POTENTIAL FOR PHENOTYPIC PLASTICITY IN THE HYPEROSMOTIC STRESS RESPONSE OF DIVERSE *E. COLI*

A Thesis
Submitted to the Graduate Faculty
of the
North Dakota State University
of Agriculture and Applied Science

By

Madelyn Ashley Schwartz

In Partial Fulfillment of the Requirements
for the Degree of
MASTER OF SCIENCE

Major Department:
Microbiological Sciences

May 2021

Fargo, North Dakota

North Dakota State University
Graduate School

Title

POTENTIAL FOR PHENOTYPIC PLASTICITY IN THE HYPEROSMOTIC STRESS RESPONSE OF DIVERSE *E. COLI*

By

Madelyn Ashley Schwartz

The Supervisory Committee certifies that this ***disquisition*** complies with North Dakota State University's regulations and meets the accepted standards for the degree of

MASTER OF SCIENCE

SUPERVISORY COMMITTEE:

Dr. Peter Bergholz
Chair

Dr. Teresa Bergholz

Dr. John Wilkinson

Approved:

June 30, 2021
Date

Dr. John McEvoy
Department Chair

ABSTRACT

Escherichia coli inhabits multiple environments that contain varying physical stresses which emphasize the importance of adaptation. The large pangenome of *E. coli* can account for some strain-to-strain variances in phenotype, but phenotypic plasticity may be another key factor. The key behind phenotypic plasticity may be transcriptional regulators such as the sigma factor RpoS. The extent of phenotypic plasticity and possible mechanisms in stress conditions has not been heavily studied in bacteria. We looked at the presence of plasticity in the growth rate of *E. coli* under a hyperosmotic condition. RNAseq was used to explore the connection between RpoS and phenotypic plasticity. The conserved genes within all the isolates and their connection to hyperosmotic stress were also explored. Phenotypic plasticity of growth rate was observed among the strains on a phenotypic level. Genes associated with osmotic stress, fermentation, and cell envelope synthesis were found to be significantly expressed within all isolates.

ACKNOWLEDGMENTS

I would like to acknowledge and express my deepest thanks to my advisor Dr. Peter Bergholz for providing me the opportunity to be a member of his lab. His guidance and support provided me with the knowledge and skills to explore the realm of microbiology with the greatest detail. I am also grateful to my committee members Dr. Teresa Bergholz, Dr. John Wilkinson, and Dr. Birgit Prüß, for their advice and aid throughout my master's studies.

I would also like to acknowledge Kaycie Schmidt, Alex Arroyo, and Allan Gramillo as their technical and motivational assistance profoundly aided my research, whether directly or indirectly. I extend my gratitude to the Microbiological Sciences Department at North Dakota State University as well. My thanks also go to my family and friends who provided me with the utmost encouragement.

Lastly, I would like to acknowledge the funding source that provided me the resources for performing my research and giving me the opportunity to study at North Dakota State University: National Science Foundation Career award no. 1453397 to P.W.B.

DEDICATION

I dedicate this thesis to my mother Kay Schwartz, who is a source of profound inspiration and encouragement not only in my graduate education, but throughout my life. She is a pillar of strength that was always there no matter how hard things got.

I also dedicate this thesis to my brothers and sisters, who shared advice and words of encouragement throughout the entirety.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDIX TABLES

LIST OF APPENDIX FIGURES

1. LITERATURE REVIEW

1.1. Adaptation of Bacteria

Microorganisms are ubiquitous throughout nature, and each has developed adaptation capabilities that can confer their survival in each unique habitat. Historically, bacteria have been unconsciously ignored in the evolutionary sense due to their microscopic nature, inadequate measuring techniques, and the Great Plate Anomaly (1). Now that microbiologists have more advanced technology (e.g. Next-generation sequencing) and have expanded microbiological specialties (e.g. microbial ecology, microbial biogeography (2)) we now know that bacteria can inhabit a multitude of environments and that interactions with environmental conditions and genes play key roles in influencing the resulting bacterial phenotypes. For instance, 12 different bacteria were tested for tolerance in a high salt environment and there was a range of inhibitory salt concentrations from 330 mM NaCl to 1,500 mM NaCl (3). Bacteria from similar environmental backgrounds showed a variation in tolerance such as the plant-associated bacteria *Sinorhizobium meliloti* and *Zymomonas mobilis* which had inhibitory salt concentrations at 330 mM NaCl to 1,500 mM, respectively. Since change is one of the few constants in life, shifting environmental factors , whether human-derived (4) or natural (5) can change the optimal survival conditions. Therefore, bacteria have two possibilities to survive: either vector-assisted transmission to a new secondary environment or adaptation to the changing primary environment.

Bacteria are great subjects of study for adaptation due to their physical growth attributes and their ubiquity in natural environments. In comparison to most eukaryotes, bacteria have a rapid generation time which allows microbiologists to observe adaptation within a few weeks instead of years (6). The generation times can differ depending upon species and culture

conditions (7), but usually multiple generations can be observed within a few days or weeks. Pathways to bacterial adaptation complicate matters especially with processes such as horizontal gene transfer (HGT). This capability of bacteria to be affected by the genetic compositions of surrounding microorganisms adds another factor to consider alongside environmental stimuli and the original genotype of the acceptor bacterium (8). For example, a strain of *Burkholderia cenocepacia* was found to have an adhesin factor that was acquired through HGT from a non-bacterial source, which improved the pathogenic phenotype of *B. cenocepacia* (9). Bacteria have developed adaptation mechanisms due to their lack of unassisted large-scale motility. Bacteria can be considered sessile, similar to plants, as they cannot physically and freely move from one macroenvironment to another without the aid of a vector (10, 11). Adaptation becomes a vital factor to maintain bacterial survival and growth as they cannot move away from non-optimal conditions. Therefore, studying bacterial adaptation can give insight to the impacts of various factors such as environmental stimuli on a bacterial cell.

Adaptation can occur when an environmental condition acts on a phenotype. While the environment plays a direct role with phenotypes, there is also the genetic foundation of an microorganism that affects the phenotype (12). The size of a bacterial genome has been shown to predict some adaptation expectations, as bacteria with larger genomes adapted better to changing environments rather than those that possessed smaller genomes due to the possession of genes and mechanisms needed to combat stressful conditions (13–15). For instance, bacteria that have smaller genomes are often associated with hosts, where survival conditions are more stable, while bacteria that have larger genomes are associated in soils, where abiotic factors can fluctuate in presence and magnitude (15). The adaptive mechanisms of bacteria can be generalized into four separate categories: gene acquisition (16), gene loss (17–19), gene

mutation, and gene regulation (20, 21). Gene acquisition is the initial mechanism of adaptation that is often evoked when discussing bacterial adaptation. The acquisition of antibiotic-resistant genes is the most prominent documentation of gene acquisition. The resistant clindamycin gene was found to be transferred from resistant *Acinetobacter baylyi* to sensitive *Escherichia coli* as seen by the persistence of *E. coli* in clindamycin-treated mice after 5 days (22). Gene loss can result in phenotypes that are adaptive as unnecessary or costly genes are removed. In a growth rate comparison, 25% of deletions in mutant *Salmonella* strains showed improved growth compared to the initial wildtype (23). Mutations change gene function by generating new alleles with differing protein chemistry. There are other transient forms of survival that bacteria use (i.e. persistor cells (24)), but they do not actively confront the environment and place higher priority on survival than growth.

While most studies have focused on the first three categories, the area of phenotypic adaptation through gene regulation has grown especially due to the creation and improvement of genetic expression techniques (25, 26). Quantitative polymerase chain reaction (qPCR) assays and microarrays have been utilized widely and are considered gold standards for gene expression analysis due to their sensitivity and accuracy (27). One major limitation of both of these methods is that the resulting gene expression profiles are limited by the specific transcript labels that are known (28). The introduction of RNA-sequencing (RNA-seq) has brought further depth into transcriptomics as it not only provides gene expression data to known coding genes, but also to new or non-coding genes which ultimately provides more information to researchers. This new technique has also brought fewer errors and its decreasing cost over the past years has made it available to more clientele and research projects (28, 29). Therefore using such techniques,

adaptation through regulation can be observed and mechanisms such as environmental stress memory (30) and phenotypic plasticity can be further defined.

1.2. Phenotypic Plasticity

Phenotypic plasticity is an adaptation method that involves phenotypic changes dependent upon environmental conditions that are specific to certain genotypes (10, 12, 31). For example a population of *Zizeeria maha* butterflies displayed different wing coloration patterns due to cold stress as their distribution was affected by climate change (32). Environment and genotype on their own can affect phenotypes, but in the absence of plasticity they generally portray the same trend for all environments or genotypes (Figure 1). Phenotypic plasticity is defined as an interaction between the environment and the genotypes thereby producing a unique phenotype for each strain. The balancing of these two factors aid in the pursuit of perfect plasticity, which is usually obstructed due to evolutionary tradeoffs (33). As perfect plasticity is almost never the case in most phenotypes, there is a gradient of expressed phenotypes. Genotypes may react differently to an environment than others even if the environment is the same across all genotypes (34, 35). For example, a collection of 41 *Staphylococcus aureus* isolates were subjected to an antibiotic stress, and the growth rates of the isolates varied with six isolates having improved growth rates in the stress condition (36). Surprisingly, the GWAS analysis did not show any antibiotic resistance genes being associated with the growth phenotype therefore plasticity was suggested as playing a role in the resulting phenotypes. Plasticity does not require the addition or deletion of genes within a genome but uses regulation changes to alter phenotypes.

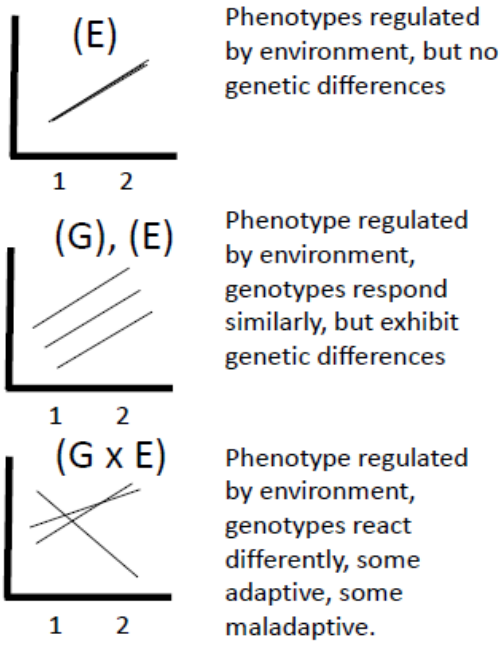

Figure 1. Models of phenotypic regulation based on environment (top), environment and genotype (middle), and the interaction of environment and genotype (i.e. phenotypic plasticity) (bottom). Modified figure (37).

While the limitations and costs of phenotypic plasticity are generally understood (12, 33), the actual biochemical mechanisms behind it are less clear. As the phenotype is specific to each genotype and to each environmental influence, the mechanism incorporates aspects of both genetic and environmental sensing machinery (11) which is a reason for its complexity. Therefore, without genetic evidence phenotypic plasticity responses can be phenotypically observed, but the specific factors that are key to the plasticity remain unknown. To determine the biochemical machinery behind plastic phenotypes, stimulus-receptor relationships and the regulation of genes are two important areas to consider since both have indirect relationships with the environment and can affect phenotype.

Adaptation by phenotypic plasticity is important for organisms that are transmitted across various environments and those that inhabit shifting environments. Sessile organisms, such as plants and microorganisms, may benefit from plasticity as they are unable to move from a

changing condition without the aid of a vector (33). Plastic responses can be triggered by environmental cues that can indicate the best phenotype. Due to their ubiquity and exposure to multiple environmental stresses, bacteria can be good models for observing phenotypic plasticity.

1.3. *Escherichia coli* Adaptation

Escherichia coli is a gram-negative bacterium that is a great model organism for adaptation as it inhabits multiple environments and has large genomic variation (38). While *E. coli* was originally documented within the intestines of warm-blooded animals and as a transient bacterium through extrahost environments, more recent studies have documented its long-term presence in the natural environment suggesting *E.coli* is free-living (39). Currently the presence of *E. coli* has been observed in natural soil and water microbiomes (40–43) beyond the expected transient population. For instance, Brennan *et al.* confirmed the presence of *E. coli* in surface soil without the introduction of enteric bacteria, via domesticates and wildlife, for over eight years (40).

In terms of genetic diversity *E. coli* has a large pangenome. A study by Rasko *et al.* classified *E. coli*'s pangenome as open due to the continuous addition of new gene families with the inclusion of each new *E. coli* genome (44). This gives *E. coli* an endless possibility of new genes being incorporated into individual strain genomes. One study analyzed *E. coli* genomes of 53 different strains which led the pangenome to encompass 13,296 gene families. While different strains can have a varying diversity of accessory genes, there are at least 1,000- 2,000 gene families that are part of the core genome of all *Escherichia coli* strains (45, 46). The genome of each *E. coli* is also rife with mobile elements and genomic islands (47) that incorporate genetic elements from the nearby environment and surrounding microbial populations.

1.4. Sigma Factors and RpoS

RpoS, a sigma factor, is a major player in *E. coli* stress responses that has effects on adaptation via plastic phenotypes. A sigma factor is a transcriptional regulator of genes that guides RNA polymerase to certain genes based on the promoters recognized by the sigma factor (48). RpoS is an alternative sigma factor that binds promoters of stress-resistance and stationary phase genes. In a single *E. coli* K12 genome RpoS controlled 140 out of 1,352 genes (~10%), either directly or indirectly (49). This study alongside others have therefore marked RpoS as one of the largest global regulators in *E. coli* (49–52). The majority of its regulon is associated with internal and external environment responses (53, 54) earning its title of general stress regulator.

There are many regulators that respond to environmental stimuli, both sigma factors and other regulators, but few cover the number of genes within the RpoS regulon. In total there are seven different sigma factors that have been identified in *E.coli* (55) . RpoS is not the only sigma factor and must therefore compete with other regulators for the limited amount of RNA polymerase (48, 55). For instance, RpoS exhibits an antagonistic relationship with RpoN, a sigma factor known for regulating nitrogen utilization genes (56). Another sigma factor, RpoE, regulates outer membrane stress responses within the cell envelope due to the presence of unfolded proteins (57). Both sigma factors, while important on their own, can be affected by RpoS regulation. In the case of RpoE, one of its promoters is regulated by RpoS during hyperosmotic conditions (58).

The RpoS regulon is a good candidate for phenotypic plasticity due to its range of functions and the impact on overall phenotype. There are a multitude of stresses that can be linked to RpoS regulon responses such as temperature, acidity, moisture, and osmotic pressure (59–62). Besides combating one stress at a time, RpoS is also a multitasking regulator that assists

in cross-protection, where one stress induces responses to a multitude of other stresses (59, 63). The influence of genotype and environment has been observed in RpoS-dependent membrane phenotypes (64–66) where mutations of *rpoS* have produced dramatic phenotypes suggesting a role of phenotypic plasticity for its regulon. One such example has been termed as the self-preservation and nutritional competence (SPANC) model (67). This model describes the inverse relationship of cell membrane permeability between nutrient-scavenging and stress tolerance conditions. The knockout of *rpoS* demonstrated low stress tolerance and high nutritional capability. The growth advantage at stationary phase (GASP), which describes the competitive advantage of growth at stationary phase caused by mutations (68), is a prime example of the ways in which *rpoS* mutations can lead to plastic phenotypes. The proportion of *rpoS*-containing isolates reduced in generations across nutrient-limiting conditions such that after 20 generations less than 1% of the isolates contained *rpoS* (69). Therefore, the presence and absence of RpoS' regulon showed vast phenotypic changes and may be involved in the plasticity of stress-related phenotypes.

While RpoS has been intensively studied in laboratory and some pathogenic strains of *E. coli* due to their direct impacts on human society, research with natural strains has become more of an underworked issue (39). For instance Carter *et al.* discovered RpoS, via positive regulation of CsgD, regulated curli production in *E. coli* K12 strains, but that was not seen in *E. coli* O157 strains (70). The use of K12 strains with RpoS, while useful in understanding basic machinery and genomic composition, is not indicative of the *E. coli* species as they have been domesticated for decades. Also previous research has found that older collections of commensal isolates, from the 1980s and 1990s, are more likely to lose their original *rpoS* gene due to their time in

insufficient storage (71). Fresher and more diverse strains of *E. coli* can further our understanding of RpoS' role with fewer worries regarding its loss in domestication.

1.5. Bacterial Membrane Stress

One of the largest systems within *E. coli* that is affected by environmental stresses is the bacterial envelope. External (i.e. temperature, osmotic, pH) and internal (i.e. oxidative) stressors impact the integrity of the bacterial membrane (21). One important stressor that is relevant to most environments that *E. coli* inhabit is osmotic stress. This stress is generally exerted through high salt concentrations and desiccation (72, 73). Osmotic stress causes a loss of water to the extracellular environment and a decrease in turgor pressure on the cell wall (Figure 2). This stress can cause a decrease in growth rates due to effects on water potential within the cell and peptidoglycan expansion (74). The loss of water due to osmotic stress also impacts protein stabilization as proteins can become misfolded or denatured (Figure 2). The osmotic balance is crucial to the maintenance of biochemical processes and especially the integrity of the bacterial membrane envelope (75). Bacteria can protect themselves from such irregular osmotic conditions through external protection (i.e. biofilms and capsules) and molecular protection. For example, during osmotic shocks *E. coli* uptake potassium ions to maintain water balance, whereas in prolonged stress *E. coli* will synthesize organic osmoprotectants (aka osmolytes), such as trehalose, to maintain turgor pressure which affects the integrity of the cell wall (76). Other important osmotic tolerance and resistance mechanisms include osmolyte synthesis, cell membrane composition, and transcriptional regulators (73, 77). The latter has been observed with IrrE, a global gene regulator, which when overexpressed in *E. coli* under high salt conditions showed increased cell growth and osmolyte synthesis in comparison to the control (77).

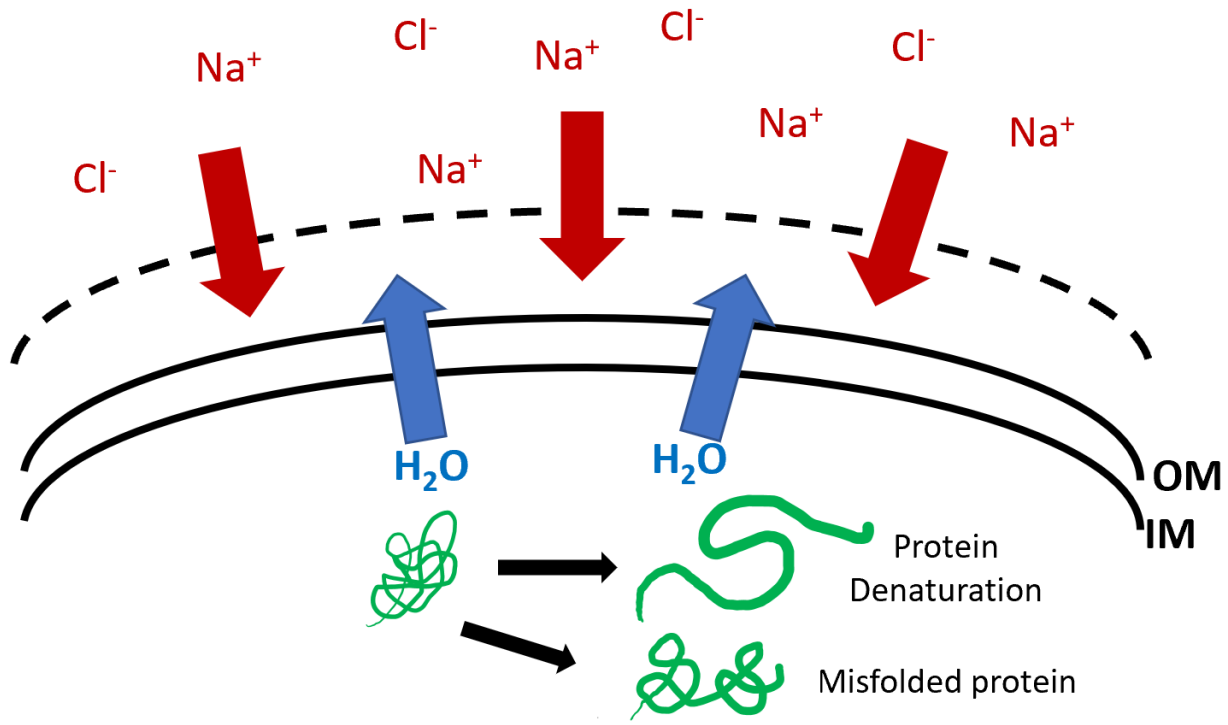

Figure 2. Osmotic stress effects on gram-negative bacterium caused by high salt. The high concentration of salt (NaCl in this case) decreases the turgor pressure of a cell (red arrows) and causes water to be lost into the extracellular environment (blue arrows). The decrease in turgor pressure and water causes the cell to shrink from its original size (dotted line). Loss of water also causes the misfolding or denaturation of proteins (green). OM= outer membrane and IM= inner membrane.

