FUNCTION AND REGULATION OF THE POLYSACCHARIDE UTILIZATION LOCUS, *DON*, IN THE GUT SYMBIONT *BACTEROIDES FRAGILIS*

By

Yanlu Cao

August, 2014

Director of Dissertation: C. Jeffrey Smith, Ph.D.

Major Department: Department of Microbiology and Immunology

**Abstract**

*Bacteroides fragilis* is the most common anaerobe isolated from clinical infections and in this report we demonstrate a novel feature of the species that is critical to their success as an opportunistic pathogen. Among the *Bacteroides spp.* in the gut, *B. fragilis* has a unique ability to efficiently harvest complex N-linked glycans from the glycoproteins common to serum and serous fluid. This activity is mediated by a Sus-like outer membrane protein complex designated as Don. Using the abundant serum glycoprotein transferrin as a model it was shown that *B. fragilis* alone can rapidly and efficiently deglycosylate this protein *in vitro* and that transferrin glycans can provide the sole source of carbon and energy for growth in defined media. We then showed that transferrin deglycosylation occurs *in vivo* when *B. fragilis* is propagated in the rat tissue cage model of extraintestinal growth and that this ability provides a competitive advantage *in vivo* over strains lacking the *don* locus. Thus, Don functionally is an extraintestinal growth factor that may contribute to *B. fragilis* opportunistic infection.

The regulation of *don* expression is controlled by two independent pathways. The first one was shown to be a typical ECF sigma/anti-sigma factor switch, commonly found in Sus-like Polysaccharide Utilization Loci (PULs), which responds to the presence of specific substrate. In

the ECF sigma factor deletion mutant, *ΔdonA*, expression of the *don* PUL was completely abolished in the presence of substrate glycans, while the cognate anti-sigma deletion strain, *ΔdonB,* expressed the *don* genes even in the absence of substrate glycans. The *donA* overexpressing strain highly expressed the *don* PUL regardless of the substrate glycan presence. The second regulatory pathway is involved with a *cis*-encoded antisense sRNA which is associated within the *don* locus, DonS. DonS was shown to negatively regulate *don* expression. In contrast, expression of the *don* genes was induced two- to six-fold in the *donS* silencing mutant and highly repressed in the *donS* overexpressing strain. Notably, this sRNA controlled regulatory pathway is not commonly found associated with *B. fragilis* PULs. Only 14 of more than 50 PULs in *B. fragilis* possess DonS-like sRNAs, but at the present time their roles in commensal colonization and opportunistic infections is not understood.

# FUNCTION AND REGULATION OF THE POLYSACCHARIDE UTILIZATION LOCUS, *DON*, IN THE GUT SYMBIONT *BACTEROIDES FRAGILIS*

A Dissertation Presented to

The Faculty of the Department of Microbiology and Immunology

Brody School of Medicine at East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Microbiology and Immunology

By

Yanlu Cao

August 15th, 2014

FUNCTION AND REGULATION OF THE POLYSACCHARIDE UTILIZATION LOCUS, *DON*, IN THE GUT SYMBIONT *BACTEROIDES FRAGILIS*

by

Yanlu Cao

APPROVED BY:

DIRECTOR OF DISSERTATION

C. Jeffrey Smith, Ph.D.

COMMITTEE MEMBER

Everett C. Pesci, Ph.D.

COMMITTEE MEMBER

R. Martin Roop, II, Ph.D.

COMMITTEE MEMBER

Rachel L. Roper, Ph.D.

COMMITTEE MEMBER

Warren Knudson, Ph.D.

CHAIR OF THE DEPARTMENT OF MICROBIOLOGY AND IMMUNOLOGY

C. Jeffrey Smith, Ph.D.

INTERIM DEAN OF THE GRADUATED SCHOOL

Paul J. Gemperline, Ph.D.

# ACKNOWLEDGEMENTS

First I would like to thank my mother Hong Cao, who had encouraged and supported me to pursue an education and career in the United States, for her constant love. I would also like to thank all my family members from my maternal side, especially my grandparents, for their support and help since my childhood. For the past six years during my high education in ECU, I foremost have to give my thanks to my advisor, Dr. C. Jeffrey Smith. For his dedication and mentorship, I can accomplish the work in this study and be qualified to have a doctoral degree in microbiology and immunology. I also must thank Dr. Edson R. Rocha, Dr. Ivan Ndamukong, Anita C. Parker and Michael Betteken for their support and cooperation, both as co-workers and great friends. I also need to acknowledge Dr. Clayton C. Caswell for his technique support during the small RNA study. I would also like to thank my committee members, Dr. Everett C. Pesci, Dr. R. Martin Roop, II, Dr. Rachel L. Roper and Dr. Warren Knudson for their contribution to my educational development. For many years I always want, but have not had a chance, to say thanks to the Biology Department of Indiana University of Pennsylvania, especially to Dr. Carl Luciano and Dr. Narayanaswamy Bharathan, for their help which are not just limited to education. There are many people in my life that I sincerely appreciate but not be able to mention every one individually here. To them, I just want to say: thank you all and God bless.

TABLE OF CONTENTS

LIST OF TABLES ..... vi

LIST OF FIGURES ..... vii

CHAPTER ONE: INTRODUCTION ..... 1

1.1 *Bacteroides* taxonomy, structure, physiology and metabolism ..... 1

1.2 The mutualistic relationship between *Bacteroides* and their hosts ..... 7

1.3 *Bacteroides* as opportunistic pathogens ..... 10

1.4 Virulence Mechanisms ..... 11

1.5 The polysaccharide utilization systems in *Bacteroides* ..... 14

CHAPTER TWO: EFFICIENT UTILIZATION OF COMPLEX N-LINKED GLYCANS IS A SELECTIVE ADVANTAGE FOR BACTEROIDES FRAGILIS IN EXTRAINTESTINAL INFECTIONS ..... 21

2.1 Abstract ..... 21

2.2 Significance ..... 22

2.3 Introduction ..... 22

2.4 Materials and methods ..... 24

2.4.1 Bacterial strains and growth ..... 24

2.4.2 Construction of a *don* operon deletion mutant ..... 27

2.4.3 Animal model of infection ..... 27

2.4.4 Total RNA extraction ..... 28

2.4.5 cDNA synthesis, qRT-PCR, and microarray analysis ..... 29

2.4.6 Whole cell deglycosylation assay ..... 30

2.4.7 Glycomics analyses of *B. fragilis* treated human transferring ..... 30

2.5 Results ..... 31

2.5.1 *In vivo*, extraintestinal PUL expression ..... 31

2.5.2 *In vitro* induction of the *donCDEFG* operon ..... 35

2.5.3 Deglycosylation of transferrin is mediated by *donCDEFG* ..... 38

2.5.4 Growth on transferrin as the sole carbon and energy source ..... 43

2.5.5 *In vivo* role of the *don* PUL ..... 46

2.5.6 Efficient transferrin deglycosylation is unique to *B. fragilis* ..... 50

2.6 Discussion ..... 53

2.7 Supplementary Figures ..... 58

2.8 Addendum ..... 67

CHAPTER THREE: *BACTEROIDES FRAGILIS* PRIMARY TRANSCRIPTOME ANALYSIS ..... 68

3.1 Introduction ..... 68

3.2 Materials and Methods ..... 69

3.2.1 Total RNA Extraction ..... 69

3.2.2 Preparation of cDNA libraries and 454 pyrosequencing ..... 69

3.2.3 Data Visualization ..... 70

3.3 Results ..... 71

3.3.1 Identification of *B. fragilis* TISs at a genome wide level ..... 71

3.3.2 Confirmation of the recognition sequence for the *B. fragilis* primary sigma factor .....74

3.3.3 Discovery of putative small non-coding RNAs in the *B. fragilis* genome.....77

3.4 Discussion.....80

CHAPTER FOUR: CHARACTERIZATION OF THE REGULATION MECHANISMS FOR THE *BACEROIDES FRAGILIS* POLYSACCHARIDE UTILIZTION LOCUS, *DON*.....82

4.1 Introduction .....82

4.2 Materials and Methods.....84

4.2.1 Bacterial strains and growth .....84

4.2.2 Construction of *donA* and *donB* deletion mutants and the *donA* over-expression strain .....85

4.2.3 Construction of DonS silencing mutant (IB561) and DonS over-expression strain (IB563).....86

4.2.4 Total RNA extraction, cDNA synthesis and qRT-PCR.....87

4.2.5 Northern blot analysis.....88

4.3 Results.....93

4.3.1 BF638R_3437 and BF638R_3438 encode regulators of the *don* operon .....93

4.3.2 A small non-coding RNA is divergently transcribed from the *donC* gene.....96

4.3.3 DonS is a *cis*-encoded antisense RNA that negatively regulates the *don* operon 99

4.3.4 DonS-like sRNAs are found in other PULs of *B. fragilis*..... 102

4.4 Discussion..... 108

REFERENCES ..... 112

APPNEDIX 1: COPYRIGHT PERMISSION POLICY OF THE AMERICAN SOCIETY FOR BIOCHEMISTRY AND MOLECULAR BIOLOGY ..... 123

APPENDIX 2: INDUCTION AND REPRESSION OF BACTEROIDES FRAGILIS 638R PUL GENES IN VIVO AND IN VITRO.....124

APPENDIX 3: PRIMARY TRANSCRIPTOME ANALYSIS .....133

APPENDIX 4: INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) APPROVAL FORM .....170

LIST OF TABLES

Table 1.1 Sus-like systems in *Bacteroidetes* ..... 19

Table 2.1 Bacterial strains and plasmids used in this study ..... 25

Table 2.2 Oligonucleotides used in this study ..... 26

Table 2.3 Quantitative glycan analysis of transferrin incubated with *B. fragilis* cells ..... 42

Table 4.1 Bacterial strains used in this study ..... 89

Table 4.2 Plasmids used in this study ..... 90

Table 4.3 Oligonucleotides used in this study ..... 91

Table 4.4 Characteristics of the DonS-like sRNAs in *B. fragilis* ..... 106

LIST OF FIGURES

Figure 1.1 A schematic diagram of two opposing half-cycles of the split TCA cycle in *B. fragilis*.....5

Figure 1.2 Functional model of the *B. thetaiotaomicron* starch utilization system (Sus).. 16

Figure 2.1 *In vivo* PUL gene expression during growth of *B. fragilis* in the rat tissue cage model ..... 33

Figure 2.2 Induction of *donC* expression during growth on mucin glycans ..... 36

Figure 2.3 Deglycosylation of transferrin is mediated by the *don* PUL ..... 40

Figure 2.4 *B. fragilis* can grow with transferrin as the sole source of carbon and energy 44

Figure 2.5 Growth *in vivo* is enhanced by deglycosylation of transferring ..... 48

Figure 2.6 Transferrin deglycosylation by medically important *Bacteroides* species ..... 51

Figure S1 Genetic map and RT-PCR analysis of the *donABCDEFG* genes (BF638R3437-43) ..... 58

Figure S2 Induction/repression patterns of SusC orthologues expressed during growth *in vivo* ..... 59

Figure S3 Deglycosylation assays using total serous fluid proteins with serum albumin removed ..... 63

Figure S4 *In vitro* competition assays in defined media with glucose ..... 64

Figure S5 Deglycosylation analysis of IgG and lactoferrin ..... 65

Figure S6 Deglycosylation analysis of transferrin by *B. thetaiotaomicron* strain 5482.... 66

Figure 3.1 The transcript signal of the BF638R_3589 gene..... 72

Figure 3.2 Logo for the promoter recognition sequence of the primary sigma factor in *B. fragilis*..... 75

Figure 3.3 The putative sRNA in the intergenic region of the *don* PUL..... 78

Figure 4.1 Relative expression of *donC* in different *B. fragilis* strains under non-inducing and inducing conditions..... 94

Figure 4.2 A small, non-coding RNA is divergently transcribe from the *donC* gene ..... 97

Figure 4.3 DonS is a negative regulator of the *don* expression..... 100

Figure 4.4 Northern blot confirmation of the DonS-like sRNAs in *B. fragilis* ..... 103

# CHAPTER ONE: INTRODUCTION

This dissertation mainly focuses on the study of a Sus-like polysaccharide utilization locus (PUL) in *Bacteroides fragilis*, named *don*. The impetus for studying this locus was derived from the fact that in the rat artificial abscess model, its expression was the most up-regulated in the *B. fragilis* transcriptome compared to in *vitro* growth. This indicates that the *don* locus may play a role in adaption to extraintestinal environments. Also, in an analysis of the *B. fragilis* primary transcriptome using RNA deep sequencing, a novel feature was observed of *cis*-encoded small RNAs associated with 14 PULs in *B. fragilis*, including *don*. Thus, the major objectives of this dissertation research were: first, characterize the role of *don* in colonization of extraintestinal sites (chapter 2); second, elucidate the role of sRNAs in the regulation of the Sus-like systems in *B. fragilis*, using the *don* locus as a model (chapter 4). Also included in this dissertation is a brief description of the *B. fragilis* primary transcriptome analysis since it led to the initial discovery of the sRNAs involved in the PUL regulation (chapter3). This chapter (chapter 1) is a general review on the current understanding of *Bacteroides* physiology and metabolism as it relates to virulence factors and the Sus-like polysaccharides utilization systems.

## 1.1 *Bacteroides* taxonomy, structure, physiology and metabolism

The genus *Bacteroides* are Gram-negative, obligate anaerobic bacteria commonly found in the gastro-intestinal tract (GI-tract) of humans and animals as members of the normal

microflora (1). The type species of *Bacteroides* is *B. fragilis*, so sometimes the genus is referred to as the *Bacteroides fragilis* group. This genus falls within the family *Bacteroidaceae* of the order *Bacteroidales* of the phylum *Bacteroidetes*. Together with the closely related genera *Porphyomonas* and *Prevotella*, they make up a major subgroup of the *Cytophaga-Flavobacter-Bacteroides* group which diverged from the primary eubacterial phylogenic lineage very early in evolution (2) and consequently are not closely related to the Gram-negative *Proteobacteria* which are more commonly studied.

All *Bacteroides* are non-motile, non-spore formers with relatively large genomes ranging in size from 4.4 to 6.7 Mb and relatively low in GC content ranging from 39% to 48% (1). The cell morphology is short rods and the doubling time under optimal growth conditions is approximately 45 minutes. *Bacteroides* have capsules which are important for their colonization in the distal gut (3). *B. fragilis* has eight different capsule polysaccharide encoding loci on its genome which are different in size, composition and staining character. Coexpression of two capsule polysaccharides is commonly observed (4-6). On the cell surface, *Bacteroides* possess pilli- or fimbrae-like structures that serve for attachment (1). Under transmission electron microscopy, numerous blebs on the cell surface and detached extracellular vesicles can be observed (7). These vesicles have been shown to have hemagglutinin function and contain sialidase activity (8, 9). The lipopolysaccharide (LPS) of *Bacteroides* is different from the canonical LPS structure represented by *E. coli* LPS. The lipid A in *Bacteroides* LPS is penta-acylated and monophosphorylated and the fatty acids chains are branched, 15 to 17 units long in contrast to the unbranched 12 or 14 unit length of canonical LPS (10). This structure probably is responsible for its low toxicity compared to the LPS of other Gram-negative bacteria (1).

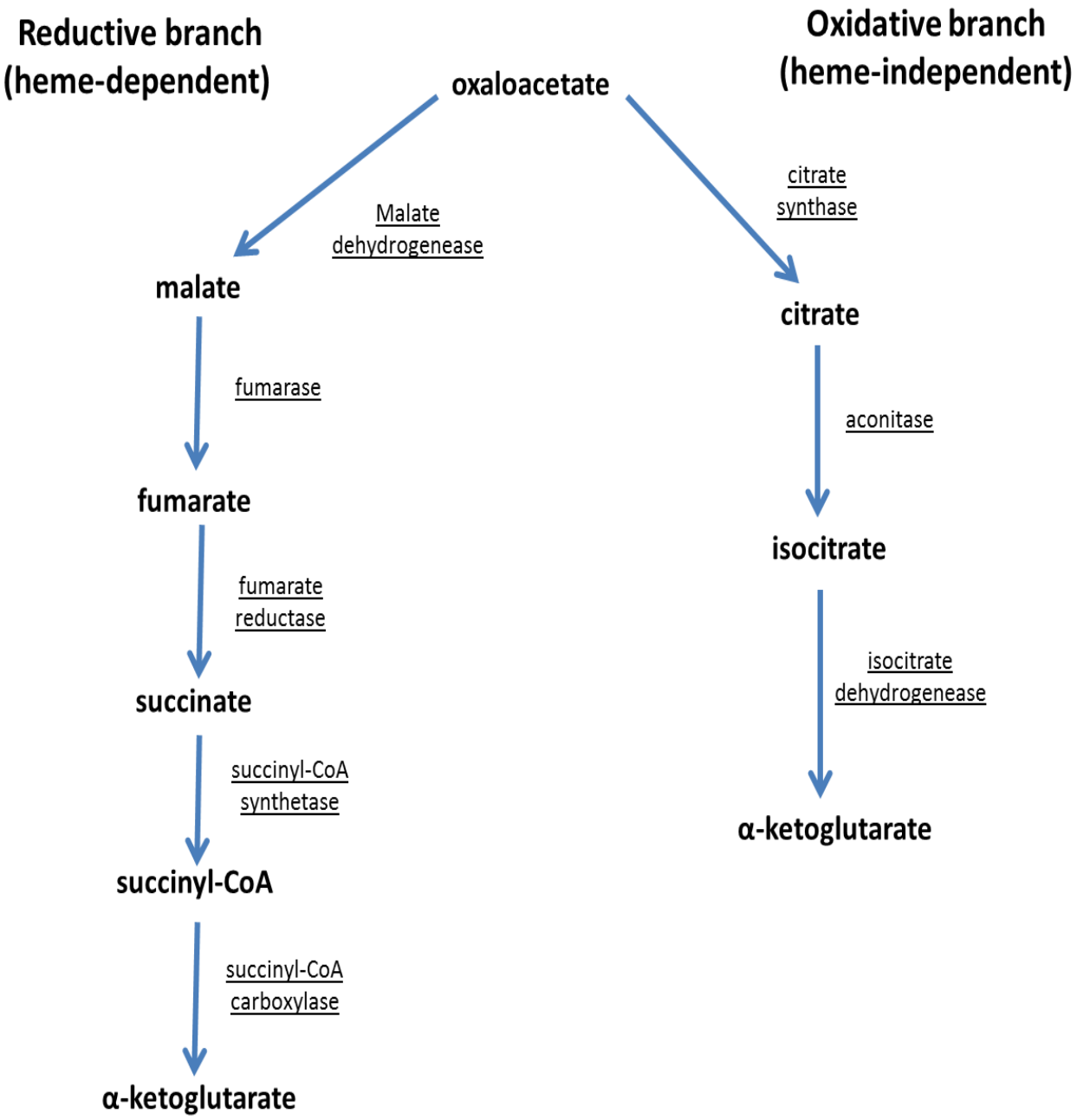
*Bacteroides* also have a rare O-glycsosylation system which enables them to synthesize fucosylated glycoproteins (11). The model for this general O-glycsosylation proposes that the glycan chain is assembled on a lipid carrier on the inner side of the cytoplasmic membrane by the sequential action of glycosyltransferases, then the glycan is flipped to the periplasmic space and glycosylation of the target protein happens there. The fucosylated O-linked glycoproteins then are inserted into the outer membrane surface and play a role in the colonization in the gut (12).

*Bacteroides* are saccharolytic organisms that primarily use carbohydrates as their carbon/energy source (1). They can express a large number of outer membrane protein complexes for polysaccharide transport and utilization (13). An in-depth review of these outer membrane complexes in *Bacteroides* will be given later in this chapter. Once the polysaccharides have been hydrolyzed to their monosaccharide units, further carbohydrate catabolism proceeds through the Embden-Meyerhof pathway followed by a split TCA cycle composed of two opposing half-cycles termed the oxidative branch and the reductive branch (1, 14). A simplified schematic diagram of this split TCA cycle is shown in Fig. 1.1. Anaerobic respiration makes use of the reductive branch of the TCA cycle with fumarate as the terminal electron receptor, reduced to succinate mediated by a membrane-bound fumarate reductase (1). Interestingly, *Bacteroides* also possess a cytochrome bd oxidase which allows them to use oxygen as a terminal receptor when oxygen is at nanomolar levels (15). Under nanomolar oxygen levels (60 nM to 1000 nM), the growth benefit of this cytochrome bd oxidase can be seen when respiration with fumarate is disrupted by deletion of the fumarate reductase (15).

However, as obligate anaerobes, *Bacteroides* cannot grow with oxygen levels close or above 2 μM (15).

Heme is important for *Bacteroides* metabolism during normal growth. However *Bacteroides* cannot synthesize heme *de novo* and it must be obtained from the environment (1). When grown in the absence of heme, their metabolism shifts from anaerobic respiration to strictly fermentation since heme is required as a cofactor by two critical components of the electron transport chains, fumarate reductase and cytochrome bd oxidase (1). Under these conditions, fumarate and lactate will accumulate as the metabolic byproducts instead of the short chain fatty acids succinate and propionate (1). During anaerobic respiration, fumarate is constantly consumed as the final electron receptor and reduced to succinate. The large demand for fumarate is provided by carboxylation of phosphoenolpyruvate to oxaloacetate, and then by enzymes of the reductive branch of the TCA cycle (1, 14). Carbon dioxide is required for the *Bacteroides* anaerobic respiration since it is the substrate for the carboxykinase to catalyze the carboxylation of phosphoenolpyruvate.

**Fig. 1.1.** A schematic diagram of two opposing half-cycles of the split TCA cycle in *B. fragilis* (14).


Reductive branch
(heme-dependent)
oxaloacetate
Oxidative branch
(heme-independent)
Malate
dehydrogenase
citrate
synthase
malate
citrate
fumarase
aconitase
fumarate
isocitrate
fumarate
reductase
isocitrate
dehydrogenase
succinate
α-ketoglutarate
succinyl-CoA
synthetase
succinyl-CoA
succinyl-CoA
carboxylase
α-ketoglutarate

As saccharolytic bacteria, *Bacteroides* cannot use proteins or peptides as the sole carbon/energy source but they can use them as the sole nitrogen source, although ammonia is the more preferred nitrogen source (1). Ammonia assimilation in *Bacteroides* is primarily through incorporation of ammonia into α-ketoglutarate to form glutamate catalyzed by glutamate dehydrogenase (1). Glutamate serves as an amino group donor for the synthesis of other amino acids. α-ketoglutarate biosynthesis is achieved through the reductive branch of the split TCA cycle in the presence of heme, alternatively, the oxidative branch in the absence of heme (14) (Fig. 1.1). The *Bacteroides*, with the possible exception of *B. ovatus*, synthesize all of their amino acids *de novo* (1).

### 1.2 The mutualistic relationship between *Bacteroides* and their hosts

The human distal gut is one of the most densely populated microbial ecosystems on earth, exceeding numbers of $10^{11}$ per gram of wet weight, and *Bacteroides* is one of the predominant species making up to 25% of the total microbial population in this ecological niche (1, 7, 16). A strong mutualistic relationship between the *Bacteroides* and their hosts has developed over the course of coevolution (16, 17).

The human large intestine provides an ideal space for the *Bacteroides* to thrive: sufficient moisture; highly reducing environment with oxygen levels less than 1.5%; carbon dioxide levels of 5% - 10%; adequate amount of ammonia, heme and vitamin $B_{12}$; and most importantly a stable source of carbon/energy from the constant influx of polysaccharides from indigestible material in the diet and host mucosal secretions (18-21). The host diet plays an

important role in shaping the gut microflora composition including the *Bacteroides* species (18-20, 22).

In return for the food and shelter, *Bacteroides* provide multiple benefits to their host. The breakdown and conversion of the indigestible dietary fiber to short chain fatty acids by the *Bacteroides* provides the host an additional nutrition source which can meet up to 10% of the human energy requirement (1, 7, 23). *Bacteroides* also play important roles in the development of the host GI tract: their presence greatly increases the volume, surface area of and villi numbers in the GI tract (1). They participate in the development of gut associated lymphatic tissue (GALT) and gut angiogenesis (24). They prevent the attachment and colonization of pathogenic bacteria by occupying the mucosal surface of the gut to form a protective barrier. B. *thetaiotaomicron* can also help prevent infection by Gram-positive species by promoting induction of a host cell secreted bactericidal lectin named RegIIIγ, which can directly bind to the peptidoglycan and result in bacterial killing (25). *B. thetaiotaomicron* also stimulates the production of a certain type of immune cell in the gut crypts called Paneth cells which secrete an antimicrobial substance to kill specific pathogens such as *Listeria monocytogenes* (26). *B. fragilis* in the gut can help the host develop a healthy immune response and prevent allergies (27). This is mediated in part by the capsule polysaccharide, polysaccharide A (PSA), which is a zwitterion, unlike most other capsule polysaccharides which are only negatively charged. PSA can bind to major histocompatibility complex II molecules in the professional antigen presenting cells and be presented to the T cells. It has been show that PSA can prevent inflammation by triggering the IL-10 mediated anti-inflammatory response at the same time repressing the Th17 mediated pro-inflammatory response (28, 29). It also can correct the CD4+

cell deficiency and establish Th1/Th2 balance in the gut (28, 29). The beneficial effect of PSA immunomodulatory properties even can be seen in the brain. A recent study using a mouse model shows that a subset of neurodevelopmental disorders, including autism spectrum disorder (ASD), could be due to dysbiosis of the gut microbiota and GI barrier impairment. Harmful bacterial metabolites, for example 4-ethylphenylsulfate (4EPS), can pass the tight junctions between intestinal epithelial cells and enter blood stream circulation. These metabolites can cause autism-like symptoms in the host such as abnormalities in sociability, communication, compulsive activity and severe stereotypic behavior. Oral treatment with *B. fragilis* can alter the gut microbial composition, restore normal GI barrier permeability and correct the ASD-related behavioral abnormalities (30). *Bacteroides* can efficiently deconjugate bile salts in the gut. Deconjugated bile salts have a much lower reabsorption rate and lowered ability to solubilize diet cholesterols. Non-reabsorbed bile will be excreted which has the indirect affect to reduce cholesterol levels by forcing the host to use more cholesterols for biosynthesis of bile salts and decrease cholesterol absorption (31). Lastly, several reports have shown that there is an association between the intestinal *Bacteroides* population and obesity. Obese individuals tend to have an abnormally low *Bacteroides* to *Firmicutes* ratio (32, 33). In some studies, it was also found that the relative abundance of *Bacteroides* increases as obese individuals lose weight (34). It is still not clear whether there is a causal relationship between obesity and the proportion of *Bacteroides* in the gut.

### 1.3 *Bacteroides* as opportunistic pathogens

Although the *Bacteroides* do provide many benefits as friendly commensals, they can transform into dangerous threats when translocated to extraintestinal sites. *Bacteroides* are frequently isolated from clinical anaerobic infections with high mortality (1, 7). *Bacteroides fragilis*, the type species of this group, only accounts for 0.5% of the colonic flora, but it is the most common anaerobe isolated from infections, being found in more than 63% of *Bacteroides* associated infections (1, 7). Thus, *B. fragilis* is regarded as the most virulent *Bacteroides* species and has served as a model for studying *Bacteroides* pathogenesis. Anaerobic infections are usually polymicrobial, and *B. fragilis* associated anaerobic infections have a mortality of more than 19%, up to 60% if left untreated (35).

*B. fragilis* opportunistic infections generally occur when the integrity of the intestinal wall is disrupted. Common predisposing conditions include GI surgery, trauma, malignancy and perforated inflammatory disease or infections in the GI tract (1, 7). *B. fragilis* associated infections include abscesses, appendicitis, bacteremia, bone and soft tissue infections, deep wound ulcers and UTI infections, but the most common and dangerous are intra-abdominal abscesses (IAA) (1, 7). IAA caused by *B. fragilis* are actually a pathologic immune response in which the immune cells are attracted to infected sites, if unable to clear the infection, additional immune cells are recruited to contain the bacterial cells and necrotic host tissue by walling them off with fibrin and connective tissues (36, 37). Once formed, antibiotic treatment has little effect on IAA and surgery is usually required. Rupture of the IAA without medical intervention usually leads to life-threatening situations.

It is also worthy to note that some strains of *B. fragilis* are able to produce the enterotoxin fragilysin. These strains are grouped as enterotoxigenic *B. fragilis* (ETBF) (92). ETBF can cause secretory diarrhea and studies have shown they also may be linked to inflammatory bowel disease and colon cancer (92).

### 1.4 Virulence Mechanisms

Following translocation from the gut to an extraintestinal site, *Bacteroides* typically face several challenges to the establishment of a successful infection. First, they need to attach at the site of infection, withstand the increased oxygen tension and resist clearing mechanisms from the host immune system. Next, they must persist in the new environment where the available nutritional resources become very limited, and where they need to shift their diet from a combination of dietary glycans and host glycans to host glycans only. *Bacteroides* need to coordinately express a variety of virulence factors to meet these challenges (1).

The *Bacteroides* capsule is a major factor associated with infection and pathogenesis. It contributes to the virulence in several ways. First, the capsule plays an important role in adherence. It has been shown that the ability of *B. fragilis* and other *Bacteroides* to adhere to rat peritoneal mesothelium depended on the presence of capsular polysaccharide, while acapsular strains were less adherent (39). Purified capsular polysaccharide alone can compete with live bacterial cells and inhibit *B. fragilis* adherence to rat peritoneal mesothelium (39). Second, the capsule can help *Bacteroides* withstand host clearance mechanisms due to its anti-phagocytic and anti-complement properties (40) and it subverts immune surveillance by

antigenic variation. For example, a single strain of *B. fragilis* has the capacity to synthesize eight different capsular polysaccharides, designated PSA to PSH (41), and expression of these capsular polysaccharides are regulated by two independent mechanisms. One is DNA inversion by which the promoter region of each individual locus can flip to give it an "on" or "off" status (42, 43). This inversion is mediated by a global DNA invertase named Mpi (42, 44, 45). The second mechanism involves *trans* locus inhibition. Within each capsule polysaccharide synthesis locus, there are two regulatory genes that encode UpxY and UpxZ, where x is replaced by "a" to "h" corresponding to PSA to PSH. UpxY is an antiterminator which is necessary for the transcription of the full length message of that polysaccharide synthesis locus, and UpxZ is a *trans* locus inhibitor which interacts with several other UpxYs to cause premature termination of the transcription of other capsule synthesis loci (46, 47). The combined effect of these two regulatory mechanisms allows *B. fragilis* to switch its capsule composition so that no more than two types of capsular polysaccharides are expressed at the same time. This ability to change the architecture on the cell surface may result in antigenicity variation and evasion of host immune surveillance (48). In addition to adherence and immune evasion, the capsule plays a vital role in abscess formation (49). Injection of capsule polysaccharide alone is sufficient to induce abscess formation, while systemic injection prevents abscess formation in rats, presumably due to the establishment of an antibody response to the capsules (49-51). It also has been demonstrated that the capsular polysaccharides do not contribute equally to abscess formation; rather, PSA is the most potent among all the eight to facilitate abscess formation (52). PSA is a zwitterionic (both positively and negatively charged) which can trigger a T-cell dependent response, and

contains the amino sugar acetamido-amino-2, 4, 6-trideoxygalactose (AATGal) as a component sugar. These two properties were thought to boost the PSA's abscess formation ability (49, 52).

Extraordinary aerotolerance is another important factor associated with *Bacteroides* extraintestinal infection. When translocated from the colon to the peritoneal cavity, bacteria have to face the challenge of increased oxygen tension up to 6% in this new environment. Oxidative stress will generate dramatic amounts of intracellular reactive oxygen species (ROS) such as hydrogen peroxide, superoxide and hydroxyl radicals, and ROS cause damage in a cell by peroxidation of membrane lipids, oxidation of amino acids in proteins, destruction of iron-sulfur clusters in enzymes, modification of DNA bases and strand breaks (53). *Bacteroides*, as strict anaerobes, cannot shift to a full aerobic metabolism but they can mount a sophisticated oxidative stress response that allows them to survive in oxygenated environments for extended periods of time. The oxidative stress response in *B. fragilis* has been well characterized, and can be divided into two phases. One is an acute response in which approximately 28 gene products involved in detoxification, repairing, and protection are quickly up-regulated within minutes of oxygen exposure (1, 54, 55). These gene products include catalase, peroxidase, superoxide dismutase, alkyl hydroperoxide reductase, and the non-specific DNA binding protein DPS, which all work synergistically to minimize the effect of ROS and to restore the reduced intracellular environment (54, 55). If the oxidative stress remains for an extended period, *B. fragilis* will mount a prolonged oxidative stress response (POST) in which there is a shift in gene expression aimed at remodeling the bacterial metabolism and physiology. This metabolic response induces genes encoding enzymes that can supply reducing power for detoxification and restore energy-generating capacity (55-59). This sophisticated oxidative stress response makes *B. fragilis* one of

the most aerotolerant anaerobes on earth and enhances its survival in extraintestinal sites to allow it to cause infection (55).

Besides its capsules and oxidative stress response, *B. fragilis* also has other virulence traits. These include the ability to release outer membrane vesicles containing a variety of hydrolases and hemagglutinases (8, 9), expression of neuraminidinases, expression of hemolysins (54, 60), and release of proteases and enterotoxin (enterotoixigenic strains only) (92).

### 1.5 The polysaccharide utilization systems in *Bacteroides*

*Bacteroides* are glycan utilization experts with extraordinary enzymatic abilities to degrade a wide variety of polysaccharides as carbon/energy sources (21). The human distal gut allows *Bacteroides* to exert their expertise by constantly providing a broad spectrum of polysaccharides, mostly undigested dietary glycans such as cellulose, pectin, xylan, and host derived glycans in the form of glycoconjugates such as glycoproteins, glycolipids and glycosaminoglycans (20, 22, 61). This ability explains in part why *Bacteroides* are so predominant in the distal gut (18, 19).

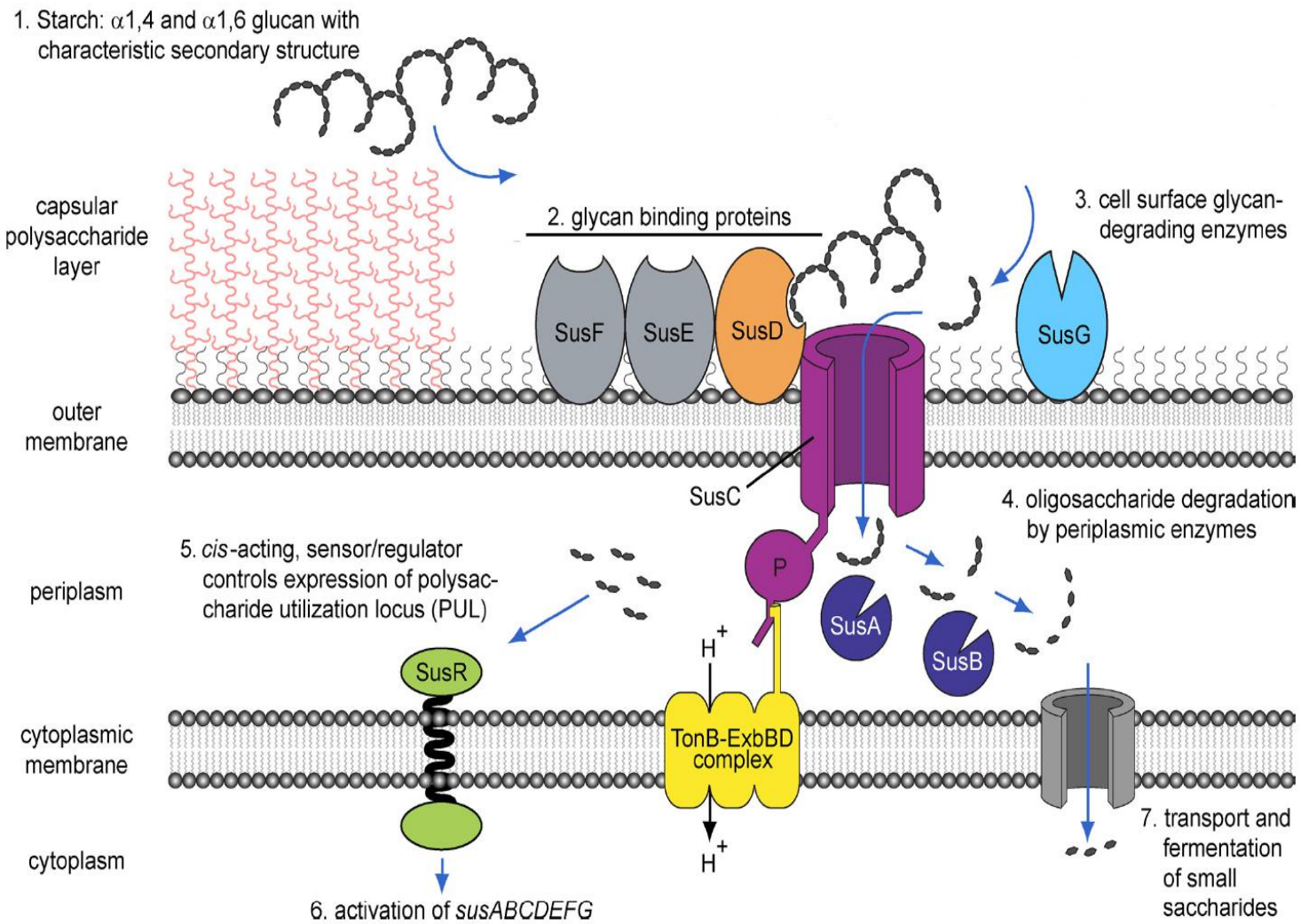
In *Bacteroides* genomes, genes encoding the machinery to degrade a specific type of polysaccharides are organized into multiple gene clusters, known as polysaccharide utilization loci (PULs). The ability to utilize such a wide variety of polysaccharides is the result of dedicating

a large portion of their genomes to these PULs. For instance, 18% of *B. thetaiotaomicron* genome is composed of these PULs (62).

*Bacteroides* PULs were first characterized in *B. thetaiotaomicron* using the starch utilization system (Sus), which specifically degrades starch, as a model. The sus locus contains eight genes *susABCDEFGR*. *SusA* and *susB* encode periplasmic amylases. *SusC* encodes a TonB-linked outer membrane porin which is required to transport the processed starch oligosaccharides from outside of the cell to the periplasmic space. The *susD* gene encodes an outer membrane protein which binds the oligosaccharides and aids its transportation through the channel formed by SusC porin. Downstream genes *susEFG* encode gene products which will form an outer membrane complex with SusC and SusD and are important for recognition and processing of large starch molecules (63-66). SusR is a transcriptional activator which responds to the substrate in the periplasmic space (67). A functional model of the *B. thetaiotaomicron* starch utilization system is represented by Fig. 1.2.

**Fig. 1.2. Functional model of the *B. thetaiotaomicron* starch utilization system (Sus) (68).**

Steps involved in processing starch are sequentially illustrated and numbered. Step 1, starch molecules transit through the capsule layer. Step 2, starch molecules are bound by outer membrane complex components such as SusD. Step 3, surface bound starch molecules are degraded and processed by outer membrane complex component glycohydrolase such as SusG, generating smaller oligosaccharides through the outer membrane porin SusC. Step 4, oligosaccharides are further degraded into component mono- or disaccharides by periplasmic glycan degrading enzymes SusA and SusB. Step 5 and 6, liberated glycan components serve as signal molecules for the transcriptional activator SusR that activate *sus* operon expression. Step 7, degraded sugar units are imported in to the cytoplasm serving as a carbon/energy source.


1. Starch: α1,4 and α1,6 glucan with characteristic secondary structure
capsular polysaccharide layer
2. glycan binding proteins
3. cell surface glycan-degrading enzymes
SusF
SusE
SusD
SusG
outer membrane
SusC
4. oligosaccharide degradation by periplasmic enzymes
5. cis-acting, sensor/regulator controls expression of polysaccharide utilization locus (PUL)
periplasm
P
SusA
SusB
H+
SusR
cytoplasmic membrane
TonB-ExbBD complex
H+
7. transport and fermentation of small saccharides
cytoplasm
6. activation of susABCDEFG

The *B. thetaiotaomicron* Sus system serves as a paradigm for all the Sus-like PULs in *Bacteroides*. These PULs encode similar outer membrane protein complexes for glycan utilization and are identified by sequence homology. Most PULs contain the SusC and SusD homolog with or without the presence of other downstream genes encoding outer membrane complex proteins, but in limited cases, only the SusC/SusD-like protein pair is present in the locus (7, 13). In contrast to the archetypical *B. thetaiotaomicron* Sus system, the most frequently observed genetic regulators of the PULs in *B. fragilis* are not SusR homologs but rather they are ECF sigma factor/anti sigma pairs, and in some cases two component regulatory systems.

Although there is a wide range of glycan substrates used by the Sus-like systems, each system generally only has one specific type of glycan as the substrate (68), and the Sus-like systems are not only limited in *Bacteroides* species but seen in the *Bacteroidetes* phylum (68). A few well-studied Sus-like systems are listed in Table 1.1.

The traditional catabolite repression/activation system which allows for the select utilization of a preferential carbon/energy source does not function in *Bacteroides* as they do not possess cyclic AMP (69). Although multiple Sus-like systems can be induced at the same time when their substrates are present, a recent study showed that there is a hierarchy in the expression of these Sus-like systems (22). The mechanism of how *Bacteriodes* define the hierarchy is not clear. Interestingly, there also appears to be a link between the expression of some specific Sus-like systems and polysaccharide capsule biosynthesis, but the mechanism is still not fully understood (62).

**Table 1.1. Sus-like systems in the *Bacteroidetes***

| Name | Species | Substrate | Special note | Ref. |
|---|---|---|---|---|
| Sus | *B. thetaiotaomicron* | Starch | Paradigm for all Sus-like systems | (68) |
| XyGULs | *B. ovatus* | Xylan glycans | Break down plant cell walls | (124) |
| Don | *B. fragilis* | Transferrin and other serum glycoproteins | Extraintestinal growth factor | This study |
| CCF | *B. fragilis* | Unknown host glycan | Commensal colonization factor | (98) |
| Osu | *B. fragilis* | Starch | Oxygen inducible | (59) |
| Gpd | *C. canimorsus* | Human IgG glycan | A Sus-like system in a *Flavobacteria* | (93) |
| Csu | *B. thetaiotaomicron* | Glycosaminoglycan | Enable the bacterium to grow on chondroitin sulfate and hyaluronic acid | (1) |

# CHAPTER TWO: EFFICIENT UTILIZATION OF COMPLEX N-LINKED GLYCANS IS A SELECTIVE ADVANTAGE FOR BACTEROIDES FRAGILIS IN EXTRAINTESTINAL INFECTIONS

## 2.1 Abstract

*Bacteroides fragilis* is the most common anaerobe isolated from clinical infections and in this report we demonstrate a novel feature of the species that is critical to their success as an opportunistic pathogen. Among the *Bacteroides spp.* in the gut, *B. fragilis* has a unique ability to efficiently harvest complex N-linked glycans from the glycoproteins common to serum and serous fluid. This activity is mediated by a Sus-like outer membrane protein complex designated as Don. Using the abundant serum glycoprotein transferrin as a model it was shown that *B. fragilis* alone can rapidly and efficiently deglycosylate this protein *in vitro* and that transferrin glycans can provide the sole source of carbon and energy for growth in defined media. We then showed that transferrin deglycosylation occurs *in vivo* when *B. fragilis* is propagated in the rat tissue cage model of extraintestinal growth and that this ability provides a competitive advantage *in vivo* over strains lacking the *don* locus.

## 2.2 Significance

The human microbiota has a huge impact on health from the proper development of the immune system to the maintenance of normal physiological processes. The largest concentration of microbes is found in the colon which is home to more than 500 bacterial species most of which are obligate anaerobes. This population also poses a significant threat of opportunistic infection and of all the species present, *Bacteroides fragilis* is the one most frequently isolated from anaerobic, extraintestinal infections. New findings presented here describe a unique ability of this species to efficiently deglycosylate complex N-linked glycans from the most abundant glycoproteins found in serum and serous fluid. This provides *B. fragilis* a competitive, nutritional advantage for extraintestinal growth.

## 2.3 Introduction

The genus *Bacteroides* are Gram-negative, obligate anaerobic bacteria that account for approximately 30% of the microbiota in the human large intestine (1, 17). This relationship benefits the host by aiding in the development of a healthy immune response and in the maintenance of many physiological and nutritional processes. One major factor that contributes to the predominance of *Bacteroides* is their capacity to utilize a wide spectrum of polysaccharides ranging from dietary compounds that cannot be digested by the host, such as xylans or pectins, to the host derived glycans in the form of glyco-conjugates (19, 20). Polysaccharide digestion in the *Bacteroides* is mediated in large part by novel outer membrane complexes that bind, cleave and transport these substrates. These systems have a similar

genetic organization characterized by a regulatory region followed by an operon coding for orthologues of the TonB-dependent transporter, SusC, and the accessory binding protein SusD. Additional genes in the operon are for specific substrate binding and glycohydrolase enzymes (68). The genes for these protein complexes are termed polysaccharide utilization loci (PULs).

In contrast to their beneficial role in the colon, *Bacteroides* species also are the most common opportunistic pathogens isolated from clinical specimens of anaerobic infections. These opportunistic infections can occur when the integrity of the intestinal wall becomes compromised. Predisposing conditions include intestinal surgery, perforated or gangrenous appendix, carcinoma, diverticulitis, trauma and inflammatory bowel diseases. Peritonitis and intra-abdominal abscesses are the most common infections associated with *Bacteroides* often leading to bacteremia (70). Notably, *Bacteroides fragilis* represents only 0.5% of the gut flora yet it is isolated in the majority of anaerobic infections, thus it is regarded as more invasive than the other *Bacteroides* (71-73). A number of factors have been identified that may contribute to its enhanced extraintestinal survival. For example, there is a variable polysaccharide capsule which interferes with immune surveillance and there is a robust oxidative stress resistance system but these attributes are not exclusive to *B. fragilis* (43, 52, 55). Another challenge faced by organisms invading from the colon is the need to adapt to different nutritional sources. In the gut, the *Bacteroides* rely on dietary polysaccharides and host derived glycans as sources of carbon/energy but when outside of the gut environment they will have more limited choices of only host derived glycans. The ability to efficiently harvest glycans present on host glycoproteins in extraintestinal sites may enhance survival and colonization.

In this report we describe a Sus-like PUL which is unique to *B. fragilis* and allows it to efficiently utilize complex N-linked glycans from the most abundant serum/serous fluid glycoproteins including transferrin. Overall the studies suggest that the ability to harvest these glycans is advantageous for growth at extraintestinal sites and can explain, in part, the success of *B. fragilis* as an opportunistic pathogen relative to the other *Bacteroides* species.

## 2.4 Materials and methods

### 2.4.1 Bacterial strains and growth.

Bacterial strains and plasmids used in this study are listed in Table 2.1. *B. fragilis* 638R was the wild type strain used for genetic analyses. *Bacteroides* strains were grown in an anaerobic chamber in Brain Heart Infusion broth supplemented with hemin and cysteine (BHIS) (55). Rifampicin (20 µg/ml), gentamicin (100 µg/ml), erythromycin (10 µg/ml) and tetracycline (5 µg/ml) were added as indicated. Minimal defined media (DM) were prepared as described previously with the specific carbon/energy sources described in the text (74). Mucin glycans were prepared by proteolysis of porcine gastric mucin (Sigma-Aldrich, Cat. M2378) followed by alkaline β-elimination to release free glycans (20). In experiments using DM-transferrin, the transferrin was >98% iron saturated holo-transferrin and the transferrin-containing media (and controls) contained 100 µM $FeSO_4$ to ensure a readily available source of iron for growth.

Table 2.1 Bacterial strains and plasmids used in this study.

| Bacterial strain or plasmid | [a] Description | Reference or source |
|---|---|---|
| *Bacteroides* strains | | |
| IB101 | *B. fragilis*, 638R, clinical isolate, $Rf^r$ | (75) |
| IB555 | 638R *Δdon*, BF638R3439-3443 genes replaced with a *tetQ* cassette, $Tc^r$, $Rf^r$ | This study |
| IB114 | *B. fragilis*, ATCC 25285, clinical isolate, $Rf^r$ | (55, 76) |
| IB116 | *Bacteroides thetaiotaomicron*, VPI strain 2302, $Rf^r$ | (77) |
| BT5482R | *Bacteroides thetaiotaomicron*, VPI strain 5482 (ATCC 29148), $Rf^r$ | (77) |
| IB102 | *Bacteroides uniformis*, VPI strain 006-1 (ATCC8492), $Rf^r$ | |
| IB103 | *Bacteroides ovatus*, VPI 0038 (ATCC 8483) | (76) |
| IB351 | *Bacteroides vulgatus*, ATCC 8482 | (76) |
| BER37 | *Parabacteroides distasonis*, clinical isolate CLA348 | (78) |
| Ber39 | *Parabacteroides merdae*, ATCC 43185 | |
| *E. coli* strains | | |
| DH10B | *E. coli*, $F^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K rpsL *nup*G $\lambda^-$ | Invitrogen |
| HB101::RK231 | *E. coli*, HB101 containing RK231, $Kn^r$, $Tc^r$, $Sm^r$ | (79) |
| Plasmids | | |
| pFD516 | *Bacteroides* suicide vector, 7.7 kb, $Sp^r$ in *E. coli*, $Em^r$ in *Bacteroides* | (80) |
| pFD516/omp117±2K/tetQ | pFD516 containing the BF638R3439-3443 deletion construct with the *tetQ* gene cassette, 14.6 kb, $Sp^r$ in *E. coli*, $Em^r$ and $Tc^r$ in *Bacteroides*. | This study |

[a] Rf, rifampicin; Tc, tetracycline; Sp, spectinomycin; Em, erythromycin; Kn, kanamycin; Sm, streptomycin.

Table 2.2. Oligonucleotides used in this study.

| Name | Sequence ( 5' → 3') | Tag | Description |
|---|---|---|---|
| UF | AGTCCTGCAGCAAACAGATGCCTTATCAGTTT | PstI | Designed to amplify 2 kb upstream of *donC* |
| UR | AGTCGGATCCCAGTGTCTTATCCAGAGATTCA | BamHI | |
| DF | AGTCGGATCCCTGGACAATATCGATTCAAGCA | BamHI | Designed to amplify 2 kb downstream for *donG* |
| DR | AGTCGAGCTCATAGCGACCAAAACCCTTCT | SacI | |
| Omp117rtL | GGTGAAGGCATTTCCGACTT | | Designed to amplify a 140 bp fragment of *donC* gene for quantitative PCR |
| Omp117rtR | TTGCCTTCCTGCCCTTTCTT | | |
| 16srL | GATGCGTTCCATTAGGTTGTTG | | Designed to amplify a 127 bp fragment of 16s ribosomal RNA gene for quantitative PCR |
| 16srR | CACTGCTGCCTCCCGTAG | | |
| PsigOKL2 | AATCACAATCAGCCTTATATTCTAC | | Designed to amplify a 375 bp fragment spanning *donA* and its potential TIS region |
| PsigOKR | CATTGCTCCTGATAGGTCTG | | |
| PantisigOKL | GGATATTTCGATCCGCACTG | | Designed to amplify a 298 bp fragment spanning the intergenic region of *donA* and *donB* |
| PantisigOKR | TTTCCTTCCGTTTCGTTCCA | | |
| Pomp117L | AGATATCCTGAACATTATGCAGG | | Designed to amplify a 339 bp fragment spanning the intergenic region of *donB* and *donC* |
| Pomp117R | GCCGTTTGTAGAATAAAACAGC | | |
| PsusDL | AATGACCTCTTCGACAGGTA | | Designed to amplify a 220 bp fragment spanning the intergenic region of *donC* and *donD* |
| PsusDR | GGTTCCGGTTGATATCTTCG | | |
| PendoSL | ATTGATCGGAAACGGATACG | | Designed to amplify a 240 bp fragment spanning the intergenic region of *donD* and *donE* |
| PendosR | CTACCAACAATGCTACTCCG | | |
| PsusEL | GAAGACGCATTTGAAGACGA | | Designed to amplify a 261 bp fragment spanning the intergenic region of *donE* and *donF* |
| PsusER | GCTCTAAAGTGAAGTTGACGA | | |
| PsusFL | GAGACCAAAGACCTGAAAGC | | Designed to amplify a 288 bp fragment spanning the intergenic region of *donF* and *donG* |
| PsusFR | GAGTTGCTCACGGATATCTT | | |

### 2.4.2 Construction of a *don* operon deletion mutant.

The *don*CDEFG genes were replaced with a *tetQ* tetracycline resistance (Fig. S1). Chromosomal fragments of about 2 kb flanking the *donC* and *donE* genes were amplified using PCR primer pairs UF/UR and DF/DR respectively (Table 2.2). The amplified DNA was cloned into the *Bacteroides* suicide vector pFD516 together with the *tetQ* gene cassette in a three-fragment ligation. The recombinant plasmid was mobilized into *B. fragilis* 638R by conjugation and exconjugants were selected on BHIS plates containing rifampicin, gentamicin and tetracycline. Candidate mutants were screened for sensitivity to erythromycin and by PCR to identify the double-crossover allelic exchange.

### 2.4.3 Animal model of infection.

The tissue cage infection model has been described previously (60, 81). Briefly, a perforated, sterilized Ping-Pong ball was surgically implanted into the peritoneal cavity of an adult male Sprague Dawley rat and allowed to encapsulate for 4-5 weeks. The ball becomes encased in connective tissue, develops a blood supply, and fills with sterile serous fluids (~25ml per ball). *B. fragilis* strains grown overnight in BHIS media were diluted in PBS buffer (50 mM $PO_4$, 150 mM NaCl, pH 7.4,) to $10^5$ CFU/ml and 4 ml of this suspension was injected into the tissue cage. Samples were aspirated at the indicated time points for viable cell counts and RNA extraction. For viable cell counts aliquots were serially diluted in PBS buffer, plated on BHIS plates and incubated in an anaerobic chamber to determine CFU. Samples were also plated on LB agar plates and incubated aerobically to check for contamination.

The *in vivo* competition assay was performed by mixing overnight cultures of 638R and *Δdon* at a one to one ratio followed by dilution in PBS buffer to a total viable cell count of $10^5$ cfu/ml. Four ml of the cell mixture was injected into the tissue cage of each animal and aliquots were plated to determine CFU/ml. Diluted samples were plated on BHIS plates with rifampicin and gentamicin. After 4-5 days incubation, 100 to 300 colonies from each sample were picked to BHIS plates with and without tetracycline to check the resistance phenotype and determine the ratio of mutant to wild type.

All procedures involving animals followed the National Institutes of Health guidelines and were approved by the Animal Care and Use Committee of East Carolina University. For each bacterial strain tested two trials of 5 animals each were performed.

### 2.4.4 Total RNA extraction.

RNA extraction from *in vivo* samples was done as described previously with a few modifications (60). Fluid was aspirated and immediately mixed with RNAprotect Cell Reagent (Qiagen, Inc.) at a 1:2 ratio. Next, 0.1% sodium deoxycholate was used to lyse host cells, then bacterial cells and debris were collected and washed in PBS containing RNAprotect Cell Reagent and sodium deoxycholate. Total RNA was extracted from the cell pellet using the hot phenol method described previously (82) and stored in 50% formamide at -80°C. RNA extraction from *in vitro* bacterial cell cultures was performed on washed cell pellets using the same hot phenol method.

### 2.4.5 cDNA synthesis, qRT-PCR, and microarray analysis.

Total RNA was purified using the RNeasy Mini Kit (Qiagen, Inc.) and DNA was removed by treatment with DNase (Ambion/Life Technologies Inc.). DNA contamination was determined by PCR using specific for the 16s ribosomal RNA gene (Table 2.2). First strand cDNA synthesis was carried out using 1 µg total RNA with random hexamer primers and Superscript III RT (Life Technologies, Inc.). For qRT-PCR, primer pair, omp117rtL and omp117rtR, were used to amplify a 140 bp fragment of the *donC* gene in a standard reaction mixture with SYBR$^{R}$ Green Supermix (Bio-Rad, Inc.). All sample reactions were in run in triplicate and RNA with no reverse transcriptase was used as a control to monitor for DNA contamination. Relative expression values were normalized to 16s rRNA and calculated by the method of Pfaffl (83). Results represent at least two independent experiments performed in triplicate.

Expression microarray analyses were performed essentially as described previously (58). Double stranded cDNA was synthesized with the SuperScript$^{R}$ Double-Stranded cDNA Synthesis Kit (Life Technologies, Inc.). One µg of purified double stranded cDNA was then labeled with cy3, hybridized to microarray slides, and processed by the Florida State University Roche/NimbleGen Microarray Facility. For each experimental condition at least two independent trials were performed. Each trial consisted of a high-density-oligonucleotide whole genome expression microarray (Roche/NimbleGen, Madison WI) with 8 technical replicates of each probe per slide. RNA for each trial was pooled from five rats prior to cDNA synthesis. Raw microarray expression data were normalized by using the RMA algorithm implemented with the

Roche DEVA 1.1 software. The microarray expression data have been deposited in the NCBI Gene Expression Omnibus database under accession number GSE53883.

### 2.4.6 Whole cell deglycosylation assay.

Cultures were grown in 5 ml of DM-mucin glycan, harvested by centrifugation at mid-logarithmic phase, washed and the cells were suspended to an OD $A_{550}$ of 1.0 in PBS. 100 µL of the cell suspension was incubated with 100 µL of tissue cage serous fluid (1/10 diluted), purified human transferrin (1 mg/ml in PBS), IgG (1 mg/ml in PBS) or IgA (1 mg/ml in PBS) as described in the text. All proteins were from Sigma Aldrich Inc. For controls 100 µL PBS replaced the bacterial cell suspension. Assays were incubated anaerobically at 37°C for 3 hours or overnight, then samples were centrifuged, supernatants collected, mixed with loading buffer and electrophoresed on 12% SDS-PAGE gels. The gels were analyzed by Coomassie blue staining or proteins were transferd to PVDF membranes for glycan determination using *Sambucus nigra* agglutinin according to the manufacturer's instructions (Roche, DIG Glycan Differentiation Kit).

### 2.4.7 Glycomics analyses of *B. fragilis* treated human transferrin.

Purified human transferrin was suspended in PBS to 2 mg/ml and 100 µL of this was used in deglycosylation assays as described above except that following incubation the assay supernatants were filter sterilized and frozen at -80°C. Samples were analyzed by Ezose Sciences Inc. (Pine Brook, NJ) to quantify *N*-linked glycans using previously reported methods

(84, 85). Briefly, samples were denatured, digested with trypsin and then heat-inactivated. The *N*-glycans were then enzymatically released from the peptides by treatment with PNGase F, captured on chemoselective beads and then processed for MALDI-TOF mass spectrometry. Mass spectra were analyzed using Ezose's proprietary bioinformatics programs. Results were normalized to 1 g/L of transferrin and reported as the average of two independent experiments.

## 2.5 Results

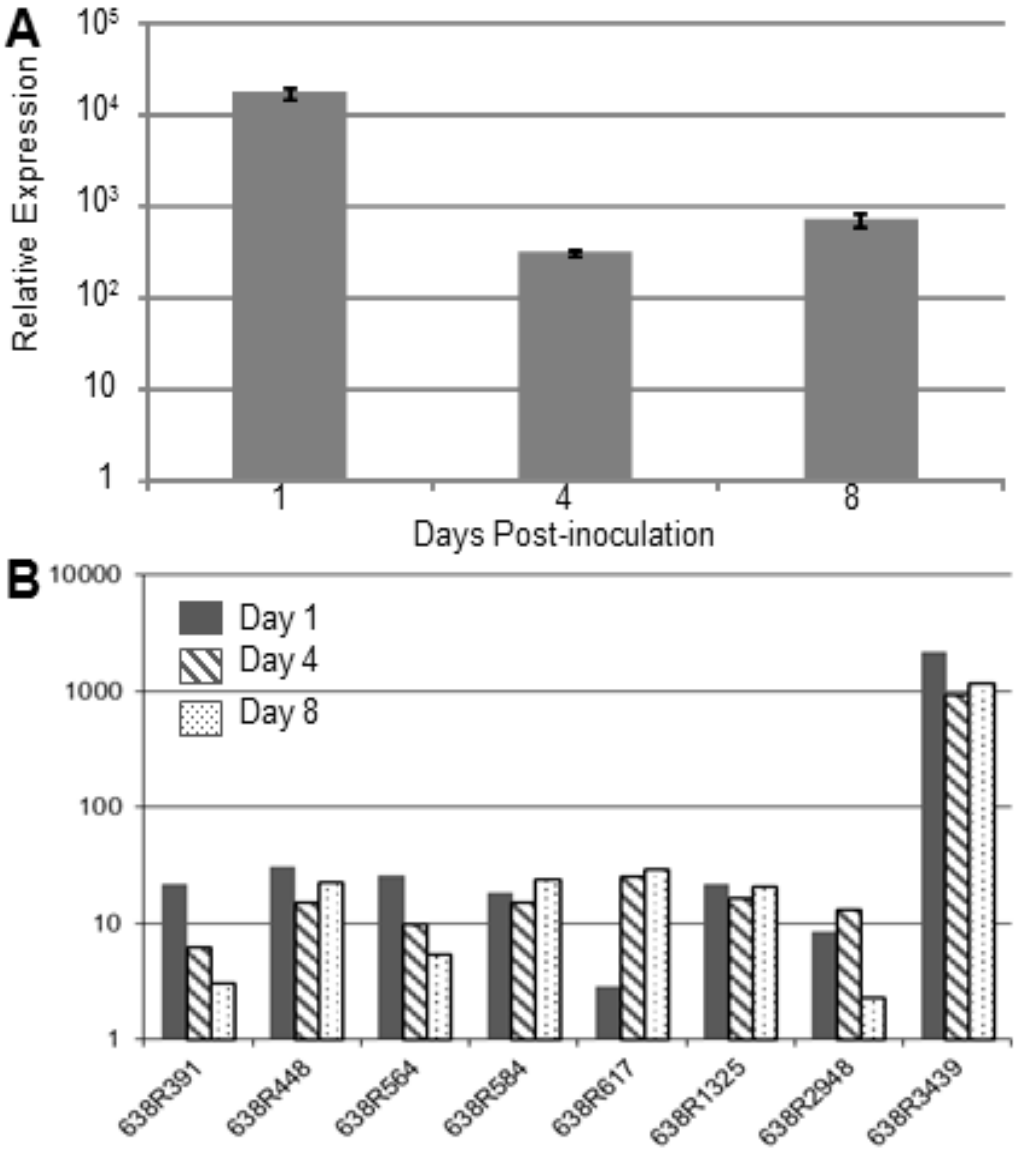
### 2.5.1 *In vivo*, extraintestinal PUL expression.

