Development of a novel system for the genetic interaction analysis of essential genes in *E. coli* K-12

Yong Han Tek

Nara Institute of Science and Technology

Graduate School of Biological Science

Systems Microbiology (Professor Hirotada Mori)

2013/06/03

Graduate School of Biological Sciences　　Doctoral Thesis Abstract

Lab name (Supervisor)	Systems Microbiology Hirotada Mori		
Name (surname) (given name)	Yong Han Tek	Date	(13/04/23)
Title	Development of a novel system for the genetic interaction analysis of essential genes in *E. coli* K-12		

Abstract

　　　Traditionally, genetic interactions are investigated by randomly combining pairs of mutations and accessing the unexpected combinatorial phenotype (e.g synthetic sickness or synthetic lethality) which is not exhibited by either mutation alone. In the past 10 years, systematic and unbiased genetic interaction analysis has become feasible after the completion of genome sequencing projects for hundreds of organisms and the availability of single gene deletion libraries. To date, there are several technologies have been developed to study genetic interactions comprehensively in different model organisms. Nevertheless, most studies only focused on non-essential genes. The reason of lagging progress in genetic interaction analysis involving essential genes is the inherent difficulty in creating and manipulating those mutants.

　　　Here I discuss a novel system based on the Hfr conjugation gene transfer system for the genetic interaction analysis of essential genes in *E.coli* K-12. The system is facilitated by a 2-in-1 complementing plasmid, pFE604T to achieve (i) systematic construction of essential gene knockdown mutant and (ii) crossing of target essential gene knockdown mutant with non-essential deletions en masse. The expression of essential gene in pFE604T is regulated by T5 promoter and *lacI^q* repressor system under the induction of IPTG. Trans-supply of low level of essential protein from pFE604T permits the removal of essential gene from the chromosome, thus creating an essential gene knockdown mutant. The *oriT* presents in the plasmid offers the convenience of one-step transfer to an array of recipient cells (Keio collection) in order to complement the essential gene deletion in the double mutation background.

　　　Several considerations regarding on the Hfr conjugation gene transfer system to construct double mutant were carefully evaluated, including the colony deposition methods (spotting or stamping), the conjugation time for the plasmid transfer and the essential gene deletion transfer, the

intermediate selection condition and the incubation time, and the IPTG concentrations for the second selection. Colony quantification methods, genetic interaction scoring and statistical analysis were also carefully addressed.

Employing such system, the pilot test of genetic interaction analysis of 5 target essential genes with varying functions was performed. They were *dnaN* which is a beta subunit of DNA polymerase, *ftsW* which is involved in cell division, *trmD* for tRNA methylation, *yjgP* as a LPS transporter and *yrfF* which is functionally unknown. Genetic interactions for each target gene were scored in different IPTG concentrations. To minimize the false positive results, only gene specific interacting candidates were included. Interacting candidates were then classified into positive or negative interaction. Functional enrichment analysis revealed that *dnaN* has positive interaction with peptidoglycan based cell wall genes and organelle envelope genes and negative interaction with genes involved in homologous recombination, pyrimidine metabolism and DNA repair. While *ftsW* negatively interacts with genes involved in oxidative phosphorylation, organelle envelope, quionone and TCA cycle genes. *trmD* shows negative interaction with genes involved in oxidative phosphorylation, organelle envelope and purine nucleotide biosynthetic process. Iron sulfur protein genes and lipopolysaccharide core region process genes interact with *yjgP* in positive and negative manner respectively. Lastly, *yrfF*, an unknown gene was found to have positive interaction with organelle membrane genes and negative interaction with organelle envelope and RNA degradation genes.

The development of the system marks the beginning for the systematic construction of essential gene knockdown mutant and genetic interaction analysis of essential genes in *E.coli* K-12. Accumulation of genetic interaction data for many combinations of essential genes and non essential genes can help to reconstruct a functional interaction network to unveil the physiological architecture of a complex system.

CONTENTS

CHAPTER 3: Establishment of genetic interaction analysis by spotting method

CHAPTER 4: Establishment of genetic interaction analysis by stamping method and pilot test of five essential query genes

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

General introduction

1.1 Systems biology

Life is a system; it is attributed from the dynamic interplay of every single individual constituent in living organism. Genes and proteins cooperate and respond to each other to perform biochemical pathways, interrelations between pathways and the intrinsic robustness feature to cope with the environmental fluctuations; all these work together in a highly structured but incredibly complex ways to give rise to a whole system (Kitano, 2002). Thus, living organism is not easily discernible if we try to investigate the fundamental constituents in isolation without considering the system-level properties (Aitchison & Schwikowski, 2002). For the past 2 centuries, the revolutionary biological research was driven by the reductionist approach, which individual constituents were analyzed in order to study a biological system (Williams, 1997). Undoubtedly, the approach has successfully revealed many chemical basis of numerous living processes, but this alone is not enough to justify the complexity of whole organism (Gershom, 2011). In a similar way, this is best illustrated in the John Godfrey Saxe's poem, "The Blind Men and the Elephant". Each blind men describes the elephant differently, because each of them assumes elephant is like the part he touched, resulting the real image of the elephant fails to emerge (Aitchison & Schwikowski, 2002).

Recently, systems biology appears as a new strategy for biological research to complement the shortcomings in reductionist approach. In contrary to reductionism, systems biology focuses on all the components of an organism and the interactions among them, all as part of one system (Kitano, 2002). Technical advances in molecular biology have spawned a variety of genomic-scale approaches like the omics techniques

(genomics, proteomics, transcriptomics and metabolomic) for quantitative global profiling. Such broad and detailed information provides us with an unprecedented insight of how the entire system model should behave (Katagiri, 2003). However, the wealth of data generated by systems biology also presents enormous computational and mathematical challenges of interpreting and analyzing (Kitano, 2002). Close collaborations and consortium efforts between systems biologists, reductionist biologists, mathematicians and computer scientists are crucial for the reconstitute of entire systems.

The ultimate goal of systems biology is to create complete quantitative models of biological system, including predictions of the effects of perturbations. In other words, virtual organism, which we can precisely simulate the behavior of the systems *in silico* (Katagiri, 2003). To achieve the goal, it is inevitable to start with simpler model organism, like micro-organisms and also integration of multidisciplinary knowledge. Although we are still far from the goal, the current achievements do represent the first few steps towards creating virtual organism.

1.2 Epistasis and systematic genetic interaction analysis

Epistasis or genetic interaction is defined as unexpected phenotype of two mutations which is not exhibited in either mutation alone (Phillips, 1998), is a powerful tool to map out the functional linkage between genes (Kelley & Ideker, 2005). A quantitative genetic interaction definition has two components, a quantitative phenotypic measure and a neutrality function that predicts the expected phenotype of an organism carrying two non-interacting mutations (Mani et al., 2008). A widely used neutrality function assumes that the combinatorial effect of two non-interacting mutations is multiplicative (Breslow et al., 2008; Elena & Lenski, 1997; Felsenstein, 1965; Jonikas et al., 2009; Mani et al. 2008; Segre et al., 2005; St Onge et al., 2007). A genetic interaction is consequently identified as the extent to which double mutant deviates from the multiplicative expectation. A double mutant with more severe phenotype than expected is

called negative interaction, which can be explained by two genes acting in complementary pathways, so that removal of either gene alone will not bring any detrimental defects, whereas removal of both impairs viability. Positive interaction, in which the double mutant phenotype is less severe than expected, often results when two genes function in concert or in series within the same pathway (Dixon et al., 2009).

Identification of genetic interactions is classically relied on genetic modifier screens, which focus on the identification of a small number of second-site mutations (Dixon et al., 2009). Recent trend has shifted to a more comprehensive manner by systematic reverse genetic approaches for simultaneous detection of thousands of interactions in parallel in prokaryotes and eukaryotes. For example, synthetic genetic array (SGA) analysis (Tong et al., 2001) and diploid synthetic lethality analysis by microarray (dSLAM) (Pan et al., 2007) in *Saccharomyces cerevisiae*, *E. coli* synthetic genetic array (eSGA) analysis (Butland et al., 2008) and GIANT-coli (Typas et al., 2008) in *Escherichia coli*, and RNA interference (RNAi) in *Caenorhabditis elegans* (Byrne et al., 2007) and *Drosophila melanogaster* (Bakal et al., 2008).

Systematic genetic interaction requires a large collection of defined mutant alleles for example, YKO library in *S. cerevisiae* (Winzeler et al., 1999), Keio Collection in *E. coli* (Baba et al., 2006) and RNAi library in *C. elegans* (Fraser et al., 2000) and in *D. melanogaster* (Dietzl et al., 2007). Besides, a single phenotype that can be scored easily in large scale, such as growth rate is desired (Dixon et al., 2009). Currently, the most extensively mapped genetic interaction network species is budding yeast, *S. cerevesiae* (Costanzo et al, 2010). Two general strategies have been applied to establish the genetic interaction mapping, SGA analysis and dSLAM. Both SGA and dSLAM exploit genome-wide yeast knockout deletion collections which are mated or transformed to generate haploid double mutants. As for SGA, the double mutant haploid is produced in an array format on plates which manipulated by robotics. Synthetic sickness or synthetic lethality double mutants are identified by the colony growth phenotype on plates. In dSLAM, barcode tagged heterozygous diploid deletion is used and then microarray profiling is

applied to identify slow growing or non-growing barcode-tagged double mutant through growth competition assay

Systematically crossing of mutant strains is not practical in multicellular organism because comprehensive collections of deletion strains are not available and cumbersome multigenerational mating and selection screens. In metazoans like *C. elegans* and *D. melanogaster*, second mutation is introduced into a query strain that carrying a mutation of interest by gene specific mRNA degradation or RNAi which triggered by the presence of double stranded RNA species. Alternatively, RNAi can be used to inhibit the expression of two genes simultaneously.

eSGA and GIANT-coli in *E. coli* are analogous to those employed in *S.cerevisiae* in which they exploit genome wide deletion collections and special mating procedures to establish comprehensive set of double mutants.

In just about 10 years since the publication of the first large-scale genetic interaction map of *S.cerevisiae* in 2001, high throughput genetic interaction analysis involving various inbred model organisms has provided a model for interpreting the relationship between genotype and phenotype and for decoding biological function. In the future, higher resolution and functionally relevant genetic interaction maps will be expected, bringing us closer to the complete understanding of all genetic interactions relevant to cell function (Dixon et al., 2009).

1.3 Bacterial conjugation

In 1946, Laderberg and Tantum first discovered the event of bacterial conjugation, the exchange of genetic material asexually in bacteria. The finding was elaborated by Bernard Davis, who showed that the physical contact of two strains is required for the genetic transfer. Further work by Hayes led to the discovery that the unidirectional transfer of genes from a donor to a recipient depends on the presence of a fertility factor

(F) in the donor and the absence of the fertility factor in the recipient (Griffiths et al., 2000).

We now know F is a small circular episome that can replicate autonomously or can integrate into chromosome of host. F^+ strains contain the F factor in the cytoplasm and Hfr strains have the F factor integrated into the bacterial chromosome. F^+ and Hfr act as donor during conjugation and can transfer DNA at high efficiency to F^- recipients.

The mechanism of genetic transfer is mediated by *tra* genes encoded by F. F pili, fibrous proteins protruding from the cell walls, promotes cell-to-cell contact. A single strand of F DNA is nicked at *oriT* site by *traYZ* and transferred in the 5' to 3' direction through pores formed in recipient cells. New F DNA strand is synthesized on both donor and recipient cells by DNA polymerase III. (Miller, 1992; Griffiths et al., 2000)

1.4 Essential genes

An essential gene is defined as one whose loss is lethal under certain environmental condition. In other words, essential gene is required for the viability of an organism. Studies pertaining essential genes are drawing a huge interest among biologist because a clear perception of gene essentiality is vital (i) for the understanding of fundamental cellular functions, (ii) for identifying drug target in pathogenic strain and (iii) for revealing evolution between species (Xu et al., 2011).

Identifying essential genes used to be a daunting task. Recently, owing to the advent of genome sequencing and other high throughput experimental technologies, systematically determine genome wide gene essentiality has been feasible. Through genome-wide gene replacement mutagenesis studies, the sets of essential genes of few species have been identifies. In *S. cerevisiae*, ~19 % of the genes are essential (Giaever et al., 2002), ~7 % in *E. coli* (Baba et al., 2006) and *Bacillus subtilis* (Kobayashi et al.,

2003), ~10% in *Spreptococcus sanguinis* (Xu et al., 2011) and 16 % in *Acinetobacter baylyi* (de Berardinis et al, 2008).

There are some discrepancies in determining essential genes even for the same species, largely due to the mutagenesis strategies used (random transposon insertion or deletion), different growth condition (LB or minimal media) or the subjectivity to distinguish essential versus non essential genes, for example slow growth mutant can be misinterpreted as essential. For example, 119 mutants out of 3888 non-essential gene deletion from Keio Collection were reproducibly incapable of growth on glycerol minimal medium, and they were termed as conditionally essential genes (Joyce et al, 2006). In recent studies, it showed that essential gene can be rendered nonessential by over expression of some non-homologous non essential genes. In a smaller number of cases, the essential gene can be fully removed from the genome, compensated by functional replacement of other non-essential gene (Bergmiller et al., 2012). Hence, gene essentiality surveys under various conditions are helpful to the deeper understanding of the complexity of living organism.

1.5 Systematic genetic interaction analysis involving essential genes

Systematic genetic interaction analysis involving essential genes is still not widely reported compared to non-essential genes regardless of species. The nature of essential genes which is required for viability demands more laborious techniques to generate mutants. So far, there are few approaches being published to create high throughput essential gene mutants for genetic interaction analysis. The most commonly known is by temperature sensitive allele, where essential gene mutant is grown under semi permissive condition to partially deactivate the gene function (Davierwala et al., 2005). The use of conditional expression allele, where the native promoter of essential gene is replaced by titratable promoter like tet promoter (Davierwala et al., 2005; Mnaimneh et al., 2004) is also being reported. Another strategy is by hypomorphic allele, such as dAMP, a

technique to destabilize mRNA transcript by integrating a selection marker into 3' UTR, downstream of the stop codon (Schuldiner et al., 2005) and C-terminal sequential affinity tag (SPA) (Babu et al., 2011; Butland et al., 2008).

Here I present a novel system that based on the Hfr conjugation gene transfer system in *E.coli* but with the aid of a 2-in-1 complementing plasmid to study the genetic interaction involving essential genes. The complementing plasmid has multi-function; it is an expression vector and also a mobile plasmid. On the plasmid, the T5 promoter and *lacI^q* repression system regulates the expression of the essential gene of interest. Upon induction, the trans supply of essential protein allows the wild type copy of essential gene to be removed from the chromosome, thus creating a conditional lethal mutant. The presence of *oriT* on the plasmid and the artificially inserted *tra* genes on the chromosome, offers the transfer of the complementing plasmid to the recipient cells during the construction of double mutant via conjugation.

CHAPTER 2

Construction of complementing plasmid (pFE604T) and conditional lethal mutants of essential query genes

2.1 Introduction

Plasmid complementation has long been utilized as a tool to identify gene essentiality in *E. coli*. Gene of interest is cloned into a suicide plasmid which contains temperature sensitive origin of replication and can be cured under the non permissive condition (Jasin and Schimmel 1984). Essential gene is determined when the strain carrying null mutation died upon the loss of the complementing plasmid.

However, until now plasmid complementation has not been described for the use of genetic interaction analysis involving essential gene. As we known, genetic interaction can be studied by combining two mutations and observing the unexpected phenotypic consequences. Thus, genetic interaction involving essential genes will be more widely studied if the essential gene mutants are easily available.

Here I demonstrate a novel 2-in-1 complementing plasmid, pFE604T (Fig. 1) which serves 2 purposes to facilitate genetic interaction analysis. (i) It allows systematic construction of conditional lethal mutants of essential genes. (ii) It also enables high throughput construction of double mutants via Hfr conjugation gene transfer system for genetic interaction analysis. The Hfr conjugation gene transfer system exploits the capacity of genetic exchange from the Hfr donor to the F$^-$ recipients to create double mutants.

Plasmid stability maintenance such as DNA replication, single copy number control, incompatibility and partition are stringently controlled by *ori2*, *incC*, *repE*, and

8

sopABC genes on the plasmid (Ogura & Hiraga, 1983; Mori et al., 1986; Uga et al., 1999). Antibiotic resistance markers (gentamycin and tetracycline) are incorporated for selection. pFE604T has two origins of replication: *ori2* and *oriRγ*. *ori2* is the default origin of replication, resulting in a single copy of the plasmid during cell division, while *oriRγ* is a conditional origin of replication that requires the *trans*-acting pi protein (encoded by *pir*) for replication and results in multiple plasmid copies each division. Due to having these two origins of replication, pFE604T can replicate at a single or medium plasmid copy number in *pir*⁻ or *pir*⁺ *E. coli* hosts, respectively. A *pir*⁻ host is used for genetic interaction experiments, while a *pir*⁺ host is used for producing additional plasmid.

The expression of essential gene in pFE604T is regulated by T5 promoter and *lacIq* repressor system under the induction of IPTG. *Sfi*I restriction sites are located downstream of the T5 promoter, enabling the facile transfer of an essential query gene from the ASKA library (Kitagawa et al., 2005). Trans-supply of of essential protein from pFE604T permits the removal of essential gene from the chromosome, thus creating a conditional lethal essential gene mutant (Fig. 2).

During the construction of double mutant, two steps are required to ensure the transfer of the complementing plasmid and the resistance marker replacing the essential gene from the donor to the recipients. The *oriT* presents in the plasmid offers the convenience of one-step transfer to an array of recipient cells (Keio collection) in order to complement the subsequent transfer of the essential gene deletion in the double mutation background (Fig. 3).
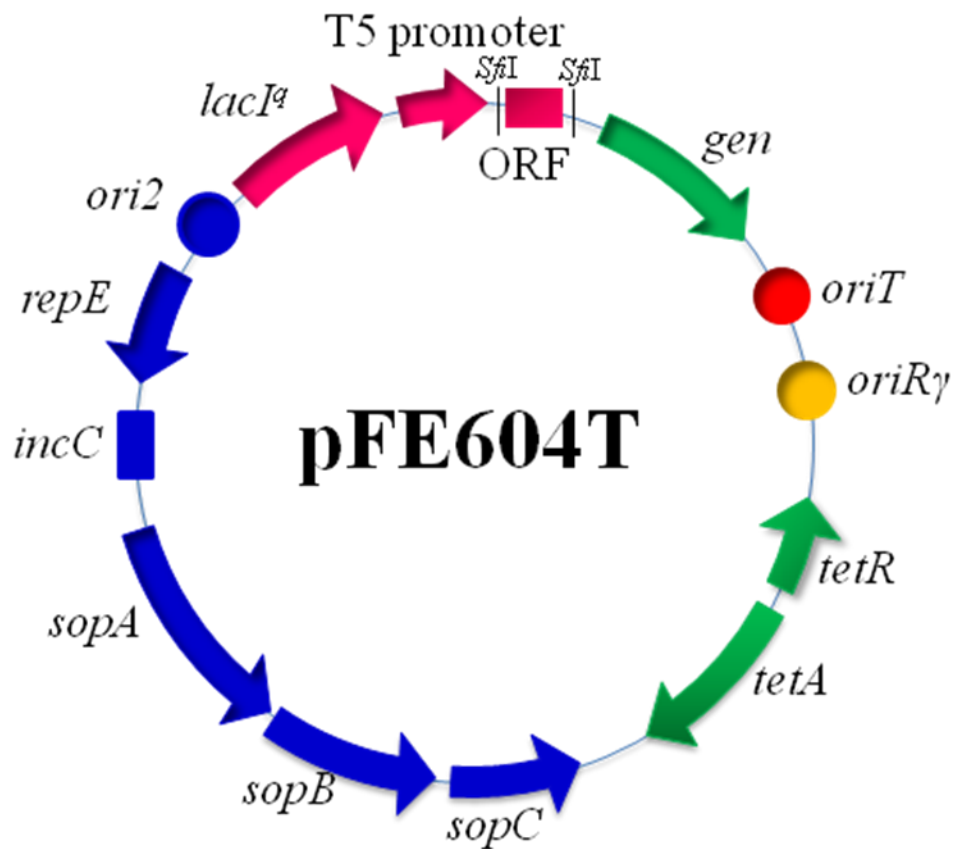
Figure 1: Diagram of the 2-in-1 complementing plasmid, pFE604T used in the construction of conditional lethal mutant of essential gene and high throughput genetic interaction analysis. pFE604T is a single-copy mini-F derivative that contains features for plasmid stability maintenance (blue colors), selection markers (green colors), expression of the essential gene (pink colors), and conditional replication (yellow colors). A target gene can be cloned into the ORF region and its expression is under the regulation of the *lacI^q* repressor system and a T5 promoter
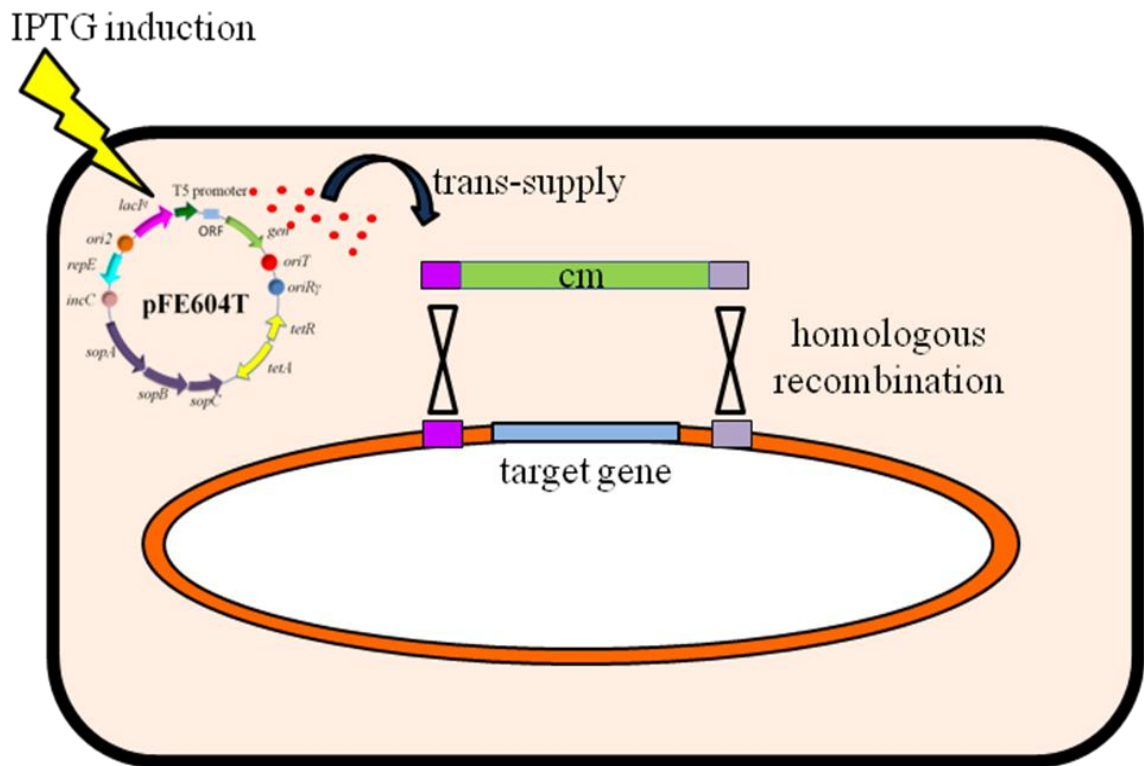
Figure 2: Deletion of chromosomal essential gene can be achieved by the provision of essential protein from the complementing plasmid. Under the induction of IPTG, the essential protein will be expressed from the T5 promoter $lacI^q$ repressor system at knock-down level, thus the essential gene on the chromosome can be replaced by the antibiotic selection marker.
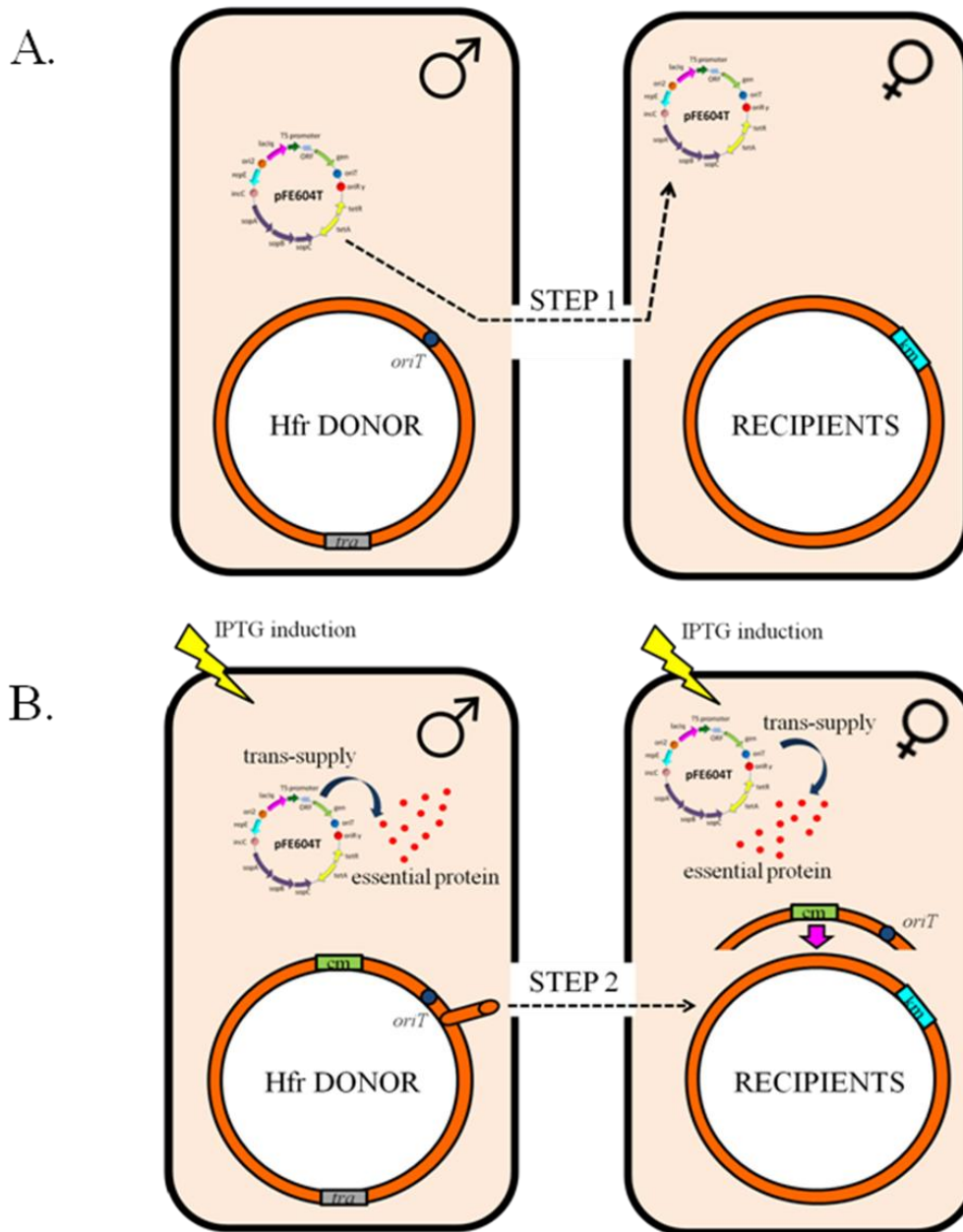
Figure 3: High throughput construction of double mutants via Hfr conjugation gene transfer system. (**A**) Donor strain carrying complementing plasmid transfers the complementing plasmid to an array of recipient strains, Keio collection marked with kanamycin via conjugation, initiated by the *oriT* on the plasmid. (**B**) Donor strain of conditional lethal mutant of essential gene transfers the essential gene deletion which is marked by chloramphenicol cassette to an array of recipient strains, Keio deletions

