UNIVERSITY OF CENTRAL OKLAHOMA

Edmond, Oklahoma

Jackson College of Graduate Studies

Genetic Evaluation of Bile Salt Resistance and Sensitivity in a Subculture of *Escherichia coli* K-12 Substrain JC3272

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

In partial fulfillment of the requirements

for the degree of

MASTER OF SCIENCE IN BIOLOGY

By

Zane A. Henderson

Edmond, Oklahoma

2024

Genetic Evaluation of Bile Salt Resistance and Sensitivity in a Subculture of *Escherichia coli* K-12 Substrain JC3272

A THESIS

APPROVED FOR THE DEPARTMENT OF BIOLOGY

December 2024

By _____
Name Committee Chairperson

_____
Name Committee Member

_____
Name Committee Member

ACKNOWLEDGEMENTS

This thesis is based on work supported by a grant from the Office of High-Impact Practices, University of Central Oklahoma. Thank you to Dr. Phillip Silverman at the Oklahoma Medical Research Foundation for providing cultures, equipment, and supplies for this research project.

Thank you to Jakob Bakhtiari, Parker Brecheen, Victoria Ekogbulu, Ethan Strange, Jennifer Vega, and other members of the Bidlack Research Group for your contributions to this work, and to Dr. James Bidlack, Dr. Paul Olson, and Dr. Matthew Parks for your guidance.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

ABSTRACT

University of Central Oklahoma

Edmond, Oklahoma

NAME: Zane A. Henderson

TITLE OF THESIS: Genetic Evaluation of Bile Salt Resistance and Sensitivity in a Subculture of *Escherichia coli* K-12 Substrain JC3272

DIRECTOR OF THESIS: James E. Bidlack, Ph.D.

PAGES: 35

ABSTRACT: Though most *Escherichia coli* strains are adapted to survival in environments with high concentrations of bile salts, some mutant strains lack this ability. Prior work has reported that a spontaneous mutant of the K-12 strain JC3272 is sensitive to bile salts. To test this, cultures of this mutant (JC3272I) and a non-mutant strain (JC3272F) were evaluated on both solid and liquid media containing bile salts. Amplification of the *lapA* and *lapB* genes—which are hypothesized to affect bile salt sensitivity—was achieved in both strains via polymerase chain reaction. A combination of Sanger and Illumina sequencing was then used to compare the amplified genes and to produce whole-genome sequences. Culturing confirmed a difference in sensitivity to bile salts between the two strains. However, sequencing did not reveal any differences between the strains in the *lapAB* region. Furthermore, no mutations found via whole-genome sequencing had known roles in bile salt resistance. These results may imply the existence of an unpublished mechanism of bile salt resistance in *E. coli*.

INTRODUCTION

Bile salts are a group of cytotoxic detergents produced by the liver and involved in the digestion of lipids (Dawson 2012; Thanassi et al. 1997). Cholic and chenodeoxycholic acid are primary acids synthesized from cholesterol in the liver, and may be modified to form secondary bile acids by bacteria (Dawson 2012). Of interest is deoxycholic acid, which is formed via the 7α-dehydroxylation of cholic acid by several bacteria including *Escherichia coli* (Yoshimoto et al. 1991; Yoshimoto et al. 1993). After this modification, the bile salt pool consists of approximately 20% cholic acid, 20% chenodeoxycholic acid, 20% deoxycholic acid, and other secondary bile acids (Carulli et al. 2000). Bile salts are known to control the growth of bacteria in the small intestine both by triggering antimicrobial action from the host and by acting as a cytotoxin (Inagaki et al. 2006). Bile salts have been administered to rats to counter bacterial overgrowth with some success (Lorenzo-Zúñiga et al. 2003).

As a species that lives in the hindgut of animals, *E. coli* has evolved mechanisms to resist the effects of bile salt exposure. Several genes have been linked to bile salt resistance, including *ptsI* and *cyaA* (Le et al. 2020), and *E. coli* has been shown to use membrane damage as a proxy for bile salt detection (Clavel et al. 1996; Gunn 2000). Of special importance are efflux pump genes, including *acrAB* (Gunn 2000)—which is regulated by *marRAB* (Okusu et al. 1996) and *tolC* (Fralick 1996)—and *emrAB* (Gunn 2000), and it is likely that more efflux genes exist (Thanassi et al. 1997). Due to the overlap in mechanisms of resistance to bile salts and antibiotics (Gipson et al. 2020), an improved understanding of these and other genes may prove invaluable in combatting

antimicrobial resistant bacteria, which are a major cause of mortality globally (Murray et al. 2022).

Lipopolysaccharides are glycolipids found in the outer membranes of Gram-negative bacteria that have been demonstrated to act as barriers to permeation of small organic molecules in computational models of *E. coli* membranes, which may contribute to antibiotic resistance (Carpenter et al. 2016). Indeed, several studies have investigated gene products related to lipopolysaccharides as targets for novel antibiotics (Erwin 2016; Srinivas et al. 2010; Werneburg et al. 2012). Lipopolysaccharides also share some important features with molecules suspected of playing a role in bile salt resistance in other organisms (Nikaido 2003), which may imply a similar role in bacteria. Indeed, the *lapAB* operon (previously *yciS*/*yciM*)—which is associated with lipopolysaccharide assembly (Klein et al. 2014; Mahalakshmi et al. 2014)—has been implicated in resistance to bile salts and osmotic stress (Nicolaes et al. 2014) and to antibiotics (Liu et al. 2010; Nicolaes et al. 2014) The operon is constitutivley expressed and is upregulated by RpoH during heat shock (Klein et al. 2014).

Previous work has suggested that *E. coli* could be sensitized to bile salts via transformation with the F plasmid (Bidlack and Silverman 2004). However, unpublished work conducted by the duo between 2004 and 2008 suggests that a subculture of *E. coli* K-12 JC3272/JCFL0 (described by Achtman et al. (1971)) which spontaneously lost the F plasmid was abnormally sensitive to bile salts compared to another F^- subculture of the same strain. As the sensitive subculture dubbed “JC3272I” was a descendant of the one

used in their previous work, the new research casts doubt on their conclusion that a gene in the F plasmid is responsible for bile salt sensitivity in JC3272/JCFL0.

Still unpublished work by Bidlack and Silverman, using the Hfr mapping set as described by Singer et al. (1989), indicated that the heightened sensitivity was possibly due to a mutation in a region between the *trp* operon and a location at about 1,692 kbp, with that range later being narrowed to between the *trp* and *psp* operons (Figure 1). The *lapAB* operon was suggested to be the site most likely to contain the hypothesized mutation.

This thesis project seeks to verify unpublished work suggesting that JC3272I is abnormally sensitive to bile salts by comparing the effect of bile salts on growth in this mutant culture to that of *E. coli* JC3272F, a wildtype of the same strain from a different laboratory. This thesis project also includes comparative DNA sequencing of JC3272I and JC3272F to determine if a mutation in *lapAB* exists that may cause bile salt sensitivity. If a mutation is detected, this research may prove valuable in identifying new mechanisms of bile salt resistance.

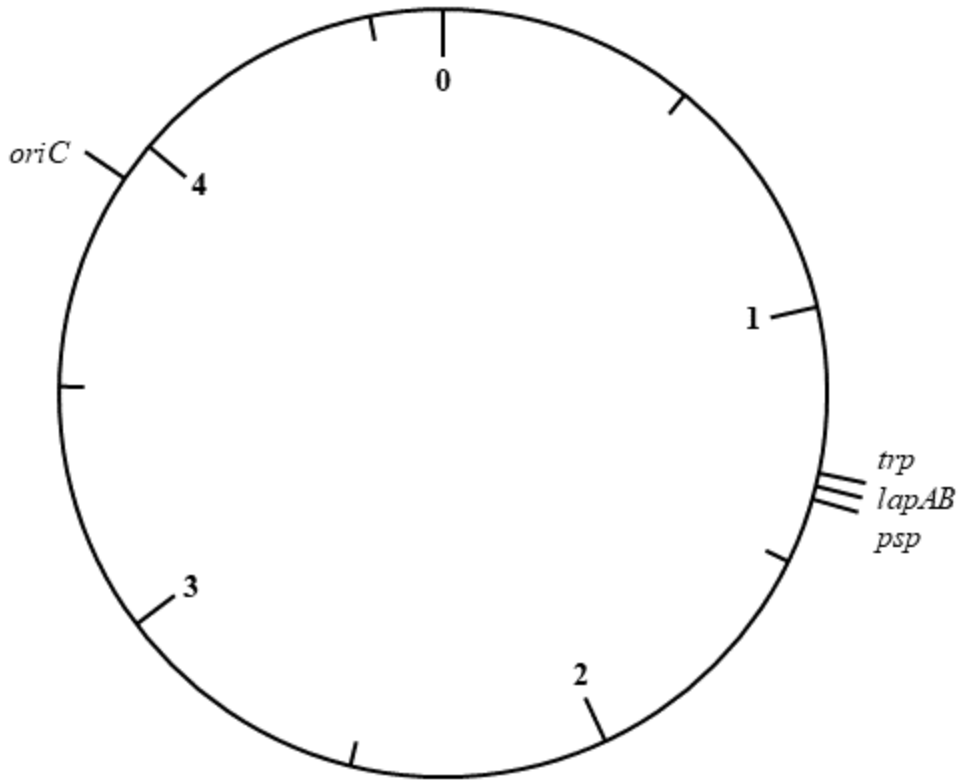

Figure 1. Locations of *trp* and *psp* operons in *E. coli* chromosome. Positions are in Mbp.

