Genome-wide analysis of *Escherichia coli* antibiotic tolerance by the Bar-code deletion library

Nurhezreen binti Md Iqbal

(Nara Institute of Science and Technology)

(Graduate School of Biological Sciences)

Supervisor: Hirotada Mori

Abstract

Background

It is often assumed that populations of isogenic bacteria are phenotypically and physiologically identical. Actually, isogenic populations displays variations in growth distribution which provide a mechanism for multiple phenotypes to arise in an isogenic bacterial population. In particular, a sub-group termed persisters which show high tolerance to antibiotics. Persisters are described as variation of dormant cells. In addition, persisters were also found in natural environment. However, the mechanism for persisters formation is still unclear. Here, I will report the application of the single-gene knockout mutant library with barcode (ASKA barcode deletion collection) to understand "persister" mechanisms.

Results

Establishment of experimental procedure for ASKA barcode deletion collection by Bar-seq

I had evaluated the ASKA barcode deletion library whether dynamic population alteration could be monitored in mixed culture of strains of the library by deep sequencing of bar-code DNA region in LB for three weeks (LTSP). I monitored bar-code frequencies from chromosomal DNAs purified directly from the time-series culture and after re-grown by refreshing of culture medium (serial passage). Almost 90% of raw sequencing reads were successfully mapped to our barcode strain database, indicate no bias or errors during sequencing. High correlation were observed between independent experiment replicates proved that experiments with ASKA barcode collection were reproducible.

Long-term stationary phase (LTSP)

In nature, bacteria are often exposed to various stress conditions such as extreme pH, temperature shift, osmotic stresses and toxic chemicals) which is akin to bacteria conditions during LTSP. During LTSP, sub-populations with growth advantage in stationary phase (GASP) will take-over the original population and this cycle kept repeating which change the population dynamic in LTSP culture. However, the survival mechanism of GASP is not yet fully understood. I had successfully monitored the *E. coli* mutant population dynamics for three weeks during LTSP. I had identified 31 mutants which showed GASP phenotype. Out of the 31 GASP mutants, 11 of these mutant have mutation involved in transcription regulation (*rpoS, mraZ, cspC, cspE, slyA, iscR, tdcA, mprA, rbsR, fimE*, and *sspA*), 10 involved in membrane transport, biosynthesis and stress (*ygaH,ygaZ, metI, metN, yrbG, hsrA, kdsC, kdsD, pldA* and *cpxA*) and others such as ATP-dependent serine protease (*clpA*), DNA-specific endonuclease

(*endA*), putative fimbrial-like adhesin protein (*yehA*), phosphodiesterase (*cpdA* and *yahA*), phospholipase (*rssA*), ribosomal protein subunit (*rimK* and *yggJ*) and 2-keto-3-deoxy-L-rhamnonate aldolase (*rhmA*). From these results, several survival mechanisms deduced *for E. coli* mutant populations to persist during LTSP; (i) down-regulation of RpoS activity (*rpoS, cspC, cspE, sspA* and *hsrA*), (ii) more efficient substrate utilization (*mraZ, rssA, rbsR* and *tdcA*) (iii) motility (mprA) and (iv) increased biofilm formation (*iscR* and *fimE*). Although the relationship of these mutant and GASP is unclear, I identified several several mutants involved in cell membrane (*kdsC, kdsD, cpxA* and *pldA*) and transpor (*metNI* and *ygaZH*). Further verification using *rssA, cspC, sspA, tdcA, hsrA, endA, ycfJ* and *ygaH* mutant showed fitness advantage over WT after 24 hours competition in LB medium.

Genome-wide analysis of drug-tolerant mutant

Genome-wide screening for drug tolerant mutant were performed to investigate bacterial persisters. Mutants associated with oxidative phosphorylation; NAD(H):ubiquinone oxidoreductase complex (*nuoACEHJMN*), cytochrome complex (*cyoABDE*), ATPase (*atpABCEF*), and Fe-S cluster (*iscUA-hscAB-fdx*) showed increased tolerance against multiple drugs. Others were involves in phosphotransferase system (PTS) such as *ptsI* and *crr* mutants, TCA cycle (*sucB*), NAD(P) transhydrogenase subunit (*pntA*) and iron transporter (*feoB*).In addition, drug tolerant mutants exhibit slow growth phenotype due to disruption in protein complexes involved in oxidative phosphorylation which leads to decrease in ATP. Susceptibility assay indicate that these mutants were susceptible to same MIC concentration as WT. Furthermore, these mutants were shown to become dormant during LTSP. This result supports the current understanding of persisters formation which strongly suggest that these mutations might contributes to persisters formation.

Conclusion

ASKA barcode deletion collection is an effective system for genome-wide screening studies (interaction between gene and specific environment) to comprehensively understand biological systems. I identified 31 mutant candidates that showed GASP phenotypes, of which two mutants (*rpoS* and *sspA*) had been reported. Using this approach, we able to obtain new novel findings which give better understanding toward survival mechanisms of *E. coli* during LTSP. Genome-wide analysis of drug tolerant strongly suggest that drug tolerance were due to formation of persisters. Mutations involved in the oxidative phosphorylation leads to decrease in ATP levels increases persisters formation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 <u>INTRODUCTION</u>

<u>1.1 Discovery of antibiotic and it's resistance</u>

Antibiotics and antimicrobial agents are drugs or chemicals that are used to kill or hinder the growth of bacteria. Antibiotics specifically target bacteria for destruction while leaving other cells of the human body unharmed. The introduction of penicillin in the 1940s, which began the era of antibiotics, has been recognized as one of the greatest advances in therapeutic medicine. Penicillin was discovered by Alexander Fleming was a Scottish physician-scientist in 1928 which was later published in the British Journal of Experimental Pathology in 1929[1], but failed to gain any interest by the scientific community due to difficulty to extract the active compound from the *Penicillium* fungi. Later, together with Howard Florey and Ernst Chain, who devised methods for the large-scale isolation and production of penicillin[2,3]. They received Nobel Prize in Physiology/Medicine in 1945 for discovery of penicillin[3]. The drug was shown to be effective in the treatment of a wide variety of infections, including streptococcal, staphylococcal and gonococcal infections[1]. Penicillin works by disrupting bacterial cell wall assembly process which leads to bacterial cell death. Today, other penicillin-related antibiotics including ampicillin, amoxicillin, methicillin and flucloxacillin are used to treat a variety of infections. Subsequent to the discovery of penicillin, sulfonamide (Prontosil) was discovered by Gerhard J. P. Domagk in 1932[4], aminoglycosides (streptomycin) in 1944[5], cephalosporin from *Cephalosporium acremonium* by Giuseppe Brotzu in 1948[6], macrolides (erythromycin) from *Streptomyces erythrew* by McGuire in 1952[7], quinolones (nalidixic acid) in 1962[8] and its derivatives (norfloxacin, ciprofloxacin and ofloxacin)[9] and carbapenems in 1974[10]. After introduction of carbapenems drugs, new drug discovery had decline and became void by 1990s. Until 2015, a new class of antimicrobial called teixobactin was isolated from soil bacteria *Eleftheria terrae* which shown to be effective in killing *Staphylococcus aureus* and *Mycobacterium tuberculosis*[11]. Teixobactin was shown to kill *S. aureus* by inhibiting it's cell wall synthesis by binding to lipid II (precursor of peptidoglycan) and lipid III (precursor of cell wall teichoic acid). Unfortunately, teixobactin did exhibit killing activity against *E. coli* which indicated that this compound may not be effective against Gram-negative bacteria[11]. The membrane structure of Gram-negative bacteria which decrease its permeability or efflux system was reasoned to failure for teixobactin[11].

Antibiotic resistance is becoming a common occurrences, due to the frequent use of antibiotics. Infections by resistant strains of bacteria are much more difficult to treat and often cause relapse. While this resistance often results from horizontal gene transfer of pre-existing resistance genes, there are also some that were due to mutations[12,13]. For example, kanamycin resistance can result from loss of a specific transport protein[14] and ampicillin resistance can result from an SOS response that halts cell division[15], while loss of nitroreductase activity can provide resistance to metronidazole and amoxicillin[16]. Resistance to quinolone drugs such as ciprofloxacin were often mediated by mutations at drug target site; gyrase (*gyrA*) and/or topoisomerase IV (*parC*)[17]. Similarly, mutation at *folA* will confer resistance against trimethoprim[18]. Meanwhile *cpxA* mutations were shown to cause increase resistance to multiple drugs such as aminoglycoside, beta-lactams and fluoroquinolones[19].These mutations usually reduce or inactivate cellular systems or functions. Thus, bacteria becomes resistance to antibiotic by elimination certain functions, such as transport proteins, enzymatic activity and binding affinity.

Extensive reviews of drug, its mechanism and resistance have been published. A review "How antibiotics kill bacteria: from target to network" by Kohanski and co-workers[20] discussed drug-target interactions and the associated mechanism for major bactericidal drug classes; quinolones, rifamycins, beta-lactams and aminoglycosides. "Mechanism of Resistance to Aminoglycoside Antibiotics: Overview and Perspectives" by Garneau-Tsodikova and Labby[21] provides an comprehensive information on resistance towards aminoglycosides via intrinsic and acquired mechanisms. "Mechanism of action of and resistance to quinolones" by Fabrega *et. al.*[22] discussed structure, mode of action and acquired resistance against quinolones by specific Gram-negative and Gram-positive bacteria (*E. coli, Salmonella enterica, Klebsiella pneumoniae, Pseudomonas aeruginosa, Acinetobacter baumannii*, *S. aureus*, and *Streptococcus pneumoniae*) while "Mechanism of Quinolone Action and Resistance" by Aldred and co-workers[17] have lengthily reviewed on the quinolone drug from its structure, drug target and mode of action and resistance against quinolone by chromosomal mutation and plasmid-mediated resistance. "Bacterial resistance to Tetracycline: Mechanisms, Transfer and Clinical Significance" by Speer and co-workers[23] gives comprehensive information of tetracycline drugs from clinical treatments, drug structure and mode of action, resistance mechanism by chromosomal mutation, plasmid-mediated and from environment (clinical surveys).

According to Centers for Disease Control and Prevention (CDC), at least 2 million people is infected by drug resistant bacteria which caused more than 23, 000 death in US

alone[24]. Among gram-positive pathogens, methicillin-resistant *Stapylococcus aureus* (MRSA) and vancomycin-resistant *Enterococcus* species (VRE) posed the biggest threat. Furthermore, gram-negative pathogens are more serious because they are becoming resistant to nearly all the antibiotic drug options available, multidrug resistance (MDR) strain which are becoming more prevalent in hospitals. These are most commonly caused by *Enterobacteriaceae* (mostly *Klebsiella pneumoniae*), *Pseudomonas aeruginosa*, *Acinetobacter,* beta-lactamase-producing *Escherichia coli* and *Neisseria gonorrhoeae*[25].

1.2 Drug persistence

Bacterial persistence are major cause of antibiotic drug treatment failure and relapsed in infection. This phenomenon has been reported as early as 1942, due to the recurrences of staphylococcal infections even after treatment with high dosage of penicillin[26] . Joseph Bigger also had observed that penicillin failed to sterilized flask cultures of *Stapylococcus aureus* at exponential phase[27]. He referred to these as "persisters" which is a small population of dormant or non-growing bacteria that have non-heritable tolerance to penicillin but have the capacity to regrow and remain susceptible to the same antibiotic. Two main characteristics of bacterial persisters are (i) drug tolerant and (ii) low metabolic activity or dormant state. There are two views on "persister cells", (i) Phenotypic heterogeneity; persister cells comprise a subpopulation of bacteria that become highly tolerant to antibiotics and reach this state without undergoing genetic change[28] (as describe by J. Bigger) and (ii) genotypic heterogeneity; persisters population arise due to genetic modification; most well-known persisters mutation is *hipA* mutation[29,30]. In my study, I use the latter view as definition of persisters population.

1.2.1 Mechanism of bacterial persistence

When microbial populations were exposed to killing concentration of antibiotic, a small fraction of the population can escape the killing action by entering a dormant state. This phenomenon is known as persistence. In addition, persistence bacterial populations also exhibit multidrug tolerance[31]. The presence of persistence during drug treatment is shown by biphasic killing pattern as shown in Figure 1.1. Consequently, when the drug is removed the persister cells will give rise to bacterial population that were sensitive to drugs.

Numerous studies indicate that toxin-antitoxin modules are required for persistence in *E. coli*. In 1983, Moyed and Bertrand had identified in *E. coli* the of high persistence (*hip*) mutants; *hipA7* mutant which is related to the toxin-antitoxin modules, that increased the

frequency of persistence by 10 000 fold[29]. Toxin HipA is a kinase that inactivates the translation factor EF-Tu by phosphorylation, leading to protein synthesis inhibition and cell
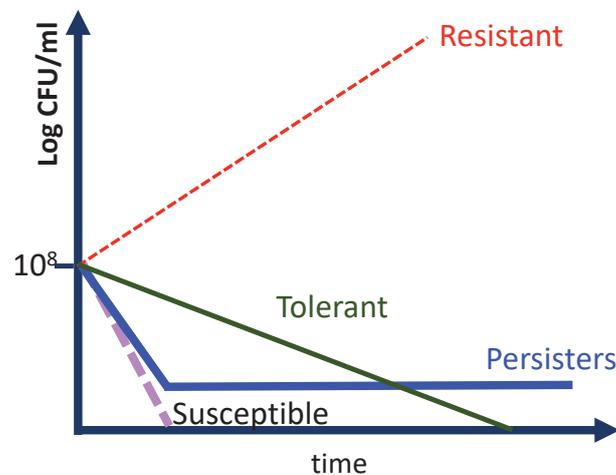
Figure 1.1 Killing kinetics during treatment with a bactericidal drug

Viable cells (CFU) were plotted against duration of drug treatment or drug concentration. Resistant bacterial population remain viable and grow during drug treatment. Susceptible bacterial population were rapidly killed when exposed to drug. Tolerant bacterial population are susceptible to the drug but need longer time to be killed. Meanwhile for persisters, increase incubation time or drug concentration leads to rapid killing of majority of bacterial population. However, after certain threshold (blue line), a killing plateau observed as persister cells remain viable. (Adapted from "Persister cells", by Lewis, K. (2010) *Annu. Rev. Microbiol.* **64,** 357–372)[28]

growth arrest. The *hipA7* mutation decreases the affinity of *HipA* to *HipB* antitoxin resulting in enhanced toxicity of HipA[32]. The discovery of Moyed and Bertrand[29] of the predicted TA module *hipBA* and its association in persister cell formation encouraged researchers to explore the role of other TA modules. Figure 1.2 showed the mechanisms of persisters formation mediated by TA modules.

Transcriptome analysis by DNA microarray was performed by Lewis and coworkers[33] using *E. coli hipA7* strain and ampicillin treatment had revealed other toxin-antitoxin (TA) genes (*yafQ/dinJ*, *relE/relB* and *mazE/mazF*) were upregulated in persisters. Other genes upregulated in persisters are members of at least two functional groups: the SOS genes (*recA, umuDC, uvrAB, sulA*) and the genes of the heat and cold shock response family (*cspH, htrA, ibpAB, htpX,* and *clpB*). Another transcriptome analysis was performed by Lewis and coworkers[34] from dormant cells isolated using fluorescence-activated cell sorting (FACS) based on diminished GFP fluorescent (metabolic inactive cells). These dormant cells were

shown to have 20-fold greater persistence to ofloxacin. They found that these cells have elevated transcription of toxin gene *mqsR*. In addition, *dinJ, yoeB* and *yefM* which are related to TA were also identified.

In addition to TA systems, Maisonneuve *et. al*[35] had shown that Lon protease and mRNA endonuclease are both required for persister cell formation. Overexpression of *lon* showed increased in persister formation in WT cells but not in mRNA endonuclease mutant. Lon activity is required to degrade labile antitoxins for type II TA systems where a protein antitoxin inactivates the protein toxin. Meanwhile, activation of the mRNA endonuclease would inhibits translation, subsequently induces dormancy and persistence.

Balaban and coworkers observed two types of persister population, type I and type II[36]. Type I persister appear only after the culture had reached the stationary phase prior to exposure to antibiotics, which result in delay in regrowth (extended lag time) while type II persister is by slow growth that are continuously generated during exponential growth. To determine whether the evolved tolerant population in their study (intermittent exposure to high concentration of ampicillin) exhibit which persister type, they monitored and compared the time of cell division using a single-cell approach and microfluidics. They found that the tolerant exhibited were due to extended lag time or type I. Furthermore, they had identified mutations in 6 genes in their evolved strain; *metG* (methionyl-tRNA synthetase), *vapB* (antitoxin homolog, *vapC* toxin), *prs* (ribose-phosphate diphosphokinase), *sspA* (stringent starvation protein), *pgm* (phosphoglucomutase) and *yeaI* (predicted diguanylate cyclase). VapB is associated with the TA module. VapC toxin acts as ribonuclease, cleaving RNA molecules and thereby reducing the rate of translation, which support TA system in persister cells formation.
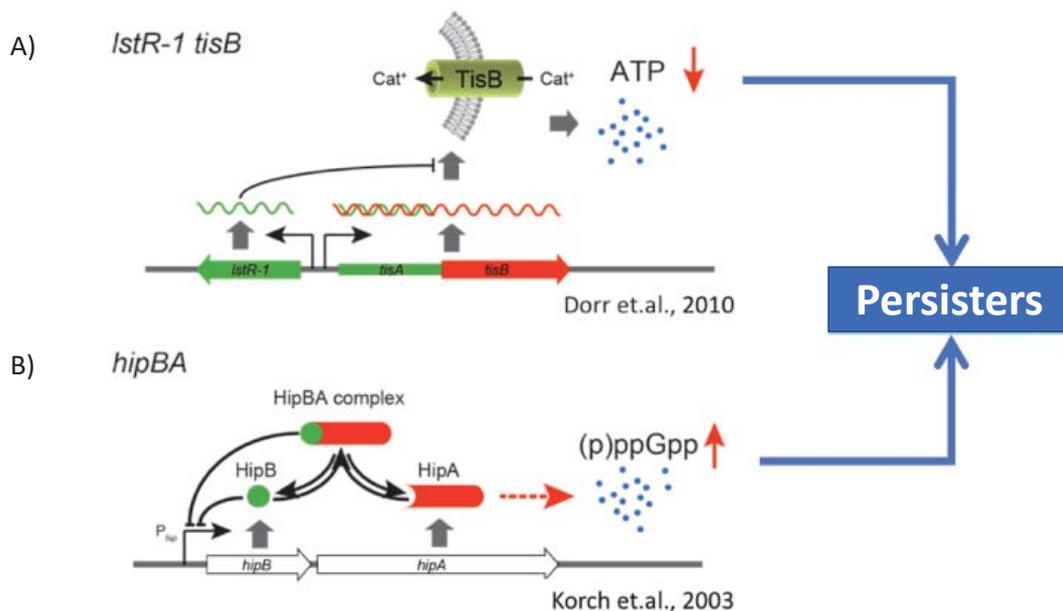
Figure 1.2 Mechanisms of persister cells formation mediated by TA system

Mechanism persister cells formation mediated by toxin-antitoxin system. (A) TisB-IstR-1 toxin antitoxin system. *IstR* is an antisense RNA antitoxin, is constitutively express from its own promoter so under normal conditions very little tisB is synthesized (bind to tisA which is contains the antisense RNA binding site as well as the ribosome binding for tisB). However, *tisB* translation is under lexA control so it is induce by DNA damage as part of SOS response. During dna damage, *tisB* is strongly induced which overrides the antitoxin. TisB insertion into the membrane result in loss of membrane potential which leads to decreased in ATP. The decrease in ATP increase persisters formation[37]. (B) HipAB toxin-antitoxin system. HipA is toxin and hipB encodes for antitoxin, persister formation is due to mutation in hipA7 cause decrease in binding interaction with hipB thus increase the intracellular level of hipA toxin. The increase of *hipA* increase the level of ppGpp synthesis. This leads to activation of stringent response. This will lead to growth inhibition and cause persisters formation[38].

Persisters are bacterial cells or population that are metabolically inactive (or reduced) and are dormant. Although many studies relates persisters as an adaptive response or survival strategy during drug treatment, persister cells also formed under different stress conditions such as starvation[39] and nutrient shift[40]. Starvation-induced persisters have been reported in *E. coli*[41,42] and *Pseudomonas aeruginosa*[43]. Shan and coworker[42] had used Tn-Seq to investigate *E. coli* tolerance against aminoglycosides by exposing the transposon library grown to stationary phase to lethal dose of gentamicin. They found that 37 genes showed 5-fold increase of tolerance to gentamicin. They had suggested that activation of motility and amino acid biosynthesis contributes to the formation of persisters tolerant to gentamicin. The author had

suggested that decreased in ATP synthesis might contribute to formation of persister cells[41]. Mlynarcik and Kolar[43] had compared survival of *P. aeruginosa* at exponential and stationary phase on several drugs. They found that stationary phase cells increase tolerance against imipenem, polymyxin B and tobramycin and nutrient starvation (serine starvation) contribute to increased tolerance in *P. aeruginosa*.

From above findings, stationary phase cultures showed increased tolerance against drug due to formation of persisters cells which in turn formed due to nutrient starvation. This indicate the importance of stationary phase in formation of persister cells. However, studies of stationary phase revealed interesting phenomena such as; viable but non-culturable (VBNC)[44,45] and growth advantage during stationary phase (GASP) phenotypes[46–48].

Bacteria displays five growth phases when grown as batch culture as shown in Figure1.3; lag, exponential, stationary, death and long-term stationary phase (LTSP)[48]. Bacteria enters stationary phase once nutrient in culture becomes depleted and accumulation of waste products from bacterial metabolism eventually leads to cell death (death phase). *E. coli* was shown to enter death phase after 3 days of incubation in LB[48]. Here, almost 99% of the population dies and nutrient released by their lysed cells became nutrient source for the survivors population[48]. During this phase, bacterial population became heterogenous as populations with GASP mutation arise[46]. It was shown that *E. coli* can survive in long-term stationary phase (LTSP), without the addition of nutrients, from days to several years[49]. Mutations in the *rpoS* gene were shown to be common for GASP phenotype[46]. Additional mutations were found in the *lrp* and *ybeJ-gltJKL*, which encodes leucine-responsive protein and aspartate/glutamate transporter[47]. Mutations in these genes results in increased ability to catabolize amino acids as carbon and energy source[47,50,51].

The viable and non-culturable (VBNC) is one of the state of dormancy when bacteria were under stress conditions. It is differed from dead cells in terms of membrane integrity; membrane of dead cells is damaged while VBNC cell still maintained complete membrane structure[45]. There have been arguments that VBNC and persisters are the same as their "dormant" state shared similar traits[52–54]. VBNC cells also shown to be drug tolerant. One of the main difference between VBNC and persister cell is their culturability. VBNC were characterized by loss of culturability in conventional culture media (i.e LB) while persister cells is able to resume growth when stress (i.e. antibiotic) were removed. However, successful resuscitation of VBNC cells brings back this question[44,55].
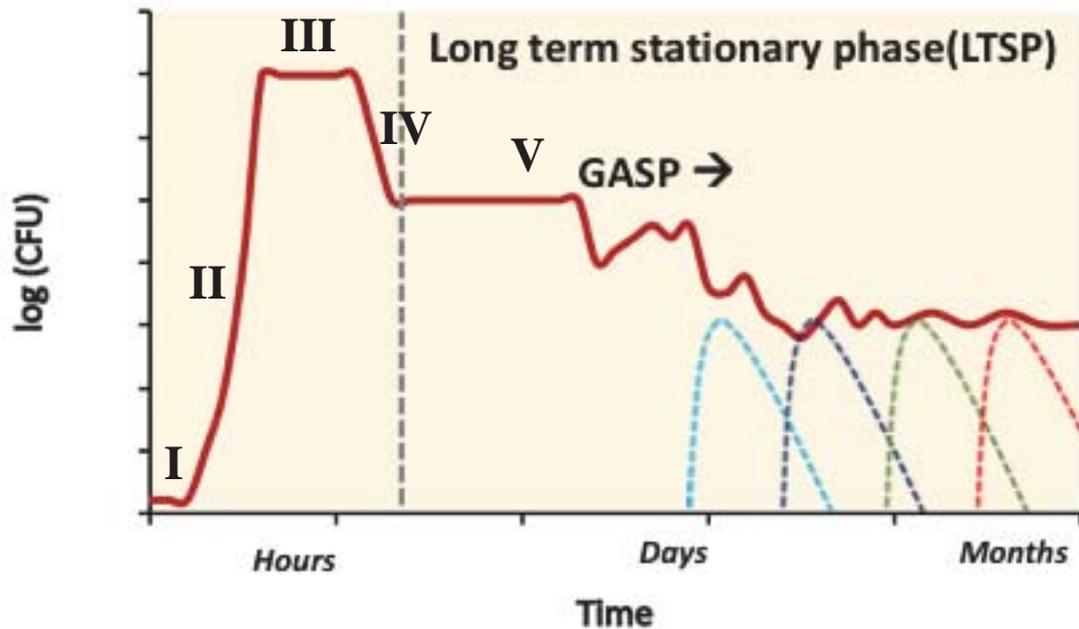
Figure 1.3 Growth phases of Escherichia coli during batch cultivation.

E. coli exhibit 5 growth phases during batch cultivation; phase I: lag phase, phase II: exponential phase, phase III: stationary phase, phase IV: death phase and phase V: long-term stationary phase. During LTSP, mutant population with growth advantage during stationary phase (GASP) population were observed. (Adapted from "Long-term survival during stationary phase: evolution and the GASP phenotype" by Finkel, S. E. (2006), *Nat.Rev.Microbiol.* **4,** 113–120)[48]

1.2.2 Bacterial persistence leads to resistance

Persisters pose significant challenges for the treatment of many chronic and persistent bacterial infections such as tuberculosis (TB)[56], Lyme disease[57] and urinary tract infections. While the most attention has been given to drug resistance, drug persistence or tolerance is equally important, if not more because prolonged and repeated treatment of persistent infections may lead to genetic drug resistance, which especially be true for TB treatment. This has been proven Levin-Reisman and co-workers[58] by intermittent exposure of *E. coli* to ampicillin had evolved the bacteria to become resistant by developing mutations in *ampC* gene, which hydrolyzed ampicillin. Furthermore, they had observed that resistant strains only emerged from tolerant strain or persister. The *E. coli* first develop mutation that slowed their growth, which allow them to persist during ampicillin treatment and then acquire *ampC* mutations that confer to resistance.

1.3 *Escherichia coli*

Escherichia coli, originally called "*Bacterium coli commune*", was first isolated from child feces in 1885 by German pediatrician Theodor Escherich[59]. *Escherichia coli* is classified under the class of gamma-proteobacteria and family of *Enterobacteriaceae*. It is facultative anaerobic, gram-negative, non-spore forming rod-shape bacteria. It size is approximately 0.5 μm in width and 1.0-3.0 μm in length. It is able to grow fermentatively on glucose or other sugars (*i.e* lactose, arabinose, maltose, D-mannitol, D-sorbitol etc.) producing acid and CO_2[60]. Using laboratory biochemical test, it shows positive for indole production and methyl red test. Meanwhile, most *E. coli* strains shows negative for oxidase, citrate, urease and hydrogen sulfide test[60]. *E. coli* are commonly motile in liquid by means of peritrichous flagella. It can grow within pH range of 5.5 to 8.0 with optimum growth at neutral (pH 6.8-7) while the optimum growth temperature range at 37-42 °C [61].

E. coli which is a member of the *Enterobacteriaceae* and is closely related to pathogens such as *Salmonella, Klebsiella, Serratia,* and the infamous *Yersinia pestis*, which causes plague. *E. coli* are commonly present in intestinal tract of humans and animals; also can be found in soil and water source as result of fecal contamination. Although *E. coli* is mostly harmless, pathogenicity islands have been identified and associated with pathogenesis in *E. coli* resulting in strains that colonize different tissues. There are some strains of *E. coli* (*i.e. E. coli* O157:H7) that are pathogenic to humans and animals. This strain produce Shiga toxins, causing severe diarrhea and a major cause of foodborne infections[62]. As mention previously, infection by these pathogens have become harder to treat as they have shown to be multidrug resistance (MDR). These are most commonly caused by *Enterobacteriaceae* (mostly *Klebsiella pneumoniae*),*Pseudomonas aeruginosa, Acinetobacter* and beta-lactamase-producing *Escherichia coli*[25].

1.3.1 E. coli research achievements

Since 1940s, *E. coli* became a prevalent choice as model organism to work with as it is easily available as an isolate. It is non-pathogenic and versatile as it grows on variety of nutrient and can be isolated from any humans. Many scientific achievements (Nobel prizes) were established using *E. coli* as a model organism; questions from fundamental aspects of life, *viz.* the genetic code[63], transcription[64], translation, replication[65] and gene regulation[66,67]. Others include discovery of restriction enzymes[68] and life cycle of lytic and lysogenic

bacteriophages[69,70]. These discoveries lead to accumulation of knowledges and development of molecular techniques for investigating and manipulating genes or cells.

1.3.2 E. coli in systems biology

Systems biology is an interdisciplinary field which integrates mathematics, computational and biology giving a holistic view of complex interactions within biological systems. Information from different cellular levels; genes (*genomics*), gene expression products (*transcriptomics*), proteins (*proteomics*), metabolites (*metabolomics*) and interactions (*interactomics*) are neccessary for a complete understanding of a given biological systems. From biochemical, physiological and genetic perspectives, *E. coli* is one of the best understood and characterized living organisms, with laboratory studies on *E. coli* K-12 for almost a century (isolated in 1922)[71] makes it a suitable model not only for molecular biology but also in systems biology.

In 1997, the complete genome sequence of the two *E. coli* K-12 sub-strain, MG1655[72] and W3110[73] was obtained. This complete genome contains a single circular duplex molecule composed of 4,639,221 bp. The genome also consist of remnants of many phages, and insertion sequences (IS) and a high transport capacity toward the cytoplasm. Regarding its structure, protein-coding regions correspond to 87.8% of the genome, while 0.8% encodes for stable RNAs, and 0.7% consists of noncoding repeats. The remaining 11% encodes for regulatory and other functions[72]. Since the sequence of *E. coli* genome has been completed, comprehensive effort has been made to annotate the sequenced genome[74]. Yet, only half (~54%) of the gene products have its function confirmed (experimental evidence). It has been reported that from the 4,225 protein coding regions in *E. coli*, 2794 (66%) had been annotated[75] while 1431 (34%) were still functionally uncharacterized genes or orphans. Out of the 1431 orphan genes, 446 (31%) have at least one putative molecular function defined on the basis of sequence (such as the presence of a predicted DNA-binding domain or an enzymatic motif) in the Clusters of Orthologous Groups (COGs) of proteins catalog[76]. Hu and co-workers[77] used combined computational and experimental data (using PI dataset) to elucidate the biological functions for these unannotated (orphan) genes. From their PI dataset, they obtained 5,993 high-confidence, nonredundant pairwise interactions among 1,757 distinct *E. coli* proteins, including 451 orphans. Furthermore, they had identified 25 orphans genes were involves in protein synthesis (i.e. *ybcJ*, *yncE*, *yfgB*, *yibL*, *ydhQ*, *yagJ*, *yjcF* and *ybeB*).

Since the availability of complete *E. coli* genome sequence, several experimental resources has been established such as ORFeome clone libraries and deletion mutant collections for elucidation of *E. coli* gene function in a systematic approach. Table 1.1 showed the list of experimental resources developed by our lab for functional genomics studies in *E. coli*. The Keio collection has been used by many groups to observe the effect of gene deletion under certain conditions; *i.e.* mutations which caused hypersensitivity to antibiotics[78,79], motility[80], biofilm formation[81], cysteine tolerance and production[82] and sensitivity to hydroxyurea (HU)[83]. Using the Keio collection, we are able to detect gene-environment interaction of individual mutants systematically with high-throughput approach.

With recent advancement of Next Generation Sequencing (NGS) technology, our lab had established another *E. coli* deletion collection named as ASKA barcode deletion collection which allows pooling of all mutant strains and grown under certain selection condition in a competitive growth approach.

Table 1.1 Escherichia coli experimental resources

Resources	Description
Plasmid clone library*	
ASKA ORF clone (A Complete Set of *E. coli* K-12 ORF Archive)[84]	A complete set of individual genes encoding histidine tagged proteins (+/- GFP) attached was constructed for functional genomic analysis. Expression of the cloned ORF is under the control of an IPTG-inducible promoter which is strictly repressed by the *LacI* (gene product of *lacI*^{q}) repressor. (sequence based on *E. coli* K-12 strain W3110)
Gateway entry clone[85]	*E. coli* K-12 ORFs (from ASKA clone) were cloned into Gateway® entry vector pENTR/Zeo. The Gateway® system facilitates the transfer of ORFs into a large range of expression vectors that are suitable for downstream studies.
TransBac	Full sequence of *E. coli* K-12 ORFs were cloned into pFE604TR-ccd cloning vector. Expression of the cloned ORF is under the control of an IPTG-inducible promoter which is strictly repressed by the *LacI* (gene product of *lacI*^{q}) repressor.
Single-gene deletion library	
KEIO collection[86,87]	A set of single-gene deletions of all nonessential genes in *E. coli* K-12. ORF regions were replaced with a kanamycin cassette flanked by FRT sequence. Resistant cassette can be excised by the Flp-recombinase (pCP20) to create in-frame deletions.
ASKA barcode collection	Construction method is similar with KEIO collection. ORF regions were replaced with chloramphenicol cassette flanked by FRT1 sequence. Additional feature is insertion ~20 nucleotides downstream to PS1 site. Resistant cassette can be excised by the Flp-recombinase (pCP20) to create in-frame deletions.

* details on plasmid clone library is underway (Yamamoto *et. al.,* in preparation)

The advancement of genomic, transcriptomic, and proteomic technologies has led to the development of several online database which focus on *E. coli* shown in Table 1.2.

Table 1.2 List of Escherichia coli databases

Name	URL	Description*
EcoCyc: Encyclopedia of *E. coli* Genes and Metabolic Pathways	https://ecocyc.org/	Comprehensive database joining together genomic information with biochemical features of *E. coli*
PortEco	http://www.porteco.org	Data of the biology of *E. coli* including plasmids, mobile genetic elements, and phages
EcoliWiki	http://ecoliwiki.net/	Community-based pages about everything related to the biology of the nonpathogenic *E. coli*
EcoGen 3.0	http://ecogene.org/	*E. coli* database dedicated to analyzing and comparing genomic and transcriptomic data
RegulonDB	http://regulondb.ccg.unam.mx	Database on transcriptional regulation in *E. coli* K-12 containing knowledge manually curated from original scientific publications, complemented with high throughput datasets and comprehensive computational predictions
GenoBase	http://ecoli.naist.jp/GB/	Comprehensive resource database of *Escherichia coli* K-12[88]
***E. coli* Genetic Stock Center**	http://cgsc.biology.yale.edu/	The CGSC Database of *E. coli* genetic information includes genotypes and reference information for the strains in the CGSC collection
*Source: database website.		

1.4 Bar-seq: Barcode analysis by sequencing

The principle of Bar-seq is that each strain carries a deletion mutation with unique random sequence (barcode). Pools of mutant strains can therefore be competed under selective conditions and the fitness of each mutant can then be determined by sequencing the barcodes of the pool before and after selection. Barcodes are then amplified using PCR and processed for deep sequencing.

Bar-code analysis by Bar-seq was first introduced using barcoded yeast deletion mutant (YKO) library (*Saccharomyces* Genome Deletion Project)[89] which were completed in 2002[90]. Since debut, this deletion collection has been a starting point for numerous large-scale "genetic-network" type studies that offer a global insight into complex genetic phenotypes. The YKO library has been used for genome-wide phenotypic analysis toward better understanding of biological function, response to stress and mechanism of drug action. One of the earliest study using the barcoded yeast deletion library were the genome-wide profiling under six stress conditions *viz.* high salt concentration, sorbitol, galactose, pH8, nutrient limitation (minimal medium) and nystatin, an antifungal drug[90]. Some of notable findings from this study was that 18.7% of 5916 yeast ORFs were identified as essential for growth. In addition, 15% of the mutant exhibit slow-growth phenotype in rich media and almost half of the "essential" mutants under minimal medium involves in known metabolic pathways. Other study using the YKO library for queries on biological function; to identify genes that contribute to budding site selection[91], genes important for cell wall synthesis and regulation[92] and respiration[93,94]. In the YKO screening genes involved in bipolar budding pattern, Ni and Synder (2001) had identified 127 genes involved in the budding of which 22 were uncharacterized genes that are important for bud site selection in yeast cells[91]. Genes involved in yeast cell wall synthesis were screened by exposing the YKO library to antifungal protein, K1 killer toxin. Here, they had identified 268 mutant; 186 showing resistance and 82 showing higher susceptibility to this toxin compared with WT. Furthermore, 42 of these genes were still unknown function[92]. The completion of the YKO collection encouraged the construction of many other genome-wide libraries using fission yeast; *Saccharomyces pombe*[95], plant; *Arabidopsis thaliana*[96] by insertion mutagenesis, bacteria; *Escherichia coli* (Keio deletion collection and ASKA barcode deletion collection)[87], *Pseudomonas aeruginosa* by transposon mutagenesis[97] (>30,000 mutants) and *Bacillus subtillis*[98] as well as novel genome-wide techniques such as protein microarray[99], digenetic interactions by SGA (systematic genetic analysis)[100], and large-scale expression studies[101]. Hughes and coworkers (2000) had collected the expression profiles of

300 mutant (unknown function) exposed to various chemicals to predict the genes function by matching the profiles pattern[101].