carrying the complementing plasmid which are made in step (A) for the construction of double mutants with the presence of 0.1 mM IPTG.

2.2 Materials and methods

2.2.1 Strains and growth conditions

BW25113($lacI^q$ $rrnB_{T14}$ $\Delta lacZ_{WJ16}$ $hsdR514$ $\Delta araBAD_{AH33}\Delta rhaBAD_{LD78}$),BW25141($lacI^q$ $rrnB_{T14}$ $\Delta lacZ_{WJ16}$ $\Delta phoBR580hsdR514$ $\Delta araBAD_{AH33}\Delta rhaBAD_{LD78}$ $galU95endA_{BT333}$ $uidA(\Delta MluI)∷:pir^+$ $recA1$) are derivatives of the F⁻, λ⁻, *E. coli* K-12 strain BD79 and have no other known mutations. Conditional replicative *oriR*γ plasmids were maintained in the *pir⁺* host BW25141 or similar ones (Datsenko & Wanner, 2000). All the strains used were routinely grown in LB medium containing 1% Bacto Tryptone, 0.5% yeast extract, and 1% NaCl with or without antibiotics at 50 µg/ml for ampicillin, 30 µg/ml for kanamycin, 5 µg/ml for gentamycin, 12.5 µg/ml for tetracycline, 25 µg/ml for chloramphenicol at 30 ℃ or 37 ℃. Conditional lethal mutant strains were grown in LB medium containing tetracycline, chloramphenicol and 0.1 mM IPTG at 37 ℃.

2.2.2 Plasmids

pAH143 (Haldiman & Wanner, 2001), pLZ2210-CAS8 (Wanner et al., unpublished data), pKD46 and pKD3 (Datsenko & Wanner, 2000) were gifts from B. Wanner (University of Purdue, Indianapolis). CIP (Takeuchi et al., unpublished data), pFE604 (Yamamto et al., unpublished data) and ASKA ORF clones (Kitagawa et al., 2005) were from the home laboratory.

2.2.3 Generation of PCR fragment

PCR reactions were carried out in 50 µl reactions containing 1.0 U of Toyobo KOD polymerase, template DNA, 0.3 µM of each reverse and forward primer as listed in Table 1, 1.0 µM of $MgSO_4$ and 1.0 µM of dNTPs. Reactions were first pre-denatured at 98ºC for 1 minute, then run for 30 cycles: 98ºC for 15 seconds, 65ºC for 2 seconds and 74ºC with 40 seconds. PCR products were digested with *Dpn*I, ethanol precipitated, and resuspended in 30 µl H_2O and analyzed by 1 % agarose gel electrophoresis.

2.2.4 One step homologous recombination

Linear transforming DNA fragment was prepared by PCR using primers with 50-nt flanking homology extensions to the target site as described above. First, host strain was transformed with pKD46 plasmid and selected on ampicilin containing plate at 30 ºC. pKD46 has a temperature sensitive origin of replication and encoded lambda red recombinase genes which are required for homologous recombination (Datsenko & Wanner, 2000). The expressions of the recombinases are inducible by arabinose. Transformant carrying pKD46 was purified by single colony isolation and was precultured overnight in liquid at 30 ºC. Next, overnight culture was subcultured by diluting 1:10 in SOB medium with 1 mM of arabinose and further incubated until OD_{600} reached around 0.6 (~3 hours). The culture was then washed carefully 3 times in 10 % ice cooled glycerol to make competent cell. 40 µl of freshly prepared competent cells were mixed with 1-2 µl of transforming DNA in an ice cold cuvvette. Cells were electroporated at 2.4 kV with 25 mF and 200 Ω, immediately followed by the addition of 1 ml of SOC medium and incubated at 37 ºC for 1-2 hours for recovery. 100 µl of recombinant was plated for selection and incubated at 37 ºC for 1 day to cure the pKD46 plasmid.

2.2.5 PCR verification of constructs

PCR reactions were performed in 20 µl PCR reactions containing 2.5 U of *TaKaRa Ex Taq* polymerase, 1.0 µM of each primer (see Table 1), and 200 µM of dNTPs. As for DNA template, it was obtained by using yellow tip to pick up colonies from single colony isolation or 1 µl of liquid culture. Reactions were run for 30 cycles: 94ºC for 30 seconds, 59ºC for 30 seconds, and 72ºC for 2 minutes and additional 2 minutes at 72ºC.

2.2.6 Construction of conditional lethal mutants

The expression of target essential gene in complementing plasmid was first induced by the addition of 0.1 mM IPTG before the replacement of chromosomal ORF of target essential gene with chloramphenical cassette by one-step homologous recombination as described in 2.2.4.

2.2.7 Growth curve analysis of conditional lethal mutants

Overnight cultures of conditional lethal mutants were inoculated into 96-well microtitre plates containing 200 µl of liquid medium with tetracycline and chloramphenicol and supplemented with 1 mM, 0.1 mM , 0.01 mM, 0.001 mM, 0.0001 mM or zero supply of IPTG. Cultures were incubated at 37°C for 24 hours with optical density (600 nm) measured every 30 min using an automated SpectraMax® GEMINI EM (Molecular Devices Inc).

2.2.8 Conjugation in liquid culture

Construction of Hfr donor strain, examination of the transfer ability of the complementing plasmid and examination of the transfer ability of Hfr donor strain of conditional lethal mutants of essential query genes were all performed by conjugation in liquid culture. Overnight culture of donor strains was subcultured by diluting 1:25 in LB medium and incubated for 3 hours in a 37 ℃ water bath. Overnight culture of recipient strain was subcultured by diluting 1:5 in LB medium and aerated at 37 ℃ until used. 200 µl of each donor and recipient strains were mixed in test tube and let sit in a 37 ℃ water bath for 90 minutes. 2 ml of LB medium was added to test tube and aerated for an additional 1-2 hours at 37 ℃. Mating mixture was plated on selection plate and further incubated overnight at 37 ℃.

Table 1: Primers used in this study

Remarks	Name	Sequence (5'-3')
PCR of *oriT* fragment for cloning	YHP1	AAGGAAAAAAGCGGCCGCCACCTCTGGTGACTTTATC
	YHP2	TTTTCCTTTTGCGGCCGCGCAGCGCCCCTAGCGGTATC
oriT structure colony PCR check	RT1	TCTCGCATAAAAAACTGCGCAGGGCGCTGAAGGCCATCACCCGTTCAGAACTGGCAGTTCCCTACTCTCG
	RT5	CCTTAGAGGCTATTTAAGTTGCTGA
PCR of *oriT* and *gen* fragment for homologous recombination between *sopC* and cat in pFE604 (inverted direction)	YHP13	TGCGGTCGCCCGCTTACAGGTGCGGCACGGCCTGATGGAGGCCGCATGTGTCTTGCGGCCGCGCAGCGCC
	YHP14	TACGCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCTCCTCTCTGGCAGTTCCCTACTCTCG
PCR of *oriT* and *gen* fragment for homologous recombination between *sopC* and *cat* in pFE604 (non-inverted direction)	YHP15	TGCGGTCGCCCGCTTACAGGTGCGGCACGGCCTGATGGAGGCCGCATGTGCTGGCAGTTCCCTACTCTCG
	YHP16	TACGCCTGAATAAGTGATAATAAGCGGATGAATGGCAGAAATTCTCCTCTTCTTGCGGCCGCGCAGCGCC
oriT and *gen* position between *sopC* and *cat* in pFE604 colony PCR check	NYP45	TATACGCAAGGCGACAAG
	NYP140	AAGTATTGACATGTCGTCGTAAC

PCR of *oriT* and *gen* fragment for homologous recombination between *ori2* and *laq1^q* in pFE604 (inverted direction)	YHP17	TACCGCACAGATGCGTAAGGAGAAAATAC CGCATCAGGCGCTCTTCCGCTTCTTGCGGC CGCGCAGCGCC
	YHP18	GTAAGCAGAATATATAAGTCCTGTTCCCTG GTGCTTCCTCGCTCACTCGACTGGCAGTTC CCTACTCTCG
oriT and *gen* position between *ori2* and *lacI^q* in pFE604 colony PCR check	NYP43	AGTCACGTAGCGATA
	NYP49	CATATCACCAGCTCACC
PCR of *oriT* and *gen* fragment for homologous recombination between atth80 and *oriγ* in pFE604 (inverted direction)	YHP19	GCAATTTTCAGTGACACAGGAACACTTAAC GGCTGACATGGGAATTAGCCTCTTGCGGCC GCGCAGCGCC
	YHP20	TATAATTCTTGAAGACGAAAGGGCCTCGTG ATACGCTCGCACGGGCCCATCTGGCAGTTC CCTACTCTCG
oriT and *gen* position between atth80 and *oriγ* in pFE604 colony PCR check	NYP244	GCTGAACGGTCTGGTTATAGG
	ECK0342-right	TGGTGGTGTCGATGGTAGAA
pFE604T *dnaN* cloning colony PCR check	NYP121	ATCAGCTGTTGCCCGTCTC
	ECK3693-right	TTGAAGCCGATTTCCATCTC
pFE604T *ftsW* cloning colony PCR check	NYP121	ATCAGCTGTTGCCCGTCTC
	ECK0090-right	TGGAAGCTAAACCAGATGCC
pFE604T *trmD* cloning colony PCR check	NYP121	ATCAGCTGTTGCCCGTCTC
pFE604T *yrfF* cloning	NYP121	ATCAGCTGTTGCCCGTCTC

colony PCR check	ECK3385-right	TGCTGAACGTAATGAGGCAC
pFE604T *yjgP* cloning colony PCR check	NYP121	ATCAGCTGTTGCCCGTCTC
	ECK4252-right	TTCGATTTCAGGGAGGTCTG
dnaN chromosomal deletion	ECK3693N	CAAAGAAGATTTTTCAAATTTAATCAGAACATTGTCATCGTAAACCTATGATTCCGGGGATCCGTCGACC
	ECK3693C	ATCGCGGATCAACAAGCGGGTGAGGGACATTACAGTCTCATTGGCATGACTGTAGGCTGGAGCTGCTTCG
ftsW chromosomal deletion	ECK0090N	GAACAACGAGGCAATGAGTTTGCCCGTCTGGCGAAGGAGTTAGGTTGATGATTCCGGGGATCCGTCGACC
	ECK0090C	CATCACCATTAATCGCTTTCCTTGACCACTCATCGTGAACCTCGTACAAATGTAGGCTGGAGCTGCTTCG
trmD chromosomal deletion	ECK2604N	GGGATCCTGGTTTTTAAACCACCGGATAAACGGTAAAAGACGGCGCTATGATTCCGGGGATCCGTCGACC
	ECK2604C	TATCCTGGGTAAACTGATATCTCGGGGGCTTACGCCATCCCATCATGTTTTGTAGGCTGGAGCTGCTTCG
yrfF chromosomal deletion	ECK3385N	CCGGTGCGACTGACCACGCCTGACAGACTAAGTAAGATGGGGAAAGCATGATTCCGGGGATCCGTCGACC
	ECK3385C	AGGGTAGCATAACCTGCCGCGCAAACGTGTTATTCGATAAGGCTTTCTGATGTAGGCTGGAGCTGCTTCG
yjgP chromosomal	ECK4254N	ACAAGCTAAAATCCTGCAAAAGACGAGTTTTTACGGGCGTATTTAAAGTGATTCCGGGGA

deletion		TCCGTCGACC
	ECK4254C	ATAGCGGTCAAGTACGCCAAAAGGTTGCAT CACACCGCTCCTTTACGCGATGTAGGCTGG AGCTGCTTCG
dnaN conditional lethal mutant partial duplication check	NYP244	GCTGAACGGTCTGGTTATAGG
	ECK3693-right	TTGAAGCCGATTTCCATCTC
	ECK3693-up	CCCTGCTGGAAGGTAATCAA
ftsW conditional lethal mutant partial duplication check	NYP244	GCTGAACGGTCTGGTTATAGG
	ECK0090-right	TGGAAGCTAAACCAGATGCC
	ECK0090-up	GATGCCTTAACAATGCCGAT
trmD conditional lethal mutant partial duplication check	NYP244	GCTGAACGGTCTGGTTATAGG
	ECK2604-right	CAGTCAGAGCCAGGTTTTCC
	ECK2604-up	TATTTCTGATCGCGTTGCTG
yrfF conditional lethal mutant partial duplication check	NYP244	GCTGAACGGTCTGGTTATAGG
	ECK3385-right	TGCTGAACGTAATGAGGCAC
	ECK3385-up	CCAGTGATTCCGGGTAGAGA

yjgP conditional lethal mutant partial duplication check	NYP244	GCTGAACGGTCTGGTTATAGG
	ECK4254-right	TTCGATTTCAGGGAGGTCTG
	ECK4254-up	CGTGCAGCTCAGTCAGAAAG

2.3 Results

2.3.1 Construction of pFE604T

There are few steps involved to construct pFE604T (Fig. 4). Important issues associated with plasmid design including plasmid copy number and stability maintenance, origin of replication, essential gene expression regulation, antibiotic selection and plasmid mobility for high throughput conjugation were considered and addressed. pFE604T is a derivative of mini F plasmid which relies on the *ori2*, *repE*, *sopA*, *sopB*, *sopC* and *incC* gene for the stringent maintenance of single plasmid copy number. Besides, two origins of replication were incorporated, *ori2*, and conditional replication, *oriRγ* from the R6K origin which requires pi replication protein. With such design, it is feasible to copy up the plasmid (Fig. 5) by transforming into pir^+ host strain. This feature greatly facilitates plasmid extraction and for subsequent manipulation like cloning.

Insertion of DNA fragment of *oriT*, *gen* and *oriRγ* from the pAH143oriT into pFE604 by homologous recombination was not successful until the *oriRγ* was inserted seperately from *oriT* and *gen*. The reasons were unknown, but the insertion location or DNA fragments containing *oriT*, *gen* and *oriRγ* might cause instability or adverse effect to the replication. As a results, insertion of DNA fragment of *oriT* and *gen* was attempted in different positions of the plasmid and also in different orientations, non-inverted or inverted to avoid potential problems (Fig. 4, B). All constructs in different positions and orientations were successfully made, finally the construct of DNA fragment of *oriT* and *gen* in non-inverted orientation and inserted between T5 promoter and *oriRγ* was selected because that position is the join of fragments from two plasmids during the construction of pFE604, so it is deemed to be most stable for modification.

In my previous experiment, gentamycin was discovered as a weak selection for single copy plasmid, this phenomenon was observed when longer incubation period (about 2 days) was needed for the clear selection. In substitute, tetracycline gene, TetAR originated from Tn10, was used for the selection of single copy plasmid. Expression of