Considerable progress has been made in understanding how the *Bacteroides* compete for and utilize limited nutrients in the gut ecosystem. In contrast, there is a paucity of knowledge concerning *Bacteroides* nutrient foraging in extraintestinal environments. As an initial strategy to identify significant pathways of catabolism, microarrays were used to measure gene expression of *B. fragilis* cells growing in a rat tissue cage abscess model. The results suggested that the organism may rely on a variety of PULs to harvest glycans from host proteins present in the serous fluid of the model abscess. *B. fragilis* has 47 putative PULs and 22 stand-alone SusC/SusD-like protein pairs and several of these were among the most highly induced genes *in vivo*. A five gene operon, BF638R3439-43 associated with one of these PULs was induced from 2147- to 765-fold along the length of the 8 kb operon. These were the most highly induced genes *in vivo* and the genes in this PUL have been designated as *donABCDEFG* (Fig. S1). To more easily compare expression of all PUL operons, the SusC orthologues were

used as proxy for expression of the entire operon (22). This approach was validated by qRT-PCR which confirmed the *in vivo* induction of *donC* (Fig. 2.1A). The results from day 1, 4, and 8 samples were compared to *in vitro* grown mid-logarithmic phase cultures and showed a high level of induction throughout the course of the experiment.

Examination of the SusC orthologue expression data by *k*-means clustering revealed four major *in vivo* PUL gene induction patterns (Appendix 2, Fig. S2). The *donC* gene was associated with the most highly induced genes, Cluster 1, which consisted of 8 PULs and 5 stand-alone SusC/SusD orthologues. As shown in Fig. 2.1B the PULs in this cluster were rapidly induced more than 10-fold and the level of induction remained high throughout the 8 days. A second cluster of 25 SusC orthologues was induced more gradually and remained induced 1.5-5-fold for the 8 days (Fig. S2B). The remaining two PUL clusters were repressed in the abscess model and may be of less importance outside of the intestinal environment.

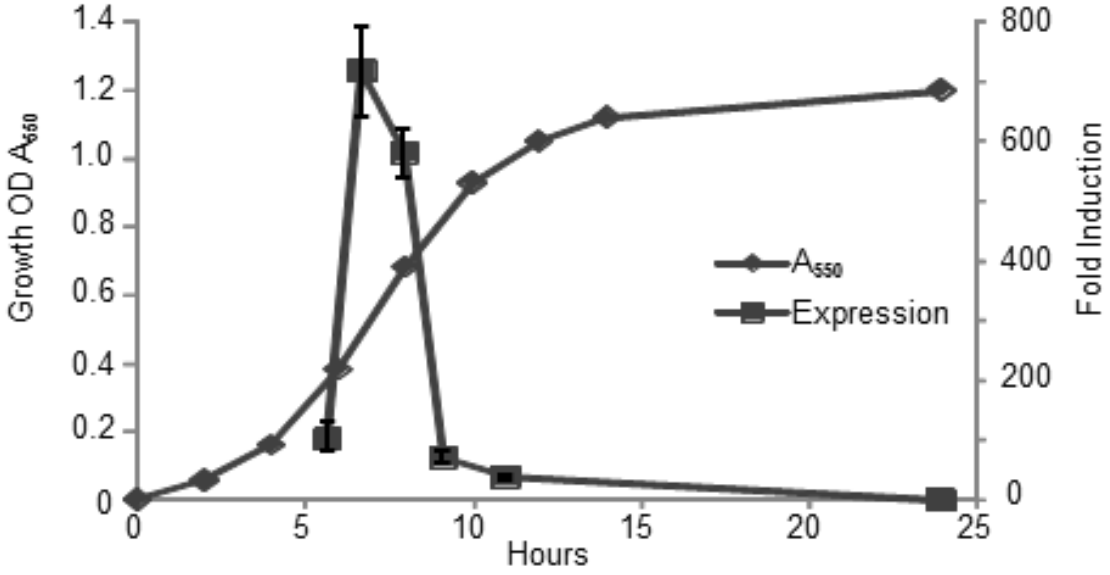
**Fig. 2.1.** *In vivo* PUL gene expression during growth of *B. fragilis* in the rat tissue cage model. Samples for analyses were pooled from 5 animals 1, 4, and 8 days post inoculation. **A)** Expression of *donC* was determined by qRT-PCR for samples from the rat tissue cage and mid-logarithmic cultures grown in DM-glucose. The results are the average of triplicate samples and are presented as the *in vivo/in vitro* ratio. **B)** Induction of SusC orthologues during growth *in vivo* relative to DM-glucose. Shown are the highly induced-*in vivo susC*-like genes associated with PULs in Cluster 1. Clusters were determined from expression microarray data analyzed by *k*-means clustering using the Standard Pearson's correlation coefficient distance metric. All induction values were significant at $p<.01$.


A
Relative Expression
$10^5$
$10^4$
$10^3$
$10^2$
10
1
1
4
8
Days Post-inoculation
B
10000
1000
100
10
1
Day 1
Day 4
Day 8
638R391
638R448
638R564
638R584
638R617
638R1325
638R2948
638R3439

### 2.5.2 *In vitro* induction of the *donCDEFG* operon.

The *don* PUL was chosen as a model for extraintestinal PUL function and it included several gene products that may be associated with N-linked glycan utilization such as an endo-S-like endo-N-acetyl-β-D-glucosaminidase (DonE) and a lectin-like product with an α-N-acetylglucosaminidase domain (DonF). In order to demonstrate a function of the PUL it was necessary to find an inducing substrate for *in vitro* expression studies. This was accomplished by using a crude glycan mixture prepared by alkaline β-elimination hydrolysis of porcine stomach mucin. The majority of mucin glycans are O-linked glycans and 10-20% are N-linked glycans (20). A defined medium containing the glycan mixture as the sole carbon/energy source supported rapid growth of the wild type strain BF638R (Fig. 2.2). Expression analysis by qRT-PCR showed that *donC* was strongly induced in this medium in a growth phase dependent manner with the highest level of induction at mid-logarithmic phase. This RNA also was used for RT-PCR with primers pairs that spanned the intergenic regions of the PUL and the data indicated linkage between the *donBCDEFG* genes (Fig. S1).

**Fig. 2.2.** Induction of *donC* expression during growth on mucin glycans. Strain BF638R was grown in defined medium containing 2% porcine gastric mucin glycans as the sole carbon/energy source. The induction of *donC* by mucin glycans relative to growth on glucose was determined by qRT-PCR. The results are overlaid on a typical growth curve with mucin glycans (OD $A_{550}$, grey diamonds). The qRT-PCR results represent 2 independent experiments (grey squares) with standard deviations shown.


Growth OD A550
Fold Induction
Hours
A550
Expression
0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
0
200
400
600
800
0
5
10
15
20
25

### 2.5.3 Deglycosylation of transferrin is mediated by *donCDEFG*.

Don PUL function was explored by incubating serous fluid from the rat tissue cage (obtained prior to inoculation) with wild type or Δ*donCDEFG* (Δ*don*) mutant cells grown to mid-logarithmic phase in the inducing, mucin glycan medium. Samples were analyzed by SDS-PAGE and the results in Fig. 2.3A show that one of the major fluid proteins was significantly reduced in size in samples treated with the wild type strain but not with the deletion mutant. Analysis by mass spectrometry revealed that this protein was rat serum transferrin, one of the most abundant glycoproteins in serum. The precise size reduction of the protein suggested the possibility that the transferrin glycans were removed. To test this, human transferrin was incubated with the wild type or Δ*don* mutant cells. Duplicate samples were analyzed by SDS-PAGE with either Coomassie blue or *Sambucus nigra* (SNA) lectin staining. As shown in Fig. 2.3B transferrin treated with the wild type strain was reduced in size after the 3 h incubation whereas the deletion strain did not have an obvious effect on transferrin size. Also, no degradation or reduced intensity of the protein band was observed in any of the samples. The SNA stained gels showed that transferrin treated with wild type cells was completely deglycosylated but in contrast an intense signal remained in samples treated with the deletion mutant and controls (Fig. 2.3B). In addition to transferrin, other serous fluid proteins were rapidly deglycosylated by *B. fragilis*. In assays using serous fluid samples the wild type strain had significant deglycosylation activity on many serous fluid proteins but the Δ*don* mutant activity was considerably reduced (Fig. S3). Although the mutant clearly retained some activity, the overall results indicate that *B. fragilis don* coded proteins play an important role in the

deglycosylation of N-linked glycans on transferrin and other glycoproteins. These genes were designated *don* for **D**eglycosylation **O**f **N**-linked glycans.

The SNA lectin recognizes terminal sialic acids linked to Gal or GalNAc which is suitable to detect N-linked glycans but is not a definitive assay. Conceivably loss of the terminal sialic acids by the action of neuraminidase could result in loss of SNA staining. Therefore a comprehensive mass spectrometric glycomics analysis of transferrin was performed. Human transferrin has two glycosylation sites and there were seven distinct oligosaccharides linked to those sites, five of which made up 98.1% of the total transferrin glycans (Table 1). Incubation of transferrin with wild type cells for 3 h resulted in essentially complete deglycosylation of N-linked glycans. By comparison, nearly 70% of the N-linked glycans remained in samples incubated with the Δ*don* mutant. There were no new oligosaccharide structures detected in reactions with the wild type strain suggesting efficient cleavage of the glycan likely at the β-1,4-di-N-acetylchitobiose core by the endo-β-N-acetylglucosaminidase S (DonE). These results confirm that transferrin is a substrate for the Don PUL but also indicate there is a second, less efficient deglycosylation system for this substrate.

**Fig. 2.3.** Deglycosylation of transferrin is mediated by the *don* PUL. **A)** Serous fluid obtained from tissues cages prior to inoculation was incubated for 3 h with PBS or *B. fragilis* cells induced by growth in DM-mucin glycan medium. Samples were analyzed by 12% SDS-PAGE and Coomassie blue staining (Co). The arrowhead indicates the location of transferrin above the abundant serum albumin protein. **B)** Deglycosylation analysis of human transferrin. Human transferrin was incubated with wild type or *Δdon* cells induced by growth in DM-mucin glycan medium. Samples were analyzed by SDS-PAGE and followed by SNA glycan-staining (SNA) to detect N-linked glycans as described in the text.

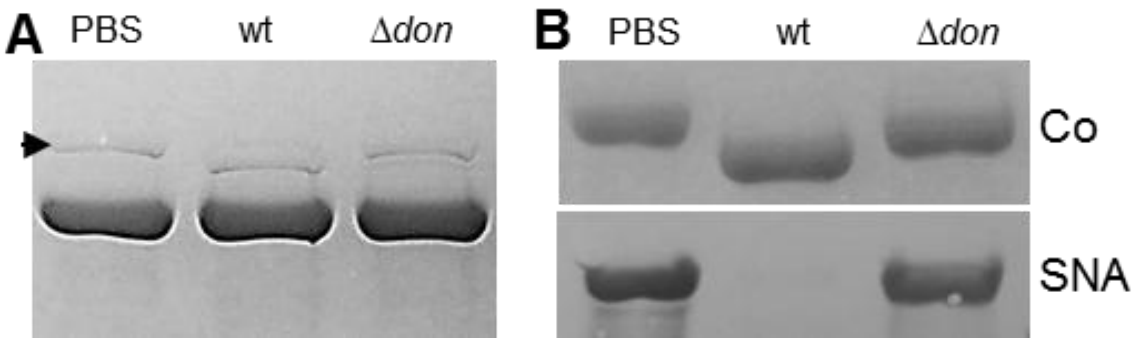
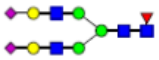
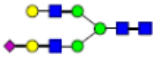
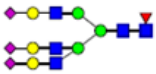

A
PBS
wt
Δdon
B
PBS
wt
Δdon
Co
SNA

Table 2.3. Quantitative glycan analysis of transferrin incubated with *B. fragilis* cells.

| | | [a] Concentration of glycan (μM) | | |
|---|---|---|---|---|
| [b]Glycan | % Total | PBS | wt | Δ*don* |
|  | 87.5% | 15.86 ±0.32 | ND | 10.43 ±1.06 |
|  | 4.0% | 0.72 ±0.03 | ND | 0.25 ±0.01 |
|  | 3.3% | 0.60 ±0.02 | ND | 0.71 ±0.11 |
|  | 1.9% | 0.34±0.01 | 0.05 | 0.27 ±0.07 |
|  | 1.4% | 0.25 ±0.01 | ND | 0.21 ±0.04 |
| Totals | 98.1% | 17.77±0.4 | 0.05 | 11.87 ±1.11 |

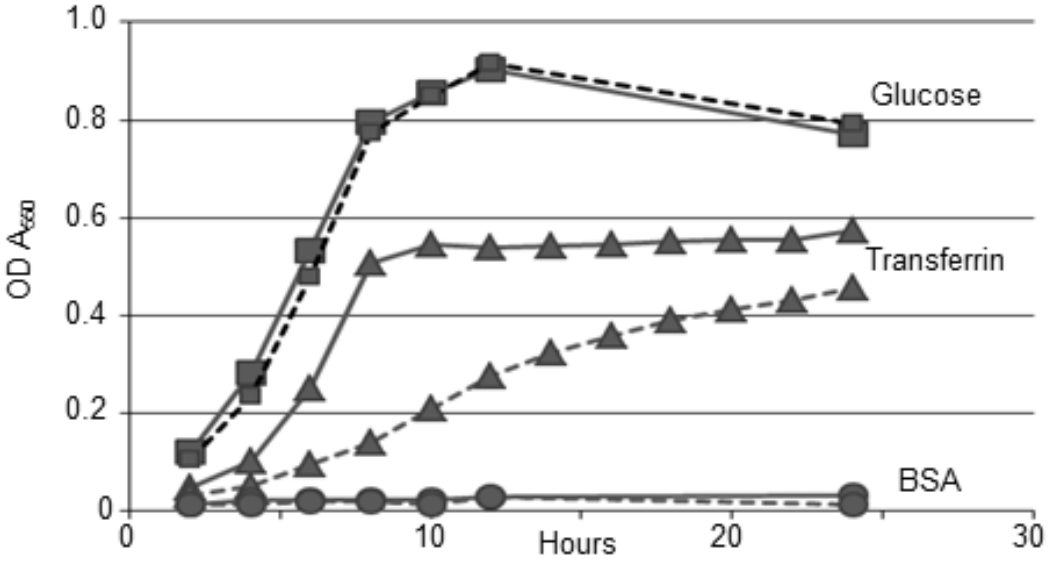
[a] Transferrin was treated with PBS, strain 638R (wt), or the Δ*don* mutant for 3 hours as described in the text. Results are normalized to 1 g/L of transferrin and are the average of 2 independent samples. ND=not detected.

[b] The five major glycans on transferrin are shown and were composed of the following sugars: 

### 2.5.4 Growth on transferrin as the sole carbon and energy source.

The ability to forage for high quality carbohydrates in extraintestinal sites would be advantageous to *Bacteroides spp.* translocated from the gut. Transferrin is the most abundant glycoprotein present in serum and is estimated to have a carbohydrate content of 5.8% so we tested if it could support the growth of *B. fragilis*. As shown in Fig. 2.4, 25 mg/ml of transferrin supported robust growth of the wild type strain but not the Δ*don* mutant. There was no significant difference in the growth rate between wild type and mutant strains grown with glucose (Fig. 2.4). Expression of the *don* operon during growth on transferrin was measured by qRT-PCR with *donC* primers and the results showed a $10^4$ –fold induction over glucose grown cells. SDS-PAGE analysis of the transferrin medium after 24 or 48 hours of growth indicated there was no degradation of the transferrin peptide. The non-glycosylated protein, bovine serum albumin, did not support growth (Fig. 2.4). Interestingly, the Δ*don* mutant was able to grow in the transferrin medium albeit at a much slower rate. This confirms the presence of a second system that enables slow utilization of transferrin glycans.

**Fig. 2.4.** *B. fragilis* can grow with transferrin as the sole source of carbon and energy. Growth curves are shown for wild type or *Δdon* strains in defined media with different carbon/energy sources. An overnight inoculum of 2% was used and the OD $A_{550}$ was measured at specific time intervals. Dashed lines are *Δdon* and solid lines are wild type. Squares, glucose (0.4%); triangles, human transferrin (25mg/ml); circles, bovine serum albumin (25 mg/ml). Each growth curve represents three biological repeats.


1.0
0.8
0.6
0.4
0.2
0
OD A600
0
10
20
30
Hours
Glucose
Transferrin
BSA

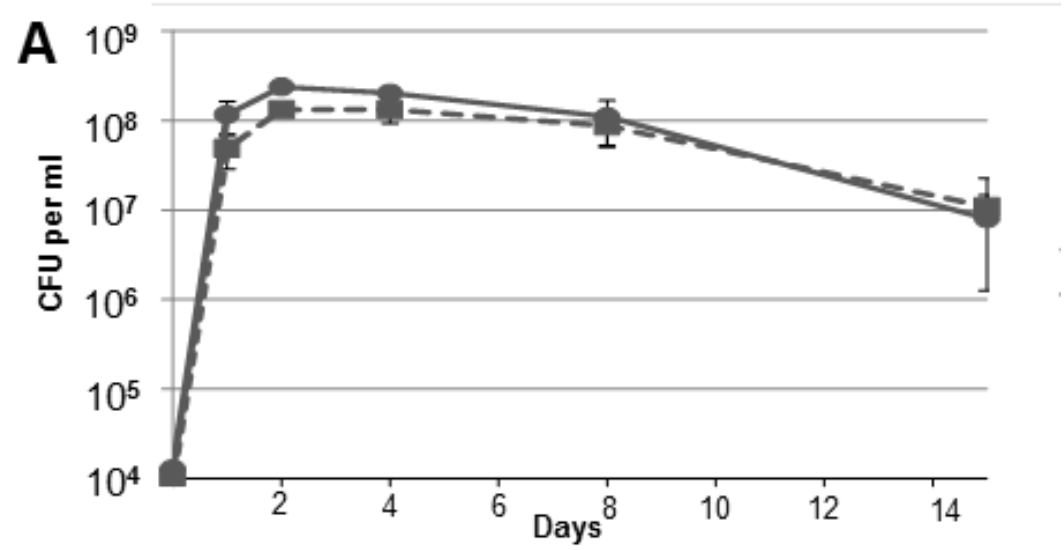
### 2.5.5 *In vivo* role of the *don* PUL.

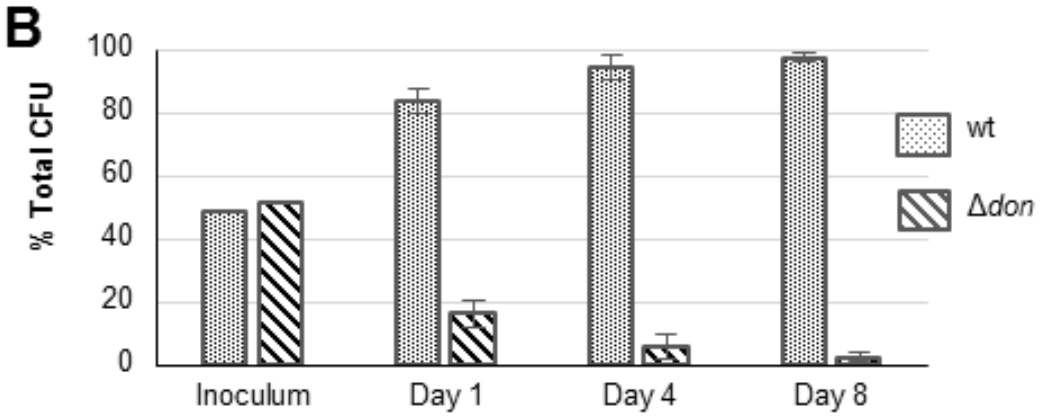
To investigate a potential role for the *don* operon in extraintestinal sites, the rat tissue cage model was used to measure growth *in vivo*. There was not a drastic difference in CFU counts between wild type and Δ*don* strains during monoculture experiments over the course of a 15 day period (Fig. 2.5A). There was however a small, two-fold advantage for the wild type in the first two days following inoculation. To determine if this difference would translate into a competitive advantage, mixed culture assays were performed using wild type and mutant strains co-inoculated into rat tissue cages. The results were clear, Δ*don* was quickly out-competed and at one day post-inoculation about 85% of the total population was wild type (Fig. 2.5B). The same trend carried through day 8 when the wild type strain reached about 97% of the total population. In control experiments (Fig. S4) with strains co-inoculated into glucose defined media there was no significant difference between the percentage of strains in the population through logarithmic phase and in fact the Δ*don* mutant had a slight advantage in stationary phase. These results suggest that access to novel carbohydrate nutrient sources *in vivo* can be advantageous for survival at extraintestinal sites.

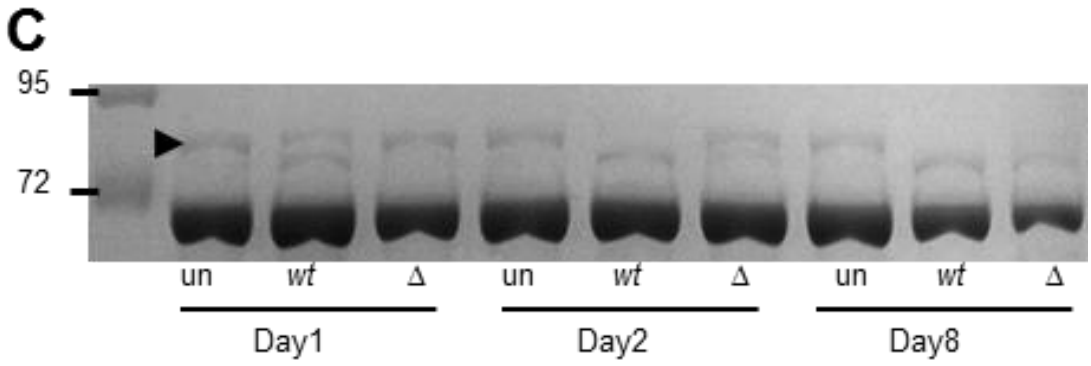
One important extraintestinal carbohydrate source for growth *in vivo* is transferrin which was rapidly deglycosylated during growth in the rat model. This was demonstrated by SDS-PAGE analysis of fluids removed from rat tissue cages over the course of the eight day experiment (Fig. 2.5C). The transferrin from uninoculated controls appeared as a single peptide species of about 80 kDa but by day one post-inoculation there were two transferrin peptides in samples from wild type infected animals and by day two nearly all of the transferrin was

converted to the smaller species. In contrast, samples from animals inoculated with the Δ*don* strain appeared unchanged at day one and by day two only a small portion of the transferrin was deglycosylated. Interestingly all of the transferrin in samples from wild type or mutant infected animals was deglycosylated by the end of the experiment on day eight (Fig. 2.5C). This result shows that the *don* locus plays an important role in the efficient deglycosylation of transferrin *in vivo*.

**Fig. 2.5.** Growth *in vivo* is enhanced by deglycosylation of transferrin. **A)** Growth curve for wild type (circles) and *Δdon* (squares) strains inoculated separately into rat tissue cages. **B)** Mixed culture competition assay. Mixtures of wild type (stippled bars) and *Δdon* (hatched marked bars) cells were prepared in a 1:1 ratio, inoculated into 5 rats and the CFU/ml determined. The percentage of wild type or *Δdon* cells was determined by screening the total cell counts for tetracycline resistant colonies. **C)** Coomassie blue- stained SDS-PAGE gel of serous fluid samples obtained following growth of strains in the rat tissue cage. Samples from the wild type (wt) and *Δdon* (*Δ*) strains were compared to uninoculated controls (un). The arrow indicates the migration of the glycosylated transferrin. The results are from two biological repeats and panel C shows a representative result.


A
CFU per ml
$10^9$
$10^8$
$10^7$
$10^6$
$10^5$
$10^4$
2
4
6
8
10
12
14
Days

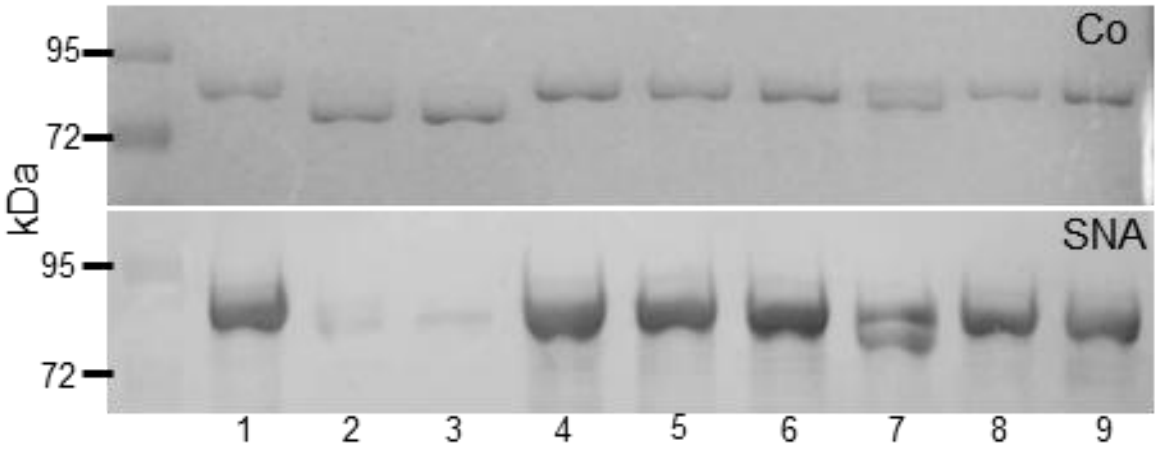

B
% Total CFU
100
80
60
40
20
0
Inoculum
Day 1
Day 4
Day 8
wt
Δdon


C
95
72
un
wt
Δ
un
wt
Δ
un
wt
Δ
Day1
Day2
Day8

### 2.5.6 Efficient transferrin deglycosylation is unique to *B. fragilis*.

*B. fragilis* is the most common *Bacteroides* species isolated from opportunistic infections at extraintestinal sites. The ability of six other intestinal *Bacteroides* species to deglycosylate human transferrin was tested by SDS-PAGE and SNA glycan staining. The results in Fig. 2.6 show that only the two *B. fragilis* strains were able to efficiently deglycosylate transferrin in the 3 hour assays. This was indicated by both the change in mass and the loss of SNA staining of the transferrin. No significant loss in SNA staining was seen for the other species except *B. vulgatus*. In this case the transferrin band appeared as a doublet with somewhat decreased staining intensity in the SNA blot (Fig. 2.6, lane 7). These data indicate that *B. vulgatus* deglycosylates transferrin inefficiently and that of the intestinal *Bacteroides* species tested only *B. fragilis* can rapidly remove the N-linked glycans from this abundant serum protein.

**Fig. 2.6.** Transferrin deglycosylation by medically important *Bacteroides* species. Human transferrin was used in standard 3 h deglycosylation assays with mid-logarithmic phase cells of *Bacteroides* species grown in DM-mucin glycan media. Samples were analyzed by SDS-PAGE with Coomassie blue (Co) staining and on duplicate SDS-PAGE gels followed by SNA glycan-staining (SNA). 1, PBS; 2, *B. fragilis* (638R); 3, *B. fragilis* (ATCC 25285); 4, *B. thetaiotaomicron* (IB116); 5, *B. uniformis*; 6, *B. ovatus*; 7, *B. vulgatus*; 8, *Parabacteroides distasonis*; 9, *Parabacteroides merdae*.


Co
SNA
kDa
95
72
95
72
1
2
3
4
5
6
7
8
9

## 2.6 Discussion

The intestinal microbiome has a complex, symbiotic relationship with its host that is maintained by a series of physical barriers and immunological processes designed to keep the microbial populations in check. When these barriers are breached, contaminating normally sterile body sites, the invading organisms must rapidly adapt to an environment with a challenging set of new physical, chemical, and nutritional parameters. In this study, we used a rat tissue cage model to gain a better understanding of the nutritional sources that opportunists such as *B. fragilis* might encounter in extraintestinal habitats. Expression microarrays showed that PUL genes were highly induced *in vivo* and the Don locus was induced more than any other genes *in vivo*. The carbon/energy sources available in the abdominal cavity are quite varied and complex so we were not surprised to see such a robust, complex response in PUL expression. A recent study of *B. thetaiotaomicron* PUL regulation during growth on mixtures of dietary glycans provides some insight into this response (22). The work showed that there is a rapid response to the influx of new glycan sources and that multiple PULs were simultaneously expressed in order to prioritize utilization of the most advantageous substrates. This required coordinated induction and repression of multiple systems working to maximize utilization of certain glycans although the rationale for prioritization is not yet entirely clear. Similarly *B. fragilis* growing in the rat tissue cage elicited a broad PUL response and in this case it seemed to be directed toward use of N-linked glycans found on host glycoproteins. The most abundant glycoprotein in tissue cage serous fluid is transferrin which is a significant substrate for the Don PUL. Many other glycoproteins present in the fluid also were targeted by the Don PUL (Fig. S3). Other PULs that were strongly induced in the microarray experiments also had

protein signatures consistent with the utilization of N-linked host glycans: BF638R0384-92 (β-N-1-4 acetylglucosaminidase + GH88 glycosidase); BF638R1323-30 (endo-β-N-acetylglucosaminidase + concanavalin A-like); BF638R0444-448 has member proteins with fibronectin binding domains.

A critical requirement for utilization of N-linked glycans is removal of terminal sialic acid residues and in a previous study it was shown that neuraminidase activity was necessary for robust growth in a rat pouch model similar to the tissue cage used in this report (86). An extensive region of the *B. fragilis* chromosome (BF638R1715-1740) seems to be devoted to sialic acid utilization and neuraminidase activity and this region has been shown to be required for optimal utilization of mucin (87). This gene cluster includes the *nanLET* operon (88), the *nanH* operon, and 5 standalone pairs of SusC/SusD orthologues. Our microarray data showed that regulation of these genes was complex but the entire region was induced on day 1 following inoculation of the tissue cage and for the most part remained elevated throughout (Appendix 2 and Fig. S2).

The hydrolysis of host glycoprotein glycans is an important characteristic of several bacterial pathogens and this has been proposed to be a mechanism for immune evasion as well as for nutrient acquisition. *Streptococcus pyogenes* secretes the archetypical GH18 family glycoprotein hydrolase, EndoS, which can efficiently hydrolyze the IgG glycan and significantly reduce IgG mediated killing of the bacteria in blood (89, 90). *Streptococcus pneumoniae* possesses several exoglycosidases that in combination deglycosylate IgA1, lactoferrin, human secretory component and α1-acid glycoprotein, and it was shown that deglycosylation of α1-

acid glycoprotein supported growth in the absence of other carbohydrates (91, 92). This ability may be linked to the persistence of *S. pneumoniae* in the nasopharynx. Likewise transferrin supports the growth of *B. fragilis* (Fig. 2.4) and is present in serum and serous fluid at 2-4 mg/ml so this would be an excellent source of fermentable carbohydrate for extraintestinal growth. Glycan harvesting also seems to be important for *in vivo* growth of *Capnocytophaga canimorsus* a member of the *Bacteroidetes* phylum and a common inhabitant of the oral cavity of dogs (93). *C. canimorsus* was able to deglycosylate fetuin and IgG *in vitro* using a Sus-like system and the deglycosylation activity was necessary to sustain growth with cultured mammalian cells.

A notable difference of the Don PUL and the glycosidase systems listed above is the inability to deglycosylate IgG or IgA. Several attempts with different sources of human IgG and IgA failed to demonstrate activity against these substrates (Fig. S5). This suggests that the Don system did not evolve for the need to inactivate immune clearance mechanisms. Further, it is well known that transferrin and lactoferrin also have antibacterial activity, largely due to their ability to sequester iron from the bacteria (94, 95). However, the glycosylation status of transferrin and lactoferrin has little effect on their ability to either bind iron or their receptors (96, 97). The Don PUL then does not seem to be designed to circumvent this element of the innate immune system.

A key finding of the work presented here was that rapid, efficient deglycosylation of transferrin and many other glycoproteins in serous fluid is mediated by the *don* genetic locus which is unique to *B. fragilis* among the medically important *Bacteroides* species. This was

demonstrated specifically for transferrin but may apply to the other serous proteins (Fig. 2.6, Fig. S3). Although not as efficient as wild type, the *Δdon* mutant was able to deglycosylate and utilize transferrin glycan at a slow rate. This was seen both *in vivo* and *in vitro* (Figs. 2.4, 2.5) and these results indicate a second transferrin deglycosylation system in *B. fragilis*. Other *Bacteroides* may have deglycosylation capabilities more similar to the *Δdon* mutant. For example, *B. vulgatus* showed some deglycosylation of transferrin after 3 h (Fig. 2.6) and results with a second strain of *B. thetaiotaomicron* showed some activity on transferrin after 3 h but failed to further deglycosylate the substrate even after overnight incubation (Fig. S6). We propose that the Don PUL is composed of 7 genes (Fig. S1). DonA and DonB are the regulators of the system encoding an ECF sigma factor and antisigma factor respectively. DonEFG are outer membrane proteins that bind and cleave glycans from the target glycoproteins. DonCD would work in tandem to transport oligosaccharides into the periplasm where the concerted effort of neuraminidases and other glycohydrolases release monosaccharides from the glycan chains (68).

The normal habitat for *B. fragilis* is the colon where it is in a persistent, mutualistic relationship with its host. The Don system must have evolved to provide a means by which to establish a specific niche in the gut or to help the organism maintain its competitive position. Several studies have indicated that *B. fragilis* are closely associated with the mucosa and that they penetrate the mucus layer to bind receptors deep in the crypts (98, 99). The *don* genes are significantly induced by a crude mixture of gastric mucin glycans in fact they were the most highly induced of all genes in mucin glycan medium (Appendix 2). One potential role for the Don PUL may be to harvest the high quality complex N-linked mucus glycans, [which make up

about 20% of mucus glycans (100)] as cells transit through the mucus layers to the crypts. Other host associated N-linked glycan sources may become available to bacteria adherent in the crypts or the *don* operon may be shut down since *don* transcription is negligible in the absence of inducer. Koropatkin et al. (19) suggested that in the colon there are many "glycan microhabitats" and any given *Bacteroides spp*. has evolved to respond to a specific subset of these. If we define the glycan habitat of *B. fragilis* then we may gain further insight as to why the Don PUL potentiates the organism to so efficiently deglycosylate proteins in extraintestinal sites.

## 2.7 Supplementary Figures

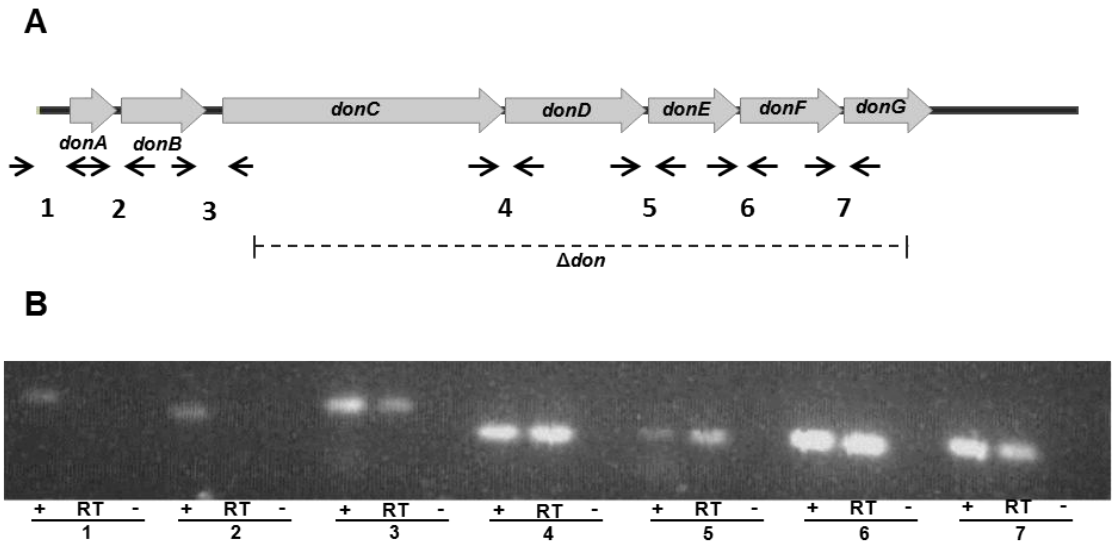



**Fig. S1.** Genetic map and RT-PCR analysis of the *donABCDEFG* genes (BF638R3437-43). **A)** Schematic diagram of the predicted *donABCDEFG* locus. The location of the Δ*don* deletion is show with the dashed line. *donAB* are regulatory genes coding an ECF sigma factor and antisigma factor respectively. *donC* and *donD* code for the SusC and SusD orthologues respectively. The *donEFG* genes code for an endo-S-like endo-N-acetyl-β-D-glucosaminidase, a Concanavalin A lectin-like product with an α-N-acetylglucosaminidase domain, and a protein with a DUF 1735 domain common to acylhydrolases respectively. **B)** Agarose gel visualization of RT-PCR products performed with oligonucleotide primers specific for each intergenic region of the *donABCDEFG* locus. Arrows below the schematic in panel A indicate the relative positions of primers used to amplify each intergenic region. "+", positive control using genomic DNA; "RT", complete RT-PCR reaction containing reverse transcriptase; "-", negative control with reverse transcriptase omitted.

The RT-PCR results show a strong signal and linkage between the *donCDEFG* genes. There also is a weaker linkage between the antisigma factor gene, *donB,* and *donC* which are separated by 182 bp. Although these two genes are linked, based on the microarray data and regulatory models for other PULs it is likely a second promoter is located upstream *donCDEFG* that would act to amplify the expression of these structural genes under inducing conditions. Regulation of this locus is currently under investigation.

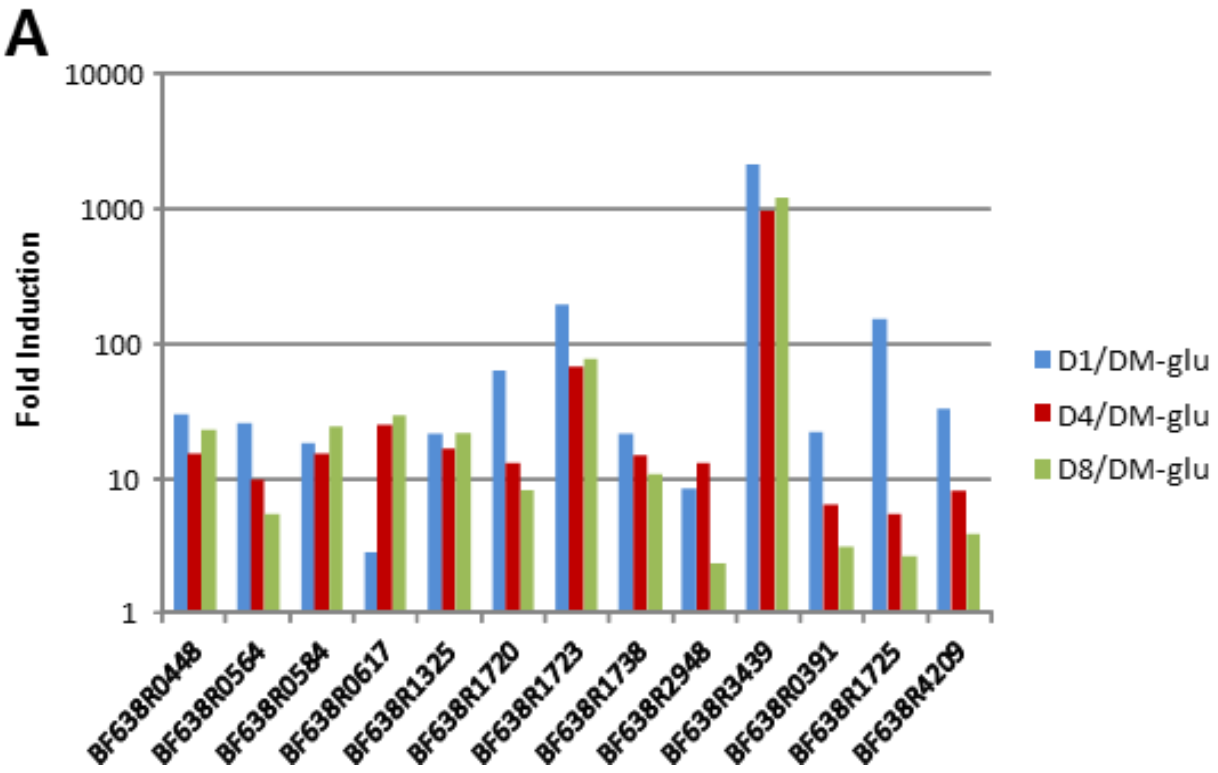



**Fig. S2.** Induction/repression patterns of SusC orthologues expressed during growth *in vivo.* Expression microarray data were used to determine induction or repression of all *susC*-like genes during growth of *B. fragilis* 638R in the rat tissue cage model relative to DM-glucose. Four induction patterns (Clusters) were identified by *k*-means clustering using the Standard Pearson's correlation coefficient distance metric. The four clusters are shown in panels A, B, C and D respectively. The fold-induction for *in vivo* growth relative to mid-logarithmic phase growth *in vitro* for 1, 4, and 8 days post-inoculation are shown by the blue, red, and green bars respectively.

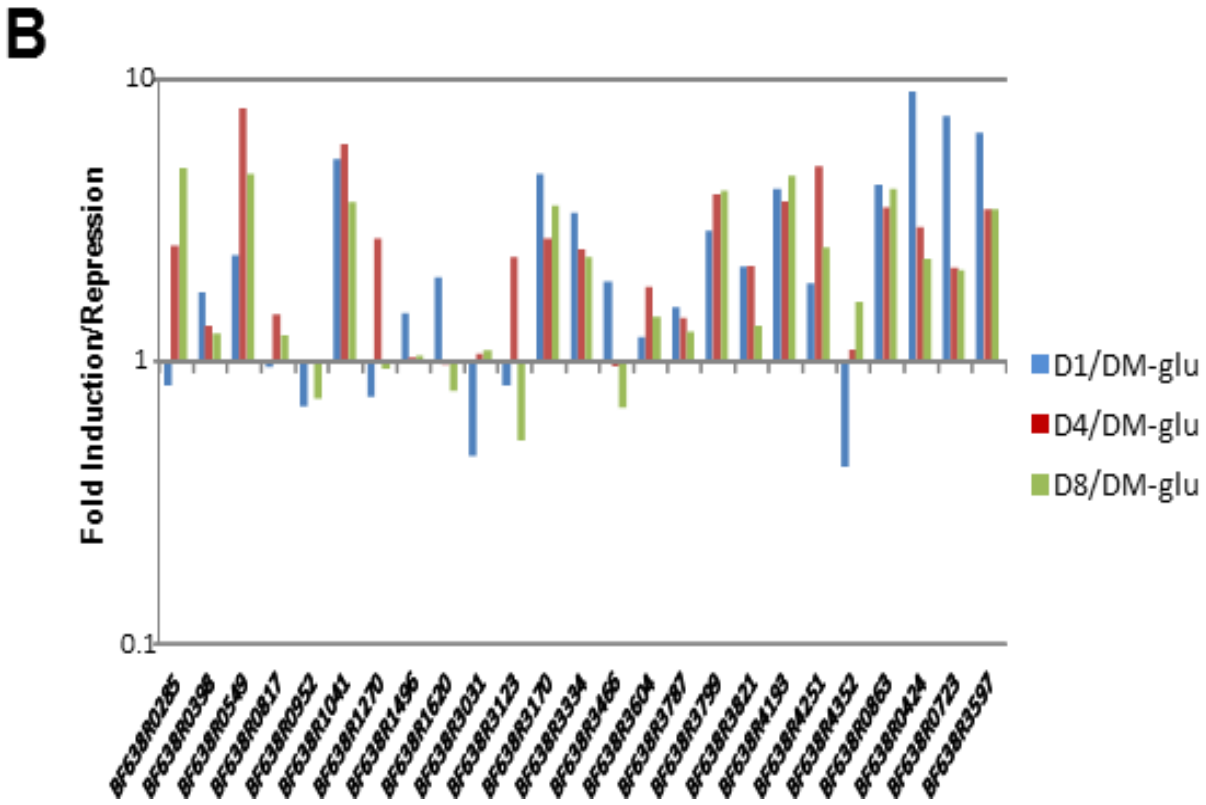



**Fig. S2.** Induction/repression patterns of SusC orthologues expressed during growth *in vivo.* Expression microarray data were used to determine induction or repression of all *susC*-like genes during growth of *B. fragilis* 638R in the rat tissue cage model relative to DM-glucose. Four induction patterns (Clusters) were identified by *k*-means clustering using the Standard Pearson's correlation coefficient distance metric. The four clusters are shown in panels A, B, C and D respectively. The fold-induction for *in vivo* growth relative to mid-logarithmic phase growth *in vitro* for 1, 4, and 8 days post-inoculation are shown by the blue, red, and green bars respectively.

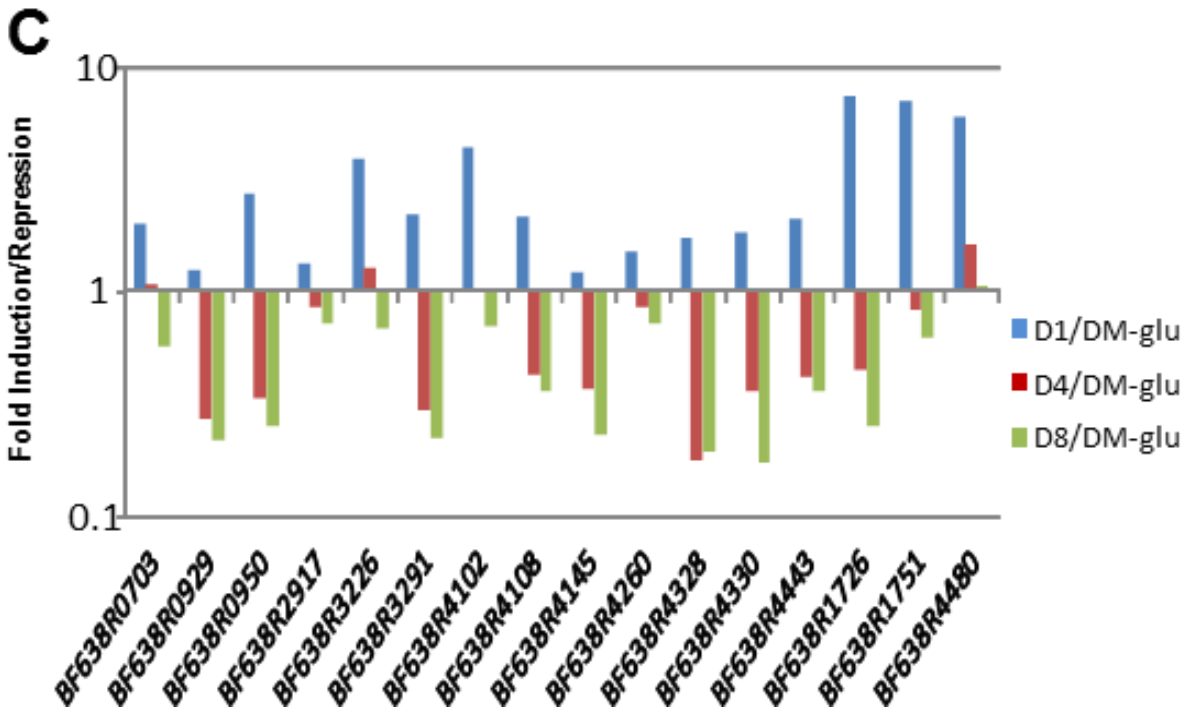



**Fig. S2.** Induction/repression patterns of SusC orthologues expressed during growth *in vivo.* Expression microarray data were used to determine induction or repression of all *susC*-like genes during growth of *B. fragilis* 638R in the rat tissue cage model relative to DM-glucose. Four induction patterns (Clusters) were identified by *k*-means clustering using the Standard Pearson's correlation coefficient distance metric. The four clusters are shown in panels A, B, C and D respectively. The fold-induction for *in vivo* growth relative to mid-logarithmic phase growth *in vitro* for 1, 4, and 8 days post-inoculation are shown by the blue, red, and green bars respectively.

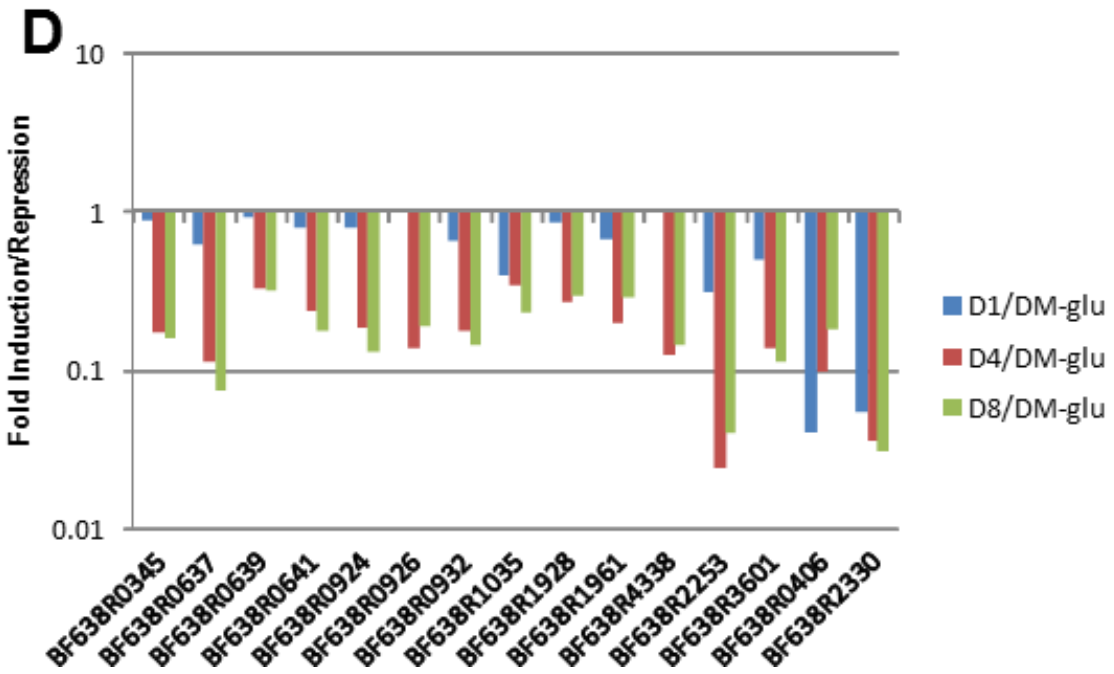



**Fig. S2.** Induction/repression patterns of SusC orthologues expressed during growth *in vivo.* Expression microarray data were used to determine induction or repression of all *susC*-like genes during growth of *B. fragilis* 638R in the rat tissue cage model relative to DM-glucose. Four induction patterns (Clusters) were identified by *k*-means clustering using the Standard Pearson's correlation coefficient distance metric. The four clusters are shown in panels A, B, C and D respectively. The fold-induction for *in vivo* growth relative to mid-logarithmic phase growth *in vitro* for 1, 4, and 8 days post-inoculation are shown by the blue, red, and green bars respectively.

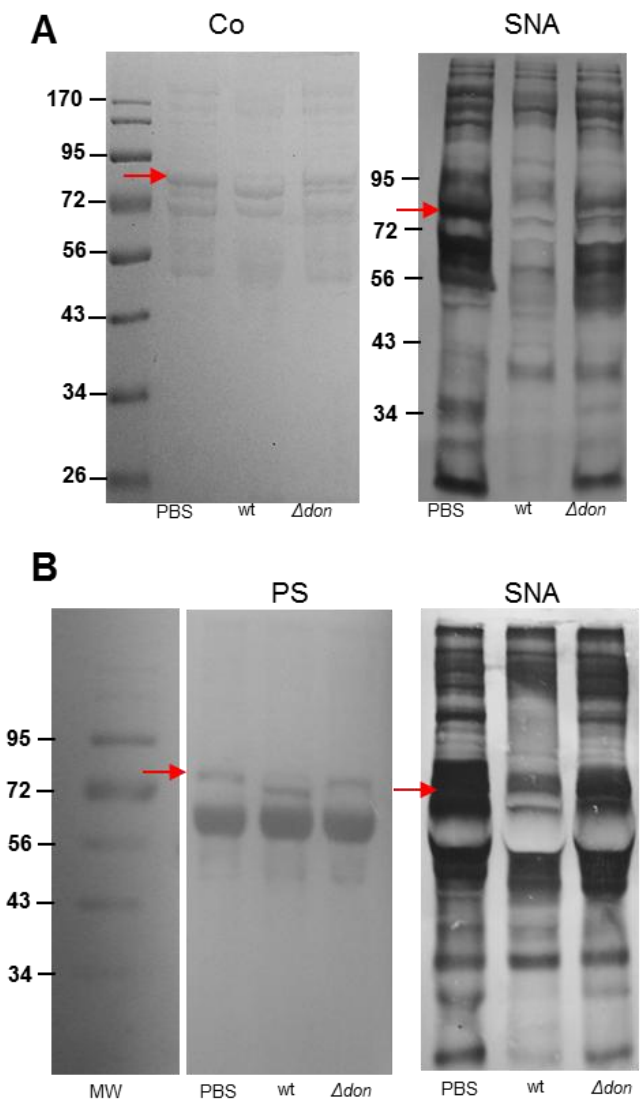



**Fig. S3.** Deglycosylation assays using total serous fluid proteins with serum albumin removed (**A**) or with serum albumin present (**B**). Standard deglycosylation analysis was performed with wild type (wt) or Δ*don* strains for 3 h. Samples were analyzed by SDS-PAGE followed by protein staining with either Coomassie blue staining (Co)or Ponceau S (PS) or glycan staining SNA lectin. The arrows indicate the transferrin bands. Serum albumin was removed with the Qproteome Murine Albumin Depletion Kit, (Qiagen, Inc.).

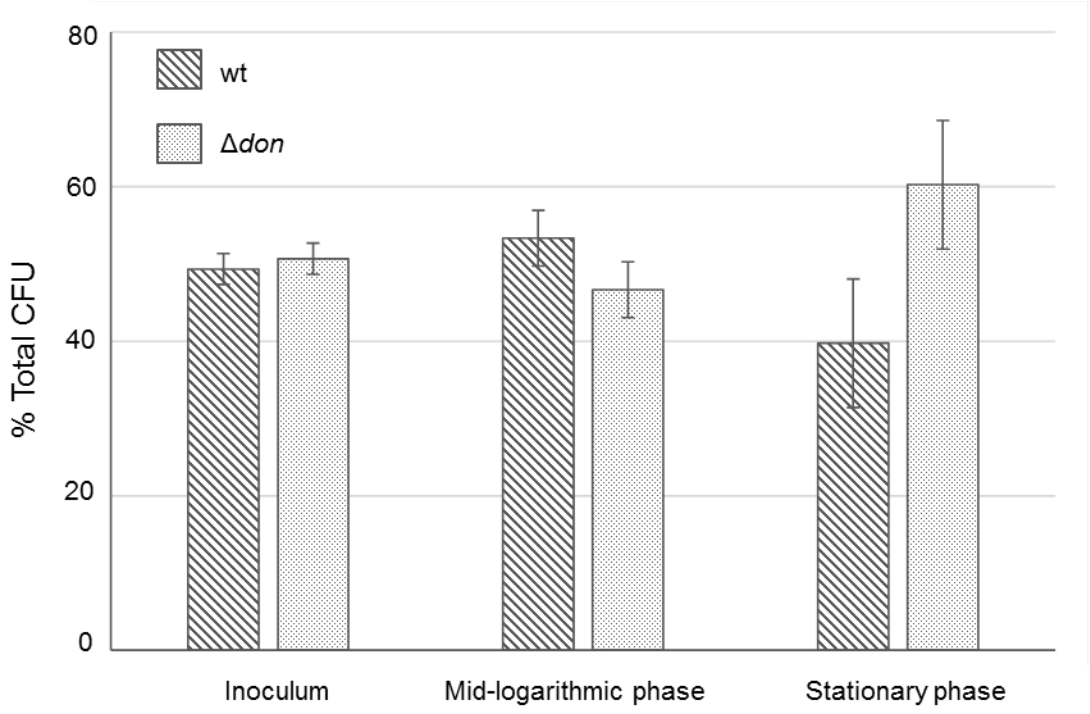



**Fig. S4.** *In vitro* competition assays in defined media with glucose. Mixtures containing a total of 4 ml of $10^5$ CFU/ml of wild type and Δ*don* cells in a 1:1 ratio were inoculated into 25 mL of defined media with glucose. Aliquots of the cell mixture were removed from the inoculum, mid-logarithmic phase cultures and the stationary phase of the cultures. The percentage of Δ*don* among the total population was determined from the percentage of tetracycline resistant CFU counts among the total CFU counts.

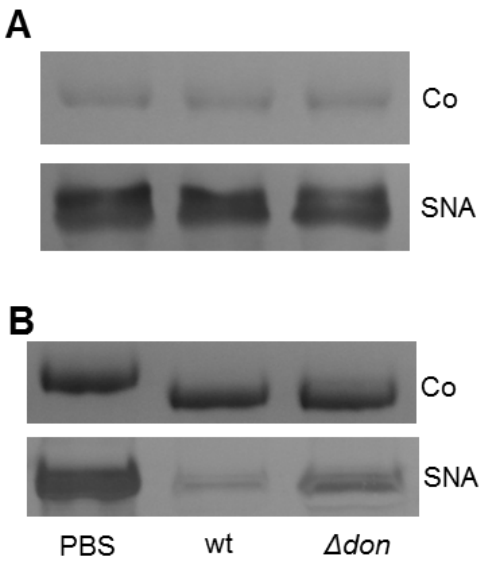



**Fig. S5.** Deglycosylation analysis of IgG and lactoferrin. **A)** Deglycosylation assay of human IgG by *B. fragilis*. Commercially prepared human IgG was mixed and incubated anaerobically at 37°C for 3 h with PBS or wild type or Δ*don* cells grown in an inducing DM-mucin glycan medium. Samples were analyzed by SDS-PAGE and subjected to Coomassie blue staining (Co) or SNA glycan-staining (SNA). **B.)** Deglycosylation analysis of lactoferrin. Commercially prepared human lactoferrin was mixed and incubated anaerobically at 37°C for 3 h with PBS or wild type or Δ*don* cells grown in an inducing DM-mucin glycan medium. . Samples were analyzed by SDS-PAGE and subjected to Coomassie blue staining (Co) or SNA glycan-staining (SNA).

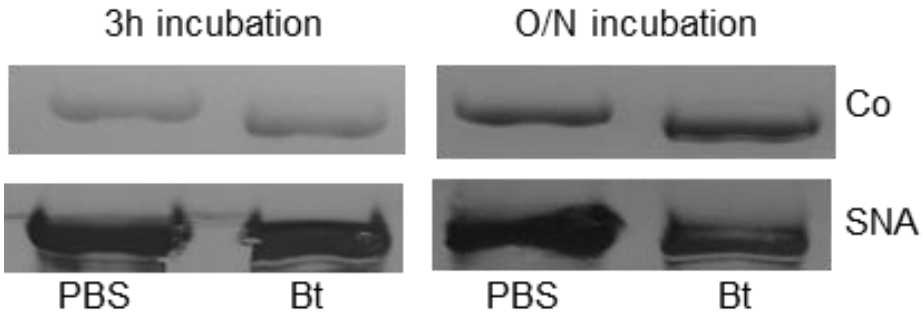



**Fig. S6.** Deglycosylation analysis of transferrin by *B. thetaiotaomicron* strain 5482. Commercially prepared human transferrin was incubated with PBS or wild type *B. thetaiotaomicron* cells (Bt) for 3 h or overnight (O/N) in standard deglycosylation assays. Cells used in the assays were grown to mid-logarithmic phase in DM-mucin glycan medium. Samples were analyzed by SDS-PAGE followed by Coomassie blue staining (Co) or SNA glycan staining (SNA).

