantitative hyaluronic acid capsule assay

The amount of HA capsular material produced by each *P. multocida* strain was determined as described previously (Chung et al., 2001). Student's unpaired *t* test was used to assess the differences in HA production between strains.

2.2.8. Serum sensitivity assays

The sensitivity of *P. multocida* strains and *E. coli* DH5α to the bactericidal complement activity of chicken serum was determined as described previously (Chung et al., 2001), with the following modifications. Approximately 5.0 x 10^5 CFU of each strain was incubated in 85% chicken serum under static growth conditions in a 96-well plate.

2.2.9. Competitive *in vivo* growth assays

Competitive growth assays in mice were performed to quantify the relative *in vivo* fitness of the *hfq* mutant and complemented strains compared with wild-type strain VP161, as described previously (Harper et al., 2003). To identify each bacterial population, after overnight incubation of input and output samples on HI medium, individual colonies were patched onto HI medium and either HI medium containing kanamycin (to select for the *hfq* mutant) or HI medium containing kanamycin and tetracycline (to select for the *hfq* mutant containing the plasmid). Mutants with a competitive index (CI) of <0.5 were identified as having significantly reduced *in vivo* fitness, and Student's unpaired two-tailed *t* test was also used to assess the statistical differences in relative CIs between each of the tested strains.

2.2.10. Virulence assays in chickens

The virulence of the *hfq* mutant in chickens was determined by direct challenge. Two groups of 10 Hy-Line Brown female chickens (14 weeks of age) were challenged via the intratracheal (IT) route to simulate the natural course of infection using the *hfq* mutant (actual dose 5.0 x 10^5 CFU) or the wild-type strain (actual dose 1.5 x 10^6 CFU). Two other groups were challenged via the intra-muscular (IM) route (pectoral muscle) with doses of 50 CFU of the *hfq* mutant or 150 CFU of the wild-type strain. Chickens were observed closely for clinical signs of fowl cholera over a 60-h time period and euthanized when deemed incapable of survival. Postmortem samples were taken from appropriate sites of a representative sample of birds and recovered colonies used to generate template DNA for PCR to confirm both the presence of *P. multocida* and the genotype of the recovered strains.

2.2.11. RNA extraction and purification, rRNA depletion, library preparation, and high-throughput RNA sequencing of the *hfq* mutant

P. multocida strains AL2521 (VP161 *hfq* mutant) and VP161 (wild type) were each grown in BHI broth to the mid-exponential growth phase (OD_{600} = 0.6; ~1.0 x 10^9 CFU/ml). A total of 2.5 ml of ice-cold killing buffer (containing 0.05 M Tris-HCl [pH 7.5], 15 mg/ml sodium azide, and 0.6 mg/ml chloramphenicol) was added to each 25-ml culture, and the mixed sample was cooled on ice before centrifugation at 9,000 x *g* for 10 min (4°C). Pelleted cells were then resuspended in 1 ml RNAlater stabilization reagent (Qiagen) and incubated for 5 min at room temperature before centrifugation at 5,000 x *g* for 10 min. Total RNA was purified by using the RNeasy minikit (Qiagen) with DNase treatment according to the manufacturer's instructions. RNA samples were quantified and analyzed for DNA contamination by using the Qubit kit (Invitrogen) and visualized by gel electrophoresis. rRNA was removed by using the Ribo-Zero magnetic kit (Illumina) according to the manufacturer's instructions. The TruSeq RNA sample preparation kit (Illumina) was used to construct cDNA libraries, with library validation, normalization, and pooling performed by Micromon Services (Monash University). All libraries were combined and sequenced on a single lane of a HiSeq 2000 instrument (Macrogen, South Korea). Mapping of the RNA-Seq reads to the draft VP161 genome was carried out by using CLC Genomics Workbench v 7.0 (CLC Bio), and statistical analyses were carried out by using voom and limma as described previously (Henry et al., 2015). Raw combined data sets of 50,071,021 and 47,130,114 total paired reads were generated for AL2521 and VP161, respectively. Of these, 17,718,215 and 14,903,935 paired reads (for AL2521 and VP161, respectively) aligned uniquely to the *P. multocida* VP161 genome (~30% unambiguous reads for each combined data set), giving high coverage of the transcriptome. Genes were identified as being differentially expressed if they displayed a ≥2.0-fold (≥1.0-$\log_2$) change in expression across the triplicate replicates at a false discovery rate (FDR) of <0.05. The differentially expressed VP161 genes identified were mapped to their closest orthologue in the fully annotated reference genome of *P. multocida* strain Pm70 (May et al., 2001) so that overrepresented gene ontology groups and pathway associations could be determined via BioCyc (Caspi et al., 2014), as we reported previously (Henry et al., 2015).

2.2.12. Quantitative proteomics and mass spectrometry

Cultures of each *P. multocida* strain were grown to either the early-exponential (OD_{600} = 0.2) or mid-exponential (OD_{600} = 0.6) growth phase and then centrifuged at 9,400 x *g* for 10 min. The pelleted cells were washed twice in 100 mM Tris-HCl (pH 7.5) and lysed via the addition of 0.35 ml of lysis buffer (4% SDS, 100 mM dithiothreitol [DTT], 100 mM Tris-HCl [pH 7.5]) and

incubation for 10 min at 99°C. The protein concentration was determined by using the Bradford assay (Bradford, 1976). Proteins (150 µg/sample) were purified and trypsin digested by using a filter-aided sample preparation (FASP) protein digestion kit (Expedeon) as described previously (Wiśniewski et al., 2009), except that triethyl ammonium bicarbonate was used instead of ammonium bicarbonate. The digested proteins were isotopically labelled by demethylation using heavy and light formaldehyde as described previously (Boersema et al., 2009). Heavy and light samples were then mixed and desalted by using an Empore 4-mm/1-ml extraction disk cartridge (C18-SD) (3M). Desalted samples were concentrated under a vacuum and resuspended in 0.1% formic acid and 2% acetonitrile. Peptides were resolved and analyzed by nano-liquid chromatography (LC)–tandem mass spectrometry using a Dionex ultrahigh- performance liquid chromatography system coupled to a Thermo Scientific Q-Exactive Orbitrap mass spectrometer (located at the Monash Biomedical Proteomics Facility). Proteins were identified and quantified with MaxQuant (Cox and Mann, 2008) and Perseus software. For differential expression testing, the protein expression ratios were log transformed. These transformed ratios were roughly normally distributed, so limma (Ritchie et al., 2015) was used to fit and test for the significance of differential expression. Those proteins showing a ≥2.0-fold (≥1.0-$\log_2$) change in expression across the triplicate replicates at an FDR of <0.05 were considered to be differentially expressed.

2.2.13. Accession numbers

The RNA-Seq data are available at the NCBI Gene Expression Omnibus under accession number GSE77721. The proteomics data have been deposited in ProteomeXchange via the PRIDE database under accession number PXD003586.

2.3. Results

2.3.1. The *hfq* mutant produces reduced hyaluronic acid capsule

To investigate the role of Hfq in *P. multocida*, a *hfq* mutant was generated in the highly virulent strain VP161. The mutant grew indistinguishably from the wild-type strain in BHI liquid broth at 37°C *in vitro* (see Figure S1 in Appendix B), but on solid medium, it produced smaller and less mucoid colonies than those of wild-type strain VP161, consistent with reduced HA capsular expression. The amount of HA was quantified in mid-exponential-phase cultures of wild-type strain VP161, the *hfq* mutant (AL2521), the complemented *hfq* mutant (AL2526), and the *hfq* mutant containing the vector only (AL2527). The *hfq* mutant produced significantly less HA than did the wild-type strain, and HA production was restored to near-wild-type levels when the mutant

was complemented with an intact copy of *hfq* (Figure 2.1A). The control *hyaD* mutant produced no detectable HA (data not shown).

The amount of HA capsule produced by the *hfq* mutant was also assessed at different growth phases *in vitro*. The wild-type strain produced small amounts of HA at the early-exponential growth phase, maximal amounts of HA at the mid-exponential growth phase, and intermediate levels of HA at the late-exponential growth phase (Figure 2.1B). No detectable HA was produced by the *hfq* mutant at the early-exponential growth phase, and significantly reduced levels of HA were produced at the mid- and late-exponential growth phases (Figure 2.1B). Again, the *hyaD* acapsular control strain produced no detectable HA at any of the tested growth phases (data not shown). These data indicate that capsule production in *P. multocida* is growth phase dependent, and inactivation of *hfq* results in significantly reduced capsule expression during all tested growth phases *in vitro*.

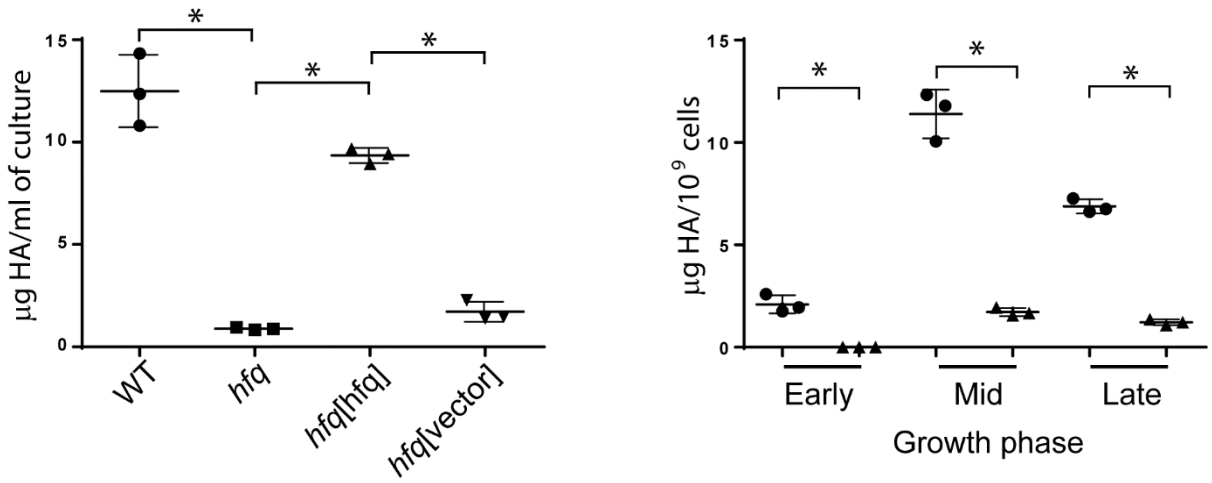


Figure 2.1. Hyaluronic acid (HA) polysaccharide capsule produced by *P. multocida* strains. **(A)** Amounts of HA produced at the mid-exponential growth phase by wild-type strain VP161 (WT), the *hfq* mutant, the complemented *hfq* mutant (*hfq*[hfq]), and the *hfq* mutant with the empty vector (*hfq*[vector]). **(B)** Amounts of HA detected (per 10^9 cells) in wild-type (●) and *hfq* mutant (▲) cells at the early-exponential (OD_{600} = 0.20 to 0.25), mid-exponential (OD_{600} = 0.65 to 0.75), and late-exponential (OD_{600} = 1.10 to 1.40) growth phases. The average amount of HA under each condition is indicated by the horizontal bars, ± standard deviations of the means. *, *P* value of < 0.05 as determined by an unpaired *t* test.

2.3.2. Sensitivity of *hfq* mutant to chicken serum

Previous work in our laboratory suggested that capsule is essential for *P. multocida* resistance to the bactericidal activity of complement in chicken serum (Chung et al., 2001). As HA capsule production was significantly reduced in the *P. multocida hfq* mutant, serum sensitivity assays were performed in 85% chicken serum to determine the level of serum resistance of the mutant. The *hfq* mutant was not serum sensitive, however the level of growth of the mutant was reduced by 2.5-fold in the presence of active chicken serum (Figure 2.2), suggesting a role for Hfq in the ability of *P. multocida* to resist the bactericidal effects of complement in chicken serum.

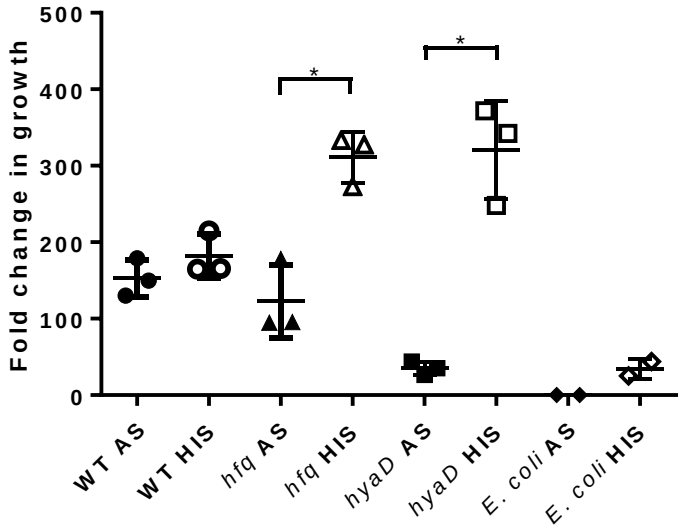


Figure 2.2. Sensitivity of *P. multocida* strains to 85% chicken serum. The fold-change in growth of wild-type strain VP161 (WT), the *hfq* mutant, the *hyaD* mutant and *E. coli* DH5α following growth of strains in active serum (AS) or heat inactivated serum (HIS) are shown. The average fold change in growth for each strain is indicated by the horizontal bars, ± standard deviations of the mean. *, *P* value of < 0.05 as determined by an unpaired *t* test.

2.3.3. The *hfq* mutant displays reduced *in vivo* fitness in mice and reduced virulence in chickens

As the HA capsule is a known virulence factor of *P. multocida* serogroup A strains (Chung et al., 2001), competitive growth assays were performed in mice to compare the *in vivo* fitness of the wild-type strain with that of the *hfq* mutant, the *hfq* mutant complemented with an intact copy of *hfq*, and the *hfq* mutant containing the vector only. The *hfq* mutant displayed a 4-fold reduction in *in vivo* growth compared to that of the wild-type parent strain VP161 (CI = 0.25 ± 0.05) (Figure 2.3A). Importantly, this loss of *in vivo* fitness was restored to near-wild-type levels when the *hfq* mutant was provided with an intact copy of *hfq* in *trans* (CI = 0.68 ± 0.06) but not when the *hfq* mutant was provided with vector only (Figure 2.3A). These data indicate that the *P. multocida* VP161 Hfq is important for *in vivo* fitness in mice.

To assess the importance of *hfq* in the virulence of *P. multocida* in a natural host, the VP161 *hfq* mutant and the wild-type strain (VP161) were separately used to challenge groups of chickens via either the IM or IT route. Although the administered doses (CFU) of wild-type and *hfq* mutant strains were not equivalent for both the IT and IM infection (~3-fold higher dose of wild-type compared to mutant), the administered doses were all well above the infective dose (ID) for the *P. multocida* VP161 strain (ID_{50}=<10 CFU) (Hatfaludi et al., 2012). Therefore, it is likely that the virulence defect observed during these infection studies is due to the inactivation of *hfq*.

All chickens challenged with the wild-type strain via the IM route developed late-stage signs of fowl cholera within 23 h (average time to death 20 h; Figure 2.3C). Seven of the ten chickens inoculated with the *hfq* mutant via the IM route also developed late-stage signs of fowl cholera but over a longer time period (average time to death 38 h; $P < 0.0001$). The three chickens that survived challenge with the *hfq* mutant showed only minor fowl cholera signs and recovered from infection by the end of the trial. A similar result was observed following IT infection; all chickens infected with the wild-type strain developed late stage fowl cholera while only five of the birds infected with the *hfq* mutant showed significant disease signs (Figure 2.3B). PCR analysis of samples taken from birds that had succumbed to fowl cholera infection after challenge with the *hfq* mutant, confirmed the retention of the *hfq* mutation (data not shown). Taken together, these data indicate that Hfq is required for full virulence and *in vivo* fitness of *P. multocida*. However, the VP161 *hfq* mutant can still cause fowl cholera disease in some birds, albeit over a longer time period compared to wild-type VP161.

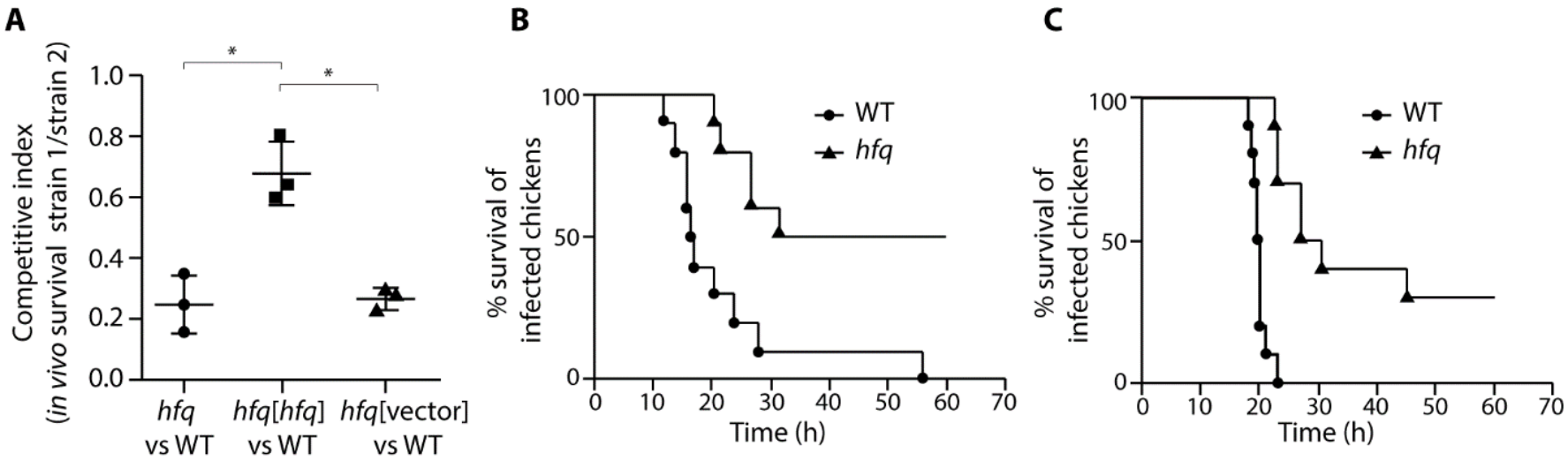


Figure 2.3. The *hfq* mutant is attenuated in mice and in chickens. (A) Competitive growth indices for the *P. multocida* strains in a mouse infection model. Each data point shows the individual CI for one pair of strains in one mouse. The average CI for each pair of competing strains is indicated by the horizontal bars, ± standard deviations of the mean. **(B)** Kaplan-Meier curve showing survival times of chickens infected with wild-type VP161 or the *hfq* mutant via the intra-tracheal route. **(C)** Kaplan-Meier curve showing survival times of chickens infected with wild-type VP161 or the *hfq* mutant via the intra-muscular route. * indicates a *P* value of <0.05 as determined by a two-tailed unpaired *t* test.

2.3.4. Global RNA expression changes in the *P. multocida hfq* mutant

Hfq is an RNA chaperone that has been shown in many other bacterial species to be essential for the action of most *trans*-encoded regulatory sRNA molecules (Chao and Vogel, 2010). Hfq and associated sRNAs can act at the level of mRNA stability or at the level of translational efficiency (Schumacher et al., 2002). To identify changes in transcript abundance, the transcriptomes of the wild-type strain and the *hfq* mutant (grown to the mid-exponential growth phase) were compared by using high-throughput RNA-Seq. The RNA-Seq reads were mapped to the *P. multocida* VP161 genome, and 128 genes were identified as being differentially expressed. More than 85% of the differentially expressed genes (109 of 128 genes) showed increased expression levels compared to those in the wild type (see Table S1 in Appendix B), suggesting that Hfq and its associated sRNAs generally act to decrease transcript levels, in keeping with the most common mode of sRNA/Hfq action at the transcriptional level, which is to increase RNase E turnover of transcripts.

Functional groups overrepresented in the upregulated gene set included those involved in cytochrome complex assembly (GO:0017004; 6 genes; $P = 2 \times 10^{-6}$), organic substance transport (GO:0071702; 13 genes; $P = 6 \times 10^{-4}$), nitrogen compound transport (GO:0071705; 7 genes; $P = 8 \times 10^{-4}$), and protein binding transcription factors (GO:0000988; 3 genes; $P = 0.001$). Among the upregulated genes were *pcgB*, *pcgD*, PM1042 (PMVP_1156, PMVP_1157, and PMVP_1057, respectively), encoding proteins involved in LPS biosynthesis (Harper and Boyce, 2017), and *hsf* and PM0855 (PMVP_0693 and PMVP_0855, respectively), predicted to encode the Hsf adhesin and a Flp pilus-like adhesin, respectively (Harper et al., 2003) The periplasmic nitrate reductase (*nap*) genes (*napFDAGHBC*) were all strongly upregulated (7.5- to 11.6-fold) in the *hfq* mutant, as were a number of genes encoding other putative electron transport proteins, including *ccmABCDEF* (2.0- to 4.1-fold) and *nrfBCD* (3-fold). In addition, numerous genes encoding proteins involved in heat shock and stress responses were also expressed at higher levels in the *hfq* mutant, including *htpX*, *uspG*, *htrA*, *dnaK*, *rpoH*, *clpB*, *rseC*, *rseB*, *mclA*, and *rpoE* (PMVP_0444, PMVP_0450, PMVP_0713, PMVP_0715, PMVP_1638, PMVP_1759, PMVP_1836, PMVP_1837, PMVP_1838, and PMVP_1839, respectively). An association of Hfq with the regulation of stress response genes/proteins has been seen in other bacteria, including *Y. pestis*, *Clostridium difficile*, and *S.* Typhimurium (Ansong et al., 2009; Boudry et al., 2014; Geng et al., 2009).

Nineteen genes displayed decreased transcript levels in the *hfq* mutant compared to the wild-type strain (see Table S2 in Appendix B). Functional groups overrepresented in this set included polysaccharide transport (GO:0015774; 2 genes; $P = 4 \times 10^{-4}$) and pyridoxal phosphate

metabolism (GO:0042822; 2 genes; $P = 0.001$). Importantly, 5 of the 10 capsule biosynthesis genes showed significantly reduced transcript expression in the *hfq* mutant ($P < 0.0001$) (see Table 2.5), correlating with the reduced HA capsule production observed for the *hfq* mutant.

2.3.5. Global changes in protein production in the *P. multocida hfq* mutant

Hfq often regulates protein production by altering mRNA translational efficiency via sRNA binding (Vogel and Luisi, 2011). Therefore, the proteomes of the VP161 *hfq* mutant and wild-type strain VP161 were compared during both early-exponential and mid-exponential growth phases by using quantitative high-throughput liquid proteomics. Totals of 1,147 and 1,041 proteins were identified during early-exponential and mid-exponential growth, respectively. During the early-exponential growth phase of the *hfq* mutant, 85 proteins were detected at increased levels and 21 proteins were detected at reduced levels compared to those in the wild-type strain (see Tables S3 and S4, respectively, in Appendix B). During the mid-exponential growth phase, 78 proteins were identified as being differentially expressed in the *hfq* mutant: 49 at increased levels (see Table S5 in Appendix B) and 29 at decreased levels (see Table S6 in Appendix B). Comparing the data from both growth phases, 37 proteins were expressed at increased levels at both the early-exponential and mid-exponential growth phases (Table 2.3), and 8 proteins were expressed at decreased levels at both growth phases (Table 2.4). Five of the capsule biosynthesis proteins (HyaE, HyaD, HyaC, HexD, and HexC) showed between 10- and 20-fold-reduced production at the early-exponential growth phase and between 1.8- and 2.7-fold-reduced production at the mid-exponential growth phase. These data correlated well with the decreased transcript levels observed for the capsule biosynthesis genes as measured by RNA-Seq (Table 2.5) and with the reduced HA capsule levels expressed by this strain. In addition, the known *P. multocida* virulence factor PfhB2 and its secretion partner LspB2 were also present at reduced levels in the *hfq* mutant at both growth phases (Table 2.4). In contrast, the levels of production of the haemoglobin binding proteins PMVP_0265 (PM0300) and PMVP_0304 (PM0337) were increased in the *hfq* mutant at both the early-exponential and mid-exponential growth phases (Table 2.3).

Table 2.3. Proteins showing increased production in the *P. multocida hfq* mutant strain in comparison to wild-type strain VP161 during both early-exponential and mid-exponential growth phases as measured by high-throughput liquid proteomics.

VP161 locus tag	Predicted function and/or product	Pm70 protein name	Early-exponential growth		Mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0194	Periplasmic dipeptide transporter	DppA	2.59	2.82E-04	2.44	1.06E-04
PMVP_0198	Dipeptide ABC transporter, ATP-binding protein	DppF	2.21	4.28E-03	1.71	1.39E-03
PMVP_0229	Murein L,D-transpeptidase, provisional	PM0270	1.57	4.66E-04	1.73	4.44E-04
PMVP_0253	Predicted membrane-bound lysozyme inhibitor of the c-type		3.30	6.48E-04	2.06	2.53E-04
PMVP_0265	Haemoglobin/haptoglobin binding protein	PM0300	1.79	5.84E-04	1.28	2.32E-03
PMVP_0267	Mrp putative ATPase	Mrp	1.18	5.19E-04	1.32	1.77E-03
PMVP_0292	Transglycosylase SLT domain protein	PM0326	1.56	1.69E-03	2.85	1.85E-03
PMVP_0304[a]	Haemoglobin-binding protein with TonB receptor	PM0337	1.49	6.40E-03	1.49	9.02E-04
PMVP_0305[a]	Haemoglobin-binding protein with TonB receptor	PM0337	1.83	1.75E-03	1.32	3.53E-03
PMVP_0327	MscS family potassium efflux protein, KefA	PM0358	1.07	8.84E-04	1.00	7.98E-03
PMVP_0376	Formate dehydrogenase accessory protein	FdhE	1.29	1.16E-03	1.00	3.62E-02
PMVP_0449	TRAP[b] transporter/C_4-dicarboxylate ABC transporter	PM0473	1.58	5.48E-04	1.78	8.79E-04
PMVP_0450	Universal stress protein, UspG	PM0474	1.21	5.03E-04	1.35	1.74E-03
PMVP_0488	Conserved protein with no known function	PM0519	2.10	4.44E-04	1.81	1.77E-03
PMVP_0516	glucose-1-phosphate adenylyltransferase	GlgC	1.46	5.48E-04	1.56	3.39E-02
PMVP_0541	Branched-chain amino acid aminotransferase	IvlE	1.39	5.84E-04	1.10	5.18E-03
PMVP_0561	Lipoprotein	Plp4	1.77	4.44E-04	1.31	1.85E-03
PMVP_0629	Tellurite resistance protein	TehB	3.50	8.84E-04	2.18	4.17E-03
PMVP_0756	ABC-ATPase-MoxR-like		1.18	1.58E-02	1.47	9.02E-04
PMVP_0837	Histidinol-phosphate aminotransferase	HisH1	1.26	1.01E-03	1.41	1.47E-03
PMVP_0852	Protein within Flp/Tight adhesion locus, homology to RpcC	PM0853	2.00	7.94E-04	1.74	5.98E-04

VP161 locus tag	Predicted function and/or product	Pm70 protein name	Early-exponential growth		Mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0939	Membrane-bound lytic murein transglycosylase	PM0928	1.66	7.94E-04	1.99	3.96E-04
PMVP_0948	Aspartate kinase, monofunctional class	PM0937	1.33	5.19E-04	1.21	2.57E-03
PMVP_0992	Predicted membrane/periplasmic protein of unknown function	PM0979	2.08	2.82E-04	2.05	4.04E-04
PMVP_1069	Dihydrodipicolinate synthase	DapA	1.11	7.94E-04	1.01	8.68E-03
PMVP_1208	Type I deoxyribonuclease HsdR (HtpX Peptidase M48 superfamily)	PM1190	1.34	7.27E-03	1.51	8.79E-04
PMVP_1233	Transglutaminase-like superfamily domain protein	PM1211	1.41	8.19E-04	1.98	2.08E-03
PMVP_1614	GTA[c] type/β-D-1,6 glucosyl transferase	PM1562	2.46	4.44E-04	2.45	1.06E-04
PMVP_1616	Large conductance mechanosensitive channel protein	MscL	2.12	7.32E-03	2.55	2.42E-04
PMVP_1649	Periplasmic nitrate reductase precursor	NapA	1.59	8.84E-04	1.29	1.81E-03
PMVP_1652	Periplasmic nitrate reductase	NapB	1.12	2.91E-02	1.75	5.47E-03
PMVP_1760	Thioredoxin	Trx	1.87	4.88E-03	1.73	1.33E-02
PMVP_1787	Outer membrane lipoprotein 2 precursor	PlpB	1.32	5.03E-04	1.09	9.91E-03
PMVP_1863	Nitrate/nitrite response regulator protein	NarP	2.51	2.82E-04	2.05	2.53E-04
PMVP_1872	SrfB homologue, uncharacterized protein conserved in bacteria	PM1819	1.21	2.04E-03	1.21	7.88E-03
PMVP_1879	Protein of unknown function	PM1826	1.86	4.66E-04	1.58	8.79E-04
PMVP_1962	Periplasmic oligopeptide-binding protein precursor	OppA	1.26	5.19E-04	1.17	3.72E-03
PMVP_2095	Glutamate dehydrogenase	GdhA	2.01	1.75E-03	2.71	1.06E-04

[a] These features likely represent broken sections of one gene (PM0337) that have not been fully assembled in the draft VP161 genome.
[b] TRAP, tripartite ATP-independent periplasmic.
[c] GTA, glycotransferase A.

Table 2.4. Proteins showing decreased production in the *Pasteurella multocida hfq* mutant strain in comparison to the wild-type strain VP161 during both early-exponential and mid-exponential growth phases as measured by high-throughput liquid proteomics.

VP161 locus tag	Predicted function and/or product	Pm70 protein name	Early-exponential growth		Mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0001	LspB	LspB2	-2.74	8.84E-04	-1.83	1.85E-03
PMVP_0002[a]	Filamentous hemmagglutinin	PfhB2	-2.81	4.44E-04	-1.85	1.92E-03
PMVP_0003[a]	Filamentous hemmagglutinin	PfhB2	-2.82	4.66E-04	-1.82	2.21E-03
PMVP_0004[a]	Filamentous hemmagglutinin	PfhB2	-2.70	4.44E-04	-1.77	3.58E-03
PMVP_0005[a]	Filamentous hemmagglutinin	PfhB2	-2.19	8.19E-04	-1.81	1.14E-02
PMVP_0006[a]	Filamentous hemmagglutinin	PfhB2	-1.69	5.03E-04	-1.86	1.85E-03
PMVP_0218	Spermidine/putrescine ABC transporter, periplasmic-binding protein	PotD1	-1.88	6.96E-04	-1.58	8.79E-04
PMVP_0767	Capsule biosynthesis, HyaE	HyaE	-4.17	2.82E-04	-1.11	2.63E-02
PMVP_0768	Capsule biosynthesis, HyaD	PM0775	-4.35	2.82E-04	-1.06	1.95E-02
PMVP_0769	Capsule biosynthesis, HyaC	PM0776	-3.86	1.22E-03	-1.41	5.60E-03
PMVP_0771	Capsule transport, HexD	HexD	-3.29	4.50E-04	-1.01	1.33E-02
PMVP_0869	DNA-binding transcriptional regulator. Fructose repressor	FruR	-1.99	1.00E-02	-1.31	4.81E-03

[a] These features likely represent broken sections of one gene (*pfhB2*) that have not been fully assembled in the draft VP161 genome.

2.3.6. Comparison of the proteomic and transcriptomic data

The proteomic analysis of the *hfq* mutant at the mid-exponential growth phase (see above) identified 78 proteins as being differentially expressed, with 49 being identified at increased levels and 29 at decreased levels. Of the 49 proteins produced at higher levels in the *hfq* mutant, 13 also showed increased transcript levels as measured by RNA-Seq; all 13 of these proteins also showed increased levels (as measured by proteomics) at the early-exponential growth phase (see Table S7 in Appendix B). The 13 genes/proteins that showed increased expression levels by both techniques included two outer membrane lipoproteins (Plp4 and PlpB), three components of the periplasmic nitrate reductase system, the universal stress protein UspG, and the oligopeptide transporter DppA. Similarly, of the 29 proteins measured at reduced levels in the *hfq* mutant by high-throughput proteomics, six were also downregulated as determined by RNA-Seq (see Table S8 in Appendix B), and five of these also showed decreased expression at the early-exponential growth phase (see Table S8 in Appendix B). The six genes/proteins determined to be downregulated in the *hfq* mutant by both proteomics and transcriptomics analyses included three capsule biosynthesis proteins, the filamentous hemagglutinin secretion partner LspB, the global regulator of carbon metabolic flux FruR (Ramseier et al., 1995), and the imidazole glycerol phosphate synthase, which is a key enzyme that links amino acid and nucleotide biosynthesis pathways (Douangamath et al., 2002). Interestingly, 59 of the 78 proteins identified as being differentially expressed in the *hfq* mutant by proteomics analysis were not determined to be differentially expressed by transcriptomics analysis. It is likely that many of these proteins are regulated by Hfq-sRNA interactions that act posttranscriptionally, a common mechanism of sRNA action (Vogel and Luisi, 2011).

Table 2.5. Differential expression of capsule biosynthesis genes/proteins in the *Pasteurella multocida hfq* mutant strain in comparison to the wild-type VP161 strain.

VP161 locus tag	PM70 name	Gene name	RNA-Seq data for mid-exponential growth		Proteomics data for early-exponential growth		Proteomics data for mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0765	*phyB*	*phyB*	-1.35	1.02E-03	ND [a]	ND	0.12	9.24E-01
PMVP_0766	*phyA*	*phyA*	-1.86	7.35E-04	-0.91	5.18E-02	-0.65	2.03E-01
PMVP_0767	*hyaE*	*hyaE*	-1.78	2.16E-03	-4.17	2.82E-04	-1.11	2.63E-02
PMVP_0768	PM0775	*hyaD*	-1.84	2.08E-03	-4.35	2.82E-04	-1.06	1.95E-02
PMVP_0769	PM0776	*hyaC*	-1.41	7.48E-03	-3.86	1.22E-03	-1.41	5.60E-03
PMVP_0770	PM0777	*hyaB*	-0.67	2.90E-02	ND	ND	ND	ND
PMVP_0771	*hexD*	*hexD*	-0.60	5.78E-02	-3.29	4.50E-04	-1.01	1.33E-02
PMVP_0772	*hexC*	*hexC*	-0.68	4.22E-02	-3.44	4.44E-04	-0.88	1.90E-02
PMVP_0773	*hexB*	*hexB*	-0.62	3.90E-02	ND	ND	ND	ND
PMVP_0774	*hexA*	*hexA*	-0.46	4.09E-02	0.30	6.42E-01	-0.71	6.89E-02

[a] ND indicates that the level of the protein was not measured under the specified conditions.

2.4. Discussion

Inactivation of *hfq* in different bacterial species has resulted in a wide range of phenotypic effects. For some species, such as *Brucella melitensis*, *Legionella pneumophila*, and *Salmonella enterica* serovar Typhimurium, *hfq* mutants show a reduced growth rate with a longer lag phase in certain *in vitro* media (Cui et al., 2013; McNealy et al., 2005; Sittka et al., 2007). In other bacterial species, such as *Listeria monocytogenes*, *Staphylococcus aureus*, *Haemophilus influenzae* (Bohn et al., 2007; Christiansen et al., 2004; Hempel et al., 2013), and *P. multocida* (as shown in this study for strain VP161), the growth rate *in vitro* in rich medium is unaffected by *hfq* inactivation. Despite this normal *in vitro* growth rate, the *P. multocida hfq* mutant showed significantly reduced fitness *in vivo*, both in mice and in chickens. In mice, the *hfq* mutant showed between 4- and 5-fold reduced growth compared to the wild-type strain in competitive growth assays. The reduced fitness in chickens was evident following both IT and IM inoculation. However, the *hfq* mutant was still able to cause disease in some animals, with 50% or 70% of inoculated birds progressing to late stage fowl cholera signs for IT and IM inoculation, respectively. Reduced *in vivo* fitness has also been demonstrated for *hfq* mutant strains of *Y. pestis*, *Klebsiella pneumoniae*, *Francisella tularensis* and *Actinobacillus pleuropneumoniae* (Chiang et al., 2011; Geng et al., 2009; Meibom et al., 2009; Subashchandrabose et al., 2013). The *hfq* mutant also showed reduced growth (up to 2.5-fold) in active chicken serum, indicating that Hfq has a role in the ability of *P. multocida* to resist the bactericidal effects of complement in chicken serum.

One of the critical *P. multocida* virulence factors is the polysaccharide capsule (Chung et al., 2001). The *P. multocida hfq* mutant showed significantly reduced HA capsular polysaccharide (CPS) production (up to 14-fold) at all tested growth phases. Transcriptomic and proteomic analyses of the *P. multocida hfq* mutant confirmed the phenotypic data, showing highly reduced transcript and protein levels for many of the capsule biosynthesis genes/proteins (Table 2.5). Thus, Hfq plays a crucial role in positively regulating capsule production. The reduced level of capsule biosynthesis gene transcripts suggests that Hfq-mediated regulation acts primarily at the level of transcription; however, we cannot rule out an additional role for posttranscriptional mechanisms. In contrast to the effect of Hfq on *P. multocida* capsule production, a *K. pneumoniae hfq* mutant is hypermucoid, with increased levels of K2-specific CPSs (Chiang et al., 2011). The reduction in capsule production in the *P. multocida hfq* mutant is likely to play a major role in its reduced ability to grow in active chicken serum, and importantly in the reduced *in vivo* fitness observed for this strain.

We have previously shown that the global transcriptional regulator Fis is essential for the expression of *P. multocida* capsule biosynthesis genes (Steen et al., 2010). Our proteomic and transcriptomic analyses indicated that Fis was produced at unchanged (proteomics data) or slightly increased (transcriptomics data) levels (1.8-fold; FDR = 0.03) in the *hfq* mutant compared to the wild-type strain, implying that the loss of capsule expression in the *hfq* mutant was not the result of a loss of Fis production. However, Fis may play a role in the regulation of *hfq* expression, as RNA-Seq analysis of a *P. multocida* VP161 *fis* mutant during the early-exponential growth phase revealed a 2.2-fold decrease in *hfq* expression (FDR = 0.001) compared to that of the wild-type strain (Harper and Boyce, unpublished). These data suggest that Fis positively regulates Hfq, which in turn positively regulates capsule expression. However, mutation of *fis* alone has a much larger effect on capsule expression than mutation of *hfq* alone, suggesting that Fis primarily acts separately from Hfq.

P. multocida expresses two filamentous hemagglutinins, PfhB1 and PfhB2, that are important for *in vivo* fitness (Fuller et al., 2000; Tatum et al., 2005). Our proteomic analyses showed that the loss of Hfq leads to a reduction in the amounts of PfhB2 and the secretion partner LspB2 during both early-exponential and mid-exponential growth. It is possible that the reduced expression of PfhB2 may have contributed, along with reduced capsule expression, to the reduced *in vivo* fitness of the *hfq* mutant. Our transcriptomic analyses showed that genes encoding these proteins were also likely downregulated (PfhB2 was 1.4-fold downregulated, with an FDR of 0.056, and LspB2 was 2.3-fold downregulated, with an FDR of 0.01), indicating that the Hfq-associated regulation of these genes is also likely to be at the transcriptional level.

LPS is another crucial *P. multocida* virulence factor; strains expressing truncated LPS are highly attenuated for virulence (Boyce et al., 2009; Harper et al., 2011). Three genes involved in LPS biosynthesis (*pcgB*, *pcgD*, and PM1042) were expressed at increased levels in the *hfq* mutant. PcgB and PcgD are required for the biosynthesis and transfer of phosphocholine residues to the terminal galactose residues of the LPS produced by strain VP161 (Harper et al., 2007b), and a functional PM1042 protein is required for the addition of phosphoethanolamine to lipid A (Harper et al., 2017). Both phosphoethanolamine and phosphocholine are important charged LPS substituents that can affect the interaction of host antimicrobial peptides with the bacterial surface (Harper et al., 2007b). Numerous genes/proteins involved in electron transport were also upregulated in the *hfq* mutant, including *napFDAGHBC*, *ccmABCDEF*, and *nrfBCD*. Interestingly, the Nap system is predicted to reduce nitrate to nitrite during anaerobic growth (Cole, 1996), and a previous microarray study of *P. multocida* (strain X73) showed that the expression of the *nap* genes was significantly upregulated during *in vivo* growth in the chicken

host, a relatively anaerobic niche (Boyce et al., 2002). Thus, appropriate regulation of these genes appears critical for full *in vivo* fitness.

In conclusion, this study defined a clear role for Hfq in the regulation of *P. multocida* gene and protein expression and allowed us to identify genes controlled (either directly or indirectly) by Hfq-sRNA interactions at both the transcriptional and translational levels. Together, transcriptional and proteomic analyses of the *hfq* mutant supported our hypothesis that Hfq is involved in the regulation of virulence- and fitness-associated genes of *P. multocida*. Hfq is clearly involved in regulating the production of the known *P. multocida* virulence factors HA capsule, filamentous hemagglutinin, and LPS, as well as accessory virulence genes such as those encoding iron uptake proteins, stress response proteins, and bacterial adhesins. We are currently focusing on identifying specific sRNA-mRNA interactions and in particular the specific mechanism(s) by which Hfq/sRNA action regulates the expression of HA capsule, LPS, and filamentous hemagglutinin.

Chapter Three: Identification of *P. multocida* sRNAs using bioinformatic analysis of whole-genome RNA-Sequencing data

3.1. Introduction

The Gram-negative bacterium *Pasteurella multocida* is the causative agent of a number of economically important animal diseases, including avian fowl cholera, haemorrhagic septicaemia and atrophic rhinitis (Wilson and Ho, 2013). A small number of *P. multocida* virulence factors have been identified, including capsule, lipopolysaccharide (LPS) and filamentous hemagglutinin, but little is known about how the expression of these virulence factors is regulated. Small non-coding RNA molecules (sRNAs) are important regulators of bacterial gene expression and protein production, with key roles in controlling a diverse range of bacterial functions and in the expression of virulence factors (Romby et al., 2006). Hfq is an RNA-binding protein that facilitates riboregulation, generally via interactions with small non-coding RNA molecules (sRNAs) and their mRNA targets (Sobrero and Valverde, 2012).

In Chapter 2, data was presented that showed that the *P. multocida* RNA chaperone Hfq plays an important role in the global regulation of *P. multocida* gene and protein expression. A *P. multocida hfq* mutant produced significantly less hyaluronic acid capsule during all growth phases, and displayed reduced *in vivo* fitness and virulence compared with the wild-type strain. The phenotype of the *hfq* mutant correlated with changes in capsule gene expression and protein production. Expression of other genes and proteins associated with *P. multocida* virulence factors, filamentous hemagglutinin and LPS were also altered in the mutant. These data indicate that Hfq, and therefore the sRNAs associated with Hfq, are likely to be critical regulators of *P. multocida* virulence. However, currently there is very little data on *P. multocida* sRNAs and their role in *P. multocida* gene regulation. Identification of the range of *P. multocida* sRNAs is crucial to understanding the *P. multocida* riboregulatory network and will therefore provide significant insights into *P. multocida* pathogenesis.

Bacterial sRNAs are *cis*-acting or *trans*-acting and are encoded on the complementary strand of their target gene, or in intergenic locations, respectively (Waters and Storz, 2009). However, some sRNAs are encoded at the 5′ or the 3′ end of untranslated regions (UTRs) of genes, such as the sRNA DapZ in *Salmonella enterica* serovar Typhimurium, or their coding region overlaps part of a gene coding sequence (Chao et al., 2012; Kawano et al., 2005). Other RNA species also have important riboregulatory functions in bacteria. These regulatory RNAs include riboswitches, thermosensors and other *cis*-acting regulatory elements of mRNAs, as well as tRNA and rRNA cleavage products (De Lay and Garsin, 2016; Gottesman and Storz, 2011; Kavita et al., 2018).

Cis-acting sRNAs share perfect complementarity with their target and consequently have a limited range of mRNA targets (Thomason and Storz, 2010). For example, the *cis*-acting sRNA GadY is known to regulate acid response genes in *E. coli* via binding to the 3′ UTR region of the *gadX* mRNA, a transcriptional regulator involved in response to acid (Opdyke et al., 2004). In contrast, *trans*-acting sRNAs regulate the degradation and translation of target mRNAs encoded at a range of locations on the genome and have limited sequence complementarity with their multiple targets (Waters and Storz, 2009). *Trans*-acting sRNAs usually require an association with the Hfq protein and a single *trans*-acting sRNA can exert a wide range of potential regulatory effects. The well-studied *trans*-acting bacterial sRNA DsrA has an important role in positively regulating the translation of the stress sigma factor RpoS, and can also act to negatively regulate translation of the histone-like protein H-NS (Lease et al., 1998; Majdalani et al., 1998). As *trans*-acting sRNAs are generally encoded in intergenic regions, they can usually be identified using transcriptomic/bioinformatic analyses. It is these molecules that will be the main focus of investigation in this chapter. There is also some evidence of RNA elements that act both in *cis* and in *trans*; these are usually other regulatory RNA elements such as *cis*-acting riboswitches, which have a dual function as a *trans*-acting sRNA, including SreA and SreB in *Listeria monocytognes* (Loh et al., 2009).

Bacterial sRNAs have been most widely identified and studied in model organisms such as *E. coli* and *S.* Typhimurium (Barquist and Vogel, 2015). Advances in bioinformatic identification tools have allowed for the prediction of many novel sRNAs across a wide range of bacterial species. In particular, the development of comprehensive high throughput RNA-Seq and transcriptome analysis methods has allowed for the detection of sRNAs in a diverse range of bacteria and for the prediction of their mRNA targets (Barquist and Vogel, 2015). These bioinformatic, online sRNA target prediction tools include TargetRNA2, IntaRNA and RNAPredator (Busch et al., 2008; Eggenhofer et al., 2011; Kery et al., 2014). Generally, these tools predict sRNA targets by taking into consideration factors such as RNA sequence and structure, and binding energy between the RNA pair (Pain et al., 2015). A previous computational-based study in *Actinobacillus pleuropneumoniae*, a member of the *Pasteurellaceae* family, identified three predicted regulatory RNAs with sequence similarity to the *P. multocida* genome; Rossi reported that these regulatory RNAs showed similarity to an RNaseP, tmRNA and RtT molecule rather than *trans*- or *cis*-encoded sRNA molecules as reported in this study (Rossi et al., 2016).

The number of sRNAs expressed by *P. multocida* is unknown and identifying the range of sRNAs and understanding their role in the *P. multocida* riboregulatory network is of great importance. At the beginning of the work described in this chapter, there were no *P. multocida* sRNAs known.

This chapter describes the identification of a range of putative *trans*-acting sRNA molecules, and some putative *cis*-acting sRNA molecules, using bioinformatic analysis of *P. multocida* RNA-Sequencing (RNA-Seq) data. To maximise sRNA discovery, we analysed transcriptomic data generated from *P. multocida* grown in a range of growth conditions. Specifically, RNA profiles of *P. multocida* grown in iron-limited and reduced oxygen conditions were generated, as these mimic some of the conditions *P. multocida* will likely encounter during colonisation of the chicken host.

Our previous *hfq* studies were performed in the poultry isolate *P. multocida* strain VP161. In this work we first analysed RNA-Seq data from VP161 that had been generated for another study to identify putative sRNAs. Following this, we chose to analyse the closely related *P. multocida* X-73 strain, grown under different conditions, for sRNA identification. *P. multocida* strain X-73 multiplies to very high levels in the blood of chickens during late infection, and this would allow for excellent recovery of RNA for sRNA analysis (Boyce and Harper, pers. comm.). Although *in vivo* studies were not conducted in this project, the use of X-73 would allow for future *in vivo* studies on any important sRNAs identified in this current study.

3.2. Materials and Methods

3.2.1. Growth of *P. multocida* under different conditions

The bacterial strains used in this study are listed in Table 3.1. *P. multocida* strains were routinely grown as described previously (Mégroz et al., 2016). For transcriptomic studies, cultures of *P. multocida* wild-type strain X-73 (biological triplicate) were grown overnight in HI broth at 37°C with constant shaking at 200 rpm. Cells were then subcultured 1:50 into 120 mL of fresh HI broth and grown at 37°C (with constant shaking at 200 rpm) to an optical density at 600 nm (OD_{600}) of 0.6 (mid-exponential growth, approximately 1.0 x 10^9 colony forming units per mL [CFU/mL]). Each mid-exponential growth phase culture was then used to inoculate cultures for three different growth conditions, aerobic growth in rich medium (HI broth), growth in iron-limited media and growth in reduced oxygen conditions. For growth under normal aerobic conditions, a 25 mL sample of each culture was transferred to separate, sterile flasks and incubated at 37°C with constant shaking at 200 rpm for 10 min. For growth in iron-limited conditions, a 25 mL sample of each mid-exponential growth phase culture was transferred to separate, sterile flasks (previously soaked in ~ 3M HCl overnight and rinsed ten times with MilliQ H_2O to remove any residual traces of HCl and available iron) then concentrated 2,2'-dipyridyl (Sigma-Aldrich) was added to each flask to a final concentration of 200 µM. Each culture was then incubated at 37°C with constant shaking at 200 rpm for 10 min. For growth in reduced oxygen conditions, a volume of approximately 50 mL of each mid-exponential growth phase culture was added to separate, RNase-free, 50 mL Falcon conical centrifuge tubes (Sigma-Aldrich). Each tube was completely filled using additional culture and then the lid was closed tightly; tubes were then incubated at 37°C for 30 min without agitation. Following incubation of all cultures, 5 mL of each culture was removed to a fresh tube; the 5 mL samples taken from the cultures under reduced oxygen conditions were carefully taken from the bottom of the tube. Each sample was immediately placed on ice for 10 min before the cells were pelleted by centrifugation at 8000 x *g* for 10 min at 4°C. The supernatant was removed, and each cell pellet was resuspended in 1 mL of TRIzol reagent (Thermo Fisher Scientific) to stabilise the RNA before incubation at 65°C for 10 min. Samples were stored at -80°C until RNA isolation.

3.2.2. RNA extraction and purification, rRNA depletion, library preparation and RNA sequencing

RNA was isolated from each sample prepared above using a phenol:chloroform extraction method as follows. A volume of 200 µL of chloroform:isoamyl alcohol solution (24:1) was added to each sample and mixed thoroughly by inverting the tube. To allow for phase separation, samples were

incubated at RT for 5-10 min and then centrifuged at 12,000 x *g* for 15 min at 4°C. The aqueous phase was carefully removed (avoiding the interface) and transferred to a separate RNase-free 1.5 mL tube (Eppendorf). To precipitate the RNA, an equal volume of ice-cold isopropanol, containing 0.5 µL of Protector RNase Inhibitor (Merck) was added to each sample. Tubes were mixed by inversion, incubated for 10 min at RT, then centrifuged at 12,000 x *g* (4°C) for 10 min. Following removal of the supernatant, each RNA pellet was washed with 500 µL of 75% EtOH, then the tubes centrifuged at 12,000 x *g* for 5 min at RT. After removal of the 75% EtOH, the precipitated RNA was dried then dissolved in 100 µL of nuclease-free H_2O (prewarmed to 50°C) containing 0.5 µL of Protector RNase Inhibitor (Merck) and incubated at 50°C for 10-15 min. A repeat of the RNA precipitation was performed as follows, 10 µL of 4M NaCl and 300 µL of 99% EtOH were added to each sample which were then incubated O/N at -20°C. RNA was pelleted by centrifugation at 12,000 x *g* for 10 min at 4°C, supernatant removed, then the RNA pellets were washed with 0.5 mL 75% EtOH before centrifugation at 12,000 x *g* for 5 min at RT. After the wash solution was removed, the RNA pellets were dried then resuspended in 50 µL of nuclease-free H_2O containing 0.5 µL of Protector RNase Inhibitor (Merck). DNA was removed from each RNA sample by treatment with DNase using the Qiagen RNeasy kit, in-tube treatment method, according to the manufacturer's instructions. Each RNA sample was further purified by phenol:chloroform extraction. Briefly, the entire volume of each sample was added to a 5 PRIME phase lock gel™ Heavy tube (Quantabio) containing 100 µL of a 1:1 phenol:chloroform solution. Tubes were vortexed vigorously and pelleted by centrifugation at 12,000 x *g* for 5 min at RT. A volume of 100 µL of chloroform was added to each tube, which was vortexed vigorously then centrifuged at 12,000 x *g* for 5 min at RT. The upper aqueous phase containing purified RNA was carefully removed and transferred to a separate RNase-free 1.5 mL tube (Eppendorf). To precipitate the RNA, 10 µL of 3M sodium acetate (pH 5.2, Thermo Fisher Scientific), 2 µL of 5 mg/mL glycogen (Thermo Fisher Scientific) and 250 µL of 100% EtOH was added to each sample (~100 µL) before incubation for 15 min at RT. RNA was pelleted by centrifugation at 12,000 x *g* for 30 min at RT, supernatant removed, then each RNA pellet was washed using 200 µL of 70% EtOH. The RNA pellet was centrifuged again (12,000 x *g* for 5 min at RT), the 70% EtOH was then removed and the RNA pellets dried. Each RNA sample was resuspended in 50 µL of nuclease-free H_2O containing 1 µL of Protector RNase Inhibitor (Merck).

Quantification of RNA in each sample and assessment of levels of DNA contamination were performed as described previously (Mégroz et al., 2016). Each total RNA sample was depleted of rRNA as described previously (Mégroz et al., 2016). For preparation of the sRNA-enriched, strand-specific libraries, RNA samples were first enriched for small RNA fragments (< 200 bp)

using the mirVana™ miRNA Isolation Kit according to the manufacturer's instructions. Libraries were prepared using the SureSelect Strand-Specific RNA Library Prep for Illumina Multiplexed Sequencing kit, mRNA library preparation protocol (Agilent), according to the manufacturer's instructions. Library validation, normalisation and pooling of rRNA depleted, total RNA libraries and sRNA-enriched libraries was performed by the Micromon Genomics Platform at Monash University. Libraries were combined and loaded into a single lane of a NextSeq instrument (Illumina) for sequencing as 150 bp single-end reads (performed by the Micromon Genomics Platform at Monash University).

Adapter sequences were removed from the raw sequencing reads using Trimmomatic (version 0.36), and all sequences less than 36 bp in length were discarded. Mapping of the trimmed RNA-Seq reads to the draft X-73 genome was performed using the Rsubread package (version 1.30.1) within RStudio. Statistical analysis was performed using the limma/voom modelling software (Law et al., 2014) via the Degust visualisation software, available through the Bioinformatics Platform at Monash University. For the total RNA libraries, approximately 9.3 million reads (98.64% of all reads) mapped uniquely to the *P. multocida* X-73 genome per sample. For the sRNA-enriched libraries, approximately 1.7 million reads (97.45% of all reads) mapped uniquely to the *P. multocida* X-73 genome per sample. Transcripts were identified as being differentially expressed if they displayed a ≥ 2.0-fold (≥ 1.0-$\log_2$) change in expression across the triplicate replicates at a false discovery rate (FDR) of < 0.05. Predicted functions for gene products were assigned by mapping the differentially expressed X-73 genes to their closest orthologue in the fully annotated *P. multocida* Pm70 reference genome (May et al., 2001).

3.2.3. Identification of sRNAs using bioinformatic analysis

Identification of putative *P. multocida* sRNAs was performed using the mapped RNA-Seq datasets as visualised in Artemis (Sanger; http://www.sanger.ac.uk/resources/software/artemis/). Reads were visually inspected for putative sRNAs, defined as highly expressed regions in the genome that were of short length (~ 50-400 bp) and located either intergenically (in *trans*), overlapping 5′ or 3′ regions of annotated genes, or in *cis* (encoded on the opposite strand of a predicted protein-encoding gene).

3.2.4. Identification of differentially expressed putative sRNAs

To identify the sRNAs that were differentially expressed during growth in iron-limited media or during growth under reduced oxygen conditions, regions encoding putative sRNAs were first precisely annotated to genomes representing *P. multocida* strains X-73 and VP161 using the Artemis (Sanger) genome browser and annotation tool. RNA-Seq reads were then re-mapped to

the re-annotated *P. multocida* X-73 genome sequence using the Rsubread package within RStudio, and statistical analyses were performed using the limma/voom modelling software (Law et al., 2014) as described in Section 3.2.2. Specific sRNAs were identified as being differentially expressed under a particular growth condition if they displayed a ≥ 2.0-fold (≥ 1.0-$\log_2$) change in expression across the replicates at a false discovery rate (FDR) of < 0.05, compared to expression in cells grown under normal aerobic laboratory conditions.

Table 3.1. Bacterial strains used in this study.

Strain or plasmid	Relevant description	Source or reference
Strains		
P. multocida		
AL2521	VP161 *hfq* TargeTron mutant; Kan^R	(Mégroz et al., 2016)
AL2677	VP161 *gcvB* TargeTron mutant; Kan^R	(Gulliver et al., 2018)
AL2973	VP161 *proQ* TargeTron mutant; Kan^R	Gulliver, Boyce Laboratory, unpublished
VP161	Serotype A:1; Vietnamese isolate from chickens	(Wilkie et al., 2000)
X-73	Serotype A:1, wild-type strain	(Heddleston and Rebers, 1972)

3.3. Results

3.3.1. Identification of *P. multocida* VP161 putative sRNAs using RNA-Seq data

In other bacteria, the Hfq RNA chaperone helps to regulate gene expression via its interaction with specific sRNAs (Vogel and Luisi, 2011). However, at the beginning of this project there was no published information on the sRNA species expressed by *P. multocida*. To initially identify putative sRNAs produced by *P. multocida*, a previously generated whole transcriptome dataset from wild-type strain VP161 was utilised (Deveson Lucas, Boyce Laboratory, unpublished); this dataset was size-selected for RNA fragments >100 bp (Boyce, pers. comm.). This dataset included RNA sequencing reads generated from *P. multocida* grown *in vitro* in Brain Heart Infusion broth (at 37°C with shaking) to early-exponential (OD_{600} = 0.2 and 0.4), mid-exponential (OD_{600} = 0.7) and late-exponential (OD_{600} = 1.0) growth phases. Using Artemis (Sanger) software, sequences were mapped to the VP161 genome and visually inspected for short regions (50-400 nucleotides) that contained high numbers of reads, particularly in predicted intergenic regions (for examples, see Figure 3.1). Using this strategy, 23 putative *P. multocida* sRNAs were identified; these were used as query sequences to interrogate the Rfam database (Sanger) (http://rfam.sanger.ac.uk/). Eleven of the 23 predicted *P. multocida* sRNAs had homologues in other species (Table 3.2), including GcvB, involved in the regulation of amino acid biosynthesis and transport in bacteria (Gulliver et al., 2018; Pulvermacher et al., 2009), and the regulatory RNA HrrF, that is involved in iron homeostasis and known to have orthologues among other members of the *Pasteurellaceae* family, including in *Haemophilus spp* (Santana et al., 2014). The *P. multocida* putative sRNAs were also used to interrogate the sRNAPredict database (http://newbio.cs.wisc.edu/sRNA/published_annotations/, no longer available) and three of the putative *P. multocida* sRNAs shared identity with sRNAs previously identified in other bacteria (the glycine riboswitch, GcvB sRNA and FMN riboswitch; Table 3.2).

Thirteen of the putative sRNAs identified in *P. multocida* strain VP161 that did not have any well characterised homologues were further investigated, as well as the GcvB and HrrF sRNAs that had previously been identified in other bacteria. As this first RNA-Seq dataset was non-stranded, both forward and reverse strands of the transcribed section of each putative sRNA were inspected for putative Rho-independent terminator sequences, typically comprising a stable hairpin followed by a string of uridine nucleotides and usually required for transcriptional termination (Lesnik et al., 2001). The sequences representing GcvB and HrrF sRNAs both contained a clear terminator sequence (Table 3.2). Four of the other putative sRNAs, PmVP161_2204, PmVP161_2205, PmVP161_2206, PmVP161_2203, also encoded a clear terminator sequence (Table 3.2).

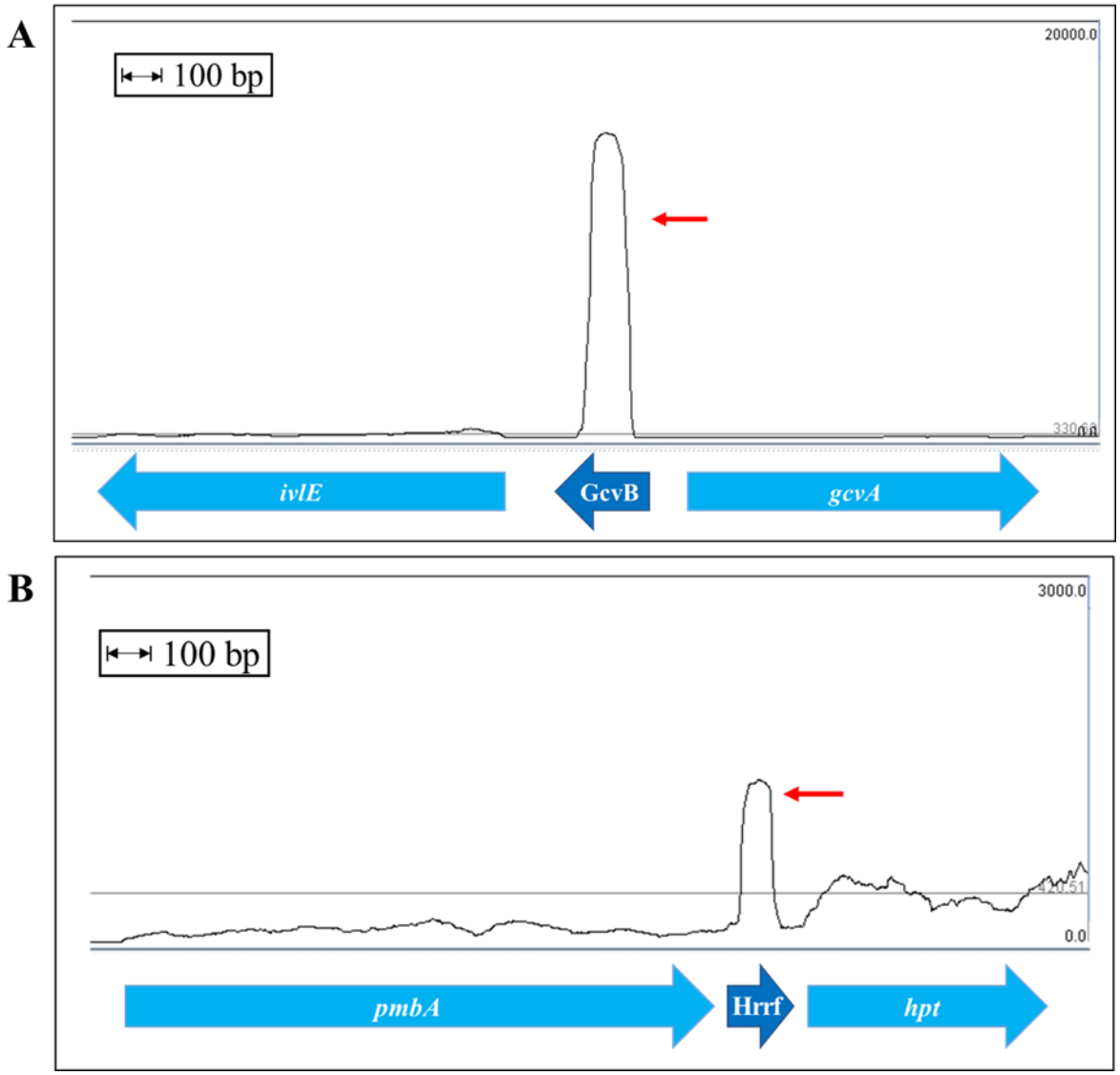


Figure 3.1. Identification of putative *P. multocida* sRNA-encoding regions by visual inspection of *P. multocida* transcriptomic data. Previously generated whole transcriptome RNA-Seq data of wild-type VP161, grown at 37°C to various growth phases, were mapped to the VP161 genome and visualised using Artemis software. Regions with high numbers of sequence reads in predicted intergenic regions were identified by visual inspection and designated as putative sRNAs (indicated by red arrows). Two putative sRNAs are shown, one upstream of *gcvA* (**A**) and one upstream of *hpt* (**B**). The sRNA upstream of *gcvA* is a homologue of a well characterised sRNA, GcvB (Pulvermacher et al., 2009), and the sRNA upstream of *hpt* is a homologue of the sRNA HrrF found in *Haemophilus* species (Santana et al., 2014).

Table 3.2. Putative sRNAs identified from visual inspection of *P. multocida* strain VP161 high throughput RNA sequencing data. Predicted Rho-independent terminator sequences, consisting of a stem loop (red and green nucleotide pairs) followed by a string of uridine nucleotides (highlighted yellow), are shown.

Designated Name[a]	Startbase[a]	Endbase[a]	Length (bp)	Known homologue[b]	Rfam number[c]	Reference	Predicted sequence
PMVP_2500	1242818	1242448	371	**Glycine riboswitch**	RF00504	(Mandal et al., 2004)	uuuuuacuauuuaucugcuaaaauuugcugugcuuuuuugaaaaau ucaacaaaaguaaguucggaugaagguagcuggagaguagggaauu gacccuacgccgacgagguaaacucuuucaggcgcugauuauaagca gggacuguuacuggacgaacccuuggagagagccguuauauuaaaaa gagauaaaacggccgccgaaggcgcaaaaagagcgguuaauuuuucc guuuuuucaaacgcucaggcaaaaggacaggggcaacaagauauucg guuugauguuuacaugcaugggcuacaugcgugcuuuucucuuuuu uaacucuuccgcaaaaugcuuuugccuccuuguaaucugugauuu
PMVP_2501 (PmVP161_ 2204)	1220886	1220731	156	No hits	None		guuuuuuaauuuugauugaaaagaugaauaaacgcacugugauuac uaugauuaugaccaccacuauuauuaccaauagugugcggguuagu gcguaacaaaaagaucgaauuccacaaaacccguacugaauaaaaagu gcgggguuuuuuuauga
sRNA HrrF PMVP_2502 (PmVP161_ 2201)	1210716	1210566	151	**HrrF**	RF02728	(Santana et al., 2014)	auuugcuaagauagccacaucuguuagagauaaaggugggaaugaua gaguagaucguuuucagucugagucauacuccaaacagguagaguag aaaguaaauucguuaggguauuuaugggcugaaggguuuaaacuuca gccuuuuuuuu
PMVP_2512 (PmVP161_ 2237)	421793	421924	132	No hits	None		uguucaccaacgucguccgucgcaguaaaucucgacaacugcuaacg aaugcauugcgcagggauugauaugugggggaaguucgagucccacc gggcgacaccuuauuaauauguuuccuuuuuucauuua
PMVP_2503 (PmVP161_ 2212)	425361	425506	146	No hits	None		augccccccuguuuuggacgaaaguaagugcaagaagacuaucauac acacuuguuuuguuaacccucaugcauuuauucguuacaggcacgga agacacgaugagcaaggaugccuuucaguuaagggcaaagcgauggc uuugc
PMVP_2504	480569	480710	142	No hits	None		auacaacgcucaaguguuuuuggguguugagugagacguuagacga agacgaagaagagagaauuacuuuagacagcuuguuuuugauuaaag ugcgagagugaagaaagcacgaaggaagaacagaagaauagagaaga ca

Designated Name[a]	Startbase[a]	Endbase[a]	Length (bp)	Known homologue[b]	Rfam number[c]	Reference	Predicted sequence
PMVP_2505	534898	535295	398	**Bacterial RNaseP classA**	RF00010	(Kazantsev and Pace, 2006)	ugcgcgccugaguuggucagacaaucgcugguuuauugaagcccuu aaccgucuuuaugacgaccuaguggggacaaauagacgagaggaaagu ccgagcuacacagggcagagugccagguaacgccuggggaggcgcgag ccuacgacaagugcaacagagagcagaccgccaguuucggcugguaa gggugaaagggugcgguaagagcgcaccgcauggcuggcaacaguac auggcaggguaaacuccacucguagcaagaccaaauaggaauccaau gaguggcccgcucaggauucgggguagguugcuugagcguaugugcg aauguacgccuagaggaaugauuguccgcgacagaacucggcuuauc gaccaacucaguacuuuucuacac
PMVP_2506 (PmVP161_ 2205)	571607	571767	161	No hits	None		uuugauuaauccguagagguaacaaugagugaauuugguuaugcaa ugaugacaguugcuuuggcgcuaggguuuaauuguuguauuggcgu ucucaauuuuuuuaagacuaaacaaaauaaaaauaacgaaagaucgag gcauaaucgccucggucuuuuuu
PMVP_2507 (PmVP161_ 2206)	574392	574552	161	***Aggregatibacter* sRNA 20 (AaHKsRNA 20)**	RF02898	(Oogai et al., 2017)	aucuuaaaagcgugaauuaguucacauuauugaaaauacccauuugc uaacuaaggguuuauuuucguauucugugucuagcuaaaaauaagug cuauaaaaaauagccgacauauuuugcuuuccgaguagugccuuguu ucaaggcacuuuuuuuuguuu
sRNA GcvB PMVP_2508 (PmVP161_2 200)	591484	591269	216	**GcvB**	RF00022	(Pulvermacher et al., 2009)	ccacucucuuauuucccucuaaaaugcgcgcauacuuaaugauuggu aauuccuuacuggguuaagaguuucggacuuaaguaugauguuguguu uugcauauuguuuggguaaccaaacagaguaacuuaucuuaaguuac uuauuuucacuuccuguauauuuauuuuagacccuuuuggguuaaac accgcucuauugagcgguguuuuuuuaucu
PMVP_2509 (PmVP161_ 2223)	631297	631437	141	No hits	None		uuuaaauuuaaucuugcaauaggauuuugcauuauuaaauauguuu aaagauaaaugaucaguuguauaaguaugaaguuuaaggaguacuua augaauuuaucuaaaaucgcaacagggaucguugcgucauuuuguu uu
PMVP_2524 (PmVP161_ 2207)	631532	631693	162	No hits	None		uuuaaugaaagaauugccgcaaaagcgggaaucgccaaaagcguaga aggugccgcguuaucuuauaacguuggugugaacuaugaguuuuaa ucguaaaaaggaguuuaacggacaacauuaauauuauuuaccucuaa acuaucccaaauuuauucaaau
PMVP_2510	658132	658292	161	**MOCO RNA motif**	RF01055	(Regulski et al., 2008)	uguuuaauuggaauuguuugcaugucacaccuuucuaaaauacgag aggauaaaccguuuccagcguaucccugaauagcacaaguuuuaugc uauuggcauuaccacccuauuucuauuguagaaacguagguuaaaua agcucggaguuaugcgauaau

Designated Name[a]	Startbase[a]	Endbase[a]	Length (bp)	Known homologue[b]	Rfam number[c]	Reference	Predicted sequence
PMVP_2511 (PmVP161_ 2215)	703897	704037	141	No hits	None		uucaaaauagaccgaaaaaacaccgcacuuuuagccauggguauaag uacgauaaaaaaauaaaaagaccaaaaucccgagcauaaaucaggauuu gaauaagucgcuuuuucguuaacuuuucugccgccauuuuucucuc
PMVP_2513	832664	832291	374	**tmRNA**	RF00023	(Keiler, 2008)	ugguggagcuggcgggaguugaacccgcguccgaaauuacucuacc uucagcacuacacguuuagucucgucuuuaauuucauuuacccacug cggacagacacgcgauagauaaacuagcuugauucaauuuaacgcuu cgauccucaagcggaggcguccacgcgaucucguuugaauuugacuc cucuugaucccgucuuacgagcggaagcuagggagagaaggcuaug agcagguuauuaagcugcuaaagcguauuguucgucguuugcgacu auuuuuuugcgguuuauuuacgaggccuaccgcaccucgacgugca ccuuggguuucgcuaaucccgucgaauccagaaucagccccaaaaag u
PMVP_2525 (PmVP161_ 2202)	908211	908026	186	No hits	None		ucaauuuuuagguauuuuuagucagaggcgaacacaaaaacacacuu augggaacgaagagccguucauuuuucugaacaacuucugcuccgua caagugaaaucguacaaaaacaguguacuguaacaaccccguguuga ucugaguaaaaaugggugggauaggggauuuaaucugaugaauaa
PMVP_2515	962664	962854	191	**Lysine riboswitch**	RF00168	(Serganov et al., 2008)	aaaaauuaauuaagaccugacccaaccaguagcgcuccauaguuuaaa aaagacuaugacaaugacguuccuuucgaauacgccaccaaccaauca aaaugacuuucaauugauuucggcggucacuccuuuccuuuuuccu cuucguuauccacucagaaaaaauacuuauagugaucgcuccucuac ac
PMVP_2516 (PmVP161_ 2246)	260499	260179	321	No hits	None		gcugucaacaaaaacgaaaagcuggauuuuauaagaaaacccagcuu uuuuaaucguuauuaucacacuaaaugcgauuuuugaacauuuagc uugugaucguuucagccuacacggcuaggcacuauuuacuuacagcc ugcucuaauuuuucacccgcguuuugaauaucuuuaccgacaccuuu cgugguuucacaagcaguuaaagcgaaagcaaaugucgcagauaaga aaauagcaaugauuuuuuucauuguaagcuccuauuuuuucuauag caucaaauaaguccucaacauaagaacuccucuucauccua
PMVP_2517	2215325	2215175	151	**Histidine operon leader**	RF00514	(Chan and Landick, 1989)	caauugaccaauaaggaaacacuaugaauuuuacucgcaugacucau caucaccaucuacccugauuaucuuucggguauuuugugaugcuug gaagauaacaaaucuuccgagugugcgauaaaaugacgaacaauaua aaccucucgga

Designated Name[a]	Startbase[a]	Endbase[a]	Length (bp)	Known homologue[b]	Rfam number[c]	Reference	Predicted sequence
PMVP_2519	1663609	1663829	221	**FMN riboswitch** (RFN element)	FMN RF00050	(Serganov et al., 2009)	cagggcugaugagaaaaguauguggcuaagaaaaucuuaaacaaaca aaaaauacagguuuugcacacaaaaccugcuuauucucuuucaucca gacuauacugucggcuuuggauucucaccaaaucugcugacuuuacu aucucuugugacaugaaauaguaaacgcucgugggcuuauguuaua cauuuaccaccgguagggaauuucacccugcccu
PMVP_2520 (PmVP161_ 2203)	1918697	1918889	191	No hits	None		ugaauauaaaacacacaacaacauacucaacgauuuggaauugacgu uauggcauuacgcccuguacuuagcauuauuauuaaccucguggug auuauuaucacugcggggcgagcgcguuaauguuugacugucaaua aacaaagaagaucaaauuuaacccucgucugauagcagaugaggguu uuuuuua
PMVP_2521 (PmVP161_ 2242)	2100866	2100660	207	No hits	None		gcagacuacgcacugcgugcuucgcuagcguaacuuggcuaaaguuu ucuuuucccuucgggcaacauucacggaaaaaagucacuaugacuaa acaaaaaucauugcuaugacagcgagaauaaugcucaaaccauagccc caguuaagcaaccugaaugaagggcuuauuuuuagaaaaguuacgcc agcgaagcacgcagugcg
PMVP_2522	2078169	2077959	211	**6S/SsrS RNA**	RF00013	(Wassarman and Storz, 2000)	caugaauauugguuguaagggaagaacccgagaugccgcugcaaguc ucaguccuugaacccaacgguucaagggaaucaguuauaacuauuuu uaggcuucuuagucaagcuaagcaugcacaccaacaauuagaaaccg auucauaaaauaaaaguaucggcucagggacauaacccacuggcaaac accucagguaaucuuuacuuaa

[a] PmVP161 locus tags and genome location are relevant to the recently sequenced and re-annotated VP161 genome (Smallman, Boyce Laboratory, manuscript in preparation; NCBI genome accession: CP048792).
[b] Where no hits are listed, there are no known homologues in the Rfam database. Hits annotated based on Rfam database search early 2020.
[c] Where there are no known homologues in the Rfam database, a Rfam number is not designated.

3.3.2. Identification of putative sRNAs produced during growth of *P. multocida* strain X-73 under iron-limited and reduced oxygen growth conditions

A number of putative *P. multocida* sRNAs were identified upon initial analysis of previously generated wild-type VP161 RNA-Seq data. However, the RNA-Seq data used in this analysis was generated from growth in only one type of media and under normal aerobic conditions, as the data was initially generated for a broader transcriptomic study. Moreover, RNA was size-selected to try and exclude very short sequences. Thus, to identify as many putative sRNAs as possible, more comprehensive experiments under different growth conditions were employed to augment these initial findings. Although 23 putative *P. multocida* sRNAs were identified upon initial analysis of previously generated wild-type VP161 RNA-Seq data (Table 3.2), eight of these putative sRNAs had Rfam homologues that indicated they were unlikely to be *trans*- or *cis*-acting sRNAs; these were mostly predicted to be riboswitches or other regulatory RNA elements. Therefore these other predicted regulatory RNAs were not studied further, and instead the focus was on the identification of the subset of *P. multocida* putative sRNAs likely to be *trans*- or *cis*-acting.

To maximise the number of *P. multocida* sRNAs identified, and to determine when they were expressed, high-throughput RNA sequencing of size-selected RNA was undertaken using RNA isolated from *P. multocida* strain X-73 grown under iron-limited conditions, under reduced oxygen conditions, and under normal aerobic *in vitro* growth conditions (in biological triplicate). To avoid the use of RNA isolation columns that are designed to size-exclude small RNA species, and to maximise recovery of sRNAs, total RNA was isolated using TRIzol reagent followed by phenol:chloroform extraction. The quality of each RNA sample was checked by agarose gel electrophoresis (Figure 3.2). A portion of each total RNA sample was then enriched for small RNA fragments (<200 bp) using the *mir*Vana miRNA Isolation Kit (Figure 3.3). The cDNA libraries from each of the total RNA samples (rRNA-depleted) and from each of the corresponding sRNA-enriched RNA samples (without rRNA depletion) were generated and sequenced. These strand-specific sequencing reads were then aligned to the *P. multocida* X-73 genome sequence.

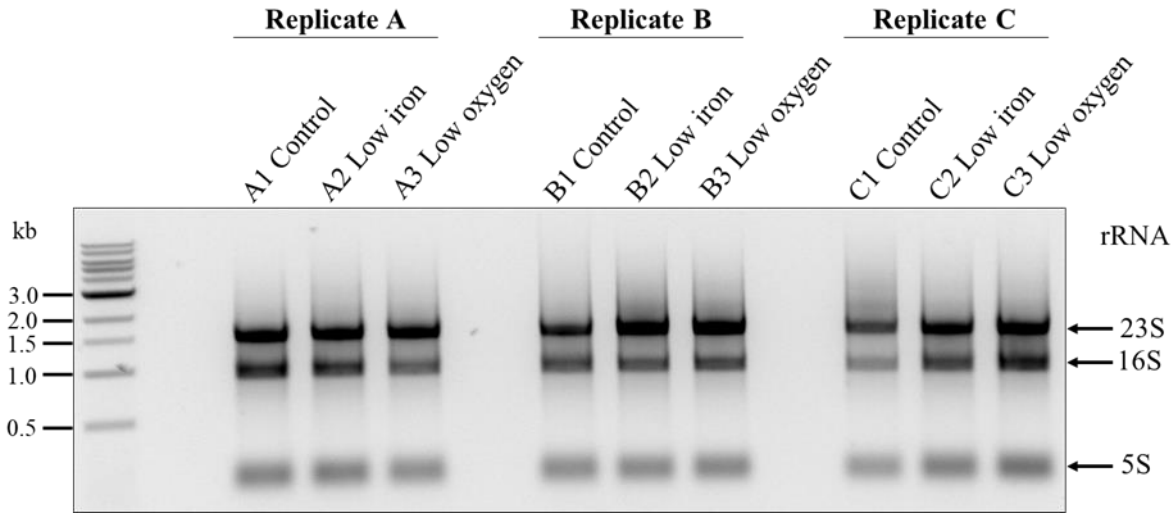


Figure 3.2. Electrophoretic separation of DNase treated RNA samples prior to size selection and rRNA depletion. Triplicate RNA samples from wild-type X-73 mid-exponential *P. multocida* cells grown *in vitro* in rich medium (control), under iron-limited conditions, and under reduced oxygen conditions are shown. Positions of 5S, 16S and 23S rRNA fragments are indicated by the arrows to the right of the gel. Sizes of the molecular weight markers (NEB 1 kb ladder) are indicated to the left of the gel.

