UNIVERSITY OF CALIFORNIA SAN DIEGO

Conservation of the Fur Regulon Persists in Non-Model, Clinical *Escherichia coli*

A Thesis submitted in partial satisfaction of the requirements
for the degree Master of Science.

in

Biology

by

Daniel Quach

Committee in charge:

Professor Bernard Palsson Chair
Professor Milton Saier Co-Chair
Professor Gen-Feng Sheng

2022

The Thesis of Daniel Quach is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2022

TABLE OF CONTENTS

THESIS APPROVAL PAGE iii

TABLE OF CONTENTS iv

LIST OF FIGURES vi

LIST OF TABLES vii

ABSTRACT OF THESIS viii

INTRODUCTION 1

CHAPTER 1 5

CHAPTER 2 10

CHAPTER 3 15

CHAPTER 4 21

CHAPTER 5 25

CONCLUSION 28

6.1 - Conclusions on Fur's Impact on Growth, Future Directions 28

6.2 - Conclusions on the Fur Regulon, Future Directions 29

6.3 - Conclusions on the Fur Regulatory Methods and Binding Motifs 31

6.4 - Conclusions on Fur-Dependent Antibiotic Resistance and Antibiotic Vulnerability 31

6.5 - Study Limitations, Final Remarks 33

METHODOLOGY 36

7.1 - Bacterial Strains, Media, and Growth Conditions 36

7.2 - ChIP-exo Experiments 37

7.3 - RNA-seq Expression Profiling 38

7.4 - Read Analysis 39

7.5 - Fur Binding Motif Survey of ChIP-exo Data 39

7.6 - Growth Curve Analysis 39

7.7 - Antibiotic Resistance Phenotypic Microarrays 40

7.8 - Phylogenetic Analysis 40

REFERENCES 42

LIST OF FIGURES

Figure 1 – Phylogenetic Comparison and Growth Analysis of E. coli 131 and E. coli 158……..5

Figure 2 – ChIP-exo and RNA-seq profiling of E. coli 131 Fur and E. coli 158 Fur…….…….....10

Figure 3 – The Fur Regulon of E. coli 131 and E. coli 158.....15

Figure 4 – Regulatory Modes in E. coli 131 and Fur Binding Motifs of E. coli 131 and E. coli 158……..21

Figure 5 – Fur knockout increases resistance to some antibiotics and increases vulnerability to others in E. coli 131 and E. coli 158……..25

LIST OF TABLES

Table 1 – List of bacterial strains used in this study…….…….....7

ABSTRACT OF THESIS

Conservation of the Fur Regulon Persists in Non-Model, Clinical *Escherichia coli*

By

Daniel Quach

Master of Science in Biology

University of California San Diego, 2022

Professor Bernard Palsson, Chair
Professor Milton Saier, Co-Chair

The ferric uptake regulator, Fur, is a highly conserved global transcriptional regulator in gram-negative bacteria that maintains iron homeostasis through its binding of Fe^{2+}. Proper management of iron ensures that pivotal processes such as DNA repair, aerobic respiration, and cell proliferation can occur fluidly without threat of iron-mediated ROS production and accumulation. Though Fur's targets, regulon, regulatory methods, and effects on phenotypes have been extensively studied in many model *E. coli* strains, its range of action in non-model

systems remains unexplored. Studying Fur in the latter scenario might provide data that is paramount towards expanding our knowledge of microbial iron metabolism. In this study, we combine genome-scale data from high-throughput sequencing (ChIP-exo) with strand-specific massively parallel complementary DNA sequencing (RNA-seq) to uncover the Fur regulon of two clinical *E. coli* strains isolated from bacteremia patients, No. 131 and No. 158. Furthermore, we assess the effect of Fur on bacterial growth and antibiotic resistance by knocking out Fur and comparing *fur* mutants to their wild-type counterparts. We find that the Fur regulon of the two strains shares major conserved regions that overlap with the closely related uropathogenic (UPEC) *E. coli* strain, CFT073, and the distantly related *E. coli* K-12 strain, MG1655. This is not without outliers, nonetheless – the unique genes that do surface in each clinical strain's respective regulons provide us with a distinct perspective by which to explore Fur's regulatory modes and regulatory scope. We also find that an absence of Fur does hinder growth of the two strains and changes the scope of their antibiotic resistances albeit in distinct fashions.

INTRODUCTION

Iron is a fundamental transition metal in nearly all biological systems. Its universal application reaches as far back as Earth's infancy where iron was coincident with the earliest precursors to modern day life in an initially anaerobic environment. Billions of years of evolution later, iron has found its way into a variety of common biological processes and structures. In the hemoglobin crucial to erythrocytes, it can bind to and aid in the transport of oxygen. It can be complexed with sulfur into Fe-S clusters capable of binding DNA repair elements for activation or continuity of their own progress. Cytochrome C, a complex featured in aerobic respiration, relies on iron in order to carry electrons through the electron transport chain of mitochondria[43]. The expression of whole operons can be markedly controlled by iron concentration as to account for iron's availability and persistence in associated downstream events[41].

Any deficit in iron can thus prove compromising to sustainability and survivability. This is especially apparent in bacteria - organisms that typically grow and thrive at a much more rapid rate than any larger eukaryotic organism can. Many of the aforementioned processes are inherent to bacteria. Given this glaring importance, bacterial species will frequently compete with other microbial species and or among themselves to secure iron. This can be done through a plethora of proteins and protein complexes, though may be more commonly overseen through means of "siderophores" - small molecules with a high affinity for binding Fe^{3+} and other forms of iron[35]. All of these mechanisms put in play cannot proceed wantonly, however. The interaction of iron with superoxide and hydrogen peroxide through Fenton and or Haber-Weiss reactions can produce threatening ROS molecules capable of damaging innate cellular machinery[3,13]. This discourages the excessive accumulation of iron in cytoplasmic space as to prevent such

unwanted side effects. It then becomes a necessity that bacteria have measures in place to properly attenuate iron concentration in order to thread the fine balance between essentiality and toxicity. Here, the ferric uptake regulator, abbreviated "Fur", makes its mark.

Fur is a transcription factor responsible for overseeing various genes tied directly to iron uptake and iron storage, as well as downstream pathways correlated to the previous. It is one of the most important transcription factors in maintaining iron homeostasis as described above. Its activity is usually dictated by binding of Fe^{2+} to an allosteric site – this interaction facilitating Fur binding to *cis*-regulatory elements on the bacterial genome. Under iron-replete conditions, or conditions where iron is plentiful, iron can bind Fur to increase Fur DNA binding activity - this model dubbed *holo*-Fur. Here, Fur is classically reported to repress most of the genes it targets including genes of iron uptake mentioned previously (*holo*-Fur repression, HR). Nonetheless, accounts of the opposite regulatory pattern, *holo*-Fur activation (HA), have also been documented[37]. Under iron-deplete conditions, or conditions where iron is lacking, iron will not bind Fur, thus decreasing Fur DNA binding activity – the model dubbed *apo*-Fur. This will allow for activation of genes that can help in compensating for decreased intracellular iron concentration, and or other pathways related to iron functionality (*apo*-Fur activation). The opposite regulatory pattern, *apo*-Fur repression, has been hypothesized to exist but has not been well documented[37].

The Fur regulon – the genes Fur directly targets – have been studied extensively across various non-pathogenic, laboratory *Escherichia coli* strains[1,3,17,37]. This was made evident through a variety of biochemical assays, whether it be mutational analyses; reporter assays; genome-wide chromatin immunoprecipitation (ChIP) followed by high-throughput sequencing (ChIP-seq); and RNA expression assays[37,38]. As alluded to, many raw siderophore genes are

repressed by Fur-Fe^{2+} binding to prevent excess iron importation, though activated in naked Fur unbinding for iron supplementation. Aerobactin (*iucABCD*), Enterobactin (*entD, entF, entCEBA*), and Enterochelin are some of the most prominent examples - the latter two siderophores procured through non-ribosomal peptide synthesis. Siderophore synthesis is naturally tied into this mixture (*fes, fepA*, *fepB*, and *fepDGC*). Non-siderophore iron transporters such as cell membrane proteins or siderophore-esque complexes are managed closely as well (*fhuABCDE*, *fecABCDE*, *feoABC*, *tonB*, *exbBD*, *cir*, *fiu*). Fur can still extend its immediate action beyond iron metabolism to activity that tries to incorporate iron into molecular structures. FeS assembly (*sufABCDES*), Superoxide Dismutase (*sodB*), and electron transport chain subunits (*sdhDC*) are all critical framework within cells. Fur can even impact metabolism of other non-transition metal nutrients - arabinose (*araC*, *araBAD*), fumarate (*fumA*), and ribonucleotide to deoxynucleotide conversion (*nrdHIEF*). Many of these genes are frequent inclusions in *Escherichia coli* Fur regulons and speak to Fur's diverse evolutionary roles.

For commonly studied *E. coli* pathogenic strains and or distantly related genera, Fur has even been shown to play a large role in the management of bacterial pathogenesis genes. In *Pseudomonas aeruginosa*, Fur regulates the additional siderophores pyoverdine and pyochelin to combat nutritional immunity mounted by the bodies of cystic fibrosis patients[7]. In *Salmonella Typhimurium*, Fur is responsible for mounting the acid response system when encountering features of the host intestinal tract[7]. In the common uropathogen, *E. coli* strain CFT073, the Fur regulon oversees additional genes pertaining to amino acid usage in order to overcome amino acid-related stressors in the iron-limiting environment of the urinary tract[6]. Elucidating Fur in its entirety in this sense – the various mechanisms it uses to regulate gene expression and the regulon it operates through– can aid in better understanding many diseases common to the

human population. Though research is evidently underway to probe Fur in familiar systems, Fur has not been extensively characterized in many non-model strains. The latter may offer additional information regarding the Fur regulon and the mechanisms of Fur, owing to their upbringing in more diverse environments and conditions.

In this study, we characterize the roles of Fur in two clinical, non-model *E. coli* strains, No. 131 and No. 158, isolated from bacteremia patients relative to two model strains, CFT073 and MG1655. We integrate the direct-monitoring of protein-to-DNA interactions provided by ChIP-exo with gene expression profiling via RNA-seq to determine which genes are directly and or indirectly regulated by Fur. We also systematically evaluate the impact of Fur on bacterial phenotypes including bacterial growth and antibiotic sensitivity. In support of our initial beliefs, we discover that, though the Fur regulon between these two strains share regions of conservation, they also boast regions of exceptionality that highlight the growing profile of Fur's capabilities. The strain's Fur binding motif and methods of regulation are also fairly stagnant. Furthermore, an absence of Fur does still hinder bacterial growth and does still induce the rise and fall of antibiotic resistances. Nonetheless, the particular patterns associated with both vary. From this, we propose that Fur retains its importance across the *Escherichia* genus, though may accommodate additional genes and or mediate certain evolutionary changes that better suit a particular strain and its origin.

