

**Studies on the adaptation of single-stranded  
RNA virus to novel environments using  
thermal adaption experimental evolution of  
RNA bacteriophage Q $\beta$  as a model system**

RNA バクテリオファージ Q $\beta$  の高温適応進化実験を  
モデルとして用いた一本鎖 RNA ウイルスの  
新規環境への適応に関する研究

**A Dissertation Submitted  
By  
Md. Tanvir Hossain**

**Doctor of Philosophy  
Bioresources Science**

**The United Graduate School of Agricultural Sciences  
Iwate University, Japan  
September 2020**

# Contents

<b>List of Figures</b>	<b>iii</b>
<b>List of Tables</b>	<b>iv</b>
<b>List of Abbreviations</b>	<b>v</b>
<b>1 General introduction</b>	<b>1</b>
<b>2 The single-stranded RNA bacteriophage Q<math>\beta</math> adapts rapidly to high temperatures</b>	<b>4</b>
2.1 Introduction . . . . .	4
2.2 Materials and methods . . . . .	6
2.2.1 Experimental materials . . . . .	6
2.2.2 Preparation of host strain 46L-1F' . . . . .	7
2.2.3 Determination of specific growth rate of 46L-1 and 46L-1F' . . . . .	8
2.2.4 Development of culture method of 46L-1F' . . . . .	9
2.2.5 Preparation of Q $\beta$ 18 mut from cDNA . . . . .	9
2.2.6 Determination of phage concentration . . . . .	10
2.2.7 Thermal adaptation experiment . . . . .	10
2.2.8 Fitness analysis . . . . .	13
2.2.9 Genome sequencing of Q $\beta$ populations . . . . .	13
2.2.10 Secondary structure prediction . . . . .	16
2.2.11 Statistical analysis . . . . .	16
2.3 Results . . . . .	16

---

2.3.1	Preparation of host strain . . . . .	16
2.3.2	Preparation of Q $\beta$ 18 mut from cDNA . . . . .	20
2.3.3	Experimental evolution . . . . .	21
2.3.4	Fitness of Q $\beta$ populations . . . . .	24
2.3.5	Molecular evolution of Q $\beta$ populations . . . . .	26
2.4	Discussion . . . . .	32
<b>3</b>	<b>Study on Q<math>\beta</math> mutants obtained in evolution experiment</b>	<b>41</b>
3.1	Introduction . . . . .	41
3.2	Materials and methods . . . . .	44
3.2.1	Strains and medium . . . . .	44
3.2.2	Construction of Q $\beta$ mutants from cDNA of genomes . . . . .	44
3.2.3	Fitness analysis of the Q $\beta$ mutants . . . . .	45
3.2.4	Statistical analysis . . . . .	45
3.3	Results . . . . .	46
3.4	Discussion . . . . .	47
<b>4</b>	<b>Conclusions</b>	<b>52</b>
	<b>Summary</b>	<b>54</b>
	<b>Acknowledgements</b>	<b>58</b>
	<b>Appendix A Nucleotide sequences of evolved Q<math>\beta</math> phages</b>	<b>59</b>
	<b>Bibliography</b>	<b>74</b>

# List of Figures

2.1	Schematic of the evolution experiment . . . . .	12
2.2	Temperature profile of adaptation experiment . . . . .	12
2.3	Location of PCR products within bacteriophage Q $\beta$ genome . . . . .	14
2.4	Electrophoresis results of <i>traU</i> and <i>leuB</i> genes in <i>E. coli</i> strains . . . . .	17
2.5	Growth curves of <i>E. coli</i> 46L-1 . . . . .	18
2.6	Growth curves of <i>E. coli</i> 46L-1F' . . . . .	19
2.7	Specific growth rates of 46L-1 and 46L-1F' . . . . .	20
2.8	Fitness of Q $\beta$ 18 mut . . . . .	21
2.9	Population dynamics of three independent Q $\beta$ lines in thermal adaptation . . . . .	23
2.10	Fitness of Q $\beta$ 18 mut and evolved populations of each line . . . . .	25
2.11	Mutation positions of the evolved population of each Q $\beta$ line . . . . .	30
2.12	Mutations introduced and reverted to original sequence with adaptation . . . . .	31
2.13	Time course of changes in cumulative mutations in the Q $\beta$ genome . . . . .	31
2.14	Mutational positions of amino acids mapped into the structure . . . . .	38
2.15	Secondary structure of Q $\beta$ Anc(P1), 18 mut and endpoint populations . . . . .	39
2.16	Secondary structure of starting and evolved (fixed mutations) Q $\beta$ genomes . . . . .	40
3.1	Position of Ile477 and Leu509 amino acids in Q $\beta$ replicase . . . . .	43
3.2	Fitness of Q $\beta$ mutants and populations . . . . .	50
3.3	Secondary structure of Q $\beta$ 18 mut and 3(43.7) mut . . . . .	51

# List of Tables

2.1	Templates, primers, and polymerase used for PCR product preparation	15
2.2	List of primers used for Q $\beta$ genome sequencing . . . . .	15
2.3	Nucleotide sequences of Anc(P1), 18 mut, and evolved phages . . . . .	28
2.4	Tendency of mutations in thermal adaptation within individual Q $\beta$ genes	31

# List of Abbreviations

BSA	bovine serum albumin
cDNA	complementary deoxyribonucleic acid
CFU	colony forming units
dATP	deoxyadenosine triphosphate
DNA	deoxyribonucleic acid
dsDNA	double-stranded deoxyribonucleic acid
dsRNA	double-stranded ribonucleic acid
EF-Tu	elongation factors thermally unstable
EF-Ts	elongation factors thermally stable
LB	Luria-Bertani
mM63l	modified-M63 medium with 1 mM L-leucine
mM63lm	modified-M63 medium with 1 mM L-leucine and 2.3 mM $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$
OD <sub>600</sub>	optical density at 600 nm
PCR	polymerase chain reaction
PFU	plaque forming units
RNA	ribonucleic acid
ssRNA	single-stranded ribonucleic acid
ssDNA	single-stranded deoxyribonucleic acid
UTR	untranslated region

# Chapter 1

## General introduction

Adaptation of living beings to their environments is integral to our understanding of organismal diversity and has been the worry of scientists for a long time [1]. Many evolutionary experiments pursue to discern how populations adapt to particular environmental conditions, conventionally defined in terms of particular factors, such as temperature [2], nutrition [3], competition [4], parasites [5], or saline stress [6]. Recognizing general adaptation patterns would improve our capacity to predict future adaptation occasions. For example, we would be better able to evaluate the risks of population extinctions and the risks of disease emergence in people, plants, and livestock due to global warming.

Viruses, the intracellular parasites, are packages of genes, the genetic information in the form of DNA or RNA, inside protective protein shells [7]. According to the genome structure, viruses are classified into double-stranded (ds)DNA, single-stranded (ss)DNA, dsRNA, or ssRNA (positive-sense or negative-sense). ssRNA viruses change maximum mutation rate compatible with maintaining the integrity of genetic information (i.e., the error threshold) because this would allow them to quickly find the beneficial mutations needed for adaptation [8–11]. It is an unquestionable fact that ssRNA virus populations exist as quasispecies, which is a population with a large number of variant genomes [12]. Due to comparatively high mutation rates ( $10^{-3}$  to  $10^{-6}$  per base per replication [13–15]) of ssRNA viruses, the evolution speed tends

to be high [16]. However, it is considered that ssRNA viruses can adapt readily to changes in the environment, it remains unclear how rapidly they can adapt to a novel environment and/or how many and what types of mutation are required to facilitate evolution. Adaptation to temperature is a useful model for studying adaptation to novel environments [17] because major living processes, such as energy transduction, reproduction, and growth, are affected by temperature change. To elucidate the mechanisms underlying ssRNA virus adaptation to the novel environment in an experimental setting, it is necessary to use a model system. This will allow to understand the relationship between genotypic and phenotypic changes easily.

The ssRNA bacteriophage Q $\beta$ , of the family *Leviviridae*, infects *Escherichia coli* strains expressing the F pili that acts as the virus receptor [18]. The Q $\beta$  has a genome of 4217 nucleotides in length encoding four proteins: the A2 protein for bacterial lysis and entry, the coat protein, the A1 protein, which is expressed through incorrect reading of the stop codon of the coat protein and present in low amounts in the capsid, and the  $\beta$  subunit for Q $\beta$  replicase [18]. This genome is replicated by the phage-specific enzyme Q $\beta$  replicase [19, 20], which is formed by the assembly of 4 subunits. The  $\beta$  subunit is phage-derived; the remaining subunits, derived from the host cell, are ribosomal protein S1 [21] and translational elongation factors EF-Tu and EF-Ts [22]. The Q $\beta$  is a congruent model, because it has small genome size, high mutation rate ( $10^{-3}$  to  $10^{-5}$  per base per replication [13, 23, 24]), and short generation time, thus facilitating genome-level evolution experiments.

A thermal adaptation experiment using ssRNA bacteriophage Q $\beta$  as a model of ssRNA virus was conducted in our laboratory, in which the culture temperature was increased in a stepwise manner from 37.2°C to 43.6°C in three independent lines using *E. coli* 43BF' as the host strain that could grow in temperatures up to 43.6°C. The effect of synonymous and nonsynonymous changes on the fitness and life history of Q $\beta$  were evaluated [25, 26]. Recently, Kishimoto et al. isolated a strain of thermally adapted *E. coli* that was capable of growing at temperatures up to 46°C by thermal

adaptation evolution experiment [27]. This technical advance can make it possible to continue the thermal adaptation experiments to explore the potential upper temperature limit.

In this study, I continued the previous thermal adaptation experiment at temperature up to 45.3°C using thermally adapted Q $\beta$  18 mut and *E. coli* 46L-1F' as the host. In Chapter 2, I constructed the host strain *E. coli* 46L-1F' and Q $\beta$  18 mut from cDNA of Q $\beta$  genomes with 18 mutations and conducted the thermal adaptation experiment at temperatures up to 45.3°C using aforementioned thermally adapted *E. coli* 46L-1F' as the host. After adaptation at 45.3°C, the culture temperature was decreased to 37.2°C to investigate whether the ancestral sequence became dominant in the population. I showed that Q $\beta$  can grow and replicate at the highest known growth temperature, 45.3°C within 114 days (including 52 days in this study and 62 days from 37.2°C to 43.6°C in a previous study [25]), without showing trade-off in the lower ranges, and adapted to elevated temperatures with only point mutations and these mutations account for 0.8% – 0.9% of the total RNA genome.

In Chapter 3, in order to investigate the improved growth characteristics of 43.65°C (hereafter 43.7°C)-adapted endpoint populations, four kinds of mutant (18 mut-A1781C, -U3784C, -C3879G, and -combined of three) were prepared through site-directed mutagenesis of the Q $\beta$  expression vector pACYCQ $\beta$ \_18 mut and the fitness was determined using the host strain 43BF' (43.6°C) and 46L-1F' (43.6°C and 43.7°C). I showed that the temperature 43.7°C acts as a selective pressure for Q $\beta$  18 mut, the mutations were introduced due to thermal adaptation, and the neutral and/or slightly deleterious mutations were important as well as beneficial mutations in thermal adaptation.

## Chapter 2

# The single-stranded RNA bacteriophage Q $\beta$ adapts rapidly to high temperatures

### 2.1 Introduction

Bacteriophages are viruses that infect and replicate in bacterial cells, are ubiquitous in the environment and are estimated as the most widely distributed and diverse biological entities on Earth [28]. Bacteria, the targets and hosts of bacteriophages, have important functions in modulating the nutrient and/or energy cycles in the global ecosystem [29]. In general, bacteriophages are highly specific for a host, may undergo a lytic and/or lysogenic cycle. The lytic bacteriophages kill the host cell in order to release their progeny [30]. Therefore, preferential replication of a specific bacteriophage may alter the composition and biochemical function of a specific bacterial community and thereby have a profound impact on the global ecosystem.

Bacteriophages are classified into dsDNA, ssDNA, dsRNA, or ssRNA according to their genome structure. Due to comparatively high mutation rates and the existence of quasispecies, the evolution speed of ssRNA bacteriophage tends to be high [12, 16]. Although it is considered that ssRNA bacteriophages can adapt readily to changes in

the environment, it is not clear how rapidly they can adapt to a novel environment and/or how many and what types of mutations are required to adapt. Every environment plays host to microorganisms that have become adapted for optimal growth. Hence, in general, when bacteriophages adapt to novel environments, this implies that they have not only adapted to the novel environment but also to adapt to bacterial hosts living there. In order to elucidate the mechanisms underlying bacteriophage adaptations to the environment in an experimental setting, it is necessary to use a single host strain that can withstand these changes with negligible interruption. This will allow the assessment of the effect of a single environmental factor on the closed bacteriophage-host system.

Temperature is an important environmental factor of great biological significance. The majority of organisms are ectothermic, and establishes organismal temperature directly from environmental temperature and thereby exerts control on biological rate processes such as energy transduction, reproduction, and growth. Therefore, it is expected that organisms acclimate phenotypically to temperature change, and populations adapt genetically to changing thermal conditions. Many laboratory experiments have been carried out using ssDNA and ssRNA bacteriophages to study the detailed relationship between fitness and the underlying genetic changes to stress caused by elevated environmental temperatures [31–36]. These early studies have focused on thermal adaptation up to 42°C or 43.5°C, the impact of short-term pulsed exposure of bacteriophages to high temperatures (above 50°C), as well as heat-pulse followed by culture at high temperature (43°C) in nutrient-rich medium, and reported the specific adaptive mutations. In these early studies, they used the ordinary laboratory bacterial strains as a host, which could not grow at temperatures above ~44°C. Therefore, it is impossible to conduct the thermal adaptation experiment higher than this temperature using ordinary laboratory bacterial strains. However, the true thermal adaptation of bacteriophages indicates that bacteriophages should be capable of growth and replication at higher temperatures, a point that has not yet been fully explored. To investigate the thermal adaptation of bacteriophages, it is necessary to

use a host strain that is capable of growth at higher temperatures as a host. A thermal adaptation experiment was conducted in our laboratory using the F pili specific ssRNA bacteriophage Q $\beta$  as a model of ssRNA virus in three independent lines using *E. coli* 43BF' as the host strain that could grow in temperatures up to 43.6°C. The effect of synonymous and nonsynonymous changes on the fitness and life history of Q $\beta$  were evaluated [25, 26]. Recently, Kishimoto et al. isolated a thermally adapted *E. coli* strain that was capable of growing at temperatures up to 46°C by thermal adaptation evolution experiment [27]. This thermally adapted strain can make it possible to continue the thermal adaptation experiments to explore the potential upper temperature limit for adaptive responses of coliphages.

In this Chapter, I conducted a thermal adaptation experiment with Q $\beta$  18 mut, which was obtained in previous evolution experiment and had adapted to 43.6°C, at temperatures up to 45.3°C using the thermally adapted *E. coli* 46L-1F' as the host strain. I showed that experimental evolution of ssRNA phage as a model system is very useful to elucidate its ability to quickly adapt to a novel environment. The collection of this fundamental knowledge is important for predicting the ssRNA virus evolution.

## 2.2 Materials and methods

### 2.2.1 Experimental materials

A thermally adapted derivative of the *E. coli* DH1 strain, 46L-1 (generated from line 1 via adaptation up to 46°C for 8,829 generations [27]), which was kindly provided by Professor Toshihiko Kishimoto, Toho University, Chiba, Japan, and *E. coli* HB2151 (K12 (lac-pro), ara, nalr, thi/F'[proAB, lacIq, lacZ M15]) strain [37] were used for preparation of host strain, 46L-1F' by conjugation. *E. coli* A/ $\lambda$  strain [38] was used as the host strain for the titer assay. The *E. coli* F<sup>-</sup> strain, DH5 $\alpha$ /pACYCQ $\beta$ -18 mut containing 18 mutations [26], was used for construction of Q $\beta$  18 mut. DH5 $\alpha$ /pACYCQ $\beta$ -18 mut was cultured in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) containing 50  $\mu$ g/mL of kanamycin to maintain their plasmids. Modified-M63 medium

(62 mM K<sub>2</sub>HPO<sub>4</sub>, 39 mM KH<sub>2</sub>PO<sub>4</sub>, 15 mM ammonium sulfate, 1.8  $\mu$ M FeSO<sub>4</sub> · 7 H<sub>2</sub>O, 15  $\mu$ M thiamine hydrochloride, 0.2 mM MgSO<sub>4</sub> · 7 H<sub>2</sub>O, and 22 mM glucose) [39] with 1 mM L-leucine (mM63l) [40] was used as minimal medium. Mg<sup>2+</sup> was used as the divalent cation for minimal medium rather than Ca<sup>2+</sup> to prevent calcium phosphate precipitation. The mM63l with 2.3 mM MgSO<sub>4</sub> · 7 H<sub>2</sub>O (total 2.5 mM MgSO<sub>4</sub> · 7 H<sub>2</sub>O) liquid medium (mM63lm) and mM63l with 1.5% agar (mM63l-agar) medium were used for culturing HB2151, 46L-1, and 46L-1F' and with 50  $\mu$ g/mL of kanamycin were used for culturing 46L-1 (during preparation of host strain via conjugation).

For static culture, a thermostatic incubator (Cool Incubator CN-25C, Mitsubishi Electric Engineering Co. Ltd., Tokyo, Japan) was used. Shaking culture of *E. coli* at 37.2°C was carried out using a culture apparatus equipped with a shaker (Eyela Multi Shaker MMS, Tokyo Rikakikai Co., Ltd., Tokyo, Japan) inside a thermostat (Eyela FMC-1000, Tokyo Rikakikai Co., Ltd.). For shaking culture of *E. coli* at 43.7°C, 44.1°C, 44.8°C, 45.3°C, and 45.9°C, a shaker (Taitec NR-1, Taitec Corporation, Tokyo, Japan) inside a thermostatic incubator (Panasonic MIR-154-PJ, Panasonic Corp., Osaka, Japan) was used. A micro-cooling centrifuge (Kubota 3740S, Kubota Corporation, Tokyo, Japan) was used for the recovery of the cell, and a thermal cycler (Astec PC 320, Astec Corporation, Fukuoka, Japan) was used for the polymerase chain reaction (PCR). Ultraviolet-Visible spectrophotometer (BioSpec-mini, Shimadzu Corporation, Tokyo, Japan) was used for determining the optical density at 600 nm (OD<sub>600</sub>) of the culture solution.

### **2.2.2 Preparation of host strain 46L-1F'**

*E. coli* 46L-1F' was constructed by conjugation with a thermally adapted derivative of the DH1 strain, 46L-1 [27], and HB2151 strain [37]. The strain 46L-1 was cultured in mM63lm medium containing 50  $\mu$ g/mL of kanamycin and HB2151 was cultured in mM63lm medium. The strain suspensions were collected from the log-phase and diluted to OD<sub>600</sub> = 0.1 in 1 mL. These bacterial solutions were centrifuged at 13,400 $\times$  *g* at room temperature for 2 min, the strains were collected and resus-

pended in 1 mL of mM63lm medium. The resuspended strains were mixed (500  $\mu$ L : 500  $\mu$ L), incubated at 37°C for 5 h and 20 min and screened for kanamycin-resistant clones on mM63l-agar medium with 50  $\mu$ g/mL of kanamycin (incubated at 37°C for about 46 h). Whether the strain, 46L-1 was converted to F<sup>+</sup> strain or not was determined by PCR using the following two kinds of primers set for 13 clones randomly selected from agar medium. The F plasmid was determined by F plasmid primers TraU<sub>f</sub> (5'-ATGAAGCGAAGGCTGTGGCT-3') and TraU<sub>r</sub> (5'-GCAGCTTGAACGCCATGCGT-3') for the *traU* gene encoded. In addition, HB2151 strains, lack the *leuB* gene and it is possible to distinguish from 46L-1 strain by examining the presence or absence of this gene. For the deletion of the *leuB* gene, the primers LeuB<sub>r</sub> (5'-TTCGGCTTCGTACACAACGTGAG-3') and PtetA (5'-CCTAATTTTTGTTGACACTCTATCATTG-3') were used. PCR was carried out using GoTaq Green Master Mix (Promega, Madison, WI, USA) with bacterial suspension suspended in autoclave water as a template. PCR cycling parameters for 751 bp (*traU*) and ~1000 bp (*leuB*) targets was as: 95°C for 2 min; 30 cycles at 95°C for 30 s, 55°C for 30 s, 72°C for 1 min; 72°C for 5 min; 15°C for  $\infty$ . The PCR products were analyzed by 1% agarose gel electrophoresis using  $\lambda$ -*Sty* I marker. Among the *E. coli* obtained, three strains (clones 1-3 (Fig. 3.2)) were selected and prepared -80°C frozen stock. The strain from clone 2 (one out of three) was selected and used as 46L-1F'.

### **2.2.3 Determination of specific growth rate of 46L-1 and 46L-1F'**

The *E. coli* 46L-1 and 46L-1F', which were stocked at -80°C, were cultured on mM63l-agar medium at 37°C for about 24 h. The strains, in agar medium, were inoculated into 3 mL mM63lm and cultured at 37.2°C, 43.7°C, 44.8°C, 45.3°C (only 46L-1F') and 45.9°C with 160 $\pm$ 1 rpm shaking for overnight (about 15 h). The overnight-cultured cells (30  $\mu$ L, for initial OD<sub>600</sub>=0.014 to 0.026) were transferred into new 3 mL mM63lm and cultured at same conditions for next 26 h. The growth curves were obtained from semi-logarithmic plots of OD<sub>600</sub>. The logarithmic growth period, the specific

growth rates and the OD<sub>600</sub> value in the stationary phase were determined from the growth curves. The specific growth rates were obtained from the fitted straight lines of logarithmic growth period of semi-logarithmic plots of OD<sub>600</sub>. The specific growth rates were determined from one to two inoculums of one or two replicates and average specific growth rates were used as the specific growth rates at that temperature.

#### **2.2.4 Development of culture method of 46L-1F'**

The *E. coli* 46L-1F', which was cultured in mM63l-agar medium at 37°C for about 24 h, was inoculated in 3 mL of mM63lm medium and cultured at 37.2°C, 43.7°C, 44.1°C, 44.8°C and 45.3°C with 160±1 shaking for 6 h. The strains were diluted and 30  $\mu$ L diluted strains were added in new 3 mL mM63lm medium and cultured at the same conditions for 15 h (initial OD<sub>600</sub> 0.00019, 0.00055 and 0.00092 for 37.2°C and 43.7°C; 0.00055, 0.00092 and 0.00129 for 44.1°C and 44.8°C; and 0.00086, 0.00144 and 0.00202 for 45.3°C). The strain with OD<sub>600</sub> value of around 0.5 was selected as the culture solution to be inoculated next and diluted to the initial OD<sub>600</sub> of 0.022 (for 37.2°C) or 0.029 (for 43.7°C, 44.1°C, 44.8°C and 45.3°C) by adding  $x$   $\mu$ L strain to (3000- $x$ )  $\mu$ L mM63lm medium. The strains were cultured at same condition for next 9 h and the OD<sub>600</sub> after 4 and 9 h were measured.

