

UCSF

UC San Francisco Electronic Theses and Dissertations

Title

Exploring Phenomics in Bacteria: High-throughput Phenotyping Drives Biological Discovery in Escherichia Coli

Permalink

<https://escholarship.org/uc/item/16g9w5dj>

Author

Nichols, Robert J.

Publication Date

2012

Peer reviewed|Thesis/dissertation

Exploring Phenomics in Bacteria: High-throughput Phenotyping
Drives Biological Discovery in Escherichia Coli

by

Robert J. Nichols

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Oral and Craniofacial Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

This work is dedicated to all of my family and friends, whose love and support inspired me to reach for new heights...

Acknowledgements

The Introduction and Future Perspectives (Chapters 1 and 3) of this thesis have been published as a book chapter in a volume entitled “Phenomics”, by Science Publishers (ISBN #978-1-57808-806-5).

The Research section of this Thesis was originally published in the January 7, 2011 edition of Cell. The article was titled “Phenotypic Landscape of a Bacterial Cell”, and is being published here under the terms of license agreement #2882741144479 with Elsevier.

Abstract

The genomic revolution of the last twenty years has launched the biosciences into a new frontier. For scientists working on many organisms, the availability of a genome sequence has dramatically accelerated research. The effects have been profound: Genetic engineering is now standard laboratory practice; evolutionary biologists can trace entire genomes; and genetic risk factors have been defined for many human conditions. However, our rapidly accelerating ability to collect and assemble DNA sequence information has greatly outpaced our ability to assign biological meaning to it. This imbalance creates a need for high-throughput methods aimed at developing leads to gene function: phenomic technologies. Phenomic technologies seek to associate DNA sequences with phenotypes in a high-throughput manner to gain a functional understanding of genetic code.

Bacteria are ideal candidates for phenomic analyses because of their amenability to genetic manipulation and ease at which they can be grown and analyzed under varying parameters *in-vitro*. In addition, even the best-studied prokaryotes like *E. coli* lack functional annotation for a large fraction of their genes. Therefore, bacteria present excellent experimental systems to establish the power of phenomic approaches to generate testable hypotheses of gene function en masse.

Using *E. coli* as proof-of-principle, we show that combining large-scale phenomics with quantitative fitness measurements provides a high quality dataset, rich in discovery. The identification of >10,000 growth phenotypes allowed us to study gene essentiality, discover leads for gene function and drug action, and understand higher-order organization of the bacterial chromosome. We highlight insights concerning a gene

involved in multiple antibiotic resistance and provide information about synergy of a broadly used combinatory antibiotic therapy, trimethoprim and sulfonamides. This dataset, publicly available at <http://ecoliwiki.net/tools/chemgen/>, is a valuable resource for both the microbiological and bioinformatic communities as it provides high confidence associations between hundreds of annotated and uncharacterized genes as well as inferences about the mode-of-action of several poorly understood drugs. This approach can be used for all culturable microbes with available ordered mutant libraries.

Table of Contents

Chapter 1: Introduction to Phenomics	1-39
Chapter 2: Phenotypic Landscape of a Bacterial Cell	40-71
List of Supplemental Material	72
Supplemental Figures	73-82
Supplemental Tables	83-168
Supplemental Text	69-179
Chapter 3: Future Perspectives	180-182
UCSF Library Release Form	183

List of Tables

Supplementary Table 1	83-88
Supplementary Table 2	89
Supplementary Table 3	90-91
Supplementary Table 4	92-94
Supplementary Table 5	95-111
Supplementary Table 6	112-149
Supplementary Table 7	150-168

List of Figures

Figure 1	46
Figure 2	49
Figure 3	51
Figure 4	53
Figure 5	57
Figure 6	60
Figure 7	63
Supplementary Figure 1	73
Supplementary Figure 2	75
Supplementary Figure 3	77
Supplementary Figure 4	79
Supplementary Figure 5	81

Chapter 1: Introduction to Phenomics

I have researched and written this chapter independently. All figures are my original creations. Carol Gross assisted with revisions and edits.

Introduction

The genomic revolution of the last twenty years has launched the biosciences into a new frontier. For scientists working on many organisms, the availability of a genome sequence has dramatically accelerated research. The effects have been profound: Genetic engineering is now standard laboratory practice; evolutionary biologists can trace entire genomes; and genetic risk factors have been defined for many human conditions. However, our rapidly accelerating ability to collect and assemble DNA sequence information has greatly outpaced our ability to assign biological meaning to it. This imbalance creates a need for high-throughput methods aimed at developing leads to gene function: phenomic technologies. Phenomic technologies seek to associate DNA sequences with phenotypes in a high-throughput manner to gain a functional understanding of genetic code. In general, phenomic technologies require: 1.) a method to create genomic variation; 2.) an ability to adjust the experimental environment, and 3.) a technique for observing and recording the phenotypes of interest (Fig. 1).

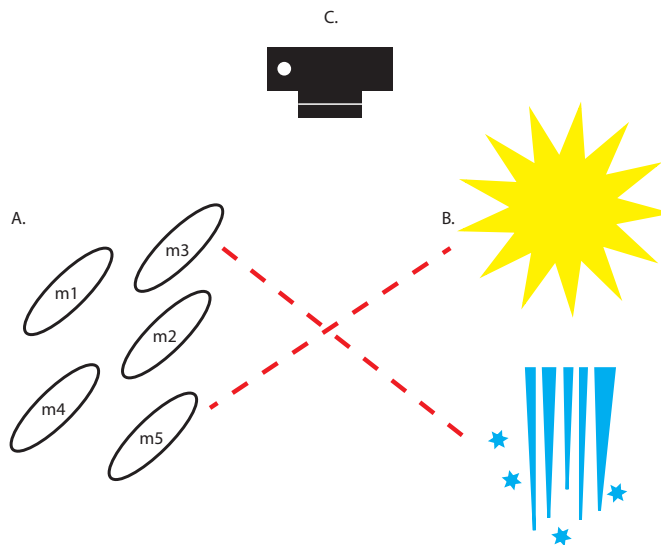


Figure 1. The cornerstones of Phenomic Analyses. All phenomic approaches require a.) a method to create a genomically-variable strain collection. Here, five isogenic mutants are depicted as m1-m5. b.) a means to vary the experimental environment, and c.) a technology to capture and record condition-gene interactions (phenotypes), depicted here as red dotted lines.

Bacteria are obvious candidates for phenomics because there are easy ways to meet each phenomics requirement. Researchers have been manipulating bacterial genomes for more than 50 years using chemical and transposon mutagenesis, and, in some species, targeted chromosomal engineering. In addition to the relative ease of genetic engineering, bacteria do not present dosage or zygosity issues, as they are haploid organisms. Thus, the interpretation of genotype-phenotype relationships is relatively straightforward. Manipulating the environment of culturable bacteria is also quite simple. Chemical and drug stresses can be mixed directly into rich media, metabolic conditions can be tuned within defined media, and environmental factors like temperature, gas, and humidity can also be easily adjusted in incubators. *In-vivo* analyses benefit from tools like germ-free and disease model mice. Techniques for capturing phenotypes resulting from genome-environment interactions have developed quickly over the last ten years. Bacterial phenomic studies have now been published using a variety of readouts: colony size, growth rate, cellular respiration, strain abundance, survival, host colonization, and host clearance are at the beginning of a growing list. In all, a variety of options exist to manipulate bacterial genomes, control growth environments, and read out phenotypes in a high-throughput manner. The suite of available tools, and the unique biological qualities of bacteria, position them as ideal subjects for developing new phenomic technologies.

This chapter will begin with a review of experimental technologies utilized in each of the three cornerstones of phenomics: creating genomic variation, adjusting growth environment, and capturing phenotypes. We briefly discuss the computational strategies used to analyze phenomic data, and then turn to a series of case studies that represent both the breadth and power of phenomic analyses in bacteria. We conclude with a look at the future and the coming advances in bacterial phenomics.

Experimental Technologies

Creating Genomic Variation

Bacterial phenomic approaches generally begin with the assembly of a genetically variable strain library, which is then screened or selected under a variety of growth conditions to search for specific condition-gene interactions, or phenotypes. Key to these strategies is that the genetic variability between strains is known or easily quantifiable, and as limited as possible. Thus, engineered libraries based on the techniques reviewed below should be constructed in a consistent strain background. Approaches based on natural genetic variation should put a premium on describing as much of the variation between strains as possible. This way, when phenotypes are detected, they can be traced to the known genetic variant(s) with high confidence. If significant genetic variability exists between strains in the same library (i.e. if mutant strains were constructed in varying strain backgrounds), it can be nearly impossible to detect meaningful condition-gene interactions.

There are several ways to establish genetic diversity for the screen, but all engineered genetic modifications can be classified as either loss-of-function (LOF) or gain-of-function (GOF). Simply put, LOF genetic approaches reduce the expression or activity of genomic DNA sequences, while GOF approaches either increase the function of native genes through overexpression or previously characterized GOF alleles, or add function by introducing non-native (heterologous) sequences into the cell. LOF mutant libraries are the most common means to establish genetic diversity for phenomic screens in bacteria, as LOF genetic analyses have a long history of revealing phenotypes that ultimately point toward gene function. GOF approaches based on overexpression have been widely used to search for drug-target relationships, while those based on

heterologous expression have been employed in functional metagenomic applications. In this section, we first review each of the seven methodologies currently used to establish genomic variation in phenomics and then explore whether the resulting mutants are archived or pooled, whether they are used for screens or selections, and the relative merits of each strategy.

1.) Transposon Mutagenesis

Transposon mutagenesis is widely employed to generate mutant libraries. Here, a transposon sequence containing a selectable marker is randomly inserted into the genome by a transposase enzyme. For competent bacteria, the selectable marker can be transposed into the host genome *in-vitro*, followed by a transformation and selection procedure. For less competent bacteria, the transposon and transposase are encoded on a plasmid transformed into host cells, where translation of the transposase enzyme, and subsequent insertion of the transposon DNA into the host genome occur *in-vivo*. selection is then carried out to purify cells that have received a transposon insertion. Transposon mutagenesis technologies are reviewed in (Mazurkiewicz et al., 2006).

Most random transposon insertions occur in a coding region because of the high density of ORFs in bacterial genomes. Occasionally, transposon insertions occur in intergenic and promoter regions. Insertions of this type may have interesting genetic effects beyond the complete loss-of-function phenotype offered by engineered deletion mutants and may also reveal previously unannotated functional elements. However, the randomness of transposon mutagenesis also presents challenges: the genomic insertion sites must be identified; follow-up experiments are necessary to determine whether insertions yield LOF or GOF proteins; and the “average” phenotype of independent insertions within a single gene must be determined. The issue of polarity is discussed in

targeted chromosomal engineering, and the strategies for mapping insertion sites are discussed in the pooled screening technology section.

2.) Targeted Chromosomal Engineering

Genetic engineering of the chromosome via natural competence or induced recombineering is possible in some bacterial species. For these, many options exist for building LOF libraries. The most straightforward is to build a single-gene knockout library (Baba et al., 2006; de Berardinis et al., 2008; Santiviago et al., 2009b). In these libraries, a single gene is deleted in each strain, covering all nonessential genes of the genome. The gene is replaced with an antibiotic resistance cassette for selection, and this cassette can also contain a DNA barcode for use in pooled approaches (discussed below). While comprehensive single-gene knockout libraries have proven to be extremely powerful tools in many organisms (Nichols et al., 2011; Roguev et al., 2008; Santiviago et al., 2009a; Schuldiner et al., 2006), they are inherently limited by the annotation of the target genome. More specifically, functional DNA sequences yet to be annotated, like those coding for small proteins or small RNAs, will not be interrogated by most targeted engineering approaches. Herein lies a central tradeoff in targeted vs. random mutagenesis: targeted knockouts are clean (complete or controlled LOF), while random mutagenesis can reveal previously undetected functional sequences. Both random mutagenesis strategies and targeted knockouts may have unanticipated transcriptional and/or translational polar effects due to operon structure, potentially clouding interpretation of phenotypes. Follow-up studies are essential to verify the suspected genotype-phenotype relationship.

Chromosomal engineering is also used to modulate gene expression. In trackable multiplex recombineering (TRMR) (Warner et al., 2010), a synthetic promoter and

ribosome-binding site (RBS) is recombined into the chromosome to replace the native promoter of a gene. These constructs may encode strong promoter/RBS sequences conferring gain-of-function (GOF), or weak ones resulting in LOF. In addition, it is possible to knock in tunable promoter sequences, allowing for analysis of mutant strains under various promoter activities on a genome-wide scale.

It is also possible to establish genetic variability in the strain library through collection and manipulation of previously constructed alleles: point mutations, small deletions, insertions, truncations, etc. Such alleles can be either LOF or GOF, and need not be previously functionally characterized. Regardless, it is essential that the alleles be transferred into a common strain background by either phage transduction or targeted chromosomal engineering to maintain an isogenic library (Nichols et al., 2011).

3.) Antisense RNA

In recent years, antisense RNA (asRNA) has been an effective tool to knockdown gene expression in some Gram-positive bacteria (Donald et al., 2009; Huber et al., 2009; Phillips et al., 2011). Very recently, the first asRNA-based screen in Gram-negative bacteria has been reported (Meng et al., 2012). Thus far, all approaches are plasmid-based, with expression of the targeted asRNA induced by an exogenous small molecule. Presumably, the asRNA transcribed from the plasmid base pairs with its complementary message, and either causes transcript degradation or interferes directly with ribosome binding. Because a variety of asRNA mechanisms exist (and many asRNAs are ineffective), asRNAs must be screened and validated before being used in a phenomic screen or selection. Because asRNA mechanisms and effectiveness are poorly understood in bacteria, there is an “off target” caveat to all phenotypes. The effects of a single asRNA may reach beyond reduction of its target mRNA levels and impact the

transcription/translation of other genes in the cellular network. Therefore, phenotypes detected using asRNA approaches must be verified through careful follow-up studies. In addition, all asRNA strains should be analyzed for response to the inducer molecule. Differential sensitivity across strains can complicate strain fitness measurements, especially in pooled approaches. One effective strategy to normalize for this effect is to pool only those strains sharing the same response kinetics to the inducer molecule (Donald et al., 2009).

4.) Chemical Mutagenesis

Many bacterial species are culturable yet lack genetic tools. For these, chemical mutagenesis can be the only laboratory means to generate genomic diversity. Chemical mutagenesis has not previously been useful for phenomics, because of the lack of a high throughput methodology to identify base changes. However, the increased efficiency and cost effectiveness of whole-genome sequencing (WGS) methodologies coupled with the small genome size of bacteria now allow chemically-induced mutations to be identified through WGS. The first report of this approach is a study of the virulence determinants of the obligate intracellular bacterium *Chlamydia trachomatis* (Nguyen and Valdivia, 2012).

5.) Overexpression

Overexpression of genes has been used in phenomic approaches known as high-copy suppressor screens. Such screens have generally been used to search for genes whose overexpression increases resistance to a drug or stress of interest. These phenotypes can suggest drug-target relationships, and have proven quite powerful in drug discovery efforts. Overexpression is usually achieved through plasmid-based systems, but

chromosomal alleles that increase the expression and/or stability of the mRNA and/or protein, or increase protein activity could also be used.

For plasmid-based systems, the plasmid bearing the ORF of interest is transformed into host cells and maintained with the appropriate selective agent. Overexpression of the ORF can be either constitutive or inducible. In the latter case, a small molecule added to the culture activates promoter expression. Overexpression screens must be optimized to find a level of overexpression that is effective, but not toxic to the cell. High levels of overexpression provide a potent selection for spontaneous mutations within the expression plasmid or elsewhere in the genome (suppressors) that can compromise the controlled genetic variability necessary to execute an effective phenomic screen. Though less common than LOF libraries, a GOF library based on overexpression has been constructed in *E. coli* (Kitagawa et al., 2005), and has been utilized for phenomic studies (Couce et al., 2012; Pathania et al., 2009).

6.) Heterologous expression

Metagenomic sequencing projects have been a major component of the bacterial sequencing boom in recent years. Using environmental samples with many species, these projects sequence all DNA to generate vast collections of ORF sequences, many of which have no known function or genomic source. However, many researchers have begun to recognize that such sequences may code for medically (antibiotic-resistance) or commercially (metabolic, remediation) valuable genes. Therefore, phenomic approaches aimed at deciphering the function of such genes are being developed. Functional metagenomic approaches rely on heterologous expression of the sequenced genes of interest in a controlled system (like *E. coli*). In these approaches, the genome

of the host strain is held constant, and expression of the unknown ORF creates genetic diversity in the strain library (McGarvey et al., 2012; Sommer et al., 2010).

7.) Natural variation

Natural variation between strains can also create genetic diversity of the strain library. In some settings, it may be desirable to understand the genetic basis of a trait possessed by only certain strains of a given organism (i.e. host colonization or metal reduction). In such cases, harnessing the natural variation between strains can be a powerful route to functional discovery. Here, isogenicity between strains becomes impossible, but screening closely related strains can allow meaningful conclusions to be drawn.

Evolved strains can also be utilized for studying the functional impact of natural variation. Bacterial strains can be evolved in the lab (Tenailon et al., 2012) or in a host (Fabich et al., 2011). Evolved strain approaches can be particularly powerful, as a high degree of isogenicity can be maintained between the evolved and parent strains, keeping phenotypic analyses relatively controlled. However, such strategies demand either whole-genome sequencing or polymorphic analyses (i.e. SNP) of all strains.

Adjusting the Experimental Environment

Phenotypes are classically defined as the output of interaction between genome and environment. By systematically varying the genome and the environment, phenotypes detected in high-throughput phenomic approaches represent functional interactions between genes and environments. A variety of approaches for modulating the genome are described above. Here, we review techniques for varying the environment.

The probability of phenotypic discovery scales with the degree of both genetic and environmental variability screened. Thus, the more unique environments under which the genetically-diverse strain library is evaluated, the greater the likelihood of detecting an interaction between a given gene and environment. Therefore, it is desirable to screen as many environments as possible. However, careful selection of those environments may also increase the odds for discovery. The establishment of genetic diversity in the strain library gives the researcher power to query the function of nearly all parts of the genome. The selection of environments allows the researcher to test what each of those parts may have evolved to do.

Phenomic studies can apply various types of stresses to the strain library. Chemical stresses including antibiotics and other drugs, mutagens, detergents, etc. can be applied (Couce et al., 2012; Donald et al., 2009; Flores et al., 2005; Gallagher et al., 2010; Girgis et al., 2009; Gomez and Neyfakh, 2006; Kohanski et al., 2008; Liu et al., 2010; Nichols et al., 2011; Pathania et al., 2009; Tamae et al., 2008), as can existing natural product libraries (Phillips et al., 2011). On the other hand, environmental stresses like temperature, pH, UV light exposure, and metabolic conditions can be effective in revealing phenotypes (Deutschbauer et al., 2011; Ishii et al., 2007). The screening environment may also be perturbed by the addition of another organism, such as the infection of bacteria with phage (Maynard et al., 2010), or the infection of an animal host (Bianconi et al., 2011; Gawronski et al., 2009; Goodman et al., 2009; Hensel et al., 1995; Hisert et al., 2004; Lawley et al., 2006; Potvin et al., 2003b; Santiviago et al., 2009a), or cell line (Camacho et al., 1999; Nguyen and Valdivia, 2012) with bacteria.

Importantly, not all stresses can be applied in all experimental systems. For example, solubility and/or volatility properties may dictate whether a chemical stress can be

assayed in solid and/or liquid-based setups. Chemicals and certain environmental conditions assayed *in-vitro* may not be tolerated by an animal host, and thus be incompatible with *in-vivo* analyses. Therefore, selection of stresses for a given phenomic study must take into account the capacity and potential limitations of the experimental system. For all types of stress, the dose or concentration used in the assay can have major effects on the results. High stress load can increase the likelihood of spontaneous suppressor mutations, complicating the interpretation of phenotypes, while low stress load may fail to separate less or more fit mutants from the population. Therefore, pre-screens aimed at quantifying and optimizing the stress load for each unique stress can be extremely beneficial for downstream phenotypic discovery.

Finding the Phenotype: Screens/Selections, Readouts, and Scoring

A multitude of readouts can be used to monitor diverse processes and states for phenomic analysis. High-throughput microscopy can be used to monitor morphological phenotypes or protein localization. Fluorescent or other colorimetric reporters can be used to monitor diverse processes including transcriptional activity, metabolic capacity, developmental processes and metal reduction. Additional phenotypes related to developmental cycles can be monitored by specialized readouts (i.e. sporulation by resistance to toxic chemicals). However, the most inclusive, and by far most-utilized readout to this point has been fitness, which itself has been quantified in a variety of ways.

The typical objective of phenomic analyses is to identify specific condition-gene interactions. By normalizing out the mutation-specific effect, it is possible to identify environments where a distinct mutation causes a more (or less) severe phenotype than

is normally observed for the same mutation under baseline conditions. Currently, there are two main ways to screen for these events. The first, called an arrayed approach, evaluates each strain in isolation under a battery of growth environments. The second, called a pooled approach, evaluates the ability of each strain in the library to compete with all other strains under the set of growth environments. In addition, it is possible to identify condition-gene interactions through a variety of selection techniques.

Just as a number of options exist for establishing genomic and environmental diversity, many experimental setups have been used successfully in bacterial phenomics, all of which can be divided roughly into forward and reverse genetic approaches. Forward genetic approaches begin with a phenotype of interest and search for genomic regions contributing to that phenotype, using technological advances that simplify identification of the genomic region(s) of interest. In general, these analyses use deletion or transposon libraries that are pooled and then subject to screens or selections. Reverse genetic approaches begin with a discrete set of genetically variant strains, and seek to associate phenotypes with each one. These analyses typically utilize arrayed screening approaches.

Arrayed Screens

In an arrayed approach, individual strains are grown in isolation, and evaluated relative to their baseline behavior. Phenotypes are defined as deviations from that baseline. The power of arrayed approaches, which can be conducted in liquid culture or on solid surfaces, comes in the control of the screen and potential robotic integration for high-throughput screening. Because the strains are grown and evaluated against themselves, the deviation of a given strain from its baseline behavior is clear: the interaction of a specific mutation and a specific environment produces the observed phenotype. While

some strain competition does exist in solid surface arrayed approaches, it is tightly controlled relative to pooled approaches, as each strain competes only mildly with the strains that surround it. The surrounding strains are normally held constant throughout the entire screen, allowing low-level “neighborhood” competition to be normalized away. Strain competition is completely absent from liquid arrayed approaches. Liquid handling (for liquid culture setups) or colony pinning (for solid surfaces) robots automate the screening process and provide technical accuracy and reproducibility (Fig. 2).

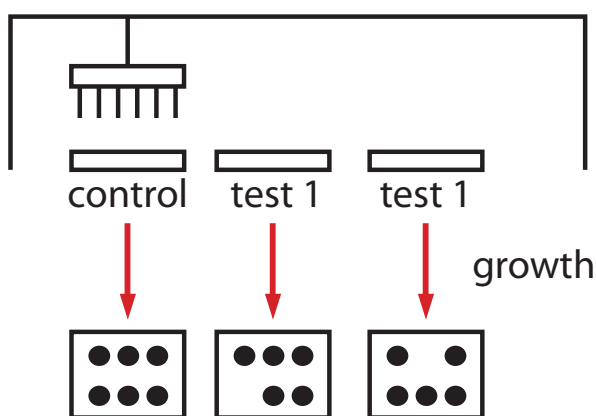


Figure 2. Arrayed approaches generally require colony-pinning or liquid-handling robotics. These platforms deliver high-throughput scaling and technical accuracy through highly-reproducible aliquoting of cells to control and test conditions.

In liquid culture applications, growth is usually read by calculating the rate of exponential growth (Nakahigashi et al., 2009), endpoint culture density (Kohanski et al., 2008), or cellular respiration (Bochner, 2009; Fabich et al., 2011). For solid surface approaches, fitness is normally read out as colony size. This non-traditional indicator of fitness is quite powerful as it reports on duration of lag phase, the rate of log phase growth, and the time of onset of stationary phase (Nichols et al., 2011).

Arrayed Screen Scoring

Arrayed approaches are based on distributing a collection of strains onto a solid surface or into microtiter plates for liquid culture. In both cases, the media is infused with a condition/stress of interest, and growth is allowed to proceed for a pre-determined incubation period. For arrayed approaches utilizing a quantitative readout (colony size, culture density, growth rate), a global scaling normalization must be employed to account for differences in stress load and absolute growth. This normalization centers the distribution of all readouts observed for each condition on a common value, making cross-condition comparisons feasible.

It is important to accurately estimate the control, or baseline behavior of each strain in order to detect specific condition-gene interactions. Two techniques exist for estimating control behavior in the context of arrayed approaches: median-centering and control conditions. A central assumption of median-centering is that condition-gene interactions (phenotypes) are rare. Therefore, if enough unique conditions are screened, it is possible to estimate the control behavior of a strain by simply identifying the median of all quantitative readouts for that strain. This approach has been effective in large-scale phenomic and genetic interaction studies of bacteria and yeast, demonstrating that the median of a large number of experiments is a robust indicator of control behavior (Collins et al., 2007; Nichols et al., 2011; Roguev et al., 2008; Schuldiner et al., 2006; Typas et al., 2008).

When a smaller number of conditions are screened, the median of all readings is not a reliable estimate of control. In these cases, an appropriate control condition must be identified and screened in parallel with the stress conditions of interest. Here, the behavior of each strain under control conditions establishes its control behavior. One disadvantage of this method compared to median centering is the control estimate is

based on a comparably low number of readings (low “n”). Therefore, it is essential that the control condition is screened a sufficient number of times for each strain to establish a statistically robust number of readings (high “n”), as estimates of control behavior should always be accompanied by corresponding measures of variance.

After control behavior and variance have been estimated for each strain, fitness scores may be calculated for each condition screened relative to the control measures.

Recently, this has been done using a modified t-statistic known as the “S score” (Nichols et al., 2011) or other statistical metrics based on the normal distribution (Kohanski et al., 2008). The S score procedure calculates fitness scores for a given strain by comparing the average growth of that strain across a replicate series of a given condition to the control growth of that strain. Variance estimates associated with the replicate series and control growth are used to modify the confidence associated with the estimated score (high variance functions as a penalty). Replicate experiments come in two forms: strain replicates and screen replicates. Strain replicates require construction and isolation of independent clones of the same mutant strain. These replicates help control for the possibility that additional mutations may arise either during strain construction or during the screening process as the cell evolves to compensate for the primary mutation. The more strain replicates screened, the lower the likelihood that the same secondary mutation(s) will arise and cloud the results. Screen replicates are technical replicates of the same screen to control for day-to-day variation of potentially variable components of the screening process (e.g. stress concentration, colony pinning accuracy, media quality, imaging parameters, human error).

Pooled Screens

Pooled screens evaluate the ability of one strain to compete with all other strains in the library in batch culture under a given growth environment. Pooled approaches work very well with forward genetic approaches like transposon mutagenesis, and thus are available for use in a wide range of bacterial organisms. However, arrayed libraries can also be used for pooled analyses, and offer the dual benefits of increased control over starting cell inoculum for individual mutants and availability of individual mutants for follow-up study after the screen.

Pooled approaches are best positioned to evaluate strain fitness, or growth, but have also been used to evaluate cellular processes like motility (Girgis et al., 2007) and biofilm formation (Amini et al., 2009). Fitness evaluation is based on reading out the relative abundance of every mutant strain in the batch culture at the start and endpoint of each competition (Fig. 3). Just as described for arrayed approaches, phenotypes are defined as deviations from the baseline behavior of each strain. In pooled approaches based on growth, baseline behavior is estimated from the typical frequency of a given strain in the pool.

Pooled approaches are powerful for several reasons. First, they scale incredibly well without the need for fancy robotics. Since each condition screened requires a single batch culture, a single person can evaluate hundreds of conditions rather easily. Second, pooled approaches can evaluate thousands of strains simultaneously in the same culture. Therefore, a tremendous amount of genomic diversity can be assayed under many conditions. Pooled approaches therefore represent the most efficient way to screen the largest amount of condition-gene interactions for potential phenotypes.

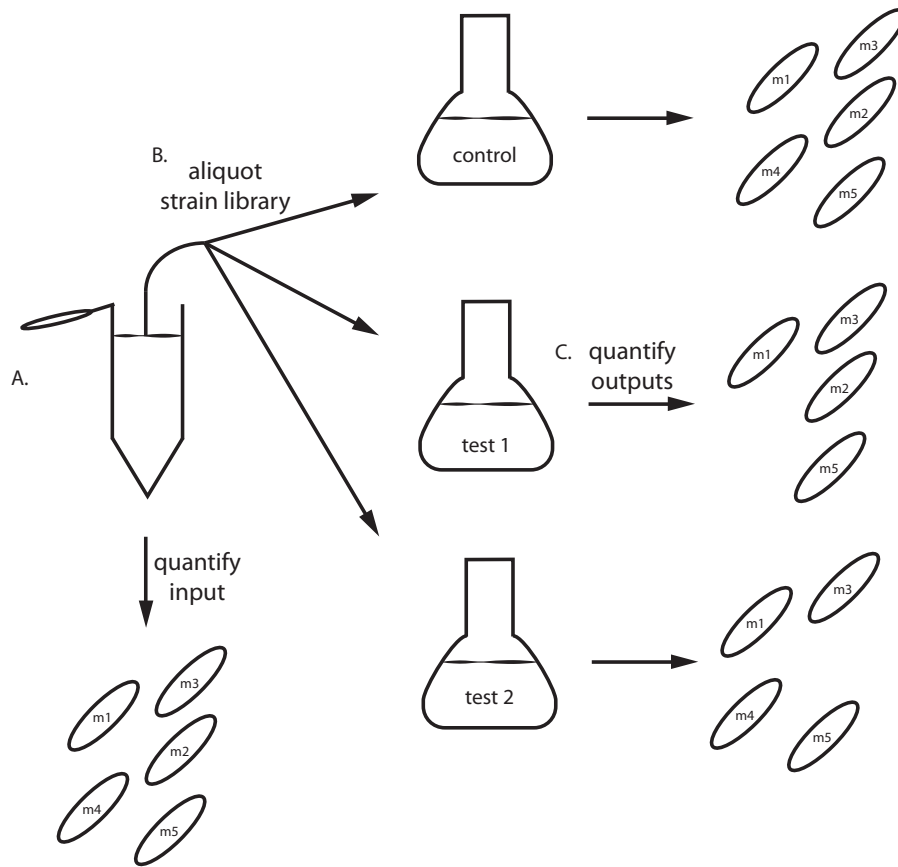


Figure 3. Workflow of Pooled Growth Approaches. A.) The starting strain library is sampled and quantified via one of the techniques discussed in the Pooled Screen section. B.) Aliquots of the strain library are distributed to test and control condition cultures. C.) Competitive cultures are grown for a set time, and then sampled and quantified according to the procedure established for the input library.

The major challenge of such approaches is teasing apart the confounding factor of strain competition on phenotypic evaluation. Because pooled approaches often compete thousands of strains against each other, the ability to quantitatively estimate changes in a given strain's fitness is influenced by the potential fitness changes of every other strain in the pool. For dramatic fitness changes, this is unlikely to prevent detection. But for more moderate effects, especially for fitness loss, competition within the pool has been reported as a potential confounding factor (Girgis et al., 2009). An additional caveat to pooled analyses is that certain mutations may exhibit phenotypes only in the context of

the complex community. For example, a LOF mutation that disables production of an essential metabolite may be tolerated, and even beneficial, if that mutant can scavenge the metabolite from other members of the community. While this scenario can complicate interpretation of pooled results, it also represents an advantage of pooled approaches. They report a physiologically relevant analysis of bacteria within a genetically-diverse community.

All pooled approaches based on transposon mutant libraries require a readout technology capable of quantifying the abundance of each unique strain in the pool. Different transposon systems have been used for phenomic studies in bacteria, and have typically been paired with complementary readout technologies. Initial transposon studies were based on Signature-Tagged Mutagenesis (STM) systems. In these systems, each insertion mutant carries a unique DNA barcode, so that a given barcode sequence can be assigned to a specific genomic insertion site. Therefore, when quantifying strains after a screen or selection, one need only to quantify the abundance of each barcode. STM approaches were the first to break through in phenomic applications, because these barcodes allowed a high-throughput readout based on DNA blotting (Hensel et al., 1995), multiplex PCR (Potvin et al., 2003a), or microarray hybridizations (Deutschbauer et al., 2011). However, the need for a large amount of unique DNA barcodes requires a great deal of transposon template synthesis, making library creation a laborious task.

Second generation transposon-based strategies focus on eliminating the need for a unique transposon sequence for each mutant. These technologies need only one transposon sequence, and rely on the genomic DNA flanking the insertion site to serve as a barcode. Three methods exist for the detection of transposon-flanking genomic

DNA signal in insertion sequencing-based approaches: 1.) type-IIIS restriction enzyme (mariner Tn), 2.) selective PCR using a Tn-specific primer, and 3.) the circle method. For all three, the resulting transposon-flanking genomic DNA fragments can be analyzed via hybridization to a full-genome tiling array or short-read sequencing with transposon and/or adapter-specific primers.

Some engineered Mariner transposons have Mmel restriction sites in their inverted repeat sequences of the transposon ends (van Opijnen et al., 2009). Mmel is a type-IIIS restriction enzyme, and cuts 20 base pairs downstream of its recognition site in the genomic DNA just outside the transposon-genome junction. Adapters can then be ligated to the cut ends, and the genomic DNA amplified with one primer specific to the transposon and the other specific to the adapter. This approach has been widely used.

Selective PCR can also be run without specific cutting by Mmel. Genomic DNA can be sheared, adapters ligated onto blunt ends, and selective PCR used to enrich transposon-genome junction fragments as described above for Mariner approaches. However, the absence of the Mmel digestion means that far more DNA fragments are ligated to adapters, and non-specific PCR can be problematic. Therefore, some studies have added an additional purification step by biotinylation of the transposon-specific primer and affinity purification following PCR (Gawronski et al., 2009).

More recently, the Tn-seq circle method has been developed (Gallagher et al., 2010). In this method, genomic DNA is sheared, and a single adapter is ligated to all free ends. After adapter ligation, fragments are digested with a restriction enzyme that cuts inside the transposon sequence (near one end), and circularized via templated ligation. Then, all circularized fragments contain part of the transposon sequence ligated to the generic

adapter, and surrounded by the genomic DNA of the transposon-genome junction. PCR is then used to amplify those fragments.

Pooled Screen Scoring

In pooled competitions, the baseline behavior of a given strain reflects the fitness associated with that mutation as indicated by the ability of the mutant to compete with all other mutants in the pool. To estimate these controls, one may simply examine the frequency of each mutant in the pool before selection, as these values reflect the relative ability of each mutant to compete during selection and outgrowth of the library (Deutschbauer et al., 2011). A second strategy is to base control estimates on the endpoint strain abundance measures in competitive cultures without stress (Langridge et al., 2009). These “control cultures” can be run in parallel to competitions under stress. It is also possible to estimate control behavior based on both time-zero and control competition results (Girgis et al., 2009).

Pooled competitions normally utilize strain libraries generated via transposon mutagenesis, though such libraries can also be arrayed (Cameron et al., 2008; Deutschbauer et al., 2011; Gallagher et al., 2007; Liberati et al., 2006). As is the case with arrayed approaches, strain and screen replicates are necessary to accurately estimate phenotypes. Transposon libraries have an advantage over arrayed engineered libraries in that saturating mutagenesis can allow for hundreds of mutants per gene to be constructed and analyzed. Therefore each experiment has many “biological replicates”, providing a very high “n” for strain replicate readings in such approaches, and thus increased confidence in the effects of mutating a given gene. However, it is still important to perform “technical” replicate screens to control for day-to-day variation, inoculum effects, and human error.

Conditional strain fitness in pooled competition experiments is usually represented with a calculated Z score. Z scores have been calculated using different metrics, but all attempt to quantify the deviation of the conditional behavior of a strain from its control behavior by calculating the mean and standard deviation of all conditional and control frequency measures for that strain (Deutschbauer et al., 2011; Girgis et al., 2009). Thus, the Z score, S score (reviewed in arrayed approaches), and T-statistic are extremely similar.

Selections – *In-vivo* Transposons and Functional Metagenomics

In-vivo phenomic approaches based on transposon mutant libraries often use positive or negative selections to identify insertion mutants of interest. The techniques used to identify the selected insertion mutants are discussed above in the pooled screen section. Negative selections can be used to identify genes necessary to establish infection or maintain host colonization, as the host organism can select against LOF insertion mutants in such genes (i.e. following infection/colonization particular transposon mutants are depleted in the population). Positive selections can identify genes detrimental to such processes and suggest host-adaptive strategies used by the bacteria.

Positive selections are ideally suited for functional metagenomic approaches.

Metgenomic ORFs of interest may be cloned into expression vectors, and then transformed into a host strain unable to grow under the selective condition. Then, the transformed strains can be pooled and plated under the given condition to select for clones that have survived. The expression vectors in the surviving clones can then be sequenced to identify the gene conferring survival.

Data Analysis - Benchmarking and Exploration

Quantitative phenomic datasets can be large and complex, and require computational strategies to: demonstrate data quality through benchmarking analyses, and then search for functional insights in the data via exploratory techniques.

Benchmarking

Upon completion of a phenomic screen and assembly of a compendium of phenotypic scores, benchmarking analyses are performed to assess data quality and potential functional predictive power. One commonly used metric to evaluate data quality is the correlation between replicate measurements. This can be calculated either as the correlation of all strain measurements across any pair of screen replicates (Deutschbauer et al., 2011; Nichols et al., 2011), or the correlation of any pair of strain replicates across all screens. Such correlations can demonstrate both the reproducibility of the data and the strength of phenotypic signal (high correlations can generally not be achieved without signal).

Receiver operating characteristic (ROC) plots are extremely powerful for determining the functional predictive power of the dataset. These plots show the relationship between the true positive rate and the false positive rate at regularly-spaced intervals along a continuous scale (in this case, the correlation coefficient scale). Therefore, the calculation of such plots requires a “gold standard” index classifying pairs of genes as true positive (functionally-related) or true negative (unrelated). For bacteria, a variety of classifiers may be used to define the gold standard index: shared operon membership can suggest a functional link between pairs of genes; protein-protein interaction datasets report on pairs of genes known to code for physically-interacting proteins; transcriptomics datasets can indicate pairs of co-expressed genes; functional annotation classifiers (Gene Ontology, TIGR, COG) can be used to identify genes in a common

class. For all of these classifiers, the precision and accuracy of prediction of true positive pairs by measures of similarity in phenomic datasets is solid evidence for potential of functional discovery.

Data Exploration

Complete phenomic datasets are large and require a strategy to systematically explore and extract testable hypotheses of gene function. Fortunately, the data itself shares many similarities with microarray data (many genes, many conditions, a null expectation of zero change) and therefore many tools have already been developed. We present a brief overview of basic techniques; see (Gentleman, 2005) for a comprehensive review.

a) Specific condition-gene interactions

The baseline level of phenomic data analysis attempts to tie genes to cellular processes manually by extrapolating from specific phenotypes of the mutant. For example, the study by Donald et. al (Donald et al., 2009) discussed below in the *in-vitro* competitive assay section generated a hypothesis that an essential gene of unknown function is a peptidoglycan flippase, based on the hypersensitivity of its knockdown strain to a set of cell wall-targeting drugs. The association of the gene with a broad process (cell wall biosynthesis) was the starting point for further exploration.

b) Unsupervised methods

The second tier of phenomic exploration approaches used unsupervised computational methods to group mutant and condition profiles by similarity. “Unsupervised” refers to the fact that no prior knowledge about the functional relatedness of genes in the dataset is fed into the algorithm. Therefore, all relationships predicted by the clustering algorithm are based only on the phenomic data. The most commonly used unsupervised method is two-dimensional (2D) hierarchical clustering and heat map visualization. In 2D hierarchical clustering, rows and columns of an input

data matrix are iteratively joined based on the similarity of their profiles. Relationships are represented by the spatial grouping and bootstrapping of rows and columns in the output matrix.

c) Machine learning

A higher level of phenomic data exploration can be achieved through machine learning, or supervised methods. Contrary to unsupervised methods, supervised methods predict relationships in the data based on prior knowledge. For example, functional annotation of genomes (Gene Ontology, TIGR, COG) can be used to “train” algorithms on functionally-related genes. Then, patterns in the data that connect these known groups can be used to predict novel members.

Phenomic Applications in Bacteria

Phenomic technologies have been implemented to understand gene function in several different bacterial applications, ranging from cellular modeling of *E. coli* to identifying novel antibiotic resistance genes in complex environmental samples. Researchers have utilized forward and reverse genetics, selections and screens, and *in-vitro* and *in-vivo* approaches. Five major areas have been developed to date, and they are reviewed below.

***In-vitro* Competitive Assays**

In-vitro competitive assays are generally based on growth, but can also be used to select for differences in traits of interest (i.e. motility). Most published studies are based on transposon mutant libraries, but other genetic techniques like asRNA have also been integrated. Regardless of genetic technique or readout, competition is used to reveal conditional phenotypic differences amongst a collection of strains. These condition-

specific phenotypes have presented powerful leads for understanding gene function in model, pathogenic, and applied species.

Girgis *et. al* conducted a series of forward genetic positive selections on an *E. coli* transposon mutant library to identify genes involved in motility (Girgis et al., 2007). By plating a pre-selected pool of mutants in the center of a soft agar plate and allowing motile mutants to chemotax away from the resource-depleted center, the authors enriched for non-motile mutants. Serial passages and selections of this non-motile pool led to further enrichment, and eventually the identification of a suite of genes implicated in motility via a microarray-based readout of the insertion-flanking genomic DNAs. Importantly, this suite contained >95% of all previously known motility genes, as well as >30 genes not previously associated with the process. By searching for suppressors of the motility-impaired phenotype of the novel candidates, the authors were able to make important biological insights regarding the involvement of the second-messenger c-di-GMP system in chemotaxis signaling. In all, this study represents an elegant fusion of classic mutagenesis and selection techniques with a high-throughput readout, and demonstrates that phenomic approaches can reveal new insights into even the most-studied processes in the best-studied organisms.

Gallagher *et. al* executed a transposon-based forward genetic screen of *P. aeruginosa* to identify genes involved in the intrinsic resistance of this opportunistic pathogen to the aminoglycoside antibiotic tobramycin (Gallagher et al., 2010). The authors competed nearly 100,000 mutants in batch culture both with and without a low concentration of tobramycin, and quantified the abundance of all mutants in the pool with a high-throughput sequencing-based readout. Because *P. aeruginosa* is normally resistant to the drug, most mutants were expected to be unaffected. Those strains harboring

mutations in genes necessary for resistance were expected to decrease in abundance, or disappear from the population under tobramycin stress. This strategy proved effective, as many previously-described tobramycin resistance genes, as well as a group of novel candidates were captured. These new candidate genes collectively suggested important roles for both cell envelope and intracellular potassium homeostasis in regulating tobramycin resistance. This study illustrates the power of pooled growth assays and high-throughput sequencing to reveal important biology in clinically-relevant species.

Deutschbauer *et. al* conducted a signature-tagged mutagenesis (STM) based screen evaluating mutant fitness in a large battery of growth conditions (Deutschbauer et al., 2011). Almost 25,000 transposon insertion mutants of *S. oneidensis*, covering ~3500 genes, were assayed for fitness under 121 growth conditions. Quantifying strain abundance with a microarray-based readout, the authors built a matrix of fitness scores for each gene under each condition. Using this matrix to identify specific mutant phenotypes and genes whose mutants behaved similarly throughout the study, the authors generated new evidence-based predictions of function for 40 genes, including many involved in metabolism and three involved in motility. Further, characterization and archiving of the transposon mutant strains created a reverse genetic resource, allowing specific mutant phenotypes to be pursued via targeted experiments at a later time. This allowed several of the 40 predictions to be validated experimentally. As *S. oneidensis* is of potential use in the bioremediation of heavy metals and energy generation, this study demonstrates the potential of transposon-based phenomic analyses to generate functional insights into important processes of applied organisms.

To this point, asRNA screens in bacteria have focused on defining drug mode-of-action (MOA), and therefore have not yet been exploited for phenomic screens. However,

Donald *et. al* conducted a drug MOA study by screening a library of asRNA essential gene knockdown strains in *S. aureus* for sensitivity to a variety of known and unknown compounds (Donald et al., 2009). While most findings emphasized using the asRNA strain response patterns to predict drug MOA, a knockdown strain of an essential gene of unknown function (SAV1754) was found to be highly sensitive to a variety of cell-wall-targeting compounds. After additional experimentation, the authors suggest that SAV1754 may function as a peptidoglycan flippase, demonstrating the potential of asRNA-based screens to inform discovery of gene function.

***In-vitro* Arrayed Approaches**

Bacterial species for which individual mutants can be isolated are amenable to arrayed phenomic analyses. To this point, the limited availability of arrayed libraries and advanced robotics have made arrayed approaches much less common than pooled ones discussed above. However, exciting biology has been revealed by arrayed approaches using deletion libraries, overexpression libraries, and evolved strains. These tools have been used for large-scale screens for gene function, smaller-scale, targeted screens, and drug discovery efforts.

Nichols *et. al* executed a large-scale reverse genetic screen of the *E. coli* Keio Collection, a comprehensive single-gene knockout library (Nichols et al., 2011). Using the endpoint colony size as a readout, almost 4000 mutants were screened under 324 growth conditions. The large number of conditions screened and the sensitivity of the reverse genetic approach allowed many uncharacterized genes to be associated with well-studied functional modules through pairwise correlation analysis of knockout strains. In all, high-confidence functional predictions were generated for more than 300 previously uncharacterized genes. As a follow-up case, one uncharacterized gene that

correlated very highly to a gene involved in peptidoglycan synthesis was characterized to be essential for peptidoglycan synthase activity in-vivo using a series of genetic and biochemical experiments (Typas et al., 2010). In addition, the phenotypes identified in this study were used to generate insights into drug action and the evolution and genomic organization of *E. coli*.

Arrayed approaches have also been used for targeted questions. A study by Kohanski *et. al* used a liquid culture array of the Keio Collection. The authors used endpoint optical density as a readout of the growth of each mutant under gentamicin stress relative to control to identify gene mutants that sensitize the cell to the drug (Kohanski et al., 2008). Integration of the phenotypic data with gene expression data collected under the same conditions, enabled the authors to make new insights into the mechanism by which aminoglycoside antibiotics trigger cell death.

Arrayed approaches can also be useful for analyzing a small number of genetically distinct strains. Biolog technology involves arraying of a single strain into microplate format, with subsequent analysis of the ability of that strain to grow under a variety of metabolic conditions. Using a colorimetric readout of cellular respiration, the Biolog system enables rapid analysis of a small number of strains under a large number of growth conditions. A recent study by Fabich *et. al* examined the relative ability of a host-adapted strain of *E. coli* to catabolize a variety of carbon sources relative to its parent strain (Fabich et al., 2011). By screening the adapted strain, which contained a loss-of-function mutation in the *flhDC* locus, a key master regulator of motility in *E. coli*, the authors were able to correlate this mutation with increased capacity to proliferate under a variety of carbon-source conditions. After constructing an *flhDC* mutant in an isogenic background to the wild-type parent strain, the authors concluded that loss of *flhDC*

actually results in increased cellular metabolic capacity, and therefore inhibition of motility may represent an adaptive response for *E. coli* host colonization. These findings demonstrate the utility of Biolog technology for rapidly evaluating cellular fitness under varying metabolic conditions, and further show the ease with which the technology can be applied to naturally-occurring or evolved bacterial strains.

Overexpression libraries can also be arrayed for phenomic analyses. Pathania and colleagues assembled a set of *E. coli* strains (Kitagawa et al., 2005), each overexpressing one of the ~300 essential genes of *E. coli* K12 (Pathania et al., 2009). The strains were arrayed into microtiter plates, grown in liquid rich media, and challenged with 49 compounds that had been selected from a library of > 8,000 small molecules for growth inhibitory effects against the wild-type parent strain. Suppressors of the drug-induced growth inhibition were identified for 33 of the 49 compounds, pointing immediately to cellular targets of those compounds (growth-inhibitory compounds generally target essential gene products). In one case, a small molecule was identified that targeted LolA, a key protein in lipoprotein trafficking in gram-negative bacteria. Importantly, no other inhibitors of this protein or process were previously known, and therefore the chemical-genetic interaction between the compound and LolA represented a new phenotype for the *lolA* gene. Although this screen was designed for drug discovery, such chemical-genetic interactions can generate new functional information about the gene, even when its function is already known. Overexpression, or high-copy suppressor screens are quite powerful for drug discovery, but can also identify chemical-genetic relationships that open new avenues of experimentation regarding both drug action and gene function.