E. coli combats against internal and external membrane stress using mechanisms such as the Cpx system, the PhoPZ Two-Component System (TCS), and the RpoE system (57, 78, 79). While each system enacts many of its own responses to specific stressors, a bacterial cell can be cross-protected (80) due to a link (i.e. gene regulator) connecting parts of each system. For example, OmpR, an important factor in regulating porins within the bacterial outer membrane, is an important regulator for both acid and osmotic stress responses (81). Sigma factors can induce osmotic stress preventatives as seen with RpoS (82) and RpoE (57). Previous research has defined important proteins (i.e. outer membrane proteins [Omps]) and specific osmotic

pathways, but the implication of these and other important osmotic proteins and pathways based on gene regulation have not been fully explored in bacteria.

The lack of research on phenotypic plasticity and on its mechanisms prompts further understanding of its extent in bacteria. As the environment and genotype of bacteria play important roles on their own, it is important to determine the interaction of these two factors on phenotype. RpoS is known for its extensive regulation in stress-related conditions, so it is a prime candidate for being a keystone in phenotypic plasticity. In this thesis, our goal is to describe the presence and extent of phenotypic plasticity among *E. coli* from soil by growth assays. For exploration into possible factors in phenotypic plasticity we used RNA-sequencing (RNAseq).

2. POTENTIAL FOR PHENOTYPIC PLASTICITY IN THE HYPEROSMOTIC STRESS RESPONSE OF DIVERSE *E. COLI*

2.1. Introduction

Bacteria, known for their ubiquity and flexible genetics, adapt to shifting environments but the extent and possible mechanisms are still being explored. Phenotypic plasticity is a form of adaptation that involves phenotypic changes dependent upon environmental conditions that are specific to certain genotypes (31, 33). Phenotypic plasticity accounts for the interaction of genotype and the environment on a resulting phenotype. For example, 41 *Staphylococcus aureus* strains displayed different growth rates in normal and vancomycin-supplemented media with six strains having higher growth rates in the vancomycin condition that was not due to vancomycin-resistant genes (36). There are a few microbial adaptation studies that define and explore phenotypic plasticity. Feugeas *et al.* found that the transcriptome profiles of four *E. coli* strains varied based on the three media conditions as well as the strains themselves (83). They observed genes associated with metabolic processes were linked with media and genes involved in environmental signaling were associated on a strain-basis. Altogether, the interaction of these two factors influenced the resulting gene expression profiles. While phenotypic plasticity has been seen in such studies, the extent and prominence of this adaptation method in bacteria has not been fully characterized and explained.

Hyperosmotic conditions, such as high salt conditions, affect the cell envelope and growth of bacteria. The cell promotes transport and synthesis of osmolytes to maintain turgor pressure. Such osmolytes include charged solutes such as potassium ions as well as compatible solutes such as trehalose. The concentration of trehalose was seen to be effective in growth of *E. coli* W3110 strains as a mutant strain that overproduces trehalose had higher growth under

increasing NaCl conditions (84). The negative effect of high salt environments on bacterial growth is generally associated with the cell's additional energy expenditure and slowdown of translation (85). On a genetic level, a strain of *E. coli* K12 highly expressed genes that were important in osmolyte synthesis and transport (e.g. *proU* and *kdpFABC*) (86). Even further the gene expression resulting from osmotic stress can vary between strains, as three out of five *E. coli* strains had significant induction of *betA*, a gene important for glycine betaine biosynthesis (87).

Sigma factors are important in gene transcription under different environmental conditions. These transcription factors bind with RNA polymerase and lead to the regulation of genes based on stress stimuli. The sigma factor RpoE is vital for the integrity of the cell envelope as cell growth decreased when *rpoE* was inhibited (88). RpoS is the general stress regulator in *E. coli* and is known to be vital in responses to stationary and environmental stresses (49–52). The functions of the RpoS' regulon vary widely and many RpoS-regulated genes showed beneficial phenotypes in stressful conditions (62, 89). The importance of RpoS in the extensive number and functions of its regulon, suggests its importance in adaptation and specifically in gene regulation.

We hypothesized that there was phenotypic plasticity of a variety of *E. coli* under hyperosmotic stress. Therefore, we expected to see differences in growth rates of each strain under each osmotic condition. We also hypothesized that the mechanism behind phenotypic plasticity may lie within sigma factors, particularly the general stress factor RpoS. To test this assumption, we looked at the transcriptomes of nine *E. coli* isolates to determine the genetic expression of osmotic responses amongst the isolates.

2.2. Materials and Methods

2.2.1. *E. coli* Isolates

The *E. coli* isolates used in the following experiments and assays are a part of an *E. coli* isolate collection which was obtained from surface soil samples along the Buffalo River of North Dakota and Minnesota (90). The collection includes 256 culturable phylogroup D isolates, select sets of which were used in our experiments.

2.2.2. qPCR

Comparing different isolates required an initial check in the baseline expression. As we were looking at RpoS as a factor for phenotypic plasticity it was important to determine any differences in baseline RpoS-regulon expression. We measured the gene expression of seven RpoS-regulated genes (*asr*, *cpxP*, *degP*, *osmC*, *osmY*, *pspA*, and *rpoH*) at three different growth phases (exponential, transitional, and stationary) through a qPCR assay.

Starting with the RNA samples, eight *E. coli* strains were selected and grown from freezer stocks onto Lysogeny Broth (LB) agar. The isolates for the qPCR experiment were selected based on their capability to culture successfully and consistently in three passages of glucose-defined minimal media (GDMM) (91) based on other studies and assays performed in our lab. Isolates were blocked into two groups of four for culturing. To acclimate the *E. coli*, 16 x 100 mm test tubes that each contained 5 mL of standard GDMM were each inoculated with one colony of its respective isolate from the LB agar. The cultures were grown overnight (~16 hours) in a shaking incubator set at 37°C and 220 rpm. Cultures were passaged two more times with RNA sampled from the tertiary culture. At the designated time points (exponential, transitionary, and stationary) when optical density readings at 600 nm (OD_{600}) reached their respective ranges (0.6-0.8, 0.9-1.1, and ≥1.2), 5 mL of tertiary culture was aspirated and placed

into sterile 16 x 100 mm tubes filled with a stop solution of 10% acid phenol in ethanol. RNA was extracted and isolated using phenol-chloroform and DNase methods as detailed by Tyagi *et al.* (92). The purity and concentrations of the RNA samples were initially checked with a spectrophotometer (Bio-Rad, Hercules, CA) and final values were verified with an RNA integrity assay (Agilent, Santa Clara, CA). Purity was calculated by the RNA integrity number (RIN) score with the Bioanalyzer 2100 v.B.02.08.SI648 (Agilent). The range of RIN scores is 1 to 10 with the lower values being associated with degraded RNA and the higher values being associated with intact RNA. In our study samples were considered pure if their RIN value was ≥7. In total, there were three biological replicates at each time point for each strain (72 RNA samples).

The qPCR experiment also required plasmids to assess PCR efficiency for each RNA sample. We generated 8 different plasmids: seven plasmids for the respective RpoS-regulated target genes as well as one plasmid for the PCR efficiency and qPCR $\Delta\Delta C_t$ calculation, *arcA*. This gene was used as our qPCR control as it is a housekeeping gene in most *E. coli* isolates and is used as one of a few housekeeping genes in *E. coli* gene expression (93). The plasmids were created by performing PCR, for gene amplification, and cloning, for inserting each target gene into their own pCR 2.1-TOPO plasmid (Life Technologies, Carlsbad, CA). These plasmids were purified via QIAprep Spin kit (Qiagen, Hilden, Germany) and verified with 1% agarose gel electrophoresis.

Gene expression was determined using SYBR-green qPCR. A total of two technical replicates were implemented with each technical replicate including all RNA samples, three negative controls for each gene, and a qPCR standard dilution curve. The PCR calibration curve is used to account for any differences in PCR efficiency between all the target genes. PCR

efficiencies for the seven target genes ranged from 2.067 ± 0.076 to 2.294 ± 0.091. These values describe that a target gene doubles in copies after each PCR cycle, which is the expected result of PCR amplification. The final gene expression values ($\Delta\Delta C_t$) were calculated and adjusted using the Pfaffl equation (94) and the raw Cycle threshold (C_t) scores. The exponential phase values were used as the controls and the *arcA* values were used as the reference. A Levene's test and a two-way ANOVA were used to determine gene expression significance between the isolates as well as phenotypic plasticity determination.

2.2.3. Growth Curves

Twenty-four selected *E. coli* isolates were inoculated onto LB agar from freezer stock cultures and incubated at 37°C for 24 hours. The isolates for the growth assays and RNAseq experiments were selected based on prior experimentation with osmotic growth conditions (0.3% NaCl, 3% NaCl, and 4% NaCl GDMM) in our laboratory. The 24 out of 256 isolates were chosen based on their growth phenotypes (growth rate and growth yield) seen in 4% NaCl GDMM. Twelve isolates were randomly selected from the top third of the total isolates (fast growth rates) and the other twelve isolates were randomly selected from the bottom third (slow growth rates). To acclimate the *E. coli*, the isolates from the LB agar were inoculated into 16 x 100 mm test tubes that each contained 5 mL of standard (0.3% NaCl [w/v]) GDMM (91). A minimal media was used to limit the amount of variation of the media on the stress condition. The cultures were placed in a shaking incubator set at 37°C and 220 rpm. These primary (1°) cultures were grown for 8 hours. For the secondary (2°) cultures, 0.1 mL of the 1° cultures were transferred to 10 mL of sterile, pre-warmed (37°C) standard GDMM. The 2° cultures were incubated at 37°C at 220 rpm for overnight (~16 hours). This transfer procedure was performed

again, for 3° cultures, with the exception that the 3° cultures were incubated for 8 hours instead of 16 hours.

In order to observe the effect of osmotic stress on these *E. coli* isolates, two conditions were used for the test cultures: a standard condition (0.3% NaCl GDMM) and a hyperosmotic condition (4% NaCl GDMM). The concentration of 0.3% NaCl is standard in the GDMM as it provides the optimal salt concentration for *Escherichia coli* growth at 37°C. The hyperosmotic condition magnitude was determined based on prior growth assays in our lab as well as through other *E. coli* osmotic growth assays in the literature (49, 82, 95). The test cultures were prepared by transferring 0.1 mL of each isolate's 3° culture into both 0.3% NaCl GDMM and 4% NaCl GDMM (10 mL in each test condition). The initial OD_{600} readings were taken after the culture transfer and all test cultures were then incubated at 37°C and 220 rpm. Optical density readings were recorded until cultures reached an OD_{600} measurement around 0.7 indicating mid-exponential phase. Our samples had final OD_{600} measurements within the range 0.688- 0.800. Nine of these cultures were further used in the RNAseq analysis.

To determine the growth rate of the isolates grown at 0.3% NaCl and at 4% NaCl, the OD_{600} readings were compiled, log-transformed, and analyzed in R using the R/nlsMicrobio "baranyi_without_Nmax" growth model and R/vegan packages (96, 97).

2.2.4. RNAseq

RNAseq was implemented to determine and explore genetic factors of phenotypic plasticity of the *E. coli* growth rates under osmotic stress. Nine of the strains, and their respective cultures in both osmotic conditions, from the hyperosmotic growth assay were used in the RNAseq analysis. When these cultures reached an OD_{600} around 0.7 (actual OD_{600} range of 0.688-0.754), 0.5 mL of each culture was transferred to 0.5 mL of sterile GDMM (0.3% or 4%

NaCl depending on culture) and centrifuged to form a cell pellet. The GDMM of each microcentrifuge tube was aspirated and the pellets were snap-frozen in liquid nitrogen.

RNA was extracted from each sample cell pellet by using a RNeasy Mini kit (Qiagen). Sample purity and concentration were initially assessed by spectrophotometer (Bio-Rad) and then ultimately verified by an RNA IQ fluorescent assay (Thermo-Fisher, Waltham, MA). The quality control parameters of a pure sample were 260/230 ratio >1.7 and 260/280 ratio >1.9. RNA samples were considered undegraded if their RNA IQ values >7.

RNA libraries were prepared using a total RNA library prep kit (Illumina, San Diego, CA). The library preparation was performed as stated by the manufacturer's protocol except for additional wash steps at the end of the protocol due to bead residue. All quality assurances of the finished library were determined through a DNA integrity assay (Agilent) and a picogreen DNA concentration assay (Invitrogen, Carlsbad, CA). Sufficient library samples were characterized by having a high DNA peak (ranging anywhere from 200-1000 bp), low background peaks, and DNA concentrations of at least 5 nM. RNA samples were sequenced using pair-end 150 bp reads on a HiSeqX platform (Psomagen Inc., Rockville MD USA). Raw reads were initially quality-checked by FastQC (98). Due to base bias at the beginning of each read and adapter content, the reads were trimmed using Trimmomatic (99). After trimming reads for each sample, FastQC was implemented again to ensure quality after trimming. The sequence counts per sample averaged around 52.7 million reads. Reads were then mapped to their respective isolate genome through bwa v.0.7.17 (100) and were checked with MultiQC v.1.9 (101). Due to lack of directly associated genomes, the samples of LGE1964 and LGE1517 were mapped to genomes LGE1946 and LGE1527, respectively. The program *featurecounts* (102) counted reads per feature with an average of 80% of the reads per sample being mapped and assigned a feature. The sample reads

where then collated with gene presence and absence files from Roary (103). Microorganisms have a vast range of gene content and a collection of isolates can have core genes, which are genes shared in all isolates. As we were observing osmotic factors among all of the isolates, we examined the conserved genes of our isolates with *edgeR* (104) for differential gene expression analysis. Also "NA" values cannot be accepted into *edgeR*, therefore genes that had a "NA" value in at least one sample were removed from the *edgeR* analysis. A total of 3,234 genes were conserved in all nine *E. coli* isolates.

Using *edgeR*, via R, the collated reads were first processed for gene coverage so that the genes that had little to no coverage, < 1 count per million in at least five samples, were removed. The library sizes of the count data were then normalized using the trimmed mean of M-values (TMM) (105). This method calculates the gene-wise $\log_2$ fold changes between samples which are the M-values. The observed M values are then trimmed, upper and lower 30%, prior to calculating the weighted average. Each sample library is then normalized by dividing the raw counts by the TMM-adjusted library sizes. The overall normalized libraries had a median of 70.3 million read counts. After normalization, the read counts were modeled according to the environment condition (High Salt [4% NaCl] and Standard [0.3% NaCl]) and were fitted with the negative binomial model. The false discovery rate (FDR) of the genes were corrected using the Benjamini and Hochberg method (106). Visualization of the data was performed with R (96) using the R/ggplot2, R/gplots and *R/edgeR* libraries. For functional analysis, the $\log_2$-fold change (logFC) and p-value of each gene from the *edgeR* analysis was inputted into the EcoCyc database (107) and was visualized with the Omics Dashboard (108). Enrichment analysis was performed by using a p-value < 0.01 resulting in enrichment scores, which were calculated by

the Grossman parent-child-union analysis (109). An enrichment value of ≥2 was considered significant, as the enrichment score was -log(p-value).

2.3. Results

2.3.1. E. coli Gene Expression of RpoS-regulated Genes

Phenotypic plasticity relies on a difference in genotypes as well as phenotypes. Therefore, it was important to investigate the baseline of plasticity through the gene expression of RpoS-regulated genes amongst a selection of *E. coli* isolates. Between all eight isolates, the RpoS-regulated genes displayed no significant variance during transitional phase (Levene Test [LT] p-value= 0.795) (Figure 3). The isolate and the gene significantly influenced the expression levels, but their interaction was insignificant (Table 1). In contrast, at least one isolate differed from the others with significant variance (LT: p-value=2.48×10^{-4}) in stationary phase (Figure 4). The difference in expression during stationary phase occurred in the genes *asr* (LT: p-value= 7.33×10^{-3}), *osmY* (LT: p-value=3.50×10^{-4}), and *degP* (LT: p-value=8.58×10^{-6}) (Figure 5). The factors of isolate and gene were both significantly important factors in the expression levels, but most notable was that the interaction between these two factors which was also significant (Table 2). Overall, we found an initial insight into phenotypic plasticity amongst the *E. coli* isolates under one condition.

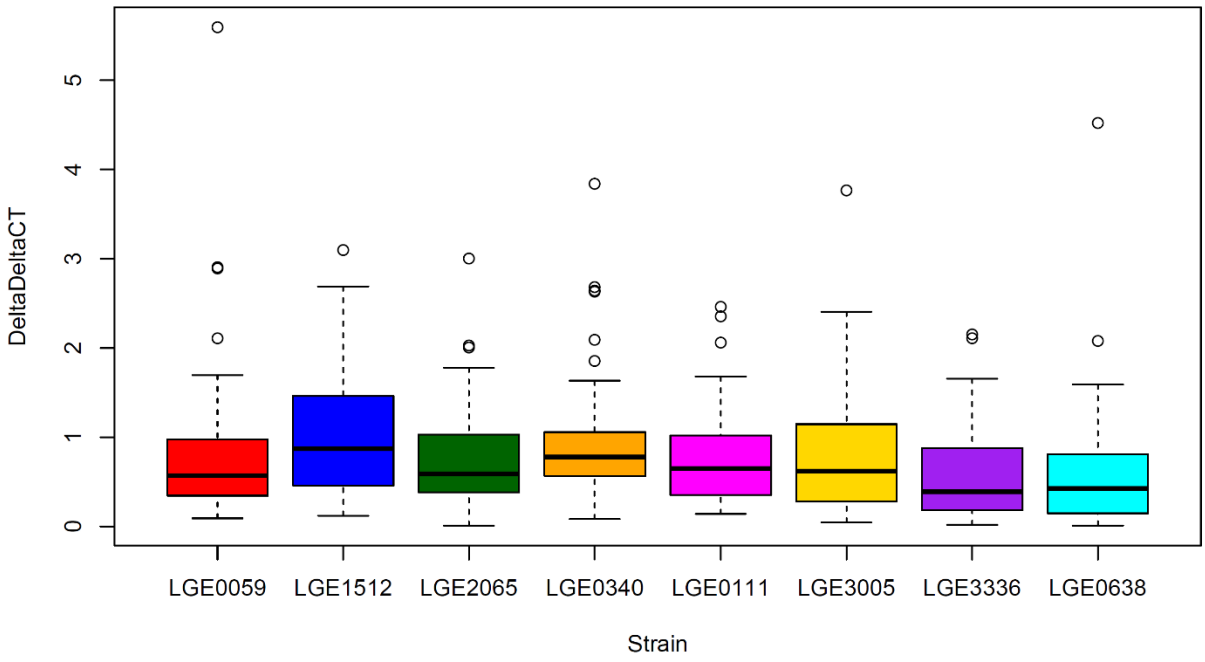

Figure 3. Box-and-whisker plot of RpoS-regulated genes across isolates in transitional phase.