the essential gene in pFE604T is induced by IPTG under the regulation of *lacI^q* repressor system and T5 promoter. *lacI^q* is a mutant *lac* repressor which produces 10 times more repressor protein than wild type *lacI* (Muller-Hill et al., 1968).
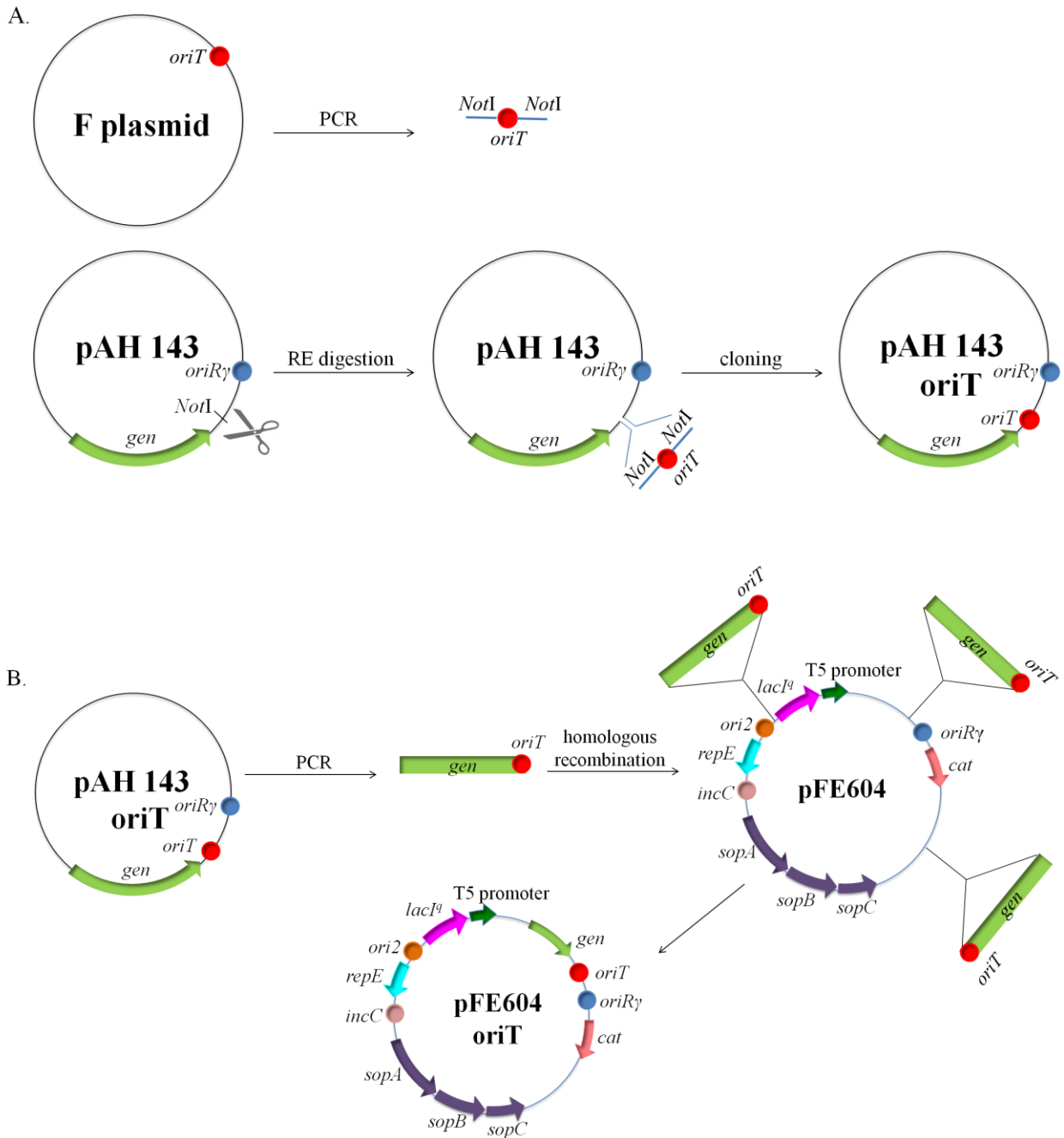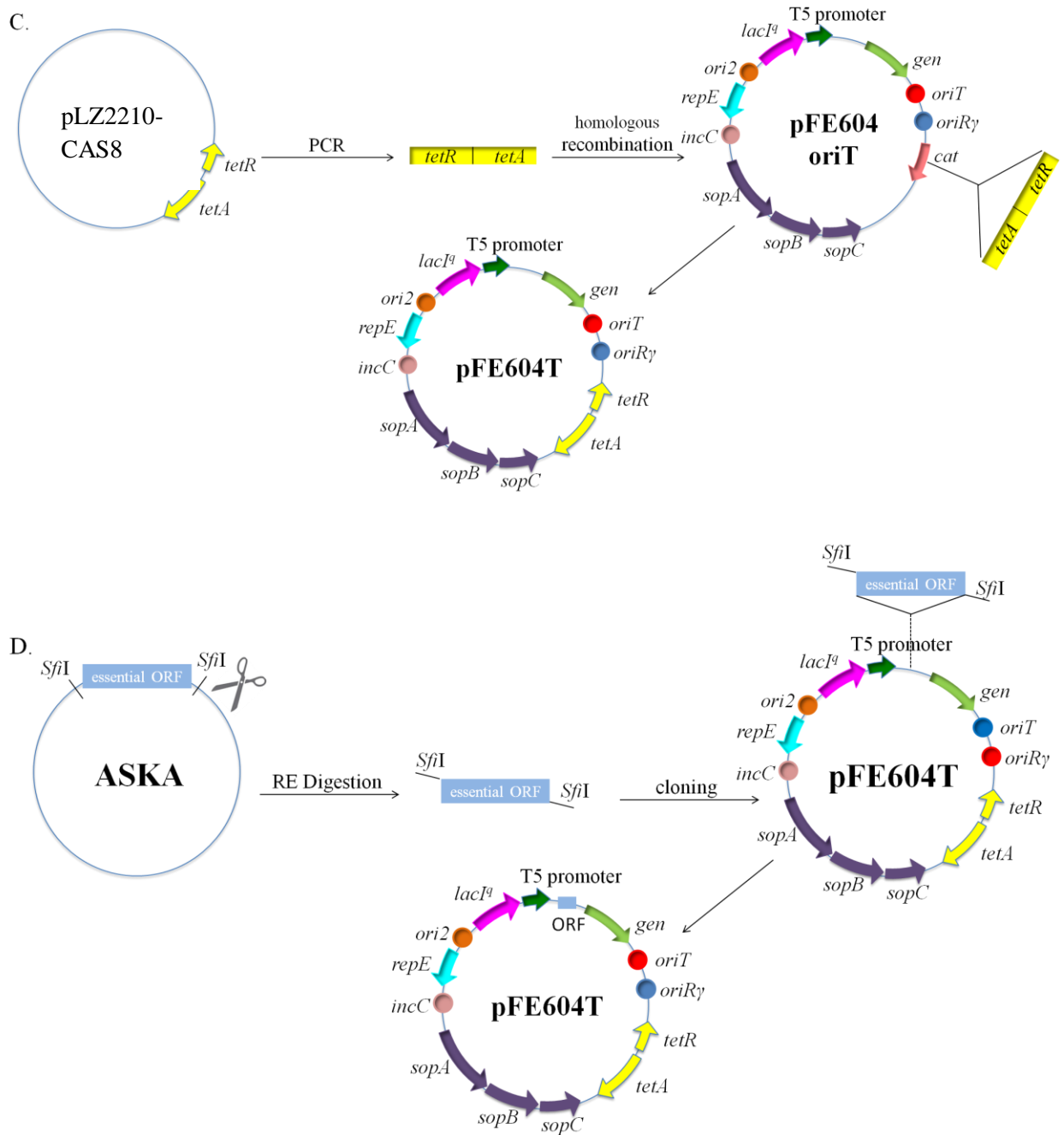
Figure 4: Construction of pFE604T. (**A**) *oriT* fragment from mini F plasmid was amplified by PCR using primers with *Not*I flanking sites. pAH143 vector was digested by *Not*I restriction enzyme and the *oriT* fragment was cloned into the vector. (**B**) The *oriT-*

24

gen fragment of pAH143oriT was amplified by PCR using primers with 50-nt homology extensions to different positions of pFE604 as shown above. Fragments were integrated into pFE604 via homologous recombination in both inverted and non-inverted direction. (**C**) Tetracycline resistance cassette from pLZ2210-CAS8 was amplified by PCR using primers with 50-nt extensions to upstream and downstream of *cat* gene and replaced the *cat* gene by homologous recombination. (**D**) Target essential genes fragments were digested from the ASKA clones by *Sfi*I and cloned into pFE604T plasmid creating essential gene complementing plasmid.
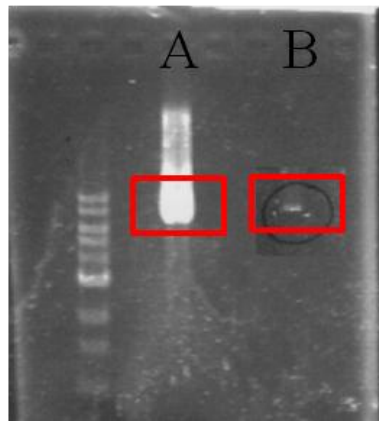
Figure 5: Plasmid yield check on electrophoresis gel. A and B lane indicate the plasmid extracted from 50 ml culture of BW25141, a pir^+ strain and BW25113, a pir^- respectively. 3 µl of each sample loaded. The presence of pi protein initiates the *oriRγ* replication together with the *ori2* replication; hence the plasmid number is copied up.

2.3.2 Conditional lethal mutants of essential query genes

Five essential query genes encoded for different functions were selected for the genetic interaction analysis. There are *dnaN*, encoding a beta subunit of DNA polymerase; *ftsW* which involved in cell division; *trmD* which takes part in tRNA

methylation; *yjgP*, a lipopolysaccharide transporter and *yrfF*, an unknown gene. The fragments of target essential genes were excised from ASKA plasmid clone library and inserted into pFE604T (Fig. 4, D). Conditional lethal essential gene mutants were generated using the system by complementing plasmid as described above (Fig. 2). Partial duplication check was performed to ensure that there is no duplication of deleted target essential genes located right next to the target gene on the chromosome (Fig. 6). The presence of the duplicated target essential gene in chromosome will affect the genetic interaction analysis as the expression of the target essential gene is not reduced, in fact similar to wild type.
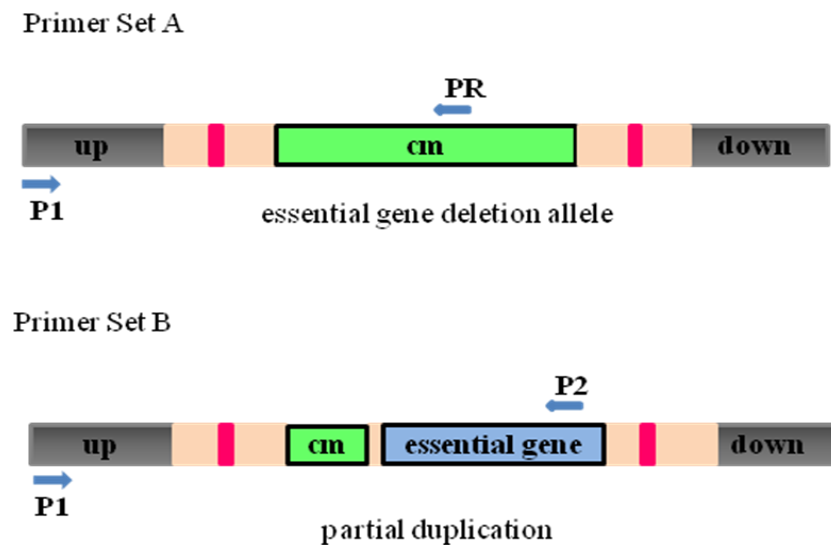
Figure 6: Partial duplication was checked using 2 sets of primers. Primer set A amplifies the upstream of target gene and within chloramphenicol cassette to confirm the correct position for the deletion. Primer set B amplifies the upstream of target gene and within target gene for partial duplication check. If partial duplication occurs in the conditional lethal mutant, we will expect to see PCR bands using both primers, because of the presence of another wild type copy located right next to the chloramphenicol casseete. Pink color indicates FRT sites.

For the purpose of genetic interaction analysis, it is very crucial to determine that the source of essential protein in the conditional lethal essential gene mutant is only trans-supplied from the complementing plasmid at knock-down level upon the induction of IPTG. Thus, the growth of conditional lethal mutants is expected to be lower than wild type strain and it is IPTG dependent.

The growth profile of each conditional lethal mutant of target essential genes in in different IPTG concentrations was examined in both on agar (Fig. 7) and in liquid (Fig. 8). As we can observe that, the IPTG dependency order of conditional lethal mutants is *yjgP*, *dnaN*, *yrfF*, *ftsW* and *trmD*. *trmD* does not show any IPTG dependency, and the mutant is able to grow even in the absence of IPTG. The failure of complete repression in the absence of IPTG could be due to the leak of *lacI*^{q} promoter. FtsW and TrmD which are required in small amount in the cell are the least IPTG dependent of all the 5 mutant strains. This could be reasoned by the leak expression from the promoter is sufficient for the viability even without the induction by IPTG.

To improve the control of the regulation system by *lacI*^{q} and T5 promoter, original pFE604T was modified by inverting the *lacI*^{q} direction or inserting a terminator after the *lacI*^{q} (Fig. 9). The new constructs with inverted *lacI*^{q} or additional terminator were tested for growth profile in different concentrations of IPTG. Based on the results, unfortunately the repression was not obviously increased when the *lacI*^{q} is inverted or with the additional terminator.
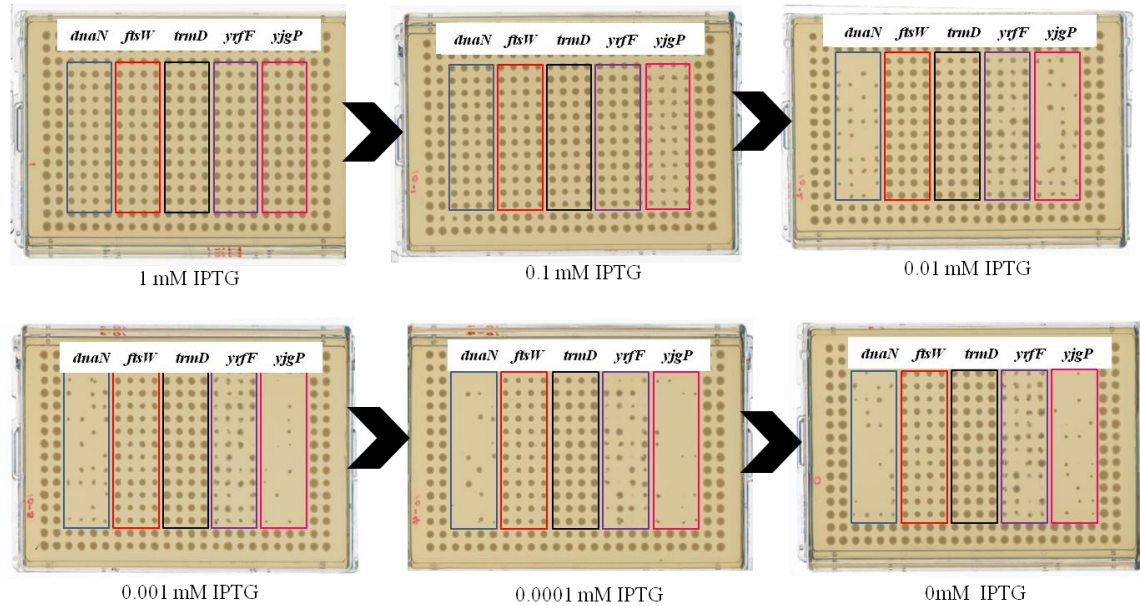
Figure 7: IPTG dependency of conditional lethal mutants of target essential genes on agar in different IPTG concentrations. *dnaN* and *yjgP* show the strongest IPTG dependency, wherein the growth corresponds with the supply of IPTG. At the zero supply of IPTG, *dnaN* and *yjgP* conditional lethal mutants fail to grow because of the absence of the respective essential protein required for survival. *yrfF* and *ftsW* show weaker dependency, whereas *trmD* is not IPTG dependent at all.
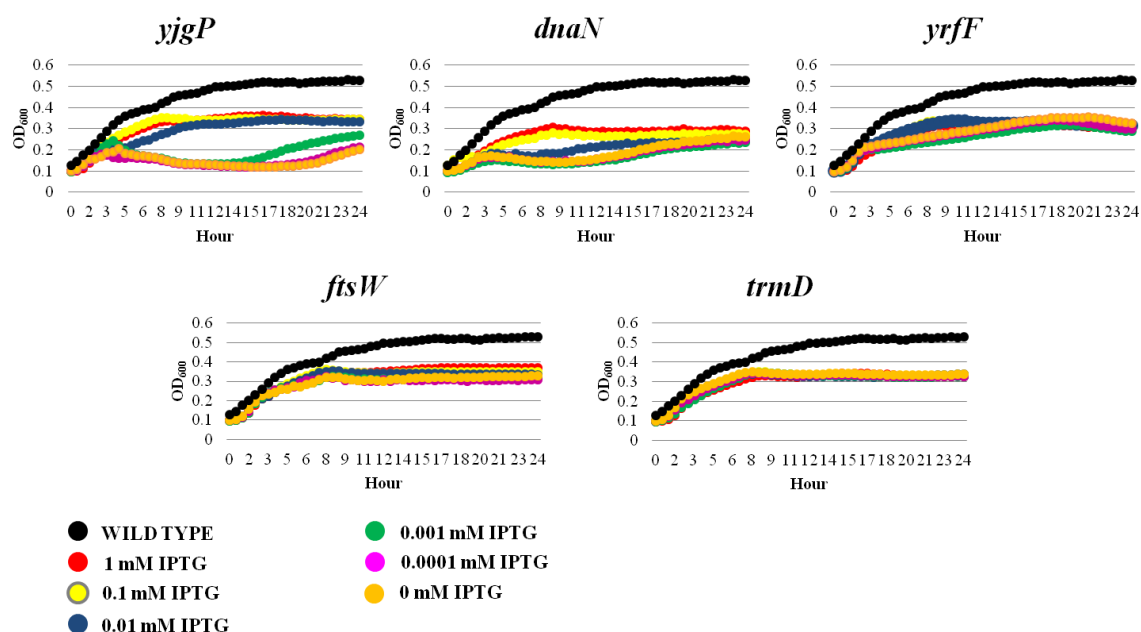
Figure 8: IPTG dependency of the conditional lethal mutants of target essential genes in liquid in different IPTG concentrations. The IPTG dependency profile of the conditional lethal mutants of target essential genes in liquid is similar to the one on agar. *yjgP* and *dnaN* show the strongest dependency followed by *yrfF* and *ftsW*. *trmD* is not IPTG at all, but the growth is slower than wild type even with the high concentration of IPTG, 1 mM. The OD was measured every 30 minutes automatically by plate reader.
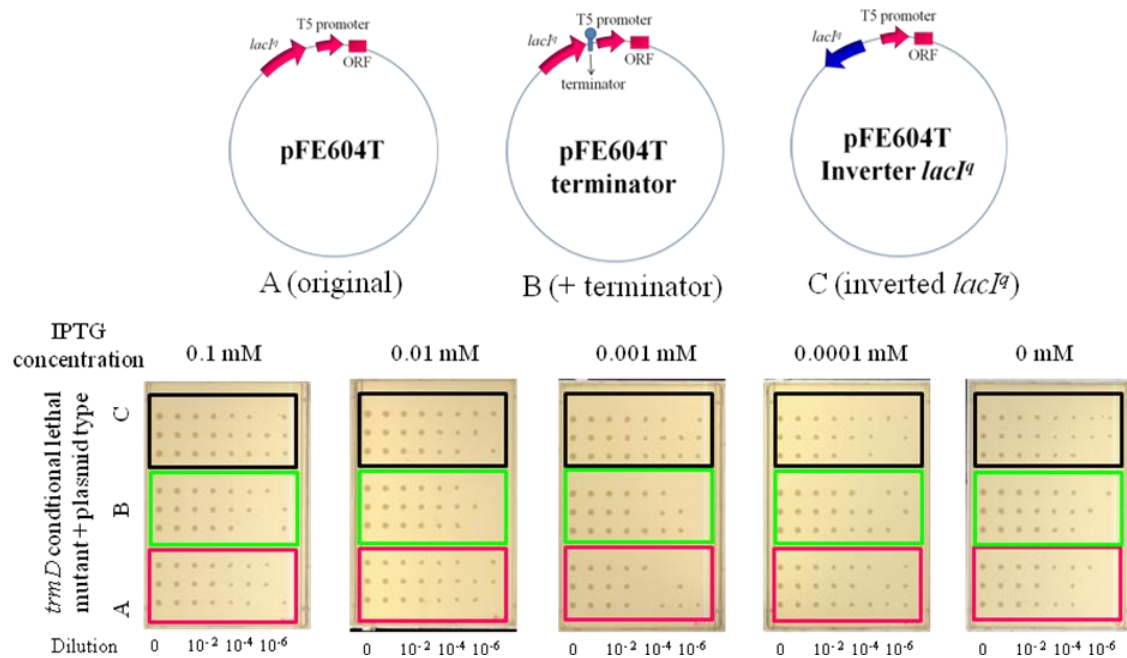
Figure 9: Modification of the original pFE604T (plasmid A) for the tighter regulation of TrmD expression from the _lacI^q_ repressor system and T5 promoter. _trmD_ was chosen because of the lack of IPTG dependency of the mutant, so it would be a good control to observe improved regulation. A terminator was inserted after the _lacI^q_ (plasmid B) to prevent the leak expression from the _lacI^q_ promoter or inverting the _lacI^q_ direction (plasmid C). Unfortunately, at the zero supply of IPTG, the _trmD_ conditional lethal mutants complemented by plasmid A or B or C did not show significant repression as desired. Triplicates of experiment was conducted and spotted on the agar plate.

Before I could start genetic interaction analysis by the Hfr conjugation gene transfer system, I had to convert the conditional lethal mutants of target essential genes into Hfr strain or literally known as artificial male strain, by insertion of F factors (*tra* genes and *oriT*) into the chromosomes. The F factors are carried in a CIP plasmid, a plasmid with a conditional origin of replication, *oriRγ*. In the absence of pi replication protein in *pir⁻* host, the plasmid will be integrated into the host chromosome's homology sequence (Takeuchi et al., unpublished data) (Fig. 10). The Hfr donor strains of all the conditional lethal mutants of target essential genes are able to transfer the genetic material to the Keio deletion by conjugation to make double mutant, as confirmed by the experiment of mating in liquid culture.
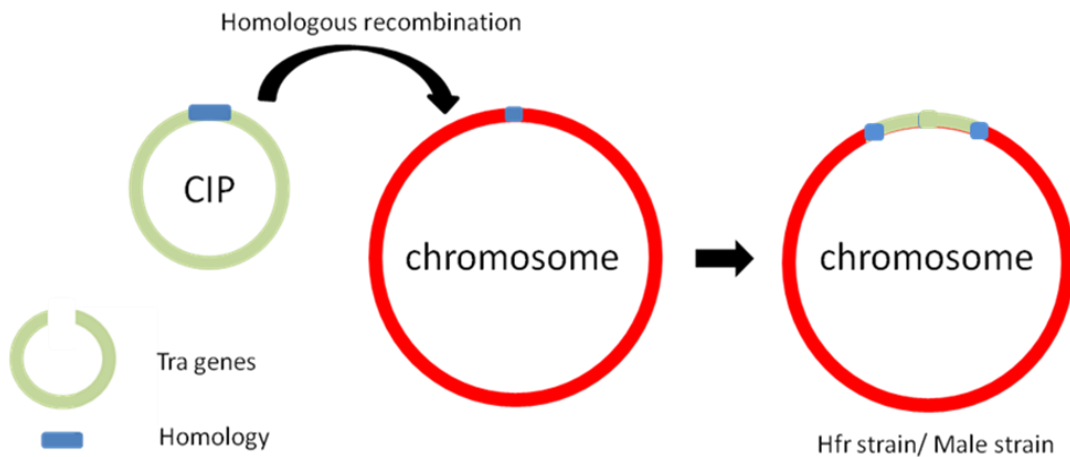
Figure 10: CIP plasmid for the making of Hfr strain. CIP plasmid carries *tra* genes which are necessary for conjugation and homology sequence to the host chromosome is integrated into the host chromosome and converting the host strain into a Hfr strain which has the capacity to transfer genetic materials to recipient cells.

2.4 Discussion

The system for the construction of conditional lethal mutant of essential gene and the construction of double mutant of essential gene and non essential gene has been successfully established using the 2-in-1 complementing plasmid, pFE604T.

Conditional lethal mutants of the 5 target essential genes (*dnaN*, *ftsW*, *trmD*, *yrfF*, *yjgP*) were generated using the strategy by the pFE604T. The inducible regulation system of *lacI*^{q} and T5 promoter controls the trans-supply of essential protein from the plasmid in order to compensate the deletion on the chromosome. However, due to the leaky promoter of *lacI*^{q}, even without the induction of IPTG, *ftsW* and *trmD* conditional lethal mutants are able to survive. Trials to tighten the regulation system by modifications of the *lacI*^{q} site were not as successful as expected. Nevertheless, the growth of all the conditional lethal mutants of target essential genes in 0.1 mM of IPTG is still lower than the wild type as the growth rates were slower compared to wild type shown in Figure 8, so it assumes that the expression of essential protein is decreased compared to wild type and the strains can be used for genetic interaction analysis. Besides, it is experimentally proven that the Hfr conditional lethal mutants can be used as a donor for genetic interaction analysis involving essential gene by crossing with the Keio single gene deletion.

The materials for genetic interaction analysis involving essential genes are now ready. So, the next challenge is the development of a system for high throughput genetic interaction analysis, including construction of double mutants, colony imaging and quantification and statistical analysis of interactions.