Natural Strain Approaches

Phenomic analyses have also been conducted on naturally-occurring bacterial isolates. While it is more difficult to associate observed phenotypes with specific genetic changes using such isolates, the increasing efficiency of genome sequencing has made the identification of genetic differences between isolates much more routine. As long as these differences are known, phenomic analyses can be extremely powerful.

A recent study by Franz *et. al* examined the relative ability of 18 strains of *E. coli* 0157 to proliferate in manure-infused soil (Franz et al., 2011). The authors were interested in the genetic basis of this phenotype, as survival in that environment is thought to be critical for strains ultimately found to contaminate agricultural crops. Amongst the 18 isolates, the authors observed a large variance in soil proliferation capacity. In parallel to the soil screen, Biolog technology was utilized to assess the metabolic capacity of the strains, and genotyping of known virulence genes was carried out. No correlation was found between the presence/absence of the virulence genes and the ability to proliferate in the soil. However, the authors observed that strains able to grow on several different acid substrates were likely to survive longer in the manure-infused soil than strains unable to grow in the acids. Therefore, it is possible that genes involved in metabolism underlie the soil survival phenotype, and sequencing of candidate loci (or of the entire genomes) may demonstrate this conclusively.

Another recent study by Kadali *et. al* utilized a selection followed by Biolog analysis to examine the ability of species able to grow on crude oil to remediate specific hydrocarbons (Kadali et al., 2012). First, a complex environmental sample from a former oil refinery site was harvested, and a selection was carried out on this sample using crude oil as a complex carbon source. A collection of strains were selected based on their ability to survive the selection, and then Biolog analysis was used to characterize

the ability of each selected strain to proliferate on specific hydrocarbons. The results represent significant progress in characterizing strains useful for the bioremediation of hydrocarbons, and mark the first step in drawing phenotypic connections between genes within the selected strains and the survival phenotype.

***In-vivo* Approaches**

In-vivo phenomic approaches based on transposon mutagenesis have been extremely powerful in studies of infection biology and the human microbiome. Both negative and positive selections have been utilized to identify genes involved in infection, virulence, and colonization. In a signature-tagged mutagenesis study of *S. typhimurium*, Hensel *et al.* used a negative selection to identify a pathogenicity island coding for a novel type three secretion system (T3SS) essential for intracellular replication (Hensel *et al.*, 1995). To identify virulence genes, the authors created a library of transposon mutants, with each mutant carrying a unique DNA barcode in the transposon. Mutants were arranged into 96-well plates, and a single pool was created from each plate. Each pool was sampled at the start point and then used to infect two BALB/c mice. After three days of infection, the mice were sacrificed, and spleens were homogenized and plated onto selective media to isolate the transposon mutants. Colonies from this plating were then scraped into a pool, which represented the “selected” population. A DNA-blotting method was used to identify the presence or absence of each DNA barcode in both the starting and selected pools. Mutants lost from the selected pool represented candidate virulence genes. Genomic DNA flanking the insertion sites in these mutants was cloned into a pUC vector and sequenced to identify the mutated genes. This study represented a major advance in *in-vivo* genetic selections, and has inspired an entire field of study over the last fifteen years (as indicated by nearly 1000 citations in PubMed). Over time, technological advances have made such approaches even more powerful by allowing

quantitative measurement of individual mutants within a pool. Rather than simply examining loss of a mutant from the pool, these more sensitive methods can assess decreased mutant abundance in the pool.

Conversely, it is also possible to employ positive selections to identify mutants that expand in frequency in the host, which could point to adaptive strategies used by the bacteria to establish and maintain infection/colonization. A recent study by Bianconi *et al.* utilized this approach to identify “patho-adaptive” mutations that promote chronic infection by *P. aeruginosa* in a mouse model of cystic fibrosis (Bianconi *et al.*, 2011). The authors used a previously-characterized *P. aeruginosa* STM mutant library, and pooled individual mutants to infect the mice. A total of 24 DNA barcode sequences from three unique vector backbones ($24 \times 3 = 72$ unique transposon sequences) were present in each pool. Following infection, mice were observed and scored for establishment of a chronic infection. Those animals in which a chronic infection was established were classified as positives, sacrificed, and lungs were homogenized and plated on selective media to count the CFU and screen for STM mutants. In this study, a previously-developed multiplex PCR strategy (Potvin *et al.*, 2003a) was used to track the abundance of each unique mutant within the pools. Mutants that expanded in the population were flagged as candidates, and re-screened in subsequent rounds. Eventually, 16 mutants were identified that significantly increased chronic infection relative to wild-type. These mutants impacted a variety of processes, suggesting that “patho-adaptive” mutations promoting chronic infection may decrease motility, alter biofilm formation, or decrease secretion of virulence factors. To extend their analysis of patho-adaptive mutations, the authors sequenced ORFs of the 16 candidate genes in a cohort of early and late CF patient samples. Interestingly, an independent clinical study found mutations in 7/16 of the candidate genes. The convergence of STM and clinical

sequence results are strong evidence that the authors have indeed identified “patho-adaptive” strategies used by *P. aeruginosa* to establish and maintain chronic infection in CF airways.

In-vivo screens of transposon mutant populations can also focus on understanding the molecular determinants of establishing the symbiotic relationships between the host and its microbiome. A study by Goodman *et. al* utilized a library of *B. thetaiotaomicron* transposon mutant strains and germ-free mice to search for candidate genes involved in colonization of the mouse gut (Goodman et al., 2009). Importantly, the authors utilized short-read sequencing technology and an engineered Mariner transposon to establish an insertion sequencing approach. Such approaches are based on a single transposon without STM. Instead, the genomic DNA flanking the insertion site is used to map and quantify the relative abundance of transposon mutants. This approach allowed the authors to carry out a forward genetic screen by assigning a quantitative measure of abundance to each unique insertion strain in the pool before and after various selections. To search for loci specifically required for *in-vivo* colonization, the authors compared all *in-vivo* strain fitness measures with a set of *in-vitro* fitness measurements calculated from the results of a competition experiment conducted in rich media. More clearly, they used an *in-vitro* competition to identify loci that affect growth generally, and therefore were able to filter for strains exhibiting fitness defects specifically *in-vivo*. After conducting a series of *in-vivo* screens in varying bacterial community contexts, the authors discovered a genomic locus containing several genes involved in vitamin B12 utilization that, when mutated, led to a significant fitness defect only when vitamin B12 was presumed to be of limited availability in the mouse gut. As vitamin B12 is essential for the mouse and is synthesized exclusively by members of the microbiome, these genes are likely to be necessary for normal colonization *in-vivo*. This finding

demonstrates the enormous potential of *in-vivo* phenomic approaches to understand symbiotic relationships between bacteria and their hosts; a topic of broad appeal that has been relatively inaccessible to investigation until now.

Functional Metagenomics

A critical and emerging area of microbiology is the study of complex microbial communities via metagenomic sequencing. Initial studies in this area have focused on surveying the species composition of varying environmental communities. To date, such projects have deposited enormous amounts of genome sequences into databases, many of which code for ORFs of both unknown function and unknown species origin. Such sequences are of extremely limited utility, and therefore approaches aimed at assigning function to them are in high demand.

A recent study by McGarvey *et. al* utilized a heterologous expression system to select for metagenomic sequences conferring antibiotic resistance to *E. coli* (McGarvey et al., 2012). Briefly, the authors cloned a metagenomic DNA library into a pUC18 vector, yielding 1.4 million clones covering 2.8×10^9 bases of DNA. By transforming the plasmid library into *E. coli* and simply plating on a variety of antibiotics, the authors were able to search for metagenomic sequences that conferred antibiotic resistance phenotypes on the *E. coli* host. In all, the authors identified 39 antibiotic-resistance genes, of which all but one coded for members of previously-defined drug-resistance protein families. However, interestingly, many of the newly-discovered members shared little nucleotide sequence homology to the previously-described family members. For example, the authors discovered one type II DHFR (conferring resistance to trimethoprim) having a maximum of 24% amino acid sequence identity to other known type II DHFR's. Evolutionary analysis of normally-conserved amino acid residues indicates that the

newly discovered DHFR is evolutionarily distant. These types of discoveries are incredibly valuable for further defining the bacterial resistome, allowing for a better understanding of drug-protein structure-function relationships, and potentially informing future drug design efforts.

An earlier study by Sommer *et. al* utilized a slightly different strategy to screen four different soil microbiomes for genetic elements that conferred resistance to growth-inhibitory compounds in *E. coli* (Sommer et al., 2010). Similar to the McGarvey study, heterologous expression of the metagenomic DNA was carried out in *E. coli*, and selective pressure consisted of high doses of industrially-relevant chemicals (organic acids, alcohols, and aldehydes). This study cloned metagenomic DNA into fosmid libraries, containing 40-50kb of DNA rather than the 1-3kb inserts in the pUC18-based approach of McGarvey. Advantages of this approach are that fewer clones need to be screened, and phenotypes based on >1 neighboring genes can be identified. Of course, in this approach followup is necessary to determine which gene(s) are responsible for a phenotype of interest. In the Sommer study, identification of candidate genes was obtained by a second transposon mutagenesis on the resistant fosmid clones, followed by re-screening for loss of the phenotype. This approach resulted in discovery of novel genes conferring tolerance to biomass-conversion-related growth inhibitory compounds. Importantly, it offers another example of a phenomic strategy to identify meaningful functional genetic elements from complex (and perhaps unculturable) microbial sources.

References

- Amini, S., Goodarzi, H., and Tavazoie, S. (2009). Genetic Dissection of an Exogenously Induced Biofilm in Laboratory and Clinical Isolates of *E. coli*. *PLoS Pathogens* 5, e1000432.
- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2, 2006 0008.
- Bianconi, I., Milani, A., Cigana, C., Paroni, M., Levesque, R.C., Bertoni, G., and Bragonzi, A. (2011). Positive Signature-Tagged Mutagenesis in *Pseudomonas aeruginosa*: Tracking Patho-Adaptive Mutations Promoting Airways Chronic Infection. *PLoS Pathogens* 7, e1001270.
- Bochner, B.R. (2009). Global phenotypic characterization of bacteria. *FEMS microbiology reviews* 33, 191-205.
- Camacho, L.R., Ensergueix, D., Perez, E., Gicquel, B., and Guilhot, C. (1999). Identification of a virulence gene cluster of *Mycobacterium tuberculosis* by signature-tagged transposon mutagenesis. *Molecular Microbiology* 34, 257-267.
- Cameron, D.E., Urbach, J.M., and Mekalanos, J.J. (2008). A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* 105, 8736.
- Collins, S.R., Miller, K.M., Maas, N.L., Roguev, A., Fillingham, J., Chu, C.S., Schuldiner, M., Gebbia, M., Recht, J., Shales, M., *et al.* (2007). Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature* 446, 806-810.
- Couce, A., Briaies, A., Rodriguez-Rojas, A., Costas, C., Pascual, A., and Blazquez, J. (2012). Genome-wide overexpression screen for fosfomycin resistance in *Escherichia coli*: MurA confers clinical resistance at low fitness cost. *Antimicrob Agents Chemother.*
- de Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., Lechaplais, C., Gyapay, G., Richez, C., *et al.* (2008). A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Mol Syst Biol* 4, 174.
- Deutschbauer, A., Price, M.N., Wetmore, K.M., Shao, W., Baumohl, J.K., Xu, Z., Nguyen, M., Tamse, R., Davis, R.W., and Arkin, A.P. (2011). Evidence-Based Annotation of Gene Function in *Shewanella oneidensis* MR-1 Using Genome-Wide Fitness Profiling across 121 Conditions. *PLoS Genetics* 7, e1002385.
- Donald, R.G.K., Skwish, S., Forsyth, R.A., Anderson, J.W., Zhong, T., Burns, C., Lee, S., Meng, X., LoCastro, L., Jarantow, L.W., *et al.* (2009). A *Staphylococcus aureus* Fitness Test Platform for Mechanism-Based Profiling of Antibacterial Compounds. *Chemistry & Biology* 16, 826-836.
- Fabich, A.J., Leatham, M.P., Grissom, J.E., Wiley, G., Lai, H., Najjar, F., Roe, B.A., Cohen, P.S., and Conway, T. (2011). Genotype and Phenotypes of an Intestine-Adapted *Escherichia coli* K-12 Mutant Selected by Animal Passage for Superior Colonization. *Infection and Immunity* 79, 2430-2439.
- Flores, A.R., Parsons, L.M., and Pavelka, M.S. (2005). Characterization of Novel *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* Mutants Hypersusceptible to β -Lactam Antibiotics. *Journal of Bacteriology* 187, 1892-1900.
- Franz, E., van Hoek, A.H.A.M., Bouw, E., and Aarts, H.J.M. (2011). Variability of *Escherichia coli* O157 Strain Survival in Manure-Amended Soil in Relation to Strain Origin, Virulence Profile, and Carbon Nutrition Profile. *Applied and Environmental Microbiology* 77, 8088-8096.
- Gallagher, L.A., Ramage, E., Jacobs, M.A., Kaul, R., Brittnacher, M., and Manoil, C. (2007). A comprehensive transposon mutant library of *Francisella novicida*, a bioweapon surrogate. *Proc Natl Acad Sci U S A* 104, 1009-1014.
- Gallagher, L.A., Shendure, J., and Manoil, C. (2010). Genome-Scale Identification of Resistance Functions in *Pseudomonas aeruginosa* Using Tn-seq. *mBio* 2, e00315-00310-e00315-00310.
- Gawronski, J.D., Wong, S.M.S., Giannoukos, G., Ward, D.V., and Akerley, B.J. (2009). Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for *Haemophilus* genes required in the lung. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16422-16427.
- Gentleman, R. (2005). *Bioinformatics and computational biology solutions using R and Bioconductor* (New York: Springer Science+Business Media).
- Girgis, H.S., Hottes, A.K., and Tavazoie, S. (2009). Genetic architecture of intrinsic antibiotic susceptibility. *PLoS ONE* 4, e5629.

Girgis, H.S., Liu, Y., Ryu, W.S., and Tavazoie, S. (2007). A Comprehensive Genetic Characterization of Bacterial Motility. *PLoS Genetics* 3, e154.

Gomez, M.J., and Neyfakh, A.A. (2006). Genes Involved in Intrinsic Antibiotic Resistance of *Acinetobacter baylyi*. *Antimicrobial Agents and Chemotherapy* 50, 3562-3567.

Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. *Cell Host and Microbe* 6, 279-289.

Hensel, M., Shea, J.E., Gleeson, C., Jones, M.D., Dalton, E., and Holden, D.W. (1995). Simultaneous identification of bacterial virulence genes by negative selection. *Science* 269, 400-403.

Hisert, K.B., Kirksey, M.A., Gomez, J.E., Sousa, A.O., Cox, J.S., Jacobs, W.R., Nathan, C.F., and McKinney, J.D. (2004). Identification of Mycobacterium tuberculosis Counterimmune (cim) Mutants in Immunodeficient Mice by Differential Screening. *Infection and Immunity* 72, 5315-5321.

Huber, J., Donald, R.G.K., Lee, S.H., Jarantow, L.W., Salvatore, M.J., Meng, X., Painter, R., Onishi, R.H., Occi, J., Dorso, K., *et al.* (2009). Chemical Genetic Identification of Peptidoglycan Inhibitors Potentiating Carbapenem Activity against Methicillin-Resistant *Staphylococcus aureus*. *Chemistry & Biology* 16, 837-848.

Ishii, N., Nakahigashi, K., Baba, T., Robert, M., Soga, T., Kanai, A., Hirasawa, T., Naba, M., Hirai, K., Hoque, A., *et al.* (2007). Multiple High-Throughput Analyses Monitor the Response of *E. coli* to Perturbations. *Science* 316, 593-597.

Kadali, K.K., Simons, K.L., Skuza, P.P., Moore, R.B., and Ball, A.S. (2012). A complementary approach to identifying and assessing the remediation potential of hydrocarbonoclastic bacteria. *Journal of Microbiological Methods*, 1-8.

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005). Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA research : an international journal for rapid publication of reports on genes and genomes* 12, 291-299.

Kohanski, M.A., Dwyer, D.J., Wierzbowski, J., Cottarel, G., and Collins, J.J. (2008). Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell* 135, 679-690.

Langridge, G.C., Phan, M.D., Turner, D.J., Perkins, T.T., Parts, L., Haase, J., Charles, I., Maskell, D.J., Peters, S.E., Dougan, G., *et al.* (2009). Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Research* 19, 2308-2316.

Lawley, T.D., Chan, K., Thompson, L.J., Kim, C.C., Govoni, G.R., and Monack, D.M. (2006). Genome-Wide Screen for *Salmonella* Genes Required for Long-Term Systemic Infection of the Mouse. *PLoS Pathogens* 2, e11.

Liberati, N.T., Urbach, J.M., Miyata, S., Lee, D.G., Drenkard, E., Wu, G., Villanueva, J., Wei, T., and Ausubel, F.M. (2006). An ordered, nonredundant library of *Pseudomonas aeruginosa* strain PA14 transposon insertion mutants. *Proc Natl Acad Sci U S A* 103, 2833-2838.

Liu, A., Tran, L., Becket, E., Lee, K., Chinn, L., Park, E., Tran, K., and Miller, J.H. (2010). Antibiotic Sensitivity Profiles Determined with an *Escherichia coli* Gene Knockout Collection: Generating an Antibiotic Bar Code. *Antimicrobial Agents and Chemotherapy* 54, 1393-1403.

Maynard, N.D., Birch, E.W., Sanghvi, J.C., Chen, L., Gutschow, M.V., and Covert, M.W. (2010). A forward-genetic screen and dynamic analysis of lambda phage host-dependencies reveals an extensive interaction network and a new anti-viral strategy. *PLoS Genet* 6, e1001017.

Mazurkiewicz, P., Tang, C.M., Boone, C., and Holden, D.W. (2006). Signature-tagged mutagenesis: barcoding mutants for genome-wide screens. *Nature Reviews Genetics* 7, 929-939.

McGarvey, K.M., Queitsch, K., and Fields, S. (2012). Wide Variation in Antibiotic Resistance Proteins Identified by Functional Metagenomic Screening of a Soil DNA Library. *Applied and Environmental Microbiology* 78, 1708-1714.

Meng, J., Kanzaki, G., Meas, D., Lam, C.K., Crummer, H., Tain, J., and Xu, H.H. (2012). A genome-wide inducible phenotypic screen identifies antisense RNA constructs silencing *Escherichia coli* essential genes. *FEMS Microbiology Letters*, n/a-n/a.

Nakahigashi, K., Toya, Y., Ishii, N., Soga, T., Hasegawa, M., Watanabe, H., Takai, Y., Honma, M., Mori, H., and Tomita, M. (2009). Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Mol Syst Biol* 5, 306.

Nguyen, B., and Valdivia, R. (2012). Virulence determinants in the obligate intracellular pathogen *Chlamydia trachomatis* revealed by forward genetic approaches. In Proceedings of the National

Nichols, R.J., Sen, S., Choo, Y.J., Beltrao, P., Zietek, M., Chaba, R., Lee, S., Kazmierczak, K.M., Lee, K.J., Wong, A., *et al.* (2011). Phenotypic landscape of a bacterial cell. *Cell* 144, 143-156.

Pathania, R., Zlitni, S., Barker, C., Das, R., Gerritsma, D.A., Lebert, J., Awuah, E., Melacini, G., Capretta, F.A., and Brown, E.D. (2009). Chemical genomics in *Escherichia coli* identifies an inhibitor of bacterial lipoprotein targeting. *Nature Chemical Biology* 5, 849-856.

Phillips, J.W., Goetz, M.A., Smith, S.K., Zink, D.L., Polishook, J., Onishi, R., Salowe, S., Wiltsie, J., Allocco, J., Sigmund, J., *et al.* (2011). Discovery of Kibdelomycin, A Potent New Class of Bacterial Type II Topoisomerase Inhibitor by Chemical-Genetic Profiling in *Staphylococcus aureus*. *Chemistry & Biology* 18, 955-965.

Potvin, E., Lehoux, D.E., Kukavica-Ibrulj, I., Richard, K.L., Sanschagrin, F., Lau, G.W., and Levesque, R.C. (2003a). In vivo functional genomics of *Pseudomonas aeruginosa* for high-throughput screening of new virulence factors and antibacterial targets. *Environ Microbiol* 5, 1294-1308.

Potvin, E., Lehoux, D.E., Kukavica-Ibrulj, I., Richard, K.L., Sanschagrin, F.o., Lau, G.W., and Levesque, R.C. (2003b). In vivo functional genomics of *Pseudomonas aeruginosa* for high-throughput screening of new virulence factors and antibacterial targets. *Environmental Microbiology* 5, 1294-1308.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J., *et al.* (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* 322, 405-410.

Santiviago, C.A., Reynolds, M.M., Porwollik, S., Choi, S.-H., Long, F., Andrews-Polymenis, H.L., and McClelland, M. (2009a). Analysis of Pools of Targeted *Salmonella* Deletion Mutants Identifies Novel Genes Affecting Fitness during Competitive Infection in Mice. *PLoS Pathogens* 5, e1000477.

Santiviago, C.A., Reynolds, M.M., Porwollik, S., Choi, S.H., Long, F., Andrews-Polymenis, H.L., and McClelland, M. (2009b). Analysis of pools of targeted *Salmonella* deletion mutants identifies novel genes affecting fitness during competitive infection in mice. *PLoS Pathog* 5, e1000477.

Schuldiner, M., Collins, S.R., Weissman, J.S., and Krogan, N.J. (2006). Quantitative genetic analysis in *Saccharomyces cerevisiae* using epistatic miniarray profiles (E-MAPs) and its application to chromatin functions. *Methods* 40, 344-352.

Sommer, M.O., Church, G.M., and Dantas, G. (2010). A functional metagenomic approach for expanding the synthetic biology toolbox for biomass conversion. *Molecular Systems Biology* 6, 1-7.

Tamae, C., Liu, A., Kim, K., Sitz, D., Hong, J., Becket, E., Bui, A., Solaimani, P., Tran, K.P., Yang, H., *et al.* (2008). Determination of Antibiotic Hypersensitivity among 4,000 Single-Gene-Knockout Mutants of *Escherichia coli*. *Journal of Bacteriology* 190, 5981-5988.

Tenaillon, O., Rodriguez-Verdugo, A., Gaut, R.L., McDonald, P., Bennett, A.F., Long, A.D., and Gaut, B.S. (2012). The molecular diversity of adaptive convergence. *Science* 335, 457-461.

Typas, A., Banzhaf, M., van den Berg van Saparoea, B., Verheul, J., Biboy, J., Nichols, R.J., Zietek, M., Beilharz, K., Kannenberg, K., von Rechenberg, M., *et al.* (2010). Regulation of peptidoglycan synthesis by outer-membrane proteins. *Cell* 143, 1097-1109.

Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L., *et al.* (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nature methods* 5, 781-787.

van Opijnen, T., Bodi, K.L., and Camilli, A. (2009). Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nat Methods* 6, 767-772.

Warner, J.R., Reeder, P.J., Karimpour-Fard, A., Woodruff, L.B.A., and Gill, R.T. (2010). Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nature Biotechnology* 28, 856-862.

Chapter 2: Phenotypic Landscape of a Bacterial Cell

This chapter is a reproduction of a Resource Article originally published in Cell in the January 7, 2011 issue. I contributed to conceiving the study, designing the research, and performing the screen. I performed all data processing independently. I also contributed to analyzing the processed data, and performed several of the follow-up experiments. I was the primary author of the manuscript, and the main creator of all figures. The relative contributions of all authors are indicated on the following page, which is the original title page of the Cell manuscript.

Phenotypic Landscape of a Bacterial Cell

Robert J. Nichols^{1,2}, Saunak Sen³, Yoe Jin Choo², Pedro Beltrao⁴, Matylda Zietek², Rachna Chaba², Sueyoung Lee², Krystyna M. Kazmierczak⁵, Karis J. Lee^{2¶}, Angela Wong^{2§}, Michael Shales⁴, Susan Lovett⁶, Malcolm E. Winkler⁵, Nevan J. Krogan⁴, Athanasios Typas^{2*} and Carol A. Gross^{2,7*}

¹Oral and Craniofacial Sciences *Graduate Program*, University of California at San Francisco, 513 Parnassus Ave, CA-94143

²Department of Microbiology and Immunology, University of California at San Francisco, San Francisco, 600 16th Street, CA-94158

³Department of Epidemiology and Biostatistics, University of California, San Francisco, CA-94107

⁴Department of Cellular and Molecular Pharmacology & The California Institute for Quantitative Biomedical Research, University of California at San Francisco, San Francisco, 1700 4th Street, CA-94158

⁵Department of Biology, Indiana University, Bloomington, Indiana 47405

⁶Department of Biology and Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, Massachusetts 02454-9110

⁷Department of Cell and Tissue Biology, University of California at San Francisco, San Francisco, California 94158

[¶]new address: UCLA School of Dentistry, 10833 Le Conte Avenue, Los Angeles, CA 90095-1762

[§]new address: Massachusetts General Hospital, 50 Staniford Street, Boston, MA 02114

*Corresponding authors:

Athanasios Typas: Phone: 415-476-1493; Fax: 415-514-4080; email:

athanasios.typas@ucsf.edu

Carol A. Gross: Phone: 415-476-1461; Fax: 415-514-4080; email:

cgrossucsf@gmail.com

Running title: Phenotypic Landscape of a Bacterial Cell

Contributions

AT, RJN, NJK and CAG conceived this study; AT and RJN designed research; AT, RJN, YJC, SL, JL and AW performed the screen; RJN did the data processing; RJN, SS, PB and AT analyzed the data; AT, RJN, CAG and SL interpreted the data; RJN, AT and MZ performed the follow up work in the anti-folate drug story, except for the work in *Streptococcus pneumoniae* which was conducted by KMK; RJN, RC, AT and MZ performed the follow up work in the *mar* story; RJN, AT and CAG wrote the manuscript; RJN and MS prepared the figures; PB, RC, MEW, KMK, SL, SS and NJK edited the manuscript; AT and CAG supervised all aspects of this project.

Summary

The explosion of sequence information on bacteria makes developing high-throughput, cost-effective phenotyping approaches imperative. Using *E. coli* as proof-of-principle, we show that combining large-scale chemical-genomics with quantitative fitness measurements provides a high quality dataset, rich in discovery. The identification of >10,000 growth phenotypes allowed us to study gene essentiality, discover leads for gene function and drug action, and understand higher-order organization of the bacterial chromosome. We highlight insights concerning a gene involved in multiple antibiotic resistance and provide information about synergy of a broadly used combinatory antibiotic therapy, trimethoprim and sulfonamides. This dataset, publicly available at <http://ecoliwiki.net/tools/chemgen/>, is a valuable resource for both the microbiological and bioinformatic communities as it provides high confidence associations between hundreds of annotated and uncharacterized genes as well as inferences about the mode-of-action of several poorly understood drugs. This approach can be used for all culturable microbes with available ordered mutant libraries.

Introduction

Before the physical basis of genes was understood, associating phenotypes with a heritable unit laid the foundation of modern genetics. Following discovery of the genetic code, linking a phenotype to the responsible gene remained the most expeditious way to unravel gene function. With the explosion of sequence information, the balance has shifted. We now have many genes of unknown function. To capitalize on the burgeoning sequence bank, it is imperative to develop high-throughput technologies that link genes to phenotypes and expedite discovery of gene function. This is particularly true for prokaryotes, which represent a major fraction of the sequenced genomes and are in the forefront of metagenomic efforts (Qin et al., 2010).

Chemical and environmental perturbations have traditionally linked phenotypes to genotypes through forward genetic screens, but reverse genetic approaches are being increasingly utilized (Barker et al., 2010). Phenotype microarrays utilize a high-resolution readout of cellular respiration to evaluate fitness of a strain in hundreds of conditions (Bochner, 2009). This approach is appropriate for studying a few strains, but is difficult to expand to genome-scale screens. In pooled growth competitions, thousands of strains are assayed in a single culture environment. Fitness values are derived from measuring strain abundance in a test relative to control condition (Giaever et al., 2004; Girgis et al., 2009; Hillenmeyer et al., 2008; Hoon et al., 2008; Lee et al., 2005; Pan et al., 2004; Parsons et al., 2006; Warner et al., 2010; Xu et al., 2007). These approaches are very efficient, but competition between strains in each condition makes it difficult to determine relative strain growth across conditions (Girgis et al., 2009), especially for strains that grow slowly even in the absence of perturbation (Lee et al., 2005). Arraying mutant strains on solid media allows independent evaluation of strain fitness, but has been used only for low-resolution measurements of entire libraries (Liu et al., 2010; Tamae et al.,

2008) or for essential genes (Pathania et al., 2009). High-throughput genetic interaction studies, pioneered in yeast (Schuldiner et al., 2005; Tong et al., 2001), are complementary to chemical genomics approaches. Such analyses *quantitatively* measure colony growth of double mutants in high-density format on agar surfaces, and have led to numerous successes in identifying gene function and network organization (Beltrao et al., 2010). Similar methodology has been developed for *E. coli* (Butland et al., 2008; Typas et al., 2008).

We use *E. coli* to illustrate the power of applying the high-resolution quantitative fitness measurements of genetic interaction analysis to high-throughput phenotypic analysis of culturable microbes. “Phenomic profiling” provides a quantitative description of the response of all single gene deletions to physiologically relevant stresses and drug challenges. By profiling ~4000 genes in >300 perturbations, we identified thousands of phenotypes and a diverse suite of conditionally essential genes. This approach provides new insights into the chromosome organization, functional landscape and evolutionary trajectory of *E. coli*. It facilitates high confidence association of genes of unknown function to those of known function, as highlighted by discovery of the role of a gene involved in multiple antibiotic resistance in this manuscript and identification of two novel lipoproteins essential for peptidoglycan synthesis (Typas et al. Cell, accepted). Finally, the degree to which various gene deletions alter toxic drug effects has led to powerful insights regarding drug mode-of-action (Kohanski et al., 2008) and we demonstrate that our analysis generates numerous leads concerning drug function.

Results and Discussion

Phenomic Profiling of *E. coli* K12 yields a robust, high-quality dataset

We determined quantitative growth scores for the Keio single-gene deletion library (Baba et al., 2006); essential gene hypomorphs [C-terminally tandem-affinity tagged (Butland et al., 2005) or specific alleles]; and a small RNA/small protein knockout library (Hobbs et al., 2010) in conditions representing the range of stresses *E. coli* encounters. Mutant strains arrayed in high-density on agar plates (1536 colonies/plate) were grown in 324 conditions covering 114 unique stresses (Figs. 1A & S1A, Table S1). Colony sizes were analyzed and converted to a drug-gene score using an approach developed for quantifying genetic interactions (see Experimental Procedures). More than half of the conditions were antibiotic/antimicrobial treatments (Fig. 1A). By using a sub-inhibitory concentration series that maximally inhibited growth of the wildtype (wt) strain $\leq 50\%$, we were able to search for specific drug-gene interactions (Fig. S1A), and reduce the ability of spontaneous suppressor mutations to overtake the colony. Two independently derived clones of each mutant strain were analyzed (for sRNA mutants, a single isolate was arrayed twice) and screens were performed at least twice, enabling scores to be based on 4-6 independent measurements. Correlation between replicate colony size measurements was very high ($r=0.77$, Fig. 1B). The final dataset (Table S2) was comprised of scores for the 3979 mutant strains passing quality control (e.g. proper normalized colony size distribution and replicate reproducibility; see Experimental Procedures). The entire dataset is available in an interactive, searchable format and as a flat file on the *E. coli* wiki website (at <http://ecoliwiki.net/tools/chemgen/>).

The entire matrix (3979 mutants X 324 conditions) was subjected to 2-D hierarchical clustering (Fig. 1C). Drugs with similar effects cluster on the X-axis; mutants that

behaved similarly cluster on the Y-axis. Notably, concentrations of the same drug, drugs of the same family and/or similar conditions clustered together as did mutants of genes known to be part of the same operon, biological pathway and/or protein complex.

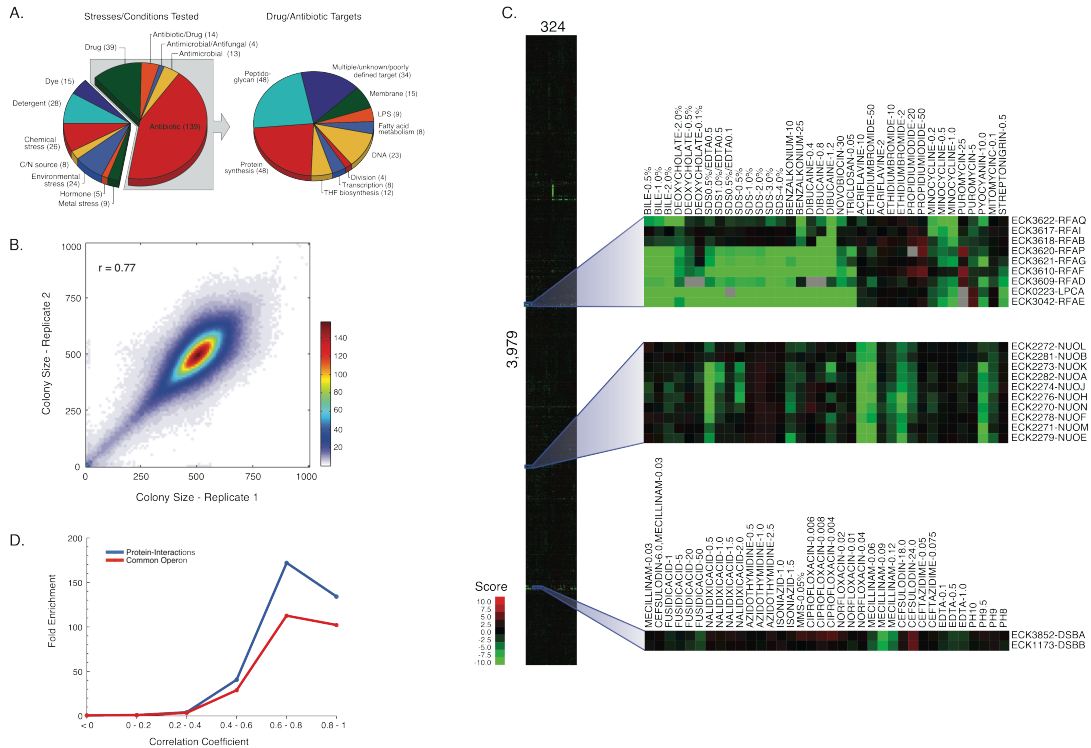


Figure 1. Phenomic Profiling of the enhanced Keio Collection yields a robust and rich dataset. (A) Classification of the 324 stresses screened (left), and cellular targets of the antibiotic/antimicrobial/drug classes (right). (B) Heat map representation of scatter plot comparing normalized colony sizes in pixels of plate replicates 1 and 2 across the entire dataset. Bins indicate the square root of the number of replicate pairs within a 10 x 10 pixel window as depicted by color scale. Note that the vast majority of the replicates have highly correlated colony sizes. (C) Clustergram of fitness scores for 3979 mutant strains in response to all 324 conditions. Zoomed insets demonstrate co-clustering of conditions (x-axis) and genes (y-axis) for a common pathway (*rfa* cluster), and protein complexes encoded in the same operon (*nuo*) or in different operons (*dsbA* and *dsbB*). Gray boxes indicate missing data. (D) High correlation between a pair of phenotypic signatures is predictive of shared protein interaction and/or operon membership.

Zoomed insets of our clustergram illustrate examples. Genes in the *rfa* operon (*rfaG*, *rfaP*, *rfaQ*, *rfaB* and *rfaI*), which encodes enzymes that synthesize the inner and outer lipopolysaccharide (LPS) core strongly cluster together with 3/4 genes responsible for the synthesis of one of the sugar building blocks, ADP-L glycerol- β -D-manno heptose

(*rfaD*, *rfaE*, *lpcA*). Importantly, clustering reflects their shared sensitivities to a concentration series of compounds known to perturb the envelope integrity of the cell, consistent with the role of LPS. *dsbA* and *dsbB*, encoded in different operons, also cluster. The DsbA/DsbB complex generates disulfide bridges in the periplasm.

The response of each mutant strain across all conditions is denoted as its “phenotypic signature”. High correlation between two phenotypic signatures is highly predictive of known indicators of functional connection between genes. Gene pairs with correlation coefficients (r) between 0.6 and 0.8 (p -value $< 10^{-34}$) are more than 100-fold enriched for genes sharing common operon membership, and 150-fold enriched for genes with known protein interactions determined from low-throughput experiments (www.ecocyc.org, Fig. 1D). This benchmarking analysis indicates that our phenomics dataset is biologically meaningful. Correlated phenotypic signatures also reproduce connections between curated biological pathways (Fig. S1B). For example, electron transfer components cluster tightly (e.g. *nuo* genes encoding NADH dehydrogenase I complex; Fig. 1C). Their clustering reflects high sensitivity to membrane-perturbing stresses including detergents, dyes and metals, and increased resistance to aminoglycosides, in agreement with early studies that illustrate decreased aminoglycoside uptake in the absence of a fully functional electron transport chain (Girgis et al., 2009; Taber et al., 1987). All three examples described in Fig. 1C are consistent with the expectation that highly correlated phenotypic signatures are biologically meaningful ($r \geq 0.6$ -0.8).

Phenomic profiling defines Responsive and Conditionally-Essential Genes

A central goal of this study was to systematically evaluate the impact of every gene deletion on *E. coli* fitness in diverse environments, as few gene deletions in *E. coli* have

robust reported growth phenotypes and only 8% of the genes are essential in rich media (Baba et al., 2006; Yamamoto et al., 2009). We used a statistical method to define a reliable phenotype. Briefly, we standardized the interquartile range of the distribution of scores for each screen and then determined the probability that each condition-gene interaction represented a true phenotype using a normal cumulative distribution function (see Experimental Procedures). Using a 5% probability that these phenotypes arose by chance as a cut-off (false discovery rate (FDR) $\leq 5\%$), 49% of all strains tested (1957/3979 strains; Fig. 2A) had one or more phenotypes. We refer to these genes as the “responsive genome”. This “responsive genome” is a work in progress, as it is limited to genes whose removal causes *growth* phenotypes in response to the stresses tested. Expanding the stresses tested and/or the readout (e.g. motility) will certainly increase this number (Girgis et al., 2007). A cumulative plot of the number of individual phenotypes per strain shows that very few genes have many phenotypes. Multi-Stress Responsive (MSR) strains (≥ 30 phenotypes; Table S3) participate in many cellular processes, suggesting that our stresses encompassed diverse cellular challenges (Fig. 2B). With a stringent cut-off of 5% FDR, the maximum number of phenotypes from a single screen was 173 ($\sim 4\%$ strains; Fig. S2A), and the total number of phenotypes (13497) represent $\sim 1\%$ of all condition-gene pairs tested. Overall, 80% of the phenotypes were negative (gene deletion more sensitive) and 20% positive (gene deletion more resistant), consistent with recent genetic interaction analyses in *S. cerevisiae* (Fiedler et al., 2009) and *S. pombe* (Roguev et al., 2008). This suggests that removal of a gene product is more likely to decrease than enhance resistance to stress (Fig. S2B). In summary, our analysis captured numerous highly specific condition-gene responses. Clearly, this dataset can be used to assign more phenotypes at a lower confidence level. Indeed, a recent chemical genomics dataset in *S. cerevisiae* reported

phenotypes for more than 95% of gene deletions tested, many stemming from a handful of severe stresses (Hillenmeyer et al., 2008).

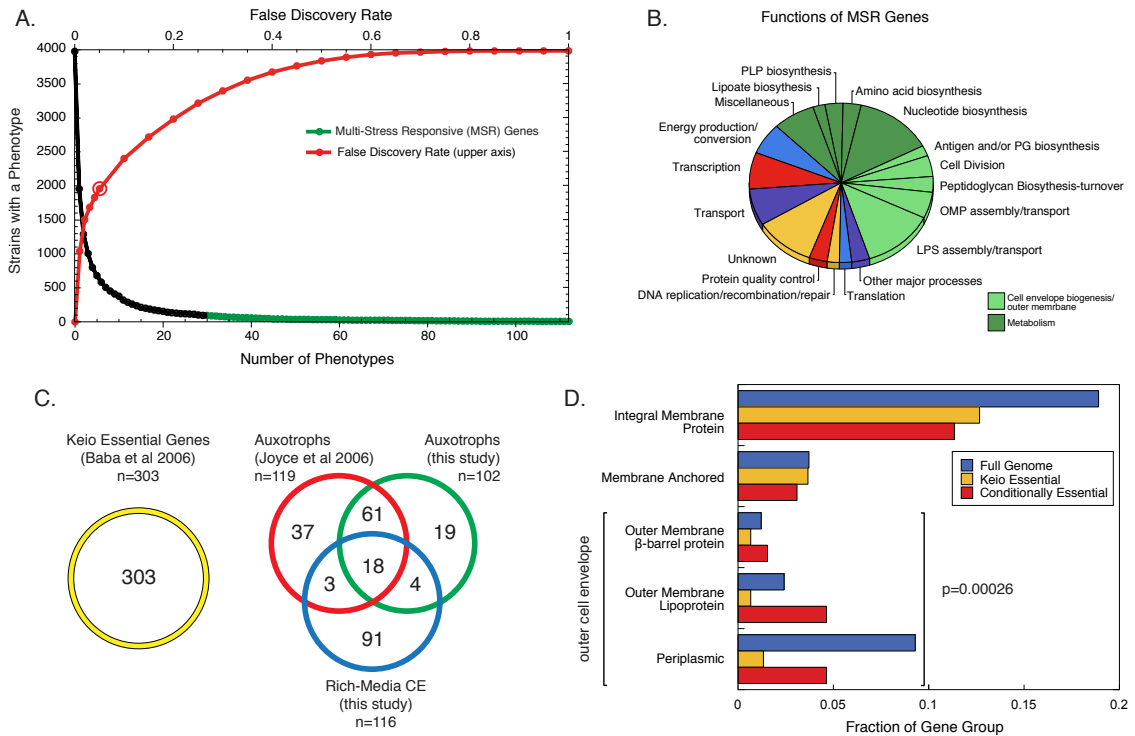


Figure 2. Identification of responsive and conditionally-essential genes. (A) Using a 5% false discovery rate (FDR), 49% of strains tested had at least one phenotype (open circle on the red line). As the FDR is relaxed, more phenotypes are identified (red line). At 5% FDR, some strains have several phenotypes (black) and very few (0.023%; 94 strains) have 30 or more phenotypes (green, Multi-Stress Responsive (MSR) genes). **(B)** MSR genes participate in a wide variety of cellular processes, particularly those related to metabolism and the cell envelope. Genes were manually curated to COG-based functions; each gene was allowed to belong only to a single function. **(C)** 196 genes are conditionally-essential (CE) in this study. Of these, roughly half have been previously described as CE due to auxotrophy. Note that some auxotrophic genes also display a no growth phenotype in at least one rich medium condition and are classified jointly as auxotroph and rich media CE. **(D)** Rich media CE gene products are enriched in the outer cell envelope (periplasm and outer membrane) relative to Keio essential genes ($p=0.00026$), highlighting the importance of this compartment in tolerating stress. The cytoplasmic gene category is not displayed here, but is not enriched for rich media CE gene products.

Conditionally essential (CE) genes are essential for growth in a particular condition.

Deletions of such genes show very small colony sizes and high confidence negative

scores in particular conditions (see Experimental Procedures). We identified 197 CE genes, comprised of auxotrophs, which exhibit no growth in minimal media, and rich-media CE genes, which exhibit no growth in at least one rich media-based stress (Table S4). Importantly, our dataset had 70% overlap with a previous study of Keio Collection auxotrophs (Fig 2C, (Joyce et al., 2006) despite significant experimental differences (e.g. growth in liquid vs. solid media). Many of the remaining 30% were extremely sick, but above the stringent threshold we used to define no growth. We also identified 23 additional auxotrophs specific to alternative carbon/nitrogen sources not tested in the Joyce *et al.* study.

Genes essential for survival in natural environments are likely to extend beyond those required for laboratory growth, and could be targets for new antimicrobials (D'Elia et al., 2009). The 116 rich-media CE genes we identified (Fig. 2C) result from physiologically relevant stresses, and increase the current number of essential genes by roughly 30%. Interestingly, many of these gene products are located in the outer cell envelope (Fig 2D), a selective permeability barrier for gram-negative bacteria that is severely underrepresented for known essential genes (Fig. 2D). Many of the stresses generating CE phenotypes are part of the natural environment of *E. coli*, e.g. bile salts (Table S5), indicating that these genes are likely indispensable for *E. coli* to survive *in vivo*. Similarly, using the largest metagenomic dataset to date, Qin *et al.* reported that envelope-specific functions, such as adhesion, were commonly required for life in the gut (Qin et al., 2010).

Phenomic profiling helps assign function to uncharacterized genes

A key motivation for our study was to provide phenotypes for mutants of genes without functional annotation. Using a recently assembled compendium of such “orphan genes” in *E. coli* (Hu et al., 2009), we find that the fraction of mutant orphan genes with

phenotypes is close to that of annotated genes (Fig. 3A), but the former tend to have fewer phenotypes, indicating the power of phenomic analysis for identifying their phenotypes. Importantly, the phenotypic profiles of >25% of all orphan genes correlate strongly with those of annotated genes ($r \geq 0.5$; Fig. 3B & Table S6), providing high confidence leads (p -value $< 10^{-22}$) for discovery of their function. As these orphans are tied to a wide variety of cellular processes (Fig. 3C), the dataset will be of broad utility.

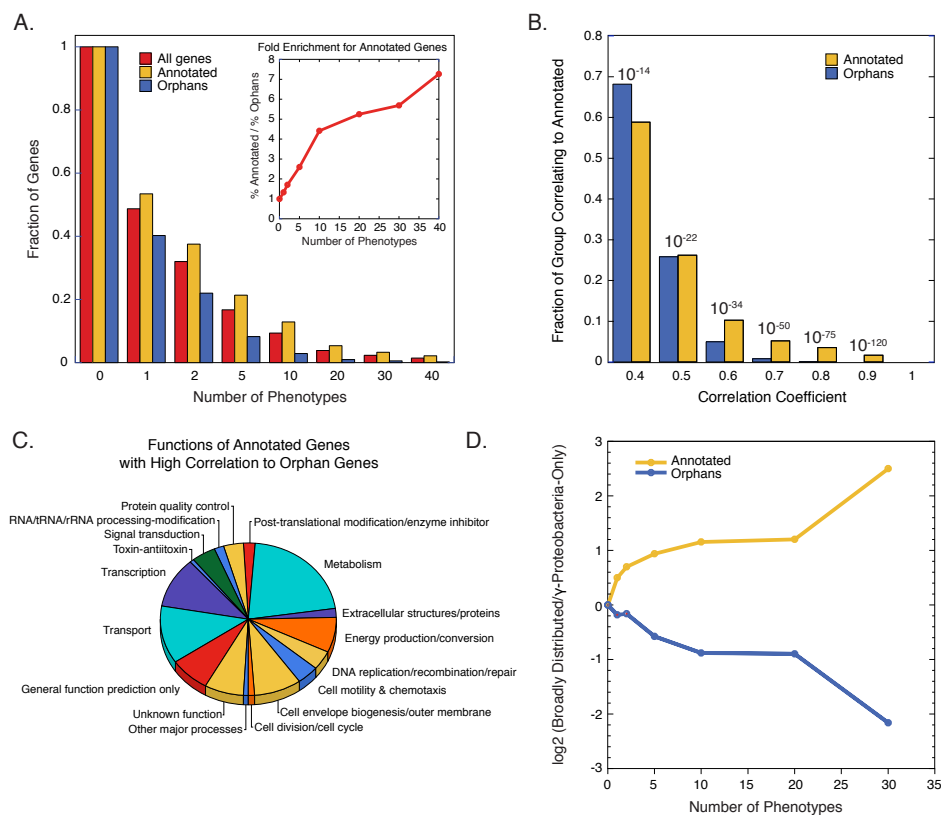


Figure 3. Phenomic profiling identifies phenotypes for orphan gene mutants. (A) Cumulative distribution of phenotypes indicating the fraction of gene mutants in each class having at least the number of phenotypes shown on the X-axis. The plot reveals that orphan gene mutants have phenotypes, but tend to have fewer phenotypes than annotated gene mutants. The insert quantifies phenotype deficit of orphan mutants. **(B)** Cumulative distribution of highly correlated pairs identifies many orphan genes that correlate highly to an annotated gene, providing high confidence clues to the function of the orphan gene. Values shown above each pair of bars are the p-values associated with pairwise correlation of any two strains at the indicated correlation coefficients. **(C)** High confidence correlations between orphans and

annotated genes ($r \geq 0.5$) provide leads related to many different cellular functions. Procedure for functional assignment is described in Fig. 2. Note that several “annotated genes” were classified as genes of “unknown function” or “general function prediction only” after manual curation. **(D)** Annotated genes responsible for many phenotypes tend to be broadly conserved, while the most responsive orphan genes tend to be restricted to γ -proteobacteria.

