博士論文番号：1481026
(Doctoral student number)

RNA guided system for genetic interaction analysis involving essential genes in *Escherichia coli*.

Yang Xi
Nara Institute of Science and Technology
Graduate School of Biological Sciences System Microbiology
Supervisor: Mori Hirotada

2018/01/22

Contents

Abstract

Escherichia coli is one of the best-studied organisms and is generally used as a model organism in fundamental and applied biological research. The study of homologous genes and proteins, and molecular mechanisms has been proven useful in the understanding of more complex systems in diverse organisms, including humans[1]. In addition, many bacteria make products of medical and/or industrial importance, which can often be synthesized in *E. coli* by introduction and engineering genes for the respective enzymes and pathways in an *E. coli* chassis, thereby creating an *E. coli* bio-factory. Thus, there is considerable interest in *E. coli* not only as a model system for elucidation of fundamental biological problems but also as a platform for production of medically or industrially important materials.

Despite studying *E. coli* extensively for several decades, the physiological function of numerous genes remains poorly understood or uncharacterized[2]. A number of powerful, high-throughput approaches are no being used to help gain a deeper understanding of gene and protein function to achieve a more complete understanding of the *E. coli* cell. These include unraveling physiological network structures, defining genetic interactions, and many others.

Synthetic lethality (or extremely poor growth) can result when two independent mutations are combined within the same strain, when neither mutation alone has an adverse effect on cell growth. Such effects can result when the mutations interrupt functionally related functions, allowing one to infer the role of a gene whose physiological roles was unknown. Analyses of such genetic interactions thus provide a powerful tool to map out the functional linkage among genes[3]. The quantification of a genetic interaction includes a quantitative measure of phenotypic change such as the growth rate and neutrality function, which can be used to predict the expected phenotype of an organism carrying two non-interacting mutations[4].

Roughly 7% of ca. 4500 *E. coli* genes are essential[5], implying that they encode a function that is critical cellular function. Genetic interaction analysis of essential genes can provide key information for more complete understanding and system-level organization of a living cell. Due to their essentiality, discovering their precise role or genetic interactions with other (essential or non-essential) genes has generally required the availability of conditional mutations, for example temperature sensitive allele, to uncover their impact on cell physiology[6].

Purpose:

The goal of my research is to develop and test a method to examine the genetic network structure of essential genes in *E. coli*, which did not depend on the availability of conditional lethal mutations. The approach that I took was to develop a tool that enabled decreasing expression of essential genes, thereby conditionally affecting cell growth. Once developing such a method, my plan was to examine quantitatively interactions between essential genes whose growth was impeded by decreased expression and non-essential *E. coli* genes.

Method:

Here, we describe a procedure to inhibit cell growth by decreasing the expression of an essential gene. Our method depends on use of a CRISPR-Cas9 gene targeting system containing a mutant Cas9 protein that binds targeted RNA-DNA complexes but does not cleave the DNA. We will engineer conjugative plasmids carrying such CRISPR-Cas9 knockdown systems designed to knockdown expression of different essential genes. By using conjugative CRISPR-Cas9 knockdown plasmids, we would be able to introduce these plasmids in a wide variety of different hosts by conjugation. We would examine quantitatively examine growth of the CRISPR-Cas9 knockdown plasmids in single-gene deletion mutants, as a means to analyze genetic interactions between essential and non-essential genes systematically.

CRISPR-Cas9 is an adaptive immune system discovered from bacterial and archaea[7] and have recently been developed as a powerful genomic editing tool. CRISPR-Cas uses small base-pairing RNAs to target and cleave foreign DNA elements in a sequence-specific manner[8]. CRISPR-Cas9 is presently the most developed DNA modification system in CRISPR family. The main components of type II CRISPR-Cas systems contain genes coding a Cas protein and two RNAs, a mature CRISPR RNA (crRNA) and a partially complementary trans-acting RNA (tracrRNA)[9]. CRISPR interference (CRISPRi) platform has recently been developed as a tool to control gene expression in *E. coli*. The CRISPRi platform encodes a mutant Cas9 protein that still binds the targeted DNA but lacks endonuclease activity and therefore does not cleave the target DNA[10].

I introduced CRSPRi on a self-transmissible conjugative plasmid in order to generate essential gene knockdown strains in a high-throughput manner by conjugation. I have demonstrated that conjugative CRISPRi plasmids can be used to generate double mutants by

crossing a CRISPRi plasmid into recipients carrying single-gene knockout mutants as a means to interrogate genetic interactions between essential and non-essential genes. In my thesis research, I developed and evaluated the performance of this system as a means to analyze the essential gene genetic network structure.

Chapter 1 Introduction

1.1 E. coli science

Escherichia coli Gram-negative, facultative anaerobic and non-sporulating intestinal bacterium that is rod shaped cell and approximately 1-3μm long and 0.5μm wide. Most *E. coli* are non-pathogenic (like *Escherichia coli* K-12 and B strains), A small number of *E. coli* are serious disease-causing pathogens, including Enterohemorrhagic *Escherichia coli*(EHEC), such as 0157/H7, *uropathogenic* (UPEC) and *enterotoxigenic* (ETEC) strains.

Escherich coli was discovered by a German pediatrician Theodor Escherich (1857-1911) in 1885 which he isolated a fast-growing bacterium that he called "*Bacterium coli commune*"[11]. He characterized its fermentation patterns and showed that grows at different rates on different carbon source[12]. This discovery was prompt applied by many laboratories, and in 1919, scientists renamed this bacterium as *Escherichia coli* to recognize Theodor Escherich's contribution on this field.

Beginning the late 1920s, *E. coli* was often used to study physiology due to its hardiness, versatility, and fast growth in different medium. These features made *E. coli* a mainstay in research in early 20th century and one of the most widely available model organisms [13] [14] [15] [16] [17]. By the 1940s, *E. coli* was used mostly in biochemistry or enzymology[18]. The discovery of conjugation in the *E. coli* K-12 by Ledeberg[19], greatly accelerated the use of *E. coli* K-12 in research for it opened the door to creating the new field of bacterial genetics. it extends dramatically its use in the field of genetics and accelerated the launch of molecular biology with monumental achievements including discovery of genetic code, molecular mechanisms of transcription, translation, DNA replication and many other biological processes[20].

Nobel prize French winner, Jacques Monod indicated that "Anything found to be true of *E. coli* must also be true of elephants"[21]. This concept has been gradually proven by scientists in many biological processes. This biological universality between bacterium and other organisms including human leads model organisms, such as *E. coli*, Yeast, *C. elegance*, drosophila and mouse, very important research targets to clear biological rules. And accumulation of biological

knowledge gives us a possible way of application, such as biosynthesis of valuable compounds like human insulin by *E. coli*.

By studying basic processes such as bacterial conjugation and phage infection with *E. coli* K-12 and other strains like *E. coli* B, phenomena of restriction and modification were uncovered, which was responsible for the discovery of restriction enzymes and the development of DNA cloning, in 1973 by Cohen and Boyer [22] and development of recombinant DNA[23]. An early application of recombinant DNA led to the production of human insulin in *E. coli* [24]. *E. coli* has been used in vaccine development, bioremediation, as well as the production of many valuable products in industry [25].

Studies of bacterial conjugation led to the discovery of multidrug-resistant plasmids carrying antibiotic resistant genes discovered. The widespread distribution of such plasmids is a major clinical problem. *E. coli* research has provided much basic knowledge on the molecular mechanisms of how such plasmids replicate and are transferred between divergent bacteria. Continued research in this area may someday provide new ways to combat the drug-resistant problem by interfering with conjugative functions.

1.2 Systems biology

Systems biology has greatly expanded with the advent of whole genome sequencing near the end of the 20th century. The concept itself, however, is not new. In 1969, Savageau showed theoretical approach for enzymatic reaction as a system[26]. The main difference between analysis in the past and those in the current concern the size of target biological systems and amount of biological experimental data being analyzed, mostly resulting from technological advances.

What is systems biology, "It is a **holistic** approach to deciphering the complexity of biological systems that starts from the understanding that the networks that form the whole of living organisms are more than the sum of their parts." defined by the institute of systems biology, Seattle USA.

In discussions of systems biology, one should mention the largest international cooperation project— Human Genome Project (HGP). HGP started since 1990 and the draft genome structure had been published at the beginning of 21st century by two groups[27], finished at 2003,

during the 13 years, the scientists from different country together figured out the hole human genome sequence, clarified approximately 22300 proteins -coding genes in human bodies, and less than 7% of the protein families are vertebrate specific[28].

HGP revealed that only 1.5% of the genome is protein coding [29], the rest of the genome is non-coding sequences, including microRNA, siRNA, piRNA, snoRNA and lncRNA, etc[29]. These non-coding sequences are part of biological processes and mostly have regulatory functions of the cell processing. In conclusion, the complexity of living cells cannot be simple as just assignment of function to each of predicted genes from the genome sequence. In a real cell, all of genes' product interact each other and show biological functions. Therefore, it is quite important to examine target genes as part of cell systems and to analyze gene-gene interactions. From that on, along with the development of computational and mathematical researches, system biology, as an interdisciplinary subject becomes more and more important.

Systems biology mostly focuses on the complex interactions within biological systems in different layers. One of the important approaches is "OMICs" type data collection to analyze global identification and measurement of unpredicted dynamic changes and interactions. Recent technology innovation is so quick and disciplinary cooperation is becoming so important to solve biological systems. Systems biology is one of the excellent examples of this successful cooperation between physics, chemistry, engineering, mathematics, computer science and experimental biology fields.

Biologists can define different areas for systems based on their research fields, purposes and methodologies. From the systems analysis, we usually get comprehensive results from hundreds and thousands discrete events, to quantitative analyze the intricate phenomena make it possible to read the results precisely and confidently. From mathematic models, we would modify or engineer some certain factors in the system to gain our purposes. In fact, most of the models are complicated and unintuitive, computational methods provide a strong support for analyzing the numerous phenomena.

One of the applied direction helping our lives is the development of synthetic biology. Synthetic biology utilizes cell pathways to produce valuable products. One example in this subject is to generate polyhydroxybutyrate (PHB) using microorganisms[30]. PHB is a material for plastic and it previously made by hydrocarbons derived from petroleum. In the early 1990s

however, Oliver Peoples developed an industrial strain which can convert corn sugar to PHB monomer[31] and extend to industrial size of production to make green biology production of plastic precursor. This success greatly encouraged people not only for the value of final products but also for advantages of environmentally clean processes to obtain the final products. Even better, this bioplastic is degradable by environmental microorganisms without any pollution.

To analyze biology on the system level, the first thing we should clear is "what to look for". System biology research focus on the structure of the system and its dynamics but not each individual factor. Identifying all the genes and proteins in an organism provides a list of all the components in the organism, for system research, we want to know how these parts are assembled to form the structure of the system and what will happen if we change one part of the system.

Robustness is an essential property of a biological system[32], it is necessary to understand the biological system robustness to detect the biology on a system level. It is a fundamental and ubiquitous feature of a system, it refers to the property of the system to maintain its functions against internal and external perturbations[33].It is an important indicator to understand the principles underlies a complex biological system based on the phenomena. The biological system usually consists of a set of highly conserved core processes that links to diverse outputs and inputs, the core processes are rigidly maintained, the inputs and outputs are very diverse. Thus, endow the system with the capacity to be robust. One of the examples of robustness is *E. coli* chemotaxis, *E. coli* can sense a wide range of attractant concentrations and change intracellular components to ensure the perfect adaptation[34].

Essential gene is the gene for survival in LB condition at 37°C in *E. coli*, there are about 300 essential genes out of 4000 genes in *E. coli* [5].A big challenge in detecting of developmental processes is the identification of all of the relevant genes and their functional relationships. Defining gene essentiality is not simple because it depends on a variety of physiological properties, such as growth media, temperature, environmental factors, and the genetic background. The study on essential genes used to be an almost impossible mission because of indispensability feature in the cell and its sensitivity to the outside condition. Recently, with the help of technology development and the accomplishment of the human genome project, especially the high-throughput technology establishment, systematically determine genome-

wide gene essentiality has been achievable. Scientists in different groups identified essential genes by various genetic tools, like transposon mutagenesis[35], genetic footprint technique[36], conditional mutation[5], in different species, for instance, *S. cerevisiae*, ~19% of the genes are essential[37]~7 % in *E. coli*[5] and *Bacillus subtilis*[38], ~10% in *Spreptococcus sanguinis*[39] and 16 % in *Acinetobacter baylyi*[40].

Elucidating the physiological functions of essential genes is crucial to a full understanding of a cell. Essential genes usually are more conserve in evolution than non-essential genes[41]. They draw a lot of interests of scientists not only because their essential roles in cell processes, but also because of their practical applications, for example, essential genes might be good targets for pathogens, and most of the essential genes are the hubs of biological groups, some precious products are achieved by modification of them.

In essential gene analysis, one troublesome is the essentiality plausible of the genes. Even in the same species, scientists would identify different essential gene sets, for instance, 620 genes in *E. coli* are identified as essential in LB medium by genetic foot printing technique[36], but 328 by deletion lethal strategy[5]; and about 119 genes in *E. coli* non-essential gene list in LB medium shows indispensable in minimal medium with glycerol[42].

1.3 Systems approach and my purpose

Since the genomic sequences are confirmed, information about the total number of genes coded on the chromosome and their predicted coding regions are available, this ORF information takes the advantage of the technology developments, scientists prepared the entire set of ORF clone and deletion mutant libraries. This type of comprehensive resources is important for high-throughput experiments and for direct comparison of the ORF deletion mutants on the same genetic background.

In *E. coli*, several comprehensive libraries are constructed.

(1). ASKA ORF library[43]. ASKA is ORF plasmid clone library, which contains all the predicted coding regions.

(2). Keio collection[5]. This is a single gene deletion library of predicted ORFs of *E. coli* K-12 expect essential genes.

(3) Random insertion mutant library of *E. coli* K-12[44]. This is a random insertion mutant

library generated by Tn10 derivative transposon and mutagenized lambda Kohara clones.

(4) Essential gene deletion library[45]. Essential genes are indispensable for cellular survival in LB medium at 37 °C. The essential gene deletion library is made by eliminating the essential genes from the chromosome under the condition of in trans complementation from a low copy plasmid.

(5) Chromosomal Fusion with GFP protein[46]. This is an in-frame chromosomal fusion of the target genes related to the central metabolic pathway with modified GFP.

These resources opened new areas especially in the research on the system level; the entire gene sets comparison on the same genetic background is possible by these comprehensive resources.

The genetic interaction which also termed "epistasis", this term was first proposed by Bateson (1909)[47]as a supplementary explanation of Mendel's principles of heredity. Bateson used this word to describe the effect of the allele on one locus mask the other locus and prevent the variant at the second locus from manifesting its effect. Later, people published an additive linear model to connect multi locus genotype values to phenotype values in the statistical way [48]. This combination makes it possible for the scientists to quantitatively measure phenotype of mutants and then trace back to evaluate the genotype.

The consequence of two mutations could be classified into 3 types. Additivity[49], magnitude epistasis[48 50 51], and sign epistasis[52]. Additivity represents two mutants have an effect on a linear model, which means two gene loci have the independent effect, when combining them together, the phenotype clearly shows the sum of two single mutants. Genes in this group usually function in different metabolic pathways, for example, the color, shape, height of pea seeds in Mendel's experiments[49 53]. Magnitude epistasis termed as the phenotype that two gene locus mutations show clear deviations from the expectation of two single mutations phenotypes. If two mutations together show enhancement of the function, it is called positive epistasis[48 54], usually, this result indicates two mutations belong to the same pathway. To the contrary, if the double mutations show weaker effect than the sum of two single mutations effect, this phenotype is called negative epistasis[55 56], the alleles in complementary pathways will present this phenotypic effect. Be attention, there is another phenotype that, when two mutations are introduced, the cell shows the opposite effect of one or both single mutations[52], for example, the toxic- antitoxic systems. If the combination of two viable single mutations shows lethal, this is called synthetic

lethality[57].

To analyze the genetic interactions, we can choose different phenotypes as the objective to introduce different gene pair combinations and evaluate the interaction types and degrees of the mutated genes or gene loci. A very common phenotype for epistasis detection is fitness, one of the advantages of using fitness as the indicator might be the universal sensitivity and quantitative ability, it could be studied by both wet experiment and computational simulation[58] [59] [60]. Quantitative to see, if we define the fitness of wild type strain as 1, and the relative fitness of gene A mutant is f(A) in the range (0<= f(A) <=1). And geneB mutant f (B) in the range (0<=f(B)<=1). If gene A and gene B double mutation strains fitness f (AB) = f(A) * f(B), it means gene A and gene B are additive epistasis; if f(AB) <f(A)*f(B), gene A and gene B are negative epistasis; if f(AB) >f(A) *F(B), gene A and geneB are positive epistasis. If f(AB) = 0, gene A and gene B can be defined as synthetic lethality.

To understand the natural mechanism of the biological behaviors, we need to collect as much as possible of the information from lives, but at the same time, a challenge is that it becomes a tough work to link the phenotype to its relative genotype. Hence, statistic and computational tools are necessary for biologists to analyze cell basic rules. For the wet experimental side, people need to provide biological data with enough size and list the potential genetic interactions base on the accumulated biological knowledge. Nowadays, a popular strategy to analyze gene-gene interactions is to monitor the phenotypes of double knockout mutants in whole genome wide by high-throughput way, and some single gene mutation libraries are constructed for different organisms, like YKO library in *S. cerevisiae*[61], Keio Collection in *E. coli* [5], RNAi library *in C. elegans*[62] and in *D. melanogaster*[63].

(I) Synthetic genetic arrays (SGA)[64], SGA is a high-throughput method to detect synthetic lethal and synthetic sick genetic interactions. It provides a strategy to combine two mutations together by mating and check the survival ability. SGA is first developed in the model organism *S. Cerevisiae*, with single gene deletion library, SGA can identify synthetic lethal or synthetic sick gene pairs in an entire *S. Cerevisiae* genome.

(II) Heterozygous diploid-based synthetic lethality analysis on microarrays (dSLAM)[65], dSLAM is another high-throughput method to analyze synthetic lethality and synthetic sickness mutations combinations. It introduced a specialized haploid-selective marker into the heterozygote diploid YKO strains. When transfer the query mutation in the genome-wide single

gene mutants, the specific molecular barcodes or tags make it possible to find out the sick or lethal gene combinations on microarray plates.

(III) Epistatic mini-array profiles(E-MAP)[58] is a way to quantities measure the phenotypes deviations of double mutants' strains, it can give us the candidates not only which have synthetic lethal or sickness effects, but also the candidates who have positive genetic interactions.

From the high throughput monitor systems (like Colony Live system for *E. coli*), we can quantitative score the phenotypes, and evaluate the underlying genetic interactions between two target genes. Colony-live system[66] combines commercial devices, like scanners and microbial incubators with in-house developed software, to monitor colony growth kinetics. The colony sizes are being captured every 30 minutes, and finally it will provide us colony mass on each time points and calculated values. The output data include lag time of growth; maximum growth rate; saturation point of growth and colony area. It is a highly compatible system, which uses standard rectangle plates with maximum 1536 mutants on each plate and is competent to different experimental conditions.

1.4 CRISPR system and my purpose

CRISPR-Cas is the abbreviation of clusters of regularly interspaced short palindromic repeats and associated proteins. CIRSPR system is an adaptive immune system that discovered from bacteria and archaea. It is an RNA directed way for prokaryotes to resist infective agents, like plasmid or virus. At the first infection, the system captures a short specific sequence, and cut a short DNA fragment in front of this sequence as a memorized marker, and combine with associated proteins, when this infective agent infects again, this motif will recognize this invader and get rid of genetic threaten.

The discovery of CRISPR system can be traced back to 30 years ago, in 1987, Ishino from Osaka University revealed fourteen repeats of a twenty-nine nucleotide, and these repeats are separated from each other by variety short DNA fragments about thirty nucleotides in *E. coli*[7]. Some years after, the similar DNA arrangement is found by some other groups in different bacterial or archaea species, like *Mycobacterium tuberculosis*[67]; Archaea *haloferax mediterranei* and *Haloferax volcanii*[68]. 10 years after that, several groups together gave a hypothesis that this repeated motif may come from plasmids or phage DNA, and this CRISPR system functions as

a prokaryotic immune system. Until recent decade, this system is well analyzed and its immunity function is confirmed by several different groups in different bacterium[69] [70].

This CRISPR system consists of a characteristic repeat - spacers array and Cas proteins. CRISPR array contains a special structure, it has a series of repeats with the same nucleotide sequence and have around 25 to 40 nucleotides distance from each other[71]. The sequences between the repeats are various to resist different invading genetic material, but they have similar size in a certain organism. The length of spacers is very different in different species, the spacer size is detected from about 20nt in bacterial to around 100nt in archaea[72]. Base on the secondary structure, size, and sequence similarity of repeats, the CRISPR arrays could be classed in 12 families[73]. A leader DNA fragment is found essential for CRISPR immune function, this part is involved in the early infection stage, to memorize the unique invader sequence, and assemble the RNA protein complex for the future resistance of the invader[73]. CRISPR associated proteins (also called Cas protein) are searched as proteins only present when CRISPR system exist in prokaryotes. More than 45 Cas9 protein families are discovered [74]. These proteins are widely spread in different organisms that with CRISPR systems, but only Cas1 and Cas2 proteins are common exist in all the organisms. Cas3, Cas4, Cas5 and Cas6 are also found in many organisms, Cas1 and Cas2 are the best-conserved Cas proteins, and known at immunity acquisition stage. Cas1 has endonuclease activity, Cas2 is a metal dependent endoribonuclease, when the foreign genetic material first invades, Cas1 and Cas2 cut the invader DNA and degrade the DNA to short fragments[75]. Cas3 is a helicase, it functions at the immune stage, Cas4 protein has similar structure to RecB exonulease, but the function is still not clear[76]. Cas5 protein is been reported as a RNA binding protein; probably function in crRNA processing procedure. Cas6 protein involved in crRNA precursor process, to generate mature protein and crRNA complex[77] [76]. Besides these common exist genes, different organisms have different CRISPR proteins, like Cse in *Escherichia coli*, Csm in *Mycobacterium*, Csa in *Aeropyrum* and so on[73]. This variety of the same function immune system suggests that, for the same purpose, different organisms chose different pathways in the evolution.

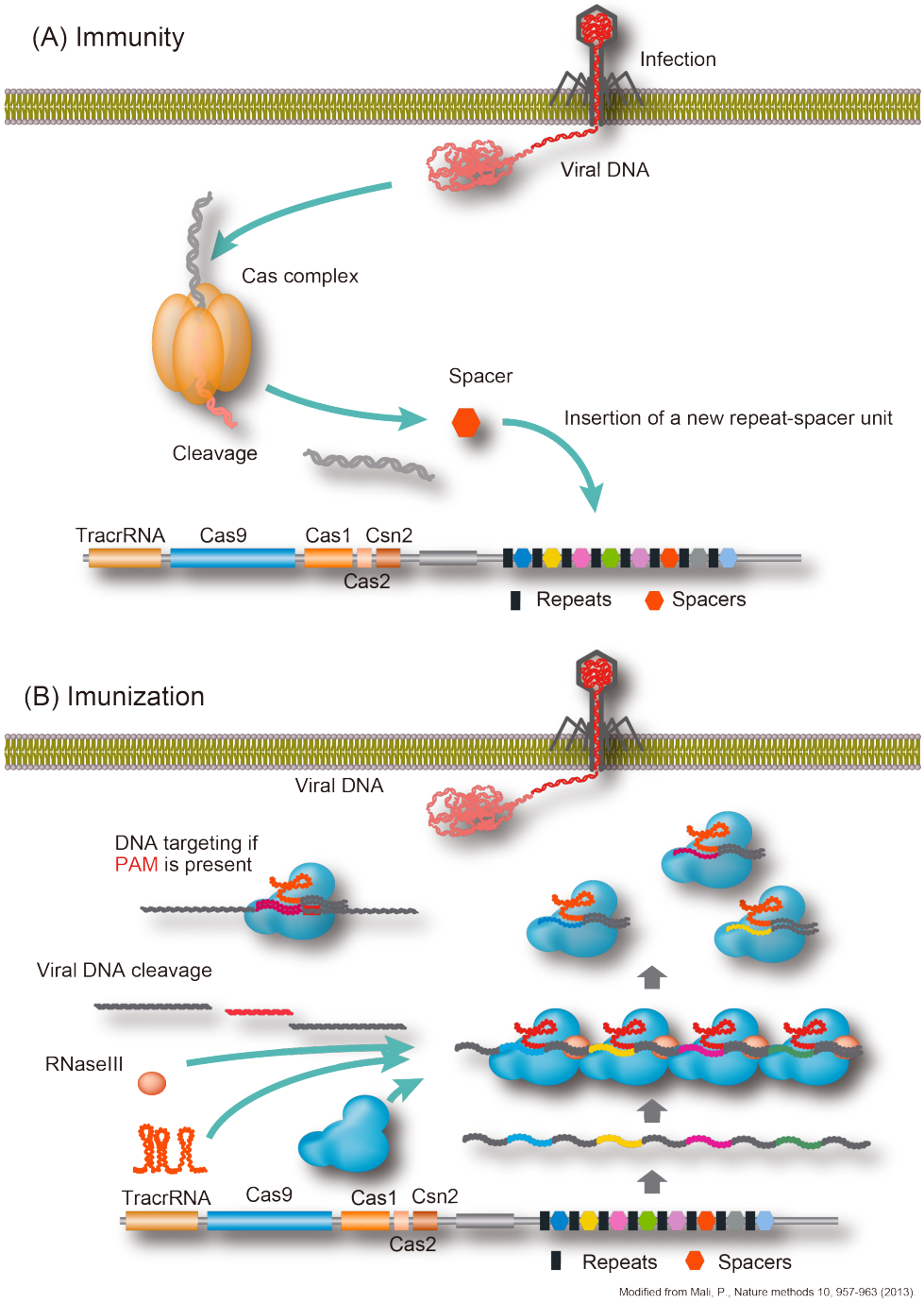


Figure 1: Type II CRISPR-mediated immunity. (A): Immunization step. CRISPRs are the adaptive immune system in bacterial and archaea. The natural CRISPR regions are composed of short DNA repeats (black boxes) and spacers (colored hexagons). When the virus first infects a bacterium, the virus DNA would be captured by the Cas proteins complex (orange polymers), a short fragment from the invading DNA will be cut and be incorporated into the spacer locus to generate one extra spacer. (B): Immunity step. The CRISPR sequence is transcribed to an RNA precursor which comprises repeats (dark grey) and spacers (colored), the RNA precursor combines with tracrRNA (red loop) and Cas9 (blue) to form a long RNA protein complex, the complex is modified by the host native RNaseIII (orange oval), and finally to generate short Cas9 protein and RNA complex. The spacer RNA guides the protein RNA complex to a matching target sequence in the invading virus, and finally cuts up the invading viral genome.

Similar to the other adaptive immune system, CRISPR system function by two stages, adaption stage; memorize and silence the invader stage (Fig 1). In the first stage, CRISPR system will search a specific motif called protospacer adjacent motif (PAM) sequence, and find the short DNA fragment in front of this PAM sequence (protospacer) on invader genome, after that, Cas1 and Cas2 proteins will cut the protospacer from invader genome and integrate the short fragments in the host genome. This insertion position is being called spacers, in this adaptive stage, the PAM sequence is essential, usually, it is a short fragment with 2-4 base pairs, and is unique for a certain type of CRISPR system. Protospacer is homologous to the spacer sequences on host genome, the research indicates that this spacer is always inserted after leading sequence and before repeat locus, so it is possible that the order of spacers on CRISPR array represents the chronological order of foreign infection [78]. The memorization of foreign DNA initiates from crRNA biosynthesis, after the insertion of the spacers in CRISPR array, transcription starts from the leader sequence in front of CRISPR array, and generate a long crRNA precursor that contains all the spacers and repeats RNA transcripts, crRNA precursor (pre-crRNA) is thereafter processed to short single spacer and repeat transcript units[73]. This process mostly depends on endonuclease homologous Cas proteins, like Cas6 in *P. Furiosus*[77]. While, in *E. coli*, this process is accomplished by a Cas protein cascade. This cascade mostly consists of 5 proteins, called Cse1, Cse2, Cse3, Cse4, and Cse5. in *E. coli*, Cse3 may function as the endonuclease to cut the pre-crRNA at 5' ends of each spacer, and then other proteins modify 3' ends of the short RNA products[79], however, the mechanism underlying this procedure is still unclear yet. When the crRNA is prepared, it associates with Cas proteins to form an RNA protein complex. The complex recognizes the foreign invader genetic elements by base-pairing of crRNA and foreign DNA[71], and cleaves the foreign DNA by host-specific Cas proteins (Fig 1), like Cas5 in *S. Thermophilus*; Cas9 in *S. Pyogenes*[80] [81]. After long-term evolution, this CRISPR array acts as a memory of the previous infection. In this system, PAM and protospacer sequences are critical, but when the invader DNA sequences have point mutations on some parts of PAM or protospacer sequence, it won't inhibit the immunity of the system[82]. Only if the mutation occurs in critical positions, CRISPR will be inhibited, however, the partial mismatch of crRNA and foreign DNA will promote the acquisition of new protospacers[83].

CRISPR systems are diverse across different species, the spacer-repeat unit in CRISPR array could be from one to more than 300[84] [72], on the other hand, PAM sequence and protospacer are also quite different both in their sequences and length in different species[85]. To date, a commonly accepted classification way to dived complicated CRISPR systems is to use crRNA

effector complex, that is when defense the foreign DNA, which Cas protein or protein cascade is associated to the specific crRNA[86]. Until 2016, about 93 CRISPR associated protein families are been discovered, they are classified into two large categories base on crRNA effector complexity. In class 1 CRISPR Cas systems, crRNA defense procedure need multi-proteins for crRNA binding and targeting, for instance, the CRISPR system in *Archaeoglobus fulgidus*; *Bacillus halodurans*; *Escherichia coli; staphylococcus epidermidis*; *cyanothece* and so on[87]. In the other class, crRNA needs only single protein to accomplish crRNA binding and targeting function, like the systems in *Legionella pneumophila*; *Streptococcus thermophilus*; *Neisseria lactamica*[87]. Base on significant cas genes in cas operons, Class 1 CRISPR system is divided into type I, type III and type IV; class 2 is divided to type II, typeV and type VI [87]. Each type has a symbolic Cas protein, like Cas3 for type I [88 89 90], Cas9 for type II[9 91 92 93 94], Cas10 for type III[95 96 97],Csf1 for typeIV[97 98] and Cpf1 for type V[98 99]. Type I CRISPR system has 7 subtypes, named type I-A, type I-B, type I-C, type I-D, type I-E, type I-F and type I-U. In type I-C, type I-D, type I-E and type I-F systems, Cas proteins are encoded from a single operon, while, in type I-A and type I-B, Cas proteins are encoded from multi-operons[98]. Type II CRISPR systems are known as the simplest CRISPR system since it has only one mark gene (cas9) for CRISPR system, it is divided into 3 subtypes base on the Cas protein besides Cas9 on the operon, or distinct locus organization[9 100]. Type III is classified into two subgroups, type III-A and type III-B, they are distinguished base on a subunit protein-encoding gene csm2 for III-A type and cmr5 for III-B type[98]. Type IV is functional unclear CRISPR system, the unique feature in this group is that they don't have wide spread Cas proteins: Cas1 and Cas2[97]. Type V is a putative CRISPR system with a significant cpf1 gene adjacent to cas1, cas2 and CRISPR array[98] (Fig 2).

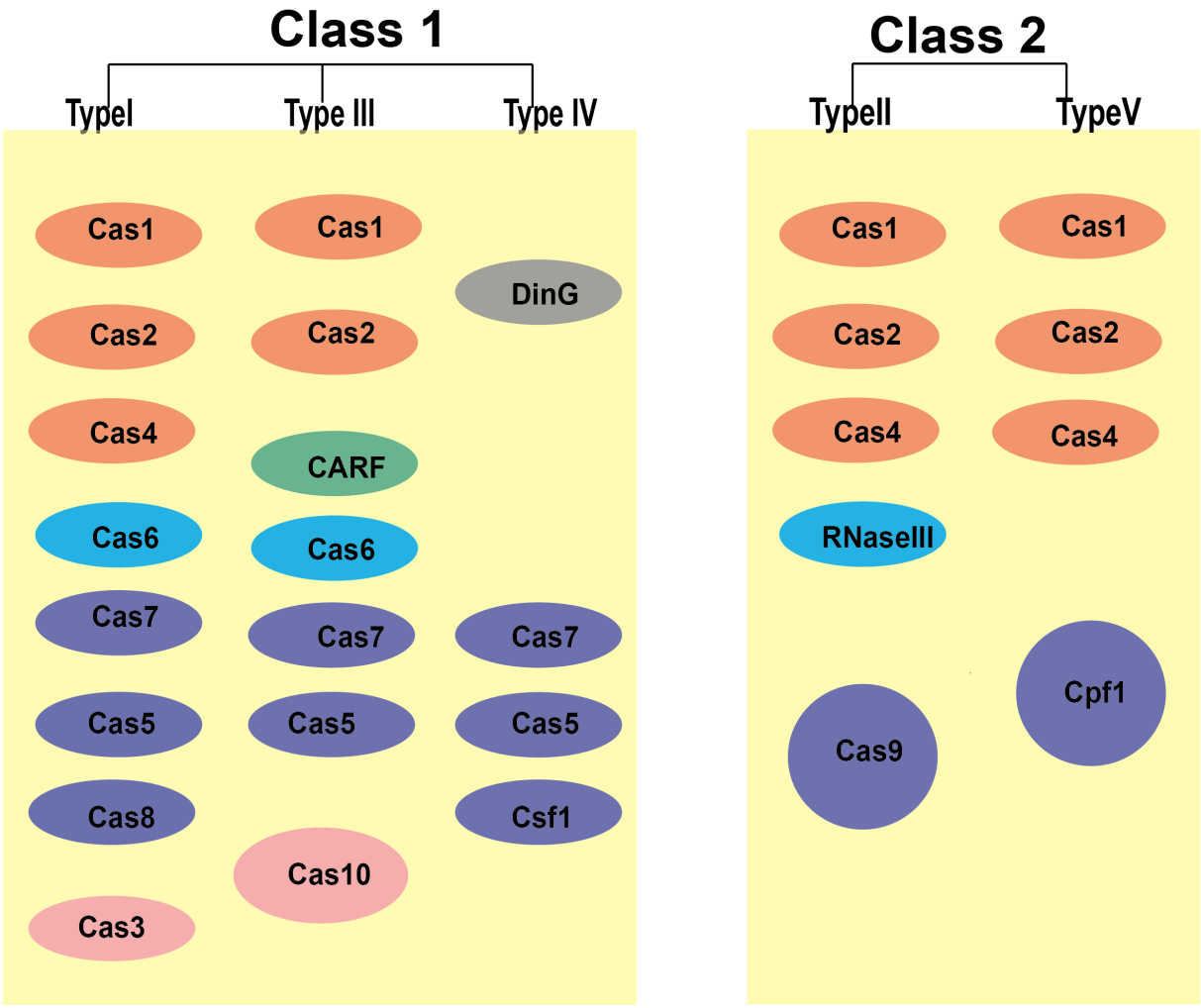


Figure 2: The types of CRISPR system. CRISPR systems are been classified into two classes (Class1 and Class2). Class 1 divided into three types (type I, type III, type IV) base on the similarity of their crRNA effector proteins. Class 2 divided to two types (type II, type V) base on the similarity of their crRNA effector proteins. The orange ovals represent proteins involved in the stage of spacer integration, blue ovals represent proteins involved in crRNA transcription stage; purple color represents the proteins in CRISPR recognition and targeting stage. Cas3 and Cas10 (pink oval) are known as nuclease proteins in type I and type III CRISPR system. CARF (dark green oval) is a regulation protein in Type III system. DinG (grey oval) is a functional unknown protein in Type IV system.

In 2011 to 2012, three groups discovered that type II system is transferrable[9] [101] [102], the purified Cas9 protein from *Streptococcus thermophilus* or *Streptococcus pyogenes* can cleave target DNA in vitro when binding with crRNA[9] [102]. For the convince of experiments, researchers fused the crRNA with another short RNA fragment -tracrRNA to facilitate the Cas9 protein binding and target DNA cleavage[9]. After that, some labs showed their results to edit genome in mammalian cells in vivo by bacterial-derived CRISPR system[103] [104].Heterologous crRNA-tracrRNA hybrids or gRNA will bring Cas9 protein to the target position and cleave double strands of the target DNA, with the help of non-homologous end joining or homologous recombination to modify mammalian cell genome. From that on, this handleable technology

becomes more and more popular for genome editing. Currently, most of the CRISPR based genome modification employed the CRISPR-Cas9 system from *Streptococcus pyogenes* (sp Cas9)[105 92 104 106]. From these studies, a lot of information about this spCas9 is accumulated including the exact PAM sequence, protospacer length and the other associated factors. It is found that, the PAM sequence for this spCas9 is a short 3 base pair sequences 5'-NGG-3', the fragment with 20nt in front of this PAM sequence is homologous to the target DNA, it will guide the Cas9 protein to the target position. It is found this protospacer and PAM sequence can tolerate some base pairs mismatch. On PAM sequence, a 5'-NAG-3' mutation also has the CRISPR recognition activity, on protospacer fragment, the nearest 8–12 nt with PAM sequence is suggested as seed sequence, this part should be unique, without any mismatch with the target position, the other base pairs are less sensitive to CRISPR motif target effect[104 107]. This feature makes a problem arose, that is the potential off-target, which means with some mismatch on crRNA, it can remain temporary bind with the DNA strands, it is observed that in this case, Cas9 has many binding sites but can only cleave a small part of them[108].

The nuclease activity of Cas9 protein cut the DNA double strands by two of its functional domains, RuvC and HNH, each of them cleave one strand of DNA (Fig 3). RuvC and HNH together will generate a blunt -ended double strand break[9]. SpCas9 is also modified to a single strand cleave version, which has one functional domain inactivated. This system will cut one strand of the target DNA[102 9], in mammalian cell, this design will repair the cleavage position via the high-fidelity base excision repair (BER) way[109] like homologous recombination. Another version of CRISPR-spCas9 is the mutants with both RuvC and HNH domains inactived, which is called dCas9 system or CRISPR interference system (CRISPRi)[10], this protein-RNA complex would keep the binding ability, but when crRNA brings the dCas9 to the binding position, it will efficiently prevent transcription of the target gene (Fig 3)[93 10].

Because of the handleability and flexibility of the system, Crispr-dCas9 system is used in many fields, like cell therapy[110 111 112]. Drug development, especially antimicrobial and antiviral drugs development[113 114]. Agriculture[115], and industrial biotechnology[116 117].

Here, we want to use a novel method for essential gene study, an RNA-base genome modification system, CRISPR-dCas9 system. As described before, CRISPR-dCas9 is a deficient version of the CRISPR-Cas9 system. This system is possible to recruit *E. coli* host cell native RNaseIII and utilize a deficient Cas9 together with specific binding gRNA and

sterically block the transcription of the target genes (Fig 3).

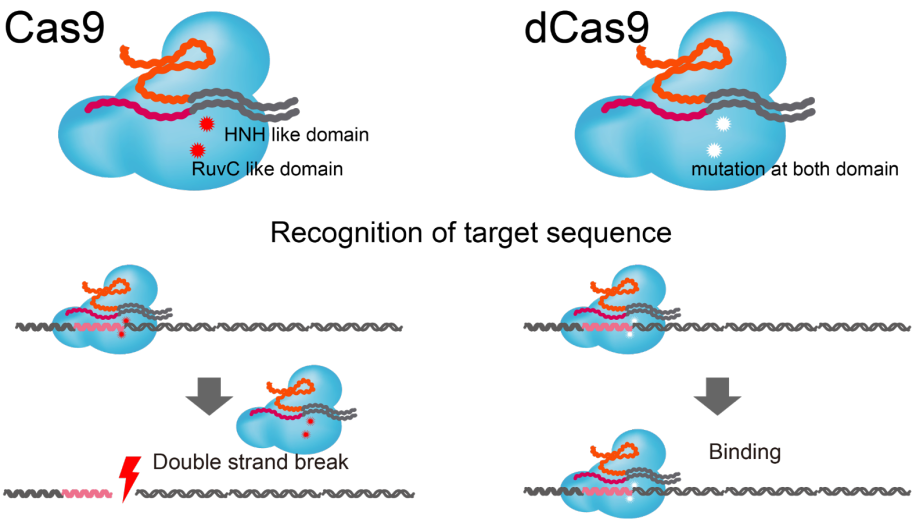


Figure 3: CRISPR-Cas9 genome editing motif. The minimal CRISPR system consists of a single protein (Cas9 protein), short tracrRNA fragment and a designed crRNA. The Cas9 protein consists of two functional domains (red asterisks), each of them response to one DNA stand cleavage. The wild-type Cas9 protein contains the nuclease activity. When the protein binds to the crRNA and forms a protein-RNA complex. The complex is guided to a specific DNA position by base pairing between the crRNA and the DNA target. In the case of wild-type Cas9, the DNA will be cleaved due to the nuclease activity of the Cas9 protein. However, the dCas9 protein is defective in nuclease activity. it is still able to form a complex with the crRNA and bind to specific DNA target. When the targeting occurs on the protein-coding region, it could block RNA polymerase and transcript elongation.

CRISPR-dCas9 is also been called CRISPRi (CRISPR interference) system. It mainly needs 3 elements to repress a target gene in the specific location in *E. coli*: tracrRNA (transcription activated RNA), gRNA (or crRNA) and pdCas9 endonuclease protein.

In this project, I want to focus on experimental essential genes in *E. coli* genome. It means the genes, which are indispensability in LB medium at 37°C. In *E. coli,* essential genes are usually in essential cellular pathways, like genome replication, cell division or cellular components synthesis (Fig 4), since the critical roles of them in the cell, it is impossible to study them by deletion mutations.

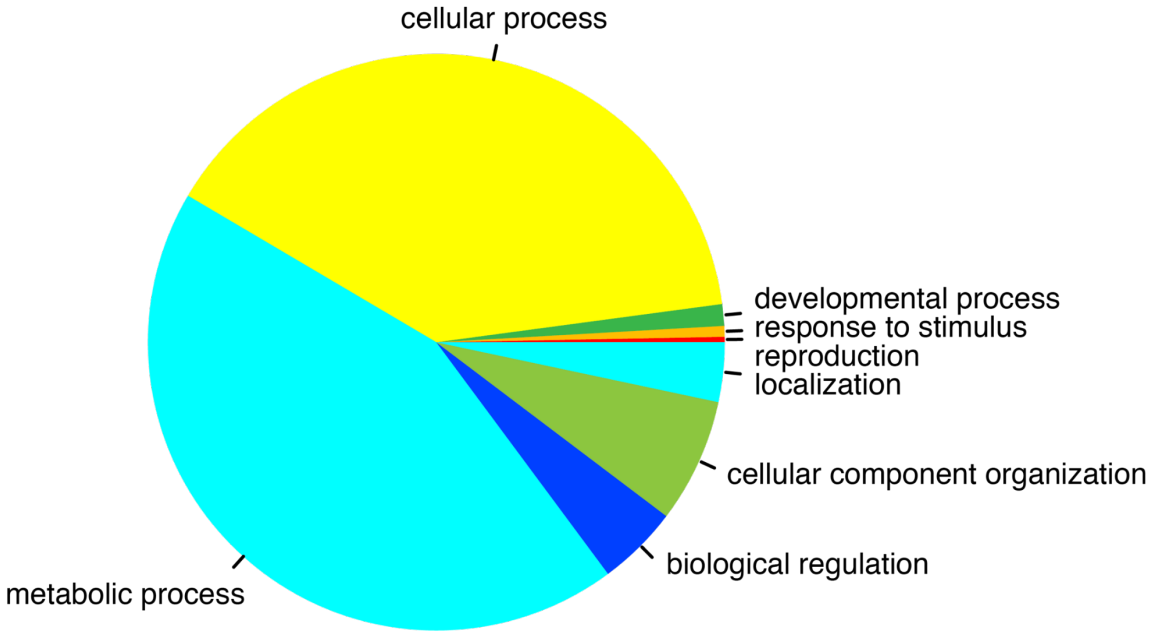


Figure 4: Essential gene classification. 325 ORF genes in *E. coli* genome is defined as essential genes by deletion mutation strategy in LB medium at 37°C. They are classed in several essential cell biological processes, like cellular process procedure, or metabolic processes or cellular component organization.

Previous studies on essential genes in our group, is to use a deletion and complementary system[118], which utilize a low copy complementary plasmid pFE604T to supply essential gene products in the essential gene deletion system. This plasmids allows people to clone essential genes ORF on the vector, and co-working with single non-essential gene deletion library Keio collection[5] to mimic a double gene knockout mutant. This strategy provides people a good way to analyze essential genes' functions and their connections with other genes in *E. coli* community, this was the first try to knockdown the essential genes in *E. coli* genome. However, there are some limits or hurdles in the mutation and complementary system, some of the essential genes show very critical to *E. coli* fitness, when they are removed, the cell cannot survive even under the supplements of these essential gene products[118]. In addition, technically, in this system, the deletion of the essential gene base on the previous prediction of essential gene ORF, the entire fragment absence may bring some affection to target gene communication with the other factors.

The development of new genome editing tool CRISPR-dCas9 system gives us a new strategy to knockdown the essential genes in *E. coli* genome, this system contains 2 plasmids,

one express pdCas9 protein, and the other one express RNA factors including tracrRNA and gRNA. The concept is utilizing the well-known spCRISPR-dCas9 system (CIRPSR -dCas9 derived from *streptococcus pyogenes*) to repress the essential genes in *E. coli*. Since the essential genes have varied functions and different essentiality levels, the gene inhibition may need different degrees, so we employed the pdCas9 plasmid with an inducible promoter to control the pdCas9 protein expression. On RNA expression plasmid, we employed a high copy plasmid vector with a constitutive promoter, and 20nt spacer sequence connects with gRNA scaffold here function as tracrRNA and repeats[10]. Theoretically, this system will repress the target essential genes of *E. coli* without killing the bacteria. Besides, combine this gene knockdown system with the single gene deletion library Keio collection, recruit high-throughput experimental method, it is possible to detect the communication of target essential genes with the entire genome.