#### **2.2.5 Preparation of Q $\beta$ 18 mut from cDNA**

*E. coli* F<sup>-</sup> strain, DH5 $\alpha$ /pACYCQ $\beta$ .18 mut containing 18 mutations was constructed previously [26], was cultured in 5 mL LB medium with 50  $\mu$ g/mL kanamycin at 30°C at 160±1 rpm for 8.5 h. 2 mL of pre-cultured strains were added in 500-mL culture flasks containing 200 mL of LB medium with 50  $\mu$ g/mL kanamycin at 30°C at 125±1 rpm for 24 h. After centrifugation at 5,000 $\times$   $g$  at 4°C for 20 min, the supernatants were filtered with 1.2- $\mu$ m-pore-size syringe filters (Minisart; Sartorius Stedim Biotech, Göttinge, Germany) and with 0.2- $\mu$ m-pore-size syringe filters (Steradisc 25; Kurabo Industries, Ltd., Osaka, Japan). The previous operation was performed two times for getting more amount of phage solution and mixed. The filtered supernatants were con-

centrated with Amicon Ultra-15 centrifugal filter units with 50,000-nominal-molecular-weight-limit membranes (EMD Millipore, Billerica, MA, USA). The concentrated supernatants including phage particles were dialyzed with P buffer [50 mM Tris-HCl (pH 7.6), 0.1 M NaCl, 5 mM MgCl<sub>2</sub>, and 0.1 mM EDTA·2Na] [41] using Slide-A-Lyzer gamma-irradiated dialysis cassettes with 10,000-nominal-molecular weight-limit membranes (Thermo Fisher Scientific, Waltham, MA, USA) and then concentrated with Amicon Ultra-4 centrifugal filter units with 50,000-nominal-molecular-weight-limit membranes (EMD Millipore). The concentrated phage particles were filtered using 0.2- $\mu$ m-pore-size syringe filters (Minisart RC15 filters; Sartorius) and stored in 40% glycerol at  $-20^{\circ}\text{C}$ .

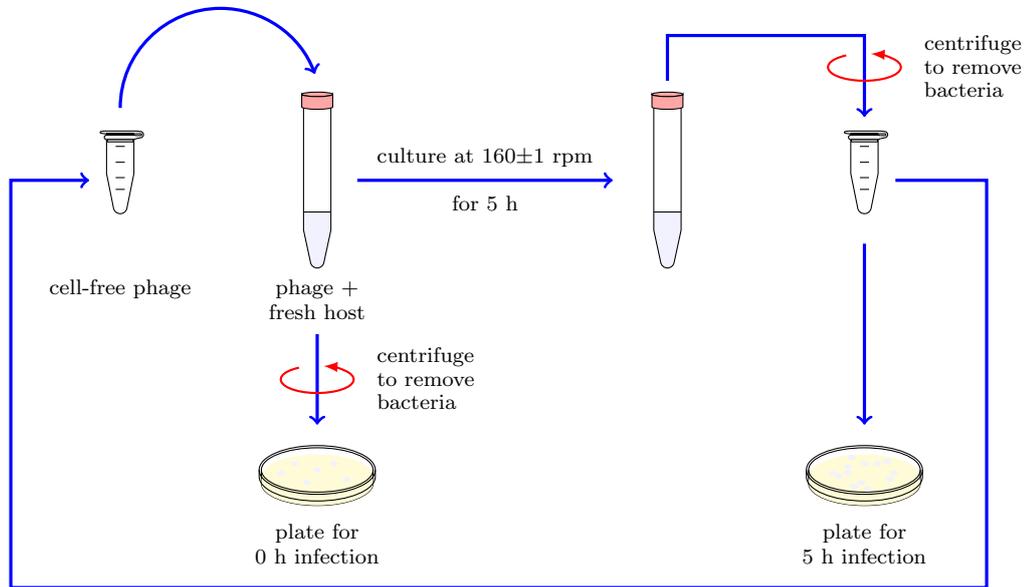
### **2.2.6 Determination of phage concentration**

The phage concentrations were determined with the plaque assay technique according to a standard method [42, 43] using LB agar and LB soft agar medium. Titer assay results were expressed in plaque forming units (PFU)/mL. Equal volume (100  $\mu$ l) of *E. coli* A/ $\lambda$  and diluted phages were added in autoclaved soft agar (0.75% agar), which was preserved on thermostated at  $50^{\circ}\text{C}$ . After gentle mixing, suspension was poured across the agar plate. Plates with solidified soft agar were incubated at  $37^{\circ}\text{C}$  for overnight. By preparing different dilutions of phage solution before being added to soft agar different number of plaques on each agar plate were formed. The plates containing countable number of plaques were used for determination of phage concentration in original solution.

### **2.2.7 Thermal adaptation experiment**

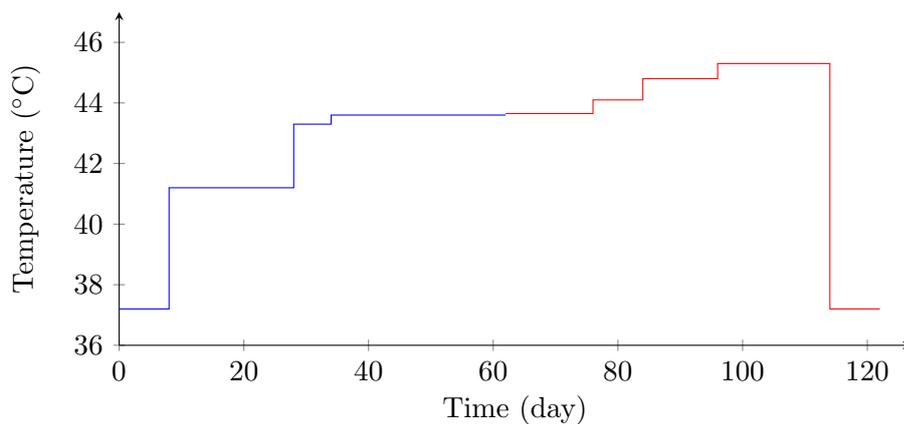
The thermal adaptation experiment was performed in three independent lines with Q $\beta$  at  $43.7^{\circ}\text{C}$ ,  $44.1^{\circ}\text{C}$ ,  $44.8^{\circ}\text{C}$ ,  $45.3^{\circ}\text{C}$  and  $37.2^{\circ}\text{C}$  (Figs. 2.1 and 2.2). Serial passages of Q $\beta$  consisted of infection of the host culture, followed by approximately 5 h of phage growth and extraction of the phage from three cultures. The culture temperature was changed when the amplification ratio at each temperature seemed to have equi-

librated. However, the passage numbers at each temperature after equilibration was reached were arbitrary. Each serial passage was performed as follows: uninfected log phase 46L-1F' cell cultures grown at each temperature at  $160 \pm 1$  rpm were transferred into fresh medium with dilution to 0.029 or 0.022 OD<sub>600</sub> units (the latter value for 37.2°C only) and cultured for 4 h at each temperature. After 4 h, the OD<sub>600</sub> was between 0.126 to 0.214 (approximately  $1 \times 10^8$  to  $2 \times 10^8$  colony forming unit (CFU)/mL calculated from  $1 \times 10^9$  CFU/mL of OD<sub>600</sub>=1), and cells were infected with phage at approximately  $9.7 \times 10^5$  to  $3.8 \times 10^6$  PFU/mL from the previous passage. An aliquot was sampled, and free bacteriophages for determination of initial titer were taken from the supernatants after centrifugation at  $13,400 \times g$  for 1 min at room temperature. The cultures were grown for approximately 5 h and divided into three portions; one was used for separation of free phages by centrifugation at  $13,400 \times g$ , the second was used for OD<sub>600</sub> analysis, and the third was used for  $-80^\circ\text{C}$  frozen (15% glycerol) stock. Free-phage was titrated and stored at  $4^\circ\text{C}$  to use for infection in the next serial passage. To prevent adsorption of Q $\beta$  to the tube wall, 0.1% bovine serum albumin (BSA)-coated 15-mL polypropylene centrifuge tubes were used [40]. The titer assay was conducted according to a standard method [42] using LB agar and LB soft agar medium described in Section 2.2.6. The replication generations ( $g$ ) of the Q $\beta$  genome was calculated as the cumulative generations of each passage,  $(N_5/N_0) = 2^g$ , where  $N_5$  and  $N_0$  represent 5 h post-infection and initial (0 h) free-phage density of each passage in PFU/mL, respectively, and  $g$  represents replication generation. The endpoint Q $\beta$  populations at  $43.7^\circ\text{C}$ ,  $44.8^\circ\text{C}$ ,  $45.3^\circ\text{C}$ , and  $37.2^\circ\text{C}$ -adapted were selected and designated as 43.7\_1, 43.7\_2, 43.7\_3, 44.8\_1, 44.8\_2, 44.8\_3, 45.3\_1, 45.3\_2, 45.3\_3, 37.2\_r1, 37.2\_r2, and 37.2\_r3 with the numbers before and after the underscore representing the passage temperature and passage line, respectively. The “r” was added to distinguish the  $37.2^\circ\text{C}$  adaptation in previous report [25].



**Figure 2.1:** Schematic of the evolution experiment.

Bacteria-free Q $\beta$  were added to culture with *E. coli* 46L-1F' and grown at  $160\pm 1$  rpm for 5 h at pre-specified temperature ( $43.7^\circ\text{C}$ ,  $44.1^\circ\text{C}$ ,  $44.8^\circ\text{C}$ ,  $45.3^\circ\text{C}$  and  $37.2^\circ\text{C}$ ). After the adaptation period, Q $\beta$  were separated from the bacteria by centrifugation at  $13,400\times g$  for 1 min and new cell-free Q $\beta$  were used for infection in next serial passage. In order to determine the amplification ratio between this 5 h adaptation, cell-free phages were plated for initial (0 h) and post-infection (5 h) concentration by mixing a dilution of the phage with A/ $\lambda$  as host in soft agar and spreading it onto LB-agar plate. After overnight incubation, plaques in the bacterial lawn were counted, each of which originated from a single viral particle. The culture temperature was elevated or decrease when the amplification ratio at each temperature seemed to have equilibrated. The experiment was continued for 60 transfers.



**Figure 2.2:** Temperature profile of adaptation experiment.

Thermal adaptation was initiated at  $37.2^\circ\text{C}$ ; culture temperatures were increased in a stepwise manner to  $45.3^\circ\text{C}$ . After adaptation to  $45.3^\circ\text{C}$ , populations were then returned to culture at  $37.2^\circ\text{C}$  to investigate whether Q $\beta$  will return to its ancestral sequence. The thermal adaptation for temperature from  $37.2^\circ\text{C}$  to  $43.6^\circ\text{C}$  (blue line) was conducted in our lab previously [25].

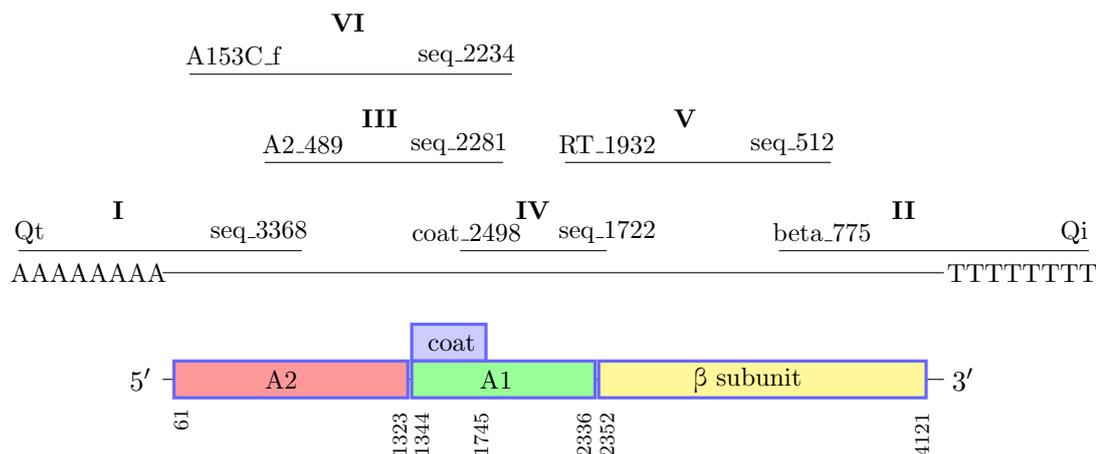
### 2.2.8 Fitness analysis

The fitness of Q $\beta$  18 mut and the aforementioned 12 endpoint populations was evaluated at 37.2°C, 43.7°C, 44.1°C (only 18 mut), 44.8°C, and 45.3°C using the host strain 46L-1F'. The strain 46L-1F' was cultured at 37.2°C, 43.7°C, 44.1°C, 44.8°C, and 45.3°C with 160 $\pm$ 1 rpm shaking according to the method described in Section 2.2.4 and 2.5 mL uninfected log phase of 46L-1F' strain (the OD<sub>600</sub> between 0.126 and 0.268 approximately 1.3 $\times$ 10<sup>8</sup> and 2.7 $\times$ 10<sup>8</sup> CFU/mL) was infected with approximately 4.1 $\times$ 10<sup>4</sup> to 2.8 $\times$ 10<sup>6</sup> PFU/mL of Q $\beta$ . Immediately after the phage infection and approximately after 5 h infection, the cultures were sampled and centrifuged at 13,400 $\times$  *g* for 1 min at room temperature to remove the bacterial cells and obtained a supernatant containing free phages. The plaque forming ability of the phage in the obtained supernatants was measured and the fitness was calculated as  $x = \log_{10}(N_5/N_0)$ , where  $N_5$  and  $N_0$  represent 5 h post-infection and initial free phage concentrations in PFU/mL, respectively. For measurement of fitness of Q $\beta$  populations, 18 mut was used as a control for each assay temperature and each line.

### 2.2.9 Genome sequencing of Q $\beta$ populations

The RNA genomes of the Q $\beta$  populations derived from approximately 1.1 $\times$ 10<sup>9</sup> to 1.1 $\times$ 10<sup>10</sup> PFU of the variants of aforementioned 12 endpoint populations were extracted using a QIAamp Viral RNA minikit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. The samples were prepared using a previously described method [44] for analyzing the full-length RNA genome sequence. Briefly, a poly(A) sequence was added at the 3' end of each RNA genome using poly(A) polymerase (Thermo Fisher Scientific). cDNA was synthesized using the primer Qt (5'-CCAGTGAGCAGAGTGACGAGGACTCGAGCTCAAGC TTTTTTTTTTTTTTTT TT-3') with SuperScript<sup>TM</sup> III Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA), and then RNA was degraded with RNaseH. The first-strand cDNA was purified and poly(A) was added at the 3' termini of the cDNA with terminal deoxynucleotidyl transferase (Promega) and dATP. The cDNA with poly(T) at the 5' terminus and

poly(A) at the 3' terminus was purified. Second-strand DNA was prepared using Qt primer for obtaining the 5' end of the genome sequence. PCR was performed separately using Phusion High-Fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA) or Pfu Ultra II Fusion HS DNA polymerase (Agilent Technologies, Santa Clara, CA, USA); the Q $\beta$  genome was divided into six regions (Fig. 2.3) using the primers shown in Table 2.1. PCR products were purified by Diffinity RapidTip<sup>®</sup> 2 D2947L-8RXN (Sigma-Aldrich, St. Louis, MO, USA) and/or 1% agarose gel electrophoresis using the Wizard<sup>®</sup> SV Gel and PCR clean-up system (Promega) according to the protocol. Sequencing was performed by the dideoxynucleotide chain termination method [45] on both strands using the primers listed in Table 2.2.



**Figure 2.3:** Location of PCR products within bacteriophage Q $\beta$  genome. The forward and reverse primers for the PCR products were presented above the horizontal lines of the PCR products location. The four genes of Q $\beta$  are represented in the lower part with numbers representing the start and end position.

**Table 2.1:** Templates, primers, and polymerase used for PCR product preparation.

PCR products	Template	DNA polymerase	Primer name	Primer sequence (5' → 3')
Qt, Qi & seq_3368 product (I)	2 <sup>nd</sup> strand cDNA	Pfu Ultra II Fusion HS DNA polymerase (Agilent Technologies)	Qt	CCAGTGAGCAGAGTGACGAG GACTCGAGCTCAAGCTTTTT TTTTTTTTTTTTT
			Qi seq_3368	ACGAGGACTCGAGCTCAAGC TATCATGACGGTTCTGCCAG
Qi & beta 775 product (II)	1 <sup>st</sup> strand cDNA with poly(T) and poly(A)	Pfu Ultra II Fusion HS DNA polymerase	Qi	ACGAGGACTCGAGCTCAAGC
			beta_775	CGGTTACACATTTCGAGCTCG
A2_489 & seq_2281 product (III)	1 <sup>st</sup> strand cDNA with poly(T) and poly(A)	Phusion <sup>TM</sup> High-Fidelity DNA polymerase (New England Biolabs)	A2_489	TTTACTGCGCAGACTGCGTG
			seq_2281	ATACTTACCTGTCCCTGGCG
coat_2498 & seq_1722 product (IV)	1 <sup>st</sup> strand cDNA with poly(T) and poly(A)	Phusion <sup>TM</sup> High-Fidelity DNA polymerase	coat_2498	CAGGCATATGCTGACGTGAC
			seq_1722	GCAATTGTGCGCTGAGAGAG
seq_512 & RT_1932 product (V)	1 <sup>st</sup> strand cDNA with poly(T) and poly(A)	Phusion <sup>TM</sup> High-Fidelity DNA polymerase	seq_512	GGACTCACTATACGGTGACG
			RT_1932	TGCTTATTGCTCTCTTAGCG
A153C_f & seq_2234 product (VI)	1 <sup>st</sup> strand TC cDNA with poly(T) and poly(A)	Phusion <sup>TM</sup> High-Fidelity DNA polymerase	A153C_f	TTTATCGACTCTTCCGACACGCA
			seq_2234	GGAGGCTCGTAAACCTCCTC

**Table 2.2:** List of primers used for Q $\beta$  genome sequencing.

PCR products	Primer name	Primer sequence (5' → 3')
Qt, Qi & seq_3368 product (I)	A2_1	ATGCCCTAAATTACCGCGTGGTCTGC
	A2_489	TTTACTGCGCAGACTGCGTG
	seq_3368	TATCATGACGGTTCTGCCAG
	seq_3752	AATCCCAGTCAACACGAGCG
	A153C_f	TTTATCGACTCTTCCGACACGCATC
Qi & beta 775 product (II)	beta_end	TTACGCCTCGTGTAGAGACGCAACC
	seq_512	GGACTCACTATACGGTGACG
	beta_775	CGGTTACACATTTCGAGCTCG
	beta_273	GTGATTACGGACCATAACAAG
	beta_561	CGGCAAGCACTACTATTCTG
A2_489 & seq_2281 product (III)	seq_2281	ATACTTACCTGTCCCTGGCG
	seq_2870	AGATCTAGTGTAGCGAACGG
	seq_3368	TATCATGACGGTTCTGCCAG
	A2_489	TTTACTGCGCAGACTGCGTG
	A2_2997	ACAGACGTGACATCCGGCTC
coat_2498 & seq_1722 product (IV)	coat_2498	CAGGCATATGCTGACGTGAC
	seq_1722	GCAATTGTGCGCTGAGAGAG
	seq_2281	ATACTTACCTGTCCCTGGCG
	seq_1722	CAGGCATATGCTGACGTGAC
	RT_1932	TGCTTATTGCTCTCTTAGCG
seq_512 & RT_1932 product (V)	seq_512	GGACTCACTATACGGTGACG
	seq_1112	TCTGTCTTACTGTTCTTAGG
	seq_1722	GCAATTGTGCGCTGAGAGAG
	RT_1932	TGCTTATTGCTCTCTTAGCG
	beta_815	CGGTAGTGTGTTACCTACGAG
A153C_f & seq_2234 product (VI)	RT_plus_f1	AAGCTGTTGCCCTGGGAGAAG
	seq_2234	GGAGGCTCGTAAACCTCCTC
	seq_2870	AGATCTAGTGTAGCGAACGG
	seq_3368	TATCATGACGGTTCTGCCAG
	A153C_f	TTTATCGACTCTTCCGACACGCATC
	A2_2997	ACAGACGTGACATCCGGCTC
	coat_2498	CAGGCATATGCTGACGTGAC
A2_489	TTTACTGCGCAGACTGCGTG	
A2_1158	ACGAGCAATCTTCCGTTTCGC	

### 2.2.10 Secondary structure prediction

The minimum free energy secondary structures of Q $\beta$  Anc(P1), 18 mut, and 12 end-point populations were predicted using RNAfold [46] after introducing mutations observed in each populations, including those that were polymorphic and at over 90% frequency in each populations.

### 2.2.11 Statistical analysis

The specific growth rates of *E. coli* were compared by two-tailed *t* test and two-way analysis of variance (ANOVA). The fitness of Q $\beta$  were compared by two-tailed *t* test and one-way ANOVA with a Tukey-HSD test using RStudio [47]. In all analysis, a *P* value of <0.01 was taken to indicate statistical significance.

