INVESTIGATING THE ROLE OF CsgD IN *SALMONELLA* BIOFILM FORMATION AND VIRULENCE

A Thesis Submitted to the College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Microbiology and Immunology
University of Saskatchewan
Saskatoon

By

MELISSA BRIANNE PALMER

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Microbiology and Immunology
2D01, Health Sciences Building
107 Wiggins Road
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5E5
Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

When exposed to environmental stress, a pure culture of *Salmonella enterica* serovar Typhimurium (*S.* Typhimurium) differentiates into two specialized cell types with 34% differential gene expression: planktonic cells and multicellular aggregates, also called biofilm. Some conditions that support phenotype switching are known, but many intrinsic and extrinsic origins of signals are unknown. *S.* Typhimurium phenotype switching may promote transmission under variable conditions; planktonic cells express virulence factors and are immediately able to infect a new host, whereas aggregates can resist harsh environmental conditions until an opportunity to infect a new host arises. The objective of these research projects was to determine whether signals in the host gut could promote phenotype switching, and to determine the suite of genes controlled during phenotype switching, to understand this phenomenon and how it contributes to transmission.

Differences in expression between biofilm aggregates and planktonic cells are directed by bistable expression of CsgD, the central biofilm regulator. CsgD is expressed at low levels in planktonic cells and at high levels in biofilm cells, and coordinates the global shift in expression . A ChIP-seq experiment was performed to identify the regulatory targets of CsgD. The technique was refined for improved antibody binding and sample consistency; however, no statistically significant regulatory regions were identified by this method.

Phenotype switching could initiate in the host gut, as a result of extrinsic signals from the host or microbiota, during infection. Gene expression of virulence- or persistence-associated genes that were differentially expressed in RNA-seq data were measured by luciferase assay with promoter-*luxCDABE* reporter in the presence of chemostat waste effluent. The only major changes to gene expression levels or times in the presence of waste effluent may be due to additional resources for growth provided by the waste effluent and chemostat media control.

ACKNOWLEDGMENTS

Thank you to my supervisor, Dr. Aaron White for seeing potential in me as an undergraduate student. Thank you for challenging me to consistently "do my best". Thank you for your encouragement through my studies that led to many diverse experiences and wonderful opportunities. I have been so privileged to be mentored by you—you taught me to do practical elements of research well, to aim for excellence in science, to communicate findings effectively, and to love it all.

Thank you to current and former labmates for creating a culture of learning and levity. Thank you to Sumudu and Dylan for training me and teaching me how to navigate life in a lab as a student. Akosiererem, thank you for your camaraderie. Keith, thank you for teaching me innumerable techniques, commiserating with me when things didn't turn out, problem-solving with me, and demonstrating diligence and an unmatched enthusiasm for finding things out. Thank you to my One Health colleagues—you have enriched this journey immensely. Thank you to many others that have helped me in some way or another in this journey: Neil, Glenn, Alex, Kim, Alyssa, Muhammad, Tyson, Shirley, Jill, Laura, Donna, and Teresia.

Thank you to my parents for your constant support, listening ear, and for instilling in me a wonder of the world. Thank you to Max and my family through Jonathan—all of you inspire me your own different ways.

Jonathan, thank you—without you this thesis would not have been possible. I love the learning we do together. Your unconditional support is liberating and such a gift. Thank you for building a home, a family, and a life with me.

TABLE OF CONTENTS

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF APPENDICES

ABBREVIATIONS

CCD	Charge Coupling Device
c-di-GMP	Bis-(3'-5')-cyclic dimeric guanosine monophosphate
ChIP-seq	Chromatin immunoprecipitation coupled with sequencing
CPS	Counts Per Second
CRISPR-Cas9	Clustered Regularly Interspaced Short Palindromic Repeats-Caspase 9
DALY	Disability Adjusted Life Years
DC	Dendritic cell
DER	Duke Excluded Regions
EMSA	Electrophoretic mobility shift assay
ENCODE	Encyclopedia of DNA Elements
FASTA	Fast-All
$FMNH_2$	Riboflavin phosphate
GEO	Gene Expression Omnibus
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
IDR	Irreproducible Discovery Rate
iNTS	Invasive nontyphoidal *Salmonella*
IP	Immunoprecipitation
IPTG	Isopropyl β-D-1-thiogalactopyranoside
LPS	Lipopolysaccharide
mAb	Monoclonal antibody
MACS2	Model-based Analysis of ChIP-seq
MDR	Multi-drug resistant
MEME	Multiple Em for Motif Elicitation
O-Ag	O-Antigen
OD	Optical density
PMN	Polymorphonuclear lymphocyte
qPCR	Quantitative real-time polymerase chain reaction
rdar	red, dry and rough
SCV	*Salmonella* Containing Vacuole

SDS	Sodium dodecyl sulfate
SPI-1	*Salmonella* pathogenicity island 1
SPI-2	*Salmonella* pathogenicity island 2
SRA	Short Read Archive
TE	Tris ethylenediamine tetraacetic acid
TF	Transcription Factor
TBS	Tris Buffered Saline
T3SS	Type three secretion system
XLD	Xylose Lysine Deoxycholate

1.0 INTRODUCTION

1.1 Introduction to *Salmonella*

Salmonellae are gram-negative, oxidase negative, catalase positive, motile, facultative anaerobic bacilli that are 0.7-1.5 µm wide and 2.0-5.0 µm long[19]. Growth can occur between 4°C and 54°C; however, optimal growth is at 37°C, the human body core temperature[156]. They are infectious and are therefore classified as Risk Group 2 and Containment Level 2 organisms[19]. Salmonellae tolerate pH between 3.8 and 9.5, with optimal growth at pH 6.5-7.5[156]. The genus *Salmonella* was named after D. E. Salmon, who first isolated *Salmonella enterica* serovar Choleraesuis from porcine intestine in 1884[108]. In the laboratory, *Salmonella* can be isolated though growth on selective agar, such as XLD agar, Hektoen agar, or selenite agar with brilliant green[76]. They can be further identified though serotyping and genetic analysis[76]. In humans, *Salmonella* infections can cause gastroenteritis, enteric (typhoid) fever, bacteremia, or an asymptomatic carrier state[46]. This thesis is primarily concerned with *Salmonella enterica* serovar Typhimurium, which causes gastroenteritis and can be transmitted from other infected individuals, animals, or environmental sources via the fecal-oral route[49,176].

1.1.1 Classification of *Salmonella*

Salmonella is a genus in the family Enterobacteriaceae, order Enterobacteriales, class Gammaproteobacteria, and phylum Proteobacteria[174]. The *Salmonella* genus is further subdivided into species, subspecies, serovars, and strains based on taxonomy and serology (Figure 1.1). Originally, species were assigned based on clinical origins, and antigenic and biochemical characteristics[102]. In the 1980's, strains in the Salmonella genus were divided into two species by nucleic acid homology: *Salmonella enterica*[36] and *Salmonella bongori*[141,173]. These species are further subdivided by the White-Kauffmann-Le Minor scheme into more than 2 600 serotypes, based on flagellar (H1 or H2), oligosaccharide (O), and capsular polysaccharide

(K) surface antigens[40,64]. The most well known serotypes are *S. enterica* serovars Typhi, Paratyphi, Enteriditis, Typhimurium and Choleraesuis[173]. Diversity in surface antigens may be driven by immune selection[44]. Serovars are antigenically and genetically distinct with independently acquired regions which may include cytolethal distending toxins, anaerobic respiratory reductases, T3SS effectors, fimbriae, and iron acquisition systems[39]. *S. enterica* is divided into six subspecies[64,178] which are reffered to by name and roman numeral: *S. enterica* subsp. *enterica* (I), *S. enterica* subsp. *salamae* (II), *S. enterica* subsp. *arizonae* (IIIa), *S. enterica* subsp. *diarizonae* (IIIb), *S. enterica* subsp. *houtenae* (IV), *S. enterica* subsp. *indica* (VI)[64]. Subspecies *enterica* are commonly isolated from infections in humans and domestic mammals, whereas the other five subspecies are commonly isolated from cold-blooded animals[39].

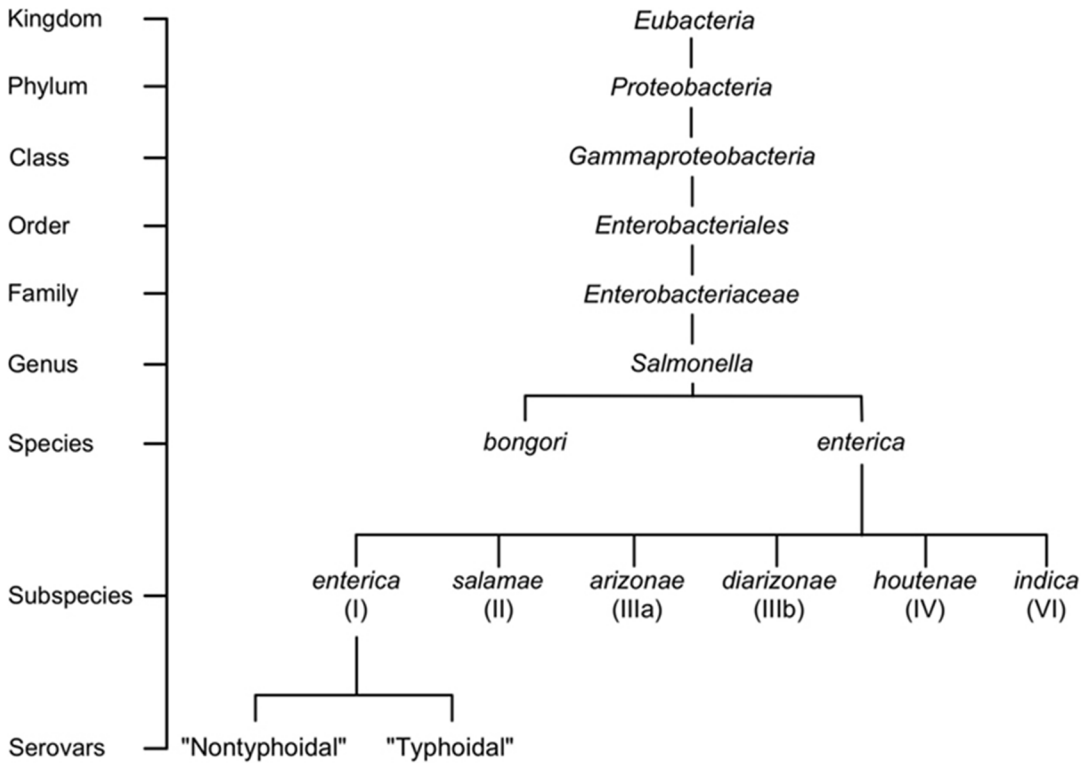


Figure 1.1. (Figure 1, MacKenzie *et al.* (2017) Frontiers in Veterinary Science[110]). *Salmonella* taxonomy and general classifications. The genus *Salmonella* is classified into species, subspecies, and serovars based on the White–Kauffman–Le Minor scheme. Serovars are often grouped into non-typhoidal or typhoidal categories; however, this referencing approach is not a part of the official *Salmonella* classification scheme.

Salmonella can also be categorized by the diseases that they cause in humans and the host organisms from which they are commonly isolated. Gastroenteritis, characterized by diarrhea and often accompanied by nausea, vomiting, and headache, is commonly caused by nontyphoidal *Salmonella* serovars *S.* Enteritidis and *S.* Typhimurium[46,176]. Typhoid or enteric fever, characterized by the invasion of extraintestinal tissue and often lacking symptoms of fever or inflammation, is caused by serovars *S.* Typhi, *S.* Paratyphi A, and *S.* Paratyphi B[176]. Recently, a third disease group, invasive nontyphoidal *Salmonella* (iNTS), has been identified as a group of NTS strains that can cause extraintestinal infections[49].

Salmonella are not limited to colonizing humans and can be described by their ability to colonize different hosts. Host generalist serovars are able to colonize several different organisms. For example, serovars *S*. Enteritidis and *S.* Typhimurium are able to colonize a broad range of organisms[22]. Serovars that are host-adapted mainly colonize one host organism and infrequently cause disease in other hosts. For example, *S.* Choleraesuis, *S.* Arizonae, and *S.* Dublin primarily colonize pigs, reptiles, and cattle, respectively[22]. Host restricted *Salmonella* cause disease in a single host. For example, host-restricted *Salmonella* serovars, *S.* Typhi and *S.* Paratyphi, cause disease in humans[92,110].

1.1.2 *Salmonella* diseases and epidemiology

The global impact of typhoidal, nontyphoidal, and invasive nontyphoidal *Salmonella* infections is considerable. There are an estimated 175 000 000 illnesses and 265 000 deaths that can be attributed to all invasive and diarrheal *Salmonella* infections[92]. *Salmonella* infections are responsible for an estimated 8.76 million disability adjusted life years (DALYs) from all transmission sources, and 6.43 million DALYs from contaminated food[92]. The global economic

costs of *Salmonella* as a foodborne disease are extensive; they are estimated from DALYs, hospitalizations, lost wages, producer litigation, premature death, lost tourism revenue, and trade embargo[175]. The economic burden of nontyphoidal *Salmonella* in the US, estimated using these parameters, was 3.67 billion dollars[75,161]. Gastroenteritis caused by nontyphoidal *Salmonella* is the most common clinical manifestation in both developed and underdeveloped nations[92]. Invasive nontyphoidal *Salmonella* is endemic in underdeveloped countries, particularly sub-Saharan African nations, with high incidence rates among infants and HIV-infected individuals[46]. Likewise, enteric fever is endemic in many African and Asian countries, with low incidence, less than 10 per 100 000 annually, in European and North American countries[46].

Salmonella is primarily transmitted through the fecal-oral route via food or water contaminated with feces from infected animals or humans. Poultry, swine, cattle, eggs, dairy products, some fresh fruits and vegetables are often sources of foodborne *Salmonella* infections[46] (Figure 1.2). Transmission from these sources may be expedited by long-term survival of persistent *Salmonella*[183]. Transmission is possible for the duration of shedding in feces, which could be up to 20 weeks for recently infected individuals, several years for chronic carriers, or intermittently or persistently for animal carriers[19]. By ingestion, the infectious dose is approximately 10^3 bacilli for nontyphoidal *Salmonella*, and 10^5 bacilli for enteric fever, although infectious dose varies with serotype for nontyphoidal *Salmonella* [19]. Of course, the infectious dose is much lower for immunocompromised individuals or individuals who are taking stomach-acid buffering medications, and onset of disease symptoms often depends on the size of the inoculum[19].

Figure 1.2. (Adapted, Figure 4, MacKenzie *et al.* (2017) Frontiers in Veterinary Science[110]*).* *Salmonella* strains that are host-generalists can encounter several host species and environments in a typical life cycle. Transfer of *Salmonella* to humans may occur zoonotically or through ingestion of contaminated vegetables (i.e., tomatoes, sprouts) or processed foods.

Gastroenteritis is usually characterized by acute onset non-bloody diarrhea, fever, abdominal cramps, nausea, vomiting, myalgia, and headache[46]. Nontyphoidal infections are localized to the ileum, but rare infections cause inflammation in the jejunum, duodenum, and stomach[16,26]. Uncommon complications of infection are hepatomegaly, splenomegaly, cholecystitis, appendicitis, and pancreatitis[76]. Nontyphoidal infections usually have a shorter incubation period than typhoid infections; symptoms usually occur 6-72 hours after ingestion[26,46]. Usually symptoms last 5-7 days and resolve spontaneously in self-limiting diarrheal disease[26]. Globally, gastroenteritis is predominantly caused by serovars *S.* Enteritidis, *S.* Typhimurium, and *S.* Newport[46]. After ingestion, virulence genes are expressed and colonization

initiates at the intestinal epithelium. Neutrophil recruitment and and inflammation in the host response leads to necrosis, edema, and fluid secretion, which manifests as diarrhea[26]. Typhoid or enteric fever is characterized by a gradually intensifying fever[46], headache, rose-coloured rash, abdominal pain, myalgia, and may be accompanied by bradycardia, hepatomegaly, splenomegaly, diarrhea (children) or constipation (adults)[96]. Infants, elderly, or immunocompromised individuals may experience complications[173] which may include hemmorhage due to gut perforations, myocarditis, encepalopathy, urinary tract infections, pancreatitis, hepatitis, and cholecystitis[19,46]. Symptom onset occurs a week or more after ingestion. Invasive disease is most commonly caused by serovars *S.* Typhi, *S.* Paratyphi, and *S.* Choleraesuis, although other serotypes are able to cause invasive salmonellosis[173]. After ingestion, *Salmonella* invade epithelial M-cells or are taken up by DCs in the intestine[26]. They translocate across the intestinal epithelium, access the host circulation, and infect other cell types and tissues[26]. Bacteremia is characterized by high fever and possibly leads to septic shock[46]. It is rare, more commonly caused by serovars *S.* Choleraesuis and *S.* Dublin, with a high mortality rate[26,46]. Comorbidity with other conditions is common in immunocompromised individuals[19]. Other extraintestinal complications of typhoid fever include urinary tract infections, pneumonia, endocarditis, meningitis, and cellulitis[46]. Asymptomatic carriers shed bacteria in feces for more than a year after acute symptoms of *Salmonella* infection[46]. Less than 5% of individuals who have experienced enteric fever, usually infants or elderly, become carriers. However, carriers of *S.* Typhi and *S.* Paratyphi are thought to be responsible for endemic transmission of enteric fever[46]. Carrier state development is even lower, approximately 0.1%, for nontyphoidal salmonellosis[46].

1.1.3 Pathogenesis of *Salmonella* infections

Pathogenesis of both nontyphoidal and typhoidal Salmonella serotypes is largely attributed to the function of two type three secretion systems (T3SS)[119]. The T3SSs mediate the transfer of virulence proteins, called effectors, into the host cytoplasm to stimulate uptake, invasion (SPI-1 T3SS), and survival in epithelial cells and macrophages (SPI-2 T3SS)[68]. The T3SS "needle" and effector proteins are encoded in large and unique chromosomal gene clusters called Salmonella pathogenicity islands[62].

Salmonellae initiate adherence and infection at the terminal ileum and colon through invasion of host intestinal epithelium[68]. After fimbrial contact with the apical surface of epithelial cells[12], salmonellae induce their own uptake[62]. SPI-1 injection of effectors SipA, SopB, SopD and SopE/E2 across the membrane activates cytoskeleton restructuring, or "membrane ruffling", which engulfs the bacterium at the epithelial cell surface[62]. SPI-1 also activates innate immune responses: NF-kB signaling, IL-1b and IL-18 activation, proinflammatory cell death, and polymorphonuclear lymphocyte (PMN) recruitment across the intestinal epithelia[26]. These inflammatory effects, in addition to tight junction disruption and SopB-mediated chloride secretion, contributes to the development of diarrhea[68,82].

Salmonellae survive inside host cells through SPI-2-mediated alteration of host cell functions, including cytoskeletal rearrangements, vesicle trafficking, signal transduction, and cytokine expression[68]. These functions are directed from Salmonella resident in the intracellular salmonella containing vacuole (SCV), distinct from phagosomes or lysosomes[59,68]. Invasive disease proceeds when salmonellae access extraintestinal tissues, through uptake by microfold (M) cells or dendritic cells (DCs), and transported throughout the body[84,144]. They ultimately reside in macrophages, DCs, PMNs, other phagocytes, and hepatocytes[123,145,193].

1.1.4 Susceptibility, treatment, and prevention of *Salmonella* infections

Effective methods of controling or eliminating microorganisms is key to preventing incidence of disease through transmission. *Salmonella* survival depends on the inoculum size and environmental conditions, particularly the surfaces on which biofilm are established, but persistence has been observed on vegetable sources for days, animal products for weeks, and in soil and water for more than a year[19]. Biofilm-forming serovars may survive for longer periods of time[32]. Salmonellae can be physically inactivated with moist heat at 121°C for at least 15 minutes, and dry heat at 170°C for at least 1 hour[19]. Incidence of infections can be decreased through proper food and water sanitation, pasteurization of dairy products, and elimination of human feces in food production[46]. Irradiation of food products is not practiced widespread, but can reduce the bacterial load and therefore the risk of foodborne illness[46]. Currently, there are vaccines to prevent enteric fever; however, vaccination does not protect against infections caused by *S.* Paratyphi and nontyphoidal serovars[46].

Established *Salmonella* infections are usually treated with fluid and electrolyte replacement, and antibiotics if necessary[19]. Antibiotics commonly prescribed are ampicillin, chloramphenicol, and trimethoprim-sulfamethoxazole[46]. However, antibiotic resistance is emerging as an issue for controlling *Salmonella* infections, since the first instance of *Salmonella* resistance to chloramphenicol was reported in the 1960s[140]. Salmonellae resistant to all three previously mentioned antibiotics are considered multi-drug resistant (MDR)[46] and are often more virulent than susceptible strains[51]. MDR strains decrease the efficacy of primary treatment defenses, and increase the morbidity and mortality of *Salmonella* infections. High frequencies of MDR *S.* Typhi have been isolated in Africa and Asia; MDR nontyphoidal serovars are increasing globally with additional resistances to cephalosporins and nalidixic acid[46]. Currently, fluoroquinolones and next generation cephalosporins have been introduced to treat MDR

serotypes[46]. Prudent use of antimicrobial agents is key; however, host immunomodulatory agents may enhance infection control in the future, with decreased risk of resistance[26].

1.2 *Salmonella* Biofilms

1.2.1 *Salmonella* biofilm characteristics

Bacterial biofilms are structural aggregates of cells embedded in a self-produced matrix, usually associated with a physical surface[32]. Bacteria in these communities are resistant to environmental stress, antibiotics, disinfectants, and immune defenses[171]. As such, they are difficult to eradicate and are major issues in treating infections and in food production. *Salmonella* biofilms may not be visible in the environment; however, their physical attributes have been described. *Salmonella* biofilm is described as the rdar (red, dry, and rough) morphotype on semisolid agar supplemented with Congo Red dye and grown on low salt media at temperatures below 30°C or on iron-depleted media at 37°C[29,155]. Congo red dye binds to cellulose and curli amyloid fibres produced by *Salmonella* biofilms, which turns colonies red[142]. In liquid media, biofilm has two forms that can be seen in Figure 1.3: 1) a pellicle at the air-liquid interface[153], and 2) aggregates in the liquid phase[111]. Biofilm aggregates in the pellicle and the liquid phase share key properties of rdar biofilm; they produce curli fimbriae and cellulose[186].

Self-produced extracellular matrices formed by *Salmonella* are composed of extracellular polysaccharides and protein components that link cells together. Protein components of biofilm include curli fimbriae, formerly named thin aggregative fimbriae (Tafi, *agf*)[150] and secreted BapA protein[101]. Curli are the major proteinaceous component of biofilm in *Salmonella* spp. and *E. coli*[11]. They are non-branching[20], amyloid structures that are resistant to proteases and sodium dodecyl sulfate[29]. Curlin protein subunits composed of repeating strand-loop-strand motifs are folded into a parallel β-helix with extensive hydrogen bonding that may contribute to its

disproportionate stability[30]. Genes encoding the curlin protein and proteins for gene regulation, secretion, and assembly are encoded by the divergent operons *csgBAC* and *csgDEFG,* formerly *agfBAC* and *agfDEFG* [28,150]. Regulation of biofilm formation, directed by CsgD, is complex and will be discussed later. Curli components encoded by the two operons are translocated through the inner membrane by Sec secretion, and assembled at the outer membrane[11]. CsgB encodes the minor nucleating subunit, onto which CsgA, the major structural subunits, are added. Curli fibers are involved in short-range adhesion to surfaces and other salmonellae in biofilm formation[182], and to host cells in invasion[11]. Initial stages of biofilm formation are mediated by curli fimbriae interactions, which also allow salmonellae to adhere to Teflon and stainless steel surfaces[8]. Curli fimbriae are also hypothesized to have an important role in invasion and dissemination, due to their ability to bind to host cells[130], fibronectin[129], laminin[128], plasminogen, and factor XII[72]. BapA, the other protein component of biofilm, is a very large (386 kDa) multidomain protein associated with the cell surface[101]. It is encoded by BapA in the *bapABCD* operon, which also encodes its own type I protein secretion system, BapBCD[101]. Experiments have demonstrated the important role of BapA in bacterial aggregation, pellicle formation, and host colonization[101]. BapA may mediate short range cell interactions in biofilm formation through direct protein-protein binding, or by supporting curli fimbriae binding[101].

Polysaccharide components of biofilm include cellulose[197], O-antigen capsule[57], and small amounts of LPS. Cellulose is made of β(1→4)-D-glucose repeating linear polysaccharide chains arranged in a rigid matrix[197]. This polysaccharide lattice facilitates long-range surface or cell-cell interactions and, together with curli fimbriae, holds cells together[197]. It is essential for biofilm formation in lab models[171], epithelial cell surfaces[103], and abiotic surfaces[139], and colonization of plant tissues[10]. The cellulose biosynthesis machinery, which produces and

assembles polysaccharides, is encoded by the divergent operons *bcsABZC* and *bcsEFG* [166,197]. BcsA and BcsB (bacterial cellulose synthesis) encode the major functions for cellulose synthesis[152]. The BcsA-BcsB complex is activated when the secondary messenger c-di-GMP binds the BcsA PilZ domain[120]. Following activation, the enzyme polymerizes uridine-5'-diphosphate α-D-glucose (UDP-glucose), originating from glycolytic intermediate glucose-6-phosphate, to cellulose polymers[152]. Growing cellulose chains are synthesized on the cytoplasmic surface of the inner membrane, and exported to the cell surface[152]. The other subunits of cellulose synthesis alter yield, quality, and enzyme activity through modulating expression of biosynthesis genes, exporting the nascent chains to the cell surface, and organizing cellulose fibers once they have been exported[152]. The O-antigen capsule is a highly hydrated lipid-anchored exopolysaccharide with more than 2300 repeating oligosaccharide units, produced during biofilm formation[57,165]. It is structurally similar to LPS O-Antigen, but differs in size, charge, lipid attachment, side chain modifications, and immunoreactivity[57,165,185]. During assembly, a carrier lipid is flipped across the inner membrane, where O-Ag oligosaccharide units are polymerized, after which the polysaccharide is translocated across the outer membrane to the cell surface[190]. Capsule surface assembly and translocation functions are encoded by *yihUTSRQPO* and *yihVW,* which are conserved in the *Salmonella* genus[57]. The O-Ag capsule allows salmonellae to tolerate desiccation, and has proposed functions in environmental persistence and, ultimately, transmission to new hosts[57,187]. In addition, the O-Ag capsule is important for *Salmonella* attachment and biofilm formation on gallstones[35] and alfalfa[10].

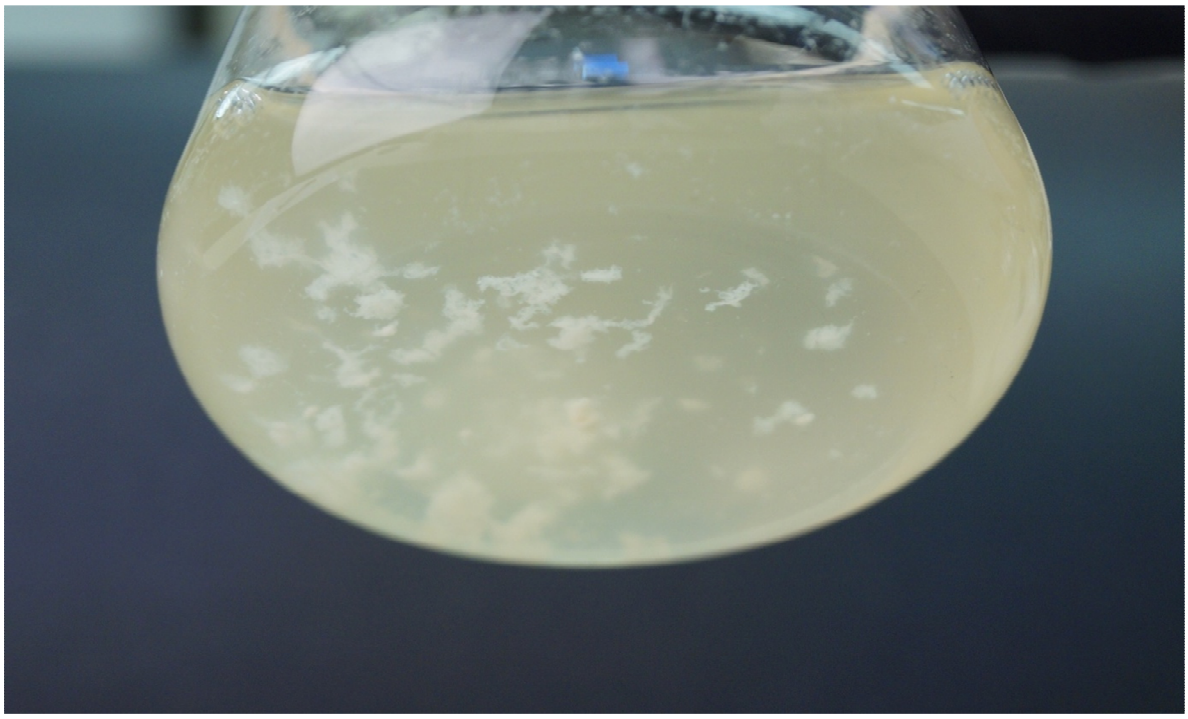


Figure 1.3. Flask culture of *S.* Typhimurium containing biofilm aggregates and planktonic cells. Salmonella Typhimurium grown in 1% tryptone for 13 hours at 28°C form phenotypically heterogeneous biofilm aggregates and planktonic cells in the liquid phase of the flask culture.

1.2.2 Regulation of *Salmonella* biofilms

The suites of genes that are expressed in biofilm are significantly different from those expressed by free-floating planktonic cells[33]. Expression and synthesis of the biofilm components described above, which are involved in stress response and carbon central metabolism[189], occur in response to environmental stress signals and the interactions of a complex regulatory network (Refer to Steenackers *et al*. for an illustrated overview of regulation[171]) *Salmonella* form biofilm at temperatures lower than 30°C, in microaerophilic environments[56], in response to iron limitation[155], in the presence of intracellular bis-(3′–5′)-cyclic dimeric guanosine monophosphate, or cyclic-di-GMP[162], and in response to nutrient limitation[110]. However, iron limitation can permit biofilm formation at temperatures that are higher than normal biofilm conditions[110].

The master biofilm regulator, CsgD, is the central modulator of gene expression in *Salmonella* biofilm formation. We now understand that biofilm cells, as differentiated from planktonic cells, arise due to bistable production of CsgD (i.e., biofilm cells are CsgD-ON; single cells are CsgD-OFF)[61]. Bimodal expression of CsgD is thought to be a result of positive autoregulation, which leads to a regulatory cascade causing broad changes in expression, and ultimately phenotype, in a subset of the population[111]. CsgD is a DNA-binding protein with a response regulator domain that binds promoter regions actively when unphosphorylated[194]. The full suite of genes that are controlled by CsgD remains to be identified; however, CsgD is known to increase expression of *adrA*, a diguanylate cyclase required for c-di-GMP production[154,162], and the *csg* operon, which produces curli fimbriae and additional CsgD in a feed-forward loop[150].

The regulation of *csgD* expression and synthesis is complex. Trans regulators of *csgD* expression include OmpR, IHF, H-NS, CpxR, and MlrA[171]. OmpR, which is phosphorylated in response to high osmolarity and can bind DNA in high and low osmolarity, regulates CsgD

expression[155]. MlrA is also a positive regulator of *csgD* expression, which is suspected to rescue the biofilm phenotype in lab-domesticated *Salmonella* strains grown in LB [188]. Histone-like proteins IHF (integrating host factor) and H-NS control several genetic processes, including CsgD regulation. IHF enhances *csgD* transcription when bound to the IHF1 site, which is also an OmpR binding site, and represses transcription when bound to the IHF3 site in the *csgD* coding sequence[55]. H-NS is an architectural protein that binds to AT-rich and bent regions of DNA; it activates or represses *csgD* transcription depending on binding location[56]. The stress response protein CpxR negatively regulates biofilm formation when phosphorylated[43]. Cis downregulation by posttranscriptional control may occur through small RNAs, OmrA and OmrB from the 174 bp 5' untranslated region of *csgD* mRNA[77]. This is in addition to cis regulation of the *csg* operon by CsgD itself.

Auxiliary regulators RpoS, Crl, c-di-GMP, PhoPQ, RstA, and Rcs also regulate biofilm formation. RpoS (σ^S) is a sigma factor that forms a holoenzyme with RNA polymerase to direct transcription of a set of promoters during stress response or stationary phase[71]. At the beginning of stationary phase, RpoS positively regulates expression of *csgD* [155,188], *mlrA*[17], *csgBAC*, and *adrA*[171]. The DNA-binding transcriptional regulator Crl acts with RpoS to increase expression of biofilm genes *csgB*, *bcsA*, *csgD*, and *adrA*[147]. Crl is a proposed temperature-sensing protein for biofilm regulation, due to its reduced activity at temperatures above 28°C[146]. The secondary messenger bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP) promotes virulence and motility at low intracellular levels, and cellulose biosynthesis, curli biosynthesis, and sessility at high intracellular levels[162,171]. Levels of c-di-GMP in the cell are adjusted through c-di-GMP synthesis and degradation. It is synthesized by diguanylate cyclase enzymes with an active GGDEF catalytic domain from two molecules of guanosine triphosphate (GTP), and

degraded by phosphodiesterase enzymes with active EAL domains into 5'-phosphoguanyl-(3'-5')-guanosine (pGpG), which is converted to two molecules of guanosine monophosphate (GMP). AdrA, which is regulated by CsgD, was the first diguanylate cyclase described in biofilm formation[154,162]. It produces c-di-GMP, which binds to the BcsA PilZ receptor, activating the enzyme to produce cellulose[120]. Two component signaling system consisting of inner membrane sensor kinase PhoQ and cytoplasmic response regulator PhoP[87] responds to low concentrations of divalent magnesium[53]. Activated PhoP controls the expression of over 100 genes[87] that are required for LPS modification, host cell invasion, survival in macrophages[21], and most importantly, repression of biofilm formation[139]. RstA, a response regulator that controls RpoS degradation, also downregulates expression of *csgD* [127] and *bapA* [18], the large biofilm protein. Repression of genes involved in flagellar synthesis and virulence with concurrent upregulation of genes involved in capsule synthesis can also occur when the Rcs phosphorelay system is activated[112,171]. The transcriptional regulator RcsA is activated through a system response to high osmolarity or cationic antimicrobial peptides[23]. Although the mechanism is unknown, quorum sensing through several density-dependent pathways are thought to regulate biofilm formation[171].

1.2.3 The role of biofilm formation in the life cycle of nontyphoidal *Salmonella*

Bacterial biofilms are sessile, highly organized, often polymicrobial communities that are inherently resistant to antimicrobials and a common cause of persistent and chronic infections[32]. Their structural organization and programmed detachment of cells facilitates cooperation and survival in otherwise uninhabitable environments. *Salmonella* biofilms are resistant to antimicrobials, bile, disinfectants, desiccation, starvation, and immune responses[171]. Biofilm formation is conserved among nontyphoidal *Salmonella* serovars[186,187], and is therefore expected

to have a significant evolutionary role. The life cycle of host generalists is dynamic and often unpredictable. Strategies of both survival and virulence are required by *Salmonella* serovars for transmission to new hosts. Phenotype switching observed in *S.* Typhimurium clonal populations, initiated by bistable expression of CsgD in response to environmental stress, produces biofilm aggregates and planktonic cells expressing virulence factors[111]. This may be a "bet-hedging" strategy whereby transmission could occur in the short-term by planktonic cells or long-term by biofilm. It was noted that strains lacking the ability to form biofilm tend to be more virulent and cause invasive disease [151]. Loss of the biofilm phenotype was observed in host-adapted serovars *S.* Typhi, *S.* Paratyphi, and NTS strains that are associated with invasive disease[110]. It appears that selection pressures for maintaining the biofilm phenotype are opposing: stress vs. nutrition, intestinal vs. systemic replication, host-restriction vs. host-adaptation, short- vs. long-term transmission.

