Evolutionary exploration of a bacterial LPS genotype to phenotype map with phages

Dissertation

in fulfilment of the requirements for the degree of **Dr. rer. nat.**

of the Faculty of Mathematics and Natural Sciences

at Christian-Albrechts-Universität zu Kiel

Submitted by Jordan Romeyer Dherbey

Kiel (2023)

First examiner: Dr Frederic Bertels

Second examiner: Prof Dr Hinrich Schulenburg

Date of the oral examination: 17/02/2023

I, **Jordan Romeyer Dherbey**, declare that apart from my supervisor's guidance, the content and design of the thesis is my research work in its entirety. The thesis has been prepared following the Rules of Good Scientific Practice of the German Research Foundation. No academic degree has ever been withdrawn.

Part of this thesis was used to write the following research article: **Stepwise evolution of *E. coli* C and ΦX174 reveals unexpected lipopolysaccharide (LPS) diversity** (submitted).

A full version of the manuscript (**Romeyer Dherbey, J.**, Parab, L., Gallie, J., and Bertels, F. *E. coli*-ΦX174 genotype to phenotype map reveals flexibility and diversity in LPS structures. *BioRxiv* 2022) can be found here: https://www.biorxiv.org/content/10.1101/2022.09.06.506728v1.full. The copyright holder for this preprint is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.

Authors' contributions:

Jordan Romeyer Dherbey: Conceptualization, Methodology, Investigation, Formal Analyses, Validation, Visualisation, Supervision, Writing - Original Draft, Writing – Review & Editing.
Lavisha Parab: Investigation.
Jenna Gallie: Validation, Visualisation, Writing – Review & Editing.
Frederic Bertels: Conceptualization, Methodology, Formal Analyses, Validation, Supervision, Writing - Original Draft, Writing – Review & Editing, Project administration.

Signature:

Table of Contents

Abstract

Multidrug-resistant bacterial infections are one of the most pressing issues in medicine, a situation that is only expected to worsen in the coming decades. An alternative to treating bacterial infections with antibiotics is phage therapy. However, the long-term success of phages as therapeutic agents stems from our ability to understand and predict the evolutionary responses of both phages and bacteria. In this thesis work, I use *Escherichia coli* C and its coliphage ΦX174 to demonstrate that our comprehension of phage-bacteria interactions is still sorely lacking, even for this well-established model system.

1. I generated a library of 31 *E. coli* C strains that resist ΦX174 infection and show that all mutants carry at least one mutation in genes linked to core lipopolysaccharide (LPS) biosynthesis, assembly or regulation. Based on which genes are mutated and on the current knowledge of LPS biology, I predict that these 31 *E. coli* C resistant strains collectively produce eight different structures.

2. I designed and performed a series of evolution experiments that successfully yielded highly LPS-specific phage strains that can, all together, infect all 31 *E. coli* C resistant strains. I find that ΦX174 overcomes LPS-based resistance via mutations in only two genes: the *F* gene involved in host recognition and the *H* gene involved in DNA injection.

3. To test my predicted LPS genotype to phenotype maps, I use multiple phage infection assays to qualitatively determine the host range of each evolved phage. If my predictions are correct, then all bacteria predicted to display the same LPS structure should be infected by the same set of evolved phages. However, this was not always the case. Instead of the eight predicted distinct LPS profiles, the infectivity patterns suggest 14 distinct LPS structures among the 31 *E. coli* C mutants.

My results demonstrate that the current understanding of LPS biology – which derives from single-gene deletion and complementation approaches – is insufficient to accurately predict the evolutionary consequences of infecting bacterial populations with phages. My approach to infer LPS structures by using an evolution experiment allows a deeper understanding of bacteria-phage interactions essential for developing efficient phage therapies in the future.

Abstrakt

Translated by DeepL and Dr Frederic Bertels

Multiresistente bakterielle Infektionen sind eines der drängendsten Probleme in der Medizin, eine Situation, die sich in den kommenden Jahrzehnten nur noch verschlimmern dürfte. Eine Alternative zur Behandlung bakterieller Infektionen mit Antibiotika ist die Phagentherapie. Der langfristige Erfolg von Phagen als Therapeutika hängt jedoch von unserer Fähigkeit ab, die evolutionären Interaktionen von Phagen und Bakterien zu verstehen und vorherzusagen. In dieser Dissertation zeige ich anhand von *Escherichia coli* C und seinem Coliphagen ΦX174, dass unser Verständnis für die Interaktionen von Phagen und Bakterien, selbst bei diesem etablierten Modellsystem, noch sehr lückenhaft ist.

1. Ich habe eine Bibliothek von 31 *E. coli* C-Stämmen erstellt, die gegen eine ΦX174-Infektion resistent sind, und gezeigt, dass alle Mutanten mindestens eine Mutation in Genen tragen, die für die Lipopolysaccharid (LPS)-Biosynthese, dessen Aufbau oder dessen Regulierung verantwortlich sind. Je nachdem, welche Gene mutiert sind, und auf der Grundlage des derzeitigen Wissens über die LPS-Biologie sage ich voraus, dass diese 31 resistenten E. coli C-Stämme zusammen acht verschiedene Strukturen bilden.

2. Ich habe eine Reihe von Evolutionsexperimenten entworfen und durchgeführt, die erfolgreich hochgradig LPS-spezifische Phagenstämme hervorgebracht haben, die zusammen alle 31 *E. coli* C resistenten Stämme infizieren können. Ich habe festgestellt, dass ΦX174 die LPS-basierte Resistenz durch Mutationen in nur zwei Genen überwindet: dem F-Gen, das an der Wirtserkennung beteiligt ist, und dem H-Gen, das an der DNA-Injektion beteiligt ist.

3. Um die von mir vorhergesagten LPS-Genotyp-Phänotyp Abbildungen zu testen, verwende ich mehrere Phageninfektionstests, um die Wirte eines jeden evolvierten Phagen qualitativ zu bestimmen. Wenn die ursprünglichen Vorhersagen richtig sind, dann sollten alle Bakterien, für die dieselbe LPS-Struktur vorhergesagt wurde, von derselben Gruppe von evolvierten Phagen infiziert werden. Dies war jedoch nicht immer der Fall. Anstelle der acht vorhergesagten unterschiedlichen LPS-Profile lassen die Infektiositätsmuster auf 14 unterschiedliche LPS-Strukturen unter den 31 *E. coli* C-Mutanten schließen.

Meine Ergebnisse zeigen, dass das derzeitige Verständnis der LPS-Biologie – das auf der Deletion einzelner Gene und auf Komplementationsansätzen beruht – nicht ausreicht, um die evolutionären Folgen der Infektion von Bakterienpopulationen mit Phagen genau vorherzusagen. Mein Ansatz zur Bestimmung von LPS-Strukturen mit Hilfe eines Evolutionsexperiments ermöglicht ein tieferes Verständnis der Wechselwirkungen zwischen Bakterien und Phagen, die für die Entwicklung effizienter Phagentherapien in der Zukunft unerlässlich sind.

Acknowledgments

My first thanks go to my supervisor Dr Frederic Bertels. He gave me the incredible chance to start my journey in the evolutionary biology world. Despite having no background in phage biology and evolution at this time, he trusted me to start his group with him. I thank him for the countless hours he spent discussing the outcomes and new perspectives yielded by the project and the ones he spent correcting my various writings to improve my scientific communication. Without him, this PhD thesis would have not been possible. It was a true pleasure to grow as a scientist under his guidance and to see him grow as a team leader through these years.

Furthermore, I would like to thank Dr Jenna Gallie for her collaboration and invaluable impact on the quality of the research article that emerge from my PhD project. I cross my fingers for this research article to be reviewed and accepted soon.

I also thank my two thesis advisory committee members, Prof. Dr Hinrich Schulenburg and Prof. Dr Daniel Unterweger, for their advice on my ongoing project and their support during these PhD years.

The COVID pandemic from 2020 impacted us all, both physically and psychically. Thankfully, I was part of a department full of great scientists that stuck together when times were difficult. They helped me to face the pandemic and the many upside-downs inherent to a PhD work. Therefore, I thank everyone from the Microbial Population Biology department (past and present) for their scientific advice, especially Dr David Rogers, Dr Loukas Theodosiou, Dr Elisa Brambilla, and Lavisha Parab for their help and support on my project, research article, presentations and job interviews. A special thanks to Prof. Dr Paul Rainey for extending my PhD contract, allowing me to finish my PhD work and summarizing it into the present thesis. I do not forget the lab technicians that make the lab run as smoothly as possible and Britta Baron, probably the best administrative assistant that someone can wish for.

The Max Planck Institute for Evolutionary Biology was a great place to carry out my PhD (even if a bit isolated from everything). I had a great time with my PhD fellows (Eric, Bilal, Nico, Aleksa, Christopher, Dana, Jule, and Carolina, to only cite a few of them), and I hope to meet them all again!

Table of Abbreviations

Abbreviation	Meaning
bp	Base pair
Cas proteins	CRISPR associated proteins
Cfu	Colony-forming unit
Core OS	Core oligosaccharide
CRISPR	Clustered regularly interspaces short palindromic repeats
dsDNA	Double stranded DNA
fw primer	Forward primer
gDNA	Genomic DNA
HGT	Horizontal gene transfer
LB	Lysogeny Broth, Miller
LPS	Lipopolysaccharide
LTEE	Long term evolution experiment
MMR	Methyl-directed mismatch repair system
MOI_{actual}	Actual multiplicity of infection: the number of phages that have *actually* adsorbed/infected the susceptible bacteria
MOI_{input}	Input multiplicity of infection: the ratio of phages *added* to susceptible bacteria
MQ water	Milli-Q water
nt	Nucleotide
OD600	Optical density at 600 nm
PBJ	Phage buffer juice
Pfu	Plaque-forming unit
rv primer	Reverse primer
SSA	Semi-solid agar
ssDNA	Single stranded DNA
s/n/c	Substitutions per nucleotide per cell infection

Chapter I – Introduction

Phages are viruses that prey upon bacterial hosts. They are the most abundant organisms and can be found everywhere on Earth. Decades of research have revealed phages' major ecological and evolutionary impacts on various microbial ecosystems. These studies greatly promoted the development of important phage-derived tools applied in molecular biology and sequencing technologies. Nevertheless, the idea of phages as therapeutic agents used to treat bacterial infections until recently has been neglected by the scientific community in favour of antibiotics. With the emergence of widespread antibiotic resistance, phages became an appealing alternative to antibiotics in the fight against multidrug-resistant bacteria. The high host specificity of phages that once was considered a weakness is now praised and sought-after. Unfortunately, bacteria can also become resistant in response to phage infection via various mechanisms. Therefore, a more in-depth understanding of the mechanisms behind the phage resistance in bacteria is required to fully harness the potential of phages in medicine.

1.1. A short tale of phages

1.1.1. Phages co-discovery

The canonical discovery of bacteriophages (or simply phages) is attributed to Frederick Twort and Félix d'Hérelle in 1915[1] and 1917[2], respectively. In 1915, Frederick Twort searched for a way to cultivate the so-called non-pathogenic filter-passing ultra-microscopic viruses *in vitro* to prove their existence. In *The Lancet,* he described a degenerative and transmissible "glassy transformation" that turned viable micrococcus colonies into dead transparent ones[1]. Independently from F. Twort's work, Félix d'Hérelle isolated an invisible anti-Shiga microbe from the stool and urine of patients recovering from bacillary dysentery[2]. He published his findings in the *Comptes rendus de l'Académie de Paris*, where he correlated the recovery of previously sick patients with the bactericidal action of the anti-Shiga microbe. To name the newly discovered microbe of immunity, he introduced the term "bacteriophage" – a bacteria-eater[2].

Older descriptions of credible or presumptive "prehistoric" phage references predating this co-discovery were recently identified[3]. One of the oldest mentions dates back to 1896 when the English chemist Ernest Hankin described the action of a "bactericidal substance" of an unknown origin and nature found in the water of the Gange and Jumna[4]. However, the veracity of these early "prehistoric" observations of antimicrobial activities is still debated today[3].

1.1.2. Important discoveries linked to phages

The field of molecular biology greatly benefited from over a century of phage research, for instance, in a form of tools used daily in molecular biology[5–7]. Phages with small genomes, such as MS2 (3569 nt) and ΦX174 (5386 nt), were the first sequenced organisms[8–10]. Shotgun sequencing methods used for larger phages, like phage λ (48502 bp), used almost exclusively phage-derived products such as T4 DNA ligase and M13 vectors. These early methods were further developed to sequence more complex organisms, ranging from *Escherichia coli* to humans[11]. Another major molecular tool originated from the CRISPR-Cas adaptative immune systems of Archaea and Eubacteria against foreign DNA, such as phages and plasmids. While CRISPR-Cas yielded an extremely sophisticated genome-editing tool[12,13], it also provided a deeper understanding of host-virus relationships and resistance mechanisms[14].

1.1.3. The renewal of interest for phages in therapeutics

After his discovery, d'Hérelle immediately saw the potential of phages as tools for therapeutic uses in animals and humans. However, debates on the controversial nature of phages slowly undermined their application and phages were rejected in favour of the newly-discovered antibiotics. Unlike antibiotics, phages lacked standardized production and controls, and their host spectrums were considered too narrow[15]. Furthermore, the association of phage therapy with German and Japanese medicine during the Second World War put an end to any further applications in the West[15]. Unfortunately, the overuse and misuse of antibiotics have driven the emergence of multidrug-resistant bacteria for years - a situation that is only expected to worsen in the coming decades[16–18]. Infections with multidrug-resistant bacterial strains are one of the most pressing issues in medicine, which resulted in a vivid interest in phages as promising agents to treat bacterial infections[19,20].

1.2. Phage morphologies and classification

Phages are simple entities. Their genome, either RNA or DNA, single- or double-stranded, is protected by a capsid. This genome-capsid complex is called a nucleocapsid. Phage capsid morphologies are highly diverse (**Fig. 1.1**), taking various shapes: polyhedral (*Microviridae, Corticoviridae, Tectiviridae, Leviviridae,* and *Cystoviridae*), filamentous (*Inoviridae, Lipothrixviridae,* and *Rudiviridae*), and pleiomorphic (*Plasmaviridae,* the curious lemon-shaped *Fuselloviridae* and *Salterprovirus* as well as the droplet-shaped *Guttaviriade*)[21–23]. The pleiomorphic and filamentous

phages (*Lipothrixviridae*), albeit not all, also have lipid envelopes which surround the nucleocapsid. The vast majority of isolated phages are dsDNA with a tail attached to the capsid (*Caudoviridae*)[22,23].

Historically, the taxonomic classification of prokaryotic viruses relied on specific characteristics, such as structural morphology, the nature of the genomic material, and host ranges. The tremendous progress made recently in sequencing technologies turned the tide. Now sequence comparisons using pairwise sequence similarities and phylogenetic relationships are used to classify phages[24].
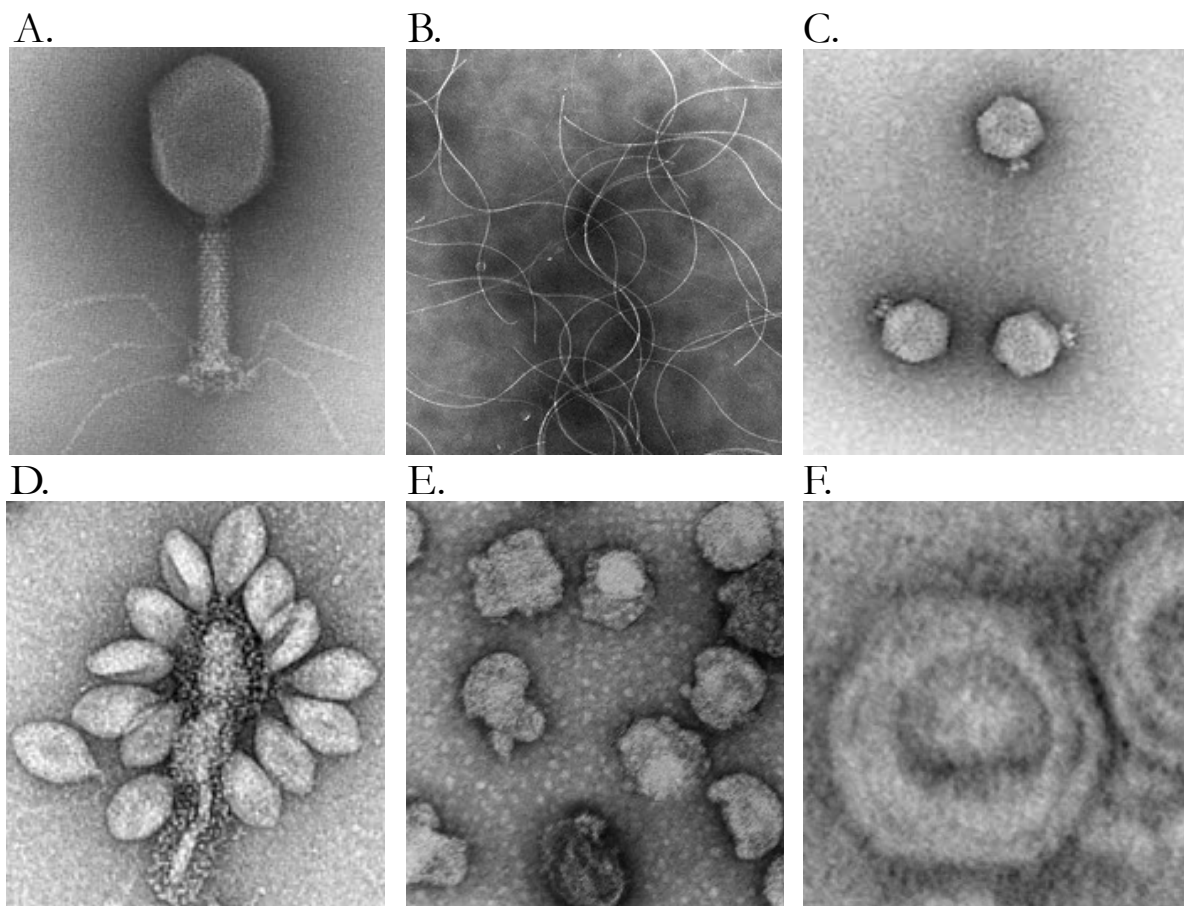
Fig. 1.1. Morphological diversity of phages. Electron-microscopy pictures of phages: **A.** *Myovirus* T4 (from R. Duda/University of Pittsburgh). **B.** *Inovirus* SW1 (from Wang *et al.*, 2007[25]). **C.** *Podovirus* P22 (from Casjens and Lenk, 1998[26]). **D.** *Fusellovirus* SSV-1 (from Stedman *et al.*, 1999[27]). **E.** *Plasmavirus* L2 (from Poddar *et al.*, 1985[28]). **F.** *Tectivirus* Wip1 (from Schuch and Fischetti, 2009[29]).

1.3. Life cycles of phages

Like any other virus, phages are obligate parasites that cannot survive on their own. To propagate and persist, they must infect a bacterial host and hijack its replicative machinery to produce new virion particles. A typical life cycle starts with a phage encountering a susceptible bacterium. The phage specifically recognizes and binds to a receptor located at the surface of the host cell envelope. Adsorption triggers specific conformational changes in the phage for the structural proteins (e.g., tails or tube-like structures) to penetrate the bacterial outer and inner membranes and deliver the phage genomic material inside the host cytoplasm[30,31].

In the lytic life cycle, the phage genome reroutes the host cell replication system (ribosomes, ATP generators) to its benefit. The host cell becomes a virus factory: expressing phage genes, replicating the phage genome and assembling new viral progenies. It ends with the bacterium's death from a burst-out of hundreds of virions[32] (**Fig. 1.2**). In the lysogenic life cycle (also called temperate), the genomic material of the phage integrates into the bacterial genome as a prophage, and a fraction of the bacterial population will survive the infection. Certain environmental conditions, such as UV radiations, chemicals or other mutagens, can trigger the excision of the prophage and its entry into the lytic life cycle (**Fig. 1.2**).
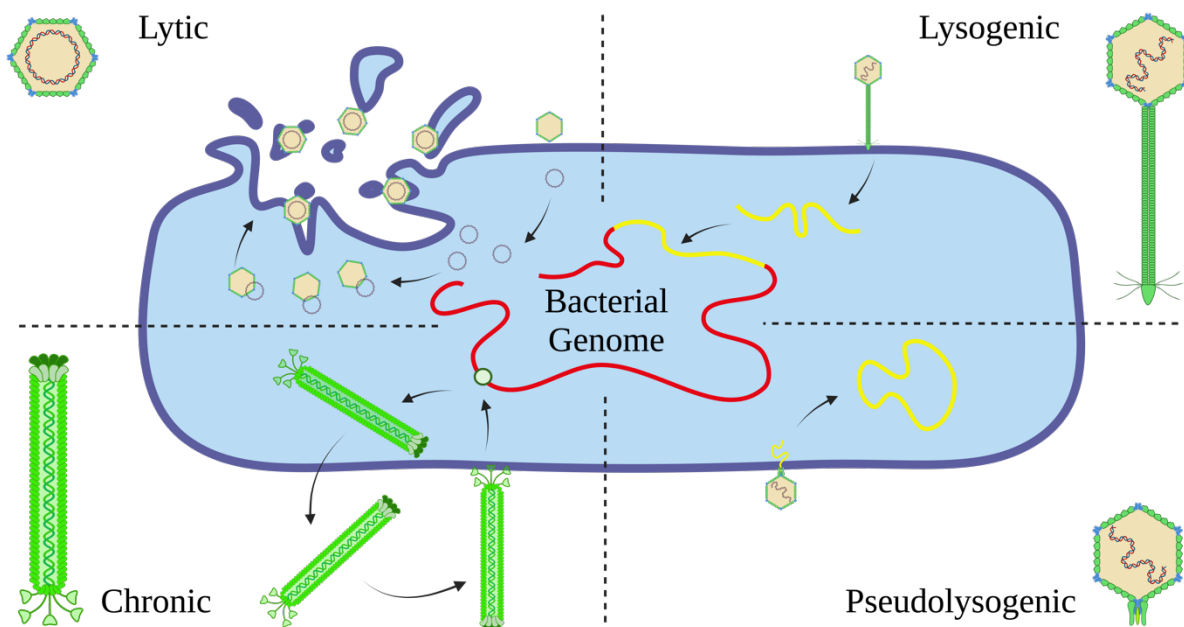
Fig. 1.2. Four known phage life cycles. The lytic life cycle consists of the active replication of the phage within its host, followed by host death triggered by the release of new phage particles

into the outer environment. In the lysogenic life cycle, the phage is integrated into the bacterial genome (prophage). The chronic life cycle allows infection and phage replication without the bacterial host's death. The new virion particles are excreted from the bacterial membrane without lysis of the host; the host is transformed into a permanent virus factory. The pseudolysogeny life cycle is a temporary state between the lytic and lysogenic life cycles. The phage genome remains dormant inside the host cytoplasm without being replicated (adapted from Lawrence *et al.*, 2019[33]).

Alternatives to the lytic-lysogenic life cycles exist. Filamentous phages from the *Inoviridae* family can establish chronic infections where new phage particles are constantly extruded into the environment without killing their hosts[34,35]. *Inoviridae* phages are large, threadlike phages whose size can sometimes exceed their susceptible hosts (~800 nm to 4 µm)[36]. Their impressive size does not allow the assembly within the bacterial cell but rather in the host membrane (**Fig. 1.2**). The last major life cycle is called pseudolysogeny. Pseudolysogenic phages can co-exist with bacteria in a relatively stable equilibrium. The phage genome is in a quiescent, temporary state that does not replicate or integrate into the host genome after its uptake by the bacterium[37–39] (**Fig. 1.2**).

