Carbon Source effects on bacterial growth and antibiotic efficacy

Effects of Carbon Metabolism on the growth of bacteria and antibiotic efficacy

By: Madeline Tong, B. Sc.

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Doctor of Philosophy

McMaster University DOCTOR OF PHILOSOPHY (2022) Hamilton, Ontario
(Biochemistry and Biomedical Sciences)

TITLE: Effects of carbon metabolism on the growth of bacteria and antibiotic efficacy

AUTHOR: Madeline Tong, B. Sc.

SUPERVISOR: Eric D. Brown, Ph. D.

NUMBER OF PAGES: xv,146

Lay Abstract

There is an urgent need for new antibiotics. Previous antibiotic discovery has primarily been conducted on bacteria growing in nutrient rich laboratory conditions. This led to antibiotics that targeted the same few bacterial processes. However, since bacteria need to survive in a host to cause an infection, there are targets that may be viable during an infection that we miss by using standard laboratory media. Bacteria need a source of carbon to survive, and each infection site contains different chemicals that bacteria can use as a source of carbon. My work studies how bacteria grow in the presence of different carbon sources. First, I systematically tested which bacterial genes are required for *E. coli* to grow in 30 different carbon sources. I then examined the effectiveness of antibiotics on bacteria grown using these different carbon sources. Together, this work helps us understand how changing carbon sources in the growth media we use to cultivate bacteria can change which genes are required and how it may change how bacteria survive antibiotic stress. When we discover the specific compositions of host infection environments, we can leverage this knowledge to find antibiotics that target these carbon acquisition pathways in bacteria.

Abstract

With the rise of antibiotic resistance, there is ongoing need to find new antibiotics. As bacteria develop resistance to the current classes of antibiotics available, it is imperative to discover new ways to target bacteria. In this thesis, I focused on one of the basic components that all bacteria need to survive: a source of carbon. Here, I explore whether we can exploit this aspect for drug discovery. For bacteria to colonize a host and cause an infection, it must first be able to meet its nutritional needs for growth. Different host infection sites will have different carbon sources available. Some sites, like the gut, will have commensal bacteria which will compete with invading pathogens for carbon sources. While we still lack understanding of the specific growth environment bacteria experience during infection, it is important to understand how bacteria grow when given different nutrients. For the first part of my work, I systematically probed the gene essentiality patterns of *E. coli* grown in different carbon sources. I generated a large dataset of growth phenotypes that I compiled into a user-friendly web-application, Carbon Phenotype Explorer (CarPE). I identified many poorly annotated genes, and further characterized the gene *ydhC* as an adenosine transporter. After characterizing how the growth of *E. coli* and the genes essential for survival change depending on each carbon source, I looked at whether antibiotic efficacy changed depending on the carbon source used. I found that growth in oxaloacetate alters the proton motive force and potentiates macrolide antibiotics. I also found that linezolid, a compound that does not work on gram-negative bacteria due to efflux, is more effective when adenosine is the carbon

source. Together, this work forms a foundation for future research into studying how carbon sources can be exploited in the field of antibiotic discovery.

Acknowledgements

Dr. Eric Brown – Thank you for giving me the chance to work in such an amazing environment to pursue my Ph. D. I appreciate all the guidance and support you have given me these past years. Thank you for knowing exactly when I needed the gentle push towards the right direction and when to let me be independent.

Dr. Lori Burrows and Dr. Gerry Wright – Thank you for being the most supportive committee members I could have ever asked for. Despite all my challenges and struggles you were always on my side and gave me invaluable guidance and suggestions.

Mom and Dad – Thank you for everything you have done for me. Mom, you've always been there to support me and can see through all my struggles. Thank you for always believing in me and giving me the freedom to pursue whatever I wanted. Dad, even though you've been halfway around the world, I've always felt your presence, support and love. Thank you for always calling and checking up on me, even though sometimes I struggle to show it, I really did appreciate knowing you were right there with me.

Boris and Lucy – I can't thank you both enough for the number of times you've supported me. It is amazing to have family here in Hamilton and you both always came through for me whenever I needed help. Thank you.

Keely – I don't even know where to find the words to explain how instrumental you have been in this journey. I can always count on you for your unwavering support and understanding. Thank you for always being there and pushing me to be my best. I could not have done this without you.

Brown Lab – To all the members past and present, thank you for all the amazing times we've had together. I've never met such an intelligent and helpful group of people. Thank you for all the great scientific (and not-so-scientific) discussions and teaching me everything I know about research. An extra thank you to Sara, Kristina, Maya, Emily, Garima, Megan, Timsy and Amelia for all the support, especially when I was struggling. Thank you to Lindsey for holding down the metabolism fort with me. Thank you, Shawn, for all the help and fostering my interest in nerdy and computer data things. Thank you, Ken and Rodion, for all the times we've had wild rants and discussions, both scientific and not. I wouldn't be here without you all.

Lynn – I have never met anyone as genuine and kind as you. You always know when to check up on me and you are always there to offer your help and support no matter what the circumstances. I have told you many times that I would not have been able to

complete this without you, and this still stands true. Thank you for believing in me even when I struggled to believe in myself.

Laura and Kirby – I cannot believe how much we've been through since we first met in high school. It's been an incredible journey and I am so proud of all of us. Thank you for all the support through all the different stages in my life and always pushing me to do my best.

Niki and Winker – Every time I start to worry about my future, you two are always there to calm me down and support me. Niki, thank you for being "my only friend" and helping me get through some of the toughest years of my life.

My pocket friends – I never dreamed that a global pandemic would lead me into meeting some of the most amazing and supportive group of friends. If there is a silver lining to one of the most stressful few years of my life, it's that I've met you all as a result of it. Thank you for always having my back and the constant encouragement.

Board game friends – I honestly never dreamed that I would have so many friends to play such esoteric board games with. Thank you for being patient with me while I puzzled and "min-maxed" (much like bacteria does with its metabolism) my way through so many games. All of you and this hobby have been instrumental in teaching me how to think in systems, networks, and rules.

To everyone I've missed because this list is already too long as it is, thank you.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMR ..... antimicrobial resistance
ADP ..... adenosine diphosphate
ATP ..... adenosine triphosphate
cAMP ..... cyclic adenosine monophosphate
CCR ..... carbon catabolite repression
CDC ..... Centre for Disease Control and Prevention
CRP ..... cAMP receptor protein
DHAP ..... dihydroxyacetone phosphate
DMSO ..... dimethyl sulfoxide
E4P erythrose-4-phosphate
ED ..... Entner-Doudoroff
EHEC ..... enterohemorrhagic *E. coli*
EIEC ..... enteroinvasive *E. coli*
EMP ..... Embden-Meyerhof-Parnas
ETC ..... electron transport chain
F6P fructose-6-phosphate
G3P glyceraldehyde-3-phosphate
GO terms ..... gene-ontology terms
Gnt-6P gluconate-6-phosphate
KDPG 2-keto-3-deoxy-6-phosphogluconate
LB ..... lysogeny broth
LEE ..... locus of enterocyte effacement
LPS ..... lipopolysaccharide
MIC ..... minimum inhibitory concentration
MHB ..... Mueller Hinton broth
MOPS ..... morpholinopropane sulfonate
NAG ..... N-acetyl glucosamine
OD ..... optical density
OM ..... outer membrane

P ..... phosphate
PAβN..... phenylalanine-arginine β-napthylamide
PC..... phosphatidylcholine
PEP..... phosphoenolpyruvate
PMF..... proton motive force
PPP..... pentose phosphate pathway
PTS..... phosphotransferase system
TCA..... tricarboxylic acid
UPEC ..... uropathogenic *E. coli*
WHO ..... World Health Organization

Chapter I: Introduction

The rise of antibiotic resistance

The discovery of antibiotics revolutionized modern medicine. Many medical procedures such as organ transplant, cancer therapy, caesarean sections, and treatment of chronic illnesses rely heavily on antibiotics to mitigate risks of infection. Prior to the discovery of penicillin in 1929[1] by Sir Alexander Fleming, infectious diseases were among the leading causes of death worldwide. Prompted by this discovery, researchers sought to discover new antibiotics and to simplify manufacturing methods in order to increase access to the general public.[2] This was well received, and antibiotics became a "panacea" that were used to treat all sorts of ailments. In 1945, Fleming warned that improper use of these drugs could lead to the spread of resistant forms of bacteria against which penicillin would no longer work[2]. Indeed, within a few short years, over 50% of *Staphylococcus aureus* infections were already resistant to penicillin[2]. This phenomenon is called antimicrobial resistance (AMR) and occurs when bacteria develop mechanisms that render previously useful antibiotics ineffective[3–6]. Although AMR is an evolutionary phenomenon that predates the clinical use of antibiotics[7], the widespread use of these drugs has added an unprecedented evolutionary pressure for bacteria to develop and spread resistance. This resistance was initially countered by discovery of numerous antibiotics soon after penicillin[6]. However, as we neared the 2000s, this pipeline of new discoveries began to dry up[6]. The increase in AMR, combined with the lack of new discoveries, is leading us towards a crisis where bacterial infections become untreatable. Numerous health organizations such as the World Health Organization (WHO)[4] and the US Centers for Disease Control and Prevention (CDC)[3] have agreed that AMR is an

urgent health issue that threatens the lives of many. In 2019 alone, an estimated 4.95 million deaths were associated with AMR, with 1.27 million deaths directly attributed to it[5]. If left unchecked, we could return to the pre-antibiotic era, where even simple infections were fatal.

A brief history of antibiotic discovery

For almost a century, we have taken advantage of the many new classes of antibiotics discovered shortly after penicillin[6]. In 1940, Selman Waksman developed a method to systematically test different soil microbes (usually streptomycetes) for their ability to produce natural antibacterial compounds[6,9]. This strategy worked well because microbes produce secondary metabolites that kill off other bacterial competitors. Using this method, Albert Schatz in Waksman's lab discovered streptomycin in 1943. This approach became known as the Waksman platform, and pharmaceutical companies who adopted this method were highly successful in discovering almost all the main classes of antibiotics that we still use today [9]. However, as more antibiotics were discovered using this platform, researchers started to repeatedly discover the same compounds[9]. To make matters worse, these secondary metabolites were created by bacteria to work in specific environments, and not meant to be drugs, meaning they were rife with pharmacological or human toxicity issues[6]. This era, deemed the golden age of discovery, eventually ended in the mid-1960s when the Waksman platform fell out of favour because it no longer yielded new antibiotics.

Following the golden age of discovery, medicinal chemists took the natural product antibiotic scaffolds previously discovered and made synthetic derivatives. These derivatives would improve upon the original antibiotics by lowering the required doses, decreasing resistance, and expanding the spectrum of bacteria on which they could work[6]. This effort was incredibly successful and most of the antibiotics we use today are synthetic derivatives of natural product antibiotics[10]. Within the last four decades, almost 90% of new antibacterial drugs, excluding vaccines, were a natural product or a derivative of one[10]. Nevertheless, since the existing chemical scaffolds remain unchanged, many antibiotics continue to target the same few bacterial cellular processes, namely DNA replication, RNA synthesis, protein synthesis, and cell wall synthesis.

In the 1990s, as resistance to antibiotics grew, scientists realized the need for novel antibiotic scaffolds and targets. Around this time, there were many scientific breakthroughs, including the sequencing of the first free-living organism: *Haemophilus influenzae*, heralding the beginning of a "post-genome era"[6,11,12]. Researchers now had access to high-throughput screening technologies, the ability to purify high yields of proteins, large synthetic chemical libraries and unprecedented amounts of genomic data[11]. Despite all this, three decades passed without many novel antibiotic scaffolds being discovered[13]. During this time, many drug discoverers in different fields were successful in using a "genes-to-drug" approach, where researchers would identify a potential target protein, then screen large libraries of compounds for one that specifically bound the target[11]. Unfortunately, despite the many attempts to try this approach, this ultimately did not work for the field of antibacterial drug discovery[6,13,14]. Compounds that bound their

protein targets exquisitely were often unable to cross the bacterial membrane and thus, lacked whole cell activity[8,13]. Drugs for gram-negative pathogens are particularly notoriously difficult to find because of their robust outer membrane that acts as a formidable permeability barrier[15]. Additionally, these bacteria have efflux pumps that extrude a variety of substrates to protect themselves from many classes of drugs. It is not surprising then, to see that gram-negative pathogens dominate the lists of priority pathogens released by many health organizations[3,8,16,17]. These intrinsic resistances make it exceptionally difficult to discover new drugs that can target these bacteria meaning new strategies are required.

After a result of the shortcomings of biochemical screening, researchers have looked towards more unconventional whole-cell screening methods [9,13,18–20]. By returning to whole-cell screening, we can alleviate some of the problems with permeability since only drugs that can accumulate in bacteria will be selected. Learning from the failures of modern drug discovery, we will have to revisit targets that were previously ignored in the early days of drug discovery. In my thesis, I will look at the work that has been done on targeting bacterial metabolism and show how bacteria growing in the host differs from bacteria growing in conventional laboratory culture methods. My focus will be on gram-negative bacteria because of the difficulty in discovering drugs that work against these organisms. Most of my work has been done on *E. coli*, a popular gram-negative model organism, because of the wealth of information available for these bacteria. By improving our understanding of how carbon availability can control *E. coli* physiology and gene

essentiality, we can contribute to the development of treatments that work well at the site of infection.

Unconventional screens targeting metabolism

During pathogenesis, the host immune system can sequester nutrients to limit their supply to invading pathogens[21,22]. Since many nutrient biosynthetic pathways exist only in bacteria and not the host, they are ideal targets to avoid toxicity. Indeed, researchers have started to investigate nutrient biosynthesis as a potential target for developing antibiotics[18–20]. Our lab has had success in targeting nutrient biosynthetic pathways by screening drugs on bacteria grown in minimal media which is devoid of amino acids, vitamins, and nucleotides[20]. By looking for drugs that show growth inhibition in minimal media but not in nutrient replete rich media, Zlitni *et al.* (2013) discovered compounds that target nutrient biosynthetic pathways[20]. When *E. coli* is grown in nutrient-rich media, there are around 300 genes that are essential for its survival[23,24]. When it is grown in minimal media, there are around 100 additional genes required for growth[25,26]. This screen targeted the subset of genes that are only essential when the bacteria is required to synthesize its own nutrients. Three probes of nutrient biosynthesis were discovered from this screen: one that affects glycine metabolism, a *p*-aminobenzoic acid biosynthesis inhibitor, and a biotin biosynthesis inhibitor[20].

With most minimal media formulations, the carbon source used is typically glucose or whatever yields the highest growth rate for the pathogen of choice in the lab. However, during a host infection, this may not be the carbon source that is present[19,27,28]. For

example, during a *Pseudomonas aeruginosa* lung infection, glucose is not the relevant carbon source. In the pulmonary environment, *P. aeruginosa* must be able to use phosphatidylcholine (PC), which is the major carbon source available in the lung[29–31]. PC is the main component of lung surfactant, which is important for pulmonary function, meaning it is ubiquitous in this infection site. *P. aeruginosa* have lipases that can break down PC into long-chain fatty acids and transport them into the cell[31]. Once in the cell, these fatty acids are reduced by β-oxidation into acetyl-CoA where they enter the TCA cycle. When bacteria are growing on fatty acids as a carbon source, the glyoxylate shunt is preferentially used in order to conserve carbon which is normally released in the form of CO_2. To mimic this environment, Fahnoe *et al.* (2012) screened compounds on minimal media containing acetate as the carbon source to identify inhibitors of the glyoxylate shunt[19]. When *P. aeruginosa* is grown on acetate as a carbon source, mutants lacking the enzymes in the glyoxylate shunt pathway, isocitrate lyase and malate synthase, show a growth defect compared to wild type[19]. This was recapitulated in a murine lung infection model, where the mutant lacking both isocitrate lyase and malate synthase was unable to infect the lung. By looking for drugs that specifically worked when acetate was the carbon source but not when glucose in the carbon source, Fahnoe *et al.* (2012) discovered eight potential lead compounds that targeted the proteins in the glyoxylate shunt[19].

These screens show a promising starting point in targeting different aspects of bacterial metabolism. The idea to screen for drugs in minimal media with acetate as a sole carbon source was a large source of inspiration for my work during my Ph. D. As a field,

we do not fully understand how bacteria grow in alternative carbon sources. Instead of using the same growth media because it is a standard, we should consider the ingredients in our media that we use to grow bacteria in the lab. I addressed this in chapter 2 of this thesis where I explored the genes required for *E. coli* to grow in 30 different carbon sources. If we want to discover new classes of antibiotics, we should prioritize researching ways to exploit how bacteria grow in the host. In the next few sections, I will give an overview of bacterial central metabolism and explore these pathways in the context of host infections and antibiotic stress.

Bacterial Central Carbon Metabolism

The minimum requirements for a bacterium to grow and survive are a source of carbon, nitrogen, water, and inorganic salts. Bacteria have robust systems that allow them to use a vast number of compounds to maintain these requirements for life. Carbon sources are vital because central carbon metabolism is required for bacteria to generate energy in the form of ATP and to provide precursors for all the biosynthetic reactions required for survival[18]. Central metabolism is a conserved process in every living cell, though some pathways may differ slightly in different phylogenetic kingdoms. The three main pathways in bacterial central metabolism include glycolysis, the pentose phosphate pathway, and the tricarboxylic acid (TCA) cycle (also known as the Krebs cycle or citric acid cycle) (Figure 1). In order for a cell to live, there must be metabolic flux through all three of these pathways since many of the intermediates in these pathways are used as precursors for molecules required for growth (Table 1). This means that for a compound to be a usable carbon source, the organism must have the enzymatic capabilities to

convert it into an intermediate in one of these three pathways. To highlight the importance of these pathways, there are extensive redundancies and isozymes so that mutants lacking single enzymes often do not kill the cell[32]. As central metabolism pathways differ slightly from organism to organism, my focus will be on *E. coli*.

In *E. coli* there are three glycolytic pathways that facilitate the conversion of glucose to pyruvate and the flux to each pathway depends on the specific requirements of the organism in the given environment. The most well-known glycolytic pathway is the Emden-Meyerhof-Parnas (EMP) pathway. This pathway consists of 10 steps that both consume and generate ATP from glucose via substrate level phosphorylation. There are two main phases of this pathway, with the first half consuming ATP to form two molecules of glyceraldehyde-3-phosphate (G3P) and the second half generating ATP as well as converting two molecules of G3P into two molecules of pyruvate[33]. In *E. coli,* glucose enters the cell using the glucose-specific phosphenolpyruvate (PEP) dependent phosphotransferase system (PTS) which phosphorylates the substrate as it enters the cell[34,35]. This means the first committed step in converting glucose to glucose-6-phosphate is usually skipped as this is the major substrate being transported into the cell, although *E. coli* also have glucokinase enzymes that can phosphorylate glucose. In the EMP pathway, most steps are reversible reactions where the same enzymes can be used to catalyze the forward reaction for glycolysis and the reverse reaction during gluconeogenesis. There are two main junctions where the reactions are irreversible and the corresponding gluconeogenic reactions require separate enzymes. The first is the irreversible reaction catalyzed by phosphofructokinase and is a key regulatory step for

this pathway. Phosphofructokinase converts F6P into fructose-1,6-bisphosphate, while the reverse reaction is catalyzed by the gluconeogenic enzyme fructose-1,6-bisphosphatase. The second irreversible step occurs during the conversion of phosphoenolpyruvate (PEP) into pyruvate by the enzyme pyruvate kinase. During glycolysis this is one of the reactions that generate ATP. The reverse reaction is catalyzed by PEP synthase during gluconeogenesis and consumes ATP to convert pyruvate into PEP. The EMP pathway of glycolysis converts one molecule of glucose-6-phosphate into two pyruvates, two net ATP, and two NADH molecules. While the EMP pathway offers the most energy generation, there are alternate glycolytic pathways which are important for generating biosynthetic precursors and for using alternate carbon sources.

The pentose phosphate pathway is a second glycolytic pathway that converts glucose-6-phosphate into precursors needed in *de novo* nucleotide, fatty acid, LPS, vitamin B6, and amino acid biosynthesis[33]. This pathway also generates NADPH as reducing equivalents needed for biosynthesis and redox homeostasis[36]. There are two parts to the pentose phosphate pathway, the first being an oxidative phase where glucose-6-phosphate is oxidized to ribulose-5-phosphate (R5P). The non-oxidative phase follows and converts R5P into F6P as well as G3P through a sequence of transaldolase and transketolase reactions. While the EMP pathway of glycolysis generates two molecules of G3P to be converted into pyruvate, the pentose phosphate pathway ends with one molecule each of F6P and G3P. These can be further catabolized by the enzymes in the EMP pathway.

In some bacteria, there is a third glycolysis pathway, the Entner-Doudoroff (ED) pathway, which is required for the metabolism of sugar acids such as gluconate[37]. The ED

pathway uses five enzymes to convert glucose-6-phosphate into one pyruvate, one G3P, and one NADPH. The major difference between the ED and EMP pathways is the intermediate prior to the aldol cleavage which generates the three carbon substrates. In the EMP pathway, fructose-1,6-bisphoshate is cleaved into G3P and dihydroxyacetone-phosphate, whereas in the ED pathway, 2-keto-3-deoxy-6-phosphogluconate (KDPG) is cleaved to form G3P and pyruvate[37]. The G3P is then metabolized through the lower half of the EMP pathway to generate one ATP and one pyruvate. Though this pathway is essential when *E. coli* are grown on sugar acids, the metabolic flux to the ED pathway is negligible during growth on glucose. However, in other organisms, such as in many pseudomonads, the ED pathway forms the core of central metabolism. It has been speculated that this pathway is an ancient pathway that predates even the EMP pathway and has evolved to help *E. coli* survive aquatic and intestinal habitats[38].

The TCA cycle is the main pathway for aerobic respiration, generating a large amount of energy and reducing power[33]. This pathway is also important in generating major biosynthetic precursors for many amino acids (Table 1). The pyruvate generated from glycolysis is first oxidized into acetyl-coA or converted into oxaloacetate in an anaplerotic reaction. When oxygen is present, the TCA cycle takes the acetyl-coA generated from glycolysis and oxidizes it to form reductants. Each turn of the TCA cycle produces one ATP, two CO_2 molecules, and reduces four molecules either in the form of NADH, NADPH or a quinone. These reduced molecules couple with the electron transport chain to move protons across the cytoplasmic membrane, generating a proton gradient called the proton motive force. These protons then travel down the pH gradient

through the enzyme ATP synthase which phosphorylates ADP molecules into ATP, generating energy for the cell.

In many prokaryotic cells like bacteria, there is an offshoot of the TCA cycle called the glyoxylate shunt, a pathway used when there is a need to preserve carbon. The glyoxylate shunt bypasses the reactions that result in a loss of carbon in the form of CO_2. It consists of two enzymes, isocitrate lyase and malate synthase. Isocitrate lyase converts isocitrate into succinate and glyoxylate. Malate synthase converts the resulting glyoxylate and acetyl-coA into malate where it gets converted back into oxaloacetate to restart the TCA cycle. This pathway is important for bacteria to preserve the carbons that would have been released as CO_2 to undergo gluconeogenesis, which allows the bacteria to generate the glucose required to fuel the other pathways of central metabolism. The glyoxylate shunt is upregulated in environments where carbon is in short supply such as on growth during acetate, but it is repressed when *E. coli* is growing on glucose.

E. coli is a facultative anaerobe, which means it can survive even in the absence of exogenous electron acceptors such as oxygen. Under anaerobic conditions, fumarate reductase is upregulated, and α-ketoglutarate dehydrogenase is repressed, which halts the TCA cycle. Instead of going through the TCA cycle, mixed acid fermentation occurs, and pyruvate is metabolized into lactate or formate and acetyl-CoA. Acetyl-coA is then converted into ethanol or acetate which are both excreted from the cell. The ability for *E. coli* to generate so many different end products through fermentation allows it to maintain redox balance while using a wide variety of substrates.

Regulation of central metabolism also controls expression of virulence genes

As central metabolism is essential to cell survival, these pathways are subjected to exquisite regulatory control. This regulatory network consists of multiple levels of control including transcription factors that alter gene expression patterns, post-transcriptional regulation by sRNAs, post-translational modifications of enzymes, and enzymatic responses to changing substrate concentrations[39]. This regulation prevents a waste of resources since there is no need to produce enzymes that metabolize a substrate that isn't present in the environment. A well-known example is the *lac* operon in *E. coli* where the genes involved in lactose metabolism are repressed until lactose is detected. While there are many levels of regulation, there are seven major global regulation factors that regulate numerous genes in *E. coli*[40]. Although there are many global regulators and transcription factors involved in metabolism, here I focus on only one. The one most important to nutrient acquisition and central metabolism is the cAMP receptor protein (CRP) which mediates carbon catabolite repression (CCR).

CCR is a regulatory system that allows bacteria to selectively use substrates as carbon sources based on what is available in the environment. In the presence of glucose, *E. coli* will activate CCR to decrease the expression of genes that encode for enzymes required to use non-glucose carbon sources. Conversely, during times of glucose starvation, *E. coli* will change its gene expression pattern to focus on scavenging and using secondary carbon sources. CCR is controlled mainly by the coordination of the EIIA component of the glucose-specific PTS transporter, the signal molecule cAMP, and the transcriptional activator cAMP receptor protein (CRP)[41]. EIIA is an intermediate

phosphotransfer protein that is involved in the uptake and phosphorylation of glucose as it enters the cell. The phosphorylation state of EIIA dictates whether carbon catabolite repression occurs. EIIA is phosphorylated through the mediation of the PtsI and PtsH proteins, using PEP as the phosphate donor. When glucose is plentiful, EIIA is dephosphorylated as it donates its phosphate to EIIB which phosphorylates glucose as it enters the cell through the PTS permease. Since transport of glucose dephosphorylates EIIA, PEP is required to replenish this phosphate. When the cellular concentration of PEP drops below the concentration of pyruvate, EIIA is predominantly dephosphorylated. Conversely, in the absence of glucose, EIIA remains in a phosphorylated state and activates a membrane bound enzyme called adenyl cyclase. This enzyme synthesizes the signal molecule cAMP which binds to CRP. The cAMP-CRP complex is the major transcriptional activator that controls as much as 10% of the genome in *E. coli*[41]. Among these genes are virulence factors that assist bacteria in colonizing the host during pathogenesis. For example, in uropathogenic *E. coli* (UPEC), CCR regulates the expression of the gene encoding for S fimbrial adhesins, a virulence factor that allows the pathogen to bind to eukaryotic receptor molecules[42].

Since CRP is crucial to nutrient acquisition, carbon metabolism, and the expression of virulence factors, many pathogens become avirulent when the gene encoding CRP is deleted[41,43,44]. CRP has also been linked to the resistance of bacteria to different antibiotics by modulating efflux pump expression[45]. In a CRP mutant, *E. coli* is more sensitive to oxacillin, azithromycin, erythromycin, and crystal violet, which are all substrates of efflux pumps. This sensitivity was supressed by the deletion of *tolC*, an

outer membrane channel that associates with multi-drug efflux pumps in gram-negative bacteria[45]. Since the primary goal for bacteria is to find nutrients in a host, it is not surprising then, that the regulator for metabolism and nutrient acquisition also regulates the genes that help bacteria survive the host environment and antibiotic treatment.

Carbon source can act as extracellular signals for virulence gene expression

The presence of certain carbon sources might act as an environmental cue for the cell to express certain genes that promote survival in the environment[43]. One of the major virulence factors that allow enterohemorrhagic *E. coli* (EHEC) to colonize the host are the ability to form attaching and effacing lesions on enterocytes[46]. The genes for these proteins are coded on a pathogenicity island called the locus of enterocyte effacement (LEE)[47]. The expression of the genes in LEE requires the DNA-binding protein regulator Ler. During growth on glycolytic carbon sources, *ler* and all the LEE genes are downregulated[48]. In contrast, when gluconeogenic carbon sources are used, two transcription factors, KdpE and Cra, bind to the *ler* promoter, upregulating the expression of all LEE genes. Indeed, deletion of *cra* and *kdpE* genes results in cells that are unable to form attaching and effacing lesions[48]. Another study discovered that another carbon source, fucose, was also able to modulate the expression of LEE genes [49]. Alline et al. (2012) showed that in EHEC, a two-component signal transduction system, FusKR senses the presence of fucose and modulates the expression of LEE[49]. They showed that deletion of FusKR impacted the ability of *E. coli* to colonize the gut. Together, these studies show that carbon sources in the host can act as a signal for virulence gene expression[48,50].

Carbon Availability in Different Host Infection Sites

In order to cause disease, bacterial pathogens must be able to survive the environment presented by the host. Different nutrients are available in different host infection sites, and pathogens often have the metabolic pathways uniquely required to survive that niche[27,51]. In a laboratory setting, bacteria are grown on a petri dish or a flask containing a nutrient rich microbiology media such as LB or MHB, which contain an abundance of nutrients that allow bacteria to proliferate very quickly allowing for time-efficient research. Bacterial growth in a laboratory setting is vastly different from bacterial growth in a host. In 1878, Louis Pasteur first described the host as a form of growth media[27,52]. In a sense, we can think of the different infection sites in our body as their own petri dish, with bacteria acquiring nutrients from this host "media" to survive. Naturally, if the bacteria are unable to use the nutrients available, they would not be able to proliferate and cause an infection. Significant research has been done on the genes required for host infection, but metabolic genes have often been overlooked due to a focus on traditional virulence factors[52]. However, the metabolic genes required for infection can yield important clues into the nutrients available and preferentially used by pathogens during an infection[27,53]. In the next few sections, we'll explore a few common infection sites in the host to see what carbon sources are available and look at what techniques pathogens are using to survive in the niche.

Gastrointestinal Tract

E. coli is a facultative anaerobe that grows efficiently within the human gastrointestinal tract. EHEC is a pathogenic strain that can colonize and cause disease.[46] In the large intestine, nutrients are diverse with many different carbohydrates available as a source of carbon[38]. This can come from food digested by the host, the host's tissues and secretions, as well as the mucus layer covering the epithelial cells[38]. For example, common foods such as fruits and proteins contain sugar acids like galacturonate, gluconate, and ketogluconate[38]. *E. coli* lacks the ability to use complex polysaccharides and starches also found in the gut because it lacks the appropriate hydrolase enzymes, but the gastrointestinal tract is home to commensal anaerobic bacteria that can break down these complex carbohydrates into other products such as short-chain fatty acids[22].The mucus layer itself is a complex mixture of glycoconjugates that can contain carbohydrates such as N-acetylglucosamine, N-acetylgalactosamine, N-acetylneuraminic acid, galactose, sialic acids, fucose, glucuronate, and galacturonate[54]. Many of these carbon sources available in the gut can be metabolized by *E. coli*.

For pathogenic strains of *E. coli* to cause disease, they must be able to colonize and outcompete commensal microorganisms present in the gut. The human gastrointestinal tract contains over 400 strains of different commensal bacteria that play vital roles in human physiology[55]. Commensal strains of *E. coli* rarely cause disease unless the host becomes immunocompromised[46]. If an invading pathogen can use a specific carbon source that commensal strains of bacteria do not use, this invading pathogen has an advantage. This forms the basis of Freter's "nutrient niche" hypothesis[56]. He hypothesized that in a community of microorganisms, if each strain preferentially uses

different nutrients, it forms a community that thrives together in a balanced ecosystem[56]. Each population size is dictated by the available concentration of the nutrients that they prefer[56]. The assortment of carbohydrates and nutrients available combined with the hundreds of strains of bacteria present makes the gastrointestinal tract a complex ecosystem.

With the plethora of carbon sources available in the gut, researchers have tried to identify the carbon source *E. coli* uses during colonization. Genes involved in the catabolism of many mucosal-available carbon sources are induced when *E. coli* is grown on mouse mucus[57]. Many strains of *E. coli* can co-exist together in the gut. Different strains can have different catabolic enzymes which dictate which substrates they can use. To see if different strains of bacteria use different carbon sources to colonize the gastrointestinal tract, researchers ran colonization experiments with different commensal, probiotic, and pathogenic strains of *E. coli*[58]. Pathways required to metabolize different carbon sources available in the intestine were deleted to determine which carbon sources were required to colonize a mouse intestine[58]. In general, pathogenic strains used ribose, hexuronates, mannose, N-acetylglucosamine, ribose, and sucrose while commensal or probiotic strains used fucose, gluconate, N-acetylgalactosamine, and N-acetylneuraminate [59]. These different patterns of carbon utilization indicates that pathogenic bacteria may be able to colonize the gut by using nutrients differently from commensal strains.