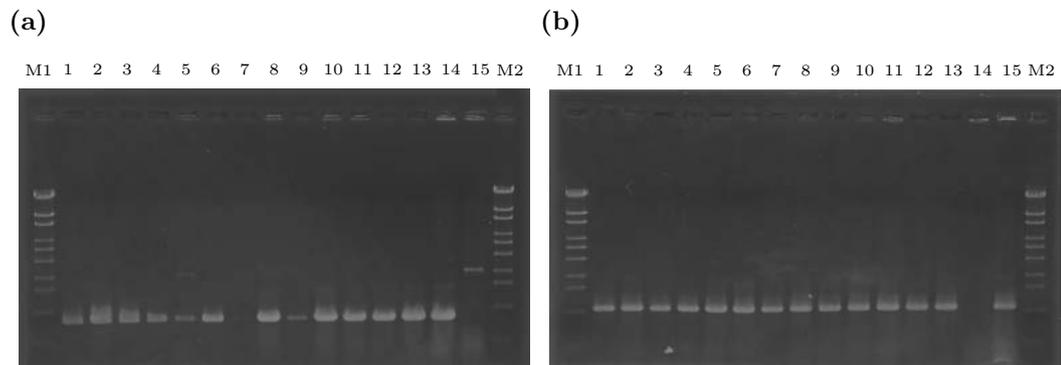
## 2.3 Results

### 2.3.1 Preparation of host strain

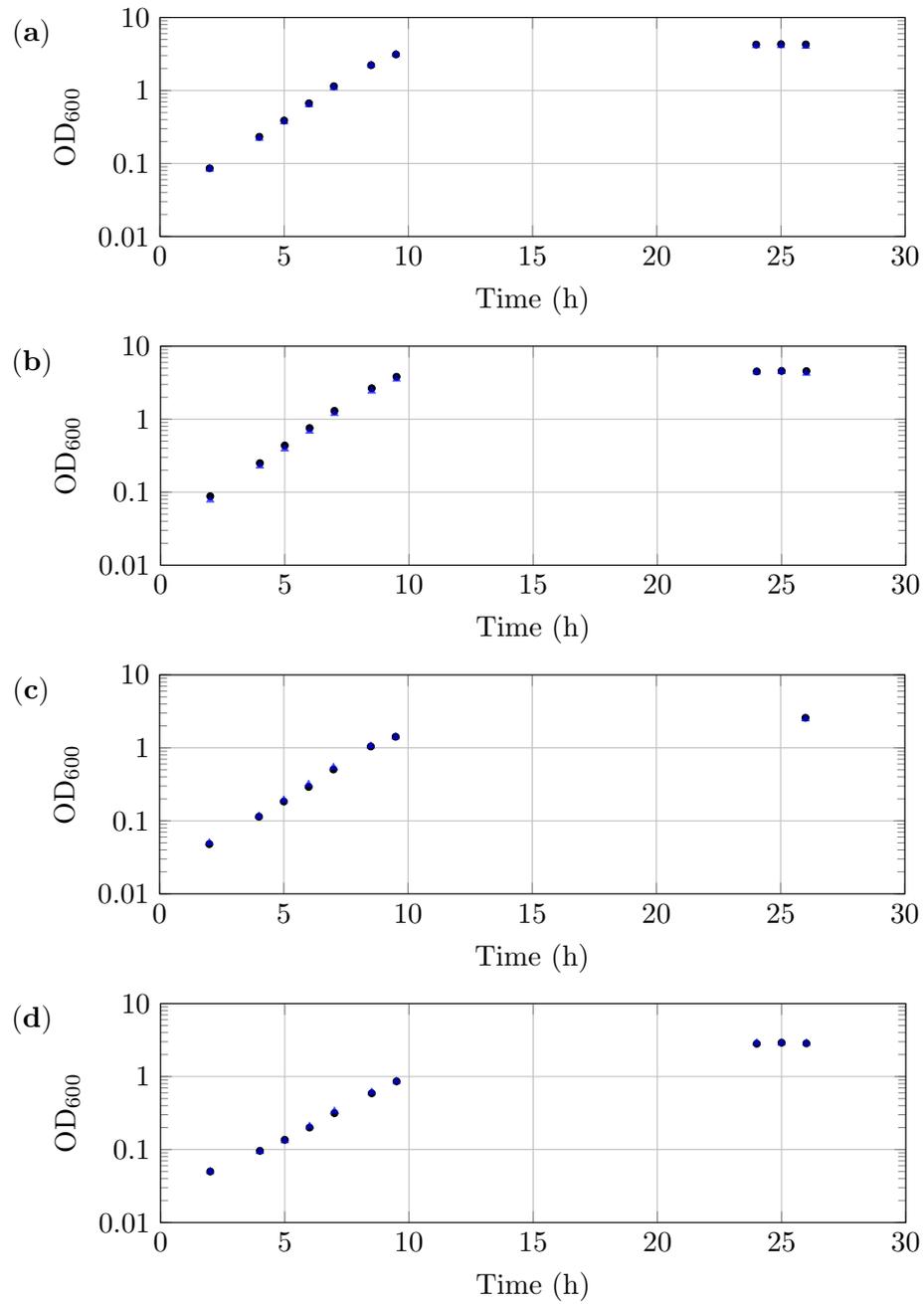
The host strain *E. coli* 46L-1F' was constructed by conjugation with strain 46L-1, which was obtained previously by thermal adaptation experimental evolution [27], and the HB2151 strain [37]. The conjugative products were checked by agarose gel electrophoresis after preparation of PCR product and clone 2 was selected as 46L-1F' (Fig. 3.2). After preparation of the strain, the growth curves of 46L-1 and 46L-1F' were prepared at increasing temperatures (Figs. 2.5 and 2.6) and obtained the specific growth rates from the slope of the fitted straight lines in semi-logarithmic plots of OD<sub>600</sub>. The specific growth rates of *E. coli* strain 46L-1 at 37.2°C, 43.7°C, 44.8°C and 45.9°C were 0.51±0.001, 0.53±0.002, 0.48±0.0002, and 0.41±0.004 h<sup>-1</sup>, respectively and the specific growth rates of 46L-1F' at 37.2°C, 43.7°C, 44.8°C, 45.3°C, 45.9°C were 0.42±0.007, 0.42±0.008, 0.42±0.015, 0.39±0.014, and 0.31±0.011 h<sup>-1</sup>, respectively (Fig. 2.7). A two-way ANOVA was conducted that examined the effect of temperatures and strains on specific growth rates. Statistically significant difference in mean specific growth rates was observed between 46L-1 and 46L-1F' strain (two-way

ANOVA,  $F_{(1,16)}=244.4779$ ,  $P < 0.001$ ) and between temperatures (two-way ANOVA,  $F_{(4,16)}=56.749$ ,  $P < 0.001$ ) (Fig. 2.7), which indicates that the F plasmid transferred during conjugation have lowering tendency of the growth of the strain. The specific growth rate of *E. coli* 46L-1F' was almost identical between 37.2°C and 44.8°C but decreased by 7% and 26% at 45.3°C and 45.9°C, respectively. Nabergoj et. al. showed that, bacteriophage population growth rate depends on bacterial growth rate [48]. Therefore, thermal adaptation experiments were considered at temperatures up to 45.3°C.

The logarithmic growth phase of *E. coli* strain 46L-1F' was considered between 2 and 9.5 hours using initial OD<sub>600</sub> of 0.014 – 0.026 (Fig. 2.6). The culture methods for evolution experiment as well as fitness assay were prepared based on the specific growth rate and logarithmic growth phase of *E. coli* strain 46L-1F'. By using the initial OD<sub>600</sub> of 0.022 at 37.2°C and 0.029 at 43.7°C, 44.1°C, 44.8°C, and 45.3°C for culture methods (9 h), the growth properties of 46L-1F' showed almost same (data not shown). Therefore, the culture methods by using initial OD<sub>600</sub> of 0.029 and 0.022 (the latter value for 37.2°C only) were used in this study.

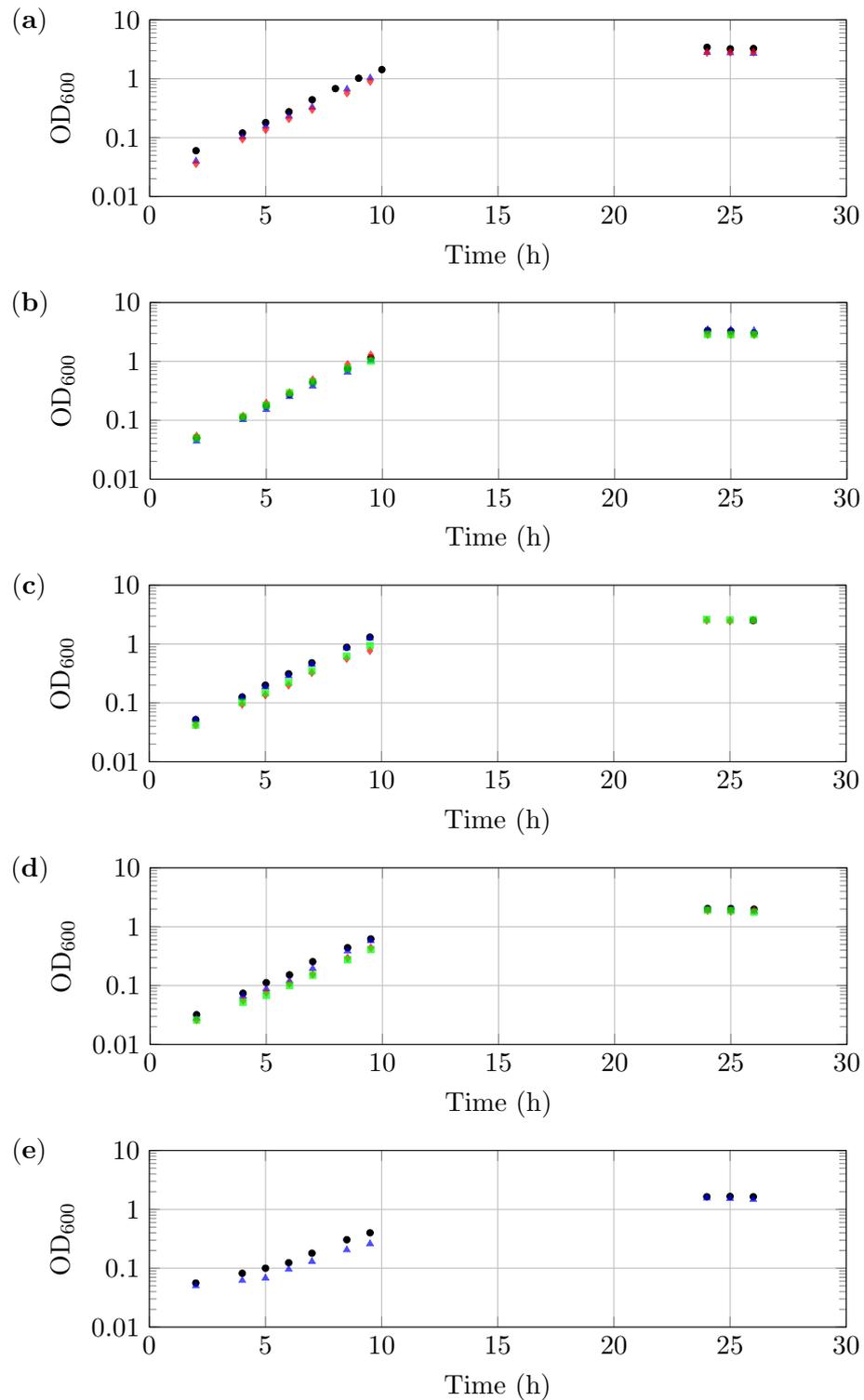


**Figure 2.4:** Electrophoresis results of *traU* and *leuB* genes in *E. coli* strains. (a) Detection of *traU* gene; (b) Detection of *leuB* gene. M1 & M2:  $\lambda$ -*Sty* I marker; 1-13: Clones 1-13; 14: HB2151; 15: 45L-1.



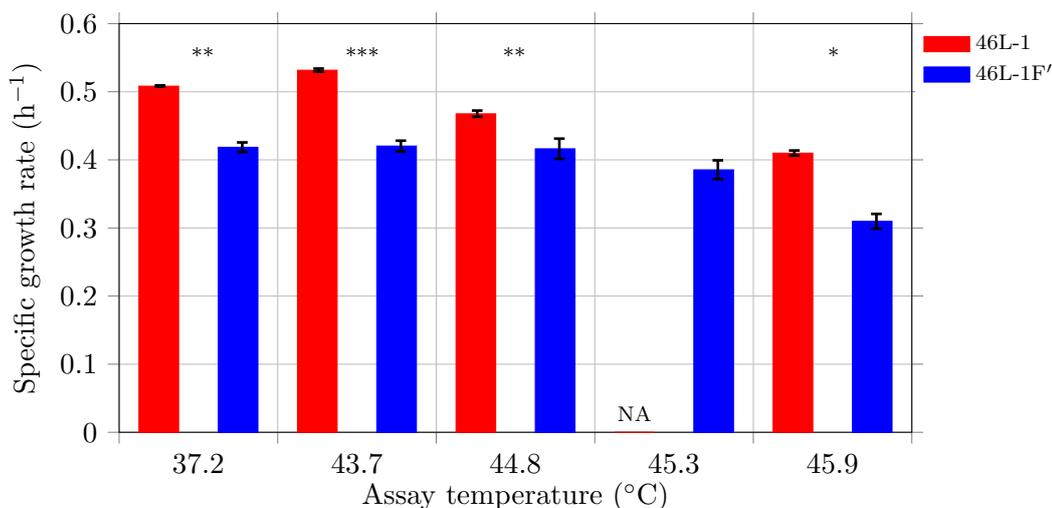
**Figure 2.5:** Growth curves of *E. coli* 46L-1.

The growth curves of *E. coli* 46L-1 were plotted based on OD<sub>600</sub>. Each experiment was conducted in duplicate. The specific growth rates were calculated from the slope of the fitting curve of logarithmic growth phase of the semi-logarithmic plots based on OD<sub>600</sub> of *E. coli* 46L-1. The growth curves a–d were prepared for 37.2°C, 43.7°C, 44.8°C, and 45.9°C, respectively.



**Figure 2.6:** Growth curves of *E. coli* 46L-1F'.

The growth curves of *E. coli* 46L-1F' were plotted based on OD<sub>600</sub>. Each experiment was conducted in one or two replicates from one or two different inoculums. The specific growth rates were calculated from the slope of the fitting curve of logarithmic growth phase of the semi-logarithmic plots based on OD<sub>600</sub> of *E. coli* 46L-1F'. The growth curves a-e were prepared for 37.2°C, 43.7°C, 44.8°C, 45.3°C, and 45.9°C, respectively.

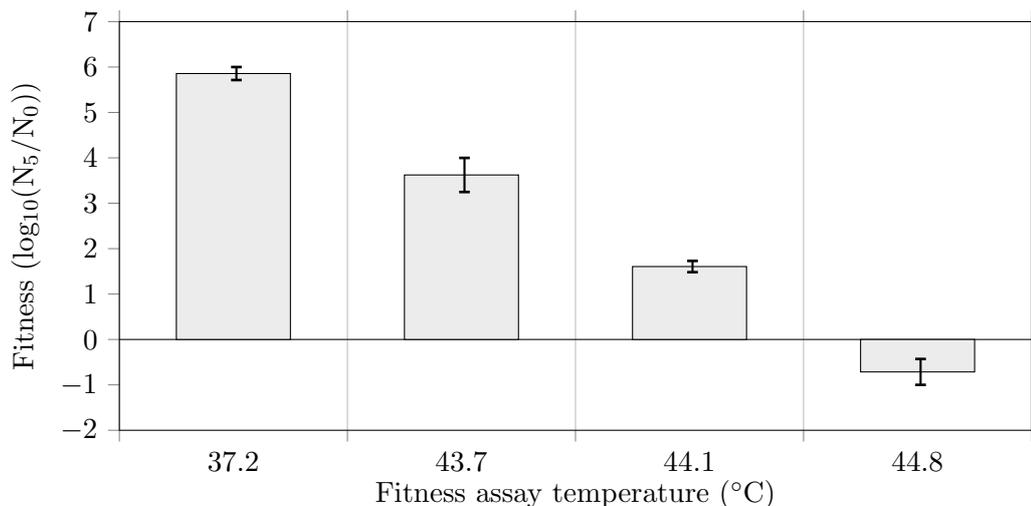


**Figure 2.7:** Specific growth rates of 46L-1 and 46L-1F'.

The means and standard deviations were obtained by repeated cultures (n=2-4). The specific growth rates were calculated from the slope of the fitting curve of logarithmic growth phase of the semi-logarithmic plots based on OD<sub>600</sub> of *E. coli*. NA represent no data for 46L-1 at 45.3°C. 2-sample *t*-test, \*\*\*  $P < 0.001$ , \*\*  $P < 0.01$ , \*  $P < 0.05$  for pairs.

### 2.3.2 Preparation of Q $\beta$ 18 mut from cDNA

As previously reported, 13, 17, and 18 mutations were observed in the final populations of the three replicates adapted at 43.6°C. The detected mutations were introduced into the cDNA and prepared mutant phages from those cDNA and designated them as 13 mut, 17 mut, and 18 mut [25, 26]. Q $\beta$  18 mut was prepared from *E. coli* DH5 $\alpha$ /pACYCQ $\beta$ .18 mut by culturing in LB medium containing 50  $\mu$ g/mL of kanamycine at 30°C to use as starting material of the present thermal adaptation experiment. The supernatants of the culture suspensions were concentrated, dialyzed against P buffer, concentrated again and stored in 40% glycerol at -20°C ( $4.90 \times 10^9$  PFU/mL was obtained). The fitness analysis at 37.2°C, 43.7°C, 44.1°C, and 44.8°C using Q $\beta$  18 mut was performed (Fig. 2.8). The fitness was defined as logarithm of the amplification ratio 5 h after the infection (Section 2.2.8). The phage titer of the culture in the 5 h infection of 18 mut decreased at 44.8°C and increased at 44.1°C. The fitness at 44.1°C is 3.6 fold lower than 37.2°C and 2.2 fold lower than 43.7°C (one-way ANOVA  $F_{3,13} = 155.5$ ,  $P < 0.001$ ; Tukey-HSD test  $P < 0.001$ ). Therefore, 44.1°C was selected as the next adaptation temperature.



**Figure 2.8:** Fitness of Q $\beta$  18 mut.

The fitness was measured at 37.2°C, 43.7°C, 44.1°C, and 44.8°C. The fitness was calculated as described in Section 2.2.8. Data are expressed as means  $\pm$  standard deviations ( $n=2-9$ ).

### 2.3.3 Experimental evolution

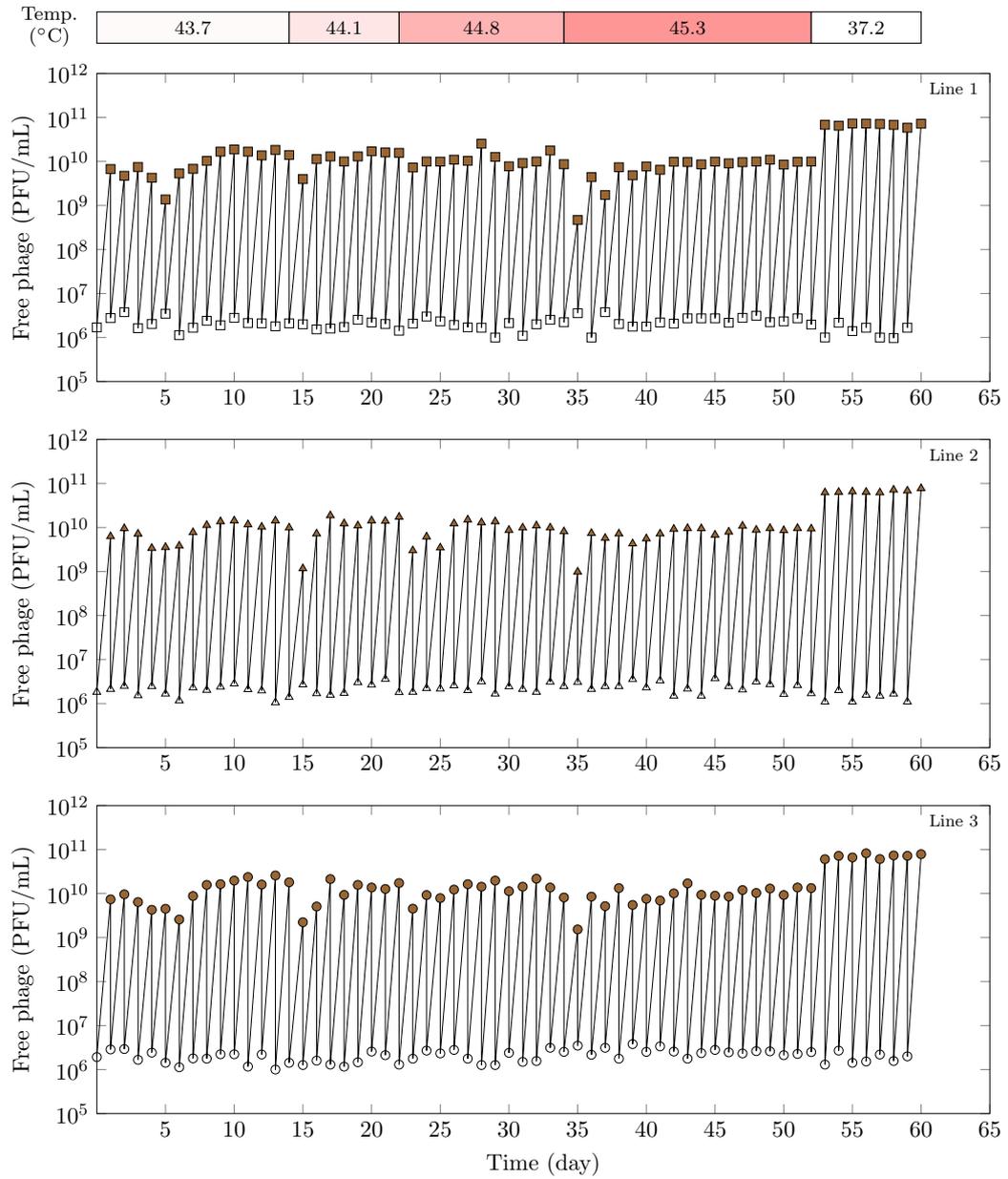
Thermal adaptation experiments were carried out from 43.7°C to 45.3°C with three independent Q $\beta$  lines using Q $\beta$  18 mut as starting materials and thermally adapted *E. coli* 46L-1F' as the host strain (Fig. 2.2). The Q $\beta$  18 mut was obtained from previous thermal adaptation experiment (from 37.2°C to 43.6°C) using *E. coli* 43BF' as the host strain, and the genome of this variant had 17 point mutations and 1 insertion when compared with the ancestral Q $\beta$  sequence [25, 26]. The additional thermal adaptation evolution started from an initial temperature of 43.7°C and was raised to 44.1°C, 44.8°C, and 45.3°C in a stepwise manner.

The Q $\beta$  18 mut was divided into three separate lines and grown in the *E. coli* 46L-1F' strain at 43.7°C for 14 days (i.e., equivalent to 167, 169, and 175 replication generations for lines 1, 2, and 3, respectively). Free Q $\beta$  was separated from the previous day's culture, diluted, and used to infect fresh *E. coli* strain 46L-1F' at logarithmic growth phase. After 14 days, thermal adaptation was carried out as follows (replication generations are given in respective order for lines 1, 2, and 3): for 8 days at 44.1°C, equivalent to 100, 97, and 101 replication generations; for 12 days at 44.8°C, equivalent

to 151, 143, and 151 replication generations; and for 18 days at 45.3°C, equivalent to 206, 207, and 211 replication generations (Fig. 2.9). Population dynamics revealed that at every temperature shift, there was a rapid recovery in the amplification ratio in the primary 1–5 days and equilibrated after an initial decrease in yield. Q $\beta$  was capable of replication at 45.3°C within 114 days (including 52 days in this study and 62 days from 37.2°C to 43.6°C in a previous study [25]) (Fig. 2.2). Diluted Q $\beta$  was added to an initial Q $\beta$  density of approximately  $1 \times 10^6$  to  $3.8 \times 10^6$  PFU/mL, which was equivalent to an effective population size ( $N_e$ ) of approximately  $2.5 \times 10^6$  to  $9.5 \times 10^6$  PFU, determined by multiplying the free-phage density immediately after infection by the culture volume of 2.5 mL. The replication generation number of the phage population ( $g$ ) was calculated as described in materials and methods section (Section 2.2.7). Q $\beta$  evolved gradually to adapt to the increasing temperature (Fig. 2.9). Significant increases in the amplification ratio were observed in the 43.7°C adaptation (approximately 3-fold), the 44.1°C adaptation (approximately 2-fold) and the 45.3°C adaptation (approximately 10-fold)(Fig. 2.9).