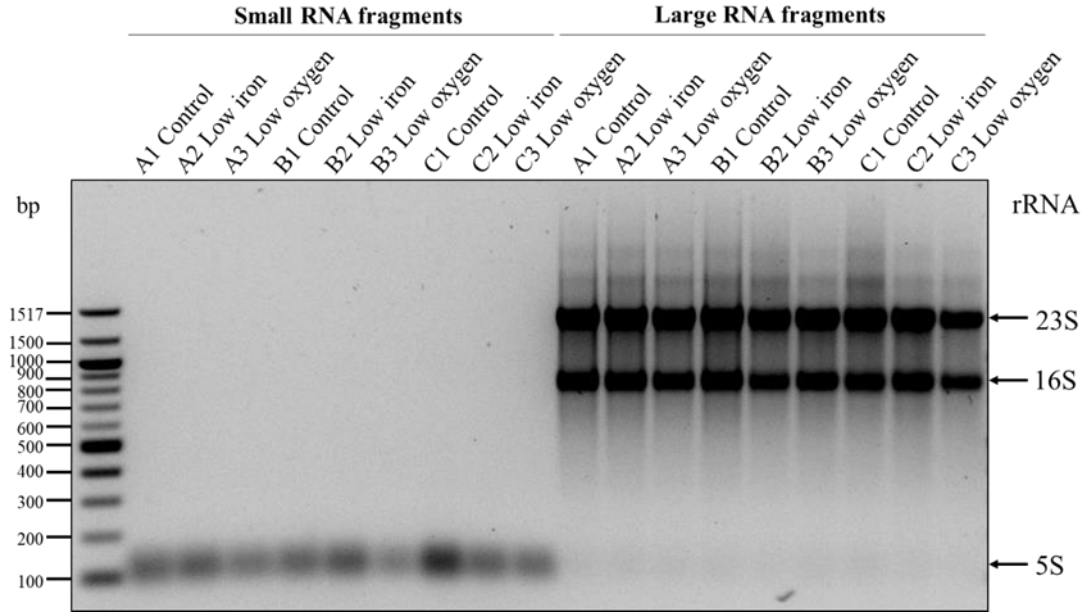


Figure 3.3. Electrophoretic separation of RNA samples following enrichment for small RNA fragments (<200 bp). Samples enriched for small RNA fragments are shown on the left of the gel, samples containing the remaining RNA fraction consisting of large RNA fragments are shown on the right; the positions of the 5S, 16S and 23S rRNA fragments are indicated by the arrows to the right of the gel. Sizes of the molecular weight markers (NEB 100 bp ladder) are indicated at the left.

Visual analysis of the mapped sequencing data was initially performed using Artemis software to identify short regions with a high number of reads in intergenic regions or as part of larger open reading frames (plus or minus strand) (Table 3.3). Regions with a high number of reads representing short RNA transcripts (50-400 nucleotides) were identified as putative sRNAs (refer to Figure 3.4 for examples) and were predominantly found in intergenic regions of the genome (Figure 3.5; category 1, Table 3.4). However, some short transcripts with a high number of reads were found within the 5′ or the 3′ region of annotated open reading frames (Figure 3.5; category 2, Table 3.4). Less commonly, some putative *cis*-acting sRNAs were identified on the opposite strand of a region encoding an open reading frame (Figure 3.5; category 3, Table 3.4).

The sequences of all identified transcripts were used to interrogate the Rfam database (Sanger) and BLAST (NCBI). They were also visually inspected for the presence of Rho-independent terminator sequences, often indicative of true sRNAs. The amount of sequence conservation for each putative sRNA in the three *P. multocida* genomes representing strains X-73, VP161 and Pm70, was also determined. In other bacteria, some sRNA species are known to target the mRNA generated from gene/s in close proximity to their genomic location, such as *cis*-encoded sRNAs, and some of these genes encode key virulence factors (Caldelari et al., 2013). Accordingly, the genes surrounding each of the *P. multocida* putative sRNA sequences were analysed but none of the putative sRNAs were encoded adjacent or in *cis* to known/predicted virulence-associated genes (e.g. capsule/LPS biosynthesis/filamentous hemagglutinin genes).

As a first attempt to identify the RNA targets of the putative sRNAs identified, three commonly used sRNA target prediction databases, Target RNA2, IntaRNA and RNApredator (Busch et al., 2008; Eggenhofer et al., 2011; Kery et al., 2014), were used. The accuracy of applying the prediction tools on experimental data was examined using data generated from a *P. multocida* GcvB sRNA study (Gulliver et al., 2018). In that study, the mRNA targets of GcvB in *P. multocida* were experimentally assessed by examining changes in protein production when GcvB was inactivated (Gulliver et al., 2018). When the entire GcvB sequence was used to interrogate the Target RNA2, IntaRNA and RNApredator sRNA target prediction databases, they were able to predict only 40%, 24% or 4% of the targets that were identified by proteomic analyses, respectively. When just the GcvB seed region sequence (5′-acacaacat-3′) +/-20 bp was used to interrogate these databases, the Target RNA2, IntaRNA and RNApredator databases predicted slightly more of the experimentally predicted GcvB targets (48%, 32% and 16%, respectively). Thus, using the GcvB data, the TargetRNA2 program had the greatest prediction accuracy of the three online prediction tools. Therefore, TargetRNA2 was employed to bioinformatically identify any potential virulence-associated mRNA targets for each sRNA identified in this current study.

The results were limited, but interestingly some predicted target mRNAs had known associations with virulence (Bosch et al., 2002a; Boyce and Adler, 2000; Chung et al., 2001; Tatum et al., 2005). These included mRNA representing an iron acquisition gene (*tonB*), one representing filamentous hemagglutinin (*pfhB2*), four mRNAs representing capsule biosynthesis and transport genes (*hyaC, phyA, hexBC*) and two representing genes encoding outer membrane lipoproteins (*plpB* and *plpE*) (Table 3.4).

All putative sRNAs identified in strain X-73 were compared to those identified in data derived from wild-type VP161 *P. multocida* (Section 3.3.1). Most of the predicted sRNAs identified in VP161 (Section 3.3.1) were also identified in X-73 (Table 3.4). In total, 38 different regions of the *P. multocida* genome were identified as encoding putative sRNAs (Table 3.4). Each putative sRNA was assigned a Prrc number (*P. multocida* regulatory RNA candidate). As it was likely that a proportion of regions identified did not encode true sRNAs, only those which best fulfilled relevant criteria were selected for further analysis (Table 3.3).

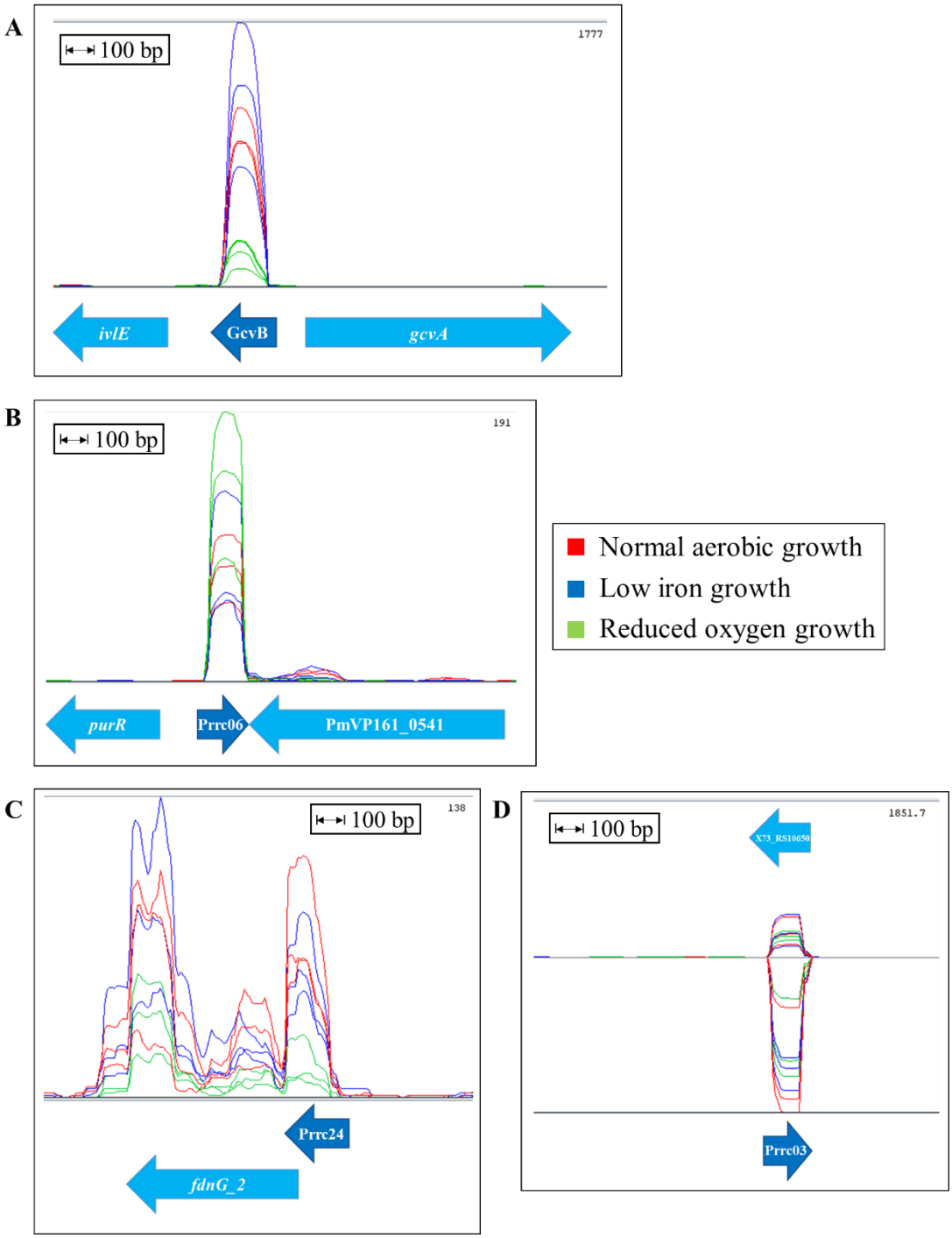


Figure 3.4. Identification of putative *P. multocida* sRNAs by visual inspection of *P. multocida* transcriptomic data. RNA isolated from wild-type X-73, grown at 37°C under normal aerobic growth conditions in HI broth, under iron-limited (low iron), or reduced oxygen conditions, was enriched for small fragments then used to generate sRNA-enriched RNA-Seq data. Transcript sequences were mapped against the X-73 genome and visualised using the Artemis genome browser. Short regions with high numbers of sequence reads were identified by visual inspection and designated as putative sRNAs. Two sRNAs encoded in intergenic regions (category 1) are shown in these examples; the GcvB sRNA (**A**) and the putative sRNA, Prrc06 (**B**). The putative sRNA, Prrc24, which overlaps the 5′ region of *fdnG_2* is shown as an example of a category 2 sRNA (**C**). Putative sRNA Prrc03 is also shown as an example of a predicted *cis*-encoded sRNA (category 3), overlapping X73_RS10650 on the opposite strand (**D**).

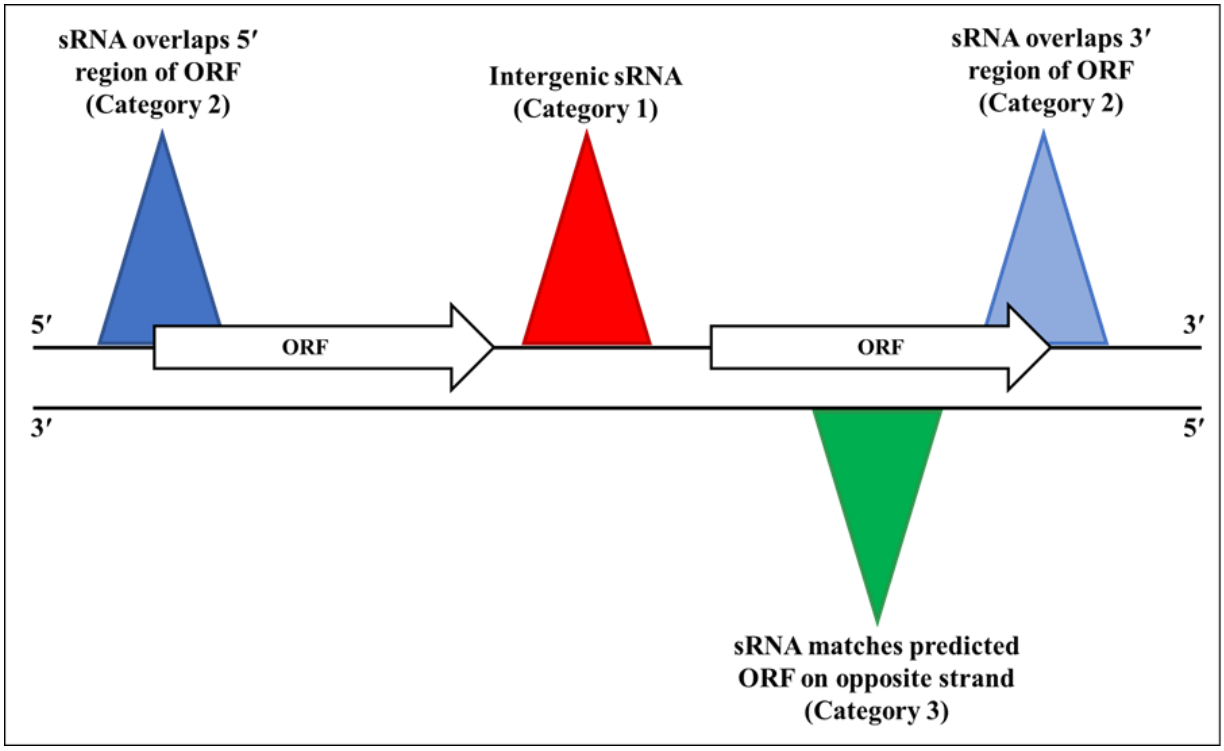


Figure 3.5. Schematic diagram showing the different categories of putative sRNAs. Category 1: putative sRNAs encoded in an intergenic region (*trans*-acting). Category 2: putative sRNAs overlapping sequences of predicted open reading frames (ORF) on the same strand, usually at the 5′ or 3′ region of the ORF. Category 3: putative sRNAs encoded on the opposite strand of a predicted ORF (*cis*-acting).

Table 3.3. Criteria used to identify putative sRNAs from RNA-Seq data.

Criteria	**Percentage of total identified putative sRNAs fulfilling the criteria**
Short sequence (between approximately 50 bp and 400 bp)	100%
Highly expressed sequence (peak height represents >50 transcript reads, as visualised using Artemis software)	60.5%
The sequence has stop codons encoded in all three reading frames	97.3%
Sequence encoded in intergenic location (*trans*-acting sRNAs) - category 1, Table 3.4	52.6%
Sequence overlaps (or partially overlaps) same strand of predicted gene in VP161 or other genome (usually 3′ or 5′ region of gene) - category 2, Table 3.4	34.2%
Sequence matches opposite strand of a strand encoding a predicted gene in VP161 or another *P. multocida* genome (*cis*-acting sRNAs) – category 3, Table 3.4	13.2%
The sequence shares no significant identity with predicted RNA species or protein encoding regions in *P. multocida* or other species (analysis excludes overlapping or putative *cis*-encoded sRNAs)	100%
High conservation of sequence between *P. multocida* strains X-73, VP161 and Pm70 (>95% sequence similarity across entire sequence)	94.7%
Clear Rho-independent terminator	39.5%
The sRNA has a high number of transcripts in three different RNA-Seq datasets: sRNA-enriched strain X-73 data, matching total RNA data, and previously generated VP161 data	13.2%
Interrogation of TargetRNA2 predicted the mRNA targets are likely to be virulence-associated	21.0%

Table 3.4. Putative sRNAs produced by *P. multocida*. The list includes all those identified in the RNA-Seq data generated from strain X-73 grown under various conditions as well as those identified using VP161 RNA-Seq data (see section 3.3.1). Predicted Rho-independent terminator sequences are defined by a string of uridine nucleotides (highlighted yellow), preceded by a stem loop (red and green nucleotide matching base pairs).

Category 1: Intergenic putative sRNA with stop codons on all three strands (*trans*-acting sRNAs)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
GcvB Prrc01 (PmVP161_2200)	183 (591453-591271)	98% X-73 99% Pm70	*plpB* (-12.01) *pfhB2* (-8.51)	1636	6054	auacuuaaugauugguaauuccuuacugguuaagaguuucgg acuuaaguaugauguuguguuugcauauuguuugggaaacca aacagaguaacuuaucuuaaguuacuuauuuucacuuccugu auauuuauuuuagacccuuuuggguuaaacaccgcucuauuga gcgguguuuuuuuau
HrrF Prrc02 (PmVP161_2201)	129 (1210694-1210566)	100% X-73 100% Pm70	None	564	ND	guuagagauaaaggugggaaugauagaguagaucguuuucag ucugagucauacuccaacagguagaguagaaaguaaauucgu uaggguauuuaugggcugaagguuuaaacuucagccuuuuuu uu Note: The underlined sequence region shows 100% sequence conservation among multiple *Pasteurellaceae (Santana et al., 2014)*
Prrc04 (PmVP161_2203)	199 (1918896-1918698)	100% X-73 98% Pm70	None	62	298	uaugauaugaauauaaaacacacaacaacauacucaacgauuu ggaauugacguuauggcauuacgcccuguacuuagcauuauu auuaaccucguggugauuauuaucacugcggggcgagcgcgu uaauguuugacugucaauaaacaaagaagaucaaauuuaaccc ucgucugauagcagaugaggguuuuuuuu
Prrc05 (PmVP161_2204)	163 (1220897-1220735)	99% X-73 100% Pm70	None	29	ND	ucuuacacacuguuuuuuaauuuugauugaaaagaugaauaa acgcacugugauuacuaugauuaugaccaccacuauuauuacc aauaguguгcgggguuagugcguaacaaaaagaucgaauuccac aaaacccguacugaauaaaaagugcgggguuuuuuu

Category 1: Intergenic putative sRNA with stop codons on all three strands (*trans*-acting sRNAs)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc06 (PmVP161_2205)	160 (571608-571767)	99% X-73 99% Pm70	None	179	359	uugauuaauccguagagguaacaaugagugaauuugguuaug caaugaugacaguugcuuuggcgcuagguuuaauuguuguau uggcguucucaauuuuuuaagacuaaacaaaauaaaaauaacg aaagaucgaggcauaaucgccucggcuuuuuuu
Prrc07 (PmVP161_2206)	191 (574358-574548)	100% X-73 100% Pm70	None	20	ND	aucacagcuaagccgaaaucauuuuuacuuuuuuaucuuaaaa gcgugaauuaguucacauuauugaaaauacccauuugcuaacu aagguuuauuuucguauucuguguguagcuaaaaauaagugc uauaaaaaauagccgacauauuuugcuuuccgaguagugccu uguuucaaggcacuuuuuuuu
Prrc11 (PmVP161_2210)	270 (1944418-1944149)	98% X-73 97% Pm70	None	608	ND	agcaaagaaaagcgacccgacuaaguuguugguccucauguca aagccauuuuuuuaacagaaaauaaguugucuucacuucguu gcgcaacguuauuuucuuaaaaaugucuuuucccuucgggca acuuacacggggaaguggguucacccaaaaauaaucgaaagugc ggugcauuuuuggaaaguuucugauuaaacacguaacaagac acuuuccaccuguacgcugcccauacgguagcaugauuguua ggaaaguggcaccaau
Prrc12 (PmVP161_2211)	158 (425188-425345)	98% X-73 99% Pm70	None	735	ND	cauggauaaggauguuugcguuuaacgcaugaaaaaauaaaag aauguauguuuuagguugauuuguauuuaauggauuaacuu gacacaacaucagcucaugccaauagucaacgcauuucauccu uuuaauuaauuugcuaaacgcaauuugauuu
Prrc13 (PmVP161_2212)	150 (425368-425517)	98% X-73 99% Pm70	None	84	ND	ccuguuuuggacgaaaguaagugcaagaagacuaucauacaca cuuguuuuguuaacccucaugcauuuauucguuacaggcacg gaagacacgaugagcaaggaugccuuucaguuaagggcaaagc gauggcuuugcccuuguuuguu Note: no stop codon in frame +3
Prrc15 (PmVP161_2213)	168 (1100818-1100985)	99% X-73 99% Pm70	None	21	ND	acugucguuugacuuaugucaguuguacacauacaauaaaaca uauuugguacgcuuuaaacagccaaucggcacaacugaacaga uuguuagccguaccauuucuguucgauaaagaauauugaaca accuuaucucaaaagccuguuucuugauacaggcuuuuuu

Category 1: Intergenic putative sRNA with stop codons on all three strands (*trans*-acting sRNAs)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc18 (PmVP161_2216)	137 (1924768-1924904)	100% X-73 100% Pm70	None	ND	278	uucucuacguguugccauagugauuacuccuaaguaauuaag acucaaauuugagcaguauuuuacgcuauuuuacugcuaaag gauaaugaauuucuugcuugucuaugaggcuugaagauucau uuucuuuuuuu
Prrc22 (PmVP161_2220)	242 (1442794-1443035)	100% X-73 99% Pm70	None	21	ND	aaauauaccucuaaaauguuauaaaaaauucuagaaacaauca uacuugauuucucuuugcacacaaucguuuucauucucuuug auguaauuuuuuuauaagaauuaaacaauuuuuugguuuaaau caaagaaaugaaauuggcgaaaaaaggagggauaagaaaaggc auaaucagaagagaaaagagaaaaaauuuaggcggugauaaca ccgccuaaauacgaauuauugauguuuu
Prrc25 (PmVP161_2223)	157 (631284-631440)	98% X-73 over last 136 bp 94% Pm70	None	405	ND	uaguagaauucuuuuuaaauuuaaucuugcaauaggauuuug cauuauuaaauauguuuaaagauaaaugaucaguuguauaag uaugaaguuuaaggaguacuuaaugaauuuaucuaaaaucgc aacagggaucguugcgucauuuuguuuuauu
Prrc26 (PmVP161_2224)	133 (188508-188640)	100% X-73 100% Pm70	None	38	ND	cacggaguugugauuugcuauuagacaacaagauagcaaguaa cgaauuguucuaaaauaaagcaauugaauaaaaaauccuuauuu aauugcuugacuuagcaagcuaaaauccguauuauucaccacg cuua
Prrc27 (PmVP161_2225)	147 (180954-181100)	95% X-73 99% Pm70	*hexC* (-11.57)	16	ND	gugauagagagcauuugaagcaugagcaggcggcacgauugg ucaccauuuggcuuguauauaaaauauuaauaaagaaaaucac ggcuaucaaauagcgacuaggugaaagauuaagaaaaacagga aauuuuagauacaaaaaaa
Prrc30 (PmVP161_2228)	153 (2043933-2044085)	97% X-73 96% Pm70	None	29	ND	aaccgcacuuagcuauucuacuuaauagauaagugcggugagc uuuugaguaauuuuuacuugaaguuuggguuguguuuaauu acaauacagccuuuucguaaacaggggagcuagcaauagcgcu auuacccauucuagaggaagcuaaaa

Category 1: Intergenic putative sRNA with stop codons on all three strands (*trans*-acting sRNAs)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc31 (PmVP161_2229)	183 (1416097-1415915)	94% X-73 96% Pm70	None	15	ND	guuuuuuggucuauuuugcuuuuuauugagaagaggacaug uuuuuuacgugaaaaaacaucuuccugguuauugaaaauagc aaaaccuuuagguauuuauuuuguacauaaaauaacaaggaaa aauaauaugaaugggugggaaaaagcucaagagaugauuguu ucaucucuugaacau
Prrc32 (PmVP161_2230)	144 (1385583-1385440)	100% X-73 100% Pm70	None	36	ND	ucaugagccucucuuuauggacggauaugucaagcauaguau ugauuggucuaucuaaucauuacauccauucagauauuaaga gggcgcuuuauugaauacauaaaacgguuuauuucaagugug aaguaaaaucguuuuuuau
Prrc34 (PmVP161_2232)	167 (1170695-1170529)	98% X-73 98% Pm70	None	30	ND	cauaccgcgcaauaacgugcgguauguauuacaacuuuuggca aggcgacagacuacgaccuuagauuugagcgaaguggguuaaaa gugcgguacauuuuuaacuacuuuuucgugucccuuuuuccuu aucucuuugauuuuuauuuucuaugauacacuucaucaaa
Prrc52 (PmVP161_2250)	171	Not identified in VP161. 98% Pm70 identity with X-73	*hexB* (-9.98)	83	417	X-73 sequence: cacuuuuuuaucaauauuaacauuguccccguguaaacugccc gaguguaguggggauuuuuaggaaaauaacguugcguagcaa ggcgaagacaacuuauuuuccgugaaaaaauccacacggaaau cggauaagcaguuuagucgggguguguuuuuugcuccuuuu uu

Category 2: putative sRNA sequence overlaps (or partially overlaps) on sense strand a predicted *P. multocida* protein-encoding gene (usually at 3′ or 5′ region of gene)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc08 (PmVP161_2207)	183 (631545-631727)	97% X-73 98% Pm70	None	8646	ND	uugccgcaaaagcgggaaucgccaaaagcguagaaggugccgc guuaucuuauaacguuggugugaacuaugaguuuuaaucgua aaaaggaguuuaacggacaacauuaauauuauuuaccucuaaa cuaucccaaauuuauucaaauagcgcgcuacguuuucguagcg cgucuuauauaa Overlaps +125 to +307 bps of PmVP161_0591 YadA adhesion protein
Prrc09 (PmVP161_2208)	173 (1410316-1410488)	100% X-73 100% Pm70	None	271	369	gguaauaugaaaaccguuacaguuucaaaacaaggauuuuuuu acuuauguuuaccauguguucuuuaucucucaaccuucugcu cacacagucauugugcggcuagguuguggauaaaaaacaguaa aauaucccacaaauuagacccgcacauguaagcgggucuuuuu uu Overlaps +178 to +350 bps of PM1963 in PM70 LeuA (not annotated in VP161 or X-73)
Prrc10 (PmVP161_2209)	199 (629944-630142)	98% X-73 98% Pm70	None	260	ND	uucguuaugaaaaacucaaaauggaucaggcaaacuacgcagc cugauuauuuuagagagaauagccugaaauacgguuacucaaa guaagaaauauucagugcuuuacaaguaaucucuucggaaauu cgccugaauaucgguguuuaacaaaaaacaaugcggaguuauu aaacaaugccuaaaauuaaaacaguac Overlaps -177 to + 22 bps of PmVP161_0589 50S ribosomal protein L35
Prrc16 (PmVP161_2214)	193 (332757-332949)	100% X-73 100% Pm70	*plpE* (-12.9)	23	ND	auuccuuuauagacaaccugaaucguaauauaauagcgaaaau uuauauacuggucgaacgaucaaauuagcacuugaaaguuaa cguuugcucguuauuuuuaucgaccauauuuuucacaauuuug aauaauugauaaggaauacuccaugaaaagagcgguuauuacu ggguuugguauuauuucgagua Overlaps +1 to +43 bps of PmVP161_0319 FabB

Category 2: putative sRNA sequence overlaps (or partially overlaps) on sense strand a predicted *P. multocida* protein-encoding gene (usually at 3′ or 5′ region of gene)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc20 (PmVP161_2218)	218 (2143627-2143844)	100% X-73 100% Pm70	None	53	ND	aaaaacugcuaaaaucucgcccacaucguuuggugagccuauc uugaaagcgcuugauugauauaacgaugcuuaacuuucgguu gugucucaccgaaaucaauaacaaauuuucagugacuaaacgc aauuagucauugacaaaagcauagaaaauaggcauauuuuacg ccuuuaauuuaucucguuuacuaacaaauaauuuuaggaguu ugacc Overlaps +156 to +216 bps of X73_RS10680 hypothetical protein
Prrc23 (PmVP161_2221)	240 (391103-391342)	100% X-73 99% Pm70	None	33	ND	cuaucuugcauuugcaaaaaugugcagacgacuuaucgucugu uagucgauuuuagguugauugagugauuggcaacaaucauuu aaucgaauuauaaaacuccaaaaacaagguagugauacuauga aaaagacaaucguagcauuagcagucgcagcaguagcagcaac uucagcaaacgcagcaacaguuuacaaucaagacgguacaaaag uugauguaaaacgguucuuuacguuu Overlaps -124 to +116 bps of PmVP161_0368 OmpH_1 Outer membrane protein H
Prrc24 (PmVP161_2222)	239 (413272-413034)	100% X-73 100% Pm70	*hyaC* (-11.63)	106	ND	gguguuuuaauugcugguaauuuacuucguaagacgcauauc aacaauucaacgcuuguggcgauuguaauacaauuuuaacaag ugaaaaacaagucaaaauauuuaauuaucuauuuaacguguuu acgacuaucaaaugcuaacaaaauuuuuguuacgcuaaggagu auuaucuaugcaggucaguagaagaaaguucuucaagaucugu gcaggggaauggcggggacaucug Overlaps +1 to +61 bps of PmVP161_0390 FdnG_2

Category 2: putative sRNA sequence overlaps (or partially overlaps) on sense strand a predicted *P. multocida* protein-encoding gene (usually at 3′ or 5′ region of gene)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc28 (PmVP161_2226)	214 (2185595-2185808)	100% X-73 100% Pm70	*pfhB2* (-11.73)	61	ND	uuuuguucauauuaaguaaaauaucagaaauaaaauaagcauuga gguugaaaaagagggcaauuuauugcgauaaucucaaauauu uuuugggugaucuuucugcggccgauaagggucgcuucucua uuuuguaaaagaucauaauuaucaaguauuuaaugaaaagaag aaaaauuagagauggcagaaaagcguaauauuuuucuuaauugg a Overlaps -181 to +33 bps of PmVP161_2039 AroK Shikimate kinase 1 protein
Prrc29 (PmVP161_2227)	379 (2048087-2048465)	99% X-73 99% Pm70	*phyA* (-8.3)	155	ND	aagaaaggggauuucuauuaaaauuucgcuccugaauuuugu cauuuuuaggugaugaaauuguugauggugcuugaccacauu caagccccguccgagaccguaggugaaaugauggauaucaucu cuuaauuuggccuacguagauggugaacagacagaauuuucu gcuucuggacaccuuagccucacaaaauagucguaccaaagug cggucauuuuuugagauuuuuuguaaauuccgcuuaggcgga guguauucaggagcuaaaagccaauggcauuaaaaucuucaaga caaacaagcaauuguugcugaaguaaacgaagcagccaaaggu gcccuaucagcugugaucgcggauucacgcgguguaaca Overlaps -277 to + 102 bps of PmVP161_1909 50S ribosomal protein L10
Prrc35 (PmVP161_2233)	187 (1144267-1144453)	99% X-73 99% Pm70	None	11	ND	gaaauaaaaucucagaccucaccgaacgcacgcaaguguugag gagguaucuuugacuucgaugcgaaaguugaacguuuagaag aagucaaugccgaacucgaacaaccugauguguggaacgaacc agaaaaagcacaagcauuagguaaagaacggguugccuuagaa ggggucgucaauacaa Overlaps -114 to + 73 bps of PmVP161_1041 PrfB

Category 2: putative sRNA sequence overlaps (or partially overlaps) on sense strand a predicted *P. multocida* protein-encoding gene (usually at 3′ or 5′ region of gene)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc36 (PmVP161_2234)	314 (886633-886946)	99% X-73 99% Pm70	None	64	ND	aauuaaauauauuuaauuaaaaauaaaauuaauaaaauguuaauaaauuuaaaauaugguuaaauuauaaacaaaaaaugugauggauaucacaaauaaaaaaauuuaucuugaaaagauaagggaaauuagcgagacuaaugacuaaaugaaaguaaggcaaucuaaauguuaaauaaauaauacacuuggagaguuuuauguuaauacacgacuacuacaaagaauaucauccugauaaaggacguuccauuauggcacauauucguagaacaagacauauuaugaugccuucguaucgcucuuguuuuucuu Overlaps -199 to + 115 bps of PM0805 in Pm70 (not annotated in VP161 or X-73)
Prrc37 (PmVP161_2235)	131 (2230876-2231006)	100% X-73 100% Pm70	None	17	ND	auuuaauuuucguuccuugcuuccgcagauggguagcuggucuaacagucggaggcugauuaacccguaaggagcaguaaugcgacacuacgaaaucguguuuaugguucauccggaccaaagcgaacaag Overlaps -79 to +52 bps of PmVP161_2087 RpsF 30S ribosomal protein S6
Prrc38 (PmVP161_2236)	211 (1978300-1978510)	100% X-73 100% Pm70	None	62	ND	auauauacagagaaaaaguaaguuuaacucucaagauuucuucgauaauugaggguuuucuuuuuguaaucaaacauuuagaggaagguuaugguaaccauucguuuaucucguggcggagcuaaaaaacgcccauucuaucaaaucguugucgcagauagccguucaccacgugacggucgcuuuauugagcgcguaggcuuuuucaac Overlaps -91 to +120 bps of PmVP161_1834 RpsP encoding a 30S ribosomal protein S16

Category 3: putative sRNA sequence overlaps (or partially overlaps) on antisense strand a predicted *P. multocida* gene (*cis*-acting sRNAs)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc03 (PmVP161_2202)	147 (908172-908026)	100% X-73 100% Pm70	None	1852	10033	acacacuuaugggaacgaagagccguucauuuuucugaacaac uucugcuccguacaagugaaaucguacaaaaacaguguacugu aacaaccccguguugaucugaguaaaaaugggugggauaggg gauuuaaucugaugaauaa Overlaps +11 to +157 of X73_RS10650, hypothetical protein
Prrc14 (PmVP161_2256)	155 (1013381-1013227)	97% X-73 98% Pm70	None	500	720	gagaggcuuaacuauacaacgcucaaguguuuuuggguguug agugagacguuagacgaagacgaagaagagagaauuacuuuag acagcuuguuuuugauuaaagugcgagagugaagaaagcacga aggaagaacagaagaauagagaagaca Overlaps +2883 to +3038 of PmVP161_0276 23S rRNA
Prrc17 (PmVP161_2215)	181 (703897-704077)	99% X-73 99% Pm70	*tonB* (-9.92)	ND	124	uucaaaauagaccgaaaaaacaccgcacuuuuagccaugggua uaaguacgauaaaaaaauaaaaagaccaaaaucccgagcauaau caggauuugaauaagucgcuuuuucguuaacuuuucugccgc cauuuuucucucucuugcuuauauuuaacacuucuuaauugu aacauuuaau Overlaps -49 to +132 of PmVP161_0655, hypothetical protein
Prrc19 (PmVP161_2217)	161 (1689286-1689126)	100% X-73 100% Pm70	None	954	ND	ucgcaagcuuucucacagagagccguuuaauguaaucauuuau uaaacgaguauuacaauagguaauacguaaaauguuuccgauu gggaaacuauaaacaggucaacuugcaccccuuuucguuuuua gaaguaagaaaaguguugguguuaguuuuuua Overlaps PMXST_01448, hypothetical protein; 117 to +115

Category 3: putative sRNA sequence overlaps (or partially overlaps) on antisense strand a predicted *P. multocida* gene (*cis*-acting sRNAs)

Designated Name/Prrc# (new locus tag) [a]	Nucleotide length[a] (Position on genome)	Percentage identity with region in other *P. multocida* genomes [b]	Predicted virulence-associated mRNA targets [c]	Peak number of transcripts in X-73 sRNA-enriched data [d]	Peak number of transcripts in X-73 in non-enriched data	Predicted sRNA sequence (5′ to 3′) [ef]
Prrc21 (PmVP161_2219)	225 (1610308-1610532)	99% X-73 99% Pm70	None	25	ND	ugcuauaaacagacauagaaacaucuauucuauuaaacaaccaa aaagagguaggaaauaugagauagaucaacuuuuaucuuuuaa gacaaaaaauaagccauauuauaugauauaaaucaaucacuua uaaaaaugauaauaauuagcauucaauuagaauaggcguaaaa aagaccaugaugugauacucauuaccacaucaugggucuguuuc augucaugc Overlaps start of PM1754, DmsA in PM70; +1 to +73

[a] Re-annotated VP161 genome (Smallman, Boyce Laboratory, manuscript in preparation; NCBI genome accession: CP048792).
[b] The fully annotated reference genome for *P. multocida* strains Pm70 and X-73 are publicly available through National Center for Biotechnology Information Search database (www.ncbi.nlm.nih.gov).
[c] Predicted virulence-associated mRNA targets were identified using TargetRNA2 (Kery et al., 2014); the predicted hybridisation thermodynamic energy (kcal/mol) value between the sRNA and target RNA molecules are reported as determined using TargetRNA2.
[d] Number of transcript reads at the height of the transcript peak.
[e] RNA sequence shown is representative of the VP161 sequence unless otherwise stated.
[f] Where known, sequence start sites are reported as determined from primer extension experiments (reported in Chapter 5); for GcvB the sequence start site was determined using 5′ RACE (Gulliver et al., 2018).

3.3.3. Changes in *P. multocida* gene expression under reduced oxygen and low iron growth conditions

The RNA-Seq reads generated from the total RNA samples (isolated from strain X-73) were first used to identify gene expression changes during growth under low iron conditions and reduced oxygen conditions, compared to gene expression during aerobic growth in rich medium (HI broth). Genes were identified as differentially expressed if they displayed a ≥ 2.0-fold (≥ 1.0-$\log_2$) change in expression across the all three replicates at a false discovery rate (FDR) of < 0.05. Differentially expressed X-73 genes that shared high identity with genes assigned a functional annotation in *P. multocida* strain Pm70 (May et al., 2001) or in other bacterial species were assigned the same gene name (see Tables S9, S10, S11 and S12 in Appendix C).

A total of 36 X-73 genes had increased transcript levels during growth under low iron conditions, compared to growth in rich medium (see Table S9 in Appendix C). As predicted, these included a number of genes associated with iron-acquisition, including *fbpA, fecBCD, yfeABCD*, many of which displayed altered expression in response to low iron in a previous *P. multocida* DNA microarray study (Paustian et al., 2001). A total of 19 genes displayed reduced transcript levels in low iron conditions (see Table S10 in Appendix C), including electron transport genes *nrfABCD* that are required for nitrate reduction in *P. multocida* (Boucher et al., 2005). Two genes encoding a transcriptional repressor of aerobic metabolism, *arcA*, and an outer membrane protein, *ompW*, also displayed reduced transcript levels in low iron conditions; both were reported to have reduced expression under low iron conditions in the previous study on the effects of iron limitation in *P. multocida* (Paustian et al., 2001).

A total of 213 genes displayed increased transcript levels and 146 genes displayed reduced transcript levels during growth under reduced oxygen conditions, compared to aerobic growth in rich medium (Tables S11 and S12 in Appendix C, respectively). Genes that displayed increased transcript levels in reduced oxygen conditions included electron transport genes (*ccmABCDEFH, nrfABCDEF*), anaerobic electron transport genes (*frdABCD, dmsABCD*), anaerobic ribonucleoside-triphosphate reductase genes (*nrdDG*), anaerobic glycerol-3-phosphate dehydrogenase genes (*glpABC*) and genes encoding components of the periplasmic nitrate reductase system (*napADH*). Genes that displayed decreased transcript levels under reduced oxygen conditions included 38 that encoded ribosomal proteins, as well as *fnr*, a transcriptional regulator involved in anaerobic metabolism.

3.3.4. Changes in expression of putative sRNAs under reduced oxygen and low iron growth conditions

In this study, each of the identified putative sRNAs was precisely annotated to the *P. multocida* X-73 and VP161 genomes. This allowed for the expression of each sRNA to be determined in the different RNA-Seq datasets, including those representing transcript changes that occurred during growth in low iron conditions and during growth under reduced oxygen conditions. Expression of these putative sRNAs was also analysed in the RNA-Seq data available for the *hfq* mutant (Chapter 2, AL2521); Hfq is involved in sRNA-mediated gene regulation (Chapter 2). In addition, RNA-Seq data generated from a *P. multocida proQ* mutant (AL2973; Gulliver, Boyce Laboratory, manuscript in preparation) was also examined for sRNA expression; ProQ is a recently described bacterial riboregulatory protein. Comparison of the RNA expression data revealed that many of the putative sRNAs displayed altered expression under one or more of the specific growth conditions (Table 3.5), indicating that they may be involved in regulating the expression of specific genes under those particular conditions. These differentially expressed putative sRNAs included HrrF and Prrc14, which showed increased and decreased expression during growth under low iron conditions, respectively. Putative sRNAs Prrc06 and Prrc34 both showed increased expression during growth under reduced oxygen conditions, while GcvB, Prrc14, Prrc19, Prrc29, Prrc30, Prrc37 and Prrc52 showed decreased expression under the same condition. One putative sRNA also displayed altered expression in the *hfq* mutant and eleven sRNAs were differentially expressed in the *proQ* mutant (ten decreased and one increased in expression) (Table 3.5). Therefore, it is possible that sRNA expression was affected by the loss of these RNA chaperone proteins.

Table 3.5. Differentially expressed putative sRNAs identified in *P. multocida* wild-type strain X-73 when grown under low iron or reduced oxygen conditions, compared to aerobic growth in rich medium. Differentially expressed putative sRNAs identified in the *P. multocida hfq* mutant and *P. multocida proQ* mutant (compared to expression in the wild-type strain VP161) are also listed.

sRNA	**Low iron growth**		**Reduced oxygen growth**		***hfq* mutant**		***proQ* mutant** [c]	
	Expression ratio (log_2) [a]	**FDR** [b]	**Expression ratio (log_2)** [a]	**FDR** [b]	**Expression ratio (log_2)** [a]	**FDR** [b]	**Expression ratio (log_2)** [a]	**FDR** [b]
GcvB	NC	NA	-2.18	2.18E-03	NC	NA	NC	NA
HrrF	1.76	8.15E-04	NC	NA	NC	NA	NC	NA
Prrc05	NC	NA	NC	NA	NC	NA	-1.25	9.43E-03
Prrc06	NC	NA	1.35	1.77E-04	NC	NA	-2.25	2.01E-03
Prrc09	NC	NA	NC	NA	NC	NA	-1.09	1.77E-02
Prrc10	NC	NA	NC	NA	NC	NA	-2.37	1.41E-02
Prrc11	NC	NA	NC	NA	2.02	9.09E-03	NC	NA
Prrc12	NC	NA	NC	NA	NC	NA	-1.43	6.15E-04
Prrc14	-1.41	3.08E-02	-2.10	1.10E-03	NC	NA	NC	NA
Prrc15	NC	NA	NC	NA	NC	NA	-1.52	3.88E-03
Prrc19	NC	NA	-1.43	4.33E-06	NC	NA	-1.56	3.20E-03
Prrc20	NC	NA	NC	NA	NC	NA	-1.28	9.22E-04
Prrc24	NC	NA	NC	NA	NC	NA	-1.20	1.89E-03
Prrc25	NC	NA	NC	NA	NC	NA	1.51	9.14E-03
Prrc29	NC	NA	-1.05	4.33E-08	NC	NA		
Prrc30	NC	NA	-1.28	1.49E-03	NC	NA	-2.48	2.23E-03
Prrc34	NC	NA	1.94	2.30E-02	NC	NA	NC	NA
Prrc37	NC	NA	-1.43	2.96E-04	NC	NA	NC	NA
Prrc52	NC	NA	-1.14	5.53E-05	NC	NA	NC	NA

[a] NC, no statistically significant change in expression ratio.
[b] NA, not applicable.
[c] Data kindly provided by Emily Gulliver, Boyce Laboratory (manuscript in preparation).

3.4. Discussion

The results presented in Chapter 2 showed that Hfq plays an important role in *P. multocida* riboregulation. Inactivation of this RNA chaperone protein, which usually acts to facilitate the interaction between sRNAs and their mRNA targets, affected the expression of genes involved in the production of the virulence factors, capsule, filamentous hemagglutinin and LPS. However, with the exception of the *P. multocida* sRNA GcvB (Gulliver et al., 2018), very little is known about the sRNAs encoded by *P. multocida* and the riboregulatory role they play in this organism. Therefore, in this chapter a number of techniques were employed, including RNA-Seq and bioinformatic analyses, to identify the full range of sRNAs produced by *P. multocida* and investigate their expression under two growth conditions, relevant to growth *in vivo*; growth in low iron and growth under reduced oxygen conditions.

Recent advances in RNA-Seq and bioinformatic techniques have allowed for the discovery of many novel sRNAs in a diverse range of bacterial species (Barquist and Vogel, 2015). Model organisms used in the initial RNA studies, *E. coli* and *S.* Typhimurium, are now known to each express more than 200 sRNAs. Large numbers of sRNAs have now also been identified in a range of other Gram-negative bacterial species including *Pseudomonas aeruginosa*, and also in Gram-positive bacterial species, including *Staphylococcus aureus* and *Listeria monocytogenes* (Gómez-Lozano et al., 2012; Howden et al., 2013; Mellin and Cossart, 2012). As the total number of sRNAs in each species is variable, and there is generally only limited conservation of sRNA primary sequence across bacteria, limiting the use of standard bioinformatic analyses, the precise number of sRNAs expressed by *P. multocida* was unknown (Lindgreen et al., 2014). In this chapter, we identified 38 putative *P. multocida* sRNAs using RNA-Seq analyses, including those predicted to be *trans*- and *cis*-acting sRNAs. Only two of the sRNAs identified in this study shared significant nucleotide identity with other well characterised sRNAs previously identified and submitted into the Rfam database, namely HrrF and GcvB; the role of GcvB in *P. multocida* has since been examined in detail in our laboratory (Gulliver et al., 2018). The total number of *P. multocida* sRNAs identified in this study was quite low compared to the total number found in other species such as *E. coli* and *S.* Typhimurium. It is possible that *P. multocida* actually does produce less sRNA species, particularly as *P. multocida* encodes less than half of the number of genes encoded by *E. coli* or *S.* Typhimurium. However, our analysis employed only a limited number of techniques to identify these molecules and the expression of sRNA was only examined under a few different bacterial growth conditions. Therefore, we predict that there are more *P. multocida* sRNAs yet to be identified.

Bacterial sRNA expression and subsequent gene regulation is often dependent on the growth phase of the bacterium, and/or on different growth conditions, including differences in the conditions that bacteria experience within host niches. The expression of these sRNA molecules allows the bacterium to make regulatory changes as required to best optimise protein production to the specific conditions. Thus, individual sRNAs may only be expressed under very specific conditions. For example, large studies involving the pathogens *S.* Typhimurium and *Burkholderia thailandensis*, grown under 22 and 54 distinct infection-relevant growth conditions, respectively, identified many sRNA candidates and demonstrated clear evidence of condition-dependent expression for each sRNA species (Kröger et al., 2013; Stubben et al., 2014). In our study, RNA-Seq analyses were performed on *P. multocida* grown under normal aerobic growth conditions in rich medium (HI broth) and two infection-relevant growth conditions; namely growth in iron-limited media and growth under reduced oxygen conditions. Iron is essential for most bacteria, including *P. multocida* (Bosch et al., 2002b) and the *P. multocida* strain Pm70 genome encodes a large number of iron acquisition and transport associated genes (May et al., 2001). However, access to free iron is limited during growth in most hosts, including chickens (Boyce et al., 2002; Wilkie et al., 2012). A DNA microarray study examining *P. multocida* gene expression under iron-limiting conditions demonstrated significant changes in the expression of a wide range of genes involved in energy metabolism and transport, as well as iron binding and transport (Paustian et al., 2001). In another DNA microarray study investigating the transcriptional response of *P. multocida* in the natural chicken host, two out of three gene expression profiles from cells harvested from the blood of infected chickens displayed changes in gene expression similar to that observed under iron-limited *in vitro* growth, highlighting the importance of iron acquisition systems during infection in the host (Boyce et al., 2002). Genes involved in iron acquisition were also identified as virulence-associated in a signature-tagged mutagenesis study in a septicemic mouse model, including *exbB* (Fuller et al., 2000).

The role of sRNAs in the regulation of iron-associated genes has previously been identified in a range of bacterial species; sRNAs include the well characterised Fur-regulated sRNA in *E. coli* called RhyB (Oglesby-Sherrouse and Murphy, 2013). It is highly likely that iron acquisition and utilization in *P. multocida*, both *in vitro* and *in vivo*, involves sRNA riboregulatory activity and our study included analysis of sRNA expression in iron-depleted media for this reason. *P. multocida* is a facultative anaerobe (Wilson and Ho, 2013), allowing the bacterium to survive in anaerobic niches (e.g. abscesses/wounds) during infection in the host. But little is known about *P. multocida* gene expression during growth in reduced oxygen conditions. Our results provide new insights into *P. multocida* gene expression under such conditions and importantly revealed

sRNA expression during this specific condition. Indeed, if additional growth phase and infection-relevant growth conditions were explored, for example analysis of *P. multocida* RNA-Seq experiments in chicken serum, other sRNAs specifically expressed under those conditions may be also identified.

To maximise sRNA discovery, different types of RNA-Seq libraries were prepared and analysed. Initially, previously generated RNA expression data in *P. multocida* strain VP161, grown in rich medium at multiple growth phases, was used to identify sRNAs and a number of putative sRNAs were identified from this data. However, these RNA-Seq data were not generated specifically for sRNA discovery and the RNA used for sequencing had been size selected for RNA fragments approximately 100 bp or larger in size, making it likely that many of the smaller sRNAs species would have been excluded. Therefore, when examining RNA expression data from *P. multocida* strain X-73 grown aerobically in rich medium, in iron-restricted medium or under reduced oxygen conditions, a total RNA library (rRNA depleted) as well as a library enriched for sRNAs (fragments < 200 bp) was prepared for each sample. In addition, the sRNA-enriched libraries were prepared using a strand specific method, providing important information about the direction of sRNA transcription and allowing for identification of *cis*-encoded putative sRNAs. A similar technique using multiple RNA-Seq libraries to maximise sRNA discovery was used to study sRNAs in *P. aeruginosa* (Gómez-Lozano et al., 2012); many novel sRNAs were identified, and the study demonstrated the importance of using multiple RNA-Seq library preparation strategies, as many sRNAs were unique to each of the different libraries. Libraries enriched for smaller RNA transcripts yielded a greater number of sRNA species and it was therefore hypothesised that larger transcripts may potentially mask the detection of sRNAs (Gómez-Lozano et al., 2012). Our study supported this hypothesis, with approximately 95% of the putative *P. multocida* sRNAs identified in the sRNA-enriched RNA-Seq library samples, compared to only 24% identified in the total RNA library samples.

The majority (approximately 53%) of putative *P. multocida* sRNAs identified in this current study were encoded in predicted intergenic regions and were annotated as putative *trans*-acting sRNAs (sRNA category 1, Table 3.4). Identifying *trans*-acting sRNAs was the focus of this study, as they are generally more easily visually identifiable from RNA-Seq data than other types of regulatory RNA molecules. A smaller number of predicted sRNAs were found in regions that overlapped the predicted 5′-UTR region upstream, the 3′ region downstream, and/or the amino acid coding region of adjacent genes (sRNA category 2, Table 3.4). Numerous putative *cis*-acting sRNAs were also identified. The sequence of *cis*-acting sRNAs partially or completely matches the opposite strand of a predicted gene and is therefore likely to bind to, and regulate, its expression (category 3, Table

3.4). The putative *cis*-acting sRNAs identified in *P. multocida* included Prrc03 (predicted to be 147 bp in size), which had 100% reverse complementarity with nucleotides 11 to 157 of a small gene (189 bp), X73_RS10650, encoding a protein of unknown function. Homologues of the X73_RS10650 gene are also present in strain VP161 and in strain Pm70, though the gene was not annotated in the available genome of the latter. Also identified was Prrc21, the coding region of which is predicted to overlap nucleotides 1 to 73 of *dmsA*, an anaerobic dimethyl sulfoxide reductase involved in anaerobic electron transport. The gene *dmsA* displayed increased RNA expression when *P. multocida* was grown under reduced oxygen conditions. Interestingly, the sRNA Prrc21 was expressed under all growth conditions but, unlike *dmsA*, was unchanged in expression when the strain was subjected to reduced oxygen conditions. Given that Prrc21 is encoded in *cis* to the 5′ region of *dmsA*, it is likely that the predicted sRNA binds to and regulates *dmsA* transcripts under some conditions. However, this is likely occurring at some other level of regulation as the expression of Prrc21 was unchanged during *P. multocida* growth under reduced oxygen conditions.

In this study, the sRNA GcvB was also identified. This sRNA shares a high level of identity with the well characterised GcvB sRNAs from *E. coli* and *S.* Typhimurium and the molecule is highly conserved across many Gram-negative species (Gulliver et al., 2018; Sharma et al., 2011). GcvB is a Hfq-dependent, *trans*-acting sRNA, with many mRNA targets. In *E. coli* and *Salmonella*, GcvB has been shown to regulate amino acid biosynthesis and transport processes by suppressing the production of proteins involved in amino acid metabolism when there is an abundance of nutrients in the environment, ultimately conserving energy for the cell (Pulvermacher et al., 2009; Sharma et al., 2011). Studies in the Boyce Laboratory, carried out in parallel with the studies reported in this thesis, have demonstrated a clear role for GcvB in amino acid biosynthesis and transport processes in *P. multocida* (Gulliver et al., 2018). *P. multocida* GcvB was shown to repress the production and transport of amino acids during the early growth stages of *P. multocida*, similar to what was observed in other bacterial species (Gulliver et al., 2018); however, the precise target genes were quite different from those observed for the *E. coli* and *Salmonella* GcvB targets. In this current study, GcvB showed decreased expression when *P. multocida* was grown under reduced oxygen conditions, compared to growth in normal aerobic conditions *in vitro*; as GcvB acts primarily to repress expression this indicated that the production and transport of amino acids was induced. The metabolic processes required for bacterial growth in anaerobic conditions are significantly different from those required for growth in aerobic conditions. An *in vivo* *P. multocida* gene expression study conducted in the chicken host showed that many genes involved in the sensing of the presence and/or absence of oxygen were differentially regulated

(Boyce et al., 2002). Interestingly, in the related bacteria *A. pleuropneumoniae*, the level of GcvB expression was seemingly unchanged under aerobic conditions compared to anaerobic conditions; however this was only determined using northern blotting (Rossi et al., 2016). GcvB is known to be a Hfq-dependent sRNA in other species. Therefore, to study the interaction between *P. multocida* GcvB and Hfq, we performed further experiments including Hfq co-immunoprecipitation; the results of these experiments are reported in Chapter 4 of this thesis.

One sRNA which displayed increased expression in iron-limited growth conditions is HrrF, orthologues of which are found in a number of *Haemophilus* species and other members of the *Pasteurellaceae* family (Santana et al., 2014). In *H. influenzae*, HrrF is involved in iron homeostasis and is regulated by Fur, a ferric uptake regulator involved in the control of iron stress (Santana et al., 2014). The *H. influenzae* HrrF sRNA shares very high sequence identity with HrrF in *P. multocida* (Santana et al., 2014). Moreover, within the *P. multocida* HrrF sequence (predicted to be 129 nucleotides long, Table 3.4) is a 32 nucleotide region that shares 100% sequence conservation with HrrF from all *Pasteurellaceae* family members examined within the *Pasteurellaceae,* indicating that this region in HrrF may be important for sRNA function (Santana et al., 2014). The *P. multocida* sRNA HrrF is encoded downstream of *pmbA* as it is in *H. influenzae*. Upstream of the HrrF promoter in *H. influenzae* (Figure. 3.6) (Santana et al., 2014) are several ferric uptake repressor (Fur) boxes and visual examination of the sequence upstream of the *P. multocida* HrrF sRNA identified a sequence that shared a high level of identity to this region. Future work could involve the use of a tool such as Visual Footprint (Münch et al., 2005) to precisely identify the boundary of each Fur box as was done for *H. influenzae* HrrF (Santana et al., 2014). Santana and co-workers demonstrated that *H. influenzae* HrrF is predicted to control the expression of genes involved with molybdate uptake, deoxyribonucleotide and amino acid biosynthesis and is most highly expressed when there are low levels of iron availability. Moreover, the stability of the *H. influenzae* HrrF molecule is not Hfq dependent (Santana et al., 2014). In this current study, HrrF in *P. multocida* strain X-73 also displayed increased expression under low iron conditions, indicating that it may play a similar role in this bacterium. Future work could include studies similar to those conducted on *H. influenzae* (Santana et al., 2014) in order to identify the specific HrrF targets in *P. multocida*. As analysis of the *P. multocida* sRNA GcvB showed that while the general function of its target genes was similar to that observed in other species (amino acid metabolism and transport), the specific gene targets were quite different. Therefore, it is likely that specific HrrF targets may also differ between *H. influenzae* and *P. multocida*.

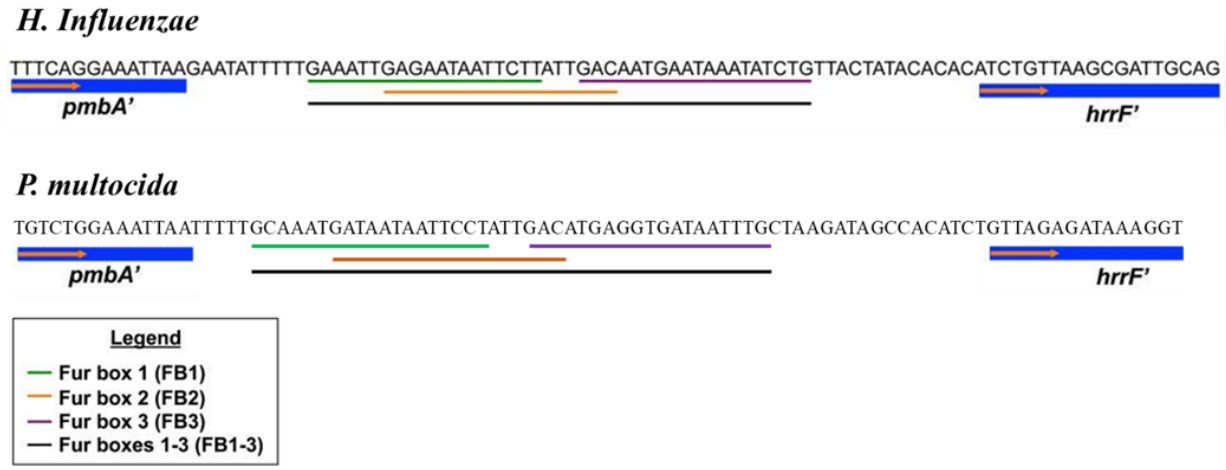


Figure 3.6. Location of the three predicted Fur boxes in the *hrrF* promoter region in *H. influenzae* and *P. multocida*. Fur boxes 1, 2, 3, and 1-3 are indicated by the green, orange, purple and black lines respectively. Modified from Santana et al., 2014.

P. multocida encodes a large number of proteins with iron-associated functions and some of these are essential for virulence (Jatuponwiphat et al., 2019). In this study, a range of *P. multocida* iron-associated genes displayed altered expression after a very short exposure to low iron conditions. Short, shock exposures to a range of growth conditions has been used previously for discovering sRNA species in *S.* Typhimurium and was found to be a very effective way to determine the immediate gene expression response to a particular environment (Kröger et al., 2013). *P. multocida* was subjected to a short, shock exposure by incubating in media depleted of iron for only ten minutes. Despite being a relatively short period, it is clear that this level of exposure was sufficient to cause an immediate regulatory response. Many of the genes identified in this study had been identified in a previous study on *P. multocida* strain Pm70 that examined gene expression changes in response to low iron using DNA microarrays (Paustian et al., 2001). In our study, we found many *P. multocida* genes with altered expression during short exposure to low iron conditions, including genes required for iron-acquisition and transport into the cytoplasm, namely *pfhR, fbpAB, fecBCDE,* and *yfeABCD*, all of which displayed increased transcript levels. The electron transport genes *nrfABCD* all displayed reduced transcript levels. This was to be expected as the products of these genes are dependent on iron for executing cellular electron transport processes. The transcriptional repressor of aerobic metabolism, *arcA*, also displayed reduced transcript levels in this study and this was also observed in the DNA microarray study on iron-depleted *P. multocida* (Paustian et al., 2001). The gene encoding the outer membrane protein, OmpW, displayed reduced transcript levels in *P. multocida* grown in iron-limited media as was observed for this gene in the *P. multocida* DNA microarray study under low iron conditions (Paustian et al., 2001). Interestingly, another DNA microarray study showed that the expression

of *ompW* was also altered during growth in minimal media supplemented with specific defined iron sources, namely transferrin and haemoglobin (Paustian et al., 2002). Corresponding to gene expression changes, OmpW production was also significantly reduced in abundance during iron-limited growth of *P. multocida* in a study analysing the components of the outer membrane sub-proteome (Boyce et al., 2006).

The *P. multocida* transcriptomic response to reduced oxygen conditions has not previously been reported. To identify genes and sRNAs expressed under such conditions, this study explored how reduced oxygen availability affected *P. multocida* total RNA expression. The *P. multocida* cells were exposed to reduced oxygen conditions for a short period of time (a 50 mL tube was completely filled with culture, avoiding agitation, then closed and incubated static at 37°C for 30 min) and a sample of cells from the bottom of the tube carefully removed for RNA isolation. For comparison, gene expression data was also prepared from cells taken from the same primary culture but grown under normal aerobic laboratory conditions for 30 min (HI broth incubated at 37°C, with shaking). The RNA-Seq data revealed that there was a clear effect on gene expression with a large number of genes displaying altered transcript levels during growth under reduced oxygen conditions compared to growth in rich medium. In total 213 genes displayed increased expression and 146 genes displayed decreased expression. Genes that displayed increased transcript levels included many that are known in other bacteria to be associated with anaerobic growth such as electron transport genes (*ccmABCDEFH, nrfABCDEF*), anaerobic electron transport genes (*frdABCD, dmsABCD*), anaerobic ribonucleoside-triphosphate reductase genes (*nrdDG*) and anaerobic glycerol-3-phosphate dehydrogenase genes (*glpABC*). Also identified as having increased expression were the genes within the periplasmic nitrate reductase system (*napADH*) which has been shown in other bacteria to reduce nitrate to nitrite during anaerobic growth (Cole, 1996). The *P. multocida nap* genes also showed significantly increased expression during *in vivo* growth in the chicken host (Boyce et al., 2002) and many of the electron transport genes listed above also showed increased expression in the *P. multocida hfq* mutant (see Chapter 2).

Interestingly, genes that displayed reduced transcript levels during growth in reduced oxygen included thirty-eight that encoded 30S or 50S ribosomal proteins. In other bacteria it has been shown that the supply of ribosomal proteins is tightly controlled to avoid unnecessary energy expenditure, but the system responds quickly to the demand for increased protein production when required e.g. during aerobic growth in rich media (Bosdriesz et al., 2015). Thus, it is likely that the demand for 30S and 50S proteins would be less when the bacteria are in sub-optimal growth conditions such as reduced oxygen, even for a facultative anaerobe such as *P. multocida*. The

fumarate and nitrate reductase gene, *fnr*, a transcriptional regulator involved in anaerobic metabolism, also displayed reduced transcript levels under reduced oxygen conditions. Although reduced expression would seem contradictory to the role of *fnr* as a regulator of anaerobic metabolism, similar results were observed in a study of *fnr* in *E. coli* (Spiro and Guest, 1987). In *E. coli*, *fnr* expression is negatively autoregulated i.e., expression is reduced under anaerobic conditions due to repression by the Fnr protein (Spiro and Guest, 1987). Studies in other bacterial species have shown that Fnr acts as a regulator of a wide range of genes associated with anaerobic growth (Constantinidou et al., 2006). Many of these genes also displayed altered transcript levels in *P. multocida*, including *frdABCD, nrfABCDEF, glpABC* and *nrdDG*, indicating that altered expression of many of the *P. multocida* genes during growth in reduced oxygen conditions is likely due to regulation by Fnr. Future experiments could assess this directly by generation of a *P. multocida fnr* mutant and identification of its regulatory targets using whole transcriptome analyses.

The genes *dppA* and *artI* and the sRNA GcvB showed reduced RNA expression under low oxygen conditions. However, it has previously been shown that the production of DppA and ArtI is increased in a *gcvB* mutant (Gulliver et al., 2018). The targets of the sRNA GcvB were identified using proteomic methods as GcvB is known to act primarily at the translational level (Gulliver et al., 2018) and therefore changes in gene expression may not directly correlate with the final levels of protein. In addition, the growth of cells in reduced oxygen conditions may result in changes to gene expression that are independent of GcvB activity. Additionally, the *dppA* transcript displayed reduced expression under the low oxygen conditions used but in a previous study the same gene displayed increased expression *in vivo* during growth in chickens (Boyce et al., 2002). While low oxygen levels may be a component of the *in vivo* environment, it is clear that growth *in vitro* under low oxygen conditions and growth under *in vivo* conditions are different and it is likely that other factors contribute to the final *dppA* expression levels observed *in vivo*.

A number of the putative *P. multocida* sRNAs showed altered transcript levels under a range of conditions. *P. multocida* GcvB and HrrF displayed altered regulation under anaerobic and low iron conditions, respectively, suggesting potential roles for the sRNAs in gene regulation under these conditions. A number of other differentially expressed putative *P. multocida* sRNAs could also play a role in adapting *P. multocida* to the specific environmental conditions used in this study. Future work should involve directed mutagenesis of these predicted sRNAs in *P. multocida* and analysis of growth of the strains under the specific conditions.

In addition to examining expression data from *P. multocida* grown to different growth phases in rich medium, in low iron and reduced oxygen, RNA-Seq data derived from *hfq* (Chapter 2) and *proQ* (Gulliver, manuscript in preparation, 2020) *P. multocida* mutants was also examined. The Hfq and ProQ proteins act as RNA chaperones and are often crucial for the interaction of particular sRNA species with their target mRNAs (Smith et al., 2007; Vogel and Luisi, 2011). When transcript levels in the mutants were compared to those in the wild-type parent strain of *P. multocida*, two putative sRNAs had altered transcript levels in the *hfq* mutant and fifteen putative sRNAs had altered transcript levels in the *proQ* mutant. These results suggest that the levels of Hfq or ProQ could affect the abundance of these sRNAs by a direct mechanism. For example, sRNAs displaying decreased transcript levels in the absence of one of the chaperones may normally be protected from RNaseE degradation by the chaperone in question. It is also possible that one or both of the RNA chaperone proteins has an indirect effect on the expression of particular sRNA species via downstream regulatory effects occurring in the regulatory network. Hfq and ProQ interaction experiments such as co-immunoprecipitations could be performed to further investigate these potential chaperone-sRNA associations (see Chapter 4).

For each of the putative sRNAs identified in this study, it will be necessary to perform experimental analyses to confirm the expression and size of each and to further characterise their function. We can say with a high degree of confidence that GcvB and HrrF in *P. multocida* are sRNAs as they share high levels of identity with sRNAs of the same names in other species. To confirm this, experimental work including northern blot analysis and/or fluorescent primer extension has since been performed in our laboratory to confirm their expression (Gulliver et al., 2018) (Wright, Boyce Laboratory, unpublished). It possible that some of the short RNA-Seq peaks generated represent 5′ UTR gene regions or degradation products from larger transcripts, rather than representing coding regions for an sRNA. Only further detailed experimental characterisation can differentiate these possibilities. Experimental validation of other *P. multocida* sRNAs will be reported in Chapter 5.

Three sRNA target prediction bioinformatic tools, namely TargetRNA2, IntaRNA and RNAPredator were tested in this study using the *P. multocida* GcvB sRNA (Busch et al., 2008; Eggenhofer et al., 2011; Kery et al., 2014). Our data, generated from GcvB proteomic analyses (Gulliver et al., 2018), indicates that all three tools have a low target prediction accuracy. Other researchers have also found that sRNA prediction tools have generally low accuracy and reported that this is likely due to the complexity of factors involved in sRNA and target mRNA binding (Pain et al., 2015). They proposed that experimental approaches for target prediction, such as pulse expression of sRNAs in combination with transcriptome analysis, as well as analysis of RNA

expression data from sRNA mutants, are more reliable (Pain et al., 2015; Sharma and Vogel, 2009). To experimentally determine interactions between Hfq and sRNAs/mRNAs, co-immunoprecipitation (Co-IP) studies using recombinant *P. multocida* Hfq were conducted and the results are presented in Chapter 4. In addition to *trans*-acting sRNAs, it is likely that there are many more additional *cis*-acting *P. multocida* sRNAs that that could be identified using other experimental methods including Hfq-Co-IP experiments. Ultimately, our aim is to identify the specific sRNA species involved in the regulation of the key *P. multocida* virulence factors; one such virulence factor is capsule and its association with Hfq is described in Chapter 2 of this thesis.

This chapter reports on the identification of 38 novel putative *P. multocida* sRNAs, including those predicted to be *trans*- and *cis*-acting sRNAs. Only two of these identified sRNAs were homologues of well-studied sRNAs in other species, namely GcvB and HrrF. Global gene expression in *P. multocida* grown under low iron and reduced oxygen conditions was investigated. A range of putative sRNAs were differentially expressed and therefore may be associated with gene regulation under the particular condition used, with or without the assistance one of the RNA chaperone proteins, Hfq and ProQ. Additionally, sRNAs with altered expression in the absence of Hfq or ProQ were identified. To try and unravel the complex *P. multocida* riboregulatory network, biological experiments were performed to confirm the identity of putative sRNAs, to investigate their association with Hfq and to identify possible mRNA targets. Selected putative sRNAs were also the subject of validation experiments prior to further characterisation. These will be reported in the following chapters.

Chapter Four: Identification of *P. multocida* sRNAs using Hfq co-immunoprecipitation analyses

4.1. Introduction

Our current understanding of how *P. multocida* virulence factors are regulated is limited. Small non-coding RNA molecules (sRNAs) have recently been recognised as important regulators of bacterial gene expression, with essential roles in controlling diverse bacterial functions (Romby et al., 2006); however, little is known about *P. multocida* sRNAs. The study reported in Chapter 2 revealed an important role for the sRNA chaperone protein, Hfq, in the production of key *P. multocida* virulence factors, including hyaluronic acid (HA) capsule, LPS and filamentous hemagglutinin (Mégroz et al., 2016), indicating that Hfq-dependent sRNAs play a key role in the regulation of these factors. Therefore, understanding the specific mechanisms by which Hfq and sRNAs act to regulate these virulence factors in *P. multocida*, as well as identifying the specific mRNA-sRNA interactions involved, is of great importance.

In Chapter 3, bioinformatic analysis of whole-genome RNA transcripts, generated from cells grown under infection-relevant growth conditions (low iron and reduced oxygen growth) allowed for the identification of 38 putative *P. multocida* sRNAs. The expression of sRNAs under these conditions was investigated, as well as how the expression of the sRNAs were affected by inactivation of Hfq or ProQ. The majority of the putative sRNAs identified were *trans*-encoded, however some putative sRNA sequences overlapped 5′ or 3′ regions of gene coding sequences or were *cis*-encoded. Two of the identified *P. multocida* sRNAs, namely GcvB and HrrF, are homologues of well-studied sRNAs in other bacterial species. GcvB is a Hfq-dependent sRNA in other bacteria that regulates amino acid metabolism and transport while HrrF, identified in a number of *Haemophilus* species, is involved in iron homeostasis and in *Haemophilus influenzae* is regulated by Fur, a ferric uptake regulator involved in the control of iron stress (Santana et al., 2014; Sharma et al., 2011). However, due to the low accuracy of sRNA target prediction databases (Pain et al., 2015) the full range of mRNA targets of the putative sRNAs could not be fully defined using *in silico* methods.

In this chapter, the investigation of sRNA regulation was continued using co-immunoprecipitation (Co-IP) experimental analyses. The bioinformatic identification of sRNAs was the focus of Chapter 3 but it was predicted that Hfq Co-IP experiments would allow for experimental validation of some of these predictions and also allow for the identification of *cis*-acting sRNA species that may have been previously missed in the bioinformatic analyses, as well as the identification of other sRNAs and target mRNAs that bind to Hfq. Our data on Hfq (Chapter 2) showed that the RNA chaperone protein Hfq plays a crucial role in regulating *P. multocida* virulence factors

including hyaluronic acid (HA) capsule, LPS and filamentous hemagglutinin. However, we did not identify the specific Hfq-dependent RNA species responsible for this regulation.

To better understand the *P. multocida* Hfq riboregulatory network, multiple experimental techniques that had been used to study Hfq-RNA interactions in other bacteria were employed, including native Hfq Co-IP assays using anti-FLAG magnetic beads, UV-crosslinking and analysis of cDNA by high-throughput sequencing (UV-CRAC) and UV-crosslinking, ligation, sequencing and analysis of hybrids (UV-CLASH) (Bilusic et al., 2014; Waters et al., 2017). In addition to identifying Hfq-associated RNA molecules, the Hfq UV-CLASH technique also allows for the identification of specific RNA:RNA and sRNA:mRNA Hfq-dependent interactions. As the GcvB sRNA has been shown to be Hfq-dependent in other bacteria, (Sharma et al., 2011) it was of particular interest to investigate the association between *P. multocida* GcvB and Hfq using these Co-IP techniques.

4.2. Materials and Methods

4.2.1. Bacterial strains, plasmids and culture conditions

The bacterial strains and plasmids used in this study are listed in Table 4.1. *P. multocida* and *E. coli* strains were grown as described previously (Mégroz et al., 2016). When required, media were supplemented with 50 µg/mL spectinomycin and/or 50 µg/mL kanamycin.

4.2.2. DNA manipulations

The following DNA manipulations were performed as described previously (Mégroz et al., 2016): restriction digests, ligations and PCR amplifications, preparation of plasmid and genomic DNA, purification of DNA, quantification of DNA and Sanger sequencing of DNA. The primers used in this study are listed in Table 4.2.

4.2.3. Transformation of *E. coli*

Aliquots of previously prepared *E. coli* DH5α $RbCl_2$ competent cells (100 µL) (Hanahan, 1983) were each transformed with 8 µL of ligation reaction or with approximately 100 ng of uncut vector (positive control) or with sterile water (negative control). Transformation mixtures were incubated on ice for 25 min and then subjected to a rapid heat shock at 37°C for 2 min. A 1 mL aliquot of LB broth was then immediately added to each transformation mixture, before incubation at 37°C for 1 h, without agitation. Following recovery, 100 µL of neat transformation mixture (or serial ten-fold dilutions if required) was spread onto LB agar plates supplemented with the appropriate antibiotic(s) to select for cells containing a plasmid. Serial ten-fold dilutions of *E. coli* DH5α cells alone were plated onto LB agar plates to determine viability of cells.

4.2.4. Transformation of *P. multocida*

Electrocompetent *P. multocida* cells were prepared by culturing the required *P. multocida* strain in 10 mL of HI broth at 37°C, with shaking at 200 rpm, until mid-exponential phase of growth (approximately an OD_{600} of 0.50). To facilitate transformation, the hyaluronic acid capsule was removed by incubating cells with hyaluronidase (final concentration of 37 µg/mL) for 1 h at 37°C without agitation. Cells were pelleted by centrifugation at 5,000 x *g* for 5 min (4°C). A volume of 20 mL ice-cold 10% (v/v) glycerol was used to wash cells, which were then pelleted by centrifugation at 5,000 x *g* for 5 min (4°C). The wash and centrifugation steps were repeated, then supernatant was removed and cells allowed to resuspend in residual liquid before storage on ice until use.

For transformation of electrocompetent *P. multocida*, 1 µL of plasmid DNA (~200-300 ng) was added to a 60 µL aliquot of cells and transferred to an ice cold 1 mm gap electro-cuvette (BTX). The DNA/competent cell mixture was subjected to an electric pulse from a BTX ECM630 Pulse Generator with the following parameters: 2.0 kV, 25 µF, 600 Ω, then allowed to recover by the immediate addition of 1 mL of HI broth followed by incubation at 37°C for 1 h without agitation. Aliquots of each transformation mix (200 µL) were then plated onto HI agar with appropriate antibiotics to select for those transformants containing plasmid.

Table 4.1. Bacterial strains and plasmids used in this study.

Strain or plasmid	Relevant description	Source or reference
Strains		
E. coli strains		
DH5α	*deoR endA1 gyrA96 hsdR17*(r_K^- m_K^+) *recA1 relA1 supE44 thi-1* (*lacZYA-argFV169*) φ80*lacZ*ΔM15, F^-	Bethesda Research Laboratories
AL177	DH5α harbouring pAL99; Kan^R	(Harper et al., 2004)
AL2708	DH5α harbouring pREXY; $Spec^R$	(Gulliver et al., 2018)
AL2747	DH5α harbouring pAL1226; $Spec^R$	This study
AL2748	DH5α harbouring pAL1225; $Spec^R$	This study
AL2754	DH5α harbouring pAL1227; $Spec^R$	This study
AL2760	DH5α harbouring pAL1104; Kan^R, $Spec^R$	(Mégroz et al., 2016)
AL2763	DH5α harbouring pAL1231; $Spec^R$	This study
AL3064	DH5α harbouring pAL1337; $Spec^R$	Boyce Laboratory, unpublished
AL3093	DH5α harbouring pAL1343; $Spec^R$	This study
AL3095	DH5α harbouring pAL1345; $Spec^R$	This study
AL3419	DH5α harbouring pAL1477; $Spec^R$	This study
P. multocida strains		
VP161	Serotype A:1: Vietnamese isolate from chickens	(Wilkie et al., 2000)
X-73	Serotype A:1, wild type strain	(Heddleston and Rebers, 1972)
AL2234	VP161 *hyaD* TargeTron mutant; Kan^R	(Mégroz et al., 2016)
AL2521	VP161 *hfq* TargeTron mutant; Kan^R	(Mégroz et al., 2016)
AL2756	VP161 *hfq* mutant AL2521 containing pAL1225; Kan^R, $Spec^R$	This study
AL2757	VP161 *hfq* mutant AL2521 containing pAL1226; Kan^R, $Spec^R$	This study
AL2758	VP161 *hfq* mutant AL2521 containing pAL1227; Kan^R, $Spec^R$	This study
AL2842	X-73 *hfq*::*hfq*FLAG strain used in Hfq co-immunoprecipitation experiments	This study

AL3089	VP161 *hfq* and *hyaD* TargeTron mutant; Kan^R	This study
AL3091	AL3089 containing the pAL1345 plasmid, which encodes His-TEV-FLAG (HTF)-tagged-Hfq. Used in Hfq UV-CRAC/CLASH experiments; Kan^R, $Spec^R$	This study
AL3092	AL3089 containing the plasmid pAL1226, which encodes wild-type *hfq*. Used in Hfq UV-CRAC/CLASH experiments; Kan^R, $Spec^R$	This study
AL3100	AL3089 containing vector only (pREXY); Kan^R, $Spec^R$	This study
AL3101	VP161 *hfq* mutant (AL2521) containing plasmid encoding HTF-tagged-Hfq (pAL1345); Kan^R, $Spec^R$	This study
AL3454	VP161 *hfq* mutant (AL2521) containing the plasmid pAL1477, encoding *glmU* to overexpress GlmU; Kan^R, $Spec^R$	This study
Plasmids		
pAL99	*P. multocida/E. coli* expression plasmid, contains the *P. multocida* constitutive *tpiA* promoter upstream of multiple cloning site (MCS); Kan^R	(Harper et al., 2004)
pAL1104	*P. multocida* TargeTron mutagenesis plasmid targeted to *hfq;* Kan^R, $Spec^R$	(Mégroz et al., 2016)
pAL1225	pREXY expression plasmid containing the VP161 *hfq* gene without its native stop codon at the 3′ end, cloned fragment amplified using BAP7676 and BAP7678 and cloned into XmaI and EcoRI sites, $Spec^R$	This study
pAL1226	pREXY containing an intact copy of the WT VP161 *hfq* gene with the native stop codon at the 3′ end, cloned fragment amplified using BAP7676 and BAP7677; cloned into XmaI and EcoRI sites, $Spec^R$	This study
pAL1227	pAL1225 containing a 3xFLAG-tag sequence in-frame at the 3′ end of *hfq*, followed by a stop codon. The 3xFLAG-tag sequence (with stop codon incorporated) was generated by annealing BAP7674 and BAP7675 primers (with EcoRI compatible ends) then cloning into EcoRI site; $Spec^R$	This study
pAL1231	Modified version of the *P. multocida* TargeTron plasmid pAL1104. Kanamycin resistance gene was replaced (via digestion of the two MluI sites in the Group II intron region) with an insert encoding 3′ 3xFLAG-tagged *hfq* (amplified from pAL1227 using primers BAP7699 and BAP7700); $Spec^R$	This study

pAL1337	*P. multocida* TargeTron plasmid targeted to *hyaD* (pAL1069) with kanamycin resistance gene (*aph3)* removed from intron; $Spec^R$	Boyce Laboratory, unpublished
pAL1343	pAL1225 containing *hfq* (minus stop codon) modified with a 3′ sequence encoding an in-frame His6-TEV-tag. The His6-TEV-tag fragment was generated by annealing BAP8098 and BAP8099 primers (with EcoRI compatible ends that regenerate an EcoRI site only at the 3′ end). Fragment was then annealed to the EcoRI site in pAL1225; $Spec^R$	This study
pAL1345	pAL1343 containing *hfq* modified with a 3′ sequence encoding an in-frame His6-TEV-FLAG-tag sequence. The DNA fragment encoding FLAG (followed by a stop codon) was generated by annealing BAP8100 and BAP8101 primers, then cloned into the EcoRI site of pAL1343 located at the 3′ end of the His6-TEV region; $Spec^R$	This study
pAL1477	pREXY containing an intact copy of VP161 *glmU*, including the predicted native *glmU* terminator, and under the control of the *tpi* promoter. Amplified using BAP8655 and BAP8656; cloned into SmaI and EcoRI sites, $Spec^R$	This study
pREXY	*P. multocida/E.coli* sRNA expression vector containing the promoter region of VP161 *tpi* immediately upstream of the cloning sites; $Spec^R$	(Gulliver et al., 2018)

Table 4.2. Primers used in this study.