1.3 ChIP-seq

1.3.1 Development of the ChIP method

Chromatin immunoprecipitation (ChIP) is a method of identifying DNA-protein interactions *in vivo*[168]. In ChIP, DNA bound by a transcription factor, nucleosome, or other DNA-associated protein is enriched by immunoprecipitation with an antibody against the protein of interest[98,133]. In the first description of this method in 1985, Solomon *et al*. leveraged the protein crosslinking characteristics of formaldehyde to map the distribution of nucleosomes on the chromosome[169]. Conventional ChIP was described later, in which formaldehyde-crosslinked interactions on fragmented genomic DNA are "pulled down" with a specific antibody[80,107,143] (Figure 1.4). Immunoprecipitated DNA was detected for analysis by radioactive labeling, quantitative real-time PCR (qPCR) and microarray (ChIP-chip) [80,107,143]. Previous methods of

studying regulation included systematic evolution of ligands by exponential enrichment (SELEX), electrophoretic mobility shift assays (EMSA), ChIP-chip, promoter deletion, and reporter analysis, which have been complemented or replaced by ChIP-seq [54]. Sequencing has replaced ChIP-chip due to decreases in cost for next-generation sequencing, accuracy and depth of data output[133], less input material required, availability of high-quality reference genome sequences[148], and opportunities for integration with other datasets[98]. Currently, there are diverse methods of performing ChIP against transcription factors, histone modifications, chromatin modifying complexes[98], enzymatic or sonication-based DNA fragmentation[126], UV or chemical crosslinking[168], and even targeting CRISPR-Cas9 chromatin complexes[195]. Additionally, there are diverse methods and tools for analyzing ChIP-seq datasets[98,121,135].

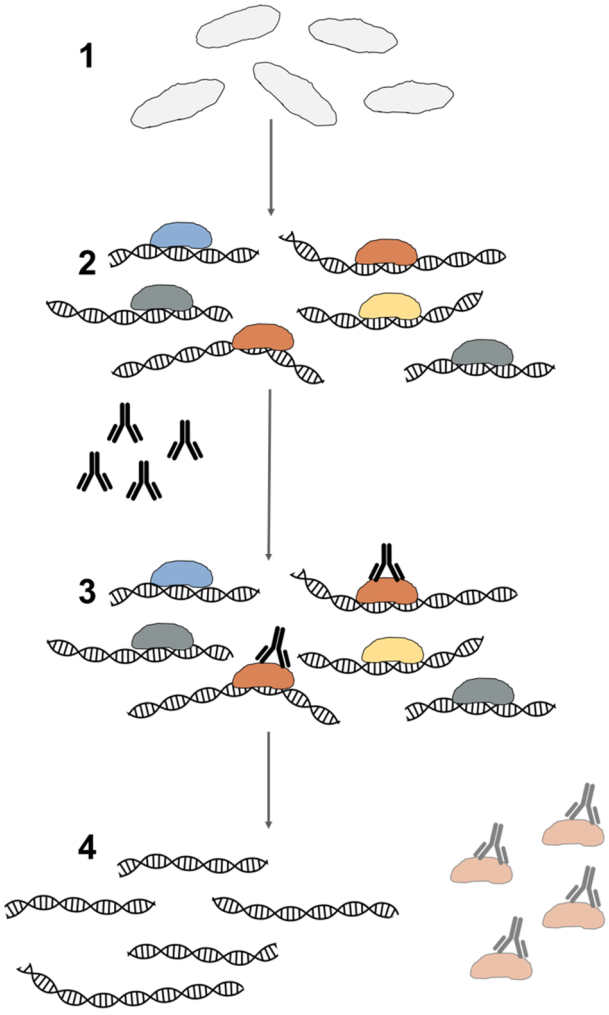


Figure 1.4. (Adapted, Hoffman & Jones. (2009) Journal of Endocrinology[74]). Simplified steps involved in Chromatin Immunoprecipitation (ChIP). In ChIP, 1) cells are harvested and 2) treated with formaldehyde crosslinker to covalently link proteins to DNA. Cells are lysed to isolate genomic DNA and chromatin is sheared to a fragment size range acceptable for downstream sequencing. 3) An antibody binds the protein of interest and the antibody-protein-DNA complex is selected by immunoprecipitation. 4) Crosslinks are reversed, protein and RNA is digested, and the resulting DNA is purified for library preparation and sequencing.

1.3.2 ChIP for identification of regulatory targets in bacteria

ChIP has been primarily used for identifying regions controlled by nucleosomes, histone modifications, and protein-DNA interactions in eukaryotes. It has also been used as a tool for identifying genomic regions controlled by DNA-binding proteins in bacteria[121]. In particular, ChIP is used to discover transcription factors within bacterial regulatory networks. ChIP has been used to analyze Sfh and OmpR regulation in *S.*Typhimurium[41,136], quorum-sensing AphA regulation in *Vibrio alginolyticus*[65], RpoS regulation in *Escherichia coli*[134], virulence regulator EspR regulation in *Mycobacterium tuberculosis*[14], potential diguanylate cyclases through AmrZ regulation in *Pseudomonas aeruginosa*[85], as well as VraSR regulation in *Staphylococcus aureus*[160]. Regulatory information from these experiments can be used to understand more about bacterial ecology, pathogenesis, and transmission.

1.3.3 Guidelines for ChIP experiments

The Encyclopedia of DNA Elements, or ENCODE, is a consortium that has developed guidelines, practices, and quality metrics to be applied to ChIP-seq experiments[133]. These guidelines account for the specificity and quality of immunoprecipitation, experimental design, sequencing depth, methods of data analysis, and reporting data[98]. The consortium has performed more than a thousand ChIP-seq experiments targeting different DNA-associated proteins in different organisms with different sequencing platforms and analysis pipelines[31]. These guidelines are built for experiments with eukaryotic organisms, specifically *Homo sapiens*, *Mus musculus*, *Caenorhabditis elegans*, and *Drosophila* spp.[98]; however, they can be useful to model experiments targeting transcription factors in bacteria.

Recommendations for pre-ChIP quality checks, experimental design, and analysis are given by the ENCODE consortium. Antibody specificity must be tested by primary and

secondary methods to rule out poor reactivity and cross-reactivity[98]. Primary characterization by immunoblot on whole-cell extracts, nuclear extracts, chromatin preparations, or immunoprecipitated material, with the major band within 20% of the expected size and representing more than 50% of the detected bands[98]. Secondary characterization may be any one of five data: 1) an immunoblot with a knockdown or knockout of the target protein, 2) mass spectrometry following immunoprecipitation, 3) immunoprecipitation with multiple antibodies against different epitopes on the target protein, 4) immunoprecipitation with an epitope-tagged target protein, and 5) motif analysis[98]. ChIP samples should be prepared in biological duplicates with a control that matches the strain, growth stage, and treatment of test samples. Control DNA may either be pre-IP sonicated DNA (i.e. "Input") or immunoprecipitated with antibody that binds irrelevant protein, type matched to the IP antibody (i.e. "Control IP")[98]. Statistical power of peak discovery is stronger with more reads; therefore ENCODE recommends 20 million uniquely mapped reads per replicate for transcription factor ChIP with mammalian genomes. Fewer uniquely mapped reads are required for smaller reference genomes. Peak calling algorithms are not specified[98], but statistical analysis of global enrichment[83], cross-correlation, and IDR (irreproducible discovery rate)[106] are specified by ENCODE. The IDR statistic compares peak pairs in replicates and is more consistent across laboratories, antibodies, and analysis protocols[106]. Analyzed data can be submitted to GEO (Gene Expression Omnibus) with metadata, and raw data can be submitted to the Short Read Archive (SRA) for public access.

1.3.4 Sequencing ChIP DNA

Sequencing is now more common than PCR or microarray methods for obtaining data from ChIP experiments. ChIP-chip and PCR-based methods yield adequate regulatory data, but often require expensive probes and introduce amplification or hybridization and array bias [91].

ChIP-seq has a larger dynamic range with increased nucleotide resolution and coverage, improved signal-to-noise ratio and fewer artefacts when compared to ChIP-chip [133].

Prior to sequencing, libraries are prepared from ChIP DNA by repairing ends and phosphorylating 5' ends, A-tailing 3' ends and ligating adapters, enriching adapter-ligated DNA by PCR, and adding index barcodes[70]. Barcode sequences allow for post-sequencing identification of replicates or samples from different growth conditions for differential analysis[133]. Libraries are then quantified and multiplexed together with differentiating library barcode sequences, before sequencing with a high-throughput next-generation sequencing platform[121]. ChIP-seq experiments with localized transcription factors should yield 10-14 million unique reads[98] or 100x coverage at each base pair on the reference genome[78].

Presently, most ChIP experiments are sequenced on the Illumina platform, which generates millions of forward and reverse reads ranging from 25-300 base pairs depending on the chemistry and kit used. Illumina sequencing can be either single-end, from one end of the DNA fragment, or paired-end, from both the 3' and 5' ends of the DNA fragment[94,133]. Sequencing data is accessed as a series of nucleotide sequences and quality scores with identifiers such as the instrument identification and the run number in FastQ file format[27]. Depending on the sequencing platform and data output, formatting may be required before analysis.

1.3.5 Bioinformatic methods for analyzing ChIP-seq data

ChIP-sequencing data output volume and complexity require stringent statistical approaches for proper analysis[179]. There are many tools available to researchers for analyzing ChIP-seq datasets, and new tools are continuously being developed. Analysis of ChIP-seq data includes obtaining sequencing data, assessing read base quality, trimming reads by base quality,

mapping sequence reads to a reference genome, calling peaks at high coverage depths, annotating peaks and associating the peaks with genes, and downstream analysis[121].

Initial quality control should be performed on the data to determine coverage of each nucleotide position on the genome and base identity by quality score at each position[121]. FastQ file formats, which have become common sequencing outputs, bundle FASTA nucleotide information with per base quality score [27]. Tools such as FastQC are available for determining data quality for other file formats[6]. Any remaining Illumina adaptors or bases that have unacceptably low quality are then removed from the reads by a trimming algorithm[15]. If the reads are from a paired-end experiment, they may be aligned at this step[15]. After trimming, the reads are then mapped to a high-quality reference genome[121]. Algorithmic tools BWA or Bowtie2 are commonly used for read mapping[100,105]. To identify areas of enrichment where a transcription factor binds or histone modification is located, peaks of high densities of reads are detected[179]. The peaks that are "called" are associated to particular genes based on proximity to the transcriptional start site of a gene[121]. Peak calling is the most important step in ChIP-seq analysis. Alignments can be visually scanned to identify peaks; however, algorithms remove subjectivity and find statistically significant peaks[121]. Peak calling algorithms use different methods of identifying peaks, but mainly do so through comparison of reads or change in read density to background (i.e. "Input-seq" or "Control-seq") read mapping[90,135,196]. With these methods, the depth of sequencing is important. Inadequate sequencing depth will not fulfill genome site saturation criteria and peaks will be indistinguishable from background, whereas adequate sequencing depth will result in saturation of background genomic sites and statistically significant peaks.

Once peaks are known, downstream analysis is performed. The data can be visualized on a genome viewer, which is useful for identifying false positives and biologically interesting sites that are not found by algorithms[121]. At this point, called peaks can be annotated and associated with genes based on proximity to the transcriptional start site of the genes[121] .The common binding sequence, or motif, is identified and analyzed against the rest of the data for over-representation and confirmation[121]. MEME is a common tool used for motif analysis[9]. Peak and motif data should be compared among biological replicates to ensure data is consistent[121]. The peak data can also be used to assess differential binding between samples harvested from different conditions, or it can be integrated with transcriptomic gene expression data to understand regulons[121,192]. Usually ChIP-seq data is complemented or validated with phenotypic or biochemical assays, such as in vitro DNA binding assays (i.e. EMSA or DNAse I footprinting) or expression assays (i.e. RT-qPCR or luciferase assays)[121,132].

1.4 Luciferase reporters

1.4.1 *Photorhabdus luminescens* bacterial luciferase

Bioluminescence is derived from chemical reactions between two or more molecules in or associated with an organism[63]. This phenomenon is observed in marine and terrestrial organisms, and most commonly in single celled organisms[24]. Light-producing bacteria have been observed for centuries, and have been characterized at a molecular level since the mid 1900s[116,172]. It is expected that bioluminescence, an energetically expensive pathway, has been retained due to its important evolutionary roles in nature[63]. Light production by symbiotic bacteria can help higher trophic host organisms attract prey, camouflage, or attract a mate, in return for a stable and nutrient rich habitat [124]. It can also be involved in bacterial communication[63]. The majority of light-producing organisms are in three bacterial genera:

Vibrio, *Photobacterium*, and *Photorhabdus* [25,177]. Bioluminescence originates from signals produced by bacterial luciferase (Lux), firefly luciferase (Luc), and green fluorescent protein[25]. In *Photorhabdus luminescens*, gene products from the *luxCDABE* operon LuxC, LuxD, and LuxE supply and regenerate the aldehyde substrate for the luciferase enzyme, a LuxAB α/β heterodimer. Blue-green light at 490 nm, and secondary emission at 590 nm is released when the luciferase enzyme (LuxAB) oxidizes a reduced flavin mononucleotide ($FMNH_2$) in the presence of a long-chain fatty acid and oxygen[25,117]. This system can produce light independently, in response to an environmental signal, and in different host organisms[24].

1.4.2 Quantitative promoter activity from luciferase assays

Bacterial bioluminescence is a popular, sensitive, and simple *in vivo* reporter for gene expression[47,63]. Over the years, bacterial luciferase technology has been improved and adapted for use in different organisms[24]. In 1985, Engebrecht *et al*. inserted the *lux* operon via mini-Mu*lux* transposon into *E. coli* [47]. Four years later, Olsson *et al*. constructed a monocistronic luxAB gene from *Vibrio harveyi*, which could be used as a reporter in both prokaryotes (e.g. *E. coli*), and eukaryotes (e.g. *Nicotiana tabacum*) [131]. The *Photorhabdus luminescens luxCDABE* genes were later cloned and expressed in *Saccharomyces cerevisiae*[67] and the cassette was inserted into a human cell line via human cytomegalovirus promoter and *tet* operator sequences[60]. Developments to luciferase reporter systems have made them excellent tools for real-time, non-invasive, sensitive detection of biological functions[63,73]. Advantages to this approach include 1) independent light production when expressed with aldehyde biosynthesis genes[24], 2) quantitation due to quantum yields of light[177], and 3) non-destructive monitoring of gene expression[131]. Luciferase bioluminescence can be detected using photographic film, camera Charge Coupling Device (CCD), or a luminometer[177]. The sensitivity of detection is dependent

the level of luciferase expression, the detection equipment, the presence of ATP or oxygen, or the optical properties of the tissue or growth media[157]. Recently, the luciferase reporter has been used as a biosensor for gene expression, *in vivo* diagnostic imaging[125], biocomputing[163], gene arrays[180], and high-throughput monitoring for environmental, agricultural, pharmacological, and clinical settings[25,67].

1.5 Origin and composition of chemostat waste effluent

The diversity and functional capacity of resident organisms of the human gut microbiome may play a role in maintaining health and preventing gastrointestinal infections[66]. Commensal gut microbiota perform this function by priming dendritic cells, producing bactericidal products that inhibit pathogenic bacteria, competing for space and resources on the intestinal epithelial and mucosal surfaces, and maintaining colonization resistance in persistent polymicrobial biofilms[33,149,170]. Molecular or pure culture techniques used to study organisms in a complex intestinal community do not provide comprehensive ecological information[3,115]. Culturing fastidious and anaerobic bacteria from the human gut microbiome is challenging and limited by end-point analysis[4,114]; however, it is essential for understanding metabolic and physical community interactions and clinically relevant or low abundance microorganisms[97]. A chemostat seeded from a homogenous fecal inoculum and brought to equilibrium can imitate the original distal gut community *in vitro*[4]. This system facilitates culturing of both biofilm and planktonic communities, and maintains species evenness and richness[4,114]. Specifically, the reactors of a chemostat are packed with a fecal inoculum from healthy donors that have not been treated with antibiotics in the previous nine months[114]. The fecal inoculum is cultured with a modified growth media containing mucin for biofilm development and antifoam B silicone emulsion for continuous flow at a rate of 400 mL/day to mimic a 24 hour transit time[115]. The entire system is

incubated at 37°C for 36 hours to reach steady-state equilibrium[114]. Effluent from the steady-state chemostat is collected into sterile bottles and filtered (0.2 μm) to produce a cell-free spent media extract [4]. This "Liquid Gold", as named by Emma Allen-Vercoe (University of Guelph) can be used as a media supplement for culturing fastidious organisms *in vitro* [4]. It is thought to contain small molecules present in the gut that support the growth of previously unculturable organisms [4,7]. Small molecules of mammalian origin are often hormone messengers[104] and small molecules from microbial origin are often used for communication[52]. Therefore, small molecules in the gut from host, microbiome, and invading pathogens may control the balance between health and disease. For example, in the presence of small molecules extracted from feces, *Salmonella* invasion gene expression is highly repressed[7]. In particular, a small molecule secreted by Clostridia acts as a repressor of SPI-I virulence genes and prevents *Salmonella* invasion of host cells[7].

2.0 HYPOTHESIS AND OBJECTIVES

2.1 Rationale and hypotheses

Salmonella enterica serovar Typhimurium (*S.* Typhimurium) is a major cause of gastroenteritis worldwide[92]. The mechanisms of *S.* Typhimurium transmission between humans, the environment, and animals are not well known, but an understanding of these mechanisms may be key for developing methods of preventing infection. Due to the dynamic environments that it experiences during its life cycle, *S.* Typhimurium requires different strategies for survival and transmission to new hosts. Mackenzie *et al*. demonstrated that a pure culture of *S.* Typhimurium exposed to environmental stress differentiates into two coexisting cell types: multicellular aggregates that are resistant to desiccation and antibiotics, and planktonic cells that express high levels of virulence factors[111]. RNA-seq analysis revealed 34% differential gene expression between the two cell types that may be directed by bi-stable expression of *csgD*, the *Salmonella* master biofilm regulator[111]. Biofilm cells express high levels of CsgD, whereas planktonic cells express low levels of CsgD[61]. We hypothesize that phenotype switching represents a bet-hedging strategy for *Salmonella*, since the infective population has limited time to react to new environments following expulsion from the host. Therefore, it is advantageous for the population to retain a portion of planktonic cells primed for immediate infection, and a portion of biofilm aggregates prepared for survival and persistence in harsh environments until an opportunity for infection arises. Accordingly, engaging phenotype switching during the *S.* Typhimurium life cycle supports several routes of infection and of course, evolutionary reproductive success.

The regulatory network that induces biofilm formation is complex, involving many extrinsic environmental input signals and intrinsic regulatory elements[171]. We do not fully understand the genetic changes associated with *Salmonella* phenotype switching. Elucidating the

regulatory network leading to genetic and phenotypic heterogeneity during the *S.* Typhimurium life cycle may be key for blocking transmission. Currently, we do not know the full suite of genes controlled by the master biofilm regulator, CsgD. I hypothesize that CsgD regulates the expression of *S.* Typhimurium genes in multicellular aggregates at high abundance and planktonic cells at low abundance. Extrinsic signals for initiation of phenotype switching are complex and may originate from cell age, stochastic events, cell-to-cell interactions, or signals received from the host during the course of infection[1]. I hypothesize that the signals for phenotype switching originate in the intestine, including signals from both the host and the microbiota.

2.2 Objectives

The objectives of my research project were as follows:

- To develop a method of preparing *S.* Typhimurium cell types for ChIP-seq

- To identify genes regulated by CsgD in the *S.* Typhimurium 14028 genome using ChIP-seq, and assess differential regulatory targets of CsgD between biofilm and planktonic cells

- To identify differences in the expression of *S.* Typhimurium virulence- or persistence-associated genes in the presence of waste effluent from a human microbiome culture using luciferase assays

3.0 CHIP-SEQ FOR IDENTIFYING THE REGULATORY REGIONS OF CsgD

3.1 Introduction

CsgD, the DNA-binding *Salmonella* master biofilm regulator, is involved in a complex regulatory network that leads to biofilm formation in limiting environmental conditions[171]. It is known that biofilm aggregate cell types have high intracellular levels of CsgD, and free-floating planktonic cells have low intracellular levels of CsgD[61]. CsgD is a transcriptional regulatory, DNA-binding protein, and the large difference in gene expression between the two specialized cell types[111] is at least in part due to genetic regulation directed by CsgD. Many of the regulatory elements that control *csg* operon expression at the intergenic region are well characterized[171], but few regulatory targets of CsgD are known. The goal of my research project was to use ChIP-seq to find the regulatory targets of CsgD, apart from known targets *adrA*[162] and the *csg* operon[154].

ChIP-seq, a method of identifying DNA-protein interactions *in vivo*[168], is an ideal technique for finding the full suite of CsgD regulatory targets on the *Salmonella* genome. In this technique, cells are harvested and DNA-binding proteins are crosslinked to DNA *in vivo* through formaldehyde cross-linker treatment[133]. Cells are lysed, releasing the cell contents, and chromatin is sheared into approximately 200-600 bp fragments. DNA fragments bound to the target protein are immunoprecipitated with antibody, and isolated by protein digestion following decrosslinking[179]. These DNA fragments represent protein binding sites matched by hybridization to a microarray (ChIP-chip) or by deep sequencing and mapping to the genome sequence[107,148]. ChIP coupled with next-generation sequencing produces massive and high-resolution genome-wide datasets that could reveal all possible regulatory targets of CsgD. Few methods in the literature describe ChIP with bacteria, and most bacterial experiments begin with a simple growth culture[14,41,65,85,136]. ChIP sample preparation with single cells is straightforward and presents fewer challenges than ChIP sample preparation with biofilm aggregates. The

extracellular material surrounding biofilm aggregates interferes with crosslinking, accurate cell number estimation, and lysis. Therefore, my objectives in this research project were to optimize ChIP experimental conditions with the goal of identifying genes that are differentially regulated in biofilm and planktonic cell types.

3.2 Materials

3.2.1 Antibodies, bacterial strains, and culture conditions

DNA samples were prepared from wild type *Salmonella enterica* serovar Typhimurium, ATCC 14028S (GenBank CP001363.1)[81]. Positive and negative controls were prepared as a comparison to wild type biofilm and planktonic cell types. A strain that overexpresses CsgD, *S.* Typhimurium ST14028 *ΔcsgD* + *pACYC csgDcompFOR3*, was built as a positive control. In the flask culture, this strain overproduces biofilm aggregates and produces few planktonic cells. The negative control, *S.* Typhimurium ST14028 *ΔcsgD*, does not produce biofilm. These cell types were prepared from flask culture in limiting conditions, 1% tryptone media at 28°C (reviewed in Appendix B). Cells were harvested from flask culture after 13 hours of incubation, which corresponds to the highest transcription levels of *csgD*[111].

Fragments of DNA with DNA-CsgD crosslinked interactions were immunoprecipitated with a monoclonal antibody specific for CsgD. Purified His-tagged CsgD was used for Rapid Prime-Custom Monoclonal Antibody Development (Immunoprecise Antibodies Ltd., Victoria, BC, Canada). Initially, clone 10G5 tissue supernatant was harvested; however, it did not contain enough anti-CsgD antibody for ChIP immunoprecipitation. Larger concentrations of anti-CsgD were generated by ascites in BALB/c mice (Immunoprecise). Ascites from clone 10G5 were used for immunoprecipitation in the first ChIP experiment, and ascites from clone 6D4 were used for the second ChIP experiment. Raw ascites required filtering and concentrating to remove

cells and fatty material prior to application for immunoprecipitation (Figure 3.1a-c). After filtration through a 0.2 µm filter, the ascites was passed through a Protein G affinity chromatography cartridge. In total, 10 volumes of 5ml ascites diluted 1:1 in binding buffer were run through the column. A BioRad DC Protein assay with bovine IgG standards was performed on all column fractions. Fractions equal to or greater than 0.2 mg/mL were pooled and concentrated on a 50K MWCO Amicon centrifugal filter. Anti-CsgD antibody was stored at 3 mg/mL in 50 mM Tris-HCl pH 8.0 (Table A1, Appendix A for details). CsgD-binding efficacy was confirmed by Western blot with purified CsgD-His.

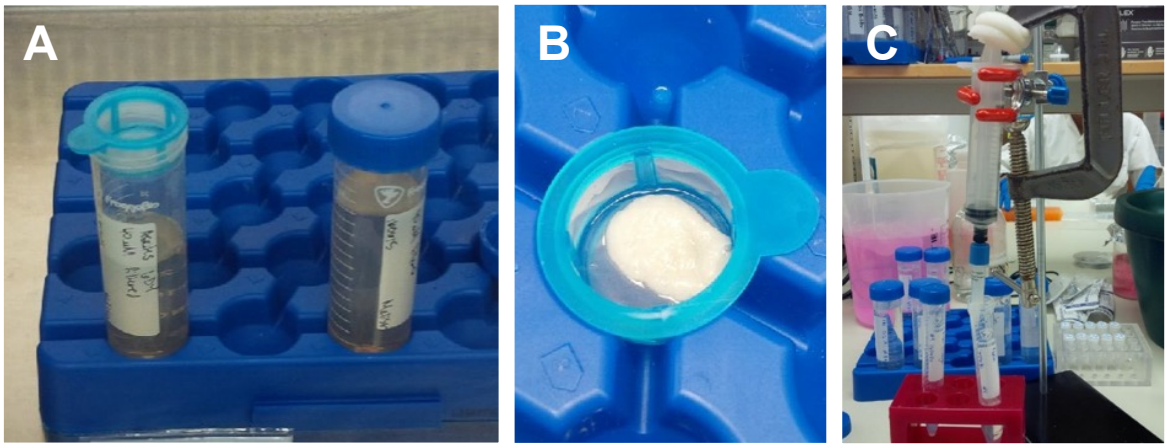


Figure 3.1a-c. Purification of monoclonal antibody clone 6D4 for ChIP from ascites. Cellular and fatty material in anti-CsgD clone 6D4 monoclonal ascites was separated by centrifugation and filtration (40 µm) (**a, b**). Antibody was purified by affinity chromatography using a Protein G cartridge (**c**) with multiple flow-throughs and elutions.

3.2.2 ChIP-seq materials and equipment

Performing ChIP-seq with *S.* Typhimurium required troubleshooting and optimization for biofilm and planktonic cell types and laboratory equipment. The different types of materials and equipment that was tested and used to perform ChIP-seq are mentioned in *3.3 Methods and Results.* A detailed list of ChIP-seq materials and equipment, which can be found in Appendix C, was required for JoVE (Journal of Visualized Experiments) manuscript preparation. Method optimization in *3.3 Methods and Results* and the materials and equipment in Appendix C are intended to be useful for researchers performing ChIP with bacterial biofilms and similar laboratory equipment.

Materials listed and described in Table C1, Appendix C are required for growing and harvesting cell types in flask culture. After harvesting the cell types from flask culture, the materials in Table C2, Appendix C are required for homogenizing cells, crosslinking proteins to DNA, and fragmenting genomic DNA. The buffers in Table C3, Appendix C are used to lyse cells after crosslinking and prior to sonication. Materials and buffer recipes for antibody selection of crosslinked CsgD-DNA fragments are listed in Tables C4 and C5, Appendix C, respectively. Immunoprecipitated DNA is purified, tested, and made into libraries for sequencing. The materials for these steps are listed and described in Table C6, Appendix C.

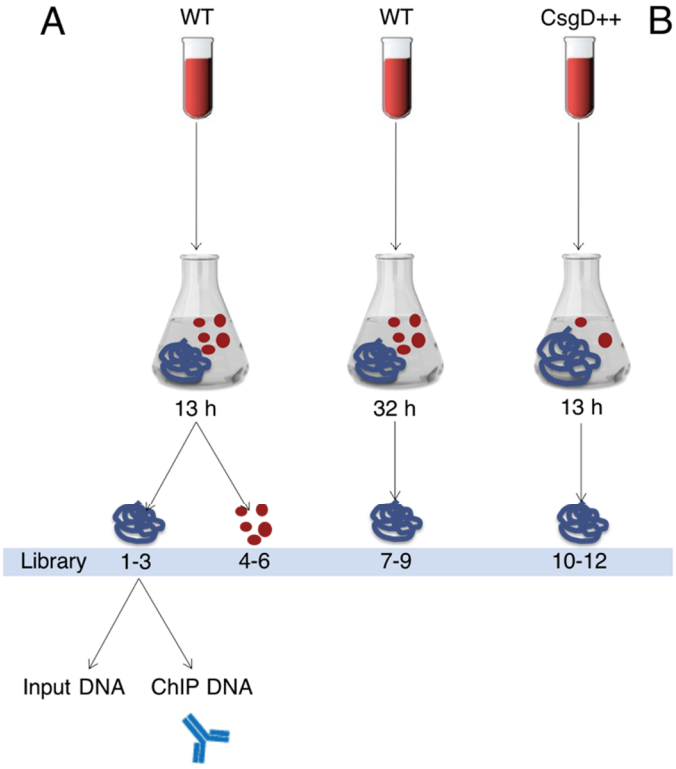
3.3 Methods and Results

3.3.1 An initial ChIP-seq experiment did not reveal regulatory peaks above background

Sequencing data from an initial ChIP-seq experiment was analyzed for CsgD-controlled regulatory peaks. This experiment was performed similarly to the ChIP-seq experiment that will be described (see Appendix B for procedure), with a few alterations. DNA samples were prepared from flask culture biofilm aggregates and planktonic cells at 13 hours, when the intracellular concentration of CsgD is highest, and from biofilm aggregates at 32 hours, when the flask culture biofilm has matured and *csgD* expression has dropped[111] (Figure 3.2). As a positive control, biofilm was harvested from the CsgD overexpresser strain, *S.* Typhimurium 14028S Δ*csgD* + *pACYC184-csgDcompFOR3* at 13 hour incubation. In this strain, it was expected that excess intracellular concentrations of CsgD will lead to occupation of all regulatory targets and maximum peak discovery sites. DNA samples were prepared from 30 mg biofilm and approximately 2.5 OD_{600} planktonic cells from flask culture. Biofilm cells were homogenized with a glass tissue homogenizer prior to formaldehyde crosslinking and lysis. Sonication proceeded with 5 sonication rounds of 30 s at 20-40% of 400 W and 2 min off on ice. A sonicated sample of input DNA was collected as background comparison, and the remainder was immunoprecipitated with anti-CsgD antibody 10G5 and Protein G magnetic beads. Input and immunoprecipitated DNA had the crosslinks reversed by incubation at 65°C, RNA was digested with RNase A, and proteins were digested with proteinase K at 45°C, prior to column-based DNA purification. Immunoprecipitated DNA was at a very low concentration, below detection by Nanodrop (<5 ng/μL), whereas Input DNA was 300-550 ng/μL. Libraries were prepared from these samples with a ThruPLEX Library Prep Kit for Illumina sequencing, and evaluated by Bioanalyzer electropherogram showing library size and a qPCR NEBNext Library Quant Kit showing library concentration. These values were used for calculations to combine libraries in

preparation for sequencing. The library pool was sequenced by MiSeq, producing 2x 75 bp reads.

A



B

Library index	Number of paired reads	Expected outcome
1	2 454 168*	Peaks at occupied biofilm regulatory sites
2	5 464 138	
3	5 542 764	
4	2 391 796*	Fewest peaks at regulatory sites
5	1 871 256	
6	1 677 440	
7	1 301 622	Fewer peaks at biofilm regulatory sites than 13 h biofilm
8	1 828 886	
9	1 802 404	
10	13 844 996	Maximum occupation of CsgD regulatory sites
11	1 983 666	
12	3 060 556	

Figure 3.2. ChIP-seq samples and sequencing reads from an initial experiment. ChIP DNA was prepared in triplicate from flask culture incubated at 28°C with shaking (200 rpm) of wild type (WT) *S.* Typhimurium biofilm and planktonic cells at 13 h (left), wild type (WT) *S.* Typhimurium biofilm at 32 h (middle), and CsgD overexpressor (CsgD++) *S.* Typhimurium 14028S Δ*csgD* + *pACYC184-csgDcompFOR3* at 13 h (right) (**a**). Sonicated DNA (Input DNA) and DNA from immunoprecipitation with anti-CsgD antibody (ChIP DNA) was prepared for each of the four cell preparations. Libraries were prepared from ChIP DNA and sequenced on an Illumina MiSeq at 2x 75 bp. The number of reads and the expected outcome for each cell type library in (a) are listed (**b**). Replicates labelled (*) denote read samples with an unassigned read pair.

Sequencing reads were acquired in FastQ file format and analyzed using Geneious 9.1.5 software. Reads from each dataset were trimmed with an error probability limit of 0.05, prior to alignment of forward and reverse read files. The aligned files were mapped to *S.* Typhimurium reference genome (GenBank CP001363.1)[81] and visualized by genome browser (Figure 3.3). Peaks were distinguished from background coverage within each read sample. In two of three *S.* Typhimurium biofilm ChIP-seq replicates harvested at 32 hours, a single peak was observed by visual scanning of mapped reads on the Geneious genome browser. This peak was composed primarily of unpaired reverse reads at the tail end of the *adrB* coding region, upstream of *psiF,* and 756 bp upstream of *adrA*. If this peak was significant, it is expected that it would be in the promoter region, accompanied by other ChIP-selected regulatory peaks, and it would be composed of paired forward and reverse reads. In *S.* Typhimurium 14028S Δ*csgD* + *pACYC184-csgDcompFOR3* (i.e. CsgD overexpresser) biofilm samples harvested at 13 hours, there is a single peak at the *csg* intergenic region. This can be explained as an excess of this sequence, "donated" to ChIP libraries by multiple copies of the plasmid. In all, no significant peaks were identified by visualization on the genome browser. A collaborator in bioinformatics, Dr. Yejun Wang, analyzed these ChIP-seq datasets and called peaks using statistical methods. His analysis concluded that peaks among replicates were inconsistent and therefore, no statistically significant peaks indicative of CsgD-regulated regions could be identified.