Table 1. Two-way ANOVA table for gene expression in transitional phase. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Variable	Df	Sum Sq	Mean Sq	F value	Significance
Isolate	7	7.93	1.133	2.257	0.0301 *
Gene	6	395.6	65.94	18.754	6.85 x 10-7 ***
Isolate x Gene	42	16.55	0.394	0.785	0.8271
Residuals	276	138.54	0.502		

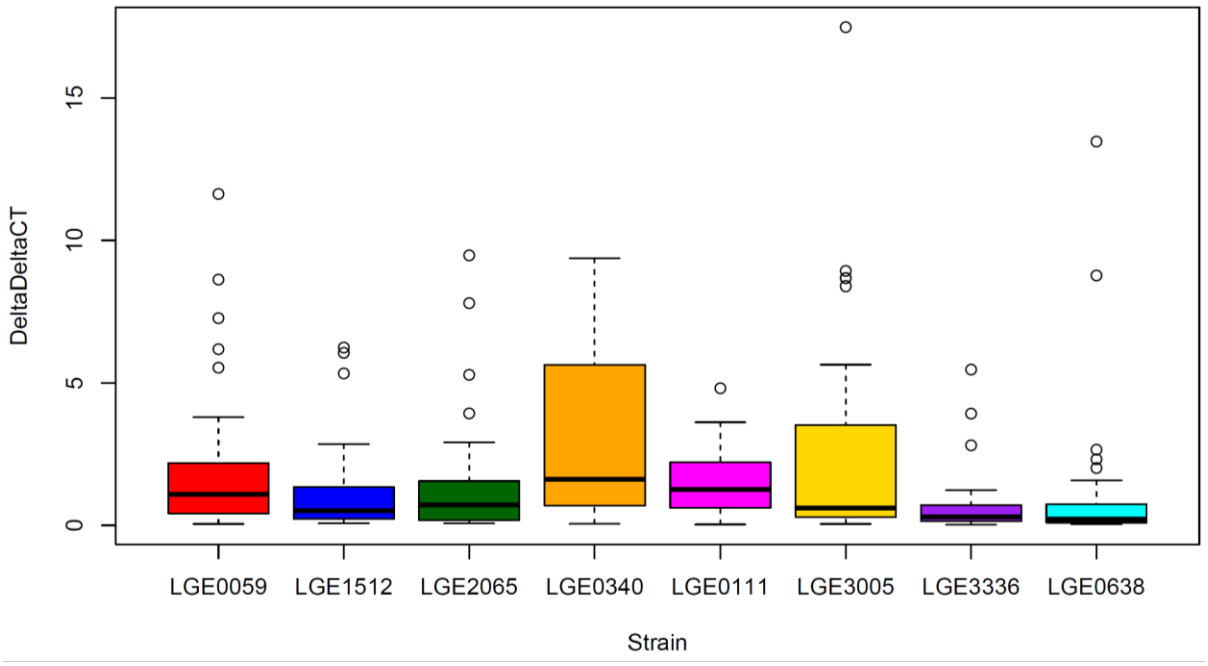

Figure 4. Box-and-whisker plot of RpoS-regulated genes across isolates in stationary phase.

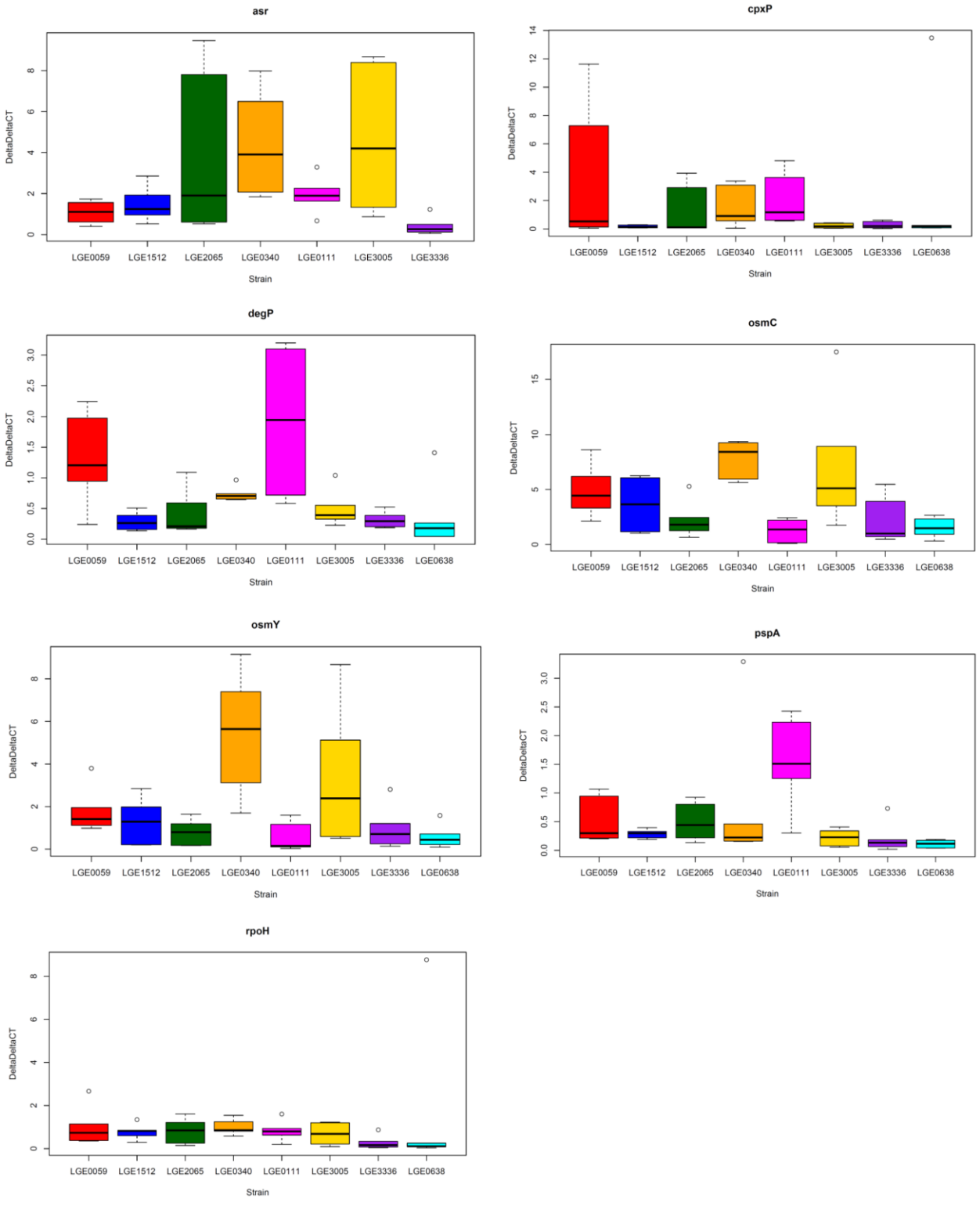

Figure 5. Box-and-whisker plots of RpoS-regulated genes across isolates in stationary phase. For the *asr* gene there was insufficient data for LGE0638 hence its absence.

Table 2. Two-way ANOVA table for gene expression in stationary phase. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Variable	Df	Sum Sq	Mean Sq	F value	Significance
Isolate	7	177	25.29	7.192	6.66×10^{-8} ***
Gene	6	395.6	65.94	18.754	$< 2.00 \times 10^{-16}$ ***
Isolate x Gene	42	414.7	9.87	2.808	2.72×10^{-7} ***
Residuals	272	956.4	3.52		

2.3.2. Growth Rates under Hyperosmotic Stress

Our first step in determining phenotypic plasticity was determining the variation in growth rates among the different isolates between two osmotic conditions. In the 0.3% NaCl condition, the growth rates of the *E. coli* isolates ranged from 0.677 ± 0.062 $\log_{10}(OD_{600})$ h^{-1} to 0.841 ± 0.013 $\log_{10}$ (OD_{600}) h^{-1} with three replicates of each isolate (Table 3). In contrast to the lower salt environment, there was a larger range of growth rates in the 4% NaCl condition as they spanned from 0.069 ± 0.021 $\log_{10}(OD_{600})$ h^{-1} to 0.323 ± 0.157 $\log_{10}(OD_{600})$ h^{-1}. The large range of growth rates in the 4% NaCl condition can be attributed to a few outlying strains having exceedingly slower growth rates such as LGE0109 and LGE1964. As expected in the 4% NaCl condition, the median growth rate of the isolates (0.218 $\log_{10}[OD_{600}]$ h^{-1}) was lower than the 0.3% NaCl rates (0.746 $\log_{10}[OD_{600}]$ h^{-1}) (Figure 6). These two conditions showed a significant difference in growth rates ($p < 2.2 \times 10^{-16}$).

Interestingly the isolates shifted their ranks according to their growth rate in the two osmotic conditions (Table 3). Isolates such as LGE1964 and LGE3205 showed large rank shifts in their growth rates from the 0.3% to 4% NaCl environments. In contrast, other isolates (e.g. LGE1523 and LGE2400) remained within the same ranking in both environments. Not only did the magnitudes of shifts occur, but the direction of the shifts also changed. For instance, some isolates that possessed higher growth rates in the 0.3% NaCl environment displayed lower

ranking growth rates in the 4% NaCl environment (e.g. LGE1576 and LGE0105). The inverse of this pattern was seen in isolates such as LGE3205 and LGE3262.

Table 3. Shift in ranks for growth rates amongst selected *E. coli* isolates in 0.3% NaCl and 4% NaCl Conditions. Ranks and growth rates of 24 *E. coli* isolates grown in 0.3% and 4% NaCl. Averages and standard deviations determined by three replicates per strain and condition.

Strain	**0.3% NaCl Condition**		**4% NaCl Condition**		**Growth Rate Magnitude between Conditions**
	Growth Rate	Rank	Growth Rate	Rank	
LGE1576	0.841 ± 0.013	1	0.214 ± 0.019	13	0.627
LGE0733	0.804 ± 0.040	2	0.238 ± 0.014	7	0.566
LGE3070	0.797 ± 0.015	3	0.239 ± 0.002	6	0.558
LGE2134	0.787 ± 0.007	4	0.205 ± 0.018	16	0.583
LGE1517	0.778 ± 0.050	5	0.323 ± 0.157	1	0.455
LGE1523	0.776 ± 0.026	6	0.241 ± 0.016	5	0.535
LGE1964	0.769 ± 0.021	7	0.069 ± 0.021	24	0.699
LGE0406	0.766 ± 0.027	8	0.210 ± 0.022	14	0.556
LGE3195	0.766 ± 0.057	9	0.254 ± 0.010	2	0.512
LGE0105	0.763 ± 0.067	10	0.119 ± 0.083	23	0.644
LGE2070	0.759 ± 0.017	11	0.249 ± 0.025	3	0.511
LGE1459	0.746 ± 0.031	12	0.177 ± 0.049	18	0.569
LGE1494	0.745 ± 0.023	13	0.126 ± 0.056	22	0.618
LGE0359	0.739± 0.023	14	0.228 ± 0.014	8	0.512
LGE1650	0.738 ± 0.035	15	0.223 ± 0.004	10	0.515
LGE2400	0.731 ± 0.018	16	0.208 ± 0.026	15	0.523
LGE0399	0.728 ± 0.035	17	0.222 ± 0.035	11	0.507
LGE1536	0.727 ± 0.078	18	0.175 ± 0.045	19	0.552
LGE0109	0.725± 0.006	19	0.156 ± 0.035	20	0.569
LGE3205	0.722 ± 0.080	20	0.226 ± 0.022	9	0.496
LGE1448	0.708 ± 0.041	21	0.192 ± 0.013	17	0.516
LGE2761	0.7 ± 0.049	22	0.247 ± 0.010	4	0.453
LGE3262	0.686 ± 0.046	23	0.221 ± 0.009	12	0.465
LGE3336	0.677 ± 0.062	24	0.135 ± 0.042	21	0.542

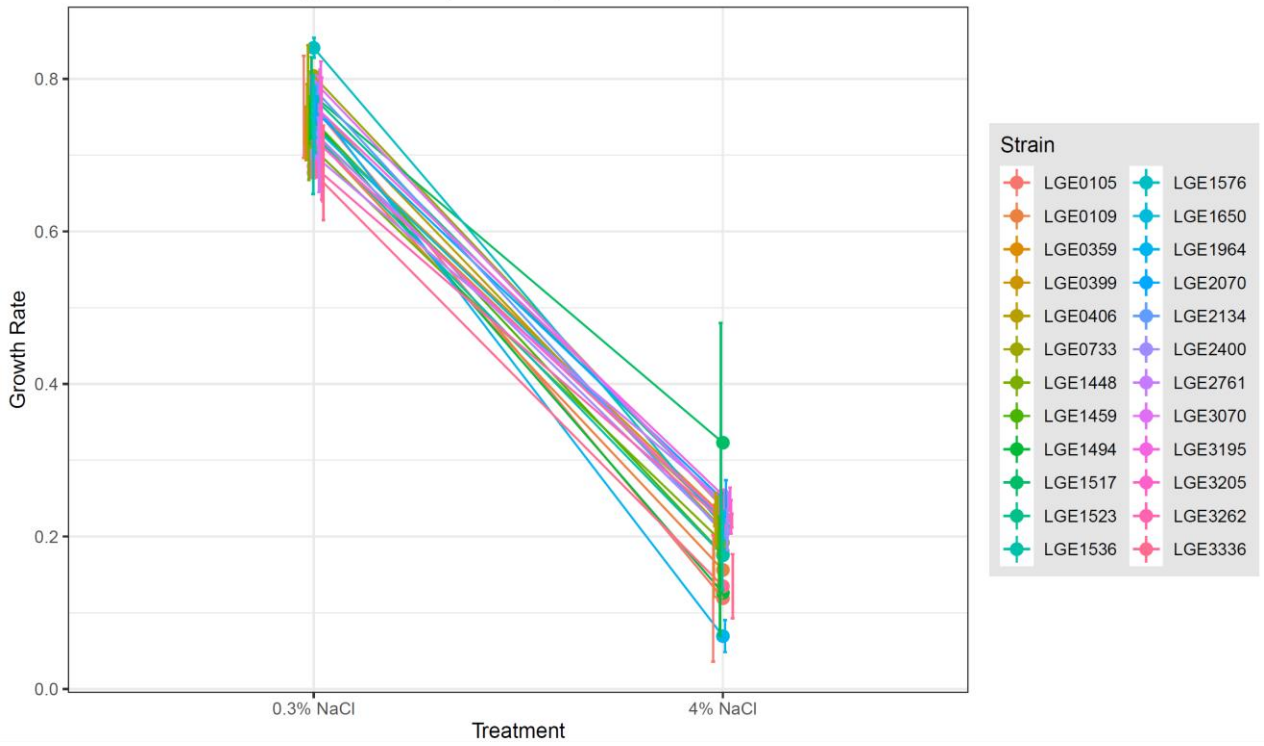

Figure 6. Norm reaction plot for growth rates of the 24 *E. coli* isolates grown in 0.3% and 4% NaCl. All growth measurements were taken in triplicate.

2.3.3. Conserved Gene Expression in E. coli Isolates

We were not able to account for phenotypic plasticity in our RNAseq analysis, but we did analyze the effect of hyperosmotic stress among different *E. coli* isolates and determine conserved responses of genes to the stress. Nine *E. coli* isolates were selected and collectively they showed the plasticity patterns that were observed in the growth rate phenotypes under the two osmotic conditions (Figure 7).

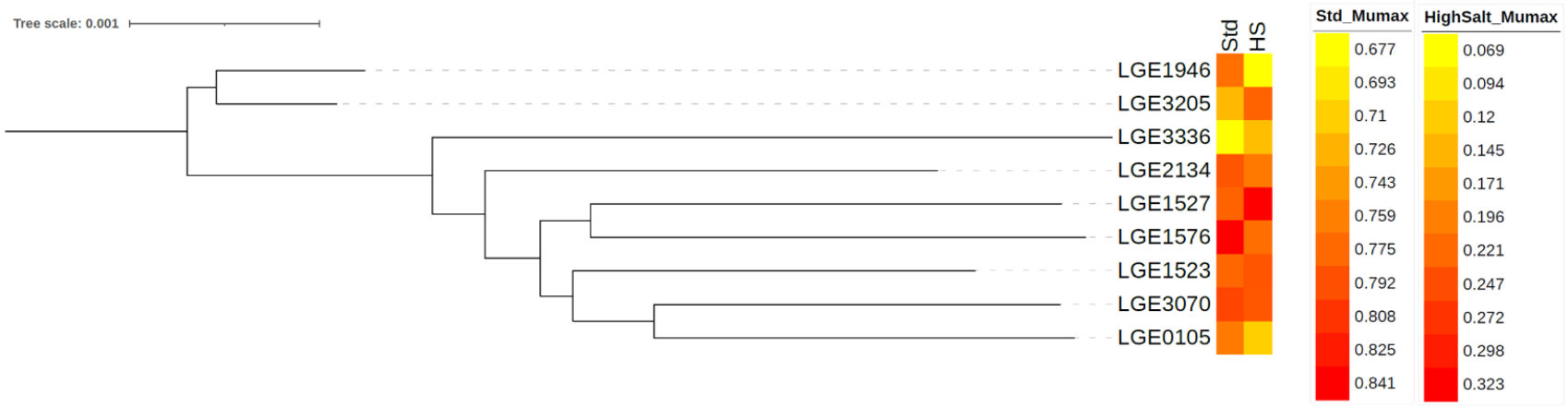

Figure 7. Phylogenetic tree with heatmap of the selected nine *E. coli* isolates for gene expression analysis. Phylogenetic relations are based on core genes. The heatmaps represent their growth rate phenotypes under each osmotic environment. Std= standard condition (0.03% NaCl), HS= high salt condition (4% NaCl), Mumax=growth rate. Visual representation performed in Interactive Tree of Life (iTOL) v6.1.2 *(110)*.

The transcriptome contained reads mapping to 3,234 genes that were present in all 9 strains in both conditions. The $\log_2$-fold change (logFC) of the genes ranged from -6.423 to 8.093 with a median logFC of -0.631. Of those conserved genes 914 genes were downregulated ($\text{logFC} \leq -1$), 814 genes were upregulated ($\text{logFC} \geq 1$), and 1,506 genes did not differ ($-1 < \text{logFC} < 1$) (Figure 8). Among the significant logFC genes 1,522 genes had a FDR (False Discovery Rate) < 0.05 and a p-value < 0.01. We then used transformed and rescaled count-per-million (CPM) data of the conserved genes to observe differences between the two osmotic conditions and differences between the isolates. Genes with conserved responses to high salt in the isolates showed distinct clustering between the standard (0.3% NaCl) and the high salt (4% NaCl) conditions (Figure 9). The dendrogram above the heatmap show that the transcriptome profiles were initially split due to the osmotic condition of the samples. Interestingly, the transcriptome profiles in the high salt condition were more clustered together than profiles in the standard condition as seen in the heat map dendrogram as well as in a PCA plot (Figure 10). The importance of the osmotic condition was the biggest factor for the separation of the transcriptome profiles (PC1= 76.7%). A secondary factor that influenced the differentiation of the transcriptome profiles was the isolates themselves (PC2= 9.76%). Amongst the isolates, the transcriptome profiles were branched differently within each osmotic condition (Figure 9). For example, the transcriptome profile of LGE1946 in the standard condition clustered with LGE0105 and LGE3336, but that was not the case in the high salt condition. Overall, the transcriptome profiles displayed distinct separation based on osmotic conditions.

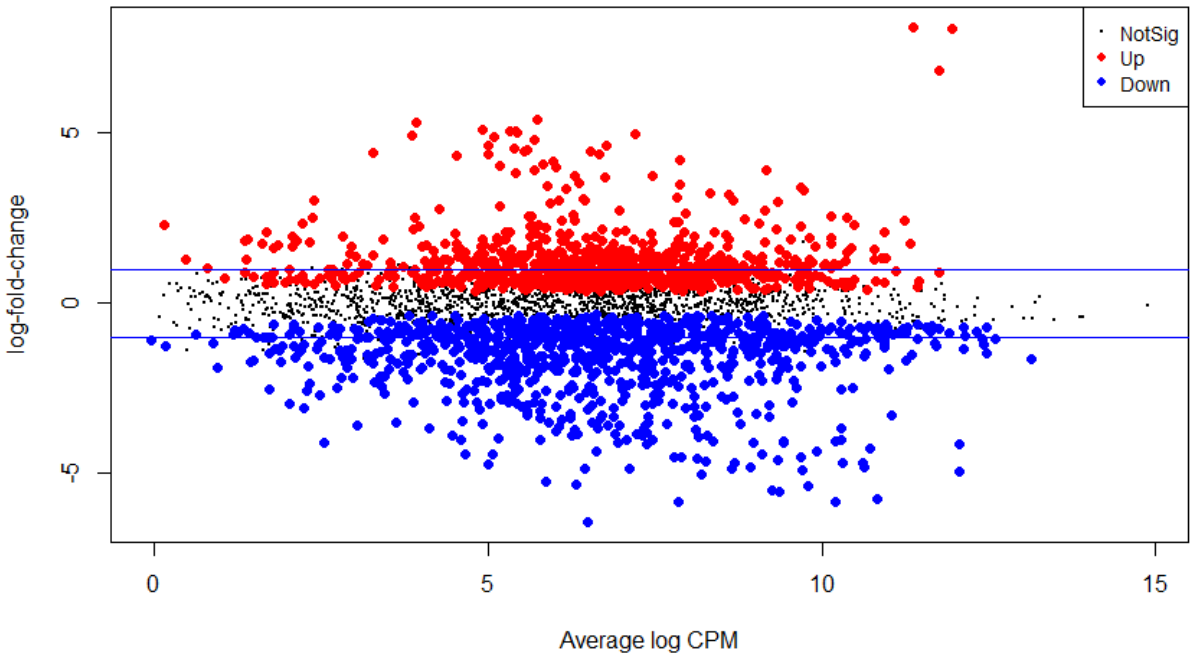

Figure 8. MA plot displaying logFC of conserved genes amongst all nine *E. coli* isolates, with default log fold-change thresholds of -1 and 1 (horizontal blue lines). Up= upregulated genes, Down= downregulated genes, NotSig= insignificant genes, CPM = Counts Per Million.