Primer	Sequence (5′ – 3′)	Description
BAP612 (Universal)	GTAAAACGACGGCCAGT	Forward primer located upstream of the of the *tpi* promoter and MCS in pREXY
BAP2067	GGAAGGAACAGTTTCTCTGGATTG	Forward primer located within VP161 *hyaD*, at approximately nucleotide 940 in gene
BAP2679	TTGTGTGGAATTGTGAGCGGA	Primer located downstream of the MCS in pREXY; located upstream of the *tpi* promoter and MCS on pAL99T
BAP5781	TTGAGCGGATCCTACTGCATTGGATAAACCAACA	Reverse primer located ~120 bp downstream of the 3′ end of *hyaD*; contains a BamHI site
BAP6735	GCATATAGCGCTAGCAGCAC	Reverse primer for amplification and sequencing of re-targeted region on TargeTron vector, located upstream of the 3′ end of the cloning region
BAP7153	AAAAAAGCTTATAATTATCCTTACTGAACCAGACT GTGCGCCCAGATAGGGTG	TargeTron IBS primer specific for inactivation of *hyaD*
BAP7154	CAGATTGTACAAATGTGGTGATAACAGATAAGTCC AGACTGTTAACTTACCTTTCTTTGT	TargeTron EBS1d primer specific for inactivation of *hyaD*
BAP7155	TGAACGCAAGTTTCTAATTTCGGTTTTCAGTCGAT AGAGGAAAGTGTCT	TargeTron EBS2 primer specific for inactivation of *hyaD*
BAP7389	TGAACGCAAGTTTCTAATTTCGATTCTTGTTCGATA GAGGAAAGTGTCT	TargeTron EBS2 primer specific for inactivation of *hfq*
BAP7390	AAAAAAGCTTATAATTATCCTTAACAAGCTCAGAT GTGCGCCCAGATAGGGTG	TargeTron IBS primer specific for inactivation of *hfq*
BAP7391	CAGATTGTACAAATGTGGTGATAACAGATAAGTCT CAGATTGTAACTTACCTTTCTTTGT	TargeTron EBS1d primer specific for inactivation of *hfq*
BAP7400	CAGAGGATCCTGTGTTTAGCTAGTTGTATC	Forward primer located upstream of VP161 *hfq*; contains BamHI site

BAP7401	AGATGTCGACCACTTGCAATAAATGTGTTC	Reverse primer located downstream of VP161 *hfq*; contains SalI site
BAP7545	TCATCCTGTCTCTTGATCAGAGCT	Reverse primer in kanamycin gene *aph3*, located at ~160 bp of *aph3*
BAP7674	AATTAGACTACAAAGACCATGACGGTGATTATAA AG ATCATGACATCGACTACAAGGATGACGATGACAA GT AGG	Forward primer for 3xFLAG-tag construct; contains EcoRI compatible ends, but reconstituting an EcoRI site only at 3′ end of fragment when ligated to EcoRI cut plasmid.
BAP7675	AATTCCTACTTGTCATCGTCATCCTTGTAGTCGATG TCATGATCTTTATAATCACCGTCATGGTCTTTGTAG TCT	Reverse primer for 3xFLAG-tag construct; contains EcoRI compatible ends, reconstituting an EcoRI site only at 3′ end of fragment when ligated to EcoRI cut plasmid
BAP7676	TTAGATCCCGGGGAAGGAAAGAAAAAATGGC	Forward primer for amplification of *hfq*; contains XmaI site
BAP7677	TTCAACGAATTCCTATTCTGCTTTTTCTAC	Reverse primer for amplification of *hfq* including the native stop codon (3′ end); contains EcoRI site
BAP7678	TTGAAAGAATTCTGCTTTTTCTACGTTGCT	Reverse primer for amplification of *hfq* without the native stop codon; contains EcoRI site
BAP7699	AGGATCACGCGTGAAGGAAAGAAAAAATGGCC	Forward primer for amplification of 3xFLAG-tagged-*hfq* region from pAL1227/pAL1228; contains MluI site
BAP7700	TTACACGCGTACTAGTCTACTTGTCATCGTCATCCT TGTAGTCGATG	Reverse primer for amplification of 3xFLAG-tagged-*hfq* region from pAL1227/pAL1228; contains MluI and SpeI sites
BAP7729	GGCTTACTAGTCGATAGCTAGACTGGG	Forward primer for amplification of kanamycin gene *aph3* from pAL953; contains SpeI site at 3′ and 5′ ends for cloning into pAL1231
BAP7730	TTGGGACTAGTTTGGTCGGTCATTTCG	Reverse primer for amplification of kanamycin gene *aph3* from pAL953; contains SpeI site at 3′ and 5′ ends for cloning into pAL1231

BAP7888	TGTTTGCATATTGTTTGGGAAA	Forward primer for amplification of *gcvB*; located at approximately nucleotide 55 in *gcvB*
BAP7889	GAGCGGTGTTTAACCAAAAGG	Reverse primer for amplification of *gcvB*; located at approximately nucleotide 162 in *gcvB*
BAP8098	AATTAATGGAGCACCATCACCATCACCATGATTATGATATTCCAACTACTGCTAGCGAGAATTTGTATTTTCAGG	Forward primer for His6-TEV-tag construct; contains EcoRI compatible ends, reconstituting an EcoRI site only at 3′ end of fragment when ligated to EcoRI cut plasmid
BAP8099	AATTCACCCTGAAAATACAAATTCTCGCTAGCAGTAGTTGGAATATCATAATCATGGTGATGGTGATGGTGCTCC	Reverse primer for His6-TEV-tag construct; contains EcoRI compatible ends, reconstituting an EcoRI site only at 3′ end of fragment when ligated to EcoRI cut plasmid
BAP8100	AATTGGACTACAAAGATGACGACGATAAAGACTACAAAGATGACGACGATA	Forward primer for FLAG-tag construct; contains EcoRI compatible ends, reconstituting an EcoRI site only at 3′ end of fragment when ligated to EcoRI cut plasmid
BAP8101	AATTCTCATTTATCGTCGTCATCTTTGTAGTCTTTATCGTCGTCATCTTTG	Reverse primer for FLAG-tag construct; contains EcoRI compatible ends, reconstituting an EcoRI site only at 3′ end of fragment when ligated to EcoRI cut plasmid
BAP8655	TATTCAGGATCCAAGAGAAGATTATGA	Forward primer located upstream of VP161 *glmU*; contains BamHI site
BAP8656	GGGATTGAATTCCACGTTTAACTAA	Reverse primer located downstream of VP161 *glmU;* contains *Eco*RI site and predicted *glmU* native terminator
EBS Universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	Universal TargeTron primer, located ~250 bp from the 5′ end of the intron

4.2.5. Quantitative hyaluronic acid (HA) capsule assay.

The amount of HA capsular material produced by each *P. multocida* strain was quantified as described previously (Mégroz et al., 2016).

4.2.6. Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie blue staining

To prepare whole cell lysate samples, *P. multocida* strains were cultured overnight (37°C, with shaking at 200 rpm) or to the desired optical density in 10 mL of HI broth. A volume of 400 µL of each culture was centrifuged at 13,000 x *g* for 1 min at RT. Following removal of supernatant, cells were resuspended in 40 µL of ddH_2O and 40 µL of 2x SDS-PAGE sample buffer (Appendix E). Samples were then heated to 99°C for 10 min, and pelleted at 13,000 x *g* for 10 min to pellet insoluble material. A volume of 10 µL of each protein sample was added onto and separated using an SDS-PAGE gel as described previously (Gallagher, 2006; Laemmli, 1970). Electrophoresis was performed at a constant voltage of 200 V for 45-55 min. Gels were then either stained with Coomassie Blue or used for western immunoblotting experiments.

Following SDS-PAGE separation, proteins were visualised by staining in Coomassie Blue solution (Appendix E). Excess stain was removed by washing the gel in destain solution (Appendix E). For gel preservation, gels were placed in fixing solution (Appendix E) for at least 2 h, before placement between two sheets of damp cellophane and drying at RT.

4.2.7. Western immunoblotting

Following protein separation by SDS-PAGE, proteins were transferred to a polyvinylidene fluoride (PVDF) membrane (Merck Millipore) by electroblotting at 100 V for 1 h at 4°C as described previously (Gallagher et al., 2008). Membranes were blocked by incubation in 5% (w/v) skim milk in Tris-buffered saline (TBS) (Appendix E) (pH 7.2) containing 0.05% v/v Tween 20 (TBS-T) at 4°C overnight with shaking. Primary antibody (either monoclonal or polyclonal anti-FLAG antibodies from either the Monash Antibody Technologies Facility or Sapphire Bioscience) were diluted in TBS-T (1:1000), then added to the membrane and incubated for 2 h at 37°C with shaking, followed by three washes in TBS-T (each 10 min, RT with shaking). Membranes were then probed with secondary HRP-conjugated antibody (either goat anti-mouse or goat anti-rabbit), diluted as recommended by manufacturer) for 1.5 h at RT with shaking, followed by three washes in TBS-T (each 10 min, RT with shaking). Amersham ECL Western Blotting Detection Reagents (GE Healthcare) were used to develop the membranes before visualisation (following the

manufacturer's instructions) by exposure to X-ray film (Fujifilm) and/or a Fujifilm LAS3000 imager with a chemiluminescence filter.

4.2.8. Reverse transcription PCR for cDNA synthesis

Reactions for cDNA synthesis, with or without (control) the addition of reverse transcriptase (+RT and -RT, respectively), were performed with random hexamer primers using the AffinityScript QPCR cDNA synthesis kit (Agilent Technologies) as per manufacturer's instructions.

4.2.9 Annealing of primers to produce short double-stranded DNA

To clone short regions of DNA, double-stranded DNA fragments were generated using two complementary primers. Primers were annealed by mixing at equimolar concentrations in annealing buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA and 50 mM NaCl), then incubating for 5 min at 100°C before gradually cooling to room temperature.

4.2.10. Hfq co-immunoprecipitation (Co-IP)

Hfq Co-IP was performed as per the protocol described by Bilusic (Bilusic et al., 2014), with minor modifications. Briefly, the *P. multocida hfq* mutant expressing 3xFLAG-tagged Hfq (*hfq*::*hfq*FLAG) and wild-type X-73 strains were grown to mid-exponential growth phase and 30mL of cells harvested from each culture by centrifugation at 8000 x *g* for 10 min at 4°C. Following removal of supernatant, cell pellets were washed in 4 mL TM buffer (10 mM Tris-HCl pH 7.0, 10mM $MgCl_2$) before cells were partially lysed by a single freeze-thaw cycle (-80°C and RT). Partially lysed cells were resuspended in 2 mL ice-cold lysis buffer (50 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 250 mM NaCl), supplemented with approximately 80 U of protector RNase inhibitor (Roche) and DNase-treated with approximately 10 U of DNase I (Qiagen) for 15 min at RT. Cell samples were then sonicated to complete cell lysis and then centrifuged at 5,400 x *g* for 10 min at 4°C before the addition of protease inhibitor cocktail (Sigma). For each sample, 15 µL of anti-FLAG M2 magnetic beads (mouse monoclonal, Sigma) were washed twice with 10 x the packed-gel volume of TBS (50 mM Tris-HCl pH 7.4, 150 mM NaCl). Lysates were then added to the washed anti-FLAG M2 magnetic beads and the mixture was then incubated overnight at 4°C with rotation to allow Hfq-RNA-complexes to bind to the beads. Magnetic beads were washed 3 times with TBS (20 x the packed-gel volume) and resuspended in a final volume of 400 µL of TBS. Each sample was then mixed with 500 µL TRIzol reagent (Ambion) before bound RNA from the beads was isolated.

4.2.11. Hfq Co-IP: RNA extraction and purification, library preparation, RNA sequencing and identification of Hfq-associated RNA species

RNA extraction was performed as described in Materials and Methods of Chapter 3 (Section 3.2.2). Quantification, and analysis of RNA samples for DNA contamination, was performed as described previously (Mégroz et al., 2016); the RNA samples were not depleted of rRNA. For the RNA samples extracted from the Hfq Co-IP, cDNA libraries were constructed using the SureSelect Strand-Specific RNA Library Prep for Illumina Multiplexed Sequencing kit (Agilent) according to the manufacturer's instructions. Validation, normalisation and pooling of the libraries was performed by Micromon services (Monash University). All libraries were then combined and sequenced on a single lane of a NextSeq instrument (Illumina).

In initial analyses, mapping of the RNA-Seq reads to the draft X-73 genome was performed using CLC Genomics Workbench v 7.0 (CLC Bio). For the *hfq*:*hfq*FLAG samples, the average number of reads that mapped to the *P. multocida* X-73 genome across all replicates was 5,075,220.0 (an average of 98.65%), giving high coverage. For the untagged-*hfq* control samples, the average number of reads that mapped to the *P. multocida* X-73 genome across all replicates was 4,671,146.33 (an average of 98.15%). Mapped RNA-Seq datasets from the Hfq Co-IP were then visually inspected using Artemis (Sanger; http://www.sanger.ac.uk/resources/software/artemis/). Statistical analysis was performed using the limma/voom modelling software (Law et al., 2014) via the Degust visualisation software, available through the Bioinformatics Platform at Monash University. Transcripts were considered to be associated with Hfq if they displayed an increased number of sequencing reads across the triplicate *hfq::hfq*FLAG replicates, relative to the untagged replicates, at a false discovery rate (FDR) of <0.05. Predicted functions for gene products were assigned by mapping the differentially expressed X-73 genes to their closest orthologue in the fully annotated *P. multocida* Pm70 reference genome (May et al., 2001). Predicted functions of gene products were also determined by interrogating the SwissProt database (excluding proteins belonging to *Pasteurellaceae*); only proteins that showed high similarity with known proteins/domains were assigned a predicted function (i.e. an alignment score of ≥ 200, graphical output = red bar). As multiple genome entries exist for strain X-73 genome online the NCBI genome accession: NZ_CM001580.1 was chosen for this study.

Analysis of the Hfq Co-IP RNA-Seq data was also performed by the Monash Bioinformatics Platform. Burrows-Wheeler Aligner (BWA-MEM) (Li, 2013) with Model-based analysis of ChIP-Seq (MACS2) (Zhang et al., 2008), was used to map the RNA-Seq reads to the *P. multocida* X-73 reference genome. For the *hfq*:*hfq*FLAG samples, the average number of reads that mapped to the

P. multocida X-73 genome across all replicates was 5,103,857.33 (an average of 96.54%), giving high coverage. For the untagged-*hfq* control samples, the average number of reads that mapped to the *P. multocida* X-73 genome across all replicates was 4,701,663.67 (an average of 97.27%). Following MACS2 (Zhang et al., 2008) analysis to call peaks, IGV (Integrative Genomics Viewer) (Robinson et al., 2011) was used to identify peak locations enriched by Hfq Co-IP.

4.2.12. Hfq UV-Crosslinking, ligation, sequencing and analysis of hybrids (Hfq UV-CRAC/CLASH) library preparation

Hfq UV-CRAC/CLASH was performed as described previously (Waters et al., 2017), with the following modifications. *P. multocida* VP161 strains AL3091, harbouring pAL1345, encoding His-TEV-FLAG (HTF)-tagged-Hfq, and AL3092, harbouring pAL1226, encoding untagged Hfq, were grown to OD_{600} = 0.8 in HI media prior to UV-crosslinking. The HTF-tagged-Hfq and untagged-Hfq strains (AL2091 and AL3042, respectively) were grown in biological duplicate for the experiments.

4.2.13. Analysis of Hfq UV-CRAC/CLASH binding and RNA hybrids

Analysis of Hfq-associated RNA molecules was performed as described previously (Holmqvist et al., 2016). Hfq UV-CLASH hybrids were also identified as described previously (Waters et al., 2017). The Multiple Em for Motif Elicitation (MEME) tool (Bailey et al., 2015) was used to examine conserved binding sequence motifs in Hfq-associated RNA sequences.

4.2.14. Generation of a *P. multocida glmU* overexpression strain

To construct the *glmU* overexpression plasmid the wild-type *glmU* sequence was amplified from *P. multocida* VP161 genomic DNA using BAP8655 and BAP8656 primers (Table 4.2), digested with SmaI and EcoRI and ligated with similarly-digested pREXY. The ligations were used to transform *E. coli* DH5α and a selected number of transformants analysed by colony PCR using pREXY-specific primers, BAP612 and BAP2679, that flanked the region of insertion (Table 4.2). Those positive for a recombinant DNA insert were separately sequenced using pREXY-specific primers that read into the cloned region, BAP612 and BAP2679 (Table 4.2). A plasmid with verified wild-type *glmU* sequence was designated pAL1477. The *hfq* mutant AL2521 was then transformed with pAL1477 to generate the *P. multocida* strain AL3454.

4.3. Results

4.3.1 Identification of Hfq-associated sRNAs and mRNAs using native co-immunoprecipitation

To experimentally identify *P. multocida* sRNAs and mRNAs that interacted with Hfq, co-immunoprecipitation (Co-IP) experiments were performed. For these experiments we chose to use recombinant derivatives of the highly virulent *P. multocida* strain X-73 (closely related to *P. multocida* strains VP161 and Pm70) as only infections with X-73 yield very high numbers of bacteria in the blood of the chicken host, providing an opportunity for future *in vivo hfq*-associated experiments (Boyce and Harper, pers. comm.). Several strains were constructed (see below) to allow Co-IP experiments to be performed. These included a *P. multocida* X-73 strain that expressed a chromosomally encoded 3xFLAG-tagged Hfq protein and X-73 wild-type that expressed an untagged version of Hfq as a control. The Hfq Co-IP method utilised anti-FLAG magnetic beads to co-immunoprecipitate Hfq and its bound target RNAs, followed by deep sequencing (Illumina NextSeq) to identify all nucleic acids that interacted with Hfq (Figure 4.1).

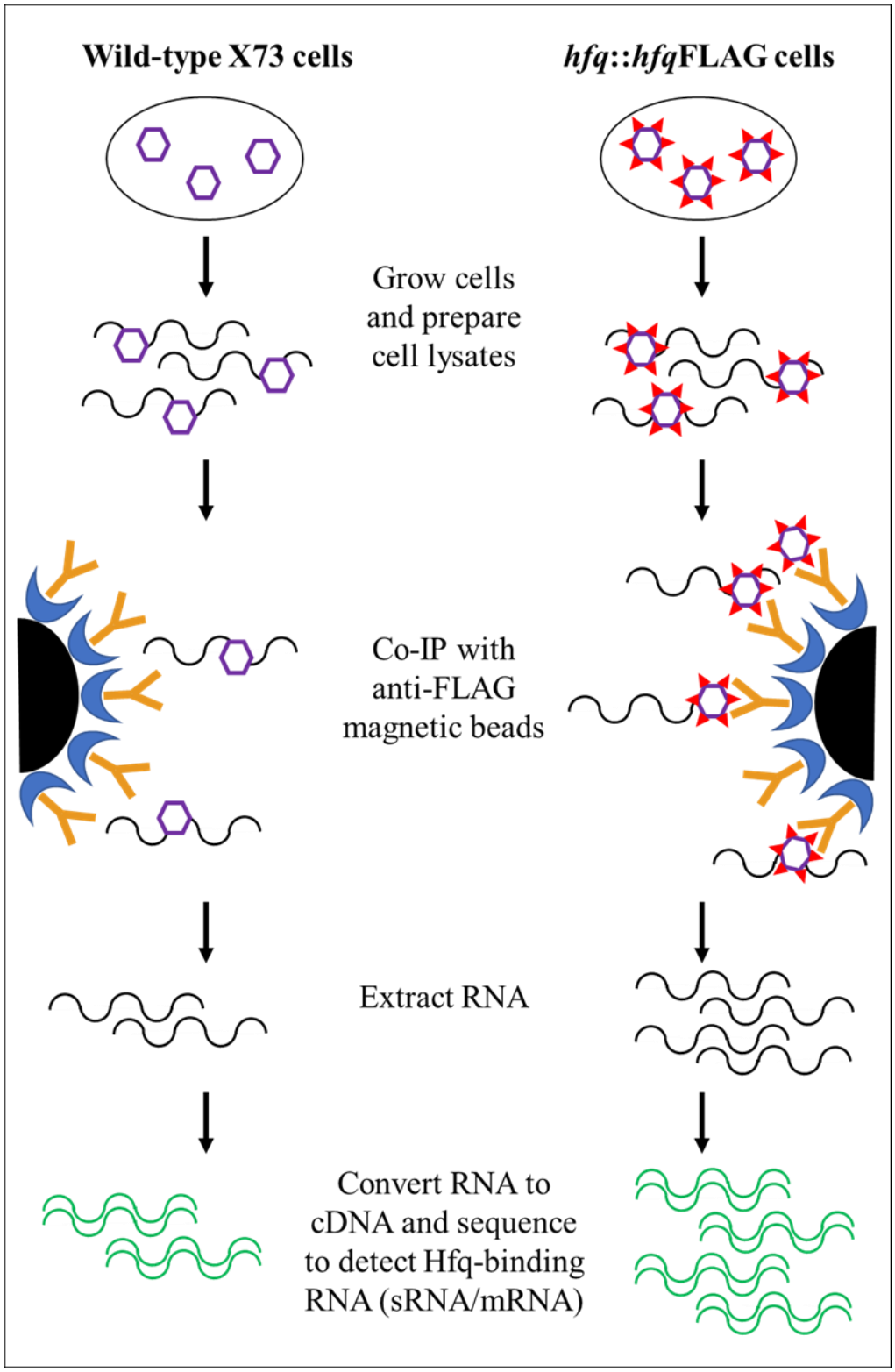


Figure 4.1. Strategy for identification of Hfq targets using Co-IP with anti-FLAG magnetic beads (closely follows protocol from (Bilusic et al., 2014)). A *P. multocida* X-73 strain expressing 3xFLAG-tagged Hfq (*hfq*::*hfq*FLAG) was used along with wild-type X-73 expressing untagged Hfq as a control. The purple hexagon represents the Hfq protein and red triangles represent FLAG epitope. Yellow 'Y' symbols and blue crescents represent anti-FLAG antibodies and receptors on the surface of the magnetic beads.

4.3.1.1. Generation of *P. multocida* strains expressing 3xFLAG-tagged Hfq, in *trans* and on the chromosome

In order to proceed with Co-IP experiments, it was important to first ensure that the Hfq protein could function when fused with the FLAG amino acids. In addition, it was anticipated that the *hfq* constructs would be used in a number of different *P. multocida* strains in future studies, so it was important that there was a high degree of identity at the amino acid level with Hfq proteins encoded by other strains. Using BLAST at NCBI, the Hfq protein was found to be 100% identical at the amino acid level across the 25 available genomes examined, including the genomes for strains VP161 and X-73 used in this study. A C-terminal 3xFLAG fusion strategy was chosen as this has been previously used in Hfq Co-IP assays for sRNA identification in other bacteria (Bilusic et al., 2014; Pfeiffer et al., 2007). To check that the C-terminal positioning of the 3xFLAG fusion had no effect on Hfq activity, the ability of the FLAG-tagged Hfq to complement a *hfq* mutant for hyaluronic acid (HA) capsule production was assessed; our previous study had shown that Hfq was required for full capsule production in *P. multocida* (Mégroz et al., 2016) (Chapter 2). To begin the analysis, a plasmid encoding a 3xFLAG-tagged *P. multocida* Hfq protein was constructed and tested for its ability to complement a *P. multocida hfq* mutant in *trans*. Once activity of the 3xFLAG-tagged Hfq was confirmed, a TargeTron mutagenesis plasmid was constructed and used to replace the wild-type *hfq* with the tagged *hfq* on the *P. multocida* genome.

Construction of plasmids encoding 3xFLAG-tagged Hfq

To construct the 3xFLAG-tagged Hfq expression plasmid, pAL1225 (Figure 4.2), the wild-type *hfq* gene, excluding the stop codon, was PCR-amplified from *P. multocida* strain VP161 genomic DNA using the primers, BAP7676 and BAP7678, that also contained XmaI and EcoRI restriction sites, respectively (Figure 4.2, Table 4.2). After digestion with XmaI and EcoRI, the DNA fragment encoding *hfq* was cloned into similarly-digested vector, pREXY (Table 4.1) such that the *P. multocida* constitutive promoter *tpiA* on the vector would drive the transcription of *hfq*. The ligated products were then used to transform *E. coli* strain DH5α and transformants containing the correct plasmid were identified by colony PCR using the pREXY-specific primer, BAP2679 with the *hfq*-specific primer BAP7676 (Figure 4.2, Table 4.2). DNA sequencing was used to confirm the fidelity of the *hfq* sequence in selected plasmids and one correct recombinant plasmid was selected and designated pAL1225 (Figure 4.2, Table 4.1). A control plasmid encoding the native *hfq* gene cloned into pREXY was also constructed. This plasmid, pAL1226 (Figure 4.2, Table 4.1), was constructed by PCR-amplifying the wild-type *hfq* gene, including the native stop codon, with BAP7676 and BAP7677 (Table 4.2), then cloning the XmaI and EcoRI-digested PCR product into similarly-digested pREXY.

To construct the plasmid that contained the 3xFLAG sequence at the 3′ end of *hfq* (pAL1227, Figure 4.2), the complementary primers BAP7674, encoding the 3xFLAG-tag amino acid sequence (DYKDHDGDYKDHDIDYKDDDDK) and BAP7675 (Table 4.2) were annealed together. The double stranded DNA fragment generated contained EcoRI compatible ends, but only the 3′ end reconstituted an EcoRI restriction site once ligated to EcoRI-digested pAL1225 (Figure 4.2). The ligated products were then used to transform *E. coli* strain DH5α and spectinomycin resistant transformants containing the correct plasmid were identified by colony PCR using pREXY-specific primers, BAP612 and BAP2679, that flanked the region of insertion (Figure 4.2, Table 4.2). DNA sequencing was used to confirm the fidelity and orientation of the 3xFLAG-tag. One correct recombinant plasmid was selected for further study and designated pAL1227 (Figure 4.2, Table 4.1).

To determine if Hfq with the addition of a C-terminal 3xFLAG-tag was still biologically active, hyaluronic acid (HA) capsule assays were performed on *P. multocida* stains harbouring the above constructs. The *P. multocida hfq* mutant (AL2521) was transformed separately with plasmid pAL1225, encoding *hfq* without a stop codon, pAL1226, encoding *hfq* with a stop codon and pAL1227, encoding *hfq* with the addition of the 3′ 3xFLAG-tag sequence, generating strains AL2756, AL2757 and AL2758, respectively (Table 4.1). As expected, the *hfq* mutant produced very low levels of HA capsule (Figure 4.3). In contrast the *hfq* mutant provided with the native *hfq* gene, with or without the stop codon (pAL1226 and pAL1225, respectively), and the *hfq* mutant provided with the 3xFLAG-tagged Hfq (encoded on pAL1227), all produced HA capsule at levels comparable to that produced by the *P. multocida* parent VP161 (Figure 4.3). Thus, Hfq activity appeared unaffected by the addition of a 3xFLAG tag at the C-terminal end of the protein.

Western blot analysis was then performed on whole cell lysates of the above strains to determine if the 3xFLAG-tagged Hfq (predicted to be 13.8 kDa in size) could be immunologically detected. Following immunoblotting with anti-FLAG antibodies, one unique protein band slightly larger than the predicted size (approximately 18 kDa) (Figure 4.4) was present in the lane containing whole cell lysate of the strain expressing the C-terminal FLAG-tagged Hfq (from plasmid pAL1227).

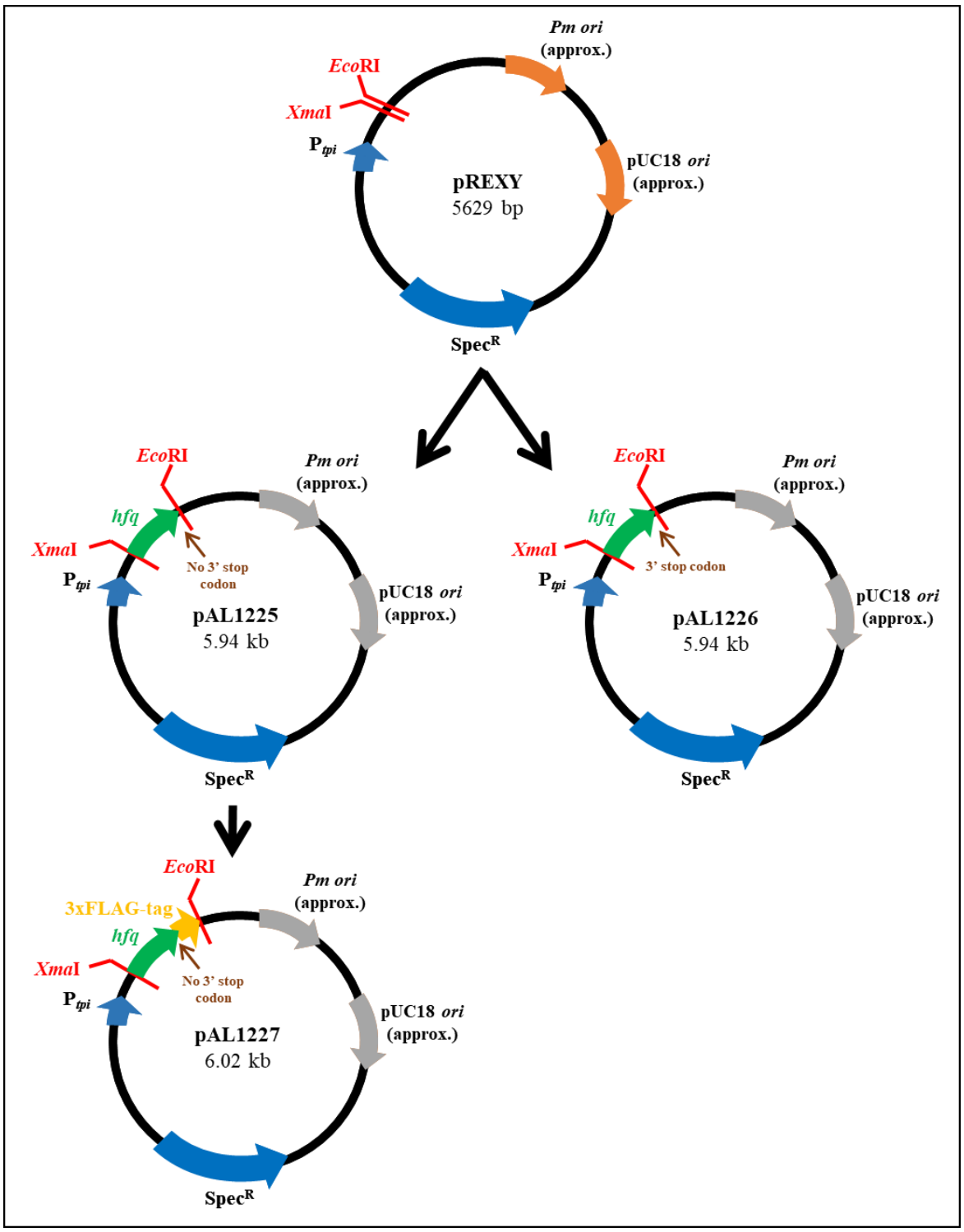


Figure 4.2. Schematic diagram showing the construction of Hfq expression plasmids used in this study. The *hfq* gene was PCR-amplified from wild-type *P. multocida* (strain VP161) with or without the stop codon. The PCR product encoding *hfq* was digested with XmaI and EcoRI then ligated to XmaI- and EcoRI-digested pREXY to generate pAL1225 (*hfq* without downstream stop codon) and pAL1226 (*hfq* with native stop codon). To generate pAL1227 encoding 3xFLAG-tagged-*hfq*, the 3xFLAG-tag sequence was digested with EcoRI and then ligated into EcoRI-digested pAL1225, in-frame at the 3′ end of the *hfq* sequence. Abbreviations: P_{tpi}, *P. multocida tpi* promoter; pUC18 *ori*, pUC18-derived *E. coli* origin of replication; Pm *ori*, *P. multocida* origin of replication; $Spec^R$ (*aadA*), spectinomycin resistance gene.

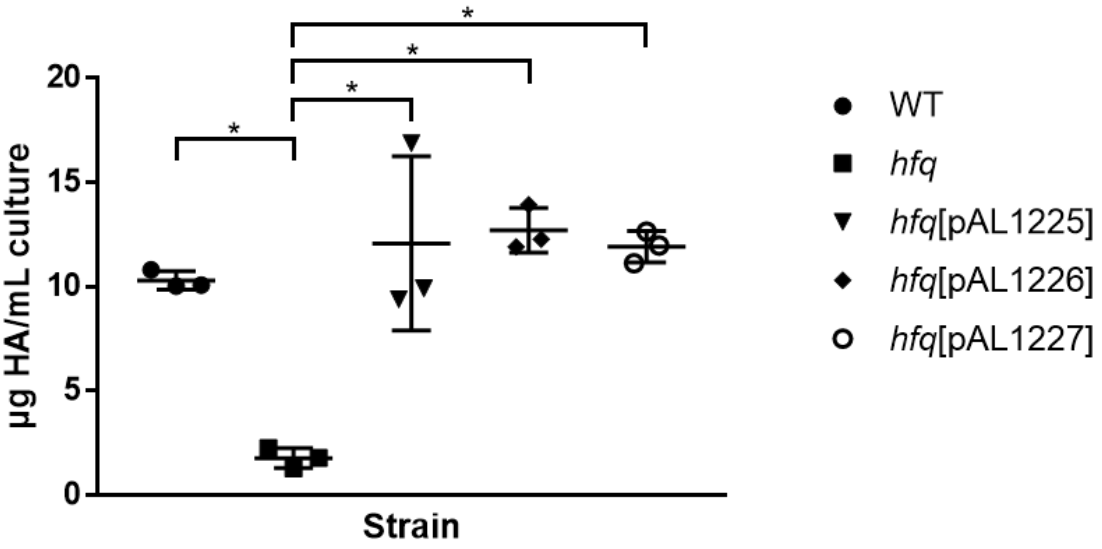


Figure 4.3. Hyaluronic acid (HA) polysaccharide capsule produced by *P. multocida* strains. Amounts of HA produced at the mid-exponential growth phase by wild-type *P. multocida* strain VP161 (WT), the *hfq* mutant, and the *hfq* mutant expressing *hfq* without downstream stop codon (*hfq*[pAL1225]), expressing *hfq* with native stop codon (*hfq*[pAL1226],) or expressing *hfq* with 3xFLAG tag (*hfq*[pAL1227],). The average amount of HA for each strain is indicated by the long horizontal bars, ± standard deviations of the means. *, *P* value of < 0.05 as determined by an unpaired *t* test.

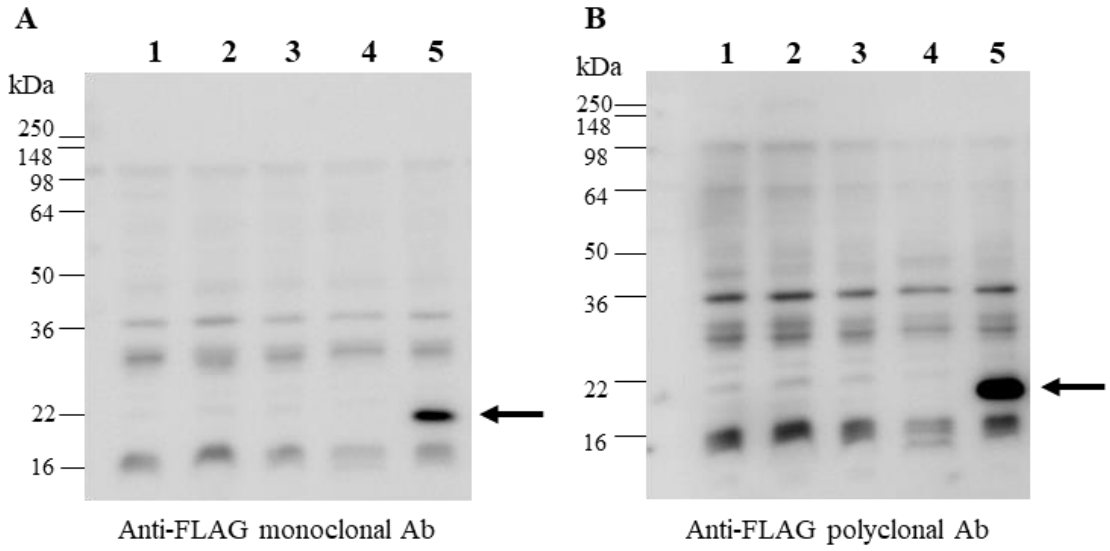


Figure 4.4. A and B. Western blot analysis of *P. multocida* strains using an anti-FLAG monoclonal antibody (A) or an anti-FLAG polyclonal antibody (B). Lane 1, wild-type VP161; lane 2, *hfq* mutant AL2521; lane 3, AL2521 harbouring pAL1225 (pREXY+*hfq* without stop codon); lane 4, AL2521 harbouring pAL1226 (pREXY+*hfq* with stop codon); lane 5, AL2521 harbouring pAL1227 (pREXY+3xFLAG-tagged-*hfq*). The protein band (~18 kDa) representing the 3xFLAG-tagged Hfq protein is indicated by the arrow to the right of each gel. Relative position of each protein marker is indicated to the left of each gel.

Generation of a *P. multocida* strain expressing chromosomally encoded 3xFLAG-tagged Hfq (*hfq*::*hfq*FLAG)

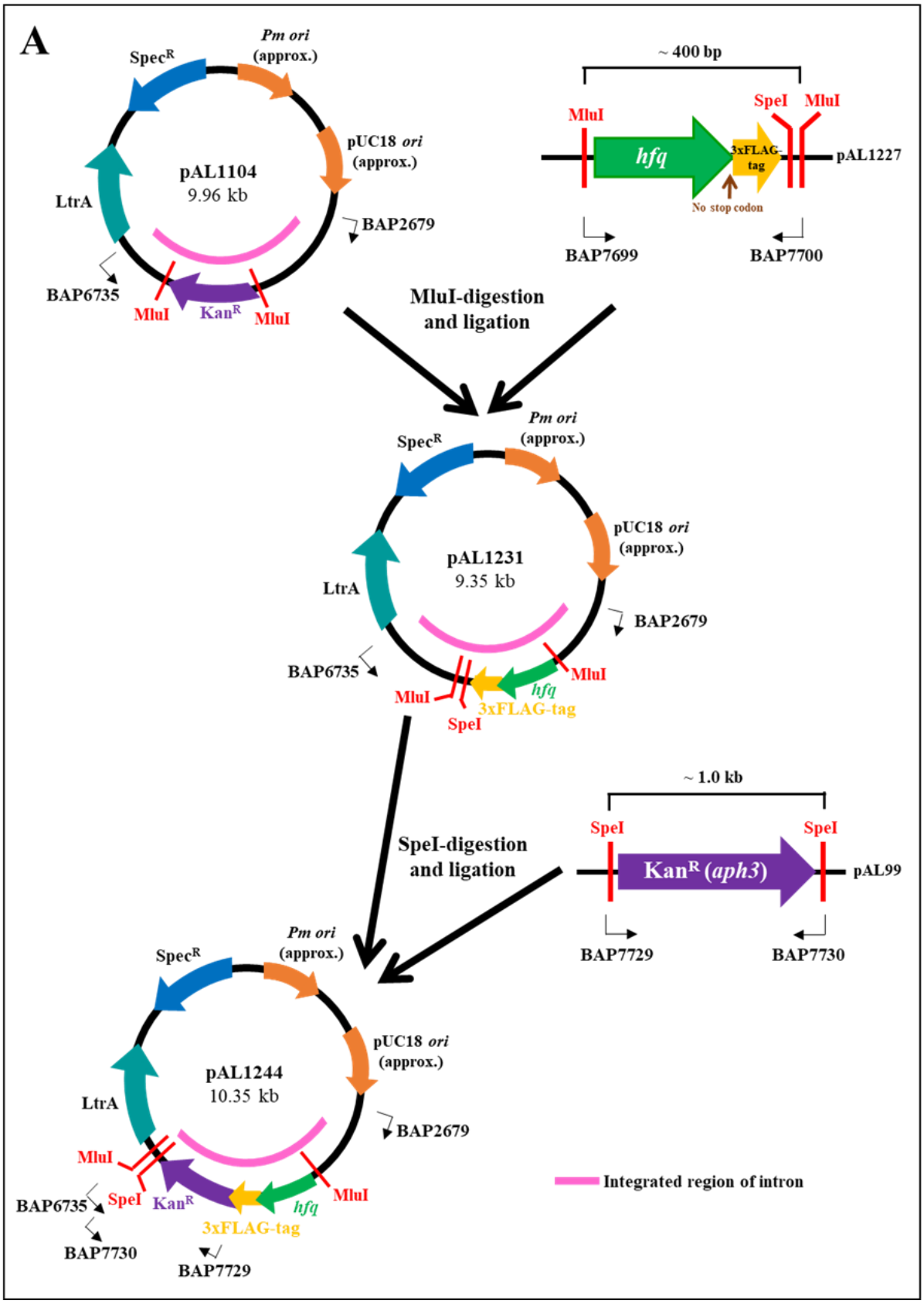
In order to conduct Hfq Co-IP assays, the construction of a *P. multocida* strain expressing chromosomally-encoded C-terminal 3xFLAG-tagged Hfq was required, ensuring stable expression of the recombinant protein without selection during the duration of the experiment. Importantly, the wild type copy of *hfq* was simultaneously inactivated by using a specially designed TargeTron plasmid that contained a *hfq*-targeted group II intron harbouring a gene encoding a 3xFLAG-tagged Hfq.

To construct the *hfq*-targeted TargeTron plasmid (Figure 4.5A), the *aph3* gene (encoding kanamycin resistance) located in the intron, was replaced with a recombinant gene encoding 3xFLAG-tagged Hfq, amplified from the plasmid pAL1227 using primers BAP7699 and BAP7700 (Figure 4.2, Table 4.2). Both primers contained an MluI site plus BAP7700 contained a unique SpeI site that would allow for the reintroduction of the *aph3* gene (see below). The plasmid pAL1104, a previously constructed, *hfq*-specific, TargeTron plasmid (Figure 4.5A, Table 4.1) (Mégroz et al., 2016) was digested with MluI to remove the *aph3* gene. The linear plasmid (now lacking *aph3*) was then ligated to MluI-digested PCR product. The ligated products were then used to transform *E. coli* strain DH5α. *E. coli* transformants resistant to spectinomycin (resistance gene present on vector backbone) but sensitive to kanamycin (due to the replacement of *aph3* with the recombinant gene encoding 3xFLAG-tagged Hfq) were chosen for colony PCR screening using BAP2679 (vector specific) and BAP6735 (intron-specific, reverse primer downstream of second MluI site) (Table 4.2). Plasmid was isolated from a selected number of transformants with the correct PCR profile and separately sequenced using BAP7699 and BAP7700 (Table 4.2). Correct nucleotide sequence of the recombinant gene was confirmed in several plasmids; one plasmid, designated pAL1231, was used for further studies (Table 4.1).

For ease of strain selection during TargeTron (Sigma-Aldrich) mutagenesis, the *aph3* gene (encoding kanamycin resistance) was reintroduced into the intron region of pAL1231 via the unique SpeI site, located immediately downstream of the 3xFLAG-tagged-*hfq* (Figure 4.5A). The *aph3* gene was amplified from the *P. multocida* expression vector pAL99 using BAP7729 and BAP7730 (Table 4.1. and 4.2), each containing a SpeI restriction site. The SpeI-digested PCR product was then ligated to SpeI-digested pAL1231 (Table 4.1). Ligated products were used to transform *E. coli* strain DH5α and kanamycin-resistant transformants containing the correct plasmid, with *aph3* in the same orientation relative to *hfq*, were identified using two separate colony PCR reactions; one reaction using the vector primer BAP2679 with BAP7730 (*aph3* reverse), and one using one BAP7729 (*aph3* forward) together with the vector primer BAP6735

(Figure 4.5A, Table 4.2). To ensure that the *aph3* fragment had inserted into the plasmid only once, the plasmid was digested with MluI. A band of the expected size (containing the *aph3* and the gene encoding 3xFLAG-tagged Hfq) was obtained (data not shown). One recombinant plasmid with *aph3* inserted in the correct orientation and downstream of the 3xFLAG-tagged-*hfq* region was selected for TargeTron mutagenesis and designated pAL1244 (Figure 4.5A, Table 4.1).

The replacement of the genomic copy of *hfq* in *P. multocida* strain X-73 with a 3xFLAG-tagged-*hfq* was then performed using the TargeTron plasmid constructed above (pAL1244) (Figure 4.5B). Following electroporation of *P. multocida* X-73 with pAL1244, kanamycin-resistant transformants were recovered on solid agar and passaged twice in LB media with no antibiotic to cure them of free-replicating plasmid as described previously (Harper et al., 2013). Putative *hfq* insertional mutants (with *aph3* on the genome and free of the plasmid pAL1244) were identified by a $Kan^R/Spec^S$ phenotype. Putative mutants (*hfq*::*hfq*FLAG strains) were initially screened by PCR using *hfq*-specific primers BAP7400/BAP7401 that flanked the recombination region (Figure 4.5C; see Figure S2 in Appendix D). For those transformants that generated a PCR product approximately 2.4 kb larger than the wild-type product, a second PCR was performed using the TargeTron-specific primer, EBS universal, together with the *hfq*-specific primer BAP7400 (Figure 4.5C, Table 4.2; see Figure S2 in Appendix D). This PCR would generate a product only if the intron had inserted in the correct region of the genome. Mutants with the correct PCR profile were then sequenced using genomic DNA isolated from each mutant as template with the intron-specific EBS universal primer (Figure 4.5C, Table 4.2) which reads out of the intron and into the genomic region. The *hfq*::*hfq*FLAG strains with the correct sequence profile (data not shown) then underwent additional sequencing analysis to confirm that the sequence of the 3xFLAG-tagged-*hfq* gene was correct; a PCR product was generated using genomic DNA from the *hfq*::*hfq*FLAG strains as template with the primers BAP7400 and BAP7401 that flanked the *hfq* region (Table 4.2). Separate sequencing reactions of this PCR product were then performed using primers BAP7400, BAP7401, BAP7700, BAP7545 (Figure 4.5C, Table 4.2). The aligned sequences were 100% identical to the expected sequence, representing the *hfq* region replaced with a TargeTron intron containing the 3xFLAG-tagged-*hfq* and *aph3* (Figure 4.5C). Southern hybridisation was also performed (data not shown) using a TargeTron intron-specific digoxigenin (DIG)-labelled probe and confirmed that the *hfq*::*hfq*FLAG strains contained only a single TargeTron insertion. One *P. multocida* X-73 *hfq*::*hfq*FLAG strain with the correct profile was selected for further study and designated AL2842 (Table 4.1)


A
Spec^R
Pm ori
(approx.)
pUC18 ori
(approx.)
pAL1104
9.96 kb
LtrA
BAP2679
BAP6735
MluI
Kan^R
MluI
~ 400 bp
MluI
SpeI
MluI
hfq
3xFLAG-tag
pAL1227
No stop codon
BAP7699
BAP7700
MluI-digestion
and ligation
Spec^R
Pm ori
(approx.)
pUC18 ori
(approx.)
pAL1231
9.35 kb
LtrA
BAP2679
BAP6735
MluI
hfq
3xFLAG-tag
SpeI
~ 1.0 kb
SpeI
SpeI
Kan^R (aph3)
pAL99
SpeI-digestion
and ligation
BAP7729
BAP7730
Spec^R
Pm ori
(approx.)
pUC18 ori
(approx.)
pAL1244
10.35 kb
LtrA
BAP2679
MluI
SpeI
Kan^R
hfq
MluI
3xFLAG-tag
BAP6735
BAP7730
BAP7729
Integrated region of intron

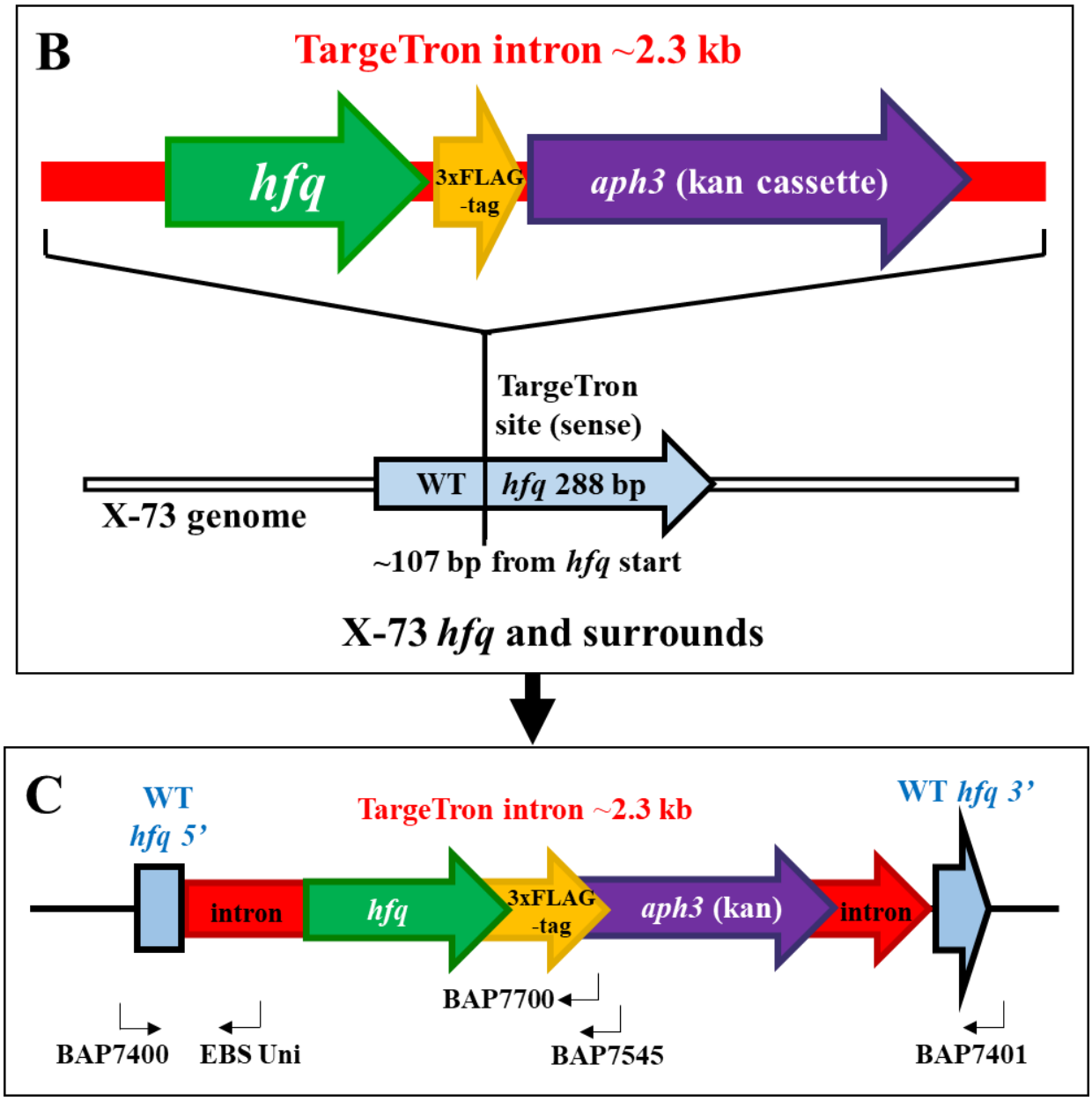


Figure 4.5. Generation of a *P. multocida* strain expressing 3xFLAG-tagged Hfq (*hfq*::*hfq*FLAG) using TargeTron mutagenesis. A. (Previous page). The 3xFLAG-tagged-*hfq* region from pAL1227 was cloned into the MluI sites of the *hfq*-specific TargeTron plasmid pAL1104, to generate pAL1231. The kanamycin gene (*aph3*) was then re-introduced downstream of the 3xFLAG-tagged-*hfq* region to generate pAL1244. The region of the intron integrated into the genome during mutagenesis is indicated by the pink bar. P_{tpi}, *P. multocida tpi* promoter; pUC18 *ori*, pUC18-derived *E. coli* origin of replication; Pm *ori*, *P. multocida* origin of replication; $Spec^R$ (*aadA*), spectinomycin resistance gene. **B.** Following construction of the TargeTron plasmid pAL1244, the intron was introduced into the *hfq* region of the *P. multocida* genome. **C.** Schematic representation of the genomic region following mutagenesis. Primers used for PCR screening and sequencing of putative *hfq*::*hfq*FLAG strains are indicated by right-angled arrows.

4.3.1.2. Phenotypic analysis of *hfq*::*hfq*FLAG strains

Three independent X-73 *hfq*::*hfq*FLAG transformants (X1, X5 and X9) were selected to determine the level of expression of the 3xFLAG-tagged Hfq protein using monoclonal and polyclonal anti-FLAG antibodies (Figure 4.6). The wild-type parent strain X-73 and the *P. multocida hfq* mutant (AL2521) were used as negative control strains. The *hfq* mutant containing the plasmid pAL1227, encoding 3xFLAG-tagged Hfq (AL2758), was used as a positive control as it had previously been shown to express an antibody-positive 3xFLAG-tagged Hfq protein. Western immunoblotting using either the monoclonal anti-FLAG antibody or the polyclonal antisera detected a unique protein band (~18 kDa) in all lanes containing whole cell lysates from transformants encoding the 3xFLAG-tagged Hfq protein (X-73 *hfq*::*hfq*FLAG X1, X5 and X9 and AL2758) confirming expression of C-terminal FLAG-tagged Hfq (Figure 4.6). One of the X-73 transformants that showed expression of the 3xFLAG-tagged Hfq protein, X5, was selected for further study and designated AL2842.

To determine if 3xFLAG-tagged Hfq protein was active in the *hfq*::*hfq*FLAG strain, the amount of Hfq-dependent capsule was measured using the hyaluronic acid assay. Capsule production in the *hfq*::*hfq*FLAG strain AL2842 was 8.3-fold higher than in the *hfq* mutant strain and 60-70% of the level of capsule produced by the wild-type parent strain, X-73 (Figure 4.7). Therefore, this strain produced active 3xFLAG-tagged Hfq with only slightly reduced activity compared to the wild-type levels of Hfq as determined by the amount of capsule produced. It was decided to use this strain for the Co-IP experiments to identify the *P. multocida* sRNAs and mRNAs that interact with Hfq.

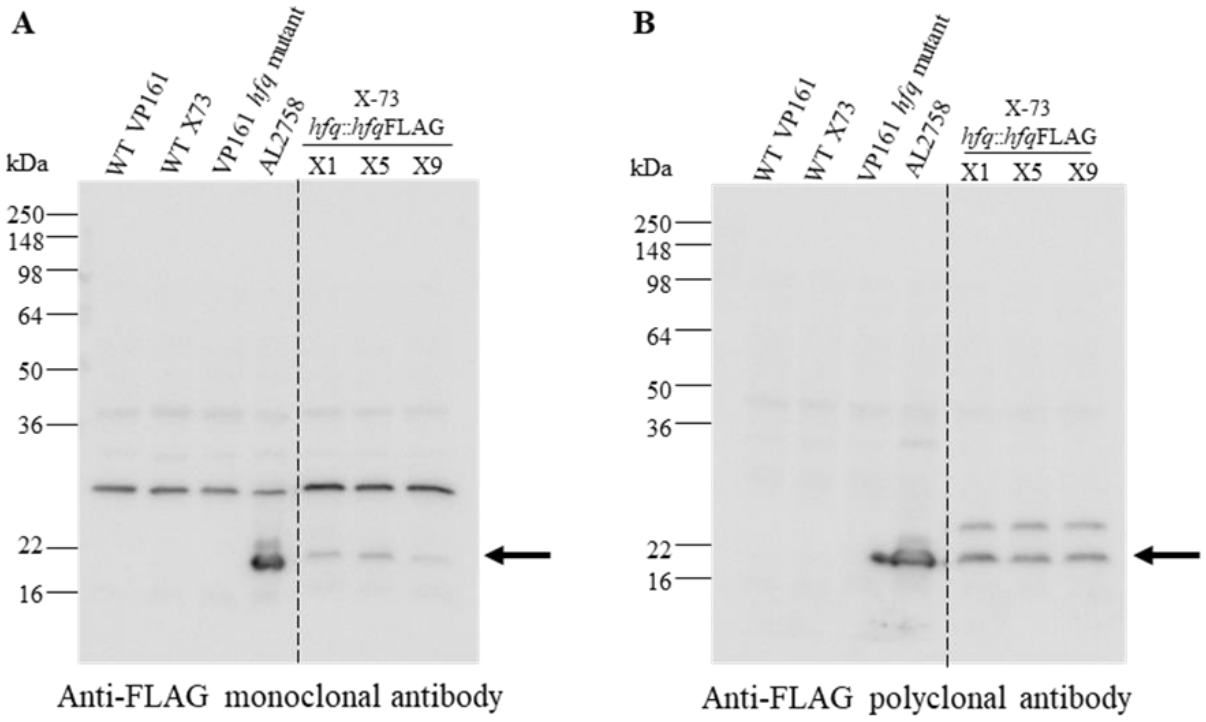


Figure 4.6. Western blot analysis of whole cell lysates derived from *P. multocida* X-73 *hfq*::*hfq*FLAG transformants X1, X5 (AL2842) and X9 using monoclonal (panel A) and polyclonal (panel B) anti-FLAG antibodies. The protein band representing the 3xFLAG-tagged Hfq protein (~18 kDa) is indicated by the arrow to the right of each gel. Sizes of protein markers are indicated to the left of each gel.

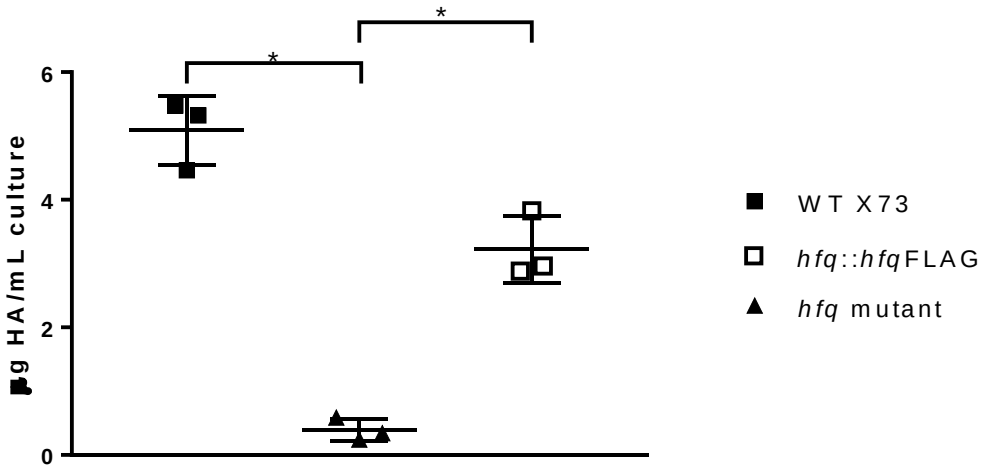


Figure 4.7. Hyaluronic acid (HA) polysaccharide capsule produced during mid-exponential growth phase by wild-type strain VP161, wild-type strain X-73, *hfq*::*hfq*FLAG X-73 strain (AL2842) and the *hfq* mutant. The average amount of HA for each strain is indicated by the horizontal bars ± standard deviations of the means. *, *P* value of < 0.05 as determined by an unpaired *t* test.

4.3.1.3. Hfq Co-IP and extraction of Hfq-associated RNA

Hfq Co-IP experiments were performed using anti-FLAG magnetic beads (Figure 4.1). For each experiment the *hfq*::*hfq*FLAG strain AL2842 (expressing 3xFLAG-tagged Hfq) (Table 4.1) was analysed alongside the wild-type X-73 strain (expressing untagged Hfq) as a negative control. Three independent experiments were performed. Following Hfq Co-IP, RNA bound to the magnetic beads was isolated using TRIzol and phenol/chloroform extraction. RNA samples were visualised by gel electrophoresis (Figure 4.8) and quantified using Qubit fluorometric quantitation (Invitrogen). One sample derived from wild-type X-73 did not have detectable levels of recovered RNA (< 50 ng/mL), while the remainder of the samples had RNA yields that ranged between 2.8-5.3 µg. The level of DNA contamination was also measured and in all samples the amount of DNA was too low to detect (below <50 ng/mL).

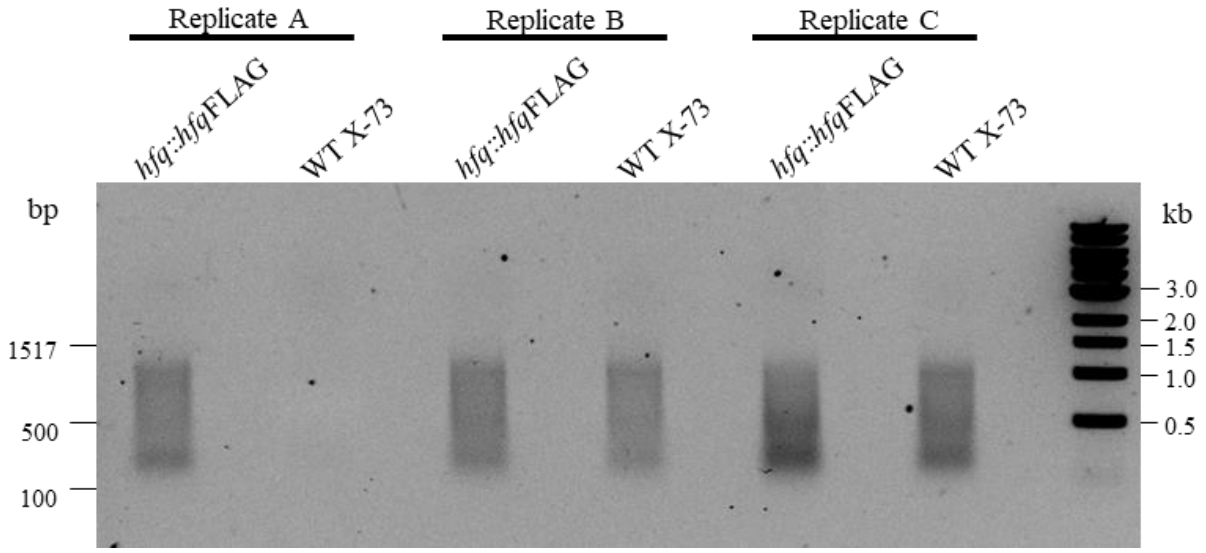


Figure 4.8. Electrophoretic separation of RNA from three independent Co-IP experiments, isolated from *hfq*::*hfq*FLAG and wild-type strains. Ten percent of the total yield recovered for each RNA sample was separated on the gel. Sizes of the NEB molecular weight markers are indicated to the left and right of the gel.

4.3.1.4. Validation of RNA isolated from Hfq Co-IP

To check for low levels of DNA contamination, PCR was employed using primers that bind to the *P. multocida gcvB* (BAP7888 and BAP7889, Table 4.2). No PCR product was detectable when any of the RNA samples was used as template with these primers, indicating that there was no significant DNA contamination (data not shown). To check for the presence of co-immunoprecipitated sRNAs within each RNA sample, RNA was first converted to cDNA then primers (BAP7888 and BAP7889, Table 4.2) specific for GcvB (the well-characterised Hfq-dependent sRNA) (Gulliver et al., 2018), were used in separate PCR reactions using cDNA generated from both *hfq*::*hfq*FLAG and wild-type RNA samples (Figure 4.9) as the template. A PCR product of the expected size for GcvB (102 bp) was amplified in all reverse transcribed (RT+) *hfq*::*hfq*FLAG RNA samples, indicating that the GcvB sRNA had been co-immunoprecipitated with the 3xFLAG-tagged Hfq as expected (Figure 4.9). No GcvB-specific PCR product was observed in the RT- samples. Furthermore, no GcvB-specific PCR product was identified in the WT X-73 replicate A RNA samples, although a very minor product was observed in the WT X-73 B and C RT+ samples. These data suggest that the co-immunoprecipitation had enriched for Hfq-dependent RNA species. Also present in most lanes was a smaller band representative of primer dimers (Figure 4.9).

To confirm expression of 3xFLAG-tagged Hfq in the Co-IP samples representing the *hfq*::*hfq*FLAG strain, western blot analysis was performed using replicate A samples with an anti-FLAG monoclonal antibody (Figure 4.10). Whole cell lysates representing wild-type X-73 (negative control), the *hfq*::*hfq*FLAG strain (positive control) and and an aliquot of each cell lysate and bead sample generated before and after the Co-IP procedure, respectively, were separated by electrophoresis on an SDS-PAGE gel, transferred to a nylon membrane then incubated with an anti-FLAG monoclonal antibody. A band corresponding to the 3xFLAG-tagged Hfq protein was observed in all lanes containing FLAG-tagged Hfq (~18 kDa) but was not present in samples representing the wild-type X-73 strain. A Coomassie-stained SDS-PAGE gel containing the identical samples was also analysed to confirm equal loading of like samples (Figure 4.10). The protein profiles were similar across like samples; each of the Co-IP samples had two distinct and strongly reacting protein bands, which likely represent the heavy and light chains of the monoclonal antibody used in the Co-IP (50 kDa and 25 kDa, respectively).

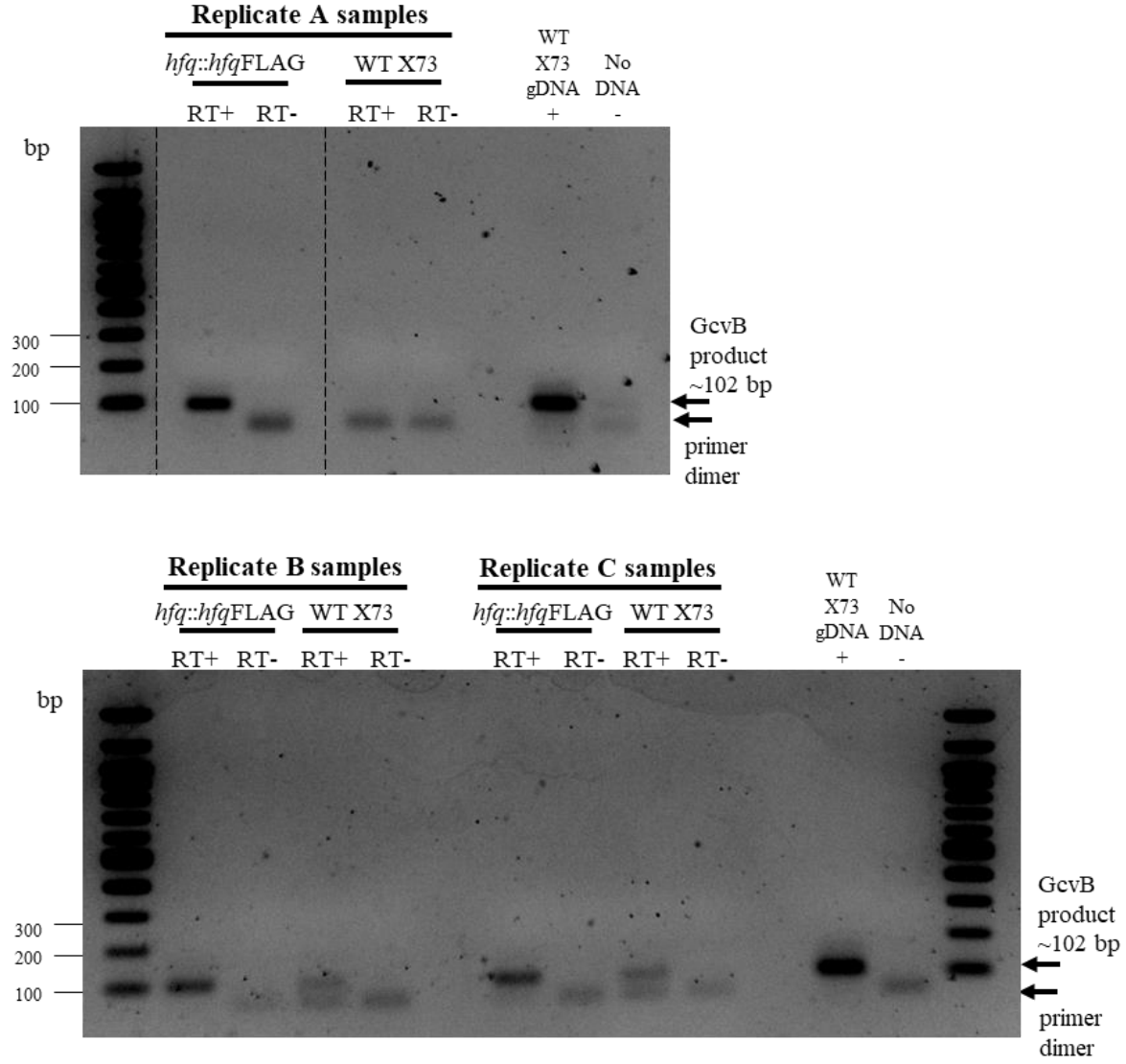


Figure 4.9. Electrophoretic separation of GcvB-specific PCR products amplified from samples derived from three independent Co-IP experiments using *hfq*::*hfq*FLAG and wild-type *P. multocida* strains. A one in ten dilution of each RNA sample was used as template directly in a PCR reaction (RT-) or converted to cDNA with addition of reverse transcriptase (RT+). Sizes of the molecular weight markers are indicated at the left of the gels. The GcvB product (~102 bp) and predicted primer dimer products are indicated by arrows to the right of the gels.

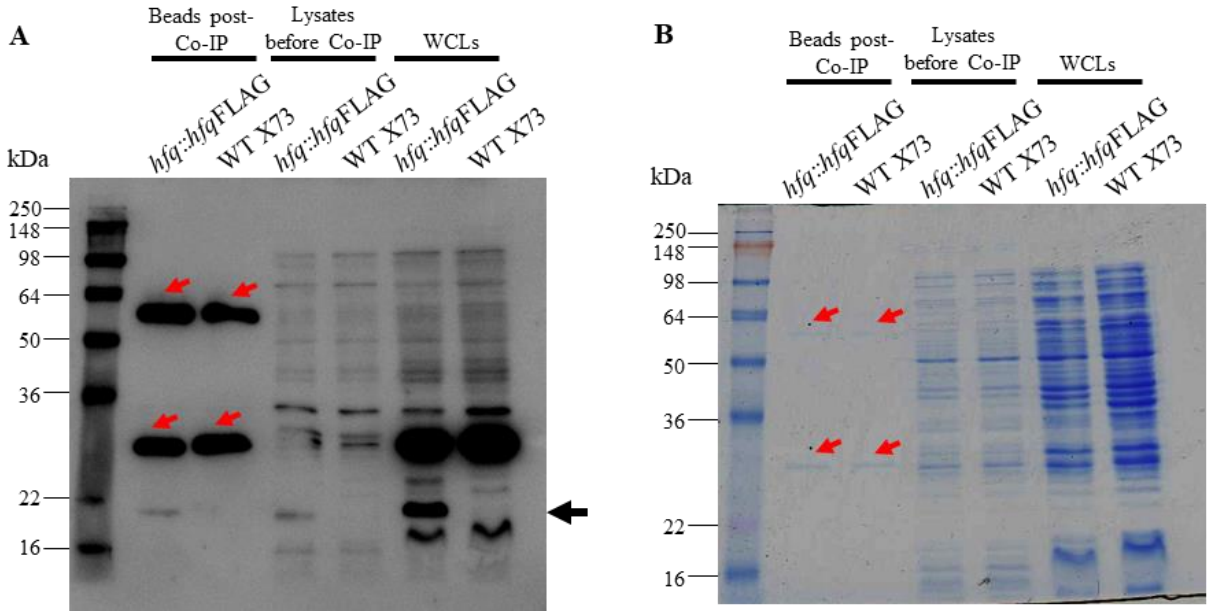


Figure 4.10. Western blot (**A**) and polyacrylamide gel electrophoretic separation (**B**) of the Co-IP samples (replicate A *hfq*::*hfq*FLAG and wild-type Hfq). From left to right; bead samples following Co-IP, cell lysates with the magnetic beads just prior to Co-IP and whole cell lysates only (WCLs). **A**. Western blot analysis using monoclonal anti-FLAG antibody (10 min exposure). The 3xFLAG-tagged Hfq protein is indicated by the black arrow and is present in all *hfq*::*hfq*FLAG samples (~18 kDa). **B** Coomassie stained polyacrylamide gel of *hfq*::*hfq*FLAG and wild-type samples following Hfq Co-IP. Protein bands in the two left lanes (red arrows) of both the immunoblot and polyacrylamide gel are predicted to represent the heavy and light chains of the antibody present on the M2 magnetic beads (**A** and **B)**. Sizes of protein markers are indicated at the left of each gel.

4.3.1.5. Sequencing of Hfq-associated RNA derived from Hfq Co-IP

All Co-IP derived RNA samples (three replicates each of RNA from wild-type and *hfq*::*hfq*FLAG paired samples) were sequenced using a strand-specific protocol (Agilent) on an Illumina NextSeq platform. All reads were aligned to the *P. multocida* X-73 genome sequence (NCBI genome accession: NZ_CM001580.1). Transcripts were identified as Hfq-associated if there was a statistically significant increase in reads recovered across all the *hfq::hfq*FLAG samples (triplicate replicates) relative to the untagged replicates at a false discovery rate (FDR) of < 0.05. Transcripts that shared high identity with genes in the *P. multocida* strain Pm70 genome (May et al., 2001) or with genes in other bacterial species were assigned the same gene name. A total of 90 transcripts were identified as associated with the Hfq protein using this analysis (Table 4.3).

The Hfq-dependent sRNA GcvB, was enriched 11.1-fold (log_2 expression ratio = 3.47; false discovery rate [FDR] = 6.7×10^{-6}) in the Hfq Co-IP (Figure 4.11 panel A, Table 4.3). However, only two of the putative *P. multocida* sRNAs identified in Chapter 3 displayed a clear association with Hfq. These were the putative sRNAs Prrc10 and Prrc19, which showed significant enrichment following Hfq Co-IP (Figure 4.11 panels B. and C, Table 4.3). No other novel putative sRNAs were identified from the analysis. However, a number of mRNA transcripts were enriched by Hfq Co-IP (Table 4.3), including the *hfq* transcript itself (14.5-fold, FDR = 1.27×10^{-6}) (Figure 4.11 panel D).

As the GcvB sRNA was strongly enriched in the Co-IP, the transcript levels of the previously predicted GcvB mRNA targets (Gulliver et al., 2018) were assessed (Table 4.3). Genes encoding one hypothetical *P. multocida* protein (X73_RS07870), a peptide methionine sulfoxide reductase (MsrA), a periplasmic dipeptide transport protein precursor (DppA), and the dihydrodipicolinate synthase protein DapA, each displayed enrichment in the Co-IP with tagged Hfq (Table 4.3).

Genes that displayed altered transcript levels in the *hfq* mutant (RNA-Seq analysis of the *hfq* mutant described in Chapter 2) were also inspected for transcript enrichment following Hfq Co-IP. Ten of the genes with increased expression in the *hfq* mutant displayed a clear association with Hfq by Co-IP, including the outer membrane lipoprotein *plpB* (X73_RS01475), the phosphocholine (PCho) transferase *pcgD* (X73_RS04500), and the preprotein translocase subunit *secG* (X73_RS09230) (Figure 4.11 panels E., F. and G, Table 4.3). A number of these genes with increased transcription in the *hfq* mutant also showed increased protein production, during early- and/or mid-exponential growth phase (Mégroz et al., 2016) (Chapter 2), including the *plpB* transcript (Table 4.3).

Model-based analysis of ChIP-Seq (MACS2) (Zhang et al., 2008) was also separately used to identify Hfq-associated transcripts from the Hfq Co-IP data. Using this analysis, >95% of the RNA-Seq reads enriched by Hfq were mapped to rRNAs, Ribonuclease P (RNase P) and 6S/SsrS RNA genes; no peaks specifically associated with any of the predicted sRNAs were identified.

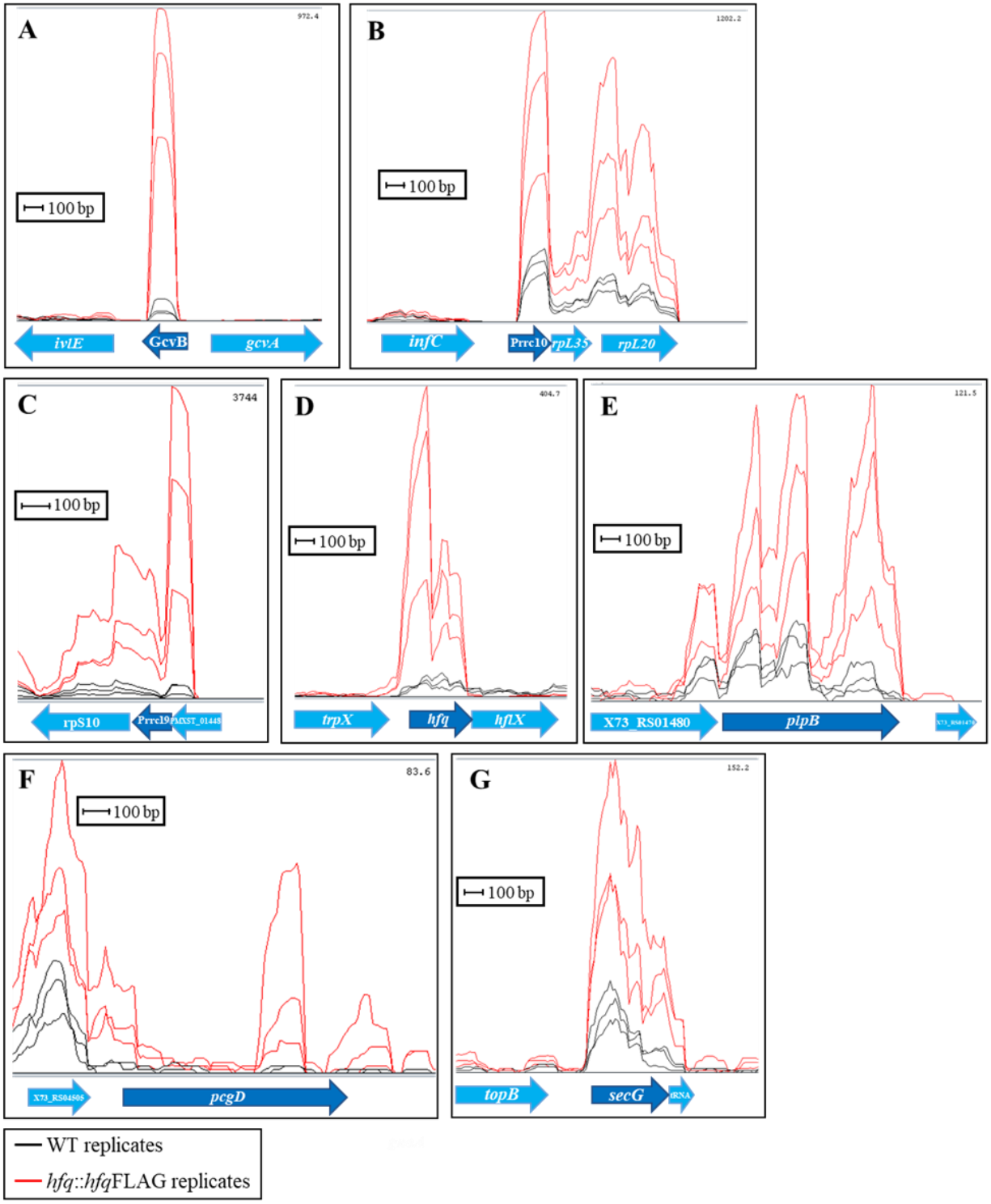


Figure 4.11. Identification of Hfq-associated RNAs by visual inspection of Hfq Co-IP RNA-Seq reads mapped to the *P. multocida* X-73 genome. Genomic regions with high numbers of reads in all *hfq*::*hfq*FLAG strain Co-IP RNA samples, and low numbers of reads in all wild-type Co-IP RNA samples were considered to be enriched by Hfq Co-IP and therefore Hfq-associated. **A-G**. RNA-Seq reads generated from wild-type (indicated by the black lines) and *hfq*::*hfq*FLAG strain (indicated by the red lines) replicates mapped to sRNA and mRNA sequences and surrounding genes following Hfq-Co-IP. Mapped reads to the GcvB sRNA (**A**), putative sRNAs Prrc10 (**B**) and Prrc19 (**C**), the *hfq* mRNA sequence (**D**) as well as some selected mRNA sequences which displayed altered expression in the *hfq* mutant strain, *plpB* (**E**), *pcgD* (**F**), and *secG* (**G**), were enriched by Hfq Co-IP.