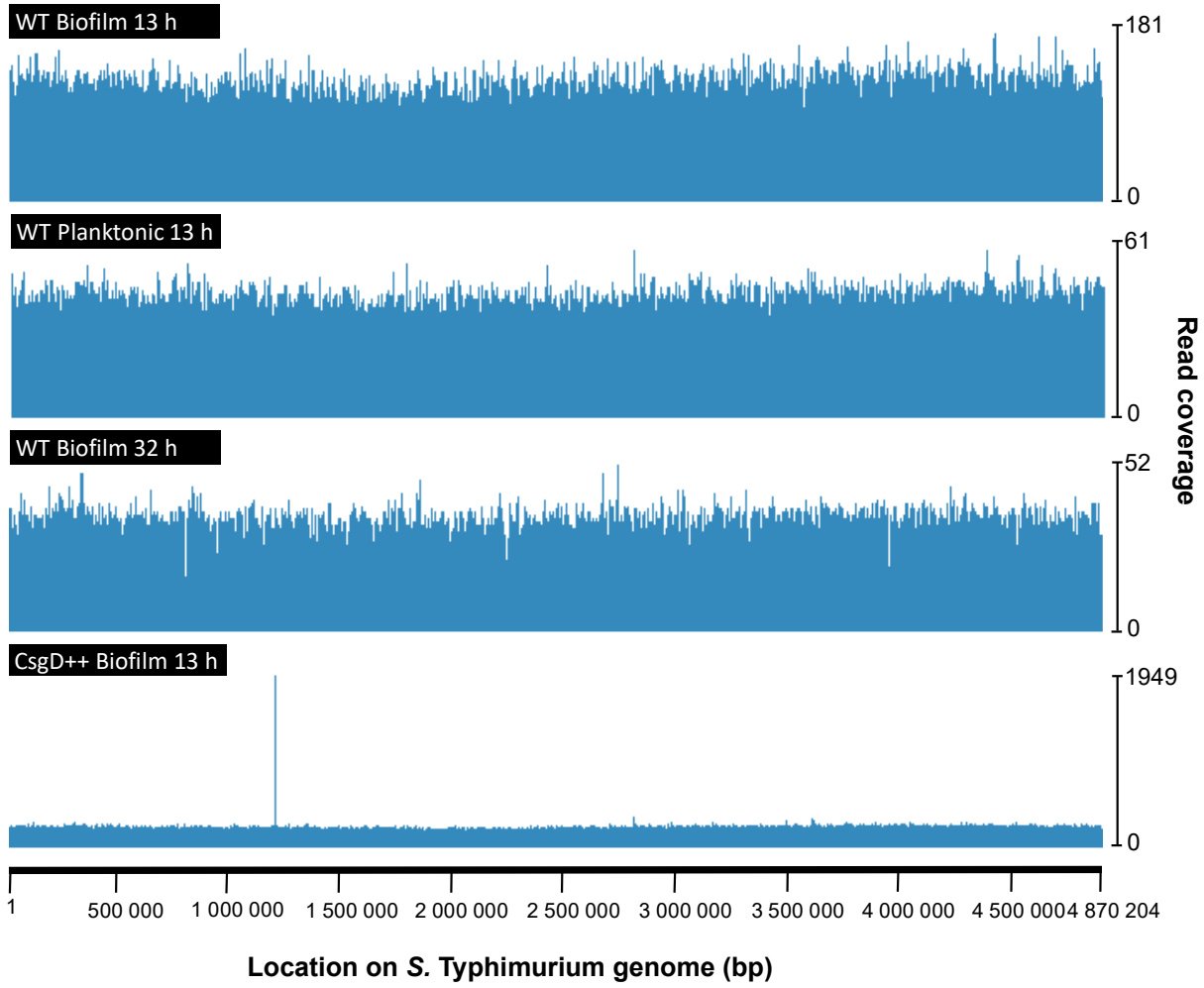


Figure 3.3. Coverage map histogram of paired-end sequencing reads from an initial ChIP-seq experiment mapped to *S.* Typhimurium reference genome and visualized by Geneious 9.1.5 genome browser. One visualized replicate was shown per cell type, strain, and time point group. Each set of paired-end sequencing reads were trimmed, forward and reverse reads were aligned, and then mapped to *S.* Typhimurium reference genome CP001363.1. Reads "stacked" at each base pair on the genome are shown as coverage in blue up to the maximum coverage in read counts at each location.

3.3.2 Optimization of ChIP-seq methods

As our initial ChIP-seq experiment for CsgD did not identify any statistically significant peaks above background, we examined the possibility that we had performed a technical error during the procedure. Technical error could have originated at nearly all procedural steps: cell harvesting, homogenization, crosslinking, sonication, antibody immunoprecipitation, washing and eluting, or DNA purification[159]. My research goal was to understand each procedural step, optimize it for laboratory working equipment and *S.* Typhimurium unique cell types, and improve consistency among ChIP-seq samples.

3.3.2.1 A comparison of bacterial ChIP-sequencing methods in the literature

Understanding experiments with successful results under similar constraints may reveal improvements that could be applied to a second ChIP-seq experiment. The majority of ChIP experiments in the literature have been performed in eukaryotes, and target abundant DNA-binding proteins, such as histones. Initially, methods and materials from ChIP experiments in *bacteria* targeting *transcription factors* were compiled. Later, all published methods and materials for ChIP-*sequencing* experiments in *bacteria* targeting *transcription factors* were compiled and assessed against the procedure followed in *3.3.1*. The results from the literature review and final protocol revisions are summarized in Table 3.1.

Table 3.1. Comparison of published ChIP methods and final methods to be used in this experiment. Published methods with bacterial cells and transcription factors were selected for their biological relevance to ChIP-seq selection of CsgD genomic targets in *Salmonella* Typhimurium.

ChIP procedural step	Literature comparison and procedural options	Final protocol
Harvest cells	Inducible expression and epitope tags are an option[a] Controls: pre-IP "input" control, parallel mock-ChIP with normal IgG, strain with gene deletion	Strains and cell types are: wild type *Salmonella* Typhimurium ST14028 biofilm and planktonic cells at 13h, CsgD overexpresser *Salmonella* Typhimurium ST14028 *ΔcsgD + pACYC csgDcompFOR3* biofilm at 13h, and CsgD mutant *Salmonella* Typhimurium ST14028 *ΔcsgD* planktonic cells at 13h
Crosslink proteins to DNA and quench crosslinking	Crosslinking for 5, 10, 20, or 30 min with 1% formaldehyde at room temperature or 37°C. Formaldehyde may be added directly to growth media. Crosslinking quenched with Glycine at final concentration of 0.1-0.5M. Cells pelleted and washed after or in lieu of quenching.	Crosslinking: 1% formaldehyde for 30min at room temperature on a rotating wheel Quenching: 0.125M Glycine for 5min at room temperature on rotating wheel. Pellet cells and wash once with PBS
Lyse cells	Lysis by osmotic shock, lysis buffer, 0.1µm zirconia glass beads, sonication, or lysis buffer with lysozyme.	In lysis buffer (SDS-based) + protease inhibitors on ice 10min. Complete lysis in IP dilution buffer + protease inhibitors on ice 0.5-2h.
Fragment DNA	DNA fragmented by sonication, bead-beating, or micrococcal nuclease.	Sonicate DNA on ice 20% 400 W 5x 30 seconds, cool 2 min between bursts.
Pre-clearing and controls	Input pre-IP sample as reference, pre-clearing with Protein G alone or Protein G and normal IgG optional. Mock IP (no antibody) or control (species-matched IgG) in parallel to IP as controls.	Pre-clearing: incubate with 50µg mouse IgG 1h at 4°C, then with Protein G magnetic beads 3h at 4°C. The precleared solution is split into *Input* (no IP), *Control* (IP with mouse IgG), and *Test* (IP with anti-CsgD mAb)
Add antibody	Monoclonal, polyclonal, commercial, or commercial antibodies against epitope-tags. Incubation with 5 or 10 µg antibody 2h at room temperature or overnight at 4°C.	Add 10 µg anti-CsgD antibody to *test* samples and 10µg normal mouse IgG to *control* samples. Incubate overnight at 4°C on a rotating wheel.
Immunoprecipitate DNA	Protein G or Protein A agarose, sepharose, or magnetic beads.	Protein G magnetic beads incubated 3 h at 4°C. Beads are separated on a magnetic stand, washed with low salt buffer, LiCl wash buffer, TE, and eluted in Elution buffer.
Reverse crosslinks and digest RNA	Rnase A treatment and concurrent crosslinking, 65°C ≥6 h or overnight. SDS may be added.	Add 2 µg Rnase A and NaCl to 0.3 M. Overnight incubation at 65°C
Digest protein	Incubation with Proteinase K or Pronase at 65°C overnight or at 45°C for ≥2h	Add 9 µg Proteinase K. Incubate 45°C for 4 hours or more.
Purify DNA	DNA purified with phenol chloroform extraction, purification columns, or gel purification.	DNA purification with magnetic beads.
QA/QC	qPCR to assess a known genomic target in comparison to Input DNA or qPCR standard curve. Agilent Bioanalyzer HS dsDNA kit to assess fragment sizes. Qubit or qPCR library quantification kit to quantify DNA.	Fluorimeter to quantify DNA. qRT-PCR to assess quantities of known targets bound by CsgD compared with Input DNA. Bioanalyzer High sensitivity DNA kit to assess library fragment size range.
Analysis	Artemis, Bowtie, or BWA for mapping to reference genome. Galaxy tools, BedTools, SAMTools for finding sequence read coverage. Peakfinder, CisGenome, Geneious, dPeak, MACS, MOSAiCS for calling peaks. YMF, BioProspector, or MEME for motif analysis.	Analysis workflow: Assess read base quality, trim reads by base quality, align sequence reads to reference genome, call peaks, associate peaks with genes, motif analysis Tools available: Geneious, Galaxy Suite

[a]references for each procedural option are in text of *3.3.2.1 A comparison of bacterial ChIP-sequencing methods in the literature*

ChIP-seq methods and materials from the surveyed published protocols were reviewed item-by-item. Firstly, cells were usually harvested from a subculture at mid-log phase. There was a consensus among methods for crosslinking cells with a final concentration of 1% formaldehyde; however, incubation times varied from 10-30 minutes. To stop formaldehyde crosslinking, glycine was added at a final concentration of 0.125-0.5 M, often without a specified incubation time or temperature [14,34,37,42,50,65,85,118,122,134,136,167]. After crosslinking, cells are usually pelleted and washed in either PBS or TBS to remove excess crosslinker and prevent overcrosslinking. The cells must be lysed prior to sonication to allow for DNA fragmentation and immunoprecipitation. The majority of surveyed protocols used a lysozyme and/or SDS-based buffer to lyse cells, with Tris-HCl or HEPES included to stabilize pH, and the addition of other components to support cell lysis, prevent protein degradation, and preserve DNA. Typically, cells were resuspended and incubated in lysis buffer for 30 minutes at 37°C [37,42,50,65,181,192]. In some cases, an IP dilution buffer was added to dilute SDS before sonication [34,86,138,167]. Cell lysis using glass beads, sonication, or mechanical disruption are alternate or parallel methods [86,118,138]. In all of the methods surveyed, DNA was fragmented to less than 1000 bp by sonication. However, sonication methods vary due to differences in sonicator equipment and the final fragment sizes required by different sequencing platforms and kits. Precipitated material and remaining whole cells were sedimented by centrifugation and separated from the cell lysate in the supernatant [65,85,181].

There are many different ways of performing immunoprecipitation (Table 3.1); different strategies and controls can influence downstream analysis. When intracellular concentration of transcription factors is low or below detection for ChIP experiments, expression may be enhanced by an inducer, such as IPTG, arabinose, or rhamnose [37,42,85,134,148,192]. Some methods

use commercial antibodies that bind to epitope-tagged proteins. These are commonly FLAG-[50,65,118,181], 3xFLAG [34,69,136], or His-tagged [14,134] proteins. After DNA-TF complexes were bound by TF-specific antibody, DNA-TF-Ab complexes were bound at the antibody Fc portion by Protein A or Protein G agarose[118,134,136,181], sepharose[34,65,69,132,192], or magnetic beads[14,37,42,85,136,167]. These beads may be pre-cleared to reduce nonspecific binding prior to immunoprecipitation[14,34,85,132,136]. Once DNA-TF-Ab complexes are bound by the beads, unbound protein, DNA, and buffer are washed away by wash buffers. When specified, wash buffers usually contain Tris-HCl and EDTA, and are described by their salt content: high salt (NaCl), low salt (NaCl), and LiCl. Elution buffers vary, but they usually contain SDS for eluting DNA from protein and EDTA for protecting DNA at an elevated pH. Transcription factors and antibody complexes were decrosslinked from DNA by heating at 65°C or higher, usually overnight [34,37,42,65,69,85,118,136,181,192]. After decrosslinking, RNA was digested with Rnase[37,42,65,167] occasionally during decrosslinking, and protein was digested with Proteinase K for more than 2 hours. Purification was performed by column purification[34,37,42,65,118,136,137,164], phenol-chloroform extraction and ethanol precipitation[34,65,69,85,167,192] or gel purification[148]. Libraries were prepared from this purified DNA using kits that are compatible with the intended sequencing platform. These libraries were checked for enrichment by qPCR with primers for housekeeping and known regulatory targets, by Bioanalyzer for library fragment size, by Qubit or by qPCR library quantification kit for library concentration. Sequencing was generally performed on the Illumina Genome Analyzer IIx[122,136,167], Illumina HiSeq 2000[69,79,85], or related platforms. ChIP-seq peaks were compared to background in several different forms. Nearly all of the surveyed methods compare ChIP-seq peaks to mapped reads from sonicated input DNA [34,37,50,65,85,122,132,134,148,164,181,192]. Other experiments compare peaks in test IP to a control "mock"

IP with species-matched antibody[14,164], background within the same read sample[136], or to immunoprecipitation with a deletion or empty vector strain[34,136,148,167,181].

3.3.2.2 Harvesting *Salmonella* cell types

Maintaining sample consistency is expected to improve the likelihood of a significant sequencing result. Planktonic and biofilm cell types are phenotypically distinct and require different handling before immunoprecipitation. Assays were performed to numerically understand flask culture biofilm and planktonic cells for the strains harvested for ChIP: wild type *S.* Typhimurium 14028, negative control *S.* Typhimurium 14028 Δ*csgD*, and CsgD overexpresser *S.* Typhimurium 14028 *ΔcsgD* + *pACYC csgDcompFOR3*. Flask culture phenotypes of these strains, grown at 28°C for 13 hours with shaking, were distinctly different. The cell types were separated by slow speed centrifugation at 1000 rpm for 2 minutes, and assessed independently. Planktonic cells were enumerated by optical density (OD_{600}) estimates and drop dilution plating. Biofilm aggregates were evaluated by wet weight (mg) and by drop dilution plating following homogenization. The wild type flask culture contained approximately $1x10^{11}$ cells split into biofilm aggregates (39%) and planktonic cells (61%) (Figure 3.5). The CsgD overexpresser flask culture contained more biofilm aggregates than wild type, suspended in media that is virtually clear (Figure 3.4a-c). The CsgD deletion flask culture contains planktonic cells, but no biofilm aggregates. This abundance of planktonic cells over the wild type and CsgD overexpresser flask culture is evident by OD_{600} in Figure 3.6d. The relationship between wet weight of biofilm and CFU (Figure 3.6a) and the relationship between volume of flask culture and wet weight of biofilm was confirmed (Figure 3.6b).

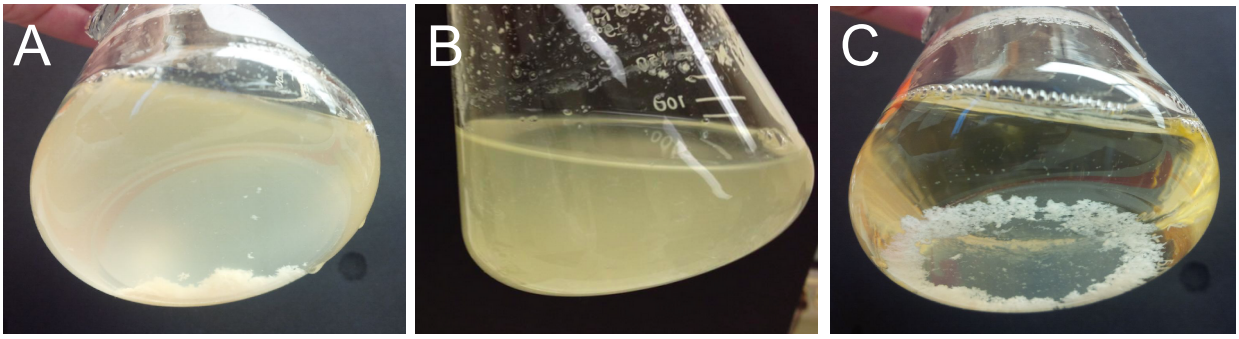


Figure 3.4a-c. Flask cultures of *S.* Typhimurium in 1% tryptone harvested for ChIP-seq. Flasks were inoculated with 1 OD_{600} log-phase culture of *S.* Typhimurium 14028 (**a**), *S.* Typhimurium 14028 Δ*csgD* (**b**), or *S.* Typhimurium 14028 *ΔcsgD + pACYC csgDcompFOR3* (**c**) and incubated for 28°C for 13 hours with shaking at 200 rpm.

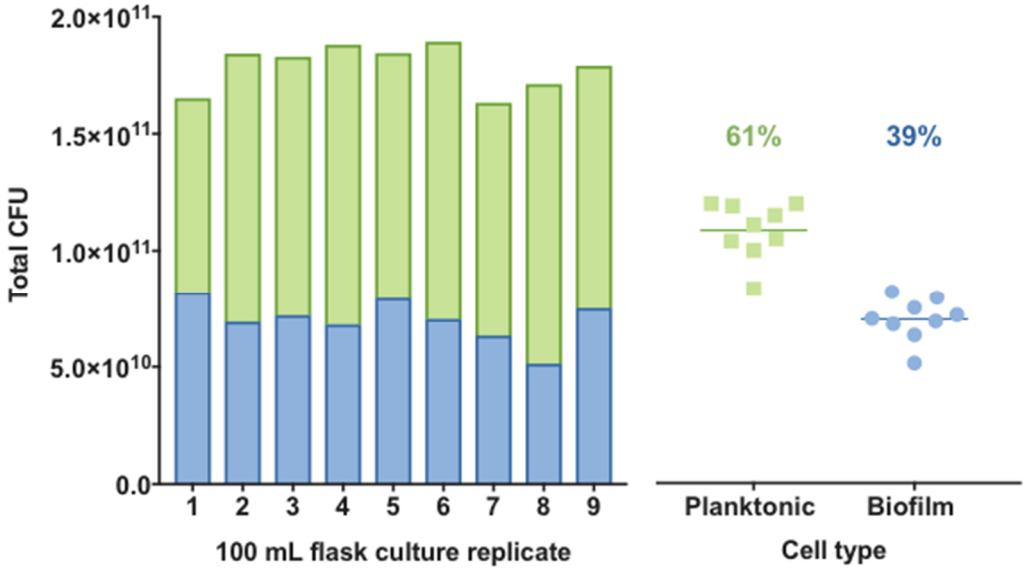


Figure 3.5. The number and proportion of planktonic and biofilm cell types in the bulk liquid phase of *S.* Typhimurium 14028 flask culture. The colony forming units (CFU) present in aggregate or planktonic cell subpopulations was calculated using conversion factors determined from serial dilution plating after homogenization (1.92 x 10^9 CFU per 1.0 OD_{600} for planktonic cells; 1.73 x 10^8 CFU/mg for aggregates). The green bars and blue bars represent the proportion of planktonic cells and aggregates comprising the total number of cells in the population; points on the right side of the graph represent total CFU values for each cell type from nine replicate flask cultures. The percentage values represent the average proportion of each cell type. (Published: MacKenzie et al. (2017) Frontiers in Veterinary Science, Figure 2d[110])

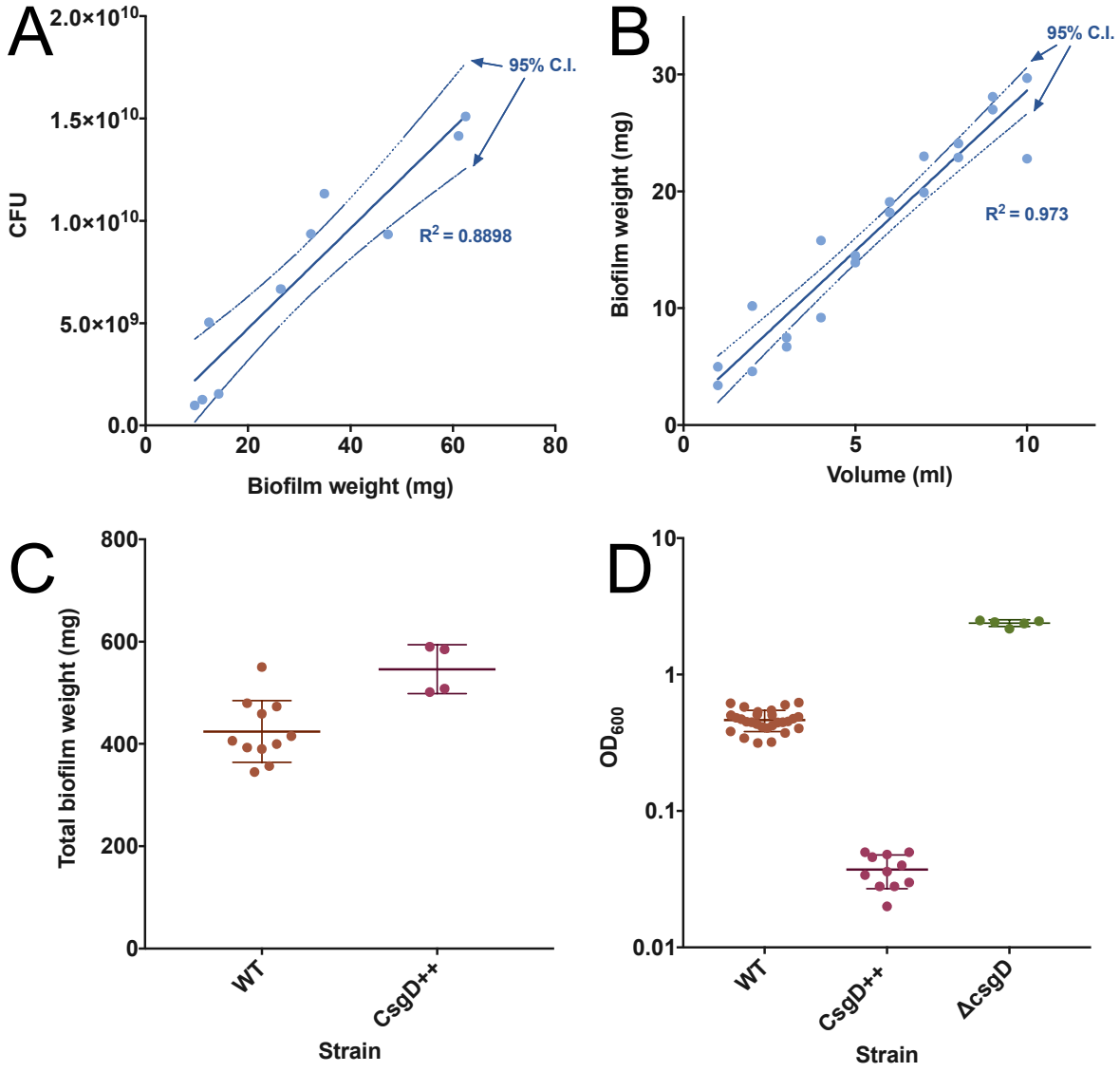


Figure 3.6a-d. Characterizing the biofilm and planktonic cell subsets in *S.* Typhimurium 14028 grown in 1% tryptone incubated at 28°C for 13 hours. Parameters were assessed for wild type *S.* Typhimurium (WT) planktonic and biofilm cells, *S.* Typhimurium 14028 *ΔcsgD* + *pACYC csgDcompFOR3* (CsgD++) biofilm cells, and *S.* Typhimurium 14028 *ΔcsgD* (ΔcsgD) planktonic cells. Biofilm and planktonic cells from the 13 hour flask culture were separated by slow speed centrifugation at 1000 rpm for 2 minutes and analyzed. The biofilm conversion factor 1.73 x 10^8 CFU/mg was confirmed by homogenizing and performing drop dilution enumeration with different weights of biofilm (**a**) and the relationship between volume of flask culture and wet weight of biofilm was confirmed (**b**). Regression analysis was performed in Prism 7, and the dotted lines represent the 95% confidence intervals of the data. Total biofilm from wild type and CsgD overexpresser flask culture was harvested and weighed (**c**); and planktonic cells from all flask cultures were measured after separation by slow speed centrifugation by OD_{600} (**d**) and CFU/ml from drop dilutions.

The recommended amount of DNA input for ChIP-seq is 10-25 µg[58]. The amount of biofilm to harvest for ChIP was calculated, using the biofilm conversion factor (BCF, 1.73 x 10^8 CFU/mg) with the following equation:

$$25\ \mu g\ DNA\ x \frac{1\ g}{1x10^{-6} \mu g} x \frac{6.022x10^{23}\ bp}{1\ mole} x \frac{1\ CFU}{Genome\ size\ (bp)} x \frac{1}{BCF} = biofilm\ (mg) \quad (3.1)$$

Therefore, 30 mg biofilm was harvested for ChIP because it yields approximately 25 µg DNA.

I hypothesized that the abundance of proteinaceous extracellular material associated with biofilm would interfere with the efficiency of crosslinking. If I were simply to normalize the different cell types by cell number, treatment of planktonic cells may result in an unequal amount of crosslinked product presented for immunoprecipitation. A protein assay was performed to determine the equivalent amount of planktonic cell material to match 30 mg of biofilm (Figure 3.7). The results from this assay indicated that approximately 6.0 OD_{600} planktonic cells harvested for ChIP will have equivalent crosslinking capacity as 30 mg biofilm. Therefore, 30 mg biofilm and 6.0 OD_{600} planktonic cells were harvested from flask culture for ChIP.

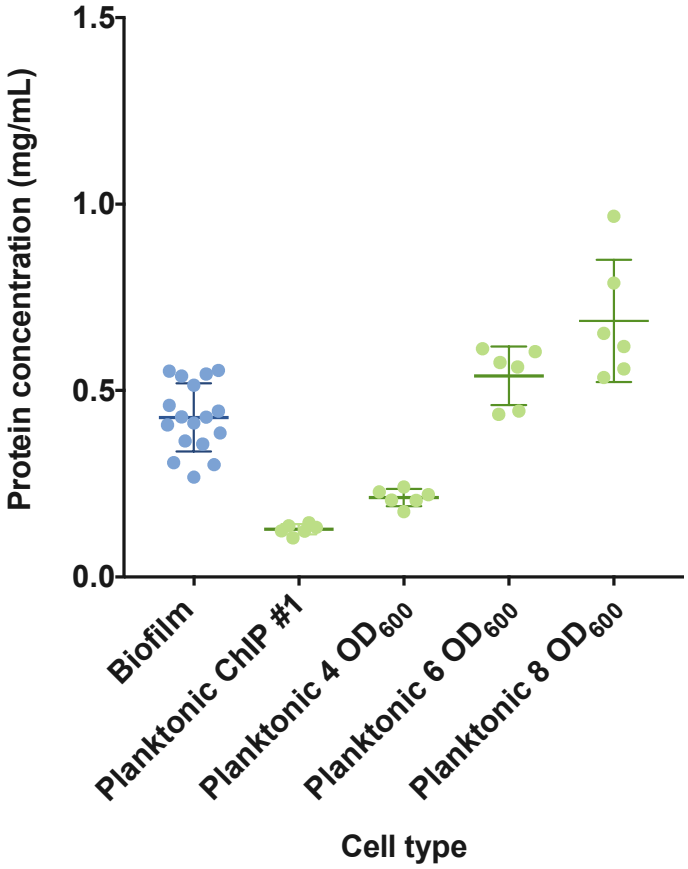


Figure 3.7. Total protein concentrations for planktonic cells and biofilm cell samples harvested from *S.* Typhimurium flask culture. BioRad DC Protein assays were performed and protein amounts relative to a BSA standard curve were measured at 750 nm. Protein content of lysed planktonic cell samples measured by OD_{600} were compared to lysed 30 mg biofilm aggregates.

3.3.2.3 Homogenizing cells

Biofilm aggregates must first be broken apart to allow equal access of formaldehyde crosslinker to all cells. Previously, I had used a glass tissue homogenizer to break apart biofilms. This method produced aerosols and was inconsistent between samples. Homogenization by mixer mill keeps tubes sealed and is expected to break apart biofilm consistently between all samples. I tested homogenization by mixer mill to determine the parameters that would break apart biofilm adequately for crosslinking.

Biofilm broken apart by glass tissue homogenizer, mixer mill (Retsch Mixer Mill MM 400) with a glass bead, and mixer mill with a metal bead was observed. The glass tissue homogenizer accumulates biofilm, tends to decrease in homogenization quality with higher volumes processed, and produces cell numbers with higher variance. Biofilm flakes are apparent in tubes homogenized by mixer mill with a glass bead, but not with a metal bead (Figure 3.8). Identical wild type biofilm samples were homogenized at different times to determine the minimum duration of homogenization at 30 Hz that would adequately break apart biofilm. Homogenized biofilm tubes containing a 5 mm steel bead that were processed with the mixer mill 5 minutes or longer appeared homogenous; therefore, 5 minutes at 30 Hz was chosen as the standard for future ChIP sample preparation (Figure 3.9). Planktonic cells are treated the same way to reduce variables in the ChIP-seq experiment.

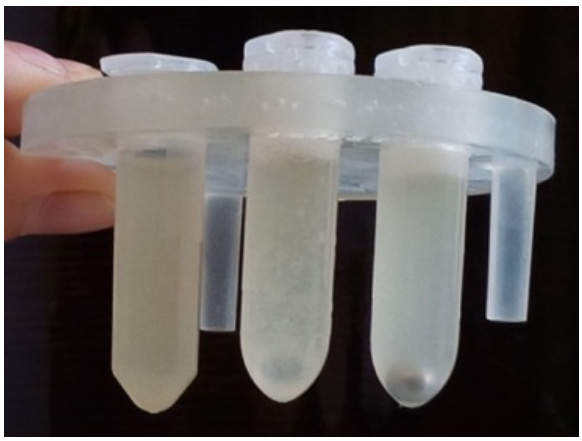


Figure 3.8. Qualitative comparison of homogenization of wild type *S.* Typhimurium 14028 biofilm by glass tissue homogenizer, mixer mill with glass bead, and mixer mill with metal bead (L to R). Thirty milligrams of biofilm were harvested from flask culture and resuspended in conditioned tryptone before homogenizing by mixer mill for 5 minutes at 30 Hz.

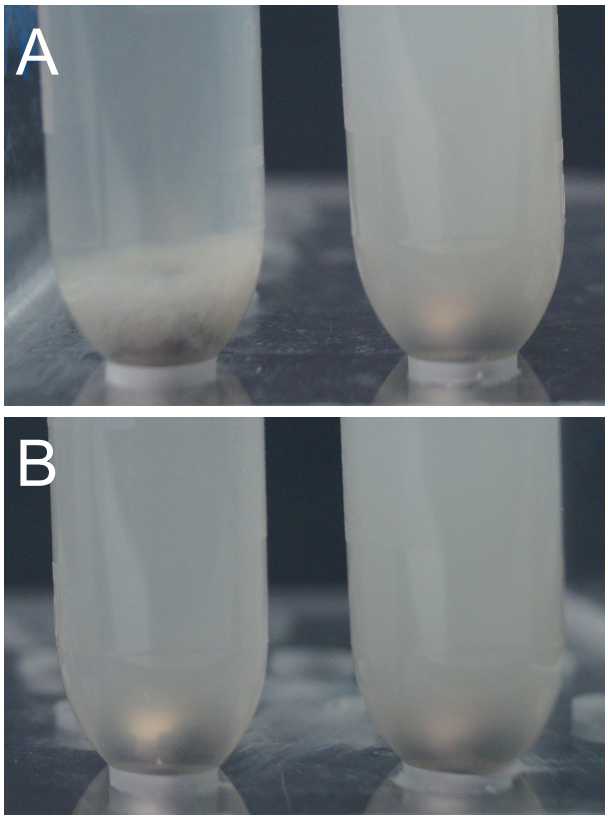


Figure 3.9. Homogenization of biofilm and planktonic cells from flask culture. Biofilm aggregates from flask culture resuspended in PBS (**a**, **left**) were homogenized using a mixer mill at 30 Hz for 5 min (**a**, **right**). Planktonic cells from flask culture resuspended in PBS (**b**) were processed by mixer mill at 30 Hz for 5 min for consistency between cell type samples.

3.3.2.4 Crosslinking proteins to DNA

Chemical crosslinking with formaldehyde is often used to fix protein-DNA interactions for chromatin immunoprecipitation[158,169]. ChIP experiments may use other chemicals or UV to immobilize protein-DNA interactions, or may forgo crosslinking entirely, as is the case with native ChIP[126,168]. Formaldehyde crosslinking was chosen for this ChIP experiment due to its availablility, ease of use, and success in published experiments. Formaldehyde is a small molecule that permeates cell membranes to crosslink reactive molecules that are in close proximity, and catalyzes the covalent linkage of amino and imino groups[158]. Crosslinking conditions are similar among published bacterial ChIP-seq methods (Table 3.1). The majority of ChIP-seq methods describe the addition of fresh formaldehyde to a final concentration of 1% in cells or growth media, incubation up to 30 minutes with gentle agitation, and the addition of glycine to stop crosslinking. During crosslinking, it is important to consider fixation time, methanol content of the formaldehyde, and how fresh the formaldehyde is; increased fixation time and methanol content can cause overcrosslinking, and old reagents can have decreased effective formaldehyde concentrations[48].

Homogenized *S.* Typhimurium cell samples were treated with 1% fresh formaldehyde and incubated with rotation for 30 minutes at room temperature to crosslink proteins to DNA. As previously described, biofilm and planktonic cell samples were harvested by protein concentration to ensure each contained equivalent crosslinking substrate. Conditioned tryptone, which contained amino acids, metabolic waste products, and other proteins, was replaced with PBSA to ensure that cellular interactions constitute a higher proportion of the total crosslinking substrate. Glycine was added to a final concentration of 0.125 M and incubated for 5 minutes with rotation at room temperature to quench crosslinking.

3.3.2.5 Lysing cells to release cell contents

After crosslinking, cell contents were released by lysis. This allows for chromatin to be sheared and for downstream immunoprecipitation with cell protein and DNA. In this ChIP-seq experiment, lysis was achieved through resuspension in an SDS-based buffer (Table A4, Appendix C), which was incubated on ice for 30 minutes. Cell lysates were then combined with 1.4 mL IP dilution buffer, which reduces the concentration of SDS before sonication. At this point, there were often visible, non-lysed particles in solution, particularly in biofilm cell samples. After correspondence with collaborators (Dr. Carsten Kröger, Trinity College Dublin, Stefani Kary, MSc.,Trinity College Dublin), the duration that samples were incubated on ice in IP dilution buffer was extended up to 4 hours, with intermittent and gentle vortexing.

3.3.2.6 Fragmenting genomic DNA by sonication

DNA must be fragmented for enrichment of fragments with crosslinked transcription factor during immunoprecipitation and for downstream library preparation and sequencing[133]. Crosslinked chromatin from lysed cells are fragmented to a size range compatible with sequencing: less than 1000 bp, with a majority 200-600 bp. Genomic DNA may be fragmented for next generation sequencing by physical, enzymatic, and chemical methods[93]. ChIP-seq experiments often use enzymatic fragmentation with micrococcal nuclease or physical fragmentation by sonication.

In this ChIP-seq experiment, *S.* Typhimurium genomic DNA from lysed cell samples was fragmented by sonication. DNA degradation by sonication occurs through cavitation, thermal, or mechanical effect by breaking hydrogen bonds[45]. Although there can be some bias, sonication fragments DNA nonspecifically. Energy transfer will vary for different sonicator units and cell samples, so a sonication assay is recommended to determine the number of sonication bursts

required to fragment DNA to less than 1 kb. Cell lysis samples that were prepared identically were fragmented for different numbers of sonication rounds at 20-40% of 400W at 30s on with a 2-minute rest on ice. Note that sonication breaks apart material and makes the solution appear clearer than before sonication (Figure 3.10). To resolve DNA fragments on an agarose gel, this sample was decrosslinked at high heat and treated with RNase A at 65 °C for 2 hours and Proteinase K at 45 °C for 3 hours. The fragments were then run on a 2% agarose gel, and show a decrease in fragment size with increasing number of sonication bursts (Figure 3.11). This assay was performed on lysed cell samples with a probe sonicator (Ultrasonic Liquid Processor, Sonics & Materials, Inc., VC300) and 3 mm probe (Sonics & Materials, Inc., 630-0418) and a cup horn sonicator (Qsonica, 431C2). The cup horn sonicator energy transfer was weaker, and inadequately fragmented DNA at up to 30 rounds of sonication (i.e. 75 minutes) (data not shown). For *S.* Typhimurium ChIP samples, 5 sonication rounds were adequate to fragment DNA to less than 1000 base pairs.