CHAPTER 3

Establishment of genetic interaction analysis by spotting method

3.1 Introduction

High throughput genetic interaction analysis of *E. coli* based on Hfr conjugation gene transfer system of non-essential deletions or hypomorphic strains to create double mutants has been reported previously (Typas et al., 2008; Butland et al., 2008; Babu et al., 2011). In general, few steps involved during the construction of double mutant. First, the marked deletion of donor strain (F^+) will be transferred to an array of single gene deletion strains of opposite mating type marked with different selectable marker. Intermediate selection which will eliminate background and to stabilize double mutants will be performed. Lastly, double mutants will be selected out in second selection, and digital image of the double mutants will be quantified and colony fitness will be determined statistically (Fig. 10). Followed by this, genetic interactions will be determined and functional analysis will be performed.

The density of starting inoculum of double mutant for the second selection is the key factor of the sensitivity of genetic interaction analysis. Two types of starting inoculum for the second selection after the intermediate selection had been compared, diluted liquid culture or dense colony. Diluted liquid culture is prepared by spotting method, in which colonies after the intermediate selection are first diluted in 1 X PBS before being spotted on the second selection plate. Whereas, the dense colony is prepared by stamping method where colonies of double mutants from the intermediate selection plate are directly transferred onto the second selection plate by replica pinning (Fig. 12) .

In short, the major biological difference of the spotting and the stamping method is the density of starting inoculum of double mutants being deposited on the second selection plate. The spotting method in general has better IPTG sensitivity than the stamping method due to the lower starting inoculum for the second selection (Fig. 13). Higher starting inoclum in stamping method presumbly causes carryover of gene product and small amount of IPTG that leads to the capability of survival of double mutants even in the absence of IPTG.

In this chapter, I will focus and present the establishment of the spotting method and the disadvantages, tehnical problems and solutions associated with the method. On the other hand, stampinng method will be discussed in the following chapter.
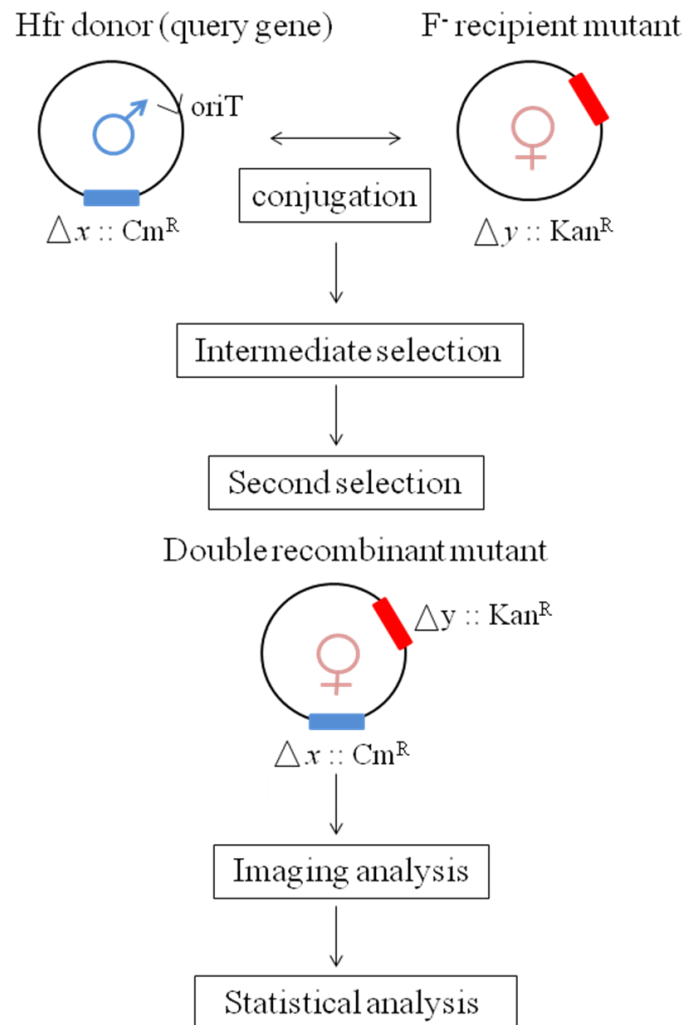
Figure 11: Prototype of genetic interaction analysis based on Hfr conjugation gene transfer system in *E. coli*. A query strain bearing a deletion of gene of interest marked with *cat* is crossed against recipient strains bearing individual gene deletions by conjugation. Intermediate selection is carried out to minimize background and to stabilize double mutants before the second selection which only selects out double mutant. Colonies of double mutants are imaged and quantified for the fitness. Statistical analysis is performed to determine genetic interactions.
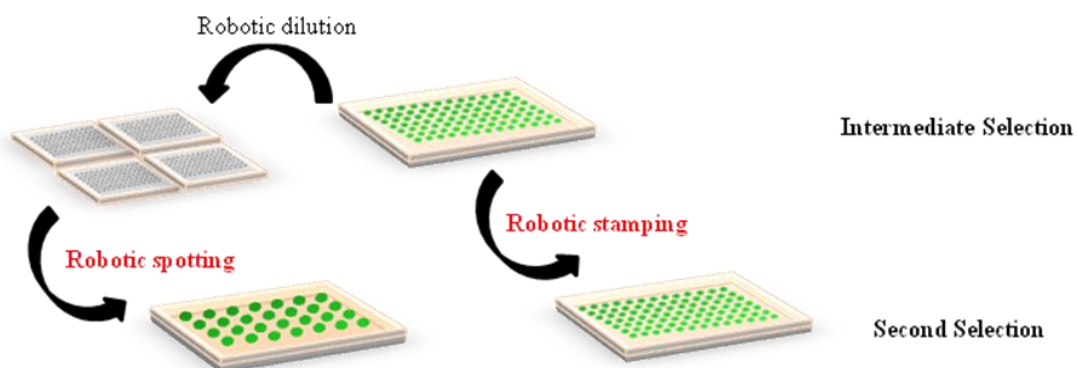
Figure 12: The mechanical difference of spotting method and stamping method. In the spotting method, colonies from the intermediate selection plate will be first diluted in 1 X PBS before being spotted onto the second selection plate. In the stamping method, colonies are directly replicated from the intermediate selection plate to the second selection plate. The biological difference of these 2 methods is the density of starting inoculum of double mutant from the intermediate selection for the second selection.
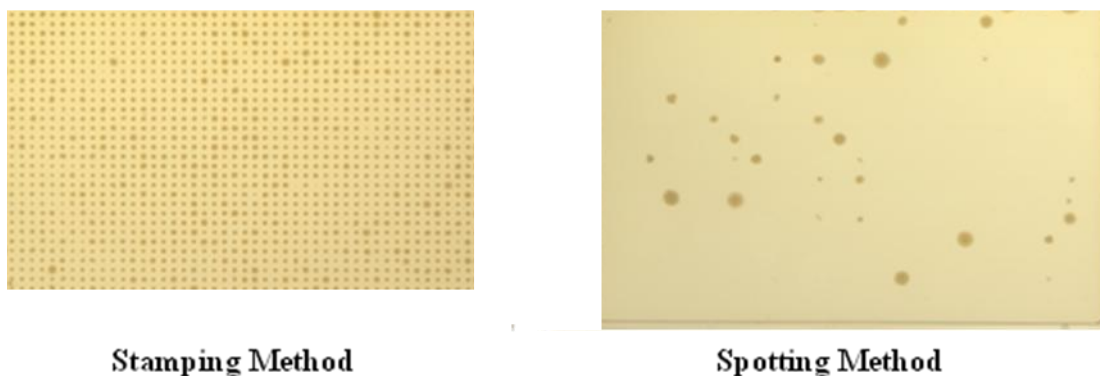
Figure 13: The spotting method demostrates better IPTG sensitivity compared to the stamping method. Double mutants of *dnaN* and nonessential genes were selected on with zero supply of IPTG. Colonies of double mutants produced from the stamping method (left) were small. Colonies of double mutants produced from the spotting method were not able to grow in the absence of IPTG.

3.2 Materials and methods

3.2.1 Strains and growth conditions

BW25113($lacI^q$ $rrnB_{T14}$ $\Delta lacZ_{WJ16}$ $hsdR514$ $\Delta araBAD_{AH33}\Delta rhaBAD_{LD78}$) is derivative of the F⁻, λ⁻, *E. coli* K-12 strain BD79 and have no other known mutations. Nonessential gene deletions were grown in LB containing 30 µg/ml of kanamycin. Conditional lethal mutant strains were grown in LB medium containing 12.5 µg/ml of tetracycline, 25 µg/ml of chloramphenicol and 0.1 mM IPTG at 37 ℃.

3.2.2 Genetic interaction analysis by spotting method

Overnight culture (16-18 hours) of donor strain carrying complementing plasmid of query essential gene was spread on LB plate and incubated for 1 hour. Recipient strains, nonessential gene deletions were arrayed on kanamycin containing LB plate in 1536 density format one night before. To transfer the complementing plasmid from donor to recipient, recipient strains were crossed to donor strain by pinning on donor lawn and allowed conjugation to take place for 6 hours. After 6 hours of conjugation, the conjugants (nonessential gene deletions carrying complementing plasmid of query essential gene) were selected on tetracycline and kanamycin containing LB plate and incubated for 24 hours for the use of next step. The following day, overnight culture (16-18 hours) of conditional lethal mutant of target essential gene was washed with LB medium and incubated for half an hour before spread on LB plate containing 0.1 mM IPTG. Nonessential gene deletions carrying complementing plasmid prepared in the previous day were pinned on the essential gene mutant lawn and mated for 6 hours. Intermediate selection was performed in LB plate containing tetracycline, kanamycin, chloramphenicol and 0.1 mM IPTG and incubated for 24 hours. Lastly, the colonies were picked up and diluted in 20 µl of 1 X PBS in 384 density well plates and spotted on second selection plates with tetracycline, kanamycin, chloramphenicol and 3 different IPTG concentrations designated for each target genes. The entire replica pinning processes was conducted by Singer Rotor HDA.

After 24 hours, images of the plates were obtained by scanner Epson GT-X970. An image processing program that gave numerical values for colony areas on the plate (Takeuchi et al., unpublished data). These raw data were normalized for the plate-by-plate variation in average colony sizes. The genetic interaction scores were calculated by dividing the normalized colony size grown in the lowest IPTG concentration with the colony size grown in the highest IPTG concentration. The top 100 lowest interaction scores were selected and checked for overlapping across all the three replicates. If the interaction appeared two times above in three replicates will be selected for individual spotting check. For the individual spotting check, the double mutants were picked up from the plates, re-grown overnight in 0.1 mM IPTG and the culture was diluted 10^{-1} to 10^{-7} times and spotted on 3 different IPTG concentrations.

3.2.3 Drug sensitivity test

To assess ampicillin and vancomycin sensitivity, overnight cultures were diluted 10^{-1} to 10^{-6} times in 1 X PBS and spotted on LB agar plate containing 3 µg/ml of ampicillin and 200 µg/ml of vancomycin.

3.2.4 Hydroxuyrea (HU) sensitivity test

To assess hydroxyurea sensitivity, overnight cultures were diluted 10^{-1} to 10^{-6} times in 1 X PBS and spotted on LB agar plate containing 2.5 mg/ml, 5 mg/ml and 7.5 mg/ml HU.

3.3 Results

3.3.1 Establishment of the genetic interaction analaysis by spotting method

Genetic interaction of essential genes using a conditional lethal mutant generated by a complementing plasmid has never been reported before. In order to transfer the antibiotic selection marker replacing the essential gene of interest to the recipients bearing individual single gene deletions, the complementing plasmid must be transferred as well to recipients for the complementation of essential gene deletion in the double mutation background. In my study, it has shown that one-step transfer of both essential gene deletion and complementing plasmid simultaneously has poor efficiency. Hence, for the construction of double mutants, complementing plasmid and essential gene deletion are transferred separately to recipients in two different events of mating, which is slightly different from the prototype method wherein only one time of mating is sufficient.

A schematic flowchart illustrating the genetic interaction procedure is shown in Fig. 14. Two steps of mating are taken place for the construction of double mutant of essential gene and non essential gene. The first mating is to transfer the complementing plasmid to the nonessential gene deletions. The second mating is to transfer the antibiotic marker replacing the target essential gene to the nonessential gene deletions carrying the complementing plasmid. Double mutants are constructed systematically on solid media in 3 plates of 1536 high-density arrays covering the genome wide interaction with 3,906 nonessential single gene replacements.
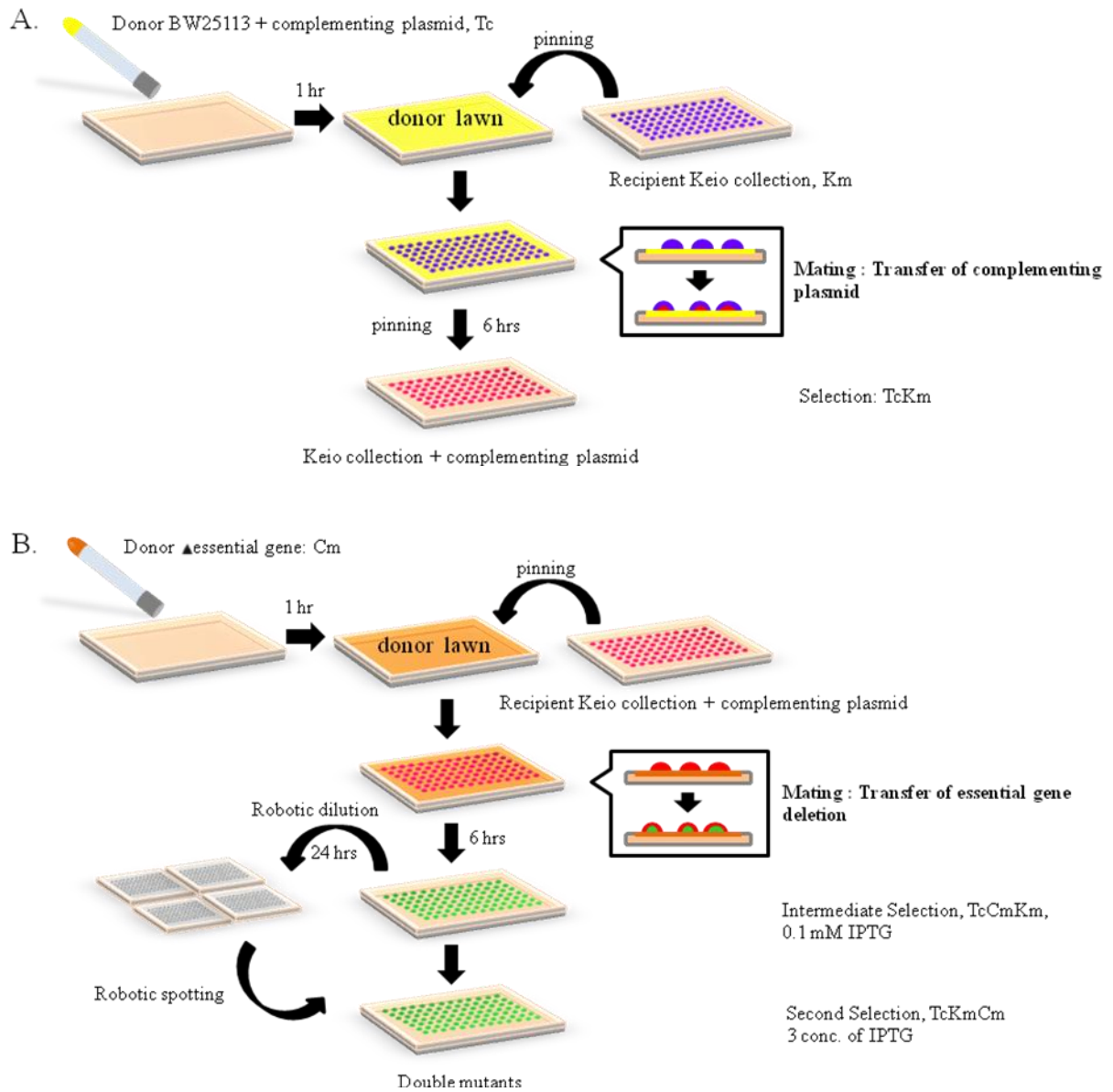
Figure 14: Schematic flowchart illustrating the complete procedure of generating double mutant by the spotting method. Two steps are required to transfer the complementing plasmid which is marked with tetracycline resistance marker and query essential gene deletion which is marked with chloramphenicol resistance cassette to an array of non-essential gene deletions, which are marked with kanamycin resistance cassette for the making of double mutants. (**A**) Complementing plasmid is first transferred from the wild type donor strain carrying the complementing plasmid to the

recipient strains, nonessential gene deletions. (**B**) Donor strain of conditional lethal mutant of query essential gene is crossed with recipients strains of nonessential gene deletions carrying the complementing plasmid that are prepared in the previous step in order to transfer the essential gene deletion marked with chloramphenicol cassette to the recipient strains. Colonies are diluted in 20 µl of 1 X PBS after the intermediate selection before spotted on the second selection plates.

3.3.2 Parameter testing and optimization

3.3.2.1 Intermediate selection

The intermediate selection which selected for only donor (chloramphenicol) or recipient (kanamycin) or no selection pressure (LB) did not yield good results compared to the selection for double mutant (TcKmCm). This is shown in the Fig. 15. The reason behind this observation can be explained by the low number of double mutants survived in the intermediate selection of only donor (Cm) or recipient (Km) or no selection pressure (LB). Hence, after the dilution, the quantity of double mutants spotted on the second selection plate was not sufficient for the subsequent growth.

From the results above, I decided to use the spotting method with the intermediate selection which selected double mutant only (TcKmCm).
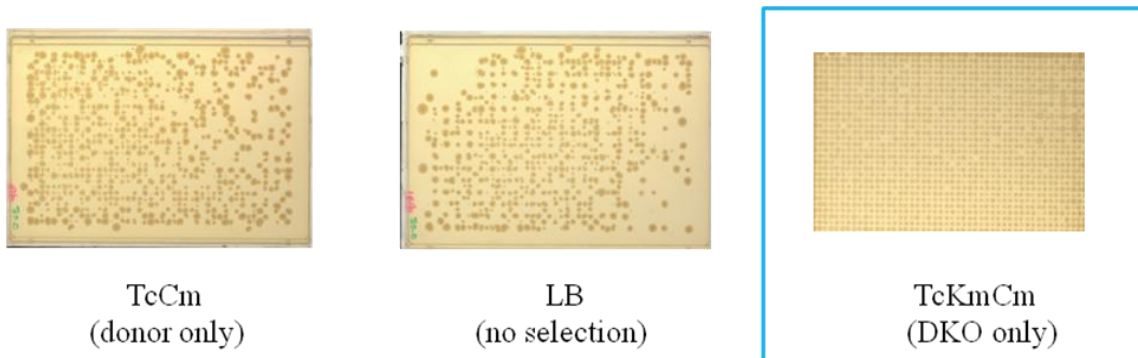
Figure 15: Evaluation of the intermediate selection condition in the spotting method. Selection of double mutants on TcKmCm in the intermediate selection gave the best results for the second selection.

3.3.2.2 Dilution volume

As mentioned above, in the spotting method, colonies of double mutant after the intermediate selection will be diluted in 1 X PBS and the liquid culture will be spotted for the second selection. The suitable dilution volume ought to be determined. I investigated the dilution of double mutants after the intermediate selection by different volumes of 1 X PBS, which were 10 μl, 20 μl and 50 μl. The evaluation was based on the fitness of double mutants in the second selection which were resulted from the different dilutions.

10 μl was too little for the proper dilution, as pins could not thoroughly reach the surface of 1 X PBS due to mechanical limitation, so only small amount of cells diluted into the liquid. Subsequently, it affected the data quality as most of the colonies are not growing well.

Statistically, dilution by 20 µl had better biological reproducibility and showed better IPTG dependency than dilution by 50 µl (Fig. 16). The correlation between replicates diluted by 20 µl was higher as the R^2 was higher, thus representing better reproducibility. Besides, the correlation of double mutants grown in different IPTG concentrations in the second selection was lower in the dilution by 20 µl, meaning that IPTG dependency was more profound in dilution by 20 µl (Fig. 17).

From the results above, I decided to use 20 µl for the dilution of double mutants from the intermediate selection plate.
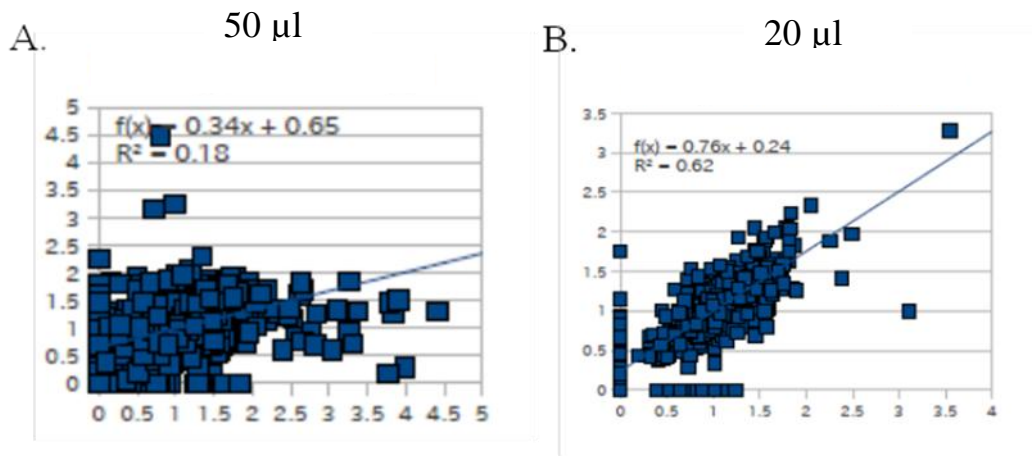
Figure 16: Evaluation of the suitable volume for the dilution of double mutants after the intermediate selection in the spotting method. (A) Dilution by 50 µl gave low reproduciblity between replicates (R^2= 0.18) compared to the (B) dilution in 20 µl (R^2 = 0.62).