## 2.8 Addendum

In this chapter, it was shown that the *B. fragilis don* locus mediates the efficient deglycosylation of N-linked glycoproteins (Fig. 2.3, 2.5C, S3; Table 2. 3) both *in vitro* and *in vivo*. Deletion of the *don* locus led to a growth defect in DM with transferrin as the sole carbon/energy source and a selective disadvantage in the rat extraintestinal infection model (Fig. 2.4, 2.5). While we have shown a correlation between deglycosylation activity and growth advantage in the rat abscess, model, we have not proven that deglycosylation is the cause of this advantage in vivo. This is especially of interest because we deleted not only donE, which has predicted deglycosylation activity, but also donF and donG. These genes have unknown functions that might affect *in vivo* growth. In the future direction, two experimental designs can strength the causal relation between efficient deglycosylation of N-linked glycoprotein and the selective advantage in extraintestional infection. First, point mutations of two active site residues on the endo-β-N-acetylglucoaminidase gene (*donE*) can abolish the deglycosylation activity of this enzyme (125) but keep the structure and other functions of the locus intact. If there is a real causal relation between the deglycosylation of N-linked glycoproteins and selective advantage in extraintestinal infection, the *donE* point mutant should show the same selective disadvantage phenotype as the *don* deletion mutant. Second, to rule out the possibility that the tetracycline resistant gene *tetQ* is responsible for the impaired in vivo growth/competency phenotype, a control strain will be constructed with only the *tetQ* gene inserted into the chromosome. If tetQ is not responsible for the selective disadvantage observed in the *don* deletion mutant, similar in vivo growth/competency should be observed between the wild type and the *tetQ* insertion mutant.

# CHAPTER THREE: *BACTEROIDES FRAGILIS* PRIMARY TRANSCRIPTOME ANALYSIS

## 3.1 Introduction

*B. fragilis* has a genome of about 5.3 Mb and encodes approximately 4,300 genes. To gain insight not only on the gene expression profile but also the overall structure of individual mRNAs, differential massively parallel cDNA sequencing (dRNA-seq) was performed using RNA obtained from mid-logarithmic phase anaerobic *B. fragilis* cultures grown in rich media. The dRNA-seq approach was first done with the human pathogen *Helicobacter pylori* which has a small genome (101). In contrast to other RNA-seq methods, terminator-5'-phosphate-dependent exonuclease is used to treat the RNA samples before conversion to cDNA. By adding this step, the processed transcripts such as mature rRNA, tRNA and partially degraded mRNAs which have a 5' mono-phosphate (5'P) group are selectively degraded leaving only the RNAs with a 5' tri-phosphate (5'ppp) group such as primary transcripts including most mRNAs and small RNAs (sRNA) (101). By comparing the exonuclease treated samples to non-treated samples, it was possible to generate a single-nucleotide resolution map of the primary transcriptome of *B. fragilis*, allowing us to locate the transcription initiation sites (TIS) of individual transcripts, distinguish potential sRNAs, and confirm the -7 promoter consensus sequence for the primary sigma factor. Analysis of this primary transcriptome identified 1,657 TISs, and confirmed the recognition sequence for the *Bacteroides* primary sigma factor. In

addition, 176 putative sRNAs were discovered. These results demonstrate that the dRNA-seq technique is a good approach and has great potential for understanding some of the molecular aspects of *B. fragilis* gene regulation.

## 3.2 Materials and Methods

### 3.2.1 Total RNA Extraction

*B. fragilis* strain IB101 (BF638R) was grown at 37°C anaerobically in Brain Heart Infusion broth supplemented with hemin and cysteine (BHIS). Cells were harvested in mid-logarithmic phase at O.D. $A_{550}$ = 0.5 for RNA extraction. RNA extractions from *in vitro* bacterial cell cultures were performed on washed cell pellets using the hot phenol method as described previously (82). RNA quality and integrity was checked by RNA-denaturing gel (1% agarose; 5.5% formaldehyde) electrophoresis.

### 3.2.2 Preparation of cDNA libraries and 454 pyrosequencing

Experiments were done by collaborator Jörg Vogel and lab members at University of Würzburg, Institute for Molecular Infection Biology, D-97080, Würzburg, Germany. The dRNA-seq was performed as previously described (101). Total RNA was treated by DNase I to eliminate residual genomic DNA. Equal amounts of IB101 RNA were either incubated with Terminator™ 5'-phosphate-dependent exonuclease (TEX) (Epicentre #TER51020) or in buffer alone for 60 min at 30°C. 1 unit TEX per µg total RNA was used. RNA was extracted with

phenol/chloroform/isoamyalcohol (25:24:1 v/v) and then precipitated overnight with 2.5 volumes of an ethanol/0.1M sodium acetate (pH 6.5) mixture. Then the TEX treated RNA was incubated with 1 unit TAP (tobacco acid pyrophosphatase) (Epicentre, #T19100) for 1 hour at 37°C to generate 5'-mono-phosphates for linker ligation, and again purified by phenol/chloroform extraction and precipitation as above.

Equal amounts of TEX treated or untreated RNA were poly(A)-tailed using poly(A) polymerase, followed by ligation of an RNA adapter to the 5'P RNA fragment. First-strand cDNA was synthesized using an oligo (dT)-adapter primer and M-MLV-RNaseH$^-$ reverse transcriptase. The reaction mix was incubated at 42°C for 20 min followed by 55°C for 5 min. The cDNA was amplified by PCR to yield a concentration of 20-30 ng/µl using a high fidelity DNA polymerase. Sequencing was performed on Roche 454 FLX machines at the MPI for Molecular Genetics (Berlin, Germany), and Roche Diagnostics GmbH (Penzberg, Germany).

### 3.2.3 Data Visualization

Linker and poly(A) sequences were removed from the sequencing results to reduce the background. cDNA inserts were mapped to the *B. fragilis* 638R genome using an error-tolerant suffix array technique (102). Graphs representing the number of mapped reads per nucleotide were calculated and visualized using the Integrated Genome Browser (IGB) version 4.56 software from Affymetrix (http://genoviz.sourceforge.net/).

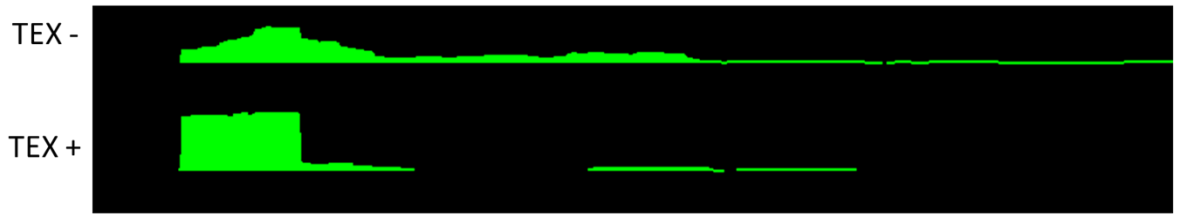
## 3.3 Results

### 3.3.1 Identification of *B. fragilis* TISs at a genome wide level

Degradation of the processed RNA (5'P transcripts) by TEX leads to an enrichment of the primary RNA (5'PPP transcripts) in the total RNA sample. Because in prokaryotic cells, RNA processing (including degradation) often causes a loss of the 5' region of a transcript (103), it is reasonable to assume that an enrichment of the primary transcripts will lead to an enrichment of the original (or intact) 5' ends of the transcripts (104). This was reflected by an increased signal at the 5' end of many transcripts on the TEX treated library compared to the untreated library. An example is shown in Fig. 3.1. Depletion of the processed RNAs by TEX leads to a characteristic change in the distribution of the total transcripts, as shown here for the BF638R_3589 mRNA. The most noticeable change is at the 5' end of the transcript where an increased number of reads are concentrated compared to the untreated library. This provides a reliable primary transcriptome map with a resolution to the single nucleotide level and thus can be used to identify the TIS of any transcript present for that specific growth condition. 1,657 primary transcripts were mapped for this growth condition (mid-logarithmic phase in rich medium), of which 838 were on the positive-strand and 819 were on the negative-strand. Analysis of the mapped primary transcripts is attached in Appendix 3.

**Fig. 3.1.** The transcript signal of the BF638R_3589 gene. *B. fragilis* strain IB101 was grown anaerobically in BHIS at 37°C to mid-logarithmic phase and RNA was prepared for dRNA-seq analysis as described in materials and methods. **A)** A comparison of the mapping signal histogram for TEX-treated and untreated samples of BF638R_3589 sRNA. **B)** A schematic diagram of the BF638R_3589 locus aligned with its signal on the transcriptome above. The TIS suggested by the dRNA-seq results is indicated by the +1 arrow and the -7 promoter consensus (TAxxTTTG) is also indicated.

**A**

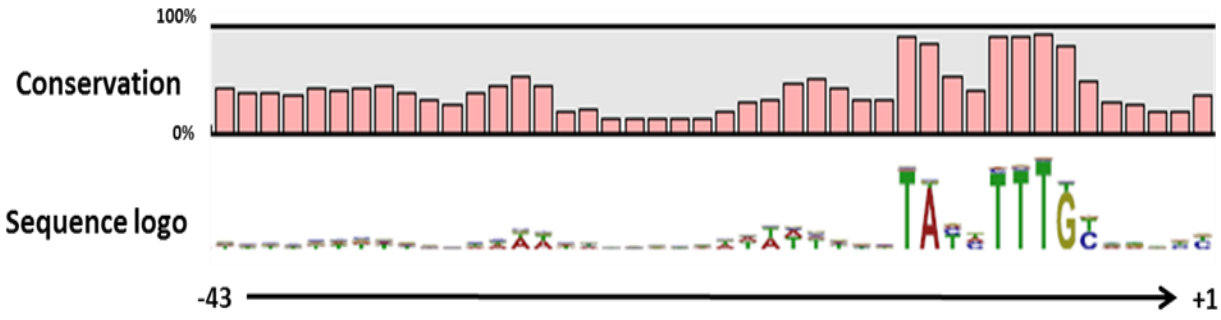



**B**

### 3.3.2 Confirmation of the recognition sequence for the *B. fragilis* primary sigma factor

The *Bacteroidetes*/*Flavobacteria* are known to have a primary sigma factor that recognizes a unique promoter structure. The *Bacteroidetes* equivalent to the classic Proteobacteria/Eubacteria -10 sequence is TAxxTTTG (105, 106). Based on the TIS identified by the dRNA-seq results, sequences about 50 bases upstream of the TIS were aligned. The conserved promoter recognition sequence "TAxxTTTG" was observed 5 to 10 bases prior the TIS in 1,428 of the 1,657 TISs observed (Fig. 3.2 and Appendix 3). These results confirmed the novel promoter consensus sequence for the *B. fragilis* primary sigma factor on a genome-wide scale.

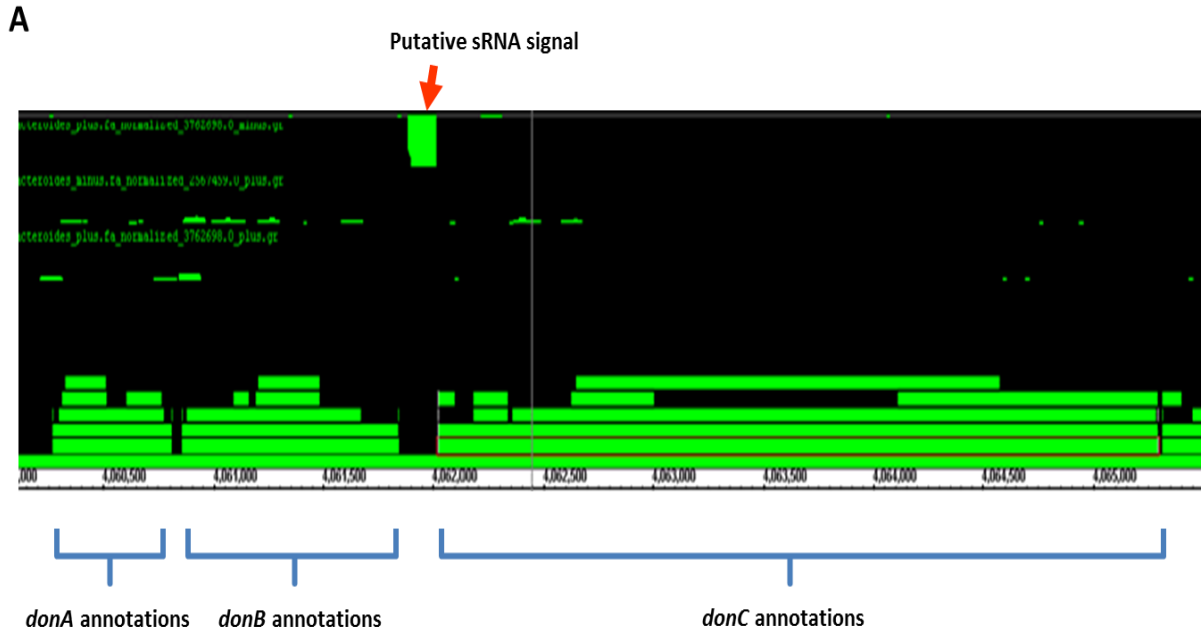
**Fig. 3.2.** Logo for the promoter recognition sequence of the primary sigma factor in *B. fragilis*. Primary transcriptome data were used to identify sequence about 50 bases upstream of the TIS sites. Sequences of 1,657 genes were aligned and analyzed using CLC Main Workbench v6.5 software.


100%
Conservation
0%
Sequence logo
-43
+1

### 3.3.3 Discovery of putative small non-coding RNAs in the *B. fragilis* genome

Small non-coding RNA transcripts will have the enriched 5' signal in the TEX (+) library and are usually located in intergenic regions. 176 putative sRNAs were observed by these criteria. Interestingly, we observed a group of 14 putative sRNAs individually located in the intergenic region of 14 Sus-like PULs, including the *don* PUL described in chapter 2. The apparent sizes of these sRNAs range from 80 bases to 250 bases and they are divergently transcribed from the putative SusC-like porin of their respective PUL. Shown in Fig. 3.3 is a putative sRNA located in the *don* PUL. Characteristics of these PUL-associated sRNAs are presented in Table 4.5 in the next chapter (chapter 4).

**Fig. 3.3.** The putative sRNA in the intergenic region of the *don* PUL. **A)** The putative sRNA located in the *don* PUL visualized on the Integrated Genome Browser. The sRNA signal is indicated by the red arrow. The green bars are local alignment annotations for *donA*, *donB*, and *donC* respectively as indicated at the bottom. **B)** A schematic diagram showing the location and orientation of the putative sRNA in the intergenic region of the *don* PUL.


A
Putative sRNA signal
4,060,500
4,061,000
4,061,500
4,062,000
4,062,500
4,063,000
4,063,500
4,064,000
4,064,500
4,065,000
donA annotations
donB annotations
donC annotations


B
Putative sRNA
donA
donB
donC (susC-homolog)

## 3.4 Discussion

One major advantage of using the dRNA-seq technique is that it can generate a primary transcriptome that precisely shows TISs across the entire genome for the condition tested. It is considerably more efficient than other TIS identification methods such as primer extension or 5' RACE. Further, based on our observation it is at least equal in accuracy compared to primer extension. As shown in the results, a majority of the TISs mapped from the primary transcriptome data have the conserved *Bacteroides* primary sigma factor recognition sequence (TAxxTTTG) 5 to 10 bases upstream. This result supports the reliability of the TIS data since it is reasonable to expect that for the growth condition tested (37°C, anaerobic, rich medium, mid-logarithmic phase), many of the genes expressed would be transcribed by the primary sigma factor.

Another important feature of the primary transcriptome map is that putative sRNAs are readily identified due to the enrichment of their 5' signal and the reduced background noise from processed RNAs. Most interesting for this work, a group of 14 putative sRNAs were observed in the intergenic region of 14 PULs, including the *don* PUL described in chapter 2. Based on the locations and orientations of these putative sRNAs, they are likely to be *cis*-encoded antisense sRNAs that play a role in the regulation of the PUL they are in proximity to. The role of these sRNAs and their regulation mechanism will be the topic of the next chapter (chapter 4).

In this study, only the primary transcriptome in one growth condition was explored: 37°C, anaerobic, rich medium, mid-logarithmic phase cells. It will be much more informative

when the transcriptomes of different growth conditions are compared. In these cases, it is reasonable to expect that there will be changes in the mRNAs such as alternative TIS and alternative transcript termination sites. By comparing the primary transcriptomes of stressed and non-stressed cells, we may also detect some stress responsive sRNAs. All in all, primary transcriptome analysis is a useful tool and has great potential in understanding the molecular mechanisms of regulating *B. fragilis* physiology and pathogenesis.

# CHAPTER FOUR: CHARACTERIZATION OF THE REGULATION MECHANISMS FOR THE *BACTEROIDES FRAGILIS* POLYSACCHARIDE UTILIZATION LOCUS, *DON*

## 4.1 Introduction

As one of the most abundant symbiotic bacteria in the human gut, *Bacteroides* species are well known for their extraordinary ability to utilize a wide range of diverse polysaccharides as their carbon/energy source. These include plant fiber glycans from dietary sources as well as host derived glycans in the form of glycoconjungates such as glycoproteins. This extensive catabolic ability is mediated by a large number of outer membrane complexes coded by discrete polysaccharide utilization loci (PULs). For example, one of the most abundant gut *Bacteroides* species, *B. thetaiotaomicrom*, dedicates about 18% of its genome to PULs (62). The PULs in *Bacteroidetes* have a unique yet well-characterized operon structure that was first documented for the starch utilization system (Sus) of *B. thetaiotaomicron* (68). This consists of genes for an outer membrane protein complex that involves recognition, binding, processing and transport of its substrate glycan molecules. This outer membrane complex contains a SusC-like TonB-dependent transmembrane transporter and a SusD-like substrate binding protein while additional genes encoding specific substrate binding and processing proteins are usually present downstream of the SusD-like gene (62-66). Please see chapter 1 for a detailed review of the structure and functional mechanism of the *Bacteroides* Sus-like PULs. These PULs are

substrate specific, for example the prototype of these systems, the *B. thetaiotaomicron* Sus, only binds, processes and transports starch into the cell (22, 68). Expression of the PULs must be under precise control to avoid futile energy spending and envelope stress. Most of the PULs have regulatory genes located upstream from the SusC-like gene in the form of a two component regulatory phosphorelay or an ECF (extracytoplasmic function) sigma factor/ anti-sigma factor pair which is the most common regulatory mechanism. The SusC-like porin, ECF sigma factor and anti-sigma factor form a trans-envelope signaling pathway that provides a positive feedback loop; when the inducing glycan substrate is present in the environment, the SusC-like transporter will sense the signal and cause a conformational change of the anti-sigma factor followed by the release of the ECF-sigma factor and specific induction of the PUL. When there is no inducer substrate glycan for that particular PUL, the SusC-like porin will interact with the anti-sigma factor so that it will sequester the sigma factor and the expression of the system will switch back to an "off" state (62, 68).

*B. fragilis* is typical of the *Bacteroides* with more than 50 PULs. Interestingly, an analysis of the primary transcriptome for possible sRNAs, intense signals were observed in the intergenic regions between the SusC-like gene and the anti-sigma factor gene of 14 PULs. Please review chapter 3 for a description of the transcriptome analysis. Each of these sRNAs is antisense to the SusC-like gene in the operon in which they were found. There are two classes of antisense sRNAs in the prokaryotic systems: one is the *cis*-encoded antisense sRNAs that are located in the same chromosomal region and fully complementary to their target mRNAs over a long sequence stretch; and another class is the *trans*-encoded antisense sRNAs which are located in another more distant chromosomal region, only partially complementary to their

targets and usually need the small RNA chaperone Hfq to aid the sRNA-target interaction (107-110). Based on the location and orientation of the 14 PUL-associated sRNAs, we hypothesized that they are *cis*-encoded regulatory sRNAs that control the expression the specific PUL in which they were found. In this study, we used the *B. fragilis don* locus (please see chapter 2 for a detailed study on the function and characterization of *don* locus) as a model to understand PUL regulation in *B. fragilis*, by first elucidating the roles of the ECF-sigma factor and anti-sigma factor, and then by testing our hypothesis that the sRNAs play important regulatory roles in the *B. fragilis* PULs.

## 4.2 Materials and Methods

### 4.2.1 Bacterial strains and growth

Bacterial strains used in this study are listed in Table 4.1. *Bacteroides* strains were grown at 37°C anaerobically in Brain Heart Infusion broth supplemented with hemin and cysteine (BHIS) for routine cultures (55). *E. coli* strains were grown at 37°C aerobically in Luria-Bertani broth for routine cultures. Ampicillin (100 µg/mL), spectinomycin (50 µg/mL), rifampicin (20 µg/mL), gentamycin (100 µg/mL), erythromycin (10 µg/mL), tetracycline (5 µg/mL), trimethoprim (100 µg/mL), thymine (50 µg/mL) were added to the media when required (74). Minimal defined media (DM) were prepared as described previously with the specific carbon/energy source as mentioned in the text. Mucin glycans were prepared from porcine stomach mucin (Sigma-Aldrich, Cat. M2378) according to Eric C. Martens (20). All subcultures were done using a 2% inoculum of an overnight BHIS culture

### 4.2.2 Construction of *donA* and *donB* deletion mutants and the *donA* over-expression strain

The *donA* and *donB* mutants, IB559 and IB560 respectively, were constructed as unmarked and in-frame deletions to avoid potential polar effects. Briefly, chromosomal fragments of about 1.2 kb flanking *donA* were PCR amplified using primer pairs sigOK+2kL/sigOK+2kR and sigOK-2kL/sigOK-2kR respectively (Table 4.3). The amplified DNA fragments were cloned into the *B. fragilis* suicide vector pYT102 (14) by three-fragment ligation. The plasmid was mobilized into *B.fragilis* ADB77 ($thyA^-$) using the *E. coli* helper strain HB101::RK231 (79), and the exconjugates were selected on BHIS plates containing rifampicin, gentamicin and tetracycline. Exconjugates were then allowed to resolve in BHIS medium supplemented with thymine for three days and then plated on trimethoprim-containing media. Sensitivity to tetracycline, resistance to trimethoprim, and PCR were used to confirm the double-crossover allelic exchange into the ADB77 chromosome. The mutant construction was finalized by reversion of the strains from $thyA^-$ to $thyA^+$ (14). The *donB* mutant was made in the same way with the primer pairs anti-sigOK+2kL/anti-sigOK+2kR and anti-sigOK-2kL/anti-sigOK-2kR (Table 4.2). The *donA* over-expression strain IB558, was made by PCR amplifying the *donA* gene using primer pairs sigOK-340L/sigOK-340R. Primer sigOK-340L contains a strong ribosomal binding sequence from the *B. fragilis ahpC* gene to facilitate expression. The amplified fragment was inserted to the *Bacteroides-E.coli* expression shuttle vector pFD340 (Table 4.2) downstream of the constitutive IS4351 promoter sequence. The recombinant plasmid was then mobilized into the wild type *B. fragilis*, BF638R, using the *E.coli* helper strain RK231.

### 4.2.3 Construction of DonS silencing mutant (IB561) and DonS over-expression strain (IB563)

DonS silencing was achieved by changing its core promoter sequence "TTTG" to "AAAC" through site-directed mutagenesis. First a 1301bp chromosomal fragment covering the entire *donS* region was amplified and the TTTG in the core promoter sequence was replaced with AAAC by using overlapping PCR with the mutagenic primer pairs sRNA117+L/sRNA117+R(AAACmut) and sRNA117-L(AAACmut)/sRNA117-R [Table 4.2; sRNA117+R(AAACmut) and sRNA117-L(AAACmut) are complementary to each other]. The mutated fragment was cloned into the suicide vector pYT102. The allelic exchange mutation was selected as described above for *donA* except that colony-PCR followed by nucleotide sequencing of the PCR fragments was used to screen for the mutation after resolving the exconjugates in BHIS plus thymine media. The DonS over-expression strain, IB563, was constructed by cloning the *donS* gene downstream of an 82 bp 16s rRNA promoter sequence into a multiple copy vector. Briefly, the primer pair sRNA117-L/sRNA117-R1 was used to amplify the *donS* sequence with 36 bp of 16s rRNA promoter sequence engineered into the sRNA117-L primer, upstream of the *donS* sequence. Then this PCR product was used as a template in a PCR amplification performed with primer pair 16sPR-36-82/sRNA117-R1. The primer 16sPR-36-82 includes an additional 46 bp adjacent to and upstream of the 36 bp fragment of the 16s rRNA promoter sequence. The final product contained the *donS* sequence downstream of the 82 bp of the 16s rRNA promoter sequence (111, 112). The final PCR product was cloned by replacing the DNA fragment between the SacI site and PstI site on the plasmid pFD340. The recombinant plasmid was mobilized into *B. fragilis* strains using the *E.coli* helper strain HB101::RK231.

Exconjugates were selected on BHIS media containing rifampicin, gentamycin and erythromycin.

### 4.2.4 Total RNA extraction, cDNA synthesis and qRT-PCR

Total RNA was extracted from cell pellets using the hot phenol method described previously (82) and stored in 50% formamide at -80°C. Total RNA was purified using the RNeasy Mini Kit (Qiagen, Inc.) and DNA was removed by treatment with DNase (Ambion/Life Technologies Inc.). DNA contamination was determined by PCR using specific primers for the 16s ribosomal RNA gene (Table 4.3). First strand cDNA synthesis was carried out using 1 µg total RNA with random hexamer primers and Superscript III RT (Life Technologies, Inc.). For qRT-PCR, primer pair, omp117rtL and omp117rtR, were used to amplify a 140 bp fragment of the *donC* gene in a standard reaction mixture with SYBR$^{R}$ Green Supermix (Bio-Rad, Inc.). All sample reactions were in run in triplicate and RNA with no reverse transcriptase was used as a control to monitor for DNA contamination. Relative expression values were normalized to 16s rRNA and calculated by the method of Pfaffl (83). Results represent at least two independent experiments performed in triplicate.

Expression microarray analyses were performed essentially as described previously (58). Double stranded cDNA was synthesized with the SuperScript$^{R}$ Double-Stranded cDNA Synthesis Kit (Life Technologies, Inc.). One µg of purified double stranded cDNA was then labeled with cy3, hybridized to microarray slides, and processed by the Florida State University Roche/NimbleGen Microarray Facility. Each trial consisted of a high-density-oligonucleotide

whole genome expression microarray (Roche/NimbleGen, Madison WI) with 8 technical replicates of each probe per slide. Raw microarray expression data were normalized by using the RMA algorithm implemented with the Roche DEVA 1.1 software.

### 4.2.5 Northern blot analysis

Northern blot analyses were performed as described previously (113). Briefly, 10 µg of RNA was separated on a 10% denaturing polyamcrylamide gel containing 7 M urea and 1X TBE (89 mM Tris-base, 89 mM boric acid and 2 mM EDTA). The low-molecular-weight DNA ladder (NEB, cat. N3233S) was labeled with [γ-$^{32}$P]-ATP and polynucleotide kinase, and used as the size standards. Following electrophoresis in 1X TBE buffer, the ladder and RNA samples were transferred to an Amersham Hybond$^{TM}$-N$^{+}$ membrane (GE Healthcare, cat. RPN303B) in 1X TBE buffer. The samples were UV-cross-linked to the membrane, and the membrane was pre-hybridized in hybridization buffer (ULTRAhyb®-Oligo Buffer, Ambion, cat. AM8863) for 3 h at 43°C on a rocking shaker. The oligonucleotide probes specific for each small RNA species were end-labelled with [γ-$^{32}$P]-ATP and polynucleotide kinase. The radiolabelled probes were incubated with the pre-hybridized membranes at 43°C on a rocking shaker overnight (~ 12 h). The membranes were then washed three times for 30 min each with 2X SSC (300 mM sodium chloride and 30 mM sodium citrate), 1X SSC and 0.5X SSC, respectively, at 43°C on a rocking shaker. All SSC wash buffers contained 0.1% sodium dodecyl sulfate (SDS). The membranes were then exposed to X-ray film and visualized by autoradiography.

Table 4.1. Bacterial strains used in this study.

| Bacterial strains | Description or genotype[a] | Reference or source |
|---|---|---|
| IB101 | *B. fragilis*, Clinical isolate 638R, Rif$^r$ | (75) |
| ADB77 | *B. fragilis*, strain 638R, ΔthyA,, Rif$^r$, Tp$^r$ | (14) |
| DH10B | *E. coli*, F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *end*A1 *ara*D139 Δ(*ara leu*) 7697 *gal*U *gal*K rpsL *nup*G λ$^-$ | Invitrogen |
| HB101::RK231 | *E. coli*, HB101 containing RK231, Kan$^r$, Tet$^r$, Sp$^r$ | (79) |
| IB558 | *B. fragilis*, strain 638R, *donA* overexpressed on pFD340, Rif$^r$, Erm$^r$ | This Study |
| IB559 (*ΔdonA*) | *B. fragilis*, strain 638R, *ΔdonA*, Rif$^r$ | This Study |
| IB560 (*ΔdonB*) | *B. fragilis*, strain 638R, *ΔdonB*, Rif$^r$ | This Study |
| IB561 | *B. fragilis*, strain 638R, DonS silenced by promoter sequence mutation, Rif$^r$ | This Study |
| IB563 | *B. fragilis*, strain 638R, DonS overexpressed on pFD1244, Rif$^r$, Erm$^r$ | This Study |

[a]Rif$^r$, rifampin resistance; Tet$^r$, tetracycline resistance; Sp$^r$, spectinomycin resistance; Erm$^r$, erythromycin resistance; Tp$^r$, trimethoprim resistance; Kan$^r$, kanamycin resistance.

Table 4.2. Plasmids used in this study

| Plasmid | Description[a] | Reference of source |
|---|---|---|
| pYT102 | P15A ori, RP4 oriT; *B. fragilis* suicide vector, thyA$^+$, (Cm$^r$), Tet$^r$ | (14) |
| pFD340 | *Bacteroides-E.coli* expression shuttle vector, (Amp$^r$), Erm$^r$ | (114) |
| pFD1243 (pFD340_*don*A) | *B. fragilis* expression vector, generated by inserting the cloned *donA* gene with ahpC ribosomal binding site downstream the IS4351 promoter on pFD340 | This Study |
| pFD1244 | *B. fragilis* expression vector, generated by replacing the IS4351 promoter on pFD340 with an 82 bp 16s rRNA promoter sequence and the *donS* gene | This Study |

[a]Cm$^r$, chloramphenicol resistance; Tet$^r$, tetracycline resistance; Amp$^r$, ampicillin resistance; Erm$^r$, erythromycin resistance. For *Bacteroides-E. coli* shuttle vectors, parentheses indicate antibiotic resistance expression in *E. coli*.

Table 4.3. Oligonucleotides used in this study[a].

| Name | Sequence ( 5’ → 3’) | Tag | Description |
|---|---|---|---|
| sigOK+2kL | AGTC GGATCC CCGATGGTTACATCTACGATC | BamHI | Designed to amplify a 1283 bp upstream fragment of *donA* gene |
| sigOK+2kR | AGTC CTGCAG CGCTTTGACTTTGGGATAAGTC | PstI | |
| sigOK-2kL | AGTC CTGCAG CACATCTATTTGGCACTAATCG | PstI | Designed to amplify a 1231 bp downstream fragment of *donA* gene |
| sigOK-2kR | GCAT AAGCTT TAACGCAAAGAATTC | HindIII | |
| Omp117rtL | GGTGAAGGCATTTCCGACTT | | Designed to amplify a 140 bp fragment of *donC* gene for quantitative PCR |
| Omp117rtR | TTGCCTTCCTGCCCTTTCTT | | |
| 16srL | GATGCGTTCCATTAGGTTGTTG | | Designed to amplify a 127 bp fragment of 16s ribosomal RNA gene for quantitative PCR |
| 16srR | CACTGCTGCCTCCCGTAG | | |
| anti-sigOK+2kL | CACG AAGCTT GCGTACAGTA | HindIII | Designed to amplify a 1393 bp upstream fragment of *donB* gene |
| anti-sigOK+2kR | AGTC CTGCAG GGCAAACAGACGGATGATTC | PstI | |
| anti-sigOK-2kL | AGTC CTGCAG GTTGTGGGAGGATTCAGTCAT | PstI | Designed to amplify a 1296 bp downstream fragment of *donB* gene |
| anti-sigOK-2kR | AGTC GGATCC CTCCGTATTGTTGGAGATCGA | BamHI | |
| sigOK-340L | AGTC GGATCC AAATAAGAAACAATT ATGATTTTAAATAACGAGTCTAATAAGAAG | BamHI | Designed to amplify the *donA* gene with the *ahpC* gene ribosomal binding site upstream |
| sigOK-340R | AGTC GAGCTC AGGTGTGACTGATTACTCAA | SacI | |
| sRNA117+L | AGTC GGATCC GTCAGCCATTCATTGTCAGA | BamHI | Designed to amplify a 693 bp DNA fragment including *donS* sequence with its -7 “TTTG” change to “AAAC” |
| sRNA117+R (AAACmut) | AAATAGACTTTAATCGATAAAAATTCATAGAAAACAAC GAATTTTAGTGTTAAATCATAATATTTATTTCC | | |
| sRNA117-L (AAACmut) | GGAAATAAATATTATGATTTAACACTAAAATTCGTTGT TT TCTATGAATTTTTATCGATTAAAGTCTATTT | | Designed to amplify a 679 bp DNA fragment including *donS* sequence with its -7 “TTTG” change to “AAAC |
| sRNA117-R | AGTC GGATCC CTGCTGAACACTGTAACTCA | BamHI | |
| sRNA117-L | TTGTCTCTTATCTCCTAATGCCTTACTTTTGCATCCCGA ATTTTAGTGTTAAATCATAATATTTATTTC | | Designed to amplify the *donS* sequence with an 82 bp 16s rRNA promoter fragment sequence tagged upstream. |
| sRNA117-R1 | AGTC GAGCTC GGATTCAGTCATGAACTGAAG | SacI | |
| 16sPR-36-82 | AGTCCTGCAGTTTACGTTTTTATTCAAAATATTTTCAAA AAAATCCCCTTTTATATTTGTCTCTTATCTCCTAATGCC | PstI | |

[a]Tag sequences are underlined.

Table 4.3. Oligonucleotides used in this study (continued).

| Name | Sequence ( 5’ → 3’) | Tag | Description |
|---|---|---|---|
| DonSP | GCCACCTGATTCCGGATATGAAATCTGAAT | | Probe for small RNA DonS |
| 0549P | GGTGGCACACTTCCGGCTCCGGGTTAACAA | | Probe for small RNA 0549 |
| 1041P | CCTTGTGAGCTTGTTTGGCGACAAGCTGAGA | | Probe for small RNA 1041 |
| 1270P | GGTCTGGTACACCTTTCCCTCCCGATAAAGTCAA | | Probe for small RNA 1270 |
| 1469P | TGAGAATGCCGATAGGTGTTCCCGCACCTA | | Probe for small RNA 1469 |
| 1035P | ATTGGCGTATCTACCTGTATGAAAGTCCACCGTC | | Probe for small RNA 1035 |
| 2917P | ACATATTGGGGTATGTCCTCCGATTCAGAA | | Probe for small RNA 2917 |
| 3123P | CACCCATTTCCGATCCCCGAAATCAATCAA | | Probe for small RNA 3123 |
| 3597P | CTGAGTTTTTTTAATCACGTCTCTTGCAGGAGGCGC | | Probe for small RNA 3579 |
| 3604P | CCTCATATCATCCAACCCCTGATTAGATTAAACTACG | | Probe for small RNA 3604 |
| 4145P | CAGATGCGCCAACACCTGCCAACCTCGGGT | | Probe for small RNA 4145 |
| 5srRNAP | CACTGTTACGCAGTACCATCGGCGTGATCA | | Probe for the 5s ribosomal RNA |

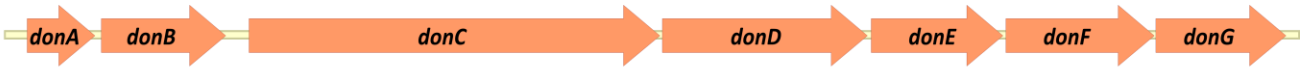
## 4.3 Results

### 4.3.1 BF638R_3437 and BF638R_3438 encode regulators of the *don* operon

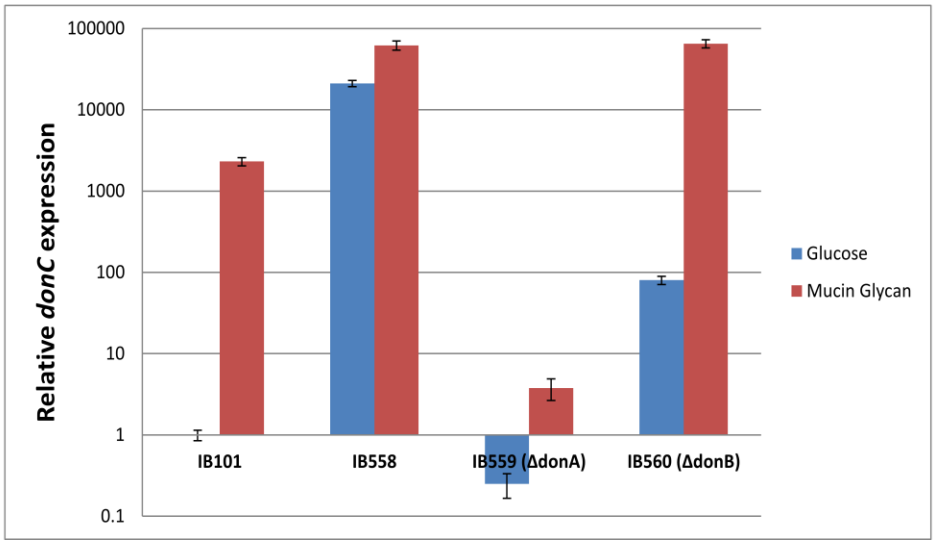
The typical genetic organization of the *Bacteroides* PULs includes regulatory genes adjacent to the *susC* orthologues. In the case of the *don* PUL, there is an ECF sigma factor homolog BF638R_3437 (*donA*) and an anti-sigma factor homolog BF638R_3438 (*donB*) which appear immediately upstream from *donC* (Fig. 4.1A), suggesting that they are the regulators of this PUL. To confirm their roles in regulation, in-frame deletion mutants for both *donA* and *donB* were constructed, and a *donA* overexpression strain was made by cloning the gene on an expression plasmid, under the control of a constitutive promoter. The relative expression level of *donC* was determined for the wild type strain and genetically modified strains in glucose defined medium (non-inducing medium) and mucin glycan defined medium (inducing medium). The expression of *donC* in the wild type strain under non-inducing conditions was as low as background levels and used as the baseline against which the expression in other strains and conditions was compared. As shown in Fig. 4.1B, *donC* expression was highly induced in the wild type when grown in mucin glycan medium. When *donA* was overexpressed on the plasmid, *donC* levels were high even in the non-inducing medium compared to the wild type, and it was induced to the highest levels observed when grown in the inducing medium. When the sigma factor homolog was deleted from the cell, induction of *donC* expression was abolished even when grown in the inducing condition. The anti-sigma factor deletion mutant, *ΔdonB*, was able to highly express *donC* in the inducing medium similar to the *donA* overexpression strain, but

**Fig. 4.1. Relative expression of *donC* in different *B. fragilis* strains under non-inducing and inducing conditions. A)** A schematic diagram of the *B. fragilis don* operon. **B)** qRT-PCR analysis of *don*C expression. RNA samples were obtained from mid-logarithmic phase cultures of the wild type strain (IB101), the *donA* overexpressing strain (IB558), *donA* deletion mutant (IB559) and *donB* deletion mutant (IB560). qRT-PCR were performed in triplicate. The *donC* expression level in the wild type strain under non-inducing condition (glucose defined medium) was used as a control. Results are the average of two biological repeats. Error bars represent standard deviations.

A


donA
donB
donC
donD
donE
donF
donG

B


Relative donC expression
100000
10000
1000
100
10
1
0.1
IB101
IB558
IB559 (ΔdonA)
IB560 (ΔdonB)
Glucose
Mucin Glycan

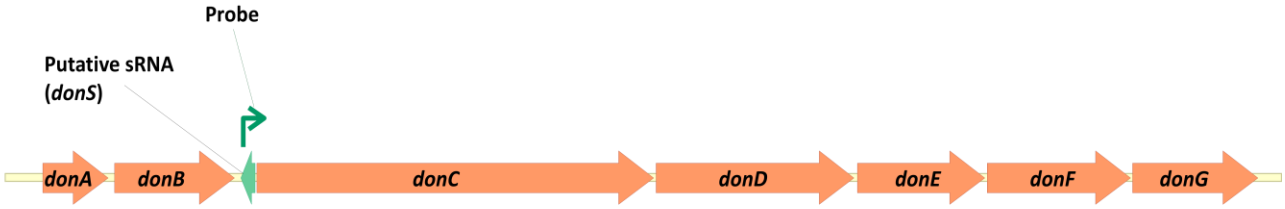
there was about an 80-fold up-regulation in the non-inducing medium compared to the wild type strain. From these data it is clear that the sigma factor homolog *donA* and anti-sigma factor homolog *donB* act as an activator and a repressor respectively and work coordinately in the regulation of the *don* operon expression in response to the presence of the substrate nutrients.

### 4.3.2 A small non-coding RNA is divergently transcribed from the *donC* gene.

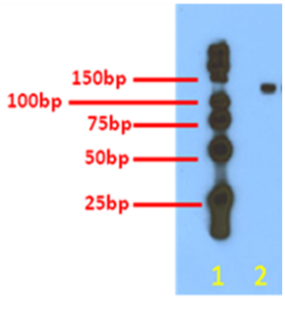
As described in chapter 3, when analyzing the primary transcriptome data of cells grown under standard growth conditions in BHIS medium, a group of putative sRNA signals in 14 PULs were identified, including one in the *don* PUL. Based on the primary transcriptome map, all these putative sRNAs are located immediately upstream from, and divergently transcribed from the SusC orthologues of each of the 14 PULs (see chapter 3 Fig. 3.3). To verify the existence of the sRNA in the *don* PUL, northern hybridization analysis was performed using an oligonucleotide complementary to the putative sRNA sequence shown in the primary transcriptome map as the probe (Fig. 4.2A). As shown in Fig. 4. 2B, a sRNA species about 125 nucleotides long was clearly detected by northern blot analysis.

**Fig. 4.2. A small, non-coding RNA is divergently transcribe from the *donC* gene. A)** A schematic diagram of the sRNA location and orientation in the *don* PUL. The probe used to detect this sRNA is indicated by the green arrow. **B)** Northern blot confirmation of the sRNA species. 10 µg of total RNA from a *B. fragilis* mid-logarithmic phase culture grown in BHIS medium was electrophoresed for analysis. Lane 1, γ-$^{32}$P labeled low molecular weight DNA ladder. Lane 2, sRNA detected by the γ-$^{32}$P labeled probe. **C)** Sequence analysis of the sRNA region in the *don* PUL. The TIS and -7 promoter sequence of the sRNA are underlined and indicated by "+1" and "-7". The DonC coding sequence is in blue and the sRNA sequence is in Red. The sRNA sequence was determined from the transcriptome analysis (see chapter 3).

A




B




C

AATATATTGATTTGTTTAACCTTAAAAACAACTGTTTAGTATGTAAAAAAGT
ATCCGGAAAGAGGCGGCCACCTGATTCCGGATATGAAATCTGAATTCTTTG
CGTTAAAGCTTATGCTTTACTTTTTGGGAAATAAATATTATGATTTAACACT

+1 -7

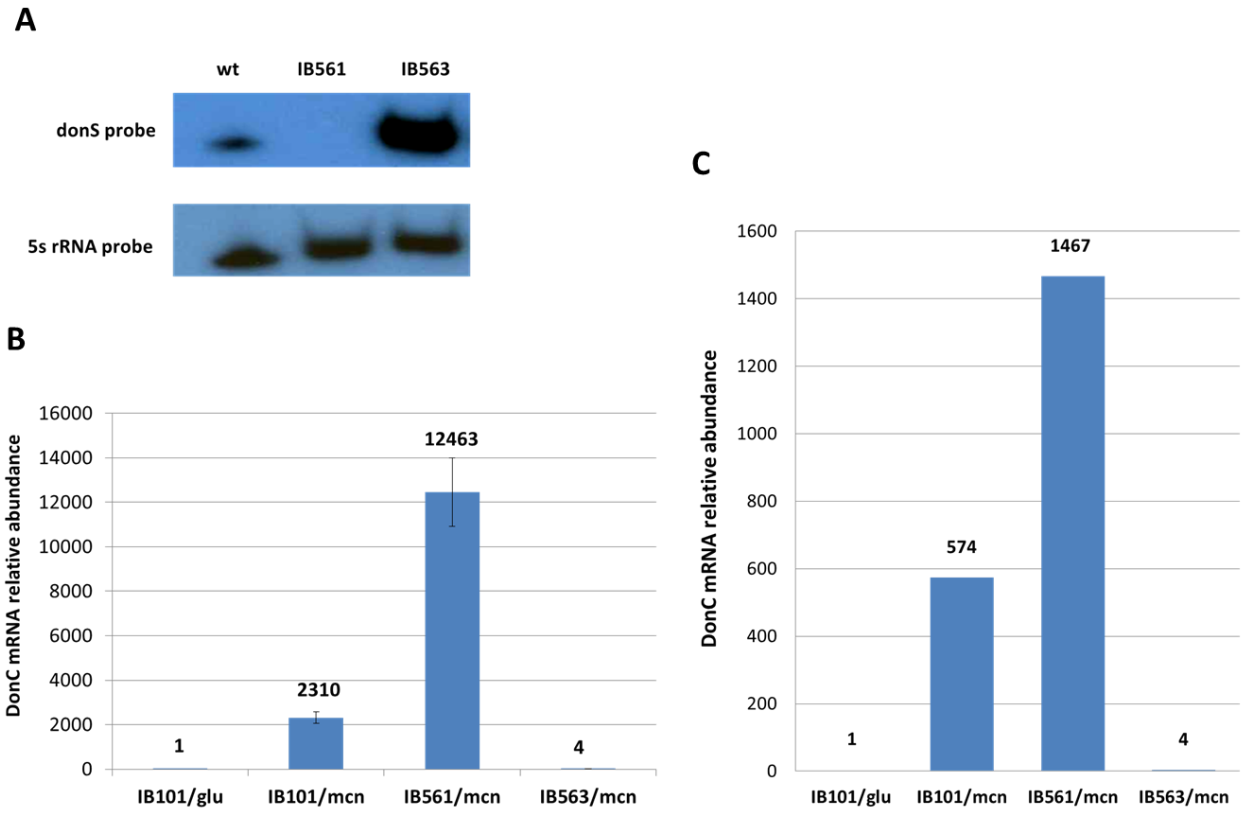
AAAATTCGTT*CAAATCT*ATGAATTTTTATCGATTAAAGTCTATTTTTTTTATA
ATTTTTAGCTGTTTTATTCTACAAACGGCTGCCTTTGCTCAAAACAACGTAA
AAATAACAATCAAGAAAAAGAATATCACGTTGCAGGAAGCATTGCGGGAG
GTTGAGAAACAATCTGATTATCTGATCGCTTTCAACGAATCCAAACTTGAG
AAAACCAAGCGCGTTAACCTGAATATTAATGCTGAATCTCTGGATAAGACA

This sRNA is divergently transcribed and located immediately upstream from *donC*. The TIS ( from primary transcriptome data) of the sRNA is only 9 bases from the *donC* start codon and is adjacent to a consensus primary sigma factor recognition sequence (Fig. 4.2C). The location and orientaton of this sRNA suggests that it is a cis-encoded antisense RNA that plays a role in regulation of the *don* PUL. Since this sRNA is in the *don* PUL, it was designated DonS.

### 4.3.3 DonS is a *cis*-encoded antisense RNA that negatively regulates the *don* operon.

The location and orientation of DonS suggests that it is a regulator for the *don* PUL. To explore this role, a DonS silencing strain (IB561) and a DonS overexpression strain (IB563) were constructed. As shown in Fig. 4.3A, the sRNA northern blot confirmed that DonS expression was abolished in IB561 and DonS was highly overexpressed in strain IB563. To test the effect of DonS on *don* expression, qRT-PCR was performed to check the *donC* transcript level. As before, the expression of *donC* in the wild type strain under non-inducing condition (glucose defined medium) was used as a baseline control. The results in Fig. 4.3B show that expression of *donC* was up-regulated further in strain IB561 in the inducing condition (mucin glycan defined medium) compared to the wild type. In contrast, *donC* expression in strain IB563 decreased to background levels even for cells grown in the inducing condition. Gene expression microarray analysis (Fig. 4.3C) showed similar trend as the qRT-PCR analysis. Taken together, these results indicated that DonS acts as a negative regulator for the *don* PUL expression.

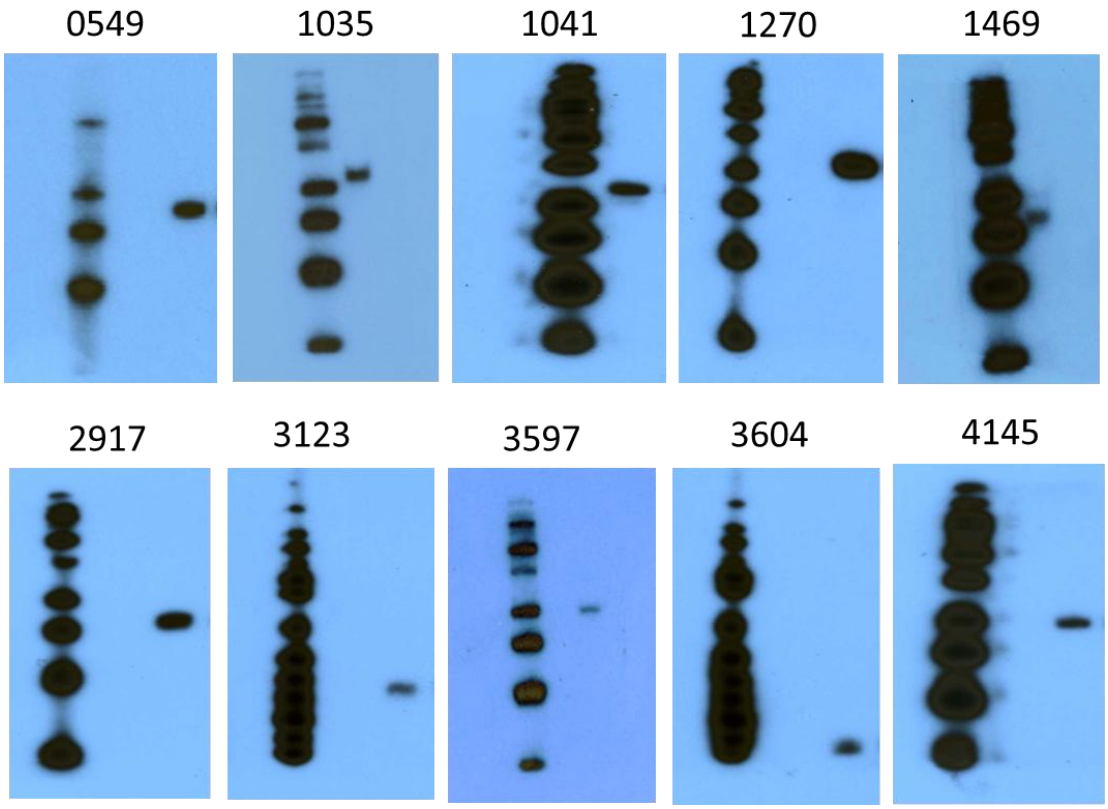
**Fig. 4.3. DonS is a negative regulator of the *don* expression. A)** sRNA northern blot. DonS expression is abolished in IB561, and highly overexpressed in IB563. 5s rRNA was used as a loading control. 10 µg of total RNA was analyzed for each sample. γ-$^{32}$P labeled probes (Table 4.3) were used for detection. **B)** qRT-PCR analysis of *donC* expression in the wild type, DonS silencing mutant and DonS overexpression mutant. qRT-PCR reactions using *donC* primers were performed in triplicate. The *donC* expression level in wild type strain grown under non-inducing condition (glucose defined medium) was used as a baseline. Results are the average of two biological repeats. Error bars represent the standard deviation. glu, cells grown in glucose defined medium; mcn, cells grown in mucin glycan defined medium. **C)** Gene expression microarray analysis result of *donC* expression in wild type, DonS silencing mutant and DonS overexpression mutant. The *donC* expression level in wild type under non-inducing condition (glucose defined medium) was used as a baseline. glu, cells grown in glucose defined medium; mcn, cells grown in mucin glycan defined medium.


A
wt
IB561
IB563
donS probe
5s rRNA probe
B
DonC mRNA relative abundance
16000
14000
12000
10000
8000
6000
4000
2000
0
1
2310
12463
4
IB101/glu
IB101/mcn
IB561/mcn
IB563/mcn
C
DonC mRNA relative abundance
1600
1400
1200
1000
800
600
400
200
0
1
574
1467
4
IB101/glu
IB101/mcn
IB561/mcn
IB563/mcn

### 4.3.4 DonS-like sRNAs are found in other PULs of *B. fragilis*

As mentioned in chapter 3, a group of putative sRNA signals in 14 PULs was identified when analyzing the primary transcriptome data of *B. fragilis*. To confirm the existence of the other sRNAs in this group, probes were designed for 10 of these (Table 4.4) and northern blots performed. As shown in Fig. 4. 4, 10 sRNAs with size range from 75 bases to 130 bases in this group were confirmed by northern blot analysis. These sRNAs have been named according to the SusC-like gene they are adjacent to. The other three putative sRNAs in this group (BF638R_3799, 3787 and 3821) will be tested in the future.

**Fig. 4.4**. **Northern blot confirmation of the DonS-like sRNAs in *B. fragilis*.** 10 µg of total RNA from mid-logarithmic phase cultures grown in BHIS medium was loaded each lane for analysis. The left lane of each panel is γ-$^{32}$P labeled low molecular weight DNA ladder. The right lane of each panel shows the specific sRNA species to that PUL using γ-$^{32}$P labeled probes specific to that sRNA (Table 4.3). The numbers above each panel are the BF638R gene designation of the SusC-like gene of each PUL.


0549
1035
1041
1270
1469
2917
3123
3597
3604
4145

This group of DonS-like sRNAs have several common features. First, they are all divergently transcribed from the *susC* homolog of the PUL. Their orientation and proximity to the *susC* homolog makes it likely that these are cis-encoded antisense sRNA that will negatively regulate that PUL expression like DonS. Second, they all have the *B. fragilis* primary sigma factor -7 promoter recognition sequence "TAxxTTTG" 5 to 9 bases upstream of their TIS and they are relatively abundant as estimated from the primary transcriptome data and the northern blot results. Third, many of the PULs in which they are located may be involved in utilization of host glycans as shown for Don. The evidence for this is that some of these PULs contain genes for glycoside hydrolase that target glycosidic linkage specific for host derived glycans. For example, *donC* in the *don* PUL encodes an endo-beta-N-acetylglucosaminidase homolog which specifically hydrolyzes the glycosidic bond between the two N-acetyglucosamine residuals on the host complex N-linked glycoproteins such as transferrin. A summary of these DonS-like sRNAs is shown in Table 4.4.

**Table 4.4. Characteristics of the DonS-like sRNAs in *B. fragilis*.** Except DonS, these sRNAs have been named according to the specific *susC* homolog from which they are divergently transcribed.

| Name | Confirmed by northern blot | Approx. Size (nucleotides) | Glycosyl hydrolase in the PUL | Potential substrate of the PUL |
|---|---|---|---|---|
| DonS | Yes | 125 | Endo-β-N-acetylglucosaminidase | Complex N-linked serum glycoproteins (confirmed) |
| 0549 | Yes | 95 | Unknown function | N/A |
| 1035 | Yes | 127 | Unknown function | N/A |
| 1041 | Yes | 103 | Arylsulfatase | Glycosaminoglycan |
| 1270 | Yes | 100 | Sulfatase | Glycosaminoglycan |
| 1496 | Yes | 100 | Unknown function | N/A |
| 2917 | Yes | 75 | β-glucoside glucohydrolase | N/A |
| 3123 | Yes | 90 | Chitobiase | Glycans contain N-acetylglucosamine repeats |
| 3597 | Yes | 97 | Sulfatase | Glycosaminoglycan |
| 3604 | Yes | 140 | Chitobiase | Glycans contain N-acetylglucosamine repeats |
| 3787 | No | 95 | Sialidase & Neuraminidase | Glycoproteins or gangliosides |
| 3799 | No | 98 | Arylsulfatase | Glycosaminoglycan |
| 3821 | No | 105 | *Laminin G | Glycosaminoglycan |
| 4145 | Yes | 97 | β-Galactosidase | N/A |

***** Laminin G is a protein domain which specifically binds to some type of glycosaminoglycans such as herparin (115).

## 4.4 Discussion

This study clearly showed that the ECF sigma factor and anti-sigma factor genes, *donA* and *donB*, encode regulators for the downstream outer membrane complex genes in the *don* PUL. Constitutive over-expression of *donA* in the absence of inducing substrate significantly induced transcription of the *donBCDEFG* operon. In contrast, deletion of *donA* completely abolished expression of the operon regardless of the presence of inducing substrate. In the anti-sigma deletion mutant *ΔdonB*, there was an 80 fold increase in *donC* expression compared to the wild type in non-inducing condition. This level of induction is significant but it is not comparable to induction in the wild type under the inducing condition (several hundred to several thousands). This low induction could be due to the fact that the anti-sigma factor DonB is a FecR type regulator which interacts with its cognate sigma factor to fully activate the sigma factor's function in a process called sigma factor maturation (116). Notably, *donC* expression in the anti-sigma deletion mutant was induced to the same level as the sigma over-expression mutant under the inducing condition, seemingly like the presence of substrate can overcome the lack of sigma maturation. The increased expression of *donC* in the presence of substrate in all of the three modified strains, the sigma factor deletion strain, anti-sigma deletion strain and sigma factor over-expressing strain, suggests a possibility that the substrate is involved in second pathway to boost the Don complex expression.

The existence of a group of *cis*-encoded antisense sRNAs associated with some of the Sus-like PULs in *B. fragilis* was demonstrated in this report. Each of these sRNAs is divergently transcribed from a SusC-like outer membrane transporter gene on a Sus-like PUL. The sRNA in

the *don* PUL, named DonS, was used as a model and we found this sRNA is likely a *cis*-acting regulatory sRNAs that negatively effects the expression of the downstream outer membrane complex genes. Small regulatory RNAs can affect gene expression in many different ways, including transcriptional and translational regulation (107-110, 117). As showed previously in chapter 2, *donBCDEFG* are transcribed in one operon. When antisense sRNAs are organized between two genes on the same messages such as *donS* in *don*, they usually repress downstream gene expression by transcriptional termination in which the sRNA binds the nascent mRNA and forms a termination loop, causing RNA polymerase to stop further transcription of the message (117, 118). It is very likely that DonS represses the *don* operon expression through a similar mechanism, since our microarray analyses showed that under inducing conditions, *donCDEFG* expression was completely abolished in the DonS over-expressing strain the while the *donB* gene was still normally induced. This could be tested by using northern blot analysis to check the effect of DonS on the *donBCDEFG* message.

Evidently, the presence of the sRNA DonS adds another layer to the regulation of the *don* PUL expression. Different from the ECF sigma/anti-sigma pair DonA/DonB, DonS expression does not respond to the presence of the *don* PUL substrate and it appeared to be constantly expressed at a relatively high level during logarithmic growth phase. However, we observed a significant decrease in DonS expression when cells were grown into stationary phase in defined glucose media (data not shown). This differential expression of DonS could be a response to the varieties of stresses when cells enter stationary phase. It has been well-known that sRNAs can be implicated in integrating environmental stress signals and regulating a plethora of stress responses including temperature stress, metabolite/nutrient stress, envelope/outer membrane

proteins (OMPs) stress, oxidative stress, iron deficiency stress and anaerobic stress (119). Here we suspect these DonS-like sRNAs respond to metabolite/nutrient starvation or perhaps envelope stress. The reason for this is that the genes under control of these sRNAs are PULs which are involved in nutrient uptake and they contain multiple OMPs.

It is also interesting to notice that many of this group of sRNAs may be specialized to regulate PULs which are involved in host derived glycan utilization. This hypothesis is supported by the presence of host derived glycan hydrolases in these PULs. As shown in Table 4. 5, among the PULs which contain a DonS-like sRNA, two encode sulfatases, two encode arysulfatases, and one encodes a laminin G. Sulfatases and arysulfatases has been implicated in hydrolysis of the sulfate linkages on various types of glycosaminoglycans (GAGs) (120), while laminin G is a protein globular domain which possess heparin binding activity (115). One PUL in this group encodes a chitobiase homolog, which specifically hydrolyzes oligomer, trimer and dimer of N-acetyglucosamine molecules into monomers (121, 122). Another PUL in this group encodes two sialidase/neuroaminidase homologs, indicating its substrates are host glycoproteins or glycolipids rather than plant derived glycans (123). The potential substrates of these glycosyl hydrolases, as well as the confirmed glycan substrates of Don (chapter 2), suggest that these DonS-like sRNA regulated PULs are specifically involved in host derived glycan utilization. Notably, by enabling *B. fragilis* to utilize different host glycans, this group of PULs may indirectly affect the body site-specific colonization and growth of the bacterium in its host. For example, we demonstrated in chapter 2 using a rat model that the Don PUL acts as an extraintestinal growth factor by enabling the bacterium efficiently utilization of N-linked serous glycoprotein glycans. Another example is the PUL BF638R_3602-3606 which has a DonS-like sRNA

divergently transcribe from the BF638R_3604 gene (Fig. 4.4; Table 4.4). A recently study showed that this PUL allows *B. fragilis* to establish a resilient colonization in a specific niche within the colonic crypts that serves as a reservoir for stable gut colonization, so it has been named the Commensal Colonization Factor (CCF) (98).