During colonization and infection of this complex ecosystem full of a diverse array of carbon sources and bacterial neighbours, *E. coli* has several techniques that help it survive. *E. coli* upregulates the genes required in glycogen synthesis and storage when

growing in the gut[60]. Glycogen can serve as an intracellular store of carbon which is useful when there is a competition for nutrients. Mutants lacking enzymes that can synthesize glycogen show a colonization defect[60]. Additionally, *E. coli* exhibits extraordinary metabolic flexibility that allows it to carve out a nutrient niche for itself. In addition to using different substrates, pathogenic *E. coli* can switch which carbon substrates it preferentially uses. Pathogenic *E. coli* uses glycolytic substrates to colonize and maintain an intestinal infection in a mouse model when it is the only strain present. However, in the presence of a commensal strain, it switches over to using gluconeogenic substrates[61]. This is especially interesting given that growth in gluconeogenic carbon sources activates the expression of virulence genes[48]. This metabolic flexibility also allows *E. coli* to use multiple carbon sources at the same time. In different mixtures of sugars *in vitro*, *E. coli* can use up to nine at a time[59]. Knockouts of multiple pathways in the pathogenic strain *in vivo* showed an additive defect in murine gut colonization, meaning *E. coli* requires the ability to use multiple carbon sources for colonization[59]. Thus, the ability for *E. coli* to synthesize and use glycogen, to change which substrates it uses in the presence of different strains, and to use several different carbon sources that commensal strains aren't using are all strategies used for pathogenesis.

Urinary Tract

The most common bacterium to infect the urinary tract is UPEC.[46] These strains can colonize the intestinal tract without causing disease but can readily infect the urinary tract. Unlike the intestinal tract, urine is poor in carbohydrates, iron limited and rich in nitrogen in the form of small peptides, amino acids and urea[62]. Although the nutrient

levels decrease when we look at urine in the bladder, the competition for nutrients also decreases as there are no longer hundreds of commensal strains of bacteria vying for survival. Naturally, since the nutrient environment is different, the genes required for infection also change. Proteins involved in the transport and metabolism of small peptides, gluconate, pentose sugars and arabinose are induced when *E. coli* is growing in urine[62]. To determine which pathways in central metabolism are required for an *in vivo* infection Alteri *et al.* (2009) systematically deleted genes required for each of the central metabolic pathways in *E. coli* and performed infection experiments with the resulting mutants[63]. They found that gluconeogenesis and the TCA cycle were required for infection and that glycolysis, pentose phosphate pathway and the ED pathway were dispensable[63]. Combined with the fact that peptide transport is upregulated in urine, these data suggest that peptides may be a major carbon source during a urinary tract infection since they are metabolized into TCA cycle intermediates. For bacteria to grow on TCA intermediates, they must be able to generate glucose through gluconeogenesis, which explains why these pathways are required for *in vivo* infection.

In addition to being able to metabolize the peptides as a carbon source, UPEC can metabolize D-serine, which is found in mammalian urine[64]. Early studies of laboratory strains of *E. coli* showed that D-serine could cause growth inhibition by inhibiting enzymes in the L-serine biosynthetic pathway[65]. However, Roesch et al. (2003) found that many UPEC strains could metabolize D-serine as a sole carbon source, even though most EHEC strains could not[66]. In a study of global gene expression of an *E. coli* strain that grows well in urine, the D-serine dehydratase gene, *dsdA*, was upregulated[66,67]. The

ability to metabolize D-serine likely gives a fitness advantage to UPEC strains since it acts as another source of carbon and prevents the growth inhibitory effects that the accumulation of D-serine may have.

Intracellular infection

Enteroinvasive *E. coli* (EIEC) can infect colonic epithelial cells and occupy an intracellular niche. Once inside these cells, *E. coli* encounters an environment drastically different from the gastrointestinal tract. Much of what's understood about this environment comes from differential gene expression profiling and ^{13}C-isotopologue-profiling analysis[68]. In an intracellular environment, glucose is available as a carbon source, as it is the carbon substrate the host cells require to survive. However, glucose concentrations may be low since it is quickly converted into glucose-6-phosphate by the host. Interestingly, EIEC mutants that lack the UhpT transporter for glucose-6-phosphate had no growth defect when grown in the host cytosol[69]. In addition to glucose, G3P, fatty acids, and lactate are available as by-products of host cell metabolism. Researchers found that some EIEC strains used glucose whereas others used the other three carbon substrates as a carbon source[69]. Even though some amino acids are available from the host cell, *E. coli* is required to synthesize the amino acids Ser, Glu, Ala, Val, Asp, Lys and Glu *de novo* since they showed an attenuation of virulence when these biosynthetic pathways are deleted[68].

Metabolism and its effect on antibiotics

Susceptibility to antibiotics is dependent on the metabolic and physiological state of the bacterial cell. There is now growing interest in how growth conditions can affect how well antibiotics work[70]. The standard growth medium used to test antibiotic potency is MHB which is a nutrient rich broth. However, many researchers are seeing evidence that antibiotic efficacy *in vitro* does not always match *in vivo* outcomes[71]. Numerous factors such as nutrient availability, oxygen levels, pH, redox potential, osmolarity can all have enormous effects on antibiotic efficacy[70,72,73]. While acute antibiotic treatment has generally been successful in the clinic for decades, we still struggle with treating chronic infections. This may be because the way we currently test antibiotics does not match the environment bacteria encounter in the host [70]. By testing antibiotics using only rapidly growing bacteria in nutrient rich media, we miss the ones that may work on slower growing bacteria in nutrient environments that are more reflective of the host. Additionally, we run the risk of missing antibiotics that may work *in vivo* even though they lack *in vitro* efficacy. In the next few sections, we'll explore how bacteria physiology *in vivo* may be different from the lab environment and how metabolism can play a role in antibiotic efficacy.

Bacterial persistence

Many chronic infections involve bacteria that can persist in the host despite multiple treatments of antibiotics[72]. Bacteria can survive treatment if they are either

resistant or tolerant to the drugs being used. While resistance is characterized by a genetic change that abrogates antibiotic treatment, tolerance is a phenomenon that usually occurs in a non-growing subpopulation and is transient in nature. The presence of the antibiotic-tolerant persister cells have been linked to chronic and recalcitrant bacterial infections often seen in the clinic. Persister cells are thought to be ubiquitous and have been studied in many different bacterial pathogens. Persister cells are often found in biofilms which are prevalent in persistent infections[73,74]. Biofilms are a protective extracellular matrix made from the secretions of polysaccharides and DNA by bacteria[74–76]. The environment found in a biofilm is very different from standard planktonic growth and may be linked to how persisters are formed.[74]

There are many factors that impact persister formation including nutrient stress, acid stress, heat shock, oxidative stress, and carbon source transitions[72,77]. As cells undergo long periods of stress and nutrient limitation, they activate a starvation response, reworking their metabolic network to promote survival[78]. During carbon and amino acid starvation, the stringent response is activated by expression of *relA* and spoT whose products synthesize the alarmone (p)ppGpp[79]. This signal regulates many genes related to bacterial growth such as genes encoding for enzymes needed for replication, translation, and transcription[79]. This leads to a cessation of growth which may be why persisters are tolerant to antibiotics. Indeed, many antibiotics only work on actively growing cells and have little effect on persisters[80]. Deletion of *relA* and *spoT* decreased the number of persister cells remaining after antibiotic treatment implying that the stringent response is required for persister formation.

The exact mechanism by which persisters are formed is still unclear; one of the models of persister formation involves the stringent response and the toxin-antitoxin expression[73]. Maisonneuve et al. (2013) showed that (p)ppGpp could trigger persistence by activating the toxin-antitoxin system[78]. During stress, bacteria can express toxins that inhibit many metabolic processes including translation, DNA metabolism, decreasing the proton motive force and inhibiting protein synthesis[72]. The expression of the corresponding antitoxin neutralizes the toxin's effect when it is no longer under stress. Early research into persisters showed that deleting an antitoxin resulted a greater proportion of persisters compared to the wild-type strain[73]. Together, these studies formed a model of persister formation that involved the stringent response and control of toxin-antitoxin study. However, other studies propose that persister formation can be controlled by cellular ATP levels[81]. As central metabolism slows down, cellular ATP levels decrease, and bacterial growth is stalled, causing the classic dormant phenotype of persisters. Although there may be multiple mechanisms involved in the formation of persisters, they all lead to slow or non-growing cells that are tolerant to antibiotic treatment.

Carbon sources can be used to increase bacterial killing by antibiotics

An interesting approach to targeting bacterial drug tolerance caused by slow growth is to jumpstart the bacteria's metabolism by adding carbon sources. Allison et. al (2011) showed that stimulating metabolism through using central carbon metabolic intermediates enabled the killing of *E. coli* persister cells by aminoglycosides[82]. Upon supplying the persister cells with glycolytic intermediates, cells were able to generate

NADH through glycolysis. NADH was then oxidized by the electron transport chain, contributing to the proton motive force, which is an important aspect of aminoglycoside uptake. Similarly, Meylan et al. (2017) showed that the addition of fumarate stimulated cellular respiration and potentiated the effect of tobramycin, an aminoglycoside, on persister cells. In contrast, they showed that the addition of glyoxylate, which shifts the metabolic flux towards the glyoxylate shunt, and away from cellular respiration, promoted persister formation and antagonized the effect of tobramycin. It is interesting to note that by shifting the metabolic flux away from cellular respiration, there should be less generation of ATP. This is consistent with the hypothesis that decreased cellular levels of ATP may be linked to increased persister formation[81]. These studies highlight how we can take advantage of carbon metabolism to enhance the activity of antibiotics on cells that are difficult to treat.

Antibiotic efficacy changes depending on the environment

The nutrients available in the growth environment play a pivotal role in how well antibiotics work[83,84]. There is growing appreciation for how antibiotics can alter the metabolic state of bacteria which changes how susceptible they are to antibiotics[85]. Metabolism and antibiotic efficacy are inextricably linked, and treatment with antibiotics drastically changes bacterial metabolism[84–87]. Nutrients in the environment directly influence how quickly bacteria grow and antibiotics generally work better on actively dividing cells[80]. In addition to affecting the physiological state of bacteria, specific metabolites in the environment potentiate the activity of antibiotics.

Purine biosynthesis is involved in the susceptibility of cells to ampicillin, ciprofloxacin, and gentamicin[88]. During antibiotic stress, purine biosynthesis tends to be induced due to adenine limitation. This increases the ATP demand, elevating central carbon metabolism and oxygen consumption eventually leading to cell death[88]. Indeed, supplementation of adenine to relieve cellular adenine limitation decreased the killing of *E. coli* by these antibiotics[88]. Conversely, supplementing pyrmidines such as uracil enhanced antibiotic killing by ampicillin, ciprofloxacin, and gentamicin[88]. Recently in our lab, Farha et al. (2018) discovered that by supplementing sodium bicarbonate, the major buffer in our body, in the growth media, antibiotics like macrolides and aminoglycosides worked better than expected [89]. Sodium bicarbonate dissipates the ΔpH gradient of the proton motive force and allows positively charged antibiotics to enter the cell through the driving force of the ΔΨ membrane potential[89,90]. We see similar results when cells are grown in oxaloacetate as a carbon source, as will be described in Chapter 3. These results indicate that changing the growth environment and the metabolites available in the growth media can alter antibiotic susceptibility. This is an interesting avenue to explore since understanding how different growth media can change antibiotic potency can give us clues as to how well antibiotics are working during a host infection.

Research Objective and Organization of Thesis

The main objective in this thesis is to understand the effects of changing carbon sources in growth media on 1) the genes required for growth and 2) the efficacy of antibiotics. This is based on the idea that we still lack understanding of all the genes in one of the most well studied model micro-organisms. This may have caused us to miss

potential therapeutic targets for antibiotics. Chapter 2 describes how I studied the growth of gene deletion mutants in 30 different carbon sources in order to identify genes that may be involved in carbon metabolism. I used the dataset I generated and compared it to the predictions from a metabolic model to find discrepancies. These genes represented the gap in knowledge we still have in metabolism. I was able to link the previously unannotated gene, *ydhC*, to its function as a purine nucleoside transporter based on its phenotype when grown on adenosine as a sole carbon source. In Chapter 3, I test the minimal inhibitory concentration of a panel of antibiotics in *E. coli* grown in media containing different carbon sources. Here, I focused on antibiotics typically not used for gram-negative bacteria in order to study how carbon sources may affect antibiotic permeability and efflux. I found that growth in oxaloacetate was able to potentiate macrolide antibiotics by altering the proton motive force to allow for increased entry. Additionally, growth in adenosine as a carbon source potentiates compounds like linezolid which are typically extruded through bacterial efflux pumps. In Chapter 4, I conclude this thesis with a brief discussion of suggestions for future research.

References

1. Fleming, A. On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. 1929. *Bulletin of the World Health Organization* **79**, 780–790 (2001).
2. Alanis, A. J. Resistance to Antibiotics: Are We in the Post-Antibiotic Era? *Archives of Medical Research* **36**, 697–705 (2005).
3. About Antibiotic Resistance | CDC. https://www.cdc.gov/drugresistance/about.html.
4. Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance.

5. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **0**, (2022).

6. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).

7. Dcosta, V. M. *et al.* Antibiotic resistance is ancient. *Nature 2011 477:7365* **477**, 457–461 (2011).

8. Butler, M. S. *et al.* Analysis of the clinical pipeline of treatments for drug resistant bacterial infections: despite progress, more action is needed. *Antimicrobial Agents and Chemotherapy* (2022) doi:10.1128/AAC.01991-21.

9. Lewis, K. Antibiotics: Recover the lost art of drug discovery. *Nature* **485**, 439–440 (2012).

10. Newman, D. J. & Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *Journal of Natural Products* **83**, 770–803 (2020).

11. MacArron, R. *et al.* Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery 2011 10:3* **10**, 188–195 (2011).

12. Fleischmann, R. D. *et al.* Whole-Genome Random Sequencing and Assembly of Haemophilus influenzae Rd. *Science* **269**, 496–512 (1995).

13. Payne, D. J., Gwynn, M. N., Holmes, D. J. & Pompliano, D. L. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nature reviews. Drug discovery* **6**, 29–40 (2007).

14. Silver, L. L. Challenges of Antibacterial Discovery. *Clinical Microbiology Reviews* **24**, 71–109 (2011).

15. Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **1794**, 808–816 (2009).

16. Boucher, H. W. *et al.* Bad Bugs, No Drugs: No ESKAPE! An Update from the Infectious Diseases Society of America. *Clinical Infectious Diseases* **48**, 1–12 (2009).

17. Tommasi, R., Brown, D. G., Walkup, G. K., Manchester, J. I. & Miller, A. A. ESKAPEing the labyrinth of antibacterial discovery. *Nature reviews. Drug discovery* **14**, 529–542 (2015).

18. Murima, P., McKinney, J. D. & Pethe, K. Targeting Bacterial Central Metabolism for Drug Development. *Chemistry & Biology* **21**, 1423–1432 (2014).

19. Fahnoe, K. C. *et al.* Non-Traditional Antibacterial Screening Approaches for the Identification of Novel Inhibitors of the Glyoxylate Shunt in Gram-Negative Pathogens. *PLoS ONE* **7**, e51732 (2012).

20. Zlitni, S., Ferruccio, L. F. & Brown, E. D. Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. *Nature Chemical Biology 2013 9:12* **9**, 796–804 (2013).

21. Hennigar, S. R. & McClung, J. P. Nutritional Immunity: Starving Pathogens of Trace Minerals. *American Journal of Lifestyle Medicine* **10**, 170 (2016).

22. Fang, F. C., Frawley, E. R., Tapscott, T. & Vázquez-Torres, A. Bacterial Stress Responses during Host Infection. *Cell Host & Microbe* **20**, 133–143 (2016).

23. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* **2**, 2006.0008 (2006).

24. Goodall, E. C. A. *et al.* The Essential Genome of *Escherichia coli*K-12. *mBio* **9**, e02096-17 (2018).

25. Joyce, A. R. *et al.* Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. *Journal of Bacteriology* **188**, 8259–71 (2006).

26. Tong, M. *et al.* Gene dispensability in *Escherichia coli* grown in thirty different carbon environments. *mBio* **11**, 1–20 (2020).

27. Brown, S. A., Palmer, K. L. & Whiteley, M. Revisiting the host as a growth medium. *Nature reviews. Microbiology* **6**, 657–66 (2008).

28. Stecher, B. Finding a Sugary Foothold: How Antibiotics Pave the Way for Enteric Pathogens. *Cell Host & Microbe* **14**, 225–227 (2013).

29. Palmer, K. L., Aye, L. M. & Whiteley, M. Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *Journal of Bacteriology* **189**, 8079–8087 (2007).

30. Fung, C. *et al.* Gene expression of *Pseudomonas aeruginosa* in a mucin-containing synthetic growth medium mimicking cystic fibrosis lung sputum. *Journal of Medical Microbiology* **59**, 1089–1100 (2010).

31. Sun, Z. *et al.* Blocking phosphatidylcholine utilization in *Pseudomonas aeruginosa*, via mutagenesis of fatty acid, glycerol and choline degradation pathways, confirms the importance of this nutrient source in vivo. *PloS one* **9**, e103778 (2014).

32. Nakahigashi, K. *et al.* Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Molecular Systems Biology* **5**, 727–738 (2009).

33. Keseler, I. M. *et al.* The EcoCyc Database in 2021. *Frontiers in Microbiology* **12**, 2098 (2021).

34. Deutscher, J., Francke, C. & Postma, P. W. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and molecular biology reviews : MMBR* **70**, 939–1031 (2006).

35. JH, T., V, N., JS, E. & Jr, S. M. The complete phosphotransferase system in *Escherichia coli*. *Journal of Molecular Microbiology and Biotechnology* **3**, 329–346 (2001).

36. Stincone, A. *et al.* The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biological Reviews* **90**, 927–963 (2015).

37. Conway, T. The Entner-Doudoroff pathway: history, physiology and molecular biology. *FEMS Microbiology Reviews* **9**, 1–27 (1992).

38. Peekhaus, N. & Conway, T. What's for dinner?: Entner-Doudoroff metabolism in *Escherichia coli*. *Journal of Bacteriology* **180**, 3495–3502 (1998).

39. Chubukov, V., Gerosa, L., Kochanowski, K. & Sauer, U. Coordination of microbial metabolism. *Nature Reviews Microbiology* **12**, 327–340 (2014).

40. Martínez-Antonio, A. & Collado-Vides, J. Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology* **6**, 482–489 (2003).

41. Görke, B. & Stülke, J. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. *Nature Reviews Microbiology* **6**, 613–624 (2008).

42. Schmoll, T., Ott, M., Oudega, B. & Hacker, J. Use of a wild-type gene fusion to determine the influence of environmental conditions on expression of the S fimbrial adhesin in an *Escherichia coli* pathogen. *Journal of Bacteriology* **172**, 5103–5111 (1990).

43. Poncet, S. *et al.* Correlations between carbon metabolism and virulence in bacteria. *Contributions to microbiology* **16**, 88–102 (2009).

44. Curtiss, R. & Kelly, S. M. Salmonella typhimurium deletion mutants lacking adenylate cyclase and cyclic AMP receptor protein are avirulent and immunogenic. *Infection and Immunity* **55**, 3035–3043 (1987).

45. Nishino, K., Senda, Y. & Yamaguchi, A. CRP Regulator Modulates Multidrug Resistance of *Escherichia coli* by Repressing the mdtEF Multidrug Efflux Genes. *The Journal of Antibiotics 2008 61:3* **61**, 120–127 (2008).

46. Kaper, J. B., Nataro, J. P. & Mobley, H. L. T. Pathogenic *Escherichia coli*. *Nature Reviews Microbiology* **2**, 123–140 (2004).

47. Mellies, J. L., Barron, A. M. S. & Carmona, A. M. Enteropathogenic and enterohemorrhagic *Escherichia coli* virulence gene regulation. *Infection and Immunity* **75**, 4199–210 (2007).

48. Njoroge, J. W., Nguyen, Y., Curtis, M. M., Moreira, C. G. & Sperandio, V. Virulence Meets Metabolism: Cra and KdpE Gene Regulation in Enterohemorrhagic *Escherichia coli*. *mBio* **3**, e00280-12-e00280-12 (2012).

49. Pacheco, A. R. *et al.* Fucose sensing regulates bacterial intestinal colonization. *Nature* **492**, 113–117 (2012).

50. Njoroge, J. W., Gruber, C. & Sperandio, V. The interacting Cra and KdpE regulators are involved in the expression of multiple virulence factors in enterohemorrhagic *Escherichia coli*. *Journal of Bacteriology* **195**, 2499–508 (2013).

51. Fuchs, T. M., Eisenreich, W., Heesemann, J. & Goebel, W. Metabolic adaptation of human pathogenic and related nonpathogenic bacteria to extra- and intracellular habitats. *FEMS Microbiology Reviews* **36**, (2012).

52. McLean, R. J. C. & Pringle, S. L. Identifying bacterial menu choices from the host buffet during infections. *Journal of Bacteriology* **195**, 4989–4990 (2013).

53. Jorth, P., Trivedi, U., Rumbaugh, K. & Whiteley, M. Probing bacterial metabolism during infection using high-resolution transcriptomics. *Journal of Bacteriology* **195**, 4991–4998 (2013).

54. Allen, A., Bell, A., Mantle, M. & Pearson, J. P. The Structure and Physiology of Gastrointestinal Mucus. *Advances in experimental medicine and biology* **144**, 115–133 (1982).

55. Macfarlane, G. T. & Macfarlane, S. Human Colonic Microbiota: Ecology, Physiology and Metabolic Potential of Intestinal Bacteria. *Scandinavian Journal of Gastroenterology* **32**, 3–9 (1997).

56. Freter, R., Brickner, H., Fekete, J., Vickerman, M. M. & Carey, K. E. Survival and Implantation of *Escherichia coli* in the Intestinal Tract. *Infection and Immunity* **39**, 686 (1983).

57. Chang, D.-E. *et al.* Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proceedings of the National Academy of Sciences* **101**, 7427–7432 (2004).

58. Conway, T. & Cohen, P. S. Commensal and Pathogenic *Escherichia coli* Metabolism in the Gut . *Microbiology Spectrum* **3**, (2015).

59. Fabich, A. J. *et al.* Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. *Infection and Immunity* **76**, 1143–1152 (2008).

60. Jones, S. A. *et al.* Glycogen and maltose utilization by *Escherichia coli* O157:H7 in the mouse intestine. *Infection and Immunity* **76**, 2531–2540 (2008).

61. Miranda, R. L. *et al.* Glycolytic and Gluconeogenic Growth of *Escherichia coli* O157:H7 (EDL933) and E. coli K-12 (MG1655) in the Mouse Intestine. *Infection and Immunity* **72**, 1666–1676 (2004).

62. Alteri, C. J. & Mobley, H. L. T. Metabolism and Fitness of Urinary Tract Pathogens. *Microbiology Spectrum* **3**, (2015).

63. Alteri, C. J. C. *et al.* Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle. *PLoS Pathogens* **5**, e1000448 (2009).

64. HUANG, Y. *et al.* Urinary Excretion of D-Serine in Human: Comparison of Different Ages and Species. *Biological & Pharmaceutical Bulletin* **21**, 156–162 (1998).

65. Cosloy, S. D. & McFall, E. Metabolism of D-serine in *Escherichia coli* K-12: mechanism of growth inhibition. *Journal of bacteriology* **114**, 685–94 (1973).

66. Roesch, P. L. *et al.* Uropathogenic *Escherichia coli* use d-serine deaminase to modulate infection of the murine urinary tract. *Molecular Microbiology* **49**, 55–67 (2003).

67. Roos, V. & Klemm, P. Global gene expression profiling of the asymptomatic bacteriuria *Escherichia coli* strain 83972 in the human urinary tract. *Infection and immunity* **74**, 3565–75 (2006).

68. Eisenreich, W., Dandekar, T., Heesemann, J. & Goebel, W. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nature Reviews Microbiology 2010 8:6* **8**, 401–412 (2010).

69. Götz, A., Eylert, E., Eisenreich, W. & Goebel, W. Carbon Metabolism of Enterobacterial Human Pathogens Growing in Epithelial Colorectal Adenocarcinoma (Caco-2) Cells. *PLOS ONE* **5**, e10586 (2010).

70. Bjarnsholt, T. *et al.* The importance of understanding the infectious microenvironment. *The Lancet Infectious Diseases* **22**, e88–e92 (2022).

71. Ersoy, S. C. *et al.* Correcting a Fundamental Flaw in the Paradigm for Antimicrobial Susceptibility Testing. *EBioMedicine* **20**, 173–181 (2017).

72. Fisher, R. A., Gollan, B. & Helaine, S. Persistent bacterial infections and persister cells. *Nature Reviews Microbiology 2017 15:8* **15**, 453–464 (2017).

73. Lewis, K. Persister cells, dormancy and infectious disease. *Nature Reviews Microbiology* **5**, 48–56 (2007).

74. Lewis, K. Multidrug Tolerance of Biofilms and Persister Cells. in 107–131 (Springer Berlin Heidelberg, 2008). doi:10.1007/978-3-540-75418-3_6.

75. Römling, U. & Balsalobre, C. Biofilm infections, their resilience to therapy and innovative treatment strategies. *Journal of Internal Medicine* **272**, 541–561 (2012).

76. Soto, S. M. & M., S. Importance of Biofilms in Urinary Tract Infections: New Therapeutic Approaches. *Advances in Biology* **2014**, 1–13 (2014).

77. Amato, S. M., Orman, M. A. & Brynildsen, M. P. Metabolic Control of Persister Formation in *Escherichia coli*. *Molecular Cell* **50**, 475–487 (2013).

78. Maisonneuve, E., Castro-Camargo, M. & Gerdes, K. (p)ppGpp Controls Bacterial Persistence by Stochastic Induction of Toxin-Antitoxin Activity. *Cell* **154**, 1140–1150 (2013).

79. Wu, J., Long, Q. & Xie, J. (p)ppGpp and drug resistance. *Journal of Cellular Physiology* **224**, 300–304 (2010).

80. Eng, R. H. K., Padberg, F. T., Smith, S. M., Tan, E. N. & Cherubin, C. E. Bactericidal effects of antibiotics on slowly growing and nongrowing bacteria. *Antimicrobial Agents and Chemotherapy* **35**, 1824–1828 (1991).

81. Shan, Y. *et al.* ATP-Dependent Persister Formation in *Escherichia coli*. *mBio* **8**, e02267-16 (2017).

82. Allison, K. R., Brynildsen, M. P. & Collins, J. J. Metabolite-enabled eradication of bacterial persisters by aminoglycosides. *Nature* **473**, 216–220 (2011).

83. Lopatkin, A. J. *et al.* Bacterial metabolic state more accurately predicts antibiotic lethality than growth rate. *Nature Microbiology* vol. 4 2109–2117 (2019).

84. Yang, J. H., Bening, S. C. & Collins, J. J. Antibiotic efficacy — context matters. *Current Opinion in Microbiology* **39**, 73–80 (2017).

85. Stokes, J. M., Lopatkin, A. J., Lobritz, M. A. & Collins, J. J. Bacterial Metabolism and Antibiotic Efficacy. *Cell Metabolism* **30**, 251–259 (2019).

86. Dwyer, D. J. *et al.* Antibiotics induce redox-related physi. *Proceedings of the National Academy of Sciences of the United States of America* **111**, (2014).

87. Radlinski, L. & Conlon, B. P. Antibiotic efficacy in the complex infection environment. *Current Opinion in Microbiology* **42**, 19–24 (2018).

88. Yang, J. H. *et al.* A White-Box Machine Learning Approach for Revealing Antibiotic Mechanisms of Action. *Cell* **177**, 1649-1661.e9 (2019).

89. Farha, M. A., French, S., Stokes, J. M. & Brown, E. D. Bicarbonate Alters Bacterial Susceptibility to Antibiotics by Targeting the Proton Motive Force. *ACS Infectious Diseases* **4**, 382–390 (2018).

90. Farha, M. A. *et al.* Overcoming Acquired and Native Macrolide Resistance with Bicarbonate. *ACS Infectious Diseases* **6**, 2709–2718 (2020).

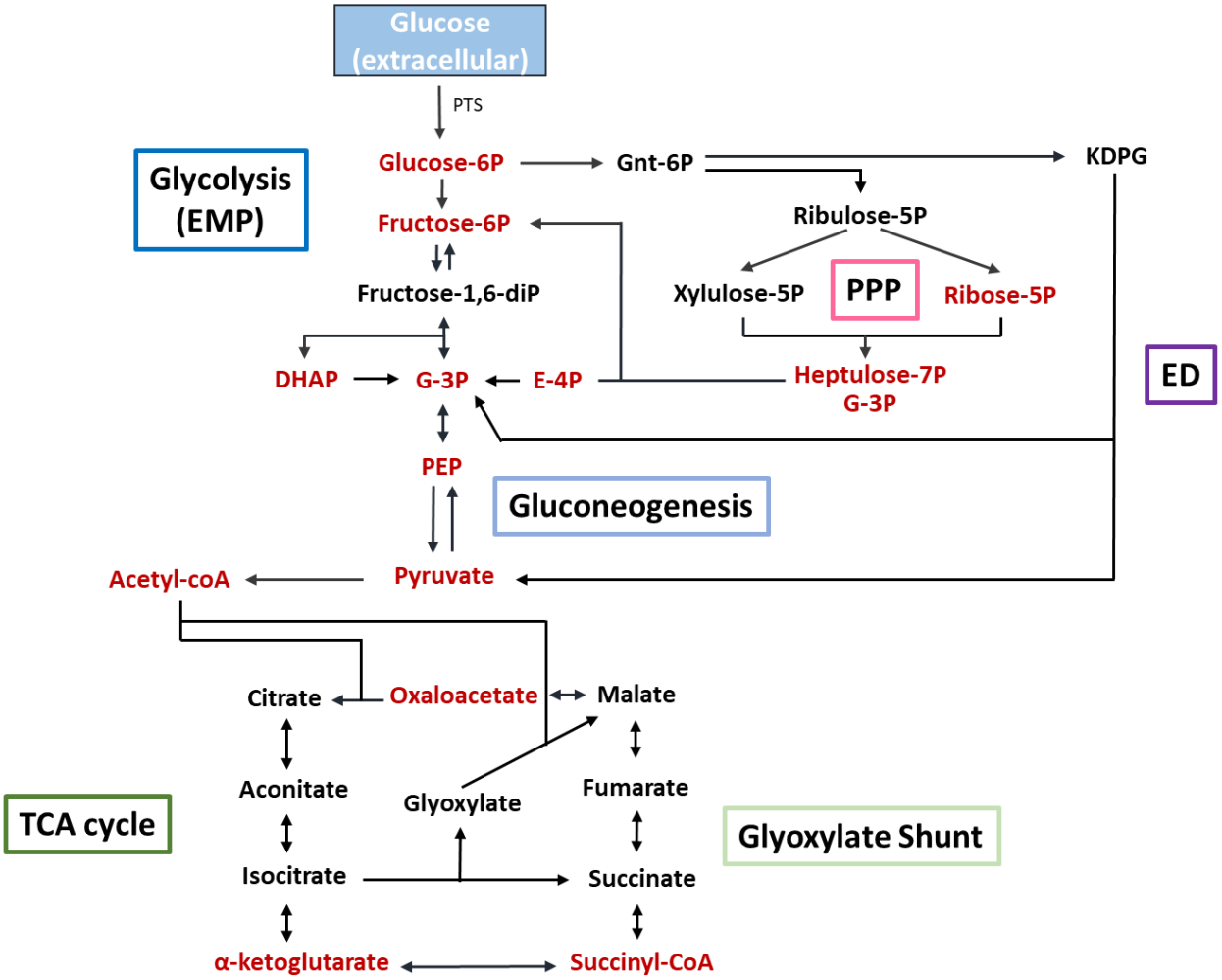
Figure Legends

Figure 1. Overview of central metabolism. Shown is a simplified overview of all the central metabolism pathways. The 13 precursor metabolites that are converted into biomass compounds are highlighted in red. NAG, N-acetyl glucosamine; P, phosphate;

DHAP, dihydroxyacetone phosphate; G-3P, glyceraldehyde-3-phosphate; E-4P, erythrose-4-phosphate; Gnt-6P, gluconate-6-phosphate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; EMP, Emden-Meyerhof-Parnas; ED, Entner Doudoroff; TCA, tricarboxylic acid; PPP, pentose phosphate pathway.

