Comparative transcriptomics and post-transcriptional regulation in *Campylobacter jejuni*

Vergleichende Transkriptomanalysen und posttranskriptionelle Regulierung in *Campylobacter jejuni*

Doctoral thesis for a doctoral degree at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Infection and Immunity

Submitted by

Gaurav Dugar

From Alipurduar, India

Würzburg, 2016

Reverse page

Submitted on: ……..

Office stamp

Members of the *Promotionskomitee*:

Chairperson: Prof. Dr. Thomas Dandekar

Primary Supervisor: Prof. Dr. Cynthia Sharma

Supervisor (Second): Prof. Dr. Jörg Vogel

Supervisor (Third): Dr. Tobias Ölschläger

Supervisor (Fourth): Dr. Martin Fraunholz

Date of Public Defence: ……

Date of Receipt of Certificates: ……..

"We are a way for the cosmos to know itself"

... Carl Sagan

Dedicated to my grandfather, Abhay Karan Doogar

and memory of a dear friend, Abhineet Bhandari …

Affidavit

I hereby confirm that my thesis entitled "Comparative transcriptomics and post-transcriptional regulation in *Campylobacter jejuni*" is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or material applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

Würzburg, Dec 16th 2016 _____

Gaurav Dugar

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Dissertation "Vergleichende Transkriptomanalysen und posttranskriptionelle Regulierung in *Campylobacter jejuni*" eigenständig, d.h. insbesondere ohne die Hilfe oder Unterstützung von kommerziellen Promotionsberatern, angefertigt habe. Ergänzend bestätige ich, dass ich keine anderen als die von mir angegebenen Quellen oder Hilfsmittel verwendet habe.

Ich erkläre außerdem, dass diese Dissertation weder in gleicher noch in ähnlicher Form bereits in einem Prüfungsverfahren vorgelegen hat.

Würzburg, den 16. Dez 2016 _____

Gaurav Dugar

SUMMARY

The transcriptome is defined as the set of all RNA molecules transcribed in a cell. These include protein-coding messenger RNAs (mRNAs) as well as non-coding RNAs, such as ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), and small non-coding RNAs (sRNAs). sRNAs are known to play an important role in regulating gene expression and virulence in pathogens. In this thesis, the transcriptome of the food-borne pathogen *Campylobacter jejuni* was characterized at single nucleotide resolution by use of next-generation sequencing approaches. The first genome of a *C. jejuni* strain was published in the year 2000. However, its transcriptome remained uncharacterized at large.

C. jejuni can survive in a variety of ecological niches and hosts. However, how strain-specific transcriptional changes contribute to such adaptation is not known. In this study, the global transcriptome maps of four closely related *C. jejuni* strains were defined using a differential RNA-seq (dRNA-seq) approach. This analysis also included a novel automated method to annotate the transcriptional start sites (TSS) at a genome-wide scale. Next, the transcriptomes of four strains were simultaneously mapped and compared by the use of a common coordinate system derived from whole-genome alignment, termed as SuperGenome. This approach helped to refine the promoter maps by comparison of TSS within strains. Most of the TSS were found to be conserved among all four strains, but some single-nucleotide-polymorphisms (SNPs) around promoter regions led to strain-specific transcriptional output. Most of these SNPs altered transcription only slightly, but some others led to a complete abrogation of transcription leading to differential molecular phenotypes. These in turn might help the strains to adapt to their specific host or microniche. The transcriptome also unveiled a plethora of sRNAs, some of which were conserved among the four strains while others were strain specific. Furthermore, a Cas9-dependent minimal type-II CRISPR-Cas system with only three Cas genes and multiple promoters to drive the transcription of the CRISPR locus was also characterized in *C. jejuni* using the dRNA-seq dataset.

Apart from sRNAs, the role of global RNA binding proteins (RBPs) is also unclear in *C. jejuni*. Aided by the global transcriptome data, the role of RBPs in post-transcriptional regulation of *C. jejuni* was studied at a global scale. Two of the most widely studied RNA binding proteins in bacteria are Hfq and CsrA. The RNA interactome of the translational regulator CsrA was defined using another global deep-sequencing technique that combines co-immunoprecipitation (coIP) with RNA sequencing (RIP-seq). Using this interactome dataset, the direct targets of this widespread global post-transcriptional regulator were defined, revealing a significant enrichment for mRNAs encoding genes involved in flagella biosynthesis. Unlike

Gammaproteobacteria, where sRNAs such as CsrB/C, antagonize CsrA activity, no sRNAs were enriched in the CsrA-coIP in *C. jejuni*, indicating absence of any sRNA antagonists and novel modes of CsrA activity regulation. Instead, the CsrA regulatory pathway revealed *flaA* mRNA, encoding the major flagellin, as a dual-function mRNA. *flaA* mRNA was the main target of CsrA but it also served to antagonize CsrA activity along with the protein antagonist FliW previously identified in the Gram-positive bacterium *Bacillus subtilis*. Furthermore, this regulatory mRNA was also shown in this thesis to localize to the poles of elongating *C. jejuni* cells in a translation-dependent manner. It was also shown that this localization is dependent on the CsrA-FliW regulon, which controls the translation of *flaA* mRNA. The role and mechanism of *flaA* mRNA localization or mRNA localization in general is not yet clear in bacteria when compared to their eukaryotic counterparts.

Overall, this study provides first insights into riboregulation of the bacterial pathogen *C. jejuni.* The work presented in this thesis unveils several novel modes of riboregulation in *C. jejuni,* which could be applicable more generally. Moreover, this study also lays out several unsolved intriguing questions, which may pave the way for interesting studies to come.

ZUSAMMENFASSUNG

Das Transkriptom ist definiert als die Summe aller RNA-Moleküle, die in einer Zelle transkribiert werden. Hierzu gehören sowohl protein-kodierende Boten-RNAs (mRNAs für „messenger RNAs“), als auch nicht-kodierende RNAs, wie ribosomale RNAs (rRNAs), transfer RNAs (tRNAs) und kleine nicht-kodierende RNAs (sRNAs für „small RNAs“). Diese sRNAs spielen eine wichtige Rolle in der Regulierung von Genexpression und Virulenz von Pathogenen. In der vorliegenden Arbeit wurde das Transkriptom des Lebensmittelkeims *Campylobacter jejuni* mit Hilfe von Next-Generation-Sequencing-Methoden charakterisiert, welche eine Auflösung des Transkriptoms auf Einzelnukleotid-Ebene ermöglichen. Obwohl eine erste Genomsequenz für *C. jejuni* bereits im Jahr 2000 veröffentlicht wurde, war das Transkriptom bisher größtenteils uncharakterisiert.

C. jejuni besitzt die Fähigkeit in vielen ökologischen Nischen und Wirten überleben zu können. Es ist jedoch bislang unbekannt, wie stammspezifische Veränderungen des Transkriptoms zu dieser Adaption beitragen. Mittels eines differenziellen RNA-Sequenzierungsansatzes wurden in dieser Arbeit globale Transkriptomkarten von vier nahverwandten *C. jejuni* Stämmen erstellt. Diese Analyse beinhaltet auch eine neue automatisierte Methode zur genomweiten Identifizierung von Transkriptionsstartstellen (TSS). Anschließend wurde aus den Genomsequenzen der vier *Campylobacter* Stämme ein SuperGenom erstellt. Dieses wiederum diente als Referenz, anhand dessen die Transkriptome kartiert und miteinander verglichen werden konnten. Dieser Ansatz ermöglichte eine verfeinerte Kartierung der Promotoren mittels des Vergleichs verschiedener Stämme. Die meisten TSS waren innerhalb der vier Stämme konserviert. Allerdings kam es durch SNPs („single-nucleotide polymorphisms“) in den Promoterregionen zu stammspezifischem Transkriptoutput. Die meisten dieser SNPs hatten nur geringe Veränderungen der Transkription zur Folge. Manche jedoch führten zu einem kompletten Verlust der Transkription und damit zu verschiedenen molekularen Phänotypen. Diese wiederum könnten es den verschiedenen Stämmen ermöglichen, sich an ihre spezifische Wirts- oder Mikronische anzupassen. Das Transkriptom wies auch eine Fülle von sRNAs auf, von denen manche in allen vier Stämmen konserviert, andere jedoch stammspezifisch waren. Zudem wurde mittels des *C. jejuni*-dRNA-seq-Datensatzes ein minimales Cas9-abhängiges CRISPR-Cas-System des Typs II entdeckt. Dieses beinhaltet lediglich drei Cas-Gene, jedoch mehrere Promotoren, die die Expression des CRISPR-Lokus antreiben.

Neben der Funktion von sRNAs ist auch die Rolle globaler RNA-Bindeproteine (RBPs) in *C. jejuni* weitestgehend unklar. Mithilfe der Transkriptomdaten wurde die Rolle von

RBPs in der posttranskriptionellen Regulierung in *C. jejuni* untersucht. Zwei der am besten untersuchten RNA-Bindeproteine in Bakterien sind Hfq und CsrA. Das RNA-Interaktom des Translationsregulators CsrA wurde mittels eines weiteren globalen Deep-Squencing-Ansatzes definiert. Bei dieser Methode werden Coimmunopräzipitation (coIP) und RNA-Sequenzierung zum so genannten RIP-seq kombiniert. Mithilfe dieses Interaktionsdatensatzes wurden die Zielgene dieses weitverbreiteten, globalen posttranskriptionellen Regulators definiert. Hierbei wurde eine signifikante Anreicherung von mRNAs, die in die Biosynthese von Flagellen involviert sind, erkennbar. Anders als in Gammaproteobakterien, in denen sRNAs wie CsrB und CsrC die CsrA-Aktivität antagonisieren, wurden in *C. jejuni* keine sRNAs in der CsrA-CoIP angereichert. Dies deutet auf das Fehlen jeglicher sRNA-Antagonisten, und damit auf eine neue Art der CsrA-Aktivitätskontrolle hin. Anstelle der sRNAs wurde die *flaA* mRNA, welche für das Hauptflagellin kodiert, als mRNA mit dualer Funktion identifiziert. Sie ist zum einen das Hauptzielgen von CsrA, fungiert aber gleichzeitig, zusammen mit dem Protein FliW, als Antagonist von CsrA. FliW wurde bereits zuvor in dem Grampositiven Bakterium *Bacillus subtilis* identifiziert. In dieser Arbeit konnte zudem gezeigt werden, dass die regulatorische *flaA* mRNA translationsabhängig an den Polen der wachsenden *C. jejuni*-Zellen lokalisiert ist. Außerdem war zu erkennen, dass diese Lokalisierung abhängig von dem CsrA-FliW-Regulon stattfindet, welches die Translation der *flaA*-mRNA kontrolliert. Im Gegensatz zu Eukaryoten ist die Rolle, die die Lokalisation der *flaA*-mRNA, oder bakterieller mRNA im Allgemeinen, spielt, sowie der Mechanismus, der zu dieser Lokalisierung führt, bisher noch unklar.

Zusammenfassend ermöglicht diese Arbeit einen ersten Einblick in die Riboregulierung des bakteriellen Pathogens *C. jejuni*. Es konnten einige neue Mechanismen dieser Art der Regulierung aufgedeckt werden, welche auch allgemeine Gültigkeit finden könnten. Zudem werden in dieser Arbeit neue, faszinierende Fragen aufgeworfen, die den Weg für weitere interessante Studien bereiten.

Table of Contents

1. AIM AND ORGANISATION OF THIS THESIS 1

2. INTRODUCTION 2

2.1. Transcriptome and RNA-based gene regulation in bacteria 2

2.2. RNA binding proteins in bacteria 3

2.3. RNA localization and its implications 5

2.4. *Campylobacter jejuni* as a model organism 7

3. RESULTS AND DISCUSSION 10

3.1. Comparative transcriptome of *C. jejuni* 10

3.1.1. dRNA-seq mapping to the SuperGenome 10

3.1.2. Comparative analysis of regulatory elements 18

3.1.3. SNPs lead to strain specific transcriptional output 22

3.1.4. The non-coding RNA repertoire of *C. jejuni* 24

3.1.5. The minimal CRISPR/Cas system of *Campylobacter jejuni* 28

3.1.6. Discussion 1 31

3.2. Exploring the translational regulator CsrA and its antagonists in *C. jejuni* 35

3.2.1. Genomic location and conservation analysis of *Campylobacter* CsrA 35

3.2.2. Transcriptome changes upon *csrA* deletion in *C. jejuni* 36

3.2.3. RIP-seq globally reveals direct mRNA targets of CsrA in *C. jejuni* 37

3.2.4. CsrA primarily binds mRNAs involved in flagella formation 39

3.2.5. CsrA binds diverse sites within target mRNAs 43

3.2.6. Validation of *C. jejuni* CsrA targets using an *E. coli* reporter system 43

3.2.7. Automated peak-detection reveals a CsrA binding motif 47

3.2.8. *flaA* mRNA is translationally repressed by CsrA 48

3.2.9. CsrA directly binds to *flaA* 5' UTR 50

3.2.10. The flagellar assembly factor FliW directly binds and antagonizes CsrA 52

3.2.11. Other flagellar mRNA targets of CsrA in *C. jejuni* 56

3.2.12. FliW and *flaA* mRNA titrate CsrA-mediated repression 58

3.2.13. *flaA* mRNA localizes to the poles of shorter *C. jejuni* cells 61

3.2.14. *flaA* mRNA localization is translation dependent 63

3.2.15. FliW affects *flaA* transcription 68

3.2.16. FlhX **–** In search for other factors involved in *flaA* mRNA localization 69

3.2.17. Discussion 2 72

4. CONCLUSION AND PERSPECTIVE 77

5. MATERIALS AND METHODS 80

5.1. Materials 80

5.2. Methods 89

6. APPENDIX 111

7. REFERENCES 195

8. LIST OF FIGURES 205

9. LIST OF TABLES 207

10. CONTRIBUTION BY OTHERS 208

11. CURRICULUM VITAE 209

12. ACKNOWLEDGEMENT 210

Abbreviation index

A	adenine
aa	amino acid
Amp	ampicillin
APS	ammonium persulfate
asRNA	antisense RNA
ATP	adenosine triphosphate
BSA	bovine serum albumin
C	cytosine
cDNA	complementary DNA
CDS	coding sequence
CFU	colony forming units
Cm	chloramphenicol
CO_2	carbon dioxide
CRISPR	clustered regularly interspaced short palindromic repeats
CTP	cytidine triphosphate
DMF	dimethyl formamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleotide
ds	double-stranded
DTT	dithiothreitol
EDTA	ethylenediamine tetraacetate
EtOH	ethanol
$Fe(II)SO_4$	iron (II)-sulfate
G	guanine
Gen	gentamicin
gDNA	genomic DNA
GFP	green fluorescent protein
H_2O	water (distilled)
H_2O_2	hydrogen peroxide
HCl	hydrochloric acid
Hyg	Hygromycin
IGB	Integrated Genome Browser
IGR	intergenic region
Kan	kanamycin
LPS	lipopolysaccharide
mRNA	messenger RNA
N_2	nitrogen
NaCl	sodium chloride
NaOH	sodium hydroxide
OD_{600}	optical density at a wavelength of 600 nm
OE	overexpression
ORF	open reading frame

P	promoter
PAA	polyacrylamide
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
P:C:I	phenol:chloroform:isoamyl alcohol
PCR	polymerase chain reaction
Phu	Phusion polymerase
RBP	RNA-binding protein
RBS	ribosome binding site
RNA	ribonucleic acid
RNase	ribonuclease
RNA-seq	RNA sequencing
RNS	reactive nitrogen species
ROS	reactive oxygen species
RT	room temperature
rRNA	ribosomal RNA
rNTP	ribonucleotide
sRNA	small regulatory/non-coding RNA
siRNA	short interfering RNA
SD	Shine-Dalgarno
SDS	sodiumdodecylsulfate
SNP	single nucleotide polymorphism
spp.	species
ss	single-standed
SSR	simple sequence repeat
Str	streptomycin
T	thymine
TAE	Tris/Acetate/EDTA
TBE	Tris/Borate/EDTA
TEMED	tetramethylethylendiamin
Tm	melting temperature
tmRNA	transfer-messanger RNA
Tris	tris-(hydroxymethyl)-aminomethan
tRNA	transfer RNA
TSS	transcriptional start site
U	uracil
UTP	uridine triphosphate
UTR	untranslated region
UV	ultraviolet
vol	volume
v/v	volume/volume
w/v	weight/volume in g/ml
WT	wildtype

Units

%	percent
° C	degree Celsius
A	ampere
Bp	base pair(s)
Ci	Curie
Da	Dalton
g	gram
h/hrs	hour(s)
l	liter
M	molar
min	minute(s)
molar	gram molecule
N	normality (volumetic)
nt	nucleotide(s)
pH	minus the decimal logarithm of the hydrogen concentration
rpm	rounds per minute
s/sec	second(s)
u	unit
V	Volt
W	Watt

Multiples

M	mega (10^{6})
k	kilo (10^{3})
c	centi (10^{-2})
m	milli (10^{-3})
µ	micro(10^{-6})
n	nano (10^{-9})
p	pico (10^{-12})

1. AIM AND ORGANISATION OF THIS THESIS

The study presented here aims to define the entire transcriptome of the human bacterial pathogen *Campylobacter jejuni* and to further explore the role of RNA binding proteins (RBPs) in its riboregulation. Firstly in Chapter 2 Introduction, this thesis provides a general background of RNA based regulation in bacteria along with recent developments in some other specialized topics such as RNA localization and flagella biosynthesis.

The second section of this thesis consists of results and discussion, which itself has been divided into two separate result parts, each containing its separate discussion at the end. The first part of this section defines the transcriptome maps of four *C. jejuni* strains using a differential RNA-seq approach and their subsequent comparison using a common coordinate system based on multiple aligned genomes. This comparison reveals how small changes or even single nucleotide polymorphisms (SNPs) on or near the promoter sequence can lead to drastic changes in the transcriptional output and may help the organism to adapt to their specific microniches. The second part of the section investigates the role of global RNA binding proteins in *C. jejuni.* Here, the role of the widespread bacterial post-transcriptional regulator, CsrA, is examined in detail. Another global deep sequencing technique, RIP-seq, is used to uncover the entire RNA interactome of CsrA. *C. jejuni* was found to lack the homologs of canonical CsrA antagonising sRNAs, such as CsrB/C, but instead revealed an mRNA, *flaA*, encoding the major flagellin as both the major target and antagonist of CsrA. Result and discussions are followed by a common conclusion and perspective section, which summarises the overall results and contemplates upon some of the novel and intriguing questions that arose during this study.

Most of the work described in this thesis has previously been published elsewhere:

- **Dugar G.**, Herbig A., Förstner K.U., Heidrich N., Reinhardt R., Nieselt K., and Sharma C.M. (2013) High resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates. **PLoS Genetics,** 9(5): p. e1003495.

- **Dugar G.**, Svensson S., Bischler T., Wäldchen S., Reinhardt R., Sauer M., Sharma C.M. (2016) The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in *Campylobacter jejuni*. **Nature Communications,** 7: p. 11667.

2. Introduction

Comparative Genomics approaches like BLAST have been used over the last decades to compare the genetic makeup of different strains or species [1]. It has allowed us to see how genes have evolved and diverged over the years to help adapt an organism to a new environment. These changes could be accounted for by acquisition or evolution of new genes performing novel functions or by improving upon the existing genes for efficient functioning. Some other genetic changes may not lead to changes in the genes themselves but to differential output in their expression. For example, an organism, which has evolved from an acidic environment to a relatively neutral environment may try to reduce the expression of genes acting upon acid stress without necessarily abrogating the response altogether. These changes at the transcriptional or post-transcriptional levels are much more subtle and can easily be overlooked.

Over the last decade, the arrival of RNA-seq has helped to analyse the transcriptome of an organism at single nucleotide resolution. It has allowed us to measure how changes in the environment or the genetic makeup of an organism can influence or alter the transcriptome of an organism. RNA-seq can be used in innovative ways to obtain novel information about the transcriptome. In its classical use, transcriptomes are compared between different environmental conditions or mutant sets to decipher the underlying genes or responses involved. The RNA samples can be pretreated using chemicals, enzymes or physical separation before sequencing to differentiate between different subsets of RNA like primary or processed transcripts. Moreover, RNAs bound to a particular protein or other RNAs can also be sequenced specifically to determine the part of the transcriptome associated with it. The plethora of information generated using such techniques can then be stitched together using classical molecular biology approaches to decipher novel pathways of gene regulation that would otherwise be hard to fathom using a rational approach based on causal-effect relationships.

2.1. Transcriptome and RNA-based gene regulation in bacteria

The blueprint of life, DNA, encodes for all the information to make an organism and passage of information to its progeny. Information from the DNA is gathered in the form of RNA via transcription. The term transcriptome describes the total set of RNA transcripts in a given organism. Most of the transcripts encode proteins, the building blocks of life, and are appropriately called messenger RNAs (mRNAs). The other transcripts, which do not encode proteins are called non-coding RNAs (ncRNAs). ncRNAs can carry out multiple functions ranging from vital to mere comforting. The vital functions provided by some of these ncRNAs are so

essential that they are conserved throughout the three kingdoms of life. These include ribosomal RNAs (rRNAs) and transfer RNAs (tRNAs), which are essential to decode or translate the mRNAs. On the other side of the spectrum are the more dynamic and locally conserved non coding RNAs called the small RNAs (sRNAs).

Bacteria live in a dynamic environment and have to constantly strive to survive. Post-transcriptional regulation provides an important layer of control to help adapt bacteria to such changing conditions. sRNAs can fine tune the rate of translation or protein synthesis to the needs of the organism through direct RNA-RNA interaction [2]. These sRNAs can be categorized into two classes based on the relative location to their targets. *cis*-encoded antisense sRNAs are encoded in the DNA strand opposite to their target and regulate the target mRNA via long stretch of complete complementarity [3]. However, the more extensively studied class of sRNAs is *trans*-encoded. These non coding RNAs are encoded separately from their targets and typically bind only with limited complementarity to their target mRNAs [4]. A typical trans-encoded sRNA can regulate several different mRNA targets by binding near their ribosome binding site (RBS) and thus, inhibiting translation initiation. Most of such trans-encoding sRNAs in Gram-negative bacteria require the protein Hfq as a chaperone for stabilization and proper function. One such global targeting sRNA is GcvB, which targets several members of the amino acid transport family [5, 6]. Most of such sRNAs repress translation but they may even lead to activation in certain cases by resolving inhibitory structures in the 5' untranslated region (5' UTR) of mRNAs, which sequester the RBS [7]. In some other cases, sRNAs may not bind the mRNAs but instead interact with other regulatory proteins to alter or sequester their activity. For instance, 6S RNA mimics an open promoter structure, which can bind and antagonize RNA polymerase [8]. In an increasing number of examples, sRNAs or even mRNAs are also known to act as a sponge by binding other sRNAs to regulate their activity. Recently, it was shown how a stable RNA (SroC) fragment derived from the mRNA decay of an amino acid transporter encoding locus could in turn bind and degrade the GcvB sRNA [9].

2.2. RNA binding proteins in bacteria

Over the recent years, a lot has been investigated on the role of sRNAs in post-transcriptional regulation. However, some RNA binding proteins are also involved in this regulatory pathway. The two most widely studied RNA binding proteins in bacteria are Hfq and CsrA [10]. Similar to sRNAs, RNA binding proteins can also function as post-transcriptional regulators using similar means. They may just bind near the RBS to repress translation or recruit other proteins such as RNases to alter the mRNA turnover. RNA binding proteins like Hfq can also stabilize the interaction between an mRNA and an sRNA to exert or promote the regulatory effect of the sRNA. Hfq has a ring like

doughnut structure composed of six monomers, which provides two RNA binding surfaces, capable of a wide range of interaction with multiple RNAs and thus leading to a complex regulation at the molecular level [11]. The current state of sequencing technology has made it possible to study the RNA interactome of RNA binding proteins to unveil such complexity at a genome-wide scale [12].

The other widespread global RNA binding protein in bacteria belongs to the CsrA/RsmA (carbon storage regulator/regulatory of secondary metabolism) family. About 75 % of all sequenced bacteria carry a homolog of the central protein regulator of this system [13, 14]. The bacterial Csr/Rsm network is a post-transcriptional regulatory system, which controls global physiological phenomena such as central carbon metabolism, motility, biofilm formation, quorum sensing, or virulence [15, 16]. CsrA acts as a translational repressor by binding to 5' regions (either 5' UTRs or the first few codons) of target mRNAs, thereby inhibiting translation initiation (Fig. 1) [15]. Examples of positive regulation have also been identified, where CsrA protects target mRNAs from degradation [17, 18]. CsrA binds target RNAs as a dimer with strong sequence- and structure-dependent specificity for GGA-rich motifs that are located in the loop of a hairpin structure and often overlap with the Shine-Dalgarno (SD) sequence [19].

In Gammaproteobacteria, the activity of CsrA itself is post-translationally regulated through the CsrB/C and RsmX/Y/Z families of small RNAs (sRNAs) [3, 15, 19]. These antagonizing sRNAs harbor varying numbers of stem-loops with GGA motifs with high affinity for CsrA/RsmA that sequester it like a sponge, thereby inhibiting its regulatory action (Fig. 1) [15, 19, 20]. They can differ in copy number, are often induced by environmental signals, and can also be connected to the regulon of the RNA chaperone, Hfq [21]. Some CsrA networks also employ the CsrD protein, which primes CsrB/C for degradation by RNaseE [22]. In the Gram-positive bacterium *Bacillus subtilis*, the flagellum assembly factor FliW directly binds to CsrA and acts as its protein antagonist [23]. FliW homologs are relatively widespread [23], indicating that protein-mediated regulation of CsrA might be more common than previously thought.

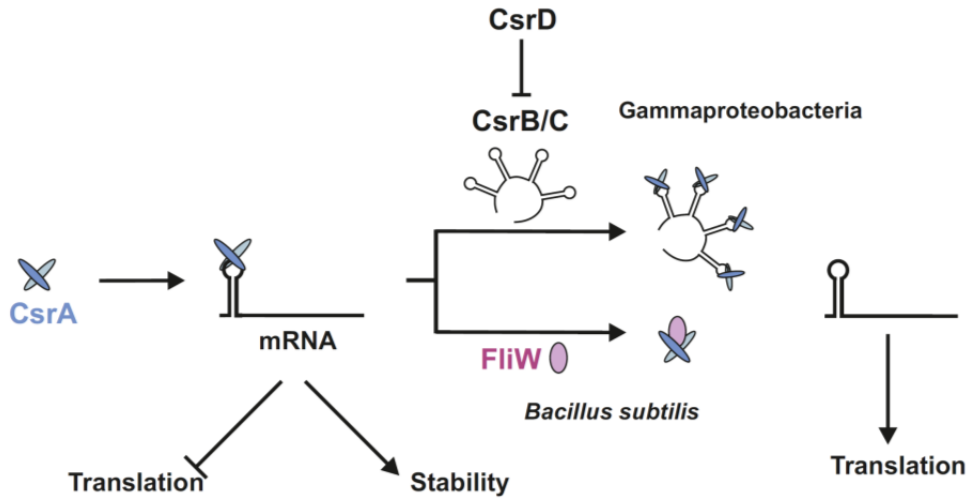


Figure 1. The current model of CsrA based regulation. mRNA targeting by CsrA along with its interplay with sRNA (CsrB/C) or protein (FliW) antagonists.

While several aspects of the regulatory circuits of CsrA have been characterized in *E. coli* and other Gammaproteobacteria, both the mRNA targets and the CsrA auto-regulation are poorly characterized outside these bacteria. Recently, it was also found that CsrA binding could also resolve an inhibitory structure for the entry site of the transcription termination factor Rho, resulting in premature transcription termination [24]. Several paralogs of CsrA/RsmA-like regulatory proteins such as CsrT, CsrR and RsmF have also been described recently [25-27]. It would be important to see if such members work alongside CsrA/RsmA or whether they have independent regulatory roles.

Along with Hfq and CsrA, several other RNA binding proteins have also been described. In *Salmonella enterica*, ProQ was recently identified as a major small RNA-binding protein by use of a gradient profiling based sequencing approach termed Grad-seq [28]. ProQ has a wide range of targets and was further shown to affect global gene expression and stability of associated sRNAs [28, 29]. In *Shigella sonnei* it was shown that the protein RodZ, which is associated with the bacterial actin-homologue MreB at the membrane, binds specifically to the mRNA of the virulence factor EnvA [30]. This raises the possibility that RNA binding proteins might also lead to localized regulation or may even lead to localization of their associated mRNA.

2.3. RNA localization and its implications

Proteins can be visualized directly by fusing them to a fluorescent protein or by immunofluorescence microscopy using labeled antibodies after fixation. Unlike protein localization, studies on RNA localization are quite challenging due to the difficulty of coupling fluorescent proteins or dyes to RNA molecules in a biological setting. Compared to protein levels,

levels of the respective mRNAs are extremely low in general, which makes it much more difficult to obtain a strong cumulative signal for the mRNA of interest, especially in prokaryotes [31]. Despite these challenges, significant progress has been made over the last two decades in visually analyzing RNA molecules first in eukaryotes, and more recently also in prokaryotes. One of the widely used method for tracking RNA in live cells requires tagging of the RNA of interest with an array of RNA aptamers, which in turn bind to a protein (*e.g.* MS2 binding aptamers bind to MS2 coat protein) that is fused to another fluorescent protein [32]. Another way to localize RNA is based on detection of multiple complementary fluorescently labeled oligonucleotides in fixed cells. This method is termed RNA FISH (Fluorescence *in situ* hybridization) [33].

In eukaryotes, certain mRNAs are localized to subcellular compartments to regulate gene expression with temporal and spatial control. Previously, only a few specialized mRNAs were thought to be localized, until an extensive study in *Drosophila* showed that a high proportion of expressed genes are in fact subcellularly localized [34]. There are several foreseen advantages of mRNA localization. Firstly, it allows gene expression to be spatially confined to the site of protein requirement. Once localized, an mRNA can undergo multiple rounds of translation, which is probably much more economical for the cell than localizing individual proteins after their synthesis. The mRNA localization in addition may also lead to quicker action in response to an external signal or stimuli [35].

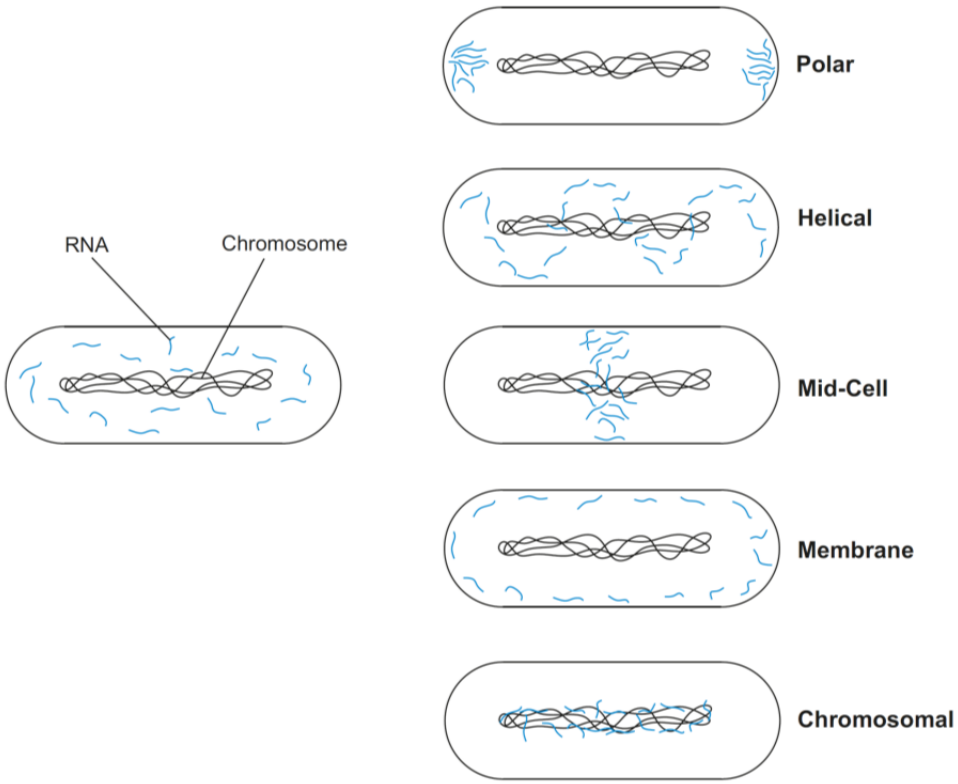


Figure 2. Localization of RNAs in bacteria. RNAs were perceived to be randomly diffused in bacteria (*Left panel*), but this picture has changed considerably in favor of defined RNA localization (*Right panel*). The figure has been adapted from Buskila *et. al.,* 2014 [36].

Compared to a eukaryotic cell, prokaryotes have roughly a 1000-fold smaller volume (*E. coli* vs. HeLa cells) [37] and poorly defined subcellular structures, which argues against the need for mRNA localization. Over the last decade, this view has been challenged and changed considerably owing to recent developments in RNA localization and microscopy techniques. One type of mRNA localization assumed to be common in bacteria is the accumulation of mRNA molecules encoding for certain membrane or secreted proteins close to the cytoplasmic membrane (Fig. 2). This is a consequence of co-translational protein translocation to the membrane translocons by targeting of translating ribosomes, a process mediated by the Signal Recognition Particle (SRP) [38]. Two more recent models have been presented for localization of RNA in bacteria. The first model by the Jacobs-Wagner group suggests that the RNA transcripts remain near their site of transcription and hence, the bacterial chromosome itself acts as a template for RNA localization [39] (Fig. 2). The second model presented by the Amster-Choder group suggests that the bacterial mRNAs localize to the sites where their encoded proteins are required [40]. The localization in the second model seems to be mediated by *cis*-encoded elements in the mRNA itself. These two models have been shown to work only for a very limited number of mRNAs and hence proposed to co-exist [36]. By using super-resolution microscopy with RNA FISH it was recently shown that mRNAs encoding for inner-membrane proteins are enriched at the membrane [41]. This enrichment was dependent on co-translational recognition of the signal-peptide and lead to selective degradation of these mRNAs due to their co-localization with the RNA degradosome. Overall, it is unclear how widespread mRNA localization is in bacterial cells and there are conflicting ideas on what mechanisms are responsible for mRNA localization. It is also unknown if RNA binding proteins play a role in this process in bacteria as they are pivotal for mRNA localization in eukaryotes.

2.4. *Campylobacter jejuni* as a model organism

Camyplobacter jejuni is a Gram-negative spiral-shaped Epsilonproteobacterium commonly associated with poultry or wild birds. It is commensal to birds and cattles and found in their digestive tracts and feces. It is regarded as one of the leading causes of human gastroenteritis in the world and is also responsible for development of secondary autoimmune disorders such as the Guillain-Barré or Miller-Fisher syndromes. It is a microaerophilic, food-borne pathogen that can survive in various stress conditions imposed by the environment and the host. Uncooked chicken, milk and contaminated water are the most common sources for infection with *C. jejuni*. The different outcome of *C. jejuni* infection in humans when compared to poultry is not properly

understood due to lack of appropriate infection models [42] . The whole genome of *C. jejuni* strain NCTC11168 was published in the year 2000 [43]. The sequencing revealed a small (1.64 Mb), AT-rich (69 %) and densely packed genome with more than 1600 genes. The genome of *C. jejuni* encodes for only three sigma factors and a very small number of transcriptional regulators. Extensive strain to strain variability is also observed among the *C. jejuni* isolates, which in part can be associated with its natural competence and genetic exchange among themselves and their environment. The genome also revealed stretches of homopolymeric tracts within hypervariable regions (Fig. 3). These hypervariable regions have hotspots within the capsule, lipooligosaccharide (LOS) and flagellum encoding regions, which often leads to phase variation in these structures by means of frameshifts and point mutations [43]. The high rate of variation in these outer structures helps in rapid adaptation of *C. jejuni* to evade the host defense or avoid bacteriophage infection [44, 45]. Such phase variation has also been frequently observed in genes encoding restriction-modification enzymes, which can lead to global changes in methylation pattern and gene expression [46]. It is currently not clear how phase variation directly or indirectly leads to changes in the transcriptional output. The *C. jejuni* genome was also found to lack many of the common metabolic pathways for the use of widely available carbohydrates such as glucose. However, it can utilize various amino acids and intermediates from the citric acid cycle as carbon source [47].

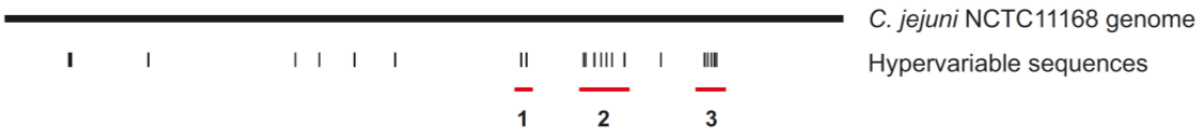


Figure 3. Hypervariable regions in the genome of *C. jejuni* NCTC11168. Three hypervariable regions representing 1) Lipooligosaccharide, 2) Flagellar modification and 3) Surface polysaccharide gene clusters are marked along the *C. jejuni* genome in red bars. The figure has been adapted from Parkhill *et. al.,* 2000 [43].

C. jejuni also lacks many of the classical virulence factors except Cytolethal Distending Toxins (CDTs). Therefore, motility is considered an important virulence trait of this pathogen. *C. jejuni* produces amphitrichous flagella *i.e.* a single flagellum at each pole. The flagellum is not just used for motility but is also an important virulence determinant factor in certain species. For example, it can secrete several virulence factors called *Campylobacter* invasion antigen (Cia) proteins. The exact function of these secretory proteins remains elusive, but they have been shown to affect the internalization of *C. jejuni* into host cells [48].

Despite the availability of whole genome sequences, the knowledge about post-transcriptional regulation and the role of RNA-binding proteins is still limited in *Campylobacter*.

C. jejuni lacks the classic sRNA chaperone Hfq, but encodes a homolog of the global post-transcriptional regulator CsrA. Deletion of *csrA* often leads to pleiotropic effects, including virulence defects in many pathogenic species [13, 14]. In *C. jejuni*, CsrA is required for motility, biofilm formation, oxidative stress response, and adherence and invasion to host cells [49]. However, its direct RNA targets in *C. jejuni* are unknown. Also, no antagonizing small RNAs like CsrB/C have been reported neither in *Helicobacter* nor *Campylobacter* despite some global transcriptome studies to identify sRNAs in these bacteria [50-53], raising the question of how CsrA itself is regulated in these bacteria. Interestingly, *C. jejuni* encodes a homolog of a flagellar assembly factor, FliW, which is known to act as a protein antagonist of CsrA in *B. subtilis* and hence may also repress CsrA activity in *C. jejuni* [23].

3. RESULTS

3.1. Comparative transcriptome of *C. jejuni*

Comparative genome analysis has revealed major strain to strain variability among several *Campylobacter* species [54]. The major species specific differences among *Campylobacter* strains are observed within gene clusters encoding for external cellular structures. These structural differences might help the bacteria in their adaptation to a specific host or environmental niche [54, 55]. Since the comparison among multiple species or strains has primarily been limited to the genomic level, the differences in transcriptome structure which may lead to further phenotypic differences, remains currently unknown. To study and define the transcriptome and TSS map of *C. jejuni,* the widely used differential RNA-seq (dRNA-seq) approach was applied [52]. dRNA–seq is based on differential treatment of total RNA sample by the enzyme terminator exonuclease, which leads to depletion of processed RNAs and hence a relative enrichment of primary transcripts. This relative enrichment in the treated library compared to an untreated library can be used to assign TSS on a genome-wide scale [56].

3.1.1. Mapping of dRNA-seq reads to the SuperGenome

Until recently, transcriptome studies have been limited to a single bacterial strain. A comparative analysis of the transcriptome can be used to measure the relative expression a gene at different conditions or of different strains or mutants. A comparative transcriptome based on multiple strains also has the additional advantage of assigning conserved TSS and transcript boundaries with a higher confidence. Here, the dRNA-seq approach was applied to four different strains of *C. jejuni*. The phylogenetic relationship between these strains and their close relatives is shown in Fig. 4.

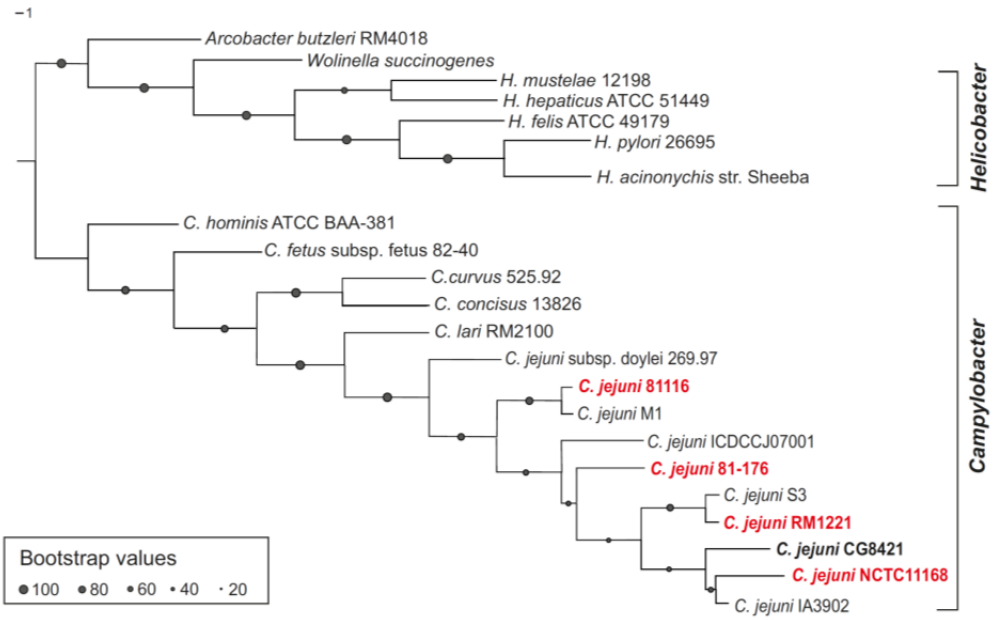


Figure 4. Phylogenetic tree of representative members of the Epsilonproteobacterial lineage. The phylogenetic tree was calculated based on concatenated sequence alignments of 12 highly conserved proteins (Ffh, FusA, GyrB, InfB, LepA, PyrG, RplB, RpsE, RpsH, RpsK, TopA, and Tuf). The four strains analyzed in this study are highlighted in red.

Strain NCTC11168 [43] was originally isolated from a case of human enteritis in the UK and was the first *Campylobacter* strain for which the whole genome was sequenced. However, this strain displays rather poor virulence. *C. jejuni* 81–176 [57], isolated from a diarrheal outbreak in the U.S., is highly pathogenic and carries two large plasmids, pVir [58] and pTet [59]. Strain 81116 [60], a human isolate from a waterborne outbreak in the UK, is a genetically stable and lab-adapted strain, which is infective for chickens [61]. The virulence potential of the chicken isolate RM1221 is not currently known in humans [54]. Several additional features of these strains are compiled in Table 1.

Table 1. Characteristics of *Campylobacter jejuni* strains used for dRNA-seq.

Feature	**NCTC11168**	**81-176**	**81116**	**RM1221**
Source	Human	Human	Human	Chicken
Country	UK	USA	UK	USA
Serotype	HS:2	HS:23,36	HS:6	HS:53
Chromosome size [Mbp]	1.64	1.62	1.63	1.78
Number of ORFs[a]	1,623	1,653	1,626	1,838
G+C content [%]	30.6	30.5	30.5	30.3
Plasmids	-	pVir, pTet	-	-
Phage/ genomic island elements		6-kb integrated element (plasmid)		CMLP1/ CJIE1 CJIE2 CJIE3 CJIE4
CRISPR/Repeats	yes/5	no	yes/8	yes/4
Genome Reference	[43, 62]	[57]	[60]	[54]

[a]Number of ORFs according to NCBI annotations (May 2012).

All four strains were grown in triplicates and samples were collected for RNA isolation at mid-exponential growth phase. RNA was isolated from two sets while the third sample was stored indefinitely at -80 °C. dRNA-seq libraries were constructed using the two replicate RNA samples from each strain. 2.3 to 5.5 Mio cDNA reads were obtained per library, and were subsequently mapped to their respective genomes. The mapping statistics for the all the dRNA-seq library have been compiled in Table 2. *C. jejuni* 81-176 also harbors two plasmids pTet and pVir and cDNA reads from the 81-176 library was also mapped to them.

A characteristic enrichment of cDNA reads was observed at the TSS in the TEX treated sample compared to the untreated library, thereby validating the dRNA-seq approach used here. These enrichment patterns allowed for determination of TSS in the four strains, and in most cases an enrichment of cDNA reads at a given TSS was observed in all strains (Fig. 5A, *rpsL*). Moreover, the comparative dRNA–seq data showed that homologous genes might share the same TSS, even if the promoter sequence of the regions is not conserved (Fig. 5B). Thus, based on sequence information alone it can be unclear whether a gene is expressed in the individual strains, whereas the comparative dRNA–seq data can provide this information. However, for several homologous genes, a clear TSS enrichment pattern (more than twofold in TEX+ vs. TEX–) was observed in only some of the strains despite highly conserved promoter regions. A clear enrichment in the TEX+ libraries was observed at the TSS of *rpsL* in all four strains. In contrast, a clear enrichment at the TSS of homologs of Cj1380 was only observed for CJE1571 and a weak enrichment for Cj1380 and CJJ81176_1382. The cDNA coverage at the TSS of C8J_1298 is even higher in the TEX- library compared to the TEX+ library, which is rather indicative of a processing site (Fig. 5A). Nevertheless, a sharp flank in the cDNA distribution and conservation of promoter region in all four strains indicates that transcription starts at exactly the same position (Fig. 5A and C). Without the comparative information from the other strains such cases could not be unambiguously defined as a TSS. Overall, a combination of sequence conservation and comparative TSS enrichment pattern analyses can be used to refine global TSS maps.

Table 2. Mapping statistics of *Campylobacter jejuni* dRNA-seq libraries. The table indicates the total number of sequenced cDNA reads considered in the analysis, the number of reads that were removed due to insufficient length (<12 nt) after poly(A)-tail clipping (before read mapping), the number of reads that were successfully mapped to the reference genomes or the pVir and pTet plasmids of strain 81-176 using *segemehl* (see Materials and Methods), the number of mappings (i.e. some reads map to different locations with the same score), and the number of uniquely mapped reads. For the number of mapped reads and number of uniquely mapped reads, the percentage values (relative to the total number of reads) are also listed.

	NCTC11168 R1 -TEX	NCTC11168 R1 +TEX	NCTC11168 R2 -TEX	NCTC11168 R2 +TEX	RM1221 R1 -TEX	RM1221 R1 +TEX	RM1221 R2 -TEX	RM1221 R2 +TEX
Total number of reads	2531653	3646875	2823393	4277090	5541367	3796307	4930476	3819843
Failed size filter after clipping	35426	108068	27541	167095	146201	38848	121493	83806
Total number of mapped reads	2463201	3373017	2749401	4014305	5344841	3696410	4742111	3679393
Total number of mappings	4778326	6051548	5642027	8734627	12716083	7651006	11133996	8154828
Uniquely mapped reads	1297582	2027267	1284067	1589563	1647729	1706330	1535896	1420358
% mappable reads	97.3	92.5	97.4	93.9	96.5	97.4	96.2	96.3
% of uniquely mapped reads	51.3	55.6	45.5	37.2	29.7	44.9	31.2	37.2
	81116 R1 -TEX	81116 R1 +TEX	81116 R2 -TEX	81116 R2 +TEX	81-176 R1 -TEX	81-176 R1 +TEX	81-176 R2 -TEX	81-176 R2 +TEX
Total number of reads	4559972	3906174	3913148	4863622	4806372	2315766	3857841	3560672
Failed size filter after clipping	119288	25786	81165	61659	41887	27390	31904	41424
Total number of mapped reads	4393451	3775032	3797781	4698430	4709413	2242991	3785770	3441085
Total number of mappings	9702783	7072857	8724948	9624104	10844930	4066371	8640160	7547467
Uniquely mapped reads	1728726	2119356	1326766	2207903	1610393	1303871	1334936	1331049
% mappable reads	96.3	97	97	97	98.0	96.9	98.1	96.6
% of uniquely mapped reads	37.9	54	34	45	33.5	56.3	34.6	37.4
Mapped reads in plasmid pVir (NC_008770)					26271	30806	21861	45196
Mapped reads in chromosome (NC_008787)					4592849	2063904	3688606	3227514
Mapped reads in plasmid pTet (NC_008790)					90293	148281	75303	168375
Mappings in plasmid pVir (NC_008770)					27054	31433	22532	45902
Mappings in chromosome (NC_008787)					10727172	3886321	8542010	7332612
Mappings in plasmid pTet (NC_008790)					90704	148617	75618	168953

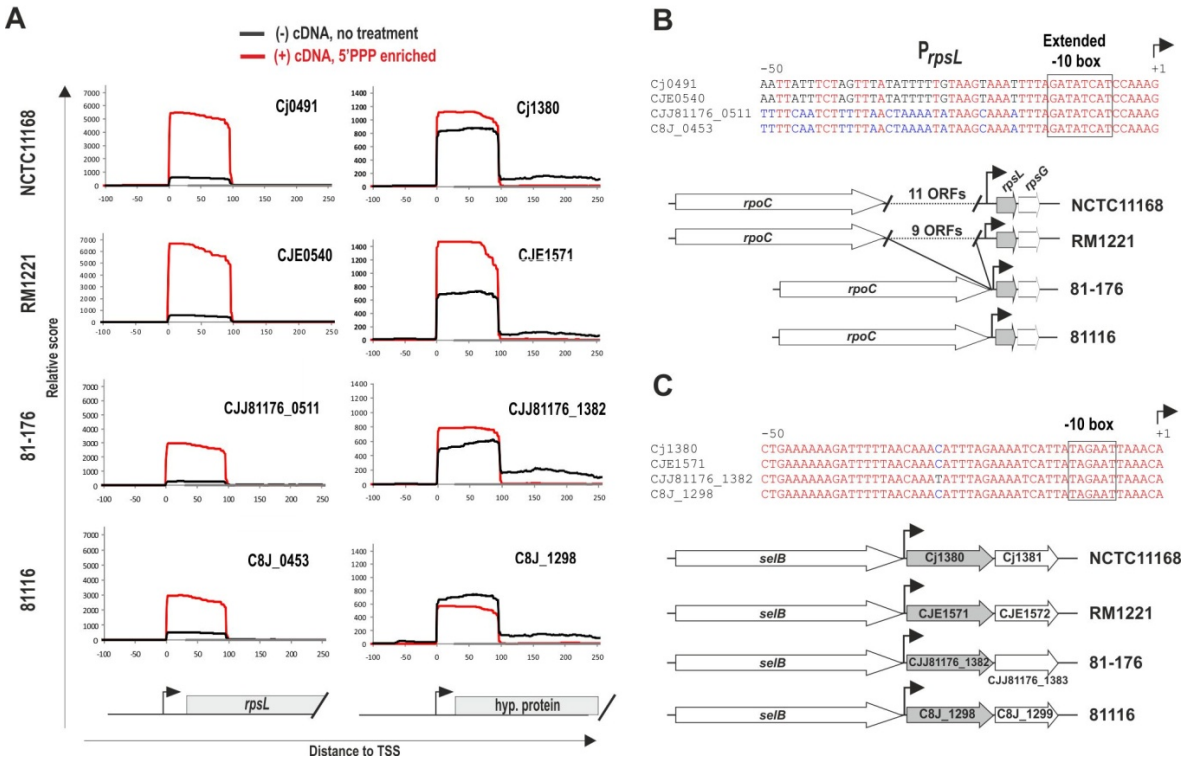


Figure 5. Comparative annotation of transcriptional start sites (TSS) based on differential RNA-seq. (A) Schematic drawing of cDNA enrichment patterns at primary 5' ends of *rpsL* mRNAs, encoding ribosomal protein S12 (left), and mRNAs of Cj1380 homologs, encoding a gene of unknown function (right). Exonuclease treatment (TEX+, red curve) shifts the cDNAs towards the nuclease-protected 5'-end, yielding a sawtooth-like profile with an elevated sharp 5' flank that corresponds to the TSS. **(B)** Nucleotide alignment of the *rpsL* promoter (-50 to +1 according to the TSS) and genomic location of *rpsL* in the four analyzed *C. jejuni* strains. Based on sequence conservation solely, it would be unclear whether the *rpsL* promoter is functional in strains NCTC11168 and RM1221, since 11 and 9 ORFs, respectively, are inserted directly upstream of its -10 box and might disrupt the promoter. However, the enrichment patterns in the dRNA-seq data on the left indicate active promoters in all four strains. **(C)** Nucleotide alignment of the promoter regions and genomic location of Cj1380 homologs in the four analyzed *C. jejuni* strains.

To facilitate TSS comparisons among the four strains, the dRNA-seq data sets were mapped to a common coordinate system (or so called SuperGenome). The SuperGenome was derived from multiple alignments of genomes from the four strains, which thus allowed for a simultaneous visualization of enrichment pattern (+TEX vs. -TEX) in the dRNA-seq dataset of all the four strains in a genome browser (Fig. 6A). The comparative dRNA–seq data show that the majority of TSS are enriched and detected in all strains (Fig. 6B, black arrows), but that there are also differences in transcriptional output among strains. For example, a TSS is observed within *kpsM*, encoding for one of the capsule export genes is only observed in two (81-176 and RM1221) of the four strains, and a TSS upstream of homologs of Cj1456c is only detected in strain 81116 (Fig. 6B, red arrow and blue arrow, respectively).

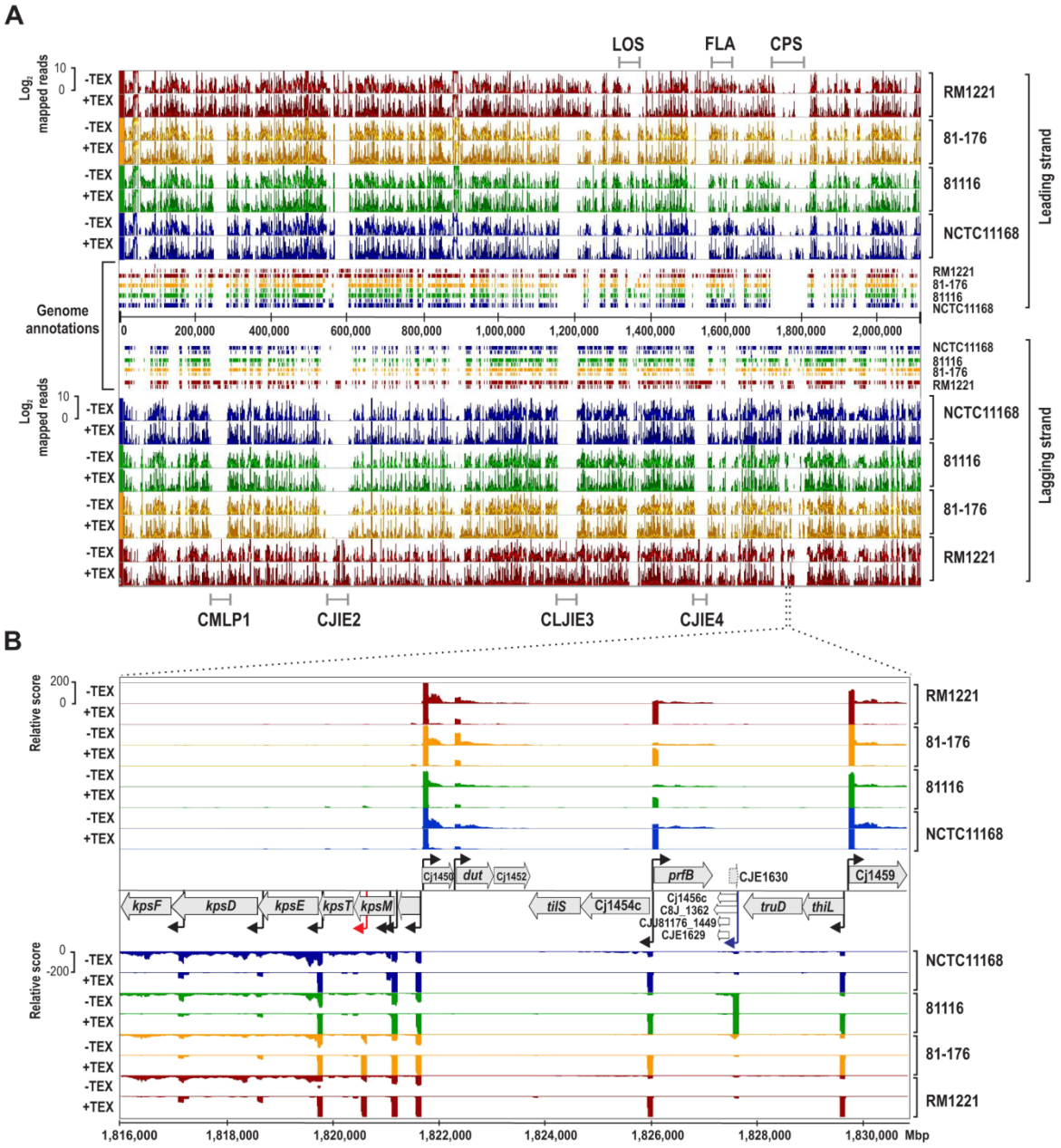


Figure 6. Differential RNA–seq and SuperGenome-based annotation of transcriptional start sites. (A) cDNA reads of −/+TEX libraries for the four *C. jejuni* strains were mapped to the SuperGenome, generated by whole genome alignment. All y-axes were set to the same scale. Horizontal bars at the top and bottom indicate the highly variable LOS (lipooligosaccharide modification), FLA (flagellin modification), and CPS (capsule) gene loci as well as the integrated genome elements (CMLP1, CJIE2, CJIE3, and CJIE4) for strain RM1221. **(B)** Example region in the SuperGenome with mapped dRNA–seq reads of −/+TEX libraries of the four strains. TSS are indicated by black arrows. Red and blue arrows indicate an internal TSS within *kpsM*, which is missing in two of the strains, and a primary TSS which was only detected in strain 81116, respectively. All y-axes were set to the same scale, which reflects a relative expression score. Note that gene CJE1630 (dotted arrow) is only annotated in RM1221 and that homologs of Cj1456c (white arrows) have different lengths in the four strains.

To automatically annotate TSS in a comparative manner, a two-step algorithm was employed: 1) TSS were detected independently for each strain based on dRNA–seq enrichment patterns and 2)

TSS were mapped to the SuperGenome to allow for comparison and assignment of TSS among strains. Subsequently, all TSS of the individual strains were automatically classified as primary TSS (pTSS; main promoter of a gene) or secondary TSS (sTSS; alternative promoter upstream of pTSS of a gene), internal TSS (iTSS; promoter within gene), antisense TSS (asTSS; promoter antisense to a gene +/−100 nts) or orphan TSS (no association with annotation) according to their location relative to annotated genes (Fig. 7A, Appendix Table 1). Around 2000 TSS were detected in each strain which also included ∼300 strain-specific TSS (Fig. 7B and Table 3). The comparative approach enabled the annotation of a total of 3,377 TSS positions in the SuperGenome. 1,035 of the TSS were detected (but not necessarily enriched) in all four strains (Fig. 8A and Table 3). The overall higher number of 2,167 TSS and 450 strain-specific TSS in RM1221 were due to the presence of four extra prophage elements (CMLP1, CJIE2-4), which are absent in the other three *C. jejuni* genomes. 70 and 58 TSS were detected on the two large plasmids, pVir (∼35 kB) and pTet (∼45 kB), respectively of *C. jejuni* strain 81–176.

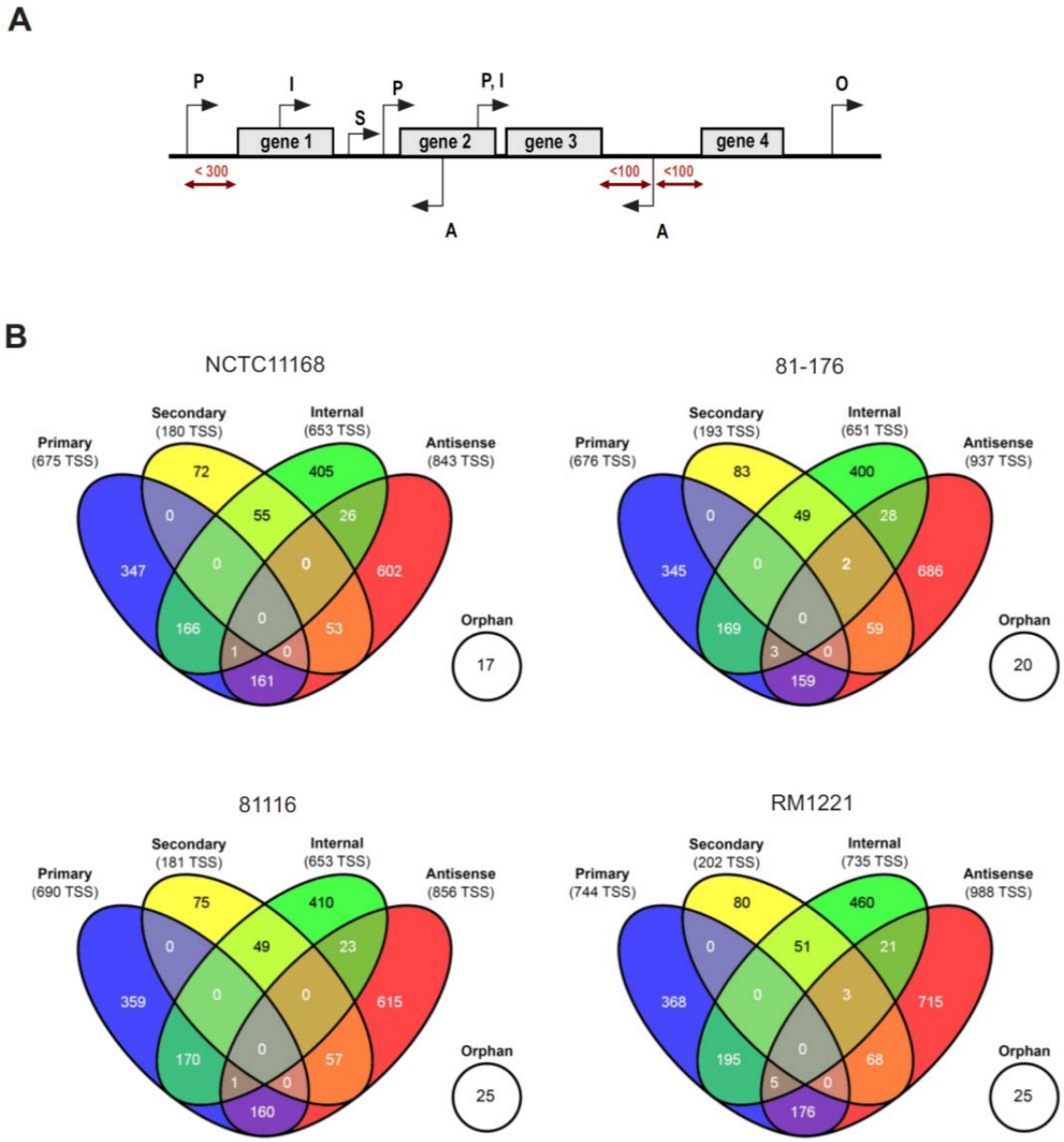


Figure 7. TSS classifications for the four *C. jejuni* strains. (A) Representation of TSS classifications based on expression strength and genomic context: primary (P), secondary (S), internal (I), antisense (A), or orphan (O). Cut-offs for the distances to flanking genes are indicated above the red arrows. **(B)** The Venn diagrams indicate the overlap between TSS classes for the four individual strains. Many TSS are assigned to more than one class. Venn diagrams were generated using VENNY (http://bioinfogp.cnb.csic.es).

Around half of TSS detected were conserved in all four strains and 60% of them were classified as either pTSS or sTSS, which mostly drives the transcription of mRNA genes. In contrast, only 21–31% of the strain-specific TSS (only present in one strain) are pTSS or sTSS and the majority of the strain-specific TSS (47–54%) are classified as antisense TSS (Table 3). Around 20 orphan TSS were also detected for each strain, 10 of which are conserved in all four strains.

Table 3. Comparative TSS annotation.
pTSS: primary TSS, sTSS: secondary TSS, iTSS: internal TSS, asTSS: antisense TSS.

	All TSS	pTSS	sTSS	iTSS	asTSS	Orphan
All TSS in SuperGenome	3377	973 (29%)	327 (10%)	1217 (36%)	1624 (48%)	61 (2%)
TSS in individual strains:						
RM1221	2167	744 (34%)	202 (9%)	735 (34%)	988 (46%)	25 (1%)
NCTC11168	1905	675 (35%)	180 (9%)	653 (34%)	843 (44%)	17 (1%)
81-176	2003	676 (34%)	193 (10%)	651 (33%)	937 (47%)	20 (1%)
81116	1944	690 (35%)	181 (9%)	653 (34%)	856 (44%)	25 (1%)
Conserved TSS						
Detected in all strains	1035	527 (51%)	92 (9%)	328 (32%)	445 (43%)	10 (1%)
Detected in 2 or 3 strains	1067	204 (19%)	118 (11%)	398 (37%)	534 (50%)	36 (3%)
Strain specific TSS						
RM1221	450	104 (23%)	39 (9%)	159 (35%)	235 (52%)	3 (1%)
NCTC11168	246	46 (19%)	19 (8%)	106 (43%)	119 (48%)	3 (1%)
81-176	260	34 (13%)	24 (9%)	98 (38%)	141 (54%)	4 (2%)
81116	319	58 (18%)	35 (11%)	128 (40%)	150 (47%)	5 (2%)
TSS missing in RM1221	99	21 (21%)	10 (10%)	35 (35%)	50 (51%)	3 (3%)
TSS missing in NCTC11168	109	22 (20%)	11 (10%)	36 (33%)	59 (54%)	4 (4%)
TSS missing in 81-176	97	25 (26%)	8 (8%)	38 (39%)	40 (41%)	4 (4%)
TSS missing in 81116	165	33 (20%)	24 (15%)	61 (37%)	83 (50%)	6 (4%)

3.1.2. Comparative analysis of regulatory elements

MEME [63] was used to predict conserved motifs in the promoter regions (−50 to +1 region) of all TSS. Sequences were isolated upstream of all the 8,019 TSS in the four strains as input sequence to predict motifs of at least 45 nt in length. This revealed a periodic A/T-rich pattern instead of a clear −35 box followed by an extended −10 box (*TGxTATAAT*) for ~89% of the promoter regions in the four strains. This -10 box sequence was recognized as a consensus motif for the housekeeping sigma factor, σ^{70} (Fig. 8B). This motif fits with a previously predicted consensus sequence for a smaller number of *Campylobacter* promoters and for σ^{70} in *H. pylori* based on dRNA–seq, indicating that transcription predominantly initiates at an extended −10 box in Epsilonproteobacteria [52], [64]. Moreover, motifs which resemble the consensus binding sites for the alternative sigma factors σ^{28} (FliA) and σ^{54} (RpoN) were also identified for 141/8,019 (1.8%) and 36/8,019 (0.4%) TSS respectively (Fig. 8B). Separate analyses of the four strains indicated that there is no strong variation in their general promoter patterns (Appendix Fig. 1). Since some genes that are known to be regulated by the alternative sigma factors were missed in the *MEME* searches, a consensus motif based on nine and eight TSS from strain NCTC11168 with a

previously described FliA- and RpoN-dependent promoter, respectively was defined (Appendix Table 2). To find genes that contain promoter sequences of the alternative sigma factors, σ^{28} and σ^{54}, that were not detected during the MEME motif generation, the -50 to +1 sequences for each TSS were scanned for σ^{54} and σ^{28} consensus motifs. This revealed additional 138 and 26 TSS fitting the FliA and RpoN motifs, respectively (Appendix Table 3).

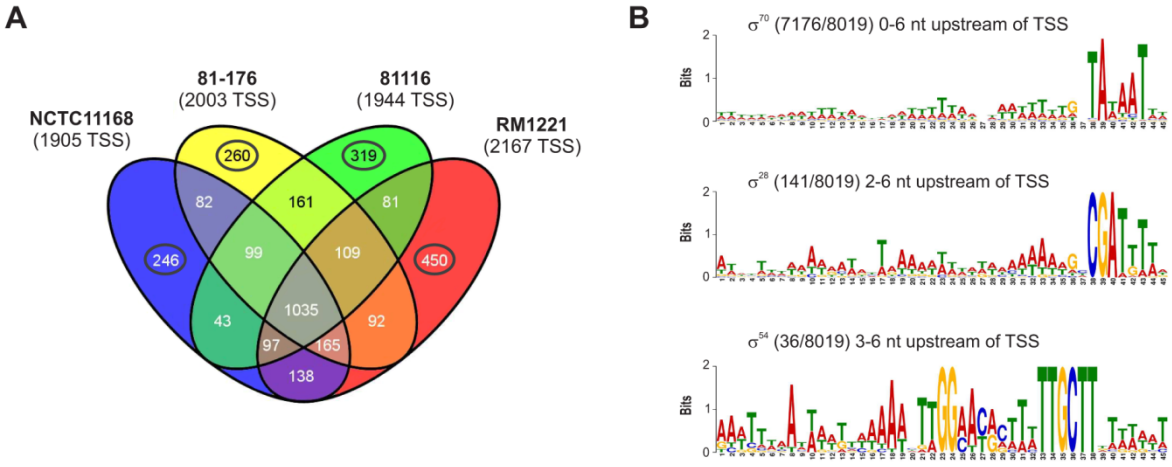


Figure 8. Transcriptome features of multiple *C. jejuni* strains. (A) Venn diagram showing the overlap among detected TSS between the four *C. jejuni* strains. The total numbers of TSS that were detected in each strain are indicated in brackets. TSS detected in individual strains are encircled. See also Table 3. (B) Motif searches in the −50 to +1 sequences upstream of the 8,019 detected TSS in all four *C. jejuni* strains reveal promoter consensus sequence motifs for the three sigma factors, σ^{70} (RpoD), σ^{28} (FliA), and σ^{54} (RpoN). The numbers of sequences which contain the motif are indicated as well as the distances to the associated TSS.

Approximately 35% and 10% of all TSS in each strain are classified as pTSS and sTSS, respectively (Table 3). The majority of the 3,241 5'UTRs defined by pTSS and sTSS of mRNA genes have a length between 20 and 50 nt (Fig. 9A). 5' UTR sequences were extracted from 3,158 of them with length ≥8 nt. *MEME* searches using these sequences as input detected an "AAGGA"-motif surrounded by A/T-rich sequence as a consensus for the ribosome binding site in the majority (95%) of sequences.

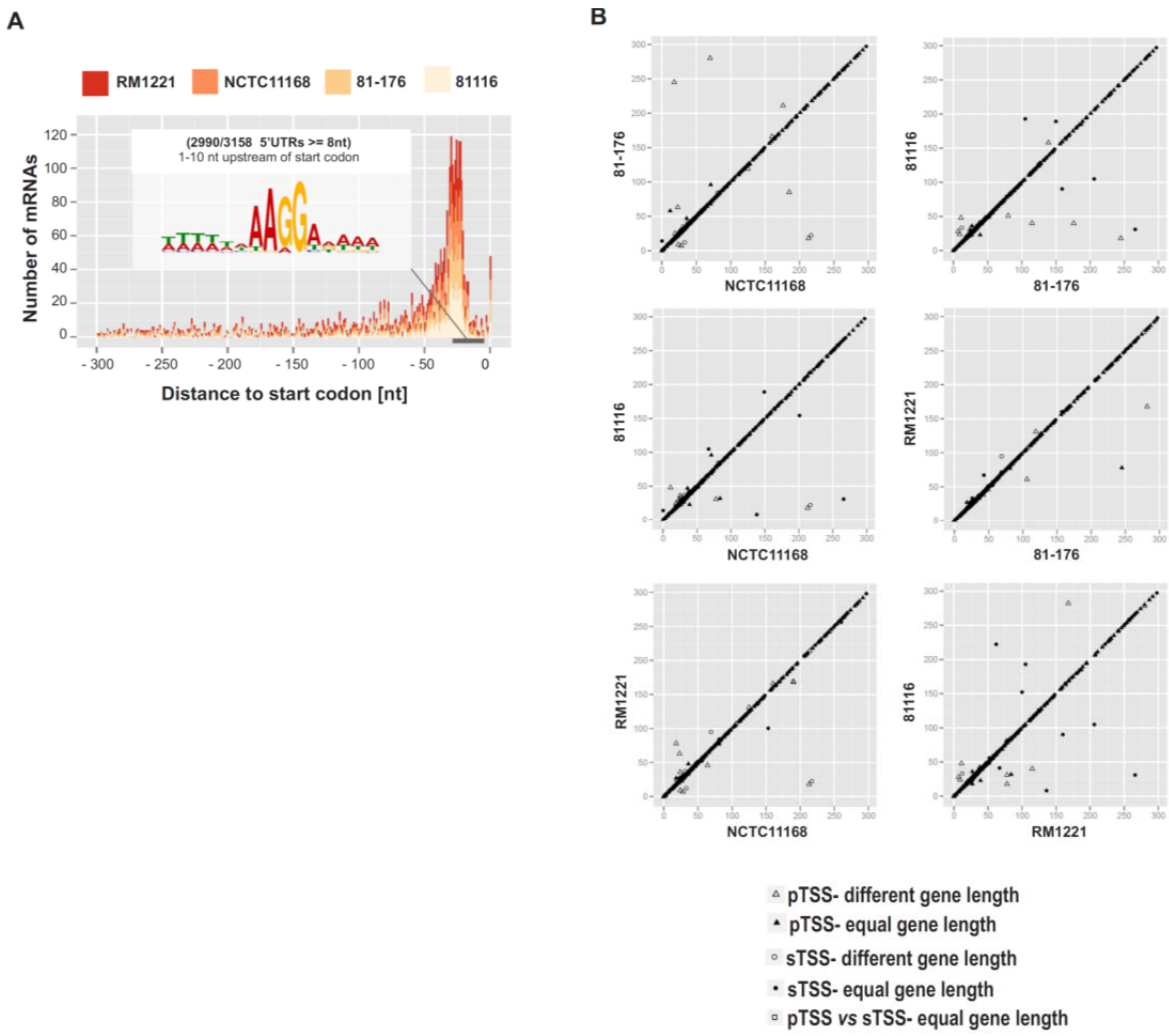


Figure 9. 5' UTR length distributions and comparisons. (A) 5' UTR length distribution for 3,241 primary and secondary SuperGenome TSS (pTSS and sTSS) of mRNAs of the four *C. jejuni* strains plotted as the frequency of the distances between primary or secondary TSS to the start codon. The inset shows the consensus ribosome binding site motif (AAGGa) that was found at a distance of 1 to 10 nt upstream of the start codon in 95 % of the 3,158 5' UTR sequences which have a minimum length of 8 nt. **(B)** For each pair of strains, the 5' UTR lengths of orthologous genes (defined via a *BLAST* based analysis – see Methods section) of the four strains were compared pairwise and visualized as a scatter plot. The dot symbols of the plots represent the different types of comparisons of the associated TSS (primary *vs.* primary = triangle, secondary *vs.* secondary = circle, primary *vs.* secondary = square) and length comparison of the two orthologs (filled symbol = genes have the same length; empty symbol = genes have different length). If one or both genes of an ortholog pair had more than one assigned 5' UTR length, i.e. in the case of multiple promoters, the 5' UTR pairs with minimal difference were compared.

Pairwise comparison of the 5'UTR lengths of genes with at least one ortholog in one of the other strains showed that the majority of the conserved genes have the same 5'UTR length but revealed also several length variations among strains (Fig. 9B). Despite a correction - where possible - for differences in start codon annotations (Appendix Tables 4 and 5), many of the 5'UTR length differences are due to different 5' ends of the CDS of the respective genes (Appendix Table 6). In some cases, different 5'UTR lengths result from insertion of new genes upstream of the start

codon and the acquisition of new promoters in certain strains (Fig. 10). For example, in three of the strains, a gene encoding for a restriction modification (RM) enzyme (CJE1195, C8J_0992, *cjeI*) is inserted upstream of *npdA*, encoding a NAD-dependent deacetylase, which is transcribed from a pTSS that was probably acquired together with *cjeI*. In contrast, strain 81–176 lacks the RM enzyme as well as the downstream promoter and *npdA* is co-transcribed as a polycistronic mRNA from a TSS upstream of *murC*.

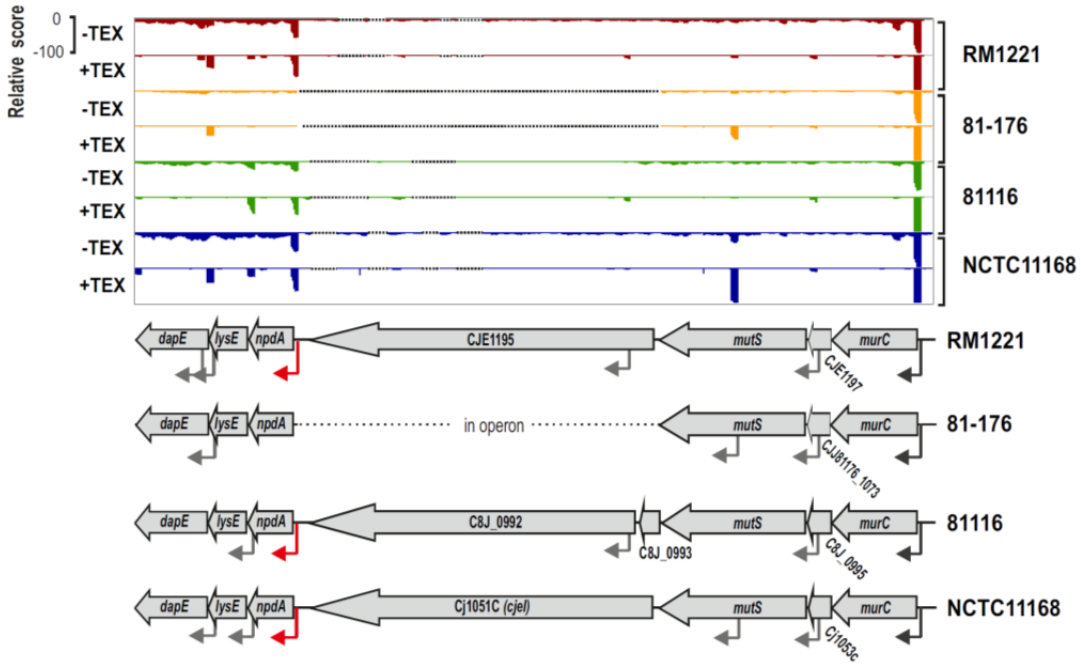


Figure 10. Effect of gene insertion on transcriptome features. dRNA–seq reads mapped to the SuperGenome between the *murC* and *dapE* genes in the four strains. In three strains, the *cjeI* gene is inserted together with a TSS (red arrow) for transcription of the downstream genes.

Besides insertions of new upstream genes with a novel promoter, different promoters can lead to 5'UTR length variation among strains. It has previously been reported that the *asnA* gene of strain 81–176 has acquired a sec-dependent secretion signal to the otherwise cytoplasmic asparaginase found in NCTC11168 and facilitates asparagine utilization in this strain [65]. The comparative dRNA–seq data indicates that the different *asnA* forms are transcribed from strain-specific promoters with ~100-fold higher cDNA read counts for the secreted asparaginase compared to the cytoplasmic form (Fig. 11).

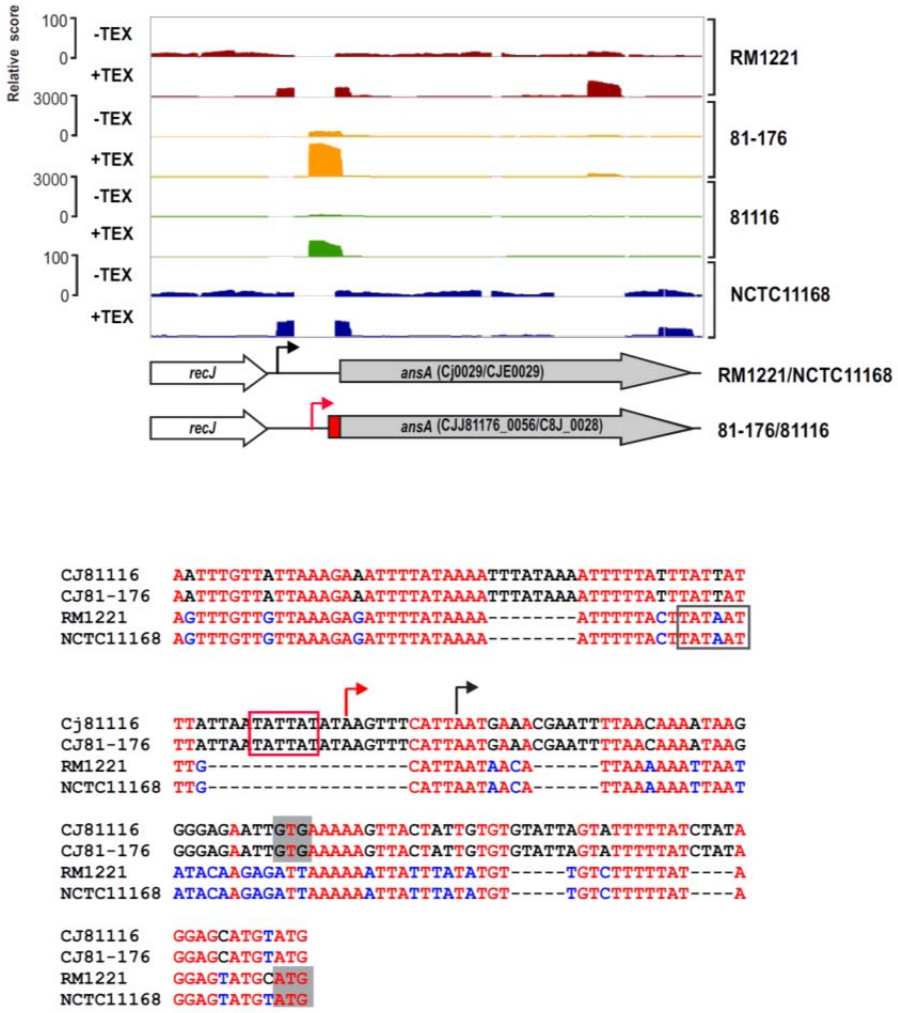


Figure 11. Orthologous genes with differences in 5' UTR length. dRNA-seq reads mapped to *ansA* encoding L-asparaginase *(Top panel)* and alignment of the *ansA* promoter and 5' UTR region *(Bottom panel)*. The *ansA* gene harbors a signal sequence in strain 81-176, which is marked by a red box upstream of the grey arrow indicating the CDS. The start codon of *ansA* in the alignment is marked in grey and differs among strains. In strain 81116 the annotation for *ansA* is too short but can be extended to the signal sequence. Note that *ansA* is only weakly expressed in NCTC11168 and RM1221 (different scales). The dRNA-seq reads reveal two different TSS: 1) Black arrow: a TSS in RM1221 and NCTC11168 leading to a long 5' UTR and 2) Red arrow: a TSS in 81-176 and 81116 leading to a short 5' UTR due the presence of the signal sequence. The alternative TSS have -10 boxes for σ^{70} at different positions. The dotted lines (in the genome browser snapshot) represent the gaps in the alignment of the four genomes.

3.1.3. SNPs lead to strain specific transcriptional output

Comparison of promoter motifs for conserved and strain-specific TSS revealed no difference in the general promoter patterns (Appendix Fig. 2). However, comparative TSS detection allowed identification of regions with strain-specific promoter usage, *e.g.* regions for which a TSS was detected in only some of the strains although the region is present in the SuperGenome in all strains (Appendix Table 1, compare columns "mapCount" and "detCount"). It was observed that

in many cases single nucleotide polymorphisms (SNP) lead to disruption of promoters in a subset of strains (Fig. 12 and Appendix Fig. 3). For example, the pTSS of *pnk*, encoding an inorganic polyphosphate/ATP-NAD kinase, is enriched and conserved in all four strains (Fig. 12). In contrast an iTSS within *pnk* is only detected in RM1221 and NCTC11168. In strains 81–176 and 81116 a C to T exchange in the "CGATTT" σ^{28} consensus sequence seems to be sufficient to abolish transcription initiation from this TSS.

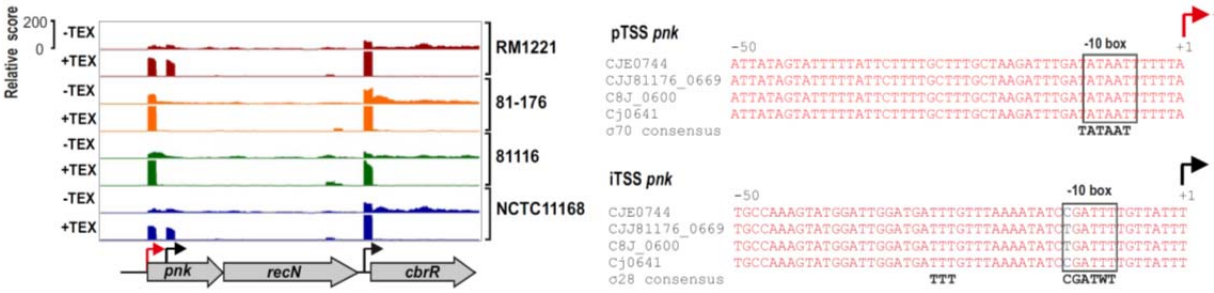


Figure 12. SNP–dependent promoter usage in *C. jejuni*. dRNA–seq reads mapped to the *pnk-recN* operon in the SuperGenome. Red and black arrows indicate a primary (pTSS) and an internal TSS (iTSS), respectively. Corresponding alignments of the promoter region of the two TSS for *pnk* are shown. The iTSS within *pnk* is not expressed in two of the strains (81–176 and 81116) which carry a T to C exchange at the conserved "T" residue of the −10 box.

Moreover, it was also noted that mutations which alter the periodic A/T-rich pattern upstream of the −10 box could also affect transcription (Fig. 13A). For example, in strain NCTC11168, the Cj0004c and Cj0005c genes encode a monoheme cytochrome *c* and molybdopterin oxidoreductase, respectively, which allows *C. jejuni* to use sulphite as a respiratory electron donor [66]. This bi-cistronic operon along with the σ^{70} −10 box is conserved in all four strains (Fig. 13A). However, mutations in the A/T rich upstream pattern of its pTSS in strain 81116 coincide with a loss of transcription either due to disturbing transcription initiation or binding of some regulatory factor. In line with this, cytochrome *c* reduction as a measure for sulphite oxidation was detected in only in only three of the strains (Fig. 13B). This indicates, that although genes are conserved among strains and also show high conservation in promoter regions, they may not necessarily be transcribed in all of them. Additional examples for mutations within conserved promoter elements and A/T-rich upstream pattern are shown in Appendix Fig. 4.

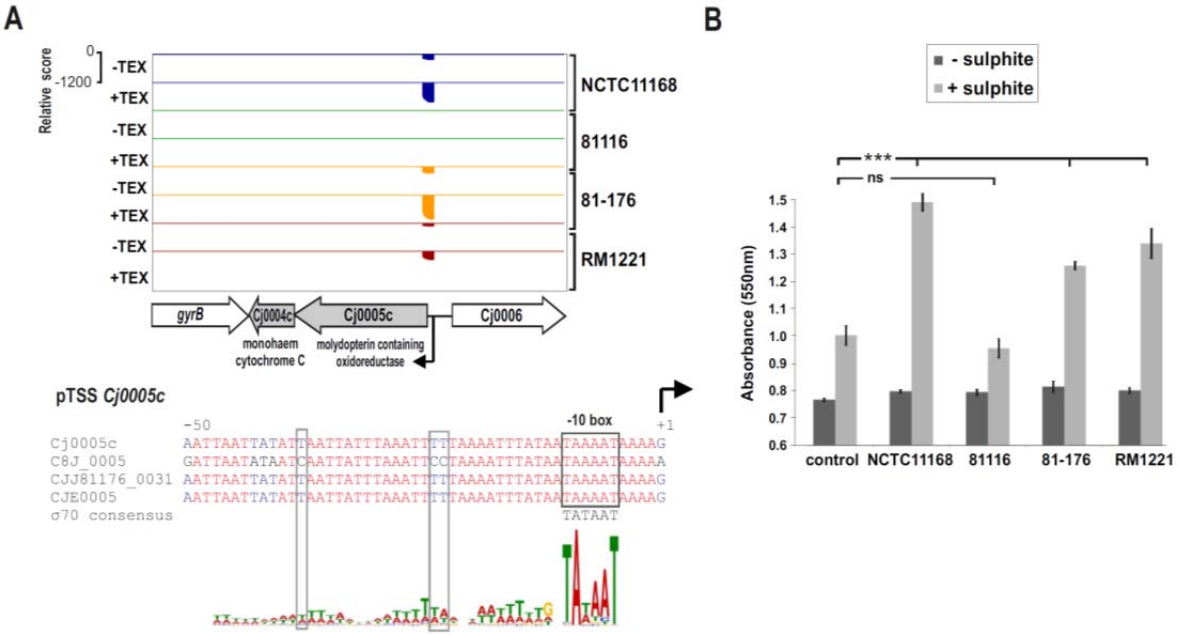


Figure 13. SNP–dependent promoter usage in *C. jejuni* affects sulphite respiration. (A) dRNA–seq reads mapped to the Cj0005c-Cj0004c operon, which encodes for a molydopterin containing oxidoreductase and monoheme cytochrome *c*. Promoter alignment of the primary TSS of Cj0005c shows disruptions in the A/T rich cyclic pattern upstream of the conserved σ^{70} −10 box shows in 81116. **(B)** Intact cells from the four *C. jejuni* strains were assayed for cytochrome *c* oxidoreductase activity in the absence (dark grey bars) and presence (light grey bars) of sulphite. The reduction of cytochrome *c* was measured as the increase in the absorbance at 550 nm. Strains NCTC11168, 81–176, and RM1221 showed a significant increase ($p<0.0001$) over basal levels (control without cells) of cytochrome *c* reduction when the samples were supplemented with sulphite.

3.1.4. The non-coding RNA repertoire of *C. jejuni*

In addition to TSS detected for genes, many candidates for non-coding RNAs were also identified in the dRNA-seq dataset. Around 45% of the TSS for each strain were classified as antisense TSS (asTSS) and 445 out of them were detected in all four strains, indicating a large fraction of conserved antisense transcription (Table 3, Appendix Fig. 5). Around 1% (20 TSS) of each strain was classified as orphan, which are a major source of small peptide encoding genes and *trans*-encoded sRNAs. Apart from this automated approach, manual inspection of the transcriptome also revealed several candidates for *trans*-encoded sRNAs in the chromosomes of the four strains and on the pVir and pTet plasmids (Fig. 14, Appendix Fig. 6, Appendix Table 7). The sequence of sRNA candidates in the strain NCTC11168 is provided in Appendix Table 8. Northern blot profiling under different growth phases confirmed expression of most of these sRNA candidates. Most of the sRNAs were transcribed by the σ^{70} housekeeping promoter, but some others like CJnc10 and CJnc170 were transcribed from the alternative σ^{28} promoter. Some sRNAs, such as CJnc60 or CJnc140 were highly conserved and show similar expression patterns in all strains. In contrast, some of the other conserved sRNAs, such as CJnc180 and CJnc190, which are encoded antisense to each other, show differential expression patterns among strains. Moreover, the Northern blots

confirmed strain-specific sRNAs, such as CJnc30, which is only present in NCTC11168, or CJnc20, which is missing in RM1221. One such strain-specific sRNA, CJnc100, is one of the most highly expressed transcripts in the transcriptome of the strain 81116. Most of the candidate sRNAs found were transcribed from their own TSS, but also examples of sRNA candidates generated by processing, e.g., from the 3' end of a mRNA were also found (CJnc120, Fig. 14A). Furthermore, the majority of the sRNA candidates were found to be accumulated in exponential or stationary growth phase.