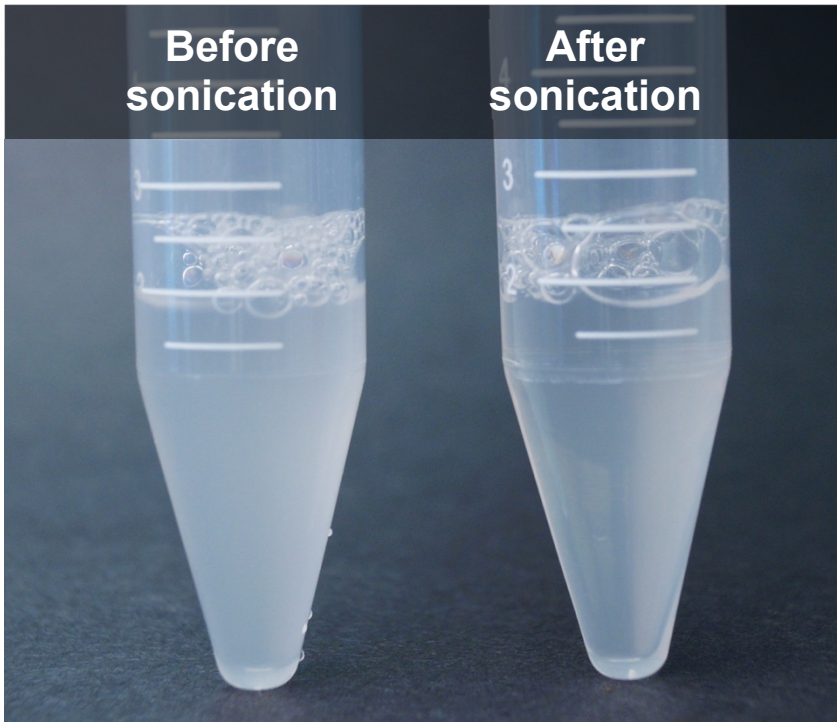


Figure 3.10. Sonication changes the opacity of lysed cell preparations. Lysed cell preparations (L) were sonicated at 20% of 400 W for 5 rounds at 30 seconds on and 2 minutes rest on ice (R).

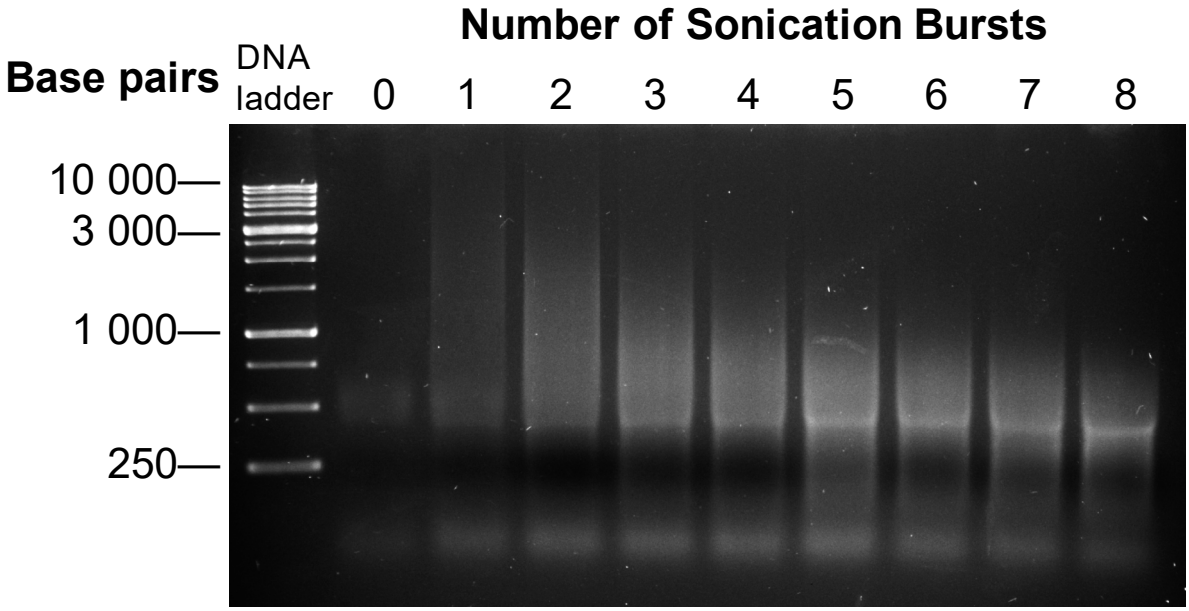


Figure 3.11. Sonication assay with pre-ChIP cell lysate. Cell samples were separated, crosslinked, and lysed. DNA-protein complexes were sonicated up to 8 times at 30 s on and 2 min off on ice to break DNA into smaller fragments. A portion of the sonicated lysate was separated on a 2% agarose gel and the number of sonication bursts to be used in future ChIP cell preparations was chosen from this assay.

3.3.2.7 Immunoprecipitation of CsgD-DNA interactions

Immunoprecipitation is the most important step in ChIP-seq. During immunoprecipitation, target transcription factors are bound by specific antibody in crosslinked DNA-TF interactions. The resulting DNA-TF-Ab complex is bound at the Fc portion of the antibody by Protein G beads, which are washed to remove nonspecifically binding DNA, protein, and other cell components. The complexes are dissociated from the Protein G beads by incubation with elution buffer, after which the TF-Ab interaction was decrosslinked from DNA at elevated heat, and contaminated RNA and protein was digested.

Antibody specificity is a common issue with ChIP; antibodies generated against purified protein may not have available epitopes to bind *in vivo*[113]. Following our first unsuccessful ChIP experiment, a second antibody clone was purified and tested for CsgD-binding efficacy. ENCODE guidelines recommend a confirmation of antibody target-binding activity with an immunoblot[98]. After antibody purification, CsgD binding activity was confirmed by western blot against purified CsgD-His protein (data not shown) and against whole cell lysates (Figure 3.12a). Despite this positive result, we wanted to assess antibody binding to the native conformation of CsgD. Native PAGE is a non-denaturing protein separation method that maintains normal folded protein "topography" that would be recognized by antibodies *in vivo*. *S.* Typhimurium CsgD has an isoelectric point (pI) of 9.13 from amino acid sequence, and is therefore expected to migrate towards the positive electrode. I attempted Native PAGE, but could not detect CsgD by this method. Lysates prepared for ChIP are most relevant for immunoprecipitation, so a dot blot was performed with purified antibody against lysates that had been crosslinked and sonicated (Figure 3.12b). The intensity of fluorescence corresponded with the amount of CsgD present, in descending order: CsgD overexpresser biofilm, wild type biofilm, wild type planktonic, and background fluorescence of *ΔcsgD* planktonic sonicated cell samples (Figure 3.12c). Compared

with the purified 10G5 antibody used in the initial ChIP-seq experiment described in *3.3.1*, the purified 6D4 antibody at the same concentration performed much better.

In the method we used to perform immunoprecipitation, DNA that binds nonspecifically to normal mouse IgG is removed from lysed and sonicated *S.* Typhimurium cell samples. The remaining DNA is immunoprecipitated with normal mouse IgG (i.e., Control IP) or the transcription-factor specific monoclonal antibody (i.e., Test IP). Preclearing was performed by incubating sonicated DNA with normal mouse IgG and protein G magnetic beads, and recovering the supernatant for immunoprecipitation. The supernatant was brought to 3 mL with IP dilution buffer and dispensed into two aliquots of 1.35 mL and one aliquot of 200 µL, stored at -80 °C as Input DNA. The Test IP aliquot was incubated with 10 µg of anti-CsgD monoclonal antibody and the Control IP aliquot was incubated with 10 µg of nonspecific normal mouse IgG overnight at 4°C on a rotating wheel. This admixture was then incubated with 50 µL Protein G magnetic beads at 4°C for 3 hours on a rotating wheel. Protein G magnetic beads and associated material were bound to the side of the tubes with a magnetic stand, and washed with cold salt buffer (IP wash buffer 1) twice, LiCl buffer (IP wash buffer 2) once, and TE twice (Table A6, Appendix C). Warmed Elution buffer was added to the washed beads, incubated at 65°C for 30 minutes with occasional, gentle vortexing, eluted on the magnetic stand, and repeated.

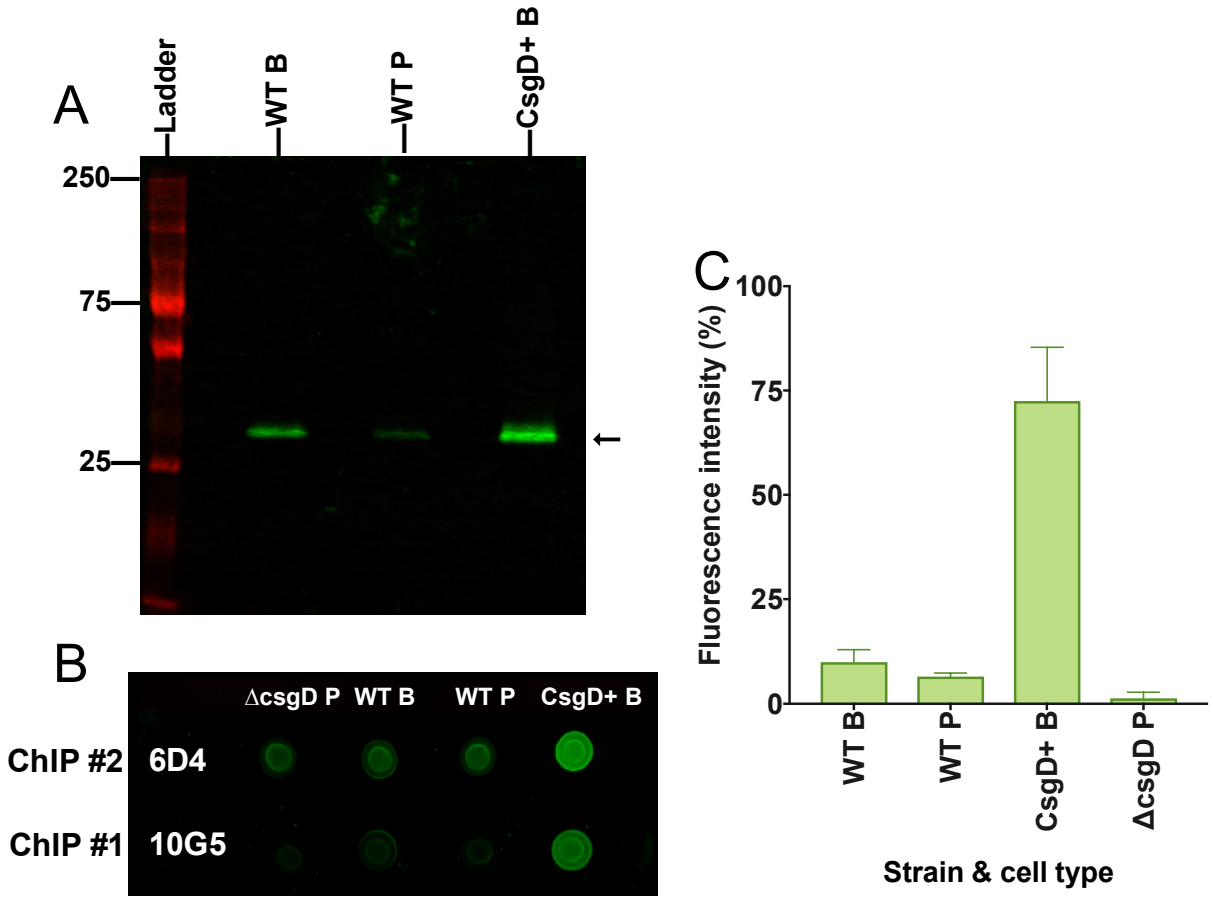


Figure 3.12a-c. Target-binding activity of ChIP monoclonal antibody by immunoblot on cell lysates. Cell lysates were prepared from flask culture inoculated with *S.* Typhimurium strains and incubated at 28°C for 13 hours with shaking. Antibody specificity was assessed by Western blot with cell lysates from flask cultures of *S.* Typhimurium biofilm (WT B) and planktonic (WT P) cells and transcription factor overexpresser (CsgD+) *S.* Typhimurium ST14028 *ΔcsgD* + *pACYC csgDcompFOR3* biofilm (**a**). Transcription factor binding activity of monoclonal antibody clones 10G5 (previous ChIP-seq) and 6D4 was assessed by fluorescent dot blot on equal amounts of sonicated preparations from *S.* Typhimurium biofilm (WT B) and planktonic (WT P) cells, CsgD overexpresser *S.* Typhimurium ST14028 *ΔcsgD* + *pACYC csgDcompFOR3* biofilm (CsgD+ B), and CsgD deletion strain *S.* Typhimurium ST14028 *ΔcsgD* planktonic (ΔcsgD P) cells from flask culture (**b**). Fluorescence intensity was read on the LiCor imager and normalized to background and to *S.* Typhimurium ST14028 *ΔcsgD* (**c**).

3.3.2.8 Purification of immunoprecipitated DNA

The small amounts of DNA from chromatin immunoprecipitation must be purified for library preparation and sequencing. The ideal purification method retains a high proportion of ChIP DNA and removes contaminants that could interfere with library preparation. I performed a dilution series of known starting Input DNA concentrations and measured the concentration of each sample before and after purification with Qiagen spin columns (QIAQuick PCR Purification Kit, Qiagen, 28104) or Axygen magnetic beads (Macherey-Nagel, 744970). Input DNA prepared for the standard curve was readily available, could be diluted to a known concentration, could be measured by Nanodrop, and approximates immunoprecipitated DNA fragment size and composition. The concentration of each standard was measured by Nanodrop. Each standard was dispensed into three replicate tubes per purification method. These three replicates were purified by column or magnetic beads according to the recommended protocol, with elution in nuclease-free water at 65°C. Figure 3.13 shows that at low concentrations of DNA, magnetic bead purification measures closer to the expected concentration from the standard curve than column purification. Some loss of DNA is apparent for both purification methods, as expected. Magnetic bead purification was chosen over column-based purification because it retained low concentrations of input ChIP DNA, removed contaminants, and measured true to standard concentrations (Figure 3.13). ChIP control and test DNA prepared from *S.* Typhimurium cells was purified using magnetic beads and DNA concentration was measured using the Qubit (Thermofisher Scientific, #Q33216) and a high sensitivity double-stranded DNA kit (Thermofisher Scientific, #Q32850).

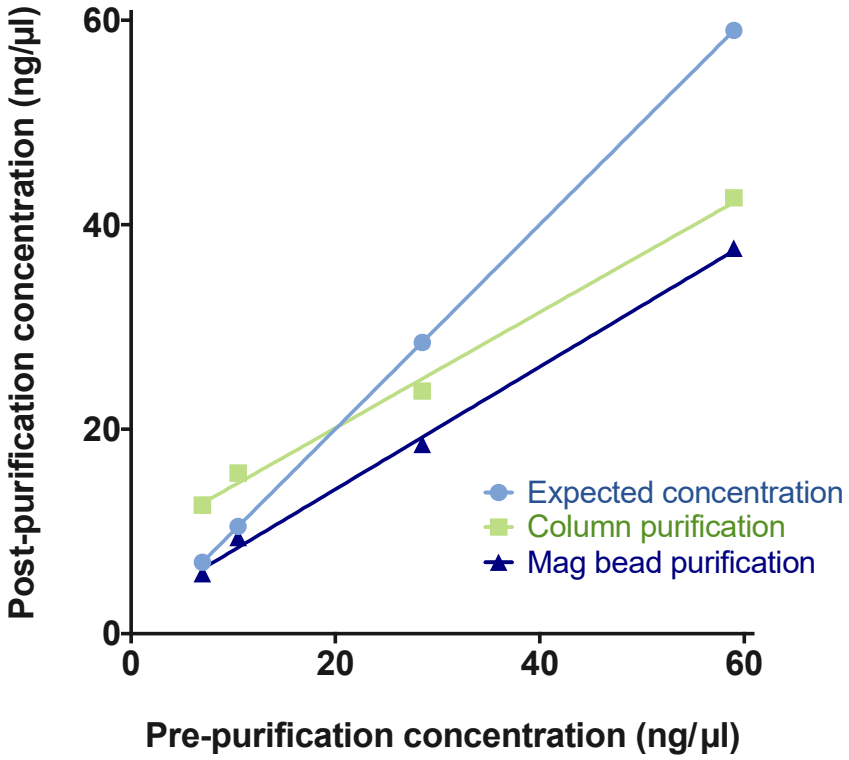


Figure 3.13. Comparison of purification strategies for small amounts of DNA resulting from ChIP-seq experiments. DNA purification of low concentrations of sonicated Input DNA measures closest to the standard curve ("Expected concentration") than column purification.

3.3.3 Confirmatory tests

Tests were performed on prepared ChIP DNA prior to library preparation to determine DNA concentration, fragment size profile, and enrichment of known regulatory targets of CsgD. Input DNA was high enough for detection by Nanodrop; however, ChIP sample DNA was below accurate detection (<5 ng/µL). Therefore, DNA concentration was measured with a Qubit double-stranded DNA High Sensitivity Assay Kit (Thermo Fisher Scientific, #Q32851) measured on the Qubit 3.0 Fluorometer (Thermo Fisher Scientific, #Q33216). Input concentrations were 15 to 35 ng/µL and test and control ChIP DNA concentrations ranged from about 10 to 270 pg/µL. After library preparation, DNA concentration was measured on the Qubit 3.0 Fluorometer with the Qubit double-stranded DNA Broad Range Assay Kit (Thermo Fisher Scientific, #Q32850)

The fragment size range of ChIP DNA, used to estimate library preparation size selection, was confirmed by Bioanalyzer (Agilent, #G2939BA) of Input DNA (Figure 3.14a). Immunoprecipitated DNA fragment sizes were below detection of the Bioanalyzer (Figure 3.14b), but amplified adapter-ligated DNA after library preparation could be detected (Figure 3.14c). The Bioanalyzer electropherogram can also reveal excess adapter dimers at approximately 120 bp that require additional clean-up (Figure 3.14d).

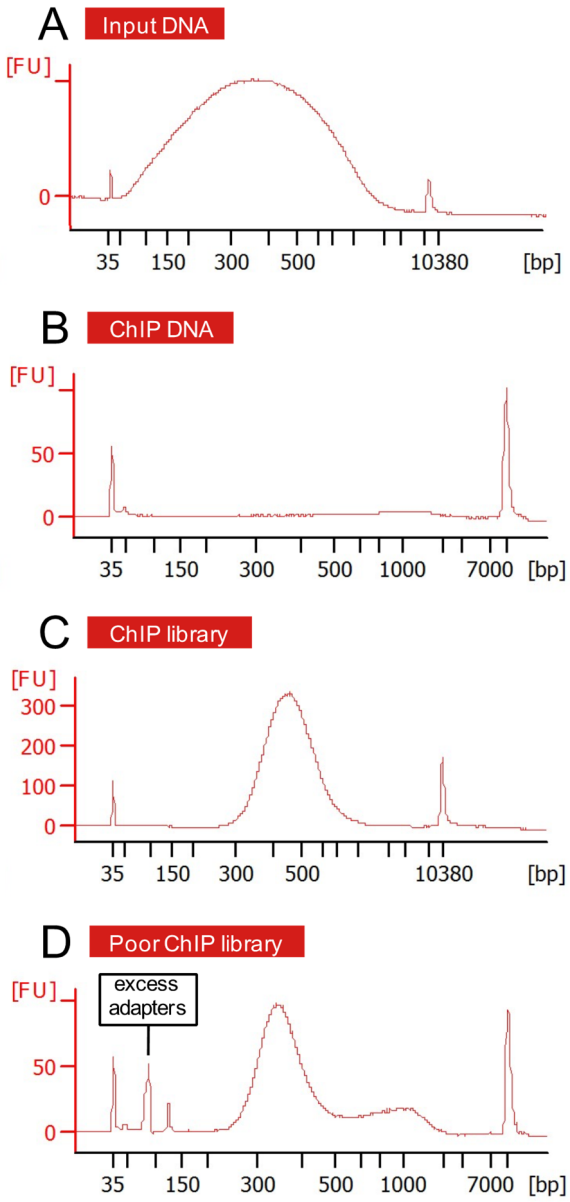


Figure 3.14a-d. ChIP DNA preparations and libraries visualized chronologically by Bioanalyzer electropherogram. Fragment sizes of genomic DNA after cell lysis and sonication were below 1000 bp, with a majority of fragments at 200-500 bp (**a**). Starting material for library preparation was below detection (**b**); however, adapter ligation and PCR enrichment allowed for visualization of libraries (**c**). A library preparation unacceptable for sequencing has low DNA peaks, oddly shaped or high molecular weight peaks, or adapter peaks at approximately 100 bp (**d**).

Enrichment of known regulatory binding regions can be assessed by qPCR, by comparing Cts between known target and reference gene sequence abundance in immunoprecipitated test and control DNA samples. These were compared to a standard curve of sonicated Input DNA. The abundance of an intergenic portion of *csgB*, a known regulatory target of CsgD, was compared to *fabG* and *groEL*, genes which are constitutively expressed and whose products are not expected to be involved in biofilm formation. We expect enrichment of *csgB* over *groEL* in the CsgD overexpresser and wild type biofilm ChIP samples and negative enrichment in the *csgD* deletion planktonic samples. Fold enrichment, calculated from these qPCR experiments, is inconsistent among replicates and does not align with expected results. Target and reference gene abundance was inconsistent in multiple replicate experiments.

Table 3.2. Fold Change ($2^{-\Delta\Delta Ct}$) from a qPCR experiment assessing enrichment of a known CsgD regulatory target, *csgB*, over a control gene, *groEL*. The average abundance of these two targets were compared between libraries prepared from ChIP DNA and a standard curve of Input DNA. Cts were recorded in duplicate and compared using the method described by Livak *et al*.[109]

ChIP Sample	Library	ΔCt Libraries	ΔCt Standard curve	ΔΔCt	Fold enrichment ($2^{-\Delta\Delta Ct}$)	Standard deviation
***S.* Typhimurium Planktonic**	1	-0.53	0.04	-0.57	1.4820	0.1152
	2	2.90	0.04	2.86	0.1378	0.2636
	3	-0.40	0.04	-0.44	1.3574	0.1826
***S.* Typhimurium Biofilm**	4	-1.22	0.04	-1.26	2.3867	0.1280
	5	-0.71	0.04	-0.74	1.6732	0.2635
	6	-0.36	0.04	-0.40	1.3174	0.1537
***S.* Typhimurium CsgD++ Biofilm**	7	0.08	0.04	0.04	0.9724	0.1050
	8	-0.35	0.04	-0.38	1.3043	0.2020
	9	-1.43	0.04	-1.46	2.7572	0.2429
***S.* Typhimurium Δ*csgD* Planktonic**	10	-0.44	0.04	-0.48	1.3934	0.0979
	11	-0.37	0.04	-0.40	1.3217	0.1146
	12	-0.60	0.04	-0.64	1.5559	0.1371

3.3.4 Library preparation

ChIP test and ChIP control DNA samples were converted into high quality libraries for Illumina sequencing. In this library preparation scheme, DNA fragments undergo end repair, 5' phosphorylation, and dA-tailing in preparation for adapter ligation and adapter loop excision (Figure 3.15). Library fragment sizes were selected by magnetic beads, after which adapter-ligated library fragments were amplified with added universal and unique index sequences during PCR enrichment. Libraries were prepared with the NEBNext Ultra II DNA Library Prep Kit for Illumina (#E7645S), with adjustments for ChIP DNA. This kit was ideal because it was designed to accommodate low DNA concentrations in its broad range of starting material, from 500 pg to 1 μg. Low concentrations of ChIP DNA starting material were not visible by Bioanalyzer electropherogram until PCR enrichment of adapter-ligated fragments (Figure 3.14)

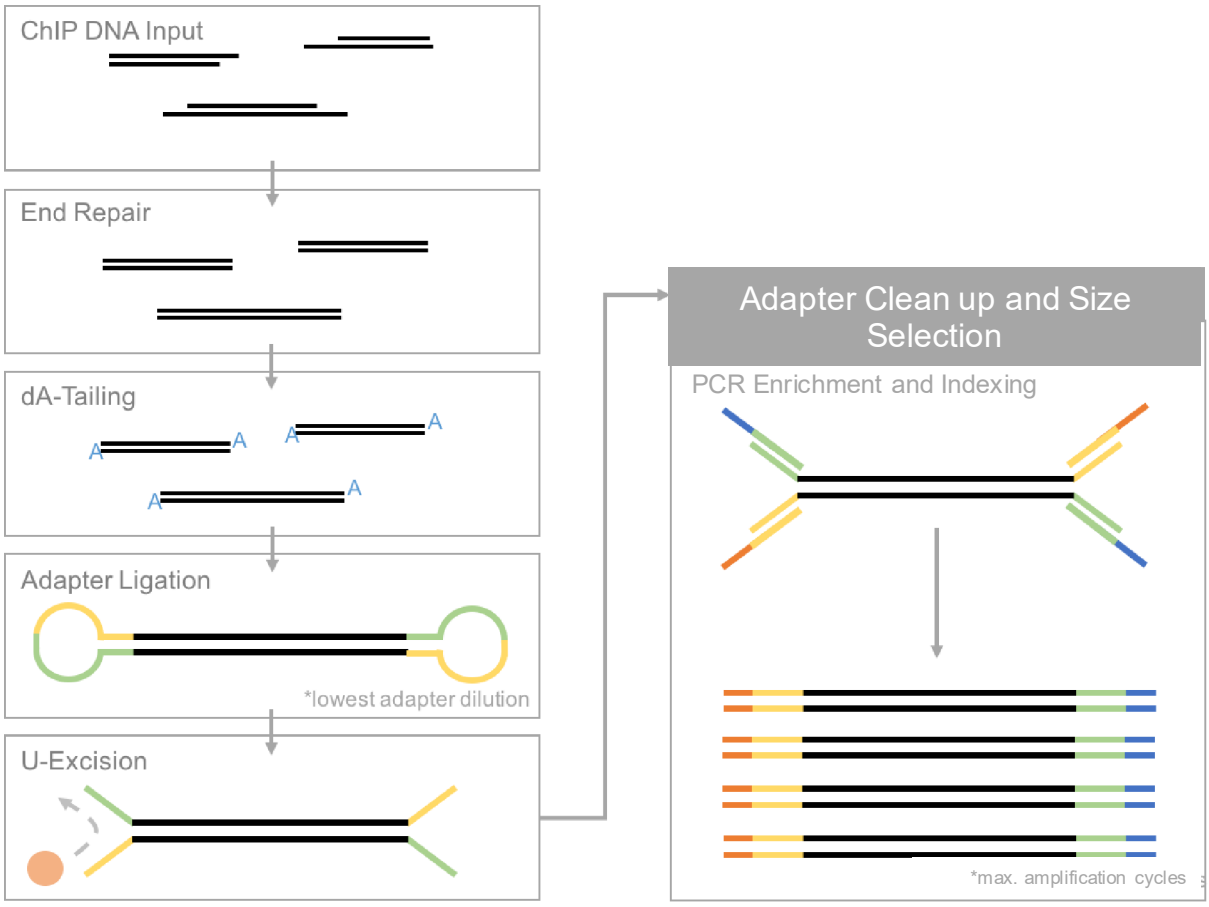


Figure 3.15. (Adapted, Figure 1 NEBNext Ultra II DNA Library Prep Kit for Illumina manual) Illustrated steps of NEBNext library preparation, with modifications for low concentration ChIP DNA starting material. Library preparation involves end repair of fragmented DNA starting material, phosphorylation and dA-Tailing, adapter ligation and U-Excision with USER enzyme, and finally, PCR enrichment and indexing with universal and unique barcode primers.

A size selection test was performed to determine magnetic bead ratios appropriate for a library size majority of 500 bp before library preparation with ChIP samples. In this assay, libraries were prepared from sonicated input DNA. These libraries were prepared in parallel and under identical conditions, apart from the ratios of magnetic beads during size selection. Magnetic bead-based size selection leverages magnetic bead affinity for high molecular weight DNA fragments. The largest DNA fragments are bound by the higher initial ratio of beads and discarded. The second ratio of beads binds to the remaining majority of higher molecular weight fragments, which are purified from the lowest molecular weight fragments. Input libraries were selected with two different ratios of magnetic beads to DNA in solution to determine the ratios to use for size selection of ChIP DNA. These libraries were compared by Bioanalyzer electropherogram (Figure 3.16). The ratios of beads to DNA that produced library fragment sizes closest to 500 bp, 0.30 and 0.15, was chosen for subsequent library preparations. This ratio produced a tight peak with a fragment size range of 300-600 bp and a median of about 450 bp.

ChIP DNA library preparation requires a couple of method-specific considerations. Low DNA concentrations after immunoprecipitation may not be visible by Bioanalyzer electropherogram, but can be visualized after library preparation due to amplification of adapter-ligated fragments (Figure 3.14b and c). Library preparation with ChIP DNA requires accurate and sensitive DNA quantification so that the right amount of starting material can be used to create a diverse library. DNA concentration is measured prior to library preparation by Qubit with a high sensitivity double stranded DNA kit, and by Qubit with a broad range double stranded DNA kit. In addition, the lowest recommended dilution of adapters was added for adapter ligation, to prevent an excess of adapters (Figure 3.14d), which interferes with cluster formation on the MiSeq flow cell. In addition, library enrichment of low yield ChIP DNA

requires the maximum number of recommended amplification cycles (i.e., 15). PCR enrichment may introduce bias; however, amplification prepares libraries for quantification, characterization, dilution, and sequencing.

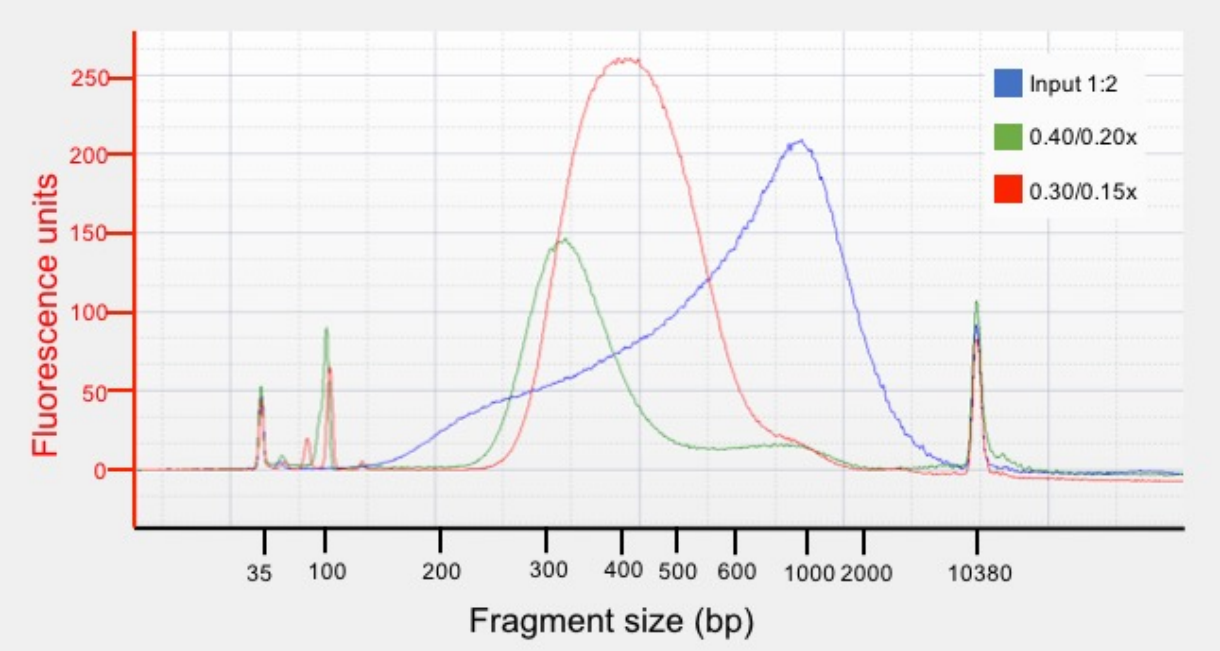


Figure 3.16. Bioanalyzer electropherogram of library magnetic bead size selection assay with input DNA. Libraries were prepared with sonicated and purified input DNA, with ratios of magnetic beads to DNA 0.40 and 0.20 (0.40/0.20) compared with 0.30 and 0.15 (0.30/0.15) in the size selection step. Magnetic bead ratios 0.30/0.15 produce libraries of fragment sizes 300-550 bp with a median of approximately 450 bp and were chosen for ChIP library preparation.

3.3.5 Sequencing ChIP-seq DNA

ChIP DNA libraries prepared from *S.* Typhimurium samples were sequenced on the Illumina MiSeq with the MiSeq Reagent Kit v3 (150-cycle, 2x75 bp, MS-102-3001). The Illumina MiSeq was chosen for its sequencing capacity, flexibility, availability, run cost per data output, ease of data acquisition, and established success in ChIP-seq and related experiments. The indexed libraries that were prepared from ChIP DNA (see *3.3.5 Library Preparation*) were multiplexed and sequenced together on the flow cell in a single lane[191]. Each unique library was tagged with a index sequence as part of the initial primers, that is detected and used to sort sequence reads by library for data acquisition and downstream analysis. In order to improve genome coverage and expected number of reads per library, two replicates were chosen from each sample type. Libraries with a single, strong Bioanalyzer peak consistent across two samples with no adapter peak were chosen. Illumina recommends around 100x coverage for ChIP-seq experiments; 8 indexed libraries are expected to yield approximately 5.6 million reads per library, which is expected for bacterial ChIP-seq experiments in the literature. Coverage was calculated using the following equation:

$$Coverage = \frac{read\ length\ (bp) x\ number\ of\ reads}{genome\ size\ (bp)} \tag{3.2}$$

i.e.,

$$\frac{2x75\ bp\ x\ 5.6\ x\ 10^6\ reads}{4.79x\ 10^6\ bp} = 175.4x$$

This is more than enough coverage. Indexed libraries were pooled to produce reads of equal proportion and adequate coverage for downstream ChIP-seq analysis. Firstly, library

concentration was converted from ng/µl to nM, and adjusted based on average fragment sizes from Bioanalyzer electropherogram (Figure 3.14c).

$$Conc.(nM) = Conc.(\frac{ng}{\mu L}) x \frac{1g}{1x10^9\, ng} x \frac{1}{Fragment\ size\ (bp)} x \frac{1mol \cdot bp}{650g}\ x \frac{1x10^9 nmol}{1mol} x\ \frac{1x10^6 \mu L}{1L} \quad (3.3)$$

i.e.,

$$38.2 \frac{ng}{\mu L} x \frac{1g}{1x10^9} x \frac{1}{450\ bp} x \frac{1mol \cdot bp}{650g}\ x \frac{1x10^9 nmol}{1mol} x\ \frac{1x10^6 \mu L}{1L} = 130.6\ nM$$

This concentration was used to find the volume required for 20 nM pooled library final concentration, or 1.67 nM per library.

$$Volume\ (\mu L) = \frac{Total\ pool\ conc.(nM)}{\#\ of\ libraries} x\ Pool\ volume\ (\mu L) x \frac{1}{Library\ conc.(nM)} \quad (3.4)$$

i.e.,

$$\frac{20\ nM}{8\ libraries} x\ 100 \mu L\ x \frac{1}{130.6\ nM\ (from\ 3.4)} = 1.91\ \mu l$$

The calculated volumes of each library were combined and the pool was brought to 1 mM Tris-HCl with 10 mM stock and nuclease free water on-site. The 20 nM pool was diluted to a final concentration of 16 pM and denatured as per kit instructions. After preparatory washes, the library pool was loaded onto the flow cell and sequenced.

3.3.6 Bioinformatic methods for analyzing ChIP-sequencing data

There are many bioinformatic tools available for processing and analyzing large amounts of heterogenous ChIP-seq data. The general steps for processing and analyzing ChIP-seq data are shown in Figure 3.17. After data is acquired from the sequencing platform, reads are scored for base quality, trimmed, aligned if paired-end, and mapped to a high quality reference genome. Trimmomatic is a common tool used for trimming reads, and BWA, Bowtie, or Bowtie2 are common tools used for aligning reads to a reference genome[100,105]. Once mapped to a reference genome, peaks are called and associated with particular genes based on proximity to the transcriptional start site of a gene[121]. At this point, peaks are visualized on a genome browser, annotated, and evaluated for consistency across biological replicates. Multiple peak sequences can be evaluated by tools such as MEME (Multiple Em for Motif Elicitation)[9] for consistency and a common binding motif. Peak data can also be compared across samples in differential analysis, or can be integrated with expression data for a deeper understanding of regulatory networks.