Another technique necessary for this project is bacterial conjugation. Bacterial conjugation is a widespread DNA horizontal transport mechanism, the conjugative DNA is transferred from one cell to another by a simple cell to cell contact or bridge connected contact[119]. This feature probably one of the most common reasons why the bacteria gains new abilities, like antibiotic resistance, xenobiotic tolerance, new metabolites synthesis abilities and so on. This mechanism is first revealed more than half century ago by Joshua Lederberg and Edward Tatum[119], they found the bacterial cells have "F" factor can transfer the DNA to the other cells. This F plasmid is a single copy plasmid in a given bacterium, it contains some important functional segments, like *oriT* (transfer origin), *oriC* (replication origin), *tra* region (transfer genes) and IS (Insertion Elements). Bacterial conjugation starts in the cytoplasm, when the DNA relaxase proteins recognize *oriT* fragment, it cut the single strand in *oriT* region, and bind on the 5 ends of the nicked strand, the nicked strand (T-strand), is unwound from the double strands and transferred to recipient cell in 5' to 3' direction. The transferred DNA can then be integrated into the recipient cells' genome by homologous recombination or cyclized to generate an isolated circular plasmid. Conjugation is a convenient way to transfer genes from cell to cell, it has been reported that it could happen from bacteria to many other organisms, including yeast[120], plants , mammalian cells , diatoms and isolated mammalian mitochondria[121] [122] [123]. This flexible activity makes it possible to recruit this system as DNA transformation tool. One of the application is Hfr strain, Hfr strain is the abbreviation for high-frequency recombination strain, it is the bacterial cells with conjugative F plasmid integrated into the chromosomes, Hfr cells can efficiently promote the transformation of its DNA to a recipient cell, which does not have F

factor.

Based on these complete resource and high-through method, I started this project to analyze the features of essential genes in *E. coli*, and the genetic interactions between the target essential genes and non-essential genes of the entire *E. coli* K-12 genome. First, I need 2 plasmids express all the elements of CRISPR-dCas9 system: dCas9 protein, crRNA and tracrRNA. Some reports suggest that high amount dCa9 accumulation in the cell may exhibit toxicity[124], to decrease the vector influence to *E. coli* growth, it is better to put this dcas9 express gene under an inducible promoter for the manipulation of dCas9 protein expression level. Another problem is that, to extend the repression to the entire genome scale, I need a F plasmid derived *oriT* fragment inserted on the plasmids. In this scenario, if these conjugative plasmids are transferred into Hfr strain, they can utilize the *tra* operon gene products and transfer themselves by a simple direct cell-to-cell contact (Fig 5).

Gibson assembly seamless cloning method is benefits to maintain the structure of crRNA and tracrRNA products. The essential gene repression strains' phenotypes in liquid could be evaluated by measure the OD600 value. And on solid plates, the phenotypes could be detected by Colony-live system[66], this system contains several scanners and the software developed by this group, it captures the images of the square plates every 30 minutes, and calculates LTG (lag time of growth), MGR (maximum growth rate), and SPG (saturation point of growth) at the end of the culture period.

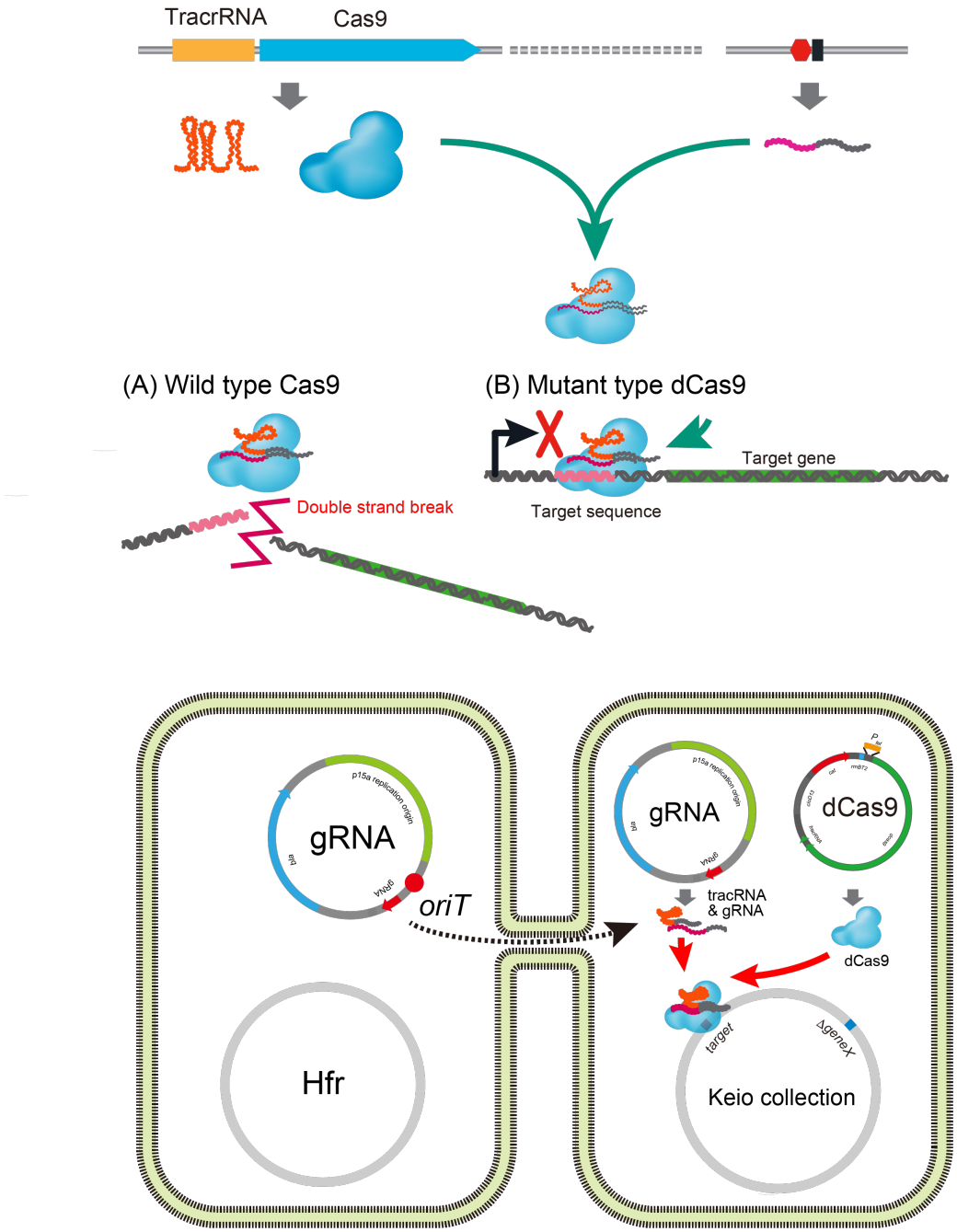


Figure 5: CRISPR-dCas9 essential gene knockdown and non-essential gene knockout system construction. 3 elements are essential for CRISPR-dCas9 system: dCas9 protein, tracrRNA and crRNA, in WT (A), the Cas9 protein, tracrRNA and crRNA complex specific recognizes the target position, and cleaves the DNA double strands at the binding locus. While, in deficiency version, the dCas9 system (B), is still keeping the target position recognition ability, but dCas9 lost the cleavage function, which means the protein RNA complex will block the RNA polymerase to stop the transcription at the target loci instead of break the DNA strands. Essential gene knockdown and non-essential gene knockout strains construction takes the advantage of bacterial conjugation system, employ a Hfr strain with conjugative CRISPR system as the donor, and contact donor with single gene deletion library Keio collection to transfer the CRISPR-dCas9 system to Keio collection, promote the CRISPR factors expression in Keio collection to gain essential gene knockdown mutants.

Chapter 2 Materials and methods

2.1 Strains and plasmids

Strains and plasmids used in this research are listed in the Table. 1.

Strains

Strains	Genotype	Reference
BW25113	*lacIq rrnBT14* D*lacZWJ16 hsdR514* D*araBADAH33* D*rhaBADLD78*	Datsenko KA, Wanner BL [125]
Mach 1	D*recA1398 endA1 fhuA Φ80*D *(lac)M15* D *(lac)X74 hsdR*($r_{K^-}m_{K^+}$)	Invitrogen
G286	*DH10B:* D*metC* D*dadX* D*alr hfr3(*D*oriT)*	Goto, S. personal communication
Keio collection	*Deletion strain collection using BW25113*	T.Baba[126]

Plasmids

Name	Features	Reference
PdCas9	PLtetO-1 Cas9 *cat p15A*	L. Qi[10], (Addgene Cat: 44249)
pgRNA-vector	D-*lac oriColE1*	L. Qi[10], (Addgene Cat: 44251)

Mach 1 (Invitrogen) is commercially available fast-growing *E. coli* cell and quite easy to use as a host strain of DNA cloning. The Mach1 colonies could be clearly detected after 8 hours, and in liquid culture, it can reach the stationary phase after 4 hours culture in LB at 37 °C.

G286 (Goto, S. personal communication) is the Hfr strain by CIP plasmids[45] with some modification. This strain was used as a donor strain of conjugative transfer of plasmids carrying *oriT* plasmids with *dal* phenotype (alanine auxotroph) by deleting *metC*, *dadX* and *alr* genes. CIP8 plasmid was integrated between *ydiO* and *ydiP* (1777kb coordinate of *E. coli* K-12 genome) to make this host strain Hfr, whose *oriT* was inactivated by deletion.

PdCas9-bacteria and pgRNA-bacteria original plasmids were purchased from Addgene (Cat:44249, 44251), the plasmids are first developed by Stanley Qi in UCSF[10], pdCas9-bacteria plasmid contains an aTc inducible promoter PLtetO-1, a chloramphenicol selectable marker, and a p15A replication origin besides dCas9 express gene. PgRNA-bacteria plasmid contains a 20nt gRNA clone space, and a minimal synthetic promoter (J23119) with an annotated

transcription start site, an ampicillin-selectable marker and a ColE1 replication origin. *oriT* fragment came from wild type F plasmid kindly provided by Dr. Sampei (personal communication).

PdCas9-*oriT*: the pdCas9- bacteria plasmid with *oriT* fragment which allows the plasmid self-transmissible. PgRNA-*oriT*: the pgRNA- bacteria plasmid with *oriT* fragment which allows the plasmid self-transmissible.

The pdCas9 plasmid has a PL-tetO promoter-operator system, it is inducible by tetracycline or more efficiently by anhydrotetracycline (aTc)[127], and has lower antibiotic activity toward *E. coli*.

DL-alanine (Sigma, Co. 302727): 5mM DL-alanine is add in the medium to support the growth of G286 strain.

Antibiotics: in this project, 100 μg/ml ampicillin was used for pgRNA plasmids selection, 30 μg/ml chloramphenicol was used for pdCas9 plasmids selection. And 30 μg/ml kanamycin was used for Keio collection selection.

2.2 Primers

Primers were synthesized by Invitrogen (XYP0096 -XYP0109) or Hokkaido System Science (XYP0110-XYP0155), without any modification. Primer YHP-1 andYHP-2 came from essential gene conditional deletion library construction project[118]. Sequences of primers are listed in the Table 2.

Table 2 Primers used		
names: primer name		
sequence: sequence		
function: purpose we made		
names	**sequences**	**funcitons**
XYP0096	ATCACGAGGCAGAATTTCAG	forward checking primers for pgRNA-bacteria(1)
XYP0097	TACGGCGTTTCACTTCTGAG	reverse checking primers for pgRNA-bacteria(1)
XYP0098	GCGGCATAAGCCAGAAAATATC	froward checking primers for pdCas9-bacteria(1)
XYP0099	TTGGCGTTTAGAAGTTGTCTCC	reverse checking primers for pdCas9-bacteria(1)
XYP0100	AAGGAAAAAAGCGGCCGCTTCCTCGCTCACTGACTCG	forward primer of oriT cloning on pgRNA-bacteria with NotI
XYP0101	TTTTCCTTTTGCGGCCGCGCCTGCAGTCTAGACTCGAG	reverse primer of oriT cloning on pgRNA-bacteria with NotI
XYP0102	AAGGAAAAAAGCGGCCGCGCTTACGAACGGGGCGGAGA	forwar primer of oriT cloning on pdCas9-bacteria with NotI
XYP0103	TTTTCCTTTTGCGGCCGCATTTCCGCTCGCCGCAGTC	reverse primer of oriT cloning on pdCas9-bacteria with NotI
XYP0104	AAAGGCTCAACAGGTTGGTG	forward checking primer of oriT
XYP0105	CACCCCTACAAAACGGTGTC	reverse checking primer of oriT
XYP0106	TAAAACGAAAGGCTCAGTCG	forward checking primer of oriT on pdCas9
XYP0107	AGGCGTGGAATGAGACAAAC	reverse checking primer of oriT on pdCas9
XYP0108	CTTGAGCGTCGATTTTTGTG	reverse checking primer of oriT on pgRNA
XYP0109	TTTTCCTTTTGCGGCCGCCACTAAAATTCATAATGCAAACAGGG	reverse primer of oriT cloning with NotI
YHP-1	AAGGAAAAAAGCGGCCGCCACCTCTGGTGACTTTATC	37nt forward primer of oriT cloning with NotI
XYP0110	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCTCACGCAGCTCTTTTACCA	tsf gRNA forward primer
XYP0111	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTGGTAAAAGAGCTGCGTGAG	tsf gRNA reverse primer
XYP0112	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCGTCAGGGTGACTTTCTTGC	lpxC gRNA forward primer
XYP0113	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGCAAGAAAGTCACCCTGACG	lpxC gRNA reverse primer
XYP0114	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTATCGCCAGGGTCTGTTTCTC	infA gRNA forward primer
XYP0115	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAGAAACAGACCCTGGCGAT	thiL gRNA reverse primer
XYP0116	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCTGGATCAGCGTTTGCTCAC	secD gRNA forward primer
XYP0117	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGTGAGCAAACGCTGATCCAG	secD gRNA reverse primer
XYP0118	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCATCAGTGGTACGGCTATTT	mrdB gRNA forward primer
XYP0119	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAAATAGCCGTACCACTGATG	mrdB gRNA reverse primer
XYP0120	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCGAAGGGTTAAAGCTTAGTA	ftsK gRNA forward primer
XYP0121	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTACTAAGCTTTAACCCTTCG	ftsK gRNA reverse primer
XYP0122	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCATTTCAATATTGTCTTCTT	infA gRNA forward primer
XYP0123	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAAGAAGACAATATTGAAATG	infA gRNA reverse primer
XYP0124	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCAATAAATAGAGCGACAGGT	lolC gRNA forward primer
XYP0125	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACACCTGTCGCTCTATTTATTG	lolC gRNA reverse primer
XYP0126	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCCGGTAGATGGAAATATTTC	yceQ gRNA forward primer
XYP0127	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGAAATATTTCCATCTACCGG	yceQ gRNA reverse primer
XYP0128	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCTGGCGCTCGAGTCGAACCT	tpr gRNA forward primer
XYP0129	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACAGGTTCGACTCGAGCGCCAG	tpr gRNA reverse primer
XYP0130	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTTCAGGTTGAGCAGGATTTCC	rpoA gRNA forward primer
XYP0131	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGGAAATCCTGCTCAACCTGA	rpoA gRNA reverse primer
XYP0132	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTGTTTGCTTCAACTTCGATTA	rplC gRNA forward primer
XYP0133	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTAATCGAAGTTGAAGCAAAC	rplC gRNA reverse primer
XYP0134	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTTTTACGCAGCGCTTGTGCAT	ftsZ gRNA forward primer
XYP0135	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACATGCACAAGCGCTGCGTAAA	ftsZ gRNA reverse primer
XYP0136	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTGCGTCGTGACCAGCTCGTCA	mukB gRNA forward primer
XYP0137	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACTGACGAGCTGGTCACGACGC	mukB gRNA reverse primer
XYP0138	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTTTCGCGACTTGAGTCATAGT	metG gRNA forward primer
XYP0139	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACACTATGACTCAAGTCGCGAA	metG gRNA reverse primer
XYP0140	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTTGGTCGCGCCTAATTGTTCG	murA gRNA forward primer
XYP0141	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCGAACAATTAGGCGCGACCA	murA gRNA reverse primer
XYP0142	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCCCGGTTTTCAAGTTCTAAT	mreC gRNA forward primer
XYP0143	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACATTAGAACTTGAAAACCGGG	mreC gRNA reverse primer
XYP0144	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTGCGATAACGCCGTCTTTCAT	mreB gRNA forward primer
XYP0145	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACATGAAAGACGGCGTTATCGC	mreB gRNA reverse primer
XYP0146	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTGTCGAGCATCTGCACCGTGC	yibJ gRNA forward primer
XYP0147	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGCACGGTGCAGATGCTCGAC	yibJ gRNA reverse primer
XYP0148	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTGACCATGAACGGATCATGTC	rpoC gRNA forward primer
XYP0149	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACGACATGATCCGTTCATGGTC	rpoC gRNA reverse primer
XYP0150	GAATTCTAAAGATCTTTGACAGCTAGCTCAGTCCTAGGTATAATACTAGTCACGTTCGAGAGCGGCGGAG	dnaC gRNA forward primer
XYP0151	GTTGATAACGGACTAGCCTTATTTTAACTTGCTATTTCTAGCTCTAAAACCTCCGCCGCTCTCGAACGTG	dnaC gRNA reverse primer
XYP0152	GTCCTAGGTATAATACTAGTGACCATGAACGGATCATGTC	rpoC gRNA forward primer (20nt)
XYP0153	GCTATTTCTAGCTCTAAAACGACATGATCCGTTCATGGTC	rpoC gRNA reverse primer (20nt)
XYP0154	GCTCAGTCCTAGGTATAATACTAGTGACCATGAACGGATCATGTC	rpoC gRNA forward primer (25nt)
XYP0155	AACTTGCTATTTCTAGCTCTAAAACGACATGATCCGTTCATGGTC	rpoC gRNA reverse primer (25nt)
XYP0158	GCGGAACTGAGCGATAACA	dnaA qrt-PCR forward primer
XYP0159	CCGCAGAAACTGGTTAGCA	dnaA qrt-PCR reverse primer
XYP0160	AGCACCTCACGGAAAAAGCA	ftsN qrt-PCR forward primer
XYP0161	CCGGTCACTTTCTGGCTTTG	ftsN qrt-PCR reverse primer
XYP0162	ACTCCTACGGGAGGCAGCAGT	16s rRNA forward q-rtPCR
XYP0163	CCTTCCTCCCCGCTGAAAG	16s rRNA reverse q-rtPCR
	(1) Qi LS. Et al. Cell 2013, 152:1173-1183	

2.3 CRISPR-dCas9 target gRNA design

crRNAs for the entire essential gene set were designed by python. The principle is 1). Search the PAM sequences (NGG) on both strands of *E. coli* genome. 2) Output the entire pool of the crRNAs. 3) Assign the crRNAs to each gene by the positions of the genes on the genome, 4) At last, decide the target DNA strand (template strand or non-template strand) by the orientation direction of the genes. The crRNAs for all the essential genes in *E. coli* is shown in supplementary Table 1.

2.4 Molecular biology experiments

Plasmids extraction follow the standard alkaline method described in Molecular Cloning[128]. For mini-preparation, 5 ml culture was used. For medium scale preparation, 50ml culture was used. All the clone steps were employed the high fidelity hot start DNA polymerase KOD-plus (Toyobo Japan), the DNA fragments are amplified the in 50μl reaction system followed by DpnI digestion, alkaline phosphatase treatment and gel purification.

PCR amplified fragments were normally purified by DpnI (Takara) to destroy template DNA for PCR and terminal dephosphorylation by BAP (Takara) to eliminate self-ligation. *Not*I digestion was performed for overnight at 37°C and inactivated at 80°C for 20 for *oriT* insertion. After enzymatic treatment, size separation of fragments was done by 1 to 2 % of low melting agarose (Lonza Co. Ltd, Japan) according to the fragment sizes.

The purified DNA fragments were sent to ligation (*oriT* insertion) or Gibson assembly (gRNA clone). For *oriT* insertion, DNA ligation Mighty Mix kit (Takara, Japan) was used. The ligation reaction followed the manufacturer's instruction. The gRNA plasmids were constructed by seamless cloning strategy Gibson assembly according to the published protocol[129]. The homemade Gibson assembly buffer was prepared by three enzymes: T5 exonuclease (New England Biolabs, USA), Phusion DNA polymerase (New England Biolabs, USA) and Taq DNA ligase (New England Biolabs, USA). In a reaction mixture, two or more double strands linear fragments with homologous regions are required. T5 exonuclease will digest linear DNA strand from 5' end to 3' end to create single DNA strand with 3' end overhangs, single stranded terminal regions anneal together and then Phusion DNA polymerase fills gaps in the annealed DNA fragments. Finally, Taq ligase seals the nicks between inserts and backbone plasmid. The DNA fragments were amplified by KOD plus DNA polymerase in

50μl reaction system, purified by low melting temperature gel[128].

Structural confirmation of cloned plasmids was performed by colony PCR with Ex-Taq (Takara, Japan) according to the instruction manual. The final plasmids 'sequences were confirmed by Big-Dye terminator sequencing (ABI 3100 Genetic analyzer).

The engineered DNA product are transferred into Mach 1 electro-competent cells by electroporation (BIO-RAD). Preparation of electro-competent cells and electroporation were performed according to Molecular Cloning[128]and manufacture's instruction (EcoliPulsor: BioRad), respectively.

Total RNA was extracted according to the conventional hot phenol method[130]. To avoid a shortage of samples of mRNA, I generally used 50 ml log phase culture with appropriate concentration of aTc (Wako, Japan) as inducer of dCas9 to activate the repression with constitutively expressed specific gRNA. Extracted RNAs were checked by neutral agarose gel electrophoresis to check the ration between 23S and 16S rRNA and their degradation.

500ng total RNAs are used to synthesis cDNA by PrimeScript reverse transcriptase (Takara, Janpan), with random hexamer (Takara, Japan) in 20μl reaction system following the user manual.

The SYBR green based real time PCR method is used to detect the relative quantification of the target genes on the basis of 16s rRNA, the wild type strain is chosen as control, the extracts concentration and quality were checked by Nano drop and non-denaturing agarose gel, finally dissolve in 25μl DEPC water.

The synthesized cDNAs were sent to qRT-PCR, and quantitatively detect the relative mRNA amount base on the standard curves from WT cDNA with different dilution times. In each target culture, 16s rRNA is employed as the reference gene, and final target gene mRNA relative expression is calculated by: Concentration of the sample/Concentration of reference gene. qRT-PCR employed SYBR® Premix Ex Taq TM (Takara, Japan) and lighCycler® 480 (Roche life science, the USA), the data statistical analysis is done by R, bar charts were generated by graphed prism.

2.5 Bacterial experiments

Conjugation was done using mid-log growing phase culture of both donor and recipient strains, 200μl mid-log phase culture of donor and recipient strains were mixed and incubated at 37°C for one hour without shaking. Plating on LB agar plates with proper antibiotics and calculate the conjugation efficiency by the function: conjugation efficiency = conjugated cells /recipient cells.

To monitor the growth in liquid culture, I measured OD600 by spectrophotometer (Jasco V-630) or micro-plate reader (Molecular Devise SpectraMax® GEMINI EM) according to sample numbers. Overnight cultures were inoculated into 96-well microtiter plates containing 100 μl of liquid medium with appropriate antibiotics supplemented with 0, 0.1, 0.2, 0.5, 1.0, 2.0 μM of aTc. Cultures were incubated at 37°C for 24 hours to measure the optical density (600 nm) every 30-min using SpectraMax® GEMINI EM (Molecular Devices Inc).

To monitor the colony growth on agar plate, I used Colony-Live scanning systems[66] with appropriate antibiotics supplemented with 0, 0.1, 0.2, 0.5, 1.0, 2.0 μM of aTc.

To check the morphology of the essential gene inhibited strains, the all-in-one fluorescence microscope (KEYENCE, BZ-9000E) was employed. For all the targeted essential genes, I checked the bacterial phenotypes at log phase with 1 μl aTc or without aTc. And for *dnaA* and *ftsN*, time series observation is detected in the medium with 1 μM aTc or without aTc.

2.6 High throughput experiment

Genetic Interaction between essential gene knockdown and non-essential gene knockout were detected by the high throughput method.

Conjugation to transfer CRISPR-dCas9 system into Keio collection by singer robot followed this protocol, also seen in Fig6:

1). pre-culture recipient strain array, here the recipient strains are single non-essential gene deletion Keio collection. Thaw the glycerol stock and stamp on LB agar plates with 30 μg/L kanamycin. Incubate at 37°C overnight to generate 384 colony arrays.

2). From 384 array stamps to LB plates with 30 μg/L kanamycin, to make 1536 colony array,

incubate at 37°C overnight.

3). Collect 200 μl log phase liquid donor strain, G286 (OD600=0.5) with pdCas9-*oriT* plasmid, spread on LB plates with DL-alanine. Stamp Keio collection plasmids array onto donor lawn. Mating at 37°C for 2-3hours, thereafter stamp the colony to LB plates with 30 μg/L kanamycin, 30 μg/L chloramphenicol, culture at 37°C overnight.

4). Pick the colonies culture in microplates, liquid LB at 37°C overnight to make glycerol stock of Keio collection strains with pdCas9-*oriT* plasmid. Store at -80°C.

5). Repeat step 1) to 4) but use G286 strain with pgRNA- *oriT* plasmids as the donor, and Keio collection strains with pdCas9-*oriT* plasmid as the recipient strains to transfer the essential gene repression gRNA plasmids into the Keio collection strains.

4). Stamp 1536 Keio collection with dCas9 expression plasmid and gRNA expression plasmids array on fresh LB agar with 30 μg/L kanamycin, 30 μg/L chloramphenicol, 100 μg/L ampicillin plates, and supply 0 μM, 0.1 μM, 0.2 μM, 0.5 μM, 1 μM, 2 μM aTc. Culture in Colony-live scanner system at 37 °C for 24 hours.

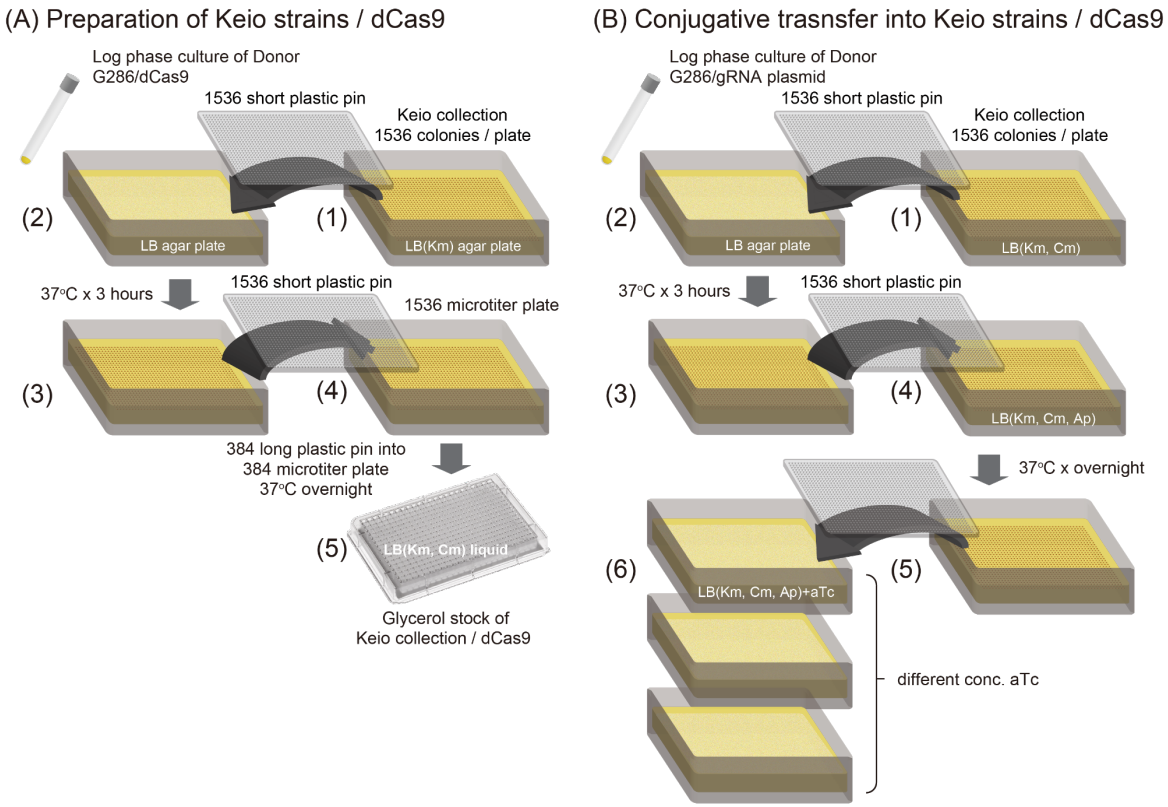


Figure 6: High throughput scheme for essential gene knockdown and non-essential gene knockout array construction. CRISPR-dCas9 elements are expressed from two plasmids, the high throughput strategy involves two stages, one is to transfer dCas9 expression plasmid to Keio collection (A); and the other stage is to transfer specific gene target crRNA to Keio collection with dCas9 (B). (A): (1) the experiment was done using a well confirmed non-essential gene deletion library Keio collection, stamp the Keio collection on LB

plate with 30 μg/ml kanamycin, culture at 37°C overnight to generate Keio collection colony array. (2) collect log phase donor strain (G286 with dCas9 expression plasmid) pure on LB plate to generate donor lawn, and stamp Keio array to donor lawn. (3) 37°C, 3 hours incubation to transfer the dCas9 expression plasmid-pdCas9 to Keio collection, (4) and stamp to LB plate with antibiotics to select the Keio mutants which gain pdCas9 plasmid. (5) transfer the Keio collection with pdCas9 to liquid LB culture with antibiotics (30 μg/ml kanamycin, 30 μg/ml chloramphenicol) in 384 well microplates, culture overnight. Add in 15% glycerol to make Keio + pdCas9 array glycerol stock. (B): (1) stamp Keio+pdCas9 array on LB plate with antibiotics (30 μg/ml kanamycin, 30 μg/ml chloramphenicol) to generate 1536 array. (2) collect log phase donor strain (G286 with crRNA expression plasmid) pure on LB plate to generate donor lawn, and stamp Keio +pdCas9 array to donor lawn. (3) 37°C, 3 hours incubation to transfer pgRNA plasmid to Keio + pdCas9 collection, (4) and stamp to LB plate with antibiotics (30 μg/ml kanamycin, 30 ug/ml chloramphenicol, and 100 μg/ml Ampicillin) to select the Keio mutants which contain both CRISPR plasmids. (5) overnight incubation to get the mutants, finally (6) stamp to the plates with or without plasmids inducer and monitor the colony growth by Colony-live system[46].

2.7 Bioinformatics

All the statistic work in this project is done by R. Growth curves of single essential genes knockdown are made by GraphPad Prism. Essential gene knockdown and non-essential gene knockout data are read and plot by R, primers are designed by primer 3 (online website). The potential target positions on *E. coli* genome is selected by python. Gene classification is made according to Panther database.

Chapter 3 Results

3.1 Knockdown efficiency of the original CRISPR-dCas9 vectors

3.1.1 Confirmation of the structure of the original vectors.

First, I confirmed the structures of the original plasmid vectors by PCR, restriction enzyme digestion and BigDye terminator sequencing, Fig. 7. The molecular verification confirmed the molecular structures of the plasmids.

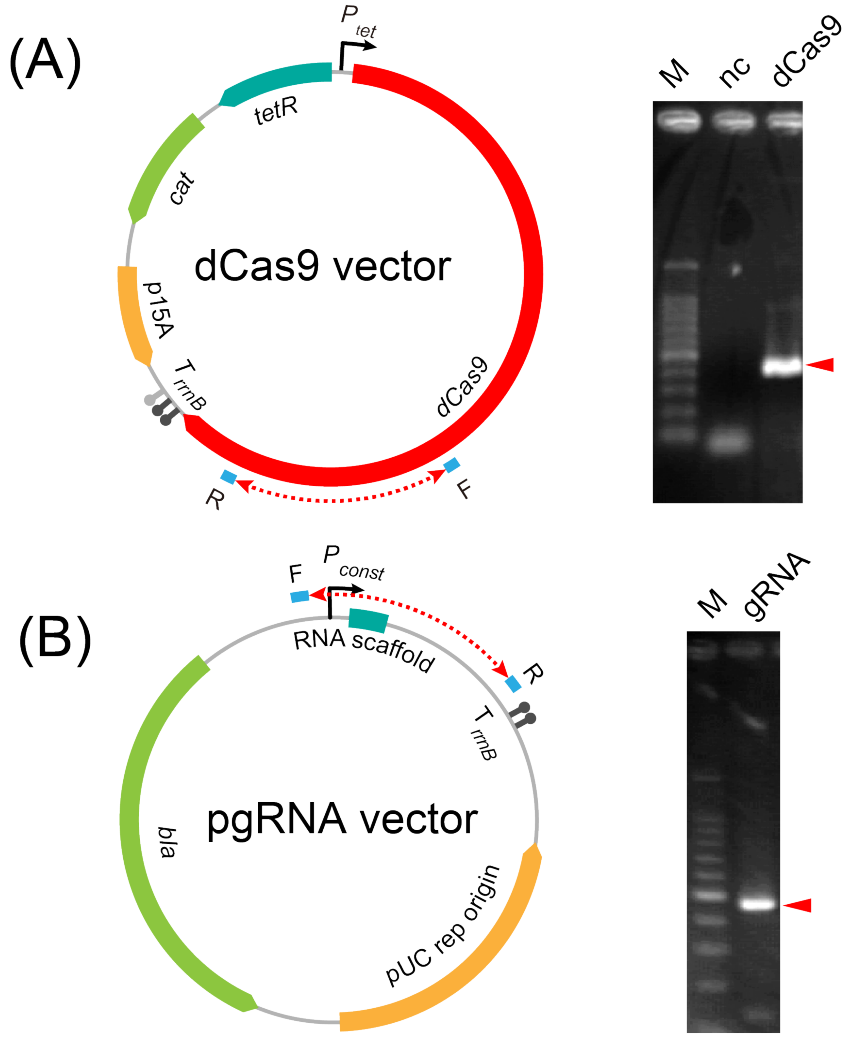


Figure 7: Plasmid structure confirmation. (A) the structure of dCas9 protein expression plasmid, pdCas9, the pdCas9 plasmid contains an aTc-inducible promoter PLtetO-1, a chloramphenicol-selectable marker (light green fragment), a p15A replication origin (yellow fragment), and dCas9 coding gene (red fragment). The blue boxes on the outside of dcas9 fragment represent primers "R" means reverse primer, and "F" means

forward primer. The dash lines in between R and F are the PCR product for plasmids structure confirmation. The 1% agarose gel images on the right side is the experimental confirmation of pdCas9 plasmid PCR products using R & F primers. (B) the structure of crRNA and tracrRNA expression plasmid, pgRNA, the pgRNA plasmid contains a minimal synthetic promoter (J23119), an ampicillin-selectable marker (green fragment), and a PUC replication origin (yellow), the RNA scaffold (dark green) is the RNA expression cassette, which contains 20nt crRNA expression region and 42 nt tracrRNA expression region. The 1% agarose gel on the right side is the experimental confirmation of pgRNA plasmid PCR products using R & F primers.

3.1.2 Growth influence of aTc to wild type *E. coli*.

aTc is the inducer of dCas9 expression plasmid, before checking the CRISPR-dCas9 inhibition, the chemical influence of aTc to wild type *E. coli* was detected. Wild type *E. coli* strain BW25113 was cultured in LB medium with different aTc concentrations at 37 °C. As shown in Fig. 8, aTc in the range between 0 μM and 2 μM in LB had not affected the cell growth, but with aTc more than 5 μM of aTc, *E. coli* growth was obvious inhibited. Based on this observation, the following experiments all employed aTc concentration less than 2 μM for dCas9 expression induction.

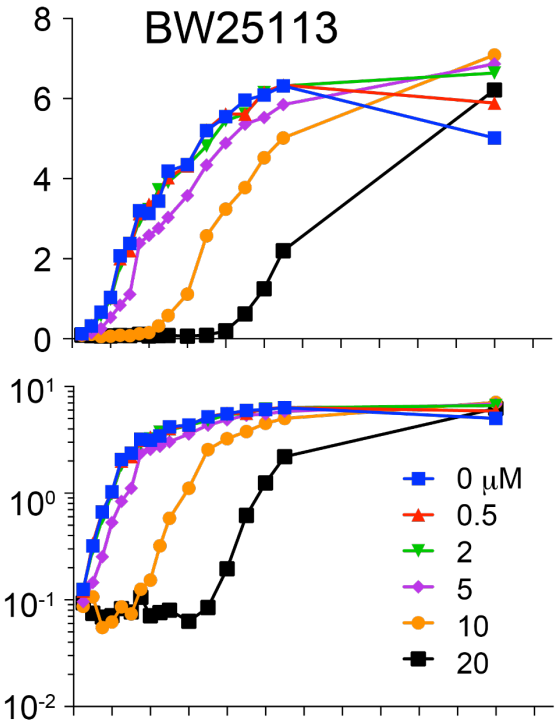


Figure 8: aTc influence to wild type growth. the 24 hours growth profile of wild type *E. coli* BW25113 strain in LB medium with 0 μM (blue), 0.5 μM (red),2 μM (green), 5 μM(purple), 10 μM (orange), or 20 μM (black) aTc. X-axis is time points (0 to 24 hours), Y-axis is linear OD600 value (top) or log OD600 value (bottom), each value came from 6 replicates. The colors indicate different aTc concentration in the medium.

3.1.3 Growth influence of dCas9 and gRNA vectors.

To measure the growth influence from pdCas9 or pgRNA vectors themselves, I performed the growth analysis of the host strain BW25113 carrying pdCas9, pgRNA, or both vector plasmids, respectively in LB at 37 °C with different concentration of aTc. As shown in Fig. 9, each of vectors or both showed no significant growth inhibition even though with inducing chemical compound.

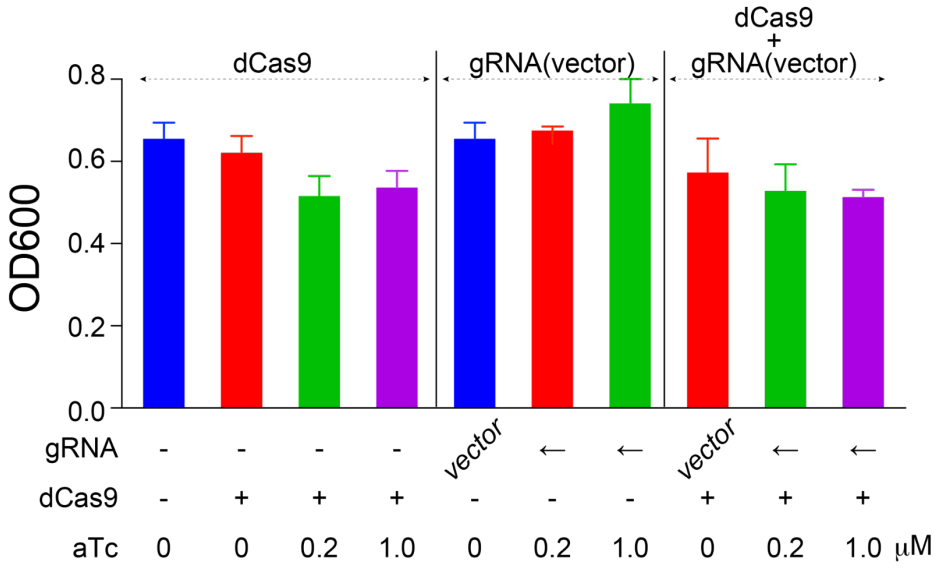


Figure 9: Vector plasmids influence to *E. coli* growth. y-axis represents OD600 value of *E. coli* strains with empty CRISPR-dCas9 vector plasmids after 6 hours culture in LB medium with antibiotics (Am, Cm), x-axis represents experimental conditions, gRNA indicates pgRNA-bacteria plasmid, dCas9 means pdCas9-bacteria plasmid, aTc line shows the concentration of aTc inducer (μM). "-" represents absent of the factor in the culture; "+" represents presence of the factor in the culture. Left arrow lines mean the same with left. "vector" means pgRNA-bacteria plasmid which has no target position in *E. coli* genome. Each value comes from 6 replicates, the bars show the mean of the test, error bars indicate SD.

3.1.4 Construction of *dnaA* and *ftsN* gRNA clones as a pilot test.

2 essential genes, *dnaA* and *ftsN* were chosen as the pilot test to optimal the experimental condition. *dnaA*, replication initiation protein and global transcription regulation[131], and *ftsN*, cell division protein[132].

The bacterial CRISPR-dCas9 system reports showed that, non-template strand and N terminus or 5' non-coding regions near 5' non-translated regions are better target sequences[10]. To test the repression efficiencies of target gene expression by CRISPR-dCas9 and gRNA combination, two primers at N-terminal region for both strand were designed and cloned into the gRNA vector by Gibson assembly. (Fig. 10 top). Repression efficiencies were measured

with different concentration of aTc in BW25113 strain, which was the host strain of Keio collection[5]. Growth were measured by OD600 in LB medium with proper antibiotics and aTc at 37°C.

OD600 values by micro-plate reader SpectraMax® GEMINI EM. Based on qRT-PCR measurements of *dnaA* and *ftsN* gene expression, difference of target DNA strand may have different efficiency. For *dnaA*, template strand showed a stronger growth inhibition. On the other hand for *ftsN*, non-template strand showed stronger repression at 37ºC in LB medium supplied with 1 μM aTc (Fig 10, bottom).

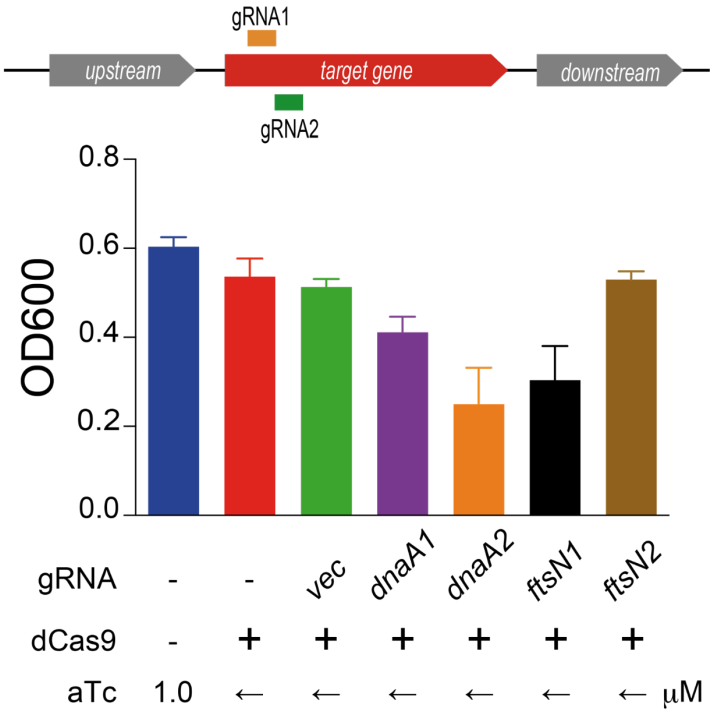


Figure 10: CRISPR-dCas9 system inhibited *E. coli* growth in liquid LB medium. top: CRISPR-dCas9 binding sites, for each target genes, design two gRNAs in the ORF region, indicated by orange and green boxes. gRNA1 (orange box) targets non-template strand, gRNA2(green box) targets template strand. bottom: Y-axis represents OD600 values; x-axis represents culture conditions, gRNA indicates pgRNA-bacteria target gene, dCas9 means pdCas9-bacteria plasmid, aTc line shows the concentration of aTc inducer (μM). "-" represents absent of the factor in the culture; "+" represents presence of the factor in the culture. Left arrow lines means the same with left. "vec" means vector system which contains an irrelevant crRNA have no target position in *E. coli* genome. "*dnaA1*" represents the gRNA target *dnaA* gene non-template strand. "*dnaA2*" represents the gRNA target *dnaA* gene template strand. "*ftsN1*" represents the gRNA target *ftsN* gene non-template strand. "*ftsN2*" represents the gRNA target *ftsN* gene template strand. Each value comes from 6 replicates, the bars show the mean of the test, error bars indicate SD.

3.1.5 Appropriate repression condition for essential genes.

To detect the appropriate induction level of dCas9 expression by addition of aTc, I checked aTc dose dependency for repression of *dnaA* and *ftsN*. Using mid-log culture in LB with

ampicillin and chloramphenicol at 37ºC, OD600 of *dnaA* and *ftsN* targeted strains with 0, 0.1, 0.2, 0.5, 1.0 and 2.0μM of aTc was measured (Fig. 11). From OD600 values of *dnaA* and *ftsN* CRISPR-dCas9 repressed strains, growth inhibition was observed depending on the concentration of aTc (Fig. 11).

For *dnaA* repression, no significant repression was observed without aTc. On the other hand, clear repression was measured at the concentration of 0.1 μM of aTc and aTc dependency was observed until 0.5 μM. Saturation was observed over 0.5 μM.

For *ftsN*, repression was observed without aTc and leak expression of dCas9 without inducer might be occurred. Over 0.1 μM up to 2.0 μM of aTc addition, same level of repression with significance was identified.

This test revealed that, my CRISPR-dCas9 and gRNA knockdown system showed inducer dependent repression and full repression level might be occurred depending on target genes or target position of gRNA.