Finally, the culture temperature was changed from 45.3°C to 37.2°C, the temperature that permits optimum growth of the original Q $\beta$  Anc(P1), to investigate whether the ancestral sequence (Anc(P1)) became dominant in the population. The 45.3°C-adapted populations were passaged at 37.2°C for 8 days at 37.2°C, equivalent to 124, 124 and 122 replication generations for line 1, 2, and 3, respectively.

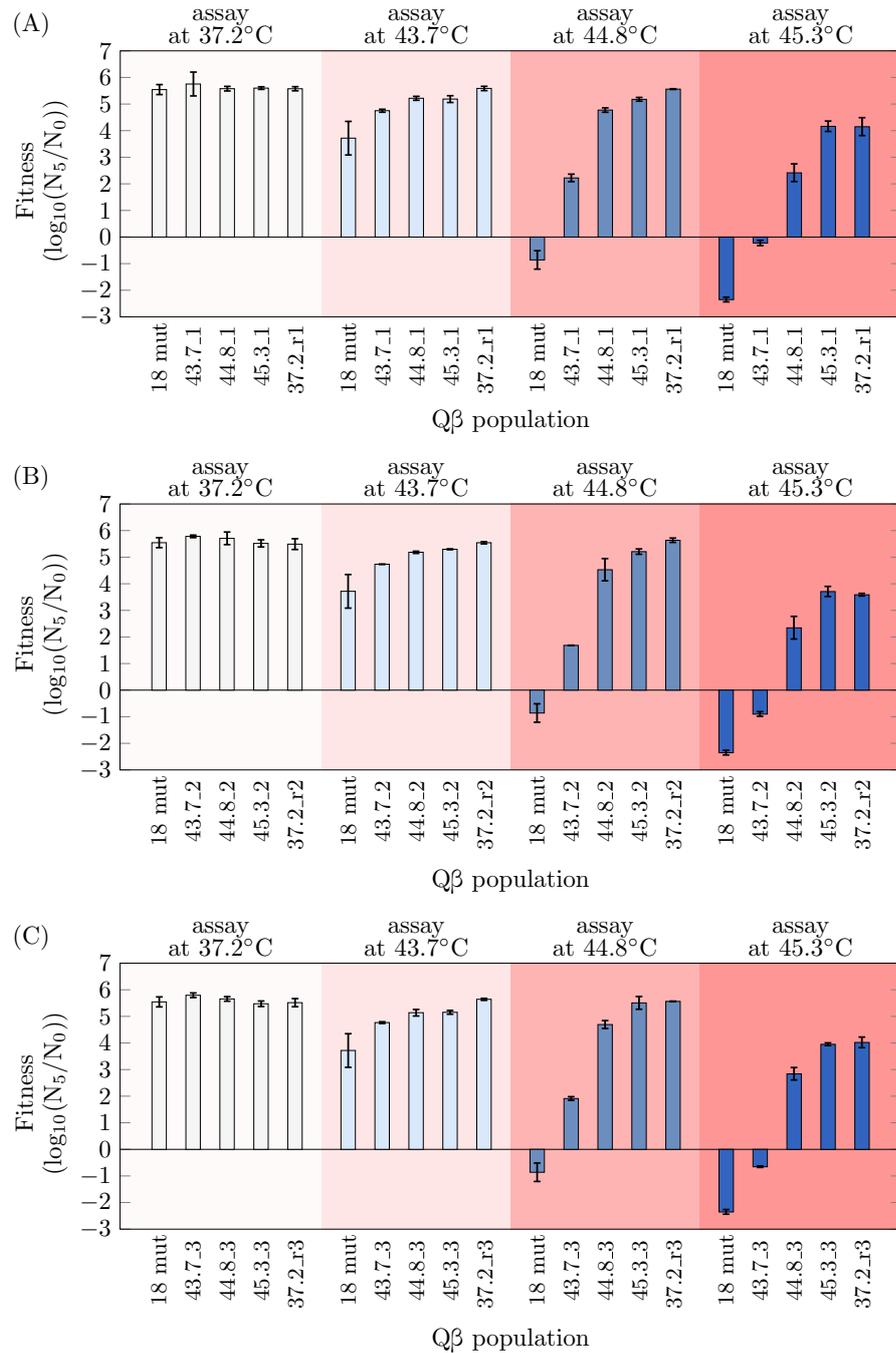
The endpoint Q $\beta$  populations at 43.7°C, 44.8°C, 45.3°C, and 37.2°C were used for further genotypic and phenotypic analysis.



**Figure 2.9:** Population dynamics of three independent Q $\beta$  lines in thermal adaptation experiments. Population dynamics of free Q $\beta$  density (PFU/mL) of line 1, line 2, and line 3 are shown. The free Q $\beta$  densities (PFU/mL) immediately after 0 h (open symbols) and after 5 h (filled symbols) infection were measured every day.

### 2.3.4 Fitness of Q $\beta$ populations

Fitness analysis showed that thermally adapted Q $\beta$  populations gained the ability to grow at high temperatures at which the 18 mut strain could not grow (Fig. 2.10). The fitness of Q $\beta$  was performed at 37.2°C, 43.7°C, 44.8°C, and 45.3°C by measuring the amplification ratio 5 h after infection to estimate the growth ability under the same conditions as used in the evolution experiment. The fitness was calculated as described in the materials and methods section (Section 2.2.8). In the 37.2°C assay, the fitness of 18 mut and populations adapted to each temperature were almost the same (one-way ANOVA, line 1,  $F_{4,16} = 0.41, P = 0.80$ ; line 2,  $F_{4,16} = 1.27, P = 0.32$ ; line 3,  $F_{4,16} = 1.81, P = 0.18$ ). In the 43.7°C, 44.8°C, and 45.3°C assays, the fitness of thermally adapted populations increased with increasing adaptation temperature (Fig. 2.10). These observations indicated that an increasing growth ability at higher temperature was not necessarily associated with decreased fitness at lower temperatures. In the 45.3°C assay, 18 mut and 43.7°C-adapted populations showed decreased fitness value, but the reduction was recovered with increasing adaptation temperature. The amplification ratios of 45.3°C-adapted populations were approximately ( $1.6 \times 10^6$ ) to ( $4 \times 10^6$ )-fold compared with 18 mut, and the fitness values of 45.3.1, 45.3.2, 45.3.3 and 37.2\_r1, 37.2\_r2, 37.2\_r3 were significantly larger than those of the other populations in each line (line 1, one-way ANOVA,  $F_{4,9} = 339.1, P < 0.001$ ; Tukey-HSD test,  $P < 0.001$ ; line 2, one-way ANOVA,  $F_{4,9} = 314, P < 0.001$ ; line 2, Tukey-HSD test,  $P < 0.001$  ( $P < 0.01$  for 44.8.2); line 3, one-way ANOVA,  $F_{4,9} = 829.7, P < 0.001$ ; line 3, Tukey-HSD test,  $P < 0.001$ ). These results indicated that the populations evolved adaptively by compensating for the inhibitory effect of the increasing temperature.



**Figure 2.10:** Fitness of Q $\beta$  18 mut and evolved populations of each line. The fitness was measured at 37.2°C, 43.7°C, 44.8°C, and 45.3°C for Q $\beta$  18 mut and endpoint 43.7°C-, 44.8°C-, 45.3°C-, and 37.2°C-adapted Q $\beta$  populations, as indicated. (A) Fitness of 18 mut and the line 1 Q $\beta$  population. (B) Fitness of 18 mut and the line 2 Q $\beta$  population. (C) Fitness of 18 mut and the line 3 Q $\beta$  population. The fitness was calculated as  $\log_{10}(N_5/N_0)$ , where  $N_5$  and  $N_0$  represent 5 h post-infection and initial (0 h) free phage density (PFU/mL), respectively. For each assay temperature and each line, 18 mut was included as a control, and at each assay temperature, all populations of each line were measured on the same day. Data are expressed as means  $\pm$  standard deviations (n=2-14).

### 2.3.5 Molecular evolution of Q $\beta$ populations

In order to identify the mutations responsible for thermal adaptation experiment, the whole-genome sequence of the aforementioned 12 endpoint Q $\beta$  populations were determined (Table 2.3, Fig. 2.11, and Appendix A). As noted earlier, the Q $\beta$  18 mut had 17 point mutation and 1-bp insertion in a untranslated region (UTR) at the 5' end of the genome [25, 26]; this resulted in an increase in genome size to 4,218 bases. The sequences of endpoint populations were compared with Q $\beta$  18 mut or Q $\beta$  Anc(P1). Many point mutations were detected, some of them were fixed and some of them were polymorphic (Table 2.3 and Fig. 2.11). The site was detected as polymorphic in sequencing chart where a double peak appeared. Total 177 mutations in 59 sites including fixed and polymorphic were observed. Three identical nonsynonymous mutations were detected in three lines with adaptation to 43.7°C; these included A1781C, U3784C, and C3879G/A. Other mutations (1–4) were detected in each of the three lines. The identical mutation at position 1781 was polymorphic for A and C at 43.7°C but this site was changed to A (44.8\_1) and C (44.8\_3) at 44.8°C adaptation and C (45.3\_2) at 45.3°C, the second identical mutation at position 3784 was polymorphic for U and C at 43.7°C but this site was changed to C (44.8\_1 and 44.8\_2) and U (44.8\_3) at 44.8°C adaptation and the third identical mutation at position 3879 was polymorphic between C and G in 43.7\_1 and 43.7\_3 but in 43.7\_2 it was polymorphic between C and A, this polymorphic mutation between C and A was changed to A and G in 44.8\_2, when adaptation temperature was increased to 45.3°C the polymorphic site was changed to monomorphic G in 45.3\_2 (Table 2.3). In response to adaptation at 44.8°C, an additional 7-, 13-, and 8-point mutations were added to lines 1, 2, and 3, respectively and 3, 0, and 2 mutations became undetectable. In the 44.8°C-adapted population, two identical mutations, i.e., C3659U and C3879G/A (appeared at 43.7°C-adapted populations), were observed in all three lines.

In adaptation at 45.3°C, an additional 13-, 6-, and 10-point mutations were added to lines 1, 2, and 3, respectively, and 4, 7, and 0 mutations became undetectable. In

total, mutations at 20, 16, and 21 sites in lines 1, 2, and 3, respectively were identified in association with the stepwise adaptation process from 43.7°C to 45.3°C. Overall, these represent 38, 34, and 39 mutations from the ancestral Q $\beta$  Anc(P1) in lines 1, 2, and 3, respectively of the 45.3°C-adapted populations. These results suggest that Q $\beta$  could adapt to these elevated temperatures with only point mutations; these mutations account for 0.8–0.9% of the total RNA genome (Table 2.3 and Figs. 2.11 and 2.12). In the 45.3°C-adapted populations, 4 specific mutations, A52G, G1494A, C3659U, and C3879G were observed in all three lines; interestingly, in all four of these mutations first appeared in the populations that had adapted to 43.7°C or 44.8°C. Nineteen additional mutational sites first appeared during adaptation at 45.3°C, but these are not common to all three lines. During the reverse adaptation at 37.2°C, an additional 3-, 4-, and 8-point mutations were introduced and 1, 1, and 5 mutations became undetectable in line 1, 2, and 3, respectively, and additional identical mutation (C3545U), which was first appeared in the populations during adaptation at 44.8°C and 45.3°C, was observed in all three lines. In addition, the mutation fixation rate was  $8.20 \times 10^{-6} \pm 8.25 \times 10^{-7}$  per base per generation (Fig. 2.13) and mutations introduced during this adaptation have a tendency to occur frequently in UTR and A1 but not randomly in all the genes (Table 2.4).

**Table 2.3:** Nucleotide sequences of Anc(P1), 18 mut, and evolved phages.

Anc(P1)	Sequence identity in the indicated phage populations <sup>a</sup>																		Gene and/or site	Genome position <sup>b</sup>	Anc(P1)	Nucleotide in: Evolved population	Gene position <sup>c</sup>	Codon change	Amino acid change
	43.7.1	44.8.1	45.3.1	37.2.1.1	43.7.2	44.8.2	45.3.2	37.2.2	43.7.3	44.8.3	45.3.3	37.2.3	UTR	4	G	A									
+	+	+	+	+	+	+	+	+	+	+	+	+	UTR	4	G	A									
+	68	+	+	+	25	+	+	+	+	+	+	+	UTR	14	U	A + I insertion									
													UTR	39	U	C									
													UTR	47	G	A									
													UTR	51	A	G									
													UTR	52	A	G									
													UTR	54	A	G									
													A2	C	U	GAC→GAU	81(26)								
													A2	A	C	GAA→GAC	93(30)		Glu→Asp						
+	+	+	+	+	+	+	+	+	+	+	+	+	A2	U	C	CGU→CGC	132(43)								
+	+	+	+	+	+	+	+	+	+	+	+	+	A2	U	C	AAU→AAC	168(55)								
													A2	C	U	ACG→ACU	360(119)								
													A2	U	C	GUU→GCU	398(132)		Val→Ala						
													A2	U	C	GUU→GUC	771(256)								
													A2	U	C	GCU→GCC	774(257)								
													A2	G	A	GCG→GGA	792(263)								
													A2	A	G	GAA→GGA	845(281)		Glu→Gly						
													A2	A	G	CAA→CAG	1005(334)								
													A2	A	G	GAU→GGU	1028(342)		Asp→Gly						
													A2	G	A	GCG→CGA	1062(353)								
													A2	G	A	GCG→GGA	1098(365)								
													A2	A	G	CAU→CGU	1137(385)		His→Arg						
													A2	U	G/(37.2+3 C)	GUU→GUG/GUC	1191(396)								
													A2/S-site	1251	C	U	ACC→ACU	1197(398)							
+	+	+	+	+	+	+	+	+	+	+	+	+	A2/S-site	1257	C	U	AGU→AGG	1206(401)		Ser→Arg					
													A2/S-site	1266	U	G	GUU→GUC	1221(406)							
													A2/S-site	1281	U	C	UUU→UGU	1235(411)		Phe→Cys					
+	+	+	+	+	+	+	+	+	+	+	+	+	A2/S-site	1295	U	G	GUA→AUA	1252(417)		Val→Ile					
+	+	+	+	+	+	+	+	+	+	+	+	+	A2/S-site	1312	G	A	AGU→AGU	28(9)		Gly→Ser					
+	+	+	+	+	+	+	+	+	+	+	+	+	Coat / A1	1371	G	A	AGU→ACC	57(18)							
+	+	+	+	+	+	+	+	+	+	+	+	+	Coat / A1	1400	U	C	GUU→AUU	151(50)		Val→Ile					
													Coat / A1	1494	G	A	CGC→CGU	261(86)							
													Coat / A1	1604	C	U	GCG→GGU	432(143)							
													A1	1775	G	U	UCA→UUA	434(144)		Ser→Leu					
													A1	1777	C	U	AAA→AAC	438(145)		Lys→Asn					
													A1	1781	A	C	GGU→GAU	488(162)		Gly→Asp					
													A1	1831	G	A	GUU→AUU	529(176)		Val→Ile					
													A1	1872	G	A	AAC→GAC	550(183)		Asn→Asp					
													A1	1893	A	G	AAA→CAA	613(204)		Lys→Gln					
													A1	1956	A	C	AGU→AGG	663(220)		Ser→Arg					
													A1	2006	U	G	UUC→GUC/CUC	673(224)		Phe→Val/Leu					
													A1	2016	U	G/(37.2+3 C)	UAU→CAU	718(239)		Tyr→His					
													A1	2061	U	C	CAG→CAA	735(244)							
													A1	2078	G	A									

Continued on next page

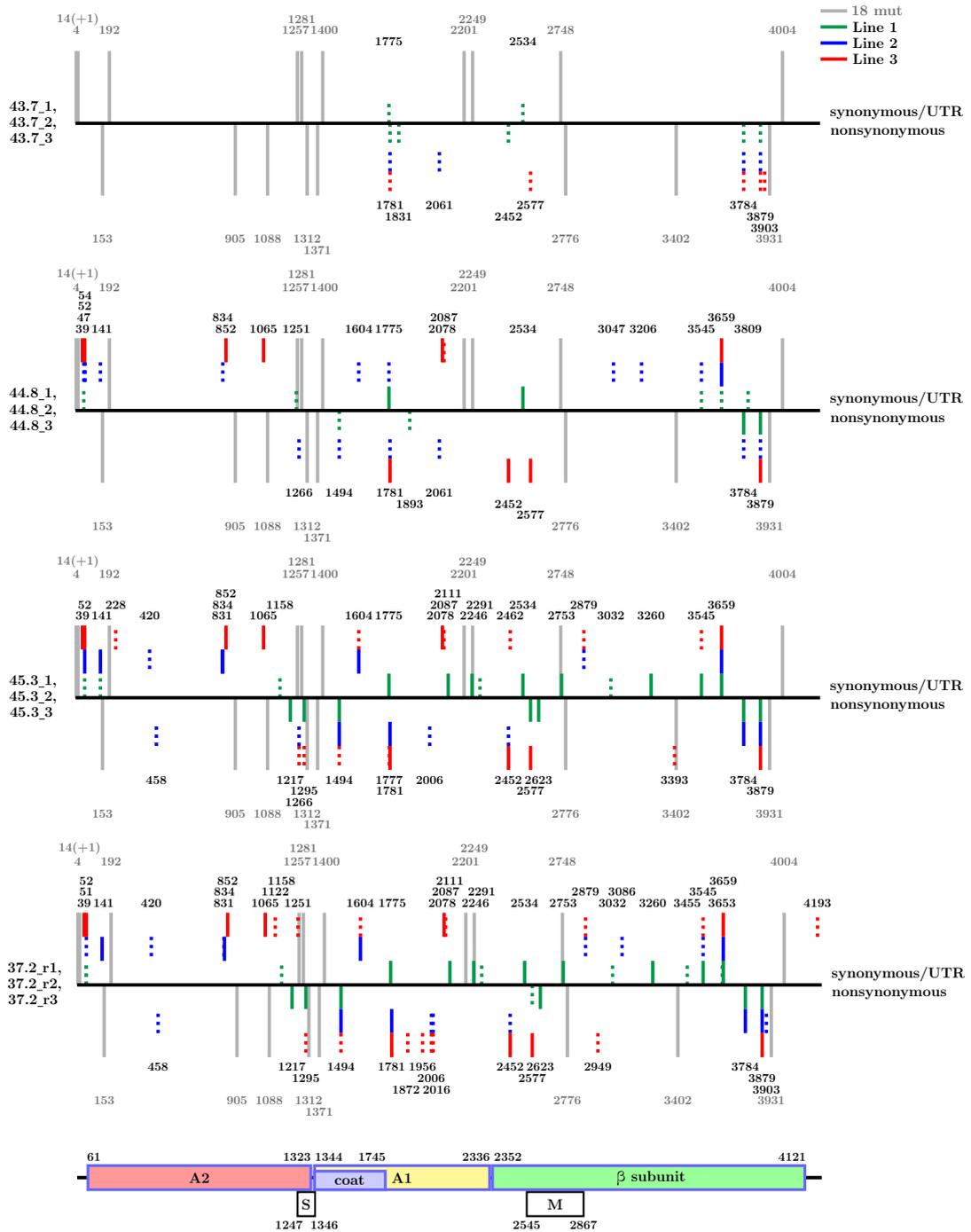
**Table 2.3:** Continued from previous page.

Anc(P1)	Sequence identity in the indicated phage populations <sup>a</sup>																		Gene position <sup>b</sup>	Gene and/or site	Genome position <sup>b</sup>	Nucleotide in:		Gene position <sup>c</sup>	Codon change	Amino acid change
	18 mut	43.7.1	44.8.1	45.3.1	37.2.F1	43.7.2	44.8.2	45.3.2	37.2.r2	43.7.3	44.8.3	45.3.3	37.2.r3	Anc(P1)	Evolved population	Gene position <sup>c</sup>	Codon change	Amino acid change								
														U	C	744(247)	CGU→CGC									
														G	A	768(255)	GAG→GAA									
														C	U	858(285)	GCC→GCU									
														C	U	903(300)	UCC→UCU									
														C/U	U	906(301)	AGC→AGU									
														U	C	948(315)	ACU→ACC									
														C	U	101(33)	GCC→GUC	Ala→Val								
														A	G	111(36)	UUA→UUG									
														G	A	183(60)	GGC→GGA									
														A	C/(45.3.1, 37.2.r1 U)	226(75)	AUG→CUG/UUG	Met→Leu								
														U	C	272(90)	GUU→GCU	Val→Ala								
														A	C	397(132)	AGA→CGA									
														A	G	402(133)	AAA→AAG									
														U	C	425(141)	GUU→GCU	Val→Ala								
														G	A	528(175)	CCG→CCA									
														A	G	598(199)	AUU→GUU	Ile→Val								
														U	C	681(226)	GGU→GGC									
														C	U	696(231)	UUC→UUU									
														U	G	735(244)	CGU→CGG									
														C	U	855(284)	GCC→GCU									
														A	G	909(302)	AGA→AGG									
														G	A	1042(347)	GAC→AAC	Asp→Asn								
														U	C	1051(350)	UCG→CCG	Ser→Pro								
														C	A	1104(367)	GUC→GUA									
														C	U	1194(397)	GGC→GGU									
														G	G	1302(433)	ACA→ACG									
														C	U	1308(435)	GAC→GAU									
														U	C	1433(477)	AUC→ACC	Ile→Thr								
														G	A	1458(485)	GGG→GGA									
														C	G/(43.7.2 A, 44.8.2 A and G)	1528(509)	CUC→GUC/AUC	Leu→Val/Ile								
														C	A	1552(517)	CUC→UUC	Leu→Phe								
														U	C	1580(526)	CUU→CCU	Leu→Pro								
														G	A	1653(550)	ACC→ACA									
														A	UTR	4163										

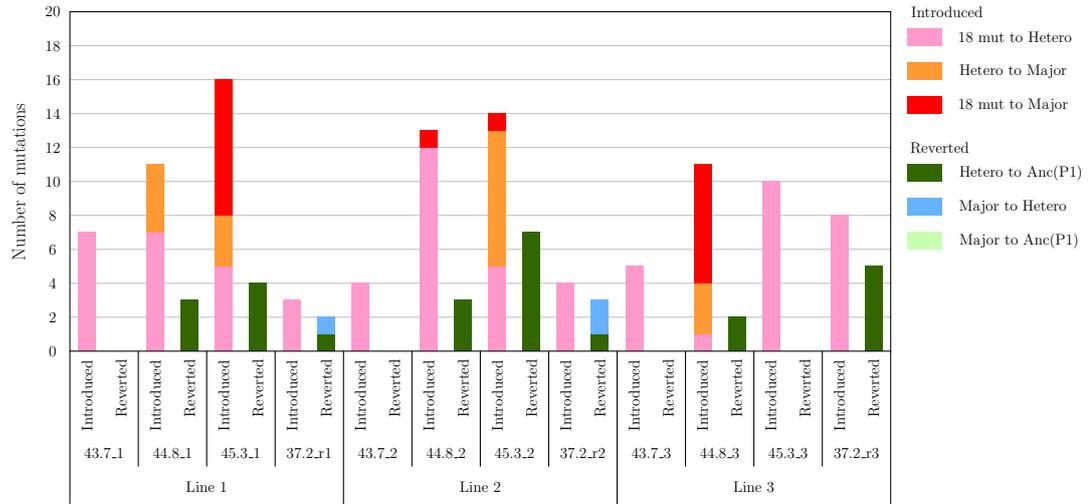
<sup>a</sup>Sequences are indicated as follows; blanks represent the same sequence as Q $\beta$  Anc(P1), + highlighted in dark blue represents change in more than 90% of the sequences, and values highlighted in light blue represent the percentage of sequences with changes from the ancestral sequence, + without highlight represents 18 mut sequence.