Figures

Figure 1



Tables

Table 1: Summary of the 13 precursor metabolites and their biosynthetic end products.

Precursor Metabolite	Pathway	Biosynthetic end products
D-glucose-6-phosphate	Glycolysis	Glycogen, LPS
D-fructose-6-phosphate	Glycolysis	Cell wall
dihydoxyacetone phosphate	Glycolysis	NAD
D-glycerate-3-phosphate	Glycolysis	Lipids
phosphoenolpyruvate	Glycolysis	Tyr, Trp
pyruvate	Glycolysis	Ala, Ile, Lys, Leu, Val
D-ribose-5-phosphate	Pentose Phosphate	His, Phe, Trp, Nucleotides
heptulose-7-phosphate	Pentose Phosphate	LPS
erythrose-4-phosphate	Pentose Phosphate	Phe, Trp, Tyr
α-ketoglutarate	TCA cycle	Glu, Gln, Arg, Pro
acetyl-CoA	TCA cycle	Leu, Lipids
succinyl-CoA	TCA cycle	Met, Lys
oxaloacetate	TCA cycle	Asn, Asp, Ile, Lys, Met, Thr, nucleotides

Chapter II - Gene Dispensability in *E. coli* Grown in Thirty Different Carbon Environments

Preface

The work presented in this chapter was previous published in:

Tong, M., French, S., El Zahed, S. S., Ong, W. K., Karp, P. D., & Brown, E. D. (2020). Gene dispensability in *Escherichia coli* grown in thirty different carbon environments. *MBio*, *11*(5), 1–20.

For this work I completed the screening, data analysis, development of the CarPE dashboard, all the experiments except for the accumulation assay and wrote the manuscript with input from French S. and Brown E.D. French S. wrote the script to analyze the image files and convert the colony sizes into integrated density values. The metabolic model predictions were done by Ong, W.K. Accumulation experiment was completed by El Zahed, S.S. Edits were provided by all authors.

Abstract

Central metabolism is a topic that has been studied for decades, and yet, this process is still not fully understood in *Escherichia coli*, perhaps the most amenable and well-studied model organism in biology. To further our understanding, we used a high throughput method to measure the growth kinetics of each of 3,796 *E. coli* single gene deletion mutants in 30 different carbon sources. In total, there were 342 genes (9.01%) encompassing a breadth of biological functions that showed a growth phenotype on at least one carbon source, demonstrating that carbon metabolism is closely linked to a large number of processes in the cell. We identified 74 genes that showed low growth in 90% of conditions, defining a set of genes which are essential in nutrient limited media, regardless of the carbon source. The data are compiled into a web-application, Carbon Phenotype Explorer (CarPE), to facilitate easy visualization of growth curves for each mutant strain in each carbon source. Our experimental data matched closely with the predictions from the EcoCyc metabolic model which uses flux balance analysis to predict growth phenotypes. From our comparisons to the model, we found that unexpectedly, phosphoenolpyruvate carboxylase (*ppc)* was required for robust growth in most carbon sources other than most TCA cycle intermediates. We also identified 51 poorly annotated genes that showed a low growth phenotype in at least one carbon source, which allowed us to form hypotheses about the functions of these genes. From this list, we further characterized the *ydhC* gene and demonstrate its role in adenosine efflux.

Importance:

While there has been much study of bacterial gene dispensability, there is a lack of a comprehensive genome-scale examinations of the impact of gene deletion on growth in different carbon sources. In this context, a lot can be learned from such experiments in the model microbe *E. coli* where much is already understood and there are existing tools for the investigation of carbon metabolism and physiology (1). Gene deletion studies have practical potential in the field of antibiotic drug discovery where there is emerging interest in bacterial central metabolism as a target for new antibiotics (2). Furthermore, some carbon utilization pathways have been shown to be critical to initiate and maintain infection for certain pathogens and sites of infection (3–5). Here, with the use of high throughput solid media phenotyping methods, we have generated kinetic growth measurements for 3,796 genes in 30 different carbon source conditions. This dataset provides a foundation for research that will improve our understanding of genes with unknown function, aid in predicting potential antibiotic targets, validate and advance metabolic models, and help to develop our understanding of *E. coli* metabolism.

Introduction

Central metabolism is an essential process that generates all of the energy and biosynthetic precursors required for bacterial survival. It is a highly regulated network of pathways that respond to extracellular concentrations of carbon and nutrients (6). *Escherichia coli* has a robust regulatory framework that allows it to best use the nutrients and carbon source available in its environment (7). This complex regulatory network has multiple levels of control mechanisms, including changes in gene expression through transcription factors and modulating enzyme activity with allosteric inhibition or post-translational modifications (8). For example, transcription factors, such as the cAMP receptor protein, regulate the usage of secondary carbon sources through a process called carbon catabolite repression (9). Carbon catabolite repression is thought to control the expression of as much as 10% of the genome (10). Using this process, *E. coli* genes are differentially regulated based on the growth medium, where specific genes become essential for bacterial survival. For example, the *E. coli* genome contains about 4,300 genes with 303 being essential for growth in the rich microbiological growth medium, LB (11, 12). However, gene essentiality is contextual and dependent on the growth condition (13). In nutrient limited media containing glucose or glycerol as a carbon source, an additional 119 genes become essential (11, 14).

The *E. coli* genome has been well studied; however, around 35% of its genes are still poorly characterized (1, 15). Identifying experimental phenotypes in gene knockout mutants is a powerful way to understand functions of specific gene products (16). For poorly annotated genes in the *E. coli* genome, such phenotypes can be used to

hypothesize the gene's function. Indeed, the development of high-throughput gene inactivation techniques has led to generation of an *E. coli* mutant collection (Keio collection) which facilitates genome-wide studies of phenotypes (11, 17, 18). Previous studies have successfully used systematic analyses of phenotypes to gain understanding of genetic interactions and to characterize genes with unknown functions. (16, 19–21)

To generate a large dataset of phenotypes, researchers have developed high-throughput experimental approaches. For example, phenotype microarrays (PMAs) can be used to evaluate the growth fitness of specific strains in hundreds of conditions (22, 23). Although PMAs have been applied to 1,400 strains from the Keio collection, it is a resource intensive process that is difficult to apply to larger libraries (24). Transposon mutagenesis and sequencing (TnSeq) studies have also been used to probe gene essentiality efficiently using one-pot selection approaches for a large number of conditions (19). While these studies generate a growth fitness score for each gene by monitoring the abundance of all mutants, they cannot generate individual kinetic measurements for each mutant. Additionally, the pooling of many mutants in a single culture has the risk of creating competition between strains making it difficult to assess the growth of a slow-growing strain across multiple conditions (25). Likewise, chemical complementation between strains can be confounding to these pooled approaches due to the provision of biosynthetic intermediates from one mutant to another (26). Pioneered first by those studying yeast genetics, arraying colonies on solid media has become a popular high-throughput method for large scale investigations of growth phenotypes (16, 27–29). This approach can be adjusted to high-throughput measurements of growth

fitness based on colony sizes of each strain. Such screens typically generate a growth fitness score for each strain. This information has successfully been used to understand gene function, chromosome organization, and provide insight to the mechanism of certain drugs (16, 27, 30).

Many studies have successfully used the *E. coli* gene knockout collection (Keio collection) to study metabolism (14, 20, 31). Phenotypic screens of this collection have studied a few carbon sources, to date, where the focus has been on validating metabolic models (14, 16, 32). *E. coli* has the remarkable ability to use many carbon sources for growth which makes it possible to exhaustively probe central metabolism pathways using different carbon sources. Since methodologies differ from lab to lab, unifying different gene essentiality datasets has been difficult (33). This speaks to the need for more comprehensive datasets that explore a large number of conditions using the same strains and methodology.

Herein, we expand on previous research and study the growth kinetics of each gene deletion mutant in the Keio collection, growing on 30 different carbon sources and probing comprehensively the pathways of central metabolism in *E. coli*. Using a chemically defined minimal medium (MOPS) and changing only the carbon source (34) we collected 24 hour growth kinetics for each deletion mutant growing on solid agar at high density, i.e., 1,536 colonies per plate. We compared these experimental data against the EcoCyc metabolic model that uses flux balance analysis to predict growth phenotypes and identified discrepancies which can help improve the accuracy of these models. This dataset presents an opportunity to investigate the gaps in knowledge of bacterial

metabolism and physiology. Our focus was on a lab strain of *E. coli* (BW25113), a representative K12 strain where extensive experimental, genome curation and metabolic modeling efforts could provide ample context for our investigations. Central metabolism is highly conserved, even between different domains of life (35, 36). Curated databases of metabolic pathways have been successfully used to predict metabolic networks of different organisms simply based on genome sequence, therefore, advancing our knowledge on a simple model microbe has profound implications in understanding metabolism as a whole (37). Here, we have identified trends in how *E. coli* growth rate changes in different carbon sources. Using this dataset, we investigated the function of the enigmatic *ydhC* gene and implicated it in adenosine efflux. Finally, we developed a web-application which can be easily used for visualization of growth curves for each mutant strain in each carbon source, allowing for a straightforward analysis of the dataset without programming expertise.

Results

A genome-wide screen of *E. coli* in different carbon sources

We first elected to do a broad survey of the carbon sources that *E. coli* is able to utilize for growth using the Biolog phenotype microarray (22). This commercially available platform contains 190 different carbon sources and can be used for the rapid identification of carbon sources for bacterial growth. We identified 74 carbon sources that could support the growth of our lab strain of *E. coli* BW25113 in MOPS minimal media (Figure S1). Based on these results, we narrowed the list to 30 carbon sources that feed different

pathways of central metabolism, highlighted in Figure 1a. We categorized each carbon source according to the pathway intermediate that *E. coli* metabolizes it to in central metabolism. The sugars which are catabolized to intermediates in the glycolysis pathway prior to glyceraldehyde-3-phosphate (G-3P) are categorized as upper glycolysis. Lower glycolysis carbon sources are metabolized to G-3P or pyruvate. Tricarboxylic acid (TCA) cycle carbon sources are intermediates within that pathway. Entner-Doudoroff carbon sources are metabolized through the metabolite 2-keto-deoxy-6-phosphogluconate (KDPG). Finally, pentose phosphate compounds are catabolized to metabolic intermediates in the pentose phosphate pathway (Figure 1a).

To evaluate the impact of each gene deletion on the growth of *E. coli* in each carbon source, we screened the Keio collection, an ordered library of single gene knockouts (11), in solid minimal media containing each carbon source. We completed this by arraying each mutant of the collection onto solid agar media in duplicate and monitoring the growth of each colony over the course of 24 hours (Figure S2). Final 24 h end-point growth was normalized using a method described by French et al. (27), and produced high-resolution colony arrays. We normalized each carbon source treatment individually, given the unique growth characteristics of *E. coli* observed in each carbon environment. These end-point biomass values formed a highly replicating final dataset of growth amplitudes for 3,796 genes tested across 30 different carbon sources (Figure S3, Table S1A). Indeed, the endpoint values were of great utility in understanding the trends in gene essentiality. The kinetic data, consisting of 72 growth measurements over 24 h, were used

to verify these growth patterns since the growth curve of *E. coli* differed depending on the carbon source.

We subjected our results to two-dimensional hierarchical clustering as a check on consistency of the data with known metabolism (Figure 1b). This produced a heat map where carbon sources (y axis) entering central metabolism in similar pathways clustered together, while on the x-axis, genes in similar pathways clustered together. For example, the carbon sources adenosine and thymidine are both nucleosides that clustered together based on their Keio collection genetic responses. TCA cycle intermediates, α-ketoglutarate, malate, succinate, and fumarate clustered together. On the x-axis, genes belonging to the same operon with similar biological function or subunits of the same protein also clustered closely together. Genes involved in thiamine biosynthesis (*thiHSCIFE*), despite being low growth in most conditions, clustered closely. As expected, *xylBARG* clustered closely as they are all required specifically for growth on xylose (Figure 1c). In *E. coli*, xylose is transported by the xylose ABC transporter. XylG is the ATP binding subunit of the transporter, while *xylHF* encodes the membrane subunit and periplasmic binding protein, respectively. As found in previous studies, we noticed that Δ*xylH* or Δ*xylF* single deletions were unable to prevent growth on xylose (38). Once in the cell, the sugar is isomerized by XylA into xylulose. XylB is a xylulokinase which converts xylulose into xylulose 5-phosphate where it can be further metabolized through the pentose phosphate pathway. XylR is a transcriptional activator that binds to the *xyl* promoters and is required for the expression of the *xylAB* (38). Interestingly, galactose clustered outside of all the other carbon sources because of gene deletions that improved

growth on that substrate. In particular, this included *galS* and *galR*, which are both DNA binding repressors of the *gal* regulon genes (39). These genes encode regulatory enzymes that repress expression of genes involved in galactose transport and metabolism pathways unless galactose is present (39). Together, hierarchical clustering showed trends that we would expect from our data based on prior knowledge in *E. coli* metabolism, indicating that our dataset was biological meaningful. Genes that are known to be in the same pathways showed a similar phenotypic pattern of growth across carbon sources.

Genes important for *E. coli* carbon metabolism

We were especially interested in identifying genes in *E. coli* required for robust growth on different carbon sources. Deletion mutants that had a growth defect when grown in a specific carbon source imply the function of the deleted gene is important for growth in that condition. To determine which gene deletions caused a low growth phenotype, we used a 3 standard deviation cut-off below the interquartile mean of our entire dataset (40) (Figure S3). On average, we identified 102 genes per carbon source that showed a growth defect, a number similar in scale to the 119 genes that were found to have an indispensable phenotype for growth in minimal media containing glucose or glycerol (Table 1) (11, 14). We note here that, while our experiments were conducted on solid media, this glucose dataset had a 98% agreement with the results of a previous study of the Keio collection grown in liquid cultures of MOPS glucose (Table S1B) (11). In total, there were 342 genes (9.01%) showing a low growth phenotype on at least one carbon source (Table S2A). These were largely enriched for genes involved in carbohydrate, amino acid, coenzyme, and nucleotide transport and metabolism. This is expected from a

screen done in minimal media, which requires bacteria to synthesize their own amino acids, nucleotides, and vitamins (Figure 2a). We also identified 51 genes with poorly annotated functions that showed a low growth phenotype in at least one carbon source (Table S2B). Finally, we identified 74 genes that showed low growth in 90% of conditions, defining a set of genes which are essential in all minimal media, regardless of the carbon source (Table S2C). It is interesting to note that while most of the genes with phenotypes are involved in metabolism, there were many genes that fell into different categories unrelated to metabolism. In all, our analysis demonstrated that carbon metabolism is closely linked to a large number of cellular processes.

To explore the functional genomics of carbon metabolism further, we compared the genes required for growth across the five different categories of carbon sources used in this study. For each of the categories described in Figure 1a, we generated a list of genes that showed a low growth phenotype when deleted in at least one carbon source in that group. We compared these gene lists with a Venn diagram analysis to compare which genes had growth phenotypes that were unique to certain carbon pathways and which were shared with multiple pathways (Figure 2b). As expected, genes encoding enzymes required to degrade a specific carbon source in a pathway showed specificity for that category. For example, the genes *edd*, *eda*, *uxaABCR*, and *idnK*, are all required to metabolize gluconate, glucoronate, and galacturonate and were all specific to the Entner-Doudoroff category (Table S3). Overall, 96 genes showed a low growth phenotype in at least one carbon source in each category (Table S3). These genes are not specific to carbon degradation pathways in central metabolism and are instead specific to the nutrient

biosynthetic pathways required in nutrient limited media. It is interesting to note that lower glycolysis carbon sources have significant overlap with the TCA cycle category as these carbon sources are closely linked in central metabolism.

The growth rate of *E. coli* varies depending on carbon source

It is well known that carbon source quality affects the growth rate of *E. coli*. We compared the growth rates of *E. coli* grown in different carbon sources in our screening conditions. In our screen, most conditions only showed around 100 phenotypes (Table 1), meaning most of the genes in *E. coli* are not directly associated with a carbon source phenotype. Based on this, we can assume that majority of mutants will grow the same as a wild-type (WT) strain. By using the interquartile mean (mean of the middle 50% of rank-ordered data for each condition – see methods) for every mutant at every time-point (n=1,898), we were able to generate a representative WT growth curve on each carbon source (Figure 3a). Indeed, this curve serves as an internal control and is representative of the exact conditions as every mutant in the screen. To confirm that this assumption was valid, we investigated the growth of WT *E. coli* in five different carbon sources and found close correspondence with the growth of the unperturbed mutants using the interquartile mean approach (Figure S3). Using this growth curve, we calculated the maximum growth rate of *E. coli* grown on each carbon source (Figure 3b). As expected, glucose as the sole carbon source resulted in the highest growth rate. For *E. coli*, catabolite repression is activated during growth on glucose to ensure it is preferentially metabolized. (9). In general, we found that *E. coli* grows well in glycolytic sugars and hexuronates that are degraded using the Entner-Doudoroff pathway. Oxaloacetate, fumarate, succinate, and

malate are all TCA cycle intermediates which had similar growth rates. Of note, L-alanine has a lower growth rate compared to D-alanine in *E. coli*. In order for *E. coli* to use L-alanine as the sole carbon source, it must first be converted into D-alanine through the activity of alanine racemase. This may be a bottleneck and account for the difference in growth rate. There are two alanine racemase enzymes in *E. coli*, which are encoded by the genes *alr* and *dadX*. While *alr* is constitutively expressed, *dadX* is the predominant isozyme that is responsible for most of the degradation activity in the cell (41). This is reflected in our data where the *Δalr* mutant is able to grow on L-alanine as the sole carbon source, but there is no growth from the *ΔdadX* mutant on L-alanine.

Carbon Phenotype Explorer

In our experiment, kinetic measurements were obtained at 20-minute intervals, adding up to over 8 million data points collected which generated 113,880 individual growth curves for each mutant in every condition. To facilitate the accessibility of this data, we created a web-application that allows for easy visualization of the growth curves for each gene deletion mutant (https://edbrownlab.shinyapps.io/CarPE/). In this application, users can visualize any of the growth curves that were generated in this study by selecting the gene and carbon sources of interest (Figure 4). There are two ways that we have plotted these data; one where the relative growth in each carbon source is on the same graph, and another where each mutant is compared to the expected WT curve for each carbon source. We observed that in certain carbon sources, the edge colonies grew larger than centre colonies. This is a phenomenon often seen in solid-media screens where neighbour effects due to competition for nutrients may occur (27, 28). To account for these edge effects,

colonies that were on the edge of the plate were instead plotted against the interquartile mean of all edge colonies in that condition.

Our data explorer also allows for visualization of the 24 h end-point growth values, which is a function of colony size (27). Users can select a carbon condition of interest and see the relative growth of all mutants represented in an index plot. This interactive plot allows the user to zoom into different sections of the graph and visualize the normalized growth value of each gene in a specific condition. Hovering above each point shows a tool-tip of the gene name and Blattner number (B-number) of the data point. Gene names can be selected and highlighted on the plot with labels so the user can see growth compared to the rest of the mutants. Based on a user-selected cut-off value, a table of gene names and their corresponding GO term annotations appear on a table underneath so users can see which gene mutants showed significantly low or high growth in each carbon source. Where valid, these genes are linked to the EcoCyc webpage which contains curated information about the gene (23). For comparing between conditions, we have included a function where the user can type in a gene and look at the normalized growth in each condition plotted as a bar graph. Since the mutants in the Keio collection contain a kanamycin cassette, this may cause polar effects on downstream genes (11). Using CarPE, the user can look up genes downstream of their gene of interest and determine whether they have a similar carbon profile. Considering these functionalities, this web-application allows other users to easily navigate and analyse our dataset.

Comparisons to the EcoCyc metabolic model

We wanted to evaluate whether the phenotypes that we observed matched the current knowledge of *E. coli* central carbon metabolism. We compared our experimental data with the predicted growth phenotypes using the EcoCyc metabolic model (23, 42). Using flux balance analysis, the media composition is considered and the model can predict growth in each gene knockout. To simulate a gene knockout in the model, reactions associated with that gene are compiled using the EcoCyc database. The model considers the relationships between a gene, its protein product, any complex(es) formed by the product, and the reaction(s) catalyzed by the product and/or complex. The model then simulates the experiment by attempting to use the remaining reactions to generate the full set of biomass metabolites required for growth. If any of these metabolites are no longer produced, or if the predicted growth rate is below our cutoff, the simulation will produce a no-growth result. There were 1,402 genes that overlapped between our dataset and the model (Table S4). On average, the model predicted the same phenotype as our experiments with 95% accuracy (Table 2). By comparing our data with the model, we can identify genes that show a growth discrepancy to focus on for further study. Genes that had a predicted growth rate above our cut-off of $1x10^{-3}$/hr were predicted to be dispensable in the model. If that same gene deletion mutant failed to show growth experimentally when deleted, it was considered a false positive prediction. Conversely, a false negative prediction means the model predicted no growth from the mutant but our experimental data showed growth. We identified 33 genes that were found to have either a false positive or negative prediction across all carbon sources (Figure S4). Of these

genes, seven were determined to be essential for growth in previous studies (12). An update to the Keio collection showed that these mutants contained a gene duplication which explains why these mutants grew in our experiment despite being essential for growth in rich media (43). We re-picked the remaining 26 genes and grew them in liquid media to determine whether these phenotypes were artifacts from high-throughput screening. Out of these, only nine genes that showed growth in solid media did not reconfirm in liquid media (Figure S4). These genes were all involved in vitamin biosynthesis pathways. It is interesting to note that apart from the Δ*pdxH* mutant, these mutants exhibited modest growth until they were passaged in minimal media, eliminating excess nutrients carried over from the previous media.

We also identified 198 genes that have a false prediction in at least one carbon source (Table S5). One of the discrepancies identified in this study is the inability of the model to correctly predict the growth patterns of a *Δppc* mutant. The enzyme encoded by *ppc*, phosphoenolpyruvate (PEP) carboxylase, catalyzes an anapleurotic reaction which converts PEP into oxaloacetate and replenishes it in the TCA cycle (44). The model predicts that the cell is able to grow without this enzyme by activating the glyoxylate shunt and generating oxaloacetate through that pathway instead (42). Our data show that the *Δppc* mutant can only grow when the sole carbon source is one of acetate, α-ketoglutarate, fumarate, malate, succinate, or ribose. This means that during growth in most carbon sources, this enzyme is vital to regulate the levels of PEP and oxaloacetate in the cell, but with different TCA cycle intermediates, notably not oxaloacetate, this is no longer necessary. Thus, by screening a large number of carbon sources, we have

characterized enigmatic patterns of regulation which may not be obvious in a smaller sample size of carbon sources. Comparing such data to a metabolic model can generate new insights for genes that have already been well studied.

YdhC is a putative transporter involved in adenosine export

Despite the many efforts to fully annotate every gene in *E. coli*, 35% of the genes still lack experimentally defined functions (45). We can use this dataset to generate hypotheses for the function of some of these genes. We identified 51 genes that were poorly annotated and exhibited a growth defect on at least one carbon source (Table S2B). One of these genes was *ydhC,* which codes for a putative transporter of the major facilitator superfamily (46). We noted slow growth for the *ΔydhC* mutant specifically when grown on adenosine as the sole carbon source (Figure 5a). This gene is found downstream of the gene *purR*, a regulator of purine biosynthesis. Interestingly, Sastry et al. (47) recently applied unsupervised learning to RNA-seq datasets and identified *ydhC* as a potential member of the *purR* regulon. We elected to follow up on *ydhC* to identify its involvement in adenosine metabolism.

To ensure the finding of slow growth for the *ydhC* mutant was not an error due to the high-throughput nature of our screen, we re-made the deletion mutant using homologous recombination with an apramycin resistant marker. We then checked to see if we could complement back this phenotype by cloning the gene onto a plasmid. Since there is no annotated promoter for *ydhC,* we cloned this gene, complete with a section 200 bp upstream, into a pBR322 plasmid. Compared to the empty plasmid vector, we found that

the growth of the Δ*ydhC* mutant returned to wild-type levels in the presence of this plasmid (Figure 5b).

Since YdhC is annotated as a transporter in the major facilitator superfamily (MFS), we postulated that YdhC may be involved in adenosine transport (46). Interestingly, two genes, *nupG* and *nupC*, have already been shown to code for nucleoside transporters. In our study, we found that *Δ*ydhC* has the strongest growth defect when using adenosine as a sole carbon source (Figure 5c). Consistent with both *nupG* and *nupC* coding for purine nucleoside transporters, a single deletion in either of these genes could be compensated by the presence of the other. The fact that the Δ*ydhC* showed such a strong growth defect by itself implied it was not a redundant function of the same adenosine transport system. Notably, a Δ*nupG*Δ*nupC* double mutant could not grow in media with adenosine as a sole carbon source (48). We reasoned that if *ydhC* were important in adenosine metabolism, it should be upregulated in the presence of adenosine. To test this hypothesis, we performed a RT-qPCR experiment to see if the *ydhC* transcript levels increased after treatment with adenosine. As a control, we also tested *nupG* and *nupC* transcript levels as these genes are expected to be upregulated in the presence of adenosine. Of course, when *E. coli* grows on glucose, there is catabolite repression and transport of other carbon sources is downregulated (9). To ensure our increased expression was not due to the lack of catabolite repression and was indeed specific to adenosine, we also tested glycerol as a control carbon source. We grew WT *E. coli* cells for two hours to mid-exponential phase, washed them, and then treated them for an hour in one of glucose, glycerol, or adenosine. We isolated the RNA from these cells and then measured the expression levels of the

genes *ydhC*, *nupG*, and *nupC*. We found that they all showed increased levels in adenosine compared to glucose and glycerol controls (Figure 5d).

Since YdhC has previously been implicated in arabinose efflux (49), we performed an accumulation assay to determine whether adenosine would be one of its efflux substrates (50). If YdhC is an efflux transporter, deletion of this protein would result in the increase of intracellular concentrations of adenosine since the cell would no longer excrete it. We grew *ΔydhC* and Δ*nupG*Δ*nupC* mutants to mid-log phase growth and treated them with adenosine for 10 min. The cells were washed and the cell lysates were then collected. We then used LC-MS/MS to quantify the amount of adenosine that was present in the cells. The Δ*ydhC* mutant accumulated more adenosine than the wild-type, implicating it in adenosine efflux (Figure 5e). This finding is a proof of principle example of the utility of our dataset in generating hypotheses for the functions of poorly annotated genes.

Discussion

Our primary goal was to generate a biologically meaningful dataset to further our understanding of *E. coli* carbon utilization patterns and central metabolism. Here, we have characterized the growth of each non-essential gene deletion mutant in *E. coli* grown on 30 different carbon sources. The carbon sources chosen probe all of the metabolic pathways within central metabolism and allow insights into aspects of *E. coli* physiology that are linked to carbon metabolism. The data collected in this study largely correlate with the EcoCyc metabolic model, which reflects our current knowledge in central metabolism. Nevertheless, we have noted several paradoxes and provide a list of genes

with discrepant phenotypes, highlighting gaps in our understanding and providing a starting point for advancing current metabolic models. We developed a web-based application that allows for easy visualization of the data for researchers interested in querying phenotypes and investigating genes of interest. We hope that this will facilitate the generation of hypotheses for the functions of enigmatic genes and help efforts in functional genomics related to carbon source utilization. In proof of principle efforts, we showed here how these phenotypes helped us to describe YdhC as an adenosine efflux transporter in the major facilitator superfamily.

Although *E. coli* can grow on a large number of carbon sources, the quality of the carbon sources will impact the growth rate (10). The preferential usage of glucose and the effect of carbon catabolite repression is well documented in previous studies (9). The latter is important for *E. coli* to grow with the preferred sugar when faced with multiple choices in a natural environment. Apart from glucose, however, the preferred carbon sources of *E. coli* are not well studied. A previous study looked at the activity of different sugar promoters in mixtures of carbon sources to determine a hierarchy of sugar utilization. The researchers noted that in a mix of two carbon sources, the promoter for the carbon source that supports a higher growth rate has higher activity while the less dominant sugar shows reduced activity (51). In this study we described a hierarchy of carbon source utilization based on growth rate. We saw that this was not dependent on which pathway of central metabolism the carbon source was metabolized to. Ribose, xylose, adenosine and thymidine are all metabolized into intermediates in the pentose phosphate pathway. However, we show that *E. coli* has different maximum growth rates on these carbon

sources. It would be interesting to combine different carbon sources systematically to discern how *E. coli* chooses its preferred carbon source given a mixture and whether this correlates with the growth rate in the sole carbon source. This would have implications in host infection environments where *E. coli* and other pathogens may encounter many different potential carbon sources.

Compared to other hexose sugars, we noticed that growth on galactose is slow. This observation was noted in a previous study characterizing the MG1655 strain of *E. coli* (52). Soupene *et al.* (52) adapted a culture of *E. coli* to grow quickly in galactose and found that the *lac* operon was induced in that culture. Since LacY, a lactose transporter, can also transport galactose, these mutations likely caused an increase of galactose uptake. They also identified an induction in the *nag* operon, a set of genes involved in N-acetylglucosamine metabolism, suggesting the cross-regulation between galactose and N-acetylglucosamine catabolic pathways. Qaidi et al. (53) further studied this phenotype and showed that NagC is involved in repressing *galP*, a gene that codes for galactose permease (53). In this study, we identified similar deletion mutants that showed improved growth including *galR*, *galS,* and *nagC*. All three of these gene products repress galactose transport, which means that deleting these genes increases galactose uptake into the cell. Based on these findings, the growth rate using galactose appears to be limited by its transport into the cell. The presence of galactose has been known to induce stress in *E. coli ΔgalE* (galactose epimerase) mutants (54). While our study only looked at galactose as the sole carbon source, other researchers have noted addition of galactose, even when another carbon source was present, can cause growth inhibition in these mutants (55). In

B. subtilis, accumulation of the toxic intermediate UDP-galactose imparts a growth defect on the cell (56). In *E. coli*, UDP-galactose is used in LPS biosynthesis, both in the core and O-antigen. *E. coli* K-12 strains do not synthesize O-antigen since there are mutations in the operon. Given that accumulation of UDP-galactose is toxic to the cell, *E. coli* that do not synthesize O-antigen may limit galactose transport into the cell since there are fewer pathways available in metabolizing UDP-galactose. Indeed, we noticed that clinical strains of *E. coli* that synthesize O-antigen have a higher growth rate on galactose than our K-12 strains (data not shown).

Our experimental data matched the predicted phenotypes from the EcoCyc metabolic model with 95.7% accuracy (Table 1). In total, we identified 198 genes with a false prediction in at least one carbon source. Resolving these inconsistencies is an opportunity to amend models that may be missing isozymes or contain incorrect annotations. Guzman et al. (57) recently followed up on a list of genes that showed a false positive prediction in a previous study. They showed that some of these mutants started growing when they increased the length of the experiment past the typical time-point of 24 hours. Furthermore, they highlight how studying inconsistencies between the model predictions and experiments can reveal important physiological features of *E. coli*, such as its genetic and metabolic flexibility in overcoming its gene deletions (57).