Since completion of the YKO libraries, analysis using this library were based on independent mutant via phenotypic profiling or pooled libraries via tag microarray. In 2009, Smith and coworkers[102] had apply deep sequencing analysis to YKO libraries and evaluates them in comparison to the barcode microarray assay. They had reported that this method, Barcode analysis by Sequencing or Bar-seq outperforms the barcode microarray method. They found that Bar-seq showed more sensitive and have wider dynamic range as it directly "count" each barcode instead of fluorescence intensity in which signal saturation caused errors during analysis of microarray data. Furthermore, Bar-seq allows multiplexing thus different samples can be combined within a single lane for sequencing. A pilot study (multiplexing), they suggested that 100-300 counts/barcode is sufficient read depth for multiplex sequencing[103]. Several studies using Bar-seq on YKO in exposed to different stress condition under competitive growth have been published[104–106]. Overall, the establishment of YKO libraries had brought new scientific discovery that greatly accelerate yeast research for better understanding of yeast cell biology.
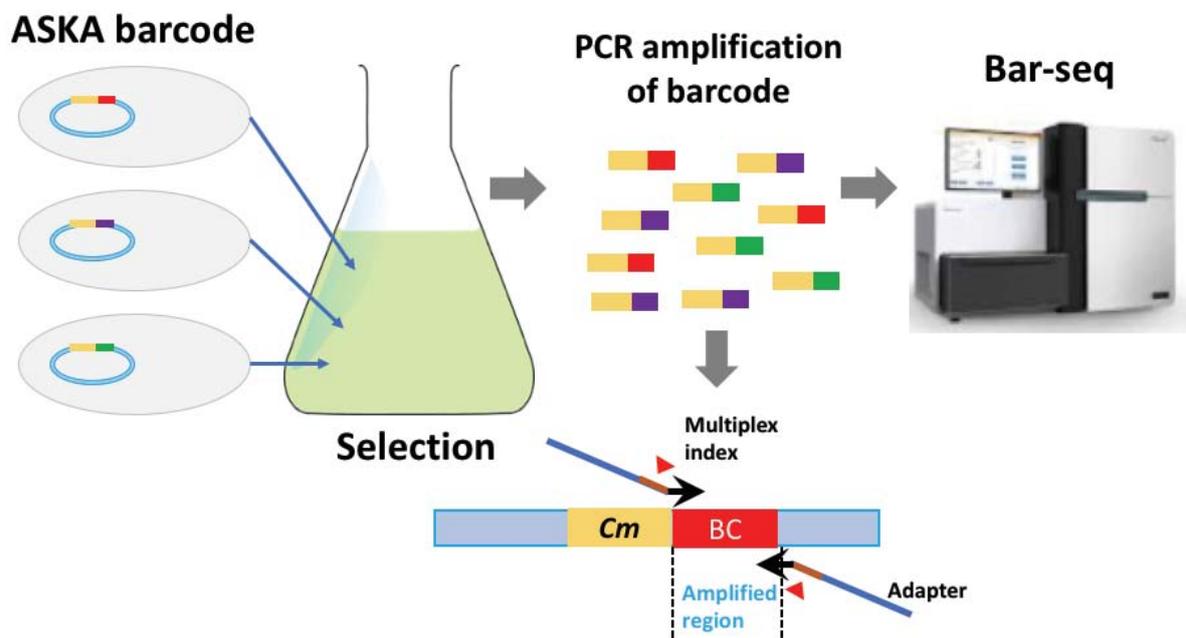
Figure 1.4 Strategy of Bar-seq using ASKA barcode deletion library

All mutant were pooled and grown under certain selection, genomic DNA before and after selection were isolated and barcode region were amplified by PCR then process for deep sequencing.

1.5 Aims of this study

Bacterial persisters came to attention as there have been growing evidence that relates persisters as main cause for recurrence infection and drug treatment failure. Furthermore, persistence is becoming increasingly recognized as an important survival strategy that bacteria used to escape killing by antibiotic drugs. To develop new drugs targeting chronic infections requires a deep understanding of the mechanisms underlying persister formation. However, the extremely low percentage of persister cells in a bacterial population and complex pathways involved in persister formation have delayed the study of this phenomenon.

It is often assumed that populations of isogenic bacteria are phenotypically and physiologically identical. However, bacteria always show considerable phenotypic variation within a population as shown in Figure 1.5(A). For example, although most cells display average growth rate, some cells exhibit lower growth rate which might leads to dormant cells or increase tendency in becoming persisters once expose to drug. The growth distribution in bacteria cells were due to cellular heterogeneity; is an important survival strategy of bacterial population during drug treatment. So, what cause increase of persister cells formation? To answer this question, I use ASKA barcode deletion collection to represent phenotypic variation within *E. coli* population. Each mutant strain will exhibit different growth fitness distribution as shown in Figure 1.5 (B) and certain mutant population may exhibit tendency to become persisters. To identify bacterial persisters, I use one of its characteristics which are drug tolerant. Hence, I perform genome-wide screening of drug tolerant mutant.

An advantage of ASKA barcode deletion collection is that each mutant strain has unique barcode that can be identified via deep-sequencing, so I can pool all mutant in a single flask and grow them under competitive growth. Using this library, I would like to identify mutant that exhibit increased tolerance against drug which likely to be drug persisters.

The ASKA barcode deletion collection is a new resource established by our lab and it's usability had yet to be validated. Therefore, I performed another experiment for evaluation. This is to check whether it is possible to monitor the population changes of pooled mutants by deep sequencing. In this experiment, I monitored the *E. coli* mutant population during long-term stationary phase in LB medium for three weeks under batch culture condition. As mentioned, persisters could be formed under various stress condition such as starvation, extreme pH, osmolarity and exposure to toxic chemicals which is akin to bacterial conditions during LTSP. In addition, LTSP is used to study bacterial adaptation in nature as the conditions the bacteria faced during LTSP is similar. Furthermore, stationary phase cells were shown to be more tolerant against antimicrobial drugs. Two phenotypes were observed during LTSP;

GASP population and VBNC. It has been shown that in *E. coli* population during long-term stationary phase, sub-populations with growth advantage during stationary phase (GASP) due to mutation will arise[46,48–50]. Meanwhile, viable and non-culturable (VBNC) state is "dormant" state of bacterial cell. Hence, I am interested to know if the persisters population during LTSP shared similar traits with drug tolerant population.
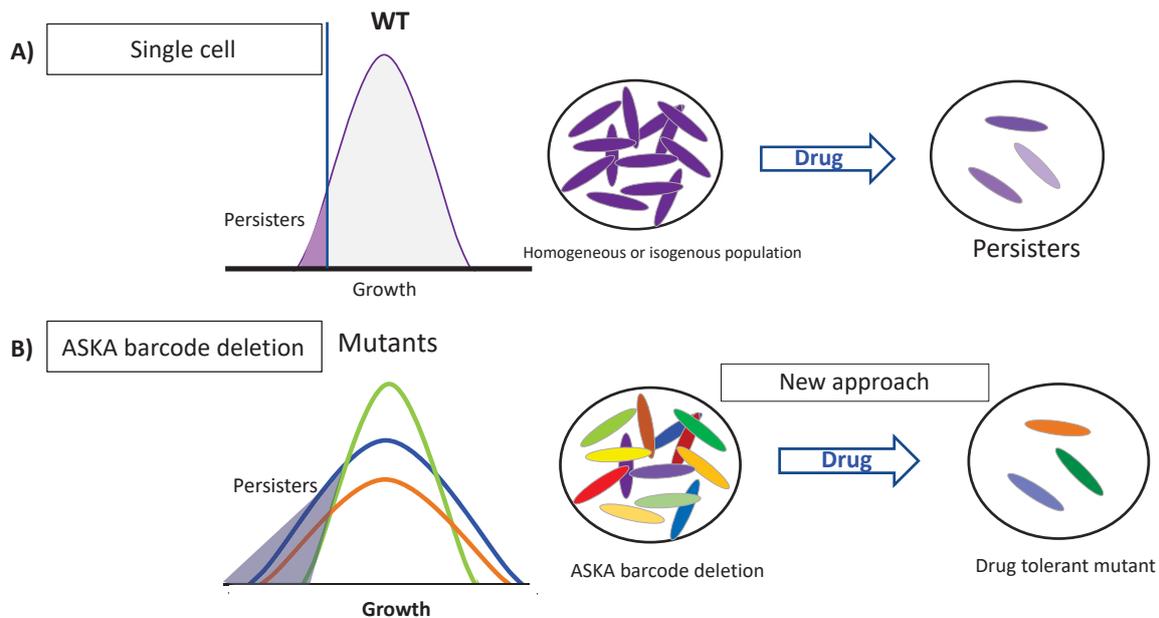
Figure 1.5 Growth fitness distribution of single bacteria population (isogenous population).

(A) Isogenous population; showed varied growth distribution in which some cells tends to become persisters when expose to drug.(B) ASKA barcode deletion collection; each mutant population showed varied growth distribution in which certain mutant population might have tendency to become persisters when expose to drug.

Research objectives:

1. Establishment of experimental procedure using Bar-seq
    a. Confirm reproducibility
    b. Monitoring growth of ASKA barcode deletion collection during LTSP
    c. Identify and validate GASP mutant candidates
2. Genome-wide screening of drug tolerant mutant by Bar-seq

Chapter 2 <u>MATERIALS AND METHODS</u>

2.1 Strain: ASKA barcode deletion collection

Escherichia coli strain *BW38028* (*F⁻, Δ(araD-araB)567, lacZp-4105(UV5)-lacY, λ⁻, hsdR514*) was used for construction of this library. The design of the this deletion library was based on annotations of *E. coli* K-12 *BW25113* using the same primer set that was used to construct Keio collection[87]. Each target ORF was replaced with chloramphenicol resistant cassette and 20 nucleotides random sequence (barcode). Figure 2.1 showed the final genomic structure of target region in the barcoded deletion library. Two independent deletion candidates were isolated; assigned as clone 1 and clone 2, each with its own unique barcode which allows mixing of both libraries in a single flask thus serve as biological replicate within the same experimental set. This library consist of 7082 SKO barcode mutants (Table 2.1); clone 1: 3598 genes and clone 2: 3484 genes were knocked out, respectively. Glycerol stocks of mix libraries mutants for clone 1 and clone 2 were prepared and could be used separately (clone 1/clone 2 only). In addition, known mutator strains (*dam, mutL, mutH, mutM, mutS, mutT, mutY, uvrD, xthA*) were removed from the mixed libraries glycerol stocks.
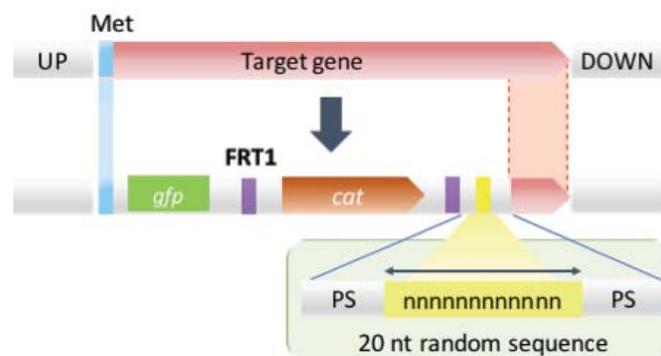
Figure 2.1 Genomic structure of target region in ASKA barcode deletion collection.

Each target ORF (open reading frame) were replaced with chloramphenicol resistant cassette flanked by FRT1 sites allowing excision using FLP recombinase (plasmid pCP20) and ~20 nucleotide random sequence (barcode).

Table 2.1 ASKA barcode collection: number of strain and barcode length distribution

Total strain	7082
Clone 1	3598
Clone 2	3484
Barcode length (20)	6489
Barcode length (>20)	24
Barcode length (<20); (<10)	569; (81)

2.2 Growth media

2.2.1 Luria Bertani (LB) medium

Dissolved 10.0 g tryptone, 5.0 g yeast extract and 10.0 g NaCl in 950 ml deionized water. Add 0.5 ml of 10N NaOH to adjust pH to 6.8 – 7 and bring volume up to 1L. Add 15.0 g of agar to the medium for LB agar plate. Autoclave on liquid cycle for 20 minutes at 15 psi. Store at room temperature. For LB agar plate preparation, allow the medium to cool to 55 °C and add antibiotic if needed. Store at 4 °C.

2.2.2 Antibiotics stocks

Chloramphenicol (Cm) [25 mg/ml] (Cas No. 56-75-7, Nacalai Tesque Inc.), Kanamycin Sulfate (KAN) (Cas No. 133-92-6, Wako Pure Chemicals), Streptomycin Sulfate (STR) (Cas No. 3810-74-0, Nacalai Tesque Inc.), Gentamicin Sulfate (GEN) (Cas No. 1405-41-0, Wako Pure Chemicals), Ciprofloxacin Hydrochloride (CIP) (Cas No. 86393-32-0, MP Biomedicals, LLC), Nalidixic acid (NAL) (Cas No. 389-08-2, Sigma-Aldrich), Tetracycline Hydrochloride (TET) (Cas No. 64-75-5, Wako Pure Chemicals) and Trimethoprim (TMP) (Cas No. 738-70-5, Nacalai Tesque Inc.). Except for chloramphenicol, all antibiotics were freshly prepared for each use.

2.3 Culture conditions: Long term stationary phase (LTSP)

Glycerol stocks of mix deletion libraries (clone 1 and clone 2) were transferred into 300 ml of LB medium with chloramphenicol [25 μg/ml] in 1L flask and incubated (Bio-Shaker BR-40LF, *TAITEC*, Japan) at 37 °C, 200 rpm under batch culture for 21 days with sampling at 0.5 (12 hours), 1, 2, 3, 4, 5, 6, 7, 8, 10, 14, 18 and 21 days. Approximately, 3 ml of culture will be taken at each time point, spin down (8 000g x 5 min), the cells pellet were washed with 0.9% NaCl solution and kept at -30 °C for DNA extraction (2 ml). Another 1 ml were used to monitor

the OD_{600nm}, CFU and for serial passage experiment. The supernatant were used to monitor the pH of culture during LTSP. For serial passage experiment, approximately 50 µl were transferred into 5 ml (100x dilution) of LB with chloramphenicol [25 µg/ml] and grown at 37 °C, 200 rpm until reached $OD_{600nm} \cong 2.0$. Then, the cells were harvested, washed and kept at-30 °C for DNA extraction. Serial passage were performed to distinguished between viable and non-viable mutant population during LTSP. Two experiment replicates were performed and referred to as R1 and R2 for LTSP, SPR1 and SPR2 for serial passage experiments. The experimental outline were as shown in Figure 2.2.

Figure 2.2 Experimental outline for LTSP experiments.

Glycerol stocks (clone 1 and clone 2) were transferred into 300 ml of LB medium with chloramphenicol (25 µg/ml) and incubate at 37 C, 200 rpm for 3 weeks with sampling (t = 0.5, 1, 2, 3, 4, 5, 6, 7, 8, 10, 14, 18 and 21 day). The optical density, cell viability (CFU count), DNA concentration and culture pH were monitored during LTSP. Serial passage experiment were performed at each time point, 50 µl were transferred into 5 ml LB and grown until reached $OD_{600nm} \cong 2.0$. Two independent replicate were carried out, assigned as R1 and R2 (LTSP) and SP_R1 and SP_R2 (serial passage).

2.4 Validation of GASP mutant candidates

2.4.1 Growth profile in LB medium

Each candidate strains were grown in 5 ml LB (25 µg/ml of chloramphenicol) at 37 °C, 200 rpm, overnight (16-18 hours). The overnight culture were diluted x100 and 200 µl were dispensed into 96-well plate. Optical density was measured at 600nm every 30 minutes for 20 hours using microplate reader (SpectraMax Plus 384, *Molecular Devices,* USA). Values were average of six replicates.

2.4.2 Complementation experiment

Candidate mutant strains were complemented with wild-type copy of the gene using IPTG inducible plasmid (TransBac plasmid clone library). Transformation were performed according to standard protocol[107]. Then, each candidate strains were grown in 5 ml LB (12.5 µg/ml of TET) at 37 °C, 200 rpm, overnight (16-18 hours). Overnight culture were diluted x100 with LB medium with (0mM or 1mM IPTG) and 200 µl were dispensed into 96-well plate. The growth profile of strains with/without IPTG induction were compared. Optical density were measured at 600nm every 30 minutes for 20 hours using microplate reader (SpectraMax Plus 384, *Molecular Devices,* USA). Values were average of six replicates.

2.4.3 Competition experiment

This experiment were performed to measure the fitness of GASP mutant candidates. Competition between candidate mutant strain (*Δgene*::Cm) and control strain (*ΔlacA*::KAN). Control strain with KAN resistance cassette were generate via lamda red homologuous recombination[86] using *E. coli K-12* strain *BW38028*. Overnight culture of each strain were diluted to OD_{600nm} ±0.2, then 100 µl of each were transferred into 5 ml of LB medium (no antibiotic) and incubated at 37 °C, 200 rpm for 24 hours. For independent growth (control, no competition), 200 µl of culture were transferred into 5 ml LB (no antibiotic) and incubated at 37 °C, 200 rpm for 24 hours. The CFU of each strain at initial (t=0) and final (t=24h) were determined by plate count on LB+Cm (25 µg/ml) and LB+KAN (30 µg/ml) agar plates. The fitness were determined by Equation 2.1. Values were average of three independent replicates.

$$Fitness\ (w) \equiv \frac{A_f/A_i}{WT_f/WT_i} \qquad\qquad Equation\ 2.1$$

A is the candidate mutant, WT is control strain, *f* is the final CFU and *i* is the initial CFU

2.5 Culture conditions: Drug treatment

Glycerol stocks of mix libraries (clone 1 and clone 2) were grown in 250 ml of LB medium in 1 L flask, 200 rpm at 37 °C until reached OD600nm $\cong$ 0.4, then selected drugs at sublethal (IC50) concentration were added to the culture and this time point were set as t = 0 hour. Approximately, 2 ml of culture were taken just before addition of drugs (t = 0) as control population (before treatment) and samplings were carried out at t = 3, 6, 9, 12, 18 and 24 hours to monitor the mutant population dynamics during sublethal drug treatment. The sublethal drug treatment were performed to (i) distinguished hypersensitive mutant population and (ii) to promote persisters formation. After 24 hours, the cells were washed with LB medium and regrown in new LB medium (20 ml in 100 ml flask) without drug (LB + chloramphenicol [25 µg/ml]) and with higher concentration of drugs (IC70, IC90, 1.5xIC90, 2xIC90 and 4xIC90). The initial cell concentration was fixed at OD600nm $\cong$ 0.4 and incubated at 37°C , 200 rpm for 24 hours. These experiments were carried out in duplicate for each drugs and the total population growth were monitored by OD$_{600nm}$ and colony forming units (CFU). However, only one replicate from each drug experiments (except control) were chosen for sequencing. The experimental outline were as shown in Figure 2.3.

2.6 Minimum inhibitory concentration (MIC)

The MIC of each drugs was determined using standard broth dilution protocol in 96 microtiter plate[108]. Overnight culture of *ΔlacA* (control strain) was grown in new LB medium until reached exponential growth phase (OD600nm $\cong$0.4), then 100 µl was dispensed into each wells containing varying concentrations of drug (100 µl) and kept overnight with shaking at 37°C. The OD value was read using microplate reader (SpectraMax Plus 384, *Molecular Devices*, USA). Dose-response curve was plotted (SoftMax Pro 4.3 Software) and the inhibitory concentrations (IC) of each drugs were calculated based on Equation 2.2. These results were shown in Supplementary Information (Supp. Figure 1 and Supp. Table 2)

$$IC(F) = (\frac{F}{100-F})^{\frac{1}{H}} \times IC50 \qquad\qquad Equation\ 2.2$$

H = Hill slope
F = IC value ($0 < F < 100$)

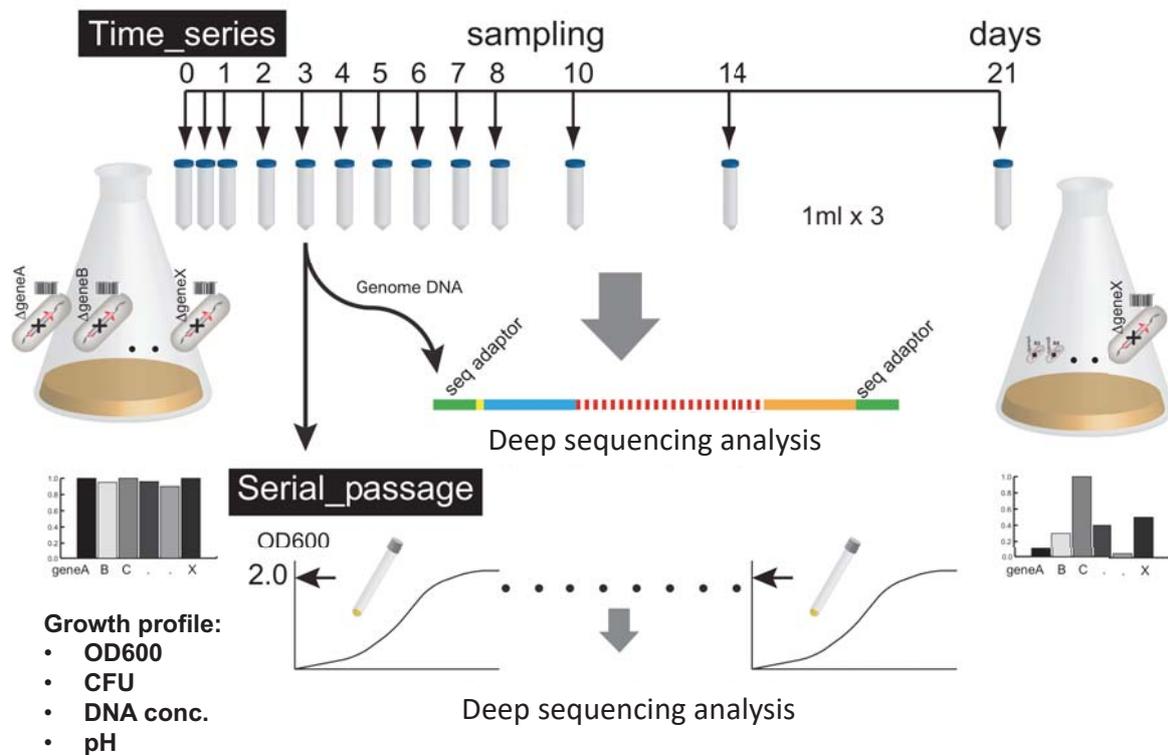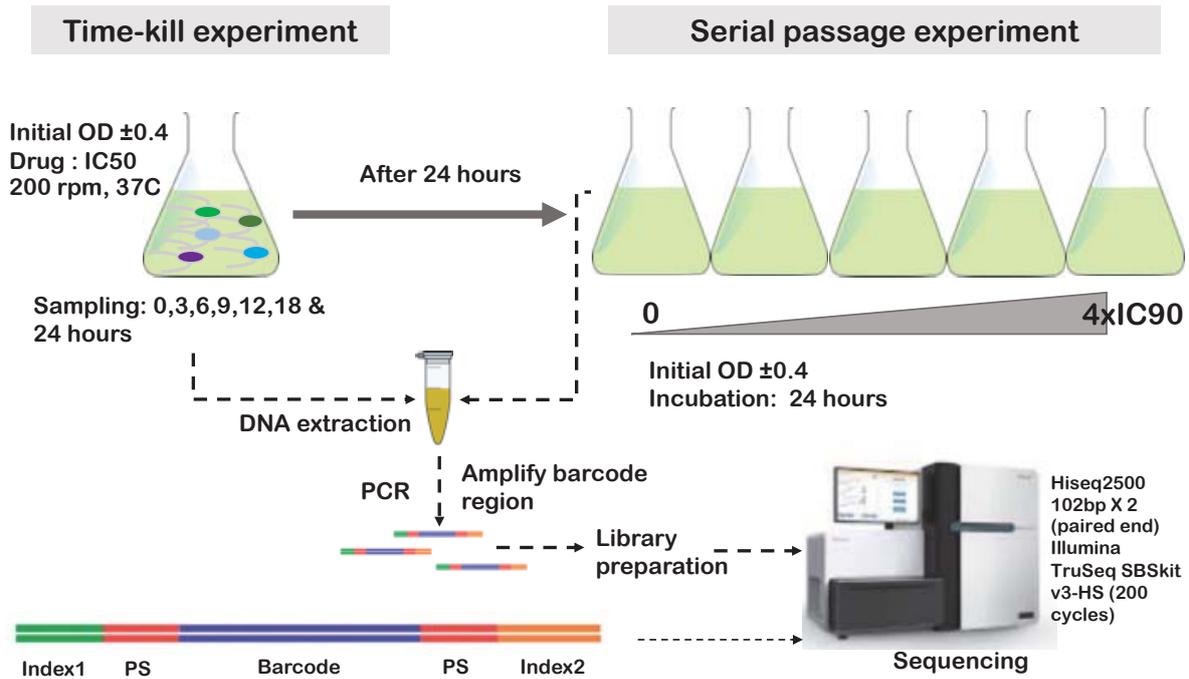*Figure 2.3 Experimental outline for drug persistence experiment.*

Glycerol stocks (clone 1 and clone 2) were transferred into 300 ml of LB medium and grown at 37 °C, 200 rpm until reached $OD_{600nm} \cong 0.4$ (2.5~3 hours), so the initial cell concentration before drug treatment is at ~$\times 10^8$ cells. First, the *E. coli* mutant population were exposed to sublethal (IC50) concentration of drugs and the population change (time-kill) during drug treatment (24 hours) were observed. Next, the cells that were exposed to IC50 concentration of drug were exposed to higher concentration of drug (serial passage) for 24 hours.

2.7 Susceptibility assay for drug tolerant mutant

This experiment were carried out using 96-well plate. Overnight culture in LB medium (25 μg/ml chloramphenicol) were regrown in new LB (no drug) until OD $\cong$ 0.4, 20 μl were dispensed into 180 μl of LB with different concentration of drug (IC0, IC50, IC70, IC90, 1.5xIC90, 2xIC90 and 4xIC90). Each plates were incubated at 37°C for 24 hours. The optical density (OD_{600nm}) were measured using microplate reader (SpectraMax Plus 384, *Molecular Devices,* USA). Three replicates were performed for each drug concentration. The OD reading were normalized (*z*-score) shown in Equation 2.3. Manhattan distance were used for clustering (Equation 2.4).

OD normalization (z-score)

$$y = \frac{(x - mean)}{sd} \qquad\qquad Equation\ 2.3$$

The *x* is OD value, mean is the average OD value in the dataset, sd is the standard deviation in dataset

Manhattan distance

$$D = \sum_{i=1}^{n} |x_i - y_i| \hspace{3cm} Equation\ 2.4$$

D is distance, x_i and y_i are the variables

2.8 Morphology

The cells morphology were observed under 1000x magnification (BZ-9000, Keyence, USA) and image processing were carried out using BZ-II Analyzer software.

2.9 Sample preparation for sequencing

2.9.1 DNA extraction

DNA extractions were carried out using CTAB method[109]. The DNA concentration of each samples were measured by fluorescence method using Quant-iT PicoGreen kit (ex:485nm/em:530nm) and fluorescence plate reader (SpectraMax Plus).

2.9.2 PCR amplification

PCR amplifications of the barcode regions were according to the Phusion High-Fidelity PCR Master Mix with HF buffer (NEB) protocol. Optimized PCR conditions (GeneAmp PCR System 9700, Applied Biosystems) were as follows; 18 cycles, initial (hot start); 98 °C (1 min), denaturation; 98 °C (15 sec), annealing; 60 °C (2 sec), elongation; 72 °C (40 sec); final elongation; 72 °C (40 sec) and final hold; 4 °C. Lower PCR cycles were used to minimize errors during amplification. The PCR product were check using gel electrophoresis, expected fragment size ~207 bp. The primers used in this study were shown in Supplementary Table 1.

2.9.3 MiSeq and HiSeq multiplexing

For normalization of each samples, we pooled equal amount of PCR products, gel electrophoresis and excise target band (~207 bp), followed by gel purification (QiaQuick Gel Extraction Kit, QIAGEN) to remove remaining primers and primer dimers within the samples. First, MiSeq was performed and the Illumina index (ID1 and ID2, Figure 2.4) for each sample were counted and number of reads per sample is set. Using this value, the concentration of each samples (PCR products) were adjusted, pooled, purified and sequenced using HiSeq. For LTSP, I had multiplex 105 samples (R1: 26, R2: 26, SP_R1: 20, SP_R2: 22 and other: 11). Two technical replicate were prepared for each sample in LTSP. For technical replicate, the same DNA template were amplified using different multiplex index. This was to check if there

are any anomalies due to sample preparation or sequencer. Meanwhile for drug experiment, I had multiplex 90 samples (No drug: 16; CIP: 12, GEN: 13, KAN: 13, STR: 12, TET: 12 and TMP: 12) per lane (Illumina HiSeq 2500; 100 bp x 2 (paired end); Illumina TruSeq SBS kit v3-HS (200 cycles)).

2.10 Sequence data analysis

The sequence reads were collected from the sequencer as fastq files. Each read contained multiplex index 1, primer 16, barcode sequence, primer 1 and multiplex index 2 in the following order (Figure 2.4) . The files were analyzed with custom Python script. Then, we had applied filtering to obtained the correct read sequence. First, we aligned the reads to common primer (primer 16 or primer 1), then we split the sequence into multiplex index tag sequences and barcode sequences. Next, the multiplex index tag was exactly mapped against the list of expected combinations of index tag. Then, we mapped each barcode to the list of barcodes in our strains database. Only those that precisely matched for index and barcode were accepted, those that did not match were discarded. Valid sequence were counted and sorted and frequency table were generated. Figure 2.5 showed the overview of sequence data analysis. Pearson's correlation coefficient were measured between the each of the technical replicates. Samples with low correlation ($r < 0.8$) were discarded. Next, correlation between clone 1 and clone 2 for each time series were measured and strain with correlation, $r < 0.7$ were discarded. This low correlation might be due to secondary mutation in one or both strain that may affect the growth pattern of that particular mutant strain, thus this mutant was discarded.
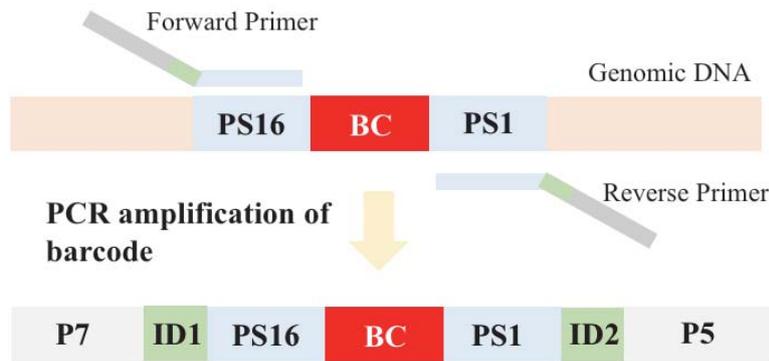
Figure 2.4 Scheme for Illumina multiplex library preparation.

The forward and reverse primer consist of P7/P5 Illumina adapters, 5 nt random sequence, multiplex index tag (ID1/ ID2, 9 nt) and common primer sequence (PS16/PS1). The P7/P5 Illumina adapter function to binds to the flow cell. The multiplex index (ID1/ID2) used to identify the sample from which it originated. The 5nt random sequence was added to increase the efficiency for template generation (increase read yields and quality scores) for amplicon sequencing on Illumina sequencer (recommended for low-diversity samples). List of primers sequences is in Supplementary Table 1.
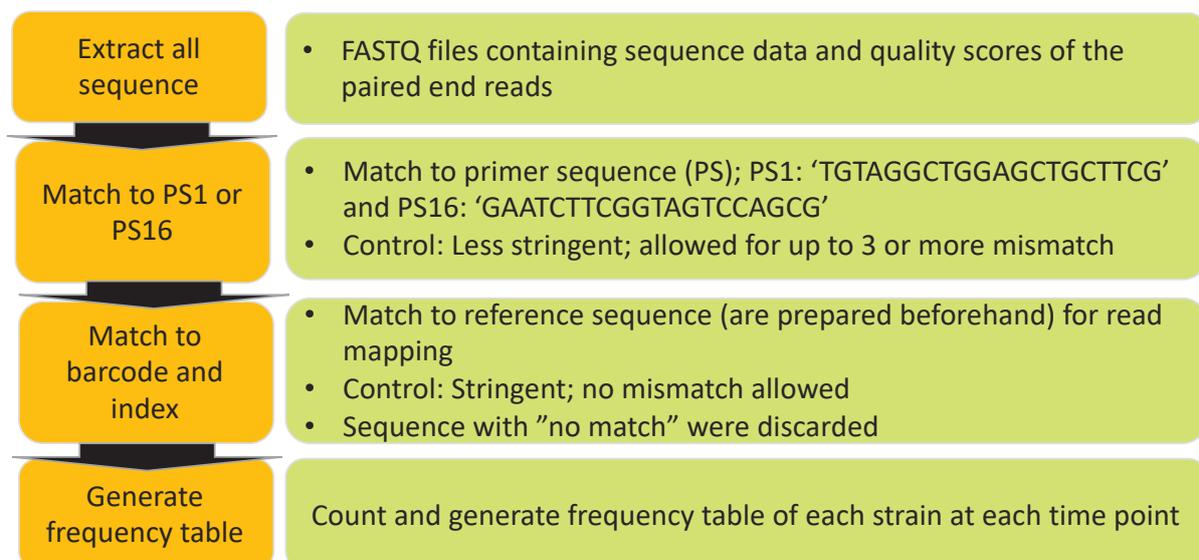
Figure 2.5 Overview for sequencing data analysis.

2.11 Statistical analysis

For statistical analysis, I used Excel, R and R packages (ggplot2, reshape2, ComplexHeatmap, corrplot) to plot graphs.

2.12 Bioinformatics analysis

Functional enrichment analyses were performed for selected gene set were analyzed using open software FunRich (funrich.org). Uniprot bacteria database (ftp://ftp.uniprot.org/pub/databases/uniprot/current_release/knowledgebase/taxonomic_divisions/uniprot_sprot_bacteria.dat.gz) and Ecocyc was used for database search. Statistical cut-off of enrichment analyses in FunRich software was kept as default with a *p*-value < 0.05 after *Bonferroni* correction.

Chapter 3 <u>RESULTS</u>

3.1 Confirm reproducibility of sequencing and experimental data

The sequenced data were processed as mentioned in Section 2.10 in Materials and methods. Almost 90% of the raw sequences were successfully mapped to reference sequences. Details were shown in Table 3.5 for LTSP and Table 3.10 for drug treatment.

The reproducibility of the sequencing results is important to ensure the reliability of the experiment. To check for the reproducibility of the sequencing and experiment data, I determined the correlation between each technical replicates and between experimental replicates. Technical replicates were performed to check for sequencing bias during Illumina sequencing. While experimental replicates were performed to check for the reproducibility of the ASKA barcode deletion collection. For LTSP experiment, I had carried out two experiment replicate; R1 and R2. For technical replicate, I used the same DNA template amplified with different Illumina index (primers). High correlation were observed for time-course experiment R2 and serial passage experiment (SP_R1 and SP_R2, respectively) as shown in Table 3.1. This indicate that there is no problem with sample preparation and sequencing run. However, Day 10-21 of experiment R1 showed low correlation and this data will not be used for further analysis. This problem may be due to insufficient genomic DNA for PCR amplification during the later stationary phase. High correlations were also observed for independent experiment replicates as shown in Table 3.2. This indicate a good experimental reproducibility using ASKA barcode deletion library. Likewise, high correlation were observed for technical (Table 3.3) and experimental replicates (Table 3.4) for drug treatment experiment. Overall, the results indicates that there were no sequencing bias between samples during sequencing and experiments performed using ASKA barcode deletion collection were reproducible.