1.4.   The ecological impact of phages

We are living in a world dominated by microbes. Phages are reported to be the most numerous organisms on Earth[40] with an estimated 10^{31} particles, outnumbering bacteria by a factor of ~10. They are ubiquitous in all natural and anthropogenic biomes (human-altered biomes or artificially created, such as wastewater treatment reservoirs or industries) on Earth[41,42]. As phages constantly infect and kill their hosts, they heavily contribute to the turnover of the microbial ecosystem by constantly shuffling microorganisms' diversity over time. The ecological impacts and relevance of phages on various microbial communities have been widely explored in aquatic[43,44], terrestrial[45], and within-organism[46–48] ecosystems, including the redistribution of nutrients from bacterial lysis and access to new ecological niches. For example, metagenomic analyses performed on the Sargasso Sea samples revealed that 80% of the viral particles present belonged to the *Microviridae* family, suggesting a preponderant role of small ssDNA viruses in the marine habitat[49]. Furthermore, by integrating into the bacterial genome via horizontal gene transfer (HGT), phages can modify the metabolism of their hosts[50–52], alter their resistance to antibiotics[53,54], or increase their pathogenicity and virulence[55,56].

1.5. Understanding evolution through phages

1.5.1. Experimenting with microbes

The co-evolutionary dynamic between phages and their hosts has been investigated to understand phages' ecological and evolutionary roles in bacterial diversity. To do so, experimental evolution approaches have been used to test evolutionary theories concerning host-pathogen relationships in controlled environments[57,58]. Experimental evolution often happens in the laboratory, where the model organism can be monitored and manipulated with a high degree of control and replication. Because of their simple nature, phages are popular models used to study central biological questions, such as the nature and structure of genes[59,60], phage-related concepts (such as adsorption, the latent period, and viral burst)[61], as well as the random nature of mutations in the absence of selection[62].

Phages are particularly appreciated in experimental evolution because they are easy to handle and propagate, often under low safety requirements. Their small genome, large population sizes, short generation times, and high mutational rates[63,64] facilitate the tracking and characterisation of many generations from short to long time scales[65,66]. Hence, robust statistical analyses can be performed to test diverse evolutionary hypotheses. Samples can be taken and frozen at different moments of the experiments to constitute banks of "fossil" records. From these records, evolution can be resumed from any timepoint (replaying the "tape of life") to better quantify and predict evolutionary outcomes[67–69]. Finally, phages can be manipulated, transformed, and eradicated without raising ethical issues.

1.5.2. Old but gold – phage ΦX174 as an experimental model organism

Among phage families, one of the most well-studied is the *Microviridae* family. It can be divided into five groups: the microviruses (genus *Microvirus*), the subfamilies *Gokushovirinae, Alpavirinae,* and *Pichovirinae*[70–72], as well as the more recent sister clade *Pecheñovirus*[73]. *Microviridae* phages are small, tailless, and carry a circular ssDNA genome (from ~4.4 to 6.1 kb) concatenated in an icosahedral capsid[30,71]. They have been found in freshwater lakes, seas, oceans, microbialites, corals, human gut, and lungs[49,71]. These widely spread phages prey upon distantly related hosts, such as *Escherichia coli* (γ-proteobacteria), *Chlamydia, Bdellovibrio,* and *Spiroplasma*[30,70].

Phage ΦX174, a coliphage from the *Microviridae* family, perfectly embodies all characteristics that make an organism suitable for exploring evolutionary theories. ΦX174 is an old model system[74]

that has been successfully used to study the adaptation to new hosts[75,76] and high temperatures[77,78], host range expension[79,80], selection for host recognition[81], parallel evolution[82,83], evolutionary dynamics[84–86], mutational and fitness landscapes[87,88], compensatory mutations[89,90], epistasis[91], and recombination[83,92].

1.6.   Evolving ΦX174 to explore resistant-host phenotypic diversity

The evolution of bacterial resistance-to-phage strategies, and the mechanisms by which phages adapt to overcome these strategies, are usually studied via evolution experiments in which phages are added to sensitive bacteria and allowed to co-evolve. While this top-down approach has improved our understanding of co-evolutionary dynamics, it underestimates the diversity of available evolutionary pathways. For instance, serial transfer experiments in a liquid environment strongly select for rapidly growing bacteria that can quickly outcompete other resistant but slower counterparts. Hence, resistant phenotypes associated with slow growth are understudied, despite their high relevance in natural environments where limited resources curb rapid growth[93].

Phage resistance can often evolve by modifying the cell envelope to avoid phage adsorption. The outer membrane of all Gram-negative bacteria is covered in lipopolysaccharides (LPS), which is consequently the first contact point between a bacterium and its environment[94]. LPS structure determines whether a bacterium will be recognized by the immune system, the efficiency of nutrient uptake as well as the susceptibility to antibiotics and, importantly, the susceptibility to phage infection[95–97]. Hence, LPS structure is an important factor influencing the success of phage therapy.

Here, I proposed an evolutionary bottom-up approach to study the natural LPS structure diversity in resistant *E. coli* C strains and ΦX174 adaptation processes to host harbouring modified LPS structures.

Objective 1 – Causes of ΦX174-resistance in *Escherichia coli* C (*E. coli* C). My first objective was to explore the evolutionary potential of *E. coli* C to become resistant to infection by ΦX174. I investigated the mutational causes responsible for ΦX174-resistance using whole-genome sequencing and tied them to specific predicted LPS phenotypes.

Objective 2 – ΦX174 adaptation to resistant *E. coli* C strains. My second objective was to determine the amino acid sites under selection for host recognition in the genome of ΦX174. Using serial transfer experiments to evolve ΦX174 against each resistant mutant, I determined the set of mutations required to overcome each resistance. I then linked these mutations to specific infection patterns.

Objective 3 – Testing the predicted LPS genotype-phenotype maps. The currently accepted model for predicting LPS structure posits that the presence or absence of each LPS gene leads to a single LPS phenotype. My third and last objective was to challenge this "LEGO® block" principle (e.g., if an enzyme that synthesizes a particular LPS component is absent or non-functional, then that LPS component cannot be synthesized). Using my evolved ΦX174 strains as biosensor tools, I was able to discriminate the phenotype of each *E. coli* C resistant strain.

I demonstrate that not only can phages be used to study LPS structures but also that ΦX174 can be effectively bred to infect a wide variety of resistant mutants. My results could potentially open up the possibility of evolving ΦX174 to infect any *E. coli* strain, including pathogenic strains. Given that ΦX174 is already approved *in vivo* as a marker of immune responses in humans[98,99], it may prove a viable alternative as a therapeutic agent to uncharacterized and potentially unsafe environmental phage isolates.

Chapter II –Methods

2.1.   General Methods

2.1.1.   Bacterial and phage strains

The ancestral *Escherichia coli* strain C (*E. coli* C) used in this study differs from the one uploaded on the NCBI website (GenBank accession number CP020543.1) by nine additional insertions and two nucleotide substitutions (c → t at position 1,1720,214 and g → t at position 1,190,560). The sequence of CP020543.1 was manually modified using Geneious Prime (version 2020.1.2). All resistant bacterial strains generated in this study are derived from our ancestral *E. coli* C strain.

All bacteriophage strains are derived from the ancestral coliphage ΦX174 (GenBank accession number AF176034). AF176034 genomic sequence was downloaded from the NCBI website and manually annotated using Geneious Prime (version 2020.1.2). Whole-genome re-sequencing on ΦX174 wildtype's glycerol stock showed no difference to the NCBI sequence. Complete genomes of *E. coli* C wildtype and ΦX174 wildtype used as references have been deposited on Zenodo (doi: 10.5281/zenodo.6952399).

The *E. coli* C and ΦX174 ancestral strains were provided by Holly A. Wichman (Department of Biological Sciences, University of Idaho). *E. coli* B strain REL606 (GenBank accession number NC_012967.1), *E. coli* K-12 strain MG1655 (GenBank accession number NC_000913.30), *E. coli* K-12 strain BZB1011 (no known accession number), and *E. coli* K-12 strain AQ4591 (no known accession number) were provided by Jenna Gallie (Max Planck Institute for Evolutionary Biology, Plön).

2.1.2.   Media and growth conditions

Phages and bacteria were grown in a shaking incubator (New Brunswick™ Innova® 44) at 37°C, 250 rpm, in Lysogeny Broth (LB), Miller medium supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively. Solid LB agar (1.5% agar) was used to plate bacteria and phages. When plating phage, top agar overlay (0.5% agar), also called semi-solid agar (SSA) in this study, was supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively. Phage buffer juice solution (PBJ: 2.03%, Tris-HCl, 0.61%, $MgCl_2.6H_2O$) was used for serial dilutions.

E. coli strains' colony plates were generated by streaking with a sterile loop a scrap of the corresponding bacterial glycerol stock on the surface of a solid LB plate and incubated inverted overnight. After incubation, the colony plates were stored at 4°C for a maximum of two weeks. Overnight bacterial liquid cultures were systematically started by inoculating a single colony in 5 ml of LB supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively, shaken at 250 rpm. Phage infections were either started from (i) a scrap of the corresponding phage glycerol stock resuspended in 100 µl of PBJ or (ii) a single plaque isolated from the top agar overlay with a sterile cut tip (a circle of agar) directly inoculated inside the bacterial culture. Overnight incubations of both plates and cultures were set up for ~16-17 hours at 37°C.

2.1.3. Data visualisation

The hierarchical agglomerative clustering analysis (Clustermap **Figs. 5.1** and **5.S1**), the heatmaps (**Figs. 4.10**, **4.12B**, **4.S1**, **5.4**, and **5.S2** to **5.S3**), and growth curves (**Figs. 3.10** and **5.5**) were generated using the default settings of the Seaborn library (version 0.11.2) for Python (version 3.7.4). All schematic drawings (**Figs. 1.1, 3.1, 3.2, 3.3, 3.7, 3.11, 4.1, 4.2, 4.4, 4.7, 4.9, 4.11** and **5.2**) were created with BioRender.com. The Lollipop plot (**Fig. 4.10A**) and growth curves (**Figs. 3.8, 3.9**, and **4.3**) were generated using trackViewer Vignette: lollipopPlot (Lolliplot) in R (version 1.27.11).

2.1.4. Data availability

Findings from my thesis work were condensed in a manuscript and deposited on *BioRxiv* (https://www.biorxiv.org/content/10.1101/2022.09.06.506728v1.full). Locations of the dataset used in the manuscript are detailed below.

Whole-genome and Sanger sequencing results. Complete genomes of *E. coli* C wildtype and ΦX174 wildtype used as references, and raw sequencing reads used to generate **Tables 3.1** to **3.3** and **Tables 4.2 to 4.5** have been deposited on Zenodo (doi: 10.5281/zenodo.6952399). *Growth curves.* Raw OD measurements used to generate **Fig. 3.10** have been deposited on Zenodo (doi: 10.5281/zenodo.6952399). *Hierarchical agglomerative clustering and heatmap data.* Pictures and raw data used to generate **Figs. 4.10** and **5.1** have been deposited on Zenodo (doi: 10.5281/zenodo.6952399).

2.2. Methods Chapter 3

2.2.1. Isolation and storage of ΦX174-resistant *E. coli* C strains

To avoid competition and produce a diverse set of phage-resistant bacterial mutants with different growth phenotypes, I generated ΦX174-resistant *E. coli* C strains on agar plates. Thirty-five *E. coli* C wildtype colonies were randomly chosen from an agar plate, and each was used to inoculate an independent overnight liquid culture. 50 µl of each stationary phase culture was mixed with 50 µl of a high titer (~10^9 plaque-forming unit pfu.ml^{-1}) stock of ΦX174 wildtype lysate, immediately plated with sterile beads on LB plates, and incubated at 37°C overnight. From each plate, half a randomly chosen colony was used to inoculate a fresh overnight LB liquid culture. Once grown, a sample from each culture was mixed with glycerol saline solution and frozen at -80°C (giving strains *E. coli* C R1 to R35). To confirm the resistance to wildtype ΦX174, the remaining half of each colony was streaked onto a LB plate soaked with 100 µl of the same high titer ΦX174 wildtype lysate and incubated at 37°C overnight.

2.2.2. Measuring the growth of *E. coli* C wildtype without ΦX174

To understand the growth dynamics of *E. coli* C wildtype, three independent cultures of *E. coli* C wildtype were grown at 37°C, 250 rpm for 17 hours. 125 µl of each overnight culture was inoculated into a 14 ml sterile tube (FALCON®) containing 4.875 ml of LB (supplemented with CaCl$_2$ and MgCl$_2$ at final concentrations of 5 and 10 mM, respectively) pre-warmed at 37°C. Incubation time was set to 5 hours at 37°C, 250 rpm. The OD600 values were measured every 30 minutes with an Ultrospec 10 (Biochrom®). Two dilutions (10^{-5} and 10^{-6}) were plated for each timepoint and each condition and incubated inverted at 37°C overnight. Phage plaques and bacterial colonies were counted the next day.

2.2.3. Impacts of the serial dilution methods and media used for bacterial enumeration

Three different methods of serial dilutions were tested: (i) 1:10 serial dilution in 180 µl of PBJ in a 96 well plate, (ii) 1:10 serial dilution in 180 µl LB supplemented with CaCl$_2$ and MgCl$_2$ (at final concentrations of 5 and 10 mM, respectively) in a 96 well plate, (iii) 1:10 serial dilution in 900 µl PBJ in Eppendorf tubes. 100 µl of dilutions 10^{-5} or 10^{-6} were plated with ~10 sterile glass beads on solid agar plates and incubated inverted at 37°C for 17 hours. Colonies were counted the next day.

2.2.4.   Measuring the growth of each resistant *E. coli* C strain

To assess the impact of LPS modifications harboured by the different *E. coli* C resistant mutants on their growth, two independent cultures of each resistant strain were grown at 37°C, 250 rpm for 17 hours. 180 µl of each overnight culture was inoculated into a 14 ml sterile tube (FALCON®) containing 5 ml of LB (supplemented with $CaCl_2$ and $MgCl_2$ at final concentrations of 5 and 10 mM, respectively) pre-warmed at 37°C. Incubation time was set to six hours at 37°C, 250 rpm. Optical density at 600 nm (OD600) was measured every 30 minutes with an Ultrospec 10 (Biochrom®). *E. coli* C wildtype was used as a reference. The mean OD600 values at each time point for each bacterial mutant were calculated and plotted.

2.2.5.   *E. coli* C whole genome re-sequencing

Samples were prepared for whole genome re-sequencing from 1 ml of stationary-phase culture. Genomic DNA was extracted using the Wizard® Genomic DNA purification Kit (Promega). Extracted DNA was tested for quality, pooled, and sequenced by the Max-Planck Institute for Evolutionary Biology (Plön, Germany) using an Illumina Nextera DNA Flex Library Prep Kit to produce 150 bp paired-end reads[100]. The quality of the sequencing output was controlled using *FastQC* version 0.11.8 (**Appendix. A1**). Reads were trimmed using *Trimmomatic* (**Appendix. A1**), assembled and analysed using the *breseq* pipeline version 0.33.2[101–103] (**Appendix. A2**) and Geneious Prime (version 2020.1.2).

2.3.   Methods Chapter 4

2.3.1.   Phage lysate preparation

To extract phages in infected host cultures, ten to twelve drops of chloroform were added to growing *E. coli* C-phage co-cultures and vortexed for at least 30 seconds to kill the bacteria. Cultures were then centrifuged at 5000 rpm for 10 min at 4°C. Supernatants (phage lysates) were transferred to sterile 5 ml Eppendorf tubes and stored at 4°C. 1 ml of each lysate was stored with glycerol saline solution and frozen at -80°C.

2.3.2.   Enumeration of phage particles

To determine the number of phage particles contained in the phage lysates, 20 µl of each phage lysate were serial diluted 1:10 up to 10^{-7} in PBJ. Top agar overlays were prepared by mixing 100 µl of dilutions 10^{-6} or 10^{-7} with 200 µl of an overnight culture of the corresponding resistant *E. coli* C

strains or 100 μl of an overnight culture of *E. coli* C wildtype in 4 ml SSA (supplemented with CaCl$_2$ and MgCl$_2$ at a final concentration of 5 and 10 mM, respectively). Top agar overlays were poured onto LB plates, dried for at least 15 minutes, and incubated inverted at 37°C for ~16-17 hours. Plaques were counted the next day.

2.3.3. Measuring the growth of *E. coli* C wildtype in the presence of ΦX174

To understand the growth dynamics of ΦX174 in a wildtype *E. coli* C culture, four independent cultures of *E. coli* C WT were grown at 37°C, 250 rpm for 17 hours. 125 μl of each ON culture was inoculated into a 14ml sterile tube (FALCON®) containing 4.875 ml of LB (supplemented with CaCl$_2$ and MgCl$_2$ at final concentrations of 5 mM and 10 mM, respectively) pre-warmed at 37°C. Incubation time was set to 7 hours at 37°C, 250 rpm. Optical density (OD) was measured at 600nm every hour with an Ultrospec 10 (Biochrom®). At t = 60, 120 and 180 min, ~1.10^6 phage particles (MOI$_{input}$ ~ 0.01) were added to the different host cultures. Two dilutions (10^{-5} and 10^{-6} in PBJ) were plated for each timepoint and each condition and incubated inverted at 37°C overnight. Phage plaques and bacterial colonies were counted the next day.

2.3.4. Isolation of ΦX174 strains from single plaques

To isolate pure, single clones from the different evolution experiments, top agar overlays were prepared by mixing 100 μl of undiluted phage lysates with 200 μl of an overnight culture of the corresponding resistant *E. coli* C strain in 4 ml SSA (supplemented with CaCl$_2$ and MgCl$_2$ at a final concentration of 5 and 10 mM, respectively). Top agar overlays were poured onto LB plates, dried for at least 15 minutes, and incubated inverted at 37°C for ~16-17 hours. Plaques were counted the next day. An isolated plaque was chosen at random for each phage lysate to infect cultures of the corresponding resistant *E. coli* C strains (in exponential growth state) incubated at 37°C, 250 rpm, for 5 hours. Phage lysates were isolated as described in **Phage lysate preparation** (section **2.3.1**). To obtain pure phage glycerol stocks (isogenic stocks which are presumed to hold one genetic clone[104]), phages were purified via a second round of plaque isolation from their first glycerol stocks. Phage lysates were isolated and stored as described in **Phage lysate preparation** (section **2.3.1**).

2.3.5. Phage evolution experiment – Standard serial transfers

The protocol from Bono *et al.*[105] was adapted to evolve phages that can infect resistant *E. coli* C strains (see **Fig. 4.4**). Thirty-five ΦX174 lineages were each founded by ΦX174 wildtype and

serially transferred daily, for up to 21 days, on non-evolving host cultures containing a mix of (i) *E. coli* C wildtype (permissive strain) and (ii) one resistant *E. coli* C strain (non-permissive strain; R1 to R35, see **Tables 3.1** and **3.2**).

For each transfer, permissive and non-permissive bacteria were grown separately until they reached their exponential growth phase. These cultures were founded with an initial inoculum of 180 µl from fast and intermediate growers' overnight cultures or 350 µl from slow growers' overnight cultures (see **Fig. 3.10**). Fast and intermediate growers' cultures were incubated at 37°C, 250 rpm for two hours, while slow growers' cultures were incubated for 3 hours. Then, the permissive and one of the non-permissive strains were mixed in a 14 ml sterile tube at a specific ratio depending on the growth of the non-permissive strain (1:1 for fast growers, 1:2 for intermediate growers, 1:3 for slow growers) to a final volume of 5 ml.

At each transfer, $\sim 10^8$ freshly prepared susceptible cells were infected by $\sim 10^7$ phage particles (MOI$_{input}$ ~ 0.1). Infected cultures were shaken at 250 rpm, 37°C for three hours. Phage lysates were isolated as described in **Phage lysate preparation** (section **2.3.1**). Phage lysates from the previous day were used to inoculate fresh bacterial cultures the next day. Transfers continued until phages were found to infect their corresponding resistant *E. coli* C strain (section **2.3.4**) or until 21 transfers were reached. Three control lines were also run in parallel by transferring ΦX174 at MOI$_{input}$ ~ 0.1 in a culture containing *E. coli* C wildtype (permissive strain) only.

2.3.6.   Increasing the mutation rate of ΦX174 with a UV-B treatment

2.3.6.1.   Effect of UV-B radiation on ΦX174 wildtype

To determine the effects of UV-B light exposure on ΦX174 wildtype decay, three independent ΦX174 wildtype lysates were generated from three independent single plaques. 1 ml of each pure lysate ($\sim 10^9$ pfu) was transferred into a 6-well plate and submitted to UV-B light exposure at 302 nm (UVP Chemstudio Plus®, 100% intensity) during a period ranging from 1 to 25 min in total. At each exposure time (each minute), 20 µl of each phage sample was taken and serially diluted 1:10 up to 10^{-8} in PBJ. Top agar overlays were prepared by mixing 100 µl of dilutions 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, or 10^{-8} of each phage sample with 100 µl of an overnight culture of *E. coli* C wildtype in 4 ml SSA (supplemented with CaCl$_2$ and MgCl$_2$ at a final concentration of 5 and 10 mM, respectively). Top agar overlays were poured onto LB plates, dried for at least 15 minutes, and incubated inverted at 37°C for $\sim$16-17 hours. Plaques were counted the next day.

This experiment was replicated with five independent ΦX174 wildtype lysates generated from five independent single plaques. UV-B light exposure times t=0, 12, 13, and 14 min were set only. At each exposure time (each minute), 20 μl of each phage sample was taken and serially diluted 1:10 up to 10^{-8} in PBJ. Top agar overlays were prepared by mixing 100 μl of dilutions 10^{-7} or 10^{-8} of each phage sample with 100 μl of an overnight culture of *E. coli* C wildtype in 4 ml SSA (supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively). Top agar overlays were poured onto LB plates, dried for at least 15 minutes, and incubated inverted at 37˚C for ~16-17 hours. Plaques were counted the next day.

2.3.6.2.   Phage evolution experiment – Serial transfers with a UV-B treatment

ΦX174 lineages that failed to infect their corresponding *galU*, *waaG*, and *rfaH* mutants after the phage evolution experiment – Standard (section **2.3.5**) were evolved for five additional transfers on non-evolving host cultures containing a mix of (i) *E. coli* C wildtype (permissive host) and (ii) one of the resistant strains of interest (non-permissive host). For the non-permissive hosts, I used the *waaG* mutants [R9, R23, R28], the *galU* mutants [R16, R17, R22, R30, R35], and the *rfaH* mutant R25 (see **Table 3.1**).

Bacterial cultures of permissive and non-permissive were grown separately until they reached their respective exponential growth phases. These cultures were founded with an initial inoculum of 180 μl of overnight culture and incubated at 37˚C, 250 rpm for two hours. Then, permissive and non-permissive strains were mixed in a 14 ml sterile tube at a specific ratio depending on the growth of the non-permissive strain (1:1 for fast growers and 1:2 for intermediate growers) to a final volume of 5 ml.

At each transfer, ~10^8 susceptible cells were infected by ~10^8 phage particles (MOI_{input} ~1). Infected cultures were shaken at 250 rpm, 37˚C, for 3 hours. To induce mutations, 2.5 ml of each lysate was submitted to UV-B light at 302 nm (UVP Chemstudio Plus ®, 100% intensity) for 13 min (see **Fig. 4.7**). UV-B treated and untreated phage lysates were isolated as described in **Phage lysate preparation** (section **2.3.1**). Only the phage lysates that were exposed to UV-B light from the previous day were used to inoculate the fresh bacterial mixes the next day. This process was repeated until phages were found to infect their corresponding resistant *E. coli* C strains (section **2.3.4**) or until five transfers had been completed. The three control lines established during the phage evolution experiment – Standard (section **2.3.5**) were also run in parallel by transferring

ΦX174 at MOI_{input} ~1 in a culture containing *E. coli* C wildtype only, and were similarly exposed to UV-B light.