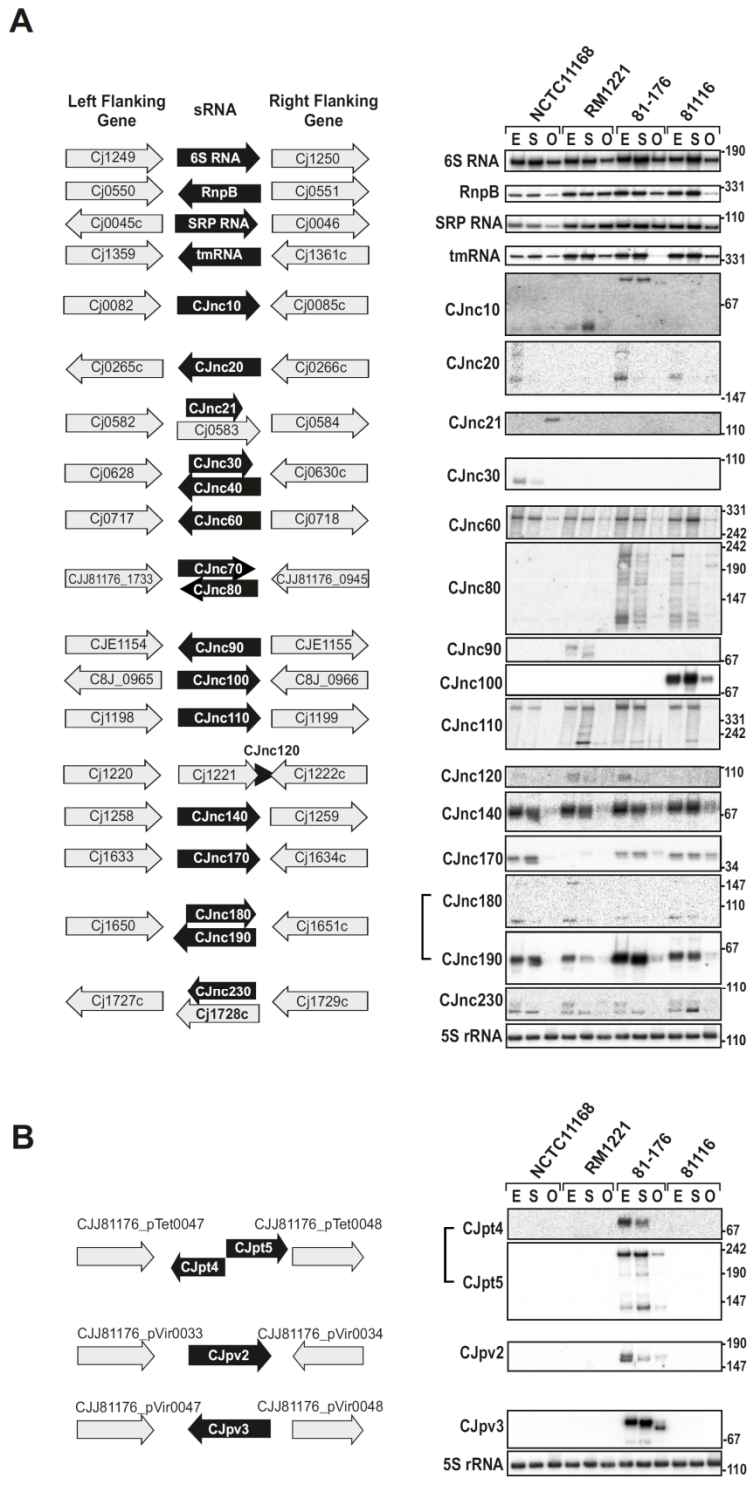


Figure 14. Small RNAs in *Campylobacter jejuni*. (A) *(Left)* Left and right flanking genes and orientation of sRNA candidates in *C. jejuni*NCTC11168. Arrows indicating the genes are not drawn to scale. Small RNA candidates were termed "CJncXXX" and numbered in steps of ten according to their genome position. *(Right)* Expression analysis of the housekeeping and candidate sRNAs during different growth phases in the four *C. jejuni* strains. Specifically, total RNA (15 µg per lane) was extracted at mid-exponential (E), stationary (S), and overnight (O) growth phase and analyzed by Northern blot using labeled DNA probes complementary to the sRNAs. The blots were probed for the housekeeping 5S rRNA as loading control. Note that the oligonucleotide probe for CJnc170 does not detect the homolog in strain RM1221 due to point mutations at probe binding site. **(B)** Genomic locations and Northern blots for sRNA candidates in the pVir and pTet plasmids of *C. jejuni* strain 81–176.

Conservation analysis of the sRNA candidates showed that the majority of them are restricted to *Campylobacter jejuni* (Fig. 15), indicating that they either have a specific regulatory function in this species or that their sequence conservation is not high enough to detect homologs by BLAST searches. Even the housekeeping RNAs (SRP RNA, tmRNA, RNase P RNA, and 6S RNA) are not conserved at the sequence level outside *Campylobacter* species. Several of the *C. jejuni* sRNA candidates, e.g., CJnc170 and CJnc190, are highly conserved in diverse strains, and could have a more general regulatory role within *C. jejuni*. In contrast, some sRNA candidates such as CJnc30 and CJnc80 or the plasmid-encoded sRNAs are found in only some of the strains, indicating strain-specific sRNA repertoires which might contribute to strain-specific regulation of gene expression.

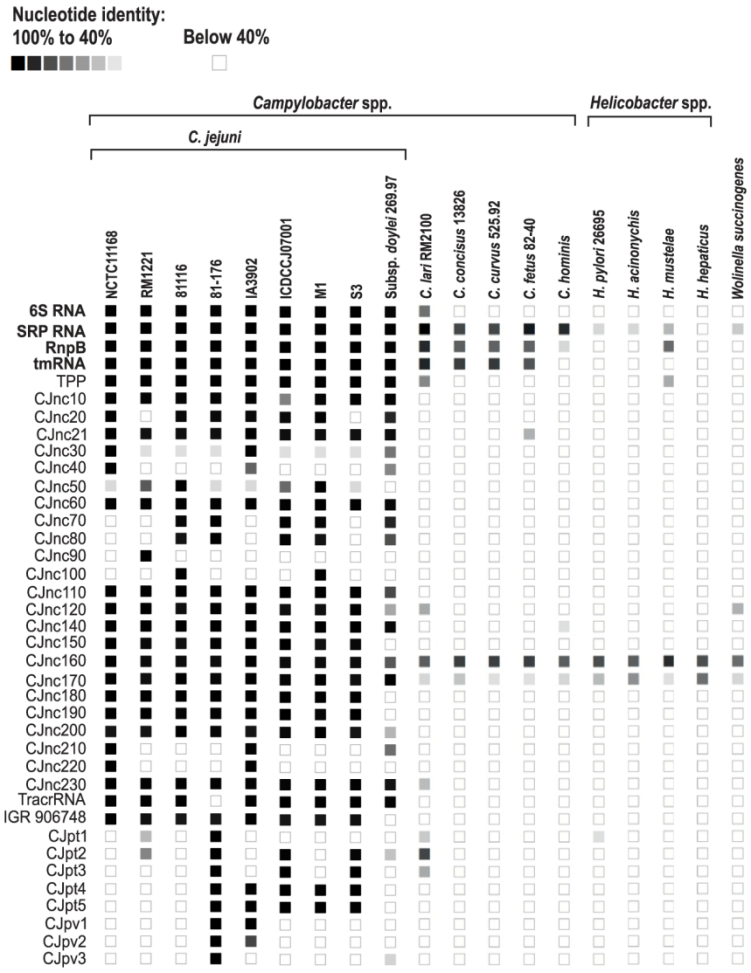


Figure 15. Conservation of *Campylobacter jejuni* sRNA candidates in Epsilonproteobacteria. Conservation of sRNA candidates in different *C. jejuni* strains and other representative Epsilonproteobacteria was analyzed by *BLAST* searches. The color intensity of the boxes represents the sequence identity to sRNAs of *C. jejuni* NCTC11168 or in case of the plasmid encoded sRNAs to strain 81–176. Identity values below 40% as well as the lack of an ortholog are symbolized by empty boxes. Housekeeping sRNAs are marked in bold.

3.1.5. The minimal CRISPR/Cas system of *Campylobacter jejuni*

The region with highest numbers of cDNA reads after rRNA in strains NCTC11168 and 81116 corresponded to the CRISPR (clustered regularly interspaced short palindromic repeats) locus (Fig. 16A and Appendix Fig. 7). Recently, the Type- II CRISPR-Cas system from *S. pyogenes* has been adapted for precise genome editing in virtually all organisms [67]. This precise genome editing method using the Cas9 protein has been heralded as one of the biggest breakthroughs in Biology research. Some of the *Campylobacter* strains also encode the type-II CRISPR/Cas system, with one of the smallest Cas9 proteins known [68].

CRISPR loci are transcribed as precursors that are processed into mature crRNAs that together with Cas proteins silence invading foreign nucleic acids such as plasmids or phages [69-71]. Three of the four *C. jejuni* strains used in this study harbor a type-II CRISPR/Cas system, which requires a trans-encoded sRNA, TracrRNA, and the host factor RNase III for proper maturation or crRNAs [72, 73]. The dRNA–seq analysis shows that the crRNAs and TracrRNA are actively transcribed and share a stretch of perfect complementarity which allows for RNase III-dependent processing and maturation of the crRNAs at the 3' end (Fig. 16A). In line with this, an accumulation of processed ~38-nt spacer-repeat units and processed ~62-nt TracrRNA was observed (Fig. 16B).

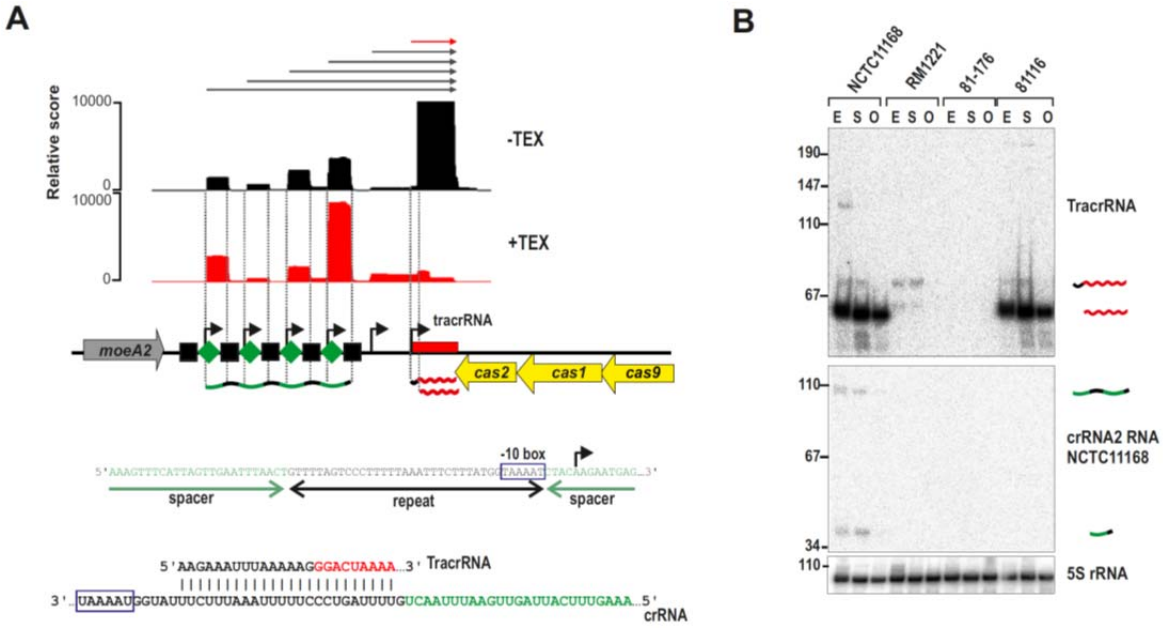


Figure 16. CRISPR locus is expressed in *C. jejuni*. (A) cDNAs reads of TEX−/+ libraries mapped to the CRISPR loci in *C. jejuni*NCTC11168. TSS are marked by black arrows. A −10 box for each crRNA is present in the 3' part of the corresponding upstream located repeat (blue box). The lower panel shows the potential base-pairing of *C. jejuni* TracrRNA with the repeat sequence. **(B)** Northern blot analysis of TracrRNA and crRNAs. Total RNA (15 µg per lane) extracted from *C. jejuni* strains at mid-exponential (E), stationary (S), and overnight (O) growth phase was analyzed on a Northern blot using labeled DNA probes complementary to

TracrRNA (two upper panels) and crRNA2 in NCTC11168. Both the mature (~62 nt) and processed (~74 nt) TracrRNA forms as well as intermediate (~103 nt) and mature (~38 nt) crRNA forms for crRNA2 in NCTC11168 were detected. 5S rRNA served as loading control.

Only weak CRISPR expression was detected in strain RM1221. Conservation analysis showed that the *cas9* gene, which is required for the stability of crRNAs and cleavage of target DNA [72, 74], carries a stop mutation in strain RM1221 leading to a truncated non-functional protein (Fig. 17A). In strain 81–176 the CRISPR region is replaced by two other genes with very low G/C content compared to the flanking genomic regions.

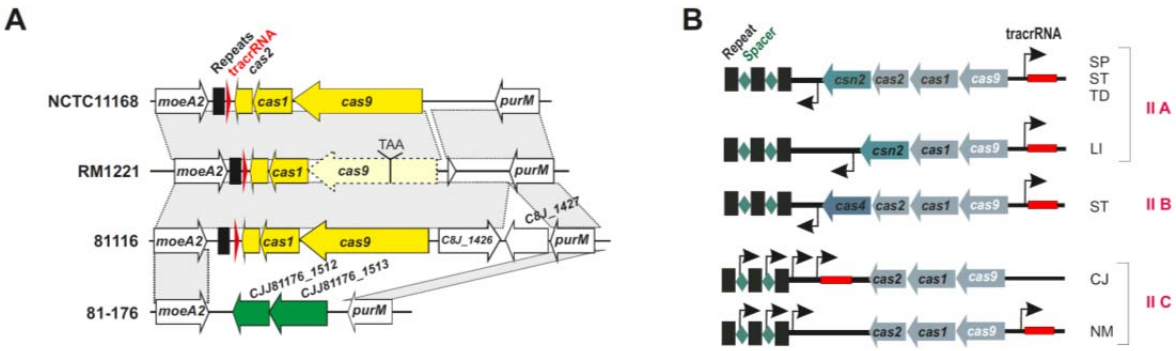


Figure 17. The minimal CRISPR-Cas IIA system of *C. jejuni*. (A) Conservation of CRISPR loci in different *Campylobacter jejuni* strains. In strain RM1221, *cas9* contains a premature-stop mutation. In strain 81–176 the whole CRISPR locus is replaced by two genes (green arrows) with low G/C-content. **(B)** Sub classification of type-II CRISPR-loci. *cas9*, *cas1*, and *cas2* are indicated with blue arrows and *tracrRNA* by a red box, respectively. TSS upstream of the repeat (black squares)-spacer (green diamonds) units in type-II A and B CRISPR loci or within each spacer in the case of the minimal type-II C CRISPR/Cas systems of *Campylobacter jejuni* (CJ) and *Neisseria meningitidis* (NM) are marked by black arrows. Representative species for each type-II class are listed (SP:*Streptococcus pyogenes*, ST: *S. thermophilus*, TD: *Treponema denticola*, LI: *Listeria innocua*, and LP: *Legionella pneumophila*).

The RNase III/TracrRNA–dependent processing leads to a cleavage event within the repeat region of the crRNAs [72]. Surprisingly, the mature spacer-repeat units were enriched in the TEX-treated libraries, indicative of primary transcripts starting at position five within each spacer. Moreover, the 3′ end of each repeat ends with "GGTAAAAT" resembling an extended −10 box. Thus, *C. jejuni* apparently employs promoter sequences within each repeat to initiate transcription of the associated spacer unit, so that only one processing event mediated by RNase III and TracrRNA in the repeat sequence is required to generate the 3′ end of the mature crRNA (Fig. 17B). In line with this model, accumulation of longer transcripts upon deletion of RNase III was observed (Fig. 18B). Primer extension analysis confirmed that these longer transcripts start at the proposed TSS within each spacer (Fig. 18C). Thus, in contrast to other type-II systems, where the crRNAs are transcribed together with a leader sequence from one upstream promoter, each repeat in the *C. jejuni* CRISPR carries its own promoter. Another novelty which differentiates *C. jejuni* CRISPR/Cas

system even from the other type-II C systems is the fact that much of the TracrRNA species are transcribed from these upstream promoters along with the crRNAs, whereas in all other systems the crRNAs and the TracrRNA are transcribed completely independently. It remains to be found whether this co-transcription helps in faster crRNA maturation due to possible intramolecular processing.

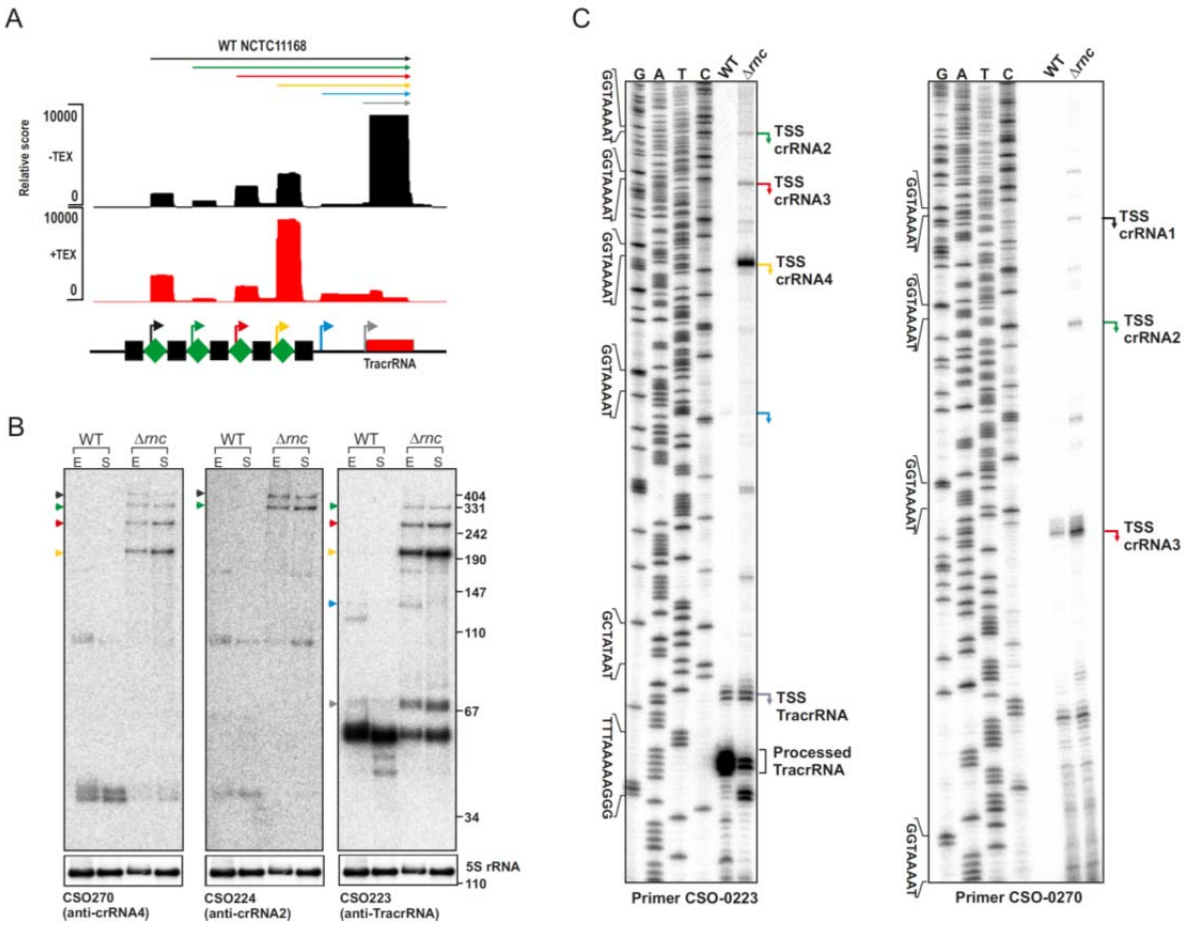


Figure 18. Deletion of RNase III leads to accumulation of longer crRNAs. (A) Sequenced reads of cDNA libraries derived from -/+ TEX-treated total RNA mapped to the CRISPR loci in *C. jejuni* NCTC11168. Arrows indicate transcripts initiating from the different TSS found in the locus. **(B)** Northern blot analysis of TracrRNA and crRNAs in NCTC11168 WT and *Δrnc*. Total RNA (10 µg per lane) extracted from NCTC11168 WT and *Δrnc* at mid-exponential (E) and stationary (S) phase was analyzed on a Northern blot using labeled DNA probes complementary to crRNA4, crRNA2 and TracrRNA in NCTC11168. **(C)** Primer extension analysis to determine the 5' end of different transcripts in the CRISPR loci. Labeled DNA probes, CSO-0223 and CSO-0270, were incubated with 5 µg of mid-exponential phase RNA from NCTC11168 WT and *Δrnc* in the RT reaction. Lanes G, A, T, C represent sequencing reactions carried out on extended CRISPR loci sequence of NCTC11168 amplified using oligos CSO-0223 and CSO-0270. The arrows indicate the identified TSS among other processed fragments. The σ^{70} -10 boxes of the TSS observed along with the TracrRNA processing site are marked on the left of the gel.

3.1.6. DISCUSSION

This chapter describes a comparative primary transcriptome analysis of four different *C. jejuni* strains using a novel common coordinate system and automated TSS annotation method. This comparative analysis revealed widespread strain-specific promoter usage and sRNA repertoires. The observed strain-specific transcriptional output revealed candidate genes which could contribute to phenotypic variation among strains and thereby facilitate adaptation to different hosts or micro-environments.

The generation of global TSS maps from RNA–seq data has primarily been based on a manual or semi-automated basis [52, 75, 76]. These strategies are typically very laborious and time-consuming with limited reproducibility, and requires even more effort when comparing transcriptomes from multiple strains or conditions. In this study, a model-based TSS prediction method based on criteria used for manual TSS annotations had been developed. It provides detailed values (such as number of read starts, enrichment factor etc.) as well as classifications (pTSS, sTSS, asTSS, iTSS) for each TSS candidate. Benchmarking of the TSS prediction method using the manually annotated TSS from *H. pylori* [52] showed high sensitivity and precision rate of prediction. Since this comparative approach also allows for the integration of data sets from different strains or conditions, its confidence level can generally be increased by the inclusion of replicates and multiple datasets. Moreover, the TSS annotation can be further improved by adjusting the parameters to additional training sets based on experimental validation. Recently, Schmidtke *et al.* presented a fully automated approach for TSS annotation of a single genome based on read count data and a sophisticated statistical model, which calculates p-values for TSS candidates and thereby provides a confidence estimation for observing a TSS [77]. However, the specific properties of the TSS candidates, such as expression height and dRNA–seq enrichment factor, cannot be directly inferred from the resulting p-values. In future, a combination of both approaches would be promising especially for weaker TSS candidates. Furthermore, the automated TSS method also allows for a comparative TSS annotation among dRNA–seq libraries from multiple strains by the integration of the SuperGenome approach. Considering the rapidly increasing number of genome sequenced as well as RNA–seq studies, the approach presented here will facilitate a systematic TSS annotation among different strains or growth conditions and can also be adapted to the analyses of eukaryotic transcriptomes as well. For the comparative transcriptome approach in this study, four widely used laboratory strains with available genome sequence were selected. However, this comparative approach can be easily adapted to the analysis of multiple isolates from different hosts or isolation sources.

The majority of TSS are conserved among the four strains, but also many instances were reported where SNPs in promoter regions apparently resulted in strain-specific promoter usage. Thus, although some promoters are highly conserved and show almost perfect overlap to promoter consensus motifs - which would be indicative for active transcription - the respective genes are not necessarily expressed at the same level among strains. Therefore, comparative transcriptomics facilitates the identification of differences in the functional output from genomes which cannot be directly inferred from apparently similar DNA sequences. SNPs can be adaptive and, thereby, lead to niche expansion [78]. Even single regulatory genes or promoter inversions can be sufficient to alter bacterial host specificity [79, 80] and promoter SNPs have been associated with overproduction of virulence genes [81, 82]. Based on the global TSS maps, several examples were found where SNPs disrupt conserved positions in promoter motifs recognized by the three sigma factors leading to abolition of transcription. Furthermore, several examples were found where disruptive SNPs in the A/T-rich upstream regions, indicating that this region is also required for transcription initiation in Epsilonproteobacteria. For example, in strain 81116 mutations were observed upstream of the promoter for monoheme cytochrome *c* and a molybdopterin oxidoreductase, which allow *C. jejuni* to use sulphite as a respiratory electron donor [66]. 81116 has lost the ability to transcribe this operon and hence even lost the ability to respire sulphite (Fig. 12). *C. jejuni* is unable to metabolize glucose since it lacks the enzyme phosphofructokinase. Instead, it has a complex branched respiratory chain and can utilize a variety of electron donors like formate, lactate, or sulphite. Sulphite respiration might also help *C. jejuni* to survive in sulphite-rich niches and foods. Since the SNPs in promoter regions could also interfere with binding of a transcriptional regulator and thereby affect transcription, such promoters with strain-specific expression patterns represent good candidates to fish for novel DNA-binding proteins. For example, a SNP in the Fur-binding site of the *sodB* promoter of certain *Helicobacter pylori* strains has been shown to affect direct binding of apo-Fur [83]. The global map of TSS and promoter SNPs might give hints as to how *C. jejuni* and other microbes with compact genomes and few transcription factors could adapt gene expression according to different environmental conditions.

The global transcriptome maps revealed several candidates for conserved and strain-specific sRNAs. Small regulatory RNAs have been implicated as key regulators in metabolic pathways and during pathogenesis [13, 84]. These newly identified sRNAs may contribute to virulence gene regulation and host adaptation by modulating metabolic pathways which are important for host colonization [65]. A study based on conventional, non strand-specific RNA–seq predicted five sRNAs in *C. jejuni* [85]. The expression of four of these sRNAs was also observed in this study (Appendix Fig. 6). The majority of the sRNA candidates are expressed as independent

transcripts in *Campylobacter*. Most of the sRNA candidates are not conserved outside the Campylobacterales. A few examples of processed sRNA species, which can be generated from the 3' end of mRNAs were also detected in this study. Recently, such 3'-end derived transcripts have been shown to stably associate with the RNA chaperone Hfq in *Salmonella* and to act as regulatory RNAs on *trans*-encoded mRNAs [86-88]. Since *Campylobacter* lacks Hfq, it will be interesting to see whether its sRNAs require a different RNA chaperone or accessory protein for their activity and stability or whether they act independently. Recently, a plasmid encoded sRNA identified in this thesis, CJpv3 (cjrA), was shown to act as an antitoxin in a Type I toxin-antitoxin system in the strain 81-176 [89]. In another recent study, two sRNAs (CJnc10 and CJnc170) transcribed from the alternative σ^{28} promoter were predicted to target several σ^{54}-dependent genes in *C. jejuni* [90]. Future studies will be required to uncover the regulatory roles of other sRNAs in *Campylobacter*.

The most abundant sRNA in strain NCTC11168 corresponds to TracrRNA, which is required together with RNase III for maturation of CRISPR RNAs [72]. Due to the high variability of the spacer sequences, CRISPR loci have been used for strain genotyping including *Campylobacter* species [91]. The crRNAs are transcribed from a leader sequence and several processing steps are required to generate the mature crRNAs. Surprisingly, the crRNAs in *C. jejuni* were found to be transcribed from individual promoters at the end of each repeat unit (Fig. 17). The CRISPR locus of *Campylobacter* with its individual crRNA promoters and only three *cas* genes represents a "minimal" system of the type-II subtype (Fig. 16B) which requires only one processing event by RNase III within the repeats to generate the mature crRNAs. Out of the three Cas proteins, Cas1 and Cas2 are known to be involved in spacer acquisition while Cas9 along with crRNA and TracrRNA is involved in DNA targeting. A similar CRISPR organization was also identified in *Neisseria meningitidis* [92], where it was also shown that RNase III based processing was dispensable for interference. Furthermore, in strain RM1221, the crRNAs and TracrRNA were only weakly expressed probably due to a stop-mutation in *cas9*, while strain 81–176 completely lacks the CRISPR locus. Interestingly, these two strains carry prophages or plasmids, indicating that these horizontally acquired genetic elements could be mutually exclusive with an active CRISPR system. Conservation analysis in additional strains showed that strains with plasmids or integrated elements very often carry degenerated CRISPR loci (Appendix Fig. 8). Moreover, it has recently been shown that ganglioside-like LOS structures of GBS-associated *C. jejuni* strains can confer efficient bacteriophage resistance and that the presence of sialyltransferases correlates significantly with an apparently non-functional CRISPR system [93]. Further studies of the influence of the CRISPR system on pathogenicity of *Campylobacter* will be required.

Overall, the high-resolution transcriptome maps revealed regulatory elements and their conservation in multiple *Campylobacter jejuni* strains on a genome-wide scale. The comparison of multiple strains improves annotation of transcriptome features such as TSS maps and reveals strain-specific TSS usage as well as sRNA repertoires. These strain-specific transcription patterns will provide new insights into genes which could promote phenotypic differences despite high conservation at the genome level.

3.2. Exploring the translational regulator CsrA and its antagonists in *C. jejuni*

CsrA (carbon storage regulator) is a global translational regulatory protein which primarily binds near the RBS of an mRNA to repress its translational [19]. The regulatory activity of CsrA is controlled by its binding to RNA or protein antagonists. sRNAs like CsrB/C which harbor multiple CsrA binding sites act as a CsrA sponge and are regulated by external cues to antagonize CsrA activity. In *Bacillus Subtilis*, a flagellin chaperone protein, FliW, can also directly bind and antagonize CsrA function [23]. CsrA antagonizing sRNAs are well conserved and studied in Gammaproteobacteria, but is not known whether they are present in all bacteria which encode for CsrA, like *C. jejuni.* Albeit previous studies have indicated a role of CsrA in virulence and stress response in *C. jejuni* and *H. pylori* [49, 94, 95], the direct RNA substrates of CsrA are largely unknown in Epsilonproteobacteria. Likewise, only a single or a handful of direct CsrA targets have been validated in other bacterial species, e.g. only *hag* mRNA, encoding flagellin, is known to be a direct target of CsrA in *B. subtilis* [23, 96]. This chapter investigates CsrA targets and its antagonists using global RNA-seq based approaches, followed by the molecular characterization of such findings.

3.2.1. Genomic location and conservation analysis of *Campylobacter* CsrA.

The C. jejuni *csrA* gene (Cj1103) is encoded between the *truB* and *ispE* genes (Fig. 19A). The comparative differential RNA-seq analysis described/ summarized in Chapter 1 revealed that *csrA* is encoded in the middle of an twelve-gene operon from *sstT* to *clpA*, with three potential σ70 promoters driving *csrA* transcription [97]. One strong promoter is located upstream of *sstT* and could drive transcription of the whole operon (Fig. 19A, blue arrows). Two additional TSS were detected within the *uvrD* and *truB* genes upstream of csrA, which might lead to uncoupling of the downstream genes from the rest of the operon. The TSS closest to the *csrA* is located 175 bp upstream of its start codon

Alignment of CsrA amino acid sequences from *Campylobacter* as well as additional Epsilonproteobacteria and other representative bacterial species showed high conservation of the two RNA binding motifs common to CsrA proteins [98] (Fig. 19B). However, in contrast to CsrA/RsmA from Gammaproteobacteria (EC: *E. coli*, SE: *Salmonella* enterica, PF: *Pseudomonas fluorescens*), the CsrA homologs from the Epsilonproteobacteria carry a C-terminal extension which is also found in *B. subtilis* (BS), *Clostridium botulinum* (CB) and *Borrelia burgdorferi* (BB). Thus, despite C. jejuni belongs to Gram-negative bacteria, CsrA homologs of

Epsilonproteobacteria seem to be more similar to their counterparts from Gram-positive bacteria compared to Gammaproteobacteria.

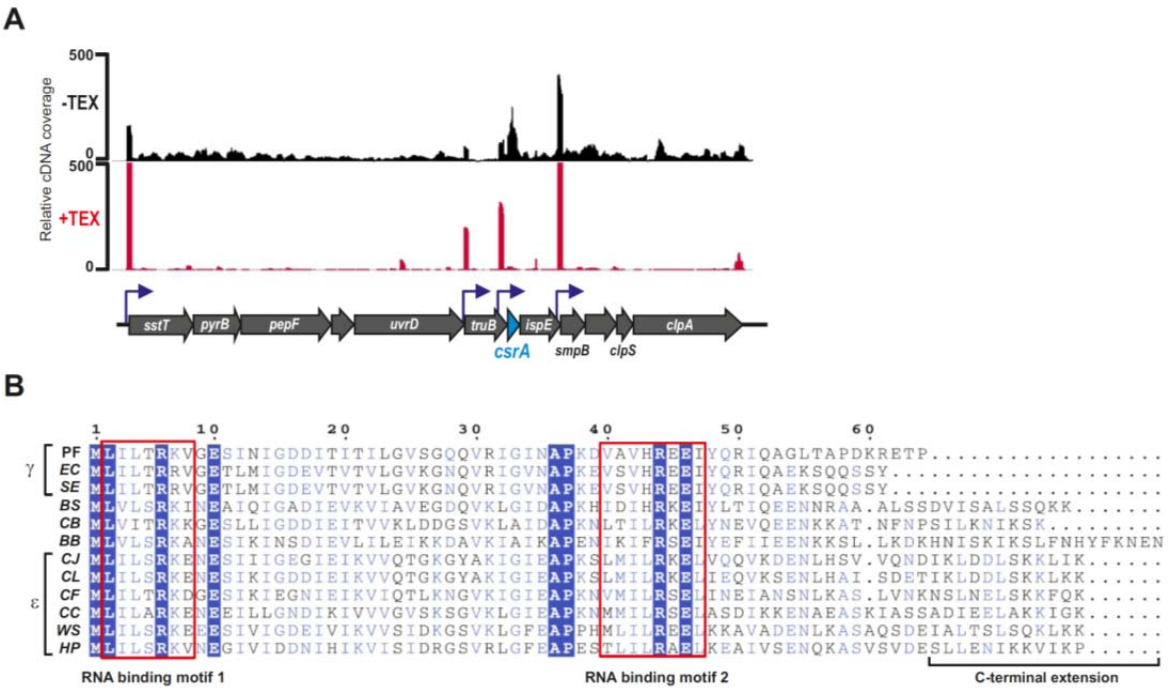


Figure 19. Genomic location and conservation of CsrA protein. (A) dRNA–seq reads mapped to the *csrA* encoding operon. The TSS detected along the operon are marked with blue arrows. **(B)** Conservation of different CsrA homologs at protein level show two conserved RNA binding motifs and a C-terminal extension in some bacteria. (γ-proteobacteria: PF, *Pseudomonas fluorescens*; EC, *Escherichia coli*; SE, *Salmonella enterica*; ε-proteobacteria: CJ, *C. jejuni*; CL, *C. lari*; CF, *C. fetus*; CC: *C. curvus*; WS, *Wolinella succinogenes*; HP, *Helicobacter pylori*; Others: BS, *Bacillus subtilis*; CB, *Clostridium botulinum*; BB, *Borrelia burgdorferi*)

3.2.2. Transcriptome changes upon *csrA* deletion in *C. jejuni*

CsrA interferes with translation initiation of many target mRNAs by sequestering the RBS, which is often coupled to a decrease in target mRNA stability since they are no longer protected from RNases by the translating ribosomes. In order to identify potential target mRNAs of CsrA in *C. jejuni*, global changes in gene expression between the *C. jejuni* NCTC11186 wildtype (WT) and a *csrA* deletion mutant were investigated by RNA-seq. Unlike *E. coli*, where deletion of *csrA* leads to a very sick phenotype [99], the deletion of *csrA* does not affect growth characteristics in *C. jejuni* [49]. Accordingly, global transcriptome changes upon *csrA* deletion might be directly associated with CsrA (*e.g.* target mRNAs), and are not due to growth defects or other pleiotropic effects (Fig. 21A). The *C. jejuni* NCTC11168 WT and Δ*csrA* mutant were grown to mid-log phase for RNA extraction and subsequent RNA-seq analysis. RNA-seq data were mapped and quantified over all ORFs, as well as previously annotated 5' UTRs. Apart from the *csrA* gene itself, a set of only five genes (ORF or 5' UTR) was regulated more than 4-fold upon *csrA* deletion (Fig. 20A). These genes

include two genes of known function – Cj0246c encodes a MCP-domain signal transduction protein and Cj1298 encodes for an N-acetyltransferase. The expression of the flagellin mRNAs is known to be translationally repressed by CsrA in other bacteria such as *Bacillus subtilis* [96] and *Borrelia burgdorferi* [100]. However, no changes in transcript levels were observed for the *flaA* (encoding major flagellin) and *flaB* (encoding major flagellin) mRNAs upon *csrA* deletion in *C. jejuni* (Fig. 20B). Also, no major difference was observed in other genes of the operon encoding *csrA*, eliminating the presence of polar effects in the Δ*csrA* strain. Since, no major changes in transcript level were observed upon *csrA* deletion in *C. jejuni*, a more direct coIP based approach was used to identify the direct targets of CsrA.

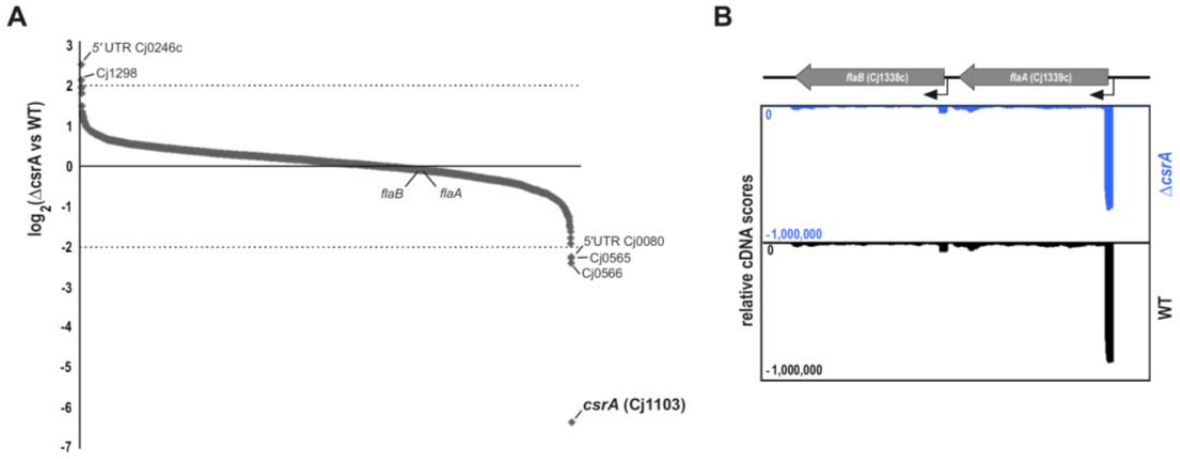


Figure 20. Changes in transcriptome upon *csrA* deletion in *C. jejuni* strain NCTC11168. (A) Changes in transcript reads mapped to ORFs or 5' UTRs in the Δ*csrA* library compared to the WT library. Each point represents an ORF or 5' UTR arranged in order of fold-change. **(B)** RNA-seq reads from the WT (black) and Δ*csrA* (blue) libraries mapped to the *flaA* and *flaB* encoding region.

3.2.3. RIP-seq revealed direct mRNA targets of CsrA in *C. jejuni*

To globally identify direct *C. jejuni* CsrA targets in an unbiased manner and uncover potential sRNA antagonists of this regulator, a RIP-seq approach was applied. RIP-seq involves epitope-tagging of the protein of interest and subsequent co-immunoprecipitation (coIP) of associated RNAs, which nature is usually identified by RNA/deep-sequencing [12, 101]. This method has previously been successfully adapted to identify RNA targets of the ribosomal protein S1 in *H. pylori* [101]. The *csrA* (Cj1103) gene was chromosomally 3xFLAG-tagged at the C-terminus in *C. jejuni* strains NCTC11168 and 81-176. The CsrA-3xFLAG protein is constitutively expressed during growth in rich medium (Fig. 21B). coIPs were performed on mid-exponential-phase (OD_{600}=0.6) lysates of the *csrA*-3xFLAG and, as control, the untagged wild-type (WT) strains using an anti-FLAG antibody. Western blot analysis confirmed successful pull-down of CsrA-3xFLAG in both strains

(Fig. 21C and Appendix Fig. 9A). After conversion of co-purified RNAs into cDNA libraries and deep sequencing thereof, between 93.2 % and 95.8 % of the 4.6 to 6.2 million sequencing reads for the individual coIP libraries could be mapped to their respective genomes (Appendix Table 9). Most of the reads from the NCTC11168 control and CsrA-3xFLAG coIP libraries were mapped to abundant classes of RNA (rRNA, tRNA, and housekeeping RNAs). These RNAs were presumably non-specifically pulled down in the RIP-seq approach (Fig. 21D and Table 4). Nevertheless, about 36-fold and about 5-fold enrichment for reads mapped to 5' UTR and ORF regions was observed in the CsrA-3xFLAG coIP library, respectively, indicating a specific enrichment for mRNAs by CsrA (Fig. 21D). Surprisingly, no specific enrichment for sRNAs was detected, suggesting the absence of CsrB/C like sRNA antagonists in *C. jejuni*. In addition, none of the mRNAs regulated upon csrA deletion in the transcriptome were enriched in the CsrA-3xFLAG coIP library (Appendix Table 10). As the RIP-seq approach from strain 81-176 showed a similar enrichment pattern (Appendix Fig. 9B), the following sections focuses on the analysis of NCTC11168 RIP-seq data.

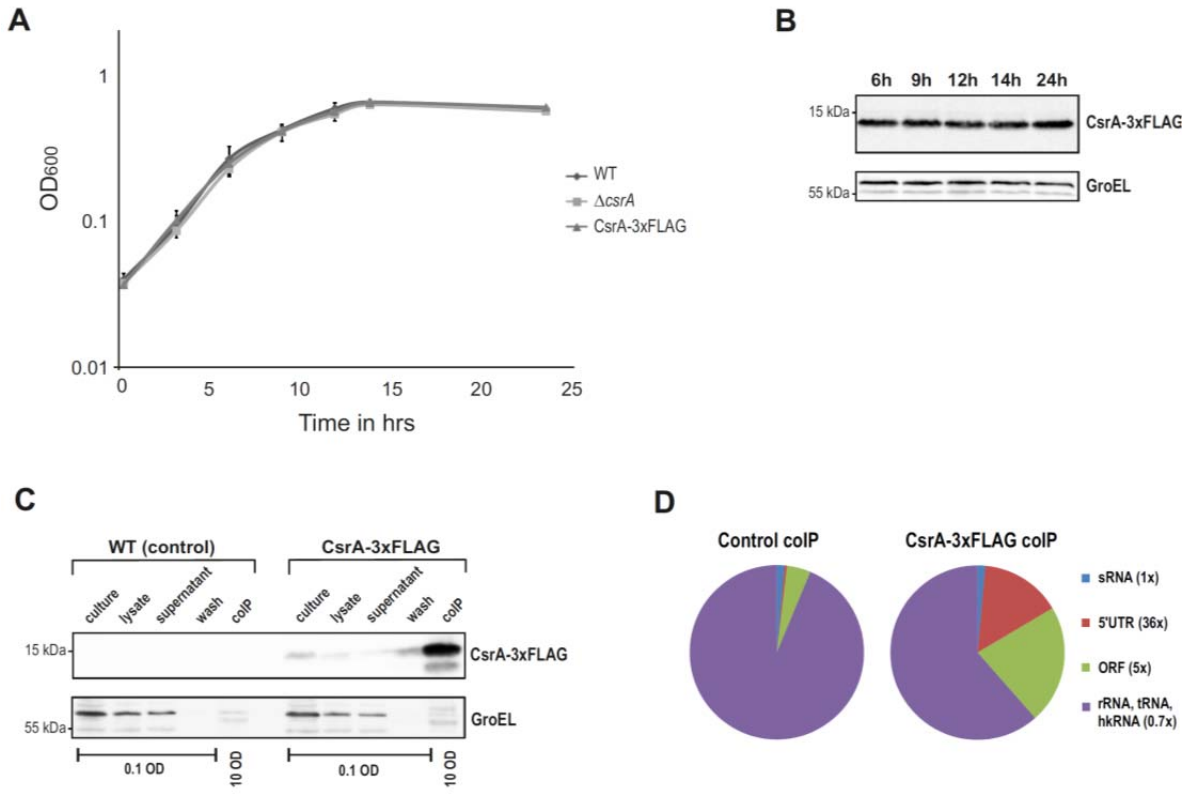


Figure 21. Growth curves and RIP-seq of CsrA in *C. jejuni* strain NCTC11168. (A) Semi-log growth curves over 24 h for *C. jejuni* NCTC11168 wild-type (WT), Δ*csrA* and CsrA-3xFLAG tagged strains grown in Brucella broth in duplicate. Error bars are mean ± s.e.m. **(B)** Western blot analysis of CsrA-3xFLAG expression during growth in liquid culture. Total protein samples corresponding to an OD_{600} of 0.05 were loaded for different time points. GroEL served as loading control. **(C)** Western blot analysis of coIP samples of *C. jejuni* NCTC11168 WT and CsrA-3xFLAG using anti-FLAG antibody confirmed a successful CsrA-3xFLAG pulldown in the tagged strain background. The amount of samples loaded (OD_{600} of bacteria) is indicated. GroEL served

as loading control. **(D)** Pie charts showing relative proportions of mapped cDNA reads of different RNA classes in the coIP libraries (hkRNAs: housekeeping RNAs). Numbers in brackets indicate the relative enrichment of the respective RNA class in the CsrA-3xFLAG vs. control coIPs.

Table 4. Distribution of mapped reads to different RNA classes. The table indicates the total number of mapped reads overlapping annotations for different RNA classes (sRNAs, 5' UTRs, ORFs, rRNAs, tRNAs, housekeeping RNAs, and pseudogenes) with a minimum overlap size of 10 nt. Absolute read numbers and percentage values based on the total number of reads overlapping all annotations are shown for all libraries.

	C. jejuni **NCTC11168 Control coIP**		*C. jejuni* **NCTC11168 CsrA-3xFLAG coIP**	
sRNAs	28,720.50	0.50%	25,507.28	0.51%
5' UTRs	23,806.78	0.41%	756,447.88	15.08%
ORFs	249,394.84	4.30%	1,092,168.37	21.77%
rRNAs	3,285,700.41	56.66%	1,593,895.71	31.78%
tRNAs	1,509,545.38	26.03%	1,191,322.49	23.75%
housekeeping RNAs	699,295.33	12.06%	351,895.93	7.02%
pseudogenes	2,893.90	0.05%	4,752.24	0.09%
total	5,799,357.14	100.00%	5,015,989.91	100.00%

	C. jejuni **81-176 Control CoIP**		*C. jejuni* **81-176 CsrA-3xFLAG CoIP**	
sRNAs	32,282.87	0.58%	22,503.67	0.53%
5' UTRs	23,432.45	0.42%	187,578.08	4.40%
ORFs	344,388.16	6.16%	509,091.48	11.94%
rRNAs	3,641,665.55	65.09%	2,630,055.93	61.69%
tRNAs	1,267,744.45	22.66%	698,004.19	16.37%
housekeeping RNAs	284,913.59	5.09%	215,899.56	5.06%
total	5,594,427.06	100.00%	4,263,132.91	100.00%

3.2.4. CsrA primarily binds mRNAs involved in flagella formation

To investigate if any functional pathways were significantly affected by CsrA in *C. jejuni*, the 154 most highly-enriched mRNAs with >5-fold enrichment in the CsrA-3xFLAG coIP were analyzed (Appendix Table 10). Functional enrichment analysis revealed an over-representation of mRNAs from the class "Surface Structures" (Fisher's exact test; Benjamini-Hochberg adjusted P-value <0.05). This functional class also includes the genes involved in flagellar biosynthesis and motility (Fig. 22).

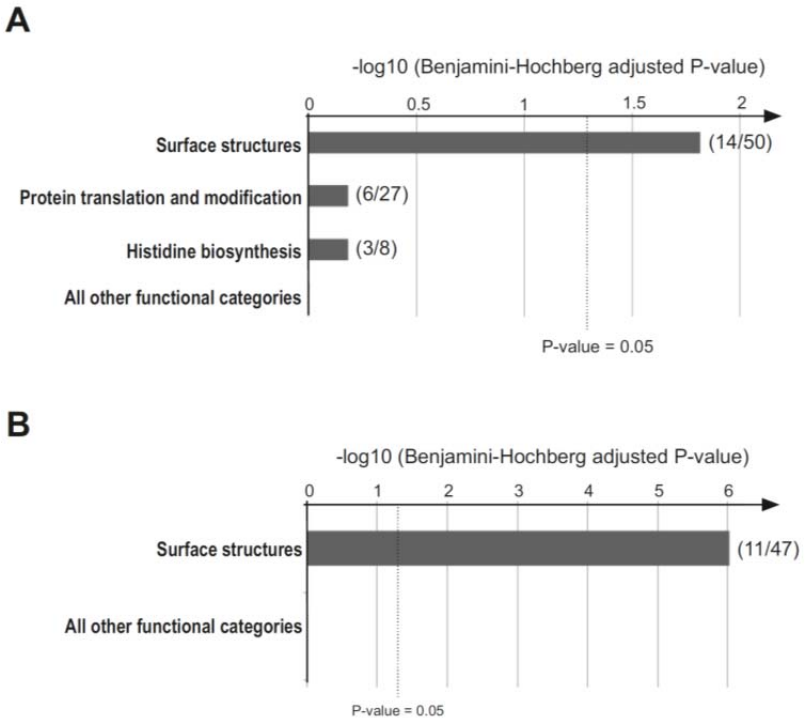


Figure 22. Statistical analysis for overrepresentation of functional categories among potential CsrA target genes. (A), (B) Statistical analysis was performed on all genes with more than 5-fold enrichment in their 5' UTR and/or coding sequence in the CsrA-3xFLAG coIP in comparison to non-enriched genes (control coIP). P-values, adjusted using the Benjamini-Hochberg method, were calculated for over-representation of enriched genes in each functional category in *C. jejuni* strains NCTC11168 **(A)** and 81-176 **(B)**. Values in brackets denote the number of enriched genes in the particular functional category (only functional categories with non-zero $\log_{10}$ P-values are shown individually for clarity). Functional categories are based on reannotation of the NCTC11168 genome [102].

In fact, 90% of the reads mapping to the >5-fold enriched CsrA targets belonged to flagella- or motility-related genes (Fig. 23A). *C. jejuni* produces amphitrichous flagella *i.e.* a single flagellum at each pole. The flagellum is a big and complex molecular structure composed of several proteins providing both the structural and functional attributes. Receptors at the cell surface sense nutrients or other molecules of interest and relay the information to the flagella which can spin at several hundred Hz in both clockwise and counterclockwise direction to effectively produce a desired motion [103]. All these proteins have to be arranged in a very ordered and timely fashion to build a functional flagellum. In *C. jejuni*, this hierarchical arrangement is largely determined at the transcription level by the activity of alternate sigma factors RpoN (σ^{54}) and FliA (σ^{28}). Unlike the filament of the widely studied organisms like *E. coli*, *Salmonella* or *Bacillus* which is composed of a single type of flagellin monomer, the filament of *C. jejuni* flagellum is composed of two different subunits – major (FlaA) and minor (FlaB) flagellin. *flaA* and *flaB* genes are encoded next to each other but are transcribed separately from FliA and RpoN dependent promoters, respectively. FlaA and FlaB have very similar amino acid composition and the role of FlaB is not

clear as its deletion still leads to a functional flagellum with comparable motility [104]. Most of the enriched flagellar genes were either class 2 or class 3 genes (Table 5 and Fig. 23B). For example, the 5' UTRs of *flaG* (346-fold), *flgH* (200-fold), and *flgI* (170-fold) (Table 5) were highly enriched. The most abundantly co-purified transcript (>300-fold enrichment) was the leader of *flaA* mRNA, encoding the major flagellin (Fig. 23A), indicating *flaA* mRNA is the major CsrA target (Fig. 23A).

Table 5. Enrichment of genes involved in flagellar biosynthesis in the CsrA-3xFLAG coIP data. Classification of flagella genes is based on Ref. [105]. Transcripts with >5-fold enrichment in cDNA read counts in the CsrA-3xFLAG vs. control coIP libraries are highlighted in bold. Numbers in brackets indicate the absolute cDNA read counts in the CsrA-3xFLAG coIP libraries.

	***C. jejuni* NCTC11168**		***C. jejuni* 81-176**	
Enrichment (Reads)	**5' UTR**	**ORF**	**5' UTR**	**ORF**
Regulation of expression (class 1)				
***rpoN* (Cj0670)**	1.5x (26)	**25x (1747)**	-	**8x (634)**
fliA (Cj0061c)	-	1.1x (121)	-	0.7x (79)
flgS (Cj0793)	-	1.7x (43)	1.3x (1)	1.7x (15)
flgR (Cj1024)	1.2x (4)	1.1x (137)	1.3x (4)	1.2x (101)
Flagellar protein secretion (class 1)				
***flgM* (Cj1464)**	-	**5x (1429)**	-	**2.3x (380)**
***fliF* (Cj0318)**	-	**6.1x (1005)**	-	**7.8x (1105)**
flhA (Cj0882c)	-	1.2x (82)	-	0.9x (53)
flhB (Cj0335)	0.6x (3)	1.2x (99)	0.7x (1)	1.0x (67)
fliO (Cj0352)	-	1.4x (41)	-	0.7x (1)
fliP (Cj0820c)	-	1.2x (29)	-	0.5x (15)
fliQ (Cj1675)	-	1.3x (45)	-	0.9x (29)
fliR (Cj1179c)	-	1.2x (6)	-	0.5x (5)
fliH (Cj0320)	-	4.8x (202)	-	1.3x (100)
fliI (Cj0195)	-	3.6x (442)	-	0.6x (86)
Basal body components (class 1 and 2)				
fliE (Cj0526c)	-	2.4x (268)	-	1.4x (120)
flgC (Cj0527c)	-	1.2x (397)	-	0.9x (217)
flgB (Cj0528c)	1.6x (7)	1.6x (181)	0.7x (2)	0.9x (61)
***flgG2* (Cj0697)**	-	**43.9x (8133)**	-	**77.4x (9670)**
***flgG* (Cj0698)**	1.2x (1)	1.4x (253)	**5.3x (4)**	1.3x (165)
flgJ (Cj1463)	-	4.4x (180)	-	1.0x (26)
***flgI* (Cj1462)**	**170.5x (5750)**	**52.7x (12087)**	**157.1x (1666)**	**61.5x (5401)**
***flgA* (Cj0769c)**	0.8x (2)	**15.8x (410)**	0.9x (4)	**3.7x (104)**
***flgH* (Cj0687c)**	**200.9x (1911)**	**20.1x (3288)**	**110.1x (917)**	**27.6x (2487)**
Flagellar hook components (class 2)				
***flgE* (Cj1729c)**	-	**68.2x (104324)**	-	**7.1x (3967)**
flgD (Cj0042)	-	3.6x (1015)	-	2.0x (284)
flgE2 (Cj0043)	-	2.5x (1045)	-	1.1x (250)
fliK (Cj0041)	-	4.1x (613)	-	1.2x (89)
Cj0040*	**356.2x (3389)**	**110.6x (9277)**	**38x (230)**	**20.4 x (727)**
flgK (Cj1466)	-	0.7x (4)	-	1.0x (141)
flgL (Cj0887c)	-	2.0x (484)	2.3x (74)	0.9x (193)
Flagellar filament components (class 2 and 3)				
***flaA* (Cj1339c)**	**304.5x (693471)**	**111x (473588)**	**324.7x (158590)**	**45.3x (138159)**
***flaB* (Cj1338c)**	**58.8x (915)**	**14.1x (17880)**	**59.4x (1170)**	**14.9x (29530)**
***fliD* (Cj0548)**	-	**6.8x (4348)**	-	**5.4x (3929)**
fliS (Cj0549)		1.6x (149)		1.3x (165)
flaC (Cj0720)	1.2x (344)	1.2x (1298)	1.3x (239)	1.2x (1237)
Other enriched genes (>5x) involved in flagella formation				
***pseB* (Cj1293)**	**119.7x (2298)**	**9.5x (2280)**	**34.5x (470)**	4.1x (759)
***pseI* (Cj1317)**	1.7x (22)	**7.2x (864)**	2.3x (14)	0.8x (111)
***flaG* (Cj0547)**	**346.1x (11077)**	**72.4x (18150)**	**168.5x (3701)**	**84.2x (16012)**
***motA* (Cj0337c)**	**10.3x (89)**	1.8x (660)	1.4x (16)	0.8x (271)
Cj0951c	-	**15.2x (79)**	2x (3)	1.3x (194)
Cj0248	**5.5x (120)**	1.8x (387)	1.1x (38)	0.9x (257)
***flhX* (Cj0848c)**	-	**7.5x (13)**	-	1.5x (7)

*Cj0040 (unknown function) is the first gene of the operon encoding the hook genes.

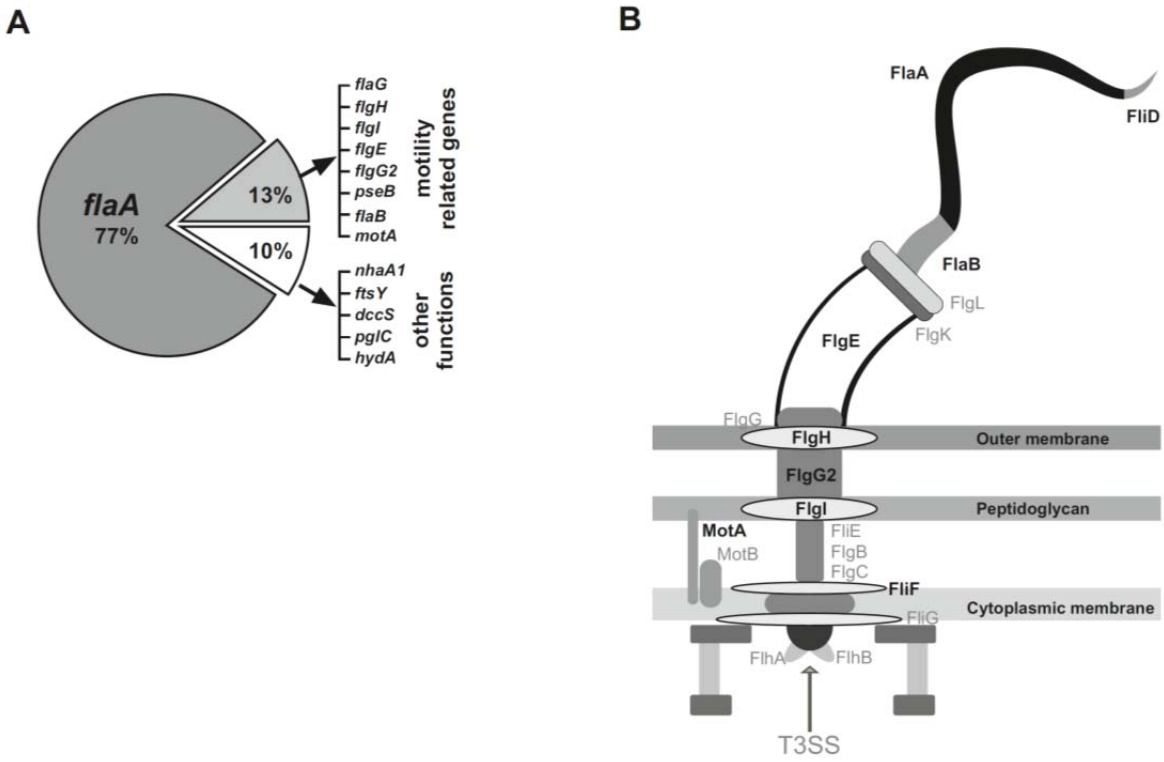


Figure 23. Flagellar mRNA targets of CsrA. (A) Pie chart showing the percentages and enriched genes of mapped reads for all >5-fold enriched CsrA target genes. **(B)** Schematic representation of structural components of the *Campylobacter* flagellum. Proteins encoded by mRNAs which showed >5-fold enrichment in the CsrA-3xFLAG coIP are marked in black and bold. T3SS: Type III secretion system.

3.2.5. CsrA binds diverse sites sites within target mRNAs

The coIP experiment revealed 154 mRNAs with >5-fold enrichment for 5' UTR- or ORF-mapped cDNA counts in the CsrA-3xFLAG vs. control coIP (Appendix Table 10). For 17 of these genes, both the 5' UTR and coding sequences (CDS) were co-purified, with some having potentially shared CsrA binding sites. Visual inspection of the cDNA read-patterns in a genome browser showed that numerous flagellar mRNAs, including *flaA*, *flaG*, and *flgI* (encoding the major flagellin, a gene involved in flagellum formation, and a P-ring component, respectively) showed strong enrichment at their 5' UTRs (Fig. 24B). Strikingly, CsrA binding was also observed between genes in polycistronic mRNAs, such as the Cj0310c-Cj0309c operon, encoding two paralogous efflux proteins, and Cj0805-*dapA*, encoding a zinc protease and dihydrodipicolinate synthase (Fig. 25B). In some cases, CsrA binding was also observed exclusively in the CDS of gene, such as *ftsY* (Fig. 25C)

3.2.6. Validation of *C. jejuni* CsrA targets using an *E. coli* reporter system

As a first screen for validation of the coIP data and investigation of regulatory consequences of CsrA binding to potential targets mRNAs, an *E. coli* translational reporter system was employed

which was originally developed to study sRNA-mediated regulation [106]. Compared to slow and tedious cloning in *C. jejuni*, this system allows for a fast and easy cloning of reporter fusions and also allows for analyzing regulation of genes within operons. Thus, putative CsrA targets from *C. jejuni* strain NCTC11168 were fused in-frame to GFP or a FLAG-*lacZ'* tag (Fig. 24A and 25A), and reporter expression in *E. coli* was compared in the presence and absence of endogenous CsrA. Deletion of *csrA* dramatically enhanced biofilm formation and poor growth in liquid culture in the *E. coli* strain, presumably due to dysregulation of the *pgaABCD* genes required for polysaccharide adhesin synthesis [107]. Accordingly, experiments with the reporter fusions were performed in an *E. coli* Δ*pgaA* background which restores wild type growth upon *csrA* deletion.

First, the effect of CsrA on translation of six putative targets with 5' UTR CsrA binding sites was examined, *flaA*, *flaG*, *flgI*, *flab, pseB* and Cj1249. The 5' UTR and first N-terminal codons (e.g. 33 aa for *flaA*) were fused in-frame to GFP in plasmid pXG10 (Fig. 24A). Expression of GFP reporter levels in the *E. coli* Δ*pgaA* vs Δ*pgaA/*Δ*csrA* strains, measured by Western blot and FACS, showed up-regulation of FlaA-GFP, FlaG-GFP, FlgI-GFP, FlaB-GFP, and PseB-GFP fusion proteins upon deletion of *csrA,* suggesting that CsrA represses translation of these genes (Fig. 24B and Appendix Fig. 10). Complementation of the *E. coli* Δ*pgaA/*Δ*csrA* mutant with a plasmid expressing *C. jejuni* CsrA-Strep from an arabinose inducible promoter (pBAD) restored repression of these target genes, indicating conserved functions of and target specificity of *C. jejuni* and *E. coli* CsrA.

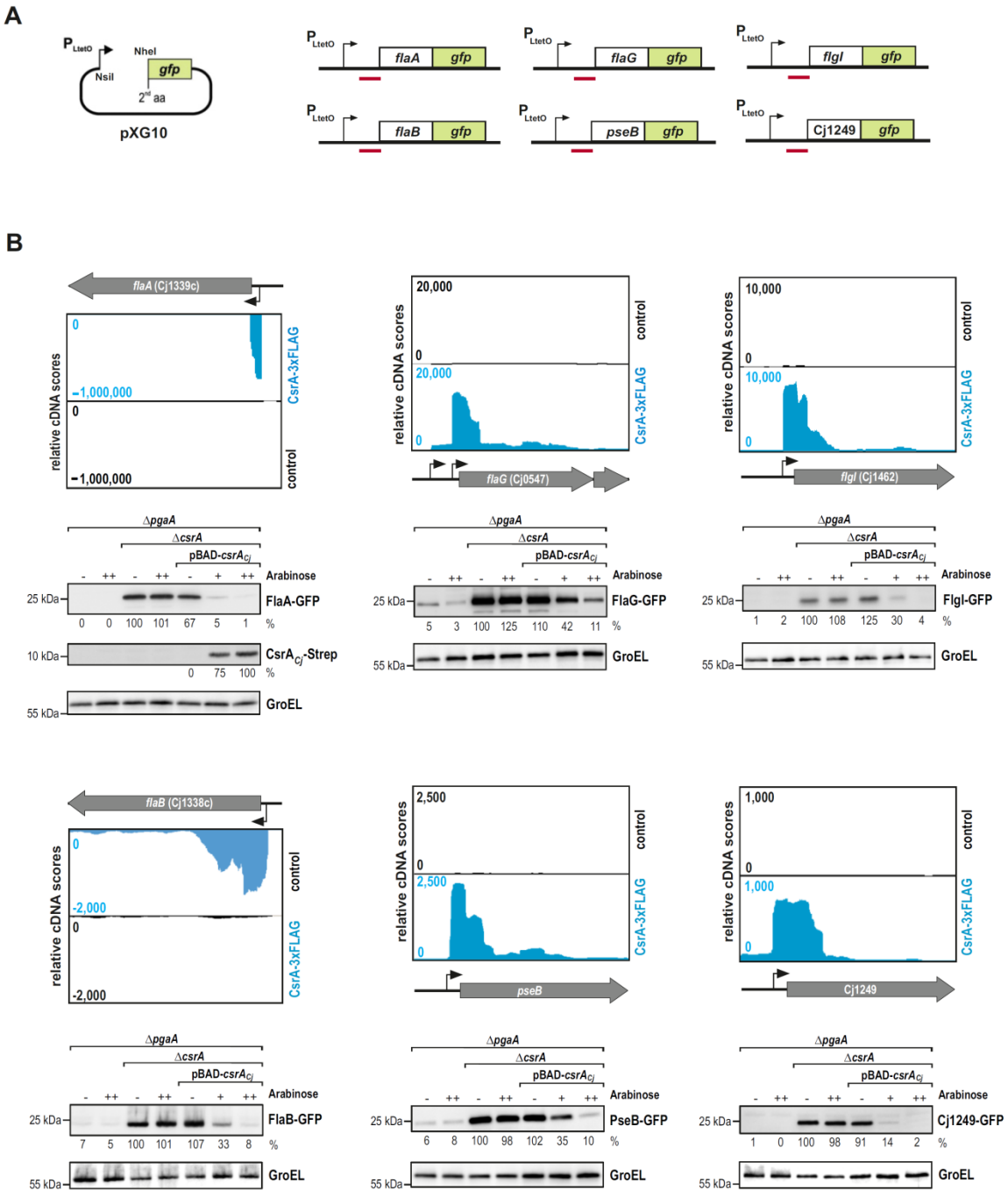


Figure 24. Validation of potential 5' UTR CsrA targets using a GFP reporter system in *E. coli*. (A) Overview of pXG10-derived GFP reporter plasmids [106] used to validate 5' UTR targets. For putative targets in 5' UTRs, the entire mRNA leader as well as the first few codons were cloned as a translational fusion to *gfp* in low-copy vector pXG10 (pSC101* origin). Fusions to *gfp* were made downstream of the +1 (TSS) site of the P_L promoter and were transcribed from a constitutive λ $P_{LtetO-1}$ promoter (P_L derivative). **(B)** *(Top panels)* Examples of enrichment patterns indicating potential CsrA binding sites in 5' UTRs (*flaA*, *flaG*, *flgI, flaB, pseB,* and Cj1249 mRNAs) that were tested in the *E. coli* reporter system. Mapped cDNA reads are shown for the control (black) and CsrA-3xFLAG coIP (blue) libraries in *C. jejuni* strain NCTC11168. ORFs are indicated by grey arrows and TSS - based on dRNA-seq [97] - by black arrows, respectively. *(Lower panels)* Western blot analysis using anti-FLAG, anti-GFP and anti-Strep antibodies of reporter translational fusions to potential *C. jejuni* CsrA target genes in *E. coli* Δ*pgaA*, Δ*pgaA*/Δ*csrA,* and Δ*pgaA*/Δ*csrA* complemented with

plasmid pGD72-3 carrying arabinose-inducible *csrA*-Strep from *C. jejuni.* Strains were grown to late exponential phase in LB medium or LB media supplemented with 0.001% (+) or 0.003% (++) L-arabinose.

Next, the regulatory effect of CsrA binding between two genes in the same polycistron (Cj0310c-Cj0309c and Cj0805-*dapA*) was examined (Fig. 25, and Appendix Fig. 10). The C-terminus of the first gene (Cj0310c or Cj0805) was fused to FLAG-*lacZ',* and the N-terminus of the second gene (Cj0309c or *dapA*) was fused to GFP (Fig. 25A). While the upstream genes (FLAG-Cj0310c and FLAG-Cj0805) were slightly or not affected by *csrA* deletion, the downstream fusions (Cj0309c-GFP and DapA-GFP) showed five-fold or higher levels in Δ*csrA,* and the levels were restored upon complementation with *C. jejuni* CsrA, suggesting that CsrA binding between genes can mediate discoordinate operon regulation (Fig. 25B). Since the potential Shine-Dalgarno sequences are located just at the end of the upstream genes and are covered by the CsrA target sites, CsrA probably interferes with ribosome binding and translation of the downstream genes.

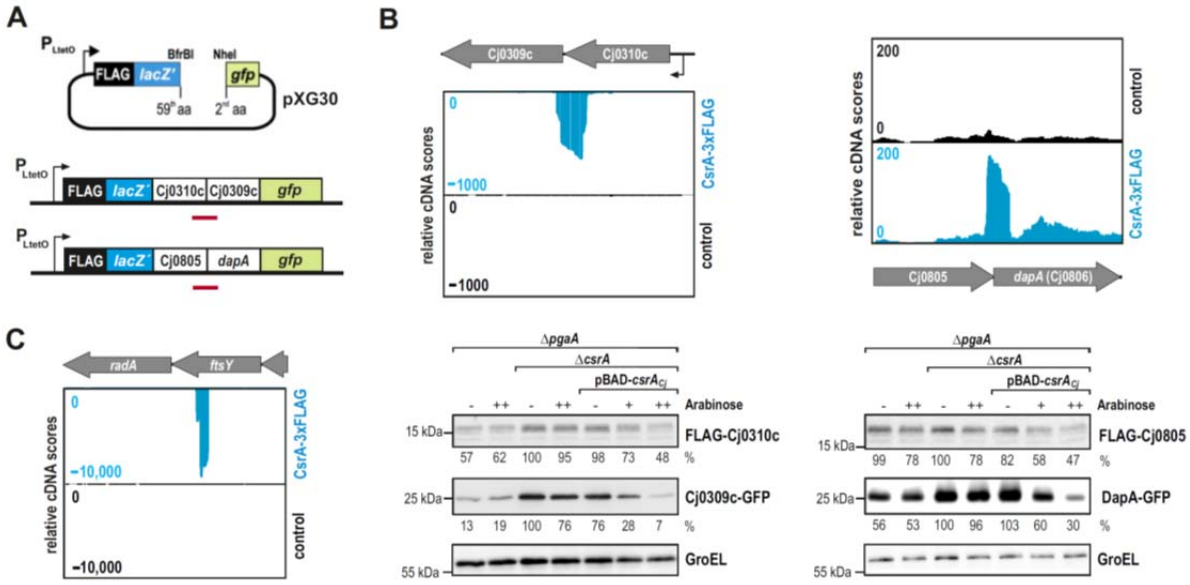


Figure 25. Validation of potential intergenic CsrA targets using a dual reporter system in *E. coli*. (A) Overview of pXG30-derived GFP reporter plasmids [106] used to validate CsrA targets between genes in polycistronic transcripts. The C-terminus of the upstream ORF was fused in frame after a short artificial reading frame composed of a FLAG epitope and truncated *lacZ* gene, whereas the N-terminus of the downstream gene is fused in frame to *gfp*, thus mimicking operon mRNA expression. Putative CsrA binding regions are marked by a red line. **(B)** *(Top panels)* Examples of enrichment patterns indicating potential CsrA targets between genes in polycistronic transcripts (Cj0805-*dapA* and Cj0310c-Cj0309c) that were tested in the *E. coli* system. *(Lower panels)* Western blot analysis using anti-FLAG and anti-GFP antibodies of reporter translational fusions to potential *C. jejuni* CsrA target genes in *E. coli* Δ*pgaA*, Δ*pgaA*/Δ*csrA,* and Δ*pgaA*/Δ*csrA* complemented with plasmid pGD72-3 carrying arabinose-inducible *csrA*-Strep from *C. jejuni.* Strains were grown to late log phase in LB medium or LB media supplemented with 0.001% (+) or 0.003% (++) L-arabinose. **(C)** Example of enrichment pattern indicating potential CsrA binding within the ORF (ftsY gene). Mapped cDNA reads are shown for the control (black) and CsrA-3xFLAG coIP (blue) libraries in *C. jejuni* strain NCTC11168. ORFs are indicated by grey arrows and TSS - based on dRNA-seq [97] by black arrows, respectively.

3.2.7. Automated peak-detection reveals a CsrA binding motif

To automatically identify CsrA binding regions and predict a CsrA binding motif from cDNA enrichment patterns in the coIPs, a peak-detection algorithm based on a sliding window approach was developed. Multiple parameters were established by iterative improvements after several rounds of peak detection accuracy by visual inspection (window size 25 bp, relative minimum expression factor of 3). Using these parameters, 328 potential CsrA binding sites with more than 5-fold enrichment in the *C. jejuni* NCTC11168 coIP were predicted (Appendix Table 11). As a control, peak-detection was performed in reverse manner by scanning for enriched regions in the control vs. CsrA-3xFLAG coIP. This revealed only five peaks without a common motif, indicating a high specificity of the peaks detected in the CsrA-3xFLAG coIP. Screening of the 328 enriched sequences using MEME [63] revealed a (C/A)A(A/T)GGA motif with high confidence (E-value = 2.1E-11) in 324/328 input sequences (Fig. 26A). Analysis of the *C. jejuni* 81-176 coIP produced a similar motif (Appendix Fig. 9C). In order to check if a similar motif can be found in non-enriched regions, the peak-detection was conducted in reverse manner but using a cutoff of only >1-fold enrichment in the control library versus the CsrA-3xFLAG coIP library. This control approach revealed 448 sites "enriched" in the control library. Subsequent motif prediction using the same parameters did not yield any significant or meaningful motifs, further supporting the specificity of the coIP approach and the detected CsrA binding motif. Screening of the 328 enriched CsrA-3xFLAG coIP sequences for consensus-structure motifs using CMfinder [108] with default parameters revealed 276/328 input sequences have an AAGGA motif in the loop of a hairpin structure (Fig. 26B). The identified *C. jejuni* CsrA sequence/structural motifs agree well with binding sites of other CsrA homologs [109], demonstrating a conserved CsrA binding motif.

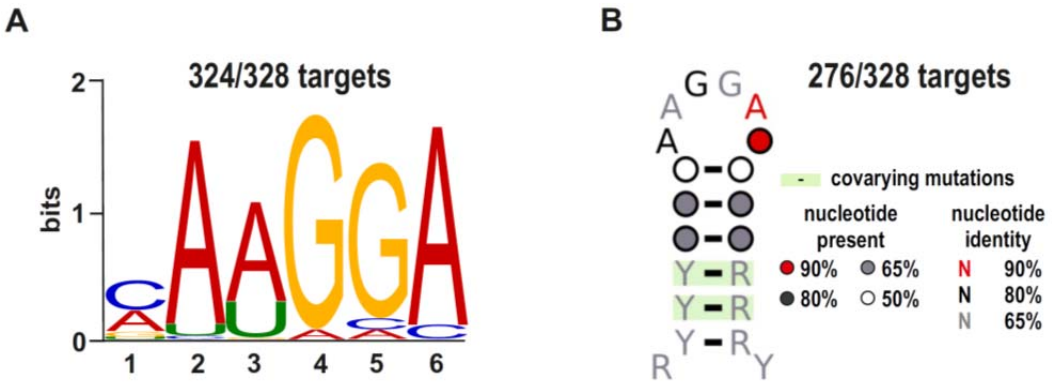


Figure 26. CsrA binding motif analysis using sequence extracted from enriched peaks. (A) CsrA binding motif predicted using MEME[63] (E-value = 2.1E-11). **(B)** Consensus secondary structure motif of *C. jejuni* CsrA binding sites predicted by CMfinder[108].

3.2.8. *flaA* mRNA is translationally repressed by CsrA

Expression of flagellar components is hierarchically controlled at the transcriptional level to ensure proper assembly of the *C.jejuni* flagellum. The filament, consisting mainly of FlaA, is among the last components produced in this cascade. In the *C. jejuni* coIP dataset, 77 % of the mapped reads from more than five-fold enriched genes (Fig. 23A) corresponded to *flaA,* indicating that it is the major CsrA target. Secondary structure predictions using Mfold [110] revealed the 45 nt-long 5' UTR of *flaA* can fold into two stem-loops (SL1 and SL2), both of which harbor an ANGGA motif in their loops (Fig. 27 and 28A). The second ANGGA motif is part of the RBS of the *flaA* mRNA and a third GGA is present as the second *flaA* codon. The *flaA* 5' UTR secondary structure is conserved and supported by compensatory base-pair changes in other *Campylobacter* species (Fig. 27 and Appendix Fig. 11).

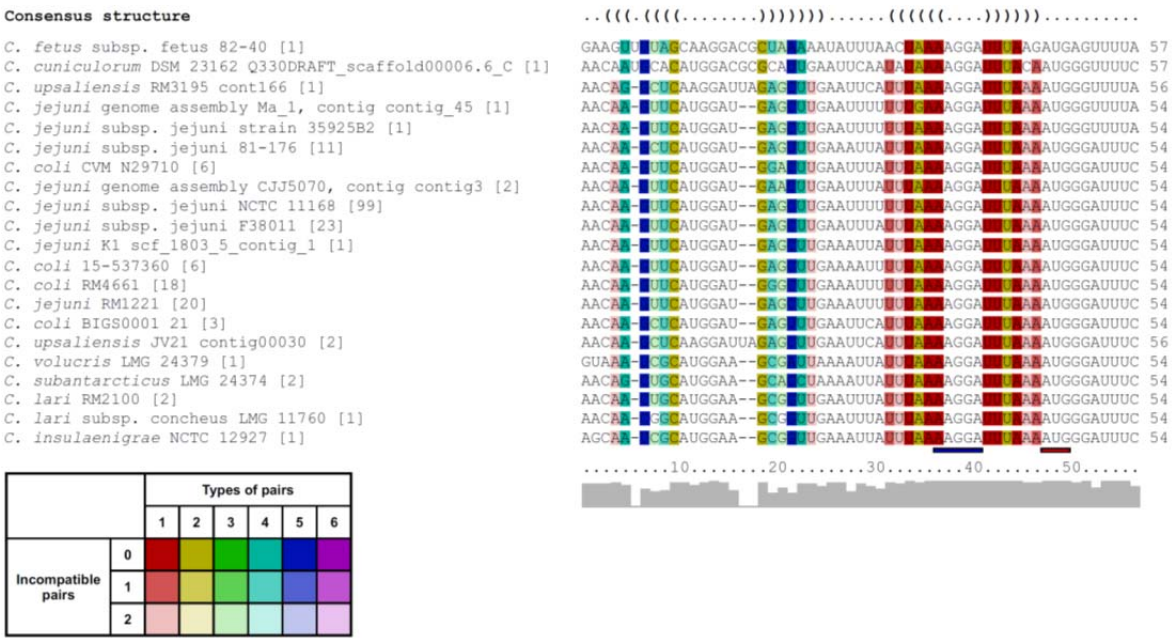


Figure 27. Structure annotated sequence alignment of the *flaA* 5' UTR. Structure-annotated sequence alignment based on representative unique sequences of the *flaA* 5' UTR and the first 10 nt of the coding region from different *Campylobacter* species. The numbers in square brackets indicate how often an identical sequence occurred in the original complete alignment (see Appendix Fig. 11). The consensus structure is shown in bracket-dot notation at the top and sequence conservation as a bar chart at the bottom. The color legend was adapted from the RNAalifold online help (http://rna.tbi.univie.ac.at/help.html) and reflects how many different kinds of nucleotide combinations support a certain base pair. Here, paler colors indicate that a base pair cannot be formed by all sequences in the alignment. The RBS and start codon of the *flaA* mRNA are indicated by blue and red bars below the alignment, respectively.

To determine if CsrA affected *flaA* translation, the C-terminus of *flaA* mRNA was tagged with a 3xFLAG epitope at its native locus in *C. jejuni* strain NCTC11168. FlaA-3xFLAG was roughly three-

fold up-regulated in a Δ*csrA* strain on Western blots, indicating translational control by CsrA (Fig. 28B). To show that CsrA affected translation by binding to the *flaA* mRNA leader, point-mutations into the two putative GGA binding motifs (M1: $SL1_{GGA->AAA}$, M2: $SL1_{GGA->UGA}$, and M3: $SL2_{GGA->GGG}$) were introduced (Fig. 28A and B, lanes 3-8). Like deletion of *csrA*, mutation of the GGA motifs resulted in two- to three-fold elevated FlaA-3xFLAG protein levels. FlaA-3xFLAG expression was not affected by deletion of *csrA* in the *flaA* leader mutants, indicating that CsrA binding was abolished in these strains. In line with no change observed for *flaA* mRNA levels upon CsrA deletion in the transcriptome dataset, Northern blot analysis also showed *flaA*-3xFLAG mRNA levels are only mildly affected in the different mutant strains, further indicating post-transcriptional regulation and translational repression of *flaA* expression by CsrA (Fig. 20B and Fig. 28B).

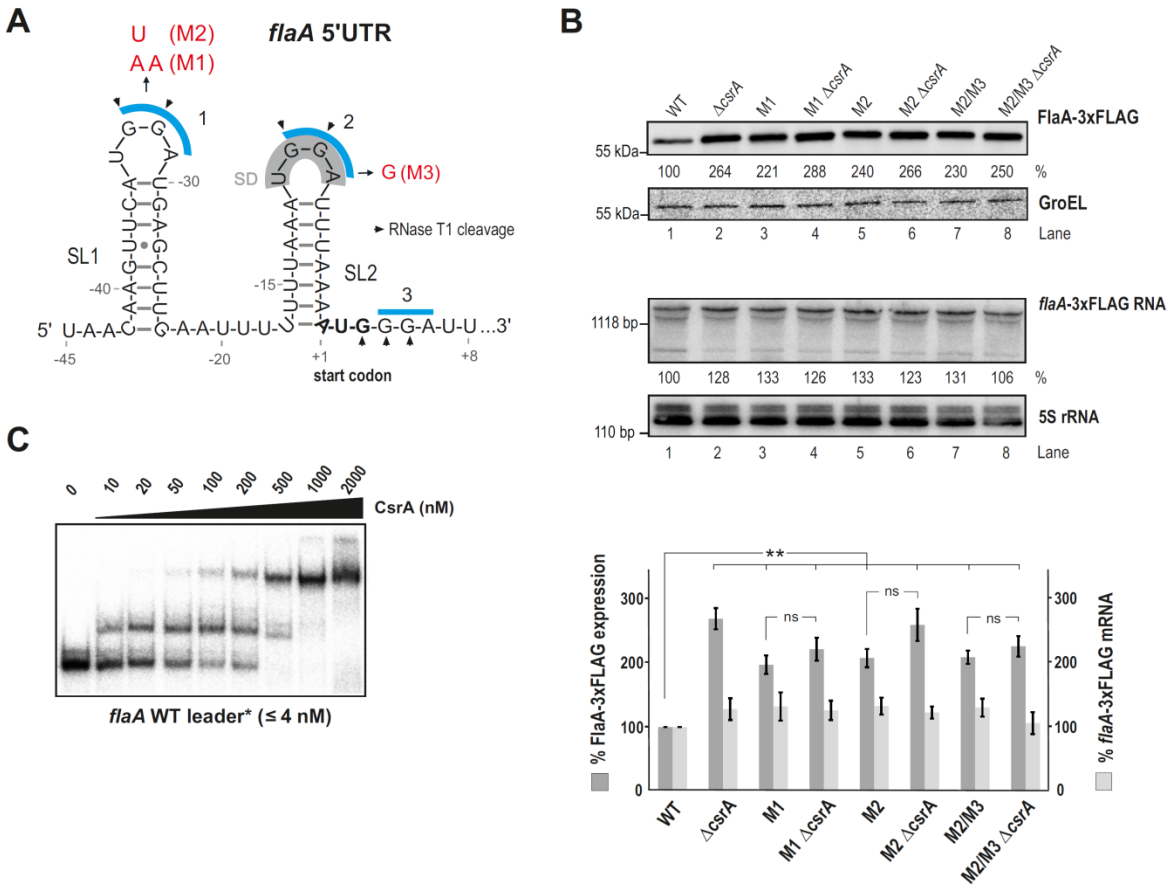


Figure 28. CsrA represses *flaA* translation by binding to its 5' UTR. (A) Predicted secondary structure of the *flaA* mRNA leader using Mfold [110]. Blue bars indicate GGA motifs; grey: SD sequence. Black triangles indicate RNase T1 cleavages from the structure probing in Fig. 30**. (B)** (*Top panel*) Representative Western and Northern blot of FlaA-3xFLAG and its mRNA in various *flaA* 5' UTR mutant strains from liquid cultures in log growth phase. GroEL and 5S rRNA served as loading control for the Western and Northern blots, respectively. (*Bottom panel*) Western blot quantification (n=5) of FlaA with a C-terminal 3xFLAG epitope tag integrated at its native locus (FlaA-3xFLAG) and Northern blot analysis of *flaA* mRNA (n=3) in Δ*csrA* and various *flaA* 5' UTR mutant strains. Shown is the mean ± s.e.m (**P<0.01 using Student's *t*-test). Mutations

are depicted in red in **(A)**. **(C)** Gel-shift assays using ~0.04 pmol *in vitro*-transcribed and 5' end-labeled *flaA* leader (-45 to +99 relative to the start codon) with increasing concentrations of CsrA-Strep.

3.2.9. CsrA directly binds to *flaA* 5' UTR

In-vitro gel-shift assays using recombinant *C. jejuni* CsrA-Strep and 5'-end radiolabeled T7-transcribed *flaA* leader variants (region -45 to +99 nt relative to the *flaA* start codon) supported that GGA mutations affect CsrA binding, rather than introducing translational bias. The *flaA* wild-type (WT) leader showed strong binding to CsrA (K_d ~50 nM) with two defined shifts indicating at least two CsrA binding sites (Fig. 28C). In contrast, the *flaA* leaders with GGA point-mutations either in SL1 (M1 and M2), SL2 (M3), or both SL1 and SL2 (M2/M3) showed four- to ten-fold higher K_d values (200 – 500 nM) (Fig. 29A and B).

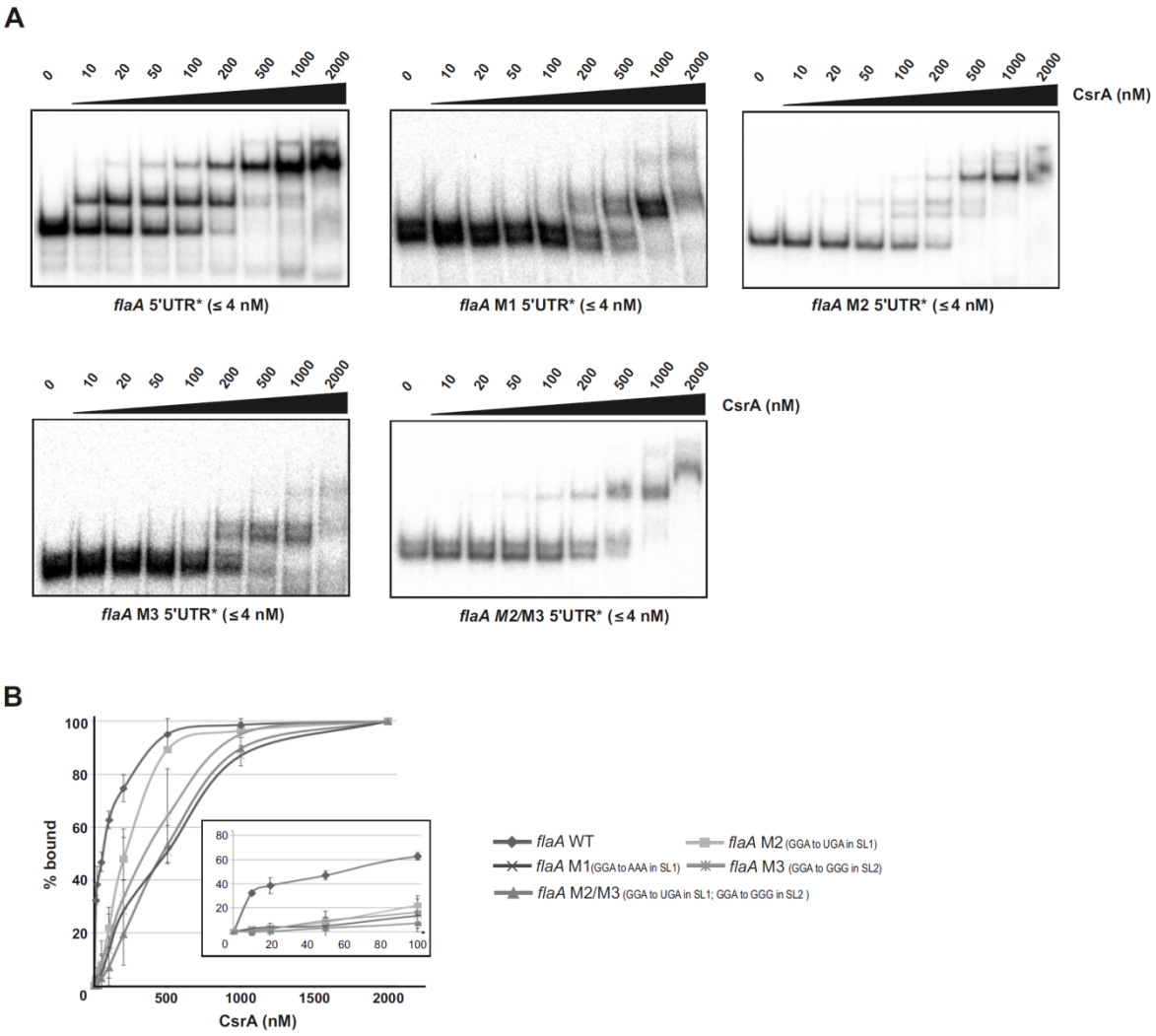


Figure 29. *In vitro* gel-shift assays of 5'-labeled T7-transcripts and purified *C. jejuni* CsrA. (A) Gel-shift assays with 5'-labeled WT *flaA* leader and its mutant variants (M1, M2, M3, and M2/M3) with increasing concentrations of CsrA-Strep. M1: GGA→AAA in stem-loop 1 (SL1, see Fig. 28A); M2: GGA>UGA in SL1; M3:

GGA>GGG in stem-loop 2 (SL2); M2/M3: combination of M2 and M3. **(B)** Affinity binding curves determined by gel-shift assays for ^{32}P-labeled *flaA* WT and mutant leaders (≤4 nM) based on three replicates. The inset represents an enlargement of the binding curves for low CsrA-Strep concentrations. Shown is the mean ± s.d.