Table 4.3. Hfq Co-IP analysis. Transcripts in the *hfq::hfq*FLAG replicate Co-IP samples with increased RNA recovery in comparison to the untagged control samples (FDR <0.05). Predicted targets of GcvB are indicated by a blue asterisk (*) beside the PMX73_RS locus tag (Gulliver et al., 2018). Transcripts which showed increased expression in transcriptomic analysis of the *P. multocida* VP161 *hfq* mutant (Chapter 2) (Mégroz et al., 2016) are indicated by the red asterisk (*) beside the PMX73_RS locus tag; from these predicted *hfq* targets, those which also showed increased protein production in the *hfq* mutant (Chapter 2) (Mégroz et al., 2016) during early- or mid-exponential growth phase, respectively, are indicated by a purple (*) or green (*) asterisk.

PMX73_RS locus tag CM001580 [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
None	putative GcvB sRNA Prrc01			3.47	6.70E-06
None	putative sRNA Prrc10			1.16	7.10E-03
None	putative sRNA Prrc19			3.28	2.39E-04
X73_RS10285	no known function			2.40	2.34E-02
PMXST_01989				2.34	3.51E-03
X73_RS10405	no known function	PM0739		1.67	4.24E-02
X73_RS00100**	export factor homologue (Skp)	PM1993	*skp*	1.36	1.96E-02
X73_RS00120	undecaprenyl diphosphate synthase	PM1989	*uppS*	1.79	2.11E-02
X73_RS00390	no known function			1.90	9.04E-03
X73_RS00625	acyl carrier protein	PM1917	*acpP*	1.15	2.55E-02
X73_RS00670	oligopeptide transport system permease protein OppC	PM1908	*oppC*	1.31	3.58E-02
X73_RS00845*	branched-chain amino acid transport system carrier protein BrnQ	PM1873	*brnQ*	1.42	2.34E-02
X73_RS01020	YihI	PM1835	*yihI*	1.63	7.59E-03
X73_RS01140	nitrate/nitrite response regulator protein	PM1810	*narL*	1.02	2.29E-02
X73_RS01460	DNA-binding protein HU-alpha	PM1732	*hupA*	1.37	4.42E-03
X73_RS01475***	outer membrane lipoprotein 2 precursor	PM1730	*plpB*	1.25	1.34E-02
X73_RS01690	Sec-independent periplasmic protein translocation protein subfamily	PM1691	*tatC*	1.28	2.65E-02
X73_RS01695	Type V Secretory Pathway or Twin Arginine Targeting (Tat) Family protein subfamily, putative	PM1690	*tatB*	1.32	2.24E-02

PMX73_RS locus tag CM001580 [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS01870*	no known function			3.21	5.67E-03
X73_RS02105	no known function			2.68	2.11E-02
X73_RS02110	transaldolase B	PM1602	*talB*	1.19	5.67E-03
X73_RS02520	FtsY	PM1519	*ftsY*	1.36	1.30E-02
X73_RS02635	TRK system potassium uptake protein	PM1498	*trkH*	1.24	4.92E-02
X73_RS02680	F0F1 ATP synthase subunit A	PM1488	*atpB*	1.26	2.03E-02
X73_RS02700	ZapB	PM1483	*zapB*	4.18	3.94E-06
X73_RS02910	anaerobic glycerol-3-phosphate dehydrogenase, subunit C	PM1440	*glpC*	1.81	6.14E-03
X73_RS02945	Rhodanese-like domain protein	PM1433	*yibN*	1.22	2.11E-02
X73_RS03030	30S ribosomal protein S10	PM1416	*rpsJ*	1.84	6.21E-04
X73_RS03190	no known function	PM1388		1.40	1.62E-02
X73_RS03445	hydrogen peroxide-inducible genes activator	PM1346	*oxyR*	2.76	3.17E-04
X73_RS03650	no known function	PM1316		1.75	4.68E-03
X73_RS03910**	thiamine biosynthesis protein, putative	PM1263		1.98	1.13E-02
X73_RS04255	inorganic pyrophosphatase	PM1191	*ppa*	1.25	1.96E-02
X73_RS04305	transcriptional repressor LexA	PM1181	*lexA*	1.24	2.11E-02
X73_RS04350	YgiW	PM1171	*ygiW*	1.48	1.30E-02
X73_RS04390	50S ribosomal protein L34	PM1162	*rpmH*	1.32	4.14E-02
X73_RS04415	cAMP-regulatory protein	PM1157	*crp*	1.17	1.00E-02
X73_RS04420	no known function	PM1156		1.61	8.04E-03
X73_RS04475	D-glycero-D-manno-heptosyl transferase	PM1144	*hptE*	0.98	4.38E-02
X73_RS04485	PcgC		*pcgC*	2.78	6.14E-03
X73_RS04490	PcgB		*pcgB*	2.31	8.04E-03
X73_RS04500*	PcgD, Phosphocholine (PCho) transferase, required for the addition of PCho residues to lipopolysaccharide		*pcgD*	2.63	8.71E-03
X73_RS04515	FtsN	PM1136	*ftsN*	2.50	9.01E-04
X73_RS04670	molybdate-binding periplasmic protein, putative	PM1108		2.09	1.12E-02
X73_RS04815	no known function	PM1077		2.05	6.14E-03

PMX73_RS locus tag CM001580 [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS04950*	dihydrodipicolinate synthase	PM1051	*dapA*	0.94	4.40E-02
X73_RS04960	Integral membrane protein domain protein	PM1049		1.24	2.15E-02
X73_RS05055	translation initiation factor IF-1	PM1031	*infA*	1.54	3.82E-02
X73_RS05255	thioredoxin	PM0994	*trxA*	1.33	2.11E-02
X73_RS05335***	no known function	PM0979		0.93	3.57E-02
X73_RS05355	no known function	PM0975		2.71	4.42E-03
X73_RS05760	RNA-binding protein Hfq	PM0906	*hfq*	3.86	1.27E-06
X73_RS05895	no known function	PM0880		2.68	2.03E-02
X73_RS05935	H-NS	PM0872	*hns*	1.29	9.52E-03
X73_RS05940	no known function	PM0871		1.20	2.03E-02
X73_RS06015	CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyltransferase	PM0856	*pgsA*	1.53	2.11E-02
X73_RS06030*	fimbrial protein-related protein	PM0855		2.74	4.07E-03
X73_RS06165	glutaredoxin 1	PM0827	*grxA*	1.76	1.84E-02
X73_RS06265	putrescine-ornithine antiporter	PM0807	*potE*	2.42	3.84E-02
X73_RS06325	cystathionine beta-lyase	PM0794	*metC*	2.25	2.11E-02
X73_RS06455	predicted membrane protein	PM0771		1.32	4.91E-02
X73_RS06795	conserved hypothetical protein	PM0703		2.91	6.14E-03
X73_RS06800	colicin V production protein	PM0702	*cvpA*	2.19	4.58E-02
X73_RS06910	no known function	PM0679	*ybaN*	2.65	1.06E-02
X73_RS07020	tellurite resistance protein TehB	PM0656	*tehB*	1.76	1.75E-02
X73_RS07275*	peptide methionine sulfoxide reductase	PM0605	*msrA*	1.99	7.59E-03
X73_RS07285	50S ribosomal protein L20	PM0604	*rplT*	1.18	6.14E-03
X73_RS07290	50S ribosomal protein L35	PM0603	*rpmI*	0.83	4.55E-02
X73_RS07690	Ribosomal protein S9/S16	PM0521	*rpsI*	2.38	1.26E-04
X73_RS07695	50S ribosomal protein L13	PM0520	*rplM*	1.23	2.37E-02
X73_RS07865	no known function	PM0473		1.71	1.29E-03
X73_RS07870*	no known function	PM0472		1.26	6.26E-03

PMX73_RS locus tag CM001580 [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS08140	5-methyltetrahydropteroyltriglutamate-homocysteine methyltransferase	PM0420	*metE*	3.16	4.12E-04
X73_RS08200	formate dehydrogenase	PM0406	*fdnI*	1.26	2.77E-02
X73_RS08325	sodium-dependent transporter, putative	PM0380		1.78	8.71E-03
X73_RS08400	Na+/H+ antiporter, putative	PM0365	*nhaP*	1.43	4.79E-02
X73_RS08520	sodium/proline symporter (proline permease)	PM0588	*putP_1*	1.50	1.13E-02
X73_RS08550	GrpE	PM0334	*grpE*	1.91	2.69E-03
X73_RS08705	30S ribosomal protein S15	PM0301	*rpsO*	2.52	9.13E-05
X73_RS08800				2.66	1.73E-03
X73_RS09075*	periplasmic dipeptide transport protein precursor	PM0236	*dppA*	1.17	4.75E-02
X73_RS09160	YkgO		*ykgO*	2.35	3.25E-02
X73_RS09230*	preprotein translocase subunit SecG	PM0208	*secG*	0.98	2.58E-02
X73_RS09245	nucleoid-associated protein YbaB	PM0205	*ybaB*	2.06	5.67E-03
X73_RS09600	transcriptional regulator MraZ	PM0133	*mraZ*	2.83	3.17E-04
X73_RS09670	no known function	PM0119		2.10	2.83E-03
X73_RS09895	probable formate transporter 1	PM0074	*focA*	1.48	1.31E-02
X73_RS09920*	6-phosphofructokinase	PM0069	*pfkA*	1.18	2.11E-02
X73_RS10385	Pentapeptide repeats (8 copies) domain protein	PM0591		2.22	1.13E-02
X73_RS10505	no known function	PM1487		2.74	4.42E-03

[a] Multiple genome entries exist for strain X-73; NCBI genome accession: NZ_CM001580.1 was chosen for this study. Where the protein has no PMX73_RS locus tag, a designated PMXST_ locus tag from the *P. multocida* X-73 draft genome is provided.
[b] The *P. multocida* sequence was used to interrogate the SwissProt database (excluding proteins belonging to *Pasteurellaceae*). Only proteins with a high similarity with known proteins/domains are assigned a predicted function i.e. an alignment score of ≥ 200 (graphical output = red bar).
[c] The Pm70 genome is available from NCBI genome database, accession: NC_002663.1 or AE004439.1.

4.3.2. Hfq Co-IP: Hfq UV-CRAC and UV-CLASH

To augment the data from initial Co-IP experiments (Section 4.3.1) and in an attempt to identify additional Hfq-associated sRNA and mRNA species, a second Hfq co-immunoprecipitation strategy was used that incorporated both Hfq-dependent UV-CRAC and UV-CLASH (Figure 4.12). These Hfq UV-CRAC/CLASH experiments were performed in the Laboratory of Dr Jai Tree at the University of New South Wales by Emily Gulliver (PhD candidate, Boyce Laboratory). Although the initial Hfq Co-IP strategy described above was performed in the *P. multocida* strain X-73, for consistency with other *P. multocida* UV-CRAC/CLASH experiments, performed in parallel in the Tree Laboratory, we used the closely related strain *P. multocida* VP161.

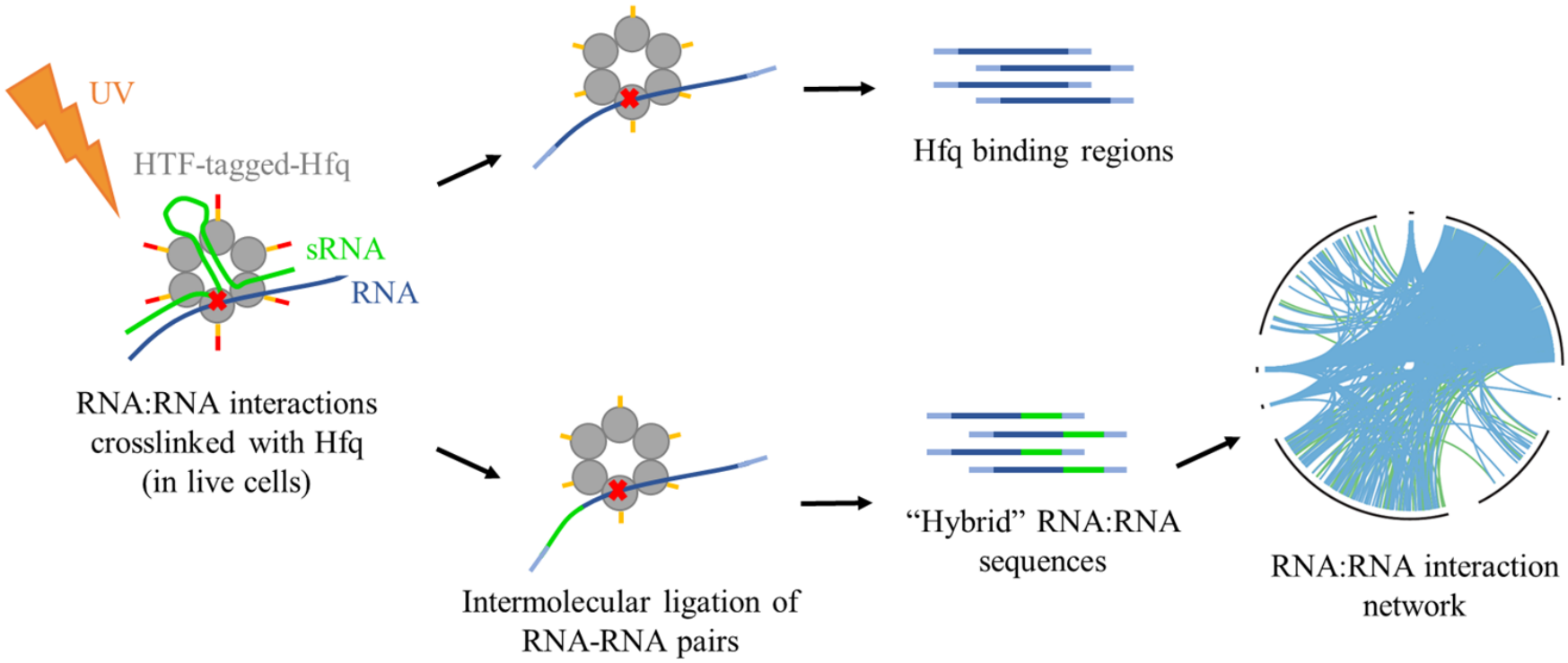


Figure 4.12. Strategy for identification of RNA species interacting with Hfq targets using Hfq UV-CRAC/CLASH (Modified from Waters et al., 2017). Both His6-TEV-FLAG-tagged Hfq (Δ*hfq*/Δ*hyaD*::*hfq*HTF) and untagged (control) (Δ*hfq*/Δ*hyaD*::*hfq*) strains are used in the analysis. The tagged Hfq hexamer (grey circles with HTF-tags indicated by the yellow and red lines) facilitates the interaction between one or more RNA molecules; an sRNA molecule is indicated by the green line and another interacting RNA molecule is indicated by the dark blue line. The interacting RNA molecules are then UV-crosslinked (orange thunderbolt symbol) to the Hfq protein. The Hfq-RNA complexes undergo co-immunoprecipitation steps, before the Hfq-associated RNA molecules are isolated (dark blue lines). Following adapter ligation (light blue lines) to allow for sequencing, the RNA molecules are sequenced, and the Hfq-associated RNA identified (UV-CRAC; top panel). To identify directly interacting RNA molecules associated with Hfq, an intermolecular ligation step is performed following UV-crosslinking, resulting in isolation of hybrid RNA sequences (UV-CLASH; lower panel). These hybrid RNA sequences may include mRNA-sRNA pairs (dark blue and green lines) and can be sequenced to generate a Hfq mediated sRNA-RNA interaction network.

4.3.2.1. Construction of an acapsular His6-TEV-FLAG-tagged-Hfq VP161 strain for use in Hfq UV-CRAC/CLASH experiments

To allow the Hfq UV-CRAC/CLASH experiments to be performed, a non-virulent, Hfq-tagged *P. multocida* strain was required for safe use in the University of New South Wales (Dr Jai Tree Laboratory) UV apparatus. As capsule is known to be a crucial *P. multocida* virulence factor (Boyce and Adler, 2000; Chung et al., 2001), a *P. multocida hyaD* mutant was constructed; *hyaD* is essential for synthesis of the capsular polysaccharide (Mégroz et al., 2016) and acapsular mutants are known to be highly attenuated in mice and chickens. Also required for the UV-CRAC/CLASH experiments was the construction of a His6-TEV-FLAG-tagged Hfq. These three tags were necessary to allow efficient two-step Co-IP. The His6-tag was necessary for the first Co-IP step using anti-His antibodies with nickel beads under denaturing conditions. The Tobacco Etch Virus (TEV) endopeptidase cleavage sequence allowed removal of the His6-tag by incubation with the TEV protease. The FLAG-tag was used for further purification in the presence of anti-FLAG resin (Waters et al., 2017).

To generate the recombinant strain, a markerless *hyaD* mutant was constructed using the *P. multocida* VP161 *hfq* mutant (AL2521) (Mégroz et al., 2016). This was achieved using a markerless TargeTron plasmid targeted to *hyaD* (pAL1337, Table 4.1). Colony PCR screening of transformants was employed using the TargeTron-specific primer EBS universal together with a *hyaD*-specific primer, BAP2067, to identify transformants with an intron in *hyaD* (Table 4.2). Three putative *hyaD* mutants were selected and cured of the TargeTron plasmid as described previously (Harper et al., 2013). Additional colony PCR reactions were performed using primers BAP2067 and BAP5781 that flanked the intron insertion point in *hyaD* (Table 4.2). Selected mutants were then sequenced using genomic DNA as template with the EBS universal primer to confirm the correct position of the intron in the genome. To confirm that both *hyaD* and *hfq* contained introns, two Southern blotting experiments were performed, one using a DIG-labelled probe specific for the kanamycin gene (present within the intron within *hfq*), and one using a DIG-labelled probe specific for the TargeTron intron (which should be present both in *hfq* and in *hyaD*); see Figure S3 in Appendix D. One of the *hyaD/hfq* double mutants with correct genetic profile was selected for further study and designated AL3089 (Table 4.1).

To engineer the HTF-tagged Hfq protein, the expression plasmid pAL1225 (Figure 4.1, Table 4.1), which encodes the *hfq* gene without its native stop codon, was modified. The DNA

fragment encoding the His6-TEV-tag, was constructed by annealing the two complementary primers, BAP8098 and BAP8099 to give a double-stranded fragment with EcoRI compatible ends that would allow for the regeneration of an EcoRI site only at the 3′ end following ligation. The fragment was ligated to EcoRI-digested pAL1225 and colonies screened via PCR using the vector-specific primers BAP612 and BAP2679 that flanked the insert (Table 4.2). One plasmid with correct sequence was designated pAL1343 (Table 4.1). To introduce the 3xFLAG-tag to the construct, a DNA fragment with EcoRI compatible ends was generated by annealing primers BAP8100 and BAP8101 (Table 4.2) then ligated to EcoRI-digested pAL1343. Colonies were screened by PCR using primers that flanked the region (BAP612 and BAP2679) and positive transformants were confirmed by DNA sequencing. One plasmid encoding the His6-TEV-FLAG-tagged Hfq (HTF-tagged Hfq) was designated pAL1345 (Table 4.1). As a negative control for UV-CRAC/CLASH, the plasmid pAL1226 encoding an untagged Hfq with its native stop codon was used (Figure 4.1., Table 4.1). For use in the Hfq UV-CRAC/CLASH experiments, the two plasmids, pAL1345 (encoding HTF-tagged Hfq) and pAL1226 (encoding untagged Hfq) were then separately used to transform the *P. multocida hyaD/hfq* double mutant (AL3089) to generate AL3091 and AL3092, respectively (Table 4.1).

To confirm HTF-tagged-Hfq expression in AL3091, before use in the Hfq UV-CRAC/CLASH experiments, western immunoblotting was performed using a monoclonal anti-FLAG antibody (Figure 4.13). A protein of ~30 kDa, corresponding to the predicted size of HTF-tagged Hfq, was detected only in the whole cell lysates of AL3091, but not in the negative control strains AL3100, containing empty vector pREXY, and AL3092, expressing untagged Hfq. It is unclear why a larger reacting band (~33 kDa) was also observed in the HTF-tagged-Hfq in the *hfq/hyaD* mutant (AL3091).

To check that the tagged Hfq was functioning, the plasmid, pAL1345, was introduced into the *P. multocida hfq* mutant (AL2521) to generate AL3101 (Table 4.1) and assessed for capsule production using a HA capsule assay. HA capsule production in AL3101 was significantly increased compared to the *hfq* mutant strain ($P = 0.026$) and was restored to near wild-type levels (~77.1 % of wild-type; $P > 0.05$), indicating that the plasmid-encoded recombinant HTF-tagged-Hfq was functional (data not shown).

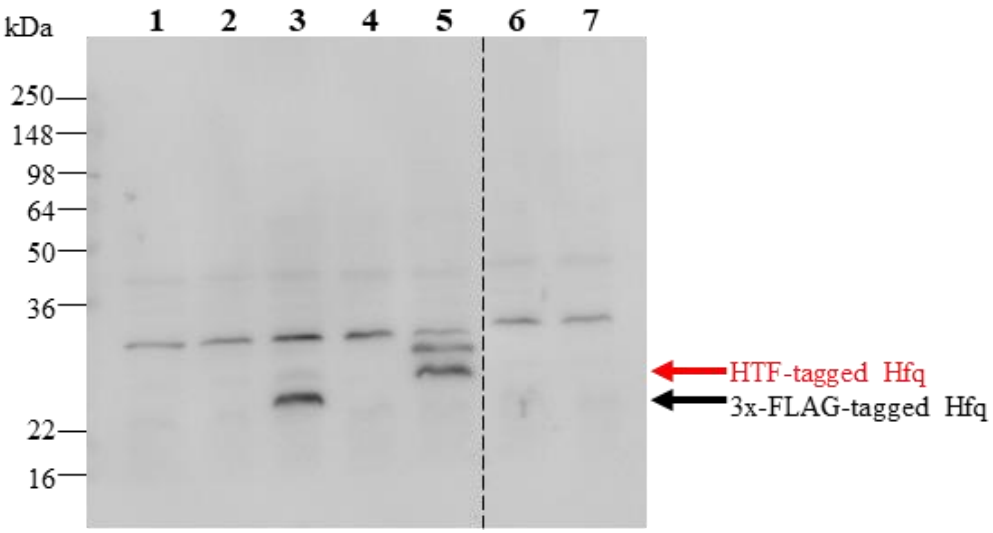


Figure 4.13. **Western immunoblot analysis of *P. multocida* strains using a monoclonal anti-FLAG antibody** (5 min exposure). Lane 1, *P. multocida* wild-type VP161; lane 2, *P. multocida* wild-type X-73; lane 3, *hfq*::*hfq*FLAG strain AL2842; lane 4, *hfq/hyaD* mutant (AL3089); lane 5, HTF-tagged-Hfq in the *hfq/hyaD* mutant (AL3091); lane 6, *hfq/hyaD* mutant harbouring pAL1226 (AL3092); lane 7, *hfq/hyaD* mutant harbouring pREXY empty vector (AL3100). Sizes of protein markers are indicated to the left of the gel. The HTF-tagged Hfq protein in AL3091 (~30 kDa, lane 5) and the 3xFLAG-tagged Hfq protein in AL2842 (~18 kDa, lane 3) are indicated by the red and black arrows, respectively, to the right of the gel.

4.3.2.2. Determination of Hfq-associated RNA species using Hfq UV-CRAC and UV-CLASH

In vivo UV-crosslinking between the *P. multocida* Hfq protein and any associating RNA was performed on strains expressing HTF-tagged-Hfq and untagged Hfq, respectively. Two biological replicates of the HTF-tagged-Hfq and untagged Hfq strains (AL3091 and AL3092, respectively) were used in the experiments. Multistep co-immunoprecipitation was performed, using anti-His antibodies, followed by TEV peptidase cleavage and then anti-FLAG resin to recover Hfq-associated RNA species. RNA molecules that had been crosslinked to Hfq were isolated from each sample, reverse transcribed into cDNA and sequenced. The RNA-Seq reads were mapped to the *P. multocida* VP161 genome and analysed for transcripts that showed increased recovery in the HTF-tagged Hfq strain compared with the untagged Hfq strain. Furthermore, to identify specific RNA:RNA interactions, an additional intramolecular RNA ligation step was performed after the UV-crosslinking. This step allowed for ligation of interacting RNA molecules, generating hybrid RNA molecules. Hybrid RNA molecules were identified from bioinformatic analysis of the cDNA sequences as continuous sequences that matched with two different regions on the *P. multocida* VP161 genome, representing two different RNA species interacting together, and/or with the same Hfq molecule.

Following analysis of the Hfq UV-CRAC sequencing data, only three RNA transcripts were recovered in both replicates of the HTF-tagged-Hfq that were not present in the untagged controls (Peaks 1, 2 and 3; Figure 4.14). Hfq showed association with the central region of *ner* (PmVP161_0399) (Peak 1), a predicted *P. multocida* transcriptional regulator (Figure 4.14A). Hfq also showed significant association with *glmU* (Peak 2), that in other bacteria has been shown to be involved in hyaluronic acid (HA) biosynthesis (de Oliveira et al., 2016) (Figure 4.14B). A third intergenic region was also identified, upstream of a gene encoding a predicted *P. multocida* protein of unknown function (PMVP_0398) (Peak 3) (Figure 4.14C). Interestingly, neither of these three transcript regions identified in the Hfq UV-CRAC experiments were identified as associated with Hfq using native Hfq Co-IP (Section 4.3.1.5).

The sRNA GcvB is known to be Hfq dependent in other species (Sharma et al., 2011) and indeed showed clear association with *P. multocida* Hfq in the initial Co-IP analysis (Section 4.3.1.5). By visually inspecting the Hfq UV-CRAC data using the Integrated Genome Browser (IGB) (Freese et al., 2016), high numbers of transcript reads mapped to the GcvB sequence in the Hfq-tagged replicates relative to the untagged control replicates (Figure 4.15), indicating

association of GcvB with Hfq via UV-CRAC; however, this did not reach statistical significance.

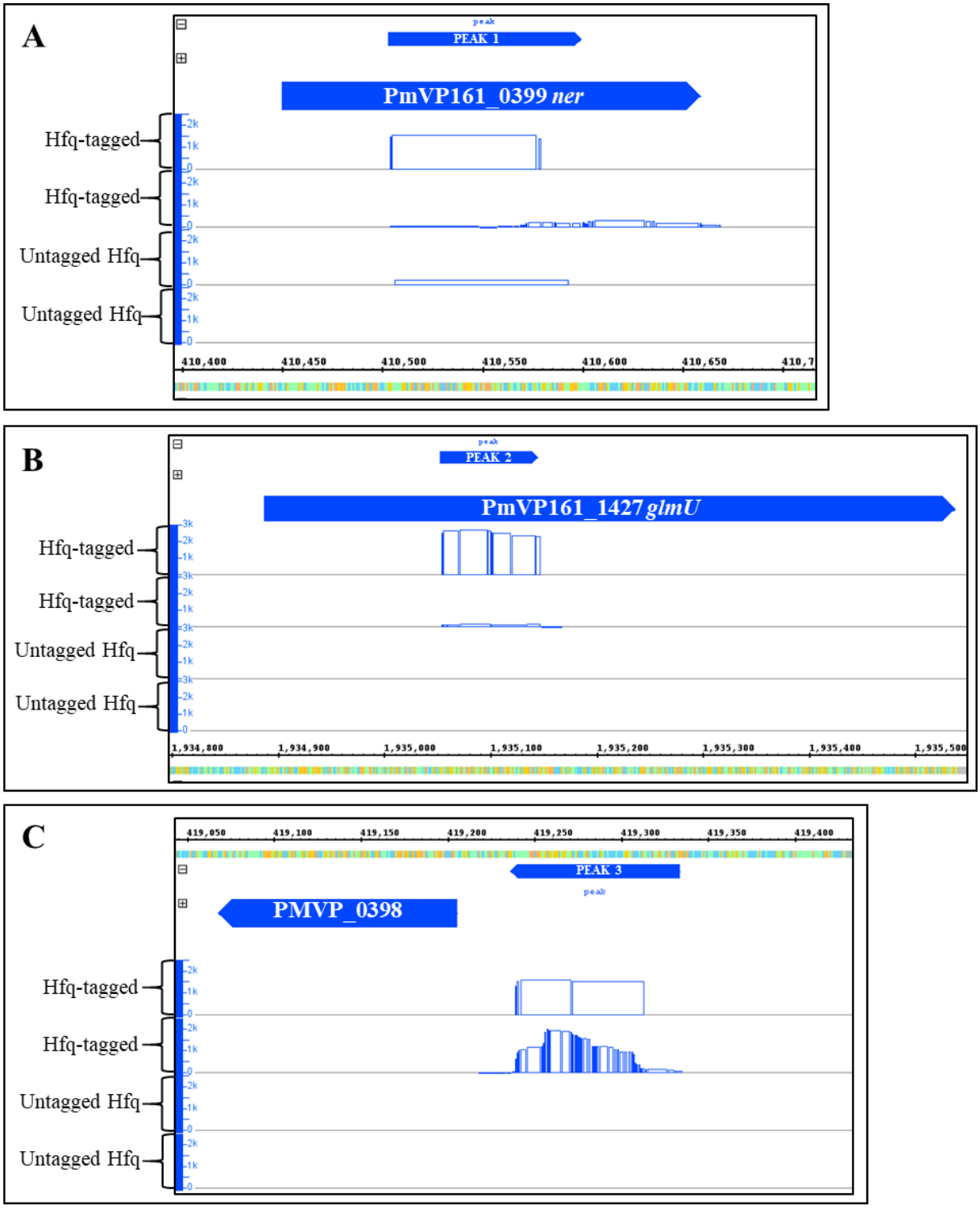


Figure 4.14. RNA-Seq reads generated from HTF-tagged Hfq (Δ*hfq*/Δ*hyaD*::*hfq*HTF) and control untagged strain replicates (biological duplicate) using UV-CRAC were mapped to the *P. multocida* VP161 genome and visualised using IGB. Reads mapping to peak 1 in **panel A** (overlapping PmVP161_0399, *ner* ; 424899-424995), peak 2 in **panel B** (overlapping PmVP161_1427, *glmU*; 1569393-1569301), and peak 3 in **panel C** (overlapping an intergenic region, upstream of PMVP_0398; 433729-433632) are shown (*P. multocida* strain VP161 NCBI genome accession: CP048792). Note: data were originally mapped and analysed using the old VP161 genome; the PMVP_0398 gene shown in panel C has no homologue in the new *P. multocida* VP161 genome.

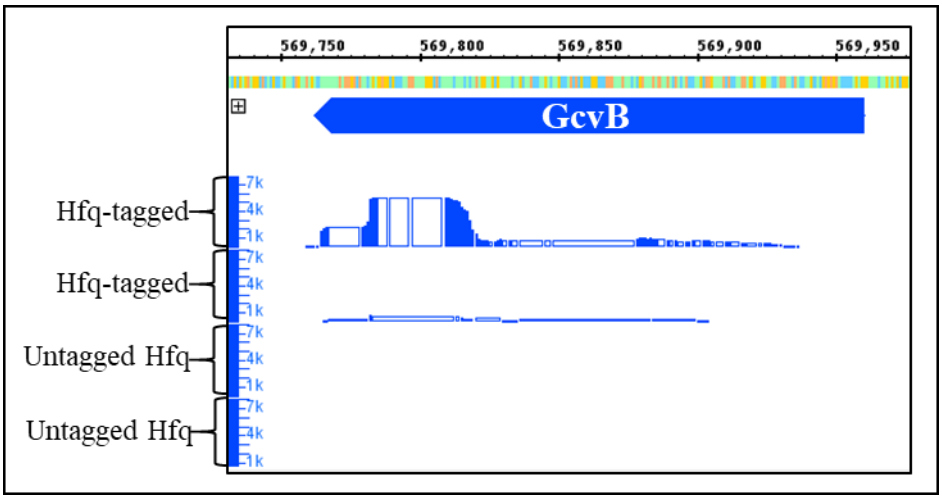


Figure 4.15. RNA-Seq reads generated from HTF-tagged Hfq (Δ*hfq*/Δ*hyaD*::*hfq*HTF) and control untagged Hfq replicates (biological duplicate) using UV-CRAC were mapped to the *P. multocida* VP161 genome and visualised using IGB. Reads mapping to the GcvB sRNA are shown.

4.3.2.3. Investigating the role of *P. multocida* GlmU

Transcripts representing the *glmU* mRNA region (peak 2, section 4.3.2.2) showed significant association with Hfq in the Hfq UV-CRAC analysis (Figure 4.14B). In other bacteria, the *P. multocida* GlmU protein is involved in the hyaluronic acid (HA) biosynthesis pathway and cell wall/membrane/envelope biogenesis (de Oliveira et al., 2016). Based on transposon-directed insertion site sequencing (TraDIS) analysis of *P. multocida* strain VP161 (Smallman, Harper, Boyce, unpublished), *glmU* was identified as an essential gene. As Hfq has previously been shown to regulate capsule production in *P. multocida* and an association of Hfq with *glmU* was observed in the Hfq UV-CRAC analysis, further investigation of this association was undertaken. However, from the initial Hfq Co-IP experiments (section 4.3.1), this region was not enriched by Hfq Co-IP and the expression of *glmU* was unaffected in the *hfq* mutant as was the production of GlmU as measured by proteomics (Chapter 2) (Mégroz et al., 2016).

To investigate the role of *P. multocida glmU,* a plasmid expressing wild-type VP161 *glmU* mRNA sequence was constructed. To construct the *glmU* overexpression plasmid, the wild-type *glmU* sequence (including the predicted native ribosome binding site) was amplified from *P. multocida* VP161 genomic DNA using BAP8655 and BAP8656 primers (Table 4.2) and cloned between SmaI and EcoRI restriction sites in pREXY. A plasmid with verified wild-type *glmU* sequence was designated pAL1477 (Table 4.2).The *hfq* mutant AL2521, which produces reduced levels of HA capsule (reported in Chapter 2) was then transformed with pAL1477 to generate strain AL3454. Capsule production was then tested in AL3454 to check for restoration of HA capsule in the presence of the *glmU* expression plasmid. However, a visual comparison of colony morphology of AL3454 with the *hfq* mutant strain AL2521 and wild-type VP161 strains indicated that capsule production was not restored in AL3454 (data not shown). A quantitative HA capsule assay was performed to quantify capsule production by AL3454 however, capsule production in AL3454 was not significantly different to the *hfq* mutant strain containing pREXY vector only.

4.3.2.4. RNA hybrids associated with Hfq in UV-CLASH

Through the addition of a ligation step (Figure 4.12), the Hfq UV-CLASH analysis allowed identification of hybrid/chimera molecules where Hfq acts to facilitate specific RNA:RNA interactions. A number of these interactions involved ribosomal RNA (rRNA) species. As rRNA species are highly abundant within the cell, these interactions were deemed unlikely to be true Hfq-dependent interactions and therefore these hybrids were excluded from the

analysis. Seventeen unique hybrid molecules were identified (Table 4.4), including one RNA pair involving the putative Prrc03 sRNA interacting with a tRNA molecule (Ser).

Table 4.4. Pairs of hybrid sequences (hybrid molecule A and B) identified via Hfq UV-CLASH analysis. Hybrid sequences identified by this method may be RNA pairs that were directly interacting with Hfq. The table cells are shaded according to the categories of RNA species listed; intergenic regions (blue), mRNAs (orange), tRNAs (yellow), putative sRNAs (green) or miscellaneous RNA sequences (grey).

Hybrid molecule A			**Hybrid molecule B**		
VP161 locus tag (Position on genome)[a]	**Predicted function/product**	**RNA Type**	**VP161 locus tag (Position on genome)[a]**	**Predicted function/product**	**RNA Type**
PmVP161_1026 (1127310 – 117283)	YbaB	mRNA	PmVP161_1451 (1590828 – 1590786)	No known function	mRNA 3′ region and UTR
none (1102329 – 1102268)	Unknown	Intergenic region	PmVP161_0424 (452172 – 452208)	PmVP161_0424	mRNA
PmVP161_0320 (334653 – 334632)	CmoB	mRNA	PmVP161_0696 (746222 – 746179)	No known function	mRNA
none (475483 – 475514)	tRNA Pro	tRNA	none (1428578 – 1428614)	tRNA Arg	tRNA
none (542330 – 542380)	no known function	Miscellaneous Located over PmVP161_0521 3′ region and UTR	PmVP161_0568 (601751 – 601722)	no known function	mRNA
none (542381 – 542350)	no known function	Miscellaneous Located over PmVP161_0521 3′ region and UTR	PmVP161_1916 (2060786 – 2060814)	*hupA*	mRNA

Hybrid molecule A			**Hybrid molecule B**		
VP161 locus tag (Position on genome)[a]	**Predicted function/product**	**RNA Type**	**VP161 locus tag (Position on genome)[a]**	**Predicted function/product**	**RNA Type**
none (708256 – 708274)	no known function	Miscellaneous Located over *sppA* (PmVP161_0659) 5′ region	none (1719818 – 1719839)	Unknown	Intergenic region
Prrc03 (908059 – 908040)	Prrc03	sRNA	none (230819 – 230798)	tRNA Ser	tRNA
PmVP161_0914 (1001344 – 1001397)	Hfq	mRNA	none (1923991 – 1924011)	Unknown	Intergenic region
PmVP161_0255 (257984 – 258044)	FimA	mRNA	none (35628 – 35605)	tRNA Val	tRNA
none (254784 – 254755)	tRNA Met	tRNA	none (254709 – 254747)	tRNA Met	tRNA
PmVP161_0199 (206331 – 206395)	CydB	mRNA	none (1617245 – 1617289)	tRNA Thr	tRNA
PmVP161_0138 (138848 – 138885)	PepB	mRNA	PmVP161_1503 (1644488 – 1644532)	XylB	mRNA
none (1922741 – 1922798)	no known function	Miscellaneous Located over PmVP161_1780 3′ region	none (2088022 – 2087967)	no known function	Miscellaneous Located over *rpiR* (PmVP161_1942)

Hybrid molecule A			**Hybrid molecule B**		
VP161 locus tag (Position on genome)[a]	**Predicted function/product**	**RNA Type**	**VP161 locus tag (Position on genome)[a]**	**Predicted function/product**	**RNA Type**
PmVP161_1963 (2108735 – 2108754)	CysK	mRNA	none (2072836 – 2072886)	tRNA Trp	tRNA
PmVP161_1414 (1546256 – 1546213)	no known function	mRNA	PmVP161_1413 (1546142 – 1546101)	no known function	mRNA
none (1428578 – 1428614)	tRNA Pro	tRNA	none (1428522 – 1428558)	tRNA Pro	tRNA

[a] The numbers of base pairs matching the sequences are included from the *P. multocida* VP161 genome (NCBI genome accession: CP048792).

4.3.2.5. Identification of a conserved tRNA binding motif in hybrid sequences

The RNA hybrid sequences identified by Hfq-associated UV-CLASH were investigated for the presence of conserved sequences using the Multiple Em for Motif Elicitation (MEME) tool (Bailey et al., 2015). All thirty-four sequences comprising the hybrid pairs, as well as separate groups containing only mRNA sequences, sRNA/intergenic sequences/miscellaneous RNA sequences, or the tRNA sequences, were separately used as input for motif identification. When all thirty-four hybrid sequences were examined together, no significant conserved sequence binding motifs were identified. However, when the tRNA hybrid sequences alone were used to interrogate the MEME tool, one predicted binding motif containing 21 nucleotides was identified in nine out of ten of the tRNA sequences (Figure 4.16).

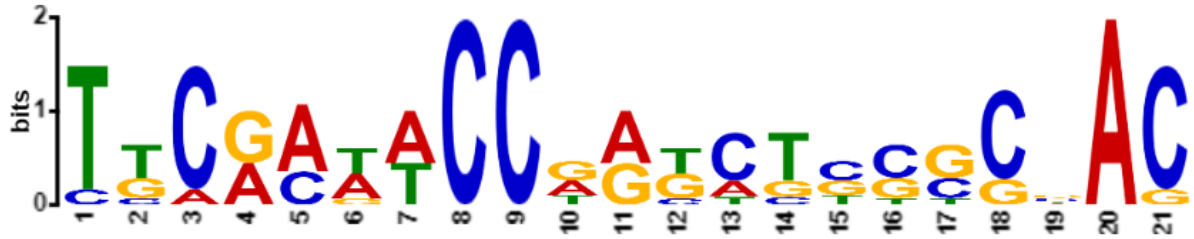


Figure 4.16. The predicted conserved sequence binding motif from MEME, derived from the tRNA sequences in hybrid RNAs associated with Hfq using UV-CLASH. The predicted motif identified had an e-value of 7.7E-011.

4.4. Discussion

Hfq Co-IP experimental analyses were performed in an attempt to identify the range of transcripts that associate with Hfq, including both interacting sRNAs and target mRNAs. Importantly, the analyses focussed on the putative sRNAs discovered from bioinformatic analysis of whole-genome RNA transcripts (Chapter 3), examining their interaction with Hfq and other RNA molecules. Two different Hfq Co-IP techniques were performed to maximise the identification of Hfq-associated RNA in *P. multocida*; Hfq Co-IP assays using anti-FLAG magnetic beads as well as Hfq UV-CRAC/CLASH experiments. Our results are the first genome-wide experimental analyses performed for the identification of Hfq-dependent sRNAs and target mRNA molecules in *P. multocida*. The Co-IP experiments were performed using *hfq* mutant strains that expressed either tagged Hfq or untagged Hfq (negative control).

The first Hfq Co-IP method employed anti-FLAG magnetic beads to co-immunoprecipitate RNA species that associated with a chromosomally encoded 3xFLAG-tagged-Hfq protein, which was expressed by *P. multocida* lacking the wild-type *hfq* gene. The method used was a modification of previous Hfq Co-IP experiments performed in *E. coli* that identified almost 300 Hfq-associated mRNAs and a range of Hfq-associated sRNAs, including many novel sRNAs (Bilusic et al., 2014). Many of the Hfq-associated RNAs identified using Co-IP in *E. coli* were shown to be Hfq regulated using deep sequencing and northern blot analysis (Bilusic et al., 2014).

Currently, GcvB is the only sRNA that has been experimentally characterised in *P. multocida* (Gulliver et al., 2018). The *P. multocida* GcvB sRNA regulates amino acid biosynthesis and transport processes in *P. multocida* and more than 20 putative targets have been identified using transcriptomics and proteomics approaches (Gulliver et al., 2018). GcvB target binding has been shown to be *hfq*-dependent in other bacterial species (Sharma et al., 2011). Therefore, we used enrichment of GcvB as a measure of the specificity of the Co-IP experiments. GcvB, was strongly enriched by Hfq Co-IP (~11.1-fold), indicating successful co-immunoprecipitation of at least one Hfq-dependent RNA species. Interestingly, four predicted GcvB target RNAs also displayed some association with Hfq; the hypothetical protein X73_RS07870, a peptide methionine sulfoxide reductase (MsrA), a periplasmic dipeptide transport protein precursor (DppA) and the dihydrodipicolinate synthase protein DapA. The production of the DapA protein was increased in *hfq* and *gcvB* mutant strains (Gulliver et al., 2018; Mégroz et al., 2016).

It therefore possible that Hfq may be acting to facilitate the degradation of DapA via GcvB targeting.

Apart from GcvB, only two other predicted *P. multocida* putative sRNAs; namely, Prrc10 and Prrc19 also showed enrichment in the Hfq Co-IP data. Although it was predicted that the Hfq Co-IP experiments would identify novel *P. multocida* sRNAs, specifically *cis*-encoded sRNAs, which were more difficult to identify from the bioinformatic studies reported in Chapter 3, no novel putative sRNAs were identified from the analysis. Bacterial sRNA expression can be dependent on the growth phase of the bacterium, and/or on different growth conditions, so that individual sRNAs may only be expressed under very specific conditions (Kröger et al., 2013; Stubben et al., 2014). Growth in rich medium (HI broth) at 37°C was the only growth condition used in our Hfq Co-IP experiment and it is possible that many sRNAs may not have been expressed and therefore not identified as Hfq-associated. However, these data may also suggest that the prediction of sRNAs, based on bioinformatic analysis of RNA-Seq data, may be quite comprehensive.

More than 80 mRNAs were enriched in the Hfq Co-IP data; including some that displayed altered regulation in the previous analysis of the *hfq* mutant strain using RNA-Seq (Chapter 2) (Mégroz et al., 2016). As RNA-Seq only measures final average transcript levels, gene expression changes observed in the *hfq* mutant cannot be directly ascribed to the action of Hfq. However, interaction with Hfq (as measured by Co-IP) gives a strong indication that Hfq facilitates the regulation of the mRNA via a direct mechanism. Gene transcripts encoding the outer membrane lipoprotein PlpB (X73_RS01475), a predicted membrane-bound lysozyme-inhibitor (X73_RS08800), PcgD, a phosphocholine (PCho) transferase (X73_RS04500) and the preprotein translocase subunit SecG (X73_RS09230) each displayed a clear association with Hfq by Co-IP. Each of these genes displayed increased expression in the *hfq* mutant as measured by RNA-Seq; PlpB and hypothetical protein X73_RS05335 also displayed increased protein production during both early- and mid-exponential growth in the absence of *hfq* (Mégroz et al., 2016). Therefore, these data suggest a potential role for Hfq in the destabilization/degradation of these transcripts either with or without the action of associated sRNAs. Interestingly, the *hfq* mRNA transcript was enriched in the Hfq Co-IP analysis (14.5-fold), indicating that the *hfq* mRNA could be interacting with the Hfq protein itself. This may either be a direct result of Hfq self-regulation or that during normal translation of the *hfq* mRNA, nascent Hfq polypeptides are associated with mRNA.

MACS2 analysis (Zhang et al., 2008) was also used to identify Hfq-associated transcripts from the Hfq Co-IP data. From this analysis, the majority (>95%) of RNA-Seq reads enriched in the *P. multocida* Hfq Co-IP samples mapped to rRNA, RNase P and 6S/SsrS RNA genes, rather than sRNA and mRNA molecules. In contrast, a number of other studies performed in *E. coli*, *Sinorhizobium meliloti and Brucella suis* have shown a lack of association between Hfq and tmRNA, 6S RNA and RNase P transcripts (Saadeh et al., 2016; Torres-Quesada et al., 2014; Zhang et al., 2003). In some other Co-IP studies ribosomal depletion was performed to overcome the abundance of rRNA (Bilusic et al., 2014). We did not perform this step as we thought it unnecessary given the high specificity of Co-IP using Hfq. However, it is possible that the reads generated from rRNA or other RNA species are due to their sheer abundance in the cell rather than these RNA species having a true association with Hfq. Future experiments should aim to use Hfq Co-IP with an additional rRNA depletion step. As discussed later in this section, direct binding assays could be performed to confirm specific interactions of each of these RNA species with *P. multocida* Hfq. It is possible that these interactions are non-specific and have been influenced by the abundance of these RNA species. Optimisation of Hfq binding conditions throughout the Hfq Co-IP experiments, such as modification of various buffer components, could be performed to identify whether altered experimental conditions resulted in altered interaction with these species.

The second Hfq Co-IP technique reported in this study was Hfq UV-CRAC/CLASH. As with the first technique, these Co-IP analyses were performed using *hfq* mutant strains that expressed either tagged or untagged Hfq. However, for these experiments the genes encoding tagged or untagged Hfq were expressed from a replicating plasmid rather than from the chromosome. Similar UV-CRAC/CLASH experiments have previously been used in *E. coli* to identify RNA species associated with RNase E (Waters et al., 2017) and Hfq (Iosub et al., 2018). This technique uses Co-IP and *in vivo* UV-crosslinking to crosslink and recover RNA species that are bound to Hfq, allowing for the identification of specific RNA-RNA hybrids; including sRNA:mRNA hybrids. Only a single growth condition was used for *P. multocida* (growth in rich medium at 37°C), limiting the identification of Hfq-associated sRNAs to only those expressed under this specific condition. Also, only two biological replicates of each strain were used in these experiments; a third replicate may have increased the significance of the data. Additionally, expression of HTF-tagged Hfq from the *P. multocida* chromosome rather than from a plasmid may have resulted in an expression of HTF-tagged Hfq that correlated better with the levels of Hfq naturally occurring in wild-type *P. multocida*. Future experiments

involving Hfq UV-CRAC/CLASH could be further optimised by using rRNA depletion steps to overcome the abundance of rRNA reads.

As performed with the Co-IP experiments, visual inspection of the mapped UV-CRAC RNA-Seq reads revealed high numbers of reads present for the sRNA GcvB in the tagged Hfq samples but not the untagged Hfq samples, confirming that GcvB activity is Hfq-dependent in *P. multocida*. Three *P. multocida* genomic regions were significantly associated with Hfq, as determined by UV-CRAC analysis; namely *glmU*, *ner* and an intergenic region (upstream of PMVP_0398). However, none of these regions were also identified as associating with Hfq in the Hfq Co-IP experiments.

Hfq UV-CRAC showed that Hfq displayed a significant association with *glmU*, known to be involved in the hyaluronic acid (HA) capsule biosynthesis and cell wall and membrane biogenesis pathways in other bacteria (de Oliveira et al., 2016). This correlates with the finding that capsule biosynthesis was strongly reduced in a *P. multocida hfq* mutant (Chapter 2). We therefore hypothesized that Hfq may act to promote GlmU translation, thereby increasing *P. multocida* capsule production. To test this hypothesis, *glmU* was overexpressed in the *hfq* mutant strain to determine if HA capsule production levels could be restored in the presence of an abundance of GlmU. However, *glmU* overexpression did not result in the restoration of capsule expression in the *hfq* mutant. It should be noted that TraDIS analysis suggests that *glmU* is an essential gene in *P. multocida* (Smallman, Harper, Boyce, unpublished), and therefore mutagenesis could not be used as a tool for *glmU* characterisation in *P. multocida*. Furthermore, expression of *glmU* was unaffected in the *hfq* mutant as measured by transcriptomics (Mégroz et al., 2016) and proteomics (Chapter 2). Thus, the predicted interaction between Hfq and *glmU* as indicated by Hfq UV-CRAC remains unvalidated.

As overexpression of GlmU did not restore capsule production in the *hfq* mutant we predict that the effect of Hfq on capsule production is likely via another pathway. In *E. coli*, *glmU* is encoded as a poly-cistronic message with *glmS*, a glucosamine-6-phosphate synthase enzyme (Salim et al., 2012). Interestingly, following cleavage of the *glmU* and *glmS* transcript by RNaseE, Hfq has been shown to activate *glmS* translation via the GlmZ sRNA (Salim et al., 2012). However, in *P. multocida*, *glmU* and *glmS* are not encoded as a poly-cistronic message and there is no known regulatory relationship with Hfq. The Co-IP methods used in this study did not demonstrate an association between Hfq and any gene transcripts from the *P. multocida* capsule biosynthetic locus. As capsule is an important virulence factor for *P. multocida* and

Hfq is known to play a key role in capsule regulation, other pathways of Hfq-mediated capsule regulation via sRNAs will be investigated in future work.

Another region that associated with Hfq using UV-CRAC, based on a significant number of reads, was *ner* (PmVP161_0399), a predicted *P. multocida* transcriptional regulator. A third region that was intergenically encoded also showed a significant association with Hfq using UV-CRAC. This intergenic region is located upstream of a predicted *P. multocida* protein of unknown function (PMVP_0398). Interestingly, the investigation of *P. multocida* RNA expression under low iron and reduced oxygen conditions (Chapter 3) revealed a low level of the sRNA-enriched RNA-Seq reads mapped to this same region. However, the limited expression, across only a few of the replicates, does not alone indicate that there is a novel sRNA transcript in this region. However further experimental analysis, including northern blotting and primer extension analysis could be used to verify the relationship in future work.

Hfq mediates interactions between RNA species and there is evidence for rapid exchange of RNAs on Hfq (Santiago-Frangos and Woodson, 2018; Wagner, 2013). *In vitro* studies have showed long-lived interactions between Hfq and some RNA substrates ($t_{½}$ ~ 100 min); however, there is only a short response time (1-2 min) between induction of an sRNA to observable effects on target gene expression (Fender et al., 2010; Massé et al., 2003; Papenfort et al., 2006; Wagner, 2013). This can be explained by the ability of sRNAs to rapidly cycle on and off Hfq, displacing Hfq-bound RNA molecules and allowing multiple RNAs to associate with Hfq simultaneously via unoccupied subunits (Wagner, 2013). Therefore, it is likely that only a snapshot of the Hfq-interactome can potentially be captured at any point in time using the Hfq Co-IP methods used in this study. Thus, data obtained from these methods is unlikely to be representative of the entire Hfq regulon. While there were some Hfq-target interactions observed using both Co-IP techniques, such as the significant enrichment with the GcvB sRNA, it was not expected that all Hfq interactions would be detected using both methods. Importantly, not all sRNAs have Hfq-dependent activity and many may act independently of Hfq or may require other RNA chaperones, such as ProQ, to mediate sRNA regulation (Smirnov et al., 2016). Therefore, other experimental methods should also be applied to investigate enrichment of specific sRNAs from the Co-IP analyses with other proteins. Additionally, qRT-PCR and/or northern blot analysis could also be performed in future work for the detection of specific sRNAs from whole recovered tagged and untagged RNA samples, using primers or probes that are specific for selected putative sRNAs.

Hfq UV-CLASH allowed for the identification of seventeen RNA:RNA hybrid molecules, indicating that these molecules interact with Hfq at the same time. These involved interactions with intergenic regions, mRNAs, tRNAs, putative sRNAs and miscellaneous RNA sequences. Interactions included one putative sRNA, Prrc03 which was predicted to associate with the tRNA Ser, and may therefore be involved in the post-transcriptional regulation of this tRNA. Previous Co-IP studies have demonstrated an association of Hfq with tRNA transcripts, suggesting that Hfq may play a role in tRNA processing (Bilusic et al., 2014; Lee and Feig, 2008; Zhang et al., 2003). A more recent study in *E. coli* has demonstrated a direct association of sRNAs RyhB and RybB with tRNA precursors that led to a reduction in the level of RyhB (Lalaouna et al., 2015). Therefore it is possible that Prrc03 is directly involved in the maturation of the tRNA Ser in *P. multocida*, or that the tRNA is acting as a sponge for the Prrc03 sRNA. As the hybrid molecules identified from the UV-CLASH may be the result of spurious ligation products, it will be important to repeat these experiments and to confirm these specific RNA:RNA interactions using future experimental analyses.

The Hfq protein functions as a homo-hexameric ring, with four regions available for interaction with RNAs, each with different binding architectures and preferences. These four regions are the proximal and distal faces, the rim and the C-terminal tail (Sobrero and Valverde, 2012). As there are no known *P. multocida* Hfq-specific binding sequences, the MEME motif identification tool (Bailey et al., 2015) was utilised to determine if a specific Hfq binding motif could be identified on the associated RNAs using information from the hybrid sequences. All thirty-four RNA hybrid sequences were initially used as the analysis set; however, no significant binding motifs were identified. RNA species were then grouped into mRNA sequences, sRNA/intergenic sequences/miscellaneous RNA sequences, and tRNAs then separately analysed using the MEME tool. One predicted binding motif consisting of 21 nucleotides was common to nine out of ten of the tRNA sequences identified in RNA hybrids. In *E. coli*, Hfq is known to interact at two independent sites on tRNAs and may play a role in tRNA metabolism (Lee and Feig, 2008). However, specific binding analyses would need to be conducted to determine if the predicted motif identified in *P. multocida* tRNAs is indeed required for association with Hfq.

To confirm the interactions between Hfq and the RNAs identified using the two different Co-IP analyses, further binding experiments should be performed. These would involve techniques such as biolayer interferometry (BLItz), electrophoretic mobility shift assay (EMSA) or surface

plasmon resonance (SPR) (Dimastrogiovanni et al., 2014; Fender et al., 2010; Osborne et al., 2014) to determine if each RNA interacts with Hfq. These analyses could also be applied to investigate the specific interactions between RNA species in the hybrid RNA:RNA molecules identified using the UV-CLASH experiments (Morita et al., 2012).

This study investigated the *P. multocida* Hfq-interactome using multiple Hfq Co-IP techniques. A number of putative sRNAs associated with Hfq, including the well characterised sRNA GcvB. A large number of other mRNAs and other RNA species also associated with Hfq, some which showed altered RNA expression in *hfq* mutant experiments (Chapter 2) and which may directly be regulated by Hfq activity. These included the transcript of the PlpB protein, an outer membrane lipoprotein. To identify the complete Hfq regulon, additional Co-IP experiments under different growth conditions could be performed. Seventeen RNA:RNA hybrids were also identified in association with Hfq, including one sRNA associating with a tRNA. Moreover, a potential tRNA-specific, conserved Hfq binding motif was identified from the tRNA hybrid sequences. Further binding analyses would need to be performed to confirm specific Hfq-RNA interactions for the putative interactions identified in these analyses. The next chapter will report on the experimental validation of a number of the putative *P. multocida* sRNAs identified from these studies. The *P. multocida* Hfq Co-IP studies did not identify any direct associations between Hfq and the transcripts of virulence-associated genes whose transcripts were affected by the loss of Hfq such as capsule and filamentous haemmaglutinin biosynthesis genes (Mégroz et al., 2016). It will be of particular importance to further elucidate the Hfq-dependent regulatory pathways that control these virulence-associated transcripts via sRNAs or other Hfq-associated regulatory pathways in future work.

Chapter Five: Experimental validation and characterisation of *P. multocida* sRNAs

5.1. Introduction

In our initial study, we showed that the *P. multocida* Hfq riboregulatory chaperone protein played an important role in the regulation of *P. multocida* gene expression and protein production (Chapter 2). This included the regulation of important virulence factors such as capsule, LPS and filamentous hemagglutinin (Mégroz et al., 2016). Given the role of Hfq in other bacteria, it was predicted that the *P. multocida* Hfq facilitated the post-transcriptional regulation of important virulence factors via the action of sRNAs. Thus, further analyses were performed to identify the range of sRNAs produced by *P. multocida* using bioinformatic and experimental analyses.

Firstly, bioinformatic analyses of whole genome RNA transcripts from *P. multocida* grown aerobically in rich medium, under low iron, or under reduced oxygen conditions, revealed 38 novel putative *P. multocida* sRNAs (Chapter 3). Hfq co-immunoprecipitation (Co-IP) experimental analyses were then performed to identify the range of Hfq-associated sRNA and mRNA species (Chapter 4). Importantly, a number of the sRNA species identified from bioinformatic analyses were shown to be associated with Hfq. The sRNA GcvB was identified from the initial bioinformatic analysis of RNA transcripts, and Co-IP analyses showed that GcvB has a direct association with Hfq. Subsequent studies showed that the *P. multocida* GcvB has a clear role in amino acid biosynthesis and transport (Gulliver et al., 2018). However, none of the other putative sRNAs have been experimentally validated or characterised.

In this chapter, we report on several putative *P. multocida* sRNAs which were chosen for further investigation. Experimental validation of the expression and sequence length of the selected putative sRNAs were investigated using fluorescent primer extension and northern blotting. The experimental validation of transcripts correlating to each predicted sRNA is the first step to determine if they are *bone fide* sRNA species. Initially, fluorescent primer extension analysis was used to identify the precise transcript start site for each selected sRNA. Northern blotting analysis was also performed to validate the size and to determine the expression of the selected sRNA transcripts. To investigate the regulatory role of each selected sRNA in *P. multocida*, each sRNA sequence was separately disrupted using TargeTron mutagenesis. TargeTron is a powerful tool for *P. multocida* mutagenesis, as double crossover allelic exchange mutagenesis has very low success rate in this species. However, as TargeTron only results in an insertion, it may not completely inactivate the activity of an sRNA if the seed region, required for binding to target RNA species, is still intact. To help determine the role of

each sRNA under investigation, *P. multocida* strains overexpressing the sRNA species or expressing the appropriate antisense RNA were also generated. The phenotype of the sRNA strains was then assessed with focus on any altered virulence phenotypes, specifically capsule production. Using these techniques, this study aimed to identify and characterise Hfq-associated sRNA species involved in the regulation of virulence factors such as capsule.

5.2. Materials and Methods

5.2.1. Bacterial strain morphology, plasmids and culture conditions.

The bacterial strains and plasmids used in this study are listed in Table 5.1. *P. multocida* and *E. coli* strains were grown as described previously (Mégroz et al., 2016). When required, media were supplemented with 50 µg/mL kanamycin and/or 50 µg/mL spectinomycin. To visually compare colony morphologies of different *P. multocida* strains, they were streaked onto different sections of a single HI agar plate and incubated at 37°C for 24 h.

5.2.2. DNA manipulations.

The following DNA manipulations were performed as described previously (Mégroz et al., 2016); restriction digests, ligations and PCR amplifications, preparation of plasmid and genomic DNA, purification of DNA, quantification of DNA and Sanger sequencing of DNA. The primers used in this study are listed in Table 5.2.

5.2.3. Transformation of *E. coli* and *P. multocida*

Transformation of *E. coli* DH5α $RbCl_2$ competent cells and preparation and transformation of electrocompetent *P. multocida* cells were performed as described in Chapter 4 (section 4.2.4).

5.2.5. Bacterial growth curve of sRNA strains: growth in rich medium

To determine the growth rate of each strain the following experiment was performed in biological triplicate. A 10 mL HI broth was inoculated with 3-4 large *P. multocida* colonies and incubated overnight at 37°C with constant shaking at 200 rpm. To maintain plasmids in specific strains, the appropriate antibiotic was added to the growth media. The overnight culture was then subcultured at a 1:50 dilution into 5 mL HI broth (without addition of antibiotic) and incubated at 37°C with constant shaking at 200 rpm until the culture reached an optical density at 600 nm (OD_{600}) of between 0.2 and 0.3. Each culture was standardised to OD_{600} = 0.2 via the addition of the appropriate amount of sterile HI broth then diluted 1:100. Aliquots (200 µL) of the diluted culture were then added to two or three wells of a single row (technical duplicate or triplicate) of an optically clear, flat-bottomed 96-well plate. The plate was then placed in an Infinite M200 plate reader (Tecan) and incubated at 37°C with constant shaking (180 rpm). Optical density (600 nm) measurements of each well was taken at until late-exponential growth phase was reached.

Table 5.1. Bacterial strains and plasmids used in this study.

Strain or plasmid	Relevant description	Source or reference
Strains		
E. coli strains		
DH5α	*deoR endA1 gyrA96 hsdR17*(r_K^- m_K^+) *recA1 relA1 supE44 thi-1* (*lacZYA-argFV169*) φ80*lacZ*ΔM15, F^-	Bethseda Research Laboratories
AL177	DH5α harbouring pAL99; Kan^R	(Harper et al., 2004)
AL2708	DH5α harbouring pREXY; $Spec^R$	(Gulliver et al., 2018)
AL3147	DH5α harbouring pAL1368; Kan^R, $Spec^R$	This study
AL3198	DH5α harbouring pAL1380; $Spec^R$	This study
AL3208	DH5α harbouring pAL1391; Kan^R, $Spec^R$	This study
AL3210	DH5α harbouring pAL1393; Kan^R, $Spec^R$	This study
AL3211	DH5α harbouring pAL1394; $Spec^R$	This study
AL3212	DH5α harbouring pAL1395; $Spec^R$	This study
AL3213	DH5α harbouring pAL1396; $Spec^R$	This study
AL3278	DH5α harbouring pAL1432; $Spec^R$	This study
AL3279	DH5α harbouring pAL1433; $Spec^R$	This study
AL3280	DH5α harbouring pAL1434; $Spec^R$	This study
P. multocida strains		
VP161	Serotype A:1: Vietnamese isolate from chickens	(Wilkie et al., 2000)
AL3188	VP161 Prrc08 TargeTron mutant; Kan^R	This study
AL3272	VP161 Prrc06 TargeTron mutant; Kan^R	This study
AL3275	VP161 Prrc25 TargeTron mutant; Kan^R	This study
AL3282	VP161 Prrc06 overexpression strain; VP161 containing pAL1394; $Spec^R$	This study

AL3283	VP161 Prrc08 overexpression strain; VP161 containing pAL1395; $Spec^R$	This study
AL3284	VP161 Prrc25 overexpression strain; VP161 containing pAL1396; $Spec^R$	This study
AL3285	VP161 Prrc06 antisense strain; VP161 containing pAL1432; $Spec^R$	This study
AL3286	VP161 Prrc08 antisense strain; VP161 containing pAL1433; $Spec^R$	This study
AL3287	VP161 Prrc25 antisense strain; VP161 containing pAL1434; $Spec^R$	This study
AL3306	VP161 Prrc06 TargeTron mutant containing pAL1394; Kan^R, $Spec^R$	This study
AL3307	VP161 Prrc08 TargeTron mutant containing pAL1395; Kan^R, $Spec^R$	This study
AL3308	VP161 Prrc25 TargeTron mutant containing pAL1396; Kan^R, $Spec^R$	This study
Plasmids		
pAL99	*P. multocida/E. coli* expression plasmid, contains the *P. multocida* constitutive *tpiA* promoter upstream of multiple cloning site (MCS); Kan^R	(Harper et al., 2004)
pAL1368	*P. multocida* TargeTron plasmid used for inactivation of Prrc08 in VP161; Kan^R, $Spec^R$	This study
pAL1380	*P. multocida* antisense sRNA expression plasmid; pREXY with *tpiA* terminator sequence cloned into EcoRI site of multiple cloning region; only a single (5′) EcoRI site is restored; $Spec^R$	
pAL1391	*P. multocida* TargeTron plasmid used for inactivation of Prrc06; Kan^R, $Spec^R$	This study
pAL1393	*P. multocida* TargeTron plasmid used for inactivation of Prrc25; Kan^R, $Spec^R$	This study
pAL1394	Prrc06 overexpression construct; pREXY with Prrc06 cloned into BamHI and EcoRI sites; $Spec^R$	This study
pAL1395	Prrc08 overexpression construct; pREXY with Prrc08 cloned into BamHI and EcoRI sites; $Spec^R$	This study
pAL1396	Prrc25 overexpression construct; pREXY with Prrc25 sequence cloned into BamHI and SmaI sites; $Spec^R$	This study

pAL1432	Prrc06 antisense construct; pAL1380 with Prrc06 antisense sequence cloned into BamHI and EcoRI sites; $Spec^R$	This study
pAL1433	Prrc08 antisense construct; pAL1380 with Prrc08 antisense sequence cloned into BamHI and EcoRI sites; $Spec^R$	This study
pAL1434	Prrc25 antisense construct; pAL1380 with Prrc25 antisense sequence cloned into SmaI and BamHI sites; $Spec^R$	This study
pREXY	*P. multocida/E.coli* sRNA expression plasmid. Recombinant gene expression driven by the constitutive *P. multocida* P_{tpi} promoter located upstream of multiple cloning site; $Spec^R$	(Gulliver et al., 2018)

Table 5.2. Primers used in this study.

Primer	Sequence (5′ – 3′)	Description
BAP612 (Universal)	GTAAAACGACGGCCAGT	Forward primer located upstream of the of the *tpi* promoter and MCS in pREXY
BAP2679	TTGTGTGGAATTGTGAGCGGA	Reverse primer located downstream of the MCS in pREXY
BAP6797	TTGGGACGCGTTTGGTCGGTCATTTCG	Reverse primer used to PCR-amplify *aph3* (kanamycin resistance gene) from pAL99. Contains flanking MluI sites
BAP8092	[6-FAM]-GAGAACGCCAATACAACAATTAAACC	Reverse primer located within VP161 Prrc06 (anneals ~68 bp downstream of the predicted transcript start). Specific for fluorescent primer extension; contains a 5′-6-fluorescein amidite (6-FAM) label
BAP8093	[6-FAM]-TAGCACTTATTTTTAGCTAGACACAGAAT	Reverse primer located within VP161 Prrc07 (anneals ~101 bp downstream of predicted transcript start). Specific for fluorescent primer extension; contains a 5′-6-fluorescein amidite (6-FAM) label
BAP8094	[6-FAM]-AGTAACCGTATTTCAGGCTATTCTC	Reverse primer located within VP161 Prrc10 (anneals ~57 bp downstream of predicted transcript start). For fluorescent primer extension; contains a 5′-6-fluorescein amidite (6-FAM) label
BAP8095	[6-FAM]-CGTTGACTATTGGCATGAGCTG	Reverse primer located within VP161 Prrc12 (anneals ~94 bp downstream of predicted transcript start). For fluorescent primer extension; contains a 5′-6-fluorescein amidite (6-FAM) label
BAP8224	GTAGAGGTAACAATGAGTGAATTTGG	Forward internal VP161 Prrc06 primer for generation of northern blot probe; anneals ~11 bp downstream of start codon
BAP8225	TAATACGACTCACTATAGGGTTAGTCTTAAAAAATTGAGAACGCCAATAC	Reverse internal VP161 Prrc06 primer for generation of northern blot probe; anneals ~50 bp upstream of 3′ end of Prrc06. Contains a 5′ T7 RNA polymerase promoter sequence (20 bp)

BAP8234	[6-FAM]-GCGCGCTATTTGAATAAATTTGGG	Reverse primer located within VP161 Prrc08 (~132 bp downstream of predicted transcript start). For fluorescent primer extension; contains a 5′-6-fluorescein amidite (6-FAM) label
BAP8235	[6-FAM]-AACGATCCCTGTTGCGATTTTAG	Reverse primer located within VP161 Prrc25 (~116 bp downstream of predicted transcript start). For fluorescent primer extension; contains a 5′-6-fluorescein amidite (6-FAM) label
BAP8236	TTATCTTATAACGTTGGTGTGAACTATG	Forward primer located within VP161 Prrc08 (~44 bp downstream of predicted transcript start). For generation of a northern blot probe
BAP8237	TAATACGACTCACTATAGGGGCGCGCTATTTGAATAAATTTGGG	Reverse primer located within VP161 Prrc08 (anneals ~27 bp upstream of 3′ end). For generation of northern blot probe. Contains a 5′ T7 RNA polymerase promoter sequence (20 bp)
BAP8238	ATTCTTTTTAAATTTAATCTTGCAATAGGATT	Forward primer located within VP161 Prrc25 (anneals ~7 bp downstream of start codon). For generation of a northern blot probe
BAP8239	TAATACGACTCACTATAGGGAACGATCCCTGTTGCGATTTTAG	Reverse primer located within VP161 Prrc25 primer (anneals ~18 bp upstream of 3′ end). For generation of a northern blot probe. Contains a 5′ T7 RNA polymerase promoter sequence (20 bp)
BAP8288	AAAAAAGCTTATAATTATCCTTAGTAGACGGTGCCGTGCGCCCAGATAGGGTG	IBS primer specific for retargeting TargeTron intron to inactivate Prrc08
BAP8289	CAGATTGTACAAATGTGGTGATAACAGATAAGTCGGTGCCGCTAACTTACCTTTCTTTGT	EBS1d primer specific for retargeting TargeTron intron to inactivate Prrc08
BAP8290	TGAACGCAAGTTTCTAATTTCGATTTCTACTCGATAGAGGAAAGTGTCT	EBS2 primer specific for retargeting TargeTron intron to inactivate Prrc08
BAP8326	ATGGAAGGATCCTTGATTAATCCGTAG	Forward primer located upstream of VP161 Prrc06; contains BamHI site
BAP8327	ATTACCGAATTCTAAAATCCGCAC	Reverse primer located downstream of Prrc06; contains EcoRI site

BAP8328	TTAATGGGATCCTTGCCGCAAAAGC	Forward primer located upstream of VP161 Prrc08; contains BamHI site
BAP8329	CTGCTGGAATTCCAATATGAGATATT	Reverse primer located downstream of Prrc08; contains EcoRI site
BAP8330	GTAGGGGGATCCTAGTAGAATTCTT	Forward primer located upstream of VP161 Prrc25; contains BamHI site
BAP8331	CAAATGCCCGGGTGAGATTAAAT	Reverse primer located downstream of Prrc25; contains SmaI site
BAP8353	AATTCGCAGATCAATGATCTGCTTTTTT TC	Encodes *tpiA* terminator sequence; used to anneal with BAP8354 to produce a double stranded fragment for cloning; has an EcoRI-specific 5′ end and an EcoRI compatible 3′ end
BAP8354	AATTGAAAAAAAGCAGATCATTGATCT GCG	Complementary sequence to *tpiA* terminator sequence; used to anneal with BAP8353 to produce a double stranded fragment; has an EcoRI-specific 3′ end and an EcoRI compatible 5′ end
BAP8414	AAAAAAGCTTATAATTATCCTTACTAGG CTTAATTGTGCGCCCAGATAGGGTG	IBS primer specific for retargeting TargeTron intron to inactivate Prrc06
BAP8415	CAGATTGTACAAATGTGGTGATAACAG ATAAGTCTTAATTGTTAACTTACCTTTCT TTGT	EBS1d primer specific for retargeting TargeTron intron to inactivate Prrc06
BAP8416	TGAACGCAAGTTTCTAATTTCGGTTCCT AGTCGATAGAGGAAAGTGTCT	EBS2 primer specific for retargeting TargeTron intron to inactivate Prrc06
BAP8417	AAAAAAGCTTATAATTATCCTTAATAGG CTTTTGCGTGCGCCCAGATAGGGTG	IBS primer specific for retargeting TargeTron intron to inactivate Prrc25
BAP8418	CAGATTGTACAAATGTGGTGATAACAG ATAAGTCTTTTGCATTAACTTACCTTTCT TTGT	EBS1d primer specific for retargeting TargeTron intron to inactivate Prrc25
BAP8419	TGAACGCAAGTTTCTAATTTCGGTTCCT ATCCGATAGAGGAAAGTGTCT	EBS2 primer specific for retargeting TargeTron intron to inactivate Prrc25
BAP8420	ATAAAGTCACCGCACGAGCACC	Forward primer, anneals ~1069 bp upstream of VP161 Prrc06

BAP8421	TTAAAACAGTACGTGGTGCTGCG	Forward primer, anneals ~1153 bp upstream of VP161 Prrc25
BAP8450	ATGGAAGAATTCTTGATTAATCCGTAG	Primer for construction of antisense expression plasmid, binds to Prrc06 nucleotides 1 to 15; contains EcoRI site for integration close to the 5′ end *tpiA* terminator in vector pAL1380
BAP8451	TCGTTAGGATCCTTTTGTTTAGTCTTAA	Primer for construction of antisense expression plasmid, binds to opposite strand relative to Prrc06 nucleotides 101 to 128; contains BamHI site for integration close to the 3′ end *tpiA* promoter in vector pAL1380
BAP8452	TTAATGGAATTCTTGCCGCAAAAGC	Primer for construction of antisense expression plasmid, binds to Prrc08 nucleotides 1 to 13; contains EcoRI site for integration close to the 5′ end *tpiA* terminator in vector pAL1380
BAP8453	GAATAAGGATCCGATAGTTTAGAGGTA A	Primer for construction of antisense expression plasmid, binds to opposite strand relative to Prrc08 nucleotides 118 to 145; contains BamHI site for integration close to the 3′ end *tpiA* promoter in vector pAL1380
BAP8454	GTAGGGCCCGGGTAGTAGAATT	Primer for construction of antisense expression plasmid, binds to Prrc25 nucleotides 1 to 10; contains SmaI site for integration close to the 5′ end *tpiA* terminator in vector pAL1380
BAP8455	TTCATTGGATCCTCCTTAAACTTCATAC	Primer for construction of antisense expression plasmid, binds to opposite strand relative to Prrc25 nucleotides 84 to 111; contains BamHI site for integration close to the 3′ end *tpiA* promoter in vector pAL1380
EBS Universal	CGAAATTAGAAACTTGCGTTCAGTAAAC	Universal TargeTron primer, reverse primer anneals ~250 bp from the 5′ end of the intron and immediately downstream of the retarget region

5.2.7. Extraction and quantification of RNA

RNA extraction was performed as described in Materials and Methods of Chapter 3 (section 3.2.2). Quantification and analysis of RNA samples for DNA contamination was performed as described previously (Mégroz et al., 2016).

5.2.8. Fluorescent primer extension analysis of sRNAs

Fluorescent primer extension analysis of RNA was performed as previously described (Gulliver et al., 2018) with the following modifications. RNA was extracted from *P. multocida* VP161 strains grown to early and mid-exponential growth phase i.e. OD_{600} of 0.2 and 0.7. The synthesis of cDNA for each gene examined was performed with the appropriate 6-carboxy fluorescein amidite (6-FAM) labelled primers (Table 5.2). For both reverse transcription reaction steps, 100 U of Super Script III reverse transcriptase (Invitrogen) was used with the addition of 0.5 µL 0.1 M Dithiothreitol (DTT). Samples were sent to the Australian Genome Research Facility (AGRF, Melbourne) for primer extension analysis using an ABI 3730xl DNA Analyzer (Thermo Fisher Scientific).

5.2.9. Northern blotting

DIG-labelled riboprobes specific to target RNA transcripts were generated using *in vitro* transcription reactions. DNA was amplified from *P. multocida* genomic DNA using primers specific for each sRNA sequence; the reverse primer in each reaction encoded a T7 RNA polymerase promoter sequence at the 5′ end (Table 5.2). To prepare enough template, multiple PCR reactions were performed then pooled and purified using phenol:chloroform extraction procedure as described in Chapter 3, section 3.2.2. The final purified DNA template (~2.5 µg) was resuspended in 25 µL nuclease free H_2O.

In vitro transcription for DIG labelling of riboprobes was performed in 20 µL of 1 x transcription optimised buffer (Promega) containing 10 mM DTT, 1 x DIG labelling RNA mix (Roche), 20 units of Protector RNase inhibitor (Merck), ~250 ng of purified DNA template and 20 units of T7 RNA polymerase (Promega). Following incubation at 37°C for 2 h the reaction was stopped with the addition of 2 µL of 0.2 M EDTA (pH 8.0) then the RNA was precipitated via the addition of 1.25 µL 8M LiCl and 75 µL of 100% EtOH. After incubation for 1 h at -80°C, the samples were centrifuged at 12, 000 x *g* (4°C) for 25 min. Following removal of supernatant, the pellets were washed in 300 µL chilled (4°C) 80% EtOH, centrifuged for 5 min at 12, 000 x *g* (4°C) then the 80% EtOH was removed and the pellets air dried. The RNA was

resuspended in 100 µL of nuclease-free H_2O containing 1 µL of Protector RNase Inhibitor (Merck). Prior to northern blotting, the generated riboprobes were tested for efficiency of DIG labelling using a dot blot assay, according to the manufacturer's instructions (Roche).