2.3.7. Preliminary phage cocktail evolution experiment

A phage cocktail made of (i) the wildtype ΦX174, (ii) all ΦX174 strains that successfully re-infected their corresponding resistant strains during the phage evolution experiment –Standard (section **2.3.5**), (iii) two ΦX174 populations that failed to re-infect their corresponding resistant strains after 21 serial transfers, and (iv) all ΦX174 UV-B irradiated populations obtained from the 5th additional serial transfers was generated by first diluting each phage lysate and mixing them at a roughly equal number of pfu (final concentration of ~ 2.10^8 pfu.ml^{-1}).

- The phage strains that successfully re-infected their corresponding resistant strains (18 in total) are namely: ΦX174 R3 T17, ΦX174 R4 T7, ΦX174 R5 T2, ΦX174 R8 T6, ΦX174 R10 T6, ΦX174 R11 T7, ΦX174 R13 T6, ΦX174 R15 T5, ΦX174 R18 T4, ΦX174 R19 T1, ΦX174 R20 T5, ΦX174 R21 T1, ΦX174 R24 T1, ΦX174 R26 T1, ΦX174 R27 T1, ΦX174 R29 T1, ΦX174 R31 T1, and ΦX174 R34 T2 (see **Tables 4.2** to **4.4**).

- The phage populations that failed to re-infect their corresponding resistant strains after 21 serial transfers (two in total) are namely: ΦX174 R2 T21 and ΦX174 R32 T21 (see **Table 4.5**).

- The UV-B irradiated phage populations obtained from the 5th additional serial transfer (12 in total) are namely ΦX174 R9 T26 UV, ΦX174 R16 T26 UV, ΦX174 R17 T26 UV, ΦX174 R22 T26 UV, ΦX174 R23 T26 UV, ΦX174 R25 T26 UV, ΦX174 R28 T26 UV, ΦX174 R30 T26 UV, ΦX174 R35 T26 UV, and the three ΦX174 control lines that grew only on *E. coli* C wildtype (ΦX174 C1 T26 UV, ΦX174 C2 T26 UV, and ΦX174 C3 T26 UV).

The phage cocktail was grown and transferred daily, four times a day, for up to four days on non-evolving host cultures containing (i) *E. coli* C wildtype, (ii) 11 host strains representative of the predicted LPS phenotypes (permissive hosts: *E. coli* C R1, R5, R8, R11, R12, R13, R15, R18, R19, R33, and R34), and (iii) one of the still-resistant strain of interest (non-permissive host). For the non-permissive host, I used R22 (*galU* mutant), R25 (*rfaH* mutant), or R28 (*waaG* mutant) (see

Tables 3.1 and **3.2**). *E. coli* C R22 and R28 were chosen as representatives of the *galU* and *waaG* mutants, respectively.

Each bacterial strain (permissive and non-permissive) was first grown separately for one hour for the fast and intermediate growers and two hours for the slow growers. After the incubation time, permissive strains were pooled at a roughly equal number of cfu in three different mixes depending on the growth categories of each bacterium (slow, intermediate, and fast growers, see **Fig. 3.10**). 714 µl of each fast grower was used to make the mix of fast growers ("fast mix" constituted of R5, R11, R12, R15, R19, R33, and R34). 1 ml of each slow grower was used to prepare the mix of slow growers ("slow mix" constituted of R1, R8, and R18. 3 ml of R1 was used instead of 1 ml). No mix was made for *E. coli* C wildtype and R13 (the only intermediate grower).

Permissive bacterial strains and mixes were independently kept in the exponential growth phase by transferring 1:4 of *E. coli* C wildtype and fast mix's volumes (1.2 ml), 1:3 of intermediate grower R13's volume (1.7 ml), and 1:2 of the slow mix's volume (2.5 ml) in fresh LB pre-heated at 37°C every hour (final volume of 5 ml). Non-permissive bacterial strains of interest were independently kept in the exponential growth phase by transferring 1:4 of R25's volume (1.2 ml) and 1:2 of R22 and R28's volumes (2.5 ml) in fresh LB pre-heated at 37°C every hour (final volume of 5 ml). To check whether the growth of the different bacteria remained consistent from one transfer to another, the OD600 values were measured before each transfer. The volume transferred at each hour was modified accordingly if the bacterial growth was too fast or too slow.

Before adding phages, all permissive bacterial strains and mixes were pooled together at a roughly equal number of cfu: 110 µl from *E. coli* C wildtype, 3.8 ml from the fast mix, 1.1 ml from *E. coli* C R13, and 1.1 ml from the slow mix (~10^8 susceptible cells present in the final mix in total). One of the resistant *E. coli* C strains of interest (non-permissive host; R22, R25 or R28) was added in excess (i.e., the number of resistant cells was comprised between ~1 and 2.10^8), for a final volume of 8 ml.

Infections were started by adding ~10^8 phages (MOI_{input} ~1) from the cocktail in the final bacterial mix. Infected cultures were shaken in 100 ml sterile Erlenmeyer at 250 rpm, 37°C, for one hour. After that, 1:15 of the volume of each infected culture was transferred to freshly mixed bacterial cultures, and the infection continued for one hour. This step was repeated for a total of four transfers. The final (fourth) transfer lasted two hours to ensure that all phages completed the

infection cycle and are present in the culture medium and not in the cell. All transfers completed on the same day involved transferring both phages and bacteria; on every fourth transfer, supernatants were collected, and only phages were transferred. Phage lysates were isolated as described in **Phage lysate preparation** (section **2.3.1**). Phage lysates from the previous day were used to inoculate the first fresh bacterial cultures on the next day. This entire process was repeated until phages that could infect their corresponding resistant *E. coli* C strains were identified (section **2.3.4**) or until four transfers had been completed.

While this method successfully yielded three evolved phages infecting respectively the *galU* mutant R22 (ΦX174 R22 T1*), the *rfaH* mutant R25 (ΦX174 R25 T1*), and the *waaG* mutant R28 (ΦX174 R28 T1*), it could not be reproduced. The evolved phages infecting respectively R3, R5, R11, R15, R19, and R34 were removed from the mutational analysis because they were either non-isogenic, adapted to a non-LPS bacterial mutant or lost for unknown reasons (impossibility of producing new phage lysates from their glycerol stocks). Furthermore, the whole-genome sequencing of the UV-B irradiated populations was missing. It made it difficult to differentiate between mutations originally present in the cocktail and *de novo* mutations due to the further adaptation of the evolved phages or potential recombination events between their genomes. The three evolved phages ΦX174 R22 T1*, ΦX174 R25 T1*, and ΦX174 R28 T1* (see **Table 5.S1**), therefore, differ from the ones obtained during the **Phage cocktail evolution experiment – Increased diversity** and **Phage cocktail evolution experiment – Increased diversity and generations** described in the subsequent sections (see sections **2.3.8** and **2.3.9**).

2.3.8.   Phage cocktail evolution experiment – Increased diversity

A phage cocktail made of (i) the wildtype ΦX174 and (ii) all isogenic ΦX174 strains that successfully re-infected their corresponding resistant strains during the phage evolution experiment – Standard was generated by first diluting each phage lysate and mixing them at a roughly equal number of pfu (final concentration of ~2.10^8 pfu.ml^{-1}). These isogenic phage strains (14 in total) are namely: ΦX174 R4 T7, ΦX174 R5 T2, ΦX174 R8 T6, ΦX174 R10 T6, ΦX174 R13 T6, ΦX174 R18 T4, ΦX174 R19 T1, ΦX174 R20 T5, ΦX174 R21 T1, ΦX174 R24 T1, ΦX174 R26 T1, ΦX174 R27 T1, ΦX174 R29 T1, and ΦX174 R31 T1 (see **Tables 4.2** and **4.3**). Despite being removed from the mutational analysis, evolved phages infecting R5 and R19, respectively, were part of the final cocktail (**Table 4.3**).

This phage cocktail was grown and transferred daily, for up to four days, on non-evolving host cultures containing (i) *E. coli* C wildtype, (ii) the 14 host strains for which resistance had been overcome (permissive hosts: *E. coli* C R4, R5, R8, R10, R13, R18, R19, R20, R21, R24, R26, R27, R29, and R31), and (iii) one of the still-resistant strain of interest (non-permissive host). For the non-permissive host, I used R6 (*waaP*/*pssA* mutant), R22 (*galU* mutant), R25 (*rfaH* mutant), or R28 (*waaG* mutant) (see **Table 3.1**). *E. coli* C R22 and R28 were chosen as representatives of the *galU* and *waaG* mutants, respectively.

Each bacterial strain (permissive and non-permissive) was first grown separately in LB liquid culture for one hour (37°C, 250 rpm). An inoculum of 180 µl from overnight cultures was used for the fast and intermediate growers, while 350 µl was used for the slow growers (see **Fig. 3.10**). After the incubation time, permissive strains were pooled at a roughly equal number of cfu for a final concentration of $\sim 10^8$ susceptible cells. One of the *E. coli* C resistant strains of interest (non-permissive host; R6, R22, R25 or R28) was then added in excess (in between ~ 1 and 5.10^8 resistant cells) in the mix of hosts to give a final volume of 4 ml.

Infections were initiated by adding $\sim 10^7$ phages (MOI$_{input}$ ~ 0.1) from the phage cocktail to the bacterial culture mix. Infected cultures were shaken in 100 ml sterile Erlenmeyer at 250 rpm, 37°C, for 5 hours. Phage lysates were isolated as described in **Phage lysate preparation** (section **2.3.1**). Phage lysates from the previous day were used to inoculate the fresh bacterial mixes the next day. This process was repeated until phages were found to infect their corresponding resistant *E. coli* C strains (section **2.3.4**) or until four transfers had been completed.

2.3.9.   Phage cocktail evolution experiment – Increased diversity and generations

The same phage cocktail used in the phage cocktail evolution experiment – Increased diversity (section **2.3.8**) was grown and transferred daily, four times a day, for up to four days on non-evolving host cultures containing (i) *E. coli* C wildtype, (ii) the 14 host strains for which resistance had been overcome (permissive hosts), and (iii) one of the still-resistant strain of interest (non-permissive host). The definitive protocol from the phage cocktail evolution experiment – Increased diversity was adjusted to reduce the amount of time necessary to retrieve the desired phages. I increased the number of transfers per day from one to four for a total of four days (16 transfers in total). Only *E. coli* C R6 (*waaP*/*pssA* mutant) and R25 (*rfaH* mutant) (see **Table 3.1**) were used as non-permissive hosts.

Each bacterial strain (permissive and non-permissive) was first grown separately for one hour. An inoculum of 180 µl from overnight cultures was used for the fast and intermediate growers, while 350 µl was used for the slow growers (see **Fig. 3.10**). After the incubation time, permissive strains were pooled at a roughly equal number of cfu in three different mixes depending on the growth categories of each bacterium (slow, intermediate, and fast growers, see **Fig. 4.9B**). 1 ml of each fast grower was used to make the "fast mix" (constituted of *E. coli* C wildtype, R5, R19, R27, and R31). 833 µl of each intermediate grower was used for the "intermediate mix" (constituted of R4, R13, R20, R21, R24, and R29). Finally, 1.25 ml of each slow grower was used to prepare the "slow mix" (constituted of R8, R10, R18, and R26).

Non-permissive strains and bacterial mixes were independently kept in the exponential growth phase by transferring 1:5 of the volume (1 ml) in fresh LB pre-heated at 37°C every hour (final volume of 5 ml). To check whether the growth of the different bacteria remained consistent from one transfer to another, the OD600 values were measured before each transfer. The volume transferred at each hour was modified accordingly if the bacterial growth was too fast or too slow.

Before adding phages, all bacterial mixes were pooled together at a roughly equal number of cfu (500 µl each, for a total of $\sim10^8$ susceptible cells), and one of the *E. coli* C resistant strains of interest (non-permissive host; R6 or R25) was added in excess (i.e., the number of resistant cells was comprised between ~1 and 5.10^8) for a final volume of 4 ml.

Infections were started by adding $\sim10^7$ phages (MOI$_{input}$ ~0.1) from the cocktail in the final bacterial mix. Infected cultures were shaken in 100 ml sterile Erlenmeyer at 250 rpm, 37°C, for one hour. After that, 1:15 of the volume of each infected culture was transferred to freshly mixed bacterial cultures, and the infection continued for one hour. This step was repeated for a total of four transfers. The final (fourth) transfer lasted two hours to ensure that all phages completed the infection cycle and are present in the culture medium and not in the cell. All transfers completed on the same day involved transferring both phages and bacteria; on every fourth transfer, supernatants were collected, and only phages were transferred. Phage lysates were isolated as described in **Phage lysate preparation** (section **2.3.1**). Phage lysates from the previous day were used to inoculate the first fresh bacterial cultures on the next day. This entire process was repeated until phages that could infect their corresponding resistant *E. coli* C strains were identified (section **2.3.4**) or until four transfers had been completed.

2.3.10. ΦX174 genome sequencing

2.3.10.1. Whole-genome re-sequencing

Sterile <u>milli-Q</u> (MQ) water was added to reconstitute each desiccated primer at a final concentration of 100 pmol.µl⁻¹ (called mother stocks of 100 µM). 10 µM working stocks were made from the mother stocks by mixing the <u>reverse</u> (rv) and <u>forward</u> (fw) primers of each corresponding set of primers (80 µl of sterile MQ water + 10 µl of the reverse primer from the mother stock at 100 µM + 10 µl of the forward primer from the mother stock at 100 µM). ΦX174 samples were prepared for whole genome re-sequencing from 1 ml of phage lysate.

Genomic DNA was extracted using the QIAprep Spin Miniprep Kit® and amplified by performing 20 cycles of PCR using Q5® High-Fidelity 2X Master Mix (NEB) (final concentration between 30-100 ng.µl⁻¹). The primers used for amplifying the genome of ΦX174 are listed in **Table 2.1**. The PCR mix and programs used for the amplification are described in **Table 2.2**. All PCR products were cleaned using the QIAquick PCR Purification Kit® (QIAGEN). The result of each PCR was visualized by loading a mix of 5 µl of PCR products with 2 µl of DNA gel loading dye 6X (ThermoFisher) on an electrophoresis gel 1% agarose (1.5 g of agarose; 150 ml of TAE 1X buffer; 5 µl of SYBR® Green) at 120V, 150A for one hour.

DNA samples were tested for quality, pooled, and sequenced by the Max-Planck Institute for Evolutionary Biology (Plön, Germany). Sequencing was performed using an Illumina MiSeq DNA Flex Library Prep Kit to produce 150 bp paired-end reads[100]. The quality of the sequencing output was controlled using *FastQC* version 0.11.8 (**Appendix. A1**). Reads were trimmed using *Trimmomatic* (**Appendix. A1**), assembled and analysed using the *breseq* pipeline version 0.33.2[101–103] (**Appendix. A2**) and Geneious Prime (version 2020.1.2).

Name	Sequence (5'→3')	Binding Region	Tm 2Q5	Ta 2Q5	Purpose
ΦX174_2361_rv	TCGCTTGGTCAACCCCTCAG	2342 - 2361	70°C	71°C	ΦX174 whole genome amplification
ΦX174_2362_fw	AGCGCGGTAGGTTTTCTGCT	2362 - 2382	70°C		
ΦX174_399_rv	CTTGACTCATGATTTCTTACC	379 - 399	58°C	59°C	ΦX174 whole genome amplification
ΦX174_400_fw	TTACTGAACAATCCGTACGTTTC	400 - 422	62°C		

Name	Sequence (5'→3')	Binding Region	Tm Phusion (buffer HF)	Ta Phusion (buffer HF)	Purpose
ΦX174_F_fw	CGCTCGTCTTTGGTATGTAGGTGG	928 - 951	65°C	68°C	Amplification of ΦX174's *F* gene
ΦX174_F_rv	AGCGGCGTTGACAGATGTATCC	2580 - 2601	65°C		Amplification of ΦX174's *F* gene
ΦX174_F_fw2	GTTGCGAGGTACTAAAGGCAAGCG	898 - 921	66°C	68°C	Amplification of ΦX174's *F* gene
ΦX174_F_rv2	CTGGAGTAACAGAAGTGAGAACCAGC	2245 - 2470	65°C		Amplification of ΦX174's *F* gene
ΦX174_F_fw_sh	CGTCTTTGGTATGTAGGTGG	932 - 951	57°C	61°C	Amplification of ΦX174's *F* gene
ΦX174_F_rv_sh	CGGCGTTGACAGATGTATC	2581 - 2599	58°C		Amplification of ΦX174's *F* gene
ΦX174_H_fw	CAAGGACTGTGTGACTATTGACGTCC	2763 - 2788	65°C	68°C	Amplification of ΦX174's *H* gene
ΦX174_H_rv	GCCTCTACGCGATTTCATAGTGGAG	4490 - 4514	65°C		Amplification of ΦX174's *H* gene
ΦX174_H_fw2	CGCTACTAAATGCCGCGGATTGG	2847 - 2869	66°C	67°C	Amplification of ΦX174's *H* gene
ΦX174_H_rv2	GCCTCCAGCAATCTTGAACACTC	4467 - 4489	64°C		Amplification of ΦX174's *H* gene
ΦX174_H_fw_sh	GACTGTGTGACTATTGACGTCC	2767 - 2788	60°C	61°C	Amplification of ΦX174's *H* gene
ΦX174_H_rv_sh	GCCTCTACGCGATTTCATAG	4495 - 4514	58°C		Amplification of ΦX174's *H* gene

Name	Sequence (5'→3')	Binding Region	Tm	Ta	Purpose
ΦX174_F_Sanger_fw	CCTCATCGTCACGTTTATGG	1217 - 1236	58°C	56°C	Sanger sequencing of ΦX174's *F* gene
ΦX174_F_Sanger_fw2	ACCGATATTGCTGGCGAC	1934 - 1951	59°C	56°C	Sanger sequencing of ΦX174's *F* gene
ΦX174_F_Sanger_rv	GGCGTTATAACCTCACACTC	2294 - 2313	58°C	56°C	Sanger sequencing of ΦX174's *F* gene
ΦX174_F_Sanger_rv2	GGCGAGAAAGCTCAGTCTC	1529 - 1547	60°C	56°C	Sanger sequencing of ΦX174's *F* gene
ΦX174_H_Sanger_fw	GTGGCGCCATGTCTAAATTG	2974 - 2993	59°C	56°C	Sanger sequencing of ΦX174's *H* gene
ΦX174_H_Sanger_fw2	CAGCAAACGCAGAATCAGC	3720 - 3738	60°C	56°C	Sanger sequencing of ΦX174's *H* gene
ΦX174_H_Sanger_rv	CCACAAGCCTCAATAGCAG	4054 - 4072	58°C	56°C	Sanger sequencing of ΦX174's *H* gene
ΦX174_H_Sanger_rv2	CTTTATCAGCGGCAGACTTG	3251 - 3270	58°C	56°C	Sanger sequencing of ΦX174's *H* gene

Table 2.1. List of all primers used in this study. *F*: encodes for the viral capsid protein. *H*: encodes for the minor spike protein. "fw": forward. "rv": reverse. Tm 2Q5: melting temperature of the primer used with the Q5® High-Fidelity 2X Master Mix (NEB). Ta 2Q5: annealing temperature of the set of primers used with the Q5® High-Fidelity 2X Master Mix (NEB). Tm Phusion (buffer HF): melting temperature of the primer used with Phusion® High-Fidelity PCR Master Mix with HF Buffer (ThermoFisher). Ta Phusion (buffer HF): annealing temperature of the primer used with Phusion® High-Fidelity PCR Master Mix with HF Buffer (ThermoFisher).

PCR Mix	Volume (μl) for one sample
2Q5 MM	25.0
10 μM mix Primers	2.5
gDNA Template	2.0
Water	20.5

Program PCR set 2361_rv/2362_fw		
Temperature (˚C)	Time	
98	30 seconds	
98	10 seconds	
71	20 seconds	Cycle 20 times
72	4 minutes	
72	2 minutes	

Program PCR set 399_rv/400_fw		
Temperature (˚C)	Time	
98	30 seconds	
98	10 seconds	
59	20 seconds	Cycle 20 times
72	4 minutes	
72	2 minutes	

Table 2.2. PCR mix and programs used for the whole genome amplification of ΦX174. "2Q5 MM": Q5® High-Fidelity 2X Master Mix (NEB). "fw": forward. "rv": reverse.

2.3.10.2. Sanger sequencing of ΦX174's *F* and *H* genes

Sterile milli-Q (MQ) water was added to reconstitute each desiccated primer at a final concentration of 100 pmol.μl^{-1} (called mother stocks of 100 μM). 10 μM working stocks were made from the mother stocks by mixing the reverse (rv) and forward (fw) primers of each corresponding set of primers (80 μl of sterile MQ water + 10 μl of the reverse primer from the mother stock at 100 μM + 10 μl of the forward primer from the mother stock at 100 μM). ΦX174 samples were prepared for whole genome re-sequencing from 1 ml of phage lysate.

Genomic DNA was extracted using the QIAprep Spin Miniprep Kit® (QIAGEN) (final concentration between 30-100 ng.μl^{-1}). ΦX174's *F* and *H* genes were amplified by performing 35 cycles of PCR using Phusion® High-Fidelity PCR Master Mix with HF Buffer (ThermoFisher). Primers used for this purpose are listed in **Table 2.1**. The PCR mix and programs used for amplifying ΦX174's *F* and *H* genes are described in **Table 2.3**. All PCR products were cleaned

using the QIAquick PCR Purification Kit®. The result of each PCR was visualized by loading a mix of 5 µl of PCR products with 2 µl of DNA gel loading dye 6X (ThermoFisher) on an electrophoresis gel 1% agarose (1.5 g of agarose; 150 ml of TAE 1X buffer; 5 µl of SYBR® Green) at 120V, 150A for one hour.

PCR Mix	Volume (µl) for one sample
Phusion HF MM	25.0
10 µM mix Primers	2.5
gDNA Template	2.0
Water	20.5

Program PCR		
Temperature (˚C)	**Time**	
98	30 seconds	
98	10 seconds	
Ta Phusion of each set of primer (see **Table 2.1**)	20 seconds	Cycle 35 times
72	4 minutes	
72	2 minutes	

Table 2.3. PCR mix and program used for the amplifying ΦX174's *F* and *H* genes of. "2Q5 MM": Q5® High-Fidelity 2X Master Mix (NEB). "fw": forward. "rv": reverse. "Ta Phusion": annealing temperature of the primer used with Phusion® High-Fidelity PCR Master Mix with HF Buffer (ThermoFisher)

Sanger sequencing was performed using the sequencing primers listed in **Table 2.1**. Reactions were set up in 96 well plates (on ice) using the cleaned PCR products. The content of the reaction mixes and program used are listed in **Table 2.4**. The X-Terminator clean-up and sequencing steps were performed at the Max-Planck Institute for Evolutionary Biology (Plön, Germany). Sequencing results were assembled and analysed with Geneious Prime (version 2020.1.2).

Content	Volume (µl) for one sample
Big Dye Terminator (BD)	0.5
5x Big Dye buffer (SD)	2.0
HPLC water	5.5
Primer (5 µM)	1.0
Purified PCR products	1.0
Total volume	10.0

Program Sanger		
Temperature (˚C)	**Time**	
96	1 minute	
96	10 seconds	
56	15 seconds	Cycle 30 times
60	4 minutes	
12	unlimited	

Table 2.4. Reaction mixes and program used for the Sanger sequencing of ΦX174's *F* and *H* genes.