CHAPTER 1

E. COLI 131 AND E. COLI 158 ARE RELATED TO CFT073 AND DEMONSTRATE PHENOTYPIC VULERNABILITY IN THE ABSENCE OF FUR

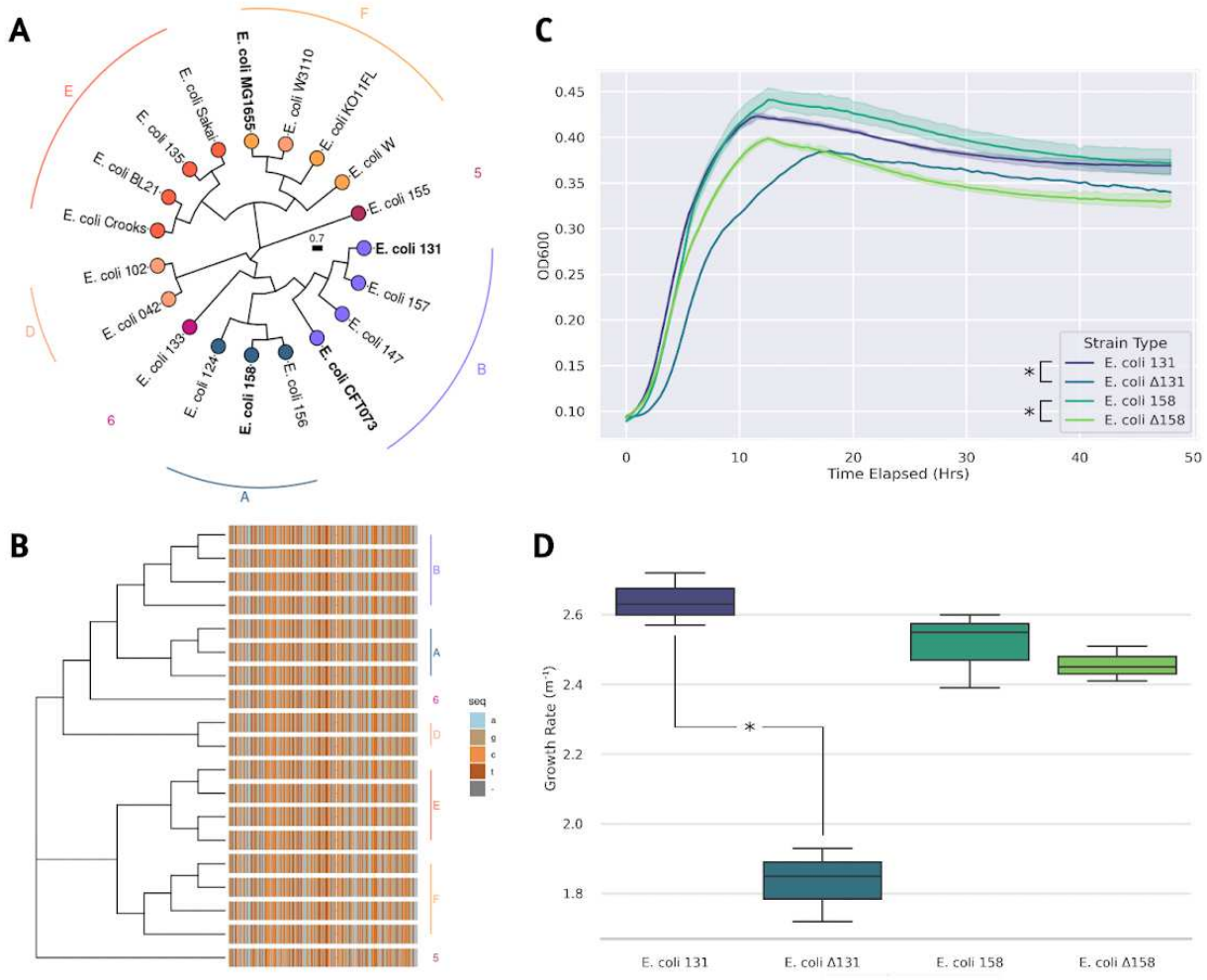

Figure 1 – Phylogenetic Comparison and Growth Analysis of E. coli 131 and E. coli 158. (**A**) A phylogenetic tree based on *rpoB* sequences showcasing the associations between a broad pool of *E. coli* strains. Major clades are labelled over the branches of the tree. Bolded names indicate strains selected for the remainder of the study. (**B**) Visualization of sequence alignment to the *rpoB* gene used to map relatedness. (**C**) OD600 for *E. coli* 131 and *E. coli* 158 across 48 hours of incubation at 37°C in M9 media. (**D**) Calculated growth rates for *E. coli* 131 and *E. coli* 158 and their respective *fur* mutants.

To assess the diversity of the two clinical strains *E. coli* 131 and *E. coli* 158, their genomes were compared to eight other clinical strains from the same batch of isolated samples (*E. coli* 102; 124; 133; 135; 147; 155; 156; 157) and nine representative model strains (*E. coli* MG1655; W3110; KO11FL; W; BL21; Crooks; CFT073; 042; Sakai). The single-copy gene *rpoB* was used as a molecular marker to compare relatedness. It was selected due to it being as efficacious as other common molecular markers like 16S rRNA gene compounded with the fact it boasts higher resolution of closely related species[8]. The resulting major phylogenetic groups are labelled A, B, D, E, F, 5, and 6 for reference (**Fig 1A**). Virulent groups are primarily localized to groups A, D, and E. A multi-sequence alignment of each strain's individual *rpoB* gene is presented below the tree as an alternative means to visualize the comparison (**Fig 1B**).

We find that *E. coli* 131 is situated in the same clade as the model UPEC *E. coli* strain CFT073. On the other hand, *E. coli* 158 is an outgroup to CFT073, though more closely related to this prominent uropathogen versus non-pathogenic strains such as the K-12 strain, MG1655 (**Fig 1A**). These associations hint towards potential trends in the Fur regulon and Fur-related phenotypes for these strains – namely that they might be expected to resemble CFT073 and or other uropathogenic *E. coli* more so than MG1655 and or other non-pathogenic *E. coli*. Consequently, from this data: alongside *E. coli* 131 and *E. coli* 158, CFT073 and MG1655 were selected as model references from which diversity of the former two could be compared to (**Table 1**).

Table 1 – List of all bacterial strains, their relevant genotypes, and their source as to be used over the course of this study.

Strain	Relevant Marker(s)/ Genotype	Source or Reference
E. coli MG1655	WT, Δ*fur*	42
E. coli CFT073	WT, Δ*fur*	42
E. coli 131	WT, Δ*fur*	This study
E. coli 158	WT, Δ*fur*	This study

To then determine the effect of Fur on growth of *E. coli* 131 and *E. coli* 158, we ran Growth Curve analysis on the two strains and their corresponding *fur* mutants. Triplicates per M9 minimal media were grown in 96-well microtiter plates at 37°C. Their growth was monitored through a BioTek LogPhase 600 microbiology reader (Agilent) every 30 minutes for 48 hours. Growth curves were visualized using Matplotlib. Growth rates per individual well were determined via the R package "Growthcurver" before being averaged based on strain type and plotted also through Matplotlib.

We found that deletion of *fur*, and consequently disruption of iron homeostasis, does affect the growth of both non-model strains, albeit with some nuances. The maximum OD both *E. coli* 131 and *E. coli* 158 *fur* mutants achieve by the time of the stationary phase - that they also maintain throughout the stationary phase - is significantly lower than that of their wild-type counterparts (**Fig 1C**). The impact on growth here is further exemplified for one of the strains by its growth rate – *E. coli* 131 *fur* mutant grew significantly slower than its wild-type counterpart (**Fig 1D**). These results are validated by the fact that iron is known to play a key role in the management of genes tied to *E. coli* growth[3,16]. A lack of Fur thereof leads to a lack of adequate growth controls and sustainability, resulting in the phenotypes visualized here.

Curious observations incidentally present here include the distinctions between the growth profiles of *E. coli* 131 and *E. coli* 158 – also, as mentioned earlier, the fact that the growth rate of *E. coli* 158 was not significantly altered. Firstly, *E. coli* 158 achieves a higher maximum OD both in wild-type and in *fur* mutants versus *E. coli* 131. This is speculated to be the consequence of the individuality of Fur in the two strains. As discovered later in **Chapter 2** and **Chapter 3**, *E. coli* 131 Fur and *E. coli* 158 Fur target a different range of genes within their respective regulons despite some more broad patterns of conservation. A subset of these genes do contribute more intensely to bacterial physiology in *E. coli* 131 versus its relative (e.g., *flu_1* and *flu*_2 for biofilm formation; *aroH_1* for amino acid metabolism). Thus, knocking out Fur in the *E.* coli 131predisposes it to worse growth outcomes versus *E. coli* 158 which may not rely on Fur nearly as much for survivability and proliferation.

Secondly, *E. coli* 158 demonstrates no significant change to its growth rate post-Fur knockout whereas *E. coli* 131 does. This is doubly unique versus the model strains MG1655 and CFT073, both of which are in line with the latter's display of effort[6]. A plausible mechanism to explain this discrepancy may involve Fur's role as a mediator of amino acid-dependent acid resistance systems and metabolism for uropathogens. Recall that we established *E. coli* 158 as being less related to *E. coli* CFT073 versus *E. coli* 131. Acetic acid build up and amino acid metabolism may be at fault. In an environment where multiple cells are undergoing glucose metabolism and producing harmful organic acids, Fur can decrease acetic acid concentration and mediate detoxification of related molecules (i.e., proton accumulation)[15]. A lack of Fur thereof would result in an increasing inability to deal with these threats, leading to a greater chance of decay as seen. Furthermore, Fur manages the existing amino acid pool for downstream protein synthesis. Inactivation of fur leads to the induction of the RpoS stress system, and consequently

degradation of cells[26]. *E. coli* 158, per **Chapter 3**, is not nearly as reliant on these pathways. It may be predicted that the strain may have these same prominent growth factors of *E. coli* 131 under the control of other transcription factors not labelled Fur. In summary, our results here demonstrate that *E. coli* 131 and *E. coli* 158 are closely related to the UPEC strain CFT073 – furthermore, they exhibit inhibited growth patterns the absence of the Fur transcription factor.