Future directions for this project will be, first, test the hypothesis that DonS has an effect on the *donBCDEFG* message using northern blot analysis. Second, characterize another sRNA-PUL system to support our idea that actually these DonS-like RNAs are regulators of their adjacent PULs, and these PULs are involved in host glycan utilization in *B. fragilis*. Third, explore the possible stress conditions that these DonS-like sRNAs respond to. Forth, since these DonS-like sRNAs have the same -7 promoter sequence, we would like to explore if there is a master regulator for all of them and the possible phenotype when the regulation of the sRNAs are abolished.

# REFERENCES


1. Smith CJ, Rocha ER, Paster BJ. The medically important bacteroides spp. in health and disease - springer. In: Martin Dworkin, Stanley Falkow, Eugene Rosenberg, Karl-Heinz Schleifer, Erko Stackebrandt, editors. The Prokaryotes. Springer New York; 2006. p. 381-427.

2. Gherna R, Woese CR. A partial phylogenetic analysis of the "flavobacter-bacteroides" phylum: Basis for taxonomic restructuring. Syst Appl Microbiol. 1992 Dec;15(4):513-21.

3. Coyne MJ, Chatzidaki-Livanis M, Paoletti LC, Comstock LE. Role of glycan synthesis in colonization of the mammalian gut by the bacterial symbiont bacteroides fragilis. Proc Natl Acad Sci U S A. 2008 Sep 2;105(35):13099-104.

4. Pantosti A, Tzianabos AO, Onderdonk AB, Kasper DL. Immunochemical characterization of two surface polysaccharides of bacteroides fragilis. Infect Immun. 1991 Jun;59(6):2075-82.

5. Pantosti A, Tzianabos AO, Reinap BG, Onderdonk AB, Kasper DL. Bacteroides fragilis strains express multiple capsular polysaccharides. J Clin Microbiol. 1993 Jul;31(7):1850-5.

6. Tzianabos AO, Pantosti A, Baumann H, Brisson JR, Jennings HJ, Kasper DL. The capsular polysaccharide of bacteroides fragilis comprises two ionically linked polysaccharides. J Biol Chem. 1992 Sep 5;267(25):18230-5.

7. Wexler HM. Bacteroides: The good, the bad, and the nitty-gritty. Clin Microbiol Rev. 2007 Oct;20(4):593-621.

8. Patrick S, McKenna JP, O'Hagan S, Dermott E. A comparison of the haemagglutinating and enzymic activities of bacteroides fragilis whole cells and outer membrane vesicles. Microb Pathog. 1996 Apr;20(4):191-202.

9. Domingues RM, Silva e Souza W das,G., Moraes SR, Avelar KE, Hirata Junior R, Fonseca ME, et al. Surface vesicles: A possible function in commensal relations of bacteroides fragilis. Zentralbl Bakteriol. 1997 Apr;285(4):509-17.

10. Berezow AB, Ernst RK, Coats SR, Braham PH, Karimi-Naser LM, Darveau RP. The structurally similar, penta-acylated lipopolysaccharides of porphyromonas gingivalis and bacteroides elicit strikingly different innate immune responses. Microb Pathog. 2009 Aug;47(2):68-77.

11. Coyne MJ, Reinap B, Lee MM, Comstock LE. Human symbionts use a host-like pathway for surface fucosylation. Science. 2005 Mar 18;307(5716):1778-81.

12. Fletcher CM, Coyne MJ, Villa OF, Chatzidaki-Livanis M, Comstock LE. A general O-glycosylation system important to the physiology of a major human intestinal symbiont. Cell. 2009 Apr 17;137(2):321-31.

13. Wexler HM. Outer-membrane pore-forming proteins in gram-negative anaerobic bacteria. Clin Infect Dis. 2002 Sep 1;35(Suppl 1):S65-71.

14. Baughn AD, Malamy MH. A mitochondrial-like aconitase in the bacterium bacteroides fragilis: Implications for the evolution of the mitochondrial krebs cycle. Proc Natl Acad Sci U S A. 2002 Apr 2;99(7):4662-7.

15. Baughn AD, Malamy MH. The strict anaerobe bacteroides fragilis grows in and benefits from nanomolar concentrations of oxygen. Nature. 2004 Jan 29;427(6973):441-4.

16. Xu J, Gordon JI. Honor thy symbionts. Proc Natl Acad Sci U S A. 2003 Sep 2;100(18):10452-9.

17. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, et al. Evolution of mammals and their gut microbes. Science. 2008 Jun 20;320(5883):1647-51.

18. Sonnenburg ED, Zheng H, Joglekar P, Higginbottom SK, Firbank SJ, Bolam DN, et al. Specificity of polysaccharide use in intestinal bacteroides species determines diet-induced microbiota alterations. Cell. 2010 Jun 25;141(7):1241-52.

19. Koropatkin NM, Cameron EA, Martens EC. How glycan metabolism shapes the human gut microbiota. Nat Rev Microbiol. 2012 Apr 11;10(5):323-35.

20. Martens EC, Chiang HC, Gordon JI. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. Cell Host Microbe. 2008 Nov 13;4(5):447-57.

21. Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. Nat Rev Microbiol. 2008 Feb;6(2):121-31.

22. Rogers TE, Pudlo NA, Koropatkin NM, Bell JS, Moya Balasch M, Jasker K, et al. Dynamic responses of bacteroides thetaiotaomicron during growth on glycan mixtures. Mol Microbiol. 2013 Jun;88(5):876-90.

23. McNeil NI. The contribution of the large intestine to energy supplies in man. Am J Clin Nutr. 1984 Feb;39(2):338-42.

24. Hooper LV, Gordon JI. Commensal host-bacterial relationships in the gut. Science. 2001 May 11;292(5519):1115-8.

25. Cash HL, Whitham CV, Behrendt CL, Hooper LV. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science. 2006 Aug 25;313(5790):1126-30.

26. Hooper LV, Stappenbeck TS, Hong CV, Gordon JI. Angiogenins: A new class of microbicidal proteins involved in innate immunity. Nat Immunol. 2003 Mar;4(3):269-73.

27. Comstock LE. Importance of glycans to the host-bacteroides mutualism in the mammalian intestine. Cell Host Microbe. 2009 Jun 18;5(6):522-6.

28. Dasgupta S, Kasper DL. Novel tools for modulating immune responses in the host-polysaccharides from the capsule of commensal bacteria. Adv Immunol. 2010;106:61-91.

29. Troy EB, Kasper DL. Beneficial effects of bacteroides fragilis polysaccharides on the immune system. Front Biosci (Landmark Ed). 2010 Jan 1;15:25-34.

30. Hsiao EY, McBride SW, Hsien S, Sharon G, Hyde ER, McCue T, et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. Cell. 2013 Dec 19;155(7):1451-63.

31. Begley M, Hill C, Gahan CG. Bile salt hydrolase activity in probiotics. Appl Environ Microbiol. 2006 Mar;72(3):1729-38.

32. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A. 2005 Aug 2;102(31):11070-5.

33. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Microbial ecology: Human gut microbes associated with obesity. Nature. 2006 Dec 21;444(7122):1022-3.

34. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006 Dec 21;444(7122):1027-31.

35. Goldstein EJ. Anaerobic bacteremia. Clin Infect Dis. 1996 Dec;23 Suppl 1:S97-101.

36. Finlay-Jones JJ, Davies KV, Sturm LP, Kenny PA, Hart PH. Inflammatory processes in a murine model of intra-abdominal abscess formation. J Leukoc Biol. 1999 Oct;66(4):583-7.

37. Comstock LE, Tzianabos AO. Abscesses. In: Nataro JP, Blaser MJ, Cunningham-Rundles S, editors. Persistent Bacterial Infections. Washington, DC: ASM Press; 2000. p. 397-407.

38. Sears CL. Enterotoxigenic bacteroides fragilis: A rogue among symbiotes. Clin Microbiol Rev. 2009 Apr;22(2):349,69, Table of Contents.

39. Onderdonk AB, Bartlett JG, Louie T, Sullivan-Seigler N, Gorbach SL. Microbial synergy in experimental intra-abdominal abscess. Infect Immun. 1976 Jan;13(1):22-6.

40. Reid JH, Patrick S. Phagocytic and serum killing of capsulate and non-capsulate bacteroides fragilis. J Med Microbiol. 1984 Jun;17(3):247-57.

41. Comstock LE, Coyne MJ, Tzianabos AO, Pantosti A, Onderdonk AB, Kasper DL. Analysis of a capsular polysaccharide biosynthesis locus of bacteroides fragilis. Infect Immun. 1999 Jul;67(7):3525-32.

42. Roche-Hakansson H, Chatzidaki-Livanis M, Coyne MJ, Comstock LE. Bacteroides fragilis synthesizes a DNA invertase affecting both a local and a distant region. J Bacteriol. 2007 Mar;189(5):2119-24.

43. Krinos CM, Coyne MJ, Weinacht KG, Tzianabos AO, Kasper DL, Comstock LE. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature. 2001 Nov 29;414(6863):555-8.

44. Weinacht KG, Roche H, Krinos CM, Coyne MJ, Parkhill J, Comstock LE. Tyrosine site-specific recombinases mediate DNA inversions affecting the expression of outer surface proteins of bacteroides fragilis. Mol Microbiol. 2004 Sep;53(5):1319-30.

45. Coyne MJ, Weinacht KG, Krinos CM, Comstock LE. Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. Proc Natl Acad Sci U S A. 2003 Sep 2;100(18):10446-51.

46. Chatzidaki-Livanis M, Weinacht KG, Comstock LE. Trans locus inhibitors limit concomitant polysaccharide synthesis in the human gut symbiont bacteroides fragilis. Proc Natl Acad Sci U S A. 2010 Jun 29;107(26):11976-80.

47. Chatzidaki-Livanis M, Coyne MJ, Comstock LE. A family of transcriptional antitermination factors necessary for synthesis of the capsular polysaccharides of bacteroides fragilis. J Bacteriol. 2009 Dec;191(23):7288-95.

48. Comstock LE, Kasper DL. Bacterial glycans: Key mediators of diverse host immune responses. Cell. 2006 Sep 8;126(5):847-50.

49. Tzianabos AO, Onderdonk AB, Rosner B, Cisneros RL, Kasper DL. Structural features of polysaccharides that induce intra-abdominal abscesses. Science. 1993 Oct 15;262(5132):416-9.

50. Coyne MJ, Kalka-Moll W, Tzianabos AO, Kasper DL, Comstock LE. Bacteroides fragilis NCTC9343 produces at least three distinct capsular polysaccharides: Cloning, characterization,

and reassignment of polysaccharide B and C biosynthesis loci. Infect Immun. 2000 Nov;68(11):6176-81.

51. Tzianabos AO, Kasper DL, Onderdonk AB. Structure and function of bacteroides fragilis capsular polysaccharides: Relationship to induction and prevention of abscesses. Clin Infect Dis. 1995 Jun;20 Suppl 2:S132-40.

52. Coyne MJ, Tzianabos AO, Mallory BC, Carey VJ, Kasper DL, Comstock LE. Polysaccharide biosynthesis locus required for virulence of bacteroides fragilis. Infect Immun. 2001 Jul;69(7):4342-50.

53. Storz G, Imlay JA. Oxidative stress. Curr Opin Microbiol. 1999 Apr;2(2):188-94.

54. Rocha ER, Herren CD, Smalley DJ, Smith CJ. The complex oxidative stress response of bacteroides fragilis: The role of OxyR in control of gene expression. Anaerobe. 2003 Aug;9(4):165-73.

55. Sund CJ, Rocha ER, Tzianabos AO, Wells WG, Gee JM, Reott MA, et al. The bacteroides fragilis transcriptome response to oxygen and H2O2: The role of OxyR and its effect on survival and virulence. Mol Microbiol. 2008 Jan;67(1):129-42.

56. Rocha ER, Tzianabos AO, Smith CJ. Thioredoxin reductase is essential for thiol/disulfide redox control and oxidative stress survival of the anaerobe bacteroides fragilis. J Bacteriol. 2007 Nov;189(22):8015-23.

57. Reott MA, Parker AC, Rocha ER, Smith CJ. Thioredoxins in redox maintenance and survival during oxidative stress of bacteroides fragilis. J Bacteriol. 2009 May;191(10):3384-91.

58. Ndamukong IC, Gee J, Smith CJ. The extracytoplasmic function sigma factor EcfO protects bacteroides fragilis against oxidative stress. J Bacteriol. 2013 Jan;195(1):145-55.

59. Spence C, Wells WG, Smith CJ. Characterization of the primary starch utilization operon in the obligate anaerobe bacteroides fragilis: Regulation by carbon source and oxygen. J Bacteriol. 2006 Jul;188(13):4663-72.

60. Lobo LA, Jenkins AL, Jeffrey Smith C, Rocha ER. Expression of bacteroides fragilis hemolysins in vivo and role of HlyBA in an intra-abdominal infection model. Microbiologyopen. 2013 Apr;2(2):326-37.

61. Ravcheev DA, Godzik A, Osterman AL, Rodionov DA. Polysaccharides utilization in human gut bacterium bacteroides thetaiotaomicron: Comparative genomics reconstruction of metabolic and regulatory networks. BMC Genomics. 2013 Dec 12;14:873,2164-14-873.

62. Martens EC, Roth R, Heuser JE, Gordon JI. Coordinate regulation of glycan degradation and polysaccharide capsule biosynthesis by a prominent human gut symbiont. J Biol Chem. 2009;284(27):18445-57.

63. Cho KH, Salyers AA. Biochemical analysis of interactions between outer membrane proteins that contribute to starch utilization by bacteroides thetaiotaomicron. J Bacteriol. 2001 Dec;183(24):7224-30.

64. Shipman JA, Cho KH, Siegel HA, Salyers AA. Physiological characterization of SusG, an outer membrane protein essential for starch utilization by bacteroides thetaiotaomicron. J Bacteriol. 1999 Dec;181(23):7206-11.

65. Shipman JA, Berleman JE, Salyers AA. Characterization of four outer membrane proteins involved in binding starch to the cell surface of bacteroides thetaiotaomicron. J Bacteriol. 2000 Oct;182(19):5365-72.

66. Reeves AR, Wang GR, Salyers AA. Characterization of four outer membrane proteins that play a role in utilization of starch by bacteroides thetaiotaomicron. J Bacteriol. 1997 Feb;179(3):643-9.

67. D'Elia JN, Salyers AA. Effect of regulatory protein levels on utilization of starch by bacteroides thetaiotaomicron. J Bacteriol. 1996 Dec;178(24):7180-6.

68. Martens EC, Koropatkin NM, Smith TJ, Gordon JI. Complex glycan catabolism by the human gut microbiota: The bacteroidetes sus-like paradigm. J Biol Chem. 2009 Sep 11;284(37):24673-7.

69. Hylemon PB, Phibbs PV,Jr. Evidence against the presence of cyclic AMP and related enzymes in selected strains of bacteroides fragilis. Biochem Biophys Res Commun. 1974 Sep 9;60(1):88-95.

70. Lassmann B, Gustafson DR, Wood CM, Rosenblatt JE. Reemergence of anaerobic bacteremia. Clin Infect Dis. 2007 Apr 1;44(7):895-900.

71. Brook I. Recovery of anaerobic bacteria from clinical specimens in 12 years at two military hospitals. J Clin Microbiol. 1988 Jun;26(6):1181-8.

72. Gorbach SL, Bartlett JG. Anaerobic infections (second of three parts). N Engl J Med. 1974 May 30;290(22):1237-45.

73. Polk BF, Kasper DL. Bacteroides fragilis subspecies in clinical isolates. Ann Intern Med. 1977 May;86(5):569-71.

74. Varel VH, Bryant MP. Nutritional features of bacteroides fragilis subsp. fragilis. Appl Microbiol. 1974 Aug;28(2):251-7.

75. Privitera G, Dublanchet A, Sebald M. Transfer of multiple antibiotic resistance between subspecies of bacteroides fragilis. J Infect Dis. 1979 Jan;139(1):97-101.

76. Johnson JL. Taxonomy of the bacteroides: I. deoxyribonucleic acid homologies among bacteroides fragilis and other saccharolytic bacteroides species. Int J Syst Bacteriol. 1978;28(2):245-56.

77. Bedzyk LA, Shoemaker NB, Young KE, Salyers AA. Insertion and excision of bacteroides conjugative chromosomal elements. J Bacteriol. 1992 Jan;174(1):166-72.

78. Jacobs MR, Spangler SK, Appelbaum PC. Beta-lactamase production, beta-lactam sensitivity and resistance to synergy with clavulanate of 737 bacteroides fragilis group organisms from thirty-three US centres. J Antimicrob Chemother. 1990 Sep;26(3):361-70.

79. Guiney DG, Hasegawa P, Davis CE. Plasmid transfer from escherichia coli to bacteroides fragilis: Differential expression of antibiotic resistance phenotypes. Proc Natl Acad Sci U S A. 1984 Nov;81(22):7203-6.

80. Smith CJ, Rollins LA, Parker AC. Nucleotide sequence determination and genetic analysis of the bacteroides plasmid, pBI143. Plasmid. 1995 Nov;34(3):211-22.

81. Bamberger DM, Herndon BL, Fitch J, Florkowski A, Parkhurst V. Effects of neutrophils on cefazolin activity and penicillin-binding proteins in staphylococcus aureus abscesses. Antimicrob Agents Chemother. 2002 Sep;46(9):2878-84.

82. Rocha ER, Smith CJ. Regulation of bacteriodes fragilis katB mRNA by oxidative stress and carbon limitation. J Bacteriol. 1997 Nov;179(22):7033-9.

83. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001 May 1;29(9):e45.

84. Furukawa J, Shinohara Y, Kuramoto H, Miura Y, Shimaoka H, Kurogochi M, et al. Comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential tag conversion. Anal Chem. 2008 Feb 15;80(4):1094-101.

85. Nishimura S, Niikura K, Kurogochi M, Matsushita T, Fumoto M, Hinou H, et al. High-throughput protein glycomics: Combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. Angew Chem Int Ed Engl. 2004 Dec 17;44(1):91-6.

86. Godoy VG, Dallas MM, Russo TA, Malamy MH. A role for bacteroides fragilis neuraminidase in bacterial growth in two model systems. Infect Immun. 1993 Oct;61(10):4415-26.

87. Nakayama-Imaohji H, Ichimura M, Iwasa T, Okada N, Ohnishi Y, Kuwahara T. Characterization of a gene cluster for sialoglycoconjugate utilization in bacteroides fragilis. J Med Invest. 2012;59(1-2):79-94.

88. Brigham C, Caughlan R, Gallegos R, Dallas MB, Godoy VG, Malamy MH. Sialic acid (N-acetyl neuraminic acid) utilization by bacteroides fragilis requires a novel N-acetyl mannosamine epimerase. J Bacteriol. 2009 Jun;191(11):3629-38.

89. Collin M, Svensson MD, Sjoholm AG, Jensenius JC, Sjobring U, Olsen A. EndoS and SpeB from streptococcus pyogenes inhibit immunoglobulin-mediated opsonophagocytosis. Infect Immun. 2002 Dec;70(12):6646-51.

90. Collin M, Olsen A. EndoS, a novel secreted protein from streptococcus pyogenes with endoglycosidase activity on human IgG. EMBO J. 2001 Jun 15;20(12):3046-55.

91. Burnaugh AM, Frantz LJ, King SJ. Growth of streptococcus pneumoniae on human glycoconjugates is dependent upon the sequential activity of bacterial exoglycosidases. J Bacteriol. 2008 Jan;190(1):221-30.

92. King SJ, Hippe KR, Weiser JN. Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by streptococcus pneumoniae. Mol Microbiol. 2006 Feb;59(3):961-74.

93. Renzi F, Manfredi P, Mally M, Moes S, Jeno P, Cornelis GR. The N-glycan glycoprotein deglycosylation complex (gpd) from capnocytophaga canimorsus deglycosylates human IgG. PLoS Pathog. 2011 Jun;7(6):e1002118.

94. Lawrence T3, Biggers C, Simonton P. Bacteriostatic inhibition of klebsiella pneumoniae by three human transferrins. 1977. annuals of human biology;4(3):281-4.

95. Martin C, Finland M, Jandl J. Enhancement of acute bacterial infections in rats and mice by iron and their inhibition by human transferrin. Journal of infectious diseases. 1963;112(2):158.

96. Padda JS, Schryvers AB. N-linked oligosaccharides of human transferrin are not required for binding to bacterial transferrin receptors. Infect Immun. 1990 Sep;58(9):2972-6.

97. Zhao N, Enns CA. N-linked glycosylation is required for transferrin-induced stabilization of transferrin receptor 2, but not for transferrin binding or trafficking to the cell surface. Biochemistry. 2013 May 14;52(19):3310-9.

98. Lee SM, Donaldson GP, Mikulski Z, Boyajian S, Ley K, Mazmanian SK. Bacterial colonization factors control specificity and stability of the gut microbiota. Nature. 2013 Sep 19;501(7467):426-9.

99. Namavar F, Theunissen EB, Verweij-Van Vught AM, Peerbooms PG, Bal M, Hoitsma HF, et al. Epidemiology of the bacteroides fragilis group in the colonic flora in 10 patients with colonic cancer. J Med Microbiol. 1989 Jul;29(3):171-6.

100. Derrien M, van Passel MW, van de Bovenkamp JH, Schipper RG, de Vos WM, Dekker J. Mucin-bacterial interactions in the human oral cavity and digestive tract. Gut Microbes. 2010 Jul;1(4):254-68.

101. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, et al. The primary transcriptome of the major human pathogen helicobacter pylori. Nature. 2010 Mar 11;464(7286):250-5.

102. Hoffmann S, Otto C, Kurtz S, Sharma CM, Khaitovich P, Vogel J, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS Comput Biol. 2009 Sep;5(9):e1000502.

103. Deana A, Celesnik H, Belasco JG. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. Nature. 2008 Jan 17;451(7176):355-8.

104. Deutscher MP. Degradation of RNA in bacteria: Comparison of mRNA and stable RNA. Nucleic Acids Res. 2006 Feb 1;34(2):659-66.

105. Bayley DP, Rocha ER, Smith CJ. Analysis of cepA and other bacteroides fragilis genes reveals a unique promoter structure. FEMS Microbiol Lett. 2000 Dec 1;193(1):149-54.

106. Vingadassalom D, Kolb A, Mayer C, Rybkine T, Collatz E, Podglajen I. An unusual primary sigma factor in the bacteroidetes phylum. Mol Microbiol. 2005 May;56(4):888-902.

107. Georg J, Hess WR. Cis-antisense RNA, another level of gene regulation in bacteria. Microbiol Mol Biol Rev. 2011 Jun;75(2):286-300.

108. Waters LS, Storz G. Regulatory RNAs in bacteria. Cell. 2009 Feb 20;136(4):615-28.

109. Brantl S. Regulatory mechanisms employed by cis-encoded antisense RNAs. Curr Opin Microbiol. 2007 Apr;10(2):102-9.

110. Liu JM, Camilli A. A broadening world of bacterial small RNAs. Curr Opin Microbiol. 2010 Feb;13(1):18-23.

111. Schneider DA, Ross W, Gourse RL. Control of rRNA expression in escherichia coli. Curr Opin Microbiol. 2003 Apr;6(2):151-6.

112. Hirvonen CA, Ross W, Wozniak CE, Marasco E, Anthony JR, Aiyar SE, et al. Contributions of UP elements and the transcription factor FIS to expression from the seven rrn P1 promoters in escherichia coli. J Bacteriol. 2001 Nov;183(21):6305-14.

113. Caswell CC, Gaines JM, Ciborowski P, Smith D, Borchers CH, Roux CM, et al. Identification of two small regulatory RNAs linked to virulence in brucella abortus 2308. Mol Microbiol. 2012 Jul;85(2):345-60.

114. Smith CJ, Rogers MB, McKee ML. Heterologous gene expression in bacteroides fragilis. Plasmid. 1992 Mar;27(2):141-54.

115. Tisi D, Talts JF, Timpl R, Hohenester E. Structure of the C-terminal laminin G-like domain pair of the laminin alpha2 chain harbouring binding sites for alpha-dystroglycan and heparin. EMBO J. 2000 Apr 3;19(7):1432-40.

116. Ochs M, Veitinger S, Kim I, Welz D, Angerer A, Braun V. Regulation of citrate-dependent iron transport of escherichia coli: fecR is required for transcription activation by FecI. Mol Microbiol. 1995 Jan;15(1):119-32.

117. Storz G, Opdyke JA, Wassarman KM. Regulating bacterial transcription with small RNAs. Cold Spring Harb Symp Quant Biol. 2006;71:269-73.

118. Brantl S, Wagner EG. An antisense RNA-mediated transcriptional attenuation mechanism functions in escherichia coli. J Bacteriol. 2002 May;184(10):2740-7.

119. Hoe CH, Raabe CA, Rozhdestvensky TS, Tang TH. Bacterial sRNAs: Regulation in stress. Int J Med Microbiol. 2013 Jul;303(5):217-29.

120. Hanson SR, Best MD, Wong CH. Sulfatases: Structure, mechanism, biological activity, inhibition, and synthetic utility. Angew Chem Int Ed Engl. 2004 Nov 5;43(43):5736-63.

121. Jenkinson HF, Shepherd MG. A mutant of cmdida albicans deficient in P-N-acetylglucosaminidase. Journal of general microbiology. 1987;133:2097-106.

122. Hanson ML, Lagadic L. Chitobiase activity as an indicator of aquatic ecosystem health. Aquat Ecosyst Health Manage. 2005;8(4):441-50.

123. Zeleny R, Kolarich D, Strasser R, Altmann F. Sialic acid concentrations in plants are in the range of inadvertent contamination. Planta. 2006 Jun;224(1):222-7.

124. Larsbrink J, Rogers TE, Hemsworth GR, McKee LS, Tauzin AS, Spadiut O, et al. A discrete genetic locus confers xyloglucan metabolism in select human gut bacteroidetes. Nature. 2014; 506(7489):498-502.

125. Rao V, Cui T, Guan C, Van Roey P. Mutations of endo-beta-N-acetylglucosaminidase H active site residueAs sp130 anG glu132: activities and conformations. Protein Science. 1998;8(11): 2338-46

# APPENDIX 1: COPYRIGHT PERMISSION POLICY OF THE AMERICAN SOCIETYFOR BIOCHEMISTRY AND MOLECULAR BIOLOGY

## Copyright Permission Policy

These guidelines apply to the reuse of articles, figures, charts and photos in the ***Journal of Biological Chemistry*** ***Molecular & Cellular Proteomics*** and the ***Journal of Lipid Research***

### For authors reusing their own material:

Authors need **NOT** contact the journal to obtain rights to reuse their own material. They are automatically granted permission to do the following:

- Reuse the article in print collections of their own writing.
- Present a work orally in its entirety.
- Use an article in a thesis and/or dissertation.
- Reproduce an article for use in the author's courses. (If the author is employed by an academic institution, that institution also may reproduce the article for teaching purposes.)
- Reuse a figure, photo and/or table in future commercial and noncommercial works.
- Post a copy of the paper in PDF that you submitted via BenchPress.
- Link to the journal site containing the final edited PDFs created by the publisher.

EXCEPTION: If authors select the Author's Choice publishing option:

- The final version of the manuscript will be covered under the Creative Commons Attribution license (CC BY), the most accommodating of licenses offered. Click here for details.
- The final version of the manuscript will be released immediately on the publisher's website and PubMed Central.

Please note that authors must include the following citation when using material that appeared in an ASBMB journal:

"This research was originally published in Journal Name. Author(s). Title. *Journal Name*. Year, Vol:pp–pp. © the American Society for Biochemistry and Molecular Biology."

### For other parties using material for noncommercial use:

Other parties are welcome to copy, distribute, transmit and adapt the work — at no cost and without permission — for noncommercial use as long as they attribute the work to the original source using the citation above.

Examples of noncommercial use include:

- Reproducing a figure for educational purposes, such as schoolwork or lecture presentations, with attribution.
- Appending a reprinted article to a Ph.D. dissertation, with attribution.

### For other parties using material for commercial use:

Navigate to the article of interest and click the "Request Permissions" button on the middle navigation bar. (See diagram at right.) It will walk you through the steps for obtaining permission for reuse.

Examples of commercial use by parties other than authors include:

- Reproducing a figure in a book published by a commercial publisher.
- Reproducing a figure in a journal article published by a commercial publisher.

Updated March 20, 2013

# APPENDIX 2: INDUCTION AND REPRESSION OF *B. FRAGILIS* 638R PUL GENES IN VIVO AND IN VITRO

*All values were calculated from the expression microarray data deposited at the NCBI GEO, accession # GSE53883. B. fragilis 638R wild type cells were grown in vitro in defined medium with either 0.4% glucose (DM-G) or 2% mucin glycans (DM-M). B. fragilis 638R cells were grown in vivo in the rat tissue cage for 1 (D1), 4 (D4), or 8 (D8) days. Values given in columns C-F are the ratios (induction or repression) of the RNA normalized Log2 expression values obtained from at least two biological replicates.

| Gene | FUNCTION | *DM-M/DM-G | D1/DM-G | D4/DM-G | D8/DM-G |
|---|---|---|---|---|---|
| BF638R0283 | exported alpha-galactosidase | 0.84 | 0.90 | 1.70 | 2.85 |
| BF638R0284 | conserved hypothetical exported protein | 0.84 | 0.60 | 1.28 | 2.08 |
| BF638R0285 | SusC-like | 0.61 | 0.82 | 2.58 | 4.84 |
| BF638R0286 | regulatory protein | 0.80 | 0.61 | 19.93 | 15.68 |
| BF638R0287 | ECF-type sigma factor | 1.10 | 0.74 | 0.83 | 1.09 |
| BF638R0288 | exported alpha-galactosidase | 1.63 | 0.94 | 1.18 | 0.86 |
| BF638R0345 | SusC-like | 2.29 | 0.85 | 0.17 | 0.16 |
| BF638R0346 | lipoprotein | 1.96 | 0.70 | 0.38 | 0.34 |
| BF638R0347 | exported protein | 2.14 | 0.40 | 0.23 | 0.24 |
| BF638R0348 | exported protein | 2.18 | 0.91 | 1.03 | 0.75 |
| BF638R0349 | exported endo-arabinase | 0.79 | 0.69 | 0.75 | 0.65 |
| BF638R0384 | conserved hypothetical exported protein | 1.99 | 1.61 | 0.40 | 0.33 |
| BF638R0385 | beta-galactosidase | 2.23 | 2.29 | 0.24 | 0.18 |
| BF638R0386 | exported beta-hexosaminidase | 5.52 | 3.61 | 0.35 | 0.32 |
| BF638R0387 | conserved hypothetical exported protein | 5.99 | 4.49 | 0.44 | 0.78 |
| BF638R0388 | exported hydrolase | 5.17 | 4.79 | 0.86 | 0.80 |
| BF638R0389 | conserved hypothetical protein | 4.17 | 6.56 | 4.08 | 3.39 |
| BF638R0390 | exported protein | 20.14 | 12.38 | 4.37 | 3.01 |
| BF638R0391 | SusC-like | 37.20 | 21.85 | 6.25 | 3.07 |
| BF638R0392 | two-component system sensor histidine kinase | 1.22 | 3.45 | 3.54 | 2.64 |
| BF638R0393 | two-component sensor histidine kinase | 1.78 | 11.73 | 3.31 | 2.65 |
| BF638R0394 | conserved hypothetical exported protein | 1.41 | 8.51 | 4.32 | 3.41 |
| BF638R0395 | conserved hypothetical exported protein | 1.27 | 2.73 | 1.72 | 1.52 |
| BF638R0397 | exported protein | 1.28 | 1.92 | 1.39 | 1.40 |
| BF638R0398 | SusC-like | 1.22 | 1.74 | 1.33 | 1.26 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R0406 | SusC-like | 0.03 | 0.04 | 0.10 | 0.18 |
| BF638R0407 | outer membrane protein | 0.06 | 0.10 | 0.21 | 0.30 |
| BF638R0421 | exported protein | 4.32 | 3.75 | 1.32 | 1.30 |
| BF638R0422 | exported protein | 5.70 | 4.33 | 1.56 | 1.23 |
| BF638R0423 | outer membrane protein | 10.43 | 6.80 | 1.52 | 1.50 |
| BF638R0424 | SusC-like | 15.08 | 8.96 | 2.96 | 2.29 |
| BF638R0425 | two-component system sensor histidine kinase/response regulator fusion protein | 1.82 | 16.18 | 2.42 | 2.01 |
| BF638R0444 | possible exported xanthan lyase/N-acetylmuramoyl-L-alanine amidase | 4.30 | 2.57 | 0.83 | 0.69 |
| BF638R0445 | cation symporter | 4.23 | 18.45 | 12.65 | 12.29 |
| BF638R0446 | conserved hypothetical exported protein | 4.93 | 6.79 | 11.61 | 14.21 |
| BF638R0447 | conserved hypothetical lipoprotein | 44.85 | 25.01 | 10.22 | 16.46 |
| BF638R0448 | SusC-like | 55.44 | 30.24 | 14.86 | 22.53 |
| BF638R0546 | melibiase | 1.65 | 0.86 | 1.79 | 1.94 |
| BF638R0547 | glycosyl hydrolase, alpha-xylosidase | 2.40 | 0.86 | 0.91 | 0.89 |
| BF638R0548 | outer membrane protein | 2.16 | 0.99 | 1.19 | 0.85 |
| BF638R0549 | SusC-like | 2.97 | 2.36 | 7.81 | 4.60 |
| BF638R0550 | anti-sigma factor | 1.40 | 0.73 | 18.21 | 6.61 |
| BF638R0551 | ECF-type RNA polymerase sigma factor | 1.42 | 1.62 | 0.50 | 0.63 |
| BF638R0555 | ATP-binding component of ABC transporter | 0.42 | 0.23 | 0.13 | 0.27 |
| BF638R0556 | ABC transporter | 0.47 | 0.32 | 0.25 | 0.51 |
| BF638R0564 | SusC-like | 21.32 | 25.78 | 9.83 | 5.33 |
| BF638R0565 | conserved hypothetical protein | 9.58 | 13.48 | 4.87 | 3.46 |
| BF638R0566 | arylsulfatase | 6.60 | 6.71 | 2.58 | 1.69 |
| BF638R0582 | RNA polymerase ECF sigma factor | 0.91 | 0.66 | 0.92 | 1.08 |
| BF638R0583 | anti-sigma factor | 1.18 | 0.79 | 2.15 | 3.34 |
| BF638R0584 | SusC-like | 3.30 | 18.29 | 14.87 | 23.79 |
| BF638R0585 | outer membrane protein | 2.65 | 8.11 | 8.98 | 14.03 |
| BF638R0586 | conserved hypothetical protein | 2.50 | 5.69 | 4.39 | 6.85 |
| BF638R0614 | RNA polymerase ECF-type sigma factor | 1.54 | 1.30 | 2.48 | 3.91 |
| BF638R0615 | anti-sigma factor | 1.33 | 0.43 | 1.62 | 1.24 |
| BF638R0616 | conserved hypothetical protein | 1.32 | 1.77 | 6.72 | 6.50 |
| BF638R0617 | SusC-like | 2.19 | 2.80 | 24.96 | 29.00 |
| BF638R0618 | outer membrane protein | 2.06 | 2.16 | 9.34 | 14.91 |
| BF638R0619 | beta-galactosidase | 1.44 | 0.71 | 1.71 | 3.31 |
| BF638R0620 | glycosyl hydrolase | 1.32 | 0.64 | 0.84 | 1.39 |
| BF638R0621 | glycosyl hydrolase/xylanase | 1.30 | 0.83 | 0.86 | 1.32 |
| BF638R0636 | conserved hypothetical protein | 0.52 | 0.64 | 0.12 | 0.08 |
| BF638R0637 | SusC-like | 0.57 | 0.60 | 0.11 | 0.07 |
| BF638R0638 | membrane protein | 0.53 | 0.99 | 0.33 | 0.29 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R0639 | SusC-like | 0.58 | 0.91 | 0.32 | 0.31 |
| BF638R0641 | SusC-like | 0.70 | 0.77 | 0.23 | 0.17 |
| BF638R0642 | lipoprotein | 0.65 | 0.85 | 0.30 | 0.23 |
| BF638R0643 | GTP-binding protein | 0.69 | 0.30 | 0.48 | 0.38 |
| BF638R0644 | conserved hypothetical protein | 0.62 | 0.52 | 0.28 | 0.24 |
| BF638R0645 | aminopeptidase | 0.96 | 0.41 | 0.16 | 0.16 |
| BF638R0703 | SusC-like | 3.60 | 1.94 | 1.04 | 0.55 |
| BF638R0704 | conserved hypothetical protein | 2.61 | 1.39 | 0.66 | 0.36 |
| BF638R0705 | conserved hypothetical exported protein | 2.20 | 1.31 | 0.80 | 0.59 |
| BF638R0706 | conserved hypothetical exported protein | 1.66 | 0.94 | 0.43 | 0.23 |
| BF638R0722 | conserved hypothetical lipoprotein | 3.17 | 6.13 | 1.76 | 1.95 |
| BF638R0723 | SusC-like | 3.85 | 7.41 | 2.15 | 2.09 |
| BF638R0724 | alpha-1,2-mannosidase precursor | 2.84 | 6.11 | 2.05 | 1.62 |
| BF638R0725 | alpha-1,2-mannosidase precursor | 3.03 | 7.13 | 3.62 | 3.16 |
| BF638R0726 | anti-sigma factor | 0.77 | 2.23 | 3.46 | 4.25 |
| BF638R0727 | RNA polymerase ECF-type sigma factor | 0.59 | 1.17 | 0.65 | 0.43 |
| BF638R0728 | alpha-1,2-mannosidase precursor | 0.64 | 1.17 | 0.96 | 0.85 |
| BF638R0816 | glycosyl hydrolase | 1.68 | 0.74 | 2.88 | 3.75 |
| BF638R0817 | SusC-like | 1.96 | 0.96 | 1.47 | 1.24 |
| BF638R0818 | outer membrane protein | 1.13 | 0.69 | 1.06 | 0.83 |
| BF638R0819 | lipoprotein | 1.08 | 0.51 | 1.00 | 0.97 |
| BF638R0858 | secreted sulfatase | 1.50 | 2.38 | 0.31 | 0.26 |
| BF638R0859 | possible alpha-galactosidase | 1.34 | 2.98 | 0.32 | 0.24 |
| BF638R0860 | conserved hypothetical protein | 2.02 | 2.38 | 0.66 | 0.43 |
| BF638R0861 | sulfatase | 4.55 | 2.10 | 2.88 | 2.00 |
| BF638R0862 | conserved hypothetical protein | 3.68 | 3.97 | 3.52 | 4.19 |
| BF638R0863 | SusC-like | 4.95 | 4.21 | 3.53 | 4.12 |
| BF638R0864 | two-component system sensor histidine kinase/response regulator fusion protein | 0.99 | 0.87 | 0.57 | 0.43 |
| BF638R0920 | beta-glucosidase | 4.59 | 1.69 | 1.59 | 1.37 |
| BF638R0921 | cytochrome c binding protein | 2.95 | 0.71 | 0.47 | 0.40 |
| BF638R0922 | transmembrane protein | 2.41 | 3.45 | 1.04 | 0.81 |
| BF638R0923 | outer membrane protein | 6.39 | 1.63 | 0.80 | 0.62 |
| BF638R0924 | SusC-like | 7.44 | 0.78 | 0.18 | 0.13 |
| BF638R0925 | conserved hypothetical protein | 1.14 | 1.45 | 0.32 | 0.38 |
| BF638R0926 | SusC-like | 1.16 | 1.01 | 0.14 | 0.18 |
| BF638R0927 | possible outer membrane protein | 2.34 | 1.34 | 0.31 | 0.37 |
| BF638R0929 | SusC-like | 2.80 | 1.21 | 0.27 | 0.22 |
| BF638R0930 | RNA polymerase ECF-type sigma factor | 0.62 | 0.58 | 0.52 | 0.63 |
| BF638R0931 | anti-sigma factor | 5.77 | 0.82 | 11.63 | 6.73 |
| BF638R0932 | SusC-like | 3.68 | 0.64 | 0.17 | 0.14 |
| BF638R0933 | possible outer membrane protein | 3.76 | 0.66 | 0.40 | 0.33 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R0949 | conserved hypothetical protein | 3.16 | 2.27 | 0.54 | 0.35 |
| BF638R0950 | SusC-like | 4.16 | 2.69 | 0.33 | 0.25 |
| BF638R0951 | conserved hypothetical protein | 1.43 | 0.82 | 1.04 | 1.61 |
| BF638R0952 | SusC-like | 2.06 | 0.69 | 1.01 | 0.74 |
| BF638R1032 | RNA polymerase ECF-type sigma factor | 0.80 | 1.33 | 2.93 | 2.41 |
| BF638R1033 | anti-sigma factor | 1.65 | 1.13 | 5.94 | 3.39 |
| BF638R1035 | SusC-like | 1.12 | 0.39 | 0.34 | 0.23 |
| BF638R1036 | conserved hypothetical protein | 1.10 | 0.37 | 0.12 | 0.10 |
| BF638R1037 | endonuclease/exonuclease/phosphatase family protein | 1.25 | 0.47 | 0.18 | 0.16 |
| BF638R1038 | exported protein | 1.93 | 0.39 | 0.32 | 0.37 |
| BF638R1039 | RNA polymerase ECF-type sigma factor | 1.47 | 1.12 | 3.34 | 2.11 |
| BF638R1040 | anti-sigma factor | 7.58 | 0.94 | 15.74 | 9.82 |
| BF638R1041 | SusC-like | 29.69 | 5.19 | 5.89 | 3.69 |
| BF638R1042 | outer membrane protein | 18.88 | 3.83 | 3.75 | 2.63 |
| BF638R1043 | arylsulfatase precursor | 14.97 | 4.08 | 4.27 | 2.98 |
| BF638R1045 | arylsulfatase | 1.77 | 0.48 | 1.58 | 1.74 |
| BF638R1264 | alpha-glucosidase, glycosylhydrolase | 2.19 | 1.49 | 1.18 | 1.11 |
| BF638R1265 | ECF-type RNA polymerase sigma factor | 1.17 | 0.34 | 0.22 | 0.23 |
| BF638R1266 | xylosidase/arabinosidase | 2.18 | 0.16 | 0.33 | 0.39 |
| BF638R1267 | beta-lactamase | 0.73 | 0.49 | 0.68 | 1.14 |
| BF638R1268 | hypothetical protein | 0.80 | 0.46 | 0.60 | 0.53 |
| BF638R1269 | anti-sigma factor | 0.56 | 0.39 | 3.33 | 2.04 |
| BF638R1270 | SusC-like | 0.79 | 0.74 | 2.71 | 0.93 |
| BF638R1271 | conserved hypothetical protein | 0.77 | 0.68 | 1.24 | 0.51 |
| BF638R1272 | phosphohydrolase, Icc family | 1.14 | 1.23 | 1.33 | 1.24 |
| BF638R1273 | possible phosphodiesterase/nucleotide pyrophosphatase-like protein | 1.67 | 0.51 | 1.08 | 0.87 |
| BF638R1274 | endonuclease/exonuclease/phosphatase family protein | 1.94 | 0.62 | 0.43 | 0.37 |
| BF638R1323 | RNA polymerase ECF-type sigma factor | 0.76 | 2.33 | 5.17 | 7.21 |
| BF638R1324 | anti-sigma factor | 2.00 | 2.03 | 13.65 | 13.92 |
| BF638R1325 | SusC-like | 20.82 | 21.43 | 16.60 | 20.89 |
| BF638R1326 | conserved hypothetical protein | 20.04 | 17.86 | 12.88 | 17.23 |
| BF638R1327 | possible endo-beta-N-acetylglucosaminidase | 19.35 | 17.54 | 11.15 | 16.70 |
| BF638R1329 | conserved hypothetical protein | 20.68 | 15.05 | 9.62 | 13.49 |
| BF638R1330 | conserved hypothetical protein | 13.81 | 13.37 | 8.34 | 10.29 |
| BF638R1494 | RNA polymerase ECF-type sigma factor | 1.07 | 7.96 | 6.75 | 13.31 |
| BF638R1495 | anti-sigma factor | 1.20 | 1.83 | 11.32 | 4.29 |
| BF638R1496 | SusC-like | 1.07 | 1.47 | 1.04 | 1.05 |
| BF638R1497 | outer membrane protein | 0.88 | 0.54 | 0.37 | 0.34 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R1498 | hypothetical protein | 0.96 | 1.08 | 0.97 | 0.79 |
| BF638R1499 | alpha-galactosidase | 1.03 | 1.68 | 0.99 | 1.04 |
| BF638R1620 | SusC-like | 1.14 | 1.99 | 0.97 | 0.79 |
| BF638R1621 | conserved hypothetical protein | 1.25 | 3.01 | 1.24 | 1.05 |
| BF638R1720 | SusC-like | 5.94 | 62.70 | 12.82 | 8.02 |
| BF638R1721 | conserved hypothetical protein | 5.43 | 74.43 | 16.70 | 11.68 |
| BF638R1722 | conserved hypothetical protein | 21.28 | 188.38 | 63.00 | 75.48 |
| BF638R1723 | SusC-like | 47.55 | 195.37 | 67.18 | 76.87 |
| BF638R1724 | conserved hypothetical protein | 13.49 | 163.55 | 7.43 | 4.49 |
| BF638R1725 | SusC-like | 16.81 | 152.89 | 5.42 | 2.61 |
| BF638R1726 | SusC-like | 3.38 | 7.18 | 0.44 | 0.25 |
| BF638R1727 | exported protein | 2.52 | 7.96 | 0.71 | 0.52 |
| BF638R1728 | neuraminidase precursor | 9.81 | 37.60 | 12.63 | 12.98 |
| BF638R1729 | beta-N-acetylhexoosaminidase | 19.66 | 128.60 | 15.90 | 15.22 |
| BF638R1730 | sialate-O-acetyltransferase | 11.99 | 75.34 | 4.85 | 7.61 |
| BF638R1731 | sialate O-acetylesterase | 2.60 | 3.43 | 1.19 | 1.28 |
| BF638R1732 | beta-mannosidase | 1.38 | 1.84 | 0.45 | 0.61 |
| BF638R1733 | beta-N-acetylhexosaminidase | 1.63 | 2.12 | 0.42 | 0.48 |
| BF638R1734 | lipoprotein | 1.29 | 2.37 | 0.27 | 0.40 |
| BF638R1735 | beta-N-acetylhexosaminidase | 1.28 | 1.18 | 0.52 | 0.43 |
| BF638R1736 | outer membrane protein | 1.09 | 1.27 | 0.37 | 0.38 |
| BF638R1737 | beta-galactosidase | 0.99 | 1.32 | 0.43 | 0.50 |
| BF638R1738 | SusC-like | 6.81 | 21.16 | 14.41 | 10.74 |
| BF638R1739 | conserved hypothetical protein | 5.40 | 6.36 | 3.37 | 1.96 |
| BF638R1740 | outer membrane protein | 4.59 | 3.95 | 2.57 | 2.02 |
| BF638R1751 | SusC-like | 23.68 | 6.85 | 0.81 | 0.62 |
| BF638R1752 | outer membrane protein | 19.34 | 6.23 | 0.68 | 0.67 |
| BF638R1753 | alpha-1,2-mannosidase precursor | 13.55 | 3.03 | 1.06 | 1.28 |
| BF638R1754 | endonuclease/exonuclease/phosphatase family protein | 4.03 | 1.65 | 0.55 | 0.58 |
| BF638R1755 | Alkaline phosphatase | 2.41 | 0.70 | 0.23 | 0.24 |
| BF638R1756 | conserved hypothetical protein | 2.22 | 1.03 | 0.29 | 0.29 |
| BF638R1757 | exported protein | 2.47 | 0.71 | 0.21 | 0.21 |
| BF638R1758 | exported protein | 1.83 | 0.92 | 0.23 | 0.18 |
| BF638R1928 | SusC-like | 1.01 | 0.85 | 0.26 | 0.29 |
| BF638R1929 | conserved hypothetical protein | 1.01 | 1.10 | 0.32 | 0.36 |
| BF638R1960 | TonB-dependent outer membrane receptor protein | 0.53 | 0.31 | 0.44 | 0.39 |
| BF638R1961 | SusC-like | 0.34 | 0.65 | 0.19 | 0.28 |
| BF638R1962 | conserved hypothetical protein | 0.32 | 0.42 | 0.09 | 0.15 |
| BF638R1963 | conserved hypothetical protein | 0.33 | 0.34 | 0.06 | 0.07 |
| BF638R1964 | hypothetical protein | 0.33 | 0.42 | 0.06 | 0.10 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R1965 | hypothetical protein | 0.37 | 0.54 | 0.09 | 0.14 |
| BF638R1966 | hypothetical protein | 0.43 | 0.59 | 0.14 | 0.17 |
| BF638R1967 | conserved hypothetical protein | 0.42 | 0.46 | 0.10 | 0.14 |
| BF638R1968 | conserved hypothetical protein | 0.46 | 0.51 | 0.07 | 0.09 |
| BF638R1969 | polysialic acid capsule transport protein | 0.52 | 0.82 | 0.16 | 0.21 |
| BF638R2253 | SusC-like | 0.13 | 0.31 | 0.02 | 0.04 |
| BF638R2254 | lipoprotein | 0.15 | 0.28 | 0.04 | 0.07 |
| BF638R2329 | lipoprotein | 0.07 | 0.08 | 0.06 | 0.07 |
| BF638R2330 | SusC-like | 0.07 | 0.05 | 0.03 | 0.03 |
| BF638R2916 | conserved hypothetical protein | 2.63 | 1.04 | 0.73 | 0.95 |
| BF638R2917 | SusC-like | 3.23 | 1.30 | 0.83 | 0.70 |
| BF638R2918 | membrane protein | 3.26 | 0.78 | 1.83 | 1.64 |
| BF638R2920 | ECF sigma factor | 1.90 | 3.73 | 13.53 | 10.60 |
| BF638R2946 | exported protein | 15.80 | 3.53 | 8.07 | 1.32 |
| BF638R2947 | membrane protein | 18.11 | 3.53 | 7.19 | 1.50 |
| BF638R2948 | SusC-like | 34.43 | 8.37 | 12.90 | 2.29 |
| BF638R2949 | conserved hypothetical protein | 15.35 | 1.25 | 12.24 | 3.46 |
| BF638R3029 | glutamine-dependent NAD+ synthetase | 0.64 | 1.25 | 1.39 | 0.97 |
| BF638R3030 | exported protein | 0.39 | 1.59 | 1.85 | 0.87 |
| BF638R3031 | SusC-like | 0.37 | 0.46 | 1.07 | 1.09 |
| BF638R3032 | exported protein | 0.50 | 0.21 | 0.40 | 0.53 |
| BF638R3033 | exported protein | 0.67 | 0.77 | 1.70 | 1.00 |
| BF638R3120 | sulfatase | 1.48 | 0.63 | 0.39 | 0.24 |
| BF638R3121 | exported hydrolase | 1.64 | 0.50 | 0.60 | 0.33 |
| BF638R3122 | lipoprotein | 1.47 | 0.34 | 0.59 | 0.24 |
| BF638R3123 | SusC-like | 1.20 | 0.82 | 2.34 | 0.52 |
| BF638R3124 | anti-sigma factor | 1.81 | 1.40 | 3.64 | 2.52 |
| BF638R3125 | sigma factor | 1.15 | 1.73 | 0.78 | 1.22 |
| BF638R3167 | conserved hypothetical lipoprotein | 8.14 | 7.10 | 2.76 | 4.31 |
| BF638R3168 | lipoprotein | 2.90 | 3.23 | 1.60 | 2.29 |
| BF638R3169 | membrane protein | 4.07 | 3.93 | 1.86 | 3.07 |
| BF638R3170 | SusC-like | 4.23 | 4.61 | 2.72 | 3.57 |
| BF638R3226 | SusC-like | 5.03 | 3.82 | 1.24 | 0.67 |
| BF638R3227 | lipoprotein | 3.44 | 3.90 | 1.43 | 0.81 |
| BF638R3228 | lipoprotein | 3.73 | 5.26 | 3.18 | 2.70 |
| BF638R3290 | exported protein | 7.72 | 1.67 | 0.27 | 0.20 |
| BF638R3291 | SusC-like | 8.76 | 2.17 | 0.29 | 0.22 |
| BF638R3331 | hypothetical protein | 0.90 | 0.56 | 0.27 | 0.32 |
| BF638R3332 | lipoprotein | 1.08 | 0.42 | 0.53 | 0.44 |
| BF638R3333 | exported protein | 3.34 | 2.11 | 1.27 | 0.87 |
| BF638R3334 | SusC-like | 5.06 | 3.38 | 2.48 | 2.34 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R3437 | RNA polymerase sigma factor | 3.60 | 23.33 | 14.85 | 22.79 |
| BF638R3438 | regulatory protein | 33.84 | 223.63 | 597.79 | 532.11 |
| BF638R3439 | SusC-like | 559.54 | 2146.72 | 933.46 | 1161.15 |
| BF638R3440 | conserved hypothetical protein | 505.26 | 1907.71 | 700.53 | 918.12 |
| BF638R3441 | exported protein | 334.26 | 1494.32 | 443.63 | 631.07 |
| BF638R3442 | lipoprotein | 233.61 | 765.28 | 367.94 | 438.35 |
| BF638R3443 | exported protein | 90.69 | 1227.85 | 235.37 | 363.24 |
| BF638R3466 | SusC-like | 0.50 | 1.92 | 0.96 | 0.68 |
| BF638R3467 | conserved hypothetical protein | 0.50 | 1.90 | 0.28 | 0.38 |
| BF638R3468 | lipoprotein | 0.51 | 0.82 | 0.08 | 0.12 |
| BF638R3469 | peptidase | 0.51 | 0.88 | 0.10 | 0.10 |
| BF638R3470 | hypothetical protein | 0.41 | 0.70 | 0.12 | 0.15 |
| BF638R3593 | exported phosphatase | 1.48 | 1.37 | 1.03 | 1.13 |
| BF638R3594 | exported protein | 1.69 | 0.99 | 0.49 | 0.38 |
| BF638R3595 | exported phosphoesterase protein | 1.81 | 0.89 | 0.72 | 0.48 |
| BF638R3596 | conserved hypothetical protein | 6.21 | 4.97 | 3.34 | 3.16 |
| BF638R3597 | SusC-like | 7.16 | 6.41 | 3.45 | 3.47 |
| BF638R3598 | anti sigma factor | 2.46 | 1.60 | 17.35 | 14.63 |
| BF638R3599 | RNA polymerase sigma factor | 1.76 | 1.46 | 0.82 | 0.69 |
| BF638R3600 | lipoprotein | 8.64 | 0.67 | 0.20 | 0.15 |
| BF638R3601 | SusC-like | 9.98 | 0.49 | 0.14 | 0.11 |
| BF638R3602 | conserved hypothetical protein | 6.63 | 0.82 | 1.21 | 1.83 |
| BF638R3603 | membrane protein | 10.52 | 0.87 | 1.31 | 1.01 |
| BF638R3604 | SusC-like | 17.96 | 1.22 | 1.82 | 1.44 |
| BF638R3605 | regulatory protein (possible anti-sigma factor) | 5.53 | 0.74 | 10.18 | 7.99 |
| BF638R3606 | RNA polymerase sigma factor | 0.87 | 0.49 | 0.83 | 1.81 |
| BF638R3783 | transport related, membrane protein | 2.67 | 1.65 | 1.55 | 1.22 |
| BF638R3784 | sialidase | 3.29 | 1.67 | 0.67 | 0.60 |
| BF638R3785 | hypothetical protein | 3.64 | 2.55 | 2.00 | 1.42 |
| BF638R3786 | lipoprotein | 2.56 | 1.55 | 1.31 | 1.32 |
| BF638R3787 | SusC-like | 2.70 | 1.54 | 1.41 | 1.28 |
| BF638R3788 | N-acetylneuraminate lyase (sialic acid lyase) | 1.90 | 2.57 | 1.51 | 1.69 |
| BF638R3789 | GntR-family regulatory protein | 1.29 | 0.38 | 1.12 | 1.22 |
| BF638R3796 | exported sulfatase | 1.59 | 0.76 | 1.08 | 0.84 |
| BF638R3797 | exported uslfatase | 2.21 | 1.11 | 1.20 | 0.82 |
| BF638R3798 | conserved hypothetical protein | 2.31 | 1.70 | 2.66 | 2.92 |
| BF638R3799 | SusC-like | 2.99 | 2.90 | 3.90 | 4.01 |
| BF638R3800 | membrane protein | 1.37 | 0.66 | 15.37 | 9.87 |
| BF638R3801 | RNA polymerase sigma factor | 0.59 | 2.08 | 3.20 | 4.32 |
| BF638R3819 | RNA polymerase sigma factor | 1.95 | 2.42 | 3.01 | 3.69 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R3820 | membrane protein | 4.68 | 4.53 | 12.17 | 6.41 |
| BF638R3821 | SusC-like | 2.13 | 2.18 | 2.17 | 1.33 |
| BF638R3822 | liporotein | 1.80 | 0.94 | 0.60 | 0.35 |
| BF638R3823 | lipoprotein | 0.95 | 1.39 | 1.07 | 0.74 |
| BF638R3824 | exported protein | 0.08 | 0.12 | 0.07 | 0.06 |
| BF638R4102 | SusC-like | 4.38 | 4.27 | 0.97 | 0.69 |
| BF638R4103 | outer membrane protein | 3.60 | 3.57 | 0.51 | 0.37 |
| BF638R4104 | endo-beta-galactosidase | 3.00 | 3.34 | 0.74 | 0.66 |
| BF638R4105 | beta-hexosaminidase precursor | 2.15 | 1.40 | 0.18 | 0.17 |
| BF638R4106 | beta-glucanase | 1.99 | 1.88 | 0.49 | 0.44 |
| BF638R4107 | beta-galactosidase | 1.92 | 2.37 | 1.23 | 0.59 |
| BF638R4108 | SusC-like | 2.63 | 2.12 | 0.42 | 0.36 |
| BF638R4109 | outer membrane protein | 2.44 | 1.50 | 0.46 | 0.42 |
| BF638R4142 | glycosyl hydrolase | 2.02 | 1.68 | 0.52 | 0.52 |
| BF638R4143 | conserved hypothetical protein | 1.71 | 0.96 | 0.30 | 0.28 |
| BF638R4144 | outer membrane protein | 2.27 | 0.84 | 0.48 | 0.42 |
| BF638R4145 | SusC-like | 2.75 | 1.19 | 0.37 | 0.23 |
| BF638R4146 | anti-sigma factor | 1.58 | 1.07 | 1.26 | 0.75 |
| BF638R4147 | ECF-type RNA polymerase sigma factor | 0.80 | 0.72 | 0.33 | 0.36 |
| BF638R4192 | Transcriptional regulator | 1.02 | 3.06 | 3.10 | 2.59 |
| BF638R4193 | SusC-like | 42.41 | 4.08 | 3.66 | 4.58 |
| BF638R4194 | lipoprotein | 32.61 | 3.23 | 1.83 | 2.95 |
| BF638R4207 | RNA polymerase ECF-type sigma factor | 2.17 | 7.16 | 12.22 | 11.38 |
| BF638R4208 | anti-sigma factor | 8.54 | 2.29 | 5.73 | 2.32 |
| BF638R4209 | SusC-like | 57.11 | 32.97 | 8.14 | 3.80 |
| BF638R4210 | conserved hypothetical protein | 24.09 | 9.37 | 2.59 | 2.31 |
| BF638R4247 | conserved hypothetical protein | 1.41 | 2.26 | 2.58 | 2.10 |
| BF638R4248 | possible xylosidase/arabinosidase | 1.71 | 2.17 | 2.75 | 2.22 |
| BF638R4249 | conserved hypothetical exported protein | 2.03 | 1.70 | 4.03 | 4.00 |
| BF638R4250 | outer membrane protein | 1.59 | 1.20 | 3.79 | 2.11 |
| BF638R4251 | SusC-like | 2.29 | 1.88 | 4.90 | 2.51 |
| BF638R4258 | hypothetical protein | 1.46 | 0.52 | 0.20 | 0.19 |
| BF638R4259 | outer membrane protein | 2.76 | 1.06 | 0.73 | 0.46 |
| BF638R4260 | SusC-like | 3.79 | 1.46 | 0.83 | 0.70 |
| BF638R4261 | two-component system sensor kinase/response regulator fusion protein | 0.89 | 5.25 | 3.59 | 2.53 |
| BF638R4328 | SusC-like | 1.91 | 1.69 | 0.17 | 0.19 |
| BF638R4329 | conserved hypothetical protein | 1.69 | 1.29 | 0.18 | 0.19 |
| BF638R4330 | SusC-like | 8.45 | 1.79 | 0.35 | 0.17 |
| BF638R4331 | outer membrane protein | 6.52 | 1.39 | 0.40 | 0.17 |
| BF638R4332 | hypothetical protein | 5.68 | 1.11 | 1.07 | 1.36 |
| BF638R4338 | SusC-like | 2.02 | 1.03 | 0.12 | 0.14 |

| | | | | | |
|---|---|---|---|---|---|
| BF638R4339 | conserved hypothetical protein | 2.18 | 1.09 | 0.12 | 0.14 |
| BF638R4348 | conserved hypothetical protein | 0.93 | 0.86 | 1.30 | 1.18 |
| BF638R4349 | arylsulfatase | 1.14 | 0.52 | 1.32 | 1.63 |
| BF638R4350 | conserved hypothetical protein; lipoprotein | 1.24 | 0.53 | 1.10 | 1.81 |
| BF638R4352 | SusC-like | 1.24 | 0.42 | 1.10 | 1.61 |
| BF638R4353 | anti-sigma factor | 1.21 | 0.28 | 12.51 | 9.78 |
| BF638R4354 | ECF-type RNA polymerase | 0.95 | 1.05 | 2.02 | 3.62 |
| BF638R4443 | SusC-like | 2.59 | 2.06 | 0.41 | 0.36 |
| BF638R4444 | lipoprotein | 0.65 | 0.46 | 0.15 | 0.21 |
| BF638R4478 | hypothetical protein | 2.23 | 2.84 | 1.45 | 2.03 |
| BF638R4479 | outer membrane protein | 4.48 | 4.32 | 0.95 | 0.72 |
| BF638R4480 | SusC-like | 4.99 | 5.82 | 1.57 | 1.04 |
| BF638R4481 | hypothetical protein | 2.09 | 1.35 | 1.07 | 0.70 |
| BF638R4482 | transmembrane sugar transporter | 2.19 | 0.83 | 2.38 | 1.95 |
| BF638R4442 | outer membrane protein | 1.00 | 1.00 | 1.00 | 1.00 |