Table 3.1 Pearson's correlation between technical replicate for LTSP

R1		R2		SP_R1		SP_R2	
Day	*r*	**Day**	*r*	**Day**	*r*	**Day**	*r*
0.5	0.98	0.5	0.96	0.5	0.98	0.5	0.97
1	0.97	1	0.94	1	0.97	1	0.98
2	0.96	2	0.96	2	0.99	2	0.99
3	0.98	3	0.99	3	0.99	3	0.99
4	0.99	4	0.99	4	0.99	4	0.99
5	0.99	5	0.99	5	0.99	5	0.99
6	0.99	6	0.99	6	0.99	6	0.99
7	0.99	7	0.97	7	0.99	7	0.99
8	0.99	8	0.98	8	0.99	8	0.99
10	0.02	10	0.99	10	0.99	10	0.99
14	0.3	14	0.99	14	NA	14	0.99
18	0.15	18	0.99	18	NA	18	NA
21	0.4	21	0.99	21	NA	21	NA

NA; no growth were detected at these time point

Table 3.2 Pearson's correlation between experimental replicate for LTSP

LTSP		Serial passage	
Day	*r*	**Day**	*r*
0.5	0.96	0.5	0.98
1	0.94	1	0.98
2	0.85	2	0.53
3	0.84	3	0.82
4	0.92	4	0.9
5	0.96	5	0.97
6	0.92	6	0.64
7	0.6	7	0.33
8	0.53	8	0.72
10	NA	10	0.4
14	NA	14	NA
18	NA	18	NA
21	NA	21	NA

NA; no growth were detected at these time point

Table 3.3 Pearson's correlation between technical replicate for drug experiment

Experiment	r
KAN_6h	0.94
KAN_4xMIC	0.98
TET_6h	0.92

Table 3.4 Pearson's correlation between experiment replicate (no drug) for drug experiment

Hour	r
0	0.98
3	0.97
6	0.97
9	0.96
12	0.94
18	0.96
24	0.97
SP_LB	0.97

3.2 Long-term stationary phase (LTSP)

3.2.1 Sequencing result

The sequenced data were processed as mentioned in Section 2.10 in Materials and methods. Details of the sequencing results were shown in Figure 3.1 and Table 3.5. I had obtained ~ 273 million valid reads (~ 90% of total sequences reads) with coverage of ~ 360 reads per strain. Next, I make a new dataset using the average of read 1 and read 2. Since technical replicate were carried out for each sample, thus I checked the pairwise Pearson's correlation for each of the sample and those with correlation coefficient, $r \geq 0.8$ were accepted. Then, I make new dataset using the average of each technical replicate. Here, the correlation coefficient for technical replicate Day10-21 (R1) of LTSP were lower than threshold and were removed from dataset. Here, I have four datasets; R1, R2, SP_R1 and SP_R2. Next, I removed outlier strains from the R1 (195 strains) and R2 (245 strains) dataset. Next, I determine the correlations between clone 1 and clone 2 in each dataset. Correlation for strains that only have no biological replicate (no clone 2) were determine using experimental replicate (R1 and R2). I also checked the outlier strains and found 166 strains were common between the two dataset. As such, I suspect that this occurrence may not be noise but may have some biological meaning thus the strains with high correlation coefficient ($r \geq 0.7$) were reintegrate to the dataset. Finally, the number of valid strains in each dataset; R1 (2458), R2 (1660), SP_R1(3412) and SP_R2 (3379), respectively.
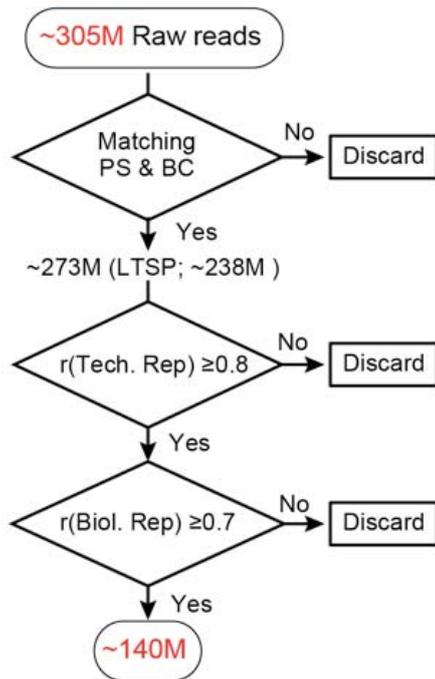
Figure 3.1 Outline of Sequencing result for LTSP

For LTSP;105 samples, ~305 million raw sequences reads has been obtained. Next, the raw sequences were map to reference, sort and count, 273, 128 700 valid reads (~ 90% of total sequences reads) were matched. The total reads for LTSP experiment only is ~238 million (read 1: 124 232 661 reads and read 2: 114 280 094 reads). Pairwise Pearson's correlation for each of the sample (technical) and those with correlation coefficient, $r \geq 0.8$ were accepted. Day10-21 (R1) of LTSP were removed from dataset. Next, I determine the correlations between clone 1 and clone 2 in each dataset. Strains with r >0.7 accepted for further analysis. Correlation for strains that only have no biological replicate (no clone 2) were determine using experimental replicate (R1 and R2). Finally, the number of valid strains in each dataset; R1 (2458), R2 (1660), SP_R1(3412) and SP_R2 (3379), respectively.

Table 3.5 Detailed sequencing result for LTSP

Experiment	Description	Result
LTSP	Multiplex	105 (94 samples LTSP)
	Total raw sequences	304 764 266
	Valid reads (Mapped to reference sequences), ALL	273 128 700 (89.6%)
	Valid reads (Mapped to reference sequences), LTSP	238 512 755
	Number of valid strain ($r \geq 0.7$) R1 R2 SP_R1 SP_R2	2458 1660 3412 3379

3.2.2 Growth profile and cell morphology of Escherichia coli mutant during LTSP

I had monitored *E. coli* mutant population dynamics during LTSP in LB medium for three weeks shown in Figure 3.2.. The OD readings (Figure 3.2A) showed fluctuation at day 2 and day 14 but mainly remains constant throughout the incubation period. These fluctuations could be due to lysis of dead cells and released of nutrients for uptake of the surviving populations. The viability of culture (Figure 3.2B) was observe using solid (LB agar plate) by colony formation and liquid medium via serial passage experiment (Figure 3.3). However, viability using solid media could be attained only up to day 10, meanwhile viability was observed up to day 14 for SPR2 using liquid medium (Figure 3.3D). The DNA concentration showed decreased throughout the LTSP (Figure 3.2C). This result correlates with the CFU result.

Growth of *E. coli* in rich media such as LB broth exposes the cells to elevated pH as organic acids are consumed[110], whereas glucose minimal medium would be acidified by acidic fermentation products, though such media are typically buffered to a near-neutral pH (pH 6.8-7). The effect of pH is important as extreme pH condition can also alter protein conformation or even denature proteins, thus the pH of the culture may shift the dynamics of the population survivability during LTSP. I observed rapid increased of pH (Figure 3.2D) from ~ 6.0 at 12 hours to pH ~ 8.5 by Day 2 and slowly increase to pH 9 by Day 10 which remains throughout the experiment period. Hence, increase in pH contributes to cell death and influence the timing of entry into the LTSP[111]. Another factor, increased glycation (non-enzymatic glycosylation) was showed to have led to cell death in *E. coli* by increased of oxidative stress levels[112]. Furthermore, changes in pH levels were found to affect the glycation levels. Finkel & Kram (2015) had observed that when pH levels rise above 8.5, the glycation levels increase greatly[113].

Serial passage experiment was carried out to distinguished the viable population and the non-viable during the LTSP shown in Figure 3.3. A fraction of the culture was diluted to 1:100 in new LB medium and grown until reached OD ~2.0. Longer recovery time (lag time) was observed when aged culture was regrown to new LB. This was assumed due to the time for the cells to synthesis cellular components necessary for growth, including repair of macromolecular damage that accumulated during stationary phase.

Overall, two significant changes that occur in the medium during stationary phase may play important roles in the precipitous loss of viability that follows; the depletion of readily available nutrients and an increase of pH in medium.
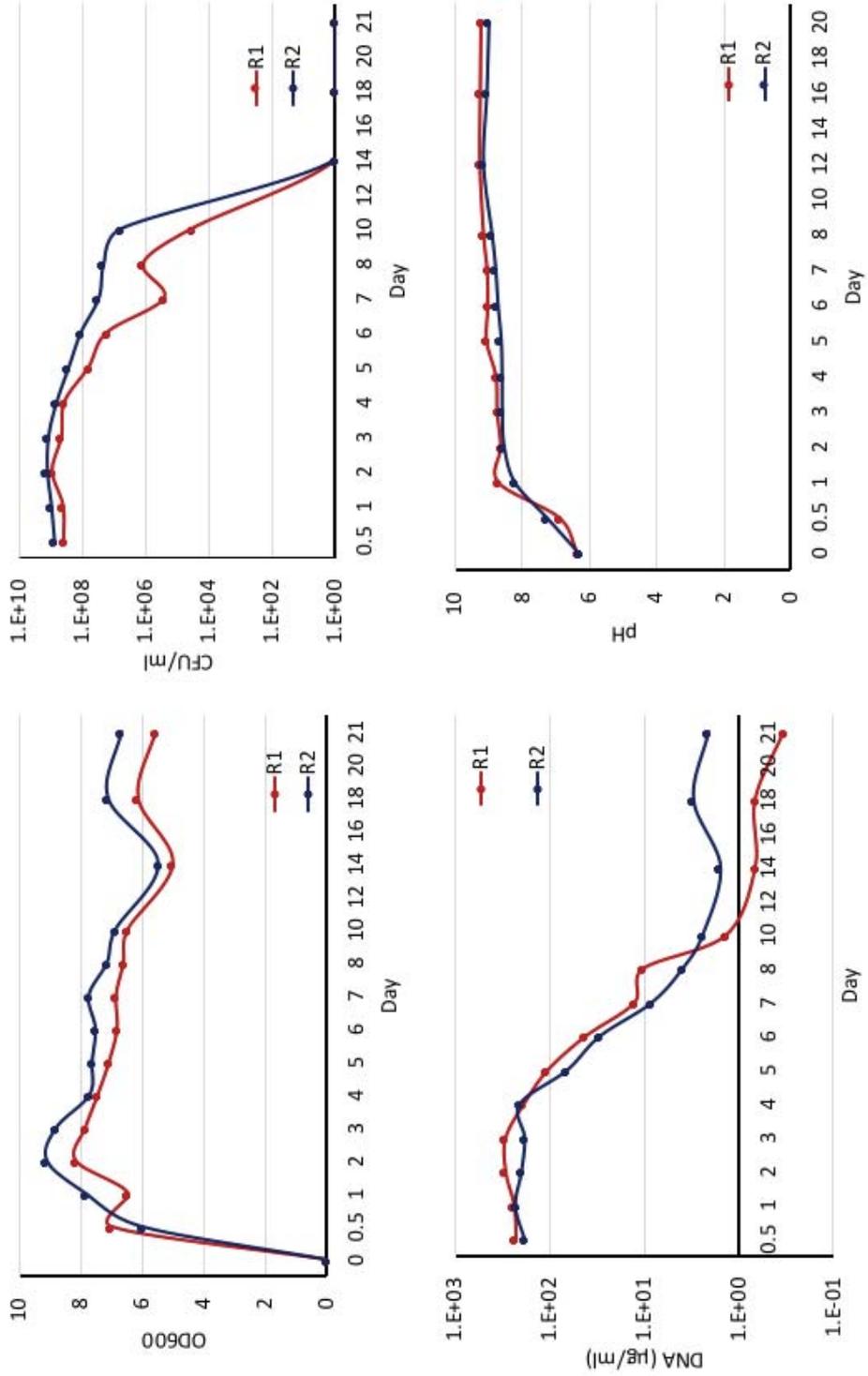
Figure 3.2 Growth profiles of E. coli mutant population (total) during LTSP.

During LTSP, the OD readings (A) showed fluctuation at day 2 and day 14 but mainly remains constant throughout the incubation period while the cells viability (B) or cfu count decreased and no growth were observed after Day 10. The DNA concentration (C) decreased during LTSP and the culture pH (D) increased from pH 6.5 to 8.5 by Day 2 and kept increasing to pH9 by Day 10. The pH maintained at pH 9 until Day 21.

41
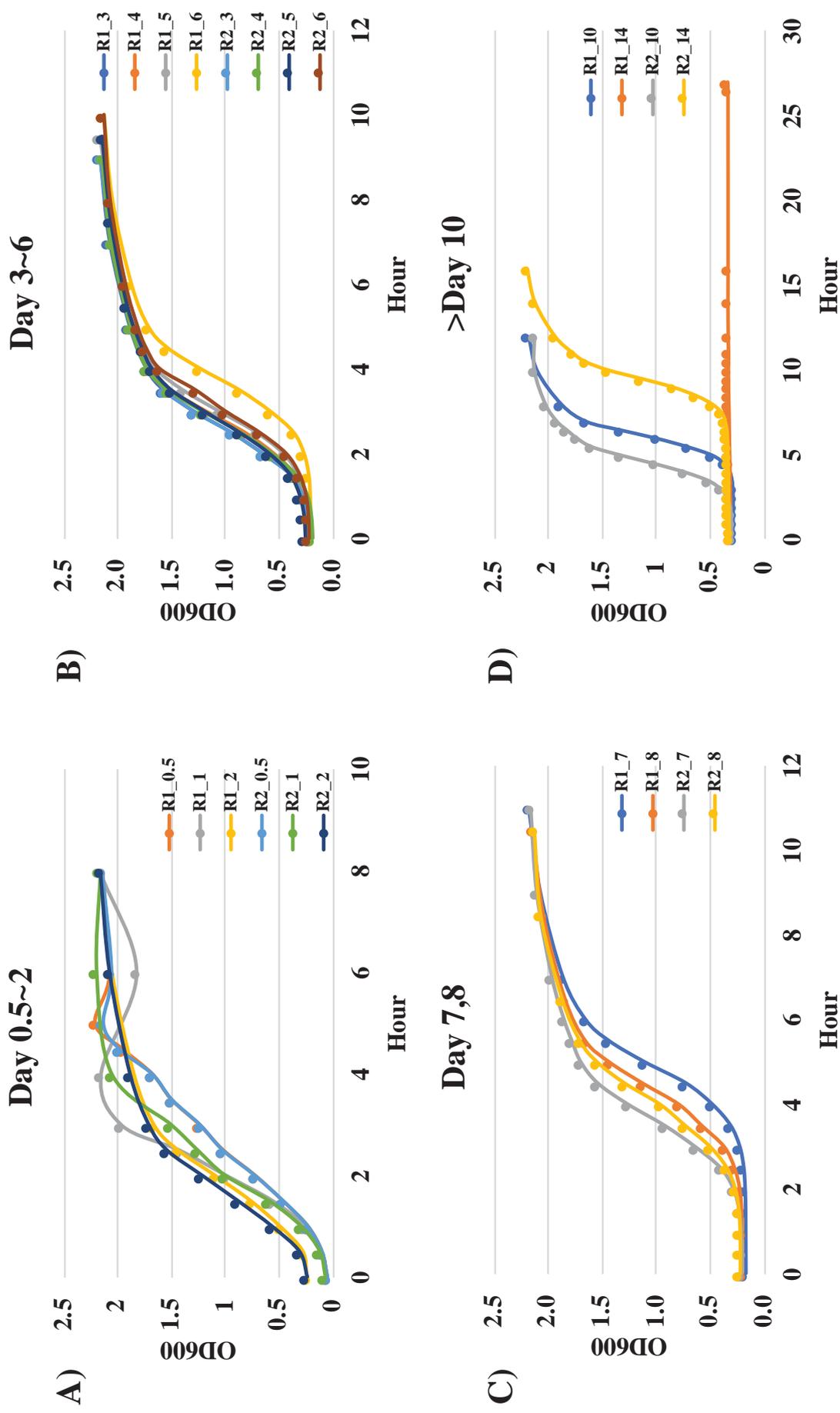
Figure 3.3 Growth profile (OD600nm) of E. coli mutant population (total) during serial passage.

Longer lag time were observed as aged cells were regrown in new LB medium. (A) Day 0.5-2, (B) Day 3-6, (C) Day 7-8 and (D) Day 10-14 .

Figure 3.4 showed the *E. coli* cells morphology during LTSP. The cells morphology from Day 0.5 until Day 2 did not show any morphological changes and is similar to WT. However, starting from Day 3 onwards (to Day 21) several morphological changes were observed; (i) formation of dense body(s) in the cell, (ii) irregular shape and (iii) decreased in cell size. Dwarfing is continuous reduction of cell size were observed after reductive division. The dwarfing process might be beneficial during starvation conditions because the surface/volume ratio of the cells increased thus improve uptake of nutrients from surrounding. Dwarfing is a form of self-digestion; whereby cells degrades its own protein for energy which is an important survival mechanism during nutrient starvation[114].
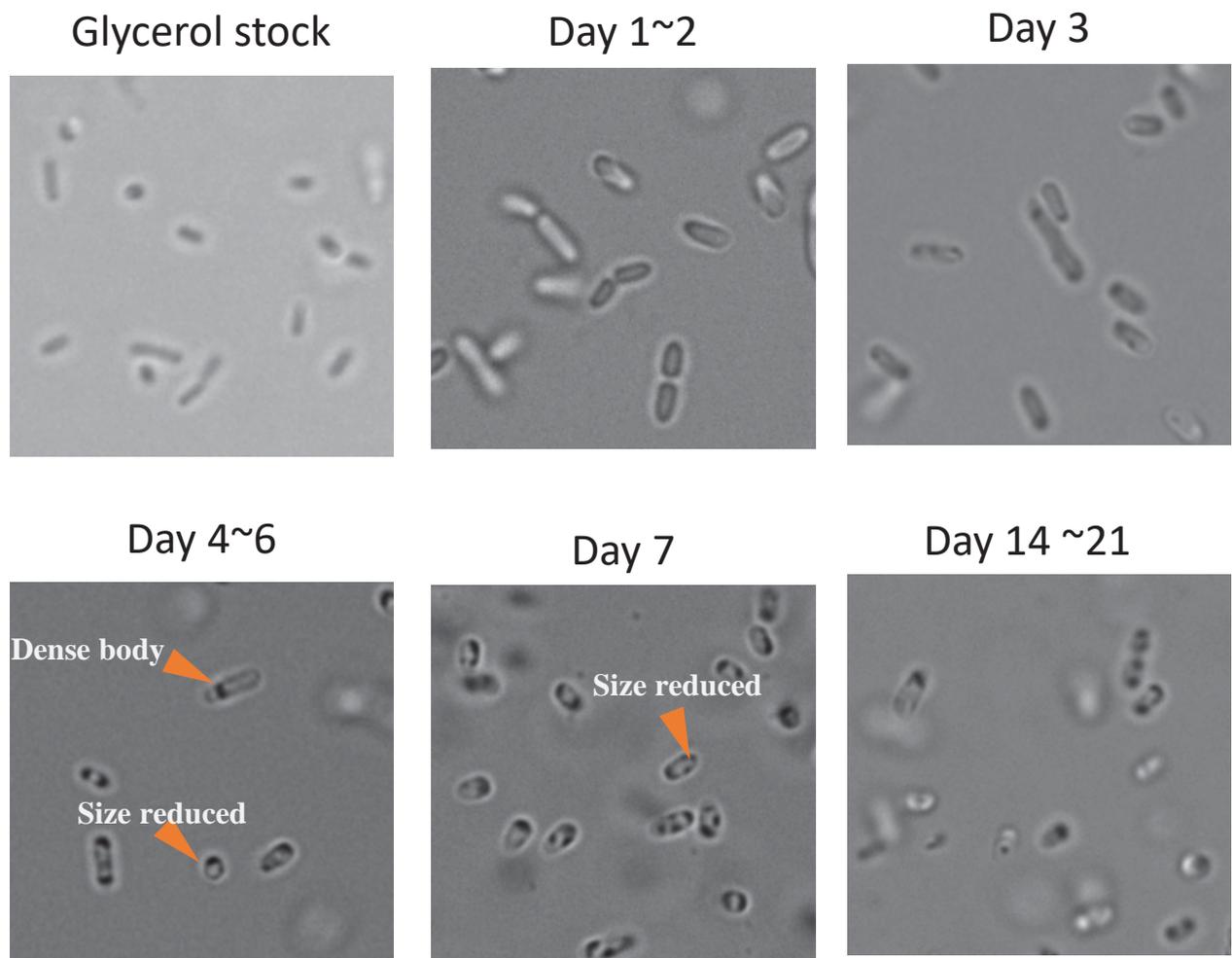
Figure 3.4 E. coli cell morphology during LTSP.

The cells morphology from Day 0.5 until Day 2 did not show any morphological changes and is similar to WT. However, starting from Day 3 onwards (to Day 21) several morphological changes were observed; (i) formation of dense body in the cell, (ii) irregular shape and (iii) decreased in cell size.

3.2.3 Survival dynamics of E. coli mutant population during LTSP

Survival has been defined as maintenance of viability under adverse circumstances. If nutrient deprivation continues, an initial loss of viability (death phase) is observed, but the population then stabilizes and enters long-term stationary phase. This is a state of dynamic equilibrium in which individual cells may die or divide but the community as a whole survives indefinitely, due in part to selection for subpopulations that are better adapted to the current nutritional conditions[48,49]. This surviving population is defined as growth advantage in stationary phase (GASP) population.

Figure 3.5 (A) showed the dynamics of each mutant population during LTSP. Each of the colored box represent one mutant strain. From the figure, change in the mutant population dynamics starts by Day 3. This is quite interesting as Finkel[48] had observed that *E. coli* cells enters death phase by Day 3 in rich medium. This suggest that this changes in mutant population were due to entry of death phase. It was noted that in batch culture, only *E. coli* cells grown after 10 days will express the GASP phenotype[46]. However, this is because mutations were only observed after 10 days. Since, I started with mutant populations from the beginning, it is possible to observed population with GASP earlier.

From Figure 3.5 (A), the first population change occurred at Day 3. Thus, I want to know if this change in population were maintained throughout the 21 days or will changes happen again. To check, I calculated pair-wise correlation between each time-point. If no population changes occurred, then the correlation between each time point should be highly correlated. Figure 3.5 (B) showed the pair-wise correlation of each time point during LTSP for experiment R1 and R2, respectively. I observed that there is a time-dependent population changes that occur during LTSP. The population changes can be seen in four clusters; Cluster 1 (Day 0.5 to Day 2), Cluster 2 (Day 3-7), Cluster 3 (Day 8-10) and Cluster 4 (Day 14-21) as shown in R2 (as mention previously, data for R1(Day 10-21) were not used for analysis). This indicate that mutant population abundance during this time were similar. Overall, these results showed that it is possible to monitor *E. coli* mutant population changes from mix culture by Bar-seq.

(A) Population dynamics of each mutant during LTSP. Changes in population dynamic occurred by Day 3 which may be due to entry of death phase[48] in rich medium. (B) Pair-wise correlation (Pearson's) between technical replicates showed time-dependent population changes that occur during LTSP. These changes can be seen in four clusters (R2); Cluster 1 (Day 0.5 to Day 2), Cluster 2 (Day 3-7), Cluster 3 (Day 8-10) and Cluster 4 (Day 14-21). Population abundance in each cluster is similar.

Figure 3.5 Survival dynamic of E. coli mutant population during LTSP.

45

3.2.4 Diversity within the culture decrease during serial passage

How changes in the population abundance would affect the mutant diversity in the LTSP culture? In experiment using WT cells, the genetic diversity is expected to increase as the mutant population with GASP arise during LTSP. Since I am using bottom-up approach, I expect that the mutant diversity would decrease due to "loss" of mutant population during competitive growth. To answer this question, I calculate the diversity index at each time point using Shannon entropy. Shannon entropy was adopted by ecologists as a diversity measure. Shannon entropy compute the uncertainty associated with identifying species in the community based on the total number of individuals, their identities, and their relative abundances[115]. A higher degree of uncertainty means greater diversity in the community.

Interestingly, the diversity during LTSP only showed slight decreased at Day 6/7, then increased and the entropy value maintained throughout the incubation. This decrease may be due to high abundance of certain mutant since Shannon entropy measures the 'species' in proportion of their population abundance. This result suggest that most *E. coli* cells enters dormant state or VBNC. Meanwhile, the mutant population diversity decreased during serial passage reflects the selection pressure as the competing low fitness mutant were displaced by mutant with higher fitness during LTSP. At the same time, using serial passage we are unable to clarify whether the cells are VBNC or not since some strains might have longer lag time or were loss due culture saturation by other mutant populations.
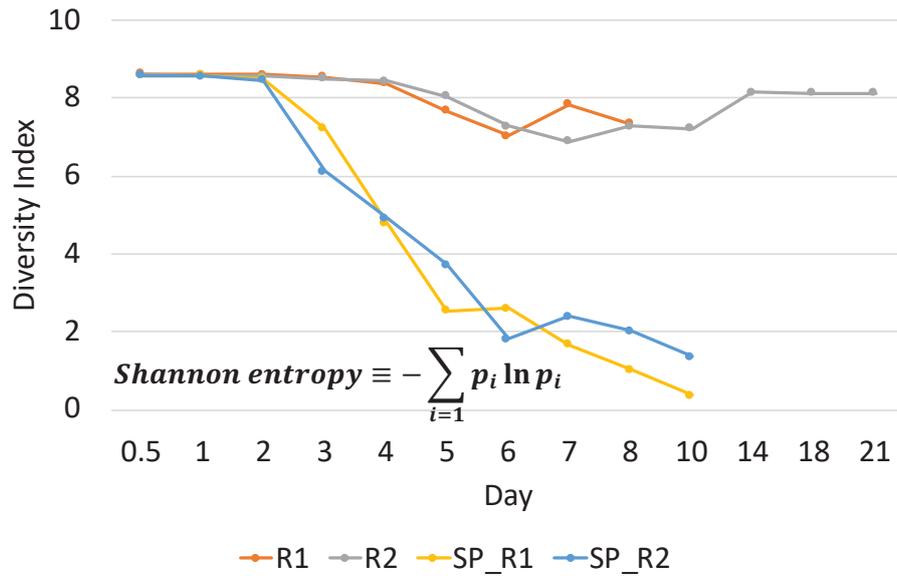
Figure 3.6 Genetic diversity of E. coli population during LTSP and serial passage.

The mutant diversity were maintained during LTSP (orange and grey line), suggesting the *E. coli* mutant population became dormant while mutant diversity decrease during serial passage (yellow and blue line) indicate surviving population during LTSP. p_i is the relative abundance of each population at each time point

3.2.5 Population size during LTSP affect the population abundance in serial passage

The serial passage experiment was carried out to differentiate between the viable and the non-viable population during LTSP. At each sampling point during LTSP, 50 μl of culture is transferred (100x dilution) in 5 ml of new LB medium and incubate at 37 °C, 200 rpm until reached OD600nm ~2. During LTSP (Figure 3.5), GASP mutant population increase in population abundance thus this population most probably be dominant population in serial passage. To check if this influence the population dynamics during serial passage, I determined the correlation between LTSP and SP. High correlation between time-course experiment and serial passage experiment were observed, as shown in Figure 3.7(A). This result could also be seen from the changes in the population dynamics during the long term stationary phase and serial passage shown in Figure 3.7(B). Although, no correlation was observed between at Day 14. This results indicate that the initial population size (time-course experiment) influenced the survival of each mutant populations during serial passage.
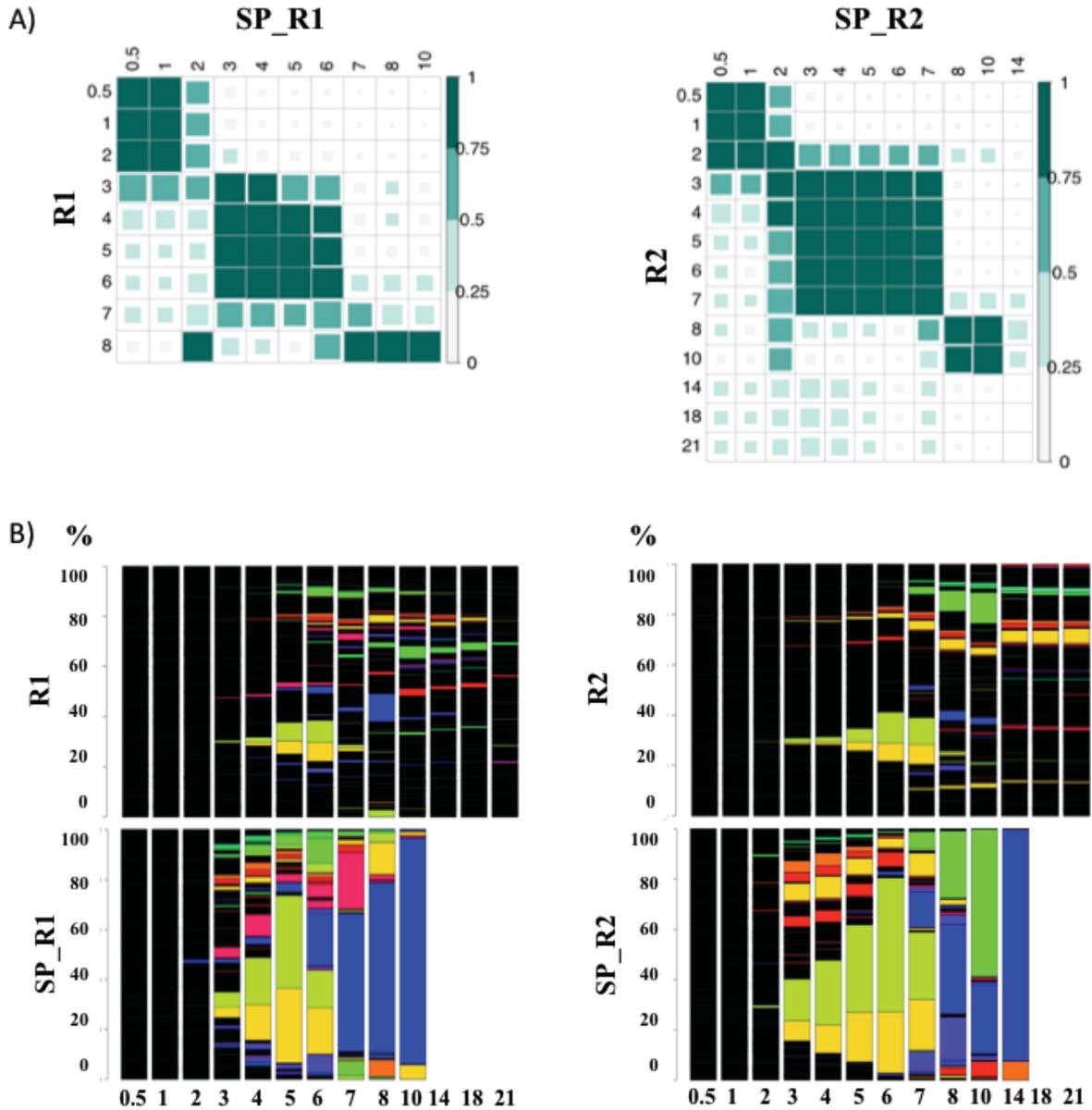
Figure 3.7 Initial population size during LTSP affect the population abundance during serial passage.

(A) Pair-wise correlation between LTSP and serial passage at each time point were highly correlated. (B) Population dynamics during LTSP (*top*) and serial passage (*bottom*) showed that mutant population with higher fraction during LTSP were also highly abundance during serial passage.

3.2.6 Enrichment analysis for top 100 rank mutant

During prolonged incubation, occurrence of secondary mutations could happen to some mutant strains which impact that population growth therefore affect the reliability of the result. As a means to improve our conception, we had two individual clone (different barcode) for each deletion mutant that will act as biological replicate in the same culture. Thus, only strain that showed consistency ($r \geq 0.7$, Pearson's correlation), which consist of 5404 mutant strains (2702 genes) and 3808 mutant strains (1904 genes) for R1 and R2, respectively that were considered valid for further analysis.

Pair-wise correlation of each time point in Figure 3.5(B) showed time-dependent population changes occurred during the 3 weeks incubation. These time-dependency can be divided into 4 clusters; day 0.5-2, day 3-7, day 8-10 and day 14-21. I categorized these clusters; cluster 1 (Day 0.5-2), cluster 2 (Day 3-7), cluster 3 (Day 8-10) and cluster 4 (Day 14-21). Using these information, I ranked the mutant based on abundance at each time point and select the top 100 mutant for gene list analysis for over-representation.

Mutant with mutation involved in transcription, transcription regulation and response to DNA damages showed to be significant (*Bonferroni*, $p < 0.05$) in all these clusters. For confirmation, I also performed enrichment analysis using the common top rank mutant during serial passage (92 strains, R1 and R2). The top ranked mutant in serial passage have mutation involved in transcription (24%), signal transduction (8%) and cellular response to DNA damage (6%). Figure 3.8(A) showed the number of genes with significant GO terms ($p < 0.05$, R2 experiment) at each time point. Similar enrichment were obtained from experiment R1 as shown in Figure 3.8(B). The enrichment results suggest that mutations in genes involved in transcription regulation gives growth advantage to cells during LTSP.

A)

Biological process
Day 0.5, Day 1 and Day 2

Biological process
Day 3 to Day 7

Biological process
Day 8 and Day 10

Biological process
Day 14, Day 18 and Day 21
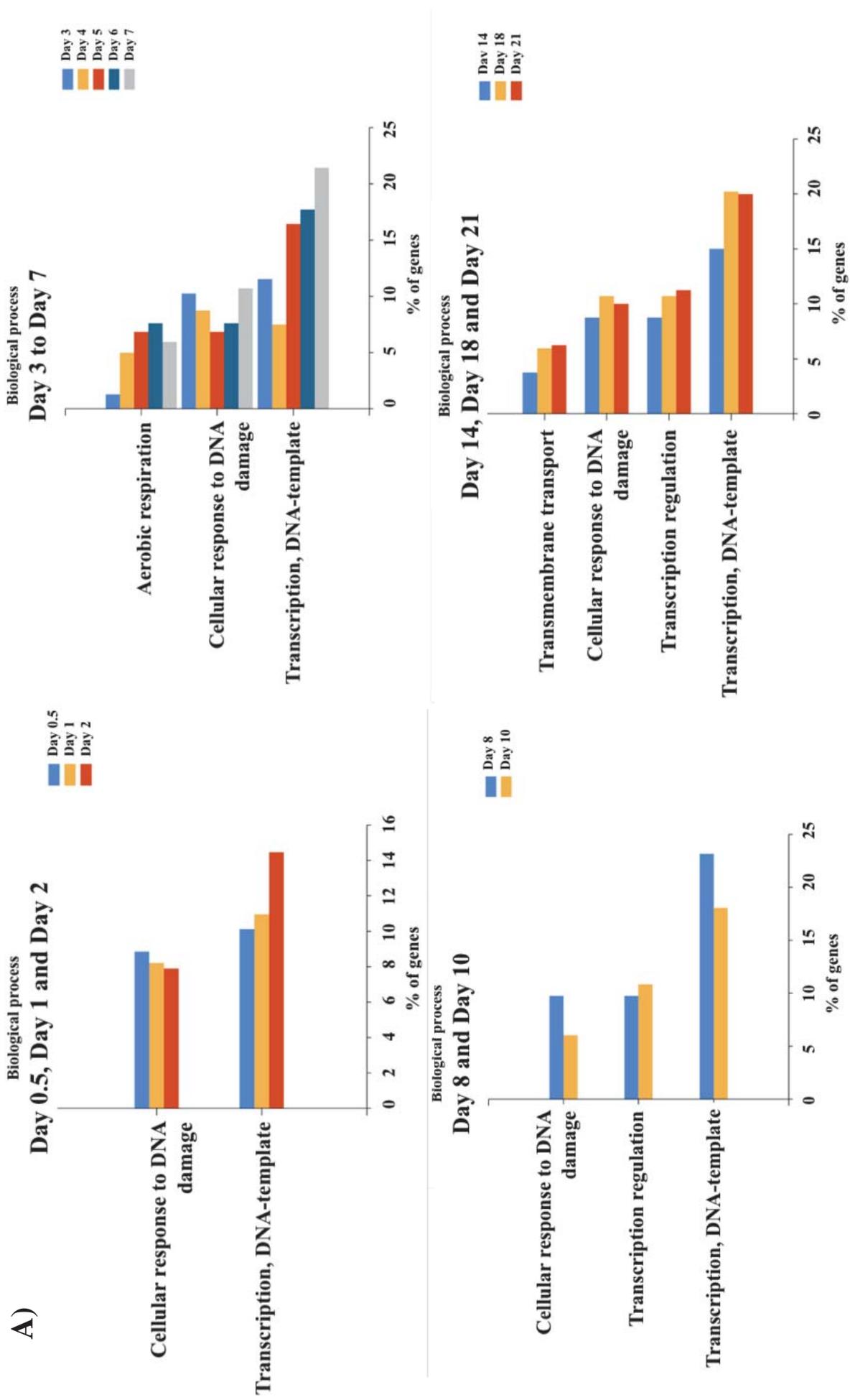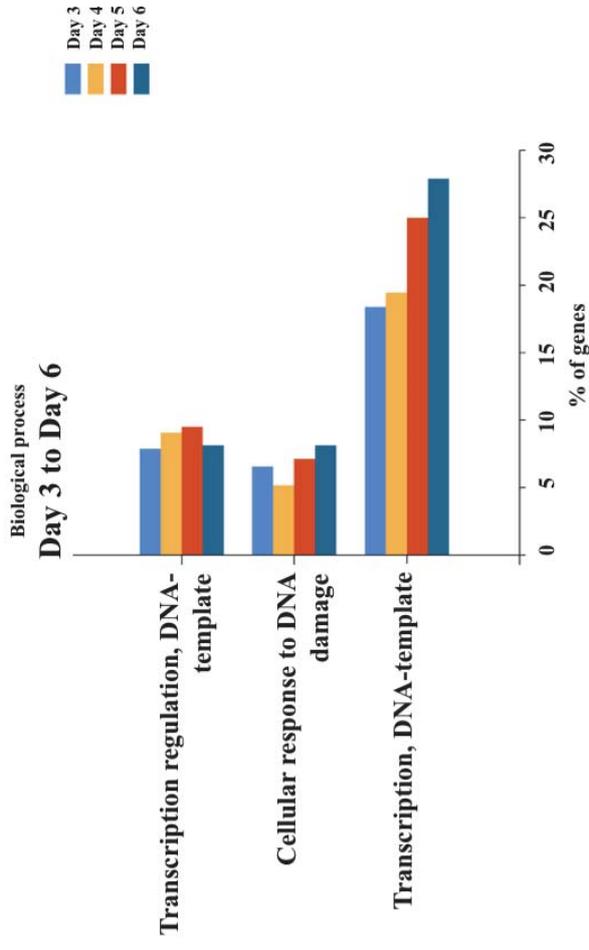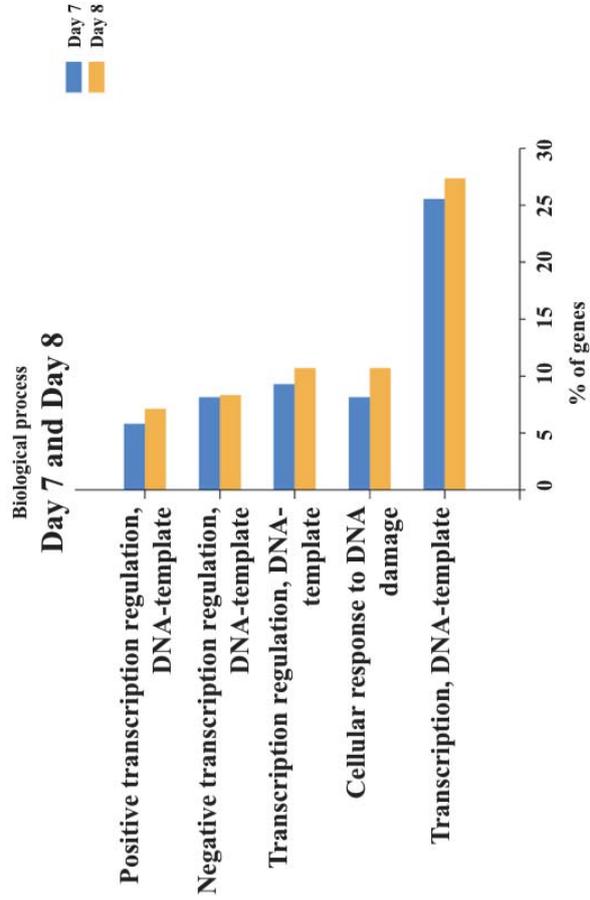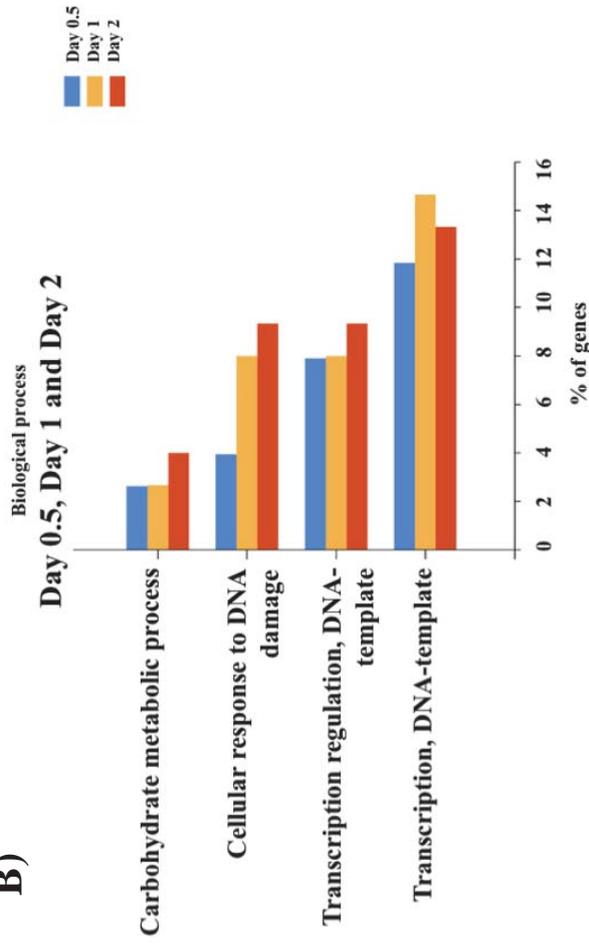
Figure 3.8 Gene enrichment analysis for each cluster during LTSP

(A) Experiment R2 and (B) Experiment R1. Number of genes (%) in the significant GO terms (p <0.05, Bonferronni correction). Mutations involved in transcription and cellular response to DNA damage were significantly enriched GO terms for all cluster.