We initially identified a set of 33 genes which do not match computational predictions in any of the carbon sources tested. Out of these paradoxes, seven were identified in gene deletion strains of the Keio collection that have since been shown to be flawed because of gene duplication events (43). When we rescreened the remaining 26 mutants, nine of

them did not grow in liquid cultures of MOPS minimal media with glucose or glycerol as a carbon source. These nine mutants all corresponded to genes involved in vitamin biosynthesis. Vitamin requirements for *E. coli* growth can be extremely low, for example, an *E. coli ΔbioA* mutant requires only 0.2-0.4 ng/mL to restore growth (58). Because trace amounts of vitamins can supplement the growth of these auxotrophs, we suggest that this can lead to variable growth findings. One of the limitations of using solid media for our screen is the necessity of growing the mutant colonies in close proximity. Because of this, true auxotrophs with low nutrient requirements may be able to grow on the trace amounts released by neighbouring colonies. Nevertheless, this problem appears to be unique to genes involved in vitamin biosynthesis pathways since amino acid auxotrophs largely showed no growth in our screen. Of the 17 remaining paradoxical genes, we compared our findings for these to two other datasets, one using glucose (11) as the sole carbon source and another using glycerol (14). Our experimental findings for these 17 genes matched the growth patterns found in at least one of these other datasets (Figure S4). Some 11 of these deletion mutants showed inconsistent growth across datasets and as a result, it is difficult to determine their phenotypes in minimal media. Six of these genes (i.e., *pabC, zupT, kdsC, fpr, ubiC, and pyrI*) reproducibly showed growth in all three datasets. These genes were incorrectly predicted by the Ecocyc metabolic model to be essential on all carbon sources and resolving these inconsistencies may help improve the model. One of the genes, *zupT*, encodes a heavy metal divalent cation transporter and was predicted to be essential in all carbon sources. It was predicted to be essential because it was set to be the only reaction which allows the organism to obtain metal cations.

However, this is surely not the case as *E. coli* can grow in absence of this transporter. In all, a careful comparison of metabolic modelling with a large number of experimental phenotypes is a promising starting point for generating hypotheses about *E. coli* physiology and identifying missing reactions in the current metabolic models.

While metabolic models are good at predicting which genes are required, they sometimes fail to account for differences in regulation. In our experimental results, for example, we found that the *ΔsrlA* and *ΔsrlE* mutants were unable to grow in sorbitol as a sole carbon source. These genes code for subunits of the sorbitol specific PTS enzyme. The model predicts growth for these genes due to the presence of the *gatABC* genes which encodes the galactitol specific PTS enzyme. This enzyme is able to transport sorbitol with low affinity, however in some *E. coli* strains this gene is induced by galactitol and is likely not expressed at a high enough level when grown in sorbitol as a sole carbon source (59). Since the model is unable to account for regulation, it predicts sorbitol can enter the cell using a secondary transporter which may not be expressed in the given condition. We noticed that the *Δppc* mutant shows a growth defect in most carbon sources except for most of the TCA cycle intermediates: succinate, fumarate, malate, acetate, and α-ketoglutarate. In contrast, the flux balance model predicts growth through activation of the glyoxylate shunt to generate oxaloacetate in all conditions (60). Ppc catalyzes an anapleurotic reaction which generates oxaloacetate from phosphoenolpyruvate. During growth on most carbon sources, this reaction is required to replenish oxaloacetate as TCA cycle intermediates get used for biosynthesis of amino acids. However, when grown on TCA cycle intermediates, there is no longer a need to generate more oxaloacetate and Ppc

is not required for robust growth. Since there are other pathways to generate oxaloacetate, it is not readily obvious when this regulatory enzyme is required for growth. By testing a *Δppc* mutant for growth across carbon conditions, we have identified specific conditions where this enzyme is important for growth and revealed a new aspect of *E. coli* metabolism.

Metabolite efflux is well-described in the literature (49, 61–64). In 1980, Huber et al. (65) found that a large proportion of β-galactosidase products, galactose, glucose, and allolactose, were found in the culture media when lactose was added. Cells were found to be intact, and no β-galactosidase was detected in the media. Furthermore, this was only observed when the lactose permease was present and lactose can be imported into the cell. This implies that lactose was imported into the cell, broken down by β-galactosidase and the products were excreted into the media as they were generated during growth. Subsequently, Liu et al. (2002) described a family of β-galactoside sugar efflux pumps (*setABC*) in *E. coli* which explained this observation. For other sugars, studies have implicated YdeA and YdhC in arabinose efflux (49, 66). Interestingly, both arabinose and adenosine are metabolized into ribulose-5-phosphate in the pentose phosphate pathway. Previous studies have shown that accumulation of ribulose-5-phosphate results in a growth defect (67). This may explain why *E. coli* has a need for controlling intracellular levels of these carbon sources. In *E. coli,* gene *nepI* has recently been characterized to efflux inosine, setting a precedent for nucleoside export (62). However, unlike *ydhC,* a *ΔnepI* mutant did not appear to have a growth defect when grown in inosine or adenosine.

Current antibiotics target a small subset of cellular processes that are deemed to be essential, however essentiality is contextual and dependent on growth environment (13). There has been growing interest in targeting genes important under nutrient-restricted conditions that may also be required for infection *in vivo* (2, 6). Additionally, there has been interest in targeting bacterial carbon utilization pathways, as there is recognition that various carbon sources may have different importance at different sites of infection (2–4, 68, 69). In our study, we identified 342 genes that were conditionally essential dependent on carbon source available and 74 genes that were essential in MOPS minimal media regardless of carbon source. Where knowledge of the nutrient composition available to pathogens at various sites of infection is an emerging field of study (6, 58, 70, 71) the data generated herein may have value in understanding the set of unique targets that are practicable in the context of antibacterial drug discovery. This information could generate new hypotheses for drug target validation studies and could catalyze the creation of unique screening platforms to identify compounds that are uniquely antibacterial at specific sites of infections (72–74). Indeed, a carbon source and site-specific approach to antibacterial therapy might have the benefit of sparing commensal bacteria that use a different carbon source. Additionally, by understanding the genes required for growth that is carbon source-specific, we might avoid prioritizing targets that are not relevant *in vivo.* For example, *Mycobacterium tuberculosis* grown in standard laboratory conditions uses glycerol as a carbon source. A screen for anti-tubercular compounds was performed in this medium and identified compounds targeting glycerol metabolism which showed potent *in vitro* activity (74). However, since glycerol was not a relevant carbon source *in*

vivo for M. *tuberculosis*, the compounds lacked any *in vivo* efficacy (74). Understanding bacterial carbon source utilization therefore has foundational value in the hunt for new antibiotics. Similar studies in various commensals and pathogens will likewise provide a useful dataset with the capacity to identify targets that are pathogen-specific.

In conclusion, we have generated an easily-accessible set of data that will be broadly useful in microbiological research. This study provides a starting point for future investigations aimed at further understanding of bacterial central metabolism.

Materials and Methods

Chemicals

Chemicals used in this study were purchased from Sigma-Aldrich unless otherwise noted. A full list of carbon sources and the concentrations can be found on the carbon conditions tab in the Carbon Phenotype Explorer (https://edbrownlab.shinyapps.io/CarPE/). Concentrations were picked so that all carbon sources resulted in the same amount of carbon added. MOPS minimal media (Teknova) was used for all work in minimal media. Antibiotics were added to the medium as required with final concentrations as follows: 50 μg/mL kanamycin, 100 μg/mL ampicillin, 25 μg/mL chloramphenicol, 100 μg/mL apramycin.

Bacterial strains and growth conditions

In this study, we screened the Keio collection of *E. coli* K-12 non-essential gene deletions (Baba et al., 2006) and 100 small RNA and small protein deletions (Hobbs et al., 2010) in

30 different carbon sources. The sRNA and small protein deletion library were generated in *E. coli* strain MG1655 [F^- LAM^- *rph-1*]), while the Keio collection was in *E. coli* strain BW25113 (F^- Δ(*araD-araB*)*567 lacZ4787*Δ::*rrnB-3* LAM^- *rph-1* Δ(*rhaD-rhaB*)*568 hsdR514).*

Strains that did not show consistent growth across repeated screens were removed from this study. In our final dataset, we have gathered information on 3,796 strains of *E. coli.*

For routine experiments, *E. coli* BW25113 strain was streaked onto LB media with appropriate antibiotic selection from a frozen glycerol stock and grown overnight in a stationary incubator at 37°C. A single colony was picked to inoculate an overnight culture which was grown shaking at 250 rpm. This was used to subculture into a new tube to mid-log phase growth. For experiments in minimal media, this subculture was washed using MOPS minimal media containing no carbon source and diluted 1:100 in final conditions. Mutants were made by homologous recombination as described by Datsenko et al. (17) Gene deletions were made by replacing the gene with an apramycin or chloramphenicol resistant cassette. The plasmids, pSET152 and pKD3 were used as templates for the apramycin and chloramphenicol resistance genes respectively. Deletions were moved into different mutant strains to generate double deletions by using P1 phage transduction.

Biolog Phenotype Microarrays

Colonies were picked and re-suspended in M9 minimal media containing no carbon source. Phenotypic microarray assay plates (Biolog Inc., Hayward, CA) were inoculated

with mid-log phase culture of *E. coli* washed with MOPS minimal media with tetrazolium dye mix A (Biolog Inc., Hayward, CA). Plates were then incubated at 37°C in a stationary incubator for 24 h and 48 h. The OD_{600} values were read on the Tecan plate reader at both time-points. Plates were also scanned using the Epson V750 scanner to visualize the plate by eye.

Screening Conditions

The *E. coli* Keio collection was pinned onto LB agar medium containing 50 µg/mL kanamycin from frozen glycerol stocks in 96-density. Colonies were grown overnight at 37°C and upscaled to 384-density using a Singer Rotor HDA (Singer Instruments) and grown overnight at 37°C. From 384-density, the colonies were once again upscaled to 1,536-density and grown overnight. Plates were duplicated at 1,536-density in LB agar, grown overnight and the resulting plates were stored at 4°C for no longer than 4 weeks.

Keio plates were pinned from LB agar plates onto solid MOPS minimal media containing a carbon source in 1,536-density. Plates were incubated for 24 h to deplete any nutrients that may be carried over from LB agar. These plates were used to inoculate new solid minimal media assay plates containing the same carbon source. The assay plates were tested in two technical replicates, both at 1,536-density. Assay plates were incubated for 24 h at 37°C and scanned once every 20 minutes using the Epson Perfection V750 scanner. The workflow is summarized in Supplemental Figure 2.

Data Analysis

Images were analyzed using ImageJ and quantified into an integrated density value using the method described by French *et al.* (27) The integrated density value, a value that tracks with cell number linearly, was used as a proxy for growth at each time-point. The resulting data was compiled using a program written in R. For the end-point values, the first time-point was subtracted from the last time-point to remove any background noise caused by a large initial inoculum. After this, the data were normalized as described previously by French *et al.* to remove edge effects (27). The raw integrated density values of each colony were divided by the interquartile mean of the row, then the column it belonged in. Since corner colonies grew differently from the edges, we normalized them by dividing the raw data of each one to the interquartile mean of all the corner colonies in the condition.

To ensure data quality, the glucose condition was repeated every week to check that the growth of certain strains did not change over time and through subsequent passaging of the library. There were a few strains that developed mutations which allowed them to grow in minimal media after 4 weeks of passaging. As these genes did not replicate in the same condition they were removed from the final dataset. In addition to this, the genes that showed no growth in LB (the media that the collection was made and grown in) were also removed.

To determine which genes were essential in each condition, we used a cut-off that was 3 standard deviations below the mean of the dataset. Mutants that grew to a normalized

value below this were considered to show a low growth phenotype. Venn diagrams were drawn using a web tool developed by the bioinformatics and evolutionary genomics group at Ghent University. (http://bioinformatics.psb.ugent.be/webtools/Venn/). The COG annotation index was used to group genes according to functions. (https://www.ncbi.nlm.nih.gov/COG/) We manually curated genes based on current COG annotations and existing literature published on the genes. Genes were assigned to only one COG category.

Growth curves were corrected using a local regression method (loess). Maximum growth rates were calculated by finding the maximum slope of the line of best fit between five points during log-phase growth of the log-transformed growth curves. The heatmap and clustering were constructed using the heatmap.2 function from the 'gplots' package in R.

Carbon Phenotype Explorer (CarPE)

The interactive web-application developed for easy visualization of our data was built using the R 'shiny' and 'shinydashboard' packages. The interactive scatterplots were made using the 'scatterD3' package. To analyze our kinetic data, we subtracted the first time-point of each curve to minimize background noise from a high inoculum. We then took the average of our two replicates at each time-point. Given that most of our data normalize to the same growth value of one, we can make the assumption that most gene mutations do not affect the growth of *E. coli*. (40, 75). If we take the interquartile mean of each time-point, we can generate an expected curve for each carbon source. Since the low and high growing mutants are removed before taking the mean, this expected curve

represents how wild-type bacteria would grow in a given condition. To confirm this, we also analyzed a plate of WT colonies arrayed on 5 different carbon sources and compared these to the growth of the unperturbed mutants using the interquartile mean approach. Since different carbon conditions grew to different final colony sizes, we normalized the data by dividing each value by the maximum value of the expected curve. This normalizes the data from each time-point to a relative growth value based on what we expect a typical colony to grow to in the condition. This allows different conditions to be compared since curves are now normalized to a similar final amplitude.

Flux Balance Analysis (FBA)

We used the FBA-based EcoCyc metabolic model for *E. coli* strain BW25113 and the data from EcoCyc version 22.0 to obtain growth/no-growth predictions on each carbon source (23, 76). The cut-off for growth for the predictions was set to anything greater than $1x10^{-3}$/hr. The carbon source uptake rate was set to 10 mmol/gDW/hr in aerobic conditions. The nutrients provided were set to be the same as the ingredients of MOPS minimal media. There were 1427 genes in the model and they were tested against the 30 carbon sources used in this study.

Liquid Growth Kinetics

Phenotypes were reconfirmed by testing their growth in liquid media in a 96-well plate. Overnight incubations of the strain of interest were sub-cultured in LB and grown until an OD_{600} of 0.4. Cells were subsequently washed in MOPS minimal media with no carbon source and diluted 1:100 into the carbon source of interest. OD_{600} measurements were

taken every 15 minutes for up to 48 hours using a Sunrise absorbance microplate reader (Tecan Life Sciences).

Construction of complementation plasmid

The construction of the pBR322 plasmid containing *ydhC* was constructed using standard DNA manipulation techniques. DNA encoding the gene was obtained by PCR amplification from *E. coli* genomic DNA (Forward: 5' *atgacc*GGATCCGAACGCGATGCAGCTCCTGTG 3', Reverse: 5'*atgacc*GAATTCGTCTATAATCCGACGTAGAAC 3'). This plasmid was constructed by inserting the *ydhC* gene with 200bp upstream of the gene into the EcoRI and BamHI sites on the plasmid. Both the PCR products and the pBR322 plasmid were digested using EcoRI and BamHI enzymes, gel purified using the QIAgen gel extaction kit (QIAgen, Hilden, Germany) and ligated together using T4 ligase.

RNA Isolation, cDNA preparation and Quantitative RT-PCR

Bacterial cultures were grown to an OD of 0.4 in LB media. Cells were washed in MOPS minimal media containing no carbon source three times. The washed culture was split into three media containing different carbon sources: glucose, glycerol, and adenosine. These were incubated at 37°C for one hour before the cells were harvested and added to ice cold 5% acid phenol in ethanol solution. Cells were centrifuged for 10 minutes at 4°C and the resulting cell pellets were used for RNA extraction. RNA was extracted using RNAzol RT RNA isolation reagent. (77) cDNA was prepared using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA, USA).

Quantitative RT-PCR reactions were run using a CFX96 Read-Time System (Bio-Rad). Reactions (20µl) contained SYBR Green qPCR master mix (Dongsheng Biotech, Guangzhou, Guangdong, China), primers at a final concentration of 400 nM, and 5 ng of cDNA. 16S rRNA (rrsA) was used as a reference gene. The following primers were used for qPCR (5' – 3'): ydhC – F*:* GCAGAAATGGCAAGGCAAGC; ydhC – R: GGCAATCGCCATCACACAGA; nupG – F: GCCAGTGCACAAGGGATGTT; nupG – R: GAACCACGGAGTAACCAGCG; nupC – F: AGCATCTCCTTCCAGGGCAT; nupC – R: AGATGATGCCTTCAGCACGC; 16S – F: GAAGACTGACGCTCAGGTGC; 16S – R: GGGCCCCCGTCAATTCATTT. Cycling conditions were as follows: 95°C for 2 min, 40 cycles of 95°C for 15s, 62°C for 40s.

Accumulation Assay

The method for quantifying L-adenosine accumulation was adapted from Richter et al. (50) with some modifications. Following equilibration of the samples at 37°C with shaking for 5 minutes, cells were treated with 12 mM L-adenosine and incubated for another 10 minutes at 37°C with shaking. 800 µL of the treated cultures were layered on 700 µL of silicone oil (9:1 AR200/Sigma High Temperature, cooled at -80°C). Cells were pelleted and lysed, and the supernatant was collected as described by Richter et al. (50). Supernatants were lyophilized overnight with the Virtis BenchTop Pro lyophilizer (SP Scientific), re-suspended in methanol, and analysed by LC-MS. The accumulation assay was performed in biological triplicates, and each sample was injected twice into the LC-MS.

Samples were analysed with the LTQ Orbitrap XL MS system (Thermo Scientific) with a 1290 infinity series HPLC system (Agilent Technologies) including a degasser, an autosampler, and a binary pump. The liquid chromatography separation was performed on a HILIC-Z column (2.1 x 150 mm, 2.7 μm) (Agilent Technologies) with mobile phase A (10 mM ammonium acetate, pH 9.8) and mobile phase B (98% acetonitrile and 2% mobile phase A). The flow rate was 0.2 mL min^{-1}. The gradient was as follows: 0-2 minutes hold 15% mobile phase A; 2-17 minutes linear gradient to 80% mobile phase A; 17.1 minutes 15% mobile phase A and hold until 25.5 minutes. The injection volume was 10 μL. Mass spectra were acquired with positive electrospray ionization at the ion spray voltage of 3,900 V. The source temperature was 250°C. The sheath gas, auxiliary gas, and sweep gas were 40, 0, and 5 arbitrary units, respectively. Single ion monitoring of 268.1 m/z was used to quantify the metabolite.

Acknowledgements

This research was supported by a Foundation grant from the Canadian Institutes for Health Research (FRN 143215), by infrastructure funding from Canada Foundation for Innovation, and by a Tier I Canada Research Chair award to E.D.B., and by NIH grant GM077678 to P.D.K. We thank Hirotada Mori (Nara Institute of Science and Technology) for providing the Keio collection and Gisela Storz (NICHD, National Institutes for Health) for providing the sRNA and small peptide deletion collection used. We would like to thank Stephanie Nagy and Ian Paulsen (Macquarie University) for productive discussions about the transport mutants. We thank all the members of the Brown lab for helpful discussions, especially Michael Ellis for experimental advice.

References

1. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Schroder 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 Res 41:D605–D612.

2. Murima P, McKinney JD, Pethe K. 2014. Targeting Bacterial Central Metabolism for Drug Development. Chem Biol 21:1423–1432.

3. Alteri CJ, Himpsl SD, Mobley HLT. 2015. Preferential use of central metabolism in vivo reveals a nutritional basis for polymicrobial infection. PLoS Pathog 11.

4. Alteri CJ, Smith SN, Mobley HLT. 2009. Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle. PLoS Pathog 5:e1000448.

5. Rohmer L, Hocquet D, Miller SI. 2011. Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. Trends Microbiol 19:341–348.

6. Brown SA, Palmer KL, Whiteley M. 2008. Revisiting the host as a growth medium. Nat Rev Microbiol 6:657–66.

7. Martínez-Antonio A, Janga SC, Salgado H, Collado-Vides J. 2006. Internal-sensing machinery directs the activity of the regulatory network in *Escherichia coli*. Trends Microbiol 14:22–27.

8. Chubukov V, Gerosa L, Kochanowski K, Sauer U. 2014. Coordination of microbial metabolism. Nat Rev Microbiol 12:327–340.

9. Görke B, Stülke J. 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat Rev Microbiol 6:613–624.

10. Liu M, Durfee T, Cabrera JE, Zhao K, Jin DJ, Blattner FR. 2005. Global Transcriptional Programs Reveal a Carbon Source Foraging Strategy by *Escherichia coli*. J Biol Chem 280:15921–15927.

11. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. 2006. Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2:2006.0008.

12. Goodall ECA, Robinson A, Johnston IG, Jabbari S, Turner KA, Cunningham AF, Lund PA, Cole JA, Henderson IR. 2018. The Essential Genome of *Escherichia coli* K-12. MBio 9:e02096-17.

13. D'Elia MA, Pereira MP, Brown ED. 2009. Are essential genes really essential? Trends Microbiol 17:433–438.

14. Joyce AR, Reed JL, White A, Edwards R, Osterman A, Baba T, Mori H, Lesely SA, Palsson BØ, Agarwalla S. 2006. Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. J Bacteriol 188:8259–71.

15. Galperin MY, Koonin E V. 2010. From complete genome sequence to “complete” understanding? Trends Biotechnol 28:398–406.

16. Nichols RJ, Sen S, Choo YJ, Beltrao P, Zietek M, Chaba R, Lee S, Kazmierczak KM, Lee KJ, Wong A, Shales M, Lovett S, Winkler ME, Krogan NJ, Typas A, Gross CA. 2011. Phenotypic Landscape of a Bacterial Cell. Cell 144:143–156.

17. Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci USA 97:6640–6645.

18. Hobbs EC, Astarita JL, Storz G. 2010. Small RNAs and Small Proteins Involved in Resistance to Cell Envelope Stress and Acid Shock in *Escherichia coli*: Analysis of a Bar-Coded Mutant Collection. J Bacteriol 192:59–67.

19. Price MN, Wetmore KM, Waters RJ, Callaghan M, Ray J, Liu H, Kuehl J V., Melnyk RA, Lamson JS, Suh Y, Carlson HK, Esquivel Z, Sadeeshkumar H, Chakraborty R, Zane GM, Rubin BE, Wall JD, Visel A, Bristow J, Blow MJ, Arkin AP, Deutschbauer AM. 2018. Mutant phenotypes for thousands of bacterial genes of unknown function. Nature 557:503–509.

20. Nakahigashi K, Toya Y, Ishii N, Soga T, Hasegawa M, Watanabe H, Takai Y, Honma M, Mori H, Tomita M. 2009. Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. Mol Syst Biol 5:727–738.

21. Côté J-P, French S, Gehrke SS, MacNair CR, Mangat CS, Bharat A, Brown ED. 2016. The Genome-Wide Interaction Network of Nutrient Stress Genes in *Escherichia coli*. MBio 7.

22. Bochner BR. 2001. Phenotype MicroArrays for High-Throughput Phenotypic Testing and Assay of Gene Function. Genome Res 11:1246–1255.

23. Karp PD, Ong WK, Paley S, Billington R, Caspi R, Fulcher C, Kothari A, Krummenacker M, Latendresse M, Midford PE, Subhraveti P, Gama-Castro S, Muñiz-Rascado L, Bonavides-Martinez C, Santos-Zavaleta A, Mackie A, Collado-Vides J, Keseler IM, Paulsen I. 2018. The EcoCyc Database. EcoSal Plus 8.

24. Ito M, Baba T, Mori H, Mori H. 2005. Functional analysis of 1440 *Escherichia coli* genes using the combination of knock-out library and phenotype microarrays. Metab Eng 7:318–327.

25. Hibbing ME, Fuqua C, Parsek MR, Peterson SB. 2010. Bacterial competition: Surviving and thriving in the microbial jungle. Nat Rev Microbiol.

26. Seth EC, Taga ME. 2014. Nutrient cross-feeding in the microbial world. Front Microbiol 5:350.

27. French S, Mangat C, Bharat A, Côté J-P, Mori H, Brown ED. 2016. A robust platform for chemical genomics in bacterial systems. Mol Biol Cell 27:1015–25.

28. Takeuchi R, Tamura T, Nakayashiki T, Tanaka Y, Muto A, Wanner BL, Mori H. 2014. Colony-live —a high-throughput method for measuring microbial colony growth kinetics— reveals diverse growth effects of gene knockouts in *Escherichia coli*. BMC Microbiol 14:171.

29. Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Pagé N, Robinson M, Raghibizadeh S, Hogue CW, Bussey H, Andrews B, Tyers M, Boone C. 2001. Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science 294:2364–8.

30. Ding T, Case KA, Omolo MA, Reiland HA, Metz ZP, Diao X, Baumler DJ. 2016. Predicting essential metabolic genome content of niche-specific enterobacterial human pathogens during simulation of host environments. PLoS One 11:e0149423.

31. Long CP, Antoniewicz MR. 2014. Metabolic flux analysis of *Escherichia coli* knockouts: lessons from the Keio collection and future outlook. Curr Opin Biotechnol 28:127–133.

32. Monk JM, Lloyd CJ, Brunk E, Mih N, Sastry A, King Z, Takeuchi R, Nomura W, Zhang Z, Mori H, Feist AM, Palsson BO. 2017. iML1515, a knowledgebase that computes *Escherichia coli* traits. Nat Biotechnol. Nature Publishing Group.

33. Mackie A, Paley S, Keseler IM, Shearer A, Paulsen IT, Karp PD. 2014. Addition of *Escherichia coli* K-12 Growth Observation and Gene Essentiality Data to the EcoCyc Database.

34. Neidhardt FC, Bloch PL, Smith DF. 1974. Culture medium for enterobacteria. J Bacteriol 119:736–47.

35. Peregrin-Alvarez JM, Tsoka S, Ouzounis CA. 2003. The phylogenetic extent of metabolic enzymes and pathways. Genome Res 13:422–427.

36. Peregrín-Alvarez JM, Sanford C, Parkinson J. 2009. The conservation and evolutionary modularity of metabolism. Genome Biol 10:R63.

37. Caspi R, Altman T, Dreher K, Fulcher CA, Subhraveti P, Keseler IM, Kothari A, Krummenacker M, Latendresse M, Mueller LA, Ong Q, Paley S, Pujar A, Shearer AG, Travers M, Weerasinghe D, Zhang P, Karp PD. 2012. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 40.

38. Song S, Park C. 1997. Organization and regulation of the D-xylose operons in *Escherichia coli* K-12: XylR acts as a transcriptional activator. J Bacteriol 179:7025–32.

39. Semsey S, Krishna S, Sneppen K, Adhya S. 2007. Signal integration in the galactose network of *Escherichia coli*. Mol Microbiol 65:465–476.

40. Mangat CS, Bharat A, Gehrke SS, Brown ED. 2014. Rank ordering plate data facilitates data visualization and normalization in high-throughput screening. J Biomol Screen 19:1314–1320.

41. Wild J, Hennig J, Lobocka M, Walczak W, Kłopotowski T. 1985. Identification of the dadX gene coding for the predominant isozyme of alanine racemase in *Escherichia coli* K12. Mol Gen Genet 198:315–22.

42. Weaver DS, Keseler IM, Mackie A, Paulsen IT, Karp PD. 2014. A genome-scale metabolic flux model of *Escherichia coli* K–12 derived from the EcoCyc database. BMC Syst Biol 8:79.

43. Yamamoto N, Nakahigashi K, Nakamichi T, Yoshino M, Takai Y, Touda Y, Furubayashi

A, Kinjyo S, Dose H, Hasegawa M, Datsenko KA, Nakayashiki T, Tomita M, Wanner BL, Mori H. 2009. Update on the Keio collection of *Escherichia coli* single-gene deletion mutants. Mol Syst Biol 5:335.

44. Kameshita I, Tokushige M, Izui K, Katsuki H. 1979. Phosphoenolpyruvate Carboxylase of *Escherichia coli*: Affinity Labeling with Bromopyruvate1. J Biochem 86:1251–1257.

45. Ghatak S, King ZA, Sastry A, Palsson BO. 2019. The y-ome defines the 35% of *Escherichia coli* genes that lack experimental evidence of function. Nucleic Acids Res 47:2446–2454.

46. Saier MH, Beatty JT, Goffeau A, Harley KT, Heijne WHM, Huang SC, Jack DL, Jähn PS, Lew K, Liu J, Pao SS, Paulsen IT, Tseng TT, Virk PS. 1999. The major facilitator superfamily. J Mol Microbiol Biotechnol.

47. Sastry A V., Gao Y, Szubin R, Hefner Y, Xu S, Kim D, Choudhary KS, Yang L, King ZA, Palsson BO. 2019. The *Escherichia coli* Transcriptome Mostly Consists of Independently Regulated Modules. bioRxiv 620799.

48. Munch-Petersen A, Mygind B. 1976. Nucleoside transport systems in*Escherichia coli* K12: Specificty and regulation. J Cell Physiol 89:551–559.

49. Koita K, Rao C V. 2012. Identification and Analysis of the Putative Pentose Sugar Efflux Transporters in *Escherichia coli*. PLoS One 7.

50. Richter MF, Drown BS, Riley AP, Garcia A, Shirai T, Svec RL, Hergenrother PJ. 2017. Predictive compound accumulation rules yield a broad-spectrum antibiotic. Nature 545:299–304.

51. Aidelberg G, Towbin BD, Rothschild D, Dekel E, Bren A, Alon U. 2014. Hierarchy of non-glucose sugars in *Escherichia coli*. BMC Syst Biol 8:133.

52. Soupene E, Van Heeswijk WC, Plumbridge J, Stewart V, Bertenthal D, Lee H, Prasad G, Paliy O, Charernnoppakul P, Kustu S. 2003. Physiological Studies of *Escherichia coli* Strain MG1655: Growth Defects and Apparent Cross-Regulation of Gene Expression. J Bacteriol 185:5611–5626.

53. El Qaidi S, Allemand F, Oberto J, Plumbridge J. 2009. Repression of *galP* , the galactose transporter in *Escherichia coli* , requires the specific regulator of *N* -acetylglucosamine metabolism. Mol Microbiol 71:146–157.

54. Kalckar HM, Kurahashi K, Jordan E. 1959. HEREDITARY DEFECTS IN GALACTOSE METABOLISM IN *ESCHERICHIA COLI* MUTANTS, I. DETERMINATION OF ENZYME ACTIVITIES. Proc Natl Acad Sci U S A 45:1776–86.

55. Lee SJ, Trostel A, Le P, Harinarayanan R, Fitzgerald PC, Adhya S. 2009. Cellular stress created by intermediary metabolite imbalances. Proc Natl Acad Sci U S A 106:19515–20.

56. Chai Y, Beauregard PB, Vlamakis H, Losick R, Kolter R. 2012. Galactose metabolism plays a crucial role in biofilm formation by Bacillus subtilis. MBio 3:e00184-12.

57. Guzmán GI, Olson CA, Hefner Y, Phaneuf P V., Catoiu E, Crepaldi LB, Micas LG, Palsson BO, Feist AM. 2018. Reframing gene essentiality in terms of adaptive flexibility.

BMC Syst Biol 12:143.

58. Carfrae LA, MacNair CR, Brown CM, Tsai CN, Weber BS, Zlitni S, Rao VN, Chun J, Junop MS, Coombes BK, Brown ED. 2019. Mimicking the human environment in mice reveals that inhibiting biotin biosynthesis is effective against antibiotic-resistant pathogens. Nat Microbiol 1–9.

59. JH T, V N, JS E, Jr SM. 2001. The complete phosphotransferase system in *Escherichia coli*. J Mol Microbiol Biotechnol 3:329–346.

60. Peng L, Arauzo-Bravo MJ, Shimizu K. 2004. Metabolic flux analysis for a *ppc* mutant *Escherichia coli* based on ^{13}C-labelling experiments together with enzyme activity assays and intracellular metabolite measurements. FEMS Microbiol Lett 235:17–23.

61. Livshits VA, Zakataeva NP, Aleshin V V., Vitushkina M V. 2003. Identification and characterization of the new gene rhtA involved in threonine and homoserine efflux in *Escherichia coli*. Res Microbiol 154:123–135.

62. Gronskiy S V., Zakataeva NP, Vitushkina M V., Ptitsyn LR, Altman IB, Novikova AE, Livshits VA. 2005. The yicM (nepI) gene of *Escherichia coli* encodes a major facilitator superfamily protein involved in efflux of purine ribonucleosides. FEMS Microbiol Lett 250:39–47.

63. Nygaard P, Saxild HH. 2005. The purine efflux pump PbuE in Bacillus subtilis modulates expression of the PurR and G-box (XptR) regulons by adjusting the purine base pool size. J Bacteriol 187:791–4.