A small fraction of orphan gene knockouts have many phenotypes. Whereas annotated genes responsible for many phenotypes are broadly distributed among bacteria, the most responsive orphans tend to be narrowly distributed (Fig. 3D). This result suggests that evolutionary conservation is not a reliable indicator of the importance of an orphan gene to the organism, and that annotating them solely by homology has limitations. Such orphans may have evolved to fulfill an important but specialized function required by the niche of the organism. In support of this idea, a multi-responsive orphan identified in this study (*IpoB*) is restricted to enterobacteria and regulates peptidoglycan synthesis, a conserved process ubiquitous among bacteria (Typas et al.; Cell, accepted).

Using phenotypic signatures to identify gene function

Both correlated phenotypic signatures (Fig. 1C&D; Typas et al.; Cell, accepted) and anticorrelated phenotypic signatures have functional significance. For example, the phenotypic signatures of deletions of a transcriptional repressor and important target genes are likely to be anticorrelated. We find that *marR*⁻ and *marB*⁻ were highly anticorrelated with *acrB*⁻, whereas *marR*⁻ and *marB*⁻ were highly correlated (Fig. 4A). *marB* is a gene of unknown function in the multiple antibiotic resistance operon (*marRAB*), which also includes the operon repressor, *marR*, and its activator, *marA*. MarA also activates genes involved in antibiotic resistance, most importantly *acrAB*, encoding the major antibiotic efflux pump in *E. coli* (Fig. 4B; (Martin and Rosner, 2002). We explored whether MarB, like MarR, repressed MarA. Because of the inherent problems of high-throughput collections (suppressors, gene duplications, cross-

contamination), we always apply stringent quality control procedures to any follow-up investigations including PCR-validation of Keio isolates, and verification that retransduced strains maintain their phenotype. As *mar* is a hotspot for adaptive mutations, we also sequenced the entire operon and promoter region of single deletions and the double mutants we constructed.

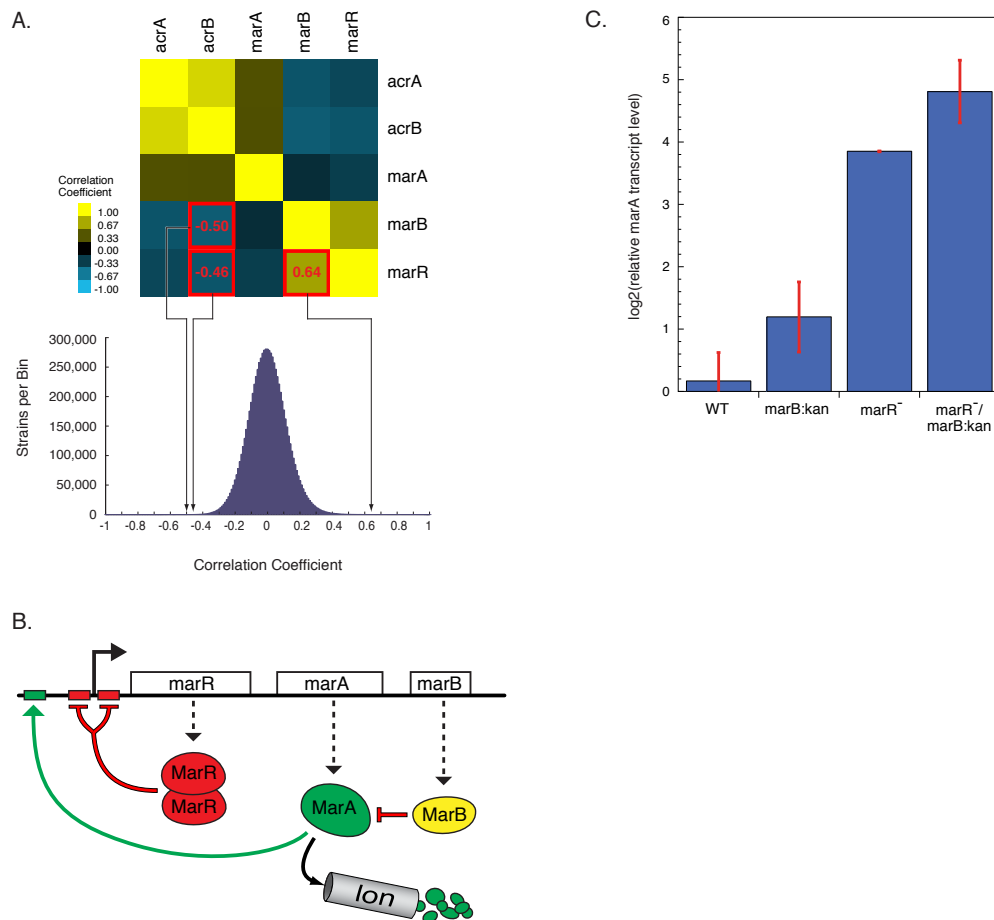


Figure 4. A function for *marB*. (A) *marR*⁻ and *marB*⁻ phenotypic signatures are highly correlated with each other, and are highly anticorrelated with that of *acrB* (top). The bottom graph positions these correlations in a histogram showing all pair-wise correlation coefficients between the 3979 mutants. (B) Schematic of the *E. coli* multiple antibiotic resistance (*mar*) operon. *marB* is a gene

of unknown function, but our results suggest it encodes a protein that inhibits MarA. (C) RT-PCR analysis shows that *marA* transcription is derepressed in *marB*⁻ cells. Derepression is independent of and additive with that of *marR*⁻.

Deletion of either *marB* or *marR* resulted in higher MarA levels, and the double *marRB* mutant showed additive effects on MarA transcript level (Fig. 4C) and protein level (data not shown). These effects were observed in both Kan-marked and clean deletions (Kan cassette excised, leaving an 82nt scar). The Δ *marR* strain exhibits ~2X more increase in MarA transcript levels than *marR::kan* (data not shown), arguing for a small polar effect of the cassette. Both *marB::kan* and Δ *marB* exhibit the same 2-fold increase in MarA levels (data not shown). These data suggest that MarB represses MarA independently of MarR. MarB does not have the signature of a DNA-binding protein, suggesting it acts post-transcriptionally. MarA level is controlled by the Lon protease (Griffith et al., 2004), but *lon*⁻ and *marB*⁻ effects are additive, indicating that MarB does not function through Lon (data not shown). MarA has been proposed to scan for activation sites while bound to RNA polymerase; by direct binding to either partner, MarB could disrupt complex formation. Alternatively, MarB may function in the periplasm. As MarB has a predicted periplasmic signal sequence, it could titrate an activating ligand for *mar* (e.g. salicylate).

Although *mar* is highly studied (~200 primary publications; Pubmed), our screen provided the first lead for MarB function. MarA targets approximately 40 genes, many of which are also co-regulated by the SoxS and Rob activators, with similar DNA-binding motifs as MarA (Martin et al., 2008; Martin and Rosner, 2002). The rules of engagement are poorly understood, but each activator responds to different environmental cues and overexpression of each leads to distinct phenotypes (Warner and Levy, 2010). It is likely that tight control of each activator impacts on the final gene expression output, which is

crucial for cellular proliferation. MarB may be an important player in fine-tuning the expression of MarA, especially since it is a conserved member of the *mar* operon, which has only recently emerged in selected pathogenic enterobacteria. Strong evidence for the importance of *mar* operon regulation in these organisms is that *mar* is a hotspot for mutations conferring higher drug resistance in *E. coli* (Nicoloff et al., 2007; Nicoloff et al., 2006).

Phenomic profiling reveals metabolic network behaviors under anti-folate drug stress

Tetrahydrofolate (THF) and its methyl/formylated derivatives are key molecules in all kingdoms of life for one-carbon metabolism. THF is used to synthesize glycine, methionine, purines and dTTP, in a process that leads to recycling of the THF species back to THF or dihydrofolate (DHF) (Fig. 5A). The bacterial THF biosynthesis pathway is targeted by two drugs: Sulfonamides (Sulfa) target FolP, and Trimethoprim (TMP) targets FolA (Fig. 5A). Dual inhibition by Sulfa and TMP is strongly synergistic, and therefore these drugs are almost exclusively administered in combination for treatment of ear, urinary tract and bronchial infections. Despite extensive clinical use and years of laboratory investigation, we lack a complete mechanistic understanding of why these drugs are strongly synergistic. A network feature identified by phenomic profiling could contribute to synergy.

We find that the two drug classes have major phenotypic differences. Sulfa and TMP treatments are highly correlated within their class ($r=0.57$ for Sulfa; 0.67 for TMP), but poorly correlated with each other ($r=0.15 \pm 0.04$), just slightly more than the correlation observed between all screens ($r=0.025 \pm 0.12$). Thus, subinhibitory TMP and Sulfa treatments have fundamentally different effects on the cell, even though both partially

block THF biosynthesis. Importantly, removing enzymes acting directly downstream of THF production resulted in opposite drug sensitivities: the serine hydroxymethyltransferase mutant (*glyA::kan*) was sensitive only to TMP; conversely, glycine cleavage (GCV) mutants (*gcvP::kan*, *gcvH::kan*, *gcvT::kan*) were sensitive only to Sulfa (Fig 5B). The mutant results were reproduced in liquid culture (Fig. 5D), where *glyA*⁻ TMP sensitivity is manifested as a growth rate phenotype (left panel), and *gcvP*⁻ sulfamethizole (SMT) sensitivity is registered as a low stationary phase density (right panel).

GlyA and GCV lie on opposite sides of a branched pathway that converts THF to 5,10-methylene THF (5,10-mTHF; Fig 5A). As *glyA* and *gcv* mutants exhibit synthetic lethality, they are the only routes to production of this essential metabolite (Fig. 5C). A simple explanation for the differential responses of *glyA*⁻ and *gcvP*⁻ is that 5,10-mTHF is predominantly produced via different branches under each drug treatment. A corollary is that combination drug treatment inhibits both branches, resulting in synergistic limitation for 5,10-mTHF, before the pools of THF are depleted. In support of this idea, despite the increased sensitivity of *glyA*⁻ and *gcvP*⁻ to single drugs, these strains grew no more poorly than wt under the drug combination (Fig. 5D). Thus, genetically eliminating either branch of the pathway reduced but did not eliminate synergy. The downstream biosynthetic reactions are also differentially affected by TMP and Sulfa (Fig. S3A), and we are currently testing whether they partially account for the residual synergy.

Streptococcus pneumoniae lacks the GCV system and exhibits significantly less drug synergy than *E. coli* across different growth conditions (Fig. 5E & data not shown). We performed our comparison using concentrations of TMP and SMT that caused the same relative growth defect in each species (Fig. 5E). These data together support the

hypothesis that simultaneous inhibition of the branched pathway for production of 5,10-mTHF contributes to the observed anti-folate synergy in *E. coli*.

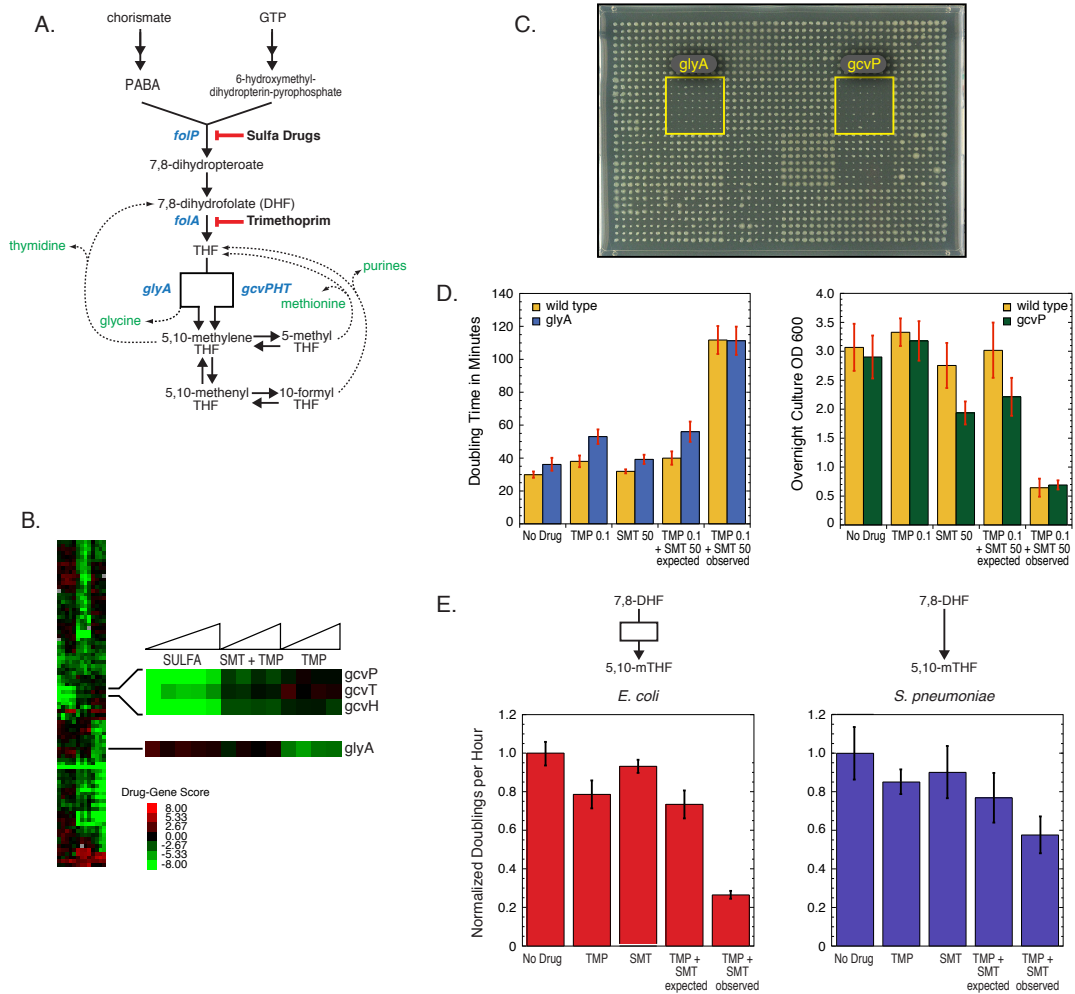


Figure 5. A new network feature contributing to anti-folate drug synergy. (A) Schematic of the *E. coli* tetrahydrofolate (THF) biosynthesis pathway and the enzymatic steps inhibited by Sulfa and TMP. (B) Clustergram of genes that respond to Sulfa, TMP, or the combination. Zoomed image indicates that *gcv* mutants are sensitive to Sulfa, *glyA* is sensitive to TMP, and that these four mutants exhibit essentially wildtype growth in response to the drug combination. (C) *glyA*⁻ and *gcvP*⁻ are a synthetic lethal pair. Image of a plate mating between the donor Hfr *gcvP::cat* and 24 *kan^R* recipients (arrayed in boxes of 8x8 colonies), grown on kanamycin/chloramphenicol medium to select for double mutant strains; position of the *glyA::kan* and *gcvP::kan* recipients is highlighted. (D) Liquid culture experiments verify growth phenotypes on agar plates shown in Fig.5B. The TMP/SMT combination has less synergy than wildtype in *glyA*⁻ and *gcvP*⁻ cells. Concentrations shown for TMP and SMT are in $\mu\text{g}/\text{mL}$. (E) Quantification of synergy in *E. coli* and *S. pneumoniae*, which lacks the branched pathway for generating 5,10-mTHF present in *E. coli*. Comparisons were performed using single drug concentrations giving

equivalent inhibition of both organisms. *S. pneumoniae* has reduced synergy compared to *E. coli*.

Our data do not indicate whether differential effects of TMP and Sulfa on GlyA and GCV result from differential inhibition of expression or activity, or the intrinsic properties of each enzyme. We favor the idea that differential metabolite accumulation and subsequent feed-forward enzymatic regulation make a contribution to the distinct cellular responses to these two drugs. Recent metabolomic flux analyses indicate that high doses of TMP lead to accumulation and depletion of select metabolites, as well as to protein-level regulation of portions of the network (Kwon et al., 2010; Kwon et al., 2008). Although a comparable analysis has not been performed for Sulfa drugs, deletion of the predicted 5-formyl-THF cycloligase, *ygfA*, which likely degrades 5-formyl-THF (Jeanguenin et al., 2010), clusters tightly with the *gcv* mutants, and exhibits sensitivity only to Sulfa drugs (Fig. S3A). That 5-formyl-THF degradation is critical only under Sulfa stress suggests differential accumulation (or requirement) of THF species under Sulfa and TMP treatments. 5-formyl-THF is a known inhibitor of several enzymes in the THF network of other organisms (Field et al., 2006; Stover and Schirch, 1993), and could act as an effective protein-level regulator. Similarly, a strain lacking a predicted alanine racemase, *yggS*, is sensitive only to Sulfa; D-alanine is known to inactivate GlyA (Schirch et al., 1985), and *yggS*⁻ and *glyA*⁻ form a synthetic lethal pair (Fig. S3B). Thus, the different cellular responses to these two drugs may be due in part to metabolite-based enzymatic regulation. An extension is that the synergy of combination therapy could rest primarily on complementary inhibition of different one-carbon biosynthesis reactions, and therefore recycling of THFs. This model would allow for synergy even with the expected additive limitation of THF production.

In summary, our results illustrate the power of phenomic profiling to yield insights into drug action and the ability of a networks view to provide new paradigms for analysis of drug interaction mechanisms, which can facilitate hypothesis-driven research on drug interactions (Bollenbach et al., 2009). This type of analysis may be generally useful in predicting drug synergies, and in explaining variable drug-drug interactions across species.

Phenomic profiling gives insights into genomic organization

The *E. coli* genome is encoded on a single, circular chromosome, with a single origin of replication, *oriC*. Essential genes are biased to the plus (+) strand, where transcription proceeds in the same direction as DNA replication. This may avert head-on collisions between RNA and DNA polymerases that would result in aborted transcripts, truncated, or frame-shifted proteins (Rocha and Danchin, 2003). Here we show that responsive and CE genes, which are important for optimal growth of the organism, also show + strand bias (Fig. 6A). Indeed, the weighted responsive genome (responsive genes weighted by number of phenotypes identified) is heavily biased to the + strand, indicating great selective pressure to place genes important for rapid growth on the + strand.

Conversely, the non-responsive genome is biased to the minus (-) strand. As our approaches expand to incorporate additional phenotypic readouts more important for cells with reduced division and DNA replication (e.g. biofilm formation), the + strand bias of responsive genes will presumably be reduced.

The chromosome is massively compacted in the cell to create the nucleoid, which is thought to contribute significantly to the organization of gene expression (Travers and Muskhelishvili, 2005; Vora et al., 2009). Chromosomal loci have spatial addresses in the cell, corresponding closely to their chromosomal position (Toro and Shapiro, 2010).

Additionally, highly expressed genes associated with transcription and translation are located near the origin of replication (*oriC*), presumably to benefit from the “gene

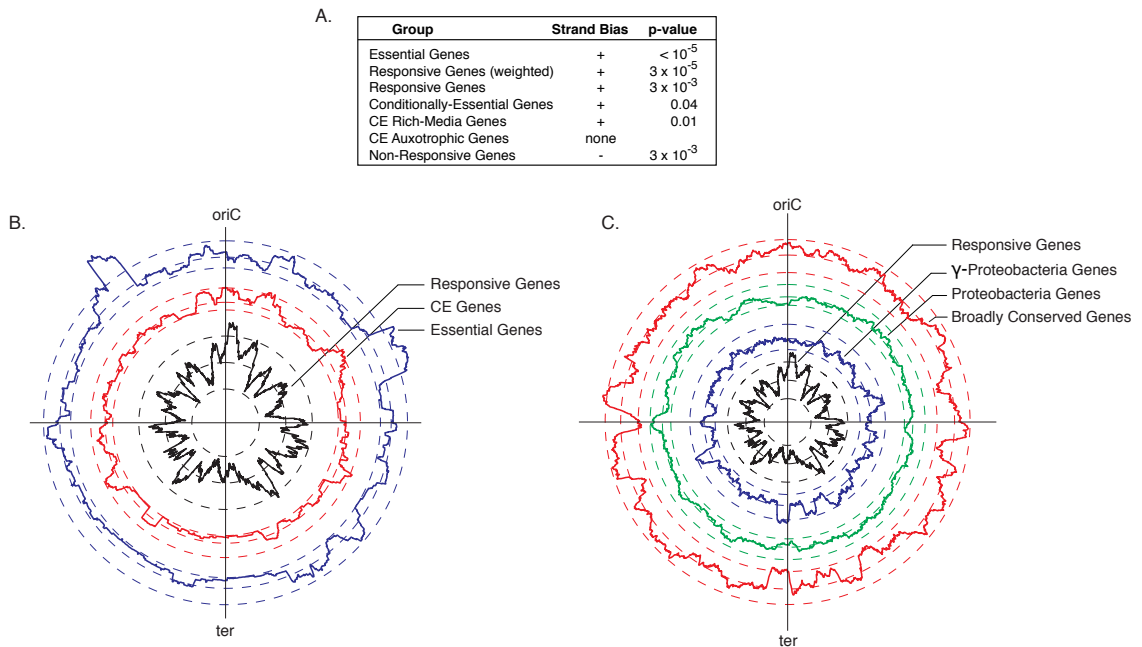


Figure 6. Phenomic profiling generates insights into genome organization. (A) Essential and responsive genes are biased to the plus strand of DNA (transcription direction coincident with replication) and the non-responsive genes are biased to the minus strand of DNA. (B-C) For each panel, circular plots depict gene position, adjusting coordinates so that the chromosome starts at the origin of replication (*oriC* = 0 bp); the terminus region (*ter*) is opposite the *oriC*. Each trace represents spatial enrichment for the variable plotted based on a 100kb sliding window. Three dashed lines of the same color accompany each trace indicating the minimum permutation threshold, the baseline representing zero enrichment, and the maximum permutation threshold (inside to outside). Permutation thresholds are the result of 1000 randomizations of gene class assignments (see Experimental Procedures), and indicate significant negative and positive spatial enrichment at a p-value of 0.05. (B) Responsive and CE genes are concentrated around the *oriC*, and scarce around the terminus. (C) The terminus is positively enriched for genes restricted to the γ -proteobacteria, and negatively enriched for broadly conserved genes.

dosage” effect created when rapidly growing cells initiate multiple rounds of DNA replication per division (Couturier and Rocha, 2006). Projecting the spatial distribution of the responsive genes onto the circular chromosome (Fig. 6B, black trace) provides us with a snapshot of the *E. coli* genome from a functional perspective. This projection is

based on a 100kb sliding window and therefore captures organization above the operon level (Fig. S4A; Experimental Procedures). A pattern of alternating peaks and valleys is clearly evident, indicating that responsive genes cluster spatially into large chromosomal regions separated by regions generally devoid of responsive genes. “Valleys” are comprised of spatially separated operons often transcribed from different strands, indicating that low responsiveness is a regional characteristic rather than an artifact due to large non-responsive operons. Our finding of clustering above the operon level is in accord with other studies showing that gene expression is broadly correlated across certain regions of the chromosome (Carpentier et al., 2005; Jeong et al., 2004).

The responsive genome is most enriched around *oriC*, which has the highest concentration of responsive genes (Fig. 6B, black trace). This area is also enriched for the most responsive genes (Fig. S4B), and for conditionally essential (CE) genes (Fig. 6B, red trace), providing strong support for the idea that the *E. coli* chromosome tends to store genes of high functional importance near the *oriC*. In contrast, the terminus region is relatively devoid of responsive genes (Fig 6B, black trace), has a paucity of broadly conserved genes (Fig 6C, red trace) and a corresponding enrichment for genes restricted to γ -proteobacteria (Fig 6C, blue trace). We postulate that the terminus region contains newly acquired genes that have yet to fully integrate into the cellular network, and tend to lack phenotypes. This could enable cells to avoid unnecessarily high expression of such genes as a consequence of the gene-dosage effect. Should this result prove true across bacterial species, it could point to a general organizing principle of circular chromosomes.

Phenomic Profiling Describes Drug Action

“Drug-centric” analyses are more complex than “gene-centric” analyses. Whereas genes

mostly participate in a single biological process, many parameters are required to describe drug action: uptake, primary/secondary targets, efflux. Therefore, pairwise relationships between drugs are more complex than those between genes. For example, two drugs may cluster based on drug uptake, even though their primary targets differ. In addition, drug signatures are an order of magnitude larger than gene signatures (3979 vs. 324). To reduce the complexity of drug signatures, we calculated Drug-Gene Ontology (GO) scores, which represent the probability that a given GO group specifically interacts with a given drug (i.e. number of phenotypes associated with genes in the GO group vs. across the entire dataset). We used these Drug-GO scores to explore drug mode-of-action through a network-based clustering strategy (see Experimental Procedures). The position of drugs in the network (Fig. 7A) is based both on the magnitude of their Drug-Gene Ontology (GO) scores (gray) and on Drug-Drug correlations (yellow). Of the 719 significant Drug-GO interactions (p -value $\leq 10^{-3}$), which include 64 drugs and 218 GO groups, 657 were negative and only 62 were positive. Thus, disrupting a linked biological process was very likely to increase drug sensitivity (Table S7). Drug-Drug correlations increased the resolution of the network and captured drug similarities that escaped the Drug-GO analysis.

We found that drugs with the same cellular target tend to cluster. For example, drugs targeting DNA (orange) fall in the lower right, those targeting THF-biosynthesis (light green) fall on the bottom edge and those targeting peptidoglycan (PG; purple) predominantly cluster in the upper left. Interestingly, β -lactams cover the center of the PG cluster, whereas drugs targeting other steps of PG synthesis are located at the periphery. The correlation coefficients between β -lactams reveal that the similarity of their phenotypic signatures is related to their respective primary target Penicillin Binding

Proteins (Fig. S5A). Interestingly, known synergistic double drug combinations (TMP/sulfonamides and mecillinam/cefsulodin) occupy spaces distinct from either individual drug, arguing that the combination elicits a different cellular response from the individual drugs. It will be interesting to determine whether this holds true for antagonistic or neutral interactions or whether these combinations elicit responses closer to one or both drugs.

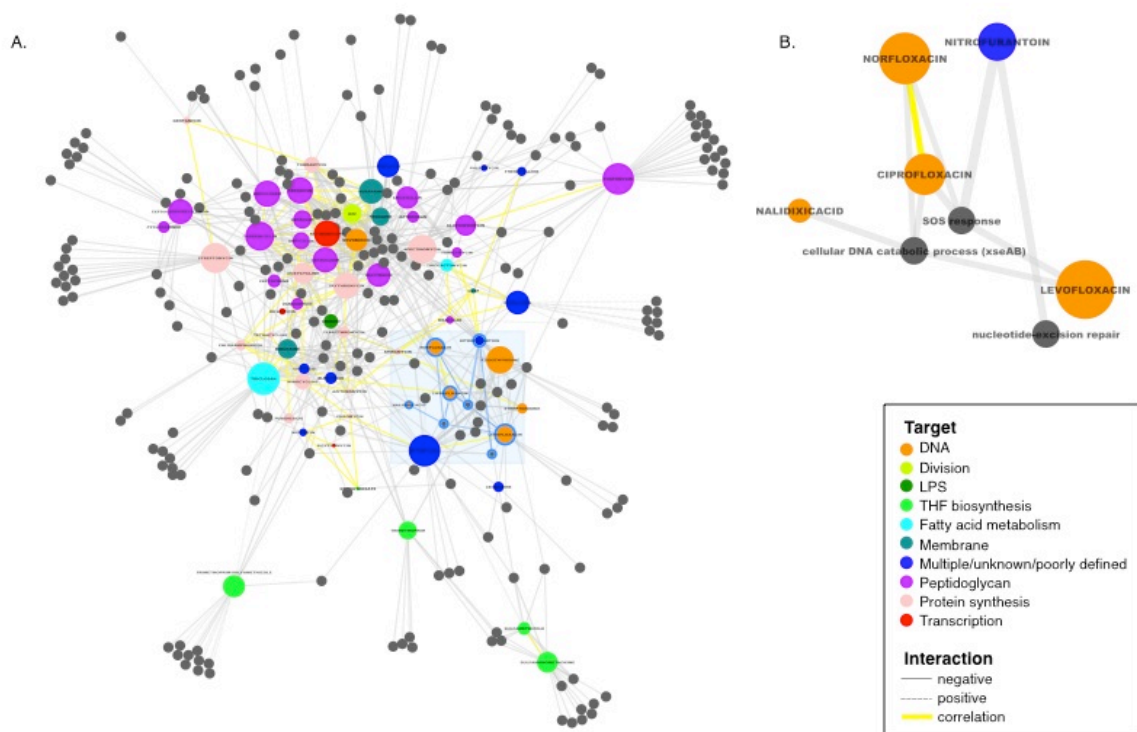


Figure 7. Network view reveals new insights into drug action. (A) Colored nodes represent all drugs profiled in this study found to have significant interactions with Gene Ontology (GO) biological process groups (gray nodes). Connections between nodes represent significant Drug-GO interactions ($p\text{-value} \leq 10^{-3}$, gray) or high Drug-Drug correlation ($r \geq 0.32$, $p\text{-value} \leq 10^{-97}$, yellow). Drug node size is based on the number of connections associated with that node, i.e. larger nodes have more Drug-GO interactions. Spatial clustering is driven by the p -values of Drug-GO interactions and Drug-Drug correlations, resulting in drugs with similar cellular action lying near each other in the network. Drugs with multiple, unknown, or poorly defined targets are shown in dark blue. (B) Zoomed view of subnetwork shadowed by light blue box in (A). All four quinolones screened (orange) interact negatively with *xseAB* (exonuclease VII) and are the only

drugs that require the exonuclease, $p\text{-value}=10^{-6}$. NTF is found to activate the SOS response, and create lesions requiring nucleotide-excision repair (Fig S5B).

Importantly, specific Drug-GO interactions suggest hypotheses for mechanism of action even for well-studied drugs. Quinolones inhibit DNA gyrase by trapping it as a quinolone adduct, whose mechanism of resolution is poorly understood (Drlica et al., 2008). One GO category, “cellular DNA catabolic process” (*xseAB*) selectively and specifically ($p\text{-value}=10^{-6}$) interacted negatively with all four quinolones screened (Fig 7B & S5B), expanding on a previous report of *xseAB* mutant sensitivity to fluoroquinolones (Tamae et al., 2008). We suggest that XseAB (exonuclease VII) is the enzyme that cleaves quinolone-bound DNA gyrase from the DNA to allow repair to proceed, a possibility we are currently exploring.

Our drug network also provides clues for the mode of action of poorly described drugs, and, conversely suggests that additional factors are required to explain the action of other drugs. Nitrofurantoin (NTF; Fig. 7B) is reported to have a multi-faceted impact on cells (McOsker and Fitzpatrick, 1994; Tu and McCalla, 1975), but our data suggests its cytotoxic effects reflect DNA damage, as it causes lesions requiring nucleotide-excision repair (NER) and activates the SOS response. NTF is the only DNA-targeting drug requiring NER, but not double-strand break repair, suggesting that its primary toxic lesion is associated with the replication fork (Fig. S5B). Additionally, our network analysis validates the idea that indolicidin, a neutrophil antimicrobial peptide, mediates its effects by compromising the inner membrane permeability of *E. coli* in a manner similar to the proton motive force uncoupler, CCCP (Falla et al., 1996). Finally, phleomycin and bleomycin do not cluster with DNA response drugs, suggesting they have broader cellular impact (Hecht, 2000; Yeh et al., 2006), from inducing DNA

scissions (Giloni et al., 1981). These insights suggest that this phenomic dataset is a rich source for discovery of drug function and interrelationships.

Perspectives

To keep pace with exploding sequence information, cost-effective, high-throughput phenotyping technologies must be developed. Here we show that phenomic profiling in *E. coli* fulfills this goal. Our dataset is of great utility in identifying the function of orphan genes. Three cases (*marB*, *lpoA*, *lpoB*) were investigated here or in a study based on this dataset (Typas et al., Cell, accepted), and we are actively pursuing functional discovery of numerous (>20) orphan genes, as well as annotated genes with previously unsuspected roles in collaboration with others. Since >25% of the orphan genes are highly correlated to an annotated gene ($r \geq 0.5$), this dataset provides a rapid method for function discovery.

An important finding is that the most responsive orphan genes tend to be narrowly distributed among bacteria. Interestingly, our results mirror initial observations from human microbiome studies. These studies found that: a) roughly half of the functions encoded in the minimal gut metagenome (ubiquitously present in all 124 individuals screened) are both unknown and of limited evolutionary conservation (Qin et al., 2010); b) across 4 pan-genome species analyzed, the vast majority of non-common genes were of either unknown function (~70%) or unique family members of functions that were part of the core gene set (Nelson et al., 2010). The latter are probably species-specific additions to conserved biological processes of the pan-genome. Together these studies argue that when computational methods based on gene conservation fail, large-scale phenomic analyses can be a second tier for assigning function. To make this approach a reality, low cost methods for developing deletion libraries must be developed (Goodman

et al., 2009). Single-gene deletion ordered libraries are currently available for only a handful of organisms [(Cameron et al., 2008; de Berardinis et al., 2008; Gallagher et al., 2007; Goodman et al., 2009; Kim et al., 2010; Liu et al., 2008; Noble et al., 2010) and references in (Barker et al., 2010)], but advances in transposon mutagenesis make it feasible to create ordered mutant libraries in most organisms. In *E. coli*, expansion of this work will rest on the ability to assess additional phenotypes through deeper exploration of phenotypic space. The greatest potential resides at the intersection of screening more diverse stresses and incorporating additional cellular readouts. Colorimetric readouts would enable measurement of transcriptional activity or biofilm formation on solid agar surfaces, and represent an immediate potential advance for phenomic profiling. High-throughput microscopy would provide a new avenue for such approaches (Werner et al., 2009).

Our dataset provides information on a substantial collection of antibiotics/antimicrobial compounds that cover a broad spectrum of drug targets, structural classes and drug generations, providing a platform for future studies focused on natural products or antimicrobials with unknown targets. Our dataset can also provide a platform for studying the mechanism behind drug interactions (Yeh et al., 2009), as shown here for the case of sulfonamides and TMP. Understanding the mechanism underlying known drug interactions may help to predict novel interactions and manipulate existing drug combinations to increase their effectiveness in the clinic.

In summary, we have generated a valuable resource for microbiologists studying a wide range of biology, and demonstrated the numerous and diverse applications of this dataset to infer information both on gene and drug function. As the most comprehensive prokaryotic chemical genomic study to date (3979 strains x 324 conditions), our dataset

will serve as a base for future studies that aim to increase information and/or resolution on both the gene and drug fronts. We hope that the usefulness of this resource will trigger analogous studies in other organisms, bringing us a step nearer to closing the gene sequence-function gap.

Experimental procedures

Experimental procedures are partially elaborated in the text and figure legends, and are fully explicated in supplementary material.

References

- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2, 2006 0008.
- Barker, C.A., Farha, M.A., and Brown, E.D. (2010). Chemical genomic approaches to study model microbes. *Chem Biol* 17, 624-632.
- Beltrao, P., Cagney, G., and Krogan, N.J. (2010). Quantitative genetic interactions reveal biological modularity. *Cell* 141, 739-745.
- Bochner, B.R. (2009). Global phenotypic characterization of bacteria. *FEMS Microbiol Rev* 33, 191-205.
- Bollenbach, T., Quan, S., Chait, R., and Kishony, R. (2009). Nonoptimal microbial response to antibiotics underlies suppressive drug interactions. *Cell* 139, 707-718.
- Butland, G., Babu, M., Diaz-Mejia, J.J., Bohdana, F., Phanse, S., Gold, B., Yang, W., Li, J., Gagarinova, A.G., Pogoutse, O., *et al.* (2008). eSGA: *E. coli* synthetic genetic array analysis. *Nat Methods* 5, 789-795.
- Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., *et al.* (2005). Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 433, 531-537.
- Cameron, D.E., Urbach, J.M., and Mekalanos, J.J. (2008). A defined transposon mutant library and its use in identifying motility genes in *Vibrio cholerae*. *Proc Natl Acad Sci U S A* 105, 8736-8741.
- Carpentier, A.S., Torresani, B., Grossmann, A., and Henaut, A. (2005). Decoding the nucleoid organisation of *Bacillus subtilis* and *Escherichia coli* through gene expression data. *BMC Genomics* 6, 84.
- Couturier, E., and Rocha, E.P. (2006). Replication-associated gene dosage effects shape the genomes of fast-growing bacteria but only for transcription and translation genes. *Mol Microbiol* 59, 1506-1518.
- D'Elia, M.A., Pereira, M.P., and Brown, E.D. (2009). Are essential genes really essential? *Trends Microbiol* 17, 433-438.
- de Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., Lechaplais, C., Gyapay, G., Richez, C., *et al.* (2008). A complete collection of single-gene deletion mutants of *Acinetobacter baylyi* ADP1. *Mol Syst Biol* 4, 174.
- Drlica, K., Malik, M., Kerns, R.J., and Zhao, X. (2008). Quinolone-mediated bacterial death. *Antimicrob Agents Chemother* 52, 385-392.
- Falla, T.J., Karunaratne, D.N., and Hancock, R.E. (1996). Mode of action of the antimicrobial peptide indolicidin. *J Biol Chem* 271, 19298-19303.

Fiedler, D., Braberg, H., Mehta, M., Chechik, G., Cagney, G., Mukherjee, P., Silva, A.C., Shales, M., Collins, S.R., van Wageningen, S., *et al.* (2009). Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* **136**, 952-963.

Field, M.S., Szebenyi, D.M., and Stover, P.J. (2006). Regulation of de novo purine biosynthesis by methenyltetrahydrofolate synthetase in neuroblastoma. *J Biol Chem* **281**, 4215-4221.

Gallagher, L.A., Ramage, E., Jacobs, M.A., Kaul, R., Brittnacher, M., and Manoil, C. (2007). A comprehensive transposon mutant library of *Francisella novicida*, a bioweapon surrogate. *Proc Natl Acad Sci U S A* **104**, 1009-1014.

Giaever, G., Flaherty, P., Kumm, J., Proctor, M., Nislow, C., Jaramillo, D.F., Chu, A.M., Jordan, M.I., Arkin, A.P., and Davis, R.W. (2004). Chemogenomic profiling: identifying the functional interactions of small molecules in yeast. *Proc Natl Acad Sci U S A* **101**, 793-798.

Giloni, L., Takeshita, M., Johnson, F., Iden, C., and Grollman, A.P. (1981). Bleomycin-induced strand-scission of DNA. Mechanism of deoxyribose cleavage. *J Biol Chem* **256**, 8608-8615.

Girgis, H.S., Hottes, A.K., and Tavazoie, S. (2009). Genetic architecture of intrinsic antibiotic susceptibility. *PLoS One* **4**, e5629.

Girgis, H.S., Liu, Y., Ryu, W.S., and Tavazoie, S. (2007). A comprehensive genetic characterization of bacterial motility. *PLoS Genet* **3**, 1644-1660.

Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying genetic determinants needed to establish a human gut symbiont in its habitat. *Cell Host Microbe* **6**, 279-289.

Griffith, K.L., Shah, I.M., and Wolf, R.E., Jr. (2004). Proteolytic degradation of *Escherichia coli* transcription activators SoxS and MarA as the mechanism for reversing the induction of the superoxide (SoxRS) and multiple antibiotic resistance (Mar) regulons. *Mol Microbiol* **51**, 1801-1816.

Hecht, S.M. (2000). Bleomycin: new perspectives on the mechanism of action. *J Nat Prod* **63**, 158-168.

Hillenmeyer, M.E., Fung, E., Wildenhain, J., Pierce, S.E., Hoon, S., Lee, W., Proctor, M., St Onge, R.P., Tyers, M., Koller, D., *et al.* (2008). The chemical genomic portrait of yeast: uncovering a phenotype for all genes. *Science* **320**, 362-365.

Hobbs, E.C., Astarita, J.L., and 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.

Hoon, S., Smith, A.M., Wallace, I.M., Suresh, S., Miranda, M., Fung, E., Proctor, M., Shokat, K.M., Zhang, C., Davis, R.W., *et al.* (2008). An integrated platform of genomic assays reveals small-molecule bioactivities. *Nat Chem Biol* **4**, 498-506.

Hu, P., Janga, S.C., Babu, M., Diaz-Mejia, J.J., Butland, G., Yang, W., Pogoutse, O., Guo, X., Phanse, S., Wong, P., *et al.* (2009). Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol* **7**, e96.

Jeanguenin, L., Lara-Nunez, A., Pribat, A., Hamner Mageroy, M., Gregory, J.F., 3rd, Rice, K.C., de Crecy-Lagard, V., and Hanson, A.D. (2010). Moonlighting glutamate formiminotransferases can functionally replace 5-formyltetrahydrofolate cycloligase. *J Biol Chem*.

Jeong, K.S., Ahn, J., and Khodursky, A.B. (2004). Spatial patterns of transcriptional activity in the chromosome of *Escherichia coli*. *Genome Biol* **5**, R86.

Joyce, A.R., Reed, J.L., White, A., Edwards, R., Osterman, A., Baba, T., Mori, H., Lesely, S.A., Palsson, B.O., and Agarwalla, S. (2006). Experimental and computational assessment of conditionally essential genes in *Escherichia coli*. *J Bacteriol* **188**, 8259-8271.

Kim, D.U., Hayles, J., Kim, D., Wood, V., Park, H.O., Won, M., Yoo, H.S., Duhig, T., Nam, M., Palmer, G., *et al.* (2010). Analysis of a genome-wide set of gene deletions in the fission yeast *Schizosaccharomyces pombe*. *Nat Biotechnol* **28**, 617-623.

Kohanski, M.A., Dwyer, D.J., Wierzbowski, J., Cottarel, G., and Collins, J.J. (2008). Mistranslation of membrane proteins and two-component system activation trigger antibiotic-mediated cell death. *Cell* **135**, 679-690.

Kwon, Y.K., Higgins, M.B., and Rabinowitz, J.D. (2010). Antifolate-induced depletion of intracellular glycine and purines inhibits thymineless death in *E. coli*. *ACS Chem Biol* **5**, 787-795.

Kwon, Y.K., Lu, W., Melamud, E., Khanam, N., Bognar, A., and Rabinowitz, J.D. (2008). A domino effect in antifolate drug action in *Escherichia coli*. *Nat Chem Biol* **4**, 602-608.

Lee, W., St Onge, R.P., Proctor, M., Flaherty, P., Jordan, M.I., Arkin, A.P., Davis, R.W., Nislow, C., and Giaever, G. (2005). Genome-wide requirements for resistance to functionally distinct DNA-damaging agents. *PLoS Genet* **1**, e24.

Liu, A., Tran, L., Becket, E., Lee, K., Chinn, L., Park, E., Tran, K., and Miller, J.H. (2010). Antibiotic sensitivity profiles determined with an *Escherichia coli* gene knockout collection: generating an antibiotic bar code. *Antimicrob Agents Chemother* **54**, 1393-1403.

Liu, O.W., Chun, C.D., Chow, E.D., Chen, C., Madhani, H.D., and Noble, S.M. (2008). Systematic genetic analysis of virulence in the human fungal pathogen *Cryptococcus neoformans*. *Cell* **135**, 174-188.

Martin, R.G., Bartlett, E.S., Rosner, J.L., and Wall, M.E. (2008). Activation of the *Escherichia coli* *marA/soxS/rob* regulon in response to transcriptional activator concentration. *J Mol Biol* **380**, 278-284.

Martin, R.G., and Rosner, J.L. (2002). Genomics of the *marA/soxS/rob* regulon of *Escherichia coli*: identification of directly activated promoters by application of molecular genetics and informatics to microarray data. *Mol Microbiol* **44**, 1611-1624.

McOsker, C.C., and Fitzpatrick, P.M. (1994). Nitrofurantoin: mechanism of action and implications for resistance development in common uropathogens. *J Antimicrob Chemother* **33 Suppl A**, 23-30.

Nelson, K.E., Weinstock, G.M., Highlander, S.K., Worley, K.C., Creasy, H.H., Wortman, J.R., Rusch, D.B., Mitreva, M., Sodergren, E., Chinwalla, A.T., *et al.* (2010). A catalog of reference genomes from the human microbiome. *Science* **328**, 994-999.

Nicoloff, H., Perreten, V., and Levy, S.B. (2007). Increased genome instability in *Escherichia coli* *lon* mutants: relation to emergence of multiple-antibiotic-resistant (Mar) mutants caused by insertion sequence elements and large tandem genomic amplifications. *Antimicrob Agents Chemother* **51**, 1293-1303.

Nicoloff, H., Perreten, V., McMurry, L.M., and Levy, S.B. (2006). Role for tandem duplication and *lon* protease in AcrAB-TolC- dependent multiple antibiotic resistance (Mar) in an *Escherichia coli* mutant without mutations in *marRAB* or *acrRAB*. *J Bacteriol* **188**, 4413-4423.

Noble, S.M., French, S., Kohn, L.A., Chen, V., and Johnson, A.D. (2010). Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet* **42**, 590-598.

Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F., and Boeke, J.D. (2004). A robust toolkit for functional profiling of the yeast genome. *Mol Cell* **16**, 487-496.

Parsons, A.B., Lopez, A., Givoni, I.E., Williams, D.E., Gray, C.A., Porter, J., Chua, G., Sopko, R., Brost, R.L., Ho, C.H., *et al.* (2006). Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* **126**, 611-625.

Pathania, R., Zlitni, S., Barker, C., Das, R., Gerritsma, D.A., Lebert, J., Awuah, E., Melacini, G., Capretta, F.A., and Brown, E.D. (2009). Chemical genomics in *Escherichia coli* identifies an inhibitor of bacterial lipoprotein targeting. *Nat Chem Biol* **5**, 849-856.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., *et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59-65.

Rocha, E.P., and Danchin, A. (2003). Essentiality, not expressiveness, drives gene-strand bias in bacteria. *Nat Genet* **34**, 377-378.

Roguev, A., Bandyopadhyay, S., Zofall, M., Zhang, K., Fischer, T., Collins, S.R., Qu, H., Shales, M., Park, H.O., Hayles, J., *et al.* (2008). Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast. *Science* **322**, 405-410.

Schirch, V., Hopkins, S., Villar, E., and Angelaccio, S. (1985). Serine hydroxymethyltransferase from *Escherichia coli*: purification and properties. *J Bacteriol* **163**, 1-7.

Schuldiner, M., Collins, S.R., Thompson, N.J., Denic, V., Bhamidipati, A., Punna, T., Ihmels, J., Andrews, B., Boone, C., Greenblatt, J.F., *et al.* (2005). Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell* **123**, 507-519.

- Stover, P., and Schirch, V. (1993). The metabolic role of leucovorin. *Trends Biochem Sci* 18, 102-106.
- Taber, H.W., Mueller, J.P., Miller, P.F., and Arrow, A.S. (1987). Bacterial uptake of aminoglycoside antibiotics. *Microbiol Rev* 51, 439-457.
- Tamae, C., Liu, A., Kim, K., Sitz, D., Hong, J., Becket, E., Bui, A., Solaimani, P., Tran, K.P., Yang, H., *et al.* (2008). Determination of antibiotic hypersensitivity among 4,000 single-gene-knockout mutants of *Escherichia coli*. *J Bacteriol* 190, 5981-5988.
- Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Page, N., Robinson, M., Raghizadeh, S., Hogue, C.W., Bussey, H., *et al.* (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294, 2364-2368.
- Toro, E., and Shapiro, L. (2010). Bacterial chromosome organization and segregation. *Cold Spring Harb Perspect Biol* 2, a000349.
- Travers, A., and Muskhelishvili, G. (2005). DNA supercoiling - a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol* 3, 157-169.
- Tu, Y., and McCalla, D.R. (1975). Effect of activated nitrofurans on DNA. *Biochim Biophys Acta* 402, 142-149.
- Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L., *et al.* (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nat Methods* 5, 781-787.
- Vora, T., Hottes, A.K., and Tavazoie, S. (2009). Protein occupancy landscape of a bacterial genome. *Mol Cell* 35, 247-253.
- Warner, D.M., and Levy, S.B. (2010). Different effects of transcriptional regulators MarA, SoxS and Rob on susceptibility of *Escherichia coli* to cationic antimicrobial peptides (CAMPs): Rob-dependent CAMP induction of the marRAB operon. *Microbiology* 156, 570-578.
- Warner, J.R., Reeder, P.J., Karimpour-Fard, A., Woodruff, L.B., and Gill, R.T. (2010). Rapid profiling of a microbial genome using mixtures of barcoded oligonucleotides. *Nat Biotechnol* 28, 856-862.
- Werner, J.N., Chen, E.Y., Guberman, J.M., Zippilli, A.R., Irgon, J.J., and Gitai, Z. (2009). Quantitative genome-scale analysis of protein localization in an asymmetric bacterium. *Proc Natl Acad Sci U S A* 106, 7858-7863.
- Xu, D., Jiang, B., Ketela, T., Lemieux, S., Veillette, K., Martel, N., Davison, J., Sillaots, S., Trosok, S., Bachewich, C., *et al.* (2007). Genome-wide fitness test and mechanism-of-action studies of inhibitory compounds in *Candida albicans*. *PLoS Pathog* 3, e92.
- Yamamoto, N., Nakahigashi, K., Nakamichi, T., Yoshino, M., Takai, Y., Touda, Y., Furubayashi, A., Kinjyo, S., Dose, H., Hasegawa, M., *et al.* (2009). Update on the Keio collection of *Escherichia coli* single-gene deletion mutants. *Mol Syst Biol* 5, 335.
- Yeh, P., Tschumi, A.I., and Kishony, R. (2006). Functional classification of drugs by properties of their pairwise interactions. *Nat Genet* 38, 489-494.
- Yeh, P.J., Hegreness, M.J., Aiden, A.P., and Kishony, R. (2009). Drug interactions and the evolution of antibiotic resistance. *Nat Rev Microbiol* 7, 460-466.