# APPENDIX 3: PRIMARY TRANSCRIPTOME ANALYSIS

| *TSS* | *Strand* | *Transcript Start locus* | *Name* | *Putative sRNA* | *Sequence -50 nt upstream + TSS (51nt)* | *TAxxT TTG* |
|---|---|---|---|---|---|---|
| 44 | + | bf638r_0001 | SpoU rRNA methylase | 0 | TTATCAACACCTATGATTGCGTTAACA AGAAAAGAATTACTTTTGCAACAT | 1 |
| 716 | + | bf638r_0002 | omp | 0 | AAATAGAAATAGTAATATTATTCTATTT ATATGGTTTTTATGATATTGCAAA | 0 |
| 3187 | + | bf638r_0004 | sugar transferase | 0 | CCGTTATCTTGGCTTTTGCGCGAATAA AAAATTTCCTTTACTTTAATAACTT | 0 |
| 14658 | - | bf638r_0012 | hypo | 0 | TTTGCAAAAATAAAGGATTAAATTTAT ATTAATGGTTTTCCTTGCAGCATTT | 1 |
| 14734 | + | bf638r_0012 | hypo | 0 | TATTAATGGTTTTCCTTGCAGCATTAAC CTATTTAAATTATCTTCGACAAT | 0 |
| 16811 | + | bf638r_0014 | SAM superfamily | 0 | AAACGGATGTACAAATGCAGCTGCGT CTAAAGTTTTAGTAATTTTGTAGAA | 1 |
| 22995 | - | bf638r_0018 | iron transport receptor | 0 | GCCGCAAAGGTATGTGTACCTGAAAG AGTCTACAATACCTATAAAAAGGTA | 1 |
| 24900 | - | bf638r_0019 | sulfate permease | 0 | TTGCAAAATTACATAAAATAAATGGCA ATATTATGTTTTTGAACATCTTAT | 1 |
| 25042 | + | bf638r_0020 | rubreythrin | 0 | TAATGGATAAAGTTCCGTGTTTTTTGA ACTAAATCCTTATATTTGCGTTAT | 1 |
| 27348 | - | bf638r_0021 | nadB | 0 | ACCGCAAAGCTACAAAATAAGTTTTTT CGTTGTGCTTTTCGTTGCCATTAA | 1 |
| 27425 | + | bf638r_0022 | mem prt | 0 | TCGTTGTGCTTTTCGTTGCCATTAATTC ATAAATTTACTAACTTTCCGCCAA | 1 |
| 32662 | - | bf638r_0026 | hypo | 0 | CCGCCAAAGTTATACATTTAATCTGAT TGATACCAATATTTTGGCATGAAA | 1 |
| 32790 | - | | | 1 | GCCCAAATGTGTGAAATAAAATTAATA GTGCCAAATCCTTACGCCATAAAT | 1 |
| 32951 | + | bf638r_0027 | SAM/TRAM methylase | 0 | TTGGTTTCATTTTACAACTTTTTGAGGA TGAAGCAGTATCTTTGCGGACGC | 1 |
| 34441 | + | bf638r_0028 | alpha fucosidase | 0 | TATGACTGCACCTTGTTTGTTATATGG CTACATTGTTTTATCTTTGTGCGT | 1 |
| 41129 | + | bf638r_0033 | mem prt | 0 | AATCTATGAGGGAACTTTGAGATAATA TTGTAAAGTATTACATTTGTTTGC | 1 |
| 42445 | + | bf638r_0034 | hypo | 0 | TAATGCAAAAAATGGTGTTATATTGAT CAATAGCTTTTAACTTTGCTGTGA | 1 |
| 44349 | + | bf638r_0037 | hypo | 0 | AATAAAATTGATAGCAGTTTATCCAAA ATCTCTTTTTACCTTTGCTGCCAG | 1 |
| 44748 | - | | | 1 | GGCACGAAGATAGCGAAAAATTTATT GGAATGAGTAAATGTAGCATTTACT | 1 |
| 46745 | + | bf638r_0040 | mem prt | 0 | ATTTTTTCTATATTTTAGCTTCTTGTTAG GAAATAATATTATTTTTGTCAG | 1 |
| 47955 | + | bf638r_0041 | hypo | 0 | AATAAAATTGATAGCAGTTTATCCAAA ATCTCTTTTTACCTTTGCTGCCAG | 1 |
| 48260 | + | | | 0 | AGTGAGGACTCTTTTATATAAATCTTA TCACATTATAATCATTCATTGCAT | 0 |
| 48357 | - | | | 1 | GGCACGAAGATAGCGAAAAATTAATT AGAAAGAGTAAATGTAGCATTTAC | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 48461 | + | bf638r_0042 | hypo | 0 | TTCTACGGAGGAGGTGGCTACCCGAG AGAGAGACACGGCAAATAAGAAAATA | 0 |
| 49845 | + | bf638r_0044 | hypo-pseudo | 0 | AGATGTGGATTGTAATATAGAATCTGT CGATTTGGCAGAGTTGTGGAAGCA | 0 |
| 50914 | + | | | 1 | CTCTGCCTGATGACAAAGGAAACGGA TGGTGGAGAAGAGCGTTATCATTAT | 0 |
| 52178 | + | bf638r_0049 | hypo | 1 | CTGTGATGCTCCGTGGTGAAAAGAGT TATATGAATATGAAAATTATGAAAT | 0 |
| 54233 | + | bf638r_0050 | hypo | 0 | CTCCTTAAGGAAATGCCTATTCGTCAC TTTTTATTTGTTATTTTGCACCAA | 1 |
| 55578 | - | | | 0 | ATTGCGAAAATAGAGAAATTGATTAA AAGAAATACCAATATACAGGATGTG | 1 |
| 59623 | - | | | 1 | GAAGCAAAAGTAGCAATAAATTCAAA ATATTATCTACCTTTGTCCCTGAAA | 1 |
| 59671 | + | bf638r_0058 | metK | 0 | ACGAAGCAAAAGTAGCAATAAATTCA AAATATTATCTACCTTTGTCCCTGA | 1 |
| 61222 | + | bf638r_0059 | lysine decarboxylase | 0 | AACAAGGCAGCCTTCTTCTTTTCTCCTT GAAATTTGGTACTTTTGCAGCGT | 1 |
| 61284 | - | | | 1 | AAGTACAAAAATAGACCAGACTTATTT TGATGCAGCCATGAAATTAGGGCA | 1 |
| 62588 | + | bf638r_0061 | uroporphyrinogen III synthetase | 0 | TCCTGTAATATTTATGGGGTTTTCCTTT TAATTTTATACTTTTGTGCCCTT | 1 |
| 64787 | + | bf638r_0063 | tyrosyl-tRNA synthetase | 0 | TTTTAGGCAATAAGAGTTGCTTGTTAC GATAAAGAGTCGTACTTTTGCCCTAA | 1 |
| 69533 | - | bf638r_0066 | peptidase | 0 | AAGTACAAAGGTAGCAATTCGCTTCG GCTAACGAGACTTATTATACAAAAA | 1 |
| 69635 | + | tRNA-gln | tRNA | 0 | GTCTAAAAATATTTGGAAAATAGCTTG CATGGCATTATCTTTGCATCGCTT | 1 |
| 70027 | + | bf638r_0067 | choline sulfatase | 0 | ATTATCTTAAATTTCTCTTCTTTCTCTTT CTTTTTACCTATTTTTGCCTAT | 1 |
| 70049 | - | | | 1 | AATCAATGAAACGTCGCGATTTTCTGA AATGTTCACTTGCCGTGGGAGCAG | 0 |
| 71701 | + | bf638r_0068 | 4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase | 0 | GGAGTATCATTTATTCCCAAATCCCGT TTTCTTTTCTTATTTTTGCGTCGT | 1 |
| 76163 | + | | | 1 | CGGATGTTGTCGCCTACCATGAAAAA GGTTGCCTTGATGTTGTGTTTGTCC | 1 |
| 79721 | - | bf638r_0073 | hypo | 0 | GCGCAAAGATAATGATTTACGTGAAG ATAGCCGTAGAAAGTGTAGAGTTTT | 1 |
| 79895 | - | | | 1 | GCAACAAAGATAAGCCGCTTTTTTCAT CTGTGCAAATATTTCTACATTGTT | 1 |
| 80013 | + | | | 1 | ATCTCTTCCTTTGGCCTTCAGTGAGCA GAGCCAAAACTGTGTTTCAAAAAA | 0 |
| 90338 | - | | | 1 | ATTTGTGTTCGGAACGGTTTAGTCCTC CTGGGGCTTTTTGATTTTGGGCTC | 0 |
| 92697 | - | bf638r_0086 | serine acetyltransferase | 0 | GGCTACAAAGATATGAATAATTGCGC ATAAAAAGTAGTACGTTTGTTATCG | 1 |
| 92778 | + | bf638r_0087 | TCS-sensor | 0 | AAGTAGTACGTTTGTTATCGAATTTAA AAAAAGGATGTATTTTTGTTGGGT | 1 |
| 95605 | + | bf638r_0089 | hypo | 0 | CGCGGATTCTTGAAAGAACTACAAGA GCGCTTTTGTATAATTTGAAGTTTTT | 1 |
| 95716 | + | bf638r_0089 | hypo | 0 | CAGATAAAAAAGGGTTGCTATGCTTTT ACTTTGATTGTATCTTTGTGAGGTA | 1 |
| 96426 | - | bf638r_0090 | exp prt | 0 | GGTGGCAAAGTTACATATTAAACGGA TATAAATGCATAAAATCGAAAGAAT | 1 |
| 96484 | + | | | 1 | AGTTACATATTAAACGGATATAAATGC ATAAAATCGAAAGAATTTTATCAT | 0 |
| 99476 | - | bf638r_0091 | polA | 0 | AAAAAAATACCGTATTAACCGTTTTAT TATTACTTTTGTGCTCAAATAGTG | 0 |
| 99520 | + | bf638r_0092 | octaprenyl- | 0 | AAGTACAAAAAAATACCGTATTAACCG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | diphosphate synthase | | TTTTATTATTACTTTTGTGCTCAA | |
| 103330 | - | bf638r_0098 | D-tyrosyl-tRNA acylase | 0 | ATTGCAAAGATAGCTATTTTTAATGAT TATCGAGACCCTCTTTTATGATT | 1 |
| 107716 | - | bf638r_0101 | gidA | 0 | CGCGCAAAGTTACGAAAAAGTTGCCA GTTGCGGGTTTTATGCAATAAGTTT | 1 |
| 108414 | + | bf638r_0102 | lipoprotein | 0 | TCCTTTGAAAAATTTGGAAGTGGCTGA AAAGAATCGTAAGTTTGTCATCAT | 1 |
| 110223 | + | bf638r_0103 | exp prt | 0 | ATTGCTTTATATCTGTTAATCTGCTATT TACAATGTTTAAGATGTTATCTC | 0 |
| 113100 | - | bf638r_0107 | hypo | 0 | AAGACAAAGATAAGAATTTTAATTGA GGATAAGCGCAAATTGCAGTAATAA | 1 |
| 113416 | + | bf638r_0109 | hypo | 0 | TACTATCTTTTGAGACCGTAACTATTTT GTTTTTTATATATTTGCAAATGA | 1 |
| 115236 | + | bf638r_0111 | hypo | 0 | CGGTACGCAAAAACAGCATTACTCGCT ATCAAAAATCAAGGCACGGAACAGG | 0 |
| 117337 | + | | | 1 | AGTCCGGAAAGACTGTTTGATACAACT TTTATTATGCGTAAGTTTGTGTAT | 1 |
| 118967 | - | bf638r_0113 | spore maturation prt | 0 | ACAAAATTAAGAATATGTTTCCACTCA TTGCGGCTTTTTTGACTATTTTTG | 1 |
| 119022 | + | bf638r_0114 | hypo | 0 | ATTAAGAATATGTTTCCACTCATTGCG GCTTTTTTGACTATTTTTGTCCGT | 1 |
| 119937 | + | | | 1 | GTTCCGGTGCAGCGAAAGAGATCAGC TGTCCGGAGTCAGACAATAAAGAAC | 0 |
| 120062 | + | bf638r_0115 | ruvB | 0 | CTCTGTGAAACTCTGCCTTACTATGTA GCGAAATCTATATCTTTGCTACCT | 1 |
| 124037 | + | bf638r_0118 | hypo | 0 | ATTGTTGAAGAGTGACGGTAAGCGTG TCGCCAGAGAAAAGAAGATGAATAA | 0 |
| 127175 | - | bf638r_0119 | transmembrane prt | 0 | CACGCAAAGGTAGCACTTTTAAAGACA ATAATAAAGAGGGGAAGGATTTTT | 1 |
| 129826 | - | bf638r_0121 | RNA helicase | 0 | GGGTGCAAAGATACGCATTAATTGTG AATAATGAATGATAACAATAAATAA | 1 |
| 130975 | - | bf638r_0122 | hypo | 0 | GATTGCAAAGATAGGAATTCTCGTTAC ATGGTATAATTACTTTGCGATAAT | 1 |
| 131662 | + | | | 0 | CGCTCTTTTTTCTTTTGCTATGTCCAGC AGTTCCAATATCTTATCACGCAA | 0 |
| 133199 | - | bf638r_0125 | lysine decarboxylase | 0 | CTTACAAATATCTGCAATTTTCTATAAA CGACAATAAGTTTTATTGTATTT | 1 |
| 133258 | + | bf638r_0126 | peptidase | 0 | ATCTGCAATTTTCTATAAACGACAATA AGTTTTATTGTATTTTTGTGTCCT | 1 |
| 135073 | - | bf638r_0127 | exp prt | 0 | GGCGGCAAATGTATGGATTAATATGG CTCTTTGGTCCGGGTGACTCTGCTT | 1 |
| 135279 | + | bf638r_0128 | acetyl transferase | 0 | AATGTTTAAAGACTGCTGTCTTCCGGC TTATTCAGCTTATCTTTGCTCGAA | 1 |
| 138145 | - | bf638r_0130 | uxuB; D-mannonate oxidoreductase | 0 | GCGCAAATATAGCTATAAAGTTGATG GGATGGTTTAAAAAAAGAGGGAAAT | 1 |
| 142122 | - | bf638r_0134 | dihydrodipicolinate reductase | 0 | AGTCGCAAAGATAAGATTTTTATTACA TTTGTCGAGCACATTAACGTCTTG | 1 |
| 142158 | + | bf638r_0136 | hypo | 0 | ATATCCAAAATATTAGTCGCAAAGATA AGATTTTTATTACATTTGTCGAGC | 1 |
| 146534 | - | bf638r_0139 | hypo | 0 | GCTGCAAAGGTAAGGAAATCATCTGA CTGACCATAAAAAAGGAGATCCGA | 1 |
| 146666 | + | bf638r_0140 | hypo | 0 | GCATCGGATGGAGAACGGAATACTGA AATAAAGTATATATTTTTGCAAGAT | 1 |
| 154115 | - | | | 1 | GCGACAAAAGTACTCTTTATTTTCTATA CCGTTGTTTTTTAATATCTGTTT | 1 |
| 157550 | + | bf638r_0146 | hypo | 0 | TTAGTGTAAAGTATAGCATTAATTAAT GGATAAGGCTTATTTTTGCAGCAT | 1 |
| 162269 | - | bf638r_0149 | cysS (tRNA synthetase) | 0 | CACAAAGTTAGAATAAATAATTTGAAT TGGGACGTGTTCTGCCGGATTATT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 162692 | - | | | 1 | AGTATGATGAATGATAATATACTTATA TATAACGCGTTAATATCATGGAG | 0 |
| 163635 | - | bf638r_0150 | haloacid dehalogenase | 0 | GGGTGCAAAGGTAGTGGTTTTGGGTA TATCACCCAAACAGGATGCGAGTAA | 1 |
| 163813 | + | bf638r_0151 | transmembrane prt | 0 | TTTTTCTTTATTTTCTTTAAAATCAGAA ATAAATACACGATATTTGTATAA | 1 |
| 165319 | + | bf638r_0152 | mem prt | 0 | TAGTAAGGAACGTTTTCTTAATAGATA GAAAAAATCGTCTATATTTGTTAT | 1 |
| 165666 | - | | | 0 | GTTACAAAGGTAATGAATTTAAGTAAG AATCAGTAGGTGCATCATTTACTT | 1 |
| 170318 | + | bf638r_0158 | TCS-regulator | 0 | TTTCTGTCCGCATGACATTCAGGATAG GAAAAAAGTCTTAACTTTGTGCCGT | 1 |
| 172531 | + | bf638r_0160 | mgtA | 0 | TAATTTGTTTCTAACTACCTTCTCATAA GGTGGTTTGTACCTTTGTGGCGT | 1 |
| 176570 | + | bf638r_0162 | lipoprotein | 0 | TAATTTGTTTCTAACTACCTTCTCATAA GGTGGTTTGTACCTTTGTGGCGT | 1 |
| 180505 | - | bf638r_0163 | SNF family helicase | 0 | GGATCAAAGATACGGGATTTTATGTA ATTTTGCGGTATGATGGAAAAAAAC | 1 |
| 183232 | - | bf638r_0165 | putative plant auxin-regulated protein | 0 | CGTGCAAAAATAGCCTTAATTATCTTTT TTTTCTATCTTTACGACGGAAAT | 1 |
| 183278 | + | bf638r_0166 | pfkA2; 6-phosphofructokinase 2 | 0 | TAGTCGTGCAAAAATAGCCTTAATTAT CTTTTTTTTCTATCTTTACGACGG | 1 |
| 186844 | - | bf638r_0169 | acpP, acyl carrier protein | 0 | GGTGCAAAGGAATAAATATTTATCGTC CTGCGCAAATTTTTGATTAAATAA | 1 |
| 186848 | + | bf638r_0170 | purN, phosphoribosylglycinamide formyltransferase | 0 | ACTTTAAGTTTTTAATTAATAATTAGTT TTATTTCTTTAAATTTGCGGTGC | 1 |
| 188620 | - | bf638r_0171 | pdxB, erythronate-4-phosphate dehydrogenase | 0 | CATGCAAAGGTAGGAAATAAAAAGCA GGAAAGATTGGCTATATGGAAAATA | 1 |
| 188689 | + | bf638r_0172 | hypo | 0 | AAAAGCAGGAAAGATTGGCTATATGG AAAATATTGCATAGATTTGCCCGGT | 1 |
| 192995 | - | bf638r_0176 | glycosyltransferase | 0 | GATACAAATGTAATACTTATTTTGAAA TAAAGGACTTTTATATTGCTCTTA | 1 |
| 194171 | - | bf638r_0177 | glycosyltransferase | 0 | GTGCAAAGATAGAACGTTTATATTAGA ATGTTACTAAAAAATGGAGAAAAG | 1 |
| 194521 | + | b638r_0178 | transport ATP-binding protein | 0 | TATGCGATGATTTATAAATTCACTAAA GATTTTCTGTTATCTTTGTGCCCT | 1 |
| 198264 | - | bf638r_0180 | rnh1, ribonuclease H1 | 0 | GGGTGGTAAAGGTACGAATATATTCC ATTACCGAACAACTTTCGTTTG | 1 |
| 198324 | + | bf638r_0181 | hypo | 0 | GGTACGAATATATTCCATTACCGAACA ACTTTCGTTTGTCAGTTGTTTTAA | 0 |
| 199787 | + | bf638r_0184 | shikimate kinase | 0 | TACCATGTGGGTCATTCGGTGTTCTCT CAATAAATATCTATCTTTGTGCGAA | 1 |
| 205871 | + | bf638r_0188 | arginine decarboxylase | 0 | TAATGTCAATTTAATTGGCAACCTGTA GATTAATACGTATTTTTGCCGTCC | 1 |
| 208769 | + | bf638r_0190 | ecf sigma factor | 0 | AAATAGGAGTACTTCTGTAACTTTTCA CCAGCTCATCCGTCAACTTAATAG | 0 |
| 212007 | - | bf638r_0194 | lipoprotein | 0 | ATTGATTTAGACGTAAGTCGGCTATGA AAAGTTGCATCTATCTTTCTTTTT | 1 |
| 212088 | + | bf638r_0195 | hypo | 0 | TTGCATCTATCTTTCTTTTTTATGCAAA ATAATCCGTTTATTTGTATTTAT | 1 |
| 213391 | - | bf638r_0196 | transmembrane prt | 0 | GTGATAAAGATAGTGATTTGTTTTATT CATTACAAGAATAACGATACCTTT | 1 |
| 213447 | + | bf638r_0197 | transmembrane prt | 0 | AAAGATAGTGATTTGTTTTATTCATTAC AAGAATAACGATACCTTTGCTGAA | 1 |
| 214664 | + | bf638r_0198 | exp prt | 0 | CCTTCTTGTTGTTTTTTTATTCCTTTTCT TTTTCCATAGATTTGTAACCGT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 225803 | - | | | 1 | AGTGGCAAAATTAGCAAATTACGGGG ACTTACCTTCGGGATTTACCCAGTT | 1 |
| 229623 | + | bf638r_0211 | tetR regulator | 0 | ACCATTTTTAGGGATTGATTATGTAAA GAGATATTGCGAACTTTGCAAACG | 1 |
| 240103 | + | bf638r_0219 | peptidase | 0 | TTAGGAGAAAATCTTA1GATGTTTTAA AATAAATTCATTTATTTTGCATTAT | 1 |
| 243575 | - | bf638r_0223 | exp prt | 0 | TATAAAGATACTTTTGTTTTAAAATTTC TGCAATACCTAAAAAAGGGTATTT | 1 |
| 245937 | - | bf638r_0227 | ecf sigma factor | 0 | CCAAATATACGAAAAGTTTATAATCCT TCCAAGAAACCGCAGTGCGTTGTG | 1 |
| 246591 | - | bf638r_0228 | mem prt | 0 | GTGTTCAAAAATACGATTTCCCACTCT CTTAATCAAATAAAAGGTCTACCC | 1 |
| 265519 | + | | | 1 | TTTCGGATAGCTCTGTATCTTTTTCGCA ATAAATCAAGTATATTTGCGTTAT | 1 |
| 271932 | + | bf638r_0200 | hypo | 0 | AATTTGGAGAAAATTCTTTTTTTTCACG ATATACTCTTATCTTTGCAACAG | 1 |
| 284377 | + | bf638r_0273 | hypo | 0 | TTTTCTCTAATATTTTGTAGTCGAACAG ATTATTTATATCTTTGCACCCGA | 1 |
| 295744 | + | bf638r_0282 | hypo | 0 | AATGAAGGATTCCTGCAACCTTTTTCTT TCTTTATCGTCTAATATTAATAT | 0 |
| 303118 | + | | | 1 | CCTCCCTGCCTTATGACAAGGAATTGC TTTCTTTTTCATACATTTACATTGA | 1 |
| 305077 | - | bf638r_0287 | ecf sigma factor | 0 | GCTGCAAAAATATAAAATATAAACTTC TTTTCTTAATATTTATTACATTGT | 1 |
| 307474 | - | bf638r_0288 | aga, alpha galactosidase | 0 | GTGTGCAAAGTAAGATAAACTTCATTA TATTCACAAATGAAACGGGAAGAA | 1 |
| 307666 | + | bf638r_0289 | lipoprotein | 0 | TCATTCCGGTACTTTGTTGTTACATAAA TAAAAATATGTAATTTTGAGGAT | 1 |
| 308332 | + | bf638r_0290 | inorganic pyrophosphatase | 0 | AAAAGTCGAAAATATTTGGTTATCTCG TTTAGATTCAATACTTTTAGGCTT | 1 |
| 310765 | + | bf638r_0291 | rnh2, ribonuclease H2 | 0 | TTCTTTTTCATCTTTTGTTCATTTCCGTT TAATTGATTACATTTGCCGTGT | 1 |
| 312450 | - | bf638r_0292 | cationic transporter | 0 | AGCGCAAAAATAAGGAAATCATTTCA AATCTGCGGTATTTTTTCTTCCTTT | 1 |
| 312573 | + | bf638r_0293 | gpmI, phosphoglycerate mutase | 0 | AACATAGATTAAACGTGTTCGTAATTA ACTTTTTTGTACCTTTGCACCCGA | 1 |
| 314252 | + | bf638r_0294 | TCS sensor | 0 | TTTATCTAAATGCTTGTTGCTTCATATT TAATGAGTTATCTTTGCTAACCT | 1 |
| 317437 | + | | | 1 | CCCGGCGAAAACGATCTGTCGAACCAT ATATTTCTGTAACTTTGTCTTTAA | 1 |
| 323879 | - | | tRNA glu | 0 | GAGTGCAAAGGTAGGGCTTTTTTTGA AACTACCAAACTATTCGCGGAAAAAT | 1 |
| 324087 | + | bf638r_0300 | recO | 0 | GATAGTAGTATTCGATTGTCTATTTCCT CTTTTTTTGTAGCTTTGTTACCG | 1 |
| 341181 | - | bf638r_0314 | mraZ | 0 | GACCGTAAAAGTAGGTATTTCCCCACA ATTCCCCACAATTTTCCACCAAAA | 1 |
| 341951 | - | bf638r_0316 | hypo | 0 | AGACAAAGGTATGTGAAATAATGGAG ATATACAAGGAAAATAAGCTGGAAA | 1 |
| 342168 | + | bf638r_0317 | exp prt | 0 | ATGCAGGAAGTAATGCTATTTATTAAA AAATATGTTATATTTACGATTTAG | 1 |
| 346772 | - | bf638r_0322 | sigE, ecf | 0 | GATGCAAGTTTACGGCATTTTTTAGAC AGTGACGGGGGAAAATTAAGCTTT | 1 |
| 346916 | + | bf638r_0323 | mem prt | 0 | TGATAGTTTTTGGGTGTTTTTGGGACG CTGGGTTTTCTATCTTTGCGTCGG | 1 |
| 350251 | + | bf638r_0326 | hypo | 0 | TTCTACCGATTTATTGCGATTAATAAA AAAGAACAGTTAATTTTGCGCAAT | 1 |
| 353464 | - | bf638r_0328 | dgtdeoxyguanosine triphosphate triphosphohydrolase | 0 | AACTGCAAAAGTATGAATTTTAGGCTC AAATTGCGTATTTTCGCCCGTTAA | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 353514 | + | bf638r_0329 | dut, deoxyuridine 5'-triphosphate nucleotidohydrolase | 0 | AACTGCAAAAGTATGAATTTTAGGCTCAAATTGCGTATTTTCGCCCGTTAA | 1 |
| 362986 | - | bf638r_0334 | exported hydrolase | 0 | ATTGCAAATTAAATACGTTTTGGGACGAACGGAGACACGTCCGGCTCAAGT | 1 |
| 366957 | - | bf638r_0336 | glycosylhydrolase | 0 | GATTTGCAAAATAATAAAATAGCAGATAACAGTTGAAAGCCTCCCTTTCAT | 1 |
| 373369 | + | bf638r_0340 | lipoprotein | 0 | TGAGCAGTATTTATCCCTTTATTCAACGAAAATGTTTTTTCTTTGTAGCAT | 1 |
| 385493 | + | bf638r_0345 | omp | 0 | TGAACGAAAATTATCGTCTTATTAGTTAAAAATAACTTTCCTTTGCAATGT | 1 |
| 394047 | + | bf638r_0352 | glycosyl hydrolase | 0 | TGAATAATAATTATCCTCCTTTCTCCAATATTTCCCCAATCTTTGTAACAT | 1 |
| 399908 | + | bf638r_0355 | exported alpha-1,2-mannosidase | 0 | ATCAAGATGCTTGAGAAGAAGTTGAAAAAATAAGATTCTGTCCACAGGTTA | 0 |
| 403846 | + | bf638r_0357 | galM, aldose 1 epimerase | 0 | TACAGTAAGAATATAACGAAACAGAAATGAATGAGATTCACCACAGAGT | 0 |
| 405168 | + | bf638r_0358 | mem prt | 0 | CAAATTTTCGGTAGATAACGGAAAATAACACAAGCGTTTTTTTAGTTGT | 0 |
| 413636 | + | bf638r_0367 | hypo | 0 | TATTGTTATTTTCTTTTAAGGATATGTTTGTAGTTATTATCTTTGTCAGAC | 1 |
| 416778 | - | bf638r_0370 | cAMP binding prt | 0 | GTTTACAAACTTACATAATTTTATCAATTCCTAAAAATTATCGATCTTAAATCCGCCCCTT | 1 |
| 417039 | + | bf638r_0371 | exp prt | 0 | TTAACGCCTCTATGTGCTTATATTCAGACCTTTCCGTTAACTTTGTCTGAA | 1 |
| 419595 | - | bf638r_0372 | regulator | 0 | GCTGCAAAGATAGGATAAAAAACGTAAAAGGGTACAGCTACTGTCCGGATA | 1 |
| 419723 | + | bf638r_0373 | hypo | 0 | ATACATACCAGTTCGTATCAGTTACGATAGGTTTCATTATCTTTGCACTGT | 1 |
| 428215 | - | bf638r_0376 | peptidase/deacetylase | 0 | GGATGCAAATGTACAAATAATACTTGAATCTTTCAGTTTGTCAGACAACTA | 1 |
| 428365 | + | bf638r_0377 | AMP binding long chain acyl-CoA synthetase | 0 | TTCAGTATTCACACCTGATAATTCCAAATTAAACGATACCTTTGCAGACGA | 1 |
| 431468 | + | bf638r_0379 | hypo | 0 | TTTTATTTCCTTTCTTTGTCCGTCATTTTATTCTCTATATTTTTGGCAGGT | 1 |
| 431496 | + | bf638r_0379 | hypo | 0 | TTATTCTCTATATTTTTGGCAGGTTAAAAGATTTGTTCTTACTTTGTTAGC | 1 |
| 433594 | + | bf638r_0382 | mem prt | 0 | AAAACCGTATACTACCTGGTCGTAGCTAACAAATAATCCGTATATTTGTTT | 1 |
| 447012 | + | | | 0 | AAAGTAACACAATTTATTGCTATTTTATCTTACTAAATGGTTCTTAATGCT | 0 |
| 461853 | - | | | 1 | CTCTTAATAATCTGGGTTCTGTACCAGTACGCCTTTTGATTTATCTATTTC | 0 |
| 468384 | + | bf638r_0402 | hypo | 0 | TGTATAGGGGCTGCGCTGGTGATACCGGAAATTATGTGCACCTTTGTGTGT | 1 |
| 469458 | + | | | 1 | TGTGTAGGGGCTGTGCTGGTGATACCAGATATTATGTGCACCTTTGTATGT | 1 |
| 469758 | + | bf638r_0403 | hypo | 0 | TTTTATCCGTTTGCTATGGTAATTCAAAAATAATCACCATCTTTGTGGTGA | 1 |
| 470398 | + | | | 1 | CTTTAATCTTTCATGATCAATTTTCTTTTATAATTGGCCTTGAATCTAAAT | 0 |
| 471463 | - | tRNA-leu | | 0 | GCTGCAAAGATAGATACTTTTTCCGATTTTACAAATTTATAGGCAATAAAG | 1 |
| 471795 | + | bf638r_0406 | exp prt | 0 | GCTTACTTCACCAGTATTTTTTGAGGTATAATTATTGGAGTATAAATGCAA | 0 |
| 481212 | - | bf638r_0410 | parC | 0 | AGGCAAAAATACGAATTTAATTGAAAAAATCATTTACTTTGCAGATTATT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 481258 | + | bf638r_0411 | glyS | 0 | TGTGCAGGCAAAAATACGAATTTAATT GAAAAAAATCATTTACTTTGCAGAT | 1 |
| 486731 | + | bf638r_0416 | nitroimidazole resistance-like protein | 0 | TAACAGTATTATATTTAGTAAATATTGT CTCATCTTTATATTTTTGCTTAT | 1 |
| 487421 | + | bf638r_0417 | napC/nirT cytochrome c-type protein | 0 | TGTAATCTGGATTACATTCTTTTTAAAT GGAAATCCCTTTCTTTGCATAAA | 1 |
| 496921 | + | bf638r_0425 | TCS-sensor | 0 | AAGTATCAAAAATAAAGTGTCAATCAT TAAAATTTCTAACTTTGCTAATAC | 1 |
| 507330 | + | bf638r_0432 | FNR family regulator | 0 | TATTTACTAAGTGGTAAGATTGGTACG GGAGAAAATGCTATCTTTGCTCAA | 1 |
| 529869 | - | bf638r_0446 | exp prt | 0 | GTAGCAAAGATACTTTTTATTTTAATAT CAAAAACAATTATTTAGAAATAT | 1 |
| 536453 | - | bf638r_0449 | hypo | 0 | GATACAAATATATAATATTTTTTGTAAT TTGATTTGTTTTGATCGGAAAG | 1 |
| 537804 | - | bf638r_0450 | asd | 0 | GCTGCAAAATTATAACATTTTGCAATA TAATCAGCATTTTAGAGAGAAGAT | 1 |
| 537921 | + | bf638r_0451 | Na+/H+ antiporter | 0 | CTTTCCTTATGAATTGGCGGTGTGATA ACTTTTTTGTTATTTTGCAGATAA | 1 |
| 543762 | - | bf638r_0455 | sodium-dependent transporter | 0 | GTGCAAATATAGAAAATGTTCCTTATA TTTGCGATTCATCTGTAATTTATT | 1 |
| 543798 | + | bf638r_0456 | mem prt | 0 | TTTCTCGCTTTAAACGTGCAAATATAG AAAATGTTCCTTATATTTGCGATTCA | 1 |
| 545537 | - | bf638r_0457 | murF | 0 | GGGTGCAAAATTACATTTTCTATTTCA GATTATCATTAAACAAGAAAGAAT | 1 |
| 546820 | - | bf638r_0458 | transporter | 0 | GTGTGAAGAACAAACGATCCGTTAGA ATGTTCGGCACACAAGCAGCAGAAA | 1 |
| 546999 | + | bf638r_0459 | gadB | 0 | ATAATGTAAAACCGAAGAACTTTCATA TCCTCTCACTGTTGTTGTGGGAA | 0 |
| 551597 | - | bf638r_0461 | TCS-sensor | 0 | AGCTGCAAAGATAACTTTATTTTTAGT GTGCTGATAGCATTAGGACATTTA | 1 |
| 551669 | + | bf638r_0462 | dhpS, dihydropteroate synthase | 0 | TTAGTGTGCTGATAGCATTAGGACATT TATTTTTCTTTATCTTTGCGTCAA | 1 |
| 553457 | + | | | 1 | TTCTTTATTTTCAAGTTGATAATTGTCC TTTTTGCCTGCTTTAATAAGAAA | 0 |
| 554213 | + | bf638r_0466 | mem prt | 0 | GAAAAATAATATTTTATTTGTTACCAA CTCTTTTTTTGATTTTCTTTGCAG | 1 |
| 555354 | + | bf638r_0467 | exp prt | 0 | ATCAGTCTTTCTCACTGGCGGTGGGCG GGATAAGCCATACCTTTGCAATAA | 1 |
| 555915 | - | | | 1 | GTGCAAAAGTAGCAAATTTTATCAAAA CAGGAAAAGGTTTTTGATTCCACC | 1 |
| 557687 | + | | | 1 | TTCGAATTTATCATTTGGTCCTGTTAAG CATCATTTGTATCATTGTTAGGT | 1 |
| 558154 | + | | | 1 | AAAGGATGTTTCCGTTTCCTTTCTTTGG TTTCTACTATATTTTTTATAACT | 0 |
| 558843 | + | bf638r_0473 | exp prt | 0 | ATCAGTCTTTCTCACTGGCGGTGGGCG GGATAAGCCATACCTTTGCAATAA | 1 |
| 561707 | + | bf638r_0477 | exp prt | 0 | TTTTATGAACTCAAATAGTTTAAATCTT ATTACTTTTTTATATTAGTACAA | 0 |
| 567895 | - | bf638r_0481 | acetyl transferase | 1 | TCAGCAAACAACTCGTTGATATACGAG CTATCATCGTGTTTATAGACAACA | 1 |
| 569507 | - | bf638r_0482 | hypo | 0 | TTATATATAACAAAGATAAAGCTATTT TGTTAGAGAATTCCTAATAATCGC | 1 |
| 570593 | - | bf638r_0484 | LrgA, holin | 0 | GTTGCAAAGTTATGAAAAATAGAGTTA GGGTAGCAACTAATAAGGCACGAA | 1 |
| 570724 | + | bf638r_0486 | hypo | 0 | AAAATAAAGAAAAAGTAAACGAAATG GATTACTTTATATATTTTTGCAAAG | 1 |
| 573968 | + | bf638r_0489 | ptaA, phosphate acetyltransferase | 0 | ATACCCCGAATTGCAATGTTGATATTG CAATATTGTGTACTTTTGCGGCGT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 576439 | + | bf638r_0491 | alkaline phosphatase | 0 | CATATCTCTTTCTCCTTTTCTTCTTTTTA AAATTTTTATCTTTGCTATCAC | 1 |
| 577928 | + | bf638r_0492 | hypo | 0 | ATTTATCCTCAATATTCTCTATCTTTCA AAGAATTTCTTTAGCTTTGTCTT | 1 |
| 579113 | + | bf638r_0494 | radC1 | 0 | ACCCAACTTTAATTTTTCCGGATTCCAA TCGTTATGCCGGTCAGATAAGCC | 0 |
| 580862 | - | bf638r_0495 | glycosyl transferase | 0 | GGCAAAGATAAATTATTATTTTCCGAT AATCTCTTTTAATGCACTCTTTT | 1 |
| 580948 | + | | | 1 | TAATGCACTCTTTTTGTCCATTATGCTT TGTCTTGAATACCTTTGGGGTCA | 1 |
| 581907 | - | bf638r_0496 | efp, elongation factor P | 0 | GGGTGCAAAAGTAATCTAAATCGGCA AAAAACACAAAATGTTTTGACTAAA | 1 |
| 582006 | + | bf638r_0497 | rpmH, ribosme prt 34 | 0 | AAAATGACTTAACTCTTGGTTTATTTAA AAAAAGTGCCTATTTTTGCAGCC | 1 |
| 582402 | + | bf638r_0498 | mem prt | 0 | ACTCGGTGAAAGTCTGTGATACTGTGT AACGAAAATAGTATCTTTGTCCGC | 1 |
| 589373 | - | bf638r_0505 | hypo | 0 | GGCGCAAAGATACGGATTTCTATTCCA AAACAAAAATAATTGTCTCATTTA | 1 |
| 590451 | - | bf638r_0506 | exp prt | 0 | GTTTTTTATAAGGAATAGCAAGGCGTT TTGTTCTCGTATTACTAATAATTT | 0 |
| 590829 | - | | | 1 | CTACAAAGATAGTACAATCCGGAGGTT AAGTCAAGCTTTGAACCGATTATTT | 1 |
| 593264 | - | bf638r_0508 | proB | 0 | CACAAATATATGGATTTTTGTCGATAT GTTACCTCTCCCATACGGATATC | 1 |
| 593625 | - | bf638r_0509 | hypo | 0 | ACAAATATAACAAAAAGAACTTAGATA AAGTTGTTTTACCGATAGATTATT | 1 |
| 595165 | - | bf638r_0511 | ompH | 0 | GGGTGCAAATGTATGAAATTAATTTGA ATATCCTAATTTCGCCAGCACTTC | 1 |
| 600505 | - | bf638r_0515 | hypo | 0 | GTTGCAAAAATAGGAACATTTTTCTCT ATATTGAACATTGAAGGGAAGTTT | 1 |
| 601701 | + | bf638r_0517 | hemK | 0 | TCTTTCATCTTTTAGGCTGATAGCTATA TTATTTGTACTTTTGCACCCCA | 1 |
| 603081 | + | bf638r_0519 | pyrE, orotate phosphoribosyltransferase | 0 | AATACGGATTCTTCTTTTTAGTCTCGTT CCATTTCCGTATTTTTGCATCCT | 1 |
| 605738 | + | | tRNA-gly | 0 | CATCGGAAAAAGTTTGGCAGAAAGGT AGAAACTACTTATCTTTGCACCCGT | 1 |
| 606731 | + | bf638r_0523 | hypo | 0 | TTGAGATTCTTCTTCTTTTTTATACCTA AAAAAGTACTACTTTTGTAATAA | 1 |
| 609522 | - | bf638r_0524 | hypo | 0 | GGATGCAAATATAGCAGGTTGTTTGTT CATTTAAGAACAATTTCTGTTCAG | 1 |
| 612732 | - | bf638r_0527 | proC | 0 | GACAAAGGTATAAGTAAAAAGTGAAA ATACGAGCAAACGTGGCAGGTATTT | 1 |
| 617244 | - | bf638r_0532 | argR | 0 | GACGCAAAAGTATGCAATATTTTGGAA ATAGTATGCATAAAAAGCTGTTTT | 1 |
| 617717 | - | | tRNA-phe | 0 | GGTGCAAAGATAGGTATAATTTTTAAT ACTGCAATAGTTTGGCAAAGAAAA | 1 |
| 618074 | + | bf638r_0534 | TCS-sensor | 0 | AACGTTTGTCTTTATTATATGTATAATG CTATTTCCTTATCTTTACCGCAA | 1 |
| 621107 | - | bf638r_0535 | gadC | 0 | AGTGTCTATAACAACGTTTAACGTGAG ATGTTTTAGTTTACTAACGTCAAA | 0 |
| 621257 | + | bf638r_0536 | ion transporter | 0 | AATCCGGGTGGGTATCAAACAAATCG TGTAGGACTCTTGTTCCTTCTGCAT | 0 |
| 623324 | + | bf638r_0538 | hypo | 0 | TTTCGGCGGCTAATCATTCTTACGTCT GTTTTTTATTTATCTTTGCAGTAA | 1 |
| 624256 | + | bf638r_0539 | hypo | 0 | ATAGATAGAATCATTAGGGTGGGGAG TTTCCTCACCCTATTTTTGTGATAT | 1 |
| 628607 | + | bf637r_0544 | mismatch repair prt | 0 | TCGGCTGCAAAGTTATACAATTCCCGG AAAAATATTATATCTTTGGGAAAT | 1 |
| 629974 | + | bf637r_0545 | hypo | 0 | TGTAACTGACTCTTATAGGAAAAATTG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | TGAATAATAGATATATTTGTAGAT | |
| 640334 | + | | | 1 | ATTTACGATTTTACTTGTTTTTATCAGA GTGTTTTTCATACATTTGTACAC | 1 |
| 642105 | - | bf638r_0551 | ecf sigma factor | 0 | GCCGCAAGGTATGAAAAAAAAAGCAA AAGAGGATTAATAGTTGTGAGTATT | 0 |
| 648454 | + | bf638r_0558 | glcK repressor | 0 | TTTTAAAAAAAGATTTTGTAGACTCAA TCTTCTTTACGTAATTTGTGTGCAA | 1 |
| 650664 | - | bf638r_0559 | 50s ribosomal prt | 0 | GGTGCAAAGTAACGAATTATTTGCCG GATAGCCAAATGTTTTTGTATATTT | 1 |
| 650721 | + | bf638r_0560 | UDP sugar hydrolase | 0 | AGTAACGAATTATTTGCCGGATAGCCA AATGTTTTTGTATATTTGTAGCAT | 1 |
| 667489 | - | bf638r+0568 | hypo | 0 | CAGAAAATAACAAGACTCTTGTTAAAA GGTTCGCAAAGAATAATGTTAAAA | 0 |
| 667586 | + | bf638r_0569 | peptidase | 0 | AAAATGCAGTATGATTACTTGGAGGCT TTGGTGAAAGATCTATCTTTGTAC | 1 |
| 668964 | + | | | 0 | TCTTGCCCCTTTGTAGATGAATAGAAA AGTTACAGAGGTGTTTTCACCACA | 0 |
| 671938 | - | bf638r_0572 | pyridine nucleotide oxidoreductase | 0 | ATGGAAATAAACGTTTGCCGACTTAAA AAGGTTCATGGAAAAGCTTAAAAA | 0 |
| 673856 | + | bf638r_0575 | atoC, TCS-regulator | 0 | TCAGCAGTCCGATCATTTGCGTTGTAG CAGAAAACCATTATATTTGCACGC | 1 |
| 683597 | + | bf638r_0582 | rpoE, ecf sigma factor | 0 | AACTACTTATGACATTATTATATTAAG GTTTTTTATTATTTTTGCCGACGT | 1 |
| 685312 | + | bf638r_0584 | frrG | 0 | ATTTATATAGATTAAGGCTTGGCTTAT ATTATCATTTGTTAACATTATATC | 1 |
| 685312 | + | | | 1 | ATTTATATAGATTAAGGCTTGGCTTAT ATTATCATTTGTTAACATTATATC | 1 |
| 685370 | - | | | 1 | TTGTTTAACTAAAAAAAATAATATTAT GAACAAAAATCGATCAATCCGAGT | 1 |
| 690348 | + | bf638r_0586 | frrI | 0 | CATTCAGAATGAAGGATACTAATGCTT CTTAAGAGATTATTATAAATTAAT | 0 |
| 692978 | + | bf638r_0587 | hypo | 0 | AACAGTTTCTGTAGTAAGCTTGAAAAA TAGTAGGAATACGTATCTTTGCAT | 1 |
| 698657 | - | | | 1 | TTATGCAAAAATAAATGTTTTGTTTTG ATATACAATATATTCTTTGGGATC | 1 |
| 699097 | - | | tRNA-met | 0 | GGGTGCAAAGATAGAGCCTTTTTTTTA ATCTCCAAAACATTTTCTATCTTT | 1 |
| 699282 | + | bf638r_0591 | hypo | 0 | AGTGAATTAATTATGCCTTGTAATTAG AAAAAAGTATTTATCTTGGCACAC | 1 |
| 699749 | + | bf638r_0592 | permease | 0 | TCGACCAAAAATTTTTCGCACTCCTTTT TTTGTGCTAATTTTGTATCTTAA | 1 |
| 701680 | + | bf638r_0593 | ribA, riboflavin biosynthesis | 0 | GGATGTGGAATTGACAATAAAAGATA TTAATAACGAAAACAAGATATAAAA | 0 |
| 702948 | + | bf638r_0594 | aspC | 0 | ACTGTCTATTTGCTTTCATATATCAAAT ATTATAGTTTATTTTGCAATCAA | 1 |
| 707327 | - | | | 0 | AGGTACAAAAATAGAGTTTTTAATCAT TAGTTGGTATCGTTTTTGATTAAA | 1 |
| 707429 | + | bf638r_0596 | metA | 0 | TGTTTTAATTTGTATATCTTCTAAATAA GGAAAACACTATTTTTGCATTGA | 1 |
| 713726 | - | bf638r_0599 | tonB receptor | 0 | GTTTGCAAAGATGCAGCAAAAAATGTT AAACGGCAATCCTCATTTATTAGG | 1 |
| 717864 | - | bf638r_0603 | radA, dna repair | 0 | GTCTCAAAGTTACGCATTTCGTCTTTTT CTGCCAAACAGAATATAGCTTTT | 1 |
| 719577 | - | bf638r_0604 | ansA, asparaginase | 0 | GTTGCAAAGATAATGATATATCTTGTA AATGGCAGCAAAGTGTCAAGAAAGT | 1 |
| 719785 | + | | | 1 | TAATCGCAAAAAGTGATAGATTGAAT GAAATTATTCTTATTTTTGCAACAT | 1 |
| 720032 | + | bf638r_0605 | aspartokinase | 0 | ATCAACAAAGATGGTTTTGTATGCCTG AAAGAGCGTACAAGGCATTCCATT | 0 |
| 723728 | + | bf638r_0607 | thrC | 1 | AATTTATTAAAGAGTTTATGAACCGTT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | AAGGAATAGGAACAGAGTTTGTTCA | |
| 725544 | + | bf638r_0608 | mem prt | 0 | TCGTCCTTCTTTCTCCGAACTTTCCGTC TCCCTCTTTTGTTGTACGGGAAT | 0 |
| 729969 | - | bf638r_0611 | tonB receptor | 1 | GGCGGCAAATTTACGCTTTTTGCTTGT GTAAAGCAAGAAACACTTCGCTAT | 1 |
| 730050 | + | bf638r_0612 | mscS, ion channel | 0 | AGCAAGAAACACTTCGCTATTTGATAA AAACATTGTAGCTTTGTCCGCTGT | 1 |
| 731322 | - | bf638r_0613 | hypo | 0 | ACAGTTAAAACATTCTCAAAGAGGATA AGGTTCAAACTGTTTTTTCTTATC | 1 |
| 731383 | + | bf638r_0614 | ecf sigma factor | 0 | ATTCTCAAAGAGGATAAGGTTCAAACT GTTTTTTCTTATCTTTGCCATTTC | 1 |
| 748134 | - | bf638r_0622 | hypo | 1 | CTTTCAATCTGTGGCGATTTCAGGTAT ATGGCACATTTACCAGAACTTATT | 0 |
| 748367 | - | bf638r_0622 | hypo | 1 | GGCTGCAAAGATACGGTAACTCTGTG TATCCTGCAAGAGCACTTCTTTCTT | 1 |
| 752285 | + | bf638r_0626 | araC | 0 | CCTATTAATGGGGATTGTACCGATGTA TGGGTTATCGTAATTTTGTATCAG | 1 |
| 752297 | - | | | 1 | GGGACAAAAGTAGAGCCTTATCGGCT TTTACTGCTTACGCGAAAAGGAACT | 1 |
| 753915 | - | | tRNA-lys | 0 | GGGTGCAAAGGTAGTGACTTTTTTTAT ATTTGCAAGCTTTTGGATGAAAAA | 1 |
| 755795 | - | bf638r_0630 | tdk, thymidine kinase | 0 | GCTTGCAAAAGTAATAGTTTTATTGGC TTCTCAGTGTTTTTTCATTATTTA | 1 |
| 755863 | + | bf638r_0631 | hypo | 0 | TTTATTGGCTTCTCAGTGTTTTTTCATT ATTTATTTCTACATTTGCCAAGT | 1 |
| 759020 | + | bf638r_0635 | amino acid transporter | 0 | GTTGATGGAACTATTGTAACATTTATA GATATATTGTATTTTTGCGTTCAA | 1 |
| 764405 | - | bf638r_0637 | surface prt | 0 | AATATGATTATTACAAATTCAGATTGA ACTTAGTGTAGAAAGAAACTGTAT | 1 |
| 769529 | - | bf638r_0639 | tonB receptor | 0 | TGCAAATATATTAAGAAATATGAAACG GACAATCATTTTTATAAAAAAATC | 1 |
| 769904 | + | bf638r_0640 | integrase | 0 | TAAATCTTATTCAATTCTAATATAGTAC CCAAGCACGTCTTATTTGTCCAT | 1 |
| 771721 | + | bf638r_0641 | tonB receptor | 0 | AAACTCTTATAATAAAGAAGTTATGAT ATAATTGTAGTTCTGCAAAGACTT | 0 |
| 776167 | + | bf638r_0643 | gtp binding prt | 0 | GGGCGATTTTTTCTTTTAGTATGTTACG AAAATAGCTATATTTGTCGCCAT | 1 |
| 779487 | + | bf638r_0645 | pepN | 1 | TAACCGCATCGGGCAATTGGCGTGAG AGAGGGTACGAGTTATATAATATTTA | 0 |
| 782395 | + | bf638r_0646 | fumB | 0 | AATCTGCCACTTTATTCATACTTTGGAA AACTATTCTTACCTTTGCCTCAC | 1 |
| 784078 | + | | | 1 | CTTTAAGCAGATAAAACCCCGTTGCGG TTCTTGCAGCAACAAATAAGAAAT | 0 |
| 789369 | - | bf638r_0649 | amino acid transporter | 0 | AGTTGCAAAGATAATTAATTCATAACA CATACGGTCATAGGATGTTCACTT | 1 |
| 800289 | - | bf638r_0657 | hypo | 0 | TTGTTTCTTGACTTTTTCGTGAAGCAGT ATTTTTTGCTTTTCGACCTGTTG | 0 |
| 803347 | - | bf638r_0659 | hypo | 0 | TTTGTTAATATATGACTTTTTTTATATG AACATCCTCCAAAAAGATTTTTT | 0 |
| 805903 | - | bf638r_0662 | frr, ribosome recycling factor | 0 | AAAACAAAAGTAGATTATTTTTTGATT ATCTGCAACTATCTATAGAGAGAA | 1 |
| 815277 | + | | | 1 | ATACCGATTATCGGGGCTATTGATGAT AAAAGGTCTGCTGTAAGAAAAAGA | 0 |
| 816122 | - | bf638r_0669 | pyrH | 0 | TGCAAAGATACGCGATTCCTTTTAATG GACTATCTTTTTTGCTAAAAAAG | 1 |
| 817461 | - | bf638r_0670 | dinF, dna repair | 0 | ATGCAAAGTTACATGAAAAACAGAGA ATATGAAAATGATGTGAACAATCAT | 1 |
| 817570 | + | | | 1 | ATTGGTTCGCAACGTTCTCGATACAGG AATTAATATTATTTTTGTCTGCAT | 1 |
| 817978 | + | bf638r_0672 | endo beta | 0 | TAACGTTGACGTAAATGATAGTATGGC | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | galactosidase | | TATTCTTTCATTAAAGACATCTTC | |
| 819068 | + | bf638r_0673 | hypo | 0 | AGTGAGTATGTGACTTGGGAATACAC AATATAAATACTATCTTTGTAGGGA | 1 |
| 821139 | - | bf638r_0674 | hypo | 0 | CAGGGTCAAAGATAATGGTTTAAACTT TAAATCTCAATTTTAAAGAATGTAT | 1 |
| 822720 | - | bf638r_0677 | hypo | 0 | TCTGCAAATTTACTCTTTTTATTCTTTG AAATTACAGAAGTAAAAGTAAAT | 1 |
| 822779 | + | bf638r_0678 | hypo | 0 | TTACTCTTTTTATTCTTTGAAATTACAG AAGTAAAAGTAAATTTGCAAAGA | 1 |
| 823794 | - | bf638r_0680 | hypo | 0 | GTAGTACAACAAACGTGCCAGAAGGA TTGTTTATTTATTTTTTCTTGTTAA | 0 |
| 825252 | - | bf638r_0682 | psd, phosphatidylserine decarboxylase proenzyme | 0 | ATCCGCAAAAGTATACTTATTTTAAAA ATCCAGCAAGCATTTGCAGGTAGA | 1 |
| 825373 | + | bf638r_0683 | dnaE | 0 | TTTTGATGTTTATCTTTGTTAGTAGCTG TAGACATTGTAATTTTATGGCTT | 1 |
| 829203 | + | bf638r_0684 | trxA | 0 | AAGGAACGTCCCGAACTTGCGTTTCAT ATAAATTGATTATCTTTGCCGCAT | 1 |
| 830906 | - | bf638r_0686 | mem prt | 0 | GCTGCAAATATATCATAATTCTCTTACT TTTGCTTCCGGTTTCGGACAAAA | 1 |
| 831785 | + | bf638r_0688 | mem prt | 0 | TTTTTAGTTTTGGTTGCAAATATATCAT AATTCTCTTACTTTTGCTTCCGG | 1 |
| 839899 | - | bf638r_0697 | hypo | 0 | GCAAAGGTACTTATTAATTTGTATTATT CCGTATATTTGTCCTTGCAATGTG | 1 |
| 839944 | + | bf638r_0698 | amino peptidase | 0 | TTTTCGCAAAGGTACTTATTAATTTGTA TTATTCCGTATATTTGTCCTTGC | 1 |
| 844648 | + | bf638r_0701 | mem prt | 0 | AATCGGCGCATACGTTTGCCAATTCTA TAATTTTTCGTATTTTTGTTGCAC | 1 |
| 847891 | + | | | 0 | TTTTGCTGTCTTTCTGAAACCTTGGTTC TTCTTTGTGCTATCTTTGTGTAT | 1 |
| 860500 | - | bf638r_0708 | aroA | 0 | GGTTGCAAAAATAGGAAATCTATTCTT TTTTGCCTCTATCCTCTTCCTAAA | 1 |
| 860742 | + | bf638r_0709 | araC | 0 | AAAGACTTGTGTAGTTACGGGGTTAAT GTAGTTATTTTACTTTTGTATCAA | 1 |
| 863049 | + | bf638r_0711 | beta-N-acetylhexosaminidase | 0 | TTTTAGACAGAATGTTGGAGAATCGGT AGATTTACACTACTTTTGGTACCA | 1 |
| 866490 | - | bf638r_0712 | acyl transferase prt | 1 | ACGCAAAGATAGAAAAAACTTGAGAT ACTGAGCAACAAACGGAAAGAATCC | 1 |
| 866804 | - | | | 0 | GGATGCAAAAGTAGGATTTTATTTTGA ATAAGCAAAATATTCGGATGGTTT | 1 |
| 868576 | - | bf638r_0714 | aroB | 0 | ACTGCAAAGATAGGAGTTCGGAGATA AATAGTGAAACTTAGCACTTAAAAA | 1 |
| 868674 | + | bf638r_0715 | mem prt | 0 | AAATGTAAAACATGGGGTTCTGTTTGT GAGAAAAAGCTATCTTTGTGACAT | 1 |
| 872220 | + | bf638r_0717 | cls2, cardiolipin synthetase | 0 | TCGGGACAGTTGTTGTTCGTATCCGAT ATTTTCCATACATTTGTTCTTTGT | 1 |
| 875823 | - | bf638r_0720 | hypo | 0 | AGCGGCAAATTAAGGAAAATTTTTCTT ATCAGCCAACTTTTGAGCAAGAAA | 1 |
| 875743 | + | bf638r_0721 | putative ATP-dependent exodeoxyribonuclease | 0 | AAATACTGTTTTCATGATAAATATGTTT TTTTTAGTTTATATTCGTAGCGC | 1 |
| 882557 | + | | | 1 | AATTAGTTAACTTTAAAAACATACAAG GTTTAAGCATTGTATTTATAATAT | 0 |
| 882584 | - | bf638r_0723 | tonB receptor | 0 | GTTAATAAAACAAGTGTCCCGAAAGA AACACTTACTTCAGTTACTAACTTT | 0 |
| 888421 | - | bf638r_0726 | anti sigma factor | 0 | ATAATAGAAACAACTAAATCTTTAAAA ACACTTATCTATTTGATAAAAAAT | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 891307 | - | bf638r_0728 | alpha mannosidase | 0 | TGTGCAAAGGTAGAGAATATAAATGT AAAACGAAAGAGCTTGTGTGCTGAT | 1 |
| 892980 | - | bf638r_0729 | efflux transporter | 0 | GGTTGCAAAATTACTCCATCTGTTAGG AATCGAATATAGGCTAATCTGTAC | 1 |
| 893093 | + | bf638r_0730 | araC | 0 | TTTCAGAACAAAATAGCCCGAATATCA AAAATAAACCTACATTTGTTGCCT | 1 |
| 896870 | - | bf638r_0731 | alaS, alanine synthetase | 0 | AGTGCAAAGATAGCTTGTTTTTATTAT TTTTGCCGTATTAACAAGCGAAAT | 1 |
| 896907 | + | bf638r_0732 | membrane peptidase | 0 | TTCTGAAAAAAGCAGTGCAAAGATAG CTTGTTTTTATTATTTTTGCCGTAT | 1 |
| 900602 | - | bf638r_0734 | relA | 0 | GCAAAAATAATCAAATTTAAACATCAG GCAAGCGAAAGGCGATTTATTTGT | 1 |
| 904646 | - | bf638r_0739 | parA | 1 | CTTGCAAATATACGTATTAAAATTGAA TATCGATTCCTTTTTTTACTCACT | 1 |
| 904970 | + | bf638r_0741 | surE, stationary phase phosphatase | 0 | TCCTATTTGGTTCTGTTGTCGATACAA GTTCTTTTTGTTACTTTTGTGCAT | 1 |
| 905741 | + | bf638r_0742 | lpxB, lipid disaccharide | 0 | ATGGATACGTAGCGATTACACCTACTG TAGTGGATATGACCGCTTATCATT | 0 |
| 911198 | - | bf638r_0746 | hypo | 0 | TTACAAATATACGCTGATTTATTGTATC AATGAAAGTAATGTGCTAAATCT | 1 |
| 911371 | + | | tRNA gln | 0 | AAGAGCGAAAAGTATTTGTGATTAAA AAAAACTTTGTACCTTTGCAACCGC | 1 |
| 913715 | - | bf638r_0747 | hypo | 0 | AGATCACAAATTAATGAAATAAAATGA TTACTTCATACTATTTTCCTTTAT | 1 |
| 916072 | - | bf638r_0749 | ksgA, dimethyladenosine transferase | 0 | AGCAGACAAAGGTAATTCTTTTAAATG AAGAATAAAGAATTAGGGATGAAG | 1 |
| 916077 | + | bf638r_0750 | membrane prt | 0 | TTAGGCTTTACTAATTTCATTTATCAAC TAAATAACTACTTTTGCAGCAGA | 1 |
| 917231 | + | bf638r_0751 | pepD, aminoacyl histidine peptidase | 0 | TCTTCCGGTAATTCTTTCTTTTCCATTCT TTTTAGTAGCTTTGACCTTTAG | 1 |
| 919023 | + | | | 1 | TATCTTGCTTTTGAGTACAACGTATCTG TTTTGGCTTTACTTTTGTGTTGT | 1 |
| 925763 | - | bf638r+_755 | hypo | 0 | ATTACAAAGATACCCGAAAAAGTTTTG CTATTTTGATTTTTGTGTTAAAAG | 1 |
| 925915 | + | bf638r_0756 | hypo | 0 | TTTAAGCTAAAAAGAACAAATCTTTAA ACTTTAATTTGTTACTTTGTTAAT | 0 |
| 931873 | + | bf638r_0757 | hypo | 0 | TATTTTTATTAAAATGGTTTTTGTTGCA AATAAAACTTATCTTTGTAGGAT | 1 |
| 934037 | + | bf638r_0760 | hypo | 0 | AATATGTAATCAATGGCAGAATCTCAC TGAAAACCCCTATTTTTGTACTCT | 1 |
| 937225 | - | bf638r_0762 | permease | 0 | GATGCAAAAGTAACCAATTTATCGTTT GAATTTTCTATATTTGTTTAATTT | 1 |
| 937272 | + | bf638r_0764 | ABC transport | 0 | ACTGATGCAAAAGTAACCAATTTATCG TTTGAATTTTCTATATTTGTTTAA | 1 |
| 940399 | + | bf638r_0766 | araC | 0 | TTCGGGAGTGGAATAGAAGAAATTAA AAAAGTGTGTTATATTTGTGGCTAA | 1 |
| 942773 | - | bf638r_0770 | dna binding prt | 0 | AAACAAAAATACTAAGTATTTGTTTGA TAAGCAAGGAAATTCCCAATATAT | 1 |
| 945255 | - | PS-G inverted promoter | | 1 | AACGCAAAGATAAGGGGTATTTTCATA ATAGCAAAATGTTTAATTAAAAA | 1 |
| 972143 | + | bf638r_0814 | hypo | 0 | ATTATGGTTAGCAAAACCGATAATTAA GAATAATCTTGTATATTTGTGAGT | 1 |
| 974815 | - | bf638r_0815 | aldo_keto-reductase | 0 | GATGCAAAGATAATCGATTTATGGAG CAACATAGGTTATCAAAATAATAAT | 1 |
| 974928 | + | bf638r_0816 | glycosyl transferase | 0 | GGATAAAATAATACTGGATGGTTTGAT GGAATCCCGATAATTTTGCATTAT | 1 |
| 1002481 | + | bf638r_0833 | membrane prt | 0 | TGTTTCCTGTTTCGTTAGATTGTTTTCC AAACTTTTTGTACTTTTGGGCAC | 1 |
| 1010920 | + | bf638r_0840 | TCS | 0 | TATTTAATCCATTCTCTTGGTGTGAATG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | GATAATTTCCTAAATTTGCGTAC | |
| 1026321 | - | bf638r_0851 | hypo | 0 | GGTTGCAAATATAGAATTTAATTTTAT CTCTTTTGCAAACTCCACGTGACA | 1 |
| 1031290 | - | bf638r_0854 | dxs, 1-deoxy-D-xylulose-5-phosphate synthase | 0 | AATGCAAATTTAGTAAAAGAAAGCGG AAGGAGCATCTTTTTGATTGGAAAT | 1 |
| 1033461 | - | bf638r_0856 | transmembrane prt | 0 | TTCACAAACTTAAAATATTTTCTTTGAA AAAGGAGCTTTTCGGACTGATTC | 1 |
| 1033595 | + | bf638r_0857 | gph, putative phosphoglycolate phosphatase | 0 | AGGTGGGATGGCAGTAAAGTTACCGC TGAATTATCCGTATCTTTGTCGTAA | 1 |
| 1033604 | - | | | 1 | GTATAAAGATATGAAGAAACTGATAA TATTCGATTTGGATGGTACTTTATT | 0 |
| 1040843 | - | bf638r_0861 | sulfatase | 0 | GCTGTAAAGTTACTCTTTTTTGGGTAG AATACAATTTTTATTGAAAATAAT | 1 |
| 1046307 | + | bf638r_0864 | TCS | 0 | TTTTTTTTGAATCTCTTTTGTATGTTTAG GCTTTATTTCGTTCTTTTGCAA | 1 |
| 1052439 | + | bf638r_0866 | fucosidase | 0 | TTCTTCCTTTTATTTTGTATTTTAAGATA GATATTTTATCTTTGCCTTAAT | 1 |
| 1055321 | - | | | 1 | AGTGCAAATATAGCGGTTTGCAATATA CTGAGGTGTTAACGAAGTTCAGAA | 1 |
| 1058848 | - | bf638r_0868 | hypo | 0 | ACAGCAAAAATACATGATAATTCCCGG ATTTTCATGCGGACGGATGAAAAT | 1 |
| 1060060 | - | bf638r_0869 | mdmB, acetyl transferase | 0 | ACTGCAAATATATGGAAAGTATAGTTA GTATGAATCAAGAAGTATGCTTTT | 1 |
| 1061326 | - | bf638r_0871 | tatA-sec independent translocase | 0 | GTGCAAAGATAGACTATTTTGAATGAA AAAAGAAAGGCAGAATCATAAAAG | 1 |
| 1065001 | - | bf638r_0873 | hypo | 0 | AGTACAAAGGTACATGATAATTTCGG GAAACGTATGCAAGCAGCTGAAAAT | 1 |
| 1065938 | + | | | 0 | GCTTTATAAGTATAGCAAAATTAAGAT GTTTATTTACTTAATTGGCAGGAT | 0 |
| 1066582 | - | | | 1 | CTTTATGACAAAGCCGTGAGCATATAA CTAAATTATATGCTCTTGGCTGTT | 0 |
| 1066689 | - | | | 1 | ACTACAAAGGTGAGAATTATTTTTGAT TTAGCATAGCAAACGGTTAAAAAA | 1 |
| 1071124 | + | nf638r_0881 | hypo | 0 | CTGATCGTGGAAATCGATCTGGAGAA TCTGGAAAAATAAAAGAGAGATAAC | 0 |
| 1074268 | - | bf638r_0884 | hypo | 0 | GGCGTAAAAGTAATAAAAAACCGAAA GATTGCAAATAAGATGGCTGGGTT | 1 |
| 1077613 | - | bf638r_0887 | lpxD | 0 | ACGCAAATAACGGACTTTTTTTTTATAT ATCCTTACTCTTTGATGAATATT | 0 |
| 1079735 | - | bf638r_0889 | pyrF | 0 | TTACAAATTTACGATTTACAATTGAAG TAACCGTCCGTTACAATTCACTCT | 1 |
| 1082137 | - | bf638r_0891 | purM | 0 | AACGGCAAAAGTACAGCTTTTTTTCCT TATAGCCAACAAATTGGAGGATAT | 1 |
| 1084954 | - | | | 1 | GACTGCCGTCTTTTTGGTTGGTTGCGG AGCTGCAATGTAATAAGCATGTAG | 0 |
| 2999613 | + | bf638r2545 | sodB | 0 | AAGTTTGGACAACTCCGAACATTTGGG CCCTATATCCTGTTATATATTCAG | 0 |
| 3000465 | + | bf638r2546 | thiamine biosynthesis | 0 | ATCCCCGATTTACTTTCTTCTTATTTTCT TTTTCACTTACTTTGCCTCCGT | 1 |
| 3006587 | + | bf638r2553 | thiamine biosynthesis | 0 | TTCACTATAAAGGTTGGTTTCTGAGTA AGTTAATTAGTAACTTTGCTAAAC | 1 |
| 3007877 | - | bf638r2554 | hypo | 0 | GGGCGCAAAAGTATATCATATTCATTA ATTAAACAACTATTTCGATCATT | 1 |