For northern blotting, a total of 12 µg of each RNA sample, extracted from *P. multocida* VP161 strains grown to mid-exponential growth phase, was separated using a high percentage, denaturing polyacrylamide-urea (12.5% acrylamide, 8 M urea) gel as previously described (Rio et al., 2010) with the exception that a Bio-Rad Mini-PROTEAN® II apparatus was used instead of a large gel format. A low range ssRNA ladder (New England Biolabs) was also electrophoresed to use as a reference. Following electrophoresis at a constant current of 25 mA for 2-3 hours (until both bromophenol blue and xylene cyanol loading buffer dye fronts had entered the lower buffer chamber), the RNA was transferred to a nylon membrane (Amersham Hybond-N+, GE Healthcare) as previously described (Rio, 2014) except that the transfer was performed at 4°C for ~18 h at 30 V (constant voltage).

For detection of the RNA, the nylon membrane was first baked in a hybridisation oven at 80°C for 2 hours to allow fixing of the RNA to the membrane. The membrane was washed in 2X SSC solution (Appendix E) before being placed in a pyrex glass tube containing prehybridization solution (DIG Easy Hyb, Merck) and incubated for 3-4 hours at 68°C with rotation. The riboprobe solution was prepared by adding 20 µL of riboprobe (prepared above) to 50 µL of MilliQ H_2O, followed by heating at 95°C for 5 min before cooling on ice. The entire volume of heated probe was added to the solution bathing the membrane and allowed to hybridize for ~18 h overnight at 68°C, with rotation. The hybridisation solution was removed, and the membrane briefly rinsed with 2X SSC followed by two 5 min washes in 15 mL of 2X SSC solution at 68°C. The membrane was then briefly rinsed with 0.1X SSC solution (Appendix E) followed by two 15 min washes in 15 mL of 0.1X SSC solution at 68°C. The membrane was then washed in washing buffer (Appendix E) for 2 min then placed in 25 mL blocking solution (Appendix E) for 1 h at RT with shaking. After blocking, the membrane was incubated with anti-digoxigenin-AP, Fab fragments (Merck) at a dilution of 1:5000 in fresh blocking solution (Appendix E) for 30 min with shaking (RT). Following incubation with the antibody solution, the membrane was washed twice in 25 mL washing buffer for 15 min with shaking (RT). For detection of the bound probe, the membrane was equilibrated in 25 mL detection buffer (Appendix E) for 3 min before the addition of 1 mL of detection buffer, with

the addition of 10 µL CDP-Star chemiluminescent substrate (Roche). Bands present on the membrane were visualised using Fuji X-ray film at a range of exposure times.

5.2.10. Construction of *P. multocida* sRNA mutants

The TargeTron mutagenesis system (Sigma-Aldrich) was used to separately inactivate *P. multocida* sRNAs Prrc06, Prrc08 and Prrc25 as previously described (Harper et al., 2013). For each mutant construction, the group II intron within the *P. multocida* TargeTron plasmid pAL953 was retargeted to the sRNA gene of interest using specific primers designed using the TargeTron design site (Sigma-Aldrich) or the ClosTron tool (Heap et al., 2010) (Table 5.2). The constructed plasmids, encoding introns retargeted to Prrc06 (pAL1391), Prrc08 (pAL1368) and Prrc25 (pAL1393) were then used to separately transform *P. multocida* VP161 by electroporation. Transformants were identified and confirmed to be correct using separate colony PCR reactions with Prrc06-, Prrc08- or Prrc25-specific forward primer (BAP8420, BAP8328 and BAP8421, respectively) (Table 5.2) together with a reverse primer specific for the intron (EBS universal or BAP6797). To confirm that only a single TargeTron insertion had occurred in the *P. multocida* genome and in the correct position, direct sequencing using genomic DNA isolated from each TargeTron mutant was employed, using the intron-specific EBS universal primer that generated sequence representing the 5′ end of the intron and the genomic region immediately adjoining the intron. Three VP161 mutants that gave the correct PCR profile and generated the correct and unambiguous sequencing data were selected for further study. These mutants were designated AL3272 (Prrc06 mutant), AL3188 (Prrc08 mutant) and AL3275 (Prrc25 mutant).

5.2.11. Construction of sRNA overexpression strains and complemented mutant strains

To generate each *P. multocida* sRNA overexpression strain, the appropriate sRNA sequence was PCR-amplified from WT VP161 genomic DNA using primers specific for each sRNA; BAP8326 and BAP8327 (Prrc06), BAP8328 and BAP8329 (Prrc08) or BAP8330 and BAP8331 (Prrc25). Each PCR fragment generated contained a BamHI site at the 5′ end and either an EcoRI (Prrc06 and Prrc08) or SmaI (Prrc25) site at the 3′ end, allowing them to be digested then ligated to similarly digested pREXY vector. Ligation mixtures were separately transformed into *E. coli* DH5α and transformants containing pREXY with an insert were identified using colony PCR analysis with pREXY primers, BAP612 and BAP2679, that flanked the cloning region. The identity of each recombinant plasmid was confirmed in

separate sequencing reactions using pREXY primers, BAP612 and BAP2679. Plasmids with the correct sRNA and *tpiA* promotor sequence were identified and designated pAL1394 (Prrc06), pAL1395 (Prrc08) and pAL1396 (Prrc25). To generate the *P. multocida* sRNA overexpression strains, each sRNA overexpression plasmid was used to separately transform electrocompetent VP161 cells. The identity of five *P. multocida* transformants from each experiment was confirmed using colony PCR with pREXY-specific primers that flanked the cloning region (BAP612 and BAP2679). *P. multocida* sRNA overexpression strains were designated AL3282 (pAL1394_Prrc06), AL3283 (pAL1395_Prrc08) and AL3284 (pAL1396_Prrc25).

The three *P. multocida* sRNA overexpression vectors were separately used to transform the appropriate sRNA mutant to generate complemented mutant strains. Thus, the overexpression vectors, pAL1394, pAL1395 and pAL1396, were used to transform AL3272 (Prrc06), AL3188 (Prrc08) and AL3275 (Prrc25) sRNA mutants, respectively. Spectinomycin resistant transformants were checked for the correct plasmid identity by colony PCR using primers that flanked the recombinant gene region (BAP612 and BAP2679). Prrc06, Prrc08 and Prrc25 sRNA complemented mutants were designated AL3306, AL3307 and AL3308, respectively.

5.2.12. Generation of an antisense sRNA *P. multocida* expression plasmid

Before generating each antisense sRNA expression plasmid, the downstream terminator sequence from the *P. multocida tpi* gene was first cloned into the EcoRI unique restriction enzyme site at the 3′ of the multiple cloning site region in the pREXY expression plasmid. This allowed for the appropriate termination of the antisense sRNA transcript. The predicted *tpiA* terminator sequence, defined as a stem loop followed by a string of uridine nucleotides, was identified by analysis of *tpiA* gene transcripts from available RNA-Seq data in the Boyce Laboratory. To generate the DNA fragment encoding the *tpiA* terminator sequence, two complementary primers BAP8353 and BAP8354 (Table 5.2) were generated such that when annealed they formed a double-stranded DNA fragment with EcoRI-compatible overhangs but when ligated to an EcoRI sticky end on the vector an EcoRI site would only be reconstituted at one end (to allow for later cloning reactions). Briefly, primers at equimolar concentrations were combined in annealing buffer (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA and 50 mM NaCl) and subjected to the following thermal cycling conditions; 95°C for 5 min, followed by 70 x one min cycles, starting at 94°C, each cycle thereafter having a temperature reduction of 1°C. The resulting double-stranded fragment encoding the *tpiA* terminator sequence and EcoRI

compatible ends was then ligated to EcoRI-digested pREXY (Table 5.1). The ligation was then used to transform competent *E. coli* DH5α cells. A transformant with the correct recombinant plasmid was identified using colony PCR with primers that flanked the multiple cloning site (BAP612 and BAP2679). To confirm identity, the plasmid was isolated and separately sequenced using the above primers. The antisense expression plasmid, encoding a *tpi* promoter upstream of the multiple cloning site and a *tpi* transcript terminator sequence downstream of the multiple cloning site, was designated pAL1380.

5.2.13. Construction of *P. multocida* sRNA antisense expression strains

Each sRNA was inserted into the appropriate site in the pAL1380 expression plasmid in the opposite orientation, relative to normal expression of the sRNA, between the *tpiA* promoter and terminator regions, allowing for expression and termination of each antisense sRNA transcript. As the antisense expression vector, pAL1380, already encoded a transcriptional terminator sequence downstream of the multiple cloning site, primers used for the amplification of Prrc06 were placed in such a way to exclude the predicted terminator sequence identified bioinformatically. For the sRNAs Prrc08 and Prrc25, a terminator sequence could not be identified. For these sRNAs, available transcriptomic data was utilized and the end of the transcript was deemed to be at the end of the RNA-Seq read peak for each region; sequence beyond this peak was excluded from primer design (~40 bp). The DNA representing each sRNA was PCR-amplified from WT VP161 genomic DNA with primers BAP8450 and BAP8451 (Prrc06), BAP8452 and BAP8453 (Prrc08) or BAP8454 and BAP8455 (Prrc25) (Figure 5.1). The amplified fragments were then digested with the appropriate enzymes and cloned into similarly digested pAL1380. Plasmids containing the correct inserts were identified from *E. coli* DH5α transformants by PCR analysis using primers BAP612 and BAP2679 that flanked the cloning region. For each plasmid, the correct antisense sRNA sequence was confirmed by separate sequencing reactions using the vector primers BAP612 and BAP2679. Antisense expression plasmids were designated pAL1432 (Prrc06 antisense), pAL1433 (Prrc08 antisense) and pAL1434 (Prrc25 antisense). Each plasmid was used to separately transform strain VP161 to generate *P. multocida* antisense sRNA strains. Following confirmation by colony PCR, the *P. multocida* strains were designated AL3285 (pAL1432_Prrc06 antisense), AL3286 (pAL1433_Prrc08 antisense) and AL3287 (pAL1434_Prrc25 antisense).

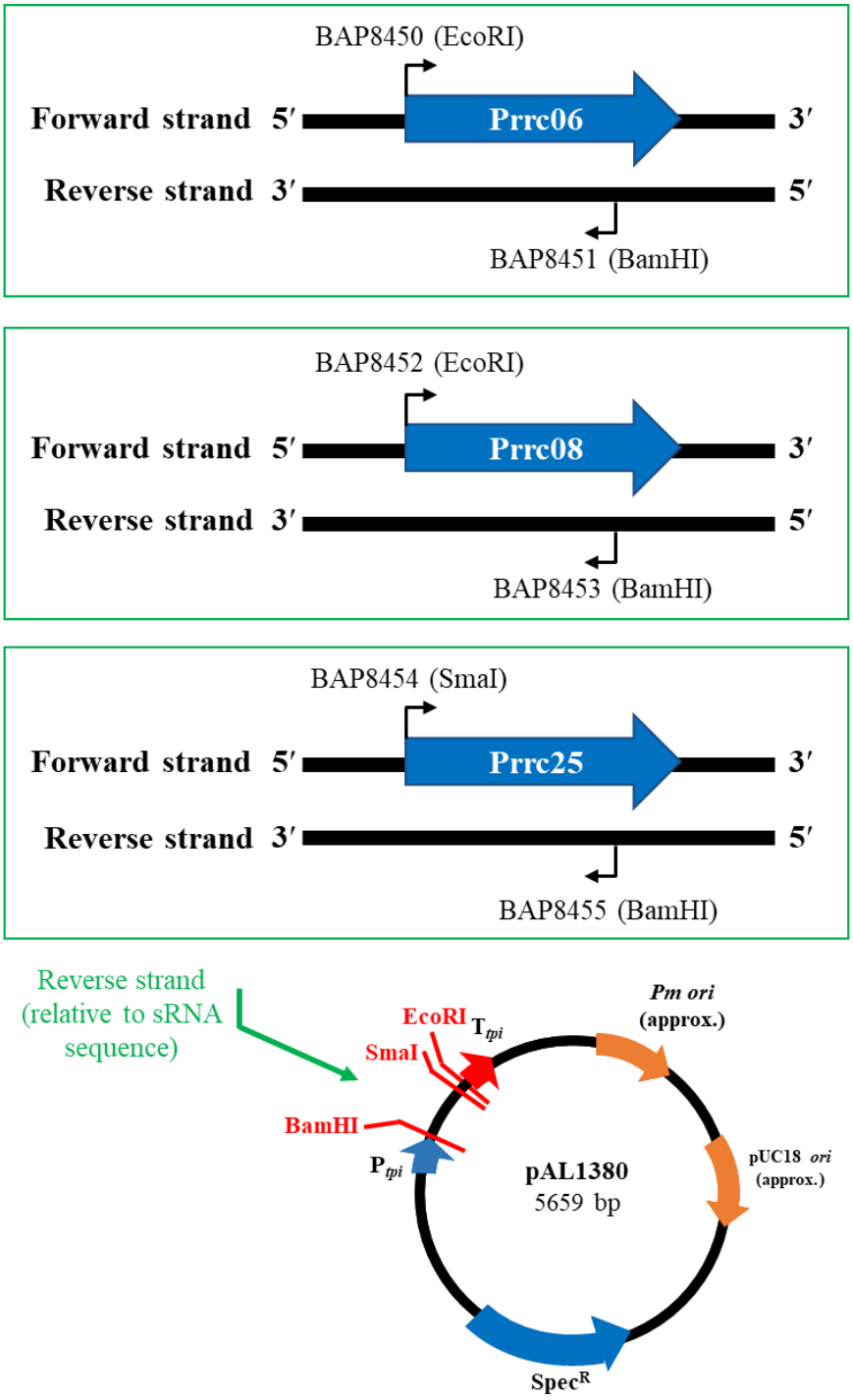


Figure 5.1. Schematic diagram showing the construction of sRNA antisense expression plasmids used in this study. The Prrc06, Prrc08 and Prrc25 sRNA sequences (excluding predicted terminator sequence) were PCR-amplified from wild-type *P. multocida* (strain VP161). The PCR product encoding each sRNA was separately digested with the appropriate enzymes then ligated to similarly digested pREXY to generate pAL1432 (Prrc06 antisense), pAL1433 (Prrc08 antisense) and pAL1434 (Prrc25 antisense). DNA fragments were cloned in the opposite orientation, relative to the sRNA and the direction of the *tpi* promoter. Abbreviations: P_{tpi} , *P. multocida tpiA* promoter; T_{tpi} , *P. multocida tpiA* terminator; pUC18 *ori*, pUC18-derived *E. coli* origin of replication; Pm *ori*, *P. multocida* origin of replication; Spec R (*aadA*), spectinomycin resistance gene.

5.2.15. Quantitative hyaluronic acid capsule assay

The amount of hyaluronic acid (HA) capsular material produced by each *P. multocida* strain was quantified as described previously (Mégroz et al., 2016).

5.2.16. Bioinformatic analysis

To determine RNA secondary structures, the Mfold Web Server was used with default parameters (Zuker, 2003).

5.3. Results

5.3.1. Selection of putative sRNAs for experimental validation and characterisation

Six putative *P. multocida* sRNAs, Prrc06, Prrc07, Prrc08, Prrc10, Prrc12 and Prrc25, identified using bioinformatic and Hfq co-immunoprecipitation analyses (Chapters 3 and 4), were selected for experimental validation. All six passed the initial criteria for sRNA identification as reported in Chapter 3, i.e. all were short in nature and in most cases highly expressed, especially in the sRNA-enriched RNA-Seq dataset. Thus, they were predicted to be *bone fide* sRNA species expressed by *P. multocida* (Table 5.3). Four of the six sRNAs, Prrc06, Prrc07, Prrc12, and Prrc25, were predicted to be *trans*-encoded sRNAs (category 1, Table 3.4, Chapter 3). The remaining sRNAs, Prrc08 and Prrc10, overlapped (on the same strand) a predicted open reading frame in one or more of the *P. multocida* genomes analysed (category 2, Table 3.4). All six sRNAs displayed high sequence conservation between the *P. multocida* strains VP161, X-73 and Pm70 (Table 5.3) and had no significant identity with known sRNA species or open reading frames present in the NCBI database. The Prrc07 putative sRNA is a homologue of sRNA 20 (AaHKsRNA20) in *Aggregatibacter actinomycetemcomitans* (Oogai et al., 2017); although this sRNA has not yet been characterised in *A. actinomycetemcomitans*, it was still of interest to investigate in *P. multocida*.

The expression of identified sRNAs during low iron and reduced oxygen growth was also investigated in Chapter 3 and considered during selection of sRNA candidates for further study. The Prrc06 sRNA displayed increased expression during growth in reduced oxygen conditions and may therefore help to regulate the expression of specific genes under this condition. In addition, the sRNAs Prrc06, Prrc10, Prrc12 and Prrc25 had altered transcript levels in the *proQ* mutant, compared to those in the wild-type parent (Gulliver, Boyce Laboratory, manuscript in preparation, 2020); the levels of the RNA chaperone ProQ could therefore have a direct effect on the abundance of these sRNAs. *P. multocida* Hfq co-immunoprecipitation (Co-IP) studies were performed to identify RNA species, including sRNAs and mRNAs that associate with the RNA chaperone Hfq (Chapter 4). The Prrc10 sRNA displayed an association with Hfq using Co-IP and may therefore be a Hfq-dependent sRNA. None of the six sRNAs chosen for further study were predicted to target mRNAs known to be associated with virulence using TargetRNA2 (Kery et al., 2014). However, other sRNAs associated with Hfq may play a role in the regulation of virulence as Hfq is essential for capsule biosynthesis (Mégroz et al., 2016).

Additional information was also gained via analysis of whole transcriptome data of wild-type VP161. Genes within the *P. multocida* capsule biosynthesis operon are highly expressed during early-exponential growth phase, but have reduced levels of expression during mid-to-late-exponential growth phase (Figure 5.2A.). Expression levels of the putative sRNAs identified in Chapters 3 and 4 were also examined and only two, Prrc08 and Prrc25, displayed a similar expression pattern to that shown by the capsule genes (Figure 5.2B) (Boyce and Harper, pers. comm.). Interestingly, the Prrc08 and Prrc25 putative sRNAs are nearly adjacent to each other on the *P. multocida* chromosome, separated only by the predicted adhesin protein, YadA (PmVP161_0591). Moreover, it was also observed that the sequence at the transcript start of both *hyaE* and *phyA* capsule genes (when reverse-complemented) shared identity to the sequences of the Prrc08 and Prrc25 putative sRNAs (Figure 5.3). Given that many sRNAs have a region of sequence similarity (called the seed region) that is used to bind to their target mRNA molecules, facilitating riboregulatory function (Peer and Margalit, 2011), it was hypothesized that Prrc08 and/or Prrc25 may bind to these regions of *hyaE* and *phyA* and therefore play a role in capsule gene expression.

Table 5.3. Criteria used to analyse and select putative sRNAs for experimental validation.

Criteria	**Prrc06**	**Prrc07**	**Prrc08**	**Prrc10**	**Prrc12**	**Prrc25**
Short sequence (~50-400 bp)	160 bp	191 bp	183 bp	199 bp	158 bp	157 bp
Highly expressed (transcript reads at peak >50) [a]	179; 359	20; ND	8646; ND	260; ND	735; ND	405; ND
trans-acting sRNA; intergenic location? Category 1, Table 3.4 (Chapter 3).	Yes	Yes	No	No	Yes	Yes
Overlaps (or partially overlaps) predicted gene? Category 2, Table 3.4 (Chapter 3)	No	No	Yes	Yes	No	No
cis-acting sRNAs; match on opposite strand of predicted gene? Category 3, Table 3.4 (Chapter 3)	No	No	No	No	No	No
VP161 sRNA conserved (94-98% identity) in *P. multocida* strains X-73 and Pm70?	Yes	Yes	Yes	Yes	Yes	Yes
Clear Rho-independent terminator (red region) / 3′-region rich in uridine? [b,c]	Yes/Yes	Yes/Yes	No/No	No/No	Yes/Yes	No/Yes
	uugauuaauccg uagagguaacaa ugagugaauuug guuaugcaauga ugacaguugcuu uggcgcuagguu uaauu*guugua uuggcguucuca auuuuuuuaagac uaaacaaaauaaa aauaacgaaagau cgaggcauaauc gccucgguc uuuuuu	aucacagcuaagccg aaaucauuuuuacuu uuuuaucuuaaaagc gugaauuaguucaca uuauugaaaauaccc auuugcuaacuaagg uuuauuuucguauu cugugucuagcuaaa aauaagugcuauaaa aaauagccgacauau uuugcuuuccgagu agugccuugu uuuca aggcacuuuuuuuu	uugccgcaaaagcgg gaaucgccaaaagcg uagaaggugcc*gc guuaucuuauaacg uuggugugaacuau gaguuuuaaucgua aaaaggaguuuaacg gacaacauuaauauu auuuaccucuaaacu aucccaaauuuauuc aaauagcgcgcuacg uuuucguagcgcgu cuuauauaa	uucguuaugaaaaac ucaaaauggaucagg caaacuacgcagccu gauuauuuuagaga gaauagccugaaaua cgguuacucaaagua agaaauauucagugc uuuacaaguaaucuc uucggaaauucgccu gaauaucgguguuu aacaaaaaacaaugc ggaguuauuaaacaa ugccuaaaauuaaaa caguac	cauggauaaggaug uuugcguuuaacgc augaaaaaauaaaag aauguauguuuuag guugauuuguauuu aauggauuaacuuga cacaacaucagcuca ugccaauagucaacg cauuucauccuuuua auuaauuugcuaaac gcaauuugauuu	uaguagaauucuu uuuaaauuuaauc uugcaauaggauu uugc*auuauuaa auauguuuaaaga uaaaugaucaguu guauaaguaugaa guuuaaggaguac uuaaugaauuuau cuaaaaucgcaaca gggaucguugcgu cauuuuguuuuau u

Criteria	Prrc06	Prrc07	Prrc08	Prrc10	Prrc12	Prrc25
High number of transcripts in: 1. sRNA-enriched strain X-73 data 2. Matching total RNA strain X-73 data 3. Strain VP161 RNA data (Chapter 3)	1, 2 and 3	1 and 3	1 and 3	1	1	1 and 3
Altered sRNA expression during: 1. Low iron growth 2. Reduced oxygen growth	2	-	-	-	-	-
Altered sRNA expression in: 1. *hfq* mutant compare to wild-type parent 2. *proQ* mutant compare to wild-type parent	2	-	-	2	2	2
sRNA sequence enriched in Hfq Co-IP?	No	No	No	Yes	No	No
Predicted to be virulence-associated?	No prediction	No prediction	Possible capsule genes	No prediction	No prediction	Possible capsule genes

[a] The peak number of reads in the sRNA-enriched data vs non-enriched data, separated by a semicolon are listed. These reads were generated using data generated with *P. multocida* strain X-73 and visualised using Artemis software (Chapter 3).
[b] *P. multocida* strain VP161 sequence, where known, sequence start sites are reported as determined from primer extension experiments (section 5.3.2).
[c] Location of TargeTron intron insertion for mutagenesis of the sRNA (section 5.3.3) is indicated by an asterisk (*) within the sequences where appropriate.

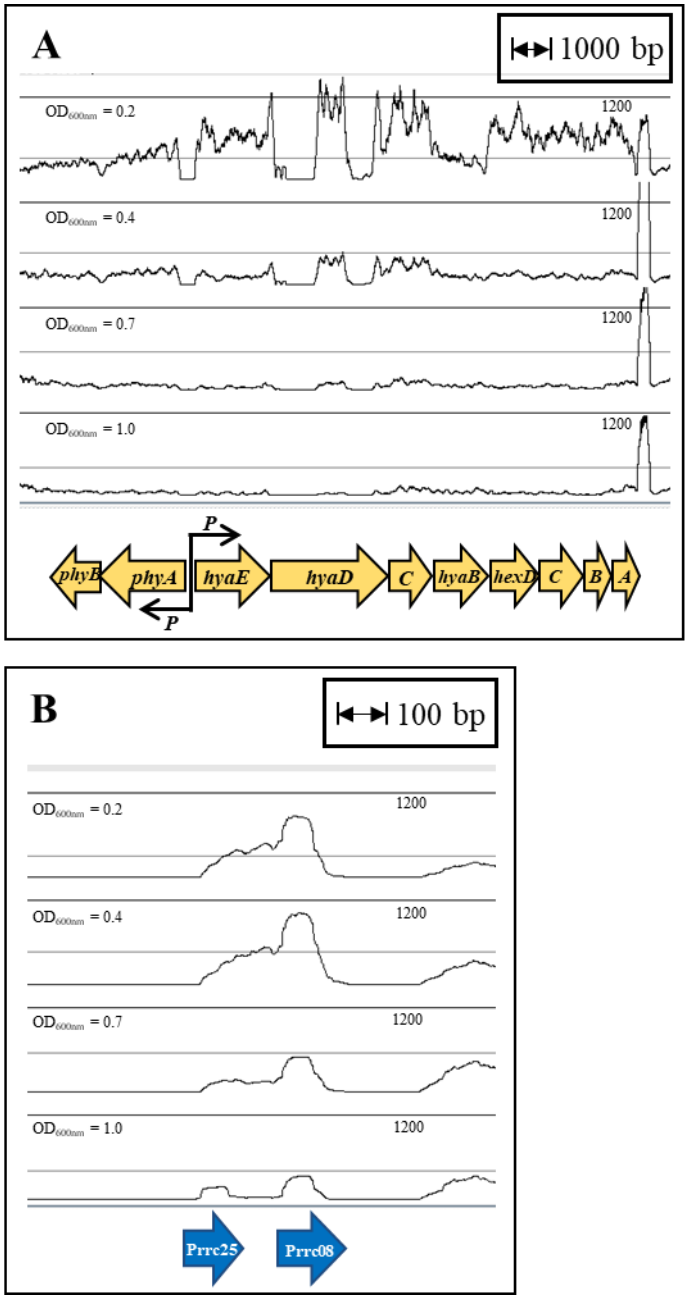


Figure 5.2. A. Relative number of RNA-Seq reads mapped to genes within the *P. multocida* serogroup A capsule biosynthetic locus, as visualised using Artemis (Sanger). Sourced from whole transcriptomic data of strain VP161 grown *in vitro* (at 37°C with shaking) in brain heart infusion broth to early-exponential (OD_{600} = 0.2 and 0.4), mid-exponential (OD_{600} = 0.7) and late-exponential (OD_{600} = 1.0) growth phases. **B.** Number of RNA-Seq reads that mapped to the putative *P. multocida* sRNAs Prrc25 and Prrc08 at the growth phases detailed above.

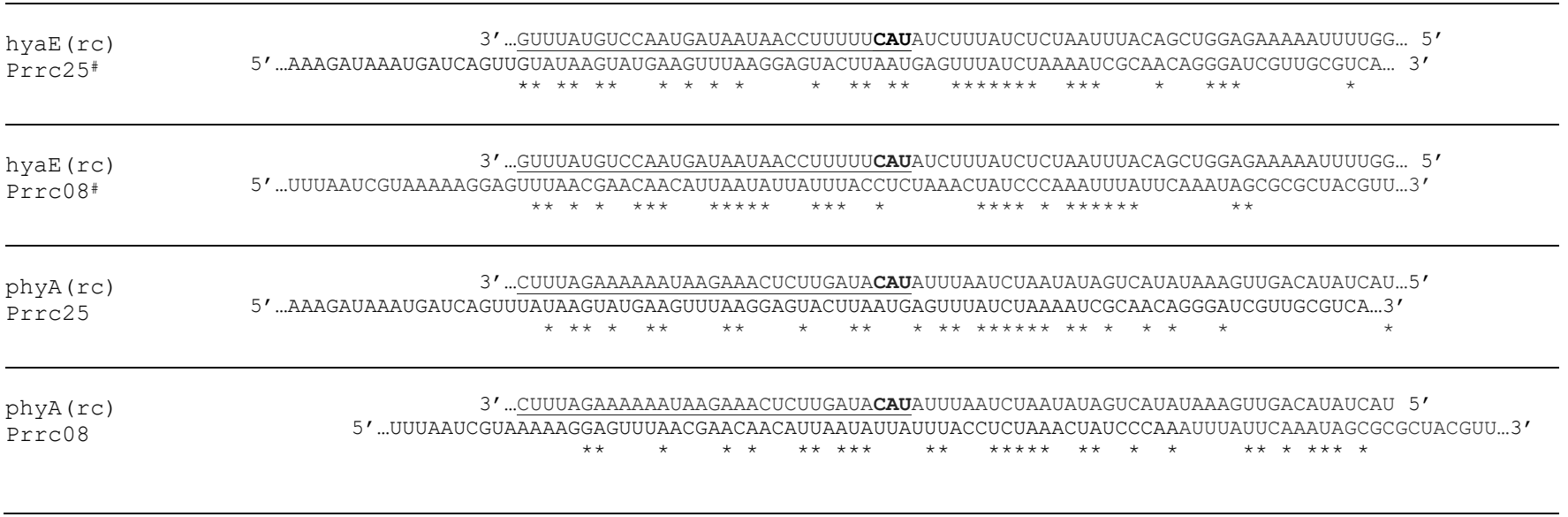


Figure 5.3. Sequence similarity between capsule transcripts and putative sRNAs Prrc08 and Prrc25. Intergenic and start regions of *hyaE* and *phyA* and relevant regions of the putative sRNAs Prrc08 and Prrc25 are shown. Capsule gene transcripts are shown 3′ to 5′ (rc) and the relevant region of each of the sRNA sequences is shown 5′ to 3′. Position of translation start codon (3′ to 5′) is indicated by the bold lettering. Nucleotides within the capsule gene open reading frame region are underlined.

5.3.2. Determination of transcript start of each putative sRNA

Each of the sRNA transcript start sites were first estimated by visual observation of previously generated RNA-Seq data (sRNA-enriched) from wild-type X-73 *P. multocida* cells (Chapter 3). The transcriptional start sites were predicted to be at the very start of the peak of RNA-Seq reads that mapped to the putative sRNA sequence. To precisely identify the 5′ start site of the six selected sRNA transcripts, fluorescent primer extension was performed using the information obtained from the above analysis. Each fluorescently labelled reverse primer was designed such that it would anneal approximately 100 bp downstream of the estimated transcript start site.

For the primer extension analysis, RNA to be used for the generation of cDNA was isolated from wild-type *P. multocida* VP161 grown to the appropriate growth phase for each sRNA to be optimally expressed, as informed by previously generated whole transcriptome data (Boyce, pers. comm.). Thus, RNA was extracted from early-exponential ($OD_{600} \approx 0.2$) growth phase for analysis of the putative sRNAs Prrc08 and Prrc25, and from mid-exponential ($OD_{600} \approx 0.7$) growth phase for analysis of sRNAs Prrc06, Prrc07, Prrc10 and Prrc12. The isolated RNA was used as a template in separate reactions to extend the fluorescently labelled reverse primers specific for Prrc06, Prrc07, Prrc08, Prrc10, Prrc12 and Prrc25 (BAP8092, BAP8093, BAP8234, BAP8094, BAP8095 and BAP8235, respectively, Table 5.2). Fragment size analysis of the generated cDNA molecules identified fragments of between 57-156 nucleotides in length. The precise start site for each sRNA transcript was identified by correlating the cDNA molecule size to the primer annealing position (Table 5.4 and Figure 5.4).

From the primer extension analysis, with the exception of Prrc07 and Prrc12, the transcript start site as determined by primer extension analysis aligned closely to the very start of the peak of RNA-Seq reads that mapped to each sRNA (Figure 5.4). However, for the sRNA Prrc12, the transcript start site determined by primer extension analysis correlated to the middle of the peak of RNA-Seq reads in that region (Figure 5.4) indicating that the extension of the primer may have been arrested resulting in a shorter cDNA molecule that predicted. For Prrc07, two separate primer extension experiments were performed but no primer extension product could be generated and therefore the precise transcript start site could not be determined.

The sequences upstream of the encoding regions for sRNAs Prrc06, Prrc08, Prrc10 and Prrc25 (all with clear primer extension products and transcript start sites that aligned with those

predicted by RNA-Seq analysis) were visually inspected and putative promoters identified (Table 5.4). Following the determination of the transcript start sites for Prrc06, Prrc08, Prrc10, Prrc12 and Prrc25, the Mfold Web Server (Zuker, 2003) was used to model the predicted secondary structures of the sRNAs (Figure 5.5). The transcriptional start sites determined from fluorescent primer extension analysis were used to define the start of each of the Prrc06, Prrc08, Prrc10 and Prrc25 sRNA sequences used as input for the modelling. For Prrc12, the original bioinformatically predicted transcript sequence start was instead used as input (Table 5.3). As no primer extension product was generated for Prrc07, the predicted secondary structure was not determined for this putative sRNA.

Table 5.4. Fluorescent primer extension analysis results for putative sRNAs Prrc06, Prrc07, Prrc08, Prrc10, Prrc12 and Prrc25.

sRNA	**cDNA molecule size generated from fragment size analysis (bp)**	**Sequence of sRNA transcript start as determined from fragment size analysis**	**Sequence of sRNA transcript with predicted promoter sequences indicated** [a]
Prrc06	93	5′-UUGAUUAAUC-3′	5′- UUUUAAAUGUUGAAAGAGACUUCUAUUUCAUAUCGGUU UAUUGAUAGAAUGGAACUCUAUUUGAUUAAUCCGUAGAG GUAACAAUGAGUGAAUUUGGUUAUGCAAUGAUGACAGUU GCUUUGGCGCUAGGUUUAAUUGUUGUAUUGGCGUUCUCAA UUUUUUAAGACUAAACAAAAUAAAAAUAACGAAAGAUCG AGGCAUAAUCGCCUCGGUCUUUUUU -3′
Prrc07	No product	N/A	
Prrc08	156	5′-UUGCCGCAAA-3′	5′- UUACCAAUCUCAAACUGCCAUUGCUAUUGGUACAGGUU AUCGUUUUAAUGAAAGAAUUGCCGCAAAAGCGGGAAUCGC CAAAAGCGUAGAAGGUGCCGCGUUAUCUUAUAACGUUGGU GUGAACUAUGAGUUUUAAUCGUAAAAAGGAGUUUAACGG ACAACAUUAAUAUUAUUUACCUCUAAACUAUCCCAAAUUU AUUCAAAUAGCGCGCUACGUUUUCGUAGCGCGUCUUAUAU AA -3′
Prrc10	82	5′-UUCGUUAUGA-3′	5′- AACAGACGAAUUCUAUAGAAUUCGUCUGUUAACAGACG AAUUCUAUAGAAUUCGUCUGUUUUUCGUUAUGAAAAACUC AAAAUGGAUCAGGCAAACUACGCAGCCUGAUUAUUUUAGA GAGAAUAGCCUGAAAUACGGUUACUCAAAGUAAGAAAUAU UCAGUGCUUUACAAGUAAUCUCUUCGGAAAUUCGCCUGAA UAUCGGUGUUUAACAAAAAACAAUGCGGAGUUAUUAAACA AUGCCUAAAAUUAAAACAGUAC -3′

sRNA	cDNA molecule size generated from fragment size analysis (bp)	Sequence of sRNA transcript start as determined from fragment size analysis	Sequence of sRNA transcript with predicted promoter sequences indicated [a]
Prrc12	57	5′-UGAUUUGUAU-3′	
Prrc25	139	5′-UAGUAGAAUU-3′	5′- UAUUUUUGUUCUAUUUUAGCUUCAUAUUUAUUCUUUUAAAUUUUAAAAUAAAUCAAAGGCUUGUAAUUUGUAGGGGAUUUUUAGUAGAAUUCUUUUUAAAUUUAAUCUUGCAAUAGGAUUUUGCAUUAUUAAAUAUGUUUAAAGAUAAAUGAUCAGUUGUAUAAGUAUGAAGUUUAAGGAGUACUUAAUGAAUUUAUCUAAAAUCGCAACAGGGAUCGUUGCGUCAUUUUGUUUUAUU -3

[a] The sRNA sequence is represented by the blue text; predicted -10 and -35 promoter regions for each sRNA are highlighted in red, respectively. Promoter sequences were only predicted for sRNAs where the transcript start site as determined by primer extension analysis aligned closely to the very start of the peak of RNA-Seq reads that mapped to each sRNA.

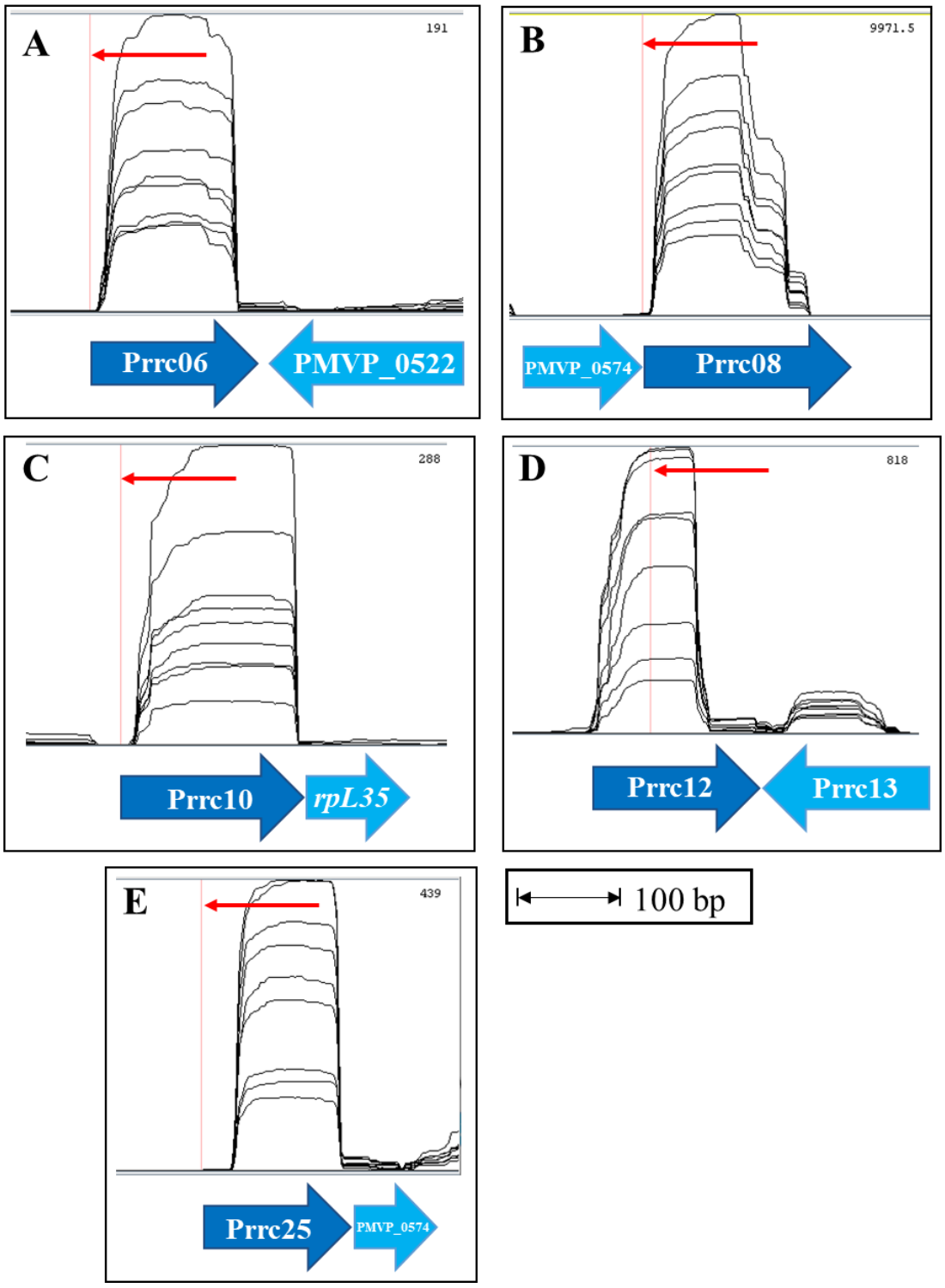


Figure 5.4. Identification of putative *P. multocida* sRNA transcriptional start sites. Predicted start sites for Prrc06 (**A**), Prrc08 (**B**), Prrc10 (**C**), Prrc12 (**D**) and Prrc25 (**E**)**.** The cDNA product sizes determined from primer extension analysis were mapped to each sRNA sequence in the *P. multocida* VP161 genome, relative to the annealing sites of the fluorescently labelled primers used in the analysis. Previously generated sRNA-enriched RNA-Seq data was used to overlay and visualise the transcript start sites as determined by primer extension. The transcript start site as determined by fluorescent primer extension for each sRNA is indicated by the pink vertical line (red arrow).

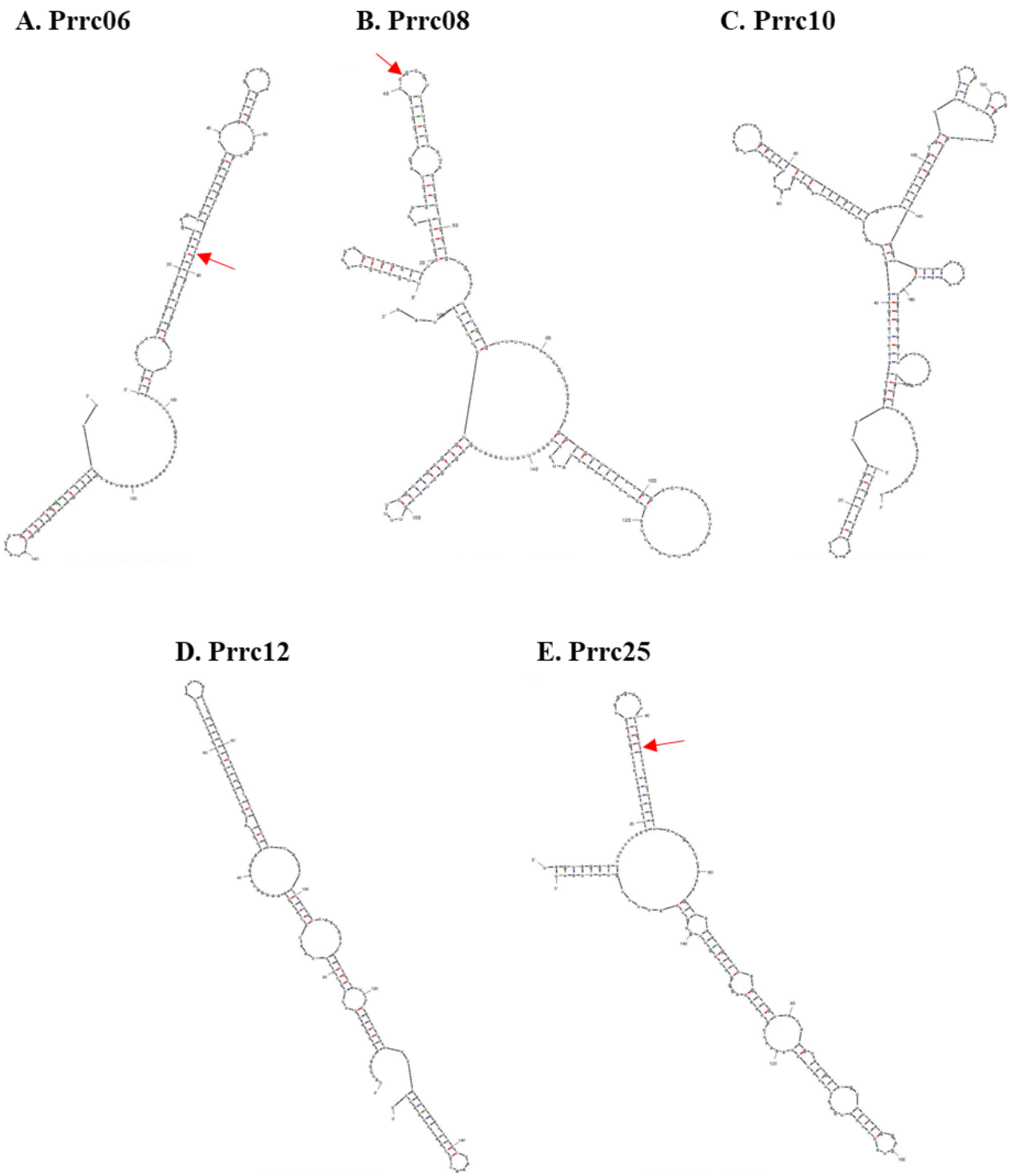


Figure 5.5. Predicted secondary structures of putative *P. multocida* sRNA molecules Prrc06 (**A**), Prrc08 (**B**), Prrc10 (**C**), Prrc12 (**D**) and Prrc25 (**E**). Red arrows indicate the location of intron insertion in the sRNA sequence structure during TargeTron mutagenesis (section 5.3.3). Modelled using the Mfold Web Server (Zuker, 2003).

5.3.3. Further experimental validation and phenotypic analysis of Prrc06, Prrc08 and Prrc25 sRNAs

Following the determination of the transcript start sites, three of the six putative sRNAs, namely Prrc06, Prrc08 and Prrc25, were selected for further investigation. Each of these sRNAs had experimentally determined transcript start sites that correlated with the transcript start predicted using RNA-Seq reads (section 5.3.2). The Prrc06 sRNA was of particular interest as it displayed altered expression under reduced oxygen conditions, and may therefore play a regulatory role under this important infection-relevant growth condition. In addition, Prrc08 and Prrc25 were of particular interest due to their growth-phase dependent expression pattern partially correlating with capsule production and potential sequence similarity with a number of the capsule biosynthesis genes (section 5.3.1).

To help determine the role of the Prrc06, Prrc08 and Prrc25 sRNAs in *P. multocida*, insertional mutants and and overexpression strains were separately generated for each of the three sRNAs in *P. multocida* strain VP161 (see materials and methods for the construction of these strains). Although it is acknowledged that insertional mutants may not be as effective as deletion mutants in abrogating sRNA function, the success rate of deletion mutants in *P. multocida* is very low, thus TargeTron intron insertion mutants were generated. The position of intron insertion within each sRNA is indicted on their predicted secondary structures and was likely to significantly disrupt the function of the region (Figure 5.5). As additional control strains for the phenotypic analyses of the Prrc06, Prrc08 and Prrc25 sRNA mutant strains, antisense sRNA strains were generated (section 5.2.13; Figure 5.1). These strains express a specific ‘antisense’ sRNA fragment with sequence complementarity to the corresponding sRNA under examination and can bind and inactivate the sRNA transcript (Figure 5.1). The antisense sRNA expression strains thus act as additional control strains and are of particular importance if the mutagenesis of the sRNA encoding region was not effective.

5.3.3.1. Northern blot analysis of sRNAs

To confirm expression of the sRNA transcripts and to experimentally determine the size of Prrc06, Prrc08 and Prrc25, northern blotting was performed using complementary, strand-specific DIG-labelled riboprobes. Complementation strains were also constructed and used in the experiment; the Prrc06, Prrc08 and Prrc25 mutants were each provided a copy of the appropriate sRNA on an RNA expression plasmid (pAL1394, pAL1395 and pAL1396, respectively, see materials and methods). In addition, sRNA overexpression strains were

produced (WT harbouring the sRNA expression plasmid pAL1394, pAL1395 or pAL1396). For each sRNA examined, RNA was isolated from WT VP161, the corresponding mutant strain, the complemented sRNA mutant strain, the sRNA overexpression strain and from WT VP161 harbouring empty expression vector. Prior to northern analysis, dot blot assays were performed to ensure all riboprobes used were efficiently DIG-labelled (data not shown).

Northern analysis for the Prrc06 sRNA, predicted to be 160 bp in length, showed the probe hybridised with a fragment approximately 200 bp in size in RNA isolated from WT VP161 and WT VP161 harbouring empty expression vector, and strongly with a band approximately 220 bp in size in RNA isolated from the complemented sRNA mutant strain and the sRNA overexpression strain (Figure 5.6A). Importantly there were no hybridizing bands in the RNA isolated from the Prrc06 mutant, indicating that insertion inactivation of the sRNA was successful. However, as not all of the standards in the ssRNA ladder were clearly visible, the sizes of these hybridised fragments will require confirmation in future northern blot analysis.

For Prrc08 northern analysis, all molecular bands in the ssRNA ladder were clearly visible, allowing for more accurate estimation of hybridised fragment sizes. The probe hybridised very weakly with a fragment of approximately 180 bp size (predicted size 183 bp) in the WT strain but there were no hybridising bands detected in the lane containing RNA isolated from WT containing the empty expression vector, both results indicating the innate level of expression of Prrc08 in WT VP161 is relatively low (Figure 5.6B). However, the probe hybridized strongly to the equivalent band in the lanes representing the Prrc08 complementation and overexpression strains. No hybridising bands were observed in the Prrc08 mutant strain indicating that inactivation of the sRNA was successful. Northern blots performed for Prrc25 analysis, predicted to be 157 bp in length, failed to detect any hybridising bands in RNA samples suggesting that expression of this sRNA was insufficient for detection or that the northern hybridisation had failed. Future work will involve optimisation and repetition of these northern blots.

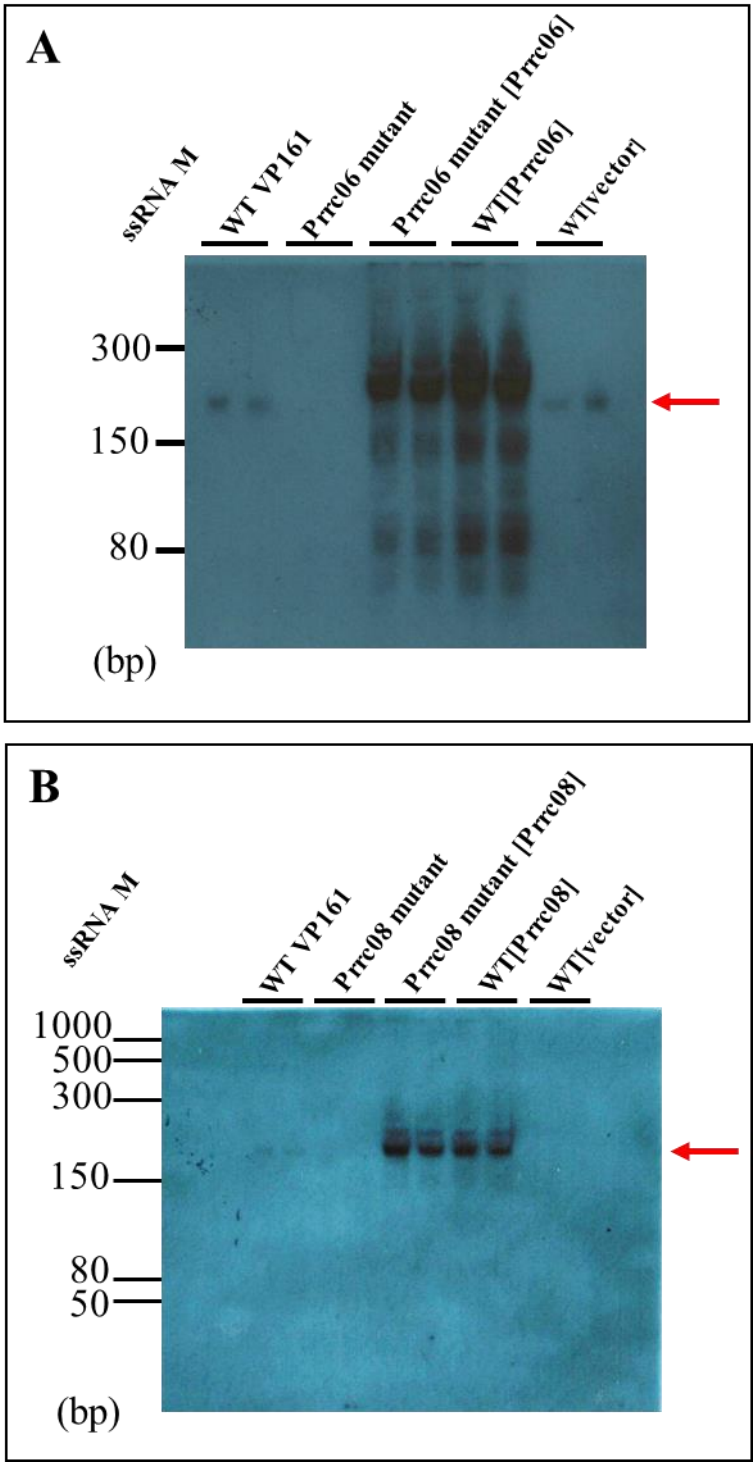


Figure 5.6. A. Northern blotting using an RNA probe specific for Prrc06. RNA was isolated from WT VP161, Prrc06 mutant, complemented Prrc06 mutant [Prrc06], overexpression strain WT [Prrc06] and VP161 with vector only, WT [vector]. All cells were grown to mid-exponential growth phase (biological duplicate) prior to RNA extraction. **B**. Northern blotting using an RNA probe specific for Prrc08. RNA was isolated RNA was isolated from WT VP161, Prrc08 mutant, complemented Prrc08 mutant [Prrc08], overexpression strain WT [Prrc08] and VP161 with vector only, WT [vector]. All cells were grown to mid-exponential growth phase (biological duplicate) prior to RNA extraction. Sizes of the molecular weight markers (NEB low range ssRNA ladder) are indicated at the left of the gels.

5.3.3.2. General growth characteristics of sRNA-modified *P. multocida* strains

The growth of the Prrc06, Prrc08 and Prrc25 sRNA mutants and corresponding overexpression strains was compared to growth of the WT VP161 parent strain. All strains were grown separately in HI broth, at 37°C with constant shaking and in biological triplicate. The growth rates of all strains examined were the same as the WT parent indicating that the sRNA mutants along with their overexpression strains display normal growth *in vitro* (data not shown).

Colony morphology changes can indicate dramatic changes in capsule production. To compare the colony morphology of the *P. multocida* sRNA-modified strains with the WT parent strain, each strain was streaked onto HI agar plates and incubated at 37°C for 24 h. Colony morphology characteristics of each of the sRNA strains were indistinguishable from the WT parent, indicating that there was no observable change in the amount of capsule produced (data not shown).

5.3.3.3. Capsule production by the Prrc06, Prrc08 and Prrc25 sRNA mutants

To determine if any of the predicted sRNAs, Prrc06, Prrc08 or Prrc25, played a role in hyaluronic acid (HA) capsule production in *P. multocida*, HA assays were performed on the sRNA mutants. Of particular interest was the role of the sRNAs Prrc08 and Prrc25 that had a similar transcript expression pattern to that of the capsule genes and whose sequence shared complementarity with genes within the capsule locus (section 5.3.1). HA capsule production during mid-exponential growth was quantified for each of the Prrc06, Prrc08 and Prrc25 sRNA mutant strains and compared to capsule produced by the WT parent and two negative control strains; *P. multocida hfq* mutant and the *P. multocida fis* mutant strain (Steen et al., 2010) (Figure 5.7). As expected, the *P. multocida hfq* mutant and the *fis* mutant did not produce significant amounts of HA capsule (Figure 5.7) and the positive control, wild type strain VP161, produced approximately 7.5 µg HA per ml of culture. All three sRNA mutants also produced wild-type levels of capsule during mid-exponential growth, indicating that the Prrc06, Prrc08 and Prrc25 putative sRNAs are not required for capsule production in *P. multocida*.

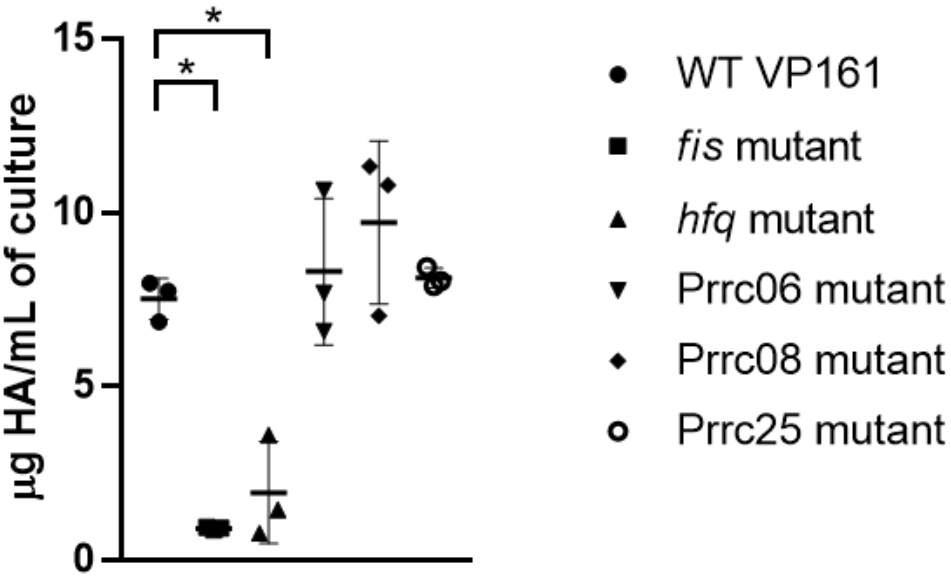


Figure 5.7. Hyaluronic acid (HA) polysaccharide capsule produced by *P. multocida* strains at mid-exponential growth phase. Strains examined were wild-type *P. multocida* strain VP161 (WT), the *P. multocida fis* mutant, the *P. multocida hfq* mutant and the Prrc06, Prrc08 and Prrc25 sRNA mutants. The average amount of HA produced by each strain is indicated by the horizontal bars, ± standard deviations of the means. *, *P* value of < 0.05 as determined by an unpaired *t* test.

5.4. Discussion

Thirty-eight putative *P. multocida* sRNAs were initially identified using bioinformatic analyses of whole genome RNA sequencing data generated from growth under three different conditions, namely aerobically in rich medium, under low iron and under reduced oxygen conditions (Chapter 3). Hfq co-immunoprecipitation analyses including native Hfq co-immunoprecipitation as well as Hfq UV-CRAC/CLASH were also performed to identify the range of Hfq-associated RNA transcripts (Chapter 4). These Hfq-associated transcripts included putative sRNA sequences, some of which were previously identified using bioinformatic analysis of RNA-Seq data, such as the *P. multocida* sRNA GcvB that has since been shown to play a role in amino acid biosynthesis and transport processes (Gulliver et al., 2018). In addition, preliminary studies on the putative *P. multocida* sRNAs HrrF and Prrc03 were also performed (Wright, Boyce Laboratory, unpublished), but no clear phenotype was observed for these sRNAs in *P. multocida* gene regulation. Aside from this preliminary work, no other putative *P. multocida* sRNA candidates had been experimentally validated or characterised.

In silico bioinformatic prediction of sRNAs and webtools for prediction of mRNAs targets can be very useful for the genome-wide investigation of bacterial riboregulatory networks (Li et al., 2012b). However experimental evidence is crucial for identifying *bone fide* sRNA species. Initial evidence includes determining if a predicted sRNA is transcribed from the genome and its actual size. Following this, experiments involving the use of mutants and complemented mutants are required to assess the biological function of a given sRNA and to show an association with specific mRNA targets.

Six putative *P. multocida* sRNAs; namely Prrc06, Prrc07, Prrc08, Prrc10, Prrc12 and Prrc25, were selected for experimental validation based on a range of criteria (Table 5.3). All had a high level of sequence conservation across genomes representing important *P. multocida* capsular serogroup A strains, VP161, X-73 and Pm70. With the exception of Prrc07, the sRNAs chosen were quite highly expressed, as determined from the sRNA-enriched RNA-Seq data from *P. multocida* X-73 cells grown under a range of growth conditions (Chapter 3). In addition, the 3′ end of sRNAs Prrc06, Prrc07 and Prrc12 had a predicted stable hairpin followed by a string of uridine nucleotides, typical of Rho-independent terminator sequences that are often required for transcriptional termination (Lesnik et al., 2001). Prrc25 lacked an obvious 3′ hairpin but had a string of uridine nucleotides at this position. Neither Prrc08 nor Prrc10 had a

clear termination region. It is likely that the sRNAs without clear Rho-independent terminator sequences could instead undergo transcriptional termination via Rho-dependent mechanisms; Rho binding sites (~70-80 nt) are generally rich in cytosine and lack strong secondary structures (Banerjee et al., 2006).

The sRNA species identified during growth in low iron and reduced oxygen (Chapter 3) was also taken into consideration during the selection process. Indeed, the Prrc06 sRNA displayed increased expression during growth in reduced oxygen conditions and may therefore play a role in the regulation of specific genes under this important, infection-relevant, growth condition. In addition, the sRNAs Prrc06, Prrc10, Prrc12 and Prrc25 had altered transcript levels in the *proQ* mutant, compared to those in the wild-type parent (Gulliver, Boyce Laboratory, manuscript in preparation, 2020). The levels of the RNA chaperone ProQ could therefore have an effect on the abundance of these sRNAs via a direct mechanism. Other important data considered during selection of sRNA candidates was interaction with the Hfq protein, as determined by Hfq Co-IP/Hfq UV-CLASH analyses (Chapter 4). As Hfq is an RNA chaperone protein that plays a crucial role in post-transcriptional regulation via sRNA activity (Vogel and Luisi, 2011), any sRNAs which showed a direct association with this protein were of particular interest. Our study showed that the sRNA GcvB clearly interacted with Hfq (Hfq Co-IP analysis), and GcvB is known to be a Hfq-dependent sRNA in other organisms (Sharma et al., 2011). The Prrc10 sRNA was also identified as Hfq-associated from the Co-IP analysis (Table 5.3); electrophoretic mobility shift assays could be employed in future experiments to determine if there is direct binding of Hfq to this sRNA transcript.

The putative sRNAs have been categorised based on their position relative to open reading frames (ORFs) on the *P. multocida* genome; *trans*-acting sRNAs (category 1, encoded in an intergenic region), overlapping sRNAs (category 2, sequence overlapping predicted open reading frames on the same strand) and *cis*-acting sRNAs (short region encoded on the opposite strand of a predicted ORF). Prrc06, Prrc07, Prrc12, Prrc25 are predicted to be *trans*-acting sRNAs (category 1) and Prrc08, Prrc10 are predicted to be overlapping (category 2). No predicted *cis*-encoded sRNAs were selected for further study. These types of sRNAs would be problematic to study as insertional mutagenesis of a *cis*-encoded sRNA would simultaneously result in disruption of the gene encoded on the opposite strand.

The RNA chaperone protein Hfq is involved in the regulation of numerous virulence-associated genes and proteins in *P. multocida* including capsule, LPS and filamentous hemagglutinin

(Chapter 2) (Mégroz et al., 2016). To facilitate this post-transcriptional regulation of virulence factors, *P. multocida* Hfq is likely acting in association with one or more sRNAs to alter mRNA stability or translation, as observed in other bacteria (Sobrero and Valverde, 2012). Secondary effects may also occur, Hfq-mediated regulation of virulence-associated genes could in turn affect the expression/production of other genes and proteins. Therefore, identifying specific sRNAs in the *P. multocida* riboregulatory network directly involved in the regulation of virulence factors, as well as determining precisely how they facilitate this regulation, will significantly contribute to understanding the pathogenesis of *P. multocida*.

The hyaluronic acid capsule is critical for virulence in *P. multocida* serotype A strains (Chung et al., 2001) and inactivation of Hfq results in reduced capsule gene expression and capsule production indicating that Hfq-associated regulation is involved (Mégroz et al., 2016). Expression of capsule genes is highest during early-exponential growth phase and lowest during mid-to-late-exponential growth phase. In turn, the amount of actual capsule produced also fluctuates and is lowest during early-exponential growth and highest during mid-exponential growth (Mégroz et al., 2016); this is likely due to a slow build up of capsule material after the biosynthesis genes are transcribed. To help identify the sRNAs that play a role in the regulation of *P. multocida* capsule expression, the expression pattern of the capsule genes and the sRNAs were compared. Two of the putative sRNAs, Prrc08 and Prrc25, showed a similar expression pattern to the capsule gene expression with highest expression in early-exponential growth phase and lowest expression during mid-exponential growth phase. Known sRNA species interact with target mRNA molecules via a short, and often highly conserved, region of sequence similarity section usually 13-24 bp in length (Peer and Margalit, 2011; Storz et al., 2011). Analysis of Prrc08 and Prrc25 sRNA sequences showed that there was some limited identity with the transcriptional start region of the capsule biosynthesis genes (when reverse complemented), *hyaE* and *phyA*. It is possible that one or both of these identified regions could represent the seed region that allows for direct interaction between the RNA species, facilitating regulation of these capsule genes. Therefore, on the basis of similar gene expression patterns and limited sequence similarity, it was hypothesised that Prrc08 and Prrc25 were involved in the regulation *P. multocida* capsule and were selected for further study.

To precisely define the transcript start site of the selected sRNAs, fluorescent primer extension was used. For four of the six sRNAs, namely Prrc06, Prrc08, Prrc10 and Prrc25, the predicted start sites determined from primer extension matched closely with the start that had been

predicted by visualizing RNA-Seq reads associated with each sRNA, allowing us to confidently define the start site. However, the transcript start site for Prrc12 could not unambiguously be defined using primer extension as the cDNA product terminated at a length that corresponded to approximately halfway through the predicted transcript peak of RNA-Seq reads mapped to the region. It is possible that the reverse transcription of the sRNA sequence during cDNA synthesis was arrested due to strong secondary structure of the Prrc12 sRNA, leading to short transcripts. To further interrogate this, a second primer for Prrc12 encoded past the pause site could be used in additional primer extension analyses, or northern blot analysis for Prrc12 could be performed to determine if the predicted sRNA product size correlates to the Prrc12 sequence length of the RNA-Seq peak. For the putative sRNA Prrc07, no product was generated using primer extension, even after repeating the experiment with freshly isolated RNA samples for cDNA synthesis. It is possible that the 6-FAM labelled primer specific for Prrc07 was not binding efficiently and failed to initiate reverse transcription of the RNA. Alternatively, the region encoding Prrc07 may not actually be transcribed in the manner predicted by the RNA-Seq data. As an addition to primer extension, rapid amplification of cDNA ends (5′ RACE) should be employed. This is considered to be a highly accurate method for defining the transcript start as the 5′ end of the transcript is protected with an adapter (Liu et al., 2018) and was used successfully to precisely define the transcript start site of *P. multocida* GcvB sRNA, which differed by three nucleotides from the predicted start site as defined by fluorescent primer extension (Gulliver et al., 2018).

Following primer extension analysis, three sRNAs were chosen for further experimental validation and phenotypic characterization, namely Prrc06, Prrc08 and Prrc25. In particular, Prrc08 and Prrc25 sRNAs were chosen due to their potential association with capsule regulation. For phenotypic analysis, *P. multocida* sRNA mutant strains were constructed using TargeTron insertional mutagenesis as this is the most effective mutagenesis method available for this bacterium. This method involves the re-targeting on the group II intron to the target gene of interest and for this reason available insertion sites are limited. The seed regions in Prrc06, Prrc08 and Prrc25 that bind to the mRNA target are unknown and while the 5′ ends of some sRNAs studied in other species are known to bind to the mRNA target, the seed region can also be located at other positions (Peer and Margalit, 2011). The seed region of *P. multocida* GcvB is encoded towards the 5′ end of the sRNA, across nucleotides 54-63 of the sequence which is 183 nucleotides in length in total (Gulliver et al., 2018). To try and inactivate the seed region in each selected sRNA, all introns designed for TargeTron mutagenesis were

targeted within the first 77 nucleotides of Prrc06, Prrc08 and Prrc25. Though it was unknown if the binding activity of each sRNA was disrupted by the intron insertion, the large size of the inserted TargeTron intron (~1.9 kb) would have significantly disrupted the secondary structure of the molecules which ranged in size from 157 to183 bps. Thus, it was likely that insertional mutagenesis would result in functional inactivation of each sRNA.

Northern blot analysis was used to confirm the expression and size of the selected putative sRNAs Prrc06, Prrc08 and Prrc25. The probe specific for sRNA Prrc06, predicted by primer extension to be 160 bp in length, hybridised strongly with a band in all lanes except for the lane containing RNA isolated from the Prrc06 mutant. The hybridizing band was very strong and appeared slightly larger (~220 bp compared to ~200 bp) in strains harbouring a copy of the sRNA on the plasmid, perhaps indicating a larger transcript size was being generated. However, this may also be an artefact of more copies of the hybridizing molecule being present in the sample. The size of the hybridised fragments was also difficult to ascertain as only some of the bands in the ssRNA ladder were clearly visible. Further analysis such as fluorescent primer extension or 5′ RACE using the plasmid as template, would be required to confirm the transcript start for the recombinant Prrc06 was the same as the wild-type copy of Prrc06 on the genome. The probe specific for sRNA Prrc08, predicted by primer extension to be 183 bp in length, hybridised strongly with a fragment of this approximate size in lanes representing strains harbouring recombinant Prrc08, and weakly hybridised with a fragment of the expected size in the WT strain but there were no hybridising bands in the lane representing the WT strain harbouring the empty expression vector. As expected, no hybridisation was observed in the lane containing RNA isolated from the Prrc08 mutant strain. Analysis of bioinformatic data revealed that Prrc08 has a higher level of expression during early-exponential growth phase in comparison to mid-exponential growth phase, when cells were harvested for RNA isolation and Prrc08 expression is naturally lower. Thus, isolating RNA from WT cells at early-exponential growth phase may have increased the abundance of this transcript and thus increased the hybridisation signal. For Prrc25, no hybridised fragments were observed in any of the strains tested. Although the Prrc25 riboprobe was efficiently DIG-labelled, the observation of lack of hybridisation to any of the RNA samples tested for Prrc25 could be the result of many experimental factors including degradation of RNA samples tested during the northen blotting procedure. It is possible that Prrc25 is not a true sRNA; however, evidence from the initial primer extension analysis strongly suggests otherwise. A clear primer extension product was generated for this transcript and the start site of this product aligned closely to the

start of the RNA-Seq read peak (page 179). Future work should involve optimisation of northern blotting for Prrc25. In future analyses, northern blots should also be performed on the strains expressing antisense sRNA to confirm that their expression inactivates Prrc06, Prrc08 or Prrc25. This is especially important if inactivation of selected sRNAs is unsuccessful.

Each of the sRNA mutants along with their overexpression strains displayed normal growth *in vitro* in rich medium at 37°C. Initial phenotypic analysis of the sRNA mutants focussed on capsule production as strains that are acapsular produce smaller and less mucoid colonies. The colony morphology of each sRNA mutant was compared to wild type parent (VP161), VP161 containing a copy of the sRNA (overexpression strain) or an antisense copy of the sRNA (antisense strain). All strains produced a normal wild-type colony morphology, indicating wild-type levels of capsule were being produced. Given the lack of any clear morphological differences in capsule production, HA assays were performed on only a sample of strains; namely the sRNA mutant strains, which each showed normal capsule production. Although it was hypothesised that sRNAs Prrc08 and Prrc25 may play a role in the regulation of *P. multocida* capsule production based on initial bioinformatic analyses, this seems unlikely given that the Prrc08 and Prrc25 mutants produced capsule equivalent to wild-type VP161. It is possible that one or more of the 35 other putative sRNA species identified may play a role in capsule production and others therefore should be investigated and characterised. In particular, though TargetRNA2 (Kery et al., 2014) is a limited prediction tool of unknown accuracy, it may be of interest to investigate Prrc24 (*hyaC*), Prrc27 (*hexC*), Prrc29 (*phyA*) or Prrc52 (*hexB*) that had capsule genes as predicted mRNA targets. Current work in the Boyce Laboratory is focussed on transposon-directed insertion site sequencing (TraDIS) as a method of identifying genes essential for normal capsule production (Smallman, Boyce Laboratory, unpublished). It is possible that this independent study will identify capsule-associated sRNAs. Though unlikely, it is also possible that Hfq could be acting independently of sRNAs to post-transcriptionally regulate capsule production in *P. multocida*. Binding analysis of the Hfq protein with specific capsule gene transcripts using techniques such as electrophoretic mobility shift assays could be used to test a possible direct association with Hfq and the capsule gene transcripts. In addition, future work could involve phenotypic analysis of the Prrc06 sRNA mutant under reduced oxygen conditions, as the sRNA displayed increased expression during reduced oxygen growth (Chapter 3) and may play a role in regulation under this specific condition.

This chapter reported the validation and characterisation of a number of identified *P. multocida* sRNAs. The transcriptional start site of a number of these sRNAs was successfully determined using fluorescent primer extension, and initial northern blot analysis was performed which confirmed the expression of sRNAs Prrc06 and Prrc08; these are only the second and third sRNAs confirmed in *P. multocida*. However, a clear role for the sRNAs Prrc06, Prrc08 and Prrc25 in *P. multocida* capsule production could not be determined following our initial phenotypic analysis and thus it is likely that they regulate other genes/systems in *P. multocida*. Future studies to determine the precise targets of the sRNAs investigated in this study will focus on both transcriptomic and proteomic analyses of the sRNA mutants and their corresponding overexpression strains. The utilisation of two methods measuring both RNA expression and protein production will be necessary to determine if there is altered gene expression and/or altered protein production in each mutant and/or overexpression strain. Using this method, it is hoped that the role of each sRNA in RNA stability and/or translational efficiency can be identified. As a single sRNA can have multiple mRNA targets, these large-scale analyses will be especially important. Pulse-expression studies of the sRNAs could also be used in parallel as a means of identifying the direct targets of the sRNAs. It will also be important to determine the precise region that each sRNA interacts with its mRNA target (seed region) so that other direct targets of each sRNA can be predicted. Future validation and characterisation studies of *P. multocida* sRNAs will still largely focus on understanding the role of sRNAs in the regulation of virulence.

Chapter Six: General discussion and future directions

Pasteurella multocida is the causative agent of a number of economically important animal diseases. While numerous *P. multocida* virulence factors have been identified, including capsule, lipopolysaccharide (LPS) and filamentous hemagglutinin, little is known about how the expression of these virulence factors is regulated. Small non-coding RNA molecules (sRNAs) are important regulators of bacterial gene expression and protein production, with essential roles in controlling diverse bacterial functions including virulence (Gottesman and Storz, 2011). The RNA chaperone Hfq interacts with small noncoding RNA molecules (sRNAs) and their mRNA targets to facilitate post-transcriptional regulation (Sobrero and Valverde, 2012). The work presented in this thesis explored the role of the Hfq protein and sRNAs in *P. multocida*, specifically focusing on riboregulatory functions that impact *P. multocida* pathogenesis.

Hfq plays an important role in mediating the various post-transcriptional riboregulatory mechanisms of sRNAs at the level of translational efficiency and RNA stability (Vogel and Luisi, 2011). Hfq can also facilitate regulation independently of sRNAs via various mechanisms. The Hfq protein was characterised in *P. multocida* (Chapter 2) as it is known that in other species this RNA chaperone plays a critical role in virulence and fitness. When compared to the wild-type *P. multocida* parent strain, the *P. multocida hfq* mutant produced almost no hyaluronic acid capsule during all growth phases tested, proving that Hfq was involved in capsule production. The mutant also displayed reduced *in vivo* fitness in mice, and had reduced levels of virulence in chickens, a natural host of this bacterium. These results indicate that the *hfq* mutant had a reduced ability to avoid the host innate immune system and to grow in the host environment. It is known that capsule is required for growth *in vivo* and therefore the lack of capsule on the *hfq* mutant is likely largely responsible for the *in vivo* phenotype displayed by the mutant. However, transcriptomic and proteomic analyses of the *hfq* mutant indicated that the expression of other known known virulence factors, including filamentous hemagglutin and the LPS modifying enzymes phosphoethanolamine transferase (PM1042), and PcgB and PcgD (phosphocholine addition) were also affected by the loss of Hfq. Therefore, reductions in the level of filamentous hemagglutinin adhesin on the bacterial surface and/or changes to LPS modifications may also have played a part in the reduced *in vivo* fitness. Taken together, these data indicate that Hfq and associated sRNAs are likely to be critical regulators of *P. multocida* virulence factors. As little was known about sRNAs or the riboregulatory network in *P. multocida* it was uncertain which of the observed changes in the

hfq mutant were the result of direct interaction of sRNAs/mRNAs with Hfq and which were secondary effects as a result of Hfq-mediated regulatory changes.

At the beginning of this project, there was very limited data on *P. multocida* sRNAs and, with the exception of the GcvB study conducted by Emily Gulliver in our laboratory (Gulliver et al., 2018), to date there is no published data on experimentally verified sRNA species in this organism. Therefore, our initial analyses focused on the identification of novel predicted *P. multocida* sRNA molecules using bioinformatic analysis of whole genome RNA-Sequencing data (Chapter 3). As sRNA expression may be growth-phase dependent or limited only to particular growth conditions, a range of *P. multocida* RNA-Seq datasets were analysed to maximise sRNA discovery. To promote sRNA discovery in these analyses, the RNA samples prepared during this work were enriched for small RNA fragments (<200 bp). In total, 38 sRNA species were identified by conducting several analyses. Firstly, RNA-Seq data from the closely related wild-type *P. multocida* strains VP161 and X-73, grown aerobically in rich medium to various growth phases were analysed. RNA-Seq data was also generated from strain X-73 grown under iron-limited and low oxygen conditions as these conditions are predicted to mimic *in vivo* niches, based on previous gene expression data generated from growth of *P. multocida* in chickens (Boyce et al., 2002). A number of putative *P. multocida* sRNAs showed altered transcript levels during growth under the iron-limited or low oxygen conditions and therefore could play a role in adapting *P. multocida* to the specific environmental conditions. The GcvB sRNA was identified during the analyses of whole genome RNA-Seq data and displayed altered expression under low oxygen conditions. Studies in the Boyce laboratory carried out in parallel with the studies reported in this thesis have since defined a clear role for GcvB in amino acid biosynthesis and transport processes in *P. multocida* (Gulliver et al., 2018) In addition, the sRNA HrrF was also identified and was most highly expressed during growth under iron-limited conditions, suggesting that this sRNA plays an important role in iron-associated regulation in *P. multocida*. Orthologues of HrrF in other members of the *Pasteurellaceae* family have been proven to be involved in iron homeostasis; *H. influenzae* Hrrf is predicted to control the expression of genes involved with molybdate uptake, deoxyribonucleotide and amino acid biosynthesis (Santana et al., 2014). Notably, the *P. multocida* HrrF sequence contains a 32 nucleotide region that shares 100% sequence conservation with HrrF from all members examined within the *Pasteurellaceae*, indicating that this region in HrrF may be important for sRNA function (Santana et al., 2014). To prove the role of HrrF in *P. multocida*, experiments should be performed to identify the mRNA targets such as additional RNA-Seq

experiments, which were used successfully to identify HrrF targets in *H. influenzae*. Many more sRNAs have been identified in other bacterial species and we therefore predict that our analysis has not captured all of the sRNAs expressed by *P. multocida*. Additional sRNAs may be discovered by growing *P. multocida* in other conditions, including growth in chemically defined media, growth in serum and growth *in vivo* in the natural chicken host. However, it is also possible that *P. multocida* has a limited regulatory network with only a small number of sRNAs. Indeed, *P. multocida* has a small genome compared to the widely studies model organisms *E. coli, S.* Typhimurium and *P. aeruginosa* (~2000 genes compared to ~4300, ~5000 and ~5500 genes, respectively). Furthermore, *P. multocida* is predicted to have only a small number of two component signal transduction systems (~8 compared with ~30, ~20 and ~50, respectively) so it may also possess a reduced number of sRNAs (Gumerov et al., 2019).

The mRNA targets of the 38 putative sRNAs were predicted using the bioinformatic tool, TargetRNA2 (Kery et al., 2014). Some of the bioinformatically predicted target mRNAs were virulence-associated, including the iron acquisition gene, *tonB*, the filamentous hemagglutinin gene, *pfhB2*, several capsule biosynthesis/transport genes (*hyaC, phyA, hexBC*) and two genes encoding outer membrane lipoproteins (*plpB* and *plpE*). However, sRNA target prediction databases are known to have a low accuracy of prediction and therefore caution is required when trying to define the mRNA targets using only *in silico* methods. Experimentally identifying the sRNA targets is crucial for defining the true regulatory network in *P. multocida*, so future analyses should include generating RNA-Seq expression data from individual sRNA mutants, the corresponding overexpression strains, as well as strains engineered for pulse expression of sRNAs; in preliminary work a range of these analyses were performed for a small seletion of sRNAs as described in Chapter 5. In addition, global small non-coding RNA target identification by ligation and sequencing (GRIL-Seq) methods could be employed to identify direct sRNA-targets (Han et al., 2016). GRIL-Seq analysis involves the *in vivo* ligation of sRNAs to their targets following the co-expression of a specific sRNA with a T4 RNA ligase protein; sRNA targets are then identified by high-throughput sequencing.