<sup>b</sup>Gene positions were counted from G of the 5' terminus of the Q $\beta$  genome as position 1.

<sup>c</sup>Gene position was counted from the A of the start codon (AUG) of each gene as position 1. The numbers in the parentheses represent the amino acid positions of each protein in which the position of second codon was counted as position 1. Position 2249 of Anc(P1) was heterogeneous with C and U although the Anc(P1) Q $\beta$  was derived from cloned cDNA in previous study [25].



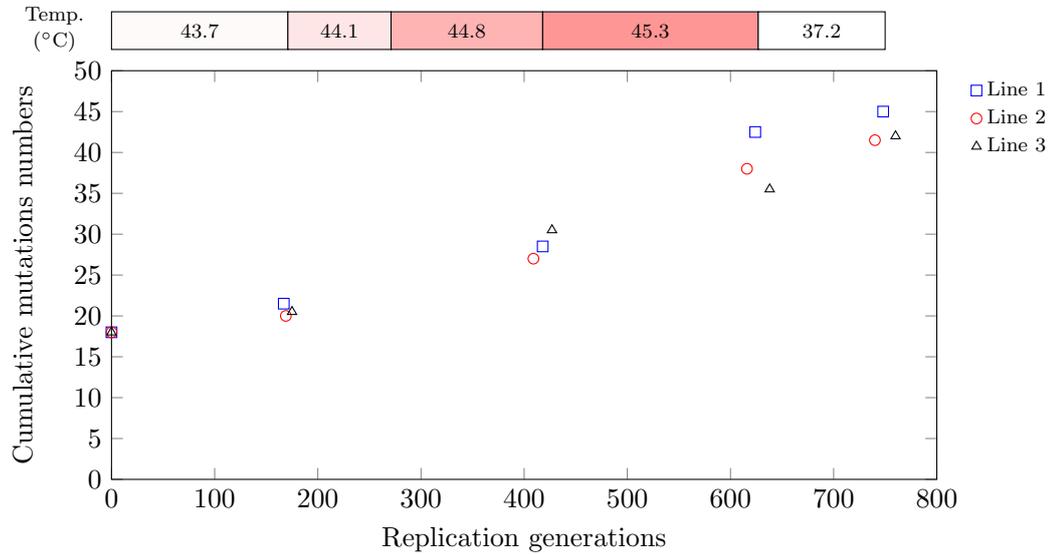
**Figure 2.11:** Mutation positions of the evolved population of each Q $\beta$  line. Marked positions are where differences between the sequence of Anc(P1) and that of each Q $\beta$  populations were observed. Positions of mutations in lines 1, 2, and 3 are indicated on the graph. Solid and dotted lines represent monomorphic and polymorphic sites, respectively. Vertical lines above the black horizontal lines represent synonymous mutations or mutations in the UTR, and vertical lines below the black horizontal line represent nonsynonymous mutations. The numbers above and below the vertical lines represent the mutation positions. The four genes, S site, and M site are represented in the lowest part, with numbers representing the start and end positions.



**Figure 2.12:** Mutations introduced and reverted to original sequence with adaptation. Hetero or major were determined as the ratio of mutant/ancestral sequences in the adapted population as ratio  $10\% < \text{ratio} < 90\%$ , or ratio  $\geq 90\%$ , respectively.

**Table 2.4:** Tendency of mutations in thermal adaptation within individual Q $\beta$  genes.

Gene	Gene length (bases)	Mutational positions		Mutations/gene length	
		from 18 mut to 45.3°C adaptation	of 45.3°C adaptation with 18 mut	from 18 mut to 45.3°C adaptation	of 45.3°C adaptation with 18 mut
UTR	192	4	6	0.021	0.031
A2	1263	13	20	0.010	0.016
Coat	402	2	4	0.005	0.010
A1	591	12	14	0.020	0.024
$\beta$	1770	18	23	0.010	0.013



**Figure 2.13:** Time course of changes in cumulative mutations in the Q $\beta$  genome. To determine mutation fixation and cumulative mutation numbers in Table 2.3 and Fig. 2.11, polymorphic changes (Anc  $\rightleftharpoons$  polymorphic  $\rightleftharpoons$  monomorphic) were counted as 0.5, and monomorphic changes (Anc  $\rightleftharpoons$  monomorphic) were counted as 1. The adaptation temperatures are presented in the upper part, where the average generation number for three lines representing the start and end positions of adaptation temperatures.

## 2.4 Discussion

In this Chapter, the patterns of thermal adaptation was explored in a population of ssRNA bacteriophage Q $\beta$  by stepwise increase to the highest known growth temperature 45.3°C, a temperature at which the Q $\beta$  18 mut strain could not grow. The results showed that Q $\beta$  can grow and replicate at this temperature within 52 days (616 generations) when starting with the Q $\beta$  18 mut variant and within 114 days (1238 generations) when starting with ancestral Q $\beta$ , which has an optimum growth temperature of  $\sim$ 37°C, was used as the starting material. The 45.3°C adapted population had at most 21 mutations from 18 mut and 39 mutations from ancestral Q $\beta$ . Overall, thermal adaptation results in mutations in 0.8–0.9% of the Q $\beta$  genome.

Although the specific growth rate of *E. coli* depends directly on the culture medium, ordinary laboratory strains can not grow in minimal media at temperatures greater than 43°C [49]; the rate of polypeptide synthesis begins to decrease dramatically at temperatures over 44°C [50]. However, the experiments featured in this study used the *E. coli* 46L-1F' strain that can grow effectively in temperatures ranging from 37.2°C to 45.3°C. This technical advance made it possible to evaluate Q $\beta$  adaptation to thermal changes without the confounding effect of host bacteria. I found that Q $\beta$  adapted to grow and replicate at 45.3°C had an overall increased temperature range because these populations could grow with equivalent fitness at 37.2°C. Intriguingly, even after returned the 45.3°C-adapted populations to 37.2°C for 122–124 generations, little to no decrease in fitness was observed. These results clearly indicate that Q $\beta$  gained the potential for growth at higher temperatures without showing trade-off in the lower ranges.

In the course of adaptation to grow at 43.7°C, the adaptation of Q $\beta$  18 mut to the thermally adapted *E. coli* 46L-1F' host, may have included not only adaptation of Q $\beta$  to this temperature because Q $\beta$  18 mut derived from cDNA did not grow on 46L-1F' before. Therefore, Q $\beta$  18 mut could take different pathways for adaptation before se-

lection at warmer temperature. Total 10 mutational sites were detected in three lines, three mutational sites were identical in three lines and non-synonymous mutations. The endpoint populations showed higher fitness than 18 mut, which is the consensus sequence of the populations that had already been adapted at 43.6°C using 43BF' as the host strain in previous study [25], at higher temperatures (Fig. 2.10). There are at least two plausible explanation for the higher fitness of the endpoint populations. The first is the temperature 43.7°C acts as a selective pressure for Q $\beta$  18 mut and mutations were accumulated. The second is the *E. coli* 46L-1F' strain, which was generate from thermal adaptation experiment after 8829 generation, was used as host in this evolutionary experiment and the *E. coli* 43BF' strain, which was generated from the thermal adaptation experiment after 4403 generation, was used in previous study [25], so the adaptive mutations can be introduced due to the use of different host strains. I shall investigate which environmental factor (temperature or host) acts as selective pressure at this temperature adaptation (Chapter 3).

Previously, Q $\beta$  was passaged at 37.2°C for  $\sim$ 120 generations in three replicates from Anc(P1) using the same method [25]. Kashiwagi et al. (2014) observed C2249U in all the three replicates and another 1 point mutation in the 5' UTR (G4G/A) in one of three replicates [25]. Therefore, in this study, most of the mutations were observed might be introduced upon thermal adaptation. However, as reported by Singhal et al. (2017) with regard to the importance of considering the effects of multiple selective pressures, even in environments where a single factor is changed [34], further analysis is required to determine the effects of each mutation on the fitness increase at 45.3°C. In addition, the introduction of some mutations was observed at the 37.2°C passage with the 45.3°C-adapted populations. To compare the result with the previous 37.2°C passage starting from Anc(P1), the mutation fixation number per generation was counted using polymorphic sites and monomorphic sites and was found to be 0.5 and 1, respectively. The mutation number per generation of the 37.2°C passage using 45.3°C-adapted populations ( $3.4 \times 10^{-2}$  mutations/generation) was larger than that of the 37.2°C passage starting from Anc(P1) ( $5.7 \times 10^{-3}$  mutations/generation). This

difference might be because of the difference in the sequences used as the starting phage population.

The Q $\beta$  genome encodes four proteins: A2, coat, A1, and  $\beta$  subunit for Q $\beta$  replicase. A2 is a multifunctional protein that binds the F pili of *E. coli* and lysis function via binding with MurA [18, 51–54]. Cryo-electron microscopic analysis has clearly revealed the Q $\beta$  structure. The  $\beta$ -sheet-rich region of A2 protrudes from the capsid, while the  $\alpha$ -helix-rich region is within the capsid [55]. The  $\beta$ -sheet-rich region interacts with the F pili of *E. coli* and MurA, while the  $\alpha$ -helix-rich region interacts with the RNA genome [56, 57]. Two amino acid mutations, Ser401Arg and Phe411Cys, in two of three lines as well as two amino acid mutations, Val132Ala and His385Arg, in one of three lines were observed. Mapping these mutational positions on the A2 structure (Protein Data Bank 5MNT) reveals the following: (i) Ser401 is located near the position where coat and A2 interact [56] (Fig. 2.14a). (ii) Phe411 is located in  $\alpha 9$  and is close to the RNA-binding region [58] (Fig. 2.14b). (iii) Val132 and His385 are closely located at the protruding region (Fig. 2.14a). The region between the 30<sup>th</sup> to 120<sup>th</sup> amino acids of the N-terminus of A2 is in contact with MurA [56], and no mutations, except Glu30Asp, are observed in this region.

In coat protein, Val50Ile mutation was observed in all three lines (Fig. 2.14c). Asn30, Thr49, Ser51, and Gln65 produce an adenine-binding pocket to fit the A nucleotide of the operator of the  $\beta$  subunit gene [59]. When it is mapped on the structure (Protein Data Bank 4L8H), Val50 is also located near the operator sequence (Fig. 2.14d), indicating that this mutation might be related to the translation of the  $\beta$  subunit. The A1 gene shares the same initiation codon as the coat gene. When the coat stop codon is suppressed at low probability, A1 is produced as the read-through protein. Thus, the N-terminus amino acid sequence of A1 is the same as that of coat, and A1 has an additional 197 amino acids in its C-terminus; the additional region is located outside the capsid [60]. A1 is essential for producing infectious phage particles; however, its detailed function is unclear [61]. Lys145Asn in lines 2 and 3 and Ser144Leu in line 3 were observed; however, the Ser144 and Lys145 positions in A1 were missing in the

crystal structure in Protein Data Bank 3RLC. The functional effects of these amino acid changes are unclear.

Q $\beta$  replicase formed by the assembly of  $\beta$  subunit from Q $\beta$ , translational elongation factors EF-Tu and EF-Ts, and ribosomal protein S1 from *E. coli* [22, 62]. The  $\beta$  subunit structure comprises finger, thumb, and palm domains; the finger and thumb domains interact with EF-Tu and EF-Ts, respectively, while active site residues are in the palm domain [63]. Eight nonsynonymous mutations in the  $\beta$  subunit gene were observed. Ala33Val, Met75Leu, and Ile477Thr were found in two of three lines, while Leu509Val (mutation 3879) was found in all three lines. Tyr510 interacts with the 5' terminus of newly synthesized RNA in the processive elongation stage, and Tyr510 mutation decreases replication activity [64]. Since Leu509Val is the next amino acid of Tyr510, this amino acid mutation might be related to replication activity. When I mapped on the structure (Protein Data Bank 4Q7J), Val141Ala, Asp347Asn, and Ser350Pro were located near the region to interact with S1 (Fig. 2.14e) [65] and Ala33, Met75, and Ile477 were located on the thumb domains. The functional effects of these amino acid changes are unclear.

To understand the mechanisms underlying adaptation and speciation, the concept of the fitness landscape has been proposed [66, 67]. Fitness landscape is essentially the relationship between genotype and response, or fitness for a given environment. The direction of evolution may be predicted with a strong understanding of the shape of a given fitness landscape [68]. Theoretical models, for example the NK model, the Mt. Fuji model, and the rough Mt. Fuji model, have been presented in support of this concept [69–71]. Several studies focused on evolution of proteins, RNA aptamers, and ribozymes have been presented with respect to the shape of fitness landscape [72–76]. For example, in a comprehensive study of short-length RNA, fitness peaks were isolated from one another [77]. In this study, when it was assumed that the mutations existing in over 90% of the sequences in a given population were fixed or nearly fixed, then most of the mutational sites were polymorphic (Table 2.3 and Fig. 2.11). Furthermore, it was found that when populations adapted to 45.3°C were re-adapted

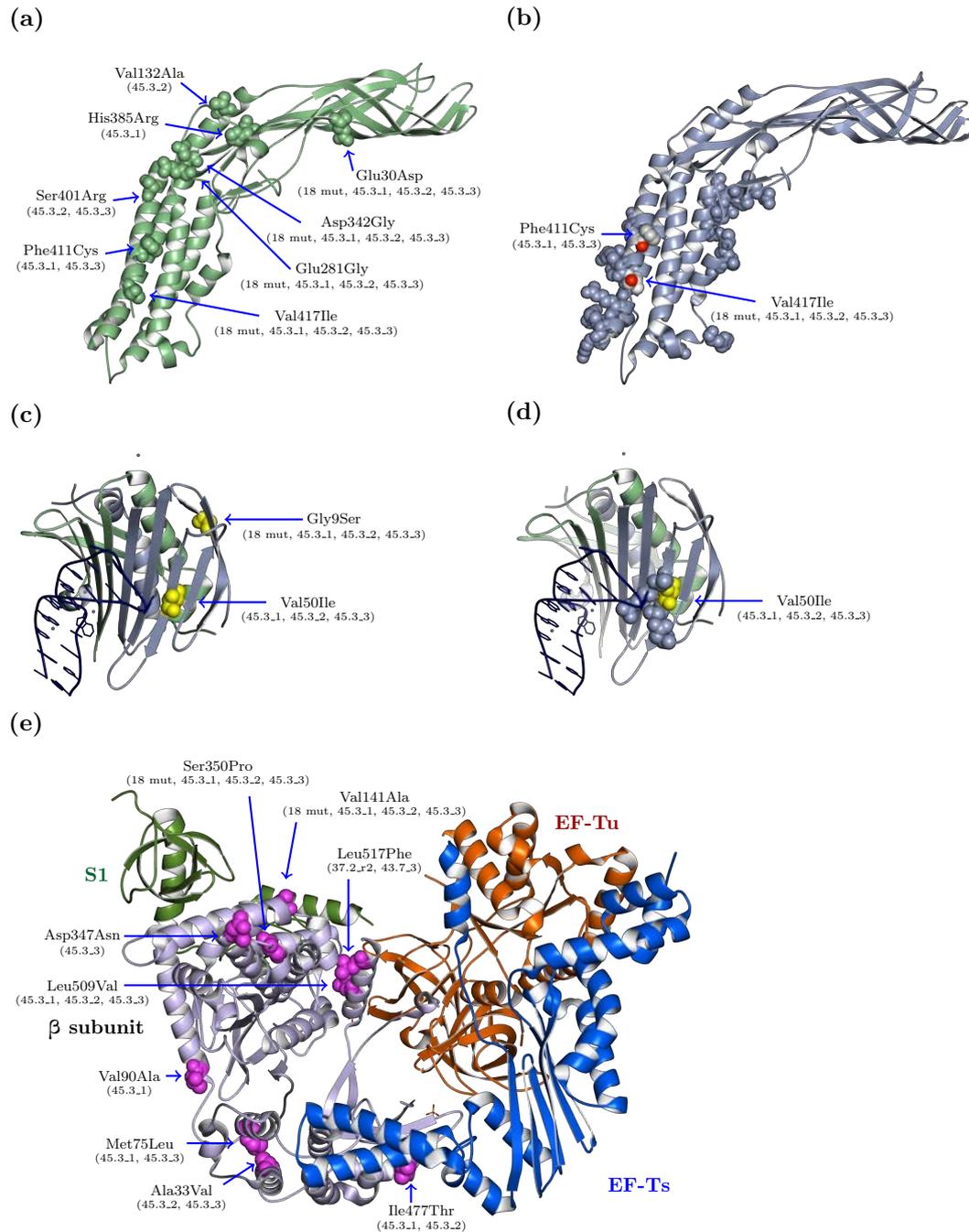
at 37.2°C, fitness at the higher temperature was maintained; under these conditions, 9 mutated sites returned to the original sequence including Hetero to Anc(P1) and Major to Hetero (Table 2.3 and Fig. 2.12). These results suggest that the fitness landscape of Q $\beta$  at 45.3°C may be flat rather than rugged. Two scenarios can be considered to explain these results. First, these findings may be due to the high fitness of the reverted mutations for the 37.2°C landscape; second, is the possibility that these are hitchhiking mutations associated with other adaptive mutations. Because there are many polymorphic sites in the 45.3°C population and existed as quasispecies, the reversions may be due to the hitchhiking.

The frequency of both nonsynonymous and synonymous mutations increased in the population (Table 2.3 and Fig. 2.11). The Q $\beta$  RNA genome is a multifunctional RNA that acts as an mRNA-encoding protein, a genome replication template, and regulatory RNA for controlling the level and timing of protein expression in the Q $\beta$  life cycle [18]. In addition, the secondary structure of ssRNA genome has important functions with respect to bacteriophage formation. Among its roles, the RNA genome has to be compact to be packaged appropriately into the capsid, it should have specific contacts with the coat protein inside the capsid, and should be able to recruit coat protein to assemble to make capsid [55, 57, 78]. The propagation processes of synonymous mutations have been extensively investigated by experimental evolution of ssDNA and ssRNA viruses, in which synonymous mutations and mutations in intergenic regions were fixed during adaptation to elevated temperatures [31–33]. In Q $\beta$  thermal adaptation, not only nonsynonymous mutations but also synonymous mutations contributed to fitness increase [25, 26]. Therefore, some, and not all, synonymous mutations introduced in this study may have been adaptive. The secondary structures predicted for the thermally-adapted RNA sequences are shown in Figs. 2.15 and 2.16 after introducing mutations observed in each populations, including those that were polymorphic and at over 90% frequency in each populations. The secondary structures prepared after including those mutational sites that were polymorphic were grouped into two types; one was ancestral Anc-type, and the other was 44.8.2-type, the 44.8.2-type

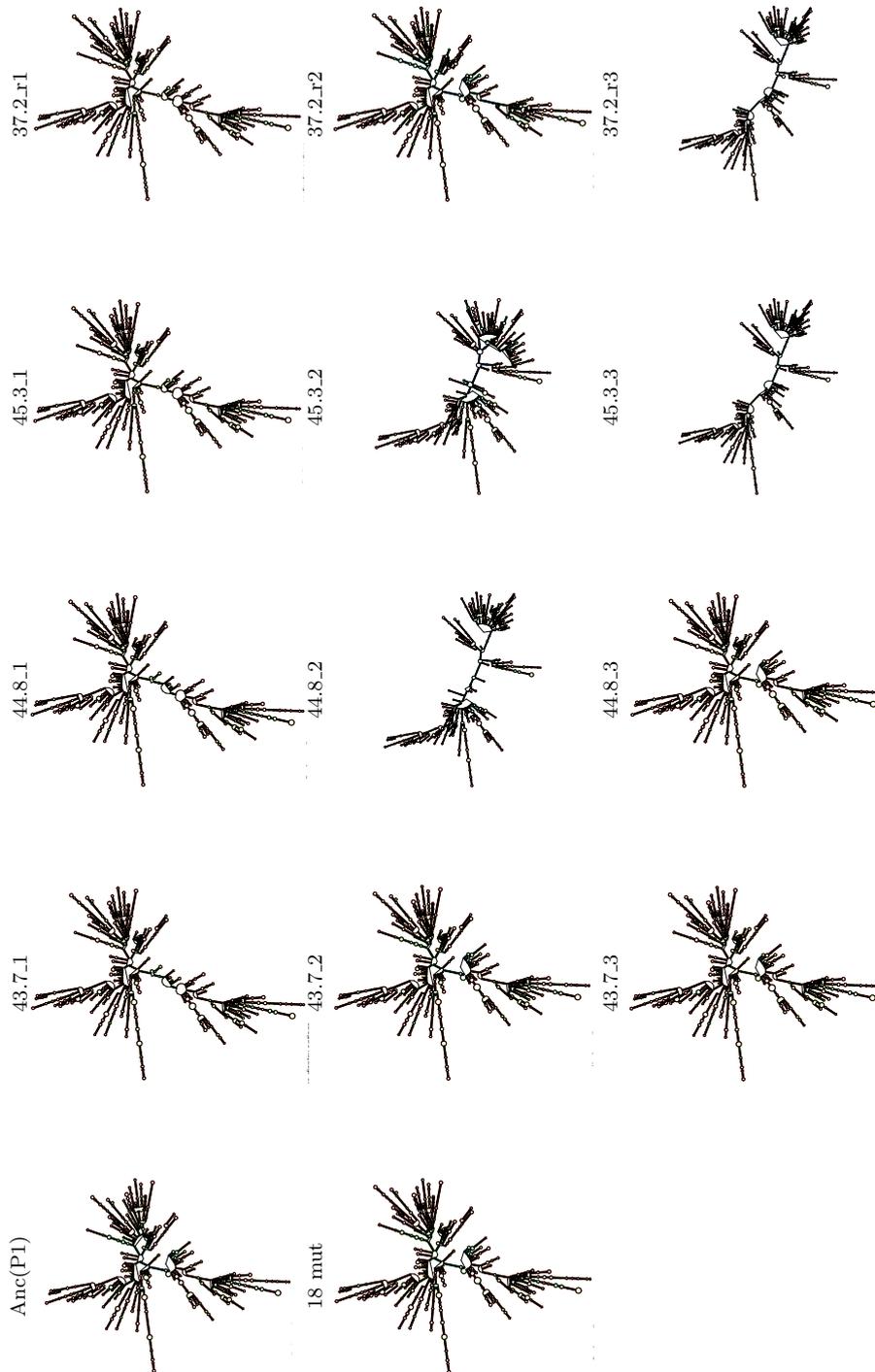
group included variants 45.3\_2, 45.3\_3, and 37.2\_r3 (Fig. 2.15). Interestingly, the secondary structure of 37.2\_r2 reverted to the Anc-type from 44.8\_2-type (Fig. 2.15). The secondary structures, prepared after introducing mutations observed at over 90% frequency in each populations, changes slightly of 44.8\_1, 45.3\_1, and 37.2\_r1 or 45.3\_2 and 37.2\_r2, which correspond to the beginning region of the  $\beta$  subunit gene or the region between the A1 gene and the beginning of the  $\beta$  subunit gene. Some nonsynonymous mutations became frequent and were lost later, possibly as a result of the following mechanisms: clonal interference, epistasis, or hitchhiking when introducing additional adaptive mutations into the population [32].