| 3008625 | - | bf638r2555 | hypo | 0 | AAGATAAAGGAAATATTTTTATTTGAA TTATACAAATATTTACTCTACATT | 1 |
| 3008936 | - | | | 1 | ACTACAAAAGTAGCATTAAGTTTTTAA TAAAATTGTTTTGGAGAAAAAGTT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3020332 | - | bf638r_2565 | ppdk | 0 | TTCGCAAATCTAAGGATATTTCGCAAA CCAGAAAACTTTTATATGAAAA | 1 |
| 3020472 | + | bf638r_2566 | tRNA Methytransferase | 0 | TTATTTCACTTCATATAGAGAAATACG AAAAAAAATACCTACTTTTGCGAAAG | 1 |
| 3023961 | - | bf638r_2569 | hypo with Signal peptide | 0 | ACGGACAAATATAAGGATTGTTTTCTA ATAAAAAGGAAATAGGATAAATAAAA | 1 |
| 3024742 | + | bf638r_2570 | Methionine aminopeptidase | 0 | ATCTTTCCTAAAATACCGCATTCCCCTT TTTATTCCCTACATTTGTACTA | 1 |
| 3032314 | - | bf638r_2575 | EF-Tu | 0 | GCGCAAATATACACAAATGCCAGAAC ATAAAAAAATGCGCTGATAACATATCA TA | 1 |
| 3032671 | - | bf638r_2576 | Unkown fuction | 0 | ACTGCAAATATAGTTATAATTATGTTA TATAACAGATTATTAGTAACTT | 1 |
| 3033241 | - | bf638r_2577 | Redox active protein | 0 | TTGCAAAATTATTAAAAAAGTATCTTT GTCGGCAACTAATCAATGATTTT | 1 |
| 3033272 | + | bf638r_2578 | Unkown fuction | 0 | TTTATATTTTTTTATTTCTTGCAAAATTA TTAAAAAAGTATCTTTGTCGG | 1 |
| 3035792 | - | bf638r_2580 | natrual resistance associated macrophage protein | 0 | AATAACAAGCGAGAACCTTGTTCGGTT CACTTTTAACAACTATCTTTGCCCA | 0 |
| 3035844 | + | bf638r_2581 | Unkown fuction | 0 | CAAGCGAGAACCTTGTTCGGTTCACTT TTAACAACTATCTTTGCCCAA | 1 |
| 3039986 | - | bf638r_2584 | phenyalanyl tRNA synthetase subunit | 0 | CCGCAAAGTTAATAAGATTTTTTGGTT GATGGGGAATGAATGGTTATAAAAAA | 1 |
| 3057177 | - | bf638r_2600 | NusG | 0 | AGCACAAAGGTATAAACTAAATTTTGA TGTGCAAAATATTTTTAGATGAAAAA | 1 |
| 3057338 | + | | | 1 | AGCACAAAGGTATAAACTAAATTTTGA TGTGCAAAATATTTTTAGATGAAAAA | 1/2 |
| 3062979 | - | bf638r_2603 | DnaB | 0 | AGGCAAAGATAATTCTCTATTATCAAT TTTCCATTATCAAATCTGAATTATT | 1 |
| 3063049 | + | bf638r_2604 | 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase | 0 | ATCAATTTTCCATTATCAAATCTGAATT ATTTATTTAACTTTGCGTCCA | 1 |
| 3067983 | - | bf638r_2607 | galE | 0 | TTGCAAATGTAACAAAATTAATTTAGA TAAATGCACAAAAAACGAATATT | 1 |
| 3075909 | - | bf638r_2616 | Unkown fuction | 0 | GGGCAAATTTACATTAAATAGTTTAGA TTACATAAAATTTGCTCGATTTTA | 1 |
| 3076222 | + | bf638r_2618 | CTP synthetase | 0 | TTTGTTTTGTAGAAAGCATCCGTTCTTC CCCATTTTTTTGTAATTTTGCGGT | 1 |
| 3081229 | + | bf638r_2621 | Peptidase | 0 | AACTGGCAGAAAAAGCAGATTTGACC TGATTTTTGCTTATCTTCGCAGCA | 1 |
| 3086078 | - | br638r_2624 | Transcriptional regulator | 0 | GCAGCAAATATAGGCATTGCTATGGT ACTATGCAATACAAAGTACCTTATTTT | 1 |
| 3086197 | + | br638r_2625 | acyl transferase | 0 | TTAAGTCCCTTTCACCTTATTATAAATG CAAATTATCACTATATTTGCCGT | 1 |
| 3088526 | - | bf638r_2626 | transmembrane | 0 | GCAAATATATACTTTTTACTGAAATTA AACAGTATATAATACACAAA | 1 |
| 3092131 | - | bf638r_2628 | uvrA1 | 0 | GCTGCAAAATTAATGTTTCCGCCCGAA ATATAGTTTTAAAAAACACAATA | 1 |
| 3092240 | + | bf638r_2629 | Unkown fuction | 0 | TTAACTCAATATATTATGTGAATGTAT GACATTATTCGTACCTTTGCCGTAA | 1 |
| 3096575 | - | bf638r_2633 | YfiO | 0 | GCGCAAAGTTAAAGATTCGCGCTGAA TAAACCAATTAATCAATGTTTTTTAAG | 1 |
| 3098526 | - | bf638r_2635 | paak1 | 0 | TGTGCAAAGTAAAGGATTTATTTTAAA ATTTGATTTGTTTATCTCAACTATT | 1 |
| 3100804 | + | bf638r_2637 | Unkown fuction | 0 | GGTTTTCTATGAAACATTTGCCAAGAT CTCAAAAATCTGTATCTTTGCATT | 1 |
| 3103067 | + | bf638r_2640 | Chaperone | 0 | | 0 |
| 3106767 | - | bf638r_2641 | Unkown fuction | 0 | TCTACAAAAATACACATAAGTTACAAG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | CTAACGTACATTTACAGAGAAAT | |
| 3114151 | - | bf638r_2647 | with signal peptide | 0 | ACTAACAAATATACGAATAGTTCGTAA CTCCGGATTGTTTGTACGGATAAA | 1 |
| 3114438 | + | bf638r_2648 | Unkown fuction | 0 | TTATATCGGGCCTGCTTATAAATCGTA TATTTTACCTAAGTTTGCCGT | 1 |
| 3123377 | - | bf638r_2652 | glutamine amidotransferase | 0 | GCGTGCAAAGATACACAATCTATATGA CATAAAGATAAGCTTTCCTATA | 1 |
| 3125819 | + | bf638r_2654 | phosphoinositide-specific phospholipase C-like phosphodiesterases superfamily | 0 | TTAGCATCCTTATCCGGCTAAATTGCG TTACTTTTTATATCTTTGAAGGC | 1 |
| 3126766 | + | bf638r_2655 | Unkown fuction | 0 | TCAAAACCAAGTCTTCCAATATTTCAA ATTAATTATCTACATTTACGGTCAT | 1 |
| 3128434 | - | bf638r_2655 | Unkown fuction | 1 | AATGCAAAGATAGACTTTTTCTGCAAT GTATGGTAGATTTTATGTAGAATA | 1 |
| 3128811 | + | bf638r_2656 | Prophage Transposase | 0 | ACTACGGGATTTATCTCGTAGTTTTTTT ATTTATCTATTAACTTTGTAAA | 1 |
| 3139492 | + | bf638r_2669 | Unkown fuction | 0 | CATTCCTCCCGTTTTTCGAGGTTATTCC AGAGATTTCTATATATTTGTAC | 1 |
| 3140799 | + | bf638r_2670 | dapF | 0 | AAGACATGCTTTTTATAAAAACAAAAG TCAAAAAGACATATCTTTGTGCGG | 1 |
| 3143704 | - | bf638r_ 2672 | Unkown fuction | 0 | AAAAAGAACAAATGAATTTTGTAATAA GTTCATAGCAAGATTGTAATATCT | 1/2 |
| 3143786 | + | bf638r_2673 | polyphosphate kninase | 0 | CAAGATTGTAATATCTATTACATCATG CAAGCGTACTTTTGCAAAAAA | 1 |
| 3148164 | - | bf638r_2675 | Butanol dehydrogenase | 0 | GATACAAAAATAATCGCTTCTTTCCAG AAAACAAAAAGATTCGCAAGGTA | 1 |
| 3148469 | - | | hypo or sRNA or else | 1 | ATCTGAGAAACAATATAAATAACAAAA TGTTTAAAAAGAGTTTCCTTTTCAA | 0 |
| 3149137 | + | br638r_2678 | trpB | 0 | AAGAAAGATTTTATTTGTTATTCCAAA CAAAGTATCTATATTTGCCCCCGT | 1 |
| 3163550 | + | | hypo or sRNA | 1 | CCCTAAAAAGGGGATTGTGGCAACCG CCGAAACCGCCTATCTTTGCAGTAT | 1 |
| 3165387 | - | bf638r_2693 | Unkown fuction | 0 | GGGTGCAAAGAAACACAATATTTTAAT GGTATACAATCACTATAAATAGGT | 1 |
| 3165554 | + | bf638r_2695 | Unkown fuction | 0 | AAATCACAGATTATTATTCATCGTTCAC CACTATTTTTTACCTTTGCACAAT | 1 |
| 3166855 | + | bf638r_2696 | Flavodoxin | 0 | TTATAAATGGGGATTTTCTATTTCTGCC GGACCGCCTATCTTTGCCGTAT | 1 |
| 3171242 | - | bf638r_2702 | hypo with Signal peptide | 0 | AGAGGTCAAAGATACATTATTTTCCGG AAAGAGATAAAAAAGTCCGGTCGA | 1 |
| 3172609 | - | bf638r_2703 | Methionine synthase | 0 | CGACAAATATAAAAACAATAAATCAG ATATTTGTGTTCAACATGATACTT | 1 |
| 3175499 | - | bf638r_2705 | hypo | 0 | GGCAAAGATAAGAAGTATTCCATAAT GGAACAATCCCCATTTATAGGGATA | 1 |
| 3176938 | - | bf638r_2707 | Unkown fuction | 0 | GGACAAAAGTAATAAAATGAATCTGA ACCGGGAAAAATAAAGTCAATAAA | 1 |
| 3181057 | - | bf638r_2710 | permease | 0 | GTGCGCAAAGGTACGTAAAAAACTGT TTTCGACTATAGCAAAAGAAGAAG | 1 |
| 3183026 | - | bf638r_2713 | hypo | 0 | GGCACAAAAGTACGCAGATCACCAAG TGAAGGCAATACCTACTAATATCT | 1 |
| 3183560 | + | bf638r_2714 | Memberane transporters | 0 | TTCTGTTTAGGTATTGTAACACCTTTTT AGTCCACCTACCTTTGCAAAC | 1 |
| 3195099 | + | bf638r_2722 | ribonuleotide diphosphate reductase | 0 | AAAACACACTCTCTGCACAACCATTCC AAATTCTATTATCTTTGCTCCCT | 1 |
| 3199219 | + | bf638r_2724 | Unkown fuction | 0 | TCTTCATTCAAAAAACTTTTGACCCCTG ATTTTTTTAATTACCTTTGTACC | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3202906 | - | bf638r_2728 | cabohydrate kinase pfkb family | 0 | GACACAAAGATACGTAAATTTTCAGGA TTACAAAAATTGTTATATATTTGT | 1 |
| 3202962 | + | bf638r_2729 | Sugar transporter or lacI regulator | 0 | AAGATACGTAAATTTTCAGGATTACAA AAATTGTTATATATTTGTGTCCAT | 1 |
| 3205126 | - | bf638r_2731 | Transcriptional regulator | 0 | GACGCAAAGATATTATTTGATTCATAA AAATAATCAAAGAGGGTATGTC | 1 |
| 3205261 | + | bf638r_2732 | hybrid Two component regulatory system | 0 | TTTAGTCTAAACAAGAGGTACAATCCA TTTATTATTAATACTTTTGTACAC | 1 |
| 3212086 | - | bf638r_2737 | Unkown fuction | 0 | GGCGGCAAAAGTAATGGTTCTTTTGAT TTTCACAAAACGAAAAGACATTTTT | 1 |
| 3213295 | - | | | 1 | GATGCAAAATTACCGTTTTTCTCCCTAT CATTCACTATCTTTGCACCTC | 1 |
| 3213343 | + | bf638r_2738 | Na+/H+ antiporter | 0 | ATGCAAAATTACCGTTTTTCTCCCTATC ATTCACTATCTTTGCACCTC | 1 |
| 3216662 | - | bf638r_2739 | Unkown fuction | 0 | AGCAAATATATGCTATTTGTGCGAGAA GAGAACCAACCTCAAAGCGG | 1 |
| 3216816 | + | bf638r_2740 | Unkown fuction | 0 | CTATCGTCTTTTCTACCGGACAACTCTC TTTATGTGTACTTTTGCCCCCAT | 1 |
| 3217306 | + | bf638r_2741 | rpsp | 0 | AAGAACCGGAAGTCTTTGTCAATCAAA ACAAAACCCTTATCTTTGCACCCG | 1 |
| 3220346 | - | bf638r_2744 | RNA binding protein | 0 | GCGGTAAATATAGACATTTACAACAGA CTGAGAAACAAAATCATCATTTTC | 1/2 |
| 3224938 | - | bf638r_2748 | ABC transport | 0 | GGTGCAAAGGTACACATTTATCCGATT CGTACATACATGTTAATCTAATAT | 1 |
| 3225036 | + | bf638r_2749 | Unkown fuction | 0 | ATAACAGAGAAATTCTATCCGGAATTA AAAAAACACCTATCTTTGCATAA | 1 |
| 3231341 | - | bf638r_2752 | glycogen synthase | 0 | AGGTGCAAAAGTACGCATTATTTATTG AAAAATAATGCTTTATGCAAAAAA | 1 |
| 3239122 | - | bf638r_2759 | ATP synthase | 0 | CGCCAAAAGTACAAAATTCAATCTGAT AATAAAAGAGAAAAGAGCAAGAA | 1 |
| 3245007 | - | | rRNA | 0 | GATGCAAAAGTAAGGCATTAGGAGAT AAGAGACAAATATAAAAGGGGAT | 1 |
| 3247835 | - | bf638r_2760 | NagA | 0 | GAAACAAAATTAGACACTTCTGTTGAG ATAGAAAAATAAAATAGCTTTT | 1 |
| 3247926 | + | bf638r_2761 | responsive regulator | 0 | CTTTTCTTTTCAGTTTTATTTCAGAAAT AAGCCTTTTCTTTGCAGGTAG | 1 |
| 3250209 | + | bf638r_2763 | Unkown fuction | 0 | GGAGAGAATTCCCAATTCTTTTCAGAA TCGACCTTCATCTTTGCATTAT | 1/2 |
| 3253291 | - | bf638r_2766 | Unkown fuction | 0 | TTCGCAAAGATAACAAAAAAAAACGC CCACCAAAAGATGAAGCGTTTTTT | 1 |
| 3254252 | - | bf638r_2767 | Sigma 70 | 0 | AGTGCAAAGTTACGACTATTTATTTGT TAACGTATAAATTTTCACAACTT | 1 |
| 3260170 | - | bf638r_2773 | Unkown fuction | 0 | GCAAAGATAGAAACTCTTTCCATTTTA GATATATATTTTCAGAAACAT | 0 |
| 3260493 | + | bf638r_2774 | ribonulase | 0 | AATCCTCCCTTCCGTTTGCTCAAACAAA ATAATTCTTGTACTTTTGCTGAC | 1 |
| 3263698 | - | bf638r_2776 | ribE | 0 | AGGGGTCAAAGATACAAAGAAAATGG GATTGGATAAAGAATCAACATAAAA | 1 |
| 3264732 | - | bf638r_2777 | phoU | 0 | AGCAAATATACAATATTAAATGATACA TCGAGCTGCATGCCGATCAA | 1 |
| 3267816 | + | bf638r_2781 | Unkown fuction | 0 | **GCGGCACGTTTTGAAACAGGAATGTA ACACGCTATTGCTTTATTTGCACAT** | 1/2 |
| 3268828 | + | bf638r_2782 | glns | 0 | TATACCCGTGAAATAAGGATATACAAA AAAAATATCCTACTTTTGCAGCAA | 1 |
| 3273158 | + | bf638r_2785 | Unkown fuction | 0 | TTTTCGTCCTCCATGCAGTACCTTTATA TTTTCTGCATTTACTTATTGT | 0 |
| 3274559 | - | bf638r_2786 | tpx | 0 | ATATTCTAAACAATCGGTGCAAGAACT TTGTTCACGATGGACAATAATCTA | 0 |
| 3274600 | + | bf638r_2787 | Unkown fuction | 0 | TTATTTACTATATTCTAAACAATCGGTG | 1/2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | CAAGAACTTTGTTCACGATGGAC | |
| 3277147 | - | bf638r_2788 | mpi recombinase | 0 | GCTGCAAAAATAGAGATAATAATTGA AACTCCACAAACAAAATGATAATT | 1 |
| 3303474 | + | bf638r_2816 | glycosyl transferase | 0 | TTTTTAATAAATACTTTCCTCTTTATCG GAAAAAGCATACCTTTGCAAAGA | 1 |
| 3318366 | - | bf638r_2830 | phosphoesterase | 0 | CCACAAAGGTAAAAAAAATATATCCAA AAGCAACGCACCACAAATAAAAGT | 1 |
| 3320356 | + | bf638r_2832 | Unkown fuction | 0 | CTTTCTCCATTTTCTCAACCTTTCAGATT TTCTCTTTATCTTTGTAGTAG | 1 |
| 3341232 | - | bf638r_2852 | signal transduction protein | 0 | GATGGCAAAAATAGAGAGTGACCGCC CCGGTTACAATTTTATATTTTAAA | 1 |
| 3345064 | - | bf638r_2856 | polysaccharide biosynthesis protein | 0 | GGTGCAAAGATACACTATTTTCACTAA CACGTGCGTTTTTATTTGACAA | 1 |
| 3346594 | - | bf638r_2857 | PhoH like | 0 | GCAACTAAGATAACAAATAAAAAGGC AAACAAAGGAAGTTGCCCGTTA | 0 |
| 3346560 | + | bf638r_2858 | folC | 0 | TTCTTTGTTCCCATGATTCTTAGATTTA AATGTTGATATTTGCAATCA | 1/2 |
| 3348458 | - | bf638r_2859 | YjgF | 0 | GTGAACAAAGATAAGAAAAATATGGA AGAGCTATGCAGTATTTGGCAAA | 1 |
| 3348527 | + | | | 1 | TATGGAAGAGCTATGCAGTATTTGGC AAACAAACGCATTTATATATATG | 0 |
| 3356145 | - | bf638r_2865 | DNA pol III | 0 | AGCACAAAGTTACAAAAAAGATCGAG GCAGCACAATAAAAGAGAAGAAAAA | 1 |
| 3356247 | + | bf638r_2866 | Unkown fuction | 0 | ACAGGGGAATATTCGTCGGATATCCCT AAAAAATTGTAACTTCGCCGCAC | 1 |
| 3358377 | - | bf638r_2867 | Unkown fuction | 0 | AGTGCAAAAATAATAATTTCTCCAAAT AAAACAAAATATACTTAAAGAAAG | 1 |
| 3362223 | - | bf638r_2872 | nuleoside-diphosphate epimerase | 0 | GCGCAAAGTAAGCACTTTTTTTTCATTT TTATACCAAATATATGTTGTTTA | 1/2 |
| 3362303 | + | bf638r_2873 | pyridoxal phosphate-dependent acyltransferase | 0 | ATACCAAATATATGTTGTTTATTCCAA GAAAAATTGTGTCCTTTGCAAAAG | 0 |
| 3377844 | + | bf638r_2887 | Unkown fuction | 0 | ACAATCCCCGGGACGATATGCTATACT CTTTTTGTATATCTTTGCCGATC | 1 |
| 3380201 | - | bf638r_2888 | AAA-ATPase | 0 | AAATACAAAGATAACATTTTCTTCTGT ATTTTACGTGTTGTACTACAAAA | 1 |
| 3380971 | - | bf638r_2890 | Unkown fuction | 0 | CCTACAAATATAAGTTTTAAAAAAGAA CTATCCAAGCATACGGCGGAAG | 1 |
| 3382635 | - | bf638r_2891 | ftnA | 0 | ACTACAAAAATAACAATTTTAAAAGAA AGGTTGTTCGGAGAAAGTGAAA | 1 |
| 3386917 | - | bf638r_2896 | hisI | 0 | GGCCGCAAAGATACGATTTTAATTCGG GAATTTTAATTTCTCCGAAGTG | 1 |
| 3389119 | - | bf638r_2899 | hisH | 0 | GTTCGCAAAAATAACGAATTATTGCAA AACTACGAAAAATCAGAGCAGGA | 1 |
| 3389414 | + | | tRNA | 0 | TTTTCGTTCAGCTATTGCAGGTTTCGG AAAAATGCCTACCTTTGCAACCG | 1 |
| 3392221 | - | bf638r_2900 | Unkown fuction | 0 | GATAACAAAAGTATTCTTTTTCCCGATT ATAAGAGAATAAAATAGGAAA | 1 |
| 3392932 | - | | | 1 | TTTTTTGCTCACTCCTGAACAAAATGAT TTCTTCTGCGTTTATAATTATGC | 0 |
| 3394604 | - | bf638r_2903 | regulator | 0 | GGGCAAATATAGTGAGAATCACTCAC ACCTGCTATTTATCCGAATACAAA | 1 |
| 3395156 | - | bf638r_2905 | Unkown fuction | 0 | ACTACAAATATATTGTTTTTTGGATCAT CCCGGCAAAGATAAGGGACGGG | 1 |
| 3415283 | + | | | 1 | ATATGTCATTCCGCTATAAAAACATTTT TTTTTCATACATTTGTACTGA | 1 |
| 3417266 | - | bf638r_2920 | Sigma Factor | 0 | AAGAACAAAGTTAATATTTTCATATAC CGTTAATCATGTTTTATCTTTATT | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3417475 | + | bf638r_2921 | Unkown fuction | 0 | AAAAGTTTTGAAGATGAATTTTATCAG TTTTTATATTTATATTTGCATTGT | 1 |
| 3419233 | - | bf638r_2923 | pyrophosphohydrol ase | 0 | AGAGGCAAAAGTAATAAAAGAATGCG GGAAAGAAAGAAATAAAAAGTAAT | 1 |
| 3419320 | + | bf638r_2924 | Unkown fuction | 0 | AATAAAAAGTAATAAAGACTAATCTTC AATATTTTTCTTACCTTTGTTAAT | 1 |
| 3420457 | | | | 1 | CCTGCAAAGATACATATTTTTTAATCA GGCCCGATTTGCCGGCCGGAAGAT | 1 |
| 3420527 | + | bf638r_2925 | methytransferase | 0 | TAATCAGGCCCGATTTGCCGGCCGGA AGATTCTCCGTATCTTTGCCACAT | 1 |
| 3428730 | - | bf638r_2930 | transporter | 0 | GGGGTGCAAAGATACGGTTTTTTAGA GATCTTGGTTACTTTTGCGTCA | 1 |
| 3428778 | + | | | 1 | GGGGTGCAAAGATACGGTTTTTTAGA GATCTTGGTTACTTTTGCGTCAA | 1 |
| 3431279 | - | bf638r_2932 | ion transporter | 0 | GAAAGGCATATAGCCAATATCGTGCC AAAACGGAAAGAAAAAGTACAAGTA | 0 |
| 3432310 | - | bf638_2934 | HTH regulator | 0 | CTGAGCAAAGATAAGCGATTCTACCG GAATAACGCCCTGCTTTTGGTTA | 0 |
| 3432428 | + | bf638r_2835 | transporter | 0 | ATGATATCGGACGAAATAGTATGCAA AAATAACGGCAACTTTGCAACCGGA | 1/2 |
| 3433714 | + | bf638r_2936 | dehydrogenase | 0 | CTTTCCACATTCCGTACCTTGTCTTATT TTTATTTCTTATCTTTGCCTCAA | 1 |
| 3442137 | - | bf638r_2943 | phosphofructokinas e | 0 | GTCGCAAATTTCGTGTTTTTTTCCGAAT TATGAAAGAAAATGATGATTTT | 1/2 |
| 3445149 | - | bf638r_2945 | coagulation factor/attachment | 0 | ATAAATAAAACAAAGGAAATACAAAT TTTGTTCTCATAAAAGACTTCAC | 1/2 |
| 3453113 | + | bf638r_2951 | Sigma Factor | 0 | CTTGTAGAAATCGCAGCAATTTAGAAC AAAAAGATTACTTTTGCACCTCT | 1 |
| 3459543 | - | bf638r_2954 | efflux transporter | 0 | AGCTACTTTACAAATCAGATACCAAAT TATACCTGAACCGATCTTTTCAAC | 0 |
| 3459804 | + | bf38r_2955 | phospholipase | 0 | ATGGCATTCGCTACTTATTTATCAAATA ATAGTATAAATTTGTAGTAA | 1 |
| 3462245 | + | bf638r_2956 | Unkown fuction | 0 | TTTCAAATTTATTACTCAGAACCAAAG AAAGTCATTATCTTTACCGAA | 0 |
| 3463976 | + | bf638r_2959 | NagA | 0 | TCTCTTGAATTATAGTATATAAATGCA AAAAAGTATCTAATTTTGCCAGAC | 1 |
| 3465412 | + | bf638r_2960 | glucosamine-6-phosphate deaminase-like protein | 0 | AATTTAATTTCCCTGTCAACCTTATCGC ATAGTTGATTGTTTTTATATAT | 0 |
| 3468363 | + | bf638r_2962 | peptidoglycan aminohydrolase | 0 | ATCGGCAATTGCCGGTTTTTCATTTATT TCTTCTACTTTTGCACCCGA | 1 |
| 3472373 | + | bf638r_2965 | Unkown fuction | 0 | ATCCTTGCCACTTACTCCATTCTTCATT TTAAATTCTATCTTTGCTCCTAG | 1 |
| 3473780 | + | bf638r_2969 | Unkown fuction | 0 | TGCTTTATCCGTGCCCTCGTATAGCGA AATTATCCGTATCTTTGTCCCCAT | 1 |
| 3479137 | - | bf638r_2970 | Transcription-repair coupling factor | 0 | ACCGCAAAAGTACTAAATTATCCTTAG TTTTAATGCGCTATCTTTGGTTAG | 1 |
| 3479216 | + | bf638r_2971 | glycosyl transferase | 0 | TTAATGCGCTATCTTTGGTTAGTAGCG GAAAAGCATTACATTTGTCTTCAA | 1 |
| 3482707 | + | bf638r_2976 | tonB like outer membrane protein | 0 | CTGGCAGCAAACTTGCAATAAAAGTTT TCTTTCGATATATTTGTAACTG | 1 |
| 3486848 | - | bf638r_2980 | glutamate cystine ligase | 0 | GTCCGCAAAGTTACAATATTTATCATA AATATAACAAAATAAAACAATT | 1 |
| 3487070 | + | bf638r_2982 | Unkown fuction | 0 | GCCCCCCGGAATAATAAATATTCAATA TTATCTCCCTATATTTACAGACAG | 1/2 |
| 3488259 | - | bf638r_2983 | Unkown fuction | 0 | GGCTTCCAAATTTAATAAATATCCGGA GACGAACCCCAATTACCACCAATT | 1 |
| 3488351 | + | bf638r_2984 | Sigma Factor | 0 | CACCAATTTTCACTCTTATTCTGTGATT ATTTCCTTTTCTTTGCGACTAA | 1/2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3499225 | - | bf638r_2990 | exonulease | 0 | AAACAAAGTTAGCAATTATTCAGGACA GCAATATTATTTTGTTTATTTTT | 0 |
| 3499281 | + | bf638r_2991 | phospholipid acyltransferase | 0 | TTAGCAATTATTCAGGACAGCAATATT ATTTTGTTTATTTTTGCACAAG | 1 |
| 3513159 | - | bf638r_2999 | Unkown fuction | 0 | ACTCGCAAAAATAATACATTTTAAATT ATAAACACATCTTATTCGTAAAAA | 1 |
| 3523831 | - | bf638r_3010 | Unkown fuction | 0 | CAGCAAAGATAAGCGAGTGCATAGAA GCATGCAAACATTCTATGTCAACA | 1 |
| 3524017 | + | bf638r_3011 | responsive regulator | 0 | TTCAACAAATAGGGTGGTAAATTCAAA CTTTTAAAGTAACTTTGCAACAC | 1 |
| 3526034 | - | bf638r_3013 | Unkown fuction | 0 | TTCGTCAAATGTAGCGCATATTTTTCA GACAGCCAAAGAAATTACCCGGA | 1 |
| 3528733 | - | bf638r_3015 | Unkown fuction | 0 | GACGACAAAGGTACAACGGATTCGTA TGCAAATGAAACATTTTATCAAT | 1 |
| 3538875 | + | bf638r_3016 | anaerobic sulfatase-maturase | 0 | ATCTACATAAGGACGGATTTTTCATTT ACTTTACCTAAATTTGCACTTTC | 1 |
| 3536911 | + | bf638r_3022 | thioredoxin | 0 | CGGGACTTCTCTGTTATTATAATAAAA CAAATAACTATATTTGCCCCCTGT | 1 |
| 3538762 | - | bf638r_3024 | nitroreductase family | 0 | AAGCACAAATATAAAAGAATCATCAGT CAACATCTTTATTTCAAAAGAAAA | 1 |
| 3539922 | - | bf638r_3026 | fur | 0 | GATGCAAAGATAAAGAGTTCTTTTATT TTCGCAAACAAAATACGAATATT | 1 |
| 3540452 | - | bf638r_3027 | Unkown fuction | 0 | GCCACAAATATACTAACTACTATTTATA TTTTTTATATTTTCCGCATTTA | 1 |
| 3542726 | - | bf638r_3028 | Zn-dependent oligopeptidases | 0 | TTTGACAAAAATACAAAGATTCGATTA AAATGTCTACTTTTGCAACAAGAA | 1 |
| 3542774 | + | bf638r_3029 | nadE | 0 | ATTTTGACAAAAATACAAAGATTCGAT TAAAATGTCTACTTTTGCAACAAG | 1 |
| 3544882 | - | | | 1 | GACAATAAGAACAAAACAGAACTCAA TTCTGTCGTCAAAAAGACATGGGA | 0 |
| 3555788 | - | bf638r_3037 | hisG | 0 | GCGACAAAAATAAAGTATTAGTTCTGA ATCACCAAACAATTATCAGAAA | 1 |
| 3556364 | - | bf638r_3038 | thioesterase | 0 | GCTCAAAGATATAAATCTTTTCGCAAA ATGCACCCATCTCACCACCAGAA | 1 |
| 3556445 | + | bf638r_3039 | pirin | 0 | CCATCTCACCACCAGAACTTTATCTATA ATTCGTTTGTTATACAGATGT | 1/2 |
| 3559725 | - | bf638r_3041 | stress responsive protein | 0 | AGTTGCAAAATTACTCTATTTGTTTCAA ATATACACTATAAAAGGATGTTA | 1 |
| 3559709 | + | bf638r_3042 | udk | 0 | GTTTTACCATTACGTTGTTATTTTAATT ATTTGTTTTTATATTTGTCAGCC | 1 |
| 3563533 | - | bf638r_3044 | symporter | 0 | TGGTGCAAATGTAAGGAATTTATTGAA ATAGAGAAACAAAATGCAGAAAT | 1 |
| 3567122 | - | bf638r_3046 | metH | 0 | GCTTTTTACCTTTAGCCAAAGCTACAA CCACTTTGGCCAAACCTCTTTCAT | 1 |
| 3568907 | - | bf638r_3049 | Unkown fuction | 0 | GTTTGCAAAGATATAAATTCTAATTCA CAATAAGTAAATGCAGAGACTT | 1 |
| 3569517 | - | bf638r_3050 | Unkown fuction | 0 | AGGAAACAAAAATACGGAAATATATG GTTCAGAAAGGCTTTTTGATAAG | 1 |
| 3570482 | - | bf638r_3051 | Unkown fuction | 0 | TCAGAAGAAACAAAGAATCAGAGTAA TAGTTCAACAAAACAAAGATTATC | 1 |
| 3570601 | + | bf638r_3052 | ion-sulfer cluster scafford like protein | 0 | TAATAAAACCCGATTTGCAAAAACAAA TAAAACGACTACCTTTGTGAGCGA | 1 |
| 3587164 | + | bf638r_3070 | Unkown fuction | 0 | AACAACCAAAGCATTGCTATTAAAAGA GGAATGATTACCTTTGTCTTAA | 1 |
| 3588559 | + | bf638r_3073 | Ketoacyl-acyl carrier protein synthase III | 0 | AGATTACAAGATATTGAATAAATCTTT CTTTTCTGTATATTTGCGCCCGG | 1 |
| 3589749 | + | bf638r_3074 | alpha amylase | 0 | TTTGGGTAAACACAAGAAACCTTATCA TACACTTCTTTGTTATTTTCCGAT | 0 |
| 3592516 | - | bf638r_3076 | Unkown fuction | 0 | TTATTTAAAACAATAATTGAGTATAAT AGTTCAGGAAACAGAATCTCCA | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3593756 | - | bf638r_3077 | Unkown fuction | 0 | GCAACAAAGGTACACTTTTATGAAGAA AAAAAGAGTAAACGAAGACAAA | 1 |
| 3593785 | + | bf638r_3078 | Rho | 0 | AAGAGTAAACGAAGACAAAAACCAAA ATTATTTATTATCTTTGCGGCGT | 1 |
| 3596101 | + | bf638r_3079 | histidine kinase | 0 | AGTGAGATTCTGTTTGATTCTCACTCTT TTTATTTAGCTTTGCACTAC | 1 |
| 3600007 | + | bf638r_3081 | Na+-driven multidrug efflux pump | 0 | TCCCCTATCCGCTCGTCTACTTCTACAG AAAGTTGTATCTTTGCAACCTG | 1 |
| 3601382 | + | bf638r_3082 | ffh | 0 | AGAACAAAAAGATATAAAGTACGAGA GATTTTTGTACATTTGCGGCTGA | 1 |
| 3603206 | + | bf638r_3083 | lpA-like family | 0 | TACTCTTGACAAAAAACAGGGGATGG TAAATGATATTACTTTTGCCCGATA | 1 |
| 3608099 | + | bf638r_ | tRNA | 0 | AGTTCGCAGAATGTTTGCAGATTCAAA AAGAATCGTTACCTTTGCATCGC | 1 |
| 3617859 | - | bf638r_3094 | porin | 0 | TAGCTCGCAAAAGTATAAGATATCCGC AGCAGAGGCAATACCAAAAATTA | 1 |
| 3619219 | - | bf638r_3096 | Unkown fuction | 0 | AGCTATTAGACGTTACAACAACACAAA ACGTTGCATGGCCGATCCTCTTTT | 0 |
| 3621169 | - | bf638r_3097 | DNA pol III | 0 | TTGTGCAAATGTACGCAAAAAAGTGC AAGGATAAAAAAGGAAGCGAAATAA | 1 |
| 3621273 | + | bf638r_3098 | Septum formation initiator | 0 | TAGAGCGGAAATGATACGTTTTAAGTT ATTAGTTTTATCTTTGCAACTGG | 1 |
| 3626394 | - | bf638r_3104 | Unkown fuction | 0 | GGGTGCAAATATAGACTTTTGTACTGA CTAAGCCTCCTTTTTTACCATA | 1 |
| 3626399 | + | bf638r_3105 | pepD, aminoacyl histidine peptidase | 0 | AACAGAGTTATCAGTATTGTATAGAGC ATTTCTATTAATTTTGTGGGTGC | 1 |
| 3630415 | - | bf638r_3107 | ion channel | 0 | ATCCAAATAACACCTAAACAGGCTTTA GGTTTAGTCCAAATAAAAAAGG | 0 |
| 3640227 | - | bf638r_3112 | sulfatase | 0 | GGTGCAAATATAAAGAGTTTTAACAGT CCGTCGAGTTCAAAAACAACTA | 1 |
| 3640403 | + | bf638r_3113 | sulfatase | 0 | AGCCTCTAACTAACAAGTCTTAGCTGA AATTTGTATATTTGCAGACAT | 1 |
| 3643936 | - | bf638r_3114 | Glycosyl hydrolase | 0 | AGTGCAAAGATAGGATTTTATTTCCAT TTTTTTGGTCACGTCTATGCAA | 1 |
| 3644004 | + | bf638r_3115 | periplasmic ligand-binding sensor | 0 | TTCCATTTTTTTGGTCACGTCTATGCAA ACTGTTACTTTTGCCGGTAA | 1 |
| 3649006 | - | bf638r_3117 | gpmA | 0 | ACTATAGAACATTTGTACTGAGTACAA AGTTCACGAAACTTTTTCTAAAT | 0 |
| 3649099 | + | bf638r_3118 | Unkown fuction | 0 | CTAAATATCTATGCTGAAGTATGCTGA TAAAAGTTACCTTTGTACACTGT | 1 |
| 3653095 | - | bf638r_3119 | ABC transport | 0 | AGGGCAAAATTATACCAAAATCCGCAC AAAGAGACGGAATAATGCTAAAA | 1 |
| 3656331 | - | bf638r3121 | Glycosyl hydrolase | 0 | GTCCAAAGATAAGATTTTCTCCGGACA GTTTTTTATCACACGGTTAAAA | 1 |
| 3661732 | + | bf638r_3124 | FecR | 1 | TAATCAATTAATACTGTTTAAAGTCTCT AAAGGTTAAAAAGTTTTTGATTG | 1 |
| 3666695 | - | bf638r_3128 | Unkown fuction | 0 | TTCCGCAAACATAGCAATTATTTCAGA ATAATACATTATATTTGCAGC | 1 |
| 3666744 | + | | tRNA | 0 | TTCCGCAAACATAGCAATTATTTCAGA ATAATACATTATATTTGCAGCAC | 1 |
| 3668975 | + | bf638r_3130 | phenyalanyl tRNA synthetase subunit | 0 | TATTTCTTTCTATCCTCTTTGATTGCAAT ATTATTGTTACTTTTGCAGCAA | 1 |
| 3672059 | + | bf638r_3133 | pgk | 0 | ATAACGATTTCATAGCAAAAAAGAGA AATTAATACTAACTTTGCGGTCAT | 1 |
| 3680816 | - | bf638r_3140 | laminarinase | 0 | TAATAAGCAACAAACCGTATATTTATT TAGTTCATCGTCCCCCATTGTT | 0 |
| 3681092 | + | bf638r_3141 | Acyl-protein synthetase, LuxE | 0 | TAATCGATTGTGCGTAGGGAGATCAG ACCATTTCACTGTCATCTTAATAA | 0 |
| 3689027 | - | bf638r_3148 | NAD(P)H:flavin | 0 | GAGCGCAAAGATAGGCAATTTATTCCA | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | oxidoreductase-like | | CTCTCGTCAACTCGCCGGTCAGT | |
| 3689639 | - | bf638r_3149 | Carbonic anhydrase | 0 | GACAGCTCAACAAAATAAGCCCTGCA AATGTTGTAAAGAGAAATAAAAAA | 1/2 |
| 3689705 | + | bf638r_3150 | Unkown fuction | 0 | GCCCTGCAAATGTTGTAAAGAGAAAT AAAAAATGTAAATTTGTCCCGAT | 1 |
| 3690871 | + | bf638r_3151 | Methylmalonyl-CoA epimerase | 0 | AATTGCAGAAAAATTTCTTTATTAAAC AAGAATTGATTTTCTTTGTGCTGT | 1/2 |
| 3698825 | + | bf638r_3158 | amylase | 0 | AAGAAAGAAAACGGAAAACAATCCGA TGATTTACCCTATATTAGCAACAT | 1/2 |
| 3700839 | + | bf638r_3159 | Transcriptional regulator | 0 | AATAGTCAAACGTTTGACTATTATCGG AGAATATTCCTACCTTTGCAGCGA | 1/2 |
| 3705113 | - | bf638r_3162 | fructose-bisphosphate aldolase | 0 | TATGCAAAAGTAAGAATAATCCTTTGA ATAGAGAAACCCTCGGGCGGAAA | 1 |
| 3706387 | - | bf638r_3163 | Endonuclease/Exonuclease/phosphatase | 0 | CCCTGCAAAAGTCTTAAAAAAAAGCG AGAATACGACTTTTTTTCCAAAAC | 1/2 |
| 3706476 | + | bf638r_3164 | rpmE | 0 | TTTCCAAAACTTTTTTCCTAAGAAACAA AAAAGCATTATCTTTGCAACCT | 1 |
| 3708371 | + | bf638r_3166 | Peptidase | 0 | ATAGGAGAAAAAACAGCGCTGTAGGA AACTTCTTTCGTATCTTCGTAACAT | 0 |
| 3719369 | - | bf638r_3170 | OsuA | 0 | GTAGCAAAAATACTATTTAAACAACAG TCGCTTATTTAAATTCAAACAAA | 1 |
| 3719517 | + | bf638r_3171 | OsuR | 0 | AACATACTATCACTGAACTTTTCAGTCT CCTTTACACTATTTTTGTGACAT | 1 |
| 3722267 | + | bf638r_3173 | maltose phosphorylase | 0 | ATCTGTTATCCACTGTCGGCACGGACT TTTTTTGTATCTTGGGTGCCGT | 1/2 |
| 3724800 | + | bf638r_3174 | efflux transporter | 0 | ACAAACCGGAAAACTGTTTTTGGTTTT TTAGTTTTATATCTTTGCGCCCTC | 1 |
| 3726055 | + | bf638r_3176 | efflux transporter | 0 | CTGAAATAAGAGTGCCCGTTCAACGCA ATAATTAATTTCACCACAGATTA | 0 |
| 3732184 | + | bf638r_3179 | phosphoribosylformylglycinamidine synthase | 0 | TATAGTATTTCCTTCTCACCACCAACGA ATTTGCTTATCTTTGCACGCAA | 1 |
| 3736041 | + | bf638r_3180 | Unkown fuction | 0 | TCCTTTCCTAAAGTTTGCACTTCTCCAC TCTTTTTATTAATTTTGTAAAAT | 1 |
| 3742924 | - | bf636r_3183 | Unkown fuction | 0 | GGTGCAAAGATACAATTTTGCTCCTCA TTTGCAACGGTTAAGCAAGCAATT | 1 |
| 3743054 | + | bf638r_3184 | uvrA2 | 0 | GTTCCTTCCACTGTTAGGTTAAACGTT GCTGACTCACTATCTTTGCAGTCC | 1 |
| 3748206 | - | bf638r_3187 | Carbon starvation protein CstA | 0 | AGCGCAAATGTAACATAATCTAAGGA AACTACACAACTTTACGAAGGGTTT | 1 |
| 3750018 | + | bf638r_3189 | Histidine kinase-like ATPases | 0 | AAATAACCATCTTTTCTTATTTATATCA TTAATATAGTATCTTTGTGAAAT | 1 |
| 3753750 | - | bf638r_3190 | Unkown fuction | 0 | CTGCAAATATATGCACAAAAACGCATA TAAAACCGGGAAACAGGTCACTT | 1 |
| 3754057 | + | bf638r_3191 | ATPase | 0 | TACCGGCAGAAACTACCGAAAACAAA AAACAATATTTATATTTGCAAGCAA | 1 |
| 3764091 | - | bf638r_3198 | opuAA | 0 | AATTAAGTAAACAAGCCGAACGTTCAA ATGGTTCATCGCATTATTTCCGG | 0 |
| 3767308 | - | bf638r_3200 | sensor kinase | 0 | GGAGGCAAATATAACAATTTCCCTCCG GAATCTCCCACCTGCTTGGTTTA | 1 |
| 3771513 | - | bf638r_3203 | Glycosyl hydrolases family | 0 | GGTGGCAAATATAGCGGAGATATCTG CCACCGGCTATTAATTTTGTTTC | 1 |
| 3773920 | - | bf638r_3204 | outer membrane receptor protein | 0 | GCCGCAAAGGAACGCGGAATCCCATT TCCCAACTAAAAGATTTGTTTCAAA | 1/2 |
| 3777198 | - | bf638r_3207 | Unkown fuction | 0 | GAAGCAAATGTAGATATTATAATTATA CTAAGCAAGAGTTAAGAGAAAAAA | 1 |
| 3778263 | + | bf638r_3208 | glucose/galactose transporter | 0 | CATAGCCTTTTCAACCTAAATTATATAT TGGTCTAACTTTGTTTATGT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3781123 | - | bf638r_3210 | glutamine amidotransferase | 0 | GATATAGAAACGTACCTACAACCGACT TGTTATCTTGCTAACCGGAGTTT | 1/2 |
| 3784358 | - | bf638r_3212 | ion transporter | 0 | AAAGATAAAACAAACCAAAATAAAAA ATAGTTTTAGAGTAAGTTTAAATC | 1/2 |
| 3794891 | - | bf638r_3220 | efflux transporter | 0 | ATATGCAAAGTTATTGGGAAGGGAGG AGATAGAAAGGTTGAAATGAATGA | 1 |
| 3794890 | + | bf638r_3221 | regulator | 0 | CCTGTCAGAATAAGATAAATCTGTTTC ATAATTTGCGTACTTTTGTATTTC | 1 |
| 3795981 | + | bf638r_3222 | Peptidase family | 0 | GATAGACGTTCATTTCCATATTTGGGG ATATTCAGCTATCTTTGTCGCAT | 1 |
| 3799443 | + | bf638r_3225 | Unkown fuction | 0 | GAATCAACGAACAGGGGGATTTTATC AGTAATTATCTATCTTTGCAGCCC | 1 |
| 3800599 | - | | | 1 | GTTGCAAATATAGCGATTATTTTAAAA ATAGATAATAATCCCAAAAGTTT | 1 |
| 3805958 | + | bf638r_3228 | Alpha-L-fucosidase | 0 | TTAGTTAGTCCCCGGAACGAGTTTCCG GGGATTTTTTTATCTTTGCACCAT | 1 |
| 3808169 | + | bf638r_3229 | electron transfer flavoprotein | 0 | TCAGATAGATTTCTTGTCAGAACCAAG CGTTTGCTTAAATTTGCAGGCTG | 1 |
| 3812026 | + | bf638r_3232 | glycosy hydrolase | 0 | ATGTCAAATGATTTGGCTATGTGGCTG GTTTTTACGAATATTGCAACAA | 0 |
| 3817577 | - | bf638r_3234 | CbpA | 0 | ATACAAACAACACGCCAACCTGAGTTT TAGATTCAAATTTTAGCACAAAT | 0 |
| 3819498 | - | bf638r_3235 | NtrY | 0 | GACAGTAAATATAGATATTATTCTCCT ACTTTCAACACAAAACACAACAAA | 1/2 |
| 3822775 | - | bf638r_3237 | surface antigen | 0 | CTCCATTTAACACTTAATAAACAGAAA ATGTTCAAAATCCAAGGTATGA | 0 |
| 3822928 | + | | | 1 | AAGCTCCCTGCTATTCCTTAAGTAAATT AAATCCGCTTTATTTGTATCAT | 1/2 |
| 3823344 | - | | | 1 | TCCGGACAAAGATAAAGTAACGTTCC GTAAGGTTTTTACCGACAAATCTTT | 1 |
| 3823407 | + | bf638r_3238 | Unkown fuction | 0 | AAGTAACGTTCCGTAAGGTTTTTACCG ACAAATCTTTATCTTTACAGTCT | 1/2 |
| 3826794 | - | bf638r_3240 | Beta-eliminating lyase | 0 | TCTGCAAAGGAAGAAAATTATTCGCA GAAAAAGAAAGGATTTCACCACAG | 1/2 |
| 3827875 | - | bf638r_3242 | Sigma factor | 0 | AGTGCAAAGAACGAAATTATTCACTTA AAAACATACGATCATGATGGATT | 1/2 |
| 3827881 | + | bf638r_3243 | Unkown fuction | 0 | TATTTTTTTCTTTGTAACCTTTCCGCAG CCTGCTGCGTCAAACAGTGCAA | 0 |
| 3828731 | - | bf638r_3244 | Unkown fuction | 0 | GCGGTGAAAGTTAACAAAATTATGCA CGCATTGTTATTACAAAGACACTAA | 1/2 |
| 3828898 | + | bf638r_3245 | poxB | 0 | GCATACGTTAAATATACAACTTTTTATC ACAACTACCTGTTATTGCAACAT | 0 |
| 3831486 | - | bf638r_3246 | NADPH-dependent FMN reductase | 0 | AATACAAATATAATCAGTTAAATATGC GATGGTTAAAGTAATATTGTTTCC | 1 |
| 3836168 | - | bf638r_3251 | groES | 0 | GCTGCAAATGTATGGATTCTTTTTAAT GTAACCAAAATATTTCATTTCTTT | 1 |
| 3836387 | + | bf638r_3252 | Unkown fuction | 0 | TATTCATTCCTAATTAAAGATACGGTTT CAGGTCAGAATCATTTGCATAG | 0 |
| 3846074 | - | bf638r_3259 | Polyphosphate kinase 2 | 0 | GTAATAAAGATACAAAAAATCGGGCA GAAAAGTTTCATCCGATTAGCTAA | 1/2 |
| 3848557 | + | | | 1 | TTTCCTACACTTCTTTTGAAGTAACACG GAAATTGTGTATCTTTGTTCTGA | 1 |
| 3848580 | - | bf638r_3262 | integrase | 0 | AATGCAAAGTTAACTAATCATGAGGA GATTTCGGTCTCCTCTTTCTGTTT | 1 |
| 3851660 | - | bf638r_3264 | glycosyl transferase | 0 | TTGCAAAGAAACAGCTTACGTATTACA ATTGTTTGACGGAAGTATGACGA | 1/2 |
| 3851838 | + | bf638r_3265 | glycosyl hydrolase | 0 | CTTCCTTTATCCTGTGCGATACTGTCAG TTTTTTCCTTATTTTTGCCCGAA | 1 |
| 3851920 | + | bf638r_3265 | glycosyl hydrolase | 0 | AATGAACGAGTTCATTTTATTCTGCCCT CAGTTTGCTTATCTTTGTCACAA | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3856930 | + | bf638r_3268 | transposon | 0 | AACAGAAAGAGGAGACCGAAATCTCC TCATGATTAGTTAACTTTGCATTGC | 1 |
| 3856953 | - | | | 1 | AGAACAAAGATACACAATTTCCGTGTT ACTTCAAAAGAAGTGTAGGAAAA | 1 |
| 3859531 | - | bf638r_3269 | histidyl-tRNA synthetase | 0 | GTTGCAAAGTTATATGAAATAATTGAT TATTGAAACAGAAACCACAGACTT | 1 |
| 3862393 | - | bf638r_3272 | Unkown fuction | 0 | ATTTACAAAGATATATATAAAAAGAGA CTATCAACATTTAGAACATAAAA | 1 |
| 3863253 | - | bf638r_3273 | Putative neutral zinc metallopeptidase | 0 | TTGCCAAAAATAAACAAAAGCGACGT ACAGGGAGTTTACCGTACATCGCT | 1 |
| 3866541 | - | bf638r_3276 | Fur | 0 | GCCAACAAAGTTAATTGTTTTTTGCGG AAAGCAAAAATATATCTGTAAATT | 1 |
| 3866600 | + | bf638r_3277 | Dolichyl-phosphate-mannose-protein mannosyltransferase | 0 | GTTAATTGTTTTTTGCGGAAAGCAAAA ATATATCTGTAAATTTGCAGCCAA | 1 |
| 3871765 | - | bf638r_3281 | AraC-type DNA-binding | 0 | GTCACAAATTTCATACATTATTTTGATA TACACAACTTTTTCGTATATAT | 1/2 |
| 3871871 | + | bf638r_3282 | recQ | 0 | AAGAATATTTAACTTTTAATAACCCGA AAATAAAGCGTACCTTTGCCTGAA | 1 |
| 3875996 | - | bf638r_3285 | cysteine synthase | 0 | GCGCAAAGATAAAGAAAAGAATCGTA CCTTTGCGACCGTTCAATCTTTTTA | 1 |
| 3876034 | + | bf638r_3286 | Phosphoglycerol transferase | 0 | ACCACAGTTAATGCGCAAAGATAAAG AAAAGAATCGTACCTTTGCGACCGT | 1 |
| 3878761 | + | bf638r_3288 | TonB-dependent heme/hemoglobin receptor | 0 | CCTACATATAGGTATTGTACACTACCG CATAAAAACCGATATTTGCCGGAC | 1/2 |
| 3887908 | + | | | 1 | AAGATTCGTATTTTGCATTTTATTAAAT CTTATTCTTATCTTTGCCGCAT | 1 |
| 3889361 | - | bf638r_3294 | Nucleoside-diphosphate-sugar epimerase | 0 | TTACCATTAACAAACTCTTCCAGTTTCT TGTTCGAACGTATTTTTTTGTTC | 0 |
| 3892265 | - | bf638r_3297 | Glycosyl transferase | 0 | CCCGCAAAAGTACAAAAAATATGAAA TAAAATAGGATTATTGAAACAACAA | 1 |
| 3898940 | - | bf638r_3302 | leuA | 0 | GATGCAAAGGTAGAGATTATTTCTTAG CCTCCAAATAATACGTTACTTTTT | 1 |
| 3899399 | - | bf638r_3303 | Unkown fuction | 0 | GATTGCAAATATAAATAAAAATAAGTA ATCTACCAAAAACAAGTCATCCCC | 1 |
| 3905042 | - | | tRNA | 0 | GGATGCAAAAGTAAGGCATTAGGAGA TAAGAGACAAATATAAAAGGGGATT | 1 |
| 3905953 | + | bf638r_3305 | ferritin | 0 | AAAATCACCCTTTCAGAAAACAAAATC CCCTACCCCATGTTACATAACCAA | 0 |
| 3906703 | + | bf638r_3306 | sspA | 0 | TATGCGACTTATGGGATGAATTTCAAC TCAATTATTTATCTTTGTTAGCTT | 1 |
| 3912751 | + | bf638r_3312 | Transcriptional regulator | 0 | ACCAAGCATATTATTTGCATGTTTCCG ACAAGTAACATATATTTGTTCCCG | 1 |
| 3918867 | + | bf638r_3322 | HTH protein | 0 | AAAAATAGTAAAAATTGCACAACTAAT AATTAATCAATATATTTGCAATAT | 1 |
| 3924581 | - | bf638r_3326 | Nitroreductase family | 0 | AAAGCAAAGTTATACCATTTCCAGATC TATTATTATGTATTACTTTTCGTT | 1 |
| 3924672 | + | bf638r_3327 | HTH protein | 0 | ACTTTTCGTTATATCGGTAAAATTACGT TTTTTAGACTATATTTGTCTTAT | 1 |
| 3929635 | - | bf638r_3328 | periplasmic ligand-binding sensor | 0 | TTCACAAAGATAATACATGCCGGACAA ATAATTCTCAAGTGCATGATAAA | 1 |
| 3931805 | - | bf638r_3329 | Glycoside hydrolase | 0 | ATTTGCAAAATTAGCACATAAATAGAA CATTCTGTTCTGCATTTGTGACAT | 1 |
| 3931948 | + | bf638r_3330 | Unkown fuction | 0 | GTCATGTATTTGGGTGCGTTTTTAAGT ATAGATTTCTATCTTTGTACAAA | 1 |
| 3943253 | - | bf638r_3335 | Glycosyl hydrolase | 0 | CCTGCAAAAGTATAAGATACACTGAAA GGTGTTTTATAAATCCATGACATA | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 3943669 | + | bf638r_3336 | Acetyltransferase | 0 | AAAAATGAATAAATATCACCCATTGGA AATAAAATGCCATATTTGCACAAT | 1/2 |
| 3945766 | - | bf638r_3337 | efflux transporter | 0 | GGAACAAAAGTATAAAAAAGAAGCGG ATAAACAATTTTCTCCCTCAAACGT | 1 |
| 3945827 | + | bf638r_3338 | cupin-like protein | 0 | TAAAAAAGAAGCGGATAAACAATTTTC TCCCTCAAACGTTTTATTCTCAC | 0 |
| 3946452 | - | | | 1 | ATGTAAAGGTAGTCATTTTAGGTCAAG ATCCATATCCTAATCCGGGACAAT | 1/2 |
| 3948124 | + | bf638r_3342 | Unkown fuction | 0 | ATCGTAATTCATAGTAATAATTATATTA ATAACACTAGTTTTGTTATCAC | 1 |
| 3950211 | + | bf638r_3343 | Unkown fuction | 0 | TTCTTATCATCCGAATAGGGAATCTTCT TATTTTTTCTGTCTTTATAATCAA | 0 |
| 3952547 | - | bf638r_3346 | Unkown fuction | 0 | CCACAAAGATATATTTTTTTTAGATAA GTTAAAATTAATCTGATCTTTTAG | 1 |
| 3953900 | - | bf638r_3348 | Unkown fuction | 0 | GCACTAATAACAATATAAAAATCATTT TGCTTTCAAAATCATCTATAAATT | 0 |
| 3954914 | - | | | 0 | ATTGCAAAGATACGGACTTTTATCAAC TTTGCAAATTTACCAAGCGAAAC | 1 |
| 3954752 | + | bf638r_3349 | hdhA | 0 | AGAGGGTTTATTCGAATAACAAATTCT GAAACAGACTGTTTATGTAAGTAT | 0 |
| 3970745 | - | bf638r_3360 | phage integrase | 0 | GTTGCAAATAAAAGATTCCTCCAGCAA TTGCGCAAAAGATCGATACCAAAT | 1/2 |
| 3972026 | - | bf638r_3361 | HTH regulator | 0 | AAGTTTTAAAACTAATAAAAAGATAGA AAAGCACCAACTTATAAGATCTTT | 1/2 |
| 3972094 | + | | | 1 | AAAAGATAGAAAAGCACCAACTTATA AGATCTTTTTTTGTTTTTAATATCA | 1/2 |
| 3972214 | + | bf638r_3362 | pncB | 0 | ATGAACCGATACGGAATGGATTACTAT AGAATATTTTGTATCTTTGACGAG | 1 |
| 3974658 | - | bf638r_3363 | TlpA-like family protein | 0 | ACACGCAAATATAACCAATGAAAACA GAAATCGACAAGCGGTATCCCAACT | 1 |
| 3977707 | - | bf638r_3364 | Unkown fuction | 0 | ACCCAAAGAACAATCGGTAGTCCGTTT GTGTTCCGTATATATGACATTTAT | 0 |
| 3977748 | + | | | 1 | TCGTTTATTGTACCCAAAGAACAATCG GTAGTCCGTTTGTGTTCCGTATAT | 1/2 |
| 3977850 | + | bf638r_3365 | acetyl-CoA carboxylase biotin carboxylase subunit | 0 | TCGTGCAAAAAGTATCCATCCTATTCA ATATTTCTCTATATTTGTCGCCGA | 1 |
| 3982904 | + | bf638r_3369 | beta-lactamase | 0 | TGTGATATCGCTATTTCCCGTGAACTA AAAGCCGCCTATCTTTGCAACAT | 1 |
| 3986462 | - | bf638r_3370 | impDH | 0 | CATGCAAAATTACATCTTTCTTTTTAAT TATGTGACTGGGGAAGGGAGATT | 1 |
| 3986982 | - | bf638r_3372 | Unkown fuction | 0 | CCCGCAAAGGTAACACAAATCGGCCT ACGTTAGGCTTTTCTTATAAATTTA | 1 |
| 3987118 | + | bf638r_3373 | Unkown fuction | 0 | AAACAACACTACTTGGCAAACACTTTC ATACCTCCGGTGTTATACATCAAC | 0 |
| 3988311 | - | bf638r_3374 | Unkown fuction | 0 | CTACAAAAGTAGGTCTTTTTAGTGAAT TAGCCCGCATACTGTTCCGATAAT | 1 |
| 3992184 | - | | | 1 | GAGAATCAGCAAAGCGGTACGACGGC CATCATCTTCGGCATCTGTTTTCTG | 0 |
| 3995708 | - | bf638r_3378 | Cys/Met metabolism PLP-dependent enzy | 0 | GACGCAAAAGTAGAGAAAATTATTGG TTTAAAAGGCTATCTTTGCGAGCGA | 1 |
| 3995758 | + | bf638r_3379 | N-acetylmuramoyl-L-alanine amidase | 0 | GACGCAAAAGTAGAGAAAATTATTGG TTTAAAAGGCTATCTTTGCGAGCGA | 1 |
| 3996968 | + | bf638r_3380 | Homocysteine S-methyltransferase | 0 | TACTGAGAGCAGATGACCGGAGTGAG CCGATAAATCCTATATTTGTTGCCG | 1 |
| 3999019 | - | bf638r_3381 | Unkown fuction | 0 | CCGCAAAGATAGTCATTTATCAGGAG GCATCCCTCCCGAATGGCTACCGGA | 1 |
| 3999489 | + | bf638r_3383 | Unkown fuction | 0 | GGTAGGCGCACGAAGAAAGATTCCCG GATTATTTCCTACTTTTGTACCCAT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4004632 | - | bf638r_3387 | haloacid dehalogenase-like hydrolase | 0 | TCCTGCAAAAATACGCTTATAAAATGA AAGCAGCCTGCATCCGGTAAATA | 1 |
| 4004740 | + | bf638r_3388 | Unkown fuction | 0 | TTCTTTTCCGGATTCAAGAACCAACAC ATACGGCTTTTGTTATTACCGCAT | 0 |
| 4007483 | - | bf638r_3390 | his kinase | 0 | TGAAGTACAACAATCGGGAAAAGAAA AGGTTCGGTAGCGGTCCTCTACTTT | 0 |
| 4015476 | - | bf638r_3396 | Unkown fuction | 0 | ATATAATTAACGTATGAACATCCTTTTT GGTTTAATCCTTTTGGTTATCTT | 0 |
| 4007695 | - | | | 1 | CTTCTATCCTACGCTGTGACAGCGTTA CTCATCCCGATATTTGCAGCAGCT | 0 |
| 4015534 | + | bf638r_3397 | transglycosylase | 0 | AACGTATGAACATCCTTTTTGGTTTAAT CCTTTTGGTTATCTTTACCCCAC | 1/2 |
| 4023430 | - | bf638r_3399 | helicase | 0 | GTGGCAAAATTACATATTTATTTATAA AGGAGACCTCTTTAATTTAAAGAA | 1 |
| 4023966 | - | bf638r_3400 | Unkown fuction | 0 | ACTGCGAATATACGTTATTTCTAAAAA ACTATTGTATCTTTAGGTAATGAA | 1/2 |
| 4024012 | + | bf638r_3401 | MgtC | 0 | TTTGACTGCGAATATACGTTATTTCTAA AAAACTATTGTATCTTTAGGTAA | 1/2 |
| 4027708 | - | bf638r_3404 | Transcriptional regulator | 0 | GCCAAAAATAAGCACAATCCTCCAAAA GATCAAAAGATAATCTCTTTTTAA | 1 |
| 4027823 | + | bf638r_3405 | peptide methionine reductase | 0 | GATACAACAGTTCTTAATAATCCGAAA ACGATCTTCGTTTGTTATTTATAA | 1/2 |
| 4027869 | + | bf638r_3405 | peptide methionine reductase | 0 | TATAAAAGAGACTGAAAATTATGAAG AAGAAGTTTACTATATTTGTCGTGC | 1 |
| 4029296 | - | | | 1 | GGGCAAAAATACATCCTTTCCCGAAAT AAACGCATAAAAACGATATTTTTT | 1 |
| 4029414 | + | bf638r_3406 | Unkown fuction | 0 | AAATCAGCAAAAAAATAAGGTTTAGTC TTTTTTATCTACTTTTGCAGGCAT | 1 |
| 4031545 | + | bf638r_3408 | Unkown fuction | 0 | TCCGTGTAATCTGTAGTAATTTAACAA CTTATTTAGTAAATTTGCAACCTC | 1 |
| 4032306 | + | bf638r_3409 | radical SAM family | 0 | ATAGCAAAGCGGAGGTTTTCTACCTCC GCTTTTTTTGTTATATTTATGCAA | 1/2 |
| 4036590 | - | bf638r_3411 | Unkown fuction | 0 | AAGGCGCAAAAGTAATTAAAAATCGT GGGAGTACGTTTCTATCAAATAGAT | 1 |
| 4036663 | + | bf638r_3412 | UDG | 0 | CGTGGGAGTACGTTTCTATCAAATAGA TATTATCACTATTTTTGCTCCCGA | 1 |
| 4037685 | + | | tRNA | 0 | CTTTGCATAAAAGCTTGCAAATTCAGA AGTTTATTCTACCTTTGCAGCCGC | 1 |
| 4041866 | - | bf638r_3418 | CDP-alcohol phosphatidyltransferase | 0 | TCGCAAATATACACATCATCTGATGAG TTGCAAAGGAATGAGTGAGAATTT | 1 |
| 4048448 | - | bf638r_3424 | Unkown fuction | 0 | ATTACAAGAACAAGGCTGAAGGAAAA AATGTTCGGCGGGAAGTTGTTCATA | 0 |
| 4049274 | - | bf638r_3426 | 6-pyruvoyl tetrahydropterin synthase | 0 | GGGCACAAAGATAATACAAAAGATAG TTTTCATTCAGCGAGAAAAGTTAT | 1 |
| 4049359 | - | | | 1 | TTATATTAGGTACCATTTCGCTGGGGC TCGGCATTCTTGGCATTTTTCTG | 0 |
| 4049343 | + | | | 1 | AAGATAGTTTTCATTCAGCGAGAAAAG TTATTTTTCCTATTTTTGCCGAAT | 1 |
| 4050271 | - | bf638r_3428 | phosphorylase | 0 | GCTGCAAAAGTAATCAATTAAAGCCAA ATTGACCAGATATTATCTTTAAT | 1 |
| 4055339 | - | bf638r_3434 | Unkown fuction | 0 | ACACAAAAATACGATTAACTCTGATAT TACAAAAACTTTTCGGCGGAAAAT | 1 |
| 4055632 | + | bf638r_3435 | maeB | 0 | TATATCAGATAGACTACGATTTCGAAC TTTTATTCCTACATTTGCGGGCAA | 1 |
| 4058231 | + | bf638r_3436 | gdhB2 | 0 | TTCTTTATAAAATGCTTGCTAGTTCGAT TTTTATGTTTACTTTTGCAACGT | 1 |
| 4062016 | - | | | 1 | GTTCAAATCTATGAATTTTTATCGATTA AAGTCTATTTTTTTTATAATTTT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4070909 | + | bf638r_3445 | HTH regulator | 0 | TAATATTCATTTTATTCATCTTATTCCAT TTTTTCGTATATTTGCTTCGAA | 1 |
| 4071946 | + | bf638r_3446 | Unkown fuction | 0 | CAATTAATAGGTATTTTTCTGACAAGA AAAAACTTGTAAATTTGCATAAAC | 1 |