Geneious 9.1.5 software[88] was used primarily for analyzing this dataset because most functions required for ChIP-seq data analysis are integrated on a single platform. Quality control by base quality score at each nucleotide is automatically assigned in FastQ file format from Illumina sequencing. The remaining Illumina adapters or bases that have unacceptably low quality were trimmed with an error probability limit of 0.05. This sequencing dataset was not paired-end and as such, forward and reverse reads did not need to be aligned. Instead, trimmed reads were immediately mapped to *S.* Typhimurium 14028 reference genome (GenBank CP001363.1)[81] using the Bowtie2 plugin for Geneious. Some analysis methods recommend visualization of mapped reads on a genome browser following peak calling with a statistical tool;

however, the mapped reads were visualized first. Due to the nature of the data, it was not necessary to call peaks and perform downstream analysis after visualization.

Peaks were called with bioinformatics tools to confirm the peaks observed visually. All raw FastQ files were uploaded to the Galaxy Tools suite[2] for complete bioinformatic processing. Reads in each file were trimmed with Trimmomatic, which used a sliding window of four bases and an average threshhold quality of 20[15]. Trimmed sequence files were then mapped to the *S.* Typhimurium 14028S reference genome (GenBank CP001363.1)[81] with Bowtie2 sensitive end-to-end alignment[99]. CsgD deletion ChIP samples were expected to return no peaks, and wild type *S.* Typhimurium planktonic ChIP samples were expected to return few peaks, if any. Therefore, peak analysis was assessed based on these two ChIP sample types as background. Peak calling was performed in Galaxy with the MACS2 (Model-based Analysis of ChIP-seq) tool. Peaks were called on a trimmed and aligned "test" file, with the background set as the "control" file input. Peak calling with MACS2 proceeded with the following parameters: effective genome size, $5x10^6$ bp; band width, 450 bp; lower mfold, 2; upper mfold, 100; q-value, 0.05; with no shifting model.

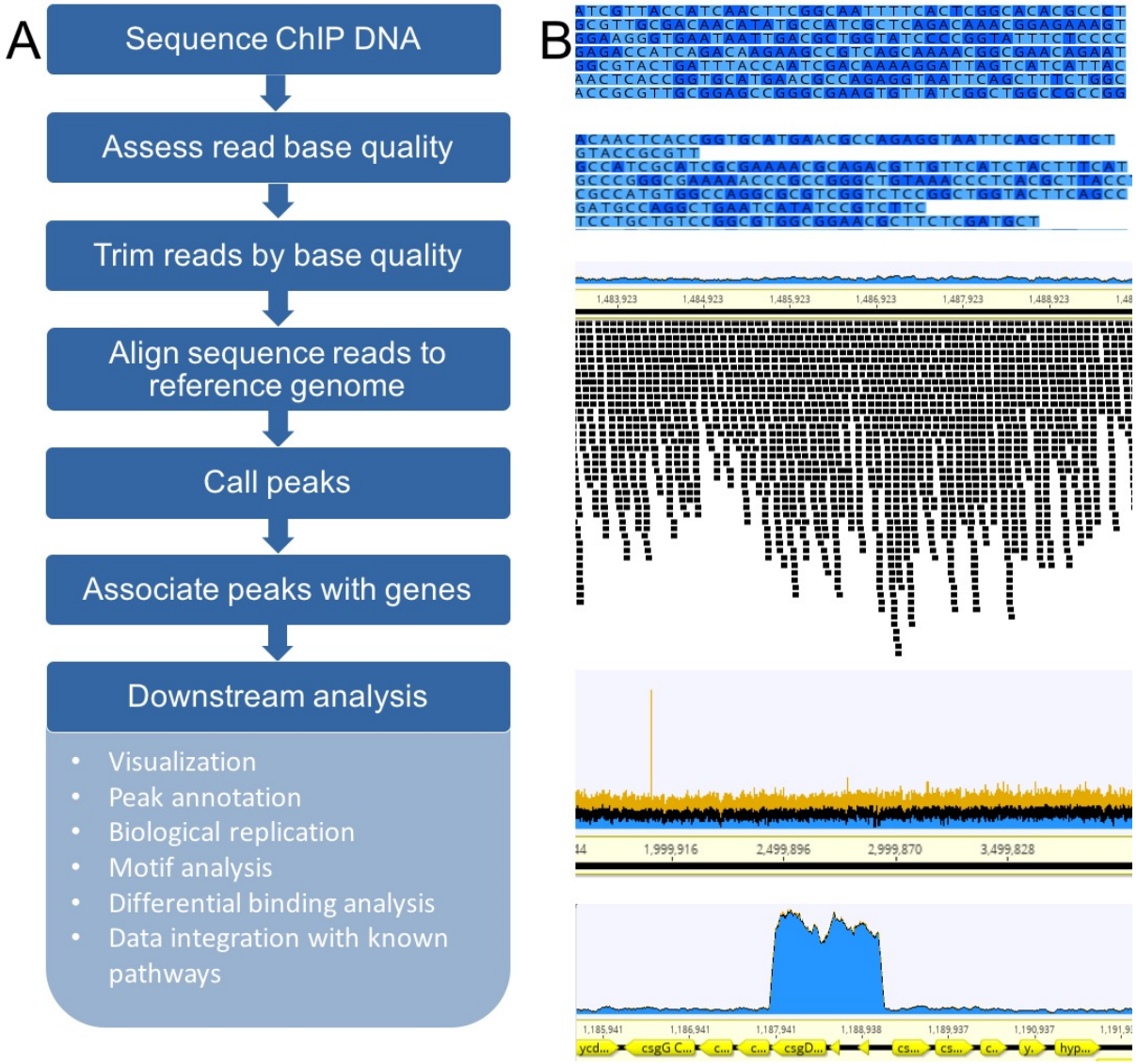


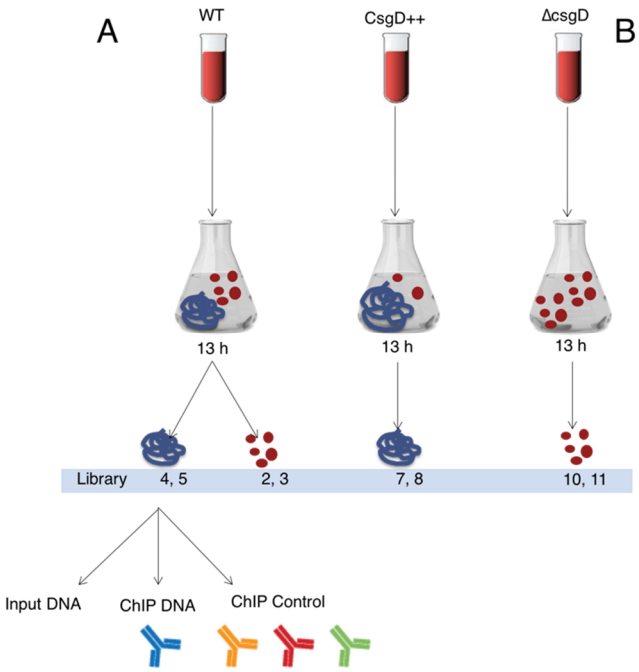
Figure 3.17a and b. Analysis of ChIP-seq data using bioinformatic methods. ChIP-seq reads are trimmed by base quality and mapped to a reference genome, where peaks can be resolved at specific genomic locations and associated with genes for biologically relevant information that can be further analyzed (**a**). Geneious 9.1.5 was used to perform most of the functions required for analysis. Figure **b** shows Geneious windows for (top to bottom) base quality in sequence reads, trimmed reads, mapped reads, visualized peaks, and peaks associated with genes.

3.3.7 ChIP-sequencing results

The sequencing run ended more than 4 hours earlier than expected with one read set sequenced, due to an error on the sample sheet that directs the MiSeq. Even so, there were approximately 2 million reads per sample (Figure 3.18), which is still adequate for a bacterial ChIP-seq experiment. The cluster density was low compared to saturated sequencing flow cells. DNA concentration by Qubit may have been overestimated, which would lead to underclustering. This may be partially why the read data was high quality (Phred quality score of Q>30), albeit half the volume than expected. Control ChIP libraries were not sequenced, after a quick check for peaks in the test ChIP sequencing reads.

Table 3.3. Metadata and summary statistics from test ChIP sequencing run. MiSeq data output for sequencing with *S.* Typhimurium test ChIP libraries.

Sequencing platform		MiSeq
Sequencing kit		MiSeq Reagent Kit v3 2x75 bp
Data output	**Cluster density**	1047 K/mm^2
	Clusters passing filter	91.4%
	Estimated yield	1820.6 Mb
	Total sequences	18 025 265



B

Library index	Index sequence read	Output sequence reads	Expected outcome
2	CGATGT	2 132 391	Fewest peaks at regulatory sites
3	TTAGGC	2 804 665	
4	TGACCA	3 376 330	Peaks at occupied biofilm regulatory sites
5	ACAGTG	2 670 410	
7	CAGATC	4 058 644	Maximum occupation of CsgD regulatory sites
8	ACTTGA	1 848 337	
10	TAGCTT	3 011 746	No regulatory peaks
11	GGCTAC	2 387 474	

Figure 3.18. ChIP-seq samples and sequencing reads. ChIP DNA was prepared in duplicate from flask culture incubated at 28°C for 13 hours with shaking (200 rpm) of wild type (WT) *S.* Typhimurium biofilm and planktonic cells (left), CsgD overexpressor (CsgD++) *S.* Typhimurium 14028S Δ*csgD* + *pACYC184-csgDcompFOR3* biofilm (middle), and CsgD deletion *S.* Typhimurium 14028S Δ*csgD* planktonic cells (right) (**a**). Sonicated DNA (Input DNA), DNA from immunoprecipitation with anti-CsgD antibody (ChIP DNA), and DNA from immunoprecipitation with normal mouse IgG (ChIP Control) was prepared for each of the cell preparations. Libraries were prepared from DNA and sequenced on an Illumina MiSeq at 2x 75 bp. The number of reads and the expected outcome for each cell type library in (a) are listed (**b**).

ChIP data was acquired from the MiSeq in FastQ file format, and processed as described in *3.3.6 Bioinformatic methods for analyzing ChIP-sequencing data*. Sequencing reads that were mapped to the *S.* Typhimurium 14028 (GenBank: CP001363.1)[81] genome can be seen in Figure 3.19. In a successful ChIP dataset, multiple, strong peaks two-fold or more above background at known and previously unknown regulatory regions should be obvious by visual scanning of mapped reads, and apparent by statistical isolation. Initially, the mapped reads were visually inspected for peaks two-fold above background reads. There were no peaks that met this criteria in ChIP samples from wild type *S.* Typhimurium biofilm and planktonic cells. As expected, there was a strong peak in both of the CsgD overexpresser biofilm replicates that spanned the *csgD* coding region and the *csg* operon intergenic region (1 188 331 to 1 189 680 bp on the genomic map), contributed by the CsgD expression plasmid. In *S.* Typhimurium *csgD* deletion replicates, there was a peak (at about 1 911 000bp on the genomic map) in the *ycgO* coding region. A statistically nonsignificant peak at *ycgO* was also seen in one wild type *S.* Typhimurium 14028 planktonic, one wild type *S.* Typhimurium 14028 biofilm, and one *S.* Typhimurium 14028 CsgD overexpresser biofilm replicate by visual scanning. Another common but statistically nonsignificant peak was seen by visual scanning (at about 1 502 500 bp on the genomic map) in a putative inner membrarne protein coding region behind *uvrA* in both replicates for wild type *S.* Typhimurium 14028 biofilm and planktonic ChIP samples. Peak calling was performed with Galaxy MACS2 callpeak to confirm the results observed visually (Table 3.3). The same peaks were called at *csgD* and the *csg* intergenic region in *S.* Typhimurium CsgD overexpresser libraries 7 (3.6-fold enrichment) and 8 (6.5-fold enrichment). New peaks at several different genomic locations were called in *S.* Typhimurium planktonic library 2 at approximately 2-fold enrichment.

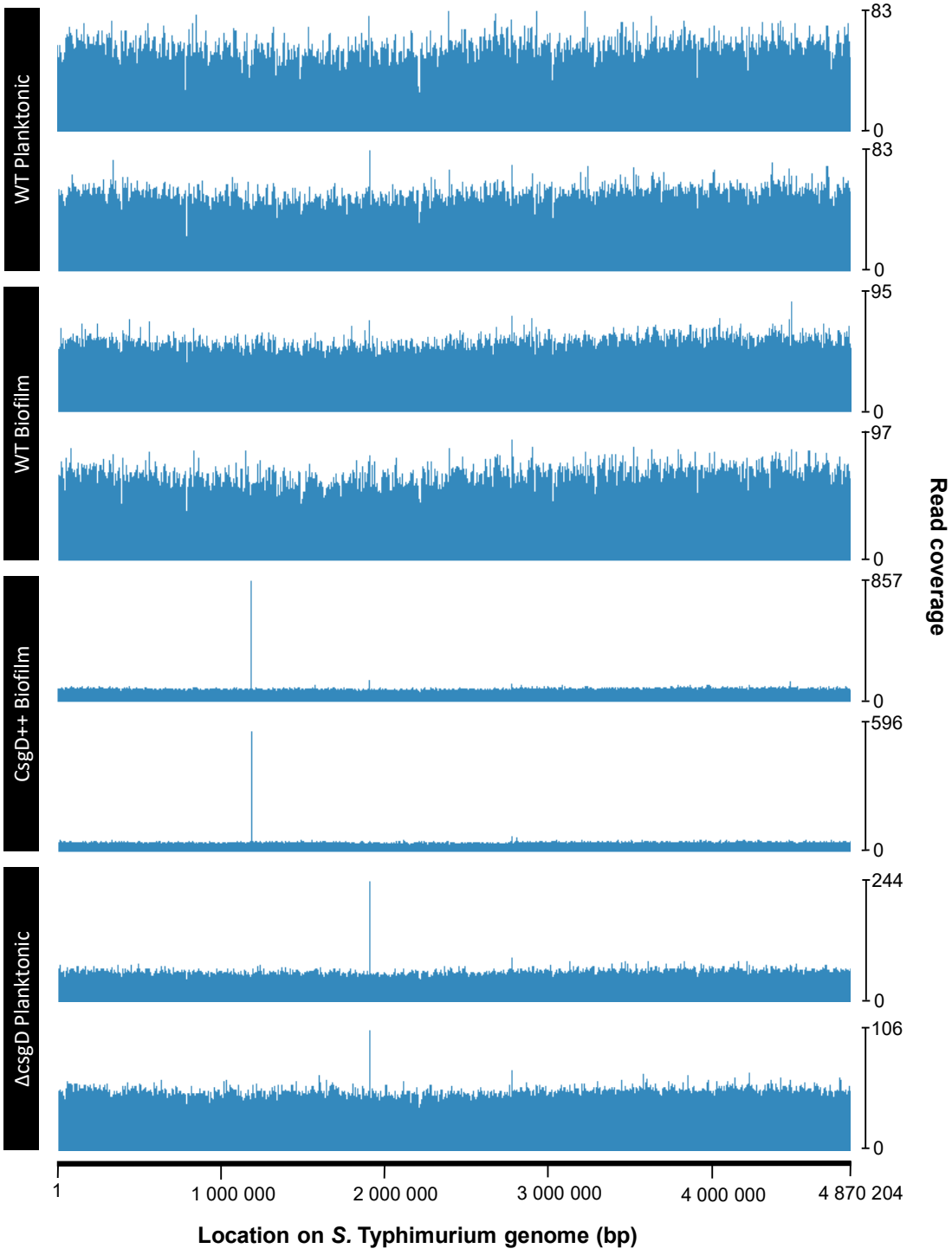


Figure 3.19. Coverage map histogram of ChIP-sequencing reads mapped to *S.* Typhimurium reference genome and visualized by Geneious 9.1.5 genome browser. Each set of reads were trimmed, forward and reverse reads were aligned, and then reads were mapped to *S.* Typhimurium reference genome CP001363.1. Coverage maps from the duplicate libraries are shown. The scale to the right of the maps indicates read count coverage at each location of the genome.

Table 3.4. Called peaks from Galaxy MACS2 analysis of ChIP-seq data. Sequence reads that were trimmed (Trimmomatic) and mapped to the *S.* Typhimurium reference genome (Bowtie2) were assessed by MACS2 analysis for peaks against *S.* Typhimurium wild type planktonic or ΔcsgD planktonic background control. MACS2 returned peaks that were less than 100-fold and greater than 2-fold above background. Peaks that were two-fold above background were recorded.

ChIP sample	Library	Background (Strain, cell type, library)	Peak start (bp)	Peak end (bp)	Length (bp)	Fold enrichment	Associated gene
S. Typhimurium 14028S Planktonic	2	WT P 3	No peak				
		ΔcsgD P 10	4434666	4435462	797	2.12	putative phage tail fiber protein H CDS, putative cytoplasmic protein CDS
		ΔcsgD P 11	2616907	2617662	756	2.20	ethanolamine utilization protein EutK, EutR
			3647225	3647678	454	2.10	putative surface-exposed virulence protein BigA
			2903965	2905614	1650	2.10	Putative permease protein, NupG
	3	No peaks for all background comparisons					
S. Typhimurium 14028S Biofilm	4	No peaks for all background comparisons					
	5	No peaks for all background comparisons					
S. Typhimurium 14028S Δ*csgD* + *pACYC184-csgDcompFOR3* Biofilm	7	WT P 2	1187828	1189268	1441	3.62	CsgD, *csg* intergenic region
		WT P 3	1187832	1189267	1436	3.62	
		ΔcsgD P 10	1187815	1189273	1459	3.63	
		ΔcsgD P 11	1187825	1189271	1447	3.63	
	8	WT P 2	1187765	1189295	1531	6.49	
		WT P 3	1187770	1189292	1523	6.49	
		ΔcsgD P 10	1187769	1189294	1526	6.49	
		ΔcsgD P 11	1187770	1189292	1523	6.49	
S. Typhimurium 14028S Δ*csgD* Planktonic	10	No peaks for all background comparisons					
	11	No peaks for all background comparisons					

3.4 Discussion

A second ChIP-seq experiment did not return any meaningful statistically significant peaks at genomic regions regulated by CsgD. There is an inherent risk for introducing error in the many procedural steps involved in ChIP-seq. However, risk of error was minimized by the use of a verified monoclonal anti-CsgD antibody and by step-wise alterations to ChIP methods based on successful published protocols and experimental optimization[159].

In my initial ChIP-seq experiment, DNA samples were harvested from wild type *S.* Typhimurium biofilm and planktonic cells at 13 hours and from biofilm at 32 hours of flask growth. In the flask culture at 13 hours, intracellular CsgD is highest and is expected to bind to its regulatory regions, whereas intracellular CsgD concentration lessens and biofilm matures at 32 hours. The *S.* Typhimurium CsgD overexpresser strain has high levels of intracellular CsgD and is expected to saturate all available binding locations to return strong, positive regulatory peaks from immunoprecipitation. ChIP DNA strains, cell types, and time points were refined for the second experiment. CsgD binding and regulatory control is expected to occur earlier in *S.* Typhimurium growth; therefore, we did not include a biofilm sample from 32 hour flask culture. Instead, *S.* Typhimurium 14028S Δ*csgD* was harvested as negative control, since it should not return any regulatory peaks from immunoprecipitation with anti-CsgD antibody. Furthermore, immunoprecipitation was designed to facilitate a different type of analysis. In the first ChIP experiment, sonicated input DNA was harvested in parallel with immunoprecipitation with the monoclonal anti-CsgD antibody. Peaks were distinguished as regions of coverage above background coverage in a single set of reads, but background also could have been referenced as Input DNA. Immunoprecipitation in the second experiment was much more stringent: sonicated input DNA and DNA from immunoprecipitation with normal mouse IgG could be used as background for immunoprecipitation with anti-CsgD antibody. This method accounts for error

that could arise from immunoprecipitation, whereas sonicated input DNA simply covers the genome and can introduce error through local differences in coverage.

The second ChIP sample preparation was improved from the last. The new monoclonal antibody (6D4) bound better to CsgD than the antibody clone used in the initial ChIP-seq experiment. Methods were also changed based on analysis of successful ChIP-seq with bacterial transcription factors. For example, a washing step was added after crosslinking to remove excess crosslinker, a preclearing step was added before adding Protein G magnetic beads to the sonicated cell lysate with antibody, and a portion of the precleared chromatin was immunoprecipitated with normal mouse IgG (i.e. "control IP") For the most part, buffer recipes varied little between published protocols and functioned as they were intended. I also changed a few methods after experimental optimization for *S.* Typhimurium cell types and the lab equipment available for preparing ChIP DNA. The goal of method optimization was to improve sample consistency and thereby reduce opportunities for introducing error. I determined the amount of planktonic cells to harvest so both cell types have an equal amount of substrate for protein crosslinking. Mechanical homogenization of cell samples in a mixer mill at 30 Hz for 5 min broke apart biofilm consistently. After consulting with collaborators collaborators (Dr. Carsten Kröger, Trinity College Dublin, Stefani Kary, MSc.,Trinity College Dublin), cell samples were incubated in lysis buffer on ice longer, which reduced the amount of visible material in the tube before sonication. Planktonic cells were processed by mixer mill at the same time to reduce variables. DNA fragmentation by sonication with a cup horn sonicator was tested, but was not as consistent and effective as probe sonication. Most protocols that were surveyed used column-based or phenol-chloroform extraction. Magnetic beads were chosen for future ChIP DNA purification after I performed an assay comparing column kits to magnetic beads.

Libraries were made from ChIP DNA, which were tested for concentration, fragment size range by Bioanalyzer, and enrichment by qPCR. After size selection during library preparation, library fragment sizes were confirmed with a range of about 300-700bp and a median of 450 bp. Tests for the enrichment of one known regulatory target of CsgD, *csgB*, against control target *groEL* returned inconsistent results. There were a couple of issues that could have been responsible for these results: 1) bias could have been introduced due to sonication, 2) tests with ChIP DNA could have been at the limits of detection, and 3) the *csgB* target could have been outside of a regulatory peak, which are often narrow for transcription factors. We concluded that the any issues of detection that were encountered with qPCR would be resolved by sequencing, since it is much more sensitive.

Despite these improvements to ChIP methods, no significant peaks at CsgD regulatory regions were identified. By visual inspecition and MACS2 peak calling, a strong peak 3.6-6.5 fold above background was identified in *S.* Typhimurium ST14028 ΔcsgD + *pACYC csgDcompFOR3* biofilm ChIP reads. This peak is not significant because the plasmid contributed extra copies of *csgD* with the intergenic region to background reads. Peaks at other known regulatory regions (i.e., *adrA*), or at unknown regions should accompany this peak, if it was truly the result of successful immunoprecipitation. In addition, the intergenic peak was abnormally broad for normal transcription factor binding sites. MACS2 analysis also identified peaks in *S.* Typhimurium ST14028 planktonic ChIP reads (Library 2) that were about 2-fold above background. Since these peaks are not called in the other planktonic ChIP reads (Library 3), are fairly low fold enrichment, and are in unrelated genes, they are likely nonsignificant. Localized low coverage in background alignments compared with Library 2 coverage at these genomic locations may have caused MACS2 to call peaks. By visual inspection, peaks were observed at

ycgO in *S.* Typhimurium ST14028 ΔcsgD planktonic ChIP reads. They were also observed to a lesser extent in *S.* Typhimurium ST14028 wild type biofilm and planktonic ChIP reads. If this peak were significant, it would be accompanied by known binding regions (i.e. *adrA, csg*) or other biologically relevant genomic sites, and would be distinguishable by differential binding analysis between strain and cell types. This peak may be an artifact from sonication or immunoprecipitation. Common artifacts in ChIP experiments with human cell lines are "blacklisted" on databases such the Duke Excluded Regions (DER)[74,89,95]. However, it appears that no such database exists for bacterial reference genomes.

The low yield of DNA that I've achieved after immunoprecipitation may still be attributed to ChIP selection of a small population of fragments. Enrichment tests by qPCR were inconsistent, but in hindsight may have indicated low enrichment. However, no change in enrichment was observed for *S.* Typhimurium ST14028 ΔcsgD + *pACYC csgDcompFOR3* biofilm ChIP DNA, which contained additional *csg* fragments from the plasmid and should show artificial enrichment. A common ChIP issue, antibody specificity, is likely the cause of immunoprecipitation failure in this experiment. Even though the second antibody appeared to bind CsgD in immunoblots with sonicated cell lysates prepared for ChIP, binding epitopes may not be available for immunoprecipitation of DNA-CsgD crosslinked fragments. Epitopes for efficient ChIP binding may differ from those predominant in the CsgD-His protein used to generate the monoclonal antibodies or they may be shielded by CsgD dimerization[184], interactions with DNA or other proteins, or protein folding *in vivo*[113]. Commercial antibodies against epitope tags may provide more consistent read peaks from immunoprecipitation. ChIP-sequencing remains a powerful tool for identifying the regulatory targets of transcription factors, and generates massive amounts of data for analysis and integration with other components of

regulatory networks. The modifications made to these ChIP-seq methods for the identification of CsgD regulatory targets in the two *S.* Typhimurium cell types may be useful if the right antibody or anti-TF construct is chosen.

TRANSITION

Origins of signals for regulation of biofilm formation

Overall, my hypotheses address the changes that *S.* Typhimurium initiates in response to environmental stress. My two main research projects concern phenotypic heterogeneity from different perspectives. The project previously described, identifying the regulatory targets of CsgD using ChIP-seq, focuses narrowly on one transcription factor and the handful of genes it controls under laboratory conditions. The research project I will describe in the next chapter considers the broader relationship between the host and *S.* Typhimurium during infection, with the goal of identifying genes that exhibit altered expression levels when exposed to factors present in the gut. These changes in gene expression may be involved in *S.* Typhimurium phenotype switching and strategies of infection and transmission. Together, these findings could identify the intrinsic and extrinsic factors that lead to population heterogeneity.

4.0 EXPRESSION OF VIRULENCE AND PERSISTENCE GENES IN THE PRESENCE OF WASTE EFFLUENT

4.1 Introduction

MacKenzie *et al*. discovered that many genes were differentially expressed in biofilm and planktonic cell types that are formed in response to environmental stress[111]. The signals for initiation of phenotype switching are complex and may originate from stochastic events, cell age, cell-to-cell interactions, or signals received from the host during the course of infection[1]. Currently, we do not know when *S.* Typhimurium experiences signals leading to bistability of CsgD, or if differential gene expression leading to population heterogeneity arises within the host due to small molecules encountered in the gut from the host or the microbiota. The goal of this research project was to determine if small molecules from the human gut microbiota change expression of genes involved in *S.* Typhimurium phenotype switching. A precedent for this experiment was set by Antunes *et al*., who demonstrated that a small molecule secreted by microbiota member Clostridia represses expression of *S.* Typhimurium SPI-1 and genes involved in host cell invasion[7]. Therefore, the expression of genes known to be involved in virulence or persistence were measured in the presence of effluent from the steady-state culture of a human intestinal microbiota. Promoters for genes involved in biofilm formation or virulence were inserted behind the *luxCDABE* operon on a plasmid introduced to *S.* Typhimurium. When the promoter is activated, *luxCDABE* gene products are made, which oxidizes a reduced flavin mononucleotide ($FMNH_2$) and produces light at 490 nm, with a secondary emission at 590 nm[25,117]. Light production due to promoter activation can be measured quantitatively as gene expression levels throughout the organism's growth[177]. The growth media for these reporter strains was supplemented with liquid gold, waste effluent from a steady-state chemostat seeded with donor feces. Liquid gold is thought to contain small molecules present in the gut that

originate from the host or microbiota[4,7], and could contain the signals for bistable expression of virulence and persistence genes. I noted whether there was any difference in promoter activity when *S.* Typhimurium was grown in the presence of waste effluent at biofilm-inducing and host-mimicking growth conditions.

4.2 Materials and Methods

4.2.1 Strains for luciferase assays

Promoters for genes that are important in virulence or persistence functions were chosen for ligation with pCS26 plasmid vector containing *luxCDABE* genes and kanamycin antibiotic resistance selection genes (Table 4.1). These ligated plasmids were electroporated into *S.* Typhimurium.

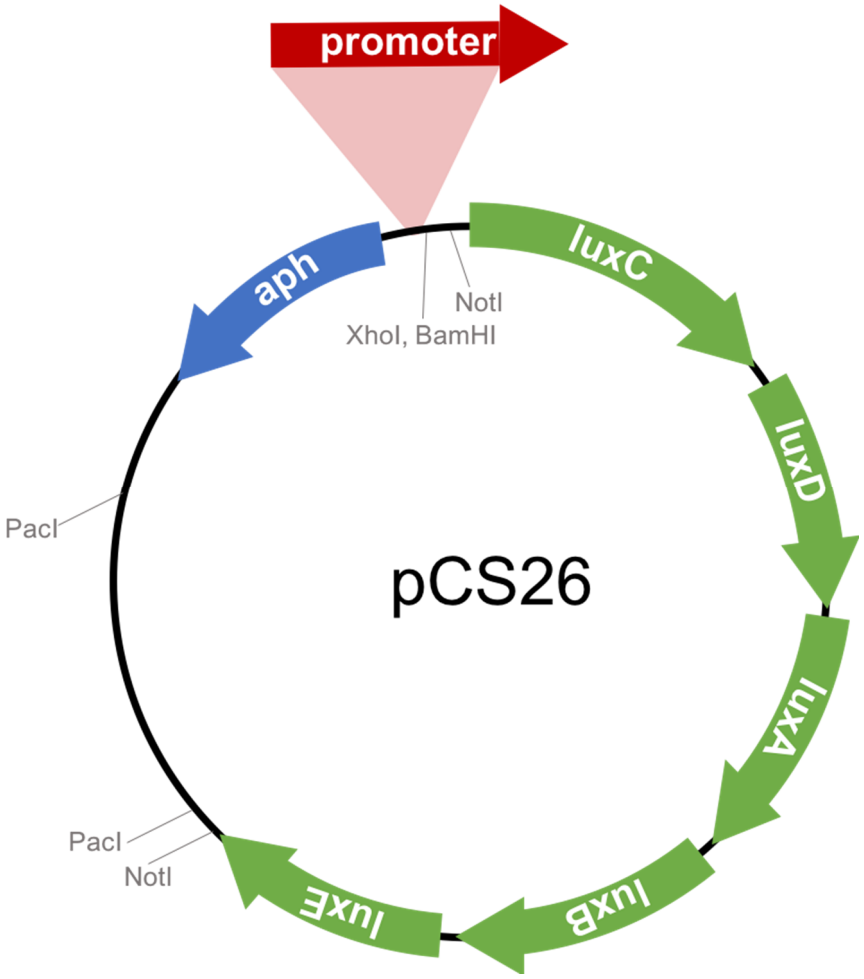


Figure 4.1. Promoters for genes primarily expressed in biofilm or planktonic cell types were cloned behind *luxCDABE* reporters in pCS26 plasmid and electroporated into *S.* Typhimurium 14028. This was maintained through kanamycin resistance (*aph*) carried on the plasmid. When the promoter is activated, luciferase genes are expressed and light measured as fluorescence counts per second (CPS) is detected proportional to the strength of promoter activation[13].

Table 4.1. ***S.* Typhimurium ST14028 strains containing pCS26 vectors with luciferase genes (*luxCDABE*) controlled by a *S.* Typhimurium promoter involved in virulence or biofilm formation.** Promoter-reporters labeled (*) were constructed and tested by White *et al*.[189] and promoter-reporters labeled ($^{+}$) were constructed for luciferase assays with waste effluent.

Plasmid *pCS26*	**Gene function**	**Gene association**
PadrA::lux*	diguanylate cyclase, produces c-di-GMP, catalyzes conversion of 2GTP into c-di-GMP, induces cellulose biosynthesis, cell adherence, swimming/swarming, activates cellulose biosynthesis (*bcsABZC*)	Persistence
PcsgB::lux*	Curli fimbriae subunit for adhesion, biofilm formation, and aggregation	
PcsgD::lux*	DNA-binding transcriptional regulator CsgD, activates *csgBA* and *csgDEFG* operons	
Psig38H4::lux*	RpoS, global regulator of gene expression during stress and starvation	
PSTM1987::lux	Diguanylate cyclase, produces c-di-GMP, involved in cellulose synthesis and biofilm formation	
***PhilA::lux*$^{+}$**	Transcriptional regulator, activates expression of invasion genes an and pathogenicity island type III secretion system	Virulence
PhilD::lux	Transcriptional regulator, helix-turn-helix, activator for invasion genes, derepresses *hilA* expression	
PinvF::lux	Transcriptional regulator for SPI-I Type III Secretion system effector proteins	
***PmisL::lux*$^{+}$**	Binds fibronectin, allows for colonization of the intestine	
***PprgH::lux*$^{+}$**	Secretion system protein, needle complex inner membrane protein, invasion type III secretion apparatus	
***PshdA::lux*$^{+}$**	Involved in intestinal persistence and prolonged shedding	
PsicA::lux	SPI-1 chaperone, regulates expression of virulence genes	
***PssrAB::lux*$^{+}$**	Secretion system sensor kinase, type III secretion system regulator, regulated by OmpR-EnvZ	
PyhjH::lux	c-di-GMP phosphodiesterase degrades c-di-GMP to counteract biofilm formation, enhances motility	

4.2.2 Preparation of waste effluent

Initial luciferase assays were performed with liquid gold in growth media; however, results were inconsistent between replicates. This may have been due to unequally distributed silt-like material in the liquid gold. After consulting Dr. Emma Allen-Vercoe, we performed separation and filtration steps on waste effluent and chemostat medium to remove this material.

Thawed waste effluent and chemostat media was spun to pellet precipitates (47 800 xg at 4°C for 45 minutes, Figure 4.2a). These tubes were placed on ice and the supernatant was filtered with a 0.45 µm syringe filter, followed by a 0.2 µm syringe filter. Filtrate of waste effluent and chemostat media was dispensed into single-use aliquots and stored in the freezer to prevent freeze-thaw degradation. Filtered waste effluent and chemostat media, shown in Figure 4.2b, was light yellow and free of debris.

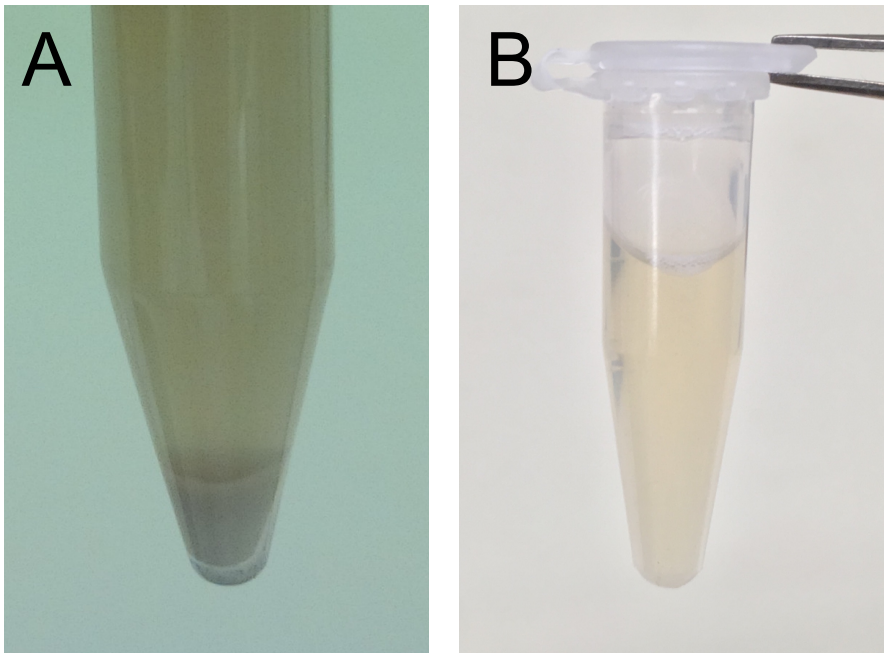


Figure 4.2a and b. Waste effluent, or "liquid gold" before (a) and after (b) separation and filtration to remove precipitate material. Precipitated material was pelleted by centrifugation, and the supernatant was filtered down to 0.2 µm.