Since the start of this project, a second RNA chaperone has been identified called ProQ. To examine the effect of inactivating these chaperones on sRNA expression, RNA-Seq data derived from *P. multocida hfq* (VP161, Chapter 2) and *proQ* (VP161, Gulliver, Boyce Laboratory, manuscript in preparation, 2020) mutants were examined. Interestingly, only two putative sRNAs had altered transcript levels in the *hfq* mutant, compared to those in the wild-

type parent. In contrast, fifteen putative sRNAs had altered transcript levels in the *proQ* mutant, compared to those in the wild-type parent (Gulliver, Boyce Laboratory, manuscript in preparation, 2020). These results indicate that the levels of the RNA chaperones Hfq and ProQ could have a direct effect on the abundance of a limited number of the identified sRNAs and that ProQ may have a more significant effect in *P. multocida* than Hfq.

As Hfq is known to directly interact with sRNAs to facilitate riboregulation, Hfq co-immunoprecipitation (Co-IP) analyses were employed (Chapter 4) to identify the range of Hfq-associated RNA species and to compare this data to the RNA-Seq expression data. RNA species that bind Hfq include sRNAs and mRNAs. A total of 90 transcripts were identified as Hfq-associated using native Hfq Co-IP, including GcvB which is known to be Hfq-dependent in other bacterial species. Two other predicted sRNAs, Prrc10 and Prrc19, co-immunoprecipitated with Hfq as well as a range of mRNA transcripts, including those encoding the outer membrane lipoprotein PlpB, a predicted membrane-bound lysozyme-inhibitor, the phosphocholine (PCho) transferase PcgD and the preprotein translocase subunit SecG. Ten of the mRNA transcripts associating with Hfq also displayed increased expression in the *hfq* mutant (Chapter 2); these included PlpB and the hypothetical protein, X-73_RS05335, that also displayed increased protein production during both early- and mid-exponential growth in the absence of *hfq* (Mégroz et al., 2016). Taken together, these data strongly suggest that Hfq destabilizes or facilitates degradation of these transcripts, with or without the action of an associated sRNA.

Hfq targets were further investigated using an independent co-immunprecipitation technique, Hfq UV-CRAC/CLASH. This method uses UV light to covalently link Hfq and any bound RNA species; it was predicted that this would increase the number of identified RNAs. Furthermore, the UV-CLASH component of these analyses allow any RNA species that simultaneously associate with Hfq to be ligated together and then identified by sequencing the hybrid molecules. Reads in the UV-CRAC data directly associated with GcvB sRNA were visually inspected and a high number of reads were generated from the tagged replicates (in comparison to untagged replicates), supporting previous evidence that indicates GcvB is Hfq-dependent in *P. multocida*. Surprisingly, the number of sRNAs identified via UV-CRAC was limited. However, three Hfq-associated transcripts were identified from this analysis, namely a *ner* transcript, encoding a predicted *P. multocida* transcriptional regulator, *glmU*, involved in the hyaluronic acid (HA) biosynthesis pathway in other bacteria, and a transcript corresponding

to an intergenic region located immediately upstream of a *P. multocida* protein of unknown function (PMVP_0398). Interestingly, none of the three transcripts identified using UV-CRAC were identified in the native Hfq Co-IP analysis; however, this may be due to the transient nature of many Hfq-RNA interactions. Additional Co-IP analyses using *P. multocida* cells grown under a range of conditions should be performed in future, such as growth in iron-limited or low oxygen conditions. These additional studies would allow for a more comprehensive understanding of the *P. multocida* Hfq-interactome and help to identify novel Hfq-associated sRNAs that may only be expressed under these specific conditions. Seventeen Hfq associated RNA:RNA hybrids were also identified using the UV-CLASH analysis, each predicted to comprise of RNA molecules interacting with each other and Hfq simultaneously. Identified interactions included the Prrc03 sRNA with the tRNA Ser, suggesting that Prrc03 may be involved in the post-transcriptional regulation of this tRNA.

It will be important in future work to experimentally confirm that Hfq can directly bind GcvB, Prrc10 and Prrc19 sRNA as well as the virulence-associated transcripts *plpB* and *pcgD*. Experiments that would help determine if Hfq (or *hfq* transcript) can directly bind these specific RNA molecules include techniques such as biolayer interferometry (BLItz), electrophoretic mobility shift assay (EMSA) or surface plasmon resonance (SPR) (Dimastrogiovanni et al., 2014; Fender et al., 2010; Osborne et al., 2014). For these analyses, Hfq would need to be expressed and purified and the putative RNAs produced by *in vitro* transcription. In preliminary work, Hfq has been expressed and purified as a soluble HexaHis-tagged protein (Wright, Boyce Laboratory, pers. comm.) so these experiments are feasible. These analyses could also be applied to investigate any specific interactions between hybrid RNA:RNA molecules identified using the UV-CLASH experiments, such as the interaction between Prrc03 and the tRNA Ser transcript.

Bioinformatic analyses of transcriptional expression alone are not sufficient to identify *bona fide* sRNA species in *P. multocida*. Accordingly, a small number of the 38 putative sRNAs identified were selected for experimental validation and functional characterization (Chapter 5). These sRNAs were selected based on strong bioinformatic data; all were short in nature and in most cases highly expressed in the sRNA-enriched RNA-Seq dataset. The putative Prrc06 sRNA showed reduced expression during growth under reduced oxygen conditions, while a number of the selected sRNAs (Prrc06, Prrc09, Prrc10 and Prrc12) also showed altered expression in the *proQ* mutant compared to the wild-type strain. The Prrc10 sRNA displayed

an association with Hfq using Co-IP and may therefore be a Hfq-dependent sRNA. The Prrc07 sRNA is a homologue of sRNA 20 (AaHKsRNA20) identified in *A. actinomycetemcomitans* (Oogai et al., 2017). Prrc08 and Prrc25 were of particular interest due to their growth-phase dependent expression pattern partially correlating with capsule production, and potential sequence similarity with a number of the capsule biosynthesis genes. The transcriptional start site for four of the six selected sRNAs was identified by fluorescent primer extension and the identified start sites closely matched the predictions from the RNA-Seq data; this is strong evidence that these predicted sRNAs are real. Northern blotting of sRNA mutants and overexpression strains, along with the appropriate control strains, was also performed to check the size and expression of some of the selected sRNAs. In future experiments, rapid amplification of cDNA ends (5′ RACE) analysis could be employed for sRNA validation in addition to primer extension, as this is considered to be a highly accurate method for defining the transcript start as the 5′ end of the transcript is protected with an adapter (Liu et al., 2018). Initial phenotypic studies of the sRNAs failed to identify any major phenotypic changes. However, the experiments were very limited and further analysis of these sRNAs and their targets is of importance. Ideally, all 38 sRNAs should be experimentally validated using these and other methods in order to determine their biological significance.

This study has clearly demonstrated that *P. multocida* Hfq plays an essential role in the regulation of HA capsular polysaccharide biosynthesis. Fis is also critical for capsule production in *P. multocida* and the loss of Fis leads to decreased transcription of the capsule locus and an acapsular phenotype (Steen et al., 2010). Therefore, a lack of Fis or Hfq leads to the loss of capsule production. Indeed, a reduction in capsule expression and production by the *hfq* mutant was evident from transcriptomic, proteomic as well as phenotypic analyses. Fis transcript expression was increased in the *hfq* mutant, relative to levels in the wild-type parent, but this did not correlate with translation as proteomics indicated that the levels of Fis protein in the *hfq* mutant were unchanged. This lack of correlation suggests that Fis acts independently of Hfq to regulate capsule production. We hypothesized that Hfq, with or without sRNA involvement, acts to positively regulate capsule production at the transcriptional level, either through degradation or destabilisation of the capsule transcripts. However, the Hfq Co-IP data did not detect any Hfq-associated transcripts corresponding to any of the ten genes within the capsule biosynthesis locus. However, *glmU* was identified as Hfq-associated in the UV-CRAC experiments and the GlmU protein is involved in the HA biosynthesis pathway. It was therefore hypothesised that Hfq may promote GlmU production, which in turn increases the expression

of HA capsule biosynthesis genes and subsequently capsule production. However, when experiments were conducted using a GlmU overexpression plasmid in the *P. multocida hfq* mutant, capsule production was not restored in the strain. However, it is possible that GlmU transcription requires Hfq for normal translation, so it be worthwhile in future work to assess if there is any direct interaction between Hfq and the GlmU transcript. Alternatively, Hfq may facilitate regulation of capsule production via another as yet unidentified mechanism.

Two of the sRNAs selected for experimental validation and characterisation in this study, Prrc08 and Prrc25, showed a generally similar growth-phase dependent expression pattern to the capsule biosynthesis genes, and also showed a limited amount of sequence identity (reverse complement) to the capsule genes *hyaE* and *phyA*. Although Co-IP analyses failed to show an association of Prrc08 and Prrc25 with Hfq, it was proposed that they may be involved in the regulation of the capsule biosynthesis genes. However, independent TargeTron mutants of the two sRNAs expressed normal levels of capsule. It is possible that Hfq could be acting independently of sRNAs to post-transcriptionally regulate capsule production in *P. multocida*. Binding analysis of the Hfq protein with specific capsule gene transcripts using techniques such as EMSA could be used to test a possible direct association with Hfq and the capsule gene transcripts. Alternatively, and perhaps more likely, capsule regulation requires Hfq to interact with another sRNA that we have not yet identified. In particular, though TargetRNA2 is a limited prediction tool of unknown accuracy, it may still be of interest to investigate Prrc24 (*hyaC*), Prrc27 (*hexC*), Prrc29 (*phyA*) or Prrc52 (*hexB*) that TargetRNA2 predicted interacted with capsule genes. Current work in the Boyce laboratory is focussed on transposon-directed insertion site sequencing (TraDIS) as a method of identifying genes essential for normal capsule production (Thomas Smallman, Boyce laboratory, unpublished). It is possible that this independent study will identify capsule-associated sRNAs.

This study has allowed us to investigate regulation of *P. multocida* gene expression via the Hfq chaperone and sRNAs. Future work will focus on the elucidating the precise mechanism by which Hfq regulates virulence factors, and in particular the regulation of capsule production in *P. multocida;* this should include both identification and characterisation of the specific sRNAs involved. The identification of sRNAs with specific virulence-associated targets may allow us to inhibit their activity using DNA mimics such as peptide-conjugated phosphorodiamidate morpholino oligomers (PPMOs) (Geller et al., 2013). Inhibition of virulence-associated sRNAs via PPMOs could be an important therapeutic tool to control *P. multocida* pathogenesis. Taken

together, this work has contributed to the understanding of the pathogenesis of *P. multocida* via riboregulation. Future studies will assist in unravelling the complexities of the *P. multocida* regulatory network.

References

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., Ishihama, A., 1999. Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. Journal of Bacteriology 181, 6361-6370.

Allam, U.S., Krishna, M.G., Lahiri, A., Joy, O., Chakravortty, D., 2011. *Salmonella enterica* serovar Typhimurium lacking *hfq* gene confers protective immunity against murine typhoid. PLoS ONE 6.

Ansong, C., Yoon, H., Porwollik, S., Mottaz-Brewer, H., Petritis, B.O., Jaitly, N., Adkins, J.N., McClelland, M., Heffron, F., Smith, R.D., 2009. Global systems-level analysis of Hfq and SmpB deletion mutants in *Salmonella*: Implications for virulence and global protein translation. PLoS ONE 4.

Attia, A.S., Sedillo, J.L., Wang, W., Liu, W., Brautigam, C.A., Winkler, W., Hansen, E.J., 2008. *Moraxella catarrhalis* expresses an unusual Hfq protein. Infection and Immunity 76, 2520-2530.

Backstrand, J.M., Botzler, R.G., 1986. Survival of *Pasteurella multocida* in soil and water in an area where avian cholera is enzootic. Journal of Wildlife Diseases 22, 257-259.

Bailey, T.L., Johnson, J., Grant, C.E., Noble, W.S., 2015. The MEME Suite. Nucleic Acids Research 43, W39-W49.

Baillot, R., Voisine, P., Côté, L.M.E.G., Longtin, Y., 2011. Deep sternal wound infection due to *Pasteurella multocida*: The first case report and review of literature. Infection 39, 575-578.

Banerjee, S., Chalissery, J., Bandey, I., Sen, R., 2006. Rho-dependent transcription termination: more questions than answers. Journal of Microbiology 44, 11-22.

Barquist, L., Vogel, J., 2015. Accelerating discovery and functional analysis of small RNAs with new technologies. Annual Review of Genetics 49, 367-394.

Beich-Frandsen, M., Večerek, B., Konarev, P.V., Sjöblom, B., Kloiber, K., Hämmerle, H., Rajkowitsch, L., Miles, A.J., Kontaxis, G., Wallace, B.A., Svergun, D.I., Konrat, R., Bläsi, U., Djinović-Carugo, K., 2011. Structural insights into the dynamics and function of the C-terminus of the *E. coli* RNA chaperone Hfq. Nucleic Acids Research 39, 4900-4915.

Benkirane, A., De Alwis, M.C.L., 2002. Haemorrhagic septicaemia, its significance, prevention and control in Asia. Veterinarni Medicina 47, 234-240.

Bilusic, I., Popitsch, N., Rescheneder, P., Schroeder, R., Lybecker, M., 2014. Revisiting the coding potential of the *E. coli* genome through Hfq co-immunoprecipitation. RNA Biology 11, 641-654.

Blackall, P.J., Miflin, J.K., 2000. Identification and typing of *Pasteurella multocida*: A review. Avian Pathology 29, 271-287.

Bobrovskyy, M., Vanderpool, C.K., 2013. Regulation of bacterial metabolism by small RNAs using diverse mechanisms, Annual Review of Genetics, pp. 209-232.

Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S., Heck, A.J.R., 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nature Protocols 4, 484-494.

Bohn, C., Rigoulay, C., Bouloc, P., 2007. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. BMC Microbiology 7.

Bosch, M., Garrido, E., Llagostera, M., Pérez De Rozas, A.M., Badiola, I., Barbé, J., 2002a. *Pasteurella multocida exbB, exbD* and *tonB* genes are physically linked but independently transcribed. FEMS Microbiology Letters 210, 201-208.

Bosch, M., Garrido, M.E., Llagostera, M., Pérez de Rozas, A.M., Badiola, I., Barbé, J., 2002b. Characterization of the *Pasteurella multocida hgbA* gene encoding a hemoglobin-binding protein. Infection and Immunity 70, 5955-5964.

Bosdriesz, E., Molenaar, D., Teusink, B., Bruggeman, F.J., 2015. How fast-growing bacteria robustly tune their ribosome concentration to approximate growth-rate maximization. The FEBS Journal 282, 2029-2044.

Boucher, D.J., Adler, B., Boyce, J.D., 2005. The *Pasteurella multocida nrfE* Gene Is Upregulated during Infection and Is Essential for Nitrite Reduction but Not for Virulence. Journal of Bacteriology 187, 2278-2285.

Boudry, P., Gracia, C., Monot, M., Caillet, J., Saujet, L., Hajnsdorf, E., Dupuy, B., Martin-Verstraete, I., Soutourina, O., 2014. Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen *Clostridium difficile*. Journal of Bacteriology 196, 3234-3248.

Boyce, J.D., Adler, B., 2000. The capsule is a virulence determinant in the pathogenesis of *Pasteurella multocida* M1404 (B:2). Infection and Immunity 68, 3463-3468.

Boyce, J.D., Adler, B., 2006. How does *Pasteurella multocida* respond to the host environment? Current Opinion in Microbiology 9, 117-122.

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, 870-880.

Boyce, J.D., Harper, M., St Michael, F., John, M., Aubry, A., Parnas, H., Logan, S.M., Wilkie, I.W., Ford, M., Cox, A.D., Adler, B., 2009. Identification of novel glycosyltransferases required for assembly of the *Pasteurella multocida* A:1 lipopolysaccharide and their involvement in virulence. Infection and Immunity 77, 1532-1542.

Boyce, J.D., Harper, M., Wilkie, I.W., Adler, B., 2010. *Pasteurella* in: Gyles, C.L., Prescott, J.F., Songer, J.G., Thoen, C.O. (Eds.), Pathogenesis of bacterial infections of animals, 4th ed. Blackwell Publishing, Iowa, pp. 325-346.

Boyce, J.D., Wilkie, I., Harper, M., Paustian, M.L., Kapur, V., Adler, B., 2002. Genomic scale analysis of *Pasteurella multocida* gene expression during growth within the natural chicken host. Infection and Immunity 70, 6871-6879.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein dye binding. Analytical Biochemistry 72, 248-254.

Brantl, S., 2007. Regulatory mechanisms employed by *cis*-encoded antisense RNAs. Current Opinion in Microbiology 10, 102-109.

Brennan, R.G., Link, T.M., 2007. Hfq structure, function and ligand binding. Current Opinion in Microbiology 10, 125-133.

Busch, A., Richter, A.S., Backofen, R., 2008. IntaRNA: efficient prediction of bacterial sRNA targets incorporating target site accessibility and seed regions. Bioinformatics 24, 2849-2856.

Caldelari, I., Chao, Y., Romby, P., Vogel, J., 2013. RNA-mediated regulation in pathogenic bacteria. Cold Spring Harbor Perspectives in Medicine 3.

Carpenter, T.E., Snipes, K.P., Wallis, D., McCapes, R., 1988. Epidemiology and financial impact of fowl cholera in turkeys: a retrospective analysis. Avian Diseases 32, 16-23.

Carter, G.R., 1952. The type specific capsular antigen of *Pasteurella multocida*. Canadian Journal of Medical Sciences 30, 48-53.

Carter, G.R., 1958. Some characteristics of Type A strains of *Pasteurella multocida*. The British Veterinary Journal 114, 356-357.

Caspi, R., Altman, T., Billington, R., Dreher, K., Foerster, H., Fulcher, C.A., Holland, T.A., Keseler, I.M., Kothari, A., Kubo, A., Krummenacker, M., Latendresse, M., Mueller, L.A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D.S., Weerasinghe, D., Zhang, P., Karp, P.D., 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. Nucleic Acids Research 42, D459–D471.

Chan, C.L., Landick, R., 1989. The *Salmonella typhimurium his* operon leader region contains an RNA hairpin-dependent transcription pause site. Mechanistic implications of the effect on pausing of altered RDNA hairpins. Journal of Biological Chemistry 264, 20796-20804.

Chao, Y., Papenfort, K., Reinhardt, R., Sharma, C.M., Vogel, J., 2012. An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs. The EMBO Journal 31, 4005-4019.

Chao, Y., Vogel, J., 2010. The role of Hfq in bacterial pathogens. Current Opinion in Microbiology 13, 24-33.

Chen, J., Gottesman, S., 2017. Hfq links translation repression to stress-induced mutagenesis in *E. coli*. Genes & Development 31, 1382-1395.

Chiang, M.K., Lu, M.C., Liu, L.C., Lin, C.T., Lai, Y.C., 2011. Impact of Hfq on global gene expression and virulence in *Klebsiella pneumoniae*. PLoS ONE 6.

Christiansen, J.K., Larsen, M.H., Ingmer, H., Søgaard-Andersen, L., Kallipolitis, B.H., 2004. The RNA-binding protein Hfq of *Listeria monocytogenes*: Role in stress tolerance and virulence. Journal of Bacteriology 186, 3355-3362.

Christiansen, K.H., Carpenter, T.E., Snipes, K.P., Hird, D.W., 1992. Transmission of *Pasteurella multocida* on California turkey premises in 1988-89. Avian Diseases 36, 262-271.

Chung, J.Y., Wilkie, I., Boyce, J.D., Adler, B., 2005. Vaccination against fowl cholera with acapsular *Pasteurella multocida* A:1. Vaccine 23, 2751-2755.

Chung, J.Y., Wilkie, I., Boyce, J.D., Townsend, K.M., Frost, A.J., Ghoddusi, M., Adler, B., 2001. Role of capsule in the pathogenesis of fowl cholera caused by *Pasteurella multocida* serogroup A. Infection and Immunity 69, 2487-2492.

Chung, J.Y., Zhang, Y.M., Adler, B., 1998. The capsule biosynthetic locus of *Pasteurella multocida* A-1. FEMS Microbiology Letters 166, 289-296.

Cole, J., 1996. Nitrate reduction to ammonia by enteric bacteria: Redundancy, or a strategy for survival during oxygen starvation? FEMS Microbiology Letters 136, 1-11.

Constantinidou, C., Hobman, J.L., Griffiths, L., Patel, M.D., Penn, C.W., Cole, J.A., Overton, T.W., 2006. A reassessment of the FNR regulon and transcriptomic analysis of the effects of nitrate, nitrite, NarXL, and NarQP as *Escherichia coli* K12 adapts from aerobic to anaerobic growth. Journal of Biological Chemistry 281, 4802-4815.

Costanzo, J.T., Wojciechowski, A.L., Bajwa, R.P.S., 2017. Urinary tract infection with *Pasteurella multocida* in a patient with cat exposure and abnormal urinary tract physiology: Case report and literature review. IDCases 9, 109-111.

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature Biotechnology 26, 1367-1372.

Cui, M., Wang, T., Xu, J., Ke, Y., Du, X., Yuan, X., Wang, Z., Gong, C., Zhuang, Y., Lei, S., Su, X., Wang, X., Huang, L., Zhong, Z., Peng, G., Yuan, J., Chen, Z., Wang, Y., 2013. Impact of Hfq on global gene expression and intracellular survival in *Brucella melitensis*. PloS ONE 8.

De Lay, N., Schu, D.J., Gottesman, S., 2013. Bacterial small RNA-based negative regulation: Hfq and its accomplices. Journal of Biological Chemistry 288, 7996-8003.

De Lay, N.R., Garsin, D.A., 2016. The unmasking of 'junk' RNA reveals novel sRNAs: from processed RNA fragments to marooned riboswitches. Current Opinion in Microbiology 30, 16-21.

de Oliveira, J.D., Carvalho, L.S., Gomes, A.M.V., Queiroz, L.R., Magalhães, B.S., Parachin, N.S., 2016. Genetic basis for hyper production of hyaluronic acid in natural and engineered microorganisms. Microbial Cell Factories 15, 119.

DeAngelis, P.L., Gunay, N.S., Toida, T., Mao, W.J., Linhardt, R.J., 2002. Identification of the capsular polysaccharides of Type D and F *Pasteurella multocida* as unmodified heparin and chondroitin, respectively. Carbohydrate Research 337, 1547-1552.

Deeb, B.J., DiGiacomo, R.F., Bernard, B.L., Silbernagel, S.M., 1990. *Pasteurella multocida* and *Bordetella bronchiseptica* infections in rabbits. Journal of Clinical Microbiology 28, 70-75.

Dietrich, M., Munke, R., Gottschald, M., Ziska, E., Boettcher, J.P., Mollenkopf, H., Friedrich, A., 2009. The effect of *hfq* on global gene expression and virulence in *Neisseria gonorrhoeae*. The FEBS Journal 276, 5507-5520.

DiGiacomo, R.F., Deeb, B.J., Giddens Jr, W.E., Bernard, B.L., Chengappa, M.M., 1989. Atrophic rhinitis in New Zealand white rabbits infected with *Pasteurella multocida*. American Journal of Veterinary Research 50, 1460-1465.

Dimastrogiovanni, D., Fröhlich, K.S., Bandyra, K.J., Bruce, H.A., Hohensee, S., Vogel, J., Luisi, B.F., 2014. Recognition of the small regulatory RNA RydC by the bacterial Hfq protein. Elife 3, e05375.

Ding, Y., Davis, B.M., Waldor, M.K., 2004. Hfq is essential for *Vibrio cholerae* virulence and downregulates σE expression. Molecular Microbiology 53, 345-354.

Douangamath, A., Walker, M., Beismann-Driemeyer, S., Vega-Fernandez, M.C., Sterner, R., Wilmanns, M., 2002. Structural evidence for ammonia tunneling across the (βα)8 barrel of the imidazole glycerol phosphate synthase bienzyme complex. Structure 10, 185-193.

Dutt, S.N., Kameswaran, M., 2005. The aetiology and management of atrophic rhinitis. Journal of Laryngology and Otology 119, 843-852.

Eggenhofer, F., Tafer, H., Stadler, P.F., Hofacker, I.L., 2011. RNApredator: fast accessibility-based prediction of sRNA targets. Nucleic Acids Research 39, W149-W154.

Faner, M.A., Feig, A.L., 2013. Identifying and characterizing Hfq–RNA interactions. Methods 63, 144-159.

Fantappiè, L., Metruccio, M.M.E., Seib, K.L., Oriente, F., Cartocci, E., Ferlicca, F., Giuliani, M.M., Scarlato, V., Delany, I., 2009. The RNA chaperone Hfq is involved in stress response and virulence in *Neisseria meningitidis* and is a pleiotropic regulator of protein expression. Infection and Immunity 77, 1842-1853.

Fekete, R.A., Miller, M.J., Chattoraj, D.K., 2003. Fluorescently labeled oligonucleotide extension: a rapid and quantitative protocol for primer extension. BioTechniques 35, 90-98.

Fender, A., Elf, J., Hampel, K., Zimmermann, B., Wagner, E.G.H., 2010. RNAs actively cycle on the Sm-like protein Hfq. Genes and Development 24, 2621-2626.

Figueroa-Bossi, N., Lemire, S., Maloriol, D., Balbontín, R., Casadesús, J., Bossi, L., 2006. Loss of Hfq activates the σE-dependent envelope stress response in *Salmonella enterica*. Molecular Microbiology 62, 838-852.

Folichon, M., Arluison, V., Pellegrini, O., Huntzinger, E., Régnier, P., Hajnsdorf, E., 2003. The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. Nucleic Acids Research 31, 7302-7310.

Franze De Fernandez, M.T., Eoyang, L., August, J.T., 1968. Factor fraction required for the synthesis of bacteriophage Qβ-RNA. Nature 219, 588-590.

Freese, N.H., Norris, D.C., Loraine, A.E., 2016. Integrated genome browser: visual analytics platform for genomics. Bioinformatics 32, 2089-2095.

Fuller, T.E., Kennedy, M.J., Lowery, D.E., 2000. Identification of *Pasteurella multocida* virulence genes in a septicemic mouse model using signature-tagged mutagenesis. Microbial Pathogenesis 29, 25-38.

Gallagher, S., Winston, S.E., Fuller, S.A., Hurrell, J.G.R., 2008. Immunoblotting and immunodetection. Current Protocols in Molecular Biology 83, 10.18.11-10.18.28.

Gallagher, S.R., 2006. One-dimensional SDS gel electrophoresis of proteins. Current Protocols in Molecular Biology 75, 10.12.11-10.12A.37.

Geinguenaud, F., Calandrini, V., Teixeira, J., Mayer, C., Liquier, J., Lavelle, C., Arluison, V., 2011. Conformational transition of DNA bound to Hfq probed by infrared spectroscopy. Physical Chemistry Chemical Physics 13, 1222-1229.

Geller, B.L., Marshall-Batty, K., Schnell, F.J., McKnight, M.M., Iversen, P.L., Greenberg, D.E., 2013. Gene-silencing antisense oligomers inhibit *Acinetobacter* growth *in vitro* and *in vivo*. The Journal of Infectious Diseases 208, 1553-1560.

Geng, J., Song, Y., Yang, L., Feng, Y., Qiu, Y., Li, G., Guo, J., Bi, Y., Qu, Y., Wang, W., Wang, X., Guo, Z., Yang, R., Han, Y., 2009. Involvement of the post-transcriptional regulator Hfq in *Yersinia pestis* virulence. PLoS ONE 4.

Glorioso, J.C., Jones, G.W., Rush, H.G., Pentler, L.J., Darif, C.A., Coward, J.E., 1982. Adhesion of type A *Pasteurella multocida* to rabbit pharyngeal cells and its possible role in rabbit respiratory tract infections. Infection and Immunity 35, 1103-1109.

Gómez-Lozano, M., Marvig, R., Molin, S., Long, K., 2012. Genome-wide identification of novel small RNAs in *Pseudomonas aeruginosa*. Environmental Microbiology 14, 2006-2016.

Gonzalez, G.M., Hardwick, S.W., Maslen, S.L., Skehel, J.M., Holmqvist, E., Vogel, J., Bateman, A., Luisi, B.F., Broadhurst, R.W., 2017. Structure of the *Escherichia coli* ProQ RNA-binding protein. RNA 23, 696-711.

Gottesman, S., Storz, G., 2011. Bacterial small RNA regulators: Versatile roles and rapidly evolving variations. Cold Spring Harbor Perspectives in Biology 3.

Graf, S., Binder, T., Heger, M., Apfalter, P., Simon, N., Winkler, S., 2007. Isolated endocarditis of the pulmonary valve caused by *Pasteurella multocida*. Infection 35, 43-45.

Guillier, M., Gottesman, S., Storz, G., 2006. Modulating the outer membrane with small RNAs. Genes and Development 20, 2338-2348.

Gulliver, E., Wright, A., Deveson Lucas, D., Mégroz, M., Kleifeld, O., B Schittenhelm, R., Powell, D., Seemann, T., B Bulitta, J., Harper, M., Boyce, J., 2018. Determination of the small RNA GcvB regulon in the Gram-negative bacterial pathogen *Pasteurella multocida* and identification of the GcvB seed binding region. RNA 24, 704-720.

Gumerov, Vadim M., Ortega, Davi R., Adebali, O., Ulrich, Luke E., Zhulin, Igor B., 2019. MiST 3.0: an updated microbial signal transduction database with an emphasis on chemosensory systems. Nucleic Acids Research 48, D459-D464.

Han, K., Tjaden, B., Lory, S., 2016. GRIL-seq provides a method for identifying direct targets of bacterial small regulatory RNA by *in vivo* proximity ligation. Nature Microbiology 2, 16239.

Hanahan, D., 1983. Studies on transformation of *Escherichia coli* with plasmids. Journal of Molecular Biology 166, 557-580.

Hansen, L.M., Hirsh, D.C., 1989. Serum resistance is correlated with encapsulation of avian strains of *Pasteurella multocida*. Veterinary Microbiology 21, 177-184.

Harper, M., Boyce, J.D., 2017. The myriad properties of *Pasteurella multocida* lipopolysaccharide. Toxins 9.

Harper, M., Boyce, J.D., Adler, B., 2006. *Pasteurella multocida* pathogenesis: 125 years after Pasteur. FEMS Microbiology Letters 265, 1-10.

Harper, M., Boyce, J.D., Adler, B., 2012. The key surface components of *Pasteurella multocida*: Capsule and lipopolysaccharide. Current Topics in Microbiology and Immunology 361, 39-51.

Harper, M., Boyce, J.D., Cox, A.D., St Michael, F., Wilkie, I.W., Blackall, P.J., Adler, B., 2007a. *Pasteurella multocida* expresses two lipopolysaccharide glycoforms simultaneously, but only a single form is required for virulence: identification of two acceptor-specific heptosyl I transferases. Infection and Immunity 75, 3885-3893.

Harper, M., Boyce, J.D., Wilkie, I.W., Adler, B., 2003. Signature-tagged mutagenesis of *Pasteurella multocida* identifies mutants displaying differential virulence characteristics in mice and chickens. Infection and Immunity 71, 5440-5446.

Harper, M., Cox, A., St Michael, F., Parnas, H., Wilkie, I., Blackall, P.J., Adler, B., Boyce, J.D., 2007b. Decoration of *Pasteurella multocida* lipopolysaccharide with phosphocholine is important for virulence. Journal of Bacteriology 189, 7384-7391.

Harper, M., Cox, A.D., Adler, B., Boyce, J.D., 2011. *Pasteurella multocida* lipopolysaccharide: the long and the short of it. Veterinary Microbiology 153, 109-115.

Harper, M., Cox, A.D., St Michael, F., Wilkie, I.W., Boyce, J.D., Adler, B., 2004. A heptosyltransferase mutant of *Pasteurella multocida* produces a truncated lipopolysaccharide structure and is attenuated in virulence. Infection and Immunity 72, 3436-3443.

Harper, M., John, M., Edmunds, M., Wright, A., Ford, M., Turni, C., Blackall, P.J., Cox, A., Adler, B., Boyce, J.D., 2016. Protective efficacy afforded by live *Pasteurella multocida* vaccines in chickens is independent of lipopolysaccharide outer core structure. Vaccine 34, 1696-1703.

Harper, M., John, M., Turni, C., Edmunds, M., St. Michael, F., Adler, B., Blackall, P.J., Cox, A.D., Boyce, J.D., 2015. Development of a rapid multiplex PCR assay to genotype *Pasteurella multocida* strains by use of the lipopolysaccharide outer core biosynthesis locus. Journal of Clinical Microbiology 53, 477-485.

Harper, M., St.Michael, F., John, M., Vinogradov, E., Steen, J.A., van Dorsten, L., Steen, J.A., Turni, C., Blackall, P.J., Adler, B., Cox, A.D., Boyce, J.D., 2013. *Pasteurella multocida* Heddleston serovar 3 and 4 strains share a common lipopolysaccharide biosynthesis locus but display both inter- and intrastrain lipopolysaccharide heterogeneity. Journal of Bacteriology 195, 4854-4864.

Harper, M., Wright, A., St. Michael, F., Li, J., Deveson Lucas, D., Ford, M., Adler, B., Cox, A.D., Boyce, J.D., 2017. Characterization of two novel lipopolysaccharide phosphoethanolamine transferases in *Pasteurella multocida* and their role in resistance to cathelicidin-2. Infection and Immunity 85, e00557-00517.

Hatfaludi, T., Al-Hasani, K., Gong, L., Boyce, J.D., Ford, M., Wilkie, I.W., Quinsey, N., Dunstone, M.A., Hoke, D.E., Adler, B., 2012. Screening of 71 *P. multocida* proteins for protective efficacy in a fowl cholera infection model and characterization of the protective antigen PlpE. PloS ONE 7, e39973.

Heap, J.T., Kuehne, S.A., Ehsaan, M., Cartman, S.T., Cooksley, C.M., Scott, J.C., Minton, N.P., 2010. The ClosTron: mutagenesis in Clostridium refined and streamlined. Journal of Microbiological Methods 80, 49-55.

Heddleston, K.L., Gallagher, J.E., Rebers, P.A., 1970. Fowl cholera: Immune response in turkeys. Avian Diseases 14, 626-635.

Heddleston, K.L., Gallagher, J.E., Rebers, P.A., 1972. Fowl cholera: gel diffusion precipitin test for serotyping *Pasteurella multocida* from avian species. Avian Diseases 16, 925-936.

Heddleston, K.L., Rebers, P.A., 1972. Fowl cholera: cross-immunity induced in turkeys with formalin-killed *in-vivo*-propagated *Pasteurella multocida*. Avian Diseases 16, 578-586.

Hempel, R.J., Morton, D.J., Seale, T.W., Whitby, P.W., Stull, T.L., 2013. The role of the RNA chaperone Hfq in *Haemophilus influenzae* pathogenesis. BMC Microbiology 13.

Henry, R., Crane, B., Powell, D., Deveson Lucas, D., Li, Z., Aranda, J., Harrison, P., Nation, R.L., Adler, B., Harper, M., Boyce, J.D., Li, J., 2015. The transcriptomic response of *Acinetobacter baumannii* to colistin and doripenem alone and in combination in an *in vitro* pharmacokinetics/pharmacodynamics model. The Journal of Antimicrobial Chemotherapy 70, 1303-1313.

Holmqvist, E., Li, L., Bischler, T., Barquist, L., Vogel, J., 2018. Global maps of ProQ binding *in vivo* reveal target recognition via RNA structure and stability control at mRNA 3′ ends. Molecular Cell 70, 971-982.

Holmqvist, E., Vogel, J., 2018. RNA-binding proteins in bacteria. Nature Reviews Microbiology 16, 601-615.

Holmqvist, E., Wright, P.R., Li, L., Bischler, T., Barquist, L., Reinhardt, R., Backofen, R., Vogel, J., 2016. Global RNA recognition patterns of post-transcriptional regulators Hfq and CsrA revealed by UV crosslinking *in vivo*. The EMBO Journal 35, 991-1011.

Hör, J., Gorski, S.A., Vogel, J., 2018. Bacterial RNA biology on a genome scale. Molecular Cell 70, 785-799.

Howden, B.P., Beaume, M., Harrison, P.F., Hernandez, D., Schrenzel, J., Seemann, T., Francois, P., Stinear, T.P., 2013. Analysis of the small RNA transcriptional response in multidrug-resistant *Staphylococcus aureus* after antimicrobial exposure. Antimicrobial Agents and Chemotherapy 57, 3864.

Iosub, I.A., Marchioretto, M., Sy, B., McKellar, S., Nieken, K.J., van Nues, R.W., Tree, J.J., Viero, G., Granneman, S., 2018. Hfq CLASH uncovers sRNA-target interaction networks enhancing adaptation to nutrient availability. bioRxiv, 481986.

Jackson, D.W., Suzuki, K., Oakford, L., Simecka, J.W., Hart, M.E., Romeo, T., 2002. Biofilm formation and dispersal under the influence of the global regulator CsrA of *Escherichia coli*. Journal of Bacteriology 184, 290.

Jarvinen, L.Z., Hogenesch, H., Suckow, M.A., Bowersock, T.L., 1998. Induction of protective immunity in rabbits by coadministration of inactivated *Pasteurella multocida* toxin and potassium thiocyanate extract. Infection and Immunity 66, 3788-3795.

Jatuponwiphat, T., Chumnanpuen, P., Othman, S., E-kobon, T., Vongsangnak, W., 2019. Iron-associated protein interaction networks reveal the key functional modules related to survival and virulence of *Pasteurella multocida*. Microbial Pathogenesis 127, 257-266.

Kajitani, M., Kato, A., Wada, A., Inokuchi, Y., Ishihama, A., 1994. Regulation of the *Escherichia coli hfq* gene encoding the host factor for phage Q beta. Journal of Bacteriology 176, 531-534.

Kalvari, I., Argasinska, J., Quinones-Olvera, N., Nawrocki, E.P., Rivas, E., Eddy, S.R., Bateman, A., Finn, R.D., Petrov, A.I., 2017. Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Research 46, D335-D342.

Katechakis, N., Maraki, S., Dramitinou, I., Marolachaki, E., Koutla, C., Ioannidou, E., 2019. An unusual case of *Pasteurella multocida* bacteremic meningitis. Journal of Infection and Public Health 12, 95-96.

Kavita, K., de Mets, F., Gottesman, S., 2018. New aspects of RNA-based regulation by Hfq and its partner sRNAs. Current Opinion in Microbiology 42, 53-61.

Kawano, M., Reynolds, A.A., Miranda-Rios, J., Storz, G., 2005. Detection of 5′- and 3′-UTR-derived small RNAs and *cis*-encoded antisense RNAs in *Escherichia coli*. Nucleic Acids Research 33, 1040-1050.

Kazantsev, A.V., Pace, N.R., 2006. Bacterial RNase P: A new view of an ancient enzyme. Nature Reviews Microbiology 4, 729-740.

Keiler, K.C., 2008. Biology of trans-translation. Annual Review of Microbiology 62, 133-151.

Kery, M.B., Feldman, M., Livny, J., Tjaden, B., 2014. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. Nucleic Acids Research 42, W124-W129.

Kimura, A., Mountzouros, K.T., Relman, D.A., Falkow, S., Cowell, J.L., 1990. *Bordetella pertussis* filamentous hemagglutinin: Evaluation as a protective antigen and colonization factor in a mouse respiratory infection model. Infection and Immunity 58, 7-16.

King, A.M., Vanderpool, C.K., Degnan, P.H., 2019. sRNA target prediction organizing tool (SPOT) integrates computational and experimental data to facilitate functional characterization of bacterial small RNAs. mSphere 4, e00561-00518.

Koo, J.T., Alleyne, T.M., Schiano, C.A., Jafari, N., Lathem, W.W., 2011. Global discovery of small RNAs in *Yersinia pseudotuberculosis* identifies *Yersinia*-specific small, noncoding RNAs required for virulence. Proceedings of the National Academy of Sciences of the United States of America 108, E709-E717.

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S., Hammarlöf, D., Canals, R., Grissom, J., Conway, T., Hokamp, K., Hinton, J., 2013. An infection-relevant transcriptomic compendium for *Salmonella enterica* serovar Typhimurium. Cell Host & Microbe 14, 683-695.

Kudla, G., Granneman, S., Hahn, D., Beggs, J.D., Tollervey, D., 2011. Cross-linking, ligation, and sequencing of hybrids reveals RNA–RNA interactions in yeast. Proceedings of the National Academy of Sciences 108, 10010-10015.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lalaouna, D., Carrier, M.-C., Semsey, S., Brouard, J.-S., Wang, J., Wade, Joseph T., Massé, E., 2015. A 3′ external transcribed spacer in a tRNA transcript acts as a sponge for small RNAs to prevent transcriptional noise. Molecular Cell 58, 393-405.

Law, C.W., Chen, Y., Shi, W., Smyth, G.K., 2014. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biology 15, R29.

Lease, R.A., Cusick, M.E., Belfort, M., 1998. Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. Proceedings of the National Academy of Sciences 95, 12456.

Lee, T., Feig, A.L., 2008. The RNA binding protein Hfq interacts specifically with tRNAs. RNA 14, 514-523.

Lenz, D.H., Mok, K.C., Lilley, B.N., Kulkarni, R.V., Wingreen, N.S., Bassler, B.L., 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell 118, 69-82.

Leonard, S., Meyer, S., Lacour, S., Nasser, W., Hommais, F., Reverchon, S., 2019. APERO: a genome-wide approach for identifying bacterial small RNAs from RNA-Seq data. Nucleic acids research 47.

Lesnik, E.A., Sampath, R., Levene, H.B., Henderson, T.J., McNeil, J.A., Ecker, D.J., 2001. Prediction of rho-independent transcriptional terminators in *Escherichia coli*. Nucleic Acids Research 29, 3583-3594.

Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM, arXiv.

Li, L., Huang, D., Cheung, M.K., Nong, W., Huang, Q., Kwan, H.S., 2012a. BSRD: a repository for bacterial small regulatory RNA. Nucleic Acids Research 41, D233-D238.

Li, W., Ying, X., Lu, Q., Chen, L., 2012b. Predicting sRNAs and their targets in bacteria. Genomics, Proteomics and Bioinformatics 10, 276-284.

Lindgreen, S., Umu, S.U., Lai, A.S.-W., Eldai, H., Liu, W., McGimpsey, S., Wheeler, N.E., Biggs, P.J., Thomson, N.R., Barquist, L., Poole, A.M., Gardner, P.P., 2014. Robust identification of noncoding RNA from transcriptomes requires phylogenetically-informed sampling. PLoS Computational Biology 10, e1003907-e1003907.

Link, T.M., Valentin-Hansen, P., Brennan, R.G., 2009. Structure of *Escherichia coli* Hfq bound to polyriboadenylate RNA. Proceedings of the National Academy of Sciences 106, 19292-19297.

Liu, F., Zheng, K., Chen, H.-C., Liu, Z.-F., 2018. Capping-RACE: a simple, accurate, and sensitive 5′ RACE method for use in prokaryotes. Nucleic Acids Research 46, e129-e129.

Liu, H., Wang, Q., Liu, Q., Cao, X., Shi, C., Zhang, Y., 2011. Roles of Hfq in the stress adaptation and virulence in fish pathogen *Vibrio alginolyticus* and its potential application as a target for live attenuated vaccine. Applied Microbiology and Biotechnology 91, 353-364.

Livny, J., Fogel, M.A., Davis, B.M., Waldor, M.K., 2005. sRNAPredict: an integrative computational approach to identify sRNAs in bacterial genomes. Nucleic Acids Research 33, 4096-4105.

Livny, J., Teonadi, H., Livny, M., Waldor, M.K., 2008. High-throughput, kingdom-wide prediction and annotation of bacterial non-coding RNAs. PloS ONE 3, e3197-e3197.

Lloyd, A.L., Marshall, B.J., Mee, B.J., 2005. Identifying cloned *Helicobacter pylori* promoters by primer extension using a FAM-labelled primer and GeneScan® analysis. Journal of Microbiological Methods 60, 291-298.

Loh, E., Dussurget, O., Gripenland, J., Vaitkevicius, K., Tiensuu, T., Mandin, P., Repoila, F., Buchrieser, C., Cossart, P., Johansson, J., 2009. A *trans*-Acting Riboswitch Controls Expression of the Virulence Regulator PrfA in *Listeria monocytogenes*. Cell 139, 770-779.

López-Gomollón, S., 2011. Detecting sRNAs by northern blotting, in: Dalmay, T. (Ed.), MicroRNAs in Development: Methods and Protocols. Humana Press, Totowa, NJ, pp. 25-38.

Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T., Gottesman, S., 1998. DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. Proceedings of the National Academy of Sciences of the United States of America 95, 12462–12467.

Mandal, M., Lee, M., Barrick, J.E., Weinberg, Z., Emilsson, G.M., Ruzzo, W.L., Breaker, R.R., 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. Science 306, 275-279.

Mandin, P., Guillier, M., 2013. Expanding control in bacteria: Interplay between small RNAs and transcriptional regulators to control gene expression. Current Opinion in Microbiology 16, 125-132.

Massé, E., Escorcia, F.E., Gottesman, S., 2003. Coupled degradation of a small regulatory RNA and its mRNA targets in *Escherichia coli*. Genes and Development 17, 2374-2383.

Massé, E., Gottesman, S., 2002. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America 99, 4620-4625.

Massé, E., Vanderpool, C.K., Gottesman, S., 2005. Effect of RyhB small RNA on global iron use in *Escherichia coli*. Journal of Bacteriology 187, 6962-6971.

Matsumoto, M., Strain, J.G., Engel, H.N., 1991. The fate of *Pasteurella multocida* after intratracheal inoculation into turkeys. Poultry Science 70, 2259-2266.

May, B.J., Zhang, Q., Li, L.L., Paustian, M.L., Whittam, T.S., Kapur, V., 2001. Complete genomic sequence of *Pasteurella multocida*, Pm70. Proceedings of the National Academy of Sciences of the United States of America 98, 3460-3465.

McNealy, T.L., Forsbach-Birk, V., Shi, C., Marre, R., 2005. The Hfq homolog in *Legionella pneumophila* demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. Journal of Bacteriology 187, 1527-1532.

Mégroz, M., Kleifeld, O., Wright, A., Powell, D., Harrison, P., Adler, B., Harper, M., Boyce, J.D., 2016. The RNA-binding chaperone Hfq is an important global regulator of gene expression in *Pasteurella multocida* and plays a crucial role in production of a number of virulence factors, including hyaluronic acid capsule. Infection and Immunity 84, 1361-1370.

Meibom, K.L., Forslund, A.L., Kuoppa, K., Alkhuder, K., Dubail, I., Dupuis, M., Forsberg, Å., Charbit, A., 2009. Hfq, a novel pleiotropic regulator of virulence-associated genes in *Francisella tularensis*. Infection and Immunity 77, 1866-1880.

Mellin, J.R., Cossart, P., 2012. The non-coding RNA world of the bacterial pathogen *Listeria monocytogenes*. RNA Biology 9, 372-378.

Mikulecky, P.J., Kaw, M.K., Brescia, C.C., Takach, J.C., Sledjeski, D.D., Feig, A.L., 2004. *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, *rpoS* and poly(A) RNAs. Nature Structural and Molecular Biology 11, 1206-1214.

Mizan, S., Henk, A., Stallings, A., Maier, M., Lee, M.D., 2000. Cloning and characterization of sialidases with 2-6' and 2-3' sialyl lactose specificity from *Pasteurella multocida*. Journal of Bacteriology 182, 6874-6883.

Mohanty, B.K., Maples, V.F., Kushner, S.R., 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. Molecular Microbiology 54, 905-920.

Moller, T., Franch, T., Hojrup, P., Keene, D.R., Bächinger, H.P., Brennan, R.G., Valentin-Hansen, P., 2002. Hfq: A Bacterial Sm-like protein that mediates RNA-RNA interaction. Molecular Cell 9, 23-30.

Morita, T., Aiba, H., 2019. Mechanism and physiological significance of autoregulation of the *Escherichia coli hfq* gene. RNA 25, 264-276.

Morita, T., Maki, K., Aiba, H., 2012. Detection of sRNA–mRNA interactions by electrophoretic mobility shift assay, in: Keiler, K.C. (Ed.), Bacterial Regulatory RNA: Methods and Protocols. Humana Press, Totowa, NJ, pp. 235-244.

Mullan, P.B., Lax, A.J., 1998. *Pasteurella multocida* toxin stimulates bone resorption by osteoclasts via interaction with osteoblasts. Calcified Tissue International 63, 340-345.

Münch, R., Hiller, K., Grote, A., Scheer, M., Klein, J., Schobert, M., Jahn, D., 2005. Virtual Footprint and PRODORIC: an integrative framework for regulon prediction in prokaryotes. Bioinformatics 21, 4187-4189.

NHMRC, 2004. Australian code of practise for the care and use of animals for scientific purposes. National Health and Medical Research Council, Canberra, Australia. 7th ed.

Oglesby-Sherrouse, A.G., Murphy, E.R., 2013. Iron-responsive bacterial small RNAs: variations on a theme. Metallomics: Integrated Biometal Science 5, 276-286.

Olsen, A.S., Møller-Jensen, J., Brennan, R.G., Valentin-Hansen, P., 2010. C-terminally truncated derivatives of *Escherichia coli* Hfq are proficient in riboregulation. Journal of Molecular Biology 404, 173-182.

Oogai, Y., Gotoh, Y., Ogura, Y., Kawada-Matsuo, M., Hayashi, T., Komatsuzawa, H., 2017. Small RNA repertoires and their intraspecies variation in *Aggregatibacter actinomycetemcomitans*. DNA Research 25, 207-215.

Opdyke, J.A., Kang, J.-G., Storz, G., 2004. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. Journal of Bacteriology 186, 6698-6705.

Orth, J.H.C., 2009. *Pasteurella multocida* toxin. Microbial Toxins: Current Research and Future Trends.

Osborne, J., Djapgne, L., Tran, B.Q., Goo, Y.A., Oglesby-Sherrouse, A.G., 2014. A method for *in vivo* identification of bacterial small RNA-binding proteins. MicrobiologyOpen 3, 950-960.

Pain, A., Ott, A., Amine, H., Rochat, T., Bouloc, P., Gautheret, D., 2015. An assessment of bacterial small RNA target prediction programs. RNA Biology 12, 509-513.

Panja, S., Schu, D.J., Woodson, S.A., 2013. Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. Nucleic Acids Research 41, 7536-7546.

Papenfort, K., Pfeiffer, V., Mika, F., Lucchini, S., Hinton, J.C.D., Vogel, J., 2006. σE-dependent small RNAs of *Salmonella* respond to membrane stress by accelerating global *omp* mRNA decay. Molecular Microbiology 62, 1674-1688.

Papenfort, K., Vogel, J., 2009. Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. Research in Microbiology 160, 278-287.

Papenfort, K., Vogel, J., 2010. Regulatory RNA in bacterial pathogens. Cell Host and Microbe 8, 116-127.

Paustian, M.L., May, B.J., Kapur, V., 2001. *Pasteurella multocida* gene expression in response to iron limitation. Infection and Immunity 69, 4109-4115.

Paustian, M.L., May, B.J., Kapur, V., 2002. Transcriptional response of *Pasteurella multocida* to nutrient limitation. Journal of Bacteriology 184, 3734-3739.

Peer, A., Margalit, H., 2011. Accessibility and evolutionary conservation mark bacterial small-RNA target-binding regions. Journal of Bacteriology 193, 1690.

Peña-Castillo, L., Grüell, M., Mulligan, M.E., Lang, A.S., 2016. Detection of bacterial small transcripts from RNA-Seq data: a comparative assessment, Biocomputing 2016, pp. 456-467.

Peng, Y., Curtis, J.E., Fang, X., Woodson, S.A., 2014. Structural model of an mRNA in complex with the bacterial chaperone Hfq. Proceedings of the National Academy of Sciences 111, 17134-17139.

Peng, Z., Wang, X., Zhou, R., Chen, H., Wilson, B.A., Wu, B., 2019. *Pasteurella multocida*: genotypes and genomics. Microbiology and Molecular Biology Reviews 83, e00014-00019.

Pfeiffer, V., Sittka, A., Tomer, R., Tedin, K., Brinkmann, V., Vogel, J., 2007. A small non-coding RNA of the invasion gene island (SPI-1) represses outer membrane protein synthesis from the *Salmonella* core genome. Molecular Microbiology 66, 1174-1191.

Pichon, C., Felden, B., 2008. Small RNA gene identification and mRNA target predictions in bacteria. Bioinformatics 24, 2807-2813.

Pischimarov, J., Kuenne, C., Billion, A., Hemberger, J., Cemič, F., Chakraborty, T., Hain, T., 2012. sRNAdb: a small non-coding RNA database for Gram-positive bacteria. BMC Genomics 13, 384.

Potts, A.H., Vakulskas, C.A., Pannuri, A., Yakhnin, H., Babitzke, P., Romeo, T., 2017. Global role of the bacterial post-transcriptional regulator CsrA revealed by integrated transcriptomics. Nature Communications 8, 1596.

Pruimboom, I.M., Rimler, R.B., Ackermann, M.R., 1999. Enhanced adhesion of *Pasteurella multocida* to cultured turkey peripheral blood monocytes. Infection and Immunity 67, 1292-1296.

Pulvermacher, S.C., Stauffer, L.T., Stauffer, G.V., 2009. Role of the sRNA GcvB in regulation of *cycA* in *Escherichia coli*. Microbiology 155, 106-114.

Ramseier, T.M., Bledig, S., Michotey, V., Feghali, R., Saier Jr, M.H., 1995. The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. Molecular Microbiology 16, 1157-1169.

Rath, E.C., Pitman, S., Cho, K.H., Bai, Y., 2017. Identification of streptococcal small RNAs that are putative targets of RNase III through bioinformatics analysis of RNA sequencing data. BMC Bioinformatics 18, 540.

Regulski, E.E., Moy, R.H., Weinberg, Z., Barrick, J.E., Yao, Z., Ruzzo, W.L., Breaker, R.R., 2008. A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. Molecular Microbiology 68, 918-932.

Rhoades, K.R., Rimler, R.B., 1990. Virulence and toxigenicity of capsular serogroup D *Pasteurella multocida* strains isolated from avian hosts. Avian Diseases 34, 384-388.

Rio, D.C., 2014. Northern blots for small RNAs and microRNAs. Cold Spring Harbor Protocols 2014, 793-797.

Rio, D.C., Ares, M., Hannon, G.J., Nilsen, T.W., 2010. Polyacrylamide gel electrophoresis of RNA. Cold Spring Harbor Protocols 2010.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research 43, e47.

Rivas, E., Eddy, S.R., 2001. Noncoding RNA gene detection using comparative sequence analysis. BMC Bioinformatics 2, 8.

Robertson, G.T., Roop, R.M.I., 1999. The *Brucella abortus* host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. Molecular Microbiology 34, 690-700.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., Mesirov, J.P., 2011. Integrative genomics viewer. Nature Biotechnology 29, 24-26.

Romby, P., Vandenesch, F., Wagner, E.G.H., 2006. The role of RNAs in the regulation of virulence-gene expression. Current Opinion in Microbiology 9, 229-236.

Romeo, T., Babitzke, P., 2018. Global regulation by CsrA and its RNA antagonists. Microbiology Spectrum 6.

Romeo, T., Gong, M., Liu, M.Y., Brun-Zinkernagel, A.M., 1993. Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. Journal of Bacteriology 175, 4744-4755.

Rossi, C.C., Bossé, J.T., Li, Y., Witney, A.A., Gould, K.A., Langford, P.R., Bazzolli, D.M.S., 2016. A computational strategy for the search of regulatory small RNAs in *Actinobacillus pleuropneumoniae*. RNA 22, 1373-1385.

Roumen, M.P., van de Goor, P.T., 1988. Atrophic rhinitis in cattle. Tijdschrift voor diergeneeskunde 113, 28.

Ruffolo, C.G., Tennent, J.M., Michalski, W.P., Adler, B., 1997. Identification, purification, and characterization of the type 4 fimbriae of *Pasteurella multocida*. Infection and Immunity 65, 339-343.

Saadeh, B., Caswell, C.C., Chao, Y., Berta, P., Wattam, A.R., Roop, R.M., Callaghan, D., 2016. Transcriptome-wide identification of Hfq-associated RNAs in *Brucella suis* by deep sequencing. Journal of Bacteriology 198, 427-435.

Sabnis, N.A., Yang, H., Romeo, T., 1995. Pleiotropic regulation of central carbohydrate metabolism in *Escherichia coli* via the gene *csrA*. Journal of Biological Chemistry 270, 29096-29104.

Salim, N.N., Faner, M.A., Philip, J.A., Feig, A.L., 2012. Requirement of upstream Hfq-binding (ARN)x elements in *glmS* and the Hfq C-terminal region for GlmS upregulation by sRNAs GlmZ and GlmY. Nucleic Acids Research 40, 8021-8032.

Salim, N.N., Feig, A.L., 2010. An upstream Hfq binding site in the *fhlA* mRNA leader region facilitates the OxyS-*fhlA* interaction. PLoS ONE 5, e13028.

Santana, E.A., Harrison, A., Zhang, X., Baker, B.D., Kelly, B.J., White, P., Liu, Y., Munson, R.S., Jr., 2014. HrrF is the Fur-regulated small RNA in nontypeable *Haemophilus influenzae*. PloS ONE 9, e105644-e105644.

Santiago-Frangos, A., Kavita, K., Schu, D.J., Gottesman, S., Woodson, S.A., 2016. C-terminal domain of the RNA chaperone Hfq drives sRNA competition and release of target RNA. Proceedings of the National Academy of Sciences 113, E6089.

Santiago-Frangos, A., Woodson, S.A., 2018. Hfq chaperone brings speed dating to bacterial sRNA. WIREs RNA 9, e1475.

Sauer, E., 2013. Structure and RNA-binding properties of the bacterial LSm protein Hfq. RNA Biology 10, 610-618.

Sauer, E., Weichenrieder, O., 2011. Structural basis for RNA 3'-end recognition by Hfq. Proceedings of the National Academy of Sciences 108, 13065-13070.

Schumacher, M.A., Pearson, R.F., Møller, T., Valentin-Hansen, P., Brennan, R.G., 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: A bacterial Sm-like protein. The EMBO Journal 21, 3546-3556.

Serganov, A., Huang, L., Patel, D.J., 2008. Structural insights into amino acid binding and gene control by a lysine riboswitch. Nature 455, 1263-1267.

Serganov, A., Huang, L., Patel, D.J., 2009. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. Nature 458, 233-237.

Sharma, C.M., Papenfort, K., Pernitzsch, S.R., Mollenkopf, H.J., Hinton, J.C.D., Vogel, J., 2011. Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. Molecular Microbiology 81, 1144-1165.

Sharma, C.M., Vogel, J., 2009. Experimental approaches for the discovery and characterization of regulatory small RNA. Current Opinion in Microbiology 12, 536-546.

Shivachandra, S.B., Viswas, K.N., Kumar, A.A., 2011. A review of hemorrhagic septicemia in cattle and buffalo. Animal health research reviews / Conference of Research Workers in Animal Diseases 12, 67-82.

Singh, M.B., Harrington, A.T., 2017. *Pasteurella multocida* urinary tract infection in a patient with cervical cancer. JMM Case Rep 4, e005082-e005082.

Sittka, A., Lucchini, S., Papenfort, K., Sharma, C.M., Rolle, K., Binnewies, T.T., Hinton, J.C.D., Vogel, J., 2008. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. PLoS Genetics 4.

Sittka, A., Pfeiffer, V., Tedin, K., Vogel, J., 2007. The RNA chaperone Hfq is essential for the virulence of *Salmonella typhimurium*. Molecular Microbiology 63, 193-217.

Sittka, A., Sharma, C.M., Rolle, K., Vogel, J., 2009. Deep sequencing of *Salmonella* RNA associated with heterologous Hfq proteins *in vivo* reveals small RNAs as a major target class and identifies RNA processing phenotypes. RNA Biology 6, 266-275.

Smirnov, A., Förstner, K.U., Holmqvist, E., Otto, A., Günster, R., Becher, D., Reinhardt, R., Vogel, J., 2016. Grad-seq guides the discovery of ProQ as a major small RNA-binding protein. Proceedings of the National Academy of Sciences of the United States of America 113, 11591-11596.

Smirnov, A., Wang, C., Drewry, L.L., Vogel, J., 2017. Molecular mechanism of mRNA repression in *trans* by a ProQ-dependent small RNA. The EMBO Journal 36, 1029-1045.

Smith, M.N., Kwok, S.C., Hodges, R.S., Wood, J.M., 2007. Structural and functional analysis of ProQ: an osmoregulatory protein of *Escherichia coli*. Biochemistry 46, 3084-3095.

Snipes, K.P., Hirsh, D.C., 1986. Association of complement sensitivity with virulence of *Pasteurella multocida* isolated from turkeys. Avian Diseases 30, 500-504.

Sobrero, P., Valverde, C., 2012. The bacterial protein Hfq: Much more than a mere RNA-binding factor. Critical Reviews in Microbiology 38, 276-299.

Sonnleitner, E., Abdou, L., Haas, D., 2009. Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. Proceedings of the National Academy of Sciences of the United States of America 106, 21866-21871.

Sonnleitner, E., Bläsi, U., 2014. Regulation of Hfq by the RNA CrcZ in *Pseudomonas aeruginosa* Carbon Catabolite Repression. PLoS Genetics 10, e1004440.

Sonnleitner, E., Hagens, S., Rosenau, F., Wilhelm, S., Habel, A., Jäger, K.E., Bläsi, U., 2003. Reduced virulence of a *hfq* mutant of *Pseudomonas aeruginosa* O1. Microbial Pathogenesis 35, 217-228.

Sonnleitner, E., Schuster, M., Sorger-Domenigg, T., Greenberg, E.P., Bläsi, U., 2006. Hfq-dependent alterations of the transcriptome profile and effects on quorum sensing in *Pseudomonas aeruginosa*. Molecular Microbiology 59, 1542-1558.

Sonnleitner, E., Wulf, A., Campagne, S., Pei, X.-Y., Wolfinger, M.T., Forlani, G., Prindl, K., Abdou, L., Resch, A., Allain, F.H.-T., Luisi, B.F., Urlaub, H., Bläsi, U., 2017. Interplay between the catabolite repression control protein Crc, Hfq and RNA in Hfq-dependent translational regulation in *Pseudomonas aeruginosa*. Nucleic Acids Research 46, 1470-1485.

Soper, T., Mandin, P., Majdalani, N., Gottesman, S., Woodson, S.A., 2010. Positive regulation by small RNAs and the role of Hfq. Proceedings of the National Academy of Sciences of the United States of America 107, 9602-9607.

Spiro, S., Guest, J.R., 1987. Regulation and over-expression of the *fnr* gene of *Escherichia coli*. Microbiology 133, 3279-3288.

Sridhar, J., Gunasekaran, P., 2013. Computational small RNA prediction in bacteria. Bioinformatics and Biology Insights 7, 83-95.

Steen, J.A., Steen, J.A., Harrison, P., Seemann, T., Wilkie, I., Harper, M., Adler, B., Boyce, J.D., 2010. Fis is essential for capsule production in *Pasteurella multocida* and regulates expression of other important virulence factors. PLoS Pathogens 6, e1000750.

Steenbergen, S.M., Lichtensteiger, C.A., Caughlan, R., Garfinkle, J., Fuller, T.E., Vimr, E.R., 2005. Sialic acid metabolism and systemic pasteurellosis. Infection and Immunity 73, 1284-1294.

Storz, G., Vogel, J., Wassarman, K., 2011. Regulation by small RNAs in bacteria: expanding frontiers. Molecular Cell 43, 880-891.

Stubben, C.J., Micheva-Viteva, S.N., Shou, Y., Buddenborg, S.K., Dunbar, J.M., Hong-Geller, E., 2014. Differential expression of small RNAs from *Burkholderia thailandensis* in response to varying environmental and stress conditions. BMC genomics 15, 385-385.

Subashchandrabose, S., Leveque, R.M., Kirkwood, R.N., Kiupel, M., Mulks, M.H., 2013. The RNA chaperone *hfq* promotes fitness of *Actinobacillus pleuropneumoniae* during porcine pleuropneumonia. Infection and Immunity 81, 2952-2961.

Takada, A., Wachi, M., Kaidow, A., Takamura, M., Nagai, K., 1997. DNA binding properties of the *hfq* gene product of *Escherichia coli*. Biochemical and Biophysical Research Communications 236, 576-579.

Talan, D.A., Citron, D.M., Abrahamian, F.M., Moran, G.J., Goldstein, E.J.C., 1999. Bacteriologic analysis of infected dog and cat bites. New England Journal of Medicine 340, 85-92.

Tatum, F.M., Tabatabai, L.B., Briggs, R.E., 2009. Sialic acid uptake is necessary for virulence of *Pasteurella multocida* in turkeys. Microbial Pathogenesis 46, 337-344.

Tatum, F.M., Yersin, A.G., Briggs, R.E., 2005. Construction and virulence of a *Pasteurella multocida fha*B2 mutant in turkeys. Microbial Pathogenesis 39, 9-17.

Thomason, M.K., Storz, G., 2010. Bacterial antisense RNAs: how many are there, and what are they doing? Annual Review of Genetics 44, 167-188.

Torres-Quesada, O., Reinkensmeier, J., Schlüter, J.-P., Robledo, M., Peregrina, A., Giegerich, R., Toro, N., Becker, A., Jiménez-Zurdo, J.I., 2014. Genome-wide profiling of Hfq-binding RNAs uncovers extensive post-transcriptional rewiring of major stress response and symbiotic regulons in *Sinorhizobium meliloti*. RNA biology 11, 563-579.

Townsend, K.M., Boyce, J.D., Chung, J.Y., Frost, A.J., Adler, B., 2001. Genetic organization of *Pasteurella multocida cap* loci and development of a multiplex capsular PCR typing system. Journal of Clinical Microbiology 39, 924.

Tsui, H.C.T., Leung, H.C.E., Winkler, M.E., 1994. Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12. Molecular Microbiology 13, 35-49.

Updegrove, T.B., Zhang, A., Storz, G., 2016. Hfq: the flexible RNA matchmaker. Current Opinion in Microbiology 30, 133-138.

Uzzau, S., Figueroa-Bossi, N., Rubino, S., Bossi, L., 2001. Epitope tagging of chromosomal genes in *Salmonella*. Proceedings of the National Academy of Sciences 98, 15264-15269.

Vakulskas, C.A., Potts, A.H., Babitzke, P., Ahmer, B.M.M., Romeo, T., 2015. Regulation of bacterial virulence by Csr (Rsm) systems. Microbiology and Molecular Biology Reviews 79, 193-224.

Valentin-Hansen, P., Eriksen, M., Udesen, C., 2004. MicroReview: the bacterial Sm-like protein Hfq: a key player in RNA transactions. Molecular Microbiology 51, 1525-1533.

Večerek, B., Moll, I., Bläsi, U., 2005. Translational autocontrol of the *Escherichia coli hfq* RNA chaperone gene. RNA 11, 976-984.

Večerek, B., Rajkowitsch, L., Sonnleitner, E., Schroeder, R., Bläsi, U., 2008. The C-terminal domain of *Escherichia coli* Hfq is required for regulation. Nucleic Acids Research 36, 133-143.

Vimr, E.R., 2013. Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. ISRN Microbiology 2013, 816713-816739.

Vimr, E.R., Kalivoda, K.A., Deszo, E.L., Steenbergen, S.M., 2004. Diversity of microbial sialic acid metabolism. Microbiology and Molecular Biology Reviews 68, 132-153.

Vogel, J., Luisi, B.F., 2011. Hfq and its constellation of RNA. Nature Reviews Microbiology 9, 578-589.

Vogel, J., Sharma, C.M., 2005. How to find small non-coding RNAs in bacteria. Biological Chemistry 386, 1219-1238.

Wagner, E.G.H., 2013. Cycling of RNAs on Hfq. RNA Biology 10, 619-626.

Wagner, E.G.H., Altuvia, S., Romby, P., 2002. Antisense RNAs in bacteria and their genetic elements, in: Dunlap, J.C., Wu, C.t. (Eds.), Advances in Genetics. Academic Press, pp. 361-398.

Wang, C., Ding, C., Meraz, R.F., Holbrook, S.R., 2006. PSoL: a positive sample only learning algorithm for finding non-coding RNA genes. Bioinformatics 22, 2590-2596.

Washietl, S., Hofacker, I.L., Stadler, P.F., 2005. Fast and reliable prediction of noncoding RNAs. Proceedings of the National Academy of Sciences of the United States of America 102, 2454-2459.

Wassarman, K.M., Storz, G., 2000. 6S RNA regulates *E. coli* RNA polymerase activity. Cell 101, 613-623.

Waters, L.S., Storz, G., 2009. Regulatory RNAs in bacteria. Cell 136, 615-628.

Waters, S.A., McAteer, S.P., Kudla, G., Pang, I., Deshpande, N.P., Amos, T.G., Leong, K.W., Wilkins, M.R., Strugnell, R., Gally, D.L., Tollervey, D., Tree, J.J., 2017. Small RNA interactome of pathogenic *E. coli* revealed through crosslinking of RNase E. The EMBO Journal 36, 374-387.

Wei, B.L., Brun-Zinkernagel, A.-M., Simecka, J.W., Prüß, B.M., Babitzke, P., Romeo, T., 2001. Positive regulation of motility and *flhDC* expression by the RNA-binding protein CsrA of *Escherichia coli*. Molecular Microbiology 40, 245-256.

Westermann, A.J., Venturini, E., Sellin, M.E., Förstner, K.U., Hardt, W.-D., Vogel, J., 2019. The major RNA-binding protein ProQ impacts virulence gene expression in *Salmonella enterica* Serovar Typhimurium. mBio 10, e02504-02518.

Wilkie, I.W., Grimes, S.E., O'Boyle, D., Frost, A.J., 2000. The virulence and protective efficacy for chickens of *Pasteurella multocida* administered by different routes. Veterinary Microbiology 72, 57-68.

Wilkie, I.W., Harper, M., Boyce, J.D., Adler, B., 2012. *Pasteurella multocida*: diseases and pathogenesis. Current Topics in Microbiology and Immunology 361, 1-22.

Williams, E.S., Runde, D.E., Mills, K., Holler, L.D., 1987. Avian cholera in a gyrfalcon (Falco rusticolus). Avian Diseases 31, 380-382.

Wilson, B.A., Ho, M., 2013. *Pasteurella multocida*: from zoonosis to cellular microbiology. Clinical Microbiology Reviews 26, 631-655.

Wiśniewski, J.R., Zougman, A., Nagaraj, N., Mann, M., 2009. Universal sample preparation method for proteome analysis. Nature Methods 6, 359-362.

Woo, Y.K., Kim, J.H., 2006. Fowl cholera outbreak in domestic poultry and epidemiological properties of *Pasteurella multocida* isolate. Journal of Microbiology 44, 344-353.

Wright, P.R., Georg, J., Mann, M., Sorescu, D.A., Richter, A.S., Lott, S., Kleinkauf, R., Hess, W.R., Backofen, R., 2014. CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. Nucleic Acids Research 42, W119-W123.

Wu, H., Sun, L., Blombach, F., Brouns, S.J.J., Snijders, A.P.L., Lorenzen, K., Van Den Heuvel, R.H.H., Heck, A.J.R., Fu, S., Li, X., Zhang, X.C., Rao, Z., Van Der Oost, J., 2010. Structure of the ribosome asociating GTPase Hflx. Proteins: Structure, Function and Bioinformatics 78, 705-713.

Yakhnin, A.V., Baker, C.S., Vakulskas, C.A., Yakhnin, H., Berezin, I., Romeo, T., Babitzke, P., 2013. CsrA activates *flhDC* expression by protecting *flhDC* mRNA from RNase E-mediated cleavage. Molecular Microbiology 87, 851-866.

Yakhnin, H., Baker, C.S., Berezin, I., Evangelista, M.A., Rassin, A., Romeo, T., Babitzke, P., 2011. CsrA represses translation of *sdiA*, which encodes the *N*-Acylhomoserine-l-Lactone receptor of *Escherichia coli*, by binding exclusively within the coding region of *sdiA* mRNA. Journal of Bacteriology 193, 6162-6170.

Zhang, A., Wassarman, K.M., Rosenow, C., Tjaden, B.C., Storz, G., Gottesman, S., 2003. Global analysis of small RNA and mRNA targets of Hfq. Molecular Microbiology 50, 1111-1124.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., Liu, X.S., 2008. Model-based analysis of ChIP-Seq (MACS). Genome Biology 9, R137.

Zuker, M., 2003. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research 31, 3406-3415.

Appendix A: Hfq publication (Chapter 2)

AMERICAN SOCIETY FOR MICROBIOLOGY | Infection and Immunity

The RNA-Binding Chaperone Hfq Is an Important Global Regulator of Gene Expression in *Pasteurella multocida* and Plays a Crucial Role in Production of a Number of Virulence Factors, Including Hyaluronic Acid Capsule

Marianne Mégroz,[a] Oded Kleifeld,[b] Amy Wright,[a] David Powell,[c] Paul Harrison,[c] Ben Adler,[a] Marina Harper,[a] John D. Boyce[a]

Department of Microbiology[a] and Department of Biochemistry and Molecular Biology,[b] Infection and Immunity Program, Monash Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia; Monash Bioinformatics Platform, Monash University, Clayton, Victoria, Australia[c]

The Gram-negative bacterium *Pasteurella multocida* is the causative agent of a number of economically important animal diseases, including avian fowl cholera. Numerous *P. multocida* virulence factors have been identified, including capsule, lipopolysaccharide (LPS), and filamentous hemagglutinin, but little is known about how the expression of these virulence factors is regulated. Hfq is an RNA-binding protein that facilitates riboregulation via interaction with small noncoding RNA (sRNA) molecules and their mRNA targets. Here, we show that a *P. multocida hfq* mutant produces significantly less hyaluronic acid capsule during all growth phases and displays reduced *in vivo* fitness. Transcriptional and proteomic analyses of the *hfq* mutant during mid-exponential-phase growth revealed altered transcript levels for 128 genes and altered protein levels for 78 proteins. Further proteomic analyses of the *hfq* mutant during the early exponential growth phase identified 106 proteins that were produced at altered levels. Both the transcript and protein levels for genes/proteins involved in capsule biosynthesis were reduced in the *hfq* mutant, as were the levels of the filamentous hemagglutinin protein PfhB2 and its secretion partner LspB2. In contrast, there were increased expression levels of three LPS biosynthesis genes, encoding proteins involved in phosphocholine and phosphoethanolamine addition to LPS, suggesting that these genes are negatively regulated by Hfq-dependent mechanisms. Taken together, these data provide the first evidence that Hfq plays a crucial role in regulating the global expression of *P. multocida* genes, including the regulation of key *P. multocida* virulence factors, capsule, LPS, and filamentous hemagglutinin.

Small noncoding RNA (sRNA) molecules play essential roles in regulating the production of a wide range of proteins, including those involved in virulence; quorum sensing; and the metabolism of carbon, amino acids, and iron (1–5). Typically, sRNA molecules are between 50 and 400 nucleotides long and act primarily by interacting with one or more target mRNAs via short imperfect base pairing (6, 7). Protein production can be controlled by sRNAs via multiple mechanisms, including inhibition of translation, activation of translation, and alteration of transcript degradation rates (8, 9). The regulation of bacterial protein expression via sRNAs often requires the activity of Hfq, an RNA-binding chaperone protein that belongs to the Sm-like RNA-binding protein family and displays a highly conserved core sequence (6, 10). Hfq monomers form a homohexameric ring-shaped structure that preferentially binds to A/U-rich sequences on target RNA molecules and mediates the various posttranscriptional regulatory mechanisms of sRNAs (11–13). Hfq can also directly protect sRNA molecules from RNase E-mediated degradation (14) and can have an sRNA-independent role in the regulation of mRNA decay, via direct interaction with poly(A) polymerase (15).

Hfq homologues have been identified in a range of Gram-positive and Gram-negative bacterial species, including many pathogenic species (16). Most *hfq* mutants display a broad range of pleiotropic phenotypes, suggesting a global role for Hfq in bacterial physiology (6, 17, 18). These pleiotropic effects have been observed more commonly in Gram-negative species, including the pathogens *Brucella abortus*, *Salmonella enterica*, *Yersinia pestis*, *Moraxella catarrhalis*, *Neisseria meningitidis*, and *Vibrio cholerae*, but also in some Gram-positive species such as *Listeria monocytogenes* (18–24). Many *hfq* mutants show reduced growth rates, reduced virulence, and increased susceptibility to host defense mechanisms and environmental stresses (6, 16).

Pasteurella multocida is an encapsulated, Gram-negative, facultative anaerobe that is the causative agent of a number of important animal diseases, including fowl cholera (25). Fowl cholera can affect most avian species, including chickens, turkeys, ducks, and wild waterfowl, with infections resulting in high mortality rates and major economic losses to poultry industries worldwide (26, 27). *P. multocida* is a heterogeneous species, with strains being classified into five serogroups (serogroups A, B, D, E, and F) based on capsular composition (28) and into eight lipopolysaccharide (LPS) genotypes (L1 through to L8) based on the LPS outer core biosynthesis genes (29). A number of *P. multocida* virulence factors have been defined, including the polysaccharide capsule, LPS,

Received 10 February 2016 **Accepted** 11 February 2016
Accepted manuscript posted online 16 February 2016
Citation Mégroz M, Kleifeld O, Wright A, Powell D, Harrison P, Adler B, Harper M, Boyce JD. 2016. The RNA-binding chaperone Hfq is an important global regulator of gene expression in *Pasteurella multocida* and plays a crucial role in production of a number of virulence factors, including hyaluronic acid capsule. Infect Immun 84:1361–1370. doi:10.1128/IAI.00122-16.
Editor: G. H. Palmer
Address correspondence to John D. Boyce, john.boyce@monash.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/IAI.00122-16.

filamentous hemagglutinin adhesins, iron-sequestering systems, and sialic acid uptake (30–33). The polysaccharide capsule is a critical virulence factor of *P. multocida* that allows the bacteria to evade host immune defense mechanisms (34, 35). Very little is known about how the expression of these virulence factors is regulated, as only one *P. multocida* regulatory protein, Fis, has been identified and characterized. Fis has been shown to be essential for the production of capsule, and a *fis* mutant displayed highly reduced expression levels of many genes, including those within the capsule biosynthesis locus and *pfhB2*, encoding filamentous hemagglutinin (36).

In this study, the role of Hfq in *P. multocida* gene expression and protein production was characterized by using an *hfq* mutant constructed in the *P. multocida* fowl cholera strain VP161. As Hfq is a crucial modulator of sRNA action, an understanding of the role of Hfq provides an initial indication of the significance of sRNAs in the regulation of *P. multocida* genes and will underpin future work on specific sRNAs in this bacterium. The VP161 *hfq* mutant displayed highly reduced hyaluronic acid (HA) capsule production as well as reduced *in vivo* fitness. RNA sequencing (RNA-Seq) and high-throughput quantitative liquid proteomics analyses indicated that Hfq plays an important role in controlling the expression of >100 genes/proteins. In particular, genes/proteins belonging to the capsule biosynthetic locus were expressed at reduced levels in the *hfq* mutant, consistent with the reduced levels of HA capsule production. These analyses also revealed that Hfq plays a role in the production of filamentous hemagglutinin and proteins involved in LPS biosynthesis. Taken together, these data indicate that Hfq plays an important role in the regulation of key *P. multocida* virulence factors and is essential for full *in vivo* fitness.

MATERIALS AND METHODS

Bacterial strains, plasmids, and culture conditions. The bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* was grown in lysogeny broth (LB), and *P. multocida* strains were grown in either heart infusion (HI) or brain heart infusion (BHI) broth (Oxoid). Liquid cultures were incubated at 37°C with shaking (200 rpm). Solid media were obtained by the addition of 1.5% agar. When required, the media were supplemented with antibiotics at the following concentrations: 10 μg/ml (*Escherichia coli*) or 2.5 μg/ml (*P. multocida*) tetracycline and 50 μg/ml kanamycin (*P. multocida*).