50 µl
fitness in 0.1 mM vs 0.05 mM

20 µl
fitness in 0.1 mM vs 0.05 mM

fitness in 0.1 mM vs 0.025 mM

fitness in 0.1 mM vs 0.025 mM
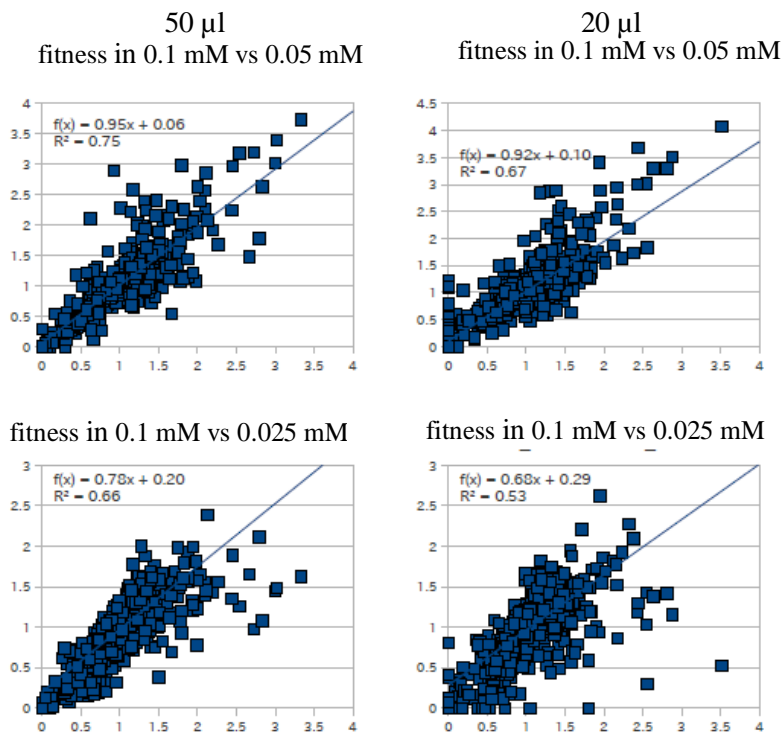
Figure 17: The correlation of the mean fitness of double mutants in different IPTG concentrations. Double mutants in the second selection which were produced from the dilution by 50 µl after the intermediate selection showed higher correlation when they were grown in different IPTG concentrations of 0.1 mM, 0.05 mM and 0.025 mM (R^2=0.75 and R^2=0.66) compared to double mutants which were produced from the dilution by 20 µl, where R^2= 0.67 and R^2=0.53. The lower correlation between different IPTG concentrations suggests the higher IPTG dependency of the double mutants.

3.3.2.3 IPTG concentrations

To generate dosage repression of essential genes in the genetic interaction analysis, suitable IPTG concentrations for each gene were determined by testing a range of IPTG concentrations and evaluating the overall results and effects (Fig. 18). For *dnaN*, the concentrations used were 0.1 mM, 0.05 mM and 0.025 mM. For *ftsW* , were 0.1 mM, 0.05 mM and 0.01 mM. For *yrfF,* were 0.1 mM, 0.05 mM and 0.02 mM and lastly for *yjgP* were 0.1 mM, 0.05 mM and 0.03 mM. The 3 IPTG concentrations selected for the genetic interaction analysis were ample to maintain the viability of the conditional lethal mutants of target essential genes, so the sickness or lethality effect was confirmed contributed by the double mutant.
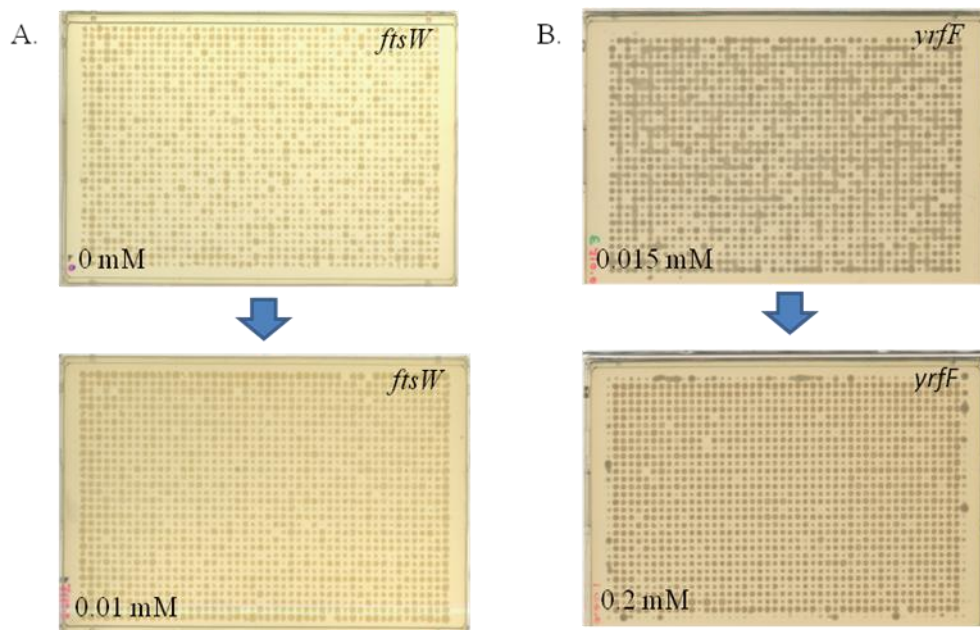
Figure 18: An example of the screening for the 3 suitable IPTG concentrations for the genetic interaction analysis. (**A**) Trials using different IPTG concentrations were tested. 0.01 mM IPTG was the lowest IPTG concentration which was suitable to observe genetic interaction for *ftsW*. (**B**) Increment to 0.2 mM IPTG was determined as the lowest IPTG concentration for the genetic interaction analysis for *yrfF*.

45

3.3.2.4 Imaging and quantification

An in-house plate-scanning system for the real-time monitoring of colony's growth has been developed. Inoculated plates are incubated in scanners which are stored inside an incubator. Images of plates are automatically scanned and saved every 30 minutes. The system offers the advantage of monitoring growth dynamics of colonies, and summarizing the growth in a more informative way than just the final phenotype at stationary phase (Takeuchi et al, unpublished data).

Unfortunately, when the colonies of double mutant measured in the plate-scanning system, uneven growth was being observed repeatedly for the double mutants consist of a single conditonal lethal mutation of essential gene that shows strong IPTG dependency, for example like *dnaN* (Fig. 19). Essentially, the aeration of the incubator and the temperature control were less optimum than the normal incubator. Strains that required more condusive environment (high IPTG dependent single mutant) might require more stringent condition to grow.

Due to the failure of growing the double mutants in the plate-scanning system as mentioned above, colonies of double mutant were scanned and measured at fixed time point after 24 hours of incubation instead. Plate images were processed and the fitness (area of the colony) of each double mutant was quantified by programming.
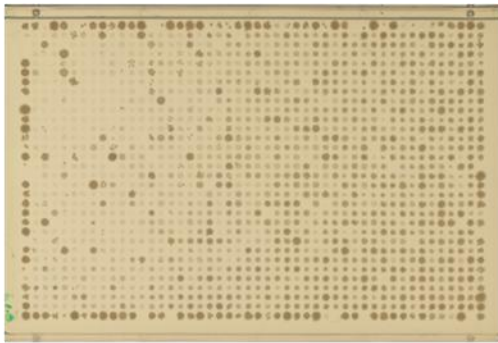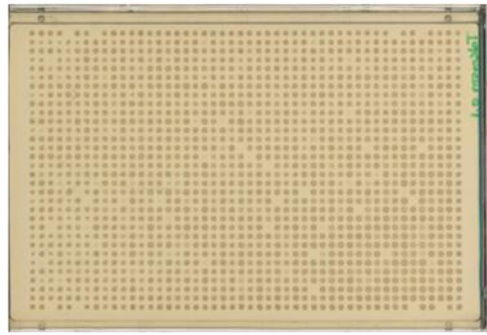
<center>*dnaN*   *ftsW*</center>

Figure 19: Double mutants of *dnaN* failed to grow when incubated in the plate-scanning system. Double mutants of *ftsW* which have lower IPTG dependency showed better growth when incubated in the plate-scanning system.

3.3.3 Identification of genetic interaction

The double mutants of potentially interacting genes were picked up from the plates of high throughput genetic interaction analysis, and grown overnight. The overnight cultures were diluted and spotted onto agar plate supplemented with different IPTG concentration to confirm the growth defect of the double mutants. To ensure the defect was resulted from genetic interaction, phenotypic deviation from the single gene mutant was compared (Fig. 20).
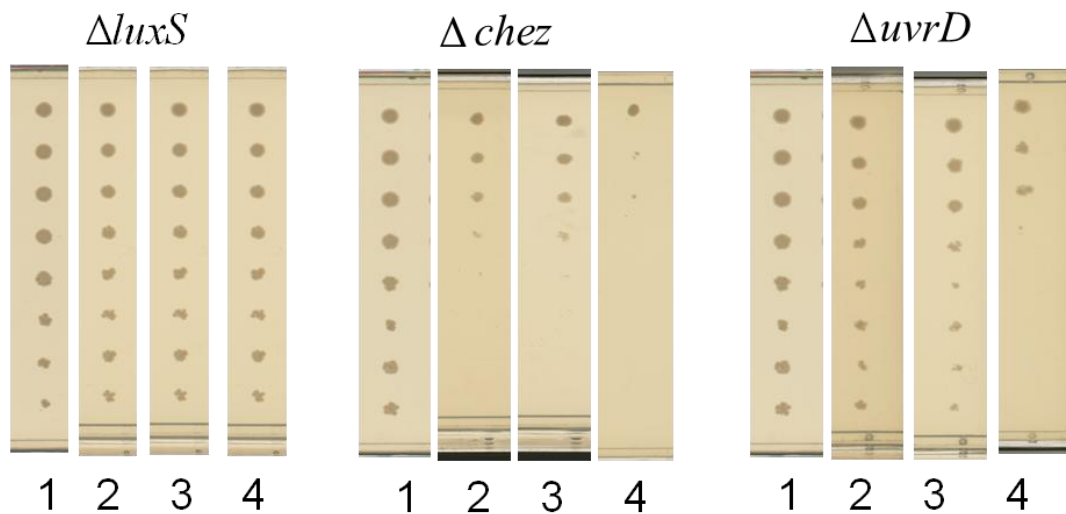
Figure 20: An example of the comparison of phenotypic deviation of double mutant from a single nonessential deletion by spotting the serially diluted liquid culture onto agar plates supplemented with different IPTG concentrations. Lane 1 shows single nonessential gene deletions, lane 2 shows double mutants grew in 0.1 mM of IPTG. Lane 3 shows double mutants grew in 0.05 mM of IPTG and lane 4 shows double mutants grew in 0.025 mM of IPTG. *luxS* shows no interaction with *dnaN* as the fitness of single mutation and double mutation are equal. *cheZ* interacts with *dnaN* in all 3 IPTG concentrations. *uvrD* interacts with *dnaN* in IPTG dependent manner, when the IPTG concentration was decreased to 0.025 mM, the double mutant exhibits sickness phenotype which is not observed in 0.1 mM and 0.05 mM IPTG.

All the listed interactions were validated interaction (Fig. 21), proven by the individual confirmation test as explained in Figure 20. In general, the interactions were gene specific with some overlappings. *dnaN,* is a beta subunit of DNA polymerase III responsible for DNA replication (LaDuca et al., 1986). 3 genes involved in DNA related function, *seqA*, *xerC* and *uvrD* were isolated. Interestingly, 2 genes in salvage pathway, *cmk* and *cdd* were found interacting with *dnaN*. Salvage pathway is an important process for synthesizing new nucleotides using recovered intermediates like bases and nucleoside from degradation of RNA and DNA. In addition, *dnaN* also showed interaction with some mRNA degradation genes like *pnp*, *ssrA*, *deaD*. From here, we postulated that *dnaN* mutant could be HU sensitive because *dnaN* might have some deficiency in dNTP production which caused sickness or lethality when combined with the dNTP synthesis genes.

It was clearly shown that the sensitivity to HU (Fig. 22). ATP-DnaA initiates replication at the origin of replication and represses the RNR. Upon the binding of beta clamp encoded by dnaN, ATP-DnaA is hydrolysed to inactive form ADP-DnaA. This conversation increases the expression of RNR to prepare for the elongation in replication. Shortage of DnaN, which prevents the conversion of ATP-DnaA to ADP-DnaA may disrupt the depression of RNR. A lack of dNTPs could lead to the HU sensitivity and replication fork arrest.

yjgP or known as *lptF* is a lipopolysaccharide transporter located in inner membrane (Ruiz et al., 2008), transporting precursors for the outer membrane LPS synthesis, in cooperation with the *lptABCFG* complex. Deletion of *yjgP* caused increased outer membrane permeability (Ruiz et al., 2008). Genetic interacting genes with *yjgP* enriched in genes which maintain membrane integrity like *tolAQR* (Lloubes et al., 2001) and *rpmFJ* (Nakayashiki and Mori, 2013). These explain the functional relationship with *yjgP*.

FtsW, a cell division protein, recently been discovered a new function as flippase of lipase II required for peptidoglycan synthesis (Mohammadi et al., 2011). It was

reasonable this gene interacted with *ldcA* and *envC* which both also linked to peptidoglycan related function (Templin et al., 1999 & Bernhardt and de Boaer, 2004).

yrfF, an unknown gene, judging from the interacting genes enriched in membrane related genes, suggesting that this gene could be important for outer membrane or peptidoglycan related function. To test the hypothesis, drug sensitivity test was performed. Ampicilin was used to test against integrity of peptidoglycan; whereas vancomycin was used to test against integrity of outer membrane. From the results, the conditional lethal mutant of *yrfF* is sensitive to vancomycin (Fig. 23), suggesting the mutant has defect in LPS and outer membrane permeability.

Figure 21: Genetic interaction network of 5 essential query genes by spotting method.
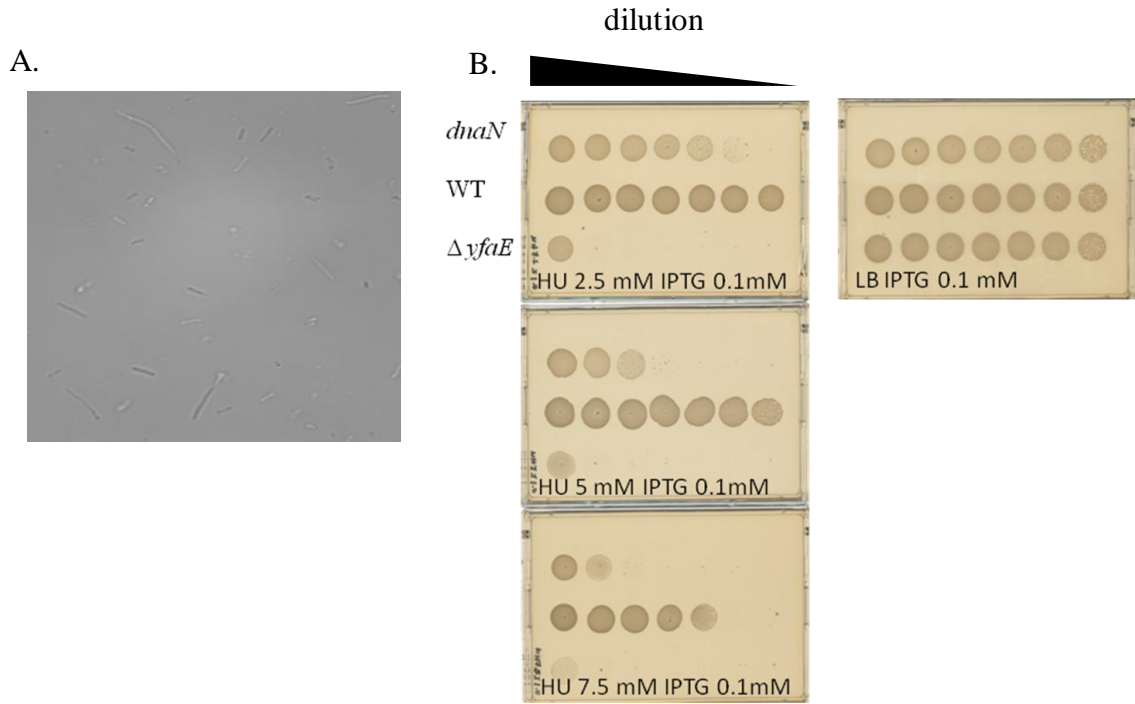
Figure 22: Phenotype of *dnaN* conditional lethal mutant. (**A**) Morphology of *dnaN* conditional lethal mutant under the microscope is filamentous. (**B**) HU sensitivity of *dnaN* conditional lethal mutant. *yfaE* deletion is a positive control.
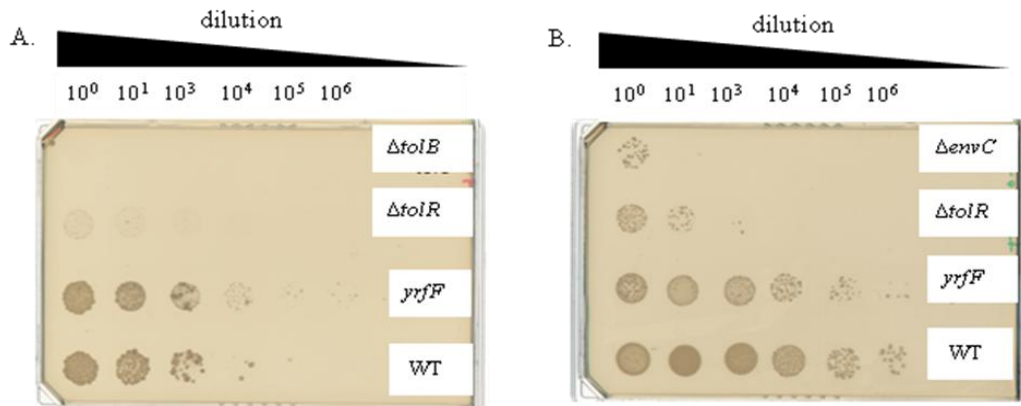
Figure 23: Antibiotic sensitivity test of *yrfF* conditional lethal mutant. (**A**) 3 μg/ml of ampicillin was used. (**B**) 200 μg/ml of vancomycin was used. *yrfF* condtional lethal mutant is resistant to ampicilin as wildtype and is sensitive to vancomycin. Deletions of *tolB*, *tolR*, and *envC* are positive control of sensitivty to ampicilin and vancomycin.

3.4 Discussion

The method has sucessfully revealed a network of genetic interactions for the 5 target essential genes. Nevertheless, the problem of the established spotting method is low global reproducibility while variations are inevitable during the experiment, causing some important interaction might be missed out during the analysis.

First and foremost, even the dilution of colonies was performed robotically, there were some limitations in controlling the starting inoculum for the second selection. The amount of cells being picked up from the intermediate selection plate for dilution and after the dilution, the number of cells from the diluted liquid culture to be spotted on the second selection plate was beyond control. Due to the difference of starting inoculum in each independent experiment, variations between replicates can be created by the strong neighboring effect, where the fast growth mutants supress the growth of the neighboring slow growth mutants. Besides, false positives in which the retarded growth was actually caused by the technical problem during the dilution step, is the another major attribution of variances. In addition, of all the multiple times of replica pinning steps throughout the procedure, variations can be introduced as well. The problem of inconsitency of starting inoclum could be solved by monitoring the maximum growth rate of colonies using the plate-scanning system for the real-time quantification of colony's growth. However, the double mutants could not grow in such condition.

Overall, the method works well to isolate potential interacting genes, and due to the high sensitivity, it is a very good method for screening. However, for the systematic analysis point of view, the method has high variations caused by the technical limitations which will hinder the inspection of the comprehensive genome wide and global level analysis.

CHAPTER 4

Establishment of genetic interaction analysis by stamping method and pilot test of five essential query genes

4.1 Introduction

Genetic interaction analysis using the spotting method as discussed in chapter 3 was useful for screening of the potential interacting candidates. However, it possesses some setbacks for the global analysis due to the part that overall lower reproducibility between replicates.

The stamping method was preferred for global statistical analysis of genetic interaction though it has lower sensitivity relatively to the spotting method, since it avoids some technical limitations that will introduce variations, and thus high reproducibility is expected. High reproducibility between replicates is very vital because it is the key issue of the reliability of high throughput analysis.

In this chapter, I will focus on the establishment of the genetic interaction analysis by stamping method and the pilot test demonstrated in the five essential query genes.

4.2 Materials and methods

4.2.1 Establishment of genetic interaction analysis by stamping method

The protocol was the same as the spotting method, with only few modifications. The mating time for essential gene deletion transfer was 12 hours instead and the intermediate selection was performed on tetracycline, chloramphenicol and 0.1 mM IPTG. After the intermediate selection, the colonies were stamped on second selection plates

with tetracycline, kanamycin, chloramphenicol and 3 different IPTG concentrations designated for each target genes. The entire replica pinning processes was conducted by Singer Rotor HDA.

After 24 hours, images of the plates were obtained by scanner Epson GT-X970. An image processing program that gave numerical values for colony sizes on the plate. These raw data were normalized for the plate-by-plate variation in average colony sizes.

4.2.2 Strains and growth conditions and plasmid

BW25113($lacI^q$ $rrnB_{T14}$ $\Delta lacZ_{WJ16}$ $hsdR514$ $\Delta araBAD_{AH33}\Delta rhaBAD_{LD78}$) is derivative of the F$^-$, λ^-, *E. coli* K-12 strain BD79 and have no other known mutations. Nonessential gene deletions were grown in LB containing 30 µg/ml of kanamycin. Conditional lethal mutants of essential genes were grown in LB medium containing 12.5 µg/ml of tetracycline, 25 µg/ml of chloramphenicol and 0.1 mM IPTG at 37 ℃. The *sulA*-GFP plasmid, pTN175 for the measurement of SOS response was obtained from Toru Nakayashiki (Nakayashiki and Mori, 2013).

4.2.3 Conjugation in liquid culture

Overnight culture of donor strain of *dnaN* conditional lethal mutant was subcultured by diluting 1:25 in LB medium and incubated for 3 hours at 37 ℃ without shaking. Overnight culture of recipient strains (nonessential gene deletions of *araC*, *feoA*, *ompT*, *rnr*, *seqA*, *ygeP*, *ybaZ*, *rnr*, *uvrD*) were subcultured by diluting 1:5 in LB medium and aerated at 37 ℃ until used. 200 µl of each donor and recipient strains were mixed in test tube and let sit in a 37 ℃ water bath for 90 minutes. 2 ml of LB medium was added to test tube and aerated for an additional 1-2 hours at 37 ℃. Mating mixtures were diluted 10^{-1} to 10^{-6} times in 1 X PBS and spotted on LB agar plate containing tetracycline, chloramphenicol, kanamycin and 0.1, 0.05 and 0.025 mM IPTG at 37 ℃.

4.2.4 Data analysis for pilot test of genetic interaction analysis

Images of each plate were scanned using an EPSON GT-X970 scanner. Raw colony densities were quantified from plate images using an in-house developed image analysis program (Takeuchi et al, unpublished data), producing a 32x48 matrix of numerical values. Data from plates was categorized as either a single knockout plus plasmid (SKOp) or as a double knockout (DKO). For each 32x48 matrix of colony densities, we define the kth "layer" as the union of the elements of the kth and $(32-k)^{th}$ rows and the kth and $(48-k)^{th}$ columns, excluding elements from layers 1, 2, …, k-1. Because colony sizes tend to be larger in outer layers due to a position effect (Butland et al., 2008), we normalized entries in layers 1, 2, and 3 by multiplying the colony density by the median colony across all plates in the category (SKOp or DKO) divided by the mean colony central density within only the particular layer the data point is located in, averaged across all plates in the category. We next normalized all colony sizes within each data matrix by dividing by the mean value of the values in the matrix.