In experiments performed with the thermally adapted *E. coli*, the bacteriophage Q $\beta$  adapted to the higher temperature (45.3°C) within 114 days (approximately 1260 generations); interestingly, the *E. coli* required almost 537 days (equivalent to 7780 generations) to reach the same endpoint [27]. These results underscore our observations regarding the rapid adaptation of RNA phages to new environments. Bacteriophages are critical components of the global ecosystem. Therefore, additional studies regarding their capacity for adaptation and the importance of the mutations, which were observed in the present study are warranted.

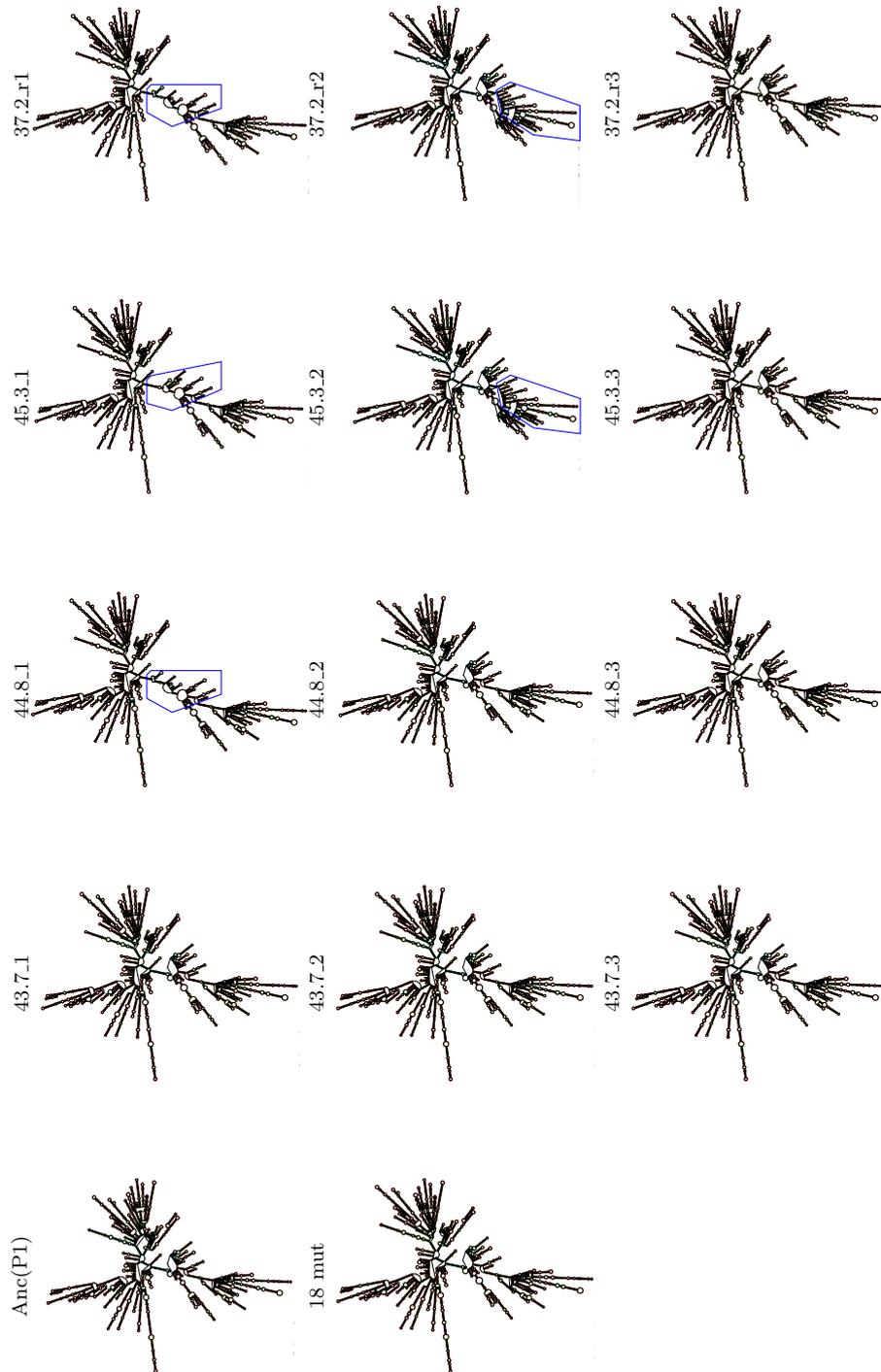
Form this Chapter, an article has been published [79] as a paper in viruses. In next Chapter, I shall clarify about selective pressure (temperature or host), which environmental factor played as selective pressure at 43.7°C, and investigate the growth characteristics of Q $\beta$  mutants, which were observed in all lines at 43.7°C-endpoint populations.



**Figure 2.14:** Mutational positions of amino acids mapped into the structure. The amino acid position and the population name in which the mutation was observed is in parentheses. (a) The positions of amino acid mutations in A2 were mapped on protein data bank 5MNT. (b) The positions of amino acid mutations in A2 were mapped in the same manner as (a), but the position shown in gray filled circles represent the amino acid that interacts with the RNA genome [58]. (c) The positions of amino acid mutations in coat protein were mapped on protein data bank 4L8H. (d) The positions of amino acid mutations in coat protein were mapped in the same manner as (c), but the position shown in gray filled circles represent the amino acids those produce an adenine-binding pocket to fit the A nucleotide of the operator of the  $\beta$  subunit gene [59]. (e) The positions of amino acid mutations in  $\beta$  subunit of Q $\beta$  with the structure of S1, EF-Tu and EF-Ts (protein data bank 4Q7J).



**Figure 2.15:** Secondary structure of Q $\beta$  Anc(P1), 18 mut and endpoint populations. The secondary structure of each RNA genome was estimated using RNAfold [46] after introducing mutations observed in each population, including those that were polymorphic. Major changes were observed in 44.8.2, 45.3.2, 45.3.3, and 37.2.r3 populations.



**Figure 2.16:** Secondary structure of starting and evolved (fixed mutations) Q $\beta$  genomes. The secondary structure of each RNA genome was estimated using RNAfold [46] after introducing mutations observed at over 90% frequency in each population. Anc(P1) is a starting phage cultured from 37.2°C in the previous study [25]; 18 mut is a starting phage used in this study. Minor differences were observed in the closed region.

## Chapter 3

# Study on Q $\beta$ mutants obtained in evolution experiment

### 3.1 Introduction

Mutations, changes in the genetic code, are the ultimate source of diversity for evolutionary processes [80–82]. Although the beneficial mutations, which increase fitness by allowing organisms to adapt in the selective environment, are rare [83], populations can adapt by fixing a series of beneficial mutations [84]. In addition, changes in environmental factors (e.g. resource availability, temperature, community dynamics, host etc.) alter the natural selection, influencing the selective benefits of subsequent adaptations that improve an organism’s fitness in the new conditions [85]. Therefore, understanding the properties of mutations is a key step toward predicting how natural populations will respond to changes in their environments.

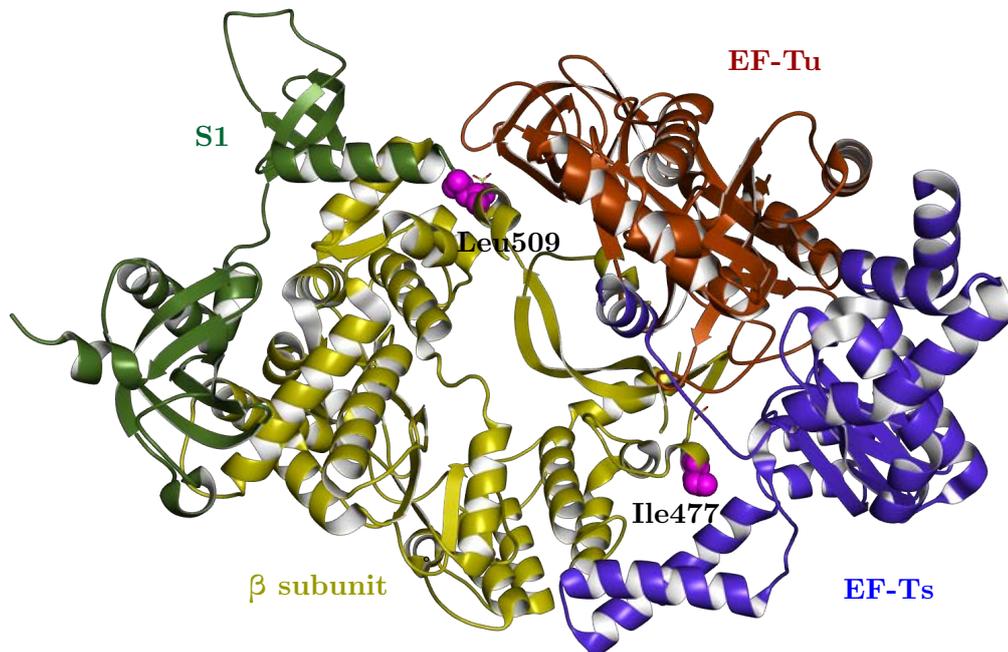
A thermal adaptation experiment of Q $\beta$  was conducted using *E. coli* 43BF’ as a host strain, which was generated from the thermal adaptation experiment after 4403 generation [49], between 37.2°C and 43.6°C with three independent lines [25]. The previous thermal adaptation was continued using Q $\beta$  18 mut, which was one of the three replicates of previous thermal adaptation that had adapted to 43.6°C, and *E. coli* 46L-1F’ strain, which was generated from thermal adaptation experiment after 8829 generation

[27], between 43.7°C and 45.3°C with three independent lines (Chapter 2, [79]). In 43.7°C adapted endpoint populations, total 10 mutational sites were observed in three lines (Table 2.3 and Fig. 2.12) and showed comparatively higher fitness than 18 mut at higher temperature (Fig. 2.10). In general, when bacteriophages adapt to new environments, this indicates that they have not only adapt to the new environment but also to bacterial hosts living there. Therefore, the adaptation temperature 43.7°C using *E. coli* 46L-1F' strain can serve as two selective pressure (temperature and host) for Q $\beta$  18 mut, it is not clear which environmental factor act as selective pressure for Q $\beta$  18 mut and/or which mutations are important for improving the fitness of the populations. In order to elucidate the selective environmental factor and growth characteristics of the endpoint populations, it is necessary to prepare the mutants and to determine the growth ability using host strains 43BF' and 46L-1F'. This will allow to determine the selective pressure and growth characteristics of the endpoint populations.

Temperature and host are critical environmental factors that has had enormous influence on the life history of bacteriophage. In the course of adaptation using *E. coli* 46L-1F' as host strain to grow at 43.7°C, three nonsynonymous mutations were observed at the same positions in the genome in all three lines; these included A1781C (Lys145Asn in A1), U3774C (Ile477Thr in  $\beta$  subunit) and C3879G/A (Leu509Val/Ile in  $\beta$  subunit) (Table 2.3 and Fig. 2.12). The endpoint populations showed comparatively higher fitness than 18 mut, which was the starting material for the present thermal adaptation experiment, at higher temperature (Fig. 2.10). There are at least two plausible explanation for the higher fitness of 43.7°C-adapted endpoint populations. The first is the temperature 43.7°C acts as a selective pressure for Q $\beta$  18 mut and mutations were accumulated. The second is the *E. coli* 46L-1F' strain, was used as host in this evolutionary experiment and the *E. coli* 43BF' strain was used in previous study [25]. When position of amino acid were mapped on the Q $\beta$  replicase structure (Protein Data Bank 4Q7J), Ile477 and Leu509 were close to the EF-Ts and EF-Tu subunits respectively (Fig. 3.1). Therefore, it is expected that these adaptive

mutations can be introduced due to the use of different host strains. In order to clarify the above two possibilities and growth characteristics of endpoint populations, it is necessary to prepare the mutants and determined the fitness using host strains 43BF' and 46L-1F'. Using these host strains and mutants, it will be possible to determine the selective pressure and growth characteristics of the endpoint populations.

In this Chapter, I prepared four kinds of mutants collaborated with Mr. Toma Yokono (three point mutations and the combination of those three mutations), analysed the fitness of Q $\beta$  mutants at 43.7°C using *E. coli* 46L-1F' and at 43.6°C using *E. coli* 46L-1F' and 43BF' as the host and showed that the adaptation temperature 43.7°C played as selective pressure for Q $\beta$  18 mut and C3879G showed beneficial function in thermal adaption.



**Figure 3.1:** Position of Ile477 and Leu509 amino acids in Q $\beta$  replicase. The structure of Q $\beta$  replicase with the  $\beta$  subunit, EF-Tu, EF-Ts and the S1 was generated from the Protein Data Bank 4Q7J [65]. The position of the amino acids were marked as space ball.

## 3.2 Materials and methods

### 3.2.1 Strains and medium

The *E. coli* 43BF' and 46L-1F' strains were used as the host strain when the thermally adapted Q $\beta$  mutants were characterized. The *E. coli* A/ $\lambda$  strain [38] was used as the host strain for the titer assay. The *E. coli* F<sup>-</sup> strain, DH5 $\alpha$ /pACYCQ $\beta$ -18 mut containing 18 mutations [26], was used for construction of Q $\beta$  18 mut\_mutants. *E. coli* F<sup>-</sup> strains, DH5 $\alpha$ /pACYCQ $\beta$ -18 mut\_mutants, were cultured in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) containing 50  $\mu$ g/mL of kanamycin to maintain their plasmids. Modified-M63 medium (62 mM K<sub>2</sub>HPO<sub>4</sub>, 39 mM KH<sub>2</sub>PO<sub>4</sub>, 15 mM ammonium sulfate, 1.8  $\mu$ M FeSO<sub>4</sub>·7H<sub>2</sub>O, 15  $\mu$ M thiamine hydrochloride, 0.2 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, and 22 mM glucose) [39] with 1 mM L-leucine and 2.3 mM MgSO<sub>4</sub>·7H<sub>2</sub>O (mM63lm) was used for culturing 43BF' and 46L-1F'. LB medium was used for titer assay.

### 3.2.2 Construction of Q $\beta$ mutants from cDNA of genomes

Bacteriophage Q $\beta$  mutants were constructed from cDNA of Q $\beta$  genomes with mutations. Each consensus mutation detected in all three populations was introduced into the cDNA genome of pACYCQ $\beta$ -18 mut using an In-Fusion cloning system (TaKaRa Bio, Inc., Shiga, Japan) and standard cloning methods [86]. Four types of cDNA: DH5 $\alpha$ /pACYC Q $\beta$ -18 mut\_C3879G, DH5 $\alpha$ /pACYCQ $\beta$ -18 mut\_ A1781C, DH5 $\alpha$ /pACYCQ $\beta$ -18 mut\_ U3784C, and DH5 $\alpha$ /pACYCQ $\beta$ -18 mut\_ A1781C\_U3785C\_C3879G were prepared in this study.

Q $\beta$  mutants were obtained from the cDNAs as follows. DHA5 $\alpha$ /pACYCQ $\beta$ -18 mut\_ A1781C, -U3784C, -C3879G, and -A1781C\_U3784C\_C3879G were each cultured in four test tube in 5 mL LB medium with 50  $\mu$ g/mL of kanamycin at 30°C at 160 $\pm$ 1 rpm for 18 hours. After centrifugation at 10,000 rpm at 4°C for 5 min, the supernatants were filtered with 0.2- $\mu$ m-pore-size syringe filters (Steradisc 25; Kurabo

Industries, Ltd.). The filtrated supernatants were concentrated and dialyzed with Amicon Ultra-4 centrifugal filter unit with 50,000 nominal-molecular weight-limit membranes (EMD Millipore). The concentrated phage particles were filtered using 0.2- $\mu$ m-pore-size syringe filters (Minisart RC15 filters; Sartorius) and stored in 40% glycerol at  $-20^{\circ}\text{C}$ . The mutants 18 mut\_A1781C, 18 mut\_U3784C, 18 mut\_C3879G, and 18 mut\_A1781C\_U3784C\_C3879G were designated as A1781C, U3784C, C3879G, and 3(43.7) mut, respectively.

### **3.2.3 Fitness analysis of the Q $\beta$ mutants**

The fitness of Q $\beta$  18 mut, A1781C, U3784C, C3879G, and 3(43.7) mut was evaluated at  $43.6^{\circ}\text{C}$  and  $43.7^{\circ}\text{C}$  using the host strain 46L-1F' according to the method described in Section 2.2.8. In this case, 2.5 mL uninfected log phase of 46L-1F' strain (the  $\text{OD}_{600}$  between 0.210 and 0.282 approximately  $2.1 \times 10^8$  and  $2.82 \times 10^8$  CFU/mL) was infected with approximately  $3.63 \times 10^4$  to  $2.6 \times 10^5$  PFU/mL of Q $\beta$ . The fitness of above Q $\beta$  mutants was also evaluated at  $43.6^{\circ}\text{C}$  using the host strain 43BF' according to the method described by Kashiwagi et al. (2014) [25]. Briefly, 2.5 mL uninfected log phase of 43BF' strain (the  $\text{OD}_{600}$  between 0.300 and 0.356 approximately  $3.0 \times 10^8$  and  $3.56 \times 10^8$  CFU/mL) was infected with approximately  $3.33 \times 10^4$  to  $2.53 \times 10^5$  PFU/mL of Q $\beta$ . The fitness was calculated as  $x = \log_{10}(N_5/N_0)$ , where  $N_5$  and  $N_0$  represent 5 h post-infection and initial free phage concentrations in PFU/mL, respectively. For each assay 18 mut was included as a control.

### **3.2.4 Statistical analysis**

The fitness of Q $\beta$  mutants and populations were compared by two-tailed  $t$  test and one-way analysis of variance (ANOVA). Two-way ANOVA was conducted to examined the effect of hosts and mutants or temperatures and mutants on fitness.

### 3.3 Results

With the aim of investigating the effect of Q $\beta$  mutants on fitness, the mutations which were detected at 43.7°C adaptation endpoint in all lines using the host strain 46L-1F', four kinds of mutant were prepared through site-directed mutagenesis of Q $\beta$  expression vector pACYCQ $\beta$ \_18 mut and the mutants were A1781C, U3784C, C3879G and 3(43.7) mut. The fitness assay was performed for four mutants at 43.7°C using the host strain 46L-1F' and compared the fitness values of 43.7°C-endpoint populations and 18 mut (Fig. 3.2a). The fitness of Q $\beta$  3(43.7) mut at 43.7°C showed comparatively higher than Q $\beta$  18 mut (one-way ANOVA,  $F_{(7,26)}=5.347$ ,  $P<0.001$ ; Tukey-HSD test  $P < 0.05$ ). The mutants U3784C and A1781C showed neutral or slightly deleterious on fitness (Fig. 3.2a). The single mutants cannot reaches the maximum fitness values separately but combined mutants (3(43.7) mut) showed maximum fitness at this temperature and the fitness of 3(43.7) mut and endpoint populations are comparable. This results clearly indicates sign epistasis in fitness.

Next, the fitness assay was performed for four mutants and 18 mut at 43.6°C using the host strain 46L-1F' and compared with the fitness values obtained at 43.7°C using the host strain 46L-1F' (Fig. 3.2b). A two-way ANOVA was conducted that examined the effect of temperature and Q $\beta$  mutants on fitness. Statistically significant difference was observed in mean fitness between 43.6°C and 43.7°C (two-way ANOVA,  $F_{(1,28)}=29.438$ ,  $P < 0.001$ ) and between Q $\beta$  mutants (two-way ANOVA,  $F_{(4,28)}=5.775$ ,  $P<0.01$ ), which indicates that the adaptation temperature 43.7°C was played as a selective pressure for Q $\beta$  18 mut.

Finally, the fitness assay was conducted for four mutants and 18 mut at 43.6°C using the host strain 43BF' and compared the values obtained using the host strain 46L-1F' at 43.6°C (Fig. 3.2c). A two-way ANOVA was conducted that examined the effect of host strain and Q $\beta$  mutants on fitness. No statistically significant difference was observed in mean fitness between host strain 43BF' and 46L-1F' (two-way ANOVA,  $F_{(1,10)}=1.50$ ,  $P = 0.25$ ), but showed statistically significant differences between Q $\beta$

mutants (two-way ANOVA,  $F_{(4,10)}=9.03$ ,  $P<0.01$ ), which indicates that the strain 46L-1F' was not behaved as a selective pressure and the mutations were not introduced due to host adaptation.

### 3.4 Discussion

In this Chapter, I quantitatively analyzed the fitness of Q $\beta$  mutants, which were obtained from thermal evolution experiment and showed that the adaptation temperature 43.7°C was behaved as a selective pressure for Q $\beta$  18 mut and all the three mutations played as a beneficial mutation but a sign epistasis was observed. The effect of A1781C and U3784C mutations on fitness were slightly deleterious or neutral ( $t = -1.25$ ,  $P = 0.225$ ) and neutral ( $t = -0.1$ ,  $P = 0.924$ ), respectively compared to 18 mut when these were introduced into the 18 mut genetic background. But these two mutations were introduced into the C3879G genetic background, these two mutations become positive.