To map CsrA binding sites on the *flaA* leader, *in-vitro* footprinting assays with labeled *flaA* leader in the absence or presence of CsrA using enzymatic and chemical cleavage was performed (RNase T1; single stranded G-residues and lead(II) acetate; single-stranded RNA). The cleavage patterns without CsrA confirmed the predicted *flaA* leader structure (Fig. 30). A clear protection was observed at the SL1 and SL2 GGA motifs of the WT leader upon addition of increasing CsrA amounts, but not for a *flaA* M2/M3 mutant in which the motifs were disrupted by point mutations. No protection was observed at the third GGA downstream of the start codon. Overall, the data suggests that *C. jejuni* CsrA represses translation of its major target *flaA* by high-affinity binding to the two GGA-containing stem-loops SL1 and SL2 in the *flaA* mRNA leader.

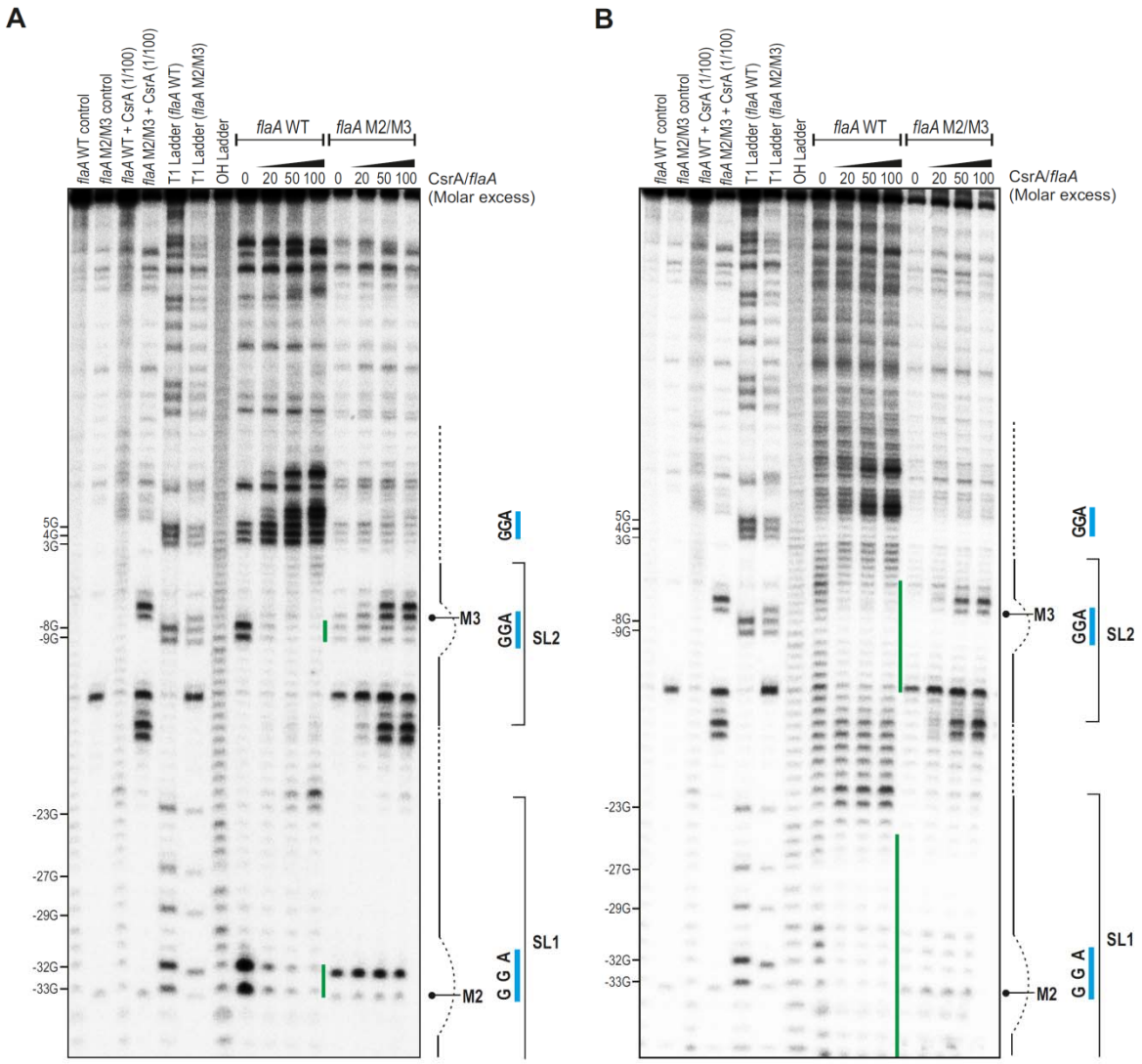


Figure 30. Footprinting assays define CsrA binding sites on *flaA* mRNA leader. Footprinting assays of ~0.2 pmol ^{32}P-labeled *flaA* WT and *flaA* M2/M3 mutant leaders in the absence or presence of increasing CsrA concentrations (molar excess of 0, 20, 50 and 100 CsrA) using RNase T1 **(A)** and lead(II) acetate **(B)**. Untreated *flaA* leader alone or incubated with 100-fold excess of CsrA served as controls and RNase T1- or alkali (OH)- digested *flaA* leader as ladders, respectively. Blue lines: GGA motifs; green lines: protection from RNA cleavage upon addition of CsrA. The secondary structure of the *flaA* leader according to panel is depicted on side of each panel.

3.2.10. The flagellar assembly factor FliW directly binds and antagonizes CsrA

The observed constitutive expression of *C. jejuni* CsrA during routine culture (Fig. 21A and B) suggested modulation of CsrA activity is important rather than its expression. However, apparent homologs of the CsrB/C sRNAs are missing in Epsilonproteobacteria [23, 111]. The more than five-fold enriched targets from the coIP dataset (Appendix Table 10) included CJnc10 sRNA. However, its relatively low abundance and a lack of GGA-motif, make CJnc10 unlikely to be a homolog of the antagonizing sRNAs. However, it was possible that other RNAs or even protein antagonists might control CsrA activity in *C. jejuni*. One candidate protein (Cj1075) is a 129-aa potential homolog of the flagella assembly factor, FliW, which is known to play a role in *C. jejuni* motility [112], but is otherwise uncharacterized. In *B. subtilis*, FliW can directly bind CsrA and antagonizes CsrA-mediated translational repression of *hag* mRNA, encoding the major flagellin [23]. FliW also binds Hag, which accumulates in the cytoplasm prior to flagellar hook completion. Hag thus sequesters FliW from CsrA, allowing CsrA to repress further unnecessary Hag synthesis. Once the hook is assembled, Hag is secreted and released FliW now binds CsrA to relieve CsrA-mediated repression of *hag* translation. While the Hag-FliW-CsrA partner-switch mechanism is thought to ensure appropriate temporal synthesis of flagellin in *B. subtilis*, it remained unclear whether this type of Csr-activity control is more general. *E. coli,* which has the antagonizing sRNAs, lacks FliW. Homologs of *fliW* are present in Epsilonproteobacteria, but unlike *Bacillus* and other bacteria, are not encoded adjacent to *csrA* (Fig. 31).

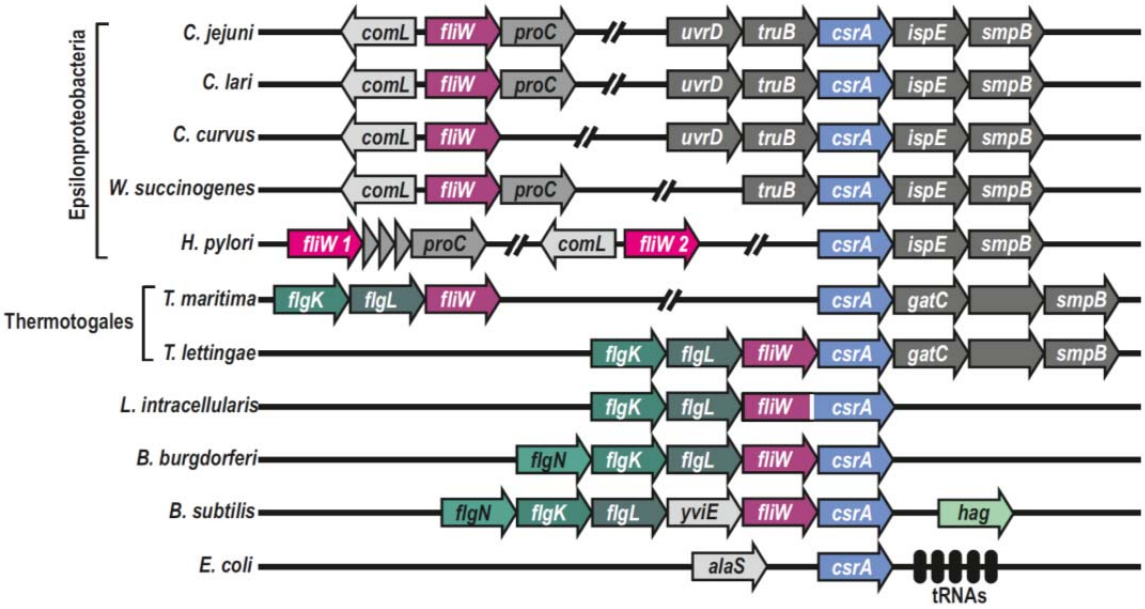


Figure 31. Genomic locations of CsrA and FliW in selected bacterial species/strains. Genomic context of *csrA* and *fliW* homologs in diverse bacterial species (*Campylobacter* spp: *C. jejuni, C. lari, C. curvus*; *Wolinella succinogenes*; *Helicobacter pylori*; Thermotogales: *T. maritima, T. lettingae*; *Lawsonia intracellularis*; *Borrellia burgdorferi*; *Bacillus subtilis*; *Escherichia coli*). Blue: *csrA* homologs; dark or light red: *fliW* homologs; shades of green: flagellar genes.

The Bacterial Two Hybrid (BACTH) system was used to assess CsrA-FliW interaction in an *E. coli* based system. BACTH is based on reconstitution of adenylate cyclase (CyaA) activity, when two domains of CyaA (T18 and T25) are fused to the candidate proteins in order to examine their interaction. CsrA and FliW from *C. jejuni* were fused to either to T18 or T25 domains of the adenylate cyclase CyaA protein in separate plasmids. Co-transformation of several CsrA and FliW fusions reconstituted CyaA activity in an *E. coli cyaA* deletion mutant, whereas CsrA and FliW fusions did not confer CyaA activity when they were co-transformed along with empty plasmids (Fig. 32A). Co-transformation of Zipper protein domain fusions to T18 and T25 domains also reconstituted CyaA activity and served as a positive control [113].

To investigate whether FliW can interact with CsrA and FlaA in *C. jejuni* itself, protein-protein coIP experiments were performed using an anti-FLAG antibody and strains with C-terminal 3xFLAG-tag fusions at their endogenous genomic location in *C. jejuni* strain NCTC11168. The anticipated interaction partners were tagged with mCherry at their C-terminus, also at the native locus. In a coIP of FliW-3xFLAG, CsrA-mCherry was successfully co-purified (Fig. 32B), indicating that the two proteins can interact. Similarly, FliW-mCherry was co-purified in a coIP of FlaA-3xFLAG (Fig. 32B), indicating conservation of the interactions between the three proteins and the potential for a partner switch mechanism. As expected, none of the proteins was co-purified in control coIPs with strains that carry the mCherry fusion proteins but not the FLAG-tagged proteins.

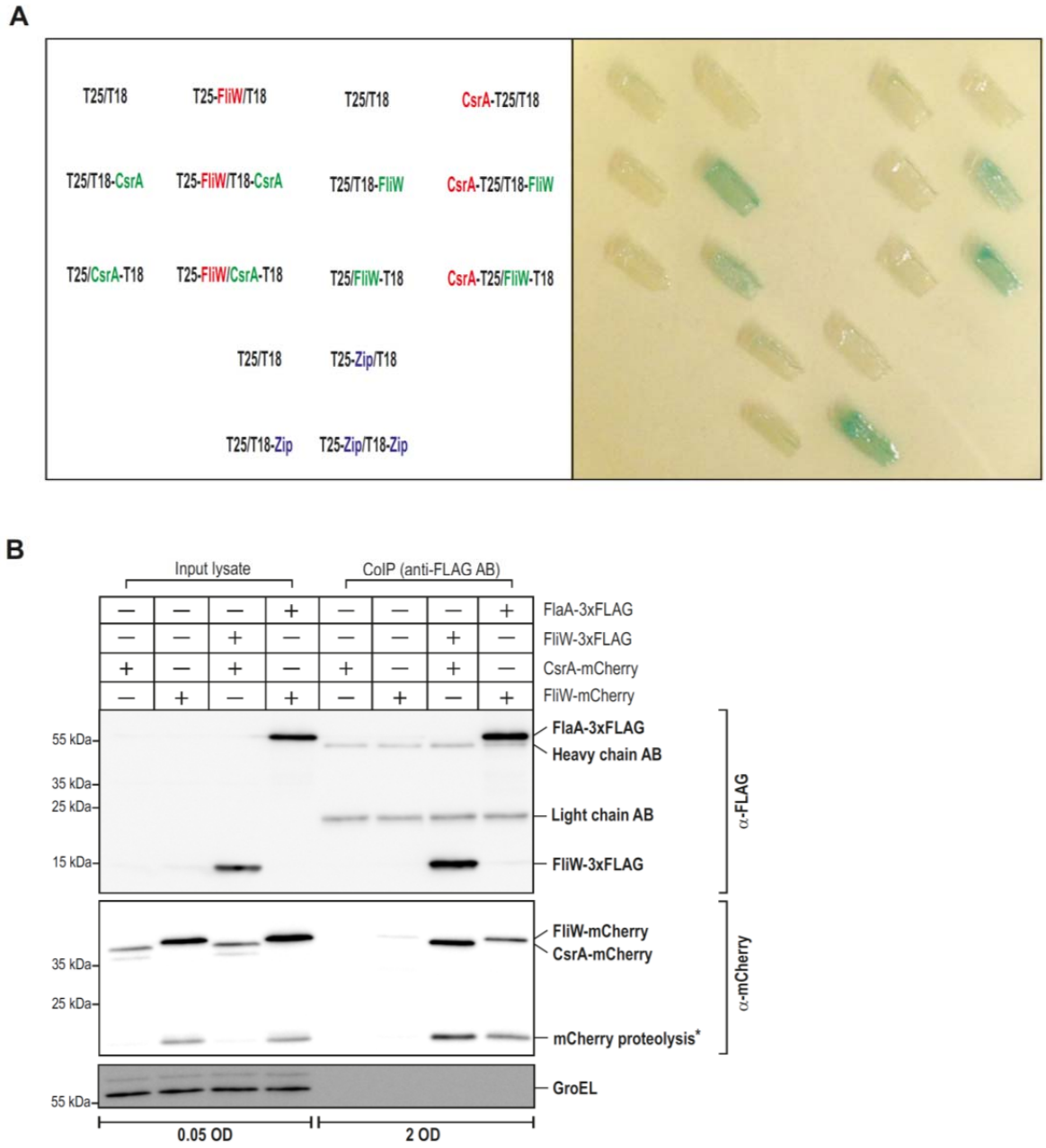


Figure 32. BACTH and protein-protein coIP confirms direct interactions of FliW with CsrA and FlaA. (A) (*Left*) The combinations of different T25 and T18 fusion plasmids introduced in *E. coli* BTH101 are indicated. (*Right*) Growth of the strains indicated on the left on X-Gal indicator plates to monitor ß-galactosidase activity. A blue signal suggests direct interaction between the fusion proteins that reconstitute the two CyaA protein domains. T25-Zip/T18-Zip serves as positive control. The strains were streaked and incubated for 1 day at 30 °C. **(B)** CsrA-mCherry and FliW-mCherry were specifically co-purified in a coIP of FliW-3xFLAG and FlaA-3xFLAG, respectively, using an anti-FLAG antibody. In a negative control reaction with non-tagged FliW and FlaA, CsrA-mCherry and FliW-mCherry were not pulled down. Western Blots were performed for the input lysates and coIP protein samples using anti-FLAG and anti-mCherry antibodies. GroEL served as loading control for the input lysate samples and was not detected in the coIP fraction. (*Please note that the lower band represents partially hydrolyzed mCherry resulting from sample preparation [114].)

To determine whether the FliW-CsrA interaction could antagonize CsrA function in Epsilonproteobacteria, FlaA levels was used as a read-out for CsrA regulation. While FlaA-3xFLAG

was roughly three-fold up-regulated in Δ*csrA* (Fig. 33, lanes 1-4), deletion of *fliW* led to around six-fold down-regulation, consistent with further repression of *flaA* translation by additional free CsrA released upon deletion of its protein antagonist. A Δ*csrA*/Δ*fliW* double deletion strain confirmed that the observed down-regulation was indeed mediated through CsrA, since FlaA-3xFLAG levels increased to levels comparable to those in the Δ*csrA* mutant. To uncouple transcriptional from translational *flaA* regulation, the σ^{28}-dependent *flaA* promoter in the FlaA-3xFLAG strain was replaced with a constitutive σ^{70}-dependent *metK* promoter. Upon deletion of *csrA* in this strain, a roughly three-fold increase FlaA-3xFLAG levels were observed, further supporting translational repression of *flaA* by CsrA (Fig. 32, lanes 5-8). Strong down-regulation of FlaA-3xFLAG was also observed upon deletion of *fliW*, which was restored in the Δ*csrA*/Δ*fliW* double mutant strain, indicating that FliW antagonizes CsrA-mediated translational repression of *flaA* mRNA.

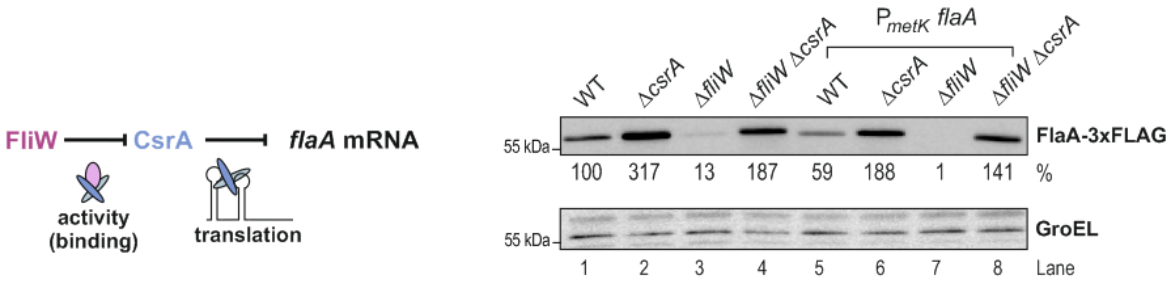


Figure 33. The flagellar assembly factor FliW binds and antagonizes CsrA. A representative Western blot of FlaA-3xFLAG in WT, Δ*csrA*, Δ*fliW* and Δ*csrA*/Δ*fliW* double mutant from liquid cultures in log growth phase (lanes 1-4). Lanes 5-8 represents the same mutations in a strain where the *flaA* promoter was exchanged with *metK* promoter. Anti-FLAG antibody was used to detect FlaA-3xFLAG. GroEL served as loading control.

In line with strong down-regulation of FlaA upon Δ*fliW* deletion, transmission electron microscopy revealed short flagella on Δ*fliW* bacteria compared to those of the WT strain (Fig. 34A and B). In fact, the flagella of Δ*fliW* appeared similar to those of a Δ*flaA* mutant strain and those of bacteria that lack the alternative sigma factor σ^{28} (Δ*fliA),* which is required for *flaA* transcription. In contrast, the Δ*csrA* and Δ*csrA*/Δ*fliW* deletion strains showed normal flagellar filaments. The short flagella of the Δ*fliW* strain are probably composed mainly of the minor flagellin FlaB, which is transcribed from an RpoN (σ^{54})-dependent promoter [115]. Upon deletion of both flagellin genes, *flaA* and *flaB*, the bacteria no longer have filaments and only a small hook structure was observed at the poles (black arrow heads, Fig. 34A). Furthermore, bacteria with a deletion of *rpoN* had neither flagella nor hooks. Motility assays revealed that compared to WT, the Δ*csrA* or Δ*fliW* strains showed a halo-radius reduction to 78% and 72%, respectively (Fig. 34B). The Δ*rpoN* and Δ*fliA* strains were non-motile. Likely due to its shorter flagella, Δ*fliW* showed slower autoagglutination than WT and intermediate to the non-motile Δ*fliA* and Δ*rpoN* mutants (Fig.

34C). Overall, these data suggest that FliW affects post-transcriptional control of FlaA and filament assembly in a CsrA-dependent manner, indicating protein-mediated control of CsrA activity is more widespread than anticipated.

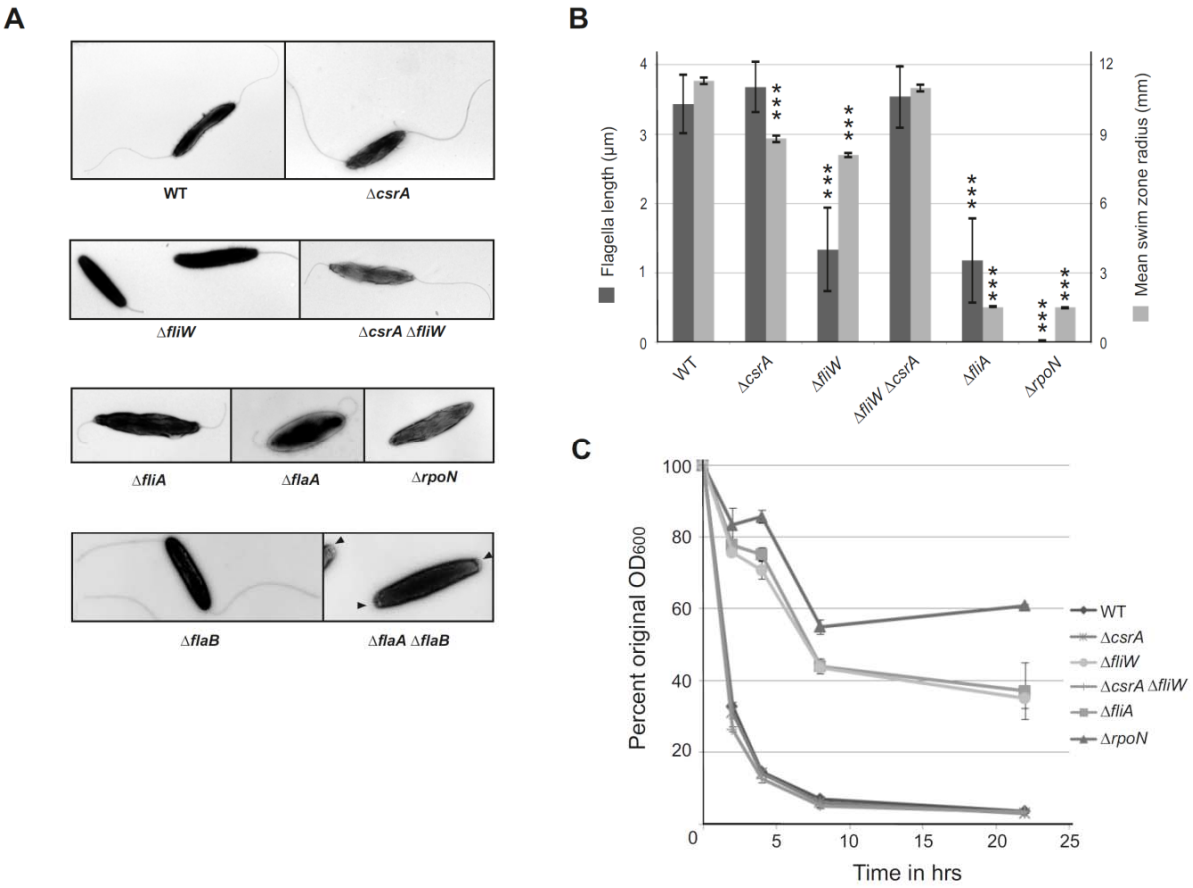


Figure 34. The flagellar assembly factor FliW affects proper flagella formation. (A) Transmission electron (TEM) micrographs of indicated strains harvested from MH agar. Black triangles indicate hook structures. **(B)** Average flagella length (dark grey bars) of indicated strains from TEM micrographs using ImageJ (n > 25 measurements). Plotted is the mean ± s.d. (***P<0.001 vs. WT using Student's *t*-test). Motility was measured as average swimming distance (light grey bars) in soft agar. Bars show the mean ± s.e.m (***P<0.001 vs. WT using Student's *t*-test). **(C)** Autoagglutination assay of *C. jejuni* WT and mutant strains (OD_{600} of supernatants of 1.0 OD bacterial suspensions grown in Brucella broth at the indicated time points) in PBS. Error bars indicate the s.e.m.

3.2.11. Other flagellar mRNAs targets of CsrA in *C. jejuni*

Besides the most abundantly co-purified transcript *flaA* mRNA, many other flagellar targets such as the 5' UTRs of *flgI* (a basal body component), *flaG* (involved in flagellum formation), and *flaB* (encoding the minor flagellin) were strongly enriched (>170-, >346- and >58-fold, respectively; Table 5). The leaders of *flgI, flaG* and *flaB* also have one or more GGA-containing motifs near their SD (Fig. 35A) and were validated to be strongly (more than ten-fold) regulated by CsrA in the *E. coli* reporter-system (Fig. 24B). *In-vitro* gel-shift assays of T7-transcribed *flgI*, *flaG,* and *flaB*

leaders, and several other flagellar mRNAs co-purified in the coIP (Cj0040, *flgA*, and *flgM*) confirmed CsrA binding to these transcripts (Fig. 35B and Appendix Fig. 12A). The Cj1324 mRNA, which encodes a gene involved in flagellin modification and was not enriched in the coIP, did not shift with CsrA, confirming specific binding of CsrA to the coIP-enriched transcripts (Appendix Fig. 12B). However, gel shift assays revealed affinity of CsrA for *flgI, flaG* and *flaB* leaders was lower (K_d values >350 nM, Fig. 35B) than for the *flaA* WT leader (K_d around 50 nM, Fig. 29). These higher K_d values are in a similar range as for *flaA* leader mutants. Unlike FlaA-3xFLAG, chromosomally-tagged FlgI-3xFLAG, FlaB-3xFLAG, and FlaG-3xFLAG levels did not change upon *csrA* deletion in *C. jejuni* strain NCTC11168 (Fig. 35A). This *in-vivo* results and higher K_d observed *in-vitro* suggest that CsrA activity under routine conditions may not be sufficient to regulate targets with lower affinity.

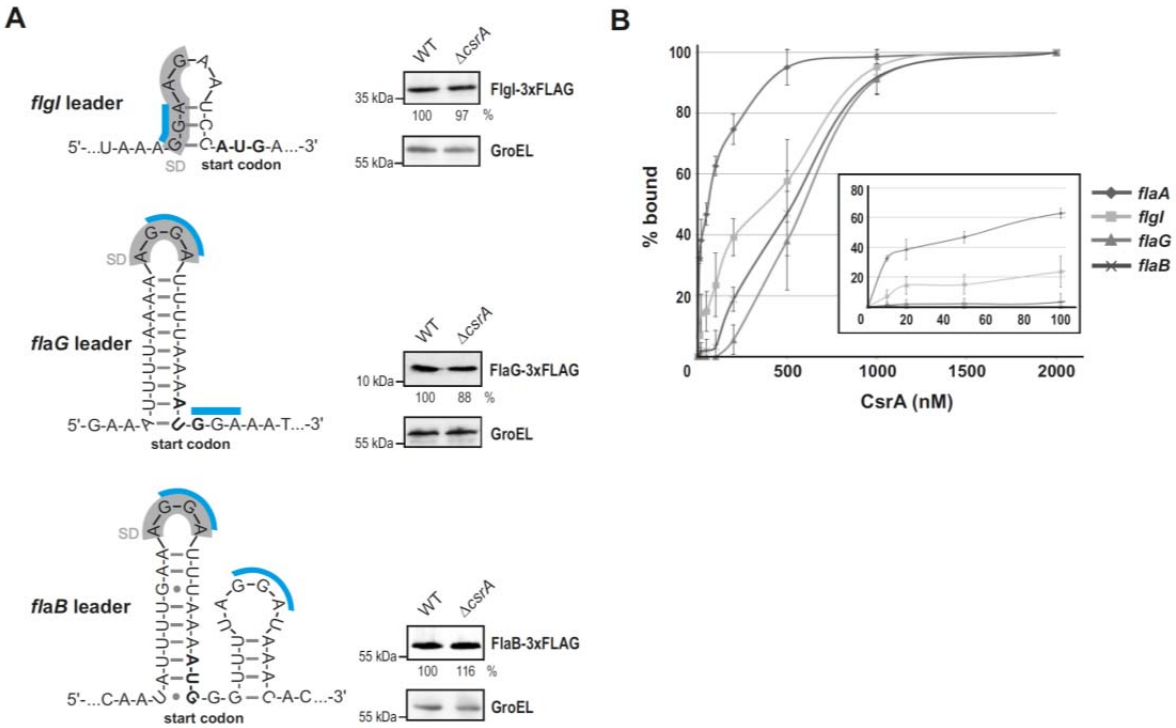


Figure 35. CsrA binds to other flagellar target mRNAs but *csrA* deletion does not affect their translation. (A) *(Left)* Predicted secondary structures of *flgI, flaG,* and *flaB* mRNA leaders using Mfold [110] with putative CsrA-binding GGA-sites (blue) and SD sequences (grey). *(Right)* Western blot analyses of FlgI-3xFLAG, FlaG-3xFLAG, and FlaB-3xFLAG in *C. jejuni* WT or Δ*csrA* strains. **(B)** CsrA binding affinities of flagella mRNA leaders (≤4 nM) determined by *in vitro* gel-shift assays. The inset represents an enlargement of the binding curves for low CsrA-Strep concentrations. Shown is the mean ± s.d.

3.2.12. FliW and *flaA* mRNA titrate CsrA-mediated repression

More than ten-fold regulation was observed for *flgI, flaB,* and *flgG* in the *E. coli* reporter system. This indicates that CsrA in principle is able to regulate these targets, however such a regulatory effect of CsrA was not observed in *C. jejuni* itself. Thus, this led to the hypothesis that FliW or even very abundant CsrA binding mRNAs might titrate CsrA under the examined conditions, thereby obscuring any regulatory effect on low affinity targets such as *flaG* and *flgI* in *C. jejuni*. As *flaA* mRNA is highly abundant [97], and is also expressed at the end of flagellar regulatory cascade by a FliA dependent promoter, it could be reasoned that the *flaA* mRNA itself may act as a regulatory RNA and could titrate CsrA along with FliW. To identify a potential effect of the *flaA* 5' UTR on regulation of the other potential CsrA targets and an interplay with the FliW protein antagonist, FlaG-3xFLAG, FlaB-3xFLAG, and FlgI-3xFLAG protein expression in mutant strains of both antagonists was analyzed. As described above, FlaA-3xFLAG levels increased 3-fold in Δ*csrA* and decreased six-fold upon deletion of *fliW* in a CsrA-dependent manner (Fig. 33). In line with FliW acting as a general CsrA antagonist, deletion of *fliW* also led to about three-fold decrease in FlaG-3xFLAG levels, which was restored to WT levels in a Δ*csrA*/Δ*fliW* double mutant (Fig. 36). Because *flaG* and *flaA* are primarily transcribed from a σ^{28}-dependent promoter [116] and thus are expressed at the same time, monitoring FlaG-3xFLAG may reveal the potential role of *flaA* 5' UTR as a CsrA antagonist. Investigation of the effect of the chromosomal *flaA* M1 leader mutation (GGA -> AAA in SL1, Fig. 28A), which leaves the coding region intact but abolishes CsrA binding, revealed about three-fold decrease in FlaG-3xFLAG levels in the *flaA*-M1 mutant (Fig. 36A). Again, upon introduction of Δ*csrA,* FlaG-3xFLAG levels were restored to that of the WT. This indicates that the *flaA* leader can indeed titrate CsrA and thereby mediate regulation of *flaG*. Interestingly, combining both Δ*fliW* and *flaA*-M1 led to a ten-fold reduction in FlaG-3xFLAG levels. The cumulative effect of *fliW* deletion and *flaA*-M1 mutation clearly shows their separate roles in antagonizing CsrA. In line with this, the M1/Δ*fliW*/Δ*csrA* triple mutant strain completely restored FlaG-3xFLAG levels back to WT levels (Fig. 36A).

Next, the effect of the two antagonists was evaluated for the RpoN-dependent genes *flaB* and *flgI*. A similar reduction, yet less pronounced compared to FlaG-3xFLAG, was also observed for FlaB-3xFLAG upon single or double mutations of *fliW* and M1 (Fig. 36B). In contrast, FlgI-3xFLAG levels were only significantly reduced upon *fliW* deletion (Fig. 36C). Overall, this reveals FliW as the major CsrA antagonist under the examined growth conditions that titrates, along with the *flaA* mRNA antagonist, CsrA from the lower affinity flagellum-related targets. This suggests the high-affinity CsrA binding site of the *flaA* 5' UTR has a regulatory function for translational control of several other flagellar genes.

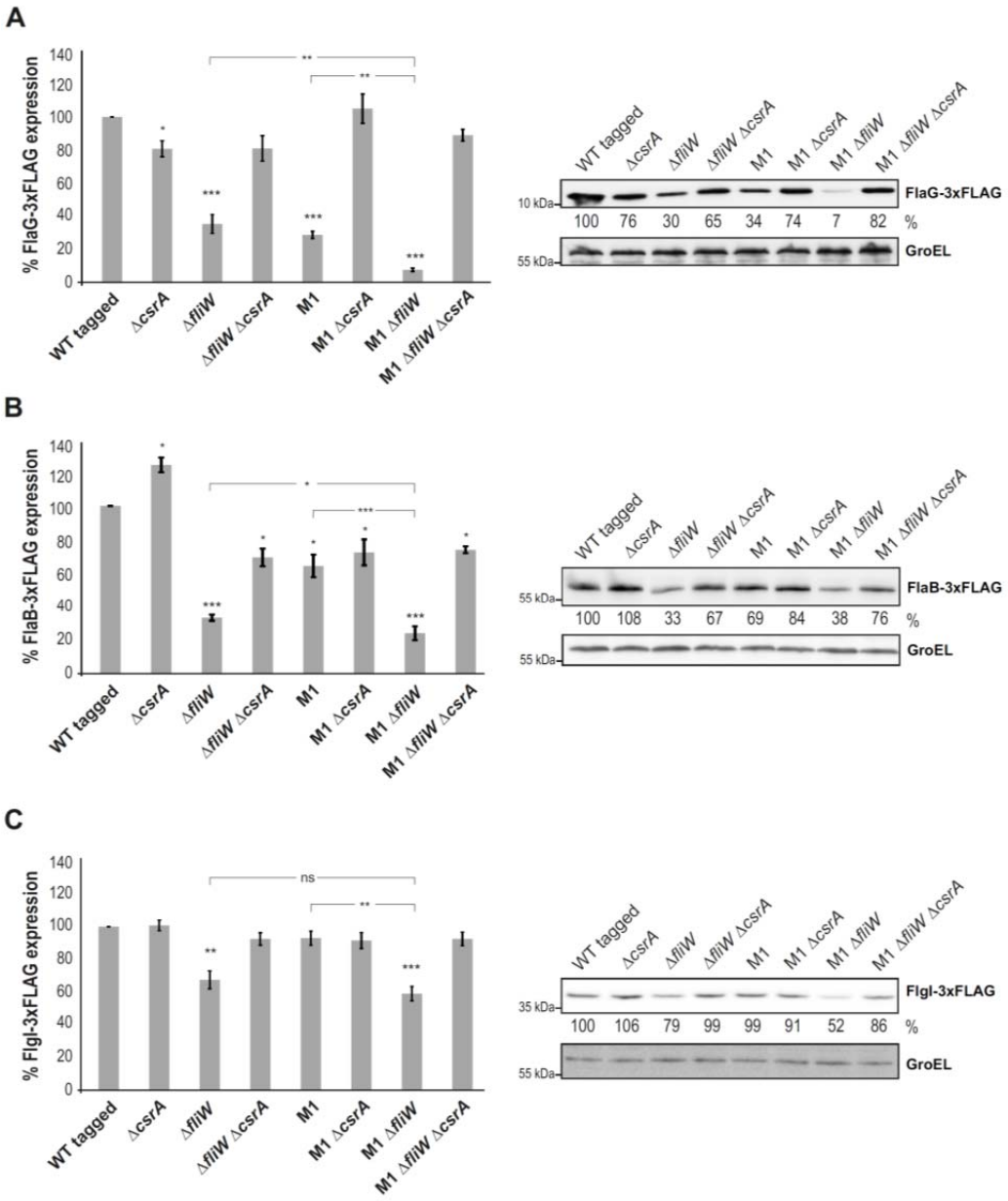


Figure 36. The flaA 5' UTR and FliW inhibit CsrA-mediated regulation of flagella genes. Quantification of FlaG-3xFLAG **(A)**, FlaB-3xFLAG **(B)**, and FlgI-3xFLAG **(C)** levels using Western blot of the indicated *C. jejuni* NCTC11168 strains grown to mid-log phase (M1: GGA->AAA in SL1 of flaA 5' UTR). Values were calculated based on at least three replicates. One representative blot is shown next to each graph. Shown is the mean ± s.e.m (*P<0.05, **P<0.01, ***P<0.001, using Student's *t*-test; n.s. – not significant).

To further confirm the role of the *flaA* 5' UTR as a bona fide CsrA antagonist, a *flaA_mini* transcript comprising the *flaA* leader and first 17 codons followed by a stable ribosomal *rrnB* terminator was inserted into the ectopic *rdxA* locus and expressed from the native σ^{28}-dependent *flaA* promoter (Fig. 37A). Northern blot analysis confirmed stable expression of the ~242-nt long *flaA_mini* transcript (Fig. 37B). Expression of the *flaA_mini* transcript in a Δ*fliW* mutant with strong CsrA-mediated translational repression of *flaA* led to an about three-fold increase of FlaA-3xFLAG protein levels. This indicates it can also bind and antagonize CsrA activity and partially

relieve CsrA-mediated translational repression of *flaA* translation. A lower, yet significant, complementation of the effect of a *fliW* deletion was also observed for FlaG-3xFLAG levels, further supporting the potential of the *flaA* 5' UTR as an mRNA-derived CsrA antagonist. It should be noted that no changes were observed in the target protein levels when flaA_mini was expressed in the WT background. This is probably due to already high sequestration of CsrA by the FliW protein and the endogenous *flaA* mRNA and the effect is only revealed in the absence of the major antagonist FliW.

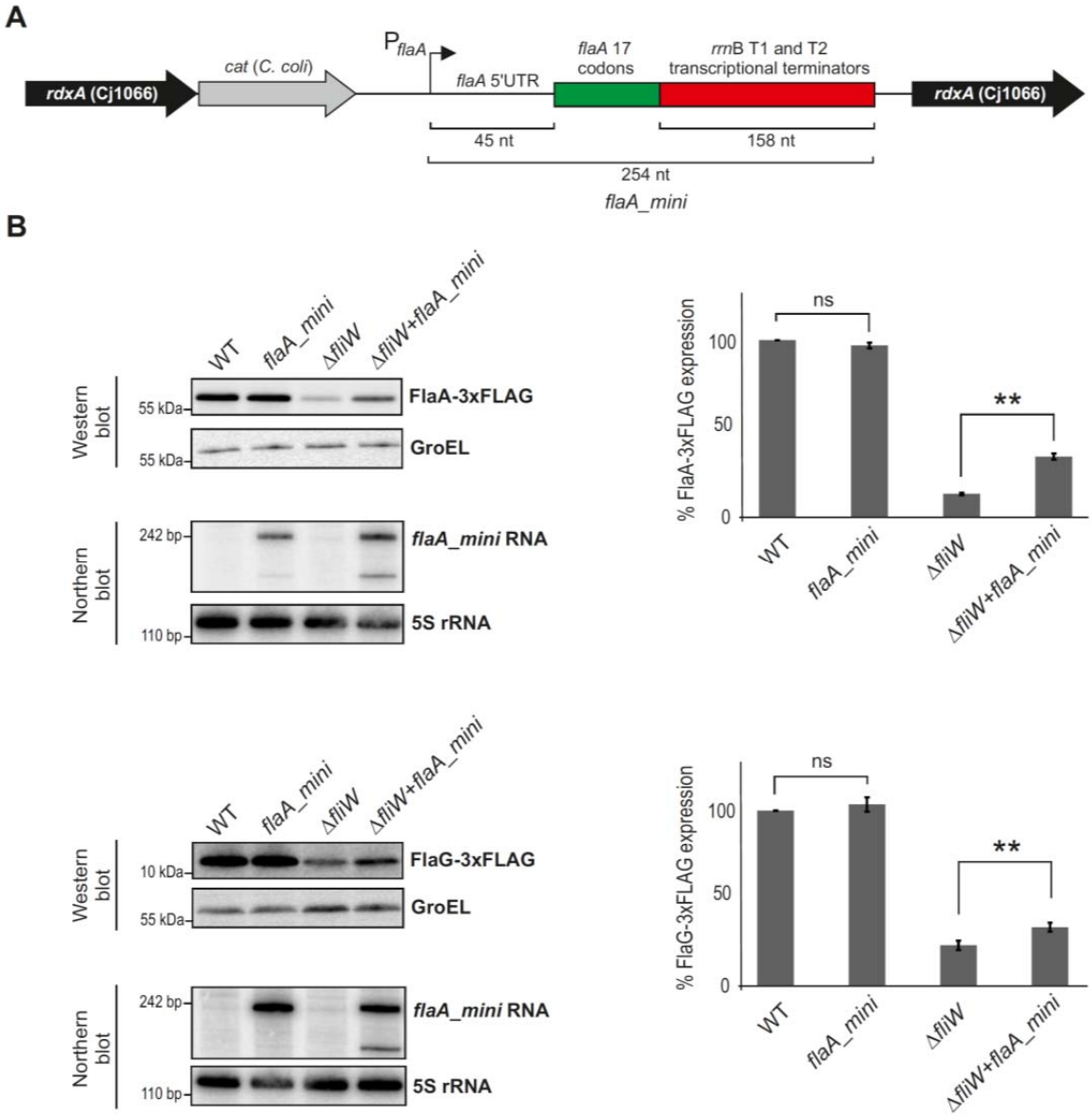


Figure 37. An ectopically-expressed *flaA* mini-gene can partially complement CsrA-mediated effects on FlaA and FlaG translation upon *fliW* deletion. **(A)** Schematic representation of the *flaA_mini* gene construct expressed from the *rdxA* locus of *C. jejuni* strain NCTC11168. **(B)** *(Left panel)* Representative Western blots of FlaA-3xFLAG and FlaG-3xFLAG used for the quantifications on the right and Northern blots of *flaA_mini* RNA expression from liquid cultures in mid-log phase. Anti-FLAG antibody was used to detect the tagged proteins. GroEL served as loading control. *(Right panel)* Quantification of FlaA-3xFLAG (*Top panel*) and FlaG-

3xFLAG (*Bottom panel*) determined by Western blot in the indicated strains (n ≥ 4). Error bars indicate mean ± s.e.m (**P<0.01). Deletion of *fliW* leads to strong CsrA-mediated translational repression of *flaA*-3xFLAG and *flaG*-3xFLAG due to release of CsrA repression in the absence of the FliW protein antagonist.

3.2.13. *flaA* mRNA localizes to the poles of shorter *C. jejuni* cells

The *flaA* 5' UTR was able to titrate CsrA activity. However, it is unclear whether/under which conditions *flaA* mRNA levels might change in order to modulate CsrA-mediated control of other flagellar genes. No strong changes in *flaA* mRNA expression were observed during growth (Fig. 38A). However, Northern blot analysis was performed on RNA harvested from bacterial populations in batch cultures. In the amphitrichous flagellated *C. jejuni*, after every cell division, a new flagellum has to be synthesized at the new pole of each daughter cell. Since bacteria in broth culture are not synchronized in cell cycle, differences in *flaA* mRNA expression might have been obscured due to this population-based analysis. To monitor *flaA* mRNA expression at the single-cell level, RNA-FISH (fluorescence *in situ* hybridization) on fixed *C. jejuni* cells from exponential phase was performed using 14 Cy5-labeled single-stranded DNA oligonucleotide probes against *flaA* mRNA (Fig. 38B, red signal). As a control, a FITC-labeled DNA probe was used to detect 16S rRNA, which was detected in all cells (Fig. 38B, green signal). In contrast, *flaA* mRNA was strongly detected in only some bacteria, indicating differential expression among cells. As a negative control, RNA-FISH for *flaA* mRNA was also performed in a Δ*fliA* mutant strain, which showed, as expected, little or no expression of *flaA* due to loss of *flaA* transcription (Fig. 40A). Furthermore, while the 16S rRNA signal was equally distributed throughout the cell, *flaA* mRNA was specifically detected at the cell poles in only about 20% of the WT cells (Fig. 38B).

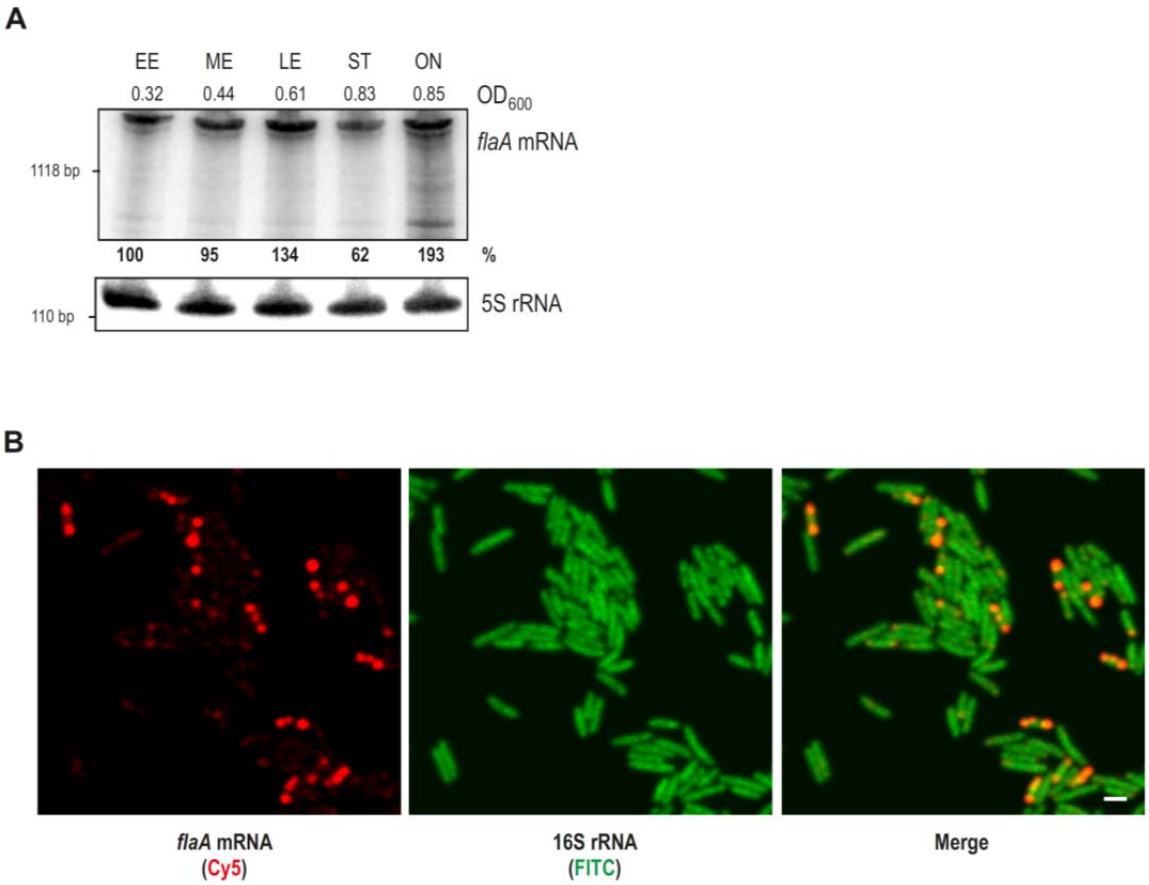


Figure 38. *flaA* mRNA localizes to the poles of about one-fifth of *C. jejuni* NCTC11168 wild-type cells. (A) Northern blot analysis of flaA mRNA using RNA extracted from C. jejuni samples collected at different growth phases (EE-Early Exponential, ME-Mid Exponential, LE-Late Exponential, ST-Stationary and ON-Overnight). The OD600 of the culture is also indicated below each phase. 5S rRNA was probed as a loading control. **(B)** RNA-FISH analysis of 16S rRNA (FITC-labeled DNA oligonucleotide probe, green) and *flaA* mRNA (14 Cy5-labeled single-stranded DNA oligonucleotide probes, red) in *C. jejuni* NCTC11168 WT cells in mid-log phase using confocal microscopy. (Scale bar, 1 µm).

Quantification of cell length across the population showed that cells with localized *flaA* mRNA are significantly shorter than cells without *flaA* expression (Fig. 39). This indicates that in shorter and presumably elongating cells, *flaA* expression and localization may be required for synthesis of a new flagellum at the new pole after cell division. Live-cell imaging of a non-motile *C. jejuni* strain (Δ*fliA*) over 2-3 division cycles showed regular patterns of an increase in cell length until cells divide at midcell, resulting in short cells (Appendix Fig. 13). This indicates that short cells correspond to cells that have just divided and are elongating. Together, these data show differential expression of *flaA* mRNA during the cell cycle and accumulation at the required place of its encoded protein. Accordingly, expression variation of this CsrA antagonist are controlled in a temporally and spatially fashion, which in turn may affect post-transcriptional regulation of the other flagellar genes.

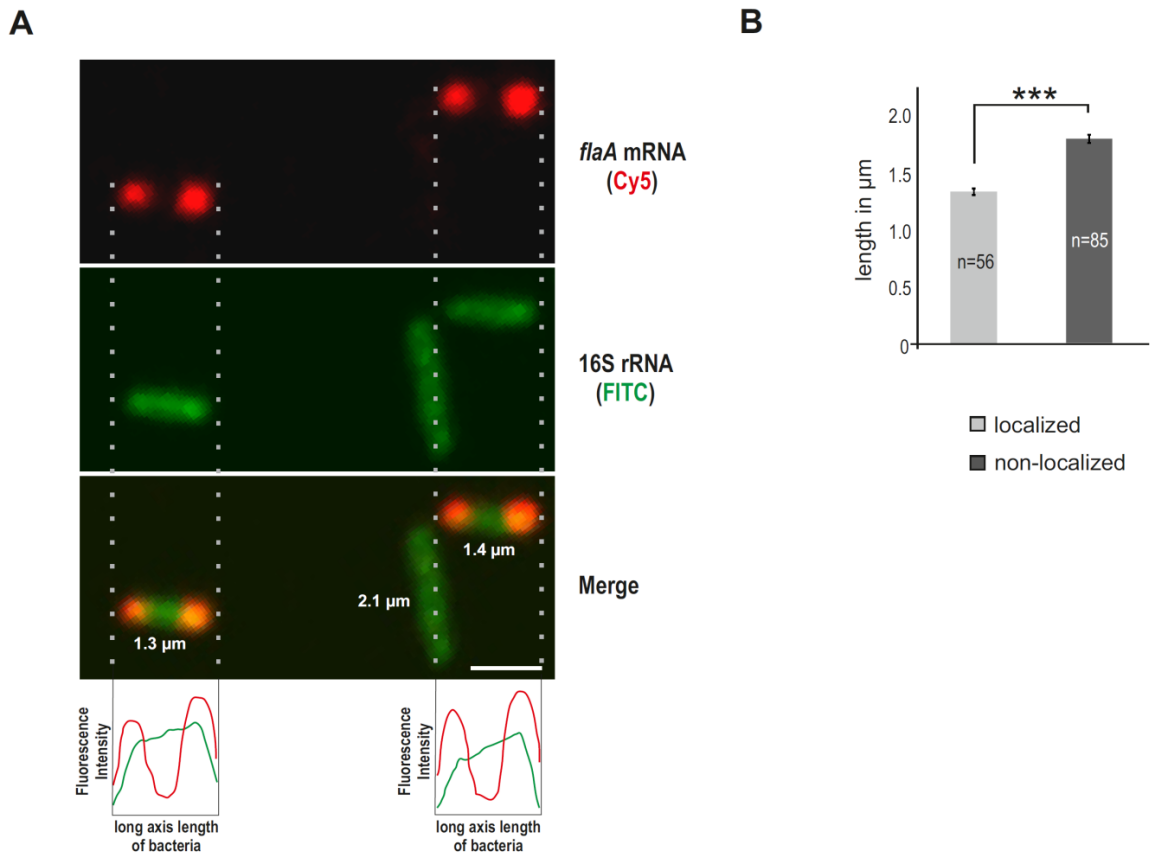


Figure 39. The *flaA* mRNA localizes to the poles of short elongating cells. (A) A magnified RNA-FISH image showing the distribution of fluorescence signals. *flaA* mRNA (Cy5) and 16S rRNA (FITC) signals were quantified along the long axis length of bacteria using ImageJ software and were subsequently merged as shown at the bottom of the panel (Scale bar, 1 µm). The length of individual cells was also quantified using ImageJ. **(B)** Average *C. jejuni* NCTC11168 WT cell lengths in bacteria where *flaA* mRNA is localized (56 cells) or non-localized (85 cells), *** $P<10^{-15}$ using Student's *t*-test.

3.2.14. *flaA* mRNA localization is translation dependent

To investigate whether the CsrA-FliW regulatory circuit impacts *flaA* mRNA localization, RNA-FISH was performed in Δ*fliW*, Δ*csrA,* and Δ*fliW*/Δ*csrA* strains. Interestingly, deletion of *csrA* had no effect on *flaA* localization, whereas deletion of *fliW* completely abolished *flaA* localization (Fig. 40A). Instead of a polar localization, *flaA* mRNA was now dispersed throughout the cell upon *fliW* deletion. Translation of *flaA,* and in turn FlaA protein levels, are strongly reduced in the Δ*fliW* mutant due to more active CsrA released upon deletion of its antagonist. Northern blot analysis showed that *flaA* mRNA expression is two to three fold higher in Δ*fliW* compared to WT, indicating that absence of *flaA* mRNA localization upon *fliW* deletion was not due to lower abundance of the *flaA* transcript (Fig. 40B). The Δ*csrA* and Δ*fliW*/Δ*csrA* strains also expressed similar levels of *flaA* compared to WT. As a control, the Δ*fliA* strain, which lacks the sigma factor for *flaA* transcription (Fig. 40B), showed no *flaA* mRNA FISH signal. Strikingly, *flaA* mRNA localization was restored to the cell poles in the Δ*fliW*/Δ*csrA* double deletion strain, showing that

CsrA affects localization of *flaA* mRNA. In wild-type cells, CsrA activity seems to be limited, allowing for translation of *flaA* mRNA and its polar localization. This suggests a model where *flaA* translation facilitates polar localization: upon deletion of *fliW,* CsrA is released and in turn strongly represses *flaA* mRNA translation to impede its localization to the poles.

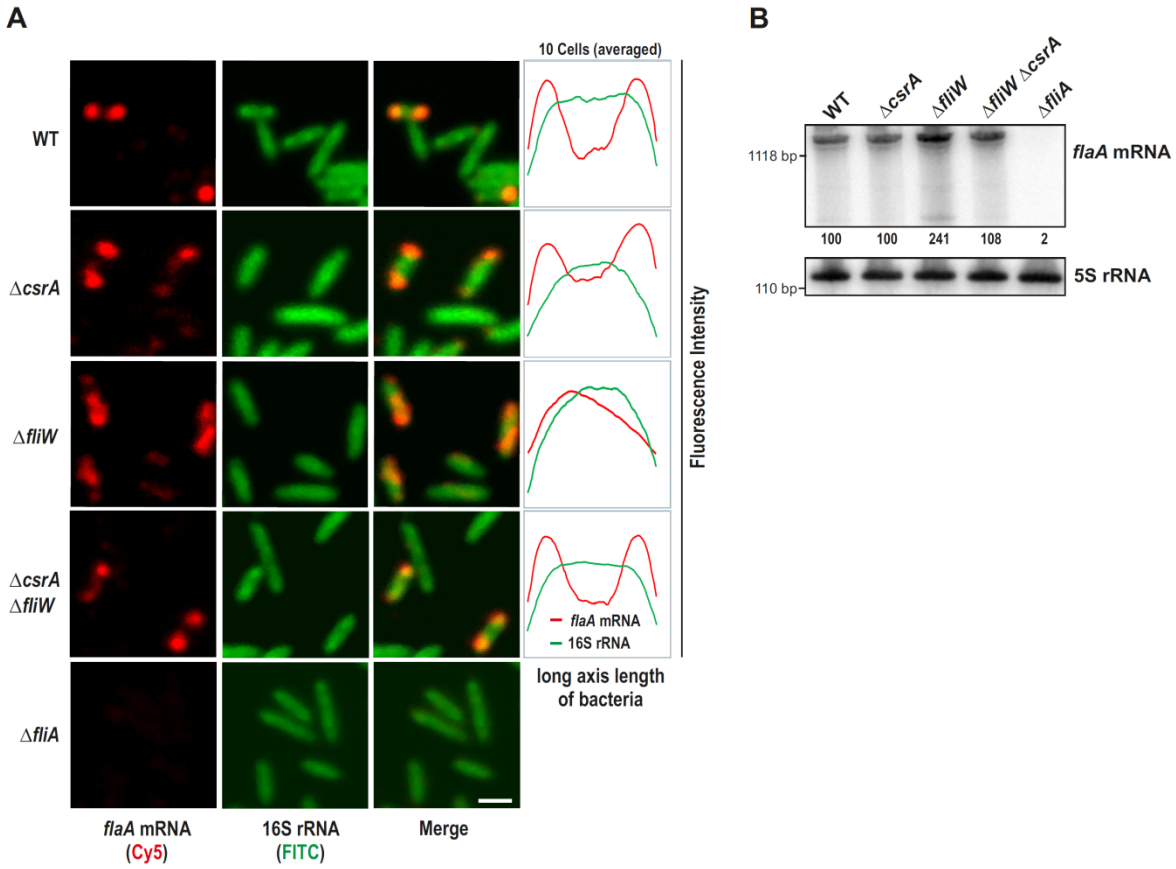


Figure 40. Influence of CsrA and FliW on *flaA* mRNA localization to the poles. (A) RNA-FISH analysis (*Left*: confocal microscopy images; *Right*: averaged fluorescence intensity along the long axis based on 10 cells) of 16S rRNA (green) and *flaA* mRNA (red) in *C. jejuni* NCTC11168 WT, Δ*csrA*, Δ*fliW*, Δ*csrA*/Δ*fliW,* and Δ*fliA* strains in mid-log phase. FITC and Cy5 channels were merged in the microcopy images in the third lanes (Scale bar, 1 µm). **(B)** Northern blot analysis of *flaA* mRNA (CSO-0486) and 5S rRNA (loading control, CSO-0192) in *C. jejuni* NCTC11168 wild-type and the indicated mutant strains grown to mid-log phase. Relative expression levels are quantified below the blot.

To support a translation-dependent model of *flaA* localization, several point mutants in the native *flaA* gene that either maintain or abolish *flaA* translation were constructed (Fig. 41A). First, the start codon of *flaA* [AUG -> AAG (X1) or AUU (X2)] was mutated to abolish translation initiation. These mutations resulted in dispersed *flaA* mRNA localization in the RNA-FISH analysis (Fig. 41B). In contrast, when the start codon was mutated to an alternative start codon [AUG -> GUG (X3)], *flaA* mRNA still localized to the cell poles, indicating that translation of *flaA* mRNA is indeed required for proper localization. Next, the 3rd codon of *flaA* was mutated to a stop codon [UUU -> UAG (X4)], which also completely abrogated polar *flaA* mRNA localization. In contrast, *flaA* mRNA

with a synonymous silent mutation [UUU -> UUC (X5); both encoding Phe] localized similarly to the WT mRNA. Northern blot analysis of the *flaA* mutant RNAs showed that some of the mutations that abolish translation (X1, X2) lead to 50-70% reduced *flaA* mRNA levels (Fig. 41C). However, since the stop mutation at the 3rd codon (X4) that abolished polar mRNA localization shows similar or higher *flaA* expression levels compared to WT, it is unlikely that reduced *flaA* mRNA levels lead to loss of localization. The results of silent and stop mutations near the translation initiation codon indicate that translation is required for *flaA* mRNA localization. To provide additional insight into the role of translation in *flaA* mRNA localization, the effect of terminating translation at a later position was determined. A strain with a stop codon at the 101st codon of *flaA* mRNA [CAA -> UAA (X6)] showed partial localization of *flaA* mRNA to the poles, indicating that the N-terminal peptide may be required for polar localization.

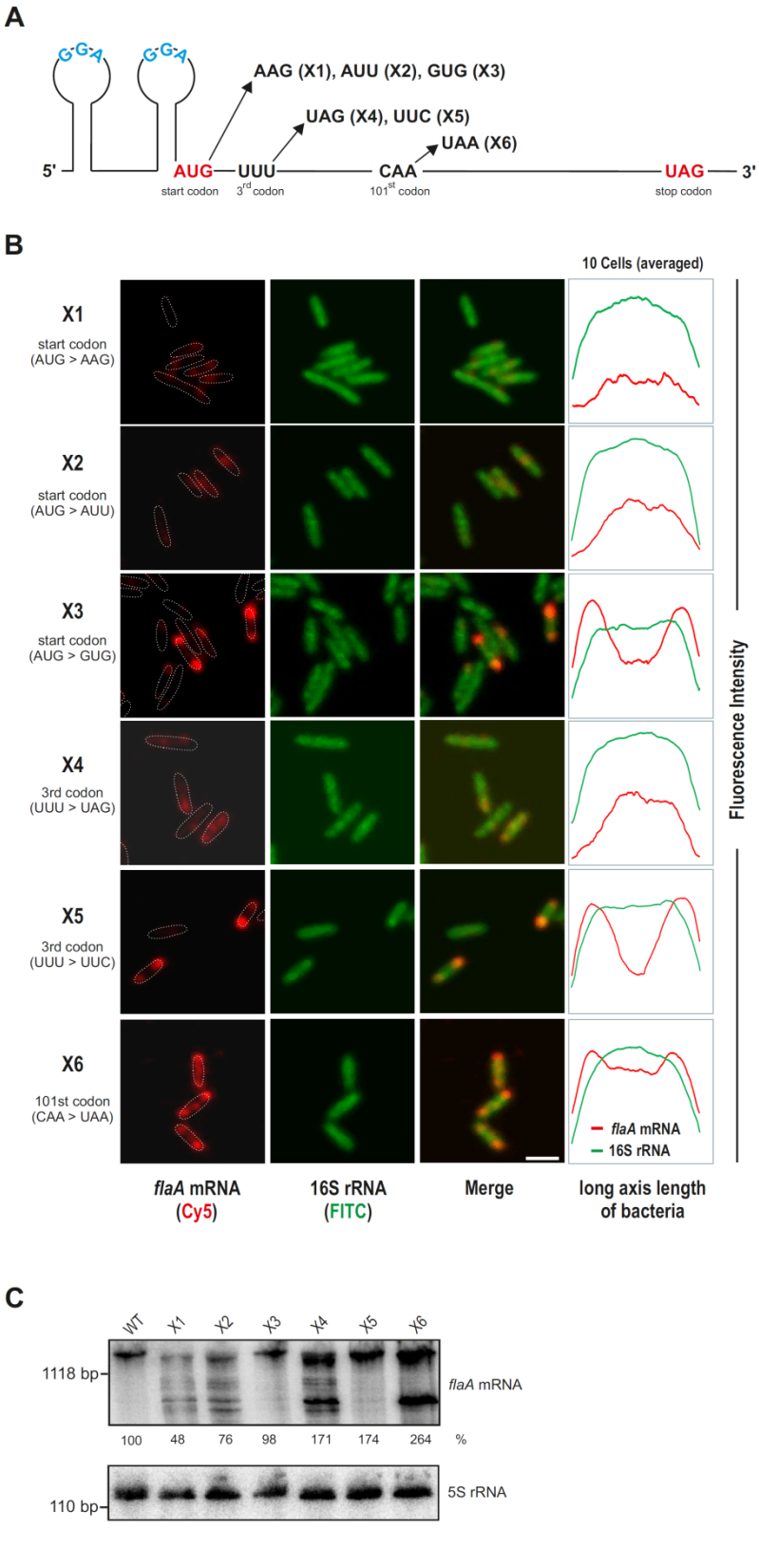


Figure 41. Translation is required for *flaA* mRNA localization to the cell poles. (A) Point mutations in *flaA* mRNA that were introduced at the native *flaA* locus**.** Mutations X1, X2, X4, and X6 abolish or prematurely stop *flaA* translation, while X3 and X5 represent silent mutations. **(B)** RNA-FISH analysis (*Left*: confocal microscopy images; *Right*: averaged fluorescence intensity along the long axis based on 10 cells) of *C. jejuni* point mutant strains depicted in **(A)**. FITC and Cy5 channels were merged in the third rows of the microscopy images (Scale bar, 1 µm). **(C)** Northern blot analysis of *flaA* wild-type (WT) and mutant RNAs

with indicated point mutations in the *flaA* coding region using RNA extracted from *C. jejuni* cells collected at log-phase growth. 5S rRNA was probed as a control on the same blot.

As a further confirmation of *flaA* mRNA localization, super-resolution imaging of *flaA* mRNA FISH by *direct* stochastic optical reconstruction microscopy (*d*STORM) was employed [117] (Fig. 42). Single-molecule localization microscopy is ideally suited to study the distribution of target molecules with virtually molecular resolution and has only recently been described for super-resolution imaging of bacterial RNA [118]. The *flaA* mRNA localization results based on *d*STORM imaging in WT and some of the mutant strains fully support the confocal microscopy analysis. Overall, these data support a role of the FliW/CsrA network in controlling translation-dependent polar *flaA* mRNA localization in *C. jejuni*. This is an unexpected function for the ubiquitous, widely-studied CsrA protein, commonly assumed to be a pleiotropic translational regulator of many biosynthetic and metabolic pathways, in the temporal and spatial control of translation.

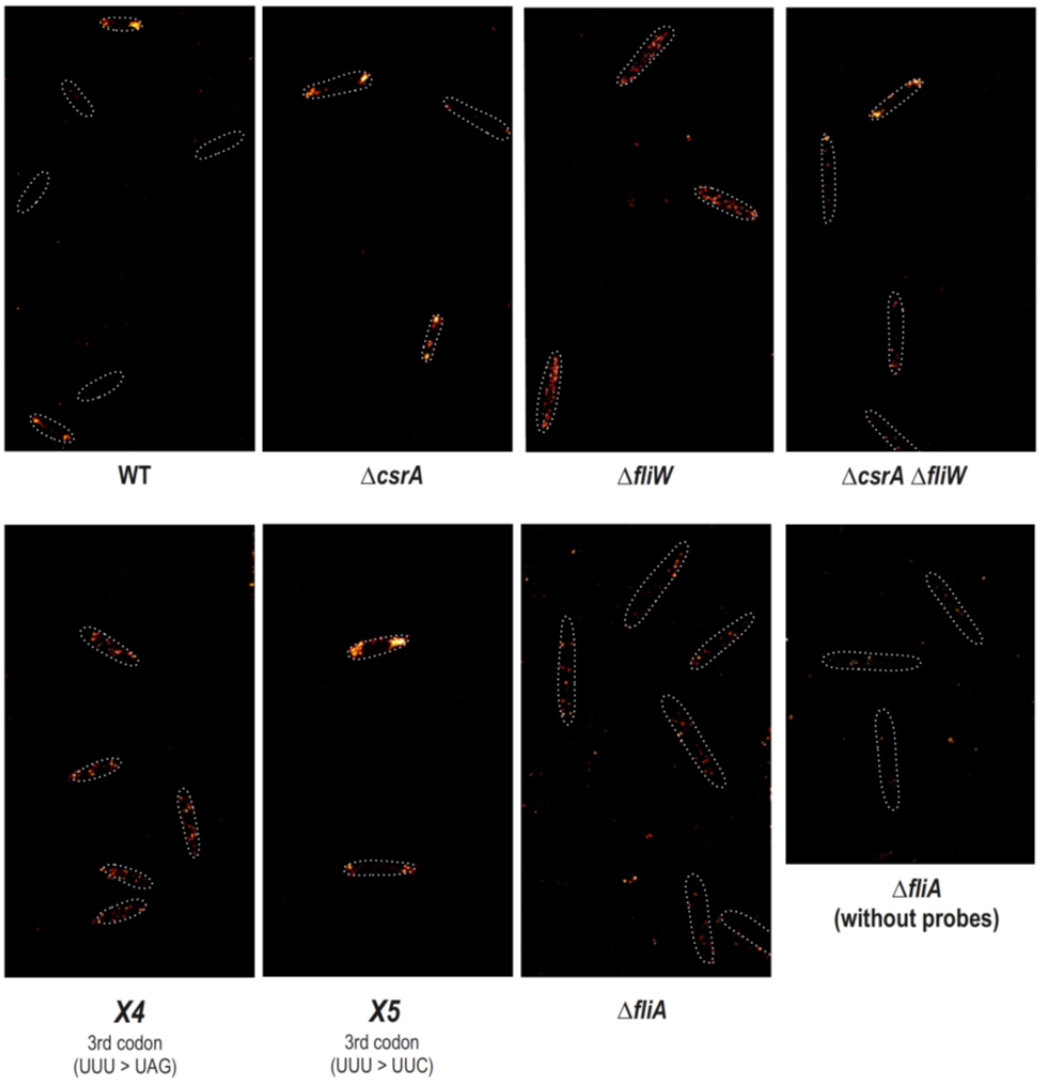


Figure 42. Super-resolution imaging of *flaA* mRNA. RNA-FISH analysis of *flaA* mRNA (14 Cy5-labeled oligos) in the indicated *C. jejuni* wild-type and mutant strains using *d*STORM imaging. Cell boundaries from bright field images are depicted by white dotted lines. As a negative control, the *C. jejuni fliA* deletion strain was analyzed with and without probes to check for background signals.

3.2.15. FliW affects *flaA* transcription

It was observed that *flaA* mRNA levels were 2 to 3 fold higher in the *fliW* deletion strain (Fig. 40B). This was unexpected as *flaA* translation is repressed upon *fliW* deletion, and hence should lead to destabilization of *flaA* mRNA. To test this, rifampicin RNA stability assay was performed and probed for *flaA* mRNA in WT, Δ*csrA*, Δ*fliW*, and Δ*fliW*/Δ*csrA* strains (Appendix Fig. 15A). The rifampicin assays revealed that *flaA* mRNA is very stable over 32 min and that there is no decrease in transcript levels in WT, Δ*csrA* and Δ*fliW*/Δ*csrA* strains (Appendix Fig. 15B). By contrast, as expected decreased stability was observed for the Δ*fliW* mutant due to translational repression, despite overall higher *flaA* expression levels. To investigate this higher steady-state levels of *flaA* mRNA upon *fliW* deletion, a transcriptional fusion of the *flaA* promoter to a mini-gene expressing a stable RNA comprising the 5' UTR and first 53 bases of the coding region of an unrelated *C. jejuni* gene (Cj1321, not a CsrA target) followed by an *rrnB* terminator was constructed. Similarly to the endogenous *flaA* mRNA, Cj1321_mini was also two-fold upregulated upon deletion of *fliW* (Fig. 43A). Since this gene is independent of CsrA translational repression, this observed upregulation indicates a potential effect of FliW on transcription from the *flaA* promoter. To further check the effect of FliW on *flaA* promoter, the *flaA* promoter at its native locus was replaced with an unrelated constitutive *metK* promoter. Compared to the observed upregulation of *flaA* mRNA expressed from its native promoter in the *fliW* mutant, *flaA* mRNA levels expressed from this constitutive promoter were rather three-fold down-regulated upon *fliW* deletion (Fig 43B). Overall, this shows that FliW affects transcription from the *flaA* promoter through an unknown mechanism.

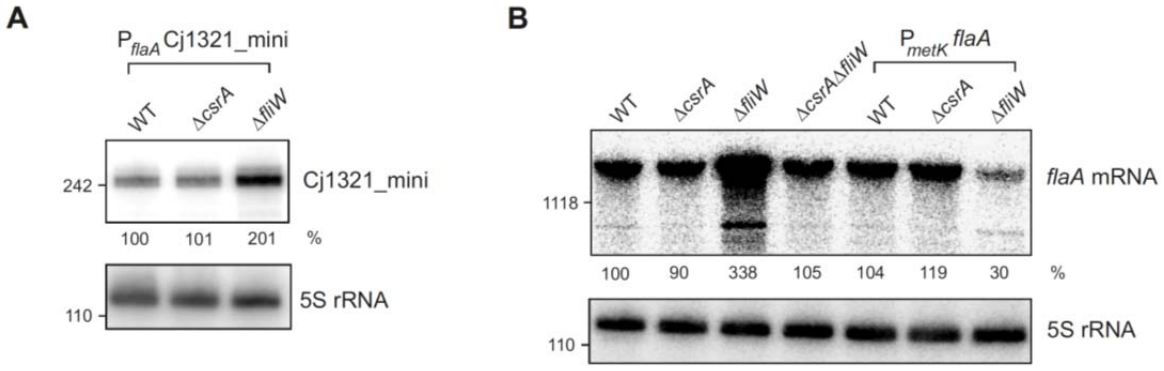


Figure 43. FliW affects transcription from *flaA* promoter. (A) The Cj1321_mini transcript was detected by Northern blot analysis of total RNA from *C. jejuni* cells expressing Cj1321_mini either in the wildtype, Δ*csrA,* or Δ*fliW* background. **(B)** Northern blot analysis to detect *flaA* mRNA levels in *C. jejuni* wildtype (WT) and

mutant strains. The *flaA* mRNA is expressed either from its native σ^{28}-dependent promoter or expressed from a constitutive σ^{70}-dependent *metK* promoter. 5S rRNA was probed as a loading control on both blots.

3.2.16. FlhX – In search for other factors involved in *flaA* mRNA localization

Translation of the *flaA* mRNA is required for its localization to the poles of short and elongating *C. jejuni* cells. Furthermore, it was also demonstrated that the N-terminus of nascent FlaA peptide plays a role in *flaA* mRNA localization. Considering this, it can be postulated that the act of flagellin subunit recognition by the T3SS machinery may also contribute to its co-translational mRNA localization. The rod and hook subunits of the flagellum are recognized by the flagellar type III export apparatus proteins FlhA and FlhB among others and transported across the cytoplasmic membrane [119].

FlhB is a transmembrane protein with a globular C-terminal domain in the cytoplasm ($FlhB_C$). The c-terminal cytoplasmic domain of FlhB can further be divided into two domains - $FlhB_{NC}$ and $FlhB_{CC}$ (Fig. 44A). The role of $FlhB_{NC}$ is unknown but a conserved hydrophobic pocket in $FlhB_{CC}$ is critical for direct interaction with export substrates [120]. Upon secretion and completion of the rod and hook, a specific interaction between $FlhB_{CC}$ and the hook length control protein, FliK, changes the specificity of FlhB from hook to flagellin subunits. FlhB is also known to undergo an autocatalytic cleavage event at a conserved NTPH motif between the $FlhB_{NC}$ and $FlhB_{CC}$ domain, which is also essential for the export specificity switch [121]. However, it is currently unknown if or how this cleavage event is related to its interaction and secretion of FliK [122]. Mutations which disrupt this conserved cleavage motif in FlhB are unable to export the flagellin subunits, but are unaffected for the secretion of rod and hook subunits. Upon FlhB cleavage, the two domains of $FlhB_C$ ($FlhB_{NC}$ and $FlhB_{CC}$) are still present in a tight complex which probably allows for structural change required for the specificity switch.

Interestingly, many bacteria including *C. jejuni* also encode a paralog of $FlhB_{cc}$ termed FlhX [123]. FlhX (Cj0848c) in *C. jejuni* shows a high degree of conservation with FlhB (Cj0335) at the aa level after the NTPH cleavage motif in FlhB (Fig. 44B). In the related bacterium, *H. pylori*, FlhX is proposed to act as a "spare part" to permit flagellar export in the absence of $FlhB_{CC}$ [124]. This replacement role of FlhX is unclear as it still requires the rest of FlhB for functioning. It is also unknown if and how FlhX interacts with FlhB, and if FlhX can still bind to export substrates in the cytoplasm without its association with FlhB.

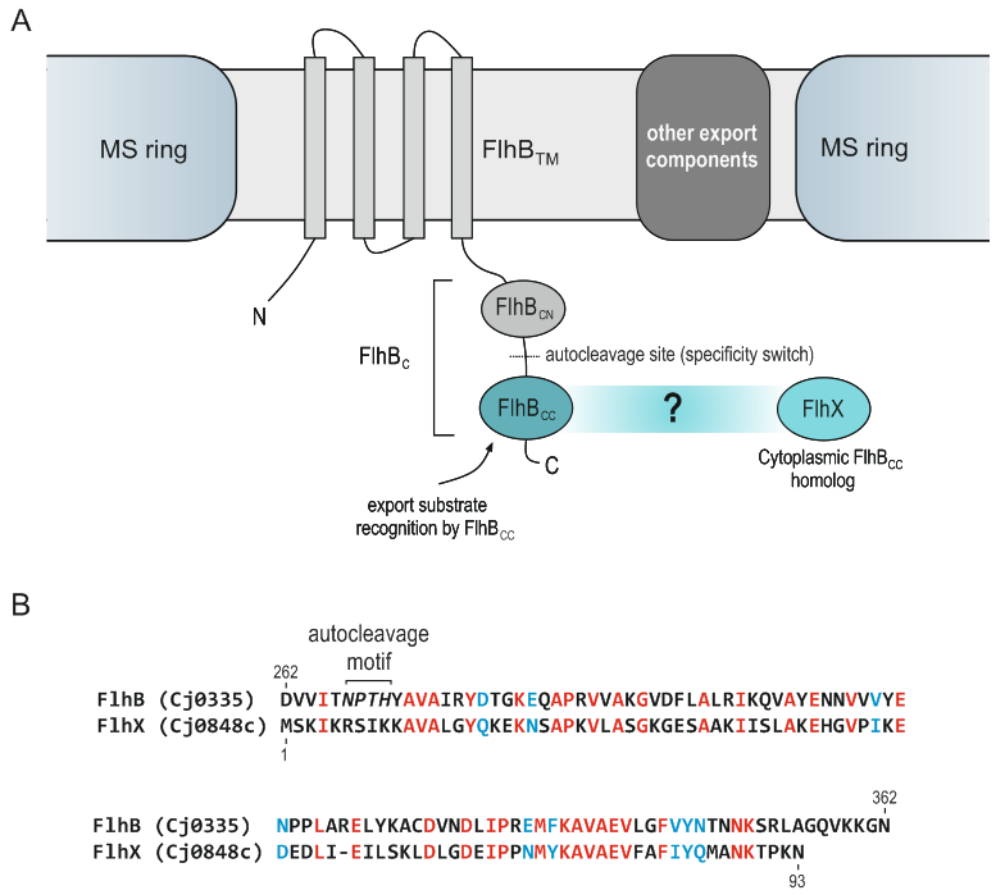


Figure 44. FlhB and FlhX in flagellar export apparatus. **(A)** Schematic cross-section of type III flagellar export apparatus with emphasis on FlhB and FlhX (adapted from Ferris *et. al.,* 2005 [121]). **(B)** Alignment of the C-terminal part of FlhB (262 – 362 aa) with full length FlhX. The NPTH cleavage motif in FlhB is marked. Red and blue represents conserved and similar aa residues, respectively.

The expression of *flaA* undergoes checks at both the transcriptional and post-transcriptional level. Transcription of FliA dependent genes like *flaA* is regulated by the anti-sigma factor FlgM, which is secreted by the export apparatus upon hook completion. The mutations in export machinery which abolish the export of rod or hook proteins also abort *flaA* expression by accumulation of FlgM. However, the processing deficient mutants of FlhB and FlhX deletion strains can still export rod and hook proteins and therefore do not affect *flaA* expression [125]. Hence, such mutant strains present a unique opportunity to study *flaA* mRNA localization when only FlaA export is affected but not its expression. An FlhB processing deficient strain was constructed in *C. jejuni* strain NCTC11168 by mutating the NTPH autocleavage motif (N267A). Preliminary results using RNA FISH revealed expression and patchy accumulation of *flaA* mRNA in most of the mutant (N267A) cells; however some cells still localized the mRNA to the poles (Fig. 45). Deletion of FlhX on the other hand, revealed a striking polar localization of the flaA mRNA in all the cells. The importance of FlhB and FlhX in *flaA* mRNA localization can only be accounted once their role in filament substrate recognition is further clarified by future studies.

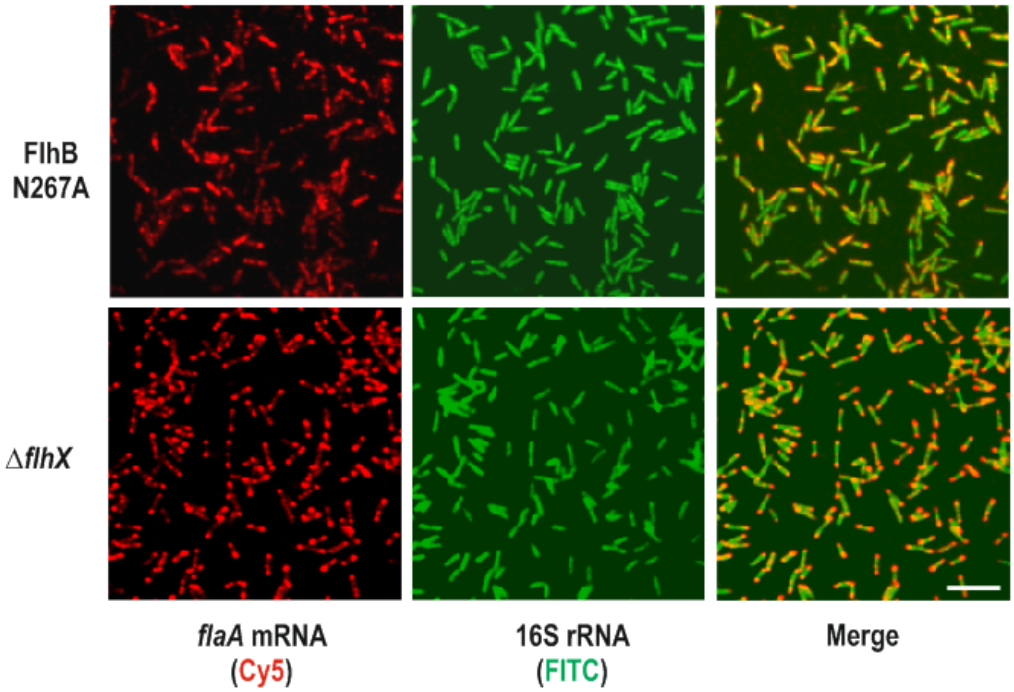


Figure 45. *flaA* mRNA localization in the mutant strains. RNA-FISH of *C. jejuni* FlhB autocleavage defective (FlhB N267A) and *flhX* deletion strains. FITC and Cy5 channels were merged in the third rows of the microscopy images (Scale bar, 5 µm).

3.2.17. DISCUSSION 2

Using a RIP-seq approach, the direct RNA targets of the translational regulator CsrA were identified on a genome-wide scale in two *C. jejuni* strains, NCTC11168 and 81-176. Previous studies in diverse Gammaproteobacteria aimed to determine the CsrA/RsmA regulon using microarray-based transcriptome analysis [126-128] or global proteome analysis [129] of WT vs. Δ*csrA/rsmA* mutant strains. While these approaches can reveal the direction of regulation, "timed-arrays" based on inducible CsrA expression helped to distinguish direct and indirect effects in *E. coli* [130]. Since CsrA can mediate translational repression without affecting mRNA levels [100], potential CsrA targets might be missed in transcriptome analysis when comparing WT vs. Δ*csrA/rsmA* mutant strains. In line with this, RNA-seq analysis of *C. jejuni* NCTC11168 WT and Δ*csrA* strains did not reveal any major changes, and only five genes (5' UTR or ORF) were found to be regulated more than four-fold in this study. In contrast, a coIP approach facilitates the identification of direct CsrA targets along with their binding sites. While Sanger sequencing of cloned cDNAs from a coIP of RsmA-His_6 identified only six novel direct target mRNAs in *Pseudomonas aeruginosa*, global transcriptomics indicated that >500 genes are part of the RsmA regulon [126]. Now, the advent of RNA-seq allows for high-throughput identification of co-purified transcripts along with the direct binding sites. 454 sequencing of RNAs from a CsrA-His_6 coIP revealed 721 co-purified transcripts in *E. coli* [129]. In this study, a coIP of the untagged WT strains was also included as a negative control to identify non-specifically bound transcripts. A high specificity of the global RIP-seq approach was confirmed by analysis of the 328 enriched CsrA binding sites using a novel peak detection algorithm which revealed that the vast majority of transcripts (324/328) are bound at "ANGGA" sequences, resembling the consensus motif determined for CsrA homologues based on *in vitro* selection using SELEX [109]. Although, strong enrichment peaks were observed in the CsrA-coIP, the specificity of the approach could be further improved by addition of a cross-linking step (CLIP-seq). Recently, CLIP-seq was used to determine the global recognition pattern of CsrA and Hfq at a single-nucleotide resolution in *Salmonella* [131]. The additional cross-linking step enables more stringent purification of associated RNA which can be further trimmed by ribonucleases to determine the precise binding site [132].

Surprisingly, the coIP data from this study revealed that CsrA binding sites occur not only in 5' regions of mRNAs but also frequently within coding regions in *C. jejuni*. Previously it was assumed that CsrA mainly inhibits translation by binding to 5' UTRs and inhibition of ribosome binding [19]. CsrA can also repress gene expression by binding to the early coding region [133], but this likely also interferes with translation initiation. Using reporter gene fusions of some of these *C. jejuni* targets in *E. coli,* this study demonstrated that CsrA binding between genes in polycistronic mRNAs can mediate dis-coordinate operon regulation. Such CsrA binding

and dis-coordinate operon regulation was also recently observed in global study in *Salmonella* [131]. The underlying mechanisms and regulatory consequences of CsrA binding to ORF regions more distant from the RBS and between co-transcribed genes need to be examined in more detail in future studies. However, the integration of CsrA binding sites within diverse regions of mRNAs and dis-coordinate operon control might be useful for the construction of synthetic genetic circuits.

The global transcriptome analysis and RIP-seq approach revealed the lack of CsrB/C like homologs of Gammaproteobacteria in *C. jejuni* [19]. Instead, transcripts of many flagellar genes were predominantly co-purified with *C. jejuni* CsrA. Moreover, the most abundantly co-purified transcript, the mRNA leader of the major flagellin *flaA*, has dual function as both a coding and regulatory transcript: it can act as both a CsrA antagonist while CsrA itself represses its translation. The interplay between an mRNA and a protein antagonist of CsrA to post-transcriptionally control expression of flagellar genes in the Epsilonproteobacterium *C. jejuni* represents an unexpected combination of regulatory themes from both Gram-positive and Gram-negative bacteria.

The *flaA* leader has higher affinity for CsrA than other flagellar targets. It has two GGA motifs in adjacent hexaloops, resembling high-affinity CANGGANG-containing apical hexaloop structures of CsrA/RsmE [109, 134]. Also, the flaA GGA motifs have a spacing of 21 nt, which is close to the optimal intersite distance of 18-nt for binding of a CsrA dimer [135]. Whereas the *flaA* leader probably binds one CsrA dimer, multiple RsmE dimers are cooperatively assembled on RsmZ sRNA [20, 134]. In agreement with this study, CsrA titration by a 5' UTR has been shown to mediate a hierarchical control of fimbriae expression in *Salmonella typhimurium* [136]. In this system, the *fimAICDHF* mRNA leader cooperates with the CsrB/C sRNAs to antagonize CsrA-mediated activation of plasmid-encoded fimbriae. Whereas *fimAICDHF* translation itself is not affected by CsrA, translation of *flaA* mRNA is repressed by *C. jejuni* CsrA. Thus, the same mRNA is both a target and an antagonist of CsrA. Several examples of bacterial mRNAs with dual function have recently been described, including mRNAs encoding for 3' UTR-derived sRNAs [137], or mRNAs that function as sRNA decoys or sponges [9, 138]. sRNAs other than CsrB/C have also been reported to sequester CsrA in addition to their function as antisense RNAs [139]. Overexpression of the *P. fluorescens infC* operon can compete with RsmA, suggesting that coding regions could also regulate CsrA/RsmA activity [140]. Global approaches such as RIP-seq are ideally suited to identify additional members of this emerging class of cross-regulating mRNAs and will no doubt help to identify additional novel RNA antagonists of regulatory proteins and to decipher complex post-transcriptional networks.

The identification of *C. jejuni* CsrA titration by FliW in this study indicates that regulation of CsrA activity control by a protein antagonist, a mechanism first identified in *B. subtilis* [23], may be more widespread and ancient than anticipated. Coexistence of *csrA* and *fliW* homologues in many bacteria [141], as well as a common co-occurrence in flagellar operons, indicates a linked function. In *B. subtilis*, *csrA* and *fliW* are encoded next to *hag* which is controlled through the CsrA-FliW-Hag partner switch mechanism [23]. Interestingly, in *Lawsonia intracellularis,* FliW and CsrA are encoded as a single fusion protein and in *H. pylori* two *fliW* homologues are present (Fig. 31). Gammaproteobacteria seem to encode only for CsrA [141] and the antagonizing sRNAs [19]. A CsrB-like sRNA was also reported in *B. subtilis* [142], but it remains to be shown whether it can sequester CsrA. Interestingly, the *csrA* gene is inserted next to a tRNA cluster in *E. coli* and the CsrA-CsrB/C system is proposed to be acquired horizontally [23, 143]. While *E. coli* CsrA also binds and inhibits translation of its own mRNA [144], the *csrA* mRNA was not enriched the coIP dataset of this study. In summary, CsrA and its antagonists seem to have diverged in different bacterial species to perform different regulatory tasks using the same basic mechanism.

The protein-protein coIP experiments performed here confirm a direct interaction between *C. jejuni* CsrA and FliW, which was also suggested by a previous yeast two hybrid screen [145]. Besides its effect on flagellin protein levels via CsrA, FliW might also impact *flaA* expression at additional levels. A two-fold upregulation of transcription of *hag* mRNA was observed in *B. subtilis* upon *fliW* deletion [23, 146]. A two to three fold higher *flaA* mRNA levels was also observed upon *fliW* deletion in this thesis. It could also be shown that FliW affects the transcription from the *flaA* promoter. One of the possible explanations is that deletion of *fliW* leads to production of short flagella and hooks or short flagella are known to secrete FlgM, anti-sigma factor for σ^{28}, much more effectively in *B. subtilis* [147]. This might then lead to release of FlgM bound FliA and hence higher transcription from σ^{28}-dependent promoters such as that of *flaA*. Further studies are required to unravel the full complexity of the CsrA/FliW regulatory network and its impact on flagellar regulation. The data indicate FliW as the major CsrA antagonist in *C. jejuni.* However, the synergistic interplay of FliW together with the *flaA* mRNA-derived antagonist on other flagellar genes showed that RNA-based regulation may also impacts CsrA activity in this type of Csr network. The conserved, or possibly more ancient, function of the CsrA/FliW system might be to mediate temporal and spatial control of proper flagellum assembly, and the pleiotropic function of CsrA in Gammaproteobacteria may have been horizontally acquired, followed by evolution of antagonizing sRNAs. Interestingly, during the conservation analysis, it was observed that certain non-flagellated *Campylobacter* species, such as *C. hominis*,

C. gracilis, and *C. ureolyticus,* which lack flagellin homologs, also lack *csrA* and *fliW* homologs (data not shown), further supporting their conserved function in flagellar regulation.