64. Liu JY, Miller PF, Gosink M, Olson ER. 1999. The identification of a new family of sugar efflux pumps in *Escherichia coli*. Mol Microbiol 31:1845–1851.

65. Huber RE, Pisko-Dubienski R, Hurlburt KL. 1980. Immediate stoichiometric appearance of beta-galactosidase products in the medium of *Escherichia coli* cells incubated with lactose. Biochem Biophys Res Commun 96:656–61.

66. Bost S, Silva F, Belin D. 1999. Transcriptional activation of ydeA, which encodes a member of the major facilitator superfamily, interferes with arabinose accumulation and induction of the *Escherichia coli* arabinose PBAD promoter. J Bacteriol 181:2185–91.

67. Englesberg E, Anderson RL, Weinberg R, Lee N, Hoffee P, Huttenhauer G, Boyer H. 1962. L-Arabinose-sensitive, L-ribulose 5-phosphate 4-epimerase-deficient mutants of *Escherichia coli*. J Bacteriol 84:137–46.

68. Fabich AJ, Jones SA, Chowdhury FZ, Cernosek A, Anderson A, Smalley D, McHargue JW, Hightower GA, Smith JT, Autieri SM, Leatham MP, Lins JJ, Allen RL, Laux DC, Cohen PS, Conway T. 2008. Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. Infect Immun 76:1143–1152.

69. Snider TA, Fabich AJ, Conway T, Clinkenbeard KD. 2009. *E. coli* O157:H7 catabolism of intestinal mucin-derived carbohydrates and colonization. Vet Microbiol 136:150–154.

70. Guirado E, Schlesinger LS. 2013. Modeling the Mycobacterium tuberculosis Granuloma - the Critical Battlefield in Host Immunity and Disease. Front Immunol 4:98.

71. Martinez-Jéhanne V, du Merle L, Bernier-Fébreau C, Usein C, Gassama-Sow A, Wane A-A, Gouali M, Damian M, Aïdara-Kane A, Germani Y, Fontanet A, Coddeville B, Guérardel Y, Le Bouguénec C. 2009. Role of deoxyribose catabolism in colonization of the murine intestine by pathogenic *Escherichia coli* strains. Infect Immun 77:1442–50.

72. Zlitni S, Ferruccio LF, Brown ED. 2013. Metabolic suppression identifies new antibacterial inhibitors under nutrient limitation. Nat Chem Biol 9:796–804.

73. Fahnoe KC, Flanagan ME, Gibson G, Shanmugasundaram V, Che Y, Tomaras AP. 2012. Non-Traditional Antibacterial Screening Approaches for the Identification of Novel Inhibitors of the Glyoxylate Shunt in Gram-Negative Pathogens. PLoS One 7:e51732.

74. Pethe K, Sequeira PC, Agarwalla S, Rhee K, Kuhen K, Phong WY, Patel V, Beer D, Walker JR, Duraiswamy J, Jiricek J, Keller TH, Chatterjee A, Tan MP, Ujjini M, Rao SPS, Camacho L, Bifani P, Mak PA, Ma I, Barnes SW, Chen Z, Plouffe D, Thayalan P, Ng SH, Au M, Lee BH, Tan BH, Ravindran S, Nanjundappa M, Lin X, Goh A, Lakshminarayana SB, Shoen C, Cynamon M, Kreiswirth B, Dartois V, Peters EC, Glynne R, Brenner S, Dick T. 2010. A chemical genetic screen in Mycobacterium tuberculosis identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy. Nat Commun 1:57.

75. Kim P-J, Lee D-Y, Kim TY, Lee KH, Jeong H, Lee SY, Park S. 2007. Metabolite essentiality elucidates robustness of *Escherichia coli* metabolism. Proc Natl Acad Sci U S A 104:13638–42.

76. Latendresse M, Krummenacker M, Trupp M, Karp PD. 2012. Construction and completion of flux balance models from pathway databases. Bioinformatics 28:388–396.

77. Chomczynski P, Wilfinger W, Kennedy A, Rymaszewski M, Mackey K. 2010. RNAzol® RT: a new single-step method for isolation of RNA. Nat Methods 7:4–5.

78. Tatusov RL. 2000. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 28:33–36.

Figure Legends

Figure 1: Summary of data collected in this study.

a. Carbon sources tested in this study. Carbon sources can enter central metabolism through different pathways. These were sorted into four different categories based on which metabolite it would be metabolized to. NAG, N-acetyl glucosamine; P, phosphate; DHAP, dihydroxyacetone phosphate; G-3P, glyceraldehyde-3-phosphate; E-4P, erythrose-4-phosphate; Gnt-6P, gluconate-6-phosphate; KDPG, 2-keto-3-deoxy-6-phosphogluconate.

b. Heatmap visualization of the normalized growth at 24 h of each gene in every carbon source. Each line on the x-axis represents the growth of a gene deletion mutant. The y-axis shows the carbon source the mutant was grown in. Red means lower growth compared to the expected growth and blue means higher growth than expected.

c. An example of a group of genes involved in xylose metabolism that clustered in the heatmap shown in c. Genes *xylB, A, R* and *G* code for proteins involved in xylose transport, regulation, and metabolism and showed growth defects specifically on xylose.

Figure 2: Analysis of the gene deletions leading to growth defects.

a. Functional analysis of genes with growth defects in any carbon source. Each gene was categorized into a single Cluster of Orthologous Groups (78) category. COG categories were based on the designations given in the current database and genes without a category were manually curated based on current known functions. The genes in the R category are

those that have predicted functions from homology or from experimental study. In the S category are genes that don't have any predicted functions. The bars represent the percentage of genes that belong in that category while the numbers show the number of genes in that category.

b. The Venn diagram assigns the 342 genes required for growth in any carbon source to the five pathways shown. Each carbon source was categorized according to that described in Figure 1a. Genes in a specific category means it was required for growth by at least one carbon source in that pathway. The names of the genes contained in each of the overlapping regions are specified in Table S3.

Figure 3: Dynamic analysis of growth in various carbon sources.

a. The growth curves of *E. coli* for each carbon source were calculated by taking the interquartile mean (40) of integrated density of every colony at each time point. We have called this the WT phenotype because it is generated from many strains that are unperturbed for growth by mutation. Each carbon source is unique in how they affect the growth of *E. coli.*

b. Average growth rate was calculated for each carbon source based on these growth curves in panel a by taking the maximum slope of the log-transformed graph. The rates are plotted from slowest to fastest, left to right.

Figure 4: Carbon Phenotype Explorer (CarPE) web-application. Example screen shots for the Carbon Phenotype Explorer web-application are shown

(https://edbrownlab.shinyapps.io/CarPE/). Each of the example plots were made based on the phenotypes for the gene *ppc*.

a. Shown are the sidebar links to the different visualization tools.

b. An example plot showing the growth curves of the *Δppc* deletion mutant in each of the 30 carbon sources. Users can select which carbon source conditions to show and download the data used to make such plots.

c. This panel shows that the growth curves of the mutant can be compared individually with the plot of the WT control for each carbon source.

d. Here, a control panel is shown that allows users to change the parameters of the index plot (panel e) and highlight genes of interest.

e. An example of an index plot is shown. The index plot shows the growth of every mutant strain in a carbon condition. The genes in the index plot are in the order our collection was arrayed, showing that positional effects were normalized out. Users can hover over the plot to identify the gene name and B-number.

f. A bar plot that shows the relative growth of the selected mutant strain across all carbon source conditions.

Figure 5: Hypothesis generation and analysis of the function of gene *ydhC*. a. Growth profile of the *ΔydhC* mutant in all carbon sources, highlighting the specific growth defect when adenosine is the carbon source.

b. Growth curves in MOPS minimal media containing adenosine as a carbon source. The growth defect of a *ΔydhC* mutant disappears when complemented with a plasmid containing the gene. Shown a representative example of many replicates.

c. Comparison of the final growth amplitudes of three deletion mutants, Δ*nupC*, Δ*nupG* and Δ*ydhC*, that have been implicated in adenosine transport.

d. qRT-PCR data showing the fold change in response (ΔΔCt values) for each gene (*nupC*, *nupG* and *ydhC*), relative to that in glucose, when cultured in glycerol and adenosine respectively. The reference gene was *E. coli* 16s rRNA.

e. Accumulation assays measuring the amount of adenosine in cells that were treated with adenosine for 10 minutes. Cells were lysed and adenosine levels were measured using LC-MS as described in Materials and Methods.

Figure S1. The biology phenotype microarray was performed on *E. coli* BW25113 using MOPS minimal media. Purple wells indicate a reaction with the tetrazolium dye and showing cellular respiration. These wells contain carbon sources that can be used by this strain of *E. coli.*

Figure S2. Workflow diagram describing the process of screening the *E. coli* single gene knockout collection. Using the Singer RoTor, plates containing the library were first grown on LB agar. Then they were pinned on to a minimal media plate and grown for 24 h to deplete any nutrients that may have been carried over. These plates were then used to inoculate two new assay plates as technical replicates. Plates were scanned every 20 min. to get a total of 73 images of the same plate at different time points over the course of 24

hours. Plates were analyzed using ImageJ in order to determine the integrated density (approximates the amount of bacteria) of each colony at each time point.

Figure S3: Summary of QC experiments for checking data quality

a. A scatterplot comparing the normalized integrated density, a value based on the final colony size relative to the rest of the condition, between two technical replicates. A correlation coefficient of 0.911 demonstrates highly reproducible results.

b. Histogram showing the distribution of all the normalized growth values based on the final colony sizes in our dataset. The red dotted line shows the cut-off of 3 standard deviations from the mean. Some 342 mutants showed growth below that cut-off in any carbon source.

c. We grew two plates of WT *E. coli* pinned in 1534-density in 5 different carbon sources: fumarate, glucose, pyruvate, sorbitol, and xylose. The black line corresponds to the average growth curve of the WT *E. coli* and the red line represents the growth curve of the unperturned Keio mutants.

Figure S4: Flow chart comparing the 33 genes we have identified in this study that do not fit in the model in any condition. Seven of the mutants contain a gene duplication which allows the mutant to grow. Nine of the genes were involved in vitamin biosynthesis and did not grow in liquid media. The remaining 17 genes were compared to two other gene essentiality datasets previously published (11, 14, 43).

Table Legends

Table S1: A) End-point biomass dataset B) Comparison to previous MOPS glucose dataset

Table S2: A) 342 genes with at least one phenotype B) 74 genes with low-growth in 90% of

conditions C) 51 poorly annotated genes with phenotypes

Table S3: Genes in Venn Diagram Intersections (Figure 2b)

Table S4: 1402 genes in the EcoCyc model compared to experimental results

Table S5: 198 genes with at least one false prediction in any carbon source

Table S6: Accuracy of the model in predictions for each carbon source

Supplemental tables can be accessed in the published journal:

https://journals.asm.org/doi/10.1128/mbio.02259-20

Figures

Figure 1

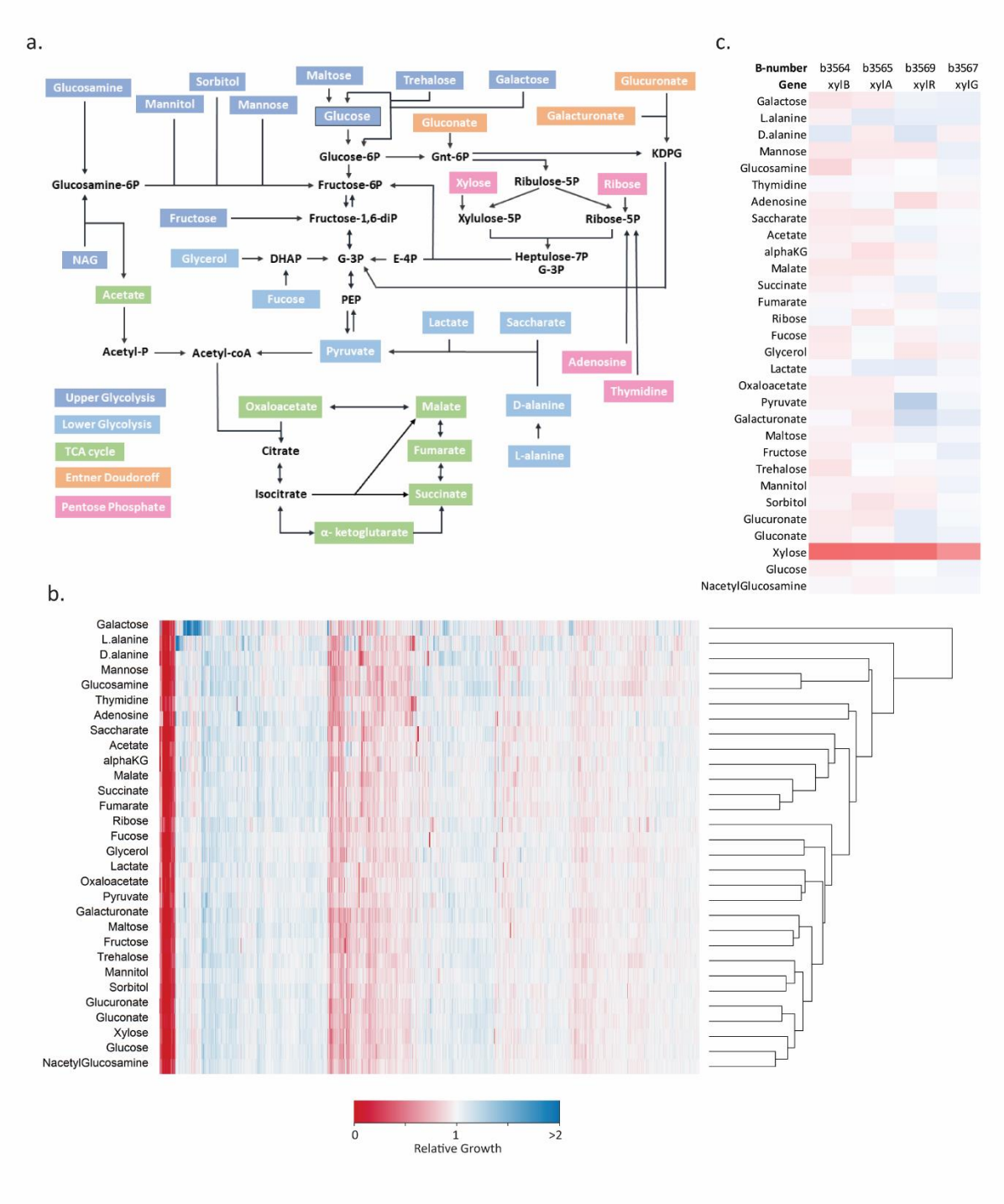


Figure 2

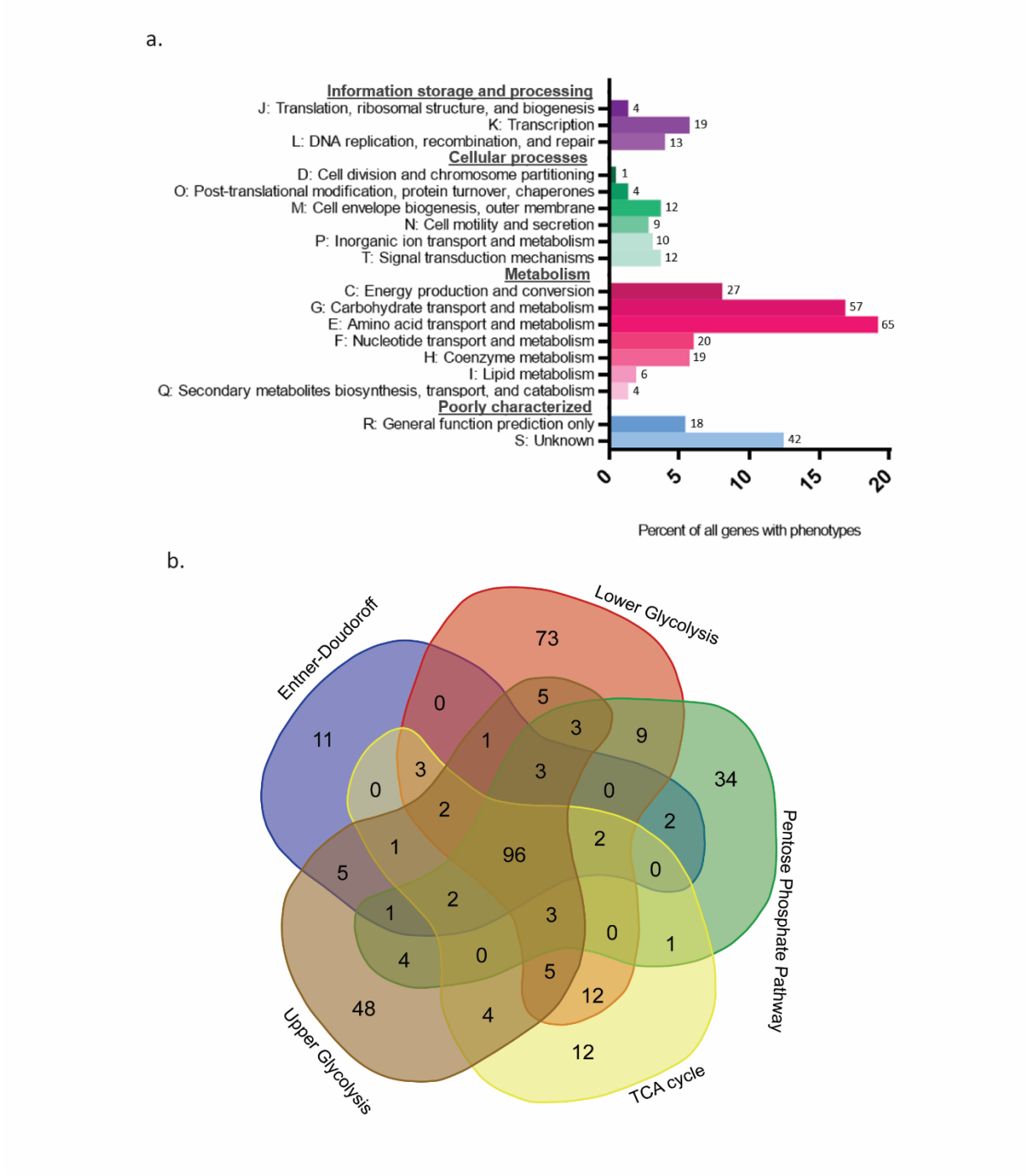

a.
Information storage and processing
J: Translation, ribosomal structure, and biogenesis
4
K: Transcription
19
L: DNA replication, recombination, and repair
13
Cellular processes
D: Cell division and chromosome partitioning
1
O: Post-translational modification, protein turnover, chaperones
4
M: Cell envelope biogenesis, outer membrane
12
N: Cell motility and secretion
9
P: Inorganic ion transport and metabolism
10
T: Signal transduction mechanisms
12
Metabolism
C: Energy production and conversion
27
G: Carbohydrate transport and metabolism
57
E: Amino acid transport and metabolism
65
F: Nucleotide transport and metabolism
20
H: Coenzyme metabolism
19
I: Lipid metabolism
6
Q: Secondary metabolites biosynthesis, transport, and catabolism
4
Poorly characterized
R: General function prediction only
18
S: Unknown
42
0
5
10
15
20
Percent of all genes with phenotypes
b.
Entner-Doudoroff
Lower Glycolysis
Pentose Phosphate Pathway
Upper Glycolysis
TCA cycle
11
73
34
48
12
96
0
5
3
9
1
3
0
2
2
1
5
2
1
3
0
1
4
0
5
12
4
12

Figure 3

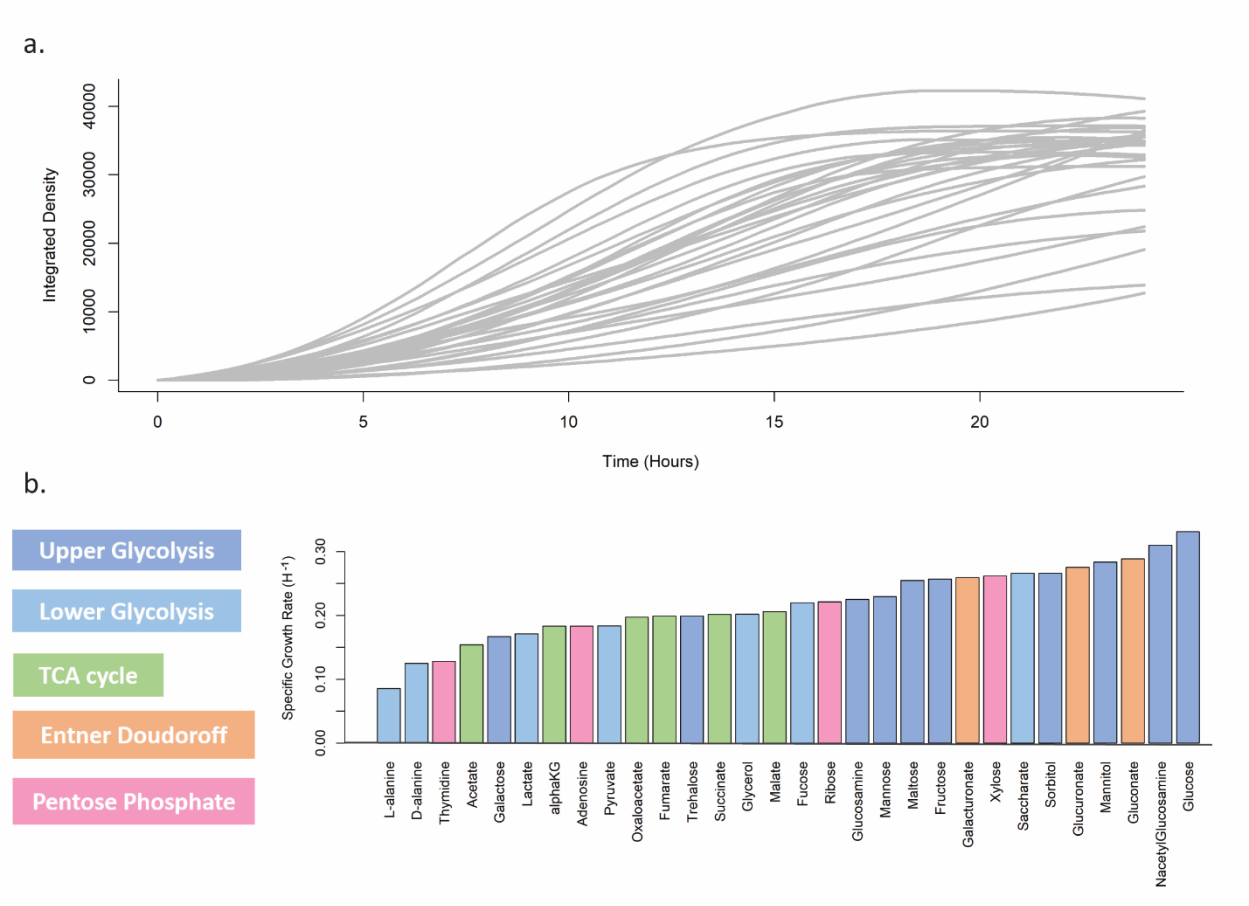

a.
Integrated Density
0
10000
20000
30000
40000
0
5
10
15
20
Time (Hours)
b.
Upper Glycolysis
Lower Glycolysis
TCA cycle
Entner Doudoroff
Pentose Phosphate
Specific Growth Rate (H-1)
0.00
0.10
0.20
0.30
L-alanine
D-alanine
Thymidine
Acetate
Galactose
Lactate
alphaKG
Adenosine
Pyruvate
Oxaloacetate
Fumarate
Trehalose
Succinate
Glycerol
Malate
Fucose
Ribose
Glucosamine
Mannose
Maltose
Fructose
Galacturonate
Xylose
Saccharate
Sorbitol
Glucuronate
Mannitol
Gluconate
Nacety/Glucosamine
Glucose

Figure 4

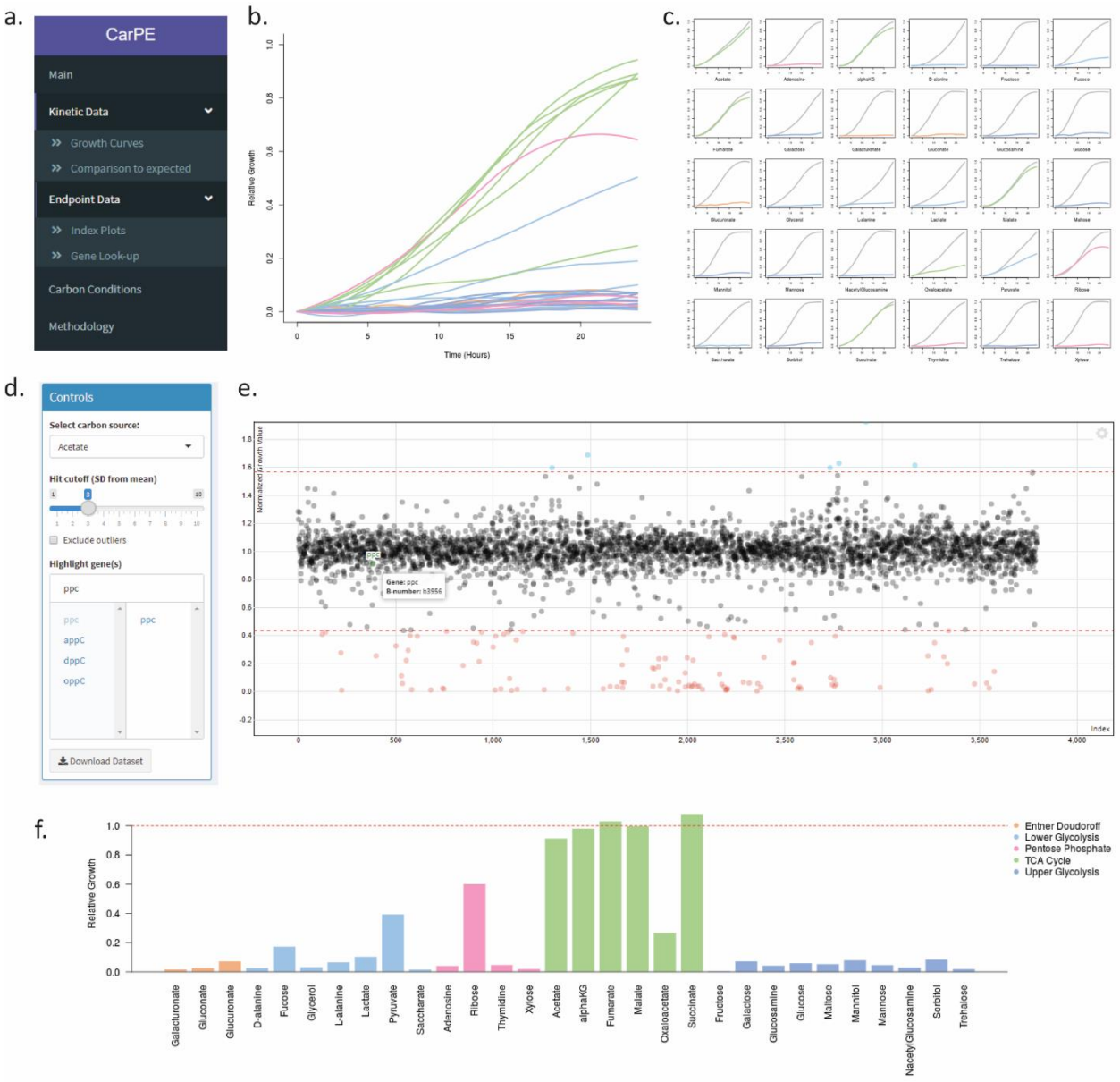

a.
CarPE
Main
Kinetic Data
Growth Curves
Comparison to expected
Endpoint Data
Index Plots
Gene Look-up
Carbon Conditions
Methodology
b.
Relative Growth
Time (Hours)
c.
d.
Controls
Select carbon source:
Acetate
Hit cutoff (SD from mean)
Exclude outliers
Highlight gene(s)
ppc
appC
dppC
oppC
Download Dataset
e.
Gene: ppc
B-number: b3956
Index
f.
Relative Growth
Entner Doudoroff
Lower Glycolysis
Pentose Phosphate
TCA Cycle
Upper Glycolysis
Galacturonate
Gluconate
Glucuronate
D-alanine
Fucose
Glycerol
L-alanine
Lactate
Pyruvate
Saccharate
Adenosine
Ribose
Thymidine
Xylose
Acetate
alphaKG
Fumarate
Malate
Oxaloacetate
Succinate
Fructose
Galactose
Glucosamine
Glucose
Maltose
Mannitol
Mannose
NacetylGlucosamine
Sorbitol
Trehalose

Figure 5

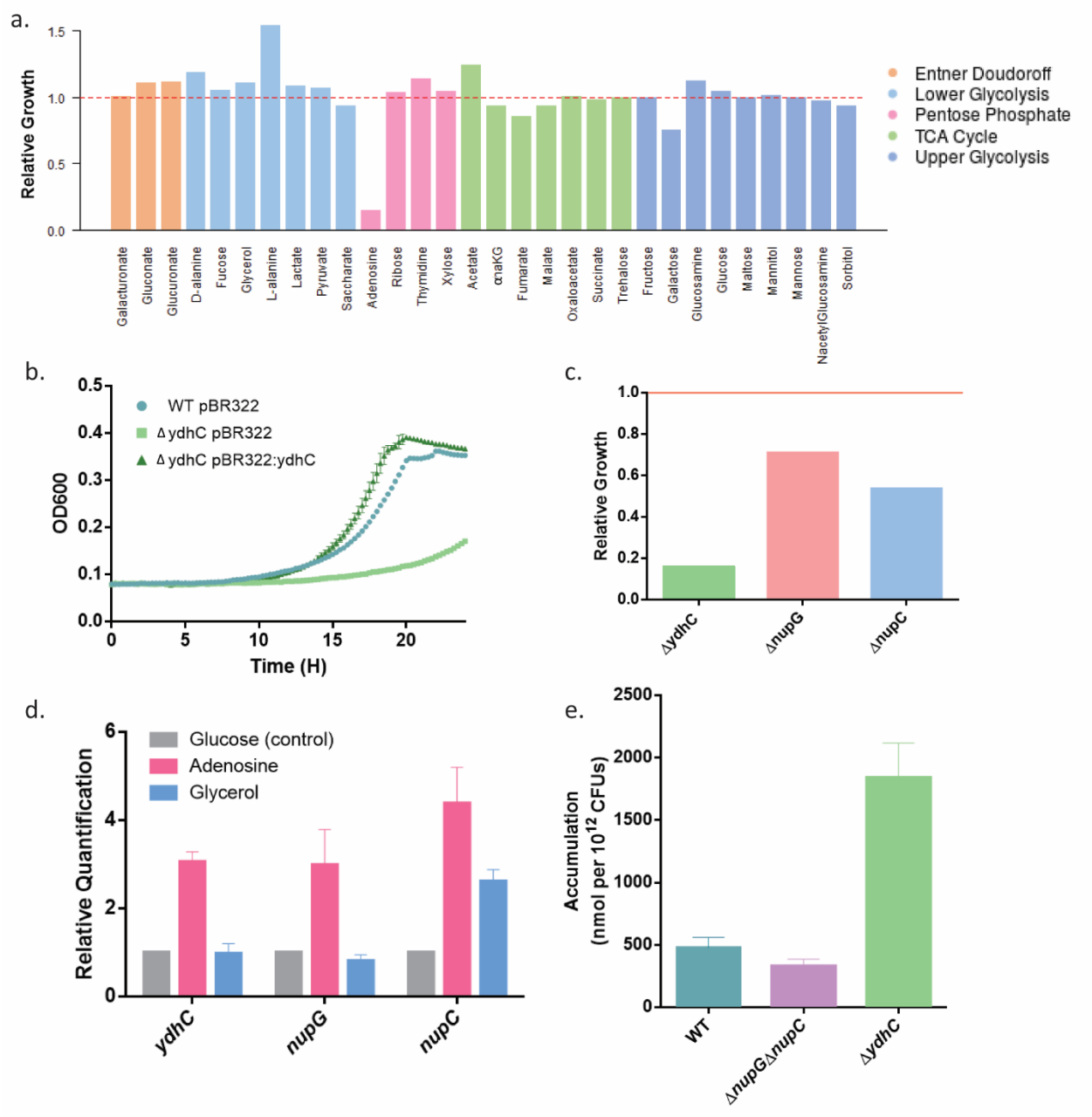
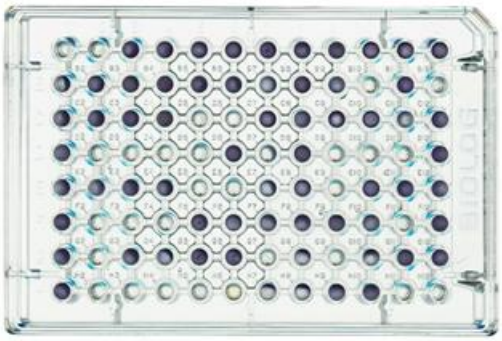
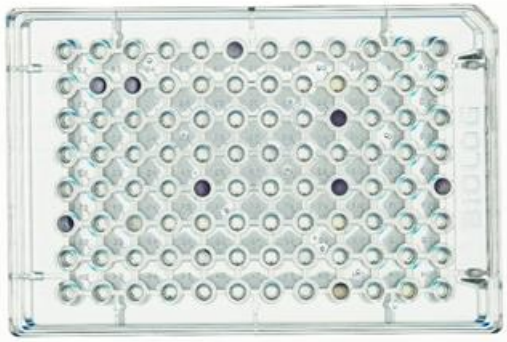

a.
Relative Growth
1.5
1.0
0.5
0.0
Entner Doudoroff
Lower Glycolysis
Pentose Phosphate
TCA Cycle
Upper Glycolysis
Galacturonate
Gluconate
Glucuronate
D-alanine
Fucose
Glycerol
L-alanine
Lactate
Pyruvate
Saccharate
Adenosine
Ribose
Thymidine
Xylose
Acetate
αnaKG
Fumarate
Malate
Oxaloacetate
Succinate
Trehalose
Fructose
Galactose
Glucosamine
Glucose
Maltose
Mannitol
Mannose
NacetylGlucosamine
Sorbitol
b.
OD600
0.5
0.4
0.3
0.2
0.1
0.0
WT pBR322
Δ ydhC pBR322
Δ ydhC pBR322:ydhC
0
5
10
15
20
Time (H)
c.
Relative Growth
1.0
0.8
0.6
0.4
0.2
0.0
ΔydhC
ΔnupG
ΔnupC
d.
Relative Quantification
6
4
2
0
Glucose (control)
Adenosine
Glycerol
ydhC
nupG
nupC
e.
Accumulation (nmol per 10^12 CFUs)
2500
2000
1500
1000
500
0
WT
ΔnupGΔnupC
ΔydhC