Acknowledgements

We thank R. Kishony, J. Weissman, A. Hochschild and RG Martin for critically reading this manuscript; J. Hu and P. Thomas for hosting these data on *E. coli* Wiki; C. Raetz for CHIR-090 and M. Gottesman for Bicyclomycin; H. Mori for the Keio Collection; G. Storz for sharing the sRNA deletion library prior to publication; J. Greenblatt and A. Emili for

SPA-tagged alleles; W. Margolin, R. Misra, T. Silhavy and B. Palsson for mutants; T. Baker and R. Sauer for controllable degradation plasmids. This work was supported by NIH R01 GM085697 & ARRA GM085697-01S1 to CAG and NJK; NIH R01 GM036278 to CAG, NIH K99GM092984 to AT, NIH AI060744 to MEW, and NIH GM078338 to SS; NIH F31 DE020206-01 and NIH T32 DE007306 (RJN support); European Molecular Biology Organization long-term fellowship (to AT); and Human Frontier Science Program Long-Term Postdoctoral Fellowship (to PB).

Phenotypic Landscape of a Bacterial Cell

Robert J. Nichols et al.

SUPPLEMENTAL INFORMATION

i. Supplemental Data

Fig. S1 Colony plates illustrate high-density arrayed screening and phenotypic signatures reproduce known connections between biological pathways; related to Fig. 1

Fig. S2 Phenotypes reported here are highly specific and the majority reflect a conditional loss of fitness; related to Fig. 2

Fig. S3 Responses of mutants involved in tetrahydrofolate and one-carbon biosynthesis indicate differential cellular effects of trimethoprim and sulfa drugs; related to Fig. 5

Fig. S4 100kb window size is optimal for avoiding operon effects and weighted responsive genome exhibits highly similar spatial enrichment to the responsive genome; related to Fig. 6

Fig. S5 Drug-Drug correlations overlap with primary target specificity for β -lactams and Drug-GO interactions describe cellular responses to DNA damaging agents; related to Fig. 7

Table S1 List of the 324 screens conducted in this study; related to Fig. 1 (separate Excel file)

Table S2 Complete dataset of fitness scores for 3979 mutants in response to 324 conditions; related to Fig. 1 (separate Excel file)

Table S3 Multi-Stress Responsive genes; related to Fig. 2 (separate Excel file)

Table S4 Conditionally-Essential genes; related to Fig. 2 (separate Excel file)

Table S5 Rich-media Conditionally-essential gene lethal interactions, related to figure 2. (separate Excel file)

Table S6 High Correlation Orphan-Annotated gene pairs; related to Fig. 3 (separate Excel file)

Table S7 Drug-GO interactions at $p < 10^{-3}$; related to Fig. 7 (separate Excel file)

Figure S1.

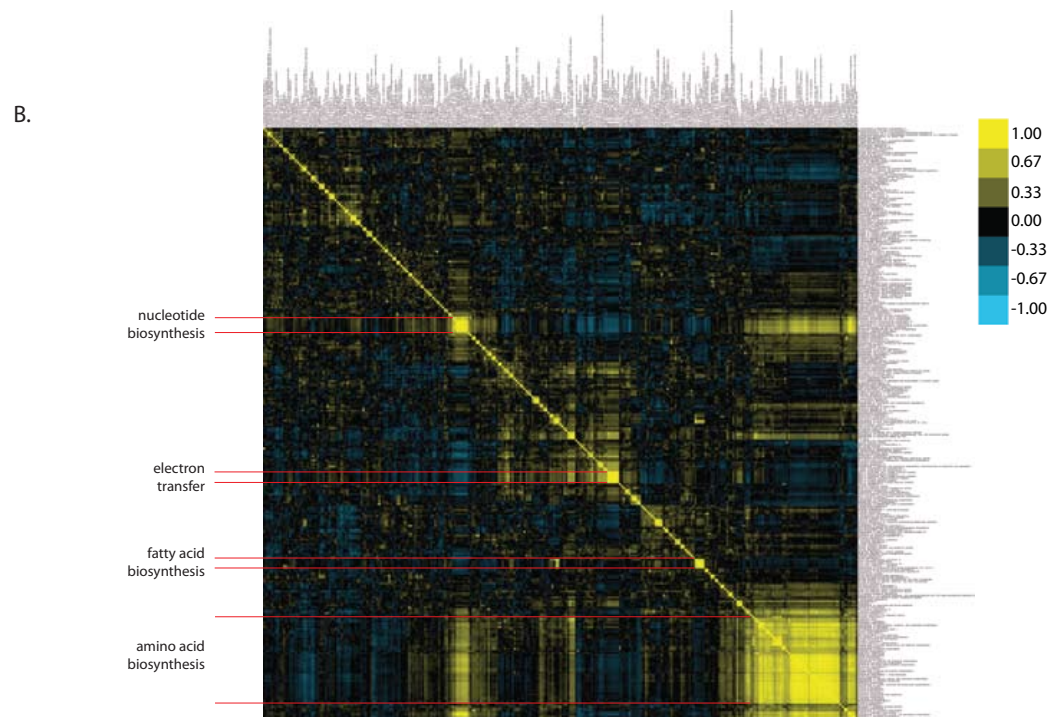
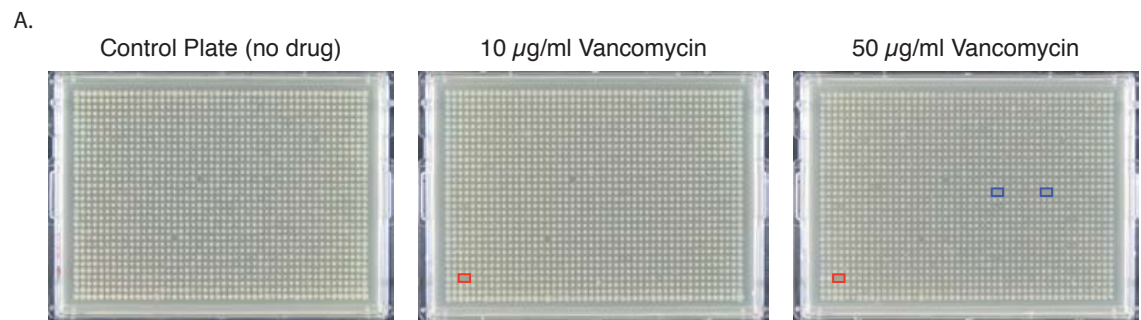
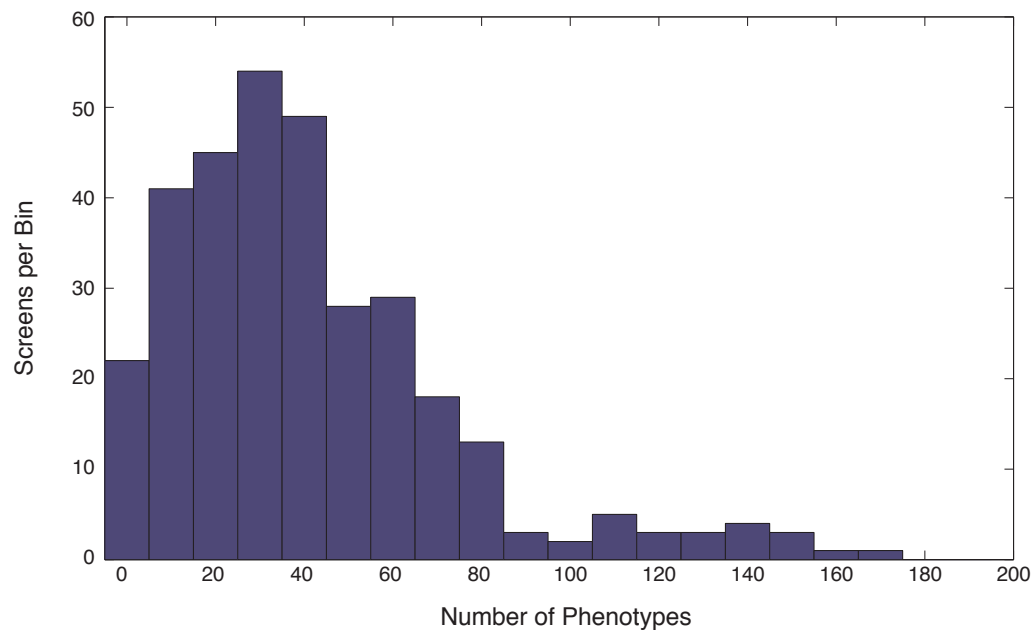


Figure S1. Colony plates illustrate high-density arrayed screening and phenotypic signatures reproduce known connections between biological pathways. (A)

Representative plates show the high-density array format used in our screens (1536 colonies/plate), and demonstrate the high-resolution condition-gene interactions (condition-gene interactions may be detected only under specific concentrations) identified by screening a range of drug concentrations (Vancomycin). **(B)** Phenotypic signatures reproduce known connections between biological pathways. Pathway signatures are generated by averaging the phenotypic signatures of all strains whose mutated gene is assigned to the given pathway in Ecocyc (www.ecocyc.org). Correlation coefficients are plotted for all pairs between 299 pathways. Clusters of known related pathways are highlighted. High off-axis correlation is observed between nucleotide and amino acid biosynthesis pathways, reflecting their shared essentiality under minimal media stresses screened in this study.

Figure S2.

A.



B.

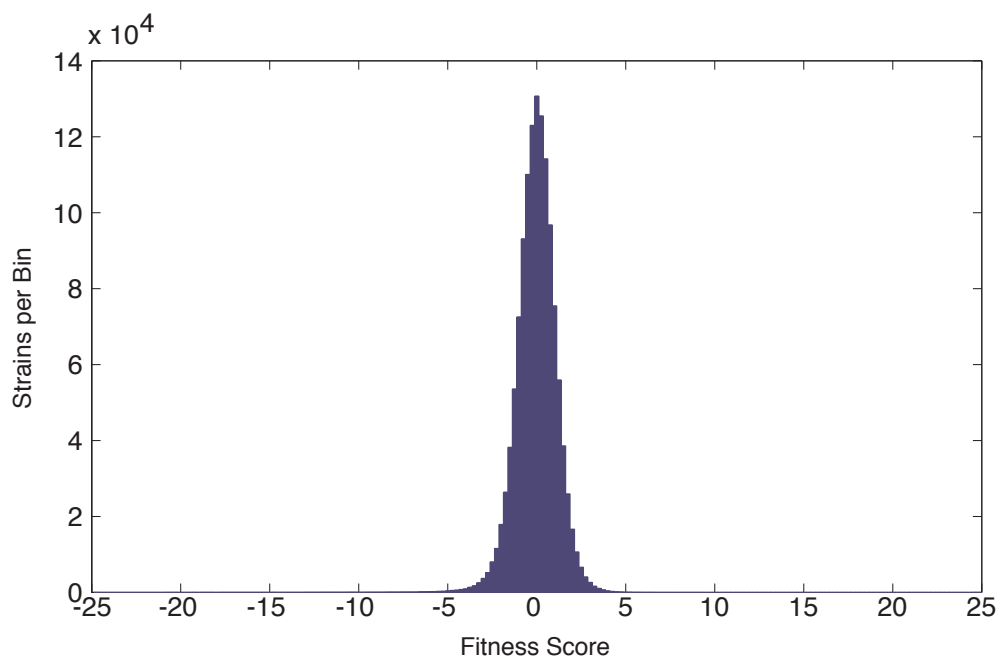
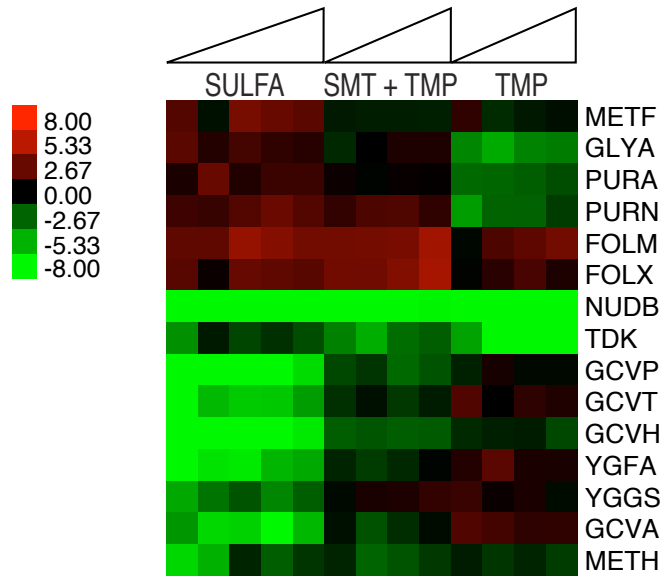


Figure S2. Phenotypes reported here are highly specific and the majority reflect a conditional loss of fitness. (A) Histogram of the number of phenotypes identified at 5% FDR per screen. Bin size = 10 phenotypes. The number of phenotypes identified for any screen ranged from 0-4% of all strains tested. (B) Histogram of all fitness scores contained in the full clustergram in Fig. 1C (3979 strains x 324 conditions). The negative tail is larger than the positive tail, reflecting the ~ 4:1 ratio of negative to positive phenotypes identified in this study.

Figure S3.

A.



B.

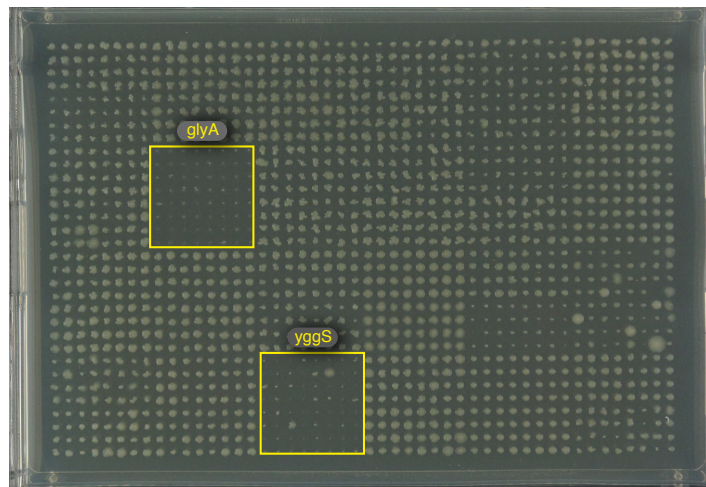


Figure S3. THF and 1C mutant responses suggest differential network effects of Sulfa and TMP. (A) Clustergram of select mutants associated with THF and one-carbon biosynthesis with multiple responses to sulfonamides (Sulfa), trimethoprim (TMP), or the combination. Profiles reveal that in addition to *gcv* and *glyA* mutants, knockouts of several other known and predicted network components exhibit differential responses to Sulfa or TMP. Mutants of *folM* and *folX* exhibit increased resistance to Sulfa, TMP (slightly), and the combination, in agreement with a previously published report (Girgis et al., 2009). Removal of *nudB*, which encodes the enzyme catalyzing the committed step in folate biosynthesis, renders the cell hypersensitive to all folate stresses tested, as expected. A deletion of *ygfA*, like the *gcv* mutants, is sensitive only to Sulfa stress. Interestingly, a recent study demonstrates that 5-formyl-THF accumulates in a *ygfA*⁻ background only in conditions of excess glycine (Jeanguenin et al., 2010), conditions in which the GCV system is presumably active. (B) *glyA*⁻ and *yggS*⁻ are a synthetic lethal pair. Image of a plate mating between the donor Hfr *yggS*::cat and 24 kan recipients (arrayed in boxes of 8x8 colonies), grown on kanamycin/chloramphenicol to select for double mutant strains; position of the *glyA*::kan and *yggS*::kan recipients is highlighted.

Figure S4.

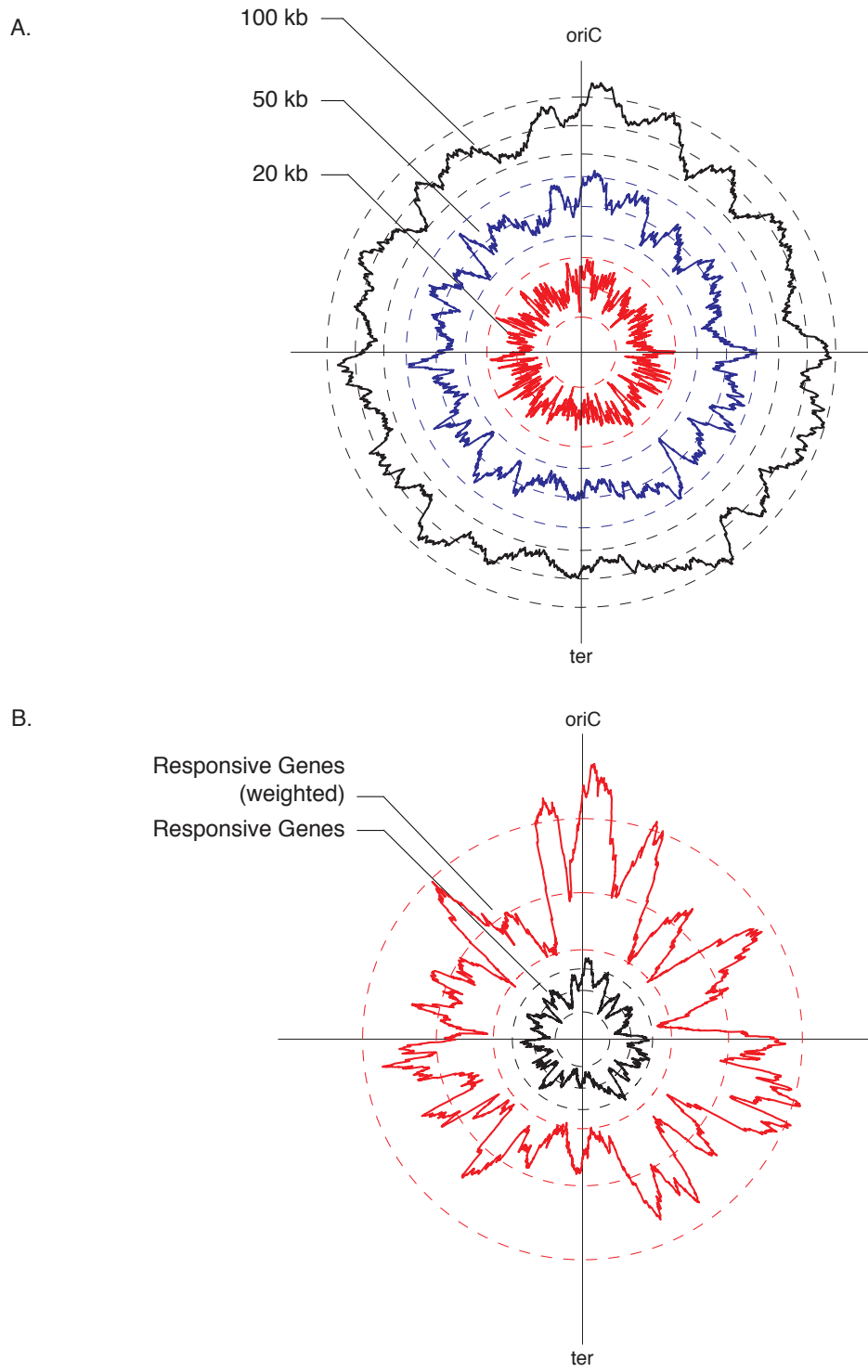
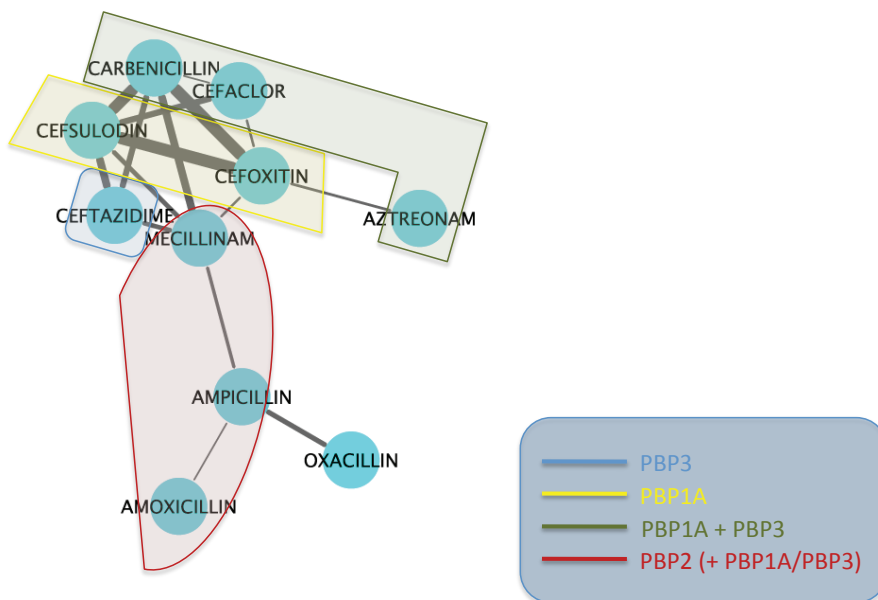


Figure S4. 100kb window size is optimal for avoiding operon effects and weighted responsive genome exhibits highly similar spatial enrichment to the responsive genome. Weighted Responsive Genome reflects Responsive Genome. Circular plots represent spatial enrichment for the variable plotted. Each trace (color) is accompanied by three dashed lines of the same color indicating the 5th percentile of all minimum negative enrichment values stored from 1000 permutations of the data, baseline representing zero enrichment, 95th percentile of all maximum positive enrichment values stored from 1000 permutations of the data (inside to outside). **(A)** Three plots of the responsive genome illustrate the operon effects associated with smaller window sizes, especially the 20kb window (red trace). **(B)** Black shows spatial enrichment for responsive genes, while red shows spatial enrichment for the weighted responsive genes. Here, responsive genes are weighted according to how many phenotypes were identified for each.

Figure S5.

A.



B.

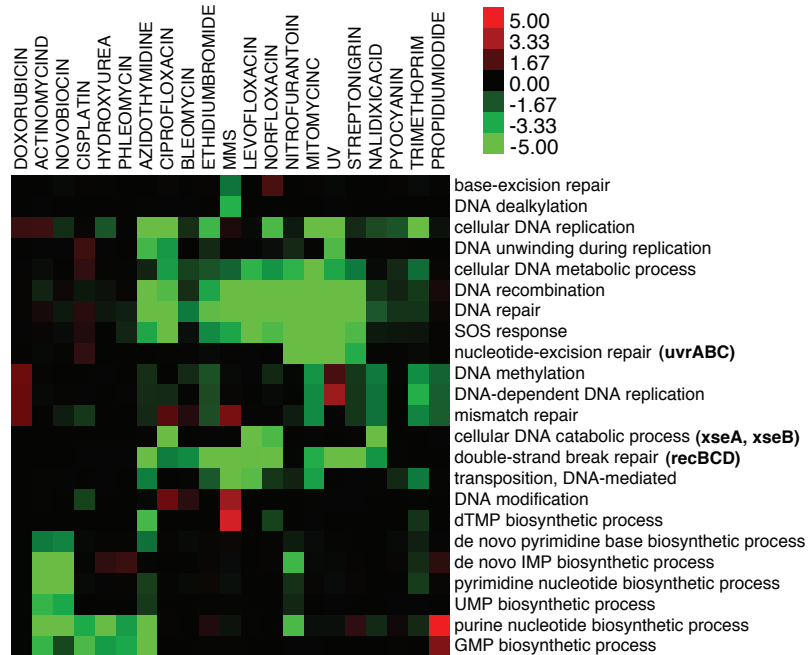


Figure S5. Drug-Drug correlations overlap with primary target specificity for β -lactams and Drug-GO interactions describe cellular responses to DNA damaging agents. (A) Correlation of drug signatures roughly overlaps with PBP target affinity for the β -lactams. Connections between nodes represent positive correlation of drug signatures ($r > 0.16$, or 1 SD). Line width is based on the magnitude of positive correlation. Color shading corresponds to the highest affinity PBP target for each drug, based on IC_{50} data (W. Vollmer, personal communication). (B) Drug-GO clustergram shows interactions between DNA damaging stresses and DNA-related GO biological processes.

Table S1 - Screen List						
Antibiotic/Condition	Supplier	Item No	Target	Concentrations Screened	Screens	
A22	Calbiochem	475951	morphogenesis; MreB	0.5,2,5,15 µg/ml	4	
Acetate (M9)	Fisher Scientific	S210-500	alternative carbon source	0.60%	1	
acidic pH (MES-HOMOPIPES)	Sigma	M2933, 53588	-	4,4.5,5,6 (100mM buffer)	4	
Acriflavine	Sigma	A8126-25G	-	2,10 µg/ml	2	
Actinomycin D	MP	104658-25MG	DNA damage-transcription	2.5,5,10,15 µg/ml	4	
Amikacin	Sigma	A2324-5G	protein synthesis-30S	0.05,0.1,0.2 µg/ml	3	
Amoxicillin	Sigma	A8523-1G	PG (transpeptidases)	0.25,0.5,1,1.5 µg/ml	4	
Ampicillin	Sigma	-	PG (transpeptidases)	1,2,4,8 µg/ml	4	
Anaerobic	-	-	-	-	1	
Azidothymidine	MP	154807-25MG	DNA damage (strand breakage)	0.5,1,2.5 ng/ml	3	
Azithromycin	Sigma	75199-25MG	protein synthesis-50S	0.02,0.1,1 µg/ml	3	
Aztreonam	MP	150415-1G	PG (transpeptidases)	0.02,0.04 µg/ml	2	
Bacitracin	MP	190301-250KU	inhibits sythesis of undecaprenyl phosphate- targets C55 PP pyrophosphatases: BacA, YbjG, PgpB and YeiU	100,200,300 µg/ml	3	
basic pH (TAPS)	Sigma	-	-	8,9,9.5,10 (100mM buffer)	4	
Benzalkonium	Sigma	B6295-100G	Membrane	1,10,25 µg/ml	3	
Bicyclomycin	gift	M. Gottesman	-	1,10 µg/ml	2	
Bile salts	Sigma	B8756-50g	Membrane	0.1,0.5,1,2 %	4	
Bleomycin	Sigma	15361-1MG	multiple (DNA/RNA degradation)	0.1,0.5,1,2 µg/ml	4	
Calcofluor (F3543 Fluorescent Brightener 28)	Sigma	F3543-1G	Biofilm-cellulose production	50 µg/ml	1	

Carbenicillin	Sigma	C1389-250MG	PG (transpeptidases)	0.5,1,1.5 µg/ml	3
CCCP (Carbonyl cyanide 3-chlorophenylhydrazone)	Sigma	C2759-100MG	pmf (proton motive force)	0.1,0.5,2 µg/ml	3
Cecropin B	GenWay	06-271-83151-1MG	membrane	0.1,0.3 µg/ml	2
Cefaclor	MP	198943-100MG	PG (transpeptidases)	1,2,3 µg/ml	3
Cefoxitin	Sigma	C4786-250MG	PG (transpeptidases)	0.25,0.5,0.75,1 µg/ml	4
Cefsulodin	Sigma	C8145-250MG	PG elong. & septation; PBP1a&b	6,12,18,24 µg/ml	4
Cefsulodin + Mecillinam	as single compounds	as single compounds	as single compounds	6 + 0.03 µg/ml	1
Ceftazidime	Sigma	63809-1G	PG (transpeptidases)	0.025,0.05,0.075 µg/ml	3
Cerulenin	MP	195098-10MG	Fatty acid biosynthesis	1,2,4,6 µg/ml	4
CHIR-090	C Raetz	-	LPS; LpxC	0.02,0.025,0.04,0.05,0.075 µg/ml	5
Chloramphenicol	Fisher Scientific	BP904-100	protein synthesis-50S	0.5,1,1.5,2 µg/ml	4
Chlorpromazine	Sigma	C8138-5G	membrane-Rcs	3,6,12,24 µM	4
Cholate	Sigma	270911-25G	Membrane	0.1,0.5,1,2 %	4
Ciprofloxacin	MP	199020_5G	DNA gyrase	0.004,0.006,0.008 µg/ml	3
Cisplatin	Sigma	479306-1G	DNA	20,50,100 µg/ml	3
Clarithromycin	Sigma	C9742-100MG	protein synthesis-50S	0.1,1,5,10 µg/ml	4
Cobalt stress-CoCl ₂	Sigma	C2644-100G	-	0.1,0.5 mM (LB)	2
Cold shock	-	-	-	16,18,20 °C	3
Copper stress-CuCl ₂	Mallinckrodt	4824	-	1,2,4 mM (LB)	3
Cycloserine D	Sigma	C6880-1G	PG racemase + ligase (Alr, DadX, DdlA/B)	16 µg/ml	1
Deoxycholate	Sigma	D6750-100G	Membrane	0.1,0.5,2 %	3
Dibucaine	Sigma	D0513-	membrane-pmf	0.4,0.8,1.2 mM	3

Doxorubicin	Sigma	5G D1515-10MG	binds ArcB in-vitro	1,10 µg/ml	2
Doxycycline	Sigma	D9891-1G	protein synthesis-	0.25,0.5,0.75,1 µg/ml	4
EDTA	Sigma	E5134-500G	Membrane-OM	0.1, 0.5, 1 mM	3
EGTA	Sigma	E4378-500G	Membrane-OM	0.1, 0.5, 1, 2 mM	4
Epigallocatechin gallate (EGCG)	Sigma	E4143-50MG	Fatty acid biosynthesis	5,20,50 µM	3
Epinephrine	Sigma	E4250-1G	QseC	50,250,1000 µM	3
Erythromycin	Sigma	E5389-1G	protein synthesis-	0.1,1,5,10 µg/ml	4
Ethidium Bromide	Sigma	lab stock	DNA	2,10,50 µg/ml	3
EtOH	Rossvill e	200 proof	protein folding	2,4,6 %	3
Fosfomycin	Sigma	P5396-1G	PG; MurA	1 µg/ml	1
Fosfomycin +Glucose 6P	as single compounds	as single compounds	as single compounds	0.05,0.2 µg/ml (in 50 mg/ml G6P)	2
Fusidic acid	Sigma	F0881-1G	protein synthesis-G factor	1,5,20,50 µg/ml	4
Gentamicin	Sigma	G3632	protein synthesis-	0.05,0.1 µg/ml	2
Glucosamine (M9)	Sigma	G4875	alternative carbon source	0.20%	1
Glucose (M9)	Fisher Scientific	D16-1	alternative carbon source	0.20%	1
Glycerol (M9)	Fisher Scientific	BP229-4	alternative carbon source	0.40%	1
Heat shock	-	-	-	40,42,43.5,45 °C	4
Hydrogen peroxide	Sigma	-	oxidative stress	0.1,0.5,1,2 mM	4
Hydroxyurea	Sigma	H8627-5G	DNA damage	1,5,10 mM	3
Indolicidin	US Biological	I7552-1MG	LPS; membranes	0.1 µg/ml	1
Iron excess-FeSO4	Mallinckrodt	5572	-	1mM (MOPS; normal 100µM)	1
Iron starvation-FeSO4	Mallinckrodt	5572	-	2µM (MOPS; normal 100µM)	1
Isoniazid	Sigma	I3377-	lhnA-mycolic acid	0.2,1,1.5 mM	3

Levofloxacin	Sigma	5G 28266-1G-F	biosynthesis in TB DNA gyrase	0.002 µg/ml	1
Maltose (M9)	Fisher Scientific	BP684-500	alternative carbon source	0.10%	1
Mecillinam	Sigma	33447-100MG	PG elongation; PBP2	0.03,0.06,0.09,0.12 µg/ml	4
Methotrexate	Sigma	M9929-25MG	folic acid biosynthesis (DHFR)	1,25 µg/ml	2
Minocycline	Sigma	M9511-100MG	protein synthesis- 16S; binds ArcB in-vitro	0.2,0.5,1 µg/ml	3
Mitomycin C	Sigma	M0440-5MG	RNApolymerase/DN Areplication	0.1 µg/ml	1
MMS	Sigma	129925	DNA damage (methylation)	0.05%	1
N-acetyl Glucosamine	Sigma	A8625	alternative carbon source	0.15%	1
NaCl	Sigma	S7653	-	150,300,450,600 mM	4
Nalidixic acid	Sigma	N4382	DNA gyrase	0.5,1,1.5,2 µg/ml	4
NH4Cl (MOPS)	Sigma	A0171-100g	alternative nitrogen source	9.5mm	1
Nickel stress- NiCl2	Sigma	N5756-100g	-	0.1,1 mM (LB)	2
Nigericin	Sigma	N7143-5MG	pmf (proton motive force)	0.1,1,5 µM	3
Nitrofurantoin	Sigma	N7878-10G	multiple (DNA, Krebs cycle)	0.1,0.5,1,1.5,2 µg/ml	5
Norepinephrine	Sigma	A7257-500MG	QseC	100,1000 µM	2
Norfloxacin	Sigma	N9890-1G	DNA gyrase	0.01,0.02,0.04 µg/ml	3
Novobiocin	Sigma	N1628-1G	DNA gyrase	4,6,8,10,12,30 µg/ml	6
Oxacillin	Sigma	28221-1G	PG (transpeptidases)	0.5,5,40 µg/ml	3
Paraquat dichloride	Sigma	36541-100MG	oxidative stress	0.2,1,5,10,18 µM	5
Phenazine methosulfate (PMS)	Sigma	P9625-1G	DNA synthesis- membrane	0.02,0.05,0.1 mM	3
Phleomycin	Sigma	P9564-5MG	DNA damage	0.2,0.5,1 µg/ml	3
Polymyxin B	Sigma	P1004	LPS	1,2,4,6 µg/ml	4
Procaine	Sigma	P9879	membrane- EnvZ/OmpR	1,5,10,30 mM	4
Propidium iodide	Sigma	P4170-100MG	DNA/RNA	1,20,50 µg/ml	3
Puromycin	Sigma	P7255-100MG	protein synthesis inhibitor	1,5,25 µg/ml	3

Pyocyanin	Cayman Chemicals	100095-94-50MG	SoxR? Superoxide stress	0.2,1,10 µg/ml	3
Radicicol	A&G Scientific	R-1130	HtpG, PhoQ	1,5,10 µM	3
Rifampicin	Sigma	R3501-5G	RNA polymerase	1,2 µg/ml	2
SDS	MP Biomedicals	811030	Membrane	0.5,1,2,3,4 %	5
SDS+EDTA	as single compounds	as single compounds	membrane	0.5%/0.1mM,0.5%/0.5mM,1%/0.5mM	3
Spectinomycin	Sigma	S9007-25G	protein synthesis-30S	4,6 µg/ml	2
Spiramycin	Sigma	S9132-1G	protein synthesis-50S	1,5,20 µg/ml	3
Streptomycin	Sigma	S6501	protein synthesis-30S	0.05 µg/ml	1
Streptonigrin	Sigma	S1014-5MG	DNA metabolism-respiration	0.1,0.2,0.5 µg/ml	3
Succinate (M9)	Fisher Scientific	A294-500	alternative carbon source	0.30%	1
Sulfamethizole	Sigma	S5632-10G	folic acid biosynthesis	100,200,300 µg/ml	3
Sulfamonomethoxine	Sigma	S0508-250MG	folic acid biosynthesis	50,100 µg/ml	2
Taurocholate	Sigma	861960-5G	Membrane	0.1,0.5,1 %	3
Tetracycline	Sigma	T3383-25G	protein synthesis-30S	0.25,0.5,0.75,1 µg/ml	4
Theophylline	Sigma	T1633-50G		- 10,100 µg/ml	2
Thiolactomycin	Sigma	T9567-10MG	Fatty acid biosynthesis	1,5,50 µM	3
Tobramycin	Sigma	T4014-100MG	protein synthesis-30S	0.05,0.1,0.2,0.4 µg/ml	4
Triclosan/Irgasan	Sigma	72779-5G-F	Fatty acid biosynthesis	0.05 µg/ml	1
Trimethoprim	Sigma	T7883-25G	folic acid biosynthesis (DHFR)	0.1,0.2,0.3,0.4 µg/ml	4
Trimethoprim + Sulfamethizole	as single compounds	as single compounds	as single compounds	0.1 + 50 µg/ml	1
Triton X-100	Fisher Scientific	BP151-500	Membrane	0.01,0.03,0.2 %	3

		^c			
Tunicamycin	Sigma	T7765-5MG	PG; MraY	1,3,7.5 µg/ml	3
	UV	-	DNA damage	6,12,18,24 sec	4
Vancomycin	Sigma	861987-1G	PG elongation	10,20,50 µg/ml	3
Verapamil	MP	195545-1G	cell division-pmf	0.1,0.5,1 mM	3

Total Screens 324

DUE TO THE LARGE DIMENSIONS OF TABLE S2 (324 X 3979), IT WILL NOT APPEAR IN THIS VOLUME. HOWEVER, IT CAN BE ACCESSED VIA THE ORIGINAL CELL PUBLICATION (NICHOLS ET. AL, CELL 2011), OR ONLINE AT <http://ecoliwiki.net/tools/chemgen/>.

**Table S3 - Multi-Stress Responsive
Genes**

ECK0095-B-FTSA(R286W)-KAN
ECK2042-CPSG
ECK0726-TOLQ
ECK0727-TOLR
ECK0729-TOLB
ECK0730-PAL
ECK0055-D-IMP4213
ECK0097-B-LPXC(G210S)-KAN
ECK0180-LPXA
ECK0223-LPCA
ECK2058-ASMA
ECK3042-RFAE
ECK3609-RFAD
ECK3610-RFAF
ECK3611-RFAC
ECK3620-RFAP
ECK3621-RFAG
ECK3622-RFAQ
ECK0054-SURA
ECK0176-B-YAET(218-9DUPL)-KAN
ECK0176-C-YAET(R64DEL)-KAN
ECK2508-BAMB
ECK2613-SMPA
ECK1091-YCFM
ECK0148-MRCB
ECK3603-ENVC
ECK3722-GLMS
ECK3374-DAM
ECK0215-DNAQ
ECK3725-ATPD
ECK3727-ATPA
ECK3729-ATPF
ECK3730-ATPE
ECK3731-ATPB
ECK0721-CYDA
ECK1080-ACPP
ECK3597-CYSE
ECK4133-ASPA
ECK2323-AROC
ECK3048-FOLB
ECK4227-FBP
ECK0114-ACEF
ECK0621-LIPA
ECK0623-LIPB
ECK2758-CYSI

ECK0654-UBIF
ECK0053-PDXA
ECK0898-SERC
ECK1634-PDXH
ECK0515-PURK
ECK0516-PURE
ECK2495-PURM
ECK2555-PURL
ECK3997-PURD
ECK3998-PURH
ECK4173-PURA
ECK2503-GUAA
ECK2504-GUAB
ECK0034-CARB
ECK0936-PYRD
ECK1047-PYRC
ECK3632-PYRE
ECK3356-CYSG
ECK3247-DUSB
ECK0015-DNAJ
ECK3852-DSBA
ECK0433-LON
ECK3904-CPXA
ECK4168-HFQ
ECK0671-FUR
ECK1523-MARR
ECK3248-FIS
ECK3953-OXYR
ECK1269-CYSB
ECK3218-SSPA
ECK3975-RPLA
ECK3156-RBFA
ECK0581-FEPC
ECK0583-FEPD
ECK0456-ACRB
ECK0457-ACRA
ECK3026-TOLC
ECK3831-TATB
ECK3832-TATC
ECK0539-NINE
ECK4401-YBHU
ECK0723-YBGT
ECK1228-YCHJ
ECK1274-YCIS
ECK2490-YFGC
ECK2614-YFJF
ECK3222-YHCB
ECK0542-YLCG
ECK3033-YQIC

**Table S4 -
Conditionally-
Essential Genes**

Rich-Media CE

ECK0034-CARB
ECK0050-APAH
ECK0054-SURA
ECK0093-DDLB
ECK0122-CUEO
ECK0144-DKSA
ECK0148-MRCB
ECK0160-DEGP
ECK0223-LPCA
ECK0456-ACRB
ECK0457-ACRA
ECK0458-ACRR
ECK0478-COPA
ECK0515-PURK
ECK0516-PURE
ECK0581-FEPC
ECK0582-FEPG
ECK0583-FEPD
ECK0585-FEPB
ECK0621-LIPA
ECK0654-UBIF
ECK0665-NAGA
ECK0722-CYDB
ECK0726-TOLQ
ECK0727-TOLR
ECK0729-TOLB
ECK0730-PAL
ECK0832-CMR
ECK0903-IHFB
ECK0936-PYRD
ECK1047-PYRC
ECK1077-FABH
ECK1091-YCFM
ECK1243-CLS
ECK1248-YCIB
ECK1260-YCIV
ECK1269-CYSB
ECK1275-YCIM
ECK1403-YNBC
ECK1544-GNSB
ECK1673-LPP
ECK1780-MIPA
ECK1862-RUVA

Auxotrophs

ECK0002-THRA
ECK0003-THRB
ECK0004-THRC
ECK0033-CARA
ECK0034-CARB
ECK0053-PDXA
ECK0073-LEUD
ECK0074-LEUC
ECK0075-LEUB
ECK0076-LEUA
ECK0108-NADC
ECK0215-DNAQ
ECK0243-PROB
ECK0244-PROA
ECK0381-PROC
ECK0515-PURK
ECK0516-PURE
ECK0581-FEPC
ECK0665-NAGA
ECK0666-NAGB
ECK0709-GLTA
ECK0710-SDHC
ECK0716-SUCC
ECK0717-SUCD
ECK0739-NADA
ECK0898-SERC
ECK0899-AROA
ECK0903-IHFB
ECK0936-PYRD
ECK1047-PYRC
ECK1048-YCEB
ECK1254-TRPA
ECK1255-TRPB
ECK1269-CYSB
ECK1634-PDXH
ECK1710-IHFA
ECK2014-HISG
ECK2015-HISD
ECK2016-HISC
ECK2017-HISB
ECK2018-HISH
ECK2019-HISA
ECK2020-HISF

ECK1864-RUVC	ECK2021-HISI
ECK2228-YFAE	ECK2306-PURF
ECK2262-RBN	ECK2323-AROC
ECK2271-NUOM	ECK2411-PTSI
ECK2273-NUOK	ECK2472-PURC
ECK2274-NUOJ	ECK2495-PURM
ECK2276-NUOH	ECK2503-GUAA
ECK2279-NUOE	ECK2504-GUAB
ECK2282-NUOA	ECK2548-GLYA
ECK2305-UBIX	ECK2555-PURL
ECK2314-PDXB	ECK2562-PDXJ
ECK2321-YFCA	ECK2572-NADB
ECK2323-AROC	ECK2596-PHEA
ECK2340-VACJ	ECK2597-TYRA
ECK2472-PURC	ECK2745-CYSC
ECK2490-YFGC	ECK2747-CYSD
ECK2495-PURM	ECK2758-CYSI
ECK2503-GUAA	ECK2759-CYSJ
ECK2504-GUAB	ECK2814-ARGA
ECK2508-BAMB	ECK2836-LYSA
ECK2555-PURL	ECK2837-LYSR
ECK2613-SMPA	ECK2909-SERA
ECK2694-RECA	ECK3000-METC
ECK2737-NLPD	ECK3048-FOLB
ECK2818-RECC	ECK3161-ARGG
ECK2901-VISC	ECK3268-AROE
ECK2903-PEPP	ECK3356-CYSG
ECK2998-EXBD	ECK3376-AROB
ECK3026-TOLC	ECK3403-MALQ
ECK3033-YQIC	ECK3405-MALT
ECK3042-RFAE	ECK3412-GLPD
ECK3048-FOLB	ECK3597-CYSE
ECK3138-YRAP	ECK3632-PYRE
ECK3156-RBFA	ECK3762-ILVE
ECK3164-SECG	ECK3763-ILVD
ECK3222-YHCB	ECK3764-ILVA
ECK3234-YHDP	ECK3766-ILVC
ECK3248-FIS	ECK3822-METR
ECK3409-GLPR	ECK3823-METE
ECK3597-CYSE	ECK3863-GLNA
ECK3602-GPMM	ECK3918-GLPK
ECK3603-ENVC	ECK3931-METB
ECK3609-RFAD	ECK3933-METF
ECK3610-RFAF	ECK3947-PPC
ECK3611-RFAC	ECK3948-ARGE
ECK3618-RFAB	ECK3949-ARGC
ECK3620-RFAP	ECK3950-ARGB
ECK3621-RFAG	ECK3951-ARGH
ECK3622-RFAQ	ECK3997-PURD

ECK3632-PYRE	ECK3998-PURH
ECK3699-MNME	ECK4005-META
ECK3713-BGLH	ECK4007-ACEA
ECK3725-ATPD	ECK4025-MALF
ECK3729-ATPF	ECK4027-MALK
ECK3730-ATPE	ECK4173-PURA
ECK3731-ATPB	ECK4210-CYSQ
ECK3734-MNMG	ECK4227-FBP
ECK3756-HDFR	ECK4240-PYRB
ECK3831-TATB	ECK4380-SERB
ECK3832-TATC	
ECK3836-FRE	
ECK3904-CPXA	
ECK3916-FPR	
ECK3997-PURD	
ECK3998-PURH	
ECK4054-SOXS	
ECK4055-SOXR	
ECK4168-HFQ	
ECK4173-PURA	
ECK4227-FBP	
ECK4393-ARCA	
ECK4419-YGDT	
ECK5005-TP2	

Table S5. Rich-media CE Lethal Events

Strain	Lethal Condition
ECK2503-GUAA	A22-0.5
ECK2504-GUAB	A22-0.5
ECK2504-GUAB	A22-15.0
ECK0148-MRCB	A22-2.0
ECK2504-GUAB	A22-2.0
ECK0148-MRCB	A22-5.0
ECK1047-PYRC	A22-5.0
ECK2504-GUAB	A22-5.0
ECK3632-PYRE	A22-5.0
ECK0034-CARB	ACETATE
ECK0515-PURK	ACETATE
ECK0516-PURE	ACETATE
ECK0581-FEPC	ACETATE
ECK0903-IHFB	ACETATE
ECK1269-CYSB	ACETATE
ECK2323-AROC	ACETATE
ECK2472-PURC	ACETATE
ECK2495-PURM	ACETATE
ECK2503-GUAA	ACETATE
ECK2504-GUAB	ACETATE
ECK2555-PURL	ACETATE
ECK3632-PYRE	ACETATE
ECK3997-PURD	ACETATE
ECK3998-PURH	ACETATE
ECK4227-FBP	ACETATE
ECK0456-ACRB	ACRIFLAVINE-10
ECK2271-NUOM	ACRIFLAVINE-10
ECK2274-NUOJ	ACRIFLAVINE-10
ECK2276-NUOH	ACRIFLAVINE-10
ECK2279-NUOE	ACRIFLAVINE-10
ECK3026-TOLC	ACRIFLAVINE-10
ECK4168-HFQ	ACRIFLAVINE-10
ECK0621-LIPA	ACRIFLAVINE-2
ECK0729-TOLB	ACTINOMYCIND-10.0
ECK2504-GUAB	ACTINOMYCIND-10.0
ECK0729-TOLB	ACTINOMYCIND-15.0
ECK2504-GUAB	ACTINOMYCIND-15.0
ECK3609-RFAD	ACTINOMYCIND-15.0
ECK2504-GUAB	ACTINOMYCIND-2.5
ECK2504-GUAB	ACTINOMYCIND-5.0
ECK3597-CYSE	AMOXICILLIN-0.25
ECK3597-CYSE	AMOXICILLIN-1.0
ECK0730-PAL	AMPICILLIN-8.0
ECK2818-RECC	AZIDOTHYIMIDINE-0.5
ECK0621-LIPA	AZITHROMYCIN-0.1
ECK0223-LPCA	AZITHROMYCIN-1.0
ECK2490-YFGC	AZITHROMYCIN-1.0