4.2.3 Luciferase assays measuring promoter activity in the presence of waste effluent

Gene expression in the presence of waste effluent was assessed at 37°C and 28°C, human core temperature and environmental biofilm temperature, respectively. The environmental temperature is expected to promote biofilm-associated gene expression, and the human core temperature is expected to more closely match host infection. Gene expression was also assessed in the presence of waste effluent and iron chelator 2,2-dipyridyl at 37°C and 28°C by luciferase assay. Overnight cultures of bacterial reporter strains (Table 4.1) were grown at 37°C with shaking in LB supplemented with 50 µg/mL kanamycin. These overnight cultures were diluted 1 in 600 in 1% tryptone supplemented with 50 µg/mL kanamycin to a final volume of 150 µl in a 96-well clear-bottom black plate (Corning, #3631). When specified, supplements such as diluted waste effluent, chemostat media, or 2,2-dipyridyl (Sigma-Aldrich, #D7505) were added to wells containing media. Waste effluent was diluted in chemostat media to 1 in 10 and added at one-tenth of the final well volume for a final dilution of 1 in 100. Growth media with kanamycin alone, or the latter with supplemented chemostat media was used as a control for waste effluent. Each well contained the same final volume, and media was adjusted such that each well was nutritionally consistent. Mineral oil was added on top of the media with cells to prevent evaporation. Absorbance was measured at 590 nm and luminescence was measured as counts per second (CPS) on the VICTOR X3 or VICTOR^{3}V 1420 (Perkin Elmer, cat. 2030-0050, 1420-040) every 30 minutes with agitation (1 min, 10 min intervals) for 48 hours at 28°C or 37°C. The run was restarted at 48 hours. After the data were exported from the Victor, any data for which growth by OD_{590} was slow or abnormal was excluded. Data for OD_{590} and luminescence over time showed expression levels during *S.* Typhimurium growth.

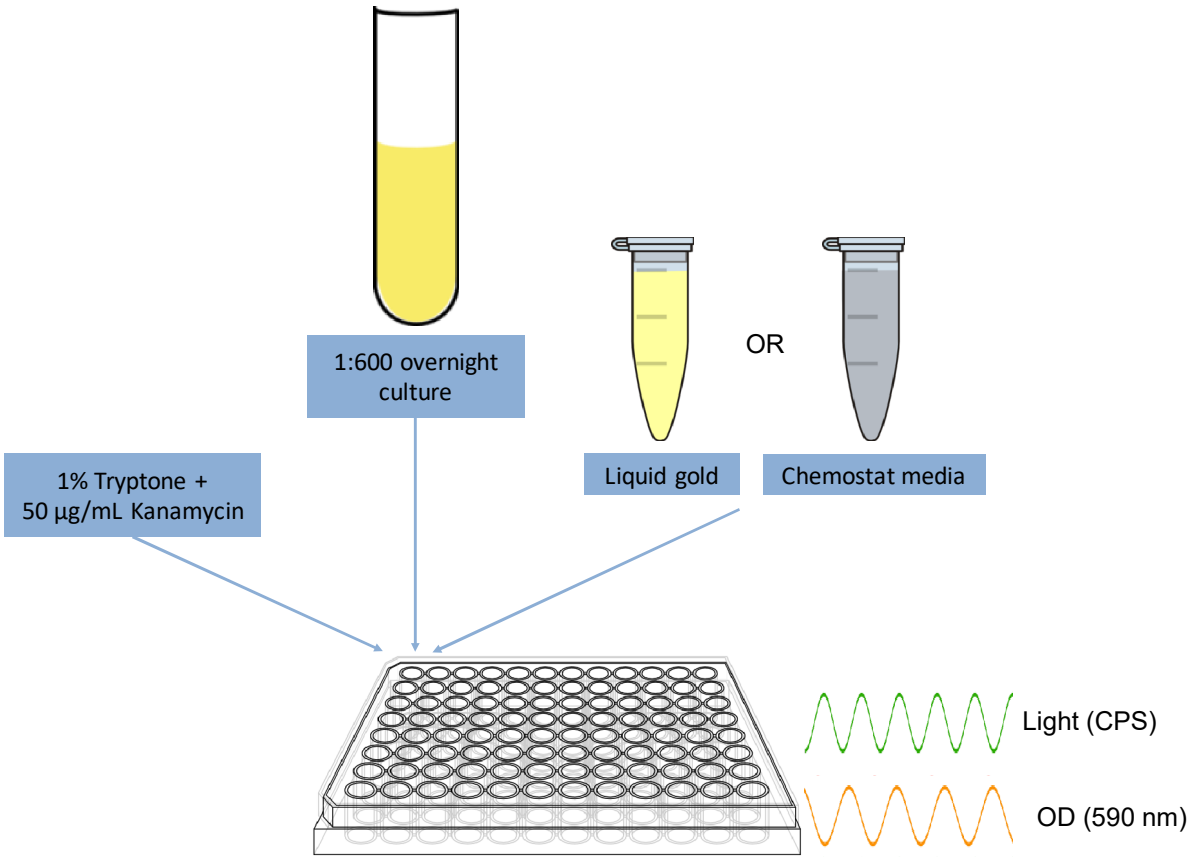


Figure 4.3. An overview of the method used to perform luciferase assays with waste effluent or chemostat media. Growth media (1% Tryptone) with 50 µg/mL Kanamycin is added to each well in a 96-well, clear bottom plate. Strains containing a pCS26 plasmid with *luxCDABE* genes behind a promoter of interest are grown overnight at 37°C with shaking and added to the appropriate wells. Liquid gold (waste effluent) or chemostat media are added to the appropriate wells at 1:100 final dilution, and mineral oil was layered on top. Cell growth was measured at 590 nm and light production was measured as counts per second (CPS) by a plate reader.

4.3 Results

The expression of promoters (listed in Table 4.1) in *S.* Typhimurium was assessed by luciferase assay in the presence of waste effluent at 28°C and 37°C and supplemented with iron chelator 2,2-dipyridyl. Expression of persistence- and virulence-associated genes was assessed in the presence of waste effluent at 28°C, introducing environmental stress, and at 37°C, reproducing conditions in the host gut. To test conditions that more closely mimicked the human gut, an additional parameter, iron limitation, was altered. The human body and microbiome has several mechanisms of sequestering iron, and as such, it is a growth-limiting nutrient[5,38].

Representative results were shown for each set of growth conditions (Figure 4.5a-e, Figure 4.6a-f). Figure 4.4 shows a normal growth plot for the strains that were tested in each set of conditions. The strains grew to about 0.6 OD_{590} in media supplemented with waste effluent, slightly less in media supplemented with chemostat media, and to 0.4 OD_{590} in unsupplemented media. There were five main trends in expression that were observed in these assays: 1) expression peaking at about 25 hours of growth, 2) time-shifted expression peaking at about 50 hours of growth, 3) no expression, 4) an early drop-off in expression, and 5) low, random, or multimodal expression. Expression peaking at about 25 hours of growth was observed for virulence-associated genes at 28°C, which is shown in representative plots of *hilD* and *yhjH* (Figure 4.5a, Figure 4.6b). Most biofilm-associated gene expression at 28°C was time-shifted and peaked at about 50 hours of growth. This is shown in representative plots of *csgD* with or without 2,2-dipyridyl (Figure 4.5b, Figure 4.6a). No expression was observed for most persistence-associated genes at 37°C, which is shown in representative plots of *csgB* and *adrA* (Figure 4.5c, Figure 4.6d). An early drop-off in expression was observed for most persistence-associated genes at 37°C, which is shown in representative plots of *misL* and *sicA* (Figure 4.5d, Figure 4.6e). Random or multimodal expression was observed for only a few genes. Multimodal

expression was observed for *csgD* at 37°C, *sig38H4* at 37°C with 2,2-dipyridyl, and for *ssrA* at 37°C without 2,2-dipyridyl and at 28°C with 2,2-dipyridyl demonstrated in Figure 4.5e and 4.6f. In general, higher expression levels by fluorescence counts were observed in the presence of waste effluent, followed by the chemostat control and finally, media without supplements.

A third luciferase assay in the presence of waste effluent, at 37°C with 0.2 mM iron chelator 2,2-dipyridyl was performed, which mimicked gut conditions with a higher amount of iron chelator. This amount was chosen because it was used by Romling *et al*. to observe *agf* (*csg*) expression and aggregative fimbriae production in environmental conditions of iron depletion[155]. Expression results were similar to the previous experiment with more random error, which could have originated from the iron chelator (data not shown). The data were largely inconsistent; however, it appears that the additional 2,2-dipyridyl did not increase expression of any genes in the presence of waste effluent.

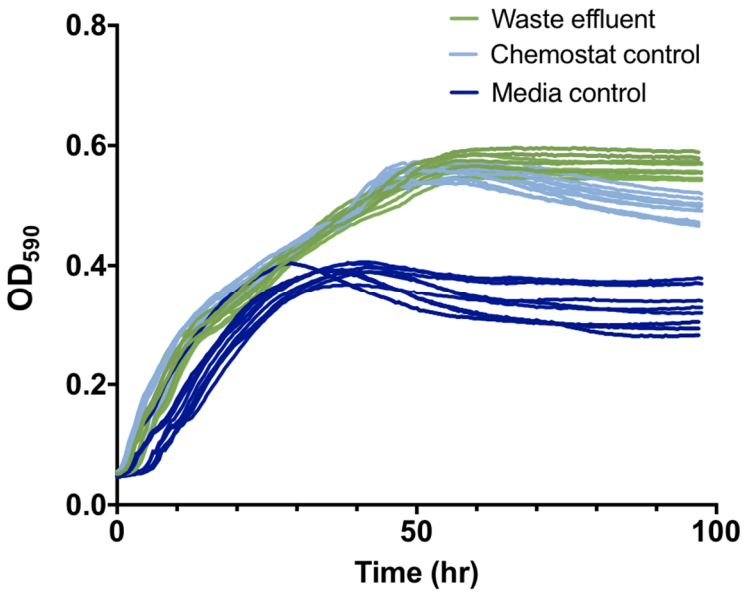


Figure 4.4. Representative growth plot for *S.* Typhimurium strains grown in 1% tryptone with waste effluent at 28 °C and 37 °C. The growth of *S.* Typhimurium strains in the presence of waste effluent, chemostat media, and a media control was measured at 590 nm over the course of the luciferase assay. The OD_{590} for the treatment groups in this growth plot of the *S.* Typhimurium 14028 *PcsgD::lux* promoter-reporter strain at 28°C is typical for other promoter-reporters.

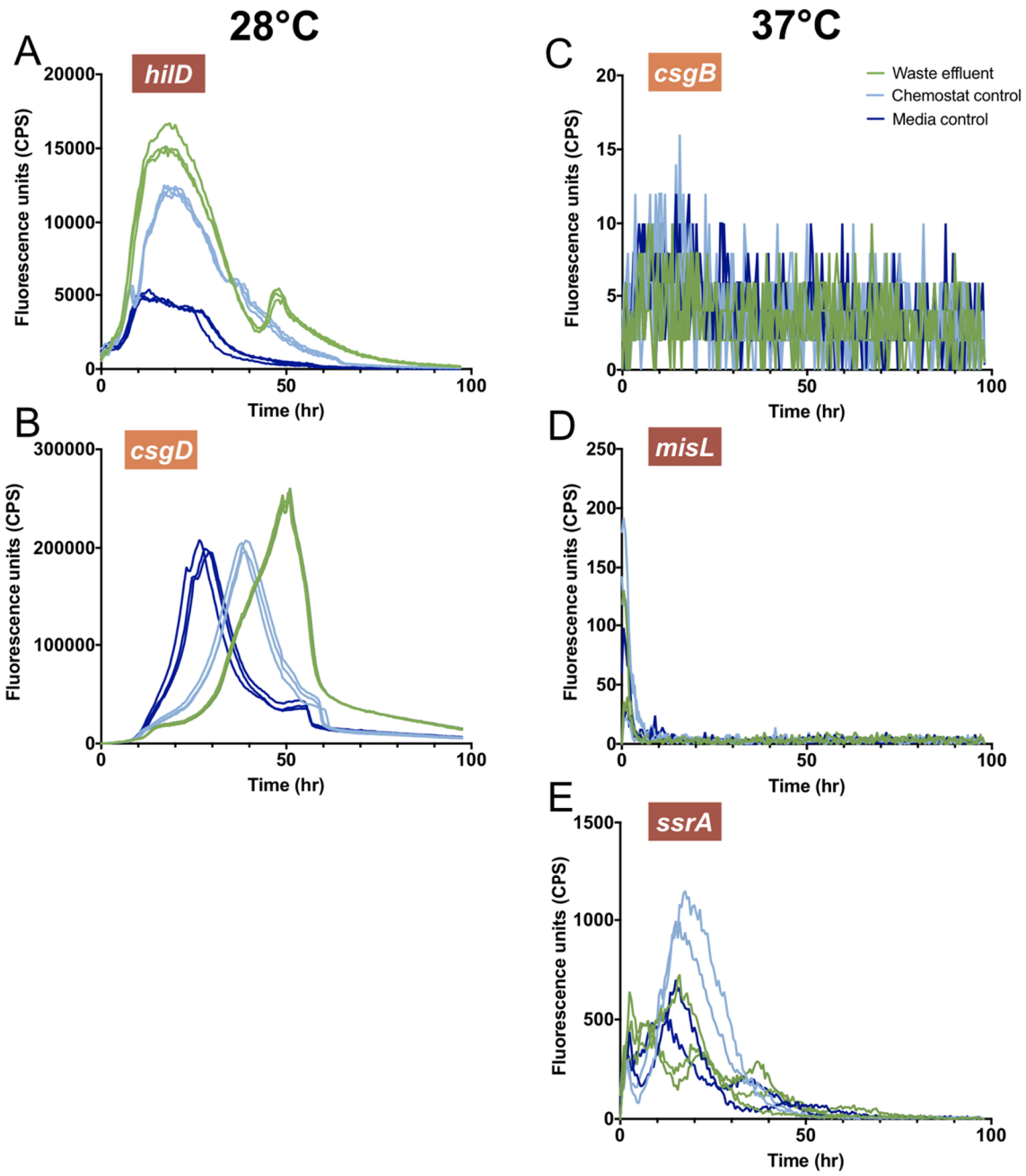


Figure 4.5a-e. Representative expression plots for *S.* Typhimurium promoter-reporter strains grown in 1% tryptone with waste effluent at 28 °C and 37 °C. Gene expression in *S.* Typhimurium promoter-*luxCDABE* reporter strains grown at 28°C (**a** and **b**) or 37°C (**c-e**) in 1% tryptone with 1:10 waste effluent measured as luciferase reporter signal over time on a Victor plate reader. Expression plots for *hilD* (**b**), *csgD* (**c**), *csgB* (**d**), *misL* (**e**), and *ssrA* (**f**) are representative of trends observed for other genes. Red labels indicate virulence-associated genes and orange labels indicate persistence-associated genes.

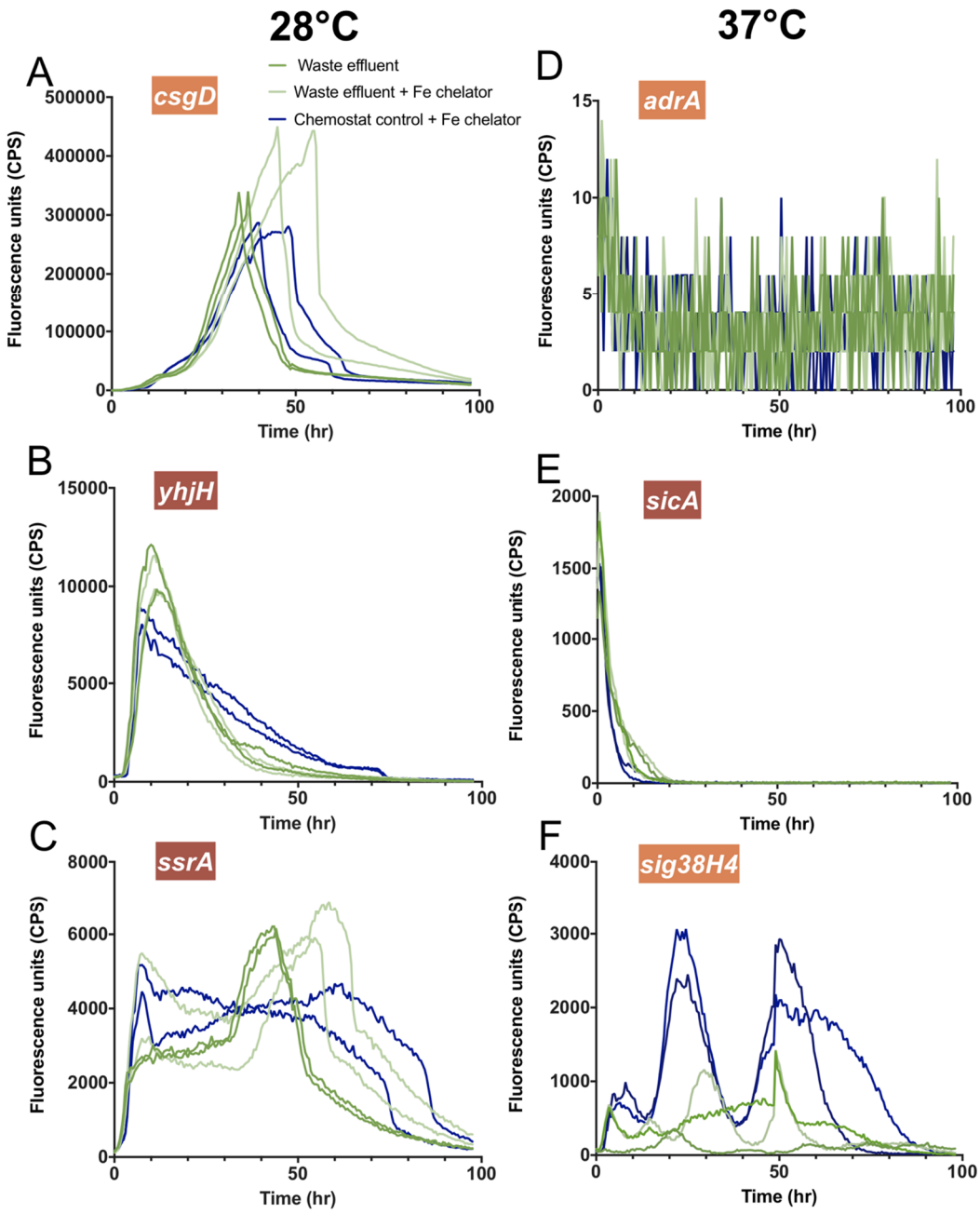


Figure 4.6a-f. Representative expression plots for *S.* Typhimurium promoter-reporter strains grown in 1% tryptone with waste effluent and 40 µM iron chelator 2,2-dipyridyl at 28 °C and 37 °C. Gene expression in *S.* Typhimurium promoter-*luxCDABE* reporter strains grown at 28°C (**a-c**) or 37°C (**d-f**) in 1% tryptone with 1:10 waste effluent and 40 µM 2,2-dipyridyl measured as luciferase reporter signal over time on a Victor plate reader. Expression plots for *csgD* (**a**), *yhjH* (**b**), *ssrA* (**c**), *adrA* (**d**), *sicA* (**e**), and *sig38H4* (**f**) are representative of trends observed for other genes. Red labels indicate virulence-associated genes and orange labels indicate persistence-associated genes.

4.4 Discussion

Luciferase assays were performed by measuring light production of strains with a plasmid-based promoter-*luxCDABE* reporter when gene expression is induced. The reporter strains harboured promoter-reporter constructs for persistence- or virulence-associated genes of interest (Table 4.1). When strains were grown in tryptone media supplemented with waste effluent at 28°C, most genes were expressed at some point during growth. Virulence-associated genes were expressed earlier during exponential growth at or before 25 hours. Conversely, persistence-associated genes were expressed at the beginning of stationary phase. These results were expected, due to the functions required by *Salmonella* to survive long-term in conditions low in nutrients. Overall, higher expression was observed in the presence of waste effluent at 28°C, but it was not significantly (i.e., 2-fold) above expression levels observed in chemostat control and tryptone. The strains also grew to different optical densities measured at 590 nm in waste effluent, chemostat control and media control, in descending order. Therefore, expression levels above media control and differences in the timing of biofilm-associated peak expression may be attributed at least partially to growth resources added by chemostat media or waste effluent. Shifts in the timing of in biofilm-associated peak expression may represent a delay in expression of genes involved in stress response until *S.* Typhimurium has run out of resources, which could occur later in the presence of resource-rich chemostat media and waste effluent. In general, growth of the *S.* Typhimurium strains in tryptone is less than in tryptone supplemented with chemostat media or Waste effluent. Waste effluent and chemostat media are hypothesized to bolster bacterial growth, with a higher concentration of resources usable by *S.* Typhimurium compared to the same volume of tryptone media.

Gene expression patterns were different when strains were grown in tryptone media supplemented with waste effluent at 37°C than expression patterns at 28°C. The majority of

virulence-associated genes shut off expression before 10 hours of growth (refer to Figure 4.5d). Virulence-associated genes are normally expressed in elevated temperatures and rich media during exponential growth; this drop off in expression could be indicative of a shift in expression from rich media in the overnight inoculum to limiting media during luciferase assays. Persistence-associated promoters *adrA*, *csgB*, and *STM1987* were inactive at 37°C (refer to Figure 4.5c). These genes are generally expressed when *S.* Typhimurium is exposed to environmental stress and temperatures below 28°C. Low, random, or multimodal expression was observed for *ssrA* and *csgD* (refer to Figure 4.5e). The reason for this is unknown; however, it may be due to additional resources for growth in chemostat media and waste effluent that were described previously. Ultimately, the presence of waste effluent at 37°C did not significantly increase expression of any of the genes surveyed.

Mammalian cells have strategies of sequestering iron, and as such, iron availability is a common limiting factor for growth of microorganisms internally. Previous luciferase assays did not indicate a major role of waste effluent in altering gene expression; however, conditions that mimic the gut in temperature, small molecule content, and iron limitation were hypothesized to alter gene expression. Promoter activity of persistence- or virulence-associated genes listed in Table 4.1 were measured by luciferase assay in the presence of waste effluent and 40 μM iron chelator 2,2-dipyridyl at 28°C, to assess the stress response, and 37°C, to assess human distal gut conditions. Expression patterns were largely the same at 28°C in the presence or absence of 40 μM 2,2-dipyridyl; however, the expression of biofilm-associated genes was slightly enhanced in the presence of waste effluent and 2,2-dipyridyl above that of media with waste effluent or media with chemostat control and 2,2-dipyridyl (refer to Figure 4.6a). This may be a response to limited resources. Expression patterns for virulence-associated genes was similar to those observed in

the absence of 2,2-dipyridyl, with a drop in expression before 25 hours of growth. However, no effect of waste effluent and iron limitation was observed. The expression of *ssrA* was inconsistent, bimodal, and showed no effect of waste effluent or iron limitation which was similar to the pattern of expression observed for *ssrA* in the absence of the iron chelator at 37°C. Iron chelation may mimic the distal gut in a similar way that temperature does for *ssrA* expression. Expression patterns were similar at 37°C in the presence or absence of 40 μM 2,2-dipyridyl. The biofilm-associated genes that were shut off in media with waste effluent at 37°C were also shut off in the presence of 2,2-dipyridyl. Likewise, virulence-associated genes were expressed initially and dropped off in media with waste effluent at 37°C also did so in the presence of 2,2-dipyridyl. Expression of *csgD*, *ssrA*, and *sig38H4* was static and multimodal at 37°C whether the iron chelator was present or absent, with no effect of waste effluent observed. Taken together, these findings did not indicate an obvious role of waste effluent and iron limitation in altering expression of persistence- and virulence-associated genes at 37°C.

I increased the amount of 2,2-dipyridyl in luciferase assays with waste effluent at 37°C from 40 μM to 200 μM (0.2 mM) to see if these conditions would mimic the distal gut more closely and alter gene expression. The increase in 2,2-dipyridyl did not change gene expression from the patterns observed in media with waste effluent and 40 μM 2,2-dipyridyl at 37°C. Expression was often shifted in time or inconsistent between replicates. This may be due to the excessive amount of iron chelator added, which did not promote normal growth. Overall, waste effluent and iron chelation at environmental and internal temperatures did not significantly alter expression of persistence- and virulence- associated genes by luciferase assay. Waste effluent may signals from the human microbiota for phenotype switching; however, they were not immediately apparent in this context and with the promoters that were assessed. These luciferase

assays should be regarded as preliminary; more research is required to understand the subtler increases in expression in the presence of waste effluent, and the trends in expression observed for *S.* Typhimurium virulence- and persistence-associated genes.

.

5.0 CONSIDERATIONS AND CONCLUSIONS

5.1 Discussion

Three key findings led to my research project to identify the regulatory targets of CsgD using ChIP-seq. Firstly, the identification of a phenotypically heterogenous population of *S.* Typhimurium biofilm and planktonic cells when exposed to environmental stress [61,186]. Secondly, that 34% of the genes in the genome were differentially expressed between the two cell types[111]. And thirdly, that the master biofilm regulator, CsgD, is at high intracellular levels in biofilm and at low intracellular levels in planktonic cells[61]. We hypothesized that bistable expression CsgD, the master biofilm regulator, was responsible for coordinating phenotype switching through its functions as a transcription factor. I intended to identify the genes regulated by CsgD and evaluate those targets by differential analysis in the two cell types. However, the *S.* Typhimurium cell types are physically very different and present a unique challenge to manipulate and evaluate as equivalents.

I performed an initial ChIP-seq experiment with a published protocol for a successful ChIP experiment in *S.* Typhimurium by Dillon *et al*.[41] We designed an experiment using the same organism and protocol, differing only by anti-CsgD monoclonal antibody, strain controls, background controls, and cell harvesting. This experiment did not return any peaks for regulatory targets of CsgD. Given the success of many ChIP experiments and the few differences between published procedures and our procedure, these results were unexpected. Nonetheless, ChIP-seq is a powerful and effective method, and produces massive amounts of high-quality data for identifying regions of DNA that are controlled by DNA-binding protein. For this reason, I diagnosed issues with this ChIP experiment and proceeded with a second, improved ChIP experiment to find the regulatory targets of CsgD.

Two main items were identified for improvement over the initial ChIP-seq experiment: 1) antibody specificity, and 2) enhanced sample consistency. Bortz and Wamhoff judiciously warned that "the assay can be cumbersome and fraught with ample opportunity to introduce technical error[159]." It is true that there are many steps involved in ChIP with their own ways to introduce error. However, the published statement has been countered with an increase in number and quality of published ChIP-seq findings over the past seven years. We proceeded with due care. I evaluated each step in our procedure evaluated against the literature to introduce improvements. I checked each testable step in our procedure for the expected result. I optimized each step in our procedure for *S.* Typhimurium cell types and the lab equipment used to sonicate, immunoprecipitate, and purify DNA. Additionally, I purified and tested a new monoclonal antibody for its ability to bind to CsgD alone and CsgD in lysates prepared for ChIP-seq, with success. I tested and optimized so many parameters—cell harvesting, biofilm homogenization, sonication, and DNA purification, that we drafted a manuscript for performing ChIP from bacterial biofilms.

Unfortunately, the second ChIP-seq experiment did not yield significant and biologically-relevant regulatory targets of CsgD. Antibody specificity was a likely cause of failure, since I had optimized all testable methods and checked non-testable items against successful experiments in the literature. Antibody specificity is the most important part of a successful ChIP-seq experiment, because it determines both the DNA fragment selection and downstream peak resolution during analysis. I had tested many steps, but immunoprecipitation is one of the only steps that is not directly observable or testable until sequencing. Looking back at methods in the literature, very few, if any, use a monoclonal antibody. As I discussed earlier, antibody specificity is a common issue in ChIP due to epitopes that may be hidden in vivo by protein

interactions, DNA interactions, or protein folding[113]. Monoclonal antibodies often have a single binding site that is unknown unless it has been characterized. Most ChIP experiments are performed with polyclonal serum or an antibody specific for an epitope-tag on the target transcription factor. With polyclonal antibodies, any background that may be introduced due to non-specific binding is countered by the diversity of binding locations on a target protein. Although an epitope tag may interfere with protein function, commercial antibodies against epitope tags are often stringently tested and their use in successful experiments is reported in the literature.

Antibody binding could have been affected by CsgD itself. CsgD is a 25 kDa response regulator protein with an N-terminal receiver domain for phosphorylation and a C-terminal LuxR-like helix-turn-helix motif for DNA binding[194]. Many binding sites for regulation of CsgD expression are known and characterized, but little is known about the regulatory targets of CsgD, and any interactions with other intracellular proteins. I have attempted ChIP-seq to address the dearth of information on the regulatory targets of CsgD; however, CsgD-protein interactions are still uncharacterized, may help coordinate phenotype switching, and may have even interfered with ChIP by obscuring epitopes. Unphosphorylated CsgD binds promoters of *csgBA* and *adrA* and controls transcription during biofilm-forming conditions. During transcriptional control, CsgD dimerizes at two main interfaces before binding regulatory sequences[184]. The anti-CsgD monoclonal antibodies used for both immunoprecipitations was raised against purified, His-tagged CsgD monomers. Therefore, the epitopes of CsgD recognized by antibodies *in vivo* could be 1) hidden by dimerization, 2) absent due to His-tag interference, 3) on CsgD in an unknown phosphorylation state, 4) hidden by unknown protein interactions, or 5) hidden by protein folding or DNA binding. A polyclonal antibody could have provided a more robust selection of CsgD-

bound DNA fragments. Even better, as long as biofilm phenotype was not impeded, a terminal epitope tag could enhance immunoprecipitation efficiently.

Our results highlighted the importance of confirming antibody binding as well as enrichment of known regulatory targets by qPCR, if known targets are available. This step is recommended to confirm target enrichment prior to sequencing if known targets are available; however, our results were inconsistent and inconclusive. Enrichment tests by qPCR may be limited by low concentrations of DNA, sonication bias, or targets which are not directly at peak sites. ChIP peaks for transcription factors tend to be narrow or split distributed into forward and reverse peaks; therefore, a target site that is not directly at a peak could be misinterpreted as nonenriched background. We proceeded with sequencing because it provides much richer information and is far more sensitive than qPCR.

Chromatin immunoprecipitation is a complex method that requires knowledge and skills in many different areas of expertise. It involves strain building, cell growth, immunochemistry, protein detection, library preparation, sequencing, and bioinformatic analysis of large amounts of data. It presents an excellent opportunity to learn and apply different techniques or to leverage the strengths of a research team. ChIP-seq methods may vary lab-to-lab with the same core requirements: protein-DNA interactions, lysis, DNA fragmentation, selection by immunoprecipitation, DNA purification, data acquisition, and analysis. Each step should be tested to control variables and avoid introducing error. Variation matters less for a few procedural items that are fairly consistent among published methods, such as the composition of lysis buffers. Compared to other methods of obtaining regulatory data, ChIP-seq is costly in time and expense. However, the cost of sequencing is lowering, and the amount and richness of ChIP-seq data is incomparable.

These ChIP experiments may not have yielded biologically significant regulatory targets of CsgD. Even so, I hope that this method of processing biofilm cells will be useful for other researchers studying *S.* Typhimurium or other biofilm-forming microbes. In particular, the techniques I have tested and described for harvesting biofilm, normalizing cell types, comparing against planktonic cells, can be used as a starting point for processing other biofilm-forming species, with adaptations. Adaptations could include using equation 3.1 to find the wet weight of biofilm to harvest, cell type separation or normalization to a control cell type if required, and mechanical separation of biofilm.

As previously discussed, my research project investigated the origins of signals for population heterogeneity: 1) molecularly and intrinsically by CsgD regulation, and 2) extrinsically with unknown signals from the human gut microbiota. Understanding more about when and why *S.* Typhimurium forms two cell types could reveal opportunities to block transmission or reduce the severity of infections.

Currently, we do not fully understand when and where signals for phenotype switching come from. We do know that the conditions that *S.* Typhimurium experiences once expelled by the host are unpredictable; it could encounter harsh conditions in the environment or a new host to infect. Therefore, *S.* Typhimurium requires strategies to survive in the face of an uncertain future. Phenotype switching produces two distinct cell types: persistent biofilm cells that are resistant to desiccation and antibiotics, and planktonic cells that express virulence factors and are prepared to infect a new host. This way, a population of *S.* Typhimurium can conserve energy and ensure survival by expressing two different suites of genes to perform two vastly different functions. It would be advantageous for phenotype switching to occur in the host gut in anticipation of unknown future conditions. I hypothesized that *S.* Typhimurium experiences

extrinsic signals for state-switching during infection in the human body. Waste effluent from steady state chemostat culture seeded from feces was thought to contain small molecules that could initiate phenotype switching, so this was used observe phenotype switching through expression of virulence- or biofilm-associated genes.

As it turns out, any increase in gene expression was likely due to addition of resources for bacterial growth in chemostat media or waste effluent. We expected that if waste effluent truly had an effect on switching cell types, a large at least 2-fold difference in expression would be observed. We did not observe an increase or decrease in expression to that scale. The differences in expression are subtler than expected, and more research is required to understand whether any increase in expression is biologically significant. Gene expression in the presence of waste effluent did shift gene expression earlier or later, most likely due to additional resources in waste effluent and chemostat media that support bacterial growth. To increase our chances of finding a difference in expression, we added an iron chelator. This would simulate iron limitation *S.* Typhimurium would experience in the human distal gut due to sequestering by the host and the microbiome. Gene expression from these assays did not differ significantly from assays without iron limitation. Instead, it introduced more noise, which Ackermann *et al*. suggested could lead to bistability[1]. Stress and iron sequestering could initiate bistability, but those effects were not observed in these tests.

These luciferase assays measured gene expression of a whole population of cells. We know that a flask culture in similar conditions is comprised of about 61% planktonic cells and 39% biofilm aggregates. Population proportions are not observable in the small wells of a luciferase assay. Therefore, during bistability, these assays would not be able to detect a gene

expression effect wherein one portion of the population drops expression and another portion of the population increases expression.

These luciferase assays with waste effluent approximated conditions that *S. Typhimurium* would encounter in the mammalian gut. There are many other complex interactions that could initiate population heterogeneity. For example, physical interactions between host or microbiota cells and *S.* Typhimurium could initiate population heterogeneity. Signals could also originate from local differences in nutrition, physical interactions with microbiota and material in the host gut, or interactions with the products and cells involved in host defense.

These assays did provide information about which genes involved in virulence or persistence were expressed in tryptone media at 37°C or 28°C. The waste effluent likely has too many unknown components to identify individual small molecules with a large influence. Waste effluent already has a function in culturing fastidious organisms from the human gut microbiome[4]; we could use waste effluent in other assays to understand more about *S.* Typhimurium infections.

5.2 Future directions

In the future, ChIP-seq could be performed with epitope-tagged CsgD and a commercial antibody to overcome limitations to antibody specificity using a monoclonal antibody. In fact, we have built a *S.* Typhimurium strain with a 3xFLAG tag on the C-terminus of CsgD for immunoprecipitation with a commercial anti-FLAG antibody. This *csgD::3xFLAG* construct was inserted chromosomally in *S.* Typhimurium Δ*csgD* to encourage physiological concentrations of CsgD. It retains the biofilm phenotype in flask culture, and has the rdar phenotype on agar supplemented with Congo Red dye, indicating that the 3xFLAG tag does not interfere with

biofilm formation (data not shown). We assume that the function of CsgD as a transcription factor is not altered by the tag; however, the 3xFLAG tag could inhibit interactions with other proteins. ChIP-seq with this strain and the methods I have checked and optimized is expected to produce strong peaks at genomic regions regulated by CsgD.

Another approach could identify the genomic locations regulated by CsgD. A full suite of CsgD binding motifs have been characterized in *E. coli*[127] by ChIP-chip, and a few binding sites have been identified in *S.* Typhimurium[194]. These known binding motifs could be used to find putative regulatory sites *in silico*. These sites could then be assessed for DNA binding by CsgD through DNA footprinting or luciferase assays measuring promoter activity in a wild type and CsgD deletion strain. A similar approach, combined with expression data, was used by Zakikhany *et al.*[194].