Figure S1





Biolog Plate	Carbon Source	Well
P1	N-Acetyl-D-Glucosamine	A03
P1	D-Saccharic Acid	A04
P1	Succinic Acid	A05
P1	D-Galactose	A06
P1	L-Aspartic Acid	A07
P1	L-Proline	A08
P1	D-Alanine	A09
P1	D-Trehalose	A10
P1	D-Mannose	A11
P1	Dulcitol	A12
P1	D-Serine	B01
P1	D-Sorbitol	B02
P1	Glycerol	B03
P1	L-Fucose	B04
P1	D-Glucuronic Acid	B05
P1	D-Gluconic Acid	B06
P1	D,L-α-Glycerol Phosphate	B07
P1	D-Xylose	B08
P1	L-Lactic Acid	B09
P1	D-Mannitol	B11
P1	D-Glucose-6-Phosphate	C01
P1	D-Galactonic Acid-γ-Lactone	C02
P1	D,L-Malic Acid	C03
P1	D-Ribose	C04
P1	D-Fructose	C07
P1	Acetic Acid	C08
P1	α-D-Glucose	C09
P1	Maltose	C10
P1	D-Melibiose	C11
P1	Thymidine	C12
P1	L-Asparagine	D01
P1	α-Keto-Glutaric Acid	D06
P1	α-Methyl-DGalactoside	D08
P1	Uridine	D12
P1	L-Glutamine	E01
P1	m-Tartaric Acid	E02

Biolog Plate	Carbon Source	Well
P1	D-Glucose-1- Phosphate	E03
P1	D-Fructose-6-Phosphate	E04
P1	α-Hydroxy Butyric Acid	E07
P1	β-Methyl-D-Glucoside	E08
P1	Maltotriose	E10
P1	2-Deoxy Adenosine	E11
P1	Adenosine	E12
P1	Glycyl-L-Aspartic Acid	F01
P1	Fumaric Acid	F05
P1	Bromo Succinic Acid	F06
P1	Propionic Acid	F07
P1	Mucic Acid	F08
P1	Glycolic Acid	F09
P1	Glyoxylic Acid	F10
P1	Inosine	F12
P1	Glycyl-L-Glutamic Acid	G01
P1	L-Serine	G03
P1	L-Threonine	G04
P1	L-Alanine	G05
P1	L-Alanyl-Glycine	G06
P1	N-Acetyl-β-DMannosamine	G08
P1	Methyl Pyruvate	G10
P1	D-Malic Acid	G11
P1	L-Malic Acid	G12
P1	Glycyl-L-Proline	H01
P1	Glucuronamide	H07
P1	Pyruvic Acid	H08
P1	L-Galactonic Acid-γ-Lactone	H09
P1	D-Galacturonic Acid	H10
P2	Dextrin	A06
P2	N-AcetylNeuraminic Acid	B02
P2	β-D-Allose	B03
P2	β-Methyl-DGlucuronic Acid	C09
P2	D-Glucosamine	E05
P2	γ-Hydroxy Butyric Acid	E09
P2	5-Keto-DGluconic Acid	E12
P2	D-Lactic Acid Methyl Ester	F01
P2	Dihydroxy Acetone	H09

Figure S2

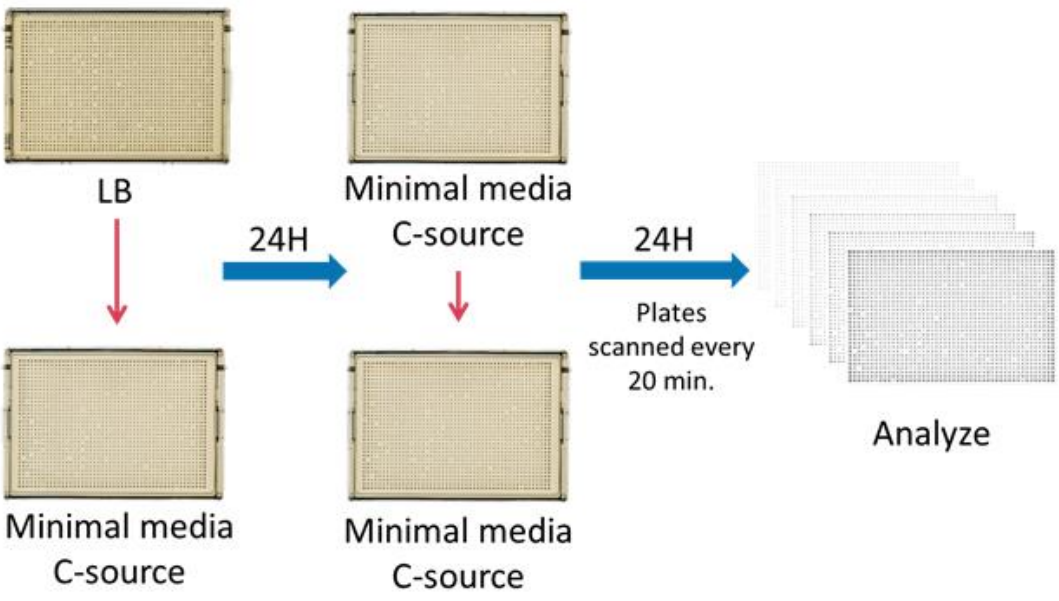

LB
24H
Minimal media
C-source
24H
Plates
scanned every
20 min.
Analyze
Minimal media
C-source
Minimal media
C-source

Figure S3

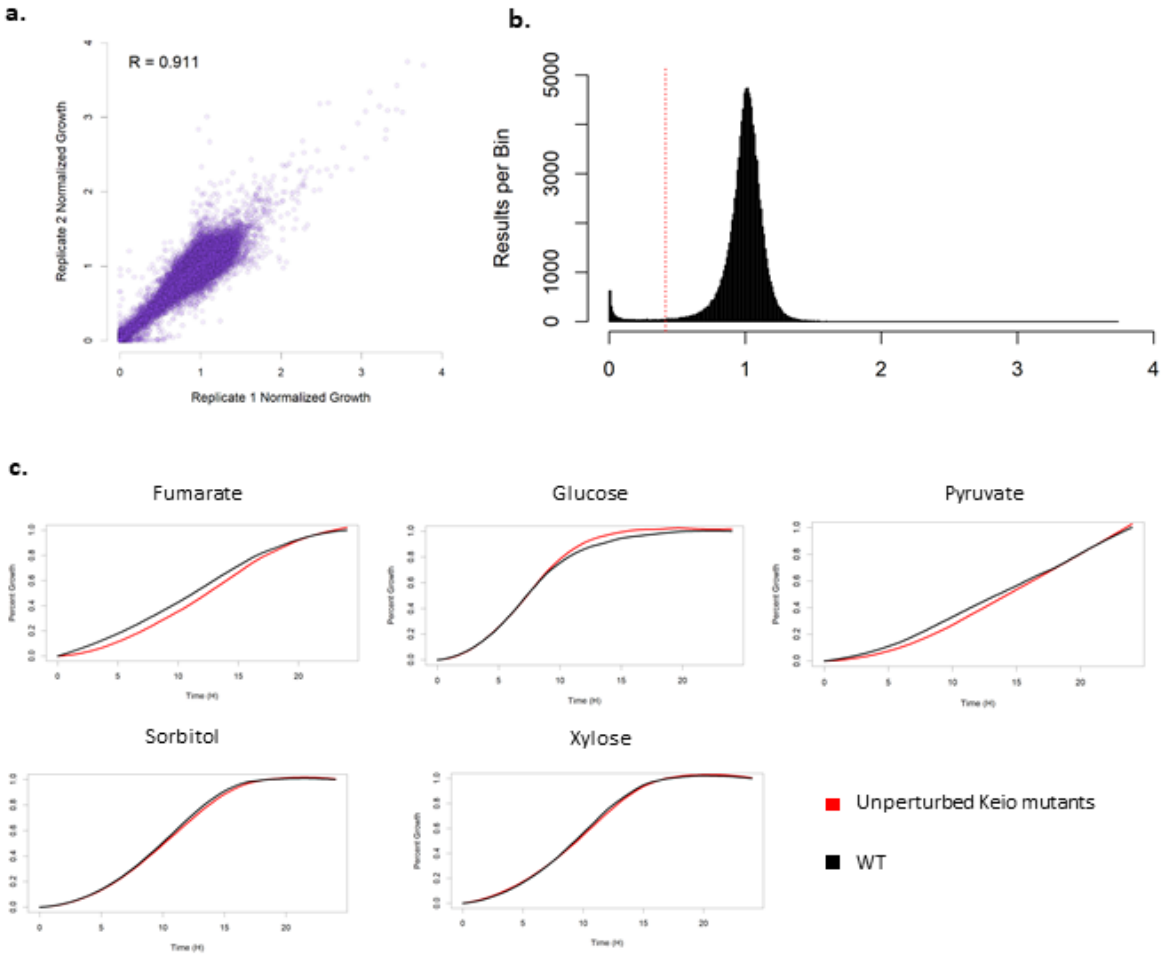
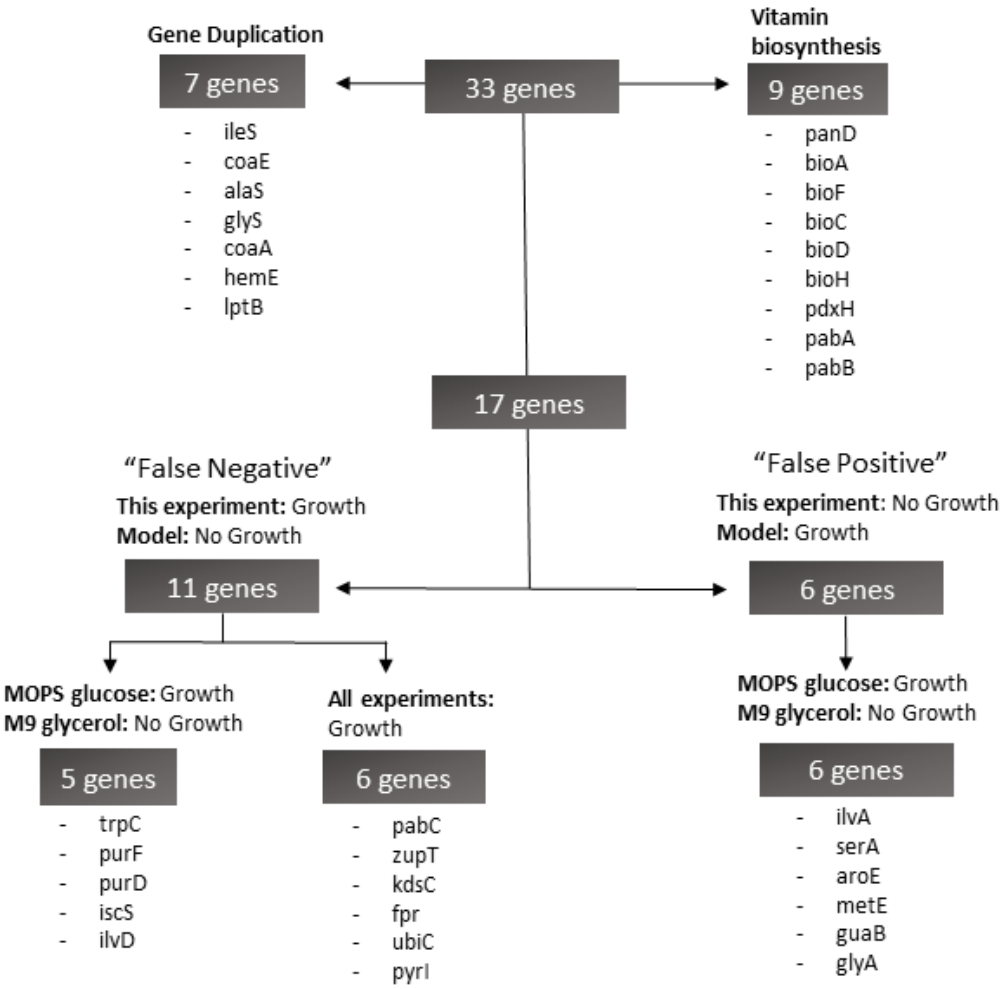

a.
R = 0.911
Replicate 2 Normalized Growth
Replicate 1 Normalized Growth
b.
Results per Bin
0
1000
3000
5000
0
1
2
3
4
c.
Fumarate
Glucose
Pyruvate
Sorbitol
Xylose
Unperturbed Keio mutants
WT

Figure S4


33 genes
Gene Duplication
7 genes
- ileS
- coaE
- alaS
- glyS
- coaA
- hemE
- lptB
Vitamin biosynthesis
9 genes
- panD
- bioA
- bioF
- bioC
- bioD
- bioH
- pdxH
- pabA
- pabB
17 genes
"False Negative"
This experiment: Growth
Model: No Growth
11 genes
"False Positive"
This experiment: No Growth
Model: Growth
6 genes
MOPS glucose: Growth
M9 glycerol: No Growth
5 genes
- trpC
- purF
- purD
- iscS
- ilvD
All experiments: Growth
6 genes
- pabC
- zupT
- kdsC
- fpr
- ubiC
- pyrI
MOPS glucose: Growth
M9 glycerol: No Growth
6 genes
- ilvA
- serA
- aroE
- metE
- guaB
- glyA

Chapter III - Evaluating antibiotic susceptibility of *Escherichia coli* grown in different carbon sources

Preface

The work presented in this chapter is in preparation for submission, as of March 2022:

Tong, M. Wang, S. Balta, K., & Brown, E. D. Evaluating antibiotic susceptibility of *Escherichia coli* grown in different carbon sources.

In preparation.

For this work, I performed all the data analysis, experiments, and writing. Wang S. and Balta K. helped optimize the methodology and did some preliminary screening. Wang S. ran the bacterial growth curves in different carbon sources.

Abstract

Carbon sources dictate how a bacterial cell generates the energy required for survival. Depending on the quality of the carbon source, the cell reroutes its cellular machinery to efficiently use what is available. Bacteria in different host infection sites encounter different carbon sources and nutrients which are not reflected in the standardized media we use in the laboratory. Our goal in this study is to understand how *E. coli* grown in these different carbon sources may have altered susceptibility to different antibiotics. We systematically changed the carbon source available in a defined minimal media and tested for the change of MICs in a diverse panel of 15 antibiotics and probes. This resulted in 465 MICs, with 196 showing a fold-change of 4-fold or higher compared to the MIC in MHB. We found that antibiotics that entered the cell through bioenergetics, such as macrolides and tetracycline, had altered susceptibility depending on the carbon source. We also discovered that cells grown in adenosine showed decreased efflux activity. We show that although many antibiotics did not show large changes in efficacy when we changed the carbon source, it can still be an important factor to consider when looking at drug uptake.

Introduction

Antibiotic resistance is a global crisis that requires the discovery of new antibiotics and a strong understanding of existing ones to preserve their effectiveness[1]. Antibiotic efficacy is currently evaluated using standard laboratory methods that include growing bacteria in a rich microbiological broth, MHB[2]. This method, developed decades ago, allows bacteria to grow in its ideal conditions and involves adding in different concentrations of antibiotic to determine how much of the drug is required to inhibit growth. However, this method does not consider the environment the bacteria are presiding in during infection and can lead to flawed results which don't match clinical outcomes[3–7]. When bacteria colonize a host infection site, they encounter different chemicals which act as signals or nutrients[4,8]. This chemical composition varies between infection sites, is dynamic, and differs vastly from the chemical composition of MHB and other microbiological media commonly used in a laboratory setting[4].

E. coli is an organism primarily found in the intestinal tract of mammalian hosts and can metabolize a variety of carbon sources[9]. This ability allows for pathogenic strains to flourish, even in areas outside of the intestine, such as the urinary tract and the bloodstream[9,10]. To cause an infection, the pathogen must be able to metabolize the available nutrients and, in some sites, outcompete resident bacteria [11]. Depending on the site of infection, the metabolic pathways required for survival will differ [11]. When uropathogenic *E. coli* (UPEC) are grown in human urine, proteins involved in the transport of small peptides, gluconate, and pentose sugars are induced compared to UPEC grown in LB media[12]. UPEC with mutations in gluconeogenesis and the TCA cycle were

outcompeted by wild-type strains *in vivo*[12]. When *E. coli* use peptides as a carbon source, they get metabolized into TCA intermediates and require gluconeogenesis for growth, suggesting peptides may be a major carbon source for growth in urinary tract infections. In contrast, pathogenic strains of *E. coli* in the gut must be able to metabolize mucosal-available carbon sources and outcompete commensal micro-organisms in the gut. Indeed, growth of *E. coli* on mouse mucus requires genes involved in the catabolism of many mucosal-available carbon sources such as gluconate and fucose [11].

While we are currently unable to precisely mimic the composition of host nutrients in each infection site, it is important to understand how changing the nutrient composition can affect antibiotic efficacy. Since the genes required and induced are different depending on the carbon source of the media, there may be changes in antibiotic susceptibility. Previous studies have shown that in *Pseudomonas aeruginosa* changing the carbon source in the media significantly affected the susceptibility of the cells to polymyxin B and colistin [13]. Fumarate has been shown to increase cellular respiration and potentiate tobramycin activity in *P. aeruginosa*[14]. In *Staphylococcus aureus* persister cells, glucose increased the killing efficiency of daptomycin[15]. With current drug testing methods, there may be drugs that will work well within the host, but they may never be tested due to the lack of *in vitro* potency. This could give mechanistic insight as to why certain antibiotics work better *in vivo* than would be expected based on *in vitro* assays. Additionally, this would help in designing new therapeutic methods aimed at improving the efficacy of current antibiotics by informing on the conditions that will change the sensitivity of bacteria to certain drugs. Herein, we evaluated the susceptibility of *E. coli* to

15 different antibiotic compounds and probes grown in media containing one of thirty different carbon sources. Our antibiotic list was enriched for ones that had a spectrum of activity typically limited to gram-positive bacteria. This way we could identify conditions that may cause them to work in a gram-negative organism. While we discovered that many of them did not change potency in any carbon source, we also found that growth in certain carbon sources can enhance uptake and decrease efflux of certain compounds. We show that growth in oxaloacetate affects the proton motive force, allowing the increased uptake of macrolides. Growth in adenosine as a carbon source decreases drug efflux resulting in the increased potency of linezolid.

Results

MICs of antibiotics are different in different Carbon Environments

To understand how growth in different carbon environments could change antibiotic efficacy, we systematically tested the MIC of 15 compounds (Table 1) on *E. coli* grown on a defined minimal media containing one of 30 different carbon sources. We included a wide range of classes of antibiotics which were typically not used to treat gram-negative infections such as macrolides, aminocoumarins, oxacillin, rifamycins, oxazolidinones, glycopeptides, and bacitracin. We wanted to determine whether some of these drugs may work in gram-negative organisms when we changed the carbon source. These antibiotics generally lack gram-negative activity due to the outer membrane permeability barrier or efflux[16]. Although most of these antibiotics have targets in *E. coli*, they have a high MIC because they are unable to accumulate in the bacteria. We tested each of these compounds

for their MICs in three mutants of *E. coli:* one lacking the outer membrane channel TolC which is required for efflux, one hyperporinated strain constitutively expressing a FhuAΔC/Δ4L pore which allows compounds through the outer membrane, and the combination of both (Table 1) [17]. By using these mutant strains, we could categorize our antibiotics based on whether they were affected by efflux, unable to cross the outer membrane, or both. This would allow us to interrogate whether carbon sources could alter permeability or efflux.

Since *E. coli* grows differently depending on the carbon source available, we first determined the growth kinetics of *E. coli* grown in MOPS minimal media containing our selected carbon sources. This analysis allowed us to select the time it took for *E. coli* to reach early stationary phase when grown on each carbon source. This way we could ensure our MICs could be comparable across our growth media. We generated a set of 465 MICs which we then compared to the MIC of the corresponding compound in MHB (Figure 1). Of the combinations tested, we found 84 showed a four-fold change, and 112 showed an eight-fold change or more in MIC relative to their MIC in MHB (Figure 1).

We performed hierarchal clustering to determine whether similar classes of antibiotics clustered together on the x-axis (Figure 1). We found that, as expected, antimetabolites such as 5-fluorouracil and sulfamethoxazole clustered closely to each other since they are both more effective in minimal media compared to rich biological media. We also noticed that nitrofurantoin, an antibiotic often used for treating urinary tract infections due to its solubility in urine, has increased susceptibility in minimal media and clustered near the antimetabolites. While the MIC of nitrofurantoin was 16 μg/ml in MHB, we found that in

most minimal media, regardless of carbon source, the MIC was 2 μg/ml, an 8-fold decrease. There are a cluster of antibiotics: oxacillin, novobiocin, and bacitracin, which showed 2-fold or less change from MHB, meaning growth in different carbon sources did not change their activity. Linezolid was uniquely potentiated by growth in adenosine as a carbon source, while rifampicin was potentiated by growth in ribose and D-alanine. Tetracycline was the only antibiotic where we observed increased MICs when *E. coli* was grown on certain carbon sources. This caused it to cluster apart from the rest of the antibiotics. CCCP, an ionophore that causes uncoupling of the proton gradient, showed a decreased MIC in several carbon sources, including many of the ones involved in the TCA cycle. Two of our macrolides, azithromycin and erythromycin clustered next to each other, however, dirithromycin did not. We were interested in further understanding why macrolide susceptibility would be changed in certain carbon sources and why adenosine was the only carbon source which could potentiate linezolid in *E. coli*.

The MICs of macrolides are lower in carbon sources such as acetate and oxaloacetate due to their effect on the proton motive force

Macrolides are antibiotics that inhibit protein synthesis by binding the 50S ribosomal subunit to inhibit the elongation step of translation[18]. They are typically used to treat gram-positive bacterial infections. In our dataset, we identified eight carbon sources (fucose, oxaloacetate, adenosine, acetate, saccharate, fumarate, L-alanine, and D-alanine) where at least two out of the three macrolides we tested showed a decrease in MIC. To ensure these phenotypes are accurate, we retested the MICs for azithromycin and erythromycin on *E. coli* grown in minimal media containing these seven carbon sources

as well as glucose as a control (Figure S1). We chose to focus on these two macrolides because dirithromycin is a pro-drug which may confound our understanding of the mechanism. Out of these carbon sources, we found that growth in media containing oxaloacetate appeared to show the largest decrease in MIC of azithryomycin and erythromycin when compared to the ones grown in glucose. We focused our studies on understanding why growth in oxaloacetate potentiates the activity of azithromycin.

To determine whether the mechanism of action of macrolides were altered when cells were grown in oxaloacetate, we compared the MICs of azithromycin in *E. coli* containing plasmids that constitutively expressed the macrolide resistance genes *mphA* and *ermC*. When *E. coli* constitutively expresses *ermC*, a 23S rRNA methylation enzyme, we found that changing the carbon source did not decrease the MIC of azithromycin (Figure 2a). Conversely, when cells are expressing the macrolide phosphatase, *mphA*, we found that oxaloacetate still decreased the MIC of azithromycin when compared to glucose. Since the MIC changes persisted in cells expressing *mphA* but not *ermC*, we determined that growth in oxaloacetate does not alter the mechanism of action of macrolides but may instead increase the concentration of macrolides in the cell.

It has recently been shown that bicarbonate can potentiate the activity of macrolides by increasing its uptake into the cell[19]. Using a checkerboard assay, we found that in minimal media containing glucose as a carbon source, we were able to see the same synergy between bicarbonate and azithromycin (Figure 2c). However, this synergy disappears when cells are grown in oxaloacetate as a carbon source (Figure 2c). Bicarbonate acts on the proton motive force (PMF) by dissipating the ΔpH component. The proton motive

force is the electrochemical gradient between cell membranes. This gradient, generated through the activity of the electron transport chain, is what allows for ATP generation using ATP synthase and transport of some chemicals between membranes. Dissipating the ΔpH component of PMF results in a compensatory increase in the electrical component, Δψ, to maintain a constant value of PMF[19,20]. Changes in the electrical potential has been shown to increase the uptake of positively charged molecules such as aminoglycosides and macrolides[14,20]. Dissipaters of ΔpH enhance the activity of azithromycin and dissipaters of Δψ antagonize the action of azithromycin[19]. To check whether oxaloacetate was also affecting the ΔpH component of PMF, we used tetracycline as a probe. Cellular uptake of tetracycline is primarily driven by the ΔpH gradient[21]. Anything that disrupted this would increase the MIC of tetracycline since it would no longer efficiently enter the cell. Indeed, the MIC of tetracycline is increased when cells are grown in oxaloacetate as a carbon source (Figure 1).

We wanted to know if this phenotype persists through pathogenic gram-negative bacteria. Macrolides are compounds which can contain different net positive charges, with higher charges being associated with enhanced accumulation when ΔpH is targeted[20]. We tested the MIC of three macrolides with different charges: erythromycin (+1), azithromycin (+2), and tulathromycin (+3) on a panel of pathogens including an uropathogenic strain of *E. coli* (UPEC), *P. aeruginosa*, *S. enterica*, *K. pneumoniae*, and *A. baumanii* (Table 2). In all pathogens tested, the MIC of all three drugs decreased when the cells were grown using oxaloacetate as a carbon source compared to MHB or glucose. The MICs between glucose and MHB were generally comparable. While *A. baumanii* are unable to utilize

glucose as a carbon source, we found that the MIC of azithromycin was 32-fold lower in minimal media containing oxaloacetate as the carbon source compared to MHB. In general, we found the largest enhancement of activity in tulathromycin, followed by azithromycin and erythromycin. This is consistent with the hypothesis that growth in oxaloacetate perturbs the ΔpH of the proton motive force. Growth in specific carbon sources can sensitize *E. coli* cells to macrolide antibiotics by altering the ΔpH component of the proton motive force.

Growth in adenosine decreases the MIC of linezolid and other efflux substrates

We noticed that the MIC of linezolid was lower when *E. coli* was grown specifically in media containing adenosine as the sole carbon source (Figure 1). The MIC of linezolid in minimal media containing glucose was found to be 1024 μg/ml but 64 μg/ml in minimal media containing adenosine (Figure 3a). Linezolid is an oxazolidinone class antibiotic typically used to treat gram-positive infections. Linezolid works by binding to the ribosome and inhibiting the initiation of translation. The lack of activity in *E. coli* is primarily due the presence of the *acrAB-tolC* drug efflux system which efficiently extrudes linezolid. Indeed, the known efflux inhibitor phenylalanine-arginine β-napthylamide (PAβN), synergizes with linezolid (Figure S2). We noticed that even in *E. coli* grown on adenosine as a carbon source, PAβN still showed synergy, even though the MIC in absence of inhibitor was lower than the MIC of *E. coli* grown in glucose (Figure S2).

PAβN has been shown to have pleiotropic effects, so we also tested our phenotypes using the *ΔtolC,* pore, and *ΔtolC* + pore strains of *E. coli* (Table 3)[22]. Since linezolid is an efflux substrate, we wanted to determine whether growth in adenosine would be able to decrease the MIC of other compounds that were subject to efflux. We confirmed that many substrates for efflux such as linezolid, tetracyclines, CCCP, chloramphenicol, lincomycins, and erythromycin all showed a lower MIC when cells were grown in adenosine compared to glucose (Table 3). These antibiotics also showed a fold-change in MICs when comparing WT with the *ΔtolC* strain indicating they are subject to efflux. In general, for these compounds, we noticed that the fold change between WT and *ΔtolC* were higher when cells were grown in glucose compared to adenosine (Table 3). Deleting the TolC efflux pump potentiates linezolid, however, growing cells on adenosine did not further potentiate this effect more than two-fold. The MICs in most of the compounds we tested were highest in cells grown in glucose, intermediate in cells grown in adenosine, and the lowest in a *ΔtolC* mutant regardless of media. This led us to believe that the lowered MIC is due to decreased efflux of the compound when the cells are grown in adenosine. To ensure the decreased MIC of tetracycline was due to increased accumulation of the drug, we employed a fluorescence-based assay to track tetracycline uptake. As expected, we found an increased accumulation of tetracycline in cells grown in adenosine compared to cells grown in glucose as a carbon source (Figure 3b). We also used a fluorescent probe, Hoechst H33342, which is an efflux substrate that increases in fluorescence intensity when bound to DNA. In this assay, the dye is added to extracellular media and diffuses into the cell where the accumulation correlates with increased

fluorescence intensity[23]. When efflux is inhibited, more dye accumulates and fluorescence increases. We grew *E. coli* cells in minimal media with either adenosine or glucose as a carbon source and found that cells grown in adenosine showed increased accumulation of the H33342 dye (Figure 3c). We used PAβN, as a positive control and cells treated with this inhibitor showed greater fluorescence than the untreated control in both conditions.

We were interested in seeing if growth in adenosine would cause other gram-negative pathogens to be more susceptible to linezolid. We determined the MIC of linezolid in an uropathogenic *E. coli* (UPEC) strain, *P. aeruginosa, K. pneumoniae, S. entericae,* and *A. baumanii* (Table 4). Notably, *P. aeruginosa* and *A. baumanii* are unable to use adenosine as a sole carbon source and did not show growth. We also confirmed that a gram-positive pathogen, *S. aureus,* which does not contain efflux pumps, showed no change in Linezolid MIC between cells grown in adenosine and glucose. However, in the other gram-negative pathogens we tested, we saw at least a 4-fold decrease in the MIC of linezolid when comparing growth in glucose to the growth in adenosine.