Based on single-cell analysis and RNA-FISH, this study shows that *flaA* mRNA localizes to the poles of elongating cells. Introduction of stop mutations that prevent translation initiation abolished polar *flaA* mRNA localization, while silent mutations that maintained translation had no influence. This indicates *flaA* mRNA localization to the cell poles requires translation. Localization of bacterial mRNAs has only recently been described and, unlike in eukaryotes, the underlying mechanisms and regulation of this process is only poorly understood [36, 148]. Both translation-independent and -dependent mechanisms of localization have been described [39, 40]. Recently, it was shown in *E.coli* that the mRNA specifically encoding for inner membrane proteins were enriched near the membrane in a co-translation dependent manner by recognition of the signal peptide by the signal recognition particle (SRP) [41]. Interestingly, a *flaA* mRNA variant with a stop-codon mutation at the 101st codon partially retains its ability to localize to the cell poles. This indicates that translation of the N-terminal peptide is required for polar localization and that translation of the N-terminus of the protein might also direct the nascent peptide along with the translating mRNA to the secretion apparatus. However, unlike the signal recognition by SRP, little is known about how flagellar filament substrates are selected for secretion, as they do not appear to have a shared secretion signal primary sequence or a cleavable signal peptide. A requirement of N-terminal domains for flagellar secretion has been reported in diverse bacteria, including *C. jejuni* [149], and both 5' UTR and N-terminal peptide secretion signals have been shown to contribute to secretion efficiency [150]. In addition, type II chaperones play a role in regulating the coupling of translation to secretion of flagellar substrates [151]. FlhB autocleavage in the flagellar export apparatus is required for the switch from export of rod or hook subunits to filament subunits [121]. Mutants defective in FlhB cleavage can still export the rod and hook substrates but not the filament substrates [125]. Such an FlhB mutant (N267A) was also found to show patchy and more dispersed *flaA* mRNA localization and hence could link FlaA export to the localization of its mRNA. Surprisingly, deletion of another related protein, FlhX- a paralog of the cleaved C-terminal domain of FlhB ($FlhB_{CC}$), was found to localize *flaA* mRNA at both poles of all cells. It is possible that FlhX co-translationally binds to nascent FlaA N-terminal peptide in the cytoplasm and hence restricts its localization along with the *flaA* mRNA to the poles. Future studies are required to clarify the role of FlhB and FlhX in filament substrate recognition and secretion to further appreciate their impact on *flaA* mRNA localization.

Translation-dependent localization of *flaA* mRNA is further supported by disruption of its localization in CsrA-dependent manner upon deletion of FliW, which antagonizes CsrA-mediated translational repression of *flaA*. Despite an increase in FlaA-3xFLAG levels, deletion

of *csrA* had no impact on polar *flaA* localization. The limited CsrA activity in WT cells under standard growth conditions, due to sequestration by the FliW protein antagonist, probably allows sufficient translation of *flaA* mRNA for its polar localization. The diffuse localization of *flaA* mRNA in a Δ*fliW* mutant indicates that strong CsrA-mediated translation repression of *flaA* upon release of CsrA from its protein antagonist interferes with mRNA localization and reveals an unexpected role for CsrA in spatial control of mRNAs. CsrA binding might mediate storage of translationally inactive *flaA* mRNA until synthesis of FlaA is required or proper localization is achieved, similar to mRNP granules in eukaryotes [152]. Future studies will show whether other flagellar mRNAs also polarly localize and if CsrA and FliW also impact their localization. In addition, besides the requirement for its translation, other factors such as *flaA* genomic location or the spatial control of the transcriptional complex might also contribute to polar localization of *flaA* mRNA. CsrA-mediated regulation of mRNA localization might also occur in *B. subtilis* and *B. burgdorferi,* where the major flagellin is also repressed upon CsrA overexpression [96, 100, 153]. Analogous systems likely have also evolved in other flagellated bacteria, including the Alphaproteobacterium *Caulobacter crescentus,* which lacks CsrA but encodes two proteins, FlbT and FlaF, which have opposing activities on flagellin regulation. Like CsrA, FlbT post-transcriptionally regulates flagellin expression by binding to the 5' UTR of its encoding mRNA[154].

The observed motility defect upon *csrA* deletion indicates that tight FlaA regulation is important for proper flagellum expression or function. *C. jejuni* Δ*flaA* expresses stubby flagella and has a motility defect [155]. Consistent with reduced FlaA expression upon deletion of the CsrA protein antagonist FliW, the Δ*fliW* mutant strain also expresses short flagella and is also defective for autoagglutination. This agrees with the previous observation of a strong motility defect of a *fliW* mutant in both *B. subtilis* and *C. jejuni* [23, 112]. While CsrA impacts motility through directly controlling flagellin expression in *C. jejuni*, *B. subtilis*, and *Borrelia*, the strong motility defect of an *E. coli csrA* transposon mutant [17] is due to a requirement for CsrA-mediated stabilization of the mRNA of the master regulator FlhDC [18]. Besides its crucial role in motility, the flagellum plays an essential, multi-factorial role in colonization and pathogenesis of *C. jejuni* [156] and is essential for proper cell division [157]. The *C. jejuni* flagellar Type III secretion system secretes Cia/Fed effectors that are dispensable for motility but influence cell invasion, intracellular survival and chick colonization [48, 116, 158-160]. CsrA did not specifically bind the characterized effector mRNAs in the CsrA coIP. However, since the regulation of basal-body and hook components is required for secretion [48], future studies may reveal whether CsrA might affect virulence phenotypes beyond motility and whether FliW is environmentally regulated to control secretion.

4. CONCLUSION AND OUTLOOK

Due to the absence of a homolog of the RNA chaperone Hfq in Epsilonproteobacteria, the role of riboregulation and sRNAs in gene expression control in these bacteria have remained a matter of debate. In 2010, this discussion was categorically resolved by the discovery of a plethora of regulatory sRNAs in the pathogenic Epsilonproteobacterium *Helicobacter pylori* using a global RNA-seq based study [52]. Moreover, functional characterization of the first *trans*-acting sRNA candidate, RepG, in *H. pylori* strain 26695 confirmed sRNAs to be functional in Epsilonproteobacteria. This highly abundant and conserved sRNA post-transcriptionally regulates expression of a chemotaxis receptor, TlpB, in *H. pylori* through direct base-pairing interactions between its C/U-rich terminator loop and a homopolymeric G-repeat in the *tlpB* mRNA leader [161]. Overall, these studies have sparked an interest in sRNA repertoire and riboregulation in other Hfq lacking bacteria like *C. jejuni*, in the hope of finding novel modes of post-transcriptional regulation in bacteria.

The genome sequence of the related pathogenic Epsilonproteobacterium *C. jejuni* has provided a wealth of information on potential virulence factors as well as phase variable regions in this bacterium; however, very little is known about gene expression and control post-transcriptional regulation [43]. Comparison of genetically closely related *C. jejuni* strains isolated from different hosts indicated that *Campylobacter* uses high phenotypic flexibility and genetic microdiversity to reversibly adapt to changing environments [162]. In addition, variation in contingency loci contributes to rapid adaptation to new environments [163]. This extraordinary high variability among *C. jejuni* strains and isolates might be supported by strain-specific differences in transcription as well as post-transcriptional regulation. Expression profiling of sRNAs in multiple strains has been reported only for a limited number of bacteria and was mainly based on Northern blot analysis [164, 165]. A global RNA–seq based comparison of regulatory elements including sRNAs was recently carried out between the two closely related species, *Escherichia coli* and *Klebsiella pneumoniae*, and revealed that the majority of orthologous operons were transcribed from different promoters [166]. In this thesis, the impact of small variations in closely related genomes of different *C. jejuni* strains on the overall transcriptional architecture was investigated. Using comparative transcriptomics this study revealed SNPs that can mildly to completely abort transcription. Such “disrupting/inhibitory” SNPs are not limited to the -10 core promoter element, but were also found in the upstream A/T-rich periodic pattern and might change the local genome structure or interfere with transcription factor binding.

This study also revealed many conserved and strain specific sRNAs, however their functions and targets are not known. These sRNAs have no significant similarity outside the

Campylobacterales, hinting at later evolution to target *Campylobacter* specific genes. Several highly abundant sRNAs were also found on the plasmids of the virulent strain 81-176, which might have a potential role in virulence or plasmid maintenance.

In this thesis, the RNA interactome of the widespread post-transcriptional regulator CsrA was examined in two *C. jejuni* strains, NCTC11168 and 81-176. The RIP-seq analysis of CsrA revealed the absence of any CsrA antagonizing sRNAs in this bacterium. CsrA binding was observed mainly in the 5' UTR region of several genes but many examples of binding within two genes in operon or within the ORF regions was also observed. Previously, CsrA has been primarily known as translational repressor that binds on or near the RBS and thus interferes with translation initiation. The consequences of CsrA binding in the ORF are unknown, but it might be involved in premature translation termination or pausing. The resulting, shortened peptides could potentially also be functional and required under certain environmental conditions. Likewise, pausing of the ribosome might allow for proper folding of the leader peptide. The use of recently established [167] Ribosome profiling to map the ribosome density and pausing sites around the CsrA binding sites within the ORF might highlight upon such auxiliary roles of CsrA.

This thesis also revealed that many genes involved in flagellar biosynthesis are direct targets of CsrA. Furthermore, it could also be shown that the most abundantly co-purified transcript, the mRNA encoding the major flagellin FlaA, can also antagonize CsrA activity. The *flaA* mRNA together with the main CsrA antagonist, the FliW protein, titrates CsrA activity to regulate expression of other flagellar genes. This dual-functional *flaA* mRNA itself was found to be localized at the poles of short and presumably newly divided cells, hinting at temporal control of mRNA synthesis and localization over cell division. Moreover, It could also be shown that FliW affects translation-dependent localization of the flagellin mRNA to the cell poles by antagonizing CsrA-mediated translational repression of *flaA* mRNA. Together, this work identifies a new role for the CsrA/FliW network in spatial control of gene expression and a new mode by which CsrA activity can be regulated by a target mRNA-derived antagonist. However, some open and intriguing questions have also been raised in this thesis. It remains to be shown what signal is required to localize the mRNA. It could be shown that the translation of the first 100 aa of *flaA* mRNA partially restores its localization. It is not known whether the leader peptide is enough to localize the translating complex to the poles or it further requires a chaperone to bind the nascent leader peptide to assist in polar localization. Preliminary results from this study show that the mutants (FlhB and FlhX) in flagellar export apparatus, which disrupt of interfere with FlaA secretion also lead to critical changes in *flaA* mRNA localization. Beyond the mechanism, the functional importance of *flaA* mRNA localization also remains unclear. The localization could promote synthesis of FlaA at the required site for proper flagella assembly, but as a CsrA antagonist *flaA*

mRNA might also sequester CsrA at the poles to spatially regulate translation of other CsrA target genes. It would be challenging to uncouple the role of *flaA* localization from its translation but with new imaging techniques being developed to visualize the mRNA along with its translation at single molecule resolution in living cells, it may soon be possible to answer these questions [168].

5. MATERIAL AND METHODS

5.1. Materials

5.1.1. Instruments and devices

Instrument / Device	Manufacturer
analytical balances TE64, TE601	Satorius
Bio-Link BLX 254 UV-Crosslinker	Peqlab
cell culture hood, HERASafe	Thermo Scientific
Celltron Shaker	Infors HT
centrifuge Eppendorf 5415C	Eppendorf
centrifuge Eppendorf 5424R	Eppendorf
centrifuge Eppendorf 5418R	Eppendorf
centrifuge Heraus Multifuge X3R	Thermo Scientific
eraser for imaging plates	GE Healthcare
gel documentation system Gel iX Imager	Intas
gel dryer Bio-Rad Model 583	Bio-Rad
heat block Eppendorf comfort	Eppendorf
horizontal electrophoresis systems PerfectBlue Mini S, M, L	Peqlab
hybridization oven HB-1000	UVP
imaging system Image Quant LAS 4000	GE Healthcare
incubator for bacterial plates (*C. jejuni*), Galaxy 170R	Eppendorf
incubator for *C. jejuni* liquid cultures, HERAcell 150i	Thermo Scientific
incubator for *E. coli*, HERAcell (Kendro)	Thermo Scientific
PCR engine, T3 thermocycler	Biometra
PhosphoImager Typhoon FLA 7000	GE Healthcare
photometer Ultrospec 3100 pro Cell Density Meter	GE Healthcare
power supplies peqPOWER E250, E300	Peqlab
Orbital Shaker NB-101S RC	N-BIOTEK
Retsch MM400 ball mill	Retsch
Rotamax 120	Heidolph
rotator – SB2 STUART	STUART
scanner for protein gels, HP Scanjet 7400c	HP
semi-dry electroblotter PerfectBlue SEDEC M	Peqlab
shaker 37 °C room, SM-30	Bühler
SORVALL centrifuge RC5B	Thieme Labortechnik
spectrophotometer NanoDrop 2000	Peqlab
tank electroblotter PerfectBlue Web S, M	Peqlab
thermal cycler MJ Mini	Bio-Rad
Thermo Mixing Block MB-102	BIOER
vaccum pump	KnF LAB
vertical electrophoresis systems PerfectBlue Twin S, ExW S, L	Peqlab
vertical sequencing gel system CBS SG-400-20	C.B.S. Scientific

Victor3 1420 multilabel counter	Perkin-Elmer
Vortex-Genie 2	Scientific Industries
waterbath, GFL	Hartenstein

5.1.2. Labware and consumables

Labware	**Manufacturer**
250 ml buckets for centrifuge	Thieme Labortechnik
Bio-spin disposable chromatography columns	BioRad
boxes (plastic), 20.5 x 20.5 cm or 9.5 x 20.5 cm	Hartenstein
boxes (metal), 10 x 21 cm	Hartenstein
cell culture flasks 25 cm^3	PAA, Corning
cell culture flasks 75 cm^3	PAA, Corning
Cellstar serological pipets (plastic) 5 ml, 10 ml, 25 ml, 50 ml	Greiner bio-one
chromatography columns	Biorad
cotton swaps	DELTALAB, Stein Labortechnik
cover slips	Hartenstein
dewar canister	Hartenstein
Erlenmeyer glass flasks 250 ml, 1l	DURAN, SIMAX
G-25, G-50 MicroSpin columns	GE Healthcare
Genomic-tip 100/G	Qiagen
Gilson pipets 10 µl, 20 µl, 200 µl, 1000 µl	Gilson
glass beads (0.1 mm) for cell lysis	Roth, Biospec
glass beads (2.85 – 3.35 mm) for plating of CFU/ml	Roth
glass bottles	Schott
glass test tubes and lids	Roth
hard-shell PCR plates 96-well WHT/WHT	Biorad
Hybond-XL membrane for nucleic acid transfer	GE Healthcare
imaging plates BAS-IP MS 2325, 2340	Fujifilm
imaging plates cassettes BAS 2325, 2340	Fujifilm
inoculation loops 10 µl	VWR
L-shape bacteriology loops	VWR
object slides	Hartenstein
PCR tubes 8 x 0.5 ml	Thermo Scientific
petri dishes	Corning
Phase Lock Gel (PLG)-tubes, 2 ml	5 Prime
Pipetboy accu-jet pro	BRAND
pipet tips	Sarstedt
PolyScreen PVDF Transfer Membrane	PerkinElmer
racks for PCR tubes / reaction tubes	Hartenstein
reaction tubes 1.5 ml, 2.0 ml	Sarstedt
reagent and centrifuge tubes 15 ml, 50 ml	Sarstedt
safe-lock tubes 1.5 ml, 2.0 ml	Eppendorf
spectrophotometer cuvettes	BRAND
sterile filters (0.20 µm pore size)	Sarstedt

tube holder 15 ml, 50 ml	Hartenstein
Whatman paper	ALBET LabScience

5.1.3. Chemicals and reagents.

Chemical/reagent/protein/size marker	**Manufacturer**
2 x gel loading buffer II (RNA)	Ambion and self-made
10 x RNA structure buffer	Ambion
1 x alkaline hydrolysis buffer	Ambion
20 x SSC (saline sodium citrate buffer)	Sigma
acetic acid (100 %)	Roth
acetone	Roth
agarose	Roth
albumin Fraktion V (BSA)	Roth
ampicillin sodium salt	Roth
BBL™ Brucella Broth	Becton, Dickinson and Company
chloramphenicol	Roth
DEPC-water	Roth
Difco-agar	BD
dimethyl sulfoxide (DMSO)	Roth
dithiothreitol (DTT)	Roth
ethanol	Roth
ethanol (absolute for analysis)	Merck
ethylenediaminetetraacetic acid disodium salt dehydrate (EDTA)	Merck
formamide (99.5 %)	Roth
formaldehyde (37 %)	Roth
Gene Ruler 1 kb plus DNA ladder	Thermo Scientific
Gene Ruler 100 bp DNA ladder	Thermo Scientific
Genomic DNA buffer set	Qiagen
gentamicin sulfate	Roth
glycerol (99 %)	Sigma
GlycoBlue™	Ambion
hydrochloric acid (HCl, 32 %)	Roth
HygromycinB	Sigma
isopropanol	Roth
isopropyl-β-D-thiogalactopyranosid (IPTG)	Roth
kanamycin sulfate	Roth
L-arginin-monohydrochloride	Sigma
lead (II)-acetate	Roth, Fluka
magensium chloride	Roth
Midori Green	Nippon Genetics GmbH
milk powder (blotting grade)	Roth
Mueller-Hinton media	Becton, Dickinson and Company
PAGE Blue staining solution (Coomassie)	Thermo Scientific

PageRulerTM Plus Prestained Protein Ladder	Thermo Scientific
PBS	Gibco
phenol	Roth
pUC Marker Mix, 8	Thermo Scientific
rifampicin	Fluka
RNA Ladder High Range	Thermo Scientific
RNA Ladder Low Range	Thermo Scientific
Roti-Aqua-P/C/I	Roth
Roti-Hybri-Quick	Roth
Rotiphorese gel 40 (19:1)	Roth
Rotiphorese gel 40 (37.5:1)	Roth
sodium carbonate (Na_2CO_3)	Roth
sodium dodecyl sulfate (SDS), powder/pellets	Roth
sodium hydroxide (NaOH)	Roth
Triton-X100	Sigma
TRIzol Reagent	Invitrogen
Tween^{20}	Roth
uracil	Sigma
vancomycin sulfate	Roth
Western Lightning chemoluminescence reagent	PerkinElmer, self-made
yeast tRNA	Ambion
γ-^{32}P-ATP (^{32}P; 222TBq (6000Ci)/mmol 370MBq (10mCi)/ml)	Hartmann Analytic

5.1.4. Commercial kits

Kit	**Manufacturer**
BioPrime DNA Labeling System	BioPrime
CycleReader™ DNA sequencing	Ambion
GeneJET™ Gel extraction	Fermentas
LightShift Chemiluminescent EMSA	Pearce
Masterpure DNA Purification	Epicentre
MEGAscript T7 *in-vitro* Transcription	Ambion
MinElute PCR Purification	Qiagen
NucleoSpin Plasmid	Macherey-Nagel
NucleoSpin Gel and PCR Clean-up	Macherey-Nagel
Power SYBR GREEN RNA-to-CT™ 1-Step	Life Technologies
PureSystem	Cosmo Bio, New England Biolabs
QIAquick Gel Extraction	Qiagen
QIAquick PCR Purification	Qiagen
SequiTherm EXCEL™ II DNA Sequencing	Epicentre
SUPERscriptII Reverse Transcription	Invitrogen
SUPERscriptIII Reverse Transcription	Invitrogen

5.1.5. Enzymes

Enzymes	Manufacturer
Antarctic Phosphatase	New England Biolabs
AMV Reverse Transcriptase	New England Biolabs
Calf Intestinal Phosphatase (CIP, 10 u/µl)	New England Biolabs
Deoxyribonuclease (DNase I, 1 u/µl)	Thermo Scientific
DpnI (20 u /µl)	New England Biolabs
Klenow enzyme	BioPrime
Phusion High-Fidelity DNA polymerase (2 u/µl)	Thermo Scientific
Protease Inhibitor Cocktail (EDTA-free)	Roche
Proteinase K	Roth
Ribonuclease III (RNase III, 1.3 u/µl)	New England Biolabs
Ribonuclease A (RNase A)	Qiagen
Ribonuclease T1 (RNase T1, 1 u/µl)	Ambion
Shrimp Alkaline Phosphatase (SAP, 1 u/µl)	Thermo Scientific
SUPERaseIN RNase Inhibitor	Ambion
SUPERscriptII Reverse Transcriptase	Invitrogen
SUPERscriptIII Reverse Transcriptase	Invitrogen
T4 DNA Ligase (5 u/µl)	Thermo Scientific
T4 Polynucleotide Kinase (PNK, 10 u/µl)	Thermo Scientific
Taq DNA polymerase (5 u/µl)	New England Biolabs
lysozyme	Roth
diverse restriction enzymes (*NheI*, *XhoI*, *XbaI*, *NdeI*, *ClaI* etc.)	New England Biolabs, Thermo Scientific

5.1.6. Antibodies

Antibody	Origin	Dilution (3 % BSA/TBS-T)	Manufacturer
polyclonal anti-GroEL antiserum	rabbit	1:10,000	Sigma-Aldrich, # G6532-5ML
polyclonal anti-mCherry antibody	goat	1:4,000	Acris,#AB0040-20
monoclonal anti-FLAG M2 antibody	mouse	1:1,000	Sigma-Aldrich
monoclonal anti-GFP antibody	mouse	1:1,000	Roche, #11814460001
monoclonal anti-Strep antibody	mouse	1:10,000	IBA GmbH, #2-1507-001
ECL anti-mouse IgG, HRP-conjugate	goat	1:10,000	GE-Healthcare, #RPN4201
ECL anti-rabbit IgG, HRP-conjugate	donkey	1:10,000	GE-Healthcare, #RPN4301
anti-goat IgG, HRP-conjugate	donkey	1:10,000	Santa Cruz Biotechnology, #sc2020

5.1.7. Oligonucleotides

Oligonucleotides used in this thesis are listed in Appendix Table 12.

5.1.8. Strains Plasmids

Plasmid constructed/used in this thesis are listed in Appendix Table 13. Detailed description of strain construction is given in the Methods section.

5.1.9. Plasmids

Plasmid constructed/used in this thesis are listed in Appendix Table 14. Detailed description of plasmid construction is given in the Methods section.

5.1.10. Media, buffer and supplements

5.1.10.1. Media and stocks

Lennox Broth (LB) medium:	10 g	tryptone or peptone
	5 g	yeast extract
	5 g	NaCl
	ad 1 l H_2O	
LB-agar plates:	LB medium (see above)	
	1.5 % (w/v)	Difco-agar
Brucella Broth (BB):	28 g	BBL™ Brucella Broth
	ad 1 l H_2O	
after autoclaving, supplement with	10 µg/ml vancomycin	
Mueller Hinton-agar plates:	21g	Mueller Hinton Broth
	15 g	Difco-agar
	ad 1 l H_2O	
after autoclaving, supplement with	10 µg/ml	vancomycin
MH soft-agar motility plates (0.4 %):	21g	Mueller Hinton Broth
	4 g	Difco-agar
	ad 1 l H_2O	
after autoclaving, supplement with	10 µg/ml	vancomycin
SOC medium (transformation *E. coli*):	1 l SOB medium	
	add 5 ml magnesium chloride	
	add 20 ml 1 M glucose	
Superbroth medium (competent *E. coli*):	35 g	tryptone
	30 g	yeast extract
	5 g	NaCl
	ad 1 l H_2O	

5.1.10.2. Buffers and solutions

30:1 ethanol/sodium acetate (pH 6.5): (for RNA precipitation)	30 parts of 100 % ethanol 1 part of 3 M sodium acetate (pH 6.5)	
10 x DNA loading dye:	1.66 ml	1 M Tris-HCl (pH 7.5)
	12 ml	0.5 M EDTA (pH 8.0)
	0.05 g	bromophenol blue
	0.05 g	xylene cyanol

	60 ml	glycerol
	ad 100 ml H_2O	
2 x gel loading buffer II (RNA, GL II):	0.13 % (w/v)	SDS
	18 µM	EDTA (pH 8.0)
	95 %	formamide
	0.025 % (w/v)	bromphenol blue
	0.025 % (w/v)	xylene cyanol
5 x native sample buffer (gel-shifts):	50 %	glycerol
	0.02 %	bromophenol blue
	0.5 x	TBE buffer
5 x protein loading dye:	10 g	SDS pellets
	31.3 ml	1 M Tris-HCl (pH 6.8)
	50 ml	glycerol
	add 2.5 ml of 2 % (w/v) bromophenol blue	
	ad 100 ml H_2O	
1 x protein loading dye:	40 ml	5 x protein loading dye
	10 ml	2 M DTT
	ad 200 ml H_2O	
agarose gel electrophoresis solution:	X % (w/v)	agarose in 1 x TAE/TBE buffer
AMV reaction buffer (1 x, primer extension):	50 mM	Tris-HCl (pH 8.3)
	75 mM	potassium
	8 mM	magnesium
	10 mM	DTT
chemiluminescence solution A:	0.1 M	Tris-HCl (pH 8.6)
	0.025 % (w/v)	luminol
chemiluminescence solution B:	0.11 % (w/v)	*p*-coumaric acid (in DMSO)

PAA gel electrophoresis solution for western blots:

PAA gel for separation gel (10 ml)	**10 %**	**12 %**	**15 %**
1 M Tris "lower" buffer (pH 8.8)	3.75 ml	3.75 ml	3.75 ml
40 % PAA solution (37.5:1 acrylamide/bisacrylamide)	2.5 ml	3 ml	3.75 ml
H_2O	3.75 ml	3.25 ml	2.5 ml
10 % (w/v) SDS	100 µl	100 µl	100 µl
10 % (w/v) APS	75 µl	75 µl	75 µl
TEMED	7.5 µl	7.5 µl	7.5 µl

PAA gel for stacking gel (10 ml)	**4 %**
1 M Tris "upper" buffer (pH 6.8)	1.25 ml
40 % PAA solution (37.5:1 acrylamide/bisacrylamide)	1 ml
H_2O	7.5 ml
10 % (w/v) SDS	100 µl
10 % (w/v) APS	150 µl

TEMED	15 µl

PAA gel electrophoresis solution for northern blots and sequencing gels:

PAA gel (500 ml, stock solution, 7 M urea)	**6 %**	**10 %**	**15 %**
40 % PAA solution (19:1 acrylamide/bisacrylamide)	75 ml	125 ml	187.5 ml
urea	210 g	210 g	210 g
10 x TBE buffer	50 ml	50 ml	50 ml

1 gel (70 ml):	70 ml	stock solution
	700 µl	10 % (w/v) APS
	70 µl	TEMED

PAA gel electrophoresis solution for gel-shift assays (native PAGE):

1 gel (70 ml, 6 % PAA):	10.5 ml	40 % PAA sol. (19:1)
	3.5 ml	10 x TBE
	56 ml	H_2O
	700 µl	10 % (w/v) APS
	70 µl	TEMED
PBS (10 x stock):	80 g	sodium chloride
	2 g	potassium chloride
	17.7 g	disodiumhydrogen phosphate
	2.72 g	monopotassium phosphate
	ad 800 ml H_2O	
	adjust to pH 7.4	
	ad 1 l H_2O	
Proteinase K solution (LPS):	20 mg/ml	Proteinase K
RNA elution buffer (*in-vitro* transcription):	0.1 M	sodium acetate
	0.1 %	SDS
	10 mM	EDTA (pH 8.0)
RNA structure buffer (10 x):	100 mM	Tris-HCl (pH 7.0)
	1 M	potassium chloride
	100 mM	magnesium chloride
SDS running buffer (10 x stock):	30.275 g	Tris base
	144 g	glycin
	10 g	SDS
	ad 1 l H_2O	
SSC (saline-sodium citrate) buffer (20 x stock):	173.5 g	sodium chloride
	88.2 g	sodium citrate
	ad 800 ml H_2O	
	adjust to pH 7.0 (using HCl)	
	ad 1 l H_2O	

Stains-All (*in-vitro* transcription):	30 ml	Stains-All stock
	90 ml	formamide
	ad 200 ml H_2O	
Stains-All stock (*in-vitro* transcription):	0.03 g dissolved in 30 ml formamide	
Stop mix (RNA preparation):	95 % (v/v)	ethanol (absolute)
	5 % (v/v)	phenol
Stop solution (LPS silver-staining):	50 mM	EDTA
TAE buffer (50 x stock):	242 g	Tris base
	51.7 ml	acetic acid
	100 ml	0.5 M EDTA (pH 8.0)
	ad 1 l H_2O	
TBE buffer (10 x stock):	108 g	Tris base
	55 g	boric acid
	40 ml	0.5 M EDTA (pH 8.0)
	ad 1 l H_2O	
Tbf I buffer (competent *E. coli* TOP 10):	1.47 g	potassium acetate
	4.975 g	manganese(II)-chloride
	3.73 g	potassium chloride
	ad 400 ml H_2O	
Tbf I buffer (competent *E. coli* TOP 10, continued):	adjust pH to 5.8 (using CH_3COOH)	
	add 75 ml glycerin	
	ad 500 ml H_2O	
Tbf II buffer (competent *E. coli* TOP 10):	2 ml	1 M MOPS
	150 ml	0.1 M calcium chloride
	8 ml	250 mM potassium chloride
	30 ml	glycerin
	ad 200 ml H_2O	
TBS buffer (10 x stock):	24.11 g Tris base	
	87.66 g NaCl	
	adjust to pH 7.4 (using HCl)	
	ad 1 l H_2O	
TBS-T buffer (1 x):	100 ml	10 x stock solution
	1 ml	Tween^{20}
	ad 1 l H_2O	
TE buffer (1 x):	100 mM	Tris-HCl (pH 8.0)
	10 mM	EDTA (pH 8.0)

Transfer buffer (10 x stock):	30 g 144 g ad 1 l H_2O	Tris base glycin
Transfer buffer (1 x):	100 ml 200 ml ad 1 l H_2O	10 x stock solution methanol
Tris “lower buffer” solution:	1.5 M 0.4% (w/v)	Tris-HCl (pH 8.8) SDS
Tris “upper buffer” solution:	0.5 M 0.4% (w/v)	Tris-HCl (pH 6.8) SDS
Western development solution:	2 mL chemiluminescence solution A 200 µl chemiluminescence solution B 6 µL 3% (v/v) H_2O_2	

5.1.10.3. Sterilization

All media and solutions used in this study were sterilized prior to use by autoclaving at 120 °C and 1 bar atmospheric pressure for 20 min. Heat-sensitive solutions were sterile filtered. Glassware was sterilized by heating to 80 °C for a minimum of three hours.

5.2. Methods

5.2.1. Phylogenetic tree of Epsilonproteobacteria

To determine the phylogenetic relationship between the analyzed *Campylobacter jejuni* strains and close relatives, the sequences of twelve conserved proteins (namely Ffh, FusA, GyrB, InfB, LepA, PyrG, RplB, RpsE, RpsH, RpsK, TopA, and Tuf), which were previously proposed for phylogenetic analyses [54], of *Campylobacter jejuni* NCTC11168 were retrieved. For each protein, the ortholog in 21 selected species (*Helicobacter acinonychis* str. Sheeba, *Helicobacter felis* ATCC 49179, *Helicobacter hepaticus* ATCC 51449, *Helicobacter mustelae* 12198, *Helicobacter pylori* 26695, *Wolinella succinogenes, Arcobacter butzleri* RM4018, *Campylobacter concisus* 13826, *Campylobacter curvus* 525.92, *Campylobacter fetus* subsp. fetus 82-40, *Campylobacter hominis* ATCC BAA-381, *Campylobacter jejuni* subsp. *doylei* 269.97, *Campylobacter lari* RM2100, and *Campylobacter jejuni* strains CG8421, RM1221, 81116, 81-176, IA3902, ICDCCJ07001, M1, and S3) was identified by *BLAST* searches [1] of the protein sequence against all protein sequences of the considered organism. The best hit of the *BLAST* search was considered as the ortholog. This search led to twelve protein collections with 22 members each. For each group of orthologs, a multiple

sequence alignment was generated using the program *muscle* [169]. These alignments were concatenated to one single alignment with a total length of 6,376 amino acid columns. This alignment was filtered by *gblocks* [170] to remove poorly aligned positions and variant regions resulting in an alignment of 4442 positions. Based on this hybrid alignment, 4,500 bootstrapped trees were generated by PhyML 3.0 [171] applying the LG matrix [172]. A consensus tree was generated using program *consense* from the Phylip package (Felsenstein, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c; distributed by the author (Department of Genetics, University of Washington, Seattle). In the next step *PhyML* estimated the branch lengths based on the topology of consensus tree and the previously described sequence alignment. The bootstrap values and branch lengths were merged into one file in the Newick format using *clann* [173]. For an easier comparison, the branch lengths were calculated as $\log_e$ values and normalized. Visualization and manual re-rooting was done in the webtool *iTOL* [174].

5.2.2. Transformation of *C. jejuni* for mutant construction

Transformation of *C. jejuni* was performed by electroporation or natural transformation as described previously [97, 175]. For electroporation, strains grown from frozen stocks until passage one or two on MH agar were harvested into cold electroporation buffer (272 mM sucrose, 15% v/v glycerol) and washed twice with the same buffer. Cells (50 µl) were mixed with 200-400 ng PCR product on ice and electroporated (Biorad MicroPulser) in a 1 mm gap cuvette (PEQLAB) at 2.5 kV. Cells were then transferred with Brucella broth to a non-selective MH plate and recovered overnight at 37 °C microaerobically before plating on the appropriate selective medium.

In some cases, *C. jejuni* double or triple mutants were constructed by natural transformation of the genomic DNA from the appropriate donor strain[176, 177]. Genomic DNA was extracted from the donor strain by phenol-chloroform extraction and ethanol precipitation. Specifically, bacteria were harvested from one-day-old selective MH plates into SET buffer (150 mM NaCl, 15 mM EDTA, 10 mM Tris-HCl; pH 8.0), collected by centrifugation, and resuspended in SET buffer. SDS and proteinase K were then added to final concentrations of 0.5% (w/v) and 100 µg/ml, respectively, and suspensions were incubated at 55 °C for 2h. Protein was then removed by extraction with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1), separation of phases by centrifugation at 13,000 rpm for 8 min, and re-extraction of the aqueous phase with an equal volume of chloroform with centrifugation at 13,000 rpm to separate phases. DNA was then precipitated from the final aqueous phase with 1/10 vol. 3M sodium acetate, pH 5.3 and 2 vol. absolute ethanol. After overnight incubation at -20 °C, precipitated DNA was collected by centrifugation at 13,000 rpm for 10 min and washed once with 75% cold ethanol. DNA pellets were resuspended in 100 µl water with shaking at 65 °C. For transformations, recipient strains

were grown from frozen stocks, patched into small circles on a non-selective MH plate, and grown for 2-3 h at 37 °C under microaerobic conditions. One hundred ng of donor genomic DNA (gDNA) was then added to the patches and plates were incubated for an additional 4-5h. Patched cells were then harvested into 1 ml Brucella broth and 10 or 100 µl was plated on the appropriate selective MH agar. Colonies were re-streaked onto selective plates, and colony PCR was performed to confirm presence of desired mutations from both donor and recipient strain.

5.2.3. Construction of *C. jejuni* deletion strains by overlap PCR

All *C. jejuni* deletion mutant strains listed in Appendix Table 13 were generated by double-crossover homologous recombination with PCR products of deletion cassettes that were constructed by overlap PCR (for details see Appendix Table 15) and electroporated into bacteria as described above. PCR products carried *aphA-3* kanamycin[178], *C. coli cat* chloramphenicol[179], *aac*(3)-IV gentamicin[180], or *aph*(7″) hygromycin[181] resistance cassettes flanked by around 500 bp of homologous sequence up- and downstream of the coding region of the target gene. Non-polar resistance cassettes were amplified from plasmids that carry the resistance markers using primers HPK1/HPK2 (Kan^R), CSO-1678/-1679 (Hyg^R), or CSO-0613/-0614 (Cm^R). The *aphA-3* (Kan^R) ORF was replaced by the *aac*(3)-IV (Gm^R) ORF (amplified using CSO-0575/-0832 and *Nde*I digested) in the plasmid pGG1 (amplified using CSO-0577/-0831 and *Nde*I digested) leaving the HPK1/HPK2 binding sites intact. The resulting plasmid pGD78-1 was used to amplify the *aac*(3)-IV (Gm^R) cassette using the same HPK1/HPK2 primers.

As an example, the construction of the chloramphenicol resistant *C. jejuni* NCTC11168 Δ*csrA*::*cat* deletion mutant is described. About 500 bp upstream of the *csrA* (Cj1103) start codon was amplified from genomic DNA (gDNA) of *C. jejuni* NCTC11168 WT using 'UP' primers (CSO-0394/-0615). Likewise, around 500bp downstream of the *csrA* stop codon was amplified using 'DN' primers (CSO-0616/-0395). The 5' ends of the antisense-UP primer and sense-DN primer contained around 25 bp of sequence homologous to the sense or antisense primer (CSO-0613/-0614), respectively, used to amplify the *cat* resistance cassette. PCR products were purified (Macherey-Nagel NucleoSpin PCR cleanup kit), and UP, DN, and resistance cassette amplicons were then added together in a ratio of 50:50:90 ng to a 100 µl Phusion polymerase PCR reaction with sense-UP and antisense-DN primers (CSO-0394/-0395) at a final concentration of 0.06 µM. Overlap PCR was performed with the following conditions: 1 cycle of [98 °C, 3 min; 61 °C, 1 min; 72 °C, 10 min; 98 °C, 1 min], 40 cycles of [98 °C, 15 s; 57 °C, 20 s; 72 °C, 1 min], followed by a 10 min final extension at 72 °C. Following verification of product size by agarose gel electrophoresis and purification (Macherey-Nagel NucleoSpin PCR cleanup kit), the resulting overlap PCR product was electroporated into the appropriate recipient *C. jejuni* strain. Deletion

mutants for *rnc*::Kan^R, *flaA*::Kan^R, *flaB*::Kan^R, *flaAB*::Kan^R, *fliW*::Gm^R, *fliW*::Hyg^R *fliA*::Gm^R, *rpoN*::Gm^R, and *csrA*::Hyg^R were constructed similarly (see Appendix Table 15).

5.2.4. Construction of chromosomal 3xFLAG epitope-tagged proteins in *C. jejuni*

C. jejuni genes were chromosomally tagged at their C-terminus either by cloning of constructs for C-terminal epitope tagging on plasmids or by construction of 3xFLAG constructs by overlap PCR.

Tagging of proteins using PCR products amplified from plasmid constructs. The CsrA, FlaA, FlgI, and FlaB proteins were fused to a 3xFLAG epitope at their C-termini by cloning regions encoding their "C-term" and "DN" regions into plasmid pGG1 to flank a 3xFLAG tag and Kan^R cassette. Afterwards, the 3xFLAG-tag constructs were amplified by PCR and introduced into the chromosome of *C. jejuni* strains by electroporation and double-crossover homologous recombination. An example of this plasmid cloning strategy is described for *csrA.* Approximately 500 bp of the region downstream of *csrA* was amplified from gDNA with primers CSO-0173/-0174. These primers included *Xba*I and *Eco*RI sites, respectively. Following cleanup, the PCR product was digested with *Eco*RI and *Xba*I and ligated into a similarly-digested pGG1 backbone, generated by inverse PCR with primers CSO-0074/-0075, to create pGD2-1. The plasmid was verified by colony PCR with primers JVO-0054/CSO-0173 and sequence verified using JVO-0054. Next, the backbone of this plasmid, including the *csrA* "DN" region, was amplified by PCR with primers CSO-0073 (*Xho*I) and JVO-5142 (blunt). The C-terminal coding region of *csrA* (around 500 bp) without the stop codon was amplified with primers CSO-0171/-0172 from NCTC11168 WT gDNA. The sense primer (CSO-0172) included an *Xho*I site, whereas the antisense primer (CSO-0171) contained a 5'-phosphate. Both the plasmid backbone with the "DN" insert and the C-term insert were digested with *Xho*I and ligated to create plasmid pGD4-1. Integration of the PCR product was confirmed by colony PCR using primers CSO-0172/-0023 and the plasmid was validated by sequencing using CSO-0023. The entire integration cassette was then amplified with Phusion polymerase (NEB) using primers CSO-0172/-0173 and electroporated into *C. jejuni* and selected on kanamycin plates. Mutants were confirmed by colony PCR with primers CSO-0196/-0023 and Western blot analysis with an anti-FLAG antibody.

*3xFLAG tagging of proteins by overlap PCR***.** Construction of a C-terminal 3xFLAG translational fusion at its native locus was performed by overlap PCR for *flaG* as described above for gene deletions, but with the following modifications. The final overlap PCR product contained around 500 bp of the C-terminal coding region of *flaG* minus the stop codon (C-term) and around 500 bp downstream of *flaG* (DN) for homologous recombination. These regions flanked an in-frame 3xFLAG tag and stop codon followed by an *aphA-3* Kan^R cassette. For example, for tagging *flaG,*

the 3xFLAG tag and Kan^R cassette was amplified from plasmid pGG1 with primers JVO-5142 and HPK2. The "C-term" region of *flaG* was amplified using primers CSO-1002/-1098, where CSO-1098 is antisense and contains region of complementarity at its 5' end to the 3xFLAG tag/JVO-5142, from NCTC11168 gDNA. The "DN" region was amplified using primers CSO-1099/-1003, where CSO-1099 is sense to *flaG* DN and contains a region of complementarity to the 3' end of the Kan^R cassette/primer HPK2. If the coding region of the target gene contained sequences required for expression of a downstream ORF (i.e., SD sequence or codons), these sequences were included in the "DN" amplicon. The three PCR products were then used for overlap PCR with primers CSO-1002/-1003, and the resulting amplicon was electroporated into *C. jejuni,* followed by selection of positive clones on kanamycin plates. Mutants were checked by colony PCR with primers CSO-1005/HPK2 and Western blot analysis with an anti-FLAG antibody.

5.2.5. Introducing chromosomal point mutations into the *flaA* 5' UTR at its native locus

To introduce point mutations in the 5' UTR of *flaA* at the native locus, a 1,100 bp region around the *flaA* promoter was amplified using oligos CSO-0752/-0753. These primers introduced *Xho*I and *Xba*I sites, respectively, into the resulting PCR product. After *Xho*I and *Xba*I digestion, the product was then ligated into a similarly-digested plasmid pJV752-1, resulting in plasmid pGD70-5. Plasmid pGD70-5 was checked by colony PCR using primers pZE-A/CSO-0753 and sequencing with pZE-A. Next, plasmid pGD70-5 was amplified by inverse PCR using primers CSO-0754/-0755, thereby introducing *Nde*I and *Bam*HI restriction sites 40 nt upstream of the *flaA* TSS. An *aac*(3)-IV gentamicin resistance cassette with its own promoter and terminator was amplified using CSO-0483/-0576 and introduced into PCR-amplified pGD70-5 in the reverse orientation to *flaA,* just upstream of its promoter, using the *Nde*I/*Bam*HI restriction sites, resulting in plasmid pGD76-1. Plasmid pGD76-1 was checked by colony PCR using primers CSO-0576/-0753 and sequencing with CSO-0753.

Point mutations were then introduced into the *flaA* 5' UTR by inverse PCR on pGD76-1 using complementary oligos harboring the desired mutation, followed by *Dpn*I digestion and transformation of the resulting purified PCR product into *E. coli* TOP10. For introduction of the *flaA* M1 mutation (GGA>AAA in stem-loop SL1 of the *flaA* leader), CSO-1114/-1115 were used for PCR on pGD76-1. The mutation was confirmed in the resulting plasmid pGD92-1 by sequencing with CSO-0753. Similarly, the *flaA* M2 (GGA>UGA in stem-loop SL1 of the *flaA* leader), M3 (GGA>GGG in stem-loop SL2 of the *flaA* leader), X1 start codon (AUG>AAG), X2 start codon (AUG>AUU), X3 start codon (AUG>GUG), X4 3rd codon (UUU>UAG), X5 3rd codon (UUU>UUC) and X6 101st codon (CAA>UAA) mutations were introduced using primer pairs CSO-0757/-0758, CSO-1116/-1117, CSO-2019/-2020, CSO-2827/-2828, CSO-2825/-2826, CSO-2829/-2830, CSO-2831/-

2832 and CSO-2833/-2834 respectively, resulting in plasmids pGD77-1, pGD93-1, pGD114-2, pGD205-1, pGD204-1, pGD206-1, pGD207-1 and pGD208-1 respectively. For combination of the *flaA* M2 and M3 mutations, a similar mutagenesis approach was performed based on PCR amplification of the M2 plasmid pGD77-1 using oligonucleotides CSO-1116/-1117, resulting in pGD95-1 harboring both the mutations. To introduce the *flaA* 5' UTR mutations into *C. jejuni,* a PCR product covering the homologous ends and the gentamicin resistance cassette was amplified from the respective WT (pGD76-1) or mutant plasmids using CSO-0752/-0850 and electroporated into *C. jejuni* as described above. To confirm introduction of point mutation in *C. jejuni,* colony PCR was performed using CSO-0576/-0753 and sequencing with CSO-0850.

5.2.6. Construction of *E. coli* mutants

The *E. coli* Δ*pgaA* and Δ*pgaA* Δ*csrA* deletion strains were constructed in the TOP10 background using the λ Red protocol[182]. Briefly, the kanamycin resistance gene, amplified using primers CSO-0652/-0653, was used to replace the entire *pgaA* ORF excluding the start and stop codon. The mutant strain was verified by colony PCR using the primer pairs CSO-0654/-0653 and CSO-0652/-0655. After verification, helper plasmid pCP20 containing FLP recombinase was introduced to remove the kanamycin resistance marker [182]. The helper plasmid, which is temperature-sensitive and carries an ampicillin resistance marker, was then cured by recovering colonies at 37 °C and confirming ampicillin sensitivity, resulting in strain CSS-0556. Similarly, the ORF of the *csrA* gene excluding the start and stop codon was then replaced by the kanamycin resistance marker (amplified using CSO-0611/-0612) in the Δ*pgaA* strain resulting in strain CSS-0557, harboring both *pgaA* and *csrA* deletions. The *csrA* deletion was verified by colony PCR using primer pairs CSO-0639/-0612 and CSO-0611/-0640.

5.2.8. Construction of dRNA-seq libraries

cDNA libraries for Solexa sequencing (HiSeq) were constructed by *vertis* Biotechnologie AG, Germany (http://www.vertis-biotech.com/) as described previously for eukaryotic microRNA [183] but omitting the RNA size-fractionation step prior to cDNA synthesis. In brief, equal amounts of RNA samples were poly(A)-tailed using poly(A) polymerase. Then, the 5'PPP structures were removed using tobacco acid pyrophosphatase (TAP). Afterwards, an RNA adapter was ligated to the 5´-phosphate of the RNA. First-strand cDNA was synthesis by an oligo(dT)-adapter primer and the M-MLV reverse transcriptase. In a PCR-based amplification step using a high fidelity DNA polymerase the cDNA concentration was increased to 20-30 ng/µl. A library-specific

barcode for multiplex sequencing was part of a 3'-sequencing adapter. The following adapter sequences flank the cDNA inserts:

TrueSeq_Sense_primer

5´AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3'

TrueSeq_Antisense_ NNNNNN_primer (NNNNNN = 6n barcode for multiplexing)

5'-CAAGCAGAAGACGGCATACGAGAT-NNNNNN-

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC(dT25)-3'

The Agencourt AMPure XP kit (Beckman Coulter Genomics) was used to purify the DNA which was analyzed by capillary electrophoresis afterwards.

5.2.9. Read mapping and coverage plot construction

To assure a high sequence quality, the Illumina reads in FASTQ format were trimmed with a cut-off phred score of 20 by the program *fastq_quality_trimmer* from FASTX toolkit version 0.0.13. After trimming, poly(A)-tail sequences were removed and a size filtering step was applied in which sequences shorter than 12 nt were eliminated. The collections of remaining reads were mapped to the respective reference genomes using *segemehl* [184]. Coverage plots representing the number of mapped reads per nucleotide were generated based on the mapped reads and visualized in the *Integrated Genome Browser* [185]. Each graph was normalized to the number of reads that could be mapped from the respective library. To restore the original data range, each graph was then multiplied by the minimum number of mapped reads calculated over all libraries.

5.2.10. Normalization of expression graphs

Prior to the comparative analysis, the expression graph data that resulted from the read mapping were further normalized. A percentile normalization step was applied to normalize the TEX+ graphs. To this end, the 90^{th} percentile of all data values was calculated for each TEX+ graph. This value was then used to normalize the TEX+ graph as well as the respective TEX- graph. Thus, the relative differences between each TEX+ and TEX- graph were not changed in this normalization step. Again, all graphs were multiplied with the overall lowest value to restore the original data range. To account for different enrichment rates, a third normalization step was applied. During this step, prediction of TSS candidates was performed for each replicate of each strain. These candidates were then used to determine the median enrichment factor for each TEX+/- library

pair. Using these medians all TEX- libraries were then normalized against the library with the strongest enrichment.

5.2.11. Transcriptional start site (TSS) annotation

Based on the normalized expression graphs, the TSS annotation approach was first applied to each strain individually. The parameters used for the TSS annotation were adjusted based on the *H. pylori* dRNA-seq training data set [52] for which TSS have been annotated manually and, thus, could be used for comparison. Based on this criterion, the threshold for the minimum flank height was set to 3.45 (see Appendix Fig. 16). If the TSS candidate reaches this threshold in at least one strain, the threshold is decreased for the other strains to 1.15. The threshold for the minimum factor of height change is set to 2.0, which, if reached in at least one strain, drops to 1.5 for the others. A TSS candidate was considered to be enriched in a strain if the respective enrichment factor is at least 2.0. A TSS candidate has to be enriched in at least one strain and is discarded otherwise. If a TSS candidate is not enriched in a strain but still reaches the other thresholds it is only indicated as “detected”. However, a TSS candidate can only be labeled as detected in a strain if its enrichment factor is above 0.66. Otherwise it was considered to be a processing site.

5.2.12. Identification/determination of orthologs

Clusters of orthologous genes among the four *Campylobacter jejuni* strains were generated as follows: 1) Each gene was *blasted* [1] against all genes of the other three strains with default parameters, except for using a word length of 10. 2) Orthologous gene pairs for each pair of strains were created by searching for best reciprocal hits of genes. Only *BLAST* matches with an e-value of 0.01 or below and an alignment length of at least 60% of the query length were taken into account. 3) These pairs of best reciprocal hits were merged into clusters of orthologs by grouping pairs that share one gene. In total there were 1,454 clusters with 4 orthologs, 78 clusters with 3 orthologs, 132 clusters with 2 orthologs and 813 genes without any ortholog (Appendix Table 4).

5.2.13. Consensus based correction of gene annotations

The length comparison of orthologous ORFs in the four strains revealed inconsistencies in the annotations for several genes. Some genes were longer or shorter than their orthologs, i.e. differed at their 5’ end, although the surrounding genomic region was conserved or at least very similar. Due to these annotation differences, a consensus based correction of such annotations was performed in which only aberrations of the 5' end were tried to be corrected. To determine if

the 5' end of a gene with aberrant length was similar to the sequence of its orthologs, the first 10 nucleotides of each ortholog-group member were extracted and the Hamming distances of the start sequence of the aberrant gene and the start sequence of each of its orthologs was calculated. If the Hamming distance exceeded the value 1, the 5' end was assumed to be different. In the case that there were three or two orthologs with the same and one with an aberrant sequence length or in the case of two with the same and two with aberrant lengths, the length of the majority was taken as target value for the aberrant ones. Based on the genomic sequences the annotation of the aberrant gene was extended or shortened to the target length. If the sequence of the modified annotation started with a potential start codon (ATG, TTG or GTG) the correction was accepted. In the case that there were two pairs of genes with the same sequence length, the length of both members of each pair to the length of the other pair was corrected as done above. If both pairs could be shortened or extended to the length of the other and fulfilled the requirement of having a valid start codon, manual inspection of the ortholog group was used to decide how to perform the correction. If all 3 or 4 members of a group of orthologs differed in length, or if there were only two orthologs in a group, no correction attempt was performed. The corrected annotations are listed in Appendix Table 5.

5.2.14. Transcriptional start site (TSS) annotation and SuperGenome approach

A whole genome alignment of the four *C. jejuni* strains was computed with *Mauve* [186]. Based on this global alignment a common genomic coordinate system, the SuperGenome was defined into which all positional information can be projected that relates to the single genomes [187]. This resulted in a consensus sequence with the coordinates of the complete alignment and a mapping of each position of each single genome to a position in the alignment. Next, all genome-specific data (expression height graphs derived from mapped read data, genomic annotations and sequences) were mapped to the common coordinate system.

The automated TSS prediction approach, which uses this SuperGenome mapping for comparative analyses, consists of several steps: The initial detection of TSS in the single strains is based on the localization of positions, where a significant number of reads start. Thus, for each position i in the RNA-seq graph corresponding to the TEX+ library the algorithm calculates e(i)-e(i-1), where e(i) is the expression height at position i (Appendix Fig. 16). In addition, the factor of height change is calculated, i.e. e(i)/e(i-1). To evaluate if the reads starting at this position are originating from primary transcripts the enrichment factor is calculated as $e_{TEX+}(i)/e_{TEX-}(i)$. For all positions where these values exceed the threshold (see Supporting information) a TSS candidate is annotated.

The TSS prediction procedure is applied to both replicates of each strain. TSS candidates, which are not detected in both replicates with a maximal positional difference of one nucleotide, are discarded. Afterwards, TSS candidates that are in close vicinity are grouped into a cluster and only the TSS candidate with the highest expression is kept. In the next step the TSS candidates of each strain are mapped to the SuperGenome to assign each TSS to the corresponding TSS in the other strains. The final TSS annotations are then characterized on the SuperGenome level with respect to their occurrence in the different strains and in which strains they appear to be enriched. In the context of the individual strains the TSS are further classified according to their location relative to annotated genes. For this a similar classification scheme was used as previously described [52]. Thus for each TSS it is decided if it is the *primary* or *secondary* TSS of a gene, if it is an *internal* TSS, an *antisense* TSS or if it cannot be assigned to one of these classes (*orphan*). A TSS is classified as *primary* or *secondary* if it is located ≤300 bp upstream of a gene. The TSS with the strongest expression considering all strains is classified as *primary.* All other TSS that are assigned to the same gene are classified as *secondary*. *Internal* TSS are located within an annotated gene on the sense strand and *antisense* TSS are located inside a gene or within ≤100 bp on the antisense strand. These assignments are indicated by a 1 in the respective column of Appendix Table 1. *Orphan* TSS, which are not in the vicinity of an annotated gene, are indicated by "0" in all four columns.

To validate the automated TSS detection used in this study, it was applied to the previously generated dRNAseq data of *Helicobacter pylori* grown under five different environmental conditions [52]. The *H. pylori* study had manually annotated the TSS based on enrichment patterns in the TEX+ compared to TEX- libraries. These hand-curated TSS positions served as benchmark and were compared to the results of the automated detection. A difference of up to one nucleotide was allowed when comparing an automatically detected TSS to a manually annotated TSS. With this threshold, the automated approach achieves a sensitivity of 82% and a precision rate of 75%.

5.2.15. Promoter and RBS motifs detection and data visualization

To detect potential promoter motifs, sequence regions corresponding to 50 nt upstream of the TSS positions and the TSS position itself were scanned by *MEME* version 4.8.1 [63] with width parameters (fixed width of 45 nt as well as flexible widths). For the detection of ribosome binding site (RBS) motifs, the 5' UTR sequences of mRNAs were inspected by *MEME*. The 5' UTR length distributions were visualized using *R* and the *ggplot2* package.

5.2.16. Comparison of 5' UTR lengths of orthologs

The 5' UTR lengths of orthologous genes were compared for pairs of strains (Appendix Table 6) and visualized as scatter plots (Fig. 9). For this purpose, the types of the originating TSS (primary or secondary) of the 5' UTRs were discriminated. If one or both genes of an ortholog pair had more than one TSS of the same type, 5' UTR pairs with the minimal distances were selected from all possible constellations. In case a 5' UTR had no partner of the same length and same type but a partner of same length and different type, the mixed type comparison was considered for the plotting.

5.2.17. Searches of alternative sigma factor promoter sequences using regular expressions

To find genes that contain promoter sequences of the alternative sigma factors, σ^{28} and σ^{54}, that were not detected during the MEME motif generation, the -50 to +1 sequences for each TSS were scanned with regular expressions: "TGG.ACA.[5]TGCTT" for σ^{54} and "TTT.[10,12]CGAT(AT|TT|TA)" for σ^{28} using the UNIX tool *grep* with the parameter "-E". In the SuperGenome, 172 sequences had hits for the σ^{28} pattern and 36 hits for the σ^{54} pattern (see Appendix Table 3).

5.2.18. Conservation analysis of sRNAs

To study the conservation of sRNAs in Epsilonproteobacteria, homologous sequences were searched with *blastn* (part of the BLAST+ package version 2.2.26 [1]; the word size parameter was set to 10 nt). The number of identical nucleotides of the best hits of each sRNA candidate was divided by the total number of nucleotides of the query sRNA and multiplied by 100 to calculate the percentage value of conserved nucleotides. For visualization, conservation values with > 40% identity were translated into a gray scale and values below 40% were depicted as white boxes.

5.2.19. Purification of recombinant *C. jejuni* CsrA-Strep from *E. coli*

The *csrA* gene, including SD sequence, was fused to a C-terminal Strep-tag in the arabinose-inducible plasmid pBAD-His-Myc (Invitrogen) for overexpression and affinity purification. The *csrA* coding region and SD were amplified from *C. jejuni* 11168 genomic DNA using primers CSO-0746/-0747, and the pBAD-His-Myc plasmid was amplified by inverse PCR with JVO-0900/-0901 as previously described [188]. CSO-0747 and JVO-0901 introduce an *Xba*I site to the insert and vector, respectively, whereas CSO-0746 has a 5'-phosphate to facilitate blunt-end ligation. *Xba*I-digested insert and vector were then ligated, resulting in pGD68-1. Plasmid pGD68-1 was checked by colony PCR using primers pBAD-FW/CSO-0747 and sequencing with pBAD-FW. A Strep-tag (WSHPQFEK) was then added at the C-terminus of *csrA* by inverse PCR using oligos CSO-0852/-

0853, resulting in plasmid pGD72-3. Plasmid pGD72-3 was checked by sequencing with pBAD-FW. Plasmid pGD72-3 was then introduced into an *E. coli* TOP10 Δ*pgaA*/Δ*csrA* deletion strain resulting in strain CSS-0931. Strain CSS-0931 was grown in 500 ml LB broth with 100 µg/ml of ampicillin at 37°C and 220 rpm shaking to an OD_{600} of 0.3, at which time L-arabinose was added to a final conc. of 0.01 % and further incubated for 8 h at 18 °C. Cells were harvested by centrifugation at 7000 x g for 30 min at 4 °C. The pellet was resuspended in 5 ml Buffer W (IBA GmbH, #2-1003-100). The rest of the protocol was followed as per manufacturer's instructions using 1 ml Gravity flow Strep-Tactin® Sepharose® (IBA GmbH, #2-1202-001). After washing steps, the CsrA-Strep protein was finally eluted using Buffer E (IBA GmbH, #2-1000-025) in three successive steps (E1 - 0.8 ml, E2 – 1.4 ml and E3 - 0.8 ml). The majority of CsrA-Strep was concentrated in the E2 fraction. Concentration was quantified using Roti®-Quant (Carl ROTH, #K015.3), and the protein was stored at -20 °C in 50 µl aliquots.

5.2.20. RIP-seq of *C. jejuni* CsrA-3xFLAG

Co-immunoprecipitation (coIP) combined with RNA-seq (RIP-seq) to identify direct RNA binding partners of CsrA-3xFLAG in *C. jejuni* was performed as previously described[101, 189] with minor modifications.

Co-immunoprecipitation of RNA with CsrA-3xFLAG: CoIP of chromosomally epitope-tagged *C. jejuni* CsrA with an anti-FLAG antibody and Protein A-Sepharose beads was performed from lysates of *C. jejuni* NCTC11168 and 81-176 WT (control) and isogenic *csrA*-3xFLAG strains grown in 100 ml (50 ml x 2 flasks) Brucella broth containing 10 µg/ml vancomycin to mid-exponential phase (OD_{600}=0.6) at 37 °C similarly as described before for *H. pylori*[101]. Cells were harvested by centrifugation at 6000 x g for 15 min at 4 °C. Afterwards, cell pellets were resuspended in 1 ml Buffer A (20 mM Tris-HCl, pH 8.0, 150 mM KCl, 1 mM $MgCl_2$, 1 mM DTT) and subsequently centrifuged (3 min, 11,000 x g, 4 °C). The pellets were shock-frozen in liquid nitrogen and stored at -80°C. Frozen pellets were thawed on ice and resuspended in 0.8 ml Buffer A. An equal volume of glass beads was then added to the cell suspension. Cells were then lysed using a Retsch MM40 ball mill (30 s^{-1}, 10 min) in pre-cooled blocks (4°C), and centrifuged for 2 min at 15,200 x g, 4 °C. The supernatant was transferred to a new tube, and additional 0.4 ml of Buffer A was added to the remaining un-lysed cells with beads. Lysis of the remaining cells was achieved by a second round of lysis at 30 s^{-1} for 5 min. Centrifugation was repeated and this second supernatant was combined with the first one. The combined supernatant was centrifuged again for 30 min at 15,200 x g, 4 °C for clarification and the resulting supernatant (lysate fraction) was transferred to a new tube. The lysate was incubated with 35 µl anti-FLAG antibody (Monoclonal ANTI-FLAG M2,

Sigma, #F 1804) for 30 min at 4 °C on a rocker. 75 µl of Protein A-Sepharose (Sigma #P6649), prewashed with Buffer A, was added and rocked for another 30 min at 4 °C. After centrifugation at 15,200 x g for 1 min, the supernatant was removed. Pelleted beads were washed 5 times with 0.5 ml Buffer A. Finally, 500 µl Buffer A was added to the beads and RNA and proteins were separated by phenol-chloroform-isoamyl alcohol extraction and precipitated as described previously[101]. From each coIP, 700-1000 ng of RNA was recovered. 100 µl of 1× protein loading buffer [62.5 mM Tris·HCl, pH 6.8, 100 mM DTT, 10% (v/v) glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue] was added to the final protein sample precipitated along with beads. This sample was termed as the coIP sample. For verification of a successful coIP, protein samples equivalent to 1.0 OD_{600} of cells were obtained during different stages of the coIP (culture, lysate, supernatant, wash, and coIP (beads)) for further Western blot analysis. One hundred microliters of 1× protein loading buffer was added to the protein samples and boiled for 8 min. Protein sample corresponding to an OD_{600} of 0.1 or 0.15 (culture, lysate, supernatant and wash fraction) and 10 or 5 (for proteins precipitated from beads) were used for Western blots as described below.

cDNA library preparation and processing of deep sequencing data: Residual gDNA was removed from the coIP RNA samples isolated from the control (WT) and CsrA-3xFLAG coIPs of the two strains *C. jejuni* NCTC11168 and 81-176 using DNase I treatment. cDNA libraries for Illumina sequencing were constructed by vertis Biotechnology AG, Germany (http://www.vertis-biotech.com/) in a strand-specific manner as described previously for eukaryotic microRNA [183] but omitting the RNA size-fractionation step prior to cDNA synthesis. In brief, equal amounts of RNA samples were poly(A)-tailed using poly(A) polymerase. Then, 5'-PPP was removed using tobacco acid pyrophosphatase (TAP), and an RNA adapter was then ligated to the resulting 5´-phosphate. First-strand cDNA was synthesized with an oligo(dT)-adapter primer and the M-MLV reverse transcriptase. In a PCR-based amplification step, using a high fidelity DNA polymerase, the cDNA concentration was increased to 20-30 ng/µL. For all libraries, the Agencourt AMPure XP kit (Beckman Coulter Genomics) was used to purify the DNA, which was subsequently analyzed by capillary electrophoresis.

A library-specific barcode for multiplex sequencing was included as part of a 3'-sequencing adapter. The following adapter sequences flank the cDNA inserts:

TrueSeq_Sense_primer

5’-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT-3’

TrueSeq_Antisense_NNNNNN_primer (NNNNNN = 6n barcode for multiplexing)

5’-CAAGCAGAAGACGGCATACGAGAT-NNNNNN-
GTGACTGGAGTTCAGACGTGTGCTCTTCCGATC(dT25)-3’

The samples were run on an Illumina HiSeq instrument with 100 cycles in single-read mode. The resulting read numbers are listed in Appendix Table 9.

Analysis of deep sequencing data: To assure high sequence quality, the Illumina reads in FASTQ format were trimmed with a cut-off phred score of 20 by the program fastq_quality_trimmer from FASTX toolkit version 0.0.13. After trimming, poly(A)-tail sequences were removed and a size filtering step was applied in which sequences shorter than 12 nt were eliminated. The collections of remaining reads were mapped to the *C. jejuni* NCTC11168 (NCBI Acc.-No: NC_002163.1) and 81–176 (NCBI Acc.-No: NC_008770.1, NC_008787.1, NC_008790.1) genomes using *segemehl* [184] with an accuracy cut-off of 95%. Mapping statistics are listed in Appendix Table 9. The raw, de-multiplexed reads as well as coverage files have been deposited in the NCBI Gene Expression Omnibus[190] under the accession number GSE58419. Coverage plots representing the numbers of mapped reads per nucleotide were generated. Reads that mapped to multiple locations contributed a fraction to the coverage value. For example, reads mapping to three positions contributed only 1/3 to the coverage values. Each graph was normalized to the number of reads that could be mapped from the respective library. To restore the original data range, each graph was then multiplied by the minimum number of mapped reads calculated over all libraries.

The overlap of sequenced cDNA reads to annotations was assessed for each library by counting all reads overlapping selected annotations on the sense strand. These annotations consist of strain-specific NCBI gene annotations complemented with annotations of previously determined 5’ UTRs and small RNAs [97]. Each read with a minimum overlap of 10 nt was counted with a value based on the number of locations where the read was mapped. If the read overlapped more than one annotation, the value was divided by the number of regions and counted separately for each region (e.g. 1/3 for a read mapped to 3 locations). To identify enriched transcripts in the CsrA-3xFLAG coIPs, differential cDNA count expression between CsrA-3xFLAG-tagged and control libraries was determined for each annotation using GFOLD version 1.0.9[191] but with manually defined normalization constants based on the number of reads that could be mapped to the respective libraries. For determination of genes enriched in the CsrA-3xFLAG-tagged library, $\log_2$ fold changes rather than GFOLD values were used.

5.2.21. Peak detection and motif analyses

In order to automatically define CsrA–bound RNA regions or peaks from the CsrA-3xFLAG coIP datasets, a simple in-house tool was developed based on a sliding window approach. In brief, the software uses normalized wiggle files of the CsrA-3xFLAG and control coIP libraries as input to determine sites showing a continuous enrichment of the CsrA-3xFLAG-tagged library compared to the control. The identification of enriched regions is based on four parameters: a minimum required fold-change (FC) for the enrichment, a factor multiplied by the 90th percentile of the wiggle graph which reflects the minimum required expression (MRE) in the tagged library, a window size in nt (WS) for which the previous two values are calculated in a sliding window approach, and a nucleotide step size (SS) which defines the steps in which the window is moved along the genomic axis. All consecutive windows that fulfill the enrichment requirements are assembled into a single peak region. The peak detection is performed separately for the forward and reverse strand of each replicon. For the CsrA-3xFLAG coIP dataset the following parameters were used: FC=5, MRE=3, WS=25 and SS=5. A detailed description of the tool will be described elsewhere (Bischler & Sharma unpublished).

For the prediction of a consensus motif based on sequences corresponding to peak regions using MEME [63], the following settings were applied: Search 0 to 1 motif of length 4 to 7 bp per sequence in the given strand only. To search for the presence of a structural motif, CMfinder 0.2.1 [108] was run on the enriched peak sequences with default parameters except for allowing a minimum single stem loop candidate length of 20 nt. The top-ranked motif incorporated 276 of the 328 sequences and was visualized by R2R [192].

5.2.22. Functional enrichment analysis

To check for overrepresentation of functional classes of CsrA-bound genes, genes with at least 5-fold enrichment in their 5' UTR and/or coding sequence in the CsrA-3xFLAG coIP library were considered (vs. control) as CsrA-bound and the remaining genes as unbound. An existing functional classification [102] of genes from strain NCTC11168 to determine statistically enriched functional classes was applied. Because a similar classification was not available for strain 81-176, a table with orthologue mappings between the two strains was downloaded from OrthologueDB [193] and used to assign the NCTC11168 functional classes to their respective 81-176 counterparts. Genes in the annotation lists without an existing functional classification in NCTC11168 or without an orthologue match were assigned to class 5.I, defined as "Unknown", in the original classification scheme. Only genes encoded on the pVir and pTet plasmids of strain 81-176 were assigned to new pVir and pTet classes, respectively. Functional overrepresentation was analyzed for each functional class via a two-sided Fisher's Exact Test followed by multiple testing

correction using the Benjamini-Hochberg method. An adjusted P-value of 0.05 was selected as significance threshold for functional overrepresentation.

5.2.23. Sequence and structure conservation of the *flaA* 5' UTR

In order to identify homologous *flaA* 5' UTR regions in different *Campylobacter* species and strains, a nucleotide blast (blastn [194]) with parameters optimized for more dissimilar sequences was performed (discontiguous megablast) using both the NCBI nucleotide collection (nr/nt) and the NCBI whole genome shotgun (wgs) contigs as databases. As query a 130 nt-long sequence encompassing the 100 nt upstream and the first 30 nt of the *C. jejuni* NCTC11168 *flaA* coding region was used. This includes the *flaA* promoter region, its 5' UTR, and the beginning of the coding sequence. Based on all hits (around 200) the 130 nt from the target sequences were extracted. If BLAST hits with an optimal score were truncated they were extended on either side to obtain 130 nt. All sequences with undefined bases or without proper species/strain association were excluded. Afterwards, sequences for *C. fetus* subsp. *fetus* 82-40 and *C. concisus* 13826, which were not found by BLAST, were added manually to the set. The sequences were aligned using MUSCLE[169] with default parameters and a conserved Sigma28 (FliA) -10 box (CGATAT) was observed in all of them. The alignment was trimmed to the region including only the *flaA* 5' UTR (based on the region according to the 5' UTR of *C. jejuni* NCTC11168) and the first 10 nt of the coding sequence. More dissimilar sequences disrupting the alignment (*C. peloridis* LMG 23910, *C. lari* NCTC11845, *C. lari* NCTC12892, *C. lari* RM16701, *C. lari* CCUG 22395, *C. lari* RM16712, *C. curvus* 525.92, *Campylobacter* sp. FOBRC14 ctg120009214739, *C. curvus* DSM 6644 C514DRAFT scaffold00004.4_C, and *C. concisus* 13826) were removed. Based on the resulting alignment all identical sequences were collapsed keeping only one representative sequence per cluster. Subsequently, a consensus structure was predicted using RNAalifold[195] with RIBOSUM scoring and default values for all other parameters. The resulting structure-annotated sequence alignment is shown in Fig. 27 and the full alignment including additional strains is shown in Appendix Fig. 11.

5.2.24. RNA isolation.

Bacteria were grown to the indicated growth phase and culture volume corresponding to a total amount of 4 OD_{600} was harvested and mixed with 0.2 volumes of stop-mix (95% ethanol and 5% phenol, vol/vol). The samples were snap-frozen in liquid nitrogen and stored at -80°C until RNA extraction. Frozen samples were thawed on ice and centrifuged at 4 °C to collect cell pellets. Cell pellets were lysed by resuspension in 600 µl of a solution containing 0.5 mg/ml lysozyme in TE

buffer (pH 8.0) and 60 µl 10% SDS. The samples were incubated for 1–2 minutes at 65 °C to ensure lysis. Afterwards, total RNA was extracted using the hot-phenol method as described previously [76, 196, 197].

5.2.25. Northern blot analysis

For Northern Blot analysis, 5 to 10 µg RNA sample was loaded per lane. After separation on 6% polyacrylamide (PAA) gels containing 7 M urea, RNA was transferred to Hybond-XL membranes (GE-Healthcare) by electroblotting. After blotting, the RNA was UV cross-linked to the membrane and hybridized with γ32P-ATP end-labeled DNA oligonucleotides (Appendix Table 12).

5.2.26. Rifampicin RNA stability assays.

To determine the stability of *flaA* mRNA in *C. jejuni* NCTC11168 WT, Δ*csrA*, Δ*fliW,* and Δ*csrA* Δ*fliW* strains, cells were grown to an OD_{600} of 0.45 (mid-log phase) and treated with rifampicin to a final concentration 500 µg/ml. Samples were harvested for RNA isolation at indicated time points following rifampicin addition (0, 4, 8, 16, and 32 min) as described above. After RNA isolation, 10 µg of each RNA sample was used for Northern Blot analysis as detailed above.

5.2.27. SDS-PAGE and immunoblotting.

Protein analyses were performed on cells collected from *C. jejuni* in mid-exponential phase (OD_{600} 0.5-0.6) or *E. coli* cultures in late-exponential phase (OD_{600} 1.0-1.5). Cells were collected by centrifugation at 11,000 x g for 3 min. Cell pellets were resuspended in 100 µl of 1× protein loading buffer [62.5 mM Tris·HCl pH 6.8, 100 mM DTT, 10% (v/v) glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue] and boiled for 8 min. For Western blot analysis, samples corresponding to an OD_{600} of 0.02 to 0.1 were separated by 12 % or 18 % (v/v) SDS-PAA gels and transferred to a PVDF membrane by semidry blotting. Membranes were blocked for 1 h with 10% (w/v) milk powder/TBS-T (Tris-buffered saline-Tween-20) and incubated overnight with primary antibody at 4 °C. Membrane was washed with TBS-T, followed by 1 h incubation with secondary antibody. After washing, the blot was developed using ECL-reagent. Proteins of interest were detected with monoclonal anti-GFP (1:1000 in 3% BSA/TBS-T) or anti-FLAG (1:1000 in 3% BSA/TBS-T; Sigma-Aldrich, #F1804-1MG) primary antibodies and secondary anti-mouse IgG (1:10,000 in 3% BSA/TBS-T; GE-Healthcare, #RPN4201). A monoclonal antibody specific for GroEL (1:10,000 in 3% BSA/TBS-T; Sigma-Aldrich, # G6532-5ML) and an anti-rabbit IgG (1:10,000 in 3% BSA/TBS-T; GE-Healthcare, #RPN4301) secondary antibody were used as a loading control.

5.2.28. Flow cytometric analysis.

FACS analysis of GFP reporter fluorescence in *E. coli* was performed as previously described[5]. Briefly, cells corresponding to 1 OD_{600} were collected from LB cultures in log phase and resuspended in 0.25 ml PBS. Cells were then fixed for 10 min with 0.25 ml 4% paraformaldehyde, collected by centrifugation, and washed twice with 0.5 ml PBS before final resuspension in 0.5 ml PBS. A 1/100 dilution of the fixed sample in PBS was used for measurement. Measurements (50,000 counts per sample) were performed on BD FACSCalibur machine and analyzed using FlowJo (V10).

5.2.29. Cytochrome *c* oxidoreductase activity assay.

The assay was adapted from Kappler *et al* [198]. The activity of the oxidoreductase enzyme was measured as the increase in cytochrome *c* absorbance at 550 nm when converted from the oxidized to the reduced form. Briefly, 2 µl cells (corresponding to 0.01 OD_{600nm} in 10 mM Tris/HCl pH 8) were added to a freshly prepared mix containing the following: 45 µl 10 mM Tris/HCl pH 8 (with or without 2 mM sodium sulphite), 5µl horse heart cytochrome *c* (10mg/ml; #C2506, Sigma-Aldrich). Absorbance was measured after 15 minutes at 550 nm using an absorption coefficient of 20 $mM^{-1}cm^{-1}$ (spectrophotometer *ND-1000; Peqlab*). In the control sample, 2 µl of 10 mM Tris/HCl pH 8 was added instead of cells.

5.2.30. Primer extension.

5 µg of DNase I digested RNA extracted from the exponential phase of *C. jejuni* NCTC11168 wild-type and Δ*rnc* strains was concentrated in 5.5 µl water. 1µl of γ^{32}P-ATP end-labeled oligodeoxyribonucleotide probe CSO-0223 (hybridizes to the 3' end of TracrRNA) or CSO-0270 (hybridizes to crRNA4 in NCTC11168) were mixed with the two RNA samples and incubated at 80°C. The temperature was gradually shifted to 42°C over a period of 1 hr. Each of the samples was then mixed with 1 µl 10 mM dNTPs, 2 µl 5xAMV RT buffer and 0.5 µl AMV RT 20 u/µl (Fermentas, #EP0641) and incubated further at 42°C for 1 hr. Samples were denatured for 3 min at 95 °C and half of the reaction was separated on a 6% polyacrylamide/8.3 M urea sequencing gel in 1X TBE. Sequencing ladders were generated according to the manufacturer's instructions by SequiTherm EXCEL™ II DNA Sequencing Kit (Epicentre Biotechnologies) using the same primers and PCR-amplified CRISPR locus from NCTC11168 (Oligos CSO-0253 and CSO-0258) as template. 2 µl of the sequencing reactions were loaded on the same gel as reference.

5.2.31. *In vitro* transcription, T7 RNA labeling, and gel mobility shift assays.

DNA templates containing the T7 promoter sequence were generated by PCR using oligos and DNA templates listed in Appendix Table 17. T7 transcription was carried out using the MEGAscript® T7 kit (Ambion) and sequences of the resulting T7 transcripts are listed in Appendix Table 17. *In-vitro* transcribed RNAs were quality checked and 5' end labeled (γ^{32}P) as previously described [188, 199]. Gel-shift assays were performed using ~0.04 pmol 5'-labeled RNA (4 nM final concentration) with increasing amounts of purified *C. jejuni* CsrA in 10 µl reactions. In brief, 5'-radiolabeled RNA (^{32}P, 0.04 pmol in 6 µl) was denatured (1 min, 95 °C) and cooled for 5 min on ice. Yeast tRNA (1 µg) and 1 µl 10x RNA Structure Buffer (Ambion: 10 mM Tris, pH 7, 100 mM KCl, 10 mM $MgCl_2$) was then added to the labeled RNA. CsrA protein (2 µl diluted in 1x Structure Buffer), was added to the desired final concentrations (0 mM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1µM or 2 µM CsrA). Binding reactions were incubated at 37 °C for 15 min. Prior to loading on a native 6 % polyacrylamide (PAA), 0.5X TBE gel, samples were mixed with 3 µl loading buffer (50% (v/v) glycerol, 0.5x TBE, 0.2% (w/v) bromophenol blue). Samples were run on native 6% (v/v) polyacrylamide gel in 0.5x TBE buffer at 300 V at 4 °C for 3 h. Gels were dried and analyzed using PhosphoImager (FLA-3000 Series, Fuji).

5.2.32. *In-vitro* structure probing assays.

In-vitro structure probing of *flaA* WT and *flaA* M1/M2 leader with RNase T1 and lead(II) acetate was performed as previously described [6]. 0.1 pmol of labeled leader variants was denatured for 1 min at 95 °C and chilled on ice for 5 min. 1 µg yeast tRNA as competitor and 10 x RNA Structure Buffer was added (provided together with RNase T1, Ambion). Unlabeled recombinant *C. jejuni* CsrA protein was then added at 0-, 20-, 50- or 100-fold molar excess. After incubation for 15 min at 37 °C, 2 µl RNase T1 (0.01 U/µl) or 2 µl freshly prepared lead(II)-acetate solution (25 mM) were added and reactions were incubated for 3 min or 90 sec, respectively. As a control, ~0.1 pmol labeled RNA with 100 fold excess CsrA was also prepared without nuclease/lead(II) treatment. The reactions were stopped by addition of 12 µl Gel loading buffer II (#AM8546G, Ambion). For RNase T1 ladders, ~0.1 pmol labeled RNA was denatured in 1x Structure Buffer for 1 min at 95 °C and afterwards incubated with 0.1 U/µl RNase T1 for 5 min. The OH ladder was generated by incubation of ~0.1 pmol labeled *flaA* WT leader RNA in 1x alkaline hydrolysis buffer (Ambion) for 5 min at 95 °C. Ladders and samples were then separated on 10% (v/v) PAA/7M urea gels in 1 x TBE buffer. Gels were dried and analyzed using a PhosphoImager (FLA-3000 Series, Fuji).

5.2.33. BACTH Assays.

The Bacterial Two Hybrid (BACTH) system was used to assess protein-protein interactions in *E. coli*. Proteins to be tested were fused to the T18 and T25 domains of the adenylate cyclase CyaA protein from *Bordetella pertussis*. The *csrA* (Cj1103) gene from *C. jejuni* NCTC11168 was amplified using CSO-0717/-0718 and introduced into pUT18, pUT18C and pKNT25 plasmids using restriction sites *Xba*I and *Bam*HI, resulting in plasmid pSSv68-1, pGD41-1 and pGD49-1 respectively. Similarly, the *fliW* gene (Cj1075) was amplified using CSO-0721/-0722 and introduced into pUT18, pUT18C and pKT25 plasmids using *Xba*I and *Bam*HI, resulting in plasmids pGD51-1, pSSv69-10 and pGD43-1 respectively. The two plasmids (T18- and T25- fusions) were then introduced into a Δ*cyaA* strain of *E. coli* (strain BTH101). Positive interaction between the proteins of interest reconstitutes CyaA activity, which can be monitored by expression of cAMP-CRP dependent genes such as *lacZ.* Relative β-galactosidase activities were visually analyzed on X-gal/ IPTG LB agar plates.

5.2.34. Protein-protein coIP.

The FliW and CsrA protein-protein coIP was performed exactly as described for the RIP-seq coIP protocol (see above) until the step where beads were washed 5 times washing with buffer A. After washing, the beads were suspended in 200 µl of 1× protein loading buffer [62.5 mM Tris-HCl, pH 6.8, 100 mM DTT, 10% (v/v) glycerol, 2% (w/v) SDS, 0.01% (w/v) bromophenol blue] and boiled for 8 min. Lysate samples corresponding to an OD_{600} of 0.05 and 2 (for proteins precipitated from beads) were used for Western blots as described before.

5.2.35. Transmission Electron Microscopy (TEM).

C. jejuni WT and mutant strains were grown for 14h on MH plates supplemented with vancomycin (10 µg/ml). Cells were resuspended gently in PBS using a cotton swab and centrifuged at 5000 x g for 5 min. The cell pellet was resuspended in 2% glutaraldehyde in 0.1 M cacodylate and incubated at 4 °C overnight. The next day, samples were stained with 2% uranyl acetate and inspected using a Zeiss EM10 transmission electron microscope.

5.2.36. *Campylobacter jejuni* motility assays.

C. jejuni strains were inoculated from the appropriate selective MH agar plates into 20 ml Brucella broth containing 10 µg/ml vancomycin and grown microaerobically with shaking at 37°C to an OD_{600} of approximately 0.5. Cells were harvested by centrifugation at 6500 x g for 5 min and resuspended at an OD_{600} of 0.5 in Brucella broth. For each strain, 0.5 µl of bacterial suspension was inoculated into motility soft-agar plates (MH broth + 0.4% agar) poured the day before. Plates were incubated right-side-up for approximately 24 h microaerobically at 37°C. Three

measurements of each motility halo were made for each inoculation, which were averaged to give the mean swim distance for each strain on a plate. All strains were inoculated together on six replicate plates and the mean swim distance +/- standard error on these plates was used to the compare motility of each strain.

5.2.37. *Campylobacter jejuni* autoagglutination assay.

Autoagglutination was determined as described previously[112]. Briefly, strains grown in liquid cultures for motility assays were resuspended in phosphate-buffered saline (PBS), pH 7.4, to an OD_{600} of 1.0. Two milliliters were placed into three replicate tubes and the OD_{600} was measured. Tubes were incubated at 37°C microaerobically without shaking, and at indicated time points, 100 µl was carefully removed from the top of the suspension, diluted 10-fold in PBS, and the OD_{600} was measured. Measurements were normalized to the optical density of each strain at the zero time point.

5.2.38. Time-lapse microscopy to monitor cell division.

C. jejuni Δ*fliA* mutant cells corresponding to an OD_{600} of 0.5 were collected from Brucella broth culture in log phase by centrifugation and resuspended in 0.5 ml Brucella broth. The cells were further serially diluted 100- and 1000-fold in Brucella broth. 5 µl of the diluted samples were spotted on a Brucella broth-agarose (1%) plate. The plate was incubated under microaerophilic conditions at 37 °C for 10 minutes. The agarose patch was excised and inverted onto a Petri dish with a glass bottom. Single cells were then monitored over time using several bright-field images in a fluorescence microscope (Leica DMI6000 B) maintained at 37 °C under aerobic conditions.

5.2.39. RNA Fluorescence *In Situ* Hybridization (FISH).

RNA FISH was performed as previously described [200] with some modifications. Cells corresponding to 2 OD_{600} were collected from Brucella broth cultures in log phase and resuspended in 0.5 ml PBS. Cells were then fixed for 3 hours with 0.5 ml 4% paraformaldehyde at room temperature, collected by centrifugation, and washed twice with 0.5 ml PBS before final resuspension in 0.5 ml 70% ethanol. After 10 min cells were collected by centrifugation and resuspended in 95% ethanol and incubated at room temperature for 1 h. Cells were collected by centrifugation, completely dried in a laminar hood, and then washed once with 2 x SSC before final resuspension in 0.5 ml 2 x SSC containing 10% formamide. Fluorescently labeled oligos (Cy5 or FITC, Sigma, Appendix Table 12) were then added at a concentration of 10 ng/µl and incubated at 37°C overnight. The next day, cells were collected by centrifugation and washed three times for

1 h at 37 °C using 0.5 ml 2x SSC containing 10% formamide before final resuspension in 2x SSC (50-250 µl). Cells were then imaged in a Leica Confocal TCS SP5 II microscope using sequential scanning mode.

5.2.40. *d*STORM super-resolution microscopy

*d*STORM was performed on a wide field setup for localization microscopy[201]. An optically pumped semiconductor laser (Genesis MX STM-Series, Coherent) with a wavelength of 639 nm (maximum power of 1 W) was used for excitation of Cy5 and a diode laser (iBeam smart Family, TOPTICA Photonics) with a wavelength of 405 nm (maximum power of 120 mW) was used for reactivation of Cy5. Laser beams were cleaned-up by bandpass filters (Semrock/Chroma) and combined by appropriate dichroic mirrors (LaserMUX™ filters, Semrock). Afterwards they were focused onto the back focal plane of the high numerical oil-immersion objective (Olympus APON 60XO TIRF, NA 1.49), which is part of an inverted fluorescence microscope (Olympus IX71). To separate the excitation light from the fluorescence light, suitable dichroic beam splitters (Semrock) were placed into the light path before the laser beams enter the objective. Fluorescence light collected by the objective was filtered by appropriate detection filters (Semrock/Chroma) and was detected by an EMCCD camera with 512 x 512 pixels (iXon Ultra 897, Andor Technology). The pixel size in the image was 129 nm/px. Cy5 was excited with the 639 nm laser at a maximum intensity of 4.19 kW/cm^2. During imaging, the 405 nm laser was switched on to keep up a suitable switching ratio. Its laser power was increased successively to a maximum intensity of 0.04 kW/cm^2. For every image 5,000 - 25,000 frames were taken with an integration time of 15 ms/frame. For every imaged area, additionally a bright field image was taken to identify single bacteria. Data analysis was performed using rapi*d*STORM open source software[202].

5.2.41. Statistical analysis.

All data for Western, Northern blot, or FISH analysis are presented as mean ± s.e.m. Statistical analysis was carried out using the online portal www.studentsttest.com. For statistical comparison of two groups, a two-tailed paired Student's *t*-test was used. A value of $P<0.05$ was considered significant and marked with an asterisk (*) as explained in the legends. For FISH analysis, fluorescence data curves from 10 cells from a single image were merged as a single averaged curve after cell length normalization (Appendix Fig. 14). The data was acquired and normalized over cell length using ImageJ and subsequently the merged average curve was generated using Microsoft Excel.

6. Appendix

Appendix Table 1. SuperGenome TSS table. The table contains information on positions and assigned classes of all annotated TSS. It lists all TSS that were detected in the SuperGenome (column "detected" = 1). In case a TSS was not detected in a certain strain the value of detected is "0". If a TSS was mapped in more than one genome via the SuperGenome, there is one row for each genome the TSS was mapped to. Also, if the TSS is assigned to more than one class, there is one row for each class assignment and each associated gene. This table is provided as a separate Excel sheet on a DVD and can also be accessed using the link – (http://journals.plos.org/plosgenetics/article/file?type=supplementary&id=info:doi/10.1371/journal.pgen.1003495.s003)

Appendix Table 2. TSS with previously described FliA and RpoN dependent promoters. This table lists genes with previously described FliA- [116] or RpoN-dependent promoters [203]. Conserved nucleotides, which are characteristic for each promoter type, are underlined.

Gene	**Name**	***-50 to +1 of TSS***	***TSS detected***
FliA-dependent genes TTT 10…12 CGAT [AT/TT/TA]			
Cj0547	*flaG*	TTTTCTTTATAAGTGTTCATAAATTTATAAATTTGTCGATATAAGCTTTTA	Yes
Cj1339c	*flaA*	ATTTTATTGCTAAAGTATAAAATATTTTTTGATTGCACGATATAGCATTT	Yes
Cj0045c	Putative iron-binding protein	ATTCATTTGTCAAATTATTATTTTATTTTCAAAAAATACGATATAAAAATA	Yes
Cj0391c	Hypothetical protein	GAATTTAATGTCCTTATAAATTTTTTAAGTCAAAAGTCGATATAAATTAAA	Yes
Cj0977	Hypothetical protein	TATACGCATTAAAATATAAATTCTTTTTATTTTTTGCCGATATTGAATTTT	Yes
Cj1464	*flgM*	TTTAAAAGAAAGGGGTTAAGTTTTTTAAATTTAGGTCGATATGGTTTGTAG	No [#]
Cj1656c	Hypothetical protein	AAGTTTAGTAAAAAATATTAAGTTTTAAAAATATCATCCGATTTAAATAAA	Yes
Cj0859c	Hypothetical protein	AATTATTTTAAATTATAGTATATTTTTTTAAAATAATGTCGATTTAGTGTA	Yes
Cj1034c	Adenylosuccin-ate lyase	TATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCT	No [##]
RpoN-dependent genes TGGNACANNNNNTGCTT			
Cj1338c	*flaB*	AAATTTTTAATTAAACTAAAACTTGGAACACTTTTTGCTTTAATCTTTTCG	Yes
Cj1462	*flgI*	TTTTTACTCTTCTTAACTCACTTGGAACACTTTTTGCTTGATCTATCATCA	Yes
Cj0528c	*flgB*	ATATTTTTATACAAAATAGTTAAATTGGAACAGTATTTGCTTGTTAATATT	Yes
Cj0887c	*flgL*	ATAAAATTATATAATTTAAATTTTTGGAACAGTTATTGCTTTTGTTTATTA	Yes
Cj0042	*flgD*	TTGTTTTAGCGAAGTATTTTTAAATTGGAACACTTATTGCTTAAATAATAA	No [#]
Cj0697	*flgG2*	TTATTTTTCTTTGTTTTTAAAAGTTGGAACACTCTTTGCTTTTATAGTTAT	No [#]
Cj0687c	*flgH*	TTTTCTCCTTGAATTTATAAATGTGGAACACTCTTTGCTTTTTCTAATTTT	No [#]
Cj1466	*flgK*	GTAAATAAAATGTTTGATATCAATGATGGAACAAATAATGCTTATGGAGAT	No [#]

[#] Reads were detected in the 5' UTR region, but no enrichment was observed in the +TEX library.
[##] Almost equal reads in untreated and TEX treated samples.

Appendix Table 3. TSS with a predicted FliA or RpoN motif.
This table lists all TSS with a predicted FliA motif. The columns "MEME" and "Regular expression" indicate by which method the motif was detected.