CHAPTER 2

PREDICTABLE CONSERVATION PATTERNS OCCUR WITHIN THE CORE GENOME

ACROSS NON-MODEL AND MODEL STRAINS

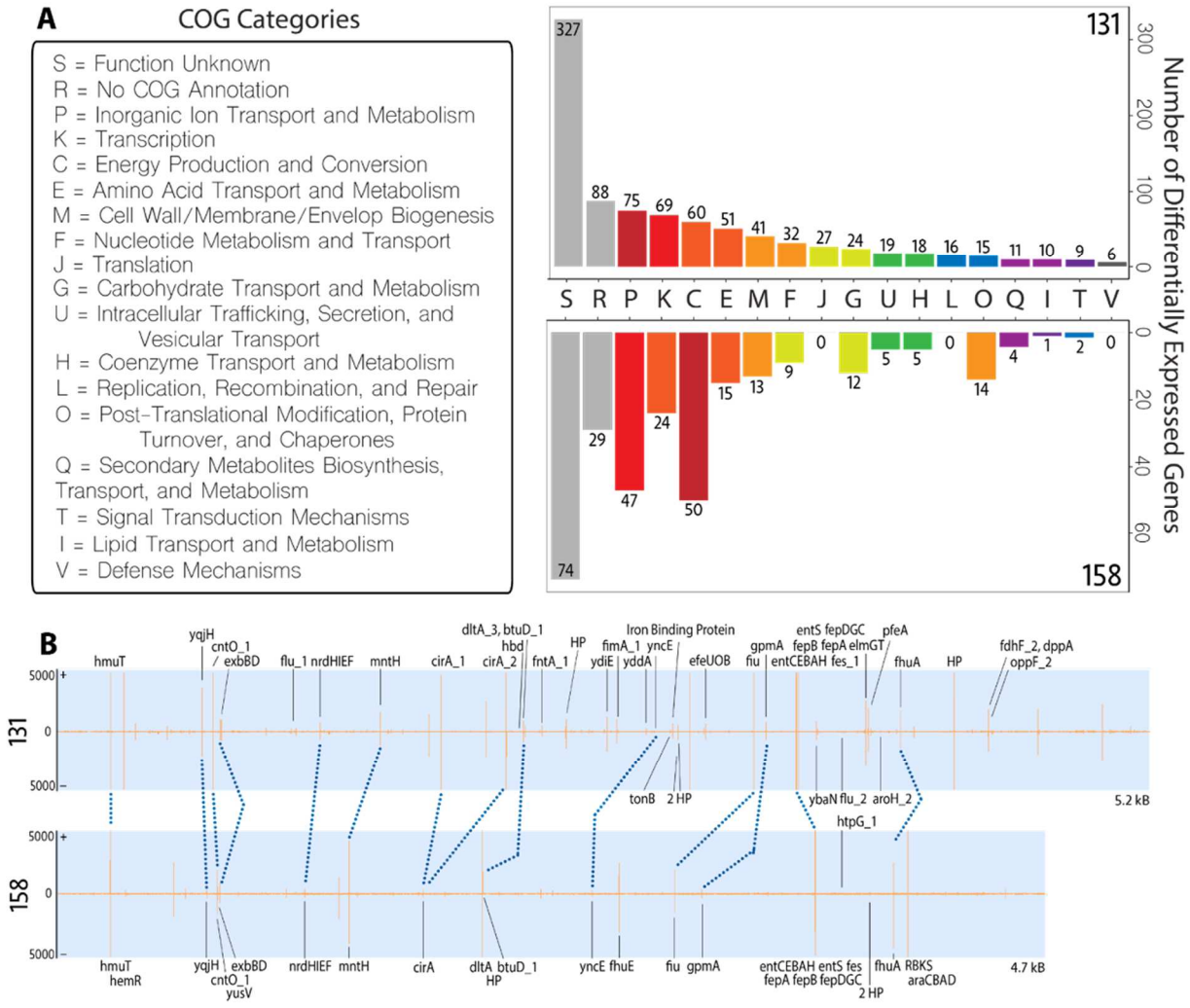

Figure 2 – ChIP-exo and RNA-seq profiling of E. coli 131 Fur and E. coli 158 Fur. (**A**) Clusters of Orthologous Genes [COG] statistics for differentially expressed genes of *E. coli* 131 and *E. coli* 158 under iron-replete conditions. Intensity of a bar's color [e.g., dark red is the most intense; purple is the least intense] except for "Function Unknown" and "No COG Annotation" [colored gray] correlates to the quantity of genes per group. (**B**) Map of genome-wide ChIP-exo binding sites for *E. coli* 131 and *E. coli* 158 under iron-replete conditions. Binding sites that occurred for both strains are correlated with one another via a dark blue dotted line between rows.

To determine the genes Fur both directly and indirectly associates with, we employed a comprehensive RNA-seq procedure to discern differentially expressed genes under iron-replete (iron-rich) and iron-deplete (iron-deficit) conditions. We first isolated the total RNA per strain and purified it exclusively for native mRNA. Libraries were assembled and amplified via real-time qPCR before submission to Illumina Sequencing Services. Results were put through the DESeq "R" package to assess differentially expressed genes - the output was labelled by eggNOG for Clusters of Orthologous Genes (COG) as to categorize gene groups[19]. A total of 1586 genes were differentially expressed with log2 fold change > 1 and false discovery rate (FDR) <0.01 in *E. coli* 131 – 886 for iron replete, and 700 for iron deplete. Meanwhile, a total of 1125 genes were differentially expressed with the same statistics in *E. coli* 158 – 299 for iron replete, and 826 for iron deplete (**Supplementary Data 1 and 2**). The number of genes overlapped between both conditions per strain were 128 and 70, respectively.

The three most prominent COG categories for genes regulated directly or indirectly by Fur under iron-replete conditions – this excluding the broad "Function Unknown (S)" and "No COG Annotation (R)" types – were Energy Production and Conversion (C); Inorganic Ion Transport and Metabolism (P); and Transcription (K). *E. coli* 131 Fur oversaw at least one gene localized to the remainder COG categories, though *E. coli* 158 Fur lacked regulation of any genes for Translation (J); Replication, Recombination, and Repair (L); and Defense Mechanisms (V) (**Fig 2A**). These overarching trends adequately speak to Fur's fundamental nature. Fur has been distinguished as a transcription factor majorly responsible for processes tied to the inorganic ion, iron. Naturally, this would include iron metabolism, iron uptake, and transcriptional functions. Nonetheless, Fur has been discovered to play a more involved role in non-iron metabolism-related processes[38, 41]. The diversity of COG categories tied to the DEGs of

these non-model strains affirms Fur's complexity in different biological systems despite our more strict, traditional definition of the popular transcription factor. Any minor inter-strain variance in COG distribution is likely a result of the evolutionary positioning of *E. coli* 131 and *E. coli* 158 relative to one another and to other strains – recall that the former is much more closely related to *E. coli* CFT073 than the latter.

To determine the genes and or operons Fur directly binds to in its regulatory regime, we employed ChIP-exo to map Fur binding regions across the genome in varying concentrations of iron. The Fur regulon was distinguished by correlating sequencing results with differentially expressed genes (DEG) discerned in RNA-seq experiments. Any genes possessing both a binding event and a DEG profile were marked as part of the Fur regulon for either iron concentration condition. Including those genes within an operon, *E. coli* 131 yielded a total of 57 genes directly regulated by Fur while *E. coli* 158 yielded a total of 42 genes (**Fig 3A**). 31 of the genes from both strains overlapped concurrently with the regulon of MG1655 and CFT073 – this dubbed the "core genome". Some of the remainder genes overlapped with at least one other strain while some genes were unique to either *E. coli* 131 or *E. coli* 158 – these dubbed the "accessory genome" and "unique genome", respectively (**Supplementary Data 1**)[6].

Fur-regulated genes of the core genome include iron homeostasis and acquisition functions – namely synthesis of the siderophore "enterobactin" (*entCEBAH, entD, entS, fes*) and siderophore-mediated ferric ion uptake (e.g., *cirA, exbBD, fepA, fepB, fepDGC, fiu, fhuA, yqjH*). In addition, the core genome includes ion transport (*mntH*), metabolic functions (*gpmA, nrdHIEF*), and proteins of unknown function and identity (Hypothetical Proteins, yncE) (**Fig. 2B, 3B, Supplementary Data 3**). The multitude of siderophore-related genes targeted by Fur is understandable. We have stated already how integral iron is towards biological systems -

assuring the transition metal can be acquired in reasonable amounts is a fundamental pillar in survivability[3,13,41]. Furthermore, for UPEC strains in particular such as CFT073 and the supposedly related *E. coli* 131 and *E. coli* 158, pathogenesis typically occurs in the extracellular environment of the urinary tract where iron is a limited resource[11]. Enterobactin biogenesis, managed through expression of *entCEBAH, entD, entS,* and *fes,* is consequently upregulated to compete with the host and its innate flora while in the process of invading and colonization[37]. Genes that sport regulatory hold on enterobactin biosynthesis or transport and or iron or iron-complex transport systems – *fhuE*, *feoABC*, and *fiu* as listed – follow suit[6,13,37]. The TonB-dependent energy transduction and receptor system, mediated by *tonB*, *exbBD*, and *fepA,* is, in a sense, interconnected to all these siderophore processes. It funnels the proton motive force into iron-siderophore and vitamin B12 transporters, thus aiding in the movement and deployment of siderophores from individual cells[18].

Extending beyond strictly iron transport, the *nrdHIEF* operon is typically Fur-associated and will be expressed in increasing amount during periods of ROS stress to minimize cellular damage[27]. Existing as a set of ribonucleotide reductases, this set of genes may collect ROS species and quench them through redox reactions in addition to their established role as converters of ribonucleotides to deoxyribonucleotides[30]. The gene *mntH* is called upon on similar circumstance – as a manganese importer, it is theorized to replace iron with manganese in otherwise iron-requiring cofactors when necessary due to the latter being more resistant to complexion with H_2O_2[2]. The activity of both the previous operon and this gene, again, are in line with Fur's common response to iron-related stressors. Next, *gpmA* is not an ROS-related gene – it encodes an isoenzyme form of phosphoglycerate mutase. Its activity has been documented in prior literature when concerning the *E. coli* Fur regulon – this extends to analogous pathogenic

species including the plant pathogen *Erwinia chrysanthemi*[6,13]. Finally, the obscurity behind the remaining Hypothetical proteins and *ycnE* leaves little to be said about their known role across the four *E. coli* strains. Nonetheless, it may be predicted that they are all crucial to some pillar of proliferation or survival per each. In conclusion, both the DEG profile and ChIP-exp binding profile support widespread conservation of the Fur regulon in these non-model strains compared to model strains.

CHAPTER 3

THE ACCESSORY GENOME SHOWCASES NUANCES IN THE FUNCTIONALITY OF FUR ACROSS NON-MODEL STRAINS AND MODEL STRAINS

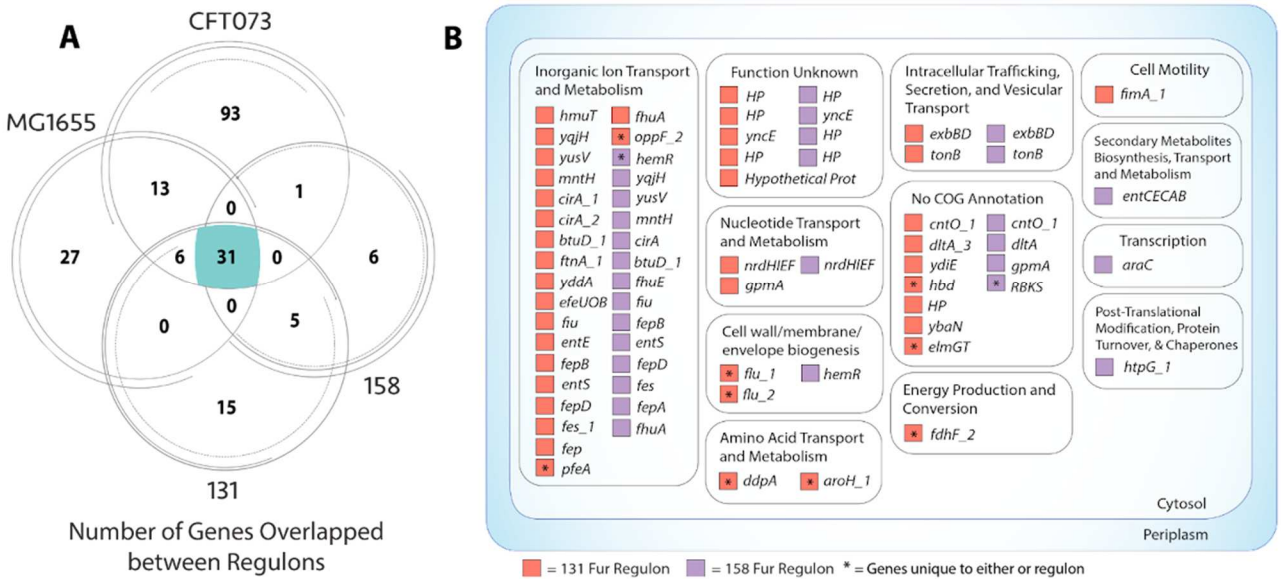

Figure 3 - The Fur Regulon of E. coli 131 and E. coli 158. (**A**) Quantitative comparison of the Fur regulon between *E. coli* 131, *E. coli* 158, MG1655, and CFT073. (**B**) COG categories for all genes within the 131 Fur regulon and the 158 Fur regulon. "HP" and "Hypothetical Prot" are abbreviations for "Hypothetical Protein".