As discussed, the methods I have described and developed could be used by other researchers searching for regulatory targets in biofilm-forming bacteria. ChIP-seq, a technique for finding sequences controlled by proteins, is continually being developed for new purposes. In the future, it could be used in conjunction with biofilm methods for discovering more about the regulatory networks in biofilm-forming species of bacteria.

Waste effluent has a described application as a media supplement for culturing fastidious organisms in the gut microbiome[4]. I used waste effluent in luciferase assays, but it could also be added to flask culture or to semisolid media to observe any phenotype changes that could occur in the presence of small molecules from the gut. As a media supplement, it's suggested concentration was 3%[4]. Luciferase assays could be performed with supplemented waste effluent at this concentration; however, the increase or time-shift of expression due to additional resources for bacterial growth is likely to change even more. This would obscure more subtle

changes to gene expression. One could characterize the small molecules in the waste effluent by fractionating and testing for an effect on gene expression and overall phenotype.

5.3 Conclusions

The methods described in this thesis for troubleshooting and optimizing parameters for ChIP-seq are useful for other researchers using ChIP-seq to find regulatory targets in other biofilm-forming bacterial species. I did not discover any significant peaks at CsgD-regulated regions; however, this could be due to epitopes that were not available for our monoclonal antibodies to bind *in vivo.* We identified antibody specificity as a key for successful ChIP experiments, and qPCR enrichment as an important confirmation prior to sequencing.

In the presence of chemostat media and waste effluent, I observed 5 main trends in expression of virulence- and persistence-associated genes. I observed that persistence-associated genes were not expressed at 37°C, virulence-associated gene expression dropped off early at 37°C, virulence-associated genes were expressed early at 28°C, persistence-associated genes at 28°C were expressed later and were time-shifted in the presence of waste effluent, and *csgD, sig38H4,* and *ssrA* exhibited multimodal expression at 37°C. Some differences in expression could be attributed to a nutritional increase provided by waste effluent and chemostat media. There may be other factors that lead to higher expression of some genes in the presence of waste effluent, but more research is needed to understand these differences.

We focused on the intrinsic changes to gene expression mediated by the *S.* Typhimurium master biofilm regulator, CsgD, using ChIP-seq, and on the extrinsic factors in the gut that could lead to differences in expression, using waste effluent containing products of the human gut microbiota. The signals for phenotype switching are complex and involve many different players in a coordinated effort to produce two cell types with different functions.

REFERENCES

1. **Ackermann, M.** A functional perspective on phenotypic heterogeneity in microorganisms. *Nature Reviews Microbiology* 13, 497-508, doi:10.1038/nrmicro3491 (2015).

2. **Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Grüning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A. & Goecks, J.** The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research* 44, W3-W10, doi:10.1093/nar/gkw343 (2016).

3. **Allen-Vercoe, E.** Bringing the gut microbiota into focus through microbial culture: recent progress and future perspective. *Current Opinion in Microbiology* 16, 625-629, doi:10.1016/j.mib.2013.09.008 (2013).

4. **Allen-Vercoe, E., McDonald, Julie.** Media Supplements and Methods to Culture Human Gastrointestinal Anaerobic Microorganisms. United States patent (2014).

5. **Andrews, N. C. & Schmidt, P. J**. Iron Homeostasis. *Annual Review of Physiology* 69, 69-85, doi:10.1146 (2007).

6. **Andrews, S. FastQC:** A Quality Control tool for High Throughput Sequence Data. (2010).

7. **Antunes, L. C., McDonald, J. A., Schroeter, K., Carlucci, C., Ferreira, R. B., Wang, M., Yurist-Doutsch, S., Hira, G., Jacobson, K., Davies, J., Allen-Vercoe, E. & Finlay, B. B.** Antivirulence activity of the human gut metabolome. *MBio* 5, e01183-01114, doi:10.1128/mBio.01183-14 (2014).

8. **Austin, J. W., Sanders, G., Kay, W. W. & Collinson, S. K.** Thin aggregative fimbriae enhance *Salmonella enteritidis* biofilm formation. *FEMS Microbiology Letters* 162, 295-301 (1998).

9. **Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., Ren, J., Li, W. W. & Noble, W. S.** MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37, W202-208, doi:10.1093/nar/gkp335 (2009).

10. **Barak, J. D., Jahn, C. E., Gibson, D. L. & Charkowski, A. O.** The Role of Cellulose and O-Antigen Capsule in the Colonization of Plants by *Salmonella enterica*. *Molecular Plant-Microbe Interactions* 20, 1083-1091 (2007).

11. **Barnhart, M. M. & Chapman, M. R.** Curli biogenesis and function. *Annual Review of Microbiology* 60, 131-147, doi:10.1146/annurev.micro.60.080805.142106 (2006).

12. **Bäumler, A. J., Tsolis, R. M., Heffron, F.** Contribution of fimbrial operons to attachment to and invasion of epithelial cell lines by *Salmonella* typhimurium. *Infection and Immunity* 64, 1862-1865 (1996).

13. **Bjarnason, J., Southward, C. M. & Surette, M. G.** Genomic Profiling of Iron-Responsive Genes in Salmonella enterica Serovar Typhimurium by High-Throughput Screening of a Random Promoter Library. *Journal of Bacteriology* 185, 4973-4982, doi:10.1128/jb.185.16.4973-4982.2003 (2003).

14. **Blasco, B., Chen, J. M., Hartkoorn, R., Sala, C., Uplekar, S., Rougemont, J., Pojer, F. & Cole, S. T.** Virulence regulator EspR of *Mycobacterium tuberculosis* is a nucleoid-associated protein. *PLoS Pathogens* 8, e1002621, doi:10.1371/journal.ppat.1002621 (2012).

15. **Bolger, A. M., Lohse, M. & Usadel, B.** Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

16. **Boyd, J. F.** Pathology of the alimentary tract in *Salmonella* typhimurium food poisoning. *Gut* 26, 935-944 (1985).

17. **Brown, P. K., Dozois, C. M., Nickerson, C. A., Zuppardo, A., Terlonge, J. & Curtiss, R.** MlrA, a novel regulator of curli (AgF) and extracellular matrix synthesis by *Escherichia coli* and *Salmonella enterica* serovar Typhimurium. *Molecular Microbiology* 41, 349-363 (2001).

18. **Cabeza, M. L., Aguirre, A., Soncini, F. C. & Vescovi, E. G**. Induction of RpoS degradation by the two-component system regulator RstA in *Salmonella enterica*. *Journal of Bacteriology* 189, 7335-7342, doi:10.1128/JB.00801-07 (2007).

19. **Government of Canada.** Pathogen Safety Data Sheets: Infectious Substances – *Salmonella enterica spp.* (2018).

20. **Chapman, M. R., Robinson, L. S., Pinkner, J. S., Roth, R., Heuser, J., Hammar, M., Normark, S. & Hultgren, S. J.** Role of *Escherichia coli* Curli Operons in Directing Amyloid Fiber Formation. *Science* 295, 851-855 (2002).

21. **Charles, R. C., Harris, J. B., Chase, M. R., Lebrun, L. M., Sheikh, A., LaRocque, R. C., Logvinenko, T., Rollins, S. M., Tarique, A., Hohmann, E. L., Rosenberg, I., Krastins, B., Sarracino, D. A., Qadri, F., Calderwood, S. B. & Ryan, E. T.** Comparative proteomic analysis of the PhoP regulon in *Salmonella enterica* serovar Typhi versus Typhimurium. *PLoS One* 4, e6994, doi:10.1371/journal.pone.0006994 (2009).

22. **Chiu, C. H., Su, L. H. & Chu, C.** Salmonella enterica Serotype Choleraesuis: Epidemiology, Pathogenesis, Clinical Disease, and Treatment. *Clinical Microbiology Reviews* 17, 311-322, doi:10.1128/cmr.17.2.311-322.2004 (2004).

23. **Clarke, D. J.** The Rcs phosphorelay: More than just a two-component pathway. *Future Microbiology* 5, 1173-1184 (2010).

24. **Close, D., Xu, T., Smartt, A., Rogers, A., Crossley, R., Price, S., Ripp, S. & Sayler, G.** The evolution of the bacterial luciferase gene cassette (*lux*) as a real-time bioreporter. *Sensors (Basel)* 12, 732-752, doi:10.3390/s120100732 (2012).

25. **Close, D. M., Ripp, S. & Sayler, G. S.** Reporter proteins in whole-cell optical bioreporter detection systems, biosensor integrations, and biosensing applications. *Sensors (Basel)* 9, 9147-9174, doi:10.3390/s91109147 (2009).

26. **Coburn, B., Grassl, G. A. & Finlay, B. B.** Salmonella, the host and disease: a brief review. *Immunology & Cell Biology* 85, 112-118, doi:10.1038/sj.icb.7100007 (2007).

27. **Cock, P. J., Fields, C. J., Goto, N., Heuer, M. L. & Rice, P. M.** The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Research* 38, 1767-1771, doi:10.1093/nar/gkp1137 (2010).

28. **Collinson, S. K., Clouthier, S. C., Doran, J. L., Kay, W. W.** *Salmonella enteritidis agfBAC* operon encoding thin, aggregative fimbriae. *Journal of Bacteriology* 178, 662-667 (1996).

29. **Collinson, S. K., Emödy, L., Müller, K. H., Trust, T. J., Kay, W. W.** Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. *Journal of Bacteriology* 173, 4773–4781 (1991).

30. **Collinson, S. P., J. M. R, Hodges, R. S. & Kay, W. W.** Structural predictions of AgfA, the insoluble fimbrial subunit of *Salmonella* thin aggregative fimbriae. *Journal of Molecular Biology* 290, 741-756 (1999).

31. **The ENCODE Project Consortium.** The ENCODE (ENCyclopedia of DNA Elements) Project. *Science* 306, 636-640, doi:10.1126/science.1105136 (2004).

32. **Costerton, J. W., Stewart, Philip S., Greenberg, E.P.** Bacterial Biofilms: A Common Cause of Persistent Infections. *Science* 284, 1318-1322 (1999).

33. **Costerton, W. J., Lewandowski, Zbigniew, Caldwell, Douglas E., Korber, Darren R. Lappin-Scott, Hilary M.** Microbial Biofilms. *Annual Review of Microbiology* 49, 711-745, doi:10.1146/annurev.mi.49.100195.003431 (1995).

34. **Crack, J. C., Munnoch, J., Dodd, E. L., Knowles, F., Al Bassam, M. M., Kamali, S., Holland, A. A., Cramer, S. P., Hamilton, C. J., Johnson, M. K., Thomson, A. J.,**

Hutchings, M. I. & Le Brun, N. E. NsrR from Streptomyces coelicolor is a nitric oxide-sensing [4Fe-4S] cluster protein with a specialized regulatory function. *J Biol Chem* 290, 12689-12704, doi:10.1074/jbc.M115.643072 (2015).

35. **Crawford, R. W., Gibson, D. L., Kay, W. W. & Gunn, J. S.** Identification of a bile-induced exopolysaccharide required for *Salmonella* biofilm formation on gallstone surfaces. *Infection and Immunity* 76, 5341-5349, doi:10.1128/IAI.00786-08 (2008).

36. **Crosa, J. H. B., D.J., Ewing, W. H., Falkow, S.** Molecular Relationships Among the Salmonelleae. *Journal of Bacteriology* 115, 307-315 (1973).

37. **Davies, B. W., Bogard, Ryan W., Mekalanos, John J.** Mapping the regulon of *Vibrio cholerae* ferric uptake regulator expands its known network of gene regulation. *PNAS* 108, 12467-12472 (2011).

38. **Deriu, E., Liu, J. Z., Pezeshki, M., Edwards, R. A., Ochoa, R. J., Contreras, H., Libby, S. J., Fang, F. C. & Raffatellu, M.** Probiotic bacteria reduce salmonella typhimurium intestinal colonization by competing for iron. *Cell Host & Microbe* 14, 26-37, doi:10.1016/j.chom.2013.06.007 (2013).

39. **Desai, P. T., Porwollik, S., Long, F., Cheng, P., Wollam, A., Bhonagiri-Palsikar, V., Hallsworth-Pepin, K., Clifton, S. W., Weinstock, G. M. & McClelland, M.** Evolutionary Genomics of *Salmonella enterica* Subspecies. *MBio* 4, doi:10.1128/mBio.00579-12 (2013).

40. **Dieckmann, R. & Malorny, B**. Rapid screening of epidemiologically important *Salmonella enterica* subsp. enterica serovars by whole-cell matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Applied and Environmental Microbiology* 77, 4136-4146, doi:10.1128/AEM.02418-10 (2011).

41. **Dillon, S. C., Cameron, A. D., Hokamp, K., Lucchini, S., Hinton, J. C. & Dorman, C. J.** Genome-wide analysis of the H-NS and Sfh regulatory networks in *Salmonella* Typhimurium identifies a plasmid-encoded transcription silencing mechanism. *Molecular Microbiology* 76, 1250-1265, doi:10.1111/j.1365-2958.2010.07173.x (2010).

42. **Dong, T. G. & Mekalanos, J. J.** Characterization of the RpoN regulon reveals differential regulation of T6SS and new flagellar operons in *Vibrio cholerae* O37 strain V52. *Nucleic Acids Research* 40, 7766-7775, doi:10.1093/nar/gks567 (2012).

43. **Dorel, C., Lejeune, P. & Rodrigue, A.** The Cpx system of *Escherichia coli*, a strategic signaling pathway for confronting adverse conditions and for settling biofilm communities? *Research in Microbiology* 157, 306-314, doi:10.1016/j.resmic.2005.12.003 (2006).

44. **Dougan, G., John, Victoria, Palmer, Sophie, Mastroeni, Pietro.** Immunity to salmonellosis. *Immunological Reviews* 240, 196-210 (2011).

45. **Elsner, H. I. & Lindblad, E. B.** Ultrasonic Degradation of DNA. *DNA* 8, doi:10.1089/dna.1989.8.697 (1989).

46. **Eng, S.-K., Pusparajah, P., Ab Mutalib, N.-S., Ser, H.-L., Chan, K.-G. & Lee, L.-H.** *Salmonella*: A review on pathogenesis, epidemiology and antibiotic resistance. *Frontiers in Life Science* 8, 284-293, doi:10.1080/21553769.2015.1051243 (2015).

47. **Engebrecht, J., Simon, M., Silverman, M**. Measuring gene expression with light. *Science* 227 (1985).

48. **Epigenie.** Getting your Fix: Optimizing Chromatin Fixation for ChIP Analysis. (2012). http://epigenie.com/getting-your-fix-optimizing-chromatin-fixation-for-chip-analysis/

49. **Feasey, N. A., Dougan, G., Kingsley, R. A., Heyderman, R. S. & Gordon, M. A**. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. *The Lancet* 379, 2489-2499, doi:10.1016/s0140-6736(11)61752-2 (2012).

50. **Fitzgerald, D. M., Bonocora, R. P. & Wade, J. T.** Comprehensive mapping of the *Escherichia coli* flagellar regulatory network. *PLoS Genetics* 10, e1004649, doi:10.1371/journal.pgen.1004649 (2014).

51. **Fluit, A. C**. Towards more virulent and antibiotic-resistant *Salmonella*? *FEMS Immunology and Medical Microbiology* 43, 1-11, doi:10.1016/j.femsim.2004.10.007 (2005).

52. **Fuqua, W. C., Winans, S. C., & Greenberg, E. P**. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of Bacteriology* 176, 269-275 (1994).

53. **Garcia Vescovi, E., Soncini, F. C., Groisman, E. A.** Mg2+ as an extracellular signal: Environmental regulation of *Salmonella* virulence. *Cell* 84, 165-174 (1996).

54. **Geertz, M. & Maerkl, S. J.** Experimental strategies for studying transcription factor-DNA binding specificities. *Briefings in Functional Genomics* 9, 362-373, doi:10.1093/bfgp/elq023 (2010).

55. **Gerstel, U., Kolb, A. & Romling, U.** Regulatory components at the *csgD* promoter--additional roles for OmpR and integration host factor and role of the 5' untranslated region. *FEMS Microbiology Letters* 261, 109-117, doi:10.1111/j.1574-6968.2006.00332.x (2006).

56. **Gerstel, U., Park, C. & Römling, U.** Complex regulation of *csgD* promoter activity by global regulatory proteins. *Molecular Microbiology* 49, 639-654, doi:10.1046/j.1365-2958.2003.03594.x (2003).

57. **Gibson, D. L., White, A. P., Snyder, S. D., Martin, S., Heiss, C., Azadi, P., Surette, M. & Kay, W. W.** *Salmonella* produces an O-antigen capsule regulated by AgfD and important for environmental persistence. *Journal of Bacteriology* 188, 7722-7730, doi:10.1128/JB.00809-06 (2006).

58. **Gill, A.** Cross-linking chromatin immunoprecipitation (X-ChIP) protocol. (2017). http://docs.abcam.com/pdf/protocols/X-ChIP_protocol.pdf

59. **Gorvel J. P., M. S.** Maturation steps of the *Salmonella*-containing vacuole. *Microbes and Infection* 3 (2001).

60. **Gossen, M., Bujard, H.** Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences* 89, 5547-4451 (1992).

61. **Grantcharova, N., Peters, V., Monteiro, C., Zakikhany, K. & Romling, U**. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar typhimurium. *Journal of Bacteriology* 192, 456-466, doi:10.1128/JB.01826-08 (2010).

62. **Grassl, G. A. & Finlay, B. B.** Pathogenesis of enteric Salmonella infections. *Current Opinion in Gastroenterology* 24, 22-26 (2008).

63. **Greer, L. F., 3rd & Szalay, A. A**. Imaging of light emission from the expression of luciferases in living cells and organisms: a review. *Luminescence* 17, 43-74, doi:10.1002/bio.676 (2002).

64. **Grimont, P. A., Weill, Francois-Xavier.** (ed WHO Collaborating Center for Reference and Research on *Salmonella*) (World Health Organization, 2007).

65. **Gu, D., Liu, H., Yang, Z., Zhang, Y. & Wang, Q.** Chromatin Immunoprecipitation Sequencing Technology Reveals Global Regulatory Roles of Low-Cell-Density Quorum-Sensing Regulator AphA in the Pathogen *Vibrio alginolyticus*. *Journal of Bacteriology* 198, 2985-2999, doi:10.1128/JB.00520-16 (2016).

66. **Guinane, C. M., Cotter, Paul D.** Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therapeutic advances in Gastroenterology* 6, 295–308, doi:10.1177/ 1756283X13482996 (2013).

67. **Gupta, R., Patterson, S., Ripp, S., Simpson, M. & Sayler, G.** Expression of the *Photorhabdus luminescens lux* genes (*luxA*, *B, C, D, and E*) in *Saccharomyces cerevisiae*. *FEMS Yeast Research* 4, 305-313, doi:10.1016/s1567-1356(03)00174-0 (2003).

68. **Haraga, A., Ohlson, M. B. & Miller, S. I.** Salmonellae interplay with host cells. *Nature Reviews Microbiology* 6, 53-66, doi:10.1038/nrmicro1788 (2008).

69. **Haycocks, J. R., Sharma, P., Stringer, A. M., Wade, J. T. & Grainger, D. C.** The molecular basis for control of ETEC enterotoxin expression in response to environment and host. *PLoS Pathogens* 11, e1004605, doi:10.1371/journal.ppat.1004605 (2015).

70. **Head, S. R., Komori, H. K., LaMere, S. A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D. R. & Ordoukhanian, P.** Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* 56, 6168, doi:10.2144/000114133 (2014).

71. **Hengge-Aronis, R.** Signal Transduction and Regulatory Mechanisms Involved in Control of the S (RpoS) Subunit of RNA Polymerase. *Microbiology and Molecular Biology Reviews* 66, 373-395, doi:10.1128/mmbr.66.3.373-395.2002 (2002).

72. **Herwald H, M. M., Olsen A, Rhen M, Dahlback B, Muller-Esterl, W, Bjorck, L.** Activation of the contact-phase system on bacterial surfaces—a clue to serious complications in infectious diseases. *Nature Methods* 4, 298-302 (1998).

73. **Hill P. J., S. G. S.** Use of *lux* genes in applied biochemistry. *Journal of Bioluminescence and Chemiluminescence* 9, 211-215, doi:10.1002/bio.1170090315 (1994).

74. **Hoffman M, E. J., Wilder S, Kundaje A, Harris R, Libbrecht M et al.** Integrative annotation of chromatin elements from ENCODE data. *Nucleic Acids Research* 41, 827-841 (2012).

74. **Hoffman, Brad G. & Jones, Steven J. M.** Genome-wide identification of DNA-protein interactions using chromatin immunoprecipitation coupled with flow cell sequencing. *Journal of Endocrinology* 201, 1-13 (2009).

75. **Hoffmann, S., Batz, M. B. & Morris, J. G., Jr.** Annual cost of illness and quality-adjusted life year losses in the United States due to 14 foodborne pathogens. *Journal of Food Protection* 75, 1292-1302, doi:10.4315/0362-028X.JFP-11-417 (2012).

76. **Hohmann, E. L.** Nontyphoidal salmonellosis. *Clinical Infectious Disease* 15, 263-269 (2001).

77. **Holmqvist, E., Reimegard, J., Sterk, M., Grantcharova, N., Romling, U. & Wagner, E. G.** Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis. *EMBO Journal* 29, 1840-1850, doi:10.1038/emboj.2010.73 (2010).

78. **Illumina**. Estimating Sequencing Coverage. (2014).

79. **Imam, S., Noguera, Daniel R., Donohue, Timothy J.** Global Analysis of Photosynthesis Transcriptional Regulatory Networks. *PLoS Genetics* 10 (2014).

80. **Iyer, V. R., Horak, Christine E., Scafe, Charles S., Botstein, David, Snyder, Michael, Brown, Patrick O.** Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* 409, doi:10.1038/35054095.

81. **Jarvik, T., Smillie, C., Groisman, E. A. & Ochman, H.** Short-term signatures of evolutionary change in the *Salmonella enterica* serovar typhimurium 14028 genome. *Journal of Bacteriology* 192, 560-567, doi:10.1128/JB.01233-09 (2010).

82. **Jepson, M. A., Collares-Buzato, C. B., Clark, M. A., Hirst, B. H., Simmons, N. L.** Rapid disruption of epithelial barrier function by *Salmonella* typhimurium is associated with structural modification of intercellular junctions. *Infection and Immunity* 63, 356-359 (1995).

83. **Ji, H., Jiang, H., Ma, W., Johnson, D. S., Myers, R. M. & Wong, W. H.** An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nature Biotechnology* 26, 1293-1300, doi:10.1038/nbt.1505 (2008).

84. **Jones B. D., G. N., Falkow S.** *Salmonella* typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *Journal of Experimental Medicine* 180, 15-23 (1994).

85. **Jones, C. J., Newsom, D., Kelly, B., Irie, Y., Jennings, L. K., Xu, B., Limoli, D. H., Harrison, J. J., Parsek, M. R., White, P. & Wozniak, D. J.** ChIP-Seq and RNA-Seq reveal an AmrZ-mediated mechanism for cyclic di-GMP synthesis and biofilm development by *Pseudomonas aeruginosa*. *PLoS Pathogens* 10, e1003984, doi:10.1371/journal.ppat.1003984 (2014).

86. **Kahramanoglou, C., Seshasayee, A. S., Prieto, A. I., Ibberson, D., Schmidt, S., Zimmermann, J., Benes, V., Fraser, G. M. & Luscombe, N. M.** Direct and indirect effects of H-NS and Fis on global gene expression control in *Escherichia coli*. *Nucleic Acids Research* 39, 2073-2091, doi:10.1093/nar/gkq934 (2011).

87. **Kato, A. & Groisman, E. A.** The PhoQ/PhoP regulatory network of *Salmonella enterica*. *Advances in Experimental Medicine and Biology* 631, 7-21 (2008).

88. **Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A.** Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647-1649 (2012).

89. **Kharchenko P, T. M., Park, P.** Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nature Biotechnology* 26, 1351-1359 (2008).

90. **Kim, N.-K., Jayatillake, R. V., & Spouge, J. L.** NEXT-peak: a normal-exponential two-peak model for peak-calling in ChIP-seq data. *BMC Genomics* 14, doi:10.1186/1471-2164-14-349 (2013).

91. **Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D. & Ren, B.** A high-resolution map of active promoters in the human genome. *Nature* 436, 876-880, doi:10.1038/nature03877 (2005).

92. **Kirk, M. D., Pires, S. M., Black, R. E., Caipo, M., Crump, J. A., Devleesschauwer, B., Dopfer, D., Fazil, A., Fischer-Walker, C. L., Hald, T., Hall, A. J., Keddy, K. H., Lake, R. J., Lanata, C. F., Torgerson, P. R., Havelaar, A. H. & Angulo, F. J.** World Health Organization Estimates of the Global and Regional Disease Burden of 22 Foodborne Bacterial, Protozoal, and Viral Diseases, 2010: A Data Synthesis. *PLoS Medicine* 12, e1001921, doi:10.1371/journal.pmed.1001921 (2015).

93. **Knierim, E., Lucke, B., Schwarz, J. M., Schuelke, M. & Seelow, D.** Systematic comparison of three methods for fragmentation of long-range PCR products for next generation sequencing. *PLoS One* 6, e28240, doi:10.1371/journal.pone.0028240 (2011).

94. **Korbel, J. O., Urban, A. E., Affourtit, J. P., Godwin, B., Grubert, F., Simons, J. F., Kim, P. M., Palejev, D., Carriero, N. J., Du, L., Taillon, B. E., Chen, Z., Tanzer, A., Saunders, A. C., Chi, J., Yang, F., Carter, N. P., Hurles, M. E., Weissman, S. M., Harkins, T. T., Gerstein, M. B., Egholm, M. & Snyder, M.** Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318, 420-426, doi:10.1126/science.1149504 (2007).

95. **Kundaje, A.** A comprehensive collection of signal artifact blacklist regions in the human genome. *ENCODE* (2013).

96. **Kuvandik C., K. I., Namiduru M., Baydar I.** Predictive value of clinical and laboratory findings in the diagnosis of the enteric fever. *New Microbiologica* 32, 25-30 (2009).

97. **Lagier, J. C., Armougom, F., Million, M., Hugon, P., Pagnier, I., Robert, C., Bittar, F., Fournous, G., Gimenez, G., Maraninchi, M., Trape, J. F., Koonin, E. V., La Scola, B. & Raoult, D.** Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical Microbiology and Infection* 18, 1185-1193, doi:10.1111/1469-0691.12023 (2012).

98. **Landt, S. G. *et al.*** ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Research* 22, 1813-1831, doi:10.1101/gr.136184.111 (2012).

99. **Langmead, B. & Salzberg, S. L.** Fast gapped-read alignment with Bowtie 2. *Nature Methods* 9, 357-359, doi:10.1038/nmeth.1923 (2012).

100. **Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L.** Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology* 10, doi:10.1186/gb-2009-10-3-r25 (2009).

101. **Latasa, C., Roux, A., Toledo-Arana, A., Ghigo, J. M., Gamazo, C., Penades, J. R. & Lasa, I.** BapA, a large secreted protein required for biofilm formation and host colonization of *Salmonella enterica* serovar Enteritidis. *Molecular Microbiology* 58, 1322-1339, doi:10.1111/j.1365-2958.2005.04907.x (2005).

102. **Le Minor, L., Popoff, Michel Y.** Designation of Salmonella enterica sp. nov., norn. rev., as the Type and Only Species of the Genus *Salmonella*. *International Journal of Systematic Bacteriology*, 465-468 (1987).

103. **Ledeboer, N. A. & Jones, B. D.** Exopolysaccharide sugars contribute to biofilm formation by *Salmonella enterica* serovar typhimurium on HEp-2 cells and chicken intestinal epithelium. *Journal of Bacteriology* 187, 3214-3226, doi:10.1128/JB.187.9.3214-3226.2005 (2005).

104. **Levine, M.** Normal mineral homeostasis. Interplay of parathyroid hormone and vitamin D. *Endocrine Development* 6, 14-33 (2003).

105. **Li, H. & Durbin, R**. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760, doi:10.1093/bioinformatics/btp324 (2009).

106. **Li, Q., Brown, J. B., Huang, H. & Bickel, P. J.** Measuring reproducibility of high-throughput experiments. *The Annals of Applied Statistics* 5, 1752-1779, doi:10.1214/11-aoas466 (2011).

107. **Lieb, J. D.** in *Functional Genomics* Vol. 224 *Methods in Molecular Biology* (ed Khodursky A.B. Brownstein M.J.) 99-109 (Humana Press, 2003).

108. **Lin-Hui Su, M. C.-H. C.** Salmonella: Clinical importance and Evolution of Nomenclature. *Chang Gung Medical Journal* 30 (2007).

109. **Livak, K. J. & Schmittgen, T. D.** Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\Delta\Delta CT}$ Method. *Methods* 25, 402-408, doi:10.1006/meth.2001.1262 (2001).

110. **MacKenzie, K. D., Palmer, M. B., Koster, W. L. & White, A. P.** Examining the Link between Biofilm Formation and the Ability of Pathogenic *Salmonella* Strains to Colonize Multiple Host Species. *Frontiers in Veterinary Science* 4, 138, doi:10.3389/fvets.2017.00138 (2017).

111. **MacKenzie, K. D., Wang, Y., Shivak, D. J., Wong, C. S., Hoffman, L. J., Lam, S., Kroger, C., Cameron, A. D., Townsend, H. G., Koster, W. & White, A. P.** Bistable

expression of CsgD in *Salmonella enterica* serovar Typhimurium connects virulence to persistence. *Infection and Immunity* 83, 2312-2326, doi:10.1128/IAI.00137-15 (2015).

112. **Majdalani, N. & Gottesman, S.** The Rcs phosphorelay: A complex signal transduction system. Annual Review of Microbiology, 59, 379−405. *Annual Review of Microbiology* 59, 379-405, doi:10.1146/ (2005).

113. **Marx, V.** Finding the right antibody for the job. *Nature Methods* 10, 703-707, doi:10.1038/nmeth.2570 (2013).

114. **McDonald, J. A., Fuentes, S., Schroeter, K., Heikamp-deJong, I., Khursigara, C. M., de Vos, W. M. & Allen-Vercoe, E.** Simulating distal gut mucosal and luminal communities using packed-column biofilm reactors and an *in vitro* chemostat model. *Journal of Microbiological Methods* 108, 36-44, doi:10.1016/j.mimet.2014.11.007 (2015).

115. **McDonald, J. A., Schroeter, K., Fuentes, S., Heikamp-Dejong, I., Khursigara, C. M., de Vos, W. M. & Allen-Vercoe, E.** Evaluation of microbial community reproducibility, stability and composition in a human distal gut chemostat model. *Journal of Microbiological Methods* 95, 167-174, doi:10.1016/j.mimet.2013.08.008 (2013).

116. **McElroy, W. D., Hastings, J. W.** The Requirement of Riboflavin Phosphate for Bacterial Luminescence. *Science* 118, 385-386, doi:10.1126/science.118.3066.385 (1953).

117. **Meighen, E. A.** Bacterial bioluminescence: organization, regulation, and application of the *lux* genes. *Federation of American Societies for Experimental Biology* 7 1016-1022 (1993).

118. **Minch, K. J., Rustad, T. R., Peterson, E. J., Winkler, J., Reiss, D. J., Ma, S., Hickey, M., Brabant, W., Morrison, B., Turkarslan, S., Mawhinney, C., Galagan, J. E., Price, N. D., Baliga, N. S. & Sherman, D. R.** The DNA-binding network of *Mycobacterium tuberculosis*. *Nature Communications* 6, 5829, doi:10.1038/ncomms6829 (2015).

119. **Moest, T. P. & Meresse, S.** Salmonella T3SSs: successful mission of the secret(ion) agents. *Current Opinion in Microbiology* 16, 38-44, doi:10.1016/j.mib.2012.11.006 (2013).

120. **Morgan, J. L., McNamara, J. T. & Zimmer, J.** Mechanism of activation of bacterial cellulose synthase by cyclic di-GMP. *Nature Structural and Molecular Biology* 21, 489-496, doi:10.1038/nsmb.2803 (2014).

121. **Myers, K. S., Park, D. M., Beauchene, N. A. & Kiley, P. J.** Defining bacterial regulons using ChIP-seq. *Methods* 86, 80-88, doi:10.1016/j.ymeth.2015.05.022 (2015).

122. **Myers, K. S., Yan, H., Ong, I. M., Chung, D., Liang, K., Tran, F., Keles, S., Landick, R. & Kiley, P. J.** Genome-scale analysis of escherichia coli FNR reveals complex features of transcription factor binding. *PLoS Genet* 9, e1003565, doi:10.1371/journal.pgen.1003565 (2013).

123. **Nakoneczna I., H. H. S.** The comparative histopathology of primary and secondary lesions in murine salmonellosis. *British Journal of Experimental Pathology* 61, 76-84 (1980).

124. **Nealson, K. H., & Hastings, J. W.** Bacterial bioluminescence: its control and ecological significance. *Microbiological Reviews* 43, 496-518 (1979).

125. **O'Connell-Rodwell, C. E., Burns, Stacy M., Bachmann, Michael H., Contag, Christopher H.** Bioluminescent indicators for *in vivo* measurements of gene expression. *Trends in Biotechnology* 20, pS19-S23, doi:10.1016/S0167-7799(02)02001-2 (2002).

126. **O'Neill, L.** Immunoprecipitation of native chromatin: NChIP. *Methods* 31, 76-82, doi:10.1016/s1046-2023(03)00090-2 (2003).

127. **Ogasawara, H., Hasegawa, A., Kanda, E., Miki, T., Yamamoto, K. & Ishihama, A.** Genomic SELEX search for target promoters under the control of the PhoQP-RstBA signal relay cascade. *Journal of Bacteriology* 189, 4791-4799, doi:10.1128/JB.00319-07 (2007).

128. **Olsen A, A. A., Hammar M, Sukupolvi S, Normark S.** The RpoS sigma factor relieves H-NS-mediated transcriptional repression of *csgA*, the subunit gene of fibronectin-binding curli in *Escherichia coli*. *Molecular Microbiology* 7, 523-536 (1993).

129. **Olsen A, J. A., Normark S.** Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. *Nature* 338, 652-655 (1989).

130. **Olsen A, W. M., Morgelin M, Bjorck L.** 1998. Infect. & 66:944–49, I. Curli, fibrous surface proteins of *Escherichia coli*, interact with major histocompatibility complex class I molecules. *Infection and Immunity* 66, 944-949 (1998).

131. **Olsson, O., Escher, Alan, Sandberg, Göran, Schell, Jeff, Koncz, Csaba, Szalay, Aladar A.** Engineering of monomeric bacterial luciferases by fusion of *luxA* and *luxB* genes in *Vibrio harveyi*. *Gene* 81, 335-347, doi:10.1016/0378-1119(89)90194-7 (1989).

132. **Park, D. M., Akhtar, M. S., Ansari, A. Z., Landick, R. & Kiley, P. J.** The bacterial response regulator ArcA uses a diverse binding site architecture to regulate carbon oxidation globally. *PLoS Genetics* 9, e1003839, doi:10.1371/journal.pgen.1003839 (2013).

133. **Park, P. J.** ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics* 10, 669-680, doi:10.1038/nrg2641 (2009).

134. **Peano, C., Wolf, J., Demol, J., Rossi, E., Petiti, L., De Bellis, G., Geiselmann, J., Egli, T., Lacour, S. & Landini, P.** Characterization of the *Escherichia coli* sigma(S) core regulon by Chromatin Immunoprecipitation-sequencing (ChIP-seq) analysis. *Scientific Reports* 5, 10469, doi:10.1038/srep10469 (2015).

135. **Pepke, S., Wold, B. & Mortazavi, A.** Computation for ChIP-seq and RNA-seq studies. *Nature Methods* 6, S22-32, doi:10.1038/nmeth.1371 (2009).