2.4.   Methods Chapter 5

2.4.1.   Determining *E. coli* C strains' susceptibility to evolved ΦX174

2.4.1.1.   Spotting assays in a semi-solid environment

Method 1. Each top agar overlay was prepared by mixing 200 µl of an *E. coli* C strain from stationary phase culture in 4 ml SSA (supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively), then poured on LB plates and dried for at least 15 minutes under a sterile laminar flow hood. Then, 3 µl of each undiluted evolved phage lysate (between 10^7 and 10^9 pfu.ml^{-1}) was dropped at the surface.

Method 2. Each top agar overlay was prepared by mixing a volume of each phage lysate in 4 ml SSA (supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively) at a final concentration of $\sim 10^7$ pfu.ml^{-1}, then poured on LB plates and dried for at least 15 minutes under a sterile laminar flow hood. Then, 3 µl of both undiluted and ten-fold diluted of each *E. coli* C strain (from overnight culture) were dropped onto the surface with a pipette.

For both methods, spots were dried for at least 30 minutes under sterile laminar flow, and the plates were subsequently incubated inverted at 37°C for ~17 hours. *E. coli* C wildtype was used as a positive control (permissive strain), and *E. coli* K-12 MG1655 as a negative control (non-permissive strain). Sterile H_2O was also spotted at the end of each plate as a control for material contamination (pipettes and tips). A bacterial host strain was classified as sensitive only when signs of lysis were detected using both methods in at least two (of three) replicates per method. A phage-bacterium combination that yielded different results between the two methods was tested in *standard plaque assays* as follows: a top agar overlay was prepared by mixing 100 μl of the phage lysate diluted to 10^{-1} and 10^{-6} with 200 μl of the resistant *E. coli* C strain from stationary phase culture in 4 ml SSA (supplemented with $CaCl_2$ and $MgCl_2$ at a final concentration of 5 and 10 mM, respectively), then poured on LB plates and dried for at least 15 minutes. The presence or absence of plaques was assessed the next day. Finally, the host strain was classified as sensitive if plaques were observed at both dilutions.

2.4.1.2. Solid environment assays

To measure the effect of media's consistency on bacterial susceptibility to phage infection, the protocol of *Method 1* in section **2.4.1.1** was adapted for solid agar plates (1.5% agar). Supplemented agar plates (1.5 % agar supplemented with $CaCl_2$ and $MgCl_2$ (at a final concentration of 5 mM and 10 mM, respectively) were prepared as described in section **2.1.2** and used for this experiment. 200 μl of an *E. coli* C strain from stationary phase culture were plated with sterile glass beads on the solid LB plates and dried for at least 30 minutes under a sterile laminar flow hood. 3 μl of each undiluted high titre evolved phage lysate (between 10^7 and 10^9 pfu.ml^{-1}) was spotted on the surface. Spots were dried for 30 minutes under a sterile laminar flow, and the plates were subsequently incubated inverted at 37°C for ~17 hours. *E. coli* C wildtype was used as a positive control (permissive strain), and *E. coli* K-12 MG1655 as a negative control (non-permissive strain). Sterile H_2O was also spotted at the end of each plate as a control for material contamination (pipettes and tips). A bacterial host strain was classified as sensitive only when signs of lysis were detected. Only one replicate was performed for this experiment.

2.4.1.3. Liquid environment assays

Bacterial susceptibility to phage infection was also measured in a well-mixed, liquid environment. Three independent cultures of each *E. coli* C strain were grown independently at 37°C, 250 rpm

for 17 hours. 9 μl of each *E. coli* C strain stationary phase culture was then mixed in a 96 well plate (one per *E. coli* C strain) with 151 μl of LB supplemented with $CaCl_2$ and $MgCl_2$ (at a final concentration of 5 and 10 mM, respectively), then incubated in a shaking incubator (New Brunswick™ Innova® 44) at 250 rpm, 37°C, for one hour. 3 μl of each undiluted high titer evolved phage lysate (between 10^7 and 10^9 pfu.ml⁻¹) was inoculated in the wells containing one of the resistant strains and shaken at 250 rpm, 37°C, for 5 hours. The OD600 values were measured using an Epoch2 microplate reader (BioTek®) before the addition of phages and at the end of the experiment. Each strain was also grown in a well without phages as a negative control. Sterile H_2O was also spotted at the end of each plate as a control for material contamination (pipettes and tips). A bacterial host strain was classified as sensitive only when signs of lysis were detected (i.e., a decrease of the OD600 values compared to the ones of the negative controls without phage) were detected in at least two (of three) replicates.

2.4.2. Generating *E. coli* C strains with resistance to a cocktail of evolved ΦX174

A phage cocktail made of (i) the wildtype ΦX174 and (ii) all ΦX174 strains (c1 clones) that successfully re-infected their corresponding resistant strains during the phage evolution experiment – Standard was generated by first diluting each phage lysate and mixing them at a roughly equal number of pfu (final concentration of ~2.10^8 pfu.ml⁻¹). These phage strains (21 in total) are namely: ΦX174 R3 T17, ΦX174 R4 T7, ΦX174 R5 T2, ΦX174 R8 T6, ΦX174 R10 T6, ΦX174 R11 T7, ΦX174 R13 T6, ΦX174 R15 T5, ΦX174 R18 T4, ΦX174 R19 T1, ΦX174 R20 T5, ΦX174 R21 T1, ΦX174 R22 T1*, ΦX174 R24 T1, ΦX174 R25 T1*, ΦX174 R26 T1, ΦX174 R27 T1, ΦX174 R28 T1*, ΦX174 R29 T1, ΦX174 R31 T1, and ΦX174 R34 T2 (see **Tables 4.2** to **4.5**, **Table 5.S1**). Despite being removed from the mutational analysis, evolved phages infecting respectively R3, R5, R11, R15, R19, and R34 were part of the final cocktail. The three evolved phages ΦX174 R22 T1*, ΦX174 R25 T1*, and ΦX174 R28 T1* (see **Table 5.S1**) differ from the ones retrieved during the phage cocktail evolution experiments – Increased diversity (section **2.3.8**) and – Increased diversity and generations (section **2.3.9**).

Eleven *E. coli* C strains resistant to the phage cocktail infection were generated and stored following the protocol described in **Isolation and storage of ΦX174-resistant *E. coli* C strains** (section **2.2.1**). Eleven additional *E. coli* C resistant strains to wildtype ΦX174 infection were also generated as controls. All new *E. coli* C mutant strains were prepared for whole genome re-sequencing and analysed as described in ***E. coli* C whole genome re-sequencing** (section **2.2.5**).

2.4.3. Infection dynamic of a phage cocktail growing on *E. coli* C wildtype

To understand the growth dynamic of a phage cocktail on a wildtype *E. coli* C culture, a phage cocktail was made by first diluting the three evolved phages ΦX174 R22 T1*, ΦX174 R25 T1*, and ΦX174 R29 and mixing them at a roughly equal number of pfu (final concentration of ~10^9 pfu.ml^{-1}). The two evolved phages ΦX174 R22 T1* and ΦX174 R25 T1* (see **Table 5.S1**) differ from the ones retrieved during the phage cocktail evolution experiments – Increased diversity (section **2.3.8**) and – Increased diversity and generations (section **2.3.9**).

Five independent cultures of *E. coli* C wildtype were grown at 37°C, 250 rpm for 17 hours. 200 μl of each overnight culture were inoculated in 4.8 ml LB pre-heated at 37°C supplemented with CaCl$_2$ and MgCl$_2$ (at a final concentration of 5 and 10 mM, respectively). Incubation was set up for 2 hours at 250 rpm, 37°C until the OD600 values reached ~1.2 – 1.3 (using an Ultrospec 10 (Biochrom®)). Infections started by adding either ~10^7 pfu of ΦX174 wildtype alone or ~10^7 pfu of the phage cocktail to each host culture (MOI$_{input}$ ~0.1). Infected cultures were shaken at 37°C, 280 rpm (double orbital) for 24 hours. The OD600 values were monitored with an Epoch2 microplate reader BioTek® every 10 minutes. The five independent cultures of *E. coli* C wildtype were also grown without the presence of phages as negative controls.

Chapter III – Investigating ΦX174-resistance pathways in *E. coli* C

3.1. Structure, biosynthesis, and regulation of the core OS LPS

3.1.1. LPS structures in *E. coli* strains

Extensive efforts have been made to understand the phenotypic diversity of lipopolysaccharide (LPS) structures through deletion and complementation experiments, and the functions of genes involved in synthesizing and assembling bacterial lipopolysaccharide (LPS) have been eluciated[106–110]. *E. coli* strains produce LPS molecules composed of three parts: the lipid A, the core oligosaccharide (OS), and the O-antigen chain. The LPS core OS itself is composed of two elements: (i) a very conserved inner core across *E. coli* species and (ii) a variable outer core that can be categorized into five types (depending on the hexose composition)[110]. Strains with the tripartite LPS molecules (lipid A, core OS, and O-antigen) are termed "smooth" types[107].

Some *E. coli* strains, however, produce LPS molecules that lack one or more components. LPS molecules lacking the O-antigen are classified as "rough", and molecules lacking both O-antigen and the outer core OS as "deep rough"[110,111,106,112–114]. *E. coli* strain C cannot produce an O-antigen due to an *IS3* insertion in the *wzzB* gene (O-antigen chain length regulator)[115]. As a result, *E. coli* C harbours a naturally truncated "rough" type LPS structure (**Fig. 3.1A**).

3.1.2. Genetics and assembly of the core OS LPS in *E. coli* C

Three major operons for the assembly of the core OS are encoded by the chromosomal locus *waa* (formerly named *rfa*) in the *E. coli* C genome: *waaA, gmhD* and *waaQ*[113] (**Fig. 3.1A**). The *waaA* operon is responsible for incorporating two Kdo (3-deoxy-D-*manno*-octulosonic acids) moieties into lipid A[116]. The *gmhD* operon is required for assembling the inner core's backbone[106], and the *waaQ* operon contains all eight genes necessary to modify the inner core and build the outer core[113] (**Fig. 3.1A**).

Inner core assembly and modification. Inner core assembly starts first with WaaA adding the two Kdo moieties to the lipid-A-IV precursor[110,106,113]. A bacterium lacking WaaA activity displays severe

membrane defects, resulting in the inability to form colonies[117]. Heptose residues I and II are then anchored to the Kdo moieties by WaaC and WaaF, respectively[106,113,118]. LPS core heptose kinases WaaP and WaaY add phosphate groups to the first and second heptose residues, respectively. WaaQ adds the last heptose to the second heptose residue[106,113,118]. *WaaY* cannot work without the activity of *waaQ*, which cannot work without *waaP* (first red ellipse in **Fig. 3.1B**).
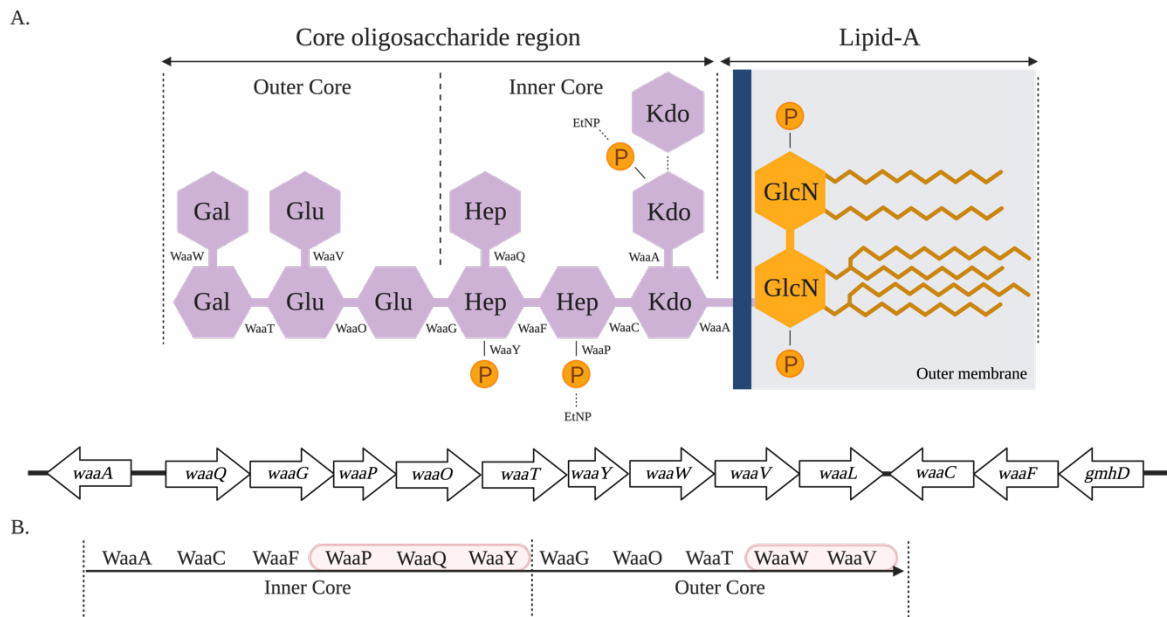
Fig. 3.1. Assembly of *E. coli* C rough type LPS. A. Organization of the chromosomal *waa* region in *E. coli* C wildtype strain. The chromosomal *waa* locus is organized into three major operons, usually designated by the first gene of each transcriptional unit: *waaA*, *waaQ* and *gmhD*[113]. **B.** Sequential steps of the core OS LPS synthesis (from WaaA to WaaV). The first red ellipse indicates that WaaY requires the prior functioning of WaaQ, which requires the prior functioning of WaaP. The second red ellipse indicates that WaaV requires the prior functioning of WaaW. WaaP, WaaQ, WaaY, and WaaV are not necessary to complete the core LPS backbone. WaaA: 3-deoxy-D-manno-octulosonic acid transferase; WaaC: ADP‑heptose‑‑LPS heptosyltransferase I; WaaF: ADP‑heptose‑‑LPS heptosyltransferase II; WaaP: Lipopolysaccharide core heptose(I) kinase; WaaQ: ADP‑heptose‑‑LPS heptosyltransferase III; WaaY: Lipopolysaccharide core heptose(II) kinase; WaaG: UDP-glucose:(heptosyl) LPS α1,3-glucosyltransferase (glucosyltransferase I); WaaO: UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase; WaaT: UTP‑‑glucose‑1‑phosphate uridylyltransferase; WaaW: UDP‑galactose‑‑(galactosyl) LPS alpha1,2-galactosyl-transferase; WaaV: UDP-glucose:(Glucosyl) LPS β1,3-Glucosyltransferase; WaaL: O-antigen ligase; GmhD: ADP-L-glycero-D-mannoheptose 6-epimerase.

Outer core assembly. The glucosyltransferases WaaG and WaaO start the formation of the outer core by anchoring the first glucose to the second heptose residue and the second glucose to the first glucose residue, respectively[119,120,113]. It has been suggested that the reaction catalyzed by WaaG might be required for WaaP and WaaY substrate specificity[121]. Then, the galactosyltransferases WaaT and WaaW add the first galactose to the second glucose residue and the second galactose to the first galactose residue, respectively[113,122]. The outer core LPS is complete when WaaV adds the third glucose to the second glucose residue. Here, *waaV* requires the activity of *waaW* and proceeds in this specific order (second red ellipse in **Fig. 3.1B**). If *waaW* is deleted, *waaV* is not functional[122,123].

3.1.3. Biosynthesis of hexose residues incorporated into the core OS

Heptose synthesis. Heptose residues are synthesised via the ADP-L-glycero-β-D-*manno*-heptose synthesis pathway[124]. It involves four genes: *gmhA*, *hldE*, *gmhB* and *gmhD* (**Fig. 3.2A**). Each gene encodes a protein catalysing the production of a specific heptose intermediate. The final product is an ADP-L-glycero-β-D-*manno*-heptose first used by the heptosyltransferase WaaC to build the inner core LPS[124]. Thus, the deletion of one of these genes is expected to result in a heptoseless, deep rough LPS[125] (**Fig. 3.2A**).

Glucose and galactose synthesis. Both glucose and galactose residues that are required to construct the outer core are produced via the Leloir pathway (Galactose degradation I)[126,127]. GalU catalyses the formation of UDP-α-D-glucose from α-D-glucopyranose 1- phosphate[124–126,128] (**Fig. 3.2B**). UDP-α-D-glucose can then be incorporated into the outer core LPS by WaaG. Therefore, Δ*galU* and Δ*waaG* mutants result in the same truncated LPS structure[129,106,128]. GalE catalyses the conversion of UDP-α-D-glucose to UDP-α-D-galactose[130]. UDP-α-D-galactose is then incorporated into the outer core LPS by WaaT[122]. Therefore, both Δ*galE* and Δ*waaT* are expected to result in the same truncated LPS[131] (**Fig. 3.2B**).
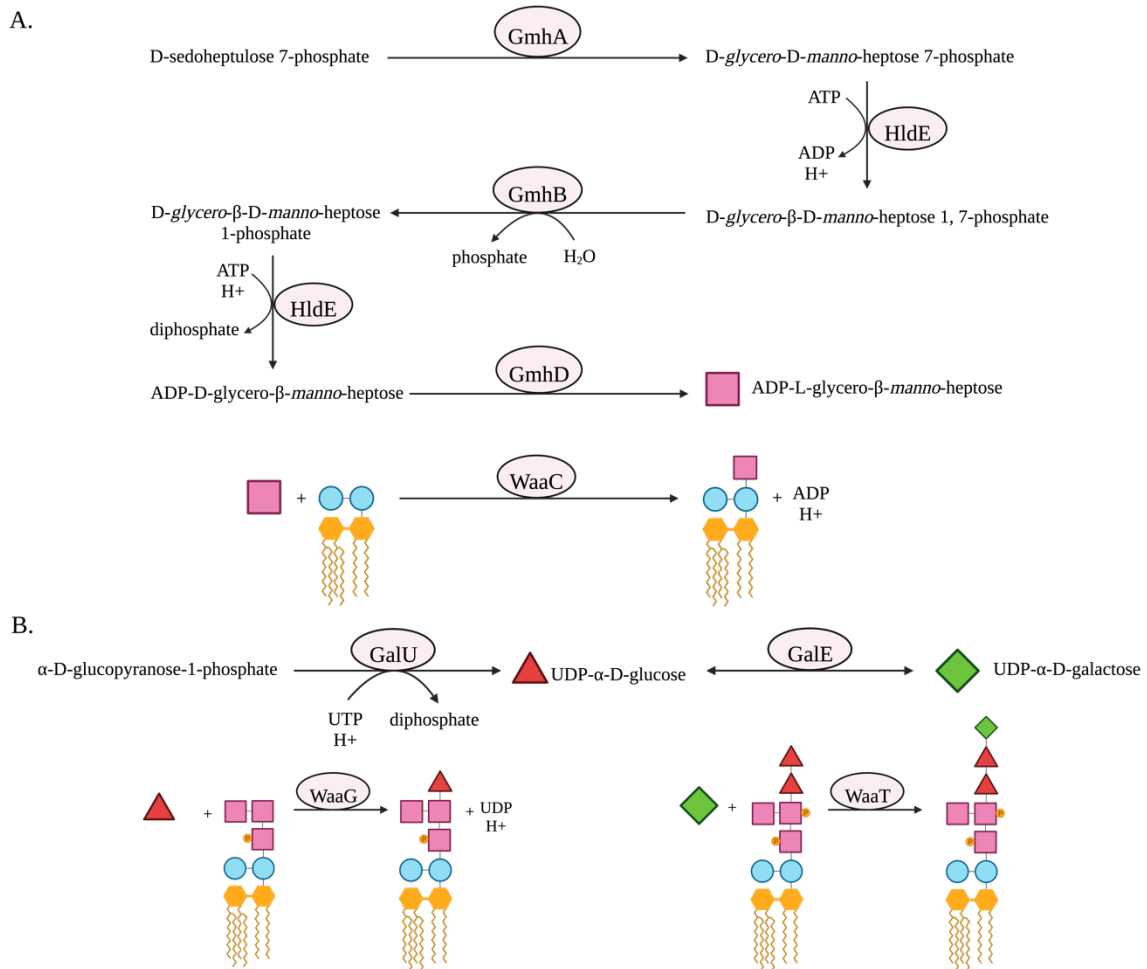
Fig 3.2. Biosynthesis of the core LPS heptose, glucose, and galactose sugar components. **A.** ADP-L-*glycero*-β-D-*manno*-heptose biosynthesis pathway. This figure is adapted from https://biocyc.org/ECOLI/NEW-IMAGE?type=PATHWAY&object=PWY0-1241[132]. GmhB: D‑glycero‑beta‑D‑manno‑heptose 1,7‑bisphosphate 7‑phosphatase; HldE: Bifunctional heptose 7‑phosphate kinase/heptose 1‑phosphate adenyltransferase; GmhD: ADP-L-glycero-D-mannoheptose 6-epimerase; GmhA: D-sedoheptulose 7-phosphate isomerase. **B.** Leloir Pathway (Galactose degradation I). This figure is adapted from https://biocyc.org/ECOLI/NEW-IMAGE?type=PATHWAY&object=GALACTMETAB-PWY[132]. GalU: UTP‑‑glucose‑1‑phosphate uridylyltransferase; GalE: UDP‑glucose 4‑epimerase.

3.1.4. Control of core LPS biosynthesis – the *rfaH* gene

RfaH participates in the control and regulation of several genes involved in the biosynthesis and assembly of the core LPS[109]. The *rfaH* gene encodes a transcriptional antiterminator homologous to *nusG*, an essential transcription elongation factor[133]. RfaH is recruited by an *ops* (operon polarity

suppressor) element and binds to a short 8-nt conserved motif 5'-GGCGGTAG-3' in *ops'* pause site[109,133]. The recruitment of RfaH by the *ops* element (and potentially other factors) to the RNA polymerase complex modifies and increases the processivity – "the number of nucleotides polymerized by a polymerase during a single association-dissociation cycle with the primer-template"[134] – of the complex, allowing transcription elongation to proceed over long operons[133]. A loss of function in *rfaH* has immediate repercussions for the transcription of the major LPS operons, resulting in a truncation of the LPS in Δ*rfaH* mutants. The resulting LPS phenotypes are often characterised by a decrease in outer membrane stability, comparable to a deep rough LPS phenotype[135].

3.2. Exploration of *E. coli* C's resistance strategies to ΦX174

In this chapter, I investigate the diversity of *E. coli* C's evolutionary pathways to become resistant to ΦX174. ΦX174 solely relies on the core LPS structure to attach to and infect its host. In return, the bacterium can quickly evolve resistance via modification of its core LPS structure to avoid phage adsorption. However, modifying the LPS structure is not without consequences for the cells and leads to various, often deleterious, pleiotropic effects. Maintenance of the native LPS structure is, therefore, crucial for the bacterium's survival[94]. Yet, we know very little about LPS structural diversity since our current knowledge exclusively derives from gene deletion and complementation experiments.

I generated spontaneous *E. coli* C mutants resistant to wildtype ΦX174 infection on solid agar plates. I isolated single colonies and identified the mutational causes responsible for ΦX174 resistance. I found that a single mutation in a gene involved in LPS biosynthesis, assembly or regulation is sufficient to alter core LPS structure and prevent phage infection. For each resistant strain, I predicted an LPS phenotype based on the current LPS model of one gene-one phenotype. My predictions suggest that many, but not all, LPS modifications can lead to phage resistance.

3.3. A Luria-Delbrück approach to generate LPS mutants

Phage resistance mechanisms have been studied mostly via "top-down" approaches using co-evolution experiments[83,136–140] in which the experimentalist propagates a phage on a permissive host for many generations[141,142]. The evolving populations (bacteria and phages) are usually cultured in a liquid environment, either continuously in a chemostat or serially transferred in vials. However, experimental evolution in well-mixed liquid environments strongly selects for fast-growing bacteria

that can quickly outcompete other resistant but slow-growing bacteria. Thus a wide variety of bacterial defence mechanisms associated with slow growth are understudied despite their relevance in natural environments where resources are limited or fast growth is disadvantageous[143].