3.2.7 Does growth phenotypes important for survivability during LTSP?

In the ASKA library, each mutant strain may exhibit different growth phenotypes; normal, slow or fast growth. Slow growth is one of the survival traits for drug tolerance population[116]. However in the case of no drug treatment does slow growth phenotype able to maintained their population or would they be displaced by the normal/faster growing mutant population. Thus, I plot the population abundance (log10) against their maximum growth rate (MGR) at Day 0.5, Day 3, Day 7, Day 8, Day 14 and Day 21 and observed the phenotypes of the surviving mutant population during LTSP as shown in Figure 3.9. The MGR for each mutant strain have been determined using Colony-Live[117].

Figure 3.9 showed the distribution of each mutant population abundance against MGR at Day 0.5, Day 3, Day 7, Day 8, Day 14 and Day 21. These time-point represent the entry and exit of each cluster shown in Figure 3.5 (B). For experiment R2 shown in Figure 3.9 (A), mutant populations that showed GASP phenotype at Day 3 were *rpoS, slyA* and *cpxA* mutant. The number of GASP mutant population increase by Day 7 and Day 8 (*cpxA, slyA, kdsC, kdsD, yfaU, hsrA*). Several mutant population showed growth advantage up to Day 7; *tdcA, rpoS, mraZ* mutant and new GASP population were observed from Day 8; *mprA, ygaH, ygaZ ,endA* mutant. Table 3.6 showed list of GASP mutant for each time points for R1 and R2 experiment during LTSP and serial passage.

Interestingly, I observed that mutant population that exhibit growth advantage during LTSP mostly showed normal growth phenotype. In addition, although some slow growth mutant were "loss" (barcode count < 10) during LTSP, most mutant population regardless of growth phenotype seems to persist in the culture (barcode count ≥100) and become dormant. Overall, the growth phenotypes did not seem to affect the *E. coli* mutant survival during LTSP.

Figure 3.9 Population abundance of each mutant log$_{10}$(barcode count) during LTSP against growth phenotype (MGR).

(A) Experiment R2: Day 0.5 showed the distribution of the initial abundance for each strain (also see histogram). Blue points were mutants with slow (MGR <0.88) or fast growth phenotype (MGR>1.12) while the orange points were mutant with normal growth phenotype. GASP population at Day 3 are *rpoS, cpxA* and *slyA* mutant (cont.)

Figure 3.9 (A) LTSP experiment (R2)

The number of GASP mutant population at Day 7 and Day 8 increased compared to Day 3. These mutant were observed for Day 7 and Day 8; *cpxA, slyA, kdsC, kdsD, yfaU, hsrA* mutant. Up to Day 7; *tdcA, rpoS, mraZ* mutant and Day 8; *mprA, ygaH, ygaZ ,endA* mutant. (cont.)

Figure 3.9 (A) LTSP experiment R2

The number of GASP mutant population at Day 14 and Day 21 increased compared to Day 3. These mutant were observed for Day 14 and Day 21; *cpxA*, *slyA*, *kdsC*, *kdsD*, *yfaU*, *ygaZ*, *endA*, *mprA*, *ygaH*, *yehA*, and *cpdA* mutant. Day 21only; *fimE* and *yrbG* mutant.

Figure 3.9 Population abundance of each mutant $\log_{10}$(barcode count) during LTSP (R1) against growth phenotype (MGR).

(B) R1: Day 0.5 showed the distribution of the initial abundance for each strain (also see histogram). Blue points were mutants with slow (MGR < 0.88) or fast growth phenotype (MGR>1.12) while the orange points were mutant with normal growth phenotype. GASP population at Day 3 are *rssA*, *cpxA*, *rbsR*, *rimK* and *yfaU* mutant (cont.)

Figure 3.9 (B) LTSP experiment R1

The number of GASP mutant population at Day 7 and Day 8 increased compared to Day 3. These mutant were observed for Day 7 and Day 8; *mraZ, metN, cspE, rimK, clpA, rssA, slyA, yehA, yfaU, nuoM, iscR, ygaZ, ygaH, endA, yggJ, cpdA, tdcA, kdsd,kdsC, rbsR,hsrA and cpxA* mutant. Day 8 only; *metI* and *pldA* mutant.

Table 3.6 List of GASP mutant population during LTSP and serial passage

LTSP (log10[bc]>3)	R1	R2
Day 3	*rimK, rssA, yfaU, rbsR, cpxA*	*slyA, rpoS, cpxA*
Day 7	*mraZ, metN, cspE, rimK, clpA, rssA, slyA, yehA, yfaU, nuoM, iscR, ygaZ, ygaH, endA, yggJ, cpdA, tdcA, kdsD, kdsC, rbsR, hsrA, cpxA*	*mraZ, slyA, yfaU, rpoS, tdcA, kdsD, kdsC, hsrA, cpxA*
Day 8	*mraZ, metI, metN, cspE, rimK, clpA, rssA, slyA, yehA, yfaU, nuoM, iscR, ygaZ, ygaH, endA, yggJ, cpdA, tdcA, kdsD, kdsC, cysE, rbsR, hsrA, pldA, cpxA*	*slyA, yfaU, ygaZ, ygaH, mprA, endA, kdsD, kdsC, cpxA,hsrA*
Day 14	*NA*	*slyA, yehA, yfaU, ygaZ, ygaH, mprA, endA, cpdA, kdsD, kdsC, cpxA*
Day 21	*NA*	*slyA, yehA, yfaU, ygaZ, ygaH, mprA, endA, cpdA, yrbG, kdsD, kdsC, cpxA, fimE*

SP (log10[bc]>2)	R1	R2
Day 3*	*cheR, cpxA, cspE, gadE, kdsC, kdsD, nfsA, nlpD, rimK, rpoS, rssA, slyA, yehA, yggJ*	*cpxA, kdsC, nfsA, nlpD, prpB, rimK, slyA, yehA*
Day 7	*clpA, cpxA, hupA, kdsC, mgsA, mraZ, proY, rimK, rssA, sanA, tdcA, yciQ, yciS, yeaA, ygfQ, yggJ*	*cpxA, cspE, fis, gadE, gadX, hupA, kdsC, nfsA, nlpD, paaX, prpB, rimK, slyA, yehA, yrbG*
Day 8	*clpA, mraZ, rssA, sanA, tdcA, yeaA*	*cspE, hupA, kdsC, mraZ, nfsA, paaX, rimK, ybiU*
Day 14	*NA*	*NA***

NA; No data

NA**; log_{10}[barcode_count]< 2

Day 3*; the threshold used for this is log_{10}[barcode_count]>3

3.2.8 GASP mutant population

From these results, I had identified 31 mutant that showed growth advantage during LTSP shown in Table 3.7. Out of the 31 GASP mutant, 11 of these mutant have mutation involved in transcription regulation (*rpoS, mraZ, cspC, cspE, slyA, iscR, tdcA, mprA, rbsR, fimE,* and *sspA*), 10 involved in membrane transport, biosynthesis and stress (*ygaH, ygaZ, metI, metN, yrbG, hsrA, kdsC, kdsD, pldA* and *cpxA*) and others such as ATP- dependent serine protease (*clpA*), DNA-specific endonuclease (*endA*), putative fimbrial-like adhesin protein (yehA), phosphodiesterase (*cpdA* and *yahA*), phospholipase (*rssA*), ribosomal protein subunit (*rimK* and *yggJ*) and 2-keto-3-deoxy-L-rhamnonate aldolase (*yfaU/rhmA*). Two mutants; *rpoS*[48,49] and *sspA*[118] has been reported as GASP were also identified in this study. This showed that the results from this study corroborates with previous studies by other researches and suggest high possibility our newly identified mutant candidates to exhibit GASP phenotype.

Table 3.7 List of GASP mutant candidates during LTSP

Gene	Gene function
clpA	ATPase and specificity subunit of ClpA-ClpP ATP-dependent serine protease, chaperone activity(clpA)
cpdA	3',5' cAMP phosphodiesterase(cpdA)
cpxA	sensory histidine kinase in two-component regulatory system with CpxR(cpxA)
cspC	stress protein, member of cspA family
cspE	constitutive cold shock family transcription antitermination protein; negative regulator of cspA transcription; RNA melting protein; ssDNA-binding protein(cspE)
endA	DNA-specific endonuclease I(endA)
fimE	tyrosine recombinase/inversion of on/off regulator of fimA(fimE)
hsrA	putative multidrug or homocysteine efflux system(hsrA)
iscR	isc operon transcriptional repressor; suf operon transcriptional activator; oxidative stress- and iron starvation-inducible; autorepressor(iscR)
kdsC	3-deoxy-D-manno-octulosonate 8-phosphate phosphatase(kdsC)
kdsD	D-arabinose 5-phosphate isomerase(kdsD)
metI	DL-methionine transporter subunit(metI)
metN	DL-methionine transporter subunit(metN)
mprA	transcriptional repressor of microcin B17 synthesis and multidrug efflux(mprA)
mraZ	DNA-binding transcriptional repressor mraZ
nuoM	NADH:ubiquinone oxidoreductase, membrane subunit M(nuoM)
pldA	outer membrane phospholipase A(pldA)
rbsR	transcriptional repressor of ribose metabolism(rbsR)
rimK	ribosomal protein S6 modification protein(rimK)
rpoS	RNA polymerase, sigma S (sigma 38) factor(rpoS)
rssA	putative patatin-like family phospholipase(rssA)
slyA	global transcriptional regulator(slyA)
sspA	stringent starvation protein A
tdcA	tdc operon transcriptional activator(tdcA)
yahA	c-di-GMP-specific phosphodiesterase(yahA)
yehA	putative fimbrial-like adhesin protein(yehA)
yfaU/ rhmA	2-keto-3-deoxy-L-rhamnonate aldolase(rhmA)
ygaH	putative L-valine exporter, norvaline resistance protein(ygaH)
ygaZ	putative L-valine exporter, norvaline resistance protein(ygaZ)
yggJ	ribosomal RNA small subunit methyltransferase E
yrbG	putative calcium/sodium:proton antiporter(yrbG)

Source: Ecocyc database

3.2.9 Verification of GASP mutant

I performed several experiment to verify if these mutation does exhibit growth advantage compared to WT. First, I compared the growth profile of *cpxA, cspC, rpoS, rssA, rbsR* and *sspA* mutant with WT strain (*lacA* mutant) grown in LB medium. The results were as shown in Figure 3.10. Strains with mutation in *cpxA, cspC, rpoS* and *rbsR* clearly showed better growth compare to WT while growth of *sspA* mutant initially increased but start to decreased after 15 hours. On the other hand, growth of *rssA* mutant is similar with WT. Meanwhile, complementation assay showed that expression of *cspC, rpoS, rssA* and *rbsR* did not inhibit cell growth. This suggest that expression of these genes is not toxic to cells. However, expression of *cpxA* clearly inhibit cell growth. This result is in agreement with report by Delhaye and co-workers[119]. Using a *cpxA* mutant (constitutively activate Cpx response), they found that this mutant exhibit abnormal growth and morphology. The mutant displayed significant growth defects: mass doubling times were at least twice as long as those of wild-type cells and filamentous morphology mixed with DNA free minicells which suggested defect in cell division. The author had conclude that the Cpx system play a role in maintaining the peptidoglycan homeostasis and mutation in *cpx* genes affect the bacterial cell wall structure.

To check the fitness of each mutant, I carried competition experiment using *rssA, cspC, sspA, tdcA, hsrA, endA, ycfJ* and *ygaH* mutant with WT(Δ*lacA*). The resistant cassette for WT were switch from Cm to KAN. The change of resistance gene did not affect the mutant fitness (Supplementary Table 3). From Figure 3.11 all mutants shown to have better fitness compared to WT. The *rssA* mutant showed highest fitness, followed by *ygaH, endA, tdcA, hsrA, cspC, sspA*, and *ycfJ*. I had performed verification only on a subset of GASP mutant candidates, however so far the results indicates that these mutations confer better fitness compared to WT. This indicates that mutation in these genes is important as survival adaptation during LTSP.
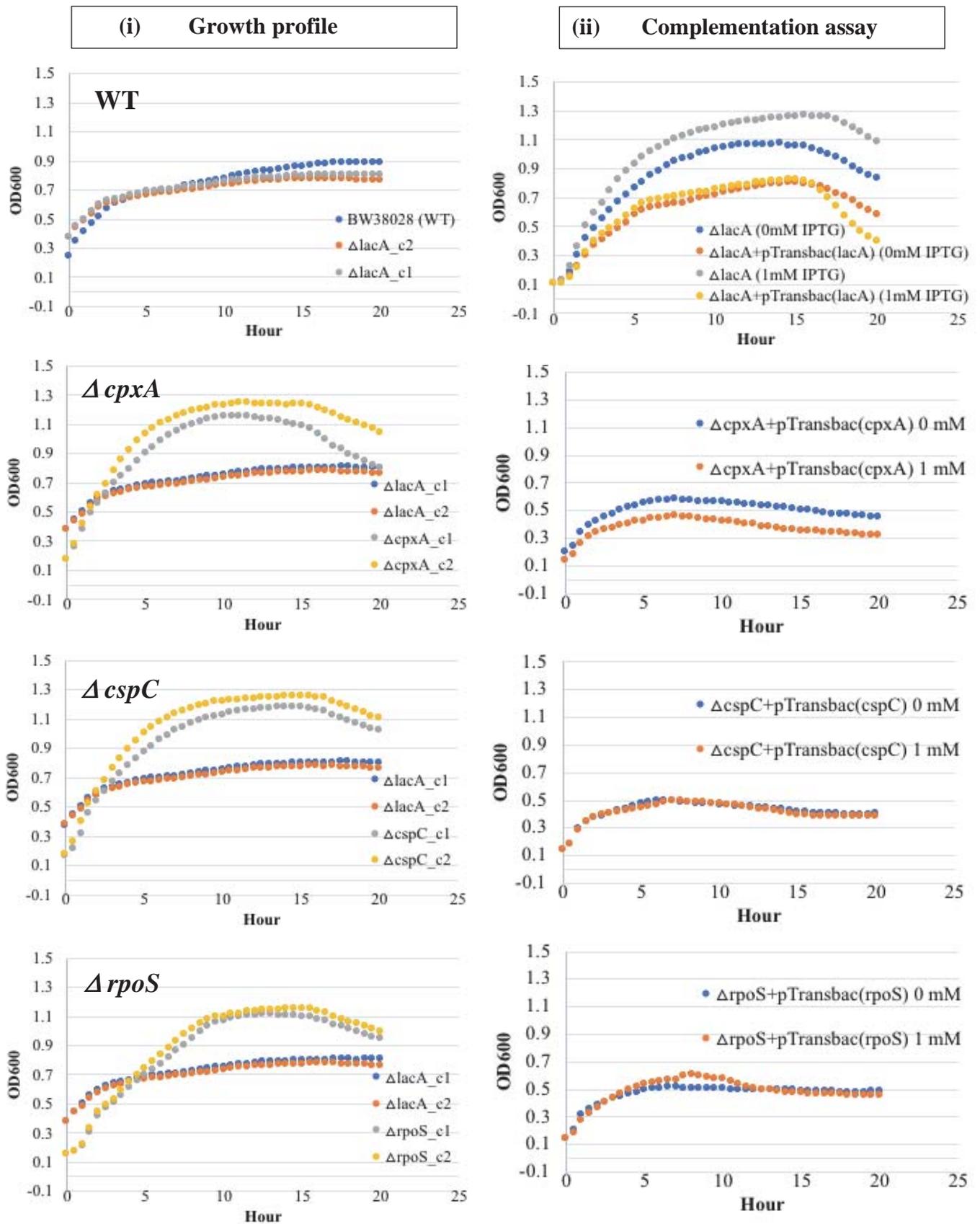
Figure 3.10 Growth profile of GASP mutant candidate.

Comparison of growth profile in (i) LB medium and (ii) complementation assay of control strain (WT), *cpxA, cspC* and *rpoS* mutant (cont.)

(ii) Growth profile	(ii) Complementation assay
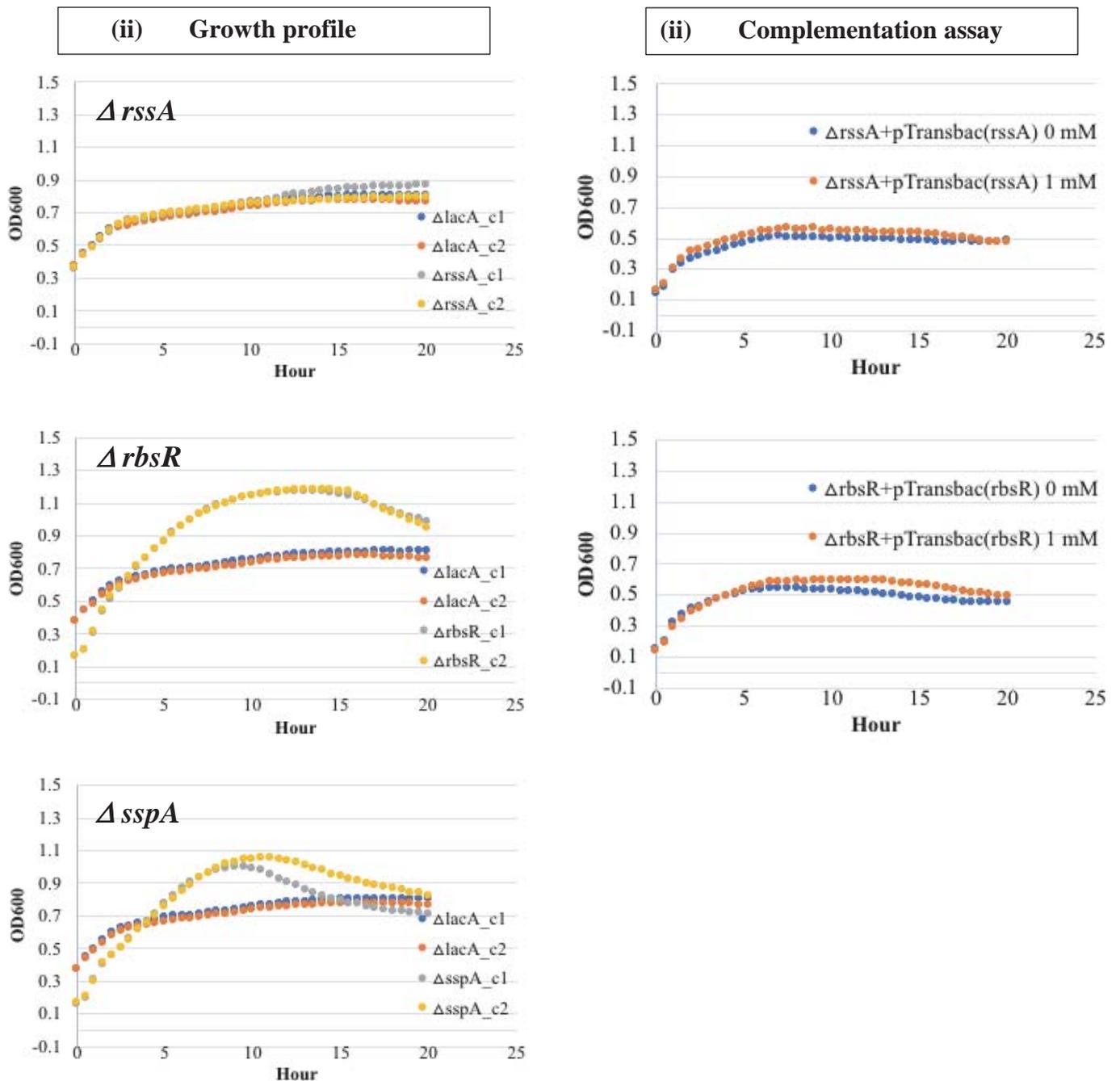
Figure 3.10 (cont.) Comparison of growth profile in (i) LB medium and (ii) complementation assay. The *cpxA, cspC, rpoS* and *rbsR* mutant exhibit better growth compare to WT while growth of *sspA* mutant initially increased but start to decreased after 15 hours became similar to WT. On the other hand, growth of *rssA* mutant is similar with WT. Meanwhile, complementation assay showed that except for *cpxA*, expression of *cspC, rpoS, rssA* and *rbsR* did not inhibit cell growth. TransBac plasmid is for *sspA* is not available. Each values were average of six replicates.
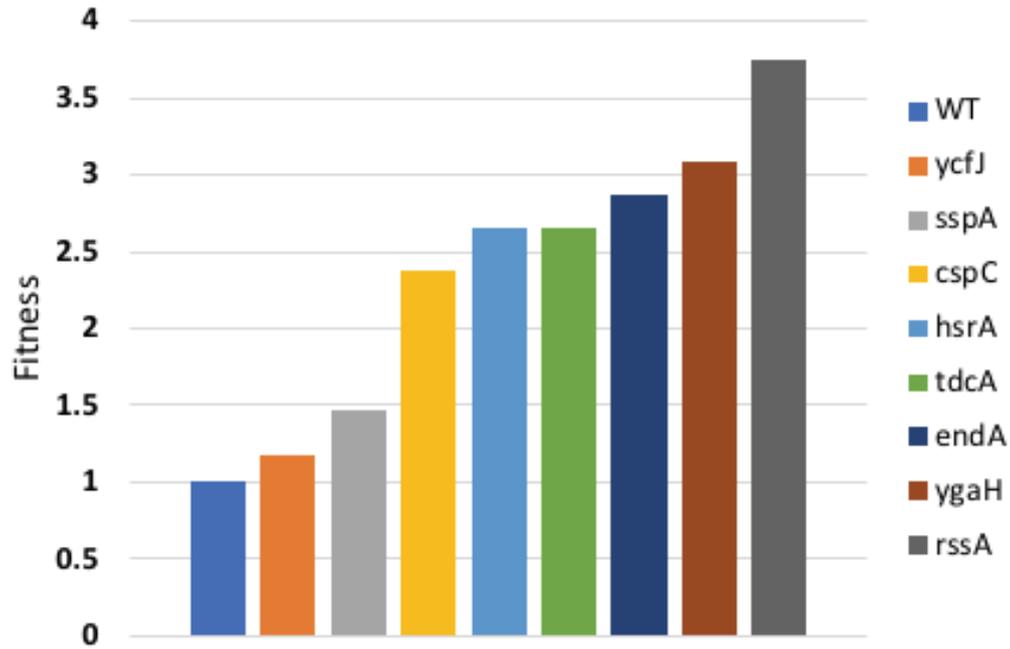
Figure 3.11 Fitness of GASP mutant candidates.

All mutants exhibit higher fitness compared to WT. The fitness is calculated using formula (Equation 2.1); WT fitness is 1 (blue bar). Strains with fitness <1 showed low fitness while >1 showed higher fitness compared to WT. The *rssA* mutant showed highest fitness, followed by *ygaH, endA, tdcA, hsrA, cspC, sspA*, and *ycfJ*. Three independent replicate were performed for each mutant.

3.3 Genome-wide screening of drug tolerant mutant population by Bar-seq

3.3.1 Minimum inhibitory concentration (MIC)

  E. coli strain BW38028 is the parental strain for this deletion library. Since this strain could not be incorporated to the mix library, I had selected *lacA* mutant as control strain as the effect of mutation is neutral. Figure 3.1 showed the growth profile of BW38028 (parental strain) and *lacA* mutant (clone 1 and clone 2). The BW38028 and control strain (WT) showed similar growth profile and the specific growth rates (h-1) were 0.1086, 0.1123 and 0.1009, respectively (Supplementary Figure 2).

  Table 3.8 showed the selected drugs used in this study. I had chosen 7 drugs from 4 drug class; aminoglycosides (GEN, KAN and STR), quinolones (NAL and CIP), tetracycline (TET) and trimethoprim (TMP). The sub-inhibitory concentration were determine from the dose response curve shown in Figure 3.12. This deletion library have resistance to chloramphenicol. Thus, WT ($\Delta lacA$ mutant) with and without Cm resistant were tested for cross-resistance to the selected drugs (Table 3.9). The inhibitory concentrations of each drug for both (+Cm and -Cm) strain showed no significant difference ($p > 0.05$). This indicate that there is no cross-resistance between Cm resistance and target drug.

Table 3.8 List of drug used in this study.

Class	Drug	Activity	Mechanism	Target
Quinolone	Ciprofloxacin (CIP)	Bactericidal	Inhibit DNA	*gyrA, parC, parE, mdtK*
	Nalidixic acid (NAL)	Bacteriostatic		*gyrA*
Trimethoprim	Trimethoprim (TMP)	Bacteriostatic		*folA*
Aminoglycoside	Gentamicin (GEN)	Bactericidal	Inhibit protein synthesis	*rpsL*, 16S rRNA
	Kanamycin (KAN)			
	Streptomycin (STR)			
Tetracycline	Tetracycline (TET)	Bacteriostatic		*rpsG, rpsN, rpsC, rpsH, rpsS,* 16S rRNA

Source: Drugbank database

Quinolone drugs (CIP and NAL) inhibit this enzyme (DNA gyrase /topoisomerase II type) by binding to the A-subunit of the enzyme due to which the bacteria is unable to replicate. CIP is bactericidal drug as it binds irreversibly to the target while NAL is bacteriostatic (at low concentration) as it binds reversibly to the target. TMP inhibits by binding to dihydrofolate reductase, a key enzyme that convert the dihydrofolic acid (DHF) to tetrahydrofolic acid (THF). THF (active form of folic acid) is an essential precursor in the thymidine synthesis pathway and disruption with this pathway indirectly inhibits bacterial DNA synthesis. TMP is bacteriostatic. The aminoglycoside drugs (GEN, KAN and STR) is bactericidal as it inhibits the protein synthesis by binding irreversibly to 30S ribosomal subunit, at the Aminoacyl-tRNA (aa-tRNA) acceptor site (A) on the 16S ribosomal RNA (rRNA), affecting protein synthesis by induction of codon misreading and inhibition of translocation. TET is bacteriostatic, time-dependent drugs that also inhibits protein synthesis by reversibly binding to the 30S ribosomal subunit.
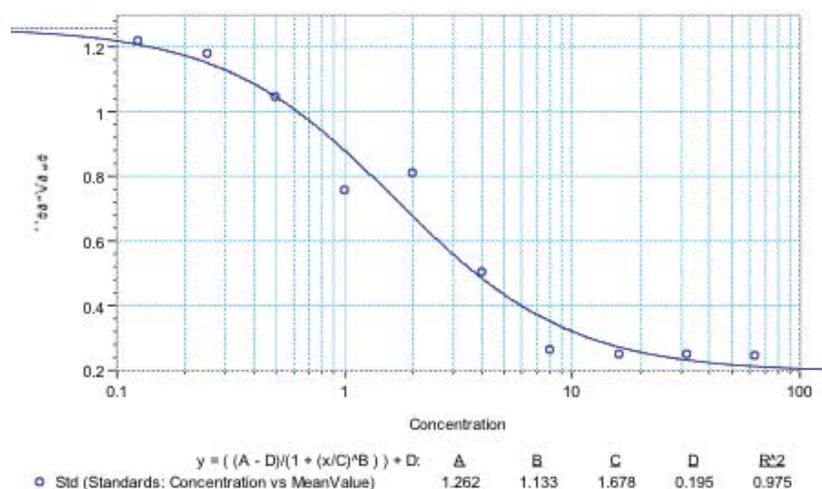
Figure 3.12 Dose-response curve.

The x-axis represents the drug concentration (log10[µg/ml]) and the y-axis represents the mean OD value. From the nonlinear regression equation; $y=((A-D)/(1+(x/C)^B))+D$, A and D represents the maximum and minimum value, B represent the Hill slope or steepness of the curve and C is the inflection point or IC50 value. The values of B and C were incorporate into Eq. 1 to calculate the drug concentrations at IC70 and IC90. WT ($\Delta lacA$::Cm strain) was used to determine the MIC concentration of each drugs used in this study. Values are mean of three replicates. Each experiment were repeated at least twice using WT strain with/ without chloramphenicol resistance (Cm^R). No significant difference ($p>0.05$) were observed between strains with/without Cm^R. (Table 3.9, Supplementary Figure 1)

Table 3.9 Drug concentration used in this study.

Drug	+Cm	-Cm	*p*-value	IC70	IC90
	IC50				
CIP	0.10	0.13	0.16	0.26	1.20
KAN	6.32	6.61	0.23	8.20	13.40
GEN	1.66	1.52	0.65	3.50	11.70
NAL	17.02	22.01	0.44	24.00	53.00
STR	3.32	4.54	0.10	7.10	25.40
TET	1.96	1.69	0.40	5.20	11.70
TMP	0.46	0.47	0.95	0.78	1.210

Drug concentration [µg/ml]; (IC50) of *lacA* mutant +/- chloramphenicol resistance. Check for cross-resistance ($p > 0.05$), chloramphenicol resistance did not cause cross-resistance to other drugs.

3.3.2 Sequencing result

The sequenced data were processed as mentioned in Section 2.10 in Materials and methods. Details of the sequencing results were shown in Figure 3.13 and Table 3.10. From LTSP sequencing result, the technical replicates of both experiments (R1 and R2) showed high reproducibility from Illumina sequencing platform. Therefore, I did not perform technical replicates for all samples in order to maximize the number of experimental samples. Technical replicates were performed for three samples (KAN: 6 hours and 4x IC90; TET: 6 hours) and the Pearson's correlation coefficient (r-value) were 0.94, 0.98 and 0.92, respectively (Table 3.3). In addition, two independent experiment were performed for control samples (no drug treatment) which showed high correlation between these two experiment (Table 3.4). Furthermore, the sequencing for the technical replicates were perform on separate lane to further confirm the reliability of these dataset.

In this study, I had multiplex 90 samples, ~ 238 million raw sequences reads has been obtained. I had obtained ~ 219 million valid reads (90% of total sequences reads) with coverage of ~ 330 reads per strain. NAL (11 samples) and three samples as technical replicates were sequence separately (different lane) obtained ~27 million valid reads and added to the dataset. Pearson's correlation between read 1 and read 2 for each samples were used to measure the reproducibility of the sequencing data. Samples with low correlation ($r < 0.8$) were discarded (5 samples; CIP: 1.5xIC90, STR: 3h, NAL: 24h, 1.5xIC90 and TET: 12h). Average value between read 1 and read 2 were used as new dataset. Next, I removed strains with inadequate barcode count (strain < mean barcode count[t=0]) at t=0 from the dataset. Next, correlations between clone 1 and clone 2 were determine and those with $r \geq 0.5$ were accepted for further analysis. Strains without biological replicate (clone 1 or clone 2) and $r < 0.5$ were removed from dataset. Finally, the number of valid strains for each dataset; LB (1108), GEN (1532), KAN (1791), STR (2294), TET (669), CIP (2478), NAL (1262) and TMP (1274), respectively.
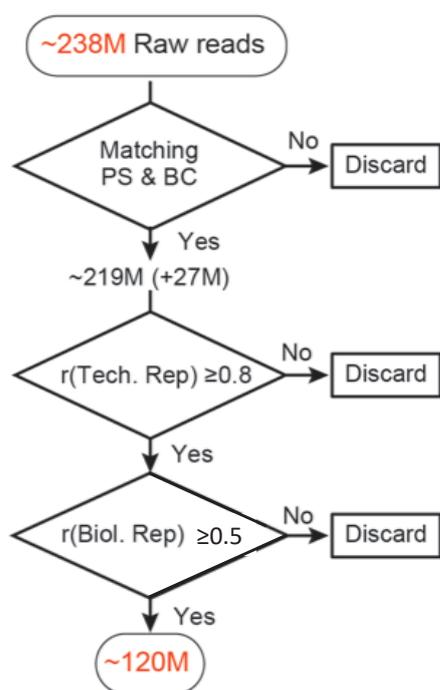
Figure 3.13 Outline of sequencing result for drug treatment

For drug experiment; multiplex 90 samples, ~ 238 million raw sequences reads has been obtained. Next, the raw sequences were map to reference, sort and count ~ 219 million valid reads (90% of total sequences reads) were matched. NAL (11 samples) and three samples as technical replicates were sequence separately (different lane) obtained ~27 million valid reads and added to the dataset. Read 1 and Read 2 were used as technical replicate and samples with $r < 0.8$ were discarded (5 samples; CIP: 1.5xIC90, STR: 3h, NAL: 24h, 1.5xIC90 and TET: 12h). Next, correlations between clone 1 and clone 2 were determine and those with $r \geq 0.5$ were accepted for further analysis. Strains without biological replicate (clone 1 or clone 2) and $r < 0.5$ were removed from dataset

Table 3.10 Detailed sequencing result for drug treatment

Experiment	Description	Result
	Multiplex	90 samples
	Total raw sequences	238 451 233
	Valid reads (Mapped to reference sequences)	218 475 869 (91%) + 27012169 (separate lane, 14 samples)
Drug treatment	Number of valid strains ($r \geq 0.5$) LB GEN KAN STR TET CIP NAL TMP	1108 1532 1791 2294 669 2478 1262 1274

3.3.3 Growth profile and cell morphology of Escherichia coli mutant during drug treatment

The ASKA collection were grown in LB medium without drug until reached OD $\cong$ 0.4 (2.5 - 3 hours), then selected drug were added at sub-inhibitory concentration (IC50) and incubated at 37 °C, 200 rpm with sampling at each time points (t = 0, 3, 6, 9, 12, 18 and 24 hour). After 24 hours, the cells were collected and washed with LB, then regrown (initial cell concentration at OD $\cong$ 0.4) in new LB medium without drug and with increasing concentration of drug (IC70, IC90, 1.5xIC90, 2xIC90 and 4xIC90) at 37 °C, 200 rpm for another 24 hours. During each sampling, I had monitored the viability of *E. coli* cells by plating on LB agar without drug (colony forming unit per ml, CFU/ml) and optical density of liquid culture at 600nm. Two independent experiments were performed for each of the drugs.