MATERIALS AND METHODS

Cultures

Two cultures of *E. coli* K-12 JC3272—which have been described elsewhere (Achtman et al. 1971)—were used: the wildtype from the lab of Laura Frost henceforth referred to as "JC3272F," and the mutant derived from a culture originating in the lab of Karen Ippen-Ihler henceforth referred to as "JC3272I" (Table 1). The JC3272I strain is a subculture of the JC3272/JCFL0 used in Bidlack and Silverman (2004) which spontaneously lost the JCFL0 strain of the F plasmid described in Achtman et. al (Achtman et al. 1971). Other *E. coli* K-12 strains used for comparison include BW25113 (Datsenko and Wanner 2000) and the Keio collection strains JW1271-1 and JW1272-3 (Baba et al. 2006). The BW25113 strain was included as a control, and the latter two were chosen due to their deletions of the *lapA* and *lapB* genes, respectively. All of these strains were obtained from Dr. Phillip Silverman's Laboratory at the Oklahoma Medical Research Foundation in Oklahoma City, Oklahoma. Frozen stocks of each culture were revived by streaking on LB Lennox base (Invitrogen) with 1.5% agarose (w/v, Phenix Research Products), with 30 μg/mL streptomycin when appropriate, and incubating overnight at 37℃.

Table 1. Summary of cultures used.

Strain	Predicted Response to Bile Salts	*lapAB* Genotype
JC3272I	Sensitive	Unknown
JC3272F	Resistant	Unknown
BW25113	Resistant	Wild Type
JW1271-1	Unknown	*ΔlapA787::kan*
JW1272-3	Unknown	*ΔlapB788::kan*

Spot Assays

As a preliminary assessment of bile salt sensitivity, JC3272I, JC3272F, BW25113, JW1271-1, and JW1272-3 were grown overnight in LB media at 37 ℃. Dilutions of each culture were made, and 3 µL of each dilution were dispensed onto LB agar plates containing either 0% or 0.5% Bile Salts no. 3 (DIFCO). Growth was observed after incubating overnight at 37℃.

Spectrophotometric Assays

To further assess the sensitivity of each culture to bile salts, a solution of Bile Salts no. 3 was prepared in water at 25% (w/v) and then sterilized via filtration through a cellulose nitrate membrane with a pore size of 0.45 µm (Nalgene). Individual colonies from the revived JC3272I and JC3272F cultures were then used to inoculate 5.0 mL LB Lennox media, which were then incubated overnight at 37℃. The number of cells/mL was approximated by measuring the OD_{600} of each culture and assuming that each 0.1 OD unit represents approximately 10^8 cells/mL. The cultures were then diluted to 10^4 cells/mL by adding LB Lennox media. Exactly 4 µL of the diluted cultures were added to a 96-well plate along with 100 µL of 2X LB Lennox media and appropriate amounts of bile salt solution and water, to achieve a final volume of 200 µL and the desired concentrations of bile salts. Negative controls were prepared as above but with the cultures being replaced with uninoculated media. The plate was incubated at 37℃ for 18 hours, then optical density was measured at 620 nm. The absorbance of each subculture was analyzed using analysis of variance (ANOVA). Tukey's test was used to determine if growth of cultures varied significantly as a result of different bile salt concentration. Two

outliers were found using the 1.5*IQR rule and removed from the JC3272I data, and five outliers were removed from the JC3272F data.

Polymerase Chain Reaction (PCR)

To screen for mutations in the *lapA* and *lapB* genes, a portion of the JC3272F, JC3272I, BW25113, JW1271-1, and JW1272-3 genomes were amplified via PCR using a *Taq* PCR Kit (New England BioLabs) and primers (Integrated DNA Technologies) with the following sequences:

YCIFI: 5'- GTG GAT GCC ATT TCG CAG ACG GC -3' (Forward)

BSRSB: 5'- AGC GCG ATC AAC TTC GCC AC -3' (Reverse)

BSRSC: 5'- GCG TGA TTA CGT AGC GGG GG -3' (Forward)

BSRSD: 5'- CGG CAC CGG TGT TCT CTT CC -3' (Reverse)

BSRSE: 5'- GGG ACG CGT GTT TAT GGC GA -3' (Forward)

YCIRO: 5'- ACG TCA TGA CCA GAC CTT CTT GAT GAT GG -3' (Reverse)

Amplifications were performed using the primers YCIFI and YCIRO with each of the five above strains. To detect insertions and deletions at a finer scale, and to facilitate later sequencing, further amplifications were performed using the strains JC3272I, JC3272F, and BW25113 and the primer pairs YCIFI/BSRSB, BSRSC/BSRSD, and BSRSE/YCIRO.

Each primer pair was designed to give amplicons of similar length (680-740 bp), all of which were under the 800 bp maximum given by Molecular Cloning Laboratories

(described below). The primers were also designed to give roughly 100 bp of overlap between consecutive amplicons, and the melting temperatures of the primers vary by no more than 2.7 ℃.

Cell counts of overnight broth cultures were estimated using a hemocytometer and diluted to $6.0*10^{-3}$ CFU/mL using TE buffer. 4 µL of this cell suspension were added to each PCR reaction, along with 1 µL of 100 µM forward primer, 1 µL of 100 µM reverse primer, 1 µL of a 1:10 Taq DNA polymerase solution, and 43 µL of master mix. The master mix contained 50 µL 10x Standard *Taq* Reaction Buffer, 10 µL 10 mM dNTPs, 2.5 µL Taq DNA Polymerase, and 407.5 µL Nanopure water. In the thermocycler, each reaction was held at 95 ℃ for five minutes to induce cell lysis before beginning the following cycle:

Denaturing: 30 s, 95 ℃

Annealing: 30 s, 58 ℃

Extension: 48 s, 68 ℃

The annealing temperature was determined using the NEB Tm Calculator (https://tmcalculator.neb.com), and the other temperatures were to the manufacturer's recommendation. The extension time is equal to the 60 s per kbp recommended by the manufacturer. After thirty cycles, a final extension step was performed for five minutes at 68 ℃, then the samples were held at 4 ℃ to terminate the reaction. Electrophoresis was then performed with 4 µL of each sample, mixed with 1 µL of loading dye. The gel was made with 2% agarose in 50 mL TBE buffer with 2 µL ethidium bromide for visualizing bands. The TBE buffer was made to 89 mM Tris base, 89 mM boric acid, with 4 mL 0.5

M EDTA per liter (Moore 1996). A DNA ladder was included on each side of the gel for estimation of size. The gel was then subjected to electrophoresis in TBE buffer at 270 volts until the dye front had traveled half the distance down the gel. Bands were cut from the gel using a gel extractor and stored in a small volume of TE buffer in 1.5 mL microcentrifuge tubes.

Sequencing of PCR Products

Extraction, cleanup, and sequencing of the PCR product were performed by Molecular Cloning Laboratories in South San Francisco, CA (https://mclab.com), with sequencing performed on an Applied Biosystems 3730xl DNA Analyzer. The constructed sequences were compared via NCBI BLAST Needleman-Wunsch Global Align. These sequences were also compared with a homologous region of the K-12 *E. coli* BW25113 (Datsenko and Wanner 2000) genome between 1,334,399 bp and 1,336,184 bp—which contains the *lapA* and *lapB* genes—being used as a reference.