To avoid competition between phage-resistant bacterial strains, I pursued a "bottom-up" approach. I generated ΦX174-resistant *E. coli* C strains on solid agar plates, similar to a Luria-Delbrück fluctuation experiment[62] (see **Chapter II** section **2.2.1**). In their original research, Luria and Delbrück elegantly showed that mutation(s) conferring resistance to phage infection arose before any encounter with it (pre-exposure mutational events)[62]. During the overnight growth of bacteria in a liquid non-selective environment, mutations occur during DNA replication at random in the genome (coloured lightning bolts in **Fig. 3.3**). By chance, one of these mutations may cause phage resistance (**Fig. 3.3**).
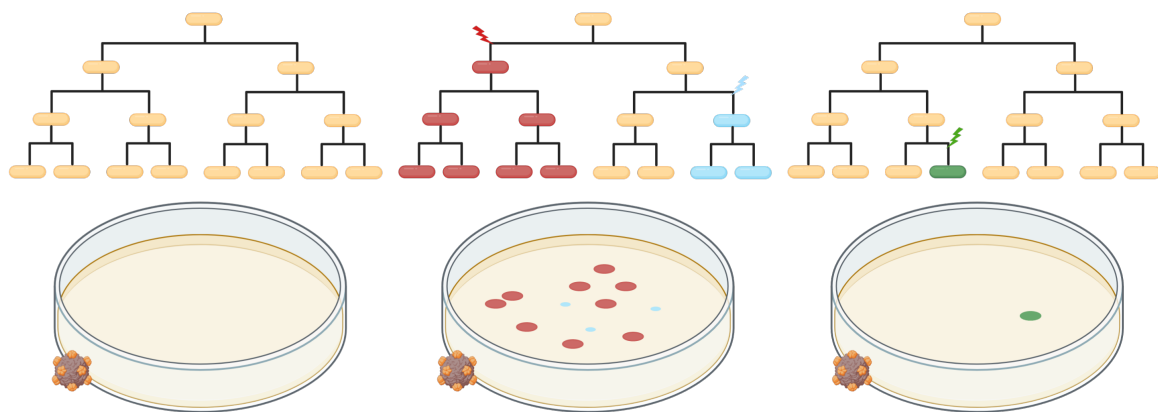
Fig. 3.3. Pre-exposure mutational events. The genealogy shows at which bacterial generation the mutation (represented by a coloured lightning bolt) occurs in the absence of ΦX174. If the mutational event occurs early during replication and in a gene relevant for phage resistance ("jackpot" mutational event), a bigger proportion of offspring will inherit the mutation and survive phage exposure. Therefore, the distribution of the number of resistant colonies varies greatly from one replicate to another, depending on which generation the mutation initially occurred.

The spatial structure of the solid LB plates creates niches where the different resistant strains can grow separately from each other. Phage-resistant strains of *E. coli* C with a wide range of LPS and growth phenotypes can co-exist without competing for space and resources. They can be isolated as long as they form visible colonies on the agar plates.

3.4.   Mutational causes of ΦX174 resistance in *E. coli* C

Using the approach described above, I generated 35 spontaneous ΦX174-resistant *E. coli* C strains, extracted their genomic DNA and used whole genome re-sequencing methods to identify the mutation(s) conferring resistance to ΦX174 infection (see **Chapter II** section **2.2.5**). In this section, however, I am only presenting whole-genome re-sequencing results of 31 *E. coli* C resistant strains (**Table 3.1**). After careful investigations, I removed four resistant strains (*E. coli* C R1, R3, R15, and R19; **Table 3.2**) from both mutational and phenotypical analyses. The reasons behind these exclusions are explained below.

3.4.1.   Exclusion of four *E. coli* C resistant strains from the dataset

I found that both *E. coli* C R3 and R15 are not isogenic. While whole genome re-sequencing from their respective glycerol stocks showed only a single mutation in the *galE* gene, whole genome re-sequencing of ten re-streaked colonies showed that additional mutations were systematically associated with the single mutation in *galE* (**Table 3.3**). Loss of function in *galE* should, in theory, lead to a partial truncation of the outer core part of the LPS (similar to the deletion of *waaT*)[131], but surprisingly, some of these additional mutations occurred in genes involved in the LPS biosynthesis or assembly (*waaT*, *waaW*, *galU*, and *gmhD*; **Table 3.3**).

Similarly, the *galE* mutation found in *E. coli* C R1 is associated with the *waaV* gene (responsible for the completion of the outer core LPS) and does not seem to lead to resistance when mutated[131] (**Table 3.2**). R1 phenotype was unstable in a semi-solid environment. I performed phenotypic assays by plating top agar overlays from independent overnight cultures of *E. coli* C R1 without phage (**Chapter II** section **2.4.1.1**). All overnight cultures were started with a randomly picked single colony obtained from the glycerol stock of R1. I observed that colonies of R1 made different lawn types: either "smooth" (no bacterial aggregate) or "granulous" (presence of numerous bacterial aggregates, **Fig. 3.4**), which could potentially impact phage infectivity.

E. coli C resistant strains	Genes	Locus tags	Synonyms	Descriptions	Positions	Nucleotide changes	Amino Acid changes	Predicted LPS structure phenotypes	"Easy" or "Hard" LPS structure phenotypes?	# Experiment during which resistance was overcome or cross-infected
R2	*waaW*	B6N50_00430 →	-	UDP-galactose--(galactosyl) LPS alpha1,2-galactosyltransferase	79,779	g→a	G267D	Rough	Easy	1
R4	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	76,970	(a)$_{7→6}$	L249fsTer#263[a]	Rough	Easy	1
R5	*gmhB*	B6N50_18960 ←	-	D-glycero-beta-D-manno-heptose 1,7-bisphosphate 7-phosphatase	3,693,747	(t)$_{8→7}$	M45fsTer#51	Deep rough	Easy	1
R6	*waaP*	B6N50_00410 →	*rfaP*	Lipopolysaccharide core heptose(I) kinase	76,163 - 76,168	Δ6 bp	Δ(A251-A253)	Deep rough	Hard	3
	pssA	B6N50_06070 ←	-	Phosphatidylserine synthase	1,190,560	t→g	V54V			
R7	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	77,219 - 77,224	(ttattt)$_{2→1}$	Δ(F332-L334)	Rough	Easy	1
R8	*hldE*	B6N50_03700 →	*rfaE/waaE*	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	721,891 - 721,892	(gc)$_{5→4}$	A79fs_Ter#83	Deep rough	Easy	1
R9	*waaG*	B6N50_00405 →	*rfaG*	UDP-glucose:(heptosyl) LPS α1,3-glucosyltransferase (glucosyltransferase I)	75,073 - 75,076	Δ4 bp	G260fs_Ter#266	Deep rough	Hard	2
R10	*hldE*	B6N50_03700 →	*rfaE/waaE*	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	721,891 - 721,892	(gc)$_{5→4}$	A79fs_Ter#83	Deep rough	Easy	1
	CRISPR	B6N50_05180 → / ← B6N50_05185	-	Repeat region	1,023,697 - 1,023,940	Δ244 bp	Non-coding region			

R11	*waaP*	B6N50_00410 →	*rfaP*	Lipopolysaccharide core heptose(I) kinase	75,562	g→a	W51stop	Deep rough	Easy	1
R12	*waaF*	B6N50_00450 ←	*rfaF*	ADP-heptose--LPS heptosyltransferase II	84,376	g→a	Q30stop	Deep rough	Easy	1
R13	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	76,255	Δ1 bp	M11fsTer#58	Rough	Easy	1
R14	*waaF*	B6N50_00450 ←	*rfaF*	ADP-heptose--LPS heptosyltransferase II	84,081	c→t	W128stop	Deep rough	Easy	1
R16	*galU*	B6N50_13325 ←	-	UTP--glucose-1-phosphate uridylyltransferase	2,611,994	Δ1 bp	V37fsTer#41	Deep rough	Hard	2
R17	*galU*	B6N50_13325 ←	-	UTP--glucose-1-phosphate uridylyltransferase	2,611,409	(g)$_{4\to3}$	P232fsTer#248	Deep rough	Hard	2
R18	*waaC*	B6N50_00445 ←	*rfaC*	ADP-heptose--LPS heptosyltransferase I	83,271	Δ1 bp	H48fsTer#71	Deep rough	Easy	1
R20	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	77,104	g→a	W294stop	Rough	Easy	1
R21	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	76,978	c→a	S252stop	Rough	Easy	1
	potH	B6N50_15295 ←	-	Spermidine/putrescine ABC transporter permease	2,990,409	g→t	Q310K			
R22	*galU*	B6N50_13325 ←	-	UTP--glucose-1-phosphate uridylyltransferase	2,611,590 - 2,611,600	INS 2,783,815 - 2,784,943	M168_V172ins(376)[b]	Deep rough	Hard	2
R23	*waaG*	B6N50_00405 →	*rfaG*	UDP-glucose:(heptosyl) LPS α1,3-glucosyltransferase (glucosyltransferase I)	74,589 - 74,606	Δ18 bp	Δ(A99-E105)	Deep rough	Hard	2
R24	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	76,970	(a)$_{7\to6}$	L249fsTer#263	Rough	Easy	1

R25	*rfaH*	B6N50_22865 →	-	Transcription/translation regulatory transformer protein	4,485,876	g→a	W4stop	Unknown	Hard	3
R26	*hldE*	B6N50_03700 →	*rfaE/waaE*	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	722,353	t→c	S233P	Deep rough	Easy	1
R27	*hldE*	B6N50_03700 →	*rfaE/waaE*	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	721,736	g→c	G27A	Deep rough	Easy	1
R28	*waaG*	B6N50_00405 →	*rfaG*	UDP-glucose:(heptosyl) LPS α1,3-glucosyltransferase (glucosyltransferase I)	74,864	g→t	G191stop	Deep rough	Hard	2
R29	*waaO*	B6N50_00415 →	*rfaI*	UDP-D-glucose:(glucosyl) LPS α-1,3-glucosyltransferase	76,970	$(a)_{7→6}$	L249fsTer#263	Rough	Easy	1
	pcoA	B6N50_01165 ←	-	Multicopper oxidase	237,963 - 237,973	INS 1,040,936 - 1,042,273	G152_G155ins(446)			
	eutM - eutN - eutE	[B6N50_06785] - [B6N50_06790] - [B6N50_06795] →	-	Ethanolamine utilization protein EutM - Ethanolamine utilization protein EutN - Aldehyde dehydrogenase EutE	1,342,381 - 1,343,448	Δ1,068 bp	Δ[EutM_A58-EutE_A181][c]			
R30	*galU*	B6N50_13325 ←	-	UTP--glucose-1-phosphate uridylyltransferase	2,611,680	c→a	E142stop	Deep rough	Hard	2
R31	*hldE*	B6N50_03700 →	*rfaE/waaE*	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	721,736	g→c	G27A	Deep rough	Easy	1
R32	*hldE*	B6N50_03700 →	*rfaE/waaE*	Bifunctional heptose 7-phosphate kinase/heptose 1-phosphate adenyltransferase	722,381 - 722,482	DUP 722,381-722,482 (102) bp	P242_A276dup(34)[d]	Deep rough	Easy	1

R33	*waaT*	B6N50_00420 →	-	UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase	78,125	g→a	W290stop	Rough	Easy	1
R34	*waaW*	B6N50_00430 →	-	UDP-galactose--(galactosyl) LPS alpha1,2-galactosyltransferase	79,820	c→t	Q281stop	Rough	Easy	1
R35	*galU*	B6N50_13325 ←	-	UTP--glucose-1-phosphate uridylyltransferase	2,611,872 - 2,611,880	INS 2,783,815 - 2,784,945	S75_L78ins(377)	Deep rough	Hard	2

Table 3.1. List of mutations found in *E. coli* C strains that are resistant to wildtype ΦX174. *E. coli* C resistant strains: R# is the strain number. Genes: name of the gene(s) in which mutations have been identified (as compared with wildtype *E. coli* C). Locus tags: identifier of each listed gene. Synonyms: alternative gene names. Descriptions: protein product encoded by each listed gene. Positions: genomic coordinates of mutation. Nucleotide changes: observed nucleotide change. Amino acid changes: resulting change in amino acid sequence. Predicted LPS structure phenotypes: based on predicted LPS structure (see **Fig. 3.7**). "Easy" or "Hard" LPS structure phenotypes?: easy LPS phenotypes were overcome or could be cross-infected after the phage evolution experiment – Standard, while hard LPS phenotypes required the phage cocktail evolution experiments – Increased diversity and – Increased diversity and generations to be infected. "→": indicates 5'→3' direction. "INS": insertion. "DUP": duplication. "stop": stop codon. "a": L249fsTer#263 indicates a frameshift (fs) leading to a premature codon stop (Ter); #: the position of the premature stop codon. "b": M168_V172ins(376) indicates an insertion; the two flanking amino acids are separated by a "_" and followed by the number of inserted amino acids in parentheses. "c": Δ[EutM_A58-EutE_A181] indicated a long deletion starting from EutM to EutE. P242_A276dup(34) indicates a duplication; the two flanking amino acids are separated by a "_" and followed by the number of duplicated amino acids in parentheses.

Excluded *E. coli* C resistant strains	Genes	Locus tags	Descriptions	Positions	Nucleotide changes	Amino Acid changes	Predicted LPS structure phenotypes
R1	*waaV*	*B6N50_00435* →	UDP-glucose:(Glucosyl) LPS β1,3-glucosyltransferase	80,534	$(t)_{7→6}$	F149fsTer#158[a]	Rough
	galE	*B6N50_15815* →	UDP-glucose 4-epimerase	3,096,098 - 3,096,107	Δ10 bp	T228fsTer#233	
R3	*galE*	*B6N50_15815* →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	Rough
R15	*galE*	*B6N50_15815* →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	Rough
R19	*yajC*	*B6N50_17610* ←	Preprotein translocase subunit	3,447,574	Δ1 bp	G18fsTer#23	Wildtype

Table 3.2. List of all *E. coli* C isolates excluded from the analysis. Excluded *E. coli* C resistant strains: R# is the excluded bacterial resistant strain number. Genes: name of the gene(s) in which mutations have been identified (as compared with wildtype *E. coli* C). Locus tags: identifier of each listed gene. Descriptions: protein product encoded by each listed gene. Positions: genomic coordinates of mutation. Nucleotide changes: observed nucleotide change. Amino acid changes: resulting change in amino acid sequence. Predicted LPS structure phenotypes: based on predicted LPS structure (see **Fig. 3.7**). "→": indicates 5'→3' direction. "stop": stop codon. "[a]": F149fsTer#158 indicates a frameshift (fs) leading to a premature codon stop (Ter); #: the position of the premature stop codon.

E. coli C resistant strains	Replicate #	Genes	Locus tags	Descriptions	Positions	Nucleotide changes	Amino Acid changes	Predicted LPS structure phenotypes
R3	1	*gltD*	B6N50_02830 ←	Glutamate synthase subunit gltD	555,954	t→c	Y318C	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	
R3	2, 3, 5, 9, 10	*WaaT*	B6N50_00420 →	UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase	77,302	(a)$_{6\to5}$	K16fsTer#23[a]	Rough
		gltD	B6N50_02830 ←	Glutamate synthase subunit gltD	555,954	t→c	Y318C	
		-	-	Hypothetical protein	2,147,627	c→t	Q793stop	
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	
R3	4	*oppB*	B6N50_13280 ←	Murein tripeptide ABC transporter/inner membrane subunit	2,600,004	g→a	A211V	Deep rough
		galU	B6N50_13325 ←	UTP--glucose-1-phosphate uridylyltransferase	2,611,933	Δ1 bp	I57fsTer#86	
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	
R3	6	*gmhD*	B6N50_00455 ←	ADP-L-glycero-D-mannoheptose-6-epimerase	84,986	g→t	Y140stop	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	
R3	7	*galE*	B6N50_15815 →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	Rough
		galT	B6N50_15820 →	Galactose-1-phosphate uridylyltransferase	3,097,396	t→c	L319P	
R3	8	*gltD*	B6N50_02830 ←	Glutamate synthase subunit gltD	555,954	t→c	Y318C	Rough
		-	-	Hypothetical protein	2,147,627	c→t	Q793stop	
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,095,443	g→a	G10D	

R15	1	*galP*	B6N50_04295 ←	Galactose-proton symporter	835,258	a→g	F405L	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	
R15	2	*WaaT*	B6N50_00420 →	UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase	77,590	g→t	E112stop	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	
R15	3 - 6	*waaW*	B6N50_00430 →	UDP-galactose--(galactosyl) LPS alpha1,2-galactosyltransferase	79,764	t→g	L262R	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	
R15	7	*galP*	B6N50_04295 ←	Galactose-proton symporter	836,442	g→t	S10stop	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	
R15	8	*galP*	B6N50_04295 ←	Galactose-proton symporter	835,353	a→g	L373P	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	
R15	9	*galE*	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	Rough
R15	10	*galP*	B6N50_04295 ←	Galactose-proton symporter	835,170	c→t	W434stop	Rough
		galE	B6N50_15815 →	UDP-glucose 4-epimerase	3,096,191	c→a	Y259stop	

Table 3.3. List of the additional mutations associated with *galE* mutations. *E. coli* C resistant strains: R# is the bacterial resistant strain number. Genes: name of the gene(s) in which mutations have been identified (as compared with wildtype *E. coli* C). Locus tags: identifier of each listed gene. Descriptions: protein product encoded by each listed gene. Positions: genomic coordinates of mutation. Nucleotide changes: observed nucleotide change. Amino acid changes: resulting change in amino acid sequence. Predicted LPS structure phenotypes: based on predicted LPS structure (see **Fig. 3.7**). "→": indicates 5'➔3' direction. "stop": stop codon. "ᵃ": K16sTer#23 indicates a frameshift (fs) leading to a premature codon stop (Ter); #: the position of the premature stop codon.
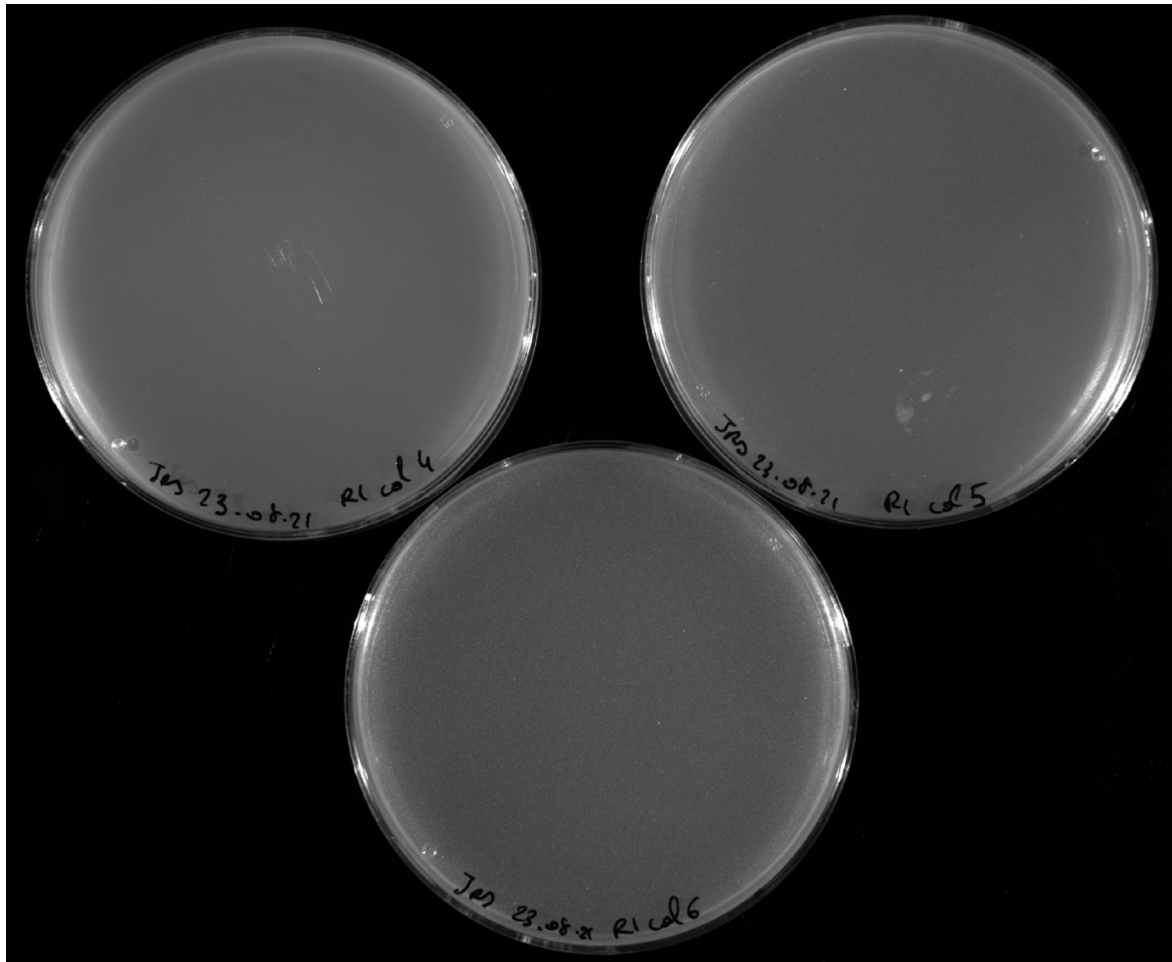
Fig. 3.4. Different lawn types made by colonies of *E. coli* C R1. Bacterial lawn formed by *E. coli* C R1 colony #4 (R1 col 4) is completely "smooth", while the bacterial lawns formed by colonies #5 and #6 (R1 col 5 and R1 col 6) are "granulous" (production of numerous aggregates).