The initial cell concentration was in the range of 10^7 to 10^8 before addition of sublethal concentration (IC50) of drugs (Figure 3.13). ASKA deletion library grown without drug treatment were used as control. For control experiment, both the optical density and CFU increased, the OD increased from 0.42 to 6.72 while the CFU increase from 1.1×10^8 to 1.11×10^9 (highest 3.07×10^9, 12h) after 24 hours. For subsequent regrowth (serial passage), the OD and CFU were measured at 5.69 and 1.14×10^9, respectively.

Similar growth profile were observed with cultures treated with GEN, KAN and STR at sublethal concentration. The growth profile by OD and CFU for each of these drugs were; 0.42 to 4.04 and 8.3×10^8 to 5.6×10^{10} for GEN, 0.39 to 4.65 and 1×10^7 to 1.12×10^9 for KAN and 0.47 to 4.97 and 1.1×10^8 to 2.2×10^9 for STR, respectively. This indicate that *E. coli* mutant populations were viable and grew when exposed to sublethal concentration of bactericidal drugs. However, exposure higher concentration of these drugs during serial passage experiments showed contradictory results between OD and CFU. The OD decreased but not much changes were found in CFU as shown in Figure 3.13 (B2, C2, D2). These indicate that these cells may have tolerance towards these drugs and growth was recovered after drug removal.

The OD increased from 0.45 to 3.98 while the CFU maintained 2.66 - 6.16 $\times 10^7$ when *E. coli* mutants were exposed to sublethal concentration of TET shown in Figure 3.13(E1). No changes of CFU might be due to the bacteriostatic effect of the drug which inhibits the cell growth.

On the other hand, the OD increased but the CFU decreased when *E. coli* mutant were exposed to sublethal concentration of CIP, NAL and TMP as shown in Figure 3.13 (F1, G1 and H1). For CIP, the OD increased from 0.39 to 1.8, while CFU decreased from 5.8×10^7 to

1.6×10^4 (F1). Subsequently, cells exposed to higher concentration of CIP were shown as not viable but OD value were maintained at $\cong 0.28$ (F2). Similarly, the OD for NAL increased from 0.35 to 3.2 while the CFU decreased from 3×10^7 to 9×10^5 at sublethal concentration (G1). For TMP, the OD increased from 0.43 to 7.49 while the CFU increased from 3.05×10^7 to 1.55×10^9 at 6h then decreased to 1×10^8 (H1).

The cell morphology before, during and after exposure to drugs were observed to be different compare with control (no drug treatment) as shown in Figure 3.14. The untreated bacterial cells (LB) in Figure 3.14 (A) showed smooth rod-shaped morphology, while the cells were exposed to GEN and KAN shown in Figure 3.14 (B & C), showed slight elongation with dense body formation after 3h. Similar morphology were observed at 24 hours. After regrowth in new LB without drug, the cells size were smaller compared to control. For cells exposed to STR shown in Figure 3.14 (D), cell size were similar to control after 3h but showed mixture of normal and reduced cell size after 24h. After regrowth in LB, cells size were smaller than control. Meanwhile, the *E. coli* cells exposed to sublethal concentration TET in Figure 3.14 (E) showed slight elongation after 3h and 24 hours. However after regrowth in LB, cells size were a mixed of small and elongated cells. The bacteria cells exposed to sublethal concentration TMP (Figure 3.14 (F)), showed morphology that is similar to the cells exposed to TET at 3h. After 24 hours, the cells size is similar to control but shape is slightly irregular.

On the other hand, *E. coli* cells exposed to sublethal concentration of CIP and NAL in Figure 3.14 (G & H, respectively) showed obvious elongated cells (filamentous) compared with control and the cells were shown to recover to normal size (CIP) or decreased in size (NAL) when these drugs were removed.

The OD or optical density represents the amount of light that is absorbed by the sample and is influenced by the size and shape of the particles (or cells). Thus, contradicting OD and CFU results for CIP, NAL and TMP were caused by the cells sizes. This suggested that OD might not be a good measurement for growth in drug studies. The cell morphology exposed to CIP, NAL and TMP were distinctive from those of KAN, GEN, STR and TET which relates the mechanism of each drugs. Interestingly, although cells morphology during drug treatment were shown to be distinct towards drug mechanism, the morphology of cells after recovery (SP_LB) showed some similarities; (i) reduced cell size and (ii) dense body formation, to cells during LTSP. This suggest that cell damages occurred during drug treatment were similar to cells during LTSP.
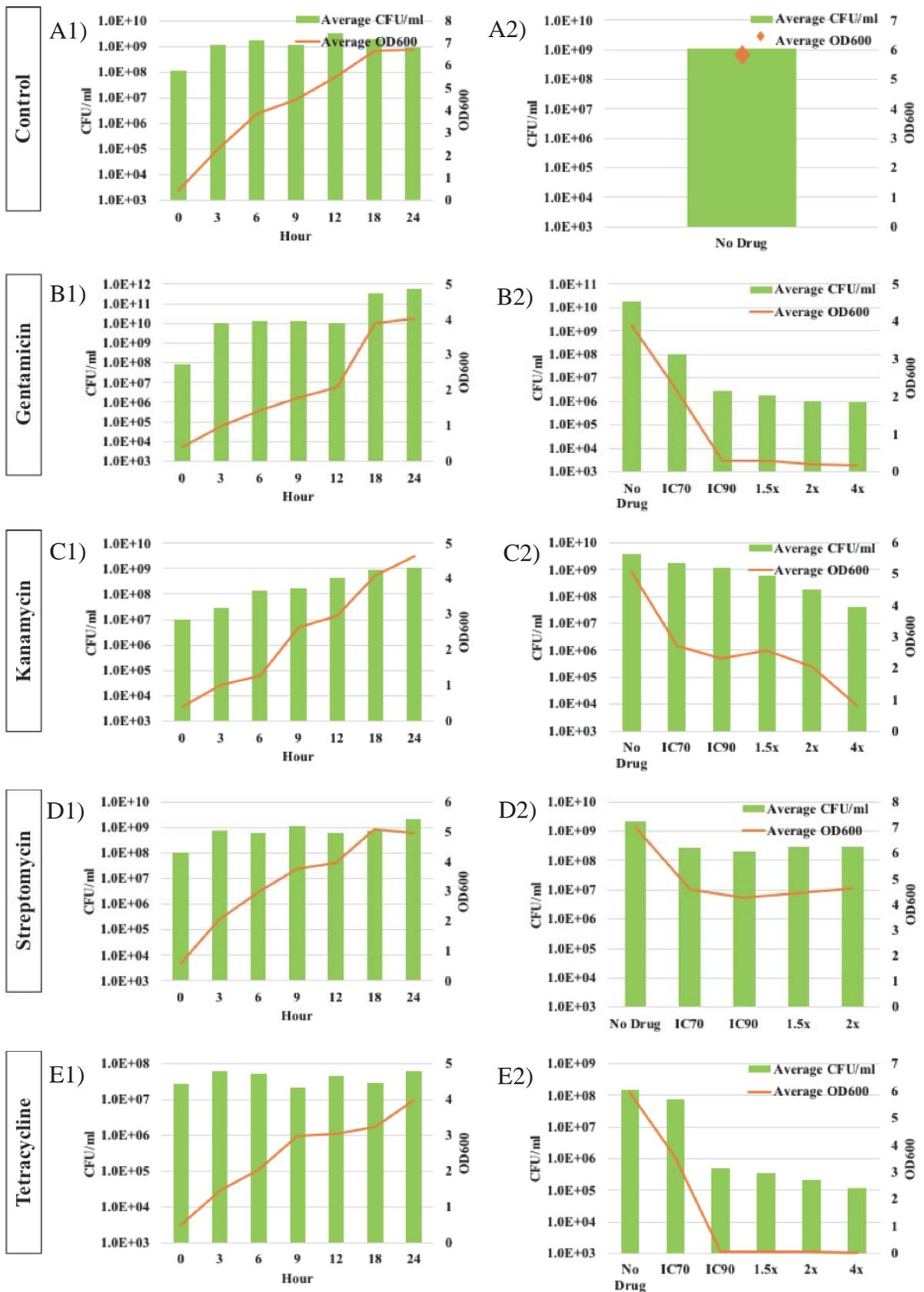
Figure 3.14 Growth profile of E. coli mutant during drug treatment.
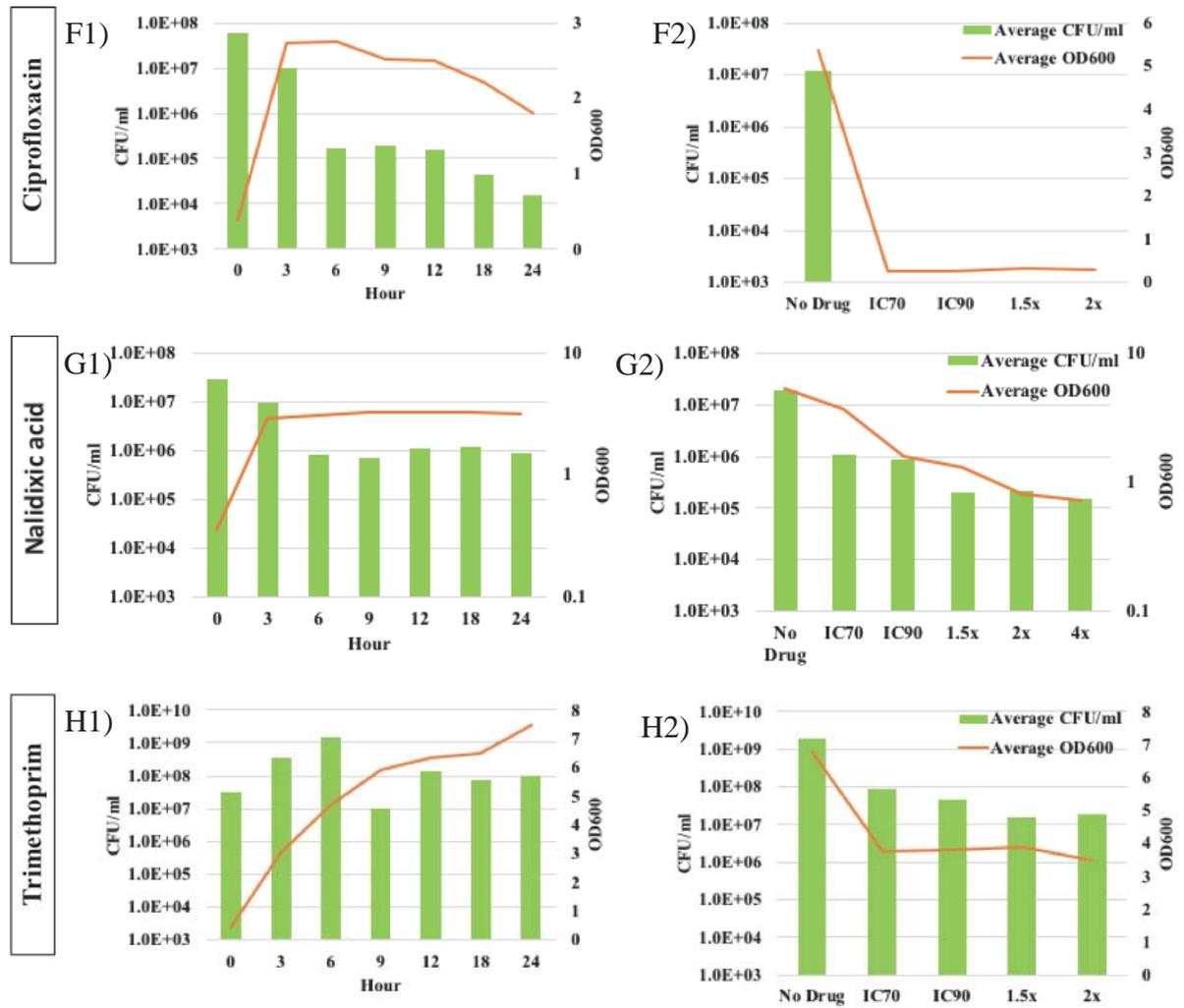
Figure 3.14 Growth profile of *E. coli* mutant during drug treatment (cont)

The profile on the left was when exposed to sub-lethal drug concentration (IC50) and (right) serial passage experiment. For serial passage experiment, cells after 24 hours exposure to drugs at IC50 were collected and regrown (initial cell concentration: OD≅0.4) in new LB medium without drugs and with increased drug concentration at 37 °C, 200 rpm for 24 hours.
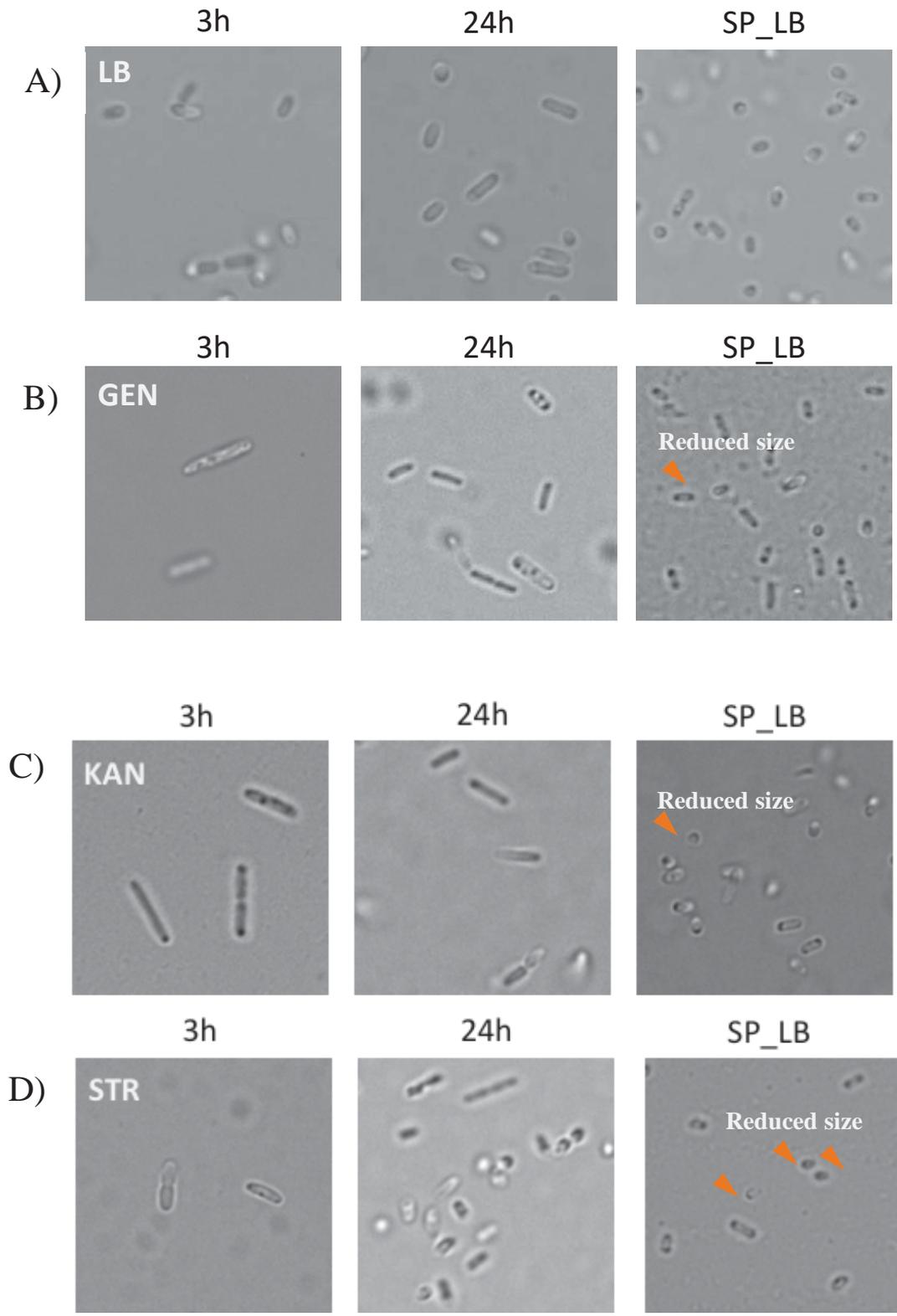
Figure 3.15 Cell morphology during and after exposure to sub-lethal concentration (IC50) of drug
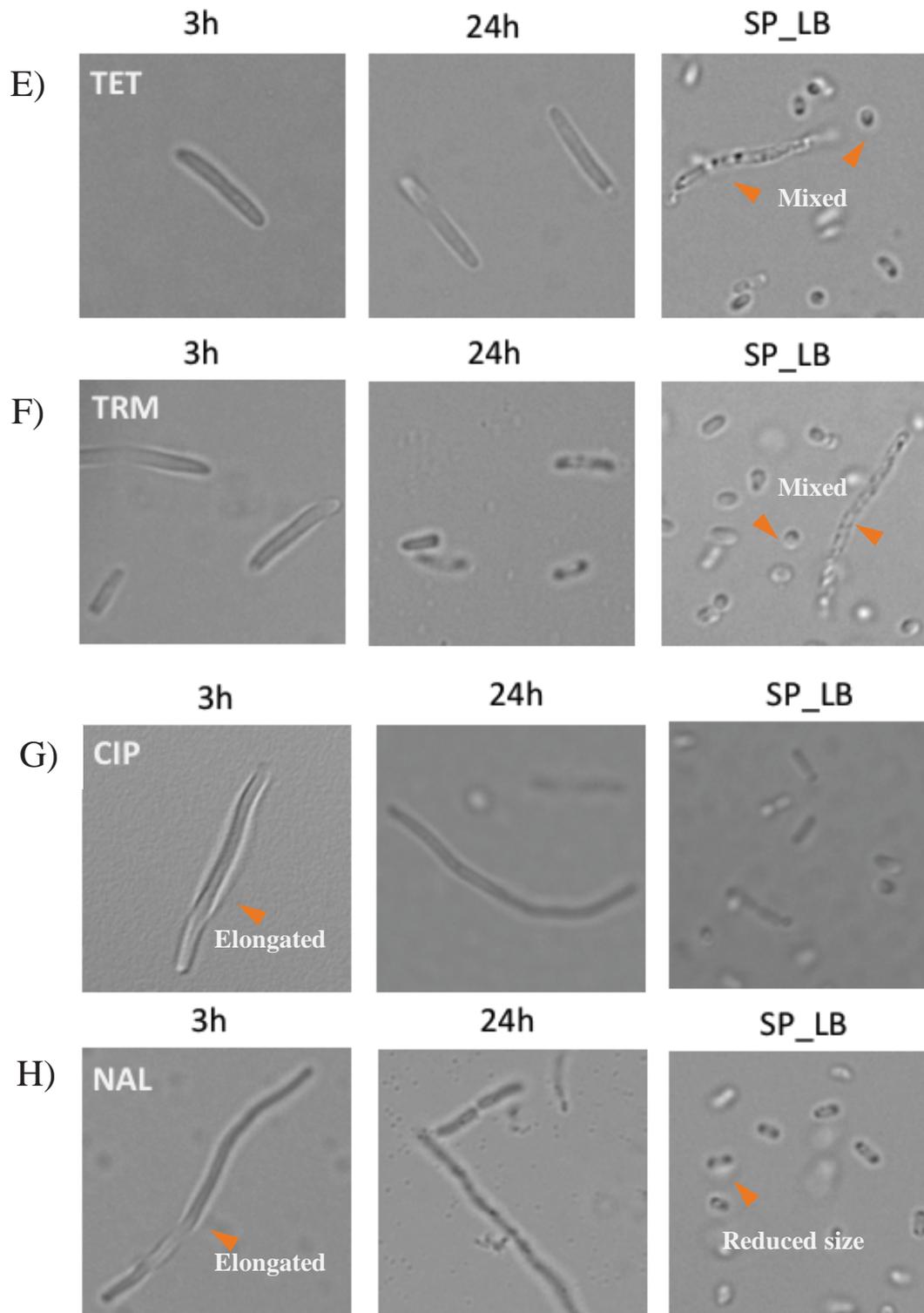
Figure 3.15 Cell morphology during and after exposure to sub-lethal concentration (IC50) of drug. (A) LB (no drug treatment), (B) GEN, (C) KAN and (D) STR, (E) TET, (F) TMP, (G) CIP and (H) NAL. Cells morphology after exposure to drug at IC50 (3 hours and 24 hours) and after drug removal (SP_LB). SP; after 24 hours exposure to drug, cells were regrown in new LB medium without drug for 24 hours.

3.3.4 Population dynamics of E. coli mutant during drug treatment

The changes in each mutant population were determined by calculating the read frequency of each strain against total reads at each time point. Figure 3.16 showed the dynamics of each population during drug treatment. Immediate changes in the *E. coli* mutant population were observed during GEN, KAN and CIP treatment at IC50 whereas no changes were observed in mutant population treated with STR, TET, NAL and TMP. NAL (inhibit DNA gyrase), TRM (inhibit folate synthesis) and TET (inhibit protein synthesis) are bacteriostatic drugs which only inhibits the population growth thus, we are unable to observed clear population alteration at IC50 for 24 hours. Meanwhile, GEN and KAN (inhibit protein synthesis) and CIP (inhibit DNA gyrase) are bactericidal drugs showed clear uneven population distributions after 6 hours incubation at IC50 drug concentration. This result showed that the drug killing activity affect the bacteria populations dynamics during drug treatment.

Meanwhile, exposure to higher concentration of STR, NAL and TRM during serial passage (Figure 3.16, right panel) lead small number of deletion mutants occupancy. This suggest that these mutant might be resistant to these drug. However, further validation showed low correlation between each strain biological replicate (clone 1 and clone 2). This suggest secondary mutations might be responsible for the increase resistance against these drugs.

Aminoglycosides
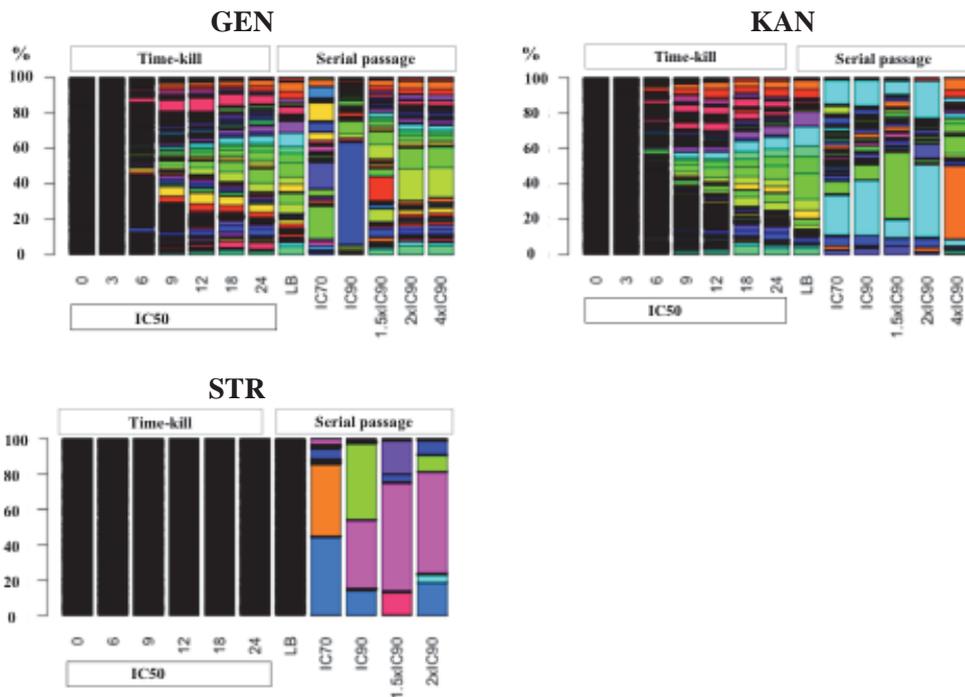
Fluoroquinolone/quinolone
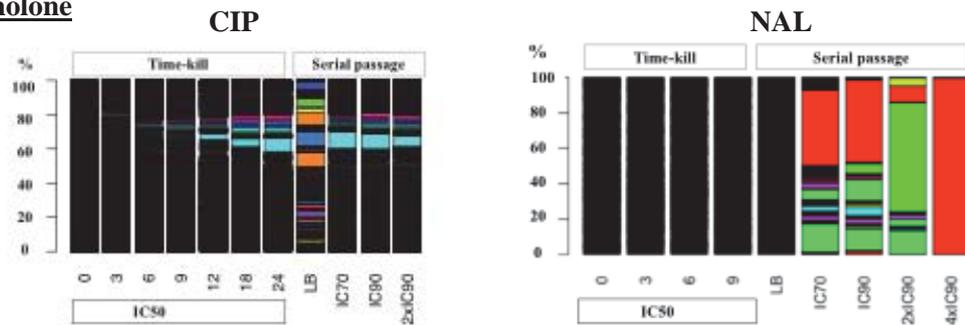
Tetracycline
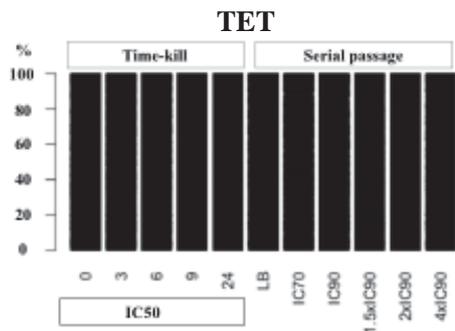
Trimethoprim
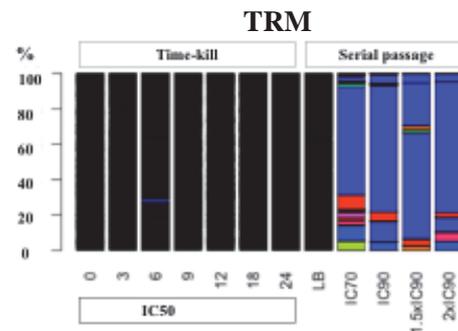
Figure 3.16 Population dynamics of E. coli mutant populations during drug treatment.

Immediate change in population dynamics were observed at sublethal concentration (IC50) of bactericidal drugs (GEN, KAN, CIP) while most can survive at sublethal concentration of bacteriostatic drugs (TET, NAL, TMP)

3.3.5 Drug tolerant mutant population

To identify the appearance of tolerant mutants, the ASKA barcode deletion collection were first challenge to sublethal (IC50) concentration of drugs for 24 hours. Next, the cells were exposed to higher concentration up to 4xIC90 for another 24 hours. I analyzed the surviving mutant population at higher drug concentration as drug tolerant population. The comparison of the distribution of each mutant population at IC90 and 2xIC90 for each drugs is shown in Figure 3.17. High abundance mutant strains with $\log_{10}$(barcode count) >3 (colored points). The gene name and their functional annotation are listed in Table 3.11.

The results from Figure 3.17 and Table 3.11 indicates that mutations involved in energy metabolism (oxidative phosphorylation); ATPase (*atpABCEF*), PTS (*ptsI, crr*), cytochrome complex (*cyoABDE*), NAD(H):ubiquinone oxidoreductase complex (*nuoACEHJMN*) showed increased tolerance against multiple drugs. In addition, mutation involved in these genes showed low MGR (Table 3.12). This indicates that these mutants exhibit slow growth phenotype. Meanwhile, mutations involved in LPS biosynthesis exhibit increased tolerance against CIP.

Overall, the results indicates that drug tolerant mutants have mutations involved in energy metabolism and exhibit slow growth phenotype. Lower energy levels have been suggested to contribute to formation of persisters[37,41]. This suggest that these mutations might contribute to persister cells formation.
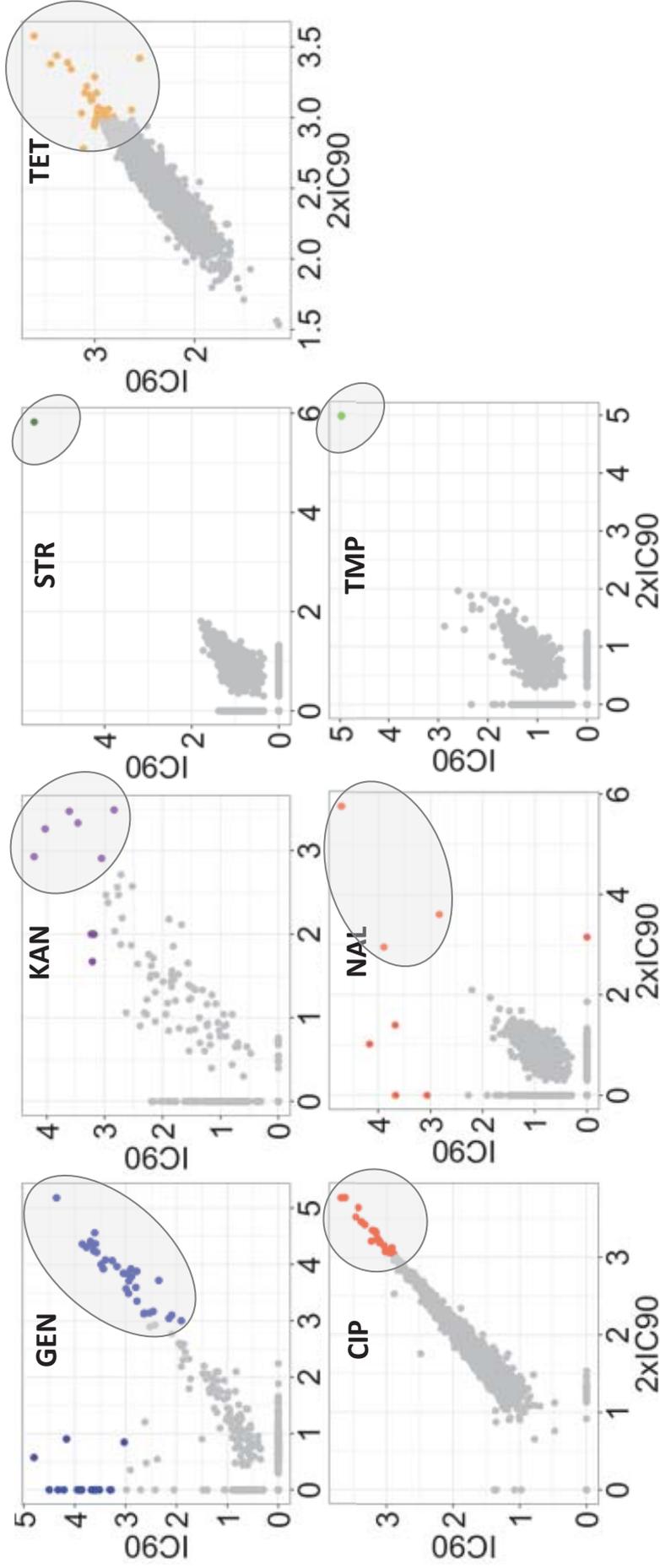
Figure 3.17 Drug tolerant mutant population for each drugs

The x and y axis represents the log$_{10}$(barcode count) of each *E. coli* mutant after exposure to drugs at IC90 and 2xIC90. Mutant strains that with log$_{10}$(barcode count) >3 (colored points) were selected as drug tolerant candidate (in grey circle).

Table 3.11 List of drug tolerant mutant candidates for each drugs

Gene	Annotation
GEN	
aceF	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex
cpxA	Sensor histidine kinase
crr	PTS system glucose-specific EIIA component
fdx	2Fe-2S ferredoxin
fre	NAD(P)H-flavin reductase
fur	Ferric uptake regulation protein
greA	Transcription elongation factor
hscA	Chaperone protein
hscB	Co-chaperone protein
iscA	Iron-binding protein
iscU	Iron-sulfur cluster assembly scaffold protein
malK	Maltose/maltodextrin import ATP-binding protein
nuoA	NADH-quinone oxidoreductase subunit A
nuoC	NADH-quinone oxidoreductase subunit C/D
nuoE	NADH-quinone oxidoreductase subunit E
nuoH	NADH-quinone oxidoreductase subunit H
nuoJ	NADH-quinone oxidoreductase subunit J
nuoM	NADH-quinone oxidoreductase subunit M
nuoN	NADH-quinone oxidoreductase subunit N
nusB	N utilization substance protein B
proC	Pyrroline-5-carboxylate reductase
ptsI	Phosphoenolpyruvate-protein phosphotransferase
rpsT	30S ribosomal protein S20
sohA	Antitoxin PrlF
ubiH	2-octaprenyl-6-methoxyphenol hydroxylase
visC	2-octaprenylphenol hydroxylase
yaeQ	Uncharacterized protein
yfjG/trxC	Thioredoxin 2
ygfZ	tRNA-modifying protein
yqiB	Uncharacterized protein
yqiC	Uncharacterized protein
KAN	
crr	PTS system glucose-specific EIIA component
feoB	Fe(2+) transporter
hscA	Chaperone protein
hscB	Co-chaperone protein

visC	2-octaprenylphenol hydroxylase
yaaA	Uncharacterized protein
ybdG	Miniconductance mechanosensitive channel
yfhD	Membrane-bound lytic murein transglycosylase F
yggL	Uncharacterized protein
STR	
pntA	NAD(P) transhydrogenase subunit alpha
TET	
atpA	ATP synthase subunit alpha
atpB	ATP synthase subunit beta
atpC	ATP synthase epsilon chain
atpE	ATP synthase subunit c
atpF	ATP synthase subunit b
baeS	Signal transduction histidine-protein kinase BaeS
clpP	ATP-dependent Clp protease proteolytic subunit
crr	PTS system glucose-specific EIIA component
cyoA	Cytochrome bo(3) ubiquinol oxidase subunit 2
cyoB	Cytochrome bo(3) ubiquinol oxidase subunit 1
cyoD	Cytochrome bo(3) ubiquinol oxidase subunit 4
cyoE	Protoheme IX farnesyltransferase
glnD	Bifunctional uridylyltransferase/uridylyl-removing enzyme
glpD	Aerobic glycerol-3-phosphate dehydrogenase
gmhB	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase
hyfC	Hydrogenase-4 component C
nuoA	NADH-quinone oxidoreductase subunit A
purA	Adenylosuccinate synthetase
rfaD	ADP-L-glycero-D-manno-heptose-6-epimerase
rfaF	ADP-heptose--LPS heptosyltransferase 2
rspB	Starvation-sensing protein RspB
sapD	Putrescine export system ATP-binding protein SapD
slyX	Protein SlyX
sucB	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex
wzxE	Lipid III flippase
yajD	Uncharacterized protein YajD
yebB	Uncharacterized protein YebB
yegR	Uncharacterized protein YegR
yfbP	Uncharacterized protein YfbP
yidE	Putative transport protein YidE
CIP	
apaH	Bis(5'-nucleosyl)-tetraphosphatase
aroB	3-dehydroquinate synthase

atpC	ATP synthase epsilon chain
crr	PTS system glucose-specific EIIA component
cysE	Serine acetyltransferase
gmhB	D-glycero-beta-D-manno-heptose-1,7-bisphosphate 7-phosphatase
nuoG	NADH-quinone oxidoreductase subunit G
nuoM	NADH-quinone oxidoreductase subunit M
pgi	Glucose-6-phosphate isomerase
pitA	Low-affinity inorganic phosphate transporter 1
pldA	Phospholipase A1
ptsI	Phosphoenolpyruvate-protein phosphotransferase
purA	Adenylosuccinate synthetase
rfaC	Lipopolysaccharide heptosyltransferase 1
rfaE	Bifunctional protein HldE
rfaF	ADP-heptose--LPS heptosyltransferase 2
rfaG	Lipopolysaccharide core biosynthesis protein RfaG
rfaH	Transcription antitermination protein RfaH
trkA	Trk system potassium uptake protein
yidG	Inner membrane protein YidG
NAL	
lysR	Transcriptional activator protein LysR
nuoM	NADH-quinone oxidoreductase subunit M
ygfU	Uric acid transporter UacT
TMP	
frlD	Fructoselysine 6-kinase

Table 3.12 List of MGR for mutants involved in energy metabolism and LPS biosynthesis

Gene	MGR	Gene	MGR	Gene	MGR
nuoN	0.73	cyoE	0.82	sucA	0.77
nuoM	0.69	cyoD	0.95	sucB	0.76
nuoL	0.86	cyoC	0.98	sucD	0.87
nuoK	0.72	cyoB	0.98	feo	0.98
nuoJ	0.73	cyoA	0.97	fre	0.77
nuoI	0.76	ptsG	0.98	fdx	0.95
nuoH	0.76	ptsH	0.99	fur	0.9
nuoG	0.88	ptsI	0.89	iscA	0.96
nuoF	0.76	ptsP	1.07	iscU	0.83
nuoE	0.78	ptsN	0.91	iscS	0.52
nuoC	0.8	ptsA	1.01	hscA	0.87
nuoB	0.74	crr	0.99	hscB	0.99
nuoA	0.81	ubiF	0.67	gmhB	0.78
atpC	0.34	ubiG	0.17	trkA	0.87
atpD	0.46	ubiX	0.72	yidG	1.05
atpG	0.4	ubiH	0.35	pldA	1.02
atpA	0.49	ubiE	0.45	rfaC	1.06
atpH	0.48	ubiC	1	rfaE	0.99
atpF	0.51	pgi	0.9	rfaF	1.25
atpE	0.4	pntA	0.9	rfaG	0.97
atpB	0.39	pntB	1.03	rfaH	1
atpI	1.01				

List of gene clusters involves in energy metabolism; NAD(H):ubiquinone oxidoreductase complex (orange), ATPase (green), cytochrome complex (grey), Fe-S cluster (yellow). The highlight color represent each function and low MGR mutant.