Whole-Genome Sequencing

Genomic sequencing was outsourced to SeqCenter in Pittsburg, PA (https://www.seqcenter.com). Extraction was performed using a ZymoBIOMICS™ DNA Miniprep Kit, and an Illumina DNA Prep kit was used to prepare the Illumina sequencing libraries. Samples of each subculture were sent twice for DNA extraction and whole-genome sequencing. FastQC (Andrews) was used for quality assessment. Cutadapt (Martin 2011) was used to clean the sequences, filtering out sequences with lengths under 20 bp or quality scores under 20. Snippy (Seemann 2015) was used with default parameters to compare the sequences to that of *E. coli* K-12 MG1655 to locate single-

nucleotide polymorphisms (SNPs). Mutations that were present in both sequences of a given strain were compiled, and lists of the genes closest to the mutations were generated using EcoCyc (Karp et al. 2023)

RESULTS & DISCUSSION

Petri Dish Assays

Growth was generally the same for all cultures on regular LB agar, with all strains exhibiting growth at the 10^6 dilution except for BW25113, which only exhibited growth up to 10^4 (Figure 2 A). However, growth was reduced in the presence of bile salts for all cultures except strains JC3272F and BW25113 (Figure 2 B).

It was expected that strains JC3272F and BW25113 would grow normally in both LB and LB + bile salts because there is no documentation in the literature suggesting these strains are sensitive to bile salts. Furthermore, it is likely that most *E. coli* strains would be resistant to bile salts since bile salts are typically found in the gastrointestinal tract. The reduction in growth for JW1271-1 and JW1272-3 in the presence of bile salts was also expected because these genetically modified strains lack the *lapAB* genes responsible, in part, for bile salt resistance (Carpenter et al. 2016; Nicolaes et al. 2014; Nikaido 2003).

Growth of JC3272I was found to be restricted in the presence of bile salts, unlike in JC3272F. The two should be identical cultures of the same strain, JC3272, so this difference may imply the existence of a mutation in JC3272I. If deletions in the *lapAB* operon interfere with bile salt resistance, as this investigation has confirmed, then the restricted growth of JC3272I in the presence of bile salts warrants a deeper look at the *lapAB* operon in the same strain. It is for this reason that DNA sequencing was implemented in this study to compare the *lapAB* operons of JC3272I and JC3272F.

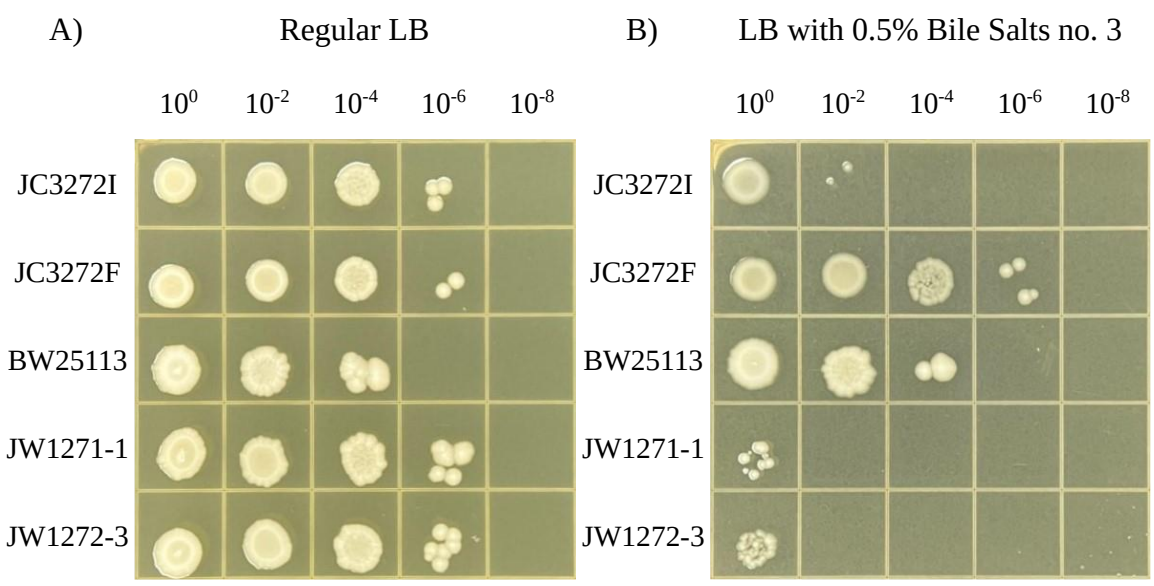

Figure 2. Growth of diluted *E. coli* cultures on LB agar (A) and LB agar with 0.5% Bile Salts no. 3 (B). Dilution factors (shown across tops of photos) increase at hundredfold increments from left (10^{0}) to right (10^{-8}). From top to bottom, cultures are JC3272I (sensitive), JC3272F (resistant), BW25113 (resistant), JW1271-1 (unknown), and JW1272-3 (unknown). No growth is observed for any culture at a dilution factor of 10^{-8}.

Spectrophotometric Assays

Response to bile salts in LB medium was found to vary significantly based on which JC3272 culture was used (Figure 3). Growth of JC3272I was completely suppressed at bile salt concentrations above 0.5%. Two-way ANOVA found that the interaction between strain and bile salt concentration was significant ($F_{1, 6} = 37.692$, $p < 0.001$). One-way ANOVA found that optical density varied significantly for both JC3272I ($F_{6, 26} = 1000.552$, $p < 0.001$) and JC3272F ($F_{6, 23} = 202.716$, $p < 0.001$). Tukey's test found that the growth of each strain was segregated into three distinct levels based on bile salt concentration.

A typical bile acid concentration within the duodenum and upper jejunum is typically 0.1-0.3% (w/v) during digestion, though it can be much higher during the initial release of bile at the beginning of digestion (Borgström et al. 1957), often exceeding 0.6% in the second part duodenum (Simmons and Bouchier 1972). Both cultures demonstrated growth at the typical post-meal bile salt concentration, but JC3272I was completely restricted at concentrations corresponding to that during the initial release of bile salts. It is unclear whether JC3272I could survive in an environment that sees such high concentrations of bile salts at short, but regular intervals.

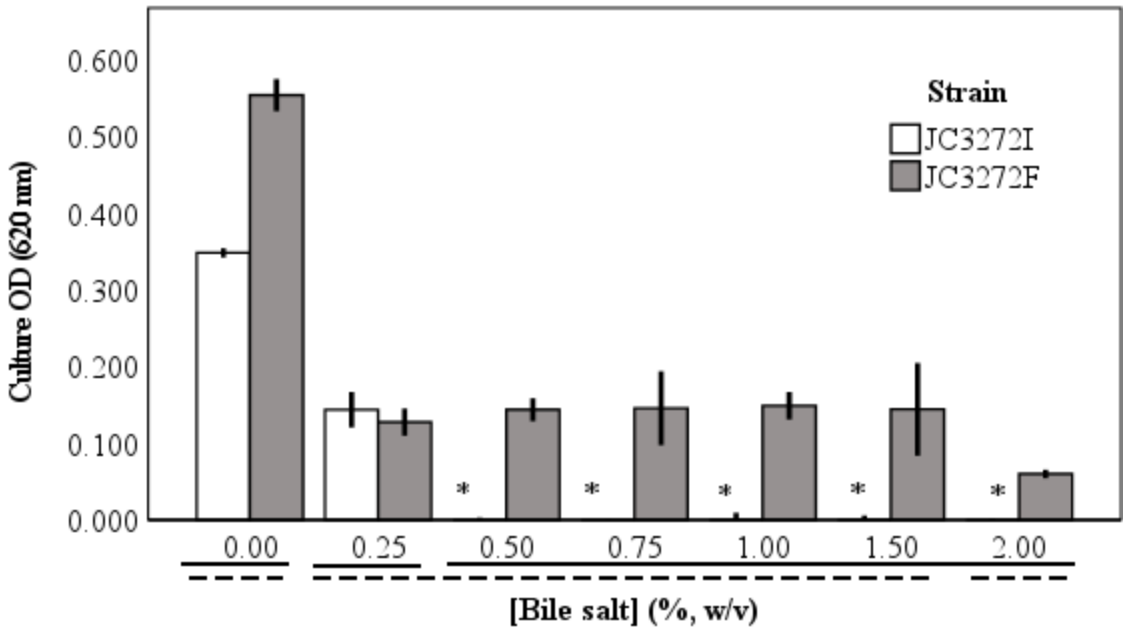

Figure 3. Optical density of E. coli strains JC3272I and JC3272F grown in bile salts. Lines represent 95% confidence limits. An asterisk indicates a negative absorbance value corresponding to no growth. Groups derived from Tukey's test are underlined, with solid lines being used for JC3272I and dashed lines being used for JC3272F.

The heightened sensitivity of JC3272I to bile salts—despite the culture's lack of the F plasmid—presents a challenge to previous work performed by Bidlack and Silverman (2004). JC3272I is a direct descendant of the F' strain used in the study, suggesting that the chromosomal mutation may have been present in that culture as well. This calls into question the assertion that a gene in the F plasmid contributed to the sensitivity to bile salts observed in the previous study (Bidlack and Silverman 2004) and highlights the need for further studies regarding the link between the F plasmid and bile salt tolerance.