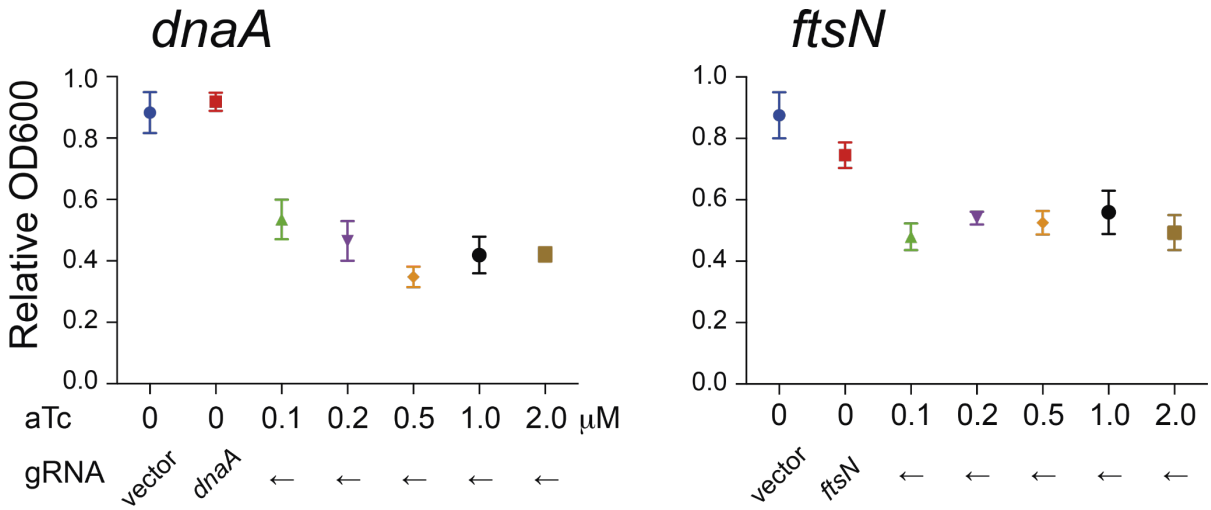


Figure 11: Inducer dose dependency of *E. coli* growth inhibition. X-axis represents the aTc concentration (μM), y-axis represents Relative OD600 value, the OD600 values are normalized by dividing the wild type strain OD600 value. "gRNA-vec" is the control system, which contains an irrelevant crRNA. Left panel is the *dnaA* target strain OD600 changes after 6 hours culture in LB with Cm and Am and different dose of aTc. Right panel is the *ftsN* target strain OD600 changes after 6 hours culture in LB with Cm and Am and different dose of aTc. vector means vector system which has an irrelevant crRNA with no target position in *E. coli* genome. *dnaA* represents the gRNA target *dnaA* gene; *ftsN* represents the gRNA target *ftsN* gene. Left arrow lines mean the same with left.

3.2 Construction of conjugation system.

3.2.1 Conjugation system construction

For high-throughput purpose, pdCas9-bacteria and pgRNA-bacteria vectors were converted to self-transmissible during conjugation by integration of *oriT* fragment from F plasmid. Two primers carrying NotI restriction enzyme adaptors were designed to amplify 400bp of *oriT* region carrying NotI restriction sites at both ends between *traM* and *ygfA* genes (see Fig. 12). Amplified *oriT* fragment was inserted into both pdCas9-bacteria and gRNA-bacteria plasmids between replication origin and transcriptional terminator with both directions. These *oriT* carrying plasmids were structurally confirmed by PCR across the cloning site, restriction pattern and BigDye terminator sequencing. Primers XYP0106 and XYP0107 for pdCas9 vector, and primers XYP100 and XYP0108 for pgRNA vector were used, respectively (Table 2). The cloning direction of *oriT* was confirmed by colony PCR using primer pairs across the cloning site.

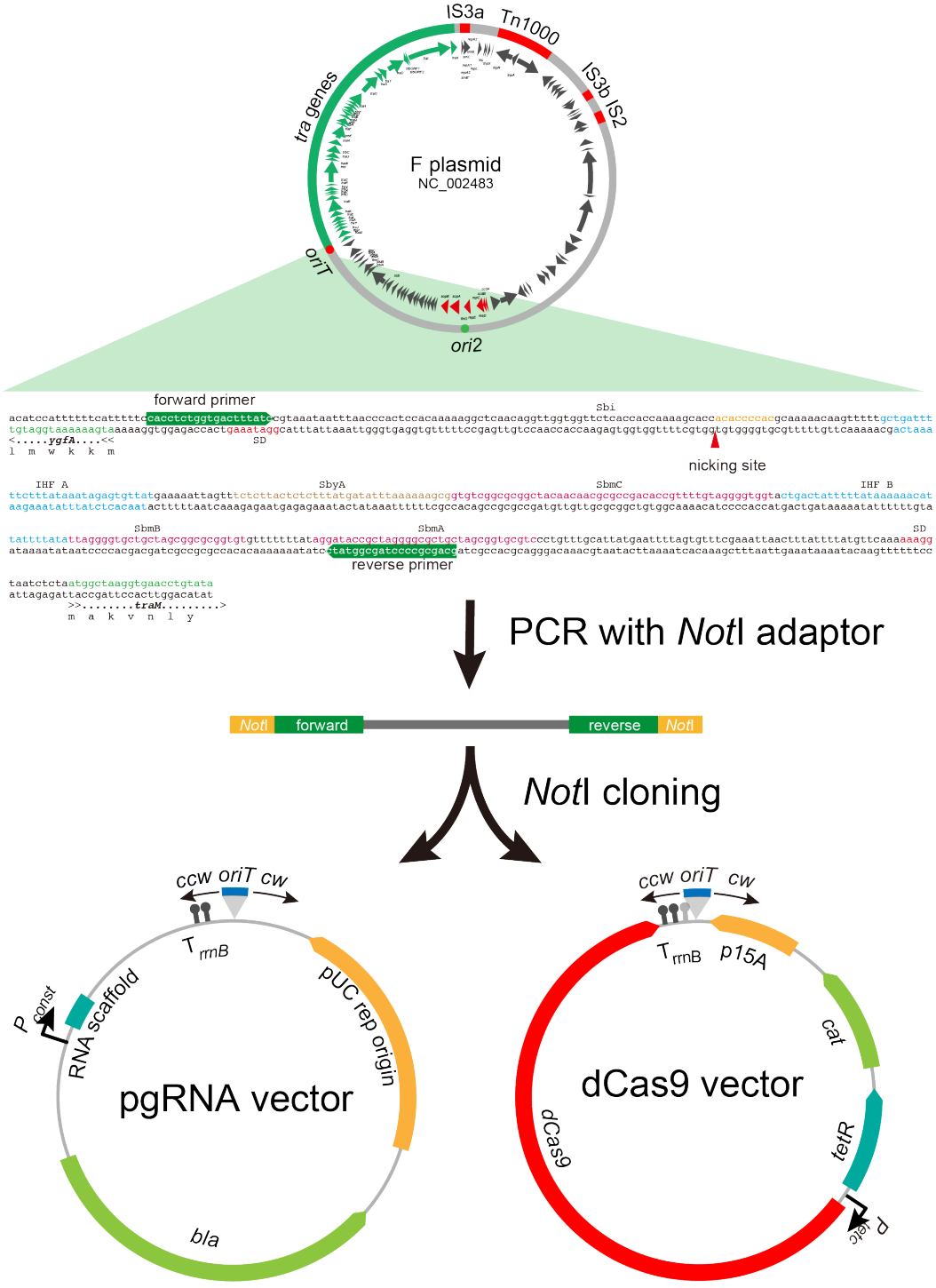


Figure 12: Conjugative CRISPR system construction. The *oriT* fragment cut from F plasmid, the full-length F-*oriT* fragment was amplified by PCR and assembled NotI adaptor on both sides. Purify the fragments and clone to insert the *oriT* fragment into pgRNA and pdCas9 vector plasmids by ligation, the finally clone products will generate pgRNA or pdCas9 plasmids with *oriT* fragment on both directions.

3.2.2 Construction of self-transmissible gRNA expression plasmids.

To insert the specific gRNA sequence into pgRNA-*oriT* vector, Gibson assembly seamless cloning method was employed here[129]. Gibson assembly doesn't need specific restriction enzyme

sites and the products have no scar sequence left. An important advantage for high-throughput construction of this method is “one tube one reaction method” needs fewer experimental steps.

For preparation of insert gRNA sequence with homology arms, I designed two 70nt synthetic oligo DNAs with 50nt as homology arm and 20nt as gRNA (Fig. 13). To prepare a double-strand DNA fragment for Gibson assembly, two synthetic DNAs were annealed to generate double-strand. The procedure is shown in Fig. 13.

Dependency of the length of homology arm for recombination was tested and I confirmed 20nt homology length had a practical efficiency for Gibson assembly for cost reduction of systematic construction of gRNA library (See supplementary Fig.1.).

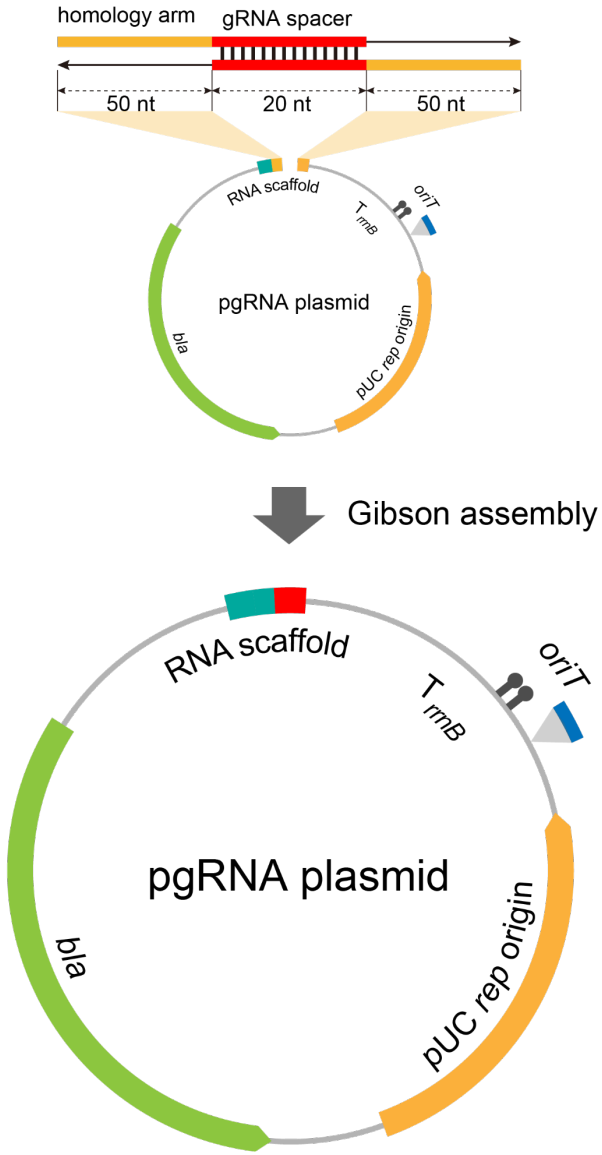


Figure 13: Specific gene target crRNA plasmids construction strategy. crRNA for essential gene target is cloned into the vector by Gibson assembly. Two 70nt single strand DNA are synthesized with 20nt overlap

(red fragment) and 50nt homologous with vector pgRNA plasmid. Annealing to generate double strands DNA with 20bp flank by 50bp homologous regions with pgRNA vector plasmid, mix the fragments with Gibson assembly buffer together to clone the specific crRNA into pgRNA-*oriT* plasmids.

3.2.3 Influence of *oriT* orientation on conjugal transfer efficiency.

To construct the donor strains of dCas9 and gRNA expression plasmids, G286 was employed, which is auxotrophic of Alanine as counter selectable marker and converted to Hfr by integration of CIP8 functioning as donor[45]. Firstly, the horizontal DNA transfer efficiency from Hfr donor cell to BW25113 recipient cells was detected.

Conjugation efficiency here is defined as the colonies appeared on the selective plates without DL-Alanine divided to recipient cells number. It clearly showed that the *oriT* direction strongly affected conjugative transmission ability. The clockwise orientation of transfer from *oriT* showed no DNA transfer ability, the counter clockwise direction, however, all showed high DNA transfer abilities. Size and copy number of the plasmids might also influence the final CFU (colony formation unites). The larger size of plasmid (pdCas9-*oriT*) had less CFU than smaller plasmid (pgRNA-*oriT*) under the same experimental conjugation condition (Fig. 14).

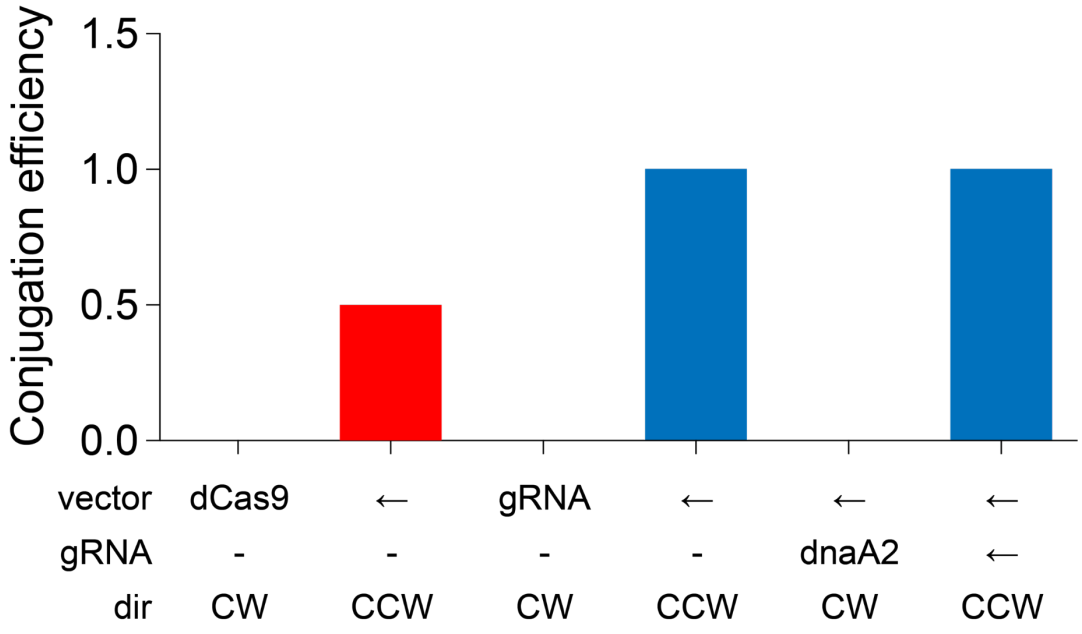


Figure 14: Conjugation efficiency of the plasmids with the F-*oriT* on different directions. y-axis represents conjugation efficiency; x-axis represents the experimental conditions, in vector row, dCas9 represents pdCas9-*oriT* plasmid, gRNA represents pgRNA-*oriT* target position, dnaA2 represents *dnaA* target pgRNA-*oriT* plasmid, "-" represents no this plasmid; Left arrow lines means the same with left. Dir represents the *oriT* direction, CW is clockwise direction, CCW is counter clockwise direction.

To optimize the conjugation time for transferring dCas9 or gRNA expression plasmids from G286 Hfr donor to BW25113 recipient strain, time series experiment of conjugal transfer was performed (see Fig. 15). After starting conjugation, the donor and recipient strains contact was stopped at indicated time points and counted conjugated cell numbers. As shown in Fig. 15, linear increase of conjugated cell number was observed until one hour. After one hour, however, dramatic increase of conjugated cells was detected probably because of cell proliferation. So, one hour for conjugation may be sufficient for plasmids transfer.

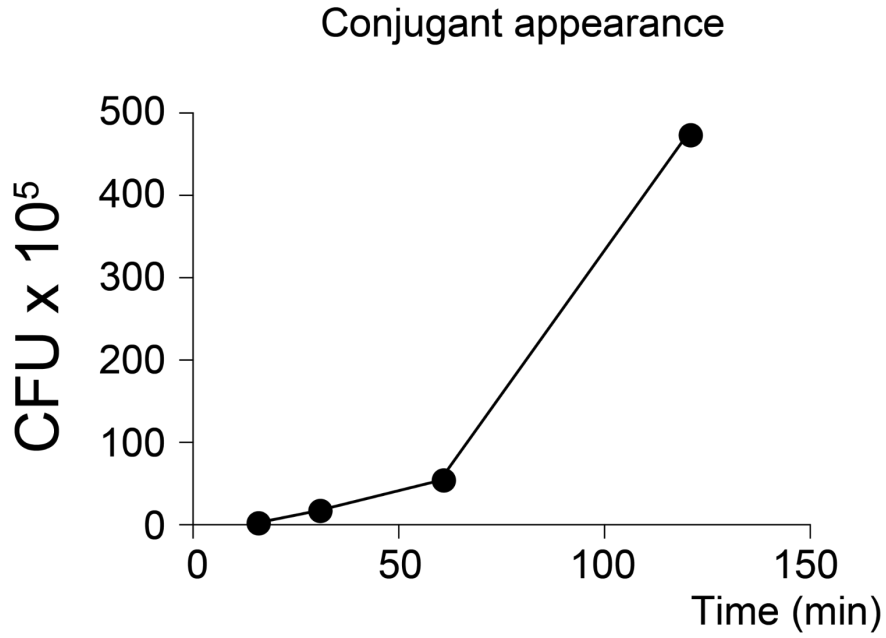


Fig 15: Relations of mating time and conjugation efficiency. X-axis is the mating time (minutes), y-axis indicates CFU.

3.3 Construction of CRISPR-dCas9 essential genes knockdown methodology.

3.3.1 CFU detection.

Colony-forming unit (CFU) is a common indicator to estimate the ability of the bacteria to give rise to colonies, theoretically, a viable cell can give rise to a colony through replication. CFU measurement for *dnaA* and *ftsN* genes in LB medium with or without inducer aTc indicates that, the repression of *dnaA* or *ftsN* gene by CRISPR-dCas9 system dramatically affected cell fitness. The *E. coli* strains colony formation ability are strongly being inhibited when *dnaA* or *ftsN* genes were repressed by supplying the inducer aTc on LB plates (Fig.16 A). CFU of BW25113 strains with CRISPR-dCas9 system have more than 10^8 times colonies in the medium without inducer than in the medium with 1 μM aTc (Fig. 16 B).

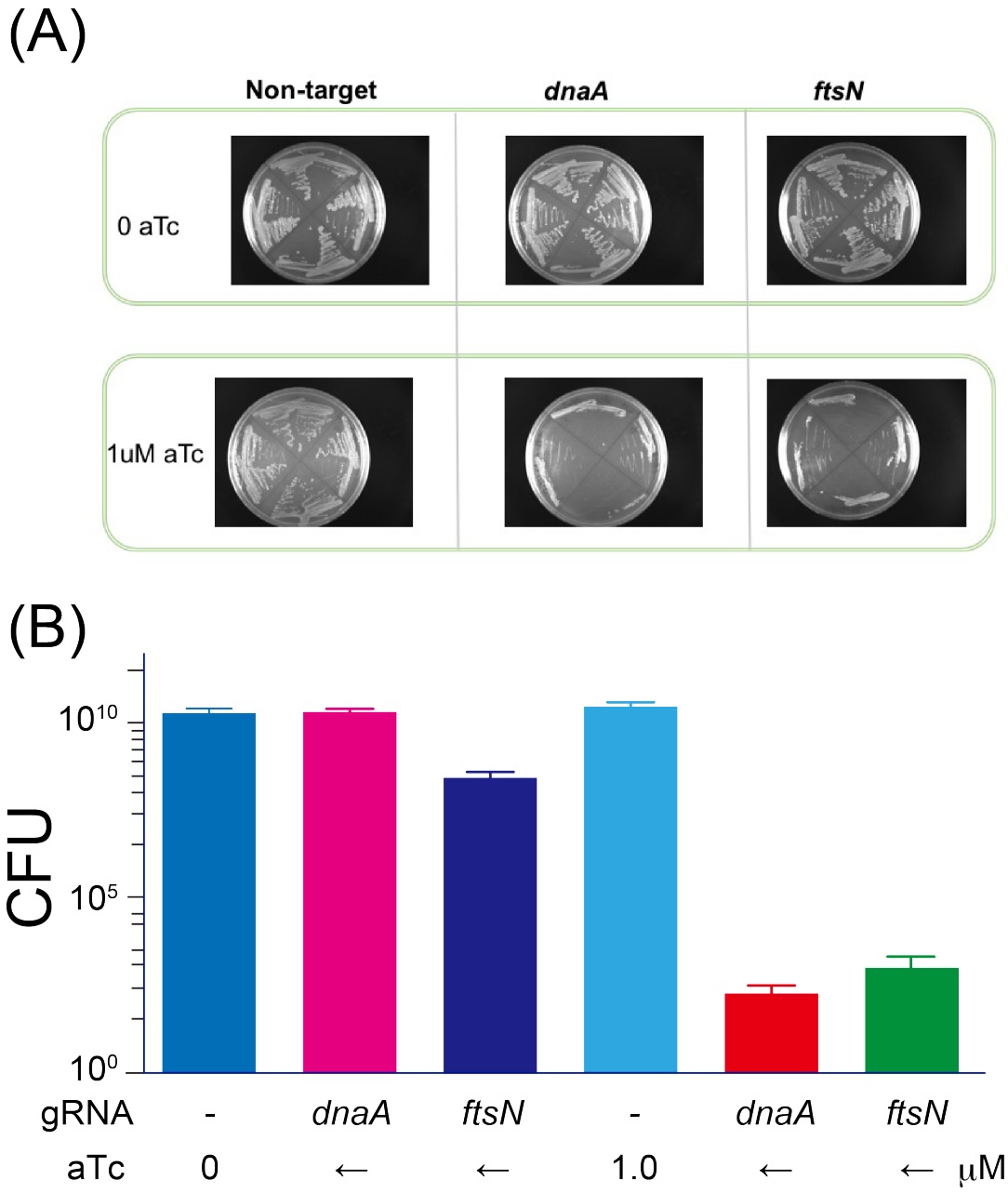


Figure 16: CRISPR-dCas9 essential gene target systems inhibited the colony formation on LB agar plates. A: growth inhibition test. Non-target represents no essential gene target crRNA, *dnaA* represents the CRISPR system inhibits *dnaA* expression, *ftsN* represents the CRISPR system inhibits *ftsN* expression. The top row is the strains grow on LB medium without aTc inducer; the second row is the strains grow on LB medium with 1 μM aTc. B: the CFU of the essential gene inhibited strains. Y-axis represents CFU, x-axis represents experimental conditions, gRNA row indicates the gRNA target gene, "-" represents absence, *dnaA* represents the CRISPR system inhibits *dnaA* expression, *ftsN* represents the CRISPR system inhibits *ftsN* expression. aTc row marks the aTc concentration (μM), left arrow lines mean the same with left.

3.3.2 Growth inhibition by conjugative CRISPR-dCas9 system.

Quantitatively checking the repression levels of this conjugative CRISPR-dCas9 system in liquid, I monitored 24 hours growth by monitoring OD600 of BW25113 strain with *dnaA* target system with or without aTc (Fig. 17). *dnaA* repressed strain shows significant weak growth when aTc was supplied, and 1 μM aTc system had stronger inhibition than 0.2 μM aTc system. One thing we should pay attention is that growth recovery in *dnaA* repressed strain by CRISPRi after longer time incubation was observed (Fig. 17C). Suppressor mutation(s) or inactivation of CRISPR-dCas9 system may be occurred. In log phase culture in LB with ampicillin and chloramphenicol at 37 °C, the *dnaA* target system shows lower cell densities compared with the control culture, and higher aTc concentration shows a stronger grow inhibition.

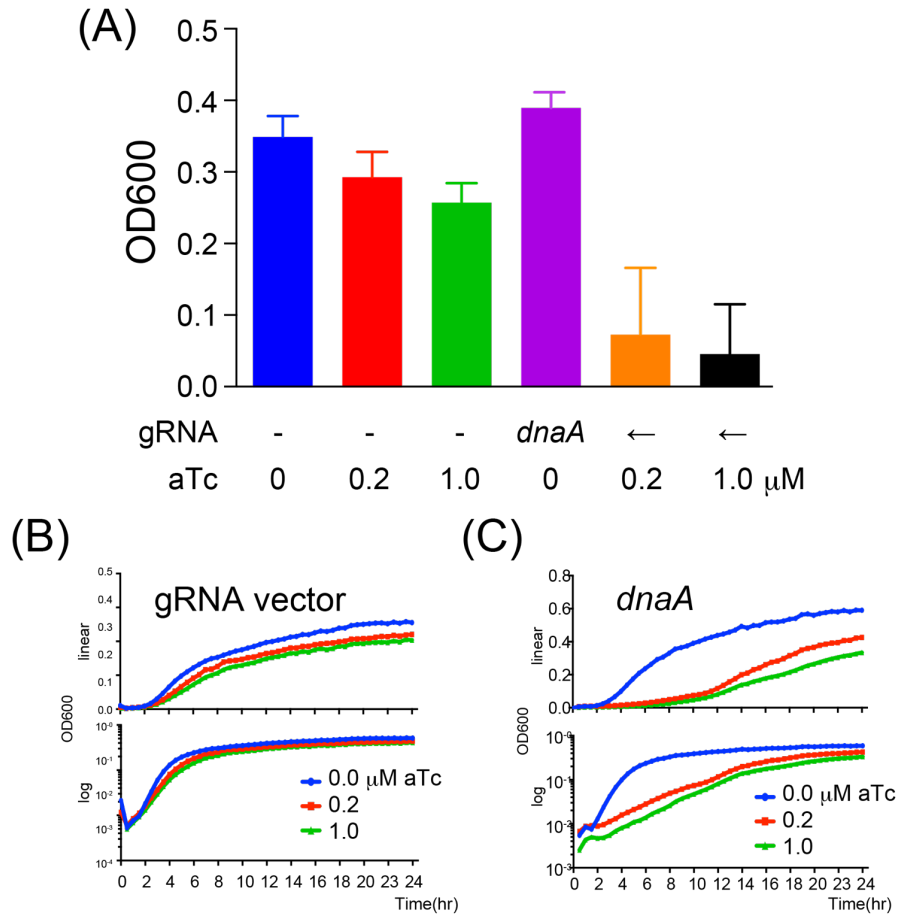


Figure17: CRISPR-dCas9 inhibits cell growth in liquid culture. (A): OD600 values of *dnaA* knockdown strains after 6 hours culture in LB medium with antibiotics (Cm, Am). y-axis represent OD600 value, x-axis represents experimental conditions, gRNA row represents gRNA target gene, "-" represents no gene target position in *E. coli* genome, *dnaA* represents the CRISPR system inhibits *dnaA* expression, aTc row marks the aTc concentration (μM), left arrow lines mean the same with left. (B): the growth profiles in LB medium with antibiotics at 37°C of control system (pgRNA vector) in 0 μM aTc (blue), 0.2 μM aTc (red) or 1.0 μM aTc (green). X-axis represents time points (hours), y-axis represents linear OD600 value (top) or log OD600 value (bottom). gRNA-vector represents CRISPR system with no gene target. (C): the growth profiles in LB

medium with antibiotics at 37°C of *dnaA* knockdown system in 0 μM aTc (blue), 0.2 μM aTc (red) or 1.0 μM aTc (green). X-axis represents culture time (hours), y-axis represents linear OD600 value (top) or log OD600 value (bottom).

3.3.3 mRNA repression.

Activation of dCas9 protein expression in *dnaA* and *ftsN* targeted strains lead obvious growth inhibition of *E. coli* cells. To confirm decrease of mRNA level by CRISPRi, quantitative measurements of mRNA levels of these genes by qRT-PCR was performed. The mRNA expression levels are shown as relative ratio between with and without gRNA target. The qRT-PCR experiments showed clear consistent results the those from growth inhibition phenotypes. A shown in Fig. 18, the growth inhibited strains shows significantly lower level of mRNA expression.

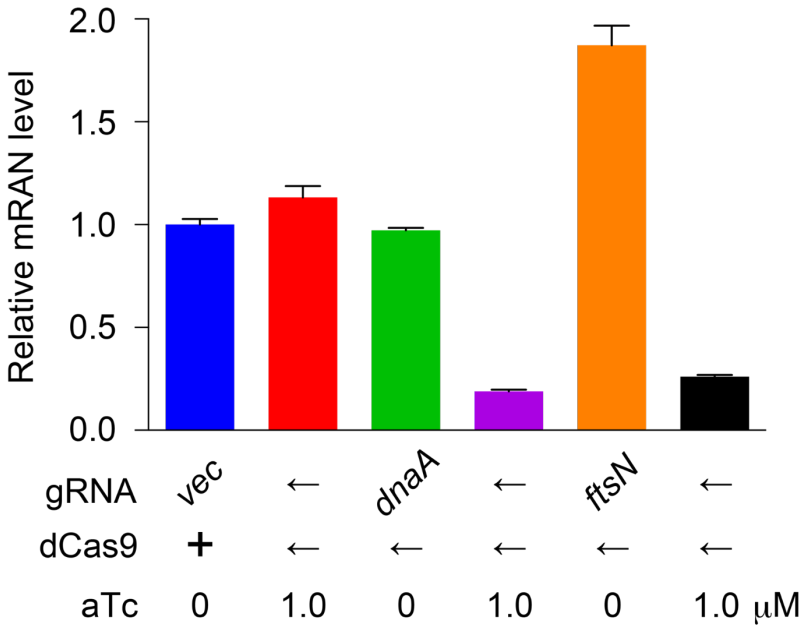


Figure 18: CRISPR-dCas9 system repressed the essential gene mRNA expression level. Y-axis represents relative mRNA level, the mRNA expression level divided to the reference gene (16s rRNA) mRNA expression level and normalized by the mRNA levels in no essential gene target system. x-axis represents the experimental conditions. gRNA row represents gRNA target gene, vec represents vector plasmid with no essential gene target, *dnaA* represents the CRISPR system inhibits *dnaA* expression, *ftsN* represents the CRISPR system inhibits *ftsN* expression. aTc row marks the aTc concentration (μM), left arrow lines mean the same with left.

3.3.4 Morphology change by CRISPR-dCas9 essential gene knockdown.

Bacterial morphology is one of the basic phenotypes to classify functions of genes. Some of the essential genes' function as the critical factor for *E. coli* cell shapes like *ftsZ* is Z-ring formation related genes, which contributes to cell division at the middle of the cells. *mreB* is

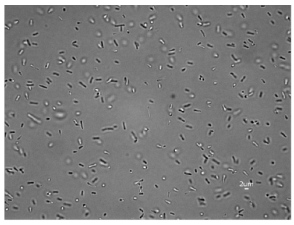
cell skeleton coding gene which contributes to cell wall elongation. *ftsN* is a member of cell division protein encoding gene, which combines with *ftsA* and *ftsQLB* to active cell wall synthesis, finally leading to the formation of the septum and the cell division of the cell[133].

Here we monitored the time series morphology changes of the essential gene repression strains in LB medium at 37 °C (Fig.19). Monitoring the morphology changes under the microscope was started using overnight pre-culture, which showed the normal short rod shape of *E. coli*.

dnaA repressed cells showed gradual increase of cell length with aTc and kept normal shapes without induction. With 1 μM aTc, the cell density of the culture did not increase until 10 hours. After 10 hours however, the short rod shape *E. coli* cells started to be appeared and the bacteria started to increase.

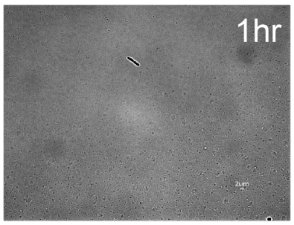
The initial culture of *ftsN* repressed strain showed normal size rod shape *E. coli* but with a small number of elongated cells. Which may be because of the leak expression of the dCas9 protein from the plasmid. While, in the LB medium without aTc, normal short rod shape *E. coli* cells were dominant in the bacterial culture. However, with 1 μM aTc in LB, the signifcantly filamentous cells became dominant in the culture until 8 hours at 37°C. After 10 hours, the normal size cells started to be appeared and became the main population in the culture.

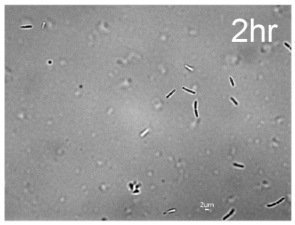
dnaA

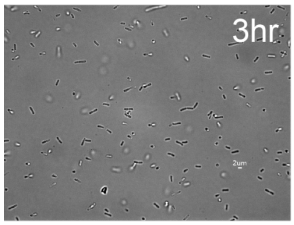


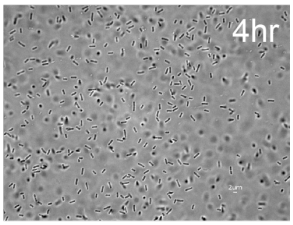
over night pre-culture without inducer, aTc.

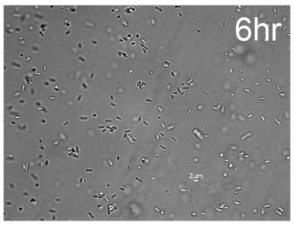
0 μM aTc

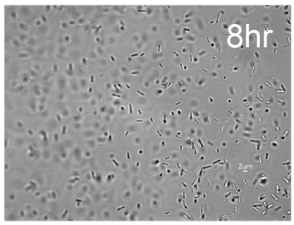


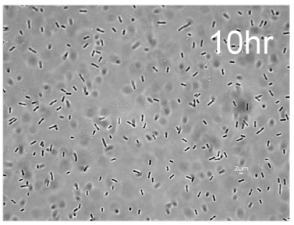


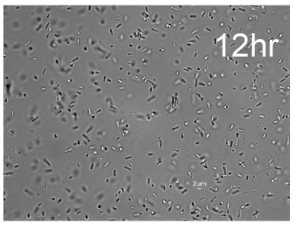


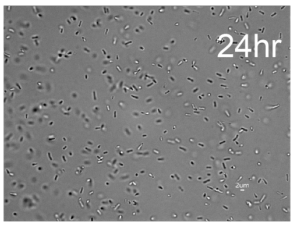




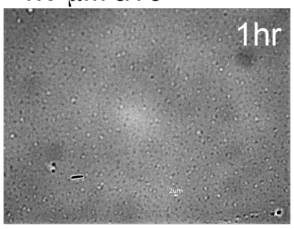


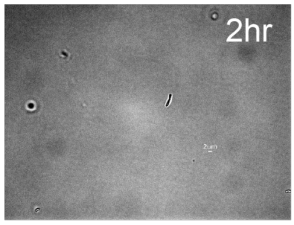


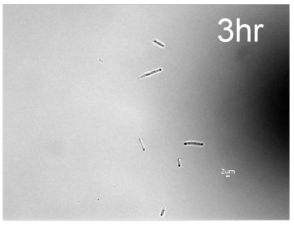


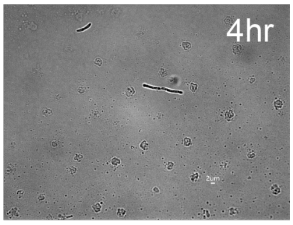


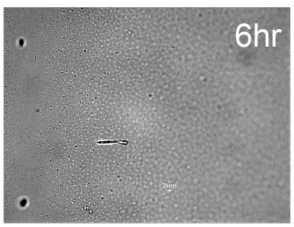
1.0 μM aTc

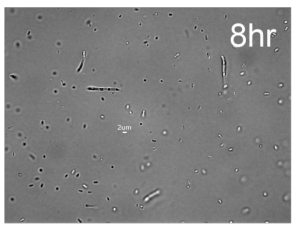


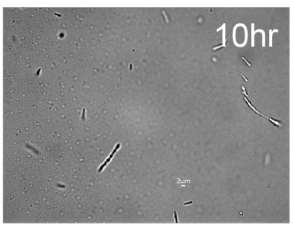


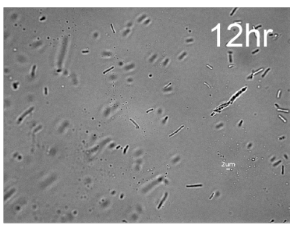


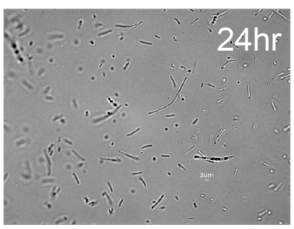












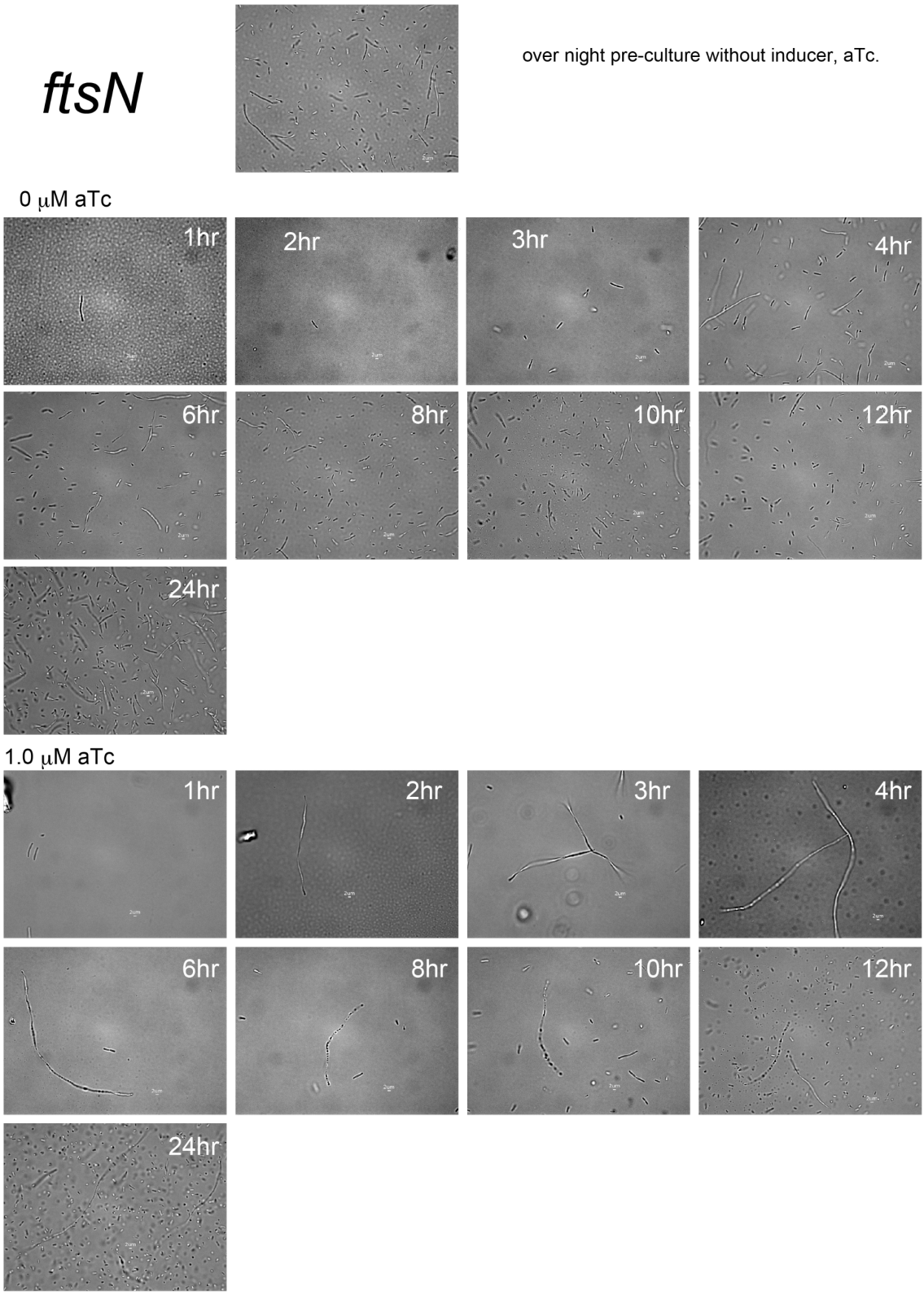


Figure 19: Time series morphology change of *dnaA* and *ftsN* repression strains. *E. coli* cells morphological detection with *dnaA* / *ftsN* inhibition system in 0 μM aTc (top) and 1 μM aTc (bottom) by microscopy at different time points. The morphology monitor initiated from overnight culture in LB medium with antibiotics (Am, Cm) but no aTc. 1/1000 diluted to fresh LB medium with antibiotics (Am, Cm) and 0 μM aTc or 1 μM aTc. *E. coli* morphology was observed under 100 times magnification by microscopy. The observation time point is marked on the picture.

3.3.5 Genetic interaction analysis of essential gene knockdown and non-essential gene knockout

The GI of essential genes and the Keio collection detection was performed follow the procedure mentioned in methods; firstly, the pdCas9-*oriT* plasmid was transferred into strains of the Keio collection by conjugation, thereafter transfer pgRNA-*oriT* plasmid into the Keio +pdCas9-*oriT* strains by conjugation.

Quantitative measurement of colonies on the plates was done by Colony-live in time-series manner (Fig. 20). The colonies at outside edges on the plates were BW25113 wild type strain with single essential gene repression system, which were transferred with CRISPR-dCas9 *dnaA* gene repression system. The rest of colonies have different single gene knockout with CRISPRi-*dnaA* repression plasmid. As shown in Fig. 20, some colonies clearly showed missing or smaller than the control plates without inducer.

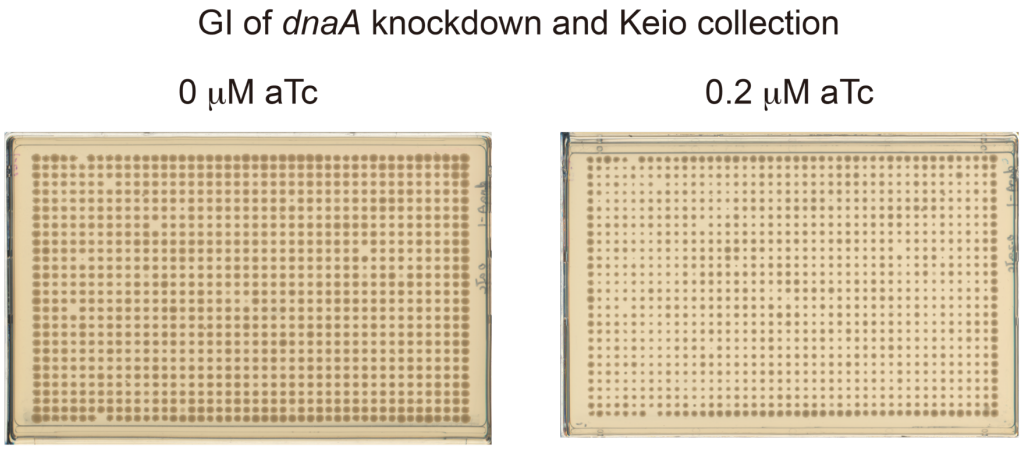


Figure 20: GI analysis of *dnaA* & Keio collection. *dnaA* colony detection on the LB plates with 0 μM aTc or 0.2 μM aTc. Each plate contains 1536 colonies, (48 columns and 32 rows), each colony contains one non-essential gene deletion, and carrying the inducible CRISPR-dCas9 system. The essential gene (*dnaA*) inhibition is promoted by supplying 0.2 μM aTc in the medium.

The double mutants' colonies distribution (Fig.21) were clearly shifted to smaller MGR (Maximum Growth Rate) when supplying the inducer aTc in *dnaA* or *ftsN* target system. Meanwhile, some synthetic lethal genes appeared in 0.2 μM aTc in *dnaA* repressed system. For *ftsN* repression, genes showing synthetic lethality were increased according to the inducer concentration.

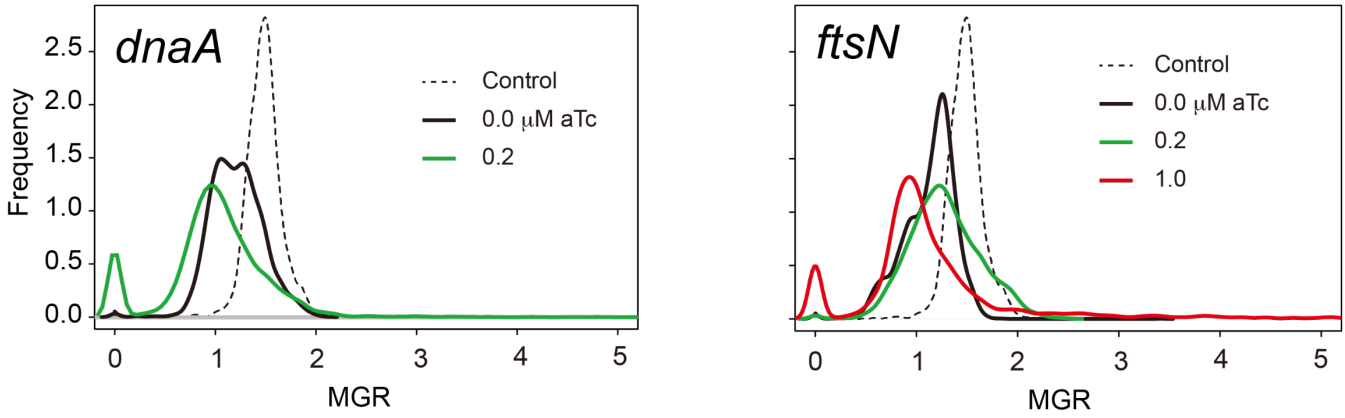


Figure 21: *dnaA* /*ftsN* knockdown and non-essential gene knockout strains maximum growth rates (MGR) distribution. The MGR distribution of *dnaA* (left) and *ftsN* (right) knockdown and Keio collection combination. Control system represents no essential gene target (black dash lines); black curves represent *dnaA* (left) or *ftsN* (right) inhibition in 0μM aTc, green curves represent *dnaA* (left) or *ftsN* (right) inhibition in 0.2 μM aTc, red curve represents *ftsN* (right) inhibition in 1 μM aTc.

The global view of essential gene knockdown and non-essential gene knockout mutants, the control system contains CRISPR-dCas9 vector system, which has no gene target position in *E. coli*. In *dnaA* repression system, synthetic lethal appeared when supply 0.2 μM aTc, for *ftsN* repression system, when it was supplied with 0.2 μM aTc, some lethal gene pairs appeared, and more lethal gene pairs arose with 1μM aTc.