A single genetic interaction score was calculated for each DKO combination by dividing the mean normalized colony density of the DKO by the mean normalized colony central density of the corresponding SKOp. We assume the genetic interaction scores of non-interacting gene pairs will form a normal distribution. We estimate the parameters of this distribution using a least-squares fit of the graph of a normal probability distribution function with the density plot of the genetic interaction scores of each query gene. Density plots are obtained using the density default function in R. Fitting only considers the central 50% density of the genetic interaction scores, with the density plot scaled so that the area under curve is the same for density plot and the fitted curve. We then separately consider for negative and positive interactions the false discovery rates calculated by this model when using various cutoffs for significance. We chose cut-offs for positive and negative interactions that result in a predicted 10% false discovery rate.

We filtered out "genetic interactions" that are non-specific to individual query genes by ignoring genes that have significant genetic interaction scores for three or more

of the query genes in downstream analysis. Additionally, we filtered out interactions with genes located within 35kb of the query, since such "interactions" are often due to overwriting of the Keio knockout and antibiotic resistance gene during recombination with DNA carrying the knockout of the query gene (Butland et al, 2008). Functional enrichment was tested for sets negative interactions and sets of positive interactions using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.7 (Dennis et al., 2003).

4.2.5 Flow cytometry analysis

The *dnaN* conditional lethal mutant harboring a *sulA*-GFP plasmid, pTN175, was grown overnight to stationary phase and diluted 1000x with 1x phosphate-buffered saline. Data was collected using Accuri^{TM} C6 flow cytometer (Becton Dickinson) with a 488-nm argon laser and a 515- to 545-nm emission filter (FL1) at high flow rate.

4.3 Results

4.3.1 Genetic interaction analaysis by stamping method

  The procedure is similar to the spotting method, the only modifications are the mating time to transfer the essential gene deletion, the intermediate selection condition, and the starting inoculum for the second selection (Fig. 24).
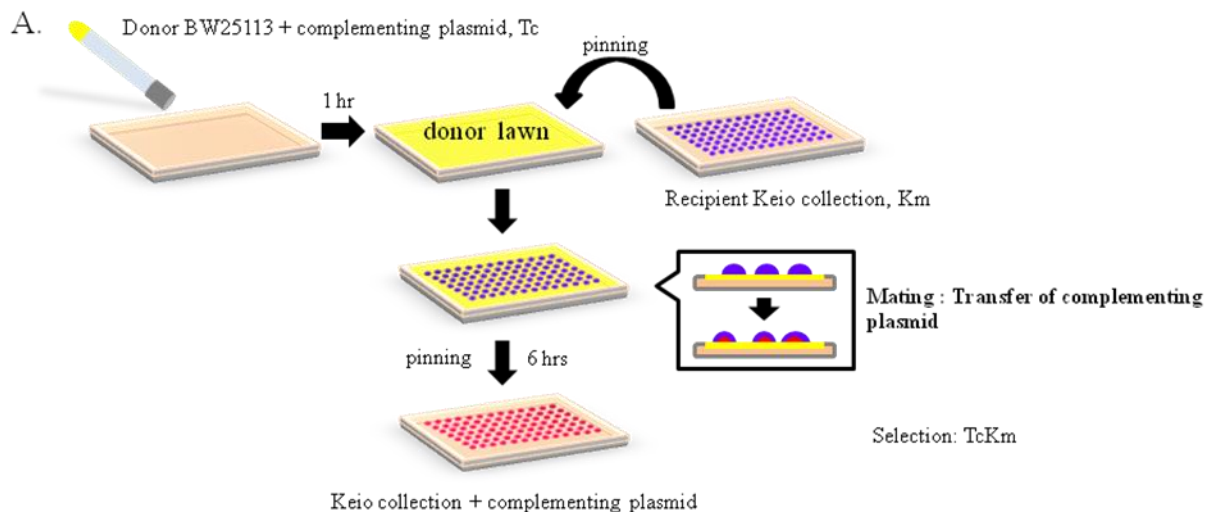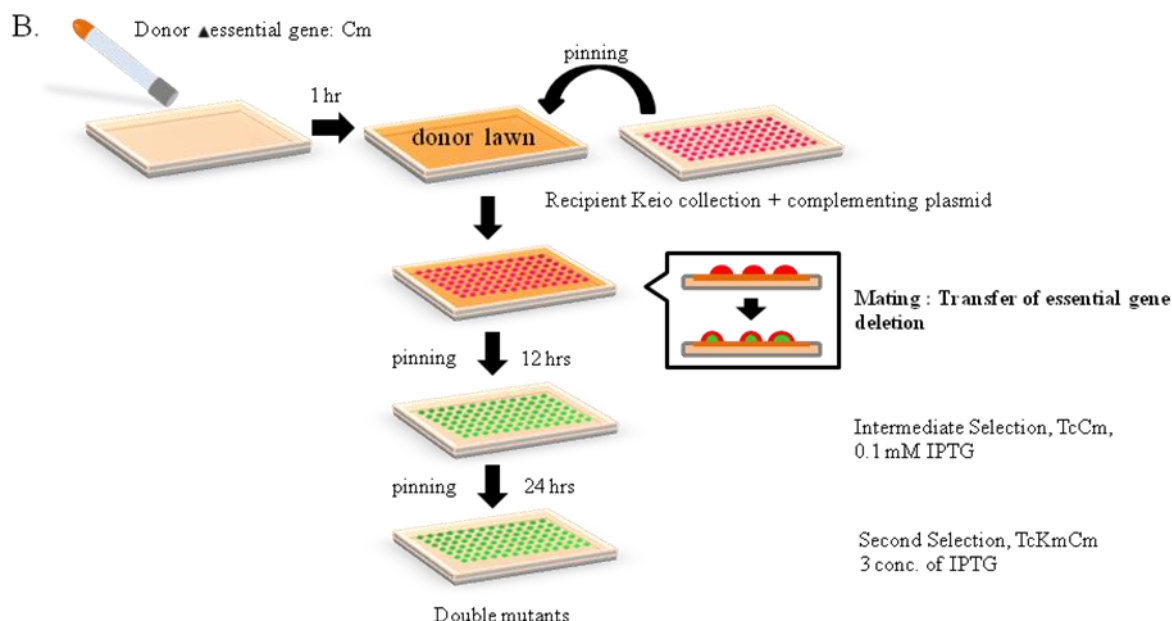
Figure 24: Schematic flowchart illustrating the complete procedure of generating double mutant by the stamping method. Two steps are required to transfer the complementing plasmid which is marked with tetracycline resistance marker and query essential gene deletion which is marked with chloramphenicol resistance cassette to an array of non-essential gene deletions, which are marked with kanamycin resistance cassette for the making of double mutants. (**A**) Complementing plasmid is first transferred from the wild type donor strain carrying the complementing plasmid to the recipient strains, nonessential gene deletions. (**B**) Donor strain of conditional lethal mutant of query essential gene is crossed with recipients strains of nonessential gene deletions carrying the complementing plasmid that are prepared in the previous step in order to transfer the essential gene deletion marked with chloramphenicol cassette to the recipient strains. Colonies are replicated onto the second selection plates supplemented with different concentrations of IPTG after the intermediate selection.

59

4.3.2 Parameter testing and optimization

Previously confirmed interacting genes with *dnaN* by the spotting method, were selected as positive control for the optimization in stamping method. They are *seqA*, *rnr*, *ybaZ* and *ypfN and uvrD*. Non interacting genes (*araC*, *feoA*, *ompT*, *prmB,ygeP*) were used as negative control (Fig. 25).

Double mutants of these genes combinations were produced under the different parameters, including plasmid transfer time, essential gene deletion transfer time or mating time, and intermediate selection condition and time. T test was used to access the fitness difference between the group of interacting gene pairs (*dnaN-seqA*, *dnaN-rnr*, *dnaN-ybaZ*,*dnaN-ypfN*,*dnaN-uvrD*) and non interacting gene pairs. Log of P value was calculated from the T test, and the smaller the P value is the greater the difference of fitness between the interacting group and non interacting group.

It is noted that *ypfN* is supposed to be interacting with *dnaN* as it had been validated several times; however the double mutants of these genes produced by the stamping method in all the parameters did not show any phenotypic defect (Fig. 25, B). This was contradicting with the previous results by the spotting method. The spotting experiment was conducted again and it was confirmed that *ypfN* did interact with *dnaN* (data not shown). This result corroborated that the spotting method was more sensitive than the stamping method as discussed. Overall the stamping method is not perfect but capable of identifying most of the genetic interactions with the positive control genes so it can be used for the further analysis.
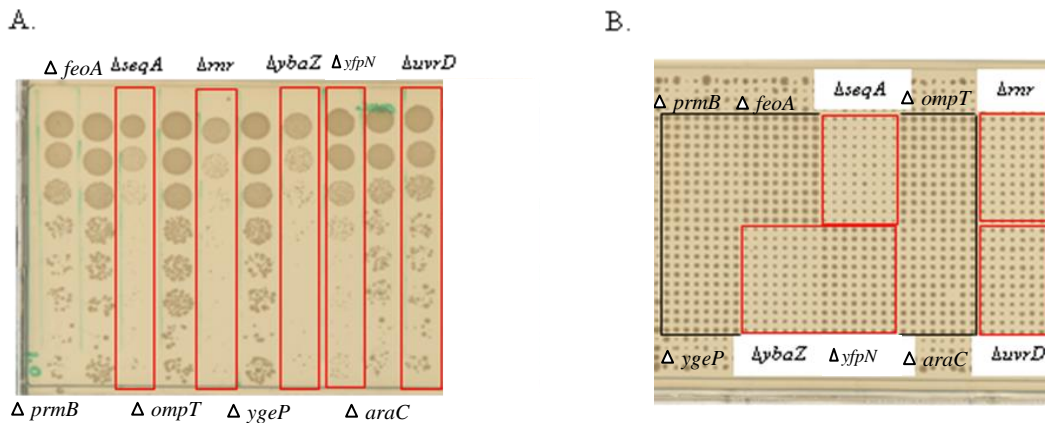
Figure 25: An example of the optimization procedure in the stamping method. (**A**) Double mutants of *dnaN* conditional lethal mutant and deletions of selected positive control genes (red column) were made by conjugation in liquid to confirm again the positive control genes are the real interacting genes with *dnaN* before the optimization. (**B**) An example of the optimization of genetic interaction analysis in different conditions. Colonies in red column are double mutants of *dnaN* conditional lethal mutants and deletions of positive control genes in which genetic interactions are expected. Colonies in black column are double mutants of *dnaN* conditional lethal mutant and deletions of negative control genes in which genetic interactions are not expected. The double mutants of this plate were generated under the intermediate selection of chloramphenicol and tetracycline at 0.1 mM of IPTG and 24-hour of incubation.

4.3.2.1 Intermediate selection

According to the graph (Fig. 26), intermediate selection of only the donor which was on chloramphenicol, and after 24 hours of incubation before the second selection on 0.025 mM IPTG yielded the best results in terms of P value. P value was calculated based on the difference of average fitness of the interacting gene pairs and (*dnaN* with the positive control genes) the non-interacting gene pairs (*dnaN* with the negative control genes). The larger the P value means the smaller the difference of fitness between

interacting and non interacting gene pairs. Using the condition of chloramphenicol for first selection and 24 hours of incubation, the ideal plasmid transfer time and essential gene deletion transfer time or mating was evaluated based on P value as well.
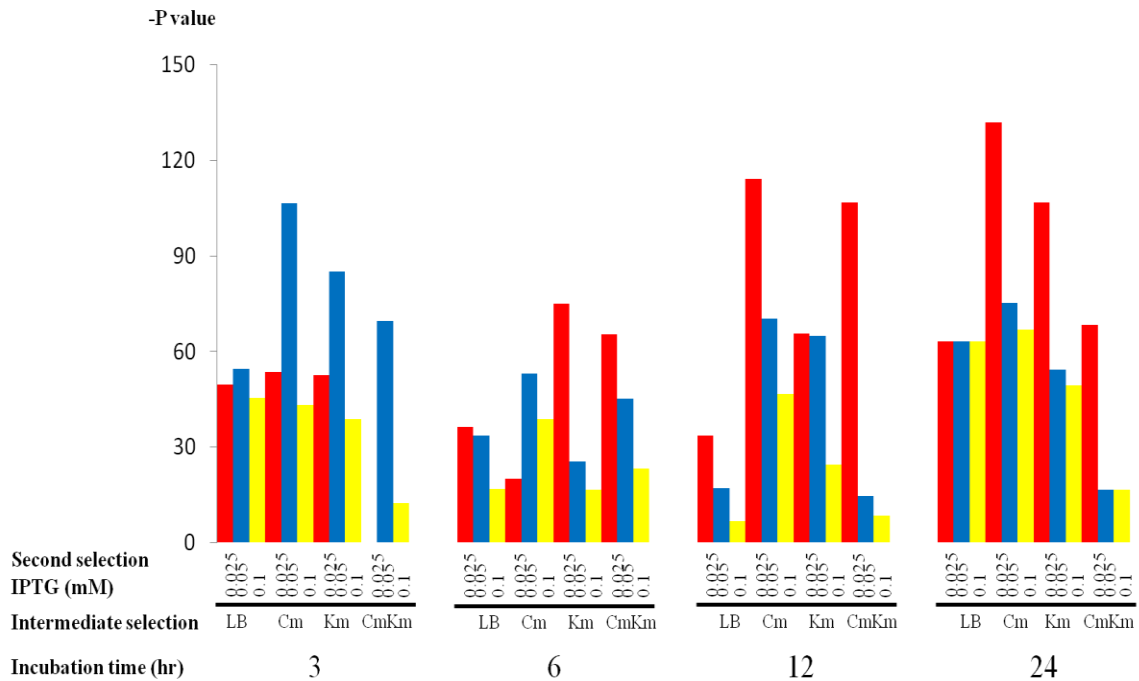
Figure 26: Comparison of the intermediate selection condition and also the incubation time for stamping method using *dnaN* knockdown as query gene. Intermediate selection on chloramphenicol and 24 hours of incubation provides the lowest P value, which is a better selection condition for genetic interaction analysis. Red bars represent second selection on 0.025 mM of IPTG, blue bars represent second selection on 0.05 mM of IPTG and yellow bars represent second selection on 0.1 mM of IPTG.

4.3.2.2 Complementing plasmid and essential gene deletion transfer time

6 hours of mating time for the transfer of complementing plasmid and 12 hours of mating time for the transfer of essential gene deletion were suggested from the data (Fig. 27). Shorter mating time for the transfer of complementing plasmid was expected to be better. It is because it can reduce the possibility of transfer of wild type copy of non essential gene from the Hfr donor strain carrying the complementing plasmid to the recipients of Keio deletions. 24 hours of mating time for the transfer of essential gene deletion produced the highest P value as predicted. Longer time might lead to IPTG accumulation, or adaptation of the double mutant, causing the interacting genes' sick phenotype less detectable.
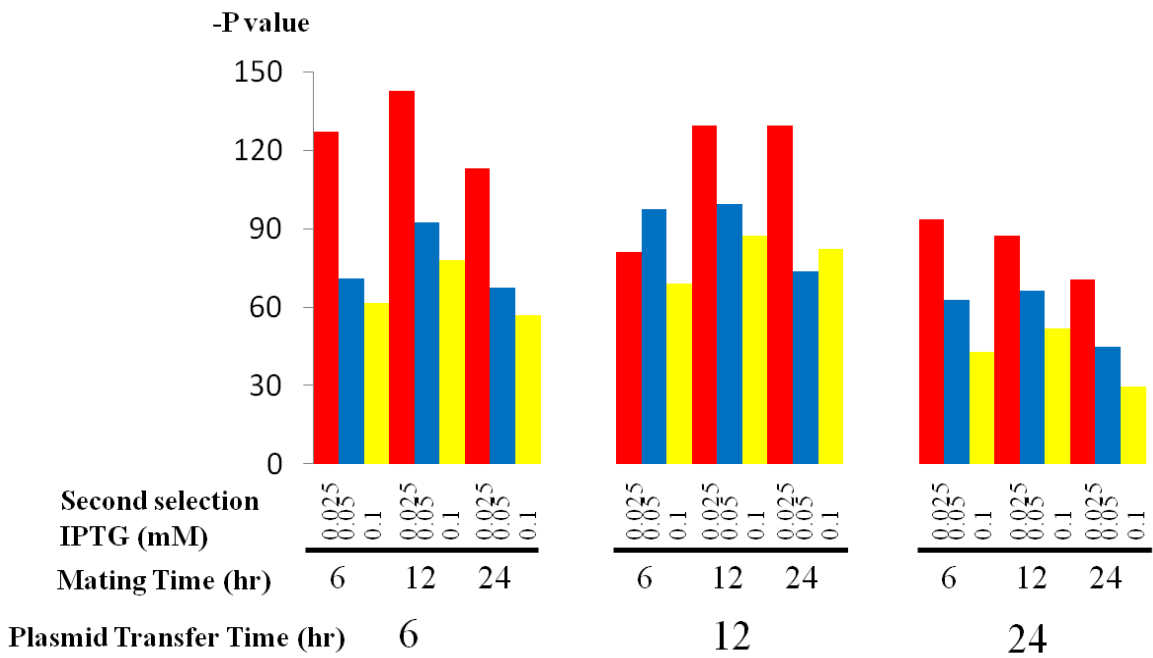
Figure 27: Comparison of different mating time to transfer essential gene deletion on chromosome and also complementing plasmid using *dnaN* conditional lethal mutant as a query strain. 6 hours of plasmid transfer time and 12 hours of mating time to transfer essential gene deletion showed the best results as indicated by the lowest P value. Red bars represent second selection on 0.025 mM of IPTG, blue bars represent second selection on 0.05 mM of IPTG and yellow bars represent second selection on 0.1 mM of IPTG.

4.3.2.3 Imaging and quantification

The plate-scanning system was used to capture the time-series growth of double mutants every 30 minutes produced by the stamping method. The growth of colonies was more stable and consistent when the growth was determined after 24 hours of incubation at fixed time point compared to the time-series by the plate-scanning system (Fig. 28 and

Fig. 29). The result is concurrent with the results of the spotting method where fix-time point measurement is better than time-series measurement.

A.

B.
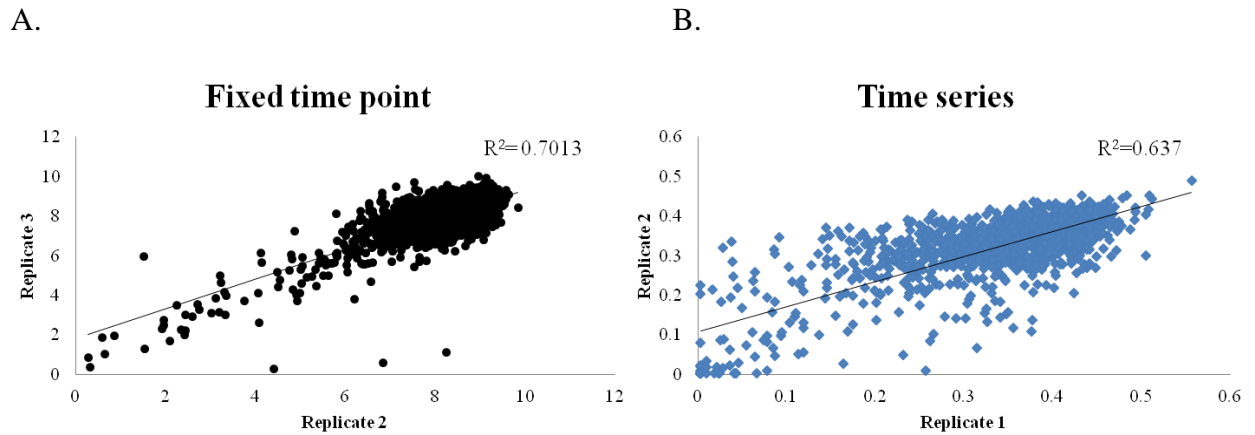
Figure 28: Fitness correlation of replicates measured by two different methods. (A) Fixed time point in normal incubator or (B) time-series in the plate-scanning system. Fixed time point measurement produced the better results in terms of reproducibility between replicates as R^2= 0.7013 versus R^2=0.637.

A.

B.
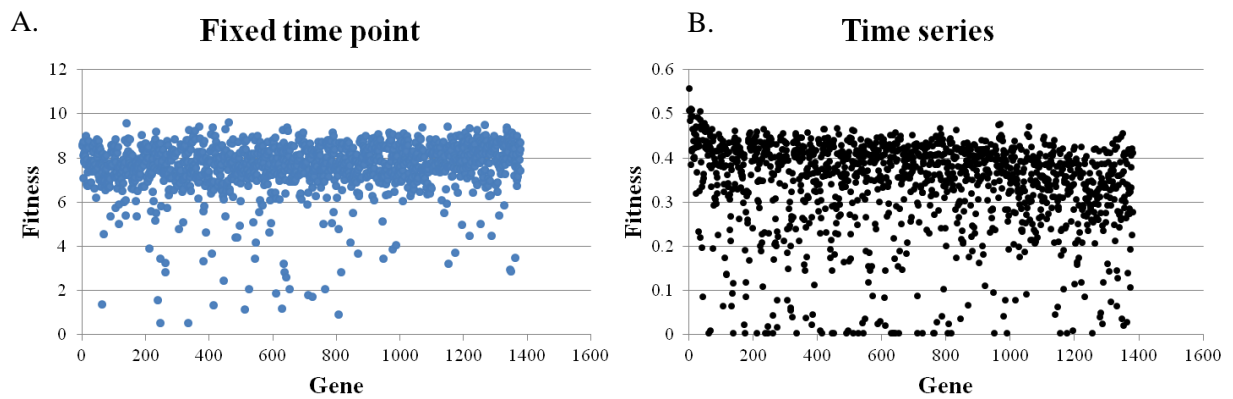
Figure 29: Fitness of double mutants measured by two different methods. (A) Fixed time point in normal incubator or (B) time-series in the plate scanning system. Fixed time point measurement produced more stable and less noisy results.

4.3.3 Computational processing

After 24 hours, images of the plates were obtained by scanning. An image-processing program computed numerical values for the density of colonies on the plate. These raw data were normalized for the plate edge effect (Butland et al., 2008) and for plate-by-plate variation in average density. The genetic interaction scores (GI scores) we next calculated controlled for differences in density due to the Keio knockout and the presence of the plasmid by normalizing average double knockout (DKO) colony densities with colony densities of the single Keio knockout plus the query-specific plasmid (SKOp). Normal distribution parameters were estimated for modeling the distribution of non-interacting gene pairs. Values deviating from this distribution, as determined by cut-offs of the genetic interaction score (using FDR=0.1), were designated as either negative interactions (low GI score) or positive interactions (high GI score). Genes that have negative interactions or positive interactions with three or more query genes were filtered out of the lists of specific genetic interactions (Supplementary Table 1). The mutants of these common interacting genes are usually slow growth and thus sensitive to random perturbation or mutations that affect the conjugation and recombination efficiency. Unfortunately, reported negative interactions of genes located nearby the query gene on the chromosome are not reliable since they could result from overwriting of the kanamycin resistance gene during incorporation of the DNA carrying the deletion of the query gene by recombination (Butland et al., 2008). Therefore, genes located within 35kb of the query gene were not included in lists of significantly interacting genes even if their GI score qualifies as significant, and they were excluded from further analysis to eliminate linkage effect.