The method of cell enumeration used in the growth assay did not account for any differences in cell viability between cultures. It is unknown if cultures of one strain accumulate dead cells more quickly than the other, which would skew any attempt to approximate CFU counts using cell counts. Future research should use plate counts to control for any such differences.

PCR & Gel Electrophoresis

In a preliminary round of PCR amplification using only the YCIFI and YCIRO primers, sufficient amplification was achieved for each strain (Figure 4). The gel showed distinct bands of DNA at approximately 1.8 kbp for strains JC3272I, JC3272F, and BW25113, which is the expected size of the amplicon in these strains. The gel also showed bands at approximately 2.8 kbp for JW1271-1 and 1.9 kbp for JW1272-3. Larger amplicons for JW1271-1 and JW1272-3 were expected since these genetically engineered strains have a larger kanamycin-resistance gene inserted where their *lapA* and *lapB* genes were, respectively. Notably, the consistent amplicon size between JC3272I, JC3272F, and

BW25113 suggests that there are no large insertions or deletions in the *lapAB* operon of JC3272I that might differentiate it from the other two strains.

Additional amplifications were performed using the BSRSB, BSRSC, BSRSD, and BSRSE primers in combination with the original YCIFI and YCIRO primers. These primers were designed to give a more detailed picture of the *lapAB* operon in JC3272I, JC3272F, and BW25113 and to be easier to sequence. With the new set of primer combinations, no difference in amplicon size was found between the three strains (Figure 5). As predicted, each amplicon was between 600 bp and 800 bp, indicating that none of the strains had any large insertions or deletions in any of the amplicons. Specifically, no large insertions or deletions that might impact the function of the *lapAB* operon in JC3272I were detected.

Sequences of lapA *and* lapB

The amplicons from JC3272I and JC3272F were cut from the agarose gel and stored in TE buffer. These samples, along with their primers, were sent to Molecular Cloning Laboratories (MCLAB), where they were sequenced. The overlapping segments of the raw sequences were then aligned visually to obtain the full sequences of *lapA* and *lapB* in both strains.

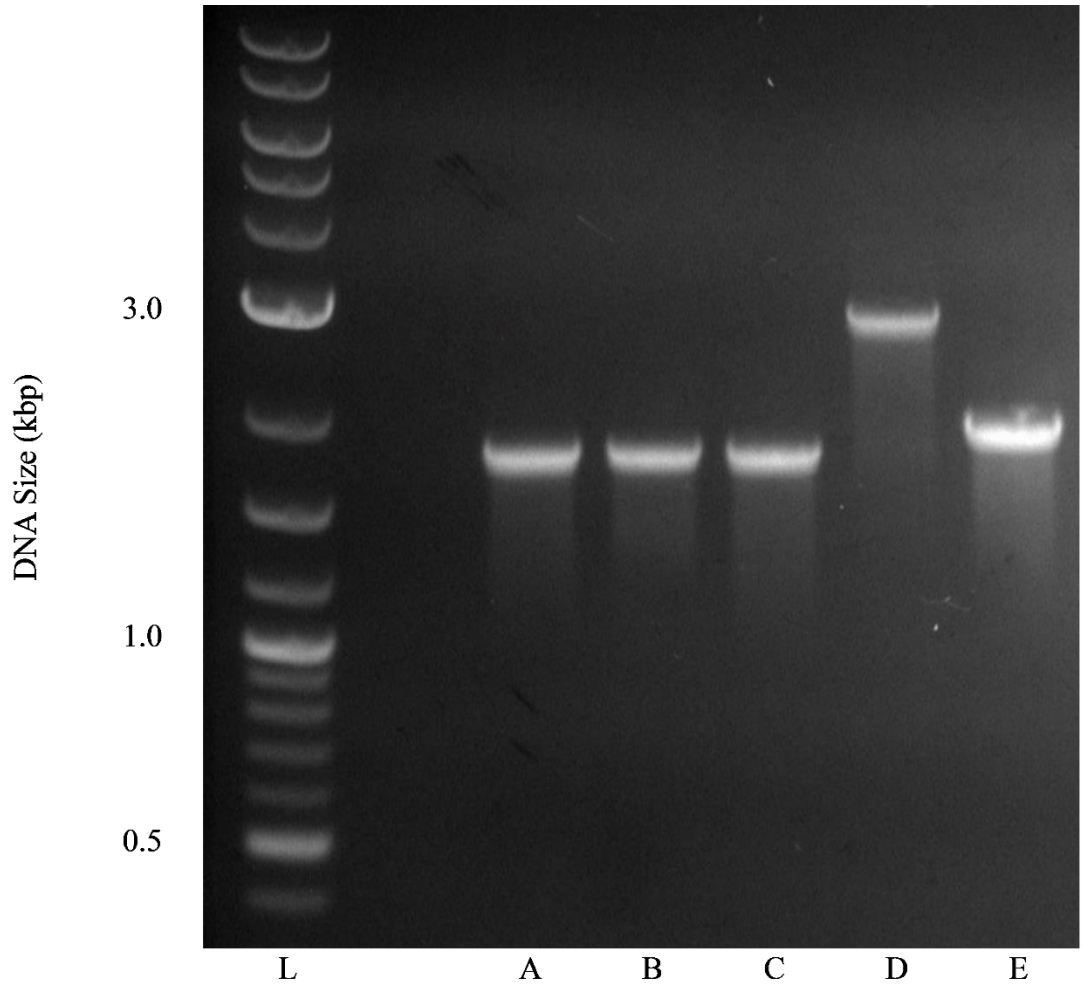

Figure 4. Agarose gel showing relative sizes of the amplicon containing *lapA* and *lapB* from the following strains: JC3272I (A), JC3272F (B), BW25113 (C), JW1271-1 (D), and JW1272-3 (E).

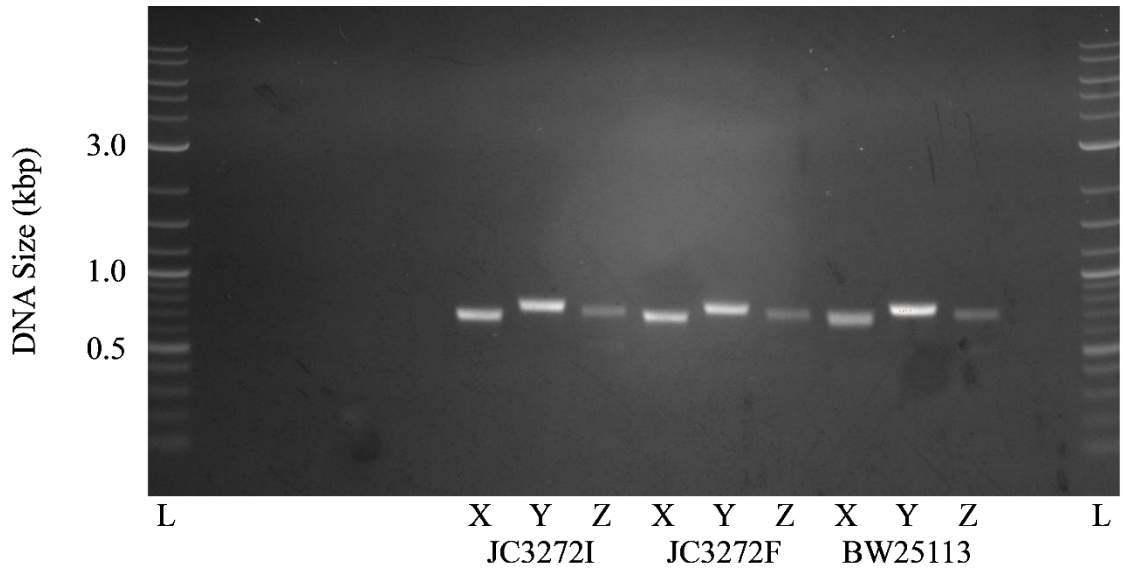

Figure 5. Agarose gel showing relative sizes of amplicons of strains JC3272I, JC3272F, and BW25113. For each strain, amplicons were derived from PCR products using the following primers: X (YCIFI/BSRSB), Y (BSRSC/BSRSD), and Z (BSRSE/YCIRO).