Discussion

With the growing need of antibiotics, it is imperative to understand how our current arsenal of antibiotics work in different nutrient environments. The environments that bacteria encounter during a host infection are dynamic and complex, with different nutrients available at different sites of infection[4,8]. This is different from the environment bacteria encounter in a lab environment, especially when using standard rich microbiological medium. Although there have been previous studies testing antibiotic

susceptibility in host-mimicking media, there lacks a study that systematically alters one specific aspect at a time[3]. In our present study, the focus was on carefully altering one part of this media, the carbon source, to see how that would affect antibiotic efficacy. Here, we show that growth in certain carbon sources alters the transmembrane pH gradient which potentiates the effect of macrolides and decreases the efficacy of tetracyclines. We also show that adenosine can decrease efflux activity, making efflux substrates like linezolid more effective. While growth in either oxaloacetate or adenosine can potentiate the activity of macrolide antibiotics, they do so in different ways. As one of the biggest hurdles to targeting gram-negative infections is difficulty in drug entry and accumulation, it is promising to see that different carbon sources can promote uptake or prevent efflux of certain antibiotics.

Since the focus of our study was on antibiotics that typically only work on gram-positive bacteria, it is not surprising that four of them didn't show any change in MIC in any carbon source. Most of these are because the MICs are close to or higher than the highest concentration we tested in our screen: 256µg/ml. In this case, we would be unable to tell if the MICs did increase in specific carbon sources since these drugs effectively do not work on *E. coli*. In addition to antibiotics, we also included various probes in our study. Lamotrigine is an anticonvulsant which has previously been shown to target ribosome biogenesis in *E. coli*[24]. Treatment with lamotrigine results in an accumulation of immature 30S and 50S ribosomal subunits at 15°C which does not occur at 37°C[24]. We found that while Lamotrigine showed an MIC of 512µg/mL in rich media at 37°C, we found it decreased to around 128µg/mL in minimal media with glucose as a carbon source, and as

low as 16 µg/mL in L-alanine as the carbon source. This is interesting since we were able to identify an environment where lamotrigine works without the requirement of cells being grown at 15°C. We tested this phenotype on several other bacterial pathogens and found that this phenotype also occurs in *S. enterica* but not *K. pneumoniae* (Data not shown). In *E. coli*, the ribosome content of cells is proportional to the growth rate; where the number of ribosomes is increased in faster growing cells to meet its protein synthesis demands[25,26]. Growth rate is linked closely to carbon metabolism and is altered when different carbon sources are available for growth [27,28]. It is difficult to ascertain the exact growth rates of bacteria within a host due to the complex nature of pathogenesis. However, growth within a host occurs much slower than in rich microbiological medium. Given the number of current antibiotics that target the ribosome and protein synthesis, it may be prudent to test antimicrobial susceptibility in media which more closely simulates the growth rate within the host.

There have been reports of carbon source affecting the susceptibility of various pathogens to antibiotics [13,29–31]. Depending on the carbon source being used, different parts of central metabolism are required, and different genes are upregulated [27,28]. The presence of TCA cycle intermediates upregulates the pathway and increases electron transport chain activity which is required from drug uptake and lethality[29,30,32]. We found that the proton motive force is altered during growth in oxaloacetate, which allowed macrolides to enter the cell. Carbon sources can affect the metabolic state and levels of cellular respiration in bacteria, which in turn affects antibiotic efficacy[33,34]. One aspect that may be overlooked

in current understanding of bacterial infections is the metabolic state of bacteria and how the proton motive force may be altered during pathogenesis.

Within the host, many compounds that can be used by bacteria as a carbon source are present. Many of these compounds have been shown to play a role as signals during an infection. Acetate and other short chain fatty acids are important in modulating *S. enterica* pathogenicity by acting as a signal for invasion gene expression[35]. Adenosine is a breakdown product of ATP. In the human host, it is a signaling molecule released by cells to promote anti-inflammatory responses, meaning it is prevalent in a typical host infection[36]. EPEC infections of cultured tissues triggers a release of ATP from the host cell which breaks down into adenosine[36]. In the case of cellular stress or tissue injury, concentrations of adenosine can increase 100-fold in the host [37,38]. In an intestinal infection, adenosine has been shown to improve the growth of EPEC compared to other nucleosides and nucleotides.[39] It is possible that this improvement of growth from adenosine may act as a double-edged sword that also increases their susceptibility to certain antibiotics.

In this study we focus on changing carbon source, however there is an opportunity to change different parts of media compositions to understand how each component affects antibiotic efficacy. The advantage to rationally altering each aspect of the media is the ease of deconvoluting what caused the effect. Our results suggest that changing the carbon source has effects on the proton motive force and drug efflux. This study provides a framework for studying how changing one aspect of metabolism can have effects on different antimicrobials.

Materials and Methods

Chemicals

Chemicals used in this study were purchased from Sigma-Aldrich unless otherwise noted. Carbon source concentrations were picked so that all carbon sources resulted in the same amount of carbon added as previously described [27]. MOPS minimal media (Teknova) was used for all work in minimal media. Kanamycin was added at a concentration of 50 μg/mL when working with strains from the Keio collection[40].

Bacterial strains and growth conditions

For routine experiments, *E. coli* BW25113 strain was streaked onto LB media with appropriate antibiotic selection from a frozen glycerol stock and grown overnight in a stationary incubator at 37°C. A single colony was picked to inoculate an overnight culture which was grown shaking at 250 rpm. This was used to subculture into a new tube to mid-log phase growth. For experiments in minimal media, this subculture was washed using MOPS minimal media containing no carbon source and diluted 1:100 in final conditions. Mutants were taken from the Keio collection and streaked on to an agar plate using kanamycin selection. The pore strain FhuAΔC/Δ4L was a gift from Helen Zgurskaya. For experiments involving gram-negative pathogens, the following strains were used: uropathogenic *E. coli* CFT073, *K. pneumoniae* ATCC 43816, *P. aeruginosa* PA01, *S. entericae* serovar Typhimurium ATCC14028, *A. baumanii* ATCC 17098, *S. aureus* strain Newman. The plasmids constitutively expressing resistance elements, *pGDP3:mphA* and *pGDP4:ermC*, were transformed into *E. coli* K-12 BW25113[41,42].

Liquid Growth Kinetics

Briefly, *E. coli* was grown overnight in LB media. This culture was diluted 1:100 into fresh LB media where it was grown to an OD600 of 0.4-0.5 (mid-log phase). We washed this culture three times using MOPS minimal media containing no carbon source and diluted it 1:100 into a 384-well assay plate, with each well containing minimal media with a different carbon source. The Biotek Epoch 2 microplate spectrophotometer was used to read OD600 measurements every 15 minutes to generate growth curves.

Minimal inhibitory concentration assays

Using the Echo Acoustic dispenser, we added antibiotics using a two-fold dilution gradient, starting at 256µg/ml as the highest concentration, to a 384-well plate. Liquid backfill was completed using the solvent the antibiotic was dissolved in to ensure the same volume of drug is dispensed across the plate. *E. coli* cells were prepared as described above and dispensed on to each plate to a final volume of 30µL.

Checkerboard Assays

Checkerboard assays were completed in 96-well plates in duplicate. On each axis, a 2-fold serial dilution series was complete of the two drugs that were tested. The control well contains only the solvent used for each drug. Plates were filled to a final volume of 200 µL using a standard inoculum preparation as described above.

Hoechst 33342 dye accumulation assay

Assay was completed as previously described. Briefly, *E. coli* BW25113 cells were grown overnight in LB media. Cells were then washed in MOPS minimal media containing no carbon source and used to inoculate fresh minimal media containing either glucose or adenosine as a sole carbon source. These cultures were grown to an OD600 of 0.6 and washed twice in PBS. Cultures were adjusted to an OD600 of 0.1 in PBS and 180 µl was added to each well of a black opaque 96-well plate (Costar). 20µL of H33342 was added to each well to give a final concentration of 2.5µM. Plates were loaded into a Biotek Synergy Neo plate reader and fluorescence was monitored from the top of the wells using excitation filters of 355nm and emission filters of 460nm. Time points were taken every 75s for 2500s. At least two biological replicates were used with 3 technical replicates of each. Appropriate no cells and no drug controls were used in the analysis. PaβN was added at the beginning of the experiment at a final concentration of 333µM.

Tetracycline uptake assay

Tetracycline uptake was measured by monitoring the fluorescence enhancement of tetracycline as it enters the cell. Assay was completed as previously described[19]. Briefly, *E. coli* was grown to mid-log phase in MOPS minimal media containing either adenosine or glucose as a carbon source. Cells were pelleted and washed in 10mM HEPES pH 7.2. Tetracycline was added to the plates for a final concentration of 125µg/ml in each well. Cell suspensions were added to the wells for a final volume of 200µL per well. Assays were performed on a black 96-well plate (Co-Star) and read on a Biotek

spectrophotometer with filters of 405nm for excitations and 535nm for emission. Plates were read every 5 minutes for 60 minutes. Results were corrected for fluorescence in a no cell control and the relative fluorescence units were plotted.

References

1. Brown, E. D. & Wright, G. D. Antibacterial drug discovery in the resistance era. *Nature* **529**, 336–343 (2016).
2. Wheat, P. F. History and development of antimicrobial susceptibility testing methodology. *Journal of Antimicrobial Chemotherapy* **48**, 1–4 (2001).
3. Ersoy, S. C. *et al.* Correcting a Fundamental Flaw in the Paradigm for Antimicrobial Susceptibility Testing. *EBioMedicine* **20**, 173–181 (2017).
4. Bjarnsholt, T. *et al.* The importance of understanding the infectious microenvironment. *The Lancet Infectious Diseases* **0**, (2021).
5. Kubicek-Sutherland, J. Z. *et al.* Host-dependent Induction of Transient Antibiotic Resistance: A Prelude to Treatment Failure. *EBioMedicine* **2**, 1169–1178 (2015).
6. Dunman, P. M. & Tomaras, A. P. Translational deficiencies in antibacterial discovery and new screening paradigms. *Current Opinion in Microbiology* vol. 27 108–113 (2015).
7. Carfrae, L. A. *et al.* Mimicking the human environment in mice reveals that inhibiting biotin biosynthesis is effective against antibiotic-resistant pathogens. *Nature Microbiology* 1–9 (2019) doi:10.1038/s41564-019-0595-2.
8. Brown, S. A., Palmer, K. L. & Whiteley, M. Revisiting the host as a growth medium. *Nature reviews. Microbiology* **6**, 657–66 (2008).
9. Kaper, J. B., Nataro, J. P. & Mobley, H. L. T. Pathogenic Escherichia coli. *Nature Reviews Microbiology* **2**, 123–140 (2004).
10. Eisenreich, W., Dandekar, T., Heesemann, J. & Goebel, W. Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nature Reviews Microbiology* **8**, 401–412 (2010).
11. Chang, D.-E. *et al.* Carbon nutrition of Escherichia coli in the mouse intestine. *Proceedings of the National Academy of Sciences* **101**, 7427–7432 (2004).
12. Alteri, C. J. C. *et al.* Fitness of Escherichia coli during urinary tract infection requires gluconeogenesis and the TCA cycle. *PLoS Pathogens* **5**, e1000448 (2009).
13. Conrad, R. S., Gary Wulf, R. & Clay, D. L. *Effects of Carbon Sources on Antibiotic Resistance in Pseudomonas aeruginosa*. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY* vol. 15 (1979).

14. Allison, K. R., Brynildsen, M. P. & Collins, J. J. Metabolite-enabled eradication of bacterial persisters by aminoglycosides. *Nature* **473**, 216–20 (2011).

15. Prax, M., Mechler, L., Weidenmaier, C. & Bertram, R. Glucose Augments Killing Efficiency of Daptomycin Challenged Staphylococcus aureus Persisters. *PLOS ONE* **11**, e0150907 (2016).

16. Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochimica et biophysica acta* **1794**, 808–16 (2009).

17. Krishnamoorthy, G. *et al.* Breaking the Permeability Barrier of Escherichia coli by Controlled Hyperporination of the Outer Membrane. *Antimicrobial Agents and Chemotherapy* **60**, 7372 (2016).

18. Vázquez-Laslop, N. & Mankin, A. S. *How Macrolide Antibiotics Work. Trends in Biochemical Sciences* vol. 43 668–684 (2018).

19. Farha, M. A., French, S., Stokes, J. M. & Brown, E. D. Bicarbonate Alters Bacterial Susceptibility to Antibiotics by Targeting the Proton Motive Force. *ACS Infectious Diseases* **4**, 382–390 (2018).

20. Farha, M. A. *et al.* Overcoming Acquired and Native Macrolide Resistance with Bicarbonate. *ACS Infectious Diseases* **6**, 2709–2718 (2020).

21. Yamaguchi, A., Ohmori, H., Kaneko-Ohdera, M., Nomura, T. & Sawai, T. *ApH-Dependent Accumulation of Tetracycline in Escherichia coli*. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY* vol. 35 (1991).

22. Lamers, R. P., Cavallari, J. F. & Burrows, L. L. The Efflux Inhibitor Phenylalanine-Arginine Beta-Naphthylamide (PAβN) Permeabilizes the Outer Membrane of Gram-Negative Bacteria. *PLOS ONE* **8**, e60666 (2013).

23. Coldham, N. G., Webber, M., Woodward, M. J. & Piddock, L. J. V. A 96-well plate fluorescence assay for assessment of cellular permeability and active efflux in Salmonella enterica serovar Typhimurium and Escherichia coli. *Journal of Antimicrobial Chemotherapy* **65**, 1655–1663 (2010).

24. Stokes, J. M., Davis, J. H., Mangat, C. S., Williamson, J. R. & Brown, E. D. Discovery of a small molecule that inhibits bacterial ribosome biogenesis. *eLife* **3**, e03574 (2014).

25. Bollenbach, T., Quan, S., Chait, R. & Kishony, R. Nonoptimal Microbial Response to Antibiotics Underlies Suppressive Drug Interactions. *Cell* **139**, 707–718 (2009).

26. Kemp, P. F. Can We Estimate Bacterial Growth Rates from Ribosomal RNA Content? in *Molecular Ecology of Aquatic Microbes* 279–302 (Springer Berlin Heidelberg, 1995). doi:10.1007/978-3-642-79923-5_16.

27. Tong, M. *et al.* Gene dispensability in escherichia coli grown in thirty different carbon environments. *mBio* **11**, 1–20 (2020).

28. Liu, M. *et al.* Global Transcriptional Programs Reveal a Carbon Source Foraging Strategy by Escherichia coli. *Journal of Biological Chemistry* **280**, 15921–15927 (2005).

29. Meylan, S. *et al.* Carbon Sources Tune Antibiotic Susceptibility in Pseudomonas aeruginosa via Tricarboxylic Acid Cycle Control. *Cell Chemical Biology* **24**, 195–206 (2017).

30. Allison, K. R., Brynildsen, M. P. & Collins, J. J. Metabolite-enabled eradication of bacterial persisters by aminoglycosides. *Nature* **473**, 216–220 (2011).

31. Pethe, K. *et al.* A chemical genetic screen in Mycobacterium tuberculosis identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy. *Nature Communications* **1**, 1–8 (2010).

32. Kohanski, M. A., Dwyer, D. J., Wierzbowski, J., Cottarel, G. & Collins, J. J. Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell* **135**, 679–90 (2008).

33. Lopatkin, A. J. *et al.* Bacterial metabolic state more accurately predicts antibiotic lethality than growth rate. *Nature Microbiology* vol. 4 2109–2117 (2019).

34. Lobritz, M. A. *et al.* Antibiotic efficacy is linked to bacterial cellular respiration. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 8173–80 (2015).

35. Lawhon, S. D., Maurer, R., Suyemoto, M. & Altier, C. Intestinal short-chain fatty acids alter Salmonella typhimurium invasion gene expression and virulence through BarA/SirA. *Molecular Microbiology* **46**, 1451–1464 (2002).

36. Crane, J. K., Olson, R. A., Jones, H. M. & Duffey, M. E. Release of ATP during host cell killing by enteropathogenic E. coli and its role as a secretory mediator. *American Journal of Physiology - Gastrointestinal and Liver Physiology* **283**, (2002).

37. Linden, J. New insights into the regulation of inflammation by adenosine. *Journal of Clinical Investigation* vol. 116 1835–1837 (2006).

38. Barletta, K. E., Cagnina, R. E., Burdick, M. D., Linden, J. & Mehrad, B. Adenosine A2B receptor deficiency promotes host defenses against gram-negative bacterial pneumonia. *American Journal of Respiratory and Critical Care Medicine* **186**, 1044–1050 (2012).

39. Crane, J. K. & Shulgina, I. Feedback effects of host-derived adenosine on enteropathogenic escherichia coli. *FEMS Immunology and Medical Microbiology* **57**, 214–228 (2009).

40. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology* **2**, 2006.0008 (2006).

41. Macnair, C. R. & Brown, E. D. Outer membrane disruption overcomes intrinsic, acquired, and spontaneous antibiotic resistance. *mBio* **11**, 1–15 (2020).

42. Cox, G. *et al.* A Common Platform for Antibiotic Dereplication and Adjuvant Discovery. *Cell Chemical Biology* **24**, 98–109 (2017).

Tables

Table 1: Antibiotics screened in this study and their MICs in *E. coli* mutants deficient in efflux (*ΔtolC*), hyperporinated (pore), or hyperporinated and deficient in efflux (*ΔtolC + pore*) grown on MHB or MOPS glucose. These are separate into four categories based on their MICs in these strains. OM, outer membrane.

		MHB				MOPS glucose			
		WT	***ΔtolC***	**pore**	***ΔtolC* + pore**	**WT**	***ΔtolC***	**pore**	***ΔtolC* + pore**
Unaffected by efflux or OM barrier	5-fluorouracil	16	16	16	16	0.06	0.06	0.06	0.06
	Sulfamethoxazole	8	16	8	16	0.25	0.25	0.25	0.25
	Lamotrigine	128	128	128	64	64	64	32	32
	Nitrofurantoin	8	4	8	4	4	2	2	2
	Tetracycline	1	0.5	0.5	0.5	4	2	0.5	0.25
Primarily efflux	CCCP	32	1	16	1	32	2	32	1
Primarily affected by OM barrier	Bacitracin	>512	>512	32	32	>512	>512	32	64
	Vancomycin	256	256	2	2	64	64	0.5	0.5
	Rifampicin	8	4	0.5	0.125	16	8	1	0.5
Combination of efflux and OM barrier	Dirithromycin	16	0.5	0.5	0.125	32	1	0.5	0.125
	Erythromycin	32	1	1	0.125	256	1	2	0.125
	Azithromycin	8	1	<.25	<.25	16	0.5	0.5	<.25
	Novobiocin	64	1	16	0.5	>1024	4	256	0.5
	Oxacillin	512	0.25	16	0.25	512	0.5	64	0.25
	Linezolid	256	8	64	8	1024	8	64	8

Table 2: The minimal inhibitory concentrations (MICs) of three macrolides of different charges: erythromycin (+1), azithromycin (+2) and tulathromycin (+3) in different strains of gram-negative bacteria.

		MHB	Glucose	Oxaloacetate
E. coli (BW25113)	Erythromycin	64	128	16
	Azithromycin	4	4	0.25
	Tulathromycin	4	2	0.06
Uropathogenic E. coli (CFT073)	Erythromycin	64	64	16
	Azithromycin	4	2	0.25
	Tulathromycin	4	1	0.06
P. aeruginosa	Erythromycin	128	64	16
	Azithromycin	32	32	2
	Tulathromycin	>32	>32	4
K. pneumoniae	Erythromycin	256	128	16
	Azithromycin	16	4	0.5
	Tulathromycin	>32	8	0.5
S. pneumoniae	Erythromycin	64	128	8
	Azithromycin	4	4	0.125
	Tulathromycin	4	2	0.06
A. baumanii	Erythromycin	16	NA	2
	Azithromycin	4	NA	0.125
	Tulathromycin	>32	NA	0.25

Table 3: The minimal inhibitory concentrations (MICs) of efflux substrates on *E. coli* mutants deficient in efflux (*ΔtolC*), hyperporinated (pore), or hyperporinated and deficient in efflux (*ΔtolC + pore*) grown on glucose and adenosine as carbon sources. FC, fold change.

	Glucose					Adenosine					
Antibiotic	WT	*ΔtolC*	*pore*	*ΔtolC + pore*	**FC:** WT/ tolC	WT	*ΔtolC*	*pore*	*ΔtolC + pore*	**FC:** WT/ tolC	**FC:** glucose/ adenosine
Doxycycline	8	0.25	2	0.25	32	0.25	0.125	0.125	0.125	2	32
Tetracycline	4	0.5	2	0.25	8	0.25	<.125	0.25	<.125	>2	16
Minocycline	4	0.25	2	0.25	8	0.5	0.125	0.25	0.125	4	4
Erythromycin	256	1	2	0.125	256	16	1	0.06	<.03	16	16
Linezolid	1024	8	64	8	128	64	4	32	4	16	16
CCCP	32	2	32	1	16	8	0.25	8	0.25	32	4
Chloramphenicol	16	2	4	2	32	4	2	2	2	2	16
Lincomycin	1024	128	1024	128	4	32	16	16	16	2	>16
Clindamycin	64	8	0.5	0.5	8	4	1	<.25	<.03	4	16

Table 4: The minimal inhibitory concentrations (MICs) of three linezolid in different strains of gram-negative bacteria grown in either glucose or adenosine as a carbon source.

	Glucose	Adenosine
E. coli (BW25113)	1024	64
Uropathogenic E. coli (CFT073)	256	64
P. aeruginosa	>1024	no growth
K. pneumoniae	>1024	128
S. entericae	512	128
A. baumanii	256	NA

Figure Legends

Figure 1: Heatmap showing the fold change in MIC in each carbon source relative to the MIC in MHB. Red depicts MICs that decreased, and blue means the MIC increased. The y-axis shows the carbon sources sorted with the lowest growth rate on the top and the highest growth rate on the bottom.

Figure 2: Growth in oxaloacetate potentiates macrolides by increasing uptake.

a. MIC curves of *E. coli* containing the plasmids *pGDP3:mphA* plasmid and *pGDP4:ermC*. Results are representative of at least two biological replicates.

b. Checkerboard assays showing increasing concentrations of Azithromycin (μg/ml) on the x-axis and increasing concentrations of bicarbonate (μM) on the y-axis. Darker colours indicate growth while white represents no growth. There is a two-fold decrease in concentrations at each step in both the x and y-axes.

Figure 3: Growth in adenosine results in accumulation of efflux substrates.

a. The MIC of linezolid is decreased when *E. coli* is grown in adenosine compared to glucose. Results are representative of at least two biological replicates.

b. Tetracycline assay showing increased accumulation of tetracycline when cells are grown in adenosine compared to glucose.

c. Hoechst33342 (H33342) (2.5μM) accumulation shown in *E. coli* grown in glucose and adenosine +/- PAβN (333 μM)

Figure S1: Reconfirmation of macrolide phenotypes

a. The fold-change of MIC of each macrolide grown in each of the carbon sources on the x-axis compared to the MIC in MOPS glucose as a control. The values represent the fold-change decreased from the MIC in glucose.

b. To ensure phenotype wasn't due to an increased buffer capacity from the carbon source, we increased the concentration of the MOPS buffer and showed that the phenotype did not change. The x-axis contains a two-fold dilution of azithromycin and the y-axis in increasing concentrations of the MOPS buffer.

Figure S2: Additional experiments on E. coli grown on adenosine as a carbon source

a. The MIC curve of linezolid against *S. aureus* grown in media with adenosine or glucose as a carbon source

b. To ensure phenotype wasn't due to an increased buffer capacity from the carbon source, we increased the concentration of the MOPS buffer and showed that the phenotype did not change. The x-axis contains a two-fold dilution of linezolid and the y-axis in increasing concentrations of the MOPS buffer.

c. The efflux pump inhibitor, PAβN showed synergy with linezolid in cells grown with both adenosine and glucose as a carbon source.

Figures

Figure 1

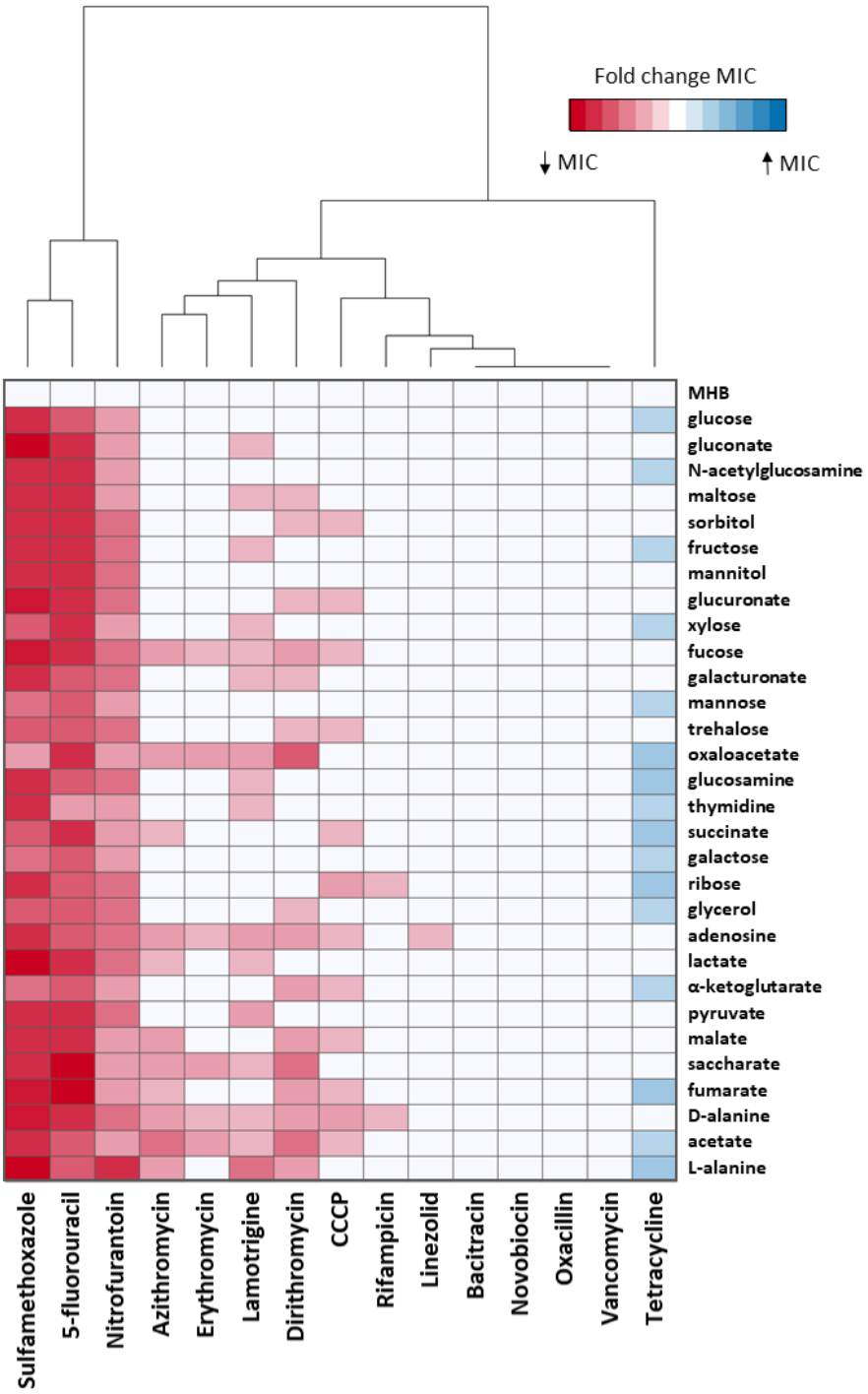


Figure 2

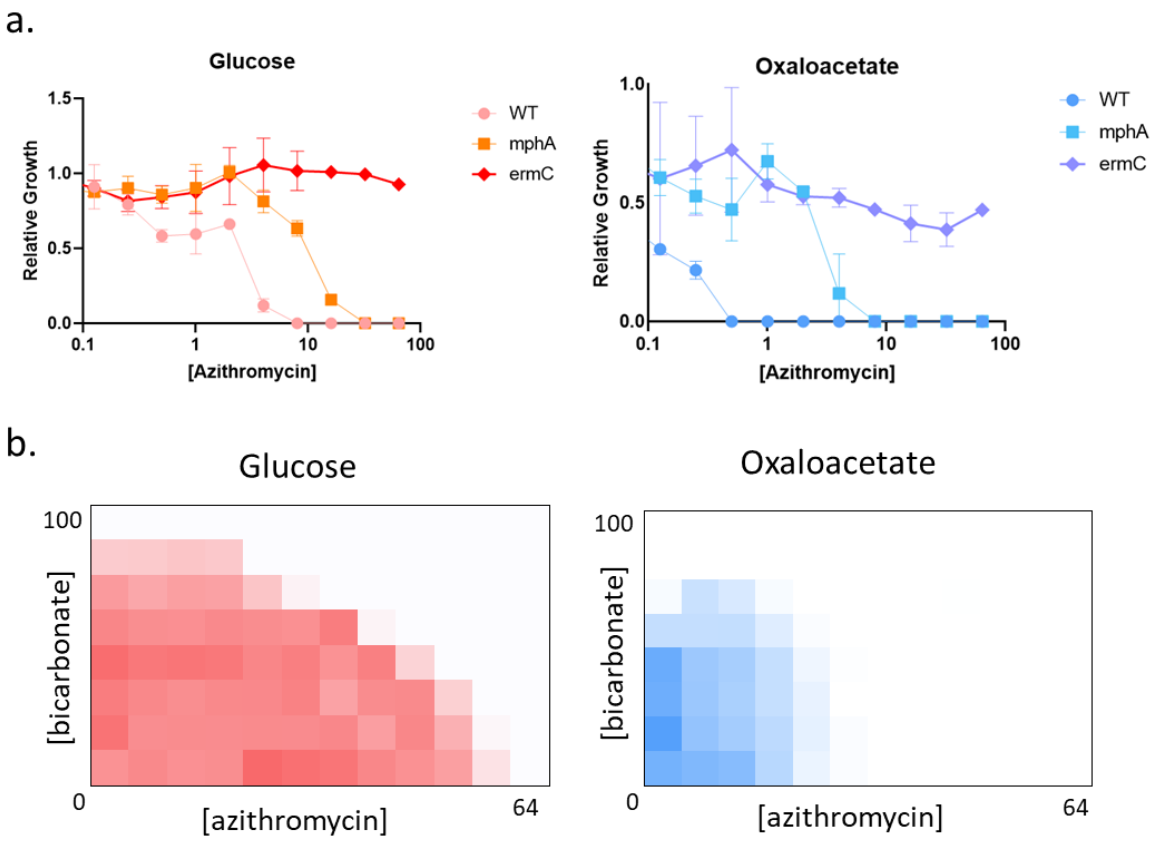

a.
Glucose
Relative Growth
1.5
1.0
0.5
0.0
0.1
1
10
100
[Azithromycin]
WT
mphA
ermC
Oxaloacetate
Relative Growth
1.0
0.5
0.0
0.1
1
10
100
[Azithromycin]
WT
mphA
ermC
b.
Glucose
100
[bicarbonate]
0
[azithromycin]
64
Oxaloacetate
100
[bicarbonate]
0
[azithromycin]
64