The existence of two different aggregation phenotypes suggests that the glycerol stock of *E. coli* C R1 is constituted of two different populations: one "smooth" and one "granulous". Unlike R3 and R15, however, no discrepancy was found between the whole genome re-sequencing results of R1 from the glycerol stock and its ten re-streaked colonies. R1 formed two different lawn types, either "smooth" (no bacterial aggregate) or an exacerbate "granulous" in top agar overlays (**Fig. 3.5**). I decided to remove *E. coli* C R1 from the dataset since I could not identify the exact cause of the phenotypic inconsistency.
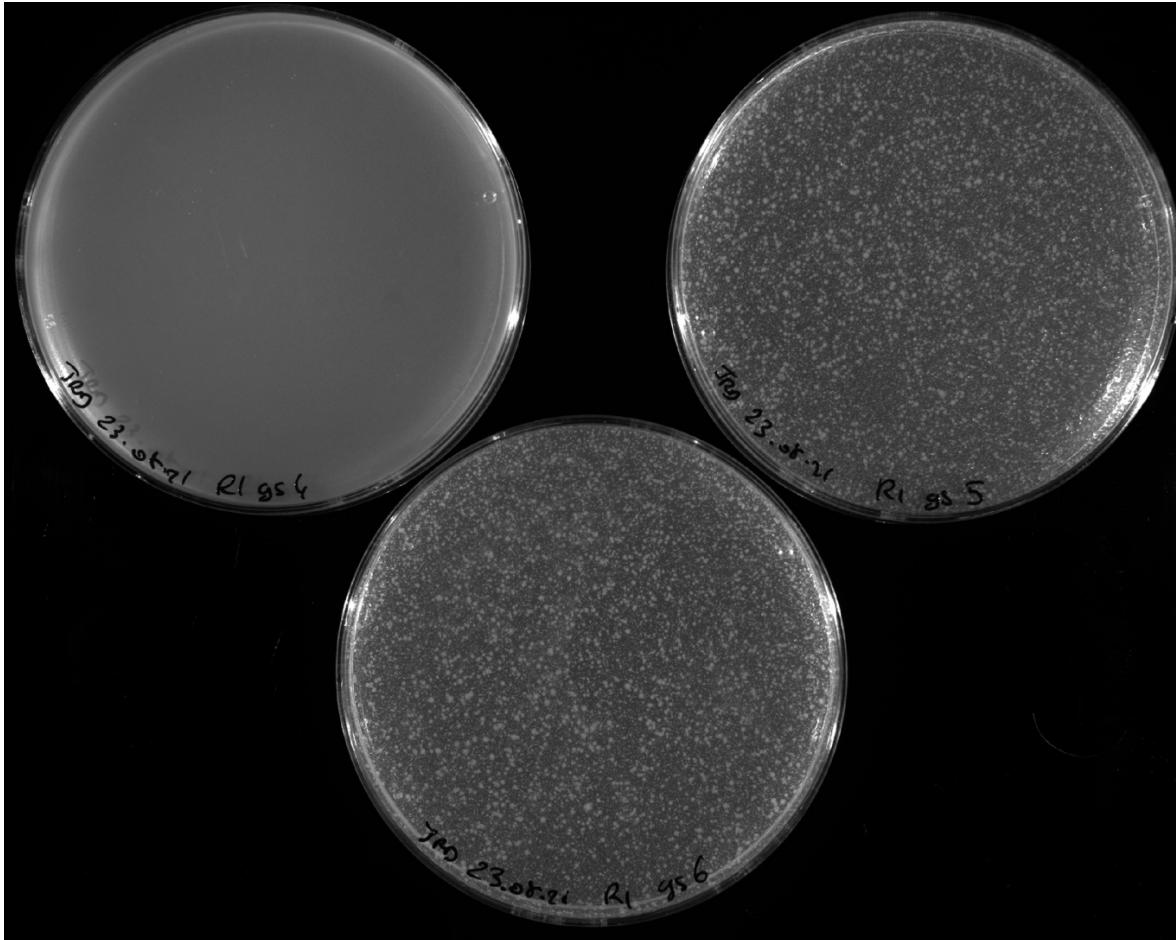
Fig. 3.5. Different lawn types made by *E. coli* C R1 glycerol overnight cultures. Bacterial lawn formed by *E. coli* C R1 #4 (R1 gs 4) is completely "smooth", while the bacterial lawns formed by #5 and #6 (R1 gs 5 and R1 gs 6) are extremely "granulous" (production of numerous and large aggregates).

Finally, I removed *E. coli* C R19 because the bacterium remains susceptible to ΦX174 wildtype. The wildtype phage can infect and produce plaques on a lawn of R19 but is not as efficient as infecting *E. coli* C wildtype (**Fig. 3.6A** and **3.6B**). When plated undiluted, a high titer ΦX174 wildtype lysate (~10^9 pfu.ml^{-1}) cannot clear *E. coli* C R19's lawn. Only a few thousand clear plaques are produced (**Fig. 3.6A**) and are very small compared to the ones produced on *E. coli* C wildtype's lawn (**Fig. 3.6B**). Thus, R19 is likely to be only partially resistant to ΦX174 wildtype infection.
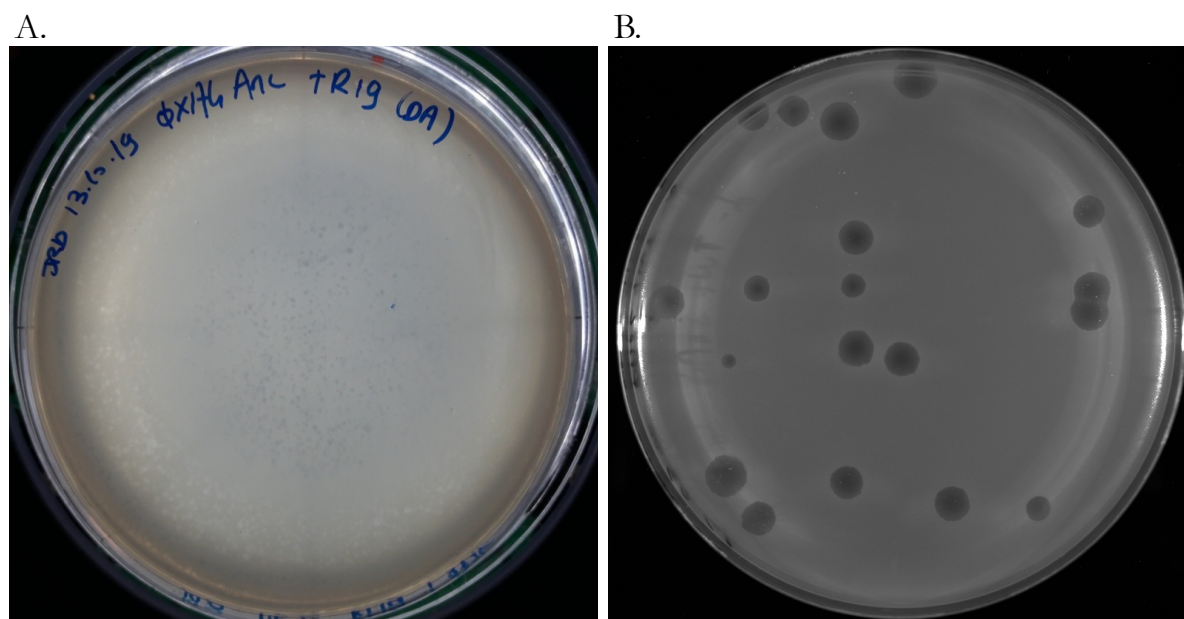
Fig. 3.6. *E. coli* C R19 remains susceptible to ΦX174 wildtype infection. A. Plaque morphotype produced by ΦX174 wildtype on a bacterial lawn of *E. coli* C R19 after 16h of incubation at 37°C. 100 μl of the undiluted wildtype phage lysate (~10^9 pfu.ml^{-1}) was mixed with an overnight culture of *E. coli* C R19. **B.** Ancestral "bull's eye" plaque morphotype produced by ΦX174 wildtype on a bacterial lawn of *E. coli* C wildtype after 16h of incubation at 37°C. Wildtype phage lysate was first serially diluted up 10^{-8}, and 100 μl of the dilution 10^{-8} was mixed with an overnight culture of *E. coli* C wildtype (see **Chapter II** section **2.3.2**).

R19 carries a single mutation *yajC* (locus tag B6N50_17610), which encodes a periplasmic protein[144] with a putative preprotein translocase subunit (Pfam e-value 2.3e-26) (see **Table 3.2**). The *yajC* gene played a crucial role in phage P22 DNA injection into the cytoplasmic membrane of *Salmonella enterica* serovar Typhimurium, perhaps via the stabilization of the SecYEG-SecDF-YajC-YidC translocon complex[145,146]. While no link to the LPS biosynthesis or assembly has been defined yet (meaning that R19 displays the native LPS structure), *yajC* could conceivably play a role in the injection of phage DNA into the bacterium's cytoplasmic membrane[145,146].

3.4.2. Mutational analysis on the remaining 31 resistant *E. coli* C strains

Of the remaining 31 strains, 27 are predicted to carry a single mutation, three carry two mutations, and one carries three mutations (see **Table 3.1**). A total of 36 mutations were identified, 32 of which are unique (including 15 nucleotide substitutions, 13 deletions, one duplication, and three

IS*4* and IS*5* insertion events). Twenty-four (of 32, or 75%) mutations introduce premature stop codons or lead to frameshifts and thus are highly likely to disrupt gene function.

Interestingly, two of these mutations are shared by two strains (R27/R31 and R8/R10), and one mutation is shared by three strains (R4/R24/R29). Given that there are almost limitless ways to disrupt gene function, this high degree of parallelism is initially surprising. The repetition of these mutations is unlikely to result from their positive selection. That is, phage-resistant *E. coli* C mutants of both low and high fitness are expected to have an equal chance of appearing on our agar plates (as long as they can form visible colonies, see **Fig. 3.3**)[62]. Presumably, the observed parallelism is, instead, due to elevated mutation rates at particular genomic positions[147]. The fact that two of three parallel mutations occur in homopolymeric tracts supports this hypothesis[148].

E. coli C R1, R3 and R15 are not the only mutant strains that have been found to carry more than one mutation. In four other *E. coli* C resistant strains included in the mutational analysis (out of 31, ~13%), whole-genome re-sequencing revealed that additional mutations with no apparent link to the LPS biosynthesis, assembly or control are also associated with the mutation responsible for the phage resistance (**Table 3.1**). *E. coli* C R8 and R10 share the same mutation in *hldE* and resist ΦX174 infection (**Table 3.1**). However, R10 also carry a 244 bp deletion in a repetitive CRISPR array sequence. A similar observation is made for R24 and R29, where both strains carry the same mutation in *waaO* which confers immunity to the phage. Yet, R29 is a triple mutant carrying an additional large 1,068 bp deletion across *EutM-EutN-EutE* genes and an *IS4* family transposase insertion in the *pcoA* gene (**Table 3.1**). Finally, *E. coli* C R6 and R21 carry an additional single nucleotide polymorphism (SNP) in *pssA* and *potH* genes, respectively (**Table 3.1**).

All 31 resistant strains carry at least one mutation in genes involved in LPS assembly, the production of the required hexoses or the regulation of the LPS core biosynthesis. Thus, I concluded that acquiring a single mutation in one of these genes is enough for *E. coli* C to become resistant to ΦX174.

Mutations in genes encoding the LPS machinery are expected to lead to changes in LPS structure and, in some cases, a switch from a rough to a deep rough phenotype[112,113,121]. Logical predictions of mutant LPS structures can be made according to a "LEGO® block" principle: if an enzyme synthetizing a particular LPS component is absent or non-functional, then that LPS component cannot be synthesized. In addition, LPS components assembled downstream of the missing

component are also absent from the final LPS structure. Under the assumption that each mutation in an LPS gene leads to a complete lack of function, I used this LEGO® block principle to predict the LPS structure and type of each of the 31 *E. coli* C phage-resistant mutants (**Figs. 3.1**, **3.2**, and **3.7**; **Table 3.1**).

3.5. Predicting the LPS phenotype of each *E. coli* C mutant

A total of seven emergent LPS structures were predicted from my set of mutants (**Fig. 3.7**). Compared with the *E. coli* C wildtype LPS, three of these seven LPS structures carry alterations in the outer-core LPS and hence are predicted to be of the rough type (as is wildtype *E. coli* C; ten mutants). The remaining four LPS structures differ from wildtype LPS in both the inner- and outer-core LPS and are presumed to be of the deep rough type (twenty mutants; see **Table 3.1**).

The LPS structure and type of the remaining mutant (*E. coli* C R25) were not predicted because the effect of the phage-resistance mutation on LPS biosynthesis remains unclear. R25 carries a nonsense mutation in *rfaH*. The *rfaH* gene encodes a transcriptional antiterminator that regulates the *waa* operon[109,133], meaning that potentially any number of LPS biosynthetic genes could be affected by this mutation.
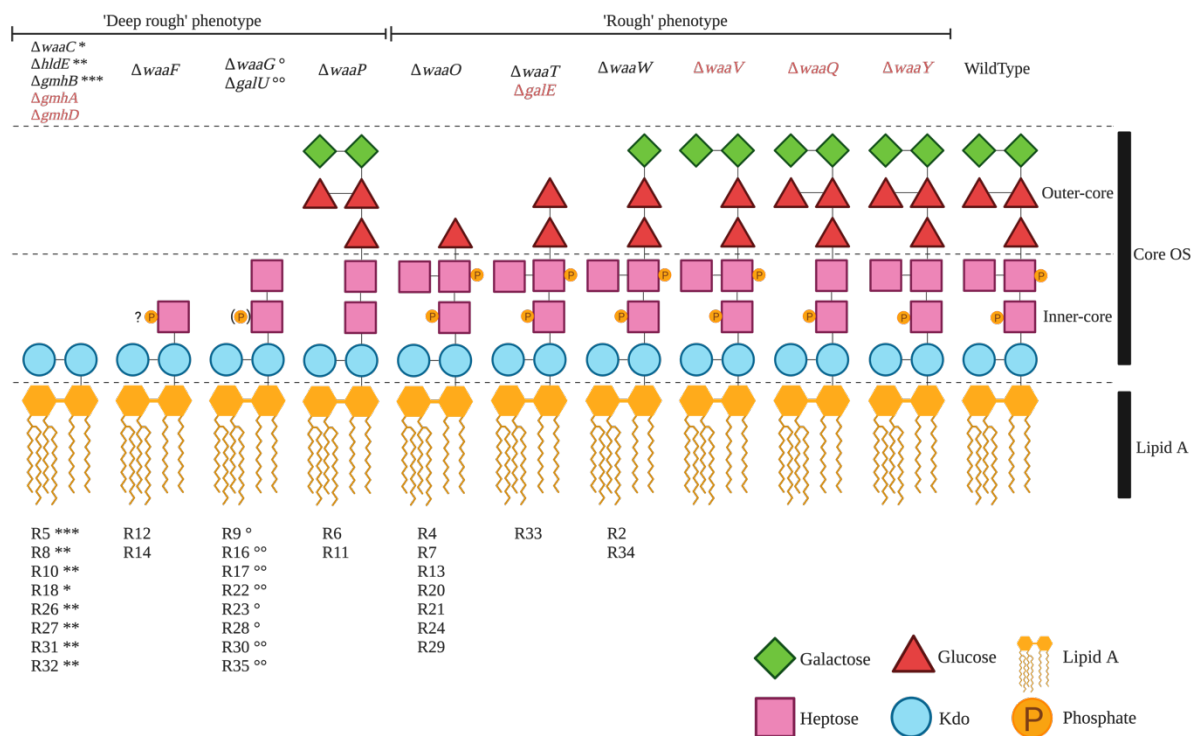
Fig. 3.7. Predicted LPS structures of 30 ΦX174-resistant *E. coli* C strains. Predicted LPS structures fall into two groups depending on their degree of truncation and/or loss of phosphate groups: the "deep rough" phenotype with a completely truncated outer core or a loss of phosphate groups (twenty *E. coli* C mutants) and the "rough" phenotype with smaller LPS truncations located in the outer core (ten *E. coli* C mutants). The LPS phenotype of the final *E. coli* C mutant (R25) could not be predicted (see text). The products of genes in red are also involved in LPS biosynthesis and although no mutations were identified in these genes during this study (the *galE* mutants R1, R3, and R15 were removed from the analyses, see text), are also potential targets for phage-resistance. Details on the functions of *gmhA*, *gmhB*, *gmhD*, *hldE*, *galU*, and *galE* are provided in **Fig. 3.2**. "?P": No information was found on the phosphorylation of Hep(I) in the absence of *waaF*; (P): only 40% of the hexose phosphorylation was observed[121]; R#: the number given to each *E. coli* C resistant strain; "*, **, ***, °, °°": identical symbols indicate that different genes are inactivated but are expected to lead to the same LPS core structure.

Each truncated LPS in the dataset is categorized into rough or deep rough phenotypes. The rough phenotype does not fundamentally change the physical and chemical properties of the LPS and still contains both an outer and an inner core[120]. However, the deep rough phenotype lacks the outer and/or the specific phosphate groups. These modifications in LPS structure can lead to a plethora of phenotypic effects[97], including (i) destabilization of the outer membrane[107], (ii) changes in the expression of some outer membrane proteins, (iii) modification of intracellular turgor pressure[149], (iv) increased susceptibility to hydrophobic compounds such as antimicrobial peptides, antibiotics, and bacteriocins[110,111,106,112–114,150], (v) altered interactions with the host immune system[151,107,152], (vi) alteration of the redox status of cells leading to an oxidative stress[150], and (vii) changes in resistance to phages[96,153–155].

Important fitness trade-offs are likely to emerge in phage-resistant bacteria carrying modified LPS structures, especially if the modifications lead to a deep rough LPS phenotype[156–159]. Based on the assumption that each mutation in an LPS gene leads to a complete loss of function (**Fig. 3.7**), phage-resistant strains that share the same predicted LPS structures should therefore experience similar fitness trade-offs, resulting in similar growth defects. Thus, the predicted LPS genotype-phenotype maps can be tested by using the growth dynamic of each resistant mutant as a proxy to discriminate between the LPS categories.

3.6.   Growth dynamics of the different *E. coli* C strains

3.6.1.   Growth dynamic of *E. coli* C wildtype without ΦX174

To test whether phage-resistant strains experienced growth defects, I first needed to assess the growth dynamic of *E. coli* C wildtype without phage. To do so, I monitored the growth of three independent *E. coli* C wildtype cultures (see **Chapter II** section **2.2.2**). The growth of *E. coli* C wildtype is characterized by a long lag phase spanning the first two hours of incubation (**Fig. 3.8**). The OD600 values increased immediately after the inoculation at t = 0 min (**Fig. 3.8A**). In contrast, the number of bacteria inside each culture tube was stable and started to increase exponentially only after two hours of inoculation (**Fig. 3.8B**).
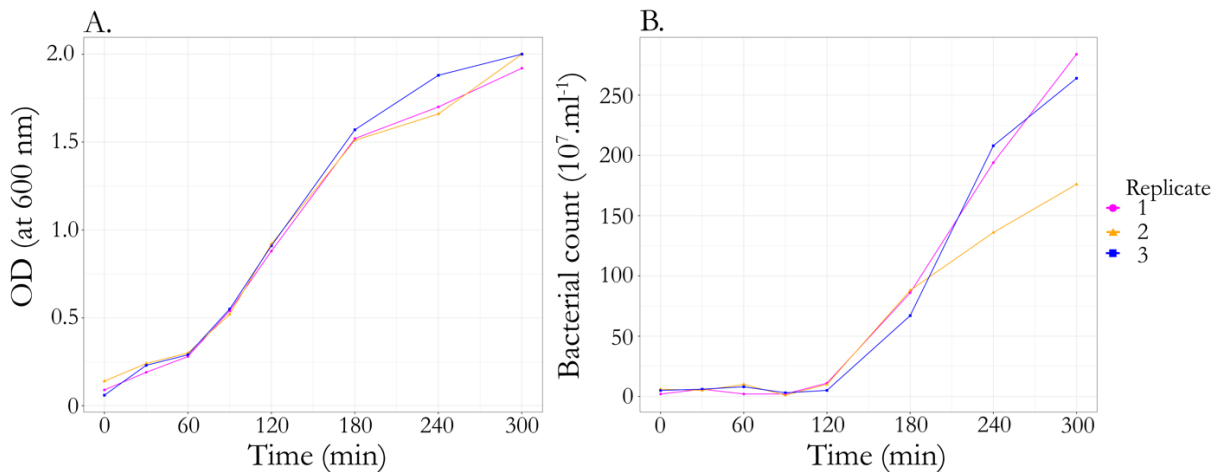
Fig 3.8. Growth curves of *E. coli* C wildtype. **A.** OD600 of three independent cultures of *E. coli* C wildtype strain. **B.** The number of cfu over time. Two dilutions (10^{-5} and 10^{-6}) were plated for each time point, in triplicate.

OD measurement and bacterial enumeration on solid agar plates presume that bacteria grow as single cells of equal size and are evenly spread in the liquid culture and solid agar plate (each visible colony is founded by a unique individual)[160,161]. However, the presence of bacterial aggregates in the sample breaks this assumption. Unbroken bacterial aggregates can remain even after vortexing the culture and not evenly dispersed in the liquid culture (sedimentation) or on the solid agar plates, resulting in an inaccurate estimation of the count of viable bacteria[162]. *E. coli* C tends to form aggregates when growing in a liquid environment[115]. The method and/or the media used to perform the serial dilutions before the plating step may affect the accuracy of the bacterial enumeration[163] and explain the disparity between the OD600 values and the colony counts.

I tested three different methods of serial dilutions: (i) in PBJ in 96 well plates, (ii) in LB supplemented with CaCl$_2$ and MgCl$_2$ (at final concentrations of 5 and 10 mM, respectively) in 96 well plates and, (iii) in PBJ in Eppendorf (Epp) tubes (see **Chapter II** section **2.2.3**). Both dilution methods (96 well plates or Eppendorf tubes) and media had no impact on the count of bacteria (**Figs. 3.9A** and **3.9B**).
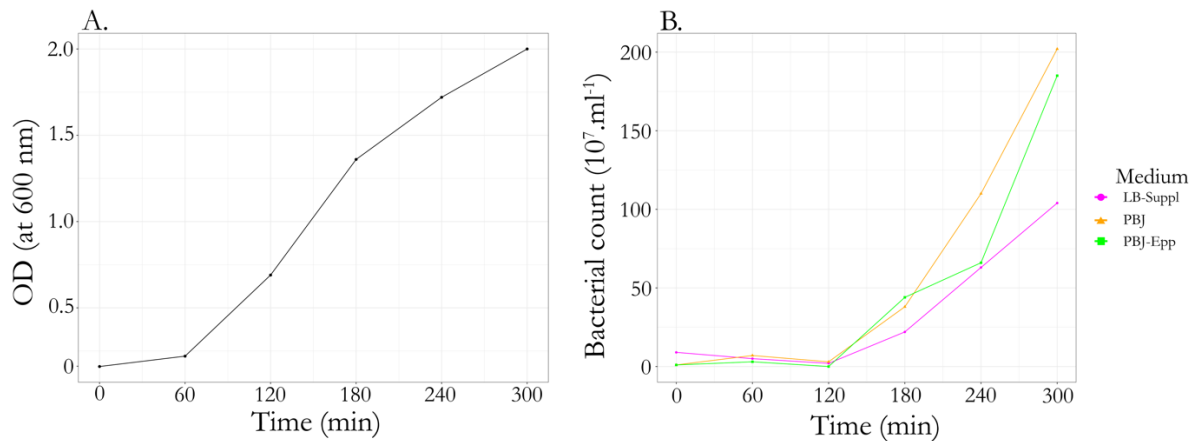
Fig 3.9. Both serial dilution methods and media have no impact on bacterial enumeration. A. OD600 a culture of *E. coli* C wildtype strain. **B.** The number of cfu over time. Two dilutions (10^{-5} and 10^{-6}) were plated for each time point. "LB-suppl": serial dilution in LB supplemented with CaCl$_2$ and MgCl$_2$ (final concentration of 5 and 10 mM, respectively) in a 96 well plate; "PBJ": serial dilution in PBJ in a 96 well plate; "PBJ-Epp": serial dilution in PBJ in Eppendorf tubes.

Ultrasounds methods (sonication) have proven their efficiency to inactivate and disrupt cell-forming aggregates in bacteria and yeast[164,165]. Similar methods could be applied to *E. coli* C wildtype before OD measurements and culture plating. Alternatively, future experiments could be done under a microscope to determine if the difference comes from the bacterial cells only increasing in size but do not divide at the beginning of the experiment (an increase in the population size versus an increase in the biomass).

3.6.2. Growth dynamic of each *E. coli* C resistant strain

To determine if (i) the resistant mutants show fitness trade-offs and (ii) if the resistant mutants sharing the same predicted LPS structures experience similar growth defects, I monitored the growth of each phage-resistant *E. coli* C resistant in my evolution experiments' conditions (**Fig.**

3.10). Resistant strains were categorized with respect to their growth compared to *E. coli* C wildtype. Mutants that grew similarly to *E. coli* C wildtype (black line) were categorized as "fast growers" (blue line). Resistant strains which grew somewhat more slowly than *E. coli* C wildtype and fast growers were categorized as "intermediate growers" (green line). Finally, resistant strains that grew more slowly than *E. coli* C wildtype and intermediate growers were categorized as "slow growers" (red line). These categories were used later to set up the ratio between the number of permissive and non-permissive cells for each evolution experiment (see **Chapter II** sections **2.3.5** to **2.3.9**).