Sequencing of PCR products (Appendices A & B) revealed no difference in DNA sequences for *lapA* or *lapB* in strains JC3272I and JC3272F obtained in this investigation as well as the sequence reported in the literature (Grenier et al. 2014). This indicates that the difference in response to bile salts between JC3272I and JC3272F (as well as BW25113) is not due to a mutation within the *lapA* or *lapB* genes but rather another mechanism. This does not exclude the promoter of the operon, which is located outside of the amplicon (Klein et al. 2014). To determine if the difference in bile salt resistance between JC3272I and JC3272F is due to a mutation in the promoter region or other gene(s), whole genome sequencing was pursued to further elucidate on the difference(s) in DNA sequences in JC3272I and JC3272F.

Whole-Genome Sequencing

Genomic analysis found 103 single nucleotide polymorphisms that differentiated the two cultures, with fifty-one alleles being unique to the JC3272I samples (Table 3) and fifty-two alleles being unique to the JC3272F samples (Table 4). In agreement with the amplification-based results, this analysis found that the *lapA* and *lapB* genes of strains JC3272I and JC3272F were identical to each other and to that of *E. coli* K-12 MG1655.

Table 2. List of mutations specific to JC3272I. These are alleles found in JC3272I but not JC3272F or MG1655.

Regions (bp)	Nucleotides	Nearest gene	Distance	Amino Acids
18,221	C -> A	*nhaA*	intragenic, + strand	R -> S
176,934	G -> A	*erpA*	intragenic, + strand	G -> S
287,313	G -> T	*yagH*	intragenic, + strand	T -> T
296,421	G -> T	*intF*	intragenic, - strand	H -> N
398,193	C -> A	*yaiW*	intragenic, + strand	R -> S
453,900	C -> A	*yajG*	intragenic, - strand	A -> S
603,054	G -> T	*pheP*	intragenic, + strand	A -> S
711,247	G -> T	*fldA*	intragenic, - strand	L -> L
931,037	A -> ATAGTC	*trxB*	48 downstream	
937,867	C -> A	*lolA*	intragenic, + strand	Q -> K
1,172,350	1 -> 13	*mfd*	intragenic, - strand	N -> GAEEN
1,177,623	G -> T	*lolE*	intragenic, + strand	G -> W
1,278,428	T -> G	*narK*	intragenic, + strand	F -> V
1,301,998	A -> G	*oppA*	intragenic, + strand	S -> G
1,329,818	G -> T	*sohB*	intragenic, + strand	D -> Y
1,332,014	G -> T	*topA*	intragenic, + strand	D -> Y
1,345,392	C -> A	*pdeR*	intragenic, - strand	G -> W
1,416,269	G -> T	*recE*	intragenic, - strand	P -> Q
1,494,488	C -> A	*ydcI*	intragenic, - strand	S -> I
1,511,550	C -> A	*ydcS*	104 upstream	
1,567,796	G -> T	*digH*	intragenic, - strand	P -> Q
1,808,427	C -> A	*yniA*	intragenic, + strand	P -> Q
1,900,571	C -> A	*yoaE*	intragenic, - strand	G -> C
1,998,806	G -> T	*dcyD*	intragenic, - strand	A -> A
2,323,872	G -> T	*atoD*	intragenic, + strand	L -> L
2,373,917	G -> T	*menE*	intragenic, - strand	Q -> K
2,396,460	G -> T	*nuoH*	5 downstream	
2,436,338	G -> T	*usg*	intragenic, - strand	P -> P
2,525,343	G -> T	*yfeN*	intragenic, + strand	G -> W
2,601,122	G -> T	*hyfA*	79 upstream	
2,687,730	G -> T	*glrR*	intragenic, - strand	N -> K
2,950,023	G -> T	*argA*	intragenic, + strand	G -> V
3,014,932	G -> T	*yqeC*	intragenic, - strand	L -> L
3,089,831	A -> G	*yggI*	intragenic, + strand	N -> S
3,149,870	C -> A	*yghA*	intragenic, + strand	A -> E
3,232,249	C -> CGTAGA	*fadH*	intragenic, + strand	frame-shift
3,318,922	G -> T	*argG*	intragenic, + strand	G -> C
3,640,506	1 -> 10	*uspA*	intragenic, + strand	V -> VDML
3,657,013	G -> T	*hdeD*	intragenic, + strand	A -> S
3,664,577	G -> T	*gadW*	intragenic, - strand	F -> L
3,683,503	T -> TTAA	*dctA*	56 upstream	
3,707,764	C -> A	*dppA*	59 upstream	
3,894,749	C -> A	*yieF*	intragenic, + strand	A -> E
3,902,759	C -> A	*bglB*	intragenic, - strand	A -> S
4,077,949	T -> TCTGC	*fabY*	intragenic, + strand	frame-shift
4,139,265	C -> A	*fsaB*	intragenic, - strand	E -> D
4,187,762	C -> A	*rpoC*	intragenic, + strand	Q -> K
4,216,749	G -> T	*aceB*	intragenic, + strand	V -> V
4,270,711	G -> T	*yjbR*	intragenic, + strand	Q -> H
4,319,120	G -> T	*phnJ*	intragenic, - strand	R -> S
4,564,563	C -> A	*yjiL*	intragenic, - strand	A -> S

Table 3. List of mutations specific to JC3272F. These are alleles found in JC3272F but not JC3272I or MG1655.

Regions (bp)	Nucleotides	Nearest gene	Distance	Amino Acids
17,274	C -> A	*nhaA*	215 upstream	
134,028	T -> A	*acnB*	intragenic, + strand	L -> Q
161,921	G -> T	*thpR*	intragenic, - strand	V -> V
244,293	G -> T	*yafJ*	34 upstream	
478,829	C -> A	*ylaC*	intragenic, - strand	L -> L
1,174,129	G -> T	*ycfT*	intragenic, - strand	P -> Q
1,264,732	G -> A	*hemA*	intragenic, + strand	S -> N
1,265,330	G -> A	*prfA*	intragenic, + strand	D -> N
1,278,282	C -> T	*narK*	intragenic, + strand	T -> I
1,281,407	C -> T	*narG*	intragenic, + strand	S -> F
1,282,565	C -> T	*narG*	intragenic, + strand	T -> I
1,301,992	A -> T	*oppA*	intragenic, + strand	N -> Y
1,306,736	T -> G	*oppF*	intragenic, + strand	S -> A
1,309,514	C -> T	*kch*	intragenic, - strand	S -> S
1,310,980	C -> T	*tonB*	109 upstream	
1,311,066	C -> T	*tonB*	23 upstream	
1,313,066	C -> T	*yciC*	intragenic, - strand	A -> T
1,314,860	C -> T	*yciE*	intragenic, - strand	G -> D
1,317,092	C -> T	*trpA*	intragenic, - strand	G -> D
1,322,022	C -> T	*trpE*	intragenic, - strand	E -> K
1,326,092	C -> T	*yciQ*	intragenic, + strand	A -> A
1,327,137	C -> T	*rluB*	intragenic, + strand	P -> S
1,351,858	C -> T	*sapF*	intragenic, - strand	L -> L
1,354,145	C -> T	*sapC*	intragenic, - strand	G -> D
1,354,984	C -> T	*sapB*	intragenic, - strand	D -> N
1,355,541	C -> T	*sapA*	intragenic, - strand	V -> I
1,367,613	G -> A	*pspF*	intragenic, - strand	D -> D
1,369,027	G -> A	*pspC*	intragenic, + strand	M -> I
1,370,067	C -> T	*pspE*	64 downstream	
1,374,240	C -> T	*ycjP*	intragenic, + strand	P -> S
1,374,838	C -> T	*ycjP*	intragenic, + strand	A -> V
1,377,813	C -> T	*ycjS*	intragenic, + strand	Q -> stop
1,472,372	G -> T			
1,646,966	G -> T	*ydfW*	208 downstream	
1,843,749	C -> A	*gdhA*	35 downstream	
1,907,379	G -> T	*cspC*	intragenic, - strand	F -> L
1,994,841	C -> A	*uvrY*	intragenic, - strand	Q -> H
2,136,096	C -> A	*wzb*	intragenic, - strand	M -> I
2,210,193	G -> T	*yehQ*	intragenic, + strand	G -> V
2,661,974	C -> A	*iscR*	intragenic, - strand	G -> V
2,706,611	C -> A	*lepA*	intragenic, - strand	V -> F
2,911,788	C -> A	*relA*	intragenic, - strand	V -> F
3,147,112	C -> A	*yghX*	intragenic, - strand	V -> F
3,158,649	G -> T	*yqhH*	intragenic, + strand	G -> V
3,311,805	G -> T	*truB*	28 downstream	
3,333,298	G -> T	*rplU*	intragenic, - strand	P -> T
3,822,645	C -> CCAGGAT	*spoT*	intragenic, + strand	Y -> YQD
3,887,992	C -> A	*mnmE*	intragenic, + strand	R -> S
3,938,365	C -> A	*rbsR*	intragenic, + strand	P -> T
4,136,004	G -> T	*katG*	intragenic, + strand	D -> Y
4,198,537	C -> A	*hemE*	intragenic, + strand	D -> E
4,286,810	C -> A	*acs*	intragenic, - strand	E -> stop

No mutation was found in either the *lapA* or *lapB* genes of JC3272I nor were there any mutations in the promoters for *lapA* and *lapB* reported by Klein et al. (2014). Further, no mutations were found in *rpoH*, which regulates one of *lapAB*'s promoters. Three mutations were found in genes elsewhere in the region implicated by Bidlack and Silverman in their unpublished work using Hfr mapping, those being *sohB* (Baird et al. 1991), *topA* (Trucksis and Depew 1981), and *pdeR* (Hengge et al. 2016). All three were non-synonymous substitutions affecting a single amino acid, though none of the genes are known to contribute to bile salt resistance. However, several mutations were found in the corresponding region of JC3272F in the genes *trpE*, *yciQ*, *rluB* and *sapF*.