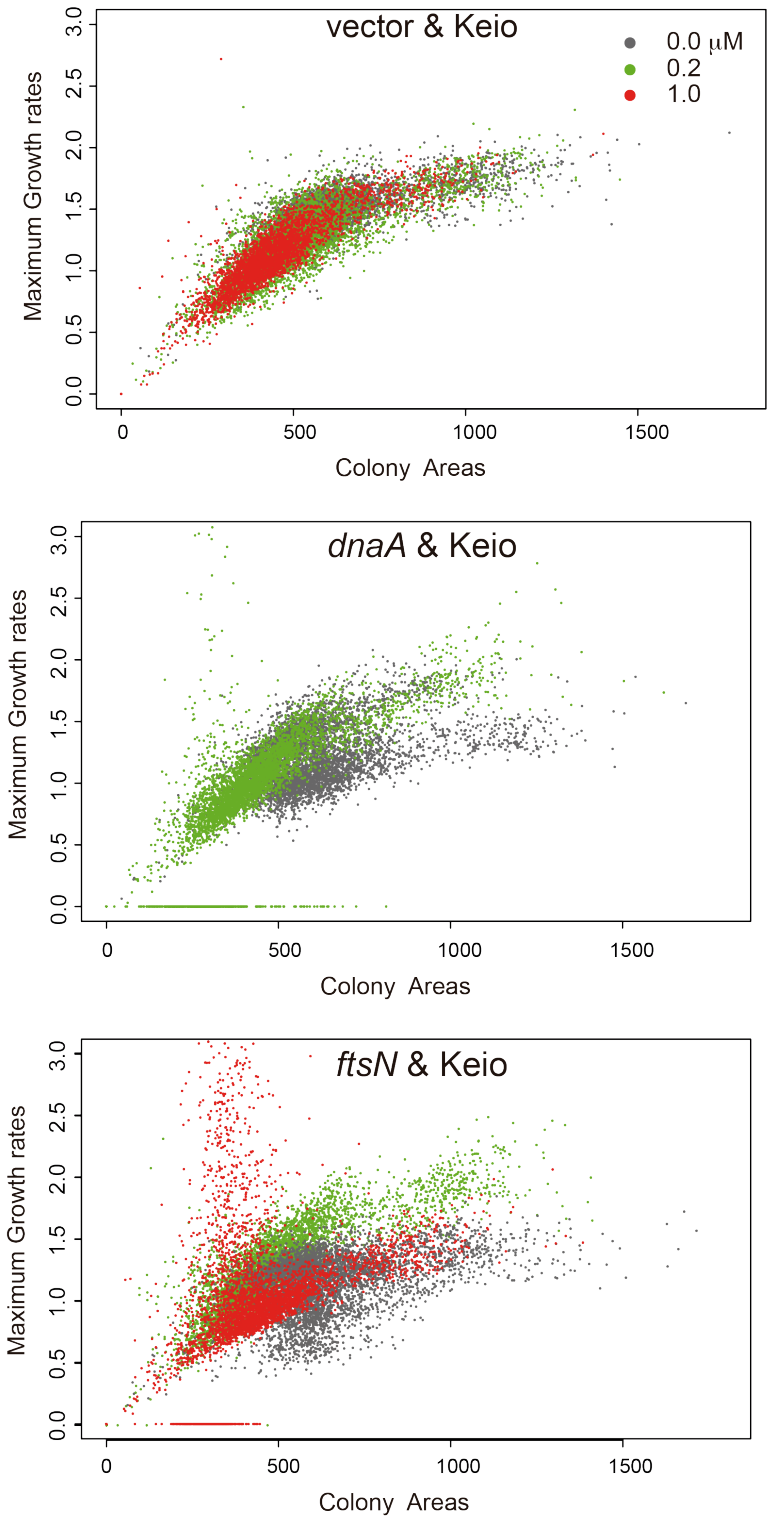


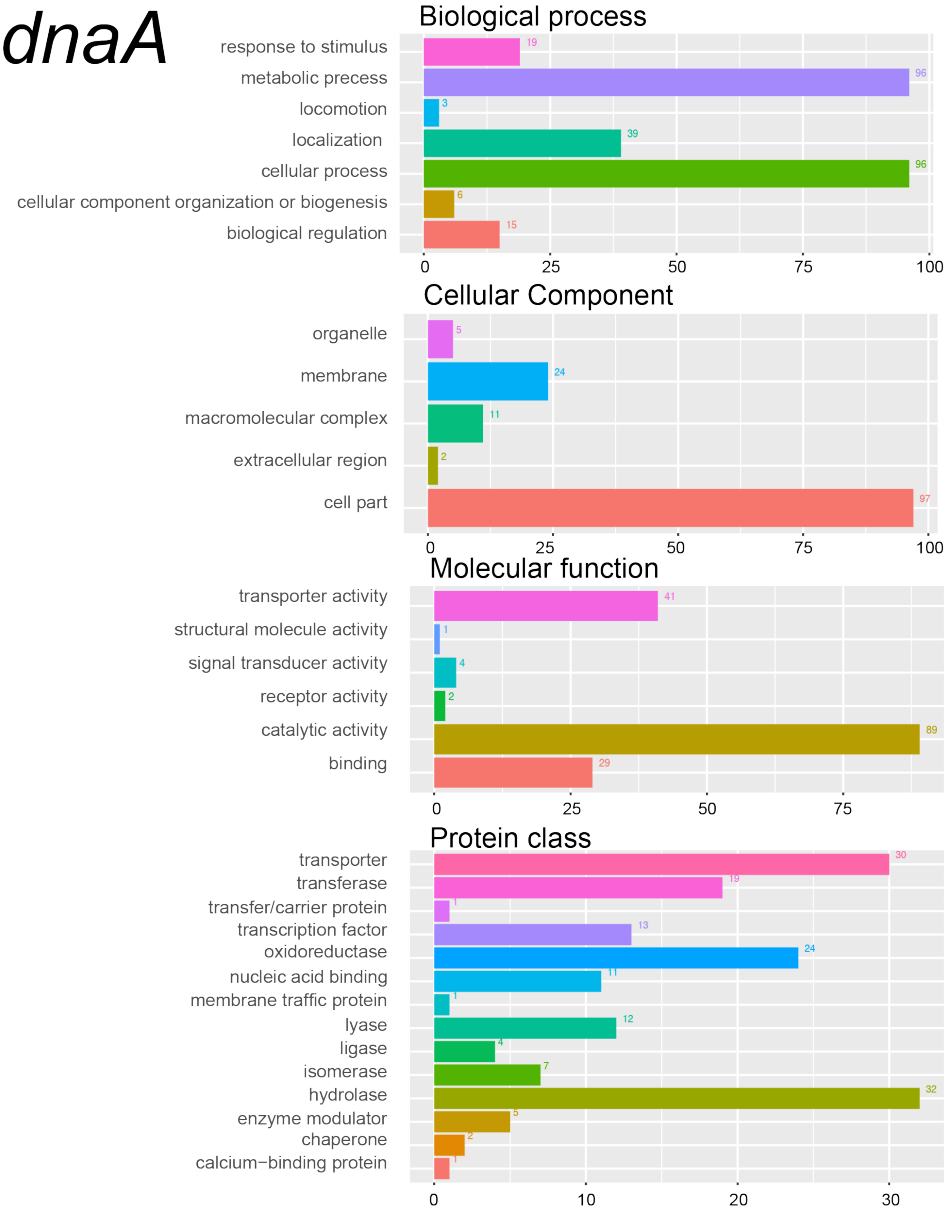
Figure 22: Essential gene knockdown and non-essential gene knockout combination. The global view of *dnaA* or *ftsN* knockdown and entire non-essential gene deletion combination. X-axis represents colony areas, y–axis represents Maximum growth rates. The Control system contains irrelevant CRISPR system which means has no gene target in *E. coli* genome. Dark grey points represent strains grow in LB medium with antibiotics (Km, Cm, Am) and 0 μM aTc. Green points represent strains grow in LB medium with antibiotics (Km, Cm, Am) and 0.2 μM aTc. Red points represent strains grow in LB medium with antibiotics (Km, Cm, Am) and 1.0 μM aTc.

3.3.6 Functional classification of *dnaA* and *ftsN* synthetic lethal genes.

Synthetic lethal interactions attracted lot of interests of geneticists, because they reveal information about the functional relationships between genes and are easy to score in genetic screens. In biological processes, many important basic cellular processes are related to cell viability, synthetic lethality screens are a good way to study a wide range of cellular processes. Here, I detected the synthetic lethal gene pairs of *dnaA* and *ftsN* repression combine with single essential gene deletion library Keio collection. The list of synthetic lethal genes of *dnaA* (supplementary Table 2) and *ftsN* (supplementary Table 3) are shown in supplementary tables.

The synthetic lethal genes of *dnaA* and *ftsN* are involved in several classes (Fig.23.), *dnaA* is a global regulator in *E. coli* cells, it plays a key role in the initiation and regulation of chromosomal replication, many *dnaA* synthetic lethal genes are in crucial metabolic process and cellular process. *ftsN* is an essential cell division protein that activates septal peptidoglycan synthesis and constriction of the cell, *ftsN* knockdown also affected many genes in metabolic process and cellular process.

dnaA


Biological process
response to stimulus
metabolic precess
locomotion
localization
cellular process
cellular component organization or biogenesis
biological regulation
19
96
3
39
96
6
15
0
25
50
75
100
Cellular Component
organelle
membrane
macromolecular complex
extracellular region
cell part
5
24
11
2
97
0
25
50
75
100
Molecular function
transporter activity
structural molecule activity
signal transducer activity
receptor activity
catalytic activity
binding
41
1
4
2
89
29
0
25
50
75
Protein class
transporter
transferase
transfer/carrier protein
transcription factor
oxidoreductase
nucleic acid binding
membrane traffic protein
lyase
ligase
isomerase
hydrolase
enzyme modulator
chaperone
calcium−binding protein
30
19
1
13
24
11
1
12
4
7
32
5
2
1
0
10
20
30

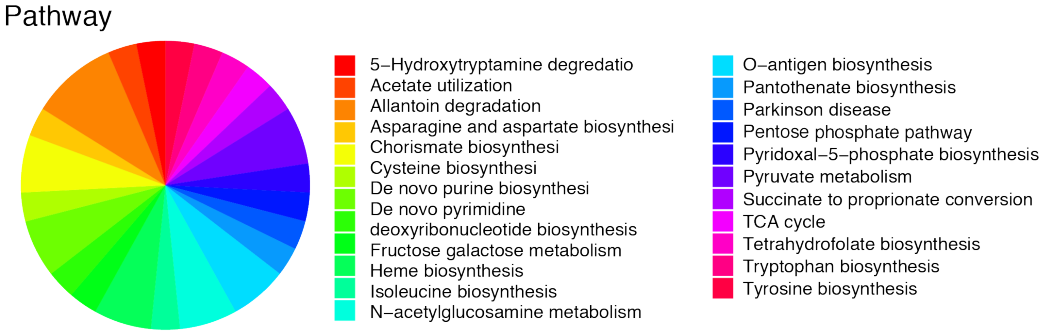

Pathway
5−Hydroxytryptamine degredatio
Acetate utilization
Allantoin degradation
Asparagine and aspartate biosynthesi
Chorismate biosynthesi
Cysteine biosynthesi
De novo purine biosynthesi
De novo pyrimidine
deoxyribonucleotide biosynthesis
Fructose galactose metabolism
Heme biosynthesis
Isoleucine biosynthesis
N−acetylglucosamine metabolism
O−antigen biosynthesis
Pantothenate biosynthesis
Parkinson disease
Pentose phosphate pathway
Pyridoxal−5−phosphate biosynthesis
Pyruvate metabolism
Succinate to proprionate conversion
TCA cycle
Tetrahydrofolate biosynthesis
Tryptophan biosynthesis
Tyrosine biosynthesis

Figure 23: Synthetic lethal genes classification. *dnaA*/ *ftsN* with Keio collection synthetic lethal genes classes in different categories. The numbers on the bars represents the genes hit the class. The pie chart indicates the pathways in which the *dnaA* /*ftsN* synthetic lethal genes involved.

3.3.7 Suppressing of growth inhibition.

The growth profiles of *dnaA* and *ftsN* inhibition exhibit that, after certain hours' culture (10 hours or 12 hours), the growth of inhibited strains started to be recovered, this phenomenon may come from the stress response of the essential gene repression strains or depletion of CRISPR-dCas9 factors. To seek the possible reason for this recovery, I collected the strains after 24 hours culture in the medium with aTc and diluted into the same new fresh medium (with 1μM aTc or without aTc) to monitor the growth curves (Fig. 24). For both *dnaA* and *ftsN* genes, the recovered cells didn't present growth inhibition even supplied the new inducer in fresh medium.

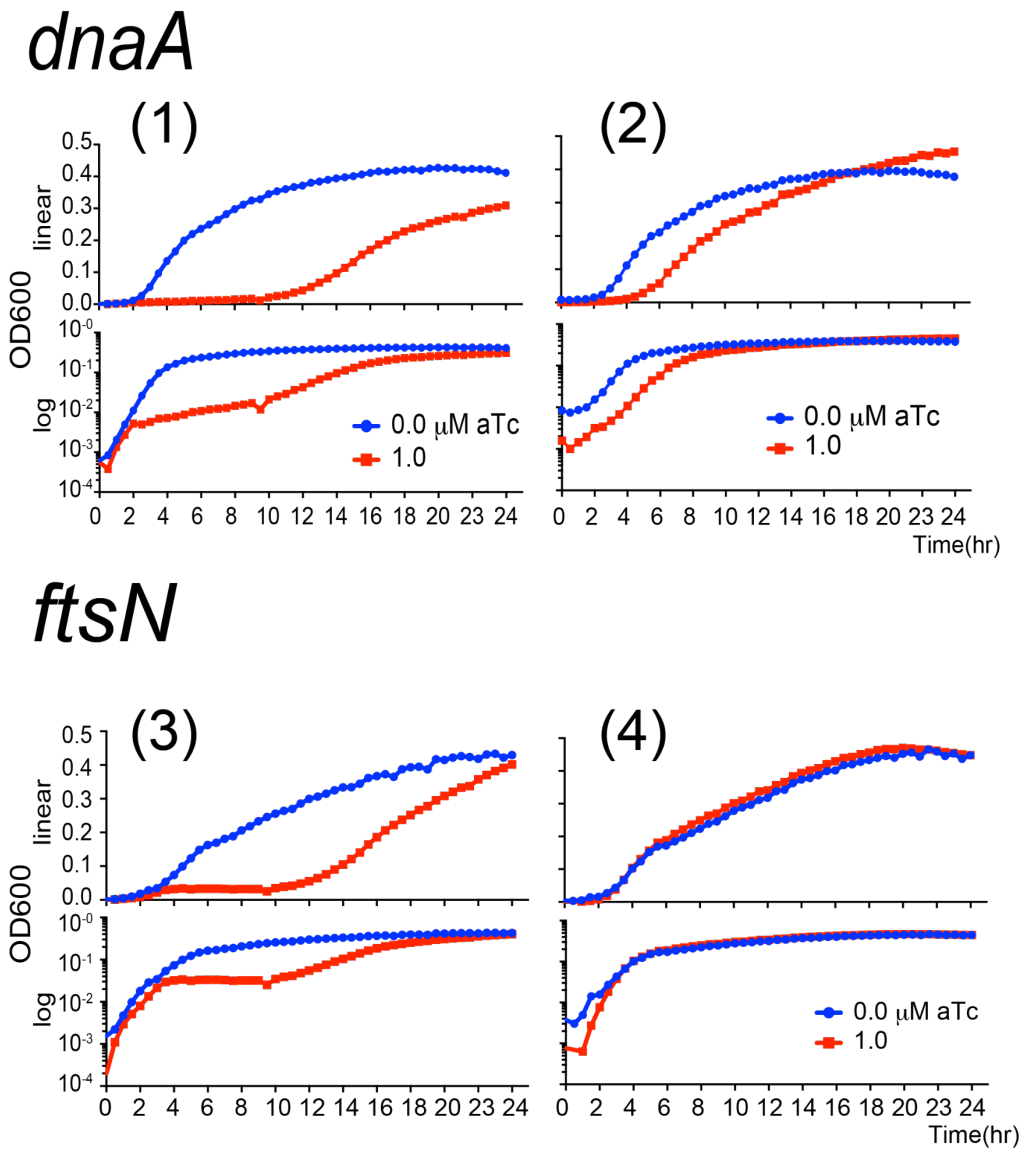


Figure 24: Suppressing of growth inhibition. Growth profiles of *dnaA* and *ftsN* inhibited strains. Blue curves represent culture in 0 μM aTc, red curves represent culture in 1.0 μM aTc. (1) shows the growth curves of the strain with *dnaA* repression system culture in the LB mediums with antibiotics (Cm, Am) and aTc (0 μM or 1.0 μM) at 37°C for 24 hours. Y-axis represent linear OD600 values (top) or log OD600 values (bottom).

X-axis represent time points(hours). (2) Re-growth of *dnaA* inhibited strain overnight culture in fresh medium with the same condition. Y-axis represent linear OD600 values (top) or log OD600 values (bottom). X-axis represent time points(hours). (3) shows the growth curves of the strain with *ftsN* repression system culture in the LB mediums with antibiotics (Cm, Am) and aTc (0 μM or 1.0 μM) at 37°C for 24 hours. Y-axis represent linear OD600 values (top) or log OD600 values (bottom). X-axis represent time points (hours). (4) Re-growth of *ftsN* inhibited strain overnight culture in fresh medium with the same condition. Y-axis represent linear OD600 values (top) or log OD600 values (bottom). X-axis represent time points (hours).

3.4 Systematic genetic interaction analysis by combination of knockdown and knockout strains

3.4.1 Construction of the first knockdown library of essential genes of *E. coli*.

From the previous trials, the function of the conjugative CRISPR-dCas9 system were confirmed, to expand the essential gene repression to systematic level, 21 essential genes were chosen as the first group to detect the high throughput strategy for essential gene knockdown library construction.

Essential genes picked here are detected by previous single ORF deletion strategy[5]. 21 targets are classified into different biological precesses. 9 of them are involved in cellular progress (*ftsK, yibJ, ftsZ, mreC, tsf, dnaC, mrdB, rpoC, mukB*). 2 are involved in biological regulation process (*mreC, mrdB*). 2 are involved in developmental process (*mreC, mrdB*). 1 gene is in localization process (*yibJ*). 6 are involved in metabolic process (*ftsZ, dnaC, thiL, rpoC, tsf, murB*). And the rest 10 of them are unclustered (Fig. 25).

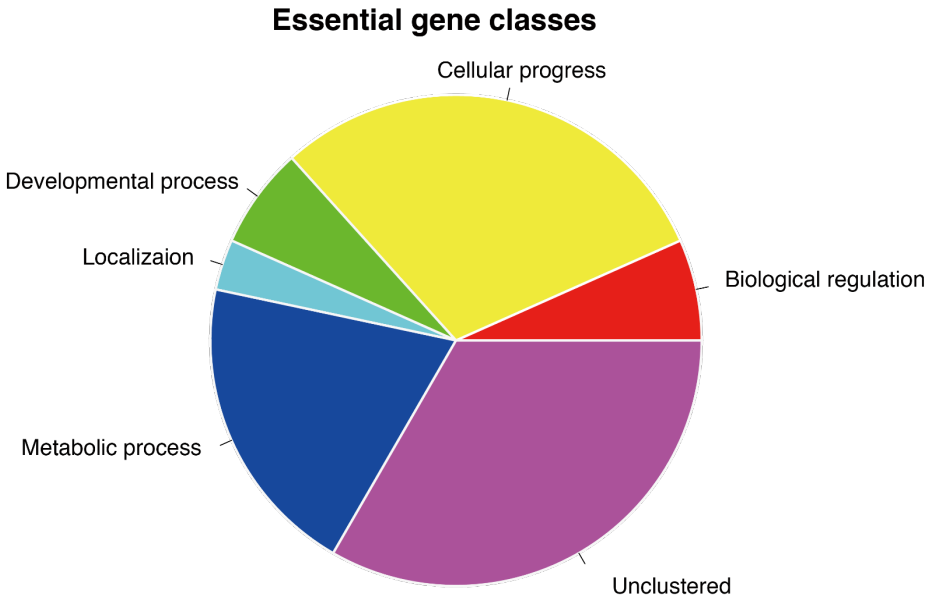


Figure 25: Essential gene classification. The entire 325 essential genes' biological process classification[5, 134].

These target genes are in different operons, some belong to single operon (like *lpxC, infA, ftsK, yceQ, metG*), some are in multiple operons (the others). The target genes in multiple operons are located in different positions, some are the first position of the operon (*lolC, dnaA*), some targets gene are the end expression unit in the operon (*tsf, murA, tpr, mukB, rpoC, ftsN*). We also picked 2 essential genes belong to the same operon (*mreB, mreC*) (Table 3).

Table 3 Operon structure of target essential genes

Gene name	Operon unit
ftsK	*ftsK*
infA	*infA*
lpxC	*lpxC*
metG	*metG*
yceQ	*yceQ*
ftsN	*cytR--ftsN*
murA	*ibaG--murA*
rpoC	*rpoB--rpoC*
tpr	*tyrV--tpr(rttR)*
dnaA	*dnaA--dnaN--recF*
dnaC	*dnaT--dnaC--yjjA*
lolC	*lolC-lolD-lolE*
tsf	*tff-- rpsB--tsf*
yibJ	*yibA--yibJ--yibG*
mukB	*cmoM--mukF--mukE--mukB*
secD	*yajC--secD--secF--yajD*
mreB	*mreB--mreC--mreD-yhdE--rng*
mreC	*mreB--mreC--mreD-yhdE--rng*
rpoA	*rpsM--rpsK--rpsD--rpoA--rplQ*
mrdB	*rsfS--rlmH--mrdA--mrdB--rlpA--dacA*
thiL	*nrdR--ribD--ribE--nusB--thiL--pgpA*
rplC	*rpsJ--rplC--rplD--rplW--rplB--rpsS--rplV--rpsC--rplP--rpmC--rpsQ*

Specific gRNA for each gene was designed. The target positions are mostly near the 5'ends in the ORF regions. PAM sequence (5'—NGG—3') was searched out and the 20nt fragment before this PAM sequence was chosen as the gRNA. In previous paper, the first 12nt close to PAM is essential for CRISPR-dCas9 repression, (which is called seed sequence), one single mutation in this region will dramatically damage the repression. To avoid off-target effect, the crRNAs were BLAST with *E. coli* K-12 genome sequence, to make sure there is no other position overlap with this target sequence (at least the seed sequence is unique). Fig. 26 shows the precise target position of CRISPR-dCas9 system, the initiation of the target gene cassettes is start codons. In this trial, all the gRNAs are target non-template strand.

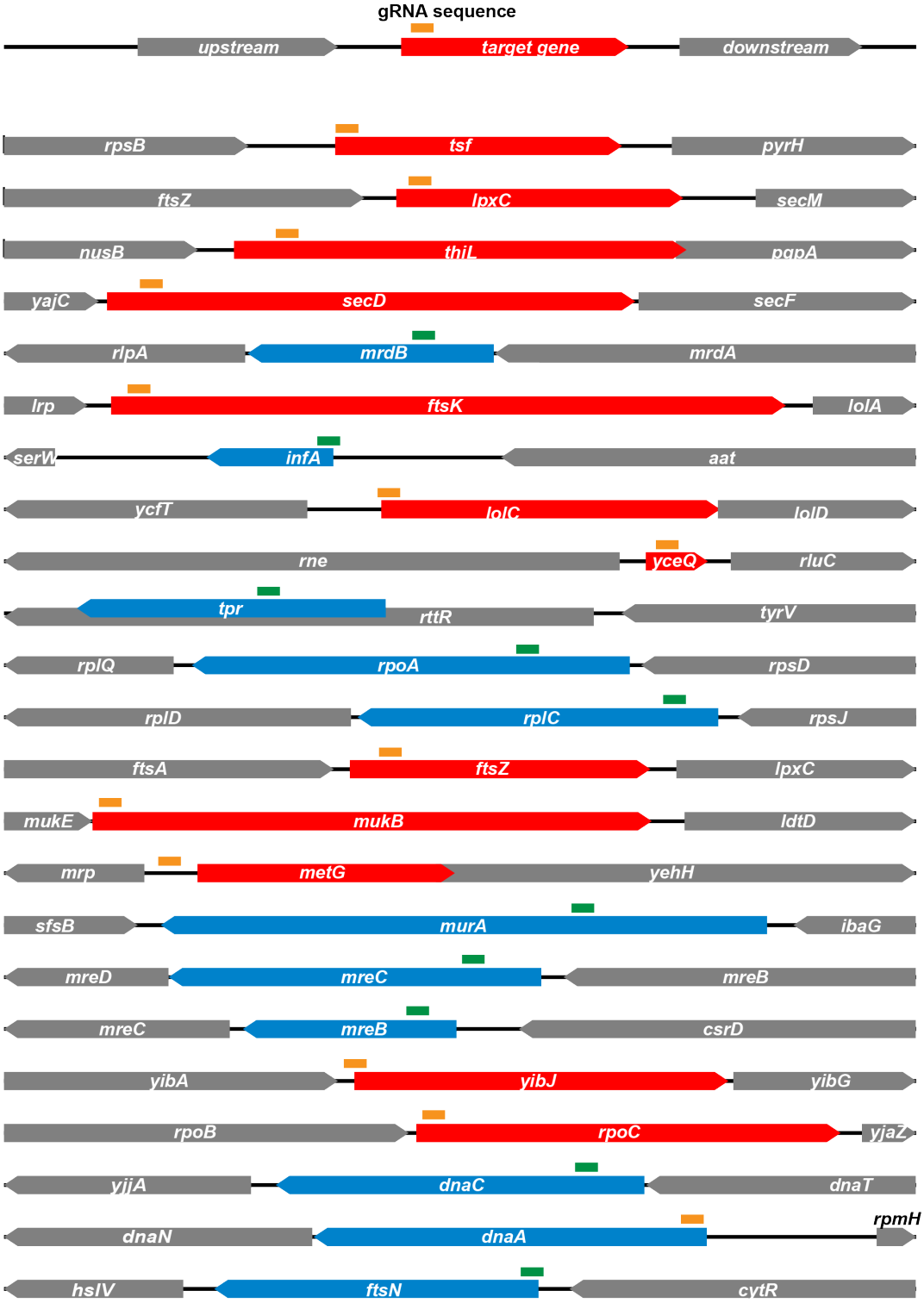


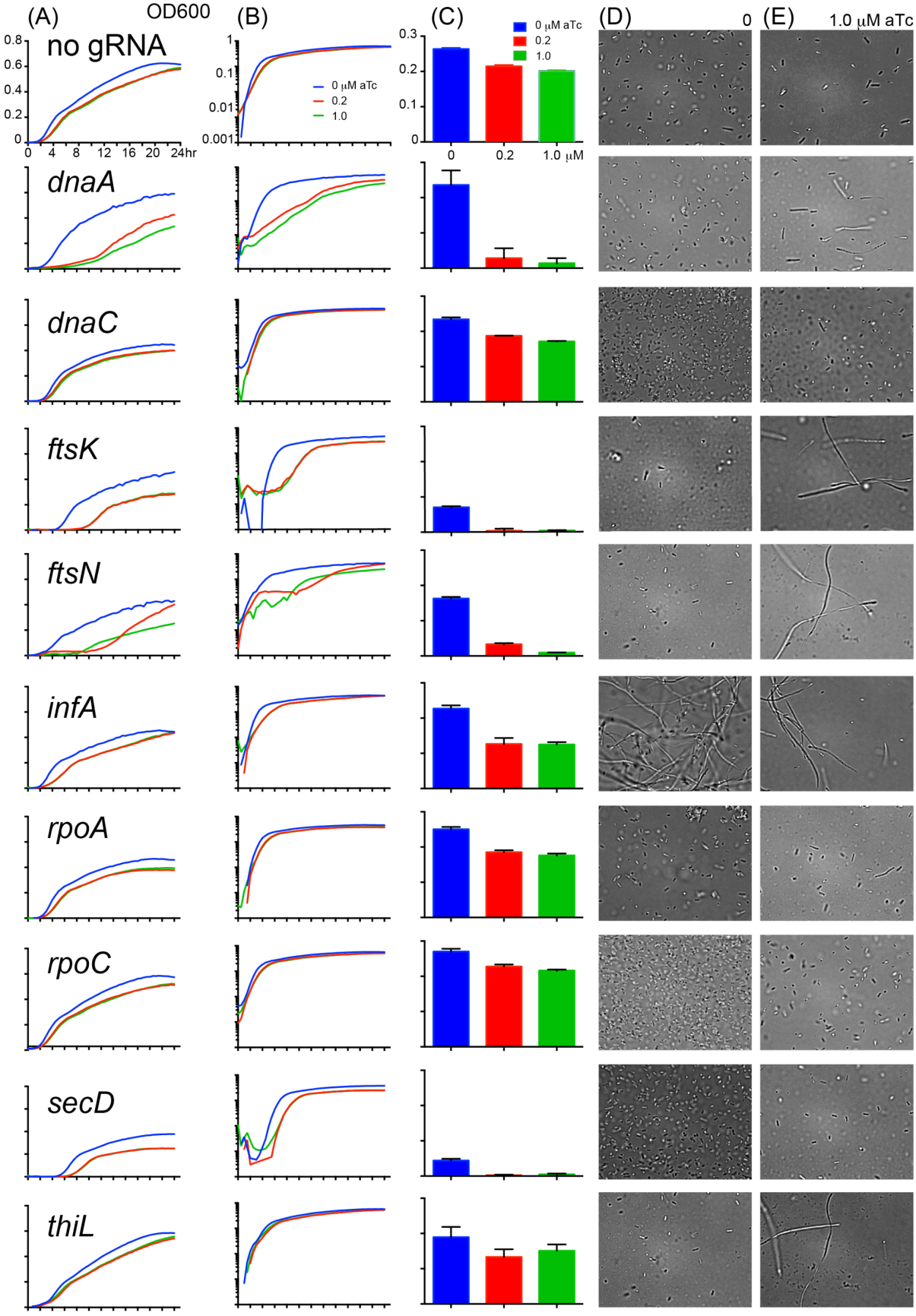
Figure 26： CRISPR-crRNA target loci of the first group essential gene targets. The top line indicates the general principle for crRNA target position design. Red fragment represents the target gene, orange box represents crRNA binding site, grey fragments represent the downstream and upstream genes. The colored fragments represent target essential genes, red and blue represent genes oriented in opposite directions. Grey fragments represent the downstream and upstream genes. orange and green boxes represent crRNA to target non-template strand of each gene.

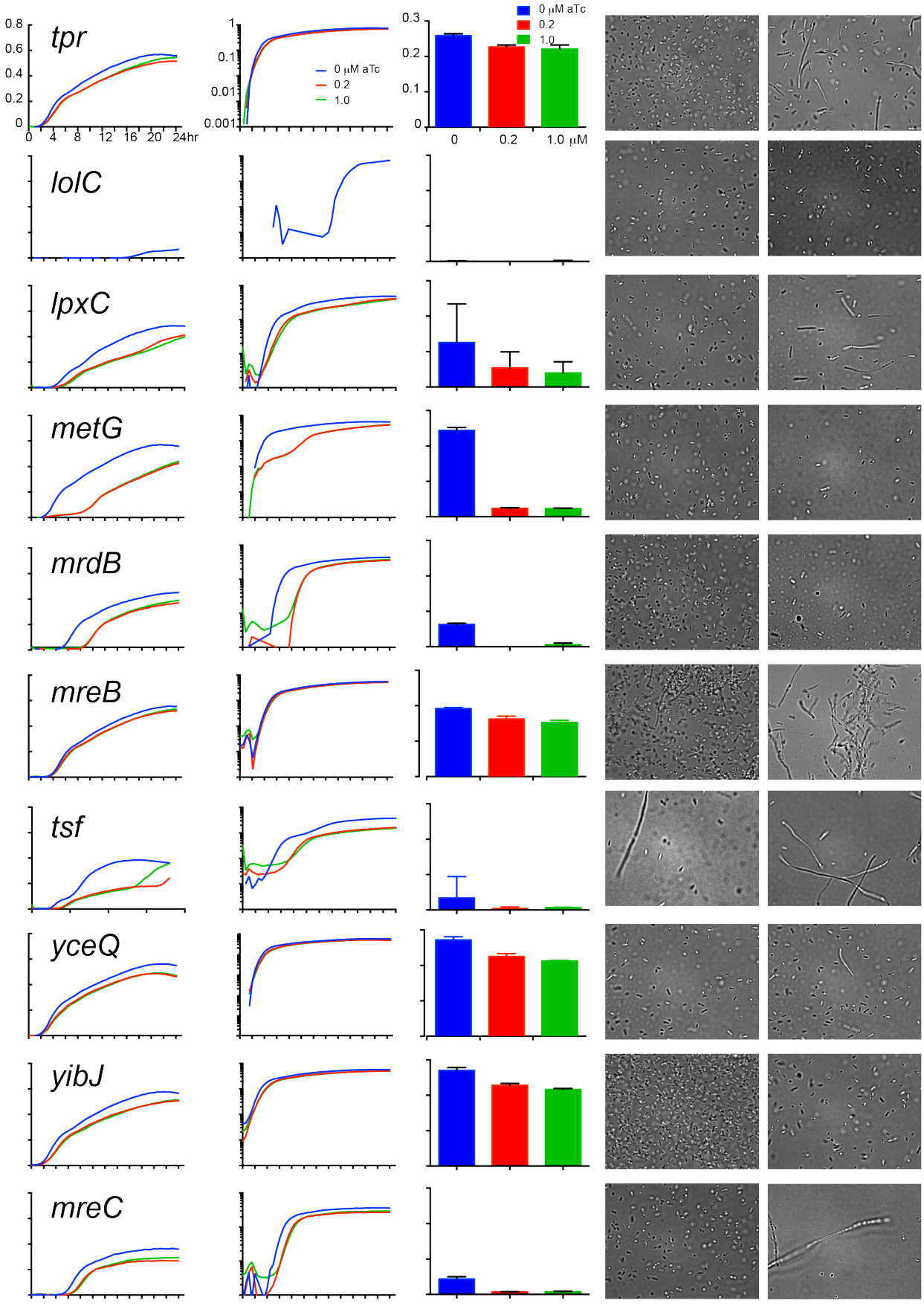
3.4.2 Knockdown efficiency in liquid medium.

To check the growth fitness changes in essential gene repression systems, we monitored the OD600 of essential gene repressed strains in (0μM, 0.2μM, 1μM aTc) by micro-plate reader for 24 hours (Fig.27 -A, Fig.27-B columns). In these targets, *lolC* and *mukB* repression systems presented high sensitivity to this system, they can hardly survive only because of the promoter leaky. The *lolC* repressed strain was not detected grow in the first 15 hours, after 15 hours, the system without aTc shows started to grow, but the strains culture in LB with aTc did not be detected grow until 24 hours. For *mukB*, I did not observe cell growth until 24 hours culture in LB medium without or with aTc at 37°C. Most of the strains show delay of growth (*tsf, murA, secD, mrdB, ftsK, metG, mreC*), some of the genes are sensitive even with a low expression of dCas9 protein by promoter leaky, which mean, without aTc, the target system still showed slight inhibition (like *tsf*), another group genes show no significant be inhibited (*thiL, yceQ, tpr*).

Fig.27-C column presents the log phase growth of the essential gene repressed strains, most of the essential gene repression mutants showed growth inhibition (like *mukB, tsf, murA, secD, mrdB, ftsK, metG, mreC, lolC, rplC, ftsN, lpxC, thiL, infA, dnaA)*, the others showed not obvious growth inhibition (like *dnaC, rpoA, rpoC, tpr, mreB, yibJ, yceQ*).

Fig.27-D and E columns show the morphology changes of the essential gene inhibited strains, the mutants with 1μM aTc (Fig.27-E column) revealed the diverse cell shapes of *E. coli* with CRISPR-dCas9 essential gene repression. The control system has no gRNA target position in *E. coli* genome. *mreC, ftsK, ftsN, infA, thiL, mreB, tsf* and *mreC* presented serious filamentous; *dnaA, rpoA, tpr, lpxC, yceQ* and *murA* showed longer cells compare with the system without aTc (Fig.27-D column). The others showed no obvious morphological changes under the CRISPR-dCas9 repression.


(A)
OD600
no gRNA
0.8
0.6
0.4
0.2
0
0 4 8 12 16 20 24hr
(B)
1
0.1
0.01
0.001
0 μM aTc
0.2
1.0
(C)
0 μM aTc
0.2
1.0
0.3
0.2
0.1
0
0 0.2 1.0 μM
(D) 0
(E) 1.0 μM aTc
dnaA
dnaC
ftsK
ftsN
infA
rpoA
rpoC
secD
thiL


tpr
lolC
lpxC
metG
mrdB
mreB
tsf
yceQ
yibJ
mreC
0 μM aTc
0.2
1.0
0 μM aTc
0.2
1.0
0
0.2
1.0 μM
0 4 8 12 16 20 24hr

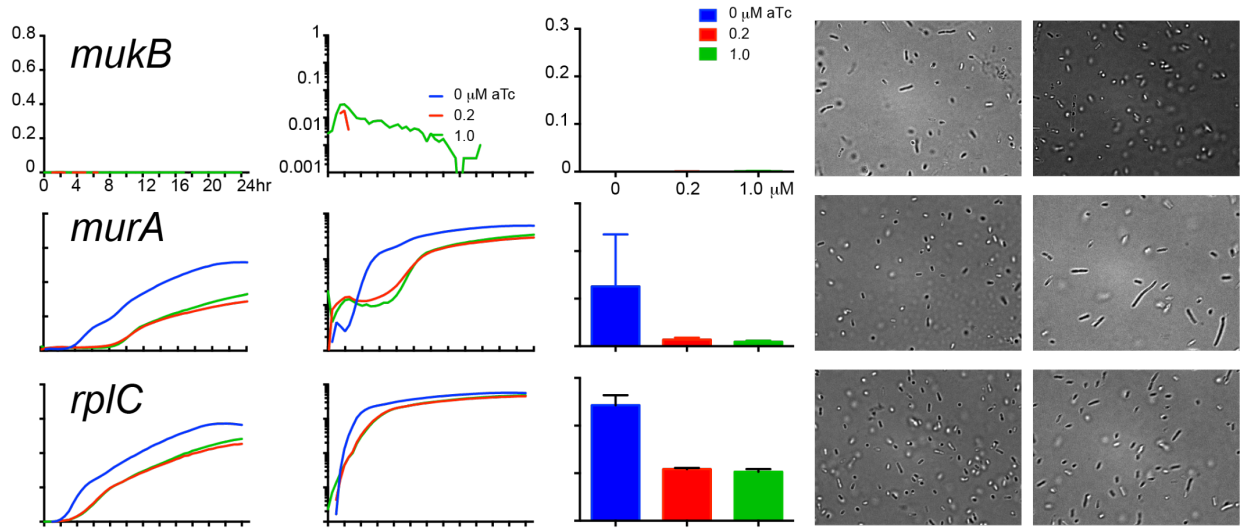


Figure 27: essential gene inhibition in liquid LB. The growth inhibition of essential genes in LB medium with antibiotics (Am, Cm). each row represents one essential gene repression. Column (A) is the growth profiles of essential gene repressed strains, x-axis represents time points (0 to 24 hours). y-axis represents linear OD600 values. Column (B) is the growth profiles of essential gene repressed strains with log OD600 values as y-axis, time points as x-axis. Column (C) shows the log phase OD600 value of essential gene repressed strains. The last two columns reveal the morphology change after 6 hours culture (log phase culture) in LB medium with 0 μM (Column (D)) or 1 μM aTc (Column (E)).

3.4.3 Knockdown efficiency in solid agar plate.

The colony fitness changes of the single essential gene repression system. In this test, Colony-live system was employed to monitor colony growth every 30 minutes. In this test, I want to find a balance that, the inhibited strains present phenotype changes, but it is still viable, i.e. Single essential gene knockdown will not lead bacterial lethal on LB plates. So, I detected the essential gene growth on LB agar plates at 37°C, 24 hours, with aTc in different concentrations. In this test, the essential genes' inhibition was detected in 3 aTc concentrations 0μM, 1μM and 2μM aTc. Each sample did 32 replicates (Fig. 28). The control system is a CRISPR-dCas9 system has no target position on *E. coli* genome. Correlation of maximum growth rates and colony areas indicated that, with aTc inducer, MGR was lower than the system without aTc, and the MGR in higher concentration of aTc (2μM aTc) was lower than in low concentration of aTc (1 μM aTc). *metG* and *rpoC* colonies showed sensitive to the induction of CRISPR-dCas9 system, however, up to 2 μM aTc, no single knockdown strains were lethal.

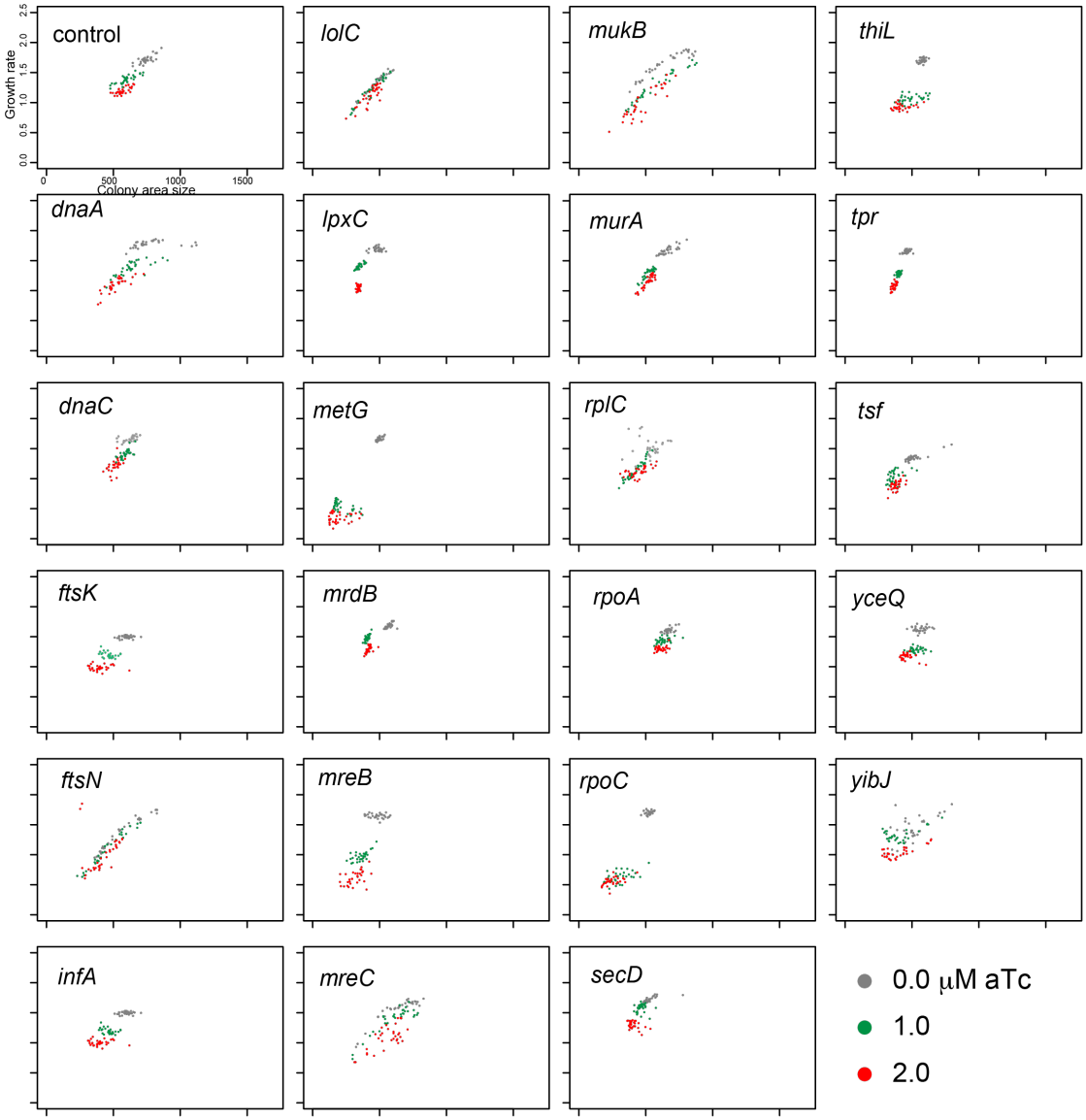
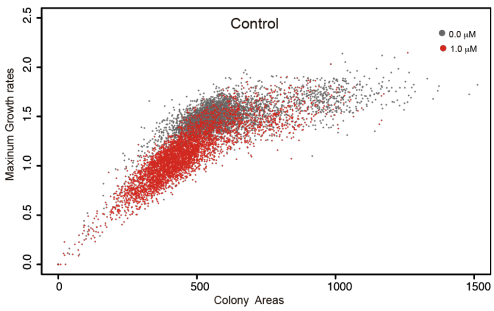


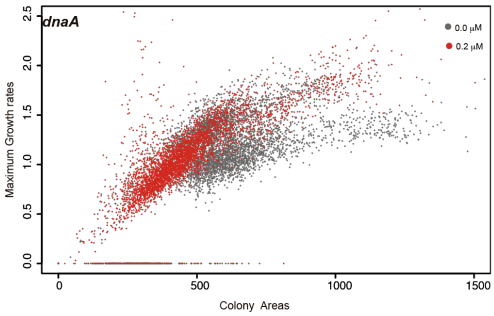
Figure 28: essential gene inhibition on LB agar plates. The essential gene knockdown colony sizes and maximum growth rates. X-axis represents colony size, y-axis represents maximum growth rates (MGR). Control system is an irrelevant CRISPR system with no essential gene target. Each point represents a strain carrying the essential gene single knockdown system, 32 replicates for each target gene are shown on the figures. Dark grey points represent colonies in LB with antibiotics (Am, Cm) medium contain 0μM aTc, green points represent colonies in LB with antibiotics (Am, Cm) medium contain 1μM aTc, red points represent colonies in LB with antibiotics (Am, Cm) medium contain 2μM aTc.

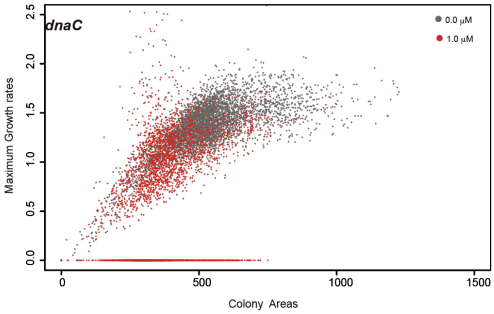
3.4.4 GI analysis of essential gene knockdown and non-essential gene knockout.