Strain	Locus	Position in strain	Strand	TSS type	MEME	Regular expression	Sequence
FliA motif							
81-176	CJJ81176_0004	1596607	-	internal	Yes	No	AAGTTTTTGTAGATAATCAAATCAACAAACAAATCGCCGATGATG
81-176	CJJ81176_0049	27632	-	internal	Yes	No	TTATAGTTGATAAAGCTATAAATTTAACTAAAAACTACGATGATG
81-176	CJJ81176_0083	77142	-	primary	No	Yes	ATTCATTTGTAAAATTATTATTTTATTTTCAAAAAATACGATATAAAAATA
81-176	CJJ81176_0179	157328	+	antisense	No	Yes	AAAATATTTCCTTCGATATAAATTTTTTAAATTATAACGATATTAAGCTTA
81-176	CJJ81176_0180	157328	+	secondary	No	Yes	AAAATATTTCCTTCGATATAAATTTTTTAAATTATAACGATATTAAGCTTA
81-176	CJJ81176_0228	198999	-	internal	Yes	No	ATTTTACATTTGATGCAAAGTGCTAGTGAAGTGATGCCGATTTTT
81-176	CJJ81176_0268	231630	+	antisense	Yes	No	AAGGATTAGATAAGGTTAAATTTTGATATGGATTTGCCGAATTTA
81-176	CJJ81176_0291	247875	-	internal	Yes	No	ATTCCAAAGATGCAAAAAAACTTGGTATAAAAAATGGCGATGTAG
81-176	CJJ81176_0295	252444	+	antisense	No	Yes	TAGCAAAGGGGTGTTTAGGCTGTTTGCGATTTCTTTGTTAGAATAAGCTTG
81-176	CJJ81176_0316	274334	-	primary	No	Yes	TTTCGTATTATTTTTAATTATTTTTTTATTTTATATCGATTAACATTATTT
81-176	CJJ81176_0329	283484	+	internal	Yes	Yes	AAAAGAAGTTATAATGCAAAATGAAATTTTAAAAAAACTCGATTTAGAGCA
81-176	CJJ81176_0346	299919	+	internal	Yes	No	GCATGCTAAATGTTGCAAAAGAAAAATTTCCAAATGTCGAATTTA
81-176	CJJ81176_0347	300874	+	internal	Yes	No	ATGCTCCTTTTTCTCTTATACAAGCTTTAAAAAAAGCCGATGATA
81-176	CJJ81176_0397	345801	+	internal	No	Yes	TTGTACATTTGGAATTAAAAGATTTATTAAAAACTTGCGATATTATAAGTA
81-176	CJJ81176_0414	361740	-	primary	No	Yes	GAATTTAATGTCCTTATAAATTTTTTAAGCCAAAAGTCGATATAAATTAAA
81-176	CJJ81176_0419	366797	-	primary	No	Yes	ATCTTTAGGGGATAAATCGATATTTTTAAAAAATTATGATAAATTATTAGA
81-176	CJJ81176_0420	366797	-	internal	No	Yes	ATCTTTAGGGGATAAATCGATATTTTTAAAAAATTATGATAAATTATTAGA
81-176	CJJ81176_0440	386757	+	primary	No	Yes	TGAAAACTCTCAAGATTATAAATTTTTAATTTTCTTCCGATATAGTATATA
81-176	CJJ81176_0466	410949	+	antisense	No	Yes	CTTGTAAGGATGATCTTTTAAATTTTCATTTTTTAAACGATTTATAATTTC
81-176	CJJ81176_0481	423800	-	primary	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
81-176	CJJ81176_0482	423800	-	internal	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
81-176	CJJ81176_0537	473639	+	antisense	No	Yes	AACGGCGTTGCAAGGCTGCATCTTTTTCAAAATATTTACGATATTCTTTTA
81-176	CJJ81176_0550	485343	+	antisense	No	Yes	CTTTTTTCGCCAAATTTAAATATTTTAAGTCCTGATACGATTTCTATATTG
81-176	CJJ81176_0564	499555	+	primary	No	Yes	TAATTATTTTAATTTAGCTACTTTTAACGATTAAAATTTTATAATCAAAAA
81-176	CJJ81176_0572	507505	+	primary	No	Yes	TTTTCTTTATAGGTGTTCATAAATTTATAAATTTGTCGATATAAGCTTTTA
81-176	CJJ81176_0600	530840	-	antisense	No	Yes	TTTTCTTCAAGCTTAATAAGGCTTTCATGTTTTAGGCGATATTCTATAAGA
81-176	CJJ81176_0697	622236	-	antisense	Yes	No	AGTGCTATCATACCTGATTTAAAAATTTCATTAGAGCCGATTTTC
81-176	CJJ81176_0697	622587	-	antisense	Yes	No	CCTGGTAAAATAAAACTATGATTGATAACAAATTTACCGATGTGA
81-176	CJJ81176_0730	659685	+	internal	Yes	No	TTTTTTAAAATCCAAATACAAAACTAGTTTAAAATCCCGATTTTT

81-176	CJJ81176_0736	664084	+	internal	Yes	No	ATCCAAGAGATTTTTCTAAAAATTCTTATCATAAAGTCGATGATT
81-176	CJJ81176_0755	682782	+	internal	Yes	No	TAAGCCCTAAAGAGCTTGATAGCTTGGGTTTAAAACACGATTTTA
81-176	CJJ81176_0761	688542	+	internal	Yes	No	ATAGTAATAAGAAGGCTAAAATTTATAATCTTAATGTCGATGGTT
81-176	CJJ81176_0791	720953	+	antisense	No	Yes	TTTGCATCTAAATCTTTTTCATTTAAAGCGATATTAGGTATATTATAGTCA
81-176	CJJ81176_0816	746454	-	internal	Yes	Yes	TCTTATTCATTCAAGTCTTGAAGGTTTAGAATTTAAAGTCGATATAGAAAA
81-176	CJJ81176_0867	798711	-	internal	Yes	No	ATTTATCAAAAATTATTCATTTTAAAGTTCTAAAAGACGATTTAA
81-176	CJJ81176_0875	805689	-	primary	No	Yes	ATTATTTTAAATTATAGTATATTTTTTTAAAATAATGTCGATTTAGTGTAA
81-176	CJJ81176_0876	805689	-	antisense	Yes	Yes	ATTATTTTAAATTATAGTATATTTTTTTAAAATAATGTCGATTTAGTGTAA
81-176	CJJ81176_0911	842070	+	internal	Yes	No	AAGTTTTAAAAAATATTAACACTACAATCTCAAAAGGCGATGTAA
81-176	CJJ81176_0996	918305	+	primary	No	Yes	TATACGCATTAAAATATAAATTCTTTTTATTTTTTGCCGATATTGAAATTT
81-176	CJJ81176_1025	944670	+	antisense	Yes	No	ATAGTTATTAGGAATAAATATAATACGCCCACTACGACGATTTTT
81-176	CJJ81176_1046	964676	-	internal	Yes	No	GTGGGGTGACGGCTATTAAGTTTAAAGAGAAAAATGACGAATTAG
81-176	CJJ81176_1053	976068	-	primary	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCCA
81-176	CJJ81176_1054	976068	-	internal	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCCA
81-176	CJJ81176_1072	990237	-	internal	Yes	No	CTAAGCTTGAAAATGTTGAGAAAAAAGAGCAAAAAGTCGATGAAA
81-176	CJJ81176_1117	1036918	-	antisense	Yes	No	AAATTAATTTGTCTATAAAAAGTTGCAAAAATATCTTCGATTTTA
81-176	CJJ81176_1119	1038647	+	internal	No	Yes	GAAAATATTTTAAATTTTCAAAACGATTTTAAAAATGTTAAACTTGTAAAA
81-176	CJJ81176_1144	1062656	-	internal	Yes	Yes	TAAATCCAAAAAAATGTGAAATTTTATACAATGCCATCGATTTAGAAAACA
81-176	CJJ81176_1245	1161252	+	internal	No	Yes	AACCGCAGTGGTGCTTAAAATTTTAAATGATAATGGCGATATCAATGAGCA
81-176	CJJ81176_1336	1249396	+	primary	Yes	No	AAATCATAAAAGAAACTTTAAAAAAATACAAAAATGACGATATTT
81-176	CJJ81176_1339	1256974	-	primary	No	Yes	TTTTTAGTTTAAGGTATAAATTTTTAAAAAAAAGTAACGATATAGCATTTA
81-176	CJJ81176_1351	1269043	+	internal	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTTCAAGA
81-176	CJJ81176_1352	1269043	+	secondary	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTTCAAGA
81-176	CJJ81176_1370	1290494	-	antisense	Yes	No	TTTTTCTATCACTTCTTCATAACTCATTTTAAAACGACGATGTAA
81-176	CJJ81176_1442	1366331	+	antisense	Yes	Yes	CCATATTATAAAAATTATACTAAAATATTTTTAAAAATAGCCGATTTTTGA
81-176	CJJ81176_1442	1366335	+	antisense	Yes	Yes	ATTATAAAAATTATACTAAAATATTTTTAAAAATAGCCGATTTTTGATTTA
81-176	CJJ81176_1443	1366331	+	secondary	No	Yes	CCATATTATAAAAATTATACTAAAATATTTTTAAAAATAGCCGATTTTTGA
81-176	CJJ81176_1443	1366335	+	primary	No	Yes	ATTATAAAAATTATACTAAAATATTTTTAAAAATAGCCGATTTTTGATTTA
81-176	CJJ81176_1535	1457962	-	internal	Yes	No	AAAAAAGGTAGAGTATTTCAAGAAATTAACACAAAACCGATTTAT
81-176	CJJ81176_1539	1462504	-	internal	Yes	No	GCTCATTTTTAAACGACGCTCATAAAGACTTAAAAGCCGATTTTA
81-176	CJJ81176_1623	1535148	-	secondary	No	Yes	TCCTTTCTGAGATTATAAAGTTTTTTAAGATTTTGTCGATTATAAAAGGTA
81-176	CJJ81176_1624	1535148	-	antisense	No	Yes	TCCTTTCTGAGATTATAAAGTTTTTTAAGATTTTGTCGATTATAAAAGGTA
81-176	CJJ81176_1625	1536225	+	antisense	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGTAAAACAAGCGATATAGTATTTG
81-176	CJJ81176_1626	1537753	-	internal	Yes	No	AGTCTTTTGCAAAAATTGCAAATAGTTTAAATTTGGGCGATTTTA

81-176	CJJ81176_1647	1556543	-	primary	Yes	Yes	GTTTTAGTAAAAAATATTAAGTTTTGAAAATATCATCCGATTTAAATAAAG
81-176	CJJ81176_1648	1556793	+	primary	No	Yes	TTTTACTTCAATAAATTTAAGTTTTATCTTATTTATACGATTTATAATATA
81-176	CjNC4	1536225	+	primary	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGTAAAACAAGCGATATAGTATTTG
81116	C8J_0137	152376	+	antisense	No	Yes	AATATTTCCTTCGATATAAATTTTTTCCAAATTATAACGATATTAAGCTTA
81116	C8J_0138	152376	+	secondary	No	Yes	AATATTTCCTTCGATATAAATTTTTTCCAAATTATAACGATATTAAGCTTA
81116	C8J_0186	196020	-	internal	Yes	No	ATTTTACATTTGATGCAAAGTGCTAGTGAAGTGATGCCGATTTTT
81116	C8J_0197	208056	-	antisense	No	Yes	TTCTTCAGGTTTTTTATAATTTGCGATTAAATTTAGCATAAAATAATTAGA
81116	C8J_0241	245067	-	internal	Yes	No	ATTCCAAAGATGCAAAAAAACTTGGTATAAAAAATGGCGATGTAG
81116	C8J_0269	278980	-	primary	No	Yes	TTTCGTATTATTTTTAATTATTTTTTTATTTTATATCGATTAATATTATTT
81116	C8J_0301	305952	+	internal	Yes	No	GCATGCTAAATGTTGCAAAAGAAAAATTTCCGAATGTCGAATTTA
81116	C8J_0302	306907	+	internal	Yes	No	ATGCTCCTTTTTCTCTTATACAAGCTTTAAAAAAAGCCGATGATA
81116	C8J_0366	367776	-	primary	No	Yes	GAATTTAATGTCCTTATAAATTTTTTAAGTCAAAAGTCGATATAAATTAAA
81116	C8J_0371	372833	-	primary	No	Yes	ATCTTTAGGGGATAAATCGATATTTTTAAAAAATTATGATAAATTATTAGA
81116	C8J_0372	372833	-	internal	No	Yes	ATCTTTAGGGGATAAATCGATATTTTTAAAAAATTATGATAAATTATTAGA
81116	C8J_0372	372901	+	antisense	No	Yes	TTTTTTAAAAATATCGATTTATCCCCTAAAGATAATGCAATAATATTTTTA
81116	C8J_0391	392557	+	primary	No	Yes	TGAAAACTCTCAAGATTATAAATTTTTAATTTTCTTCCGATATAGTATATA
81116	C8J_0429	431442	-	primary	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
81116	C8J_0430	431442	-	internal	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
81116	C8J_0473	481491	+	antisense	No	Yes	AACGGCGTTGCAAGGCTGCATCTTTTTCAAAATATTTACGATATTCTTTTA
81116	C8J_0500	507407	+	primary	No	Yes	TAATTATTTTAATTTAGCTACTTTTAACGATTAAAATTTTATAATCAAAAA
81116	C8J_0508	515357	+	primary	No	Yes	TTTTCTTTATAGGTGTTCATAAATTTATAAATTTGTCGATATAAGCTTTTA
81116	C8J_0626	629247	-	antisense	Yes	No	CCTGGTAAAATAAAACTATGATTGATAACAAATTTACCGATGTGA
81116	C8J_0646	648285	+	internal	No	Yes	GAAGAATTTGAAAGAATGGTCGATTTACATATAAATGATAAAATATCTATA
81116	C8J_0647	648285	+	primary	No	Yes	GAAGAATTTGAAAGAATGGTCGATTTACATATAAATGATAAAATATCTATA
81116	C8J_0674	673458	+	internal	Yes	No	TTTTTTAAAATCCAAATACAAAACTAGTTTAAAATCCCGATTTTT
81116	C8J_0680	677857	+	internal	Yes	No	ATCCAAGAGATTTTTCTAAAAATTCCTATCATAAAGTCGATGATT
81116	C8J_0699	696543	+	internal	Yes	No	TAAGTCCTAAAGAGCTTGATAACTTGGGTTTAAAACACGATTTTA
81116	C8J_0704	700058	-	antisense	No	Yes	TTTTCAAAATTTACTTCATTATTGTTTGATTCATCGGCGATATAGTCTATA
81116	C8J_0705	702257	+	internal	Yes	No	ATGGTAACAATAAGGCTAAAATTTATAATCTTAATGTCGATGGTT
81116	C8J_0710	717110	+	internal	Yes	No	AGCCAAATACAGCTAATGATAAAAAGAAAAAAGATGACGATGTAA
81116	C8J_0721	726993	+	antisense	No	Yes	TTTGCATCTAAATCTTTTTCATTTAAAGCGATATTAGGTATATTATAGTCA
81116	C8J_0729	737429	+	internal	Yes	No	AAAAAGGTTTAGAAAATAAAAAAGCTTTAGAATTTGTCGATTTTA
81116	C8J_0743	748795	+	internal	Yes	No	AAGGTTTTGAAAAATTAAACATTTTAAATTTTGAAGACGATGTAT
81116	C8J_0746	752531	-	internal	Yes	Yes	TCTTATTCATTCAAGTCTTGAAGGTTTAGAATTTAAAATCGATATAGAAAA

81116	C8J_0783	790487	+	internal	Yes	Yes	TATCTTTACATAGTGATGGAAAATTTTTAACTAGAGTCGATTTTTGTGATA
81116	C8J_0798	804789	-	internal	Yes	No	ATTTATCAAAAATTATTCATTTTAAAGTTCTAAAAGACGATTTAA
81116	C8J_0806	811767	-	primary	No	Yes	ATTATTTTAAATTATAGTATATTTTTTTTAAATAATGTCGATTTAGTGTAA
81116	C8J_0807	811767	-	antisense	Yes	Yes	ATTATTTTAAATTATAGTATATTTTTTTTAAATAATGTCGATTTAGTGTAA
81116	C8J_0839	848140	+	internal	Yes	No	AAGTTTTAAAAAATATTAACACTACAATCTCAAAAGGCGATGTAA
81116	C8J_0937	945227	-	antisense	No	Yes	AATTATAATTTTTTCTTTTTAAGGTATAAAATTTACGCTTTAGGTCGATAT
81116	C8J_0938	945342	+	primary	Yes	No	AAATTTGAAAAAATATTCAAATCATAAGCATAATAACCGATTTAT
81116	C8J_0944	952209	+	antisense	Yes	No	ATAGTTATTAGGAATAAATATAATACGCCCACTACGACGATTTTT
81116	C8J_0964	972213	-	internal	Yes	No	GTGGGGTGACGGCTATTAAGTTTAAAGAGAAAAATGACGAATTAG
81116	C8J_0971	983606	-	primary	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCCA
81116	C8J_0972	983606	-	internal	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCCA
81116	C8J_1040	1051307	-	antisense	Yes	No	AAATTAATTTGTCTATAAAAAGTTGCAAAAATATCTTCGATTTTA
81116	C8J_1042	1053036	+	internal	No	Yes	GAAAATATTTTAAATTTTCAAAACGATTTTAAAAATGTTAAACTTGTAAAA
81116	C8J_1174	1180429	+	internal	No	Yes	AACCGCAGTGGTGCTTAAAATTTTAAATGATAATGGCGATATAAATGAGCA
81116	C8J_1237	1247282	+	internal	Yes	No	GTGCTTTGTATAAAAGTGAAAAAGTAGGAAAAAAAGCCGATCTTT
81116	C8J_1238	1249083	+	internal	Yes	No	CTATAAATAAAGGAGTTAAACAAAAACCTATAAGTGCCGATTTTC
81116	C8J_1253	1261818	+	primary	Yes	No	AAATCATAAAAGAAACTTTAAAAAAATACAAAAATGACGATATTT
81116	C8J_1256	1269396	-	primary	No	Yes	TTTTTAGTTTAAGGTATAAATTTTTAAAAAAAAGTAACGATATAGCATTTA
81116	C8J_1268	1281457	+	internal	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTTCAAGA
81116	C8J_1269	1281457	+	secondary	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTTCAAGA
81116	C8J_1355	1373838	+	antisense	Yes	Yes	CCATATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAA
81116	C8J_1355	1373842	+	antisense	Yes	Yes	ATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAATTTA
81116	C8J_1356	1373838	+	secondary	No	Yes	CCATATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAA
81116	C8J_1356	1373842	+	primary	No	Yes	ATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAATTTA
81116	C8J_1534	1546677	-	secondary	No	Yes	TCCTTTCTGAGATTATAAAGTTTTTTAAGATTTTGTCGATTATAAAAGGTA
81116	C8J_1535	1546677	-	antisense	No	Yes	TCCTTTCTGAGATTATAAAGTTTTTTAAGATTTTGTCGATTATAAAAGGTA
81116	C8J_1536	1547754	+	antisense	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGCAAAACAAGCGATATAGTATTTG
81116	C8J_1537	1549282	-	internal	Yes	No	AGTCTTTTGCAAAAATTGCAAATAGTTTAAATTTGGGCGATTTTA
81116	C8J_1543	1554909	+	internal	Yes	No	AAGCTATAAAAAAAGCTTTAGAACTAAAAGAAAATGACGATTTAG
81116	C8J_1558	1568072	-	primary	Yes	Yes	AGTTTAGTAAAAAATATTAAGTTTTAAAAATATCATCCGATTTAAATAAAG
81116	C8J_1559	1568321	+	primary	No	Yes	TTTTACTTCAATAAATTTAAATTTTATCTTATTTATACGATTTATAATATA
81116	C8J_1614	1610161	-	internal	Yes	No	AAGTTTTTGTGGATAATCAAATCAACAAACAAATTGCCGATGATG
81116	C8J_1629	1626223	-	internal	Yes	No	ATAATGCCGATGGTACAACAGATAATATGAAAAATATCGATTTAA
81116	CjNC4	1547754	+	primary	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGCAAAACAAGCGATATAGTATTTG

NCTC11168	Cj0022c	27767	-	internal	Yes	No	TTATAGTTGATAAAGCTATAAATTTAACTAAAAACTACGATGATG
NCTC11168	Cj0045c	66492	-	primary	No	Yes	ATTCATTTGTCAAATTATTATTTTATTTTCAAAAAATACGATATAAAAATA
NCTC11168	Cj0122	124762	+	primary	Yes	No	ATATATGAGTAAAAATATTAAAAAAAATCATTTTTGCCGATTTGC
NCTC11168	Cj0143c	146556	+	antisense	No	Yes	AAATATTTCCTTTCGATATAAATTTTTTAAATTATAACGATATTAAGCTTA
NCTC11168	Cj0144	146556	+	secondary	No	Yes	AAATATTTCCTTTCGATATAAATTTTTTAAATTATAACGATATTAAGCTTA
NCTC11168	Cj0197c	193327	-	internal	Yes	No	ATTTTACATTTGATGCAAAGTGCTAGTGAAGTGATGCCGATTTTT
NCTC11168	Cj0223	208840	+	internal	No	Yes	GGTTATAGAAAAGCTTATGATAAATTTTATCTTGAGCCGATTTGGAATTTA
NCTC11168	Cj0264c	242190	-	internal	Yes	No	ATTCTAAAGATGCAAAAAAACTTGGTATAAAAAATGGCGATGTAG
NCTC11168	Cj0268c	246758	+	antisense	No	Yes	TAGCAAAGGGGTGTTTAGGCTGTTTGCGATTTCTTTGTTAGAATAAGCTTG
NCTC11168	Cj0292c	268924	-	primary	No	Yes	TTTCGTATTATTTTTAATTATTTTTTTATTTTATATCGATTAACATTATTT
NCTC11168	Cj0324	295898	+	internal	Yes	No	GCATGCTAAATGTTGCAAAAGAAAAATTTCCAAATGTCGAATTTA
NCTC11168	Cj0325	296853	+	internal	Yes	No	ATGCTCCTTTTTCTCTTATACAAGCTTTAAAAAAAGCCGATGATA
NCTC11168	Cj0373	341793	+	internal	No	Yes	TTGTACATTTGGAATTAAAAGATTTATTAAAAACTTGCGATATTATAAGTA
NCTC11168	Cj0391c	358674	-	primary	No	Yes	GAATTTAATGTCCTTATAAATTTTTTAAGTCAAAAGTCGATATAAATTAAA
NCTC11168	Cj0416	383457	+	primary	No	Yes	TGAAAACTCTCAAGATTATAAATTTTTAATTTTCTTCCGATATAGTATATA
NCTC11168	Cj0440c	409254	+	antisense	No	Yes	CTTGTAAGGATGATCTTTTAAATTTTCATTTTTTAAACGATTTATAATTTC
NCTC11168	Cj0456c	422128	-	primary	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
NCTC11168	Cj0457c	422128	-	internal	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
NCTC11168	Cj0509c	476292	+	antisense	No	Yes	GGCGTTGCAAGGCTGCATCTTTTTCAAAATATTTACGATATTCTTTTAAAG
NCTC11168	Cj0525c	487985	+	antisense	No	Yes	CTTTTTTCGCCAAATTTAAATATTTTAAGTCCTGATACGATTTCTATATTG
NCTC11168	Cj0539	502198	+	primary	No	Yes	TAATTATTTTAATTTAGCTACTTTTAACGATTAAAATTTTATAATCAAAAA
NCTC11168	Cj0547	510149	+	primary	No	Yes	TTTTCTTTATAAGTGTTCATAAATTTATAAATTTGTCGATATAAGCTTTTA
NCTC11168	Cj0572	533951	-	antisense	No	Yes	TTTTCTTCAAGCTTAATAAGGCTTTCATGTTTTAGGCGATATTCTATAAGA
NCTC11168	Cj0638c	599888	+	antisense	No	Yes	ATTTCGATTACGGCATTGATTTTATTTGGAATATCTCCGATTTTGATTTTG
NCTC11168	Cj0639c	599888	+	antisense	Yes	Yes	ATTTCGATTACGGCATTGATTTTATTTGGAATATCTCCGATTTTGATTTTG
NCTC11168	Cj0641	602533	+	internal	Yes	Yes	TGCCAAAGTATGGATTGGATGATTTGTTTAAAATATCCGATTTTGTTATTT
NCTC11168	Cj0671	626508	-	antisense	Yes	No	AGTGCTATCATACCTGATTTAAAAATTTCATTAGAGCCGATTTTC
NCTC11168	Cj0671	626859	-	antisense	Yes	No	CCTGGTAAAATAAAACTATGATTGATAACAAATTTACCGATGTGA
NCTC11168	Cj0713	668350	+	internal	Yes	No	ATCCAAGAGATTTTTCTAAAAATTCTTATCATAAAGTCGATGATT
NCTC11168	Cj0732	687045	+	internal	Yes	No	TAAGCCCTAAAGAGCTTGATAGCTTGGGTTTAAAACACGATTTTA
NCTC11168	Cj0736	690552	-	antisense	No	Yes	TTTTCAAAATTTACTTCATTATTGTTTGATTCATCGGCGATATAGTCTATA
NCTC11168	Cj0795c	746335	-	internal	Yes	Yes	TCTTATTCATTCAAGTCTTGAAGGTTTAGAATTTAAAGTCGATATAGAAAA
NCTC11168	Cj0805	758052	+	internal	No	Yes	ATTATCAACCTAAAAATGCGATTTTACTTGTAAGTGGCGATATAGAAAGTA
NCTC11168	Cj0836	784817	+	internal	Yes	Yes	TATCTTTACATAGTGATGGAAAATTTTTAACTAGAGTCGATTTTTGTGATA

NCTC11168	Cj0851c	799029	-	internal	Yes	No	ATTTATCAAAAATTATTCATTTTAAAGTTCTAAAAGACGATTTAA
NCTC11168	Cj0859c	806007	-	primary	No	Yes	ATTATTTTAAATTATAGTATATTTTTTTAAAATAATGTCGATTTAGTGTAA
NCTC11168	Cj0860	806007	-	antisense	Yes	Yes	ATTATTTTAAATTATAGTATATTTTTTTAAAATAATGTCGATTTAGTGTAA
NCTC11168	Cj0887c	824919	-	internal	Yes	No	TTCAAATTACAGATCTAAAGCAAGGTAATAATAAACTCGATTTTC
NCTC11168	Cj0902	841592	+	internal	Yes	No	AAGTTTTAAAAAATATTAACACTACAATCTCAAAAGGCGATGTAA
NCTC11168	Cj0929	863452	+	internal	Yes	No	ACCCTCACTTAAGAGAACTTATAAAATCAAATATTGCCGATGTTA
NCTC11168	Cj0944c	884000	+	antisense	Yes	No	ATCTAAACAATCAAATTAAAAATCCAAAAATTAATTCCGATGTTT
NCTC11168	Cj0967	906078	+	internal	Yes	No	ATAATACTATAAATTATAACAATGTTCAAAGCATTACCGATGTAA
NCTC11168	Cj0977	911633	+	primary	No	Yes	TATACGCATTAAAATATAAATTCTTTTTATTTTTTGCCGATATTGAATTTT
NCTC11168	Cj1001	931132	+	primary	Yes	No	AAATTTGAAAAAATATTCAAATCATAAACATAATAACCGATTTAT
NCTC11168	Cj1007c	938002	+	antisense	Yes	No	ATAGTTATTAGGAATAAATATAATACGCCCACTACGACGATTTTT
NCTC11168	Cj1027c	958008	-	internal	Yes	No	GTGGGGTGACGGCTATTAAGTTTAAAGAGAAAAATGACGAATTAG
NCTC11168	Cj1034c	969400	-	primary	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCTA
NCTC11168	Cj1035c	969400	-	internal	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCTA
NCTC11168	Cj1052c	987649	-	internal	Yes	No	CTAAGCTTGAAAATGTTGAGAAAAAAGAGCAAAAAGTCGATGAAA
NCTC11168	Cj1099	1034336	-	antisense	Yes	No	AAATTAATTTGTCTATAAAAAGTTGCAAAAATATCTTCGATTTTA
NCTC11168	Cj1127c	1060678	-	internal	Yes	Yes	TAAGTCCAAAAAAATGTGAAATTTTATACAATGCCATCGATTTAGAAAACA
NCTC11168	Cj1231	1159589	+	internal	No	Yes	AACCGCAGTGGTGCTTAAAATTTTAAATGATAATGGCGATATAAATGAGCA
NCTC11168	Cj1294	1226376	+	internal	Yes	No	GTGCTTTGTATAAAAGTGAAAAAGTAGGAAAAAAAGCCGATCTTT
NCTC11168	Cj1330	1258037	+	internal	Yes	No	AAATTATCATCCATACTTTAGAAAAAAGCTTTAAAGCCGATTTGA
NCTC11168	Cj1339c	1270994	-	primary	No	Yes	TTTTATTGCTAAAGTATAAAATATTTTTTTGATTGCACGATATAGCATTTA
NCTC11168	Cj1352	1284874	+	internal	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTTCAAGA
NCTC11168	Cj1353	1284874	+	secondary	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTTCAAGA
NCTC11168	Cj1449c	1388231	+	antisense	Yes	Yes	CCATATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAA
NCTC11168	Cj1449c	1388235	+	antisense	Yes	Yes	ATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAATTTA
NCTC11168	Cj1450	1388231	+	secondary	No	Yes	CCATATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAA
NCTC11168	Cj1450	1388235	+	primary	No	Yes	ATTGTAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTAATTTA
NCTC11168	Cj1550c	1483227	-	internal	Yes	No	AAAAAAGGTAGAGTATTTCAAGAAATTAACACAAAACCGATTTAT
NCTC11168	Cj1553c	1487769	-	internal	Yes	No	GCTCATTTTTAAACGACGCTCATAAAGACTTAAAAGCCGATTTTA
NCTC11168	Cj1632c	1558601	-	primary	No	Yes	ATTCCTTTCTAAGAATATAAAATTTTTTAAGATTTTGTCGATTATAAAAGA
NCTC11168	Cj1633	1558601	-	antisense	Yes	Yes	ATTCCTTTCTAAGAATATAAAATTTTTTAAGATTTTGTCGATTATAAAAGA
NCTC11168	Cj1634c	1559676	+	antisense	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGCAAAACAAGCGATATAGTATTTG
NCTC11168	Cj1635c	1561203	-	internal	Yes	No	AGTCTTTTGCAAAAATTGCAAATAGTTTAAATTTGGGCGATTTTA
NCTC11168	Cj1641	1566830	+	internal	Yes	No	AAGCTATAAAAAAAGCTTTAGAACTAAAAGAAAATGACGATTTAG

NCTC11168	Cj1656c	1579989	-	primary	Yes	Yes	AGTTTAGTAAAAAATATTAAGTTTTAAAAATATCATCCGATTTAAATAAAG
NCTC11168	Cj1710c	1622000	-	internal	Yes	No	AAGTTTTTGTAGATAATCAAATCAACAAACAAATCGCCGATGATG
NCTC11168	CjNC4	1559676	+	primary	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGCAAAACAAGCGATATAGTATTTG
RM1221	CJE0105	115605	+	internal	Yes	No	CGGTTTTTATACGCGCTGATAAAAGCTTAAAATATGACGATGTTA
RM1221	CJE0117	123769	+	primary	Yes	No	ATATATGAGTAAAAATATTAAAAAAAATCATTTTTGCCGATTTAC
RM1221	CJE0138	145257	+	antisense	No	Yes	AAAATATTTCCTTCGATATAAATTTTTTAAATTATAACGATATTAAGCTTA
RM1221	CJE0139	145257	+	primary	No	Yes	AAAATATTTCCTTCGATATAAATTTTTTAAATTATAACGATATTAAGCTTA
RM1221	CJE0190	191356	-	internal	Yes	No	ATTTTACATTTGATGCAAAGTGCTAGTGAAGTGATGCCGATTTTT
RM1221	CJE0293	261216	+	antisense	Yes	No	AAGGATTAGATAAGGTTAAATTTTGATATGGATTTGCCGAATTTA
RM1221	CJE0338	305237	-	primary	No	Yes	TTTCGTATTATTTTTAATTATTTTTTTATTTTATATCGATTAACATTATTT
RM1221	CJE0352	315300	+	internal	Yes	Yes	AAAAGAAGTTATAATGCAAAATGAAATTTTAAAAAAACTCGATTTAGAGCA
RM1221	CJE0369	331732	+	internal	Yes	No	GCATGCTAAATGTTGCAAAAGAAAAATTTCCAAATGTCGAATTTA
RM1221	CJE0370	332687	+	internal	Yes	No	ATGCTCCTTTTTCTCTTATACGAGCTTTAAAAAAAGCCGATGATA
RM1221	CJE0422	379478	+	internal	No	Yes	TTGTACATTTGGAATTAAAAGATTTATTAAAAACTTGCGATATTATAAGTA
RM1221	CJE0440	396360	-	primary	No	Yes	GAATTTAATGTCCTTATAAATTTTTTAAGTCAAAAGTCGATATAAATTAAA
RM1221	CJE0465	421139	+	antisense	No	Yes	TTTGAAAACTCTCAAGATTATAAATTTTTAATTTTCTTCCGATATAGTATA
RM1221	CJE0466	421139	+	primary	No	Yes	TTTGAAAACTCTCAAGATTATAAATTTTTAATTTTCTTCCGATATAGTATA
RM1221	CJE0491	447431	+	antisense	No	Yes	CTTGTAAGGATGATCTTTTAAATTTTCATTTTTTAAACGATTTATAATTTC
RM1221	CJE0506	460286	-	primary	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
RM1221	CJE0507	460286	-	internal	No	Yes	CACTTAGAGCTGCATTTAAAGTTTTAGAACAAAATGACGATATAGTTATCA
RM1221	CJE0616	554713	+	antisense	No	Yes	AACGGCGTTGCAAGGCTGCATCTTTTTCAAAATATTTACGATATTCTTTTA
RM1221	CJE0629	566427	+	antisense	No	Yes	CTTTTTTCGCCAAATTTAAATATTTTAAGCCCTGATACGATTTCTATATTG
RM1221	CJE0643	580638	+	primary	No	Yes	TAATTATTTTAATTTAGCTACTTTTAACGATTAAAATTTTATAATCAAAAA
RM1221	CJE0651	588587	+	primary	No	Yes	TTTTCTTTATAGGTGTTCATAAATTTATAAATTTGTCGATATAAGCTTTTA
RM1221	CJE0725	659059	-	antisense	Yes	No	AGCAAAAGATGTAAAGGCGTATAAGCTAGCATAATGCCGATTTTA
RM1221	CJE0733	668685	+	antisense	Yes	Yes	CTTTTAAGCAAAAGAGTGAAAATTTCTTTATATTGTTCGATTTTTAAATTA
RM1221	CJE0741	676896	+	antisense	No	Yes	ATTTCAATTACGGCATTGATTTTATTTGGAATATCTCCGATTTTGATTTTG
RM1221	CJE0742	676896	+	antisense	Yes	Yes	ATTTCAATTACGGCATTGATTTTATTTGGAATATCTCCGATTTTGATTTTG
RM1221	CJE0744	679541	+	internal	Yes	Yes	TGCCAAAGTATGGATTGGATGATTTGTTTAAAATATCCGATTTTGTTATTT
RM1221	CJE0772	705340	-	antisense	Yes	No	AGAGCTATCATACCTGATTTAAAAATTTCATTAGAGCCGATTTTC
RM1221	CJE0807	742689	+	internal	Yes	No	TTTTTTAAAATCCAAATACAAAACTAGTTTAAAATCCCGATTTTT
RM1221	CJE0813	747089	+	internal	Yes	No	ATCCAAGAGATTTTTCTAAAAATTCCTATCATAAAGTCGATGATT
RM1221	CJE0832	765786	+	internal	Yes	No	TAAGCCCTAAAGAGCTTGATAGCTTGGGTTTAAAACACGATTTTA
RM1221	CJE0850	789484	+	internal	Yes	No	AGCCAAACACAGCAAATGATAAAAAGAAAAAAGATGACGATGTAA

RM1221	CJE0883	821081	+	internal	Yes	No	AAGGTTTTGAAAAATTAAACATTTTAAATTTTGAAGACGATGTAT
RM1221	CJE0886	824799	-	internal	Yes	Yes	TCTTATTCATTCAAGTCTTGAAGGTTTAGAATTTAAAGTCGATATAGAAAA
RM1221	CJE0923	862676	+	internal	Yes	Yes	TATCTTTACATAGTGATGGAAAATTTTTAACTAGAGTCGATTTTTGTGATA
RM1221	CJE0932	870201	+	antisense	No	Yes	TTTTGATATGTTCGCGATTAATTTGCAAAAGTTTTTTAAGATCAAATCTTG
RM1221	CJE0938	876840	-	internal	Yes	No	ATTTATCAAAAATTATTCATTTTAAAGTTCTAAAAGACGATTTAA
RM1221	CJE0946	883818	-	primary	No	Yes	AATTATTTTAAATTATAGTATATTTTTTAAAATAATGTCGATTTAGTGTAA
RM1221	CJE0947	883818	-	antisense	Yes	Yes	AATTATTTTAAATTATAGTATATTTTTTAAAATAATGTCGATTTAGTGTAA
RM1221	CJE0966	903107	-	internal	Yes	No	TTCAAATTACAGATCTAAAGCAAGGTAATAATAAACTCGATTTTC
RM1221	CJE0981	919778	+	internal	Yes	No	AAGTTTTAAAAAATATTAACACTACAATCTCAAAAGGCGATGTAA
RM1221	CJE1007	941638	+	internal	Yes	No	ACCCTCACTTAAAAGAACTTATAAAATCAAATATTGCCGATGTTA
RM1221	CJE1051	985719	-	primary	No	Yes	GCTAGTTTCTTTATAGTTTTCACGATTTTTTATCGCTCCTGTTCCAACTAA
RM1221	CJE1052	985719	-	antisense	No	Yes	GCTAGTTTCTTTATAGTTTTCACGATTTTTTATCGCTCCTGTTCCAACTAA
RM1221	CJE1059	989727	+	primary	No	Yes	TATACGCATTAAAATATAAATTCTTTTTATTTTTTACCGATATTGAAAATT
RM1221	CJE1080	1008505	-	antisense	No	Yes	AATTATAATTTTTTCTTTTTAAGGTATAAAATTTACGCTTTAGGTCGATAT
RM1221	CJE1081	1008620	+	primary	Yes	No	AAATTTGAAAAAATATTCAAATCATAAACATAATAACCGATTTAT
RM1221	CJE1084	1012127	-	antisense	Yes	No	CAGAGCTTTTGCTTGAAAAAGAAAAGATAGAATATGACGATGTTA
RM1221	CJE1087	1015489	+	antisense	Yes	No	ATAGTTATTAGGAATAAATATAATACGTCCACTACGACGATTTTT
RM1221	CJE1146	1062423	+	internal	Yes	No	AATAATCAAAAGAAAAGAATATATTATTTTAATGGGACGATTTTT
RM1221	CJE1171	1086268	-	internal	Yes	No	GTGGGGTGACGGCTATTAAGTTTAAAGAGAAAAATGACGAATTAG
RM1221	CJE1178	1097659	-	primary	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCCA
RM1221	CJE1179	1097659	-	internal	No	Yes	ATTTTGTAAAAAAGTGTCAATCTTTATCTTATAAAGCCGATTATACACCCA
RM1221	CJE1242	1160684	-	antisense	Yes	No	AAATTAATCTGTCTATAAAAAGTTGCAAAAATATCTTCGATTTTA
RM1221	CJE1269	1186421	-	internal	Yes	Yes	TAAATCCAAAAAAATGTGAAATTTTATACAATGCCATCGATTTAGAAAACA
RM1221	CJE1295	1213935	-	internal	Yes	No	GTATCAAAGAAATTTCAAGTTCTTTTGTCAAAAGTGGCGATGTAC
RM1221	CJE1366	1280369	+	internal	No	Yes	AACCGCAGTGGTGCTTAAAATTTTAAATGATAATGGCGATATCAATGAGCA
RM1221	CJE1486	1383179	+	internal	Yes	No	GTGCTTTGTATAAAAGTGAAAAAGTAGGAAAAAAAGCCGATCTTT
RM1221	CJE1500	1394768	+	internal	Yes	No	AACAACTCTTTGAAGGTTATTATAAAGAACTAGATGACGATTTAG
RM1221	CJE1519	1410788	+	internal	Yes	No	AAATGACCATCCATACTTTAGAAAAAAGCTTTAAAGCCGATTTGA
RM1221	CJE1528	1423719	-	primary	No	Yes	TTTTGACACCAAAGTATAAAATATTTTTTTGATTGCACGATATAGCATTTA
RM1221	CJE1541	1437599	+	internal	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTGCAAGA
RM1221	CJE1542	1437599	+	secondary	No	Yes	ATAGCACTTTGTGGAGCTTTGTTTTTACTTGTTTGCGATATCATTGCAAGA
RM1221	CJE1622	1526803	+	antisense	Yes	Yes	CCATATTATAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTGA
RM1221	CJE1622	1526807	+	antisense	Yes	Yes	ATTATAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTGATTTA
RM1221	CJE1623	1526803	+	secondary	No	Yes	CCATATTATAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTGA

RM1221	CJE1623	1526807	+	primary	No	Yes	ATTATAAAAATTATACTAAAATATTTTTAAAAATAACCGATTTTTGATTTA
RM1221	CJE1804	1699881	-	primary	No	Yes	ATAATTTTGTTAGTATAAATAAATTTGATTTTTTTGCCGATATCAAATTTG
RM1221	CJE1804	1699888	-	secondary	No	Yes	TTAAAAAATAATTTTGTTAGTATAAATAAATTTGATTTTTTTGCCGATATC
RM1221	CJE1805	1699881	-	antisense	No	Yes	ATAATTTTGTTAGTATAAATAAATTTGATTTTTTTGCCGATATCAAATTTG
RM1221	CJE1805	1699888	-	antisense	No	Yes	TTAAAAAATAATTTTGTTAGTATAAATAAATTTGATTTTTTTGCCGATATC
RM1221	CJE1806	1700974	+	antisense	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGCAAAACAAGCGATATAGTATTTG
RM1221	CJE1807	1702502	-	internal	Yes	No	AGTCTTTTGCAAAAATTGCAAATAGTTTAAATTTGGGCGATTTTA
RM1221	CJE1813	1708129	+	internal	Yes	No	AAGCTATAAAAAAAGCTTTAGAACTAAAAGAAAATGACGATTTAG
RM1221	CJE1828	1721288	-	primary	Yes	Yes	AGTTTGGTAAAAAATATTAAGTTTTAAAAATATCATCCGATTTAAATAAAG
RM1221	CJE1829	1721537	+	primary	No	Yes	ATTTATTTCAATAAATTTAAGTTTTATCTTATTTATACGATTTATAATATA
RM1221	CJE1879	1757847	-	internal	Yes	No	AAGTTTTTGTAGATAATCAAATCAACAAACAAATCGCCGATGATG
RM1221	CjNC4	1700974	+	primary	No	Yes	TTTATTAGTTAAAGTATAAAGTATTTAGCAAAACAAGCGATATAGTATTTG
RpoN motif							
81-176	CJJ81176_0078	69745	+	primary	Yes	Yes	TAAACTACAAAGCATCAAAACTTGGAACACTTCTTGCTTTAAAGACTTGTG
81-176	CJJ81176_0267	231655	-	secondary	Yes	Yes	CTGAATTTAATTTTTAAAAAACTTGGCACACCTTTTGCTTATAAAAAATCA
81-176	CJJ81176_0448	392170	+	primary	Yes	No	ACATTAATTTTTGCATAAAAAAGGAACACTTTTTGCTTGTATAGT
81-176	CJJ81176_0553	487774	-	primary	No	Yes	TATTTTTGTACAAAATAGTTAAATTGGAACAGTATTTGCTTGTTAATATTA
81-176	CJJ81176_0572	507438	+	secondary	No	Yes	ATTGAGATTTATTTTAAAATTTTGGCACAGTTTTTGCTTATTATTTTTTTA
81-176	CJJ81176_1045	964262	-	secondary	Yes	No	GATTTGTAATAAGTTTAAAGTTGGAACGCTTATTGCTTTTTATTT
81-176	CJJ81176_1048	968179	-	secondary	Yes	No	ATTGATGATAAATTTAAAACTTGGCTTGAAATTTGCTTGTATTTT
81-176	CJJ81176_1093	1012173	+	primary	Yes	No	TAATTAAATAATTTAAAATTTTGGAATGATTATTGCTTATTTTAT
81-176	CJJ81176_1257	1172053	+	primary	Yes	No	AAGTTTAAAAAAGATATAAATTGGAATAGAAATTGCTTGTTTGTT
81-176	CJJ81176_1338	1255052	-	primary	No	Yes	AAATTTTTAATCAAATTAAAACTTGGAACACTTCTTGCTTTAATCTTTTCG
81-176	CJJ81176_1431	1351346	-	internal	No	Yes	ATGTCCAACTCTTCTAAAATTTATGGAACATATAGTGCTTTGGTTTTATTA
81-176	CJJ81176_1455	1376542	+	primary	No	Yes	CTTTTACGCTTCTTAACTTACTTGGAACACTTCTTGCTTGATCTATCATCA
81-176	CJJ81176_1458	1378649	+	internal	No	Yes	TAAATAAAATGTTTGATATCAATGATGGAACAAATAATGCTTATGGAGATA
81-176	CJJ81176_1459	1378649	+	primary	No	Yes	TAAATAAAATGTTTGATATCAATGATGGAACAAATAATGCTTATGGAGATA
81-176	CJJ81176_1641	1550995	+	primary	Yes	No	AATTCTAACTTATCCTAATTTAGGAACACTTTTTGCTTTTTAAAT
81116	C8J_0043	68335	+	primary	Yes	Yes	TAAACTACAAAGCATCAAAACTTGGAACACTTCTTGCTTTAAAGACTTGTG
81116	C8J_0193	203749	+	internal	Yes	No	ACTTGCAAGTTGTGAGCAAAGAGGCACGCTTTTTGCTTCAGGGCT
81116	C8J_0220	228850	-	primary	Yes	Yes	GAATTTAATTTTTAAAAAACTTGGCACACCTTTTGCTTATAAAAAATAGCA
81116	C8J_0403	399840	+	primary	Yes	No	ACATTAATTTTTGCATAAAAAAGGAACACTTTTTGCTTGTATAGT
81116	C8J_0489	495627	-	primary	No	Yes	TATTTTTATACAAAATAGTTAAATTGGAACAGTATTTGCTTGTTAATATTA
81116	C8J_0508	515290	+	secondary	No	Yes	ATTGAGATTTATTTTAAAATTTTGGCACAGTTTTTGCTTATTATTTTTTTA

81116	C8J_0963	971798	-	primary	Yes	No	GATTTGCAATAAGTTTAAAGTTGGAACGCTTATTGCTTTTTATTT
81116	C8J_0966	975716	-	secondary	Yes	No	ATTGATGATAAATTTAAAACTTGGCTTGAAATTTGCTTGTATTTT
81116	C8J_1016	1026559	+	primary	Yes	No	TAATTAAATAATTTAAAATTTTGGAATGATTATTGCTTATTTTAT
81116	C8J_1185	1191234	+	primary	Yes	No	AAGTTTAAAAAAGATATAAATTGGAATAGAAATTGCTTGTTTATT
81116	C8J_1255	1267474	-	primary	No	Yes	AAATTTTTAATCAAATTAAAACTTGGAACACTTCTTGCTTTAATCTTTTCG
81116	C8J_1368	1384049	+	primary	No	Yes	TTTTTACTCTTCTTAACTCACTTGGAACACTTTTTGCTTGATCTATCATCA
81116	C8J_1371	1386158	+	internal	No	Yes	TAAATAAAATGTTTAATATCAATGATGGAACAAATAATGCTTATGGAGATA
81116	C8J_1372	1386158	+	primary	No	Yes	TAAATAAAATGTTTAATATCAATGATGGAACAAATAATGCTTATGGAGATA
81116	C8J_1552	1562524	+	primary	Yes	No	AATTCTAACTTATCCTAATTTAGGAACACTTTTTGCTTTTTAAAT
NCTC11168	Cj0040	59125	+	primary	Yes	Yes	TAAACTACAAAGTATCAAAACTTGGAACACTTCTTGCTTTAAAAACTTATG
NCTC11168	Cj0204	201047	+	internal	Yes	No	ACTTGCAAGTTGTGAGCAAAGAGGCACGCTTTTTGCTTCAGGGCT
NCTC11168	Cj0243c	225984	-	secondary	Yes	Yes	CTGAATTTAATTTTTAAAAAACTTGGCACACCTTTTGCTTATAAAAAATCA
NCTC11168	Cj0528c	490416	-	primary	No	Yes	TATTTTTATACAAAATAGTTAAATTGGAACAGTATTTGCTTGTTAATATTA
NCTC11168	Cj0547	510082	+	secondary	No	Yes	TTGAGATTTATTTTAAAATTTTGGCACAGTTTTTGCTTATTATTTTTTTTA
NCTC11168	Cj0887c	825930	-	secondary	Yes	Yes	TAAAATTATATAATTTAAATTTTTGGAACAGTTATTGCTTTTGTTTATTAG
NCTC11168	Cj1026c	957593	-	primary	Yes	No	GATTTGCAATAAGTTTAAAGTTGGAACGCTTATTGCTTTTTATTT
NCTC11168	Cj1029c	961511	-	secondary	Yes	No	ATTGATGATAAATTTAAAACTTGGCTTGAAATTTGCTTGTATTTT
NCTC11168	Cj1075	1009588	+	primary	Yes	No	TAATTAAATAATTTAAAATTTTGGAATGATTATTGCTTATTTTAT
NCTC11168	Cj1242	1170393	+	primary	Yes	No	AAGTTTTACAAAGGTATAAATTGGAATAGAAATTGCTTGCTTATT
NCTC11168	Cj1338c	1269112	-	primary	No	Yes	AAATTTTTAATTAAACTAAAACTTGGAACACTTTTTGCTTTAATCTTTTCG
NCTC11168	Cj1462	1398444	+	primary	No	Yes	TTTTTACTCTTCTTAACTCACTTGGAACACTTTTTGCTTGATCTATCATCA
NCTC11168	Cj1650	1574446	+	primary	Yes	No	AATTCTAACTTATCCTAATTTAGGAACACTTTTTGCTTTTTAAAT
RM1221	CJE0039	58116	+	primary	Yes	Yes	TAAACTACAAAGTATCAAAACTTGGAACACTTCTTGCTTTAAAAACTTATG
RM1221	CJE0197	199081	+	internal	Yes	No	ACTTGCAAGTTGTGAGCAAAGAGGCACGCTTTTTGCTTCAGGGCT
RM1221	CJE0293	261241	-	secondary	Yes	Yes	TGAATTTAATTTTTAAAAAACTTGGCACACATTTTGCTTATAAAAAAATCA
RM1221	CJE0569	513394	-	primary	No	Yes	GTTTCTAATTATGGAACATTATTTGCTTGAAAAAAGTGTATAATTTTTATA
RM1221	CJE0570	513394	-	antisense	No	Yes	GTTTCTAATTATGGAACATTATTTGCTTGAAAAAAGTGTATAATTTTTATA
RM1221	CJE0632	568858	-	primary	No	Yes	TATTTTTATACAAAACAGTTAAATTGGAACAGTATTTGCTTGTTAATATTA
RM1221	CJE0651	588520	+	secondary	No	Yes	AATTGAGATTTATTTTAAAATTTTGGCACAGTTTTTGCTTATTATTTTTTA
RM1221	CJE0966	904118	-	secondary	Yes	Yes	TAAAATTATATAATTTAAATTTTTGGAACAGTTATTGCTTTTGTTTATTAG
RM1221	CJE1170	1085853	-	primary	Yes	No	GATTTGTAATAAGTTTAAAGTTGGAACGCTTATTGCTTTTTATTT
RM1221	CJE1173	1089771	-	secondary	Yes	No	ATTGATGATAAATTTAAAACTTGGCTTGAAATTTGCTTGTATTTT
RM1221	CJE1218	1135937	+	primary	Yes	No	TAATTAAATAATTTAAAATTTTGGAATGATTATTGCTTATTTTAT
RM1221	CJE1378	1291172	+	primary	Yes	No	AAGTTTTACAAAGGTATAAATTGGAATAGAAATTGCTTGCTTATT

RM1221	CJE1440	1345730	+	primary	Yes	No	ATACAATTTTAATTTAAAACTTGGCATACTTATTGCTTAGTTAAA
RM1221	CJE1526	1421835	-	primary	No	Yes	AGAATTTTTAATCAAACTAAACTTGGAACACTTCTTGCTTTAATCTTTTCG
RM1221	CJE1527	1421835	-	internal	No	Yes	AGAATTTTTAATCAAACTAAACTTGGAACACTTCTTGCTTTAATCTTTTCG
RM1221	CJE1636	1537014	+	primary	No	Yes	CTTTTACGCTTCTTAACTTACTTGGAACACTTCTTGCTTGATCTATCATCA
RM1221	CJE1639	1539121	+	internal	No	Yes	TAAATAAAATGTTTGATATCAATGATGGAACAAATAATGCTTATGGAGATA
RM1221	CJE1640	1539121	+	primary	No	Yes	TAAATAAAATGTTTGATATCAATGATGGAACAAATAATGCTTATGGAGATA
RM1221	CJE1822	1715744	+	primary	Yes	No	AATTCTAACTTATCCTAATTTAGGAACACTTTTTGCTTTTTAAAT

Appendix Table 4. Comparison of gene lengths for annotation correction.

This table lists all ORFs of the four strains, their length, and whether a correction of the gene length is required. Orthologous genes are clustered. Only if orthologs in 3 or more strains were identified, which show a length difference, a correction attempt was performed. This table is provided as a separate Excel sheet on a DVD and can also be accessed using the link – (http://journals.plos.org/plosgenetics/article/file?type=supplementary&id=info:doi/10.1371/journal.pgen.1003495.s008)

Appendix Table 5. Orthologs with corrected annotation

This table lists orthologs for the four strains for which the annotation was corrected based on comparison of ORF length among different strains.

Locus tag	Strand	Original start position	Original end position	Corrected start position	Corrected end position	Description
NCTC11168						
Cj0006	+	6702	8010	6708	8010	Na+/H+ antiporter family protein
Cj0061c	-	75409	76126	75409	76102	flagellar biosynthesis sigma factor
Cj0182	+	177994	179200	178000	179200	transporter
Cj0204	+	199235	201233	199238	201233	oligopeptide transporter, OPT family
Cj0226	+	210766	211612	210772	211612	acetylglutamate kinase
Cj0234c	-	217474	218038	217474	218032	ribosome recycling factor
Cj0249	+	229830	230310	229833	230310	hypothetical protein
Cj0254	+	232543	233704	232567	233704	hypothetical protein
Cj0275	+	251783	253037	251813	253037	ATP-dependent protease ATP-binding subunit ClpX
Cj0427	+	390161	390497	390164	390497	hypothetical protein
Cj0436	+	404584	404995	404587	404995	pyridoxamine 5'-phosphate oxidase
Cj0440c	-	409014	409683	409014	409680	transcriptional regulator
Cj0454c	-	419856	420426	419856	420375	hypothetical protein
Cj0459c	-	423760	424027	423760	424003	hypothetical protein
Cj0555	+	517059	518352	517065	518352	dicarboxylate carrier protein MatC

Cj0563	+	526555	527158	526597	527158	hypothetical protein
Cj0590	+	549824	550535	549827	550535	SAM-dependent methyltransferase
Cj0591c	-	550524	550746	550524	550740	lipoprotein
Cj0669	+	623409	624138	623400	624138	ABC-transporter ATP-binding protein
Cj0693c	-	649945	650884	649945	650878	S-adenosyl-methyltransferase MraW
Cj0695	+	652529	653918	652535	653918	cell division protein FtsA
Cj0729	+	683847	684657	683850	684657	type I phosphodiesterase/nucleotide pyrophosphatase
Cj0782	+	735707	736490	735710	736490	quinol dehydrogenase membrane component
Cj0787	+	738525	738774	738528	738774	hypothetical protein
Cj0791c	-	741166	742441	741166	742435	aminotransferase
Cj0976	+	910710	911601	910725	911601	methyltransferase
Cj1052c	-	987018	989229	987018	989226	recombination and DNA strand exchange inhibitor protein
Cj1054c	-	989561	990860	989561	990857	UDP-N-acetylmuramate--L-alanine ligase
Cj1088c	-	1019785	1020970	1019785	1020958	folylpolyglutamate synthase/dihydrofolate synthase
Cj1155c	-	1085975	1087700	1085975	1087694	cation-transporting ATPase
Cj1196c	-	1125010	1125907	1125010	1125904	NAD(P)H-dependent glycerol-3-phosphate dehydrogenase
Cj1279c	-	1211676	1212912	1211676	1212909	fibronectin domain-containing lipoprotein
Cj1407c	-	1339758	1341147	1339758	1341129	phospho-sugar mutase
Cj1461	+	1397725	1398409	1397722	1398409	DNA methylase
Cj1486c	-	1422227	1422449	1422227	1422446	periplasmic protein
Cj1488c	-	1423308	1423572	1423308	1423575	cb-type cytochrome C oxidase subunit IV
Cj1502c	-	1434178	1435444	1434178	1435426	sodium/proline symporter
Cj1539c	-	1471603	1472041	1471603	1472008	anion-uptake ABC-transport system permease
Cj1564	+	1492002	1493991	1492035	1493991	methyl-accepting chemotaxis signal transduction protein
Cj1577c	-	1508130	1508925	1508130	1508922	NADH dehydrogenase subunit C
Cj1582c	-	1511378	1511981	1511378	1511975	peptide ABC transporter permease
Cj1629	+	1556442	1556832	1556421	1556832	exbD/tolR family transport protein
Cj1661	+	1584934	1586227	1584937	1586227	ABC transporter permease
Cj1710c	-	1621695	1623690	1621695	1623684	metallo-beta-lactamase family protein
Cj1717c	-	1627920	1629333	1627920	1629327	isopropylmalate isomerase large subunit
RM1221						
CJE0011	-	14999	15551	14999	15647	rubrerythrin
CJE0056	-	72496	73318	72496	73339	flagellar motor switch protein FliY
CJE0083	+	96513	97842	96504	97842	anaerobic C4-dicarboxylate transporter

CJE0176	+	177399	178686	177327	178686	transporter
CJE0316	-	282325	282835	282325	282856	hypothetical protein
CJE0461	+	416071	417781	415951	417781	GTP-binding protein
CJE0488	+	443614	445345	443509	445345	succinate dehydrogenase, flavoprotein subunit
CJE0622	+	559220	560330	559133	560330	hypothetical protein
CJE0628	+	564132	565785	564015	565785	Na/Pi-cotransporter
CJE0633	-	568914	569817	568914	569916	hypothetical protein
CJE0723	+	655963	656626	655978	656626	hypothetical protein
CJE0751	+	686343	686778	686262	686778	hypothetical protein
CJE0755	+	688311	690111	688305	690111	penicillin-binding protein 2
CJE0880	+	817775	818807	817688	818807	polyA polymerase
CJE0910	+	849197	849809	849191	849809	hypothetical protein
CJE0965	-	898870	901696	898870	901711	cell division protein FtsK
CJE1034	-	972211	972850	972211	972982	DnaJ domain-containing protein
CJE1069	+	998543	998882	998675	998882	hypothetical protein
CJE1271	-	1187873	1188950	1187873	1188953	general glycosylation pathway protein
CJE1326	+	1240824	1242123	1240746	1242123	C4-dicarboxylate transport protein
CJE1348	-	1263796	1264507	1263796	1264522	hypothetical protein
CJE1403	-	1317797	1318880	1317797	1318937	quinone-reactive Ni/Fe-hydrogenase, small subunit
CJE1677	-	1577732	1578755	1577732	1578659	selenide, water dikinase
CJE1686	-	1587878	1588580	1587878	1588592	hypothetical protein
CJE1729	-	1629154	1629730	1629154	1629790	conserved hypothetical protein, degenerate; this region contains one or more premature stops and/or frameshifts
CJE1794	+	1691439	1692420	1691409	1692420	riboflavin biosynthesis protein RibD
81-176						
CJJ81176_0038	-	16738	17290	16738	17386	rubrerythrin
CJJ81176_0046	-	23530	25285	23530	25309	methyl-accepting chemotaxis protein
CJJ81176_0056	+	37521	38568	37572	38568	L-asparaginase
CJJ81176_0097	-	84135	84957	84135	84978	flagellar motor switch protein FliY
CJJ81176_0123	+	108152	109481	108143	109481	anaerobic C4-dicarboxylate transporter
CJJ81176_0132	+	116565	117357	116601	117357	gamma-glutamyl kinase
CJJ81176_0152	+	132334	133021	132331	133021	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase
CJJ81176_0159	-	137568	138429	137568	138555	hypothetical protein
CJJ81176_0161	-	138918	139368	138918	139380	hypothetical protein

CJJ81176_0214	+	185043	186330	184971	186330	transporter, putative
CJJ81176_0294	-	251697	252207	251697	252228	hypothetical protein
CJJ81176_0304	+	259802	260513	259763	260513	rod shape-determining protein MreC
CJJ81176_0324	-	279745	280051	279745	280147	TOBE domain-containing protein
CJJ81176_0383	+	331136	332492	331154	332492	phosphoglucosamine mutase
CJJ81176_0436	+	381687	383397	381567	383397	GTP-binding protein
CJJ81176_0491	+	432712	433318	432721	433318	putative transcriptional regulator
CJJ81176_0518	+	453835	455074	453796	455074	putative lipoprotein
CJJ81176_0543	+	478137	479247	478050	479247	hypothetical protein
CJJ81176_0554	-	487830	488919	487830	488832	hypothetical protein
CJJ81176_0555	+	488785	491317	488746	491317	hypothetical protein
CJJ81176_0556	+	491334	493638	491433	493638	isocitrate dehydrogenase, NADP-dependent
CJJ81176_0575	+	510298	510490	510214	510490	hypothetical protein
CJJ81176_0600	+	530213	531281	530261	531281	bifunctional 3,4-dihydroxy-2-butanone 4-phosphate synthase/GTP cyclohydrolase II protein
CJJ81176_0615	+	544200	545148	544143	545148	hypothetical protein
CJJ81176_0674	+	601862	602663	601835	602663	rare lipoprotein A
CJJ81176_0676	+	603210	603663	603147	603663	hypothetical protein
CJJ81176_0711	+	637593	639129	637623	639129	phosphate acetyltransferase
CJJ81176_0785	-	714900	716733	714900	716736	arginine decarboxylase
CJJ81176_0840	+	770942	771554	770936	771554	hypothetical protein
CJJ81176_0856	-	787606	787771	787606	787804	hypothetical protein
CJJ81176_0868	-	799002	799218	799002	799332	hypothetical protein
CJJ81176_0890	-	818056	820192	818056	820231	flagellar biosynthesis protein FlhA
CJJ81176_0904	-	835304	836576	835304	836591	3-phosphoshikimate 1-carboxyvinyltransferase
CJJ81176_0967	+	889986	890481	889971	890481	outer-membrane lipoprotein carrier protein
CJJ81176_0977	-	900470	901109	900470	901241	DnaJ domain-containing protein
CJJ81176_0986	+	908592	909171	908562	909171	hypothetical protein
CJJ81176_0987	+	909206	910358	909167	910358	hypothetical protein
CJJ81176_1033	-	954009	954708	954009	954705	high affinity branched-chain amino acid ABC transporter, ATP-binding protein
CJJ81176_1044	-	963292	963610	963292	963742	hypothetical protein
CJJ81176_1108	-	1024901	1025246	1024901	1025414	putative lipoprotein
CJJ81176_1177	-	1095566	1095731	1095566	1095761	hypothetical protein

CJJ81176_1218	-	1134831	1134978	1134831	1135023	hypothetical protein
CJJ81176_1227	-	1144678	1145389	1144678	1145404	hypothetical protein
CJJ81176_1235	-	1152815	1153877	1152815	1154006	sensor histidine kinase
CJJ81176_1240	-	1155927	1157100	1155927	1157175	sensor histidine kinase
CJJ81176_1249	+	1165548	1166283	1165461	1166283	M24/M37 family peptidase
CJJ81176_1268	+	1183152	1185180	1183131	1185180	organic solvent tolerance protein, putative
CJJ81176_1278	+	1194174	1195350	1194114	1195350	sensor histidine kinase
CJJ81176_1283	-	1198872	1199955	1198872	1200012	quinone-reactive Ni/Fe-hydrogenase, small subunit
CJJ81176_1288	-	1206253	1208425	1206253	1208449	RelA/SpoT family protein
CJJ81176_1292	-	1210628	1211387	1210628	1211435	cell division protein FtsX, putative
CJJ81176_1418	-	1340208	1340988	1340208	1340970	putative methyltransferase
CJJ81176_1473	-	1393482	1393926	1393482	1393908	50S ribosomal protein L13
CJJ81176_1489	-	1408760	1409042	1408760	1409189	hypothetical protein
CJJ81176_1499	-	1421191	1421893	1421191	1421929	molybdate transport repressor
CJJ81176_1503	-	1424437	1426678	1424437	1427242	formate dehydrogenase, alpha subunit, selenocysteine-containing
CJJ81176_1506	-	1427407	1428109	1427407	1428121	hypothetical protein
CJJ81176_1560	-	1481425	1481566	1481425	1481653	hypothetical protein
CJJ81176_1614	+	1527774	1528287	1527765	1528287	hypothetical protein
CJJ81176_1619	+	1532554	1532992	1532566	1532992	TonB system transport protein ExbB
CJJ81176_1626	-	1537360	1538038	1537360	1538035	ribonuclease III
CJJ81176_1629	+	1539477	1541322	1539504	1541322	DNA primase
CJJ81176_1634	+	1543816	1544887	1543792	1544887	PDZ domain-containing protein
81116						
C8J_0021	-	25676	26522	25676	26579	putative ribosomal pseudouridine synthase
C8J_0063	-	88895	89081	88895	89054	hypothetical protein
C8J_0068	-	92663	93401	92663	93404	hypothetical protein
C8J_0092	+	114179	114815	114161	114815	biotin--protein ligase
C8J_0109	+	127162	128179	127258	128179	malonyl CoA-acyl carrier protein transacylase
C8J_0110	+	128178	128865	128175	128865	5'-methylthioadenosine/S-adenosylhomocysteine nucleosidase
C8J_0119	-	134844	135294	134844	135306	hypothetical protein
C8J_0126	+	142340	142769	142367	142769	hypothetical protein
C8J_0167	-	176862	177768	176862	177771	iron ABC transporter, ATP binding subunit
C8J_0190	-	199232	199892	199232	199838	putative integral membrane protein
C8J_0287	-	291682	292057	291682	292021	SMR family multidrug efflux pump

C8J_0351	+	352740	353133	352656	353133	hypothetical protein
C8J_0362	+	361919	362414	361916	362414	shikimate kinase
C8J_0419	+	421673	423746	421631	423746	hypothetical protein
C8J_0430	-	431163	431745	431163	431814	hypothetical protein
C8J_0451	+	448314	452442	448305	452442	DNA-directed RNA polymerase subunit beta
C8J_0504	+	510423	512130	510420	512130	prolyl-tRNA synthetase
C8J_0518	-	524366	525437	524366	525452	hypothetical protein
C8J_0551	+	553709	554561	553706	554561	bifunctional riboflavin kinase/FMN adenylyltransferase
C8J_0595	+	598338	599217	598377	599217	NOL1/NOP2/sun family protein
C8J_0622	+	625128	625401	625155	625401	S4 domain-containing protein
C8J_0624	+	625797	626526	625788	626526	ABC transporter, ATP-binding protein
C8J_0627	+	629523	629619	629505	629619	hypothetical protein
C8J_0656	+	655263	656766	655260	656766	phosphate acetyltransferase
C8J_0715	-	720940	722773	720940	722776	arginine decarboxylase
C8J_0724	-	728956	729883	728956	729868	ABC transporter permease
C8J_0743	+	748693	749527	748648	749527	hypothetical protein
C8J_0764	+	770664	771396	770676	771396	3-deoxy-manno-octulosonate cytidylyltransferase
C8J_0772	+	778321	779065	778288	779065	A24 family peptidase
C8J_0793	+	799007	800099	798974	800099	Ser/Thr protein phosphatase family protein
C8J_0827	-	836033	836705	836033	836696	DNA-binding response regulator
C8J_0829	-	838262	838703	838262	838751	hypothetical protein
C8J_0832	-	841375	842647	841375	842662	3-phosphoshikimate 1-carboxyvinyltransferase
C8J_0837	-	847048	847195	847048	847228	hypothetical protein
C8J_0901	-	913245	913617	913245	913587	hypothetical protein
C8J_0927	-	935610	936366	935610	936372	hypothetical protein
C8J_0953	-	962997	964047	962997	964050	high affinity branched-chain amino acid ABC transporter, permease protein
C8J_0962	-	970831	971149	970831	971281	hypothetical protein
C8J_0965	-	974467	974896	974467	975043	transformation system protein
C8J_0983	-	993758	994559	993758	994562	thiamine biosynthesis protein ThiF
C8J_1020	+	1028726	1029119	1028666	1029119	hypothetical protein
C8J_1029	-	1036756	1037917	1036756	1037929	folC bifunctional protein
C8J_1052	-	1061564	1062203	1061564	1062191	putative integral membrane protein
C8J_1063	-	1072033	1072624	1072033	1072645	general glycosylation pathway protein

C8J_1069	-	1078497	1079574	1078497	1079577	general glycosylation pathway protein
C8J_1144	+	1149963	1150734	1149945	1150734	NLPA family lipoprotein
C8J_1190	-	1196669	1197140	1196669	1197164	hypothetical protein
C8J_1276	-	1287805	1288297	1287805	1288321	cytochrome c-type protein nrfH
C8J_1279	-	1293552	1293963	1293552	1294005	hypothetical protein
C8J_1367	+	1383330	1384014	1383327	1384014	DNA methylase
C8J_1369	+	1385151	1385460	1385118	1385460	hypothetical protein
C8J_1376	-	1390754	1392302	1390754	1392314	general secretory pathway protein E
C8J_1380	-	1394885	1395182	1394885	1395197	transformation system protein
C8J_1393	-	1408910	1409174	1408910	1409177	cytochrome c oxidase, cbb3-type, subunit IV
C8J_1414	-	1431937	1434151	1431937	1434742	hypothetical protein
C8J_1478	-	1497833	1498502	1497833	1498544	peptide ABC transporter, ATP-binding protein
C8J_1516	+	1530966	1531974	1530987	1531974	putative haemin transport system permease protein
C8J_1519	-	1533549	1534428	1533549	1534467	hypothetical protein
C8J_1531	+	1544520	1544910	1544499	1544910	biopolymer transport exbD protein
C8J_1557	-	1566568	1567699	1566568	1567717	Na+/H+ antiporter NhaA
C8J_1566	+	1576105	1576558	1576069	1576558	periplasmic thioredoxin

Appendix Table 6. 5' UTR comparisons

For all possible combinations of the four strains, the CDS (coding sequence) length and 5' UTR length of orthologous genes were compared. Genes were considered as orthologous if they were best reciprocal hits in a BLAST search on DNA level (i.e. *blastn*) in two compared strains. The gene names, lengths of the CDS of the two genes, the type of the TSS (pTSS = both primary; sTSS = both secondary; mixed_pTSS_sTSS = one of two orthologs has a primary TSS, the other ortholog has a secondary TSS), and the lengths of the 5' UTR of each gene of a homolog pair as well as the comparison status of these ("equal" or "different") are listed. This table is provided as a separate Excel sheet on a DVD and can also be accessed using the link – (http://journals.plos.org/plosgenetics/article/file?type=supplementary&id=info:doi/10.1371/journal.pgen.1003495.s010)

Appendix Table 7. sRNA candidates in the four *C. jejuni* strains.

This table lists sRNA candidates in the chromosome of *C. jejuni* strains NCTC11168, RM1221, 81116, 81-176 and the pVir and pTet plasmids of *C. jejuni* 81-176.

sRNA candidates in *C. jejuni* NCTC11168

sRNA	Strand	Start	End	Length	Orientation	LFG	RFG	LFG Description	RFG Description	Comments
SRP RNA	+	66641	66748	108	<>>	Cj0045c	Cj0046	putative iron-binding protein	transport protein pseudogene	Detected on Northern blot
CJnc10	+	94344	94249	96	>><	Cj0082	Cj0085c	*cydB*	aspartate racemase	Detected on Northern blot
CJnc11	-	237792	237728	65	>><	Cj0259	Cj0261c	dihydroorotase (*pyrC*)	putative SAM-dependent methyltransferase	Detected in dRNA-seq data. Not probed on Northern blot. Derived from 3' end of Cj0260c.
CJnc20	-	245380	245225	156	<<<	Cj0265c	Cj0266c	putative cytochrome C-type haem-binding periplasmic protein	putative integral membrane protein	Detected on Northern blot
TPP	+	418415	418550	136	>>>	Cj0452	Cj0453	DNA polymerase III subunit epsilon	thiamine biosynthesis protein (*thiC)*	Detected in dRNA-seq data. Not probed on Northern blot.
CJnc22	+	430270	430366	97	>><	Cj0463	Cj0465c	*zinc protease-like protein*	*group III truncated haemoglobin*	Detected in dRNA-seq data. Not detected on Northern blot
RnpB	-	513453	513137	317	><>	Cj0550	Cj0551	hypothetical protein	elongation factor P	Detected on Northern blot
CJnc21	+	543101	543269	169	>>>	Cj0582	Cj0584	*aspartate kinase (lysC)*	*DNA polymerase III subunit delta*	Detected on Northern blot
CJnc30	+	591313	591426	99	>><	Cj0628	Cj0630c	putative lipoprotein	DNA polymerase III subunit delta	Detected on Northern blot
CJnc40	-	591456	591356	101	>><	Cj0628	Cj0630c	putative lipoprotein	DNA polymerase III subunit delta	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc60	-	671798	671518	281	><>	Cj0717	Cj0718	putative ArsC family protein	DNA polymerase III subunit alpha	Detected on Northern blot. Note: Probably corresponds to a small mRNA. Contains non-annotated *selW* ORF [204].
CJnc110	+	1127994	1128130	137	>>>	Cj1198	Cj1199	S-ribosylhomocysteinase	putative iron/ascorbate-dependent oxidoreductase	Detected on Northern blot
CJnc120	+	1151053	1151149	97	>><	Cj1220	Cj1222c	*groES*	sensor histidine kinase (*dccS*)	Detected on Northern blot (processed from 3' end of Cj1221)
6S RNA	+	1179587	1179772	186	>>>	Cj1249	Cj1250	hypothetical protein	phosphoribosylamine--glycine ligase	Detected on Northern blot.

(CJnc130)										
pRNA	-	1179623	1179610	14	><	Cj1249	Cj1250	hypothetical protein	phosphoribosylamine-glycine ligase	Detected in dRNA-seq data. Not probed on Northern blot.
CJnc140	+	1188925	1188997	73	>>>	Cj1258	Cj1259	putative phosphotyrosine protein phosphatase	major outer membrane protein (*porA*)	Detected on Northern blot
tmRNA	-	1293658	1293300	359	><	Cj1359	Cj1361c	polyphosphate kinase (*ppk*)	hypothetical protein	Detected on Northern blot
crRNA1	+	1455167	1455204	38	>>>	Cj1519	crRNA2	putative molybdopterin biosynthesis protein (*moeA2*)	crRNA2	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA2	+	1455232	1455269	38	>>>	crRNA1	crRNA3	crRNA1	crRNA3	Detected on Northern blot
crRNA3	+	1455299	1455336	38	>>>	crRNA2	crRNA4	crRNA2	crRNA4	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA4	+	1455363	1455400	38	>>>	crRNA3	crRNA5	crRNA3	Tracr RNA	Detected on Northern blot
Tracr RNA	+	1455497	1455570	74	>><	crRNA4	Cj1521c	crRNA4	CRISPR-associated protein 2 (*cas2*)	Detected on Northern blot
CJnc150	-	1393825	1393745	81	><	Cj1455	Cj1456c	peptide chain release factor 2	putative periplasmic protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc160	+	1554542	1554568	27	<><	Cj1625c	Cj1626c	amino acid transporter	putative periplasmic protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc170	+	1559676	1559722	47	>><	Cj1633	Cj1634c	putative ATP-binding protein	chorismate synthase (*aroC*)	Detected on Northern blot
CJnc180	+	1575014	1575112	99	>><	Cj1650	Cj1651c	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc190
CJnc190	-	1575257	1575042	216	><	Cj1650	Cj1651c	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc180
CJnc210	+	1600382	1600468	87	>>>	Cj1677	Cj1679	putative lipoprotein	hypothetical protein	Detected in dRNA-seq data
CJnc230	-	1638039	1637942	98	<<<	Cj1727c	Cj1729c	putative O-acetylhomoserine (thiol)-lyase	flagellar hook protein (*flgE*)	Detected on Northern blot (processed)
IGR_906748_907066	-	907066	906748	319	><	Cj0967	Cj0969	putative periplasmic protein	probable periplasmic protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJas_Cj0363c	+	329341	329423	83	>>>	Cj0361	Cj0364	lipoprotein signal peptidase	hypothetical protein	Detected in dRNA-seq data. Transcribed from the 3' end of Cj0362. Antisense to Cj0363c (coproporphyrinogen III oxidase). Not detected on Northern blot.
CJas_Cj0566	-	528700	-	>100	><	Cj0565	Cj0567	probable pseudogene	hypothetical protein	Detected in dRNA-seq data. Antisense to Cj0566 (hypothetical protein). Not detected on Northern blot.
CJas_Cj0704	-	661794	-	>100	><	Cj0703	Cj0705	hypothetical protein	hypothetical protein	Detected on Northern blot. Antisense to Cj0704 (glycyl-tRNA synthetase subunit alpha)
CJas_Cj1667c	+	1589598	1589666	69	<><	Cj1666c	Cj1668c	putative periplasmic protein	putative periplasmic protein	Detected on Northern blot. Antisense to Cj1667c (*repA*).

CJas_Cj0168c	+	165893	-	>100	>>>	Cjp03	Cj0169	tRNA-Glu	superoxide dismutase	Detected in dRNA-seq data. Antisense to Cj0168c (putative periplasmic protein). Not detected on Northern blot.

sRNA candidates in *C. jejuni* RM1221

sRNA	Strand	Start	End	Length	Orientation	LFG	RFG	LFG Description	RFG Description	Comments
SRP RNA	+	65630	65737	108	<>>	CJE0044	CJE0045	hypothetical protein	hypothetical protein	Detected on Northern blot.
CJnc10	+	93253	93348	96	>><	CJE0079	CJE0080	hypothetical protein	aspartate racemase, putative	Detected on Northern blot.
TPP	+	456573	456708	136	>>>	CJE0502	CJE0503	DNA polymerase III subunit epsilon	thiamine biosynthesis protein (*thiC)*	Detected on dRNA-seq data. Not probed on Northern blot.
CJnc22	+	468429	-	>100	>><	CJE0513	CJE0515	zinc protease-like protein	group III truncated haemoglobin	Detected in dRNA-seq data. Not detected on Northern blot.
RnpB	-	591873	591557	317	><>	CJE0654	CJE0655	hypothetical protein	elongation factor P	Detected on Northern blot.
Cjnc60	-	750538	750258	281	><>	CJE0817	CJE0818	hypothetical protein	DNA polymerase III subunit alpha	Detected on Northern blot. Note: Probably corresponds to a small mRNA. Contains non-annotated *selW* ORF [204].
CJnc90	-	1071289	1071193	96	><>	CJE1154	CJE1155	hypothetical protein	hypothetical protein	Detected on Northern blot.
CJnc110	+	1248773	1248909	137	>>>	CJE1332	CJE1333	S-ribosylhomocysteinase	2OG-Fe(II) oxygenase family oxidoreductase	Detected on Northern blot.
CJnc120	+	1271831	1271927	97	>>>	CJ1355	CJ1357	*groES*	sensor histidine kinase	Detected on Northern blot (processed from 3' end of CJE1356)
6S RNA (CJnc130)	+	1300366	1300551	186	>>>	CJE1386	CJE1386	hypothetical protein	phosphoribosylamine-glycine ligase	Detected on Northern blot.
pRNA	-	1300402	1300389	14	><>	CJE1386	CJE1386	hypothetical protein	phosphoribosylamine-glycine ligase	Detected on dRNA-seq data. Not probed on Northern blot.
CJnc140	+	1309487	1309559	73	>>>	CJE1394	CJE1395	putative phosphotyrosine protein phosphatase	major outer membrane protein	Detected on Northern blot.
tmRNA	-	1448922	1448564	359	<<<	CJE1552	CJE1553	hypothetical protein	hypothetical protein	Detected on Northern blot.
crRNA1	+	1594145	1594182	38	>>>	CJE1693	crRNA2	Putative molybdopterin biosynthesis MoeA protein	crRNA2	Detected on dRNA-seq data. Not probed on Northern blot.
crRNA2	+	1594211	1594248	38	>>>	crRNA1	crRNA3	crRNA1	crRNA3	Detected on dRNA-seq data. Not probed on Northern blot.
crRNA3	+	1594277	1594314	38	>>>	crRNA2	Tracr RNA	crRNA2	Tracr RNA	Detected on dRNA-seq data. Not probed on Northern blot.

Tracr RNA	+	1594409	1594482	74	>><	crRNA3	CJE1694	crRNA3	CRISPR-associated protein 2 (*cas2*)	Detected on Northern blot.
CJnc160	+	1695716	1695742	27	<><	CJE1797	CJE1798	serine transporter	hypothetical protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc170	+	1700974	1701021	48	>><	CJE1805	CJE1806	hypothetical protein	chorismate synthase	Detected on dRNA-seq data. Oligo did not bind to RM1221 CJnc170 in Northern blot.
CJnc180	+	1716312	1716466	99	>><	CJE1822	CJE1823	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc190
CJnc190	-	1716340	1716556	217	>><	CJE1822	CJE1823	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc180
CJnc230	-	1774449	1774352	98	<<<	CJE1895	CJE1896	homoserine O-acetyltransferase	flagellar hook protein (*flgE)*	Detected on Northern blot (processed)
CJnc160	+	1695716	1695742	27	<><	CJE1797	CJE1798	serine transporter	hypothetical protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJas_Cj0363c	+	367027	367109	83	>>>	CJE0410	CJE0413	lipoprotein signal peptidase	hypothetical protein	Detected in dRNA-seq data. Transcribed from the 3' end of CJE0411. Antisense to CJE0412 (coproporphyrinogen III oxidase). Not detected on Northern blot.
CJas_Cj0704	-	740532	-	>100	><	CJE0803	CJE0805	hypothetical protein	hypothetical protein	Detected on Northern blot. Antisense to CJE0804 (glycyl-tRNA synthetase subunit alpha).
CJas_Cj1667c	+	1730939	-	>100	<><	CJE1838	CJE_tRNA-Leu-4	hypothetical protein	tRNA-Leu	Detected on Northern blot. Antisense to 5' end of CJE1839 and 3' of CJE1840.
CJas_Cj0168c	+	164655	-	>100	>>>	CJE_tRNA-Glu-1	CJE0164	tRNA-Glu	superoxide dismutase	Detected in dRNA-seq data. Antisense to CJE0163. Not detected on Northern blot.

sRNA candidates in *C. jejuni* 81116

sRNA	Strand	Start	End	Length	Orientation	LFG	RFG	LFG Description	RFG Description	Comments
SRP RNA	+	75877	75984	108	<>>	C8J_0048	C8J_0049	hypothetical protein	hypothetical protein	Detected on Northern blot
CJnc10	+	100738	100833	96	>><	C8J_0076	C8J_0077	hypothetical protein	putative aspartate racemase	Detected on Northern blot
CJnc20	-	248257	248102	156	<<<	C8J_0242	C8J_0243	putative cytochrome C-type haem-binding periplasmic protein	hypothetical protein	Detected on Northern blot
TPP	+	427729	427864	136	>>>	C8J_0425	C8J_0426	DNA polymerase III subunit epsilon	thiamine biosynthesis protein (*thiC)*	Detected in dRNA-seq data. Not probed on Northern blot.
CJnc22	+	439584	-	>100	>><	C8J_0436	C8J_0438	zinc protease-like protein	group III truncated haemoglobin	Detected in dRNA-seq data. Not Detected on Northern blot.

RnpB	-	518664	518348	317	><>	C8J_0511	C8J_0512	hypothetical protein	elongation factor P	Detected on Northern blot
CJnc50	+	650864	650959	96	>>>	C8J_0651	C8J_0652	hypothetical protein	invasion phenotype protein	Detected in dRNA-seq data. Not probed on Northern blot.
Cjnc60	-	681305	681025	281	><>	C8J_0684	C8J_0685	hypothetical protein	DNA polymerase III subunit alpha	Detected on Northern blot. Note: Probably corresponds to a small mRNA. Contains non-annotated *selW* ORF [204].
CJnc70	+	879931	880026	96	>>>	C8J_t0018	C8J_0874	tRNA-Leu	hypothetical protein	Detected in dRNA-seq data. Also antisense to CJnc50. Not Detected on Northern blot.
CJnc80	-	880052	879955	>100	><>	C8J_t0018	C8J_0874	tRNA-Leu	hypothetical protein	Detected on Northern blot. Also antisense to CJnc70
CJnc100	+	996243	996319	77	<><	C8J_0985	C8J_0986	hypothetical protein	hypothetical protein	Detected on Northern blot.
CJnc110	+	1148713	1148849	137	>>>	C8J_1142	C8J_1143	S-ribosylhomocysteinase	2OG-Fe(II) oxygenase family oxidoreductase	Detected on Northern blot
6S RNA (CJnc130)	+	1200515	1200700	186	>>>	C8J_1192	C8J_1194	hypothetical protein	phosphoribosylamine-glycine ligase	Detected on Northern blot
CJnc140	+	1209854	1209926	73	>>>	C8J_1202	C8J_1203	putative phosphotyrosine protein phosphatase	major outer membrane protein	Detected on Northern blot
tmRNA	-	1293552	1293194	359	<<<	C8J_1278	C8J_1279	hypothetical protein	hypothetical protein	Detected on Northern blot. Not probed on Northern blot.
crRNA1	+	1440760	1440797	38	>>>	C8J_1422	crRNA2	Putative molybdopterin biosynthesis MoeA protein	crRNA2	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA2	+	1440826	1440863	38	>>>	crRNA1	crRNA3	crRNA1	crRNA3	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA3	+	1440893	1440930	38	>>>	crRNA2	crRNA4	crRNA2	crRNA4	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA4	+	1440958	1440995	38	>>>	crRNA3	crRNA5	crRNA3	crRNA5	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA5	+	1441025	1441062	38	>>>	crRNA4	crRNA6	crRNA4	crRNA6	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA6	+	1441090	1441127	38	>>>	crRNA5	crRNA7	crRNA5	crRNA7	Detected in dRNA-seq data. Not probed on Northern blot.
crRNA7	+	1441156	1441193	38	>>>	crRNA6	Tracr RNA	crRNA6	Tracr RNA	Detected in dRNA-seq data. Not probed on Northern blot.
Tracr RNA	+	1441289	1441362	74	>><	crRNA7	C8J_1423	crRNA7	CRISPR-associated protein 2 (*cas2*)	Detected on Northern blot
CJnc170	+	1542619	1542645	27	<><	C8J_1527	C8J_1528	serine transporter	putative periplasmic protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc180	+	1563092	1563246	155	>><	C8J_1552	C8J_1553	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc190
CJnc190	-	1563337	1563121	217	><<	C8J_1552	C8J_1553	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc180
CJnc200	+	1568600	1568750	151	>>>	C8J_1559	C8J_1560	hypothetical protein	FTR1 family iron permease	Detected in dRNA-seq data. Not probed on Northern blot.

CJnc230	-	1624711	1624613	99	<<<	C8J_1628	C8J_1629	7-cyano-7-deazaguanine reductase	flagellar hook protein (*flgE)*	Detected on Northern blot (processed)
CJas_Cj0704	-	671301	-	>100	>< >	C8J_0670	C8J_0672	hypothetical protein	NGG1p interacting factor 3	Detected on Northern blot. Antisense to C8J_0671 (glycyl-tRNA synthetase subunit alpha)
CJas_Cj0168c	+	174436	-	>100	>>>	C8J_t0003	C8J_0165	tRNA-Glu	superoxide dismutase	Detected in dRNA-seq data. Antisense to C8J_0164. Not Detected on Northern blot.
CJas_CJJ81176_1020	+	947560	-	>100	>><	C8J_0938	C8J_0940	RNA polymerase sigma factor (*rpoD)*	rhomboid family protein	Detected in dRNA-seq data. Antisense to C8J_0939 (phosphohistidine phosphatase SixA). Not Detected on Northern blot.

sRNA candidates in *C. jejuni* 81-176

sRNA	Strand	Start	End	Length	Orientation	LFG	RFG	LFG Description	RFG Description	Comments
SRP RNA	+	77292	77399	108	<>>	CJJ81176_0083	CJJ81176_0084	hypothetical protein	hypothetical protein	Detected on Northern blot
CJnc10	+	104892	104987	96	>><	CJJ81176_0119	CJJ81176_0120	*cydB*	putative aspartate racemase	Detected on Northern blot
CJnc20	-	251065	250910	156	<<<	CJJ81176_0292	CJJ81176_0292	putative cytochrome C-type haem-binding periplasmic protein	hypothetical protein	Detected on Northern blot
TPP	+	420087	420222	136	>>>	CJJ81176_0477	CJJ81176_0478	DNA polymerase III subunit epsilon	thiamine biosynthesis protein (*thiC)*	Detected in dRNA-seq data. Not Probed on Northern blot.
RnpB	-	510809	510493	317	>< >	CJJ81176_0575	CJJ81176_0576	hypothetical protein	elongation factor P	Detected on Northern blot
Cjnc60	-	667532	667252	281	>< >	CJJ81176_0740	CJJ81176_0741	hypothetical protein	DNA polymerase III subunit alpha	Detected on Northern blot. Note: Probably corresponds to a small mRNA. Contains non-annotated *selW* ORF [204].
CJnc70	+	873863	873958	96	>>>	CJJ81176_1733	CJJ81176_0945	tRNA-Leu	hypothetical protein	Detected in dRNA-seq data. Also antisense to CJnc50. Not Probed on Northern blot.
CJnc80	-	873984	873887	98	>< >	CJJ81176_1733	CJJ81176_0945	tRNA-Leu	hypothetical protein	Detected on Northern blot. Also antisense to CJnc70
CJnc110	+	1129654	1129790	137	>>>	CJJ81176_1213	CJJ81176_1214	S-ribosylhomocysteinase	2OG-Fe(II) oxygenase family oxidoreductase	Detected on Northern blot
CJnc120	+	1152714	1152810	97	>><	CJJ81176_1233	CJJ81176_1235	*groES*	sensor histidine kinase	Detected on Northern blot (processed from 3' end of CJJ81176_1234)
6S RNA	+	1181243	1181428	186	>>>	CJJ81176_1265	CJJ81176_1266	hypothetical protein	phosphoribosylamine--glycine ligase	Detected on Northern blot

(CJnc130)										
pRNA	-	1181279	1181266	14	><>	CJJ81176_1265	CJJ81176_1266	hypothetical protein	phosphoribosylamine--glycine ligase	Detected in dRNA-seq data. Not Probed on Northern blot.
CJnc140	+	1190581	1190653	73	>>>	CJJ81176_1274	CJJ81176_1275	putative phosphotyrosine protein phosphatase	major outer membrane protein (*porA*)	Detected on Northern blot
tmRNA	-	1278617	1278259	359	><<	CJJ81176_1361	CJJ81176_1363	polyphosphate kinase (*ppk*)	hypothetical protein	Detected on Northern blot
CJnc150	-	1371923	1371843	81	><<	CJJ81176_1448	CJJ81176_1449	peptide chain release factor 2	hypothetical protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc160	+	1531082	1531108	27	<><	CJJ81176_1616	CJJ81176_1617	serine transporter	hypothetical protein	Detected in dRNA-seq data. Not probed on Northern blot. (low reads)
CJnc170	+	1536225	1536271	47	>><	CJJ81176_1624	CJJ81176_1625	hypothetical protein	chorismate synthase (*aroC*)	Detected on Northern blot
CJnc180	+	1551563	1551717	100	>><	CJJ81176_1641	CJJ81176_1642	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc190
CJnc190	-	1551808	1551592	217	><<	CJJ81176_1641	CJJ81176_1642	hypothetical protein	methionine aminopeptidase	Detected on Northern blot. Also antisense to CJnc180.
CJnc230	-	1613203	1613106	98	<<<	CJJ81176_0023	CJJ81176_0025	homoserine O-acetyltransferase	flagellar hook protein (*flgE)*	Detected in dRNA-seq data. Processed sRNA sense to CJJ81176_0024. Not Probed on Northern blot.
CJas_Cj0363c	+	333353	333435	83	>>>	CJJ81176_0384	CJJ81176_0387	lipoprotein signal peptidase	hypothetical protein	Detected in dRNA-seq data. Transcribed from the 3' end of CJJ81176_0385. Antisense to CJJ81176_0386 (coproporphyrinogen III oxidase). Not Detected on Northern blot.
CJas_Cj0704	-	657528	-	>100	><>	CJJ81176_0726	CJJ81176_0728	hypothetical protein	hypothetical protein	Detected on Northern blot. Antisense to CJJ81176_0727 (glycyl-tRNA synthetase subunit alpha)
CJas_Cj1667c	+	1566196	-	>100	<><	CJJ81176_1657	CJJ81176_1659	hypothetical protein	hypothetical protein	Detected on Northern blot. Antisense to CJJ81176_1658 (hypothetical protein)
CJas_Cj0168c	+	176666	-	>100	>>>	CJJ81176_1710	CJJ81176_0205	tRNA-Glu	superoxide dismutase	Detected in dRNA-seq data. Antisense to Cj0168c (putative periplasmic protein). Not Detected on Northern blot.
CJas_CJJ81176_1020	+	940021	-	>100	>><	CJJ81176_1019	CJJ81176_1021	RNA polymerase sigma factor (*rpoD*)	rhomboid family protein	Detected in dRNA-seq data. Antisense to CJJ81176_1020 (phosphohistidine phosphatase SixA). Not Detected on Northern blot.

sRNA candidates on the pVir and pTet plasmids of *C. jejuni* 81-176.

sRNA	Strand	Start	End	Length	Orientation	LFG	RFG	LFG Description	RFG Description	Comments
CJpt1	-	14943	14881	63	><	CJJ81176_pTet0016	CJJ81176_pTet0017	*cpp21*	*cpp22*	Detected on dRNA-seq data. Not detected on Northern blot.
CJpt 2	+	26519	26600	82	>>>	CJJ81176_pTet0032	CJJ81176_pTet0034	*cpp35*	*cmgb6*	Detected on dRNA-seq data. Not detected on Northern blot. Also internal TSS to CJJ81176_pTet0033 (*cmgB5)*
CJpt 3	-	36046	-	>100	><	CJJ81176_pTet0042	CJJ81176_pTet0044	*cpp45*	*cpp47*	Detected on dRNA-seq data. Not detected on Northern blot. Antisense RNA to CJJ81176_pTet0043 (*cpp46*)
CJpt 4	-	40999	40919	81	><	CJJ81176_pTet0047	CJJ81176_pTet0048	hypothetical protein	tetracycline resistance gene (*tetO*)	Detected on Northern blot.
CJpt 5	+	41001	-	>100	>>>	CJJ81176_pTet0047	CJJ81176_pTet0048	hypothetical protein	tetracycline resistance gene (*tetO*)	Detected on Northern blot.
CJpv1	+	16269	-	>100	>>>	CJJ81176_pVir0024	CJJ81176_pVir0025	hypothetical protein	para protein	Detected on dRNA-seq data. Not detected on Northern blot.
CJpv 2	+	25268	-	>100	>><	CJJ81176_pVir0033	CJJ81176_pVir0034	hypothetical protein	hypothetical protein	Detected on Northern blot.
CJpv 3	-	31119	31043	77	><	CJJ81176_pVir0047	CJJ81176_pVir0048	hypothetical protein	hypothetical protein	Detected on Northern blot.