Exclusive to the *E. coli* 131, CFT073, and MG1655 Fur regulons – the accessory genome of *E. coli* 131 versus *E. coli* 158 – were Fur-regulated genes of iron uptake (*efeUOB*), ferritin production (*ftnA_1*), ion transport (*yddA*), and more proteins of unknown function (*ydiE, ybaN*). *E. coli* 158 did not have any Fur-regulated genes shared with both CFT073 and MG1655; however, it did share a lone gene with CFT073 in the form of transcription of the L-arabinose operon (*araCBAD*). Exclusive between the *E. coli* 131 and *E. coli* 158 Fur regulons – the accessory genome relative to the two – were genes of heme metabolism (*hmuT*), ion transport

(*cntO_1, btuD_1*), lipoteichoic acid biosynthesis (*dltA/ dltA_3*), and genes of unknown function (*yusV*) (**Fig. 2B, 3D, Supplementary Data 5**).

Excluding the two proteins of unknown function, the other five genes present in the accessory genome between *E. coli* 131, CFT073, and MG1655 fall within a similar line of explanation to that of iron-related genes in the core genome. Namely: the functionality of these genes and their presence in one or more *E. coli* Fur regulon have already been documented. These genes serve to take up iron and integrate them into the metabolic processes of *E.* coli, though are prevented in excess to stop ROS production. The lack of these same genes in the Fur regulon of *E. coli* 158 does devalue this interpretation, nonetheless. To speculate beyond what context is provided in this study, it is possible that *E. coli* 158 still regulates these same genes through Fur, albeit indirectly. It is also plausible this strain has other regulatory systems in place to manage these genes and or supplements iron with distantly related metal ions. The presence of an *araCBAD* binding site in *E. coli* 158, mirroring that of CFT073, is also worth noting. It is possible *E. coli* 158 has additional regulation of L-arabinose metabolism either to supplement a greater demand for carbon in general and or this particular carbon source – e.g., energy production or nucleotide production – or to counteract stressors given arabinose's downstream involvement with the Pentose Phosphate Pathway and NADPH reduction[35].

Shifting gears now to the accessory genome of *E. coli* 131 and *E. coli* 158 bar the ambiguous *yusV* gene– this data speaks to the differentiation of these strains as non-model strains versus the two ascribed model strains. The presence of *hmuT* in the *E. coli* 131 and *E. coli 158* Fur regulons is not too far-fetched. Hemin is an alternative heme compound that integrates iron within its own structure. The presence of a binding site for the associated gene indicates that cell-bound transport systems allowing utilization of host iron and heme compounds may play a

significant role in the pathogenesis of these two strains. *cntO_1* is a documented metallophone pseudopaline key towards acquisition of nickel and zinc. It would be predicted to fall in line with the ascribed role of *hmuT* and with aforementioned genes in the core genome. *btuD_1*, being a Vitamin B12 import ATP-binding protein, likely assists in mediating energy production processes like *gpmA* in the core regulon. Finally, *dltA* and its involvement in lipoteichoic acid biosynthesis may signal the critical nature of cell wall synthesis to the integrity of and or later pathogenesis of the two strains. All in all, these results indicate that portions of the Fur regulon to the non-model strains *E. coli* 131 and *E. coli* 158 are tailored in a similar way to model strains to serve specific functions, and or speak to the relevant distinctions between a non-model and model strain.

At this point, we have evaluated the core genome and accessory genome of the four strains centered as primary subjects of this study. To now dive into the unique genome: the genes *hbd, flu_1, flu_2*, *oppF_2*, *aroH_1*, *dppA, elmGT*, *pfeA,* and *fdhF_2* were unique to the *E. coli* 131 Fur regulon (**Fig. 2B, 3B, Supplementary Data 3**). As to provide commentary on each gene individually: *hbd* encodes the oxidoreductase 3-hydroxybutyryl-CoA dehydrogenase, an enzyme tied to fatty acid metabolism. Most reported literature features the gene in context of humans – it is hard to gauge the reason for its homologous presence in the regulon of *E. coli* 131. It could be possible that regulation of *hbd* allows for modulation of lipid synthesis for the production of lipopeptides involved downstream pathogenesis pathway – nonetheless, this is a largely vague prediction. *flu_1* and *flu_2*, on the other hand, correspond to Antigen 43, a crucial component to auto-aggregation in *E. coli* cells. This gene and other genes related to biofilm formation and surface adhesion were discovered to be downregulated by a similar iron-dependent system, IscR, under iron-deficient conditions in MG1655[44]. It is entirely possible that Fur functions in a similar

manner. Per its description as a repressor, Fur may disfavor *flu* expression when iron is lacking in order to promote movement of cells to a more iron-plentiful environment.

oppF_2 is a part of the *Opp* operon responsible for the transfer of oligopeptides in bacteria. Given the wide range and functionality of peptides, it is hard to pinpoint a particular reason to why Fur targets this gene. It might be reasoned that the *E. coli* 131 regulon includes *oppF_2* as an extension of its regulatory control over amino acid metabolism genes per the requirements of UPEC *E. coli* versus other strains[6]. *aroH_*1 encodes Phospho-2-dehydro-3-deoxyheptonate aldolase, an enzyme key to the synthesis of aromatic amino acid precursor, chorismic acid. Again, given the importance of amino acid collection to UPEC *E. coli*, it is possible Fur upregulates this gene in order to ensure corrective amino acid production. *ddpA* functions as a periplasmic di-peptide transport protein in all four studied strains. Its regulation in 131 may, once more, be a homologous coincident with the related CFT073's dependence on amino acid metabolism and uptake for pathogenicity[6]. *elmGT* occurs exclusively between *E. coli* 131 and CFT073 (not MG1655) though is only directly regulated in *E. coli* 131. Encoding an Elloramycin glycosyltransferase, this gene has been implicated in the biosynthesis of the antitumor drug Elloramycin and in the recognition of deoxysugars for *Streptomyces olivaceus*[32]. It may follow that the gene functions similarly in our strain - Fur may regulate it in order to consequently dictate the uptake of deoxysugars for DNA synthesis or repair. Alternatively, Elloramycin production, if discovered in this *E. coli*, may serve as a safety net against harmful progenitors and or excrements of tumors.

fdhF_2 is a formate dehydrogenase that, like others of its kind, helps pass electrons to existing quinone pools or other proteins that later require reduction. In essence, the gene bears ties to cell energy conservation and ATP generation via ATPase[42]. Along the lines of *gpmA* from

the established core genome and of *btuD_1* in the accessory genome between 131, CFT073, and MG1655, this *fdhF_2* binding event can be interpreted as an expansion of Fur's role in energy management in bacteria. *pfeA* is noted to encode a ferric enterobactin receptor in *P. aeruginosa* – its normal homolog in *E. coli* is *fepA*[11,12]. Assuming this binding event should be treated separately to the already documented *fepA* binding sites in *E. coli* 131, it is possible 131 retains this gene separate to its co-occurring homolog as an additional method to gain iron – functional redundancy. If so, *E. coli* 131 may be more dependent on access to iron for proliferation versus its more common counterpart. This may tie into its identity as a relative to CFT073, as constantly restated at this point – the fact that urinary tract infection can grow dependent on iron concentration.

Excluding Hypothetical Proteins, the genes *hemR* and *RBKS* were the only genes seemingly unique to the *E. coli* 158 Fur regulon despite the fact they were still present in the genome of one or more of the other surveyed strains. *hemR* may not stray too far from *hmuT* in terms of its general purpose – as a proponent to heme metabolism, it likely helps the strain accrue additional iron as does *hmuT* in *E. coli* 131. *RBKS* is described as Ribokinase by Prokka annotation. As such, it may more commonly refer to the *rbsK* gene, Ribokinase, known to help catalyze the conversion of ribose to ribose-5-phosphate during the Pentose Phosphate Pathway (PPP)[39]. This gene has previously been linked to stress-induced mutagenesis, though the mechanism of how it actually participates in this process is unclear[24]. We propose two possible hypotheses arise from this knowledge: (1) *E. coli* 158 Fur may play a role in modulating the activity of the PPP reaction due to PPP's positive sensitivity to the Fe(II) co-substrate[21]; (2) Fur may modulate Ribokinase activity as to manage mutagenesis pathways similarly to how it modulates ROS stress response. In summary, our results here demonstrate that despite the

conservation of certain elements to the *E. coli* 131 and *E. coli* 158 Fur regulons, both strains possess a special set of genes that widens the boundaries to our understanding of Fur targetability.