The fast growers' category consists of *E. coli* C mutants (14 out of 31, ~45%) with no growth alterations compared to the wildtype strain (**Figs. 3.10A** and **3.10B**). Surprisingly, both rough and deep rough resistant strains belong to this group and collectively carry all seven predicted LPS structures (eight with the unpredicted LPS structure displayed by R25) (**Fig. 3.7**). This result suggests that not all mutations involved in the LPS biosynthesis or core assembly result in fitness trade-offs during bacterial growth in the absence of phage.

Bacterial mutants in the intermediate grower category (13 out of 31, ~42%) are also rough and deep rough but only collectively display two LPS structures (*waaO* and *galU/waaG* mutants; **Fig. 7**). After growth normalization with respect to their corresponding *E. coli* C wildtype control, intermediate growers are characterised by a slower mid and late-exponential growth phase (**Fig. 10C**).

Finally, the slow grower category is composed of four mutants (out of 31, ~13%) with a heptoless "deep-rough" LPS structure (**Fig. 3.7**). *E. coli* C R8, R10, R18, and R26 growth is characterised by a very long lag phase that ends after 60 to 90 minutes of incubation (**Fig. 3.10C**). Interestingly, the other resistant strains predicted to be heptoseless mutants (R5, R27, R31, and R32; **Fig. 3.7**) do not group with the slow growers but instead group with the fast growers (**Fig. 3.10A**).

My results suggest that the growth categories of the *E. coli* C resistant strains (especially the fast growers) poorly reflect the predicted LPS classes (**Fig. 3.7**). While the LPS classes generated based on the mutated genes may be incorrect, other parameters (such as the culture media's composition) may have attenuated potential fitness trade-offs of bacterial mutants, allowing them to grow normally. Further experiments could investigate the growth dynamics in another defined nutrient-limited media (e.g., minimum M9 media).
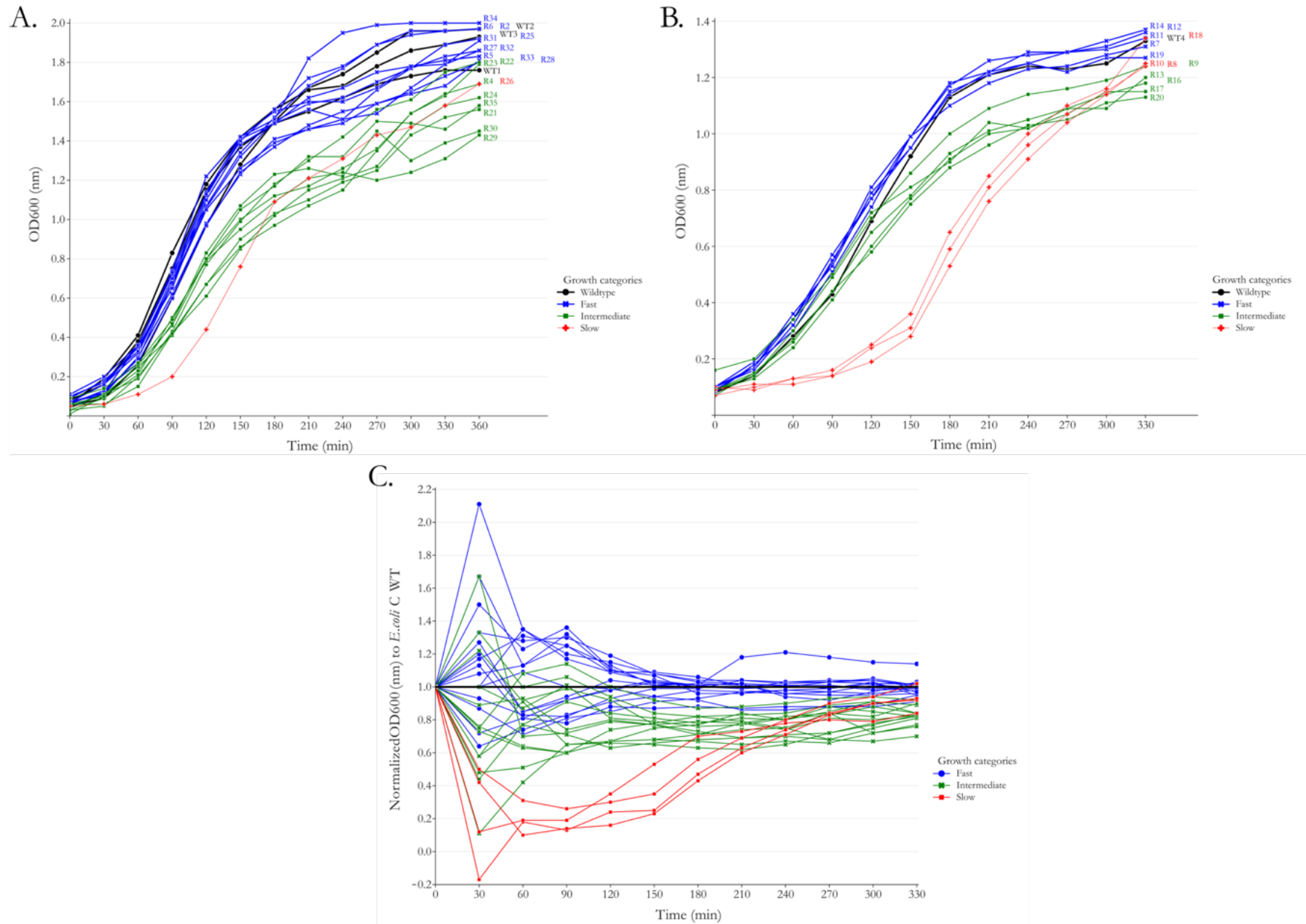
Fig. 3.10. *E. coli* C resistant strains can be categorized as fast, intermediate, or slow growers. The mean OD600 values of two independent cultures grown in 5ml LB are shown. Resistant strains are categorized with respect to their growth compared to *E. coli* C wildtype. Black lines: *E. coli* C wildtype. Blue lines: resistant bacteria that grew similarly to *E. coli* C wildtype ("fast growers": R2, R5, R6, R7, R11, R12, R14, R19, R25, R27, R28, R31, R32, R33, and R34). Green lines: resistant strains grew somewhat more slowly than *E. coli* C wildtype and fast growers ("intermediate growers": R4, R9, R13, R16, R17, R20, R21, R22, R23, R24, R29, R30, and R35). Red lines: resistant strains that grew more slowly than *E. coli* C wildtype and intermediate growers ("slow growers": R8, R10, R18, and R26). **A.** Growth curves of resistant strains R2-R6 and R21-R35. **B.** Growth curves of resistant strains R7-R20. These data (mutants and controls) were collected in a separate block from those in panel A. **C.** Mean OD600 values of all resistant strains in panels A and B, each normalized to their corresponding *E. coli* C wildtype control $((OD600\_R\#_{ti} - ODR\#_{t0})/(OD600\_wildtype_{ti} - OD600\_wildtype_{t0}))$. R# is the bacterial resistant strain number.

3.7. Discussion

3.7.1. Not all LPS genes are involved in phage defence mechanisms

No mutation was present in the *waaA* operon despite its role in the LPS biosynthesis. In *E. coli*, the 3-deoxy-manno-octulosonic acid (KDO) transporter transferase WaaA incorporates two KDO residues into the lipid A precursor IV$_A$[166]. However, bacteria lacking Kdo biosynthesis or Kdo-transferase genes exhibit severe membrane defects and depressed growth[117,167]. An *E. coli* C mutant defective of *waaA* might not be viable and able to form colonies even in a rich medium. Therefore, the minimal core LPS structure I should observe among the resistant strains is only composed of two hexaacylated lipid-A-Kdo$_2$.

There was also no mutation found in *waaQ* or *waaY* genes. *WaaPQY* mediates modifications that are not essential for the final core biosynthesis[112]. Moreover, deleting these genes individually does not result in a deep rough phenotype[113]. Evolution experiments performed by Pepin *et al.* also suggest that ΦX174 can still infect an *E. coli* Δ*waaQ* mutant[81]. Therefore, I assumed that neither *waaQ* nor *waaY* confers complete resistance when mutated.

However, the lack of spontaneous phage-resistant strains with a mutation in *gmhA* or *gmhD* genes is more surprising. GmhA is a D-sedoheptulose 7-phosphate isomerase, and GmhD is an ADP-L-glycero-D-mannoheptose 6-epimerase. They catalyse respectively the first and last steps of the ADP-L-*glycero*-β-D-*manno*-heptose production (**Fig. 3.2**), a heptose precursor utilized in the assembly of the inner core LPS[124,168]. Consequently, a loss of function in *waaC*, *gmhA*, *hldE*, *gmhB* and *gmhD* results in a similar LPS phenotype[96] (**Figs. 3.1, 3.2** and **3.7**). In the following paragraph, I speculated on possible explanations for the lack of these mutants in my dataset.

It is possible that (i) by chance I did not pick colonies of *gmhA* or *gmhD* mutants that grew on a solid agar plate while selecting for my 35 ΦX174-resistant strains (see **Chapter II** section **2.2.1**). (ii) The fitness cost of lacking GmhA and GmhD activities may be greater than lacking WaaC, HldE or GmhB activities. Thus, *gmhA* and *gmhD* mutants might not form visible colonies on solid agar plates. (iii) Concerning the *gmhD* gene only, it might be possible that ADP-D-*glycero*-β-D-*manno*-heptose can be substituted for ADP-L-*glycero*-β-D-*manno*-heptose and incorporated into the inner core LPS. ΦX174 might still be able to recognize and adsorb to this modified LPS structure.

Parab and colleagues recently carried out additional fluctuation experiments following the protocol described in **Chapter II** section **2.2.1**. The authors successfully isolated 50 additional resistant *E. coli* C strains to ΦX174 wildtype infection. Of these 50 mutants, 11 carry at least one mutation in the *gmhD* gene, while one carries a mutation in the *gmhA* gene (Parab *et al.,* in prep). Thus, hypothesis (i) likely explains the absence of these two types of bacterial mutants in my dataset.

3.7.2. Emergence of mutants with too many mutations

Neither *potH*, *pcoA*, [*EutM-EutN-EutE*] or *pssA* have been reported to affect phage resistance when mutated. PotH is a channel-forming protein of one of the major putrescine uptake systems *PotFGHI*, an ABC (ATP-binding cassette) membrane transporter[169,170]. PcoA is a multicopper oxidase expressed in the periplasm as part of the copper-resistance response in *E. coli*[171,172] and oxidases Cu(I) into the less toxic form Cu(II)[173], offering periplasmic protection from extreme copper stress[174]. EutM is one of the shell proteins required for the metabolism of ethanolamine. EutN is believed to be a minor structural component of the carboxysome[175–177]. EutE is an acetaldehyde dehydrogenase which generates acetyl-coenzyme A (acetyl-CoA) from acetaldehyde[176,178]. Δ*EutE*, Δ*EutM*, and Δ*EutN* mutants failed to use ethanolamine as a carbon source and, in particular, the Δ*EutE* mutant showed an accumulation of endogenous acetaldehyde, toxic for the cell at high concentration[178,179]. Finally, PssA is an essential phosphatidylserine synthase[180] catalysing the synthesis of phosphatidylserine[181]. Bacterial strains lacking phosphatidylserine synthase activity (encodes by *pss*) have cell membranes deficient in phosphatidylethanolamine, a major phospholipid[182,183].

Similarly, the deletion of one of the R10's CRISPR arrays may not be linked to a ΦX174 resistance mechanism. After screening *E. coli* C's genome for CRISPRs and *cas* genes – well-known integrative parts of the CRISPR-Cas adaptive immune system of bacteria and archaea[14,184,185] – using CRISPRCasFinder[115,186], I found that no *cas* genes were present. CRISPR arrays in *E. coli* C are orphans[187]. Consequently, they may not be functional since no new protospacers matching ΦX174's genome can be incorporated into CRISPR arrays by Cas proteins[188].

One explanation for the presence of these multiples is that these additional mutations were hitchhiked and driven to fixation by the mutations in the gene involved in the LPS biosynthesis or assembly. It may be the case for R6 and R10. *E. coli* C R6 carries one synonymous mutation (V54V) in *pssA* (**Table 3.1**), reversing this site to its ancestral state[115]. The deletion in R10 is also present in R8 but is not associated with additional mutations (**Table 3.1**). The single deletion $(gc)_{5\to4}$ in the

hldE gene is sufficient to confer immunity to wildtype phage infection. The deletion of R10's CRISPR array may have been lost (potentially during recombination events[189,190]) and was hitchhiked with the mutation in the *hldE* gene (**Table 3.1**).

Loss of function of *potH*, *pcoA* or [*EutM-EutN-EutE*] activities often introduces changes in the composition of the outer membrane of Gram-negative bacteria and growth defects[173,174,179,191]. Alternatively, *E. coli* C strains carrying one of these mutations before being exposed to the phage might then grow slower on solid selective agar plates (in the presence of the phage) thanks to an accumulation of toxic intermediate products (e.g., acetaldehyde, Cu(I)), or the destabilisation of certain OM proteins (e.g., Sec complex translocon complex, polyamine uptake). Therefore, slow-growing phenotypes might be beneficial in specific experimental conditions[143,192] since the phage replication is intimately linked to the metabolism of its host[193–195]. Slow-growing but still permissive mutants could survive and persist after the initial wave of infection while limiting the replication and diffusion of the phage[196,197], allowing the bacteria to accumulate mutations until one makes them fully resistant (e.g., via the modification of the core LPS structure, **Fig. 3.11**).
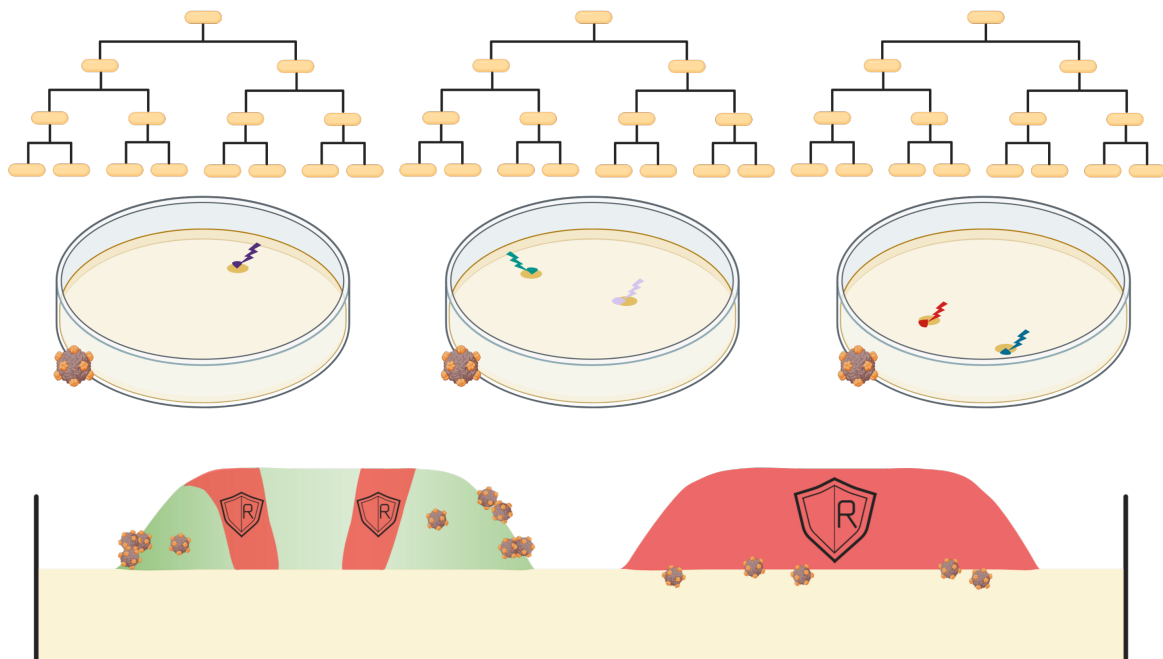
Fig. 3.11. An excess of jackpot mutations can lead to phage persistence. The genealogy shows that no mutation conferring phage resistance arises in the liquid culture without the presence of ΦX174. However, mutations resulting in a slow-growing phenotype may happen. During the growth of *E. coli* C on the solid agar plate, slow-growing bacteria are trapped as

"bubbles" inside the colony or can surf and grow at the edges of the colony[198]. Mutations can continue to accumulate in a fraction of the slow-growing bacterial population within each colony (coloured lightning bolts and sectors). One of these *de novo* mutations may confer resistance to phage. In the presence of ΦX174, fast-growing but susceptible bacteria are quickly killed by the phage. The trapped and slow-growing resistant bacteria are released into the environment, rescuing the colony. Adapted from Testa *et al.*, 2019[199].

An excess of jackpot mutations is usually massively enriched in a bacterial population growing in a structured environment compared with a well-shaken liquid environment[198]. Bacterial subpopulations within a single colony may display different frequencies of mutation and phenotypes[200,201] (coloured sectors) that will not be killed at the same rate in the presence of ΦX174[200]. This subset of the bacterial population may persist through the phage infection[202,203]. Because of the spatial competition, some will stay trapped inside the colony in a quiescent state, while others could continue accumulating mutations (coloured lightning bolts) via surfing at the edge of the colony[198]. By killing the fast-growing but susceptible bacteria in the colony, the phage might promote the dissemination of phage-resistant bacterial strains from the release of the resistant mutants trapped inside the colony[198]. Further experiments should, therefore, focus on engineering all possible single mutant *E. coli* C strains and developing assays to test their persistence or resistance to ΦX174.

3.7.3. Composition of the media might diminish fitness trade-offs

Media composition has a strong influence on phage production. For example, the presence of divalent cations Ca^{2+} and Mg^{2+} stimulates phage production and plaque morphogenesis[204,205]. They are required for some phages to successfully adsorb to their hosts and for their DNA to correctly penetrate through the bacterial OM[206–208]. In the case of ΦX174, Ca^{2+} and Mg^{2+} neutralise the negative and repulsive charges from the LPS, allowing interactions with *E. coli* C[209,210]. Hence, supplementing culture media with $CaCl_2$ and $MgCl_2$ at millimolar concentration is common to isolate, enumerate, and propagate ΦX174 and other bacteriophages[211–213].

Not only do divalent cations impact phage parameters, but they also impact the physiology of the bacterium. Ca^{2+} and Mg^{2+} can stabilize and protect the bacterial OM membrane while increasing the rigidity of the LPS layer[214–216]. Therefore, growth defects resulted from remodelling the LPS structure in my *E. coli* C mutants (i.e., a potential destabilization of the OM in strains with deep

rough LPS phenotypes) may have been attenuated or hidden by growing with membrane defects in LB media supplemented with $CaCl_2$ and $MgCl_2$.

3.7.4.   Conclusion

To study LPS diversity in *E. coli* C, I generated 35 spontaneous phage-resistant strains and used whole genome re-sequencing to identify the mutation(s) conferring resistance (see **Chapter II** sections **2.2.1** and **2.2.5**). Due to a lack of isogeneity or only partial resistance to wildtype ΦX174, four of these strains (*E. coli* C R1, R3, R15, and R19; **Figs. 3.4** to **3.6**, **Table 3.1**) were excluded from downstream analyses. All the 31 remaining *E. coli* C mutants carry at least one mutation in genes involved in the LPS biosynthesis, assembly, or regulation (**Table 3.1**), collectively producing eight distinct LPS structures (**Fig. 3.7**). The currently accepted model for predicting LPS structure posits that the presence or absence of each LPS gene leads to a single LPS phenotype. Based on the position of each mutation, the one-gene-one-phenotype model allows us to predict the LPS structure of each of our LPS mutants.

LPS predictions are usually tested quantitatively via mass spectrometry and silver-staining methods [155,217–220]. Unfortunately, none of these methods was accessible or suitable. In addition, I found that while the silver-staining methods can give a high resolution of the composition of the LPS O-antigen part, it performs poorly for the core OS part[155]. Hence, I tested my LPS predictions qualitatively via the specific sensitivity and resistance to phage by evolving wildtype ΦX174 to specifically infect *E. coli* C cells with modified LPS structures. That is, a phage that can infect a given modified LPS structure is expected to be able to infect all bacterial strains displaying this LPS structure, regardless of the underlying mutations. Alternatively, the inability of mutant phages to cross-infect bacterial strains of the same predicted LPS structural class (but via different mutations) would indicate that these mutations lead to distinct LPS structures. The series of evolution experiments developed and performed to quickly adapt ΦX174 to each *E. coli* C resistant strain are presented in the next chapter.

Chapter IV – Adaptation of the coliphage ΦX174 to the resistant strains of *E. coli* C

4.1.   Characteristics of the coliphage ΦX174

4.1.1.   Genomic features

ΦX174 is a small (~30 nm) tailless coliphage belonging to the *Microviridae* family. It carries a 5386 bases-long ssDNA genome engulfed by a strict T=1 icosahedral nucleocapsid[221]. Its genome comprises 11 genes (some of which overlap): nine essential (*A, B, C, D, E, J, F, G,* and *H*) and two non-essential (*A** and *K*)[9,32,222,223] genes. The functions of each gene are listed in **Table 4.1**.

Genes	Descriptions	Essentiality?	Functions
A	Replication-associated protein	Yes	Stage II and stage III DNA replication
A*	-	No	May play a role in the inhibition of host cell DNA replication and superinfection exclusion[a]
B	Internal scaffolding protein	Yes	Procapsid morphogenesis and assembly of early morphogenic intermediates
C	-	Yes	DNA replication; facilitates the switch from stage II to stage III. Required for stage III
D	External scaffolding protein	Yes	Procapsid morphogenesis
E	Lysis protein	Yes	Host cell lysis
F	Capsid protein	Yes	Major coat protein; host recognition
G	Major spike protein	Yes	Host recognition
H	Minor spike protein (pilot)	Yes	DNA injection
J	DNA-binding protein	Yes	DNA packaging
K	-	No	May play a role optimizing burst size and phage production[b]

Table 4.1. **Functions of genes encoded by phage ΦX174 genome.** Adapted from Calendar, 2006, Oxford University Press[32] and Wichman and Brown, 2010[74]. [a]Eisenberg and Ascarelli, 1991)[224]; [b]Gillam *et al.*, 1985)[225].

The genomic mutation rate of ΦX174 – the product of the per-nucleotide site mutation rate and the genome size which determines the average number of mutations each offspring will have compared to the ancestral genome[226] – is one of the highest among ssDNA viruses, with 1.1×10^{-6} substitutions per <u>n</u>ucleotide site per <u>c</u>ell infection (s/n/c)[227]. Typically, DNA viruses show

mutation rates ranging from 10^{-8} to 10^{-6} (s/n/c), while the ones of RNA viruses range around 10^{-6} to 10^{-4} (s/n/c)[64]. The high mutation rate ΦX174 may be explained by the absence of GATC patterns in its genome[228,229]. These GATC motifs are recognized and methylated by the methyl-directed mismatch repair system (MMR) for the correction of post-replicative errors[230,231].