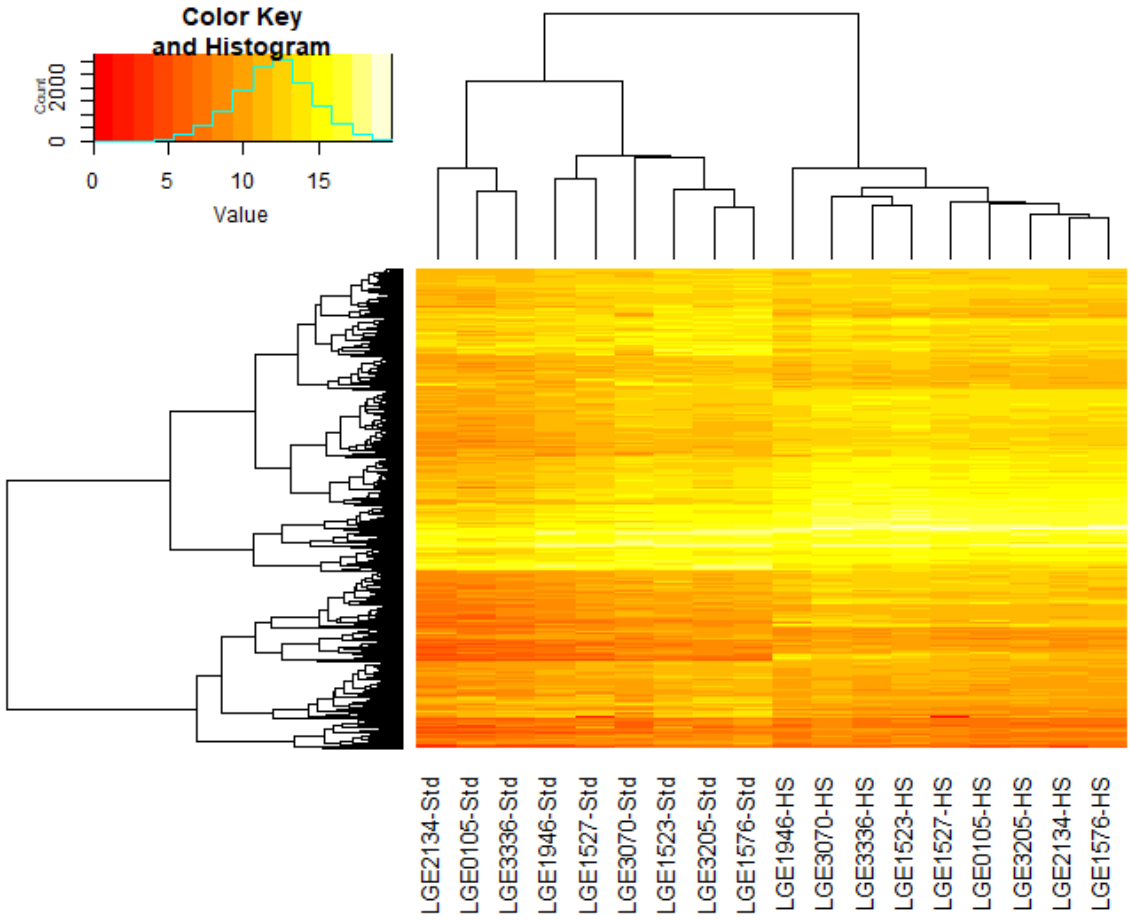

Figure 9. Clustered heatmap of conserved genes in the nine selected *E. coli* isolates. Map based on transformed and rescaled CPM data. Std=Standard condition, HS= High Salt condition.

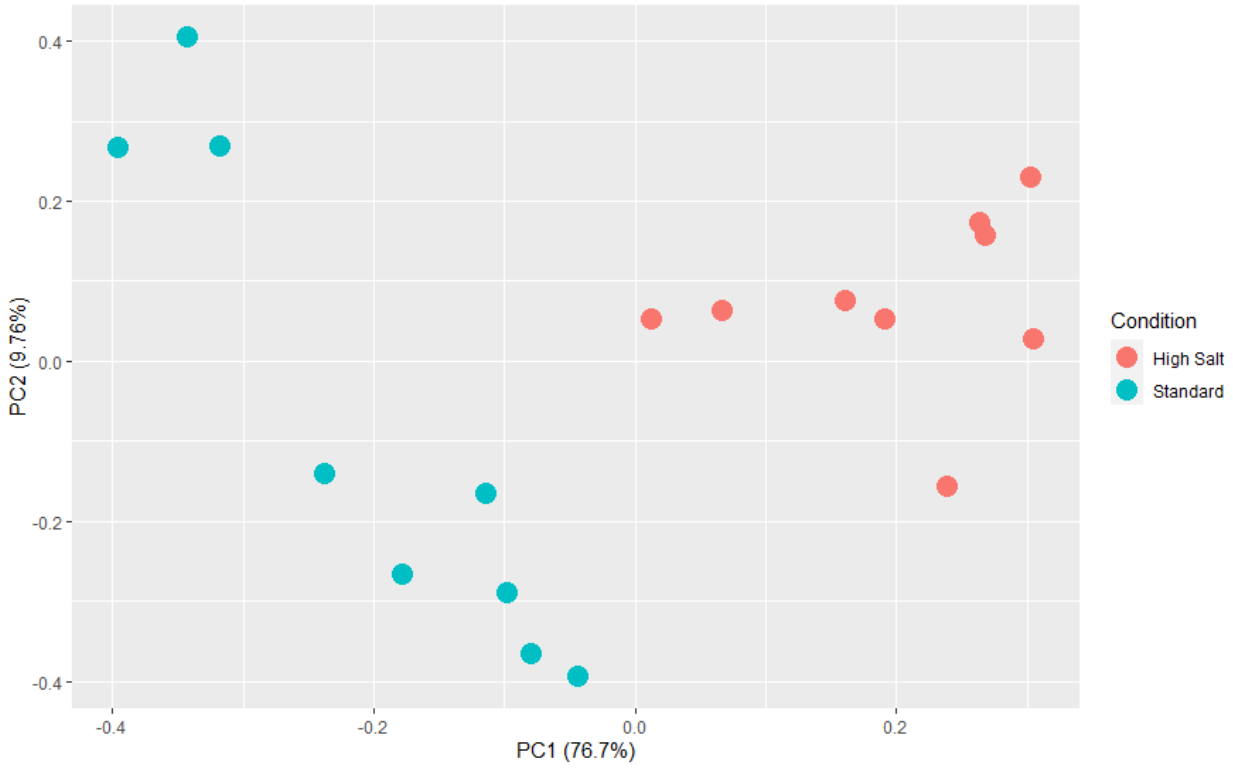

Figure 10. PCA plot of *E. coli* isolates in high salt and standard environments. Plot based on transformed and rescaled CPM data.

Delving into the significant upregulated and downregulated genes, we observed some clear patterns. In the upregulated genes, most of the genes were associated with osmolyte synthesis, exopolysaccharide synthesis, and membrane transporters (Table 4). Under high osmotic conditions, osmolytes are synthesized to assist the structural integrity of the cell wall (73). The genes of two osmolytes, glycine betaine and trehalose, were seen to be highly expressed. The genes *proV*, *proW*, and *proX* (logFC= 6.841, 8.093, and 8.047) are all part of the ProU operon that is known to be upregulated under osmotic conditions (76, 111). This operon has a high affinity for exogenous glycine betaine transport (73). The genes *betA* and *betB* (logFC=2.702 and 2.174, respectively) also aid in glycine betaine synthesis, but are known for synthesis from exogenous choline. The high levels of *otsA* and *otsB* (logFC= 3.405 and 3.880) indicate the synthesis of trehalose, an osmolyte, to aid the cell under the osmotic conditions.

In the highly downregulated genes, most of the genes were associated with carbohydrate transport, glutamate/glutamine transport, phage shock stress, and carbon starvation (Table 5). The most downregulated genes *hycE* and *hycF* play important roles in mixed-fermentation respiration (112). Specifically, these genes and others in its operon are important in the catabolism of formate to carbon dioxide and hydrogen. Therefore it is a possibility that the low expression of these formate dehydrogenase genes suggests an increase of formate within the cell that may function in some osmotic response capacity. Genes involved in maltose transport were also significantly downregulated. While maltose can be used as an osmolyte, it is not as efficient as others as it is metabolized into secondary metabolites and is not stored within the cell like other osmolytes (73, 113). Therefore, the high upregulation of glycine betaine and trehalose genes may provide more effective and long-term protection, so upregulation of maltose genes would be redundant and costly.

Table 4. Top 100 upregulated genes of selected *E. coli* under the high salt condition. LogFC= log_2-fold change.

Gene/ Operon	**Product**	**logFC**
	Amino Acid Biosynthesis & Transport	
tauABCD	taurine ABC transporter	4.617, 4.145, 3.916, 2.583
leuABC	leucine biosynthesis	2.505, 2.305, 2.063
YgaHZ	L-valine efflux transporter	2.436, 2.358
SdaB	L-serine deaminase II	2.035
	Cofactors	
bioBCD	Biotin biosynthesis	2.062, 2.561, 2.225
GrxB	Glutaredoxin 2	2.459
nrdIEF	Ribonucleotide reductase	2.317, 2.062, 2.166
	Carbohydrate Metabolism	
SgrT	SgrT, regulator of PtsG activity	3.076
YidA	sugar phosphatase	3.027
talA-tktB	pentose phosphate pathway	2.340, 2.556
dkgA_2	methylglyoxal reductase [multifunctional]	2.300

Table 4. Top 100 upregulated genes of selected *E. coli* under the high salt condition (continued).

Gene/ Operon	Product	logFC
	Exopolysaccharide Biosynthesis	
wzabc-wcaAB	Colanic acid biosynthesis	3.434, 5.279, 4.462, 4.524, 4.424
wcaCEF-fcl-wcaI-cpsBG-wcaJ-wzxC	Colanic acid biosynthesis	4.631, 4.321, 4.922, 5.392, 5.038, 4.498, 5.019, 4.883, 4.374
yjbEFH	putative protein	4.974, 5.034, 3.735
wcaKLM	colanic acid biosynthesis	4.455, 4.018, 2.836
rutG	putative xanthine/uracil transporter	3.521
ycfT	putative transport protein	2.262
yebA	putative peptidase	2.234
yqjF	putative quinol oxidase subunit	2.053
lpxB	lipid A disaccharide synthase	2.011
	Membrane Transport	
kdpABC	K transporting ATPase	4.777, 4.06, 2.497
aaeAB	hydroxylated, aromatic carboxylic acid efflux transporter	3.823, 3.369
ygaY	putative transporter	3.698
phnCDFGI	phosphonate ABC transporter	3.002, 2.498, 2.094, 2.285, 1.968
ddpF	YddO	2.200
torY	trimethylamine N-oxide reductase III, c-type cytochrome subunit	1.955
	Osmolyte Synthesis & Transport	
proVWX	glycine betaine / proline ABC transporter	6.841, 8.093, 8.047
otsAB	trehalose synthesis	3.405, 3.880
katE	heme d synthase / hydroperoxidase	3.317
yehYXW	glycine betaine ABC transporter	3.012, 2.991, 2.304
betAB	glycine betaine biosynthesis	2.702, 2.174
osmY	hyperosmotically inducible periplasmic protein	2.419
	Stress Responses	
ssuEDCB	NAD(P)H-dependent FMN reduction (ssuED) & aliphatic sulfonate ABC transporter (ssuCB)	4.369, 3.710, 3.985, 3.033
pphA	protein phosphatase 1 modulates phosphoproteins, signals protein misfolding	2.423
Bdm	biofilm-dependent modulation protein	2.175
soda	superoxide dismutase (Mn)	2.070
yhcN	stress-induced protein	2.057
	Transcription	
sgrR	SgrR DNA-binding transcriptional dual regulator	3.188

Table 4. Top 100 upregulated genes of selected *E. coli* under the high salt condition (continued).

Gene/ Operon	Product	logFC
	Transcription (continued)	
StpA	H-NS-like DNA-binding protein with RNA chaperone activity	2.922
NemR	NemR DNA-binding transcriptional repressor	2.564
RcsA	positive DNA-binding transcriptional regulator of capsular polysaccharide synthesis	2.068
	Hypothetical	
group_4320	hypothetical protein	5.086
group_11845	hypothetical protein	4.189
YghA	putative glutathionylspermidine synthase	3.487
group_4850	hypothetical protein	3.227
group_27307	hypothetical protein	2.956
AaeX	hypothetical protein	2.756
group_9571	hypothetical protein	2.698
group_9724	hypothetical protein	2.610
group_9337	hypothetical protein	2.396
group_2384	hypothetical protein	2.228
YdiH	putative protein	2.124
YgdI	putative lipoprotein	2.065
YhaK	bicupin-related protein	2.038
EcpBC	hypothetical proteins	1.955, 2.019
YaeF	putative lipoprotein	1.942

Table 5. Top 100 downregulated genes of selected *E. coli* under the hyperosmotic condition. LogFC= log_2-fold change.

Gene/ Operon	Product	logFC
	Amino Acid Biosynthesis & Transport	
yhdWXYZ	YhdW/YhdX/YhdY/YhdZ ABC transporter	-4.872, -3.489, -3.659, -4.104
sstT	serine / threonine:Na symporter	-4.106
tdcB	catabolic threonine dehydratase	-4.033
puuA_1	glutamate-putrescine ligase	-4.027
yahN	putative neutral amino acid efflux system	-3.999
yhaO	YhaO STP transporter	-3.845
alaE	L-alanine exporter	-3.500
puuD	gamma-glutamyl-gamma-aminobutyrate hydrolase	-3.368
gdhA	glutamate dehydrogenase	-3.332
artJ	arginine ABC transporter - periplasmic binding protein	-3.301
argT	lysine / arginine / ornithine ABC transporter - periplasmic binding protein	-3.242
	Carbohydrate Metabolism	
rbsD	ribose pyranase	-4.515
garL	alpha-dehydro-beta-deoxy-D-glucarate aldolase	-4.452
fucAO	L-fuculose-phosphate aldolases	-3.459, -3.883
garD	galactarate dehydratase	-3.666
lsrK	autoinducer-2 kinase	-3.326
	Cell Wall	
fimFH	fimbrial morphology	-3.170, -3.123
murQ	N-acetylmuramic acid 6-phosphate etherase	-3.865
csgF	curli assembly component	-3.085
	Cellular Respiration: Anaerobic	
nrfABCD	formate dependent nitrite reductase	-3.646, -3.492, -3.583, -3.676
nirD	nitrite reductase, small subunit	-3.545
hyaC	hydrogenase 1, b-type cytochrome subunit	-3.306
fdhF_2	formate dehydrogenase H	-3.157
	Cellular Respiration: Fermentation	
hycABCDEFGHI	formate hydrogenlyase complex	-3.892, -4.456, -4.855, -5.229, -5.834, -6.423, -4.872, -3.901, -3.609
	Lipid Metabolism	
fadB	dodecenoyl-CoA delta-isomerase, enoyl-CoA hydratase, 3-hydroxybutyryl-CoA epimerase, 3-hydroxyacyl-CoA dehydrogenase	-4.268
fade	acyl-CoA dehydrogenase	-4.045

Table 5. Top 100 downregulated genes of selected *E. coli* under the hyperosmotic condition (continued).

Gene/Operon	Product	logFC
	Membrane Transport	
mglBAC	galactose ABC transporter	-5.830, -5.031, -4.017
malK-lamB-malM	maltose ABC transporter (malKM) and porin (lamB)	-5.536, -5.749, -5.371
UhpT	hexose-6-phosphate:phosphate antiporter	-4.924
DctA	C4 dicarboxylate / orotate:H symporter	-4.815
malEFG	maltose ABC transporter	-4.690, -4.070, -3.531
CstA	peptide transporter induced by carbon starvation	-4.689
ytfQRT	galactofuranose / galactopyranose ABC transporter	-4.649, -3.805, -3.190
GlpTQ	GlpR DNA-binding transcriptional repressor & glycerophophoryl diester phosphodiesterase	-4.432, -4.505
YjiY	putative inner membrane protein	-4.350
LsrA	AI-2 ABC transporter - ATP binding subunit	-4.129
CusF	copper / silver efflux transport system - periplasmic binding protein	-4.094
rbsA_2	ribose ABC transporter - putative ATP binding subunit	-4.062
GlpFK	glycerol channel and kinase	-4.018, -3.298
SrlA	glucitol/sorbitol PTS permease - SrlA subunit	-3.964
NupG	nucleoside:H symporter NupG	-3.928
GlpC	glycerol-3-phosphate dehydrogenase (anaerobic), small subunit	-3.857
PreA	NADH-dependent dihydropyrimidine dehydrogenase subunit	-3.557
UraA	uracil:H symporter	-3.276
NagE	N-acetylglucosamine PTS permease	-3.204
DcuB	dicarboxylate transporter DcuB	-3.082
	Stress Responses	
PspG	phage shock protein G	-5.486
pspABCDE	phage shock protein	-4.959, -4.806, -4.616, -4.580, -4.108
YchH	stress-induced protein	-3.686
Slp	starvation lipoprotein	-3.342
	Transcription & Translation	
TdcA	TdcA DNA-binding transcriptional activator	-4.719
YgeV	putative transcriptional regulator	-4.673
RaiA	stationary phase translation inhibitor and ribosome stability factor	-4.155
CaiF	CaiF-L-carnitine DNA-binding Transcriptional Activator	-3.746
YjjM	putative DNA-binding transcriptional regulator	-3.704

Table 5. Top 100 downregulated genes of selected *E. coli* under the hyperosmotic condition (continued).

Gene/Operon	Product	logFC
	Transcription & Translation (continued)	
galS	GalS DNA-binding transcriptional dual regulator	-3.702
lsrR_2	LsrR DNA-binding transcriptional repressor	-3.598
melR	MelR DNA-binding transcriptional dual regulator	-3.533
yidF	putative DNA-binding transcriptional regulator	-3.430
mhpR	MhpR transcriptional activator	-3.267
glnK	Nitrogen regulatory protein P-II 2	-3.134
	Hypothetical	
group_4297	hypothetical protein	-5.343
group_28086	hypothetical protein	-4.513
group_4748	hypothetical protein	-4.363
ydcH	putative protein	-3.881
ygjR	putative NAD(P)-binding dehydrogenase	-3.584
Yahoo	putative protein	-3.496
yjfN	putative protein	-3.356

2.3.4. Characterization of Significantly Enriched Functional Groups

Functional groups of genes have large enrichment scores when the probability that the subsystem possesses upregulated or downregulated genes that are significantly differentially expressed. There were five functional subsystems that had significant enrichment scores: “Amino Acid Synthesis” (3.92), “Metabolic Regulation Synthesis” (2.31), “Sigma Regulons” (6.74), “Transcription Factor Regulons” (3.24), and “Osmotic Stress” (2.66). While each group is functionally important and can influence each other, as some genes can be categorized in multiple subsystems, we focused specifically on the “Sigma Regulons” and “Osmotic Stress” subsystems. The sigma regulons subsystem had the highest enrichment score (6.74). Among the detailed sigma factors, only the RpoE and RpoN regulons showed significance in their enrichment scores (Figure 11), 5.31 and 2.37, respectively. Other sigma regulons possessed genes that were highly expressed, but the consensus of expression was neither significantly upregulated nor downregulated.