3.3.6 Susceptibility assay of drug tolerant candidates in liquid medium

From the killing kinetics shown in Figure 1.1, drug tolerant bacterial population are susceptible to drug at same MIC as susceptible population but needed longer time to be killed. To check if these mutation confer to drug tolerance or resistance, I performed susceptibility assay using individual isolate and expose them to varying concentration of drugs and measured the OD after 24 hours. I had tested on 42 drug tolerant mutants against bactericidal drugs (CIP, GEN, KAN, STR) shown in Figure 3.18. The OD were normalized by z-score. Manhattan distance were used for clustering as it is based on absolute value distance, as opposed to squared error (Euclidean distance) and less influence by outliers. Most mutant were slightly tolerant against GEN at IC90 however all growth were inhibit at 1.5xIC90. Mutation in *ptsN, visC, iscA, cyoA, fur, yaaA, yqiC*, *nagA, ydaT, nuoG, cysE, sucB, ubiH,* and *aceF* showed increase tolerance against KAN. All mutant were susceptible to STR and CIP. Overall, most mutant were susceptible at IC90 and all mutants were killed at 4xIC90, this indicates that these mutation did not confer to resistance and survival of these mutants were due to tolerance.
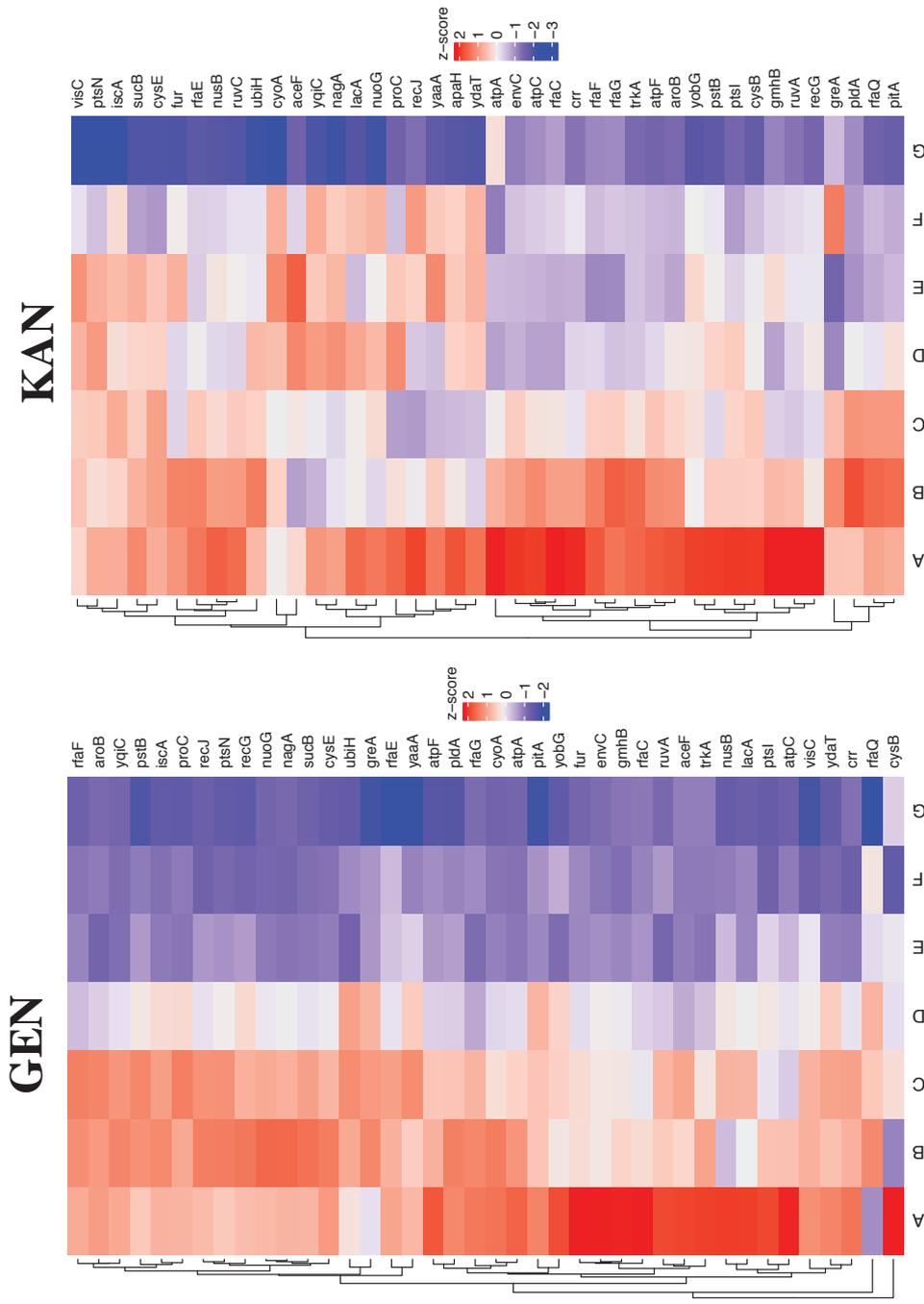
Figure 3.18 Validation of drug tolerant mutant.

The OD value were normalize [z-score] and Manhattan distance were used for clustering. Inhibitory concentration (A: IC0, B: IC50, C: IC70, D: IC90, E: 1.5xIC90, F: 2xIC90, G: 4xIC90). KAN tolerant mutant: *ptsN,visC, iscA,cyoA, fur, yaaA, yqiC, nagA, ydaT, nuoG, cysE, sucB, ubiH, aceF*. Values were average of six replicates.
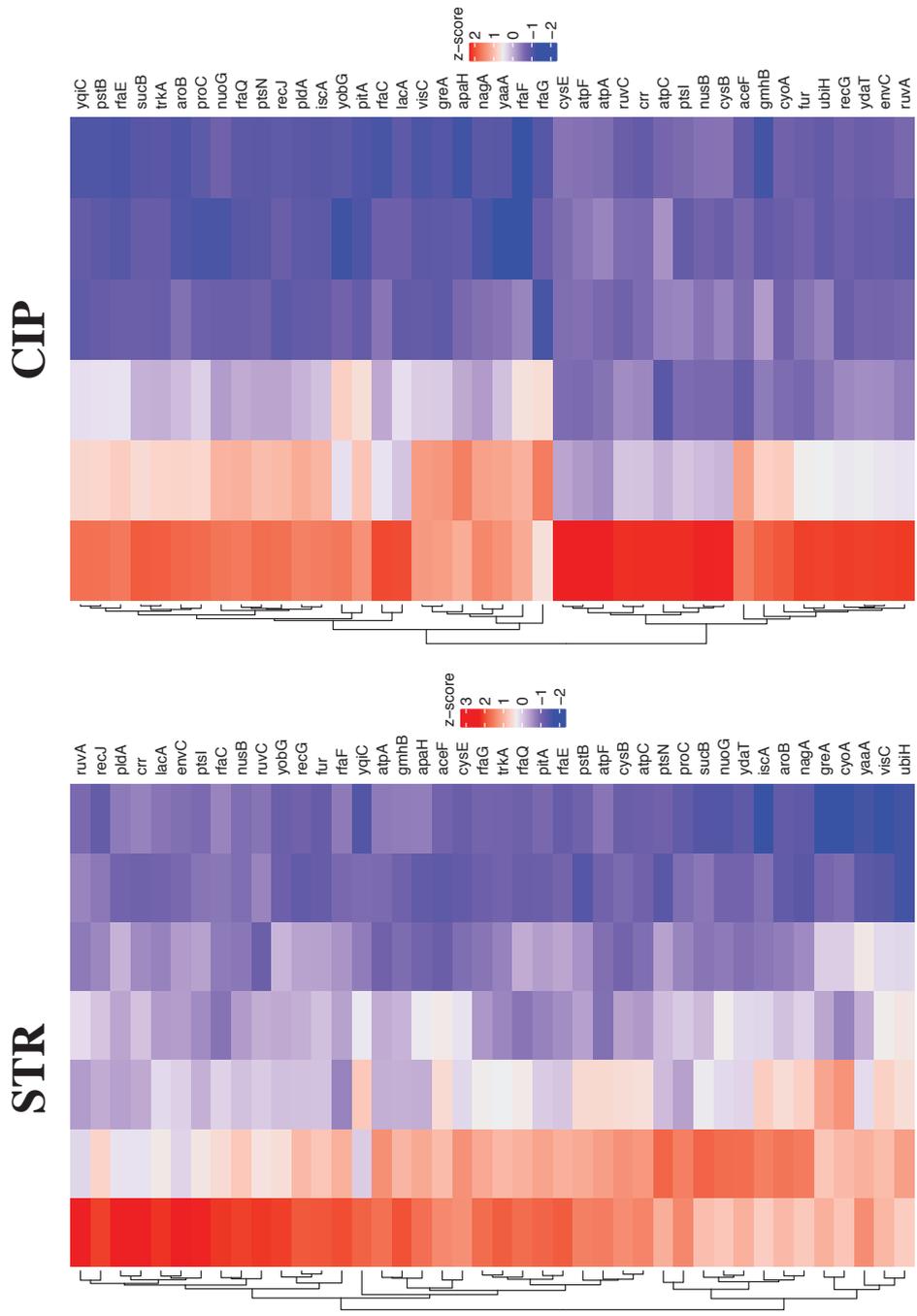
Figure 3.18 (cont.) Validation of drug tolerant mutant. The OD value were normalize [z-score] and Manhattan distance were used for clustering. Inhibitory concentration (A: IC0, B: IC50, C: IC70, D: IC90, E: 1.5xIC90, F: 2xIC90, G: 4xIC90). All mutant were susceptible to STR and CIP. Values were average of six replicates.

3.3.7 Comparison between LTSP and drug tolerant mutant

During LTSP experiment, bacteria were exposed to various stress conditions such as extreme pH, osmotic stress, extreme temperature, nutrient limitation and exposure to toxic compounds. LTSP conditions is akin to bacteria conditions in nature environment and adaptation or mutation that increase survivability during LTSP might give insight on bacteria survival mechanisms in nature. During LTSP, mutant population with growth advantage will arise and displace parental populations. This phenotype is known as GASP (growth advantage during stationary phase)[49]. From the population dynamics shown in Figure 3.19 (A), we observed the appearance of several mutant population with GASP phenotype. Beside GASP, another occurrence that has been reported is VBNC (viable but non-culturable) phenotype[45,53]. VBNC is one of "dormant" state of bacterial cells. From the LTSP experiment, I had observed mutant populations that did not show changes in the population abundance as shown in Figure 3.19(B and C). This result suggest that these *E. coli* mutant populations might become dormant. Dormant is one of characteristic of persisters. Thus, I am interested to compare *E. coli* mutant populations during LTSP and drug tolerant population. Persister cells occurred when exposed to bactericidal drugs, thus I compared the *E. coli* mutant populations at Day 14 of LTSP with *E. coli* mutant populations exposed to 2xIC90 of GEN (inhibit protein synthesis) and CIP (inhibit DNA replication). The results were shown in Figure 3.20 (A and B). Comparison of *E. coli* mutant populations during LTSP and GEN tolerant mutants as in Figure 3.20 (A) presents several interesting findings; (i) GASP mutant were shown to be highly sensitive to GEN, (ii) *cpxA* mutant showed both GASP and GEN tolerant and (iii) GEN tolerant mutants were observed in dormant state during LTSP. Comparison of *E. coli* mutant populations during LTSP and CIP tolerant mutants as in Figure 3.20 (B) also showed similar result whereby CIP tolerant mutant candidates were observed in dormant state during LTSP. GASP mutants also showed to be sensitive to CIP. Interestingly, mutations involved in LPS biosynthesis were observed as low fitness or "loser" population during LTSP however these mutations were shown to be tolerant to CIP. This result suggest that there is trade-off effect between fitness and drug tolerant. Trade-off means a mutation advantageous to one condition might be detrimental to another.

Overall, these results showed that drug tolerant mutant exist in dormant state during LTSP. These characteristics were similar to persisters population which suggest that these mutations might increase persister cells formation and drug tolerance were due to persisters.
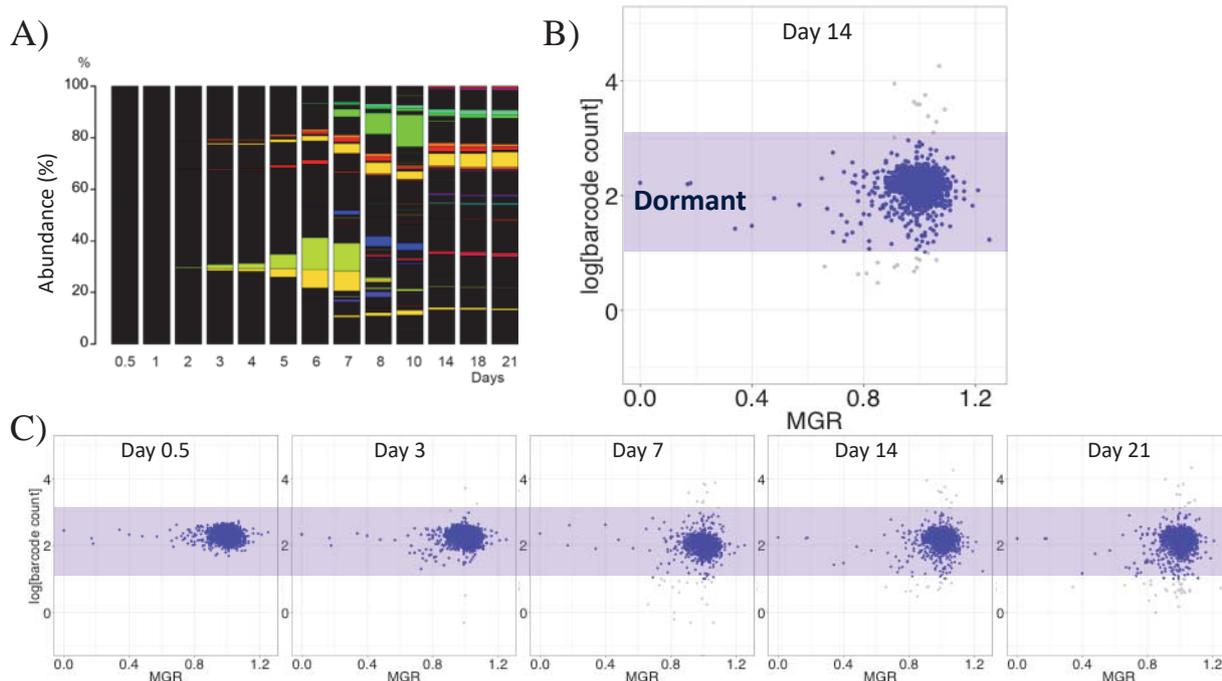
Figure 3.19 Dormant E. coli mutant populations during LTSP.

(A) Population dynamics of E. coli mutant population during LTSP. The colored bar represent specific mutant population with high abundance. (B) Distribution of *E. coli* mutant populations at Day 14. The x-axis represent the maximum growth rate (MGR) of each mutant. The MGR was measured from colony growth on LB agar plate. The y-axis represent the log barcode count of *E. coli* mutants at each time point. (C) Distribution of *E. coli* mutant populations at Day 0.5, Day 3, Day 7, Day 14 and Day 21. Dormant *E. coli* mutant populations were observed during LTSP.

A) — Day 14 vs 2xIC90 plot

GASP mutants

cpxA

Dormant

DRUG tolerant

| GASP mutant |
| yehA, yfaU, ygaZ, ygaH, mprA, endA , cpdA, kdsD and kdsC |

| Energy metabolism |
| nuoN, nuoM, nuoH, nuoC, ptsI, crr,malK |

| Fe-S cluster |
| fdx, hscA, hscB, iscU |

B) — Day 14 vs 2xIC90 plot

GASP mutant

Dormant

Trade-off

DRUG tolerant

| GASP mutant |
| yfaU, ygaZ, ygaH, mprA, endA, cpxA, slyA, yehA,cpdA |

| Energy metabolism |
| nuoM, nuoG, ptsI, crr, atpC |

| LPS biosynthesis |
| **gmhB, rfaC, rfaE, rfaF, rfaG, rfaH** |

Figure 3.20 Comparison of E. coli mutants during LTSP and drug tolerant

(A) Comparison of *E. coli* mutant population during LTSP and GEN tolerant mutant population. (i) The GASP mutant were shown to be highly sensitive to GEN, (ii) *cpxA* mutant showed both GASP and GEN tolerant and (iii) GEN tolerant mutants were observed in dormant state during LTSP. (B) Comparison of *E. coli* mutant population during LTSP and KAN tolerant mutant population. Similarly, GASP mutant were sensitive to CIP and CIP tolerant mutant were observed in dormant state during LTSP. Mutation in LPS biosynthesis showed trade-off effect between fitness and drug tolerance.

Chapter 4 <u>DISCUSSION</u>

<u>4.1 Establishment of experimental procedure by Bar-seq</u>

Our lab had established a new experimental resource; an *E. coli* single-gene deletion library with barcode called as "ASKA barcode deletion collection". In this library, each of the target ORF was replaced with chloramphenicol resistant cassette and 20 nucleotides of random sequence which referred to as barcode. The advantage of barcode is now I can pool all the mutant strains and grow them under competitive growth approach which allows me to truly assessed the fitness of each mutant.

Several key points were considered for analysis using Bar-seq. First, the read depth or coverage for each strain during experiment. Although there is no specific threshold for coverage, I used 300 counts/barcode or strain as my threshold based on recommendation by Smith and coworker[102] from their sequencing result using YKO. Second, I used lower PCR cycle to amplify the barcode fragments. This step is taken to minimized errors during PCR amplification. Third, technical replicates were performed to assess problems that might occur during sequencing (errors or biases). Fourth, experimental replicates were performed to assess the reproducibility of experiments using ASKA barcode deletion collection. Fifth, one limitation of this library is that I am unable to distinguished if the mutant strain acquire secondary or more mutations that may affect the its growth during selections. To overcome this problem, our lab had two independent deletion isolates; assigned as clone 1 and clone 2, each with its own unique barcode which allows mixing of both clones in a single flask thus serve as biological replicate within the same experimental set. Only clones that showed good correlation were valid for further analysis.

From the sequencing results, almost 90% of the raw sequences were successfully mapped to our reference sequences. This indicate that our sequencing run were successful (low sequencing errors) as not many sequences were discarded. In addition, high correlation were observed for most technical replicates and experimental replicates. This indicates that no sequencing bias between samples and experiments using ASKA barcode deletion collection are reproducible.

However, I had identified a problem with mix mutant library glycerol stocks. Here, several mutant strains identified as essential gene deletion strains (present due to gene duplication) and mutator strains (*dam, mutL, mutH, mutM, mutS, mutT, mutY, uvrD, xthA*) were actually removed prior to making the mix mutant library glycerol stock. Yet, these strains were still detected during sequencing (most at low abundance). Since independent isolates stock

were kept in 96-well plate format, this problem might be due to contamination between wells of each independent isolate. To solve this problem, our lab had kept glycerol stock of each independent isolate in vials instead of 96-well plate. New mix library glycerol stock will need to be made for future studies.

4.2 Long-term stationary phase: GASP mutant candidates

Bacteria are often exposed to harsh environmental conditions (i.e. nutrient limitation, extreme pH and temperature, osmolarity, toxic compounds, etc.). To survive these stressful environments, most bacteria adopted several survival mechanisms; for example, Gram-positive bacteria enters dormancy by formation of spores or cyst. On the other hand, non-sporulating bacteria such as Gram-negative bacteria can also survive such conditions without entering this state.

Analysis of bacteria during long-term stationary phase (LTSP) condition which mimics such stress conditions, thus gives insights on survival strategies adopted by bacteria in nature. Two common phenomena observed during LTSP are GASP population and VBNC. During LTSP, mutations conferring to better fitness appear in population which would become dominant and displaced the original population[46–48,51,120]. This phenotype is defined as growth advantage in stationary phase (GASP). I had observed changed in the *E. coli* population dynamics occurred by Day 3 during LTSP (Figure 3.5). This was due to changes in the medium condition (decreased nutrient) and medium condition with increased pH ~9 (Figure 3.2) could not support the normal populations' growth and leads to cell death. This result corroborate with Finkel's report that *E. coli* enters death phase after 3 days prolonged incubation in LB medium[48]. Here onwards, the surviving *E. coli* populations are mutants that have growth advantage during stationary phase (GASP). I had identified 31 mutant population that showed GASP phenotype (Table 3.6). From my result, only two mutations; *rpoS*[48,49] and *sspA*[118] were reported as GASP phenotype.

Out of the 31 GASP mutant, 11 of these mutant have mutation involved in transcription regulation (*rpoS, mraZ, cspC, cspE, slyA, iscR, tdcA, mprA, rbsR, fimE,* and *sspA*), 10 involved in membrane transport, biosynthesis and stress (*ygaH, ygaZ, metI, metN, yrbG, hsrA, kdsC, kdsD ,pldA* and *cpxA*) and others such as ATP- dependent serine protease (*clpA*), DNA-specific endonuclease (*endA*), putative fimbrial-like adhesin protein (yehA), phosphodiesterase (*cpdA*

and *yahA*), phospholipase (*rssA*), ribosomal protein subunit (*rimK* and *yggJ*) and 2-keto-3-deoxy-L-rhamnonate aldolase (*yfaU/rhmA*). Using this mutant information, I proposed several survival mechanisms for GASP mutant population during LTSP shown in Figure 4.1.

4.2.1 Down regulation of rpoS confer growth advantage during stationary phase (GASP)

Mutation in *rpoS* are known to confer GASP phenotype[48,111,121] which suggest that down regulation or loss of function of *rpoS* would confer to growth advantage to *E. coli* population during LTSP. In this study, I found that *rpoS* mutant and several genes (*cspC*, *cspE*, *sspA* and *hsrA*) that were involved in *rpoS* regulation showed growth advantage during LTSP. It is interesting that although *rpoS* is required for survival and adaptation of *E. coli* under various stress conditions (during stationary phase), many strains have acquired mutations in the *rpoS* gene[46,111,121–124]. The relationship between loss/decreased of RpoS activity and GASP phenotype is still unclear. Zambrano and co-workers[46], had suggested that mutation in *rpoS* leads to misregulation of members in the *rpoS* regulon. Cho et. al.[125], shown that in *rpoS* null mutant most of σ^S-dependent genes were found to be down-regulated. As the σ^S competes with other sigma factors to bind to the core RNA polymerase complex, this loss will affect the balance in the competition among sigma factors. Using a theoretical model, Mauri and co-worker[126] had shown that the up-regulated genes in *rpoS* mutant were found to be genes that were transcribed by both σ^S and σ^{70} (*rpoD*) which was proved by these theory.

Mutations in *cspC* and *cspE* showed growth advantage during LTSP, both were detected at Day 3 (top 100 mutant population). *cspC* and *cspE* (transcriptional anti-terminator and regulator of RNA stability) belong to the cold shock protein *cspA* family (9 genes, *cspA* to *cspI*). Rath and Jawali[127] has reported that loss-of-function of *cspC* confers fitness advantage over WT after 24 hours competition. I confirmed this by performing competition experiment of *cspC* deletion mutant and WT (Figure 3.11). Overexpression of *cspC* and *cspE* upregulate the expression of *rpoS*, by increasing stability of *rpoS* mRNA[128]. Meanwhile, deletion of *cspC* and *cspE* genes decreased the expression level of *rpoS, uspA* (universal stress protein), *mopAB* (GroEL and GroES), *dps* (DNA binding protein) and *katG* (catalase)[129].

Stringent starvation protein A, *sspA* is a global regulator important in *E. coli* stress response during stationary phase[130]. The known phenotypes of the *sspA* mutant in *E. coli* are lack of P1 phage growth[131]. Hansen and co-workers[130] had shown that *sspA* mutant are sensitive to acid stress and increase motility (by 2-fold compared to WT). Furthermore, the *rpoS* expression decreased in *sspA* mutant due to increase of intracellular level of H-NS. The *sspA*

negatively regulates the intracellular level of H-NS while H-NS negatively regulates expression of *rpoS*, thus loss-of-function of *sspA* increased H-NS levels leads to decrease the expression of *rpoS*[130]. In their latest evolution experiment using LB medium, Kram and co-workers[118] had found mutation in *sspA* to confer advantage in long-term cultures. In addition, competition experiment showed that *sspA* mutant have better fitness compared to WT[118].

HsrA is predicted multi-drug efflux pump. A study by Goodrich-Blair and Kolter[132], showed that overexpression of *hsrA* leads to accumulation of homocysteine, causing increased in homocysteine thiolactone (HCTL). They suggest that HCTL can act as a signal molecule to increase expression levels of RpoS. Also, HTCL accumulation were shown to increase at the onset of stationary phase which correlates with *rpoS* induction. In addition, *hsrA* mutant has reduced level of RpoS compared to WT. In this study, *hsrA* mutant were shown to have better fitness compared to WT (Figure 3.11).

4.2.2 Adaptation to efficient substrate utilization

Luria-Bertani (LB) medium is widely used as it allows fast growth and compatible for many species[133]. Sezonov and co-workers[133] had studied the composition of spent LB medium (*E. coli* was grown in LB medium for 24 hours; OD at 6.49 and ~pH9). They had shown that stop growth of *E.coli* (stationary phase) was due to limiting carbon source as addition glucose allows *E. coli* to grow in the spent medium. *E. coli* had consumed all fermentable carbon sources in LB and need to switch to amino acids as carbon source using the easier to utilize amino acids until they are depleted, then switching to the harder to utilize amino acids (sequential catabolism of amino acid; aspartate, arginine, serine, threonine, proline and glycine were consumed first, while other amino acid such as leucine and valine were used in the later phase). Beside mutation in *rpoS*, subsequent mutations that conferred GASP phenotype have been reported[50,51]. These additional mutations have were *lrp*, coding the leucine-responsive regulator protein as a global regulator[51], or to the *ybeJ-gltJKL* cluster, encoding a high affinity aspartate and glutamate transporter[47]. A mutation in the DNA-binding domain of *lrp* has been shown to cause a GASP phenotype by increase amino acid catabolism during starvation, and mutants having *ybeJ-gltJKL* also showed GASP phenotypes by increase in amino acid utilization[51]. Furthermore, nutrients provision during LTSP most probably are lysed cells of pre-existing population. Hence, mutation that increase nutrient uptake and catabolism of available nutrients would have strong fitness advantage during LTSP.

The *mraZ* gene is a transcriptional repressor involved in control of cell division and cell wall genes[134], is highly conserved in prokaryotes[135]. Loss-of-function of *mraZ* did not show any significant phenotype and viability compared to WT during stationary phase. Meanwhile, overexpression of *mraZ* inhibit cell division (becomes filamentous) and were found to be lethal at high induction levels[134]. Furthermore, *mraZ* mutant significantly increased expression of genes involved in arginine catabolism (*astA, astB, astC* and *astD*), putrescine utilization pathway (*aldA, puuA, puuB, puuD* and *puuP*) and fatty acid catabolism (*fadA, fadB, fadD* and *fadE*)[134]. This suggest that *mraZ* mutant confer growth advantage by better utilization of nutrient.

Another mutant that showed growth advantage during LTSP is *rssA* mutant. *rssA,* hypothetical protein (putative patatin-like phospholipase) is part of *rssA-rssB* operon in which *rssB* has been identified as a sigma S- specific recognition factor involves in *rpoS* proteolysis but the role of *rssA* is still unclear. Ruiz and co-workers[136] had shown that *rssA* does not regulate either *rssB* or *rpoS* expression. Mutation in *rssA* had shown growth advantage during LTSP in both R1 and R2 experiment (appears by Day 3). Competition experiment shown that *rssA* mutant higher fitness than WT (Figure 3.11), although growth rate of *rssA* mutant is similar as WT. In addition, mutation in *rssA* were reported to increase specific growth rates of strain evolved for fast growth in medium with glucose as sole carbon source[137]. This suggest that mutation in *rssA* increases the efficiency of nutrient transport and metabolism.

rbsR is a transcriptional repressor of ribose metabolism. Shimada *et. al.*[138] had shown that *rbsR* not only involves in transport and metabolism of D-ribose but also regulate the *de novo* synthesis of purine nucleotide (*purD* and *purH*) and act as an activator salvage pathways of purine nucleotide synthesis (*add* and *udk*). *rbsR* mutant were shown to activates the ribose metabolism (increased level of *rbsD* expression) and purine biosynthesis pathway (increased level of *purH* expression).

tdcA is a transcriptional activator of *tdc* operon which is involved in the transport and metabolism of L-threonine and L-serine during anaerobic growth. It is the first gene in the operon and was shown that mutation in *tdcA* reduced the expression of its downstream gene (*tdcABCDEFG*)[139]. Although, I did not find much information of *tdcA* mutant using *E. coli*. Kim and co-workers[140] had studied the *tdc* operon using *Salmonella typhimurium*. Sequence analysis of *tdc* operon in *S. typhimurium* shown about ~80% nucleotide and ~95% amino acid homology with *E. coli*. The big difference is that *S. typhimurium* lacks *tdcR* and *tdcF* genes. They had shown that *tdcA* mutant increased expression of genes involved in transport and degradation of carbohydrate and sugars such as propanediol (*pduB, pduC, pduD, pduF, pduL,*

pduM, pduO), maltose (*lamB*, *malE, malK, malM , malF, malP*) and mannose (*manY*), thus enable the bacteria to utilize versatile carbon sources.

4.2.3 Mutation involves in membrane and membrane transport might facilitates nutrient uptake

Although how these mutation confer to growth advantage during stationary phase is unclear, several mutants with mutation involved in cell membranes and transport were detected as GASP candidates. These are *kdsC* and *kdsD* (LPS core biosynthesis), *cpxA* (envelope stress), *pldA* (phospholipase A), *metNI* (methionine transporter) and *ygaZH* (putative valine transporter).

The *kdsC* and *kdsD* gene encoding a 3-deoxy-D-manno-octulosonate-8-phosphate phosphatase and D-arabinose 5-phosphate isomerase, respectively are involved in lipopolysaccharide core biosynthesis were assumed to be essential genes. However, Sperandeo and co-workers[141] shown that deletion of *kdsC* and *kdsD* to be non-essential, indicating genetic redundancy for these two functions. On the other hand, these mutants showed increased sensitivity to hydrophobic toxic chemicals, which suggest the mutants have defective outer membrane thus affect the cell membrane integrity. However, I found mutant of *kdsC* and *kdsD* to show high population abundance during LTSP.

Delhaye and co-workers[119], reported that *cpxA* mutant (exhibit abnormal growth and morphology. The mutant displayed significant growth defects: mass doubling times were at least twice as long as those of wild-type cells and filamentous morphology mixed with DNA free minicells which suggested defect in cell division. The author had conclude that the Cpx system play a role in maintaining the peptidoglycan homeostasis and mutation in *cpx* genes affect the bacterial cell wall structure. On the other hand, *cpxA* mutant was known to exhibit resistance against beta-lactams and aminoglycosides[19]. This indicate that *cpxA* mutant have fitness advantage under stress condition.

metN and *metI* encode components of an ABC-type methionine transporter (a putative ATPase and membrane-spanning region, respectively). Mutation in *met N* and *metI* destroy the transporter activity, mutants are unable to transport methionine. Furthermore, methionine is important to protect the cells from NO^- (nitric oxide) stress, mutants were shown to be sensitive to NO^- [142]. Another study shown that *metNI* mutant were hypersensitive to CO (carbon monoxide)[143]. Meanwhile, another two genes (*ygaZH*) encoding putative L-valine transporter

which located upstream from *mprA* also showed GASP phenotype. Mutation in *ygaZH* shown sensitive to DL-norvaline (L-valine analogue)[144].

4.2.4 Increase motility gives advantage during LTSP

MprA (also known as *EmrR*) is a transcriptional regulator that belongs to the MarR family DNA-binding proteins, which control the expression of a range of bacterial genes involved in virulence, resistance to antibiotics, response to oxidative stresses and the catabolism of environmental aromatic compounds[145]. The *mprA* gene is located in an operon together with the *ermAB* genes that encode a multidrug resistance pump. *MprA* represses transcription of *ermAB* by direct binding to its promoter region. Kakkanat et al.[146] reported that *mprA* mutant showed increase hypermotility compared with WT. *mprA* mutant increase transcription of *fliDC* which are involved in flagella formation. Increased motility might be another survival strategy of GASP phenotype, that allows better scavenging of nutrient.

4.2.5 Increase in biofilm gives advantage during LTSP

Biofilm environment provides nutrient and protection which improves bacterial survival in unfavorable conditions[28,40,81]. The *icsR* and *fimE*, encodes, iron-sulfur cluster regulator and type I fimbriae regulatory protein, respectively. Wu and Outten[147] reported that deletion of either genes were shown to increase of type I fimbriae expression which leads to enhance of biofilm production. They reported that increase expression of genes involved in surface adhesion and motility such as *fimAICDFGH* operon (encoding type I fimbriae) and *flu* gene (biofilm formation autotransporter) were observed in *iscR* mutant. they indicate that increase biofilm formation in *iscR* mutant was due to decreased expression of *fimE*. Meanwhile, *fimE* is repressor of the *fimAICDFGH* operon. Hence, loss-of-function of *fimE* would turn on the *fimAICDFGH* operon function thus increase biofilm formation.
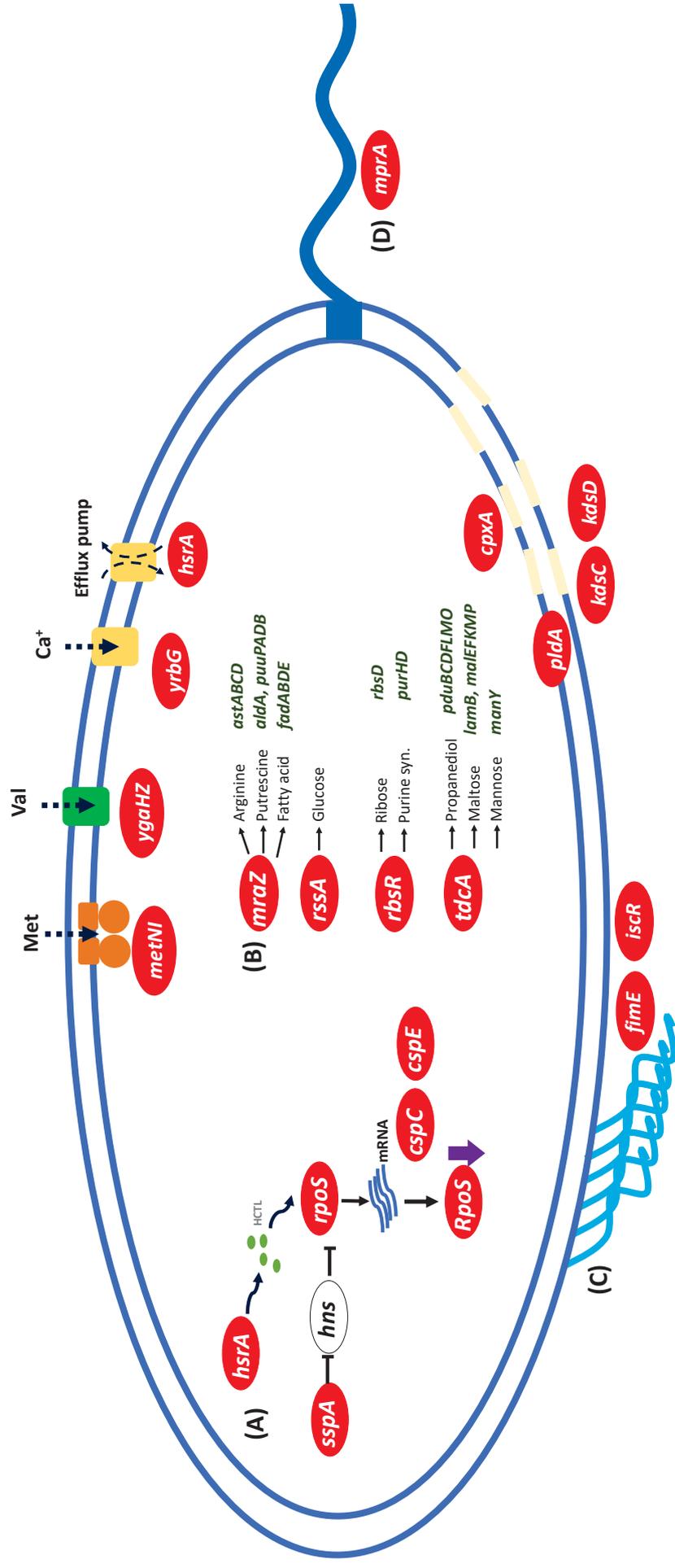
Figure 4.1 Proposed survival mechanisms of GASP mutant populations during LTSP.