Measurement of bacterial growth rates in rich medium. Cultures of *P. multocida* strains grown overnight were subcultured at a 1:50 dilution in BHI broth and grown until cultures reached an optical density at 600 nm (OD_{600}) of 0.2 to 0.3. A 1-ml aliquot (standardized to an OD_{600} of 0.2) of each culture was transferred to 50 ml of sterile BHI broth and incubated at 37°C with shaking. To determine the growth rate, optical density measurements of the liquid cultures (at 600 nm) were taken at regular intervals by using the WPA CO8000 Biowave cell density meter, until the late exponential growth phase was reached.

Ethics statement. All animal experiments were carried out in accordance with the provisions of the Prevention of Cruelty to Animal Act, 1986; the *Australian Code of Practice for the Care and Use of Animals for Scientific Purposes* (37); and Monash University Animal Welfare Committee guidelines and policies. The protocols were approved by the Monash Animal Research Platform 2 (MARP-2) Animal Ethics Committee (AEC) of Monash University (AEC number MARP/2011/066, Understanding How *Pasteurella multocida* Causes Disease in Animals).

DNA manipulations. Restriction digests, ligations, and PCR amplifications were performed by using enzymes obtained from NEB or Roche, according to the manufacturer's instructions. Plasmid DNA was prepared by using the NucleoSpin plasmid kit (Macherey-Nagel), while genomic DNA was prepared by using the HiYield genomic DNA minikit (Real Genomics). PCR amplification of DNA was performed by using *Taq* DNA polymerase (Roche) or Phusion high-fidelity DNA polymerase (NEB) on an Eppendorf Mastercycler. Amplified DNA was purified by using the NucleoSpin gel or PCR cleanup kit (Macherey-Nagel). The primers used in this study are listed in Table 2. Column-purified DNA samples were quantified by using the NanoDrop 1000 spectrophotometer (Thermo Scientific) and/or by gel electrophoresis. Sequencing reactions were performed by using Applied Biosystems Prism BigDye terminator mix, version 3.1. Electropherograms were generated on a capillary platform Applied Biosystems 3730 genetic analyzer and analyzed by using Vector NTI Advance, version 11.5 (Life Technologies).

TABLE 1 Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant description	Source or reference
Strains		
E. coli DH5α	*deoR endA1 gyrA96 hsdR17*(r_K^- m_K^+) *recA1 relA1 supE44 thi-1* (*lacZYA-argFV169*) ϕ80*lacZ*ΔM15 F^-	Bethesda Research Laboratories
P. multocida		
VP161	Serotype A:1; Vietnamese isolate from chickens	66
AL2234	VP161 *hyaD* TargeTron mutant; Kan^r	This study
AL2521	VP161 *hfq* TargeTron mutant; Kan^r	This study
AL2526	AL2521 containing pAL1108; Kan^r Tet^r	This study
AL2527	AL2521 containing pAL99T; Kan^r Tet^r	This study
Plasmids		
pAL99T	*P. multocida-E. coli* expression plasmid (Tet^r); contains the constitutive *tpiA* promoter upstream of the cloning site; derivative of pAL99	38
pAL953	*P. multocida*-specific TargeTron plasmid (Kan^r $Spec^r$); intron targeted to the *gatA* gene	38
pAL1069	*P. multocida* TargeTron plasmid for inactivation of *hyaD*	This study
pAL1104	*P. multocida* TargeTron plasmid for inactivation of *hfq*	This study
pAL1108	Wild-type *hfq* cloned into pAL99T (Tet^r)	This study

Construction of *P. multocida hfq* and *hyaD* mutants. The TargeTron (Sigma-Aldrich) mutagenesis system was used to separately inactivate *hfq* and, as a control for HA assays, the capsular biosynthesis gene *hyaD* in *P. multocida* A:1 strain VP161, as described previously (38). The group II intron within *P. multocida* TargeTron plasmid pAL953 (Table 1) was retargeted to *hfq* or *hyaD* according to the TargeTron user manual and with specific primers designed by using the TargeTron design site (Sigma-Aldrich) (Table 2). The retargeted TargeTron plasmids pAL1104 (*hfq* targeted) and pAL1069 (*hyaD* targeted) were then used to separately transform *P. multocida* VP161 by electroporation, and transformants were selected on solid agar containing kanamycin. Following transformation, strains were cured of free-replicating plasmid as described previously (38), and putative *hfq* and *hyaD* mutants were identified by patching for Kan^r/$Spec^s$ colonies. Colony PCRs using the TargeTron-specific primer EBS universal, together with an *hfq*- or *hyaD*-specific primer (BAP7400 and BAP2067, respectively) (Table 2), were used to initially identify TargeTron mutants (data not shown). To confirm that *P. multocida hfq* or *hyaD* mutants contained only a single TargeTron insertion, Southern hybridization was performed (data not shown), using a TargeTron intron-specific digoxigenin (DIG)-labeled probe. In addition, the exact position of the intron within the genome was determined by direct sequencing using genomic DNA isolated from each TargeTron mutant as the template with the intron-specific EBS universal primer (Table 2). Two mu-

TABLE 2 Primers used in this study

Primer	Sequence (5′–3′)	Description
EBS Universal	GTAAAACGACGGCCAGT	Universal TargeTron primer, located ~250 bp from the 5′ end of the intron
BAP2067	GGAAGGAACAGTTTCTCTGGATTG	Forward primer located within VP161 *hyaD*
BAP2679	TTGTGTGGAATTGTGAGCGGA	Primer located upstream of the *tpi* promoter and MCS[a] on pAL99T
BAP7153	AAAAAAGCTTATAATTATCCTTACTGAACCAGACTGTGCGCCCAGATAGGGTG	TargeTron IBS primer specific for inactivation of *hyaD*
BAP7154	CAGATTGTACAAATGTGGTGATAACAGATAAGTCCAGACTGTTAACTTACCTT TCTTTGT	TargeTron EBS1d primer specific for inactivation of *hyaD*
BAP7155	TGAACGCAAGTTTCTAATTTCGGTTTTCAGTCGATAGAGGAAAGTGTCT	TargeTron EBS2 primer specific for inactivation of *hyaD*
BAP7389	TGAACGCAAGTTTCTAATTTCGATTCTTGTTCGATAGAGGAAAGTGTCT	TargeTron EBS2 primer specific for inactivation of *hfq*
BAP7390	AAAAAAGCTTATAATTATCCTTAACAAGCTCAGATGTGCGCCCAGATAGGGTG	TargeTron IBS primer specific for inactivation of *hfq*
BAP7391	CAGATTGTACAAATGTGGTGATAACAGATAAGTCTCAGATTGTAACTTACCTT TCTTTGT	TargeTron EBS1d primer specific for inactivation of *hfq*
BAP7400	CAGAGGATCCTGTGTTTAGCTAGTTGTATC	Forward primer located upstream of VP161 *hfq*; contains BamHI site
BAP7401	AGATGTCGACCACTTGCAATAAATGTGTTC	Reverse primer located downstream of VP161 *hfq*; contains SalI site

[a] MCS, multiple-cloning site.

tants that generated unambiguous sequencing data identical to data for the correct target gene sequence (immediately adjacent to the intron insertion site) were selected for further study and designated AL2521 (VP161 *hfq* mutant) and AL2234 (VP161 *hyaD* mutant).

Complementation of the *hfq* mutant. For complementation of the *hfq* mutant, an intact copy of the *hfq* gene was amplified from VP161 genomic DNA by using primers BAP7400 and BAP7401 that contained BamHI and SalI restriction sites, respectively (Table 2). After digestion with BamHI and SalI, the amplified fragment was cloned into BamHI- and SalI-digested pAL99T (Table 1), such that the *P. multocida* constitutive promoter *tpiA* would drive the transcription of *hfq*. The ligated products were then used to transform *E. coli* strain DH5α. *E. coli* transformants containing the correct plasmid were identified by colony PCR using pAL99T-specific and *hfq*-specific primers BAP2679 and BAP7401, respectively. DNA sequencing was used to confirm the fidelity of the *hfq* sequence, and one correct recombinant plasmid was selected and designated pAL1108. The vector pAL99T and the complementation plasmid pAL1108 were then used to separately transform the VP161 *hfq* mutant (AL2521) via electroporation, generating strains AL2527 and AL2526, respectively (Table 1).

Quantitative hyaluronic acid capsule assay. The amount of HA capsular material produced by each *P. multocida* strain was determined as described previously (34). Student's unpaired *t* test was used to assess the differences in HA production between strains.

Competitive *in vivo* growth assays. Competitive growth assays in mice were performed to quantify the relative *in vivo* fitness of the *hfq* mutant and complemented strains compared with wild-type strain VP161, as described previously (39). To identify each bacterial population, after overnight incubation of input and output samples on HI medium, individual colonies were patched onto HI medium and either HI medium containing kanamycin (to select for the *hfq* mutant) or HI medium containing kanamycin and tetracycline (to select for the *hfq* mutant containing the plasmid). Mutants with a competitive index (CI) of <0.5 were identified as having significantly reduced *in vivo* fitness, and Student's unpaired two-tailed *t* test was also used to assess the statistical differences in relative CIs between each of the tested strains.

RNA extraction and purification, rRNA depletion, library preparation, and high-throughput RNA sequencing. *P. multocida* strains AL2521 (VP161 *hfq* mutant) and VP161 (wild type) were each grown in BHI broth to the mid-exponential growth phase (OD_{600} = 0.6; ~1.0 × 10^9 CFU/ml). A total of 2.5 ml of ice-cold killing buffer (containing 0.05 M Tris-HCl [pH 7.5], 15 mg/ml sodium azide, and 0.6 mg/ml chloramphenicol) was added to each 25-ml culture, and the mixed sample was cooled on ice before centrifugation at 9,000 × *g* for 10 min (4°C). Pelleted cells were then resuspended in 1 ml RNAlater stabilization reagent (Qiagen) and incubated for 5 min at room temperature before centrifugation at 5,000 × *g* for 10 min. Total RNA was purified by using the RNeasy minikit (Qiagen) with DNase treatment according to the manufacturer's instructions. RNA samples were quantified and analyzed for DNA contamination by using the Qubit kit (Invitrogen) and visualized by gel electrophoresis. rRNA was removed by using the Ribo-Zero magnetic kit (Illumina) according to the manufacturer's instructions. The TruSeq RNA sample preparation kit (Illumina) was used to construct cDNA libraries, with library validation, normalization, and pooling performed by Micromon Services (Monash University). All libraries were combined and sequenced on a single lane of a HiSeq 2000 instrument (Macrogen, South Korea). Mapping of the RNA-Seq reads to the draft VP161 genome was carried out by using CLC Genomics Workbench v 7.0 (CLC Bio), and statistical analyses were carried out by using voom and limma as described previously (40). Raw combined data sets of 50,071,021 and 47,130,114 total paired reads were generated for AL2521 and VP161, respectively. Of these, 17,718,215 and 14,903,935 paired reads (for AL2521 and VP161, respectively) aligned uniquely to the *P. multocida* VP161 genome (~30% unambiguous reads for each combined data set), giving high coverage of the transcriptome. Genes were identified as being differentially expressed if they displayed a ≥2.0-fold (≥1.0-$\log_2$) change in expression across the triplicate replicates at a false discovery rate (FDR) of <0.05. The differentially expressed VP161 genes identified were mapped to their closest orthologue in the fully annotated reference genome of *P. multocida* strain Pm70 (41) so that overrepresented gene ontology groups and pathway associations could be determined via BioCyc (42), as we reported previously (40).

Quantitative proteomics and mass spectrometry. Cultures of each *P. multocida* strain were grown to either the early exponential (OD_{600} = 0.2) or mid-exponential (OD_{600} = 0.6) growth phase and then centrifuged at 9,400 × *g* for 10 min. The pelleted cells were washed twice in 100 mM Tris-HCl (pH 7.5) and lysed via the addition of 0.35 ml of lysis buffer (4% SDS, 100 mM dithiothreitol [DTT], 100 mM Tris-HCl [pH 7.5]) and

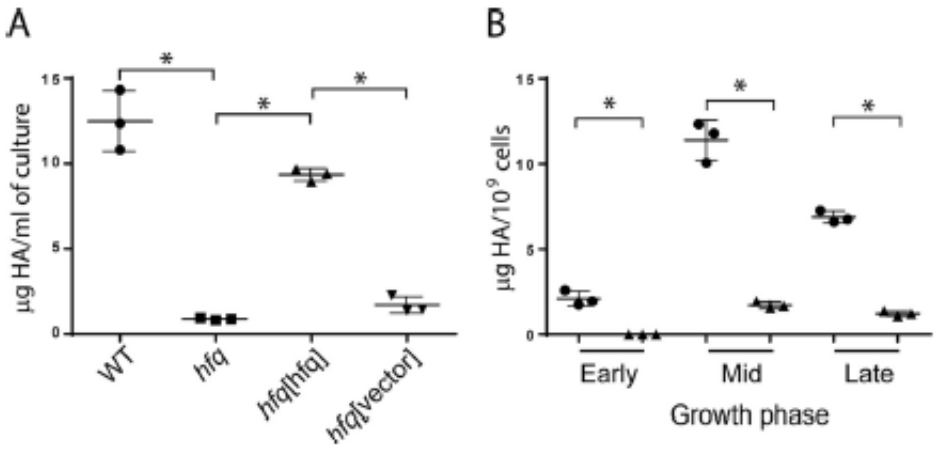


FIG 1 Hyaluronic acid (HA) polysaccharide capsule produced by *Pasteurella multocida* strains. (A) Amounts of HA produced at the mid-exponential growth phase by wild-type strain VP161 (WT), the *hfq* mutant, the complemented *hfq* mutant (*hfq*[hfq]), and the *hfq* mutant with the empty vector (*hfq*[vector]). (B) Amounts of HA detected (per 10^9 cells) in wild-type (●) and *hfq* mutant (▲) cells at the early exponential (OD_{600} = 0.20 to 0.25), mid-exponential (OD_{600} = 0.65 to 0.75), and late exponential (OD_{600} = 1.10 to 1.40) growth phases. The average amount of HA under each condition is indicated by the horizontal bars, ± standard deviations of the means. *, *P* value of <0.05 as determined by an unpaired *t* test.

incubation for 10 min at 99°C. The protein concentration was determined by using the Bradford assay (43). Proteins (150 µg/sample) were purified and trypsin digested by using a filter-aided sample preparation (FASP) protein digestion kit (Expedeon) as described previously (44), except that triethyl ammonium bicarbonate was used instead of ammonium bicarbonate. The digested proteins were isotopically labeled by dimethylation using heavy and light formaldehyde as described previously (45). Heavy and light samples were then mixed and desalted by using an Empore 4-mm/1-ml extraction disk cartridge (C_{18}-SD) (3M). Desalted samples were concentrated under a vacuum and resuspended in 0.1% formic acid and 2% acetonitrile. Peptides were resolved and analyzed by nano-liquid chromatography (LC)–tandem mass spectrometry using a Dionex ultra-high-performance liquid chromatography system coupled to a Thermo Scientific Q-Exactive Orbitrap mass spectrometer (located at the Monash Biomedical Proteomics Facility). Proteins were identified and quantified with MaxQuant (46) and Perseus software. For differential expression testing, the protein expression ratios were log transformed. These transformed ratios were roughly normally distributed, so limma (47) was used to fit and test for the significance of differential expression. Those proteins showing a ≥2.0-fold (≥1.0-log_2) change in expression across the triplicate replicates at an FDR of <0.05 were considered to be differentially expressed.

Accession numbers. The RNA-Seq data are available at the NCBI Gene Expression Omnibus under accession number GSE77721. The proteomics data have been deposited in ProteomeXchange via the PRIDE database under accession number PXD003586.

RESULTS

The *hfq* mutant produces reduced hyaluronic acid capsule. To investigate the role of Hfq in *P. multocida*, an *hfq* mutant was generated in the highly virulent strain VP161. The mutant grew indistinguishably from the wild-type strain in BHI liquid broth at 37°C *in vitro* (see Fig. S1 in the supplemental material), but on solid medium, it produced smaller and less mucoid colonies than those of wild-type strain VP161, consistent with reduced HA capsular expression. The amount of HA was quantified in mid-exponential-phase cultures of wild-type strain VP161, the *hfq* mutant (AL2521), the complemented *hfq* mutant (AL2526), and the *hfq* mutant containing the vector only (AL2527). The *hfq* mutant produced significantly less HA than did the wild-type strain, and HA production was restored to near-wild-type levels when the mutant was complemented with an intact copy of *hfq* (Fig. 1A). The control *hyaD* mutant produced no detectable HA (data not shown).

The amount of HA capsule produced by the *hfq* mutant was also assessed at different growth phases *in vitro*. The wild-type strain produced small amounts of HA at the early exponential growth phase, maximal amounts of HA at the mid-exponential growth phase, and intermediate levels of HA at the late exponential growth phase (Fig. 1B). No detectable HA was produced by the *hfq* mutant at the early exponential growth phase, and significantly reduced levels of HA were produced at the mid- and late exponential growth phases (Fig. 1B). Again, the *hyaD* acapsular control strain produced no detectable HA at any of the tested growth phases (data not shown). These data indicate that capsule production in *P. multocida* is growth phase dependent, and inactivation of *hfq* results in significantly reduced capsule expression during all tested growth phases *in vitro*.

The *hfq* mutant displays reduced *in vivo* fitness in mice. As the HA capsule is a known virulence factor of *P. multocida* serogroup A strains (34), competitive growth assays were performed in mice to compare the *in vivo* fitness of the wild-type strain with that of the *hfq* mutant, the *hfq* mutant complemented with an intact copy of *hfq*, and the *hfq* mutant containing the vector only. The *hfq* mutant displayed a 4-fold reduction in *in vivo* growth compared to that of the wild-type parent strain VP161 (CI = 0.25 ± 0.05) (Fig. 2). Importantly, this loss of *in vivo* fitness was restored to near-wild-type levels when the *hfq* mutant was provided with an intact copy of *hfq* in *trans* (CI = 0.68 ± 0.06) (Fig. 2). These data indicate that the *P. multocida* VP161 Hfq is important for *in vivo* fitness in mice.

Global RNA expression changes in the *P. multocida hfq* mutant. Hfq is an RNA chaperone that has been shown in many other bacterial species to be essential for the action of most *trans*-encoded regulatory sRNA molecules (16). Hfq and associated sRNAs can act at the level of mRNA stability or at the level of translational

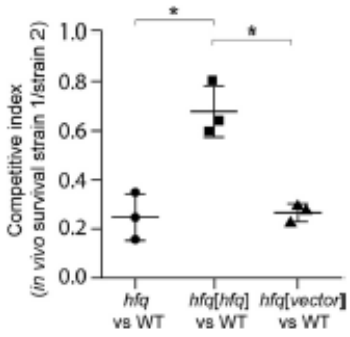


FIG 2 The *hfq* mutant is attenuated for growth in mice. Shown are competitive growth indices for the *P. multocida* strains in a mouse infection model. Each data point shows the individual CI for one pair of strains in one mouse. The average CI for each pair of competing strains is indicated by the horizontal bars, ± standard deviations of the means. * indicates a *P* value of <0.05 as determined by a two-tailed unpaired *t* test.

efficiency (11). To identify changes in transcript abundance, the transcriptomes of the wild-type strain and the *hfq* mutant (grown to the mid-exponential growth phase) were compared by using high-throughput RNA-Seq. The RNA-Seq reads were mapped to the *P. multocida* VP161 genome, and 128 genes were identified as being differentially expressed. More than 85% of the differentially expressed genes (109 of 128 genes) showed increased expression levels compared to those in the wild type (see Table S1 in the supplemental material), suggesting that Hfq and its associated sRNAs generally act to decrease transcript levels, in keeping with the most common mode of sRNA/Hfq action at the transcriptional level, which is to increase RNase E turnover of transcripts.

Functional groups overrepresented in the upregulated gene set included those involved in cytochrome complex assembly (GO:0017004; 6 genes; $P = 2 \times 10^{-6}$), organic substance transport (GO:0071702; 13 genes; $P = 6 \times 10^{-4}$), nitrogen compound transport (GO:0071705; 7 genes; $P = 8 \times 10^{-4}$), and protein-binding transcription factors (GO:0000988; 3 genes; $P = 0.001$). Among the upregulated genes were *pcgB*, *pcgD*, PM1042 (PMVP_1156, PMVP_1157, and PMVP_1057, respectively), encoding proteins involved in LPS biosynthesis (48; M. Harper and J. D. Boyce, unpublished data), and *hsf* and PM0855 (PMVP_0693 and PMVP_0855, respectively), predicted to encode the Hsf adhesin and a Flp pilus-like adhesin, respectively (39). The periplasmic nitrate reductase (*nap*) genes (*napFDAGHBC*) were all strongly upregulated (7.5- to 11.6-fold) in the *hfq* mutant, as were a number of genes encoding other putative electron transport proteins, including *ccmABCDEF* (2.0- to 4.1-fold) and *nrfBCD* (3-fold). In addition, numerous genes encoding proteins involved in heat shock and stress responses were also expressed at higher levels in the *hfq* mutant, including *htpX*, *uspG*, *htrA*, *dnaK*, *rpoH*, *clpB*, *rseC*, *rseB*, *mclA*, and *rpoE* (PMVP_0444, PMVP_0450, PMVP_0713, PMVP_0715, PMVP_1638, PMVP_1759, PMVP_1836, PMVP_1837, PMVP_1838, and PMVP_1839, respectively). An association of Hfq with the regulation of stress response genes/proteins has been seen in other bacteria, including *Y. pestis*, *Clostridium difficile*, and *S. enterica* (20, 49, 50).

Nineteen genes displayed decreased transcript levels in the *hfq* mutant compared to the wild-type strain (see Table S2 in the supplemental material). Functional groups overrepresented in this set included polysaccharide transport (GO:0015774; two genes; $P = 4 \times 10^{-4}$) and pyridoxal phosphate metabolism (GO:0042822; two genes; $P = 0.001$). Importantly, 5 of the 10 capsule biosynthesis genes showed significantly reduced transcript expression in the *hfq* mutant ($P < 0.0001$) (see Table 5), correlating with the reduced HA capsule production observed for the *hfq* mutant.

Global changes in protein production in the *P. multocida hfq* mutant. Hfq often regulates protein production by altering mRNA translational efficiency via sRNA binding (7). Therefore, the proteomes of the VP161 *hfq* mutant and wild-type strain VP161 were compared during both early exponential and mid-exponential growth phases by using quantitative high-throughput liquid proteomics. Totals of 1,147 and 1,041 proteins were identified during early exponential and mid-exponential growth, respectively. During the early exponential growth phase of the *hfq* mutant, 85 proteins were detected at increased levels and 21 proteins were detected at reduced levels compared to those in the wild-type strain (see Tables S3 and S4, respectively, in the supplemental material). During the mid-exponential growth phase, 78 proteins were identified as being differentially expressed in the *hfq* mutant: 49 at increased levels (see Table S5 in the supplemental material) and 29 at decreased levels (see Table S6 in the supplemental material). Comparing the data from both growth phases, 37 proteins were expressed at increased levels at both the early exponential and mid-exponential growth phases (Table 3), and 8 proteins were expressed at decreased levels at both growth phases (Table 4). Five of the capsule biosynthesis proteins (HyaE, HyaD, HyaC, HexD, and HexC) showed between 10- and 20-fold-reduced production at the early exponential growth phase and between 1.8- and 2.7-fold-reduced production at the mid-exponential growth phase. These data correlated well with the decreased transcript levels observed for the capsule biosynthesis genes as measured by RNA-Seq (Table 5) and with the reduced HA capsule levels expressed by this strain. In addition, the known *P. multocida* virulence factor PfhB2 and its secretion partner LspB2 were also present at reduced levels in the *hfq* mutant at both growth phases (Table 4). In contrast, the levels of production of the hemoglobin-binding proteins PMVP_0265 (PM0300) and PMVP_0304 (PM0337) were increased in the *hfq* mutant at both the early exponential and mid-exponential growth phases (Table 3).

Comparison of the proteomic and transcriptomic data. The proteomic analysis of the *hfq* mutant at the mid-exponential growth phase (see above) identified 78 proteins as being differentially expressed, with 49 being identified at increased levels and 29 at decreased levels. Of the 49 proteins produced at higher levels in the *hfq* mutant, 13 also showed increased transcript levels as measured by RNA-Seq; all 13 of these proteins also showed increased levels (as measured by proteomics) at the early exponential growth phase (see Table S7 in the supplemental material). The 13 genes/proteins that showed increased expression levels by both techniques included 2 outer membrane lipoproteins (Plp4 and PlpB), 3 components of the periplasmic nitrate reductase system, the universal stress protein UspG, and the oligopeptide transporter DppA. Similarly, of the 29 proteins measured at reduced levels in the *hfq* mutant by high-throughput proteomics, 6 were also downregulated as determined by RNA-Seq (see Table S8 in the supplemental material), and 5 of these also showed decreased expression at the early exponential growth phase (see Table S8 in the supplemental material). The six genes/proteins determined to be downregulated in the *hfq* mutant by both proteomics and tran-

TABLE 3 Proteins showing increased production in the *P. multocida hfq* mutant strain in comparison to wild-type strain VP161 during both early exponential and mid-exponential growth phases as measured by high-throughput liquid proteomics

VP161 locus tag	Predicted function and/or product	Pm70 protein name	Early exponential growth		Mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0194	Periplasmic dipeptide transporter	DppA	2.59	2.82E−04	2.44	1.06E−04
PMVP_0198	Dipeptide ABC transporter, ATP-binding protein	DppF	2.21	4.28E−03	1.71	1.39E−03
PMVP_0229	Murein L,D-transpeptidase, provisional	PM0270	1.57	4.66E−04	1.73	4.44E−04
PMVP_0253	Predicted membrane-bound lysozyme inhibitor of the c type	None	3.30	6.48E−04	2.06	2.53E−04
PMVP_0265	Hemoglobin/haptoglobin-binding protein	PM0300	1.79	5.84E−04	1.28	2.32E−03
PMVP_0267	Mrp putative ATPase	Mrp	1.18	5.19E−04	1.32	1.77E−03
PMVP_0292	Transglycosylase SLT domain protein	PM0326	1.56	1.69E−03	2.85	1.85E−03
PMVP_0304	Hemoglobin-binding protein with TonB receptor	PM0337	1.49	6.40E−03	1.49	9.02E−04
PMVP_0305	Hemoglobin-binding protein with TonB receptor	PM0337	1.83	1.75E−03	1.32	3.53E−03
PMVP_0327	MscS family potassium efflux protein, KefA	PM0358	1.07	8.84E−04	1.00	7.98E−03
PMVP_0376	Formate dehydrogenase accessory protein	FdhE	1.29	1.16E−03	1.00	3.62E−02
PMVP_0449	TRAP[a] transporter/C_4-dicarboxylate ABC transporter	PM0473	1.58	5.48E−04	1.78	8.79E−04
PMVP_0450	Universal stress protein, UspG	PM0474	1.21	5.03E−04	1.35	1.74E−03
PMVP_0488	Conserved protein with no known function	PM0519	2.10	4.44E−04	1.81	1.77E−03
PMVP_0516	Glucose-1-phosphate adenylyltransferase	GlgC	1.46	5.48E−04	1.56	3.39E−02
PMVP_0541	Branched-chain amino acid aminotransferase	IlvE	1.39	5.84E−04	1.10	5.18E−03
PMVP_0561	Lipoprotein	Plp4	1.77	4.44E−04	1.31	1.85E−03
PMVP_0629	Tellurite resistance protein	TehB	3.50	8.84E−04	2.18	4.17E−03
PMVP_0756	ABC-ATPase-MoxR-like	None	1.18	1.58E−02	1.47	9.02E−04
PMVP_0837	Histidinol-phosphate aminotransferase	HisH1	1.26	1.01E−03	1.41	1.47E−03
PMVP_0852	Protein within Flp/tight adhesion locus, homology to RpcC	PM0853	2.00	7.94E−04	1.74	5.98E−04
PMVP_0939	Membrane-bound lytic murein transglycosylase	PM0928	1.66	7.94E−04	1.99	3.96E−04
PMVP_0948	Aspartate kinase, monofunctional class	PM0937	1.33	5.19E−04	1.21	2.57E−03
PMVP_0992	Predicted membrane/periplasmic protein of unknown function	PM0979	2.08	2.82E−04	2.05	4.04E−04
PMVP_1069	Dihydrodipicolinate synthase	DapA	1.11	7.94E−04	1.01	8.68E−03
PMVP_1208	Type I DNase HsdR (HtpX peptidase M48 superfamily)	PM1190	1.34	7.27E−03	1.51	8.79E−04
PMVP_1233	Transglutaminase-like superfamily domain protein	PM1211	1.41	8.19E−04	1.98	2.08E−03
PMVP_1614	GTA[b] type/β-D-1,6-glucosyl transferase	PM1562	2.46	4.44E−04	2.45	1.06E−04
PMVP_1616	Large conductance mechanosensitive channel protein	MscL	2.12	7.32E−03	2.55	2.42E−04
PMVP_1649	Periplasmic nitrate reductase precursor	NapA	1.59	8.84E−04	1.29	1.81E−03
PMVP_1652	Periplasmic nitrate reductase	NapB	1.12	2.91E−02	1.75	5.47E−03
PMVP_1760	Thioredoxin	Trx	1.87	4.88E−03	1.73	1.33E−02
PMVP_1787	Outer membrane lipoprotein 2 precursor	PlpB	1.32	5.03E−04	1.09	9.91E−03
PMVP_1863	Nitrate/nitrite response regulator protein	NarP	2.51	2.82E−04	2.05	2.53E−04
PMVP_1872	SrfB homologue, uncharacterized protein conserved in bacteria	PM1819	1.21	2.04E−03	1.21	7.88E−03
PMVP_1879	Protein of unknown function	PM1826	1.86	4.66E−04	1.58	8.79E−04
PMVP_1962	Periplasmic oligopeptide-binding protein precursor	OppA	1.26	5.19E−04	1.17	3.72E−03
PMVP_2095	Glutamate dehydrogenase	GdhA	2.01	1.75E−03	2.71	1.06E−04

[a] TRAP, tripartite ATP-independent periplasmic.
[b] GTA, glycosyltransferase A.

scriptomics analyses included three capsule biosynthesis proteins, the filamentous hemagglutinin secretion partner LspB, the global regulator of carbon metabolic flux FruR (51), and the imidazole glycerol phosphate synthase, which is a key enzyme that links amino acid and nucleotide biosynthesis pathways (52). Interestingly, 59 of the 78 proteins identified as being differentially expressed in the *hfq* mutant by proteomics analysis were not determined to be differentially expressed by transcriptomics analysis. It is likely that many of these proteins are regulated by Hfq-sRNA interactions that act posttranscriptionally, a common mechanism of sRNA action (7).

DISCUSSION

Inactivation of *hfq* in different bacterial species has resulted in a wide range of phenotypic effects. For some species, such as *Brucella melitensis*, *Legionella pneumophila*, and *Salmonella enterica*, *hfq* mutants show a reduced growth rate with a longer lag phase in certain *in vitro* media (53–55). In other bacterial species, such as *Listeria monocytogenes*, *Staphylococcus aureus*, *Haemophilus influenzae* (24, 56, 57), and *P. multocida* (as shown in this study for strain VP161), the growth rate *in vitro* in rich medium is unaffected by *hfq* inactivation. Despite this normal *in vitro* growth rate, competition assays in mice revealed that the *P. multocida hfq* mutant showed significantly reduced fitness *in vivo*. Reduced *in vivo* fitness has also been demonstrated for *hfq* mutant strains of *Y. pestis*, *Klebsiella pneumoniae*, *Francisella tularensis*, and *Actinobacillus pleuropneumoniae* (20, 58–60).

One of the critical *P. multocida* virulence factors is the polysaccharide capsule (34). The *P. multocida hfq* mutant showed significantly reduced HA capsular polysaccharide (CPS) production

TABLE 4 Proteins showing decreased production in the *P. multocida hfq* mutant strain in comparison to wild-type strain VP161 during both early exponential and mid-exponential growth phases as measured by high-throughput liquid proteomics

VP161 locus tag	Predicted function and/or product	Pm70 protein name	Early exponential growth		Mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0001	LspB	LspB2	−2.74	8.84E−04	−1.83	1.85E−03
PMVP_0002[a]	Filamentous hemagglutinin	PfhB2	−2.81	4.44E−04	−1.85	1.92E−03
PMVP_0003[a]	Filamentous hemagglutinin	PfhB2	−2.82	4.66E−04	−1.82	2.21E−03
PMVP_0004[a]	Filamentous hemagglutinin	PfhB2	−2.70	4.44E−04	−1.77	3.58E−03
PMVP_0005[a]	Filamentous hemagglutinin	PfhB2	−2.19	8.19E−04	−1.81	1.14E−02
PMVP_0006[a]	Filamentous hemagglutinin	PfhB2	−1.69	5.03E−04	−1.86	1.85E−03
PMVP_0218	Spermidine/putrescine ABC transporter, periplasmic-binding protein	PotD1	−1.88	6.96E−04	−1.58	8.79E−04
PMVP_0767	Capsule biosynthesis, HyaE	HyaE	−4.17	2.82E−04	−1.11	2.63E−02
PMVP_0768	Capsule biosynthesis, HyaD	PM0775	−4.35	2.82E−04	−1.06	1.95E−02
PMVP_0769	Capsule biosynthesis, HyaC	PM0776	−3.86	1.22E−03	−1.41	5.60E−03
PMVP_0771	Capsule transport, HexD	HexD	−3.29	4.50E−04	−1.01	1.33E−02
PMVP_0869	DNA-binding transcriptional regulator, fructose repressor	FruR	−1.99	1.00E−02	−1.31	4.81E−03

[a] These features likely represent broken sections of one gene (*pfhB2*) that have not been fully assembled in the draft VP161 genome.

(up to 14-fold) at all tested growth phases. Transcriptomic and proteomic analyses of the *P. multocida hfq* mutant confirmed the phenotypic data, showing highly reduced transcript and protein levels for many of the capsule biosynthesis genes/proteins (Table 5). Thus, Hfq plays a crucial role in positively regulating capsule production. The reduced level of capsule biosynthesis gene transcripts suggests that Hfq-mediated regulation acts primarily at the level of transcription; however, we cannot rule out an additional role for posttranscriptional mechanisms. In contrast to the effect of Hfq on *P. multocida* capsule production, a *K. pneumoniae hfq* mutant is hypermucoid, with increased levels of K2-specific CPSs (58). The reduction in capsule production in the *P. multocida hfq* mutant is likely to play a major role in the reduced *in vivo* fitness observed for this strain.

We have previously shown that the global transcriptional regulator Fis is essential for the expression of *P. multocida* capsule biosynthesis genes (36). Our proteomic and transcriptomic analyses indicated that Fis was produced at unchanged (proteomics data) or slightly increased (transcriptomics data) levels (1.8-fold; FDR = 0.03) in the *hfq* mutant compared to the wild-type strain, implying that the loss of capsule expression in the *hfq* mutant was not the result of a loss of Fis production. However, Fis may play a role in the regulation of *hfq* expression, as RNA-Seq analysis of a *P. multocida* VP161 *fis* mutant during the early exponential growth phase revealed a 2.2-fold decrease in *hfq* expression (FDR = 0.001) compared to that of the wild-type strain (Harper and Boyce, unpublished). These data suggest that Fis positively regulates Hfq, which in turn positively regulates capsule expression. However, mutation of *fis* alone has a much larger effect on capsule expression than mutation of *hfq* alone, suggesting that Fis primarily acts separately from Hfq.

P. multocida expresses two filamentous hemagglutinins, PfhB1 and PfhB2, that are important for *in vivo* fitness (31, 61). Our proteomic analyses showed that the loss of Hfq leads to a reduction in the amounts of PfhB2 and the secretion partner LspB2 during both early exponential and mid-exponential growth. It is possible that the reduced expression of PfhB2 may have contributed, along with reduced capsule expression, to the reduced *in vivo* fitness of the *hfq* mutant. Our transcriptomic analyses showed that genes encoding these proteins were also likely downregulated

TABLE 5 Differential expression of capsule biosynthesis genes/proteins in the *P. multocida hfq* mutant strain in comparison to wild-type strain VP161

VP161 locus tag	Pm70 name	Gene name	RNA-Seq data for mid-exponential growth		Proteomics data for early exponential growth		Proteomics data for mid-exponential growth	
			Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0765	*phyB*	*phyB*	−1.35	1.02E−03	ND[a]	ND	0.12	9.24E−01
PMVP_0766	*phyA*	*phyA*	−1.86	7.35E−04	−0.91	5.18E−02	−0.65	2.03E−01
PMVP_0767	*hyaE*	*hyaE*	−1.78	2.16E−03	−4.17	2.82E−04	−1.11	2.63E−02
PMVP_0768	PM0775	*hyaD*	−1.84	2.08E−03	−4.35	2.82E−04	−1.06	1.95E−02
PMVP_0769	PM0776	*hyaC*	−1.41	7.48E−03	−3.86	1.22E−03	−1.41	5.60E−03
PMVP_0770	PM0777	*hyaB*	−0.67	2.90E−02	ND	ND	ND	ND
PMVP_0771	*hexD*	*hexD*	−0.60	5.78E−02	−3.29	4.50E−04	−1.01	1.33E−02
PMVP_0772	*hexC*	*hexC*	−0.68	4.22E−02	−3.44	4.44E−04	−0.88	1.90E−02
PMVP_0773	*hexB*	*hexB*	−0.62	3.90E−02	ND	ND	ND	ND
PMVP_0774	*hexA*	*hexA*	−0.46	4.09E−02	0.30	6.42E−01	−0.71	6.89E−02

[a] ND indicates that the level of the protein was not measured under the specified conditions.

(PfhB2 was 1.4-fold downregulated, with an FDR of 0.056, and LspB2 was 2.3-fold downregulated, with an FDR of 0.01), indicating that the Hfq-associated regulation of these genes is also likely to be at the transcriptional level.

LPS is another crucial *P. multocida* virulence factor; strains expressing truncated LPS are highly attenuated for virulence (62, 63). Three genes involved in LPS biosynthesis (*pcgB*, *pcgD*, and PM1042) were expressed at increased levels in the *hfq* mutant. PcgB and PcgD are required for the biosynthesis and transfer of phosphocholine residues to the terminal galactose residues of the LPS produced by strain VP161 (48), and a functional PM1042 protein is required for the addition of phosphoethanolamine to lipid A (Harper and Boyce, unpublished). Both phosphoethanolamine and phosphocholine are important charged LPS substituents that can affect the interaction of host antimicrobial peptides with the bacterial surface (48). Numerous genes/proteins involved in electron transport were also upregulated in the *hfq* mutant, including *napFDAGHBC*, *ccmABCDEF*, and *nrfBCD*. Interestingly, the Nap system is predicted to reduce nitrate to nitrite during anaerobic growth (64), and a previous microarray study of *P. multocida* (strain X73) showed that the expression of the *nap* genes was significantly upregulated during *in vivo* growth in the chicken host, a relatively anaerobic niche (65). Thus, appropriate regulation of these genes appears critical for full *in vivo* fitness.

In conclusion, this study defined a clear role for Hfq in the regulation of *P. multocida* gene and protein expression and allowed us to identify genes controlled (either directly or indirectly) by Hfq-sRNA interactions at both the transcriptional and translational levels. Together, transcriptional and proteomic analyses of the *hfq* mutant supported our hypothesis that Hfq is involved in the regulation of virulence- and fitness-associated genes of *P. multocida*. Hfq is clearly involved in regulating the production of the known *P. multocida* virulence factors HA capsule, filamentous hemagglutinin, and LPS as well as accessory virulence genes such as those encoding iron uptake proteins, stress response proteins, and bacterial adhesins. We are currently focusing on identifying specific sRNA-mRNA interactions and in particular the specific mechanism(s) by which Hfq/sRNA action regulates the expression of HA capsule, LPS, and filamentous hemagglutinin.

ACKNOWLEDGMENTS

We thank Deanna Deveson Lucas for assistance with RNA-Seq sample preparation and Marietta John for construction of the VP161 *hyaD* mutant.

This work was supported in part by the Australian Research Council.

REFERENCES

1. **Romby P, Vandenesch F, Wagner EGH.** 2006. The role of RNAs in the regulation of virulence-gene expression. Curr Opin Microbiol 9:229–236. http://dx.doi.org/10.1016/j.mib.2006.02.005.
2. **Lenz DH, Mok KC, Lilley BN, Kulkarni RV, Wingreen NS, Bassler BL.** 2004. The small RNA chaperone Hfq and multiple small RNAs control quorum sensing in *Vibrio harveyi* and *Vibrio cholerae*. Cell 118:69–82. http://dx.doi.org/10.1016/j.cell.2004.06.009.
3. **Sonnleitner E, Abdou L, Haas D.** 2009. Small RNA as global regulator of carbon catabolite repression in *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 106:21866–21871. http://dx.doi.org/10.1073/pnas.0910308106.
4. **Sharma CM, Papenfort K, Pernitzsch SR, Mollenkopf HJ, Hinton JCD, Vogel J.** 2011. Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA. Mol Microbiol 81:1144–1165. http://dx.doi.org/10.1111/j.1365-2958.2011.07751.x.
5. **Massé E, Gottesman S.** 2002. A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli*. Proc Natl Acad Sci U S A 99:4620–4625. http://dx.doi.org/10.1073/pnas.032066599.
6. **Sobrero P, Valverde C.** 2012. The bacterial protein Hfq: much more than a mere RNA-binding factor. Crit Rev Microbiol 38:276–299. http://dx.doi.org/10.3109/1040841X.2012.664540.
7. **Vogel J, Luisi BF.** 2011. Hfq and its constellation of RNA. Nat Rev Microbiol 9:578–589. http://dx.doi.org/10.1038/nrmicro2615.
8. **Gottesman S, Storz G.** 2011. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol 3:a003798. http://dx.doi.org/10.1101/cshperspect.a003798.
9. **Bobrovskyy M, Vanderpool CK.** 2013. Regulation of bacterial metabolism by small RNAs using diverse mechanisms. Annu Rev Genet 47:209–232. http://dx.doi.org/10.1146/annurev-genet-111212-133445.
10. **Moller T, Franch T, Hojrup P, Keene DR, Bächinger HP, Brennan RG, Valentin-Hansen P.** 2002. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. Mol Cell 9:23–30. http://dx.doi.org/10.1016/S1097-2765(01)00436-1.
11. **Schumacher MA, Pearson RF, Møller T, Valentin-Hansen P, Brennan RG.** 2002. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. EMBO J 21:3546–3556. http://dx.doi.org/10.1093/emboj/cdf322.
12. **De Lay N, Schu DJ, Gottesman S.** 2013. Bacterial small RNA-based negative regulation: Hfq and its accomplices. J Biol Chem 288:7996–8003. http://dx.doi.org/10.1074/jbc.R112.441386.
13. **Brennan RG, Link TM.** 2007. Hfq structure, function and ligand binding. Curr Opin Microbiol 10:125–133. http://dx.doi.org/10.1016/j.mib.2007.03.015.
14. **Folichon M, Arluison V, Pellegrini O, Huntzinger E, Régnier P, Hajnsdorf E.** 2003. The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. Nucleic Acids Res 31:7302–7310. http://dx.doi.org/10.1093/nar/gkg915.
15. **Mohanty BK, Maples VF, Kushner SR.** 2004. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*. Mol Microbiol 54:905–920. http://dx.doi.org/10.1111/j.1365-2958.2004.04337.x.
16. **Chao Y, Vogel J.** 2010. The role of Hfq in bacterial pathogens. Curr Opin Microbiol 13:24–33. http://dx.doi.org/10.1016/j.mib.2010.01.001.
17. **Tsui HCT, Leung HCE, Winkler ME.** 1994. Characterization of broadly pleiotropic phenotypes caused by an *hfq* insertion mutation in *Escherichia coli* K-12. Mol Microbiol 13:35–49. http://dx.doi.org/10.1111/j.1365-2958.1994.tb00400.x.
18. **Robertson GT, Roop RMI.** 1999. The *Brucella abortus* host factor I (HF-I) protein contributes to stress resistance during stationary phase and is a major determinant of virulence in mice. Mol Microbiol 34:690–700. http://dx.doi.org/10.1046/j.1365-2958.1999.01629.x.
19. **Figueroa-Bossi N, Lemire S, Maloriol D, Balbontín R, Casadesús J, Bossi L.** 2006. Loss of Hfq activates the σE-dependent envelope stress response in *Salmonella enterica*. Mol Microbiol 62:838–852. http://dx.doi.org/10.1111/j.1365-2958.2006.05413.x.
20. **Geng J, Song Y, Yang L, Feng Y, Qiu Y, Li G, Guo J, Bi Y, Qu Y, Wang W, Wang X, Guo Z, Yang R, Han Y.** 2009. Involvement of the post-transcriptional regulator Hfq in *Yersinia pestis* virulence. PLoS One 4:e6213. http://dx.doi.org/10.1371/journal.pone.0006213.
21. **Attia AS, Sedillo JL, Wang W, Liu W, Brautigam CA, Winkler W, Hansen EJ.** 2008. *Moraxella catarrhalis* expresses an unusual Hfq protein. Infect Immun 76:2520–2530. http://dx.doi.org/10.1128/IAI.01652-07.
22. **Fantappiè L, Metruccio MME, Seib KL, Oriente F, Cartocci E, Ferlicca F, Giuliani MM, Scarlato V, Delany I.** 2009. The RNA chaperone Hfq is involved in stress response and virulence in *Neisseria meningitidis* and is a pleiotropic regulator of protein expression. Infect Immun 77:1842–1853. http://dx.doi.org/10.1128/IAI.01216-08.
23. **Ding Y, Davis BM, Waldor MK.** 2004. Hfq is essential for *Vibrio cholerae* virulence and downregulates σE expression. Mol Microbiol 53:345–354. http://dx.doi.org/10.1111/j.1365-2958.2004.04142.x.
24. **Christiansen JK, Larsen MH, Ingmer H, Søgaard-Andersen L, Kallipolitis BH.** 2004. The RNA-binding protein Hfq of *Listeria monocytogenes*: role in stress tolerance and virulence. J Bacteriol 186:3355–3362. http://dx.doi.org/10.1128/JB.186.11.3355-3362.2004.
25. **Boyce JD, Harper M, Wilkie IW, Adler B.** 2010. *Pasteurella*, p 325–346. *In* Gyles CL, Prescott JF, Songer JG, Thoen CO (ed), Pathogenesis of bacterial infections of animals, 4th ed. Blackwell Publishing, Ames, IA.
26. **Blackall PJ, Miflin JK.** 2000. Identification and typing of *Pasteurella*

multocida: a review. Avian Pathol 29:271–287. http://dx.doi.org/10.1080/03079450050118395.
27. **Wilson BA, Ho M.** 2013. *Pasteurella multocida*: from zoonosis to cellular microbiology. Clin Microbiol Rev 26:631–655. http://dx.doi.org/10.1128/CMR.00024-13.
28. **Carter GR.** 1952. The type specific capsular antigen of *Pasteurella multocida*. Can J Med Sci 30:48–53.
29. **Harper M, John M, Turni C, Edmunds M, St Michael F, Adler B, Blackall PJ, Cox AD, Boyce JD.** 2015. Development of a rapid multiplex PCR assay to genotype *Pasteurella multocida* strains by use of the lipopolysaccharide outer core biosynthesis locus. J Clin Microbiol 53:477–485. http://dx.doi.org/10.1128/JCM.02824-14.
30. **Harper M, Boyce JD, Adler B.** 2012. The key surface components of *Pasteurella multocida*: capsule and lipopolysaccharide. Curr Top Microbiol Immunol 361:39–51. http://dx.doi.org/10.1007/82_2012_202.
31. **Tatum FM, Yersin AG, Briggs RE.** 2005. Construction and virulence of a *Pasteurella multocida fhaB2* mutant in turkeys. Microb Pathog 39:9–17. http://dx.doi.org/10.1016/j.micpath.2005.05.003.
32. **Bosch M, Garrido E, Llagostera M, Pérez De Rozas AM, Badiola I, Barbé J.** 2002. Pasteurella multocida exbB, exbD and tonB genes are physically linked but independently transcribed. FEMS Microbiol Lett 210:201–208. http://dx.doi.org/10.1111/j.1574-6968.2002.tb11181.x.
33. **Vimr ER.** 2013. Unified theory of bacterial sialometabolism: how and why bacteria metabolize host sialic acids. ISRN Microbiol 2013:816713. http://dx.doi.org/10.1155/2013/816713.
34. **Chung JY, Wilkie I, Boyce JD, Townsend KM, Frost AJ, Ghoddusi M, Adler B.** 2001. Role of capsule in the pathogenesis of fowl cholera caused by *Pasteurella multocida* serogroup A. Infect Immun 69:2487–2492. http://dx.doi.org/10.1128/IAI.69.4.2487-2492.2001.
35. **Boyce JD, Adler B.** 2000. The capsule is a virulence determinant in the pathogenesis of *Pasteurella multocida* M1404 (B:2). Infect Immun 68:3463–3468. http://dx.doi.org/10.1128/IAI.68.6.3463-3468.2000.
36. **Steen JA, Steen JA, Harrison P, Seemann T, Wilkie I, Harper M, Adler B, Boyce JD.** 2010. Fis is essential for capsule production in *Pasteurella multocida* and regulates expression of other important virulence factors. PLoS Pathog 6:e1000750. http://dx.doi.org/10.1371/journal.ppat.1000750.
37. **National Medical Health and Research Council.** 2004. Australian code of practice for the care and use of animals for scientific purposes, 7th ed. National Medical Health and Research Council, Canberra, Australia.
38. **Harper M, St Michael F, John M, Vinogradov E, Steen JA, van Dorsten L, Steen JA, Turni C, Blackall PJ, Adler B, Cox AD, Boyce JD.** 2013. *Pasteurella multocida* Heddleston serovar 3 and 4 strains share a common lipopolysaccharide biosynthesis locus but display both inter- and intrastrain lipopolysaccharide heterogeneity. J Bacteriol 195:4854–4864. http://dx.doi.org/10.1128/JB.00779-13.
39. **Harper M, Boyce JD, Wilkie IW, Adler B.** 2003. Signature-tagged mutagenesis of *Pasteurella multocida* identifies mutants displaying differential virulence characteristics in mice and chickens. Infect Immun 71:5440–5446. http://dx.doi.org/10.1128/IAI.71.9.5440-5446.2003.
40. **Henry R, Crane B, Powell D, Deveson Lucas D, Li Z, Aranda J, Harrison P, Nation RL, Adler B, Harper M, Boyce JD, Li J.** 2015. The transcriptomic response of *Acinetobacter baumannii* to colistin and doripenem alone and in combination in an *in vitro* pharmacokinetics/pharmacodynamics model. J Antimicrob Chemother 70:1303–1313. http://dx.doi.org/10.1093/jac/dku536.
41. **May BJ, Zhang Q, Li LL, Paustian ML, Whittam TS, Kapur V.** 2001. Complete genomic sequence of *Pasteurella multocida*, Pm70. Proc Natl Acad Sci U S A 98:3460–3465. http://dx.doi.org/10.1073/pnas.051634598.
42. **Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Subhraveti P, Weaver DS, Weerasinghe D, Zhang P, Karp PD.** 2014. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 42:D459–D471. http://dx.doi.org/10.1093/nar/gkt1103.
43. **Bradford MM.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–254. http://dx.doi.org/10.1016/0003-2697(76)90527-3.
44. **Wiśniewski JR, Zougman A, Nagaraj N, Mann M.** 2009. Universal sample preparation method for proteome analysis. Nat Methods 6:359–362. http://dx.doi.org/10.1038/nmeth.1322.
45. **Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJR.** 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc 4:484–494. http://dx.doi.org/10.1038/nprot.2009.21.
46. **Cox J, Mann M.** 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26:1367–1372. http://dx.doi.org/10.1038/nbt.1511.
47. **Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK.** 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47. http://dx.doi.org/10.1093/nar/gkv007.
48. **Harper M, Cox A, St Michael F, Parnas H, Wilkie I, Blackall PJ, Adler B, Boyce JD.** 2007. Decoration of *Pasteurella multocida* lipopolysaccharide with phosphocholine is important for virulence. J Bacteriol 189:7384–7391. http://dx.doi.org/10.1128/JB.00948-07.
49. **Boudry P, Gracia C, Monot M, Caillet J, Saujet L, Hajnsdorf E, Dupuy B, Martin-Verstraete I, Soutourina O.** 2014. Pleiotropic role of the RNA chaperone protein Hfq in the human pathogen *Clostridium difficile*. J Bacteriol 196:3234–3248. http://dx.doi.org/10.1128/JB.01923-14.
50. **Ansong C, Yoon H, Porwollik S, Mottaz-Brewer H, Petritis BO, Jaitly N, Adkins JN, McClelland M, Heffron F, Smith RD.** 2009. Global systems-level analysis of Hfq and SmpB deletion mutants in *Salmonella*: implications for virulence and global protein translation. PLoS One 4:e4809. http://dx.doi.org/10.1371/journal.pone.0004809.
51. **Ramseier TM, Bledig S, Michotey V, Feghali R, Saier MH, Jr.** 1995. The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. Mol Microbiol 16:1157–1169. http://dx.doi.org/10.1111/j.1365-2958.1995.tb02339.x.
52. **Douangamath A, Walker M, Beismann-Driemeyer S, Vega-Fernandez MC, Sterner R, Wilmanns M.** 2002. Structural evidence for ammonia tunneling across the (βα)8 barrel of the imidazole glycerol phosphate synthase bienzyme complex. Structure 10:185–193. http://dx.doi.org/10.1016/S0969-2126(02)00702-5.
53. **Cui M, Wang T, Xu J, Ke Y, Du X, Yuan X, Wang Z, Gong C, Zhuang Y, Lei S, Su X, Wang X, Huang L, Zhong Z, Peng G, Yuan J, Chen Z, Wang Y.** 2013. Impact of Hfq on global gene expression and intracellular survival in *Brucella melitensis*. PLoS One 8:e71933. http://dx.doi.org/10.1371/journal.pone.0071933.
54. **McNealy TL, Forsbach-Birk V, Shi C, Marre R.** 2005. The Hfq homolog in *Legionella pneumophila* demonstrates regulation by LetA and RpoS and interacts with the global regulator CsrA. J Bacteriol 187:1527–1532. http://dx.doi.org/10.1128/JB.187.4.1527-1532.2005.
55. **Sittka A, Pfeiffer V, Tedin K, Vogel J.** 2007. The RNA chaperone Hfq is essential for the virulence of *Salmonella* Typhimurium. Mol Microbiol 63:193–217. http://dx.doi.org/10.1111/j.1365-2958.2006.05489.x.
56. **Bohn C, Rigoulay C, Bouloc P.** 2007. No detectable effect of RNA-binding protein Hfq absence in *Staphylococcus aureus*. BMC Microbiol 7:10. http://dx.doi.org/10.1186/1471-2180-7-10.
57. **Hempel RJ, Morton DJ, Seale TW, Whitby PW, Stull TL.** 2013. The role of the RNA chaperone Hfq in *Haemophilus influenzae* pathogenesis. BMC Microbiol 13:134. http://dx.doi.org/10.1186/1471-2180-13-134.
58. **Chiang MK, Lu MC, Liu LC, Lin CT, Lai YC.** 2011. Impact of Hfq on global gene expression and virulence in *Klebsiella pneumoniae*. PLoS One 6:e22248. http://dx.doi.org/10.1371/journal.pone.0022248.
59. **Meibom KL, Forslund AL, Kuoppa K, Alkhuder K, Dubail I, Dupuis M, Forsberg Å, Charbit A.** 2009. Hfq, a novel pleiotropic regulator of virulence-associated genes in *Francisella tularensis*. Infect Immun 77:1866–1880. http://dx.doi.org/10.1128/IAI.01496-08.
60. **Subashchandrabose S, Leveque RM, Kirkwood RN, Kiupel M, Mulks MH.** 2013. The RNA chaperone *hfq* promotes fitness of *Actinobacillus pleuropneumoniae* during porcine pleuropneumonia. Infect Immun 81:2952–2961. http://dx.doi.org/10.1128/IAI.00392-13.
61. **Fuller TE, Kennedy MJ, Lowery DE.** 2000. Identification of *Pasteurella multocida* virulence genes in a septicemic mouse model using signature-tagged mutagenesis. Microb Pathog 29:25–38. http://dx.doi.org/10.1006/mpat.2000.0365.
62. **Boyce JD, Harper M, St Michael F, John M, Aubry A, Parnas H, Logan SM, Wilkie IW, Ford M, Cox AD, Adler B.** 2009. Identification of novel glycosyltransferases required for assembly of the *Pasteurella multocida* A:1 lipopolysaccharide and their involvement in virulence. Infect Immun 77:1532–1542. http://dx.doi.org/10.1128/IAI.01144-08.
63. **Harper M, Cox AD, Adler B, Boyce JD.** 2011. *Pasteurella multocida*

lipopolysaccharide: the long and the short of it. Vet Microbiol 153:109–115. http://dx.doi.org/10.1016/j.vetmic.2011.05.022.

64. **Cole J.** 1996. Nitrate reduction to ammonia by enteric bacteria: redundancy, or a strategy for survival during oxygen starvation? FEMS Microbiol Lett 136:1–11. http://dx.doi.org/10.1111/j.1574-6968.1996.tb08017.x.
65. **Boyce JD, Wilkie I, Harper M, Paustian ML, Kapur V, Adler B.** 2002. Genomic scale analysis of *Pasteurella multocida* gene expression during growth within the natural chicken host. Infect Immun 70:6871–6879. http://dx.doi.org/10.1128/IAI.70.12.6871-6879.2002.
66. **Wilkie IW, Grimes SE, O'Boyle D, Frost AJ.** 2000. The virulence and protective efficacy for chickens of *Pasteurella multocida* administered by different routes. Vet Microbiol 72:57–68. http://dx.doi.org/10.1016/S0378-1135(99)00187-X.

Appendix B: Supplementary tables and figures (Chapter 2)

From publication: *The RNA-binding chaperone Hfq is an important global regulator of gene expression in Pasteurella multocida and plays a crucial role in production of a number of virulence factors, including hyaluronic acid capsule (Mégroz et al., 2016; Infection and Immunity 84: 1361-1370).*

Table S1. Genes showing increased RNA expression during mid-exponential growth phase in the *P. multocida* VP161 *hfq* mutant strain in comparison to expression in the VP161 parent strain.

VP161 locus tag	Predicted function/product	Pm70 locus tag	PM70 gene name	Over-represented groups (GO categories or pathways)	Expression ratio (log_2)	FDR
PMVP_0016	6-phosphofructokinase, PkfA	PM0069	*pfkA*		1.34	1.63E-02
PMVP_0099	Cell division septal protein, FtsQ	PM0145	*ftsQ*		1.13	4.12E-03
PMVP_0130	ABC Transporter substrate binding protein - in MCE superfamily	PM0176	*PM0176*		1.28	1.02E-03
PMVP_0131	Ttg2 superfamily protein. Auxiliary component for an ABC-type transport system involved in resistance to organic solvents	PM0177	*PM0177*		1.49	1.23E-03
PMVP_0132	STAS superfamily protein-predicted NTP binding protein/anti-anti-sigma factor	PM0178	*PM0178*		1.29	1.23E-03
PMVP_0133	BolA superfamily protein-role in general stress response	PM0179	*PM0179*		1.25	1.58E-03
PMVP_0134	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	PM0180	*murZ*		1.06	3.52E-03
PMVP_0162	Preprotein translocase subunit, SecG	PM0208	*secG*	GO:0071702; Organic substance transport	1.11	7.48E-03
PMVP_0194	Periplasmic dipeptide transporter, DppA	PM0236	*dppA*		1.17	8.47E-03
PMVP_0201	Peptidase family protein with M23/M37 domain	PM0243	*PM0243*		1.54	2.41E-03

PMVP_0219	Spermidine/putrescine ABC transporter, periplasmic-binding protein, PotD	PM0261	*potD_2*	GO:0071702,/GO:0071705 Organic substance/ Nitrogen compound transport	2.56	1.39E-03
PMVP_0223	Peptidase T, PepT	PM0265	*pepT*		1.30	1.05E-02
PMVP_0230 and _0231	Peptidase family M15_3 protein	PM0271	*PM0271*		1.49	9.74E-04
PMVP_0232	Metallo-beta-lactamase superfamily, putative	PM0272	*PM0272*		1.24	7.97E-03
PMVP_0253	Predicted membrane-bound lysozyme-inhibitor of c-type				2.29	2.16E-03
PMVP_0292	Transglycosylase SLT domain protein	PM0326	*PM0326*		2.10	1.93E-03
PMVP_0297	OmpW	PM0331	*ompW*		1.06	1.25E-02
PMVP_0304	Haemoglobin-binding protein with TonB receptor	PM0337	*PM0337*		1.15	3.45E-02
PMVP_0341	Saccharopine dehydrogenase	PM0372	*PM0372*		1.40	6.23E-03
PMVP_0349	Sodium-dependent transporter, putative	PM0380	*PM0380*		1.34	2.18E-03
PMVP_0350	Nth endonuclease	PM0381	*nth*		1.33	4.65E-03
PMVP_0359	OmpH_2	PM0389	*ompH_2*		1.58	4.69E-03
PMVP_0444	Heat shock protein, HtpX	PM0468	*htpX*		1.34	1.02E-03
PMVP_0449	TRAP transporter/C4-dicarboxylate ABC transporter	PM0473	*PM0473*		1.19	2.41E-02
PMVP_0450	Universal stress protein, UspG	PM0474	*PM0474*		1.20	3.12E-02
PMVP_0561	Lipoprotein, Plp4	PM0586	*plp4*		1.00	7.71E-03
PMVP_0611	50S ribosomal protein Rpl25	PM0639	*rpL25*		1.13	8.53E-03
PMVP_0613	Dithiobiotin synthetise	PM0641	*bioD1*		1.17	1.02E-02
PMVP_0677	Amidophosphoribosyltransferase	PM0701	*purF*		1.04	4.69E-03
PMVP_0678	Colicin V production protein	PM0702	*cvpA*		1.35	2.35E-03
PMVP_0693	Hsf- putative adhesin protein	PM0714	*hsf*		1.26	1.90E-02
PMVP_0697	Sodium-dependent transporter, putative	PM0718	*PM0718*		1.15	4.68E-02
PMVP_0713	HtrA Serine peptidase	PM0734	*htrA*		2.18	3.41E-03

PMVP_0715	DnaK translation initiation factor IF-2	PM0736	*DnaK*		1.00	1.63E-02
PMVP_0740	Translation initiation factor IF-2	PM0759	*infB*		1.01	4.51E-03
PMVP_0742	Predicted ribosome maturation factor RimP	PM0761	*PM0761*		1.04	2.02E-02
PMVP_0777	Membrane protein belonging to TLP18.3, Psb32 and MOLO-1 founding proteins of phosphatase; This family has Rossmann-like folds.	PM0784	*PM0784*		1.15	2.21E-03
PMVP_0778	LemA-like family-amino terminal has predicted transmembrane helix	PM0785	*PM0785*		1.02	3.17E-03
PMVP_0800	TonB dependent receptor C-terminal region subfamily	PM0803	*PM0803*		1.30	1.98E-03
PMVP_0827	Amino-acid acetyltransferase	PM0828	*argA*		1.10	2.62E-02
PMVP_0835	Conserved protein of unknown function	PM0836	*PM0836*		1.16	4.62E-02
PMVP_0855	Fimbrial-related protein within Flp/Tight adhesion locus, homology to Flp-1	PM0855	*PM0855*		2.10	1.98E-03
PMVP_0865	PcnB poly(A) polymerase	PM0864	*pcnB*		1.17	2.16E-03
PMVP_0942	Conserved protein of unknown function	PM0931	*PM0931*		1.13	2.10E-02
PMVP_0958	Conserved protein of unknown function	PM0946	*PM0946*		1.71	1.30E-03
PMVP_0960	Sulfate transporter family protein	PM0948	*PM0948*		1.07	2.91E-02
PMVP_0969	AfuB_1 iron(III) ABC transporter, permease protein	PM0956	*afuB_1*		1.03	1.02E-03
PMVP_0981	Cell envelope integrity inner membrane protein TolA	PM0968	*tolA*		1.09	6.25E-03
PMVP_0982	Colicin uptake protein TolR	PM0969	*tolR*	GO:0071702; Organic substance transport	1.36	3.17E-03
PMVP_0983	TolQ Integral membrane protein belonging to MotA/TolQ/ExbB proton channel family	PM0970	*tolQ*	GO:0071702; Organic substance transport	1.57	1.23E-03
PMVP_0991	Crossover junction endodeoxyribonuclease RuvC	PM0978	*ruvC*		1.86	1.23E-03
PMVP_0992	Predicted membrane/periplasmic protein of unknown function	PM0979	*PM0979*		1.62	2.35E-03
PMVP_0993	Predicted transcriptional regulator PFAM01709 (negatively regulates the quorum-sensing response regulator by binding to its promoter region)	PM0980	*PM0980*		1.16	5.98E-03

PMVP_1057	Transferase predicted to add phosphoethanolamine to Lipid A	PM1042	*PM1042*		1.94	1.31E-03
PMVP_1089	Long-chain fatty acid outer membrane transporter PM1069	PM1069	*PM1069*		1.25	6.33E-03
PMVP_1156	Phosphocholine (PCho) transferase *pcgD*, required for the addition of PCho residues to lipopolysaccharide.	None			1.58	5.85E-04
PMVP_1157	Phosphocholine biosynthsesis protein pcgB	None			1.38	2.25E-03
PMVP_1180	Ribonuclease P protein component, RpnA	PM1163	*rnpA*		1.08	7.42E-03
PMVP_1181	Unknown protein with haemolytic domain. Predicted size 9.5kDa	PM1164	*PM1164*		1.11	2.36E-03
PMVP_1204	Biopolymer transport protein ExbB	PM1186	*exbB*	GO:0071702; Organic substance transport	1.03	4.75E-02
PMVP_1230	Survival protein SurA precursor	PM1208	*surA*	GO:0071702; Organic substance transport	1.44	2.18E-03
PMVP_1233	Transglutaminase-like superfamily domain protein	PM1211	*PM1211*		1.06	1.78E-02
PMVP_1281	Predicted Inner membrane protein	PM1258	*PM1258*		1.15	2.64E-03
PMVP_1285	Protein of unknown function	None			1.51	2.99E-02
PMVP_1286	Hydroxyethylthiazole kinase ThiM	PM1262	*thiM*		1.42	4.22E-02
PMVP_1287	Thiamine biosynthesis protein, putative	PM1263	*PM1263*		1.83	9.53E-03
PMVP_1319	30S ribosomal protein S16, RpsP	PM1294a	*rpsP*		1.25	9.45E-03
PMVP_1320	16s rRNA processing protein, RimM	PM1296	*rimM*		1.25	9.65E-03
PMVP_1321	tRNA (guanine-N1)-methyltransferase, TrmD	PM1297	*trmD*		1.12	1.77E-02
PMVP_1322	50S ribosomal protein RpL19	PM1298	*rpL19*		1.19	9.46E-03
PMVP_1575	Outer membrane lipoprotein PlpP	PM1518	*plpP*		1.33	2.44E-03
PMVP_1618	Predicted periplasmic/secreted protein	PM1566	*PM1566*		2.42	7.72E-04
PMVP_1638	RNA polymerase factor sigma-32, RpoH	PM1584	*rpoH*	GO:0000988; Protein binding transcription factors	1.53	2.64E-03
PMVP_1647	Ferredoxin-type protein NapF	PM1592	*napF*		3.54	1.74E-02

PMVP_1648	NapD protein	PM1593	*napD*		3.46	1.57E-02
PMVP_1649	Periplasmic nitrate reductase precursor, NapA	PM1594	*napA*		3.26	2.95E-03
PMVP_1650	Quinol dehydrogenase periplasmic component, NapG	PM1595	*napG*		3.11	4.12E-03
PMVP_1651	Quinol dehydrogenase membrane component, NapH	PM1596	*napH*		3.01	1.23E-03
PMVP_1652	Periplasmic nitrate reductase, NapB	PM1597	*napB*		2.88	7.89E-04
PMVP_1653	Cytochrome C-type protein, NapC	PM1598	*napC*		2.89	5.63E-04
PMVP_1670	Hypothetical protein PM1614	PM1614	*PM1614*		1.07	7.42E-03
PMVP_1678	NlpD-like lipoprotein with lysine motif	PM1623	*PM1623*		1.30	1.23E-03
PMVP_1706	Transglycosylase associated protein, predicted integral membrane protein.	None			2.29	1.02E-03
PMVP_1757	Protein of unknown function	None			1.49	2.38E-02
PMVP_1759	ATP-dependent Clp protease, ATPase subunit, ClpB	PM1704	*clpB*		1.12	3.25E-03
PMVP_1787	Outer membrane lipoprotein 2 precursor, PlpB	PM1730	*plpB*	GO:0071702 /GO:0071705 Organic substance /Nitrogen compound transport	1.15	1.94E-02
PMVP_1796	50S ribosomal protein L7/L12, Rpl12	PM1738	*rpL12*		1.16	1.13E-02
PMVP_1797	50S ribosomal protein L10, Rpl10	PM1739	*rpL10*		1.08	1.57E-02
PMVP_1836	Sigma-e factor regulatory protein RseC	PM1786	*PM1786*		1.31	1.54E-03
PMVP_1837	Periplasmic negative regulator of sigmaE, RseB	PM1787	*rseB*		1.72	1.23E-03
PMVP_1838	Sigma-E factor negative regulatory protein	PM1788	*mclA*	GO:0000988; Protein binding transcription factors	1.85	1.79E-03
PMVP_1839	RNA polymerase sigma factor, RpoE	PM1789	*rpoE*	GO:0000988; Protein binding transcription factors	1.61	2.53E-03
PMVP_1842	Nitrate/trimethylamine N-oxide reductase NapE/TorE				1.52	1.30E-02
PMVP_1925	Branched-chain amino acid transport system carrier protein, BrnQ	PM1873	*brnQ*	GO:0071702 /GO:0071705 Organic substance /Nitrogen compound transport	1.34	7.64E-04
PMVP_1972	Transcription termination factor Rho	PM1920	*rho*		1.29	3.26E-03
PMVP_1978	RlpA lipoprotein	PM1926	*PM1926*		1.65	7.72E-04

PMVP_1979	Penicillin-binding protein 5, DacA	PM1927	*dacA*		1.46	1.02E-03
PMVP_2013	Trigger factor, Tig	PM1975	*tig*	GO:0071702; Organic substance transport	1.15	7.90E-03
PMVP_2031	Skp export factor homologue	PM1993	*skp*		1.18	1.36E-02
PMVP_2032	FirA UDP-3-O-(R-3-hydroxymyristoyl)-glucosamine N-acyltransferase (lipid biosynthesis protein)	PM1994	*firA*		1.13	6.92E-03
PMVP_2058	Cytochrome c biogenesis protein, CcmA	PM0005	*ccmA*	GO:0017004; Cytochrome complex assembly	1.77	1.63E-02
PMVP_2059	Heme exporter protein B, CcmB	PM0006	*ccmB*	GO:0017004; Cytochrome complex assembly GO:0071702/ GO:0071705 Organic substance/Nitrogen compound transport	2.03	9.45E-03
PMVP_2060	Heme exporter protein C, CcmC	PM0007	*ccmC*	GO:0017004; Cytochrome complex assembly GO:0071702/ GO:0071705 Organic substance/Nitrogen compound transport	1.89	4.24E-03
PMVP_2061	Heme exporter protein D-related protein, CcmD	PM0008	*ccmD*	GO:0017004; Cytochrome complex assembly GO:0071702/ GO:0071705 Organic substance/Nitrogen compound transport	1.57	2.02E-02
PMVP_2062	Cytochrome c-type biogenesis protein, CcmE	PM0009	*ccmE*	GO:0017004; Cytochrome complex assembly	1.40	7.57E-03
PMVP_2063	C-type cytochrome biogenesis protein, CcmF	PM0010	*ccmF*	GO:0017004; Cytochrome complex assembly GO:0071702/ GO:0071705 Organic substance/Nitrogen compound transport	1.03	2.62E-03
PMVP_2077	Nitrite reductase, cytochrome C-type protein, NrfB	PM0024	*nrfB*		1.68	3.46E-02
PMVP_2078	Nitrite reductase, Fe-S protein, NrfC	PM0025	*nrfC*		1.65	2.25E-02
PMVP_2079	Nitrite reductase, transmembrane protein, NrfD	PM0026	*nrfD*		1.67	1.22E-03

Table S2. Genes showing decreased RNA expression during mid-exponential growth phase in the *P. multocida* VP161 *hfq* mutant strain in comparison to expression in the VP161 parent strain.

VP161 locus tag	Predicted function/product	Pm70 locus tag	PM70 gene name	Over-represented groups (GO categories or pathways)	Expression ratio (log_2)	FDR
PMVP_0001	LspB	PM0058	*lspB_2*		-1.19	1.05E-02
PMVP_0574	Predicted small protein (< 8 kDa). Partial match with typically large Yad A-like proteins, which are surface-exposed bacterial (adhesion) proteins.	None			-1.39	3.38E-03
PMVP_0765	Capsule biosynthesis, PhyB	PM0772	*phyB*	GO:0015774; Polysaccharide transport HA capsule biosynthesis	-1.35	1.02E-03
PMVP_0766	Capsule biosynthesis, PhyA	PM0773	*phyA*	GO:0015774; Polysaccharide transport HA capsule biosynthesis	-1.86	7.35E-04
PMVP_0767	Capsule biosynthesis, HyaE	PM0774	*hyaE*	HA capsule biosynthesis	-1.78	2.16E-03
PMVP_0768	Capsule biosynthesis, HyaD	PM0775	*PM0775*	HA capsule biosynthesis	-1.84	2.08E-03
PMVP_0769	Capsule biosynthesis, HyaC	PM0776	*PM0776*	HA capsule biosynthesis	-1.41	7.48E-03
PMVP_0869	DNA-binding transcriptional regulator. Fructose repressor, FruR	PM0868	*fruR*		-1.43	4.08E-03
PMVP_0910	HflX GTP-binding protein, *hflX* gene located adjacent to *hfq* gene. In *E.coli* predicted to interact with the 50S subunit,	PM0907	*hflX*		-1.69	9.74E-04
PMVP_1254	Pyridoxine biosynthesis protein	PM1232	*PM1232*	GO:0042822; Pyridoxal phosphate metabolism	-2.53	4.69E-03
PMVP_1255	Imidazole glycerol phosphate synthase, subunit H	PM1233	*PM1233*	GO:0042822; Pyridoxal phosphate metabolism	-2.50	6.25E-03
PMVP_1267	Putative L-xylulose 5-phosphate 3-epimerase	PM1245	*PM1245*		-1.21	2.75E-03
PMVP_1268	Hexulose-6-phosphate synthase SgbH, putative	PM1246	*PM1246*		-1.14	2.28E-03
PMVP_1310	Nucleoside triphosphate pyrophosphohydrolase	PM1285	*mazG*		-1.01	2.18E-03

PMVP_1480	Shikimate 5-dehydrogenase	PM1429	*PM1429*	-1.00	2.16E-03
PMVP_1529	Gram-positive signal peptide, YSIRK family domain protein	PM1477	*PM1477*	-1.24	2.65E-02
PMVP_1624	Uridine phosphorylase	PM1572	*udp*	-1.44	4.69E-03
PMVP_1625	DNA-binding transcriptional regulator AsnC	PM1573	*asnC*	-1.31	2.38E-03
PMVP_2053	Small hypothetical protein	None		-1.00	1.02E-02

Table S3. Proteins showing increased production in the *P. multocida* VP161 *hfq* mutant strain in comparison to the parent strain, VP161, during early-exponential growth phase as measured by high throughput liquid proteomics.