| 4075207 | - | bf638r_3448 | Unkown fuction | 0 | CATTCAAATATACTACAAAAATACATT ATTAACCCGTTTTCTCACAAAAGT | 1 |
| 4075218 | - | bf638r_3448 | Unkown fuction | 0 | ACTACAAAAATACATTATTAACCCGTTT TCTCACAAAAGTTAATAAAATAA | 1 |
| 4084876 | - | bf638r_3456 | Unkown fuction | 0 | AACGCAAAGTTATAGTTTTCCTTTATTA TAATAAATGCACAACATGCGATT | 1 |
| 4088052 | - | bf638r_3457 | ppdk | 0 | ACAGACAAAAGTCGCTGTTTTTATTGG GAAAAACAAGAAAAAACAATTCAC | 1/2 |
| 4088395 | + | bf638r_3459 | gdhB2 | 0 | AATACGCCATTCCGAATGACATCAGAA AACTTTTCTTATATTTGCATATGA | 1 |
| 4089928 | + | bf638r_3460 | aminopeptidase | 0 | TTTTTATAAAAAGGTTTCATTTTCAACT AAAATCGCCTAACTTTGCATCTC | 1 |
| 4096176 | - | bf638r_3462 | Unkown fuction | 0 | TTTGCAAAATAAGCACTTATTCCGCAA TCTATCACTAACAGAAACTCCTTT | 1/2 |
| 4097938 | - | bf638r_3464 | tRNA adenylytransferase | 0 | TCTGCAAAGGTACAACTTTTCGCTCTTC ACTTTCCACTCTTCACTCTAAAT | 1 |
| 4098005 | + | bf638r_3465 | Unkown fuction | 0 | TTTCGCTCTTCACTTTCCACTCTTCACTC TAAATTTGTATTTTTGCCCTCT | 1 |
| 4099216 | + | bf638r_3466 | TonB-linked OMP | 0 | TATGTGTAGCGATTACACAGACTATTC ATAGTATCTATCTTTGTCCCCCGA | 1 |
| 4109706 | - | bf638r_3471 | Cold Shock protein | 0 | AACGGCAAATATACACATTTTATTCGG TTATTCCTAATTCTCCGGCCCCAA | 1 |
| 4126588 | - | bf638r_3489 | NusG | 0 | TGAAACATCATTTTTGATGAAATTCAG AACCGGACAAAAAGAAAAATTCAA | 0 |
| 4126772 | + | bf638r_3490 | Unkown fuction | 0 | ATTTATAAAAATATTTTGCTATAGTAAA AATACTGTTTACCTTTGTTCCAT | 1 |
| 4129916 | + | bf638r_3492 | phage derived protein | 0 | AAGTCCTAATAAAGTTGCATCTTAATA CAACCTTTATTATCTTTGCGTCAT | 1 |
| 4131450 | - | bf638r_3495 | Unkown fuction | 0 | AGCATAAAGATAACAAAAACAAACGA CATATCATAGCAGAAGTCATACTAT | 1/2 |
| 4134236 | - | bf638r_3499 | Unkown fuction | 0 | GGAAACAAATGTACTTCTTTTCTGATA AGATTGGCTACATTTGCACCAATT | 1 |
| 4134284 | + | bf638r_3500 | Unkown fuction | 0 | ACGGAAACAAATGTACTTCTTTTCTGA TAAGATTGGCTACATTTGCACCAA | 1 |
| 4134959 | + | bf638r_3501 | HTH regulator | 0 | CGTGTGGCAATATTTGTATATACAAGA AATTATCGGTACTTTTGCCACAGA | 1 |
| 4137300 | - | bf638r_3504 | ruvA | 0 | GAACGCAAAGATAATAAATATTGTGCT TCACGAACCTTGAATAACAAAAAC | 1 |
| 4137408 | + | bf638r_3505 | diaminopimelate dehydrogenase | 0 | AACACCCTTAAGAGACCTGAAAACAA AGAAAATGACTAATTTTGCAACCGA | 1 |
| 4139271 | + | bf638r_3508 | anaerobic ribonucleoside triphosphate reductase | 0 | TCTTGTTAAAACAAGAACAACTTATTG CTTTAACTGTTACCTTTGTAGCAG | 1 |
| 4142550 | + | bf638r_3510 | transporter | 0 | CACTACCTCCCCCAACCGAAGAAGTCC GATTTTGTTTATTTTTGTAGCACT | 1 |
| 4146200 | + | bf638r_3513 | Unkown fuction | 0 | GTAACATTCATATCATGATTATATAATC CAAAAGTCATATATTTGCAATGT | 1 |
| 4150646 | - | bf638r_3516 | peptidase | 0 | ACGACAAATGTAAGCATTTCGCTTTAC TTGCCGTCTGTTTCTTTATGCTAT | 1 |
| 4154017 | - | bf638r_3520 | protease | 0 | GACGCAAATTTACATGATTATTTTCAG ATTACCACACCGAGATATAAATAA | 1 |
| 4155697 | + | bf638r_3523 | Unkown fuction | 0 | TTTTATCCATTTACAGCGATTATCTTAA AATTATCCCGTATTTTTGCACAA | 1 |
| 4158235 | - | bf638r_3526 | Unkown fuction | 0 | GAGCGCAAAGATAGACAAAAAAGCAA AAATATACGTTTTTTGCCCTATCTT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4158244 | + | bf638r_3527 | ubiA | 0 | AGATAGACAAAAAAGCAAAAATATAC GTTTTTTGCCCTATCTTTGTTCTCT | 1 |
| 4161317 | - | bf638r_3529 | glucose-1-phosphate thymidylyltransferase | 0 | GTGACAAAGAACGGAAAGAATTCTGA CGTGACAAAATAAAAATACAAATTT | 0 |
| 4161877 | - | bf638r_3530 | Unkown fuction | 0 | TCTGCAAAGGTAACTATATTTGGGCAA ACCTTACGTGCAAATGGAAAACAA | 1 |
| 4161903 | + | bf638r_3531 | ppk | 0 | TTTGTCATACTTTCTTTGTTTATCTCTGC AAAGGTAACTATATTTGGGCAA | 1 |
| 4166914 | - | bf638r_3533 | responsive regulator | 0 | GCCCGCAAGGTATAAAAAAAAATCAT AGATTTGATTATCTTTGCGGCTAAA | 0 |
| 4166963 | + | bf638r_3534 | ABC transport | 0 | AGCCCGCAAGGTATAAAAAAAAATCA TAGATTTGATTATCTTTGCGGCTAA | 1 |
| 4169215 | + | bf638r_3537 | efflux transporter | 0 | TTAAGTAACACTGTTTGTTTTTTATCCA ATGATAGACTATCTTTGCTTCGC | 1 |
| 4179445 | - | bf638r_3543 | Unkown fuction | 0 | ACCTGCAAATATCAAAAATATTTCGTC TCTTTGCAACTTTCTGTGGCATGA | 1/2 |
| 4179483 | + | bf638r_3544 | Unkown fuction | 0 | TTGTGTATTAATACCTGCAAATATCAA AAATATTTCGTCTCTTTGCAACTT | 1/2 |
| 4180444 | + | bf638r_3545 | sigma factor | 0 | TGGCTTTACAGAATTGCTTATAACGTA TTTTATGATTATATTCGCAGCCGA | 1/2 |
| 4181485 | + | bf638r_3547 | Phospholipase D | 0 | GCTCTTGACAAAAAGAAGGCAAGAAC AGAGATGTTGTATCTTTGTTTCCTT | 1 |
| 4182817 | + | bf638r_3548 | thyB | 0 | TTCACACCATTCCTGTAAGTTATTCAG GCTGTTTTTGTAACTTTGGCAGGA | 1 |
| 4184666 | - | bf638r_3550 | Transcriptional regulator | 0 | TATGCAAAGTTAGCTATTTCCCCGGCA ATAAATACTATAGTTTATTACTTT | 1 |
| 4184762 | + | bf638r_3551 | peptidase | 0 | ACTTTTATTTGTAAAAATAGAATAAAT CGGCAATAGTTAGTACCTTTGTAA | 1 |
| 4189148 | - | | tRNA | 0 | GGGTGCAAATTACAAAAAAAATGCGA ACCACAAGCGTTTTCAGCTACAAAT | 0 |
| 4189189 | - | | tRNA | 0 | AGCTACAAATGTATTCTTTTTTTTCTTC TTCTGCAGCATTATCCTTCTTTT | 1 |
| 4189148 | - | bf638r_3555 | biopolymer transporter | 0 | GGGTGCAAATTACAAAAAAAATGCGA ACCACAAGCGTTTTCAGCTACAAAT | 0 |
| 4189189 | - | bf638r_3555 | biopolymer transporter | 0 | AGCTACAAATGTATTCTTTTTTTTCTTC TTCTGCAGCATTATCCTTCTTTT | 1 |
| 4192582 | - | bf638r_3560 | tonB | 0 | GGCACAAATATAACGCAGATATTTTAA AAATACAGTATATTTGGCGAAAAA | 1 |
| 4192629 | + | bf638r_3561 | cytidylate kinase | 0 | TGTGGCACAAATATAACGCAGATATTT TAAAAATACAGTATATTTGGCGAA | 1 |
| 4194406 | + | bf638r_3563 | 6-phosphofructokinase 1 | 0 | TTAACGTTATTTAATCATACAGGGCAC TCCTTTTTGTTATCTTTGCAACGT | 1 |
| 4198359 | - | bf638r_3566 | Unkown fuction | 0 | AAGGACAAAGTTAAAAGAAAAACTTT AAAGGGACAAATAGTTTACGAAAAT | 1 |
| 4203528 | - | bf638r_3569 | aconitate hydratase | 0 | ATTTATTATAACCAATAATATGAAAGA AAGGTTTAACAAATTGAAATGTTT | 0 |
| 4203616 | + | bf638r_3570 | DNA helicase | 0 | AATTGAAATGTTTATTCCAACCCGGAA ACAGAATTGTTATCTTTGCGGGAA | 1 |
| 4207725 | - | bf638r_3571 | Unkown fuction | 0 | GGCACAAAGATATATATTAAACTTGAT TTATCTACCCAAACAATCTATCTA | 1 |
| 4214693 | - | bf638r_3578 | ilvD | 0 | GCTGCAAAGGTAAGTGCTTTTTTATAA ACTCCAAACAAATTGAAGAAAAAA | 1 |
| 4225284 | + | bf638r_3585 | peptidyl-prolyl cis-trans isomerase | 0 | TCCCCACCAGTCCATAAATCTCTGACG ATTAATTCGTATATTTGTCCCCAC | 1 |
| 4227935 | - | bf638r_3587 | L-cysteine desulfidase | 0 | GCACAAAAGTAGAAAATATACTCTTAT AAACCATCCATTTAACCAAAGAAT | 1 |
| 4228061 | + | bf638r_3588 | Glycosyl hydrolase | 0 | TTACACTCTATCCTCTGGTTATTTAATA ATTATCATGTATCTTTGCCTCTC | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4229313 | + | bf628r_3589 | Unkown fuction | 0 | GTAGGCATTCAAATGTATCTTTTCAAA AAAGAAACTTAACTTTGCAGCCAA | 1 |
| 4233575 | - | bf638r_3592 | Unkown fuction | 0 | GGAGCAAAGTAACAGCTTTATTTCCAT CCAAACAAGGTCTATCTTACGTAT | 1/2 |
| 4238737 | - | bf638r_3595 | Unkown fuction | 0 | GATACAAAAGTATACTTATTATGTTGG GCAACCAAATAGATATAGCCTTTT | 1 |
| 4243920 | + | | | 1 | AAAAAGCTTAAAAAGCTTCTCCATTTA AATTTATTCATATCTTTGTAAAAG | 1 |
| 4245766 | - | bf638r_3599 | sigma factor | 0 | ATTGACAAAGTAACAAATAATAATAAT AATCTCAAAATTTTAAGAAACGGT | 1/2 |
| 4252148 | + | | | 1 | AGATATGCTTAAATAATTGTTTTATTAA AGTTATTAGTATATTTGCCACGA | 1 |
| 4259157 | + | | | 1 | AAAACAAAAAAATACTTTTTATTGATA ATAAAGTTCTTACTTTAGTCCTAT | 1/2 |
| 4259198 | + | | | 1 | TTAGTCCTATCGGGACTAAAAATTCCA TGATTATTCATACATTTGTACGAT | 1 |
| 4261212 | - | bf638r_3606 | sigma factor | 0 | AGTTGCGCAAAATAAATAAAAAATCAT CAAATGTAGTTAATGCAGCTCTCT | 1/2 |
| 4261423 | + | bf638r_3607 | aroC | 0 | AAAACTCGAGAGAACAACTCTTTCCTG AATTTTATTGTTATATTTGCATAT | 1 |
| 4268866 | - | bf638r_3613 | DNA Topoisomerase | 0 | GCAGACAAAGGTACAAAAAAATAGGA AGGGAGAAGTAGTAATCCGTCTCCA | 1 |
| 4268938 | + | bf638r_3614 | fructosidase | 0 | GGAAGGGAGAAGTAGTAATCCGTCTC CATTTATTTTGTAGCTTTGCTTCGT | 1 |
| 4274596 | - | bf638r_3616 | mutA | 0 | GACTTGCAAATATACTGAAATACTCTA AACCTCACGCTTTTTTAGGGAAAA | 1 |
| 4274737 | + | bf638r_3617 | permease | 0 | GGCAGGCAAATATTTGATTACCTCCGA GGTTATAGCTAATTTTGCACCCGT | 1 |
| 4277168 | - | bf638r_3618 | Cytidylate kinase | 0 | AATAGCAAACATACACATTATTTCCGA AAGAAGAAAATAAATAATGTACCA | 1 |
| 4277290 | + | bf638r_3619 | matE | 0 | TTATCAATTGGTAACTTTTATCCATTTT CAGAGTCGTATCTTTGCAGCCGC | 1 |
| 4278683 | + | bf638r_3620 | transcription factor | 0 | TACCTGAAGGACTCTTTCGCCGTTCGC CGATGATATGTATCTTTGCATCAT | 1 |
| 4282508 | - | bf638r_3622 | Unkown fuction | 0 | AGATAAAAGACAAATGCCGACTGCAA TTTGTTTAGCCCACCGGCATGAAAG | 0 |
| 4284044 | - | bf638r_3623 | vmrA | 0 | GGGTGCAAAGATACAAATCTATTTAGA ATTAAAGGTGAAGAATAAAGAATA | 1 |
| 4284144 | + | bf638r_3624 | lysU | 0 | AAGAGTTACAATGCCGTATTTCTAAGA AAAGTTTCGTATTTTTGCCTCCAA | 1 |
| 4288988 | - | bf638r_3627 | Unkown fuction | 0 | TTTGCAAAAATACATATTCTGTTCGATT CACCAAAATAATCGTATCTTTGC | 1 |
| 4289044 | + | bf638r_3628 | Haloacid dehalogenase-like hydrolase | 0 | AAAATACATATTCTGTTCGATTCACCA AAATAATCGTATCTTTGCACTCCA | 1 |
| 4290005 | + | bf638r_3629 | phage integrase | 0 | TGTTGTGCAATATTGCATATATGTTGT GCAAAAGCCTTATATTTGCACAAC | 1 |
| 4293965 | - | bf638r_3630 | EpsE | 0 | AAAAGCAAAGGAAGTGAAAATATCTT TTGCGACTATTATCAAGATACAAAA | 1/2 |
| 4295109 | + | | | 1 | TTTTTTATTAGTTTATAAAACTTTTCAA GAGTTAAGTTGTTAAGGAGATAA | 0 |
| 4295824 | + | | | 1 | TAGCGGTTTTATTAACACTTTTATTAAA AATAAAAATCGACCTTTGTATTC | 1/2 |
| 4296550 | + | | | 1 | TTTTATTAGAATGTACAACAATTATAC GAATTTACGTGTTATTGAGAGAAC | 0 |
| 4314841 | - | bf638r_3651 | Unkown fuction | 0 | ACTACAAATATAATAACATTATTGTTAT TAAGCAAATATTCCAATAACTTT | 1 |
| 4325279 | + | bf638r_3676 | phage genes | 0 | TTTTGGTCACGAAATTTGTTGTTTCCAT AATTTAATATTATAATTGTTGAT | 1/2 |
| 4329161 | - | | tRNA | 0 | GACGACAAAAGTATAAAAAATATTCG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | CATTTTCCTAATGATTCACCTTTTT | |
| 4329348 | + | bf638r_3681 | asnA | 0 | TTAGCCTATTATTCATACAGATATAGA TATTATTCATATCTTTGCCGACAA | 1 |
| 4333939 | - | bf638r_3683 | ostA | 0 | AGGGGCAAAAGTACACAATCTTCACC AAATATACTTGAGAATCTTCATTTT | 1 |
| 4334917 | + | bf638r_3685 | glycosyl hydrolase | 0 | TTTTAATACGGAATGAAGATAATCAAA AGAATTGATTGTATATTTGTGCAT | 1 |
| 4339111 | - | bf638r_3687 | membrane transporter | 0 | AGGTGCAAAGTAACGTAAAAAACTCA AAACTTAACACTAAATAGTGAGTTT | 1/2 |
| 4339189 | + | bf638r_3688 | N-acetylmuramoyl-L-alanine amidase | 0 | ACTTAACACTAAATAGTGAGTTTTCAA CAAATTTCTGCTACATTTGCACAT | 1 |
| 4341665 | + | bf638r_3690 | dnaA | 0 | AGAAAAAGGAGATTTGTTTTTCTCGAA TAAGATTCCCAATTTTGCAACCGT | 1/2 |
| 4344024 | - | bf638r_3691 | Nitroreductase family | 0 | TTTCTGCAAAGATATAAAAAGAATAAT CATTCTATCTCATTTATAAAATAG | 1 |
| 4344181 | + | bf638r_3692 | Ribonucleotide reductase | 0 | ATCTCGTTTTTATTTTGATATACCAAAA AACATTCATAGCTTTGCAGGCAC | 1 |
| 4346916 | - | | | 1 | GAAAGTAAAGATACGACATATGTAGA ATATTTCTATACTTGCAAAATCATT | 1/2 |
| 4347792 | + | bf638r_3693 | 4-alpha-glucanotransferase | 0 | CGACGAAAGTAAAGATACGACATATG TAGAATATTTCTATACTTGCAAAAT | 1/2 |
| 4352805 | + | | | 1 | AACGGGCAAAAGCTTGACTTCCGCAA AGAAAAAGAGTATCTTTGCCTATGT | 1 |
| 4353861 | - | bf638r_3697 | mgsA | 0 | TGCACTAAGATACTACTTTTTATCCGTT CGGGAATAAAGTGTGAGGAAGTT | 0 |
| 4357458 | + | bf638r_3701 | luxE | 0 | TTAACACCTAAAGGTTTCTTAATCCGG AAGTTTATACCTACTTTTGCACAA | 1 |
| 4347070 | - | bf638r_3700 | Unkown fuction | 0 | AGCGCAAAGTTACACATTATTATTAGG AGTAAAGAGGTGAAAGGAGTTAAA | 1 |
| 4359263 | + | bf638r_3702 | Unkown fuction | 0 | AGTTGGGCTTAGGAGTACAACATATA CAAAAAGTACATACCTTTGCAGCGT | 1 |
| 4360971 | - | bf638r_3705 | rpsF | 0 | GGTGCAAAGGTAGGCATTTTATTTTGA ACTGCAAACATTTAATCTTTTTTT | 1 |
| 4361069 | + | bf638r_3706 | MarR regulator | 0 | TTTGTGTTACGAGATAGGATTTGTCAA AGAAGACATTTATCTTTGCAAAGT | 1 |
| 4361646 | + | | | 1 | TTGCGAAAAATATTTCGTCCGTTAAAG AATTATGCGTTAATTTGTGACTGG | 1/2 |
| 4364326 | - | bf638r_3709 | rprX | 0 | GTTTGCAAATTTAATAAAAAATTAAAG ATTTCTCTTCAGAAAATCTACTTT | 1 |
| 4367547 | + | bf638r_3713 | Fe-S oxidoreductases | 0 | TCTACCCTACAGGGTTGTTTACTAACA CGAAAACCTTACTTTTGCAATTTG | 1 |
| 4369380 | + | bf638r_3716 | FecR | 0 | TTGCACACTTCCGATAGGGAGATTCTT CAAACGCTCTGTCTTTATCCATGA | 0 |
| 4374489 | + | bf638r_3719 | Unkown fuction | 0 | GTTCCCCCAATGGTACCGGATTTATAA AGAAACCACTATCTTTGTGCCCTA | 1 |
| 4375537 | + | bf638r_3721 | Unkown fuction | 0 | TTTCTTTCTTTGCGCTTTTATGTTGAAT AACATATTTGTATATTTGCATAG | 1 |
| 4377286 | - | bf638r_3722 | ATPase | 0 | ATTGCAAATATAATAATTATCTTTTAG GTAAAAGATGCAAGTGAAAATATT | 1 |
| 4379407 | - | bf638r_3724 | Unkown fuction | 0 | AAGGCAAATATAACTACTTTTGTCGTC GCAAAGAAAGAAATGATGAAAGAA | 1 |
| 4379381 | + | bf638r_3725 | Queuosine biosynthesis protein | 0 | TTACCATTTACAATCTACAAATTAAGG CAAATATAACTACTTTTGTCGTCG | 1 |
| 4384298 | - | br638r_3728 | mannosidase | 0 | ATGCAAAGATAGTTTTTTTTCAGCTATT CCAATTCACTTCATCTCCACATT | 1 |
| 4391563 | + | bf638r_3733 | Iron only hydrogenase large subunit | 0 | TATTAATGCACGCTTGCAGAATACTCC TACAAGCCGTATCTTTGTAACCCT | 1 |
| 4398009 | - | bf638r_3737 | Unkown fuction | 0 | TTGGTCAAAGATACAGAACATTTGTTA TTTACCATTTCTTTTTCCGGATAT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4403054 | - | bf638r_3740 | Unkown fuction | 0 | CGTACAAATGTATTTTAAATTAGTCGA ATCTTTAGATTTAAAAGCAGAAAG | 1 |
| 4404229 | - | bf638r_3741 | DHH | 0 | GTTGGCAAAAATACGATAAATCTTTGT ACTTAGAGCAAAAACCTCACAGAA | 1 |
| 4407039 | - | bf638r_3743 | ribulose-phosphate 3-epimerase | 0 | AGCACAAAATTAACAAAAAAATATCA GGGCTCCCCCTCAATAACAATTTAT | 1 |
| 4410631 | - | bf638r_3746 | Unkown fuction | 0 | CGTGCAAAAATACTAATTTTGCCCAAT AATAACAAAGGTTAAGAGAAAACA | 1 |
| 4410657 | + | bf638r_3747 | 4-hydroxybenzoyl-CoA thioesterase | 0 | TTGATCTATTTCTTGTGAGAATAACGT GCAAAAATACTAATTTTGCCCAAT | 1 |
| 4413301 | - | bf638r_3748 | Unkown fuction | 0 | AATACAAAAATAACAAAAAAACAGAA ACTTACAAAATCTAATCATCTTAAT | 1 |
| 4413717 | - | | | 1 | ACTCACAAATATACAAGATTATTCCTA ATTACCGATTCCGTAAACCATATA | 1 |
| 4441521 | - | bf638r_3775 | updY | 0 | GAAACACACTTTTTTAAAGATCCAGAC AATTAAAATACATATAAAAGCAGA | 0 |
| 4441706 | + | bf638r_3776 | Unkown fuction | 0 | TTTTATTAAACATTTTGCAGTTATGAAA ATACCCTCTATCTTTGCGTTCAA | 1 |
| 4444661 | - | bf638r_3779 | Methyladenine glycosylase | 0 | GATACAAAGATAGTAGATTCACCATGT CGCACCAAATATCGATAAACAAAA | 1 |
| 4444726 | + | bf638r_3780 | recJ | 0 | TTTTTCAGTATATCCCCCACACTTCCGG TTTTTTTCGTACTTTTACAGCTT | 1/2 |
| 4448451 | + | bf638r_3782 | Tetratricopeptide repeat | 0 | TTCCTGAAAAGCAACAGGGTTATGTG ATGAAATTACTACATTTGCAACCAA | 1 |
| 4458396 | + | | | 1 | AAAACAACAATGAGACACTTAATTCGA GTAATTTTTCTCATCTTTGTATGC | 1/2 |
| 4460358 | - | bf638r_3789 | transcription regulator | 0 | CTGACAAATAAAAGAAGTTTTTAGGA GAAAAACTAGCATCTTAAGAACTGA | 1/2 |
| 4460570 | + | bf638r_3790 | Prephenate dehydratase | 0 | AAAGTTTGTTTTTTATTTGCGTTTAAGA TAAAAGCATTAATTTTGCCTCAG | 1 |
| 4465389 | + | bf638r_3795 | dnaG | 0 | ACTGAAAAGTCATTTGATTATCCAAGA AATATCACGTACATTTGCATCCGC | 1 |
| 4475635 | + | | | 1 | TTTTTAGGTACACAAAAATCCGTGCAA CGAGTATTTTTCATATCTTTGCTT | 1 |
| 4477406 | - | bf638r_3801 | sigma factor | 0 | GCGGCAAAAATAGATAACTATAACTAT ATAACCAAACTAAAATATAAGTAA | 1 |
| 4481300 | - | bf638r_3806 | Unkown fuction | 0 | TGTGCAAAGGTAACAATAATTAAAGTA AAAAGAAAAGAGGGCACTCTAAAT | 1 |
| 4482805 | - | bf638r_3808 | Nucleoside diphosphate kinase | 0 | CCCCAAAGATAGCATTTCCAGACGAAA GAGAAGGCATGTTCTTAACTATTT | 1 |
| 4489173 | - | bf638r_3815 | pdx | 0 | GCGCAAAAGTACAAATATTAGCGATA AGATAACAGCAATACCGAAAAAAAT | 1 |
| 4489240 | + | bf638r_3816 | ATP-NAD kinase | 0 | TTAGCGATAAGATAACAGCAATACCG AAAAAAAATCCCAAATTTGCGACAT | 1/2 |
| 4495995 | - | | rRNA | 0 | GATGCAAAAGTAAGGCATTAGGAGAT AAGAGACAAATATAAAAGGGGATTT | 1 |
| 4497233 | - | bf638r_3818 | Unkown fuction | 0 | AAGTACAAAGATAAAAAAAAATCAAA TATCGGTTCTTCAATCTAAAGTATT | 1 |
| 4497490 | + | bf638r_3819 | sigma factor | 0 | AAAATTAAAACTACTCTCTCACATTGC AAACAATTTAGTTATATTTGCATT | 1 |
| 4499324 | - | | | 1 | ACGACAAAGTTATGAAAAAAAACTTGT TGCAGAAATGGTATGTCCATCAAC | 1 |
| 4509602 | - | bf638r_3826 | Transcriptional regulator | 0 | GACACAAATATATAACTATAAATATAG ATATAGAACTATATTTATAGTATT | 1 |
| 4514138 | - | bf638r_3828 | mdh | 0 | GACGACAAAAGTACAACGATAATCGA AATTAAAGAAATTTTTCCATAATAA | 1 |
| 4514252 | + | bf638r_3829 | Unkown fuction | 0 | AAAGTGAATAGACTGTGTTACTCTGTA TTGAAATGATTACTTTTGCAGCGT | 1 |
| 4516748 | - | bf638r_3831 | sigma factor | 0 | AAGCCGACAAAATTAGCGATAATTCTT TTATCAGGCAGACTAAAGGCGCTG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4519813 | + | bf638r_3833 | efflux transporter | 0 | AAAAATTAAATCGTTTTGCTTTTTAAG ATTAACCCTTTATATTTGCATCGT | 1 |
| 4528792 | + | bf638r_3840 | Transcriptional regulator | 0 | TTAACACACTTCACTTTGTTTAGTTACG ACCGTTCCGTAGCTTTGTAATAT | 1 |
| 4536718 | - | bf638r_3846 | aminotransferase | 0 | GCAACAAAAATAACCTTTATTCTTCAG ATAGCCTCTTTTCTTGCAGTAATT | 1 |
| 4540823 | - | bf638r_3850 | tRNA Methytransferase | 0 | GAATGCAAATGTACAAGATTCCCGGC AGAATACAATCAATTGGACAAGTTA | 1 |
| 4541682 | + | bf638r_3853 | Unkown fuction | 0 | ATAGCTAAAAGCATTTGCCTTCTTCAA AAAGAATCAGTATATTTGAAGTAT | 1 |
| 4545414 | - | bf638r_3857 | Rex | 0 | GATTGACAAAATTACACTATTTTTTCCG AAATGCCATAAGGATGTCATAGA | 1 |
| 4545551 | + | bf638r_3858 | Translation intiation factor | 0 | GACACTTAATAGTTTGCAAGGTTCGGA ACAATATATTATTTTTGAACTTAA | 1 |
| 4549158 | - | bf638r_3862 | rplM | 0 | GCTTGCAAAAGTACGGCTTTTCTTTGA ATTGGCAAACTATTTCTGTCTTTT | 1 |
| 4549899 | - | bf638r_3864 | Unkown fuction | 0 | GTCACAAATATAAAGATTTCTTTTATA AAAATAGCAATCTAAAGACATAAA | 1 |
| 4554377 | - | bf638r_3867 | purB | 0 | AGGCTGCAAAATTAACTCTTTTTCCTG AGTTATAAGGCATAAGGGGAGAGA | 1 |
| 4556488 | - | bf638r_3868 | beta-gal | 0 | GAAGCAAAGTAAAGAAATAAATCACT TTCCCGCCAAATTTCCCTTTGTTTT | 1/2 |
| 4556963 | - | | tRNA | 0 | GATGCAAAGGTATGGCTTTTCTTGAGA CCTACAAATTATTTGCGATATTTT | 1 |
| 4558490 | - | bf638r_3870 | Unkown fuction | 0 | AAAGGAACAAAATTAATAGAAAAACG GCAGTATGCACTTTTTTTCTTATTT | 1 |
| 4558548 | + | bf638r_3871 | beta-N-acetylglucosaminidase | 0 | AAAATTAATAGAAAAACGGCAGTATG CACTTTTTTTCTTATTTTTGTGCAG | 1 |
| 4573047 | - | bf638r_3879 | aapA | 0 | AAATTCCGAGATAAAGAAGTGTCTTTA AAGGAGTTGTTTTTACTAAAATGA | 0 |
| 4573281 | + | bf638r_3881 | aapI | 0 | TAATTGTTTTATTATGCGTAATTCAAAA AATATATTTAAATTTGCCGCCGA | 1 |
| 4574434 | + | | | 1 | TAACAAACCAACTGCATGATTATCTCT TTTTTTTAGATAACTTTGCAGCCT | 1 |
| 4574992 | - | bf638r_3882 | Unkown fuction | 0 | ATTACAAATATAAGTTAATTATAAGAC AGGCAGAAAAAAAGACGGGAAAAT | 1 |
| 4578554 | - | bf638r_3883 | carB2 | 0 | TTGTGCAAAGTAACGACTTATTTCTAT AATTACAAAGGAATCTGGTAAGAA | 1/2 |
| 4578654 | + | bf638r_3884 | trpS | 0 | ATATTCGTTTTATTTGTTACTTCTGCCG ATTCTGACTATCTTTGCCCATTG | 1 |
| 4581111 | - | bf638r_3885 | ompA | 0 | ATTTAAATAACACCAACGAAACTAAAA GGTTTTGGATAAATAGAAATTTCC | 0 |
| 4601065 | - | bf638r_3903 | Unkown fuction | 0 | TCCTCAAAAGTACAATTATATTTTTAAA TATCAACAATAAGTAAAATAAAG | 1 |
| 4608419 | + | bf638r_3913 | two component regulatory systme | 0 | TACAATAATTCTTTCCATAAAAATCAAA TACAAAAACTATCTTTACAATAT | 1/2 |
| 4619354 | + | bf638r_3923 | ubb | 0 | TACTGTAATATATAATAATAAAGAAAA ACAGATAACTTATATTTGTCCCAT | 1 |
| 4632883 | - | bf638r_3932 | recQ | 0 | CTAACAAAGTTATAAACAACTTCTTAA ATATCAAGCATAATCAGACTAATA | 1 |
| 4634929 | - | bf638r_3934 | clpP | 0 | AACTGCAAAGAAACAAAAAAGAACAC AGAGACACAAAAAAACACAGAGTTA | 1/2 |
| 4636439 | - | bf638r_3935 | trigger factor | 0 | GGGTGCAAAATTAGTGCTAATCTTTCA AATTTCAAAAAACATTGCAGAATT | 1 |
| 4637475 | + | bf638r_3937 | RNA binding protein | 0 | AAAGAGAAAAAGTTATTGCAATTCCG GTTTTTATGCGTTTCTTTGCTGCGG | 1/2 |
| 4638771 | - | bf638r_3938 | ABC transport | 0 | GCGGCAAATGTAGCAAAAATAAGCCG AAAAAACGTACATTTGCAAACAAAT | 1 |
| 4638819 | + | bf638r_3939 | Unkown fuction | 0 | TTGCGGCAAATGTAGCAAAAATAAGC | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | CGAAAAAACGTACATTTGCAAACAA | |
| 4641261 | - | bf638r_3941 | responsive regulator | 0 | TGTACAAATATAAATAAAAAGCATTAC ATAAGAAACAAAAAGAATCAATTA | 1 |
| 4645712 | - | bf638r_3946 | Unkown fuction | 0 | GCGACAAAGATACGAATAATAATCCTC ATAACATACAAATTATCAGTAAAT | 1 |
| 4651357 | - | | rRNA | 0 | ACTGCAAAGATAACAACTTTTATAATA ATGATCCAAATAAAAAGGAAATTA | 1 |
| 4651484 | - | | rRNA | 0 | GGATGCAAAAGTAAGGCATTAGGAGA TAAGAGACAAATATAAAAGGGGATT | 1 |
| 4653034 | + | bf638r_3947 | Methylenetetrahyd rofolate reductase | 0 | CAGGTTCTACAACGATTTCGTTTCATCT TTTTTTCGTATTTTTGTCCCAAT | 1 |
| 4664440 | - | bf638r_3956 | purH | 0 | GCGCAAAGTTACGAAATATTGTTTATC GTACGTAATAAATAAGGTGTGAAT | 1 |
| 4668598 | - | bf638r_3958 | ABC transport | 0 | GCAGCAAAGATACTCATTTATACAGAC TGCGAAAAAAGAAACGGAATAAAT | 1 |
| 4671218 | - | bf638r_3959 | Unkown fuction | 0 | ATTTCAAATATAAACACATATTCTGAA ATAAGCAACGAAAAGAATATAAAT | 1 |
| 4673013 | - | bf638r_3960 | glycosytransferase | 0 | GTTGCAAAAGTACAAAATAGAATGAA GAATGAAGAACGAAGTATGAAGAAT | 1 |
| 4680045 | - | bf638r_3962 | Unkown fuction | 0 | GATACAAAGGAAACAAAATTTTCCGG AAAAAGATTAAAGTGACGACGATTT | 1/2 |
| 4685872 | - | bf638r_3966 | cyclophilin | 0 | CTGCAAATATACGACTTTACCTCAAAA CCTCTTACCTTTGTAGCTCACAAA | 1 |
| 4685917 | + | bf638r_3967 | responsive regulator | 0 | TTAAGTCTGCAAATATACGACTTTACC TCAAAACCTCTTACCTTTGTAGCT | 1 |
| 4686719 | - | | unknown transcript | 0 | ACTACAGGAACGTAAAATACGCCCTGT TGGTTCAAACTCAGAAATTACTGT | 0 |
| 4687443 | + | bf638r_3969 | unknown function | 0 | CAGATTAGTTCACACGTATTCTTTTTGG CACAAGATTTGGGAAATACTCAA | 1/2 |
| 4689990 | + | bf638r_3972 | Transcriptional regulator | 0 | GTGCAAAAGTAATTCTTTTCTTTCGTCC ACACAAGAAAACAAATAAATTAT | 1 |
| 4690186 | + | | tRNA | 0 | AAATGCCGAAATATTTGGTAGTTCAAA TAAAAGCCGTACCTTTGCACTCAC | 1 |
| 4690526 | + | bf638r_3973 | unknown function | 0 | AAGATGTAAATTGTGGTGACTCTATGC AATTAGTTACTATTTTTGTCCCGA | 1 |
| 4696773 | + | bf638r_3980 | ABC transport | 0 | ACATTGCCAATTCGGTTGTAATATTCA TATATTTATCTACTTTTGTGGTGT | 1 |
| 4702910 | + | bf638r_3987 | unknown | 0 | AATGAAGTGATTTTGTTCGGCAAGTG GTAAGAATGCTTATATTTGCAATAT | 1 |
| 4710608 | - | | rRNA | 0 | ACTGCAAAGATAACAACTTTTATAATA ATGATCCAAATAAAAAGGAAATTA | 1 |
| 4710736 | - | | rRNA | 0 | GATGCAAAAGTAAGGCATTAGGAGAT AAGAGACAAATATAAAAGGGGATTT | 1 |
| 4711814 | + | | | 1 | TATACAGCTTATAGGTTGTGATTTGCT TTCAAATTAGTATCTTTGAACCAT | 1 |
| 4720701 | - | bf638r_3991 | Endonuclease | 0 | ACCCCAAATATATTAATATTTTTCCATA AATAAACATGATAATAAAAATTA | 1 |
| 4720770 | + | | | 1 | TTTTCCATAAATAAACATGATAATAAA AATTATTCATTATCTTTGGGTCAC | 1 |
| 4725389 | - | bf638r_3997 | cas1 | 0 | GTTGCAAAGATAGATAAAACCCATAA AAAGAGAGGAGAATCTGCATGAAAT | 1 |
| 4732211 | - | bf638r_4002 | Transcriptional regulator | 0 | GAACAAAGATAAGCAGAAAACTCGAT GAGTTAATTCAGTTTCCTGATTTAA | 1 |
| 4738880 | - | bf638r_4007 | hutU | 0 | CTTCGCAATATAGAAAAAAGGTGGGA ATCTACCAAGATTTCAGCAGATTTA | 0 |
| 4739687 | - | bf638r_4008 | unknown | 0 | GGCGGAAAGGTACACCTTTTAACGAA TCGGAGCAAGTTTTCAGAAGAAAAG | 1/2 |
| 4739797 | + | | | 1 | AAAGAGAAAACACTTGACTTTTGCTCT CCAATGTCGTATCTTTACTATCAG | 1 |
| 4740205 | + | bf638r_4010 | mutS | 0 | CATGCCAAGAGAGCGCCGTACTCACG | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | CTAAATATATTATTTTTGCGGGTCA | |
| 4743127 | - | bf638r_4012 | unknown | 0 | GTCTGCAAATATAACAGATTATTCAGT CCGACTGTTCACCTTCAAATGTTT | 1 |
| 4751778 | - | bf638r_4027 | rpsN | 0 | GTTACAAAGGTAATATTCATTCCGAGA ATTCTGGTAATACTATCGATATTA | 1 |
| 4756629 | - | bf638r_4039 | rplW | 0 | GCTGACAAATATACATTTTTATTTGCTT CTGGTAAAACTACAAGCAGTTTT | 1 |
| 4761422 | - | bf638r_4045 | rpsL | 0 | GCTGCAAAAATACGAATAATATTTGAA TAATCAATGCACTACATCTTTTTT | 1 |
| 4762828 | - | bf638r_4048 | unknown | 0 | ATATAGGAACGCTGCAGGCACAGGAA TGTTCGGTAAATAAAGAGGGAATTA | 0 |
| 4779516 | - | bf638r_4063 | ribosomal protein | 0 | GCCGCAAAATTACACATACTTTCTTTAT TCACAAAACTTTCGGCAATTATT | 1 |
| 4784988 | - | bf638r_4067 | gltX | 0 | ATTTCTCAAAATCAATGCAAAAGGTAC AGAAAGTAGTTATCTTTGGAGAGT | 1 |
| 4785043 | + | bf638r_4068 | metal dependent phosphohydrolase | 0 | AAAATTAAACCTTTTTTTCCACACTATT GTTTTTTTTTGTACTTTTCGCAA | 1/2 |
| 4791982 | - | bf638r_4071 | priA | 0 | GAGGACAAAGATACAAAAAATCCCCG ACGCAAGCGCCGGGGATTCGTTTAT | 1 |
| 4792702 | - | bf638r_4072 | unknown | 0 | GGGACAAAGATATAGTTTTAAATTAAA AAAACATTAAATCTGTTTGCATAT | 1 |
| 4793082 | + | | tRNA | 0 | AAACACCTTTTCCGGAACAACGGACCG AAAATCCATGTACCTTTGTGGTAC | 1 |
| 4794538 | - | bf638r_4074 | unknown | 0 | GACGGGAAAGATACATATTATTACTGA AACCGATGTACTACTATACCCTAA | 1/2 |
| 4796142 | - | bf638r_4075 | aspA | 0 | AATAACCAAACTTAAGCAATATTGTTT TGAAAGCATACATATTCGCAATTT | 1 |
| 4796359 | + | bf638r_4076 | Anaerobic c4-dicarboxylate membrane transporter | 0 | ATATACTAATAGACAAAAACCTATTAA CTTCTTTTTTGTTGATTAGCATAT | 0 |
| 4800423 | + | | | 1 | ATCGCCTGTTCTATCAAGTATATTCTCT GAATTATTAGTATCTTTGCCTCT | 1 |
| 4805380 | + | bf638r_4087 | unknown | 0 | TAGATACATCCACTCCGATAAAATAAA AATTTTCCATACTTTTGTCTTCAC | 1 |
| 4846574 | - | | | 1 | GGTTACAAAATTACTATAAGTTTAAAA TATTCATTTATCTTGTCTCATTTT | 1 |
| 4850609 | - | bf638r_4125 | ATPase | 0 | GCAAAGATAGGGAAATAGTTTATGAA TCATATAAATTGCAACACAAATTTG | 1 |
| 4857615 | - | bf638r_4131 | phage integrase | 0 | GAGAAGCAAAGGTAATGTATTTTAGA CAAACAAACCCAGTTTTTGAGAAGA | 1 |
| 4858053 | - | bf638r_4132 | unknown | 0 | CGGCAAAAGTATAATAAAATATTTCAC TTGCAACTATTATTCCAAAATATT | 1 |
| 4859342 | - | bf638r_4134 | unknown | 0 | AACGCACAAAGATATATAAAACCATTT GAGAAGTAAAGTGTTATATGCAAG | 1 |
| 4859393 | + | bf638r_4135 | unknown | 0 | ACGCACAAAGATATATAAAACCATTTG AGAAGTAAAGTGTTATATGCAAGT | 1 |
| 4865126 | + | | | 1 | AGATGTTTTTTAATTTGCTTTTTTGTGA AAAATGGTTTATGTTTGTCGGAC | 1 |
| 4877735 | + | | | 1 | GGCACAAAAGATGTGCTAACAAAGGT TCTATTTTCATATCTTTGTTTCCGA | 1 |
| 4879513 | - | bf638r_4147 | sigma factor | 0 | GAATGCAAATATAGCTGTTTATTTGTT AATAATACCAATCGGCATCTAATT | 1 |
| 4883323 | - | bf638r_4150 | GDSL-like Lipase/Acylhydrolase | 0 | CCCCAAAGATACACAACGATTTTCAAT AAACGGACGCAATCCTCCTCAAAT | 1 |
| 4892611 | + | | | 1 | TGTAATACCTCTGACCATAATTATCTTA TATTTATTAATTTTGTCCACCAA | 1 |
| 4894054 | - | bf638r_4157 | unknown | 0 | ATTCGCAAAAATAGATAAAAAAAAGC AAGTTGCAAAAAATTCCCTCAAAAT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4894249 | - | | | 1 | ATCACAAATATACAATAATTAAACAAG CAACAAAGATATAAAAAAGGAAAA | 1 |
| 4894275 | - | | | 1 | GCAACAAAGATATAAAAAAGGAAAAG AAGAATAAATGGTAATAAAAAGGGG | 1 |
| 4894452 | + | bf638r_4158 | ADP-ribose pyrophosphatase | 0 | TTTATTAATGATTCTGCACCCGTCAAA GAATTAGCACTATATTTGCCCTAC | 1 |
| 4895091 | + | bf638r_4159 | two component regulatory systme for cell autolysis | 0 | AAAAGCCTTGGAGCGTAGGGAGTTTC TCATCGGGTGTTGTCTTTATAGGAA | 0 |
| 4904249 | + | bf638r_4167 | Transcriptional regulator | 0 | CTTTTTCACTTTCCATAAATAACTATGT TTTATGATTACCTTTGCAAAGGT | 1 |
| 4905713 | - | bf638r_4168 | Methylated DNA-protein cysteine methyltransferase | 0 | ACAGCAAATGTACAGAATTTCCGGATG AAACCAGCCGGAGAGTTACTAAAA | 1 |
| 4906398 | - | bf638r_4169 | unknown | 0 | GATACAAAGGTAAGTAGTTTATAATCA TCTAGCAATAAGGTACTTCTATGT | 1 |
| 4906691 | + | bf638r_4170 | Unknown | 0 | TAAGTGACAAGGAAGAACAATATTGG CCGGAACATCCTATCTTCGCAGGAA | 1/2 |
| 4911085 | + | bf638r_4174 | unknown | 0 | TCCGTTCTATATAATACAAAAGATCAA TTAAATAGCTAATTTTGTATTCGG | 1 |
| 4914543 | - | bf638r_4178 | unknown | 0 | TTTTCAAAGGTAGTGTATTTACTATGTT CAGCCAAAGAAAAACATTCATTT | 1 |
| 4914929 | - | | | 1 | ATGGCAAAGATAGAGAATAAAACGAA AGAAAACCCCAAGTTAGAGCAAAAT | 1 |
| 4916183 | + | | | 1 | AGTTTGCCCGAGTTAAAGCTTTGAGAA TAAAAATAGTAACTTTGCACAAAT | 1 |
| 4917938 | + | bf638r_4183 | integrase | 0 | TATTTAGCAGATGTGGGGAAAATGTG GGGAAAATATCTATATTTGCAGCAG | 1 |
| 4925911 | + | | | 1 | TTCCTACATCAAGACCTACAAAATGAG AATAATTCATAACTTTGTATTTAC | 1 |
| 4926064 | - | | | 1 | CTTTTCAAATATACACTTATATTTTGTT GTGATATTTAATACTGCAAATGT | 1 |
| 4926104 | - | | | 1 | ACTGCAAATGTAAATGTGTCAAAACTA AGGCACATCCCCTTCTTTTTTTAA | 1 |
| 4927845 | + | | | 1 | AAAGAACTGTTTGGTTGGGAATGTAA GGACGAAAATGTAACTTTGCCATAA | 1 |
| 4927859 | - | bf638r_4190 | transposase | 0 | GCCTAAAGATACAAAAACTTTGGGTAA TAACAAAGCCCGGGCTTGTGAAAG | 1 |
| 4929096 | + | bf638r_4192 | Transcriptional regulator | 0 | ATCTCTTCTTTTTGTAAGAGTCTGTTTG ATTTTTACTATGTTTGCAGATGG | 1 |
| 4930194 | + | bf638r_4193 | TonB-linked OMP | 0 | GGTGTATCGGATATTTGTAAACTCCGA ATATTGTCGTATATTTACATCCGT | 1/2 |
| 4939436 | - | bf638r_4200 | Transcriptional regulator | 0 | TTCGCAAATATACGAACAATATTATAA AAACCAACAACACTTCTATAATTT | 1 |
| 4939738 | - | | tRNA | 0 | GGTGCAAAGATACGGCTTTTTTTTGAA CTTGCAAGAGTTTTATAAAAAAAA | 1 |
| 4939857 | + | bf638r_4202 | Mannose-1-phosphate guanylyltransferase | 0 | AGATCGCATAAACACTTCTATTTTCAG AGAATAAACCGTATTTTAGCAGAA | 1/2 |
| 4942141 | - | bf638r_4204 | greA | 0 | TCGCAAAGATAACCTTTTAAATGACAT GTGTCAAGCAAAAAGCAATGCTAT | 1 |
| 4943403 | - | bf638r_4205 | unknown | 0 | GCTGCAAATATACTTCTTTTCAAGATTA GTCAAAGCATTTTGCCGCGTTCT | 1 |
| 4943543 | + | bf638r_4206 | pnpA | 0 | TTTATCTCATTTTCACGATATTAGTGCA ATAAATGTATATCTTTGCGGGAA | 1 |
| 4947642 | - | | | 1 | GTTACAAAGATATGAGAAAAATTTCTT TGCCGAGGCATTTATGCCCAAAAA | 1 |
| 4952886 | - | | | 1 | GCAACAAAGATAAGATGCTTTTTTTAC TTGTGCAAATCTTTTCGCATCGTT | 1 |
| 4953002 | + | bf638r_4212 | unknown | 0 | TTATCTATTCCCTTTGCATTTTGTACAG AGAGCAAGAATCCTGTTTCAAAA | 0 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 4955226 | - | | | 1 | GCAACAAAGATAAGGCGCTTTTTTTAT CTGTGCAAATTTTTCTACATCGTT | 1 |
| 4960102 | - | bf638r_4218 | Phosphate acetyl/butaryl transferase | 0 | AGCAAATATCAGAAAAAGGATCGAGA AGTACATCATGGAATACTGAAAAAA | 1/2 |
| 4960242 | + | bf638r_4219 | Laminarinase | 0 | TCATTTCTACGGATGCAGGTAATTTCG TTTTTTAATGTAAATTTGTGTCCT | 1 |
| 4962079 | + | bf638r_4221 | Unknown | 0 | AACCCTCCTATCTGACTCACCCTACAGT GAAAATCATTAATTTTGAGGCTT | 1 |
| 4965410 | - | bf638r_4225 | protease | 0 | GTACAAAGATAACGGTTTTAATTGACT ACCGCTTCGGTAGAAGAGCTATTT | 1 |
| 4970714 | - | bf638r_4229 | plasma-membrane calcium-translocating P-type ATPase | 0 | TTGCAAAAGTACAACATAAATCCGGTA TAAGATGTTTCTTAAATCTTTTTT | 1 |
| 4972459 | - | bf638r_4230 | Two component regulatory System | 0 | ACAGACAAAGGTAGTGCTTTATTAATG AAAAAAGAAAAAATAAATGAACAA | 1 |
| 4972538 | + | bf638r_4231 | Rubredoxin | 0 | AAAAGAAAAAATAAATGAACAATACG TCCGATACTTTTGTTATATTTGCAT | 1 |
| 4974603 | - | bf638r_4232 | Na+/phosphate symporter | 0 | ATTGCAAAAATAATAATATTTCTAAAG AATTCCCTATTGGGAGCTCACTTT | 1 |
| 4974557 | + | bf638r_4233 | threonyl-tRNA synthetase | 0 | TTTTAGTGGTTGTTTGTCTTCTTAGCAA TTATGTGATATATTTATCCGTAC | 1/2 |
| 4976476 | + | bf638r_4235 | BaeS | 0 | AAGAGGGGGAAATGAGAGGTTTCTAC AAAGATTTTTTAACTTTGTCACTGT | 1 |
| 4980424 | - | bf638r_4236 | unknown | 0 | TATATTTTGAGTTATTAAATCTGAATAA TAGCACATATACTGTCTCTTCCT | 0 |
| 4984080 | - | bf638r_4238 | unknown | 0 | AACGGACAAAAGTAGTGTATCGATGC ATAAAAAACAAGAATTATCAGAAGA | 1 |
| 4984947 | - | bf638r_4240 | unknown | 0 | TCTCCCAAAGGTATGCCCTTAAAATTTT AAAAACAACTCTTTCTTCTCTTT | 1 |
| 4985265 | - | | | 1 | CAACAAAAGTATCTTATTAAAATCAGA AAAACAATCCTTTCCTGATTTTAT | 1 |
| 4986879 | - | bf638r_4241 | proline aminopeptidase | 0 | GATGGCAAAGATAGGAACAAAGGATG AATTGACAAAGGAGATAAAGCAAAT | 1 |
| 4987093 | + | bf638r_4242 | sigma 54 | 0 | TATCGGATAAATCATTTCGCTATATCG AACAATTCATTACTTTTGTAGCGA | 1 |
| 5001873 | - | | | 1 | TCTACAAAGATAGGACAATAAATCAAA AATACAATCTTATTTTCTTTATTA | 1 |
| 5009116 | - | bf638r_4254 | peptidase | 0 | GACGTAAAAGTAACTCATTGCCTGCTT ACTCTGACAGAACTTCGTTAATAT | 1/2 |
| 5010142 | - | bf638r_4255 | adenosine kinase | 0 | TTTTTGCAAAGATATTGCATATTTCAAA ATATCCTATTACTTTTGCATCGC | 1 |
| 5010192 | + | | tRNA | 0 | TTTTTGCAAAGATATTGCATATTTCAAA ATATCCTATTACTTTTGCATCGC | 1 |
| 5034403 | + | bf638r_4267 | Antibiotic biosynthesis monooxygenase | 0 | TATTCAAGATGAAAAAGAAAGTATCAC CATTTATACGTAAATTTGCGGAAC | 1 |
| 5036313 | - | bf638r_4269 | suhB | 0 | TTCGGCAAAGATAAATGATTCTTTTGA AAAAACATTGCAGCATTCGGCATA | 1 |
| 5041092 | - | bf638r_4274 | unknown | 0 | GTATACAAAGATACATTTTCAATTTGA AAAATCACTTAAAATCAATCATTT | 1 |
| 5066012 | + | bf638r_4299 | NADH pyrophosphatase | 0 | TGGCGGACATATCACCTGAAACTGCTT CTTTTTCGTTATTTTTGCCTGCAC | 1 |
| 5066578 | + | bf638r_4300 | unknown | 0 | AAAGATTCCTGAAAGAGAGGTAAAAT ACATAAAAACAAGTACTTTTGTCAA | 1 |
| 5071560 | + | bf638r_4302 | topA | 0 | TTCTCTCTATATTATAAGGTAAAGAAA ACGTTTTTTTCTACTTTTGCGGGC | 1 |
| 5077084 | - | bf638r_4304 | argS | 0 | GCGCGCAAAGATACAAAAAAAGGAGA CGCTAAAGCGCCTCCTCTCTTAATT | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 5077479 | - | bf638r_4305 | hup3 | 0 | CTCACAAATATAAAGTAACAAAACAAT ATATCAAATAAAAACATCAAAAAA | 1 |
| 5077611 | + | bf638r_4306 | unknown | 0 | TAGATTATTTTTCTGTTGTAGAACAGA ATTTATTCGTATTTTTGCGTTCAT | 1 |
| 5081090 | + | bf638r_4310 | secD-secF | 0 | TTGCCTCCTAATTACTTCTGAATGTGAA AAAAAAACTATATTTGCAGCCTT | 1 |
| 5098651 | + | | | 1 | GGTTGCTTTGGATAAGTGATTCTTTTT GCAGGTGCATTATCTTTGCCCTGT | 1 |
| 5100538 | - | | tRNA | 0 | GGCTGCAAAGATAGCGTATTCTCCTGA TTTGGCAAACCCGAAAGAACAGAT | 1 |
| 5100923 | + | bf638r_4321 | porA | 0 | AATCAGAAAAGATTTTGTTAAGTCACT TAACTTTGCGAACTTTGCCTGCCA | 1/2 |
| 5107036 | - | bf638r_4324 | ugpQ | 0 | AGCGGCAAAGATACACAAATCTCTCCT TGCTTGTTTACCTTCAGAAAGCAT | 1 |
| 5107865 | - | bf638r_4325 | Phosphoribosyl transferase | 0 | GGTGCAAAGTAAGTGATAATCTTTCAA TTTGCAACGTTCAGTTTGCAGATA | 1/2 |
| 5108005 | + | bf638r_4326 | pckA | 0 | TTTCTCTAATATATTCTTTGTTTCCAAA GAAAGCATTACTTTTGTAGTCGA | 1 |
| 5110380 | - | bf638r_4327 | unknown | 0 | GGGGCAAAAATAAAACGAATTTTAAT AAGAAGCAAGAGATATTAATAAAAT | 1 |
| 5116732 | - | | | 1 | GGGGTCAAAGATAATAAGAAGTTAAT ATCCCCGCAACTTTTTACTATATTT | 1 |
| 5129750 | + | bf638r_4337 | unknown | 0 | AACGTCACATAGTCTGCACAAGATTCA AGGTTTTTCTTTATCTTTGTGCAG | 1 |
| 5140713 | - | bf638r_4340 | typA | 0 | GCGACAAAGGTACAAAAATTGGATTA GAAAAAAGGGATACTATAGAAAATA | 1 |
| 5140814 | + | bf638r_4341 | rpsO | 0 | TTTCCTGCAATTCATTGTGGGATAAAA GATAATGCCTATCTTTGCACGCTC | 1 |
| 5141232 | + | bf638r_4342 | transcriptional repressor puuR | 0 | TTTTATCTTTCTCCCTTGCTTGGATATA AAATTTCATTAATTTTGCAGCAA | 1 |
| 5145212 | - | bf638r_4344 | arcB | 0 | ACTTGCAAATATAATACATATATCCAA AAAATATCCTACAAATCGAAGTTA | 1 |
| 5158644 | + | bf638r_4354 | sigma factor | 0 | TTTTCGGCTTTTTTTCTGTTTTATAAATT GAGTATGCTATATTTGTAATTC | 1 |
| 5161007 | - | bf638r_4356 | ribosome-binding factor | 0 | AACACAAAGTTATAGAAATATATCGTT TGATAGGCCAATAATATGTTATTT | 1 |
| 5163945 | - | bf638r_4360 | aroQ | 0 | GGTGCAAAGGTAGAAAAAAGAAATGA AATTAGAGGAAAAAACAAAAAAGAA | 1 |
| 5163930 | + | bf638r_4361 | xerD | 0 | ATTTGTATCCTCATAACACAATCCTGCT ATTTTTAATACCTTTGTCTTCGA | 1 |
| 5164953 | + | bf638r_4362 | unknown | 0 | AAAAAGCCGTTTCGGTAAACAATAGA GGCTGATTTGGTGTTGTTTAAGAAG | 1 |
| 5178847 | - | bf638r_4371 | unknown | 0 | GGTTTCAAAATTACGGAAAAGAGTCAT AATTCTGATAACTTTTAAGAAATA | 1 |
| 5209276 | + | bf638r_4408 | unknown | 0 | GAGCTGGAATGAAAACATGTGTTTTTG AATAAATTTCGATTTTTGCAAAAT | 1/2 |
| 5219062 | + | bf638r_4418 | unknown | 0 | ATTCAAGTTTTTATATCCGTAATCCCGG CAATAATCCCTATATTTGCCAAT | 1 |
| 5220380 | + | | | 1 | TAGTATTCTAATATTTGTCGTTCCAACT ATTTCCGTTATCTTTGCAATCAA | 1 |
| 5221227 | - | bf638r_4419 | dihydrofolate reductase | 0 | CACAAAGGTACGGGAAACTTCCCGAA TAACAAAACAATAGCCCTGCCTAAT | 1 |
| 5236429 | - | | | 1 | AGGAACAAATATACGTATATTATGATA ATATACTATAATTAAAGTGAAAAA | 1 |
| 5242238 | - | bf638r_4437 | TonB-linked OMP | 0 | AGTTTGCAAAATTAACGGTTTTAATCA TCACTCGCAATACCTACAAATAAT | 1 |
| 5250746 | + | | | 1 | AAAATCTCCTGTTTTTTTCTTTTTCGGT CAAAATTAGTATATTTGTATCAG | 1 |
| 5259326 | - | bf638r_4449 | acetyl coenzyme A synthetase | 0 | ATTCACAAATATATTGTTTTAGAAACA ACCGACAAACAAATCTCACCTAAA | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 5260928 | - | bf638r_4450 | unknown | 0 | TTCCAAAGATATATCGTTTTTTGCAATT ATCCGGAGAAAATACAGCCTTTT | 1 |
| 5264408 | + | bf638r_4456 | metG | 0 | CATCCATCTTATTTATAAAGAATTCCTC AGAAAGTCGTTATCTTTGCACGT | 1 |
| 5271471 | + | bf638r_4461 | unknown | 0 | TTATGCAGTAAAAATGGAGTACGACC AAAAAGAATGATTATCTTTAGGCAG | 1/2 |
| 5274611 | - | bf638r_4463 | UDP-N-acetyl-D-mannosamine dehydrogenase | 0 | GCCGCAAAAGTACAAATAATATGTGA CTTTGTACGGTTTCTCACCGAAAAC | 1 |
| 5274678 | + | bf638r_4464 | Glycosyltransferase | 0 | AATATGTGACTTTGTACGGTTTCTCAC CGAAAACCACTACTTTTGCCAAAA | 1 |
| 5277774 | + | bf638r_4467 | fklB | 0 | TTTAAAACCAAAAAGTAAAATCATTCA AATAAAACACTTACTTTTGCCGCT | 1 |
| 5279465 | + | bf638r_4469 | asnC | 0 | TTTTCTGACTTTTATATAGAGAAATTAA AATAAATCTCTATATTTGCTTAC | 1 |
| 5280354 | + | | tRNA | 0 | TTGACTAAAATGCTTGCATAATTCATC AGAAAGCATTAATTTTGTCCCCGC | 1 |
| 5282309 | - | bf638r_4472 | Two component regulatory system | 0 | GCGAACAAAGCTATCATATTTCGGGTA ATTATCATAAACTATTTAATATTT | 1 |
| 5293542 | - | | rRNA | 0 | GGATGCAAAAGTAAGGCATTAGGAGA TAAGAGACAAATATAAAAGGGGATT | 1 |
| 5294434 | + | bf638r_4476 | D-ala-D-ala dipeptidase | 0 | CTCCCATAACTTATCAAATATGCGCTTA TGTTTTTATACTTTTGCATTCAT | 1 |
| 5305199 | - | bf638r_4482 | Sugar transporter | 0 | GCAACAAAGATAAGACGCTTTTTTCAT ATGTGCAAATATTTCTACATTATT | 1 |
| 5305317 | + | bf638r_4483 | unknown | 0 | ATCTCTTCTTCTTATATTCAGTGCACAG AGCCAAAACAGTGTTTCAAAAAA | 0 |
| 5320761 | - | bf638r_4491 | unknown | 0 | GTGCCAAACATACAGATTGTTTTTGAC ATAGCATAGATCGAAATCATTATA | 1 |
| 5320822 | + | bf638r_4492 | calcium/sodium:proton antiporter | 0 | ACAGATTGTTTTTGACATAGCATAGAT CGAAATCATTATATTTGCACCCGA | 1 |
| 5321874 | + | bf638r_4493 | unknown | 0 | TTAGTTTTCGTACGTGGTAATTTAAAG ATTATCACGTATCTTTGCCCCCTC | 1 |
| 5322922 | + | bf638r_4494 | parE | 0 | TTTTCCGATTTAATTCCGGTTTATCAAA AATTAGGTTTTATCTTTGCAGAA | 1 |
| 5327728 | - | bf638r_4497 | unknown | 0 | CAGCAAAGATAGATTAAAAAGACAGA GTTTTCGCAAAAAAAACCTATCTTT | 1 |
| 5327787 | + | bf638r_4498 | transcriptional regulator | 0 | TAGATTAAAAAGACAGAGTTTTCGCAA AAAAAACCTATCTTTGTTCCCTAA | 1 |
| 5329308 | + | bf638r_4499 | succinate dehydrogenase (or fumarate reductase) cytochrome b subuni | 0 | ATTTCTTTAATTCTTTGGTTAATCAATG TATTTCGTGTACTTTTGCACAAT | 1 |
| 5337717 | - | bf638r_4504 | TonB-linked OMP | 0 | AGTCCGCAAAGGTACGTATATCGACAC CGGACGACAATCACCATATTTAGG | 1 |
| 5338706 | - | bf638r_4505 | Transcriptional regulator | 0 | ACGACAAATGTAAGAGAATTATTTAGA AAGACAATAAAAGAAGGAATAAAA | 1 |
| 5342315 | + | | | 1 | TATTAAGAAGAATATTTGGATAATAGA TAAGAGAAACCTACATTTGTACAG | 1 |
| 5345706 | - | bf638r_4512 | unknown | 0 | AGGTGCAAAGGTAGGGGTTTTTTCCG TACTTTTGCAGCCAAATAATGAATA | 1 |
| 5345746 | + | bf638r_4513 | Dihydrouridine synthase | 0 | TAAATGAATGAGGTGCAAAGGTAGGG GTTTTTTCCGTACTTTTGCAGCCAA | 1 |
| 5348043 | - | bf638r_4514 | unknown | 0 | GCGAGCAAAGATACGACTTTTAGCAC AATTAACGGACATTTGACGTGAATA | 1 |
| 5349803 | - | bf638r_4516 | cation efflux pump | 0 | GAGGGGCAAATGTACATGAAAATCTT GAATTATTAATATTTAATCGCTTAC | 1 |
| 5349863 | + | bf638r_4517 | ribonuclease R | 0 | ATGTACATGAAAATCTTGAATTATTAA TATTTAATCGCTTACTTTGTACTG | 1/2 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 5353979 | + | bf638r_4522 | Pyridoxamine 5'-phosphate oxidase | 0 | ATGCCCTGTTGCTGCACGGATATCTAAAATAAATTTCCATCTTTGTAGTTC | 1/2 |
| 5356501 | - | bf638r_4524 | Esterase/lipase | 0 | TGGGCAAATATACTACAATAACGACACTGTTTGAAAAATAATGCGTTCCTT | 1 |
| 5356622 | + | bf638r_4525 | unknown | 0 | AGTATAGAGAAATAACTGTATTAATCCCAATTTAATTACCTTTGTCAATAT | 1 |
| 5358148 | + | bf638r_4526 | YchF | 0 | GAGAAAGCAGATGACTGCATTAGCCTAAAATTGATTATCTTTGCACCCCAA | 1 |
| 5359613 | + | bf638r_4527 | apbA_panE | 0 | CTCCCGGTTTTCTACTACTATCTTCAAGCAAATCATTATCTTTGCGACCGG | 1 |
| 5364713 | - | bf638r_4530 | mutS | 0 | AAGGCAAAGATAGCGGTTTCCCTCAAATCTACCTATCCTAACAGACAGTTT | 1 |
| 5364947 | + | | | 1 | TATTATAAAACACCGATTTTTATCCGAATTACTTGCTATTTTCGCAGGAAG | 1/2 |
| 5364964 | + | | | 1 | TTTTATCCGAATTACTTGCTATTTTCGCAGGAAGTCGTACCTTTGCAATGT | 1 |
| 5365193 | + | | | 1 | CGTGTTATAAAACATACTAATTTACTTGGTTTATTTGCTATTTTCGCAGAA | 1/2 |
| 5365212 | + | | | 1 | ATTTACTTGGTTTATTTGCTATTTTCGCAGAAAGCCGTACCTTTGCAATGT | 1 |
| 5365389 | + | bf638r_4531 | unknown | 0 | TTGGTCGGTTAATGATTTCCCGATAATTTAAATACCTATATTAGATCCCTT | 1/2 |
| 5366653 | - | | | 1 | ATCAACAAATATAAGATTTTTATTTTCTGATTCTTCCTTTTCATACCTATT | 1 |
| 5366622 | - | | | 1 | GATTACAAATATAAATACTAAAATTTAAATTATCAACAAATATAAGATTTT | 1 |
| 5366760 | + | bf638r_4532 | leuS | 0 | TTTAGCTATCTTCGTGGTTATCTGAAAATAATCATGTAAATTTGCTCCCTC | 1 |
| 5372929 | - | bf638r_4535 | nadA | 0 | AAAGTGCAAAGTTAAAAACTTTATGGATATAAACAGGCTAATTAGTACTGA | 1 |