4.3.4 Pilot Study of Five Genes

In our pilot study of the genetic interaction method using pFE604T, we studied five essential query genes: *dnaN*, *ftsW*, *trmD*, *yjgP*, and *yrfF*. *dnaN* is the beta subunit ofDNA polymerase III, *ftsW* is involved in cell division, *trmD* is required for tRNA methylation, *yjgP* is a LPS transporter and *yrfF* has unknown function. There were a large number of genes found genetically interacting with the 5 query essential genes in the pilot study (Supplementary Table 2). Each screen resulted in an average of 77 interactions. Thus, these essential genes are required for buffering many cellular processes.

Functional enrichment analysis (Table 2) showed some expected and unprecedented regulation mechanisms for functions of the query genes. *dnaN*, showed negative interactions with genes that are enriched in DNA-related functions. *yjgP*, a gene that encodes a LPS transporter, shows negative interactions with genes enriched in lipopolysaccharide core region biosynthetic process. Other observations of functional enrichment are not readily apparent from the primary function of the essential query gene.Some interactions may be due to secondary effects caused by low levels of the essential protein.

We found that the *recC* and *dnaT* genes, which are important for the repair of double strand breaks in DNA by recombination, show negative interaction with *dnaN*. This result suggests that mutation in *dnaN* might result in double stand breaks, resulting in induction of the SOS response. To investigate this hypothesis, a plasmid expressing a GFP-*sulA* fusion protein (pTN175) was used for monitoring the SOS response (Nakayashiki & Mori, 2013). We found that the fluorescence signals of *sulA*-GFP in the *dnaN* conditional lethal mutant are about 2020 ± 20 compared to 670 ± 15 in wild type, suggesting an approximately 3-fold increase of SOS response. The induction of SOS response was further evidenced by the observation of filamentous morphology under the microscope of the cells by microscopy.

Query genes	Functional enrichment of interacting genes	Type of interaction
dnaN	Peptidoglycan based cell wall Organelle envelope	Positive
dnaN	Homologous recombination Pyrimidine metabolism DNA replication	Negative
ftsW	Oxidative phosphorylation Organelle envelope Quinone TCA cycle	Negative
trmD	Oxidative phosphorylation Organelle envelope Purine nucleotide biosynthetic process	Positive
yjgP	Iron sulfur protein ATP-biosynthesis	Positive
yjgP	Lipopolysaccharide core region process	Negative
yrfF	Organelle inner membrane Cell membrane	Positive
yrfF	Organelle envelope RNA degradation	Negative

Table 2: Functional enrichment analysis of the interacting genes with essential query genes.

4.4 Discussion

In short, the best experimental parameters for the genetic interaction analysis of the stamping method are, 6 hours of plasmid transfer time and 12 hours of mating time to transfer essential gene deletion, first selection on chloramphenicol containing LB plate and incubation for 24 hours before the second selection. The biological reproducibility using stamping method was satisfactory, confirming the reliability of the method. Even the sensitivity of the stamping method is less than the spotting method as pointed out, stamping method could confirm most of the interactions determined previously by spotting method in the optimization process.

Thus, pilot test for five essential query genes was conducted using the stamping method with the optimized condition as discussed above.

Our system was demonstrated in five essential query genes encoding proteins involved in different cellular functions. We found that these essential genes interact with high number of nonessential genes with similar or dissimilar functions. In the future, it is readily possible to obtain lists of genetic interactions for all 303 essential genes of *E. coli* using the pFE604T plasmid. These lists can serve as an interaction catalogue for biologists to look for potential candidates to study further. Also, the construction of a library of conditional lethal mutants with down regulated of essential query gene will be a valuable experimental resource for the *E. coli* community.

Genetic interaction analysis of three out of our five query genes have been studied previously using SPA-tagging (Butland et al., 2008; Babu et al., 2011). Surprisingly, there is no overlap between the genetic interaction lists of these studies. An explanation for the lack of consistency might be that the perturbation of the essential gene used in this study is vastly different than the previously used perturbation of SPA-tagging. Furthermore, the laboratory environment external to experiment might affect the outcome of the genetic interaction analysis (Michaut and Bader 2012). Although the two methods

do not produce reproducible lists of interactions, the lists may be thought of as complementary and providing insight into different dimensions of genetic interaction.

The average genetic interactions per query gene for essential genes in this study are approximately 77, as compared to 20 in the non-essential genes (Butland et al., 2008). Thus, essential genes might be much more functionally involved in the cellular processes than what we have thought.

CHAPTER 5

Final discussion

E. coli, due to its simplicity and amenable to manipulation, it has been used as a premier model to help scientists to understand life's processes. Although this single cell organism cannot be used to directly study the manifestation of human disease, knowledge gained in the *E. coli* model can often be applied to homologous proteins in more complex higher organism (Rea et al., 2010). Jacques Monad, a French Nobel Laureate who worked on *E.coli* stated that 'Tout ce qui est vrai pour le Colibacille est vrai pour l'éléphant' (All that is true of *E. coli* is also true for the elephant) (Friedmann, 2004). This widely quoted epigram reflects the enormous impact of research based on *E.coli* in molecular biology. Even with the great versatility for biological application, *E. coli* K-12, the most widely studied organism is estimated that almost half of the genes have yet experimentally characterized (Mori et al., 2000). Consequently, *E. coli* will be a more valuable resource for studying cellular processes once they are thoroughly understood (Hunter et al., 2008).

In post genomic era, effort toward a better understanding of living organism can be achieved by systems biology approach. Systematic genetic interaction analysis has emerged as an informative tool to map the genetic cross talks and cellular architecture in several model organisms. These interaction networks have shed light on the global modular organization of gene products and protein complexes in an organism (Babu et al., 2011). Most importantly, the power of genetic interaction is not only confined to studying the organism of interest, indeed, cross species approach to identify potential clinical relevant genetic interactions with therapeutic value has been proven successful (McManus et al., 2009). Furthermore, the identification of key pathways that resulting

synthetic lethality in simple model organism can be extrapolated to higher organism even if the specific genes are not well conserved between two species (Mclellan et al., 2012).

Essential genes which are indispensible in standard laboratory condition are highly conserved within the *E. coli* species (Takeuchi et al., unpublished data), supporting the irreplaceable roles of essential genes in cell activities. Hence, genetic interaction analysis of essential genes is vital for the complete understanding of the system level organization of living cell. The genetic interaction of essential genes in *E. coli* has not been broadly examined, owing, at least in part, to the difficulty of the essential gene perturbation. The reported genetic interaction involving essential genes in *E. coli* utilized SPA-tagged mutants, in which the nature of the nature of the hypomorphic effect is not really known, rendering the difficult inference of the observed genetic interactions than in null mutations (Butland et al., 2008).

Here I present the establishment of a method for the systematic construction of essential gene mutants by a 2 in 1 complementing plasmid. These essential gene mutants can either be used to study gene functionality for reductionist biologist; they also can be used for systematic analysis including genetic interaction and chemical genetic screens for small molecules suppressors. The establishment of the method is beneficial for the entire *E. coli* community as it will prompt the further characterization of essential genes. Here I also developed a novel system for the systematic genetic interaction analysis of essential genes in *E. coli* K-12. The regulatable promoter in the complementing plasmid allows controlling the expression of essential protein. This feature is particularly advantageous for the whole new genetic interaction analysis in which genetic interactions can be examined at different levels of essential proteins. The interpretation of the interactions is more straightforward as compared to studies relied on hypomorphic allele, because the mutation effect is the shortage of the query essential protein.

Despite pioneering the research, there are some hurdles and drawbacks in the system. Since there are no literature reviews on the genetic interaction of essential genes using knockdown mutants, the establishment of method was time consuming as there

were no controls that could be used to verify the method's authentication. The initial stage of the development was based on trials and errors to screen for potential interacting candidates that could be used as control. Meanwhile, there were some technical issue dealing with the newly devised image processing and quantification system and also the HDA Robot by Singer, these obstacles had made the process of the development more challenging. The leaky promoter is the major setback of the system. Especially essential proteins which are required in small amount, for example TrmD and FtsW, the knockdown mutants of these genes show poor IPTG dependency. Besides, the complementing plasmid poses some disadvantages for genetic interaction analysis. The plasmid while being maintained in the single gene deletions, Keio collection, it aggravates the fitness of some of the single gene deletions. In addition, the system requires the separate transfer of the plasmid and the essential gene deletion. Speaking of economically efficiency, the extra step to transfer the plasmid makes the routine procedure more cumbersome and costly (extra 32 hours and more reagents required). As for experimentally wise, this step increases the frequency of biological variations due to the increased times of replica pinning procedure. In order to overcome the problems, a new version of the complementing plasmid with tighter regulation is under the way; in addition, another approach of regulating the essential protein by replacing the authentic promoter with regulatable promoter is also ongoing. Comparisons will be made in the future to determine which system is the best for the purpose of genetic interaction analysis.

Technically, two methods were developed for the genetic interaction analysis, the spotting method and the stamping method. Data produced from each method has advantages and disadvantages. However, the results are not directly comparable because, the statistical analysis to identify genetic interaction was different (only IPTG dependent genes were selected in the case of the spotting method), and the phenotypic readout differed (colony area was determined in the spotting method, and saturated point of growth was determined in the stamping method).

In spite of the imperfections, by exploiting the capacity of this novel system using the stamping method, both predicted and new functional relationships were successfully detected. The first characteristic of these interactions is the distinction from the protein-protein interactions. Non-overlapping of genetic interaction and protein-protein interaction is often observed in the previous genetic interaction assays in yeast (Tong et al., 2004; Davierwala et al., 2005). Protein interactions dictate the architecture of the cell in terms of how direct associations between molecules constitute protein complexes, while genetic interactions define functional relationships through cause-and-effect relationships between genes. The interactions that occur among genes of disparate functions are not uncommon, as they were also evident in yeast (Davierwala et al., 2005). The root mechanism that underlines genetic interaction is typically thought to be functional overlap or redundancy between the involved genes for negative interaction and members of protein complexes or downstream pathway and suppressors for positive interaction. In fact, genetic interaction can be interpreted beyond that. Indirect mechanistic links can occur because the deletion of a gene does not only represent the absence of a particular gene, but also response of the cell to the absence, involving diverse pathways from feedback regulation to repair. So, unbiased genome wide genetic interaction study in yeast and *E. coli* revealed that genetic interactions often span seemingly disparate cellular functions and may give new insights into the global functioning of a given gene and the contexts in which the gene is utilized.

In fact, essential genes in both yeast and *E. coli* tend to have more functional links than non-essential genes (Davierwala et al., 2005; Butland et al., 2008), and the tendency of having proximate functional relationship between interacting genes is not as strong as it is for non-essential genes (Davierwala et al., 2005). Essential genes include the core functions upon which basic cellular processes are based and play an essential role in an organism's functional repertoire. The indispensability of essential genes suggests that essential genes tend to have more functional links than non-essential genes in order to buffer the system level organization of living cells. This makes sense for the observation of our study that essential genes in *E. coli* interact with genes of functionally indirectly

linked is reasonable as it was also indicated in yeast. In yeast, it was five times the frequency obtained in a previously described network (Davierwala et al., 2005). As in *E. coli* it is around 4 times the frequency comparing with the nonessential gene network. However, our five screens may not be representative of *E. coli* genes in general, with results likely underestimating for two reasons: first, we used a stringent cutoff to minimize false positives, which may have eliminated genuine interactions; second, even though functionally related genes tend to cluster in the bacterial chromosome, we removed genetically linked genes (within 35 kbp) from consideration to eliminate the linkage effect as described in the materials and methods.

Thus, our results indicate that essential genes are highly significant on the genetic interaction network, and essential pathways. Finally, this system can potentially be expanded to examine genetic interaction of the combinations of 303 essential genes and 4400 non-essential genes, contributing to the significant understanding of *E. coli* as a whole system.

Acknowledgement

ive thanks in all circumstances, for this is God's will for you in Christ Jesus.

(1Thessalonians 5:18)

First of all, I am indebted to many people for making my journey of pursuing the PhD an unforgettable one. I am deeply grateful to my supervisor, Mori sensei who has added considerably to my graduate experience. I would like to thank him for teaching me to think systematically, which has remarkedly changed my perspective on research and life

Thank to Nakayashiki sensei for managing the lab and maintaining the lab in such good condition. I am thankful for his technical advice, scientific discussion and his countless meal treats to the students. My thanks extend to all my lab mates, for their generosity in lending a helping hand and friendship, especially Rikiya for his help in statistical analysis.You guys have made my stay very comfortable and care-free. Special thanks to Dr Wanner for the opportunity to conduct experiment in Purdue University. It was indeed an eye opening experience for me.

This dissertation would not have been possible without my committee members; I would like to acknowledge Maki sensei, and Takagi sensei, for their constructive suggestions and encouragement throughout my study. My sincere thanks to 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.

To all my friends, is my pleasure to know you all, I will remember the laughter and great time we have had together. Last but not least, I would like to express my heartfelt gratitude to my family members for their unconditional love and support, and I have to say I am so blessed to have my husband, Tom for being such a wonderful partner and soul mate.

References

Aitchison, J., & Schwikowski, B. (2003). Systems biology and elephants, (April), 98103.

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., et al. (2006). Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular systems biology*, *2*, 2006.0008.

Babu, M., Díaz-Mejía, J. J., Vlasblom, J., Gagarinova, A., Phanse, S., Graham, C., Yousif, F., et al. (2011). Genetic interaction maps in Escherichia coli reveal functional crosstalk among cell envelope biogenesis pathways. (W. F. Burkholder, Ed.)*PLoS genetics*, *7*(11), e1002377.

Bakal, C., Linding, R., Llense, F., Heffern, E., Martin-Blanco, E., Pawson, T., & Perrimon, N. (2008). Phosphorylation networks regulating JNK activity in diverse genetic backgrounds. *Science (New York, N.Y.)*, *322*(5900), 453–6.

Bergmiller, T., Ackermann, M., & Silander, O. K. (2012). Patterns of evolutionary conservation of essential genes correlate with their compensability. (I. Matic, Ed.)*PLoS genetics*, *8*(6), e1002803.

Bernhardt, T. G., & De Boer, P. A. J. (2004). Screening for synthetic lethal mutants in Escherichia coli and identification of EnvC (YibP) as a periplasmic septal ring factor with murein hydrolase activity. *Molecular microbiology*, *52*(5), 1255–69.

Breslow, D. K., Cameron, D. M., Collins, S. R., Schuldiner, M., Stewart-Ornstein, J., Newman, H. W., Braun, S., et al. (2008). A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nature methods*, *5*(8), 711–8.

Butland, G., Babu, M., Díaz-Mejía, J. J., Bohdana, F., Phanse, S., Gold, B., Yang, W., et al. (2008). eSGA: E. coli synthetic genetic array analysis. *Nature methods*, *5*(9), 789–95.

Byrne, A. B., Weirauch, M. T., Wong, V., Koeva, M., Dixon, S. J., Stuart, J. M., & Roy, P. J. (2007). A global analysis of genetic interactions in Caenorhabditis elegans. *Journal of biology*, *6*(3), 8.

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(12), 6640–5.

Davierwala, A. P., Haynes, J., Li, Z., Brost, R. L., Robinson, M. D., Yu, L., Mnaimneh, S., et al. (2005). The synthetic genetic interaction spectrum of essential genes. *Nature genetics*, *37*(10), 1147–52.

De Berardinis, V., Vallenet, D., Castelli, V., Besnard, M., Pinet, A., Cruaud, C., Samair, S., et al. (2008). A complete collection of single-gene deletion mutants of Acinetobacter baylyi ADP1. *Molecular systems biology*, *4*, 174.

Dennis, G., Sherman, B., Hosack, D., Yang, J., Gao, W., Lane, H. C., & Lempicki, R. (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, *4*(5), P3.

Dietzl, G., Chen, D., Schnorrer, F., Su, K.-C., Barinova, Y., Fellner, M., Gasser, B., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in Drosophila. *Nature*, *448*(7150), 151–6.

Dixon, S. J., Costanzo, M., Baryshnikova, A., Andrews, B., & Boone, C. (2009). Systematic mapping of genetic interaction networks. *Annual review of genetics*, *43*, 601–25.

Edwin, G. (n.d.). The Science Creative Quarterly » SYSTEMS BIOLOGY: A SYSTEMS APPROACH TO UNDERSTANDING THE COMPLEXITY OF BIOLOGY. Retrieved December 11, 2012, from http://www.scq.ubc.ca/systems-biology-a-systems-approach-to-understanding-the-complexity-of-biology/

Elena, S. F., & Lenski, R. E. (1997). Test of synergistic interactions among deleterious mutations in bacteria. *Nature*, *390*(6658), 395–8.

Felsenstein, J. (1965). The effect of linkage on directional selection. *Genetics*, *52*(2), 349–63. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1210855&tool=pmcentrez&rendertype=abstract

Fraser, A. G., Kamath, R. S., Zipperlen, P., Martinez-Campos, M., Sohrmann, M., & Ahringer, J. (2000). Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. *Nature*, *408*(6810), 325–30.

Friedmann, H. C. (2004). From "butyribacterium" to "E. coli": an essay on unity in biochemistry. *Perspectives in biology and medicine*, *47*(1), 47–66. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15061168

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Véronneau, S., Dow, S., et al. (2002). Functional profiling of the Saccharomyces cerevisiae genome. *Nature*, *418*(6896), 387–91.

Griffiths, A. F. J., Miller, J. H., Suzuki, D. T. (2000). An Introduction to Genetic Analysis. 7th Edition. New York: W.H. Freeman.

Hunter, P. (2008). The paradox of model organisms. The use of model organisms in research will continue despite their shortcomings. *EMBO reports*, *9*(8), 717–20. doi:10.1038/embor.2008.142

Jasin, M., & Schimmel, P. (1984). Deletion of an essential gene in Escherichia coli by site-specific recombination with linear DNA fragments. *Journal of bacteriology*, *159*(2), 783–6. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=215717&tool=pmcentrez&rendertype=abstract

Jonikas, M. C., Collins, S. R., Denic, V., Oh, E., Quan, E. M., Schmid, V., Weibezahn, J., et al. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science (New York, N.Y.)*, *323*(5922), 1693–7. doi:10.1126/science.1167983

Joyce, A. R., Reed, J. L., White, A., Edwards, R., Osterman, A., Baba, T., Mori, H., et al. (2006). Experimental and computational assessment of conditionally essential genes in Escherichia coli. *Journal of bacteriology*, *188*(23), 8259–71.

Katagiri, F. (2003). Perspectives on Systems Biology Attacking Complex Problems with the Power of, *132*(June), 417–419.

Kelley, R., & Ideker, T. (2005). Systematic interpretation of genetic interactions using protein networks. *Nature biotechnology*, *23*(5), 561–6.

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., & Mori, H. (2005). Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research. *DNA research : an international journal for rapid publication of reports on genes and genomes*, *12*(5), 291–9.

Kitano, H. (2002). Systems biology: a brief overview. *Science (New York, N.Y.)*, *295*(5560), 1662–4.

Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., et al. (2003). Essential Bacillus subtilis genes. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(8), 4678–83.

LaDuca, R. J., Crute, J. J., McHenry, C. S., & Bambara, R. A. (1986). The beta subunit of the Escherichia coli DNA polymerase III holoenzyme interacts functionally with the

catalytic core in the absence of other subunits. *The Journal of biological chemistry*, *261*(16), 7550–7. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3519609

Lloubès, R., Cascales, E., Walburger, A., Bouveret, E., Lazdunski, C., Bernadac, A., & Journet, L. (n.d.). The Tol-Pal proteins of the Escherichia coli cell envelope: an energized system required for outer membrane integrity? *Research in microbiology*, *152*(6), 523–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11501670

Mani, R., St Onge, R. P., Hartman, J. L., Giaever, G., & Roth, F. P. (2008). Defining genetic interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(9), 3461–6.

McLellan, J. L., O'Neil, N. J., Barrett, I., Ferree, E., Van Pel, D. M., Ushey, K., Sipahimalani, P., et al. (2012). Synthetic lethality of cohesins with PARPs and replication fork mediators. (D. E. Koshland, Ed.)*PLoS genetics*, *8*(3), e1002574.

McManus, K. J., Barrett, I. J., Nouhi, Y., & Hieter, P. (2009). Specific synthetic lethal killing of RAD54B-deficient human colorectal cancer cells by FEN1 silencing. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(9), 3276–81.

Michaut, M., & Bader, G. D. (2012). Multiple genetic interaction experiments provide complementary information useful for gene function prediction. *PLoS computational biology*, *8*(6), e1002559.

Miller, J. H., (1992). A Short Course in Bacterial Genetics – A Laboratory Manual and Handbook for *Escherichia coli* and Related Bacteria. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Mnaimneh, S., Davierwala, A. P., Haynes, J., Moffat, J., Peng, W.-T., Zhang, W., Yang, X., et al. (2004). Exploration of essential gene functions via titratable promoter alleles. *Cell*, *118*(1), 31–44.

Mohammadi, T., Van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., et al. (2011). Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *The EMBO journal*, *30*(8), 1425–32.

Mori, H., Isono, K., Horiuchi, T., & Miki, T. (2000). Functional genomics of Escherichia coli in Japan. *Research in microbiology*, *151*(2), 121–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10865957

Müller-Hill, B., Crapo, L., & Gilbert, W. (1968). Mutants that make more lac repressor. *Proceedings of the National Academy of Sciences of the United States of America*,

59(4), 1259–64. Retrieved from
http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=224860&tool=pmcentre
z&rendertype=abstract

Nakayashiki, T., & Mori, H. (2013). Genome-wide screening with hydroxyurea reveals a
link between nonessential ribosomal proteins and reactive oxygen species
production. *Journal of bacteriology*, *195*(6), 1226–35.

Pan, X., Yuan, D. S., Ooi, S.-L., Wang, X., Sookhai-Mahadeo, S., Meluh, P., & Boeke, J.
D. (2007). dSLAM analysis of genome-wide genetic interactions in Saccharomyces
cerevisiae. *Methods (San Diego, Calif.)*, *41*(2), 206–21.

Phillips, P. C. (1998). The Language of Gene Interaction. *Genetics*, *149*(3), 1167–1171.
Retrieved from http://www.genetics.org/content/149/3/1167.full

Rea, S. L., Graham, B. H., Nakamaru-Ogiso, E., Kar, A., & Falk, M. J. (2010). Bacteria,
yeast, worms, and flies: exploiting simple model organisms to investigate human
mitochondrial diseases. *Developmental disabilities research reviews*, *16*(2), 200–18.

Ruiz, N., Gronenberg, L. S., Kahne, D., & Silhavy, T. J. (2008). Identification of two
inner-membrane proteins required for the transport of lipopolysaccharide to the
outer membrane of Escherichia coli. *Proceedings of the National Academy of
Sciences of the United States of America*, *105*(14), 5537–42.