CHAPTER 4

FUR BINDING MOTIF STRUCTURE AND REGULATORY METHODS ARE CONSERVED BETWEEN MODEL AND NON-MODEL STRAINS

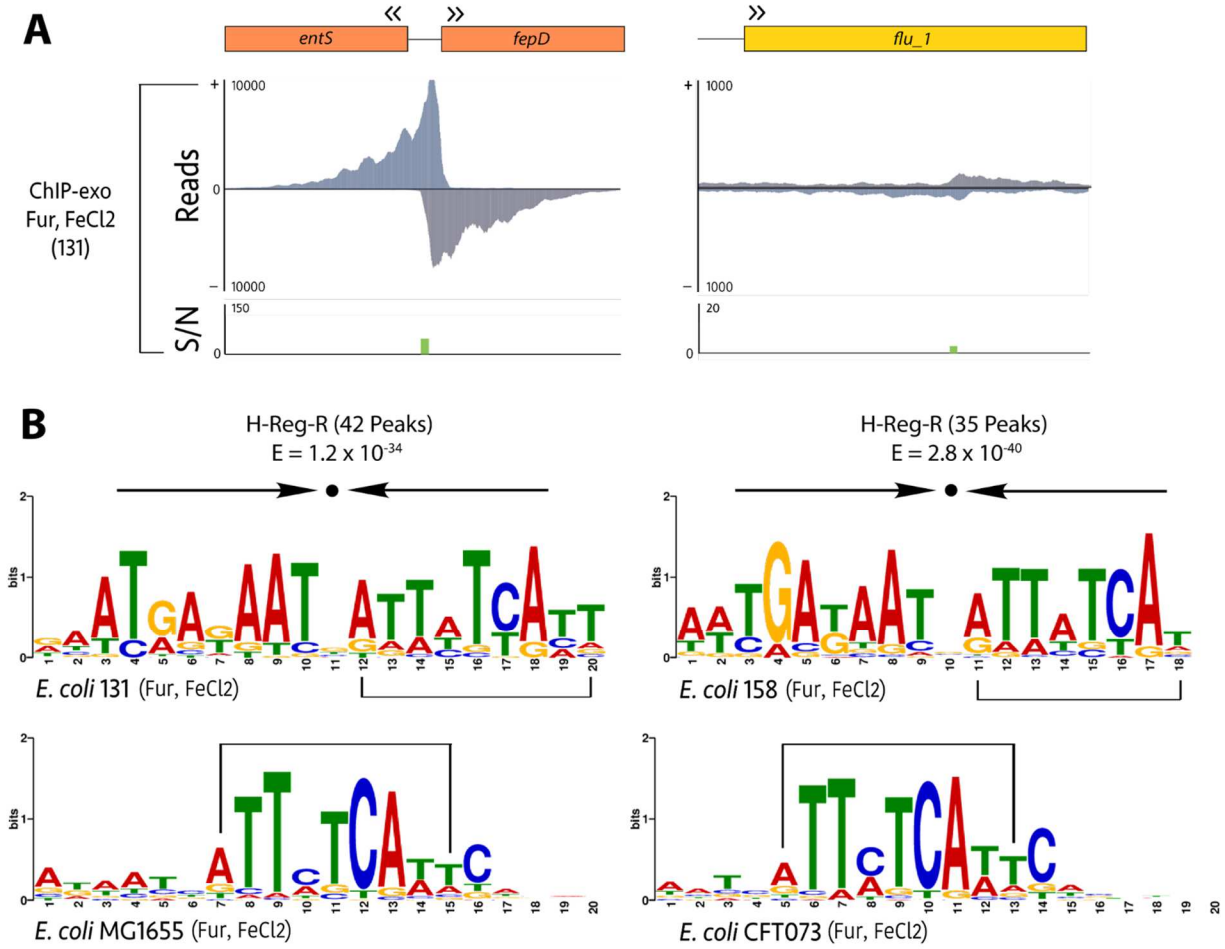

Figure 4 – Regulatory Modes in E. coli 131 and Fur Binding Motifs of E. coli 131 and E. coli 158. Reads, after being processed through Bowtie2, were mapped on Metascope. Signal to Noise ratio was determined via the peak-calling algorithm MACE. Two examples of regulatory modes of Fur are presented, including control of a divergent promoter (**A**, left) in which Fur regulates the two genes adjacent to it; and a singular promoter (**A**, right) in which Fur regulates a gene downstream to its binding site. (**B**) The average Fur binding motif for *E. coli* 131, *E. coli* 158, *E. coli* MG1655, and *E. coli* CFT073. The number of peaks the average motif was mapped to is noted above each MEME output for only the former two. Arrows showcase the consensus sequence - regions with resemblance are highlighted with black bars.

As critically important as the Fur regulon is towards understanding how Fur operates as a transcription factor, being able to uncover Fur's regulatory modes is equally vital. Recall that Fur is proven and or speculated to provide different transcriptional responses based on the availability of iron. These include HR, HA, *apo*-Fur activation, and *apo*-Fur repression[38]. Also, recall that Fur's binding motif in a strain can reveal how it physically attaches to the genome and how tightly it might associate with certain genes versus others. This is telling of how important Fur regulation of certain target genes can be versus other target genes for the same strains and or between strains. A plethora of unique sequences to the Fur binding motif have been reported, yet the most common is an 18 to 22 base pair palindromic sequence linked to HR[22,41]. To evaluate the regulatory modes and Fur binding motifs of *E. coli* 131 and *E. coli* 158, we employed Metascope to map out and curate our ChIP-exo sequencing reads. We followed this up with the MEME suite to determine the average binding motif across all peaks extracted. The resulting information was compared to existing binding motifs encountered for MG1655 and CFT073.

Both *E. coli* 131 and *E. coli* 158 demonstrate palindromic binding motifs (**Fig 4B**). The size of the binding motif per strain differs – the Fur binding motif of *E. coli* 131 is two bases larger than that of *E. coli* 158, 10-1-9 versus 9-1-8. Nonetheless, the internal sequence of the motifs bar the outside present or absent two bases is near identical with few base substitutions. Furthermore, the reported motifs were linked to all binding sites attributed to those genes reported of each strain's respective Fur regulon. The nature of these motifs is characteristic of canonic *holo*-Fur repression and has previously been reported in literature regarding K-12 model strains[23,38,41]. In essence, both halves to the motif are strongly palindromic, though one may alternatively interpret the bases as a set of hexamer sequence per a hexamer model[23]. Comparing the average binding motifs of MG1655 and CFT073 to those of *E. coli* 131 and *E. coli* 158 - the

former bears resemblance to the latter. Though the motifs of MG1655 and CFT073 are not perceivably palindromic in entirety for our case, this is speculated to be a result of the two strains having a larger quantity of peaks. This greater number of peaks requires MEME to average across more samples, inviting unique motifs the clinical strains do not possess but the model strains evidently do possess. After all, unlike MG1655 and CFT073, no motifs for any individual gene in *E. coli* 131 and *E. coli* 158 corresponded to *holo*-Fur activation, *apo*-Fur activation, and or *apo*-Fur repression. This is true even with the presence of a previously predicted *apo*-Fur activated gene, *ftnA_1* (previously reported in MG1655), in the *E. coli* 131 regulon[38]. MEME would thus only report a singular motif across all genes in the *E. coli* 131 and *E. coli* 158 regulons. Overall, these results demonstrate the highly conserved nature of Fur binding even if the Fur regulon diverges between strains.

Building upon the concept of conservation evident through binding motifs, *E. coli* 131 and *E. coli* 158 both show instances of two distinct but common Fur regulatory modes across their regulon. Fur, like most transcription factors, can act upon either a divergent promoter – it regulates two genes and or operons adjacent to binding location in opposite directions – or a singular promoter – it regulates one gene by binding the gene upstream. An example of said divergent promoter is exemplified by Fur binding downstream and upstream of *entS* and *fepD*, respectively, in *E. coli* 131; a single promoter is exemplified by Fur binding *flu*_1 (**Fig 4A**). Though not shown in the figure, similar events occur for select genes in *E. coli* 158. Again, this speaks to the conserved nature of overall regulator methods by Fur. Even in non-model strains that may be expected to have stemmed diverse conditions of growth and evolution, Fur abides by similar mechanics to well-studied and controlled laboratory strains. In conclusion, we discover

that the binding motif and regulatory modes of Fur in *E. coli* 131 and *E. coli* 158 are conserved relative to model strains.

CHAPTER 5

THE ABSENCE OF FUR PROMOTES ANTIBIOTIC RESISTANCE IN E. COLI 131 AND E. COLI 158

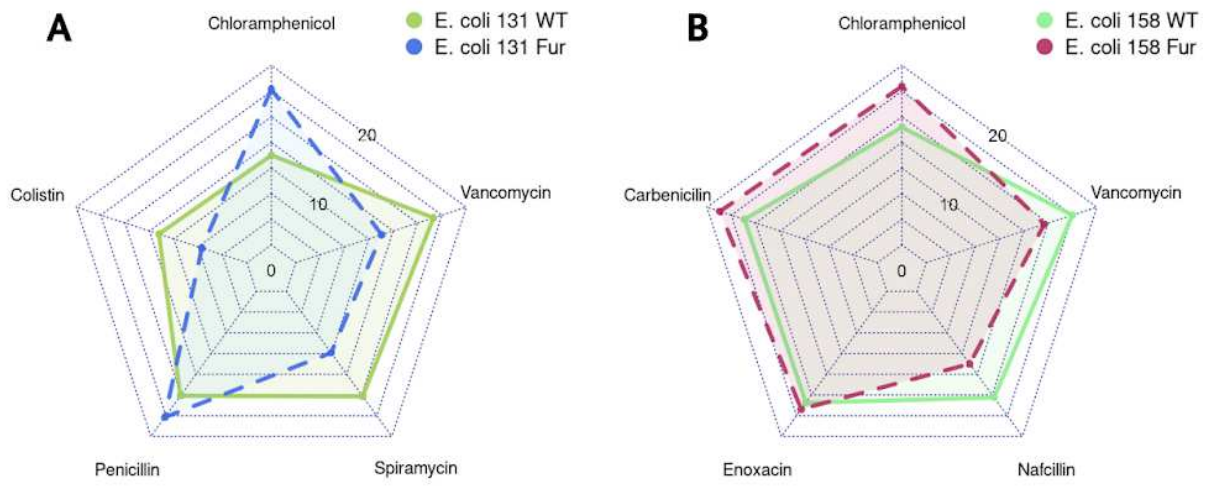

Figure 5 – Fur knockout increases resistance to some antibiotics and increases vulnerability to others in E. coli 131 and E. coli 158. "Simulated OD" is being shown - this calculated via the quotient of the experimental average Y-value (via Omnilog) over the average Y-value of a blank well. All antibiotics shown above were antibiotics the two strains were shown to have had a significant change in resistance to.

Previous literature had noted the potential for *E. coli* to develop ciprofloxacin antibiotic resistance when iron homeostasis was perturbed[26]. To see if this trend and or other antibiotic resistance developments might carry over to the non-model strains of focus in this study, we performed a Phenotype MicroArray Assay through the Biolog Omnilog system. 1 uL aliquots of cells grown in M9 media seeding cultures were washed and inoculated into 96-well microtiter plates pre-configured with varying concentrations of unique antibiotics – PM-11 and PM-12. These plates were placed inside the Omnilog instrument and monitored over 48 hours every 15 minutes for growth patterns. A Biolog- associated data analysis program "Data Analysis" was

used to derive trends of resistance per growth patterns of the *E. coli.* Results were visualized on Matplotlib. Our results indicate that *E. coli* 131 and *E. coli* 158 both develop and loosen resistance to a handful of antibiotics. Among those most prominent for *E. coli* 131, this strain tends to become more resistant to Chloramphenicol and Penicillin G while growing more vulnerable to Vancomycin, Spiramycin, and Colistin (**Fig 5A**). For *E. coli* 158, this strain tends to become slightly more resistant to Carbenicillin, Chloramphenicol, and Enoxacin – more vulnerable Nafcillin and Vancomycin (**Fig 5B**). The remainder of antibiotics surveyed did not yield major differences and are thus not included.

It is well-known that a lack of Fur protein can decrease iron-storage proteins and activate iron-uptake proteins. Being unable to adequately manage this transition metal can cultivate a high intracellular concentration of reactive ferrous iron[1]. Excessive iron can accelerate the formation of ROS species which can subsequently damage the existing nucleotide pool and DNA in a random fashion to promote mutations[3,40]. Antibiotics themselves may exacerbate this through perturbation of the Citric Acid Cycle, leading to a more rapid rate of occurrence for the damaging Fenton and Haber-Weiss reactions in line with ROS production[22]. Alternatively, downregulation and or inactivation of ROS-processing genes such as the superoxide mutase are associated with greater mutation rates. We discovered in **Chapter 2** and **Chapter 3** that, in *E. coli* 131 and *E. coli* 158, Fur directly regulates ferritin - also, siderophores like enterobactin and siderophore-related genes. Anti-ROS genes are indirectly regulated by Fur per RNA-seq results. Given this information in tandem with our own data, an ROS-mediated mutagenesis may be at fault for what is observed in *E. coli* 131 and *E. coli* 158 when concerning the development of antibiotic resistance. These mutations are a double-edged sword, nonetheless, in that they can give birth to newfound capabilities or cripple existing pillars of fortitude. Antibiotic resistance is

typically facilitated by efflux pumps capable of exporting incoming drugs out of the sensitive intracellular space - alternatively, by enzymes or proteins capable of modulating drugs as to neuter their effectiveness[28]. Mutations of these types of genes could deactivate them or affect their binding affinity to their target drug, resulting in a decline in bacterial defenses. In essence, the same mutagenesis proposed to confer beneficial mutations to *E. coli* 131 and *E. coli* 158 could also be damaging existing antibiotic resistance regimes in favor of susceptibility.