(A) Down-regulation of rpoS activity, (B) efficient substrate utilization, (C) increase biofilm formation and (D) increase motility

4.3 Drug tolerant mutant

Antibiotic tolerance is linked with the failure of antibiotic treatment and the setback of many bacterial infections. Tolerance is different from resistance as "resistance" is used to describe the inherited ability of microorganisms to grow at high concentrations of an antibiotic, irrespective of the duration of treatment, and is quantified by the minimum inhibitory concentration (MIC), whereas 'tolerance' is more generally used to describe the ability, whether inherited or not, of bacteria to survive transient exposure to high concentrations of an antibiotic without a change in the MIC[58]. The difference of resistance and tolerance is shown from killing kinetics as in Figure 1.1.

In this study, genome-wide screening for drug tolerant mutant showed that mutations involved in the energy metabolism increase tolerance against multiple drugs. Table 4.1 listed the common function of drug tolerant mutants. I found that these mutations were mainly associated with oxidative phosphorylation (Figure 4.2); NAD(H):ubiquinone oxidoreductase complex (*nuoACEHJMN*), cytochrome complex (*cyoABDE*), ATPase (*atpABCEF*), and Fe-S cluster (*iscU, iscA, hscAB, fdx*) showed increased tolerance against multiple drugs. Others were involves in phosphotransferase system (PTS) such as *ptsI* and *crr* mutants, TCA cycle (*sucB*), NAD(P) transhydrogenase subunit (*pntA*) and iron transporter (*feoB*).

Oxidative phosphorylation is an important process for energy production in cells. It is made up of two components; the electron transport chain and chemiosmosis. In the electron transport chain, electrons are passed from one molecule to another, and energy released in these electron transfers is used to form an electrochemical gradient. In chemiosmosis, the energy stored in the gradient is used to make ATP[148]. Oxidative phosphorylation consist of a series of membrane-embedded proteins and organic molecules, which are organized into five large complexes; complex I (NAD(H):ubiquinone oxidoreductase complex), complex II (succinate dehydrogenase), complex III (cytochrome bcl complex, only in eukaryotes) and complex IV (cytochrome c oxidase) and complex V (ATP synthase)[148,149].

The *nuo* genes cluster encode for NAD(H):ubiquinone oxidoreductase complex (also called complex I, is the main entry point for electrons from NADH into the respiratory chains) as shown in Figure 4.2. Erhardt et al.[150] had reported that deletion of each *nuo* genes will leads to partially assembly of complex I. These mutants were unable to assemble a functional complex I indicating that all *nuo* genes are required for assembly and maintaining the stability of complex I. The NADH oxidase activity decreased from 0.55 U/mg to 30-35 U/mg. This indicate less energy were being converted thus, *nuo* mutants showed decrease growth rate

(liquid LB). This result correlates with MGR (colony growth on LB agar) of *nuo* mutants as shown in Table 3.12. They conclude that the deletion of any of the *nuo*-genes resulted in a complete loss of complex I activity in the membrane due to the loss of a structurally intact complex I.

The conversion of ATP from ADP can occur in two different ways; either by oxidative phosphorylation or substrate level phosphorylation. Substrate-level phosphorylation direct phosphorylation of ADP to ATP with a phosphate group by using the energy obtained from a coupled reaction whereas oxidative phosphorylation is the production of ATP from the oxidized NADH and FADH. Oxidative phosphorylation requires ATPase synthase complex which are encoded by *atp* genes cluster. Jensen and Michelsen [151] had reported that *atp* mutants are dependent solely on substrate-level phosphorylation for generation of ATP which leads to decrease in growth rates and growth yield compared with WT. They had shown that *atp* deletion mutants were affected by deficiency in oxidative phosphorylation of ADP (decrease by 25% compared to WT) and ATP synthesis (decrease by 45% compared to WT) which caused the decrease in growth rates and growth yield. This result correlates with MGR (colony growth on LB agar) of *atp* mutants as shown in Table 3.12.

E. coli has two distinct cytochrome oxidases; cytochrome o oxidase complex, encoded by *cyoABCDE*, and the cytochrome d oxidase complex, encoded by *cydAB*. The difference between these two is its affinity to oxygen; cytochrome o oxidase has a relatively low affinity for oxygen and functions therefore mainly under aerobic conditions, whereas the cytochrome d oxidase complex, which has higher affinity for oxygen, accumulates as oxygen becomes limiting[152]. In this study, I found that mutation in *cyoABDE* increase tolerance against TET. Mempin and co-workers[153] had reported that *cyo* deletion mutants did not affect the growth of the mutants and have no effect on the ATP levels in the cells. This result correlates with MGR (colony growth on LB agar) of *cyo* mutants as shown in Table 3.12. Except *cyoE* mutant which showed slow growth (MGR < 0.88), all *cyo* mutants showed normal growth phenotype.

Another important component of the electron transport chain is the Fe-S proteins. Besides, Fe-S proteins are found to participate in diverse biological process such as respiration, central metabolism, DNA repair and gene regulation[154–157] as shown in Figure 4.2 (B, right panel). Two pathways were known involved in the synthesis of Fe-S complex; suf operon (*sufABCDSE*) and isc operon (*iscRSUA-hscBA-fdx*). In this study, I found that mutation in the isc operon showed increase tolerance against aminoglycoside (KAN and GEN). The mutants were shown in Figure 4.2 (B, left panel). The isc operon of *E. coli* is important for Fe-S cluster assembly in a variety of enzymes, and the *iscS* deletion mutant shows a range of growth defects

including a slow growth phenotype and several auxotrophies[158,159]. This result correlates with MGR (colony growth on LB agar) of *iscU* mutant as shown in Table 3.12. In contrast, no obvious phenotypes for a *sufABCDSE* deletion mutant under normal growth in rich or minimal media have been reported. However, deletion of the two operons is synthetically lethal indicate that the both are complementary pathway[160].Takumoto and Takahashi[161] had reported that mutation in *iscU*, *hscB*, *hscA* and *fdx* exhibit slow growth (doubling time of 55 min). Furthermore, the enzyme activities of Fe-S proteins, glutamate synthase (GltS) and succinate dehydrogenase (SDH) decreased in *iscS, iscU, hscB, hscA* and *fdx* mutant strains which indicate requirement of *iscS, iscU, iscA, hscB, hscA*, and *fdx* for the activity of cellular Fe-S proteins and generation of Fe-S protein. Their result suggest that loss-of-function either one gene causes the loss of function for the remaining proteins. Another drug tolerant mutant candidate, *feoB* mutant were reported to be incapable of transporting ferrous iron[162]. Thus, decrease of intracellular level of iron might affect the formation of Fe-S proteins.

Overall, these results indicates these mutations leads to decrease in the energy levels by disrupting the ATP synthesis. Most of these mutants exhibit slow growth phenotypes. This indicate that drug tolerance is mediated by slow growth. Slow growth has been implicated for bacterial drug tolerance[116]. This is because most drug mechanism work by killing bacteria that are actively growing and multiplying[163]. Thus, slow-growing (or dormant) bacteria having a low metabolic activity are partly or completely unaffected to killing by most antibiotics.

In this study, mutation involved in LPS biosynthesis (*gmhB, rfaC, rfaE, rfaF, rfaG, rfaH),* inner membrane protein (*yidG),* potassium transport (*trkA*) and outer membrane phospholipase (*pldA*) also showed increase tolerance against CIP. Increase resistance to ciprofloxacin had been reported for *E. coli* mutants selected with tigecycline that had defects in the core LPS genes *lpcA, rfaC, rfaE*, and *rfaF*[164]. Another study had also reported that mutations in *rfaD* and *rfaE* showed low-level resistance against CIP[165]. They had suggest that decreased susceptibility against CIP is related to porin regulation. Since, OmpF (major) and OmpC is the channel for CIP entry in *E. coli*, decreased expression levels of both ompF and ompC were detected in many of the ciprofloxacin-resistant mutants[166]. These genes are transcriptionally regulated by the two-component systems OmpR-EnvZ[167] and CpxA-CpxR[168] in response to effects on the cell membrane due to such conditions as temperature, osmolarity, pH, and chemical stress. As mutations in LPS biosynthesis cause alteration in the cell wall structure and membrane homeostasis which lead to activation of membrane stress response. In addition, LPS mutations leads deep rough phenotype, dramatic reduction (>90%) in porin proteins, including OmpF, OmpC, PhoE and LamB and decrease in buoyant density

Table 4.1 List of common drug tolerant mutants

Drug	Energy metabolism
Gentamicin	*crr, fdx, fre, fur, hscA, hscB, iscA, iscU, malK, nuoA, nuoC, nuoE, nuoH, nuoJ, nuoM, nuoN, ptsI*
Kanamycin	*crr, feoB, hscA, hscB*
Streptomycin	*pntA*
Tetracycline	*atpA, atpB, atpC, atpE, atpF, crr, cyoA, cyoB, cyoD, cyoE, nuoA, sucB*
Ciprofloxacin	*atpC, crr, nuoG, nuoM, pgi, ptsI*
Nalidixic acid	*nuoM*

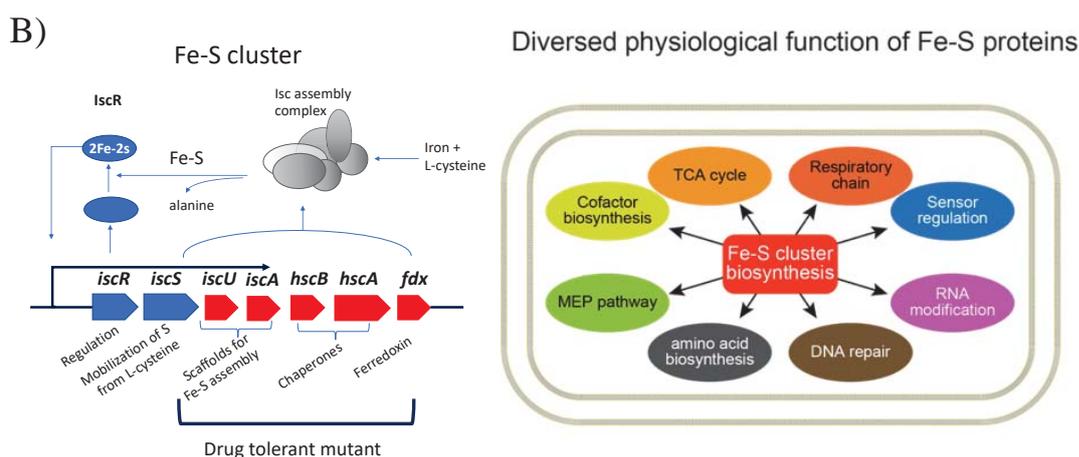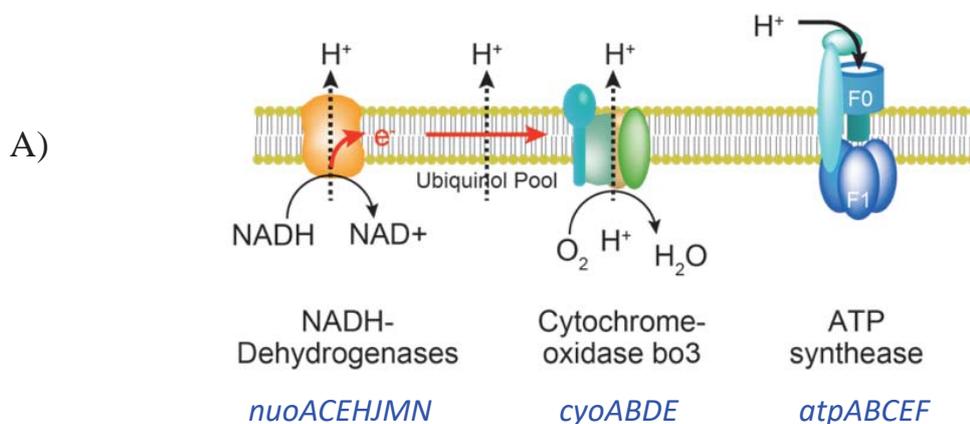| Drug	LPS biosynthesis
Ciprofloxacin	*gmhB, rfaC, rfaE, rfaF, rfaG, rfaH, trkA, yidG, pldA*

A)

B)

Figure 4.2 Mutations conferring to drug tolerant

(A) Mutations related to oxidative phosphorylation; complex I (NADH dehydrogenase), complex IV (cytochorome o oxidase) and complex V (ATPase) (B) Mutations in isc operon for assembly of Fe-S cluster.

of the outer membrane[169]. From these findings, they suggested that membrane stress from the mutant cell wall decreases porin expression in LPS mutants which leads to decreased CIP susceptibility.

4.4 Drug persistence and dormancy

In nature, microorganisms are exposed to various stress conditions which harmful and detrimental to growth. To survive, many bacteria enter a transient state of dormancy that is characterized by a decrease in growth rate and metabolic function[36,116]. Once the stress is alleviated or encounters favorable environmental conditions, these dormant subpopulations would resume growth[44,55]. Two dormancy phenomena have been seen in Gram-negative bacteria which are bacterial persistence[27] and the viable but non-culturable (VBNC) state[44,45]. Although these were studied independently, these dormancy phenomena share several characteristics that demonstrate their relatedness; (i) antibiotic tolerant and (ii) shared molecular mechanisms. Korch and Hill[170] reported that *hipB* (antitoxin) mutant *E. coli* was non-culturable. In addition, overexpression of HipA toxin was shown to induce the VBNC state in *E. coli*[170]. Overexpression of VapC toxin in *Mycobacterium smegmatis* induces the VBNC state[171]. This dormancy state allows bacteria to become tolerant against drug and is associated with the failure of antibiotic treatment and relapse of bacterial infections. This suggest that both persisters and VBNC cells are likely to be source of drug persistence. Furthermore, bacterial persistence have been shown to lead to the emergence of antibiotic resistance[58,172].

4.5 Drug tolerance were due to persisters

Studies on drug persistence had revealed several mechanisms involved in persisters formation or survival such as toxin-antitoxin system by overexpression of toxin[32,173,174], stringent response, SOS response[28] and nutrient stress[40,43,175]. Persister cells are antibiotic drug tolerant by definition because they are isolated in cultures exposed to high-dose of drug[31,33]. Two main characteristics of persisters are drug tolerant and dormant (low metabolic activity). To understand the mechanism of persisters formation, I use the characteristics of persister cells, to find mutation that have tendency to become persisters.

First, I performed genome-wide screening for drug tolerant mutant. Here, mutations involves in energy synthesis and conversion are shown to be drug tolerant. From the sequencing result, mutation involves in energy metabolism; ATPase (*atpABCEF*), PTS (*ptsI* and *crr*), cytochrome complex (*cyoABDE*), NAD(H):ubiquinone oxidoreductase complex

(*nuoACEHJMN*) and Fe-S cluster (*iscUA-hscAB-fdx*) showed increased tolerance against multiple drugs. This result correlates with findings by Lewis's group, which indicates that persister formation was dependent on ATP levels[41]. Low ATP level decreases the antibiotic target activity which leads to persisters formation. This is because most bactericidal antibiotics kill bacteria by inhibiting cellular target that is required during active cell growth. In addition, TisB-IstR-1 (toxin-antitoxin system) showed that increase expression of tisB leads to incorporation of TisB protein into the membrane. Thus, cause loss of membrane potential leading to decrease in ATP levels and inhibited cell growth (dormant)[37]. This strongly suggest that decreased ATP levels leads to increase persisters formation. Since, drug tolerant mutants identified in this study exhibit slow growth phenotype (due to low ATP) this indicate that these mutants became tolerant due to persisters formation.

Figure 1.1 showed the killing curves during treatment with bactericidal drug, the drug tolerant and persisters population demonstrate the same susceptibility (MIC) as the original population. I had performed susceptibility assay and confirmed that most of these drug tolerant mutants were still susceptible to MIC concentration of drugs as shown in Section 3.3.6. This indicates that these mutations did not confer to resistance.

Persister cells arise due to dormancy, in which the cells are metabolically inactive or very low activity. To investigate if the drug tolerant mutant also exhibit dormant state, I compare the distribution of *E.coli* mutants during LTSP and 2xIC90 (GEN or CIP) as shown in Figure 3.20. Here, I showed that both GEN and CIP tolerant mutants were observed in dormant state during LTSP. Furthermore, I found that mutations involved in LPS biosynthesis were observed as low fitness or "loser" population during LTSP however these mutations were shown to be tolerant to CIP. This result suggest that there is trade-off effect between fitness and drug tolerant. This suggest that drug tolerance against CIP due to mutations in LPS biosynthesis might not due to persister cells formation.

These results indicates that these mutants shared similar characteristics with persisters. First, these mutants are drug tolerant. Second, mutants exhibit slow growth due to decrease in ATP, corroborate with other researchers[37,41] linked to persisters formation. Third, most mutations did not confer to resistance. Lastly, mutants exhibit dormant state. Overall, these results strongly suggest that drug tolerance were due to persister cells.

4.6 Concluding remarks

We had established the experimental procedure for Bar-seq analysis using *E. coli* ASKA barcode deletion collection. High correlation between independent experiment replicates proved that experiments with ASKA barcode collection were reproducible. Thus, ASKA barcode deletion collection is an effective system for genome-wide screening studies (interaction between gene and specific environment) to comprehensively understand biological systems. In addition, we showed that we are able to monitor *E. coli* mutant population dynamics during LTSP for 3 weeks under batch culture conditions by Bar-seq. We observed appearance of GASP mutant populations during LTSP. We identified 31 mutant candidates that showed GASP phenotypes, of which two mutants (*rpoS* and *sspA*) had been reported. Using this approach, we able to obtain new novel findings which give better understanding toward survival mechanisms of *E. coli* during LTSP.

Genome-wide screening for drug tolerant mutant were performed to investigate bacterial persisters phenomenon. I found that drug tolerant mutants exhibit slow growth phenotype due to disruption in protein complexes involved in oxidative phosphorylation which leads to decrease in ATP. Susceptibility assay showed that these mutants were susceptible to same MIC concentration as WT. Furthermore, these mutants were shown to become dormant during LTSP. This result supports the current understanding of persisters formation which strongly suggest that these mutations might contributes to persisters formation.

However, validation experiment needs to be carried out to confirm that drug tolerant mutants were persisters. Although there are good evidence for several strains from literature, I think that transcriptome analysis should be perform on these GASP mutants to understand the regulatory changes that occur to adapt in natural environment. Another is genetic interaction using GASP mutant as query to find other genes that have functional relationships and/or shared pathways. The change in the nutrient composition is one of main reason for selection of GASP mutant thus metabolomics analysis of spent LB during LTSP might give reason of why these mutant were selected.

Acknowledgements

First and foremost, I would like to express my sincere gratitude and appreciation to my supervisor, Prof. Hirotada Mori for his guidance, help and critical idea throughout my research. Also thank you to Dr. Muto-Fujita for her advice; especially on the statistics and bioinformatic analysis during my study.

This dissertation would not have been possible without my committee members; I would like to acknowledge Prof. Takagi and Assc. Prof. Akiyama, for their insightful comments and encouragement throughout my study.

My sincere thanks to Dr. Nozomu Yachie from RCAST, Tokyo University for his advices on primer design for Illumina sequencer and our many discussion on LTSP project. I also gratefully thank Dr. Kubo from NAIST, Dr. Mutoaki Seki from RCAST, Tokyo University and Assc. Prof. Satoru Watanabe and Dr. Yu Kanesaki. from Tokyo Univ. of Agriculture for the use of the sequencing services and facilities.

I thank my fellow labmates in Laboratory of Systems Microbiology; Toda san, Shelley, Taiki Akiba , Kohei Kondo, Yuichiro Tanaka, Kazuichi Makishi, Steven Bowden, Wataru Nomura and Yuta Otsuka for their wonderful friendship and support in the lab.

Also, thanks to MEXT and NAIST for providing the financial support for my study and conferences, and the staffs whom I owe so much, for helping the international students in one way or another.

Finally, I would also like to express my appreciation to my family who had continually given me their endless supports, advices and encouragement.

Thank you.

References

1. Fleming, A. On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzae*. *Br. J. Exp. Pathol.* **10,** 226–236 (1929).

2. Lee Ligon, B. Sir Howard Walter Florey—the force behind the development of penicillin. *Semin. Pediatr. Infect. Dis.* **15,** 109–114 (2004).

3. Tan, S. Y. & Tatsumura, Y. Alexander Fleming (1881–1955): Discoverer of penicillin. *Singapore Medical Journal* **56,** 366–367 (2015).

4. Jeśman, C., Młudzik, A. & Cybulska, M. [History of antibiotics and sulphonamides discoveries]. *Pol. Merkur. Lekarski* **30,** 320–2 (2011).

5. Schatz, A. & Waksman, S. A. Effect of Streptomycin and Other Antibiotic Substances upon *Mycobacterium tuberculosis* and Related Organisms.,. *Exp. Biol. Med.* **57,** 244–248 (1944).

6. Bo, G. Giuseppe Brotzu and the discovery of cephalosporins. *Clin. Microbiol. Infect.* **6,** 6–8 (2000).

7. MCGUIRE, J. M. *et al.* Ilotycin, a new antibiotic. *Schweiz. Med. Wochenschr.* **82,** 1064–1065 (1952).

8. Lesher, G. Y., Froelich, E. J., Gruett, M. D., Bailey, J. H. & Brundage, R. P. 1,8-Naphthyridine Derivatives. A New Class of Chemotherapeutic Agents. *J. Med. Pharm. Chem.* **5,** 1063–1065 (1962).

9. Mitscher, L. A. Bacterial topoisomerase inhibitors: Quinolone and pyridone antibacterial agents. *Chemical Reviews* **105,** 559–592 (2005).

10. Birnbaum, J., Kahan, F. M., Kropp, H. & Macdonald, J. S. Carbapenems, a new class of beta-lactam antibiotics. *Am. J. Med.* **78,** 3–21 (1985).

11. Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517,** 455–459 (2015).

12. Anderson, K. L. The complex world of gastrointestinal bacteria. *Can. J. Anim. Sci.* **83,** 409–427 (2003).

13. Anderson, K. Is Bacterial Resistance To Antibiotics An Appropriate Example Of Evolutionary Change? *Creat. Res. Soc. Quartely* **41,** 318–326 (2005).

14. Kashiwagi, K. *et al.* Relationship between spontaneous aminoglycoside resistance in *Escherichia coli* and a decrease in oligopeptide binding protein. *J. Bacteriol.* **180,** 5484–5488 (1998).

15. Miller, C. *et al.* SOS response induction by beta-lactams and bacterial defense against antibiotic lethality. *Science* **305,** 1629–31 (2004).

16. Paul, R., Postius, S., Melchers, K. & Schäfer, K. P. Mutations of the *Helicobacter pylori* genes *rdxA* and *pbp1* cause resistance against metronidazole and amoxicillin. *Antimicrob. Agents Chemother.* **45,** 962–965 (2001).

17. Aldred, K. J., Kerns, R. J. & Osheroff, N. Mechanism of quinolone action and resistance. *Biochemistry* **53,** 1565–1574 (2014).

18. Watson, M., Liu, J.-W. & Ollis, D. Directed evolution of trimethoprim resistance in

Escherichia coli. *FEBS J.* **274,** 2661–2671 (2007).

19. Kohanski, M. A., Dwyer, D. J., Wierzbowski, J., Cottarel, G. & Collins, J. J. Mistranslation of Membrane Proteins and Two-Component System Activation Trigger Antibiotic-Mediated Cell Death. *Cell* **135,** 679–690 (2008).

20. Kohanski, M. A., Dwyer, D. J. & Collins, J. J. How antibiotics kill bacteria: From targets to networks. *Nature Reviews Microbiology* **8,** 423–435 (2010).

21. Garneau-Tsodikova, S. & Labby, K. J. Mechanisms of resistance to aminoglycoside antibiotics: Overview and perspectives. *MedChemComm* **7,** 11–27 (2016).

22. Fabrega, A., Madurga, S., Giralt, E. & Vila, J. Mechanism of action of and resistance to quinolones. *Microb Biotechnol* **2,** 40–61 (2009).

23. Speer, B. S., Shoemaker, N. B. & Salyers, A. A. Bacterial resistance to tetracycline: Mechanisms, transfer, and clinical significance. *Clinical Microbiology Reviews* **5,** 387–399 (1992).

24. Centers for Disease Control and Prevention (CDC). Centers for Disease Control and Prevention. at <https://www.cdc.gov/>

25. CDC, C. for D. C. and P. *et al. Antibiotic resistance threats in the United States, 2013. Current* (2013). doi:CS239559-B

26. Hobby, G. L., Meyer, K. & Chaffee, E. Observations on the Mechanism of Action of Penicillin. *Exp. Biol. Med.* **50,** 281–285 (1942).

27. Bigger, J. W. Treatment of staphylococcal infections with penicillin by intermittent sterilization. *Lancet* **244,** 497–500 (1944).

28. Lewis, K. Persister Cells. *Annu. Rev. Microbiol.* **64,** 357–372 (2010).

29. Moyed, H. S. & Bertrand, K. P. *hipA*, a newly recognized gene of *Escherichia coli* K-12 that affects frequency of persistence after inhibition of murein synthesis. *J. Bacteriol.* **155,** 768–75 (1983).

30. Edo Kussell and Stanislas Leibler. Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. *Science (80-.).* **309,** 2075–2078 (2005).

31. Keren, I., Kaldalu, N., Spoering, A., Wang, Y. & Lewis, K. Persister cells and tolerance to antimicrobials. *FEMS Microbiol. Lett.* **230,** 13–18 (2004).

32. Schumacher, M. A. *et al.* Molecular mechanisms of HipA-mediated multidrug tolerance and its neutralization by HipB. *Science (80-.).* **323,** 396–401 (2009).

33. Keren, I., Shah, D., Spoering, A., Kaldalu, N. & Lewis, K. Specialized Persister Cells and the Mechanism of Multidrug Tolerance in Escherichia coli Specialized Persister Cells and the Mechanism of Multidrug Tolerance in Escherichia coli. *J. Bacteriol.* **186,** 8172–8180 (2004).

34. Shah, D. *et al.* Persisters: A distinct physiological state of *E. coli.* *BMC Microbiol.* **6,** (2006).

35. Maisonneuve, E., Shakespeare, L. J., Jorgensen, M. G. & Gerdes, K. Bacterial persistence by RNA endonucleases. *Proc. Natl. Acad. Sci.* **108,** 13206–13211 (2011).

36. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L. & Leibler, S. Bacterial persistence as a phenotypic switch. *Science (80-.).* **305,** 1622–1625 (2004).

37. Dörr, T., Vulić, M. & Lewis, K. Ciprofloxacin causes persister formation by inducing

the TisB toxin in Escherichia coli. *PLoS Biol.* **8,** (2010).

38.   Korch, S. B., Henderson, T. A. & Hill, T. M. Characterization of the hipA7 allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. *Mol. Microbiol.* **50,** 1199–1213 (2003).

39.   Nguyen, D. *et al.* Active starvation responses mediate antibiotic tolerance in biofilms and nutrient-limited bacteria. *Science (80-.).* **334,** 982–986 (2011).

40.   Amato, S. M. & Brynildsen, M. P. Nutrient transitions are a source of persisters in *Escherichia coli* biofilms. *PLoS One* **9,** 1–9 (2014).

41.   Shan, Y. *et al.* ATP-Dependent persister formation in *Escherichia coli. MBio* **8,** (2017).

42.   Shan, Y., Lazinski, D., Rowe, S., Camilli, A. & Lewis, K. Genetic basis of persister tolerance to aminoglycosides in Escherichia coli. *MBio* **6,** (2015).

43.   Mlynarcik, P. & Kolar, M. Starvation- and antibiotics-induced formation of persister cells in Pseudomonas aeruginosa. *Biomed. Pap.* **161,** 58–67 (2017).

44.   Oliver, J. D. Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS Microbiol. Rev.* **34,** 415–425 (2010).

45.   Oliver, J. D. The viable but nonculturable state in bacteria. *J. Microbiol.* **43 Spec No,** 93–100 (2005).

46.   Zambrano, M. M., Siegele, D. a, Almirón, M., Tormo, a & Kolter, R. Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. *Science* **259,** 1757–1760 (1993).

47.   Zinser, E. R. & Kolter, R. Mutations enhancing amino acid catabolism confer a growth advantage in stationary phase. *J. Bacteriol.* **181,** 5800–5807 (1999).

48.   Finkel, S. E. Long-term survival during stationary phase: evolution and the GASP phenotype. *Nat.Rev.Microbiol.* **4,** 113–120 (2006).

49.   Finkel, S. E. & Kolter, R. Evolution of microbial diversity during prolonged starvation. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4023–7 (1999).

50.   Zinser, E. R. & Kolter, R. *Escherichia coli* evolution during stationary phase. *Res. Microbiol.* **155,** 328–336 (2004).

51.   Zinser, E. R. & Kolter, R. Prolonged stationary-phase incubation selects for lrp mutations in *Escherichia coli* K-12. *J. Bacteriol.* **182,** 4361–4365 (2000).

52.   Kim, J.-S., Chowdhury, N. & Wood, T. K. Viable But Non-Culturable Cells Are Persister Cells. *bioRxiv* 186858 (2017). doi:10.1101/186858

53.   Kim, J. S., Chowdhury, N., Yamasaki, R. & Wood, T. K. Viable but non-culturable and persistence describe the same bacterial stress state. *Environmental Microbiology* (2018). doi:10.1111/1462-2920.14075

54.   Bamford, R. A. *et al.* Investigating the physiology of viable but non-culturable bacteria by microfluidics and time-lapse microscopy. *BMC Biol.* **15,** (2017).

55.   Asakura, H. *et al.* Differential expression of the outer membrane protein W (OmpW) stress response in enterohemorrhagic *Escherichia coli* O157:H7 corresponds to the viable but non-culturable state. *Res. Microbiol.* **159,** 709–717 (2008).

56.   Zhang, Y., Yew, W. W. & Barer, M. R. Targeting persisters for tuberculosis control. *Antimicrobial Agents and Chemotherapy* **56,** 2223–2230 (2012).

57. Stricker, R. B. & Johnson, L. The pain of chronic Lyme disease: moving the discourse backward? *FASEB J.* **25,** 4085–7 (2011).

58. Levin-Reisman, I. *et al.* Antibiotic tolerance facilitates the evolution of resistance. *Science (80-.).* **355,** 826–830 (2017).

59. Shulman, S. T., Friedmann, H. C. & Sims, R. H. Theodor Escherich: The First Pediatric Infectious Diseases Physician? *Clin. Infect. Dis.* **45,** 1025–1029 (2007).

60. S., F. A., Buchanan, R. E. T., Gibbons, N. E. & Bergey. Bergey's Manual of Determinative Bacteriology. *Taxon* **24,** 377 (1975).

61. Ingraham, J. L. and A. G. M. in *Escherichia coli and Salmonella typhimurium* (ed. Neidhardt, F. C.) 1570–1572 (ASM Press, 1987).

62. Sperandio, V. & Hovde, C. J. Enterohemorrhagic *Escherichia coli* and other Shiga toxin-producing *E. coli*. (2015).

63. Crick, F. H. C., Barnett, L., Brenner, S. & Watts-Tobin, R. J. General nature of the genetic code for proteins. *Nature* **192,** 1227–1232 (1961).

64. Stevens, A. Incorporation of the adenine ribonucleotide into RNA by cell fractions from *E. coli* B. *Biochem. Biophys. Res. Commun.* **3,** 92–96 (1960).

65. Bessman, M. J., Kornberg, A. & Lehman, I. R. Enzymatic synthesis of deoxyribonucleic acid. *J. Biol. Chem.* (1956). at <http://lehman.stanford.edu/pdfs/4. A. Enzymatic Sythesis of Deoxyribonucleic Acid. II. General Properties of the Reaction..pdf>

66. Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3,** 318–356 (1961).

67. Englesberg, E., Irr, J., Power, J. & Lee, N. Positive control of enzyme synthesis by gene C in the L-arabinose system. *J. Bacteriol.* **90,** 946–957 (1965).

68. Meselson, M. & Yuan, R. DNA restriction enzyme from *E. coli*. *Nature* **217,** 1110–1114 (1968).

69. Ellis, E. L. & Delbruck, M. The growth of bacteriophage. *J. Gen. Physiol.* **22,** 365–384 (1939).

70. Lwoff, A. Lysogeny. *Bacteriol. Rev.* **17,** 269–337 (1953).

71. Lederberg, J. in *Genetics in the 20th Century* (ed. Dunn, L. C.) 263–289 (Macmillan, 1951).

72. Blattner, F. R. The Complete Genome Sequence of *Escherichia coli* K-12. *Science (80-.).* **277,** 1453–1462 (1997).

73. Yamamoto, Y. *et al.* Construction of a contiguous 874-kb sequence of the *Escherichia coli* K-12 genome corresponding to the 50.0-68.8 min on the linkage map and analysis of its sequence features (supplement). *DNA Res.* **4,** 169–178 (1997).

74. Riley, M. *et al. Escherichia coli* K-12: A cooperatively developed annotation snapshot - 2005. *Nucleic Acids Res.* **34,** 1–9 (2006).

75. Berlyn, M. K. B. Linkage map of *Escherichia coli* K-12, edition 10: the physical map. *Microbiol. Mol. Biol. Rev.* **62,** 985–1019 (1998).

76. Tatusov, R. L., Koonin, E. V. & Lipman, D. J. A genomic perspective on protein families. *Science (80-.).* **278,** 631–637 (1997).

77. Hu, P. *et al.* Global functional atlas of *Escherichia coli* encompassing previously

uncharacterized proteins. *PLoS Biol.* **7,** 0929–0947 (2009).

78. Tamae, C. *et al.* Determination of antibiotic hypersensitivity among 4,000 single-gene-knockout mutants of *Escherichia coli*. *J. Bacteriol.* **190,** 5981–5988 (2008).

79. Liu, A. *et al.* Antibiotic sensitivity profiles determined with an *Escherichia coli* gene knockout collection: Generating an antibiotic bar code. *Antimicrob. Agents Chemother.* **54,** 1393–1403 (2010).

80. Inoue, T. *et al.* Genome-wide screening of genes required for swarming motility in *Escherichia coli* K-12. *J. Bacteriol.* **189,** 950–957 (2007).

81. Niba, E. T. E., Naka, Y., Nagase, M., Mori, H. & Kitakawa, M. A genome-wide approach to identify the genes involved in biofilm formation in *E. coli*. *DNA Res.* **14,** 237–46 (2007).

82. Wiriyathanawudhiwong, N., Ohtsu, I., Li, Z. Di, Mori, H. & Takagi, H. The outer membrane TolC is involved in cysteine tolerance and overproduction in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **81,** 903–913 (2009).

83. Nakayashiki, T. & Mori, H. Genome-Wide screening with hydroxyurea reveals a link between nonessential ribosomal proteins and reactive oxygen species production. *J. Bacteriol.* **195,** 1226–1235 (2013).

84. Kitagawa, M. *et al.* Complete set of ORF clones of *Escherichia coli* ASKA library (A complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* **12,** 291–299 (2005).

85. Rajagopala, S. *et al.* The *Escherichia coli* K-12 ORFeome: a resource for comparative molecular microbiology. *BMC Genomics* **11,** 470 (2010).

86. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6640–5 (2000).

87. Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2,** 2006.0008 (2006).

88. Otsuka, Y. *et al.* GenoBase: comprehensive resource database of *Escherichia coli* K-12. *Nucleic Acids Res.* **43,** D606–D617 (2014).

89. Winzeler, E. A. *et al.* Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science (80-.).* **285,** 901–906 (1999).

90. Giaever, G. *et al.* Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418,** 387–391 (2002).

91. Ni, L. & Snyder, M. A genomic study of the bipolar bud site selection pattern in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **12,** 2147–2170 (2001).

92. Pagé, N. *et al.* A *Saccharomyces cerevisiae* genome-wide mutant screen for altered sensitivity to K1 killer toxin. *Genetics* **163,** 875–894 (2003).

93. Dimmer, K. S. *et al.* Genetic basis of mitochondrial function and morphology in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **13,** 847–53 (2002).

94. Merz, S. & Westermann, B. Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in *Saccharomyces cerevisiae*. *Genome Biol.* **10,** R95 (2009).

95. Han, T. X., Xu, X. Y., Zhang, M. J., Peng, X. & Du, L. L. Global fitness profiling of fission yeast deletion strains by barcode sequencing. *Genome Biol.* **11,** (2010).

96. Alonso, Stepanova, Leisse & Al, E. Genome-Wide Insertional Mutagenesis of *Arabidopsis thaliana. Science (80-.).* **301,** 653 (2003).

97. Jacobs, M. A. *et al.* Comprehensive transposon mutant library of *Pseudomonas aeruginosa. Proc. Natl. Acad. Sci.* **100,** 14339–14344 (2003).

98. Koo, B. M. *et al.* Construction and Analysis of Two Genome-Scale Deletion Libraries for *Bacillus subtilis. Cell Syst.* **4,** 291–305.e7 (2017).

99. Zhu, H. *et al.* Global analysis of protein activities using proteome chips. *Science (80-.).* **293,** 2101–2105 (2001).

100. Tong, A. H. Y. *et al.* Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science (80-.).* **294,** 2364–2368 (2001).

101. Hughes, T. R. *et al.* Functional discovery via a compendium of expression profiles. *Cell* **102,** 109–126 (2000).