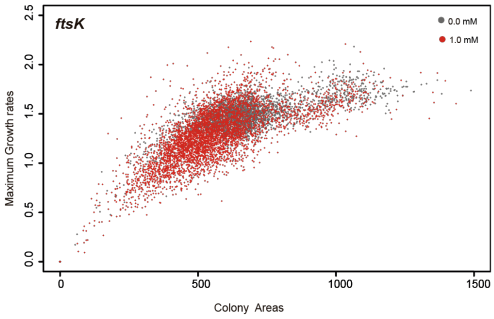
When the essential gene knockdown combines with non-essential gene knockout, some of them showed significant growth deficiency, for examples, *dnaC, lolC, metG, rpoA* and *rpoC* with Keio collection (Fig. 29).

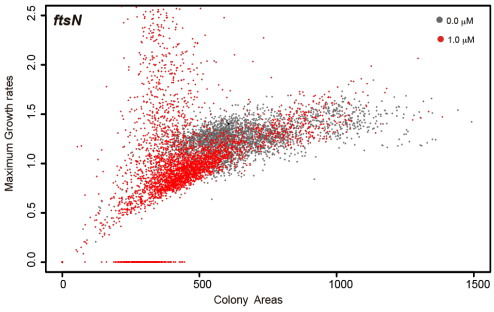
Some of the genes repression have no obvious synthetic lethal or sickness, like *mrdB, mreB, murA, secD, thiL* and *yceQ* (Fig. 29). For *mukB* repression with Keio collection, the mutants turn to higher MGR in the medium with 1μM aTc than in the medium without aTc.

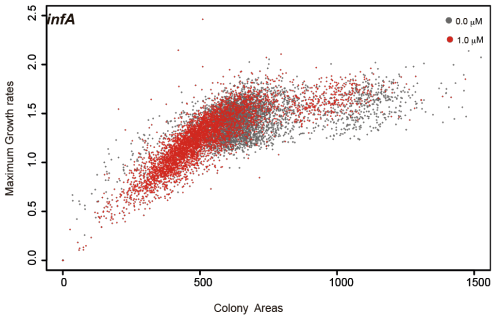

Control
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

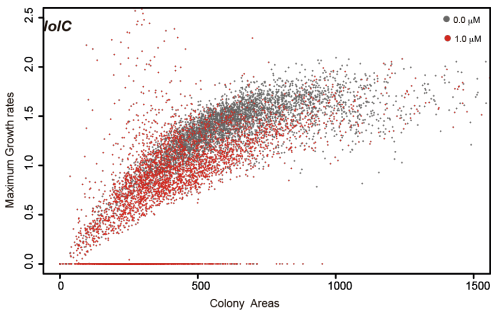

dnaA
0.0 μM
0.2 μM
Maximum Growth rates
Colony Areas

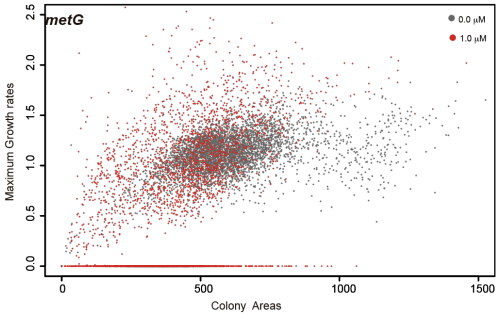

dnaC
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

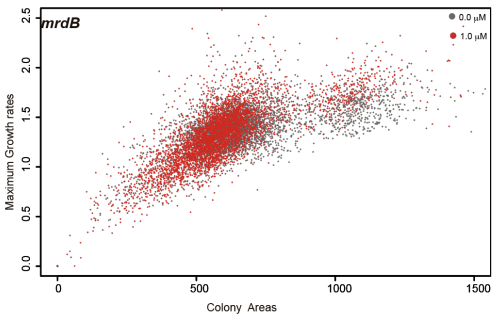

ftsK
0.0 mM
1.0 mM
Maximum Growth rates
Colony Areas

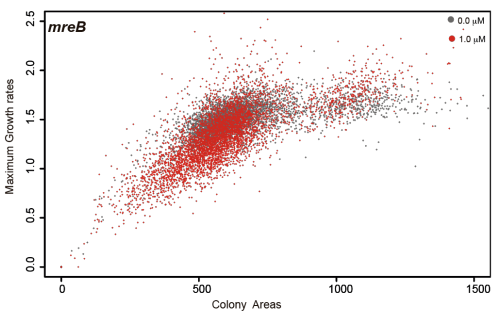

ftsN
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

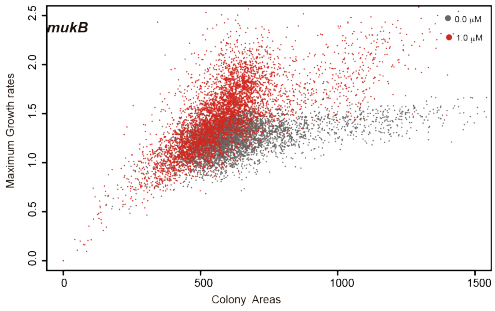

infA
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

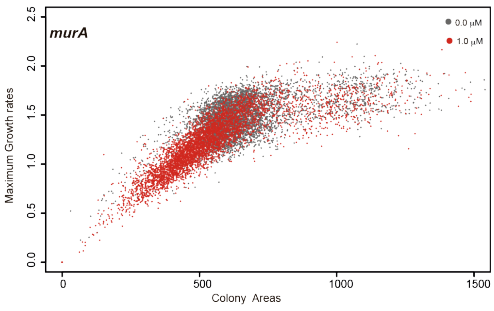

lolC
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

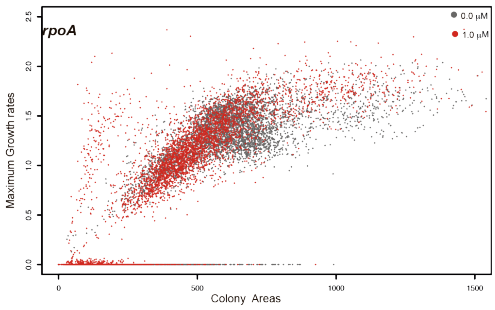

metG
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

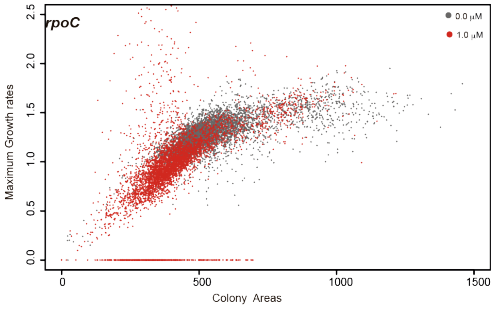

mrdB
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

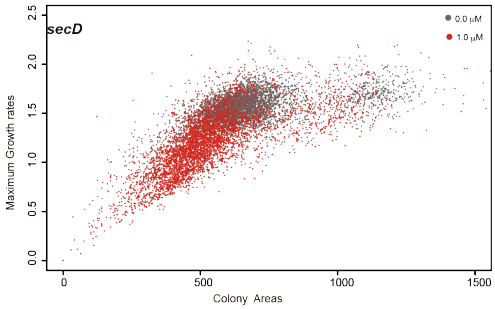

mreB
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

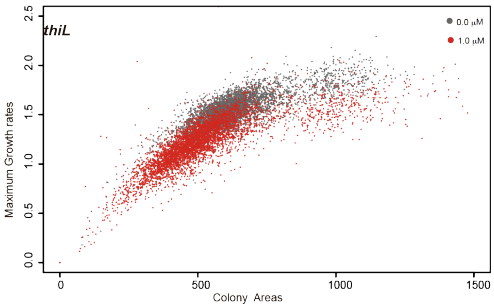

mukB
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

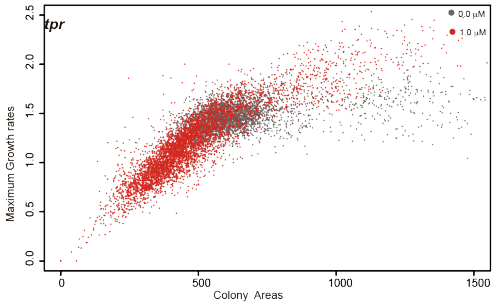

murA
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

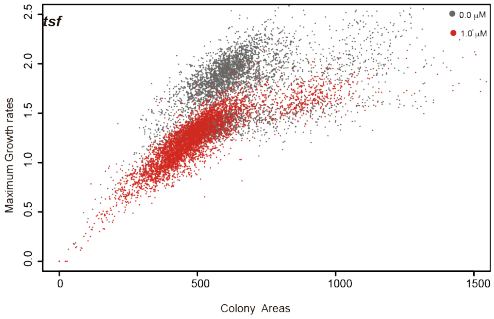

rpoA
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

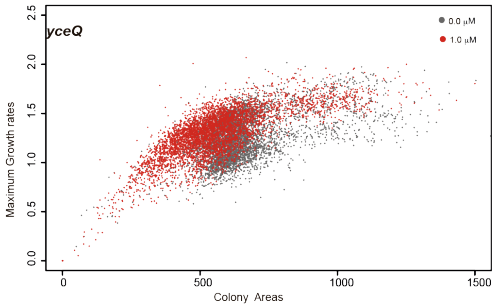

rpoC
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas


secD
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas


thiL
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas


tpr
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas


tsf
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas


yceQ
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

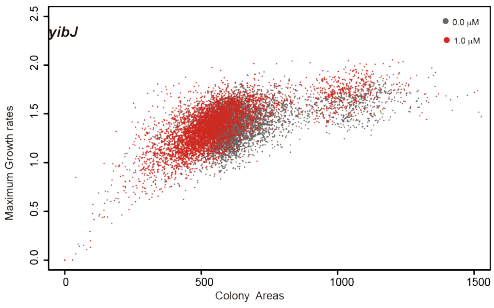

yibJ
0.0 μM
1.0 μM
Maximum Growth rates
Colony Areas

Figure 29: Essential gene knockdown and non-essential gene knockout. The global view of essential gene knockdown and entire non-essential gene deletion combination. X-axis represents colony areas, y –axis represents Maximum growth rates. The Control system contains irrelevant CRISPR system which means has no essential gene target. Dark grey points represent strains grow in LB medium with antibiotics (Km, Cm, Am) and 0μM aTc. Red points represent strains grow in LB medium with antibiotics (Km, Cm, Am) and 1.0μM aTc.

3.4.5 Global view of 19 essential gene GI with Keio collection.

The global view of essential gene knockdown and Keio collection mutants. The heat map indicates the MGR of the essential gene knockdown and non-essential gene knockout mutants. The control system is the CRISPR-dCas9 system with no target position in *E. coli* genome. The right side of x-axis represents low MGR, left side represents high MGR. *rpoC*, *dnaA*, *ftsN*, *metG*, *dnaC*, *lolC*, *rpoA* present strong MGR inhibition combine with Keio collection, some genes repression like *mukB*, *tpr*, *secD*, *ftsK* and *infA* present higher MGR than the control system.

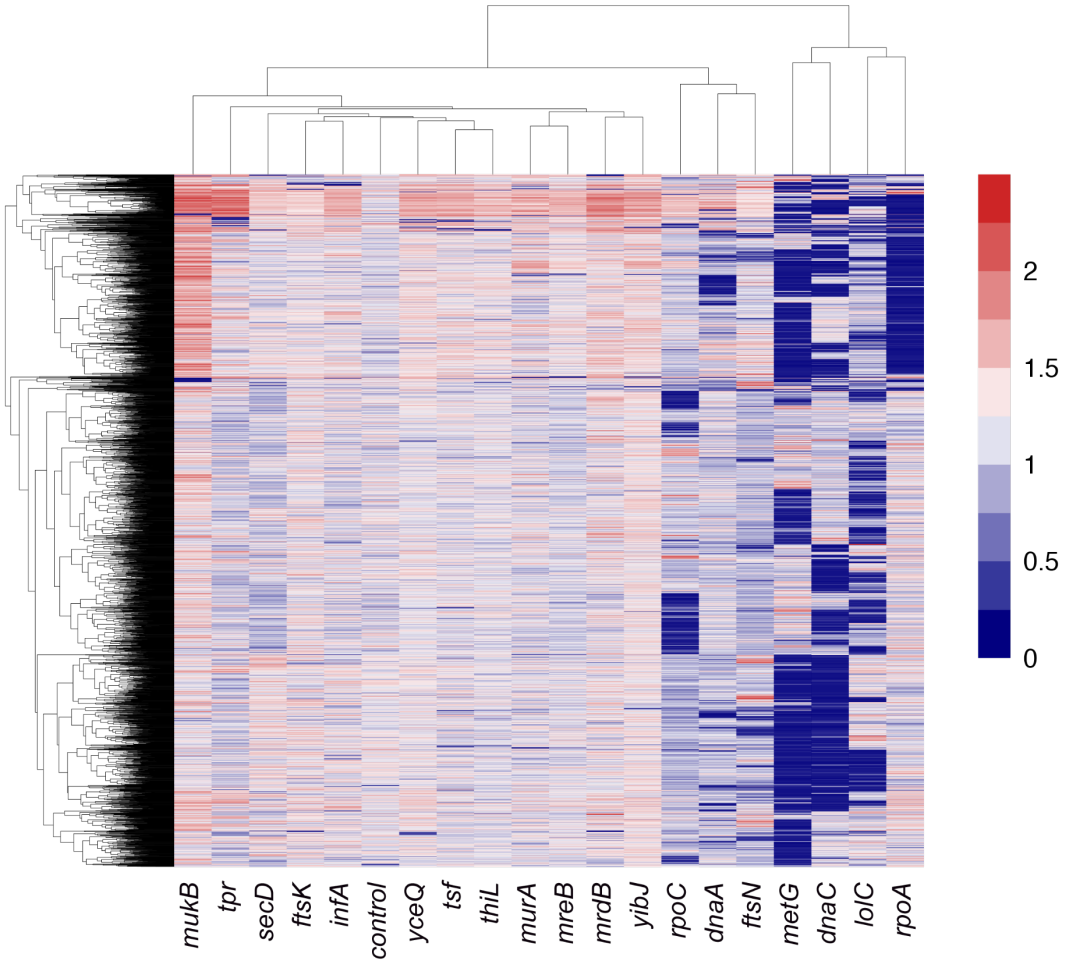


Figure 30: Heat map diagram indicates the MGR of essential gene knockdown and non-essential gene knockout strains. The MGR data was obtained and normalized by Colony-live system [66]. The x-axis represents essential gene targets, and y-axis is the entire *E. coli* genome non-essential gene deletion library.

Chapter 4 Discussion

Three major tools have been developed for genetic engineering (editing) of eukaryotes: (1) Zinc finger proteins (ZNFs), (2) transcription activator-like effector nucleases (TALENs), and (3) CRISPR/Cas9. All provide scientists methods for more precise and faster genome modification and share similar properties as well as differences. All contain nuclease proteins and means for specific targeting DNA sequence motifs. The DNA binding motifs are designed to guide the nuclease to genomic regions containing the specific DNA sequence motif.

ZNFs and TALENs generally employ an endonuclease that makes double-stranded DNA breaks (like FokI, Sp1)[135]. ZNFs contain a recognition module of about thirty amino acid residues, in which each residue interacts with a nucleotide triplet and three to six triplets form a ZNF motif. ZNF motifs have been developed that bind all 64 possible nucleotide triplets. Accordingly, one can construct ZNFs that target a specific DNA motif by combining different zinc finger motifs. While TALENs also use DNA binding motifs to target the endonuclease to specific DNA sequence, individual TALEN domains recognize single nucleotides. This feature makes it easier to engineer TALENs to target a specific DNA sequence. DNA binding domains of TALENs usually contain about 34 highly conserved residues; binding specificity is determined by the 12th and 13th residues within the module[136]. Because ZNFs and TALENs only function as dimers, the genomic region targeted must contain two appropriately spaced DNA sequence motifs, in which one lies upstream and the other downstream of the site targeted for cleavage.

CRISPR/Cas9 system is an alternative gene-editing platform. This system has three components: (1) the endonuclease protein Cas9, (2) a 20-nt RNA sequence crRNA, and (3) a trans-activating crRNA (tracrRNA). CRISPR/Cas9 functions by the tracrRNA and crRNA binding to form an R-loop structure that binds with Cas9 protein and guides it to specific DNA sequence by base-pairing of crRNA with the targeted DNA[102] [10].

Compare to ZNFs and TALENs, CRISPR/Cas9 platform offers several advantages, the first one is, it is easy to design and cheap to express, because CRISPR system recognize the target

positions by base pairing but not protein DNA binding, crRNA is easy to design and clone to an RNA expression plasmid, which makes the experimental costs decrease a lot. Secondly, CRISPR/Cas9 is more efficient since it need transfer only one RNA fragment and one protein, this eliminates the long and laborious processes of transfection and finally selection. Another important advantage of CRISPR/Cas9 system is its ability to target multiple genes simultaneously by transfer multiple crRNAs to host cell simultaneously[137], or cut multiple sites in one gene by transferring several crRNA fragments binding on different locations in one gene. CRISPR/Cas9 system gives a powerful tool for microbe genome editing, especially for gene transcription repression. Since the traditional RNA interference (RNAi) method is unavailable in prokaryotes, this deficiency Cas9 (dCas9) offers a good tool to interfere *E. coli* gene transcription.

However, CRISPR/Cas9 system has some complications or disadvantages. One of the most worrying problems is the off-target effect. Because CRISPR target need only a short fragment about 20nt, some research indicates that designed crRNA would recognize non-specific genome loci partially homology with target sites, this problem requires careful design of crRNA and blast seed crRNA on the host genome to get rid of non-specific binding of cRNA.

The research on essential genes attracts a lot of biologists in the post-genome era, essential genes are required for the survival of an organism, so they are considered as the foundation of lives. Base on the indispensability for bacterial cell survival, essential genes are taken as targets for new drugs by the pharmaceutical industry[138], and some other groups analyze essential genes for knowing more understanding of universal principles of lives[139]. While the emerging field of synthetic biology is looking for essential genes as a way to build and study living minimal cells.

E. coli as one of the most characterized model organism provides us a lot of useful information about cell processes and important pathways of any living organisms; at the same time, *E. coli* research helps us to understand the pathogenesis of harmful pathogens and to develop precise medicine therapy. *E. coli* helped people revealed a lot of cell fundamental knowledge, knowledge gained in the *E. coli* model can often be applied to reveal the mechanism in more complex higher organisms[1]. *E. coli* K-12, although is the most widely studied organism, it still has yet been experimentally characterized. The deeper understanding of the organism will make *E. coli* K-12 a more versatile experimental resource for cellular process or industrial use[2] [140].

Genetic interaction provides a powerful tool to reveal the network of essential genes and the other genes. Combine with high throughput methods, the system analysis of genetic interactions between essential genes and the other genes reveals the cross-interaction of the essential genes and the other essential genes in an organism, and give us the biological information as a global module of proteins or other factors in the organism [141]. In addition, the study of synthetic lethal or sickness gene pairs in the pathways involving essential genes from simple model organisms could reveal gene connections in a more complex organism even there is no sequence conservation of the specific genes between two spices [142].

The previous study about essential genes in *E. coli* utilized a conditional lethal and complementary system. This system consists of essential gene deletion mutants, and a low copy conjugative plasmid which supply the essential gene products under an IPTG inducible promoter. This deletion mutation is a standard strategy to classify essential genes in the genome, the construction of the library may give another resource for the study of gene function for a reductionist biologist, and they are available for systematic analysis or chemical screens. This system provides a strategy to analyze the genetic interactions involving essential genes in *E. coli* K-12 genome. The inducible ability of the complementary plasmids makes it possible to control the essential products expression level and then to examine the genetic interaction in a controllable level. However, one of the shortages of the system is the effect of the essential gene mutant. Probably because some of the essential products are too critical to be removed, the deletion of this essential ORF fragment is not acceptable, even with the supply of the essential products, bacteria still cannot survive[118].

In this project, I developed a novel strategy to systematically analyze the essential gene in *E. coli*. Use a deficiency Cas9 protein, which lost the double DNA strands cleavage ability, and the RNA fragments to repress the target essential genes specifically. The construction of essential gene repression library gives us a resource and method to study essential genes in *E. coli* K-12, different with previous strategy, this resource will give us some clues about essential genes' functions, morphology, and other features. Besides, it also has the ability to perform systematic analysis to detect the communication of the target essential genes with the other genes in *E. coli* k-12 genome since this system can transfer horizontally by bacterial conjugation. The *oriT* fragment on the CRISPR/dCas9 plasmids offers the fast and high throughput ability to systematically detect essential genes repression. In this project, we employed *oriT* from original F plasmid, which is function in gene transfer between *E. coli* to

E. coli. In the future, we could change this *oriT* fragment from other conjugative plasmids to vest the cross-species conjugation ability to the system, like *oriT* from RP4, R388 or R46. These broad host conjugation *oriT* accompany with Hfr strains, will expand the essential gene repression system to the other bacteria or yeast in a high throughput way. This application will give people a novel and direct way to attack the pathogen or reduce the antibiotic resistance of the bacteria by targeting the resistance genes.

From our pioneering experiments, we observed some obstacles of the project. firstly, the leaky of the promoter is one of the most important problems in the project, since the essential genes are critical for cell survive; some of them showed lethal even with a slight repression by the promoter leaky. In this project, we failed to isolate the strain with *ftsZ* essential gene repression system, *ftsZ* is the essential cell division protein-encoding gene related to Z-ring formation[143]. That may be because the extreme essentiality of the gene to cell survival or the repression is too tight to allow the low-level expression of this essential gene. To resolve this problem, we could change another crRNA for this *ftsZ* gene to target this ORF on another position. Probably we can try to design another crRNA binding on the middle region or latter region of this gene to gain a moderate repression. Secondly, there is no standard law for crRNA design to repress a target gene by CRISPR/dCas9 system. Previous reports tried to find some common rules for finding an optimal crRNA for the target gene by detecting several target locations on fluorescence coding genes (*gfp* and *rfp*)[10]. They indicated that, target different strand of the target gene shows different repression levels, for *gfp* and *rfp*, it seems the non-template DNA strand is more stringent as the target, and the closer to 5' end after start codon the tighter down-regulation effect presented. But from our pioneering test, for gene fitness changes by targeting essential genes in *E. coli* genome, for *dnaA*, target template strand is more efficient to inhibit the cell growth, but for *ftsN*, non-template strand is a tighter target location. This result is also supported by some recent published paper, which use this CRISPR/dCas9 system to repression essential genes in other organisms[144] [145]. Based on this result, we designed the crRNA for each essential gene to target both template and on-template DNA strands. At the first batch, we constructed the crRNA to target non-template only, to check the repression effect. Technically, the essential gene repression library is better to contain 2 crRNAs for each gene, this will help the future scientist to gain a proper inhibition of the cell growth, neither too strong nor to weak. Thirdly, in our test, some of the target systems were easy to lose in the mediums without antibiotic selection. We randomly picked several essential gene repressed colonies and isolated on LB medium without antibiotics overnight at 37°C. And picked the

colonies on LB plates streak on LB plates with antibiotics (LB+Am; LB+Cm), some of the strains lost both antibiotic resistance (like *dnaA* repression system), which means in a loose condition, the system trend to remove the plasmids. Some target systems showed less rejection to the CRISPR/dCas9 plasmids, for example in *murB* inhibition strain, after overnight culture on LB plates without any antibiotic, about 50% (26/50) *murB* inhibited strains lost the pdCas9 plasmids but all the strains kept pgRNA target plasmids. This result indicates the system is a kind of stress to the cell, and the bacterial cells have different sensitivity to different essential gene target systems. This situation suggests us that when we perform the experiments using CRISPR/dCas9 system, it is better to culture the strains with repression system in the medium with antibiotics to keep the system present in the cells.

Some other hurdles we need to overcome are 1) verification of the results. Because of the difficulty of analyzing essential genes, there are very few reports about essential genes genetic interactions with the other genes in *E. coli* genome, it is not easy to confirm our final data by comparing them with previous research. To overcome this problem, it is better to do more repeats and analyze the data very carefully. 2) For high throughput experiments, one problem we should concern is the technical problem. The colony-live system captures the plates images every 30 min, and after the fixed time period, detect the colony centers and colony areas, then trace back to fix the positions of the colonies, finally to measure the areas of the colonies, and calculate the maximum growth rates, lag time growth, and saturation point of growth and colony area. In this process, the detection of colony areas is defined to recognize the standard colonies, some abnormal colony morphologies are difficult to precisely detected and calculated, like the mucoid colony, filamentous colony, and irregular colony and so on. This gives us more challenging to quantitative analyze the final data.

In the prior experiments, the synthetic lethal genes of *dnaA* and *ftsN* were detected by the self-transmissible CRISPRi tools. *dnaT* is a functional unclear gene, it is required for primosome formation, primosome is a protein complex which is capable of priming phiX174 DNA replication *in vitro*. *dnaT* is suspected of being involved in the restart of stalled replication forks *in vivo* [146] [147]. DnaT *and* PriA combine together to form the primosome for the rescue of the replication initiation [148], and both of *dnaT* and *priA* are in the *dnaA* synthetic lethal gene list. The result supports the previous hypothesis and gives us the confidence of the new CRISPRi tools for genetic interaction analysis. For quantification analysis, data filtering is always an important step. In this test, when supplied 1μM aTc, the non-target CRISPRi system, which means

CRISPRi system has no repression gene in *E. coli* genome, still has 14 synthetic lethal genes with Keio collection mutants. This may come from the stress of aTc, or CRISPRi vector plasmids. In genetic interaction quantification analysis, we should normalize the data with the non-target system to remove the system vector interference.

mreB and *mreC* genes were reported in the same transcription unit through Northern blot and gene expression analysis. This operon has three genes only, *mreBCD, mreB* located on the upstream of *mreC* [149]. In this test, *mreB* repression didn't obtain the stronger growth inhibition than *mreC* repression. One possibility might be operon effect is not serious in *mreB* and *mreC* growth inhibition; another possibility is may be *mreC* gene has its own promoter, which means it is single operon gene.

Our system could be employed for essential gene single knockdown analysis and essential gene knockdown with non-essential gene knockout research. From my work, the essential gene knockdown library could be used both in liquid culture and on solid agar plates. This ability makes the platform flexible to detect the phenotype changes, drug screen, resistance gene detection or other application or investigation related to essential genes. In the project, I detected essential genes in *E. coli*, but this strategy is not only limited to essential genes, the other non-essential genes repression is also possible by the CRISPR/dCas9 system. Technically, we can repress any gene we want in the genome by designing specific crRNAs binding on the target gene.

As we introduced before, compared to other genome editing tools, CRISPR/dCas9 provides a good way to disturb the multiple genes simultaneously or inhibit one gene on several positions. This capability is especially benefits to pathway analysis. In biosynthesis research field, people usually need to modify several genes in the synthetic pathways in bacteria, my self-transmissible CRISPR/dCas9 system offers a simple and prominent tool to shut down or reduce the expression of pathway genes, and all the process need only several specific short crRNA fragments, the strategy saves a lot of experimental costs and time.

Acknowledgments

I would like to acknowledge my coworkers, family, and associates. First, I am deeply grateful to my supervisor, Dr. Mori Hirotada who has added considerably to my graduate experience. I would like to thank him for teaching me not only how to complete this project, but also how to gain the scientific mind and how to deal with the faced problems in a scientific way. I am sure this will benefit a lot about my perspectives on my future career and life.

I am thankful the assistant professor Ai Muto sensei in Mori lab for managing the serves and high throughput device in a good condition. Thank the laboratory technician Yai Touda san for preparing the experiments and taught me the techniques about the routine experiments in the lab. I also want to show my respects to the previous colleagues for their donation of facility development and biological data accumulation.

This project would not have been possible without my committee members; I would like to acknowledge Maki sensei, and Takagi sensei, for their constructive suggestions and encouragement. My sincere thanks to the staffs in Naist for helping the international students in one way or another.

I will not forget the concerns came from the previous postdoc in Mori lab, Steven Bowden, and Katsushi Yokoyama. Thank them for the technique advice, professional comments, and scientific discussion. Furthermore, I appreciate all the lab members for their generosity in lending a helping hand and friendship; they make my experience of being in Japan and pursuing the Ph.D. unforgettable in my whole life.

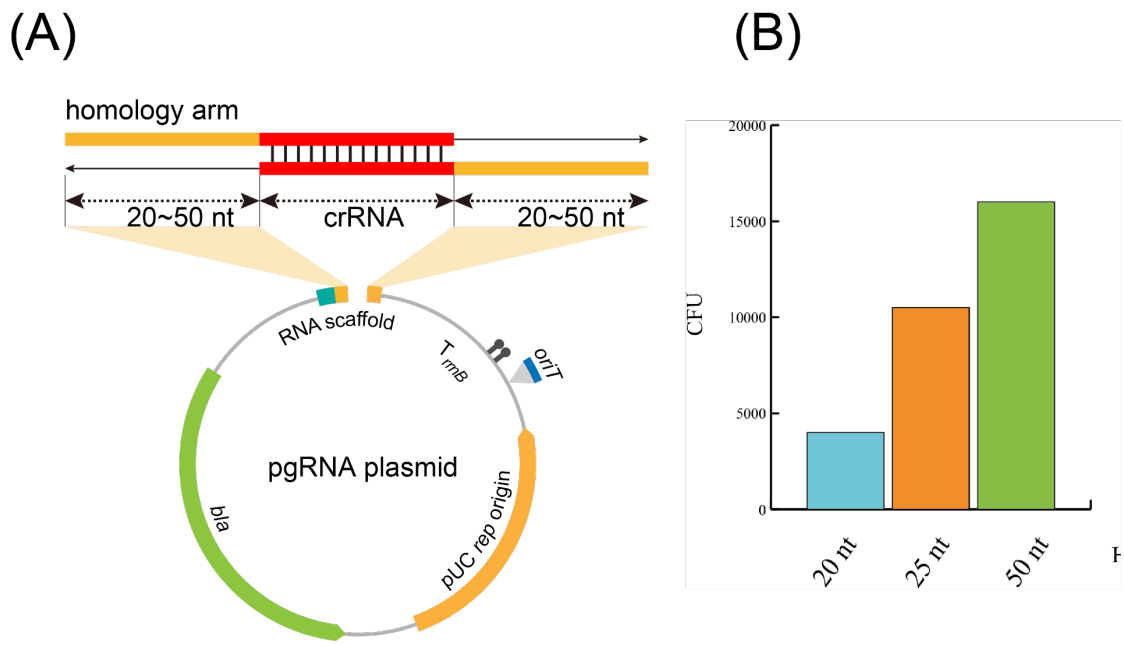
Special thanks to Dr. Barry Wanner for lots of project instruction and consultation.

To all my friends, is my pleasure to know you all, I will remember the laughter and great time we have had together. Last but not least, I would like to express my heartfelt gratitude to my family members for their unconditional love and support.

Supplementary information

Supplementary Fig 1. Homologous arm length and Gibson assembly efficiency.

rpoC gene target crRNA cloning was taken as the example, I designed the gRNA insertion fragment with 20nt, 25nt or 50nt homologous regions (Supplementary Fig 1 A). All the experimental procedures were as before described, the CFU indicated that, up to 50nt, longer homologous region will get more candidates, the ratio of CFU came from 20nt:25nt:50nt homologous lengths is 1:3:4. 20nt homologous overlap regions was sufficient for the purpose to construct specific pgRNA plasmids. The reduce of homologous regions' length from 50nt to 20nt will save the experimental costs, which benefits a lot for high throughput experiments.



Supplementary Fig1: homologous regions for crRNA insertion. (A) the gRNA insertion is done via Gibson assembly; the test detected the different homologous arm (gold) lengths to the efficiency of crRNA (red fragment) insertion into pgRNA vector. 20nt, 25nt and 50nt homologous arms are detected to check the recombination efficiency. (B) the CFU gained from different homologous arms, x-axis indicates homologous arm length. Y-axis represents CFU gained from the related homologous arm.

Supplementary Table 1. CRISPR- dCas9 target crRNA (gRNA) list for all the essential genes in *E. coli.*

Supplementary Table 1: gRNA candidate sequences of all essential genes of E. coli K-12								
ECK_id: if of Escherichia coli K-12 annotation. Riley et al. 2006								
gene_name: gene name								
non_template_strand: gRNA sequence candidate of non-template strand of the target gene								
template_strand: gRNA sequence candidate of template strand of the target gene								
left: left coordinate of BW25113 genome sequence								
right: left coordinate of BW25113 genome sequence								
orientation: orientation of target gene. "+": clockwise, "-": counter clockwise direction								
description: product description								
ECK_id	**gene**	**synonymn**	**non_template_strand**	**template_strand**	**left**	**right**	**orientation**	**description**
ECK0026	***ribF***	***yaaC***	AATTTCGACGGCGTGCATCGCGG	GTCAGCACACACCCTTCTTGCGG	21,407	22,348	-	bifunctional riboflavin kinase/FAD synthetase enamine/imine deaminase;reaction intermediate detoxification
ECK0027	***ileS***	***ilvS***	CTCGCCGTATGTGCCTGGCTGGG	TGCTGCCATTCGCATTAAGGAGGG	22,391	25,207	-	isoleucyl-tRNA synthetase
ECK0028	***lspA***		GATCCTCCAGAACTTTCGTCTGG	AACAGATCGATTGACTCATCAGG	25,207	25,701	-	prolipoprotein signal peptidase (signal peptidase II) Spase II
ECK0030	***ispH***	***lytB,yaaE***	GTCGATAGCTTGCGTGAGCGTGG	GCTGATAGCGCGGTCTACCCCGG	26,277	27,227	-	4-hydroxy-3-methylbut-2-enyl diphosphate reductase 4Fe-4S protein
ECK0032	***dapB***		CCGGTGAGCTGGCCGGAGCCGGG	ACGCCCCCCGGCTCCCGCGATGG	28,374	29,195	-	dihydrodipicolinate reductase
ECK0049	***folA***	***tmrA***	CTCAATGATCAGTCTGATTGCGG	GGCAGGCAGGTTCCACGGCATGG	49,823	50,302	-	dihydrofolate reductase;trimethoprim resistance
ECK0084	***ftsL***	***mraR,yabD***	AAGCAAAGTTAAAGGATCGATGG	CAGGCAATGCATGGCGCTCGTGG	87,519	87,884	-	cell division and growth;membrane protein
ECK0085	***ftsI***	***pbpB,sep***	CCGTTCTGGTCGCCCGTTAGCGG	AGTAATCATGCCGCGGGAGGTGG	87,900	89,666	-	transpeptidase;PBP3;penicillin-binding protein 3 involved in septal peptidoglymay synthesis septum formation penicillin-binding protein 3 peptidoglymay synthetase
ECK0086	***murE***		GACAGCCGTGTGGCTGCGGCGGG	GAAGGTGCGTCTGGCACCCACGG	89,653	91,140	-	UDP-N-acetylmuramoyl-L-alanyl-D-glutamate:meso-diaminopimelate ligase
ECK0087	***murF***	***mra***	ATTCTCAACGGTGAACTGCAAGG	GCAGGTCGATGTCCAGCGGACGG	91,137	92,495	-	UDP-N-acetylmuramoyl-tripeptide;D-alanyl-D-alanine ligase
ECK0088	***mraY***	***murX***	CATTTGCAAAAACTTTCCTTTGG	GATGGCGCGAAACGTCAGATAGG	92,489	93,571	-	phospho-N-acetylmuramoyl-pentapeptide transferase
ECK0089	***murD***		CGAAGCCGTAGAACGCCACACGG	AGGAAAGCCCGGTGAGGCCCAGG	93,574	94,890	-	UDP-N-acetylmuramoyl-L-alanine:D-glutamate ligase
ECK0090	***ftsW***		AGGGCTGGGTGATGGGCTCGCGG	AGGCGCGGCATTTTCAGGCGAGG	94,890	96,134	-	Lipid II flippase;divisome protein recruiting FtsI;SEDS protein
ECK0091	***murG***		GCTGGGGACTGCCGACCGTATGG	TCCCGGGAATACATGTCCACCGG	96,131	97,198	-	N-acetylglucosaminyl transferase;UDP-N-acetylglucosamine:N-acetylmuramyl-(pentapeptide) pyrophosphoryl-undecaprenol N-acetylglucosamine transferase
ECK0092	***murC***		ACGCAGCAGTTAATGAATCTGGG	TTCGGCAATACCGCCCATACCGG	97,252	98,727	-	UDP-N-acetylmuramate:L-alanine ligase;L-alanine adding enzyme
ECK0094	***ftsQ***		GGTGAGCGGCTGGGTCGTGTTGG	CCAGACGCGTTCCATTATTGCGG	99,642	100,472	-	divisome assembly protein;cell division and growth of wall at septum
ECK0095	***ftsA***	***divA***	TCGCGTGGTATGGATAAAGCCGG	TACTAAAGCGGCAACCTTCGCGG	100,469	101,731	-	cell division and septation protein;specific role unknown;recruited to FtsZ ring
ECK0096	***ftsZ***	***sfiB,sulB***	GTTGGACAGACGATTCAAATCGG	TTTACGCAGCGCTTGTGCATCGG	101,792	102,943	-	septal ring GTPase;required for cell division and growth;initiation of septation;tubulin family protein
ECK0097	***lpxC***	***asmB,envA***	ACGGGTGTCGGTTTACATACCGG	CGTCAGGGTGACTTTCTTGCCGG	103,044	103,961	-	UDP-3-O-acyl N-acetylglucosamine deacetylase lipid A biosynthesis
ECK0098	***secM***	***srrA,yacA***	CGGGATGGCAAATAACGTGAGTGG	ACCATCCCTAATAAGAGATGCGG	104,192	104,704	-	regulator of secA translation
ECK0099	***secA***	***azi,pea,prlD***	GAGGCAAGTAAGCGCGTCTTTGG	TCGGAGAGTTTTTCCATCTCCGG	104,766	107,471	-	preprotein translocase subunit ATPase preprotein translocase secretion protein
ECK0103	***coaE***	***yacE***	AGCGCCGGGCCTTGCGCGAGCGG	GCATGTAGCGCAGGTGCACCTGG	109,086	109,706	-	dephospho-CoA kinase;final step in CoA synthesis;predicted DNA repair protein
ECK0125	***can***	***yadF***	TGACCTGAACTGCCTTTCCGTGG	TTCTGCAGGAACGCGACTGTCGG	138,495	139,157	-	predicted carbonic anhydrase;beta class

ECK0414	***dxs***	***yajP***	ATGTCGATTTCCGAAAATGTCGG	CGCCAATGACACAGACGGTGCGG	434,320	436,182	+	1-deoxy-D-xylulose-5-phosphate synthase thiamine-requiring FAD-requiring flavoprotein 1-deoxyxylulose-5-phosphate synthase
ECK0415	***ispA***		GATGATGACGATCTGCGTCGCGG	TATGACCGGTGGCATAAACCAGG	436,207	437,106	+	geranyltranstransferase
ECK0464	***dnaX***	***dnaZ***	CTTTGCTGACGTCGTCGGCCAGG	TTCAACTTTGGTGCGCGAGGCGG	488,097	490,028	-	DNA polymerase III holoenzyme;tau and gamma ATPase subunits;gamma chain (aa 1-431) is main subunit of clamp loader complex
ECK0468	***adk***	***dnaW,plsA***	CTGCTTGGCGCTCCGGGCGCGGG	ATATCGCCAGTGGAGATTTGCGG	493,180	493,824	-	adenylate kinase
ECK0469	***hemH***	***popA,visA***	GACGCGTGGTTGATACCTCACGG	CAGGTTTGCCAGCAGGATACCGG	494,060	495,022	-	ferrochelatase
ECK0517	***lpxH***	***ybbF***	GTTCCCTGTTATTTCATTCATGG	CATCTTGCGATGGAGTGGGTTGG	549,223	549,945	+	UDP-2,3-diacylglucosamine pyrophosphatase
ECK0519	***cysS***		GCGCGCTATCTGCGTTTCCTCGG	GATATGACAGAGATCGTAAACGG	550,616	552,001	-	cysteine--tRNA ligase;binds Zn(II)
ECK0522	***folD***	***ads***	GTTCAACTGCCGTTACCGGCGGG	TTTCCGGGAGGTCATAAGAGCGG	552,880	553,746	+	Methenyltetrahydrofolate dehydrogenase/cyclohydrolase
ECK0575	***entD***		GTTTATGCTTTGCGGGAATATGG	TGTTGCAGTTGTGCGTAGTGCGG	605,464	606,084	+	enterochelin synthase;component D;EntB(ArCP)/EntF-CoA phosphopantetheinyltransferase;facilitates secretion of enterobactin peptide
ECK0627	***mrdB***	***rodA***	GAAGAACACTGGCATCGCGCTGG	CATCAGTGGTACGGCTATTTTGG	661,206	662,318	+	cell wall shape-determining protein;recruits transpeptidase MrdA;SEDS protein+J1896
ECK0628	***mrdA***	***pbpA***	ACGCATGATATCGGTAAGCTGGG	TGGGTCAACGCCGAACCGTAAGG	662,321	664,222	+	transpeptidase involved in peptidoglymay synthesis (penicillin-binding protein 2)
ECK0632	***nadD***	***fusB,ybeN***	AAAGAGTGGCGGCAGGAACAAGG	AAGAGTAAATAATGGCTTGTCGG	665,936	666,577	+	nicotinic acid mononucleotide adenylyltransferase;NAD(P) biosynthesis
ECK0633	***holA***		GACCTGCTGTTGATCGTCCGCGG	GCCGCATTCGGTCCGTTTTCTGG	666,579	667,610	+	DNA polymerase III;delta subunit
ECK0634	***lptE***	***rlpB***	GGACGTTCCATCCTTGCGTTTGG	GCACGGCTTAATGGCCCGTTCGG	667,610	668,191	+	LPS assembly OM complex LptDE;lipoprotein component
ECK0635	***leuS***		GGCAGCAAACTTCGTATTGATGG	CTTCGGCAACTTTGGTGTTACGG	668,206	670,788	+	leucyl-tRNA synthetase
ECK0639	***djlB***	***ybeS***	AGGCGTATGAGGAGGCGCTACGG	CAGTCGTTTCCTCTATATCGAGG	673,420	674,847	-	DnaJ homolog HscC co-chaperone;J domain-containing protein
ECK0649	***lnt***	***cutE***	CAGATTGATGGCCCGTTAAAAGG	GGTCACTTGCCAGAGGGCAGGGG	685,348	686,886	+	apolipoprotein N-acyltransferase
ECK0668	***glnS***		CGATCAAAAACGACGTAGAGTGG	GCACTGGCCTTTATAGTCCTGGG	702,098	703,762	-	glutamyl-tRNA synthetase
ECK0672	***fldA***		CGGCGAAGCGCAGTGTGACTGGG	TGCTGCTTTTTGCAATGTCATGG	706,940	707,470	+	flavodoxin I
ECK0721	***cydA***		TTTGGTATCAACTTCGCTCTGGG	TCAGTGGCACAAAAAGGAAGTGG	767,463	769,031	-	cytochrome d terminal oxidase;subunit I
ECK0875	***infA***	***bypA1***	AGAGTTAGAAAACGGTCACGTGG	CATTTCAATATTGTCTTCTTTGG	922,230	922,448	+	translation initiation factor IF-1
ECK0877	***cydC***	***mdrA,mdrH,surB,ycaB***	CCACCGCTGTTTTATCGTGCGGG	GCCGCCCAGCGTAAAGGCGAGGG	923,479	925,200	+	glutathione;cysteine export permease/ATP-binding protein;ABC superfamily transporter;bifunctional ABC subunits: N-terminal fragment permease domain and C-terminal fragment cytoplasmic ATP-binding domain
ECK0881	***ftsK***	***dinH***	AACCTATCCATAATTTAGGTGGG	CGAAGGGTTAAAGCTTAGTAAGG	929,229	933,218	-	DNA translocase at septal ring sorting daughter chromsomes
ECK0882	***lolA***	***lplA,yzzV***	GCTTAGTAGCAAGCAGCGTTTGG	GCGGCTTTTCAGATCGCTTGCGG	933,377	933,988	-	lipoprotein chaperone
ECK0884	***serS***		TAAGCTGGATGTAGATAAGCTGG	AGTTTTTCAGCGACTGCGTCTGG	935,433	936,725	-	seryl-tRNA synthetase;also charges selenocysteinyl-tRNA with serine
ECK0902	***rpsA***	***ssyF***	CCAGGGCGAGCTGGAAATCCAGG	CACGAACGATAGAACCCGGGCGG	958,000	959,673	-	30S ribosomal subunit protein S1
ECK0905	***msbA***		TATTACCTACGATTCCGAACAGG	TACCACCTTTCCTGATACCCAGG	962,626	964,374	-	bifunctional lipid transport protein subunits of ABC superfamily protein: membrane component/ATP-binding component
ECK0906	***lpxK***	***ycaH***	CGCCTGGCGTGCCCCCGTACCGG	AGGCCATACAACCAGGAGAGTGG	964,371	965,357	-	lipid A 4'kinase
ECK0909	***kdsB***		CTTGAACGCGCGCGTGAATCAGG	AATATCAACCAATGGTTTACCGG	966,857	967,603	-	3-deoxy-manno-octulosonate cytidylyltransferase
ECK0913	***mukF***	***kicB***	AGACCGACTCTCTTTTCTGCTGG	CAACCAGTTCGGGGACTGTCTGG	970,324	971,646	-	chromosome condensin MukBEF;kleisin family protein subunit;binds calcium
ECK0914	***mukE***	***kicA,ycbA***	GAGCTTATTCGCGCACCAGAAGG	AATATGGCGTCCTGAACGTAAGG	971,627	972,331	-	chromosome condensin MukBEF;MukE localization factor
ECK0915	***mukB***		GCTGCATTTCCGTAACACTACGG	GCGTCGTGACCAGCTCGTCAAGG	972,331	976,791	-	chromosome condensin MukBEF;ATPase and DNA-binding subunit
ECK0921	***asnS***	***lcs,tss***	TAACGAGCAGGGATTCTTCTGGG	CACGCAGATACTCAATGCTGTGG	983,590	984,990	+	asparaginyl tRNA synthetase