Appendix Table 8. Sequence of sRNA candidates in *C. jejuni* NCTC11168

sRNA	Sequence
SRP RNA	GCCTTTTCTTAGACCTGTGCAATGCTATTTTTAAGCACCGCTTCAGGGTGGGAACACAGCAGAGCACTTGATTTTAGTGTGTGCCGCAGTTATCTGGGGAAGGGTTTT
CJnc10	AAATCTTTTCAAAATATTGCAATTTGCCCATTTTTGGGCATCTTTTAATCATAAATCTTTTAAAGCTTCATTAGCAGCATCAAGGGCGTGAATAAG
CJnc11	GTTATAGATATCAGAGATGCTCTTGTAAATAAAGCTTAAATTTTCCACCTTTAAGGTGGAAAATC
CJnc20	AATATAGAAGATAAGGAGTCAATCAACCTTTAAATCTAAAAGCACTTTGCTAAATGAGAGTTAAAAATGGGAGGGAAAATAAAGCTATAAAAAGCTTATATCTTAAAACTCAACGAGATCGAAATGGGCTCAGAAGTATTAATCTGGGCTCAAAAC
TPP	ATAAATGACGGGAGCTTGTGTAACAGGCTGAGAGTAAGCTAAAAGCTTAGACCGAACCGGATCTGGATAATACCAGCGTCGGGAAGATTTATAAATTTCAATACTTTAAAAATTTATACTCACCCTTCATTTATCC
CJnc22	GTAGGAGCGGAACGTTTAGGACTTGCTACTTTACATCAACTTCGTGGCCGTGTAGGTCGTGTAGGGCTTAAAAGTGCTTGCTATCTTTATACTAAGC
RnpB	TAAAGCATAGTAAATGCTCGCTTCTTTTTAGGAGAGGAAAGTCCGAGCTGCTAAAGACAAACATTCCATCTAACAGATGGCTAGGGTAACCTAAGGGATAGTGCAACAGAAAGAAAACTACCACGCAAGTGGAAAAGGTGAAACGGCGGGGTAAAAGCCCACCAGCGATTTTGGTAACAATTTCGGCTATGTAAACCCAATGTGCAGCAAGAAGGGATGGTTAGCGTCTTTGTTTTAACCCTTCGCTTGATTTTGTTTGCAAAAACAAAACTAGATAAATGAGCATTCAAGACAGAACTCGGCTTATCGCTATGCTT
CJnc21	TGGCATGGTTGGAGTCGCGTCGTTTAGAATGGGCTCCTTTGATGGCTGCAAGACTTAAATATCTTTTAGAAGGACGCACTTTTGTATTAATGTGTGATGAACAAAGATCTTGGTATGAAGAGTATTTTTTAAAAAATATAAATTCAAAAGCCACTCGTCCCATGCTCCC
CJnc30	AGATAATACTTAGTTTAATTAGTTTACTCTCCTACTTGTTAAAAATTTCATCCTGATAATTTTAAAATTATTAGGATGGATGTTTTTAAGGTTATTAAACTCTTTTTAGATTGC
CJnc40	GAGTTTTTAAAATACTTTAAAAGAATGATTGCAATCTAAAAAGAGTTTAATAACCTTAAAAACATCCATCCTAATAATTTTAAAATTATCAGGATGAAATT
CJnc60	TAAGATTTTAATCTTAATAAAAAGAACTTCTCAAAGGATGAAATGATGAAAGTAAAAATTGCTTATTGCAATCTTTGAAATTATCGCCCACAAGCTGCAAGGGTTGCAGAAGAACTACAAAGCGATTTCAAAGATGTTGAAGTAGAATTTGAAATTGGCGGCAGAGGTGATTTTATCGTTGAAGTAGACGGAAAAGTTATCTTCTCTAAAACACAACTTATTAATTGCGAAAGTGAAAGATTTCCATATCAAAATGAAATTAATCAATTAATAAAAAATAG
CJnc110	AATTTTATTATTTTAAAAGGAGAAAAAATGAAACTGATGTGCGAGATGTGTAGTATCTAGTGCGTCCTTTGGCTTAACATTATGTTTCCAAAGGTCCAAAAAGTGGATAAGGAGATATAATGTGTAAGAAAGTCAAA
CJnc120	CCTACTATGCCAGATATGAGCGGTATGGGAGGAATGGGTGGCATGGGCGGAATGATGTAATATCTGCTCTAACCTCTTATCAGCAAGGATTTTCCTT
6S RNA (CJnc130)	GCCTAAACAGTTTGTAGTATTTAAAGTGATCCAACACATTTTTATTAACTGCGATTGTGTGATTTACCGTGTTCTGTGGCATCGTTTGAGCTTTGAAAAAAGCGAGAA

	GTTGCAGCCTTTAAAAATTACCTAGCGGTTTTCTTTGACTTTTTGGGGTCAATCTTTGTTAAACGGCTGTTTGGGTTT
pRNA	GTGTTGGATCACTT
CJnc140	GAACCGAAAAACATTCATAAGAAAAACTCCTTAAAATTACACGCCTAGCTTAAATCCCTTTAGCTAGGCTTTT
tmRNA	GGGAGCGACTTGGCTTCGACAGGAGTAAGTCTGCTTAGATGGCATGTCGCTTTGGGCAAAGCGTAAAAAGCCCAAATAAAATTAAACGCAAACAACGTTAAATTC GCTCCTGCTTACGCTAAAGCTGCGTAAGTTCAGTTGAGCCTGAAATTTAAGTCATACTATCTAGCTTAATTTTCGGTCATTTTTGATAGTGTAGCCTTGCGTTTGACA AGCGTTGAGGTGAAATAAAGTCTTAGCCTTGCTTTTGAGTTTTGGAAGATGAGCGAAGTAGGGTGAAGTAGTCATCTTTGCTAAGCATGTAGAGGTCTTTGTGGGA TTATTTTTGGACAGGGGTTCGATTCCCCTCGCTTCCACCA
crRNA1	AGTTTTTAAAAGAGCTTGGCGGTTGTTTTAGTCCCTTT
crRNA2	AAAGTTTCATTAGTTGAATTTAACTGTTTTAGTCCCTT
crRNA3	GAATGAGGATGATGATATTTTACAGTTTTAGTCCCTTT
crRNA4	GTGTGCTAAAAAAAATGGACTTAAATGTTTTAGTCCCT
Tracr RNA	AAGAAATTTAAAAAGGGACTAAAATAAAGAGTTTGCGGGACTCTGCGGGGTTACAATCCCCTAAAACCGCTTTT
CJnc150	AAGAAAAGAGTTTATTATTCTGTGAATGATGCTCTTAGGTTTTATGCAAGATAATAAAACAAGCGGGAAAATCCCGCTATC
CJnc160	ACTAAATATATTTTGAGTTTTGCAATC
CJnc170	GAATCTTTTCAAAATATTGCAATCAAGCCCATGAAAATGGGCTTTTT
CJnc180	TAGCTTAAAGAGATTTTCCAAACTCTAAAAGAGAGTTAGGCTACGAATAAAAAGGGGGAGGGAAATAGCCTAACCCAAACGAGATCTTAAAAGATCTCT
CJnc190	AATAAGTTAAGTTACTAAATGATATTTTATAGCAGATTAAGTGTTTTTAGAATAAAATGTCACATCGTATTAAACTAAAACATAAATATGGTTTAGATAATTTAATAAAA ATTGTTAAATTATCTAAATATTTATTTTAGTTTAAAGAGATCTTTTAAGATCTCGTTTGGGTTAGGCTATTTCCCTCCCCCTTTTTATTCGTAGCCTAACTCTCTT
CJnc210	AGATAATATTTAGTTTAATTAGTTTACTCTCCTACTTGTTAAAAATTTCATCCTGATAATTTTAAAATTATTAGGATGGATGTTTTT
CJnc230	CAAGCCTATCTTGGTAGTCTTGTAATACTAGGCTTAGGCTTGAACTGCCTTCAATTTATTGCTAACTCACTCTTGCTAGGGGTGGGTTGCAATAAAAT
IGR_906748_907066	TGACTTGCTCCTTTTTGTAAGATTAAGATTTAAGAATCTTTTGTTTATATTTATAAAAATACAAATAAATCATTCCTATCTTAAAAATAGGAATGATTTATTTTATCCTTG CATAAATGTTTTTGTCATCTTGAGATATTCTTGCTTCATAAGAAAGCTTTTGTAGGTTCATATTTTCTAATCTATCTATAGAACTTAGTTTAAAATCTTCTTTGTTTTTTG CATTTTTTCAATAAAGCTTTGAAGTTCTTCTTGATTTAAAATATTAGAGTTTTCATCTATAAAACCTTGCTCACTTTATAAGATTAAGATTTCTTTA
CJas_Cj0363c	ATTATCATCGCTTTAATGATGGTTGTTCGTCCCTTTTAGTTTGTCTCCGCTTTAAATTCCATACTAGCGGAGTTAAGCGGATC
CJas_Cj0566	AATACTTTCGTTTGTTTTTTCTAAGGCTTGTGGAAACATACTTTCATTTTTAAAACTTTCTATGTTATTTTTTGAAGTTAAGATGTGTAAAAAATTTATT (3' end not known)

CJas_Cj0704	ATATTATCAGGACTAGGTTTAATTAAAACTTGAAATTGATAATAAGCACCTAAGCGATTAGGATTTTCACCATAACGTCCATCAGTAGGCCTTCTACTTG (3' end not known)
CJas_Cj1667c	TTACTATGTTTTTTCATAGTATGTTCTAATGGTGGATTTCCGCCAAGAGTGAAAACTCAAGGCGGATTT
CJas_Cj0168c	AGCTTTTGATATAAGTCTATTATAGACTTATATCAAGGCTGGGCATTATTGAGCTGGTGTTTGCTCTTTTGTTTTTTTGTGAGTTTTAGCTTCATGTTTT (3' end not known)

Appendix Table 9. Mapping statistics of *C. jejuni* CsrA coIP RNA-seq libraries. The table indicates the total number of sequenced cDNA reads considered in the analysis, the number of reads that were removed due to insufficient length (<12 nt) after poly(A)-tail clipping (before read mapping), the number of reads that were successfully mapped to the reference genomes (or the pVir and pTet plasmids of strain 81-176) using *segemehl* (see Methods), the number of mappings (i.e. some reads map to different locations with the same score), and the number of uniquely-mapped reads. For the number of mapped reads and number of uniquely mapped reads, the percentage values (relative to the total number of reads) are also listed.

	***C. jejuni* NCTC11168 Control coIP**	***C. jejuni* NCTC11168 CsrA-3xFLAG coIP**
Total number of reads	6,214,261	5,389,919
Failed size filter after clipping	144,133	113,458
Total number of mapped reads	5,933,127	5,164,774
Total number of mappings (NC_002163)	13,207,712	8,897,385
Uniquely mapped reads	2,079,985	3,103,694
% mappable reads	95.48	95.82
% of uniquely mapped reads	33.47	57.58

	***C. jejuni* 81-176 Control coIP**	***C. jejuni* 81-176 CsrA-3xFLAG coIP**
Total number of reads	6,053,715	4,605,355
Failed size filter after clipping	295,445	223,949
Total number of mapped reads	5,641,676	4,299,931
Total number of mappings	13,439,403	9,866,887
Uniquely mapped reads	1,545,585	1,385,653
% mappable reads	93.19	93.37
% of uniquely mapped reads	25.53	30.09
Mapped reads in plasmid pVir (NC_008770)	16,383	9,753
Mapped reads in chromosome (NC_008787)	5,601,153	4,268,583
Mapped reads in plasmid pTet (NC_008790)	24,140	21,595
Mappings in plasmid pVir (NC_008770)	17,832	10,780
Mappings in chromosome (NC_008787)	13,396,129	9,833,383
Mappings in plasmid pTet (NC_008790)	25,442	22,724

Appendix Table 10. Potential CsrA targets from *C. jejuni* NCTC11168. 5' UTRs and ORFs from *C. jejuni* NCTC11168 with more than 5-fold enrichment in the CsrA coIP compared to the WT coIP (control). For 17 of the 154 mRNAs with >5-fold enrichment, both the 5' UTR and coding sequences (CDS) were co-purified , with some having potentially shared CsrA binding sites.

						cDNA reads mapped			
Gene	**Name**	**Strand**	**Type**	**Start**	**End**	**WT coIP**	**CsrA-3F coIP**	**Enrichm-ent**	**log2 Enricm-ent**
Cj0040		+	5'-UTR	59125	59153	9.5	3389	356.2	8.48
Cj0547	*flaG*	+	5'-UTR	510149	510183	32.0	11077	346.1	8.44
Cj1249		+	5'-UTR	1178066	1178086	1.7	566	327.2	8.35
Cj1339c	*flaA*	-	5'-UTR	1270994	1270951	2277.3	693471	304.5	8.25
Cj0687c	*flgH*	-	5'-UTR	641959	641801	9.5	1911	200.9	7.65
Cj1462	*flgI*	+	5'-UTR	1398444	1398467	33.7	5750	170.5	7.41
Cj1655c		-	5'-UTR	1579683	1579635	23.4	3208	137.4	7.10
Cj1339c	*flaA*	-	ORF	1270950	1269232	4264.8	473588	111.0	6.80
Cj1293	*pseB*	+	5'-UTR	1224823	1224848	20.8	2298	110.7	6.79
Cj0040		+	ORF	59154	59477	83.9	9277	110.6	6.79
Cj1720		+	5'-UTR	1632224	1632247	13.8	1255	90.7	6.50
Cj1206c	*ftsY*	-	ORF	1136315	1135449	210.2	18420	87.6	6.45
Cj0547	*flaG*	+	ORF	510184	510549	250.8	18150	72.4	6.18
Cj1222c	*dccS*	-	ORF	1152345	1151155	46.7	3222	69.0	6.11
Cj1729c	*flgE*	-	ORF	1640701	1638104	1530.0	104324	68.2	6.09
Cj1124c	*pglC*	-	ORF	1056846	1056244	27.7	1772	64.0	6.00
Cj1338c	*flaB*	-	5'-UTR	1269112	1269087	15.6	915	58.8	5.88
Cj1462	*flgI*	+	ORF	1398468	1399514	229.2	12087	52.7	5.72
Cj1406c		-	5'-UTR	1339742	1339713	0.9	45	52.0	5.70
Cj1277c		-	5'-UTR	1210458	1210430	3.5	167	48.3	5.59
Cj1267c		-	5'-UTR	1198385	1198354	62.3	2814	45.2	5.50
Cj0697	*flgG2*	+	ORF	655202	656014	185.1	8133	43.9	5.46
Cj0016		+	5'-UTR	21135	21158	1.7	76	43.9	5.46
Cj1431c	*hddC*	-	ORF	1366419	1364671	122.8	5349	43.6	5.44
Cj1655c	*nhaA1*	-	ORF	1579634	1578486	154.8	6462	41.7	5.38
Cj1249		+	ORF	1178087	1179547	23.4	804	34.4	5.11
Cj0309c		-	ORF	281625	281311	14.7	485	33.0	5.04
Cj1416c		-	5'-UTR	1349684	1349613	2.6	78	30.1	4.91
Cj0435		+	5'-UTR	403804	403834	5.2	150	28.9	4.85
Cj0738		+	5'-UTR	692570	692725	0.9	25	28.9	4.85
Cj0950c		-	5'-UTR	890826	890790	2.6	72	27.7	4.79
Cj0842		+	ORF	789636	790121	36.3	968	26.6	4.74
Cj1437c		-	ORF	1375020	1373917	157.4	4070	25.9	4.69
Cj0670	*rpoN*	+	ORF	624138	625388	70.1	1747	24.9	4.64
Cj0508	*pbpA*	+	ORF	472703	474634	73.5	1673	22.8	4.51
Cj0172c		-	ORF	169012	167807	423.8	9459	22.3	4.48
Cj0509c	*clpB*	-	ORF	477235	474662	59.7	1329	22.3	4.48

Cj1152c	*gmhB*	-	ORF	1085297	1084737	58.8	1302	22.1	4.47
Cj0467		+	5'-UTR	431609	431661	6.1	132	21.8	4.45
Cj0310c		-	ORF	281967	281629	28.5	606	21.2	4.41
Cj1064		+	5'-UTR	1001190	1001217	6.1	128	21.1	4.40
Cj0414		+	5'-UTR	380911	380936	2.6	54	20.8	4.38
Cj0029		+	5'-UTR	37596	37666	0.9	18	20.8	4.38
Cj0886c		-	5'-UTR	823610	823524	22.5	460	20.5	4.35
Cj1720		+	ORF	1632248	1632877	76.1	1544	20.3	4.34
Cj0824	*uppS*	+	ORF	771964	772632	36.3	731	20.1	4.33
Cj0687c	*flgH*	-	ORF	641800	641102	163.5	3288	20.1	4.33
Cj0208		+	ORF	204603	205694	9.5	190	20.0	4.32
Cj0068	*pspA*	+	ORF	81044	81940	25.9	510	19.7	4.30
Cj0380c		-	ORF	347908	347123	20.8	396	19.1	4.25
Cj1432c		-	ORF	1369517	1366422	424.7	7531	17.7	4.15
Cj1064		+	ORF	1001218	1001837	25.1	441	17.6	4.14
Cj1555c		-	ORF	1489596	1488961	24.2	411	17.0	4.09
Cj0694		+	5'-UTR	651019	651042	13.8	223	16.1	4.01
Cj0769c	*flgA*	-	ORF	720139	719477	25.9	410	15.8	3.98
Cj0435	*fabG*	+	ORF	403835	404578	68.3	1079	15.8	3.98
Cj0823		+	ORF	771343	771960	22.5	347	15.4	3.95
Cj0394c		-	ORF	362335	361706	13.0	199	15.3	3.94
Cj0951c		-	ORF	891473	890799	5.2	79	15.2	3.93
Cj0763c	*cysE*	-	ORF	714776	714138	44.1	661	15.0	3.91
Cj0464	*recG*	+	ORF	428752	430575	110.7	1647	14.9	3.90
Cj1599	*hisB*	+	ORF	1527630	1528688	126.3	1843	14.6	3.87
Cj1338c	*flaB*	-	ORF	1269086	1267368	1265.4	17880	14.1	3.82
Cj1155c		-	ORF	1088264	1085907	63.1	880	13.9	3.80
Cj1394		+	ORF	1328292	1329659	6.1	83	13.7	3.78
Cj0374		+	ORF	342115	342606	70.9	946	13.3	3.74
Cj0172c		-	5'-UTR	169136	169013	3.5	44	12.7	3.67
Cj1052c		-	5'-UTR	989338	989230	1.7	22	12.7	3.67
Cj0369c		-	ORF	338911	337535	208.4	2650	12.7	3.67
Cj0818		+	ORF	767555	767782	9.5	118	12.4	3.63
Cj0238		+	ORF	220010	221893	66.6	820	12.3	3.62
Cj0718	*dnaE*	+	ORF	671946	675548	367.6	4487	12.2	3.61
Cj0386	*engA*	+	ORF	351446	352828	95.1	1153	12.1	3.60
Cj0995c	*hemB*	-	ORF	927152	926169	486.1	5693	11.7	3.55
Cj0977		+	ORF	911674	912252	96.0	1109	11.6	3.53
Cj0108	*atpC*	+	ORF	115313	115702	67.5	756	11.2	3.49
Cj1406c		-	ORF	1339712	1339362	12.1	134	11.1	3.47
Cj0200c		-	ORF	196470	196123	117.6	1297	11.0	3.46
Cj0186c		-	5'-UTR	182820	182788	0.9	9	10.4	3.38
Cj0337c		-	5'-UTR	305783	305715	8.6	89	10.3	3.36
Cj0952c		-	ORF	892382	891663	67.5	688	10.2	3.35
Cj0197c		-	5'-UTR	193645	193624	3.5	35	10.1	3.34

Cj0107		+	5'-UTR	113705	113911	28.5	287	10.1	3.33
Cj0853c	*hemL*	-	ORF	800921	799647	300.1	2975	9.9	3.31
Cj0499		+	ORF	463617	464102	33.7	332	9.8	3.30
Cj1218c	*ribA*	-	ORF	1146483	1145872	22.5	220	9.8	3.29
Cj1071	*ssb*	+	ORF	1005538	1006089	75.2	728	9.7	3.27
Cj1293	*pseB*	+	ORF	1224849	1225853	240.4	2280	9.5	3.25
Cj1171c	*ppi*	-	ORF	1100674	1100192	14.7	139	9.5	3.24
Cj1380		+	5'-UTR	1319534	1319560	7.8	73	9.4	3.23
Cj1340c		-	ORF	1272864	1271047	13.0	120	9.2	3.21
Cj1445c		-	5'-UTR	1386337	1386265	9.5	88	9.2	3.21
Cj1109		+	5'-UTR	1042435	1042553	2.6	24	9.2	3.21
Cj0173c	*cfbpC*	-	ORF	169962	169054	6.9	63	9.1	3.19
Cj1275c		-	ORF	1208974	1207781	25.1	223	8.9	3.15
Cj0466	*nssR*	+	ORF	431050	431646	17.3	152	8.8	3.14
Cj1420c		-	ORF	1354148	1353375	485.2	4176	8.6	3.11
Cj0081	*cydA*	+	ORF	91417	92979	18.2	156	8.6	3.10
Cj0860		+	ORF	806075	806947	45.8	388	8.5	3.08
Cj1006c		-	ORF	937496	936246	126.3	1059	8.4	3.07
Cj1083c		-	ORF	1014869	1014183	25.1	207	8.3	3.04
Cj1295		+	ORF	1226978	1228285	74.4	607	8.2	3.03
Cj0167c		-	ORF	165580	165017	0.9	7	8.1	3.02
Cj0122		+	ORF	124789	125478	5.2	42	8.1	3.02
Cj0481		+	5'-UTR	448661	448696	0.9	7	8.1	3.02
Cj0975		+	ORF	908891	910615	0.9	7	8.1	3.02
Cj0605		+	ORF	563205	564395	170.4	1373	8.1	3.01
Cj0270		+	ORF	248747	248953	16.4	130	7.9	2.98
Cj0005c		-	5'-UTR	6525	6499	3.5	27	7.8	2.96
Cj0036		+	ORF	53970	55319	31.1	242	7.8	2.96
Cj0373		+	ORF	341190	342125	122.0	941	7.7	2.95
Cj0855		+	5'-UTR	801310	801344	2.6	20	7.7	2.95
Cj0109	*exbB3*	+	ORF	115702	116256	37.2	283	7.6	2.93
Cj0848c		-	ORF	795504	795223	1.7	13	7.5	2.91
Cj1417c		-	ORF	1350216	1349614	13.8	103	7.4	2.90
Cj0317	*hisC*	+	ORF	287311	288405	58.8	436	7.4	2.89
Cj0136		+	5'-UTR	136816	136855	9.5	69	7.3	2.86
Cj1317	*pseI*	+	ORF	1245650	1246681	120.2	864	7.2	2.85
Cj0737		+	ORF	691596	692672	5.2	37	7.1	2.83
Cj1138		+	ORF	1072461	1073630	16.4	117	7.1	2.83
Cj1079		+	ORF	1011696	1012148	0.9	6	6.9	2.79
Cj1248	*guaA*	+	ORF	1176466	1178001	96.9	664	6.9	2.78
Cj1347c	*cdsA*	-	ORF	1280643	1279918	80.4	550	6.8	2.77
Cj1077	*ctsT*	+	ORF	1010738	1011040	6.9	47	6.8	2.76
Cj0548	*fliD*	+	ORF	510552	512480	643.5	4348	6.8	2.76
Cj0423		+	ORF	387168	387395	8.6	55	6.4	2.67
CJnc10		+	sRNA	94249	94344	21.6	137	6.3	2.66

Cj0086c	*ung*	-	ORF	95803	95108	33.7	211	6.3	2.65
Cj1445c	*kpsE*	-	ORF	1386264	1385146	152.2	949	6.2	2.64
Cj1523c		-	ORF	1459834	1456880	97.7	608	6.2	2.64
Cj0798c	*ddl*	-	ORF	749297	748257	44.1	273	6.2	2.63
Cj0318	*fliF*	+	ORF	288457	290139	163.5	1005	6.1	2.62
Cj1207c		-	5'-UTR	1136906	1136873	9.5	58	6.1	2.61
Cj1073c	*lon*	-	ORF	1008796	1006421	134.1	807	6.0	2.59
Cj0668		+	ORF	623010	623417	33.7	202	6.0	2.58
Cj1337	*pseE*	+	ORF	1265451	1267337	55.4	325	5.9	2.55
Cj0117	*pfs*	+	ORF	121680	122369	69.2	405	5.9	2.55
Cj0019c		-	ORF	25443	23665	71.8	418	5.8	2.54
Cj1711c	*ksgA*	-	ORF	1624459	1623659	45.8	266	5.8	2.54
Cj0034c		-	5'-UTR	52688	52669	0.9	5	5.8	2.53
Cj1601	*hisA*	+	ORF	1529269	1530000	45.0	260	5.8	2.53
Cj0569		+	ORF	530261	531130	0.9	5	5.8	2.53
Cj0818		+	5'-UTR	767524	767554	0.9	5	5.8	2.53
Cj0872		+	5'-UTR	813646	813674	0.9	5	5.8	2.53
Cj1405		+	5'-UTR	1338959	1338976	0.9	5	5.8	2.53
Cj1665		+	5'-UTR	1588342	1588430	0.9	5	5.8	2.53
Cj0472	*secE*	+	ORF	435925	436104	37.2	213	5.7	2.52
Cj1413c	*kpsS*	-	ORF	1346286	1345102	75.2	429	5.7	2.51
Cj0652	*pbpC*	+	ORF	611298	613103	22.5	128	5.7	2.51
Cj0762c	*aspB*	-	ORF	713966	712797	295.8	1675	5.7	2.50
Cj1112c		-	ORF	1045593	1045234	48.4	273	5.6	2.49
Cj1270c		-	ORF	1203382	1202291	206.7	1163	5.6	2.49
Cj1613c		-	ORF	1540613	1539858	92.5	519	5.6	2.49
Cj0361	*lspA*	+	ORF	328473	328943	37.2	208	5.6	2.48
Cj0248		+	5'-UTR	228913	228945	21.6	120	5.5	2.47
Cj0780		+	5'-UTR	732153	732178	91.7	506	5.5	2.46
Cj1690c	*rpsE*	-	ORF	1613061	1612618	34.6	190	5.5	2.46
Cj0186c		-	ORF	182787	182068	13.8	76	5.5	2.46
Cj0375		+	ORF	342615	343091	15.6	84	5.4	2.43
Cj1053c		-	ORF	989569	989219	6.1	32	5.3	2.40
Cj0628		+	ORF	587868	591303	786.2	4144	5.3	2.40
Cj1156	*rho*	+	ORF	1088373	1089671	102.1	536	5.3	2.39
Cj1508c		-	5'-UTR	1444630	1444608	1.7	9	5.2	2.38
Cj0227		+	5'-UTR	211456	211615	1.7	9	5.2	2.38
Cj0739		+	ORF	692878	693051	1.7	9	5.2	2.38
Cj1273c	*rpoZ*	-	ORF	1206999	1206775	25.1	130	5.2	2.37
Cj1181c	*tsf*	-	ORF	1108792	1107719	88.2	456	5.2	2.37
Cj0416		+	5'-UTR	383457	383558	50.2	259	5.2	2.37
Cj0991c		-	ORF	923420	922155	66.6	341	5.1	2.36
Cj1464	*flgM*	+	ORF	1399918	1400115	286.3	1429	5.0	2.32
Cj1552c	*mloB*	-	ORF	1487092	1485626	36.3	181	5.0	2.32
Cj0039c	*typA*	-	ORF	59019	57211	33.7	168	5.0	2.32

Appendix Table 11. List of potential CsrA binding sites. CsrA binding sites or peaks predicted by the peak detection algorithm for *C. jejuni* strains NCTC11168 and 81-176 using RIP-seq dataset.

Strain NCTC11168							
Peak ID	**Start**	**End**	**Strand**	**Overlapping genes (locus tags)**	**Overlapping genes (gene names)**	**Enrich-ment factor**	**Peak Size (bp)**
1	17526	17660	+	Cj0013	ilvD	30.1	135
2	21121	21235	+	Cj0016	Cj0016	37.9	115
3	29621	29855	+	Cj0023/Cj0024	purB/nrdA	142.4	235
4	35796	35930	+	Cj0027	pyrG	9.7	135
5	37676	37735	+	Cj0029	ansA	20.8	60
6	55201	55345	+	Cj0036	Cj0036	34.8	145
7	59101	59565	+	Cj0040/Cj0041	Cj0040/fliK	128.8	465
8	59566	59610	+	Cj0041	fliK	5.2	45
9	61181	61250	+	Cj0041	fliK	17.4	70
10	61301	61430	+	Cj0042	flgD	9.3	130
11	62446	62585	+	Cj0043	flgE	7.2	140
12	81441	81715	+	Cj0068	pspA	35.8	275
13	81761	81795	+	Cj0068	pspA	704.5	35
14	91876	91930	+	Cj0081	cydA	32.6	55
15	94236	94305	+	CJnc10	CJnc10	6.1	70
16	97461	97555	+	Cj0087/Cj0088	aspA/dcuA	6.2	95
17	97986	98170	+	Cj0088	dcuA	7.8	185
18	113686	113820	+	Cj0106	atpG	20.5	135
19	113881	113930	+	Cj0106/Cj0107	atpG/atpD	5.9	50
20	115381	115785	+	Cj0108/Cj0109	atpC/exbB3	19.6	405
21	115806	115860	+	Cj0109	exbB3	12.3	55
22	122061	122255	+	Cj0117	pfs	13.6	195
23	136811	136940	+	Cj0135/Cj0136	Cj0135/infB	22.2	130
24	138576	138630	+	Cj0136	infB	7.0	55
25	138796	138930	+	Cj0136	infB	19.9	135
26	167771	167795	+	Cj0170	Cj0170	10.0	25
27	180271	180525	+	Cj0183	Cj0183	47.1	255
28	190551	190595	+	Cj0195	fliI	5.1	45
29	190861	191015	+	Cj0195	fliI	13.2	155
30	199591	199620	+	Cj0204	Cj0204	8.1	30
31	202586	202700	+	Cj0206	thrS	17.2	115
32	204411	204455	+	Cj0207	infC	7.3	45
33	205401	205515	+	Cj0208	Cj0208	Undefi-ned*	115
34	213651	213690	+	Cj0229	Cj0229	10.1	40
35	213706	213815	+	Cj0229	Cj0229	7.6	110
36	221171	221395	+	Cj0238	cynT	63.5	225
37	228901	228970	+	Cj0248	Cj0248	5.8	70
38	233226	233280	+	Cj0254	Cj0254	9.9	55

39	248741	248835	+	Cj0270	Cj0270	13.5	95
40	288296	288610	+	Cj0317/Cj0318	hisC/fliF	47.3	315
41	291536	291665	+	Cj0320	fliH	30.8	130
42	291996	292125	+	Cj0320/Cj0321	fliH/dxs	35.3	130
43	298856	298900	+	Cj0327	serC	5.6	45
44	302766	302865	+	Cj0334	ahpC	5.4	100
45	310586	310660	+	Cj0340	Cj0340	6.6	75
46	321186	321245	+	Cj0350/Cj0351	Cj0350/fliN	5.3	60
47	328736	328885	+	Cj0361	lspA	22.4	150
48	339861	339960	+	Cj0371	Cj0371	17.3	100
49	340481	340600	+	Cj0372	Cj0372	48.6	120
50	341156	341280	+	Cj0372/Cj0373	Cj0372/Cj0373	13.2	125
51	342001	342465	+	Cj0373/Cj0374	Cj0373/Cj0374	26.4	465
52	352081	352330	+	Cj0386	engA	60.9	250
53	354626	354800	+	Cj0389	serS	11.1	175
54	355126	355230	+	Cj0389	serS	7.8	105
55	355276	355380	+	Cj0389	serS	6.0	105
56	365196	365335	+	Cj0399/Cj0400	Cj0399/fur	11.0	140
57	368966	368995	+	Cj0403	glyA	7.0	30
58	373296	373620	+	Cj0408/Cj0409	frdC/frdA	12.7	325
59	380896	381010	+	Cj0414	Cj0414	22.8	115
60	383441	383545	+			5.9	105
61	390406	390505	+	Cj0427	Cj0427	6.0	100
62	402796	402865	+	Cj0434	pgm	16.2	70
63	403336	403435	+	Cj0434	pgm	6.5	100
64	403536	403685	+	Cj0434	pgm	8.1	150
65	403781	403980	+	Cj0435	fabG	66.4	200
66	407056	407085	+	Cj0437	sdhA	5.1	30
67	407106	407275	+	Cj0437/Cj0438	sdhA/sdhB	9.3	170
68	427426	427460	+	Cj0462	Cj0462	5.5	35
69	428076	428130	+	Cj0463	Cj0463	5.3	55
70	428331	428355	+	Cj0463	Cj0463	5.0	25
71	430246	430600	+	Cj0464/CJnc22	recG/	24.6	355
72	431591	431730	+	Cj0466/Cj0467	nssR/Cj0467	37.2	140
73	435501	435580	+	Cj0470	tuf	7.7	80
74	435991	436260	+	Cj0472/Cj0473	secE/nusG	13.3	270
75	436611	436745	+	Cj0473/Cj0474	nusG/rplK	9.6	135
76	442201	442385	+	Cj0478	rpoB	16.4	185
77	442626	442665	+	Cj0478	rpoB	7.1	40
78	442731	442780	+	Cj0478	rpoB	11.0	50
79	444066	444265	+	Cj0479	rpoC	29.3	200
80	445761	445865	+	Cj0479	rpoC	11.9	105
81	463881	464035	+	Cj0499	Cj0499	24.3	155
82	472836	473105	+	Cj0508	pbpA	91.2	270
83	480451	480560	+	Cj0514	purQ	8.9	110

84	494041	494140	+	Cj0531	icd	7.9	100
85	495116	495230	+	Cj0531	icd	14.1	115
86	510056	511000	+	Cj0547/Cj0548	flaG/fliD	74.2	945
87	512056	512185	+	Cj0548	fliD	7.5	130
88	512966	512995	+	Cj0550	fliS	5.7	30
89	515701	515835	+	Cj0553	Cj0553	17.4	135
90	521706	521730	+	Cj0559	Cj0559	5.3	25
91	535416	535730	+	Cj0574	ilvI	13.8	315
92	536011	536085	+	Cj0574	ilvI	5.9	75
93	543076	543100	+	Cj0583	Cj0583	5.3	25
94	564066	564365	+	Cj0605	Cj0605	31.4	300
95	567101	567140	+	Cj0607	Cj0607	32.0	40
96	588356	588730	+	Cj0628	Cj0628	13.8	375
97	589531	589630	+	Cj0628	Cj0628	31.2	100
98	604626	604660	+	Cj0642	recN	22.2	35
99	612076	612125	+	Cj0652	pbpC	29.1	50
100	623041	623185	+	Cj0668	Cj0668	9.2	145
101	624476	624790	+	Cj0670	rpoN	97.1	315
102	651001	651190	+	Cj0694	Cj0694	15.3	190
103	655156	655850	+	Cj0697	flgG2	56.4	695
104	655861	655925	+	Cj0697	flgG2	6.4	65
105	666346	666405	+	Cj0709	ffh	5.7	60
106	667616	667710	+	Cj0711/Cj0712	Cj0711/rimM	5.5	95
107	674096	674830	+	Cj0718	dnaE	67.4	735
108	707786	707930	+	Cj0757	hrcA	6.6	145
109	732131	732220	+	Cj0780	napA	5.7	90
110	758666	758790	+	Cj0805/Cj0806	pyrD/dapA	9.4	125
111	758896	758945	+	Cj0806	dapA	5.2	50
112	759511	759685	+	Cj0806/Cj0807	dapA/dapA	9.6	175
113	759986	760020	+	Cj0807	dapA	5.1	35
114	767621	767735	+	Cj0818	Cj0818	29.4	115
115	771826	772090	+	Cj0823/Cj0824	dfp/uppS	63.1	265
116	789906	790100	+	Cj0842	Cj0842	56.9	195
117	806721	806840	+	Cj0860	Cj0860	39.9	120
118	862461	862510	+	Cj0929	pepA	7.7	50
119	911701	911960	+	Cj0977	Cj0977	26.8	260
120	927461	927545	+	Cj0996	ribA	7.1	85
121	1001176	1001375	+	Cj1064	Cj1064	25.5	200
122	1005706	1005850	+	Cj1071	ssb	20.9	145
123	1006076	1006105	+	Cj1071/Cj1072	ssb/rpsR	13.1	30
124	1010626	1010750	+	Cj1076/Cj1077	proC/ctsT	30.9	125
125	1040991	1041040	+	Cj1108	clpA	7.9	50
126	1072876	1072970	+	Cj1138	Cj1138	42.5	95
127	1089211	1089415	+	Cj1156	rho	21.1	205
128	1121066	1121160	+	Cj1192	dctA	6.1	95

129	1138696	1138780	+	Cj1209	Cj1209	6.9	85
130	1162076	1162195	+	Cj1234	glyS	9.8	120
131	1171651	1171775	+	Cj1243	hemE	6.6	125
132	1176501	1176635	+	Cj1248	guaA	27.7	135
133	1177826	1178225	+	Cj1248/Cj1249	guaA/Cj1249	86.6	400
134	1224801	1225100	+	Cj1293	pseB	36.7	300
135	1226926	1227110	+	Cj1294/Cj1295	pseC/pseC	15.4	185
136	1231061	1231110	+	Cj1301	Cj1301	6.0	50
137	1246176	1246335	+	Cj1317	pseI	41.2	160
138	1248756	1248840	+	Cj1319	Cj1319	24.2	85
139	1250886	1250945	+	Cj1321	Cj1321	13.3	60
140	1262766	1262825	+	Cj1334	maf3	12.2	60
141	1262851	1262940	+	Cj1334	maf3	10.2	90
142	1266431	1266565	+	Cj1337	pseE	31.5	135
143	1319521	1319635	+	Cj1380	Cj1380	8.1	115
144	1329231	1329305	+	Cj1394	Cj1394	3763.6	75
145	1398421	1399005	+	Cj1462	flgI	93.8	585
146	1399756	1399945	+	Cj1463/Cj1464	flgJ/flgM	6.6	190
147	1402066	1402175	+	Cj1466	flgK	9.0	110
148	1452806	1452915	+	Cj1516	Cj1516	8.8	110
149	1474796	1474905	+	Cj1543	Cj1543	7.1	110
150	1528151	1528480	+	Cj1599	hisB	33.7	330
151	1529541	1529580	+	Cj1601	hisA	19.7	40
152	1529581	1529705	+	Cj1601	hisA	31.5	125
153	1546866	1546915	+	Cj1619	kgtP	5.4	50
154	1598601	1598700	+	Cj1677	Cj1677	31.0	100
155	1632201	1632460	+	Cj1720	Cj1720	63.4	260
156	6411	6495	-	Cj0005c	Cj0005c	10.1	85
157	24026	24150	-	Cj0019c	Cj0019c	21.8	125
158	33886	33980	-	Cj0026c	thyX	7.0	95
159	52561	52650	-	Cj0034c	Cj0034c	12.9	90
160	58061	58140	-	Cj0039c	typA	40.4	80
161	85456	85570	-	Cj0074c	Cj0074c	7.5	115
162	87391	87495	-	Cj0076c	lctP	6.2	105
163	88626	88740	-	Cj0076c	lctP	5.9	115
164	90616	90740	-	Cj0079c	cdtA	8.0	125
165	95356	95475	-	Cj0086c	ung	16.8	120
166	152706	152925	-	Cj0149c	hom	9.4	220
167	156766	156885	-	Cj0153c	Cj0153c	8.8	120
168	156911	156950	-	Cj0153c	Cj0153c	4.9	40
169	167766	167995	-	Cj0172c	Cj0172c	56.0	230
170	169041	169130	-	Cj0173c	cfbpC	29.1	90
171	193581	193645	-	Cj0197c	dapB	10.6	65
172	196176	196420	-	Cj0200c	Cj0200c	19.7	245
173	214916	214950	-	Cj0230c	Cj0230c	36.4	35

174	223461	223525	-	Cj0240c	iscS	5.4	65
175	242491	242550	-	Cj0264c	Cj0264c	8.4	60
176	243211	243385	-	Cj0264c	Cj0264c	24.7	175
177	244831	244865	-	Cj0265c	Cj0265c	5.2	35
178	263081	263110	-	Cj0284c	cheA	5.1	30
179	263436	263730	-	Cj0285c	cheV	14.3	295
180	265081	265170	-	Cj0287c	greA	5.9	90
181	281531	281680	-	Cj0309c/Cj0310c	Cj0309c/Cj0310c	136.7	150
182	300451	300575	-	Cj0329c	plsX	43.7	125
183	305681	305770	-	Cj0337c/Cj0338c	motA/polA	7.8	90
184	307746	307865	-	Cj0338c	polA	6.1	120
185	338041	338340	-	Cj0369c	Cj0369c	30.9	300
186	338856	338925	-	Cj0369c	Cj0369c	6.7	70
187	347201	347325	-	Cj0380c	Cj0380c	81.4	125
188	348261	348340	-	Cj0381c	pyrF	45.0	80
189	362021	362130	-	Cj0394c	Cj0394c	87.7	110
190	475256	475310	-	Cj0509c	clpB	64.8	55
191	475641	475800	-	Cj0509c	clpB	69.0	160
192	476286	476470	-	Cj0509c	clpB	74.4	185
193	489256	489325	-	Cj0526c	fliE	5.5	70
194	599571	599665	-	Cj0638c	ppa	8.7	95
195	619356	619470	-	Cj0662c	hslU	26.8	115
196	641256	641315	-	Cj0687c	flgH	6.4	60
197	641371	641845	-	Cj0687c	flgH	39.0	475
198	657496	657570	-	Cj0699c	glnA	5.6	75
199	680321	680455	-	Cj0725c/Cj0726c	mogA/corA	8.4	135
200	713226	713335	-	Cj0762c	aspB	5.3	110
201	713736	713995	-	Cj0762c	aspB	20.1	260
202	714421	714615	-	Cj0763c	cysE	32.0	195
203	719471	719670	-	Cj0768c/Cj0769c	Cj0768c/flgA	63.9	200
204	723991	724120	-	Cj0774c	Cj0774c	16.3	130
205	748996	749110	-	Cj0798c	ddl	28.7	115
206	799686	799755	-	Cj0853c	hemL	6.8	70
207	799791	800215	-	Cj0853c	hemL	31.9	425
208	823526	823785	-	Cj0887c	flgL	15.7	260
209	833091	833185	-	Cj0893c	rpsA	10.4	95
210	833526	833570	-	Cj0893c	rpsA	5.6	45
211	839916	839990	-	Cj0899c	thiJ	5.8	75
212	840081	840235	-	Cj0899c	thiJ	9.1	155
213	846076	846195	-	Cj0906c	Cj0906c	6.7	120
214	853956	854055	-	Cj0917c	cstA	13.6	100
215	854076	854185	-	Cj0917c	cstA	12.8	110
216	890711	890835	-	Cj0950c/Cj0951c	Cj0950c/Cj0951c	26.1	125
217	891696	891790	-	Cj0951c	Cj0951c	25.0	95
218	922746	922860	-	Cj0991c	Cj0991c	28.1	115

219	923591	923745	-	Cj0992c	hemN	7.4	155
220	926346	926705	-	Cj0995c	hemB	18.5	360
221	936571	936810	-	Cj1006c	Cj1006c	38.6	240
222	938551	938600	-	Cj1007c	Cj1007c	10.5	50
223	945901	945990	-	Cj1013c	Cj1013c	46.8	90
224	954866	954985	-	Cj1023c	asd	15.6	120
225	978051	978135	-	Cj1044c	thiH	8.8	85
226	981166	981365	-	Cj1048c	dapE	11.1	200
227	981371	981455	-	Cj1048c	dapE	6.2	85
228	993961	994335	-	Cj1058c	guaB	7.0	375
229	994336	994450	-	Cj1058c	guaB	6.3	115
230	994556	994665	-	Cj1058c	guaB	8.3	110
231	1008336	1008590	-	Cj1073c	lon	27.3	255
232	1014751	1014915	-	Cj1083c/Cj1084c	thiD/Cj1084c	36.0	165
233	1017031	1017130	-	Cj1085c	mfd	16.9	100
234	1026501	1026690	-	Cj1093c	secD	26.2	190
235	1030401	1030655	-	Cj1096c	metK	7.8	255
236	1045241	1045480	-	Cj1112c	Cj1112c	8.4	240
237	1049491	1049570	-	Cj1116c	ftsH	5.7	80
238	1049591	1049615	-	Cj1116c	ftsH	5.1	25
239	1050436	1050550	-	Cj1117c	prmA	11.5	115
240	1050556	1050715	-	Cj1117c/Cj1118c	prmA/cheY	7.4	160
241	1050961	1051030	-	Cj1118c	cheY	5.1	70
242	1056631	1056885	-	Cj1124c/Cj1125c	pglC/pglA	95.4	255
243	1085006	1085265	-	Cj1152c	gmhB	33.4	260
244	1086376	1086615	-	Cj1155c	Cj1155c	2001.2	240
245	1087736	1087870	-	Cj1155c	Cj1155c	49.8	135
246	1094181	1094300	-	Cj1163c	Cj1163c	23.6	120
247	1103356	1103405	-	Cj1175c	argS	6.5	50
248	1103456	1103505	-	Cj1175c	argS	12.7	50
249	1107826	1107990	-	Cj1181c	tsf	47.4	165
250	1116776	1116995	-	Cj1188c	gidA	25.6	220
251	1124266	1124310	-	Cj1195c	pyrC2	6.1	45
252	1126341	1126470	-	Cj1197c	gatB	14.1	130
253	1135551	1135880	-	Cj1206c	ftsY	170.4	330
254	1136261	1136305	-	Cj1206c	ftsY	6.4	45
255	1136381	1136470	-	Cj1207c	ftsY	13.2	90
256	1136801	1136915	-	Cj1207c	ftsY	9.3	115
257	1146136	1146260	-	Cj1218c	ribA	38.1	125
258	1151111	1151445	-	Cj1222c	dccS	384.5	335
259	1198156	1198405	-	Cj1267c	hydA	19.2	250
260	1202311	1202550	-	Cj1270c	amiA	16.6	240
261	1203251	1203350	-	Cj1270c	amiA	7.9	100
262	1206456	1206545	-	Cj1272c	spoT	11.1	90
263	1206576	1206750	-	Cj1272c	spoT	21.5	175

264	1207501	1207530	-	Cj1274c	pyrH	5.2	30
265	1208106	1208210	-	Cj1275c	Cj1275c	42.7	105
266	1210326	1210475	-	Cj1277c	Cj1277c	47.8	150
267	1220851	1221020	-	Cj1288c	gltX	6.5	170
268	1222501	1222595	-	Cj1290c	accC	7.1	95
269	1239036	1239090	-	Cj1309c	Cj1309c	10.3	55
270	1239806	1239940	-	Cj1310c	Cj1310c	12.5	135
271	1267611	1268275	-	Cj1338c	flaB	9.9	665
272	1268341	1268540	-	Cj1338c	flaB	8.1	200
273	1268541	1269135	-	Cj1338c	flaB	34.0	595
274	1269156	1271035	-	Cj1339c	flaA	179.4	1880
275	1279901	1280120	-	Cj1346c/Cj1347c	dxr/cdsA	19.7	220
276	1288416	1288445	-	Cj1356c	Cj1356c	5.1	30
277	1300506	1300640	-	Cj1365c	Cj1365c	11.7	135
278	1304471	1304590	-	Cj1367c	Cj1367c	32.3	120
279	1338301	1338385	-	Cj1403c	gapA	5.8	85
280	1339636	1339745	-	Cj1406c	Cj1406c	40.6	110
281	1342231	1342300	-	Cj1410c	Cj1410c	9.3	70
282	1342301	1342325	-	Cj1410c	Cj1410c	5.0	25
283	1345366	1345500	-	Cj1413c	kpsS	38.5	135
284	1349566	1349695	-	Cj1416c/Cj1417c	cysC/Cj1417c	58.1	130
285	1353296	1353515	-	Cj1419c/Cj1420c	Cj1419c/Cj1420c	38.4	220
286	1353641	1353735	-	Cj1420c	Cj1420c	5.3	95
287	1353836	1353890	-	Cj1420c	Cj1420c	5.2	55
288	1366296	1366565	-	Cj1431c/Cj1432c	hddC/Cj1432c	157.7	270
289	1368496	1368540	-	Cj1432c	Cj1432c	21.6	45
290	1368551	1368750	-	Cj1432c	Cj1432c	17.6	200
291	1368786	1368840	-	Cj1432c	Cj1432c	5.8	55
292	1368966	1369035	-	Cj1432c	Cj1432c	14.0	70
293	1369066	1369435	-	Cj1432c	Cj1432c	80.9	370
294	1373841	1374070	-	Cj1436c/Cj1437c	Cj1436c/Cj1437c	115.2	230
295	1377561	1377695	-	Cj1439c	glf	26.6	135
296	1386021	1386160	-	Cj1445c	kpsE	16.0	140
297	1386206	1386350	-	Cj1445c/Cj1447c	kpsE/kpsT	14.9	145
298	1413111	1413180	-	Cj1476c	Cj1476c	5.8	70
299	1418586	1418705	-	Cj1481c/Cj1482c	Cj1481c/Cj1482c	56.8	120
300	1423261	1423340	-	Cj1487c/Cj1488c	ccoP/ccoQ	12.2	80
301	1423511	1423605	-	Cj1488c/Cj1489c	ccoQ/ccoO	12.2	95
302	1427991	1428030	-	Cj1493c	Cj1493c	70.1	40
303	1435441	1435605	-	Cj1502c/Cj1503c	putP/putA	12.1	165
304	1438356	1438385	-	Cj1503c	putA	5.1	30
305	1445121	1445185	-	Cj1509c	fdhC	5.9	65
306	1447781	1447820	-	Cj1511c	fdhA	5.0	40
307	1457156	1457285	-	Cj1523c	Cj1523c	54.7	130
308	1457316	1457455	-	Cj1523c	Cj1523c	71.6	140

309	1468396	1468475	-	Cj1537c	acs	5.7	80
310	1469846	1469965	-	Cj1537c	acs	11.2	120
311	1470201	1470335	-	Cj1537c	acs	7.3	135
312	1485996	1486050	-	Cj1552c	mloB	11.2	55
313	1489281	1489415	-	Cj1555c	Cj1555c	64.5	135
314	1501851	1501915	-	Cj1570c/Cj1571c	nuoJ/nuoI	5.3	65
315	1501946	1502085	-	Cj1571c	nuoI	6.2	140
316	1540281	1540410	-	Cj1613c	Cj1613c	40.6	130
317	1560391	1560475	-	Cj1634c	aroC	17.8	85
318	1579431	1579480	-	Cj1655c	nhaA1	5.6	50
319	1579516	1579705	-	Cj1655c	nhaA1	203.8	190
320	1594706	1594765	-	Cj1673c	recA	6.2	60
321	1603981	1604100	-	Cj1682c	gltA	6.6	120
322	1605971	1606035	-	Cj1684c	Cj1684c	12.6	65
323	1612716	1612795	-	Cj1690c	rpsE	15.3	80
324	1623991	1624125	-	Cj1711c	ksgA	20.6	135
325	1639411	1639485	-	Cj1729c	flgE	7.5	75
326	1639521	1640140	-	Cj1729c	flgE	15.7	620
327	1640151	1640285	-	Cj1729c	flgE	5.9	135
328	1640311	1640795	-	Cj1729c	flgE	203.7	485

***** Undefined value for enrichment factor as no reads are mapped to this peak in the control coIP library

Strain 81-176							
Peak ID	**Start**	**End**	**Strand**	**Overlapping genes (locus tags)**	**Overlapping genes (gene names)**	**Enrichm-ent factor**	**Peak Size (bp)**
1	29546	29755	+	CJJ81176_0050/ CJJ81176_0051	purB-1/nrdA	22.4	210
2	69726	70100	+	CJJ81176_0078	CJJ81176_0078	26.1	375
3	71966	72050	+	CJJ81176_0080	flgD	7.3	85
4	126221	126515	+	CJJ81176_0143/ CJJ81176_0144	atpC/ CJJ81176_0144	11.7	295
5	149136	149335	+	CJJ81176_0171	infB	10.9	200
6	185951	186230	+	CJJ81176_0214	CJJ81176_0214	24.9	280
7	210096	210135	+	CJJ81176_0239	infC	5.7	40
8	211091	211135	+	CJJ81176_0240	CJJ81176_0240	252.6	45
9	292316	292695	+	CJJ81176_0339/ CJJ81176_0340	hisC/fliF	34.2	380
10	295611	295635	+	CJJ81176_0342	fliH	5.1	25
11	332796	332830	+	CJJ81176_0384	lspA	7.3	35
12	343891	343945	+	CJJ81176_0395	CJJ81176_0395	5.3	55
13	344511	344630	+	CJJ81176_0396	CJJ81176_0396	10.9	120
14	346206	346295	+	CJJ81176_0398	CJJ81176_0398	5.7	90
15	405546	405665	+	CJJ81176_0461	fabG	7.0	120
16	437191	437240	+	CJJ81176_0499	tuf	5.8	50
17	437826	437930	+	CJJ81176_0503	nusG	9.4	105
18	470286	470480	+	CJJ81176_0536	pbpA	17.4	195

19	498336	498430	+	CJJ81176_0562	oorB	11.4	95
20	507341	507385	+	CJJ81176_0571	CJJ81176_0571	12.8	45
21	507416	508440	+	CJJ81176_0572/ CJJ81176_0573	flaG/fliD	59.4	1025
22	508551	508595	+	CJJ81176_0573	fliD	5.1	45
23	564436	564520	+	CJJ81176_0636/ CJJ81176_0637	CJJ81176_0636/ CJJ81176_0637	12.6	85
24	596896	596960	+	CJJ81176_0669	pnk	12.5	65
25	620216	620535	+	CJJ81176_0696	rpoN	31.4	320
26	646741	646855	+	CJJ81176_0717	CJJ81176_0717	5.7	115
27	650891	651725	+	CJJ81176_0720	CJJ81176_0720	80.3	835
28	669816	669900	+	CJJ81176_0741	dnaE	7.1	85
29	669926	670665	+	CJJ81176_0741	dnaE	39.5	740
30	670666	670705	+	CJJ81176_0741	dnaE	8.1	40
31	698831	698870	+	CJJ81176_0772	CJJ81176_0772	11.5	40
32	701251	701345	+	CJJ81176_0774	grpE	7.8	95
33	742896	742965	+	CJJ81176_0813	CJJ81176_0813	110.1	70
34	758726	759240	+	CJJ81176_0826/ CJJ81176_0827	CJJ81176_0826/ dapA	29.7	515
35	771541	771640	+	CJJ81176_0840/ CJJ81176_0841	CJJ81176_0840/ uppS	10.1	100
36	789551	789635	+			16.1	85
37	813451	813480	+	CJJ81176_0882	CJJ81176_0882	11.2	30
38	918506	918600	+	CJJ81176_0996	CJJ81176_0996	9.8	95
39	1008321	1008485	+	CJJ81176_1089	ssb	9.7	165
40	1179561	1180085	+	CJJ81176_1264/ CJJ81176_1265	guaA/ CJJ81176_1265	169.3	525
41	1226461	1226730	+	CJJ81176_1310	flmA	16.4	270
42	1376481	1377125	+	CJJ81176_1454/ CJJ81176_1455	CJJ81176_1454/flgI	112.3	645
43	1504806	1505100	+	CJJ81176_1586	hisB	11.5	295
44	1505121	1505170	+	CJJ81176_1586	hisB	5.2	50
45	1607966	1608125	+	CJJ81176_0018	CJJ81176_0018	17.7	160
46	178636	178790	-	CJJ81176_0208	CJJ81176_0208	12.4	155
47	188296	188340	-	CJJ81176_0217	CJJ81176_0217	13.0	45
48	201921	202005	-	CJJ81176_0231	CJJ81176_0231	6.5	85
49	244231	244270	-	CJJ81176_0288	CJJ81176_0288	15.7	40
50	244271	244325	-	CJJ81176_0288	CJJ81176_0288	10.8	55
51	269121	269295	-	CJJ81176_0311	cheV	10.5	175
52	285586	285700	-	CJJ81176_0331/ CJJ81176_0332	CJJ81176_0331/ CJJ81176_0332	17.4	115
53	339641	339950	-	CJJ81176_0389/ CJJ81176_0390	cmeB/cmeA	44.5	310
54	339956	339985	-	CJJ81176_0390	cmeA	14.1	30
55	342096	342185	-	CJJ81176_0392	CJJ81176_0392	6.7	90
56	351346	351390	-	CJJ81176_0404	pyrF	9.1	45
57	472976	473035	-	CJJ81176_0537	clpB	26.5	60
58	637021	637580	-	CJJ81176_0710	flgH	58.2	560

59	653866	653890	-	CJJ81176_0722	glnA	5.0	25
60	713871	713960	-	CJJ81176_0783	aspC	8.6	90
61	717211	717245	-	CJJ81176_0786	hisS	5.6	35
62	719636	719730	-	CJJ81176_0790	flgA	17.9	95
63	748826	748860	-	CJJ81176_0819	ddl	6.8	35
64	748911	749390	-	CJJ81176_0819	ddl	20.9	480
65	786586	786710	-	CJJ81176_0855	metS	7.8	125
66	786776	786880	-	CJJ81176_0855	metS	7.4	105
67	799496	799685	-	CJJ81176_0869	hemL	12.2	190
68	826406	826430	-			5.3	25
69	854351	854515	-	CJJ81176_0924	cstA	5.8	165
70	943256	943420	-	CJJ81176_1024	CJJ81176_1024	19.9	165
71	1029156	1029190	-	CJJ81176_1111	secD	30.1	35
72	1058601	1058780	-	CJJ81176_1141	pglC	34.4	180
73	1085431	1085495	-	CJJ81176_1167/ CJJ81176_1168	hldE/waaD	7.6	65
74	1088096	1088220	-	CJJ81176_1172	CJJ81176_1172	58.5	125
75	1109356	1109455	-	CJJ81176_1195/ CJJ81176_1196	CJJ81176_1195/tsf	6.6	100
76	1109481	1109660	-	CJJ81176_1196	tsf	20.4	180
77	1118441	1118530	-	CJJ81176_1203	gidA	22.8	90
78	1137176	1137405	-	CJJ81176_1221	ftsY	49.4	230
79	1152771	1153100	-	CJJ81176_1235	CJJ81176_1235	126.5	330
80	1253526	1254205	-	CJJ81176_1338	CJJ81176_1338	6.8	680
81	1254276	1255075	-	CJJ81176_1338	CJJ81176_1338	29.9	800
82	1255096	1257205	-	CJJ81176_1339/ CJJ81176_1340	CJJ81176_1339/ CJJ81176_1340	92.1	2110
83	1332456	1332515	-	CJJ81176_1411	CJJ81176_1411	6.0	60
84	1355681	1355735	-	CJJ81176_1434	CJJ81176_1434	6.0	55
85	1355746	1355800	-	CJJ81176_1434	CJJ81176_1434	5.2	55
86	1453816	1453840	-	CJJ81176_1533	CJJ81176_1533	8.8	25
87	1464046	1464070	-	CJJ81176_1541	CJJ81176_1541	593.0	25
88	1522736	1522830	-	CJJ81176_1605	CJJ81176_1605	10.4	95
89	1555996	1556280	-	CJJ81176_1646	nhaA-2	27.6	285
90	1615406	1615470	-	CJJ81176_0025	flgE	5.0	65
91	1615471	1615875	-	CJJ81176_0025	flgE	47.9	405
pVir plasmid							
92	20986	21060	+	CJJ81176_pVir0028/ CJJ81176_pVir0029	CJJ81176_pVir0028/ CJJ81176_pVir0029	18.3	75
93	24316	24550	+	CJJ81176_pVir0032	CJJ81176_pVir0032	42.7	235
94	1831	1855	-			5.2	25
95	33111	33200	-	CJJ81176_pVir0049	CJJ81176_pVir0049	26.4	90
96	33221	33300	-	CJJ81176_pVir0049	CJJ81176_pVir0049	11.3	80
97	33351	33410	-	CJJ81176_pVir0049	CJJ81176_pVir0049	207.7	60
pTet plasmid							
98	41516	41645	+	CJJ81176_pTet0048	CJJ81176_pTet0048	23.3	130

Appendix Table 12. DNA oligonucleotides. List of all DNA oligonucleotides used in this study for PCR amplification, Northern blot hybridization, and FISH assays. DNA sequences are given in 5' to 3' direction; P- denotes a 5' monophosphate.

Name	**Sequence** (5' → 3')	**Description**
CSO-0023	CCACCAGCTTATATACCTTAGCA	Antisense to *aphA-3* for verification
CSO-0073	CTAACAAGCTTTCATCTACGCA	3xFLAG Tagging using pGG1
CSO-0074	GTTTTTGAATTCTATTCCCTCCAGGTACTAAAACA	3xFLAG Tagging using pGG1
CSO-0075	TCCTTCACAAAGAAGGGG	3xFLAG Tagging using pGG1
CSO-0171	P-TTTGATTAGTTTTTTGCTTAAGTCAT	Cloning of *csrA*-3xFLAG in pGG1
CSO-0172	GTTTTTCTCGAGCTCTTTAGAGCGCATTAAAGAA	Cloning of *csrA* -3xFLAG in pGG1
CSO-0173	GTTTTTTCTAGACAAGATATTTGTGGAAAAGTCC	Cloning of *csrA*-3xFLAG in pGG1
CSO-0174	GTTTTTGAATTCATCAAATGAAAGCTTACGCTAA	Cloning of *csrA*-3xFLAG in pGG1
CSO-0180	CAAAAATGGGCAAATTGCAA	Northern blot probe CJnc10
CSO-0181	TTTTCTTATGAATGTTTTTCGGTTC	Northern blot probe CJnc140
CSO-0182	ATTTTCATGGGCTTGATTGCA	Northern blot probe CJnc170
CSO-0183	TCAGGATGAAATTTTTAACAAGTAGG	Northern blot probe CJnc30
CSO-0184	CGATGTGACATTTTATTCTAAAAACAC	Northern blot probe CJnc190 (full length)
CSO-0185	GGAAATAGCCTAACCCAAACG	Northern blot probe CJnc190 (processed)
CSO-0187	CATTTCGATCTCGTTGAGTTTTAAG	Northern blot probe CJnc20
CSO-0188	CCTTTTTATTCGTAGCCTAACTCTCT	Northern blot probe CJnc180
CSO-0190	GCTAGGTAATTTTTAAAGGCTGCA	Northern blot probe 6S RNA (CJnc130)
CSO-0191	AAGTGCTCTGCTGTGTTCCCA	Northern blot probe SRP RNA
CSO-0192	CTACTTTCCCCCTGCCAGTAA	Northern blot probe 5S rRNA
CSO-0193	TTTTACTTTCATCATTTCATCCTTTG	Northern blot probe CJnc60
CSO-0196	GTATTTGATTGCAAGATCTTAAGC	Verification of *csrA*-3xFLAG in *C. jejuni*
CSO-0212	CCTCGAATTTAAAAGGTGGG	Northern blot probe CJpt1
CSO-0213	TGGATATAATTTTGTGGGGTTAGC	Northern blot probe CJpt2
CSO-0214	TCAAATCACACTACAAGGAGGTG	Northern blot probe CJpt3
CSO-0215	AGCTTGACAAATAAAGGGTTAAGG	Northern blot probe CJpt4
CSO-0216	GTTTATTTAAGAATACCTTGCCGC	Northern blot probe CJpt5
CSO-0217	ACAAAGACAAAGGATTAAAGATGGA	Northern blot probe CJpv1
CSO-0218	CAAGTGTTTTATGGTTATGGGGT	Northern blot probe CJpv2
CSO-0223	AGCGGTTTTAGGGGATTGTAA	Northern blot probe Tracr RNA
CSO-0224	AGGGACTAAAACAGTTAAATTCAACTAA	Northern blot probe crRNA2 (NCTC11168)
CSO-0240	TATCACTGTCCTGTTTGTTATGAG	Sense aligo binds to upstream of upstream region of *rnc*
CSO-0241	TCCTAGTTAGTCACCCGGGTACAACTTTTATGGGTTAAA GCATGT	Antisense oligo for amplification of upstream region of *rnc*; incl. complementary region to HPK1
CSO-0242	GTGAAAAACAATGAAACACATTG	Sense oligo for amplification of upstream region of *rnc*
CSO-0243	TCTTCTTCTTTCAAATTTGATACAA	Antisense oligo for amplification of downstream region of *rnc*
CSO-0244	AATTGTTTTAGTACCTGGAGGGAATACAAAAAAGAAGCA CAGCAAAT	Sense oligo for amplification of downstream region of *rnc* incl. complementary region to entire HPK2
CSO-0270	AGGGACTAAAACATTTAAGTCCATTT	Northern blot probe crRNA4 (NCTC11168)
CSO-0296	CTAAAAGGGACGAACAACCATC	Northern blot probe CJas_Cj0363c
CSO-0297	GAAAGTATGTTTCCACAAGCCTTA	Northern blot probe CJas_Cj0566

CSO-0298	CGCTTAGGTGCTTATTATCAATTTC	Northern blot probe CJas_Cj0704
CSO-0299	ACTCTTGGCGGAAATCCAC	Northern blot probe CJas_Cj1667c
CSO-0300	ATGCCCAGCCTTGATATAAGTC	Northern blot probe CJas_Cj0168c
CSO-0302	CCATCTCTTTCTACCACCGG	Northern blot probe CJnc100
CSO-0303	AGATCATTGGATTTTTAGTTCTTGTC	Northern blot probe CJnc80
CSO-0392	TACTCCTTAAGTCTTGATGATCAA	Verification of *csrA* deletion in *C. jejuni*
CSO-0393	TCCTAGTTAGTCACCCGGGTACCTTGATAATATTAACAT TTTTCAACCT	Deletion of *csrA* using *hyg* in *C. jejuni*
CSO-0394	TGCAAGGAATTATCTCCTATACAC	Deletion of *csrA* using *hyg/cat* in *C. jejuni*
CSO-0395	ATCATAAACAGCTTTAGTTTGGC	Deletion of *csrA* using *hyg/cat* in *C. jejuni*
CSO-0396	ATTGTTTTAGTACCTGGAGGGAATAGCAAAAAACTAATC AAATGAAAG	Deletion of *csrA* using *hyg* in *C. jejuni*
CSO-0397	CCTACGGCTAAATTGTGAAACC	Northern blot probe CJnc90
CSO-0398	CTTGCAACCGAAAGAGTAGCC	Northern blot probe CJnc70
CSO-0399	GCCAATTCTTCGCTAAACTTTCA	Northern blot probe CJpv3
CSO-0483	GTTTTTGGATCCTTTTATGGATAATTTTTAAAATCATTTG	Cloning of *aac*(3)-IV upstream of *flaA* 5' UTR
CSO-0486	GTGTTAATACGAAATCCCATTTTAAATC	NB detection *flaA* mRNA (Binds 5' UTR)
CSO-0496	GCACATCAGTTTCATTTTTTCTCC	Northern blot probe CJnc110
CSO-0497	GTTAGATGGAATGTTTGTCTTTAGCAG	Northern blot probe RnpB
CSO-0498	TTGGGCTTTTTACGCTTTGC	Northern blot probe tmRNA
CSO-0514	CGCTAAAACAGCGAAAAAAATAGC	Northern blot probe CJas_CJJ81176_1020
CSO-0525	AAGCCCTACACGACCTACACG	Northern blot probe CJnc22
CSO-0526	CCCATTCTAAACGACGCGAC	Northern blot probe CJnc21
CSO-0527	TACATCATTCCGCCCATGC	Northern blot probe CJnc120
CSO-0536	AGCTTTATTTACAAGAGCATCTCTGAT	Northern blot probe CJnc11
CSO-0537	AGGCAGTTCAAGCCTAAGCC	Northern blot probe CJnc230
CSO-0553	P-CTGTAGTAATCTTAAAACATTTTGTTGA	Cloning of *flaA*-3xFLAG in pGG1
CSO-0554	GTTTTTCTCGAGTGGTTATTCTTCTGTTAGTGCCT	Cloning of *flaA*-3xFLAG in pGG1
CSO-0555	GTTTTTTCTAGAGCGATATTGTCAAGTTCTTCC	Cloning of *flaA*-3xFLAG in pGG1
CSO-0556	GTTTTTGAATTCTTTACAAAAGCTGCAATATATACAAA	Cloning of *flaA*-3xFLAG in pGG1
CSO-0557	CTCTCAAGCTTCTGTTTCTTTAAG	Verification of *flaA*-3xFLAG in *C. jejuni*
CSO-0558	P-TTGAAGAAGTTTTAAAACATTTTGC	Cloning of *flaB*-3xFLAG in pGG1
CSO-0559	GTTTTTCTCGAGTTAGTGCCTATATGAGTAGCGC	Cloning of *flaB*-3xFLAG in pGG1
CSO-0560	GTTTTTTCTAGAGTGCTAGGATAGAAAGCGCT	Cloning of *flaB*-3xFLAG in pGG1 / Overlap PCR construction of *flaB* deletion with *aphA-3*
CSO-0561	GTTTTTGAATTCTTTCTTAGATGCTTTTATGCATCT	Cloning of *flaB*-3xFLAG in pGG1
CSO-0562	GATGCTAATATCGCTGATGC	Verification of *flaB*-3xFLAG
CSO-0575	CAATACGAATGGCGAAAAG	*aac*(3)-IV cloning in pGG1
CSO-0576	GTTTTTCATATGAAACACCCCCATAAGTGCAATTATGGG GATAAATCATCTCGTTCTCCGCTC	Cloning of *aac*(3)-IV upstream of *flaA* 5' UTR
CSO-0577	P-CATTTATTCCTCCTAGTTAGTCACC	*aac*(3)-IV cloning in pGG1
CSO-0606	GTTTTTATGCATTTTATTCAAGAAAATTCAACTACGG	Cj0805-Cj0806 (*dapA*) cloning in pXG30
CSO-0607	GTTTTTGCTAGCTTGCTCATCAACTTTTCCAT	Cj0805-Cj0806 (*dapA*) cloning in pXG30
CSO-0608	GTTTTTATGCATGCAATTTTACTTTTAAGTATTATAGCCC	Cj0310c-Cj0309c cloning in pXG30
CSO-0609	GTTTTTGCTAGCAAGTTCTTTCATGATCACCACG	Cj0310c-Cj0309c cloning in pXG30
CSO-0611	TACAGAGAGACCCGACTCTTTTAATCTTTCAAGGAGCAA AGAATGGTGTAGGCTGGAGCTGCTTC	Deletion of *csrA* in *E. coli* (using λ-red system)

CSO-0612	TTTGAGGGTGCGTCTCACCGATAAAGATGAGACGCGGAAAGATTAGGTCCATATGAATATCCTCCTTAG	Deletion of *csrA* in *E. coli* (using λ-red system)
CSO-0613	AACAAATCGGAATTTACGGA	Amplification of *C. coli cat* cassette
CSO-0614	GGCACCAATAACTGCCTTAA	Amplification of *C. coli cat* cassette
CSO-0615	CTCCGTAAATTCCGATTTGTTCTTGATAATATTAACATTTTTTCAACCT	Deletion of *csrA* using *cat* in *C. jejuni*
CSO-0616	TTTTAAGGCAGTTATTGGTGCCGCAAAAAACTAATCAAATGAAAG	Deletion of *csrA* using *cat* in *C. jejuni*
CSO-0621	GTTTTTATGCATTAACAAGTTCATGGATGAGCTT	*flaA* cloning in pXG10
CSO-0622	GTTTTTGCTAGCACTAAGTCTGCTTAAAGAAGCATC	*flaA* cloning in pXG10
CSO-0639	GATGTAATGTGTTTGTCATTGCT	Verification of *csrA* deletion in *E. coli*
CSO-0640	GAGACTTAAGTTGAATGAACGG	Verification of *csrA* deletion in *E. coli*
CSO-0652	AGATACAGAGAGAGATTTTGGCAATACATGGAGTAATACAGGATGGTGTAGGCTGGAGCTGCTTC	Deletion of *pgaA* in *E.coli* (using λ-red system)
CSO-0653	GCATCAGGAGATATTTATTTCCATTACGTAACATATTTATCCTTAGGTCCATATGAATATCCTCCTTAG	Deletion of *pgaA* in *E.coli* (using λ-red system)
CSO-0654	TCTCTCTTCCGCGTTTAATAAC	Verification of *pgaA* deletion in *E.coli*
CSO-0655	CTGTGGCGGTATAAATGATG	Verification of *pgaA* deletion in *E.coli*
CSO-0694	GTTTTTATGCATACAATAGATTAAAGGAAGAATCCAT	*flgI* cloning in pXG10
CSO-0695	GTTTTTGCTAGCACCTATAAGTTGGTTATCTCTTACACC	*flgI* cloning in pXG10
CSO-0701	GTTTTTTTTAATACGACTCACTATAGGAAAGCTGGTGGCGCTG	*in vitro* transcription of *hopB 3'*end, carries T7 promoter
CSO-0702	GTTAAATCAAAGCCTATAAAAGGCC	*in vitro* transcription of *hopB 3'*end
CSO-0709	GTTTTTTTTAATACGACTCACTATAGGTAACAAGTTCATGGATGAGCTT	*in vitro* transcription of *flaA* leader, carries T7 promoter
CSO-0710	ACTAAGTCTGCTTAAAGAAGCATCT	*in vitro* transcription of *flaA* leader
CSO-0713	GTTTTTTTTAATACGACTCACTATAGGACAATAGATTAAAGGAAGAATCCAT	*in vitro* transcription of *flgI* leader, carries T7 promoter
CSO-0714	ACCTATAAGTTGGTTATCTCTTACACC	*in vitro* transcription of *flgI* leader
CSO-0748	P-TAACAAGTTCATGGATGAGCTT	*flaA* leader cloning in pBAD plasmid
CSO-0749	GTTTTTTCTAGAGTTTGCTTTTGCATTTAAAGCT	*flaA* leader cloning in pBAD plasmid
CSO-0752	GTTTTTCTCGAGAAGGTGGAGCAAGGATTAA	*flaA* cloning in pJV752.1 / Overlap PCR construction of *flaA* deletion with *aphA-3*
CSO-0753	GTTTTTTCTAGATCTTAGAAGATTGAGTTGCTCC	*flaA* cloning in pJV752.1
CSO-0754	GTTTTTCATATGCAATAAAATTTCATACTTTTGACA	Cloning of *aac*(3)-IV upstream of *flaA* 5' UTR
CSO-0755	GTTTTTGGATCCTAAAGTATAAAATATTTTTTTGATTGCA	Cloning of *aac*(3)-IV upstream of *flaA* 5' UTR
CSO-0756	TATGCAGGCAAAGGTGAAG	Verification of *flaA* deletion in *C. jejuni*
CSO-0757	TAACAAGTTCATTGATGAGCTTGAATTTTTTTAAAAG	Introduction of SL1$^{GGA>UGA}$ (M2) mutation into *flaA* 5' UTR
CSO-0758	TTCAAGCTCATCAATGAACTTGTTAAATGCTATATCGT	Introduction of SL1$^{GGA>UGA}$ (M2) mutation into *flaA* 5' UTR
CSO-0831	GTTTTTCATATGTGTTTTAGTACCTGGAGGGAATA	*aac*(*3*)-*IV* cloning in pGG1
CSO-0832	GTTTTTCATATGTCATCTCGTTCTCCGCTC	*aac*(*3*)-*IV* cloning in pGG1
CSO-0852	CGGGTGGCTCCATTTGATTAGTTTTTTGCTTAAGT	Addition of Strep-tag to *csrA* in pBAD plasmid
CSO-0853	P-CAGTTCGAAAAATGAAAGCTTACGCTCTAGA	Addition of Strep-tag to *csrA* in pBAD plasmid
CSO-0997	GATAACGAATATAATCAGCATTGC	Deletion of *fliW* using *aac*(3)-IV in *C. jejuni*
CSO-0998	TCCTAGTTAGTCACCCGGGTACGCATTTCAGCTAGGGTCATG	Deletion of *fliW* using *aac*(3)-IV in *C. jejuni*
CSO-0999	TGTGTTTTAGTACCTGGAGGGAATACCGACTTTTTTCAAGCTGATC	Deletion of *fliW* using *aac*(3)-IV in *C. jejuni*
CSO-1000	GACAAACCTTCATAAACTCCAG	Deletion of *fliW* using *aac*(3)-IV in *C. jejuni*
CSO-1002	GTTTTTCTCGAGAAATTTTGGCACAGTTTTTGCTTA	Overlap PCR construction of *flaG*-3xFLAG with *aphA-3* cassette

CSO-1003	GTTTTTTCTAGACCTGTGTTTACAATCTTAGCAAC	Overlap PCR construction of *flaG*-3xFLAG with *aphA -3* cassette
CSO-1005	GTGATAGAAGATTTGATCTTGC	Verification of *flaG*-3xFLAG in *C. jejuni*
CSO-1011	P-GATGATCTCCAAATCCGCGT	Cloning of *flgI*-3xFLAG in pGG1
CSO-1012	GTTTTTCTCGAGAAATTCCACAAAATTTTAGCC	Cloning of *flgI*-3xFLAG in pGG1
CSO-1013	GTTTTTTCTAGATGCGATTTTACTCGCTTTATCA	Cloning of *flgI*-3xFLAG in pGG1
CSO-1014	GTTTTTGAATTCACGCGGATTTGGAGATCATC	Cloning of *flgI*-3xFLAG in pGG1
CSO-1015	AACTGTAATGGGCGGAGCTA	Verification of *flgI*-3xFLAG in *C. jejuni*
CSO-1072	TAAAGCCTGATTACGATTTGGC	Verification of *fliW* deletion in *C. jejuni*
CSO-1081	GTTTTTTTTAATACGACTCACTATAGGTAACAAGTTCATTGATGAGCTTG	*in vitro* transcription of *flaA* M1/M2 variant, carries T7 promoter
CSO-1082	GTTTTTTTTAATACGACTCACTATAGGAAATTAAAATTTTAAAAAGGAAGTTAAA	*in vitro* transcription of Cj0040, carries T7 promoter
CSO-1083	TCTAAAACTTGAAGCAAAACTTC	*in vitro* transcription of Cj0040
CSO-1084	GTTTTTTTTAATACGACTCACTATAGGACTAGCAATAGGAAATTTTAAAAAG	*in vitro* transcription of *flaG*, carries T7 promoter
CSO-1085	TGTGTCTCACTTGTTCTTTGG	*in vitro* transcription of *flaG*
CSO-1088	GTTTTTTTTAATACGACTCACTATAGGCAAATGTTGATGTTTTAATCGAA	*in vitro* transcription of *flgA*, carries T7 promoter
CSO-1089	TTTACCCACTACGATACCTTG	*in vitro transcription of flgA*
CSO-1092	GTTTTTTTTAATACGACTCACTATAGGATTATAACTAAGATCAAGGAG	*in vitro* transcription of *flgM*, carries T7 promoter
CSO-1093	CTTTATCTATTCTATTTGTATTTAATG	*in vitro* transcription of *flgM*
CSO-1098	TCACCGTCATGGTCTTTGTAGTCACTCTCCTTATCAAATATCATTCC	Overlap PCR construction of *flaG*-3xFLAG with *aphA -3* cassette
CSO-1099	ATTGTTTTAGTACCTGGAGGGGGAATGATATTTGATAAGGAGAGT	Overlap PCR construction of *flaG*-3xFLAG with *aphA -3* cassette
CSO-1114	TAACAAGTTCATAAATGAGCTTGAATTTTTTTAAAAGG	Introduction of SL1$^{GGA>AAA}$ (M1) mutation into *flaA* 5' UTR
CSO-1115	ATTCAAGCTCATTTATGAACTTGTTAAATGCTATATCG	Introduction of SL1$^{GGA>AAA}$ (M1) mutation into *flaA* 5' UTR
CSO-1116	TTTTTTAAAAGGGTTTAAAATGGGATTTCGTATTAACA	Introduction of SL2$^{GGA>GGG}$ (M3) mutation into *flaA* 5' UTR
CSO-1117	TCCCATTTTAAACCCTTTTAAAAAAATTCAAGCTCAT	Introduction of SL2$^{GGA>GGG}$ (M3) mutation into *flaA* 5' UTR
CSO-1138	GTTTTTCATATGTATAAAATATTTTTTTGATTGCACGATATAGCATTTAACAAGTTCATGGATGAGCTT	Cloning of *flaA_mini* in *Campylobacter rdxA* complementation plasmid
CSO-1139	GTTTTTATCGATAAGGCCCAGTCTTTCGACT	Cloning of *flaA_mini* in *Campylobacter rdxA* complementation plasmid
CSO-1144	AGTGGAAAAGTTCTTTTAGACGG	Overlap PCR construction of *rpoN* deletion with *aac*(3)-IV
CSO-1145	TCCTAGTTAGTCACCCGGGTACCTTGGGTGATTTTTTGCTTTAACA	Overlap PCR construction of *rpoN* deletion with *aac*(3)-IV
CSO-1146	TGTGTTTTAGTACCTGGAGGGAATATCTATCAATCTATCAAACCCATTAC	Overlap PCR construction of *rpoN* deletion with *aac*(3)-IV
CSO-1147	CATTGGACGCTCAGGACG	Overlap PCR construction of *rpoN* deletion with *aac*(3)-IV
CSO-1148	AACAACTTTTTATATGATATGTGGAC	Verification of *rpoN* deletion in *C. jejuni*
CSO-1149	GAATTCTTAGGTCATTTAAGCGC	Overlap PCR construction of *fliA* deletion with *aac*(3)-IV
CSO-1150	TCCTAGTTAGTCACCCGGGTACTTTCTTTAGCATTTGTGCATAAGC	Overlap PCR construction of *fliA* deletion with *aac*(3)-IV
CSO-1151	TGTGTTTTAGTACCTGGAGGGAATATAAAAAACTTAGAGAAAGGCTAGTG	Overlap PCR construction of *fliA* deletion with *aac*(3)-IV
CSO-1152	GATAACAATCTCATTTTGAGATACG	Overlap PCR construction of *fliA* deletion with *aac*(3)-IV
CSO-1153	TGCAGATGCAAACATTAAAAATCC	Verification of *fliA* deletion in *C. jejuni*
CSO-1407	CAGCCAAACAACTTGGACTT	Verification of Cj0529c-3xFLAG in *C. jejuni*
CSO-1408	TCACCGTCATGGTCTTTGTAGTCTTTTCCTTTTTGTAAATTTATGGCTT	Overlap PCR construction of Cj0529-3xFLAG with *aphA -3* cassette

CSO-1409	GCTCCTTATGATGAAGGAGT	Overlap PCR construction of Cj0529-3xFLAG with *aphA -3* cassette
CSO-1410	CTTTAACTTAATTTAGAGCTTGC	Overlap PCR construction of Cj0529-3xFLAG with *aphA -3* cassette
CSO-1411	ATTGTTTTAGTACCTGGAGGGAATATTTTTGATATTTTTA TACAAAATAGTTAA	Overlap PCR construction of Cj0529-3xFLAG with *aphA -3* cassette
CSO-1471	GTTTTTTTTAATACGACTCACTATAGGTAACAAGTTCATA AATGAGCTTGA	*in vitro* transcription of *flaA* 5' UTR M1 variant, carries T7 promoter
CSO-1548	TCCTAGTTAGTCACCCGGGTATTTAAATCCTTTTAAAAAA ATTCAAGCT	Overlap PCR construction of *flaA* deletion with *aphA-3*
CSO-1549	ATTGTTTTAGTACCTGGAGGGAATATTTACAAAAGCTGC AATATATACAAAT	Overlap PCR construction of *flaA* deletion with *aphA-3*
CSO-1550	ATAGCTTGACCTAAAGTGGCT	Overlap PCR construction of *flaA* deletion with *aphA-3*
CSO-1665	GTTTTTTTTAATACGACTCACTATAGGATTTTTATTAAATT GAAGGGGTGGG	*in vitro* transcription of Cj1324, carries T7 promoter
CSO-1666	TACCTTCTTTATCTTTTGTAAAATTAATAC	*in vitro* transcription of Cj1324
CSO-1678	GTACCCGGGTGACTAACTAGGGTGACTAACTAGGAGGA ATAAATG	Amplification of Hyg^R cassette
CSO-1679	TATTCCCTCCAGGTACTAAAACAGTCATATTCCCTCCAG GTATCA	Amplification of Hyg^R cassette
CSO-1815	GTTTTTATGCATCGATGCAATATTTTGAAAGGATT	*flaB* cloning in pXG10
CSO-1816	GTTTTTGCTAGCACCTGAACTAAGTCTGCTTAAA	*flaB* cloning in pXG10
CSO-1817	GTTTTTTTTAATACGACTCACTATAGGCGATGCAATATTT TGAAAGGATT	*in vitro* transcription of *flaB* leader, carries T7 promoter
CSO-1818	ACCTGAACTAAGTCTGCTTAAA	*in vitro* transcription of *flaB* leader
CSO-1819	TAGAAATTTCAAGGAAGAAATATGCATGGAAAAATAGCT ATTTATATGGATTCTACAGGACGTGGAACCG	Cj1249 cloning in pXG10
CSO-1820	CTAGCGGTTCCACGTCCTGTAGAATCCATATAAATAGCT ATTTTTCCATGCATATTTCTTCCTTGAAATTTCTATGCA	Cj1249 cloning in pXG10
CSO-1823	GTTTTTATGCATACTAGCAATAGGAAATTTTAAAAAG	*flaG* cloning in pXG10
CSO-1824	GTTTTTGCTAGCCTGTGTCTCACTTGTTCTTTG	*flaG* cloning in pXG10
CSO-1825	GTTTTTATGCATAAAAAACTTAAGCAAAGGAAGGC	*pseB* cloning in pXG10
CSO-1826	GTTTTTGCTAGCTTCTAGCAAAACCTTAGTATAAGTT	*pseB* cloning in pXG10
CSO-1895	[CY5] ATTGGTGTTAATACGAAATCCCATTT	FISH oligo 1 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-1896	[CY5] AATTCAAGCTCATCCATGAACTTGT	FISH oligo 2 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-1963	[CY5] CGTTTGCTTTTGCATTTAAAGCTG	FISH oligo 3 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-1964	[CY5] CTTAAAGAAGCATCTAAACTTTTACTAT	FISH oligo 4 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2006	[FITC] GCTGCCTCCCGTAGGAGT	Universal FISH oligo to detect 16s rRNA (FITC-labeled 5' end)
CSO-2019	AAGGATTTAAAAAGGGATTTCGTATTAACACCAAT	Introduction of start $codon^{AUG>AAG}$ mutation into *flaA* 5' UTR
CSO-2020	ATACGAAATCCCTTTTTAAATCCTTTTAAAAAAATTCAAG C	Introduction of start $codon^{AUG>AAG}$ mutation into *flaA* 5' UTR
CSO-2023	[CY5] GAGCTAAGATATTTGCTTTAGAGTAG	FISH oligo 5 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2024	[CY5] TGCCATGGCATAAGAGCCGCT	FISH oligo 6 to detect *flaA* mRNA (Cy- labeled 5' end)
CSO-2150	GTGAGCAAGGGCGAGGA	Amplification of *mCherry* (2^{nd} codon to stop)
CSO-2151	TTACTTGTACAGCTCGTCCAT	Amplification of *mCherry* (2^{nd} codon to stop)
CSO-2155	CCTCCTCGCCCTTGCTCACTTTTTTAATATAATTAGCAAT TTGATCA	Overlap PCR construction of *fliW-mCherry* with *aac*(3)-IV cassette
CSO-2156	CCTCCTCGCCCTTGCTCACTTTGATTAGTTTTTTGCTTAA GTCAT	Overlap PCR construction of *csrA-mCherry* with *aphA -3* cassette
CSO-2746	GTAAAGCCACCCGCTCCTATG	NB oligo to detect Cj1321_mini
CSO-2809	[Cy5] CACGGATTTGCGATTCTGCTG	FISH oligo 7 to detect *flaA* mRNA (Cy5-labeled 5'

		end)
CSO-2810	[Cy5] GTGATGTTGTTTATAGTTGATGTAAC	FISH oligo 8 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2811	[Cy5] AACTAAGGCTCCATTAGCATCAC	FISH oligo 9 to detect *flaA* mRNA (Cy5 -labeled 5' end)
CSO-2812	[Cy5] GTAATCTACTTTACCGATTTTTACCC	FISH oligo 10 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2813	[Cy5] AGATCTTAAACTATCTGCTATCGC	FISH oligo 11 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2814	[Cy5] TAGATATAGCTTGACCTAAAGTATTAG	FISH oligo 12 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2815	[Cy5] TATTAGCATCAATTTGTCCTTTTGAC	FISH oligo 13 to detect *flaA* mRNA (Cy5 -labeled 5' end)
CSO-2816	[Cy5] GTAATCTTAAAACATTTTGTTGAACAGAA	FISH oligo 14 to detect *flaA* mRNA (Cy5-labeled 5' end)
CSO-2821	P-TTTGATAAGTTTATTTGGATACAATTGTGGTAACAAGT TCATGGATGAGCTT	Exchange of *flaA* promoter with *metK* promoter
CSO-2841	CTTCAGGTTCAGGTTATTCTG	Verification of *flaB* deletion in *C. jejuni*
CSO-2842	GGCTCAGGTTTTTCAAGTGG	Overlap PCR construction of *flaB* deletion with *aphA-3*
CSO-2843	TCCTAGTTAGTCACCCGGGTAATCCTAAAACCCATTTTA AATCCTT	Overlap PCR construction of *flaB* deletion with *aphA-3*
CSO-2844	ATTGTTTTAGTACCTGGAGGGAATACAGCAAAATGTTTT AAAACTTCTTC	Overlap PCR construction of *flaB* deletion with *aphA-3*
CSO-2853	GCTTGATAAATTAAAAATTTACTAAAATTAGGATCCTTTT ATGGATAATTTTTAAA	Exchange of *flaA* promoter with *metK* promoter
CSO-2825	AAAGGATTTAAAGTGGGATTTCGTATTAACACCAA	Introduction of start codon$^{AUG>GUG}$ mutation into *flaA* 5' UTR
CSO-2826	TACGAAATCCCACTTTAAATCCTTTTAAAAAAATTCAAGC	Introduction of start codon$^{AUG>GUG}$ mutation into *flaA* 5' UTR
CSO-2827	AGGATTTAAAATTGGATTTCGTATTAACACCAATG	Introduction of start codon$^{AUG>AUU}$ mutation into *flaA* 5' UTR
CSO-2828	AATACGAAATCCAATTTTAAATCCTTTTAAAAAAATTCAA GC	Introduction of start codon$^{AUG>AUU}$ mutation into *flaA* 5' UTR
CSO-2829	TTAAAATGGGATAGCGTATTAACACCAATGTTGCAG	Introduction of 3rd codon$^{UUU>UAG}$ mutation into *flaA* 5' UTR
CSO-2830	GGTGTTAATACGCTATCCCATTTTAAATCCTTTTAAAAAA AT	Introduction of 3rd codon$^{UUU>UAG}$ mutation into *flaA* 5' UTR
CSO-2831	TTAAAATGGGATTCCGTATTAACACCAATGTTGCAG	Introduction of 3rd codon$^{UUU>UUC}$ mutation into *flaA* 5' UTR
CSO-2832	GGTGTTAATACGGAATCCCATTTTAAATCCTTTTAAAAAA AT	Introduction of 3rd codon$^{UUU>UUC}$ mutation into *flaA* 5' UTR
CSO-2833	ACTCAAGCGGCTTAAGATGGACAAAGTTTAAAAACAAG	Introduction of 101st codon$^{CAA>UAA}$ mutation into *flaA* 5' UTR
CSO-2834	TTTGTCCATCTTAAGCCGCTTGAGTTGCCTTA	Introduction of 101st codon$^{CAA>UAA}$ mutation into *flaA* 5' UTR
CSO-2835	GCTGCAACATTGGTGTTAATACG	NB oligo to detect *flaA* mRNA in all 5' UTR point mutants and *flaA_mini*
CSO-3068	TGCCAAATTTAGGAAAAATCAATC	Overlap PCR construction of N267A mutation in FlhB with *aphA-3*
CSO-3069	TGTGATCACCACATCCGCA	Overlap PCR construction of N267A mutation in FlhB with *aphA*
CSO-3070	GGTGCGGATGTGGTGATCACAGCCCCAACGCATTATGC CGTA	Overlap PCR construction of N267A mutation in FlhB with *aphA*
CSO-3073	TCCTAGTTAGTCACCCGGGTATTAAAAGTTTTAGTTTCCT TTTTTTAC	Overlap PCR construction of N267A mutation in FlhB with *aphA*
CSO-3074	ATTGTTTTAGTACCTGGAGGGAATAAAAACTTTTAAGGT GCTTTAGTTC	Overlap PCR construction of N267A mutation in FlhB with *aphA*
CSO-3075	CGATGAACTCAAATCGTTAAAAC	Overlap PCR construction of N267A mutation in FlhB with *aphA*
CSO-3077	CCAAAAACAGAACAAGATATAGC	Overlap PCR construction of *flhX* deletion with *aac*(3)-IV
CSO-3078	TCCTAGTTAGTCACCCGGGTAGATTTTGCTCATACTTTTT TATCCA	Overlap PCR construction of *flhX* deletion with *aac*(3)-IV
CSO-3079	ATTGTTTTAGTACCTGGAGGGAATATGCCTTTATCTATCA AATGGCAA	Overlap PCR construction of *flhX* deletion with *aac*(3)-IV

CSO-3080	GGAGCTTTTAAAAGATAGCTTTG	Overlap PCR construction of *flhX* deletion with *aac*(3)-IV
CSO-3076	CAAGCAGAACTTAAAAATCTCGA	Verification of *flhX* deletion in *C. jejuni*
HPK1	GTACCCGGGTGACTAACTAGG	Amplification of *aphA-3* cassette
HPK2	TATTCCCTCCAGGTACTAAAACA	Amplification of *aphA-3* cassette
JVO-0054	GGGATCAAGCCTGATTG	Sense to *aphA-3* for verification
JVO-0900	GGAGAAACAGTAGAGAGTTGC	Antisense oligo for inverse PCR on pBAD/*Myc*-His A
JVO-0901	TTTTTTCTAGATTAAATCAGAACGCAGA	Sense oligo for inverse PCR on pBAD /*Myc*-His A
JVO-5142	GACTACAAAGACCATGACGG	Sense oligo to 3xFLAG tag
pBAD-FW	ATGCCATAGCATTTTTATCC	Verification of insert in pBAD/*Myc*-His A plasmid
pZE-A	GTGCCACCTGACGTCTAAGA	Verification of insert in pJV752.1

Appendix Table 13. List of all *C. jejuni* and *E. coli* strains. All strains were generated in this thesis unless otherwise stated. All *C. jejuni* strains correspond to NCTC11168 background unless otherwise stated.