Tables 3 and 4 list the mutations that differentiate JC3272I and JC3272F, respectively, from each other and from MG1655. Both JC3272 derivatives used in this study share mutations that differentiate them from MG1655 but not from each other; those mutations are excluded from the tables. Comparison of the JC3272 derivatives to a third strain was necessary since they could not be compared directly without assembling their genomes.

Looking outside of the original study region, none of the mutations for JC3272I (Table 4) were found to have well-published effects on bile salt resistance. This may point to one of the unidentified efflux pumps speculated upon by Thanassi et al. (1997). However, one gene found to contain a mutation—*nhaA* (Figure 6)—is known to play a role as a sodium antiporter (Taglicht et al. 1991). Since this thesis did not control for the increase in sodium concentration from the bile salts in the medium, future studies should investigate the role of sodium content and osmotic stress on the growth of these cultures.

Areas of study which may provide some insight into the origin of bile salt sensitivity in JC3272I include the role of *sohB*, *topA*, and *pdeR* in bile salt resistance. Higher-coverage sequencing is also a worthwhile effort, as it would allow for the assembly of the JC3272I and JC3272F genomes and possibly facilitate the description of these strains (Riesco and Trujillo 2024). RNA sequencing could also be used to determine if the difference is sensitivity is due to a difference in gene expression, and this could be followed by gene network analysis.

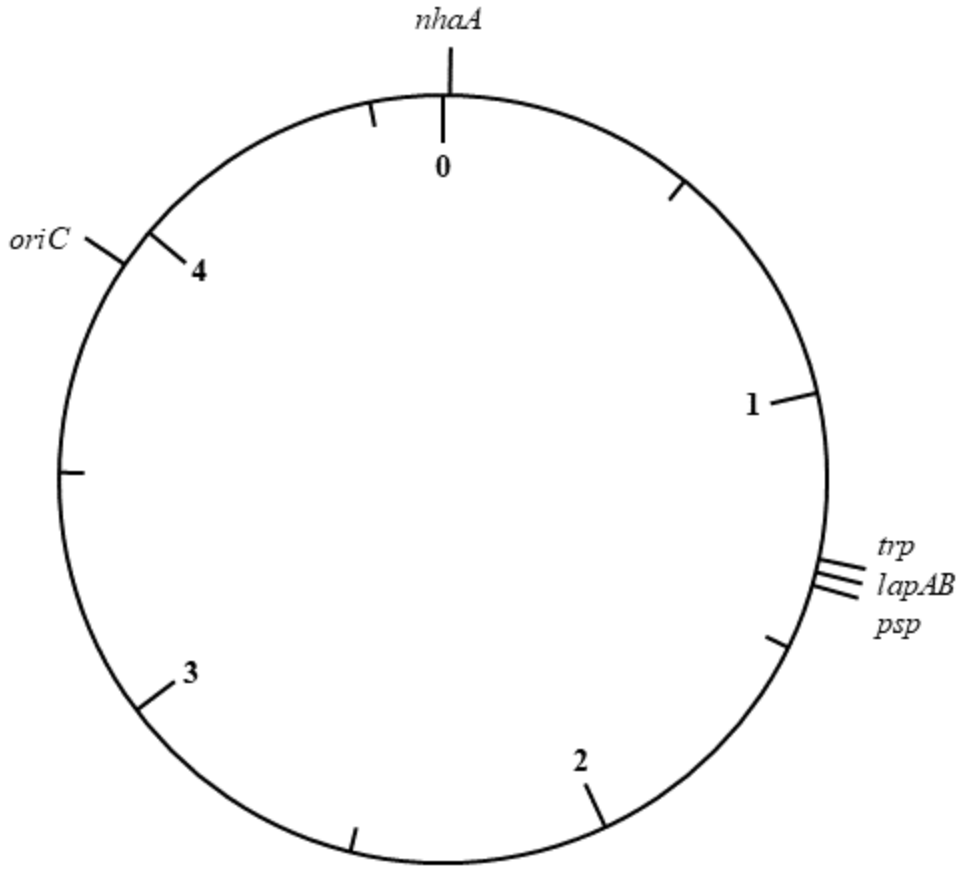

Figure 6. Locations of *nhaA* and the *trp*, and *psp* operons in the *E. coli* chromosome. Positions are in Mbp.

LITERATURE CITED

Achtman M, Willetts N, Clark AJ. 1971. Beginning a genetic analysis of conjugational transfer determined by the F factor in *Escherichia coli* by isolation and characterization of transfer-deficient mutants. Journal of Bacteriology. 106(2):529-538.

Fastqc a quality control tool for high throughput sequence data. [accessed]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* k-12 in-frame, single-gene knockout mutants: The Keio collection. Molecular Systems Biology. 2(1):2006.0008.

Baird L, Lipinska B, Raina S, Georgopoulos C. 1991. Identification of the *Escherichia coli sohB* gene, a multicopy suppressor of the HtrA (DegP) null phenotype. Journal of Bacteriology. 173(18):5763-5770.

Bidlack JE, Silverman PM. 2004. An active type IV secretion system encoded by the F plasmid sensitizes *Escherichia coli* to bile salts. Journal of Bacteriology. 186(16):5202-5209.

Borgström B, Dahlqvist A, Lundh G, Sjövall J. 1957. Studies of intestinal digestion and absorption in the human. The Journal of Clinical Investigation. 36(10):1521-1536.

Carpenter TS, Parkin J, Khalid S. 2016. The free energy of small solute permeation through the *Escherichia coli* outer membrane has a distinctly asymmetric profile. The Journal of Physical Chemistry Letters. 7(17):3446-3451.

Carulli N, Bertolotti M, Carubbi F, Concari M, Martella P, Carulli L, Loria P. 2000. Review article: Effect of bile salt pool composition on hepatic and biliary functions. Alimentary Pharmacology & Therapeutics. 14(s2):14-18.

Clavel T, Lazzaroni JC, Vianney A, Portalier R. 1996. Expression of the *tolQRA* genes of *Escherichia coli* K-12 is controlled by the RcsC sensor protein involved in capsule synthesis. Molecular Microbiology. 19(1):19-25.

Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences. 97(12):6640-6645.

Dawson PA. 2012. Bile formation and the enterohepatic circulation. In: Said HM, editor. Physiology of the gastrointestinal tract. Fifth ed. San Diego, UNITED STATES: Elsevier Science & Technology. p. 1461-1484.

Erwin AL. 2016. Antibacterial drug discovery targeting the lipopolysaccharide biosynthetic enzyme LpxC. Cold Spring Harb Perspect Med. 6(7).

Fralick JA. 1996. Evidence that TolC is required for functioning of the Mar/AcrAB efflux pump of *Escherichia coli*. Journal of Bacteriology. 178(19):5803-5805.

Gipson KS, Nickerson KP, Drenkard E, Llanos-Chea A, Dogiparthi SK, Lanter BB, Hibbler RM, Yonker LM, Hurley BP, Faherty CS. 2020. The great ESKAPE: Exploring the crossroads of bile and antibiotic resistance in bacterial pathogens. Infection and Immunity. 88(10):10.1128/iai.00865-00819.

Grenier F, Matteau D, Baby V, Rodrigue S. 2014. Complete genome sequence of *Escherichia coli* BW25113. Genome Announcements. 2(5):10.1128/genomea.01038-01014.

Gunn JS. 2000. Mechanisms of bacterial resistance and response to bile. Microbes and Infection. 2(8):907-913.

Hengge R, Galperin MY, Ghigo J-M, Gomelsky M, Green J, Hughes KT, Jenal U, Landini P. 2016. Systematic nomenclature for GGDEF and EAL domain-containing cyclic di-GMP turnover proteins of *Escherichia coli*. Journal of Bacteriology. 198(1):7-11.