Figure 3

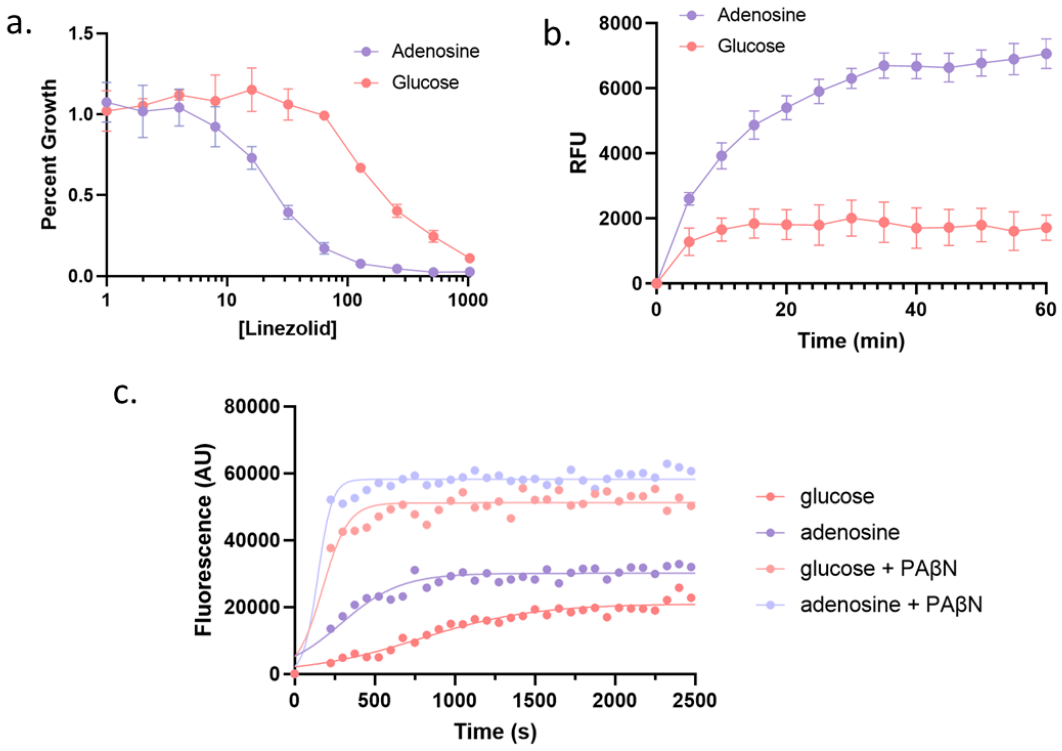

a.
Percent Growth
[Linezolid]
Adenosine
Glucose
b.
RFU
Time (min)
Adenosine
Glucose
c.
Fluorescence (AU)
Time (s)
glucose
adenosine
glucose + PAβN
adenosine + PAβN

Figure S1

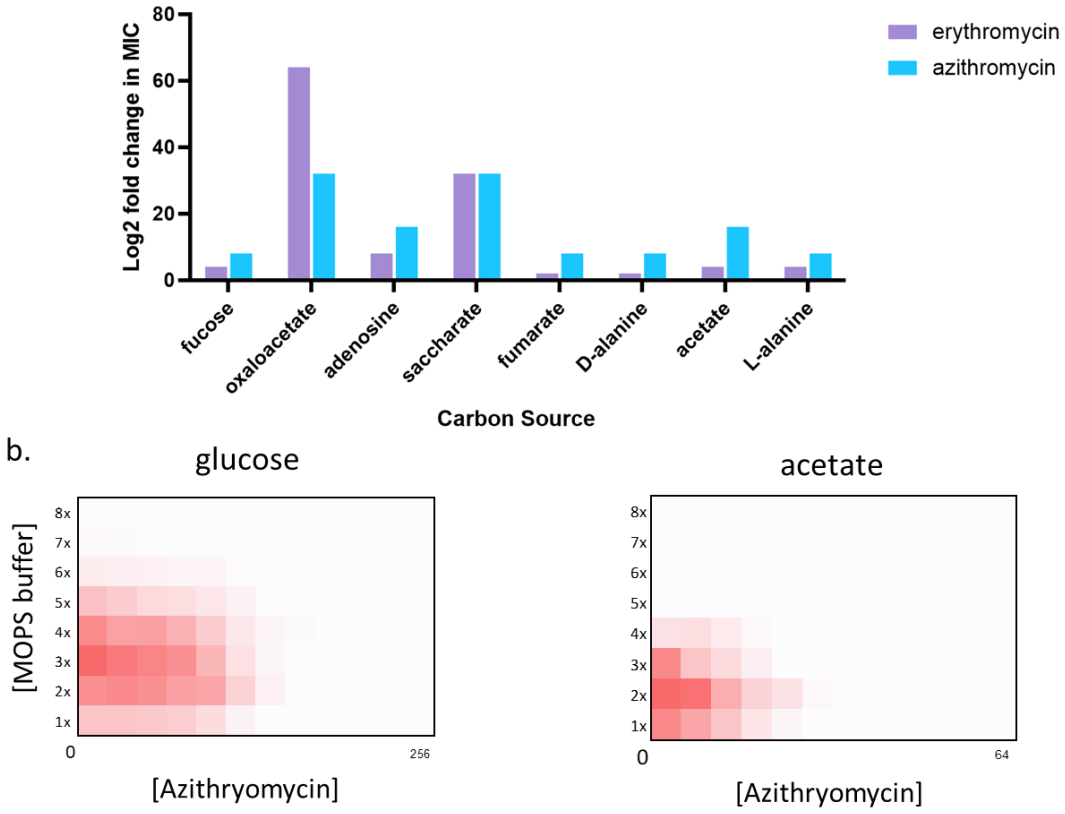
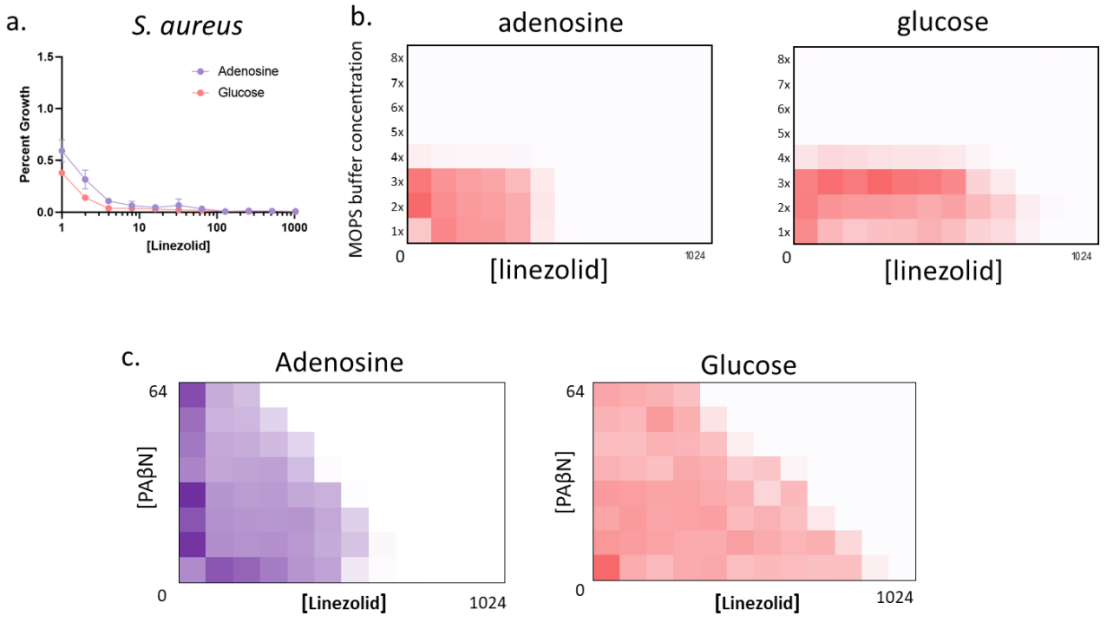

Log2 fold change in MIC
80
60
40
20
0
erythromycin
azithromycin
fucose
oxaloacetate
adenosine
saccharate
fumarate
D-alanine
acetate
L-alanine
Carbon Source
b.
glucose
acetate
[MOPS buffer]
8x
7x
6x
5x
4x
3x
2x
1x
0
256
64
[Azithryomycin]
[Azithryomycin]

Figure S2


a.
S. aureus
Adenosine
Glucose
Percent Growth
1.5
1.0
0.5
0.0
1
10
100
1000
[Linezolid]
b.
adenosine
glucose
MOPS buffer concentration
8x
7x
6x
5x
4x
3x
2x
1x
0
1024
[linezolid]
c.
Adenosine
Glucose
64
[PAβN]
0
[Linezolid]
1024

Chapter IV - Conclusion

Summary

The contents of this thesis highlight the importance of revisiting central carbon metabolism. In Chapter 1, I scratched the surface of over 100 years' worth of biochemical studies elucidating all the metabolic pathways that we understand today. I've shown in Chapter 2 that despite over a century of work in the field, there is still so much we don't understand about the *E. coli* genome and how changing the carbon source can affect this. Even though most nutrient and metabolic pathways appear to be exhaustively studied, almost 35% of the *E. coli* genome remains poorly characterized. In that percentage, 51 of the genes show a phenotype when grown in a different carbon source which implies some sort of connection to carbon metabolism[1]. The first chapter describes what is currently known about which carbon sources may be relevant *in vivo*. My work in Chapter 2 provides a list of potential gene targets for pathogens that use those carbon sources for infection. If we can identify the exact composition of carbon sources in an infection site, we can find genes that are required for growth in all those carbon sources. In Chapter 3 I investigated how antibiotic efficacy can change in bacteria growing in different carbon sources. Our understanding for how effective an antibiotic may be in a host could be skewed due low *in vitro* efficacy. If simply changing the carbon source of the media can have such profound changes to the MIC of antibiotics, it is likely that the MIC in a host environment would also be different. Based on my knowledge of metabolism, I will discuss some ideas and strategies for future research in this field.

What are the carbon sources relevant for infection?

It is difficult to determine the exact composition of the host due to the variable nature of human bodies. Chapter 1 covered how researchers have been able to study potential carbon sources that pathogens consume during an infection. For example, there has been success in determining which carbon sources different strains of *E. coli* prefer during murine gut colonization and urinary tract infections. These studies showed that different strains have variable patterns of nutrient utilization, meaning we should be trying new strains in infection experiments. There is also evidence that pathogens behave in different ways in the presence of other bacteria[2,3]. While more complex, there is an opportunity to test co-cultures of different bacteria to see how gene regulations or the transcriptome changes when bacteria are infected with another bacterial strain compared to by itself[2]. In addition to this, we can learn from all colonization and infection experiments conducted on the murine gut and do similar experiments with different pathogens in different host infection sites. We can build upon what we know about which central metabolism pathways are required in each infection site by conducting similar studies using other pathogens of interest. By systematically deleting genes in different central metabolism pathways and doing *in vivo* competition experiments in different pathogens and different infection models, we can exhaustively determine which pathways are required for survival in the host. This data, combined with transpon-insertion sequencing approaches to determine which genes are required in different infection sites can be a powerful tool to understand what nutrients are vital for pathogen survival in the host.

Other bacterial stresses that occur during host infection

While nutrient availability is something to consider during a host infection, there are other stresses that bacteria encounter during an infection. These should be considered when trying to recreate an infection environment. For example, depending on the infection site, pH can vary significantly[4]. In the gastrointestinal tract, dental plaque, skin and some phagosomes, pH is much lower than what's found in blood[4]. In contrast, MOPS minimal media is buffered to a pH of 7.4 to match the pH in serum[5]. In certain host infection sites, the oxygen levels are low, and bacteria often go through anaerobic fermentation instead of aerobic growth[6,7]. When I was testing which carbon sources *E. coli* could use, I found certain carbon sources could not be used when I removed oxygen from the environment (data not shown). While most infection sites are low in oxygen due to host inflammation, we rarely test antibiotics or screen drugs in an anaerobic or microaerophilic environment.

Media Matters

When designing an experiment, one should consider details as basic as the media that the organism is growing in. In the laboratory, we often fall into the trap of using the tried-and-true methods without question. In our pursuit of new and exciting ideas we often forget to revisit the basics. Consider that the basic needs for an organism to survive are as follows: a source of carbon, nitrogen, trace minerals, metal cofactors, and water. When considering media formulation, a buffer is required as micro-organisms excrete compounds that alter the pH. We often use an undefined rich medium because it is the

standard procedure. However, I have noticed that even between nutrient rich growth media, bacteria grow differently. I discovered that various gene deletion mutants showed different growth phenotypes when grown on six different rich laboratory growth media (Figure 1). When I compared the two minimal medias we frequently used in the lab, I also noticed several minor differences (Figure 2). While these may not affect most experiments, reports have suggested there are circumstances where they matter[8,9]. For example, a ΔtolC mutant grown in M9 minimal media show a growth defect due to a lack of iron but show no such defect when grown on MOPS minimal media[8]. When designing MOPS minimal media, Neidhardt *et al.* (1974) placed careful consideration into every aspect of the formulation, with the goal to create a simple standard defined minimal media for the growth of Enterobacteriaceae[5]. It is interesting to note that it has been nearly 50 years since this formulation was first published, and yet it is still one of the most common minimal medias used.

When re-evaluating the media I used for my studies, I recognized that there are many factors that are difficult to account for, especially when using a diverse set of carbon sources. Although I was careful to standardize the pH of my media prior to use, addition of hydrochloric acid and sodium hydroxide would have slightly altered the chemical composition of my media. Many of the carbon sources I used were acids that could affect the buffering of the media. Indeed, when I checked the pH of the media with *E. coli* that has grown overnight, the pH changes dramatically depending on the carbon source. This phenomenon is unlikely to happen in the host where a variety of factors are involved in keeping pH constant. I ensured I had the same concentration of carbon in each

media by adjusting the concentration of my carbon source depending on the number of carbon atoms it had, but this would not matter if I was above saturating concentrations for the transporters for that compound. That being said, I believe it is worthwhile for every researcher to understand the different aspects of the media they are working with and its potential limitations. As a cautionary tale, researchers have dedicated considerable resources to study leads that ended up having no *in vivo* activity because the media they used did not represent the environment in the host[9]. In our lab, we showed that even the infection models we use may be flawed. Carfrae *et al.* (2019) demonstrated that although the biotin biosynthesis inhibitor, MAC-13772, showed no activity in a mouse model, when the biotin levels were changed to mimic the concentrations of biotin in humans, the drug showed efficacy on many pathogens[10]. Although macrolides typically have a spectrum of activity limited to gram-positive bacteria, Farha et al. (2020) showed that macrolides also work on gram-negative pathogens in the presence of sodium bicarbonate, the buffer ubiquitous in our body[11]. I showed in Chapter 3 that changing the carbon source alters the susceptibility of *E. coli* to many antibiotics by affecting drug uptake. By testing in the wrong media or using the wrong infection models, we are potentially missing antibiotics that would work *in vivo* on humans.

Designing a host mimicking screening media

In an ideal world, we would be able to have screening campaigns on pathogens in an environment mimicking their host infection site. Our lab has seen success screening *K. pneumoniae* in serum and *S.* Typhimurium in macrophages and macrophage-mimicking media[12,13]. Moving forward, drug screening campaigns should be completed on host

mimicking media or with certain aspects of the host physiology in mind. There are different formulations used to simulate nutrients in the host including ones for the oral cavity, artificial urine, artificial sputum, and ones that mimic the macrophage environment. We should be testing different pathogens in these media and comparing how gene expression *in vitro* compares to *in vivo* infections and modifying accordingly until they begin to match. This effort would be helped by pathogens that are well studied so that we know what each gene does and the best approach for how to alter its expression. While this idea would be nearly impossible to do by hand, computational models are already being developed to help facilitate this in *P. aeruginosa*[14].

When designing a screen, there is promise in picking the correct set of carbon and nitrogen sources, controlling the concentrations of the micronutrients, picking the correct pH and buffer, and limiting oxygen levels to the ones that exist in the host. A recent transcriptomics analysis on 32 human bacterial pathogens in 11 different human host mimicking stress conditions showed that *in vitro* stresses were able to recapitulate similar responses when compared to the transcriptomes of pathogens in *in vivo* infection models[15]. This means that with careful consideration when determining screening conditions, it could be possible to mimic some aspects of the relevant infection environment for a pathogen of interest. As we continue to build our understanding of host-pathogen interactions, I believe we can get to a point where drug screening is done on a synthetic media that mimics the host environment.

Potential targets for drug development

Based on our understanding of metabolism, and the genes that are involved in carbon acquisition and catabolism, there are a few candidate targets that could be ideal targets following more extensive studies. Bacteria need to have extensive regulatory control over their processes to survive, and many have global regulators that affect numerous genes. Targeting these regulators will cause pleiotropic effects because they affect many downstream reactions simultaneously, but this may also make it more difficult to develop resistances to. Although these transcription factors were often thought of as "undruggable", we can look to cancer research studies for inspiration[16]. There is now considerable research on metabolic targets to treat cancer, and many of these ideas could be applied to developing antibiotics[17]. There are seven major global regulators in *E. coli* that control numerous other downstream transcription factors and genes: CRP, ArcA, FNR, Fis, integration host factor (INF), H-NS, and Lrp[18]. Many of these regulators are also involved in the expression of virulence genes and will often affect the pathogenicity of bacteria. Although most of these regulators show minimal *in vitro* phenotypes, many pathogens show a defect in host infection when these genes are deleted. In this thesis I have focused on CRP for its role on CCR, however, there are many other transcription factors that regulate metabolism. If we think in terms of evolution, bacteria in the host are likely just looking to acquire the nutrients required to survive and proliferate. By this logic, if we target their ability to regulate their metabolism, they should no longer be able to infect a host.

If we assume that bacteria are primarily looking for food to survive, carbon acquisition and catabolic pathways become an attractive target. Bacteria can use a variety of substrates as carbon sources because they have a variety of enzymes that allow them to metabolize them into central metabolism intermediates. Since the human host lacks many of these enzymes, these can be ideal targets to avoid cytotoxicity. As discussed in Chapter 1, there is ongoing research in understanding how different strains of bacteria may use different carbon sources to infect a host[3,19]. Targeting the pathways only used by pathogenic bacteria has the additional advantage of sparing the commensal bacteria which provide benefits to the host. Targeting individual carbon pathways may not be viable because bacteria often show metabolic flexibility and may use multiple different carbon substrates[20]. However, carbon sources must be metabolized into a central carbon metabolism intermediate. This is where there is opportunity to block several carbon sources from being used at the same time. The proof of concept for this has already seen success as shown by the researchers who used media containing acetate to search for inhibitors of the glyoxylate shunt[21]. Another opportunity for a similar study may be to look for gluconeogenesis inhibitors to target UPEC. I've shown in Chapter 2 that Phosphoenolpyruvate carboxylase deletions show a growth defect when using TCA cycle intermediates as a carbon source[1]. If UPEC strains require a functional TCA cycle and gluconeogenesis for infection, there are many genes in these pathways that could be potential drug targets. As we start to learn more about the nutrients and carbon sources used by bacteria in the host, there will be more opportunities and pathways revealed for antibiotics to target.

Combination therapy

Perhaps a missed opportunity in targeting metabolism is to look for metabolic inhibitors that would synergize with inhibitors with other processes. Combination therapy is not a novel concept, and one commonly prescribed antibiotic is a combination of sulfamethoxazole and trimethoprim[22,23]. We have learned that metabolic enzymes are highly redundant and often targeting one enzyme rarely leads to growth defects[24]. However, in the field of bioengineering, deletions of metabolic enzymes have been used for decades to alter metabolic flux to different pathways to improve the yields of biomolecules[25]. Targeting enzymes that alter the metabolic flux to certain pathways may not be lethal in different *in vitro* environments but may cause the cell to be sensitive to other perturbations. In a study that looked for antibiotic hypersensitivity on the *E. coli* Keio gene deletion collection, many gene deletions in central metabolism sensitized *E. coli* to different antibiotics[26].

An example of a potential combination could be to direct flux away from the pentose phosphate pathway, then target something that requires precursors from that pathway. The pentose phosphate pathway generates a key metabolite required in LPS biosynthesis. By itself, this may not halt cellular growth, however, this may sensitize the bacteria to another compound that targets LPS biosynthetic pathways. By depriving a pathway of its precursor metabolites through altering central metabolism, you open up opportunities to find weak inhibitors that may not work alone. Of course, since metabolic flux is heavily dependent on the environment the bacteria is grown in, understanding the central nutrients available during infection is crucial in order to take advantage of this

technique. It is interesting to note that mutants in the LPS biosynthesis pathway such as *ΔlpcA*, *ΔrfaE,* and *ΔwaaF* grow very differently depending on the carbon source available[1]. During growth on adenosine as a carbon source, these mutants grow even better than wildtype, but show a growth defect when grown on TCA cycle intermediates as a carbon source[1]. This implies that altering metabolism and metabolic flux to pathways may affect downstream biosynthesis reactions that require intermediates from central metabolism.

Our lab has previously explored how nutrient biosynthetic genes can interact with the rest of the *E. coli* genome[27]. This study revealed interactions between nutrient transport and biosynthesis, as well as redundancies in many pathways[27]. Intriguingly there were also a large number of genes with unknown functionality that interacted with these pathways[27]. Following this, our lab has also explored these interactions under nutrient stress by using chemical probes[28]. This work revealed 81 synergistic and 105 antagonistic interactions between antibiotics and probes of diverse function[28]. As we now understand, treatment with antibiotics results in major metabolic changes in the bacteria[29]. By probing bacteria compounds that target various processes, we can induce metabolic changes which can make bacteria more susceptible to antibiotics. It will be interesting to revisit these studies through the lens of central metabolism to see how deletions in central metabolic pathways can interact with the rest of the genome and various antibiotic probes.

Final Remarks

With the urgent need for new antibiotics, researchers need to identify new unconventional screening conditions and targets. In my thesis I have outlined how simply changing the carbon source used in the media can vastly alter the genes required for growth as well as how antibiotics work on a micro-organism. Moving forward, I believe that these data will be valuable for researchers to identify potential targets for drug discovery. I am hopeful that with the vast amount of data generated every day by researchers around the world, we will one day arrive at a unified model of bacterial life, with complete understanding of regulation, genomes, and the requirements for survival in all environments.

Thinking back, it is perhaps not a coincidence that I identified phenotypes as a result of using adenosine as a carbon source in both my studies. Adenosine is an important molecule that is both essential and ubiquitous due to the prevalence of ATP. However, free adenosine base is often overlooked because of the focus on all the other metabolites it is incorporated into. Additionally, adenosine is typically thought of as a nitrogen source instead of a carbon source which makes it a peculiar choice in a study of carbon metabolism. I believe that the way forward in the study of metabolism is to challenge the norms, such as in media formulation. We should look in esoteric places for things that have often been overlooked or ignored. Bacteria grow everywhere, not just in the laboratory.

References

1. Tong, M. *et al.* Gene dispensability in *Escherichia coli* grown in thirty different carbon environments. *mBio* **11**, 1–20 (2020).
2. Ramsey, M. M., Rumbaugh, K. P. & Whiteley, M. Metabolite Cross-Feeding Enhances Virulence in a Model Polymicrobial Infection. *PLOS Pathogens* **7**, e1002012 (2011).
3. Conway, T. & Cohen, P. S. Commensal and Pathogenic *Escherichia coli* Metabolism in the Gut . *Microbiology Spectrum* **3**, (2015).
4. Lund, P., Tramonti, A., de Biase, D., Pasteur-Fondazione, I. & Bolognetti, C. Coping with low pH: molecular strategies in neutralophilic bacteria. *FEMS Microbiology Reviews* **38**, 1091–1125 (2014).
5. Neidhardt, F. C., Bloch, P. L. & Smith, D. F. Culture medium for enterobacteria. *J Bacteriol* **119**, 736–47 (1974).
6. Fang, F. C., Frawley, E. R., Tapscott, T. & Vázquez-Torres, A. Bacterial Stress Responses during Host Infection. *Cell Host & Microbe* **20**, 133–143 (2016).
7. Schaffer, K. & Taylor, C. T. The impact of hypoxia on bacterial infection. *FEBS J* **282**, 2260–2266 (2015).
8. Vega, D. E. & Young, K. D. Accumulation of periplasmic enterobactin impairs the growth and morphology of Escherichia colitolC mutants. *Molecular Microbiology* **91**, 508–521 (2014).
9. Pethe, K. *et al.* A chemical genetic screen in Mycobacterium tuberculosis identifies carbon-source-dependent growth inhibitors devoid of in vivo efficacy. *Nature Communications* **1**, 1–8 (2010).
10. Carfrae, L. A. *et al.* Mimicking the human environment in mice reveals that inhibiting biotin biosynthesis is effective against antibiotic-resistant pathogens. *Nature Microbiology* 1–9 (2019) doi:10.1038/s41564-019-0595-2.
11. Farha, M. A. *et al.* Overcoming Acquired and Native Macrolide Resistance with Bicarbonate. *ACS Infectious Diseases* **6**, 2709–2718 (2020).
12. Weber, B. S. *et al.* Genetic and Chemical Screening in Human Blood Serum Reveals Unique Antibacterial Targets and Compounds against Klebsiella pneumoniae. *Cell Reports* **32**, (2020).
13. Ellis, M. J. *et al.* A macrophage-based screen identifies antibacterial compounds selective for intracellular Salmonella Typhimurium. *Nature Communications 2019 10:1* **10**, 1–14 (2019).

14. Cornforth, D. M. *et al.* Pseudomonas aeruginosa transcriptome during human infection. *Proc Natl Acad Sci U S A* **115**, (2018).

15. Avican, K. *et al.* RNA atlas of human bacterial pathogens uncovers stress dynamics linked to infection. *Nature Communications 2021 12:1* **12**, 1–14 (2021).

16. Lambert, M., Jambon, S., Depauw, S. & David-Cordonnier, M. H. Targeting Transcription Factors for Cancer Treatment. *Molecules 2018, Vol. 23, Page 1479* **23**, 1479 (2018).

17. Murima, P., McKinney, J. D. & Pethe, K. Targeting Bacterial Central Metabolism for Drug Development. *Chemistry & Biology* **21**, 1423–1432 (2014).

18. Martínez-Antonio, A. & Collado-Vides, J. Identifying global regulators in transcriptional regulatory networks in bacteria. *Current Opinion in Microbiology* **6**, 482–489 (2003).

19. Rohmer, L., Hocquet, D. & Miller, S. I. Are pathogenic bacteria just looking for food? Metabolism and microbial pathogenesis. *Trends in Microbiology* **19**, 341–348 (2011).

20. Miranda, R. L. *et al.* Glycolytic and Gluconeogenic Growth of Escherichia coli O157:H7 (EDL933) and E. coli K-12 (MG1655) in the Mouse Intestine. *Infection and Immunity* **72**, 1666–1676 (2004).

21. Fahnoe, K. C. *et al.* Non-Traditional Antibacterial Screening Approaches for the Identification of Novel Inhibitors of the Glyoxylate Shunt in Gram-Negative Pathogens. *PLoS ONE* **7**, e51732 (2012).

22. Cottarel, G. & Wierzbowski, J. Combination drugs, an emerging option for antibacterial therapy. *Trends Biotechnol* **25**, 547–55 (2007).

23. Tamma, P. D., Cosgrove, S. E. & Maragakis, L. L. Combination therapy for treatment of infections with gram-negative bacteria. *Clin Microbiol Rev* **25**, 450–70 (2012).

24. Nakahigashi, K. *et al.* Systematic phenome analysis of Escherichia coli multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Molecular Systems Biology* **5**, 727–738 (2009).

25. Keasling, J. D. Manufacturing Molecules Through Metabolic Engineering. *Science (1979)* **330**, 1355–1358 (2010).

26. Liu, A. *et al.* Antibiotic Sensitivity Profiles Determined with an Escherichia coli Gene Knockout Collection: Generating an Antibiotic Bar Code. *Antimicrobial Agents and Chemotherapy* **54**, 1393 (2010).

27. Côté, J.-P. *et al.* The Genome-Wide Interaction Network of Nutrient Stress Genes in Escherichia coli. *mBio* **7**, (2016).

28. El Zahed, S. S. & Brown, E. D. Chemical-Chemical Combinations Map Uncharted Interactions in Escherichia coli under Nutrient Stress. *iScience* **2**, 168–181 (2018).

29. Stokes, J. M., Lopatkin, A. J., Lobritz, M. A. & Collins, J. J. Bacterial Metabolism and Antibiotic Efficacy. *Cell Metabolism* **30**, 251–259 (2019).

Figure Legends

Figure 1: A heatmap summarizing the growth of the *E. coli* non-essential gene deletion collection (Keio) grown on six different rich microbiological growth medias.

Figure 2: A scatterplot showing the growth of the *E. coli* non-essential gene deletion collection (Keio) grown on M9 minimal media with glucose (x-axis) compared to the growth in MOPS minimal media with glucose (y-axis).

Figures

Figure 1

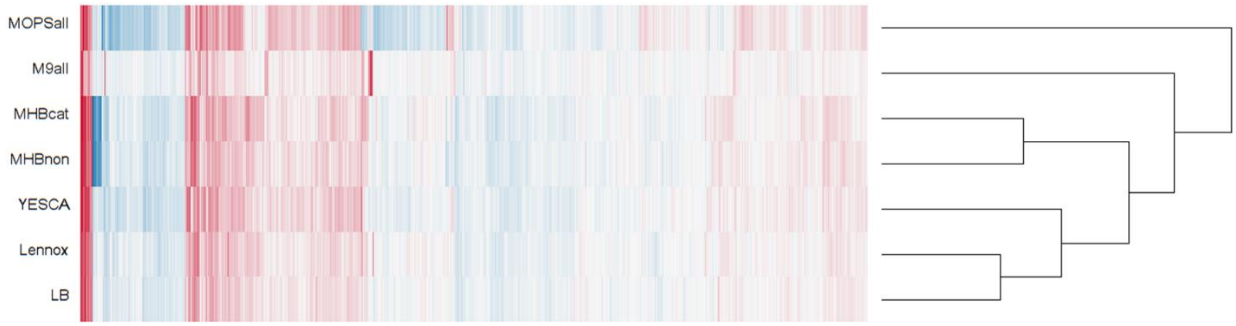
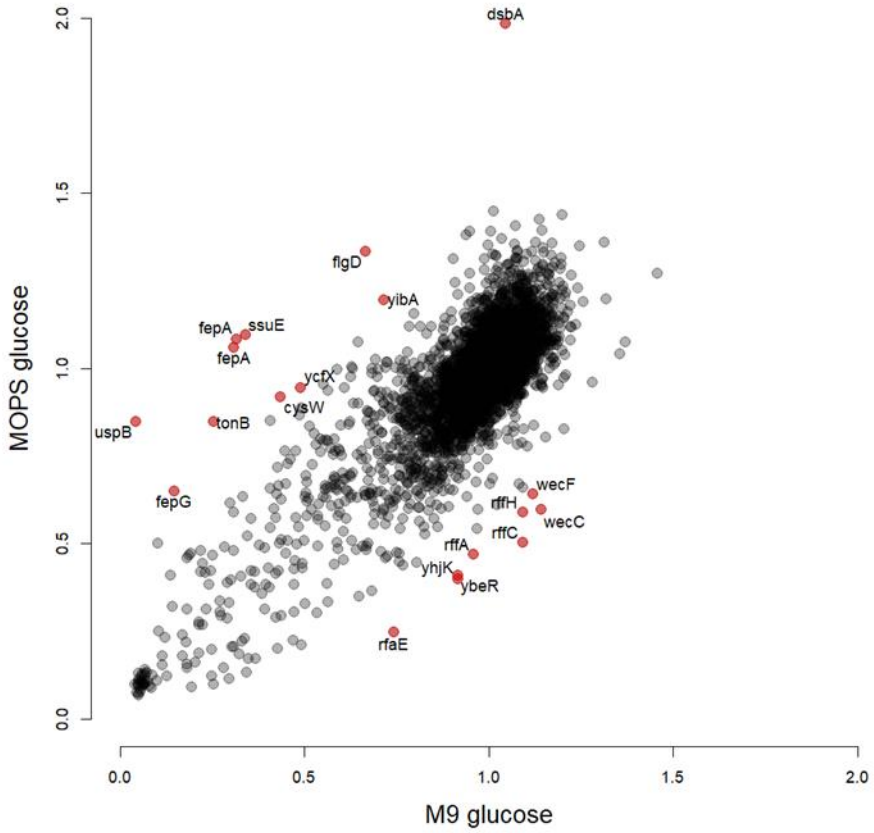


Figure 2


dsbA
flgD
yibA
fepA
ssuE
fepA
ycfX
cysW
uspB
tonB
fepG
wecF
rffH
wecC
rffC
rffA
yhjK
ybeR
rfaE
MOPS glucose
M9 glucose
0.0
0.5
1.0
1.5
2.0