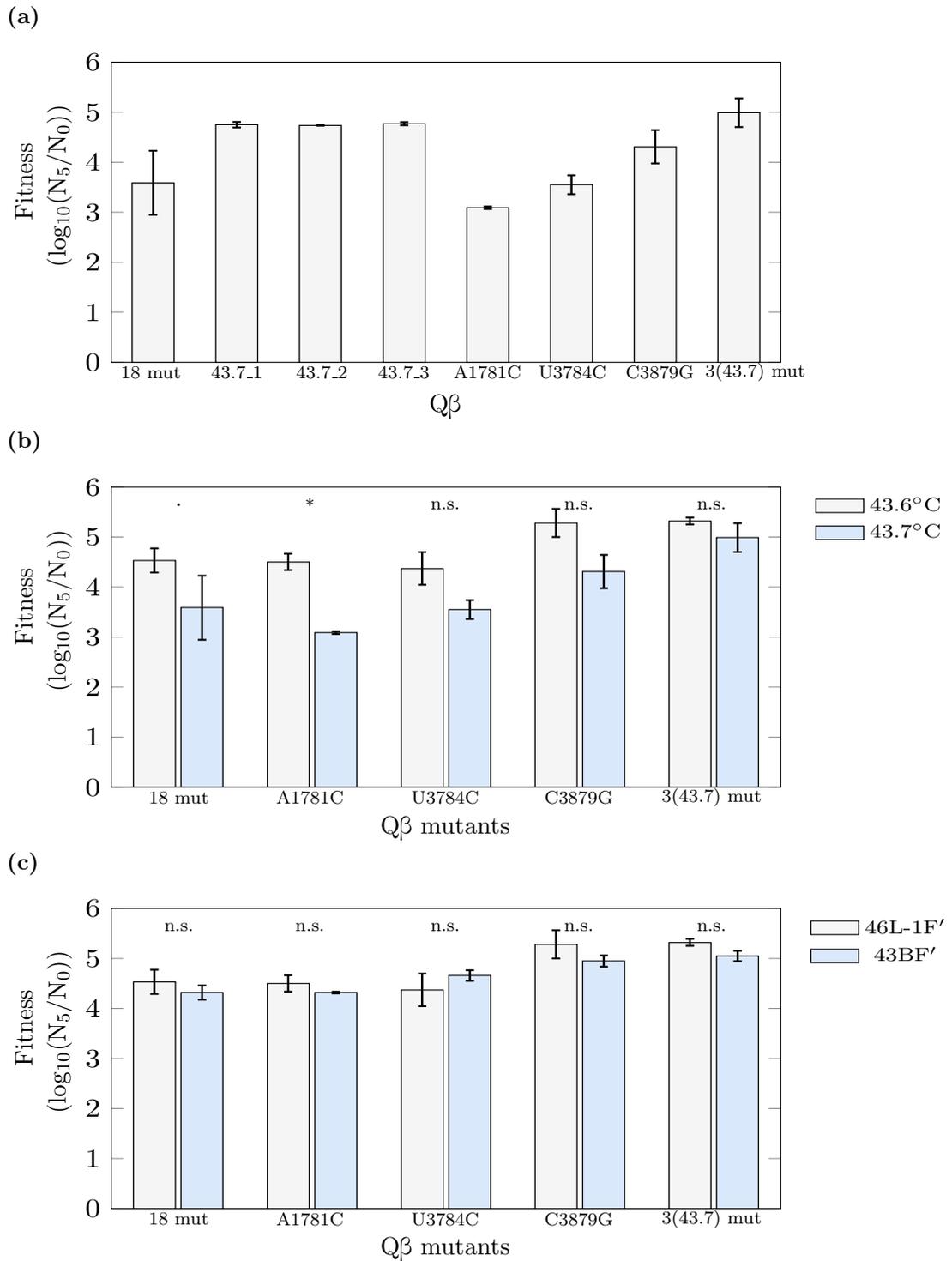
The present study identified amino acid changes that contributed to the phenotypic changes. Lys145Asn in the A1 showed a trend toward reduced fitness, Ile477Thr in the  $\beta$  subunit showed normal effect on fitness, and Leu509Val in the  $\beta$  subunit showed a trend toward improved fitness when compared with the 18 mut at 43.7°C (Fig. 3.2a). The A1 gene shares the same initial codon as the coat gene and produced as the read-through protein due to suppress at low probability of the coat stop codon. Thus, the N-terminus amino acid sequence of A1 is the same as that of coat, and the additional 197 amino acids region of the A1 in its C-terminus is located outside the capsid [60]. A1 is essential for producing infectious phage particles; however, its detailed function is unclear [61]. The position of Lys145Asn in A1 was missing in the crystal structure in Protein Data Bank 3RLC and the functional effects of this amino acid change is unclear. Q $\beta$  replicase comprises a  $\beta$  subunit from Q $\beta$ , translational elongation factors EF-Tu and EF-Ts, and ribosomal protein S1 from *E. coli* [22, 62]. Tyr510 interacts with the 5' terminus of newly synthesized RNA in the processive elongation stage,

and Tyr510 mutation decreases replication activity [64]. Since Leu509Val is the next amino acid of Tyr510, this amino acid mutation might be related to replication activity. Leu509Val was observed in another thermal adaptation experiment with the presence of the mutagenic nucleoside analogue 5-azacytidine and found beneficial effects on growth up to 43°C [87]. When I mapped on the structure (Protein Data Bank 4Q7J) Ile477 and Leu509 were close to the EF-Ts and EF-Tu subunits respectively (Fig. 3.1). The Ile477Thr was observed in line 2 of 43.6°C adapted population in previous thermal adaptation experiments [25]. The functional effects of this amino acid change are unclear.

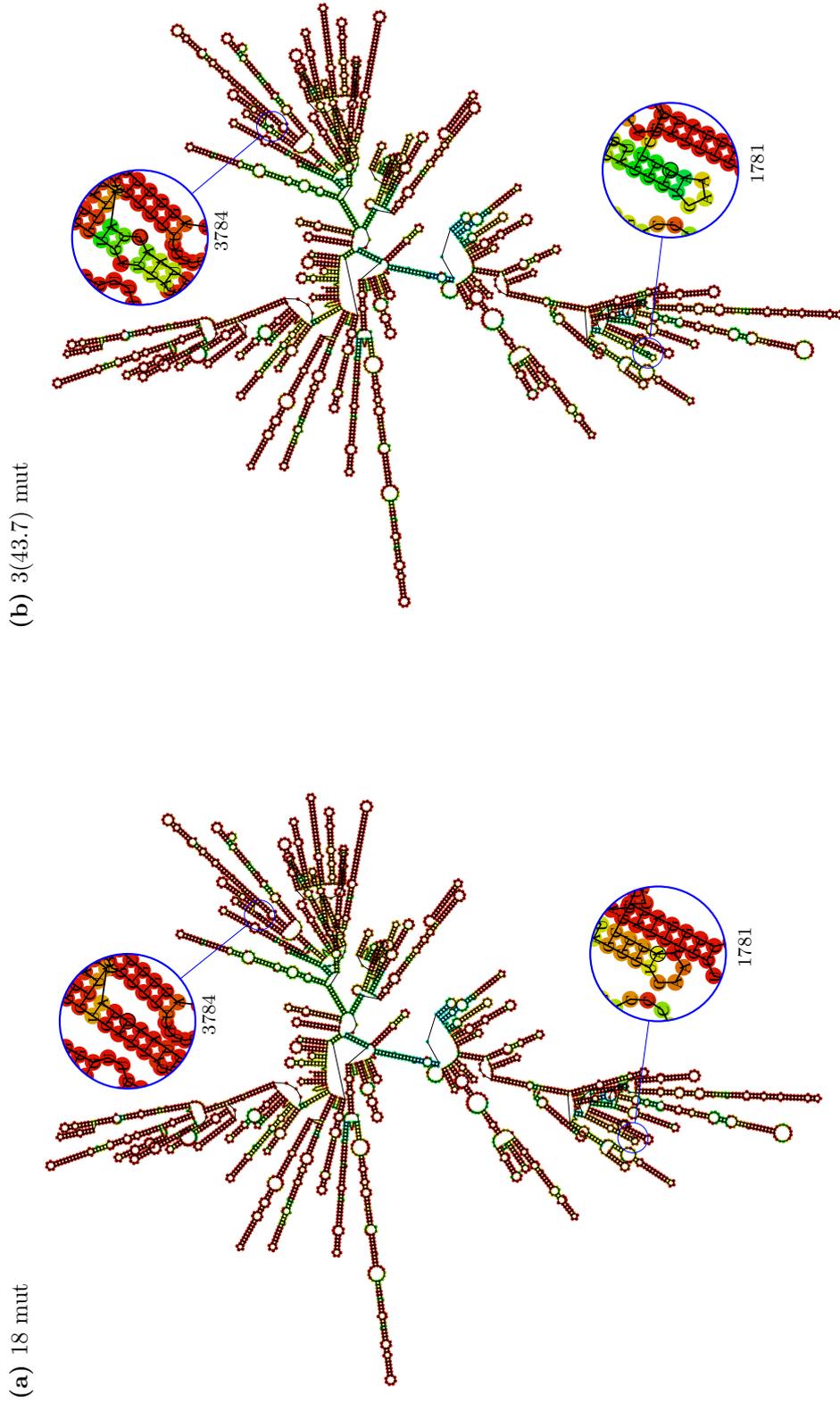
Sign epistasis interaction of mutations was observed in 43.7°C fitness assay. Sign epistasis in fitness analysis reflected the evolutionary path in this thermal adaptation evolution. The epistasis of mutations has been investigated extensively to elucidate the predictability and repeatability of evolutionary adaptation [88–96]. When the single mutations were introduced into 18 mut genetic background separately then A1781C showed slightly deleterious, U3784C showed neutral, and C3879G showed beneficial effect on fitness at 43.7°C. When those three mutations were introduced together into 18 mut genetic background then the mutant (3(43.7) mut) showed more beneficial effect on fitness at 43.7°C. Therefore, if A1781C mutation was introduced into the genome of 18 mut without U3784C and/or C3879G mutations during thermal adaptation, the mutant would not survive due to decrease in fitness. The secondary structure was predicted using RNAfold [46] with the whole-genome sequences of 18 mut and 3(43.7) mut and showed that A1781C and U3784C mutations alter the base-pair-making pattern compared to 18 mut (Fig. 3.3). These finding suggested that slightly deleterious and/or neutral mutations are also important for thermal adaptation.

The fitness of four mutants and 18 mut at 43.6°C using the host strain 43BF' and 46L-1F', and at 43.7°C using 46L-1F' as host strain showed no statistically significant difference in mean fitness between host strain 43BF' and 46L-1F' (Fig. 3.2c) but mean fitness at 43.7°C are comparatively lower than the mean fitness at 43.6°C (Fig. 3.2b).

These results suggested that the adaptation temperature 43.7°C was act as selective pressure for Q $\beta$  18 mut and the mutations were occurred due to thermal adaptation. In addition, C3879G mutation showed beneficial effect on fitness; interestingly, this mutations was observed in another thermal adaptation experiment with the presence of the mutagenic nucleoside analogue 5-azacytidine and showed beneficial effects on growth up to 43°C [87]. These results underscore my observation regarding the fixation of C3879G and other mutations due to thermal adaptation. Therefore, further studies regarding the importance of other mutations, which are not common in three lines and observed in other endpoint populations, in this experiments are required.



**Figure 3.2:** Fitness of Q $\beta$  mutants and populations. (a) Fitness of Q $\beta$  populations and mutants at 43.7°C using *E. coli* 46L-1F' as the host. (b) Fitness of Q $\beta$  mutants at 43.6°C and 43.7°C using *E. coli* 46L-1F' as the host. (c) Fitness of Q $\beta$  mutants at 43.6°C using *E. coli* 46L-1F' and 43BF' as the host. Data are expressed as average values  $\pm$  standard deviations (n=2 to 17). 2-sample *t*-test, \*  $P < 0.05$ , ·  $P < 0.1$ , and non-significant (n.s.)  $P > 0.1$  for pairs.



**Figure 3.3:** Secondary structure of Q $\beta$  18 mut and 3(43.7) mut. A1781C and U3784C mutations alter the base-pair-making pattern, but C3879G has no effect on changing the secondary structure.

## Chapter 4

# Conclusions

The overarching theme of the work presented in this thesis was “Studies on the adaptation of single-stranded RNA virus to novel environments using thermal adaptation experimental evolution of RNA bacteriophage Q $\beta$  as a model system”.

Chapter 2 focused on adaptation of ssRNA virus to novel environment using a model system. In this Chapter, I explored the patterns of thermal adaptation in a population of ssRNA bacteriophage Q $\beta$  by stepwise increase to the highest known growth temperature 45.3°C, a temperature at which the Q $\beta$  18 mut strain could not grow. I showed that Q $\beta$  can grow and replicate at this temperature within 52 days (616-638 generations) when starting with the Q $\beta$  18 mut variant and within 114 days (1238-1260 generations) when starting with ancestral Q $\beta$ , which has an optimum growth temperature of  $\sim$ 37°C, was used as the starting material. I also showed that thermally adapted Q $\beta$  populations could grow and replicate at 45.3°C and these populations could grow with equivalent fitness at 37.2°C, even after returned the 45.3°C-adapted populations to 37.2°C for 122–124 generations. These results clearly indicate that Q $\beta$  gained the potential for growth at higher temperatures without showing trade-off in the lower ranges. The genotypic analysis showed that Q $\beta$  adapted in novel environment with accumulation of point mutations. In total 34–39 mutations were identified in 45.3°C-adapted populations, these mutations accounts for 0.8% – 0.9% of the total RNA genome. In addition, mutations introduced during this adaptation have a tendency to occur frequently in UTR and A1 but not randomly in all the genes.

Chapter 3 presented the importance of Q $\beta$  mutants obtained in thermal adaptations experiment. In this Chapter, my effort was to investigate the growth characteristics of Q $\beta$  mutants and to clarify about the selective pressure (temperature or host), which was responsible for occurring those mutations, for Q $\beta$  18 mut at 43.7°C adaptation using 46L-1F' as a host strain. I showed that the temperature 43.7°C played as a selective pressure for Q $\beta$  18 mut, the mutations were introduced due to thermal adaptation, and the neutral and/or slightly deleterious mutations are important as well as beneficial mutation in thermal adaptation.

Collectively, the body of work detailed in this thesis presents many interesting and exciting opportunities for further exploration. The evolution of virus diseases, both their emergence and disappearance, involves complex interaction between the agent, the host, and the environment [97, 98]. The global warming is a great problem in modern time in the world. The new environment for living organism are creating due to global warming. The living organism adapted and evolved in these new environments by changing their genetic information and phenotypes and produces new type or mutants type organism. Our present evolution experiment focus on temperature on ssRNA bacteriophage Q $\beta$ , displaying the adaptation mechanism that can be readily explored for applications in tracking the emergence of new type virus. Recently, COVID-19 was spread all over the world within short time from China through infected human, which indicates that this virus adapted to new climatic environment in different countries quickly. The reduction or disappearance of virus diseases usually involves human intervention, as exemplified by immunization or an eradication process for diseases. Changes in lifestyle can control the rapid spread of virus diseases. The virus disease can fadeout after many replication generations for changing the genotypes due to high mutation rate and for naturally occurring immunity of the host. Future studies might include DNA as well as RNA bacteriophages are required for understanding of adaptation in more complex systems.

# Summary

The dissertation related to evolution experiment has been presented in two parts. The larger part of this thesis is devoted for investigating adaptation of single-stranded (ss)RNA virus to novel environments using thermal adaptation of ssRNA bacteriophage Q $\beta$  as a model system. The second and smaller part of this thesis is studied for investigating the growth characteristics of Q $\beta$  mutants.

ssRNA viruses change at high mutations rates to maintain the integrity of genetic information, which allows them to find the beneficial mutations needed for adaptation quickly. In spite of the fact that it has been considered that ssRNA virus can adapt readily to changes in the environment, it remains unclear how quickly they can adapt to a novel environment and/or how many and what types of mutation are required to facilitate evolution. To elucidate the mechanisms underlying ssRNA virus adaptations to the novel environment, I conducted thermal adaptation evolution experiments of ssRNA bacteriophage Q $\beta$  as a model system. The ssRNA bacteriophage Q $\beta$ , of the family *Leviviridae*, which infects *Escherichia coli* strains expressing the F pili that acts as the virus receptor. The Q $\beta$  has short generation time and a genome of 4217 nucleotides in length that encodes only four proteins: the A2 protein for bacterial lysis and entry, the coat protein, the A1 protein, which is expressed through incorrect reading of the stop codon of the coat protein and present in low amounts in the capsid, and the  $\beta$  subunit for Q $\beta$  replicase. These make us understand the relationship between genotypic and phenotypic changes easily. In addition, evolution experiments with thermal selection of pressure is suitable for analyzing the process of adaptation because vital living processes, such as energy transduction, reproduction, and growth,

are affected by temperature change and the temperature is precisely controllable in the laboratory. Previously, a thermal adaptation experiment using Q $\beta$  as a model of ssRNA virus was conducted in our laboratory, in which the culture temperature was increased in a stepwise manner from 37.2°C to 43.6°C in three independent lines using *E. coli* 43BF' as the host strain that could grow in temperatures up to 43.6°C. The effect of synonymous and nonsynonymous changes on the fitness and life history of Q $\beta$  were evaluated. To investigate the thermal adaptation of phages at higher temperatures, it is necessary to use the host strain that is capable of growth at higher temperatures. Recently, Kishimoto et al. in Toho University, Japan isolated a strain of thermally adapted *E. coli* 46L-1 that was capable of growing at temperatures up to 46°C by thermal adaptation evolution experiment. This technical advance made it possible to evaluate Q $\beta$  adaptation to thermal changes at higher temperatures.

To monitor thermal adaptation of Q $\beta$ , I used the thermally adapted *E. coli* strain 46L-1F' as the host, which was constructed via conjugation with strain 46L-1 and HB2151, and Q $\beta$  18 mut as the starting phage, which was one of three replicates that had adapted to 43.6°C and was prepared from cDNA. Because the growth of Q $\beta$  depends on the growth of the host *E. coli* strain and the specific growth rate of *E. coli* 46L-1F' strain was almost identical between 37.2°C and 44.8°C but decreased by 7% and 26% at 45.3°C and 45.9°C, respectively. Therefore, I conducted the present thermal adaptation experiments at temperatures up to 45.3°C using the *E. coli* 46L-1F' strain as the host. After adaptation at 45.3°C, the culture temperature was reverted back to 37.2°C to investigate whether the ancestral sequence became dominant in the population. In this study, I explored the adaptation process of ssRNA bacteriophage Q $\beta$  via stepwise increases to the highest known growth temperature, 45.3°C and showed that Q $\beta$  can grow and replicate at this temperature within 52 days (616-638 generations) when the Q $\beta$  18 mut variant was used as the starting material and within 114 days (1238-1260 generations) when Q $\beta$  ancestral was used as the starting material. Fitness analysis revealed that Q $\beta$  adapted to grow and replicate at 45.3°C had an overall increased temperature range because these populations could grow with

equivalent fitness at 37.2°C. Intriguingly, the reverse-adapted Q $\beta$  populations showed little to no decrease in fitness after adaptation of the 45.3°C adapted populations to 37.2°C for 8 days (122-124 generations). These results clearly demonstrate that Q $\beta$  acquired the potential for growth at higher temperatures without showing trade-off in the lower ranges. The 45.3°C-adapted population had at most 21 mutations from Q $\beta$  18 mut and 39 mutations from Q $\beta$  ancestral. In addition, the mutation introduced during this adaptation tend to increase the frequency in UTR and A1 but not randomly in all the genes. These results indicate that Q $\beta$  could adapt to these elevated temperatures with only point mutations and these mutations account for 0.8% - 0.9% of the total RNA genome.

Finally, in order to investigate the improved growth characteristics of 43.7°C-adapted endpoint populations, four kinds of mutant (18 mut-A1781C, -U3784C, -C3879G, and -combined of three) were prepared through site-directed mutagenesis of the Q $\beta$  expression vector pACYCQ $\beta$ \_18 mut. The fitness assay was performed for four mutants and 18 mut at 43.6°C using the host strains 43BF' and 46L-1F' and at 43.7°C using the host strain 46L-1F' and compared to determine the selective pressure of 43.7°C-adaptation using 46L-1F' and Q $\beta$  18 mut. The fitness of four mutants and 18 mut at 43.6°C using the host strain 43BF' and 46L-1F' showed no statistically significant difference in mean fitness but the mean fitness at 43.7°C is comparatively lower than that of at 43.6°C. The slightly deleterious and/or neutral mutants (A1781C and U3784C) and beneficial mutant (C3879G) cannot reach the maximum fitness value separately but combined mutants showed maximum fitness at 43.7°C. These results suggest that the adaptation temperature 43.7°C acts as a selective pressure for Q $\beta$  18 mut, the mutations were introduced due to thermal adaptation, and the neutral and/or slightly deleterious mutations are important as well as the beneficial mutations for thermal adaptation.

The evolution experiment results of ssRNA bacteriophage Q $\beta$  performed with thermally adapted *E. coli* underscore the observation regarding the rapid adaptation of

ssRNA virus to novel environments. This evolution experiments focus on temperature on ssRNA bacteriophage Q $\beta$ , displaying the adaptation mechanism that can be readily explored for applications in tracking the emergence of new type of virus. Studies such as these are not only important for the fields related to evolutionary biology, but also they will improve our understanding of adaptation in more complex systems.

# Acknowledgements

First and foremost, I would like to acknowledge my advisor, Prof. Akiko Kashiwagi, for her active involvement in all the studies described herein and for her patience, mentoring, and general good nature throughout these last 3 years. I would like to thank all those in the lab who not only made this work possible but also made my time spent in the lab very enjoyable, especially Toma Yokono, Takahiro Iwamoto, and Takuya Kuramoto. I also thank the authorities of the Noakhali Science and Technology University, Bangladesh for granting me study leave to pursue doctoral studies in Japan. Furthermore, I would like to thank Prof. Toshihiko Kishimoto (Toho University, Chiba, Japan) for providing the *E. coli* strain. Finally, I would like to thank my associate advisors, Profs. Teruo Sano (Hirosaki University, Aomori, Japan) and Miwa Yamada (Iwate University, Iwate, Japan) for their guidance and critical consideration of my research.































# Bibliography

- [1] Bennett, A. F., Lenski, R. E., and Mittler, J. E., 1992. Evolutionary adaptation to temperature. i. fitness responses of *Escherichia coli* to changes in its thermal environment. *Evolution*, **46**(1):16–30. doi:10.1111/j.1558-5646.1992.tb01981.x.
- [2] Bennett, A. F. and Lenski, R. E., 1993. Evolutionary adaptation to temperature II. Thermal niches of experimental lines of *Escherichia coli*. *Evolution*, **47**(1):1–12. doi:10.1111/j.1558-5646.1993.tb01194.x.
- [3] Kolss, M., Vijendravarma, R. K., Schwaller, G., and Kawecki, T. J., 2009. Life-history consequences of adaptation to larval nutritional stress in *Drosophila*. *Evolution*, **63**(9):2389–2401. doi:10.1111/j.1558-5646.2009.00718.x.
- [4] terHorst, C. P., 2011. Experimental evolution of protozoan traits in response to interspecific competition. *Journal of Evolutionary Biology*, **24**(1):36–46. doi:10.1111/j.1420-9101.2010.02140.x.
- [5] Zbinden, M., Haag, C. R., and Ebert, D., 2008. Experimental evolution of field populations of *Daphnia magna* in response to parasite treatment. *Journal of Evolutionary Biology*, **21**(4):1068–1078. doi:10.1111/j.1420-9101.2008.01541.x.
- [6] Dhar, R., Sägesser, R., Weikert, C., Yuan, J., and Wagner, A., 2011. Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. *Journal of Evolutionary Biology*, **24**(5):1135–1153. doi:10.1111/j.1420-9101.2011.02249.x.
- [7] Clark, D. P., Pazdernik, N. J., and McGehee, M. R., 2019. Viruses, viroids, and prions. In *Molecular Biology*, chapter 24, pages 749–792. Academic Cell, third edition. doi:10.1016/B978-0-12-813288-3.00024-0.