ECK3042-RFAE	AZITHROMYCIN-1.0
ECK3609-RFAD	AZITHROMYCIN-1.0
ECK3621-RFAG	AZITHROMYCIN-1.0
ECK0054-SURA	BACITRACIN-300
ECK0726-TOLQ	BACITRACIN-300
ECK0729-TOLB	BACITRACIN-300
ECK0730-PAL	BACITRACIN-300
ECK1077-FABH	BACITRACIN-300
ECK2508-BAMB	BACITRACIN-300
ECK0223-LPCA	BENZALKONIUM-10
ECK0456-ACRB	BENZALKONIUM-10
ECK0457-ACRA	BENZALKONIUM-10
ECK3026-TOLC	BENZALKONIUM-10
ECK3042-RFAE	BENZALKONIUM-10
ECK3610-RFAF	BENZALKONIUM-10
ECK3620-RFAP	BENZALKONIUM-10
ECK4168-HFQ	BENZALKONIUM-10
ECK0054-SURA	BENZALKONIUM-25
ECK0223-LPCA	BENZALKONIUM-25
ECK0456-ACRB	BENZALKONIUM-25
ECK0457-ACRA	BENZALKONIUM-25
ECK0832-CMR	BENZALKONIUM-25
ECK1077-FABH	BENZALKONIUM-25
ECK1243-CLS	BENZALKONIUM-25
ECK2901-VISC	BENZALKONIUM-25
ECK3026-TOLC	BENZALKONIUM-25
ECK3033-YQIC	BENZALKONIUM-25
ECK3042-RFAE	BENZALKONIUM-25
ECK3222-YHCB	BENZALKONIUM-25
ECK3610-RFAF	BENZALKONIUM-25
ECK3620-RFAP	BENZALKONIUM-25
ECK3621-RFAG	BENZALKONIUM-25
ECK4168-HFQ	BENZALKONIUM-25
ECK2504-GUAB	BILE-0.1%
ECK3026-TOLC	BILE-0.1%
ECK0223-LPCA	BILE-0.5%
ECK0727-TOLR	BILE-0.5%
ECK0729-TOLB	BILE-0.5%
ECK0730-PAL	BILE-0.5%
ECK3026-TOLC	BILE-0.5%
ECK3042-RFAE	BILE-0.5%
ECK0223-LPCA	BILE-1.0%
ECK0726-TOLQ	BILE-1.0%
ECK0727-TOLR	BILE-1.0%
ECK0729-TOLB	BILE-1.0%
ECK0730-PAL	BILE-1.0%
ECK3026-TOLC	BILE-1.0%
ECK3042-RFAE	BILE-1.0%
ECK0223-LPCA	BILE-2.0%

ECK0457-ACRA	BILE-2.0%
ECK0729-TOLB	BILE-2.0%
ECK0730-PAL	BILE-2.0%
ECK2504-GUAB	BILE-2.0%
ECK3026-TOLC	BILE-2.0%
ECK3042-RFAE	BILE-2.0%
ECK3609-RFAD	BILE-2.0%
ECK3610-RFAF	BILE-2.0%
ECK3042-RFAE	BLEOMYCIN-1.0
ECK4168-HFQ	BLEOMYCIN-1.0
ECK4168-HFQ	BLEOMYCIN-2.0
ECK0729-TOLB	CARBENICILLIN-0.5
ECK2321-YFCA	CARBENICILLIN-0.5
ECK2504-GUAB	CARBENICILLIN-0.5
ECK3632-PYRE	CARBENICILLIN-0.5
ECK0148-MRCB	CARBENICILLIN-1.0
ECK0726-TOLQ	CARBENICILLIN-1.0
ECK0729-TOLB	CARBENICILLIN-1.0
ECK1091-YCFM	CARBENICILLIN-1.0
ECK2504-GUAB	CARBENICILLIN-1.0
ECK0729-TOLB	CARBENICILLIN-1.5
ECK1047-PYRC	CARBENICILLIN-1.5
ECK2495-PURM	CARBENICILLIN-1.5
ECK2504-GUAB	CARBENICILLIN-1.5
ECK3632-PYRE	CARBENICILLIN-1.5
ECK3026-TOLC	CCCP-2.0
ECK0148-MRCB	CEFOXITIN-0.25
ECK0516-PURE	CEFOXITIN-0.25
ECK1047-PYRC	CEFOXITIN-0.25
ECK1091-YCFM	CEFOXITIN-0.25
ECK2495-PURM	CEFOXITIN-0.25
ECK2504-GUAB	CEFOXITIN-0.25
ECK3632-PYRE	CEFOXITIN-0.25
ECK3997-PURD	CEFOXITIN-0.25
ECK0148-MRCB	CEFOXITIN-0.5
ECK1091-YCFM	CEFOXITIN-0.5
ECK2504-GUAB	CEFOXITIN-0.5
ECK0148-MRCB	CEFOXITIN-0.75
ECK1091-YCFM	CEFOXITIN-0.75
ECK2504-GUAB	CEFOXITIN-0.75
ECK0148-MRCB	CEFOXITIN-1.0
ECK1091-YCFM	CEFOXITIN-1.0
ECK2504-GUAB	CEFOXITIN-1.0
ECK0148-MRCB	CEFSULODIN-12.0
ECK1091-YCFM	CEFSULODIN-12.0
ECK2504-GUAB	CEFSULODIN-12.0
ECK0148-MRCB	CEFSULODIN-18.0
ECK0726-TOLQ	CEFSULODIN-18.0
ECK0729-TOLB	CEFSULODIN-18.0

ECK1091-YCFM	CEFSULODIN-18.0
ECK2504-GUAB	CEFSULODIN-18.0
ECK3156-RBFA	CEFSULODIN-18.0
ECK3603-ENVC	CEFSULODIN-18.0
ECK3632-PYRE	CEFSULODIN-18.0
ECK0148-MRCB	CEFSULODIN-24.0
ECK0729-TOLB	CEFSULODIN-24.0
ECK0730-PAL	CEFSULODIN-24.0
ECK1091-YCFM	CEFSULODIN-24.0
ECK3603-ENVC	CEFSULODIN-24.0
ECK3620-RFAP	CEFSULODIN-24.0
ECK0148-MRCB	CEFSULODIN-6.0
ECK1091-YCFM	CEFSULODIN-6.0
ECK2504-GUAB	CEFSULODIN-6.0
ECK0726-TOLQ	CEFSULODIN-6.0,MECILLINAM-0.03
ECK0727-TOLR	CEFSULODIN-6.0,MECILLINAM-0.03
ECK1091-YCFM	CEFSULODIN-6.0,MECILLINAM-0.03
ECK0148-MRCB	CEFTAZIDIME-0.05
ECK1269-CYSB	CEFTAZIDIME-0.05
ECK0148-MRCB	CEFTAZIDIME-0.075
ECK4168-HFQ	CERULENIN-6.0
ECK3026-TOLC	CHIR090-0.02
ECK2504-GUAB	CHIR090-0.025
ECK0054-SURA	CHIR090-0.04
ECK0456-ACRB	CHIR090-0.04
ECK3026-TOLC	CHIR090-0.04
ECK3042-RFAE	CHIR090-0.04
ECK3156-RBFA	CHIR090-0.04
ECK0223-LPCA	CHIR090-0.05
ECK3026-TOLC	CHIR090-0.05
ECK3156-RBFA	CHIR090-0.05
ECK0223-LPCA	CHIR090-0.075
ECK2321-YFCA	CHIR090-0.075
ECK3610-RFAF	CHIR090-0.075
ECK2504-GUAB	CHLORAMPHENICOL-1.0
ECK0456-ACRB	CHLORAMPHENICOL-1.5
ECK0457-ACRA	CHLORAMPHENICOL-1.5
ECK3026-TOLC	CHLORAMPHENICOL-1.5
ECK0456-ACRB	CHLORAMPHENICOL-2.0
ECK0457-ACRA	CHLORAMPHENICOL-2.0
ECK3026-TOLC	CHLORAMPHENICOL-2.0
ECK0457-ACRA	CHLOROPROMAZINE-24
ECK1544-GNSB	CHLOROPROMAZINE-24
ECK2340-VACJ	CHLOROPROMAZINE-24
ECK2613-SMPA	CHLOROPROMAZINE-24
ECK3156-RBFA	CHLOROPROMAZINE-24
ECK0582-FEPG	CHOLATE-0.1%
ECK2472-PURC	CHOLATE-0.1%
ECK2503-GUAA	CHOLATE-0.1%

ECK2504-GUAB	CHOLATE-0.1%
ECK0034-CARB	CHOLATE-0.5%
ECK0054-SURA	CHOLATE-0.5%
ECK0936-PYRD	CHOLATE-0.5%
ECK1047-PYRC	CHOLATE-0.5%
ECK2504-GUAB	CHOLATE-0.5%
ECK2555-PURL	CHOLATE-0.5%
ECK3026-TOLC	CHOLATE-0.5%
ECK3632-PYRE	CHOLATE-0.5%
ECK3997-PURD	CHOLATE-0.5%
ECK0457-ACRA	CHOLATE-1.0%
ECK0729-TOLB	CHOLATE-1.0%
ECK0730-PAL	CHOLATE-1.0%
ECK2495-PURM	CHOLATE-1.0%
ECK2504-GUAB	CHOLATE-1.0%
ECK2555-PURL	CHOLATE-1.0%
ECK3026-TOLC	CHOLATE-1.0%
ECK3997-PURD	CHOLATE-1.0%
ECK0034-CARB	CHOLATE-2.0%
ECK0457-ACRA	CHOLATE-2.0%
ECK0729-TOLB	CHOLATE-2.0%
ECK0730-PAL	CHOLATE-2.0%
ECK1673-LPP	CHOLATE-2.0%
ECK2504-GUAB	CHOLATE-2.0%
ECK3026-TOLC	CHOLATE-2.0%
ECK3042-RFAE	CHOLATE-2.0%
ECK3620-RFAP	CHOLATE-2.0%
ECK3997-PURD	CHOLATE-2.0%
ECK2818-RECC	CIPROFLOXACIN-0.004
ECK1403-YNBC	CIPROFLOXACIN-0.006
ECK2694-RECA	CIPROFLOXACIN-0.006
ECK3248-FIS	CIPROFLOXACIN-0.008
ECK1269-CYSB	CISPLATIN-100
ECK1269-CYSB	CISPLATIN-20
ECK3597-CYSE	CISPLATIN-20
ECK1269-CYSB	CISPLATIN-50
ECK3597-CYSE	CISPLATIN-50
ECK0223-LPCA	CLARYTHROMYCIN-10.0
ECK3026-TOLC	CLARYTHROMYCIN-10.0
ECK3042-RFAE	CLARYTHROMYCIN-10.0
ECK3026-TOLC	CLARYTHROMYCIN-5.0
ECK3042-RFAE	CLARYTHROMYCIN-5.0
ECK0223-LPCA	DEOXYCHOLATE-0.1%
ECK0729-TOLB	DEOXYCHOLATE-0.1%
ECK0730-PAL	DEOXYCHOLATE-0.1%
ECK3026-TOLC	DEOXYCHOLATE-0.1%
ECK3042-RFAE	DEOXYCHOLATE-0.1%
ECK0223-LPCA	DEOXYCHOLATE-0.5%
ECK0726-TOLQ	DEOXYCHOLATE-0.5%

ECK0727-TOLR	DEOXYCHOLATE-0.5%
ECK0729-TOLB	DEOXYCHOLATE-0.5%
ECK0730-PAL	DEOXYCHOLATE-0.5%
ECK3026-TOLC	DEOXYCHOLATE-0.5%
ECK3042-RFAE	DEOXYCHOLATE-0.5%
ECK0223-LPCA	DEOXYCHOLATE-2.0%
ECK0457-ACRA	DEOXYCHOLATE-2.0%
ECK0726-TOLQ	DEOXYCHOLATE-2.0%
ECK0727-TOLR	DEOXYCHOLATE-2.0%
ECK0729-TOLB	DEOXYCHOLATE-2.0%
ECK0730-PAL	DEOXYCHOLATE-2.0%
ECK1275-YCIM	DEOXYCHOLATE-2.0%
ECK1673-LPP	DEOXYCHOLATE-2.0%
ECK3026-TOLC	DEOXYCHOLATE-2.0%
ECK0223-LPCA	DIBUCAINE-0.4
ECK0456-ACRB	DIBUCAINE-0.4
ECK0457-ACRA	DIBUCAINE-0.4
ECK3026-TOLC	DIBUCAINE-0.4
ECK3042-RFAE	DIBUCAINE-0.4
ECK3621-RFAG	DIBUCAINE-0.4
ECK0054-SURA	DIBUCAINE-0.8
ECK0223-LPCA	DIBUCAINE-0.8
ECK0456-ACRB	DIBUCAINE-0.8
ECK0730-PAL	DIBUCAINE-0.8
ECK2503-GUAA	DIBUCAINE-0.8
ECK3026-TOLC	DIBUCAINE-0.8
ECK3042-RFAE	DIBUCAINE-0.8
ECK3610-RFAF	DIBUCAINE-0.8
ECK3620-RFAP	DIBUCAINE-0.8
ECK3621-RFAG	DIBUCAINE-0.8
ECK0054-SURA	DIBUCAINE-1.2
ECK0223-LPCA	DIBUCAINE-1.2
ECK0456-ACRB	DIBUCAINE-1.2
ECK0457-ACRA	DIBUCAINE-1.2
ECK0726-TOLQ	DIBUCAINE-1.2
ECK0727-TOLR	DIBUCAINE-1.2
ECK0729-TOLB	DIBUCAINE-1.2
ECK0730-PAL	DIBUCAINE-1.2
ECK3026-TOLC	DIBUCAINE-1.2
ECK3042-RFAE	DIBUCAINE-1.2
ECK3610-RFAF	DIBUCAINE-1.2
ECK3618-RFAB	DIBUCAINE-1.2
ECK3620-RFAP	DIBUCAINE-1.2
ECK3621-RFAG	DIBUCAINE-1.2
ECK3622-RFAQ	DIBUCAINE-1.2
ECK3026-TOLC	DOXORUBICIN-10.0
ECK2504-GUAB	DOXYCYCLINE-0.25
ECK3048-FOLB	DOXYCYCLINE-0.25
ECK2504-GUAB	DOXYCYCLINE-0.5

ECK0456-ACRB	DOXYCYCLINE-0.75
ECK2504-GUAB	DOXYCYCLINE-0.75
ECK3026-TOLC	DOXYCYCLINE-0.75
ECK3621-RFAG	DOXYCYCLINE-0.75
ECK0457-ACRA	DOXYCYCLINE-1.0
ECK2504-GUAB	DOXYCYCLINE-1.0
ECK3026-TOLC	DOXYCYCLINE-1.0
ECK0582-FEPG	EDTA-1.0
ECK0729-TOLB	EDTA-1.0
ECK2321-YFCA	EGCG-20
ECK0516-PURE	EGCG-5
ECK1047-PYRC	EGCG-5
ECK2321-YFCA	EGCG-5
ECK2495-PURM	EGCG-5
ECK3632-PYRE	EGCG-5
ECK3997-PURD	EGCG-5
ECK2262-RBN	EGCG-50
ECK2321-YFCA	EGCG-50
ECK2504-GUAB	EGCG-50
ECK2998-EXBD	EGCG-50
ECK2503-GUAA	ERYTHROMYCIN-0.1
ECK2504-GUAB	ERYTHROMYCIN-0.1
ECK3632-PYRE	ERYTHROMYCIN-0.1
ECK2504-GUAB	ERYTHROMYCIN-1.0
ECK3632-PYRE	ERYTHROMYCIN-1.0
ECK0457-ACRA	ERYTHROMYCIN-10.0
ECK2504-GUAB	ERYTHROMYCIN-10.0
ECK3026-TOLC	ERYTHROMYCIN-10.0
ECK3611-RFAC	ERYTHROMYCIN-10.0
ECK2504-GUAB	ERYTHROMYCIN-10.0
ECK3026-TOLC	ERYTHROMYCIN-5.0
ECK3611-RFAC	ERYTHROMYCIN-5.0
ECK3597-CYSE	ERYTHROMYCIN-5.0
ECK0160-DEGP	ETHANOL-2.0
ECK0223-LPCA	ETHANOL-6.0
ECK3042-RFAE	ETHANOL-6.0
ECK0456-ACRB	ETHANOL-6.0
ECK0457-ACRA	ETHIDIUMBROMIDE-10
ECK0621-LIPA	ETHIDIUMBROMIDE-10
ECK3026-TOLC	ETHIDIUMBROMIDE-10
ECK3725-ATPD	ETHIDIUMBROMIDE-10
ECK0456-ACRB	ETHIDIUMBROMIDE-2
ECK0457-ACRA	ETHIDIUMBROMIDE-2
ECK0621-LIPA	ETHIDIUMBROMIDE-2
ECK3026-TOLC	ETHIDIUMBROMIDE-2
ECK3725-ATPD	ETHIDIUMBROMIDE-2
ECK3729-ATPF	ETHIDIUMBROMIDE-2
ECK3730-ATPE	ETHIDIUMBROMIDE-2
ECK4168-HFQ	ETHIDIUMBROMIDE-2

ECK0456-ACRB	ETHIDIUMBROMIDE-50
ECK0457-ACRA	ETHIDIUMBROMIDE-50
ECK2271-NUOM	ETHIDIUMBROMIDE-50
ECK2273-NUOK	ETHIDIUMBROMIDE-50
ECK2276-NUOH	ETHIDIUMBROMIDE-50
ECK2279-NUOE	ETHIDIUMBROMIDE-50
ECK2282-NUOA	ETHIDIUMBROMIDE-50
ECK2818-RECC	ETHIDIUMBROMIDE-50
ECK3026-TOLC	ETHIDIUMBROMIDE-50
ECK3725-ATPD	ETHIDIUMBROMIDE-50
ECK4168-HFQ	ETHIDIUMBROMIDE-50
ECK3409-GLPR	FOSFOMYCIN-1.0
ECK0223-LPCA	FUSIDICACID-20
ECK3026-TOLC	FUSIDICACID-20
ECK3042-RFAE	FUSIDICACID-20
ECK0223-LPCA	FUSIDICACID-50
ECK3026-TOLC	FUSIDICACID-50
ECK3042-RFAE	FUSIDICACID-50
ECK0034-CARB	GLUCOSAMINE
ECK0515-PURK	GLUCOSAMINE
ECK0516-PURE	GLUCOSAMINE
ECK1269-CYSB	GLUCOSAMINE
ECK2323-AROC	GLUCOSAMINE
ECK2472-PURC	GLUCOSAMINE
ECK2495-PURM	GLUCOSAMINE
ECK2503-GUAA	GLUCOSAMINE
ECK2555-PURL	GLUCOSAMINE
ECK3632-PYRE	GLUCOSAMINE
ECK3997-PURD	GLUCOSAMINE
ECK3998-PURH	GLUCOSAMINE
ECK4173-PURA	GLUCOSAMINE
ECK0034-CARB	GLUCOSE
ECK0515-PURK	GLUCOSE
ECK0516-PURE	GLUCOSE
ECK0936-PYRD	GLUCOSE
ECK1047-PYRC	GLUCOSE
ECK1269-CYSB	GLUCOSE
ECK2323-AROC	GLUCOSE
ECK2472-PURC	GLUCOSE
ECK2495-PURM	GLUCOSE
ECK2503-GUAA	GLUCOSE
ECK2504-GUAB	GLUCOSE
ECK2555-PURL	GLUCOSE
ECK3632-PYRE	GLUCOSE
ECK3997-PURD	GLUCOSE
ECK3998-PURH	GLUCOSE
ECK0034-CARB	GLYCEROL
ECK0515-PURK	GLYCEROL
ECK0516-PURE	GLYCEROL

ECK1269-CYSB	GLYCEROL
ECK2323-AROC	GLYCEROL
ECK2472-PURC	GLYCEROL
ECK2495-PURM	GLYCEROL
ECK2504-GUAB	GLYCEROL
ECK2555-PURL	GLYCEROL
ECK3632-PYRE	GLYCEROL
ECK3997-PURD	GLYCEROL
ECK3998-PURH	GLYCEROL
ECK4173-PURA	GLYCEROL
ECK4227-FBP	GLYCEROL
ECK4227-FBP	HIGHCOPPER-1.0
ECK0122-CUEO	HIGHCOPPER-4.0
ECK0478-COPA	HIGHCOPPER-4.0
ECK0726-TOLQ	HIGHCOPPER-4.0
ECK0727-TOLR	HIGHCOPPER-4.0
ECK1248-YCIB	HIGHCOPPER-4.0
ECK3831-TATB	HIGHCOPPER-4.0
ECK3832-TATC	HIGHCOPPER-4.0
ECK0034-CARB	HIGHFE
ECK0515-PURK	HIGHFE
ECK0516-PURE	HIGHFE
ECK0936-PYRD	HIGHFE
ECK1047-PYRC	HIGHFE
ECK1269-CYSB	HIGHFE
ECK2323-AROC	HIGHFE
ECK2472-PURC	HIGHFE
ECK2495-PURM	HIGHFE
ECK2503-GUAA	HIGHFE
ECK2504-GUAB	HIGHFE
ECK2555-PURL	HIGHFE
ECK3597-CYSE	HIGHFE
ECK3632-PYRE	HIGHFE
ECK3997-PURD	HIGHFE
ECK3998-PURH	HIGHFE
ECK4173-PURA	HIGHFE
ECK1269-CYSB	HIGHNICKEL-0.1
ECK0144-DKSA	HIGHNICKEL-1.0
ECK2228-YFAE	HYDROXYUREA-10.0
ECK2228-YFAE	HYDROXYUREA-5.0
ECK0034-CARB	LOWFE
ECK0515-PURK	LOWFE
ECK0516-PURE	LOWFE
ECK0936-PYRD	LOWFE
ECK1269-CYSB	LOWFE
ECK2323-AROC	LOWFE
ECK2472-PURC	LOWFE
ECK2495-PURM	LOWFE
ECK2504-GUAB	LOWFE

ECK2555-PURL	LOWFE
ECK3597-CYSE	LOWFE
ECK3632-PYRE	LOWFE
ECK3997-PURD	LOWFE
ECK3998-PURH	LOWFE
ECK0034-CARB	MALTOSE
ECK0515-PURK	MALTOSE
ECK0516-PURE	MALTOSE
ECK1269-CYSB	MALTOSE
ECK2323-AROC	MALTOSE
ECK2472-PURC	MALTOSE
ECK2495-PURM	MALTOSE
ECK2503-GUAA	MALTOSE
ECK2504-GUAB	MALTOSE
ECK2555-PURL	MALTOSE
ECK3632-PYRE	MALTOSE
ECK3997-PURD	MALTOSE
ECK3998-PURH	MALTOSE
ECK2504-GUAB	MECILLINAM-0.03
ECK0729-TOLB	MECILLINAM-0.06
ECK2504-GUAB	MECILLINAM-0.06
ECK0729-TOLB	MECILLINAM-0.09
ECK1047-PYRC	MECILLINAM-0.09
ECK2504-GUAB	MECILLINAM-0.09
ECK2504-GUAB	MECILLINAM-0.12
ECK3026-TOLC	MINOCYCLINE-0.5
ECK0456-ACRB	MINOCYCLINE-1.0
ECK0457-ACRA	MINOCYCLINE-1.0
ECK3026-TOLC	MINOCYCLINE-1.0
ECK3042-RFAE	MINOCYCLINE-1.0
ECK0054-SURA	MITOMYCINC-0.1
ECK0456-ACRB	MITOMYCINC-0.1
ECK2694-RECA	MITOMYCINC-0.1
ECK1862-RUVA	MMS-0.05%
ECK2694-RECA	MMS-0.05%
ECK0034-CARB	N-ACETYLGLUCOSAMINE
ECK0515-PURK	N-ACETYLGLUCOSAMINE
ECK0516-PURE	N-ACETYLGLUCOSAMINE
ECK0665-NAGA	N-ACETYLGLUCOSAMINE
ECK1269-CYSB	N-ACETYLGLUCOSAMINE
ECK2323-AROC	N-ACETYLGLUCOSAMINE
ECK2472-PURC	N-ACETYLGLUCOSAMINE
ECK2495-PURM	N-ACETYLGLUCOSAMINE
ECK2503-GUAA	N-ACETYLGLUCOSAMINE
ECK2504-GUAB	N-ACETYLGLUCOSAMINE
ECK2555-PURL	N-ACETYLGLUCOSAMINE
ECK3632-PYRE	N-ACETYLGLUCOSAMINE
ECK3997-PURD	N-ACETYLGLUCOSAMINE
ECK3998-PURH	N-ACETYLGLUCOSAMINE

ECK1091-YCFM	NACL-600
ECK2321-YFCA	NACL-600
ECK3597-CYSE	NALIDIXICACID-0.5
ECK0223-LPCA	NALIDIXICACID-2.0
ECK3042-RFAE	NALIDIXICACID-2.0
ECK3597-CYSE	NALIDIXICACID-2.0
ECK0034-CARB	NH4CL
ECK0515-PURK	NH4CL
ECK0516-PURE	NH4CL
ECK0936-PYRD	NH4CL
ECK1047-PYRC	NH4CL
ECK1269-CYSB	NH4CL
ECK2323-AROC	NH4CL
ECK2472-PURC	NH4CL
ECK2495-PURM	NH4CL
ECK2503-GUAA	NH4CL
ECK2504-GUAB	NH4CL
ECK2555-PURL	NH4CL
ECK3048-FOLB	NH4CL
ECK3597-CYSE	NH4CL
ECK3632-PYRE	NH4CL
ECK3997-PURD	NH4CL
ECK3998-PURH	NH4CL
ECK4173-PURA	NH4CL
ECK0621-LIPA	NITROFURANTOIN-0.1
ECK2694-RECA	NITROFURANTOIN-2.0
ECK2694-RECA	NORFLOXACIN-0.01
ECK2818-RECC	NORFLOXACIN-0.01
ECK3731-ATPB	NORFLOXACIN-0.01
ECK2694-RECA	NORFLOXACIN-0.02
ECK2818-RECC	NORFLOXACIN-0.02
ECK1269-CYSB	NOVOBIOCIN-10
ECK3026-TOLC	NOVOBIOCIN-10
ECK3611-RFAC	NOVOBIOCIN-10
ECK0034-CARB	NOVOBIOCIN-12
ECK0515-PURK	NOVOBIOCIN-12
ECK0516-PURE	NOVOBIOCIN-12
ECK0936-PYRD	NOVOBIOCIN-12
ECK1047-PYRC	NOVOBIOCIN-12
ECK2495-PURM	NOVOBIOCIN-12
ECK3026-TOLC	NOVOBIOCIN-12
ECK3042-RFAE	NOVOBIOCIN-12
ECK3611-RFAC	NOVOBIOCIN-12
ECK3632-PYRE	NOVOBIOCIN-12
ECK3997-PURD	NOVOBIOCIN-12
ECK4173-PURA	NOVOBIOCIN-12
ECK3026-TOLC	NOVOBIOCIN-30
ECK1269-CYSB	NOVOBIOCIN-4
ECK2504-GUAB	NOVOBIOCIN-4

ECK3632-PYRE	NOVOBIOCIN-4
ECK1269-CYSB	NOVOBIOCIN-6
ECK2504-GUAB	NOVOBIOCIN-6
ECK3026-TOLC	NOVOBIOCIN-6
ECK3632-PYRE	NOVOBIOCIN-6
ECK0516-PURE	NOVOBIOCIN-8
ECK0936-PYRD	NOVOBIOCIN-8
ECK1047-PYRC	NOVOBIOCIN-8
ECK1269-CYSB	NOVOBIOCIN-8
ECK2495-PURM	NOVOBIOCIN-8
ECK2555-PURL	NOVOBIOCIN-8
ECK3026-TOLC	NOVOBIOCIN-8
ECK3632-PYRE	NOVOBIOCIN-8
ECK3997-PURD	NOVOBIOCIN-8
ECK4173-PURA	NOVOBIOCIN-8
ECK0457-ACRA	OXACILLIN-40.0
ECK3026-TOLC	OXACILLIN-40.0
ECK2504-GUAB	PARAQUAT-10.0
ECK2504-GUAB	PARAQUAT-18.0
ECK0722-CYDB	PH10
ECK5005-TP2	PH10
ECK0050-APAH	PH4
ECK0054-SURA	PH4
ECK0223-LPCA	PH4
ECK0726-TOLQ	PH4
ECK0727-TOLR	PH4
ECK0729-TOLB	PH4
ECK0730-PAL	PH4
ECK1091-YCFM	PH4
ECK1269-CYSB	PH4
ECK3026-TOLC	PH4
ECK3164-SECG	PH4
ECK3597-CYSE	PH4
ECK3699-MNME	PH4
ECK3734-MNMG	PH4
ECK3904-CPXA	PH4
ECK0050-APAH	PH4.5
ECK0726-TOLQ	PH4.5
ECK0729-TOLB	PH4.5
ECK1269-CYSB	PH4.5
ECK3026-TOLC	PH4.5
ECK3597-CYSE	PH4.5
ECK3904-CPXA	PH4.5
ECK1269-CYSB	PH5
ECK3597-CYSE	PH5
ECK1269-CYSB	PH9.5
ECK3026-TOLC	PMS-0.05
ECK3026-TOLC	PMS-0.1
ECK3042-RFAE	PMS-0.1

ECK3836-FRE	PMS-0.1
ECK4054-SOXS	PMS-0.1
ECK2504-GUAB	PROCAINE-1
ECK2504-GUAB	PROCAINE-10
ECK3620-RFAP	PROCAINE-10
ECK2504-GUAB	PROCAINE-30
ECK2504-GUAB	PROCAINE-5
ECK3597-CYSE	PROPIDIUMIODIDE-1
ECK0456-ACRB	PROPIDIUMIODIDE-20
ECK0457-ACRA	PROPIDIUMIODIDE-20
ECK3026-TOLC	PROPIDIUMIODIDE-20
ECK0456-ACRB	PROPIDIUMIODIDE-50
ECK0457-ACRA	PROPIDIUMIODIDE-50
ECK3026-TOLC	PROPIDIUMIODIDE-50
ECK0457-ACRA	PUROMYCIN-25
ECK3026-TOLC	PUROMYCIN-25
ECK0456-ACRB	PUROMYCIN-5
ECK0457-ACRA	PUROMYCIN-5
ECK3026-TOLC	PUROMYCIN-5
ECK2305-UBIX	PYOCYANIN-1.0
ECK3048-FOLB	PYOCYANIN-1.0
ECK0144-DKSA	PYOCYANIN-10.0
ECK0456-ACRB	PYOCYANIN-10.0
ECK0457-ACRA	PYOCYANIN-10.0
ECK0458-ACRR	PYOCYANIN-10.0
ECK0665-NAGA	PYOCYANIN-10.0
ECK1260-YCIV	PYOCYANIN-10.0
ECK1864-RUVC	PYOCYANIN-10.0
ECK2271-NUOM	PYOCYANIN-10.0
ECK2314-PDXB	PYOCYANIN-10.0
ECK2901-VISC	PYOCYANIN-10.0
ECK2903-PEPP	PYOCYANIN-10.0
ECK3026-TOLC	PYOCYANIN-10.0
ECK3033-YQIC	PYOCYANIN-10.0
ECK3621-RFAG	PYOCYANIN-10.0
ECK3756-HDFR	PYOCYANIN-10.0
ECK3836-FRE	PYOCYANIN-10.0
ECK3916-FPR	PYOCYANIN-10.0
ECK4054-SOXS	PYOCYANIN-10.0
ECK4055-SOXR	PYOCYANIN-10.0
ECK4168-HFQ	PYOCYANIN-10.0
ECK4393-ARCA	PYOCYANIN-10.0
ECK0054-SURA	SDS-0.5%
ECK0223-LPCA	SDS-0.5%
ECK0457-ACRA	SDS-0.5%
ECK0726-TOLQ	SDS-0.5%
ECK0727-TOLR	SDS-0.5%
ECK0729-TOLB	SDS-0.5%
ECK0730-PAL	SDS-0.5%

ECK1673-LPP	SDS-0.5%
ECK3026-TOLC	SDS-0.5%
ECK3042-RFAE	SDS-0.5%
ECK3138-YRAP	SDS-0.5%
ECK3609-RFAD	SDS-0.5%
ECK3610-RFAF	SDS-0.5%
ECK3620-RFAP	SDS-0.5%
ECK3621-RFAG	SDS-0.5%
ECK4168-HFQ	SDS-0.5%
ECK0223-LPCA	SDS-1.0%
ECK0457-ACRA	SDS-1.0%
ECK0726-TOLQ	SDS-1.0%
ECK0727-TOLR	SDS-1.0%
ECK0729-TOLB	SDS-1.0%
ECK0730-PAL	SDS-1.0%
ECK0903-IHFB	SDS-1.0%
ECK3042-RFAE	SDS-1.0%
ECK3610-RFAF	SDS-1.0%
ECK3620-RFAP	SDS-1.0%
ECK3621-RFAG	SDS-1.0%
ECK0223-LPCA	SDS-2.0%
ECK0457-ACRA	SDS-2.0%
ECK0726-TOLQ	SDS-2.0%
ECK0727-TOLR	SDS-2.0%
ECK0729-TOLB	SDS-2.0%
ECK0730-PAL	SDS-2.0%
ECK1673-LPP	SDS-2.0%
ECK3026-TOLC	SDS-2.0%
ECK3042-RFAE	SDS-2.0%
ECK3610-RFAF	SDS-2.0%
ECK3621-RFAG	SDS-2.0%
ECK0223-LPCA	SDS-3.0%
ECK0726-TOLQ	SDS-3.0%
ECK0727-TOLR	SDS-3.0%
ECK0729-TOLB	SDS-3.0%
ECK0730-PAL	SDS-3.0%
ECK1673-LPP	SDS-3.0%
ECK3026-TOLC	SDS-3.0%
ECK3042-RFAE	SDS-3.0%
ECK3609-RFAD	SDS-3.0%
ECK3610-RFAF	SDS-3.0%
ECK3620-RFAP	SDS-3.0%
ECK3621-RFAG	SDS-3.0%
ECK3998-PURH	SDS-3.0%
ECK0223-LPCA	SDS-4.0%
ECK0456-ACRB	SDS-4.0%
ECK0457-ACRA	SDS-4.0%
ECK0726-TOLQ	SDS-4.0%
ECK0727-TOLR	SDS-4.0%

ECK0729-TOLB	SDS-4.0%
ECK1673-LPP	SDS-4.0%
ECK3026-TOLC	SDS-4.0%
ECK3042-RFAE	SDS-4.0%
ECK3609-RFAD	SDS-4.0%
ECK3610-RFAF	SDS-4.0%
ECK3620-RFAP	SDS-4.0%
ECK3621-RFAG	SDS-4.0%
ECK3831-TATB	SDS-4.0%
ECK3832-TATC	SDS-4.0%
ECK4168-HFQ	SDS-4.0%
ECK0457-ACRA	SDS0.5%/EDTA0.1
ECK0726-TOLQ	SDS0.5%/EDTA0.1
ECK0729-TOLB	SDS0.5%/EDTA0.1
ECK1673-LPP	SDS0.5%/EDTA0.1
ECK3026-TOLC	SDS0.5%/EDTA0.1
ECK3042-RFAE	SDS0.5%/EDTA0.1
ECK3138-YRAP	SDS0.5%/EDTA0.1
ECK3610-RFAF	SDS0.5%/EDTA0.1
ECK3621-RFAG	SDS0.5%/EDTA0.1
ECK0054-SURA	SDS0.5%/EDTA0.5
ECK0223-LPCA	SDS0.5%/EDTA0.5
ECK0457-ACRA	SDS0.5%/EDTA0.5
ECK0583-FEPD	SDS0.5%/EDTA0.5
ECK0585-FEPB	SDS0.5%/EDTA0.5
ECK0654-UBIF	SDS0.5%/EDTA0.5
ECK0665-NAGA	SDS0.5%/EDTA0.5
ECK0726-TOLQ	SDS0.5%/EDTA0.5
ECK0727-TOLR	SDS0.5%/EDTA0.5
ECK0729-TOLB	SDS0.5%/EDTA0.5
ECK0730-PAL	SDS0.5%/EDTA0.5
ECK1673-LPP	SDS0.5%/EDTA0.5
ECK2273-NUOK	SDS0.5%/EDTA0.5
ECK2274-NUOJ	SDS0.5%/EDTA0.5
ECK2276-NUOH	SDS0.5%/EDTA0.5
ECK2282-NUOA	SDS0.5%/EDTA0.5
ECK2323-AROC	SDS0.5%/EDTA0.5
ECK2737-NLPD	SDS0.5%/EDTA0.5
ECK3026-TOLC	SDS0.5%/EDTA0.5
ECK3042-RFAE	SDS0.5%/EDTA0.5
ECK3234-YHDP	SDS0.5%/EDTA0.5
ECK3603-ENVC	SDS0.5%/EDTA0.5
ECK3610-RFAF	SDS0.5%/EDTA0.5
ECK3831-TATB	SDS0.5%/EDTA0.5
ECK4168-HFQ	SDS0.5%/EDTA0.5
ECK0054-SURA	SDS1.0%/EDTA0.5
ECK0223-LPCA	SDS1.0%/EDTA0.5
ECK0581-FEPC	SDS1.0%/EDTA0.5
ECK0583-FEPD	SDS1.0%/EDTA0.5

ECK0585-FEPB	SDS1.0%/EDTA0.5
ECK0665-NAGA	SDS1.0%/EDTA0.5
ECK0726-TOLQ	SDS1.0%/EDTA0.5
ECK0727-TOLR	SDS1.0%/EDTA0.5
ECK0729-TOLB	SDS1.0%/EDTA0.5
ECK0730-PAL	SDS1.0%/EDTA0.5
ECK1673-LPP	SDS1.0%/EDTA0.5
ECK1780-MIPA	SDS1.0%/EDTA0.5
ECK3026-TOLC	SDS1.0%/EDTA0.5
ECK3042-RFAE	SDS1.0%/EDTA0.5
ECK3234-YHDP	SDS1.0%/EDTA0.5
ECK3602-GPMM	SDS1.0%/EDTA0.5
ECK3610-RFAF	SDS1.0%/EDTA0.5
ECK3620-RFAP	SDS1.0%/EDTA0.5
ECK3621-RFAG	SDS1.0%/EDTA0.5
ECK3831-TATB	SDS1.0%/EDTA0.5
ECK3832-TATC	SDS1.0%/EDTA0.5
ECK4168-HFQ	SDS1.0%/EDTA0.5
ECK0093-DDLB	SPECTINOMYCIN-6.0
ECK2504-GUAB	SPECTINOMYCIN-6.0
ECK3156-RBFA	SPECTINOMYCIN-6.0
ECK3713-BGLH	SPECTINOMYCIN-6.0
ECK4419-YGDT	SPECTINOMYCIN-6.0
ECK0223-LPCA	SPIRAMYCIN-20
ECK0457-ACRA	SPIRAMYCIN-20
ECK2490-YFGC	SPIRAMYCIN-20
ECK2508-BAMB	SPIRAMYCIN-20
ECK3026-TOLC	SPIRAMYCIN-20
ECK3042-RFAE	SPIRAMYCIN-20
ECK3609-RFAD	SPIRAMYCIN-20
ECK0223-LPCA	STREPTONIGRIN-0.5
ECK0456-ACRB	STREPTONIGRIN-0.5
ECK0034-CARB	SUCCINATE
ECK0515-PURK	SUCCINATE
ECK0581-FEPC	SUCCINATE
ECK1269-CYSB	SUCCINATE
ECK2472-PURC	SUCCINATE
ECK2504-GUAB	SUCCINATE
ECK2555-PURL	SUCCINATE
ECK3997-PURD	SUCCINATE
ECK3998-PURH	SUCCINATE
ECK4227-FBP	SUCCINATE
ECK2504-GUAB	TAUROCHOLATE-0.1%
ECK0034-CARB	TAUROCHOLATE-0.5%
ECK0516-PURE	TAUROCHOLATE-0.5%
ECK0936-PYRD	TAUROCHOLATE-0.5%
ECK1047-PYRC	TAUROCHOLATE-0.5%
ECK2495-PURM	TAUROCHOLATE-0.5%
ECK2504-GUAB	TAUROCHOLATE-0.5%

ECK2555-PURL	TAUROCHOLATE-0.5%
ECK3026-TOLC	TAUROCHOLATE-0.5%
ECK3632-PYRE	TAUROCHOLATE-0.5%
ECK3997-PURD	TAUROCHOLATE-0.5%
ECK0457-ACRA	TAUROCHOLATE-1.0%
ECK2504-GUAB	TAUROCHOLATE-1.0%
ECK3026-TOLC	TAUROCHOLATE-1.0%
ECK2504-GUAB	TOBRAMYCIN-0.05
ECK2504-GUAB	TOBRAMYCIN-0.1
ECK2504-GUAB	TOBRAMYCIN-0.2
ECK2504-GUAB	TOBRAMYCIN-0.4
ECK0223-LPCA	TRICLOSAN-0.05
ECK0456-ACRB	TRICLOSAN-0.05
ECK0457-ACRA	TRICLOSAN-0.05
ECK3026-TOLC	TRICLOSAN-0.05
ECK3042-RFAE	TRICLOSAN-0.05
ECK3610-RFAF	TRICLOSAN-0.05
ECK2504-GUAB	TRITONX-0.01%
ECK3597-CYSE	TRITONX-0.01%
ECK2504-GUAB	TRITONX-0.03%
ECK2613-SMPA	TRITONX-0.03%
ECK3026-TOLC	TRITONX-0.03%
ECK3042-RFAE	TRITONX-0.03%
ECK0223-LPCA	TRITONX-0.2%
ECK2504-GUAB	TRITONX-0.2%
ECK3026-TOLC	TRITONX-0.2%
ECK3042-RFAE	TRITONX-0.2%
ECK3156-RBFA	TRITONX-0.2%
ECK2321-YFCA	TUNICAMYCIN-7.5
ECK3725-ATPD	TUNICAMYCIN-7.5
ECK2508-BAMB	VANCOMYCIN-10
ECK0729-TOLB	VANCOMYCIN-20
ECK0730-PAL	VANCOMYCIN-20
ECK1269-CYSB	VANCOMYCIN-20
ECK2508-BAMB	VANCOMYCIN-20
ECK3597-CYSE	VANCOMYCIN-20
ECK0726-TOLQ	VANCOMYCIN-50
ECK0727-TOLR	VANCOMYCIN-50
ECK0729-TOLB	VANCOMYCIN-50
ECK0730-PAL	VANCOMYCIN-50
ECK1077-FABH	VANCOMYCIN-50
ECK2490-YFGC	VANCOMYCIN-50
ECK3597-CYSE	VANCOMYCIN-50
ECK2504-GUAB	VERAPAMIL-0.1
ECK3597-CYSE	VERAPAMIL-0.1
ECK2504-GUAB	VERAPAMIL-0.5
ECK3597-CYSE	VERAPAMIL-0.5
ECK2504-GUAB	VERAPAMIL-1.0
ECK3597-CYSE	VERAPAMIL-1.0

Table S6 - Highly Correlated Orphan-Annotated Gene Pairs ($r \geq 0.5$)

Orphan Gene	Annotated Gene	Correlation Coefficient	Annotated Gene Function	Annotated Gene Function - more detail
ECK3654- YICN	ECK3631 -SLMA	0.5106	cell division/cell cycle	
ECK3647- YICJ	ECK0881 -FTSK	0.6015	cell division/cell cycle	
ECK4135- YJEH	ECK3550 -WECH	0.5308	Cell envelope biogenesis/outer membrane	Antigen biosynthesis
ECK3874- YIHT	ECK3780 -RFFG	0.501	Cell envelope biogenesis/outer membrane	Antigen biosynthesis
ECK3993- YJAH	ECK3780 -RFFG	0.5076	Cell envelope biogenesis/outer membrane	Antigen biosynthesis
ECK1600- YDGI	ECK1231 -GALU	0.6289	Cell envelope biogenesis/outer membrane	Antigen/LPS biosynthesis
ECK2333- YFCV	ECK2047 -GMD	0.5171	Cell envelope biogenesis/outer membrane	Antigen/LPS biosynthesis
ECK2684- YQAA	ECK2047 -GMD	0.5461	Cell envelope biogenesis/outer membrane	Antigen/LPS biosynthesis
ECK0835- YBJJ	ECK1028 -CSGA	0.5116	Cell envelope biogenesis/outer membrane	curli
ECK1400- YDBD	ECK1028 -CSGA	0.5007	Cell envelope biogenesis/outer membrane	curli
ECK1428- YDCN	ECK1028 -CSGA	0.5293	Cell envelope biogenesis/outer membrane	curli
ECK0983- YCCM	ECK0523 -SFMA	0.5269	Cell envelope biogenesis/outer membrane	fimbriae
ECK1428- YDCN	ECK0523 -SFMA	0.528	Cell envelope biogenesis/outer membrane	fimbriae
ECK1748- YDJX	ECK0523 -SFMA	0.5051	Cell envelope biogenesis/outer membrane	fimbriae
ECK1275- YCIM	ECK2058 -ASMA	0.5044	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration

ECK2490- YFGC	ECK2058 -ASMA	0.615	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK1274- YCIS	ECK3620 -RFAP	0.5373	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3914- YIIS	ECK3609 -RFAD	0.6445	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK1274- YCIS	ECK3622 -RFAQ	0.5635	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3181- YRBC	ECK3622 -RFAQ	0.5982	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3182- YRBD	ECK3622 -RFAQ	0.5778	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK2490- YFGC	ECK1856 -LPXM	0.515	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3181- YRBC	ECK3617 -RFAI	0.5862	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3182- YRBD	ECK3617 -RFAI	0.5928	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3914- YIIS	ECK3042 -RFAE	0.5281	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK4370- YJJV	ECK2374 -LPXP	0.5156	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3147- YHBV	ECK0517 -LPXH	0.5145	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK2490- YFGC	ECK0180 -LPXA	0.5683	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport

ECK2490- YFGC	ECK0097 -B- LPXC(G2 10S)- KAN	0.5593	Cell envelope biogenesis/outer membrane	t/decoration LPS biosynthesis/as sembly/transport/ decoration
ECK1419- YDCH	ECK0055 -C-IMP- DAS- KAN	0.5826	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK2153- YEII	ECK2250 -ARNT	0.5328	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK0780- YBHQ	ECK2028 -WBBI	0.5232	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK2064- YEGI	ECK2028 -WBBI	0.5978	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3914- YIIS	ECK0223 -LPCA	0.5304	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK3875- YIHU	ECK2036 -GALF	0.5306	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK0420- YAJQ	ECK2703 -GUTQ	0.5747	Cell envelope biogenesis/outer membrane	LPS biosynthesis/as sembly/transport/ decoration
ECK2490- YFGC	ECK2508 -BAMB	0.5752	Cell envelope biogenesis/outer membrane	OMP assembly/transport
ECK2490- YFGC	ECK0176 -B- YAET(21 8- 9DUPL)- KAN	0.678	Cell envelope biogenesis/outer membrane	OMP assembly/transport
ECK1091- YCFM (LPOB)	ECK0148 -MRCB	0.8984	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK0252- YKFB	ECK2696 -MLTB	0.5476	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK1040- YCEA	ECK1180 -LDCA	0.618	Cell envelope biogenesis/outer	Peptidoglycan Biosynthesis-

ECK1090- YCFL	ECK1180 -LDCA	0.5507	membrane Cell envelope biogenesis/outer membrane	turnover Peptidoglycan Biosynthesis- turnover
ECK1228- YCHJ	ECK1180 -LDCA	0.6602	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK3562- YSAA	ECK1180 -LDCA	0.5752	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK4243- YJGI	ECK1180 -LDCA	0.5525	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK3234- YHDP	ECK1780 -MIPA	0.5807	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK1440- YDCY	ECK1322 -MPAA	0.5194	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK0868- YBJX	ECK1985 -ERFK	0.5118	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK1086- YCFH	ECK1985 -ERFK	0.5002	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK0650- YBEX	ECK0858 -AMID	0.5436	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK1228- YCHJ	ECK0858 -AMID	0.5175	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK1267- YGIN	ECK0858 -AMID	0.5249	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK0480- YBAT	ECK1706 -NLPC	0.5007	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK1099- YCFS	ECK1181 -EMTA	0.518	Cell envelope biogenesis/outer membrane	Peptidoglycan Biosynthesis- turnover
ECK3181- YRBC	ECK3184 -YRBF	0.6799	Cell envelope biogenesis/outer membrane	phospholipid biosynthesis/tra nsport
ECK3182- YRBD	ECK3184 -YRBF	0.6797	Cell envelope biogenesis/outer membrane	phospholipid biosynthesis/tra nsport
ECK3181- YRBC	ECK3183 -YRBE	0.7059	Cell envelope biogenesis/outer membrane	phospholipid biosynthesis/tra nsport
ECK3182- YRBD	ECK3183 -YRBE	0.7116	Cell envelope biogenesis/outer	phospholipid biosynthesis/tra