The basis for the seemingly unique combinations of drug *E. coli* 131 and *E. coli* 158 grow resistant to and or susceptible to - the magnitude by which they become resistant and or susceptible in absence of Fur - is unclear. Following from the ROS-mediated mutagenesis explanation: it could be possible that *E. coli* 158 has evolved in a fashion to better counteract ROS and any consequent mutagenesis. As seen in its Fur regulon, Pentose Phosphate Pathway genes are uniquely upregulated. The resulting pool of NADPH and related anti-stressor molecules could be larger versus the other strain. On the other hand, *E. coli* 131 may be more susceptible to mutations overall. This strain may be slightly more deficient in Fur regulation of genes tied to the appropriate stress response. Alternatively, given the variability of these strains, it is possible one of these strains inherently possesses antibiotic resistance genes that the other does not have. In summary, our results demonstrate that *E. coli* 131 and *E. coli* 158 can develop certain antibiotic resistances in the absence of Fur potentially through a mechanism like that for model strains.

CONCLUSION

Iron is a key metal to the livelihood of many species via it roles in DNA synthesis, aerobic respiration, growth, and many more essential biological processes. In bacteria, maintaining an appropriate concentration of intracellular iron is critical as to meet normal cellular demands while avoiding the overproduction of toxic ROS species. The corresponding transcription factor in control of iron metabolism, Fur, has been extensively studied across various model bacterial strains in order to determine its regulatory scope and regulatory methods. However, very little research has been done on Fur's activity and mechanisms of action in non-model strains – strains that may not confer to the idealism of a laboratory-cultivated candidate. Such data may more accurately elucidate Fur to not only better our understanding of bacterial systems, but to also contribute to future studies aimed at combatting iron-dependent pathogens. In this study, we discover major conservation of the Fur regulon, Fur regulatory modes, and Fur binding motifs in two non-model, clinical strains, *E. coli* 131 and *E. coli* 158, relative to CFT073 and MG1655. We also validate previously encountered phenotypic effects of an absence of Fur – reduced growth output and development of antibiotic resistance.

6.1 - Conclusions on Fur's Impact on Growth, Future Directions

E. coli 131 and *E. coli* 158 *fur* mutants both showcased a significantly lowered maximum OD versus their wild-type counterparts. However, *E. coli* 131 was the only strain to actually exhibit significantly decreased growth rates. The former observation is upheld by existing literature for the two model strains used in this study, MG1655 and CFT073, as well as many other *E. coli* strains. Fur is naturally essential to combatting iron-related ROS stress - its absence spells disaster for the involved participants[3,13]. Fur occasionally even encompasses some genes vital to other pathways pertaining to growth and proliferation (e.g., the *araC*/ *araBAD* operon,

responsible for arabinose metabolism)[6]. The latter observation (that only *E. coli* 131 showcased decreased growth rates) was predicted to be a result of the different Fur regulons between *E. coli* 131 and *E. coli* 158. Per **Chapter 2** and **Chapter 3**, we had discerned that the Fur of *E. coli* 131 was more involved in many of the strain's non-iron metabolism related processes. For example, biofilm formation (*flu*), aromatic amino acid metabolism (*aroH_1*), and non-iron-based electron transfer (*btuD_1*, *gpmA*) were all direct targets of Fur. Following this, a lack of Fur cripples a broader range of physiological variables for *E. coli* 131 unlike *E. coli* 158, making the former more vulnerable overall.

Future studies interested in investigating this distinction might probe some of the unique Fur regulon genes in *E. coli* 131 to see if the aforementioned hypotheses actually amount to anything. For example, individual genes within the Fur regulon could be knocked out via the same lambda red recombinase system used in this study or through other molecular methods such as shRNA - this independent of any direct *fur* knockouts - in an effort to study consequences on growth patterns. Alternatively, the same *fur* mutants employed in this study could have target genes supplemented in whatever way they were presumably affected. For example, if *btuD_1* was downregulated post Fur knockout, gene expression could be remediated through a transgenic model to see if this rescues our strains.

6.2 - Conclusions on the Fur Regulon, Future Directions

The Fur regulon of *E. coli* 131 and *E. coli* 158 did overlap with portions of *E. coli* MG1655 and *E. coli* CFT073, though they both possessed their own sub-shared and unique portions. Many of the genes in the "core genome" were those in COG categories of Energy Production and Conversion (C), Inorganic Ion Transport and Metabolism (P), and Transcription (K). These included synthesis of the siderophore "enterobactin" (*entCEBAH, entD, entS, fes*) and

siderophore-mediated ferric ion uptake (e.g., *cirA, exbBD, fepA, fepB, fepDGC, fiu, fhuA, yqjH*). In addition, the core genome included ion transport (*mntH*), metabolic functions (*gpmA, nrdHIEF*), and proteins of unknown function and identity (Hypothetical Proteins, yncE). All of these were predictable binding sites given Fur's designated role as a mediator to iron homeostasis. Genes in the "accessory genome" stemmed similar COG categories, and individual genes did not stray too far in functionality from those above (e.g., many were extensions of already labelled iron-related processes, including iron uptake [*efeUOB*], and ferritin production [*ftnA_1*]). Genes in the "unique genome" spoke more to the diverse nature of the clinical isolates versus common laboratory-cultivated samples.

We already provided some light speculation on why certain genes were targeted by Fur for *E. coli* 131 and *E. coli* 158 in **Chapter 3** based on cited literature. We did not verify in this study if any of the proposed hypotheses are correct, however. It would thus be of interest to probe these genes for their relationship to Fur, and or probe these genes for their ties to iron concentration and or iron metabolism. For example, a variety of genes were seemingly implicated in the amino acid-dependent pathogenesis of the two strains given their evolutionary relatedness to the uropathogen CFT073. Using the same Fur knockouts administered here within the context of a simulated urinary tract environment deficit of amino acids could better confirm if any of the genes are genuinely involved in the way we coin them to be. Alternatively, knocking out these amino-acid-related genes as opposed to Fur could provide more intimate analyses on amino acid importance separate to iron concentration. Furthermore, both *E. coli* 131 and *E. coli* 158 Fur exhibited direct regulation of heme metabolism proteins separate to either MG1655 or CFT073. Observing growth patterns and associated phenotypic behavior of the strains in the presence or absence of heme would attest to or provide dismissal of our

connections. A similar process to how iron-replete and iron-deplete conditions were assembled in our experiments could be employed - an iron chelator could remove innate iron in solution, then heme could be physically added.

6.3 - Conclusions on the Fur Regulatory Methods and Binding Motifs

We established that all genes in the regulons of both *E. coli* 131 and *E. coli* 158 are seemingly under the control of *holo*-Fur repression via a canonic palindromic binding motif. Furthermore, we found instances of Fur acting both on singular and divergent promoters. This speaks to the conserved nature of Fur in our clinical isolates despite our hypothesis that their unique upbringing might spell distinct properties. Nonetheless, our survey of these regulatory elements within the *E. coli* 131 and *E. coli* 158 Fur regulons could be explored more intimately. How tightly Fur binds to these motifs and the specificity the motif offers per gene relative to these strains, for example, was not discussed despite being alluded to as a contributing factor for survey. Discovering binding strength can be done on two fronts. The first front is to calculate binding affinity through the bases of each individual motif. Guanine to Cytosine interactions within DNA tend to be stronger than Adenine to Thymine as a result of an additional hydrogen bond between pyrimidine and purine. Using a related algorithm might help discern the contribution of each individual bases to the whole sequence. The second front is to correlate gene expression levels with binding motifs. RNA-seq results can be indicative of binding motif strength in addition to sorting real ChIP-exo binding events from background noise.

6.4 - Conclusions on Fur-Dependent Antibiotic Resistance and Antibiotic Vulnerability

We had found that, for both *E. coli* 131 and *E. coli* 158, an absence of Fur cultivated both significantly stronger resistances to certain antibiotics and significantly weaker resistances to others. The five antibiotics highlighted for both strains, respectively, were Chloramphenicol

(more resistant), Colistin (less), Penicillin (more), Spiramycin (less), and Vancomycin (less) - Chloramphenicol (more), Carbenicillin (more), Enoxacin (more), Nafcillin (less), and Vancomycin (less). Though each individual antibiotic type and the trends they underwent may not necessarily be affirmed in existing literature, the overarching trends that occurred are. *E. coli* K-12 *fur* mutants have been documented to have altered antibiotic resistance patterns (i.e., Ciprofloxacin resistance) in response to losing the key iron-regulating transcription factor[26]. The mechanism behind these shifting patterns, however, remains up in the air. Long standing tradition in the scientific community has it that antibiotic resistance may be a result of ROS-mediated mutagenesis of the bacterial genome. A lack of Fur leads to dysregulation of iron homeostasis and ROS stress responses, thus allowing accumulation of the products to Fenton and or Haber Weiss reactions that can directly damage DNA[3,40]. Nonetheless, some studies have found that specific antibiotic resistances may not be acquired or lost through this path. A plausible alternative thus coined was the fact that DNA Polymerase III is capable of accidentally incorporating oxidized nucleotides into its repair mechanisms during DNA replication[46]. On top of this, DNA polymerases IV and V are typically error prone especially in the context of ROS stress. Any of these three hypotheses could very well apply to the findings presented in this study.

A natural course of action here, in the case this aspect of the survey proves the most enticing, would be to verify if any of our two clinical's strains antibiotic resistance patterns abide to any of the previous three hypotheses. Knockouts could be administered to those ROS-mediating genes discovered in either the regulon of *E. coli* 131 or *E. coli* 158 to see what outcome is had on the mutation rates of the clinical strain's genomes. Should there be a significant trend between set gene(s) and mutation rates, an iron chelator could be used to inhibit

the Fenton and Haber-Weiss reactions and discover if newfound antibiotic resistances are affected as a result. Should ROS-mediated mutagenesis not be the correct mechanism, double mutants of *fur* and the genes to either DNA Polymerase III, IV, or V could be assembled to test the effect each polymerase has in Fur's absence. It might seem natural that at least one of the three hypotheses would explain the case presented in this paper; nonetheless, if this were not the case, this would serve as a potential groundbreaking discovery for how antibiotic resistance arises through Fur.

6.5 - Study Limitations, Final Remarks

As with most scientific studies, our own research is not without limitations. The most glaring limitation at hand is the very small sample size of clinical isolates employed within this comparative survey. *E. coli* 131 and *E. coli* 158 are only two strains among many non-model strains available for research including the larger set of clinical isolates initially alluded to in **Figure 1**. Thus, their Fur regulons and *fur* mutant phenotypes may not be representative of non-model strains as a whole. This is compounded by the fact that they were isolated from very specific patient types rather than a myriad of diverse locations and environments. Even if these strains were discovered to be indicative of their perceived serotype (uropathogens), other non-conforming, non-model strains might still be more telling in the vast potential of Fur as a transcription factor. Future studies, therefore, might be aimed at expanding upon what work was done here. Said studies could employ the same experimental techniques (ChIP-exo, RNA-seq, various phenotypic assays) on other non-model strains either isolated in the same conditions or similar conditions to distinguish their Fur regulons and Fur regulatory methods. These experiments could also be detailed further through the suggestions discussed in each preceding Chapter of the conclusion section - for example, probing the relationship between individual

genes and Fur through specific gene knockouts or correlating mRNA expression with Fur binding motifs.