Name	Description	Strain number	Resistance
C. jejuni **strains**	All *C. jejuni* strains have NCTC11168 background unless otherwise stated		
NCTC11168	WT strain; (Kindly provided by Arnoud van Vliet, Institute of Food Research, Norwich, UK)	CSS-0032	-
81-176	WT strain; (Patricia Guerry, Naval Medical Research Center, Silver Spring, MD, USA)	CSS-0063	Tet^R
RM1221	WT strain (Dirk Hofreuter, Hannover Medical School, Hannover, Germany)	CSS-0088	-
81116	WT strain (Steffen Backert, University College Dublin, Dublin, Ireland)	CSS-0129	-
Δ*rnc*	*rnc*:: *aphA-3* Deletion of *rnc*	CSS-0606	Kan^R
CsrA-3xFLAG NCTC11168	*csrA*-3xFLAG::*aphA-3* C-terminal 3xFLAG tag at native locus (Cj1103) in NCTC11168 background	CSS-0625	Kan^R
CsrA-3xFLAG 81-176	81-176, C-terminal 3xFLAG tag in 81-176 background	CSS-0604	Tet^R Kan^R
Δ*csrA*	*csrA*::*cat* Deletion of *csrA* (Cj1103)	CSS-0643	Cm^R
Δ*fliW*	*fliW*::*aac*(3)-IV Deletion of *fliW* (Cj1075)	CSS-0820	Gm^R
Δ*csrA* **Δ***fliW*	*csrA*::*cat fliW*::*aac(3)-IV* Deletion of *csrA* and *fliW*	CSS-1134	Cm^R Gm^R
Δ*rpoN*	*rpoN*::*aac*(3)-IV Deletion of *rpoN* (Cj0670)	CSS-1141	Gm^R
Δ*fliA*	*fliA*::*aac*(3)-IV Deletion of *fliA* (Cj0061c)	CSS-1133	Gm^R
Δ*flaA*	*flaA*:: *aphA-3* Deletion of *flaA* (Cj1339c)	CSS1512	Kan^R
Δ*flaB*	*flaB*:: *aphA-3* Deletion of *flaB* (Cj1338c)	CSS-2892	Kan^R
Δ*flaA* **Δ***flaB*	*flaAB*:: *aphA-3* Deletion of *flaA* (Cj1339c) and *flaB* (Cj1338c)	CSS-2891	Kan^R
flaA-**3xFLAG**	*flaA*-3xFLAG::*aphA-3* C-terminal 3xFLAG tag of *flaA* at native locus (Cj1339c)	CSS-0640	Kan^R
flaA-**3xFLAG Δ***csrA*	*flaA*-3xFLAG::*aphA-3 csrA*::*cat* Deletion of *csrA* in *flaA*-3xFLAG background	CSS-0644	Kan^R Cm^R
flaA-**3xFLAG Δ***fliW*	*flaA*-3xFLAG::*aphA-3 fliW*::*aac*(3)-IV Deletion of *fliW* in *flaA*-3xFLAG background	CSS-1100	Gm^R Kan^R
flaA-**3xFLAG Δ***csrA* **Δ***fliW*	*flaA*-3xFLAG::*aphA-3 csrA*::*cat fliW*::*aac*(3)-IV Deletion of *csrA* and *fliW* in *flaA*-3xFLAG background	CSS-1107	Kan^R Cm^R Gm^R
FlaA-3xFLAG M1	*flaA*[5' UTR $SL1^{GGA\rightarrow AAA}$::*aac*(3)-IV]-3xFLAG::*aphA-3* $SL1^{flaA}$ 5' UTR point mutant in *flaA*-3xFLAG background	CSS-0991	Kan^R Gm^R
FlaA-3xFLAG M1 Δ*csrA*	*flaA*[5' UTR $SL1^{GGA\rightarrow AAA}$::*aac*(3)-IV]-3xFLAG::*aphA-3 csrA*::*cat* $SL1^{flaA}$ point mutant and *csrA* deletion in *flaA*-3xFLAG background	CSS-1410	Kan^R Cm^R Gm^R
FlaA-3xFLAG M1 Δ*fliW*	*flaA*[5' UTR $SL1^{GGA\rightarrow AAA}$::*cat*]-3xFLAG::*aphA-3 fliW*::*aac*(3)-IV Deletion of *fliW* in *flaA* 5' UTR point mutant/3xFLAG-tag	CSS-1418	Kan^R Cm^R Gm^R
FlaA-3xFLAG M1 Δ*fliW* **Δ***csrA*	*flaA*-3xFLAG::*aphA-3 flaA*[5' UTR $SL1^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV *csrA*:: *aph(7")* $SL1^{flaA}$ point mutant, *fliW* deletion and *csrA* deletion in *flaA*-3xFLAG background	CSS-1554	Kan^R Cm^R Gm^R Hyg^R
FlaA-3xFLAG M2	*flaA*[5' UTR $SL1^{GGA\rightarrow UGA}$::*aac*(3)-IV]-3xFLAG::*aphA-3* $SL1^{flaA}$ point mutant in *flaA*-3xFLAG background	CSS-0955	Kan^R Gm^R
FlaA-3xFLAG M2 Δ*csrA*	*flaA*[5' UTR $SL1^{GGA\rightarrow UGA}$::*aac*(3)-IV]-3xFLAG::*aphA-3 csrA*::*cat* $SL1^{flaA}$ point mutant and *csrA* deletion in *flaA*-3xFLAG background	CSS-1105	Kan^R Cm^R Gm^R
FlaA-3xFLAG M2/M3	*flaA*[5' UTR $SL1^{GGA\rightarrow UGA}$ $SL2^{GGA\rightarrow GGG}$::*aac*(3)-IV]-3xFLAG::*aphA-3* $SL1^{flaA}$ and $SL2^{flaA}$ double point mutant in *flaA*-3xFLAG background	CSS-1095	Kan^R Gm^R

FlaA-3xFLAG M2/M3 Δ*csrA*	*flaA*[5' UTR SL1$^{GGA\rightarrow UGA}$ SL2$^{GGA\rightarrow GGG}$::*aac*(3)-IV]-3xFLAG::*aphA-3* *csrA*::*cat* SL1^{flaA} and SL2^{flaA} double point mutant and *csrA* deletion in *flaA*-3xFLAG background	CSS-1106	Kan^R Cm^R Gm^R
FlaG-3xFLAG	*flaG*-3xFLAG::*aphA-3*; C-terminal 3xFLAG-tag of *flaG* at native locus (Cj0547)	CSS-0968	Kan^R
FlaG-3xFLAG Δ*csrA*	*flaG*-3xFLAG::*aphA-3* *csrA*::*cat* Deletion of *csrA* in *flaG*-3xFLAG background	CSS-0983	Kan^R Cm^R
FlaG-3xFLAG Δ*fliW*	*flaG*-3xFLAG::*aphA-3* *fliW*::*aac*(3)-IV Deletion of *fliW* in FlaG-3xFLAG background	CSS-1112	Kan^R Gm^R
FlaG-3xFLAG Δ*csrA* Δ*fliW*	*flaG*-3xFLAG::*aphA-3* *csrA*::*cat* *fliW*::*aac*(3)-IV Deletion of *csrA* and *fliW* in FlaG-3xFLAG background	CSS-1204	Kan^R Cm^R Gm^R
FlaG-3xFLAG M1	*flaG*::3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] SL1^{flaA} point mutant in *flaG*-3xFLAG background	CSS-1399	Kan^R Gm^R
FlaG-3xFLAG M1 Δ*csrA*	*flaG*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] *csrA*::*cat* SL1^{flaA} point mutant and *csrA* deletion in *flaG*-3xFLAG background	CSS-1411	Kan^R Cm^R Gm^R
FlaG-3xFLAG M1 Δ*fliW*	*flaG*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV SL1^{flaA} point mutant and *fliW* deletion in *flaG*-3xFLAG background	CSS-1421	Kan^R Cm^R Gm^R
FlaG-3xFLAG M1 Δ*fliW* Δ*csrA*	*flaG*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV *csrA*::*aph(7'')*; SL1^{flaA} point mutant, *fliW* deletion and *csrA* deletion in *flaG*-3xFLAG background	CSS-1435	Kan^R Cm^R Gm^R Hyg^R
FlgI-3xFLAG	*flgI*-3xFLAG::*aphA-3* C-terminal 3xFLAG tag of *flgI* at native locus (Cj1462)	CSS-0967	Kan^R
FlgI-3xFLAG Δ*csrA*	*flgI*-3xFLAG::*aphA-3* *csrA*::*cat* Deletion of *csrA* in *flgI*-3xFLAG background	CSS-0982	Kan^R Cm^R
FlgI-3xFLAG Δ*fliW*	*flgI*-3xFLAG::*aphA-3* *fliW*::*aac*(3)-IV Deletion of *fliW* in *flgI*-3xFLAG background	CSS-1114	Kan^R Gm^R
FlgI-3xFLAG Δ*csrA* Δ*fliW*	*flgI*-3xFLAG::*aphA-3* *csrA*::*cat* *fliW*::*aac*(3)-IV Deletion of *csrA* and *fliW* in *flgI*-3xFLAG background	CSS-1203	Kan^R Cm^R Gm^R
FlgI-3xFLAG M1	*flgI*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] SL1^{flaA} point mutant in *flgI*-3xFLAG background	CSS-1426	Kan^R Gm^R
FlgI-3xFLAG M1 Δ*csrA*	*flgI*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] *csrA*::*cat* SL1^{flaA} point mutant and *csrA* deletion in *flgI*-3xFLAG background	CSS-1542	Kan^R Cm^R Gm^R
FlgI-3xFLAG M1 Δ*fliW*	*flgI*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV SL1^{flaA} point mutant and *fliW* deletion in *flgI*-3xFLAG background	CSS-1420	Kan^R Cm^R Gm^R
FlgI-3xFLAG M1 Δ*fliW* Δ*csrA*	*flgI*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV *csrA*::*aph(7'')* SL1^{flaA} point mutant, *fliW* deletion and *csrA* deletion in *flgI*-3xFLAG background	CSS-1436	Kan^R Cm^R Gm^R Hyg^R
FlaB-3xFLAG	*flaB*-3xFLAG::*aphA-3;* C-terminal 3xFLAG tag of *flaB* at native locus (Cj1338c)	CSS-0641	Kan^R
FlaB-3xFLAG Δ*csrA*	*flaB*-3xFLAG::*aphA-3* *csrA*::*cat* Deletion of *csrA* in *flaB*-3xFLAG background	CSS-0645	Kan^R Cm^R
FlaB-3xFLAG Δ*fliW*	*flaB*-3xFLAG::*aphA-3* *fliW*::*aac*(3)-IV Deletion of *fliW* in *flaB*-3xFLAG background	CSS-1201	Kan^R Gm^R
FlaB-3xFLAG Δ*csrA* Δ*fliW*	*flaB*-3xFLAG::*aphA-3* *csrA*::*cat* *fliW*::*aac*(3)-IV Deletion of *csrA* and *fliW* in *flaB*-3xFLAG background	CSS-1202	Kan^R Cm^R Gm^R
FlaB-3xFLAG M1	*flaB*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] SL1^{flaA} point mutant in *flaB*-3xFLAG background	CSS-1405	Kan^R Gm^R
FlaB-3xFLAG M1 Δ*csrA*	*flaB*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] SL1^{flaA} point mutant and *csrA* deletion in *flaB*-3xFLAG background	CSS-1543	Kan^R Cm^R Gm^R
FlaB-3xFLAG M1 Δ*fliW*	*flaB*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV SL1^{flaA} point mutant and *fliW* deletion in *flaB*-3xFLAG background	CSS-1419	Kan^R Cm^R Gm^R
FlaB-3xFLAG M1 Δ*fliW* Δ*csrA*	*flaB*-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV *csrA*::*aph(7'')* SL1^{flaA} point mutant, *fliW* deletion and *csrA* deletion in *flaB*-3xFLAG background	CSS-1438	Kan^R Cm^R Gm^R Hyg^R
Cj0529-3xFLAG	Cj0529-3xFLAG::*aphA-3* C-terminal 3xFLAG tag of Cj0529 at native locus	CSS-1541	Kan^R
Cj0529-3xFLAG Δ*csrA*	Cj0529-3xFLAG::*aphA-3* *csrA*::*cat* Deletion of *csrA* in Cj0529-3xFLAG background	CSS-1431	Kan^R Cm^R
Cj0529-3xFLAG Δ*fliW*	Cj0529-3xFLAG::*aphA-3* *fliW*::*aac*(3)-IV Deletion of *fliW* in Cj0529-3xFLAG background	CSS-1430	Kan^R Gm^R
Cj0529-3xFLAG Δ*csrA* Δ*fliW*	Cj0529-3xFLAG::*aphA-3* *csrA*::*cat* *fliW*::*aac*(3)-IV Deletion of *csrA* and *fliW* in Cj0529-3xFLAG background	CSS-1437	Kan^R Cm^R Gm^R

Cj0529-3xFLAG M1	Cj0529-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] SL1^{flaA} point mutant in Cj0529-3xFLAG background	CSS-1432	Kan^R Gm^R
Cj0529-3xFLAG M1 Δ*csrA*	Cj0529-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*aac*(3)-IV] *csrA*::*cat* SL1^{flaA} point mutant and *csrA* deletion in Cj0529-3xFLAG background	CSS-1576	Kan^R Cm^R Gm^R
Cj0529-3xFLAG M1 Δ*fliW*	Cj0529-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV SL1^{flaA} point mutant and *fliW* deletion in Cj0529-3xFLAG background	CSS-1575	Kan^R Cm^R Gm^R
Cj0529-3xFLAG M1 Δ*fliW* Δ*csrA*	Cj0529-3xFLAG::*aphA-3* *flaA*[5' UTR SL1$^{GGA\rightarrow AAA}$::*cat*] *fliW*::*aac*(3)-IV *csrA:: aph(7")* SL1^{flaA} point mutant, *fliW* deletion and *csrA* deletion in Cj0529-3xFLAG background	CSS-1582	Kan^R Cm^R Gm^R Hyg^R
flaA X1	*flaA*[Start Codon$^{AUG\rightarrow AAG}$::*aac*(3)-IV]	CSS-1586	Gm^R
flaA X2	*flaA*[Start Codon$^{AUG\rightarrow AUU}$::*aac*(3)-IV]	CSS-3089	Gm^R
flaA X3	*flaA*[Start Codon$^{AUG\rightarrow GUG}$::*aac*(3)-IV]	CSS-3087	Gm^R
flaA X4	*flaA*[3rd Codon$^{UUU\rightarrow UAG}$::*aac*(3)-IV]	CSS-3091	Gm^R
flaA X5	*flaA*[3rd Codon$^{UUU\rightarrow UUC}$::*aac*(3)-IV]	CSS-3093	Gm^R
flaA X6	*flaA*[101st Codon$^{CAA\rightarrow UAA}$::*aac*(3)-IV]	CSS-3095	Gm^R
P$_{MetK}$-*flaA*	Exchange of *flaA* native promoter with *metK* promoter	CSS-3096	Gm^R
P$_{MetK}$-*flaA* Δ*csrA*	P$_{MetK}$-*flaA:: aac*(3)-IV *csrA*::*cat* Exchange of *flaA* native promoter with *metK* promoter	CSS-3100	Gm^R Cm^R
P$_{MetK}$-*flaA* Δ*fliW*	P$_{MetK}$-*flaA:: aac*(3)-IV *fliW*::*hyg* Exchange of *flaA* native promoter with *metK* promoter	CSS-3116	Gm^R Hyg^R
P$_{MetK}$-*flaA*-3xFLAG	*flaA*-3xFLAG::*aphA-3* P$_{MetK}$-*flaA:: aac*(3)-IV C-terminal 3xFLAG tag of *flaA* in *flaA* promoter exchanged strain	CSS-3098	Gm^R Kan^R
P$_{MetK}$-*flaA*-3xFLAG Δ*csrA*	*flaA*-3xFLAG::*aphA-3* P$_{MetK}$-*flaA:: aac*(3)-IV *csrA*::*cat* C-terminal 3xFLAG tag of *flaA* in *flaA* promoter exchanged strain	CSS-3102	Gm^R Kan^R Cm^R
P$_{MetK}$-*flaA*-3xFLAG Δ*fliW*	*flaA*-3xFLAG::*aphA-3* P$_{MetK}$-*flaA:: aac*(3)-IV *fliW*::*hyg* C-terminal 3xFLAG tag of *flaA* in *flaA* promoter exchanged strain	CSS-3104	Gm^R Kan^R Hyg^R
P$_{MetK}$-*flaA*-3xFLAG Δ*csrA* Δ*fliW*	*flaA*-3xFLAG::*aphA-3* P$_{MetK}$-*flaA:: aac*(3)-IV *csrA*::*cat* *fliW*::*aph(7")* C-terminal 3xFLAG tag of *flaA* in *flaA* promoter-exchanged strain	CSS-3124	Gm^R Kan^R Cm^R Hyg^R
FliW-3xFLAG	*fliW*-3xFLAG::*aac*(3)-IV C-terminal 3xFLAG tag of *fliW* at native locus (Cj1075)	CSS-0962	Gm^R
FliW-mCherry	*fliW*-mCherry::*aac*(3)-IV C-terminal mCherry tag of *fliW* at native locus (Cj1075)	CSS-3073	Gm^R
CsrA-mCherry	*csrA*-mCherry::*aphA-3* C-terminal mCherry tag of *csrA* at native locus (Cj1103)	CSS-3071	Kan^R
FliW-3xFLAG CsrA-mCherry	*fliW*-3xFLAG::*aac*(3)-IV *csrA*-mCherry::*aphA-3*	CSS-3126	Gm^R Kan^R
FlaA-3xFLAG FliW-mCherry	*flaA*-3xFLAG::*aphA-3* *fliW*-mCherry::*aac*(3)-IV	CSS-3128	Gm^R Kan^R
***flaA*_mini**	*flaA*_mini::*rdxA* Introduction of *flaA_mini* in *rdxA* (Cj1066) complementation locus	CSS-3075	Cm^R
FlaA-3xFLAG *flaA*_mini	*flaA*-3xFLAG::*aphA-3* *flaA_mini*::*rdxA* Introduction of *flaA*_mini in *rdxA* (Cj1066) in *flaA*-3xFLAG background	CSS-3076	Cm^R Kan^R
FlaA-3xFLAG Δ*fliW* *flaA*_mini	*flaA*-3xFLAG::*aphA-3* *fliW*::*aac*(3)-IV *flaA_mini*::*rdxA* Introduction of *flaA_mini* in *rdxA* (Cj1066) in *flaA*-3xFLAG background with *fliW* deletion	CSS-3106	Cm^R Kan^R Gm^R
FlaG-3xFLAG *flaA*_mini	*flaG*-3xFLAG::*aphA-3* *flaA_mini*::*rdxA* Introduction of *flaA_mini* in *rdxA* (Cj1066) in *flaG*-3xFLAG background	CSS-3080	Cm^R Kan^R
FlaG-3xFLAG Δ*fliW* *flaA*_mini	*flaG*-3xFLAG::*aphA-3* *fliW*::*aac*(3)-IV *flaA_mini*::*rdxA* Introduction of *flaA_mini* in *rdxA* (Cj1066) in *flaG*-3xFLAG background with *fliW* deletion	CSS-3112	Cm^R Kan^R Gm^R
Cj1321_mini	Cj1321_mini::*rdxA* Introduction of *flaA_mini* in *rdxA* (Cj1066) complementation locus	CSS-3130	Cm^R
Cj1321_mini Δ*csrA*	Cj1321_mini::*rdxA csrA*::*aph(7")* Introduction of *flaA_mini* in *rdxA* (Cj1066) complementation locus	CSS-3131	Cm^R Hyg^R
Cj1321_mini Δ*fliW*	Cj1321_mini::*rdxA fliW*::*aac*(3)-IV Introduction of *flaA_mini* in *rdxA* (Cj1066) complementation locus	CSS-3132	Cm^R Gm^R
Δ*flhX*	*flhX*::*aac*(3)-IV Deletion of *flhX* (Cj0848c)	CSS-3860	Gm^R
***flhB* N267A**	N267A (AAC to GCC) mutation in the native locus of *flhB* (Cj0335)	CSS-3864	Kan^R

E. coli **strains**			
TOP10	*mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ* ΔM15 Δ*lacX74 deoR recA1 araD139* Δ(*ara-leu*)7697 *galU galK rpsL* endA1 nupG (from Invitrogen)	CSS-0070	Str^R
Δ*pgaA*	*pgaA* deletion in TOP10 background	CSS-0556	Str^R
Δ*pgaA* **Δ***csrA*	*pgaA* and *csrA* deletion in TOP10 background	CSS-0557	Str^R Kan^R
BTH101	F-, *cya-99*, *araD139, galE15, galK16, rpsL1*, *hsdR2*, *mcrA1*, *mcrB1* (from Euromedex)	CSS-0840	Str^R

Appendix Table 14. List of all plasmids used in this thesis.

Name	Description/Generation	Origin/ Marker	Reference
pJV752.1	Cloning vector, pZE12-*luc* with modified p15A origin	p15A mod/ Amp^R	[6]
pUC1813 *apra*	Carries *aac*(*3*)-*IV* gene	pBR322/ Gm^R	[180]
pAC1H	Carries *aph*(7") gene	ColE1/pBR322/Hyg^R	[181]
pBAD/Myc-His A	pBAD expression plasmid	pBR322/ Amp^R	Invitrogen
pZE12-luc	General expression plasmid	ColE1/ Amp^R	[205]
pXG-10	Standard plasmid for directional cloning of a target mRNA as N-terminal translational fusion to GFP	pSC101*/ Cm^R	[106]
pXG-30	Plasmid for cloning operon fusions with the N-terminus of downstream gene fused to GFP and the C-terminus of upstream gene fused to a short artificial reading frame composed of a FLAG epitope and truncated *lacZ* gene	pSC101*/ Cm^R	[106]
pKT25	Cloning vector, for fusing C-terminus of T25 fragment to gene of interest	P15A/ Kan^R	[113, 206]
pKNT25	Cloning vector, for fusing N-terminus of T25 fragment to gene of interest	P15A/ Kan^R	[113, 206]
pUT18	Cloning vector, for fusing N-terminus of T18 fragment to gene of interest	ColE1/ Amp^R	[113, 206]
pUT18C	Cloning vector, for fusing C-terminus of T18 fragment to gene of interest	ColE1/ Amp^R	[113, 206]
pKT25-zip	Leucine zipper fragment fused to T25 in the plasmid pKT25	P15A/ Kan^R	[113, 206]
pUT18C-zip	Leucine zipper fragment fused to T18 in the plasmid pKT18C	ColE1/ Amp^R	[113, 206]
pGD41-1	18T-CsrA (*fliW* fused to the C-terminal of T18 fragment in pUT18C)	ColE1/ Amp^R	This Study
pGD43-1	25T-FliW (*fliW* fused to the C-terminal of T25 fragment in pKT25)	P15A/ Kan^R	This Study
pGD49-1	CsrA-25T (*csrA* fused to the N-terminal of T25 fragment in pKNT25)	P15A/ Kan^R	This Study
pGD51-1	FliW-18T (*fliW* fused to the N-terminal of T18 fragment in pUT18)	ColE1/ Amp^R	This Study
pSSv68-1	CsrA-18T (*csrA* fused to the N-terminal of T18 fragment in pUT18)	ColE1/ Amp^R	This study
pSSv69-10	18T-FliW (N-terminal fusion of *fliW* to the T18 fragment in pUT18C)	ColE1/ Amp^R	This study
pGD68-1	pBAD::$CsrA_{Cj,}$ based on pBAD/Myc-His A	pBR322/ Amp^R	This study
pGD72-3	pBAD::$CsrA_{Cj}$-Strep, based on pGD68-1	pBR322/ Amp^R	This study
pGG1	Plasmid (based on pZE12-luc) harbouring 3xFLAG and non-polar *aphA-3* cassette. Used for introduction of 'UP' and 'DN' regions of a gene of interest to be FLAG-tagged	ColE1/ Kan^R Amp^R	Sharma lab
pGD78-1	*aphA -3* ORF in pGG1 replaced by *aac*(*3*)-*IV* ORF	ColE1/ Gm^R Amp^R	This study
pGD4-1	Plasmid harbouring *csrA*-3xFLAG C-terminal translational fusion, *csrA* upstream and downstream regions, and *aphA-3* cassette in pGG1 for chromosomal epitope tagging at native locus	ColE1/ Kan^R Amp^R	This study
pMW5-2	Plasmid harbouring *flaA*-3xFLAG C-terminal translational fusion, *flaA* upstream and downstream regions, and *aphA-3* cassette in pGG1 for chromosomal epitope tagging at native locus	ColE1/ Kan^R Amp^R	This study
pMW6-1	Plasmid harbouring *flaB*-3xFLAG C-terminal translational fusion, *flaB* upstream and downstream regions, and *aphA-3* cassette in pGG1 for chromosomal epitope tagging at native locus	ColE1/ Kan^R Amp^R	This study
pSSv1-2	Plasmid harbouring *flgI*-3xFLAG C-terminal translational fusion, *flgI* upstream and downstream regions, and *aphA-3* cassette in pGG1 for chromosomal epitope tagging at native locus	ColE1/ Kan^R Amp^R	This study
pGD70-5	Plasmid harbouring 1,100 bp region around *flaA* promoter; based on pJV752.1	p15A mod/ Amp^R	This study
pGD76-1	*aac*(3)-IV gentamicin cassette introduced upstream of *flaA* promoter in pGD70-5 in reverse orientation to *flaA*	p15A mod/ Gm^R Amp^R	This study
pGD92-1	M1 (SL1 GGA>AAA) mutation in *flaA* 5'UTR in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD77-1	M2 (SL1 GGA>UGA) mutation in *flaA* 5'UTR in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD93-1	M3 (SL2 GGA>GGG) mutation in *flaA* 5'UTR in pGD76-1	p15A mod/ Gm^R Amp^R	This study

pGD95-1	M2 (SL1 GGA>UGA) /M3 (SL2 GGA>GGG) mutation in *flaA* 5'UTR in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD25-1	pJV752 derived *C. jejuni* complementation plasmid for introduction in rdxA locus of strain NCTC11168 using Cm selection	p15A mod/ Amp^R/ Cm^R	This study
pGD98-12	pGD25-1 based complementation plasmid to introduce *flaA*_mini in *C. jejuni* NCTC11168	p15A mod/ Amp^R/ Cm^R	This study
pGD117.1	pGD25-1 based complementation plasmid to introduce Cj1321_mini in *C. jejuni* NCTC11168	p15A mod/ Amp^R/ Cm^R	This study
pGD114-2	Start codon mutation in *flaA* (AUG→AAG) in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD205-1	Start codon mutation in *flaA* (AUG→AUU) in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD204-1	Start codon mutation in *flaA* (AUG→GUG) in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD206-1	3rd codon mutation in *flaA* (AUG→UAG) in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD207-1	3rd codon mutation in *flaA* (AUG→UUC) in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD208-1	101st codon mutation in *flaA* (CAA→UAA) in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD209-1	*flaA* promoter replaced by *metK* promoter in pGD76-1	p15A mod/ Gm^R Amp^R	This study
pGD107-1	*aac(3)*-IV gentamicin cassette replaced by *cat* cassette in pGD92-1	p15A mod/ Cm^R Amp^R	This study
pGD31-1	5'UTR along with first 33 codons of *flaA* fused in-frame to *gfp* in pXG-10	pSC101*/ Cm^R	This study
pGD111-1	5'UTR along with first 25 codons of *flaG* fused in-frame to *gfp* in pXG-10	pSC101*/ Cm^R	This study
pGD38-1	5'UTR along with first 35 codons of *flgI* fused in-frame to *gfp* in pXG-10	pSC101*/ Cm^R	This study
pGD109-1	5'UTR along with first 35 codons of *flaB* fused in-frame to *gfp* in pXG-10	pSC101*/ Cm^R	This study
pGD112-1	5'UTR along with first 26 codons of *pseB* fused in-frame to *gfp* in pXG-10	pSC101*/ Cm^R	This study
pGD110-1	5'UTR along with first 16 codons of Cj1249 fused in-frame to *gfp* in pXG-10	pSC101*/ Cm^R	This study
pGD28-1	Last 17 codons of Cj0310c and first 23 codons of Cj0309c fused in-frame to FLAG-*lacZ* and *gfp,* respectively, in pXG-30	pSC101*/ Cm^R	This study
pGD27-1	Last 17 codons of Cj0805 and first 25 codons of *dapA* fused in-frame to FLAG-*lacZ* and *gfp,* respectively, in pXG-30	pSC101*/ Cm^R	This study

Appendix Table 15. Construction of *C. jejuni* mutants.

Mutation	'UP' PCR primers	'DN' PCR primers	Cassette/ Primers	UP-cassette-DN amplification primers	Mutant validation primers
Single gene deletions (overlap PCR)					
Δ*rnc*(Kan^R)	CSO-0241 CSO-0242	CSO-0243 CSO-0244	*aphA-3* (HPK1/HPK2)	CSO-0242 CSO-0243	CSO-0240 CSO-0023
Δ*csrA* (Cm^R)	CSO-0394 CSO-0615	CSO-0616 CSO-0395	*cat* (CSO-0613/-0614)	CSO-0394 CSO-0395	CSO-0392 CSO-0614
Δ*csrA* (Hyg^R)	CSO-0393 CSO-0394	CSO-0395 CSO-0396	*aph(7'')* (CSO-1678/-1679)	CSO-0394 CSO-0395	CSO-0392 HPK2
Δ*fliW* (Gm^R)	CSO-0997 CSO-0998	CSO-0999 CSO-1000	*aac*(3)-IV (HPK1/HPK2)	CSO-0997 CSO-1000	CSO-1072 HPK2
Δ*fliW* (Hyg^R)	CSO-0997 CSO-0998	CSO-0999 CSO-1000	*aph(7'')* (CSO-1678/-1679)	CSO-0997 CSO-1000	CSO-1072 HPK2
Δ*flaA* (Kan^R)	CSO-0752 CSO-1548	CSO-1549 CSO-1550	*aphA-3* (HPK1/HPK2)	CSO-0999 CSO-1000	CSO-0756 CSO-0023
Δ*flaB* (Kan^R)	CSO-2842 CSO-2843	CSO-2844 CSO-0560	*aphA-3* (HPK1/HPK2)	CSO-0999 CSO-1000	CSO-2841 CSO-0023
Δ*flaAB* (Kan^R)	CSO-0752 CSO-1548	CSO-1549 CSO-1550	*aphA-3* (HPK1/HPK2)	CSO-0999 CSO-1000	CSO-0756 CSO-0023
Δ*rpoN* (Gm^R)	CSO-1144 CSO-1145	CSO-1146 CSO-1147	*aac*(3)-IV (HPK1/HPK2)	CSO-1144 CSO-1147	CSO-1148 HPK2
Δ*fliA* (Gm^R)	CSO-1149 CSO-1150	CSO-1151 CSO-1152	*aac*(3)-IV (HPK1/HPK2)	CSO-1149 CSO-1152	CSO-1153 HPK2
Δ*flhX* (Gm^R)	CSO-3077 CSO-3078	CSO-3079 CSO-3080	*aac*(3)-IV (HPK1/HPK2)	CSO-1149 CSO-1152	CSO-3076 HPK2
3xFLAG tags (cloned in pGG1)					
csrA-**3xFLAG** (NCTC11168 and 81-176) (pGD4-1)	CSO-0171 CSO-0172	CSO-0173 CSO-0174	*aphA-3* (from pGG1)	CSO-0172 CSO-0173	CSO-0023 CSO-0196
flaA-**3xFLAG** (pMW5.2)	CSO-0553 CSO-0554	CSO-0555 CSO-0556	*aphA-3* (from pGG1)	CSO-0554 CSO-0555	CSO-0023 CSO-0557
flgI-**3xFLAG** (pSSv1.2)	CSO-1011 CSO-1012	CSO-1013 CSO-1014	*aphA-3* (from pGG1)	CSO-1012 CSO-1013	CSO-1015 HPK2
flaB-**3xFLAG** (pMW6.1)	CSO-0558 CSO-0559	CSO-0560 CSO-0561	*aphA-3* (from pGG1)	CSO-0559 CSO-0560	CSO-0562 HPK2
3xFLAG tags (overlap PCR)					
flaG-**3xFLAG**	CSO-1002 CSO-1098	CSO-1099 CSO-1003	*aphA-3* (HPK1/HPK2)	CSO-1002 CSO-1003	CSO-1005 HPK2
Cj0529-3xFLAG	CSO-1408 CSO-1409	CSO-1410 CSO-1411	*aphA-3* (HPK1/HPK2)	CSO-1409 CSO-1410	CSO-1407 CSO-0023
For construction of *flaA* 5' UTR point mutations refer to Methods section.					

Appendix Table 16. GFP fusions for validating CsrA-target interactions in *E. coli*.

Plasmid	*C. jejuni* target gene(s)	Primers used for target amplification	Plasmid backbone	Colony PCR	Description
pGD31-1	*flaA*	CSO-0621 CSO-0622	pXG10	CSO-0621 CSO-0155	5' UTR along with first 33 codons of *flaA* fused in-frame to *gfp*
pGD111-1	*flaG*	CSO-1823 CSO-1824	pXG10	CSO-1823 CSO-0155	5' UTR along with first 25 codons of *flaG* fused in-frame to *gfp*
pGD38-1	*flgI*	CSO-0694 CSO-0695	pXG10	CSO-0694 CSO-0155	5' UTR along with first 35 codons of *flgI* fused in-frame to *gfp*
pGD109-1	*flaB*	CSO-1815 CSO-1816	pXG10	CSO-1815 CSO-0155	5' UTR along with first 35 codons of *flaB* fused in-frame to *gfp*
pGD112-1	*pseB*	CSO-1825 CSO-1826	pXG10	CSO-1825 CSO-0155	5' UTR along with first 26 codons of *pseB* fused in-frame to *gfp*
pGD110-1	Cj1249	CSO-1819* CSO-1820*	pXG10	CSO-1819 CSO-0155	5' UTR along with first 16 codons of Cj1249 fused in-frame to *gfp*
pGD28-1	Cj0310c-Cj0309c	CSO-0608 CSO-0609	pXG30	CSO-0608 CSO-0155	Last 17 codons of Cj0310c and first 23 codons of Cj0309c were fused in-frame to FLAG-*lacZ* and *gfp,* respectively.
pGD27-1	Cj0805-*dapA*	CSO-0606 CSO-0607	pXG30	CSO-0606 CSO-0155	Last 17 codons of Cj0805 and first 25 codons of *dapA* were fused in-frame to FLAG-*lacZ* and *gfp,* respectively

Sequencing on plasmids was performed using CSO-0155 which binds antisense to *gfp.*
*CSO-1819/-1820 were annealed together to yield insert (without PCR) for direct introduction into pXG10.

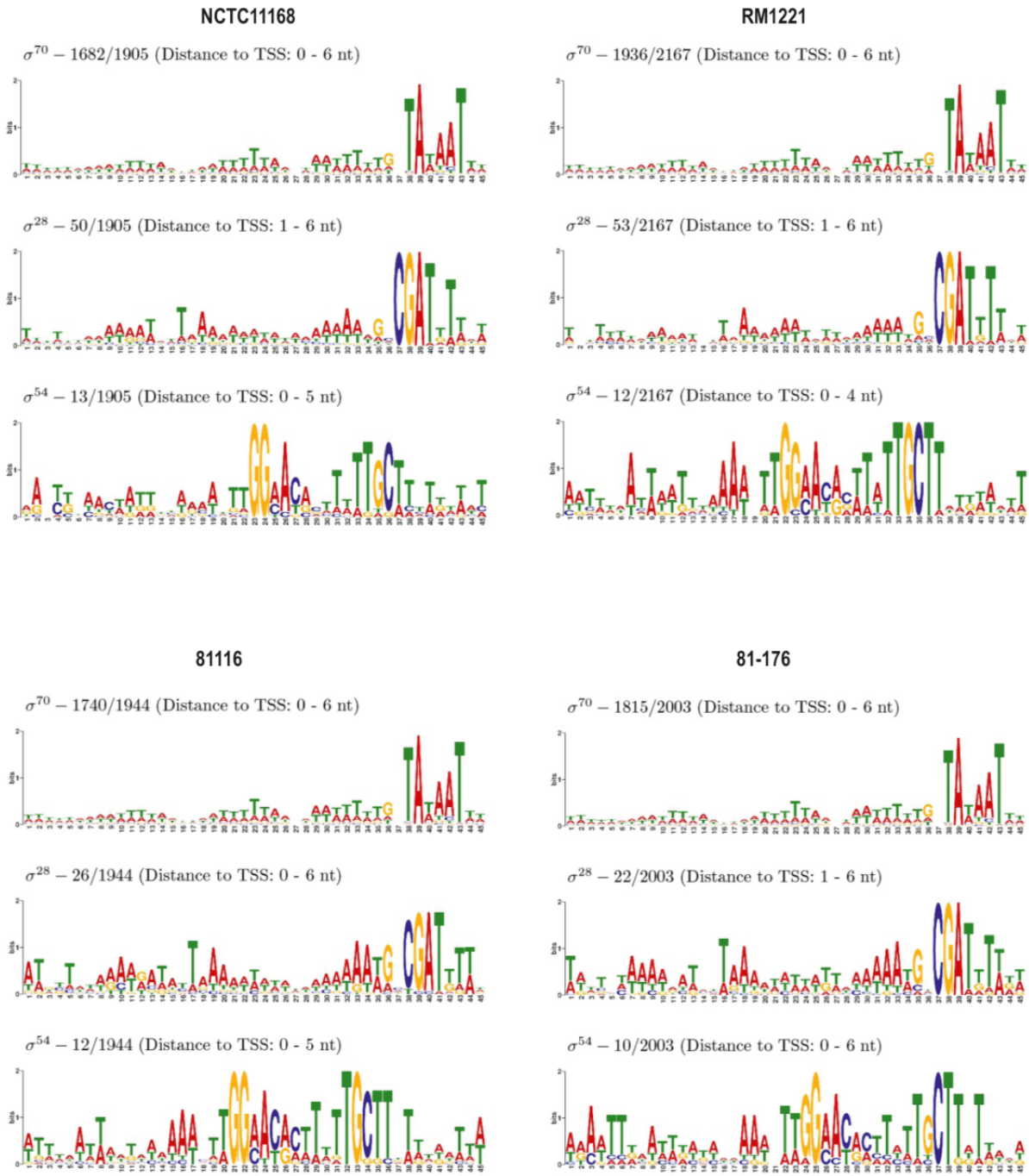
Appendix Table 17. Details of RNAs used for *in vitro* work. WT GGA motifs are marked in blue and introduced point mutations in the leader variants are marked in red. Start (ATG) and stop (TAA) codons are underlined.

Name	DNA template (plasmid or gDNA)	Primers	Size of T7-transcript [nt]	Sequence (5' → 3')
***flaA* WT leader**	gDNA NCTC11168	CSO-0709 CSO-0710	144	UAACAAGUUCAU**GGA**UGAGCUUGAAUUUUUUUAAAA**GGA**U UUAAA**AUG**GGAUUUCGUAUUAACACCAAUGUUGCAGCUUU AAAUGCAAAAGCAAACGCUGAUUUAAAUAGUAAAAGUUUAG AUGCUUCUUUAAGCAGACUUAGU
***flaA* M1 leader**	pGD92-1	CSO-1656 CSO-0710	144	UAACAAGUUCAU**AAA**UGAGCUUGAAUUUUUUUAAAA**GGA**U UUAAA**AUG**GGAUUUCGUAUUAACACCAAUGUUGCAGCUUU AAAUGCAAAAGCAAACGCUGAUUUAAAUAGUAAAAGUUUAG AUGCUUCUUUAAGCAGACUUAGU
***flaA* M2 leader**	pGD77-1	CSO-1081 CSO-0710	144	UAACAAGUUCAU**UGA**UGAGCUUGAAUUUUUUUAAAA**GGA**U UUAAA**AUG**GGAUUUCGUAUUAACACCAAUGUUGCAGCUUU AAAUGCAAAAGCAAACGCUGAUUUAAAUAGUAAAAGUUUAG AUGCUUCUUUAAGCAGACUUAGU
***flaA* M3 leader**	pGD93-1	CSO-0709 CSO-0710	144	UAACAAGUUCAU**GGA**UGAGCUUGAAUUUUUUUAAAA**GGG**U UUAAA**AUG**GGAUUUCGUAUUAACACCAAUGUUGCAGCUUU AAAUGCAAAAGCAAACGCUGAUUUAAAUAGUAAAAGUUUAG AUGCUUCUUUAAGCAGACUUAGU
***flaA* M2/M3 leader**	pGD95-1	CSO-1081 CSO-0710	144	UAACAAGUUCAU**UGA**UGAGCUUGAAUUUUUUUAAAA**GGG**U UUAAA**AUG**GGAUUUCGUAUUAACACCAAUGUUGCAGCUUU AAAUGCAAAAGCAAACGCUGAUUUAAAUAGUAAAAGUUUAG AUGCUUCUUUAAGCAGACUUAGU
***flgI* leader**	gDNA NCTC11168	CSO-0713 CSO-0714	128	ACAAUAGAUUAAAGGAGAAUCC**AUG**AGAGUUUUAACGAUA UUUUUACUCUUUAUGACAAGCAUUUUUGCAGUGCAAAUCAA GGAUGUAGCAAAUACUGUAGGUGUAAGAGAUAACCAACUU AUAGGU
***flaG* leader**	gDNA NCTC11168	CSO-1084 CSO-1085	108	ACUAGCAAUAGGAAAUUUUAAAAAGGAUUUUAAA**AUG**GAA UAUCGAAGGCAAAUGGGCAAAUGGAUACAGCUUUGGCAAA CAUUAGCCAAAGAACAAGUGAGACACA
***flaB* leader**	gDNA NCTC11168	CSO-1817 CSO-1818	132	CGAUGCAAUAUUUUGAAAGGAUUUAAA**AUG**GGUUUUAGGA UAAACACCAACAUCGGUGCAUUAAAUGCACAUGCAAAUUCA GUUGUUAAUGCUAGAGAACUGGAUAAGUCUUUAAGCAGAC UUAGUUCAGGU
Cj0040 leader	gDNA NCTC11168	CSO-1082 CSO-1083	117	AAAUUAAAAUUUUAAAAAGGAAGUUAAA**AUG**UCAAAACCAU UAAAUGAAGAGAUUUUUGUUGAAUUUAAAAGUGAUCUAGCU GAAAGAAAAAAUGAAGUUUUGCUUCAAGUUUUAGA
***flgA* 3'end**	gDNA NCTC11168	CSO-1088 CSO-1089	113	CAAAUGUUGAUGUUUUAAUCGAACUUGUGGCUUUGCAAAG UGCAAAUAUGGGCGAAAGGAUUCGUGCAAAAAACAAAGAAG GUAAAGUUAUGCAAGGUAUCGUAGUGGGUAAA
***flgM* leader**	gDNA NCTC11168	CSO-1092 CSO-1093	97	AGGAUUAUAACUAAGAUCAAGGAGGCAGAA**AUG**AUCAAUCC UAUACAACAAAGUUAUGUGGCAAAUACCGCAUUAAAUACAA AUAGAAUAGAUAAAG
Cj1324 leader	gDNA NCTC11168	CSO-1665 CSO-1666	100	AUUUUUAUUAAAUUGAAGGGGUGGGGAGGA**AUG**AUUUAUUGUG AUCACUGCGUGAUGCCAAAUACUAGACCUGGUAUUAAUUU UACAAAAGAUAAAGAAGGUA
***hopB* 3'end with UTR**	gDNA *H. pylori* G27	CSO-0701 CSO-0702	107	AAAGCUGGUGGCGCUGAAGUGAAAUACUUCCGCCCUUAUA GCGUGUAUUGGGUCUAUGGCUACGCCUUC**UAA**AAAAGCUC AAGGCCUUUUAUAGGCUUUGAUUUAAC

Appendix Table 18. *flaA*_mini and Cj1321_mini transcript expressed in *C. jejuni*. *flaA*_mini and Cj1321_mini were introduced in rdxA locus of *C. jejuni* NCTC11168 driven by *flaA* promoter (in lower case). The sequence in grey is from the pBAD/Myc-His A plasmid retaining the *rrn*B T1 and T2 transcriptional terminator sequence. The start codon are marked in red.

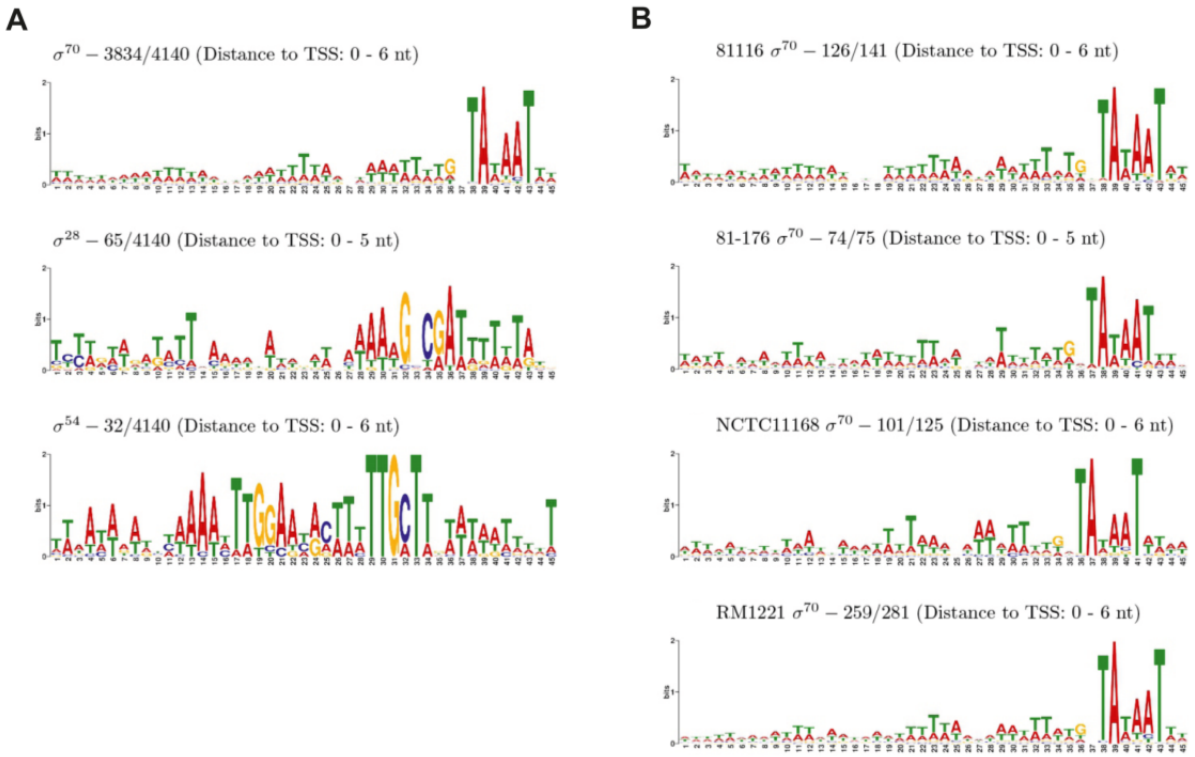
Name	Sequence (5' → 3')
*flaA*_mini	tataaaatatttttttgattgcacgatatagcattTAACAAGTTCATGGATGAGCTTGAATTTTTTTAAAAGGATTTAAAATGGGATTTC GTATTAACACCAATGTTGCAGCTTTAAATGCAAAAGCAAACTCTAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACA GAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACTCAGAAGTGAAACGCCGTAGCGCCGA TGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCGAAAGACTG GGCCTT
Cj1321_mini	tataaaatatttttttgattgcacgatatagcattTGGGATTTGATTTTTTTACTTCTTCTTTTAGGAGAAGTACGTGAAAAATGTCATC ATCATAGGAGCGGGTGGCTTTACTAGAGCAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCC GAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGGCATC AAATAAAACGAAAGGCTCAGTCGAAAGACTGGGCCTT

Appendix Figure 1



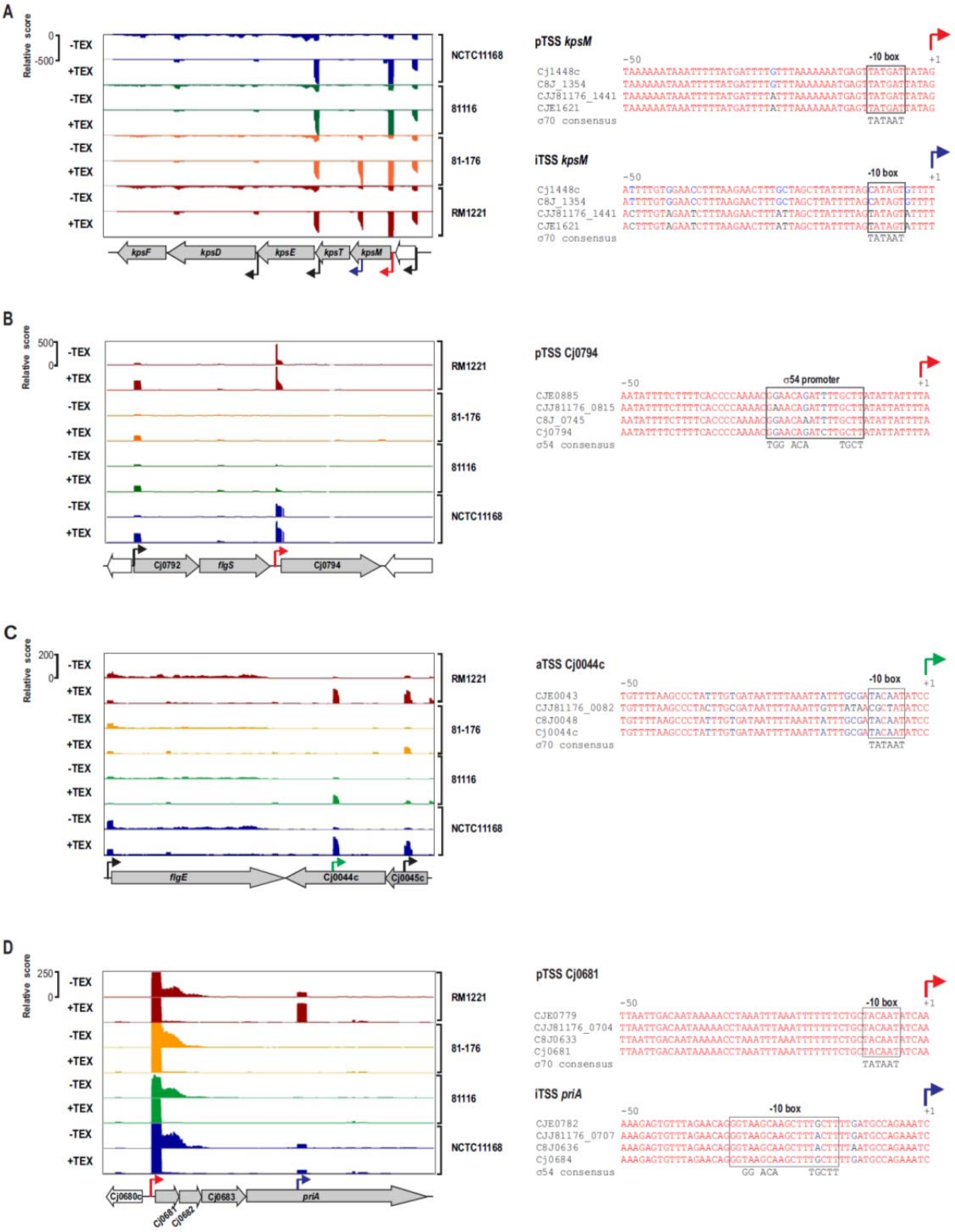
Promoter motifs detected for TSS of the four individual strains. Promoter motifs detected for the three sigma factors, σ^{70}, σ^{28} and σ^{54} in the -50 to +1 sequences of the TSS for each of the four *C. jejuni* strains. The number of occurrences for each motif is indicated. The search was performed using *MEME* and a fixed motif size of 45 nt. The range of distances to the TSS for each motif is given. The results are based on 1,905 TSS upstream sequences of *C. jejuni* NCTC11168, 2,167 of *C. jejuni* RM1221, 1,944 of *C. jejuni* 81116 and 2,003 of *C. jejuni* 81-176.

Appendix Figure 2

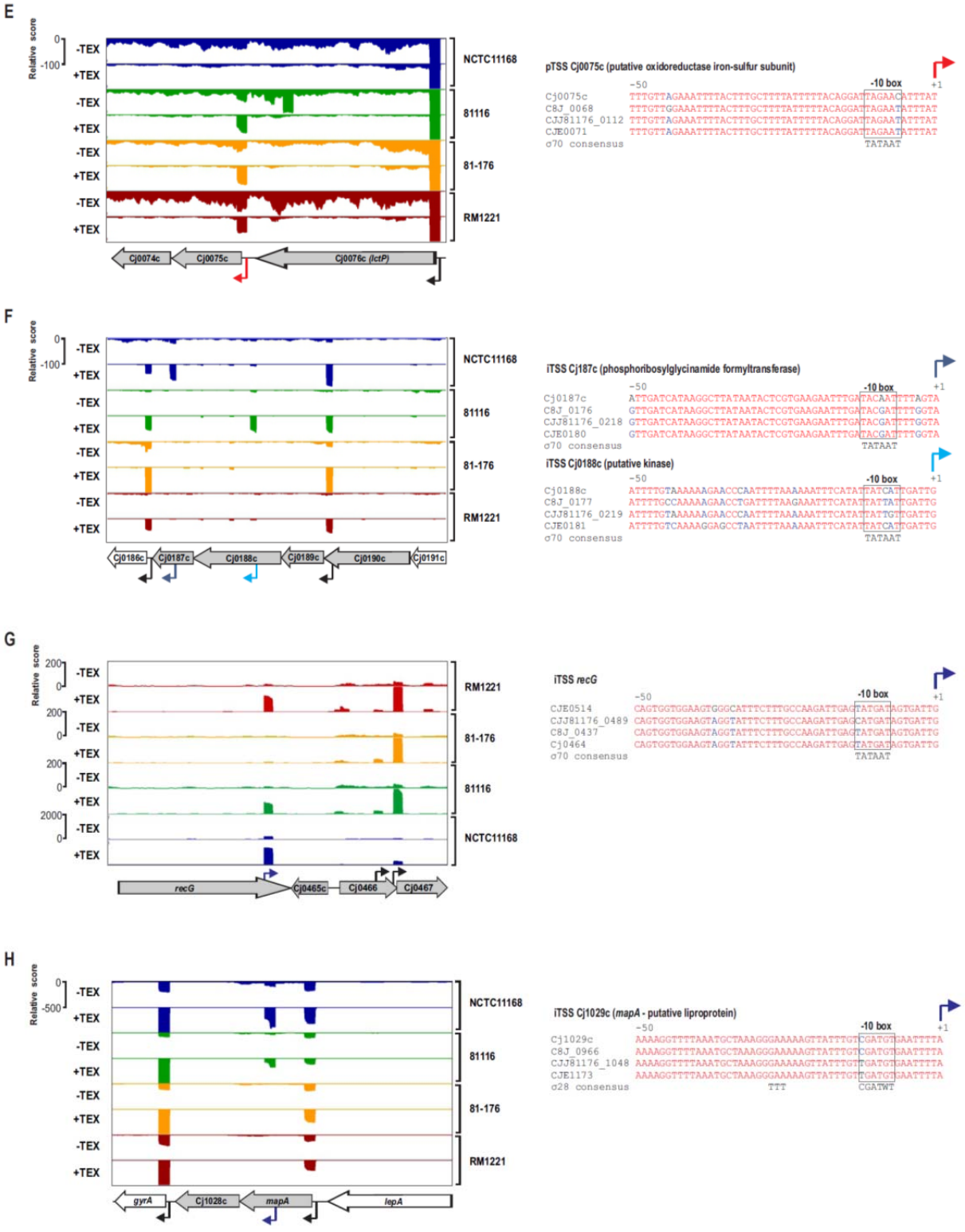


Promoter Motifs detected by *MEME* for conserved and strain-specific TSS. (A) Promoter motifs detected by *MEME* searches for the three sigma factors, σ^{70}, σ^{28} and σ^{54}, which were detected in the -50 to +1 sequences upstream of conserved TSS, i.e. TSS which were detected in all four strains. The number of hits, the total number of analyzed sequences (4,140 TSS in total which corresponds to 1,035 TSS in each strain) and the range of distances of the motif to the TSS are indicated. The motif size was set to 45 nt. **(B)** σ^{70} promoter motifs found in the -50 to +1 sequences upstream of the TSS, which were detected only in one strain. The number of sequences containing the motif and the total number of analyzed sequences as well as the distance range of the motif to the TSS are indicated. The motif size was set to 45 nt.

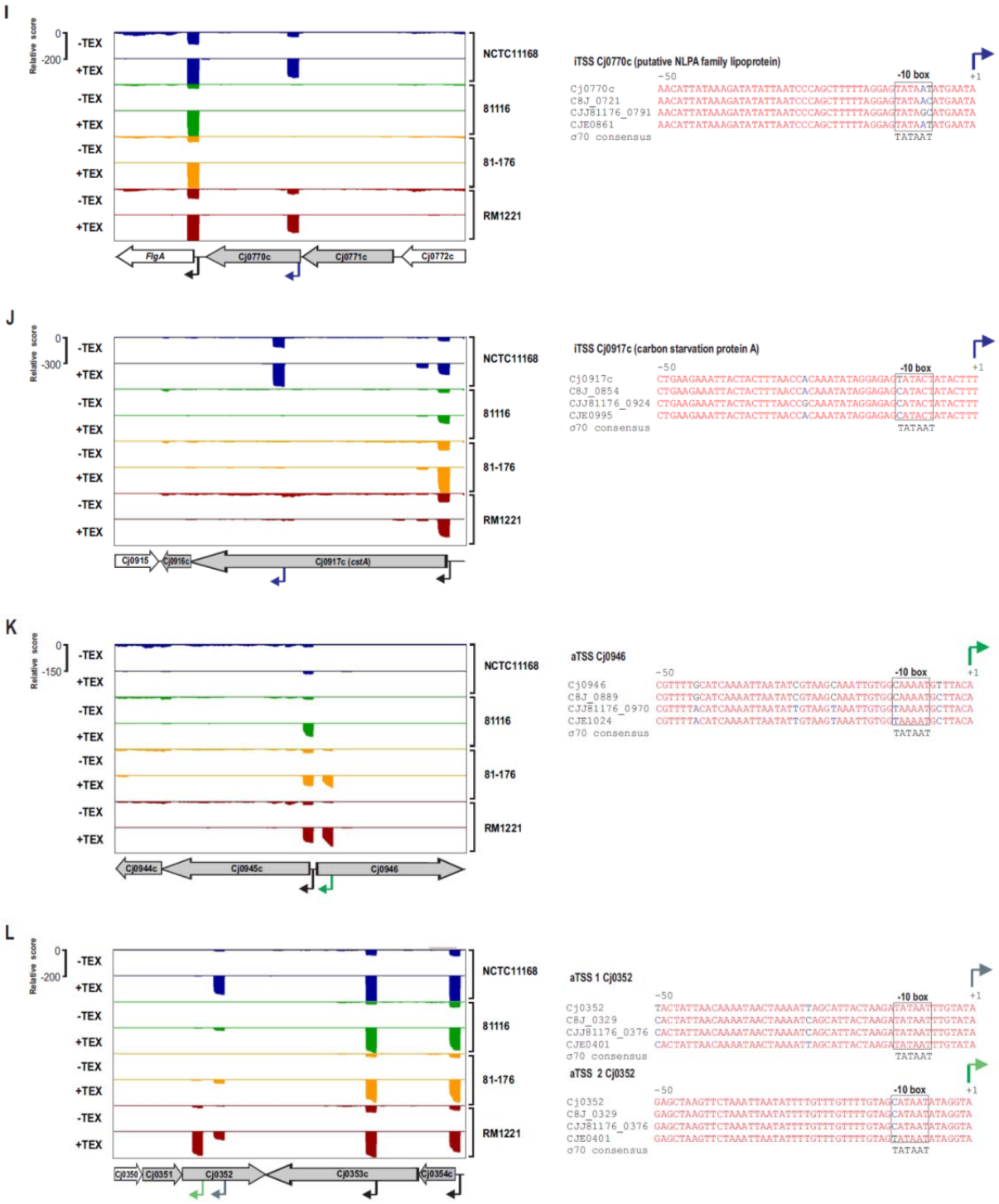
Appendix Figure 3A-D



Appendix Figure 3E-H



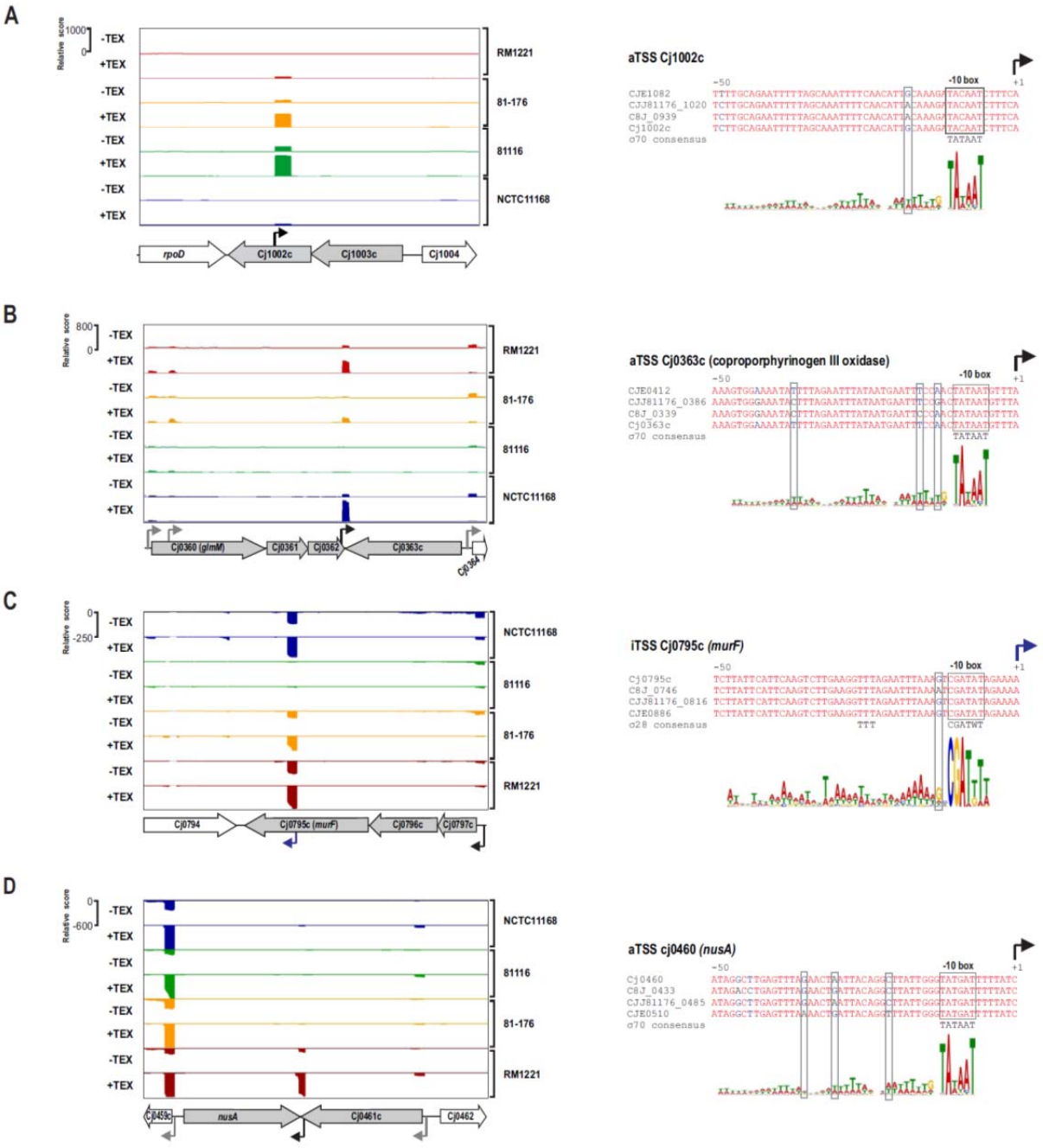
Appendix Figure 3I-L



dRNA-seq data and promoter alignments for genes with differentially expressed TSS among four *C. jejuni* strains. *(Left)* dRNA-seq reads mapped to different regions of the SuperGenome with differential TSS usage among strains. *(Right)* Alignments of the -50 to +1 regions upstream of the TSS shown in the dRNA-seq data on the left for the four different *C. jejuni* strains. **(A)** dRNA-seq data of the four strains mapped onto the region corresponding to the capsule export genes *kpsMTEDF*. Black arrows indicate TSS that were detected within this region based on enrichment in the TEX+ libraries of all four strains. The primary TSS of *kpsM* (red arrow) is enriched in all four strains. In contrast, the internal TSS within *kpsM* (blue arrow) is present in only two strains

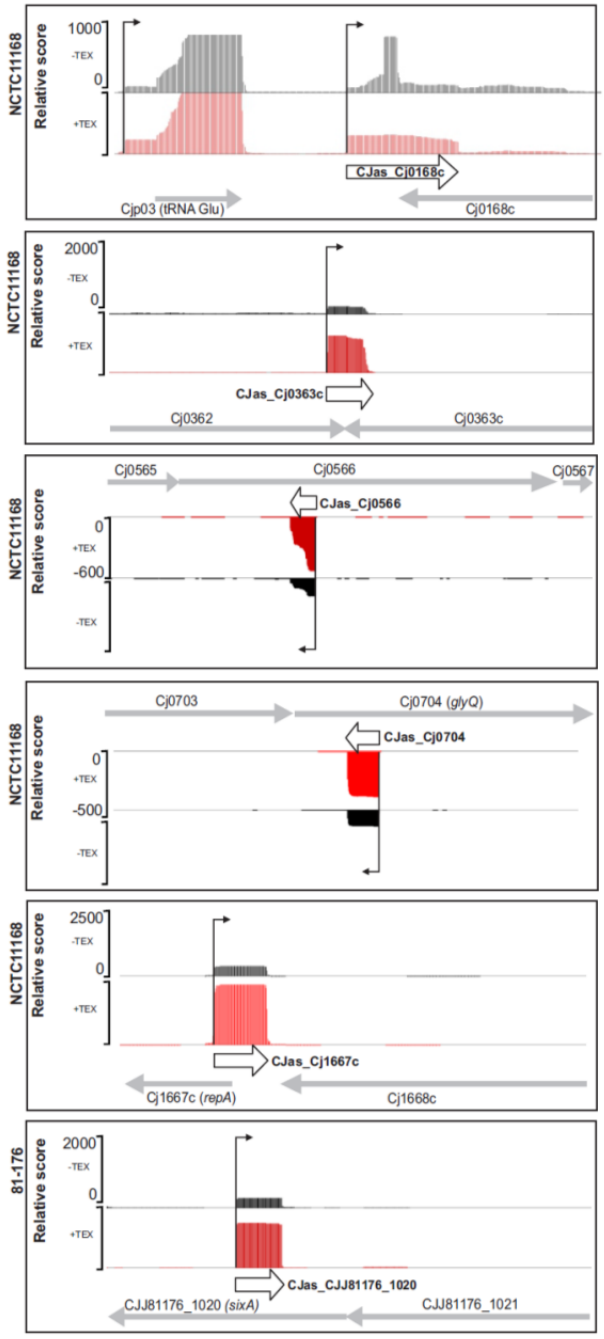
(81-176 and RM1221), probably due to point mutations in the respective promoter region of NCTC11168 and 81116 as visible in the alignment on the right**. (B)** dRNA-seq reads mapped to *Cj0794,* which encodes a hypothetical protein, indicate differential expression of its primary TSS. The promoter region with the non-detected TSS in strain 81-176 shows mutations within a σ^{54} consensus sequence. **(C)** A *cis*-encoded antisense RNA to *Cj0044c* is differentially expressed among the strains and has a disrupted σ^{70} -10 box in strain 81-176. **(D)** Differential expression of an internal TSS within *priA* coincides with mutations in a σ^{54} box in strains 81-176 and 81116. **(E)** Differential expression of the pTSS of Cj0075c due to a T to C exchange at the conserved last "T" of the σ^{70} -10 box in strain NCTC11168. **(F)** Two internal TSS were detected in *Cj0187c* and *Cj0188c*. These iTSS (*Cj0187c*, dark blue; *Cj0188c*, light blue) are only transcribed in strains NCTC11168 or 81116. **(G)** Differential expression of an iTSS at the 3' end of *recG* which shows a T to C exchange at the conserved first "T" residue of the σ^{70} -10 box in 81-176. Note that the scale for NCTC11168 has been changed to 0-2,000 due to high levels of this internal transcript compared to 81116 and RM1221. **(H)** Differential expression of an iTSS in *mapA* is probably due to disruption of a putative σ^{28} box in 81-176 and RM1221. **(I)** Differential expression of an iTSS at the 5' end of cj0770c coincides with disruption of the σ^{70} -10 box in strains 81116 and 81-176. **(J)** An iTSS within *cstA* is only transcribed in NCTC11168 which carries an intact σ^{70} -10 box upstream. **(K)** A *cis*-encoded antisense RNA (green arrow) at the 5' end of *Cj0946* is transcribed only in strains 81-176 and RM1221. **(L)** Two *cis*-encoded antisense RNAs to C*j0352* with differential expression among strains*.* The aTSS 1 (dark green arrow) along with its promoter is conserved in all strains. In contrast, aTSS 2 (green arrow) is only transcribed in strain RM1221, whereas all other strains have a point mutation in the -10 box.

Appendix Figure 4



Differentially expressed TSS with SNPs in the periodic A/T-rich pattern or in the extended -10 promoter box. dRNA-seq reads mapped to different loci in the genome. **(A)** *(Left)* dRNA-seq reads mapped antisense to the *Cj1003c-Cj1002c* operon reveal a novel conserved *cis*-encoded antisense RNA to *Cj1002c*. (*Right*) The promoter corresponding to this asTSS has a conserved σ^{70} -10 box but shows an A to G exchange in the A/T rich cyclic pattern upstream of the -10 box in strains RM1221 and NCTC11168, which apparently leads to loss of transcription from this aTSS in these strains. **(B)** A differentially expressed *cis*-encoded antisense RNA to *Cj0363c* (coproporphyrinogen III oxidase) shows mutations in the A/T-rich upstream region in strain 81116. **(C)** A differentially transcribed iTSS in *Cj0795c (murF)* is not transcribed in strain 81116 probably due to a mutation at a conserved G residue two nucleotides upstream of a σ^{28} box. **(D)** A differentially expressed *cis*-encoded antisense RNA to *Cj0460 (nusA)* is only expressed in strain RM1221 and carries point mutations in the A/T-rich upstream region in the three other strains.

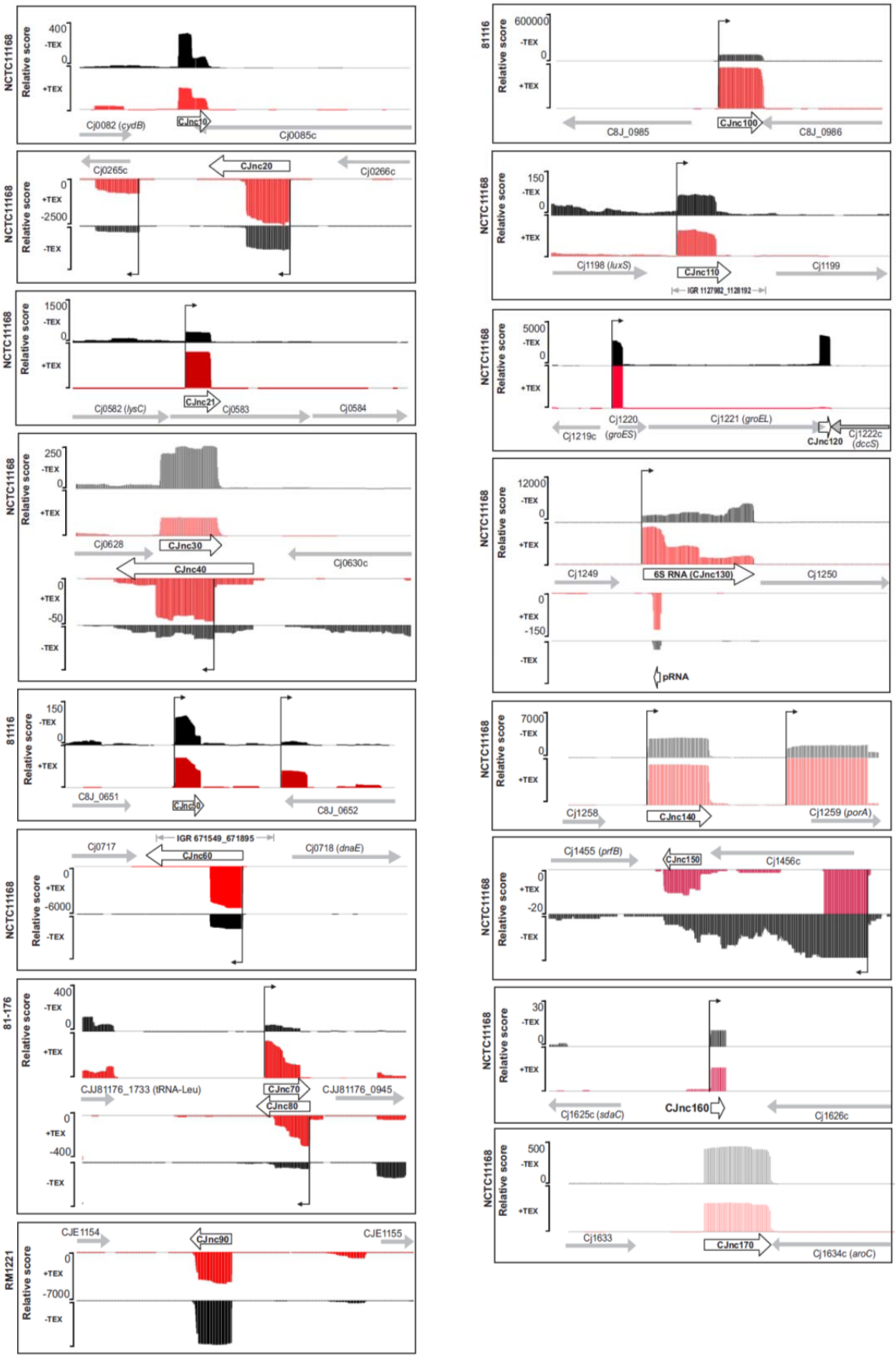
Appendix Figure 5



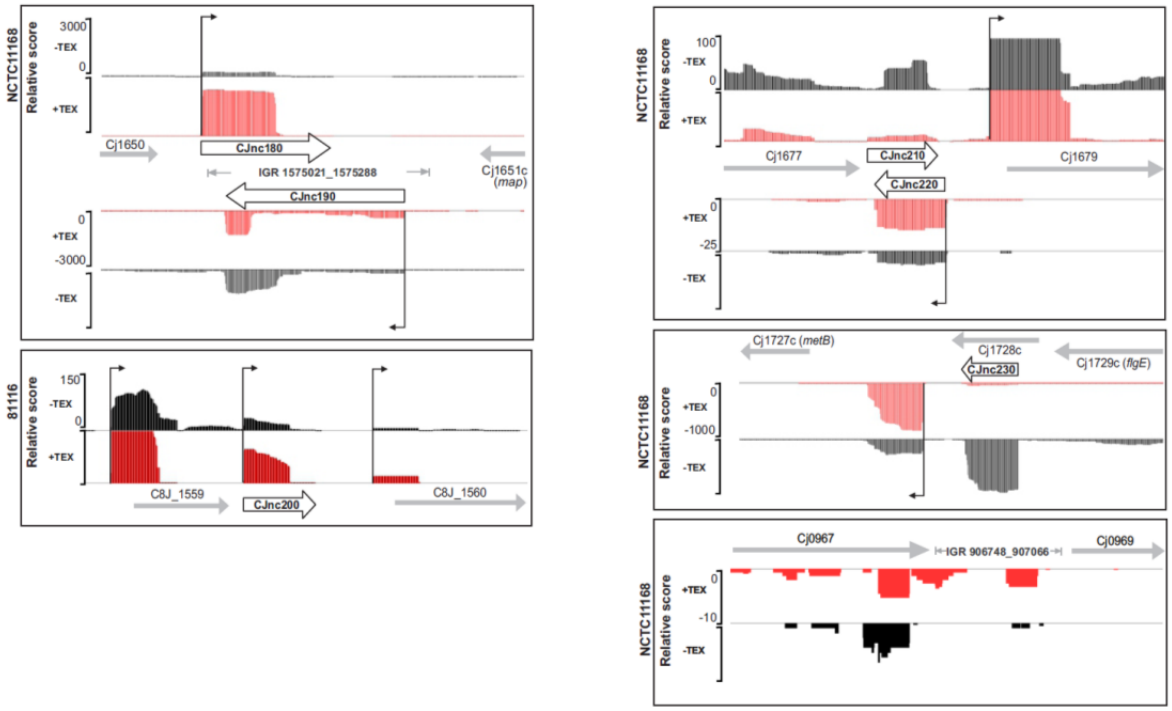
Examples for genes with *cis*-encoded antisense RNAs. Sequenced reads of cDNA libraries derived from -/+ TEX-treated total RNA mapped to several antisense RNA loci in the chromosome of different strains of *C. jejuni.* Strain names are indicated in the left of the screen shot. The antisense transcripts were named CJas_"X" according to the name of gene "X" encoded on the opposite strand. The selected examples had at least 100 cDNA reads in the TEX- library.

Appendix Figure 6

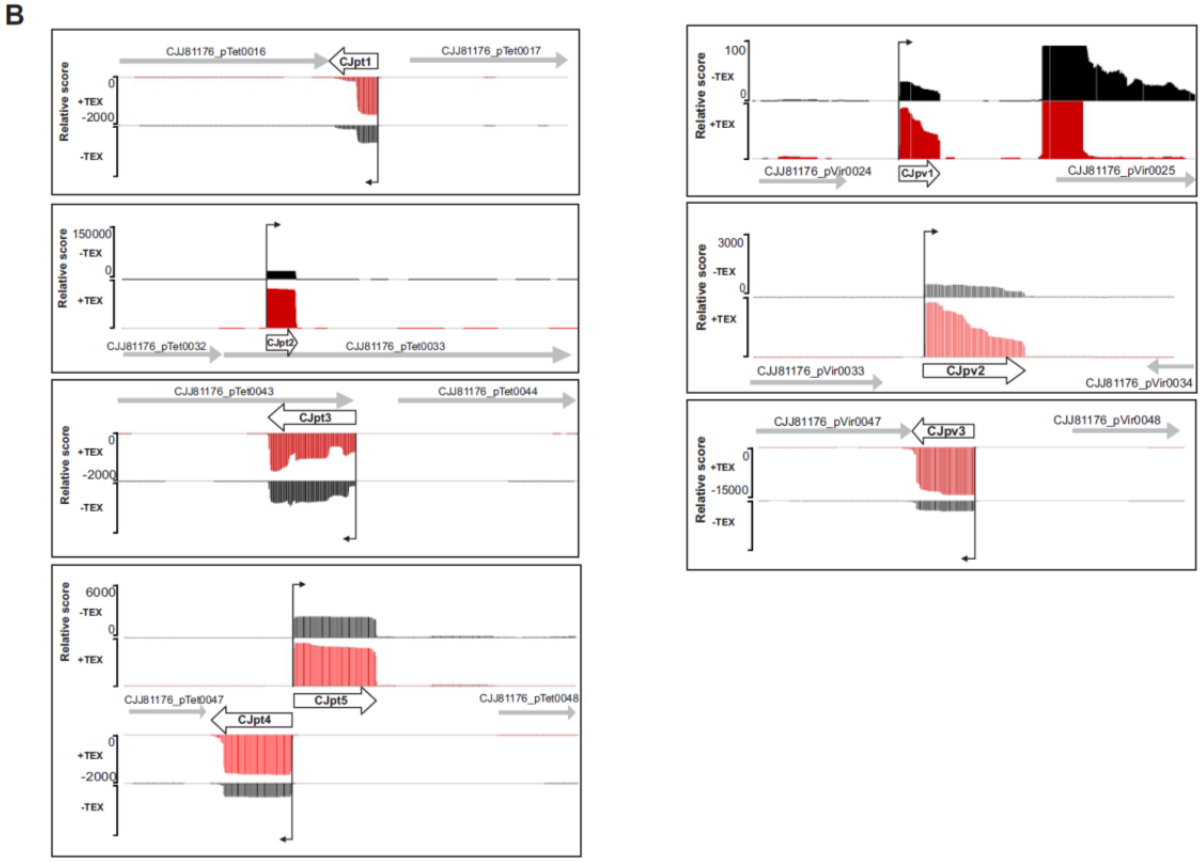
A



Appendix Figure 6A (cont.)-B

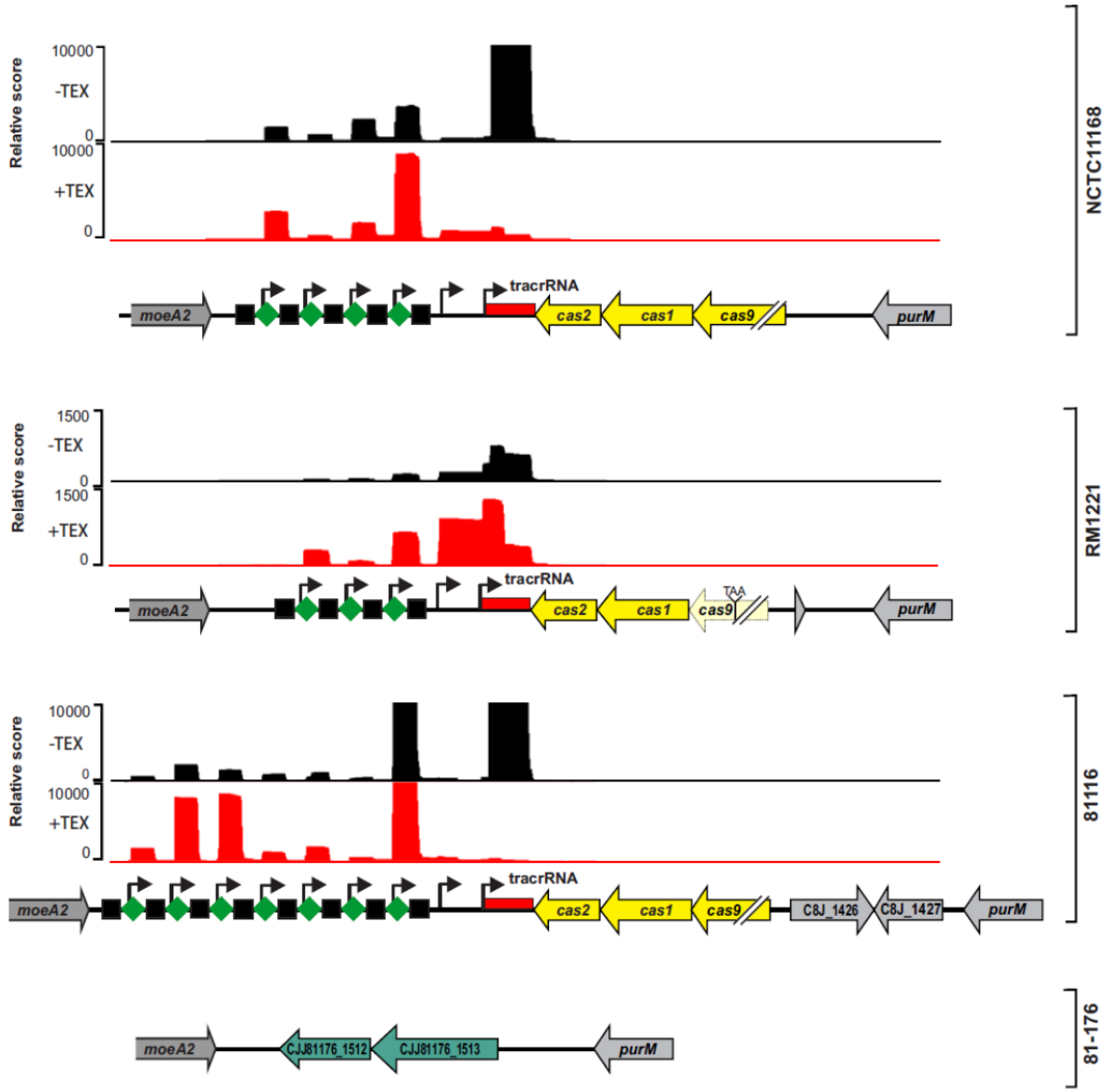


B



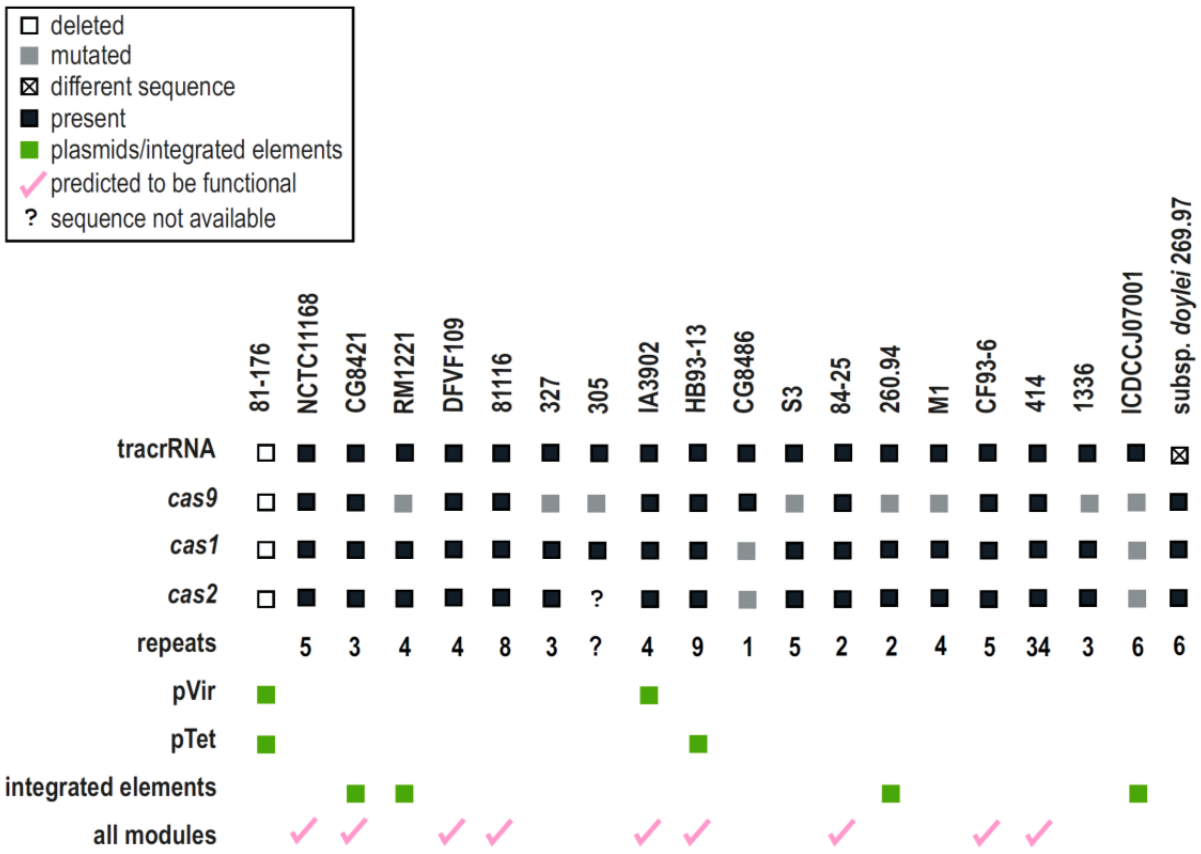
dRNA-seq reveals several sRNA candidates in *C. jejuni.* Sequenced reads of cDNA libraries derived from -/+ TEX-treated total RNA mapped to candidate sRNA loci. Regarding nomenclature, sRNA candidates were termed "CJncXXX" and numbered in steps of ten according to the genome position of the TSS of putative small RNA. Exact coordinates are listed in Appendix Table 8. Steps of ten were chosen to accommodate additional sRNAs. TSS with enrichment in the TEX+ library are indicated by black arrows. **(A)** Screen-shots of cDNA reads mapped to sRNAs loci in the chromosomes of different *C. jejuni* strains. The respective strain name is indicated on the left of each screen shot. Note that IGR regions with potential sRNAs candidates identified by Chaudhuri *et al.* [85] which overlap with sRNA candidates identified in this study or for which cDNA reads were detected are also indicated in the screenshots. Most sRNAs are transcribed from their own TSS. However, for example a sRNA candidate, CJnc120, downstream of *groEL* seems to be generated by processing from the 3'end of the *groEL* mRNA since it is enriched in the TEX- library. Moreover, also CJnc230 seems to be generated by processing. **(B)** Screen shots of dRNA-seq reads mapped to plasmid encoded sRNA loci in the pTet and the pVir plasmids from 81-176. TSS of sRNAs are indicated by black arrows. Exact coordinates are listed in Appendix Table 8.