ECK3181- YRBC	ECK2340 -VACJ	0.5596	membrane Cell envelope biogenesis/outer membrane	nsport phospholipid biosynthesis/tra nsport
ECK3182- YRBD	ECK2340 -VACJ	0.5394	Cell envelope biogenesis/outer membrane	phospholipid biosynthesis/tra nsport
ECK0793- YBIX	ECK1076 -PLSX	0.5122	Cell envelope biogenesis/outer membrane	phospholipid biosynthesis/tra nsport
ECK3234- YHDP	ECK1673 -LPP	0.5452	Cell envelope biogenesis/outer membrane	
ECK3138- YRAP	ECK1673 -LPP	0.6451	Cell envelope biogenesis/outer membrane	
ECK0996- YMDF	ECK2045 -GMM	0.5003	Cell envelope biogenesis/outer membrane	Antigen/LPS biosynthesis
ECK0725- YBGC	ECK0730 -PAL	0.7772	Cell envelope biogenesis/outer membrane-cell division	
ECK3138- YRAP	ECK0730 -PAL	0.5433	Cell envelope biogenesis/outer membrane-cell division	
ECK0725- YBGC	ECK0726 -TOLQ	0.623	Cell envelope biogenesis/outer membrane-cell division	
ECK0725- YBGC	ECK0727 -TOLR	0.6438	Cell envelope biogenesis/outer membrane-cell division	
ECK3138- YRAP	ECK0727 -TOLR	0.5513	Cell envelope biogenesis/outer membrane-cell division	
ECK0725- YBGC	ECK0729 -TOLB	0.6978	Cell envelope biogenesis/outer membrane-cell division	
ECK3138- YRAP	ECK0729 -TOLB	0.5099	Cell envelope biogenesis/outer membrane-cell division	
ECK1040- YCEA	ECK1061 -FLGE	0.5418	Cell motility & chemotaxis	
ECK1228- YCHJ	ECK1061 -FLGE	0.5372	Cell motility & chemotaxis	

ECK1439- YDCX	ECK1061 -FLGE	0.5386	Cell motility & chemotaxis
ECK0046- YAAU	ECK1939 -FLII	0.5358	Cell motility & chemotaxis
ECK2684- YQAA	ECK1939 -FLII	0.5126	Cell motility & chemotaxis
ECK1491- YDEM	ECK1415 -TRG	0.5274	Cell motility & chemotaxis
ECK1664- YDHS	ECK1415 -TRG	0.5243	Cell motility & chemotaxis
ECK1036- YCEK	ECK1890 -MOTB	0.5102	Cell motility & chemotaxis
ECK1664- YDHS	ECK1890 -MOTB	0.5594	Cell motility & chemotaxis
ECK0996- YMDF	ECK1890 -MOTB	0.5414	Cell motility & chemotaxis
ECK2911- YQFE	ECK1890 -MOTB	0.562	Cell motility & chemotaxis
ECK1440- YDCY	ECK1063 -FLGG	0.5614	Cell motility & chemotaxis
ECK3591- YIBT	ECK1063 -FLGG	0.6949	Cell motility & chemotaxis
ECK1513- YNEE	ECK1063 -FLGG	0.5045	Cell motility & chemotaxis
ECK2911- YQFE	ECK1063 -FLGG	0.5177	Cell motility & chemotaxis
ECK2925- YGGD	ECK1944 -FLIN	0.5323	Cell motility & chemotaxis
ECK1245- YCII	ECK1067 -FLGK	0.5531	Cell motility & chemotaxis
ECK1440- YDCY	ECK1067 -FLGK	0.5522	Cell motility & chemotaxis
ECK2850- YGEH	ECK1067 -FLGK	0.5014	Cell motility & chemotaxis
ECK3028- YGIB	ECK1067 -FLGK	0.5047	Cell motility & chemotaxis
ECK2925- YGGD	ECK1946 -FLIP	0.5011	Cell motility & chemotaxis
ECK2926- YGGF	ECK1946 -FLIP	0.5314	Cell motility & chemotaxis
ECK2684- YQAA	ECK1946 -FLIP	0.5334	Cell motility & chemotaxis
ECK1933	ECK1880 -FLHA	0.5339	Cell motility & chemotaxis
ECK3670- YIDK	ECK1880 -FLHA	0.5065	Cell motility & chemotaxis
ECK1491- YDEM	ECK1938 -FLIH	0.561	Cell motility & chemotaxis
ECK1664- YDHS	ECK1938 -FLIH	0.5149	Cell motility & chemotaxis
ECK4152-	ECK1945	0.6588	Cell motility &

YJEM	-FLIO		chemotaxis	
ECK0307-	ECK1945	0.644	Cell motility &	
YKGG	-FLIO		chemotaxis	
ECK1453	ECK1945	0.6548	Cell motility &	
	-FLIO		chemotaxis	
ECK1253-	ECK1940	0.5194	Cell motility &	
YCIG	-FLIJ		chemotaxis	
ECK1245-	ECK1940	0.504	Cell motility &	
YCII	-FLIJ		chemotaxis	
ECK3022-	ECK2995	0.5317	defense mechanisms	detoxification of
YQIA	-YGHZ			methylglyoxal
ECK2660-	ECK2295	0.5288	defense mechanisms	glutathione
YQAE	-YFCF			transferase
ECK1245-	ECK1331	0.5403	DNA	DNA
YCII	-OGT		replication/recombina	methylation
			tion/repair	
ECK1428-	ECK1331	0.5617	DNA	DNA
YDCN	-OGT		replication/recombina	methylation
			tion/repair	
ECK1553-	ECK1331	0.5818	DNA	DNA
QUUQ	-OGT		replication/recombina	methylation
			tion/repair	
ECK1689-	ECK1331	0.5577	DNA	DNA
YDIN	-OGT		replication/recombina	methylation
			tion/repair	
ECK2850-	ECK1331	0.5956	DNA	DNA
YGEH	-OGT		replication/recombina	methylation
			tion/repair	
ECK1130-	ECK1331	0.5059	DNA	DNA
YMFJ	-OGT		replication/recombina	methylation
			tion/repair	
ECK1370-	ECK1331	0.532	DNA	DNA
YNAE	-OGT		replication/recombina	methylation
			tion/repair	
ECK1443-	ECK1331	0.5305	DNA	DNA
YNCB	-OGT		replication/recombina	methylation
			tion/repair	
ECK0465-	ECK4050	0.505	DNA	DNA repair
YBAB	-UVRA		replication/recombina	
			tion/repair	
ECK1978-	ECK1346	0.5371	DNA	recombination
YEEN	-RECT		replication/recombina	& repair
			tion/repair	
ECK2333-	ECK1346	0.541	DNA	recombination
YFCV	-RECT		replication/recombina	& repair
			tion/repair	
ECK2992-	ECK1346	0.5675	DNA	recombination
YGHW	-RECT		replication/recombina	& repair
			tion/repair	
ECK2750-	ECK4253	0.5304	DNA	
YGBT	-PEPA		replication/recombina	

ECK0465- YBAB	ECK3808 -UVRD	0.5069	tion/repair DNA replication/recombina tion/repair
ECK0465- YBAB	ECK1862 -RUVA	0.5441	DNA replication/recombina tion/repair
ECK0465- YBAB	ECK1864 -RUVC	0.5694	DNA replication/recombina tion/repair
ECK0465- YBAB	ECK0466 -RECR	0.7298	DNA replication/recombina tion/repair
ECK0465- YBAB	ECK0768 -UVRB	0.62	DNA replication/recombina tion/repair
ECK0542- YLCG	ECK0541 -RUSA	0.6758	DNA replication/recombina tion/repair
ECK1040- YCEA	ECK1144 -PINE	0.5183	DNA replication/recombina tion/repair
ECK1090- YCFL	ECK1144 -PINE	0.5094	DNA replication/recombina tion/repair
ECK0465- YBAB	ECK1912 -UVRC	0.5801	DNA replication/recombina tion/repair
ECK3110- YHAC	ECK3637 -LIGB	0.5032	DNA replication/recombina tion/repair
ECK3803- YIFL	ECK3990 -NFI	0.5122	DNA replication/recombina tion/repair
ECK0465- YBAB	ECK2563 -RECO	0.7379	DNA replication/recombina tion/repair
ECK1426- YDCM	ECK2578 -UNG	0.5737	DNA replication/recombina tion/repair
ECK1827- YEBQ	ECK2578 -UNG	0.5916	DNA replication/recombina tion/repair
ECK1878- YECT	ECK1739 -CHO	0.5025	DNA replication/recombina tion/repair
ECK2614- YFJF	ECK2728 -MUTS	0.5293	DNA replication/recombina tion/repair
ECK2373- YFDY	ECK2956 -MUTY	0.6296	DNA replication/recombina

ECK3671- YIDL	ECK0100 -MUTT	0.5899	tion/repair DNA replication/recombina tion/repair	
ECK0465- YBAB	ECK3692 -RECF	0.5632	tion/repair DNA replication/recombina tion/repair	
ECK1228- YCHJ	ECK1468 -FDNG	0.5881	energy production and conversion	anaerobic respiration
ECK1318- YCJF	ECK1468 -FDNG	0.5494	energy production and conversion	anaerobic respiration
ECK2684- YQAA	ECK2478 -HYFB	0.5264	energy production and conversion	anaerobic respiration
ECK2735- YGBN	ECK1472 -ADHP	0.6138	energy production and conversion	anaerobic respiration
ECK0814- YBIY	ECK0887 -DMSC	0.5083	energy production and conversion	anaerobic respiration
ECK1169- YCGN	ECK0887 -DMSC	0.5268	energy production and conversion	anaerobic respiration
ECK0679	ECK0885 -DMSA	0.5226	energy production and conversion	anaerobic respiration
ECK0958- YCCW	ECK0885 -DMSA	0.5006	energy production and conversion	anaerobic respiration
ECK1036- YCEK	ECK0885 -DMSA	0.6281	energy production and conversion	anaerobic respiration
ECK0996- YMDF	ECK0885 -DMSA	0.5395	energy production and conversion	anaerobic respiration
ECK3671- YIDL	ECK3886 -FDOH	0.502	energy production and conversion	anaerobic respiration
ECK3266- YRDA	ECK3886 -FDOH	0.506	energy production and conversion	anaerobic respiration
ECK4250- YJGN	ECK3732 -ATPI	0.521	energy production and conversion	ATP synthesis
ECK1400- YDBD	ECK1744 -ASTD	0.5112	energy production and conversion	dehydrogenase
ECK4213- YTFC	ECK1518 -SAD	0.5063	energy production and conversion	dehydrogenase
ECK0934- YCBV	ECK2485 -HYFI	0.501	energy production and conversion	hydrogenase
ECK2763- YGCP	ECK0842 -NFSA	0.5339	energy production and conversion	nitroreductase
ECK2892- YGFY	ECK2278 -NUOF	0.561	energy production and conversion	oxidoreducatas e activity
ECK0857- YBJQ	ECK2706 -NORW	0.5414	energy production and conversion	oxidoreducatas e activity
ECK3033- YQIC	ECK3836 -FRE	0.5844	energy production and conversion	oxidoreducatas e activity
ECK3654- YICN	ECK3682 -CBRA	0.5298	energy production and conversion	oxidoreductase
ECK4269- YJHB	ECK3682 -CBRA	0.5104	energy production and conversion	oxidoreductase

ECK1036- YCEK	ECK0863 -HCR	0.6613	energy production and conversion	oxidoreductase
ECK1664- YDHS	ECK0863 -HCR	0.5112	energy production and conversion	oxidoreductase
ECK0793- YBIX	ECK0864 -HCP	0.5263	energy production and conversion	oxidoreductase
ECK1022- YCDZ	ECK0864 -HCP	0.5325	energy production and conversion	oxidoreductase
ECK3266- YRDA	ECK1095 -NDH	0.53	energy production and conversion	oxidoreductase
ECK0690- YBFB	ECK0423 -CYOD	0.5379	energy production and conversion	
ECK0669- YBFM	ECK0423 -CYOD	0.5518	energy production and conversion	
ECK0868- YBJX	ECK0423 -CYOD	0.5622	energy production and conversion	
ECK1086- YCFH	ECK0423 -CYOD	0.5938	energy production and conversion	
ECK0690- YBFB	ECK0424 -CYOC	0.5187	energy production and conversion	
ECK0669- YBFM	ECK0424 -CYOC	0.536	energy production and conversion	
ECK0868- YBJX	ECK0424 -CYOC	0.559	energy production and conversion	
ECK1086- YCFH	ECK0424 -CYOC	0.5947	energy production and conversion	
ECK0669- YBFM	ECK0425 -CYOB	0.5258	energy production and conversion	
ECK0868- YBJX	ECK0425 -CYOB	0.5503	energy production and conversion	
ECK1086- YCFH	ECK0425 -CYOB	0.5496	energy production and conversion	
ECK0983- YCCM	ECK0970 -APPB	0.5508	energy production and conversion	
ECK1214- YCHO	ECK0970 -APPB	0.5151	energy production and conversion	
ECK1253- YCIG	ECK0970 -APPB	0.5137	energy production and conversion	
ECK1245- YCII	ECK0970 -APPB	0.5391	energy production and conversion	
ECK1466- YDDL	ECK0970 -APPB	0.5687	energy production and conversion	
ECK1443- YNCB	ECK0970 -APPB	0.5332	energy production and conversion	
ECK1578- YNFB	ECK0970 -APPB	0.5005	energy production and conversion	
ECK0868- YBJX	ECK0426 -CYOA	0.5792	energy production and conversion	
ECK1086- YCFH	ECK0426 -CYOA	0.5724	energy production and conversion	
ECK1688-	ECK0426	0.5533	energy production	

YDIM	-CYOA			and conversion	
ECK1253-	ECK0722	0.5272		energy production	
YCIG	-CYDB			and conversion	
ECK2892-	ECK2271	0.5027		energy production	
YGFY	-NUOM			and conversion	
ECK2892-	ECK2279	0.516		energy production	
YGFY	-NUOE			and conversion	
ECK3592-	ECK4072	0.6952		energy production	
YIBL	-FDHF			and conversion	
ECK1040-	ECK0964	0.5315		energy production	
YCEA	-HYAB			and conversion	
ECK1090-	ECK0964	0.5035		energy production	
YCFL	-HYAB			and conversion	
ECK1228-	ECK0964	0.6249		energy production	
YCHJ	-HYAB			and conversion	
ECK1318-	ECK0964	0.5233		energy production	
YCJF	-HYAB			and conversion	
ECK3562-	ECK0964	0.6035		energy production	
YSAA	-HYAB			and conversion	
ECK2965-	ECK3519	0.5081		energy production	
YGHF	-YHJQ			and conversion	
ECK3865-	ECK3519	0.5027		energy production	
YIHL	-YHJQ			and conversion	
ECK2333-	ECK2050	0.5025		Extracellular	colanic acid
YFCV	-WCAD			structures/proteins	biosynthesis
ECK1090-	ECK1011	0.5597		Extracellular	exopolysachhar
YCFL	-PGAD			structures/proteins	ide production
ECK1228-	ECK1011	0.5896		Extracellular	exopolysachhar
YCHJ	-PGAD			structures/proteins	ide production
ECK1267-	ECK1011	0.588		Extracellular	exopolysachhar
YCIN	-PGAD			structures/proteins	ide production
ECK1931-	ECK2044	0.5411		Extracellular	exopolysachhar
YEDL	-WCAI			structures/proteins	ide production
ECK0705-	ECK1012	0.5186		Extracellular	exopolysachhar
YBGO	-PGAC			structures/proteins	ide production
ECK1040-	ECK1012	0.5554		Extracellular	exopolysachhar
YCEA	-PGAC			structures/proteins	ide production
ECK1090-	ECK1012	0.5144		Extracellular	exopolysachhar
YCFL	-PGAC			structures/proteins	ide production
ECK1228-	ECK1012	0.572		Extracellular	exopolysachhar
YCHJ	-PGAC			structures/proteins	ide production
ECK1312-	ECK1012	0.5362		Extracellular	exopolysachhar
YCJU	-PGAC			structures/proteins	ide production
ECK0691-	ECK1013	0.5579		Extracellular	exopolysachhar
YBFO	-PGAB			structures/proteins	ide production
ECK1169-	ECK1013	0.5817		Extracellular	exopolysachhar
YCGN	-PGAB			structures/proteins	ide production
ECK1032-	ECK1013	0.5546		Extracellular	exopolysachhar
YMDC	-PGAB			structures/proteins	ide production
ECK0679	ECK1170	0.5321		Extracellular	secreted factor
	-HLYE			structures/proteins	

ECK2684- YQAA	ECK1170 -HLYE	0.506	Extracellular structures/proteins	secreted factor
ECK2002- YEEA	ECK2456 -EUTP	0.6084	general function prediction only	
ECK1090- YCFL	ECK0291 -MATC	0.5223	general function prediction only	
ECK1136- YMFR	ECK0291 -MATC	0.5595	general function prediction only	
ECK0769- YBHK	ECK1266 -SOHB	0.5199	general function prediction only	
ECK1074- YCED	ECK1266 -SOHB	0.51	general function prediction only	
ECK1042- YCEJ	ECK1266 -SOHB	0.5209	general function prediction only	
ECK2319- YFCL	ECK2446 -EUTA	0.5133	general function prediction only	
ECK1451- YDCD	ECK2454 -EUTT	0.6286	general function prediction only	
ECK0373- YAIW	ECK2455 -EUTQ	0.508	general function prediction only	
ECK0857- YBJQ	ECK2455 -EUTQ	0.5092	general function prediction only	
ECK4004- YJAB	ECK2455 -EUTQ	0.5025	general function prediction only	
ECK3472- YHII	ECK0533 -REND	0.5999	general function prediction only	
ECK1167- YCGL	ECK0617 -CRCB	0.5008	general function prediction only	
ECK0809- YBIS	ECK0812 -YBIV	0.5241	general function prediction only	
ECK1374- YDBK	ECK0888 -YCAC	0.5304	general function prediction only	
ECK1428- YDCN	ECK0888 -YCAC	0.5475	general function prediction only	
ECK1689- YDIN	ECK0888 -YCAC	0.5297	general function prediction only	
ECK0943- YMBA	ECK0888 -YCAC	0.5408	general function prediction only	
ECK3033- YQIC	ECK0888 -YCAC	0.554	general function prediction only	
ECK0284- YAGS	ECK1900 -FTNB	0.5266	general function prediction only	
ECK3131- YRAI	ECK3871 -YIHQ	0.5055	general function prediction only	
ECK3520- YHJR	ECK3559 -BAX	0.6057	general function prediction only	
ECK3563- YIAJ	ECK3559 -BAX	0.5864	general function prediction only	
ECK3563- YIAJ	ECK4085 -PHNP	0.6065	general function prediction only	
ECK3503- YIAJ	ECK0138	0.5646	general function	

YHJA	-HTRE		prediction only
ECK4181-	ECK3313	0.5854	general function
YJFM	-GSPE		prediction only
ECK0958-	ECK2186	0.521	general function
YCCW	-CCMH		prediction only
ECK1467-	ECK2186	0.559	general function
YDDG	-CCMH		prediction only
ECK1549-	ECK2186	0.5054	general function
YDFR	-CCMH		prediction only
ECK2684-	ECK2186	0.527	general function
YQAA	-CCMH		prediction only
ECK3551-	ECK3740	0.6044	general function
YIAA	-RAVA		prediction only
ECK3677-	ECK3740	0.5436	general function
YIDE	-RAVA		prediction only
ECK3654-	ECK3571	0.5083	general function
YICN	-YIAR		prediction only
ECK1784-	ECK4295	0.5602	general function
YEAJ	-SGCX		prediction only
ECK1836-	ECK4295	0.5288	general function
YEBW	-SGCX		prediction only
ECK1792-	ECK0012	0.5124	general function
YEAP			prediction only
ECK0983-	ECK2901	0.6153	general function
YCCM	-VISC		prediction only
ECK1495-	ECK2901	0.5015	general function
YDEQ	-VISC		prediction only
ECK3028-	ECK2901	0.5491	general function
YGIB	-VISC		prediction only
ECK0943-	ECK2901	0.502	general function
YMBA	-VISC		prediction only
ECK3033-	ECK2901	0.7757	general function
YQIC	-VISC		prediction only
ECK2764-	ECK1610	0.5666	general function
YGCQ	-UIDC		prediction only
ECK2762-	ECK1830	0.5057	general function
YGCO	-PROQ		prediction only
ECK0220-	ECK1575	0.5273	general function
YAFV	-RSPB		prediction only
ECK2320-	ECK1575	0.513	general function
YFCM	-RSPB		prediction only
ECK1036-	ECK0544	0.5723	general function
YCEK	-NMPC		prediction only
ECK1789-	ECK0544	0.5514	general function
YEAN	-NMPC		prediction only
ECK0929-	ECK1143	0.5691	general function
YCBQ	-STFE		prediction only
ECK1245-	ECK1143	0.5141	general function
YCII	-STFE		prediction only
ECK1343-	ECK1143	0.5406	general function
YDAQ	-STFE		prediction only

ECK0302- YKGI	ECK4264 -INTB	0.5096	general function prediction only	
ECK3134- YRAL	ECK3869 -YIHO	0.5444	general function prediction only	
ECK3824- YSGA	ECK3869 -YIHO	0.583	general function prediction only	
ECK3671- YIDL	ECK3399 -BIOH	0.5785	general function prediction only	
ECK2081- YEGR	ECK4086 -PHNO	0.5092	general function prediction only	
ECK2766- YGCS	ECK4086 -PHNO	0.5275	general function prediction only	
ECK2356- YFDS	ECK1991 -YEPP	0.6666	general function prediction only	
ECK1419- YDCH	ECK4047 -APHA	0.5272	general function prediction only	
ECK1954- YEDQ	ECK4047 -APHA	0.5372	general function prediction only	
ECK3430- YRHB	ECK2884 -IDI	0.5611	lipid metabolism	
ECK2893- YGFZ	ECK2310 -ACCD	0.5487	lipid metabolism	
ECK0447- YBAY	ECK0479 -YBAS	0.5947	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2018 -HISH	0.7174	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3949 -ARGC	0.6894	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3161 -ARGG	0.6665	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK4380 -SERB	0.5895	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK1255 -TRPB	0.7306	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3931 -METB	0.634	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2759 -CYSJ	0.6841	metabolism	amino acid biosynthesis & utilization
ECK3526- YHJV	ECK2952 -ANSB	0.5056	metabolism	amino acid biosynthesis & utilization
ECK3757- YIFE	ECK2952 -ANSB	0.5555	metabolism	amino acid biosynthesis &

ECK1048- YCEB	ECK0243 -PROB	0.7279	metabolism	utilization amino acid biosynthesis & utilization
ECK0404- YAJD	ECK1257 -TRPD	0.5587	metabolism	amino acid biosynthesis & utilization
ECK1428- YDCN	ECK1257 -TRPD	0.5225	metabolism	amino acid biosynthesis & utilization
ECK1840- YEBY	ECK1257 -TRPD	0.5098	metabolism	amino acid biosynthesis & utilization
ECK1370- YNAE	ECK1257 -TRPD	0.5005	metabolism	amino acid biosynthesis & utilization
ECK1443- YNCB	ECK1257 -TRPD	0.5419	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2909 -SERA	0.6777	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3948 -ARGE	0.6823	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0003 -THRB	0.7479	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0244 -PROA	0.7134	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0898 -SERC	0.5589	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3000 -METC	0.529	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3951 -ARGH	0.6915	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3933 -METF	0.657	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0899 -AROA	0.6133	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2014 -HISG	0.6777	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2596 -PHEA	0.6275	metabolism	amino acid biosynthesis &

ECK1048- YCEB	ECK4005 -META	0.6545	metabolism	utilization amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0073 -LEUD	0.6773	metabolism	amino acid biosynthesis & utilization
ECK3208- YHCF	ECK0274 -ARGF	0.5809	metabolism	amino acid biosynthesis & utilization
ECK3241- YHDH	ECK0274 -ARGF	0.6005	metabolism	amino acid biosynthesis & utilization
ECK2892- YGFY	ECK0919 -ASPC	0.508	metabolism	amino acid biosynthesis & utilization
ECK1021- YCDY	ECK2015 -HISD	0.5429	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2015 -HISD	0.6427	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2323 -AROC	0.6038	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2597 -TYRA	0.7072	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3268 -AROE	0.6056	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3762 -ILVE	0.6924	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0074 -LEUC	0.7421	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0381 -PROC	0.7158	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2016 -HISC	0.7057	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3764 -ILVA	0.6682	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0076 -LEUA	0.6446	metabolism	amino acid biosynthesis & utilization
ECK2895- YQFB	ECK0383 -AROL	0.529	metabolism	amino acid biosynthesis &

ECK1048- YCEB	ECK2020 -HISF	0.7296	metabolism	utilization amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2814 -ARGA	0.7092	metabolism	amino acid biosynthesis & utilization
ECK1021- YCDY	ECK3823 -METE	0.5134	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3823 -METE	0.6576	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK1254 -TRPA	0.6647	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2021 -HISI	0.7064	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2836 -LYSA	0.6948	metabolism	amino acid biosynthesis & utilization
ECK3645- YICH	ECK4046 -TYRB	0.6246	metabolism	amino acid biosynthesis & utilization
ECK4182- YJFC	ECK4046 -TYRB	0.5076	metabolism	amino acid biosynthesis & utilization
ECK3551- YIAA	ECK0079 -ILVI	0.5732	metabolism	amino acid biosynthesis & utilization
ECK3677- YIDE	ECK0079 -ILVI	0.5054	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2017 -HISB	0.673	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3376 -AROB	0.6446	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK2019 -HISA	0.6948	metabolism	amino acid biosynthesis & utilization
ECK1316- YCJW	ECK1256 -TRPC	0.5919	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0004 -THRC	0.6778	metabolism	amino acid biosynthesis & utilization
ECK2764- YGCQ	ECK0662 -ASNB	0.5168	metabolism	amino acid biosynthesis &

ECK1048- YCEB	ECK3763 -ILVD	0.6885	metabolism	utilization amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3766 -ILVC	0.6641	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK3950 -ARGB	0.6801	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0002 -THRA	0.5775	metabolism	amino acid biosynthesis & utilization
ECK2066- YEGK	ECK2656 -GABT	0.6187	metabolism	amino acid biosynthesis & utilization
ECK1021- YCDY	ECK0075 -LEUB	0.5126	metabolism	amino acid biosynthesis & utilization
ECK1048- YCEB	ECK0075 -LEUB	0.7047	metabolism	amino acid biosynthesis & utilization
ECK1022- YCDZ	ECK1392 -PAAH	0.5021	metabolism	aromatic compound metabolism
ECK1228- YCHJ	ECK1392 -PAAH	0.6331	metabolism	aromatic compound metabolism
ECK2066- YEGK	ECK0345 -MHPB	0.7922	metabolism	aromatic compound metabolism
ECK0404- YAJD	ECK1853 -ZWF	0.5524	metabolism	carbohydrate metabolism glycolysis + Entner- Doudorof pathway, pentose phospahte pathway
ECK1723- YNIA	ECK1853 -ZWF	0.512	metabolism	carbohydrate metabolism glycolysis + Entner- Doudorof pathway, pentose phospahte pathway
ECK3472- YHII	ECK3917 -GLPX	0.5061	metabolism	carbohydrate metabolism glycolysis, gluconeog enesis
ECK0462- YBAN	ECK0113 -ACEE	0.5417	metabolism	carbohydrate metabolism
ECK0835-	ECK0113	0.6238	metabolism	carbohydrate

YBJJ	-ACEE			metabolism
ECK0983-	ECK0113	0.6773	metabolism	carbohydrate
YCCM	-ACEE			metabolism
ECK1748-	ECK0113	0.5672	metabolism	carbohydrate
YDJX	-ACEE			metabolism
ECK2614-	ECK0113	0.5146	metabolism	carbohydrate
YFJF	-ACEE			metabolism
ECK1031-	ECK0113	0.5399	metabolism	carbohydrate
YMDB	-ACEE			metabolism
ECK1451-	ECK0745	0.5929	metabolism	carbohydrate
YDCD	-GALM			metabolism
ECK1903-	ECK0745	0.5712	metabolism	carbohydrate
YECR	-GALM			metabolism
ECK0943-	ECK2382	0.5024	metabolism	carbohydrate
YMBA	-FRYC			metabolism
ECK1402-	ECK2382	0.5473	metabolism	carbohydrate
YNBB	-FRYC			metabolism
ECK3033-	ECK2382	0.5584	metabolism	carbohydrate
YQIC	-FRYC			metabolism
ECK3173-	ECK3689	0.5503	metabolism	carbohydrate
YHBE	-YIDA			metabolism
ECK4186-	ECK3689	0.5784	metabolism	carbohydrate
YJFP	-YIDA			metabolism
ECK4216-	ECK3689	0.552	metabolism	carbohydrate
YTFM	-YIDA			metabolism
ECK2925-	ECK0185	0.5495	metabolism	carbohydrate
YGGD	-LDCC			metabolism
ECK1773-	ECK3560	0.6148	metabolism	carbohydrate
YDJK	-MALS			metabolism
ECK3677-	ECK3066	0.6005	metabolism	carbohydrate
YIDE	-EBGA			metabolism
ECK3179-	ECK3417	0.5135	metabolism	carbohydrate
YRBA	-GLGX			metabolism
ECK3551-	ECK3418	0.5385	metabolism	carbohydrate
YIAA	-GLGB			metabolism
ECK3677-	ECK3418	0.5052	metabolism	carbohydrate
YIDE	-GLGB			metabolism
ECK1036-	ECK1408	0.5061	metabolism	carbohydrate
YCEK	-ALDA			metabolism
ECK2735-	ECK2928	0.5179	metabolism	carbohydrate
YGBN	-CMTA			metabolism
ECK3757-	ECK4190	0.5116	metabolism	carbohydrate
YIFE	-ULAB			metabolism
ECK4267-	ECK4193	0.5551	metabolism	carbohydrate
YJGZ	-ULAE			metabolism
ECK4152-	ECK4194	0.504	metabolism	carbohydrate
YJEM	-ULAF			metabolism
ECK3091-	ECK4189	0.5371	metabolism	carbohydrate
YQJK	-ULAA			metabolism
ECK4202-	ECK4189	0.5898	metabolism	carbohydrate
YTFB	-ULAA			metabolism

ECK0290- YAGX	ECK0329 -PRPB	0.5086	metabolism	carbohydrate metabolism	
ECK0793- YBIX	ECK1388 -PAAD	0.5763	metabolism	carbohydrate metabolism	
ECK1022- YCDZ	ECK1388 -PAAD	0.5536	metabolism	carbohydrate metabolism	
ECK0521- YBCJ	ECK1187 -DHAL	0.5722	metabolism	carbohydrate metabolism	
ECK3369- YHFY	ECK1187 -DHAL	0.5097	metabolism	carbohydrate metabolism	
ECK4284- YJHU	ECK1187 -DHAL	0.5528	metabolism	carbohydrate metabolism	
ECK2872- YQEC	ECK3511 -KDGK	0.5394	metabolism	carbohydrate metabolism	
ECK4103- YJCZ	ECK3939 -PTSA	0.535	metabolism	carbohydrate metabolism	
ECK3671- YIDL	ECK3646 -YICI	0.5072	metabolism	carbohydrate metabolism	
ECK3671- YIDL	ECK3404 -MALP	0.5414	metabolism	carbohydrate metabolism	
ECK3671- YIDL	ECK3361 -FRLD	0.5781	metabolism	carbohydrate metabolism	
ECK3645- YICH	ECK4259 -IDNO	0.5436	metabolism	carbohydrate metabolism	
ECK1140- YCFK	ECK4260 -IDND	0.5058	metabolism	carbohydrate metabolism	
ECK3645- YICH	ECK4260 -IDND	0.5062	metabolism	carbohydrate metabolism	
ECK1933	ECK4261 -IDNK	0.5052	metabolism	carbohydrate metabolism	
ECK1048- YCEB	ECK2458 -MAEB	0.5444	metabolism	central metabolism	gluconeog enesis
ECK2541- YPHB	ECK2090 -FBAB	0.6603	metabolism	central metabolism	glycolysis, gluconeog enesis
ECK0404- YAJD	ECK1271 -ACNA	0.5086	metabolism	central metabolism	TCA
ECK1466- YDDL	ECK1271 -ACNA	0.6352	metabolism	central metabolism	TCA
ECK2615- YFJG	ECK1271 -ACNA	0.5338	metabolism	central metabolism	TCA
ECK1048- YCEB	ECK0709 -GLTA	0.6581	metabolism	central metabolism	TCA
ECK1503- YDEK	ECK0711 -SDHD	0.5455	metabolism	central metabolism	TCA
ECK2892- YGFY	ECK0713 -SDHB	0.5378	metabolism	central metabolism	TCA
ECK1048- YCEB	ECK3947 -PPC	0.5917	metabolism	central metabolism	TCA, gluconeog enesis
ECK0814-	ECK1606	0.5843	metabolism	central	

YBIY	-FUMC			metabolism		
ECK0857-	ECK1606	0.5507	metabolism	central		
YBJQ	-FUMC			metabolism		
ECK1169-	ECK1606	0.5944	metabolism	central		
YCGN	-FUMC			metabolism		
ECK4179-	ECK1606	0.5829	metabolism	central		
YJFK	-FUMC			metabolism		
ECK1503-	ECK0716	0.5169	metabolism	central		
YDEK	-SUCC			metabolism		
ECK2614-	ECK0716	0.5619	metabolism	central		
YFJF	-SUCC			metabolism		
ECK2615-	ECK0716	0.5381	metabolism	central		
YFJG	-SUCC			metabolism		
ECK2892-	ECK0716	0.5332	metabolism	central		
YGFY	-SUCC			metabolism		
ECK1317-	ECK2577	0.6625	metabolism	central		
YCJX	-YFID			metabolism		
ECK1960-	ECK2577	0.6349	metabolism	central		
YEDJ	-YFID			metabolism		
ECK3033-	ECK0954	0.532	metabolism	central		
YQIC	-MGSA			metabolism		
ECK1048-	ECK0764	0.5341	metabolism	cofactor	biotin	
YCEB	-BIOB			biosynthesis		
ECK2925-	ECK0631	0.5643	metabolism	cofactor	cobalini	
YGGD	-COBC			biosynthesis		
ECK1096-	ECK1082	0.5894	metabolism	cofactor	folate	
YCFJ	-PABC			biosynthesis		
ECK2768-	ECK1082	0.5478	metabolism	cofactor	folate	
YGCW	-PABC			biosynthesis		
ECK1048-	ECK2548	0.6479	metabolism	cofactor	folate	
YCEB	-GLYA			biosynthesis		
ECK2892-	ECK0621	0.5902	metabolism	cofactor	liptoate	
YGFY	-LIPA			biosynthesis	biosynthes	
					is	
ECK2614-	ECK0654	0.5874	metabolism	cofactor	menaquin	
YFJF	-UBIF			biosynthesis	one	
ECK2892-	ECK0654	0.5482	metabolism	cofactor	menaquin	
YGFY	-UBIF			biosynthesis	one	
ECK4394-	ECK0654	0.5528	metabolism	cofactor	menaquin	
YJJY	-UBIF			biosynthesis	one	
ECK3428-	ECK2255	0.5046	metabolism	cofactor	menaquin	
YHHZ	-MENC			biosynthesis	one	
ECK2892-	ECK3827	0.559	metabolism	cofactor	menaquin	
YGFY	-UBIE			biosynthesis	one	
ECK4394-	ECK3827	0.5474	metabolism	cofactor	menaquin	
YJJY	-UBIE			biosynthesis	one	
ECK4076-	ECK2258	0.5089	metabolism	cofactor	menaquin	
YJCS	-MEND			biosynthesis	one	
ECK0835-	ECK4031	0.5682	metabolism	cofactor	menaquin	
YBJJ	-UBIC			biosynthesis	one	
ECK0983-	ECK4031	0.5557	metabolism	cofactor	menaquin	

YCCM	-UBIC				biosynthesis	one
ECK1253-	ECK4031	0.5063	metabolism		cofactor	menaquin
YCIG	-UBIC				biosynthesis	one
ECK1245-	ECK4031	0.5143	metabolism		cofactor	menaquin
YCII	-UBIC				biosynthesis	one
ECK2614-	ECK4031	0.5215	metabolism		cofactor	menaquin
YFJF	-UBIC				biosynthesis	one
ECK3028-	ECK4031	0.5498	metabolism		cofactor	menaquin
YGIB	-UBIC				biosynthesis	one
ECK3159-	ECK4031	0.5449	metabolism		cofactor	menaquin
RIMP	-UBIC				biosynthesis	one
ECK1099-	ECK0415	0.5625	metabolism		cofactor	menaquin
YCFS	-ISPA				biosynthesis	one
ECK2218-	ECK1632	0.5125	metabolism		cofactor	pyridoxine
YFAP	-PDXY				biosynthesis	
ECK1048-	ECK2562	0.533	metabolism		cofactor	pyridoxine
YCEB	-PDXJ				biosynthesis	
ECK0890-	ECK0440	0.5103	metabolism		cofactor	thiamin
YCAM	-COF				biosynthesis	
ECK1148-	ECK0440	0.5227	metabolism		cofactor	thiamin
YCGX	-COF				biosynthesis	
ECK1310-	ECK0440	0.5207	metabolism		cofactor	thiamin
YCJS	-COF				biosynthesis	
ECK1440-	ECK0440	0.5084	metabolism		cofactor	thiamin
YDCY	-COF				biosynthesis	
ECK1665-	ECK0440	0.5712	metabolism		cofactor	thiamin
YDHT	-COF				biosynthesis	
ECK2551-	ECK0440	0.5067	metabolism		cofactor	thiamin
YFHA	-COF				biosynthesis	
ECK2911-	ECK0440	0.519	metabolism		cofactor	thiamin
YQFE	-COF				biosynthesis	
ECK1040-	ECK1120	0.5488	metabolism		cofactor	thiamin
YCEA	-NUDJ				biosynthesis	
ECK0428-	ECK1092	0.5526	metabolism		cofactor	thiamin
YAJG	-THIK				biosynthesis	
ECK1048-	ECK3984	0.5861	metabolism		cofactor	thiamine
YCEB	-THIF				biosynthesis	
ECK1048-	ECK2096	0.5971	metabolism		cofactor	thiamine
YCEB	-THID				biosynthesis	
ECK0281-	ECK1766	0.5381	metabolism		cofactor	
YAGP	-PNCA				biosynthesis	
ECK1048-	ECK3356	0.5482	metabolism		cofactor	
YCEB	-CYSG				biosynthesis	
ECK0228-	ECK0068	0.5545	metabolism		cofactor	
YAFL	-THIP				biosynthesis	
ECK1960-	ECK0009	0.5068	metabolism		cofactor	
YEDJ	-MOG				biosynthesis	
ECK1048-	ECK2572	0.6389	metabolism		cofactor	
YCEB	-NADB				biosynthesis	
ECK1048-	ECK0053	0.5131	metabolism		cofactor	
YCEB	-PDXA				biosynthesis	

ECK1048- YCEB	ECK0108 -NADC	0.6093	metabolism	cofactor biosynthesis
ECK1048- YCEB	ECK0739 -NADA	0.6431	metabolism	cofactor biosynthesis
ECK2402- YFEN	ECK0408 -RIBD	0.6197	metabolism	cofactor biosynthesis
ECK3147- YHBV	ECK0049 -FOLA	0.5018	metabolism	cofactor biosynthesis
ECK2109- YEHI	ECK2713 -HYCH	0.514	metabolism	fermentation
ECK2619- YFJH	ECK2713 -HYCH	0.5808	metabolism	fermentation
ECK0236- YKFJ	ECK0240 -FRSA	0.5239	metabolism	fermentation
ECK1728- YDJO	ECK0510 -ALLD	0.5161	metabolism	inorganic ion metabolism
ECK2897- YGFF	ECK0337 -CYNS	0.57	metabolism	inorganic ion metabolism
ECK3210- YHCH	ECK0337 -CYNS	0.5012	metabolism	inorganic ion metabolism
ECK1048- YCEB	ECK2747 -CYSD	0.5192	metabolism	inorganic ion metabolism
ECK1048- YCEB	ECK2745 -CYSC	0.5083	metabolism	inorganic ion metabolism
ECK2349- YFDL	ECK2517 -SSEA	0.5195	metabolism	inorganic ion metabolism
ECK2627- YFJO	ECK2519 -SSEB	0.5256	metabolism	inorganic ion metabolism
ECK1048- YCEB	ECK2758 -CYSI	0.5862	metabolism	inorganic ion metabolism
ECK3587- YIBI	ECK4087 -PHNN	0.5072	metabolism	inorganic ion metabolism
ECK3666- YIDG	ECK3656 -ADE	0.5794	metabolism	nucleotide biosynthesis
ECK4021- YJBH	ECK3988 -NUDC	0.5304	metabolism	nucleotide biosynthesis
ECK1048- YCEB	ECK0515 -PURK	0.5569	metabolism	nucleotide biosynthesis
ECK0677- YBFP	ECK1977 -AMN	0.5009	metabolism	nucleotide biosynthesis
ECK0758- YBHH	ECK1977 -AMN	0.5859	metabolism	nucleotide biosynthesis
ECK1428- YBHH	ECK1977 -AMN	0.5036	metabolism	nucleotide

YDCN	-AMN				biosynthesis
ECK2850-	ECK1977	0.5088	metabolism		nucleotide
YGEH	-AMN				biosynthesis
ECK1443-	ECK1977	0.5158	metabolism		nucleotide
YNCB	-AMN				biosynthesis
ECK1723-	ECK1977	0.5458	metabolism		nucleotide
YNIA	-AMN				biosynthesis
ECK3671-	ECK3825	0.5034	metabolism		nucleotide
YIDL	-UDP				biosynthesis
ECK1048-	ECK0033	0.7248	metabolism		nucleotide
YCEB	-CARA				biosynthesis
ECK1048-	ECK3998	0.6266	metabolism		nucleotide
YCEB	-PURH				biosynthesis
ECK2341-	ECK2879	0.5462	metabolism		nucleotide
YFDC	-GUAD				biosynthesis
ECK1048-	ECK2472	0.6074	metabolism		nucleotide
YCEB	-PURC				biosynthesis
ECK0868-	ECK1001	0.5176	metabolism		nucleotide
YBJX	-RUTC				catabolism
ECK0983-	ECK1001	0.6107	metabolism		nucleotide
YCCM	-RUTC				catabolism
ECK0955-	ECK1001	0.5217	metabolism		nucleotide
YCCT	-RUTC				catabolism
ECK0705-	ECK1002	0.5632	metabolism		nucleotide
YBGO	-RUTB				catabolism
ECK0679	ECK1292	0.515	metabolism		polyamine
	-PUUA				catabolism
ECK1036-	ECK1292	0.513	metabolism		polyamine
YCEK	-PUUA				catabolism
ECK2853-	ECK1292	0.513	metabolism		polyamine
YGEK	-PUUA				catabolism
ECK3476-	ECK3382	0.5026	metabolism		DNA
YHIM	-HOFM				catabolism
ECK3266-	ECK3382	0.5938	metabolism		DNA
YRDA	-HOFM				catabolism
ECK1931-	ECK1106	0.5281	post-translational		deacetylation
YEDL	-COBB		modification/enzyme		
			inhibitor		
ECK2755-	ECK1504	0.5639	post-translational		kinase
YGCL	-LSRK		modification/enzyme		
			inhibitor		
ECK2092-	ECK4008	0.514	post-translational	kinase/phospha	
YEGU	-ACEK		modification/enzyme	tase	
			inhibitor		
ECK1545-	ECK1500	0.5841	post-translational	serine/threonin	
YNFN	-HIPA		modification/enzyme	e kinase	
			inhibitor		
ECK2661-	ECK1837	0.531	post-translational	serine/threonin	
YGAV	-PPHA		modification/enzyme	e phosphatase	
			inhibitor		
ECK2134-	ECK2054	0.5623	post-translational	tyrosine kinase	

YOHJ	-WZC		modification/enzyme inhibitor	
ECK1466-	ECK0893	0.5053	post-translational modification/enzyme inhibitor	
YDDL	-PFLA			
ECK0286-	ECK0990	0.5964	post-translational modification/enzyme inhibitor	
YAGU	-CBPM			
ECK2337-	ECK1627	0.5475	post-translational modification/enzyme inhibitor	
YFCZ	-RSXG			
ECK0437-	ECK0600	0.5622	post-translational modification/enzyme inhibitor	
YBAW	-AHPF			
ECK0572-	ECK0600	0.5152	post-translational modification/enzyme inhibitor	
YBDJ	-AHPF			
ECK2684-	ECK0600	0.5821	post-translational modification/enzyme inhibitor	
YQAA	-AHPF			
ECK1497-	ECK1631	0.5065	post-translational modification/enzyme inhibitor	
YDES	-GST			
ECK2614-	ECK1624	0.5071	post-translational modification/enzyme inhibitor	
YFJF	-RSXB			
ECK3266-	ECK3324	0.5127	post-translational modification/enzyme inhibitor	
YRDA	-BFD			
ECK2871-	ECK3192	0.5312	post-translational modification/enzyme inhibitor	
YQEB	-HPF			
ECK3805-	ECK4248	0.506	post-translational modification/enzyme inhibitor	
YIGA	-RRAB			
ECK0813-	ECK1963	0.5505	protein quality control	chaperone
YBIW	-HCHA			
ECK1960-	ECK1963	0.5197	protein quality control	chaperone
YEDJ	-HCHA			
ECK2101-	ECK1963	0.5323	protein quality control	chaperone
YEHA	-HCHA			
ECK3586-	ECK4069	0.5224	protein quality control	chaperone
YIBH	-NRFG			
ECK3551-	ECK3678	0.5104	protein quality control	chaperone
YIAA	-IBPB			
ECK3458-	ECK3493	0.5323	protein quality control	chaperone
YHHT	-HDEB			
ECK1571-	ECK2984	0.5175	protein quality control	chaperone
YDFE	-HYBG			
ECK3945-	ECK4067	0.5231	protein quality control	chaperonin
YIJO	-NRFE			

ECK1664- YDHS	ECK2187 -CCMG	0.5546	protein quality control	disulphide reductase
ECK0319- YAHG	ECK2967 -PPPA	0.5075	protein quality control	endopeptidase
ECK1534- YDFZ	ECK2712 -HYCI	0.5598	protein quality control	endopeptidase
ECK2614- YFJF	ECK2903 -PEPP	0.5982	protein quality control	exopeptidase
ECK2892- YGFY	ECK2903 -PEPP	0.5451	protein quality control	exopeptidase
ECK3159- RIMP	ECK2903 -PEPP	0.5228	protein quality control	exopeptidase
ECK3033- YQIC	ECK2903 -PEPP	0.5477	protein quality control	exopeptidase
ECK1253- YCIG	ECK1924 -FLIS	0.5039	protein quality control	flagella
ECK1245- YCII	ECK1924 -FLIS	0.5464	protein quality control	flagella
ECK1400- YDBD	ECK1924 -FLIS	0.5589	protein quality control	flagella
ECK1428- YDCN	ECK1924 -FLIS	0.527	protein quality control	flagella
ECK1495- YDEQ	ECK1924 -FLIS	0.5053	protein quality control	flagella
ECK2614- YFJF	ECK1924 -FLIS	0.5396	protein quality control	flagella
ECK2892- YGFY	ECK1924 -FLIS	0.5363	protein quality control	flagella
ECK1443- YNCB	ECK1924 -FLIS	0.5783	protein quality control	flagella
ECK2614- YFJF	ECK3401 -NFUA	0.5761	protein quality control	maturation
ECK1031- YMDB	ECK3401 -NFUA	0.5178	protein quality control	maturation
ECK4163- YJEF	ECK1125 -LIT	0.5416	protein quality control	peptidase
ECK2639- YFJX	ECK3924 -HSLV	0.5473	protein quality control	protease
ECK1905- YECH	ECK1113 -PEPT	0.5999	protein quality control	protease
ECK2333- YFCV	ECK2788 -SYD	0.5639	protein quality control	
ECK0814- YBIY	ECK3247 -DUSB	0.6544	RNA/tRNA/rRNA processing- modification	tRNA modification
ECK0857- YBJQ	ECK3247 -DUSB	0.6405	RNA/tRNA/rRNA processing- modification	tRNA modification
ECK1169- YCGN	ECK3247 -DUSB	0.5443	RNA/tRNA/rRNA processing- modification	tRNA modification