In a similar line of thought to that above: a very small sample size of reference strains is also employed. This choice is more grounded in reason - we had mentioned how we had chosen *E. coli* MG1655 and *E. coli* CFT073 as a result of their opposing and or associated relatedness, respectively, to *E. coli* 131 and *E. coli* 158. However, choosing additional related and or contrasting strains might be more revealing of which differences uncovered in this study are genuinely meaningful versus others. For example, a lot of our findings were dependent on the fact that *E. coli* 131 and *E. coli* 158 were truly the most closely related to the UPEC *E. coli* strain CFT073. However, this could potentially be proven false with ease given the inclusion of more strains, which would then put into question much of the hypotheses coined throughout the study. In order to resolve this issue, future studies, again, may still utilize the same methodology outlined in this study, though with the inclusion of extra model strains either situated in the phylogeny of **Chapter 1** and or documented to have some place within said phylogeny. This provides more coverage that could lead to better resolution of extrapolated data. This would also make pan-genomic studies more available and within reason, enabling better survey of widespread patterns of iron homeostasis processes and or related processes dictating bacterial livelihood and or pathogenicity.

Following from the conceptual aspects of a small sample size, this study is more an entry-level dive into the topic of non-model *E. coli* versus model *E. coli* rather than a comprehensive, intricate analysis of the fundamental divergence in Fur between model and non-model strains. In essence, we present a myriad of interesting avenues by which we found Fur to be unique in unordinary circumstances but fail to thoroughly understand why. For example, the centerpiece of

this study is **Chapter 3**, which highlights all the genes within the Fur regulon of the four strains featured in this study. Commentary on each individual gene, however, is grounded on speculation rather than raw fact - we do not ascertain the role of certain genes being present in one strain versus the other. In **Chapter 4**, we offer the binding motifs for Fur in our four strains and qualitatively assess their similarities and differences. However, we don't deeply explore why this is the case nor why certain nuances take place. This paper offers foundational information that would need to be expanded upon in future studies for better efficacy and solidarity. Many of these novel avenues have already been described in the previous conclusion sections.

With the end of this study, we have described here an integrative analysis of various types of genome-scale experimental data applied to non-model *E. coli* strains to better understand the potentially complex roles of the Fur regulatory network in bacteria. By combining the high-resolution ChIP-exo[29,33] with the intimately sensitive RNA-seq-based transcriptome analysis, we have provided insight into the Fur regulon of; the effects *fur* mutants have on phenotypes for; and the Fur regulatory patterns associated with clinical, non-model *E. coli* strains versus common laboratory strains. In the future, we hope to incorporate this data into systems-level approaches tasked at computing more intricate genomic and phenotypic interactions - also towards better understanding the action of Fur in pathogenesis.

METHODOLOGY

7.1 - Bacterial Strains, Media, and Growth Conditions

All strains (wild-type and knockout) used in this study are listed in **Table 1**. The model strains *E. coli* MG1655 and *E. coli* CFT073 were obtained from in-house stock[45]. The non-model, clinical isolate strains *E. coli* 131 and *E. coli* 158 were provided from a collaboration with the University of California Irvine School of Medicine. These two strains were documented to have been isolated from bacteremia patients in a clinical environment. Details on isolation procedures and patient information is currently redacted. *fur* mutants were constructed through a λ Red-mediated site-specific recombination system[10].

Strains were grown in M9 minimal media (47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 8.6 mM NaCl, 18.7 mM NH_4Cl, 2 mM $MgSO_4$, and 0.1 mM $CaCl_2$) with 0.2% (w/v) glucose. 5 mL seeding cultures were consistently used as a basis for subsequent molecular biology assays – these were grown at 30°C overnight in Falcon™ Round-Bottom Polystyrene test tubes with vigorous agitation. Said cultures were supplemented with 1 mL trace element solution (100X) containing 1 g EDTA, 29 mg $ZnSO_4.7H_2O$, 198 mg $MnCl_2.4H_2O$, 254 mg $CoCl_2.6H_2O$, 13.4 mg $CuCl_2$, and 147 mg $CaCl_2$.

For both ChIP-exo and RNA-seq experiments, strains were grown in either iron-replete conditions or iron-deplete conditions. The former consisted of the aforementioned M9 minimal media supplemented with 0.1 mM $FeCl_2$; (Fisher Scientific, USA); the latter M9 minimal media supplemented with 0.2 mM 2,2'-dipyridyl (Sigma Aldrich, USA). Trace element solution was not provided in either case. Strains were inoculated from seeding cultures so that the initial OD_{600} was within the range of 0.05 to 0.1. They were then incubated at 37°C with vigorous

agitation to the mid-log phase ($OD_{600} \approx 0.5$). Subsequent biological replicates were derived from the same 10 mL fresh culture for ChIP-exo or from two separate 10 mL cultures for RNA-seq.

7.2 - ChIP-exo Experiments

Chromatin immunoprecipitation (ChIP) followed by lambda exonuclease activity and high-throughput sequencing (ChIP-exo) experimentation granted us the capability to obtain high-resolution binding peaks for the Fur transcription factor. Only *E. coli* 131 and *E. coli* 158 required live experiments; previously sequenced ChIP-exo reads provided by the Systems Biology Research Group were used to supplement *E. coli* MG1655 and *E. coli* CFT073. Cells grown to mid-log phase were crosslinked through use of formaldehyde. They were then lysed via lysozyme to free Fur-bound DNA – said DNA was later fragmented through sonification for 50 minutes at 0°C. Fur-specific antibodies followed by Dynabeads Pan Mouse IgG magnetic beads (Invitrogen) were employed to isolate the now free genomic material for later complexion.

ChIP materials (chromatin-beads) were used to perform on-bead enzymatic reactions of the ChIP-exo method[29,33]. In short, the sheared DNA of chromatin-beads was repaired by the NEBNext End Repair Module (New England Biolabs) followed by the addition of a single dA overhang and ligation of the first adaptor (5'-phosphorylated). This was done through the NEBNext Quick Ligation Module (New England Biolabs) and the dA-Tailing Module (New England Biolabs), respectively. Nick Repair was performed by using PreCR Repair Mix (New England Biolabs). Lambda exonuclease- and RecJf exonuclease-treated chromatin was eluted from the beads; overnight incubation at 65°C reversed the protein-DNA cross-link.

At this point, the samples underwent primer extension and second adaptor ligation with the following modifications. The DNA samples incubated for primer extension as described previously were treated with dA-Tailing Module (New England Biolabs) and NEBNext Quick

Ligation Module (New England Biolabs) for second adaptor ligation. The DNA samples were subsequently purified by a GeneRead Size Selection Kit (Qiagen) and enriched with polymerase chain reaction (PCR) using a Phusion High-Fidelity DNA Polymerase (New England Biolabs). The now-amplified libraries were purified again by a GeneRead Size Selection Kit (Qiagen) and quantified using a Qubit dsDNA HS Assay Kit (Life Technologies). Quality of the DNA sample was verified by running Agilent High Sensitivity DNA Kit using Agilent 2100 Bioanalyzer (Agilent) before HiSeq 2500 sequencing (Illumina) following the manufacturer's instructions. Each modified step was also performed following the manufacturer's instructions. ChIP-exo experiments were performed in biological duplicates.

7.3 - RNA-seq Expression Profiling

3 mL of cells grown to mid-log phase were complexed with 6 mL RNAprotect Bacteria Reagent (Qiagen). Samples were mixed immediately by vortexing for 30 seconds, incubated for 5 minutes at room temperature, and centrifuged at 5000*g* for 10 minutes. The supernatant was decanted, and any residual supernatant was removed by inverting the tube once onto a paper towel. Total RNA samples were then isolated using a Quick-RNA Fungal/Bacterial Kit (Zymogen) following the manufacturer's instructions. Samples were quantified using a Nanodrop 1000 spectrophotometer (Thermo Scientific), and quality of the isolated RNA was assessed by running an RNA 6000 Pico Kit in tandem with the Agilent 2100 Bioanalyzer (Agilent). Ribosomal RNA was removed using an in-house procedure involving hybridization of complementary RNA to excess rRNA followed by contaminant removal with an RNA Clean and Concentrator-5 kit (Zymogen)[9]. Paired-end, strand-specific RNA-seq libraries were prepared using a KAPA RNA Hyper Prep kit (KAPA Biosystems), following the manufacturer's

instructions[31,34]. The resulting libraries were analyzed on an Agilent Bioanalyzer DNA 1000 chip (Agilent). Sequencing was performed on a HiSeq 2500.

7.4 - Read Analysis

ChIP-exo reads, post-sequencing, were aligned to their respective genomes via Bowtie2. The data was then subject to Samtools and Bamtools for processing – peak calling was done using the MACE algorithm. Genes associated with peaks were manually determined through Metascope visualization. RNA-seq reads, post-sequencing, were also aligned to their respective genomes via Bowtie2 – this followed up, again, with Samtools and Bamtools processing. The derived data was then run through the DESeq package to determine differentially expressed genes. The Fur regulon was determined by manually discovering genes co-occurring in both ChIP-exo and RNA-seq datasets.

7.5 - Fur Binding Motif Survey of ChIP-exo Data

The sequence motif analysis for Fur binding sites was performed using the MEME software suite[5]. For each strain, sequences in binding regions were extracted from the reference genome after curation through MACE (**Dataset S1**). To locate an accurate motif, the sequence of each binding site was extended by 10bp at each end. The width parameter was fixed at 20 bp and the min-sites parameter was fixed at 90% of the total number of the sequence. All other parameters followed the default setting.

7.6 - Growth Curve Analysis

In order to determine differences in growth should Fur be dysfunctional and or absent, we performed Growth Curve Analysis through the Agilent Biotecan system. 1 uL aliquots of cells grown in M9 media seeding cultures were washed and inoculated into 96-well microtiter plates

containing more M9 media. These plates were placed inside the Biotecan instrument and monitored OD600 every 15 minutes across 24 hours for growth patterns. Output data was subject to the R package Growthcurver in an to derive major statistics such as the growth rate and carry capacity. Growth curves and growth rates were visualized using the Python package Matplotlib.

7.7 - Antibiotic Resistance Phenotypic Microarrays

The phenotypic fingerprints of cells grown on FL-BSG medium and MRS were recorded using the Omnilog Phenotype MicroArray (PM) platform (Biolog, Hayward, CA, USA). Each PM assay was performed on two biological replicates, in accordance with the manufacturer's instructions. Briefly, cells at the late exponential (LE) growth (after ca. 14 h on FL-BSG medium and ca. 8 h on MRS) were collected, washed in sterile potassium phosphate buffer (50 mM, pH 7.0), and inoculated in PM1 and PM2 microplates, which account for 190 different carbon sources. Kinetic data from PM panels were automatically recorded by the Omnilog reader (Biolog) during incubation at 33 °C for 48 h. Generated longitudinal data were analyzed using the Micro4Food PM pipeline. Briefly, blank subtraction was performed, and metabolic profiles were categorized as active and non-active. Metabolic signals were normalized per replicate and array. After removal of common non-active profiles, metabolic parameters were computed using a free splines method and confidence intervals (CI) were determined through bootstrapping.

7.8 - Phylogenetic Analysis

The genomes for model strains indicated in Dataset S1 were obtained through the NCBI database. The genomes for non-model strains were obtained through hybrid sequencing - complexion of long-read sequencing reads and short-read sequencing reads. Short-read sequencing reads were provided on request similarly to model strain ChIP-exo data (via Systems

Biology Research Group). Long-read sequencing reads were manually constructed through a combination of a modified CTAB Genomic DNA extraction procedure and standard Nanopore sequencing[4]. Hybrid sequencing was done through Unicycler – annotation with Prokka.