- 
- [8] Domingo, E., 2000. Viruses at the edge of adaptation. *Virology*, **270**(2):251–253. doi:10.1006/viro.2000.0320.
- [9] Domingo, E. and Holland, J. J., 1997. RNA virus mutations and fitness for survival. *Annual Review of Microbiology*, **51**(1):151–178. doi:10.1146/annurev.micro.51.1.151.
- [10] Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol, S., 1982. Rapid evolution of RNA genomes. *Science*, **215**(4540):1577–1585. doi:10.1126/science.7041255.
- [11] Novella, I. S., Elena, S. F., Moya, A., Domingo, E., and Holland, J. J., 1995. Size of genetic bottlenecks leading to virus fitness loss is determined by mean initial population fitness. *Journal of Virology*, **69**(5):2869–2872.
- [12] Domingo, E. and Perales, C., 2019. Viral quasispecies. *PLoS Genetics*, **15**(10):e1008271.
- [13] Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M., and Belshaw, R., 2010. Viral mutation rates. *Journal of Virology*, **84**(19):9733–9748. doi:10.1128/JVI.00694-10.
- [14] Sanjuán, R. and Domingo-Calap, P., 2016. Mechanisms of viral mutation. *Cellular and Molecular Life Sciences*, **73**(23):4433–4448. doi:10.1007/s00018-016-2299-6.
- [15] Peck, K. M. and Lauring, A. S., 2018. Complexities of viral mutation rates. *Journal of Virology*, **92**(14). doi:10.1128/JVI.01031-17.
- [16] Sanjuán, R., 2012. From molecular genetics to phylodynamics: Evolutionary relevance of mutation rates across viruses. *PLoS Pathogens*, **8**(5):1–5. doi:10.1371/journal.ppat.1002685.
- [17] Knies, J. L., Izem, R., Supler, K. L., Kingsolver, J. G., and Burch, C. L., 2006.

- The genetic basis of thermal reaction norm evolution in lab and natural phage populations. *PLoS Biology*, **4**(7). doi:10.1371/Journal.pbio.0040201.
- [18] Van Duin, J. and Tsareva, N., 2006. Single-stranded RNA phages. In Calendar, R., editor, *The Bacteriophages*, chapter 15, pages 175–196. Oxford University Press, New York, USA, 2 edition.
- [19] Weissmann, C., Billeter, M. A., Goodman, H. M., Hindley, J., and Weber, H., 1973. Structure and function of phage RNA. *Annual Review of Biochemistry*, **42**(1):303–328. doi:10.1146/annurev.bi.42.070173.001511.
- [20] Weissmann, C., Feix, G., Slor, H., and Pollet, R., 1967. Replication of viral RNA. XIV. Single-stranded minus strands as template for the synthesis of viral plus strands in vitro. *Proceedings of the National Academy of Sciences*, **57**(6):1870–1877. doi:10.1073/pnas.57.6.1870.
- [21] Wahba, A. J., Miller, M. J., Niveleau, A., Landers, T. A., Carmichael, G. G., Weber, K., Hawley, D. A., and Slobin, L. I., 1974. Subunit I of Q $\beta$  replicase and 30 S ribosomal protein S1 of *Escherichia coli*: Evidence for the identity of the two proteins. *Journal of Biological Chemistry*, **249**(10):3314–3316.
- [22] Blumenthal, T., Landers, T. A., and Weber, K., 1972. Bacteriophage Q $\beta$  replicase contains the protein biosynthesis elongation factors EF Tu and EF Ts. *Proceedings of the National Academy of Sciences*, **69**(5):1313–1317. doi:10.1073/pnas.69.5.1313.
- [23] Drake, J. W., 1993. Rates of spontaneous mutation among RNA viruses. *Proceedings of the National Academy of Sciences*, **90**(9):4171–4175. doi:10.1073/pnas.90.9.4171.
- [24] Bradwell, K., Combe, M., Domingo-Calap, P., and Sanjuán, R., 2013. Correlation between mutation rate and genome size in riboviruses: Mutation rate of bacteriophage Q $\beta$ . *Genetics*, **195**(1):243–251. doi:10.1534/genetics.113.154963.

- 
- [25] Kashiwagi, A., Sugawara, R., Tsushima, F. S., Kumagaia, T., and Yomo, T., 2014. Contribution of silent mutations to thermal adaptation of RNA bacteriophage Q $\beta$ . *Journal of Virology*, **88**(19):11459–11468. doi:10.1128/JVI.01127-14.
- [26] Kashiwagi, A., Kadoya, T., Kumasaka, N., Kumagai, T., Tsushima, F. S., and Yomo, T., 2018. Influence of adaptive mutations, from thermal adaptation experiments, on the infection cycle of RNA bacteriophage Q $\beta$ . *Archives of Virology*, **163**:2655–2662. doi:10.1007/s00705-018-3895-6.
- [27] Kishimoto, T., Ying, B.-W., Tsuru, S., Iijima, L., Suzuki, S., Hashimoto, T., Oyake, A., Kobayashi, H., Someya, Y., Narisawa, D., and Yomo, T., 2015. Molecular clock of neutral mutations in a fitness-increasing evolutionary process. *PLoS Genetics*, **11**(7):1–18. doi:10.1371/Journal.pgen.1005392.
- [28] Abedon, S. T., editor, 2008. *Bacteriophage Ecology: Population Growth, Evolution, and Impact of Bacterial Viruses*. Advances in Molecular and Cellular Microbiology. Cambridge University Press. doi:10.1017/CBO9780511541483.
- [29] Abedon, S. T., 2006. Phage ecology. In Calendar, R., editor, *The Bacteriophages*, chapter 5, pages 37–46. Oxford University Press, New York, USA, 2 edition.
- [30] Clokie, M. R., Millard, A. D., Letarov, A. V., and Heaphy, S., 2011. Phages in nature. *Bacteriophage*, **1**(1):31–45. doi:10.4161/bact.1.1.14942.
- [31] Bull, J. J., Badgett, M. R., Wichman, H. A., Huelsenbeck, J. P., Hillis, D. M., Gulati, A., Ho, C., and Molineux, I. J., 1997. Exceptional convergent evolution in a virus. *Genetics*, **147**(4):1497–1507.
- [32] Holder, K. K. and Bull, J. J., 2001. Profiles of adaptation in two similar viruses. *Genetics*, **159**(4):1393–1404.
- [33] Bollback, J. P. and Huelsenbeck, J. P., 2009. Parallel genetic evolution within and between bacteriophage species of varying degrees of divergence. *Genetics*, **181**(1):225–234. doi:10.1534/genetics.107.085225.

- 
- [34] Singhal, S., Leon Guerrero, C. M., Whang, S. G., McClure, E. M., Busch, H. G., and Kerr, B., 2017. Adaptations of an RNA virus to increasing thermal stress. *PLoS ONE*, **12**(12):1–26. doi:10.1371/Journal.pone.0189602.
- [35] Lázaro, E., Arribas, M., Cabanillas, L., Román, I., and Acosta, E., 2018. Evolutionary adaptation of an RNA bacteriophage to the simultaneous increase in the within-host and extracellular temperatures. *Scientific Reports*, **8**(1):8080. doi:10.1038/s41598-018-26443-z.
- [36] Somovilla, P., Manrubia, S., and Lázaro, E., 2019. Evolutionary dynamics in the RNA bacteriophage Q $\beta$  depends on the pattern of change in selective pressures. *Pathogens*, **8**(2). doi:10.3390/pathogens8020080.
- [37] Carter, P., Bedouelle, H., and Winter, G., 1985. Improved oligonucleotide site-directed mutagenesis using M13 vectors. *Nucleic Acids Research*, **13**(12):4431–4443. doi:10.1093/nar/13.12.4431.
- [38] Watanabe, I., Toshizo, S., Kohsuke, F., and Asako, A., 1979. “Pseudolysogenization” by RNA phage Q $\beta$ . *Microbiology and Immunology*, **23**(11):1077–1083. doi:10.1111/j.1348-0421.1979.tb00539.x.
- [39] Kashiwagi, A., Sakurai, T., Tsuru, S., Ying, B.-W., Mori, K., and Yomo, T., 2009. Construction of *Escherichia coli* gene expression level perturbation collection. *Metabolic Engineering*, **11**(1):56–63. doi:10.1016/j.ymben.2008.08.002.
- [40] Inomata, T., Kimura, H., Hayasaka, H., Shiozaki, A., Fujita, Y., and Kashiwagi, A., 2012. Quantitative comparison of the RNA bacteriophage Q $\beta$  infection cycle in rich and minimal media. *Archives of Virology*, **157**(11):2163–2169. doi:10.1007/s00705-012-1419-3.
- [41] Tsukada, K., Okazaki, M., Kita, H., Inokuchi, Y., Urabe, I., and Yomo, T., 2009. Quantitative analysis of the bacteriophage Q $\beta$  infection cycle. *Biochimica et Biophysica Acta (BBA)-General Subjects*, **1790**(1):65–70. doi:10.1016/j.bbagen.2008.08.007.

- 
- [42] Carlson, K., 2005. Working with bacteriophages: common techniques and methodological approaches. In Kutter, E. and Sulakvelidze, A., editors, *Bacteriophages biology and applications*, chapter Appendix, pages 437–494. CRC Press, Boca Raton, FL.
- [43] Kropinski, A. M., Mazzocco, A., Waddell, T. E., Lingohr, E., and Johnson, R. P., 2009. Enumeration of bacteriophages by double agar overlay plaque assay. In Clokie, M. R. and Kropinski, A. M., editors, *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*, pages 69–76. Humana Press, Totowa, NJ. doi:10.1007/978-1-60327-164-6\_7.
- [44] Kashiwagi, A. and Yomo, T., 2011. Ongoing phenotypic and genomic changes in experimental coevolution of RNA bacteriophage Q $\beta$  and *Escherichia coli*. *PLoS Genetics*, **7**(8):1–13. doi:10.1371/Journal.pgen.1002188.
- [45] Sanger, F., Nicklen, S., and Coulson, A. R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, **74**(12):5463–5467. doi:10.1073/pnas.74.12.5463.
- [46] Lorenz, R., Bernhart, S. H., Höner zu Siederdisen, C., Tafer, H., Flamm, C., Stadler, P. F., and Hofacker, I. L., 2011. Viennarna package 2.0. *Algorithms for Molecular Biology*, **6**(1):26. doi:10.1186/1748-7188-6-26.
- [47] RStudio Team, 2015. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.
- [48] Nabergoj, D., Modic, P., and Podgornik, A., 2018. Effect of bacterial growth rate on bacteriophage population growth rate. *MicrobiologyOpen*, **7**(2):e00558. doi:10.1002/mbo3.558.
- [49] Kishimoto, T., Iijima, L., Tatsumi, M., Ono, N., Oyake, A., Hashimoto, T., Matsuo, M., Okubo, M., Suzuki, S., Mori, K., Kashiwagi, A., Furusawa, C., Ying, B.-W., and Yomo, T., 2010. Transition from positive to neutral in mutation

- fixation along with continuing rising fitness in thermal adaptive evolution. *PLoS Genetics*, **6**(10):1–10. doi:10.1371/Journal.pgen.1001164.
- [50] Farewell, A. and Neidhardt, F. C., 1998. Effect of temperature on in vivo protein synthetic capacity in *Escherichia coli*. *Journal of bacteriology*, **180**(17):4704–4710.
- [51] Young, R. and Wang, I.-N., 2006. Phage lysis. In *The Bacteriophages*, chapter 10, pages 104–125. Oxford University Press, New York, USA, 2 edition.
- [52] Bernhardt, T. G., Wang, I.-N., Struck, D. K., and Young, R., 2001. A protein antibiotic in the phage Q $\beta$  virion: Diversity in lysis targets. *Science*, **292**(5525):2326–2329. doi:10.1126/science.1058289.
- [53] Bernhardt, T. G., Wang, I.-N., Struck, D. K., and Young, R., 2002. Breaking free: “Protein antibiotics” and phage lysis. *Research in Microbiology*, **153**(8):493–501. doi:10.1016/S0923-2508(02)01330-X.
- [54] Karnik, S. and Billeter, M., 1983. The lysis function of RNA bacteriophage Q $\beta$  is mediated by the maturation (A2) protein. *The EMBO Journal*, **2**(9):1521–1526.
- [55] Gorzelnik, K. V., Cui, Z., Reed, C. A., Jakana, J., Young, R., and Zhang, J., 2016. Asymmetric cryo-EM structure of the canonical allovivivirus Q $\beta$  reveals a single maturation protein and the genomic ssRNA in situ. *Proceedings of the National Academy of Sciences*, **113**(41):11519–11524. doi:10.1073/pnas.1609482113.
- [56] Cui, Z., Gorzelnik, K. V., Chang, J.-Y., Langlais, C., Jakana, J., Young, R., and Zhang, J., 2017. Structures of Q $\beta$  virions, virus-like particles, and the Q $\beta$ –MurA complex reveal internal coat proteins and the mechanism of host lysis. *Proceedings of the National Academy of Sciences*, **114**(44):11697–11702. doi:10.1073/pnas.1707102114.
- [57] Dai, X., Li, Z., Lai, M., Shu, S., Du, Y., Zhou, Z. H., and Sun, R., 2017. *In situ* structures of the genome and genome-delivery apparatus in a single-stranded RNA virus. *Nature*, **541**:112.

- 
- [58] Rumnieks, J. and Tars, K., 2017. Crystal structure of the maturation protein from bacteriophage Q $\beta$ . *Journal of Molecular Biology*, **429**(5):688–696. doi:10.1016/j.jmb.2017.01.012.
- [59] Rumnieks, J. and Tars, K., 2014. Crystal structure of the bacteriophage Q $\beta$  coat protein in complex with the RNA operator of the replicase gene. *Journal of Molecular Biology*, **426**(5):1039–1049. doi:10.1016/j.jmb.2013.08.025.
- [60] Rumnieks, J. and Tars, K., 2011. Crystal structure of the read-through domain from bacteriophage Q $\beta$  A1 protein. *Protein Science*, **20**(10):1707–1712. doi:10.1002/pro.704.
- [61] Hofstetter, H., Monstein, H.-J., and Weissmann, C., 1974. The readthrough protein A1 is essential for the formation of viable Q $\beta$  particles. *Biochimica et Biophysica Acta (BBA)-Nucleic Acids and Protein Synthesis*, **374**(2):238–251. doi:10.1016/0005-2787(74)90366-9.
- [62] Blumenthal, T. and Carmichael, G. G., 1979. RNA replication: Function and structure of Q $\beta$ -replicase. *Annual Review of Biochemistry*, **48**(1):525–548. doi:10.1146/annurev.bi.48.070179.002521.
- [63] Takeshita, D. and Tomita, K., 2010. Assembly of Q $\beta$  viral RNA polymerase with host translational elongation factors EF-Tu and -Ts. *Proceedings of the National Academy of Sciences*, **107**(36):15733–15738. doi:10.1073/pnas.1006559107.
- [64] Takeshita, D. and Tomita, K., 2012. Molecular basis for RNA polymerization by Q $\beta$  replicase. *Nature Structural & Molecular Biology*, **19**(2):229–237. doi:10.1038/nsmb.2204.
- [65] Takeshita, D., Yamashita, S., and Tomita, K., 2014. Molecular insights into replication initiation by Q $\beta$  replicase using ribosomal protein S1. *Nucleic Acids Research*, **42**(16):10809–10822. doi:10.1093/nar/gku745.
- [66] Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics*, **1**:356–366.

- 
- [67] Maynard Smith, J., 1970. Natural selection and the concept of a protein space. *Nature*, **225**(5232):563–564. doi:10.1038/225563a0.
- [68] de Visser, J. A. G. M. and Krug, J., 2014. Empirical fitness landscapes and the predictability of evolution. *Nature Reviews Genetics*, **15**(7):480–490. doi:10.1038/nrg3744.
- [69] Kauffman, S. A. and Weinberger, E. D., 1989. The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology*, **141**(2):211 – 245. doi:10.1016/S0022-5193(89)80019-0.
- [70] Neidhart, J., Szendro, I. G., and Krug, J., 2014. Adaptation in tunably rugged fitness landscapes: The rough Mount Fuji model. *Genetics*, **198**(2):699–721. doi:10.1534/genetics.114.167668.
- [71] Fragata, I., Blanckaert, A., Louro, M. A. D., Liberles, D. A., and Bank, C., 2019. Evolution in the light of fitness landscape theory. *Trends in Ecology & Evolution*, **34**(1):69 – 82. doi:10.1016/j.tree.2018.10.009.
- [72] Hayashi, Y., Aita, T., Toyota, H., Husimi, Y., Urabe, I., and Yomo, T., 2006. Experimental rugged fitness landscape in protein sequence space. *PLoS ONE*, **1**(1):1–8. doi:10.1371/journal.pone.0000096.
- [73] Athavale, S. S., Spicer, B., and Chen, I. A., 2014. Experimental fitness landscapes to understand the molecular evolution of RNA-based life. *Current Opinion in Chemical Biology*, **22**:35–39. doi:10.1016/j.cbpa.2014.09.008.
- [74] Blanco, C., Janzen, E., Pressman, A., Saha, R., and Chen, I. A., 2019. Molecular fitness landscapes from high-coverage sequence profiling. *Annual Review of Biophysics*, **48**(1):1–18. doi:10.1146/annurev-biophys-052118-115333.
- [75] Pitt, J. N. and Ferré-D’Amaré, A. R., 2010. Rapid construction of empirical RNA fitness landscapes. *Science*, **330**(6002):376. doi:10.1126/science.1192001.

- 
- [76] Held, D. M., Travis Greathouse, S., Agrawal, A., and Burke, D. H., 2003. Evolutionary landscapes for the acquisition of new ligand recognition by rna aptamers. *Journal of Molecular Evolution*, **57**(3):299–308. doi:10.1007/s00239-003-2481-y.
- [77] Jiménez, J. I., Xulvi-Brunet, R., Campbell, G. W., Turk-MacLeod, R., and Chen, I. A., 2013. Comprehensive experimental fitness landscape and evolutionary network for small rna. *Proceedings of the National Academy of Sciences*, **110**(37):14984–14989. doi:10.1073/pnas.1307604110.
- [78] Koning, R. I., Gomez-Blanco, J., Akopjana, I., Vargas, J., Kazaks, A., Tars, K., Carazo, J., and Koster, A. J., 2016. Asymmetric cryo-EM reconstruction of phage MS2 reveals genome structure *in situ*. *Nature Communications*, **7**:12524.
- [79] Hossain, M. T., Yokono, T., and Kashiwagi, A., 2020. The single-stranded RNA bacteriophage Q $\beta$  adapts rapidly to high temperatures: An evolution experiment. *Viruses*, **12**(6):638. doi:10.3390/v12060638.
- [80] Sung, W., Ackerman, M. S., Miller, S. F., Doak, T. G., and Lynch, M., 2012. Drift-barrier hypothesis and mutation-rate evolution. *Proceedings of the National Academy of Sciences*, **109**(45):18488–18492. doi:10.1073/pnas.1216223109.
- [81] Wielgoss, S., Barrick, J. E., Tenaillon, O., Wisner, M. J., Dittmar, W. J., Cruveiller, S., Chane-Woon-Ming, B., Médigue, C., Lenski, R. E., and Schneider, D., 2013. Mutation rate dynamics in a bacterial population reflect tension between adaptation and genetic load. *Proceedings of the National Academy of Sciences*, **110**(1):222–227. doi:10.1073/pnas.1219574110.
- [82] Shibai, A., Takahashi, Y., Ishizawa, Y., Motooka, D., Nakamura, S., Ying, B.-W., and Tsuru, S., 2017. Mutation accumulation under UV radiation in *Escherichia coli*. *Scientific Reports*, **7**(1):14531. doi:10.1038/s41598-017-15008-1.
- [83] Eyre-Walker, A. and Keightley, P. D., 2007. The distribution of fitness effects of new mutations. *Nature Reviews Genetics*, **8**(8):610–618. doi:10.1038/nrg2146.

- 
- [84] Morley, V. J. and Turner, P. E., 2017. Dynamics of molecular evolution in RNA virus populations depend on sudden versus gradual environmental change. *Evolution*, **71**(4):872–883. doi:10.1111/evo.13193.
- [85] Cox, J., Schubert, A. M., Travisano, M., and Putonti, C., 2010. Adaptive evolution and inherent tolerance to extreme thermal environments. *BMC Evolutionary Biology*, **10**(1):75. doi:10.1186/1471-2148-10-75.
- [86] Sambrook, J., Fritsh, E. F., and Maniatis, T., 1989. *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.
- [87] Arribas, M., Cabanillas, L., Kubota, K., and Lázaro, E., 2016. Impact of increased mutagenesis on adaptation to high temperature in bacteriophage Q $\beta$ . *Virology*, **497**:163–170. doi:10.1016/j.virol.2016.07.007.
- [88] Chou, H.-H., Chiu, H.-C., Delaney, N. F., Segrè, D., and Marx, C. J., 2011. Diminishing returns epistasis among beneficial mutations decelerates adaptation. *Science*, **332**(6034):1190–1192. doi:10.1126/science.1203799.
- [89] Kvittek, D. J. and Sherlock, G., 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genetics*, **7**(4):1–11. doi:10.1371/Journal.pgen.1002056.
- [90] Schrag, S. J., Perrot, V., and Levin, B. R., 1997. Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. *Proceedings of the Royal Society of London B: Biological Sciences*, **264**(1386):1287–1291. doi:10.1098/rspb.1997.0178.
- [91] Mammano, F., Trouplin, V., Zennou, V., and Clavel, F., 2000. Retracing the evolutionary pathways of human immunodeficiency virus type 1 resistance to protease inhibitors: Virus fitness in the absence and in the presence of drug. *Journal of Virology*, **74**(18):8524–8531. doi:10.1128/JVI.74.18.8524-8531.2000.

- 
- [92] Khan, A. I., Dinh, D. M., Schneider, D., Lenski, R. E., and Cooper, T. F., 2011. Negative epistasis between beneficial mutations in an evolving bacterial population. *Science*, **332**(6034):1193–1196. doi:10.1126/science.1203801.
- [93] Maharjan, R. P. and Thomas, F., 2013. Epistatic interactions determine the mutational pathways and coexistence of lineages in clonal *Escherichia coli* populations. *Evolution*, **67**(9):2762–2768. doi:10.1111/evo.12137.
- [94] Weinreich, D. M., Delaney, N. F., DePristo, M. A., and Hartl, D. L., 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, **312**(5770):111–114. doi:10.1126/science.1123539.
- [95] Salverda, M. L. M., Dellus, E., Gorter, F. A., Debets, A. J. M., van der Oost, J., Hoekstra, R. F., Tawfik, D. S., and de Visser, J. A. G. M., 2011. Initial mutations direct alternative pathways of protein evolution. *PLoS Genetics*, **7**(3):1–11. doi:10.1371/Journal.pgen.1001321.
- [96] Betancourt, A. J., 2010. Lack of evidence for sign epistasis between beneficial mutations in an RNA bacteriophage. *Journal of Molecular Evolution*, **71**(5):437–443. doi:10.1007/s00239-010-9397-0.
- [97] Krause, R. M., 1992. The origin of plagues: Old and new. *Science*, **257**(5073):1073–1078. doi:10.1126/science.257.5073.1073.
- [98] Nathanson, N., McGann, K. A., Wilesmith, J., Desrosiers, R. C., and Brookmeyer, R., 1993. The evolution of virus diseases: their emergence, epidemicity, and control. *Virus Research*, **29**(1):3–20. doi:10.1016/0168-1702(93)90122-4.