136. **Perkins, T. T., Davies, M. R., Klemm, E. J., Rowley, G., Wileman, T., James, K., Keane, T., Maskell, D., Hinton, J. C., Dougan, G. & Kingsley, R. A.** ChIP-seq and transcriptome analysis of the OmpR regulon of Salmonella enterica serovars Typhi and Typhimurium reveals accessory genes implicated in host colonization. *Molecular Microbiology* 87, 526-538, doi:10.1111/mmi.12111 (2013).

137. **Petrone, B. L., Stringer, Anne M., Wadea, Joseph T.** Identification of HilD-Regulated Genes in *Salmonella enterica* Serovar Typhimurium. *Journal of Bacteriology* 196, 1094-1101 (2014).

138. **Prieto, A. I., Kahramanoglou, C., Ali, R. M., Fraser, G. M., Seshasayee, A. S. & Luscombe, N. M.** Genomic analysis of DNA binding and gene regulation by homologous nucleoid-associated proteins IHF and HU in *Escherichia coli* K12. *Nucleic Acids Research* 40, 3524-3537, doi:10.1093/nar/gkr1236 (2012).

139. **Prouty, A. M. & Gunn, J. S.** Comparative Analysis of *Salmonella enterica* Serovar Typhimurium Biofilm Formation on Gallstones and on Glass. *Infection and Immunity* 71, 7154-7158, doi:10.1128/iai.71.12.7154-7158.2003 (2003).

140. **Ramsey, C. H., Edwards, P. R.** Resistance of Salmonellae Isolated in 1959 and 1960 to Tetracyclines and Chloramphenicol. *Applied and Environmental Microbiology* 9, 389-391 (1961).

141. **Reeves, M. W., Evins, G. M., Heiba, A. A., Plikaytis, B. D. & Farmer, J. J. I.** Clonal nature of Salmonella typhi and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis and proposal of *Salmonella bongori* comb. nov. *Journal of Clinical Microbiology* 27, 313-320 (1989).

142. **Reichhardt, C., McCrate, O. A., Zhou, X., Lee, J., Thongsomboon, W. & Cegelski, L.** Influence of the amyloid dye Congo red on curli, cellulose, and the extracellular matrix in E. coli during growth and matrix purification. *Analytical and Bioanalytical Chemistry* 408, 7709-7717, doi:10.1007/s00216-016-9868-2 (2016).

143. **Ren, R., Wyrick, Aparicio, Jennings, Simon, Young.** Genome-wide location and function of DNA binding proteins. *Science* 290, 2306-2309 (2000).

144. **Rescigno M., U. M., Valzasina B., Francolini M., Rotta G., Bonasio R. et al.** Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature Immunology* 2, 361-367 (2001).

145. **Richter-Dahlfors A., B. A. M., Finlay B. B.** Murine salmonellosis studied by confocal microscopy: *Salmonella* typhimurium resides intracellularly inside macrophages and exerts a cytotoxic effect on phagocytes in vivo. *Journal of Experimental Medicine* 186, 569-580 (1997).

146. **Robbe-Saule, V., Carreira, I., Kolb, A. & Norel, F.** Effect of growth temperature on Crl-dependent regulation of sigmaS activity in *Salmonella enterica* serovar Typhimurium. *Journal of Bacteriology* 190, 4453-4459, doi:10.1128/JB.00154-08 (2008).

147. **Robbe-Saule, V., Jaumouille, V., Prevost, M. C., Guadagnini, S., Talhouarne, C., Mathout, H., Kolb, A. & Norel, F.** Crl activates transcription initiation of RpoS-regulated genes involved in the multicellular behavior of *Salmonella enterica* serovar Typhimurium. *Journal of Bacteriology* 188, 3983-3994, doi:10.1128/JB.00033-06 (2006).

148. **Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O. L., He, A., Marra, M., Snyder, M. & Jones, S.** Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods* 4, 651-657, doi:10.1038/nmeth1068 (2007).

149. **Rolfe, R. D.** Role of Volatile Fatty Acids in Colonization Resistance to *Clostridium dificile*. *Infection and Immunity* 45, 185-191 (1984).

150. **Romling, U., Bian, Z., Hammar, M., Sierralta, W. D. & Normark, S.** Curli Fibers Are Highly Conserved between *Salmonella* typhimurium and *Escherichia coli* with Respect to Operon Structure and Regulation. *Journal of Bacteriology* (1998).

151. **Romling, U., Bokranz, W., Rabsch, W., Zogaj, X., Nimtz, M. & Tschape, H.** Occurrence and regulation of the multicellular morphotype in *Salmonella* serovars important in human disease. *International Journal of Food Microbiology* 293, 273-285, doi:10.1078/1438-4221-00268 (2003).

152. **Romling, U. & Galperin, M. Y.** Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. *Trends in Microbiology* 23, 545-557, doi:10.1016/j.tim.2015.05.005 (2015).

153. **Römling, U. & Rohde, M.** Flagella modulate the multicellular behavior of *Salmonella* typhimurium on the community level. *FEMS Microbiology Letters* 180, 91-102 (1999).

154. **Romling, U., Rohde, M., Olsen, A., Normark, S. & Reinkoster, J.** AgfD, the checkpoint of multicellular and aggregative behaviour in *Salmonella* typhimurium regulates at least two independent pathways. *Molecular Microbiology* 36, 10-23 (2000).

155. **Romling, U., Sierralta, W. D., Eriksson, K. & Normark, S.** Multicellular and aggregative behaviour of *Salmonella* typhimurium strains is controlled by mutations in the *agfD* promoter. *Molecular Microbiology* 28, 249-264 (1998).

156. **Ryan, M. P., O'Dwyer, J. & Adley, C. C.** Evaluation of the Complex Nomenclature of the Clinically and Veterinary Significant Pathogen *Salmonella*. *BioMed Research International* 2017, doi:10.1155/2017/3782182 (2017).

157. **Sadikot, R. T. & Blackwell, T. S**. Bioluminescence imaging. *Proceedings of the American Thoracic Society* 2, 537-540, 511-532, doi:10.1513/pats.200507-067DS (2005).

158. **Schmiedeberg, L., Skene, P., Deaton, A. & Bird, A.** A temporal threshold for formaldehyde crosslinking and fixation. *PLoS One* 4, e4636, doi:10.1371/journal.pone.0004636 (2009).

159. **Schoppee Bortz, P. D. & Wamhoff, B. R.** Chromatin immunoprecipitation (ChIP): revisiting the efficacy of sample preparation, sonication, quantification of sheared DNA, and analysis via PCR. *PLoS One* 6, e26015, doi:10.1371/journal.pone.0026015 (2011).

160. **Sengupta, M., Jain, V., Wilkinson, B. J. & Jayaswal, R. K.** Chromatin immunoprecipitation identifies genes under direct VraSR regulation in *Staphylococcus aureus*. *Canadian Journal of Microbiology* 58, 703-708, doi:10.1139/w2012-043 (2012).

161. **USDA Economic Research Service.** Cost Estimates of Foodborne Illnesses (2014). http://ers.usda.gov/data-products/cost-estimates-of-foodborne-illnesses.aspx.

162. **Simm, R., Morr, M., Kader, A., Nimtz, M. & Romling, U.** GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Molecular Microbiology* 53, 1123-1134, doi:10.1111/j.1365-2958.2004.04206.x (2004).

163. **Simpson, M. L., Sayler, Gary S., Fleming, James T., Applegate, Bruce.** Whole-cell biocomputing 19 (2001).

164. **Singh, S. S., Singh, N., Bonocora, R. P., Fitzgerald, D. M., Wade, J. T. & Grainger, D. C.** Widespread suppression of intragenic transcription initiation by H-NS. *Genes & Development* 28, 214-219, doi:10.1101/gad.234336.113 (2014).

165. **Snyder, D. S., Gibson, D., Heiss, C., Kay, W. & Azadi, P.** Structure of a capsular polysaccharide isolated from *Salmonella* enteritidis. *Carbohydrate Research* 341, 2388-2397, doi:10.1016/j.carres.2006.06.010 (2006).

166. **Solano, C., García, Begoña, Valle, Jaione, Berasain, Carmen, Ghigo, Jean-Marc, Gamazo, Carlos, Lasa, Iñigo.** Genetic analysis of *Salmonella enteritidis* biofilm formation: critical role of cellulose. *Molecular Microbiology* 43, doi:10.1046/j.1365-2958.2002.02802.x (2002).

167. **Solans, L., Gonzalo-Asensio, J., Sala, C., Benjak, A., Uplekar, S., Rougemont, J., Guilhot, C., Malaga, W., Martin, C. & Cole, S. T.** The PhoP-dependent ncRNA Mcr7 modulates the TAT secretion system in *Mycobacterium tuberculosis*. *PLoS Pathogens* 10, e1004183, doi:10.1371/journal.ppat.1004183 (2014).

168. **Solomon, L. & Varshavsky, A.** Mapping protein-DNA interactions in vivo with formaldehyde: Evidence that histone H4 is retained on a highly transcribed gene. *Cell* 53, 937-947 (1988).

169. **Solomon, M. J. & Varshavsky, A.** Formaldehyde-mediated DNA-protein cross-linking: a probe for in vivo chromatin structures. *Proceedings of the National Academy of Sciences* 82, 6470-6474 (1985).

170. **Stecher, B. & Hardt, W. D.** Mechanisms controlling pathogen colonization of the gut. *Current Opinion in Microbiology* 14, 82-91, doi:10.1016/j.mib.2010.10.003 (2011).

171. **Steenackers, H., Hermans, K., Vanderleyden, J. & De Keersmaecker, Sigrid C. J.** Salmonella biofilms: An overview on occurrence, structure, regulation and eradication. *Food Research International* 45, 502-531, doi:10.1016/j.foodres.2011.01.038 (2012).

172. **Strehler, B. L., Harvey, E. N., Chang, J. J., Cormier, M. J.** The luminescent oxidation of reduced riboflavin or reduced riboflavin phosphate in the bacterial luciferin-luciferase reaction. *Proceedings of the National Academy of Sciences USA* 40, 10-12 (1954).

173. **Su, L. H., Chiu, C.H.** Salmonella: clinical importance and evolution of nomenclature. *Chang Gung Medical Journal* 30, 210-219 (2007).

174. **Integrated Taxonomic Information System (ITIS).** *Salmonella* Lignieres, 1900. (National Museum of Natural History, USA, 2012).

175. **Tauxe, R. V., Doyle, M. P., Kuchenmuller, T., Schlundt, J. & Stein, C. E**. Evolving public health approaches to the global challenge of foodborne infections. *International Journal of Food Microbiology* 139, S16-28, doi:10.1016/j.ijfoodmicro.2009.10.014 (2010).

176. **Tennant, S. M., MacLennan, C. A., Simon, R., Martin, L. B. & Khan, M. I**. Nontyphoidal *Salmonella* disease: Current status of vaccine research and development. *Vaccine* 34, 2907-2910, doi:10.1016/j.vaccine.2016.03.072 (2016).

177. **Thouand, G., Daniel, P., Horry, H., Picart, P., Durand, M. J., Killham, K., Knox, O. G., DuBow, M. S. & Rousseau, M.** Comparison of the spectral emission of lux

recombinant and bioluminescent marine bacteria. *Luminescence* 18, 145-155, doi:10.1002/bio.716 (2003).

178. **Tindall, B. J., Grimont, P. A., Garrity, G. M. & Euzeby, J. P.** Nomenclature and taxonomy of the genus *Salmonella*. *International Journal of Systematic and Evolutionary Microbiology* 55, 521-524, doi:10.1099/ijs.0.63580-0 (2005).

179. **Valouev, A., Johnson, D. S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., Myers, R. M. & Sidow, A.** Genome-wide analysis of transcription factor binding sites based on ChIP-Seq data. *Nature Methods* 5, 829-834, doi:10.1038/nmeth.1246 (2008).

180. **Van Dyk, T. K., Wei, Yan, Hanafey, Michael K., Dolan, Maureen, Reeve, Mary Jane G., Rafalski, J. Antoni, Rothman-Denes, Lucia B., LaRossa, Robert A.** A genomic approach to gene fusion technology. *PNAS* 98, 2555-2560, doi:10.1073/pnas.041620498 (2001).

181. **van Kessel, J. C., Ulrich, L. E., Zhulin, I. B. & Bassler, B. L.** Analysis of activator and repressor functions reveals the requirements for transcriptional control by LuxR, the master regulator of quorum sensing in *Vibrio harveyi*. *MBio* 4, doi:10.1128/mBio.00378-13 (2013).

182. **Vidal, O., Longin, R., Prigent-Combaret, C., Dorel, C., Hooreman, M. & Lejeune, P.** Isolation of an *Escherichia coli* K-12 mutant strain able to form biofilms on inert surfaces: involvement of a new ompR allele that increases curli expression. *Journal of Bacteriology* 180, 2442-2449 (1998).

183. **Waldner, L. L., MacKenzie, K. D., Koster, W. & White, A. P.** From Exit to Entry: Long-term Survival and Transmission of Salmonella. *Pathogens* 1, 128-155, doi:10.3390/pathogens1020128 (2012).

184. **Wen, Y., Ouyang, Z., Devreese, B., He, W., Shao, Y., Lu, W. & Zheng, F**. Crystal structure of master biofilm regulator CsgD regulatory domain reveals an atypical receiver domain. *Protein Science* 26, 2073-2082, doi:10.1002/prot.3245 (2017).

185. **White, A. P., Gibson, D. L., Collinson, S. K., Banser, P. A. & Kay, W. W**. Extracellular Polysaccharides Associated with Thin Aggregative Fimbriae of *Salmonella enterica* Serovar Enteritidis. *Journal of Bacteriology* 185, 5398-5407, doi:10.1128/jb.185.18.5398-5407.2003 (2003).

186. **White, A. P., Gibson, D. L., Grassl, G. A., Kay, W. W., Finlay, B. B., Vallance, B. A. & Surette, M. G.** Aggregation via the red, dry, and rough morphotype is not a virulence adaptation in *Salmonella enterica* serovar Typhimurium. *Infection and Immunity* 76, 1048-1058, doi:10.1128/IAI.01383-07 (2008).

187. **White, A. P., Gibson, D. L., Kim, W., Kay, W. W. & Surette, M. G.** Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of *Salmonella*. *Journal of Bacteriology* 188, 3219-3227, doi:10.1128/JB.188.9.3219-3227.2006 (2006).

188. **White, A. P. & Surette, M. G.** Comparative genetics of the rdar morphotype in *Salmonella*. *Journal of Bacteriology* 188, 8395-8406, doi:10.1128/JB.00798-06 (2006).

189. **White, A. P., Weljie, A. M., Apel, D., Zhang, P., Shaykhutdinov, R., Vogel, H. J. & Surette, M. G.** A global metabolic shift is linked to *Salmonella* multicellular development. *PLoS One* 5, e11814, doi:10.1371/journal.pone.0011814 (2010).

190. **Whitfield, C. & Roberts, I. S.** Structure, assembly and regulation of expression of capsules in *Escherichia coli*. *Molecular Microbiology* 31, 1307-1319 (1999).

191. **Wong, K. H., Jin, Y. & Moqtaderi, Z.** in *Current Protocols in Molecular Biology*. Ch. 7, Unit 7.11, (2013).

192. **Wu, D. Y., Bittencourt, D., Stallcup, M. R. & Siegmund, K. D.** Identifying differential transcription factor binding in ChIP-seq. *Frontiers in Genetics* 6, 169, doi:10.3389/fgene.2015.00169 (2015).

193. **Yrlid U., S. M., Hakansson A., Chambers B. J., Ljunggren H. G., Wick M. J.** In vivo activation of dendritic cells and T cells during *Salmonella enterica* serovar Typhimurium infection. *Infection and Immunity* 69 (2001).

194. **Zakikhany, K., Harrington, C. R., Nimtz, M., Hinton, J. C. & Romling, U.** Unphosphorylated CsgD controls biofilm formation in *Salmonella enterica* serovar Typhimurium. *Molecular Microbiology* 77, 771-786, doi:10.1111/j.1365-2958.2010.07247.x (2010).

195. **Zhang, Y., Hu, Ji-Fan, Wang, Hong, Cui, Jiuwei, Gao, Sujun, Hoffman, Andrew R., Li, Wei.** CRISPR Cas9-guided chromatin immunoprecipitation identi es miR483 as an epigenetic modulator of IGF2 imprinting in tumors. *Oncotarget* 8, 34177-34190 (2017).

196. **Zhang, Y., Lin, Y. H., Johnson, T. D., Rozek, L. S. & Sartor, M. A**. PePr: a peak-calling prioritization pipeline to identify consistent or differential peaks from replicated ChIP-Seq data. *Bioinformatics* 30, 2568-2575, doi:10.1093/bioinformatics/btu372 (2014).

197. **Zogaj, X., Nimtz, M., Rohde, M., Bokranz, W. & Römling, U.** The multicellular morphotypes of *Salmonella* typhimurium and *Escherichia coli* produce cellulose as the second component of the extracellular matrix. *Molecular Microbiology* 39, 1452-1463, doi:10.1046/j.1365-2958.2001.02337.x (2001).

APPENDIX A

Anti-CsgD monoclonal antibody purification and characteristics.

Table A1. Anti-CsgD mouse monoclonal antibody initial characteristics, purification methods, and final stocks for use in ChIP-seq in *Salmonella* Typhimurium.

Ascites received	Antibody product	Raw mouse ascites
	Received from	Immunoprecise
	Cell line	Mouse anti-CsgD His-RP 6D4
	Volume	36.5ml + 1ml stock
	Mouse strain	5x BALB/c
	Initial concentration	0.9-10mg/ml
Purification	Method	Protein G chromatography cartridge, syringe
	Solutions	Pierce Protein G Binding buffer Elution buffer 0.1M Glycine pH 2.3 Neutralization buffer 1M Tris-HCl pH 8.0
	Ascites flow-through volume and elutions	~5ml per flow-through 10 elutions
Processing	Protein concentration cut-off	>0.12mg/ml
	Combined and concentrated mAb	50K MWCO centrifugal filter
Final antibody stocks	Final volume and storage solution	10.3ml in 50mM Tris-HCl pH 8.0 + 0.02% Sodium azide
	Aliquots	6x 1.5ml 1x 1.3ml
	Concentration	3mg/ml
	Tested efficacy	Yes, by Western blot and dot blot

APPENDIX B

ChIP-seq protocol from a manuscript prepared for the Journal of Visualized Experiments (JoVE).

1. Flask culture cell growth

1.1. From frozen stocks, streak Salmonella serovar Typhimurium 14028 on LB agar and incubate at 37 °C for 16-20 h to obtain isolated colonies.

1.2. Inoculate 5 mL LB broth with 1-3 colonies from streak plates and incubate at 37 °C with shaking for 7 h

1.3. Find OD_{600} of broth culture using a spectrophotometer

1.4. Add 1.0 OD_{600} volume equivalent of cell culture (i.e. 10^9 cells) to an Erlenmeyer flask containing 100 mL 1% tryptone. Incubate at 28 °C with shaking for 13 h

2. Collect cell-free conditioned 1% tryptone

2.1. Collect flask culture for cell-free conditioned 1% tryptone.

2.2. Pipette flask culture several times to mix before adding to a centrifuge tube. Centrifuge at 12 000 xg for 10 min at 10 °C.

2.3. Decant supernatant into a 0.2 µm filter unit, vacuum filter, and dispense into a new tube.

Note: Biofilm are highly resistant in conventional solutions (e.g. PBS) and will adhere to itself and the sides of tubes and pipette tips. Resuspension in conditioned tryptone allows for manipulation and homogenization.

3. Separate biofilm and planktonic cells from flask cultures [111].

3.1. Using a sterile 25 mL pipette, aliquot flask culture into 15 mL tubes. Centrifuge at 210 xg for 2 min, to separate the two cell types.

Note: Biofilm cells should be a loose pellet at the bottom of the tube

3.2. Pipette the supernatant containing the planktonic cells into a centrifuge tube. Do not disturb the pellet. The supernatant will be used later. Remove all remaining liquid from the pellet.

4. Prepare biofilm aggregates

Note: Biofilm is harvested by wet weight to yield 25 µg of DNA starting material. The CFU of *S.* Typhimurium biofilm can be found from biofilm weight by this conversion factor: 1.63×10^8 CFU/mg. Researchers are advised to determine the weight of biofilm required using this equation:

$$25\ \mu g\ DNA\ x \frac{1\ \text{g}}{1\text{x}10^6\ \text{g}} x \frac{6.022\text{x}10^{23}\ \text{bp}}{1\ \text{mole}} x \frac{1\ \text{mole}}{650\ \text{g}} x \frac{1\ \text{CFU}}{\text{genome size (bp)}} xx \frac{1\ g}{1000\ mg} x \frac{Biofilm\ conversion\ factor\ mg}{1\ CFU} x \frac{1000\ mg}{1\ g}$$

Example:

$$25\ g\ DNA\ x \frac{1\ g}{1x10^6\ g} x \frac{6.022x10^{23}\ bp}{1\ mole} x \frac{1\ mole}{650\ g} x \frac{1\ CFU}{4\ 870\ 265\ bp} x \frac{1\ g}{1000\ mg} x \frac{1\ mg}{1.63x10^8\ CFU} x \frac{1000\ mg}{1\ g}$$

4.1. Resuspend the biofilm pellet in 1 mL conditioned tryptone. Move the resuspended

biofilm into a pre-weighed 2 mL snap cap (i.e., Safe-Lock) or screw cap tube.

4.2. Centrifuge for 5 s, rotate and centrifuge for 1 min at 11 000 xg at 20 °C

4.3. Remove all supernatant from the tube and weigh the tube accurately. Subtract the tube weight from the weight of the tube with biofilm to find the weight of aggregates. It should be +/-10% of the target biofilm weight.

4.4. Wash the biofilm pellet in PBS. Add 1ml PBS to the tube and vortex to resuspend the pellet.

5. Prepare planktonic cells

5.1. Dispense supernatant from slow speed centrifugation into 40 mL centrifuge tubes.

5.2. Measure the OD_{600} of the planktonic cells using a spectrophotometer and calculate the required volume for a final OD_{600} of 6.0.

$$Final\ vol.(mL) = Vol.of\ planktonic\ supernatant\ (mL)\ x\ \frac{Supernatant\ OD_{600}}{6.0\ OD_{600}}$$

5.3. Pellet planktonic cells using the floor centrifuge. Cool the floor centrifuge to 10 °C and spin at 10 000 xg for 10 min at 10 °C.

5.4. Remove the supernatant and resuspend the pellet in the calculated final volume of PBS. Re-measure OD_{600} of the planktonic cells using a spectrophotometer.

5.5. Dispense the volume of 6.0 OD_{600} planktonic cells into a 2 mL snap cap (i.e., Safe-Lock) or screw cap tube . Bring volume to 1 mL with conditioned tryptone.

6. Homogenize cells

6.1. Aseptically add one sterilized metal bead to each of the tubes.

6.2. Homogenize using a mixer mill (i.e., Qiagen TissueLyser II) for 5 min at 30 Hz. Observe aggregate tubes to confirm that biofilm has been broken apart.

6.3. Transfer the homogenized cells to a new 1.5 mL tube, avoiding the metal bead, Bring the volume to 1 mL with PBS.

Note: perform drop dilutions to enumerate input cells and check that the final cell number is close to the desired or predicted number

7. Cross-linking of proteins to DNA

7.1. Dispense fresh formaldehyde into sample tubes to a final concentration of 1%. Incubate for 30 min at room temperature on a rotating wheel.

Caution: Formaldehyde is corrosive, a skin, eye, and respiratory irritant, and flammable. Dispense in a fume hood.

7.2. Add glycine to a final concentration of 125 mM to stop crosslinking. Incubate for 5 min at room temperature on a rotating wheel.

8. Wash cells to remove excess crosslinker

8.1. Centrifuge tube for 3 min at 8000 xg and remove the supernatant

8.2. Resuspend the pellet in 20 µL 25x protease inhibitors (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Sigma-Aldrich, *COEDTAF-RO ROCHE)* and 500 µL filter sterilized PBS.

8.3. Centrifuge tube for 3 min at 8000 xg and remove the supernatant.

9. Lyse cells

9.1. Resuspend the pellet in 600 µL Lysis buffer and incubate on ice for 10 min.

9.2. Add 1.4 mL IP dilution buffer to a sterile 15 mL tube, and move lysed cells to the 15 mL tube. Keep the tube on ice for 1.5-2 h, and vortex occasionally.

Note: Some resistant material may remain in tubes. A long incubation period on ice with occasional vortexing will break apart material. The remainder will be broken up through sonication.

10. Sonicate to fragment DNA

10.1. Tune the sonicator (Vibra-Cell Ultrasonic Processor, VC300, 3mm probe). Place the 15 mL tube in a beaker of ice, and place the probe inside the tube.

10.2. Perform 5 sonication "bursts" of 30 s on at 20-40% of 400 W with 2 min cooling on ice between bursts.

10.3. Remove precipitated material by centrifuging at 15800 xg for 10 min at 4 °C. Transfer the supernatant to a new tube.

10.4. OPTIONAL: Rapidly decrosslink at 65°C and digest RNA and protein at 45°C before running on a 2% agarose gel gel to check for proper size fragments.

Note: Immerse the probe into the cell lysate. Keep the solution on ice while sonicating and resting. Do not remove the tube while the sonicator probe is pulsing. The solution should appear cloudy pre-sonication and clear post-sonication.

11. Immunoprecipitate DNA-protein-antibody complexes

In this experimental format, DNA that binds non-specifically to normal mouse IgG is removed. The remaining DNA is immunoprecipitated with normal mouse IgG (control IP) or the transcription-factor specific monoclonal antibody (test IP). Pre-cleared input DNA without immunoprecipitation selection is also a suggested sequencing control.

11.1. Pre-clear IP samples

11.1.1. Add 50 µg normal mouse IgG to sonicated DNA and incubate for 1 h at 4 °C on a rotating wheel.

11.1.2. Add 100 µl Protein G magnetic beads (ChIP-grade, Cell Signaling Technology, #9006S) to sonicated DNA containing normal mouse IgG. Incubate 3 h at 4 °C on a rotating wheel.

11.1.3. Separate the bead complexes on a magnetic stand. Dispense the supernatant into a new 15 mL tube containing 1 mL IP Dilution buffer to bring it to a total volume of 3 mL.

11.1.4. Dispense two 1.35 mL aliquots for immunoprecipitation and one 200 µL aliquot as an input control. Keep the input control at -80 °C until decrosslinking and digesting steps.

11.2. Immunoprecipitation with primary antibody.

11.2.1. Add 10 µg purified protein-specific primary antibody to one of the 1.35 µL aliquots as a test sample. Add 10 µg non-specific normal mouse IgG to one of the 1.35 µL aliquots as a control sample. Incubate overnight at 4 °C on a rotating wheel.

Note: The primary antibody can be a monoclonal antibody, high-quality polyclonal serum or commercial epitope antibody (i.e., anti-FLAG)

11.2.2. Add 50 µL Protein G magnetic beads to the precleared DNA with primary antibody and incubate at 4 °C for 3 h on a rotating wheel.
11.2.3. Place IP wash buffer 1, IP wash buffer 2, and TE pH 8.0 in an ice bucket. Warm elution buffer to 65 °C.

11.2.4. Bind beads to the side of tubes using a magnetic stand
11.2.5. Perform washes: wash twice with 750 µL cold IP wash buffer 1, wash once with 750 µL cold IP wash buffer 2, and twice with 750 µL cold TE at pH 8.0. Keep the tubes on the magnetic stand during washes.
11.2.6. Add 450 µL IP Elution buffer to each tube and incubate at 65 °C for 30 min with gentle vortexing every 5 min.
11.2.7. Bind beads to the side of tubes using a magnetic stand. Wait at least 2 min until the solution is clear.
11.2.8. Dispense the cleared solution to a new 1.5 mL tube.

Note: Do not put solution containing elution buffer on ice.

12. Reverse crosslinks and digest RNA

12.1. Add 2 µg RNase A and NaCl to a final concentration of 0.3 M to each tube. Incubate at 65 °C, for ≥6 h, or overnight

13. Digest protein

13.1. Add 180 µg Proteinase K to each tube
13.2. Incubate at 45 °C for 3-5 h.

14. Purify DNA.

Note: Magnetic beads are preferred for isolating the small amounts of DNA usually recovered during ChIP with bacterial cells.

15. Prepare libraries using a kit that is compatible with your selected sequencing platform (NEBNext Ultra DNA Library Prep Kit for Illumina, E7370S).

Check library concentration with a Qubit (Thermo Fisher Scientific, Q33216) and BR dsDNA kit (Thermo Fisher Scientific, Invitrogen, Q32850) or a qRT-PCR library quantification kit.

Note: Qubit measurements may overestimate DNA concentration slightly.

16. **Pool libraries to 20 pM in 100 µL, accounting for adequate coverage for each library sample.** For *Salmonella*, we pooled 10-12 libraries. Add Tris-HCl pH 8.0 to a final concentration of 1 mM, or TE pH 8.0 to a final concentration of 1 mM.

16.1. Dilute the library pool and sequence according to the selected platform's specifications (MiSeq, MiSeq Reagent Kit v3 (150 cycle) Illumina, MS-102-3001).

APPENDIX C

Materials required for performing ChIP-seq with *S.* Typhimurium biofilm and planktonic cells from 13 hour flask culture.

Table C1. Materials and equipment required for culturing and harvesting *S.* Typhimurium cell types

Name of Material or Equipment	Company	Catalog Number	Description/Procedural step
Luria-Bertani broth	BD Biosciences		Overnight culture in 5 mL LB for inoculating flask culture
BD Bacto™ Tryptone	BD Biosciences	211705	Media for the growth of *S.* Typhimurium cell types in flask culture
Serological pipettes, 10 and 25 mL	FroggaBio	2507646 2507645	Transfer of flask culture (biofilm adheres to glass pipettes)
Spectrophotometer and 2.0 mL cuvettes			Measure OD_{600} of planktonic cells
Nalgene® 115 mL Filter Units, Sterile, 0.2 µm pore size	Thermo Scientific	73520-980	Filtration unit for preparing conditioned tryptone from a portion of flask culture.
40 mL centrifuge tubes	Sigma Aldrich	T1418	Centrifugation of flask culture for conditioned tryptone and slow speed centrifugation supernatant for planktonic cells
15 mL tubes	FroggaBio	116930000	Collection and separation of cell types

Table C2. Materials and equipment required for homogenizing, crosslinking, lysing, and sonicating cell samples

Name of Material/ Equipment	Company	Catalog Number	Description/Procedural step
Safe-Lock Tubes, 2.0 mL, colorless	Eppendorf	22363344	Small tubes with a secure, snapping lid for use during homogenization.
Mixer mill	Retsch	MM 400	Equipment for homogenizing cells in Safe-Lock tubes
5mm Stainless Steel Beads	Qiagen	69989	Placed into Safe-Lock tubes with cell samples to aid homogenization with Retsch mixer mill
Formaldehyde	Sigma-Aldrich	252549	ACS reagent, 37 % in H2O, contains 10-15% methanol for crosslinking.
Ultrasonic Liquid Processor	Sonics & Materials, Inc.	VC300	Probe sonicator for fragmenting DNA after lysis.
Tapered microtip 1/8" (3mm)	Sonics & Materials, Inc.	630-0418	Probe for sonication-based DNA fragmentation.
Tube Revolver	Crystal Industries	HYQ-1130A	Portable rotating wheel for distributing and incubating ChIP reagents in Eppendorf tubes.

			Can be substituted for tilting or "Belly Dancer" platform.
Mouse Gamma Globulin	Jackson ImmunoResearch Laboratories, Inc	015-000-002	Normal species-matched IgG for control (i.e. "mock") IP
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Sigma-Aldrich	COEDTAF-RO ROCHE	Prevents protein digestion prior to immunoprecipitation. Can be substituted for other protease inhibitor cocktails.
RNase A, DNase and protease-free (10 mg/mL)	ThermoFisher Scientific	EN0531	Enzyme for digesting RNA prior to DNA purification.
Proteinase K powder	ThermoFisher Scientific	AM2542	Enzyme for digesting protein prior to DNA purification
Axygen Magnetic beads	Macherey-Nagel	744970	DNA purification
ChIP-Grade Protein G Magnetic Beads	Cell Signaling Technology	9006	Immunoprecipitation.
DynaMag™-2 Magnet	Life Technologies	12321D	Separation of magnetic beads during DNA immunoprecipitation and purification

Table C3. Buffer recipes for cell lysis.

Lysis buffer	**IP dilution buffer**
50mM Tris-HCl pH 8.1	20mM Tris-HCl pH 8.1
10mM EDTA	2mM EDTA
1% SDS	0.01% SDS
protease inhibitors	protease inhibitors
	150mM NaCl
	1% Triton X-100

Table C4. Materials and equipment required for immunoprecipitation and DNA purification

Name of Material/ Equipment	Company	Catalog Number	Description/Procedural step
ChIP-Grade Protein G Magnetic Beads	Cell Signaling Technology	9006	Immunoprecipitation.
NGS Cleanup and Size Select	Machery-Nagel	744970.5	Purify DNA after immunoprecipitation
Anti-CsgD Antibody	See above (Immunoprecise antibody)		
DynaMag™-2 Magnet	Life Technologies	12321D	Separation of magnetic beads during DNA immunoprecipitation and purification
Tube Revolver	Crystal Industries	HYQ-1130A	Portable rotating wheel for distributing and incubating ChIP reagents in Eppendorf tubes. Can be substituted for tilting or "Belly Dancer" platform.

Table C5. Buffer recipes for immunoprecipitation washes and elution

IP wash buffer 1	IP wash buffer 2	TE ($T_{10}E_1$) pH 8.0	Elution buffer
20mM Tris-HCl pH 8.1	10mM Tris-HCl pH 8.1	10mM Tris-HCl	1% SDS
2mM EDTA	1mM EDTA	1mM EDTA	100mM NaHCO3
50mM NaCl	250mM LiCl		
1% Triton X-100	1% NP-40		
0.1% SDS	1% deoxycholic acid		

Table C6. Materials and equipment required for confirmatory tests, library preparation, and sequencing

Name of Material/ Equipment	Company	Catalog Number	Description/Procedural step
Qubit 3.0	Thermo Fisher Scientific	Q33216	Measure DNA concentration of ChIP DNA accurately
BR dsDNA kit	ThermoFisher Scientific	Q32850	
2100 Bioanalyzer	Agilent	G2939BA	Observe fragment size range of DNA before and after library preparation
High Sensitivity DNA kit	Agilent	5067-4626	
NEBNext Ultra DNA Library Prep Kit for Illumina	New England BioLabs	E7370S	Prepare libraries from ChIP DNA
MiSeq Reagent Kit v3 (150 cycle)	Illumina	MS-102-3001	Sequence DNA libraries

APPENDIX D

Materials required for performing luciferase assays with *S.* Typhimurium with *luxCDABE* promoter-reporters for virulence- and persistence-associated genes.

Table D1. Materials and equipment required for luciferase assays measuring gene expression in the presence of waste effluent.

Name of Material or Equipment	Company	Catalog Number	Description
96-Well Clear Bottom Black or White Polystyrene Microplates	Corning	07-200-567	Plates for detection of bacterial growth and fluorescence
Multi-Detector Microplate Reader VICTOR™ X3, VICTOR^{3}V 1420	Perkin Elmer	2030-0050 1420-040	Equipment for detecting bacterial growth and fluorescence intensity
Mineral oil	Sigma-Aldrich	M5904	Applied to plate wells to prevent evaporation
Waste effluent or chemostat media	Given by Dr. Emma Allen-Vercoe		
Luria-Bertani broth	BD Biosciences	244620	Overnight culture in 5 mL LB for inoculating plates, growth media for luciferase assays
BD Bacto™ Tryptone	BD Biosciences	211705	Growth media for luciferase assays
2,2-dipyridyl (2,2′-Bipyridyl)	Sigma-Aldrich	D7505	Iron chelator
Kanamycin	Sigma-Aldrich	60615	Maintenance of the pCS26-*luxCDABE* plasmid during growth