To detail collection of long-read sequencing reads: 5 mL worth of cells grown in mid-log phase were pelleted through centrifugation at 3000*g* for 10 minutes. The supernatant was decanted before incubation with 50 uL CTAB for 60 minutes at 70°C. 500 uL of chloroform was then added, and the mixture was incubated on ice for 30 minutes. After subsequent centrifugation at 10,000*g* for 1 minutes, the liquid phase was extracted and complexed with 500 uL phenol:chloroform and 500 chloroform. The liquid layer was extracted once again, and the DNA was precipitated through a Na-acetate and ethanol solution. The DNA was pelleted and placed through an ethanol wash. Elution Buffer was used to dissolve all the DNA before preparation for Nanopore.

The DNA was properly prepared per manufacturer's instructions and subject to sequencing on a Flongle adapter in tandem with MinION per Oxford Nanopore Technologies. Basecalling and De-multiplexing was done after sequencing through Guppy. The data was concatenated and then subject to Unicycler alongside the obtained short-read sequencing reads. The finalized, raw genome was annotated with Prokka and used system-wide for experimentation.

REFERENCES

1. Abdul-Tehrani H, Hudson A, Chang Y, Timms A, Hawkins C, Williams J, Harrison P, Guest J, Andrews S. (1999). *Ferritin Mutants of* Escherichia coli *are Iron Deficient and Growth Impaired, and* fur *Mutants are Iron Deficient.* American Society for Microbiology 181, 1415-1428.

2. Anjem A, Varghese S, Imlay J. (2009). *Manganese import is a key element of the OxyR response to hydrogen peroxide in* Escherichia coli. Molecular Microbiology 72, 844-858.

3. Andres S, Robinson A, Rodríguez-Quiñones F. (2003). *Bacterial Iron Homeostasis*. FEMS microbiology Reviews 27, 215-237.

4. Arseneau J, Steeves R, Laflamme M. (2017). *Modified low-salt CTAB extraction of high-quality DNA from contaminant-rich tissues.* Molecular Ecology Resources 17, 686-693.

5. Bailey T, Boden M, Buske F, Frith M, Grant C, Clementi L, Ren J, Li W, Noble W. (2009). *MEME SUITE: tools for motif discovery and searching*. Nucleic Acids Research 37, 202-208.

6. Banerjee R, Weisenhorn E, Schwartz K, Myers K, Glasner J, Perna N, Coon J, Welch R, Kiley P. (2020). *Tailoring a Global Iron Regulon to a Uropathogen.* American Society for Microbiology 11, 2.

7. Carpenter B, Whitmire J, Merrell D. (2020). *This Is Not Your Mother's Repressor: The Complex Role of Fur in Pathogenesis*. Infection and Immunity 77, 2590-2601.

8. Case R, Boucher Y, Dahllöf I, Holmström C, Doolittle W, Kjelleberg S. (2007) *Use of 16S rRNA and* rpoB *Genes as Molecular Markers for Microbial Ecology Studies*. Applied and Environmental Microbiology 73, 278-288.

9. Choe D, Szubin R, Poudel S, Sastry A, Song Y, Lee Y, Cho S, Palsson B, Cho B. (2021). *RiboRid: A low cost, advanced, and ultra-efficient method to remove ribosomal RNA for bacterial transcriptomics.* PLOS Genetics 17, 9.

10. Datta S, Costantino N, Court D. (2006). *A set of recombineering plasmids for gram-negative bacteria*. Gene 379, 109-115.

11. Dean, C, Neshat S, Poole K. (1996). *PfeR, an Enterobactin-Responsive Activator of Ferric Enterobactin Receptor Gene Expression in* Pseudomonas aeruginosa. Journal of Bacteriology 178, 5361-5369

12. Dean C, Poole K. (1993). *Cloning and characterization of the ferric enterobactin receptor (pfeA) of* Pseudomonas aeruginosa. Journal of Bacteriology 175, 317-324.

13. Desnoyers G, Morissette A, Prevost K, Masse E. (2009). *Small RNA-induced differential degradation of the polycistronic mRNA iscRSUA.* The EMBO Journal 28, 1551-1561.

14. Franza T, Sauvage C, Expert D. (1998). *Iron Regulation and Pathogenicity in* Erwinia chrysanthemi *3937: Role of the Fur Repressor Protein*. The American Phytopathological Society 12, 119.

15. Garcia E, Brumbaugh A, Mobley H. (2011). *Redundancy and specificity of* Escherichia coli *iron acquisition systems during urinary tract infection*. Infection and Immunity 79, 1225-1235.

16. Guillement M, Moreau P. (2008). *Fur-dependent detoxification of organic acids by rpoS mutants during prolonged incubation under aerobic, phosphate starvation*. Journal of Bacteriology 190, 5567-5575.

17. Hartmann A, Braun V. (1981). *Iron uptake and iron limited growth of* Escherichia coli K-12. Archives of Microbiology 130, 353-356.

18. Higgs P, Larsen R, Postle K. (2002). *Quantification of known components of the* Escherichia coli *TonB energy transduction system: TonB, ExbB, ExbD, and FepA*. Molecular Microbiology 44, 271-281.

19. Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund S, Cook H, Mende D, Letunic I, Rattei T, Jensen L, Mering C, Bork P. (2018). *eggNOG 5.0: a hierarchical, functionally, and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses.* Nucleic Acids Research 47, 1.

20. Jin D, Cagliero C, Zhou Y. (2012). *Growth rate regulation in* Escherichia coli. FEMS Microbiology Reviews 36, 269-287.

21. Keller M, Zylstra A, Castro C, Turchyn A, Griffin J, Ralser M. (2016). *Conditional iron and pH-dependent activity of a non-enzymatic glycolysis and pentose phosphate pathway.* Science Advances 2, 1.

22. Kohanski M, Dwyer D, Hayete B, Lawrence C, Collins J. (2007). *A Common Mechanism of Cellular Death Induced by Bactericidal Antibiotics.* Cell 130, 781-783.

23. Lavrrar J, McIntosh M. (2003). *Architecture of a Fur Binding Site: A Comparative Analysis.* Journal of Bacteriology 185, 2194-2202.

24. Mamun A, Lombardo M, Shee C, Lisewski A, Gonzalez C, Lin D, Nehring R, Saint-Ruf C, Gibson J, Frisch R, Lichtarge O, Hastings P, Rosenberg S. (2012). *Identity and Function of a Large Gene Network Underlying Mutagenic Repair of DNA Breaks.* Science 338, 1344-1348.

25. McHugh J, Rodríguez-Quiñones F, Abdul-Tehrani H, Svistunenko D, Poole R, Cooper C, Andres S. (2003). *Global Iron-Dependent Gene Regulation in* Escherichia coli. Journal of Biological Chemistry 278, 32.

26. Mehi O, Bogos B, Csorgo B, Pal F, Nyerges A, Papp B, Pal C. (2014). *Perturbation of Iron Homeostasis Promotes the Evolution of Antibiotic Resistance.* Molecular Biology and Evolution 31, 10.

27. Monje-Casas F, Jurado J, Prieto-Alamo M, Holmgren A, Pueyo C. (2001). *Expression Analysis of the* nrdHIEF *operon from* Escherichia coli. Journal of Biological Chemistry 276, 18031-18037

28. Munita J, Arias C. (2016). *Mechanisms of Antibiotic Resistance.* Microbiology Spectrum 4, 2.

29. Nakato R, Sakata T. (2021). *Methods for ChIP-seq analysis: A practical workflow and advanced applications*. Methods 187, 44-53.

30. Nordlund P, Reichard P. (2006). *Ribonucleotide reductases*. Annual Review of Biochemistry 75, 681-706.

31. Quail M, Otto T, Gu Y, Harris S, Skelly T, McQuillan J, Swerdlow H, Oyola S. (2011). *Optimal enzymes for amplifying sequencing libraries*. Nature Methods 9, 10-11.

32. Ramos A, Olano C, Brana A, Mendez C, Salas J. (2009). *Modulation of deoxysugar transfer by the Elloramycin glycosyltransferase ElmGT through site-directed mutagenesis.* Journal of Bacteriology 191, 8.

33. Rhee H, Pugh B. (2012). *ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy*. Current Protocols in Molecular Biology 24, 24.

34. Ross M, Russ C, Costello M, Hollinger A, Lennon N, Hegarty R, Nusbaum C, Jaffe D. (2013). *Characterizing and measuring bias in sequence data*. Genome Biology 14, 51.

35. Salvail H, Lanthier-Bourbonnais Pascale, Sobota J, Masse E. (2010). *A small RNA promotes siderophore production through transcriptional metabolic remodeling.* PNAS 107, 15223-15228.

36. Schleif, R. (2000). *Regulation of the L-arabinose operon of* Escherichia coli. Trends in Genetics 16, 559-565.

37. Synder J, Haugen B, Buckles E, Lockatell C, Johnson D, Donnenberg M, Welch R, Mobley H. (2004). *Transcriptome of Uropathogenic* Escherichia coli *during urinary tract infection.* Infection and Immunity 72, 6373-6381

38. Seo S. W, Kim D, Latif H, O'Brian E, Szubin R, Palsson B. (2014). *Deciphering Fur transcriptional regulatory network highlights its complex role beyond iron metabolism in Escherichia coli.* Nature Communications, 5, 4910.

39. Sigrell J, Cameron A, Jones T, Mowbray S. (1998). *Structure of Escherichia coli Ribokinase in complex with ribose and dinucleotide determined to 1.8 A resolution: insights into a new family of kinase structures.* Structure 6, 183-193.

40. Touati D, Jacques M, Tardat B, Bouchard L, Despied S. (2021). *Lethal oxidative damage and mutagenesis are generated by iron in delta fur mutants of* Escherichia coli*: protective role of superoxide dismutase.* American Society for Microbiology 177, 9.

41. Troxell B, Hassan H. (2013). *Transcription regulation by Ferric Uptake Regulator in pathogenic bacteria.* Frontiers in Cellular and Infection Microbiology 3, 59.

42. Wang Henian, Gunsalus R. (2003). *Coordinate Regulation of the* Escherichia coli *Formate Dehydrogenase* fdnGHI *and* fdhF *Genes in Response to Nitrate, Nitrite, and Formate: Roles for NarL and NarP.* Journal of Bacteriology 185, 5076-5085.

43. Winter W, Bazydlo L, Harris N. (2014). *The molecular biology of human iron metabolism*. Laboratory Medicine 45, 92-102.

44. Wu Y, Outten F. (2009). *IscR Controls Iron-Deficient Biofilm Formation in* Escherichia coli *Type I Fimbria Expression.* Journal of Bacteriology 191, 4.

45. Gao Y, Bang I, Seif Y, Gayoung N, Sastry A, Chen K, Monk J, Choudhary K, Seo S, Lee E, Kim D, Palsson B. (2020). *Reconstruction of Fur pan-regulon uncovers the complexity and diversity of transcriptional regulation in* E. coli. BioRxiv 10, 1101.

46. Yamada M, Shimizu M, Katafuchi A, Gruz P, Fujii S, Usui Y, Nohmi T. (2012). Escherichia coli *DNA polymerase III is responsible for the high level of spontaneous mutations in* mutT *strains*. Molecular Microbiology 86, 1364-1375.