ECK4179- YJFK	ECK3247 -DUSB	0.5789	RNA/tRNA/rRNA processing- modification	tRNA modification
ECK3367- YHFW	ECK0399 -QUEA	0.5006	RNA/tRNA/rRNA processing- modification	
ECK3757- YIFE	ECK0399 -QUEA	0.5774	RNA/tRNA/rRNA processing- modification	
ECK3159- RIMP	ECK3633 -RPH	0.5786	RNA/tRNA/rRNA processing- modification	
ECK2373- YFDY	ECK3733 -GIDB	0.727	RNA/tRNA/rRNA processing- modification	
ECK1176- YCGB	ECK2740 -TRUD	0.5224	RNA/tRNA/rRNA processing- modification	
ECK0705- YBGO	ECK0939 -RLML	0.5169	RNA/tRNA/rRNA processing- modification	
ECK1040- YCEA	ECK0939 -RLML	0.547	RNA/tRNA/rRNA processing- modification	
ECK1228- YCHJ	ECK0939 -RLML	0.736	RNA/tRNA/rRNA processing- modification	
ECK1267- YCIN	ECK0939 -RLML	0.5987	RNA/tRNA/rRNA processing- modification	
ECK3562- YSAA	ECK0939 -RLML	0.5304	RNA/tRNA/rRNA processing- modification	
ECK0630- YBEB	ECK0620 -TATE	0.5073	secretion	
ECK0314- YAHB	ECK0313 -YAHA	0.5208	signal transduction	c-di-GMP phosphodiester ase
ECK2634- YFJU	ECK0313 -YAHA	0.6608	signal transduction	c-di-GMP phosphodiester ase
ECK1140- YCFK	ECK2212 -ATOS	0.5595	signal transduction	sensor kinase
ECK2002- YEEA	ECK2780 -BARA	0.5693	signal transduction	sensor kinase
ECK1404- YNBD	ECK1889 -CHEA	0.6046	signal transduction	sensor kinase
ECK0814- YBIY	ECK4118 -DCUS	0.5349	signal transduction	sensor kinase
ECK0404- YAJD	ECK1216 -NARX	0.5522	signal transduction	sensor kinase

ECK0462-	ECK1216	0.6288	signal transduction	sensor kinase
YBAN	-NARX			
ECK0835-	ECK1216	0.629	signal transduction	sensor kinase
YBJJ	-NARX			
ECK0983-	ECK1216	0.6629	signal transduction	sensor kinase
YCCM	-NARX			
ECK1253-	ECK1216	0.5918	signal transduction	sensor kinase
YCIG	-NARX			
ECK1400-	ECK1216	0.5006	signal transduction	sensor kinase
YDBD	-NARX			
ECK1495-	ECK1216	0.5488	signal transduction	sensor kinase
YDEQ	-NARX			
ECK1748-	ECK1216	0.6401	signal transduction	sensor kinase
YDJX	-NARX			
ECK2614-	ECK1216	0.6538	signal transduction	sensor kinase
YFJF	-NARX			
ECK2615-	ECK1216	0.5708	signal transduction	sensor kinase
YFJG	-NARX			
ECK2868-	ECK1216	0.5046	signal transduction	sensor kinase
YGEY	-NARX			
ECK2892-	ECK1216	0.5109	signal transduction	sensor kinase
YGFY	-NARX			
ECK0943-	ECK1216	0.5206	signal transduction	sensor kinase
YMBA	-NARX			
ECK1031-	ECK1216	0.5482	signal transduction	sensor kinase
YMDB	-NARX			
ECK1363-	ECK1216	0.5138	signal transduction	sensor kinase
YNAK	-NARX			
ECK1443-	ECK1216	0.5429	signal transduction	sensor kinase
YNCB	-NARX			
ECK3033-	ECK1216	0.5727	signal transduction	sensor kinase
YQIC	-NARX			
ECK2931-	ECK4391	0.5092	signal transduction	sensor kinase
YGGG	-CREC			
ECK1467-	ECK1604	0.514	signal transduction	sensor kinase
YDDG	-RSTB			
ECK1549-	ECK1604	0.5589	signal transduction	sensor kinase
YDFR	-RSTB			
ECK2329-	ECK1604	0.6011	signal transduction	sensor kinase
YFCR	-RSTB			
ECK2101-	ECK3659	0.5292	signal transduction	sensor kinase
YEHA	-UHPB			
ECK3874-	ECK4105	0.5078	signal transduction	sensor kinase
YIHT	-BASS			
ECK0705-	ECK0562	0.6783	signal transduction	sensor kinase
YBGO	-CUSS			
ECK1169-	ECK0562	0.5617	signal transduction	sensor kinase
YCGN	-CUSS			
ECK1228-	ECK0562	0.5098	signal transduction	sensor kinase
YCHJ	-CUSS			
ECK3562-	ECK0562	0.5226	signal transduction	sensor kinase

YSAA	-CUSS				
ECK0896-	ECK0601	0.5169	signal transduction	stress response	
YCAO	-USPG				
ECK0134-	ECK2681	0.5494	signal transduction		
YADC	-LUXS				
ECK0190-	ECK0246	0.5256	toxin-antitoxin		
YAEQ	-YKFI				
ECK3757-	ECK4143	0.5148	toxin-antitoxin		
YIFE	-ECNB				
ECK1499-	ECK1412	0.5724	toxin-antitoxin		
YNEL	-HOKB				
ECK2064-	ECK1998	0.566	toxin-antitoxin		
YEGI	-YEEV				
ECK2752-	ECK1998	0.6293	toxin-antitoxin		
YGCI	-YEEV				
ECK3883-	ECK2642	0.5114	toxin-antitoxin		
YIIF	-YFJZ				
ECK0236-	ECK0188	0.5051	transcription	antitermination	
YKFJ	-ROF				
ECK1267-	ECK3660	0.5356	transcription	transcriptional activator	
YCIN	-UHPA				
ECK3367-	ECK4109	0.5078	transcription	transcriptional activator	
YHFW	-ADIY				
ECK1440-	ECK1982	0.5616	transcription	transcriptional activator	
YDCY	-CBL				
ECK1466-	ECK1982	0.5315	transcription	transcriptional activator	
YDDL	-CBL				
ECK1840-	ECK1982	0.5975	transcription	transcriptional activator	
YEBY	-CBL				
ECK1443-	ECK1982	0.5127	transcription	transcriptional activator	
YNCB	-CBL				
ECK1363-	ECK3279	0.5015	transcription	transcriptional activator	
YNAK	-ZNTR				
ECK1228-	ECK3861	0.5094	transcription	transcriptional dual regulator	
YCHJ	-GLNG				
ECK1267-	ECK3861	0.5416	transcription	transcriptional dual regulator	
YCIN	-GLNG				
ECK0404-	ECK1215	0.5522	transcription	transcriptional dual regulator	
YAJD	-NARL				
ECK0462-	ECK1215	0.6147	transcription	transcriptional dual regulator	
YBAN	-NARL				
ECK0835-	ECK1215	0.5617	transcription	transcriptional dual regulator	
YBJJ	-NARL				
ECK0983-	ECK1215	0.5952	transcription	transcriptional dual regulator	
YCCM	-NARL				
ECK1400-	ECK1215	0.5205	transcription	transcriptional dual regulator	
YDBD	-NARL				
ECK1495-	ECK1215	0.5087	transcription	transcriptional dual regulator	
YDEQ	-NARL				
ECK1748-	ECK1215	0.6084	transcription	transcriptional dual regulator	
YDJX	-NARL				

ECK2175- YEJG	ECK1215 -NARL	0.532	transcription	transcriptional dual regulator
ECK2614- YFJF	ECK1215 -NARL	0.5927	transcription	transcriptional dual regulator
ECK2615- YFJG	ECK1215 -NARL	0.5532	transcription	transcriptional dual regulator
ECK2868- YGEY	ECK1215 -NARL	0.5276	transcription	transcriptional dual regulator
ECK2892- YGFY	ECK1215 -NARL	0.5175	transcription	transcriptional dual regulator
ECK1031- YMDB	ECK1215 -NARL	0.5581	transcription	transcriptional dual regulator
ECK3033- YQIC	ECK1215 -NARL	0.55	transcription	transcriptional dual regulator
ECK1964- YEDV	ECK1913 -UVRV	0.5243	transcription	transcriptional dual regulator
ECK3874- YIHT	ECK0112 -PDHR	0.5394	transcription	transcriptional dual regulator
ECK1137- YMFO	ECK0112 -PDHR	0.5071	transcription	transcriptional dual regulator
ECK1136- YMFR	ECK0112 -PDHR	0.5723	transcription	transcriptional dual regulator
ECK1048- YCEB	ECK3822 -METR	0.6444	transcription	transcriptional dual regulator
ECK4394- YJJY	ECK4393 -ARCA	0.6396	transcription	transcriptional dual regulator
ECK1048- YCEB	ECK2837 -LYSR	0.6663	transcription	transcriptional dual regulator
ECK1901- YECJ	ECK2185 -NARP	0.5038	transcription	transcriptional dual regulator
ECK2601- YFIN	ECK2185 -NARP	0.512	transcription	transcriptional dual regulator
ECK0814- YBIY	ECK3248 -FIS	0.6071	transcription	transcriptional dual regulator
ECK0857- YBJQ	ECK3248 -FIS	0.7185	transcription	transcriptional dual regulator
ECK1169- YCGN	ECK3248 -FIS	0.5718	transcription	transcriptional dual regulator
ECK4004- YJAB	ECK3248 -FIS	0.5282	transcription	transcriptional dual regulator
ECK4179- YJFK	ECK3248 -FIS	0.5343	transcription	transcriptional dual regulator
ECK0275- YAGJ	ECK0986 -TORR	0.6118	transcription	transcriptional dual regulator
ECK1022- YCDZ	ECK0986 -TORR	0.516	transcription	transcriptional dual regulator
ECK1040- YCEA	ECK0986 -TORR	0.5738	transcription	transcriptional dual regulator
ECK1090- YCFL	ECK0986 -TORR	0.5992	transcription	transcriptional dual regulator
ECK1228-	ECK0986	0.6683	transcription	transcriptional

YCHJ	-TORR			dual regulator
ECK1267-	ECK0986	0.6339	transcription	transcriptional
YCIN	-TORR			dual regulator
ECK3562-	ECK0986	0.5474	transcription	transcriptional
YSAA	-TORR			dual regulator
ECK4243-	ECK0986	0.5167	transcription	transcriptional
YJGI	-TORR			dual regulator
ECK1158-	ECK0986	0.5087	transcription	transcriptional
YMGD	-TORR			dual regulator
ECK0824-	ECK0807	0.5071	transcription	transcriptional
YLIF	-MNTR			dual regulator
ECK2802-	ECK0144	0.5576	transcription	transcriptional
YGDD	-DKSA			dual regulator
ECK3757-	ECK0144	0.5513	transcription	transcriptional
YIFE	-DKSA			dual regulator
ECK1748-	ECK0380	0.5299	transcription	transcriptional
YDJX	-ADRA			dual regulator
ECK1031-	ECK0380	0.5234	transcription	transcriptional
YMDB	-ADRA			dual regulator
ECK1930-	ECK1335	0.5618	transcription	transcriptional
YEDK	-ABGR			dual regulator
ECK0775-	ECK1381	0.5275	transcription	transcriptional
YBHL	-FEAR			dual regulator
ECK1169-	ECK1381	0.5088	transcription	transcriptional
YCGN	-FEAR			dual regulator
ECK0679	ECK0682	0.5001	transcription	transcriptional
	-KDPE			dual regulator
ECK0795-	ECK3953	0.5148	transcription	transcriptional
MCBA	-OXYR			dual regulator
ECK2996-	ECK3496	0.5202	transcription	transcriptional
YQHA	-GADE			dual regulator
ECK1571-	ECK1004	0.5132	transcription	transcriptional
YDFE	-RUTR			dual regulator
ECK1842-	ECK1004	0.5189	transcription	transcriptional
YOBA	-RUTR			dual regulator
ECK0762-	ECK0830	0.5068	transcription	transcriptional
YBHB	-DEOR			repressor
ECK1228-	ECK0830	0.6604	transcription	transcriptional
YCHJ	-DEOR			repressor
ECK3424-	ECK0354	0.5544	transcription	transcriptional
YHHW	-FRMR			repressor
ECK3757-	ECK0354	0.5529	transcription	transcriptional
YIFE	-FRMR			repressor
ECK1774-	ECK4187	0.5389	transcription	transcriptional
YDJL	-ULAR			repressor
ECK1228-	ECK3594	0.5052	transcription	transcriptional
YCHJ	-LLDR			repressor
ECK3551-	ECK4082	0.5254	transcription	transcriptional
YIAA	-RPIR			repressor
ECK1773-	ECK3955	0.5745	transcription	transcriptional
YDJK	-FABR			repressor

ECK0630- YBEB	ECK1523 -MARR	0.5263	transcription	transcriptional repressor
ECK1160	ECK1501 -HIPB	0.5386	transcription	transcriptional repressor
ECK3092- YQJF	ECK2709 -ASCG	0.5318	transcription	transcriptional repressor
ECK3757- YIFE	ECK3756 -HDFR	0.5076	transcription	transcriptional repressor
ECK2102- YEHB	ECK1621 -CNU	0.5418	transcription	
ECK2634- YFJU	ECK0070 -SGRR	0.5728	transcription	
ECK3757- YIFE	ECK3491 -DCTR	0.5269	transcription	
ECK0322- YAHJ	ECK3232 -AAER	0.5722	transcription	
ECK2489- YFGO	ECK4289 -SGCR	0.5702	transcription	
ECK1048- YCEB	ECK2475 -GCVR	0.5212	transcription	
ECK0193- YAEF	ECK0161 -CDAR	0.5253	transcription	
ECK0209- YAFD	ECK0161 -CDAR	0.5393	transcription	
ECK2389- YFEA	ECK0161 -CDAR	0.5926	transcription	
ECK0536- YBCL	ECK3397 -FEOC	0.5295	transcription	
ECK1778- YEAD	ECK3397 -FEOC	0.5164	transcription	
ECK1448- YNCG	ECK3397 -FEOC	0.5323	transcription	
ECK3671- YIDL	ECK0718 -MNGR	0.5101	transcription	
ECK1036- YCEK	ECK1617 -MALY	0.5259	transcription + metabolism	amino acid transport & biosynthesis
ECK1664- YDHS	ECK1617 -MALY	0.5001	transcription + metabolism	amino acid transport & biosynthesis
ECK3367- YHFW	ECK4169 -HFLX	0.5188	translation	GTPase
ECK3092- YQJF	ECK3290 -RPSE	0.6456	translation	ribosome protein
ECK4053- YJCC	ECK1506 -LSRA	0.5137	transport	AI2
ECK3677- YIDE	ECK2671 -PROV	0.5987	transport	amino acid transport
ECK3551- YIAA	ECK3439 -LIVG	0.5121	transport	amino acid transport
ECK0691-	ECK1479	0.5025	transport	amino acid

YBFO	-DDPC			transport	
ECK0775-	ECK1479	0.536		transport	amino acid
YBHL	-DDPC				transport
ECK1228-	ECK1479	0.5025		transport	amino acid
YCHJ	-DDPC				transport
ECK0775-	ECK0798	0.5436		transport	amino acid
YBHL	-GLNQ				transport
ECK1265-	ECK0852	0.5244		transport	amino acid
YCIK	-ARTM				transport
ECK0775-	ECK1006	0.5536		transport	amino acid
YBHL	-PUTP				transport
ECK0785-	ECK1006	0.5058		transport	amino acid
YBIH	-PUTP				transport
ECK0282-	ECK3441	0.5019		transport	amino acid
YAGQ	-LIVH				transport
ECK1265-	ECK0853	0.5307		transport	amino acid
YCIK	-ARTQ				transport
ECK3551-	ECK1486	0.5288		transport	amino acid
YIAA	-GADC				transport
ECK3562-	ECK0645	0.5876		transport	amino acid
YSAA	-GLTL				transport
ECK1427-	ECK4132	0.5082		transport	C4-
YDCO	-DCUA				dicarboxylate
ECK0420-	ECK3813	0.5241		transport	drug
YAJQ	-RARD				
ECK3650-	ECK3813	0.5925		transport	drug
YICL	-RARD				
ECK1443-	ECK0372	0.5173		transport	drug transporter
YNCB	-SBMA				
ECK0493-	ECK2232	0.646		transport	G3P
YLBH	-GLPT				
ECK2609-	ECK2193	0.5454		transport	heme
YFJD	-CCMA				
ECK3671-	ECK3461	0.5247		transport	inorganic ion
YIDL	-NIKB				nickel
ECK0251-	ECK1463	0.5277		transport	transport
YKFF	-NARU				inorganic ion
ECK2376-	ECK1463	0.5307		transport	nitrite/nitra
YPDA	-NARU				te
ECK2546-	ECK1463	0.575		transport	transport
YPHG	-NARU				inorganic ion
ECK4246-	ECK3741	0.5077		transport	nitrite/nitra
YJGL	-KUP				te
ECK0252-	ECK0149	0.5076		transport	transport
YKFB	-FHUA				inorganic ion
ECK1040-	ECK1008	0.5706		transport	transport
YCEA	-EFEO				inorganic ion
ECK1228-	ECK1008	0.5583		transport	transport
YCHJ	-EFEO				inorganic ion
ECK0314-	ECK0150	0.5151		transport	transport
YAHB	-FHUC				inorganic ion
					transport

ECK2925-	ECK0150	0.5324	transport	inorganic ion
YGGD	-FHUC			transport
ECK2926-	ECK0150	0.5204	transport	inorganic ion
YGGF	-FHUC			transport
ECK3267-	ECK0048	0.5007	transport	inorganic ion
YRDB	-KEFC			transport
ECK3576-	ECK0152	0.5057	transport	inorganic ion
YIAW	-FHUB			transport
ECK2954-	ECK3323	0.5614	transport	inorganic ion
YGGL	-BFR			transport
ECK3677-	ECK3395	0.5084	transport	inorganic ion
YIDE	-FEOA			transport
ECK3367-	ECK3396	0.5108	transport	inorganic ion
YHFW	-FEOB			transport
ECK3757-	ECK3396	0.5071	transport	inorganic ion
YIFE	-FEOB			transport
ECK1400-	ECK0925	0.5037	transport	inorganic ion
YDBD	-SSUC			transport
ECK2614-	ECK0925	0.5113	transport	inorganic ion
YFJF	-SSUC			transport
ECK4186-	ECK0580	0.5588	transport	iron
YJFP	-FEPE			
ECK2105-	ECK2148	0.5365	transport	iron transport
YEHE	-CIRA			
ECK1419-	ECK3593	0.5308	transport	lactate
YDCH	-LLDP			
ECK3244-	ECK3352	0.5009	transport	MFS
YHDT	-TSGA			
ECK3865-	ECK3352	0.5893	transport	MFS
YIHL	-TSGA			
ECK4377-	ECK4328	0.5088	transport	MFS
YJJJ	-MDTM			
ECK4186-	ECK3649	0.5502	transport	MFS
YJFP	-SETC			transporter
ECK0951-	ECK0443	0.614	transport	multi-drug
YCCS	-MDLB			
ECK1036-	ECK0443	0.5314	transport	multi-drug
YCEK	-MDLB			
ECK1262-	ECK0443	0.5069	transport	multi-drug
YCIQ	-MDLB			
ECK1343-	ECK0443	0.5133	transport	multi-drug
YDAQ	-MDLB			
ECK2893-	ECK0443	0.5057	transport	multi-drug
YGFZ	-MDLB			
ECK3654-	ECK3664	0.5094	transport	multi-drug
YICN	-EMRD			
ECK4246-	ECK3664	0.531	transport	multi-drug
YJGL	-EMRD			
ECK4284-	ECK3664	0.5019	transport	multi-drug
YJHU	-EMRD			
ECK3876-	ECK4144	0.6271	transport	multi-drug

YIHV	-SUGE				
ECK1158-	ECK4144	0.5174	transport	multi-drug	
YMGD	-SUGE				
ECK3586-	ECK0333	0.5082	transport	nucleotide transporter (cytocine)	
YIBH	-CODB				
ECK0814-	ECK1239	0.5767	transport	oligopeptide	
YBIY	-OPPC				
ECK1169-	ECK1239	0.6143	transport	oligopeptide	
YCGN	-OPPC				
ECK1752-	ECK1239	0.5362	transport	oligopeptide	
YNJB	-OPPC				
ECK0705-	ECK4302	0.5264	transport	oligosaccharide	
YBGO	-NANC				
ECK4344-	ECK4302	0.5019	transport	oligosaccharide	
YJIY	-NANC				
ECK1090-	ECK1289	0.512	transport	peptide	
YCFL	-SAPA				
ECK3562-	ECK1289	0.527	transport	peptide	
YSAA	-SAPA				
ECK1040-	ECK0847	0.5583	transport	polyamine	
YCEA	-POTH				
ECK1228-	ECK0847	0.5055	transport	polyamine	
YCHJ	-POTH				
ECK1267-	ECK0847	0.5113	transport	polyamine	
YGIN	-POTH				
ECK0679	ECK1112	0.5406	transport	polyamine transport	
	-POTA				
ECK0758-	ECK0845	0.5665	transport	polyamine transport	
YBHH	-POTF				
ECK2513-	ECK0848	0.5054	transport	polyamine transport	
RLMN	-POTI				
ECK1040-	ECK1109	0.53	transport	polyamine transport	
YCEA	-POTD				
ECK0368-	ECK0846	0.545	transport	polyamine transport	
YAIT	-POTG				
ECK1494-	ECK0846	0.5134	transport	polyamine transport	
YDEP	-POTG				
ECK1496-	ECK0846	0.532	transport	polyamine transport	
YDER	-POTG				
ECK1965-	ECK0846	0.5142	transport	polyamine transport	
YEDW	-POTG				
ECK2064-	ECK0846	0.513	transport	polyamine transport	
YEGI	-POTG				
ECK2367-	ECK0846	0.5049	transport	polyamine transport	
YFDE	-POTG				
ECK2404-	ECK0846	0.5234	transport	polyamine transport	
YFEH	-POTG				
ECK2768-	ECK0846	0.521	transport	polyamine transport	
YGCW	-POTG				
ECK1449-	ECK0846	0.6088	transport	polyamine	

YNCH	-POTG			transport	secretion
ECK3173-	ECK3315	0.6514		transport	
YHBE	-GSPG				
ECK4186-	ECK3315	0.6796		transport	secretion
YJFP	-GSPG				
ECK2373-	ECK1594	0.6041		transport	spermidine
YFDY	-MDTI				
ECK1150-	ECK2171	0.5375		transport	
YCGF	-YEJA				
ECK2082-	ECK2171	0.512		transport	
YEGS	-YEJA				
ECK2101-	ECK0870	0.5693		transport	
YEHA	-MACB				
ECK2790-	ECK1507	0.5584		transport	
YGDH	-LSRC				
ECK1495-	ECK0445	0.5016		transport	
YDEQ	-AMTB				
ECK1933	ECK0445	0.5705		transport	
	-AMTB				
ECK4186-	ECK0445	0.5346		transport	
YJFP	-AMTB				
ECK0252-	ECK0126	0.5139		transport	
YKFB	-YADG				
ECK4243-	ECK3568	0.5057		transport	
YJGI	-YIAO				
ECK3179-	ECK1437	0.5336		transport	
YRBA	-YDCV				
ECK1090-	ECK1307	0.5404		transport	
YCFL	-YCJP				
ECK1267-	ECK1307	0.5136		transport	
YCIN	-YCJP				
ECK3562-	ECK1307	0.535		transport	
YSAA	-YCJP				
ECK3654-	ECK3567	0.5644		transport	
YICN	-YIAN				
ECK1245-	ECK2237	0.518	unknown function		pseudogene
YCII	-YPAA				
ECK1086-	ECK1252	0.5082	unknown function		
YCFH	-YCIF				
ECK1411-	ECK1731	0.5487	unknown function		
YDCA	-CHBG				
ECK0542-	ECK0539	0.673	unknown function		
YLCG	-NINE				
ECK1689-	ECK1153	0.5267	unknown function		
YDIN	-ARIR				
ECK0437-	ECK2296	0.5225	unknown function		
YBAW	-YFCG				
ECK2265-	ECK2296	0.6203	unknown function		
YFBL	-YFCG				
ECK2684-	ECK2296	0.5213	unknown function		
YQAA	-YFCG				

ECK0361-	ECK1367	0.5425	unknown function
YAIS	-STFR		
ECK0758-	ECK1367	0.5567	unknown function
YBHH	-STFR		
ECK0983-	ECK1367	0.5131	unknown function
YCCM	-STFR		
ECK1748-	ECK1367	0.5751	unknown function
YDJX	-STFR		
ECK2614-	ECK1367	0.5862	unknown function
YFJF	-STFR		
ECK0210-	ECK0356	0.543	unknown function
YAFE	-YAIX		
ECK3033-	ECK1260	0.5754	unknown function
YQIC	-YCIV		
ECK2268-	ECK1737	0.5756	unknown function
YFBO	-OSME		
ECK3757-	ECK3709	0.5388	unknown function
YIFE	-CBRB		
ECK3563-	ECK3335	0.5018	unknown function
YIAJ	-SLYX		
ECK3956-	ECK3335	0.5651	unknown function
YIJD	-SLYX		
ECK2093-	ECK3363	0.5092	unknown function
YEGV	-YHFS		
ECK3669-	ECK3710	0.5319	unknown function
YIDJ	-CBRC		
ECK4370-	ECK4367	0.5012	unknown function
YJJV	-OSMY		
ECK0931-	ECK1875	0.5084	unknown function
YCBS	-CUTC		
ECK1491-	ECK1875	0.6001	unknown function
YDEM	-CUTC		
ECK2316-	ECK1875	0.5173	unknown function
YFCJ	-CUTC		
ECK3677-	ECK0549	0.5311	unknown function
YIDE	-BORD		
ECK0630-	ECK1525	0.5097	unknown function
YBEB	-MARB		
ECK1545-	ECK1525	0.5044	unknown function
YNFN	-MARB		
ECK1148-	ECK0941	0.5007	unknown function
YCGX	-PQIA		
ECK1253-	ECK0941	0.5287	unknown function
YCIG	-PQIA		
ECK1245-	ECK0941	0.511	unknown function
YCII	-PQIA		
ECK1428-	ECK0941	0.5616	unknown function
YDCN	-PQIA		
ECK2850-	ECK0941	0.5356	unknown function
YGEH	-PQIA		
ECK1443-	ECK0941	0.643	unknown function

YNCB	-PQIA		
ECK1228-	ECK4392	0.5256	unknown function
YCHJ	-CRED		
ECK3109-	ECK1342	0.6331	unknown function
YHAB	-INTR		
ECK3109-	ECK1546	0.6224	unknown function
YHAB	-CSPI		
ECK1304-	ECK1146	0.5463	unknown function
YCJM	-ICDC		
ECK3935-	ECK4042	0.5102	unknown function
YIJE	-PSPG		
ECK0320	ECK4272	0.5114	unknown function
	-YJHE		
ECK1155-	ECK1361	0.5471	unknown function
YCGG	-RZOR		
ECK1036-	ECK1157	0.5347	unknown function
YCEK	-YCGH		
ECK1789-	ECK1157	0.5064	unknown function
YEAN	-YCGH		
ECK0369	ECK2653	0.5676	unknown function
	-CSID		
ECK0678-	ECK2856	0.5175	unknown function
YBFG			
ECK2402-	ECK3231	0.5932	unknown function
YFEN	-AAEX		
ECK1496-	ECK1592	0.5095	unknown function
YDER	-ASR		
ECK2081-	ECK1592	0.5175	unknown function
YEGR	-ASR		
ECK1449-	ECK2637	0.594	unknown function
YNCH			
ECK1674-	ECK2637	0.5413	unknown function
YNHG			
ECK1685-	ECK0231	0.6207	unknown function
YDIK	-LAFU		
ECK4300-	ECK0231	0.5347	unknown function
YJHS	-LAFU		
ECK0260-	ECK0231	0.5773	unknown function
YKFC	-LAFU		
ECK1096-	ECK1016	0.5202	unknown function
YCFJ	-YMDE		
ECK1419-	ECK0553	0.5081	unknown function
YDCH	-TFAD		
ECK1954-	ECK0553	0.5526	unknown function
YEDQ	-TFAD		
ECK0493-	ECK0553	0.5688	unknown function
YLBH	-TFAD		

**Table S7 - Drug-GO Interactions
at p<= 10⁻³
GO Biological Process**

	Drug	Drug-GO Score
ribosomal small subunit assembly	CISPLATIN	-4.2929
ribosomal small subunit assembly	SPECTINOMYCIN	-3.542
ribosomal small subunit assembly	THIOLACTOMYCIN	-4.5803
sulfate assimilation	AMOXICILLIN	-5.6992
sulfate assimilation	CISPLATIN	-23.5158
sulfate assimilation	GLUFOSFOMYCIN	-3.5682
sulfate assimilation	LEVOFLOXACIN	4.2058
sulfate assimilation	VERAPAMIL	-6.8348
rRNA modification	SPECTINOMYCIN	-4.2201
peptidoglycan metabolic process	CEFSULODIN6.0MECILLINAM	-3.2152
polysaccharide biosynthetic process	A22	-6.7125
polysaccharide biosynthetic process	CYCLOSERINED	-3.2906
polysaccharide biosynthetic process	STREPTOMYCIN	-6.6375
enzyme-directed rRNA pseudouridine synthesis	CHLORAMPHENICOL	-3.1568
conjugation	DOXYCYCLINE	-3.0959
cell morphogenesis	TRIMETHOPRIM0.1SULFAMETHIZOLE	-3.5328
barrier septum formation	GENTAMICIN	-4.0262
barrier septum formation	MECILLINAM	-5.7042
selection of site for barrier septum formation	FOSFOMYCIN	-4.0693
selenocysteine incorporation	SPECTINOMYCIN	5.8646
pseudouridine synthesis	CHLORAMPHENICOL	-3.4814
carbohydrate metabolic process	AZITHROMYCIN	-3.4065
carbohydrate metabolic process	BLEOMYCIN	-3.745
carbohydrate metabolic process	CHIR090	-4.2852
carbohydrate metabolic process	DIBUCAINE	-3.3945
carbohydrate metabolic process	MINOCYCLINE	-6.2018
carbohydrate metabolic process	SPECTINOMYCIN	-3.5648
carbohydrate metabolic process	SPIRAMYCIN	-3.9811
carbohydrate metabolic process	TRICLOSAN	-4.7106
fructose catabolic process	GENTAMICIN	-3.0469
glucose catabolic process	STREPTOMYCIN	-4.3828
D-ribose metabolic process	MINOCYCLINE	-3.1584
D-ribose metabolic process	MITOMYCINC	3.108
UDP-N-acetylglucosamine metabolic process	TRICLOSAN	-3.8943
UDP-N-acetylglucosamine biosynthetic process	AMPICILLIN	-3.2349
UDP-N-acetylglucosamine biosynthetic process	CEFACLOR	-3.0098
UDP-N-acetylglucosamine biosynthetic process	CHIR090	-3.1114
UDP-N-acetylglucosamine biosynthetic process	CLARYTHROMYCIN	-3.3866

UDP-N-acetylglucosamine biosynthetic process	ERYTHROMYCIN	-3.3649
UDP-N-acetylglucosamine biosynthetic process	MINOCYCLINE	-3.3457
UDP-N-acetylglucosamine biosynthetic process	NIGERICIN	3.1058
UDP-N-acetylglucosamine biosynthetic process	RIFAMPICIN	-4.568
UDP-N-acetylglucosamine biosynthetic process	SPECTINOMYCIN	-3.1516
UDP-N-acetylglucosamine biosynthetic process	SPIRAMYCIN	-3.7257
UDP-N-acetylglucosamine biosynthetic process	TETRACYCLINE	-3.1043
gluconeogenesis	CECROPINB	-3.718
gluconeogenesis	LEVOFLOXACIN	-3.8071
gluconeogenesis	PROCAINE	3.2121
glycolysis	THEOPHYLLINE	-3.2601
tricarboxylic acid cycle	TRICLOSAN	-7.4789
malate metabolic process	GLUFOSFOMYCIN	-5.6714
purine nucleotide biosynthetic process	A22	-12.3568
purine nucleotide biosynthetic process	ACTINOMYCIND	-20.5143
purine nucleotide biosynthetic process	AZIDOTHYMIDINE	-4.9978
purine nucleotide biosynthetic process	AZTREONAM	-4.6394
purine nucleotide biosynthetic process	CARBENICILLIN	-15.4486
purine nucleotide biosynthetic process	CEFOXITIN	-22.6225
purine nucleotide biosynthetic process	CEFSULODIN	-12.8416
purine nucleotide biosynthetic process	CHIR090	-4.8099
purine nucleotide biosynthetic process	CISPLATIN	-3.3132
purine nucleotide biosynthetic process	DIBUCAINE	-3.0769
purine nucleotide biosynthetic process	DOXYCYCLINE	-12.0291
purine nucleotide biosynthetic process	ERYTHROMYCIN	-23.792
purine nucleotide biosynthetic process	GLUFOSFOMYCIN	-4.8789
purine nucleotide biosynthetic process	INDOLICIDIN	5.4742
purine nucleotide biosynthetic process	MECILLINAM	-10.9059
purine nucleotide biosynthetic process	NITROFURANTOIN	-4.2747

process		
purine nucleotide biosynthetic process	NOVOBIOCIN	-17.6638
process		
purine nucleotide biosynthetic process	PHLEOMYCIN	-3.0347
process		
purine nucleotide biosynthetic process	PROCAINE	-18.5108
process		
purine nucleotide biosynthetic process	THIOLACTOMYCIN	3.4514
process		
purine nucleotide biosynthetic process	TOBRAMYCIN	-22.2549
process		
purine nucleotide biosynthetic process	VERAPAMIL	-30.1201
process		
GMP biosynthetic process	A22	-4.8333
GMP biosynthetic process	ACTINOMYCIND	-3.8028
GMP biosynthetic process	AZIDOTHYIMIDINE	-5.7691
GMP biosynthetic process	CARBENICILLIN	-3.7605
GMP biosynthetic process	CEFOXITIN	-4.9182
GMP biosynthetic process	CISPLATIN	-4.2929
GMP biosynthetic process	DOXYCYCLINE	-3.4944
GMP biosynthetic process	ERYTHROMYCIN	-8.9017
GMP biosynthetic process	GLUFOSFOMYCIN	-9.5513
GMP biosynthetic process	INDOLICIDIN	-4.3828
GMP biosynthetic process	MECILLINAM	-3.0762
GMP biosynthetic process	PHLEOMYCIN	-3.2721
GMP biosynthetic process	PROCAINE	-5.091
GMP biosynthetic process	SPECTINOMYCIN	-3.9759
GMP biosynthetic process	STREPTOMYCIN	-4.3828
GMP biosynthetic process	TOBRAMYCIN	-4.5119
GMP biosynthetic process	VERAPAMIL	-3.4769
IMP biosynthetic process	VERAPAMIL	-3.4769
de novo IMP biosynthetic process	A22	-4.1656
de novo IMP biosynthetic process	ACTINOMYCIND	-11.7317
de novo IMP biosynthetic process	CARBENICILLIN	-11.9935
de novo IMP biosynthetic process	CEFOXITIN	-12.4069
de novo IMP biosynthetic process	CEFSULODIN	-7.6039
de novo IMP biosynthetic process	CLARYTHROMYCIN	-3.9995
de novo IMP biosynthetic process	DOXYCYCLINE	-8.6991
de novo IMP biosynthetic process	ERYTHROMYCIN	-9.4043
de novo IMP biosynthetic process	INDOLICIDIN	3.3084
de novo IMP biosynthetic process	MECILLINAM	-9.4362
de novo IMP biosynthetic process	NITROFURANTOIN	-3.9991
de novo IMP biosynthetic process	NOVOBIOCIN	-13.3498
de novo IMP biosynthetic process	PROCAINE	-9.4654
de novo IMP biosynthetic process	TOBRAMYCIN	-9.976
de novo IMP biosynthetic process	VERAPAMIL	-14.1084
de novo pyrimidine base biosynthetic process	PROCAINE	-4.6341
de novo pyrimidine base biosynthetic process	VERAPAMIL	-3.9101

pyrimidine nucleoside metabolic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	4.2112
pyrimidine nucleotide biosynthetic process	A22	-11.5
pyrimidine nucleotide biosynthetic process	ACTINOMYCIND	-8.8097
pyrimidine nucleotide biosynthetic process	BACITRACIN	-3.8124
pyrimidine nucleotide biosynthetic process	CARBENICILLIN	-12.6193
pyrimidine nucleotide biosynthetic process	CEFOXITIN	-10.2715
pyrimidine nucleotide biosynthetic process	CEFSULODIN	-6.8542
pyrimidine nucleotide biosynthetic process	CLARYTHROMYCIN	-4.7679
pyrimidine nucleotide biosynthetic process	DOXYCYCLINE	-8.8838
pyrimidine nucleotide biosynthetic process	ERYTHROMYCIN	-8.6761
pyrimidine nucleotide biosynthetic process	MECILLINAM	-9.1308
pyrimidine nucleotide biosynthetic process	NOVOBIOCIN	-11.9065
pyrimidine nucleotide biosynthetic process	PROCAINE	-9.892
pyrimidine nucleotide biosynthetic process	SPECTINOMYCIN	-5.3937
pyrimidine nucleotide biosynthetic process	TOBRAMYCIN	-9.4265
pyrimidine nucleotide biosynthetic process	VERAPAMIL	-12.7205
UMP biosynthetic process	A22	-3.486
UMP biosynthetic process	ACTINOMYCIND	-3.8028
UMP biosynthetic process	CARBENICILLIN	-3.7605
UMP biosynthetic process	CEFOXITIN	-3.6488
UMP biosynthetic process	DOXYCYCLINE	-3.4944
UMP biosynthetic process	ERYTHROMYCIN	-3.7587
UMP biosynthetic process	MECILLINAM	-3.6706
UMP biosynthetic process	NOVOBIOCIN	-3.2935
UMP biosynthetic process	PROCAINE	-3.8001
UMP biosynthetic process	TOBRAMYCIN	-3.1314
UMP biosynthetic process	VERAPAMIL	-3.4769
dTMP biosynthetic process	AZIDOTHYIMIDINE	-4.1336
cellular DNA metabolic process	CIPROFLOXACIN	-3.0309
cellular DNA metabolic process	LEVOFLOXACIN	-3.704
cellular DNA metabolic process	MITOMYCINC	-5.2244
cellular DNA metabolic process	NITROFURANTOIN	-3.6714
cellular DNA replication	AZIDOTHYIMIDINE	-6.194
cellular DNA replication	CIPROFLOXACIN	-7.7875

cellular DNA replication	MITOMYCINC	-5.009
cellular DNA replication	NORFLOXACIN	-4.2989
cellular DNA replication	TRIMETHOPRIM	-8.7447
DNA-dependent DNA replication	TRIMETHOPRIM	-3.3991
DNA unwinding during replication	AZIDOTHYIMIDINE	-4.0226
DNA repair	AZIDOTHYIMIDINE	-11.6469
DNA repair	CIPROFLOXACIN	-14.2606
DNA repair	LEVOFLOXACIN	-11.2348
DNA repair	MITOMYCINC	-17.171
DNA repair	NITROFURANTOIN	-14.2334
DNA repair	NORFLOXACIN	-11.3863
DNA repair	STREPTONIGRIN	-10.4412
nucleotide-excision repair	MITOMYCINC	-11.1653
nucleotide-excision repair	NITROFURANTOIN	-8.5473
nucleotide-excision repair	STREPTONIGRIN	-3.3933
cellular DNA catabolic process	CIPROFLOXACIN	-5.6337
cellular DNA catabolic process	LEVOFLOXACIN	-5.2953
cellular DNA catabolic process	NALIDIXICACID	-5.621
cellular DNA catabolic process	NORFLOXACIN	-4.2952
DNA recombination	AZIDOTHYIMIDINE	-5.9495
DNA recombination	CIPROFLOXACIN	-4.4406
DNA recombination	FOSFOMYCIN	-3.6068
DNA recombination	LEVOFLOXACIN	-13.1591
DNA recombination	MITOMYCINC	-9.371
DNA recombination	NITROFURANTOIN	-8.1066
DNA recombination	NORFLOXACIN	-7.3911
DNA recombination	STREPTONIGRIN	-6.1956
transposition, DNA-mediated	LEVOFLOXACIN	-4.6067
transposition, DNA-mediated	MITOMYCINC	-3.108
cellular transcription	FOSFOMYCIN	-3.8885
cellular transcription	TRICLOSAN	-3.6818
cellular transcription, DNA-dependent	FOSFOMYCIN	-3.3931
transcription initiation	CARBENICILLIN	-3.091
transcription termination	SPECTINOMYCIN	-3.542
regulation of cellular transcription, DNA-dependent	DOXYCYCLINE	-3.4165
regulation of cellular transcription, DNA-dependent	FOSFOMYCIN	-3.54
regulation of cellular transcription, DNA-dependent	TRICLOSAN	-3.3355
mRNA processing	SPECTINOMYCIN	-3.1516
translation	SPECTINOMYCIN	-6.849
translational initiation	MITOMYCINC	-3.3843
regulation of translation	A22	5.7899
regulation of translation	CARBENICILLIN	-3.6772
regulation of translation	CEFTAZIDIME	-4.356
regulation of translation	SPECTINOMYCIN	-4.9839
phenylalanyl-tRNA aminoacylation	FUSIDICACID	-3.5429
phenylalanyl-tRNA aminoacylation	NIGERICIN	-3.374

regulation of translational termination	FOSFOMYCIN	-4.0693
protein folding	BACITRACIN	-3.06
protein folding	CECROPINB	-4.5956
protein folding	CHIR090	-3.8068
protein folding	DIBUCAINE	-6.4379
protein folding	TRICLOSAN	-4.524
protein folding	VANCOMYCIN	-3.5471
protein amino acid O-linked glycosylation	TRIMETHOPRIM0.1SULFAMETHIZOLE	-4.2112
proteolysis	TRICLOSAN	-3.7564
alanine metabolic process	CEFSULODIN6.0MECILLINAM	-4.1072
alanine metabolic process	CYCLOSERINED	-4.7665
alanine metabolic process	TRICLOSAN	-3.2215
aspartate metabolic process	METHOTREXATE	-3.1047
aspartate metabolic process	MITOMYCINC	-4.4601
aspartate metabolic process	SPECTINOMYCIN	4.2201
aspartate metabolic process	SULFAMETHIZOLE	5.8062
aspartate biosynthetic process	STREPTOMYCIN	-3.9834
cysteine biosynthetic process from serine	AMOXICILLIN	-3.702
cysteine biosynthetic process from serine	CARBENICILLIN	-3.7714
cysteine biosynthetic process from serine	CEFSULODIN6.0MECILLINAM	-3.7098
cysteine biosynthetic process from serine	CISPLATIN	-4.1459
cysteine biosynthetic process from serine	VANCOMYCIN	-3.5333
cysteine biosynthetic process from serine	VERAPAMIL	-3.0876
glutamine metabolic process	ACTINOMYCIND	-3.9463
glutamine metabolic process	CARBENICILLIN	-3.5836
glutamine metabolic process	CEFOXITIN	-5.9265
glutamine metabolic process	ERYTHROMYCIN	-10.2546
glutamine metabolic process	INDOLICIDIN	5.7392
glutamine metabolic process	NOVOBIOCIN	-3.4171
glutamine metabolic process	PUROMYCIN	4.8543
glutamine metabolic process	RIFAMPICIN	-3.0636
glutamine metabolic process	TOBRAMYCIN	-3.1338
glutamine metabolic process	VERAPAMIL	-6.6892
glycine metabolic process	SULFAMETHIZOLE	-5.1158
glycine metabolic process	SULFAMONOMETHOXINE	-4.2408
glycine metabolic process	TRIMETHOPRIM	-3.6513
glycine catabolic process	CECROPINB	-4.8015
glycine catabolic process	MITOMYCINC	-3.3843
glycine catabolic process	SULFAMETHIZOLE	-10.6035
glycine catabolic process	SULFAMONOMETHOXINE	-8.7972
L-serine metabolic process	TRIMETHOPRIM	-4.3306
L-serine biosynthetic process	BACITRACIN	-3.6335