Inagaki T, Moschetta A, Lee Y-K, Peng L, Zhao G, Downes M, Yu RT, Shelton JM, Richardson JA, Repa JJ et al. 2006. Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. Proceedings of the National Academy of Sciences. 103(10):3920-3925.

Karp PD, Paley S, Caspi R, Kothari A, Krummenacker M, Midford PE, Moore LR, Subhraveti P, Gama-Castro S, Tierrafria VH et al. 2023. The Ecocyc database (2023). EcoSal Plus. 11(1):eesp-0002-2023.

Klein G, Kobylak N, Lindner B, Stupak A, Raina S. 2014. Assembly of lipopolysaccharide in *Escherichia coli* requires the essential LapB heat shock protein. Journal of Biological Chemistry. 289(21):14829-14853.

Le VVH, Biggs PJ, Wheeler D, Davies IG, Rakonjac J. 2020. Novel mechanisms of tolc-independent decreased bile-salt susceptibility in *Escherichia coli*. FEMS Microbiology Letters. 367(10):fnaa083.

Liu A, Tran L, Becket E, Lee K, Chinn L, Park E, Tran K, Miller JH. 2010. Antibiotic sensitivity profiles determined with an *Escherichia coli* gene knockout collection: Generating an antibiotic bar code. Antimicrobial Agents and Chemotherapy. 54(4):1393-1403.

Lorenzo-Zúñiga V, Bartolí R, Planas R, Hofmann AF, Viñado B, Hagey LR, Hernández JM, Mañé J, Alvarez MA, Ausina V et al. 2003. Oral bile acids reduce bacterial overgrowth, bacterial translocation, and endotoxemia in cirrhotic rats. Hepatology. 37(3):551-557.

Mahalakshmi S, Sunayana MR, SaiSree L, Reddy M. 2014. *YciM* is an essential gene required for regulation of lipopolysaccharide synthesis in *Escherichia coli*. Molecular Microbiology. 91(1):145-157.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011. 17(1):3.

Moore DD. 1996. Commonly used reagents and equipment. Current Protocols in Molecular Biology. 35(1):A.2.1-A.2.8.

Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Robles Aguilar G, Gray A, Han C, Bisignano C, Rao P, Wool E et al. 2022. Global burden of bacterial antimicrobial resistance in 2019: A systematic analysis. The Lancet. 399(10325):629-655.

Nicolaes V, Hajjaji HE, Davis RM, Henst CVd, Depuydt M, Leverrier P, Aertsen A, Haufroid V, Choudens SOd, Bolle XD et al. 2014. Insights into the function of YciM, a heat shock membrane protein required to maintain envelope integrity in *Escherichia coli*. Journal of Bacteriology. 196(2):300-309.

Nikaido H. 2003. Molecular basis of bacterial outer membrane permeability revisited. Microbiol Mol Biol Rev. 67(4):593-656.

Okusu H, Ma D, Nikaido H. 1996. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (MAR) mutants. Journal of Bacteriology. 178(1):306-308.

Riesco R, Trujillo ME. 2024. Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. International Journal of Systematic and Evolutionary Microbiology. 74(3).

Riley M, Abe T, Arnaud MB, Berlyn MKB, Blattner FR, Chaudhuri RR, Glasner JD, Horiuchi T, Keseler IM, Kosuge T et al. 2006. *Escherichia coli* K-12: A cooperatively developed annotation snapshot—2005. Nucleic Acids Research. 34(1):1-9.

Seemann T. 2015. Snippy: Fast bacterial variant calling from NGS reads.

Simmons F, Bouchier IAD. 1972. Intraluminal bile salt concentrations and fat digestion after cholecystectomy. South African Medical Journal. 46:2089-2092.

Singer M, Baker TA, Schnitzler G, Deischel SM, Goel M, Dove W, Jaacks KJ, Grossman AD, Erickson JW, Gross CA. 1989. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. Microbiological Reviews. 53(1):1-24.

Srinivas N, Jetter P, Ueberbacher BJ, Werneburg M, Zerbe K, Steinmann J, Van der Meijden B, Bernardini F, Lederer A, Dias RLA et al. 2010. Peptidomimetic antibiotics target outer-membrane biogenesis in *Pseudomonas aeruginosa*. Science. 327(5968):1010-1013.

Taglicht D, Padan E, Schuldiner S. 1991. Overproduction and purification of a functional Na+/H+ antiporter coded by *nhaA* (*ant*) from *Escherichia coli*. Journal of Biological Chemistry. 266(17):11289-11294.

Thanassi DG, Cheng LW, Nikaido H. 1997. Active efflux of bile salts by *Escherichia coli*. Journal of Bacteriology. 179(8):2512-2518.

Trucksis M, Depew RE. 1981. Identification and localization of a gene that specifies production of *Escherichia coli* DNA topoisomerase I. Proceedings of the National Academy of Sciences. 78(4):2164-2168.

Werneburg M, Zerbe K, Juhas M, Bigler L, Stalder U, Kaech A, Ziegler U, Obrecht D, Eberl L, Robinson JA. 2012. Inhibition of lipopolysaccharide transport to the outer membrane in *Pseudomonas aeruginosa* by peptidomimetic antibiotics. ChemBioChem. 13(12):1767-1775.

Yoshimoto T, Higashi H, Kanatani A, Lin XS, Nagai H, Oyama H, Kurazono K, Tsuru D. 1991. Cloning and sequencing of the 7 alpha-hydroxysteroid dehydrogenase gene from *Escherichia coli* HB101 and characterization of the expressed enzyme. Journal of Bacteriology. 173(7):2173-2179.

Yoshimoto T, Nagai H, Ito K, Tsuru D. 1993. Location of the 7 alpha-hydroxysteroid dehydrogenase gene (*hdhA*) on the physical map of the *Escherichia coli* chromosome. Journal of Bacteriology. 175(17):5730.

APPENDICES

Appendix A: DNA Sequences for *lapA* from the literature for strain BW25113 and from Zane Henderson at the University of Central Oklahoma for strain JC3272I.

Escherichia coli lapA (also known as *yciS* (Klein et al. 2014), Genbank locus tag b1280 (Riley et al. 2006)) lipopolysaccharide assembly protein A gene sequence from *Escherichia coli* str. K-12 substr. BW25113 according to Grenier et al. (2014). This is for base pairs 1334500-1334808 on the chromosome:

```
1       GTGAAATATT TACTCATTTT CTTACTGGTG TTAGCGATCT TCGTGATTTC
51      GGTCACGTTG GGTGCGCAGA ACGATCAACA GGTGACGTTT AATTATCTGT
101     TAGCGCAAGG GGAGTACCGT ATTTCCACAT TGCTGGCGGT ATTGTTTGCT
151     GCGGGGTTTG CTATCGGTTG GTTGATTTGT GGCCTGTTCT GGCTGCGAGT
201     TCGTGTTTCC CTGGCGCGCG CTGAACGTAA AATAAAGCGA CTGGAAAACC
251     AGCTTTCACC CGCGACTGAC GTGGCTGTAG TGCCGCACTC GTCAGCGGCG
301     AAGGAATAA
```

Escherichia coli lapA sequence obtained by Zane Henderson for *Escherichia coli* strain JC3272I using the amplicon YCIFI/BSRSB on August 1, 2022:

```
1       GTGAAATATT TACTCATTTT CTTACTGGTG TTAGCGATCT TCGTGATTTC
51      GGTCACGTTG GGTGCGCAGA ACGATCAACA GGTGACGTTT AATTATCTGT
101     TAGCGCAAGG GGAGTACCGT ATTTCCACAT TGCTGGCGGT ATTGTTTGCT
151     GCGGGGTTTG CTATCGGTTG GTTGATTTGT GGCCTGTTCT GGCTGCGAGT
201     TCGTGTTTCC CTGGCGCGCG CTGAACGTAA AATAAAGCGA CTGGAAAACC
251     AGCTTTCACC CGCGACTGAC GTGGCTGTAG TGCCGCACTC GTCAGCGGCG
301     AAGGAATAA
```

Appendix B: DNA sequences for *lapB* from the literature for strain BW25113 and from Zane Henderson at the University of Central Oklahoma in strains JC3272I and JC3272F.