Schuldiner, M., Collins, S. R., Thompson, N. J., Denic, V., Bhamidipati, A., Punna, T.,
Ihmels, J., et al. (2005). Exploration of the function and organization of the yeast
early secretory pathway through an epistatic miniarray profile. *Cell*, *123*(3), 507–19.

Segrè, D., Deluna, A., Church, G. M., & Kishony, R. (2005). Modular epistasis in yeast
metabolism. *Nature genetics*, *37*(1), 77–83.

St Onge, R. P., Mani, R., Oh, J., Proctor, M., Fung, E., Davis, R. W., Nislow, C., et al.
(2007). Systematic pathway analysis using high-resolution fitness profiling of
combinatorial gene deletions. *Nature genetics*, *39*(2), 199–206.

Templin, M. F., Ursinus, A., & Höltje, J. V. (1999). A defect in cell wall recycling
triggers autolysis during the stationary growth phase of Escherichia coli. *The EMBO
journal*, *18*(15), 4108–17. doi:10.1093/emboj/18.15.4108

Tong, A. H., Evangelista, M., Parsons, A. B., Xu, H., Bader, G. D., Pagé, N., Robinson,
M., et al. (2001). Systematic genetic analysis with ordered arrays of yeast deletion
mutants. *Science (New York, N.Y.)*, *294*(5550), 2364–8.

Tong, A. H. Y., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., Young, J., et al. (2004). Global mapping of the yeast genetic interaction network. *Science (New York, N.Y.)*, *303*(5659), 808–13.

Typas, A., Nichols, R. J., Siegele, D. A., Shales, M., Collins, S. R., Lim, B., Braberg, H., et al. (2008). High-throughput, quantitative analyses of genetic interactions in E. coli. *Nature methods*, *5*(9), 781–7. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2700713&tool=pmcentr ez&rendertype=abstract

Williams, N. (1997). PHILOSOPHY OF SCIENCE: Biologists Cut Reductionist Approach Down to Size. *Science*, *277*(5325), 476–477.

Winzeler, E. A. (1999). Functional Characterization of the S. cerevisiae Genome by Gene Deletion and Parallel Analysis. *Science*, *285*(5429), 901–906.

Xu, P., Ge, X., Chen, L., Wang, X., Dou, Y., Xu, J. Z., Patel, J. R., et al. (2011). Genome-wide essential gene identification in Streptococcus sanguinis. *Scientific reports*, *1*, 125.

Supplementary Table 1. List of common interacting genes across 5 essential query genes.
Yellow box indicates negative GI score and blue box indicates positive GI score.

Array genes	Query genes and GI score[a]				
	dnaN	*ftsW*	*trmD*	*yjgP*	*yrfF*
cyaA	0.547156	0.413158	1.538139	0.497372	0.504344
ybgT	0.623392	0.539995	0.341628	0.487762	0.612032
recA	0.192606	0.403211	0.520303	0.303338	0.387962
recB	0.425735	0.530546	1.327501	0.665911	1.027006
rfaH	0.518262	0.584189	0.651204	0.737041	0.425311
rplA	0.725374	0.480046	0.594118	0.614985	0.421183
acrB	0.69281	0.420781	0.657236	0.625674	0.608246
dnaJ	0.201871	0.692733	0.699684	0.367285	0.739755
atpF	0.504571	0.782754	0.787409	0.678032	0.594932
lipA	0.40844	0.417669	1.128206	0.570562	0.534327
aceE	0.541481	0.202512	0.477756	0.617911	0.549536
ubiX	0.24167	0.463589	0.753179	0.406729	0.270172
hns	0.672961	0.591938	0.878916	0.664637	0.734875
yhcB	0.250243	1.065534	0.47081	0.5226	0.258745
rfaC	0.355004	0.391658	0.575235	0.676594	0.33175
ECS078	0.617132	0.490072	0.686871	0.468241	0.366757
rfaG	0.402572	0.457095	0.859989	0.751709	0.487413
cydB	0.51806	0.445555	0.30137	0.337878	0.546338
gmhB	0.320765	0.350485	0.50868	0.681415	0.373434
asmA	0.615709	0.596528	0.770869	0.604216	0.662706
tolC	0.377377	0.237298	0.632857	0.365351	0.405117
rfaD	0.612676	0.45384	0.772468	0.747618	0.333244
folB	2.832676	0.966715	6.649504	0.835835	1.617577
oxyR	3.867612	1.689056	6.207046	0.872143	2.161481
atpG	3.66952	1.001973	2.106018	0.985994	1.574775
ycdB	0.967402	1.311941	1.393971	1.100668	1.295818
cysB	1.544483	1.485409	6.914718	1.855535	1.958739
cysE	2.006963	1.038577	10.08754	2.471522	1.487754
C0293	1.521173	1.338289	0.841737	1.462995	1.071003
prc	1.549158	1.480237	1.176596	1.559581	1.376014
dapF	0.923862	2.941394	5.737403	3.59853	0.87135
dnaK	1.350437	1.196223	1.493034	1.251705	1.402103
lpd	1.408295	3.397967	6.478767	4.121719	1.873293
sufA	1.32576	1.472137	1.522977	1.482749	1.284299
rimM	1.392781	1.760482	1.810096	1.517464	1.414651
tpiA	1.445196	1.154762	1.435464	1.314052	1.430128
tolA	0.898649	1.346931	1.436898	1.349229	1.059403
ubiH	0.818678	0.568286	1.924339	1.411921	1.282191
guaA	0.995337	1.537375	1.479568	1.373984	1.204983

[a]GI socre was calcuated as in Materials and Methods

Supplementary Table 2. List of genetic interactions filtered after counting for linkage effect (35kbp-window) and common interactions. Yellow box indicates negative GI score and blue box indicates positive GI score.

Array genes	Querygenes	GI score	Type of interaction
holC	dnaN	0.301195277	negative
holD	dnaN	0.357744611	negative
pdxH	dnaN	0.360588521	negative
nudH	dnaN	0.440547819	negative
fabH	dnaN	0.464775545	negative
dnaT	dnaN	0.485387263	negative
cmk	dnaN	0.49142478	negative
recC	dnaN	0.524046628	negative
iscS	dnaN	0.583227365	negative
ptsI	dnaN	0.593509709	negative
smpB	dnaN	0.613229252	negative
mrcB	dnaN	2.691997408	positive
ycfM	dnaN	2.274544003	positive
glyA	dnaN	2.217022871	positive
potG	dnaN	2.105126849	positive
rpsF	dnaN	2.053572637	positive
ygjH	dnaN	1.700041527	positive
ycjN	dnaN	1.669914951	positive
hscB	dnaN	1.565755329	positive
lpp	dnaN	1.537048799	positive
talB	dnaN	1.511990556	positive
trpB	dnaN	1.500437321	positive
ydiI	dnaN	1.473469338	positive
yddW	dnaN	1.456183547	positive
yegR	dnaN	1.451388316	positive
araG	dnaN	1.448764832	positive
aroG	dnaN	1.447922356	positive
ychN	dnaN	1.442413498	positive
rfaE	dnaN	1.442091408	positive
purF	dnaN	1.438031865	positive
yegE	dnaN	1.428126458	positive
endA	dnaN	1.42635983	positive
yebW	dnaN	1.422721392	positive
ycaM	dnaN	1.413242076	positive
ydhA	dnaN	1.407360459	positive
ybeY	dnaN	1.390127678	positive
yadN	dnaN	1.389591456	positive
deoR	dnaN	1.385106057	positive
yfcM	dnaN	1.376484463	positive
rbfA	dnaN	1.37175529	positive
gutM	dnaN	1.371381605	positive
yedA	dnaN	1.37125883	positive
yliK	dnaN	1.369135007	positive
yfdH	dnaN	1.367719057	positive
ychM	dnaN	1.364964089	positive

rzpD	dnaN	1.3625052	positive
ykfF	dnaN	1.361824116	positive
yafM	dnaN	1.361266268	positive
hsdM	dnaN	1.360735067	positive
rsxA	dnaN	1.358574574	positive
fliR	dnaN	1.355050516	positive
sdhD	dnaN	1.354622486	positive
glxR	dnaN	1.354334644	positive
btuC	dnaN	1.34939033	positive
yeeY	dnaN	1.346721763	positive
ycfJ	dnaN	1.3461287	positive
rmf	dnaN	1.345271248	positive
ksgA	dnaN	1.344063834	positive
cpsG	dnaN	1.341785829	positive
glcD	dnaN	1.341748462	positive
nupG	dnaN	1.340042365	positive
yfgL	ftsW	0.312557917	negative
yciM	ftsW	0.358108535	negative
yheM	ftsW	0.358639642	negative
yjjY	ftsW	0.389579063	negative
aceF	ftsW	0.407329512	negative
lipB	ftsW	0.409795422	negative
hflK	ftsW	0.425521465	negative
lpcA	ftsW	0.426922827	negative
fimZ	ftsW	0.452675398	negative
yacL	ftsW	0.458309847	negative
yciS	ftsW	0.472401543	negative
hfq	ftsW	0.488476806	negative
sucB	ftsW	0.489919276	negative
hflD	ftsW	0.491059913	negative
sdhA	ftsW	0.495994056	negative
nuoM	ftsW	0.498274085	negative
hdeA	ftsW	0.520411721	negative
ddpC	ftsW	0.526582149	negative
rsgA	ftsW	0.532181822	negative
hofQ	ftsW	0.5331975	negative
yjeK	ftsW	0.544199452	negative
csgB	ftsW	0.547061454	negative
acrA	ftsW	0.554858928	negative
glyA	ftsW	0.562155567	negative
ubiH	ftsW	0.568285938	negative
yihW	ftsW	0.569451144	negative
nuoL	ftsW	0.57059058	negative
dctA	ftsW	0.573139772	negative
dinF	ftsW	0.580277485	negative
cpdB	ftsW	0.58064367	negative
tauB	ftsW	0.583412387	negative
ubiE	ftsW	0.584891314	negative
sdhC	ftsW	0.585440789	negative
iscS	ftsW	0.59384378	negative

rffH	ftsW	0.595155538	negative
pspE	ftsW	0.595840387	negative
cydD	ftsW	0.600227704	negative
atpH	ftsW	0.601017226	negative
deoA	ftsW	0.603989567	negative
yhcD	ftsW	0.605817426	negative
nuoJ	ftsW	0.606672144	negative
rfaE	ftsW	0.607519467	negative
yibA	ftsW	0.611498283	negative
ruvC	ftsW	0.615416662	negative
araH	ftsW	0.615451291	negative
hupA	ftsW	0.618565198	negative
ybeA	ftsW	0.618933505	negative
flgE	ftsW	0.619032121	negative
ytfE	ftsW	0.619227231	negative
yoaH	ftsW	0.620865083	negative
livH	ftsW	0.622882383	negative
holC	ftsW	0.623329419	negative
iscU	ftsW	0.623797122	negative
ldcA	ftsW	0.624652968	negative
yheN	ftsW	0.62571121	negative
selD	ftsW	0.627526623	negative
rzoD	ftsW	0.628695294	negative
clpP	ftsW	0.631877785	negative
ybjR	ftsW	0.63190629	negative
yhhI	ftsW	0.633630616	negative
yqcD	ftsW	0.635161423	negative
yhjV	ftsW	0.637727397	negative
trpR	ftsW	0.640767846	negative
dksA	ftsW	0.642196952	negative
ymcA	ftsW	0.642893209	negative
yidH	ftsW	0.64377681	negative
nusB	ftsW	0.645417505	negative
yjiP	ftsW	0.64568484	negative
tnaA	ftsW	0.646389919	negative
htrE	ftsW	0.648213622	negative
nuoA	ftsW	0.650174745	negative
yebG	ftsW	0.650276519	negative
priA	ftsW	3.282075782	positive
ubiE	ftsW	2.239723387	positive
ubiG	ftsW	1.689056328	positive
tolQ	ftsW	1.616076098	positive
yliF	ftsW	1.431874527	positive
yeaT	ftsW	1.427444554	positive
ycbC	ftsW	1.422129936	positive
yncJ	ftsW	1.419769212	positive
hisD	ftsW	1.417064061	positive
dedD	ftsW	1.408488204	positive
aroD	ftsW	1.399578075	positive
fabH	ftsW	1.39762403	positive

ycgZ	ftsW	1.393363791	positive
yodA	ftsW	1.39173286	positive
trmU	ftsW	1.380062868	positive
rstB	ftsW	1.373798007	positive
ycjW	ftsW	1.37219877	positive
racC	ftsW	1.347890465	positive
yeaM	ftsW	1.347173348	positive
fliP	ftsW	1.345719595	positive
ybdD	ftsW	1.343527299	positive
ynaA	ftsW	1.334446805	positive
yddE	ftsW	1.33349825	positive
yfjK	ftsW	1.332753615	positive
yeaW	ftsW	1.328959338	positive
yciF	ftsW	1.32503292	positive
fruA	ftsW	1.324735486	positive
ydiJ	ftsW	1.320054247	positive
ydhY	ftsW	1.316144796	positive
ycjT	ftsW	1.315337267	positive
acnA	ftsW	1.314985354	positive
ytfF	ftsW	1.311400124	positive
yceB	ftsW	1.307607316	positive
dnaT	ftsW	1.30760107	positive
ubiE	trmD	6.093944224	positive
ubiE	trmD	2.856045609	positive
nhaA	trmD	2.490452365	positive
atpC	trmD	1.773814919	positive
atpD	trmD	1.714994321	positive
tig	trmD	1.713040406	positive
ycgY	trmD	1.573324736	positive
yajD	trmD	1.556203503	positive
atpB	trmD	1.544807872	positive
cyaA	trmD	1.538139252	positive
hslU	trmD	1.463761764	positive
ydhS	trmD	1.458729717	positive
seqA	trmD	1.420816817	positive
atpB	trmD	1.414736673	positive
nuoF	trmD	1.399334807	positive
ychH	trmD	1.396174551	positive
ydeI	trmD	1.386579415	positive
hyaD	trmD	1.38526315	positive
ynfM	trmD	1.380764505	positive
yddJ	trmD	1.379239301	positive
atpD	trmD	1.375702733	positive
ygeH	trmD	1.369847254	positive
ydjM	trmD	1.361721022	positive
pabA	trmD	1.356771781	positive
ybdD	trmD	1.346418978	positive
yniD	trmD	1.341606151	positive
yaiY	trmD	1.340844629	positive
baeS	trmD	1.338092234	positive

paaH	trmD	1.336201731	positive
ybiA	trmD	1.335866436	positive
yccZ	trmD	1.335601548	positive
ydgH	trmD	1.333384252	positive
yhhY	trmD	1.332470661	positive
hslJ	trmD	1.331397645	positive
ybdB	trmD	1.33018481	positive
recB	trmD	1.327501327	positive
holC	yjgP	0.304271691	negative
rfaE	yjgP	0.396017402	negative
lipA	yjgP	0.538558272	negative
lipB	yjgP	0.551896726	negative
nhaA	yjgP	0.577442353	negative
tatC	yjgP	0.589498319	negative
cydD	yjgP	0.589511258	negative
yciM	yjgP	0.624859008	negative
envC	yjgP	0.647588017	negative
glyA	yjgP	0.663622338	negative
lpcA	yjgP	0.66870461	negative
rfaP	yjgP	0.681934246	negative
nudH	yjgP	0.688467332	negative
priA	yjgP	0.701658834	negative
racC	yjgP	0.705174936	negative
ruvC	yjgP	0.705763351	negative
lpxL	yjgP	0.709123754	negative
hisD	yjgP	1.627258087	positive
atpE	yjgP	1.497977939	positive
puuD	yjgP	1.453194984	positive
ynjA	yjgP	1.414286216	positive
atpB	yjgP	1.402103084	positive
hscA	yjgP	1.400290355	positive
malF	yjgP	1.386829922	positive
yraM	yjgP	1.38054226	positive
atpD	yjgP	1.375293054	positive
gst	yjgP	1.371201289	positive
ycbC	yjgP	1.364876203	positive
glnA	yjgP	1.357468118	positive
cysG	yjgP	1.352573938	positive
ybfH	yjgP	1.349912751	positive
yeaO	yjgP	1.347997267	positive
wcaJ	yjgP	1.344903297	positive
yphA	yjgP	1.343738218	positive
pspF	yjgP	1.34271203	positive
yebE	yjgP	1.342289914	positive
cysI	yjgP	1.32801013	positive
cysM	yjgP	1.3268131	positive
degP	yjgP	1.325747728	positive
ynfC	yjgP	1.324200445	positive
ubiF	yjgP	1.318556667	positive
ydbK	yjgP	1.316966955	positive

ssrA	yjgP	1.313469253	positive
hscB	yjgP	1.312582609	positive
cpsB	yjgP	1.310333497	positive
recD	yjgP	1.304766472	positive
eutP	yjgP	1.302139333	positive
pphB	yjgP	1.290841501	positive
atpB	yjgP	1.289921033	positive
hscC	yjgP	1.287405659	positive
acnA	yjgP	1.285142343	positive
yhbJ	yjgP	1.284203575	positive
ynfF	yjgP	1.283243658	positive
yncA	yjgP	1.282899188	positive
flgI	yjgP	1.28183526	positive
pqqL	yjgP	1.274961882	positive
yfbR	yjgP	1.274797165	positive
nudB	yjgP	1.263076319	positive
hofQ	yrfF	0.326692077	negative
nuoM	yrfF	0.434956789	negative
rpmJ	yrfF	0.509963048	negative
ubiF	yrfF	0.531729104	negative
envC	yrfF	0.537817528	negative
rbfA	yrfF	0.547514912	negative
hflD	yrfF	0.557002624	negative
minC	yrfF	0.558262648	negative
nuoJ	yrfF	0.568051182	negative
ubiE	yrfF	0.572929402	negative
lipA	yrfF	0.573950051	negative
hfq	yrfF	0.57765812	negative
aceF	yrfF	0.585530984	negative
flgL	yrfF	0.58588876	negative
secB	yrfF	0.590239974	negative
tatC	yrfF	0.59548425	negative
yjeF	yrfF	0.596532917	negative
nuoB	yrfF	0.599250212	negative
talA	yrfF	0.614337178	negative
glyA	yrfF	0.618231331	negative
yheM	yrfF	0.624774872	negative
ydaE	yrfF	0.63284227	negative
nhaA	yrfF	0.636771701	negative
yhaH	yrfF	0.640762677	negative
recC	yrfF	0.644833652	negative
deaD	yrfF	0.646898117	negative
ynjA	yrfF	0.653349822	negative
yajC	yrfF	0.655932567	negative
nikD	yrfF	0.656997726	negative
yheN	yrfF	0.659188425	negative
ynbC	yrfF	0.661898938	negative
nikE	yrfF	0.665741362	negative
rfaI	yrfF	0.670241317	negative
ykfJ	yrfF	0.684160288	negative

yrbC	yrfF	0.69079751	negative
mdtE	yrfF	0.691951554	negative
pnp	yrfF	0.702023857	negative
atpH	yrfF	0.702056142	negative
acrA	yrfF	0.704779662	negative
yfgL	yrfF	0.707504757	negative
ygeN	yrfF	0.707780043	negative
yhcG	yrfF	0.708229157	negative
cpxR	yrfF	0.714824302	negative
kbaZ	yrfF	0.716229747	negative
icd	yrfF	0.71861566	negative
pflB	yrfF	0.719711551	negative
rffA	yrfF	0.720013499	negative
xylF	yrfF	0.720124366	negative
pdxH	yrfF	0.720317764	negative
yhhP	yrfF	0.722651275	negative
gltA	yrfF	0.724135938	negative
yjjY	yrfF	0.73177701	negative
rnr	yrfF	0.732558441	negative
yhiL	yrfF	0.732952294	negative
asnB	yrfF	0.733182048	negative
bioF	yrfF	0.735703733	negative
chiX	yrfF	0.738156184	negative
argR	yrfF	0.738411609	negative
rfaP	yrfF	0.738452454	negative
sbcB	yrfF	0.742648022	negative
glgA	yrfF	0.74367606	negative
cdaR	yrfF	0.745965289	negative
dppD	yrfF	0.748488062	negative
yhdA	yrfF	0.749168142	negative
gspO	yrfF	0.75071331	negative
rpsF	yrfF	2.414859803	positive
mrcB	yrfF	2.183120857	positive
ubiG	yrfF	2.162280247	positive
potG	yrfF	1.922461063	positive
ycfM	yrfF	1.736337597	positive
yfjK	yrfF	1.575932614	positive
ansP	yrfF	1.371004142	positive
ybjC	yrfF	1.365092799	positive
ykiA	yrfF	1.351694443	positive
atpC	yrfF	1.350743305	positive
iscS	yrfF	1.330952393	positive
ygeH	yrfF	1.329495456	positive
ygaM	yrfF	1.326929977	positive
betT	yrfF	1.319308707	positive
ycgI	yrfF	1.31756334	positive
ycdL	yrfF	1.314542807	positive
yeiB	yrfF	1.31202554	positive
yfdX	yrfF	1.307375696	positive
yegQ	yrfF	1.298182295	positive

stfE	yrfF	1.29664903	positive
cysM	yrfF	1.291259932	positive
etp	yrfF	1.28921651	positive
lsrD	yrfF	1.288534062	positive
ygjH	yrfF	1.288010253	positive
yliE	yrfF	1.286930174	positive
ydfH	yrfF	1.28437215	positive
yfeH	yrfF	1.276326393	positive
atoS	yrfF	1.27295259	positive
yeeV	yrfF	1.269674375	positive
rseB	yrfF	1.262946518	positive
ydjM	yrfF	1.256064969	positive
ycbT	yrfF	1.256041192	positive
yohC	yrfF	1.252239811	positive
nrdE	yrfF	1.250921845	positive
ybiX	yrfF	1.250599293	positive
pykF	yrfF	1.248217809	positive
serC	yrfF	1.24763577	positive
yncJ	yrfF	1.24515586	positive
yeaE	yrfF	1.243056629	positive
yohL	yrfF	1.239477612	positive
ymjA	yrfF	1.236376556	positive
aroD	yrfF	1.233252601	positive
ypdJ	yrfF	1.229983573	positive
yejE	yrfF	1.228439366	positive
yedW	yrfF	1.226485932	positive
topB	yrfF	1.225789533	positive
pepB	yrfF	1.221581685	positive
yfcV	yrfF	1.221372236	positive
fsaA	yrfF	1.22030765	positive
chaB	yrfF	1.219008881	positive
ynjB	yrfF	1.218998095	positive
ydhW	yrfF	1.218783007	positive
hcaD	yrfF	1.216613927	positive
hyaD	yrfF	1.215797764	positive
pgpB	yrfF	1.215252709	positive
ydhL	yrfF	1.213735273	positive
yaiU	yrfF	1.213595722	positive
sfcA	yrfF	1.213569823	positive
yeaY	yrfF	1.213249398	positive
ychM	yrfF	1.213126152	positive