ECK0141	***folK***		CTTAAACGCAGCCGTGGCGCTGG	GCGGCCCCAGCGGTGGGGTGCGG	153,740	154,219	-	6-Hydroxymethyl-7,8-dihydropterin pyrophosphokinase;monomeric 7,8-dihydro-6-hydroxymethylpterin-pyrophosphokinase
ECK0153	***hemL***	***gsa?,popC***	ACAAAATTATTAAATTTGAAGGG	TGCGCACCATATCCATGGTCGGG	170,089	171,369	-	glutamate-1-semialdehyde aminotransferase (aminomutase)
ECK0155	***erpA***	***yadR***	ATCACCGGTGGCGGTTGCAGCGG	AACTTTGTTGGCTGCTGCGTCGG	173,097	173,441	-	Iron-sulfur cluster insertion protein;A family protein Fe-S protein;essential for respiratory growth
ECK0164	***dapD***	***ssa***	CTTCCGCGTTGTGCCACCAGCGG	TCGGCACTTTGTCGAAGTAGCGG	181,610	182,434	-	2,3,4,5-tetrahydropyridine-2-carboxylate N-succinyltransferase;mutations suppress growth defects of strains lacking superoxide dismutase
ECK0166	***map***	***pepM(S.t.)***	AGTGGTGTGCCACGGTATCCCGG	ATAGCCGTGATAGCCGAGGCAGG	185,199	185,993	-	methionine aminopeptidase
ECK0168	***rpsB***		AAAATGAAGCCGTTCATCTTCGG	TCGGGTTCCAGTAACGGGTCTGG	186,361	187,086	-	30S ribosomal subunit protein S2
ECK0169	***tsf***		GCGTACTGGCGCAGGCATGATGG	CTCACGCAGCTCTTTTACCAGGG	187,344	188,195	-	protein chain elongation factor EF-Ts
ECK0170	***pyrH***	***smbA,umk***	CGGTATTGATGCAAGCATACTGG	AGCAGAATGCGTTTATAGACGGG	188,342	189,067	-	uridylate kinase
ECK0171	***frr***	***rrf***	TCGTGCTTCTCCCAGCCTGCTGG	CCGTGCGTATTTTGCTGATTTGG	189,359	189,916	-	Ribosome recycling factor (RRF);dissociates ribosomes from mRNA after termination of translation;tRNA mimic
ECK0172	***dxr***	***ispC,yaeM***	AGGCAAAAATGTCACTCGCATGG	CGCGGAAGTGTTCGGGATTATGG	190,008	191,204	-	1-deoxy-D-xylulose 5-phosphate (DXP) reductoisomerase;NAPDH-dependent;2-C-methyl-D-erythritol 4-phosphate (MEP) synthase;alternative nonmevalonate (DXP) pathway for terpenoid biosynthesis
ECK0173	***ispU***	***rth,uppS,yaeS***	GTGCCTTTGGGCATAAAGCCGGG	GCTGGCAATTTTTCGCTAAGTGG	191,390	192,151	-	undecaprenyl pyrophosphate synthase
ECK0174	***cdsA***		GTTAATACCCGTCGTCATCGCGG	GTTACAATGGCGAACCCCACCGG	192,164	193,021	-	CDP-diglyceride synthase;integral membrane protein with eight transmembrane helices;also known as phosphatidate cytidylyltransferase
ECK0175	***rseP***	***ecfE,yaeL***	GAGCGTTTCTCAATAGGGTTTGG	GAAATGACCAAATTCATGCACGG	193,033	194,385	-	inner membrane zinc RIP metalloprotease;RpoE activator;by degrading RseA
ECK0176	***bamA***	***ecfK,omp85(N.m.),yaeT,yzzN,yzzY***	GTGAAAGATATTCATTTCGAAGG	CCCTTCAGCACCGTATACGGTGG	194,415	196,847	-	OM (outer membrane) protein required for OM biogenesis;compnent of BamABCDE complex
ECK0178	***lpxD***	***firA,omsA,ssc***	CTACCTGACTTACGCGCGCATGG	CGCGAAAGGAAGATCGTCCTGGG	197,458	198,483	-	UDP-3-O-(3-hydroxymyristoyl)-glucosamine N-acyltransferase
ECK0179	***fabZ***	***sabA,sefA,sfhC,yaeA***	GCACCGTTTCCCGTTCTTACTGG	CCACCAGTAAGAACGGGAAACGG	198,588	199,043	-	3R-hydroxymyristoyl acyl carrier protein (ACP) dehydratase
ECK0180	***lpxA***		TCCAACCGCCATTGTGGAAGAGG	TGACGCGCCCTCTTCCACAATGG	199,047	199,835	-	UDP-N-acetylglucosamine acetyltransferase
ECK0181	***lpxB***	***pgsB***	ATGGGCATTGTTGAAGTGCTCGG	GCACATGTTCTTTCAGAGCGCGG	199,835	200,983	-	tetraacyldisaccharide-1-P synthase
ECK0183	***dnaE***	***polC,sdgC***	TGTGGTCTGGTGAAGTTCTACGG	CCAACGGTGCGGTTTTGGCCAGG	201,613	205,095	-	DNA polymerase III;alpha subunit;suppressor of dnaG-Ts
ECK0184	***accA***		TCTGACTGCGGTTAGCCGTCAGG	TTCGCTTCCAGCTCTGCAATCGG	205,108	206,067	-	acetyl-CoA carboxylase;carboxytransferase;alpha subunit
ECK0187	***tilS***	***mesJ,yaeN***	TCAGTTGGTGCAGTGGCGGACGG	GCAATGCGTAACCCAGGCATCGG	208,818	210,116	-	tRNA(Ile)-lysidine synthetase
ECK0194	***proS***	***drpA***	AACCTACGATGCAATGTATGCGG	CTTCGTCGCGGAACTTGGTCTGG	213,544	215,262	+	prolyl-tRNA synthetase
ECK0219	***yafF***		GACGACTACAAAATAAGAAGAGG	TGCTGTGGCGAACTTCTCTGCGG	235,593	235,865	-	pseudogene;H repeat-associated protein
ECK0271	***yagG***		ACACCGCCTGCGGCTTCGTCTGG	GTAGCCAATTTTGTCTTTCATGG	281,106	282,488	-	CP4-6 cryptic prophage predicted sugar transport protein
ECK0366	***hemB***	***ncf***	TTCTGTGAATACACTTCTCACGG	CGCTGCCGGTTTCATCGGTATGG	384,758	385,732	+	5-aminolevulinate dehydratase (porphobilinogen synthase)
ECK0402	***secD***		GAAGTACGTCATGCTGATCGTGG	CTGGATCAGCGTTTGCTCACTGG	423,652	425,499	-	SecYEG protein translocase auxillary subunit protein secretion membrane protein part of channel
ECK0403	***secF***		GCGCGGCTTTAACTGGGGGCTGG	AGTCATAGACTTTACGGCCGTGG	425,510	426,481	-	SecYEG protein translocase auxillary subunit protein secretion membrane protein
ECK0408	***ribD***	***ribG,ybaE***	CGCATCCCAACCCGAATGTCGGG	CATGTGGTTCACCCGCACGTTGG	429,460	430,563	-	fused diaminohydroxyphosphoribosylaminopyrimidine deaminase and 5-amino-6-(5-phosphoribosylamino) uracil reductase
ECK0409	***ribE***	***ribH,ybaF***	TGAAGCTAACGTTGCTACCCCGG	GATGGCGACGCGAGCGTCCGGGG	430,652	431,122	-	riboflavin synthase beta chain
ECK0411	***thiL***		TCTCGTCTTGATGTCGAACTGGG	ATCGCCAGGGTCTGTTTCTCGGG	431,639	432,616	-	thiamin-monophosphate kinase

ECK0945	***fabA***		TGGATATCAATCCGGATCTGTGG	CGGTCCATCATCAGCATGTTCGG	1,011,957	1,012,475	+	beta-hydroxydecanoyl thioester dehydrase
ECK1054	***murJ***	***mviN(S.t.),mviS(S.t.),yceN***	GCAATTGTCGCCAGAATCTTTGG	GCCAAGCACACGCGAAAACATGG	1,123,844	1,125,379	-	predicted peptidoglycan lipid II flippase;required for murein synthesis;predicted MviN family transcription factor with helix-turn-helix motif
ECK1069	***rne***	***ams,hmp1,smbB***	CTGCGCCTGCGTGACCTCGGCGG	GCCGCCGCGGGTCGCGCGTGCGG	1,137,187	1,140,372	+	bifunctional ribonuclease E;RNase E;component of RNA degradosome;mRNA turnover;5S and 16S RNA maturation;endoribonuclease/RNA-binding protein/RNA degradosome binding protein
ECK1070	***yceQ***		CTCAAAGAGGCGCAGAGTGTCGG	CCGGTAGATGGAAATATTTCTGG	1,140,507	1,140,827	-	predicted protein
ECK1078	***fabD***	***tfpA***	GCTGAAGCTTCTGCGGCGCTGGG	ACGGTTTGAGAACCCTGTCCAGG	1,145,733	1,146,662	-	malonyl-CoA-acyl carrier protein transacylase
ECK1079	***fabG***		GCTGAAACGCTCGCAGCCCGTGG	GCCAATTCCGCGGCTTGCACCGG	1,146,675	1,147,409	-	3-oxoacyl-[acyl-carrier-protein] reductase
ECK1080	***acpP***		GAACGCGTTAAGAAAATTATCGG	ATCAAAACTCGCTTTCGCGATGG	1,147,620	1,147,856	-	acyl carrier protein ACP
ECK1084	***tmk***	***ycfG***	AACTACCGCGCGTAATGTGGTGG	TCGGCAAGTTGCGTACCGCCAGG	1,151,129	1,151,770	-	thymidylate kinase
ECK1085	***holB***		GCATCCCGATTACTACACCCTGG	CCGCAACTTTTGTGGCCCTGCGG	1,151,767	1,152,771	-	DNA polymerase III;delta prime subunit
ECK1102	***lolC***	***ycfU***	GCTTCGGTCGTTTCGTCTCCTGG	CAATAAATAGAGCGACAGGTTGG	1,171,432	1,172,631	-	lipoprotein-releasing system transmembrane protein
ECK1103	***lolD***	***ycfV***	ATGGCGATCGTCGGTAGCTCTGG	CTTCCTGATAGCGTTTGCACAGG	1,172,624	1,173,325	-	OM (outer membrane)-specific lipoprotein transport protein subunit
ECK1104	***lolE***	***ycfW***	TCTACCATTGGCATTGCCCTCGG	GCCGTCCGCGACTAAAACGCAGG	1,173,325	1,174,569	-	lipoprotein-releasing system transmembrane protein
ECK1117	***purB***	***ade(h)***	ATGAAGTGATCCTGCCATACTGG	ATTGATATCTTCCGAAGTACAGG	1,186,621	1,187,991	+	adenylosuccinate lyase;purine synthesis
ECK1131	***cohE***	***ymfK***	ATTAGTCCCTGTCATATCATGGG	GATCGAGCCATCCTACAGGCAGG	1,198,264	1,198,938	+	e14 prophage;predicted CI family protein transcription regulator (repressor)
ECK1162	***minE***	***minB***	GTTGCGTAAAGATATTCTTGAGG	CAGCCGTTCTTTTGCAATGTTGG	1,220,284	1,220,550	+	cell division topological specificity factor
ECK1163	***minD***	***minB***	ATTTATCGTTTGTGACTCCCCGG	TTATCGCGTGTTTGCGATGCCGG	1,220,554	1,221,366	+	inhibitor of FtsZ ring polymerization;chromosome-membrane tethering protein;membrane ATPase of MinCDEE system
ECK1192	***pth***	***asuA?,rap***	CATTAATCCGGACGAAATTCTGG	TAAATGTAGTCGGGACTAACAGG	1,253,934	1,254,518	+	peptidyl-tRNA hydrolase
ECK1195	***prs***	***dnaR,prsA***	CTACTTTGGCTATGCGCGCCAGG	AACAGCGGTGATACGACCTGCGG	1,256,933	1,257,880	+	phosphoribosylpyrophosphate synthase
ECK1196	***ispE***	***ipk,ychB***	GACAAGCGTTTGCCGATGGGCGG	TCATGTTCCACGCCTTCAACGGG	1,258,031	1,258,882	+	4-diphosphocytidyl-2-C-methylerythritol kinase
ECK1197	***lolB***	***hemM,ychC***	TTTCTGGCAGCAAACCGGCCAGG	GACGCCATTGTGGCGAATCCGGG	1,258,882	1,259,505	+	OM (outer membrane) lipoprotein;required for localization of lipoproteins
ECK1198	***hemA***	***gtrA***	CAGCCTGCTTGCGCAGCCGATGG	GTTGATACCGAGTGCTAAAAGGG	1,259,719	1,260,975	-	glutamyl tRNA reductase
ECK1199	***prfA***	***sueB,uar,ups?***	CGCGCAAACTATCGCCGACCAGG	TCCAGTTTGGCAACGATAGAAGG	1,261,017	1,262,099	-	translation termination peptide chain release factor 1;RF-1;recognizes UAG and UAA
ECK1200	***prmC***	***hemK***	CGTACTTTTATTCTCGCCTTTGG	GGCTTTCGCTCGCCTGAAGTTGG	1,262,099	1,262,932	-	N5-glutamine methyltransferase;modifies release factors RF-1 and RF-2
ECK1203	***kdsA***		AAACAAAAAGTGGTTAGCATTGG	AAGAGGCTTTGAACACGTAAGGG	1,264,170	1,265,024	-	3-deoxy-D-manno-octulosonate 8-phosphate synthase
ECK1224	***tpr***		TGCCGCAGGCAACGACCCGAAGG	CTGGCGCTCGAGTCGAACCTTGG	1,283,092	1,283,181	+	protamine family protein
ECK1268	***topA***	***asuA?,supX***	GGTTGACCCGTGGCACAATTGGG	AGCCACTGGTCGGCAAATCGCGG	1,325,854	1,328,451	-	DNA topoisomerase I;omega subunit
ECK1272	***ribA***		TTCAGCTTGCGCTGCGATTGTGG	CACCGGTCAGACATTCGGAATGG	1,333,376	1,333,966	+	GTP cyclohydrolase II
ECK1283	***fabI***	***envM,gts,qmeA***	AAAGTTTGGCCGAAATTTGACGG	GATGTCAGAACCCAATTGAGCGG	1,345,057	1,345,845	+	enoyl-[acyl-carrier-protein] reductase;NADH-dependent
ECK1354	***racR***	***cohR,ydaR***	CGTTTCTTTTCTGATGTTGTTGG	GACCTCCGCCTTTGATTTGACGG	1,414,571	1,415,047	+	Rac cryptic prophage;predicted DNA-binding YdaR family transcription regulator
ECK1564	***dicA***	***ftsT***	CCATGTGTCTGTATCACAATGGG	GGCTTTAGCAAGAGACCTTTGGG	1,642,740	1,643,147	-	Qin prophage;predicted DNA-binding transcription regulator (repressor) of dicB
ECK1566	***ydfB***		TATGGATTTCGACACAATCATGG	GAGCTTCTTCGCCTTCGGCAAGG	1,643,471	1,643,599	-	Qin prophage;predicted expressed protein
ECK1633	***tyrS***		GCGAACAACTATGACTGGTTCGG	CCACAGTCGAAATCGAGGAACGG	1,710,754	1,712,028	+	Tyrosine--tRNA ligase
ECK1658	***ribC***	***ribE***	ACGTTACGCATTACCAATCTTGG	GTTATGCGCCACGGAAGCACCGG	1,737,407	1,738,048	+	riboflavin synthase;alpha subunit
ECK1687	***ydiL***		GATGAACGCTTATGAACTCCAGG	CAAAAATATGGCGGAGTGCCTGG	1,765,421	1,765,777	-	predicted HTH domain DNA-binding protein
ECK1711	***pheT***		AAGAAAAACTGCGTCGTTGCGGG	CGTTAATGCCTTTTACCACACGG	1,790,363	1,792,750	+	phenylalanine tRNA synthetase;beta subunit
ECK1712	***pheS***	***phe-act***	GACCGTATCGAAAGTTTCTTCGG	CCGTTTTCAATGCGACGACCTGG	1,792,765	1,793,748	+	phenylalanine tRNA synthetase;alpha subunit
ECK1714	***rplT***	***pdzA***	AACGCAGCAGCACGTCAGAACGG	AGCTTTGATAACAGCCTGGAAGG	1,794,199	1,794,555	+	50S ribosomal subunit protein L20

ECK1716	***infC***	***fit,srjA***	GCCGCCGGTTTGTCGTATAATGG	GACCTGTTAAGCGAACTTCCTGG	1,794,902	1,795,444	+	protein chain initiation factor IF-3
ECK1717	***thrS***		AAAATGTTGCAACGTATTTACGG	TGTTCGGTACGTGCGGACCGCGG	1,795,448	1,797,376	+	threonyl-tRNA synthetase
ECK1738	***nadE***	***efg,ntrL***	TCACTGGTGCTCGGGATCAGCGG	TCCTCTTCAGCATTAATCTGCGG	1,817,264	1,818,091	-	NAD synthetase;NH3/glutamine-dependent
ECK1777	***gapA***		CTGAAATATGACTCCACTCACGG	GGAAAACAATGCGACCGATACGG	1,857,577	1,858,572	-	glyceraldehyde 3-phosphate dehydrogenase A
ECK1805	***tsaB***	***rpf(S.t.),yeaZ***	AGCTTTACTGGCGTGCGCATTGG	TTCGTTGAGTATGTTCACGAGGG	1,885,378	1,886,073	+	tRNA(NNU) t(6) A37 threonylcarbamoyladenosine modification;binding partner and protease for TsaD
ECK1867	***aspS***	***tls***	GTGCCTTCTCGTGTGCACAAAGG	GCATCGGAGTTTCGATGTCGAGG	1,943,556	1,945,328	+	aspartyl-tRNA synthetase
ECK1877	***argS***	***lov***	GATTGCGGCAGGCGCGCCTGCGG	CGCAGAGTAGTCAACCACAATGG	1,954,868	1,956,601	-	arginine-tRNA synthetase
ECK1911	***pgsA***		TGGTGCTTTCCTTGACCCTGTGG	AAAATGAGCGCGGCGGCAAACGG	1,986,299	1,986,847	+	phosphatidylglycerophosphate synthetase
ECK2012	***yefM***		ACTCGAAGAATACAACTCGCTGG	CTGACGAGTAATAAGGATCGGGG	2,083,492	2,083,743	+	antitoxin of YoeB-YefM toxin-antitoxin system;antitoxin component of a toxin-antitoxin (TA) module;antitoxin that counteracts the effect of YoeB toxin;YefM binds to the promoter region of yefM-yeoB operon to repress transcription;YeoB acts as a corepressor
ECK2107	***metG***		GCTAACGGCTCAATCCACCTCGG	TTCGCGACTTGAGTCATAGTAGG	2,188,328	2,190,361	-	methionyl-tRNA synthetase
ECK2146	***folE***		TATGTCGATGAAATTTTCTCCGG	CGGTCATATGACCAGCAATAAGG	2,237,012	2,237,680	+	GTP cyclohydrolase I
ECK2179	***rplY***		CGCAGAAGTACGTAAAGAGCAGG	ACTTGTTAGCGGCACGCAGGCGG	2,276,545	2,276,829	-	50S ribosomal subunit protein L25
ECK2182	***yejM***	***yejN***	GAGGTTTTTGTCCGTCATTCTGG	ATGGCCGATAATGCTTACGTAGG	2,278,404	2,280,164	-	predicted hydrolase;inner membrane
ECK2223	***gyrA***	***hisW,nalA,nfxA,norA,parD***	CGTCCACGAAATTCCGTATCAGG	GATATACACCTTGCCGCGACCGG	2,330,821	2,333,448	+	DNA gyrase (type II topoisomerase);subunit A
ECK2226	***nrdA***	***dnaF***	CGACAAAATCCATCGCGTTCTGG	CCGGCGCATCACGGGAGATCAGG	2,338,893	2,341,178	-	ribonucleoside-diphosphate reductase 1;alpha subunit
ECK2227	***nrdB***	***ftsB***	CTCAAAGAACCGATGTTCTTTGG	TCGTAGCGAGCCACGTTGACCGG	2,341,412	2,342,542	-	ribonucleoside-diphosphate reductase 1;beta subunit;ferritin family protein
ECK2309	***folC***	***dedC***	ATTCTGGAAGTAGGGCTGGGCGG	CGATTCAATCTCCGCAAAAGAGG	2,425,702	2,426,970	+	dihydrofolate:folylpolyglutamate synthase
ECK2310	***accD***	***dedB,usg***	AGGCACTCTGTATGGAATGCCGG	CAGACGGTCTTTATACTTCTTGG	2,427,040	2,427,954	+	acetyl-CoA carboxylase;beta (carboxyltransferase) subunit
ECK2317	***fabB***	***fabC***	GACGCAATGCGCGGCCCGCGCGG	GGGTTATTCTGGTAAGCTTCCGG	2,434,413	2,435,633	+	3-oxoacyl-[acyl-carrier-protein] synthase I
ECK2394	***gltX***		GTTTTTGACGATCAGATCCGTGG	CAGCATGATGCTCATGGCTGTGG	2,513,285	2,514,700	+	glutamyl-tRNA synthetase
ECK2406	***ligA***	***dnaL,lig,lop,pdeC***	GAGCTGCCGGATACTCATCTTGG	GGTCGCTTCGCTGTAATACGCGG	2,522,069	2,524,084	+	DNA ligase;NAD(+)-dependent
ECK2407	***zipA***		GTCTGCGCAGCCGCGTCAACCGG	GGCGGTTGGTACTGGTGTTGCGG	2,524,155	2,525,141	+	cell division protein involved in Z ring assembly;FtsZ stabilizer
ECK2467	***dapE***	***msgB***	GTGAAACGTTAGCCTTTGCCGGG	TCCTGCATCATCAGGACTCAGGG	2,585,515	2,586,642	-	N-succinyl-diaminopimelate desuccinylase;DAP/lysine biosynthesis;contains Zn(2+)/Co(2+)
ECK2474	***dapA***		TACAATCGTCCGTCGCAAGAAGG	AGTAGCGTTAGCGCCGGTCCCGG	2,592,790	2,593,668	+	dihydrodipicolinate synthase
ECK2507	***der***	***engA,yfgK***	ATCGCCGCGTCTCACGGTCGTGG	CCAGGAAGGTCGGTTTTTCACGG	2,629,792	2,631,264	+	multicopy suppressor of ftsJ;GTPase;ribosome biogenesis;depleted cells form filaments with defective chromosome segregation;Der-YhiI complex
ECK2510	***hisS***		TGCGAAGTTTTCGGTCTGCAAGG	GGACGCTCGTGACGGAACATCGG	2,633,209	2,634,483	+	histidyl tRNA synthetase
ECK2511	***ispG***	***gcpE***	AAAAACATTCCGATCCGTATTGG	GCGCAATGCGATAGTCGAAGTGG	2,634,594	2,635,712	+	1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate synthase;alternative nonmevalonate (DXP) pathway for terpenoid biosynthesis;[4Fe-4S] protein
ECK2530	***suhB***	***ssyA***	TGCCAAAAACTATGAAACCCCGG	TTCGGCAGCTTTATCTACGTTGG	2,657,350	2,658,153	-	inositol monophosphatase
ECK2557	***tadA***	***yfhC***	ATGCAGAAATCATGGCCCTGCGG	TGCACTAATACCGCGCCGACCGG	2,691,262	2,691,765	+	tRNA-specific adenosine deaminase
ECK2561	***ohsC***	***ryfC***	CGCTTAATACGCGGCGTGCCAGG	CAGCAAAAGCGTTTGGCACCGGG	2,694,428	2,694,504		sRNA antisense regulator of shoB toxin
	acpS	***dpj***			2,694,526	2,694,906	+	ACP-CoA phosphopantetheinyltransferase;Holo-ACP synthase;4'-phosphopantetheinyl transferase
ECK2564	***era***	***sdgE***	TGAAGGCACCCGCTGGACGCCGG	CGCGTTTTTCTTCCATATGCAGG	2,696,389	2,697,294	+	ribosome-associated GTPase essential for growth and adaptation to thermal stress;GTP-dependent autophosphorylating protein kinase activity;membrane-associated
ECK2565	***rnc***	***ranA***	ATGCGCGCCACGCTGGTCCGTGG	ACGCTCGTTATGTTTACTGCTGG	2,697,291	2,697,971	+	ribonuclease III;RNase III;cleaves double-stranded RNA

ECK2566	***lepB***	***lep***	GTTAATTGGTGATTTTATTCTGG	GTTTCCAGCCAGCCAGGCTTCGG	2,698,243	2,699,217	+	leader peptidase (signal peptidase I)
ECK2571	***rpoE***	***sigE***	TCGTGCGCTGGATTCGTTCCGGG	GCGCACTACCAGTAAGTTAAAGG	2,703,345	2,703,920	+	RNA polymerase;sigma 24 (sigma E) factor
ECK2583	***pssA***	***pss***	TGTCGCCCTGTATCTCGAACAGG	TTGAGAAATCTTGGGTAGTTGGG	2,716,635	2,717,990	-	phosphatidylserine synthase (CDP-diacylglycerol-serine O-phosphatidyltransferase)
ECK2593	***bamD***	***ecfD,yfiO***	AAAAGCTGCAGGACGGTAACTGG	GTCGCGTAAATTTCATTTGGCGG	2,730,054	2,730,791	-	TPR-repeat lipoprotein;required for OM biogenesis;BamABCDE complex
ECK2603	***rplS***		GTTATCGCTATTCGTAACCGCGG	GTATCACCCGGACGGAAGGAAGG	2,738,091	2,738,438	+	50S ribosomal subunit protein L19
ECK2604	***trmD***		GTGATTTATCTGTCACCACAGGG	TGCACCATCATTAACATCCCCGG	2,738,480	2,739,247	+	tRNA m(1) G37 methyltransferase;SAM (S-adenosyl-L-methionine)-dependent
ECK2606	***rpsP***		AGAAGGCACTCGCCTGGATCTGG	GGCTGTCAGCGACAACAACCTGG	2,739,845	2,740,093	+	30S ribosomal subunit protein S16
ECK2607	***ffh***		GATTTCTTCCCTTCTGATGTTGG	TCTTCTTGTGCTTCTCGCGCAGG	2,740,342	2,741,703	+	signal recognition particle (SRP) protein;with 4;5S RNA;GTPase involved in co-translational protein translocation into and through membranes
ECK2610	***grpE***		TTTGCGTGTAAAAGCCGAAATGG	TGCCGTCACGTTCACGGGTCTGG	2,744,023	2,744,616	+	nucleotide exchange factor for the DnaKJ chaperone;heat shock protein;mutant survives lambda induction;stimulates DnaK and HscC ATPase
ECK2611	***nadK***	***yfjB,yfjE***			2,744,739	2,745,617	-	ATP-NAD kinase
ECK2691	***csrA***	***zfiA***	GGTCACCGTGACAGTTTTAGGGG	TGACCTCATCCCCAATCATGAGG	2,812,869	2,813,054	+	global regulator of carbon source metabolism;RNA binding protein
ECK2692	***alaS***	***act,ala-act,lovB***	GATCGACGTTATGGGCTCTGCGG	TAACCACGGTTTTCATTGGACGG	2,813,289	2,815,919	+	alanyl-tRNA synthetase
ECK2741	***ispF***	***ygbB***	CGATGCATTGCTTGGCGCGGCGG	GCCAGCAATCCTTTTTCGTAAGG	2,865,209	2,865,688	+	2-C-methyl-D-erythritol 2,4-cyclodiphosphate (MECP) synthase;alternative nonmevalonate (DXP) pathway for terpenoid biosynthesis
ECK2742	***ispD***	***ygbP***	TGAGCGTGCCGATTCCGTGCTGG	AATGACGACACGTTTCACCCGGG	2,865,688	2,866,398	+	4-diphosphocytidyl-2C-methyl-D-erythritol synthase
ECK2743	***ftsB***	***ygbQ***	TGACGATCTCAATGGCGGCCAGG	CGCCACATCATCATTGACGCGGG	2,866,417	2,866,728	+	cell division protein for septum localization dependent on FtsI and FtsQ;membrane protein
ECK2773	***eno***		GCTGAACTGAACGGTACTCCGGG	GCCGAATTTGGATTTGTTTTCGG	2,900,551	2,901,849	+	enolase;phosphoprotein;component of RNA degradosome
ECK2774	***pyrG***		CCTGACGCTGGTGCCGTACATGG	TGGCGAATCGCTTCGAGGAACGG	2,901,937	2,903,574	+	CTP synthetase
ECK2777	***mazE***	***chpAI,chpR,tasA***	TGATGAAGTGAAGATTGACCTGG	CCCAACGCTTTACGCTACTGTGG	2,904,999	2,905,247	+	antitoxin of ChpA-ChpR toxin-antitoxin system
ECK2824	***lgt***	***umpA***	CGCTGTATCTGTTCCGTGTCTGG	GTCCCCCGAGGAAGACGCCGAGG	2,959,070	2,959,945	+	Phosphatidylglycerol:prolipoprotein diacylglycerol transferase
ECK2886	***prfB***	***supK***	AGCTGGAACAGCCGGATGTCTGG	AAGATACCCCCTAAGAACGTCGG	3,029,092	3,030,190	+	translation termination peptide chain release factor 2;RF-2;recognizes UGA and UAA
ECK2921	***fbaA***	***ald,fba,fda***	ATCGACGGTCTGTTGGACGCGGG	GCGCACCAGAGATCGCGCCCAGG	3,064,073	3,065,152	+	fructose 1,6-bisphosphate aldolase;class II;binds Zn(II);homodimeric
ECK2922	***pgk***		CAGGCTTCTACTCACGGTATCGG	TGCAGCGTATTTTTTGGACAGGG	3,065,367	3,066,530	+	phosphoglycerate kinase
ECK2937	***metK***	***metX***	CGAAATCACCACCAGCGCCTGGG	AACCATGCCGGTTTTTACGTAGG	3,080,614	3,081,768	-	S-adenosyl-L-methionine synthetase
ECK2944	***yqgF***		CCTGCAATTAAAGCACAGGACGG	CGAGTAAGGTTCCACTCATCAGG	3,087,408	3,087,824	-	predicted Holliday junction resolvase;may be a nuclease that resolves Holliday junction intermediates in genetic recombination
ECK3009	***plsC***	***parF***	CCACCAGAACAACTATGACATGG	ACGGCCAAACATATGCCCAAAGG	3,156,652	3,157,389	+	1-acyl-sn-glycerol-3-phosphate acyltransferase
ECK3010	***parC***		ATCCGTAAAATCTACGAGAACGG	AGCTGATCGAGCGTGGTTTTCGG	3,157,623	3,159,881	+	DNA topoisomerase IV;subunit A
ECK3012	***mqsA***	***ygiT***	AGTAAAGGCATTTCGGGCTTCGG	TACTGTTTTTCGTCCACGGAAGG	3,161,759	3,162,154	+	antitoxin for MqsR toxin;predicted transcription regulator
ECK3021	***parE***	***nfxD***	GCCAAAGCGGTATTGTGCCCTGG	ATCCGGCCAGAAGTGCACACTGG	3,167,412	3,169,304	+	DNA topoisomerase IV;subunit B
ECK3032	***ribB***	***htrP,luxH***	GAAAATAACACCAGCGCCTATGG	CGCCATCTGCTCAACAGTCATGG	3,177,721	3,178,374	+	3,4-dihydroxy-2-butanone-4-phosphate synthase
ECK3046	***cca***		GCACGTACCGAACGGAAATCCGG	CCCGCGTCGAGCATCTCCTGTGG	3,195,799	3,197,037		tRNA nucleotidyltransferase;repairs terminal CCA of tRNAs
ECK3054	***tsaD***	***gcp,ygjD***	TCCCTGTACACCATATGGAAGGG	GCCGACTAATCCAGGGCCTGCGG	3,203,438	3,204,451	+	tRNA(NNU) t(6) A37 threonylcarbamoyladenosine modification;glycation binding protein
ECK3055	***rpsU***		CAAAGGTGAGAGGCACATGCCGG	CGACGCAGAGCTACGTCGAACGG	3,204,689	3,204,904	-	30S ribosomal subunit protein S21
ECK3056	***dnaG***	***dnaP,parB,sdgA***	AAACAGTTTTACCACTGCTTTGG	AGGACGGGGTTTTCTCGTTGTGG	3,205,015	3,206,760	-	DNA primase
ECK3057	***rpoD***	***alt***	AGGCTATCTGACCTATGCCGAGG	GCGGGTTTTGCTCCATAAGACGG	3,206,955	3,208,796	-	RNA polymerase;sigma 70 (sigma D) factor

ECK3102	***tdcF***	***yhaR***	CCCACAGACCGGTGAGATCCCGG	AATCAACGCCCTGAACATAAGGG	3,253,629	3,254,018	+	predicted L-PSP (mRNA) endoribonuclease
ECK3134	***rsmI***	***yraL***	ACGCCGCTAATTAACGATCCTGG	GGTCGTGCAGCGCAAACAACCGG	3,286,383	3,287,243	+	16S rRNA C1402 ribose 2'-O-methyltransferase;SAM (S-adenosyl-L-methionine)-dependent
ECK3147	***yhbV***		AGCGCCGACGTGATTTATCTTGG	GTACCACAGCACTGGCCCTAAGG	3,296,397	3,297,275	-	predicted U32 peptidase family protein
ECK3154	***rpsO***	***secC***	TGCACAGATCAACCACCTGCAGG	CAGCAGTGCTACCTGAACTTCGG	3,305,323	3,305,592	+	30S ribosomal subunit protein S15
ECK3157	***infB***	***gicD,ssyG***	GAAGCACGCGCAGCAGTACGTGG	TCGCGTTACGACCACGGCCACGG	3,307,250	3,309,922	+	bifunctional protein chain initiation factor 2;IF2: membrane protein/predicted protein
ECK3158	***nusA***		GACAACATCTCTCTGGATCTGGG	TGATTTCACCTTCGTGTTCACGG	3,309,947	3,311,434	+	transcription termination/antitermination L factor
ECK3165	***glmM***	***mrsA,yhbF***	GGAGATCAGCTGCGTTGATTCGG	CGATGCAGATATCACAATTCCGG	3,316,641	3,317,978	+	phosphoglucosamine mutase;UDP-GlcNAc pathway;peptidoglycan;lipopolysaccharide synthesis;mRNA stability effects
ECK3166	***folP***	***dhpS***	TTGCTCAACGCTTCGAAGTCTGG	CGCTAACTTCCGCCGCCCCTGGG	3,317,971	3,318,819	+	7,8-dihydropteroate synthase
ECK3167	***ftsH***	***hflB,mrsC,std,tolZ***	ATGGTCGGTCCTCCGGGTACCGG	GGATCTTACCGCCGAGTTTCTGG	3,318,909	3,320,852	+	shifted translation start -9 nt;UUG start codon;ATP-dependent membrane protease;binds Zn(II)
ECK3172	***obgE***	***cgtA,obg,yhbZ***	CGTCTGCTGGTTGCTAAGGGCGG	CGCGCTTACCGGTACAGTCGCGG	3,324,490	3,325,662	+	GTPase involved in cell partioning and DNA repair
ECK3174	***rpmA***	***rpz***	TTTCGGTGGCGAATCCGTTCTGG	TGAATCGCGACCGTTACGTGTGG	3,326,770	3,327,027	+	50S ribosomal subunit protein L27
ECK3175	***rplU***		GAAGTGCTGATGATCGCAAACGG	GGTGTTGTTTACCACCACTTTGG	3,327,048	3,327,359	+	50S ribosomal subunit protein L21
ECK3176	***ispB***	***cel,yhbD***	GTACTGGCTGCACGAGCTGTTGG	CTGATTGATCAGTTGGACGTCGG	3,327,618	3,328,589	-	octaprenyl diphosphate synthase
ECK3178	***murA***	***mrbA,murZ***	TACGTTAAAGCTTCCGTCGATGG	TGGTCGCGCCTAATTGTTCGAGG	3,329,143	3,330,402	+	UDP-N-acetylglucosamine 1-carboxyvinyltransferase
ECK3180	***mlaB***	***yrbB***	ACCGCTTTGGGAAATGCGTGAGG	AACGTCCTGATCCAGCTCTCCGG	3,330,871	3,331,164	+	ABC superfamily transporter maintaining OM lipid asymmetry;cytoplasmic STAS component
ECK3188	***lptC***	***yrbK***	ATTCCGATCAGGCCGTTTCGTGG	CGTATGCTCGCTTTTATAGGTGG	3,336,744	3,337,319	-	IM-tethered periplasmic lipopolysaccharide export protein of LptBFGC export complex
ECK3189	***lptA***	***yhbN***	GTGGTCGTTACCCGTCCGGGCGG	TGGTCCGATTCAATGTGGATCGG	3,337,288	3,337,845	-	periplasmic LPS-binding protein
ECK3190	***lptB***	***yhbG***	CAGCCTGACCGTCAACTCCGGGG	GGCCTTTATAGGCTTTTGCAAGG	3,337,852	3,338,577	-	lipopolysaccharide export ABC superfamily transporter ATP-binding protein of LptBFGC export complex
ECK3219	***rpsI***		CGGTCGTGAAACTGCCCGCATGG	CGCGAGCTGCGGAACTTTTGCGG	3,371,723	3,372,115	+	30S ribosomal subunit protein S9
ECK3220	***rplM***		GCTGACAAAGTTGCTGTAACCGG	CAGACGGCCCAGAGTTTTACCGG	3,372,131	3,372,559	+	50S ribosomal subunit protein L13
ECK3224	***degS***	***hhoB,htrH***	CAGCTTGAGATCCGCACCCTGGG	ATAGCTGGCAGGCGTCTCATCGG	3,376,108	3,377,175	-	serine protease;degrades periplasmic RseA;activating RpoE;multicopy suppressor of prc;periplasmic stress sensor for unfolded or misfolded OMPs
ECK3237	***mreD***		GACGCTTGGCGTACGCGTATTGG	TGAGTAACACCCAGTTTGGCCGG	3,392,295	3,392,783	+	cell wall structural complex MreBCD transmembrane component MreD
ECK3238	***mreC***		GCGTCAGGATGAGCAGAAAATGG	CCCGGTTTTCAAGTTCTAATTGG	3,392,783	3,393,886	+	cell wall structural complex MreBCD transmembrane component MreC
ECK3239	***mreB***	***envB,mon,rodY***	GCGCGTTCTGGTTTGTGTGCCGG	GCGATAACGCCGTCTTTCATTGG	3,393,952	3,394,995	+	cell wall structural complex MreBCD;actin family protein component MreB
ECK3242	***accB***	***fabE***	ATCGAGCTGGTTGAAGAATCAGG	TCTTACGAATATCCATGAGTGGG	3,398,546	3,398,939	-	acetyl CoA carboxylase;BCCP subunit
ECK3243	***accC***	***fabG***	AGCGGGCGTCCCTTGCGTACCGG	CAAAGTTGGCGTTCTCGGAGAGG	3,399,825	3,401,174	-	acetyl-CoA carboxylase;biotin carboxylase subunit
ECK3269	***tsaC***	***rimN,yrdC***	GTGAAACCATTTTTTCCCGCTGG	AACATGGTGTCATCAATATAGGG	3,424,751	3,425,323	+	tRNA(NNU) t(6) A37 threonylcarbamoyladenosine modification;threonine-dependent ADP-forming ATPase
ECK3273	***def***	***fms***	ATGTACGCAGAAGAAGGTATTGG	ACTTTGCGAAGCCGCTCGTCCGG	3,427,598	3,428,107	-	peptide deformylase;N-formylmethionylaminoacyl-tRNA deformylase
ECK3274	***fmt***	***yhdD***	GTGCTGGAGATGCCGCGTCTTGG	ACCAGTTGCTGGTTTTCTTGTGG	3,428,122	3,429,069	-	Methionyl-tRNA formyltransferase
ECK3281	***rplQ***		TAGCGTTGCTAATCGTCGTCTGG	CCAGTGAACCTGCCATATTGCGG	3,433,524	3,433,907	+	50S ribosomal subunit protein L17
ECK3282	***rpoA***	***pez,phs,sez***	GCAGCCGATATCACCCACGACGG	TCAGGTTGAGCAGGATTTCCAGG	3,433,948	3,434,937	+	RNA polymerase;alpha subunit
ECK3283	***rpsD***	***ramA,sud(2)***	CAAAAAGTTCGCCGTATCTATGG	TCGCGCGAACGCCAGACTTAAGG	3,434,963	3,435,583	+	30S ribosomal subunit protein S4
ECK3284	***rpsK***		TCAGGGTAACGCGTTGGGTTGGG	ACACGTTTACGTGCACGAATTGG	3,435,617	3,436,006	+	30S ribosomal subunit protein S11
ECK3285	***rpsM***		GTTGCCAAATTTGTCGTTGAAGG	CAGCCAGGATGGCTTTAGAACGG	3,436,023	3,436,379	+	30S ribosomal subunit protein S13
ECK3287	***secY***	***prlA***	CCGAATATGCCTGGTATGCAAGG	TCTTAATTTCTGCCAACGTTGGG	3,436,674	3,438,005	+	preprotein translocase membrane subunit

ECK3288	***rplO***		TCTGGCGGTGGCGTACGTCGCGG	ACCCGCCTTTTTGGAGCCTTCGG	3,438,013	3,438,447	+	50S ribosomal subunit protein L15
ECK3289	***rpmD***		GAGGATACTCCTGCTATTCGCGG	GCAGACGACCGATTGCACTGCGG	3,438,451	3,438,630	+	50S ribosomal subunit protein L30
ECK3290	***rpsE***	***eps,spc,spcA***	AGCAGCGATCCAGAAAGCGATGG	CTTTAACGGTTTTAGATACGCGG	3,438,634	3,439,137	+	30S ribosomal subunit protein S5
ECK3291	***rplR***		ACCGAACGGTTCTGAAGTTCTGG	CGCCCAGCTCCTGGAGCTTGCGG	3,439,152	3,439,505	+	50S ribosomal subunit protein L18
ECK3292	***rplF***		GCGTGCCCTGCTGAACTCAATGG	CATCACGCGGACCGAAGGTCAGG	3,439,515	3,440,048	+	50S ribosomal subunit protein L6
ECK3293	***rpsH***		GCCAACGTGCTGAAGGAAGAAGG	TTTCAGCTTGGAGGAAGGCATGG	3,440,061	3,440,453	+	30S ribosomal subunit protein S8
ECK3294	***rpsN***		CGTCAAACAGGTCGTCCGCATGG	AGCGTTCCAACGATCTTCGTCGG	3,440,487	3,440,792	+	30S ribosomal subunit protein S14
ECK3295	***rplE***		AAAGCACGCAAATCTGTTGCAGG	AGGGTGATCTTCTCGACCCGAGG	3,440,807	3,441,346	+	50S ribosomal subunit protein L5
ECK3296	***rplX***		GGTTAAGAAACATCAGAAGCCGG	CGATAACTTCGTCATCACGACGG	3,441,361	3,441,675	+	50S ribosomal subunit protein L24
ECK3297	***rplN***		ATCAAAGAAGCAATTCCGCGTGG	CGTTCAGCATAGTCTGTTCTTGG	3,441,686	3,442,057	+	50S ribosomal subunit protein L14
ECK3298	***rpsQ***	***neaA***	TTTGTGAAACACCCGATCTACGG	TTGCAGAGTACGGATTTTATCGG	3,442,222	3,442,476	+	30S ribosomal subunit protein S17
ECK3299	***rpmC***		GCAGTTCAACCTGCGTATGCAGG	ACGCAGCAGGTTCAGCAGCTCGG	3,442,476	3,442,667	+	50S ribosomal subunit protein L29
ECK3300	***rplP***		ACCCGTGCAGTTAAGCGTCAAGG	CCTGCGCCAGACCGCGGTTACGG	3,442,667	3,443,077	+	50S ribosomal subunit protein L16
ECK3301	***rpsC***		AGACGTAGAAAAACTGCGTAAGG	CTTTAGCCAGTTCCTTAGTCAGG	3,443,090	3,443,791	+	30S ribosomal subunit protein S3
ECK3302	***rplV***	***eryB***	CAACAAGAAAGCGGCTGTACTGG	ACACTTTCTTACCGCGAATCAGG	3,443,809	3,444,141	+	50S ribosomal subunit protein L22
ECK3303	***rpsS***		ACAAGAAGCCCCTGCGCACTTGG	GGACCTTTCTTGAGAGAACGTGG	3,444,156	3,444,434	+	30S ribosomal subunit protein S19
ECK3304	***rplB***		CGGTGAACGCCGTTACATCCTGG	CAATACGGTAAGCCTGCTTGTGG	3,444,451	3,445,272	+	50S ribosomal subunit protein L2
ECK3305	***rplW***		AGTCGAAGTCGTTAACACCCTGG	AACTTTGAGTACGATGGTGTTGG	3,445,290	3,445,592	+	50S ribosomal subunit protein L23
ECK3306	***rplD***	***eryA***	CCGTGGCGCCAGAAAGGCACCGG	CTGCATAAGCAACAACAACCTGG	3,445,589	3,446,194	+	50S ribosomal subunit protein L4
ECK3307	***rplC***		GCTGGCCACTTCGCTAAAGCTGG	GTTTGCTTCAACTTCGATTACGG	3,446,205	3,446,834	+	50S ribosomal subunit protein L3
ECK3308	***rpsJ***	***nusE***	CGCACTGGTGCGCAGGTCCGTGG	TCAGGCGGATACGGATTCTTTGG	3,446,867	3,447,178	+	30S ribosomal subunit protein S10
ECK3327	***fusA***	***far,fus***	CCCGGCAGACATGGTTGAACTGG	TGATAGCTTTCATTTTCACCAGG	3,465,308	3,467,422	+	protein synthesis Elongation Factor EF-G;GTP-binding;fusidic acid resistance
ECK3328	***rpsG***		GCGCTCTGGTAAATCTGAACTGG	AGCAGTTCTGATCCGAACTTCGG	3,467,450	3,467,989	+	30S ribosomal subunit protein S7
ECK3329	***rpsL***	***asuB,strA***	TGCCGTGTTCGTCTGACTAACGG	TGCGGGCATGCTTCCAGCCCAGG	3,468,086	3,468,460	+	30S ribosomal subunit protein S12
ECK3371	***trpS***		CGCTGGTCTGTTTGACTATCCGG	CTGTGCATGTTCCGGCACGTGGG	3,506,542	3,507,546	+	tryptophanyl-tRNA synthetase
ECK3385	***yrfF***	***igaA(S.t.),mucM(S.t.),umoB(P.m.)***	ATGGCGTTACTACCTCGATTCGG	GTGCCGCGCTGGCTCCCGTTGGG	3,520,377	3,522,512		predicted inner membrane protein;RcsCDB-response attenuator
ECK3419	***asd***	***dap,hom***	AATGGCATCAGGACTTTTGTTGG	GTAATGACGTCCTGATTGACGGG	3,567,684	3,568,787	+	aspartate-semialdehyde dehydrogenase;NAD(P)-binding
ECK3445	***rpoH***	***fam,hin,htpR***	CGCCGTTTCAACCCGGAAGTGGG	CAACAAACCGCAGGTGAGACAGG	3,593,838	3,594,692	+	RNA polymerase;sigma 32 (sigma H) factor
ECK3446	***ftsX***	***ftsS***	TGACGCTGCTGCGGGCGTGGTGG	TACACCATATAACAGACGCTGGG	3,594,937	3,595,995	+	Integral membrane protein invoved in cell division;binds FtsE to the inner membrane
ECK3447	***ftsE***		CCGTTTCTGCGCCGCCAGATTGG	GCCGCTAAACCAGATTTTCCCGG	3,595,988	3,596,656	+	ATP-binding protein associated with the inner membrane via FtsX;cell division;null mutant has filamentous growth and requires high salt for viability
ECK3448	***ftsY***		GGCGGTGATGGTGGTTCCTCCGG	CGCCGCTTCGACAATCTCTACGG	3,596,659	3,598,152	+	signal recognition particle (SRP) receptor;GTPase
ECK3455	***yhhQ***		CTCGCAAACTCAACGCTATAAGG	CGGAAAGCTAAACGCGCCCCAGG	3,603,126	3,603,791	-	predicted DUF165 family inner membrane protein
ECK3517	***bcsB***	***yhjN***	CACCTTGCCGATGGTCTTTGCGG	CTTCACATTCAGGGTCTGATAGG	3,684,177	3,686,516	+	Cellulose synthase;regulatory subunit;may bind cyclic-di-GMP;may be periplasmic
ECK3539	***yiaD***		CGCAGGTCTGGGCTCTCTCGTGG	TTTACCTGCTTCGCGTTCGCCGG	3,710,456	3,711,115	-	multicopy suppressor of bamB;OM (outer membrane) lipoprotein;suppresses temperature-sensitive mutations in BamB when overexpressed
ECK3547	***glyS***	***gly-act,glyS(B)***	CCGGAAATTCTGCGTGAGCGTGG	TGAATGCCCAGAATGGTTGCCGG	3,716,237	3,718,306	+	glycine tRNA synthetase;beta subunit
ECK3548	***glyQ***	***cfcA,glyS(A)***	GCCATCGCCGGACAATATTCAGG	TAGTAGTGCTGTAAACGGTTGGG	3,718,316	3,719,227	+	glycine tRNA synthetase;alpha subunit
ECK3584	***rhsJ***	***yibJ***	GATGAAAAACCAGATGGACCCGG	GTCGAGCATCTGCACCGTGCCGG	3,761,130	3,762,074	-	predicted Rhs-family protein
ECK3598	***gpsA***		TCCTGATGCGCGTCTGGTGTGGG	TGGCGAGATCGCTTTCAAGATGG	3,776,551	3,777,570	+	sn-Glycerol-3-phosphate dehydrogenase [NAD(P) +]
ECK3613	***waaU***	***rfaK,waaK(S.t.)***			3,792,148	3,793,221	+	lipopolysaccharide core biosynthesis
ECK3623	***waaA***	***kdtA***	GGTGCGCGGACGTAAGGCTCCGG	CGGCTGAATAAGGTAGAGAAGGG	3,802,449	3,803,726	-	3-deoxy-D-manno-octulosonic-acid transferase (KDO transferase)
ECK3624	***coaD***	***kdtB,yicA***	CAGCAGGCAACCGCGCATCTGGG	TAAACATCGGTTTTTTACTGGGG	3,803,734	3,804,213	-	phosphopantetheine adenylyltransferase