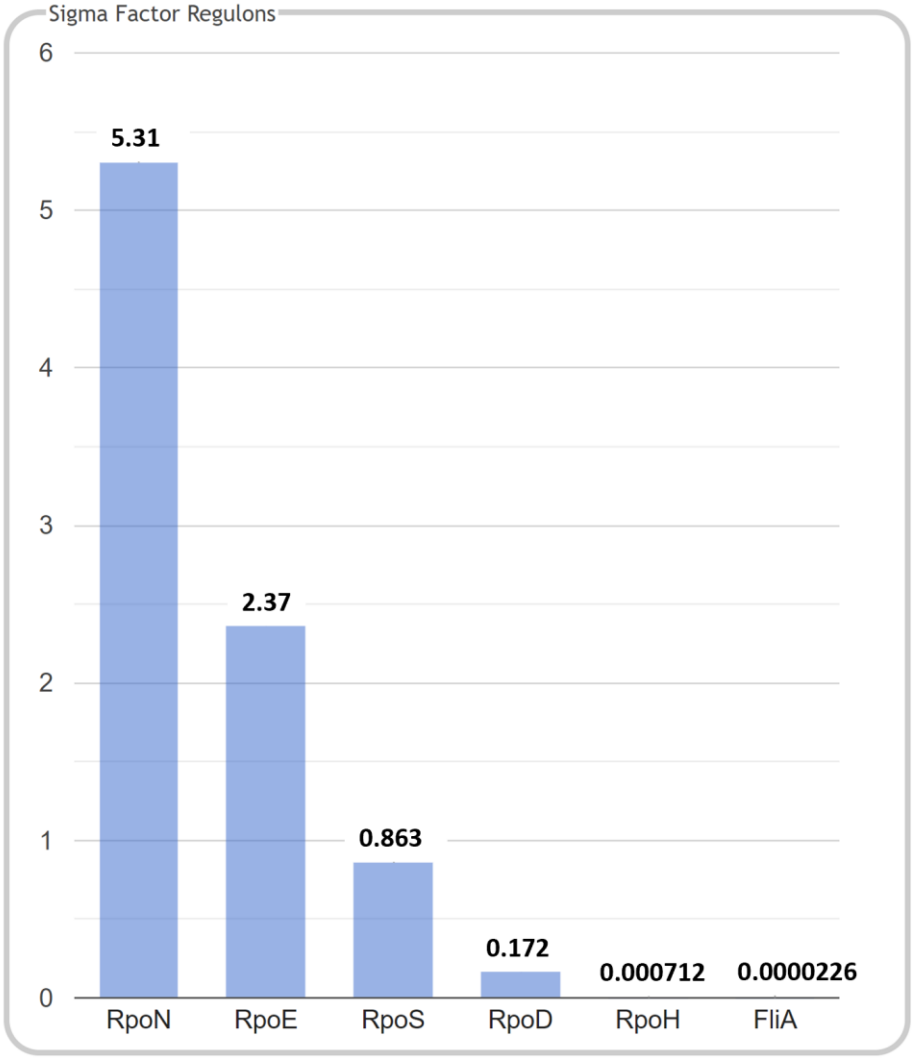

Figure 11. Enrichment analysis of sigma factor regulons. Enrichment values are labeled on the y-axis. Genes were sorted based on BioCyc database categorization. Image formatted by Ecocyc Omics Dashboard tool.

The RpoN regulon was the most significantly enriched of the sigma factor regulons (enrichment value= 5.31). RpoN is a sigma factor that mostly controls nitrogen-associated metabolism and amino acid transport (114). Overall, many of the genes in this regulon were downregulated (Table 6). Most genes were associated with the catabolism and transport of the amino acids glutamate, glutamine, and arginine. Some stress-related genes such as the *pspA* and *ibpB*, were also shown to be downregulated. The few genes that were upregulated included

regulation genes (RpoN, RpoE, and rseBC). The downregulation of the regulon, but the upregulation of *rpoN* transcription brings into question whether RpoN is being translated. RpoN is the only sigma factor that requires an additional transcriptional factor for activation, the transcriptional factor NtrC (115). The expression of the NtrC regulon in the subsystem of "Transcription Factor Regulons" was significant with an enrichment score of 5.99. Further investigation showed that many of the NtrC-associated genes were downregulated. The downregulation of both RpoN- and NtrC-regulated expression suggested that RpoN was being post-transcriptionally inhibited or transcribed passively by another upregulated transcription factor.

Table 6. Significant differentially expressed genes in the RpoN regulon. Significant genes were determined on a logFC value of ≥ 1 or ≤-1. Biocyc database used for gene product annotation. logFC= log_2-fold change.

Gene/ Operon	Product	logFC
	Metabolism	
hycBCDEFGHI	formate hydrogenlyase complex	-4.457, -4.855, -5.229, -5.835, -6.423, - 4.873, -3.902, -3.609
GlpQ	glycerophosphoryl diester phosphodiesterase GlpQ	-4.506
xdhABC	putative xanthine dehydrogenase	-1.874, -2.225, -2.436
Pup	putrescine:H+ symporter PuuP	-2.291
astCADBE	L-arginine degradation	-2.000, -1.489, -1.407, -1.241, -1.183
rutABCD	pyrimidine monooxygenase (rutA), peroxyureidoacrylate/ureidoacrylate amidohydrolase (rutB), aminoacrylate peracid reductase (rutC), aminoacrylate hydrolase (rutD)	-1.768, -1.383, -1.196, -1.229
NorVW	anaerobic nitric oxide reductase	-1.567, -1.083
YgfK	putative oxidoreductase, Fe-S subunit	-1.009
DcuD	putative transporter DcuD	1.552
HyfA	hydrogenase 4 component A	1.840
	Protein Maturation	
HypA	hydrogenase 3 nickel incorporation protein HypA	-1.399
HypF	carbamoyl--[HypE] ligase	-1.255
	Regulation	
hycA	regulator of the transcriptional regulator FhlA	-3.892
flhDC	DNA-binding transcriptional dual regulators	-1.428, -1.904

Table 6. Significant differentially expressed genes in the RpoN regulon (continued).

Gene/Operon	Product	logFC
rpoN	RNA polymerase, sigma 54 (sigma N factor)	1.083
rpoE	RNA polymerase sigma E factor	1.224
rseBC	anti-sigma factor (RseB) & protein (RseC)	1.244, 1.54
	Stress Responses	
pspG	phage shock protein G	-5.486
pspABCDE	phage shock proteins	-4.959, -4.807, -4.617, -4.581, -4.109
ddpX	D-alanyl-D-alanine dipeptidase	-1.653
ibpB	small heat shock protein IbpB	-1.525
yeaG	protein kinase YeaG	-1.358
	Transport	
yhdWXYZ	YhdW/YhdX/YhdY/YhdZ ABC transporter	-4.873, -3.489, -3.659, -4.105
argT	lysine/arginine/ornithine ABC transporter periplasmic binding protein	-3.242
glnK-amtB	ammonium transporter	-3.134, -2.963
hydN	putative electron transport protein HydN	-2.847
gltIJKL	glutamate/aspartate ABC transporter	-2.527, -1.094, -1.287, -1.726
potFGHI	putrescine ABC transporter	-2.469, -1.707, -1.818, -1.626
hisJQMP	lysine/ arginine/ ornithine/ histidine ABC transporter	-2.268, -1.471, -1.241, -1.406
nikR	DNA-binding transcriptional repressor NikR	-1.565
glnH	L-glutamine ABC transporter periplasmic binding protein	-1.241
ddpF	putative D,D-dipeptide ABC transporter ATP-binding subunit DdpF	2.200
rutG	pyrimidine:H+ symporter	3.521

The other significantly enriched sigma factor regulon was RpoE. This regulator is responsible for many genes associated with membrane stress and maintenance. Some specific stresses include cold shock and high osmolarity (58). Most of the genes in the RpoE regulon were upregulated and functioned in membrane synthesis, biofilm formation, DNA repair, oxidative stress, and regulation (Table 7). The *rpoE* gene itself was upregulated possibly due to its own positive feedback loop or by the regulation of other sigma factors (58). A suppressor of RpoE, *ptsN*, was significantly expressed (logFC=1.018). The upregulation of this gene may indicate that RpoE was overexpressed (88).

Table 7. Significant differentially expressed genes of the RpoE regulon. Significant genes were determined on a logFC value of ≥ 1 or ≤-1. Biocyc database used for gene product annotation. logFC= log_2-fold-change.

Gene/ Operon	Product	logFC
	Biofilm	
yggN	DUF2884 domain-containing protein YggN	1.155
	DNA/RNA Repair	
rnhB	RNase HII	1.764
recJ	ssDNA-specific exonuclease RecJ	1.291
recR	DNA repair protein RecR	1.194
Cca	fused CCA tRNA nucleotidyltransferase/phosphohydrolase	1.087
dnaE	DNA polymerase III subunit α	1.002
	Membrane Synthesis	
Wzabc	colanic acid biosynthesis	3.434, 5.279, 4.463
lpxB	lipid A disaccharide synthase	2.011
eptB	Kdo2-lipid A phosphoethanolamine 7"-transferase	1.658
waaFC	ADP-heptose--LPS heptosyltransferases	1.486, 1.539
yfeY	DUF1131 domain-containing lipoprotein YfeY	1.362
yfeK	DUF5329 domain-containing protein YfeK	1.242
ygiM	putative signal transduction protein (SH3 domain)	1.139
lpxP	palmitoleoyl acyltransferase	1.112
ydhI	DUF1656 domain-containing protein YdhI	1.057
bepA	β-barrel assembly-enhancing protease	1.053
	Metabolism	
pdxA	4-hydroxythreonine-4-phosphate dehydrogenase	1.397
malQ	4-α-glucanotransferase	-1.041
sgbE	L-ribulose-5-phosphate 4-epimerase SgbE	-1.224
yiaK	2,3-diketo-L-gulonate reductase	-1.59
lgoR	putative DNA-binding transcriptional regulator LgoR	-3.704
	Regulation	
rseBC	anti-sigma factor (RseB) & protein (RseC)	1.244, 1.54
rpoE	RNA polymerase sigma E factor	1.224
rpoN	RNA polymerase, sigma 54 (sigma N) factor	1.083
ptsN	PtsN-phosphorylated	1.018
rapZ	RNase adaptor protein RapZ	1.001
	Stress Responses	
degP	periplasmic serine endoprotease DegP	1.257
uspD	universal stress protein D	-1.691
	Transport	
sbmA	peptide antibiotic/peptide nucleic acid transporter	1.646
	Unknown	
yoaC	DUF1889 domain-containing protein YoaC	-2.54

The conserved osmotic responses in all of the selected *E. coli* isolates show responses such as osmolyte synthesis and extracellular protection. The upregulation of osmolyte synthesis genes suggest the increase in the solutes to increase turgor pressure and stabilize proteins. The increased expression of extracellular synthesis genes such as the *wca* operon indicate external protection from the osmotic stress. The downregulation of the *hyc* operon suggest the accumulation of formate being a benefit to the osmotic stress response. Also, sigma factors RpoN and RpoE showed significant enrichment in their regulons.

2.4. Discussion

2.4.1. Phenotypic Plasticity is Observed in Soil *E. coli* Growth Rate under a Hyperosmotic Condition

Our results showed plasticity in the growth phenotypes of *E. coli* isolates between a standard condition (0.3% NaCl GDMM) and a hyperosmotic condition (4% NaCl GDMM). An increased osmotic concentration in the environment has been documented as being a negative influence on *E. coli* growth (116–119). Cells maintain turgor pressure to maintain cell growth, elongation, and division (75). The increase of salts outside of the cell causes the initiation of osmoregulation responses to prevent plasmolysis. First, charged solutes such as potassium ions and glutamate are transported to counter initial loss of turgor pressure. Then compatible solutes, such as glycine betaine and proline, are synthesized or imported to sustain the cell against longer or larger magnitudes of hyperosmotic stress (73). Cell growth is negatively affected due to energy costs as well as the adjustment to cellular metabolism.

Prior data has depicted the reductive effect that osmotic stress has on growth rate. Three enterohemorrhagic *E. coli* (EHEC) strains' growth rates decreased from about 0.28 h^{-1} to 0.16 h^{-1} between a non-supplemented to a 4.5% NaCl supplemented media, respectively (117). While

there was a decrease in all three, only the growth rates of STEC2 were significantly different. The small difference in growth rates may be due to the use of rich media (Brain Heart Infusion [BHI] broth) which contains more salts and minerals than defined media, thus providing some additional osmotic protection, as seen with the growth of *E. coli* MG1655 in minimal and LB media (120). In contrast, our study showed all the strains' growth rates were distinctly significant between the two environments with rates in the range of 0.6-0.8 $\log_{10}(OD_{600})$ h^{-1} (0.3% NaCl) reduced to 0.069 – 0.323 $\log_{10}(OD_{600})$ h^{-1} (4% NaCl). In a study by Zhu and Dai, *E. coli* K12 NCM3722 showed a significant decrease in growth rate from 1.0 h^{-1} to 0.5 h^{-1} in 0.1 M and 0.4 M NaCl defined media, respectively (119). This difference between osmotic conditions coincides with our data, but the high salt condition in the study had a concentration of about 2% NaCl which is lower than our high salt condition of 4% NaCl. Overall, our data between environments agrees with the previous studies in response to increased osmotic conditions. All isolates exhibited decreased growth rates in the 4% NaCl condition than in the 0.3% NaCl condition, which was expected.

The growth rates vary not only between environments, but also amongst the strains suggesting the influence of genotype on the phenotype. Past studies have observed ranges in phenotypes between different strains of *E. coli* under a specific condition (62, 72, 89, 121). In a study by Somorin *et al.*, six *E. coli* isolates from soil ranged from 3 hours of *Acanthamoeba polyphaga* predation survival to more than 30 hours (62). In a single patient, eight strains of *E. coli* had a wide range of hydrogen peroxide tolerance as 80% of Strain 52 cells survived after 5 minutes whereas <20% of Strain 54 cells survived (121). Our data coincided with the previous research as growth rates within our isolates varied greatly under the hyperosmotic condition. The variation of growth rates suggests that the effect of genotypes played a role in the resulting

phenotypes. This is further supported by our qPCR work where we found that the expression profiles differed amongst our eight selected isolates in the same condition in the same stationary growth phase.

Phenotypic plasticity describes the interaction of the environment with the genotype on a phenotype, and this was seen in our data at as the growth rate rankings of the isolates were not all parallel between the two conditions. While we did confirm plasticity on a phenotypic level, we could not verify it on a transcriptomic level due to the one replicate in our RNAseq study. We were able to distinguish the effect of the environment on the transcriptomes, but not the effect of the isolates and ultimately the interaction between the two. We did, however, find conserved elements of osmotic stress responses amongst all our isolates such as the expression of sigma factors RpoN and RpoE.

2.4.2. The Downregulation of RpoN Regulon Expression and the Upregulation of RpoE Regulon Expression are Important for *E. coli* under a Hyperosmotic Condition

Sigma factors are important in the regulation of genes especially during stress conditions. In our work we found RpoN and RpoE regulons were significantly downregulated and upregulated, respectively. RpoN is an alternative sigma factor responsible for nitrogen metabolism and is most active during nitrogen starvation (114). Prior studies found that the downregulation of RpoN-regulated genes during antibiotic stress in *Salmonella enterica* (122), where the Δ*rpoN* mutants had increased resistance to the antibiotic polymyxin B. In *E. coli* O157:H7, lack of RpoN resulted in increased protein expression in acid resistance and virulence factors (123). All of our selected isolates had a significantly downregulated RpoN regulon suggesting that the minimal effects of RpoN had a positive impact on cell growth rate. Therefore,

it seems that the downregulation or absence of RpoN-regulated genes provides advantages to cells under stressful conditions.

RpoE is one of many regulators of membrane stress in *E. coli*. Response systems such as the Psp, Cpx, Bae, and Rcs systems are among the stress pathways that combat physical and chemical membrane stresses. Each of these systems have specific operons that they regulate, but there are not functionally isolated from each other as some genes are regulated by more than one system (124). While crosstalk can be observed, in a study against four other membrane stress systems, the RpoE regulon was seen to be the most upregulated and significant membrane system as it had the highest count of dependent genes (69 total) (88, 125). Klein *et al.* also found that *rpoE* transcription was increased by 2-3 fold increase induction of the *rpoeP4* promoter under high osmolarity of 0.25 M sucrose (58). Our data showed that the RpoE regulon was significantly upregulated, suggesting that it played a key role in the response to the osmotic stress. There was an upregulation of genes associated with other membrane stress regulons specifically the *Rcs* genes, so the other membrane systems may be in effect as well.

2.4.3. RpoS May Directly and Indirectly Affect Growth Rate Variation under a Hyperosmotic Condition

RpoS is a strong transcriptional regulator and its regulon is expansive in size and in functions. The regulon of RpoS in our data was insignificantly enriched, but RpoS is known to have dual regulator properties (126). Genes such as *proVWX* and *malBC* were significantly expressed, of which all are regulated by RpoS (69, 86). While the expression of RpoS-regulated genes was promising for the presence of RpoS regulation, confirmation of RpoS significance requires *rpos* knockouts. Therefore, we can only suggest that RpoS is present and may be at work in gene regulation based on the expression of its regulon.

The two significantly enriched sigma factors that had significant regulon expression, RpoE and RpoN, both have regulation relationships with RpoS. In the case of RpoN, it was found that RpoN and RpoS had an antagonistic relationship as genes that were upregulated by RpoN were downregulated by RpoS (56, 127). The gene expression of *rpoE* is regulated by other sigma factors including RpoN and RpoS. In this study, RpoS was shown to aid transcription of *rpoE* under osmotic conditions (58). In *Salmonella enterica* Typhi, 38 genes were shown to be under expressed in a *ΔrpoEΔrpoS* mutant that was not affected by either of the single sigma factor mutants (128). Overall, the interplay between sigma factors and the expression of known regulated genes may suggest a connection between RpoS and the observed phenotypic plasticity.

2.4.4. Conserved Osmotic Stress Response Elements are Observed Across All *E. coli* Isolates

Escherichia coli is classified into seven phylogroups (A, B1, B2, C, D, E, and F) that are each associated with certain characteristics (129). The isolates used in our study were all from phylogroup D which is known to be the most genetically diverse phylogroup. In a study by Touchon *et al.*, phylogroup D *E. coli* isolates had the highest genetic diversity with nucleotide diversity of >0.005 in contrast to the other phylogroups (<0.005) (130). The genetic diversity of phylogroup D may be a reason behind its variety of phenotypes. Within 35 *E. coli* isolates that stemmed from three phylogroup D strains (ST405, ST69, & ST393), there was a range of adherence phenotypes with 26 isolates classified as weak, 6 isolates classified as moderate, and 3 isolates classified as strong (131). Since its genetic diversity is vast, finding conserved elements in all the isolates was valuable.

General osmotic responses were all highly expressed in the isolates. The genes responsible for osmolyte synthesis and transport were highly expressed. The *proVWX* genes

belong to the ProU operon, which is responsible for osmolyte transport specifically glycine betaine and proline. The genes *otsAB* are responsible for the synthesis of trehalose (75). The upregulation of these genes have been observed in *E. coli* K12 when subjected to long term osmotic stress (86). The *wca* operon, responsible for extrapolysaccharide synthesis of colanic acid, was also upregulated. In a study by Chen *et al.*, they also saw the impact of colanic acid production to increase survival of *E. coli* under 1.5 M and 2.5 M NaCl stress (132). The *hyc* operon was highly downregulated which is important for formate degradation. Therefore, an increase in formate could be a possible osmotic response. This formate accumulation was seen to increase the intake of proline, an osmolyte, in potato plant cells which were subjected to drought stress (133).

2.5. Conclusion

Phenotypic plasticity in a growth rate phenotype was observed among soil *E.coli* isolates under hyperosmotic stress at a phenotypic level. The growth rate rankings of the isolates portrayed the interaction of the hyperosmotic environment and the genotypes on the overall growth rate phenotype. Conserved genes in all isolates showed upregulation of multiple osmotic stress genes and downregulation of nitrogen-associated genes. The sigma factors RpoE, and indirectly RpoS, may be important factors of observed phenotypes under osmotic stress. Our study suggests a role of phenotypic plasticity in *E. coli* adaptation to stress environments. Future work in other *E. coli* isolates and other environmental conditions can expand the significance of phenotypic plasticity in bacterial adaptation.