102. Smith, A. M. *et al.* Quantitative phenotyping via deep barcode sequencing. *Genome Res.* **19,** 1836–1842 (2009).

103. Smith, A. M. *et al.* Highly-multiplexed barcode sequencing: An efficient method for parallel analysis of pooled samples. *Nucleic Acids Res.* **38,** (2010).

104. Payen, C. *et al.* High-Throughput Identification of Adaptive Mutations in Experimentally Evolved Yeast Populations. *PLoS Genet.* **12,** (2016).

105. Qian, W., Ma, D., Xiao, C., Wang, Z. & Zhang, J. The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast. *Cell Rep.* **2,** 1399–1410 (2012).

106. Acton, E. *et al.* Comparative functional genomic screens of three yeast deletion collections reveal unexpected effects of genotype in response to diverse stress. *Open Biol.* **7,** (2017).

107. Sambrook, J. & Russell, D. W. Molecular Cloning - Sambrook & Russel - Vol. 1, 2, 3. *Cold Spring Harb. Lab. Press* **3th Editio,** (2001).

108. Andrews, J. M. & Andrews, J. M. Determination of minimum inhibitory concentrations. *J. Antimicrob. Chemother.* **48 Suppl 1,** 5–16 (2001).

109. Ausubel, F. M. *et al. Current Protocols in Molecular Biology. Molecular Biology* **1,** (2003).

110. Stancik, L. M. *et al.* pH-dependent expression of periplasmic proteins and amino acid catabolism in *Escherichia coli. J. Bacteriol.* **184,** 4246–4258 (2002).

111. Farrell, M. J. & Finkel, S. E. The Growth Advantage in Stationary-Phase Phenotype Conferred by *rpoS* Mutations Is Dependent on the pH and Nutrient Environment. *J. Bacteriol.* **185,** 7044–7052 (2003).

112. Pepper, E. D., Farrell, M. J., Nord, G. & Finkel, S. E. Antiglycation effects of carnosine and other compounds on the long-term survival of *Escherichia coli. Appl. Environ. Microbiol.* **76,** 7925–7930 (2010).

113. Kram, K. E. & Finkel, S. E. Rich Medium Composition Affects *Escherichia coli* Survival, Glycation, and Mutation Frequency during Long-Term Batch Culture. *Appl.*

Environ. Microbiol. **81,** 4442–50 (2015).

114. Nyström, T. Stationary-Phase Physiology. *Annu. Rev. Microbiol.* **58,** 161–181 (2004).

115. Sherwin, W. B., Jabot, F., Rush, R. & Rossetto, M. Measurement of biological information with applications from genes to landscapes. *Mol. Ecol.* **15,** 2857–2869 (2006).

116. Fridman, O., Goldberg, A., Ronin, I., Shoresh, N. & Balaban, N. Q. Optimization of lag time underlies antibiotic tolerance in evolved bacterial populations. *Nature* **513,** 1–9 (2014).

117. Takeuchi, R. *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* **14,** 171 (2014).

118. Kram, K. E. *et al.* Adaptation of *Escherichia coli* to Long-Term Serial Passage in Complex Medium: Evidence of Parallel Evolution. *mSystems* **2,** e00192-16 (2017).

119. Delhaye, A., Collet, J. F. & Laloux, G. Fine-tuning of the Cpx envelope stress response is required for cell wall homeostasis in *Escherichia coli*. *MBio* **7,** (2016).

120. Bacun-Druzina, V., Cagalj, Z. & Gjuracic, K. The growth advantage in stationary-phase (GASP) phenomenon in mixed cultures of enterobacteria. *FEMS Microbiol. Lett.* **266,** 119–127 (2007).

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

122. Yang, H., Wolff, E., Kim, M., Diep, A. & Miller, J. H. Identification of mutator genes and mutational pathways in *Escherichia coli* using a multicopy cloning approach. *Mol. Microbiol.* **53,** 283–295 (2004).

123. Engelberg-Kulka, H. & Kolodkin-Gal, I. The stationary-phase sigma factor S is responsible for the resistance of *Escherichia coli* stationary-phase cells to mazEF-Mediated cell death. *J. Bacteriol.* **191,** 3177–3182 (2009).

124. Dukan, S. & Nyström, T. Bacterial senescence: Stasis results in increased and differential oxidation of cytoplasmic proteins leading to developmental induction of the heat shock regulon. *Genes Dev.* **12,** 3431–3441 (1998).

125. Cho, B. K., Kim, D., Knight, E. M., Zengler, K. & Palsson, B. O. Genome-scale reconstruction of the sigma factor network in *Escherichia coli*: Topology and functional states. *BMC Biol.* **12,** (2014).

126. Mauri, M. & Klumpp, S. A Model for Sigma Factor Competition in Bacterial Cells. *PLoS Comput. Biol.* **10,** (2014).

127. Rath, D. & Jawali, N. Loss of expression of *cspC*, a cold shock family gene, confers a gain of fitness in *Escherichia coli* K-12 strains. *J. Bacteriol.* **188,** 6780–6785 (2006).

128. Phadtare, S. & Inouye, M. Role of CspC and CspE in regulation of expression of RpoS and UspA, the stress response proteins in *Escherichia coli*. *J. Bacteriol.* **183,** 1205–1214 (2001).

129. Phadtare, S., Tadigotla, V., Shin, W. H., Sengupta, A. & Severinov, K. Analysis of *Escherichia coli* global gene expression profiles in response to overexpression and

deletion of CspC and CspE. *J. Bacteriol.* **188,** 2521–2527 (2006).

130. Hansen, A. M. *et al.* Structural basis for the function of stringent starvation protein A as a transcription factor. *J. Biol. Chem.* **280,** 17380–17391 (2005).

131. Williams, M. D., Fuchs, J. A. & Flickinger, M. C. Null mutation in the stringent starvation protein of *Escherichia coli* disrupts lytic development of bacteriophage P1. *Gene* **109,** 21–30 (1991).

132. Goodrich-Blair, H. & Kolter, R. Homocysteine thiolactone is a positive effector of σ(S) levels in *Escherichia coli*. *FEMS Microbiol. Lett.* **185,** 117–121 (2000).

133. Sezonov, G., Joseleau-Petit, D. & D'Ari, R. *Escherichia coli* physiology in Luria-Bertani broth. *J. Bacteriol.* **189,** 8746–8749 (2007).

134. Eraso, J. M. *et al.* The highly conserved MraZ protein is a transcriptional regulator in *Escherichia coli*. *J. Bacteriol.* **196,** 2053–2066 (2014).

135. Szklarczyk, D. *et al.* The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* **39,** (2011).

136. Ruiz, N., Peterson, C. N. & Silhavy, T. J. RpoS-dependent transcriptional control of sprE: Regulatory feedback loop. *J. Bacteriol.* **183,** 5974–5981 (2001).

137. Aguilar, C. *et al.* Genetic changes during a laboratory adaptive evolution process that allowed fast growth in glucose to an *Escherichia coli* strain lacking the major glucose transport system. *BMC Genomics* **13,** 385 (2012).

138. Shimada, T., Kori, A. & Ishihama, A. Involvement of the ribose operon repressor RbsR in regulation of purine nucleotide synthesis in *Escherichia coli*. *FEMS Microbiol. Lett.* **344,** 159–165 (2013).

139. Ganduri, Y. L., Sadda, S. R., Datta, M. W., Jambukeswaran, R. K. & Datta, P. TdcA, a transcriptional activator of the tdcABC operon of *Escherichia coli*, is a member of the LysR family of proteins. *MGG Mol. Gen. Genet.* **240,** 395–402 (1993).

140. Kim, M. J., Lim, S. & Ryu, S. Molecular analysis of the *Salmonella typhimurium tdc* operon regulation. *J. Microbiol. Biotechnol.* **18,** 1024–1032 (2008).

141. Sperandeo, P., Pozzi, C., Dehò, G. & Polissi, A. Non-essential KDO biosynthesis and new essential cell envelope biogenesis genes in the *Escherichia coli* yrbG-yhbG locus. *Res. Microbiol.* **157,** 547–558 (2006).

142. Flatley, J. *et al.* Transcriptional responses of *Escherichia coli* to S-nitrosoglutathione under defined chemostat conditions reveal major changes in methionine biosynthesis. *J. Biol. Chem.* **280,** 10065–10072 (2005).

143. Nobre, L. S., Al-Shahrour, F., Dopazo, J. & Saraiva, L. M. Exploring the antimicrobial action of a carbon monoxide-releasing compound through whole-genome transcription profiling of *Escherichia coli*. *Microbiology* **155,** 813–824 (2009).

144. Park, J. H., Lee, K. H., Kim, T. Y. & Lee, S. Y. Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proc. Natl. Acad. Sci.* **104,** 7797–7802 (2007).

145. Lomovskaya, O. & Lewis, K. Emr, an *Escherichia coli* locus for multidrug resistance. *Proc. Natl. Acad. Sci. U. S. A.* **89,** 8938–42 (1992).

146. Kakkanat, A., Phan, M.-D., Lo, A. W., Beatson, S. A. & Schembri, M. A. Novel genes

associated with enhanced motility of *Escherichia coli* ST131. *PLoS One* **12,** e0176290 (2017).

147. Wu, Y. & Outten, F. W. IscR controls iron-dependent biofilm formation in *Escherichia coli* by regulating type I fimbria expression. *J. Bacteriol.* **191,** 1248–1257 (2009).

148. Henkel, S. G. *et al.* Basic regulatory principles of *Escherichia coli*'s electron transport chain for varying oxygen conditions. *PLoS One* **9,** (2014).

149. Unden, G. & Bongaerts, J. Alternative respiratory pathways of *Escherichia coli*: Energetics and transcriptional regulation in response to electron acceptors. *Biochimica et Biophysica Acta - Bioenergetics* **1320,** 217–234 (1997).

150. Erhardt, H. *et al.* Disruption of individual nuo-genes leads to the formation of partially assembled NADH:ubiquinone oxidoreductase (complex I) in *Escherichia coli*. *Biochim. Biophys. Acta - Bioenerg.* **1817,** 863–871 (2012).

151. Jensen, P. R. & Michelsen, O. Carbon and energy metabolism of *atp* mutants of *Escherichia coli*. *J. Bacteriol.* **174,** 7635–7641 (1992).

152. Rice, C. W. & Hempfling, W. P. Oxygen-limited continuous culture and respiratory energy conservation in *Escherichia coli*. *J. Bacteriol.* **134,** 115–24 (1978).

153. Mempin, R. *et al.* Release of extracellular ATP by bacteria during growth. *BMC Microbiol.* **13,** (2013).

154. Kiley, P. J. & Beinert, H. The role of Fe-S proteins in sensing and regulation in bacteria. *Current Opinion in Microbiology* **6,** 181–185 (2003).

155. Fontecave, M. Iron-sulfur clusters: Ever-expanding roles. *Nat. Chem. Biol.* **2,** 171–174 (2006).

156. Lill, R. Function and biogenesis of iron-sulphur proteins. *Nature* **460,** 831–838 (2009).

157. Py, B., Moreau, P. L. & Barras, F. Fe-S clusters, fragile sentinels of the cell. *Current Opinion in Microbiology* **14,** 218–223 (2011).

158. Takahashi, Y. & Nakamura, M. Functional assignment of the ORF2-iscS-iscU-iscA-hscB-hscA-fdx-ORF3 gene cluster involved in the assembly of Fe-S clusters in *Escherichia coli*. *J. Biochem.* **126,** 917–926 (1999).

159. Schwartz, C. J., Djaman, O., Imlay, J. A. & Kiley, P. J. The cysteine desulfurase, IscS, has a major role in in vivo Fe-S cluster formation in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **97,** 9009–9014 (2000).

160. Takahashi, Y. & Tokumoto, U. A third bacterial system for the assembly of iron-sulfur clusters with homologs in Archaea and plastids. *J. Biol. Chem.* **277,** 28380–28383 (2002).

161. Tokumoto, U. & Takahashi, Y. Genetic analysis of the isc operon in *Escherichia coli* involved in the biogenesis of cellular iron-sulfur proteins. *J. Biochem.* **130,** 63–71 (2001).

162. Kammler, M., Schön, C. & Hantke, K. Characterization of the ferrous iron uptake system of *Escherichia coli*. *J. Bacteriol.* **175,** 6212–6219 (1993).

163. Maisonneuve, E. & Gerdes, K. Molecular mechanisms underlying bacterial persisters. *Cell* **157,** 539–548 (2014).

164. Linkevicius, M., Sandegren, L. & Andersson, D. I. Mechanisms and fitness costs of

tigecycline resistance in *Escherichia coli*. 2809–2819 (2013). doi:10.1093/jac/dkt263

165. Vinué, L., Corcoran, M. A., Hooper, D. C. & Jacoby, G. A. Mutations that enhance the ciprofloxacin resistance of *Escherichia coli* with qnrA1. *Antimicrob. Agents Chemother.* **60,** 1537–1545 (2016).

166. Huguet, A., Pensec, J. & Soumet, C. Resistance in *Escherichia coli*: Variable contribution of efflux pumps with respect to different fluoroquinolones. *J. Appl. Microbiol.* **114,** 1294–1299 (2013).

167. Cohen, B. E. Functional linkage between genes that regulate osmotic stress responses and multidrug resistance transporters: Challenges and opportunities for antibiotic discovery. *Antimicrob. Agents Chemother.* **58,** 640–646 (2014).

168. Batchelor, E., Walthers, D., Kenney, L. J. & Goulian, M. The *Escherichia coli* CpxA-CpxR envelope stress response system regulates expression of the porins OmpF and OmpC. *J. Bacteriol.* **187,** 5723–5731 (2005).

169. Schnaitman, C. a & Klena, J. D. Genetics of lipopolysaccharide biosynthesis in enteric bacteria. *Microbiol. Rev.* **57,** 655–682 (1993).

170. Korch, S. B. & Hill, T. M. Ectopic overexpression of wild-type and mutant *hipA* genes in *Escherichia coli*: Effects on macromolecular synthesis and persister formation. *J. Bacteriol.* **188,** 3826–3836 (2006).

171. Demidenok, O. I., Kaprelyants, A. S. & Goncharenko, A. V. Toxin-antitoxin *vapBC* locus participates in formation of the dormant state in *Mycobacterium smegmatis*. *FEMS Microbiology Letters* **352,** 69–77 (2014).

172. Cohen, N. R., Lobritz, M. A. & Collins, J. J. Microbial persistence and the road to drug resistance. *Cell Host and Microbe* **13,** 632–642 (2013).

173. Chopra, N., Saumitra, Pathak, A., Bhatnagar, R. & Bhatnagar, S. Linkage, mobility, and selfishness in the MazF family of bacterial toxins: A snapshot of bacterial evolution. *Genome Biol. Evol.* **5,** 2268–2284 (2013).

174. Neumann, G. F. & Jetschke, G. BioSystems Evolutionary classification of toxin mediated interactions in microorganisms. *BioSystems* **99,** 155–166 (2010).

175. Schmidt, A. Bacterial persistence is an active σ S stress response to metabolic flux limitation Bacterial persistence is an active σ S stress response to metabolic flux limitation. *Mol. Syst. Biol.* **12,** 882 (2016).

Supplementary Information

Supplementary Table 1 List of primers for Illumina sequencing

Primer	Sequence			
	|------------------ PE-1.0 ------------------|	N*5	Index	PS16
PE-1.0_tag01_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	CAAGTG TTC	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag02_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	AGGACA TTC	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag03_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	CACTAA TGG	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag04_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	AGCCTG ATG	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag05_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	TTACGCT AA	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag06_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	ACTCTCC GT	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag07_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	GTCGAT GCA	GAATCTTCGGTAGT CCAGCG
PE-1.0_tag08_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA CGCTCTTCCGATCT	NNN NN	ACGGGA ATT	GAATCTTCGGTAGT CCAGCG

			Index	PS1
PE-1.0_tag09_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	NNNNN	CGCGCCCAG	GAATCTTCGGTAGTCCAGCG
PE-1.0_tag10_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	NNNNN	ACTAGTTTG	GAATCTTCGGTAGTCCAGCG
PE-1.0_tag11_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	NNNNN	AGTATTACA	GAATCTTCGGTAGTCCAGCG
PE-1.0_tag12_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	NNNNN	AGGTTGGGT	GAATCTTCGGTAGTCCAGCG
PE-1.0_tag13_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	NNNNN	GTGAACCGA	GAATCTTCGGTAGTCCAGCG
PE-1.0_tag14_PS16	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT	NNNNN	GCACAAAAC	GAATCTTCGGTAGTCCAGCG
	|---------------- PE-2.0 ----------------|	N*5	Index	PS1
PE-2.0_tag15_PS1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGGCTCTTCCGATCT	NNNNN	CTGTCTTCG	TGTAGGCTGGAGCTGCTTCG
PE-2.0_tag16_PS1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGGCTCTTCCGATCT	NNNNN	GACGCGACT	TGTAGGCTGGAGCTGCTTCG
PE-2.0_tag17_PS1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGGCTCTTCCGATCT	NNNNN	CAGCCCATA	TGTAGGCTGGAGCTGCTTCG
PE-2.0_tag18_PS1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGGCTCTTCCGATCT	NNNNN	AGATATCTG	TGTAGGCTGGAGCTGCTTCG

PE-2.0_tag19_PS 1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTG AACCGCTCTTCCGATCT	NNN NN	AATACG CAC	TGTAGGCTGGAGCT GCTTCG
PE-2.0_tag20_PS 1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTG AACCGCTCTTCCGATCT	NNN NN	TATCGTG CC	TGTAGGCTGGAGCT GCTTCG
PE-2.0_tag21_PS 1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTG AACCGCTCTTCCGATCT	NNN NN	TCCTGGT AT	TGTAGGCTGGAGCT GCTTCG
PE-2.0_tag22_PS 1	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTG AACCGCTCTTCCGATCT	NNN NN	GGCAGA GGA	TGTAGGCTGGAGCT GCTTCG

Supplementary Table 2 MIC (IC50) concentration for control strain with/ without chloramphenicol resistance (raw data)

| Drug	CIP		KAN		GEN		NAL		STR		TET		TMP	
Experiment	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm
Exp1	0.073	0.09	5.96	6.14	1.678	1.766	19.42	28.586	3.2	5.17	3.148	3.336	0.542	0.692
Exp2	0.15	0.16	6.68	7.08	1.65	1.28	14.616	15.433	2.823	3.474	1.15	0.85	0.5	0.32
Exp3									3.94	4.97	1.57	0.90	0.33	0.38
	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm	+Cm	-Cm
Mean	0.112	0.125	6.320	6.610	1.664	1.523	17.018	22.010	3.321	4.538	1.955	1.694	0.458	0.465
Variance	0.003	0.002	0.259	0.442	0.000	0.118	11.539	86.501	0.323	0.859	1.109	2.023	0.012	0.039
Observations	2.000	2.000	2.000	2.000	2.000	2.000	2.000	2.000	3.000	3.000	3.000	3.000	3.000	3.000
Pearson Correlation	1.000		1.000		1.000		1.000		0.684		0.984		0.545	
Hypothesized Mean Difference	0		0		0		0		0		0		0	
df	1		1		1		1		2		2		2	
t Stat	-3.857		-2.636		0.616		-1.196		-3.104		1.053		-0.073	
P(T<=t) one-tail	0.081		0.115		0.324		0.222		0.045		0.201		0.474	
t Critical one-tail	6.314		6.314		6.314		6.314		2.920		2.920		2.920	
P(T<=t) two-tail	0.161		0.231		0.649		0.443		0.090		0.403		0.949	
t Critical two-tail	12.706		12.706		12.706		12.706		4.303		4.303		4.303	

The strain in the ASKA barcode deletion library carries chloramphenicol resistance. Control ($\Delta lacA$) with/ without Cm resistance is used to check for cross-resistance between drugs. $p>0.05$, accept H0: no significant difference between +Cm and -Cm. Each experiment is average of three replicates. At least two independent experiment were performed for each drug.

Table 3 Fitness for control strains (WT)

Competition	0h		24h		Fitness
	ΔlacA::Km	*ΔlacA*::Cm	*ΔlacA*::Km	*ΔlacA*::Cm	(Km/Cm)
Control 1	2060000	2280000	1260000000	1300000000	1.07274085
Control2	2060000	2180000	1370000000	1470000000	0.98626247
Control3	1900000	1930000	1240000000	1300000000	0.96890688

Competition experiment for control strain. To check effect of resistance cassette on mutant fitness. The average fitness is 1.01, thus resistance cassette did not affect the fitness of control strain
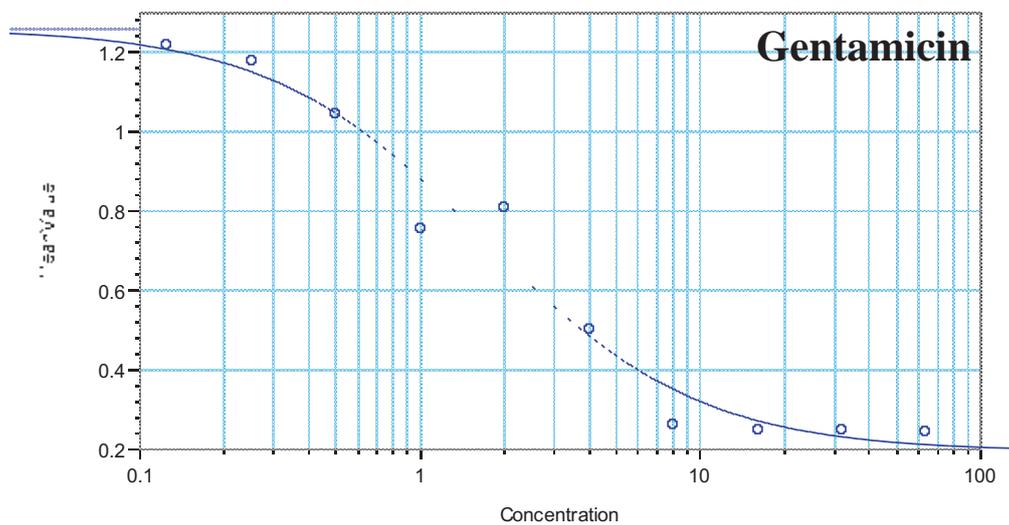
Supplementary Figure 1 Dose-response curve

$y = ((A - D)/(1 + (x/C)^B)) + D:$
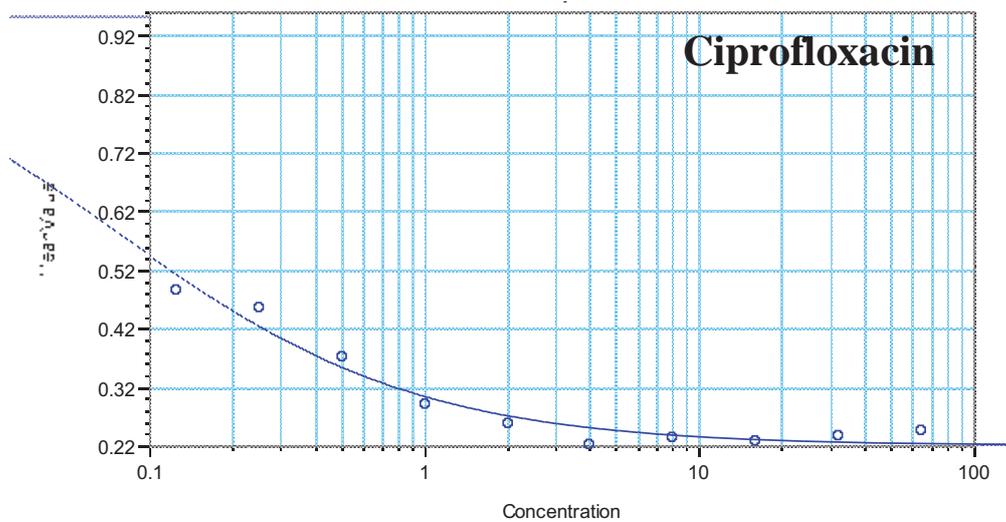
	A	B	C	D	R^2
Std (Standards: Concentration vs MeanValue)	0.955	0.784	0.073	0.221	0.991

$y = ((A - D)/(1 + (x/C)^B)) + D:$

	A	B	C	D	R^2
Std (Standards: Concentration vs MeanValue)	1.262	1.133	1.678	0.195	0.975

Supplementary Figure 1 (cont.)

Kanamycin

y = ((A - D)/(1 + (x/C)^B)) + D:	A	B	C	D	R^2
Std (Standards: Concentration vs MeanValue)	0.952	2.709	5.96	0.21	0.991

Nalidixic acid

y = ((A - D)/(1 + (x/C)^B)) + D:	A	B	C	D	R^2
Std (Standards: Concentration vs MeanValue)	1.043	1.704	14.616	0.082	0.973

Supplementary Figure 1 (cont.)

Streptomycin

y = ((A - D)/(1 + (x/C)^B)) + D:	A	B	C	D	R^2
Std (Standards: Concentration vs MeanValue)	1.116	1.061	3.2	0.225	0.997

Tetracycline

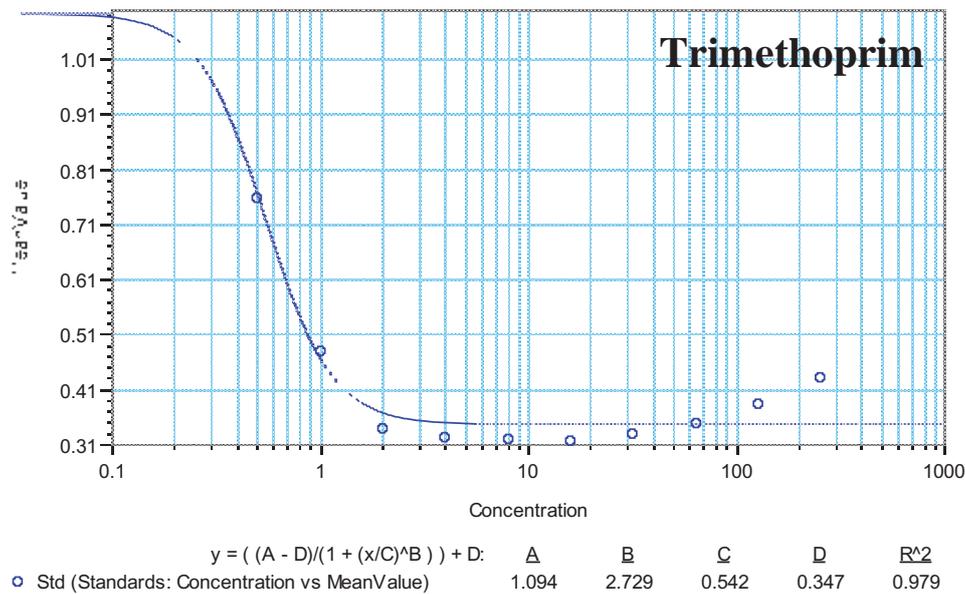| y = ((A - D)/(1 + (x/C)^B)) + D:	A	B	C	D	R^2
Std (Standards: Concentration vs MeanValue)	1.224	1.675	3.148	0.291	0.991
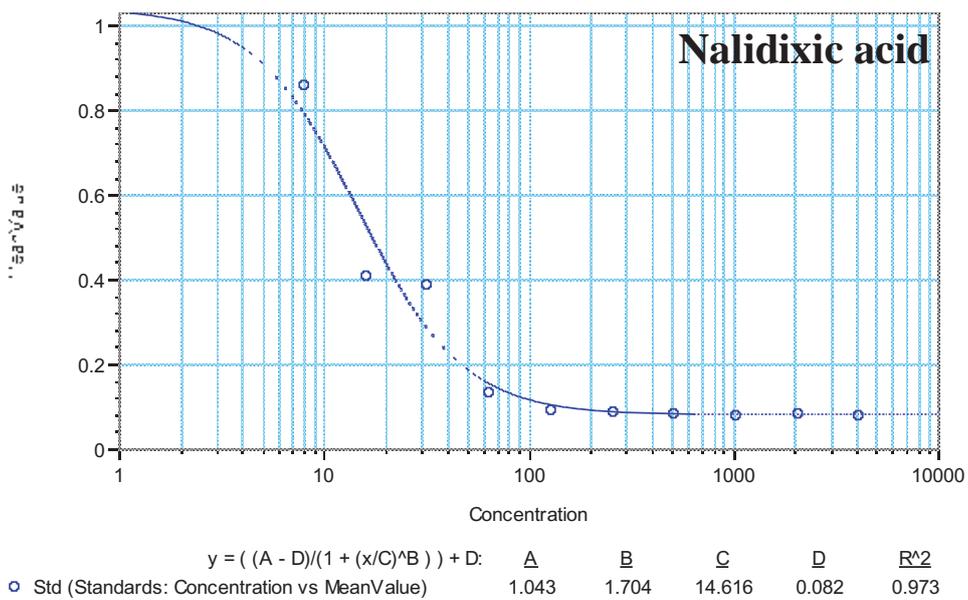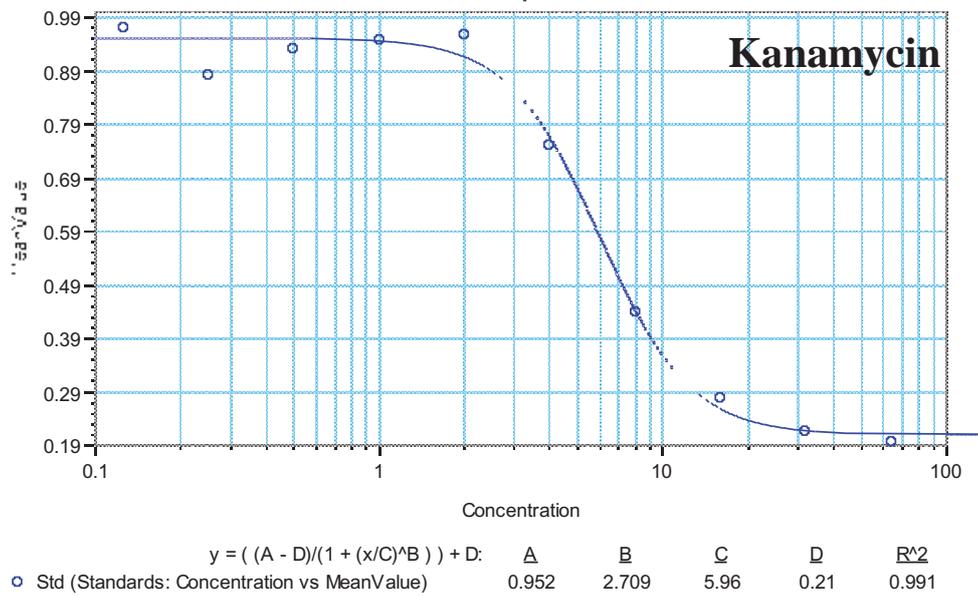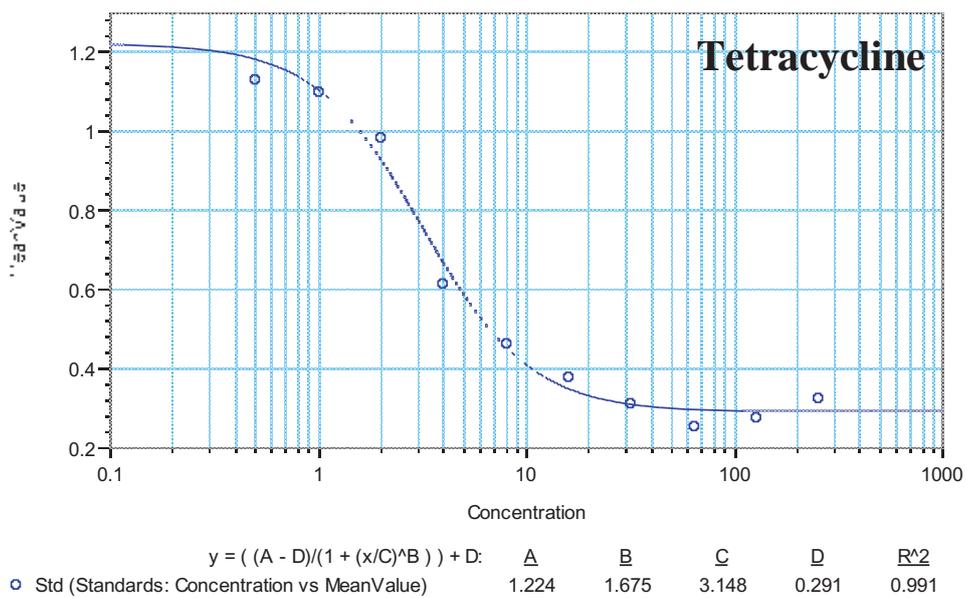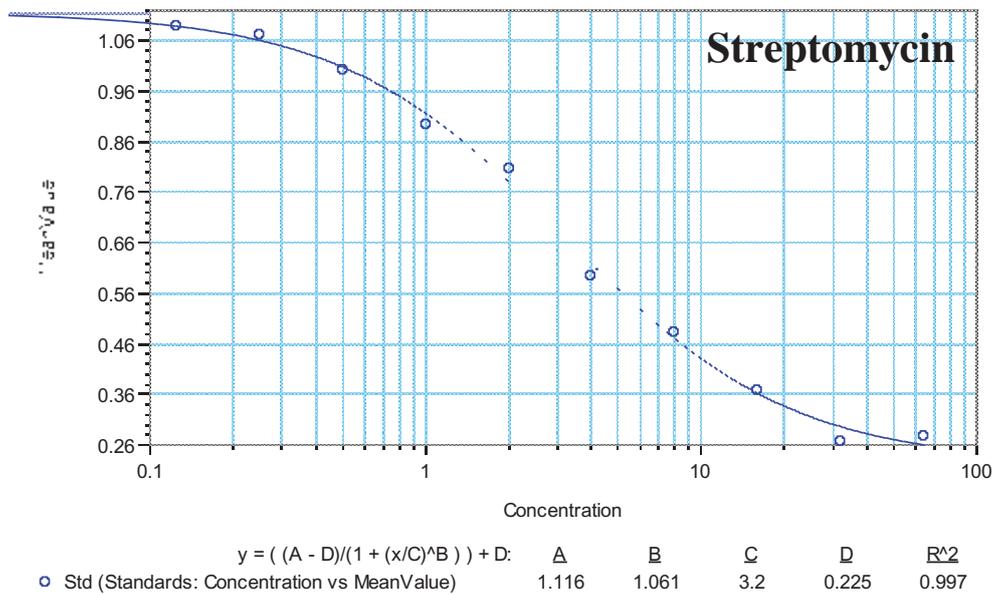
Figure 1 Dose-response curve

Dose response for each drugs; CIP, GEN, KAN, STR, TET, TMP and NAL. The *x*-axis represents the drug concentration ($\log_{10}[\mu g/ml]$) and the *y*-axis represents the mean OD value.

From the nonlinear regression equation; $y=\left(\dfrac{A-D}{1+\left(\frac{x}{C}\right)^{B}}\right) + D$, A and D represents the maximum

and minimum value, B represent the Hill slope or steepness of the curve and C is the inflection point or IC50 value. The values of B and C were incorporate into *Eq.* 1 to calculate the drug concentrations at IC70 and IC90. WT ($\Delta lacA$::Cm strain) was used to determine the MIC concentration of each drugs used in this study. Values are mean of three replicates. Each experiment were repeated at least twice using WT strain with/ without chloramphenicol resistance (Cm^{R}). No significant difference ($p > 0.05$) were observed between strains with/without Cm^{R}.

Supplementary Figure 2 Growth profile of BW38028 and *ΔlacA* mutant
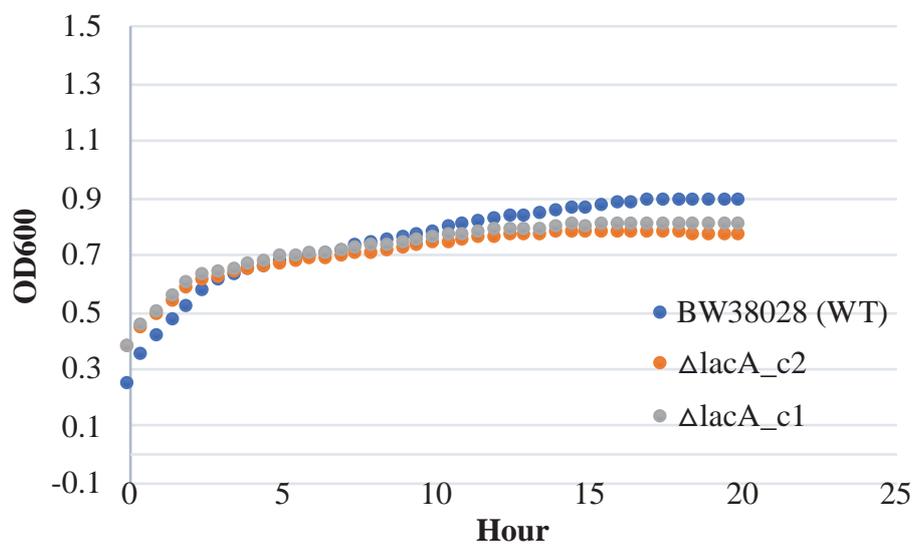
Growth profile of BW38028 and *ΔlacA* mutant

The WT and control strain showed similar growth profile and the specific growth rates (h-1) were 0.1086, 0.1123 and 0.1009, respectively. Value were average of three replicates.