ECK3627	***rpmB***		CCTGCACTCTCACCGTTTCTGGG	TGCCAGTAACTTGGCAGACTCGG	3,805,347	3,805,583	+	50S ribosomal subunit protein L28
ECK3629	***dfp***	***coaBC***	TTCCGACAGTCTGCTGGACCCGG	GCAAACGACGCACCAGTTCAGGG	3,806,640	3,807,860	-	coenzyme A biosynthesis;bifunctional enzyme;phosphopantothenoylcysteine decarboxylase (N) and phosphopantothenoylcysteine synthase (C)
ECK3630	***dut***	***dnaS,sof***	AAGATTCTGGACCCGCGCGTTGG	GAGGTGGCATAAGTCGGGAGCGG	3,807,838	3,808,296	-	shifted translation start -3 nt;AUG start codon;dUTP pyrophosphatase;dUTPase
ECK3638	***gmk***	***spoR***	CCAACCGTTGTATGACACCCAGG	TGGATTTACCCGCGCCACTGGGG	3,815,337	3,815,960	-	guanylate kinase
ECK3640	***spoT***		CGTTGCACGTGATGCTCACGAGG	AGACGCTTGATTTGGTCTTCCGG	3,816,309	3,818,417	-	bifunctional (p) ppGpp synthetase II/guanosine-3';5'-bis pyrophosphate 3'-pyrophosphohydrolase
ECK3691	***gyrB***	***acrB,Cou,himB,hisU,hopA,nalC,parA,pcbA,pcpA***	AGACGGTATTGGCGTCGAAGTGG	AGAAGTAGAAGATATTCGGGTGG	3,871,614	3,874,028	+	DNA gyrase;subunit B
ECK3693	***dnaN***		GCTGTCTACCCTGCCAGCGGCGG	CCGCGCCTTCAGGCAGACCACGG	3,875,130	3,876,230	+	DNA polymerase III sliding clamp beta subunit;required for high processivity;required for regulatory inactivation of DnaA
ECK3694	***dnaA***	***hsm-2***	CAGCCACAGAATTCAGTATGTGG	CACATACTGAATTCTGTGGCTGG	3,876,235	3,877,638	+	chromosomal replication initiator protein DnaA;DNA-binding transcription regulator
ECK3695	***rpmH***	***rimA,ssaF***	CTGAAGCGCAACCGTTCTCACGG	CGGTTGAAAAGTGCGTTTCATGG	3,878,245	3,878,385	-	50S ribosomal subunit protein L34
ECK3696	***rnpA***		GGTTAAGCTCGCATTTCCCAGGG	AAGACGAATGTGAATTGACTGGG	3,878,402	3,878,761	-	protein C5 component of RNase P;involved in tRNA and 4;5S RNA-processing
ECK3698	***yidC***		TGGCCAGGGGAAACTGATCTCGG	CGGGTTTTTATCCTGCTCCCAGG	3,878,985	3,880,631	-	membrane protein insertase;inner membrane protein integration factor;binds transmembrane (TM) regions of nascent inner membrane proteins (IMPs);required for Sec-independent IMP integration;associated with the Sec translocase
ECK3702	***tnaB***	***tnaP,trpP***	AAAAAAAGCACTCTGCATTTTGG	CAGGCACCGGCAAGATCAACAGG	3,884,145	3,885,392	-	tryptophan transport protein of low affinity
ECK3722	***glmS***		GCTGGTGATTGGCCTGGGGATGG	GTGCCGCCAGCAGGGTATCCGGG	3,905,748	3,907,577	+	L-glutamine:D-fructose-6-phosphate aminotransferase
ECK3723	***glmU***	***tms(B.s.),yieA***	CTGGATGATCCGACCGGTTATGG	AGCAGACCAATGCCACCCTGCGG	3,907,739	3,909,109	+	bifunctional glucosamine-1-phosphate acetyltransferase and N-acetylglucosamine-1-phosphate uridylyltransferase;hexameric
ECK3775	***rho***	***hdf,nitA,nusD,psuA,rnsC,sbaA,sun,tabC,tsu***	TCTCGGCGAAAATATGGGGCTGG	GGTGTGAGTTCTTAAACTTGGGG	3,960,326	3,961,585	-	transcription termination factor
ECK3786	***wzyE***	***rffT***	GTTGCAGGCGTTGCTTTCTGCGG	ATACCAGCACGCTGGTCAGCGGG	3,972,561	3,973,862	-	shifted translation start +51 nt;CUG start codon;predicted Wzy protein involved in ECA polysaccharide chain elongation
ECK3798	***hemD***		CTACCTGATTATTTCGCCATTGG	AGCAACCGCGTGTTGCGAGAGGG	3,982,997	3,983,737	+	uroporphyrinogen III synthase
ECK3799	***hemC***	***popE***	GTCCAATAACTATGACAGTCTGG	CACATCTTTCATTGAGTGTACGG	3,983,734	3,984,675	+	hydroxymethylbilane synthase
ECK3828	***ubiJ***	***yigP***	ACGGCCCGCTCGCGTCTGCTGGG	TCAATTCCTGCCGTCACTAAAGG	4,013,533	4,014,138	-	aerobic ubiquinone synthesis protein;SCP2 family protein
ECK3829	***ubiB***	***aarF,yigQ,yigR,yigS***	CGACTGGCCCTGCAAGAACTGGG	AGTTTGTCTTTATGCCGATTTGG	4,014,135	4,015,775	-	2-octaprenylphenol hydroxylase
ECK3835	***ubiD***	***yigC,yigY***	CTGTTGTTCGAAAACCCTAAAGG	ATTTCCAGATGCGGATCCACCGG	4,018,906	4,020,390	-	shifted translation start +9 nt;AUG start codon;3-octaprenyl-4-hydroxybenzoate decarboxylase
ECK3842	***hemG***	***yihB***	GAAAGAACTGGGGATCCAGGCGG	CAGTTCCGAAGCCAGGTAGGAGG	4,028,517	4,029,062	-	protoporphyrin oxidase;flavoprotein
ECK3855	***polA***	***resA***	GAGCCGACCGGTGCGATGTATGG	GAACCATCTACAAGGATAAGTGG	4,040,875	4,043,661	-	5' to 3' DNA polymerase and 3' to 5' and 5' to 3' exonuclease
ECK3857	***yihA***	***engB***	AGTGGCTGACGGCAAGCGTCTGG	TGGATTTACCTGCGTTGGAACGG	4,044,042	4,044,674	+	predicted GTP-binding protein required for normal cell division and maintenance of normal septation
ECK3925	***ftsN***	***msgA***	AAAGCAACGAAATCTGCCTGCGG	TTCGCCGCGAAGGTGCCGGTTGG	4,112,857	4,113,816	+	cell division and growth
ECK3958	***btuB-NTB***	***bfe,cer,dcrC***	TGATGTGCTGCGCCGTCTTCCGG	CAGTAACGACGAGAGTATCCGGG	4,154,116	4,155,960	-	vitamin B12 OM (outer membrane) receptor;receptor for E colicins;phages BF23 and C1; variation from MG1655 U00096;3;C>G;btuB (A162G)
ECK3959	***murI***	***dga,glr,yijA***	CTCCGGCGTCGGTGGGTTGTCGG	AACACCAGCACGGTGGGACGTGG	4,155,905	4,156,762	-	glutamate racemase;D-glutamate synthesis
ECK3964	***murB***	***yijB***	GGACAACCCGTTCTTATTCTGGG	GAGTAATTGTTGTTCGTCTTCGG	4,162,534	4,163,562	-	UDP-N-acetylenolpyruvoylglucosamine reductase;FAD-binding

ECK3965	***birA***	***bioR,dhbB***	GAGCAGTTGGGTGAAACGCTGGG	AAATTCACCGTTCGCTAACAGGG	4,163,559	4,164,524	-	bifunctional biotin protein ligase;biotin operon regulator (repressor);biotin-[acetyl-CoA carboxylase] holoenzyme synthase;monomeric
ECK3966	***coaA***	***panK,rts,ts-9***	TATTGCTGGCAGTGTCGCGGTGG	TGACAAAGGTAAATAGATCTCGG	4,164,553	4,165,503	+	pantothenate kinase
ECK3972	***secE***	***mbrC,prlG***	GATGAAGTGGGTCGTTGTGGTGG	GCGCCCGCTTCCTTGAGCTTCGG	4,167,835	4,168,218	-	preprotein translocase membrane subunit
ECK3973	***nusG***		GAAGTGGTTGAAATCCGTGGCGG	CACGCAGCGACGTTGCTACGCGG	4,168,220	4,168,765	-	transcription termination factor
ECK3974	***rplK***	***relC***	CCGCCAGTAGGTCCGGCTCTGGG	TGCAACCTGCAGCTTGACATAGG	4,168,924	4,169,352	-	50S ribosomal subunit protein L11
ECK3976	***rplJ***		CTGCTGCGCCGTGCTGTTGAAGG	TACTGCAGACAGCGCGCCTTTGG	4,170,473	4,170,970	-	50S ribosomal subunit protein L10
ECK3977	***rplL***		GCTGTTATCAAAGCAGTACGTGG	TTTCTTCAGCAGCTTCAACCGGG	4,171,037	4,171,402	-	50S ribosomal subunit protein L7/L12
ECK3978	***rpoB***	***ftsR,groN,mbrD?,nitB,rif,ron,sdgB,stl,stv,tabD,tabG***	TCCCAGCTGCACCGTAGTCCGGG	GTACCGTTGTCTGTCATGAGCGG	4,171,722	4,175,750	-	RNA polymerase;beta subunit
ECK3979	***rpoC***	***tabB***	AAAGATTACGAGTGCCTGTGCGG	GACCATGAACGGATCATGTCTGG	4,175,827	4,180,050	-	RNA polymerase;beta prime subunit
ECK3989	***hemE***	***hemC***	TGCCGGACGCGATGGGGTTAGGG	CGACATAAAATCGCCTGCCTGGG	4,188,193	4,189,257	-	uroporphyrinogen decarboxylase
ECK4032	***ubiA***	***cyr,sdgG?***	CACCGGGCGTTCCCCAGCTCTGG	ACCCACAACGCCCATAATGTTGG	4,243,493	4,244,365	-	p-hydroxybenzoate octaprenyltransferase
ECK4033	***plsB***		CACCTGGAACCGACTTTACCAGG	AGCCCAGAATACGGTCAGTCAGG	4,244,520	4,246,943	+	glycerol-3-phosphate O-acyltransferase
ECK4035	***lexA***	***exrA,spr,tsl,umuA***	CGGAAATCGCGCAGCGTTTGGGG	GCGTCGGCGGCATACCTGTCTGG	4,247,592	4,248,200	-	global regulator (repressor) of SOS regulon;dimeric
ECK4044	***dnaB***	***groP,grpA,grpD***	CGAAGCGGAGCAGTCGGTGTTGG	ACTTTCAGCCCGGCAACTTGTGG	4,254,791	4,256,206	-	replicative DNA helicase
ECK4051	***ssb***	***exrB,lexC***	TCTCGTTGGTAATCTGGGTCAGG	AACCTTGTTTACGCCTCTGCTGG	4,264,602	4,265,138	-	single-stranded DNA-binding protein
ECK4077	***alsK***	***yjcT***	CGTTAACCTGCAACTCTCCTGGG	GCCGCTGTTAACGGCAGGTTAGG	4,297,236	4,298,165	+	D-allose kinase
ECK4136	***groS***	***groES,mopB,TabB***	GCTGGCGGCATCGTTCTGACCGG	AGCCAGCACTTCGCCGCGGGTGG	4,361,054	4,361,347	-	chaperonin Cpn10;GroESL small subunit GroES;phage morphogenesis
ECK4137	***groL***	***groEL,mopA***			4,361,391	4,363,037	-	Cpn60 chaperonin GroEL;large subunit of GroESL
ECK4156	***psd***		GATAAAATCCTGCAAGCCAAAGG	GGGCGTACTTCGTCACGCAGCGG	4,379,758	4,380,726	+	phosphatidylserine decarboxylase
ECK4158	***orn***	***yjeR***	CGATCTTGAGATGACCGGTCTGG	CCAAATCAGGTTGTTTTCATTGG	4,381,970	4,382,515	-	oligoribonuclease
ECK4164	***tsaE***	***yjeE***	CGGGGCTTTTTACAGGCTCTGGG	ATCGCCATACAGATAGATTACGG	4,385,951	4,386,412	-	tRNA(NNU) t(6) A37 threonylcarbamoyladenosine modification;ADP binding protein
ECK4197	***priB***		GTTGTCCGGCACCGTGTGCAGGG	TGATGGACTGACCTTTCGAAGGG	4,415,886	4,416,200	-	primosomal protein N;ssDNA-binding protein
ECK4198	***rpsR***		TTCTGCCGTTTCACCGCGGAAGG	GACGGAAATAACGTGCCATATGG	4,416,205	4,416,432	-	30S ribosomal subunit protein S18
ECK4220	***chpS***	***chpBI,yjfB***	GAAAGAACTTAACTTACAGCCGG	ACTGTTCCCCCATCTTTTTATGG	4,438,813	4,439,064	-	antitoxin of ChpBS toxin-antitoxin system
ECK4222	***ppa***		CATCAACCACACCCTGTCTCTGG	ACATCGCGGTGGACATGAAGCGG	4,439,488	4,440,018	-	inorganic pyrophosphatase
ECK4251	***valS***	***val-act***	CGACTTTAACGACTATGAAGTGG	CGCCAACGATCGGAATACGACGG	4,471,348	4,474,203	+	valyl-tRNA synthetase
ECK4254	***lptF***	***yjgP***	CTGCCATTAAGCCTGTTCCTCGG	CGGCACGCCCAACCCGAGAAGGG	4,476,584	4,477,684	-	lipopolysaccharide export ABC permease of LptBFGC export complex
ECK4255	***lptG***	***yjgQ***	CAGGGGAGTTACGACGCGTTAGG	AAATAGTTTTACCGATATAGCGG	4,477,684	4,478,766	-	lipopolysaccharide export ABC permease of LptBFGC export complex
ECK4351	***dnaC***	***dnaD***	AACTATCGCGTTGAGTGTGAAGG	CACGTTCGAGAGCGGCGGAGCGG	4,590,604	4,591,341	+	DNA biosynthesis protein
ECK0058			TGAATATCTTTCAACGCTTGTGG	AAAAGTGTAACCGGCATGCCTGG				
ECK2936								

Supplementary Table 2. *dnaA* synthetic lethal genes with Keio collection.

Supplementary table 2: Synthetic lethal combination of *dnaA* knockdown induced by 0.2 aTc

ECK_id: eck id of E. coli K-12 gene annotation
name: gene name
description: description of E. coli gene annotation

ECK_id	name	description
ECK1525	***marB***	periplasmic mar operon regulator
ECK2307	***cvpA***	membrane protein;required for colicin V production
ECK0741	***zitB***	zinc Zn(II) efflux transport protein;zinc-inducible
ECK4258	***idnT***	L-idonate and D-gluconate transport protein
ECK3566	***yiaM***	2,3-diketo-L-gulonate TRAP (TRipartite ATP-independent periplasmic) family protein small permease transport protein system YiaMNO involved in the uptake of 2,3-diketo-L-gulonate
ECK2005	***sbcB***	exodeoxyribonuclease I;exonuclease I
ECK1843	***holE***	DNA polymerase III;theta subunit
ECK4087	***phnN***	ribose 1;5-bisphosphokinase
ECK2381	***ypdF***	Xaa-Pro aminopeptidase;hydrolyzes N-terminal methionine when the next amino acid is alanine;proline or serine;substrate preference for methionyl aminopeptidase activity is Pro > Ala > Ser;can hydrolyze the Xaa-Pro peptide bond when the first amino acid is alanine, asparagine or methionine
ECK2392	***yfeC***	shifted translation start -15 nt;AUG start codon;predicted DUF1323 YfeC/YfeD family DNA-binding transcription regulator
ECK0867	***ybjD***	predicted protein with nucleoside triphosphate hydrolase domain
ECK2552	***yfhG***	predicted OM (outer membrane) protein modulating the QseEF response
ECK2600	***yfiR***	predicted periplasmic inhibitor of YfiN activity;required for swarming motility
ECK3342	***yheU***	predicted UPF0270 family protein
ECK2496	***purN***	phosphoribosylglycinamide formyltransferase 1
ECK4278	***fecD***	ferric citrate ABC superfamily transporter;permease
ECK3430	***yrhB***	predicted stable heat shock chaperone protein;monomeric
ECK3437	***ugpB***	glycerol-3-phosphate transport protein subunit
ECK4369	***yjjU***	predicted phospholipase;patatin family protein;predicted lipid hydrolase
ECK1938	***fliH***	negative regulator of FliI ATPase activity involved in flagellar assembly and export
ECK4381	***radA***	DNA repair protein
ECK1684	***ydiJ***	predicted FAD-linked oxidoreductase
ECK2175	***yejG***	predicted protein
ECK1152	***ymgA***	predicted connector protein for RcsB/C regulation of biofilm formation
ECK3375	***damX***	cell division protein;binds septal ring;bile salts resistance
ECK1946	***fliP***	flagellin export apparatus;integral membrane protein
ECK2567	***lepA***	back-translocating elongation factor EF4;GTPase
ECK3250	***yhdU***	predicted membrane protein
ECK2477	***hyfA***	hydrogenase 4;4Fe-4S subunit
ECK0133	***panB***	3-methyl-2-oxobutanoate hydroxymethyltransferase
ECK3462	***nikC***	nickel ABC superfamily transporter permease
ECK1761	***topB***	DNA topoisomerase III
ECK3891	***frvX***	predicted peptidase homolog
ECK1958	***vsr***	DNA mismatch endonuclease of very short patch repair
ECK2619	***yfjH***	CP4-57 cryptic prophage predicted protein
ECK1202	***ychA***	transglutaminase family protein TPR-repeat protein
ECK2782	***gudX***	glucarate dehydratase-related protein;substrate unknown
ECK0814	***ybiY***	predicted pyruvate formate lyase activating enzyme
ECK2543	***yphD***	predicted sugar transport protein subunit: membrane component of ABC superfamily protein;may be part of binding-protein-dependent transport system YphDEF;may be responsible for the translocation of substrate across the membrane
ECK0425	***cyoB***	cytochrome O oxidase subunit I;quinone-binding domain;cytochrome bo(3) ubiquinol oxidase subunit I
ECK0631	***cobC***	predicted alpha-ribazole-5'-phosphate phosphatase;potential partial cobalamin biosynthesis pathway
ECK0422	***cyoE***	cytochrome O oxidase protoheme IX farnesyltransferase subunit protoheme IX farnesyltransferase (haeme O biosynthesis)
ECK2350	***yfdM***	CPS-53 (KpLE1) cryptic prophage predicted methyltransferase
ECK1848	***yebF***	predicted expressed extracellular Colicin M immunity family protein

ECK2712	***hycI***	protease involved in processing C-terminal fragment end of HycE
ECK0476	***ybaP***	predicted TraB family protein
ECK2694	***recA***	DNA recombination and repair protein;ssDNA-dependent ATPase;synaptase;ssDNA and dsDNA binding protein;ATP-dependent homologous DNA strand exchanger;recombinase A;LexA autocleavage cofactor
ECK0009	***mog***	molybdochelatase incorporating molybdenum into molybdopterin;chlorate resistance;molybdopterin adenylyltransferase
ECK1927	***yedD***	predicted/verified lipoprotein
ECK2867	***ygeX***	2,3-diaminopropionate (DAPA) ammonia lyase;pyridoxal phosphate (PLP)-dependent
ECK3717	***phoU***	negative regulator of PhoR/PhoB two-component regulatory system;Bag domain, chaperone Hsp70 family protein suggests PhoU promotes dephosphorylation of phospho-PhoR/PhoB via protein-protein interaction
ECK0504	***ybbW***	predicted allantoin transport protein
ECK4279	***fecC***	ferric citrate ABC superfamily transporter;permease
ECK1768	***ydjF***	predicted DNA-binding DeoR family transcription regulator
ECK0538	***ybcN***	DLP12 cryptic prophage;predicted recombination protein;SSB and ssDNA binding protein
ECK1466	***yddL***	predicted lipoprotein
ECK1554	***ydfU***	Qin prophage;predicted DUF968 family protein
ECK0308	***ykgH***	predicted inner membrane protein
ECK3452	***yhhN***	predicted TMEM86 family inner membrane protein
ECK3885	***fdoI***	formate dehydrogenase-O cytochrome b556 subunit;aerobic
ECK4305	***fimA***	Fimbrin type 1;major structural subunit;phase variation
ECK3480	***uspA***	universal stress global response regulator
ECK0237	***prfH***	predicted defective peptide chain release factor
ECK1728	***ydjO***	predicted protein
ECK0225	***yafK***	predicted L;D-transpeptidase-related protein
ECK3677	***yidE***	predicted transport protein
ECK3495	***hdeD***	acid-resistance membrane protein
ECK1394	***paaJ***	3-oxoadipyl-CoA/3-oxo-5;6-dehydrosuberyl-CoA thiolase
ECK4067	***nrfE***	heme lyase (NrfEFG) for insertion of heme into c552;subunit NrfE
ECK3217	***sspB***	ClpXP protease specificity-enhancing factor
ECK4392	***creD***	inner membrane protein of unknown function
ECK0395	***brnQ***	branched-chain amino acid transport system 2 carrier protein;LIV-II transport system for Ile;Leu;and Val;mutants are valine and o-methylthreonine resistant;glycylvaline sensitive
ECK1474	***sra***	stationary-phase-induced ribosome-associated protein
ECK2319	***yfcL***	predicted protein
ECK1535	***ydfI***	predicted D-mannonate oxidoreductase;NAD-dependent
ECK0596	***ybdN***	predicted PAPS reductase family domain protein
ECK0923	***pepN***	aminopeptidase N
ECK4134	***fxsA***	suppressor of F exclusion of phage T7
ECK3344	***yhfA***	predicted OsmC family protein
ECK3690	***yidB***	predicted DUF937 family protein
ECK1948	***fliR***	flagellin export apparatus;integral membrane protein
ECK0048	***kefC***	potassium:proton antiporter NEM-activatable K+/H+ antiporter
ECK0851	***artJ***	arginine ABC superfamily transporter periplasmic binding protein
ECK2519	***sseB***	rhodanase family protein enzyme;sulfur transfer from thiosulfate
ECK1220	***narJ***	molybdenum-cofactor-assembly chaperone subunit (delta subunit) of nitrate reductase 1
ECK0562	***cusS***	two-component regulatory system sensory histidine kinase protein (autophosphorylation site,His271);senses copper;CusR is partner response regulator
ECK3981	***thiH***	tyrosine lyase;involved in thiamin-thiazole moiety synthesis
ECK2539	***hcaD***	phenylpropionate dioxygenase;ferredoxin reductase subunit
ECK3628	***yicR***	predicted UPF0758 family protein;not radC
ECK1993	***flu***	CP4-44 cryptic prophage antigen 43;phase-variable bipartite OM (outer membrane) protein;Ag43 affects surface properties;piliation;colonial morphology
ECK2126	***dld***	D-lactate dehydrogenase;FAD-binding;NADH independent;FAD enzyme
ECK3355	***nirC***	nitrite uptake transport protein
ECK4015	***yjbD***	predicted DUF3811 family protein

ECK3874	***yihT***	predicted LacD aldolase family protein of unknown function;6-deoxy-6-sulphofructose-1-phosphate aldolase;cleaves 6-deoxy-6-sulfo-D-fructose 1-phosphate (SFP) to form dihydroxyacetone phosphate (DHAP) and 3-sulfolactaldehyde (SLA)
ECK0305	***ykgE***	predicted electron transport chain YkgEFG component;cysteine-rich LutA family protein;lactate-related compound oxidative catabolism
ECK3227	***yhcN***	predicted cadmium and peroxide resistance protein;stress-induced
ECK1419	***ydcH***	predicted DUF465 family protein
ECK0773	***moaD***	molybdopterin synthase;small subunit
ECK0502	***glxR***	tartronate semialdehyde reductase;glyoxylate-inducible;NADH-dependent
ECK0157	***btuF***	vitamin B12-binding protein;periplasmic
ECK2421	***ucpA***	predicted short-chain oxidoreductase
ECK2706	***norW***	NADH:flavorubredoxin oxidoreductase
ECK3113	***garR***	tartronate semialdehyde reductase
ECK3177	***sfsB***	DNA-binding NIp family transcription regulator of maltose malPQ operon
ECK0451	***ylaB***	predicted membrane-anchored cyclic-di-GMP phosphodiesterase that controls cell surface-associated traits
ECK0797	***ybiO***	predicted intermediate (MscS family protein) mechanosensitive channel
ECK3712	***yieL***	predicted xylanase;bgl operon;endo-1,4-beta-xylanase homolog
ECK0948	***ompA***	OM (outer membrane) protein A with weak porin activity;T-even phage receptor;homodimer, abundant cell surface protein
ECK3941	***frwB***	predicted fructose family protein PTS system enzyme IIB;component 1
ECK3987	***rsd***	stationary phase protein;binds sigma 70 RNA polymerase subunit
ECK1889	***cheA***	chemotaxis two-component regulatory system sensory histidine kinase protein;responds to signals via interactions with MCPs;CheY is partner response regulator
ECK3880	***dtd***	D-tyrosine tRNA deacylase
ECK2435	***eutC***	ethanolamine ammonia lyase;small subunit;adenosylcobalamine-dependent;concerted induction requires both B12 and ethanolamine;heterodimeric
ECK3991	***yjaG***	predicted DUF416 domain protein
ECK1307	***ycjP***	predicted sugar ABC superfamily transporter permease;may be part of binding-protein-dependent transport system YcjNOP;may be responsible for the translocation of substrate across the membrane
ECK0505	***allB***	allantoinase: allantoin for anaerobic nitrogen;tetrameric
ECK3805	***yigA***	predicted DUF484 family protein
ECK0815	***fsaA***	fructose-6-phosphate aldolase A;aldolase 1
ECK2562	***pdxJ***	pyridoxine 5'-phosphate synthase
ECK0996	***ymdF***	predicted KGG family protein;PubMed:25546632
ECK1334	***abgA***	p-aminobenzoyl-glutamate hydrolase;A subunit
ECK1476	***osmC***	lipoyl-dependent Cys-based peroxidase;hydroperoxide resistance;salt-shock inducible membrane protein;peroxiredoxin
ECK2191	***ccmC***	heme export ABC superfamily transporter permease;CcmABCD transport protein complex;delivers heme to and interacts with CcmE
ECK4228	***mpl***	UDP-N-acetylmuramate:L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase
ECK0880	***lrp***	transcription regulator;leucine-binding;DNA-binding
ECK0445	***amtB***	ammonia gas channel;sequesters GlnK;a negative regulator of AmtB activity;to the inner membrane;bi-directional facilitated diffusion
ECK3079	***alx***	predicted SanA family protein;DUF218 superfamily inner membrane transport protein;alkali-induced riboswitch
ECK2661	***ygaV***	predicted HTH family transcription regulator;tributyltin-inducible repressor of ygaVP
ECK2525	***iscA***	FeS cluster assembly protein
ECK3982	***thiG***	thiamin biosynthesis Thigh complex subunit
ECK4147	***frdD***	fumarate reductase (anaerobic);membrane anchor subunit
ECK3680	***yidQ***	predicted DUF1375 family protein OM (outer membrane) protein
ECK1626	***rsxD***	SoxR iron-sulfur cluster reduction factor component;predicted membrane protein of electron transport complex
ECK3533	***eptB***	KDO phosphoethanolamine transferase;Ca(2+)-inducible;required for Ca(2+) tolerance of heptose-deficient cells
ECK0773	***moaD***	molybdopterin synthase;small subunit
ECK2845	***yqeI***	predicted OmpR family transcription regulator;ToxR homolog;part of T3SS PAI ETT2 remnant
ECK2447	***eutH***	ethanolamine transport
ECK1273	***pgpB***	phosphatidylglycerophosphatase B

ECK4021	***yjbH***	predicted DUF940 family protein extracellular polysaccharide porin protein
ECK0754	***modC***	molybdate ABC superfamily transporter ATPase;chlorate resistance protein
ECK3626	***rpmG***	50S ribosomal subunit protein L33
ECK1081	***fabF***	3-oxoacyl-[acyl-carrier-protein] synthase II
ECK1886	***tap***	methyl-accepting chemotaxis protein (MCP) IV
ECK3145	***yhbT***	SCP-2 sterol transfer family protein
ECK0416	***xseB***	exonuclease VII small subunit
ECK3854	***yihG***	shifted translation start +27 nt;AUG start codon;predicted inner membrane acyltransferase
ECK0800	***glnH***	glutamine-binding protein;periplasmic component
ECK3198	***elbB***	Involved in isoprenoid biosynthesis;also known as sigma cross-reacting protein 27A
ECK4062	***acs***	acetyl CoA synthase;acetate-scavenging enzyme;improves carbon starvation survival
ECK4071	***yjcO***	predicted Sel1 family protein TPR family protein repeat protein
ECK0987	***torC***	trimethylamine N-oxide (TMAO) reductase I;cytochrome c family protein subunit
ECK0050	***apaH***	dinucleoside tetraphosphatase;symmetrical;diadenosine tetraphosphatase;bis(5'-nucleosyl)-tetraphosphatase;symmetrical
ECK3913	***yiiR***	predicted DUF805 family inner membrane protein
ECK3212	***nanE***	predicted N-acetylmannosamine-6-P epimerase
ECK0682	***kdpE***	two-component regulatory system DNA-binding response regulator protein (autophosphorylation site,Asp52);KpdD is partner histidine kinase protein
ECK1410	***cybB***	cytochrome b561
ECK2766	***ygcS***	predicted major facilitator superfamily protein (MFS) membrane transport protein
ECK3836	***fre***	NAD(P) H-flavin reductase;ferrisiderophore reductase C
ECK0474	***ushA***	bifunctional UDP-sugar hydrolase/5'-nucleotidase
ECK3922	***menA***	1,4-dihydroxy-2-naphthoate octaprenyltransferase
ECK1974	***yeeJ***	predicted adhesin;overproduction increases adhesion and biofilm formation;intimin/invasin homolog
ECK2498	***ppx***	exopolyphosphatase
ECK2321	***yfcA***	predicted UPF0721 family inner membrane protein
ECK2965	***yghF***	predicted defective general secretion pathway protein;N-terminal fragment
ECK1721	***pfkB***	6-phosphofructokinase II
ECK0506	***ybbY***	predicted uracil/xanthine transport protein
ECK3100	***yhaO***	predicted amino acid:H+ symport permease transport protein
ECK3256	***yhdX***	predicted amino-acid transport protein subunit;may be part of binding-protein-dependent transport system YdhWXYZ for an amino acid;responsible for the translocation of substrate across the membrane
ECK2862	***xdhA***	xanthine dehydrogenase;molybdenum binding subunit
ECK3459	***dcrB***	predicted lipoprotein;required for phage C1 adsorption;periplasmic;Cys predicted 4'-phosphopantetheinyl transferase
ECK0383	***aroL***	shikimate kinase II
ECK3320	***gspL***	general secretory pathway component;cryptic
ECK2409	***cysK***	cysteine synthase A;O-acetylserine sulfhydrylase A;homodimeric;selenate;azaserine;chromate resistance;alkali-inducible;sulfate starvation-inducible protein SSI5;cysteine desulfhydrase
ECK3348	***pabA***	aminodeoxychorismate synthase;subunit II
ECK3366	***php***	predicted hydrolase
ECK4024	***malG***	maltose ABC superfamily transporter;permease subunit
ECK3410	***glpG***	Rhomboid intramembrane serine protease;complements the phenotypic defects of P;stuartii aarA;mutant is slightly cefotaxime resistant
ECK4110	***adiA***	arginine decarboxylase;acid-inducible;arginine-dependent acid resistance
ECK2191	***ccmC***	heme export ABC superfamily transporter permease;CcmABCD transport protein complex;delivers heme to and interacts with CcmE
ECK0748	***galE***	UDP-galactose 4-epimerase;hexose-1-phosphate uridylyltransferase
ECK4182	***yjfC***	predicted ATP-Grasp family protein ATPase of unknown function;helps mitigate tolC mutant defects;tolC operon
ECK2875	***ssnA***	predicted chlorohydrolase/aminohydrolase
ECK1408	***aldA***	aldehyde dehydrogenase A;NAD-linked
ECK4390	***creB***	two-component regulatory system DNA-binding response regulator protein (autophosphorylation site,Asp54);CreC is partner histidine kinase protein
ECK2486	***hyfJ***	predicted processing element hydrogenase 4
ECK0043	***fixB***	predicted flavoprotein subunit carnitine metabolism required for anaerobic carnitine reduction

ECK3748	***hsrA***	predicted multidrug or homocysteine efflux system
ECK3721	***pstS***	periplasmic phosphate binding of phosphate-specific ABC superfamily transporter
ECK3104	***tdcD***	propionate kinase/acetate kinase C;anaerobic
ECK1038	***mdtG***	predicted drug efflux system
ECK3666	***yidG***	inner membrane protein
ECK0446	***tesB***	acyl-CoA thioesterase II
ECK2747	***cysD***	sulfate adenylyltransferase subunit 2
ECK3797	***hemX***	predicted uroporphyrinogen III methyltransferase
ECK0793	***ybiX***	predicted Fe(II)-dependent oxygenase superfamily protein
ECK0429	***bolA***	stationary-phase morphogene;transcription regulator (repressor) of mreB;dacA;dacC;and ampC;predicted reductase
ECK0015	***dnaJ***	DnaK co-chaperone HSP40;DNA-binding protein;stress-related DNA biosynthesis;responsive to heat shock;binds Zn(II)
ECK0880	***lrp***	transcription regulator;leucine-binding;DNA-binding
ECK4333	***yjiT***	predicted defective predicted protein;N-terminal fragment 488 residue fragment
ECK1644	***ydhL***	predicted DUF1289 family protein
ECK4095	***phnF***	predicted DNA-binding transcription regulator of phosphonate uptake and biodegradation
ECK3806	***xerC***	site-specific tyrosine recombinase
ECK3476	***yhiM***	predicted DUF1323 family protein acid resistance inner membrane protein
ECK0452	***ylaC***	inner membrane protein;DUF1449 family protein
ECK2057	***yegH***	predicted inner membrane protein
ECK2113	***yehP***	predicted VMA domain stimulator of YehL ATPase;required for swarming phenotype
ECK1771	***ydjI***	predicted sugar 1,6-bisphosphate aldolase
ECK1311	***ycjT***	predicted family protein 65 glycosyl hydrolase
ECK1830	***proQ***	RNA chaperone;predicted regulator of ProP translation
ECK2134	***yohJ***	inner membrane protein;UPF0299 family protein
ECK1401	***ynbA***	predicted inner membrane protein;phosphatidylglycerophosphate synthase homolog
ECK1859	***znuC***	ABC superfamily protein zinc transport protein;ATP-binding component
ECK4028	***lamB***	maltose OM (outer membrane) porin (maltoporin)
ECK3813	***rarD***	predicted chloramphenical resistance permease
ECK2715	***hycF***	formate hydrogenlyase complex iron-sulfur protein
ECK3051	***ttdA***	L-tartrate dehydratase;alpha subunit
ECK3321	***gspM***	general secretory pathway component;cryptic
ECK0160	***degP***	membrane-associated periplasmic serine endoprotease;protease Do heat shock protein HtrA;required for high-temperature growth and the degradation of damaged proteins
ECK3697	***yidD***	predicted UPF0161 family protein;may be involved in insertion of integral membrane proteins
ECK3400	***gntX***	DNA catabolic protein for the utilization of DNA as a carbon source;H;influenzae competence protein ComF homolog
ECK0603	***rnk***	regulator of nucleoside diphosphate kinase
ECK3850	***yihD***	predicted DUF1040 family protein YihD
ECK0989	***torD***	TorA-maturation chaperone
ECK0610	***citD***	citrate lyase;acyl carrier (gamma) subunit
ECK4053	***yjcC***	predicted cyclic-di-GMP phosphodiesterase of unknown function;may function as a c-di-GMP phosphodiesterase to control cell surface-associated traits
ECK2142	***mglA***	bifunctional methyl-galactoside transport protein subunits of ABC superfamily protein: ATP-binding components
ECK0561	***nfrB***	bacteriophage N4 receptor;inner membrane subunit
ECK0344	***mhpA***	3-(3-hydroxyphenyl) propionate hydroxylase
ECK4039	***yjbL***	predicted protein
ECK3910	***cdh***	CDP-diacylglycerol phosphotidylhydrolase
ECK3703	***mdtL***	multidrug efflux system protein
ECK4118	***dcuS***	two-component regulatory system sensory histidine kinase protein (autophosphorylation site,His349);senses fumurate and succinate;DcuR is partner response regulator
ECK3213	***nanT***	sialic acid transport protein
ECK4178	***yjfJ***	predicted PspA/IM30 family protein
ECK2885	***lysS***	lysine tRNA synthetase;constitutive
ECK3801	***cyaY***	iron-dependent inhibitor of iron-sulfur cluster formation;frataxin;iron-binding and oxidizing protein

ECK0489	***ybbA***	predicted transport protein subunit: ATP-binding component of ABC superfamily protein
ECK1880	***flhA***	flagellar export pore protein;integral membrane protein
ECK2305	***ubiX***	3-octaprenyl-4-hydroxybenzoate carboxy-lyase
ECK2351	***yfdN***	CPS-53 (KpLE1) cryptic prophage predicted protein
ECK3325	***chiA***	periplasmic endochitinase/lysozyme;Hns-repressible
ECK2314	***pdxB***	erythronate-4-phosphate dehydrogenase
ECK0325	***yahM***	predicted protein
ECK3824	***ysgA***	predicted carboxymethylenebutenolidase
ECK3315	***gspG***	pseudopilin;cryptic;type II secretion pathway protein
ECK1027	***csgB***	curlin nucleator protein;minor subunit in curli complex
ECK0421	***yajR***	predicted transport protein
ECK4335	***mcrC***	5-methylcytosine-specific restriction enzyme McrBC;subunit McrC
ECK0951	***yccS***	predicted FUSC superfamily inner membrane protein
ECK3318	***gspJ***	predicted general secretory pathway component;cryptic
ECK2090	***fbaB***	fructose 1,6-bisphosphate aldolase;class I
ECK3075	***ygjP***	predicted metal dependent hydrolase
ECK1690	***ydiB***	quinate/shikimate 5-dehydrogenase;NAD(P)-binding;NAD-dependent shikimate 5-dehydrogenase
ECK1640	***ydhJ***	predicted membrane fusion protein (MFP) of YdhJK efflux pump
ECK1970	***yodB***	predicted cytochrome b561 homolog
ECK3083	***uxaC***	uronate isomerase
ECK3777	***wzzE***	entobacterial common antigen (ECA) polysaccharide chain length modulation protein
ECK3214	***nanA***	N-acetylneuraminate lyase
ECK1390	***paaF***	2,3-dehydroadipyl-CoA hydratase;enoyl-CoA hydratase;phenylacetic acid degradation
ECK3711	***yieK***	predicted 6-phosphogluconolactonase
ECK0919	***aspC***	aspartate aminotransferase;pyridoxal phosphate (PLP)-dependent
ECK1386	***paaB***	predicted ring 1,2-phenylacetyl-CoA epoxidase subunit
ECK2190	***ccmD***	heme export ABC superfamily transporter holo-CcmE release factor;CcmABCD transport protein complex
ECK3317	***gspI***	general secretory pathway component;cryptic
ECK2945	***yggR***	predicted pilus retraction ATPase
ECK0047	***kefF***	predicted NAD(P) H oxidoreductase potassium-efflux system ancillary protein for KefC glutathione-regulated quinone oxidoreductase FMN-dependent
ECK3745	***rbsB***	D-ribose ABC superfamily transporter periplasmic binding protein;ribose chemotaxis receptor
ECK2943	***yqgE***	predicted protein
ECK1265	***yciK***	predicted COG1028 EmrKY-TolC system oxoacyl-(acyl carrier protein) reductase
ECK0709	***gltA***	citrate synthase
ECK0470	***aes***	short acyl chain and triacetyl glycerol esterase;affects maltose gene expression
ECK4117	***dcuR***	two-component regulatory system DNA-binding response regulator protein (autophosphorylation site,Asp56);DcuS is partner histidine kinase protein
ECK0604	***rna***	ribonuclease I;RNase I;periplasmic endoribonuclease;RNase A;RNase M;RNase T2
ECK0342	***lacI***	DNA-binding transcription repressor of lacZYA operon
ECK3689	***yidA***	erythrose 4-P and mannose 1-P phosphatase;catalyzes dephosphorylation of different sugar phosphates including erythrose-4-phosphate (Ery4P);ribose-5-phosphate (Ribu5P);fructose-1-phosphate (Fru1P);fructose-6-phosphate (Fru6P);glucose-6-P (Glu6P);and also imidodiphosphate (Imido-di-P) and acetyl phosphate (Acetyl-P);Selectively hydrolyzes alpha-D-glucose-1-phosphate (Glu1P) and has no activity with the beta form
ECK3486	***arsR***	DNA-binding ArsR family transcription factor with helix-turn-helix motif
ECK0707	***ybgQ***	predicted OM (outer membrane) fimbrial subunit export usher protein;may be involved in the export and assembly of fimbria
ECK3888	***fdhD***	formate dehydrogenase formation protein
ECK2942	***gshB***	glutathione synthetase
ECK1387	***paaC***	ring 1,2-phenylacetyl-CoA epoxidase subunit
ECK0673	***ybfE***	LexA-regulated protein;CopB family protein
ECK2841	***kdul***	4-deoxy-L-threo-5-hexosulose-uronate ketol-isomerase

Supplementary Table 3. *ftsN* synthetic lethal genes with Keio collection.