REFERENCES

1. Dykhuizen DE. 1990. Experimental Studies of Natural Selection in Bacteria. Annu Rev Ecol Syst 21:373–398.
2. Chu H, Gao G-F, Ma Y, Fan K, Delgado-Baquerizo M. 2020. Soil Microbial Biogeography in a Changing World: Recent Advances and Future Perspectives. mSystems 5.
3. Sévin DC, Stählin JN, Pollak GR, Kuehne A, Sauer U. 2016. Global Metabolic Responses to Salt Stress in Fifteen Species. PLOS ONE 11:e0148888.
4. Poursat BAJ, Spanning RJM van, Voogt P de, Parsons JR. 2019. Implications of microbial adaptation for the assessment of environmental persistence of chemicals. Critical Reviews in Environmental Science and Technology 49:2220–2255.
5. Classen AT, Sundqvist MK, Henning JA, Newman GS, Moore JAM, Cregger MA, Moorhead LC, Patterson CM. 2015. Direct and indirect effects of climate change on soil microbial and soil microbial-plant interactions: What lies ahead? Ecosphere 6:art130.
6. McDonald MJ. 2019. Microbial Experimental Evolution – a proving ground for evolutionary theory and a tool for discovery. EMBO Rep 20.
7. Gibson B, Wilson DJ, Feil E, Eyre-Walker A. 2018. The distribution of bacterial doubling times in the wild. Proc Biol Sci 285.
8. Aujoulat F, Roger F, Bourdier A, Lotthé A, Lamy B, Marchandin H, Jumas-Bilak E. 2012. From Environment to Man: Genome Evolution and Adaptation of Human Opportunistic Bacterial Pathogens. Genes (Basel) 3:191–232.
9. Holden MTG, Seth-Smith HMB, Crossman LC, Sebaihia M, Bentley SD, Cerdeño-Tárraga AM, Thomson NR, Bason N, Quail MA, Sharp S, Cherevach I, Churcher C,

Goodhead I, Hauser H, Holroyd N, Mungall K, Scott P, Walker D, White B, Rose H, Iversen P, Mil-Homens D, Rocha EPC, Fialho AM, Baldwin A, Dowson C, Barrell BG, Govan JR, Vandamme P, Hart CA, Mahenthiralingam E, Parkhill J. 2009. The Genome of Burkholderia cenocepacia J2315, an Epidemic Pathogen of Cystic Fibrosis Patients. Journal of Bacteriology 191:261–277.

10. Bradshaw AD. 1965. Evolutionary Significance of Phenotypic Plasticity in Plants, p. 115–155. *In* Caspari, EW, Thoday, JM (eds.), Advances in Genetics. Academic Press.
11. Kleunen MV, Fischer M. 2005. Constraints on the evolution of adaptive phenotypic plasticity in plants. New Phytologist 166:49–60.
12. Andrew Hendry. 2017. Plasticity, p. 276–303. *In* Eco-Evolutionary Dynamcis. Princeton University Press, Princeton, New Jersey.
13. Bentkowski P, van Oosterhout C, Ashby B, Mock T. 2017. The effect of extrinsic mortality on genome size evolution in prokaryotes. ISME J 11:1011–1018.
14. Bentkowski P, Van Oosterhout C, Mock T. 2015. A Model of Genome Size Evolution for Prokaryotes in Stable and Fluctuating Environments. Genome Biol Evol 7:2344–2351.
15. Cobo-Simón M, Tamames J. 2017. Relating genomic characteristics to environmental preferences and ubiquity in different microbial taxa. BMC Genomics 18:499.
16. Wiedenbeck J, Cohan FM. 2011. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. FEMS Microbiology Reviews 35:957–976.
17. Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. mBio 3.

18. Hottes AK, Freddolino PL, Khare A, Donnell ZN, Liu JC, Tavazoie S. 2013. Bacterial Adaptation through Loss of Function. PLoS Genet 9.

19. Finkel SE. 2006. Long-term survival during stationary phase: evolution and the GASP phenotype. Nat Rev Microbiol 4:113–120.

20. Ryall B, Eydallin G, Ferenci T. 2012. Culture History and Population Heterogeneity as Determinants of Bacterial Adaptation: the Adaptomics of a Single Environmental Transition. Microbiol Mol Biol Rev 76:597–625.

21. Alvarez-Ordóñez A, Broussolle V, Colin P, Nguyen-The C, Prieto M. 2015. The adaptive response of bacterial food-borne pathogens in the environment, host and food: Implications for food safety. Int J Food Microbiol 213:99–109.

22. Maeusli M, Lee B, Miller S, Reyna Z, Lu P, Yan J, Ulhaq A, Skandalis N, Spellberg B, Luna B. 2020. Horizontal Gene Transfer of Antibiotic Resistance from Acinetobacter baylyi to Escherichia coli on Lettuce and Subsequent Antibiotic Resistance Transmission to the Gut Microbiome. mSphere 5.

23. Koskiniemi S, Sun S, Berg OG, Andersson DI. 2012. Selection-Driven Gene Loss in Bacteria. PLOS Genetics 8:e1002787.

24. Harms A, Maisonneuve E, Gerdes K. 2016. Mechanisms of bacterial persistence during stress and antibiotic exposure. Science 354.

25. Costa-Silva J, Domingues D, Lopes FM. 2017. RNA-Seq differential expression analysis: An extended review and a software tool. PLoS ONE 12:e0190152.

26. Sorek R, Cossart P. 2010. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. 1. Nature Reviews Genetics 11:9–16.

27. Derveaux S, Vandesompele J, Hellemans J. 2010. How to do successful gene expression analysis using real-time PCR. Methods 50:227–230.

28. Mantione KJ, Kream RM, Kuzelova H, Ptacek R, Raboch J, Samuel JM, Stefano GB. 2014. Comparing Bioinformatic Gene Expression Profiling Methods: Microarray and RNA-Seq. Med Sci Monit Basic Res 20:138–141.

29. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. 2008. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18:1509–1517.

30. Lambert G, Kussell E. 2014. Memory and Fitness Optimization of Bacteria under Fluctuating Environments. PLOS Genetics 10:e1004556.

31. Schlichting CD, Smith H. 2002. Phenotypic plasticity: linking molecular mechanisms with evolutionary outcomes. Evolutionary Ecology 16:189–211.

32. Otaki JM, Hiyama A, Iwata M, Kudo T. 2010. Phenotypic plasticity in the range-margin population of the lycaenid butterfly Zizeeria maha. BMC Evol Biol 10:252.

33. Thomas J. DeWitt, Samuel M. Scheiner. 2004. Phenotypic Plasticity: Functional and Conceptual Approaches. Oxford University Press, New York.

34. Pigliucci M. 2004. Phenotypic Plasticity and Integration, p. 155–175. *In* Pigliucci, M, Preston, K (eds.), Phenotpyic Integration: Studying the Ecology and Evolution of Complex Phenotypes. Oxford University Press, New York, NY.

35. Chevin L-M, Gallet R, Gomulkiewicz R, Holt RD, Fellous S. 2013. Phenotypic plasticity in evolutionary rescue experiments. Philos Trans R Soc Lond B Biol Sci 368.

36. Rong M, Zheng X, Ye M, Bai J, Xie X, Jin Y, He X. 2019. Phenotypic Plasticity of Staphylococcus aureus in Liquid Medium Containing Vancomycin. Front Microbiol 10.

37. Rodríguez RL. 2012. Grain of environment explains variation in the strength of genotype × environment interaction. Journal of Evolutionary Biology 25:1897–1901.

38. Turrientes M-C, González-Alba J-M, Campo R del, Baquero M-R, Cantón R, Baquero F, Galán JC. 2014. Recombination Blurs Phylogenetic Groups Routine Assignment in Escherichia coli: Setting the Record Straight. PLOS ONE 9:e105395.

39. Schellhorn HE. 2020. Function, Evolution, and Composition of the RpoS Regulon in Escherichia coli. Front Microbiol 11.

40. Brennan FP, Abram F, Chinalia FA, Richards KG, O'Flaherty V. 2010. Characterization of Environmentally Persistent Escherichia coli Isolates Leached from an Irish Soil. Appl Environ Microbiol 76:2175–2180.

41. Elsas JD van, Semenov AV, Costa R, Trevors JT. 2011. Survival of *Escherichia coli* in the environment: fundamental and public health aspects. The ISME Journal 5:173–183.

42. Walk ST, Alm EW, Calhoun LM, Mladonicky JM, Whittam TS. 2007. Genetic diversity and population structure of Escherichia coli isolated from freshwater beaches. Environmental Microbiology 9:2274–2288.

43. Ishii S, Ksoll WB, Hicks RE, Sadowsky MJ. 2006. Presence and Growth of Naturalized Escherichia coli in Temperate Soils from Lake Superior Watersheds. Appl Environ Microbiol 72:612–621.

44. Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, Crabtree J, Sebaihia M, Thomson NR, Chaudhuri R, Henderson IR, Sperandio V, Ravel J. 2008. The Pangenome Structure of Escherichia coli: Comparative Genomic Analysis of E. coli Commensal and Pathogenic Isolates. Journal of Bacteriology 190:6881–6893.

45. Lukjancenko O, Wassenaar TM, Ussery DW. 2010. Comparison of 61 Sequenced Escherichia coli Genomes. Microb Ecol 60:708–720.

46. Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, Calteau A, Chiapello H, Clermont O, Cruveiller S, Danchin A, Diard M, Dossat C, Karoui ME, Frapy E, Garry L, Ghigo JM, Gilles AM, Johnson J, Bouguénec CL, Lescat M, Mangenot S, Martinez-Jéhanne V, Matic I, Nassif X, Oztas S, Petit MA, Pichon C, Rouy Z, Ruf CS, Schneider D, Tourret J, Vacherie B, Vallenet D, Médigue C, Rocha EPC, Denamur E. 2009. Organised Genome Dynamics in the Escherichia coli Species Results in Highly Diverse Adaptive Paths. PLOS Genetics 5:e1000344.

47. Dobrindt U, Chowdary MG, Krumbholz G, Hacker J. 2010. Genome dynamics and its impact on evolution of Escherichia coli. Med Microbiol Immunol 199:145–154.

48. Mauri M, Klumpp S. 2014. A Model for Sigma Factor Competition in Bacterial Cells. PLoS Comput Biol 10.

49. Weber H, Polen T, Heuveling J, Wendisch VF, Hengge R. 2005. Genome-Wide Analysis of the General Stress Response Network in Escherichia coli: σS-Dependent Genes, Promoters, and Sigma Factor Selectivity. Journal of Bacteriology 187:1591–1603.

50. Christian E. W. Steinberg. 2012. Stress Ecology: Environmental Stress as Ecological Driving Force and Key Player in Evolution. Springer, New York.

51. Battesti A, Majdalani N, Gottesman S. 2011. The RpoS-Mediated General Stress Response in Escherichia coli. Annual Review of Microbiology 65:189–213.

52. Mika F, Hengge R. 2014. Small RNAs in the control of RpoS, CsgD, and biofilm architecture of Escherichia coli. RNA Biol 11:494–507.

53. Patten CL, Kirchhof MG, Schertzberg MR, Morton RA, Schellhorn HE. 2004. Microarray analysis of RpoS-mediated gene expression in Escherichia coli K-12. Mol Genet Genomics 272:580–591.

54. Layton JC, Foster PL. 2003. Error-prone DNA polymerase IV is controlled by the stress-response sigma factor, RpoS, in Escherichia coli. Molecular Microbiology 50:549–561.

55. Cho B-K, Kim D, Knight EM, Zengler K, Palsson BO. 2014. Genome-scale reconstruction of the sigma factor network in Escherichia coli: topology and functional states. BMC Biol 12:4.

56. Dong T, Yu R, Schellhorn H. 2011. Antagonistic regulation of motility and transcriptome expression by RpoN and RpoS in Escherichia coli. Mol Microbiol 79:375–386.

57. Konovalova A, Grabowicz M, Balibar CJ, Malinverni JC, Painter RE, Riley D, Mann PA, Wang H, Garlisi CG, Sherborne B, Rigel NW, Ricci DP, Black TA, Roemer T, Silhavy TJ, Walker SS. 2018. Inhibitor of intramembrane protease RseP blocks the σE response causing lethal accumulation of unfolded outer membrane proteins. PNAS 201806107.

58. Klein G, Stupak A, Biernacka D, Wojtkiewicz P, Lindner B, Raina S. 2016. Multiple Transcriptional Factors Regulate Transcription of the rpoE Gene in Escherichia coli under Different Growth Conditions and When the Lipopolysaccharide Biosynthesis Is Defective. J Biol Chem 291:22999–23019.

59. Gottesman S. 2019. Trouble is coming: Signaling pathways that regulate general stress responses in bacteria. J Biol Chem jbc.REV119.005593.

60. Barnett MJ, Bittner AN, Toman CJ, Oke V, Long SR. 2012. Dual RpoH Sigma Factors and Transcriptional Plasticity in a Symbiotic Bacterium. Journal of Bacteriology 194:4983–4994.

61. Dong T, Chiang SM, Joyce C, Yu R, Schellhorn HE. 2009. Polymorphism and selection of rpoS in pathogenic Escherichia coli. BMC Microbiology 9:118.

62. Somorin Y, Bouchard G, Gallagher J, Abram F, Brennan F, O'Byrne C. 2017. Roles for RpoS in survival of Escherichia coli during protozoan predation and in reduced moisture conditions highlight its importance in soil environments. FEMS Microbiol Lett 364.

63. Rodríguez-Verdugo A, Gaut BS, Tenaillon O. 2013. Evolution of Escherichia coli rifampicin resistance in an antibiotic-free environment during thermal stress. BMC Evol Biol 13:50.

64. Ferenci T, Galbiati HF, Betteridge T, Phan K, Spira B. 2011. The constancy of global regulation across a species: the concentrations of ppGpp and RpoS are strain-specific in Escherichia coli. BMC Microbiology 11:62.

65. King T, Ishihama A, Kori A, Ferenci T. 2004. A Regulatory Trade-Off as a Source of Strain Variation in the Species Escherichia coli. Journal of Bacteriology 186:5614–5620.

66. Phan K, Ferenci T. 2013. A design-constraint trade-off underpins the diversity in ecologically important traits in species Escherichia coli. ISME J 7:2034–2043.

67. Ferenci T. 2005. Maintaining a healthy SPANC balance through regulatory and mutational adaptation. Molecular Microbiology 57:1–8.

68. Zambrano MM, Siegele DA, Almirón M, Tormo A, Kolter R. 1993. Microbial competition: Escherichia coli mutants that take over stationary phase cultures. Science 259:1757–1760.

69. Notley-McRobb L, King T, Ferenci T. 2002. rpoS mutations and loss of general stress resistance in Escherichia coli populations as a consequence of conflict between competing stress responses. J Bacteriol 184:806–811.

70. Carter MQ, Louie JW, Huynh S, Parker CT. 2014. Natural rpoS mutations contribute to population heterogeneity in Escherichia coli O157:H7 strains linked to the 2006 US spinach-associated outbreak. Food Microbiology 44:108–118.

71. Bleibtreu A, Clermont O, Darlu P, Glodt J, Branger C, Picard B, Denamur E. 2014. The rpoS Gene Is Predominantly Inactivated during Laboratory Storage and Undergoes Source-Sink Evolution in Escherichia coli Species. J Bacteriol 196:4276–4284.

72. Zhang Q, Yan T. 2012. Correlation of Intracellular Trehalose Concentration with Desiccation Resistance of Soil Escherichia coli Populations. Appl Environ Microbiol 78:7407–7413.

73. Sleator RD, Hill C. 2002. Bacterial osmoadaptation: the role of osmolytes in bacterial stress and virulence. FEMS Microbiology Reviews 26:49–71.

74. Rojas E, Theriot JA, Huang KC. 2014. Response of Escherichia coli growth rate to osmotic shock. PNAS 111:7807–7812.

75. Chung HJ, Bang W, Drake MA. 2006. Stress Response of Escherichia coli. Comprehensive Reviews in Food Science and Food Safety 5:52–64.

76. Flint A, Butcher J, Stintzi A. 2016. Stress Responses, Adaptation, and Virulence of Bacterial Pathogens During Host Gastrointestinal Colonization. Microbiology Spectrum 4.

77. Zhang W, Zhu J, Zhu X, Song M, Zhang T, Xin F, Dong W, Ma J, Jiang M. 2018. Expression of global regulator IrrE for improved succinate production under high salt stress by Escherichia coli. Bioresource Technology 254:151–156.

78. Grabowicz M, Silhavy TJ. 2017. Envelope Stress Responses: An Interconnected Safety Net. Trends in Biochemical Sciences 42:232–242.

79. Spöring I, Felgner S, Preuße M, Eckweiler D, Rohde M, Häussler S, Weiss S, Erhardt M. 2018. Regulation of Flagellum Biosynthesis in Response to Cell Envelope Stress in Salmonella enterica Serovar Typhimurium. mBio 9:e00736-17.

80. Dragosits M, Mozhayskiy V, Quinones-Soto S, Park J, Tagkopoulos I. 2013. Evolutionary potential, cross-stress behavior and the genetic basis of acquired stress resistance in Escherichia coli. Mol Syst Biol 9:643.

81. Chakraborty S, Kenney LJ. 2018. A New Role of OmpR in Acid and Osmotic Stress in Salmonella and E. coli. Front Microbiol 9.

82. Cheung KJ, Badarinarayana V, Selinger DW, Janse D, Church GM. 2003. A Microarray-Based Antibiotic Screen Identifies a Regulatory Role for Supercoiling in the Osmotic Stress Response of Escherichia coli. Genome Res 13:206–215.

83. Feugeas J-P, Tourret J, Launay A, Bouvet O, Hoede C, Denamur E, Tenaillon O. 2016. Links between Transcription, Environmental Adaptation and Gene Variability in Escherichia coli: Correlations between Gene Expression and Gene Variability Reflect Growth Efficiencies. Mol Biol Evol 33:2515–2529.

84. Purvis JE, Yomano LP, Ingram LO. 2005. Enhanced Trehalose Production Improves Growth of Escherichia coli under Osmotic Stress. Appl Environ Microbiol 71:3761–3769.

85. Dai X, Zhu M, Warren M, Balakrishnan R, Okano H, Williamson JR, Fredrick K, Hwa T. 2018. Slowdown of Translational Elongation in Escherichia coli under Hyperosmotic Stress. mBio 9.

86. Gunasekera TS, Csonka LN, Paliy O. 2008. Genome-Wide Transcriptional Responses of Escherichia coli K-12 to Continuous Osmotic and Heat Stresses. J Bacteriol 190:3712–3720.

87. Peng S, Stephan R, Hummerjohann J, Tasara T. 2014. Transcriptional analysis of different stress response genes in *Escherichia coli* strains subjected to sodium chloride and lactic acid stress. FEMS Microbiol Lett 361:131–137.

88. Hayden JD, Ades SE. 2008. The Extracytoplasmic Stress Factor, σE, Is Required to Maintain Cell Envelope Integrity in Escherichia coli. PLOS ONE 3:e1573.

89. Somorin Y, Abram F, Brennan F, O'Byrne C. 2016. The General Stress Response Is Conserved in Long-Term Soil-Persistent Strains of Escherichia coli. Appl Environ Microbiol 82:4628–4640.

90. Dusek N, Hewitt AJ, Schmidt KN, Bergholz PW. 2018. Landscape-Scale Factors Affecting the Prevalence of Escherichia coli in Surface Soil Include Land Cover Type, Edge Interactions, and Soil pH. Appl Environ Microbiol 84:e02714-17.

91. Neidhardt FC, Bloch PL, Smith DF. 1974. Culture Medium for Enterobacteria. J Bacteriol 119:736–747.

92. Tyagi D, Kraft AL, Bergholz TM. 2019. Isolation of Bacterial RNA from Foods Inoculated with Pathogens, p. 129–137. *In* Bridier, A (ed.), Foodborne Bacterial Pathogens: Methods and Protocols. Springer, New York, NY.

93. Jandu N, Ho NKL, Donato KA, Karmali MA, Mascarenhas M, Duffy SP, Tailor C, Sherman PM. 2009. Enterohemorrhagic Escherichia coli O157:H7 Gene Expression Profiling in Response to Growth in the Presence of Host Epithelia. PLOS ONE 4:e4889.

94. Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT–PCR. Nucleic Acids Res 29:e45.

95. Ionescu M, Belkin S. 2009. Overproduction of Exopolysaccharides by an Escherichia coli K-12 rpoS Mutant in Response to Osmotic Stress. Appl Environ Microbiol 75:483–492.

96. R Core Team. 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

97. Baranyi J, Roberts TA. 1994. A dynamic approach to predicting bacterial growth in food. Int J Food Microbiol 23:277–294.

98. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data.

99. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

100. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760.

101. Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32:3047–3048.

102. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923–930.

103. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

104. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

105. Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology 11:R25.

106. Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57:289–300.

107. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martínez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muñiz-Rascado L, Ong Q, Paley S, Schröder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, Karp PD. 2013. EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Research 41:D605–D612.

108. Paley S, Parker K, Spaulding A, Tomb J-F, O'Maille P, Karp PD. 2017. The Omics Dashboard for interactive exploration of gene-expression data. Nucleic Acids Res 45:12113–12124.

109. Grossmann S, Bauer S, Robinson PN, Vingron M. 2007. Improved detection of overrepresentation of Gene-Ontology annotations with parent–child analysis. Bioinformatics 23:3024–3031.

110. Letunic I, Bork P. 2007. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23:127–128.

111. Rajkumari K, Gowrishankar J. 2001. In Vivo Expression from the RpoS-Dependent P1 Promoter of the Osmotically Regulated proU Operon inEscherichia coli and Salmonella entericaSerovar Typhimurium: Activation by rho andhns Mutations and by Cold Stress. Journal of Bacteriology 183:6543–6550.

112. McDowall JS, Murphy BJ, Haumann M, Palmer T, Armstrong FA, Sargent F. 2014. Bacterial formate hydrogenlyase complex. PNAS 111:E3948–E3956.

113. Gouffi K, Pica N, Pichereau V, Blanco C. 1999. Disaccharides as a New Class of Nonaccumulated Osmoprotectants for Sinorhizobium meliloti. Applied and Environmental Microbiology 65:1491–1500.

114. Reitzer L, Schneider BL. 2001. Metabolic Context and Possible Physiological Themes of ς54-Dependent Genes in Escherichia coli. Microbiol Mol Biol Rev 65:422–444.

115. Klose KE, Weiss DS, Kustu S. 1993. Glutamate at the site of phosphorylation of nitrogen-regulatory protein NTRC mimics aspartyl-phosphate and activates the protein. J Mol Biol 232:67–78.

116. Arense P, Bernal V, Iborra JL, Cánovas M. 2010. Metabolic adaptation of Escherichia coli to long-term exposure to salt stress. Process Biochemistry 45:1459–1467.

117. Olesen I, Jespersen L. 2010. Relative gene transcription and pathogenicity of enterohemorrhagic Escherichia coli after long-term adaptation to acid and salt stress. International Journal of Food Microbiology 141:248–253.

118. Wu X, Altman R, Eiteman MA, Altman E. 2014. Adaptation of Escherichia coli to Elevated Sodium Concentrations Increases Cation Tolerance and Enables Greater Lactic Acid Production. Appl Environ Microbiol 80:2880–2888.

119. Zhu M, Dai X. 2018. High Salt Cross-Protects Escherichia coli from Antibiotic Treatment through Increasing Efflux Pump Expression. mSphere 3.

120. Cayley S, Record MT. 2003. Roles of Cytoplasmic Osmolytes, Water, and Crowding in the Response of Escherichia coli to Osmotic Stress: Biophysical Basis of Osmoprotection by Glycine Betaine. Biochemistry 42:12596–12609.

121. Levert M, Zamfir O, Clermont O, Bouvet O, Lespinats S, Hipeaux MC, Branger C, Picard B, Saint-Ruf C, Norel F, Balliau T, Zivy M, Le Nagard H, Cruvellier S, Chane-Woon-

Ming B, Nilsson S, Gudelj I, Phan K, Ferenci T, Tenaillon O, Denamur E. 2010. Molecular and Evolutionary Bases of Within-Patient Genotypic and Phenotypic Diversity in Escherichia coli Extraintestinal Infections. PLoS Pathog 6.

122. Barchiesi J, Espariz M, Checa SK, Soncini FC. 2009. Downregulation of RpoN-controlled genes protects Salmonella cells from killing by the cationic antimicrobial peptide polymyxin B. FEMS Microbiology Letters 291:73–79.

123. Riordan JT, Tietjen JA, Walsh CW, Gustafson JE, Whittam TS. 2010. Inactivation of alternative sigma factor 54 (RpoN) leads to increased acid resistance, and alters locus of enterocyte effacement (LEE) expression in Escherichia coli O157:H7. Microbiology (Reading) 156:719–730.

124. Becker LA, Bang I-S, Crouch M-L, Fang FC. 2005. Compensatory role of PspA, a member of the phage shock protein operon, in rpoE mutant Salmonella enterica serovar Typhimurium. Molecular Microbiology 56:1004–1016.

125. Bury-Moné S, Nomane Y, Reymond N, Barbet R, Jacquet E, Imbeaud S, Jacq A, Bouloc P. 2009. Global Analysis of Extracytoplasmic Stress Signaling in Escherichia coli. PLOS Genetics 5:e1000651.

126. Storvik KAM, Foster PL. 2010. RpoS, the Stress Response Sigma Factor, Plays a Dual Role in the Regulation of Escherichia coli's Error-Prone DNA Polymerase IV. Journal of Bacteriology 192:3639–3644.

127. Mitra A, Fay PA, Morgan JK, Vendura KW, Versaggi SL, Riordan JT. 2012. Sigma Factor N, Liaison to an ntrC and rpoS Dependent Regulatory Pathway Controlling Acid Resistance and the LEE in Enterohemorrhagic Escherichia coli. PLOS ONE 7:e46288.

128. Du H, Wang M, Luo Z, Ni B, Wang F, Meng Y, Xu S, Huang X. 2011. Coregulation of Gene Expression by Sigma Factors RpoE and RpoS in Salmonella enterica Serovar Typhi During Hyperosmotic Stress. Curr Microbiol 62:1483–1489.

129. Clermont O, Christenson JK, Denamur E, Gordon DM. 2013. The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ Microbiol Rep 5:58–65.

130. Touchon M, Perrin A, Sousa JAM de, Vangchhia B, Burn S, O'Brien CL, Denamur E, Gordon D, Rocha EP. 2020. Phylogenetic background and habitat drive the genetic diversification of Escherichia coli. PLOS Genetics 16:e1008866.

131. Novais Â, Vuotto C, Pires J, Montenegro C, Donelli G, Coque TM, Peixe L. 2013. Diversity and biofilm-production ability among isolates of Escherichia coli phylogroup D belonging to ST69, ST393 and ST405 clonal groups. BMC Microbiol 13:144.

132. Chen J, Lee SM, Mao Y. 2004. Protective effect of exopolysaccharide colanic acid of Escherichia coli O157:H7 to osmotic and oxidative stress. International Journal of Food Microbiology 93:281–286.

133. Ambard-Bretteville F, Sorin C, Rébeillé F, Hourton-Cabassa C, Colas des Francs-Small C. 2003. Repression of formate dehydrogenase in *Solanum tuberosum* increases steady-state levels of formate and accelerates the accumulation of proline in response to osmotic stress. Plant Molecular Biology 52:1153–1168.

134. Topp E, Welsh M, Tien Y-C, Dang A, Lazarovits G, Conn K, Zhu H. 2003. Strain-dependent variability in growth and survival of Escherichia coli in agricultural soil. FEMS Microbiology Ecology 44:303–308.

135. Jang J, Hur H-G, Sadowsky MJ, Byappanahalli MN, Yan T, Ishii S. 2017. Environmental Escherichia coli: ecology and public health implications—a review. Journal of Applied Microbiology 123:570–581.

136. Brennan FP, Grant J, Botting CH, O'Flaherty V, Richards KG, Abram F. 2013. Insights into the low-temperature adaptation and nutritional flexibility of a soil-persistent Escherichia coli. FEMS Microbiology Ecology 84:75–85.

137. Polissi A, De Laurentis W, Zangrossi S, Briani F, Longhi V, Pesole G, Dehò G. 2003. Changes in Escherichia coli transcriptome during acclimatization at low temperature. Res Microbiol 154:573–580.

138. Whitman RL, Shively DA, Pawlik H, Nevers MB, Byappanahalli MN. 2003. Occurrence of Escherichia coli and enterococci in Cladophora (Chlorophyta) in nearshore water and beach sand of Lake Michigan. Appl Environ Microbiol 69:4714–4719.

139. Ishii S, Yan T, Vu H, Hansen DL, Hicks RE, Sadowsky MJ. 2010. Factors Controlling Long-Term Survival and Growth of Naturalized Escherichia coli Populations in Temperate Field Soils. Microb Environ 25:8–14.

140. White-Ziegler CA, Um S, Pérez NM, Berns AL, Malhowski AJ, Young S. 2008. Low temperature (23 °C) increases expression of biofilm-, cold-shock- and RpoS-dependent genes in Escherichia coli K-12. Microbiology 154:148–166.

141. Vidovic S, Mangalappalli-Illathu AK, Korber DR. 2011. Prolonged cold stress response of Escherichia coli O157 and the role of rpoS. International Journal of Food Microbiology 146:163–169.

142. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G,

Alekseyev MA, Pevzner PA. 2012. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol 19:455–477.

143. Gardner SN, Slezak T, Hall BG. 2015. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics 31:2877–2878.

144. Collins C, Didelot X. 2018. A phylogenetic method to perform genome-wide association studies in microbes that accounts for population structure and recombination. PLOS Computational Biology 14:e1005958.

145. Barria C, Malecki M, Arraiano CMY 2013. Bacterial adaptation to cold. Microbiology 159:2437–2443.

146. Phadtare S. 2004. Recent developments in bacterial cold-shock response. Curr Issues Mol Biol 6:125–136.

147. Broeze RJ, Solomon CJ, Pope DH. 1978. Effects of low temperature on in vivo and in vitro protein synthesis in Escherichia coli and Pseudomonas fluorescens. Journal of Bacteriology 134:861–874.

148. Farewell A, Neidhardt FC. 1998. Effect of Temperature on In Vivo Protein Synthetic Capacity in Escherichia coli. Journal of Bacteriology 180:4704–4710.

149. Sant'Ana AS, Franco BDGM, Schaffner DW. 2012. Modeling the growth rate and lag time of different strains of Salmonella enterica and Listeria monocytogenes in ready-to-eat lettuce. Food Microbiology 30:267–273.

150. Seel W, Derichs J, Lipski A. 2016. Increased Biomass Production by Mesophilic Food-Associated Bacteria through Lowering the Growth Temperature from 30°C to 10°C. Appl Environ Microbiol 82:3754–3764.

151. Andrews SC, Robinson AK, Rodríguez-Quiñones F. 2003. Bacterial iron homeostasis. FEMS Microbiology Reviews 27:215–237.

152. Adler C, Corbalan NS, Peralta DR, Pomares MF, de Cristóbal RE, Vincent PA. 2014. The Alternative Role of Enterobactin as an Oxidative Stress Protector Allows Escherichia coli Colony Development. PLoS One 9.

153. Abdul-Tehrani H, Hudson AJ, Chang Y-S, Timms AR, Hawkins C, Williams JM, Harrison PM, Guest JR, Andrews SC. 1999. Ferritin Mutants of Escherichia coli Are Iron Deficient and Growth Impaired, and fur Mutants are Iron Deficient. Journal of Bacteriology 181:1415–1428.

154. Wu Y, Outten FW. 2009. IscR Controls Iron-Dependent Biofilm Formation in Escherichia coli by Regulating Type I Fimbria Expression. Journal of Bacteriology 191:1248–1257.

155. Babu M, Díaz-Mejía JJ, Vlasblom J, Gagarinova A, Phanse S, Graham C, Yousif F, Ding H, Xiong X, Nazarians-Armavil A, Alamgir M, Ali M, Pogoutse O, Pe'er A, Arnold R, Michaut M, Parkinson J, Golshani A, Whitfield C, Wodak SJ, Moreno-Hagelsieb G, Greenblatt JF, Emili A. 2011. Genetic Interaction Maps in Escherichia coli Reveal Functional Crosstalk among Cell Envelope Biogenesis Pathways. PLoS Genet 7.

APPENDIX. ASSOCIATION OF LOW TEMPERATURE GROWTH PHENOTYPES WITH GENOMIC VARIANTS IN *E. COLI* ISOLATES FROM SOIL

A.1. Introduction

Escherichia coli is one of a few bacteria that possess a wide variation in phenotypic traits and therefore a large capability for adaptation. The pangenome of *E. coli* is open even as more genomes are sequenced, with core genome sizes of at least 2,000 genes (41, 45, 46). The inclusion of genes through horizontal gene transfer and other gene acquisition processes only increase the adaptive qualities of *E. coli* (44). Therefore *E. coli* has a wide range of phenotypic and genetic variation that is constantly being explored and explained.

Extrahost environments such as soils have been suggested as selective pressures on stress and survival phenotypes for *E. coli* (40, 43, 89, 134). In most cases, *E. coli* from natural environments often possess higher tolerance than host- or lab-associated strains when grown under stressful conditions (62, 89) which can be attributed to the multiple stresses that are present in their habitats (135). One important stress that is observed in soil *E. coli* is cold temperature (~15°C) (136). Commonly the cold stress responses of *E. coli* are induced around 20°C (137). In soil and water environments, *E. coli* is known to survive and grow in low temperatures, some as low as 4°C as seen in *Cladophora* algae mats in the Great Lakes (43, 138, 139). Past studies have observed phenotypes under suboptimal growth, but most have been in K12, commensal, and pathogenic strains (140, 141). Characterization of environmental *E. coli* can expand our understanding of cold stress phenotypes as well as the effects of the environment on adaptation.

We strived to look at the variation of growth phenotypes among phylogroup D *E. coli* isolates at a low temperature (15°C) and determine genetic factors that are associated with this

variation. We observed the extent of growth rates and yields of 244 *E. coli* in optimal (37°C) and suboptimal temperatures (15°C). A GWAS analysis using TreeWAS was performed to look at underlying genetic variants associated with the phenotypes at low temperature. The genes of the significant variants were then functionally classified to aid what pathways were being affected most by genetic variation under low temperature.

A.2. Materials and Methods

The *E. coli* isolates used in the following experiments and assays are a part of an *E. coli* isolate collection which were obtained from surface soil samples along the Buffalo River of North Dakota and Minnesota (90). The collection includes 256 culturable phylogroup D isolates, of which 244 were used in our study.

Isolates were inoculated onto LB agar from freezer stock cultures and incubated at 37°C for 24 hours. Isolates were inoculated, via a pin replicator, into wells of four 96-well plates (1° plates) where each well contained 150 µL of sterile GDMM. Negative control wells were present in each plate throughout the acclimation process. The 1° plates were sealed and incubated for 8 hours at 37°C. After 8 hours, 3 µL of the 1° cultures were transferred to secondary plates (2° plates) that contained 147 µL of sterile GDMM. The 2° plates were then sealed and incubated for 16 hours at 37°C.

Tertiary cultures were made by transferring 20 µL of 2° cultures to 180 µL of pre-chilled (15°C) GDMM within 96-well microtiter plates. Wells that were used as negative controls received 20 µL of sterile GDMM. Plates were then sealed with optically clear sealing film. The OD_{600} was read by a plate reader (Biotek, Winnoski, VT) immediately after plates were sealed. After a OD_{600} reading, plates were incubated at 15°C so that the isolates could grow in low temperature conditions. Additional readings every 3 hours were taken until 56 hours was

reached. This time limit was set as most of the cultures reached a stable stationary phase by 56 hours based on prior tests cultures (data not shown). For these additional readings, heating blocks were placed on top of the test plates for 1 minute to dissipate condensation. To determine the growth rate and growth yield of the isolates at 15°C, the OD_{600} readings were compiled, log-transformed, and analyzed in R using the R/nlsMicrobio "baranyi" growth model and R/vegan packages (96, 97).

The genomes of the 244 phylogroup D *E. coli* were sequenced using paired-end 150bp reads on an Illumina HiSeq 4000 (Macrogen Clinical Labs, Seoul, KR). Genomes were assembled using SPAdes v. 3.10 (142) and variants were called using kSNP v. 3.1 (143). The resulting variants were filtered for presence in 10% to 80% of genomes.

The association between the growth phenotypes and variants in 15°C was performed by the software TreeWAS (144). This program accounted more than original GWAS methods as it provides a phylogeny-based process that enhances the associations between the phenotype and variants due to a deeper insight of population structure. The programs RAxML and ClonalFrameML were implemented for the associations made for this study's isolates. The strain UMN026 was used as the reference sequence for the TreeWAS analysis.

A.3. Results

A.3.1. Growth Phenotypes of *E. coli* in 37°C and 15°C Conditions

Under the 15°C condition the *E. coli* isolates grew at slower rates than in the 37°C condition (0.139 ± 0.016 and 1.123 ± 0.151 $\log_{10}(OD_{600})$ h^{-1}, respectively.) There was a wide variability in the 37°C while the variability was limited in the 15°C condition. The difference in growth rates was significant ($p < 2.2 \times 10^{-16}$) between the two temperature conditions (Figure A1). In contrast, growth yields of the low temperature and optimal temperature were not significant as the

medians of the isolates were -0.519 ± 0.043 and -0.556 ± 0.054 $\log_{10}$ (OD_{600}), respectively. These observations detail that while the growth yield was unaffected by the lower temperature, the growth rate decreased drastically in comparison to the normal optimal temperature.

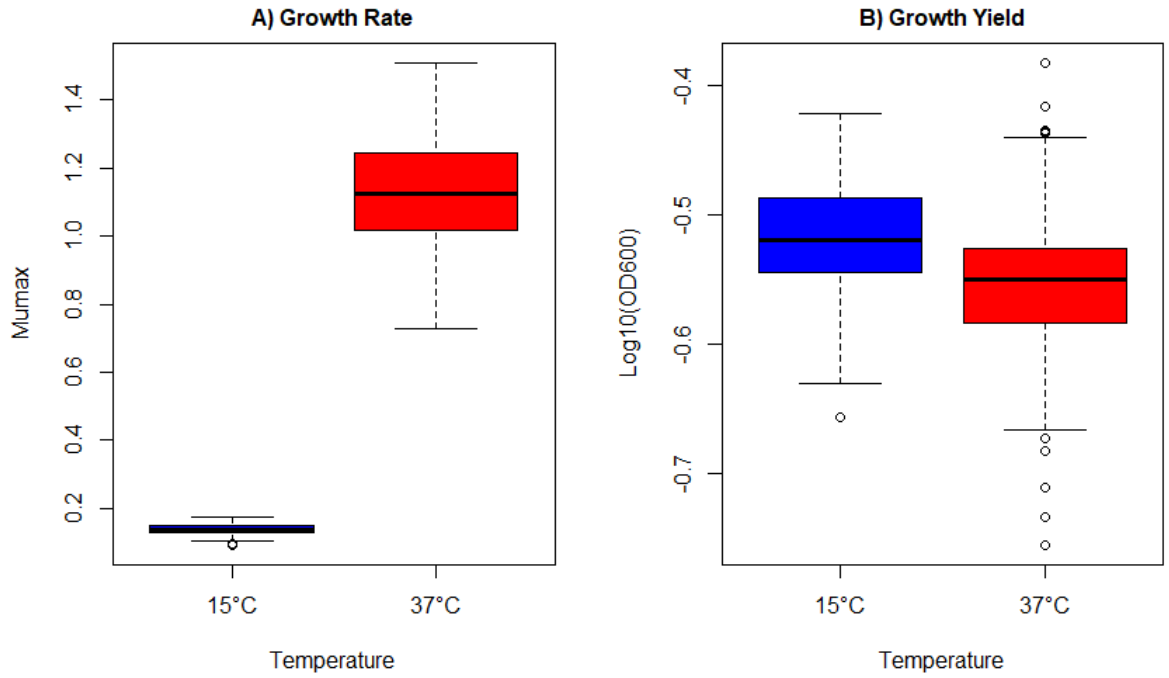

Figure A1. Growth phenotypes of soil phylogroup D *E. coli* isolates at 15°C and 37°C. A) Growth rate and B) growth yield. All growth measurements were taken in triplicate for each isolate.

A.3.2. TreeWAS Analysis

We implemented TreeWAS to perform a GWAS to delve into the association between genetic variants and the growth phenotypes at 15°C. We observed 1,043 variants that were associated with growth phenotypes at 15°C in the core and accessory genome levels (Figure A2). A total of 454 variants within the core genomes were associated with 15°C growth rate were called by TreeWAS: 42 intergenic variants, 95 missense variants, and 317 synonymous variants. Within the missense variants, 72 variants were significantly associated with growth rate at 15°C (Figure A3A). A total of 589 genetic variants in the core genomes were associated with 15°C

growth yield (Figure A3B). Further classification found that 64 variants were intergenic, 94 were missense, 429 were synonymous, and one was a splice region variant. There were 64 missense variants that were significantly associated with growth yield at 15°C.