Appendix Figure 7



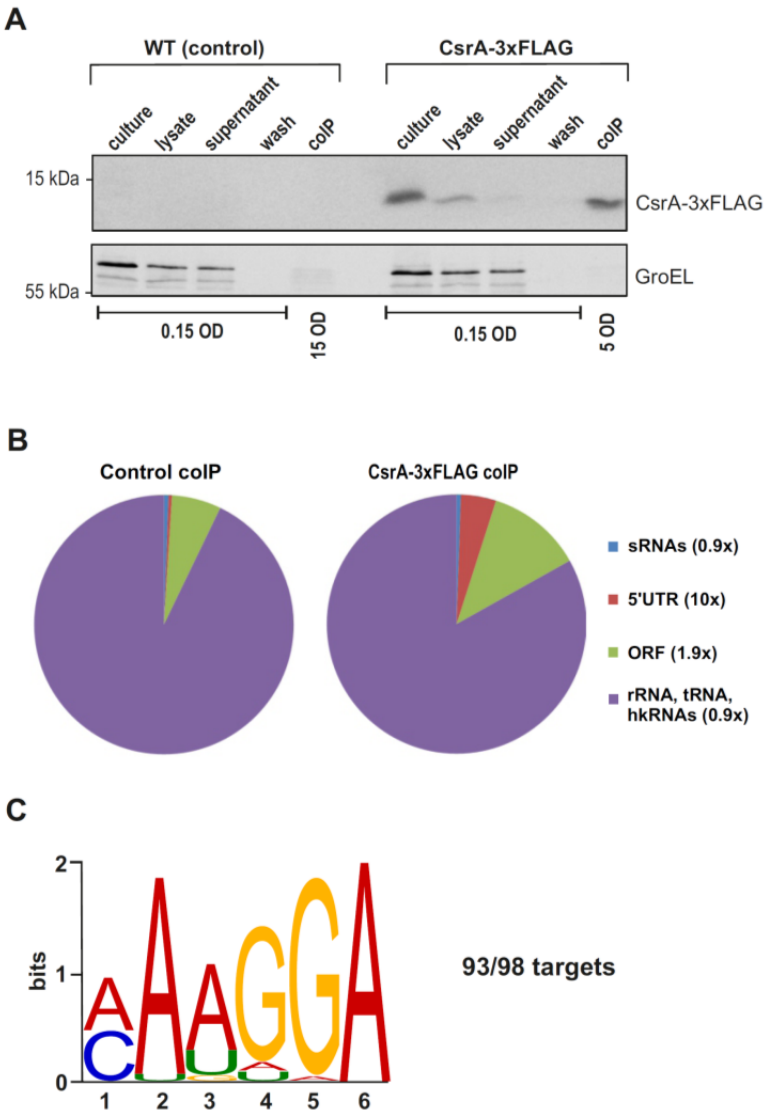
Transcription and genomic organization of type-II C CRISPR/*cas* (Nmeni/CASS4) loci in *C. jejuni*. Sequenced cDNAs reads derived from -/+ TEX-treated total RNA that were mapped to the CRISPR loci in *C. jejuni* NCTC11168, RM1221, and 81116. Black squares indicate the CRISPR repeats, green diamonds the CRISPR spacers, and the red rectangle the tracrRNA gene, respectively. Enrichment of the 5'ends of mature crRNAs in the TEX+ treated cDNA library indicates the presence of a TSS within each spacer in the CRISPR locus. The processed TracrRNA is enriched in the untreated cDNA library, whereas enrichment of the 5'-end of the TracrRNA precursor shows the presence of a promoter upstream of the tracrRNA gene. Locations of the crRNA promoters and the tracrRNA promoter are indicated by black arrows. The CRISPR *cas* genes (*cas9-cas1-cas2*) are encoded in opposite direction to tracrRNA and the CRISPR spacer-repeat array. Note that the CRISPR/*cas* locus in *C. jejuni* 81-176 is replaced by two A/T-rich genes encoding for hypothetical proteins.

Appendix Figure 8



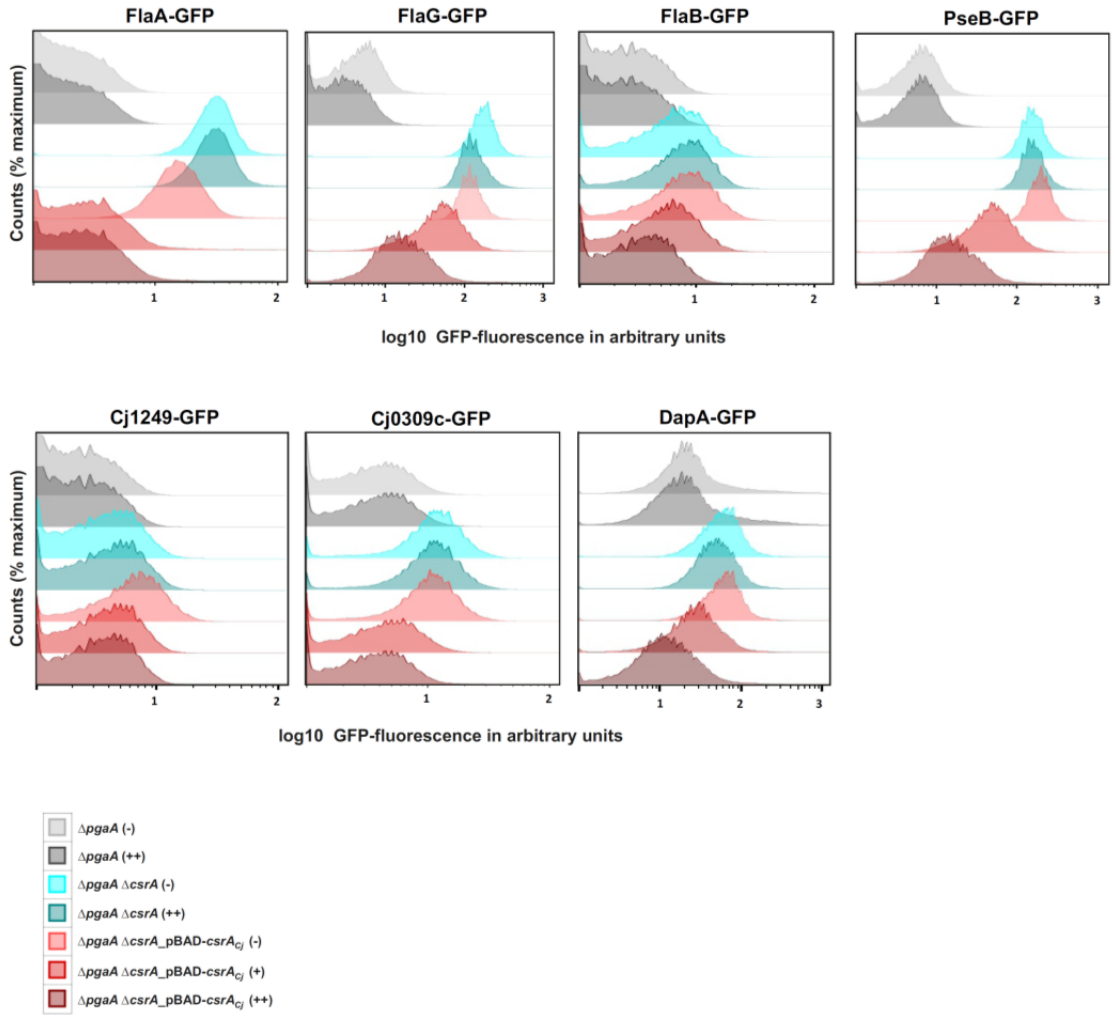
Correlation analysis of intact CRISPR loci and the presence/absence of plasmids or integrated elements in diverse *C. jejuni* clinical isolates. Only nine out of twenty *C. jejuni* strains with available genome sequence are predicted to have functional CRISPR/Cas modules and TracrRNA sequences based on conservation analysis. Interestingly, in the majority of strains with predicted functional CRISPR/Cas and tracrRNA genes no extra-chromosomal elements can be found. However, for three strains, *C. jejuni* CG8421, IA3902, and HB93-13, that seem to have functional CRISPR/Cas loci, also plasmids and integrated elements are present. Intact TracrRNA homologs can be predicted for nine genomes that harbor only deleted or mutated CRISPR/Cas loci. This indicates that there is no correlation between the presence of a conserved TracrRNA and a functional CRISPR/Cas locus in those species. For strain *C. jejuni* subsp. *doylei,* the TracrRNA sequence is not conserved, indicating that the CRISPR/Cas system might not be active in this strain.

Appendix Figure 9



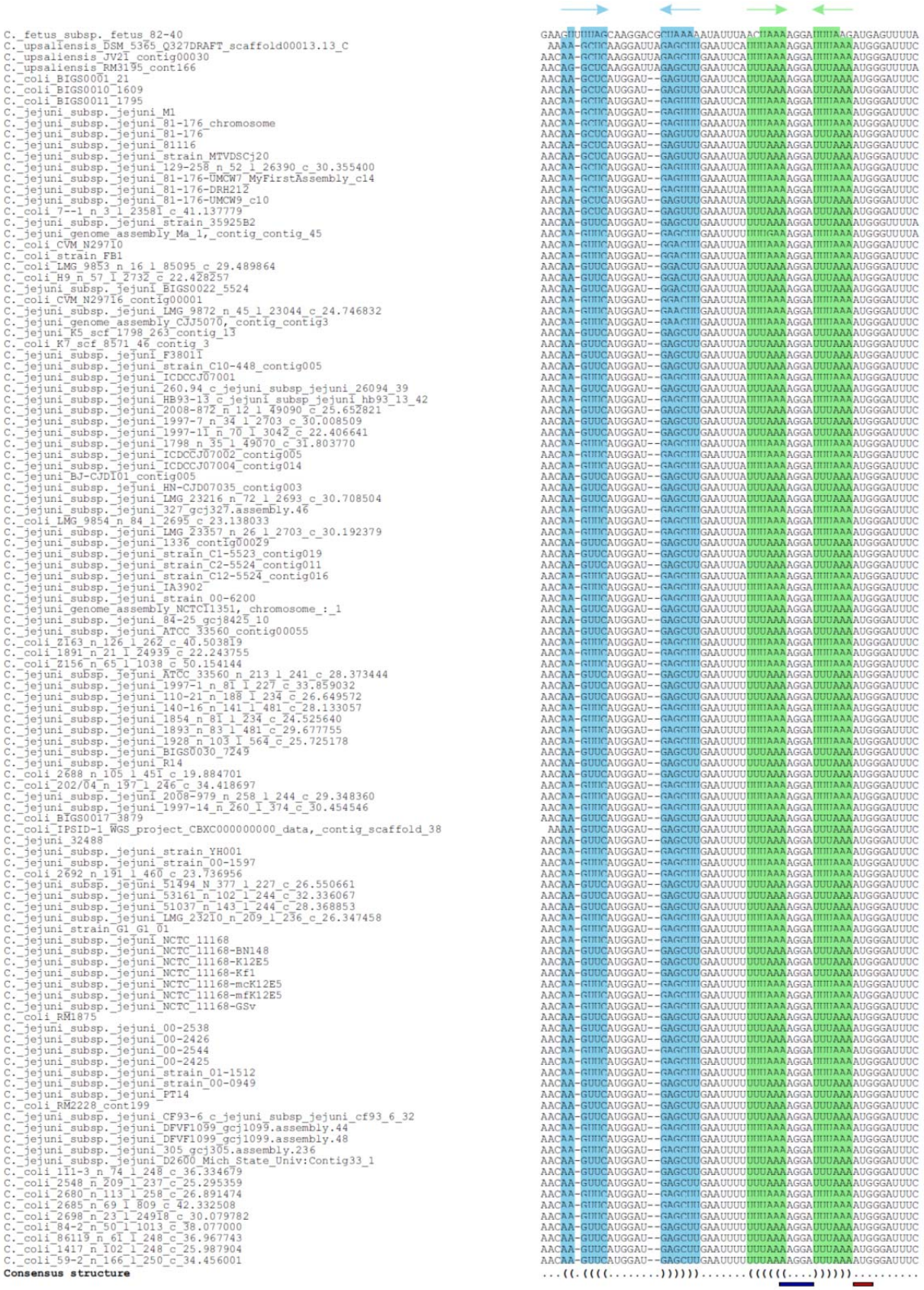
RIP-seq of CsrA-3xFLAG in strain 81-176. (A) Western blot of protein samples from CsrA-3xFLAG and control coIPs from strain 81-176. The protein amount loaded in each lane corresponds to the OD_{600} of cells as indicated below. GroEL served as a loading control. **(B)** Pie charts showing the relative proportions of mapped cDNAs of different RNA classes in the control and CsrA-3xFLAG coIP libraries from strain 81-176. Values in brackets for each RNA class denote its relative enrichment in the CsrA-3xFLAG vs. control coIP. **(C)** Consensus motif for CsrA determined by MEME using peak sequences enriched more than 5-fold in the CsrA-3xFLAG coIP from *C. jejuni* strain 81-176.

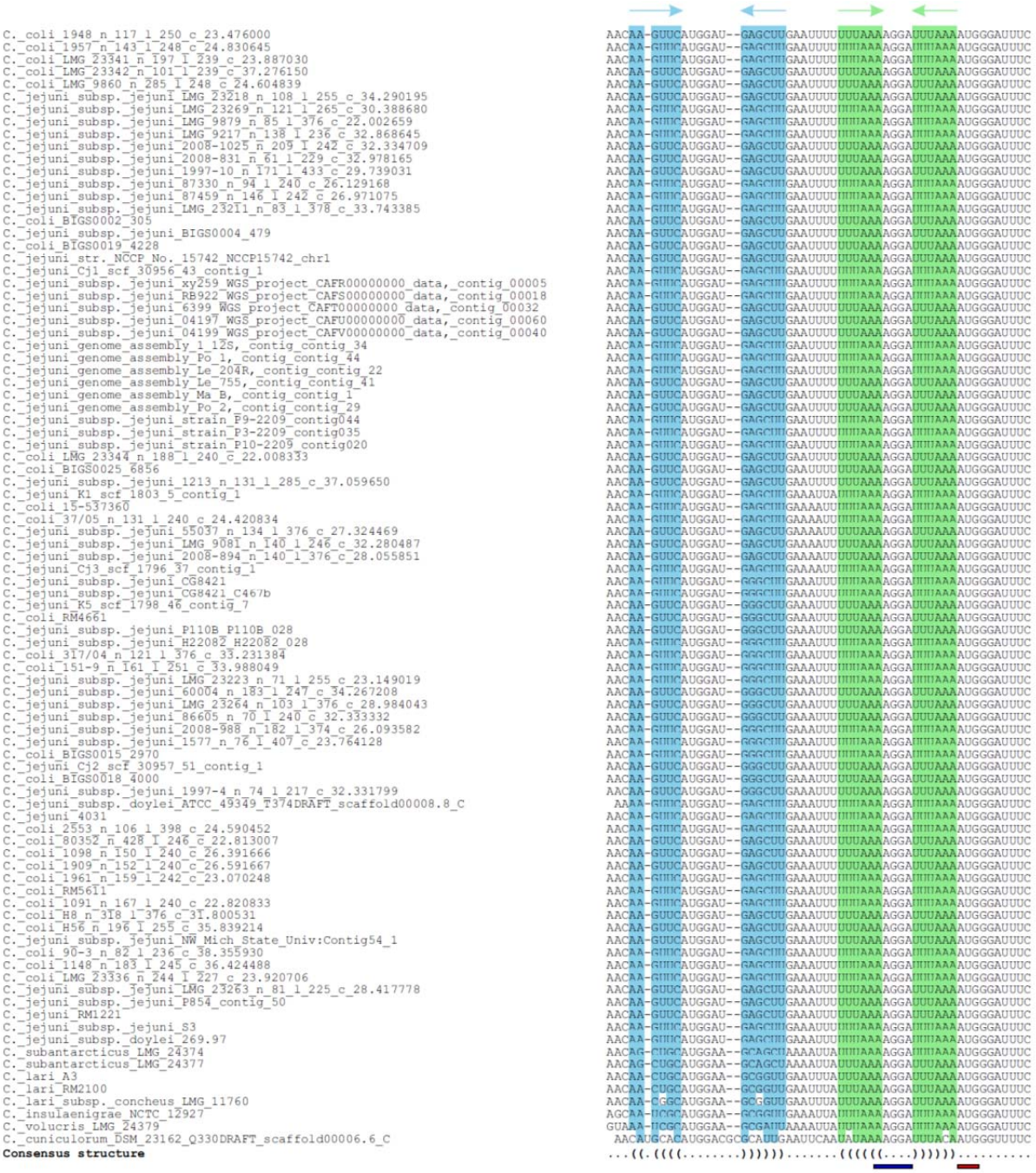
Appendix Figure 10



FACS analysis of GFP reporter fusions. pXG10- or pXG30-based GFP reporter plasmids from Fig. 24 anf Fig. 25 were introduced into *E. coli* strains Δ*pgaA,* Δ*pgaA/*Δ*csrA,* and Δ*pgaA/*Δ*csrA* complemented with arabinose-inducible *C. jejuni csrA*-Strep. All strains were grown to late log phase in LB or LB supplemented with 0.001% (+) or 0.003% (++) L-arabinose, and GFP levels were measured by flow cytometry. Data acquired in each experiment is plotted in fluorescence histograms generated from all events measured (50,000 events). Cellular fluorescence is given in arbitrary units (GFP intensity). Regulation by CsrA is visible as a shift of the peak of fluorescence curves to the right (higher GFP intensity) in the Δ*pgaA/*Δ*csrA* background and a shift to the left upon complementation with CsrA-Strep. Please note that levels of FlgI-GFP were not detectable in FACS due to low expression or fluorescence of this fusion.

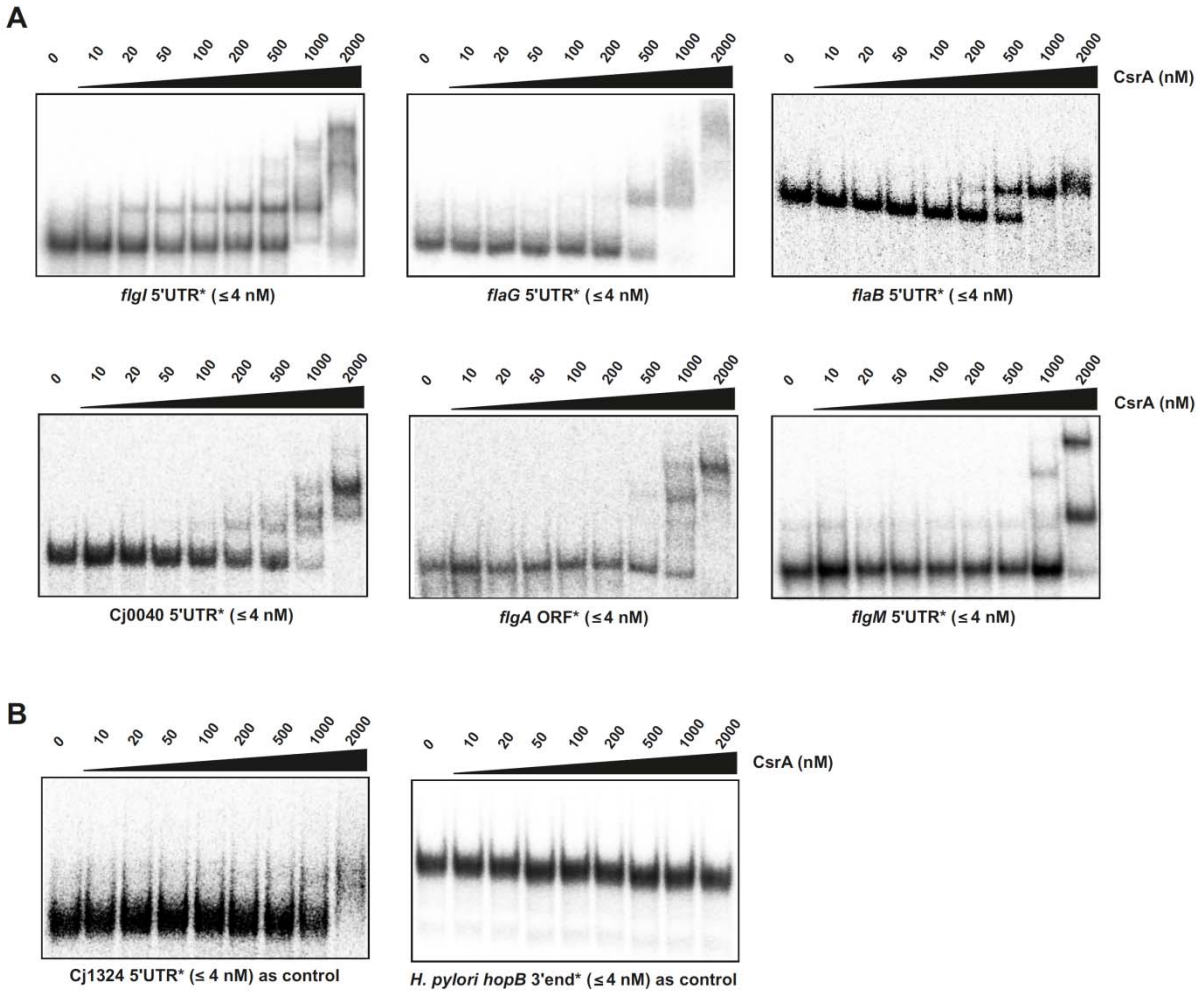
Appendix Figure 11





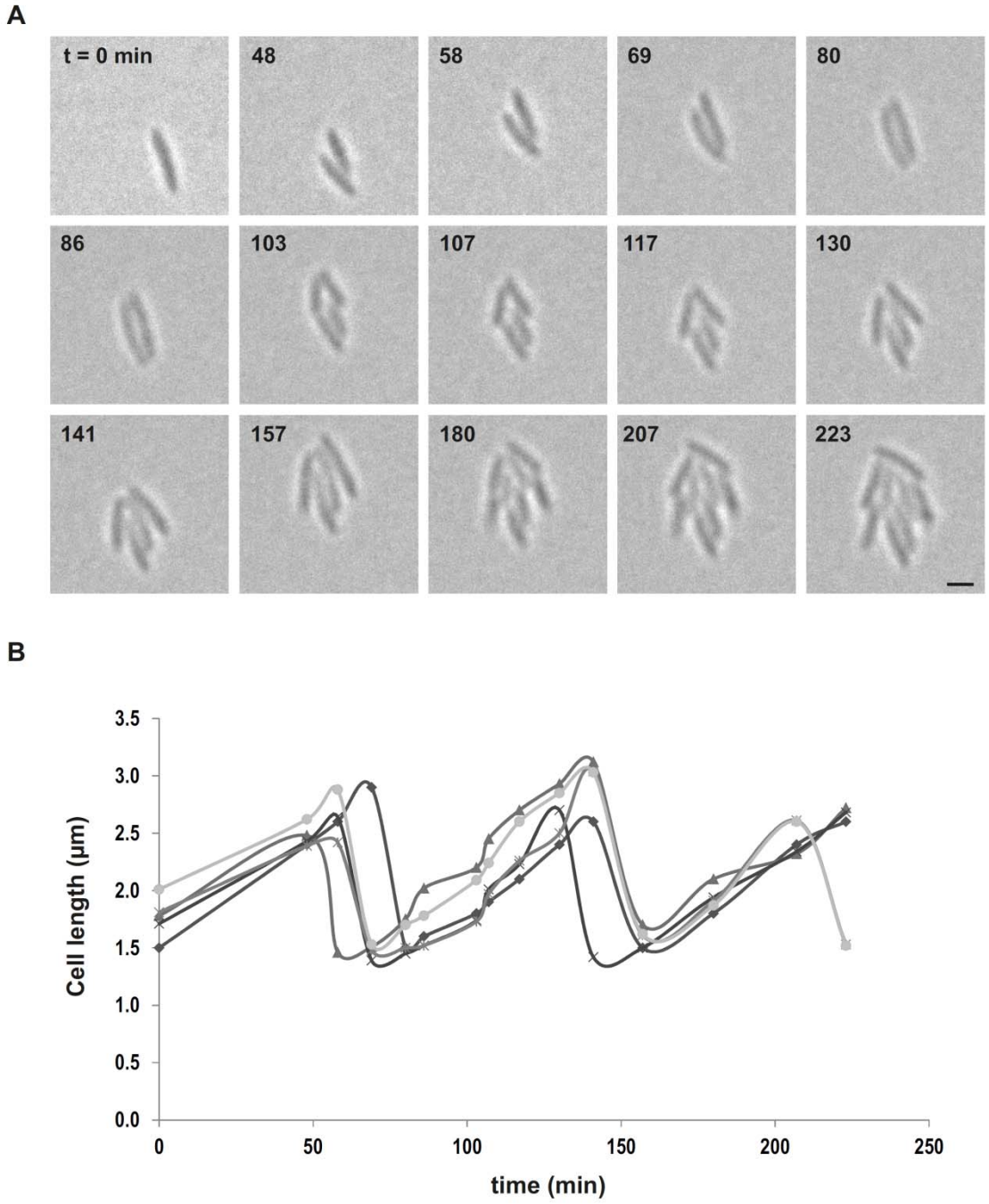
Sequence alignment of *flaA* leaders from *Campylobacter* species. Sequence alignment of redundant sequences for the *flaA* 5' UTR and the first 10 nt of the coding region from different *Campylobacter* species. The consensus structure calculated based on the collapsed version of the alignment (see Fig. 27) is shown at the bottom of each page. The blue and green arrows at the top mark stem loops in the consensus structure while pointing to the unpaired loop region, respectively. The nucleotides of each sequence below the stems are marked in the respective color if the base pair can be formed and have a white background otherwise. The RBS and start codon of *flaA* mRNA are indicated by blue and red bars below the alignment, respectively.

Appendix Figure 12



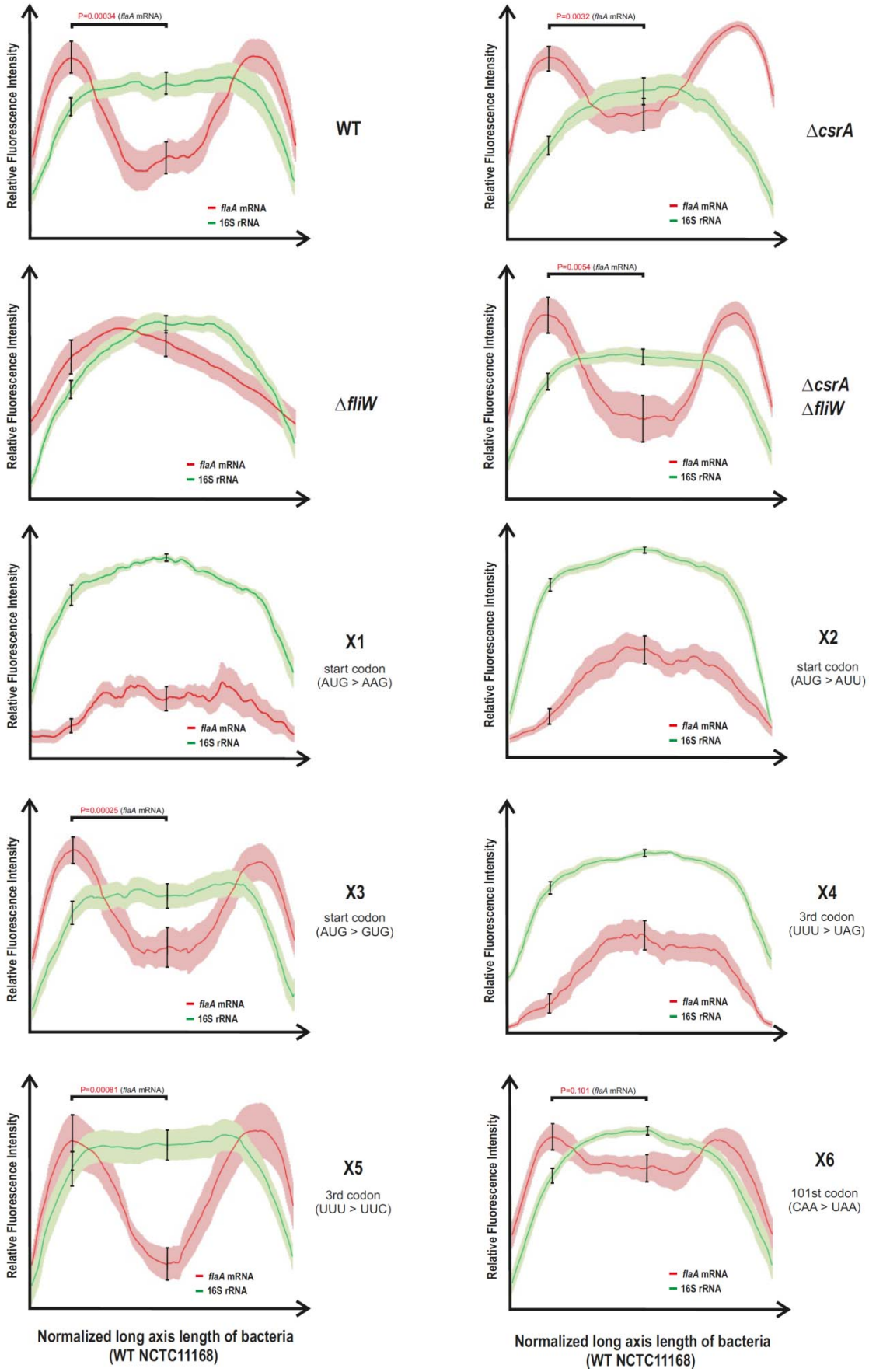
***In vitro* gel-shift assays of 5'-labeled T7-transcripts and purified *C. jejuni* CsrA. (A)** Gel-shift assays of T7-transcribed, 5'-labeled RNAs of flagellar targets with increasing concentrations of CsrA. **(B)** Gel-shift assays with negative controls using the leader of Cj1324 from *C. jejuni* (one GGA, not enriched in either coIP) and a fragment of the unrelated *hopB* mRNA (no GGAs, from *Helicobacter pylori* G27).

Appendix Figure 13



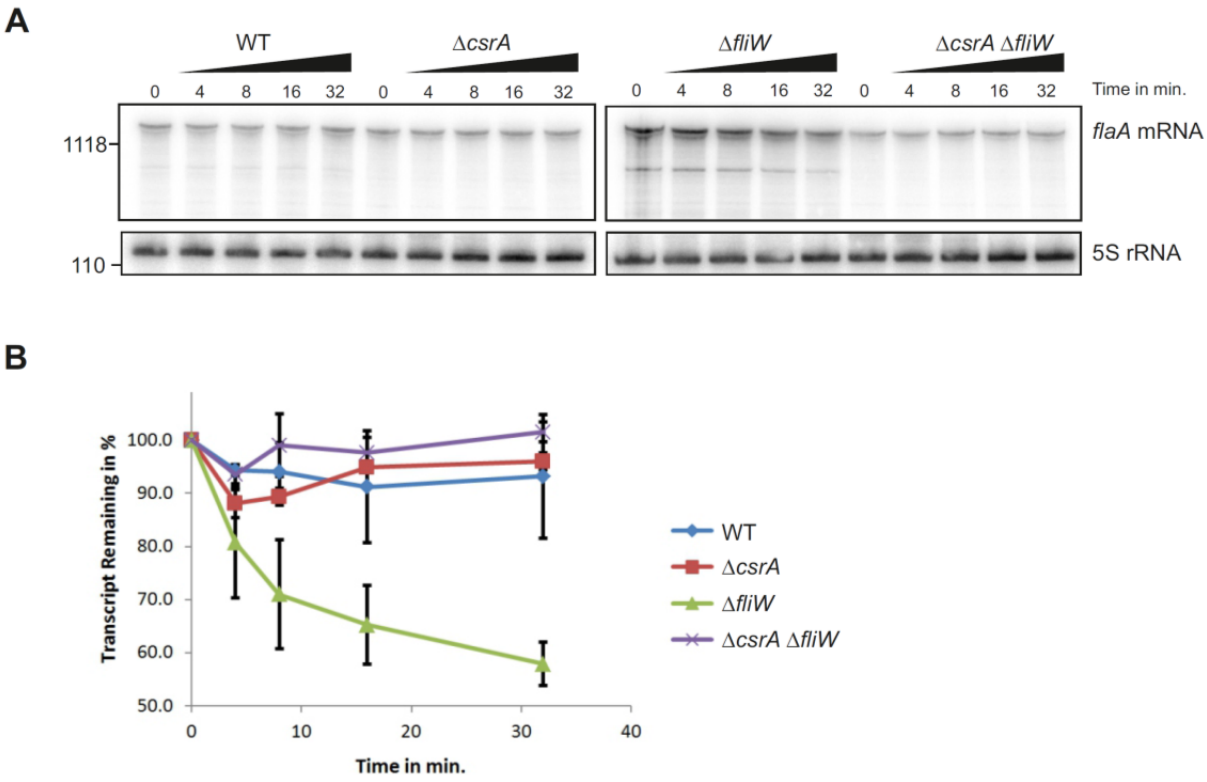
Live-cell imaging of growth of non-motile *C. jejuni* over 2-3 generations. **(a)** A dividing non-motile (Δ*fliA*) *C. jejuni* cell (collected from Brucella broth culture in log phase) was imaged under a fluorescence microscope in bright field mode. Fifteen images were taken over a period of 223 min and 3 cell divisions. The black bar (lower right) represents 1 µm in length. **(b)** Cell lengths were measured over 2-3 divisions for five representative cells using ImageJ and were plotted along the time frame (each curve represents one cell). Please note that the cell lengths were longer (up to 3 µm) compared to those of the fixed WT cells used in the FISH analysis (Fig. 39, up to around 2 µm). This length difference could be due to slightly different morphology of the non-motile strain or different growth conditions (aerobic) used during microscopy.

Appendix Figure 14



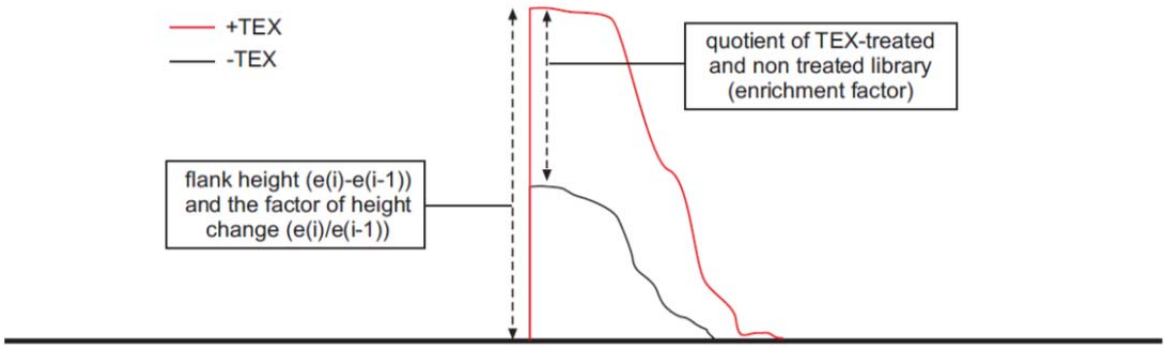
Averaged fluorescence intensity curves for *flaA* mRNA and 16S rRNA FISH signals. Averaged fluorescence intensities (based on 10 cells) from RNA-FISH analysis of 16S rRNA (green) and *flaA* mRNA (red) plotted along the long cell axis of indicated *C. jejuni* strains. The shaded regions along the curves mark the boundary of errors bars (±s.e.m.) of 320 points along the long axis. The points taken for statistical analysis are marked by error bars (Student's *t*-test). These points correspond near the pole and mid-cell of the bacterium.

Appendix Figure 15



***flaA* mRNA rifampicin stability assay. (A)** Northern blot probed for *flaA* mRNA in *C. jejuni* wildtype (WT) and the mutant strains over a time course after rifampicin addition (0-32 min). **(B)** Averaged quantification of *flaA* mRNA transcript levels over time from two independent rifampicin stability assays.

Appendix Figure 16



Schematic representation of basic TSS detection criteria based on dRNA-seq data. The dRNA-seq expression graphs from the exonuclease (TEX) treated (red) and untreated (black) cDNA libraries are the basis for the TSS detection procedure. The most important parameters that are considered during the process are the flank height at position i of the potential TSS (e(i)-e(i-1)) and the factor of height change at that position (e(i)/e(i-1)), where e(i) is the expression value at position i. Additionally, the enrichment factor at the same position is taken into account (i.e. the quotient of the expression values from the treated and untreated library).

7. References

1. Camacho, C., et al., *BLAST+: architecture and applications.* BMC Bioinformatics, 2009. **10**: p. 421.
2. Wassarman, K.M., *Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes.* Cell, 2002. **109**(2): p. 141-4.
3. Waters, L.S. and G. Storz, *Regulatory RNAs in bacteria.* Cell, 2009. **136**(4): p. 615-28.
4. Storz, G., J. Vogel, and K.M. Wassarman, *Regulation by small RNAs in bacteria: expanding frontiers.* Mol Cell, 2011. **43**(6): p. 880-91.
5. Sharma, C.M., et al., *Pervasive post-transcriptional control of genes involved in amino acid metabolism by the Hfq-dependent GcvB small RNA.* Mol Microbiol, 2011. **81**(5): p. 1144-65.
6. Sharma, C.M., et al., *A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites.* Genes Dev, 2007. **21**(21): p. 2804-17.
7. Papenfort, K. and C.K. Vanderpool, *Target activation by regulatory RNAs in bacteria.* FEMS Microbiol Rev, 2015. **39**(3): p. 362-78.
8. Cavanagh, A.T. and K.M. Wassarman, *6S RNA, a global regulator of transcription in Escherichia coli, Bacillus subtilis, and beyond.* Annu Rev Microbiol, 2014. **68**: p. 45-60.
9. Miyakoshi, M., Y. Chao, and J. Vogel, *Cross talk between ABC transporter mRNAs via a target mRNA-derived sponge of the GcvB small RNA.* EMBO J, 2015.
10. Van Assche, E., et al., *RNA-binding proteins involved in post-transcriptional regulation in bacteria.* Front Microbiol, 2015. **6**: p. 141.
11. Vogel, J. and B.F. Luisi, *Hfq and its constellation of RNA.* Nat Rev Microbiol, 2011. **9**(8): p. 578-89.
12. Sittka, A., et al., *Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq.* PLoS Genet, 2008. **4**(8): p. e1000163.
13. Papenfort, K. and J. Vogel, *Regulatory RNA in bacterial pathogens.* Cell Host Microbe, 2010. **8**(1): p. 116-27.
14. Heroven, A.K., K. Bohme, and P. Dersch, *The Csr/Rsm system of Yersinia and related pathogens: a post-transcriptional strategy for managing virulence.* RNA Biol, 2012. **9**(4): p. 379-91.
15. Babitzke, P. and T. Romeo, *CsrB sRNA family: sequestration of RNA-binding regulatory proteins.* Curr Opin Microbiol, 2007. **10**(2): p. 156-63.
16. Seyll, E. and L. Van Melderen, *The ribonucleoprotein Csr network.* Int J Mol Sci, 2013. **14**(11): p. 22117-31.
17. Wei, B.L., et al., *Positive regulation of motility and flhDC expression by the RNA-binding protein CsrA of Escherichia coli.* Mol Microbiol, 2001. **40**(1): p. 245-56.
18. Yakhnin, A.V., et al., *CsrA activates flhDC expression by protecting flhDC mRNA from RNase E-mediated cleavage.* Mol Microbiol, 2013. **87**(4): p. 851-66.
19. Romeo, T., C.A. Vakulskas, and P. Babitzke, *Post-transcriptional regulation on a global scale: form and function of Csr/Rsm systems.* Environ Microbiol, 2013. **15**(2): p. 313-24.
20. Duss, O., et al., *Structural basis of the non-coding RNA RsmZ acting as a protein sponge.* Nature, 2014.
21. Sonnleitner, E., et al., *Hfq-dependent alterations of the transcriptome profile and effects on quorum sensing in Pseudomonas aeruginosa.* Mol Microbiol, 2006. **59**(5): p. 1542-58.

22. Suzuki, K., et al., *Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E.* Genes Dev, 2006. **20**(18): p. 2605-17.
23. Mukherjee, S., et al., *CsrA-FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in Bacillus subtilis.* Mol Microbiol, 2011. **82**(2): p. 447-61.
24. Figueroa-Bossi, N., et al., *RNA remodeling by bacterial global regulator CsrA promotes Rho-dependent transcription termination.* Genes Dev, 2014. **28**(11): p. 1239-51.
25. Marden, J.N., et al., *An unusual CsrA family member operates in series with RsmA to amplify posttranscriptional responses in Pseudomonas aeruginosa.* Proc Natl Acad Sci U S A, 2013. **110**(37): p. 15055-60.
26. Abbott, Z.D., et al., *csrT Represents a New Class of csrA-Like Regulatory Genes Associated with Integrative Conjugative Elements of Legionella pneumophila.* J Bacteriol, 2016. **198**(3): p. 553-64.
27. Abbott, Z.D., et al., *csrR, a Paralog and Direct Target of CsrA, Promotes Legionella pneumophila Resilience in Water.* MBio, 2015. **6**(3): p. e00595.
28. Smirnov, A., et al., *Grad-seq guides the discovery of ProQ as a major small RNA-binding protein.* Proc Natl Acad Sci U S A, 2016. **113**(41): p. 11591-11596.
29. Attaiech, L., et al., *Silencing of natural transformation by an RNA chaperone and a multitarget small RNA.* Proc Natl Acad Sci U S A, 2016. **113**(31): p. 8813-8.
30. Mitobe, J., et al., *RodZ regulates the post-transcriptional processing of the Shigella sonnei type III secretion system.* EMBO Rep, 2011. **12**(9): p. 911-6.
31. Taniguchi, Y., et al., *Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells.* Science, 2010. **329**(5991): p. 533-8.
32. Broude, N.E., *Analysis of RNA localization and metabolism in single live bacterial cells: achievements and challenges.* Mol Microbiol, 2011. **80**(5): p. 1137-47.
33. Brown, J.M. and V.J. Buckle, *Detection of nascent RNA transcripts by fluorescence in situ hybridization.* Methods Mol Biol, 2010. **659**: p. 33-50.
34. Lecuyer, E., et al., *Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function.* Cell, 2007. **131**(1): p. 174-87.
35. Liao, G., et al., *Control of cell migration through mRNA localization and local translation.* Wiley Interdiscip Rev RNA, 2015. **6**(1): p. 1-15.
36. Buskila, A.A., S. Kannaiah, and O. Amster-Choder, *RNA localization in bacteria.* RNA Biol, 2014. **11**(8): p. 1051-60.
37. Moran, U., R. Phillips, and R. Milo, *SnapShot: key numbers in biology.* Cell, 2010. **141**(7): p. 1262-1262 e1.
38. Driessen, A.J. and N. Nouwen, *Protein translocation across the bacterial cytoplasmic membrane.* Annu Rev Biochem, 2008. **77**: p. 643-67.
39. Montero Llopis, P., et al., *Spatial organization of the flow of genetic information in bacteria.* Nature, 2010. **466**(7302): p. 77-81.
40. Nevo-Dinur, K., et al., *Translation-independent localization of mRNA in E. coli.* Science, 2011. **331**(6020): p. 1081-4.
41. Moffitt, J.R., et al., *Spatial organization shapes the turnover of a bacterial transcriptome.* Elife, 2016. **5**.
42. Young, K.T., L.M. Davis, and V.J. Dirita, *Campylobacter jejuni: molecular biology and pathogenesis.* Nat Rev Microbiol, 2007. **5**(9): p. 665-79.
43. Parkhill, J., et al., *The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.* Nature, 2000. **403**(6770): p. 665-8.
44. Holst Sorensen, M.C., et al., *Phase variable expression of capsular polysaccharide modifications allows Campylobacter jejuni to avoid bacteriophage infection in chickens.* Front Cell Infect Microbiol, 2012. **2**: p. 11.
45. Kim, J.S., et al., *Passage of Campylobacter jejuni through the chicken reservoir or mice promotes phase variation in contingency genes Cj0045 and Cj0170 that strongly associates with colonization and disease in a mouse model.* Microbiology, 2012. **158**(Pt 5): p. 1304-16.

46. Anjum, A., et al., *Phase variation of a Type IIG restriction-modification enzyme alters site-specific methylation patterns and gene expression in Campylobacter jejuni strain NCTC11168.* Nucleic Acids Res, 2016. **44**(10): p. 4581-94.
47. Stahl, M., J. Butcher, and A. Stintzi, *Nutrient acquisition and metabolism by Campylobacter jejuni.* Front Cell Infect Microbiol, 2012. **2**: p. 5.
48. Konkel, M.E., et al., *Secretion of virulence proteins from Campylobacter jejuni is dependent on a functional flagellar export apparatus.* J Bacteriol, 2004. **186**(11): p. 3296-303.
49. Fields, J.A. and S.A. Thompson, *Campylobacter jejuni CsrA mediates oxidative stress responses, biofilm formation, and host cell invasion.* J Bacteriol, 2008. **190**(9): p. 3411-6.
50. Porcelli, I., et al., *Parallel evolution of genome structure and transcriptional landscape in the Epsilonproteobacteria.* BMC Genomics, 2013. **14**: p. 616.
51. Butcher, J. and A. Stintzi, *The transcriptional landscape of Campylobacter jejuni under iron replete and iron limited growth conditions.* PLoS One, 2013. **8**(11): p. e79475.
52. Sharma, C.M., et al., *The primary transcriptome of the major human pathogen Helicobacter pylori.* Nature, 2010. **464**(7286): p. 250-5.
53. Taveirne, M.E., et al., *The complete Campylobacter jejuni transcriptome during colonization of a natural host determined by RNAseq.* PLoS One, 2013. **8**(8): p. e73586.
54. Fouts, D.E., et al., *Major structural differences and novel potential virulence mechanisms from the genomes of multiple campylobacter species.* PLoS Biol, 2005. **3**(1): p. e15.
55. Biggs, P.J., et al., *Whole-Genome Comparison of Two Campylobacter jejuni Isolates of the Same Sequence Type Reveals Multiple Loci of Different Ancestral Lineage.* Plos One, 2011. **6**(11).
56. Bischler, T., et al., *Differential RNA-seq (dRNA-seq) for annotation of transcriptional start sites and small RNAs in Helicobacter pylori.* Methods, 2015. **86**: p. 89-101.
57. Hofreuter, D., et al., *Unique features of a highly pathogenic Campylobacter jejuni strain.* Infect Immun, 2006. **74**(8): p. 4694-707.
58. Bacon, D.J., et al., *DNA sequence and mutational analyses of the pVir plasmid of Campylobacter jejuni 81-176.* Infect Immun, 2002. **70**(11): p. 6242-50.
59. Batchelor, R.A., et al., *Nucleotide sequences and comparison of two large conjugative plasmids from different Campylobacter species.* Microbiology, 2004. **150**(Pt 10): p. 3507-17.
60. Pearson, B.M., et al., *The complete genome sequence of Campylobacter jejuni strain 81116 (NCTC11828).* J Bacteriol, 2007. **189**(22): p. 8402-3.
61. Manning, G., et al., *Evidence for a genetically stable strain of Campylobacter jejuni.* Appl Environ Microbiol, 2001. **67**(3): p. 1185-9.
62. Gaynor, E.C., et al., *The genome-sequenced variant of Campylobacter jejuni NCTC 11168 and the original clonal clinical isolate differ markedly in colonization, gene expression, and virulence-associated phenotypes.* J Bacteriol, 2004. **186**(2): p. 503-17.
63. Bailey, T.L., et al., *MEME SUITE: tools for motif discovery and searching.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W202-8.
64. Petersen, L., et al., *RpoD promoters in Campylobacter jejuni exhibit a strong periodic signal instead of a -35 box.* J Mol Biol, 2003. **326**(5): p. 1361-72.
65. Hofreuter, D., V. Novik, and J.E. Galan, *Metabolic diversity in Campylobacter jejuni enhances specific tissue colonization.* Cell Host Microbe, 2008. **4**(5): p. 425-33.
66. Myers, J.D. and D.J. Kelly, *A sulphite respiration system in the chemoheterotrophic human pathogen Campylobacter jejuni.* Microbiology, 2005. **151**(Pt 1): p. 233-42.
67. Doudna, J.A. and E. Charpentier, *Genome editing. The new frontier of genome engineering with CRISPR-Cas9.* Science, 2014. **346**(6213): p. 1258096.
68. Ran, F.A., et al., *In vivo genome editing using Staphylococcus aureus Cas9.* Nature, 2015. **520**(7546): p. 186-91.

69. Brouns, S.J., et al., *Small CRISPR RNAs guide antiviral defense in prokaryotes.* Science, 2008. **321**(5891): p. 960-4.
70. Hale, C.R., et al., *RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex.* Cell, 2009. **139**(5): p. 945-56.
71. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA.* Science, 2008. **322**(5909): p. 1843-5.
72. Deltcheva, E., et al., *CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III.* Nature, 2011. **471**(7340): p. 602-7.
73. Makarova, K.S., et al., *Evolution and classification of the CRISPR-Cas systems.* Nat Rev Microbiol, 2011. **9**(6): p. 467-77.
74. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.
75. Mitschke, J., et al., *An experimentally anchored map of transcriptional start sites in the model cyanobacterium Synechocystis sp. PCC6803.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 2124-9.
76. Albrecht, M., et al., *Deep sequencing-based discovery of the Chlamydia trachomatis transcriptome.* Nucleic Acids Res, 2010. **38**(3): p. 868-77.
77. Schmidtke, C., et al., *Genome-wide transcriptome analysis of the plant pathogen Xanthomonas identifies sRNAs with putative virulence functions.* Nucleic Acids Res, 2012. **40**(5): p. 2020-31.
78. Perutz, M.F., *Species adaptation in a protein molecule.* Mol Biol Evol, 1983. **1**(1): p. 1-28.
79. Mandel, M.J., et al., *A single regulatory gene is sufficient to alter bacterial host range.* Nature, 2009. **458**(7235): p. 215-8.
80. Somvanshi, V.S., et al., *A single promoter inversion switches Photorhabdus between pathogenic and mutualistic states.* Science, 2012. **337**(6090): p. 88-93.
81. Liang, X., et al., *Identification of single nucleotide polymorphisms associated with hyperproduction of alpha-toxin in Staphylococcus aureus.* PLoS One, 2011. **6**(4): p. e18428.
82. Okuda, J. and M. Nishibuchi, *Manifestation of the Kanagawa phenomenon, the virulence-associated phenotype, of Vibrio parahaemolyticus depends on a particular single base change in the promoter of the thermostable direct haemolysin gene.* Mol Microbiol, 1998. **30**(3): p. 499-511.
83. Carpenter, B.M., et al., *A single nucleotide change affects fur-dependent regulation of sodB in H. pylori.* PLoS One, 2009. **4**(4): p. e5369.
84. Beisel, C.L. and G. Storz, *Base pairing small RNAs and their roles in global regulatory networks.* FEMS Microbiol Rev, 2010. **34**(5): p. 866-82.
85. Chaudhuri, R.R., et al., *Quantitative RNA-seq analysis of the transcriptome of Campylobacter jejuni.* Microbiology, 2011. **157**(10): p. 2922-2932.
86. Chao, Y. and J. Vogel, *A 3' UTR-Derived Small RNA Provides the Regulatory Noncoding Arm of the Inner Membrane Stress Response.* Mol Cell, 2016. **61**(3): p. 352-63.
87. Chao, Y., et al., *An atlas of Hfq-bound transcripts reveals 3' UTRs as a genomic reservoir of regulatory small RNAs.* EMBO J, 2012. **31**(20): p. 4005-19.
88. Miyakoshi, M., Y. Chao, and J. Vogel, *Regulatory small RNAs from the 3' regions of bacterial mRNAs.* Curr Opin Microbiol, 2015. **24**: p. 132-9.
89. Shen, Z., et al., *Identification and functional analysis of two toxin-antitoxin systems in Campylobacter jejuni.* Mol Microbiol, 2016.
90. Le, M.T., et al., *Conservation of sigma28-Dependent Non-Coding RNA Paralogs and Predicted sigma54-Dependent Targets in Thermophilic Campylobacter Species.* PLoS One, 2015. **10**(10): p. e0141627.
91. Schouls, L.M., et al., *Comparative genotyping of Campylobacter jejuni by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: strain diversity, host range, and recombination.* J Clin Microbiol, 2003. **41**(1): p. 15-26.

92. Zhang, Y., et al., *Processing-independent CRISPR RNAs limit natural transformation in Neisseria meningitidis.* Mol Cell, 2013. **50**(4): p. 488-503.
93. Louwen, R., et al., *A novel link between Campylobacter jejuni bacteriophage defence, virulence and Guillain-Barre syndrome.* Eur J Clin Microbiol Infect Dis, 2012.
94. Barnard, F.M., et al., *Global regulation of virulence and the stress response by CsrA in the highly adapted human gastric pathogen Helicobacter pylori.* Mol Microbiol, 2004. **51**(1): p. 15-32.
95. Kao, C.Y., B.S. Sheu, and J.J. Wu, *CsrA regulates Helicobacter pylori J99 motility and adhesion by controlling flagella formation.* Helicobacter, 2014. **19**(6): p. 443-54.
96. Yakhnin, H., et al., *CsrA of Bacillus subtilis regulates translation initiation of the gene encoding the flagellin protein (hag) by blocking ribosome binding.* Mol Microbiol, 2007. **64**(6): p. 1605-20.
97. Dugar, G., et al., *High-Resolution Transcriptome Maps Reveal Strain-Specific Regulatory Features of Multiple <italic>Campylobacter jejuni</italic> Isolates.* PLoS Genet, 2013. **9**(5): p. e1003495.
98. Mercante, J., et al., *Comprehensive alanine-scanning mutagenesis of Escherichia coli CsrA defines two subdomains of critical functional importance.* J Biol Chem, 2006. **281**(42): p. 31832-42.
99. Romeo, T., et al., *Identification and molecular characterization of csrA, a pleiotropic gene from Escherichia coli that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties.* J Bacteriol, 1993. **175**(15): p. 4744-55.
100. Sze, C.W., et al., *Carbon storage regulator A (CsrA(Bb)) is a repressor of Borrelia burgdorferi flagellin protein FlaB.* Mol Microbiol, 2011. **82**(4): p. 851-64.
101. Rieder, R., et al., *Experimental tools to identify RNA-protein interactions in Helicobacter pylori.* RNA Biology, 2012. **9**(4).
102. Gundogdu, O., et al., *Re-annotation and re-analysis of the Campylobacter jejuni NCTC11168 genome sequence.* BMC Genomics, 2007. **8**: p. 162.
103. Berg, H.C., *The rotary motor of bacterial flagella.* Annu Rev Biochem, 2003. **72**: p. 19-54.
104. Guerry, P., et al., *Role of two flagellin genes in Campylobacter motility.* J Bacteriol, 1991. **173**(15): p. 4757-64.
105. Hendrixson, D.R., *Regulation of Flagellar Gene Expression and Assembly*, in *Campylobacter*. 2008, ASM Press.
106. Urban, J.H. and J. Vogel, *Translational control and target recognition by Escherichia coli small RNAs in vivo.* Nucleic Acids Res, 2007. **35**(3): p. 1018-37.
107. Wang, X., et al., *CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a biofilm polysaccharide adhesin of Escherichia coli.* Mol Microbiol, 2005. **56**(6): p. 1648-63.
108. Yao, Z., Z. Weinberg, and W.L. Ruzzo, *CMfinder--a covariance model based RNA motif finding algorithm.* Bioinformatics, 2006. **22**(4): p. 445-52.
109. Dubey, A.K., et al., *RNA sequence and secondary structure participate in high-affinity CsrA-RNA interaction.* RNA, 2005. **11**(10): p. 1579-87.
110. Zuker, M., *Mfold web server for nucleic acid folding and hybridization prediction.* Nucleic Acids Res, 2003. **31**(13): p. 3406-15.
111. Dugar, G., et al., *The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in Campylobacter jejuni.* Nat Commun, 2016. **7**: p. 11667.
112. Golden, N.J. and D.W. Acheson, *Identification of motility and autoagglutination Campylobacter jejuni mutants by random transposon mutagenesis.* Infect Immun, 2002. **70**(4): p. 1761-71.
113. Battesti, A. and E. Bouveret, *The bacterial two-hybrid system based on adenylate cyclase reconstitution in Escherichia coli.* Methods, 2012. **58**(4): p. 325-34.
114. Zhang, B., et al., *A dual-intein autoprocessing domain that directs synchronized protein co-expression in both prokaryotes and eukaryotes.* Sci Rep, 2015. **5**: p. 8541.
115. Hendrixson, D.R. and V.J. DiRita, *Transcription of sigma54-dependent but not sigma28-dependent flagellar genes in Campylobacter jejuni is associated with formation of the flagellar secretory apparatus.* Mol Microbiol, 2003. **50**(2): p. 687-702.

116. Barrero-Tobon, A.M. and D.R. Hendrixson, *Identification and analysis of flagellar coexpressed determinants (Feds) of Campylobacter jejuni involved in colonization.* Mol Microbiol, 2012. **84**(2): p. 352-69.
117. Heilemann, M., et al., *Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes.* Angew Chem Int Ed Engl, 2008. **47**(33): p. 6172-6.
118. Fei, J., et al., *RNA biochemistry. Determination of in vivo target search kinetics of regulatory noncoding RNA.* Science, 2015. **347**(6228): p. 1371-4.
119. Minamino, T., *Protein export through the bacterial flagellar type III export pathway.* Biochim Biophys Acta, 2014. **1843**(8): p. 1642-8.
120. Evans, L.D., et al., *A chain mechanism for flagellum growth.* Nature, 2013. **504**(7479): p. 287-90.
121. Ferris, H.U., et al., *FlhB regulates ordered export of flagellar components via autocleavage mechanism.* J Biol Chem, 2005. **280**(50): p. 41236-42.
122. Morris, D.P., et al., *Kinetic characterization of Salmonella FliK-FlhB interactions demonstrates complexity of the Type III secretion substrate-specificity switch.* Biochemistry, 2010. **49**(30): p. 6386-93.
123. Pallen, M.J., C.W. Penn, and R.R. Chaudhuri, *Bacterial flagellar diversity in the post-genomic era.* Trends Microbiol, 2005. **13**(4): p. 143-9.
124. Wand, M.E., et al., *Helicobacter pylori FlhB function: the FlhB C-terminal homologue HP1575 acts as a "spare part" to permit flagellar export when the HP0770 FlhBCC domain is deleted.* J Bacteriol, 2006. **188**(21): p. 7531-41.
125. Smith, T.G., L. Pereira, and T.R. Hoover, *Helicobacter pylori FlhB processing-deficient variants affect flagellar assembly but not flagellar gene expression.* Microbiology, 2009. **155**(Pt 4): p. 1170-80.
126. Brencic, A. and S. Lory, *Determination of the regulon and identification of novel mRNA targets of Pseudomonas aeruginosa RsmA.* Mol Microbiol, 2009. **72**(3): p. 612-32.
127. Lawhon, S.D., et al., *Global regulation by CsrA in Salmonella typhimurium.* Mol Microbiol, 2003. **48**(6): p. 1633-45.
128. Burrowes, E., et al., *Influence of the regulatory protein RsmA on cellular functions in Pseudomonas aeruginosa PAO1, as revealed by transcriptome analysis.* Microbiology, 2006. **152**(Pt 2): p. 405-18.
129. Edwards, A.N., et al., *Circuitry linking the Csr and stringent response global regulatory systems.* Mol Microbiol, 2011. **80**(6): p. 1561-80.
130. Jonas, K., et al., *The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins.* Mol Microbiol, 2008. **70**(1): p. 236-57.
131. Konig, J., et al., *Protein-RNA interactions: new genomic technologies and perspectives.* Nat Rev Genet, 2012. **13**(2): p. 77-83.
132. Zhang, C. and R.B. Darnell, *Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data.* Nat Biotechnol, 2011. **29**(7): p. 607-14.
133. Yakhnin, H., et al., *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.* J Bacteriol, 2011. **193**(22): p. 6162-70.
134. Duss, O., et al., *Molecular basis for the wide range of affinity found in Csr/Rsm protein-RNA recognition.* Nucleic Acids Res, 2014. **42**(8): p. 5332-46.
135. Mercante, J., et al., *Molecular geometry of CsrA (RsmA) binding to RNA and its implications for regulated expression.* J Mol Biol, 2009. **392**(2): p. 511-28.
136. Sterzenbach, T., et al., *A novel CsrA titration mechanism regulates fimbrial gene expression in Salmonella typhimurium.* EMBO J, 2013. **32**(21): p. 2872-83.
137. Miyakoshi, M., Y. Chao, and J. Vogel, *Regulatory small RNAs from the 3' regions of bacterial mRNAs.* Curr Opin Microbiol, 2015. **24C**: p. 132-139.
138. Figueroa-Bossi, N., et al., *Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target.* Genes Dev, 2009. **23**(17): p. 2004-15.
139. Jorgensen, M.G., et al., *Dual function of the McaS small RNA in controlling biofilm formation.* Genes Dev, 2013. **27**(10): p. 1132-45.

140. Blumer, C. and D. Haas, *Multicopy suppression of a gacA mutation by the infC operon in Pseudomonas fluorescens CHA0: competition with the global translational regulator RsmA.* FEMS Microbiol Lett, 2000. **187**(1): p. 53-8.
141. Snel, B., et al., *STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene.* Nucleic Acids Res, 2000. **28**(18): p. 3442-4.
142. Kulkarni, P.R., et al., *Prediction of CsrA-regulating small RNAs in bacteria and their experimental verification in Vibrio fischeri.* Nucleic Acids Res, 2006. **34**(11): p. 3361-9.
143. Agaras, B., P. Sobrero, and C. Valverde, *A CsrA/RsmA translational regulator gene encoded in the replication region of a Sinorhizobium meliloti cryptic plasmid complements Pseudomonas fluorescens rsmA/E mutants.* Microbiology, 2013. **159**(Pt 2): p. 230-42.
144. Yakhnin, H., et al., *Complex regulation of the global regulatory gene csrA: CsrA-mediated translational repression, transcription from five promoters by Esigma(7)(0) and Esigma(S), and indirect transcriptional activation by CsrA.* Mol Microbiol, 2011. **81**(3): p. 689-704.
145. Parrish, J.R., et al., *A proteome-wide protein interaction map for Campylobacter jejuni.* Genome Biol, 2007. **8**(7): p. R130.
146. Mukherjee, S., P. Babitzke, and D.B. Kearns, *FliW and FliS function independently to control cytoplasmic flagellin levels in Bacillus subtilis.* J Bacteriol, 2013. **195**(2): p. 297-306.
147. Calvo, R.A. and D.B. Kearns, *FlgM is secreted by the flagellar export apparatus in Bacillus subtilis.* J Bacteriol, 2015. **197**(1): p. 81-91.
148. Keiler, K.C., *RNA localization in bacteria.* Curr Opin Microbiol, 2011. **14**(2): p. 155-9.
149. Barrero Tobon, A.M. and D.R. Hendrixson, *Flagellar biosynthesis exerts temporal regulation of secretion of specific Campylobacter jejuni colonization and virulence determinants.* Mol Microbiol, 2014. **93**(5): p. 957-74.
150. Singer, H.M., M. Erhardt, and K.T. Hughes, *Comparative analysis of the secretion capability of early and late flagellar type III secretion substrates.* Mol Microbiol, 2014. **93**(3): p. 505-20.
151. Karlinsey, J.E., et al., *Translation/secretion coupling by type III secretion systems.* Cell, 2000. **102**(4): p. 487-97.
152. Buchan, J.R., *mRNP granules: Assembly, function, and connections with disease.* RNA Biol, 2014. **11**(8).
153. Ouyang, Z., J. Zhou, and M.V. Norgard, *CsrA (BB0184) is not involved in activation of the RpoN-RpoS regulatory pathway in Borrelia burgdorferi.* Infect Immun, 2014. **82**(4): p. 1511-22.
154. Anderson, P.E. and J.W. Gober, *FlbT, the post-transcriptional regulator of flagellin synthesis in Caulobacter crescentus, interacts with the 5' untranslated region of flagellin mRNA.* Mol Microbiol, 2000. **38**(1): p. 41-52.
155. Wassenaar, T.M., N.M. Bleumink-Pluym, and B.A. van der Zeijst, *Inactivation of Campylobacter jejuni flagellin genes by homologous recombination demonstrates that flaA but not flaB is required for invasion.* EMBO J, 1991. **10**(8): p. 2055-61.
156. Guerry, P., *Campylobacter flagella: not just for motility.* Trends Microbiol, 2007. **15**(10): p. 456-61.
157. Balaban, M. and D.R. Hendrixson, *Polar flagellar biosynthesis and a regulator of flagellar number influence spatial parameters of cell division in Campylobacter jejuni.* PLoS Pathog, 2011. **7**(12): p. e1002420.
158. Konkel, M.E., et al., *Identification of proteins required for the internalization of Campylobacter jejuni into cultured mammalian cells.* Adv Exp Med Biol, 1999. **473**: p. 215-24.
159. Buelow, D.R., et al., *Campylobacter jejuni survival within human epithelial cells is enhanced by the secreted protein CiaI.* Mol Microbiol, 2011. **80**(5): p. 1296-312.
160. Christensen, J.E., S.A. Pacheco, and M.E. Konkel, *Identification of a Campylobacter jejuni-secreted protein required for maximal invasion of host cells.* Mol Microbiol, 2009. **73**(4): p. 650-62.

161. Pernitzsch, S.R., et al., *A variable homopolymeric G-repeat defines small RNA-mediated posttranscriptional regulation of a chemotaxis receptor in Helicobacter pylori.* Proc Natl Acad Sci U S A, 2014. **111**(4): p. E501-10.
162. Gripp, E., et al., *Closely related Campylobacter jejuni strains from different sources reveal a generalist rather than a specialist lifestyle.* BMC Genomics, 2011. **12**: p. 584.
163. Jerome, J.P., et al., *Standing genetic variation in contingency loci drives the rapid adaptation of Campylobacter jejuni to a novel host.* PLoS One, 2011. **6**(1): p. e16399.
164. DiChiara, J.M., et al., *Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis.* Nucleic Acids Res, 2010. **38**(12): p. 4067-78.
165. Pichon, C. and B. Felden, *Small RNA genes expressed from Staphylococcus aureus genomic and pathogenicity islands with specific expression among pathogenic strains.* Proc Natl Acad Sci U S A, 2005. **102**(40): p. 14249-54.
166. Kim, D., et al., *Comparative Analysis of Regulatory Elements between Escherichia coli and Klebsiella pneumoniae by Genome-Wide Transcription Start Site Profiling.* PLoS Genet, 2012. **8**(8): p. e1002867.
167. Ingolia, N.T., et al., *Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.* Science, 2009. **324**(5924): p. 218-23.
168. Halstead, J.M., et al., *Translation. An RNA biosensor for imaging the first round of translation from single cells to living animals.* Science, 2015. **347**(6228): p. 1367-671.
169. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Res, 2004. **32**(5): p. 1792-7.
170. Talavera, G. and J. Castresana, *Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments.* Syst Biol, 2007. **56**(4): p. 564-77.
171. Guindon, S., et al., *New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0.* Syst Biol, 2010. **59**(3): p. 307-21.
172. Le, S.Q. and O. Gascuel, *An improved general amino acid replacement matrix.* Mol Biol Evol, 2008. **25**(7): p. 1307-20.
173. Creevey, C.J. and J.O. McInerney, *Clann: investigating phylogenetic information through supertree analyses.* Bioinformatics, 2005. **21**(3): p. 390-2.
174. Letunic, I. and P. Bork, *Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy.* Nucleic Acids Res, 2011. **39**(Web Server issue): p. W475-8.
175. Hansen, C.R., et al., *Rapid construction of Campylobacter jejuni deletion mutants.* Lett Appl Microbiol, 2007. **45**(6): p. 599-603.
176. Wassenaar, T.M., B.N. Fry, and B.A. van der Zeijst, *Genetic manipulation of Campylobacter: evaluation of natural transformation and electro-transformation.* Gene, 1993. **132**(1): p. 131-5.
177. McLennan, M.K., et al., *Campylobacter jejuni biofilms up-regulated in the absence of the stringent response utilize a calcofluor white-reactive polysaccharide.* J Bacteriol, 2008. **190**(3): p. 1097-107.
178. Skouloubris, S., et al., *The Helicobacter pylori UreI protein is not involved in urease activity but is essential for bacterial survival in vivo.* Infect Immun, 1998. **66**(9): p. 4517-21.
179. Boneca, I.G., et al., *Development of inducible systems to engineer conditional mutants of essential genes of Helicobacter pylori.* Appl Environ Microbiol, 2008. **74**(7): p. 2095-102.
180. Bury-Mone, S., et al., *Presence of active aliphatic amidases in Helicobacter species able to colonize the stomach.* Infect Immun, 2003. **71**(10): p. 5613-22.
181. Cameron, A. and E.C. Gaynor, *Hygromycin B and apramycin antibiotic resistance cassettes for use in Campylobacter jejuni.* PLoS One, 2014. **9**(4): p. e95084.
182. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.

183. Berezikov, E., et al., *Diversity of microRNAs in human and chimpanzee brain.* Nat Genet, 2006. **38**(12): p. 1375-7.
184. Hoffmann, S., et al., *Fast mapping of short sequences with mismatches, insertions and deletions using index structures.* PLoS Comput Biol, 2009. **5**(9): p. e1000502.
185. Nicol, J.W., et al., *The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets.* Bioinformatics, 2009. **25**(20): p. 2730-1.
186. Darling, A.C., et al., *Mauve: multiple alignment of conserved genomic sequence with rearrangements.* Genome Res, 2004. **14**(7): p. 1394-403.
187. Herbig, A., et al., *GenomeRing: alignment visualization based on SuperGenome coordinates.* Bioinformatics, 2012. **28**(12): p. i7-15.
188. Papenfort, K., et al., *SigmaE-dependent small RNAs of Salmonella respond to membrane stress by accelerating global omp mRNA decay.* Mol Microbiol, 2006. **62**(6): p. 1674-88.
189. Sittka, A., et al., *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 Biol, 2009. **6**(3): p. 266-75.
190. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.* Nucleic Acids Res, 2002. **30**(1): p. 207-10.
191. Feng, J., et al., *GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data.* Bioinformatics, 2012. **28**(21): p. 2782-8.
192. Weinberg, Z. and R.R. Breaker, *R2R--software to speed the depiction of aesthetic consensus RNA secondary structures.* BMC Bioinformatics, 2011. **12**: p. 3.
193. Whiteside, M.D., et al., *OrthologueDB: a bacterial and archaeal orthology resource for improved comparative genomic analysis.* Nucleic Acids Res, 2013. **41**(Database issue): p. D366-76.
194. Altschul, S.F., et al., *Basic local alignment search tool.* J Mol Biol, 1990. **215**(3): p. 403-10.
195. Bernhart, S.H., et al., *RNAalifold: improved consensus structure prediction for RNA alignments.* BMC Bioinformatics, 2008. **9**: p. 474.
196. Blomberg, P., E.G. Wagner, and K. Nordstrom, *Control of replication of plasmid R1: the duplex between the antisense RNA, CopA, and its target, CopT, is processed specifically in vivo and in vitro by RNase III.* EMBO J, 1990. **9**(7): p. 2331-40.
197. Dugar, G., et al., *High-Resolution Transcriptome Maps Reveal Strain-Specific Regulatory Features of Multiple Campylobacter jejuni Isolates.* PLoS Genet, 2013. **9**(5): p. e1003495.
198. Kappler, U., et al., *Sulfite:Cytochrome c oxidoreductase from Thiobacillus novellus. Purification, characterization, and molecular biology of a heterodimeric member of the sulfite oxidase family.* J Biol Chem, 2000. **275**(18): p. 13202-12.
199. Sittka, A., et al., *The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium.* Mol Microbiol, 2007. **63**(1): p. 193-217.
200. Russell, J.H. and K.C. Keiler, *Subcellular localization of a bacterial regulatory RNA.* Proc Natl Acad Sci U S A, 2009. **106**(38): p. 16405-9.
201. van de Linde, S., et al., *Direct stochastic optical reconstruction microscopy with standard fluorescent probes.* Nat Protoc, 2011. **6**(7): p. 991-1009.
202. Wolter, S., et al., *rapidSTORM: accurate, fast open-source software for localization microscopy.* Nat Methods, 2012. **9**(11): p. 1040-1.
203. Wosten, M.M., J.A. Wagenaar, and J.P. van Putten, *The FlgS/FlgR two-component signal transduction system regulates the fla regulon in Campylobacter jejuni.* J Biol Chem, 2004. **279**(16): p. 16214-22.
204. Gursinsky, T., et al., *Factors and selenocysteine insertion sequence requirements for the synthesis of selenoproteins from a gram-positive anaerobe in Escherichia coli.* Appl Environ Microbiol, 2008. **74**(5): p. 1385-93.
205. Lutz, R. and H. Bujard, *Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements.* Nucleic Acids Res, 1997. **25**(6): p. 1203-10.

206. Gopel, Y., et al., *Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA.* Genes Dev, 2013. **27**(5): p. 552-64.

8. List of Figures

Figure 1. The current model of CsrA based regulation **5**
Figure 2. Localization of RNAs in bacteria **6**
Figure 3. Hypervariable regions in the genome of *C. jejuni* NCTC11168 **8**
Figure 4. Phylogenetic tree of representative members of the Epsilonproteobacterial lineage **11**
Figure 5. Comparative annotation of transcriptional start sites (TSS) based on differential RNA-seq **14**
Figure 6. Differential RNA–seq and SuperGenome-based annotation of transcriptional start sites **15**
Figure 7. TSS classifications for the four *C. jejuni* strains **17**
Figure 8. Transcriptome features of multiple *C. jejuni* strains **19**
Figure 9. 5' UTR length distributions and comparisons **20**
Figure 10. Effect of gene insertion on Transcriptome features **21**
Figure 11. Orthologous genes with differences in 5' UTR length **22**
Figure 12. SNP–dependent promoter usage in *C. jejuni* **23**
Figure 13. SNP–dependent promoter usage in *C. jejuni* affects sulphite respiration **24**
Figure 14. Small RNAs in *Campylobacter jejuni* **26**
Figure 15. Conservation of *Campylobacter jejuni* sRNA candidates in Epsilonproteobacteria **27**
Figure 16. CRISPR locus is expressed in *C. jejuni* **28**
Figure 17. The minimal CRISPR-Cas IIA system of *C. jejuni* **29**
Figure 18. Deletion of RNase III leads to accumulation of longer crRNAs **30**
Figure 19. Genomic location and conservation of CsrA protein **36**
Figure 20. Changes in transcriptome upon *csrA* deletion in *C. jejuni* strain NCTC11168 **37**
Figure 21. Growth curves and RIP-seq of CsrA in *C. jejuni* strain NCTC11168 **38**
Figure 22. Statistical analysis for overrepresentation of functional categories among potential CsrA target genes **40**
Figure 23. Flagellar mRNA targets of CsrA **43**
Figure 24. Validation of potential 5' UTR CsrA targets using a GFP reporter system in *E. coli* **44**
Figure 25. Validation of potential intergenic CsrA targets using a dual reporter system in *E. coli* **45**
Figure 26. CsrA binding motif analysis using sequence extracted from enriched peaks **47**
Figure 27. Structure annotated sequence alignment of the *flaA* 5' UTR **48**
Figure 28. CsrA represses *flaA* translation by binding to its 5' UTR **49**
Figure 29. *In vitro* gel-shift assays of 5'-labeled T7-transcripts and purified *C. jejuni* CsrA **50**
Figure 30. Footprinting assays define CsrA binding sites on *flaA* mRNA leader **51**
Figure 31. Genomic locations of CsrA and FliW in selected bacterial species/strains **53**
Figure 32. BACTH and protein-protein coIP confirms direct interactions of FliW with CsrA and FlaA **54**
Figure 33. The flagellar assembly factor FliW binds and antagonizes CsrA **55**
Figure 34. The flagellar assembly factor FliW affects proper flagella formation **56**
Figure 35. CsrA binds to other flagellar target mRNAs but *csrA* deletion does not affect their translation **57**
Figure 36. The flaA 5' UTR and FliW inhibit CsrA-mediated regulation of flagella genes **59**
Figure 37. An ectopically-expressed *flaA* mini-gene can partially complement CsrA-mediated effects on FlaA and FlaG translation upon *fliW* deletion **60**
Figure 38. *flaA* mRNA localizes to the poles of about one-fifth of *C. jejuni* NCTC11168 wild-type cells **62**
Figure 39. The *flaA* mRNA localizes to the poles of short elongating cells **63**
Figure 40. Influence of CsrA and FliW on *flaA* mRNA localization to the poles **64**
Figure 41. Translation is required for *flaA* mRNA localization to the cell poles **66**
Figure 42. Super-resolution imaging of *flaA* mRNA **67**
Figure 43. FliW affects transcription from *flaA* promoter **68**

Figure 44. FlhB and FlhX in flagellar export apparatus **70**
Figure 45. *flaA* mRNA localization in the mutant strains **71**

Appendix Figure 1. Promoter motifs detected for TSS of the four individual strains **172**
Appendix Figure 2. Promoter Motifs detected by *MEME* for conserved and strain-specific TSS **173**
Appendix Figure 3. dRNA-seq data and promoter alignments for genes with differentially expressed TSS among four *C. jejuni* strains **174**
Appendix Figure 4. Differentially expressed TSS with SNPs in the periodic A/T-rich pattern or in the extended -10 promoter box **178**
Appendix Figure 5. Examples for genes with *cis*-encoded antisense RNAs **179**
Appendix Figure 6. dRNA-seq reveals several sRNA candidates in *C. jejuni* **180**
Appendix Figure 7. Transcription and genomic organization of type-II C CRISPR/*cas* (Nmeni/CASS4) loci in *C. jejuni* **183**
Appendix Figure 8. Correlation analysis of intact CRISPR loci and the presence/absence of plasmids or integrated elements in diverse *C. jejuni* clinical isolates **184**
Appendix Figure 9. RIP-seq of CsrA-3xFLAG in strain 81-176 **185**
Appendix Figure 10. FACS analysis of GFP reporter fusions **186**
Appendix Figure 11. Sequence alignment of *flaA* leaders from *Campylobacter* species **187**
Appendix Figure 12. *In vitro* gel-shift assays of 5'-labeled T7-transcripts and purified *C. jejuni* CsrA **189**
Appendix Figure 13. Live-cell imaging of growth of non-motile *C. jejuni* over 2-3 generations **190**
Appendix Figure 14. Averaged fluorescence intensity curves for *flaA* mRNA and 16S rRNA FISH Signals **191**
Appendix Figure 15. *flaA* mRNA rifampicin stability assay **193**
Appendix Figure 16. Schematic representation of basic TSS detection criteria based on dRNA-seq data **194**

9. List of Tables

Table 1. Characteristics of *Campylobacter jejuni* strains used for dRNA-seq **11**
Table 2. Mapping statistics of *Campylobacter jejuni* dRNA-seq libraries **13**
Table 3. Comparative TSS annotation **18**
Table 4. Distribution of mapped reads to different RNA classes **39**
Table 5. Enrichment of genes involved in flagellar biosynthesis in the CsrA-3xFLAG coIP data **42**

Appendix Table 1. SuperGenome TSS table **111**
Appendix Table 2. TSS with previously described FliA and RpoN dependent promoters **111**
Appendix Table 3. TSS with a predicted FliA or RpoN motif **112**
Appendix Table 4. Comparison of gene lengths for annotation correction **122**
Appendix Table 5. Orthologs with corrected annotation **122**
Appendix Table 6. 5' UTR comparisons **128**
Appendix Table 7. sRNA candidates in the four *C. jejuni* strains **129**
Appendix Table 8: Sequence of sRNA candidates in *C. jejuni* NCTC11168 **137**
Appendix Table 9. Mapping statistics of *C. jejuni* CsrA coIP RNA-seq libraries **140**
Appendix Table 10. Potential CsrA targets from *C. jejuni* NCTC11168 **141**
Appendix Table 11. List of potential CsrA binding sites **145**
Appendix Table 12. DNA oligonucleotides **155**
Appendix Table 13. List of all *C. jejuni* and *E. coli* strains **162**
Appendix Table 14. List of all plasmids used in this thesis **166**
Appendix Table 15. Construction of *C. jejuni* mutants **168**
Appendix Table 16. GFP fusions for validating CsrA-target interactions in *E. coli* **169**
Appendix Table 17. Details of RNAs used for *in vitro* work **170**
Appendix Table 18. *flaA*_mini and Cj1321_mini transcript expressed in *C. jejuni* **171**

10. Contribution by others

The work described in this doctoral thesis was conducted under the supervision of Prof. Dr. Cynthia Sharma at the Institute of Molecular Infection Biology (IMIB), University of Würzburg, Würzburg, Germany.

Several parts of the work described here have been contributed by others and are listed below.

- The SuperGenome construction together with the prediction of TSS using an automated approach was performed by Dr. Alexander Herbig under the supervision of Prof. Dr. Kay Nieselt at the University of Tübingen.
- Processing of RNA-seq data and bioinformatics analyses was performed by Dr. Konrad Förstner for the transcriptome datasets.
- Correlation analysis for CRISPR loci in *Campylobacter* was performed by Dr. Nadja Heidrich
- The Peak detection approach along with the bioinformatic analyses for the CsrA study was performed by Thorsten Bischler with inputs from Dr. Konrad Förstner and Dr. Lars Barquist.
- Wet lab experiments and strain construction for the CsrA study were performed together with Dr. Sarah Svensson with technical assistance from Belinda Aul.
- High-throughput sequencing was performed by Dr. Richard Reinhardt at the Max-Planck Institute for Plant Breeding Research, Cologne, Germany.
- cDNA libraries were constructed by the vertis Biotechnologie AG, Germany.
- Electron and confocal microscopy was performed with help from Hilde Merkert.
- *d*STORM super-resolution microscopy was performed by Sina Wäldchen under the supervision of Prof. Dr. Markus Sauer at the University of Würzburg.

11. Curriculum Vitae

Gaurav Dugar

29 yrs, +49-15774806990, gaurav.dugar@uni-wuerzburg.de

Rimparerstr. 4, 97080 Würzburg, Germany

Education

University of Würzburg – **PhD Candidate (2011 – present)**

Indian Institute of Technology, Delhi – **M.Tech** in Biochemical Engineering and Biotechnology **(2009-2010)**
Thesis - Screening of novel phosphate solubilizing bacteria from rhizosphere.

Indian Institute of Technology, Delhi – **B.Tech** in Biochemical Engineering and Biotechnology **(2005-2009)**

Research/Work Experience

Institute Cochin, Paris (May´09 – July´09)
Characterization of Secretory proteins from *Taxoplasma gondii*

Nanyang Technological University (NTU), Singapore (Dec´08-Jan´09)
Novel biomaterials with layered protein epitopes for better attachment and proliferation of mammalian cells

USV Ltd., Mumbai (May´08 – July´08)
Optimization of a 10 stage protein refolding and purification process of PDGF-BB for therapeutical application

Brain and Mind Research Institute (BMRI), Sydney (May´07 – July´07) & (Dec´07-Jan´08)
Study the role of amyloid-beta and BACE1 in Alzheimer's disease

The learning Foundation India, India (May´06 – July´06)
Part of 20-member international team to teach unprivileged children in Thar desert, India

Scholastic Achievements

- Awarded fellowship by **Graduate School of Life Sciences (GSLS), University of Würzburg** for PhD studies
- Awarded scholarship by **French Embassy in India** for summer research training in Paris, 2009
- Secured **99.7 percentile (1 out of 400 applicants)** in Graduate Aptitude Test in Engineering **(GATE)**, 2009
- Awarded fellowship by the **Government of India** for the duration of Master's degree
- Secured **99.9 percentile (1 out of 1000 applicants)** in Indian Institute of Technology Screening examination, 2005 **(IIT-JEE)**.

Publications

- **Dugar G**, Svensson S, Bischler T, Wäldchen S, Reinhardt R, Sauer M, Sharma CM. **The CsrA-FliW network controls polar localization of the dual-function flagellin mRNA in *Campylobacter jejuni*.** Nature Comm. 2016.
- Heidrich N, **Dugar G**, Vogel J, Sharma CM**. Investigating CRISPR RNA Biogenesis and Function Using RNA-seq.** Methods Mol. Biol. 2015.
- **Dugar G**, Herbig A, Förstner KU, Heidrich N, Reinhardt R, Nieselt K, Sharma CM. **High resolution transcriptome maps reveal strain-specific regulatory features of multiple *Campylobacter jejuni* isolates.** PLoS Genet. 2013.
- **Dugar G**, Gopinath B, Arun B, Sharma S. **Plant growth promoting abilities of phosphate solubilizers from the rhizosphere of *Parthenium hysterophorus.*** Afr. J. Microbiol. Res. 2013.
- Sharma S, Gupta R, **Dugar G**, Srivastava AK. **Impact of Application of Biofertilizers on Soil Structure and Resident Microbial Community Structure and Function.** Bacteria in Agrobiology 2012.

12. Acknowledgement

At the end of my thesis, I would like to thank all the people who have helped me along the way, both at the academic and personal fronts.

Firstly, I would like to thank my primary supervisor Prof. Dr. Cynthia M. Sharma for giving me the opportunity to work in her group and her wholesome support during my PhD thesis. She has guided and supported me through our frequent discussions, as well as encouraged me to participate in numerous international conferences.
Likewise, I am thankful to Prof. Dr. Jörg Vogel for his continuous guidance and encouragement, which helped me to focus my work on the most relevant questions in my research. His frequent reminders to 'think out of the box' and 'use proper controls' has enhanced my scientific skills to become a more mature and critical scientist.
I would also like to thank Dr. Tobias Ölschläger for being a member of my PhD thesis committee and for our fruitful discussions on *Campylobacter* and flagella. I would like to thank Dr. Martin Fraunholz for kindly agreeing on being an external reviewer for appraisal of this thesis on a short notice. Moreover, I would also like to thank Dr. Gabriele Blum-Oehler and the GSLS staff for their ongoing support. They have always lend an open ear for all my questions and concerns.

Next, I would like to thank all the members of the Sharma Lab who have continuously maintained an amicable atmosphere, inside and outside the lab. Ranging from scientific methods to chilling at the Main, I have learned a lot from all of them. I have been through high highs and low lows, and it would be hard for me to keep my spirit alive without being part of such a wonderful team. We all look out for each other as part of a big family. Sandy, Anika, Patrick, Thorsten, Mona, Marcus, Sarah, Sara, Eli and Beli – it is hard to put in words … THANK YOU again for everything. Sarah, my *Campy* expert, has taught me through long discussions the art of taming *Campylobacter.* I would like to thank Beli and her amazing pair of hands for the numerous experiments she has tirelessly performed for me. Sandy, Mona and Sara – Thank you for your prompt help on suggestions and corrections of this thesis. I would also like to thank the Bachelor and Master students who worked with me – Richie, Michi, Marta and Tina, for teaching me again that science is fun.

I would especially like to thank Konrad and Hilde for their numerous scientific help, but also and even more so for their infectious positive attitude and tranquil words. These words have often calmed my nerves and helped me grow as a person.

Furthermore, I would like to thank my friends, which I have met during my stay here. One can only dream to have the kind the friends I have been lucky to be associated with. Anna, Berny, Eggi, Gudrun, Jorge, Malvika, Nikola and Rabea – you have been my anchor and my force.

Lastly and most importantly, I would like to thank my family, my parents and my grandparents for their love, unmitigated support and their virtues.