Supplementary table 3: Synthetic lethal combination of *ftsN* knockdown induced by 0.2 aTc

ECK_id: eck id of E. coli K-12 gene annotation
name: gene name
description: description of E. coli gene annotation

ECK_id	*name*	description
ECK3781	***rffH***	glucose-1-phosphate thymidylyltransferase
ECK2478	***hyfB***	hydrogenase 4;membrane subunit
ECK0207	***dkgB***	2;5-diketo-D-gluconate reductase B;glyoxal reductase;glyoxal resistance protein
ECK1641	***ydhK***	predicted efflux PET family protein component of YdhJK efflux pump
ECK2333	***yfcV***	predicted fimbrial family protein adhesin protein of yfcOPQRSUV fimbrial operon
ECK3800	***cyaA***	adenylate cyclase
ECK3724	***atpC***	F1 sector of membrane-bound ATP synthase;epsilon subunit
ECK3402	***gntT***	high-affinity gluconate transport
ECK4166	***mutL***	methyl-directed mismatch repair protein
ECK3383	***mrcA***	bifunctional penicillin-binding protein 1a: murein transglycosylase/murein transpeptidase
ECK2393	***yfeD***	predicted DUF1323 YfeC/YfeD family DNA-binding transcription regulator;N-terminal fragment HTH domain of MerR superfamily protein
ECK1519	***yneJ***	predicted LysR family protein DNA-binding transcription regulator of lamB
ECK0252	***ykfB***	CP4-6 cryptic prophage predicted protein
ECK4008	***aceK***	Isocitrate lyase;acetate utilization;glyoxylate shunt;tetrameric Isocitrate dehydrogenase kinase/phosphatase
ECK0116	***yacH***	predicted DUF3300 family protein membrane protein
ECK0240	***frsA***	Fermentation-respiration switch protein;PTS Enzyme IIAGlc-binding protein;binds unphosphorylated form;has pNP-butyrate esterase activity in vitro
ECK3642	***recG***	ATP-dependent DNA helicase
ECK2887	***recJ***	ssDNA exonuclease;5' -> 3'-specific
ECK0971	***appA***	acid phosphatase;pH 2;5;exopolyphosphatase;phytase
ECK3731	***atpB***	F0 sector of membrane-bound ATP synthase;subunit a
ECK1779	***yeaE***	aldo-keto reductase;methylglyoxal to acetol;NADPH-dependent
ECK3512	***yhjJ***	predicted periplasmic M16 family protein chaperone;zinc-dependent peptidase
ECK4324	***yjiK***	SdiA-regulated family protein;membrane-anchored protein;predicted phytase family protein esterase
ECK0247	***yafW***	CP4-6 cryptic prophage antitoxin of YkfI-YafW toxin-antitoxin system
ECK0677	***ybfP***	predicted/verified lipoprotein
ECK3394	***yhgF***	predicted transcription accessory protein
ECK0985	***torT***	periplasmic sensory protein associated with TorS/TorR two-component regulatory system
ECK4233	***nrdD***	anaerobic ribonucleoside-triphosphate reductase
ECK2710	***ascF***	bifunctional cellobiose/arbutin/salicin-specific phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) enzymes: IIB component/IC component
ECK3422	***gntK***	gluconokinase 2;thermoresistant
ECK0739	***nadA***	quinolinate synthase;[4Fe-4S] cluster subunit, A protein
ECK2903	***pepP***	proline aminopeptidase P II
ECK0123	***gcd***	glucose dehydrogenase (pyrroloquinoline-quinone) quinoprotein;inner membrane protein
ECK1445	***yncD***	predicted TonB-dependent OM (outer membrane) iron transport receptor protein
ECK0490	***ybbP***	predicted ABC superfamily transporter permease
ECK3488	***arsC***	arsenate reductase
ECK3487	***arsB***	arsenite/antimonite transport protein
ECK3347	***argD***	acetylornithine aminotransferase;succinyldiaminopimelate aminotransferase;pyridoxal phosphate (PLP)-dependent
ECK3459	***acpT***	predicted lipoprotein;required for phage C1 adsorption;periplasmic;Cys predicted 4'-phosphopantetheinyl transferase
ECK3634	***yicC***	predicted UPF0701 family protein
ECK3508	***yhjE***	predicted inner membrane transport protein
ECK3551	***yiaA***	predicted YiaAB family inner membrane protein
ECK2495	***purM***	phosphoribosylaminoimidazole synthetase

ECK0150	***fhuC***	ABC superfamily transporter ATPase;Fe(3+)-ferrichrome uptake
ECK0878	***cydD***	glutathione;cysteine export permease/ATP-binding protein;ABC superfamily transporter;N-terminal fragment permease domain and C-terminal fragment cytoplasmic ATP-binding domain
ECK0199	***metN***	ABC superfamily transporter ATPase;L;D-methionine uptake;methionine sulfoximine sensitivity
ECK1507	***lsrC***	autoinducer 2 import system permease protein
ECK2428	***yfeZ***	predicted inner membrane protein
ECK0742	***ybgS***	predicted periplasmic protein
ECK2643	***ypjF***	CP4-57 cryptic prophage toxin of YpjF-YfjZ toxin-antitoxin system
ECK4356	***bglJ***	LuxR family protein transcription regulator (activator) of silent bgl operon
ECK1176	***ycgB***	predicted SpoVR family protein stationary phase protein
ECK1558	***relB***	antitoxin of RelE-RelB toxin-antitoxin system;transcription regulator
ECK0232	***dinB***	DNA polymerase IV;capable of translesion synthesis;overproduction enhances mutagenesis;mediates targeted mutagenesis by 4-NQO;intrinsic AP lyase activity
ECK1939	***fliI***	cytoplasmic membrane ATPase involved in flagellar assembly;involved in export of flagellar axial protein subunits
ECK4009	***arpA***	ankyrin repeat protein
ECK3187	***kdsC***	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase
ECK2016	***hisC***	histidinol-phosphate aminotransferase
ECK1912	***uvrC***	excinuclease UvrABC;endonuclease subunit C
ECK4395	***yjtD***	predicted rRNA methyltransferase
ECK1028	***csgA***	curlin subunit;amyloid curli fibers;cryptic
ECK2322	***mepA***	murein DD-endopeptidase
ECK4017	***pgi***	glucosephosphate isomerase
ECK2269	***yfbP***	TPR family protein repeats-containing protein
ECK1682	***ydiH***	predicted protein
ECK0213	***yafS***	predicted S-adenosyl-L-methionine (SAM)-dependent methyltransferase
ECK0600	***ahpF***	alkyl hydroperoxide reductase;F52a subunit;FAD/NAD(P)-binding
ECK0645	***gltL***	glutamate/aspartate ABC superfamily transporter ATPase subunit
ECK3335	***slyX***	phi X174 lysis protein
ECK3498	***mdtF***	anaerobic multidrug efflux transport protein;ArcA-regulated
ECK3428	***yhhZ***	predicted Hcp1 family protein polymorphic toxin protein with colicin family protein DNAse/tRNase activity
ECK2229	***inaA***	predicted Kdo/WaaP family protein kinase;acid-inducible
ECK1300	***pspB***	transcription regulator of psp operon;DNA-binding
ECK1472	***adhP***	ethanol-active dehydrogenase/acetaldehyde-active reductase
ECK2488	***focB***	predicted bidirectional formate transport protein 2
ECK0712	***sdhA***	succinate dehydrogenase;flavoprotein subunit
ECK1560	***flxA***	Qin prophage member of FliA regulon
ECK0148	***mrcB***	fused glycosyl transferase and transpeptidase peptidoglymay synthetase penicillin-binding protein 1B
ECK0643	***hscC***	Hsp70 family protein chaperone Hsc62;RpoD-binding transcription inhibitor
ECK1926	***amyA***	flagellar synthesis;predicted chaperone;role unknown alpha-amylase;cytoplasmic
ECK1462	***narZ***	nitrate reductase 2 (NRZ);alpha subunit
ECK1516	***yneG***	predicted DUF4186 family protein
ECK1577	***ynfA***	predicted UPF0060 family inner membrane protein
ECK0013	***yaaI***	predicted UPF0412 family protein periplasmic protein transport protein
ECK2461	***ypfG***	predicted DUF1176 family protein
ECK1639	***ydhI***	predicted DUF1656 family inner membrane efflux pump associated protein
ECK3357	***yhfL***	predicted small lipoprotein
ECK3208	***yhcF***	predicted transcription regulator
ECK3528	***dppD***	dipeptide/heme transport;ATP-binding protein;also transports 5-aminolevulinic acid
ECK0549	***borD***	DLP12 cryptic prophage;predicted lipoprotein involved in bacterial virulence
ECK0353	***frmA***	Glutathione-dependent formaldehyde dehydrogenase
ECK3872	***yihR***	predicted aldose-1-epimerase
ECK2954	***yggL***	predicted DUF469 family protein
ECK4092	***phnI***	carbon-phosphorus lyase complex subunit
ECK2982	***gss***	Glutathionylspermidine synthase/amidase;bifunctional protein
ECK3951	***argH***	argininosuccinate lyase

ECK0241	***crl***	sigma factor-binding protein;stimulates RNAP holoenzyme formation and RpoS activity during stationary phase
ECK0956	***yccU***	predicted CoA-binding protein with NAD(P)-binding domain
ECK4127	***cadC***	transcription regulator (activator) of cadBA operon
ECK0307	***ykgG***	predicted LutC family protein electron transport chain YkgEFG component
ECK0935	***ycbF***	predicted periplasmic pilini chaperone of elfADCG-ycbUVF fimbrial operon;promotes adhesion of bacteria to different abiotic surfaces;may be required for the biogenesis of fimbriae
ECK0477	***ybaQ***	predicted DNA-binding YbaQ family transcription regulator
ECK2195	***napB***	nitrate reductase;small;cytochrome C550 subunit;periplasmic
ECK2076	***yegP***	predicted UPF0339 family protein
ECK2366	***evgS***	two-component regulatory system hybrid sensory histidine kinase protein (autophosphorylation sites:His721, Asp1009, and His1137);signal unknown;EvgA is partner response regulator
ECK3710	***cbrC***	predicted;UPF0167 family protein glutaredoxin/thioredoxin thiol-disulfide oxidoreductase protein;required for colicin E2 tolerance
ECK3378	***hofQ***	DNA catabolic protein;required for the utilization of DNA as a carbon source;fimbrial transport protein homolog
ECK0591	***cstA***	carbon starvation protein involved in peptide utilization
ECK4018	***yjbE***	extracellular polysaccharide production threonine-rich protein
ECK1060	***flgD***	flagellar hook assembly protein
ECK2369	***oxc***	oxalyl CoA decarboxylase;ThDP-dependent
ECK0651	***ybeY***	metal-binding heat shock protein;required for rRNA maturation;ssRNA-specific endoribonuclease;16S rRNA 3' end maturation and quality control co-endoribonuclease working with RNase R;rRNA transcription antitermination factor;involved in late-stage 70S ribosome quality control and maturation of 3' terminus of 16S rRNA
ECK2985	***hybF***	protein involved with the maturation of hydrogenases 1 and 2
ECK4103	***yjcZ***	mutational suppressor of yhjH motility mutation of unknown function
ECK0924	***ssuB***	alkanesulfonate transport protein ATPase subunit
ECK3392	***ompR***	two-component regulatory system DNA-binding response regulator protein (autophosphorylation site,Asp55);EnvZ is partner histidine kinase protein
ECK3162	***yhbX***	predicted EptAB family protein phosphoethanolamine transferase;inner membrane protein
ECK1167	***ycgL***	predicted UPF0745 family protein
ECK3763	***ilvD***	dihydroxyacid dehydratase
ECK1107	***ycfZ***	predicted inner membrane protein
ECK3867	***yihN***	predicted inner membrane transport protein
ECK4037	***yjbJ***	predicted UPF0337 family protein stress-induced protein
ECK4011	***metH***	homocysteine-N5-methyltetrahydrofolate transmethylase;B12-dependent
ECK0868	***ybjX***	shifted translation start +42 nt;AUG start codon;predicted DUF535 family protein
ECK2991	***hybO***	hydrogenase 2;small subunit
ECK1049	***grxB***	glutaredoxin 2 (Grx2)
ECK0968	***hyaF***	hydrogenase-1 protein nickel incorporation factor
ECK0435	***ppiD***	periplasmic folding chaperone;has an inactive PPIase domain
ECK3944	***frwD***	predicted fructose family protein PTS system enzyme IIB;component 2
ECK0080	***ilvH***	acetolactate synthase III thiamin-dependent small subunit acetolactate synthase III valine sensitive small subunit
ECK3123	***agaA***	predicted defective N-acetyl-D-galactosamine-6-phosphate deacetylase;C-terminal fragment
ECK3514	***yhjK***	predicted defective cyclic-di-GMP phosphodiesterase associated with cellulose production
ECK3358	***frlA***	predicted fructoselysine transport protein
ECK0336	***cynT***	carbonic anhydrase;beta class
ECK0008	***talB***	transaldolase B trans-aconitate 2-methyltransferase;SAM (S-adenosyl-L-methionine)-dependent
ECK4047	***aphA***	acid phosphatase/phosphotransferase;class B;non-specific
ECK2243	***yfaZ***	predicted OM (outer membrane) lipoprotein;may be porin
ECK3222	***yhcB***	predicted DUF1043 family inner membrane-anchored protein
ECK2813	***amiC***	N-acetylmuramyl-L-alanine amidase;periplasmic;recruited to the septal ring by FtsN during cell division;overproduction causes lysis;activated by NlpD

ECK2095	***yegX***	shifted translation start -9 nt;AUG start codon;predicted family protein 25 glycosyl hydrolase
ECK3911	***tpiA***	triosephosphate isomerase
ECK4132	***dcuA***	C4-dicarboxylate antiporter;anaerobic
ECK3571	***sgbU***	predicted L-xylulose 5-phosphate 3-epimerase;required for the aerobic utilization of L-ascorbate but not for fermentation;L-xylulose catabolism
ECK4196	***rpsF***	30S ribosomal subunit protein S6
ECK4061	***yjcH***	predicted DUF485 family inner membrane protein
ECK0959	***yccX***	predicted acylphosphatase;weak acylphosphatase activity
ECK4342	***yjiA***	predicted metal-binding GTPase;binds GTP;may function as GTP-dependent regulator
ECK1175	***fadR***	transcription regulator;iclR family protein repressor and activator;of fabAB regulon
ECK1097	***ycfQ***	HTH family TetR family transcription regulator (repressor) of bhsA(ycfR);copper OM (outer membrane) regulator
ECK1881	***flhB***	flagellin export apparatus;substrate specificity protein;determines the order of subunit export
ECK1309	***ycjR***	predicted TIM alpha/beta barrel enzyme
ECK0614	***dcuC***	anaerobic C4-dicarboxylate transport
ECK0142	***pcnB***	poly(A) polymerase I
ECK3006	***yqhH***	predicted OM (outer membrane) lipoprotein;Lpp paralog
ECK4091	***phnJ***	carbon-phosphorus lyase complex;SAM (S-adenosyl-L-methionine)-dependent
ECK4086	***phnO***	predicted acyltransferase with acyl-CoA N-acyltransferase domain
ECK0193	***yaeF***	predicted lipoprotein
ECK2528	***iscR***	transcription regulator of isc operon;DNA-binding
ECK2974	***glcD***	glycolate oxidase subunit
ECK1486	***gadC***	glutamate:gamma-aminobutyric acid antiporter
ECK3603	***envC***	activator of AmiB,C murein hydrolases;septal ring factor
ECK2772	***ygcG***	predicted TPM domain phosphatase protein
ECK1479	***ddpC***	D;D-dipeptide ABC superfamily transporter;membrane translocase
ECK2559	***yfhH***	predicted RpiR family DNA-binding transcription regulator
ECK2531	***yfhR***	predicted S9 family protein prolyl oligopeptidase
ECK2531	***yfhR***	predicted S9 family protein prolyl oligopeptidase
ECK3438	***livF***	branched-chain amino acid ABC superfamily transporter ATPase
ECK1606	***fumC***	fumarate hydratase (fumarase C);aerobic Class II
ECK2946	***yggS***	predicted UPF0001 family protein;pyridoxal phosphate (PLP)-binding;mutant has perturbed isoleucine/valine metabolism
ECK3201	***yhcC***	predicted Fe-S oxidoreductase;radical SAM (S-adenosyl-L-methionine) superfamily protein
ECK4085	***phnP***	carbon-phosphorus lyase complex accessory protein;5-phospho-alpha-D-ribosyl 1,2-cyclic phosphate phosphodiesterase
ECK2767	***ygcU***	predicted FAD containing dehydrogenase;may it be related to PAF PubMed: 22771767
ECK0567	***cusA***	copper/silver efflux system;membrane component
ECK2855	***ygeN***	predicted defective protein of type three secretion system (T3SS) pathogenicity island ETT2
ECK1247	***yciA***	predicted acyl CoA esterase;catalyzes the hydrolysis of thioester bond in palmitoyl-CoA and malonyl-CoA
ECK0006	***yaaA***	predicted UPF0246 Grp1_Fun34_YaaH family inner membrane protein;peroxide resistance protein
ECK2730	***ygbI***	predicted DeoR family transcription regulator
ECK1405	***azoR***	NADH-azoreductase;FMN-dependent
ECK0447	***ybaY***	predicted/verified lipoprotein
ECK2990	***hybA***	hydrogenase 2 4Fe-4S ferredoxin family protein component
ECK4150	***frdA***	fumarate reductase and NAD/flavoprotein flavoprotein subunit
ECK0730	***pal***	peptidoglycan-associated OM (outer membrane) lipoprotein
ECK1082	***pabC***	4-amino-4-deoxychorismate lyase component of para-aminobenzoate synthase multienzyme complex
ECK1988	***cobU***	bifunctional cobinamide kinase and cobinamide phosphate guanylyltransferase
ECK2814	***argA***	N-acetylglutamate synthase;first step in arginine biosynthesis;amino-acid acetyltransferase;L-glutamate N-acetyltransferase;growth on acetylornithine;feedback inhibited by L-arginine
ECK4306	***fimI***	fimbrial protein involved in type 1 pilus biosynthesis;FimA homolog
ECK3840	***yigZ***	predicted UPF0029 IMPACT family protein

References

1. Rea SL, Graham BH, Nakamaru-Ogiso E, Kar A, Falk MJ. Bacteria, yeast, worms, and flies: Exploiting simple model organisms to investigate human mitochondrial diseases. *Dev Disabil Res Rev*. 2010;16(2):200-218. doi:10.1002/ddrr.114.

2. Mori Y, Sato Y, Takamatsu S. Molecular phylogeny and radiation time of Erysiphales inferred from the nuclear ribosomal DNA sequences. *Mycoscience*. 2000;41(5):437-447. doi:10.1007/BF02461662.

3. Kelley R, Ideker T. Systematic interpretation of genetic interactions using protein networks. *Nat Biotechnol*. 2005;23(5):561-566. doi:10.1038/nbt1096.

4. Karas BJ, Suzuki Y, Weyman PD. Strategies for cloning and manipulating natural and synthetic chromosomes. *Chromosom Res*. 2015;23(1):57-68. doi:10.1007/s10577-014-9455-3.

5. Baba T, Ara T, Hasegawa M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006;2:2006.0008. doi:10.1038/msb4100050.

6. Davierwala AP, Haynes J, Li Z, et al. The synthetic genetic interaction spectrum of essential genes. *Nat Genet*. 2005;37(10):1147-1152. doi:10.1038/ng1640.

7. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata a. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol*. 1987;169(12):5429-5433. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=213968&tool=pmcentrez&rendertype=abstract.

8. Wiedenheft B, Sternberg SH, Doudna J a. RNA-guided genetic silencing systems in bacteria and archaea. *Nature*. 2012;482(7385):331-338. doi:10.1038/nature10886.

9. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna J a, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821. doi:10.1126/science.1225829.

10. Qi LS, Larson MH, Gilbert L a, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 2013;152(5):1173-1183. doi:10.1016/j.cell.2013.02.022.

11. Escherich T. The intestinal bacteria of the neonate and breast-fed infant. *Rev Infect Dis*. 1989;11(2):352-356.

12. Shulman ST, Friedmann HC, Sims RH. Theodor Escherich: The First Pediatric Infectious Diseases Physician? *Clin Infect Dis*. 2007;45(8):1025-1029. http://dx.doi.org/10.1086/521946.

13. Bordet J. The theories of blood coagulation. *Bull Johns Hopkins Hosp*. 1921;32:213.

14. Werkman CH, Weaver HJ. Studies in the bacteriology of sulphur stinker spoilage of canned sweet corn. *State Coll J Sci*. 1927;2:57-67.

15. Wollman E. Recherches sur la bactériophagie (phénomène de Twort-d'Hérelle). *Ann Inst Pasteur*. 1925;39:789-832.

16. Wollman E, Giroud A, Ratsimamanga R. Synthesis of vitamin C in an orthopterous insect, Blattella germanica, reared aseptically. *Compte rendu des seances la Soc Biol*. 1937;124:434-435.

17. Bronfenbrenner JJ, Korb C. Studies on the bacteriophage of d'Herelle. *J Exp Med*. 1925;42(4):483-497.

18. Stoner EC, 1899-1968. Magnetism and matter. 1934.

19. <Tatum-1947-Gene Recombination in the Bacterium.pdf>.

20. Moore JH, Editors SMW, Walker JM. *Epistasis Series Editor*.

21. Friedmann HC. From "butyribacterium" to "E. coli": an essay on unity in biochemistry. *Perspect Biol Med*. 2004;47(1):47-66.

22. Cohen SN, Chang ACY, Boyer HW, Helling RB. Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc Natl Acad Sci U S A*. 1973;70(11):3240-3244. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC427208/.

23. Russo TA, Johnson JR. Medical and economic impact of extraintestinal infections due to Escherichia coli: focus on an increasingly important endemic problem. *Microbes Infect*. 2003;5(5):449-456. doi:https://doi.org/10.1016/S1286-4579(03)00049-2.

24. Benninghoven A. Chemical analysis of inorganic and organic surfaces and thin films by static time-of-flight secondary ion mass spectrometry (TOF-SIMS). *Angew Chemie Int Ed*. 1994;33(10):1023-1043.

25. Cornelis GR, Van Gijsegem F. Assembly and Function of Type III Secretory Systems. *Annu Rev Microbiol*. 2000;54(1):735-774. doi:10.1146/annurev.micro.54.1.735.

26. Savageau MA. Introduction to S-systems and the underlying power-law formalism. *Math Comput Model*. 1988;11:546-551. doi:https://doi.org/10.1016/0895-7177(88)90553-5.

27. Venter JC, Adams MD, Myers EW, et al. The Sequence of the Human Genome. *Science (80-)*. 2001;291(5507):1304 LP-1351. http://science.sciencemag.org/content/291/5507/1304.abstract.

28. Pertea M, Salzberg SL. Between a chicken and a grape: estimating the number of human genes. *Genome Biol*. 2010;11(5):206. doi:10.1186/gb-2010-11-5-206.

29. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921. doi:10.1038/35057062.

30. DiGregorio BE. Biobased Performance Bioplastic: Mirel. *Chem Biol*. 2009;16(1):1-2.

doi:https://doi.org/10.1016/j.chembiol.2009.01.001.

31. Peoples OP, Sinskey AJ. Poly-??-hydroxybutyrate (PHB) biosynthesis in Alcaligenes eutrophus H16. Identification and characterization of the PHB polymerase gene (phbC). *J Biol Chem*. 1989;264(26):15298-15303.

32. Csete ME, Doyle JC. Reverse Engineering of Biological Complexity. *Science (80-)*. 2002;295(5560):1664 LP-1669. http://science.sciencemag.org/content/295/5560/1664.abstract.

33. Kitano H. Biological robustness. *Nat Rev Genet*. 2004;5:826. http://dx.doi.org/10.1038/nrg1471.

34. Wadhams GH, Armitage JP. Making sense of it all: bacterial chemotaxis. *Nat Rev Mol Cell Biol*. 2004;5:1024. http://dx.doi.org/10.1038/nrm1524.

35. Salama NR, Shepherd B, Falkow S. Global Transposon Mutagenesis and Essential Gene Analysis of Helicobacter pylori . *J Bacteriol*. 2004;186(23):7926-7935. doi:10.1128/JB.186.23.7926-7935.2004.

36. Gerdes S, Scholle MD, Campbell JW, et al. Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. *J Bacteriol*. 2003;185(19):5673-5684. doi:10.1128/JB.185.19.5673.

37. Giaever G, Chu AM, Ni L, et al. Functional profiling of the Saccharomyces cerevisiae genome. *Nature*. 2002;418:387. http://dx.doi.org/10.1038/nature00935.

38. Kobayashi K, Ehrlich SD, Albertini A, et al. Essential Bacillus subtilis genes. *Proc Natl Acad Sci*. 2003;100(8):4678-4683. doi:10.1073/pnas.0730515100.

39. Xu P, Ge X, Chen L, et al. Genome-wide essential gene identification in Streptococcus sanguinis. *Sci Rep*. 2011;1:125. http://dx.doi.org/10.1038/srep00125.

40. de Berardinis V, Vallenet D, Castelli V, et al. A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1. *Mol Syst Biol*. 2008;4(1). http://msb.embopress.org/content/4/1/174.abstract.

41. Jordan IK, Rogozin IB, Wolf YI, Koonin E V. Essential Genes Are More Evolutionarily Conserved Than Are Nonessential Genes in Bacteria Essential Genes Are More Evolutionarily Conserved Than Are Nonessential Genes in Bacteria. *Genome Res*. 2002:962-968. doi:10.1101/gr.87702.

42. Butland G, Babu M, Díaz-Mejía JJ, et al. eSGA: E. coli synthetic genetic array analysis. *Nat Methods*. 2008;5(9):789-795.

43. Kitagawa S, Uemura K. Dynamic porous properties of coordination polymers inspired by hydrogen bonds. *Chem Soc Rev*. 2005;34(2):109-119. doi:10.1039/B313997M.

44. Kang Y, Kang Y, Durfee T, et al. Systematic Mutagenesis of the. *Microbiology*. 2004;186(15):4921-4930. doi:10.1128/JB.186.15.4921.

45. Typas A, Nichols RJ, Siegele DA, et al. High-throughput , quantitative analyses of

genetic interactions in E . coli. 2008;5(9):781-787. doi:10.1038/NMETH.1240.

46. Fish J, Chen W, Nagai G. Non-local dispersive model for wave propagation in heterogeneous media: One-dimensional case. *Int J Numer Methods Eng*. 2002;54:331-346. doi:10.1002/nme.423.

47. Bateson W, Waunders ER, Punnett RC. Experimental studies in the physiology of heredity. *Mol Gen Genet MGG*. 1909;2(1):17-19.

48. Phillips PC. The Language of Gene Interaction. *Genetics*. 1998;149(3):1167 LP-1171. http://www.genetics.org/content/149/3/1167.abstract.

49. Kauffman SA. The origin of order. 1993.

50. Charlesworth B. *Elements of Evolutionary Genetics*. Roberts Publishers; 2010.

51. Azevedo RBR, Lohaus R, Srinivasan S, Dang KK, Burch CL. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature*. 2006;440(7080):87-90.

52. DePristo MA, Weinreich DM, Hartl DL. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat Rev Genet*. 2005;6(9):678-687.

53. Bornscheuer UT, Huisman GW, Kazlauskas RJ, Lutz S, Moore JC, Robins K. Engineering the third wave of biocatalysis. *Nature*. 2012;485(7397):185-194.

54. Domingo E, Sheldon J, Perales C. Viral quasispecies evolution. *Microbiol Mol Biol Rev*. 2012;76(2):159-216.

55. Hmelo-Silver CE, Azevedo R. Understanding Complex Systems: Some Core Challenges. *J Learn Sci*. 2006;15(1):53-61. doi:10.1207/s15327809jls1501_7.

56. Bonhoeffer S, Chappey C, Parkin NT, Whitcomb JM, Petropoulos CJ. Evidence for positive epistasis in HIV-1. *Science (80-)*. 2004;306(5701):1547-1550.

57. Dame TM, Hartmann D, Thaddeus P. The Milky Way in molecular clouds: a new complete CO survey. *Astrophys J*. 2001;547(2):792.

58. Schuldiner M, Collins SR, Thompson NJ, et al. Exploration of the function and organization of the yeast early secretory pathway through an epistatic miniarray profile. *Cell*. 2005;123(3):507-519. doi:10.1016/j.cell.2005.08.031.

59. Segrè D, DeLuna A, Church GM, Kishony R. Modular epistasis in yeast metabolism. *Nat Genet*. 2004;37:77. http://dx.doi.org/10.1038/ng1489.

60. Collins SR, Miller KM, Maas NL, et al. Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature*. 2007;446:806. http://dx.doi.org/10.1038/nature05649.

61. Winzeler EA, Shoemaker DD, Astromoff A, et al. Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. *Science (80-)*. 1999;285(5429):901-906.

62. Fraser N. Rethinking recognition. *New left Rev*. 2000;3:107.

63. Dietzl G, Chen D, Schnorrer F, et al. A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature*. 2007;448(7150):151-156.

64. Tong AHY, Evangelista M, Parsons AB, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science (80-)*. 2001;294(5550):2364-2368.

65. Ooi SL, Pan X, Peyser BD, et al. Global synthetic-lethality analysis and yeast functional profiling. *Trends Genet*. 2006;22(1):56-63.

66. Takeuchi R, Tamura T, Nakayashiki T, et al. Colony-live --a high-throughput method for measuring microbial colony growth kinetics-- reveals diverse growth effects of gene knockouts in Escherichia coli. *BMC Microbiol*. 2014;14(1):171. doi:10.1186/1471-2180-14-171.

67. van Soolingen D, Hermans PW, De Haas PE, Soll DR, Van Embden JD. Occurrence and stability of insertion sequences in Mycobacterium tuberculosis complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol*. 1991;29(11):2578-2586.

68. Mojica FJM, Ferrer C, Juez G, Rodríguez-Valera F. Long stretches of short tandem repeats are present in the largest replicons of the Archaea Haloferax mediterranei and Haloferax volcanii and could be involved in replicon partitioning. *Mol Microbiol*. 1995;17(1):85-93.

69. Barrangou R, Fremaux C, Deveau H, et al. CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science (80-)*. 2007;315(5819):1709 LP-1712. http://science.sciencemag.org/content/315/5819/1709.abstract.

70. Deveau H, Barrangou R, Garneau JE, et al. Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J Bacteriol*. 2008;190(4):1390-1400. doi:10.1128/JB.01412-07.

71. Gudbergsdottir S, Deng L, Chen Z, et al. Dynamic properties of the Sulfolobus CRISPR/Cas and CRISPR/Cmr systems when challenged with vector-borne viral and plasmid genes and protospacers. *Mol Microbiol*. 2011;79(1):35-49. doi:10.1111/j.1365-2958.2010.07452.x.

72. Garrett RA, Vestergaard G, Shah SA. Archaeal CRISPR-based immune systems: exchangeable functional modules. *Trends Microbiol*. 2011;19(11):549-556. doi:https://doi.org/10.1016/j.tim.2011.08.002.

73. Terns MP, Terns RM. CRISPR-based adaptive immune systems. *Curr Opin Microbiol*. 2011;14(3):321-327. doi:https://doi.org/10.1016/j.mib.2011.03.005.

74. Haft DH, Selengut J, Mongodin EF, Nelson KE. A Guild of 45 CRISPR-Associated (Cas) Protein Families and Multiple CRISPR/Cas Subtypes Exist in Prokaryotic Genomes. *PLOS Comput Biol*. 2005;1(6):e60. https://doi.org/10.1371/journal.pcbi.0010060.

75. Wiedenheft B, Zhou K, Jinek M, Coyle SM, Ma W, Doudna JA. Structural Basis for

DNase Activity of a Conserved Protein Implicated in CRISPR-Mediated Genome Defense. *Structure*. 2009;17(6):904-912. doi:https://doi.org/10.1016/j.str.2009.03.019.

76. Deveau H, Garneau JE, Moineau S. CRISPR/Cas System and Its Role in Phage-Bacteria Interactions. *Annu Rev Microbiol*. 2010;64(1):475-493. doi:10.1146/annurev.micro.112408.134123.

77. Carte J, Wang R, Li H, Terns RM, Terns MP. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev*. 2008;22(24):3489-3496. doi:10.1101/gad.1742908.

78. Lillestøl RK, Shah SA, Brügger K, et al. CRISPR families of the crenarchaeal genus Sulfolobus: Bidirectional transcription and dynamic properties. *Mol Microbiol*. 2009;72(1):259-272. doi:10.1111/j.1365-2958.2009.06641.x.

79. Agari Y, Sakamoto K, Tamakoshi M, Oshima T, Kuramitsu S, Shinkai A. Transcription Profile of Thermus thermophilus CRISPR Systems after Phage Infection. *J Mol Biol*. 2010;395(2):270-281. doi:https://doi.org/10.1016/j.jmb.2009.10.057.

80. Garneau JE, Dupuis M-È, Villion M, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature*. 2010;468:67. http://dx.doi.org/10.1038/nature09523.

81. Pennisi E. The CRISPR craze. *Science (80-)*. 2013;341(6148):833-836.

82. Semenova E, Jore MM, Datsenko KA, et al. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc Natl Acad Sci*. 2011;108(25):10098-10103. doi:10.1073/pnas.1104144108.

83. Datsenko KA, Pougach K, Tikhonov A, Wanner BL, Severinov K, Semenova E. Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nat Commun*. 2012;3:945. http://dx.doi.org/10.1038/ncomms1937.

84. Lillestøl R, Redder P, Garrett RA, Brügger KIM. A putative viral defence mechanism in archaeal cells. *Archaea*. 2006;2(1):59-72.

85. Mojica FJM, García-Martínez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol*. 2005;60(2):174-182.

86. Makarova KS, Wolf YI, Alkhnbashi OS, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol*. 2015;13(11):722-736.

87. Koonin E V, Makarova KS, Zhang F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol*. 2017;37:67-78. doi:https://doi.org/10.1016/j.mib.2017.05.008.

88. Sinkunas T, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J*. 2011;30(7):1335-1342.

89. Swarts DC, Mosterd C, van Passel MWJ, Brouns SJJ. CRISPR interference directs

strand specific spacer acquisition. *PLoS One*. 2012;7(4):e35888. doi:10.1371/journal.pone.0035888.

90. Brouns SJJ, Jore MM, Lundgren M, et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*. 2008;321(5891):960-964. doi:10.1126/science.1159689.

91. Ran FA, Hsu PD, Wright J, Agarwala V, Scott D a, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-2308. doi:10.1038/nprot.2013.143.

92. Hsu PD, Scott D a, Weinstein J a, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31(9):827-832. doi:10.1038/nbt.2647.

93. Jiang W, Bikard D, Cox D, Zhang F, Marraffini L a. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*. 2013;31(3):233-239. doi:10.1038/nbt.2508.

94. Yu Z, Ren M, Wang Z, et al. Highly Efficient Genome Modifications Mediated by CRISPR/Cas9 in Drosophila *Genetics*. 2013;195(1):289 LP-291. http://www.genetics.org/content/195/1/289.abstract.

95. Zhu X, Ye K. Crystal structure of Cmr2 suggests a nucleotide cyclase-related enzyme in type III CRISPR-Cas systems. *FEBS Lett*. 2012;586(6):939-945. doi:10.1016/j.febslet.2012.02.036.

96. Hatoum-Aslan A, Samai P, Maniv I, Jiang W, Marraffini L a. A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. *J Biol Chem*. 2013;288(39):27888-27897. doi:10.1074/jbc.M113.499244.

97. Makarova KS, Aravind L, Wolf YI, Koonin E V. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol Direct*. 2011;6(1):38. doi:10.1186/1745-6150-6-38.

98. Vestergaard G, Garrett RA, Shah SA. CRISPR adaptive immune systems of Archaea. *RNA Biol*. 2014;11(2):156-167. doi:10.4161/rna.27990.

99. Schunder E, Rydzewski K, Grunow R, Heuner K. First indication for a functional CRISPR/Cas system in Francisella tularensis. *Int J Med Microbiol*. 2013;303(2):51-60. doi:https://doi.org/10.1016/j.ijmm.2012.11.004.

100. Chylinski K, Makarova KS, Charpentier E, Koonin E V. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Res*. 2014;42(10):6091-6105. doi:10.1093/nar/gku241.

101. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The Streptococcus thermophilus CRISPR/Cas system provides immunity in Escherichia coli. *Nucleic Acids Res*. 2011;39(21):9275-9282. doi:10.1093/nar/gkr606.

102. Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc Natl Acad Sci U S A*. 2012;109(39):E2579-86. doi:10.1073/pnas.1208507109.

103. Ran FA, Cong L, Yan WX, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015. doi:10.1038/nature14299.

104. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. *Nat Methods*. 2013;10(10):957-963. doi:10.1038/nmeth.2649.

105. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. CRISPR RNA-guided activation of endogenous human genes. *Nat Methods*. 2013;10(10):977-979. doi:10.1038/nmeth.2598.

106. Davis KM, Pattanayak V, Thompson DB, Zuris J a, Liu DR. Small molecule–triggered Cas9 protein with improved genome-editing specificity. *Nat Chem Biol*. 2015;(April):1-4. doi:10.1038/nchembio.1793.

107. Fu Y, Foden JA, Khayter C, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*. 2013;31:822. http://dx.doi.org/10.1038/nbt.2623.

108. Wu Y, Gao T, Wang X, et al. TALE nickase mediates high efficient targeted transgene integration at the human multi-copy ribosomal DNA locus. *Biochem Biophys Res Commun*. 2014;446(1):261-266. doi:https://doi.org/10.1016/j.bbrc.2014.02.099.

109. Dianov GL, Hübscher U. Mammalian Base Excision Repair: the Forgotten Archangel. *Nucleic Acids Res*. 2013;41(6):3483-3490. doi:10.1093/nar/gkt076.

110. Firth AL, Menon T, Parker GS, et al. Functional Gene Correction for Cystic Fibrosis in Lung Epithelial Cells Generated From Patient iPSCs. *Cell Rep*. 2015;12(9):1385-1390. doi:10.1016/j.celrep.2015.07.062.

111. Wu Y, Liang D, Wang Y, et al. Correction of a Genetic Disease in Mouse via Use of CRISPR-Cas9. *Cell Stem Cell*. 2013;13(6):659-662. doi:https://doi.org/10.1016/j.stem.2013.10.016.

112. Long C, McAnally JR, Shelton JM, Mireault AA, Bassel-Duby R, Olson EN. Prevention of muscular dystrophy in mice by CRISPR/Cas9–mediated editing of germline DNA. *Science (80-)*. 2014;345(6201):1184 LP-1188. http://science.sciencemag.org/content/345/6201/1184.abstract.

113. Beisel CL, Gomaa AA, Barrangou R. A CRISPR design for next-generation antimicrobials. *Genome Biol*. 2014;15(11):516. doi:10.1186/s13059-014-0516-x.

114. Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *MBio*. 2014;5. doi:10.1128/mBio.00928-13.

115. Peng Z, Richardson S, Robinson D, Deutsch S, Cheng J. Genome Editing in Escherichia coli with Cas9 and synthetic CRISPRs. 2014;(March).

116. Barrangou R, Marraffini LA. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol Cell*. 2014;54(2):234-244. doi:10.1016/j.molcel.2014.03.011.

117. Shipman SL, Nivala J, Macklis JD, Church GM. Molecular recordings by directed

CRISPR spacer acquisition. *Science (80-)*. 2016;353(6298). http://science.sciencemag.org/content/353/6298/aaf1175.abstract.

118. Yong HT, Yamamoto N, Takeuchi R, et al. Development of a system for discovery of genetic interactions for essential genes in Escherichia coli K-12. *Genes Genet Syst*. 2013;88(4):233-240. http://www.ncbi.nlm.nih.gov/pubmed/24463526.

119. LEDERBERG J, TATUM EL. Gene recombination in Escherichia coli. *Nature*. 1946;158(4016):558. http://www.ncbi.nlm.nih.gov/pubmed/21001945.

120. Heinemann JA, Sprague Jr GF. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature*. 1989;340:205. http://dx.doi.org/10.1038/340205a0.

121. Waters VL. Conjugation between bacterial and mammalian cells. *Nat Genet*. 2001;29(4):375-376. doi:10.1038/ng779.

122. Karas BJ, Diner RE, Lefebvre SC, et al. Designer diatom episomes delivered by bacterial conjugation. *Nat Commun*. 2015;6:6925. doi:10.1038/ncomms7925.

123. Yoon YG, Koob MD. Transformation of isolated mammalian mitochondria by bacterial conjugation. *Nucleic Acids Res*. 2005;33(16):e139. doi:10.1093/nar/gni140.

124. Nielsen AA, Voigt CA. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol Syst Biol*. 2014;10:763. doi:10.15252/msb.20145735.

125. Datsenko K a, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A*. 2000;97(12):6640-6645. doi:10.1073/pnas.120163297.

126. Baba T, Ara T, Hasegawa M, et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol Syst Biol*. 2006;2(1). http://msb.embopress.org/content/2/1/2006.0008.abstract.

127. Degenkolb J, Takahashi M, Ellestad GA, Hillen W. Structural requirements of tetracycline-Tet repressor interaction: determination of equilibrium binding constants for tetracycline analogs with the Tet repressor. *Antimicrob Agents Chemother*. 1991;35(8):1591-1595. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC245224/.

128. Sambrook J. *Molecular Cloning : A Laboratory Manual*. Third edition. Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, [2001] ©2001 https://search.library.wisc.edu/catalog/999897924602121.

129. Gibson DG, Young L, Chuang R, et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. 2009;6(5):12-16. doi:10.1038/NMETH.1318.

130. Watve MG, Tickoo R, Jog MM, Bhole BD. How many antibiotics are produced by the genus Streptomyces? *Arch Microbiol*. 2001;176(5):386-390. doi:10.1007/s002030100345.

131. Donczew R, Mielke T, Jaworski P, Zakrzewska-Czerwińska J, Zawilak-Pawlik A. Assembly of Helicobacter pylori Initiation Complex Is Determined by Sequence-Specific and Topology-Sensitive DnaA–oriC Interactions. *J Mol Biol*.

2014;426(15):2769-2782. doi:https://doi.org/10.1016/j.jmb.2014.05.018.

132. Dai K, Xu Y, Lutkenhaus J. Cloning and characterization of ftsN, an essential cell division gene in Escherichia coli isolated as a multicopy suppressor of ftsA12(Ts). *J Bacteriol*. 1993;175(12):3790-3797. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC204796/.

133. Weiss DS. Last but not least: new insights into how FtsN triggers constriction during Escherichia coli cell division. *Mol Microbiol*. 2015;95(6):903-909. doi:10.1111/mmi.12925.

134. Yamamoto N, Nakahigashi K, Nakamichi T, et al. Update on the Keio collection of Escherichia coli single-gene deletion mutants. *Mol Syst Biol*. 2009;5(1):335.

135. Laity JH, Lee BM, Wright PE. Zinc finger proteins: new insights into structural and functional diversity. *Curr Opin Struct Biol*. 2001;11(1):39-46. doi:https://doi.org/10.1016/S0959-440X(00)00167-6.

136. Boch J, Scholze H, Schornack S, et al. Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science (80-)*. 2009;326(5959):1509 LP-1512. http://science.sciencemag.org/content/326/5959/1509.abstract.

137. Yang H, Wang H, Shivalila CS, Cheng AW, Shi L, Jaenisch R. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*. 2013;154(6):1370-1379. doi:10.1016/j.cell.2013.08.022.

138. Clatworthy AE, Pierson E, Hung DT. Targeting virulence: a new paradigm for antimicrobial therapy. *Nat Chem Biol*. 2007;3:541. http://dx.doi.org/10.1038/nchembio.2007.24.

139. Glass JI, Assad-Garcia N, Alperovich N, et al. Essential genes of a minimal bacterium. *Proc Natl Acad Sci U S A*. 2006;103(2):425-430. doi:10.1073/pnas.0510013103.

140. Hunter MS, Gupta P, Papitsch-Clark A, Bhugra D, Sturdee D. Culture, country of residence and subjective well-being: a comparison of South Asian mid-aged women living in the UK, UK Caucasian women and women living in Delhi, India. *Int J Cult Ment Health*. 2008;1(1):44-57. doi:10.1080/17542860802121000.

141. Babu M, Díaz-Mejía JJ, Vlasblom J, et al. Genetic interaction maps in Escherichia coli reveal functional crosstalk among cell envelope biogenesis pathways. *PLoS Genet*. 2011;7(11):e1002377. doi:10.1371/journal.pgen.1002377.

142. McLellan JL, O'Neil NJ, Barrett I, et al. Synthetic Lethality of Cohesins with PARPs and Replication Fork Mediators. *PLOS Genet*. 2012;8(3):e1002574. https://doi.org/10.1371/journal.pgen.1002574.

143. Ehrt S. Controlling gene expression in mycobacteria with anhydrotetracycline and Tet repressor. *Nucleic Acids Res*. 2005;33:e21-e21. doi:10.1093/nar/gni013.

144. Verma R, Singh AK, Jaiswal K. Preliminary study on diminution level of RNA/DNA ratio in tissue of Labeo rohita by exposure to some endocrine disrupting compounds (EDCs). *Aceh J Anim Sci*. 2016;1(1):16-20. doi:10.13170/ajas.1.1.3800.

145. Elhadi D, Lv L, Jiang X-R, Wu H, Chen G-Q. CRISPRi engineering E. coli for morphology diversification. *Metab Eng*. 2016;38:358-369. doi:https://doi.org/10.1016/j.ymben.2016.09.001.

146. Arai K, McMacken R, Yasuda S, Kornberg A. Purification and properties of Escherichia coli protein i, a prepriming protein in phi X174 DNA replication. *J Biol Chem* . 1981;256(10):5281-5286. http://www.jbc.org/content/256/10/5281.abstract.

147. Mnaimneh S, Davierwala AP, Haynes J, et al. Exploration of essential gene functions via titratable promoter alleles. *Cell*. 2004;118(1):31-44. doi:10.1016/j.cell.2004.06.013.

148. Liu J, Nurse P, Marians KJ. The Ordered Assembly of the ϕX174-type Primosome: III. PriB FACILITATES COMPLEX FORMATION BETWEEN PriA AND DnaT . *J Biol Chem* . 1996;271(26):15656-15661. doi:10.1074/jbc.271.26.15656.

149. WACHI M, OSAKA K, KOHAMA T, et al. Transcriptional Analysis of the Escherichia coli mreBCD Genes Responsible for Morphogenesis and Chromosome Segregation. *Biosci Biotechnol Biochem*. 2006;70(11):2712-2719. doi:10.1271/bbb.60315.