In terms of the accessory genome, 9 genes were associated with growth rate and 13 genes were associated with growth yield at 15°C (Figure A4).

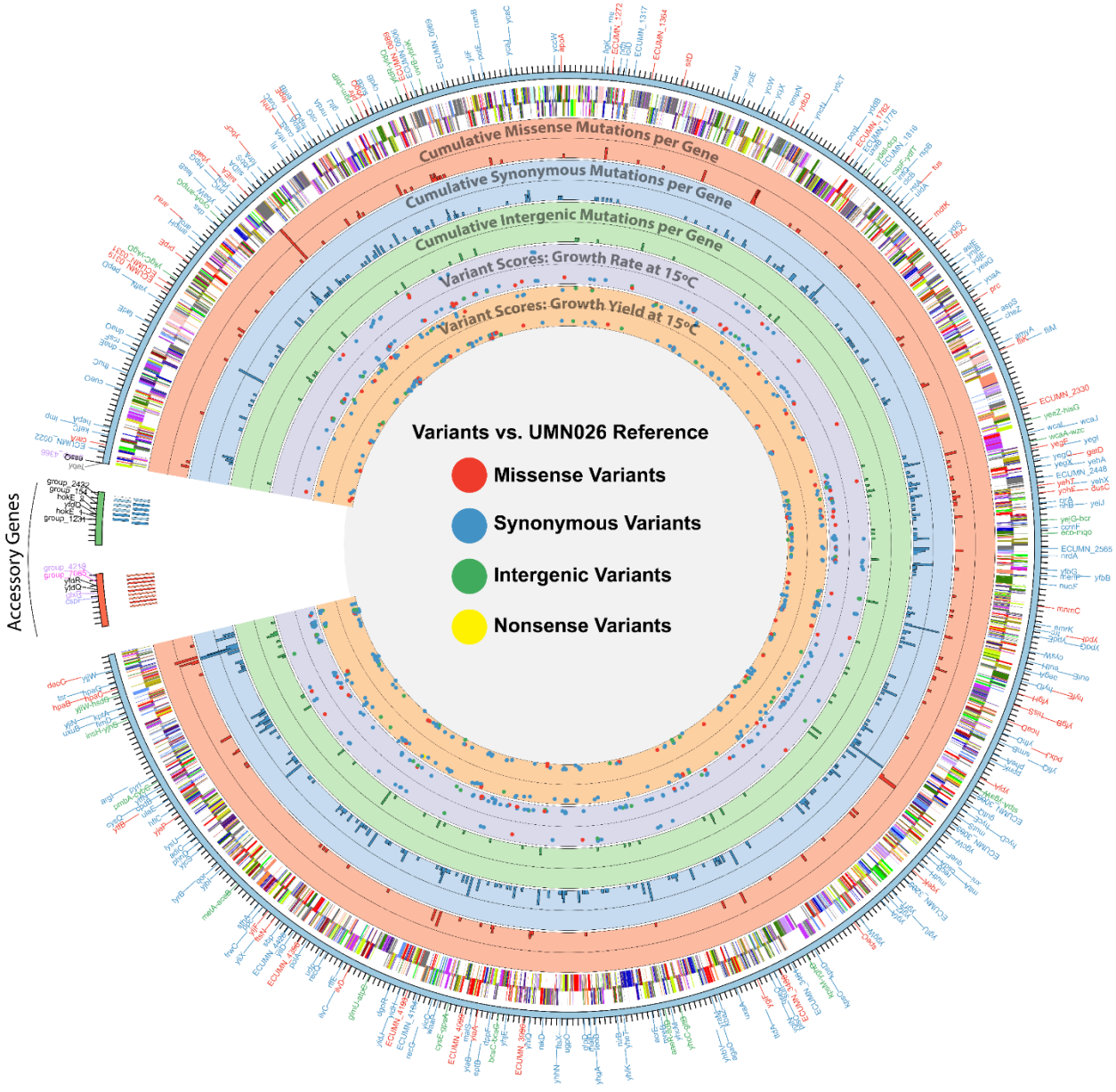

Figure A2. Map of variants associated with growth rate and yield at 15ºC in MOPS minimal media with 0.1% glucose (PBonferroni < 0.5). Outer track: Gene names colored by variant

type, Track 2: Core genome SNP variant chromosome called using the complete *E. coli* UMN026 genome as a reference (light blue), accessory gene content variants associated with growth rate (red), accessory gene content variants associated with growth yield (green). **Accessory gene tiles** are lightly shaded when associated with decreased growth. Tracks 3 & 4: Gene positions in the complete *E. coli* UMN026 genome, colored by COG on the forward and reverse strand, respectively. Tracks 5, 6, and 7: Cumulative counts of variants in each gene for missense (red), synonymous (blue) and intergenic (green) variants, respectively. Tracks 8 and 9: Variant locations, types, and associations with log-growth rate per h (purple) and growth yield (orange), respectively.

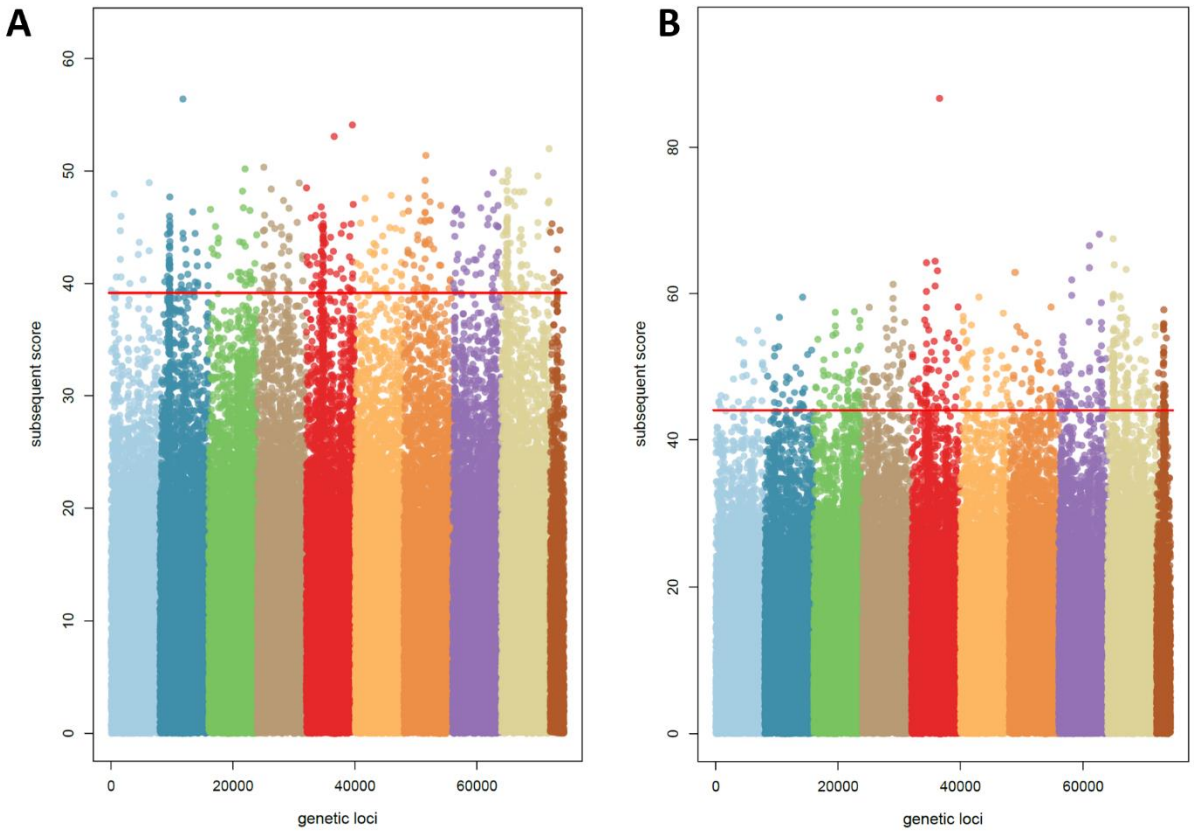

Figure A3. Manhattan plots of core genome SNP variants associated with 15°C A) growth rate and B) growth yield. Points above the significance threshold (red line) show significant association.

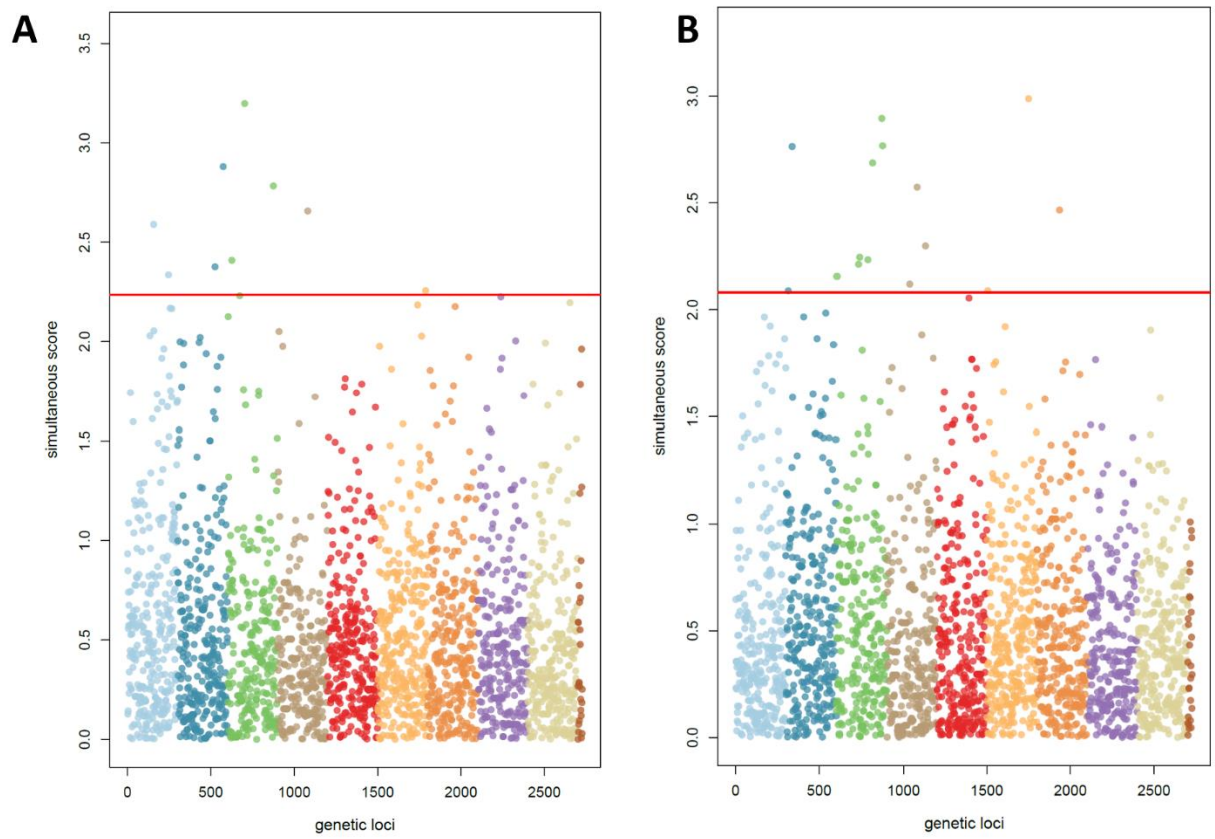

Figure A4. Manhattan plots of accessory genome SNP variants associated with 15°C A) growth rate and B) growth yield. Points above the significance threshold (red line) show significant association.

The missense variables that were significantly associated with both growth phenotypes functioned in aromatic carbon metabolism, iron acquisition, central metabolism, and membrane/capsule synthesis (Table A1). Most of the variants were negatively associated with the growth phenotypes in both temperature conditions although there were eight variants that showed positive associations with at least one phenotype in one condition. There were five genetic variants that had negative association on both phenotypes in 15°C belonging to *hpaC*, *hpaB*, *hpaG*, and *adrB*. There were also 19 variants that were associated with either growth rate or growth yield at 15°C. Some variants also showed associations based on growth phenotype: seven variants associated with growth rate and five variants associated with growth yield.

Table A1. Missense variants of core and accessory genomes with significant associations (subsequent scores [core] & simultaneous scores [accessory]) in 15°C growth phenotypes. Mumax= growth rate, Nmax= growth yield.

				Mumax		Nmax	
Gene	**Product**	**Position**	**Variant**	**15°C**	**37°C**	**15°C**	**37°C**
	Aromatic Carbon Metabolism						
hpaC	4-hydoxyphenylacetate 3-monooxygenase small component	5136662	T→C; Thr30Ala	-44.6	NA	-49.5	NA
hpaB	4-hydroxyphenylacetate 3-monooxygenase large component	5137714	C→T; Val206Ile	-45.5	NA	-50.2	NA
		5137785	G→A; Thr182Ile	-42.9	NA	-47.3	NA
		5138307	G→A; Ala8Val	NA	NA	-44.4	NA
hpaG	4-hydroxyphenylacetate degradation bifunctional isomerase/decarboxylase	5147983	G→C; Pro53Arg	-40.8	NA	-51.1	-39.1
		5148047	C→T; Ala32Thr	-41.9	NA	-49.2	NA
	Iron Acquisition						
sitD	Iron transport protein, inner membrane component	1452588	T→A; Lys193Ile	-49.0	NA	NA	NA
entF	Enterobactin synthase subunitF	731497	G→A; Gly739Ser	39.3	NA	NA	NA
fepE	ferric enterobactin transport protein FepE	733310	A→T; Lys17Asn	45.4	NA	NA	45.7
		733480	A→C; Glu74Ala	-44.2	NA	NA	NA
entB	Isochorimatase; aids in siderophore formation	742794	A→C; Glu54Ala	NA	-34.0	-50.8	NA
	Central Metabolism						
yadE	Putative exported polysaccharide deacetylase	149736	T→A; Phe114Tyr	NA	NA	-55.0	NA
prpE	Propinoyl-CoA synthetase	412437	C→T; Arg575Cys	NA	NA	55.5	NA
adrB	Putative cyclic-di-GMP phosphodiesterase	2132279	T→C; Val141Ala	-44.0	-36.5	-44.4	NA
ydcM	Putative transposase	2665175	A→C,T; Glu67Gly	-39.8	-36.9	NA	NA
		4913409	A→C,T; Glu67Ala	-39.8	-36.9	NA	NA
pdxJ	Pyridoxine 5'-phosphate synthase	2992908	G→T; Asp114Glu	NA	NA	-52.7	-40.1
Wag	Glucosyltransferase I	4265431	T→G; His307Pro	-42.2	NA	NA	NA
treC	Trehalose-6-phosphate hydrolase	4951024	T→C; Ile471Val	NA	NA	45.1	NA
		4951791	A→T; Leu215Gln	-45.2	NA	NA	NA
		4951796	T→A; Glu213Asp	41.8	33.1	NA	NA
	Membrane & Capsule Composition						
siiEA	Adhesin for cattle intestine colonization	584893	G→A; Ala3803Thr	NA	NA	-56.5	-38.7
yehX	Putative ABC transporter ATP-binding protein	2519791	T→C; Asn55Ser	NA	NA	-52.9	-41.1

Table A1. Missense variants of core and accessory genomes with significant associations (subsequent scores [core] & simultaneous scores [accessory]) in 15°C growth phenotypes (continued).

				Mumax		Nmax	
Gene	**Product**	**Position**	**Variant**	**15°C**	**37°C**	**15°C**	**37°C**
	Membrane & Capsule Composition (continued)						
emrK	EmrKY-TolC multidrug resistance efflux pump, membrane fusion protein component	2775313	A→G; Val205Ala	NA	NA	-50.1	-51.2
yfgH	Putative outer membrane lipoprotein	2909149	C→T; Thr13Met	45.4	NA	NA	NA
kpsD	Polysialic acid transport protein	3533251	T→C; Ser377Pro	-54.1	NA	NA	NA
ytfB	Putative cell envelope opacity-associated protein	4914582	G→A; Pro204Ser	-39.2	-33.6	NA	NA
	Hypothetical proteins						
ybaP	Hypothetical protein	566195	G→T; Pro109Thr	NA	NA	56.8	45.6
ydbD	Hypothetical protein	1669291	C→A; Ala525Glu	-40.0	-33.7	NA	NA
yjeP	Hypothetical protein	4872361	C→G; Trp567Ser	45.2	34.2	NA	NA
ydfU	Hypothetical protein	1873521	G→A,T; Pro234Ser	39.3	NA	NA	NA

A.4. Discussion

A.4.1. Growth Rates Differ between 37°C and 15°C Conditions

Under cold temperatures bacterial growth is reduced. Cold temperatures cause a decrease in membrane fluidity and stabilization of nucleic acids especially mRNA (145). These two functions are important in downstream processes such as transcription and translation so ultimately cellular growth rate is dramatically decreased or even halted (146). The concentration of non-translating ribosomes increased in *E. coli* grown at 15°C suggested that the growth rate was affected by the lack of translating ribosomes (147, 148). Therefore, most genes are not being transcribed except for the small group of cold shock proteins. The negative impact of cold temperature on growth rate was seen in three strains of *Salmonella enterica* that had reduced growth rates from 0.32-0.44 h^{-1} to 0.094-0.149 h^{-1} in 30°C and 15°C environments (149). Our data agreed with this study as growth rates of our isolates reduced due to the decrease in

temperature. The range of growth rates in the study and in our data were similar at 15°C, but the range in the warmer temperatures differed. This difference may be due to our control temperature being at 37°C instead of the 30°C detailed in the study. Similarly, the growth of two *E. coli* strains from 30°C to 15°C saw a drastic reduction from a growth rate of ~0.80 h^{-1} to < 0.15 h^{-1} (150). The reduction between environments and the growth rate range showed similarities with our data. Overall, our *E. coli* isolates demonstrated a significant decrease in growth rate from the 37°C to the 15°C condition.

The growth yields between the two temperature environments were insignificant between our isolates. A study looking at a variety of mesophiles at 10°C, showed that *E. coli* isolates had significant increased biomass when compared with the bacteria at 30°C (150). This increase was attributed to the cell surface area increasing to promote the influx of nutrients to increase cellular functionality. The insignificant growth yields of our isolates between the conditions are opposite of what has been noted in this study and could be attributed to the observation of large variances of isolates in both conditions. While the growth rate decreased significantly from 37°C to 15°C, there was not significance in growth yield.

A.4.2. Variants Involved in Iron Acquisition and Membrane Composition are Associated with Low Temperature Phenotypes

Iron is a necessary element in bacterial processes such as gene regulation, metabolic respiration, and DNA biosynthesis (151). The uptake and sequestration of iron allows for continuous bacterial growth. Prior studies have observed lowered growth when iron transport and storage genes are mutated (152–154). In the significant genetic variants, there were four genes that impacted the growth phenotypes negatively. These variants were missense mutations,

and likely the function of these genes was inhibited or detrimentally altered resulting in lowered growth.

Five missense mutations that affected membrane composition and functionality were negatively associated with the growth phenotypes at 15°C. In low temperatures, the cell membrane becomes rigid and therefore mutations in the composition and transport adds another stress. There was one gene, *yfgH*, that had a missense mutation that was positively associated with growth rate at 15°C. This gene has been shown as an important factor in membrane integrity as a *yfgH* mutant displayed lysis of the cell and increase oxidative stress (155). The membrane is one of the most affected systems in low temperatures, so the positive association of *yfgH* to the 15°C growth phenotypes coincide with the need for membrane stability.

A.5. Conclusion

Cold stress and genetic variation influenced the varying growth phenotypes of *E. coli* grown at a low temperature when compared to 37°C. There was a significant decrease in growth rate but not in growth yield from the 37°C to the 15°C condition among all isolates. TreeWAS identified 72 and 64 missense genetic variants in the core genomes that were associated to 15°C growth rate and yield, respectively. Additionally, 9 and 13 accessory genes were also associated with growth rate and yield at 15°C. A large percentage of missense variants were involved in iron acquisition, membrane composition, and aromatic carbon metabolism. Overall, the association between the observed growth phenotypes and genetic variants suggests environmental adaptation to soil, but future work in genome editing experiments is needed to ascertain causation.