# Appendix 4: INSTITUTIONAL ANIMAL CARE AND USE COMMITTEE (IACUC) APPROVAL FORM

East Carolina University

**Animal Care and Use Commitee**
212 Ed Warren Life Sciences Building
East Carolina University
Greenville, NC 27834

**252-744-2436** office
**252-744-2355** fax

October 9, 2012

C. Jeffrey Smith, Ph.D.
Department of Micro/Immuno
Brody 5E-106
ECU Brody School of Medicine

Dear Dr. Smith:

Your Animal Use Protocol entitled, "Role of B. Fragilis Oxygen Stress Response in Infection" (AUP #K155a) was reviewed by this institution's Animal Care and Use Committee on 10/9/12. The following action was taken by the Committee:

"Approved as submitted"

***Please contact Dale Aycock at 744-2997 prior to hazard use***

A copy is enclosed for your laboratory files. Please be reminded that all animal procedures must be conducted as described in the approved Animal Use Protocol. Modifications of these procedures cannot be performed without prior approval of the ACUC. The Animal Welfare Act and Public Health Service Guidelines require the ACUC to suspend activities not in accordance with approved procedures and report such activities to the responsible University Official (Vice Chancellor for Health Sciences or Vice Chancellor for Academic Affairs) and appropriate federal Agencies.

Sincerely yours,

Susan McRae, Ph.D.
Chair, Animal Care and Use Committee

SM/jd

enclosure

**EAST CAROLINA UNIVERSITY**
**ANIMAL USE PROTOCOL (AUP) FORM**
**LATEST REVISION JULY, 2012**

**Project Title:**

Click here to enter text.
Role of B. Fragilis Oxygen Stress Response in Infection

| | Principal Investigator | Secondary Contact |
|---|---|---|
| **Name** | C. Jeffrey Smith | Edson R. Rocha |
| **Office Ph #** | 4-2700 | 4-9563 |
| **Cell Ph #** | 252-714-8466 | |
| **Pager #** | | |
| **Home Ph #** | 252-756-8131 | |
| **Email** | smithcha@ecu.edu | rochae@ecu.edu |

***For IACUC Use Only***

| | | | |
|---|---|---|---|
| AUP # K155a | | | |
| New/Renewal Renewal 10/5/12 | | | |
| Full Review/Date | DR/Date | | |
| Approval Date 10/9/12 | | | |
| Study Type B. fragilis | | | |
| Pain/Distress Category D | | | |
| Surgery ✓ | Survival ✓ | Multiple | abdominal ocm |
| Prolonged Restraint | | | |
| Food/Fluid Regulation | | | |
| Other | | | |
| Hazard Approval/Dates | Rad | IBC ✓ bacteroides fragilis | EHS |
| OHP Enrollment | | | |
| Mandatory Training | | | |
| Amendments Approved | | | |
| | | | |

## I. Personnel

**A. Principal Investigator(s):**
Charles Jeffrey Smith

**B. Department(s):**
Microbiology & Immunology

**C. List all personnel (PI's, co-investigators, technicians, students) that will be working with live animals and describe their qualifications and experience with these specific procedures. If people are to be trained, indicate by whom:**

| Name | Role(s) and Responsibilities for this Project | Required ECU Training (Yes/No) | Other Relevant Animal Experience/Training |
|---|---|---|---|
| Charles Jeffrey Smith | PI | yes | 8 years of animal research and completed the ECU animal welfare training |
| Edson R. Rocha | Co-PI | yes | 8 years of experience and training in performing the procedures detailed in this application. The PI has developed the implanted tennis "ping pong" ball in the rat peritoneal cavity to study B. fragilis intra-abdominal infection. The PI is acquainted with animal care procedures. Completed ECU training. |
| Yanlu Cao | graduate student | yes | several months experience with this model. Has received training from Dr. Rocha and has completed the online animal welfare course. |
| Matthew Rosenbaum | Consultant/Collaborator | yes | Veterinarian, Assistant Professor of Comparative Medicine |
| | | | |
| | | | |
| | | | |

| | | | |
|---|---|---|---|
| | | | |
| | | | |

## II. Regulatory Compliance

### A. Non-Technical Summary

**Using language a non-scientist would understand, please provide a clear, concise, and sequential description of animal use. Additionally, explain the overall study objectives and benefits of proposed research or teaching activity to the advancement of knowledge, human or animal health, or good of society. (More detailed procedures are requested later in the AUP.)**

***Do not cut and paste the grant abstract.***

Bacteroides fragilis is the most frequent pathogen isolated from anaerobic infections in humans such as intra-abdominal or pelvic abscesses. In the present study we propose experiments to identify and characterize the genes involved in the pathogenicity and virulence of B. fragilis. The animal experiments will specifically test the ability of B. fragilis mutants to survive in the rat peritoneal cavity. In addition, we will test the expression of virulence genes during the growth of these organisms inside the peritoneal cavity. To accomplish this, rats will be surgically implanted with an intra-abdominal tissue cage as a model for intra-abdominal infection. The cage will be inoculated with bacterial strains and then periodically sampled to determine viability of bacteria and gene expression. We also will perform "competition" studies in which both the wild-type and mutant strains are inoculated into the same 'ping pong ball. Then we will measure the differential survival over time. The idea here is that, in some cases differences in growth rate and final population level of the mutants may not be great enough to easily distinguish from the wild-type parent strain except in the case of very severe defects. Thus we will use mixed culture "competition" experiments plus the mono-culture experiments for a more complete picture of the role of virulence factors.

### B. Duplication

**Does this study duplicate existing research?** No

**If yes, why is it necessary? (note: teaching by definition is duplicative)**

Click here to enter text.

### C. Alternatives to the Use of Live Animals

**Are there less invasive procedures, other species, isolated organ preparation, cell or tissue culture, or computer simulation that can be used in place of the live vertebrate species proposed here?** No

**If yes, please explain why you cannot use these alternatives.**

Click here to enter text.

## D. Literature search to ensure that there are no alternatives to all potentially painful and/or distressful procedures

**1. Please list the potentially painful or distressful procedures in the protocol:**

surgery, implantation of tissue cage (ping pong ball), inoculation of tissue cage with bacteria, application of anesthesia and subsequent sampling of the tissue cage using a syringe and needle to aspirate material from inside the tissue cage, euthanasia of animals at the end of the experiment.

**2. For the procedures listed above, provide the following information (please do not submit search results but retain them for your records):**

| | |
|---|---|
| **Date Search was performed:** | August 16, 2012 |
| **Database searched:** | Medline via ECU OVID program |
| **Period of years covered in the search:** | 1946-August 2012 |
| **Keywords used and strategy:** | Bacteroides fragilis; experimental intra-abdominal infection, experimental model of anaerobic infections, animal models, and alternatives. I performed an individual search for each of the keywords above then combined each of the data sets with the connector "And"<br><br>Also searched the database with the entire phrase without the punctuation. |
| **Other sources consulted:** | NCBI Pubmed and Google |

**3. Narrative indicating the results of the search (2-3 sentences) and explaining why there are no alternatives to your proposed procedures that have the potential to cause pain and/or distress:**

There are no alternative procedures to in vivo experimental intra-abdominal infection. The procedure using tissue cage implanted in the rat abdominal cavity will reduce the number of animals required for this project by allowing multiple samples from the same animal. The use of culture cell models will not provide the complex immune system response that is essential for the recruitment of PMNs and macrophages in the initial stages of B. fragilis infection and abscess formation. The use of adult Sprague Dawley rats is appropriate for the tissue cage "pig-pong-tennis ball" model of intra-peritoneal infection. In addition, this model will allow us to obtain multiple sampling of intra-

abdominal exudates without the need to sacrifice the animal in each time point. This procedure will greatly reduce the number of animals to be used in this study. This model has advantages over direct inoculation of bacteria into the peritoneal cavity because the infection process is confined to inside of the encapsulated implanted ping-pong ball tissue cage. Our experience (Rocha and Smith combined) using this model with more than 75 rats is that we have observed six systemic infections due and have lost one animal to causes we could not directly attribute to the procedures. Generally the animals do not exhibit any signs of discomfort. Other methods of abscess formation do not allow multiple sampling at different time points and would require a large number of animals for each experiment.

## E. Hazardous agents

**1. Protocol related hazards (chemical, biological, or radiological):**
**Please indicate if any of the following are used in animals and the status of review/approval by the referenced committees:**

| HAZARDS | Oversight Committee | Status (Approved, Pending, Submitted)/Date | AUP Appendix I Completed? |
|---|---|---|---|
| **Radioisotopes** | **Radiation** | NA | |
| **Ionizing radiation** | **Radiation** | NA | |
| **Infectious agents (bacteria, viruses, richettsia, prions, etc.)** | **IBC** | approved 11/2009 | yes |
| **Toxins of biological orgins (venoms, plant toxins, etc.)** | **IBC** | NA | |
| **Transgenic, Knock In, Knock Out Animals---breeding, cross breeding or any use of live animals or tissues** | **IBC** | NA | |
| **Human tissues, cells, body fluids, cell lines** | **IBC** | NA | |
| **Viral/Plasmid Vectors/Recombinant DNA or recombinant techniques** | **IBC** | approved 11/2009 | yes |
| **Oncogenic/toxic/mutagenic chemical agents** | **EH&S** | NA | |
| **Nanoparticles** | **EH&S** | NA | |

| | | | |
|---|---|---|---|
| **Cell lines, tissues or other biological products injected or implanted in animals** | **DCM** | NA | |
| **Other agents** | | NA | |

**2. Incidental hazards**

**Will personnel be exposed to any incidental zoonotic diseases or hazards during the study (field studies, primate work, etc)? If so, please identify each and explain steps taken to mitigate risk:**

Bacteroides fragilis is a normal gut flora bacterium. This is not a contagious agent and is classified at BSL-1. The only chance of infection is from direct injection of the organism and even in that case during my 30+ years of working with these organisms I do not know of any case of laboratory acquired infection. However, since these organisms have been genetically modified, the experiments will be conducted under ABSL-2 conditions to minimize chances of contamination. A Biological Safety Registration has been approved for this work.

## III. Animals and Housing

**A. Species and strains**:

Sprague Dawley outbred rats.

**B. Weight, sex and/or age:**

>350g/11-13 weeks old male.

**C. Animal numbers:**

**1. Please complete the following table:**

| **Total number of animals in treatment and control groups** | **Additional animals (Breeders, substitute animals)** | **Total number of animals used for this project** |
|---|---|---|
| 198 | + 2 | =200 |

**2. Justify the species and number (use statistical justification when possible) of animals requested:**

The species chosen for these experiments is the adult Sprague Dawley rat. The use of these animals is appropriate for the tissue cage "pig-pong-tennis ball" model of intra-abdominal abscess infection because of their size and historical use of rats in intra-abdominal infection studies. In addition, this model will allow us to obtain multiple sampling of intra-abdominal exudates without the need to sacrifice the animal in each time point. This procedure will reduce the number of animals to be used in this study. There is

no alternative in vitro model to study abscess formation or response to an abscess inducing organism such B. fragilis because no in vitro system has been able to duplicate the complex mammalian immune response. Also, this model has been used previously with a different species of bacteria (Bamberger et al. 2002, Antimicrob. Agents Chemother, 46:2878-2884).

There are two factors that must be considered to determine the number of animals needed for the study. The first factor is that we have a maximum of 8 candidate genes that we propose to test as virulence factors. For each candidate gene we will construct the corresponding mutant strain thus we will need to have 8 independent trials (including both competition assays and monoculture assay). In addition each trial will be composed of two identical experiments in order to demonstrate that the results are repeatable. The second factor is that we must determine the lowest number of animals needed in each experiment that can provide statistical significance for our results. Based on previous work with this model by Dr. Rocha and myself (see AUP #K146 and# K155) we can estimate the number of wild-type or mutant bacteria found per ml of fluid in the artificial abscess and use these numbers in power calculations to determine the sample size needed for statistical significance in each experiment.For these calculations we will be using a study design based on use of the unpaired t-test on group means. We have set P = .01 and power probability = 0.9. The bacterial numbers from previous work used for this analysis are:

wild type strain: 3.28 x 10exp9 (SD 2.07 x 10exp9) mutant strain: 6.43 x 10exp7 (SD 3.56 x 10exp7). Based on these numbers we have calculated that statistical significance can be achieved using groups of 5 animals and we do not see any difference in using the competition or monoculture approaches. In experimental design section (section IV.A) we present a table for the total number of animals to be used based on groups of 5 animals per experiment. Briefly, we propose to study 8 different mutants so that is 10 animals per mutant in monoculture experiments and 10 animals per mutant in competition assay experiment plus one uninoculated control per mutant for a total of 178 animals. The remaining animals are accounted for by the need to run two trials of monoculture experiments with the wild type parent strain to obtain gene expression data to use as baseline for when we perform the competition assays (20 animals).

**3. Justify the number and use of any additional animals needed for this study:**

**a. For unforeseen outcomes/complications:**

two additional animals should be ordered if there are any unexpected deaths that lead to invalid statistics.

**b. For refining techniques:**

n/a

**c. For breeding situations, briefly justify breeding configurations and offspring expected:**

n/a

**d. Indicate if following IACUC tail snip guidelines:** Choose an item.

**(if no, describe and justify)**

Click here to enter text.

**4. Will the phenotype of mutant, transgenic or knockout animals predispose them to any health, behavioral, physical abnormalities, or cause debilitating effects in experimental manipulations?** No **(if yes, describe)**

Click here to enter text.

**5. Are there any deviations from standard husbandry practices?**

Yes **If yes, then describe conditions and justify the exceptions to standard housing (temperature, light cycles, sterile cages, special feed, prolonged weaning times, wire-bottom cages, etc.):**

Animals will be housed individually on soft bedding. Tubes are not allowed due to surgery so alternative enrichments such as softer bedding, crinkle nesting will be provided.

**6. Is it necessary for animals to be singly housed?**

No **(If yes, describe housing and justify the need to singly house social species)**

Click here to enter text.

**7. Are there experimental or scientific reasons why routine environmental enrichment should not be provided?** No

**(If yes, describe and justify the need to withhold enrichment)**

Click here to enter text.

**8. If wild animals will be captured or used, provide permissions (collection permit # or other required information):**

NA

**9. List all laboratories or locations outside the animal facility where animals will be used. Note that animals may not stay in areas outside the animal facilities for more than 12 hours without prior IACUC approval. For field studies, list location of work/study site.**

none

## IV. Animal Procedures

**A. Outline the Experimental Design including all treatment and control groups and the number of animals in each. Tables or flow charts are particularly useful to communicate your design.**

The objective is to determine if various mutant strains lacking various oxidative stress resistance genes can survive as well in the abscess as the wild type parent strain. We also need to

determine the overall gene expression patterns of the bacteria in the abscess and if the oxidative stress gene mutations modulate these patterns. For this work we have a bank of 8 mutant strains or proposed mutant strains that lack a variety of genes known to be important for in vitro resistance to oxidative stress. We now will test the role of these genes in vivo using the rat model. Briefly we will implant a sterile tissue cage into the peritoneal cavity of the rat. Then allow the animals to recover for 3-4 weeks to allow for encapsulation to occur according to Bamberger et al. 2002 (Antimicrob. Agents Chemother, 46:2878-2884). Next five rats will be inoculated with about 1,000,000 c.f.u./ml B. fragilis wild type strain in monoculture (Control) by injection into the implanted tissue cage. Samples will be aspirated from infected cages at 2, 4, 8, and 15 days post-infection. These control experiments will be repeated four times over the course of the entire project. For each mutant strain five rats will be used in monoculture experiments to determine survival and gene expression patterns. Also for each mutant strain 5 animals will be used in competition experiments in which mutant and wild-type strain are co-cultured in the tissue cage in order to determine if the mutant has a defect relative to the parent. The idea here is that, in the absence of competition, differences in growth rate and final population level may not be great enough to distinguish at a statistically significant level except in the case of very major defects. The competition assay is a tried and true method in bacterial pathogenesis studies. These experiments with the mutant will be repeated once. One uninoculated control will be used for each mutant strain and 2 for the wild type parent strain. Two animals may be needed to protect against adverse outcomes due to unforeseen complications (See Table Below).

| Bacterial Strain/treatment | Animals per group | # experiments | # trials | total # animals |
|---|---|---|---|---|
| **Wild-type monoculture** | **5** | **2** | **2** | **20** |
| Mutant monoculture | 5 | 2 | 8 | 80 |
| **Competition assay** | **5** | **2** | **8** | **80** |
| Added for unforeseen complications | 1 | 2 | 1 | 2 |
| **Uninoculated control** | **1** | **1** | **18** | **18** |
| Totals | | | | 200 |

***In sections IV.B-IV.S below, please respond to all items relating to your proposed animal procedures. If a section does not apply to your experimental plans, please leave it blank.***

***Please refer to DCM and IACUC websites for relevant guidelines and SOPs.***

**B. Anesthesia/Analgesia/Tranquilization/Pain/Distress Management For Procedures Other than Surgery:**
***Adequate records describing anesthetic monitoring and recovery must be maintained for all species.***

**If anesthesia/analgesia must be withheld for scientific reasons, please provide compelling scientific justification as to why this is necessary:**
Click here to enter text.

1. **Describe the pre-procedural preparation of the animals:**
   a. **Food restricted for** Not needed **hours**
   b. **Food restriction is not recommended for rodents and rabbits and must be justified:**
   n/a
   c. **Water restricted for** not needed **hours**
   d. **Water restriction is not recommended in any species for routine pre-op prep and must be justified:**
   n/a

2. **Anesthesia/Analgesia *for Procedures Other than Surgery***

| | Agent | Concentration | Dose (mg/kg) | Max Volume | Route | Frequency | Number of days administered |
|---|---|---|---|---|---|---|---|
| Pre-procedure analgesic | | | | | | | |
| Pre-anesthetic | | | | | | | |
| Anesthetic | isoflurane | 2-4% | | | inhalation | 6 | <5 min |
| Post procedure analgesic | | | | | | | |
| Other | | | | | | | |

**3. Reason for administering agent(s):**
To sedate rats for injection and aspiration.

**4. For which procedure(s):**
Sample injection into and aspiration from the implanted ping pong ball. In all cases aseptic technique will be used for these procedures. The skin is prepped prior to insertion of the needle by swabbing with alcohol.

**5. Methods for monitoring anesthetic depth:**
Foot withdraw

**6. Methods of physiologic support during anesthesia and recovery:**
n/a

**7. Duration of recovery:**
< 5 min

**8. Frequency of recovering monitoring:**
Click here to enter text.

**9. Specifically what will be monitored?**
respiration and mobility

**10. When will animals be returned to their home environment?**
when they are conscious and mobile

**11. Describe any behavioral or husbandry manipulations that will be used to alleviate pain, distress, and/or discomfort:**
Animals typically do not exhibit distress or discomfort following procedure

## C. Use of Paralytics

**1. Will paralyzing drugs be used?** NO

**2. For what purpose:**
Click here to enter text.

**3. Please provide scientific justification for paralytic use:**
Click here to enter text.

**4. Paralytic drug:**
Click here to enter text.

**5. Dose:**
Click here to enter text.

**6. Method of ensuring appropriate analgesia during paralysis:**
Click here to enter text.

## D. Blood or Body Fluid Collection

**1. Please fill out appropriate sections of the chart below:**

| | Location on animal | Needle/catheter size | Volume collected | Frequency of procedure | Time interval between collections |
|---|---|---|---|---|---|
| **Blood Collection** | | | | | |
| **Body Fluid Collection** | Intra-abdominal from within the tissue cage | 25-20 g | 0.5-2 ml | once per sampling period | 2, 4, 8, 15 days post infection |
| **Other** | | | | | |

## E. Injections, Gavage, & Other Substance Administration

**1. Please fill out appropriate sections of the chart below:**

| | Compound | Location & Route of admin | Needle/catheter/gavage size | Max volume admin | Freq of admin (ie two times per day) | Number of days admin (ie for 5 days) | Max dosages (mg/kg) |
|---|---|---|---|---|---|---|---|
| **Injection/ Infusion** | Live Bacteroides fragilis. 4 ml at $10^6$ to $10^8$ cells/ml | Intra-abdominal injection into the tissue cage | 25-20 g | 4 ml | once to initiate the experiment | n/a | n/a |
| **Gavage** | | | | | | | |
| **Other** | | | | | | | |

**2. For all injections and infusions, PHARMACEUTICAL GRADE compounds should be used. If not available, refer to IACUC Guidelines for non-pharmaceutical grade compound use and provide required information below:**
n/a

## E. Prolonged restraint with mechanical devices

**Prolonged restraint in this context means *beyond routine care and use procedures* for rodent and rabbit restrainers, and large animal stocks.**

**Prolonged restraint also includes *any* use of slings, tethers, metabolic crates, inhalation chambers, primate chairs and radiation exposure restraint devices.**

1. **For what procedure(s):**

n/a

2. **Explain why non-restraint alternatives cannot be utilized:**

n/a

3. **Restraint device(s):**

n/a

4. **Duration of restraint:**

n/a

5. **Frequency of observations during restraint/person responsible:**

n/a

6. **Frequency and total number of restraints:**

n/a

7. **Conditioning procedures:**

n/a

8. **Steps to assure comfort and well-being:**

n/a

9. **Describe potential adverse effects of prolonged restraint and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

## G. Tumor Studies, Disease Models, Toxicity Testing, Vaccine Studies, Trauma Studies, Pain Studies, Organ or System Failure Studies, Shock Models, etc.

1. **Describe methodology:**

The wild type and mutant bacteria will be injected into the sterile, artificial abscess and their ability to survive in the abscess will be determined. Samples will be taken on days 2,4,8, and 15 by aspiration using a sterile syringe and then plated on bacteriological growth media to determine the viable cell count. Some material also will be used for the extraction of RNA to be used in gene expression studies to determine which genes are expressed by the bacteria during the course of "infection".

2. **Expected model and/or clinical/pathological manifestations:**

It is expected that wild type bacteria will survive at a higher rate than mutant bacteria in the artificial abscess in the co-culture competition assays and we expect to see more rapid

clearing of the mutant strains in the monoculture assays. We do not expect any obvious tissue pathology.

**3. Signs of pain/discomfort:**

We have had several years of experience with this model and during this time seven animals became "sick" and displayed a hunched posture, depressed attitude and loss of appetite. These animals were provided analgesics but eventually had to be euthanized. This is a rare occurrence.

**4. Frequency of observations:**

daily

**5. Describe potential adverse side effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

If animals show signs of peritonitis, hunched posture, depressed attitude and loss of appetite, as described above we will first try to treat with analgesics but if no change after a day they will be euthanized.

## H. Treadmills/Swimming/Forced Exercise

**1. Describe aversive stimulus (if used):**

n/a

**2. Conditioning:**

n/a

**3. Safeguards to protect animal:**

n/a

**4. Duration:**

n/a

**5. Frequency:**

n/a

**6. Total number of sessions:**

n/a

**7. Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

## I. Projects Involving Food and Water Regulation or Dietary Manipulation

(*Routine pre-surgical fasting not relevant for this section)*

**1. Food Regulation**

**a. Amount regulated and rationale:**

n/a

**b. Frequency and duration of regulation (hours for short term/weeks or months for long term):**

n/a

**c. Frequency of observation/parameters documented (i.e. recording body weight, body condition, etc.):**

n/a

**d. Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

**2. Fluid Regulation**

**a. Amount regulated and rationale:**

n/a

**b. Frequency and duration of regulation (hours for short term/weeks or months for long term):**

n/a

**c. Frequency of observation/parameters documented (body weight, hydration status, etc.):**

n/a

**d. Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

**3. Dietary Manipulations**

**a. Compound supplemented/deleted and amount:**

n/a

**b. Frequency and duration (hours for short term/week or month for long term):**

n/a

c. **Frequency of observation/parameters documented:**
n/a

d. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**
n/a

## J. Endoscopy, Fluoroscopy, X-Ray, Ultrasound, MRI, CT, PET, Other Imaging

1. **Describe animal methodology:**
n/a
2. **Duration of procedure:**
n/a
3. **Frequency of observations during procedure:**
n/a
4. **Frequency/total number of procedures:**
n/a
5. **Method of transport to/from procedure area:**
n/a
6. **Describe potential adverse side effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**
n/a
7. **Please provide or attach appropriate permissions/procedures for animal use on human equipment:**
n/a

## K. Polyclonal Antibody Production

1. **Antigen/adjuvant used and justification for adjuvant choice:**
n/a
2. **Needle size:**
n/a
3. **Route of injection:**
n/a
4. **Site of injection:**

n/a

5. **Volume of injection:**
   n/a
6. **Total number of injection sites:**
   n/a
7. **Frequency and total number of boosts:**
   n/a
8. **What will be done to minimize pain/distress:**
   n/a
9. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**
   n/a

## L. Monoclonal Antibody Production

1. **Describe methodology:**
   n/a

2. **Is pristane used:** No
   **Volume of pristane:**
   n/a

3. **Will ascites be generated:** No
   i. **Criteria/signs that will dictate ascites harvest:**
      Click here to enter text.
   ii. **Size of needle for taps:**
      Click here to enter text.
   iii. **Total number of taps:**
      Click here to enter text.
   iv. **How will animals be monitored/cared for following taps:**
      Click here to enter text.
4. **What will be done to minimize pain/distress:**
   Click here to enter text.
5. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

Click here to enter text.

## M. Temperature/Light/Environmental Manipulations

1. **Describe manipulation(s):**
n/a
2. **Duration:**
n/a
3. **Intensity:**
n/a
4. **Frequency:**
n/a
5. **Frequency of observations/parameters documented:**
n/a
6. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**
n/a

## N. Behavioral Studies

1. **Describe methodology/test(s) used:**
n/a
2. **Will conditioning occur? If so, describe:**
n/a
3. **If aversive stimulus used, frequency, intensity and duration:**
n/a
4. **Length of time in test apparatus/test situation:** ***(i.e., each test is ~10 mins)***
n/a
5. **Frequency of testing and duration of study:** ***(i.e., 5 tests/week for 6 months)***
n/a
6. **Frequency of observation/monitoring during test:**
n/a
7. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

## O. Capture with Mechanical Devices/Traps/Nets

1. **Description of capture device/method:**
   n/a
2. **Maximum time animal will be in capture device:**
   n/a
3. **Frequency of checking capture device:**
   n/a
4. **Methods to ensure well-being of animals in capture device:**
   n/a
5. **Methods to avoid non-target species capture:**
   n/a
6. **Method of transport to laboratory/field station/processing site and duration of transport:**
   n/a
7. **Methods to ensure animal well-being during transport:**
   n/a
8. **Expected mortality rates:**
   n/a
9. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**
   n/a

## P. Manipulation of Wild-Caught Animals in the Field or Laboratory

1. **Parameters to be measured/collected:**
   n/a
2. **Approximate time required for data collection per animal:**
   n/a
3. **Method of restraint for data collection:**
   n/a
4. **Methods to ensure animal well-being during processing:**
   n/a
5. **Disposition of animals post-processing:**
   n/a

6. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

## Q. Wildlife Telemetry/Other Marking Methods

1. **Describe methodology (including description of device):**

   n/a
2. **Will telemetry device/tags/etc be removed?** n/a **If so, describe:**

   n/a
3. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

## R. Other Animal Manipulations

1. **Describe methodology:**

   n/a
2. **Describe methods to ensure animal comfort and well-being:**

   n/a
3. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

n/a

## S. Surgical Procedures

***All survival surgical procedures must be done aseptically, regardless of species or location of surgery. Adequate records describing surgical procedures, anesthetic monitoring and postoperative care must be maintained for all species.***

1. **Location of Surgery (Building & Room #):**

   ECU/SOM Animal care facility operating room

**2. Type of Surgery (check all that are appropriate):**

☐ **Non-survival surgery (animals euthanized without regaining consciousness)**

X☐ **Major survival surgery (major surgery penetrates and exposes a body cavity or produces substantial impairment of physical or physiologic function)**

☐ **Minor survival surgery**

☐ **Multiple survival surgery**

**If yes, provide scientific justification for multiple survival surgical procedures:**
n/a

**3. Describe the pre-op preparation of the animals:**

a. **Food restricted for** 0 **hours**
b. **Food restricted is not recommended for rodents and rabbits and must be justified:**
   n/a
c. **Water restricted for** 0 **hours**
d. **Water restriction is not recommended in any species for routine pre-op prep and be justified:**
   n/a

**4. Minimal sterile techniques will include (check all that apply):**

***Please refer to DCM Guidelines for Aseptic Surgery for specific information on what is required for each species and type of surgery (survival vs. non-survival).***

X☐ **Sterile instruments**

**How will instruments be sterilized?**
Ethylene oxide for ping pong balls, autoclave for instruments

**If serial surgeries are done, how will instruments be sterilized between surgeries:**
Instruments will be sterilized by autoclave. For serial surgery, instruments will be sterilized by dry hot beads sterilizer.

x☐ **Sterile gloves**
x☐ **Mask**
x☐ **Cap**
☐ **Sterile gown**
x☐ **Sanitized operating area**
x☐ **Clipping or plucking of hair or feathers**
x☐ **Skin preparation with a sterilant such as betadine**
x☐ **Practices to maintain sterility of instruments during surgery**
x☐ **Non-survival (clean gloves, clean instruments, etc.)**

**5. Describe all surgical procedures:**

**a. Skin incision size and site on the animal:**

An abdominal incision of about 3 cm

**b. Describe surgery in detail (include size of implant if applicable):**

Click here to enter text.

**c. Method of wound closure:**

The rats will be given pre-emptive analgesia Buprenex prior to the procedure and then anesthetized with gas (isoflurane) anesthesia. Ketamine and xylazine (9:1) may be substituted for gas if instructed by veterinarian. Aseptic technique (sterile instruments, surgical gloves, masks and surgical prep) will be used. An abdominal incision of about 3 cm will be performed and a single sterile table-tennis ball with 300 1.5 mm-diameter holes will be implanted in the peritoneal cavity by sterile techniques, abdominal closure with at least two layers, 3-0 absorbable followed by 3-0 nylon skin. Animal will receive buprenex at 0.1mls/100 grams bw (0.03mg/ml conc) or an NSAID carprofen or meloxicam for post-op analgesia and every 8-12 hours post as determined by DCM veterinarian. Animal will be allowed to recover, for 5 weeks to allow for encapsulation to occur according to Bamberger et al. 2002 (Antimicrob. Agents Chemother, 46:2878-2884). Rats will be housed under standard laboratory housing conditions and receive food and drink at libitum and support care for anesthetized rodents at the animal housing facility of the Department of Comparative Medicine, East Carolina University, Greenville, North Carolina. The rats will be inoculated with about 108 c.f.u./ml B. fragilis strains into the implanted ball tissue cage. Samples will be aspirated from infected cages at 2, 4, 8, and 15 or 21 days post-infection.

**i. Number of layers**

Abdominal closure with at least two layers

**ii. Type of wound closure and suture pattern:**

Interrupted, 3-0 absorbable, and 3-0 nylon

**iii. Suture type/size/wound clips/tissue glue:**

Click here to enter text.

**iv. Plan for removing of skin sutures/wound clip/etc:**

Skin sutures removed after 7 – 10 days

6. **Anesthetic Protocol:**

   a. **If anesthesia/analgesia must be withheld for scientific reasons, please provide compelling scientific justification as to why this is necessary:**

   n/a

   b. **Anesthesia/Analgesia For Surgical Procedures**

| | Agent | Dose (mg/kg or %) | Volume | Route | Frequency | Number of days administered |
|---|---|---|---|---|---|---|
| **Pre-operative analgesic** | Buprenex | 0.1 ml/100 g BW | 1 ml depending on BW | SC | once | prior to anesthesia |
| **Pre-anesthetic** | | | | | | |
| **Anesthetic** | isoflurane | 2-4% | | inhalation | once | 6 times; at insertion of tissue cage, inoculation day, sample days 2, 4, 8, 15 |
| **Post-operative Analgesic** | Buprenex | 0.1 mg/100g BW | 1 ml depending on BW | SC | once but more in indicated | 8-12 h |
| **Other** | Alternative Anesthetic - Alternative- ketamine plus xylazine (9:1)<br><br>Alternative Analgesic – meloxica | 0.1 ml/100g BW of Ketamine (90mg/ml) + Xylazine (10mg/ml)<br><br>1.5 mg/kg | 1 ml depending on BW<br><br>1 ml depending on BW | Ip<br><br>PO | once<br><br>SID, PRN | 30-60 min<br><br>24 hr |

   c. **Methods that will be used to monitor anesthetic depth (include extra measures employed when paralyzing agents are used):**

Standard procedures for anesthetic monitoring, surgical plane of anesthesia. Animals will be adequately anesthetized prior to initiation of surgery and then maintained in a surgical plane throughout procedure the will be done by monitoring toe pinch and pedal reflexes to ensure adequate depth. Pulse oximetry will monitor HR and O2 saturation. Respiration Rate will be monitored by watching thoracic cavity as well.

**d. Methods of physiologic support during anesthesia and immediate post-op period (fluids, warming, etc.):**

On pad to conserve body heat (surgery takes about 20 min per rat). If needed warm 0.9% NaCl will be administered SQ (3-7mL) along with soft and caloric dense food being accessible on the cage floor.

**e. List what parameters are monitored during immediate post-op period. Provide the frequency and duration:**

Following surgery procedures in the "tissue cage model of infection", animals will be examined for distress and pain on an hourly basis for 4 hours on the day of surgery. The animals will be kept warm, given fluid and nutritional supplementation as required until recovery.

**f. Describe any other manipulations that will be used to alleviate pain, distress, and/or discomfort during the immediate post-op period (soft bedding, long sipper tubes, food on floor, dough diet, etc.):**

If necessary, food will be placed on cage bottom, and gel packs used for water (usually not required) will be provided.

**g. List criteria used to determine when animals are adequately recovered from anesthesia and when the animals can be returned to their home environment:**

monitor continuously until mobile and alert.

**7. Recovery from Surgical Manipulations (after animal regains consciousness and is returned to its home environment)**

**a. What parameters (behavior, appetite, mobility, wound healing, etc.) will be monitored:**

Animals will be assessed for hunched posture, depressed attitude, off feed. Animals also will be monitored for overt clinical signs of illness (distress, sick, scruffy appearance, immobile, anorectic, moribund.

**b. How frequently (times per day) will animals be monitored:**

once

**c. How long post-operatively (days) will animals be monitored:**

Animals will be monitored for eating, drinking, defecating, behavior and appearance (body condition scoring) for 4 hours post operatively that day until normal signs appear. Then daily for about 4 weeks until encapsulation of the implanted tissue cage.

**8. Surgical Manipulations Affecting Animals**

**a. Describe any signs of pain/discomfort/functional deficits resulting from the surgical procedure:**

Hunched posture, depressed attitude, off feed, distress, sick-looking, scruffy appearance, immobile, anorectic, moribund.

b. **What will be done to manage any signs of pain or discomfort (include pharmacologic and non-pharmacologic interventions):**
Analgesic will be provided as described above in analgesia/anesthesia table.

c. **Describe potential adverse effects of procedures and provide humane endpoints (criteria for either humanely euthanizing or otherwise removing from study):**

Abscesses are a self-contained "walled" off infectious process. Nonetheless, intra-abdominal abscesses can eventually rupture and release its purulent pus contents into the peritoneal cavity. This may lead to a disseminated systemic infection. No antibiotic can be administered since these will interfere with experimental infections. Thus, any animal showing clinical signs of illness (distress, sick, scruffy appearance, immobile, anorectic, moribund) will be euthanized prior to the end of the experiment. We have done this procedure on more than 75 rats and have lost six to sepsis and one to an unknown cause (likely intestinal strangulation).

## V. Euthanasia

***Please refer to the AVMA Guidelines on Euthanasia and DCM Guidelines to determine appropriate euthanasia methods.***

A. **Euthanasia Procedure.** ***All investigators, even those conducting non-terminal studies, must complete this section in case euthanasia is required for humane reasons.***

**1. Physical Method- If a physical method is used, the animal should be first sedated/anesthetized with $CO_2$ or other anesthetic agent. If prior sedation is not possible, a scientific justification must be provided:**
n/a

**2. Inhalant Method**- Carbon Dioxide
**(if other, describe the agent and delivery method)**
Click here to enter text.

**3. Non-Inhalant Pharmaceutical Method (injectables, MS-222, etc.)- Please provide the following:**

a. **Agent:**
n/a
b. **Dose or concentration:**
n/a
c. **Route:**
n/a

25

**B. Method of ensuring death (can be physical method, such as pneumothorax or decapitation for small species and assessment method such as auscultation for large animals):**

check for signs of breathing and open the thoracic cavity

**C. Describe disposition of carcass following euthanasia:**

Each animal is placed in a disposal plastic bag and every three bags are wrapped up in another bag. Then all carcasses are put in a biohazard red bag, closed and they are put in the cold room container available at the animal house facility for disposal of carcasses.

**I acknowledge that humane care and use of animals in research, teaching and testing is of paramount importance, and agree to conduct animal studies with professionalism, using ethical principles of sound animal stewardship. I further acknowledge that I will perform only those procedures that are described in this AUP and that my use of animals must conform to the standards described in the Animal Welfare Act, the Public Health Service Policy, The Guide For the Care and Use of Laboratory Animals, the Association for the Assessment and Accreditation of Laboratory Animal Care, and East Carolina University.**

**Please submit the completed animal use protocol form via e-mail attachment to iacuc@ecu.edu. You must also carbon copy your Department Chair.**

**PI Signature:** C. Kelly Smith email **Date:** 10/5/12 ~~10/4/2012~~

**Veterinarian:** _____ **Date:** 10/9/12

**IACUC Chair:** S B McRae jh **Date:** 10/9/12

27

| APPENDIX 1-HAZARDOUS AGENTS | | | |
|---|---|---|---|
| Principal Investigator:<br>Charles Jeffrey Smith | Campus Phone:<br>744-2700 | | Home Phone:<br>252-756-8131 |
| IACUC Protocol Number:<br>K155 | Department:<br>Microbiology & Immunology | | E-Mail:<br>smithcha@ecu.edu |
| Secondary Contact:<br>Department: Edson Rocha | Campus<br>Phone: 744-9563 | Home<br>Phone: 756-8538 | E-Mail:<br>rochae@ecu.edu |
| Chemical Agents used: N/A | | Radioisotopes used: N/A | |
| Biohazardous Agents used:<br>Bacteroides fragilis | Animal<br>Biosafety Level: ABSL2 based on use of recombinant DNA | | Infectious to humans? Probably if large numbers are directly injected into the blood stream. |
| **PERSONAL PROTECTIVE EQUIPMENT REQUIRED:** | | | |
| Route of Excretion: There is no excretion of the organisms since they are confined to the artificial abscess | | | |
| Precautions for Handling Live or Dead Animals: No special precautions needed. Lab coats, eye protection, and gloves will be worn during injection of animals and sampling procedures. Bacteria closely related to B. fragilis are present in all rodent feces. . | | | |
| Animal Disposal: incinerate | | | |
| Bedding/Waste Disposal: normal procedures | | | |
| Cage Decontamination: normal procedures | | | |
| Additional Precautions to Protect Personnel, Adjacent Research Projects including Animals and the Environment: B. fragilis is a BSL-1 organism but due to the fact that we have genetically manipulated the organisms the experiments are elevated to ABSL-2 but risk of infection is minimal. Thus there are no special "Additional" precautions needed. Organisms closely related to B. fragilis are found in the feces of all rodents. | | | |
| **Initial Approval**<br>Safety/Subject Matter Expert Signature & Date<br><br>_____ | | | |

**The Brody School of Medicine**
**Office of Prospective Health**
East Carolina University
188 Warren Life Sciences Building • Greenville, NC 27834
252-744-2070 office • 252-744-2417 fax

COPY

Occupational Medicine
Employee Health
Radiation Safety
Infection Control
Biological Safety

TO: Dr. C. Jeffrey Smith
Department of Microbiology & Immunology

FROM: Eddie Johnson/John Williams
Biological Safety Officers

RE: Registration Approval

Date: December 10, 2009

Your Biological Safety Protocol J Smith, 09-01 "*Role of B. fragilis Oxygen Stress Response in Infection*" to use genetically modified micro organisms has received **approval** to be conducted under BSL-1 and ABSL-2 based on your registration submitted. This registration was approved by the ECU Biological Safety Committee with Dr. Greg Smith as acting chair.

This approval is effective for a period of 3 years and may be renewed with an updated registration if needed. Please notify the Animal Care staff before or if you begin work with Biohazard agents in animals. Also please keep in mind all individuals who will be exposed to or handle biohazardous agents in your work will be due for Blood Borne Pathogens refresher training annually.

Please do not hesitate to contact Biological Safety at 744-2070 if you have any questions, concerns, or need any additional information. Best wishes on your research.

cc: Dr. C. Jeffrey Smith, Chair, Biosafety Committee
Dr. Greg Smith, Community Member
Janine Davenport, IACUC
Dr. Robert Carroll, IACUC
Dale Aycock, IACUC