Escherichia coli lapB (also known as *yciM* (Klein et al. 2014), Genbank locus tag b1280 (Riley et al. 2006)) *lapB* – lipopolysaccharide assembly protein B gene sequence from *Escherichia coli* str. K-12 substr. BW25113 according to Grenier et al. (2014). This is for base pairs 1334815-1335984 on the chromosome:

```
1       ATGCTGGAGT TGTTGTTTCT GCTTTTGCCT GTAGCCGCTG CCTATGGCTG
51      GTATATGGGC CGCAGAAGTG CGCAACAAAA CAAGCAAGAT GAAGCCAACC
101     GCTTGTCGCG TGATTACGTA GCGGGGGTTA ACTTCCTGCT TAGTAATCAA
151     CAGGATAAAG CGGTAGACCT GTTTCTCGAT ATGCTTAAAG AGGATACAGG
201     CACCGTTGAA GCCCACCTTA CGCTCGGAAA CCTGTTCCGT TCGCGTGGCG
251     AAGTTGATCG CGCTATTCGC ATCCATCAGA CCCTAATGGA AAGCGCCTCG
301     CTGACCTATG AACAGCGTCT GTTGGCGATT CAACAACTGG GGCGTGATTA
351     CATGGCCGCC GGGTTATATG ACCGCGCGGA AGACATGTTC AATCAGCTGA
401     CCGATGAAAC TGACTTCCGC ATTGGCGCGC TGCAACAGTT GCTACAAATC
451     TACCAGGCTA CCAGTGAGTG GCAGAAAGCA ATTGATGTTG CCGAACGCCT
501     GGTGAAGCTG GGTAAAGATA AACAGCGCGT CGAAATTGCC CATTTCTACT
541     GTGAGTTAGC CCTGCAGCAT ATGGCCAGCG ACGATCTCGA TCGTGCCATG
601     ACCTTGCTAA AAAAAGGGGC GGCGGCAGAT AAAAACAGCG CCCGCGTATC
651     CATAATGATG GGACGCGTGT TTATGGCGAA AGGAGAATAC GCCAAAGCCG
701     TCGAAAGTCT GCAACGCGTG ATATCCCAGG ACAGAGAACT GGTCAGCGAA
751     ACGCTGGAAA TGTTGCAAAC CTGCTATCAG CAGTTGGGTA AAACTGCCGA
801     ATGGGCAGAA TTCCTGCAGC GTGCGGTGGA AGAGAACACC GGTGCCGATG
851     CTGAATTGAT GCTTGCTGAT ATCATCGAAG CGCGCGACGG TAGTGAGGCC
901     GCACAGGTCT ATATTACGCG CCAGCTTCAG CGTCATCCGA CCATGCGTGT
951     GTTCCATAAG TTAATGGATT ACCACTTAAA TGAAGCGGAA GAAGGGCGTG
1001    CCAAAGAGAG CCTGATGGTG CTGCGTGACA TGGTTGGCGA GAAGGTACGT
1051    AGTAAGCCTC GTTATCGTTG CCAGAAATGT GGTTTTACCG CATACACCCT
1101    CTACTGGCAT TGTCCGTCTT GTCGGGCCTG GTCAACCATT AAACCGATTC
1151    GCGGTCTTGA TGGCCTGTAA
```

Escherichia coli lapB sequence obtained by Zane Henderson for *Escherichia coli* strain JC3272I using amplicons YCIFI/BSRSB, BSRSC/BSRSD, and BSRSE/YCIRO on August 1, 2022:

```
1       ATGCTGGAGT TGTTGTTTCT GCTTTTGCCT GTAGCCGCTG CCTATGGCTG
51      GTATATGGGC CGCAGAAGTG CGCAACAAAA CAAGCAAGAT GAAGCCAACC
101     GCTTGTCGCG TGATTACGTA GCGGGGGTTA ACTTCCTGCT TAGTAATCAA
151     CAGGATAAAG CGGTAGACCT GTTTCTCGAT ATGCTTAAAG AGGATACAGG
201     CACCGTTGAA GCCCACCTTA CGCTCGGAAA CCTGTTCCGT TCGCGTGGCG
251     AAGTTGATCG CGCTATTCGC ATCCATCAGA CCCTAATGGA AAGCGCCTCG
301     CTGACCTATG AACAGCGTCT GTTGGCGATT CAACAACTGG GGCGTGATTA
351     CATGGCCGCC GGGTTATATG ACCGCGCGGA AGACATGTTC AATCAGCTGA
401     CCGATGAAAC TGACTTCCGC ATTGGCGCGC TGCAACAGTT GCTACAAATC
451     TACCAGGCTA CCAGTGAGTG GCAGAAAGCA ATTGATGTTG CCGAACGCCT
501     GGTGAAGCTG GGTAAAGATA AACAGCGCGT CGAAATTGCC CATTTCTACT
551     GTGAGTTAGC CCTGCAGCAT ATGGCCAGCG ACGATCTCGA TCGTGCCATG
601     ACCTTGCTAA AAAAAGGGGC GGCGGCAGAT AAAAACAGCG CCCGCGTATC
651     CATAATGATG GGACGCGTGT TTATGGCGAA AGGAGAATAC GCCAAAGCCG
701     TCGAAAGTCT GCAACGCGTG ATATCCCAGG ACAGAGAACT GGTCAGCGAA
751     ACGCTGGAAA TGTTGCAAAC CTGCTATCAG CAGTTGGGTA AAACTGCCGA
801     ATGGGCAGAA TTCCTGCAGC GTGCGGTGGA AGAGAACACC GGTGCCGATG
851     CTGAATTGAT GCTTGCTGAT ATCATCGAAG CGCGCGACGG TAGTGAGGCC
901     GCACAGGTCT ATATTACGCG CCAGCTTCAG CGTCATCCGA CCATGCGTGT
951     GTTCCATAAG TTAATGGATT ACCACTTAAA TGAAGCGGAA GAAGGGCGTG
1001    CCAAAGAGAG CCTGATGGTG CTGCGTGACA TGGTTGGCGA GAAGGTACGT
1051    AGTAAGCCTC GTTATCGTTG CCAGAAATGT GGTTTTACCG CATACACCCT
1101    CTACTGGCAT TGTCCGTCTT GTCGGGCCTG GTCAACCATT AAACCGATTC
1151    GCGGTCTTGA TGGCCTGTAA
```

Escherichia coli lapB sequence obtained by Zane Henderson for *Escherichia coli* strain JC3272F using amplicons BSRSC/BSRSD and BSRSE/YCIRO on August 1, 2022:

```
1       NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
51      NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
101     NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
151     NNNNNNNAAG CGGTAGACCT GTTTCTCGAT ATGCTTAAAG AGGATACAGG
201     CACCGTTGAA GCCCACCTTA CGCTCGGAAA CCTGTTCCGT TCGCGTGGCG
251     AAGTTGATCG CGCTATTCGC ATCCATCAGA CCCTAATGGA AAGCGCCTCG
301     CTGACCTATG AACAGCGTCT GTTGGCGATT CAACAACTGG GGCGTGATTA
351     CATGGCCGCC GGGTTATATG ACCGCGCGGA AGACATGTTC AATCAGCTGA
401     CCGATGAAAC TGACTTCCGC ATTGGCGCGC TGCAACAGTT GCTACAAATC
451     TACCAGGCTA CCAGTGAGTG GCAGAAAGCA ATTGATGTTG CCGAACGCCT
201     GGTGAAGCTG GGTAAAGATA AACAGCGCGT CGAAATTGCC CATTTCTACT
551     GTGAGTTAGC CCTGCAGCAT ATGGCCAGCG ACGATCTCGA TCGTGCCATG
601     ACCTTGCTAA AAAAAGGGGC GGCGGCAGAT AAAAACAGCG CCCGCGTATC
651     CATAATGATG GGACGCGTGT TTATGGCGAA AGGAGAATAC GCCAAAGCCG
701     TCGAAAGTCT GCAACGCGTG ATATCCCAGG ACAGAGAACT GGTCAGCGAA
751     ACGCTGGAAA TGTTGCAAAC CTGCTATCAG CAGTTGGGTA AAACTGCCGA
801     ATGGGCAGAA TTCCTGCAGC GTGCGGTGGA AGAGAACACC GGTGCCGATG
851     CTGAATTGAT GCTTGCTGAT ATCATCGAAG CGCGCGACGG TAGTGAGGCC
901     GCACAGGTCT ATATTACGCG CCAGCTTCAG CGTCATCCGA CCATGCGTGT
951     GTTCCATAAG TTAATGGATT ACCACTTAAA TGAAGCGGAA GAAGGGCGTG
1001    CCAAAGAGAG CCTGATGGTG CTGCGTGACA TGGTTGGCGA GAAGGTACGT
1051    AGTAAGCCTC GTTATCGTTG CCAGAAATGT GGTTTTACCG CATACACCCT
1101    CTACTGGCAT TGTCCGTCTT GTCGGGCCTG GTCAACCATT AAACCGATTC
1151    GCGGTCTTGA TGGCCTGTAA
```