L-serine biosynthetic process	CARBENICILLIN	-3.3669
L-serine biosynthetic process	NOVOBIOCIN	-3.5315
tyrosine biosynthetic process	STREPTOMYCIN	-4.3828
valine metabolic process	CEFSULODIN6.0MECILLINAM	-3.4302
cellular biogenic amine metabolic process	MITOMYCINC	4.4601
cellular biogenic amine metabolic process	STREPTOMYCIN	-5.0699
polyamine biosynthetic process	FOSFOMYCIN	-3.0046
fatty acid biosynthetic process	SPECTINOMYCIN	3.8657
fatty acid biosynthetic process	TRICLOSAN	-4.5891
phospholipid metabolic process	ERYTHROMYCIN	3.2927
phospholipid metabolic process	SPECTINOMYCIN	-4.2201
ubiquinone biosynthetic process	AZTREONAM	5.0691
ubiquinone biosynthetic process	FOSFOMYCIN	-4.131
ubiquinone biosynthetic process	GLUFOSFOMYCIN	-3.8229
ubiquinone biosynthetic process	MITOMYCINC	-4.8637
glutathione metabolic process	THEOPHYLLINE	-5.2931
ATP biosynthetic process	A22	-4.1546
ATP biosynthetic process	ACTINOMYCIND	-3.658
ATP biosynthetic process	AMOXICILLIN	-4.8452
ATP biosynthetic process	AMPICILLIN	-3.4268
ATP biosynthetic process	AZIDOTHYIMIDINE	-4.0332
ATP biosynthetic process	AZTREONAM	7.406
ATP biosynthetic process	BACITRACIN	-6.808
ATP biosynthetic process	CEFACTOR	-3.6185
ATP biosynthetic process	CEFSULODIN	-4.0158
ATP biosynthetic process	CISPLATIN	-3.8099
ATP biosynthetic process	GLUFOSFOMYCIN	-5.7194
ATP biosynthetic process	NORFLOXACIN	-5.0257
ATP biosynthetic process	THIOLACTOMYCIN	-4.7851
ATP biosynthetic process	TUNICAMYCIN	-3.5482
folic acid and derivative metabolic process	CEFACTOR	-3.2765
Mo-molybdopterin cofactor biosynthetic process	SPECTINOMYCIN	3.6848
nitrogen compound metabolic process	TOBRAMYCIN	-3.6847
transport	AMOXICILLIN	-3.0316
transport	AZIDOTHYIMIDINE	-4.0647
transport	BACITRACIN	-3.0055
transport	CEFSULODIN6.0MECILLINAM	-3.3028
transport	ERYTHROMYCIN	-3.4257
ion transport	ACTINOMYCIND	-3.127
ion transport	AZTREONAM	3.2663
ion transport	BACITRACIN	-6.5179
ion transport	GLUFOSFOMYCIN	-5.9937
ion transport	NORFLOXACIN	-3.0114
ion transport	STREPTOMYCIN	-4.0648
ion transport	THIOLACTOMYCIN	-3.1419

iron ion transport	BACITRACIN	-3.2826
iron ion transport	FOSFOMYCIN	-4.1436
iron ion transport	NOVOBIOCIN	-3.3451
iron ion transport	VERAPAMIL	-3.1939
zinc ion transport	GLUFOSFOMYCIN	-3.0112
extracellular transport	TRIMETHOPRIM0.1SULFAMETHIZOLE	-3.5328
response to stress	BLEOMYCIN	-3.0396
response to stress	TRICLOSAN	-3.7698
response to stress	TUNICAMYCIN	-3.213
response to osmotic stress	STREPTOMYCIN	-3.9834
response to DNA damage stimulus	AZIDOTHYIMIDINE	-9.7688
response to DNA damage stimulus	CIPROFLOXACIN	-15.4632
response to DNA damage stimulus	LEVOFLOXACIN	-11.8224
response to DNA damage stimulus	MITOMYCINC	-18.5001
response to DNA damage stimulus	NITROFURANTOIN	-15.1025
response to DNA damage stimulus	NORFLOXACIN	-11.7494
response to DNA damage stimulus	STREPTONIGRIN	-9.424
cellular cell wall organization	AMPICILLIN	-3.3323
cellular cell wall organization	CEFACTOR	-3.5234
cellular cell wall organization	CEFOXITIN	-3.5411
cellular cell wall organization	CEFTAZIDIME	-3.1449
cellular cell wall organization	CYCLOSERINED	-3.9032
cellular cell wall organization	MECILLINAM	-5.8177
cell cycle	GENTAMICIN	-4.2469
cell cycle	MECILLINAM	-4.654
chromosome segregation	LEVOFLOXACIN	-4.2058
mitotic chromosome condensation	CIPROFLOXACIN	-3.0603
sensory perception	SULFAMONOMETHOXINE	-3.3457
tRNA processing	ACTINOMYCIND	3.5159
tRNA processing	CEFOXITIN	3.7451
tRNA processing	ERYTHROMYCIN	3.6152
tRNA processing	TOBRAMYCIN	3.8929
metabolic process	CEFOXITIN	-3.1363
metabolic process	CISPLATIN	-3.2052
regulation of cell shape	CARBENICILLIN	-3.3063
regulation of cell shape	CEFOXITIN	-4.147
regulation of cell shape	CYCLOSERINED	-5.1352
regulation of cell shape	MECILLINAM	-3.6198
lipid biosynthetic process	ERYTHROMYCIN	3.0899
lipid biosynthetic process	FUSIDICACID	-3.7896
lipid biosynthetic process	NIGERICIN	-3.1421
lipid biosynthetic process	TRICLOSAN	-5.4754
lipid biosynthetic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	3.7436
lipid biosynthetic process	VANCOMYCIN	-3.3601
pyridoxine biosynthetic process	A22	-4.4028
pyridoxine biosynthetic process	INDOLICIDIN	3.0329
pyridoxine biosynthetic process	MECILLINAM	-3.5507
pyridoxine biosynthetic process	NOVOBIOCIN	-5.1682
pyridoxine biosynthetic process	OXACILLIN	3.25

cellular amino acid biosynthetic process	CARBENICILLIN	-3.7263
cellular amino acid biosynthetic process	CISPLATIN	-4.5203
cellular amino acid biosynthetic process	NOVOBIOCIN	-3.3613
cellular amino acid biosynthetic process	STREPTOMYCIN	-4.204
cellular amino acid biosynthetic process	VERAPAMIL	-3.2784
biosynthetic process	DIBUCAINE	-3.7296
biosynthetic process	MINOCYCLINE	-3.0862
biosynthetic process	STREPTOMYCIN	-3.1562
aerobic respiration	FOSFOMYCIN	-5.4937
anaerobic respiration	AMOXICILLIN	-3.548
anaerobic respiration	CCCP	-5.4282
anaerobic respiration	FOSFOMYCIN	-6.1096
anaerobic respiration	NITROFURANTOIN	-3.0675
anaerobic respiration	SPECTINOMYCIN	4.145
anaerobic respiration	THEOPHYLLINE	-5.4461
anaerobic respiration	TUNICAMYCIN	-3.4464
serine family amino acid metabolic process	TRIMETHOPRIM	-4.3306
aromatic amino acid family biosynthetic process	AMPICILLIN	-4.5241
aromatic amino acid family biosynthetic process	FOSFOMYCIN	-5.6397
aromatic amino acid family biosynthetic process	STREPTOMYCIN	-5.1375
methionine biosynthetic process	SULFAMONOMETHOXINE	-4.5155
L-phenylalanine biosynthetic process	STREPTOMYCIN	-3.7018
leucine biosynthetic process	STREPTOMYCIN	-3.4847
lipopolysaccharide biosynthetic process	ACTINOMYCIND	-3.3389
lipopolysaccharide biosynthetic process	AZIDOTHYIMIDINE	-3.9938
lipopolysaccharide biosynthetic process	AZITHROMYCIN	-5.6612
lipopolysaccharide biosynthetic process	BACITRACIN	-7.9785
lipopolysaccharide biosynthetic process	BLEOMYCIN	-5.7182
lipopolysaccharide biosynthetic process	CEFSULODIN	-3.5536
lipopolysaccharide biosynthetic process	CHIR090	-8.4132
lipopolysaccharide biosynthetic process	CLARYTHROMYCIN	-5.5868
lipopolysaccharide biosynthetic process	DIBUCAINE	-9.1555

process		
lipopolysaccharide biosynthetic process	DOXYCYCLINE	-6.3373
process		
lipopolysaccharide biosynthetic process	ERYTHROMYCIN	-3.2948
process		
lipopolysaccharide biosynthetic process	FUSIDICACID	-4.3977
process		
lipopolysaccharide biosynthetic process	MECILLINAM	4.692
process		
lipopolysaccharide biosynthetic process	MINOCYCLINE	-21.3471
process		
lipopolysaccharide biosynthetic process	NALIDIXICACID	-9.3207
process		
lipopolysaccharide biosynthetic process	NOVOBIOCIN	-10.0812
process		
lipopolysaccharide biosynthetic process	SPECTINOMYCIN	-4.2406
process		
lipopolysaccharide biosynthetic process	TRICLOSAN	-16.8175
process		
lipopolysaccharide biosynthetic process	TRIMETHOPRIM	-6.5428
process		
lipopolysaccharide biosynthetic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	-8.0133
process		
lipoate biosynthetic process	GLUFOSFOMYCIN	4.076
lipoate biosynthetic process	ISONIAZID	-3.0475
lipoate biosynthetic process	LEVOFLOXACIN	-5.2953
lipoate biosynthetic process	MITOMYCINC	-8.942
lipoate biosynthetic process	THIOLACTOMYCIN	-4.5803
lipoate biosynthetic process	TRICLOSAN	-3.2215
purine base biosynthetic process	A22	-3.0967
purine base biosynthetic process	ACTINOMYCIND	-6.7694
purine base biosynthetic process	AZTREONAM	-4.5528
purine base biosynthetic process	CARBENICILLIN	-3.3669
purine base biosynthetic process	CEFOXITIN	-6.5001
purine base biosynthetic process	CEFSULODIN	-3.0923
purine base biosynthetic process	DOXYCYCLINE	-3.1044
purine base biosynthetic process	ERYTHROMYCIN	-6.7336
purine base biosynthetic process	PROCAINE	-5.7832
purine base biosynthetic process	TOBRAMYCIN	-5.1541
purine base biosynthetic process	VERAPAMIL	-7.2416
nucleoside metabolic process	ACTINOMYCIND	-4.2037
nucleoside metabolic process	CEFOXITIN	-4.7208
nucleoside metabolic process	ERYTHROMYCIN	-4.1724
nucleoside metabolic process	PROCAINE	-4.1626
nucleoside metabolic process	TOBRAMYCIN	-5.1599
nucleoside metabolic process	VERAPAMIL	-4.2937
nucleotide biosynthetic process	AZIDOTHYMIDINE	-3.0683
lipopolysaccharide core region biosynthetic process	CLARYTHROMYCIN	-3.2173
lipopolysaccharide core region	DIBUCAINE	-3.5143

lipopolysaccharide core region biosynthetic process	INDOLICIDIN	-3.0329
lipopolysaccharide core region biosynthetic process	MINOCYCLINE	-7.9747
lipopolysaccharide core region biosynthetic process	MITOMYCINC	-5.6766
lipopolysaccharide core region biosynthetic process	NALIDIXICACID	-4.592
lipopolysaccharide core region biosynthetic process	TRICLOSAN	-4.5891
lipopolysaccharide core region biosynthetic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	-5.1938
lipid A biosynthetic process	FUSIDICACID	-3.685
lipid A biosynthetic process	MINOCYCLINE	-4.4148
enterobacterial common antigen biosynthetic process	CYCLOSERINED	-4.7665
glucan biosynthetic process	DOXYCYCLINE	3.4784
peptidoglycan biosynthetic process	AMPICILLIN	-4.1112
peptidoglycan biosynthetic process	CARBENICILLIN	-3.558
peptidoglycan biosynthetic process	CEFACLOR	-4.1885
peptidoglycan biosynthetic process	CEFOXITIN	-4.4261
peptidoglycan biosynthetic process	CEFSULODIN	-3.4673
peptidoglycan biosynthetic process	CEFSULODIN6.0MECILLINAM	-5.9262
peptidoglycan biosynthetic process	CEFTAZIDIME	-4.0182
peptidoglycan biosynthetic process	CYCLOSERINED	-7.8786
peptidoglycan biosynthetic process	MECILLINAM	-4.9988
10-formyltetrahydrofolate biosynthetic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	3.142
deoxyribonucleotide catabolic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	7.3552
peptidoglycan-based cell wall biogenesis	CARBENICILLIN	-3.3669
peptidoglycan-based cell wall biogenesis	CEFACLOR	-3.0098
peptidoglycan-based cell wall biogenesis	CEFOXITIN	-3.2564
peptidoglycan-based cell wall biogenesis	CEFSULODIN	-3.0923
peptidoglycan-based cell wall biogenesis	CEFSULODIN6.0MECILLINAM	-3.7098
flagellum assembly	FOSFOMYCIN	-4.4805
protein secretion	PROCAINE	3.8624
protein secretion	STREPTOMYCIN	-4.3312
protein secretion	TRICLOSAN	-6.0661
phospholipid catabolic process	TRICLOSAN	-3.2215
folic acid and derivative biosynthetic process	SULFAMETHIZOLE	-4.4282
folic acid and derivative biosynthetic process	SULFAMONOMETHOXINE	-8.1108

response to heat	AZIDOTHYMIDINE	-3.0683
response to heat	CECROPINB	-4.8015
response to heat	SPECTINOMYCIN	-3.1516
response to cold	INDOLICIDIN	-4.3828
response to cold	MITOMYCINC	-3.7784
response to cold	TRICLOSAN	3.2215
SOS response	AZIDOTHYMIDINE	-3.2487
SOS response	CIPROFLOXACIN	-7.9925
SOS response	LEVOFLOXACIN	-6.7687
SOS response	MITOMYCINC	-12.7685
SOS response	NITROFURANTOIN	-13.2018
SOS response	NORFLOXACIN	-4.3039
SOS response	STREPTONIGRIN	-4.3184
glyoxylate catabolic process	CEFSULODIN6.0MECILLINAM	3.4302
allantoin assimilation pathway	CEFSULODIN6.0MECILLINAM	-3.4302
putrescine biosynthetic process	FOSFOMYCIN	-3.3931
putrescine biosynthetic process	PROCAINE	-3.0179
RNA modification	CECROPINB	-5.2044
RNA modification	MITOMYCINC	-3.7784
detection of virus	ACTINOMYCIND	-3.658
detection of virus	CARBENICILLIN	-3.0473
detection of virus	CIPROFLOXACIN	-4.7631
cellular response to iron ion	FOSFOMYCIN	-4.0693
starvation		
response to organic cyclic	BLEOMYCIN	-4.7733
substance		
response to organic cyclic	CCCP	-3.44
substance		
response to organic cyclic	CHIR090	-4.1759
substance		
response to organic cyclic	CHLORAMPHENICOL	-5.0641
substance		
response to organic cyclic	CLARYTHROMYCIN	-3.2207
substance		
response to organic cyclic	DIBUCAINE	-3.871
substance		
response to organic cyclic	DOXYCYCLINE	-4.1721
substance		
response to organic cyclic	ERYTHROMYCIN	-3.2876
substance		
response to organic cyclic	FUSIDICACID	-4.1309
substance		
response to organic cyclic	MINOCYCLINE	-4.4195
substance		
response to organic cyclic	MITOMYCINC	-4.4601
substance		
response to organic cyclic	NOVOBIOCIN	-4.6103
substance		
response to organic cyclic	OXACILLIN	-5.496
substance		

response to organic cyclic substance	PUROMYCIN	-3.3259
response to organic cyclic substance	SPECTINOMYCIN	-4.2201
response to organic cyclic substance	TRICLOSAN	-3.8943
response to organic cyclic substance	TRIMETHOPRIM	-4.3306
protein transport	ACTINOMYCIND	-3.1177
protein transport	AMOXICILLIN	-3.3099
protein transport	AMPICILLIN	-7.477
protein transport	BACITRACIN	-5.2845
protein transport	BLEOMYCIN	-3.4307
protein transport	CARBENICILLIN	-4.8294
protein transport	CEFSULODIN	-5.2085
protein transport	CEFSULODIN6.0MECILLINAM	-10.3319
protein transport	CEFTAZIDIME	-5.5461
protein transport	DIBUCAINE	-6.5644
protein transport	MECILLINAM	-4.8381
protein transport	STREPTOMYCIN	-5.885
protein transport	TRICLOSAN	-12.6653
protein transport	TRIMETHOPRIM	-3.2069
protein transport	VANCOMYCIN	-8.6886
DNA integration	LEVOFLOXACIN	-3.4485
lactate transport	FOSFOMYCIN	-3.3931
fumarate transport	FOSFOMYCIN	-4.0693
fumarate transport	SPECTINOMYCIN	4.2201
proline transport	STREPTOMYCIN	-4.3828
peptidoglycan transport	CEFSULODIN6.0MECILLINAM	-4.7923
cobalamin transport	SULFAMONOMETHOXINE	-3.8424
lipopolysaccharide export	TRICLOSAN	-3.8943
carbon utilization by utilization of organic compounds	FOSFOMYCIN	-3.0046
ATP synthesis coupled proton transport	A22	-6.647
ATP synthesis coupled proton transport	ACTINOMYCIND	-4.1215
ATP synthesis coupled proton transport	AMOXICILLIN	-6.5602
ATP synthesis coupled proton transport	AMPICILLIN	-3.0905
ATP synthesis coupled proton transport	AZIDOTHYIMIDINE	-3.3325
ATP synthesis coupled proton transport	AZTREONAM	8.096
ATP synthesis coupled proton transport	BACITRACIN	-9.4535
ATP synthesis coupled proton transport	CEFACLOR	-5.8133
ATP synthesis coupled proton	CEFSULODIN	-5.6048

transport		
ATP synthesis coupled proton transport	CEFTAZIDIME	-3.2719
ATP synthesis coupled proton transport	CISPLATIN	-7.011
ATP synthesis coupled proton transport	GLUFOSFOMYCIN	-6.3943
ATP synthesis coupled proton transport	NORFLOXACIN	-6.4016
ATP synthesis coupled proton transport	STREPTOMYCIN	-6.883
ATP synthesis coupled proton transport	THIOLACTOMYCIN	-7.0939
ATP synthesis coupled proton transport	TUNICAMYCIN	-5.6105
electron transport coupled proton transport	AMOXICILLIN	-3.9872
electron transport coupled proton transport	FOSFOMYCIN	-3.3931
proton transport	A22	-5.3142
proton transport	ACTINOMYCIND	-3.3189
proton transport	AMOXICILLIN	-5.2996
proton transport	AZTREONAM	6.8993
proton transport	BACITRACIN	-7.9262
proton transport	CEFACTOR	-4.8183
proton transport	CEFSULODIN	-4.4146
proton transport	CISPLATIN	-5.8036
proton transport	GLUFOSFOMYCIN	-5.2279
proton transport	NORFLOXACIN	-5.4037
proton transport	STREPTOMYCIN	-5.7392
proton transport	THIOLACTOMYCIN	-5.8966
proton transport	TUNICAMYCIN	-4.7288
carbohydrate biosynthetic process	AMPICILLIN	-3.2349
carbohydrate biosynthetic process	CEFACTOR	-3.0098
carbohydrate biosynthetic process	CHIR090	-3.1114
carbohydrate biosynthetic process	CLARYTHROMYCIN	-3.3866
carbohydrate biosynthetic process	ERYTHROMYCIN	-3.3649
carbohydrate biosynthetic process	MINOCYCLINE	-4.7012
carbohydrate biosynthetic process	NIGERICIN	3.1058
carbohydrate biosynthetic process	RIFAMPICIN	-4.568
carbohydrate biosynthetic process	SPECTINOMYCIN	-3.1516
carbohydrate biosynthetic process	SPIRAMYCIN	-3.7257
carbohydrate biosynthetic process	TETRACYCLINE	-3.1043
carbohydrate catabolic process	INDOLICIDIN	3.7018
antibiotic catabolic process	CEFACTOR	-3.7333
cytochrome complex assembly	CEFSULODIN6.0MECILLINAM	-4.82
cytochrome complex assembly	ISONIAZID	-4.7555
protein import	AMPICILLIN	-5.2431
protein import	CARBENICILLIN	-4.4419
protein import	CEFSULODIN6.0MECILLINAM	-4.7923

protein import	MECILLINAM	-3.743
protein import	TRICLOSAN	-3.8943
regulation of fatty acid metabolic process	TRICLOSAN	-3.8943
glycine biosynthetic process from serine	TRIMETHOPRIM	-4.3306
siderophore biosynthetic process	FOSFOMYCIN	-4.0693
tyrosine biosynthetic process from chorismate via 4-hydroxyphenylpyruvate	STREPTOMYCIN	-5.0699
D-ribose catabolic process	MITOMYCINC	-3.7784
cysteine biosynthetic process	AMOXICILLIN	-6.9294
cysteine biosynthetic process	BICYCLOMYCIN	-3.4415
cysteine biosynthetic process	CISPLATIN	-15.9598
cysteine biosynthetic process	ISONIAZID	-3.2392
cysteine biosynthetic process	NALIDIXICACID	-3.6513
cysteine biosynthetic process	NORFLOXACIN	-3.4145
cysteine biosynthetic process	STREPTOMYCIN	-3.0329
cysteine biosynthetic process	THEOPHYLLINE	-3.2561
cysteine biosynthetic process	VANCOMYCIN	-7.5494
cysteine biosynthetic process	VERAPAMIL	-7.4248
siroheme biosynthetic process	CISPLATIN	-4.9791
siroheme biosynthetic process	VERAPAMIL	-4.1539
nicotinate nucleotide biosynthetic process	FOSFOMYCIN	-4.0693
sulfate assimilation, phosphoadenylyl sulfate reduction by phosphoadenylyl-sulfate reductase (thioredoxin)	CISPLATIN	-4.9791
sulfate reduction	CISPLATIN	-4.9791
glycine decarboxylation via glycine cleavage system	SULFAMETHIZOLE	-17.6887
glycine decarboxylation via glycine cleavage system	SULFAMONOMETHOXINE	-14.8876
D-gluconate metabolic process	AMOXICILLIN	-3.0803
propionate catabolic process	INDOLICIDIN	5.0699
propionate catabolic process, 2-methylcitrate cycle	INDOLICIDIN	3.7018
propionate metabolic process, methylcitrate cycle	INDOLICIDIN	4.3828
photosynthesis, light reaction	AMOXICILLIN	-3.2774
L-ascorbic acid metabolic process	DIBUCAINE	3.423
L-ascorbic acid catabolic process	FOSFOMYCIN	-3.3931
pyrimidine base biosynthetic process	A22	-4.1631
pyrimidine base biosynthetic process	CARBENICILLIN	-4.4419
pyrimidine base biosynthetic process	CEFOXITIN	-4.3289
pyrimidine base biosynthetic process	CEFSULODIN	-4.1556

process		
pyrimidine base biosynthetic	DOXYCYCLINE	-3.1357
process		
pyrimidine base biosynthetic	ERYTHROMYCIN	-4.4401
process		
pyrimidine base biosynthetic	MECILLINAM	-3.743
process		
pyrimidine base biosynthetic	NOVOBIOCIN	-3.8651
process		
pyrimidine base biosynthetic	VERAPAMIL	-4.1539
process		
dormancy process	AZIDOTHYMIDINE	-3.1881
dormancy process	SULFAMETHIZOLE	-4.7131
dormancy process	SULFAMONOMETHOXINE	-3.8424
lipid modification	CEFSULODIN6.0MECILLINAM	-4.7923
lipid modification	DIBUCAINE	-3.871
establishment of competence for	AZIDOTHYMIDINE	-4.1336
transformation		
maturation of SSU-rRNA	CEFSULODIN	-4.1556
maturation of SSU-rRNA	ISONIAZID	-3.5046
maturation of SSU-rRNA	MITOMYCINC	-4.4601
maturation of SSU-rRNA	STREPTOMYCIN	-5.0699
organic acid phosphorylation	FOSFOMYCIN	-3.3931
transcription antitermination	FOSFOMYCIN	-3.0046
transcription antitermination	PHLEOMYCIN	-3.0068
negative regulation of nuclease	FOSFOMYCIN	-4.0693
activity		
transposition	LEVOFLOXACIN	-3.2476
GMP salvage	SULFAMONOMETHOXINE	-4.2408
IMP salvage	SULFAMONOMETHOXINE	-4.2408
ribosome disassembly	CEFTAZIDIME	-3.3857
ribosome disassembly	MITOMYCINC	-4.4601
L-phenylalanine biosynthetic	STREPTOMYCIN	-5.0699
process from chorismate via		
phenylpyruvate		
negative regulation of translation,	CHLORAMPHENICOL	-3.241
ncRNA-mediated		
negative regulation of translation,	CYCLOSERINED	4.7665
ncRNA-mediated		
negative regulation of translation,	TRICLOSAN	3.2215
ncRNA-mediated		
regulation of cell proliferation	FOSFOMYCIN	-4.0693
plasmid recombination	ISONIAZID	-3.5015
plasmid recombination	LEVOFLOXACIN	-5.9862
plasmid recombination	MITOMYCINC	-4.4601
response to drug	AMPICILLIN	-3.426
response to drug	CARBENICILLIN	-3.0349
response to drug	PROCAINE	3.3561
response to drug	TRICLOSAN	-3.5311
hyperosmotic salinity response	STREPTOMYCIN	-5.0699

biofilm formation	CECROPINB	-3.6233
plasma membrane ATP synthesis coupled proton transport	AMOXICILLIN	-5.6222
plasma membrane ATP synthesis coupled proton transport	BACITRACIN	-3.9245
plasma membrane ATP synthesis coupled proton transport	GLUFOSFOMYCIN	-3.7408
vitamin B6 metabolic process	INDOLICIDIN	5.0699
vitamin B6 metabolic process	OXACILLIN	3.4269
pyridoxal phosphate biosynthetic process	BACITRACIN	-3.3558
pyridoxal phosphate biosynthetic process	CARBENICILLIN	-3.091
pyridoxal phosphate biosynthetic process	ERYTHROMYCIN	-3.0889
pyridoxal phosphate biosynthetic process	MECILLINAM	-3.9121
pyridoxal phosphate biosynthetic process	NOVOBIOCIN	-5.6051
pyruvate biosynthetic process	FOSFOMYCIN	-4.0693
pyruvate biosynthetic process	GLUFOSFOMYCIN	-4.7607
pyruvate biosynthetic process	INDOLICIDIN	-5.0699
pyruvate biosynthetic process	THIOLACTOMYCIN	-3.1983
antibiotic transport	TRIMETHOPRIM0.1SULFAMETHIZOLE	4.2112
enterobactin transport	BLEOMYCIN	-4.0884
enterobactin transport	CHIR090	-3.4998
enterobactin transport	CHLORAMPHENICOL	-4.3774
enterobactin transport	DIBUCAINE	-3.2004
enterobactin transport	DOXYCYCLINE	-3.4944
enterobactin transport	FUSIDICACID	-3.615
enterobactin transport	MINOCYCLINE	-3.7388
enterobactin transport	MITOMYCINC	-3.7784
enterobactin transport	NOVOBIOCIN	-3.9271
enterobactin transport	OXACILLIN	-4.8075
enterobactin transport	SPECTINOMYCIN	-3.542
enterobactin transport	TRICLOSAN	-3.2215
enterobactin transport dipeptide transport	TRIMETHOPRIM	-3.6513
cellular metabolic compound salvage	INDOLICIDIN	3.4847
Gram-negative-bacterium-type cell outer membrane assembly	TRIMETHOPRIM0.1SULFAMETHIZOLE	3.5328
Gram-negative-bacterium-type cell outer membrane assembly	BLEOMYCIN	-3.0153
Gram-negative-bacterium-type cell outer membrane assembly	DIBUCAINE	-3.2004
Gram-negative-bacterium-type cell outer membrane assembly	MITOMYCINC	-3.7784
Gram-negative-bacterium-type cell outer membrane assembly	NALIDIXICACID	-3.0464
Gram-negative-bacterium-type cell outer membrane assembly	TRICLOSAN	-3.2215

Gram-negative-bacterium-type cell	TRIMETHOPRIM0.1SULFAMETHIZOLE	-3.5328
outer membrane assembly		
bacteriocin transport	ACTINOMYCIND	-8.527
bacteriocin transport	AMOXICILLIN	-6.301
bacteriocin transport	AMPICILLIN	-9.6992
bacteriocin transport	BACITRACIN	-7.6391
bacteriocin transport	CARBENICILLIN	-11.3158
bacteriocin transport	CEFSULODIN	-4.496
bacteriocin transport	CEFSULODIN6.0MECILLINAM	-10.9443
bacteriocin transport	CEFTAZIDIME	-7.6884
bacteriocin transport	DIBUCAINE	-5.8324
bacteriocin transport	MECILLINAM	-7.037
bacteriocin transport	TRICLOSAN	-8.2436
bacteriocin transport	VANCOMYCIN	-8.6649
cellular metabolic process	SULFAMONOMETHOXINE	-3.8657
cellular carbohydrate metabolic process	TRICLOSAN	-3.2215
extracellular polysaccharide biosynthetic process	STREPTOMYCIN	-3.1603
negative regulation of fatty acid biosynthetic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	4.2112
negative regulation of cellular transcription, DNA-dependent	SULFAMONOMETHOXINE	-3.2217
negative regulation of translational initiation	CEFSULODIN	-4.1556
negative regulation of translational initiation	ISONIAZID	-3.5046
negative regulation of translational initiation	STREPTOMYCIN	-5.0699
ATP metabolic process	AMOXICILLIN	-5.5181
ATP metabolic process	BACITRACIN	-4.7329
ATP metabolic process	CEFACTOR	-3.3408
ATP metabolic process	CEFSULODIN	-3.5823
ATP metabolic process	GLUFOSFOMYCIN	-4.6198
ATP metabolic process	NORFLOXACIN	-3.0976
adenosine biosynthetic process	AZIDOTHYIMIDINE	-4.1336
adenosine biosynthetic process	CEFOXITIN	-3.2526
adenosine biosynthetic process	NOVOBIOCIN	-3.8651
thymidine metabolic process	TRIMETHOPRIM0.1SULFAMETHIZOLE	4.2112
L-idoate catabolic process	DOXYCYCLINE	-3.1098
glyoxylate metabolic process	CEFSULODIN6.0MECILLINAM	-4.1072
tetrahydrofolate biosynthetic process	SULFAMETHIZOLE	-5.1158
tetrahydrofolate biosynthetic process	SULFAMONOMETHOXINE	-4.2408
tetrahydrofolate biosynthetic process	TRIMETHOPRIM	-3.6513
folic acid biosynthetic process	SULFAMETHIZOLE	-5.4729
folic acid biosynthetic process	SULFAMONOMETHOXINE	-3.0234
folic acid biosynthetic process	TRIMETHOPRIM	-4.9594

response to antibiotic	CEFACLOR	-3.0007
response to antibiotic	CHLORAMPHENICOL	-3.9441
response to antibiotic	SPECTINOMYCIN	-3.0029
response to antibiotic	TETRACYCLINE	-3.4553
response to arsenic	INDOLICIDIN	3.4847
response to cadmium ion	INDOLICIDIN	4.3828
entry of virus into host cell	ACTINOMYCIND	-3.658
entry of virus into host cell	CARBENICILLIN	-3.0473
entry of virus into host cell	CIPROFLOXACIN	-4.7631
protein stabilization	MITOMYCINC	-3.3843
protein stabilization	TRIMETHOPRIM0.1SULFAMETHIZOLE	-3.142
detection of stimulus involved in sensory perception	CEFOXITIN	3.2379
chaperone mediated protein folding requiring cofactor	MITOMYCINC	-3.108
cell division	GENTAMICIN	-4.1689
cell division	MECILLINAM	-4.4753
regulation of cell cycle	FOSFOMYCIN	-4.0693
oxidation reduction	CISPLATIN	-3.9877
oxidation reduction	THEOPHYLLINE	-3.7967
oxidation reduction	TUNICAMYCIN	-3.1295
intracellular protein transmembrane transport	BACITRACIN	-3.1752
intracellular protein transmembrane transport	CEFOXITIN	3.1586
intracellular protein transmembrane transport	CEFSULODIN	-5.7207
intracellular protein transmembrane transport	CEFSULODIN6.0MECILLINAM	-6.9383
intracellular protein transmembrane transport	CEFTAZIDIME	-3.0095
intracellular protein transmembrane transport	CHIR090	4.3195
intracellular protein transmembrane transport	STREPTOMYCIN	-12.3541
intracellular protein transmembrane transport	TRICLOSAN	-8.7834
intracellular protein transmembrane transport	VANCOMYCIN	-3.8944
double-strand break repair	AZIDOTHYIMIDINE	-7.2012
double-strand break repair	LEVOFLOXACIN	-4.8921
double-strand break repair	MITOMYCINC	-3.3843
double-strand break repair	NORFLOXACIN	-6.1722
double-strand break repair	STREPTONIGRIN	-5.2647

Experimental Procedures

Bacterial strains

The screened collection includes 4334 deletions of non-essential genes (Keio Collection; (Baba et al., 2006), 148 SPA-tagged derivatives of essential genes (Butland et al., 2005), 9 point-mutant alleles of essential genes (plus corresponding “linked” strains for each of the 9 alleles, in which the antibiotic-resistance cassette was linked to the wild-type allele as a control), 5 DAS-tagged essential genes (McGinness et al., 2006), 2 truncations of essential genes, and 100 deletions of small RNAs (sRNAs) and small proteins (Hobbs et al., 2010). All mutant strains are marked with or linked to kanamycin. All linked alleles and controls, DAS-tags, and truncations were constructed in BW25113 (Keio Collection strain background) using standard recombineering methods. Primers are available upon request. SPA-tagged essential genes, originally constructed in W3110 (Butland et al., 2005), were transduced into BW25113 using P1 phage. sRNA and small protein deletions were constructed in MG1655 (Hobbs et al., 2010) and used in that background. For all mutants, except for the sRNA/small protein library, two independently derived clones were screened and evaluated for reproducibility of growth. For the sRNA/small protein mutants a single isolated clone was arrayed twice. Of the 4607 strains screened, 628 were eliminated from the study either because of consistently discrepant growth between clones or because updated curation of the Keio collection reclassified them as incorrect strains, leaving a total of 3979 strains in the final dataset (3737 Keio, 117 SPA, 5 DAS, 9 alleles, 9 linked controls, 2 truncations, and 100 sRNAs). A list of these strains is available online at <http://ecoliwiki.net/tools/chemgen/>.

All individual mutants used for follow up biology stories were re-transduced to BW25113 before further experiments. During the entire study, BW25113 is referred to as wildtype.

Double mutants used in the experiments dissecting the function of MarB were constructed using P1 transduction or recombineering.

Growth conditions for drug synergy experiments

Wildtype and mutant strains were grown overnight to stationary phase in LB broth, and diluted into fresh drug-containing LB broth (0.1 µg/ml TMP and/or 50 µg/ml Sulfamethizole) at 1:1000 v/v. Cultures were then grown aerobically in 150 mL flasks at 37°C and optical densities were recorded at 600nm every 20-30 minutes. Final culture density was recorded at a 24-hour endpoint. The *Streptococcus pneumoniae* experiments used *S. pneumoniae* serotype 2 strain D39 (IU1690; (Lanie et al., 2007), which was cultured statically in Brain Heart Infusion broth (Bacto BHI, Becton Dickinson) at 37°C in an atmosphere of 5% CO₂. Growth was monitored by optical density (OD₆₂₀). Bacterial cultures were started from frozen stocks and propagated overnight in BHI as described (Ramos-Montanez et al., 2008). Overnight cultures in exponential phase (OD₆₂₀ = 0.1-0.3) were diluted back to a density of OD₆₂₀~0.001 to start final cultures. Final cultures were grown to a density of OD₆₂₀~0.1 and divided among 16 mm glass tubes containing no antibiotic, 7.8 µg/ml TMP, 1000 µg/ml SMT or both antibiotics.

Screen Conditions and Data Processing

The collection was arrayed in 1536-format onto six LB-agar plates using a Singer Rotor robot. The independently derived or duplicate clones of each strain were arrayed side-by-side. Glycerol stocks of each plate were kept at -80°C, and re-arrayed at least once every four weeks in effort to reduce the passage and accumulation of spontaneous suppressor mutants throughout the study. For the same reasons and to minimize the exposure time of a given plate at 4°C, the collection was passaged to fresh plates at least once a week.

For each screen, the chemical/stress was mixed into LB agar (except where indicated) at the concentrations listed in Table S1. Plates were allowed to dry for two days, and then pinned with the collection. For the majority of conditions, plates were removed from the incubator following 14-16 hrs growth at 37°C. At this timepoint, fitness differences were apparent but growth had not saturated. Several stresses, e.g. low temperature, required longer incubation times. Plates were then imaged using a Canon G10 digital camera, and the images processed into colony size data, which was used to generate Condition-Gene scores. All image processing, colony size measurements, and score generation was carried out as described previously for genetic interactions (Collins et al., 2006; Typas et al., 2008).

Two key quality control principles utilized in the S-score fitness measurement procedure for genetic interaction analysis are not applicable to our chemical profiling analysis. First, in genetic interaction analysis, each gene-gene interaction is constructed twice, once with each mutation as the donor, and those scores are averaged. Here, condition-gene interactions are constructed only one way, (i.e. stress applied to single mutants), so the scores generated here are analogous to “unaveraged” S-scores (Collins et al., 2006). Second, a key quality control principle underlying the S-score procedure is based on genetic linkage, which is not applicable to Condition-Gene interactions. Therefore, two alternative quality control measures were implemented: (1) the interquartile range (IQR) of the distribution of normalized colony sizes for each screen; and (2) the correlation between replicates of the same screen. Screens with high IQR, or with poor correlation within a series of technical replicates were eliminated from the dataset to minimize noise. High IQR or low correlation of technical replicates was found to be due to a variety of factors including high rates of spontaneous suppressor mutants (due to

selection under high stress), poor or saturated growth and human error in the generation of the data.

The sRNA/protein deletion library was constructed in MG1655, whereas the rest of the collection was constructed in BW25113. As these strains were reasonably close, we did not transduce the sRNA/small protein deletions into BW25113. Therefore, raw colony sizes of all sRNA/small protein mutants (which were arrayed on the same plate of the collection) were normalized separately to eliminate the relative comparisons with the BW25113-based mutants arrayed on the plate. This approach reduced complications from strain background differences, but it should be noted that sRNA/small protein clustering patterns are less reliable, unless the particular mutant was found to have strong phenotypes.

Phenotype Analysis

For each screen, we estimated the mean and IQR of Condition-Gene scores and re-scaled the data to a standard normal distribution (IQR=1.35). We reasoned that if there were no phenotypes (outliers) in a given screen, all scores would conform to a standard normal distribution. For example, we expect 5% of the scores to be less than the 5th normal percentile. More generally, we expect a fraction of scores, p , to be less than the p^{th} quantile of the normal distribution. If a fraction of scores, x ($x > p$), is found to be less than the p^{th} quantile, then the false discovery rate (FDR) associated with the p^{th} quantile would be p/x (Efron et al., 2001).

Using this approach, we classified all scores with FDR 5% or less as phenotypes.

Responsive genes were defined as strains for which at least one phenotype was identified. It is important to note that the 5% FDR attached to each phenotype was

calculated individually for each phenotype. As each phenotype was deemed to be of independent interest, we did not apply a multiple testing correction across phenotypes.

Conditionally-essential genes were identified using a combination of the probability scores and raw colony sizes. Because cells are always transferred to the screen plate during the pinning process, lethal events cannot be scored as colony sizes of zero, unless the stress/drug is bactericidal. Therefore, we defined conditionally-essential events as those in which the probability score was 1 (definite outlier), and raw colony size was less than 60 pixels. These parameters were empirically defined based on known lethal interactions, and likely exclude other true conditionally-lethal events.

Gene Annotation

The COG annotation index (<http://www.ncbi.nlm.nih.gov/COG/>; (Tatusov et al., 2003) was used as the basis for grouping genes according to function. We applied minimal changes to the main COG function families, manually curating genes of interest based on existing COG annotations, published literature (we sampled Pubmed and Ecocyc) and whether evidence was experimental/computational. Genes with purely computationally predicted functions were mostly assigned to the "general function prediction only" or the "unknown function" families. Additionally, each gene was assigned to only a single function.

Protein localization of *E. coli* K12 genes has been reported previously (Riley et al., 2006). Conditionally-essential gene enrichment for envelope proteins was calculated relative to the Keio Collection essential genes (Baba et al., 2006) using Fisher's exact test.

Orphan genes were identified previously, as were ortholog counts across bacterial clades for all *E. coli* K12 genes (Hu et al., 2009). These counts were used to assign all strains in our dataset to γ -proteobacteria-only, proteobacteria-only, or broadly-conserved categories. Genes assigned to the γ -proteobacteria-only category were those for which no orthologs were identified outside the γ -proteobacteria. The proteobacteria-only class contains genes with orthologs outside the γ -proteobacteria, but not outside the proteobacteria. Broadly conserved genes were defined as those for which orthologs had been identified outside the proteobacteria.

RNA isolation, cDNA preparation and Quantitative RT-PCR

Cultures were grown to OD₆₀₀ of 0.3. Samples (8 ml) were harvested and added to ice-cold 5% water-saturated phenol in ethanol solution, centrifuged for 2 min, and the cell pellets were flash-frozen in liquid nitrogen before storing at -80 °C. RNA was extracted using the hot-phenol technique, with modifications. Briefly, cell pellets were resuspended in 500 μ l of lysis solution (320 mM Na acetate, pH 4.6, 8% SDS, 16 mM EDTA), and mixed with 1 ml of water-buffered phenol. The samples were incubated at 65 °C for 5 min with intermittent mixing. The samples were then placed on ice for 5 min and centrifuged for 10 min at 4 °C. The supernatant was extracted twice with phenol-chloroform, precipitated with 2.5 volumes of 100% ethanol and washed with 70% ethanol. The RNA pellet was air dried and resuspended in 85 μ l of RNase free water. Genomic DNA was removed from the samples using Turbo DNA-free DNase Treatment according to the manufacturer's directions for rigorous DNase treatment (Applied Biosystems, Foster City, CA, USA). cDNA was prepared for qRT-PCR as previously described using 5 μ g of input RNA (Cummings et al., 2006).

Quantitative RT-PCR reactions were carried out using Stratagene Brilliant II Sybrgreen master mix according to the manufacturer's directions (Agilent Technologies, La Jolla, CA, USA), and 6 pmol of each forward and reverse primer (Integrated DNA Technologies). Primer sequences are available upon request. Real-time PCR was performed with a Stratagene Mx3000P sequence-detection system (Agilent Technologies). Data were analyzed as described (Vandesompele et al., 2002) with *recA* and *gyrA* as references.

Genomic Context Analysis

For all genomic context plots, base pair coordinates for each gene were adjusted to center the chromosome around the origin of replication (*oriC* = 0 base pairs). Each trace (solid line) represents relative enrichment for the designated gene class in 100 kb windows beginning at the origin of replication and sliding by a 1 kb interval. Enrichment values were calculated using the following approach:

Let x be the number of genes in a given class (i.e. responsive genes) in a 100 kb window out of m genes in that window; let f be the fraction of genes in that class in the whole genome. Our enrichment value is $(x/m - f)$. If we wish to test the hypothesis that $x/m = f$, the chi-squared test statistic would be $m \cdot (x/m - f)^2$. Thus, we used it as our enrichment measure. High (or low) enrichment values would indicate association between class membership and genomic position. Note that this is based on a one-sample test for a proportion. For visual purposes, we plotted the square root of the enrichment values.

The significance of the enrichment is not straightforward to establish since there are correlations between neighboring enrichment values (because of overlapping windows).

To find out how high (or low) enrichment values would be under the null hypothesis of no association between genomic position and class membership, we used a permutation test. Under the null hypothesis, class membership and genomic position would be exchangeable. Therefore, we created 1000 random permutations of the data where class membership of genes was permuted. The max (and min) enrichment values from each permutation were recorded. These provide us with the distribution of the max (and min) enrichment values under the null hypothesis. We used the 95th and 5th percentiles of the max and min enrichment values, respectively, to establish our thresholds of positive and negative spatial enrichment corresponding to a p-value of 0.05.

Coding strand assignments were reported previously (Baba et al., 2006), and were inverted for all genes on the left arc of the *oriC*-centered chromosome to standardize the meaning of the strand designation for all genes. Plus strand indicates the gene is transcribed in the same direction as DNA replication, whereas minus strand indicates the gene is transcribed in the opposite direction of DNA replication. Strand bias values were calculated using the following approach:

Let x be the fraction of genes in a given class (i.e. responsive genes) on the plus strand, and y be the fraction of genes in that category on the minus strand. Also let m be the number of genes on the plus strand in the whole genome, and n be the number of genes on the minus strand in the whole genome. To test for strand bias in a given class, we used the formula $(x-y) / \sqrt{((1/m)+(1/n))}$. This statistic is based on a two-sample test of a proportion.

To establish statistical significance of the strand bias value, we used a permutation test similar to the one described above for genomic context enrichment values. P-values

were calculated directly based on 100,000 permutations of the class designations as described above for genomic context enrichment calculations.

Drug Network Analysis

Gene Ontology Biological Process associations were obtained from the Gene Ontology and EcoliWiki collaborative annotation for *E. coli* K-12 version 1.21 (www.ecocyc.org).

Drug-GO scores used to build the network in Figure 7 were generated using the hypergeometric probability density function in Matlab. Using the binary phenotype/no phenotype data produced as described above (Phenotype Analysis section), this procedure tested for relative enrichment of phenotypes within a GO biological process relative to the entire dataset for each of the 324 screens. Scores were produced by calculating the negative natural logarithm of the resulting p-values, and the sign of the interaction (positive or negative) was re-attached by calculating the mean of all condition-gene scores underlying the phenotypes in a given group. All conditions tested in a concentration series were then collapsed into a single value by taking the mean of Drug-GO scores for each GO group across the concentration series. All connections used to build the network represent Drug-GO interactions with a p-value of 10^{-3} or lower.

Drug-Drug correlations in the network represent Pearson correlation coefficients of 0.32 or higher, which corresponds to a threshold of +2 standard deviations in the distribution of all pairwise Drug-Drug correlation coefficients. Here, condition-gene scores for drugs screened in a concentration series were averaged to generate a composite signature for each drug. These composite signatures were used for calculation of the correlation coefficients.

The network in Figure 7 was generated using Cytoscape version 2.6.3 (Shannon et al., 2003). Clustering was based on the Edge-weighted Spring-Embedded algorithm, using the P-values of Drug-GO interactions and high Drug-Drug correlations to drive the positions of nodes in the network.

References

- Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol* 2, 2006 0008.
- Butland, G., Peregrin-Alvarez, J.M., Li, J., Yang, W., Yang, X., Canadien, V., Starostine, A., Richards, D., Beattie, B., Krogan, N., *et al.* (2005). Interaction network containing conserved and essential protein complexes in *Escherichia coli*. *Nature* 433, 531-537.
- Collins, S.R., Schuldiner, M., Krogan, N.J., and Weissman, J.S. (2006). A strategy for extracting and analyzing large-scale quantitative epistatic interaction data. *Genome Biol* 7, R63.
- Cummings, C.A., Bootsma, H.J., Relman, D.A., and Miller, J.F. (2006). Species- and strain-specific control of a complex, flexible regulon by *Bordetella* BvgAS. *J Bacteriol* 188, 1775-1785.
- Efron, B., Tibshirani, R., Storey, J.D., and Tusher, V. (2001). Empirical Bayes Analysis of a Microarray Experiment. *Journal of the American Statistical Association* 96, 1151-1160.
- Girgis, H.S., Hottes, A.K., and Tavazoie, S. (2009). Genetic architecture of intrinsic antibiotic susceptibility. *PLoS One* 4, e5629.
- Hobbs, E.C., Astarita, J.L., and 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.
- Hu, P., Janga, S.C., Babu, M., Diaz-Mejia, J.J., Butland, G., Yang, W., Pogoutse, O., Guo, X., Phanse, S., Wong, P., *et al.* (2009). Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS Biol* 7, e96.
- Jeanguenin, L., Lara-Nunez, A., Pribat, A., Hamner Mageroy, M., Gregory, J.F., 3rd, Rice, K.C., de Crecy-Lagard, V., and Hanson, A.D. (2010). Moonlighting glutamate formiminotransferases can functionally replace 5-formyltetrahydrofolate cycloligase. *J Biol Chem*.
- Lanie, J.A., Ng, W.L., Kazmierczak, K.M., Andrzejewski, T.M., Davidsen, T.M., Wayne, K.J., Tettelin, H., Glass, J.I., and Winkler, M.E. (2007). Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. *J Bacteriol* 189, 38-51.
- McGinness, K.E., Baker, T.A., and Sauer, R.T. (2006). Engineering controllable protein degradation. *Mol Cell* 22, 701-707.
- Ramos-Montanez, S., Tsui, H.C., Wayne, K.J., Morris, J.L., Peters, L.E., Zhang, F., Kazmierczak, K.M., Sham, L.T., and Winkler, M.E. (2008). Polymorphism and regulation of the *spxB* (pyruvate oxidase) virulence factor gene by a CBS-HotDog domain protein (SpxR) in serotype 2 *Streptococcus pneumoniae*. *Mol Microbiol* 67, 729-746.
- Riley, M., Abe, T., Arnaud, M.B., Berlyn, M.K., Blattner, F.R., Chaudhuri, R.R., Glasner, J.D., Horiuchi, T., Keseler, I.M., Kosuge, T., *et al.* (2006). *Escherichia coli* K-12: a cooperatively developed annotation snapshot--2005. *Nucleic Acids Res* 34, 1-9.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-2504.
- Tatusov, R.L., Fedorova, N.D., Jackson, J.D., Jacobs, A.R., Kiryutin, B., Koonin, E.V., Krylov, D.M., Mazumder, R., Mekhedov, S.L., Nikolskaya, A.N., *et al.* (2003). The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4, 41.
- Typas, A., Nichols, R.J., Siegele, D.A., Shales, M., Collins, S.R., Lim, B., Braberg, H., Yamamoto, N., Takeuchi, R., Wanner, B.L., *et al.* (2008). High-throughput, quantitative analyses of genetic interactions in *E. coli*. *Nat Methods* 5, 781-787.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3, RESEARCH0034.

Chapter 3: Future Perspectives

I have written this chapter independently. Carol Gross assisted with revisions and edits.

The tremendous impact of the genomic revolution has inspired the burgeoning field of phenomics. Following the traditional model, the gateway technologies for high-throughput discovery of gene function were developed primarily in the model organisms, and recently have expanded into many applied systems. Going forward, the field will advance on three major fronts: screening readout technology, monoculture species application, and mixed community phenomics.

As more approaches are adapted to sequencing platforms, researchers will have the ability to survey increasing swaths of phenotypic space efficiently. Especially for arrayed approaches, the addition of diverse phenotypic readouts will push the field beyond its current “fitness-centric” state. Transcriptional reporters, indicator dyes for developmental processes, and metabolomic integration represent coming advances, but are likely just the tip of the iceberg.

Sequencing and genetic technologies will open up a growing list of bacterial genes and species for phenomic analyses. The recent demonstrations of phenomic analyses based on antisense RNA should allow LOF studies of essential genes in gram-positive and gram-negative species. Further engineering of the Mariner transposons will continue to expand their utility and allow for cutting edge phenomic studies of many applied and environmental species.

Last, and perhaps most excitingly, phenomics will expand into mixed community applications. Exciting work has already been done examining the ability of genetically variant bacteria to colonize the mouse gut in controlled mixed communities (Goodman et

al., 2009). *In-vitro* coculture models also present exciting opportunities for phenomics aimed at deciphering complex interspecies and interkingdom relationships.

Overall, the future is extremely bright for phenomics. As genomics technologies continue to inspire and drive the development of functional genomics technologies, phenomic science will take center stage, as rapid and efficient techniques to assign meaning to sequence will be in greater demand than ever before.

References

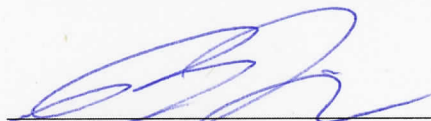
Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. (2009). Identifying Genetic Determinants Needed to Establish a Human Gut Symbiont in Its Habitat. *Cell Host and Microbe* 6, 279-289.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.



Author Signature

April 9, 2012

Date