VP161 locus tag	Predicted function/product	Pm70 locus tag	PM70 protein name	Expression ratio (log_2)	FDR
PMVP_0131	Ttg2 superfamily protein. Auxiliary component for an ABC-type transport system involved in resistance to organic solvents	PM0177	PM0177	1.01	7.94E-04
PMVP_0194	Periplasmic dipeptide transporter, DppA	PM0236	DppA	2.59	2.82E-04
PMVP_0195	Dipeptide ABC transporter, permease protein, DppB	PM0237	DppB	2.55	6.90E-03
PMVP_0198	Dipeptide ABC transporter, ATP-binding protein, DppF	PM0240	DppF	2.21	4.28E-03
PMVP_0229	Murein L,D-transpeptidase; Provisional	PM0270	PM0270	1.57	4.66E-04
PMVP_0236	Citrate synthase, GltA	PM0276	GltA	1.30	8.84E-04
PMVP_0253	Predicted membrane-bound lysozyme-inhibitor of c-type			3.30	6.48E-04
PMVP_0265	Haemoglobin/haptoglobin binding protein	PM0300	PM0300	1.79	5.84E-04
PMVP_0267	ATPase, Mrp	PM0302	Mrp	1.18	5.19E-04
PMVP_0292	Transglycosylase SLT domain protein	PM0326	PM0326	1.56	1.69E-03
PMVP_0301	Proton glutamate symport protein, putative	PM0335	PM0335	1.22	2.83E-02
PMVP_0304	Haemoglobin-binding protein with TonB receptor	PM0337	PM0337	1.49	6.40E-03
and _0305				1.83	1.75E-03
PMVP_0327	MscS family protein potassium efflux protein, KefA	PM0358	PM0358	1.07	8.84E-04
PMVP_0349	Sodium-dependent transporter, putative	PM0380	PM0380	1.33	3.51E-03
PMVP_0376	Formate dehydrogenase accessory protein FdhE	PM0405	FdhE	1.29	1.16E-03
PMVP_0449	TRAP transporter/C4-dicarboxylate ABC transporter	PM0473	PM0473	1.58	5.48E-04
PMVP_0450	Universal stress protein, UspG	PM0474	PM0474	1.21	5.03E-04
PMVP_0457	Cold shock-like protein-related protein	PM0481	CspD	1.13	3.09E-02

PMVP_0487	Glutaredoxin 2, Grx2	PM0518	Grx2	1.11	7.20E-03
PMVP_0488	Conserved protein with no known function	PM0519	PM0519	2.10	4.44E-04
PMVP_0516	Glucose-1-phosphate adenylyltransferase, GlgC	PM0543	GlgC	1.46	5.48E-04
PMVP_0518	Carbohydrate phosphorylases subfamily	PM0545	GlgP	1.08	2.40E-03
PMVP_0523	Arginine repressor	PM0549	ArgR	1.06	3.34E-02
PMVP_0541	Branched-chain amino acid aminotransferase, IlvE	PM0566	IvlE	1.39	5.84E-04
PMVP_0561	Lipoprotein, Plp4	PM0586	Plp4	1.77	4.44E-04
PMVP_0629	TehB SAM-dependent methyltransferase/tellurite resistance protein, TehB	PM0656	TehB	3.50	8.84E-04
PMVP_0647	Hypothetical protein	PM0674	PM0674	1.67	4.44E-04
PMVP_0713	Serine peptidase, HtrA	PM0734	HtrA	2.38	2.82E-04
PMVP_0756	ABC-ATPase-MoxR-like			1.18	1.58E-02
PMVP_0795	Protein of unknown function	PM0798	PM0798	1.40	2.43E-02
PMVP_0837	Histidinol-phosphate aminotransferase, His_H1	PM0838	HisH_1	1.26	1.01E-03
PMVP_0842	Protein within Flp/Tight adhesion locus, homology to TadG	PM0843	TadG	1.12	7.94E-04
PMVP_0845	Protein within Flp/Tight adhesion locus, homology to TadD	PM0846	TadD	1.52	6.08E-03
PMVP_0849	Protein within Flp/Tight adhesion locus, homology to ATPase CpaE	PM0850	PM0850	1.07	1.05E-03
PMVP_0851	Protein within Flp/Tight adhesion locus, RcpA	PM0852	RcpA	1.39	5.16E-03
PMVP_0852	Protein within Flp/Tight adhesion locus, homology to RpcC	PM0853	PM0853	2.00	7.94E-04
PMVP_0870	Acetolactate synthase 3 regulatory subunit, IlvH	PM0869	IlvH	1.44	7.94E-04
PMVP_0871	Acetolactate synthase, large subunit, biosynthetic type, IlvI	PM0870	IlvI	1.76	1.18E-02
PMVP_0939	Membrane-bound lytic murein transglycosylase	PM0928	PM0928	1.66	7.94E-04
PMVP_0948	Aspartate kinase, monofunctional class	PM0937	PM0937	1.33	5.19E-04
PMVP_0958	Protein of unknown function	PM0946	PM0946	1.66	4.44E-04

PMVP_0992	Predicted membrane/periplasmic protein of unknown function	PM0979	PM0979	2.08	2.82E-04
PMVP_1008	Cystathionine gamma-synthase, MetB	PM0995	MetB	1.08	8.19E-04
PMVP_1069	Dihydrodipicolinate synthase, DapA	PM1051	DapA	1.11	7.94E-04
PMVP_1162	Formamidopyrimidine-DNA glycosylase, Fpg	PM1145	Fpg	1.22	2.83E-02
PMVP_1208	Type I deoxyribonuclease HsdR (HtpX Peptidase_M48 superfamily)	PM1190	PM1190	1.34	7.27E-03
PMVP_1220	Histidinol-phosphate aminotransferase	PM1199	HisC	1.03	6.28E-03
PMVP_1233	Transglutaminase-like superfamily domain protein	PM1211	PM1211	1.41	8.19E-04
PMVP_1239	Cellulose synthase catalytic subunit, putative	PM1217	PM1217	2.27	1.03E-02
PMVP_1533	Bacterial regulatory protein, arsR family, putative	PM1481	PM1481	1.10	5.63E-03
PMVP_1575	Outer membrane lipoprotein PlpP	PM1518	PlpP	1.68	7.39E-03
PMVP_1614	Glycosyltransferase GTA type/β-D-1,6 glucosyl transferase	PM1562	PM1562	2.46	4.44E-04
PMVP_1616	Large conductance mechanosensitive channel protein, MscL	PM1564	MscL	2.12	7.32E-03
PMVP_1618	Predicted periplasmic/secreted protein	PM1566	PM1566	2.30	4.44E-04
PMVP_1623	Conserved unknown protein TIGR01125	PM1571	PM1571	1.13	1.88E-03
PMVP_1649	Periplasmic nitrate reductase precursor, NapA	PM1594	NapA	1.59	8.84E-04
PMVP_1650	Quinol dehydrogenase periplasmic component, NapG	PM1595	NapG	1.49	1.51E-02
PMVP_1652	Periplasmic nitrate reductase, NapB	PM1597	NapB	1.12	2.91E-02
PMVP_1678	Protein of unknown function	PM1623	PM1623	1.45	5.03E-04
PMVP_1683	Acetolactate synthase, large subunit, biosynthetic type, IlvG	PM1628	IlvG	1.16	3.93E-03
PMVP_1687	Aspartate-semialdehyde dehydrogenase, Asd	PM1632	Asd	1.03	6.96E-04
PMVP_1759	ATP-dependent Clp protease, ATPase subunit, ClpB	PM1704	ClpB	1.42	4.66E-04
PMVP_1760	Thioredoxin, Trx	PM1705	Trx	1.87	4.88E-03
PMVP_1787	Outer membrane lipoprotein 2 precursor, PlpB	PM1730	PlpB	1.32	5.03E-04

PMVP_1832	Protein of unknown function	None		1.06	4.88E-03
PMVP_1837	Periplasmic negative regulator of sigmaE, RseB	PM1787	RseB	1.16	2.15E-03
PMVP_1838	Sigma-E factor negative regulatory protein, MclA	PM1788	MclA	1.06	4.28E-03
PMVP_1863	Nitrate/nitrite response regulator protein, NarP	PM1810	NarP	2.51	2.82E-04
PMVP_1864	Adenylate cyclase, CyaA	PM1811	CyaA	1.36	2.50E-02
PMVP_1872	SrfB homologue, uncharacterized protein conserved in bacteria	PM1819	PM1819	1.21	2.04E-03
PMVP_1874	Protein of unknown function	PM1821	PM1821	1.00	2.94E-03
PMVP_1879	Protein of unknown function	PM1826	PM1826	1.86	4.66E-04
PMVP_1880	Superfast myosin heavy chain, putative	PM1827	PM1827	1.64	2.96E-03
PMVP_1881	Protein of unknown function	PM1828	PM1828	1.44	3.31E-02
PMVP_1882	Protein of unknown function	PM1829	PM1829	1.03	2.14E-02
PMVP_1884	Protein of unknown function	None		1.16	1.85E-02
PMVP_1949	Protein of unknown function	PM1897	PM1897	1.15	3.01E-02
PMVP_1958	Oligopeptide ABC transporter, ATP-binding protein	PM1906	OppF	1.43	3.79E-02
PMVP_1959	Oligopeptide transport ATP-binding protein Oppd	PM1907	OppD	1.32	6.31E-03
PMVP_1962	Periplasmic oligopeptide-binding protein precursor, OppA	PM1910	OppA	1.26	5.19E-04
PMVP_1999	3-isopropylmalate dehydratase, large subunit, LeuC	PM1960	LeuC	1.15	4.45E-02
PMVP_2031	Export factor homologue, Skp	PM1993	Skp	1.09	1.72E-03
PMVP_2065	Conserved hypothetical protein	PM0012	CcmH_1	1.18	2.93E-02
PMVP_2095	Glutamate dehydrogenase	PM0043	GdhA	2.01	1.75E-03
PMVP_2109	LspB	PM0056	LspB_1	1.00	4.88E-03

Table S4. Proteins showing reduced production in the *P. multocida hfq* mutant strain in comparison to the wild-type VP161 strain during early-exponential growth phase as measured by high throughput liquid proteomics.

VP161 locus tag	Predicted function/product	Pm70 locus tag	PM70 protein name	Expression ratio (log_2)	FDR
PMVP_0001	Hemolysin activation/secretion protein, LspB	PM0058	LspB_2	-2.74	8.84E-04
PMVP_0002	Filamentous hemagglutinin, PfhB2	PM0059	PfhB2	-2.81	4.44E-04
and _0003	Filamentous hemagglutinin, PfhB2			-2.82	4.66E-04
and _0004	Filamentous hemagglutinin, PfhB2			-2.70	4.44E-04
and _0005	Filamentous hemagglutinin, PfhB2			-2.19	8.19E-04
and _0006	Filamentous hemagglutinin, PfhB2			-1.69	5.03E-04
PMVP_0218	Spermidine/putrescine ABC transporter, periplasmic-binding protein, PotD	PM0260	PotD_1	-1.88	6.96E-04
PMVP_0273	CRISPR-associated protein, Csy2 family	PM0306	PM0306	-1.34	2.55E-02
PMVP_0303	Protein with TonB-dependent receptor plug domain	None		-1.00	6.92E-03
PMVP_0767	Capsule biosynthesis, HyaE	PM0774	HyaE	-4.17	2.82E-04
PMVP_0768	Capsule biosynthesis, HyaD	PM0775	PM0775	-4.35	2.82E-04
PMVP_0769	Capsule biosynthesis, HyaC	PM0776	PM0776	-3.86	1.22E-03
PMVP_0771	Capsule transport, HexD	PM0778	HexD	-3.29	4.50E-04
PMVP_0772	Capsule transport, HexC	PM0779	HexC	-3.44	4.44E-04
PMVP_0869	DNA-binding transcriptional regulator. Fructose repressor, FruR	PM0868	FruR	-1.99	1.00E-02
PMVP_0909	RNA-binding protein, Hfq	PM0906	Hfq	-5.02	2.19E-03
PMVP_0950	Protein of unknown function	PM0939	PM0939	-1.23	3.93E-03
PMVP_1012	Protein of unknown function	PM0998	PM0998	-1.36	2.49E-03
PMVP_1051	Galactose-1-phosphate uridylyltransferase	PM1036	GalT	-1.16	1.72E-02
PMVP_1053	Galactose ABC transporter, periplasmic-binding protein, MglB	PM1038	MglB	-1.77	4.44E-04

PMVP_1142	N-acetyl-gamma-glutamyl-phosphate reductase, ArgC, ArgC	PM1118	ArgC	-1.44	5.38E-03
PMVP_1216	Histidinol dehydrogenase, HisD	PM1198	HisD	-1.66	1.90E-02
PMVP_1255	Imidazole glycerol phosphate synthase, subunit H	PM1233	PM1233	-1.50	2.47E-02
PMVP_1330	3-deoxy-D-manno-octulosonic-acid kinase	PM1303	PM1303	-1.21	2.84E-02
PMVP_1573	Outer membrane lipoprotein, PlpE	PM1517	PlpE	-1.95	5.37E-04

Table S5. Proteins showing increased production in the *P. multocida hfq* mutant strain in comparison to the wild-type VP161 strain during mid-exponential growth phase as measured by high throughput liquid proteomics.

VP161 locus tag	Predicted function/product	Pm70 locus tag	PM70 protein name	Expression ratio (log_2)	FDR
PMVP_0040	AlsT Sodium/Alanine Symporter	PM0091	PM0091	1.58	3.29E-02
PMVP_0160	Recombination protein RecR	PM0206	RecR	1.44	4.82E-02
PMVP_0194	Periplasmic dipeptide transporter, DppA	PM0236	DppA	2.44	1.06E-04
PMVP_0198	Dipeptide ABC transporter, ATP-binding protein, DppF	PM0240	DppF	1.71	1.39E-03
PMVP_0229	Murein L,D-transpeptidase; Provisional	PM0270	PM0270	1.73	4.44E-04
PMVP_0243	Adenylate kinase, Adk	PM0284	Adk	1.14	8.68E-03
PMVP_0253	Predicted membrane-bound lysozyme-inhibitor of c-type			2.06	2.53E-04
PMVP_0265	Haemoglobin/haptoglobin binding protein	PM0300	PM0300	1.28	2.32E-03
PMVP_0267	ATPase, Mrp	PM0302	Mrp	1.32	1.77E-03
PMVP_0292	Transglycosylase SLT domain protein	PM0326	PM0326	2.85	1.85E-03
PMVP_0304	Haemoglobin-binding protein with TonB receptor	PM0337	PM0337	1.49	9.02E-04
PMVP_0327	MscS family protein potassium efflux protein, KefA	PM0358	PM0358	1.00	7.98E-03
PMVP_0376	Formate dehydrogenase accessory protein, FdhE	PM0405	FdhE	1.00	3.62E-02
PMVP_0449	TRAP transporter/C4-dicarboxylate ABC transporter	PM0473	PM0473	1.78	8.79E-04
PMVP_0450	Universal stress protein, UspG	PM0474	PM0474	1.35	1.74E-03
PMVP_0488	Conserved protein with no known function	PM0519	PM0519	1.81	1.77E-03
PMVP_0516	Glucose-1-phosphate adenylyltransferase, GlgC	PM0543	GlgC	1.56	3.39E-02
PMVP_0541	Branched-chain amino acid aminotransferase, IlvE	PM0566	IlvE	1.10	5.18E-03
PMVP_0561	Lipoprotein, Plp4	PM0586	Plp4	1.31	1.85E-03
PMVP_0629	TehB SAM-dependent methyltransferase/tellurite resistance protein, TehB	PM0656	TehB	2.18	4.17E-03
PMVP_0638	Ferritin, RsgA_1	PM0666	RsgA_1	1.46	1.39E-03
PMVP_0639	Ferritin, RsgA_2	PM0667	RsgA_2	1.17	6.19E-03
PMVP_0756	ABC-ATPase-MoxR-like			1.47	9.02E-04

PMVP_0780	FAD-binding, FAD-linked oxidase family membrane protein	PM0787	PM0787	1.92	1.39E-03
PMVP_0811	Phosphoribosylaminoimidazolesuccinocarboxamide synthase, PurC	PM0815	PurC	1.99	6.28E-04
PMVP_0837	Histidinol-phosphate aminotransferase, HisH_1	PM0838	HisH_1	1.41	1.47E-03
PMVP_0852	Protein within Flp/Tight adhesion locus, homology to TadC	PM0853	PM0853	1.74	5.98E-04
PMVP_0939	Membrane-bound lytic murein transglycosylase	PM0928	PM0928	1.99	3.96E-04
PMVP_0948	Aspartate kinase, monofunctional class	PM0937	PM0937	1.21	2.57E-03
PMVP_0992	Predicted membrane/periplasmic protein of unknown function	PM0979	PM0979	2.05	4.04E-04
PMVP_1000	Glyoxylase I, GloA	PM0987	GloA	1.67	2.63E-02
PMVP_1069	Dihydrodipicolinate synthase, DapA	PM1051	DapA	1.01	8.68E-03
PMVP_1153	Cell division protein FtsN	PM1136	PM1136	2.08	9.40E-03
PMVP_1208	Type I deoxyribonuclease HsdR (HtpX Peptidase_M48 superfamily)	PM1190	PM1190	1.51	8.79E-04
PMVP_1233	Transglutaminase-like superfamily domain protein	PM1211	PM1211	1.98	2.08E-03
PMVP_1287	Thiamine biosynthesis protein, putative	PM1263	PM1263	1.03	1.98E-02
PMVP_1309	Ketol-acid reductoisomerase, IlvC	PM1284	IlvC	1.01	7.15E-03
PMVP_1614	Glycosyltransferase GTA type/β-D-1,6 glucosyl transferase	PM1562	PM1562	2.45	1.06E-04
PMVP_1616	Large conductance mechanosensitive channel protein, MscL	PM1564	MscL	2.55	2.42E-04
PMVP_1649	Periplasmic nitrate reductase precursor, NapA	PM1594	NapA	1.29	1.81E-03
PMVP_1652	Periplasmic nitrate reductase, NapB	PM1597	NapB	1.75	5.47E-03
PMVP_1760	Thioredoxin, Trx	PM1705	Trx	1.73	1.33E-02
PMVP_1787	Outer membrane lipoprotein 2 precursor, PlpB	PM1730	PlpB	1.09	9.91E-03
PMVP_1863	Nitrate/nitrite response regulator protein, NarP	PM1810	NarP	2.05	2.53E-04
PMVP_1872	SrfB homologue, uncharacterized protein conserved in bacteria.	PM1819	PM1819	1.21	7.88E-03
PMVP_1879	Hypothetical protein PM1826	PM1826	PM1826	1.58	8.79E-04
PMVP_1906	Conserved protein of unknown function	PM1853	PM1853	1.00	1.68E-02
PMVP_1962	Periplasmic oligopeptide-binding protein precursor, OppA	PM1910	OppA	1.17	3.72E-03
PMVP_2095	Glutamate dehydrogenase, GdhA	PM0043	GdhA	2.71	1.06E-04

Table S6. Proteins showing reduced production in the *P. multocida hfq* mutant strain in comparison to the wild-type VP161 strain during mid-exponential growth phase as measured by high throughput liquid proteomics.

VP161 locus tag	Predicted function/product	Pm70 locus tag	PM70 protein name	Expression ratio (log_2)	FDR
PMVP_0001	Hemolysin activation/secretion protein, LspB	PM0058	LspB_2	-1.83	1.85E-03
PMVP_0002	Filamentous hemagglutinin, PfhB2	PM0059	PfhB2	-1.85	1.92E-03
and_0003	Filamentous hemagglutinin, PfhB2			-1.82	2.21E-03
and_0004	Filamentous hemagglutinin, PfhB2			-1.77	3.58E-03
and_0005	Filamentous hemagglutinin, PfhB2			-1.81	1.14E-02
and_0006	Filamentous hemagglutinin, PfhB2			-1.86	1.85E-03
PMVP_0154	Fumarate reductase iron-sulfur subunit	PM0200	FrdB	-1.44	3.38E-03
PMVP_0155	Fumarate reductase, flavoprotein subunit	PM0201	FrdA	-1.36	1.07E-02
PMVP_0178	Bifunctional phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase	PM0222	PurH	-2.76	2.10E-03
PMVP_0218	Spermidine/putrescine ABC transporter, periplasmic-binding protein	PM0260	PotD_1	-1.58	8.79E-04
PMVP_0297	OmpW	PM0331	OmpW	-1.48	4.61E-02
PMVP_0506	Trimethylamine-N-oxide reductase	PM1793	TorA	-2.29	1.90E-02
PMVP_0519	Protein with periplasmic binding domain PBP2_Fbp_like_1	None		-2.08	5.49E-04
PMVP_0524	Malate dehydrogenase	PM0550	Mdh_2	-1.44	9.40E-03
PMVP_0744	Putative L-ascorbate 6-phosphate lactonase	PM0763	PM0763	-5.19	2.09E-04
PMVP_0767	Capsule biosynthesis, HyaE	PM0774	HyaE	-1.11	2.63E-02
PMVP_0768	Capsule biosynthesis, HyaD	PM0775	PM0775	-1.06	1.95E-02
PMVP_0769	Capsule biosynthesis, HyaC	PM0776	PM0776	-1.41	5.60E-03
PMVP_0771	Capsule biosynthesis, HexD	PM0778	HexD	-1.01	1.33E-02
PMVP_0788	TRAP transporter used for the uptake of organic acids	None		-1.71	2.42E-02
PMVP_0802	Ornithine decarboxylase	PM0806	SpeF	-3.09	8.79E-04

PMVP_0803	Putrescine-ornithine antiporter	PM0807	PotE	-2.58	3.98E-04
PMVP_0822	Fumarate hydratase, class II, FumC	PM0823	FumC	-1.97	9.02E-04
PMVP_0869	DNA-binding transcriptional regulator. Fructose repressor, FruR	PM0868	FruR	-1.31	4.81E-03
PMVP_1013	Putative nicotinate phosphoribosyltransferase	PM0999	PM0999	-1.16	2.47E-02
PMVP_1255	Imidazole glycerol phosphate synthase, subunit H	PM1233	PM1233	-1.50	2.47E-02
PMVP_1407	Glucuronate isomerise, UxaC	None		-3.97	5.19E-04
PMVP_1485	Anaerobic C4-dicarboxylate membrane transporter, DcuB	PM1434	DcuB	-1.50	3.99E-02
PMVP_1521	Phosphatidylethanolamine-binding protein, COG1881	PM1470	PM1470	-2.85	1.85E-03
PMVP_1594	Phosphoenolpyruvate carboxykinase, PckA	PM1542	PckA	-1.91	1.39E-03
PMVP_1746	Cysteine synthase, CysK	PM1693	CysK	-1.02	2.37E-02
PMVP_1965	Fatty acid/phospholipid biosynthesis enzyme, PlsX	PM1913	PlsX	-1.56	3.39E-02
PMVP_1970	Anaerobic C4-dicarboxylate transporter, DcuB1	PM1918	PM1918	-1.13	3.17E-02

Table S7. Genes/proteins showing increased expression in the *P. multocida hfq* mutant strain in comparison to the wild-type VP161 strain during mid-exponential growth phase as measured by RNA-Seq and high-throughput liquid proteomics analysis. The protein expression levels at early-exponential phase are included for comparison.

VP161 Feature ID	Predicted function/product	RNA-Seq (mid-exponential)		Proteomics (early-exponential)		Proteomics (mid-exponential)	
		Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0194	Periplasmic dipeptide transporter, DppA	1.17	8.47E-03	2.44	1.06E-04	2.59	2.82E-04
PMVP_0253	Predicted membrane-bound lysozyme-inhibitor of c-type	2.29	2.16E-03	2.06	2.53E-04	3.30	6.48E-04
PMVP_0292	Transglycosylase SLT domain protein	2.10	1.93E-03	2.85	1.85E-03	1.56	1.69E-03
PMVP_0305	TonB receptor (haemoglobin-binding protein)	1.15	3.45E-02	1.32	3.53E-03	1.83	1.75E-03
PMVP_0449	TRAP transporter/C4-dicarboxylate ABC transporter	1.19	2.41E-02	1.78	8.79E-04	1.58	5.48E-04
PMVP_0450	Universal stress protein, UspG	1.20	3.12E-02	1.35	1.74E-03	1.21	5.03E-04
PMVP_0561	Lipoprotein, Plp4	1.00	7.71E-03	1.31	1.85E-03	1.77	4.44E-04
PMVP_0992	Predicted membrane/periplasmic protein of unknown function	1.62	2.35E-03	2.05	4.04E-04	2.08	2.82E-04
PMVP_1233	Transglutaminase-like superfamily domain protein	1.06	1.78E-02	1.98	2.08E-03	1.41	8.19E-04
PMVP_1649	Periplasmic nitrate reductase precursor, NapA	3.26	2.95E-03	1.29	1.81E-03	1.59	8.84E-04
PMVP_1650	Quinol dehydrogenase periplasmic component, NapG	3.11	4.12E-03	1.42	5.07E-02	1.49	1.51E-02
PMVP_1652	Periplasmic nitrate reductase, NapB	2.88	7.89E-04	1.75	5.47E-03	1.12	2.91E-02
PMVP_1787	Outer membrane lipoprotein 2 precursor, PlpB	1.15	1.94E-02	1.09	9.91E-03	1.32	5.03E-04

Table S8. Genes/proteins showing reduced expression in the *P. multocida hfq* mutant strain in comparison to the wild-type VP161 strain during mid-exponential growth phase as measured by RNA-Seq and high-throughput liquid proteomics analysis. The protein expression levels at early-exponential phase are included for comparison.

VP161 Feature ID	Predicted function/product	RNA-Seq (mid-exponential)		Proteomics (early-exponential)		Proteomics (mid-exponential)	
		Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR	Expression ratio (log_2)	FDR
PMVP_0001	Filamentous hemagglutinin outer membrane transporter, LspB	-1.19	1.05E-02	-2.74	8.84E-04	-1.83	1.85E-03
PMVP_0767	Capsule biosynthesis, HyaE	-1.78	2.16E-03	-4.17	2.82E-04	-1.11	2.62E-02
PMVP_0768	Capsule biosynthesis, HyaD	-1.84	2.08E-03	-4.35	2.82E-04	-1.06	1.95E-02
PMVP_0769	Capsule biosynthesis, HyaC	-1.41	7.48E-03	-3.86	1.22E-03	-1.41	5.60E-03
PMVP_0869	DNA-binding transcriptional regulator. Fructose repressor, FruR	-1.43	4.08E-03	-1.99	1.00E-02	-1.31	4.81E-03
PMVP_1255	Imidazole glycerol phosphate synthase, subunit H	-2.50	6.25E-03	-0.28	1.35E-01	-1.50	2.47E-02

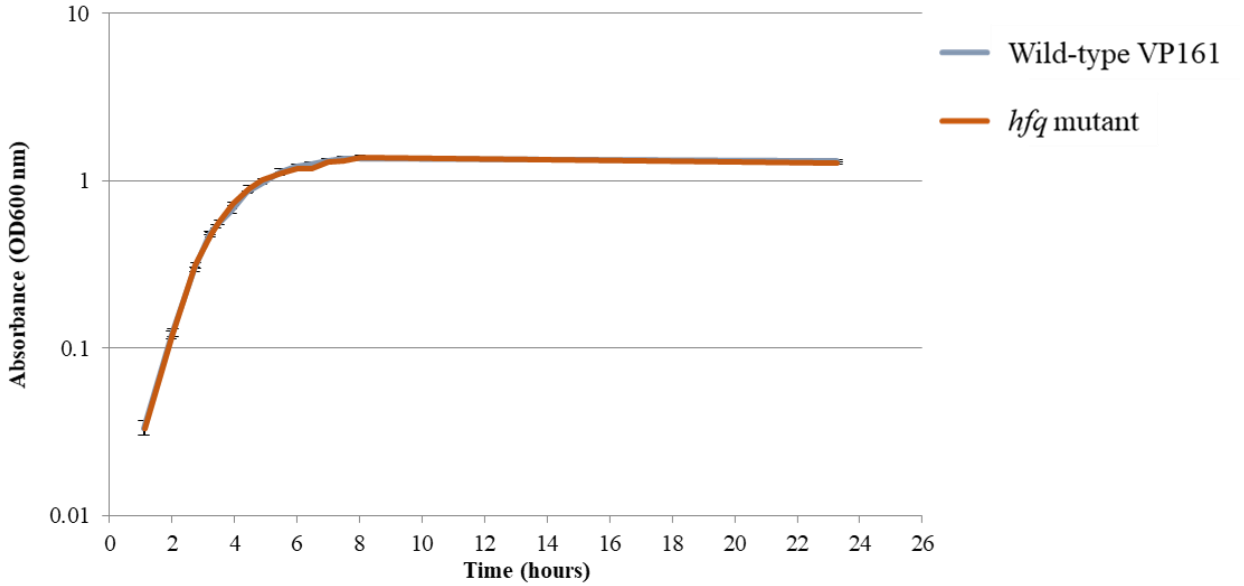


Figure S1. Growth rate of *P. multocida* wild-type VP161 and the *hfq* mutant in BHI broth at 37°C with constant shaking at 200 rpm. Optical density readings were taken for each culture at appropriate intervals over a 24 h time period. Data are presented as mean OD_{600} ± standard error of the mean (determined from three biological replicates).

Appendix C: Supplementary tables (Chapter 3)

Table S9. Genes in the wild-type *P. multocida* X-73 strain showing increased RNA expression during growth in low iron conditions, in comparison to gene expression during growth in rich medium.

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS09630	no known function	PM0127		1.26	2.21E-02
X73_RS09625	iron-dicitrate transporter ATP-binding subunit	PM0128	*fecE*	1.65	4.31E-04
X73_RS09620	iron-dicitrate transporter subunit FecD	PM0129	*fecD*	1.81	5.41E-06
X73_RS09615	iron-dicitrate transporter permease subunit	PM0130	*fecC*	1.95	5.55E-05
X73_RS09610	iron-dicitrate transporter substrate-binding subunit	PM0131	*fecB*	2.05	2.11E-09
X73_RS08825	L-lactate dehydrogenase LldD	PM0288	*lldD*	3.42	6.24E-10
X73_RS08720	no known function	PM0298		1.47	8.15E-04
X73_RS08715	no known function	PM0299		1.51	3.70E-04
X73_RS08710	no known function	PM0300		1.56	2.60E-07
X73_RS08710b	no known function			3.17	2.60E-07
X73_RS08540	Hemoglobin-haptoglobin utilization protein B	PM0337		2.23	9.25E-07
X73_RS08240	YfeD	PM0397	*yfeD*	1.01	9.28E-03
X73_RS08235	YfeC	PM0398	*yfeC*	1.53	2.00E-05
X73_RS08230	YfeB	PM0399	*yfeB*	1.86	6.95E-08
X73_RS08225	YfeA	PM0400	*yfeA*	1.91	1.42E-05
X73_RS08010	ResA	PM0447	*resA*	1.02	2.63E-02
X73_RS08005	Uncharacterized ABC transporter ATP-binding protein MJ1508	PM0448		1.22	5.09E-04
X73_RS08000	no known function	PM0449		1.38	9.22E-03
X73_RS07995	no known function	PM0450		1.64	1.40E-03
X73_RS07990	no known function	PM0451		1.70	7.89E-06
X73_RS07985	no known function	PM0452		2.02	3.71E-05

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS07980	no known function	PM0453		1.90	1.06E-07
X73_RS06910	no known function	PM0679		1.26	2.42E-06
X73_RS06615	no known function	PM0741		3.09	1.69E-08
X73_RS06280	Uncharacterized protein YddB	PM0803		5.17	2.07E-09
X73_RS06275	Probable zinc protease PqqL	PM0804	*pqqL*	3.09	5.34E-06
X73_RS06265	putrescine transporter	PM0807	*potE*	1.24	4.17E-04
X73_RS06270	ornithine decarboxylase	PM0806	*speF*	1.82	1.98E-06
X73_RS03610	TcmP	PM1322	*tcmP*	1.19	1.11E-02
X73_RS02750	Putative multidrug export ATP-binding/permease protein SAS1788	PM1473		1.03	3.36E-06
X73_RS02745	CydD	PM1474	*cydD*	1.59	6.95E-08
X73_RS10070	PfhR	PM0040	*pfhR*	2.16	3.95E-06
X73_RS10065	Inner membrane ABC transporter ATP-binding protein YddA	PM0041		4.74	2.35E-06
X73_RS10060	no known function	PM0042		4.24	7.89E-06
X73_RS10020	FbpB	PM0050	*fbpB*	1.38	6.55E-05
X73_RS10015	FbpA	PM0051	*fbpA*	1.98	6.95E-08

[a] Multiple genome entries exist for strain X-73; NCBI genome accession: NZ_CM001580.1 was chosen for this study. Differentially expressed open reading frames identified in this study that have not been assigned a locus tag in the genome entry NZ_CM001580.1 have been assigned a "b" suffix to the name of the gene tag numerically preceding it in the genome.

[b] The *P. multocida* sequence was used to interrogate the SwissProt database (excluding proteins belonging to *Pasteurellaceae*). Only proteins with a high similarity with known proteins/domains are assigned a predicted function i.e. an alignment score of ≥ 200 (graphical output = red bar).

[c] The Pm70 genome is available from NCBI genome database, accession: NC_002663.1 or AE004439.1.

Table S10. Genes in the wild-type *P. multocida* X-73 strain showing decreased RNA expression during growth in low iron conditions, in comparison to gene expression during growth in rich medium.

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS09265	fumarate reductase flavoprotein subunit	PM0201	*frdA*	-1.21	5.59E-03
X73_RS09165	two-component response regulator	PM0219	*arcA*	-1.43	7.68E-04
X73_RS08960	cytidine deaminase	PM0259	*cdd*	-1.29	2.56E-02
X73_RS08930	peptidase T	PM0265	*pepT*	-2.53	2.31E-02
X73_RS08565	OmpW	PM0331	*ompW*	-1.18	2.89E-03
X73_RS07615	Cytochrome c-type protein TorY	PM1792		-2.71	3.88E-02
X73_RS03735	no known function	PM1299		-1.40	3.67E-02
X73_RS03245	Ribose import permease protein RbsC	PM1378		-1.04	5.50E-05
X73_RS03240	Ribose import ATP-binding protein RbsA 2	PM1379		-1.37	7.11E-03
X73_RS01225	Cytochrome c-type protein TorC	PM1792		-1.70	7.39E-03
X73_RS01220	TorA	PM1793	*torA*	-1.21	2.59E-03
X73_RS00940	Uncharacterized protein YkgE	PM1853		-1.18	3.59E-05
X73_RS00935	Uncharacterized electron transport protein YkgF	PM1854		-1.39	1.84E-05
X73_RS00930	Uncharacterized protein YkgG	PM1855		-1.36	1.30E-04
X73_RS00245	3-isopropylmalate dehydrogenase	PM1961	*leuB*	-1.08	5.59E-03
X73_RS10150	cytochrome c nitrite reductase	PM0023	*nrfA*	-1.93	1.40E-03
X73_RS10145	cytochrome c-type protein NrfB	PM0024	*nrfB*	-1.40	3.30E-03
X73_RS10140	NrfC	PM0025	*nrfC*	-1.57	2.21E-02
X73_RS10135	NrfD	PM0026	*nrfD*	-1.04	8.06E-03

[a] Multiple genome entries exist for strain X-73; NCBI genome accession: NZ_CM001580.1 was chosen for this study. Differentially expressed open reading frames identified in this study that have not been assigned a locus tag in the genome entry NZ_CM001580.1 have been assigned a "b" suffix to the name of the gene tag numerically preceding it in the genome.
[b] The *P. multocida* sequence was used to interrogate the SwissProt database (excluding proteins belonging to *Pasteurellaceae*). Only proteins with a high similarity with known proteins/domains are assigned a predicted function i.e. an alignment score of ≥ 200 (graphical output = red bar).
[c] The Pm70 genome is available from NCBI genome database, accession: NC_002663.1 or AE004439.1.

Table S11. Genes in the wild-type *P. multocida* X-73 strain showing increased RNA expression during growth in reduced oxygen conditions, in comparison to gene expression during aerobic growth.

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS09945	autonomous glycyl radical cofactor GrcA	PM0064		1.38	4.52E-04
X73_RS09920	6-phosphofructokinase	PM0069	*pfkA*	1.46	3.13E-07
X73_RS09895	Probable formate transporter 1	PM0074		2.80	3.55E-05
X73_RS09890	PflB	PM0075	*pflB*	4.21	7.26E-10
X73_RS09890b	no known function			1.94	1.65E-02
X73_RS09805	no known function	PM0092		1.11	6.09E-05
X73_RS09805b	no known function			1.28	1.99E-02
X73_RS09505	ribokinase	PM0152	*rbsK*	1.04	1.96E-05
X73_RS09500	D-ribose transporter subunit RbsB	PM0153	*rbsB_1*	1.82	1.44E-03
X73_RS09495	ribose ABC transporter permease protein	PM0154	*rbsC_1*	2.13	1.13E-03
X73_RS09490	D-ribose transporter ATP binding protein	PM0155	*rbsA_1*	2.37	6.71E-05
X73_RS09485	D-ribose pyranase	PM0156	*rbsD*	1.98	4.43E-04
X73_RS09465	allantoate amidohydrolase	PM0160		2.53	5.08E-11
X73_RS09280	fumarate reductase subunit D	PM0198	*frdD*	3.18	5.40E-09
X73_RS09275	fumarate reductase subunit C	PM0199	*frdC*	3.20	3.00E-09
X73_RS09270	fumarate reductase iron-sulfur subunit	PM0200	*frdB*	3.02	5.38E-10
X73_RS09265	fumarate reductase flavoprotein subunit	PM0201	*frdA*	2.75	7.35E-08
X73_RS08990	no known function	PM0253		2.52	4.28E-05
X73_RS08960	cytidine deaminase	PM0259	*cdd*	3.18	5.28E-08
X73_RS08930	peptidase T	PM0265	*pepT*	3.26	7.74E-06
X73_RS08820	MesJ	PM0289	*mesJ*	1.12	1.06E-02
X73_RS08815	pyridoxamine kinase	PM0290	*pdxY*	1.28	3.42E-05
X73_RS08810	no known function	PM0291		1.37	6.57E-06
X73_RS08565	OmpW	PM0331	*ompW*	1.49	1.09E-05
X73_RS10595	no known function			1.42	4.81E-03

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS08155	glucose-6-phosphate isomerase	PM0416	*pgi*	2.34	5.38E-10
X73_RS08150	no known function	PM0417		2.02	5.82E-04
X73_RS08010	ResA	PM0447	*resA*	1.26	1.50E-03
X73_RS08005	Uncharacterized ABC transporter ATP-binding protein MJ1508	PM0448		1.20	1.12E-04
X73_RS08000	no known function	PM0449		1.47	1.29E-03
X73_RS07995	no known function	PM0450		1.47	8.46E-04
X73_RS07990	no known function	PM0451		1.44	1.07E-05
X73_RS07985	no known function	PM0452		1.97	8.07E-06
X73_RS07980	no known function	PM0453		1.41	8.20E-07
X73_RS07975	peptidase E	PM0454	*pepE*	1.43	4.18E-05
X73_RS07945	PII uridylyl-transferase	PM0460	*glnD*	1.03	2.66E-07
X73_RS10315	no known function			1.60	7.20E-06
X73_RS07830	CspD	PM0481	*cspD*	1.46	5.42E-03
X73_RS07620	TorA	PM1793	*torA*	4.59	9.73E-09
X73_RS07615	Cytochrome c-type protein TorY			3.93	7.57E-06
X73_RS07600	putative sulfite oxidase subunit YedY	PM0537		1.17	2.69E-02
X73_RS07595	putative sulfite oxidase subunit YedZ	PM0538		1.07	1.07E-02
X73_RS07585	4-alpha-glucanotransferase	PM0540	*malQ*	2.01	1.68E-06
X73_RS07580	glycogen branching enzyme	PM0541	*glgB*	2.39	9.77E-07
X73_RS07575	Glgx	PM0542	*glgx*	2.30	4.08E-06
X73_RS07570	glucose-1-phosphate adenylyltransferase	PM0543	*glgC*	1.63	4.33E-08
X73_RS07565	glycogen synthase	PM0544	*glgA*	1.18	6.89E-03
X73_RS07555	phosphoenolpyruvate carboxylase	PM0546	*ppc*	2.35	7.77E-11
X73_RS07550	DNA-binding transcriptional repressor PurR	PM0547	*purR*	1.76	2.18E-07
X73_RS07535	malate dehydrogenase	PM0550	*mdh_2*	1.42	2.91E-11
X73_RS07450	DNA-binding transcriptional activator GcvA	PM0567	*gcvA*	1.74	6.43E-06
X73_RS07445	putative RNA 2'-O-ribose methyltransferase	PM0568		1.36	6.07E-08
X73_RS07445b	no known function			1.13	6.27E-03

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS07355	Putative kinase VV1_2562	PM0587		3.23	1.56E-07
X73_RS07190	no known function	PM0622	*moaE*	2.56	3.30E-08
X73_RS07185	molybdopterin synthase small subunit	PM0623	*moaD*	2.55	8.67E-08
X73_RS07180	molybdenum cofactor biosynthesis protein C	PM0624	*moaC*	2.42	1.02E-08
X73_RS07175	molybdenum cofactor biosynthesis protein A	PM0625	*moaA*	2.13	4.84E-10
X73_RS07095	dithiobiotin synthetase	PM0641	*bioD1*	3.25	9.09E-06
X73_RS07090	Mlc	PM0642	*mlc*	1.02	2.89E-03
X73_RS07020	tellurite resistance protein TehB	PM0656	*tehB*	1.50	5.29E-05
X73_RS06935	no known function	PM0674		1.10	1.49E-06
X73_RS06930	N-acetyl-D-glucosamine kinase	PM0675		1.51	3.48E-08
X73_RS06850	no known function	PM0688		2.04	2.67E-05
X73_RS06710	TtrA	PM0721	*ttrA*	1.97	1.17E-04
X73_RS06705	TtrC	PM0722	*ttrC*	2.59	3.07E-05
X73_RS06700	TTRB	PM0723	*ttrB*	1.87	8.46E-04
X73_RS06650	Putative NAD(P)H nitroreductase SERP2086	PM0733		1.80	1.67E-04
X73_RS06615	no known function	PM0741		1.38	6.32E-05
X73_RS06470	no known function	PM0768		1.29	1.36E-02
X73_RS06345	COG1063, Tdh, Threonine dehydrogenase and related Zn-dependent dehydrogenases			3.77	2.42E-03
X73_RS06340	3-ketoacyl-			1.37	1.10E-02
X73_RS06335	no known function			1.03	1.49E-02
X73_RS06280	Uncharacterized protein YddB	PM0803		2.56	2.31E-06
X73_RS06275	Probable zinc protease PqqL	PM0804	*pqqL*	1.36	5.47E-03
X73_RS05915	PTS system N-acetylglucosamine-specific EIICBA component	PM0876		1.06	2.02E-05
X73_RS05810	phosphoenolpyruvate-protein phosphotransferase	PM0897	*ptsI*	1.36	5.06E-04
X73_RS05805	PtsH	PM0898	*ptsH*	1.46	4.24E-05
X73_RS05680	D-allose-binding periplasmic protein			1.04	1.95E-02
X73_RS05640	methionine sulfoxide reductase B	PM0923		1.70	2.81E-04

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS05635	glyceraldehyde-3-phosphate dehydrogenase	PM0924	*gapdH*	2.74	1.99E-11
X73_RS05595	fructose-1,6-bisphosphatase	PM0930	*fbp*	1.30	7.55E-05
X73_RS05550	Probable cytochrome c peroxidase	PM0939		2.32	4.62E-05
X73_RS05545	anaerobic ribonucleoside triphosphate reductase	PM0940	*nrdD*	1.74	3.49E-06
X73_RS05545b	anaerobic ribonucleotide reductase-activating protein	PM0941	*nrdG*	1.87	4.78E-06
X73_RS05540	no known function			1.58	2.07E-03
X73_RS05510	keto-hydroxyglutarate-aldolase/keto-deoxy- phosphogluconate aldolase	PM0947	*eda*	2.41	8.78E-08
X73_RS05505	putative sulfate transporter YchM	PM0948		2.54	2.64E-08
X73_RS05475	AfuA	PM0953	*afuA_1*	1.73	3.62E-06
X73_RS05470	AfuA	PM0954	*afuA_2*	2.02	2.04E-07
X73_RS05465	AfuA	PM0955	*afuA_3*	2.11	1.01E-06
X73_RS05460	AfuB	PM0956	*afuB_1*	1.92	8.67E-08
X73_RS05455	ferric transporter ATP-binding subunit	PM0957	*afuC*	1.81	5.06E-08
X73_RS05415	no known function	PM0965		1.19	2.24E-02
X73_RS06125	no known function	PM0836		1.38	1.27E-06
X73_RS06130	no known function	PM0835		1.33	7.33E-03
X73_RS06135	PTS system mannose-specific EIIAB component	PM0834		3.55	5.49E-09
X73_RS06140	PtnC	PM0833	*ptnC*	3.52	5.38E-10
X73_RS06145	mannose-specific PTS system protein IID	PM0832	*ptnD*	3.57	8.93E-11
X73_RS06150	no known function	PM0830		1.87	3.31E-05
X73_RS06155	Pmi	PM0829	*pmi*	1.91	5.13E-07
X73_RS06185	fumarate hydratase	PM0823	*fumC*	4.41	6.32E-11
X73_RS06265	putrescine transporter	PM0807	*potE*	2.78	2.32E-09
X73_RS06270	ornithine decarboxylase	PM0806	*speF*	2.96	4.38E-10
X73_RS05215	putative nicotinate phosphoribosyltransferase	PM0999		1.37	2.18E-07
X73_RS05160	no known function	PM1010		1.01	3.12E-03
X73_RS05040	aldose 1-epimerase	PM1034	*galM*	1.25	4.17E-02

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS05035	galactokinase	PM1035	*galK*	1.17	6.57E-06
X73_RS05020	MglB	PM1038	*mglB*	2.47	4.54E-06
X73_RS05015	galactose/methyl galaxtoside transporter ATP-binding protein	PM1039	*mglA*	1.56	2.19E-03
X73_RS05010	beta-methylgalactoside transporter inner membrane component	PM1040	*mglC*	1.37	1.04E-07
X73_RS04870	no known function	PM1064		1.74	1.14E-08
X73_RS04840	no known function	PM1070		1.18	7.81E-05
X73_RS10430	Putative transport protein AHA_3492	PM1071		1.70	4.63E-05
X73_RS10430b	no known function			1.64	3.55E-02
X73_RS04825	YhxB	PM1074	*yhxB*	1.63	1.14E-10
X73_RS04695	aspartate ammonia-lyase	PM1103	*aspA*	1.34	9.14E-07
X73_RS04570	no known function	PM1125		1.29	8.81E-03
X73_RS04565	CRISPR-associated endonuclease Cas1	PM1126		1.59	9.24E-04
X73_RS04555	no known function	PM1128		1.27	9.84E-03
X73_RS04410	no known function	PM1158		1.18	2.37E-07
X73_RS04280	ExbB	PM1186	*exbB*	1.54	3.71E-05
X73_RS04275	no known function	PM1187	*exbD*	1.51	2.44E-03
X73_RS04270	TonB	PM1188	*tonB*	1.79	2.04E-04
X73_RS03950	no known function	PM1255		1.01	3.20E-02
X73_RS03795	Universal stress protein A UspA	PM1286	*uspA*	1.49	4.66E-05
X73_RS03735	no known function	PM1299		3.56	6.57E-08
X73_RS03730	Indole-3-acetyl-aspartic acid hydrolase	PM1300		3.94	1.35E-09
X73_RS03675	triosephosphate isomerase	PM1311	*tpiA_1*	2.51	1.99E-11
X73_RS03595	no known function	PM1325		1.76	3.61E-08
X73_RS03590	Putative ribose/galactose/methyl galactoside import ATP-binding protein	PM1326		2.25	2.88E-07
X73_RS03585	RbsC	PM1327	*rbsC_3*	2.01	6.12E-09
X73_RS03440	Glutathione amide-dependent peroxidase	PM1347		1.17	6.91E-06
X73_RS03435	no known function	PM1348	*slyX*	1.25	2.99E-03

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS03015	TnaB	PM1419	*tnaB*	1.03	1.73E-02
X73_RS03010	tryptophanase	PM1420	*tnaA*	1.72	5.35E-04
X73_RS03005	oxaloacetate decarboxylase subunit gamma	PM1421		1.47	2.12E-04
X73_RS03000	pyruvate carboxylase subunit B	PM1422		1.89	3.90E-10
X73_RS02995	Oxaloacetate decarboxylase beta chain 2	PM1423		2.06	1.49E-11
X73_RS02940	anaerobic C4-dicarboxylate transporter	PM1434	*dcuB*	2.79	5.38E-10
X73_RS02910	anaerobic glycerol-3-phosphate dehydrogenase subunit C	PM1440	*glpC*	1.54	4.48E-09
X73_RS02905	anaerobic glycerol-3-phosphate dehydrogenase subunit B	PM1441	*glpB*	1.36	3.07E-09
X73_RS02900	anaerobic-glycerol-3-phosphate dehydrogenase subunit A	PM1442	*glpA*	1.02	9.67E-08
X73_RS02895	GlpT	PM1443	*glpT*	1.29	1.38E-08
X73_RS02890	glycerophosphodiester phosphodiesterase	PM1444	*glpQ*	2.63	2.13E-10
X73_RS02885	Glycerol uptake facilitator protein	PM1445	*glpF*	2.33	2.67E-09
X73_RS02880	glycerol kinase	PM1446	*glpK*	2.21	6.22E-12
X73_RS02855	XynC	PM1451	*xynC*	1.18	2.89E-03
X73_RS02845	bifunctional acetaldehyde-CoA/alcohol dehydrogenase	PM1453	*adh2*	5.67	3.52E-08
X73_RS02820	PgtC	PM1458	*pgtC*	2.27	1.36E-06
X73_RS02815	PgtB	PM1459	*pgtB*	2.18	2.56E-05
X73_RS02810	Phosphoglycerate transport system transcriptional regulatory protein PgtA	PM1460		1.33	3.07E-05
X73_RS02620	no known function			1.24	2.79E-02
X73_RS02585	phosphoglyceromutase	PM1506	*gpmA*	2.38	8.93E-11
X73_RS02410	phosphoenolpyruvate carboxykinase	PM1542	*pckA*	4.51	7.21E-12
X73_RS02195	no known function	PM1585		1.92	8.40E-06
X73_RS02155	NapD	PM1593	*napD*	3.23	2.84E-03
X73_RS02150	nitrate reductase	PM1594	*napA*	1.53	1.14E-03
X73_RS02140	quinol dehydrogenase membrane component	PM1596	*napH*	1.13	2.21E-02
X73_RS01755	Anaerobic sulfatase-maturating enzyme homologue AslB	PM1678		2.18	1.43E-05
X73_RS01750	no known function	PM1679		1.79	8.95E-04

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS01745	no known function	PM1680		1.56	3.79E-03
X73_RS01740	Uncharacterized 52.8 kDa protein in TAR-I ttuC' 3′ region	PM1681		1.68	3.32E-04
X73_RS01735	no known function	PM1682		1.52	1.54E-05
X73_RS01690	Sec-independent protein translocase protein TatC	PM1691	*tatC*	1.11	8.78E-06
X73_RS01420	GntP	PM1740	*gntP*	1.09	1.33E-05
X73_RS01350	PTS system glucose-specific EIICBA component	PM1752	*ptsG*	1.76	2.58E-06
X73_RS01340	DmsA	PM1754	*dmsA*	5.67	5.68E-08
X73_RS01335	DmsB	PM1755	*dmsB*	5.89	1.64E-07
X73_RS01330	DmsC	PM1756	*dmsC*	5.94	4.30E-12
X73_RS01325	twin-argninine leader-binding protein DmsD	PM1757		6.10	5.38E-10
X73_RS01320	no known function	PM1758		4.24	4.30E-12
X73_RS01310	IunH	PM1767	*iunH*	1.78	8.04E-10
X73_RS01230	no known function			2.44	2.82E-03
X73_RS01225	Cytochrome c-type protein TorC	PM1792		2.55	1.65E-06
X73_RS01220	TorA	PM1793	*torA*	3.08	7.85E-10
X73_RS01215	chaperone protein TorD	PM1794	*torD*	3.28	3.69E-09
X73_RS01005	Nicotinamide riboside transporter PnuC	PM1838		1.57	1.14E-08
X73_RS01005b	no known function			2.33	2.78E-02
X73_RS00965	ScrB	PM1848	*scrB*	1.45	8.47E-05
X73_RS00960	aminoimidazole riboside kinase	PM1849		1.30	2.31E-06
X73_RS00905	phosphoglycerate kinase	PM1860	*pgk*	1.75	3.00E-08
X73_RS00900	fructose-bisphosphate aldolase	PM1861	*fba*	1.88	3.02E-10
X73_RS00855	phosphopyruvate hydratase	PM1871	*eno*	2.17	8.25E-10
X73_RS00740	no known function	PM1894		2.78	2.33E-05
X73_RS00735	Inner membrane protein YjjP	PM1895		2.12	3.37E-04
X73_RS00620	anaerobic C4-dicarboxylate transporter	PM1918		2.03	5.38E-10
X73_RS00615	no known function	PM1919		1.22	1.94E-03
X73_RS00375	COG1414, IclR, Transcriptional regulator			1.83	1.63E-02

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS00205	PTS system glucitol/sorbitol-specific EIIC component	PM1971		1.44	1.72E-02
X73_RS10255	malic enzyme	PM0002	*mdh_1*	2.12	1.04E-08
X73_RS10240	cytochrome c biogenesis protein CcmA	PM0005	*ccmA*	1.42	5.73E-05
X73_RS10235	CcmB	PM0006	*ccmB*	1.45	4.23E-04
X73_RS10230	CcmC	PM0007	*ccmC*	1.23	5.99E-06
X73_RS10225	CcmD	PM0008	*ccmD*	1.26	1.37E-02
X73_RS10220	cytochrome c-type biogenesis protein CcmE	PM0009	*ccmE*	1.16	1.18E-02
X73_RS10215	CcmF	PM0010	*ccmF*	1.30	2.78E-06
X73_RS10210	DsbE	PM0011	*dsbE_1*	1.25	2.85E-02
X73_RS10205	CcmH	PM0012	*ccmH_1*	1.05	2.06E-03
X73_RS10200	CcmH	PM0013	*ccmH_2*	1.20	3.28E-03
X73_RS10150	cytochrome c nitrite reductase	PM0023	*nrfA*	4.16	5.38E-10
X73_RS10145	cytochrome c-type protein NrfB	PM0024	*nrfB*	4.58	1.43E-11
X73_RS10140	NrfC	PM0025	*nrfC*	4.64	5.58E-10
X73_RS10135	NrfD	PM0026	*nrfD*	4.64	4.30E-12
X73_RS10130	NrfE	PM0027	*nrfE*	1.69	2.12E-09
X73_RS10125	DsbE	PM0028	*dsbE_2*	1.36	1.98E-03
X73_RS10120	NrfF	PM0029	*nrfF_1*	1.41	5.26E-04
X73_RS10115	NrfF	PM0030	*nrfF_2*	1.15	3.22E-03
X73_RS10105	HktE	PM0032	*hktE*	1.13	2.14E-03
X73_RS10090	L-serine dehydratase 1	PM0036	*sdaA*	1.13	3.75E-08
X73_RS10085	SdaC	PM0037	*sdaC*	1.20	1.19E-04
X73_RS10040	disulfide bond formation protein B	PM0046	*dsbB*	1.00	2.55E-05
X73_RS10035	sodium/proton antiporter	PM0047	*nhaB*	1.15	3.49E-06

[a] Multiple genome entries exist for strain X-73; NCBI genome accession: NZ_CM001580.1 was chosen for this study. Differentially expressed open reading frames identified in this study that have not been assigned a locus tag in the genome entry NZ_CM001580.1 have been assigned a "b" suffix to the name of the gene tag (closest numerically) preceding it in the genome.

[b] The *P. multocida* sequence was used to interrogate the SwissProt database (excluding proteins belonging to *Pasteurellaceae*). Only proteins with a high similarity with known proteins/domains are assigned a predicted function i.e. an alignment score of ≥ 200 (graphical output = red bar).
[c] The Pm70 genome is available from NCBI genome database, accession: NC_002663.1 or AE004439.1.

Table S12. Genes in the wild-type *P. multocida* X-73 strain showing decreased RNA expression during growth in reduced oxygen conditions, in comparison to gene expression during aerobic growth.

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS09650	arginine transporter ATP-binding subunit	PM0123	*artP*	-2.24	9.79E-06
X73_RS09645	Putative ABC transporter arginine-binding protein 2 ArtI	PM0124	*artI*	-1.77	4.72E-04
X73_RS09640	arginine transporter permease subunit ArtQ	PM0125	*artQ*	-1.73	1.60E-03
X73_RS09635	arginine transporter permease subunit ArtM	PM0126	*artM*	-1.32	1.06E-02
X73_RS09370	UDP-N-acetylglucosamine 1-carboxyvinyltransferase	PM0180	*murZ*	-1.01	2.54E-07
X73_RS09255	no known function	PM0203		-1.04	5.73E-05
X73_RS09250	bifunctional aconitate hydratase 2/2-methylisocitrate dehydratase	PM0204	*acnB*	-1.07	1.23E-04
X73_RS09110	queuine tRNA-ribosyltransferase	PM0229	*tgt*	-1.06	4.88E-05
X73_RS09090	Uncharacterized protease YegQ	PM0233		-1.45	7.53E-04
X73_RS09075	Periplasmic dipeptide transport protein DppA	PM0236	*dppA*	-1.50	6.57E-08
X73_RS08955	PotD	PM0260	*potD_1*	-2.05	2.91E-04
X73_RS08950	PotD	PM0261	*potD_2*	-1.56	5.51E-06
X73_RS08890	C4-dicarboxylate TRAP transporter large permease protein DctM	PM0273		-1.71	5.43E-04
X73_RS08885	C4-dicarboxylate TRAP transporter small permease protein DctQ	PM0274		-2.29	2.65E-04
X73_RS08880	C4-dicarboxylate-binding periplasmic protein DctP	PM0275		-1.90	1.15E-04
X73_RS08785	inositol-5-monophosphate dehydrogenase	PM0295	*guaB*	-1.02	1.37E-06
X73_RS08785b	Carboxynorspermidine synthase	PM0372		-1.21	1.85E-04
X73_RS08785c	no known function	PM0372		-1.56	7.44E-03
X73_RS08310	electron transport complex protein RnfG	PM0383		-1.17	1.56E-03
X73_RS08280	OmpH	PM0389	*ompH_2*	-1.27	2.71E-06
X73_RS08185	formate dehydrogenase accessory protein	PM0410	*fdhD*	-1.16	2.24E-03
X73_RS07870	no known function	PM0472		-1.58	1.18E-08
X73_RS07865	no known function	PM0473		-1.67	5.03E-08
X73_RS07860	no known function	PM0474		-1.87	1.18E-05
X73_RS07650	Amino-acid carrier protein AlsT	PM0529		-1.42	1.79E-05

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS07330	threonyl-tRNA synthetase	PM0593	*thrS*	-1.22	5.95E-09
X73_RS07105	50S ribosomal protein L25	PM0639	*rpL25*	-1.22	8.47E-08
X73_RS06985	bifunctional chorismate mutase/prephenate dehydrogenase	PM0664	*tyrA*	-1.48	8.39E-04
X73_RS06980	phospho-2-dehydro-3-deoxyheptonate aldolase	PM0665	*aroF*	-2.09	8.22E-05
X73_RS06975	RsgA	PM0666	*rsgA_1*	-1.40	1.02E-03
X73_RS06970	RsgA	PM0667	*rsgA_2*	-1.69	8.59E-05
X73_RS06965	fumarate/nitrate reduction transcriptional regulator	PM0668	*fnr*	-1.71	7.64E-05
X73_RS06900	no known function	PM0681		-1.01	1.11E-03
X73_RS06830	acetyl-CoA synthetase	PM0692	*acsA*	-1.30	1.94E-02
X73_RS06720	Uncharacterized sodium-dependent transporter MJ1319	PM0718		-3.25	3.34E-07
X73_RS06660	riboflavin synthase subunit beta	PM0731	*ribH*	-1.11	1.76E-04
X73_RS06520	translation initiation factor IF-2	PM0759	*infB*	-2.02	6.12E-09
X73_RS06515	transcription elongation factor NusA	PM0760	*nusA*	-1.58	2.18E-05
X73_RS06430	UDP-glucose 6-dehydrogenase	PM0776		-1.07	1.54E-05
X73_RS05830	dihydrolipoamide dehydrogenase	PM0893	*lpdA*	-2.59	6.18E-12
X73_RS05825	dihydrolipoamide acetyltransferase	PM0894	*aceF*	-2.42	6.18E-12
X73_RS05820	pyruvate dehydrogenase subunit E1	PM0895	*aceE*	-2.29	1.17E-11
X73_RS05620	high-affinity zinc transporter periplasmic component	PM0926	*fimA*	-1.05	2.16E-05
X73_RS05620b	no known function	PM0972		-1.45	1.88E-04
X73_RS05365	CydB	PM0973	*cydB*	-1.83	1.49E-11
X73_RS05360	CydA	PM0974	*cydA*	-1.70	1.23E-07
X73_RS05285	Glutamine transporter 1	PM0989		-1.08	1.24E-05
X73_RS06160	N-acetylglutamate synthase	PM0828	*argA*	-1.35	5.35E-04
X73_RS06165	glutaredoxin 1	PM0827	*grx*	-1.58	1.05E-03
X73_RS06260	ornithine carbamoyltransferase	PM0808	*arg*	-1.26	4.23E-06
X73_RS05075	S-adenosylmethionine synthetase	PM1027	*metX*	-1.28	5.53E-03
X73_RS05000	Phosphoethanolamine transferase EptA	PM1042		-1.04	2.78E-03

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS04845	Long-chain fatty acid transport protein or 47 kDa outer membrane protein	PM1069		-1.20	8.42E-08
X73_RS04810	no known function	PM1078		-1.54	4.63E-05
X73_RS04805	no known function	PM1079		-1.62	4.09E-05
X73_RS04800	no known function	PM1080		-1.46	3.14E-03
X73_RS04755	sodium/panthothenate symporter	PM1089	*panF*	-1.03	6.57E-06
X73_RS04690	MtrF	PM1104	*mtrF*	-1.86	2.03E-06
X73_RS04650	DeaD	PM1112	*deaD*	-1.39	5.02E-04
X73_RS04595	N-acetyl-gamma-glutamyl-phosphate reductase	PM1118	*argC*	-1.32	1.18E-04
X73_RS04505	50S ribosomal protein L31	PM1142	*rpL31_2*	-1.10	3.04E-06
X73_RS04325	50S ribosomal protein L9	PM1177	*rpL9*	-1.67	5.03E-08
X73_RS04320	30S ribosomal protein S18	PM1178	*rpS18*	-1.56	1.00E-05
X73_RS04315	primosomal replication protein N	PM1179	*priB*	-1.54	7.64E-05
X73_RS04310	30S ribosomal protein S6	PM1180	*rpS6*	-1.48	1.76E-04
X73_RS04160	no known function	PM1211		-1.99	3.33E-06
X73_RS10470	MerT	PM1212	*merT*	-1.39	5.20E-03
X73_RS04155	MerP	PM1213	*merP*	-1.62	2.51E-04
X73_RS04065	pyridoxine biosynthesis protein	PM1232		-1.76	8.67E-08
X73_RS04060	Pyridoxal 5′-phosphate synthase subunit PdxT	PM1233		-1.53	8.67E-08
X73_RS03930	Inner membrane metabolite transport protein YhjE	PM1259		-1.65	5.13E-03
X73_RS03925	thiamine-phosphate pyrophosphorylase	PM1260	*thiE*	-1.88	3.35E-03
X73_RS03920	phosphomethylpyrimidine kinase	PM1261	*thiD*	-1.84	9.34E-03
X73_RS03915	hydroxyethylthiazole kinase	PM1262	*thiM*	-2.69	1.36E-03
X73_RS03915b	no known function			-2.18	1.44E-02
X73_RS03910	no known function	PM1263		-2.81	2.02E-07
X73_RS03905	no known function	PM1264		-1.32	1.63E-02
X73_RS03805	ketol-acid reductoisomerase	PM1284	*ilvC*	-1.80	3.97E-06
X73_RS03755	30S ribosomal protein S16	PM1294a	*rpsP*	-1.35	2.27E-04

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS03750	16S rRNA-processing protein	PM1296	*rimM*	-1.40	5.56E-07
X73_RS03745	tRNA	PM1297	*trmD*	-1.53	5.38E-10
X73_RS03740	50S ribosomal protein L19	PM1298	*rpL19*	-1.47	1.11E-08
X73_RS03395	elongation factor G	PM1356	*fusA*	-1.01	2.23E-08
X73_RS10485	elongation factor Tu	PM1357	*tufA*	-1.00	8.24E-08
X73_RS03260	fructokinase	PM1375		-1.38	2.58E-06
X73_RS03245	Ribose import permease protein RbsC	PM1378		-1.37	6.04E-07
X73_RS03240	Ribose import ATP-binding protein RbsA 2	PM1379		-1.32	2.61E-03
X73_RS03225	agmatinase	PM1381	*speE*	-1.22	2.23E-08
X73_RS03220	arginine decarboxylase	PM1382	*speA*	-1.15	5.06E-08
X73_RS03185	50S ribosomal protein L17	PM1389	*rpL17*	-1.38	5.95E-09
X73_RS03180	DNA-directed RNA polymerase subunit alpha	PM1390	*rpoA*	-1.30	2.97E-09
X73_RS03175	30S ribosomal protein S4	PM1391	*rpS4*	-1.19	2.74E-04
X73_RS03170	30S ribosomal protein S11	PM1392	*rpS11*	-1.23	1.29E-08
X73_RS03165	30S ribosomal protein S13	PM1393	*rpS13*	-1.05	9.73E-09
X73_RS03160	50S ribosomal protein L36	PM1394	*rpL36*	-1.28	9.73E-09
X73_RS03155	preprotein translocase subunit SecY	PM1395	*secY*	-1.40	3.00E-08
X73_RS03150	50S ribosomal protein L15	PM1396	*rpL15*	-1.49	3.77E-07
X73_RS03145	30S ribosomal protein S5	PM1398	*rpS5*	-1.49	9.46E-08
X73_RS03140	50S ribosomal protein L18	PM1399	*rpL18*	-1.51	6.54E-08
X73_RS03135	50S ribosomal protein L6	PM1400	*rpL6*	-1.41	5.21E-08
X73_RS03130	30S ribosomal protein S8	PM1401	*rpS8*	-1.29	1.36E-06
X73_RS03125	30S ribosomal protein S14	PM1402	*rpS14*	-1.34	2.21E-08
X73_RS10495	50S ribosomal protein L5	PM1403	*rpL5*	-1.28	1.63E-08
X73_RS03120	50S ribosomal protein L24	PM1404	*rpL24*	-1.13	2.98E-05
X73_RS03115	50S ribosomal protein L14	PM1405	*rpL14*	-1.17	1.55E-07
X73_RS03075	30S ribosomal protein S17	PM1406	*rpS17*	-1.47	1.18E-08
X73_RS03070	50S ribosomal protein L29	PM1407	*rpL29*	-1.80	1.28E-06

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS03065	50S ribosomal protein L16	PM1408	*rpL16*	-1.76	5.90E-07
X73_RS03060	30S ribosomal protein S3	PM1409	*rpS3*	-1.75	4.12E-09
X73_RS03055	50S ribosomal protein L22	PM1410	*rpL22*	-1.71	8.77E-07
X73_RS03055b	30S ribosomal protein S19	PM1411	*rpS19*	-1.70	3.98E-09
X73_RS03050	50S ribosomal protein L2	PM1412	*rpL2*	-1.67	3.11E-09
X73_RS03045	50S ribosomal protein L23	PM1413	*rpL23*	-1.61	7.39E-09
X73_RS03040	50S ribosomal protein L4	PM1414	*rpL4*	-1.56	3.30E-08
X73_RS03035	50S ribosomal protein L3	PM1415	*rpL3*	-1.46	8.24E-08
X73_RS03030	30S ribosomal protein S10	PM1416	*rpS10*	-1.53	1.68E-08
X73_RS02730	no known function	PM1477		-2.35	7.74E-06
X73_RS02725	no known function	PM1478		-2.62	1.11E-03
X73_RS02720	Probable metallo-hydrolase BURPS1710b_2304	PM1479		-1.03	4.21E-02
X73_RS02715	no known function	PM1480		-1.09	2.32E-05
X73_RS02605	carbamoyl phosphate synthase small subunit	PM1502	*carA*	-1.05	1.71E-02
X73_RS02230	no known function	PM1578		-1.05	6.16E-03
X73_RS02005	threonine dehydratase	PM1624	*ilvA*	-1.09	6.07E-06
X73_RS02000	dihydroxy-acid dehydratase	PM1625	*ilvD*	-1.12	1.17E-03
X73_RS01440	DNA-directed RNA polymerase subunit beta'	PM1736	*rpoC*	-1.50	5.38E-10
X73_RS01435	DNA-directed RNA polymerase subunit beta	PM1737	*rpoB*	-1.26	6.57E-08
X73_RS01430	50S ribosomal protein L7/L12	PM1738	*rpL12*	-1.13	9.73E-09
X73_RS01425	50S ribosomal protein L10	PM1739	*rpL10*	-1.08	1.67E-04
X73_RS01410	50S ribosomal protein L1	PM1742	*rpL1*	-1.85	5.05E-09
X73_RS01405	50S ribosomal protein L11	PM1743	*rpL11*	-1.75	1.74E-07
X73_RS01170	tRNA	PM1803	*trmA*	-1.04	2.16E-02
X73_RS00950	adenylosuccinate lyase	PM1851	*purB*	-1.24	1.36E-06
X73_RS00945	LctP	PM1852	*lctP*	-1.19	1.71E-05
X73_RS00940	Uncharacterized protein YkgE	PM1853		-1.17	7.74E-06
X73_RS00935	Uncharacterized electron transport protein YkgF	PM1854		-1.39	3.93E-06

PMX73 locus tag [a]	Predicted function/product [b]	Pm70 locus tag [c]	Annotated gene name	Expression ratio (log_2)	FDR
X73_RS00930	Uncharacterized protein YkgG	PM1855		-1.35	2.94E-05
X73_RS00560	lipoyl synthase	PM1930	*lipA*	-1.59	4.24E-05
X73_RS00255	isopropylmalate isomerase small subunit	PM1959	*leuD*	-1.19	2.95E-04
X73_RS00250	isopropylmalate isomerase large subunit	PM1960	*leuC*	-1.54	5.30E-06
X73_RS00245	3-isopropylmalate dehydrogenase	PM1961	*leuB*	-1.71	3.06E-05
X73_RS00240	2-isopropylmalate synthase	PM1962	*leuA*	-1.76	1.47E-06
X73_RS00190	trigger factor	PM1975	*tig*	-1.38	1.63E-08
X73_RS00145	30S ribosomal protein S2	PM1984	*rpS2*	-1.18	3.84E-08
X73_RS00140	elongation factor Ts	PM1985	*tsf*	-1.67	5.38E-10
X73_RS00015	histidyl-tRNA synthetase	PM2011	*hisS*	-1.01	3.85E-06
X73_RS10260	SodA	PM0001	*sodA*	-1.82	5.19E-10

[a] Multiple genome entries exist for strain X-73; NCBI genome database, accession: NZ_CM001580.1 was chosen for this study. Differentially expressed open reading frames identified in this study that have not been assigned a locus tag in the genome entry NZ_CM001580.1 have been assigned a "b" or "c" suffix to the name of the gene tag (closest numerically) preceding it in the genome.
[b] The *P. multocida* sequence was used to interrogate the SwissProt database (excluding proteins belonging to *Pasteurellaceae*). Only proteins with a high similarity with known proteins/domains are assigned a predicted function i.e. an alignment score of ≥ 200 (graphical output = red bar).
[c] The Pm70 genome is available from NCBI genome database, accession: NC_002663.1 or AE004439.1.

Appendix D: Supplementary figures (Chapter 4)

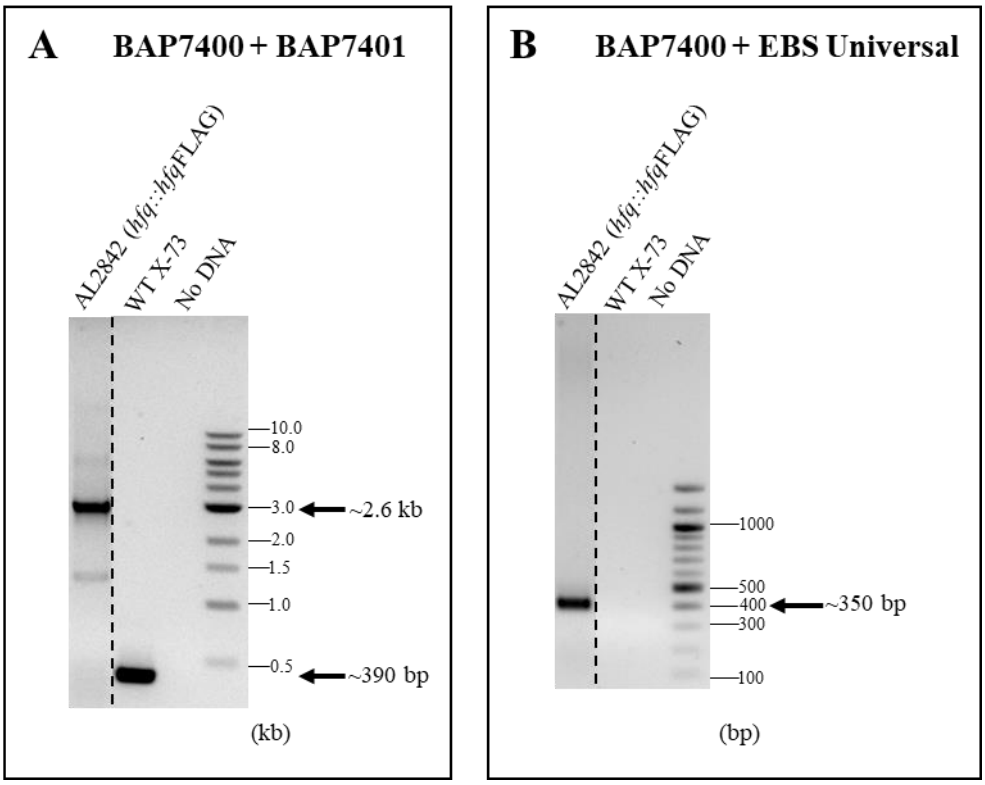


Figure S2. Electrophoretic separation of PCR products generated from analysis of the *hfq::hfq*FLAG strain AL2842 (expressing 3xFLAG-tagged Hfq). PCRs contained *hfq::hfq*FLAG (AL2842) genomic DNA, wild-type X-73 genomic DNA or no DNA (negative control). Primers BAP7400 and BAP7401 are *hfq*-specific forward and reverse primers, respectively, located either side of the recombination region (Figure 4.5C, Chapter 4) (**A**). BAP7400 is a *hfq*-specific forward primer located upstream of *hfq* and EBS Universal is an intron-specific reverse primer (**B**). Sizes of the NEB molecular weight markers are indicated to the right of the gels.

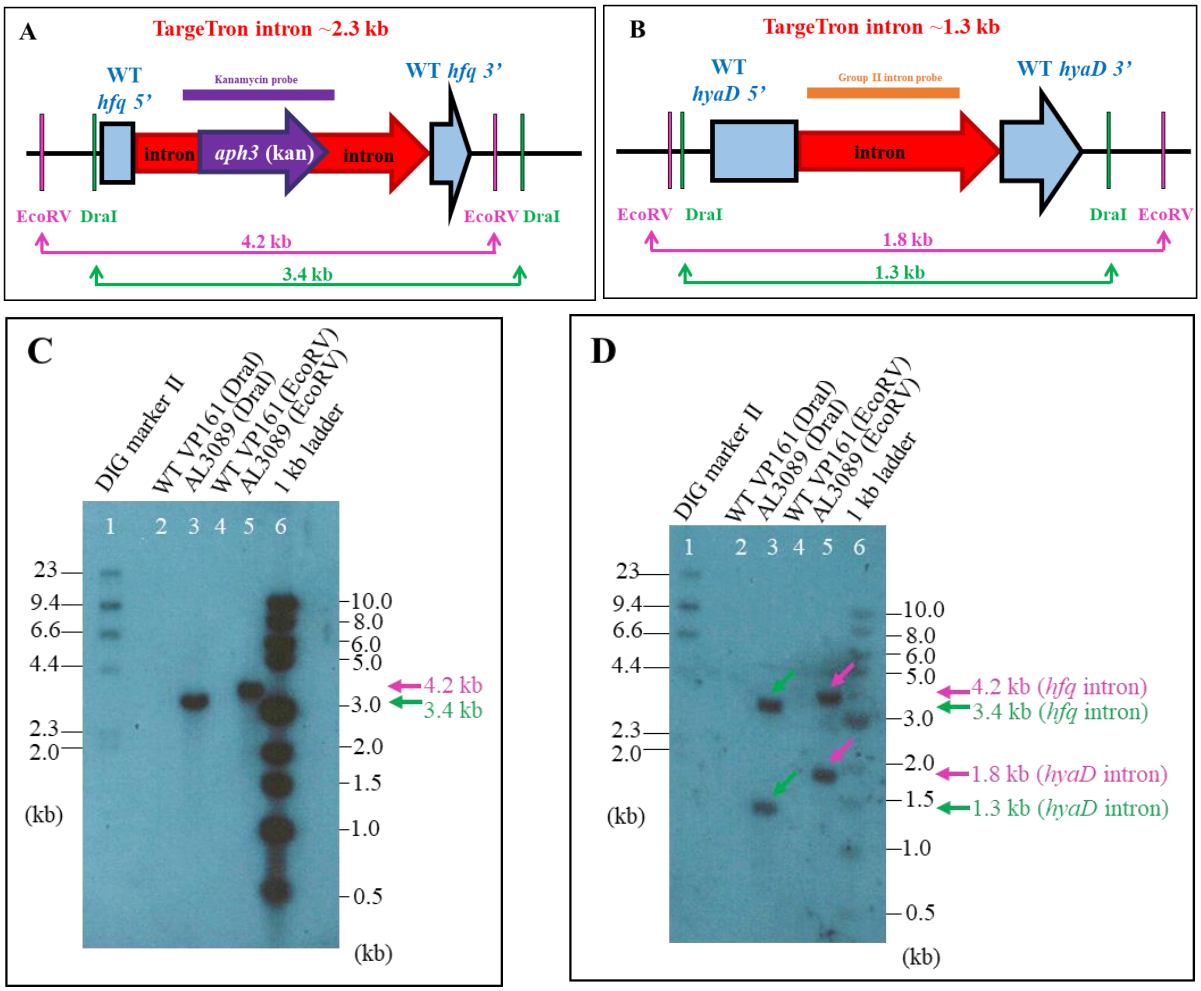


Figure S3. Southern hybridisation analysis of *hyaD/hfq* strain AL3089.
A. Schematic representation of the *hfq* region in AL3089 with intron (containing kanamycin cassette) insertion showing positions of EcoRV (pink) and DraI (green) restriction sites. The disrupted *hfq* gene is located on one EcoRV fragment (4.2 kb in size), or one DraI fragment (3.4 kb in size). The purple bar indicates the location at which the DIG-labelled kanamycin cassette probe will anneal to the TargeTron insertion.
B. Schematic representation of the *hyaD* region in AL3089 with intron insertion showing positions of EcoRV (pink) and DraI (green) restriction sites. The disrupted *hyaD* gene is located on one EcoRV fragment (1.8 kb in size), or one DraI fragment (1.3 kb in size). The orange bar indicates the location at which the DIG-labelled GroupII intron probe will anneal to the TargeTron insertion.
C and D. Southern hybridisation using either a DIG-labelled kanamycin cassette probe (**C**) or a DIG-labelled GroupII intron probe (**D**). Each blot contains EcoRV- and DraI-digested genomic DNA from *P. multocida* wild-type VP161 and *hyaD/hfq* (AL3089). Lane 1, DIG-labelled DNA molecular weight marker II; lane 2, DraI-digested wild-type VP161 genomic DNA; lane 3, DraI-digested AL3089 genomic DNA; lane 4, EcoRV-digested wild-type VP161 genomic DNA, lane 5, EcoRV-digested AL3089 genomic DNA; lane 6, NEB molecular weight marker. Fragment sizes are indicated to the left and right of the images, and arrows to the right indicate the sizes of hybridising fragments of EcoRV- and DraI-digested AL3089 genomic DNA. The Group II intron probe bound to two bands (**D**); the larger band represents the kanamycin cassette intron in *hfq* and the smaller band represents the markerless intron in *hyaD*.

Appendix E: Formulation and preparation of culture media and buffers

AE.1 Blocking solution for northern blot

- 10X DIG block buffer (Merck; Cat. No. 11 585 762 001) — 5 mL
- 1X Maleic acid buffer — 45 mL

AE.2 Coomassie Blue stain

- Coomassie Brilliant Blue R-250 — 2.0 g
- Glacial acetic acid — 200 mL
- Absolute EtOH — 800 mL

Make volume up to 2 L with ddH_2O.

AE.3 Destain solution

- Glacial acetic acid — 140 mL
- Absolute EtOH — 800 mL

Make volume up to 2 L with ddH_2O.

AE.4 Detection buffer for northern blot

- 1 M Tris (pH 8.0) — 10 mL
- 5 M NaCl — 2 mL
- MilliQ H_2O — 88 mL

pH to 9.5 with NaOH, sterile filter into RNase-free 50 mL tubes.

AE.5 Fixing solution

- Glycerol — 40 mL
- Absolute EtOH — 900 mL

Make volume up to 2 L with ddH_2O.

AE.6 Heart Infusion (HI) / Brain Heart Infusion Broth (BHI)

- Heart Infusion / Brain Heart Infusion 14.8 g

The media was made up to 400 mL with ddH_2O and autoclaved at 100 kPa and 121°C for 15 min.

For solid medium, an additional 6 g of agar was added prior to autoclaving. Once set, agar plates were stored at 4°C.

AE.7 Lysogeny (Luria) Broth (LB)

- Tryptone 4.0 g
- Yeast extract 2.0 g
- NaCl 2.0 g

The media was made up to 400 mL with ddH_2O and autoclaved at 100 kPa and 121°C for 15 min.

For solid medium, an additional 6 g of agar was added prior to autoclaving. Once set, agar plates were stored at 4°C.

AE.8 5X Maleic acid buffer

- 0.5 M Maleic acid 29.0 g
- 0.75 M NaCl 21.92 g

Add 500 mL MilliQ H_2O, pH to 7.5 with NaOH. Diluted 1:5 with ddH_2O for a working solution.

AE.9 5X PAGE Buffer

- Tris base 15 g
- Glycine 72 g
- SDS 5 g

The solution was made up to 1 L with ddH_2O and diluted 1:5 with ddH_2O for a working solution.

AE.10 Phosphate Buffered Saline (PBS)

- NaCl 8.0 g
- KCl 0.2 g
- Na_2HPO_4 1.44 g
- KH_2PO_4 0.24 g

The pH was adjusted to 7.4 with HCl. The solution was made up to 1 L with ddH_2O and autoclaved at 100 kPa and 121°C for 15 min.

AE.11 2X SDS-PAGE Sample buffer

- 0.5M Tris-HCl, pH 6.8 200 mM
- Glycerol 20%
- 10% (w/v) SDS 2.5%
- 2-mercaptoetanol 10%
- 1% Bromophenol blue 0.004% (w/v)

Buffer was prepared and stored at -20°C.

AE.12 10% Sodium dodecyl sulfate (SDS)

- SDS 20g

Add 180 mL MilliQ H_2O and mix; make volume up to 200 mL with MilliQ H_2O.

AE.13 20X Saline sodium citrate solution (SSC)

- Sodium chloride 350.6g
- Tri-sodium citrate 176.5g

Add 1800 mL MilliQ H_2O and pH to 7.0 with HCl; make volume up to 2 L and autoclave.

AE.14 2X SSC solution for northern blot

- 20X SSC 20.0 mL
- 10% SDS 2.0 mL
- MilliQ H_2O 178.0 mL

AE.15 0.1X SSC solution for northern blot

- 20X SSC 1.0 mL
- 10% SDS 2.0 mL
- MilliQ H_2O 197.0 mL

AE.16 10X Tris buffered saline (TBS), pH 7.4

- Tris base 121.2 g
- NaCl 175.4 g

Add 1800 mL MilliQ and pH to 7.4; make volume up to 2 L

AE.17 Washing buffer for northern blot

- 1X Maleic acid buffer 300 mL
- Tween 20 900 µL