4.1.2. Host recognition and DNA injection

ΦX174 is a lytic phage that solely relies on attaching to the core oligosaccharide (OS) of the host's LPS for infection. In the laboratory, ΦX174 infects - and hence is usually grown on - *E. coli* C, which produces rough type (*i.e.*, lacking in O-antigen) LPS molecules[232–235] (**Fig. 4.1**). ΦX174 does not possess a tail that mediates the host recognition and attachment. The attachment is instead performed by one of its twelve spike proteins (protein G, also called a vertex)[32] with a high degree of specificity (**Figs. 4.2A** and **4.2B**). For instance, among 783 different *E. coli* isolates, only six (0.8 %) could be infected by wildtype ΦX174[79].
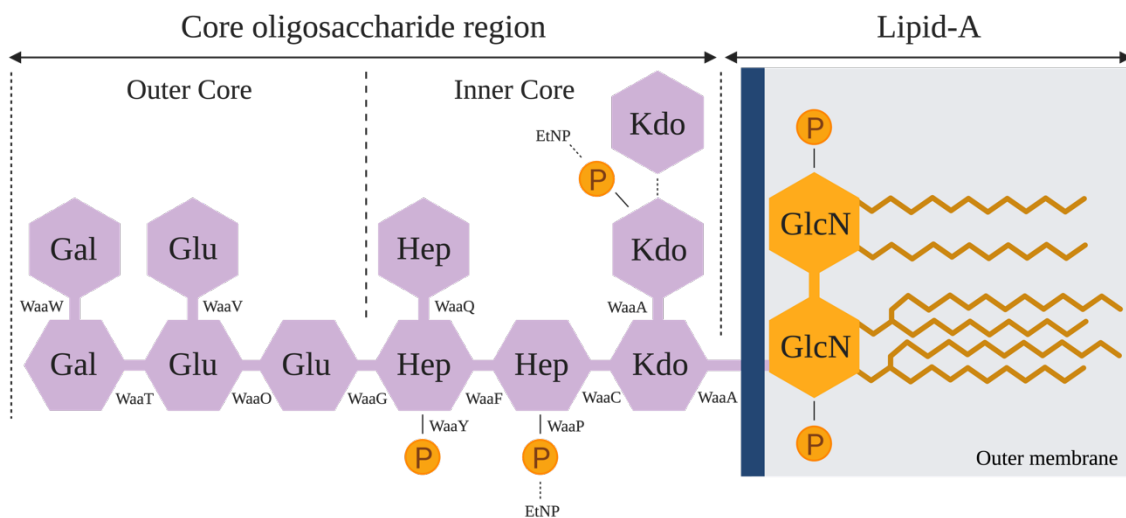
Fig. 4.1. Structure of *E. coli* C's rough type LPS. The rough type LPS of *E. coli* C is composed of two parts: lipid A (composed of an acetylated and 1,4'-diphosporylated β(1→6)-linked glucosamine (GlcN) disaccharide), and the core oligosaccharide[107]. The core oligosaccharide is subdivided into a structurally conserved inner core and an outer core. The dashed lines show non-stochiometric substitutions of phosphate (P), ethanolamine (EtNP), and 3-deoxy-D-*manno*-octulosonic acid (Kdo) residues on the LPS[236]. Details of the LPS assembly process are provided in **Chapter III** sections **3.1.1** to **3.1.4**).

Quickly after the attachment, the spike protein that mediated the LPS recognition is dissociated from the capsid. The F proteins previously in contact with the spike protein change their conformations, particularly in their EF and FG loops[237] (**Fig. 4.2C**). The EF loop maintains a stable interaction between the phage capsid and the bacterial outer membrane, while the FG loop opens a way for the spike (made of H proteins) and prepares the virus' DNA translocation[237,238] (**Fig. 4.2C**). A tube-like structure of H proteins is formed inside the host cell's membrane, facilitating the DNA ejection through the vertex (**Fig. 4.2D**). In the end, the H proteins disappear into the cytoplasm, while the hollow capsid stays outside (**Fig. 4.2E**).
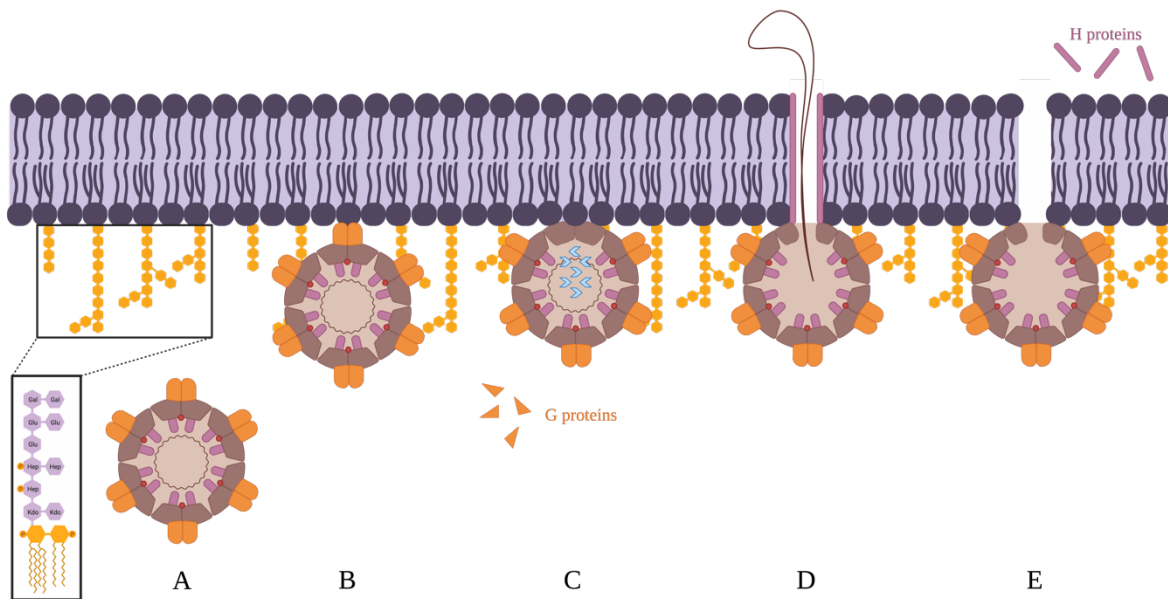
Fig. 4.2. Latest model for ΦX174 host recognition and DNA injection. **A**. and **B**. ΦX174 recognizes *E. coli* C wildtype's native rough type LPS (see **Fig. 4.1**) and attaches to the outer membrane surface of the bacterium via one of its major spike protein G (coloured in orange). **C**. The interaction destabilizes the G proteins, which quickly dissociate from the capsid. The loops on the surface of the F proteins (coloured in brown) stabilize and maintain the interaction between the capsid and the cell wall. The dissociation of the G proteins also triggers conformation changes (blue arrows) in the F and H proteins (coloured in pink). **D**. H proteins assemble into a tube-like structure and cooperate with the F proteins to translocate the ssDNA (coloured in black) across the periplasm. **E.** The exit used by DNA remains open after ejection. H proteins are ejected and degraded in the host cytoplasm. The empty capsid remained attached to the host surface. Adapted from Sun *et al.*, 2017[238].

4.1.3. ΦX174 host adaptation

Co-evolution experiments of ΦX174 with its host showed that substitutions mainly arose in structural proteins involved in the host recognition (F capsid protein) and DNA injection (H minor spike protein) systems (**Table 4.1**)[74–76,91]. Furthermore, the evolution of ΦX174 is strikingly parallel at the nucleotide level. Wichman and Brown summarized the results of over ten evolution experiments carried out on ΦX174 (for a total of 58 lineages) and demonstrated that, for any two evolution experiments performed under the same experimental conditions, ~50% of all substitutions arose in both experiments[74]. Interestingly, high degrees of parallelism was also observed between evolution experiments performed under different experimental conditions[74,75,77].

These findings strongly support the idea that parallel evolution was adaptive evolution and did not occur by chance due to the large dataset examined[82,83]. A significant proportion of substitution sites were common to strains that evolved in the laboratory and wild isolates[74] and may suggest that ΦX174 mechanisms of adaptation in laboratory conditions could recapitulate phage adaptation in nature[75] or that very few adaptive pathways are available for ΦX174 (possibly because of its small genome size).

4.2. Exploration of ΦX174 overcoming strategies to LPS-based phage resistance

None of these past evolution experiments has investigated how ΦX174 adapts to phage-resistant bacterial strains that harbour modified LPS structures. Thus, the overall aim of the chapter was to evolve highly-specific ΦX174 mutants infecting every LPS phenotype (see **Fig. 3.7**) and determining the mutations required for overcoming the different predicted LPS phenotypes. In a series of evolution experiments involving serial transfers of single and phage cocktails, I evolved ΦX174 to re-infect each *E. coli* C resistant strain. As observed previously, ΦX174 adapted to its resistant hosts by only modifying its host recognition (the F capsid protein) and DNA injection system (the H minor spike protein). I also showed that, under certain circumstances, increasing host and phage diversities could accelerate phage adaptation.

4.3.   Evolving ΦX174 in a liquid, well-mixed environment

4.3.1.   Growth dynamic of *E. coli* C wildtype during ΦX174 infection

This experiment aimed to maximize the production of ΦX174. To do so, I needed to find the best inoculation time to start the infection of *E. coli* C wildtype with the phage. I monitored the growth dynamic of *E. coli* C wildtype after adding the wildtype ΦX174 at three different time points (red pins in **Fig.4.3**, see **Chapter II** section **2.3.3**).
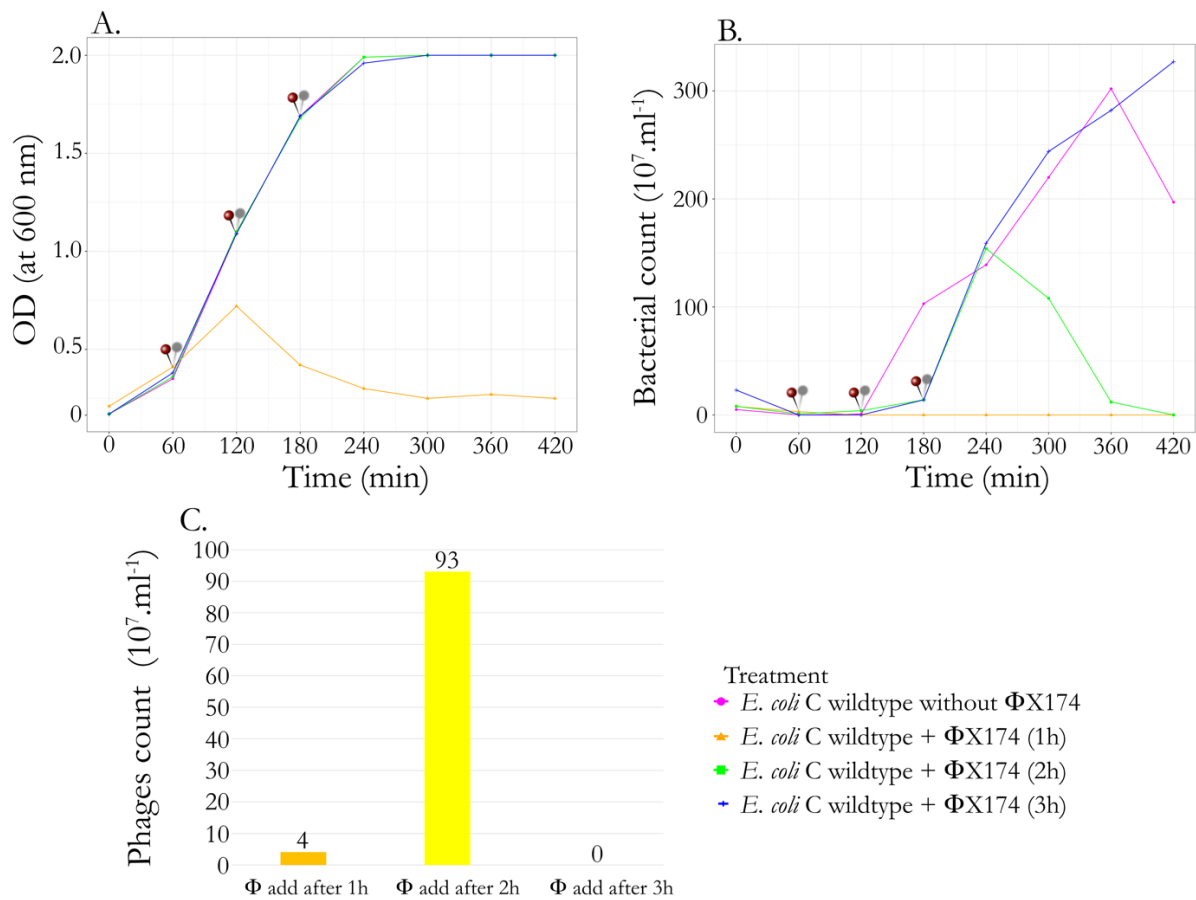
Fig 4.3. Growth curves of *E. coli* C wildtype in the presence of the wildtype ΦX174. **A.** OD600 of *E. coli* C wildtype strain after infection by ~10^6 phages (MOI$_{input}$ ~ 0.01) after one, two or three hours of incubation (red pins). **B.** The number of cfu over time. Two dilutions (10^{-5} and 10^{-6}) were plated for each time point. **C.** Number of ΦX174 present in each sample at the end of the experiment.

The entire *E. coli* C population quickly decreased when ΦX174 was added to the bacterial culture after one hour of incubation (orange line, **Figs 4.3A** and **4.3B**). In contrast, *E. coli* C populations

that grew for two hours only showed signs of killing after two hours post-infection (yellow line, **Fig. 4.3B**) and yielded the biggest number of phages at the end of the experiment (~$9.3*10^8$ pfu/ml, **Fig. 4.3C**). Phage production was increased by ~23-fold compared to infecting *E. coli* C wildtype culture after one hour of incubation.

Infecting bacteria after three hours of incubation was the worst-case scenario. It resulted in a late killing (blue line, **Fig. 4.3B**) and not enough phage was produced (**Fig. 4.3C**). The lack of sufficient phage production after 3h of incubation is likely due to the bacteria entering the stationary phase, which cannot be infected anymore by ΦX174[239]. Hence, I chose to infect *E. coli* C wildtype cultures after growing them for two hours without phage for my evolution experiments since it successfully maximized the number of phages retrieved at the end of the experiment.

4.3.2. Phage evolution experiment via serial transfers (Standard)

To understand how ΦX174 wildtype evolves to overcome LPS-based phage resistance in *E. coli* C, I serially transferred the phage on bacterial cultures containing: (i) *E. coli* C wildtype and (ii) one of the 35 resistant strains (R1 to R35, **Fig. 4.4;** see **Chapter II** section **2.3.5**). Phage lineages were founded for R1, R3, R15, and R19 despite bacterial strains being removed later from the different analyses (**Chapter III** section **3.4.1**). The presence of *E. coli* C wildtype was necessary to propagate ΦX174 and avoid its dilution from one transfer to another. Phages were inoculated into fresh, exponentially growing host cultures at each transfer to avoid the co-evolution of bacteria.
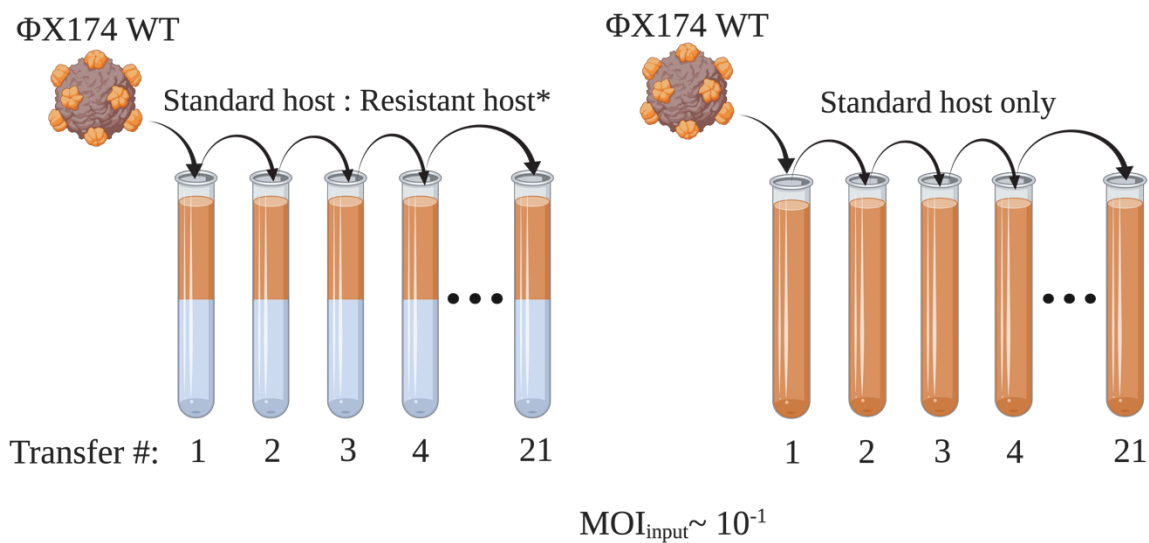
Fig. 4.4. Phage evolution experiment used to overcome *E. coli* C resistance to wildtype ΦX174. Phage evolution experiment – Standard. Thirty-five independent lineages, each founded by ΦX174 wildtype (WT), were serially transferred daily for up to 21 days (21 transfers). Phages were grown on non-evolving (i.e., freshly prepared at each transfer) host cultures containing a mixture of *E. coli* C wildtype and one of the 35 initially resistant strains (see **Table 3.1**). Phages capable of infecting the initially resistant host strain were isolated from 18 lineages. Three control lineages were included, where ΦX174 was cultured on only *E. coli* C wildtype. Details of the evolved phages are presented in **Tables 4.2** and **4.3**, and further methodological details are in **Chapter II** section **2.3.5**.

The experiment gave rise to 18 evolved phages that successfully infected their corresponding resistant strains. However, the evolved phages respectively infecting R3, R5, R11, R15, R19, and R34 were removed from the mutational and phenotypical analyses. They were found to be either non-isogenic (phage infecting R5), adapted to bacterial strains that were non-isogenic (phages infecting R3 and R15, respectively), adapted to a non-LPS bacterial mutant (phage infecting R19) or lost for unknown reasons with the impossibility of producing new phage lysates from their glycerol stocks (phages infecting R11 and R34, respectively). Details on the exclusion of the evolved phages respectively infecting R5 and R19 are provided in section **4.5.1**.

The remaining 12 evolved phages were able to collectively infect 21 initially resistant *E. coli* C strains (of the 31, ~68%; *E. coli* C R3, R5, R15, and R19 were removed from the further experiments, see **Chapter III** section **3.4.1**). Ten (of 31, 32%) resistant *E. coli* C strains remained uninfected (the "hard" resistant *waaP/pssA*, *waaG*, *galU*, and *rfaH* mutants, **Table 3.1**) after 21 transfers. However, I needed each *E. coli* C resistant strain to be infected or cross-infected by at least one evolved phage to determine the amino acid under selection for host recognition and later to challenge my predicted LPS genotype-phenotype maps.

Overcoming the phage resistance of this subset of "hard" resistant strains – the bacterial strains that remained uninfected after the phage evolution experiment – Standard may not be achievable for the phage in a few mutational steps. Otherwise, I would have identified ΦX174 mutants with single mutations infecting the "hard" resistant strains in a phage population of size $\sim 10^9$, where all non-deleterious point mutations should be present (mutational rate of $\sim 1.1 * 10^{-6}$ s/n/c[63,227]). ΦX174 may undergo several mutational steps to infect the strains that remained uninfected. Theory and previous research suggest that adaptation can be sped up by increasing genetic

variation[240,241]. Phage diversity can be increased by elevating the mutation rate via UV mutagenesis[242]. In the following section (**4.3.3**), I described another serial transfer experiment in which I included UV-B light exposure steps to increase the mutation rate of ΦX174.

Interestingly, I observed a new plaque morphotype that differs from the ancestral "bull's eye" plaque morphotype (**Fig. 4.5A**) in several evolved and control lineages after the 5th serial transfer. Evolved phages' plaques were small and turbid, without defined edges (**Fig. 4.5B**). After a few more transfers, the phage mutants producing small plaques took over the entire population (**Fig. 4.5B**).
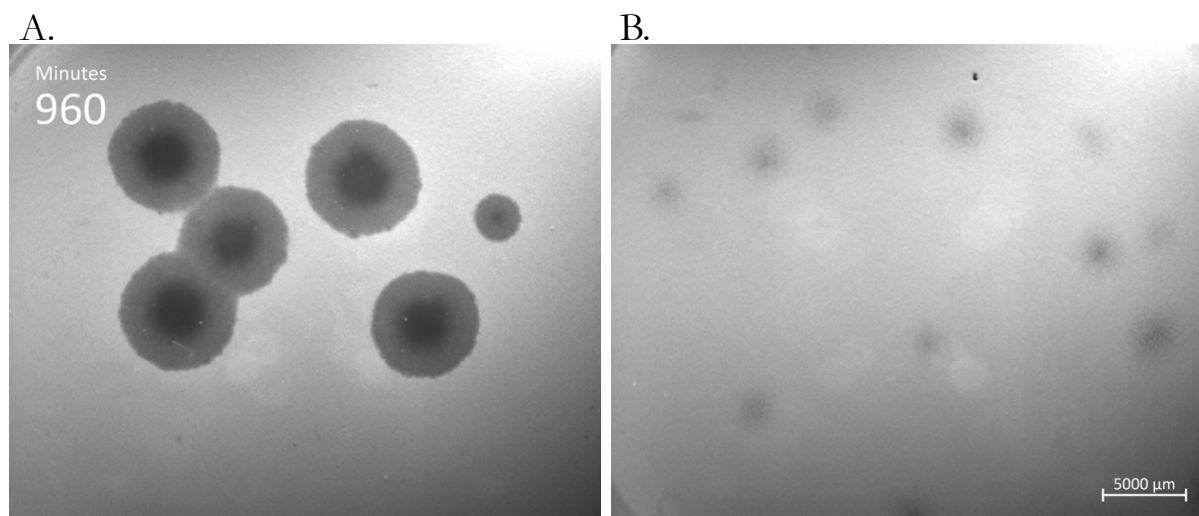
Fig 4.5. Ancestral and evolved plaque morphologies produced by ΦX174 before and after the phage evolution experiment – Standard. **A.** Ancestral "bull's eye" plaque morphotype produced by ΦX174 wildtype on a bacterial lawn of *E. coli* C wildtype after 16h of incubation at 37°C. **B.** Evolved plaque morphotype formed by an evolved ΦX174 population after 21 serial transfers (ΦX174 R7 T21) on a lawn of *E. coli* C wildtype after 16h incubation at 37°C. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number.

The appearance of plaques produced by phages can be influenced by many parameters, such as virion morphologies (sizes and shapes), adsorption rates, burst sizes, or the time spent by viruses inside bacteria before lysis (lysis time)[243,244]. During its evolution in liquid LB, ΦX174 might have increased its affinity for *E. coli* C wildtype but, in return, might have compromised its ability to diffuse through the bacterial population in the top agar overlay[243,245,246]. Adsorption for its host became stronger, leading to lower diffusion. A large proportion of the viral progeny likely adsorbed

and killed cells in their immediate vicinity, which slowed the expansion of the lysis zone (the plaque size). Bacterial hosts entered the stationary phase, stopping phage infection and lysis, resulting in the formation of small plaques.

4.3.3.   Increasing ΦX174 sequence diversity by UV-B mutagenesis

4.3.3.1.   Effect of UV-B radiations on the number of phages

To test the effect of UV radiations on phage populations, Christopher Böhmker (my intern and later a master student at this time) and I exposed 1 ml of three independent wildtype ΦX174's phage lysates to UV-B light at 302 nm during different periods ranging from 1 to 25 minutes (see **Chapter II** section **2.3.6.1**). The exposure time that inactivates ~90% of the overall phage population was used later to set up the UV-B treatment step carried out after each additional serial transfer (see **Chapter II** section **2.3.6.2**).

Exposing ΦX174 wildtype to UV-B radiations quickly reduced the number of phages, and this reduction was positively correlated with the exposure time (**Fig. 4.6A**). Viable phages could barely be detected after 16 minutes of exposure (**Fig. 4.6A**). We replicated this experiment with five independent populations of ΦX174 wildtype but only tested exposure times of 12, 13, and 14 minutes (see **Chapter II** section **2.4.6.1**). Exposing the phage population to 13 min of UV-B radiation led to the best result by significantly decreasing the mean number of phages by 88% (**Fig. 4.6B**).
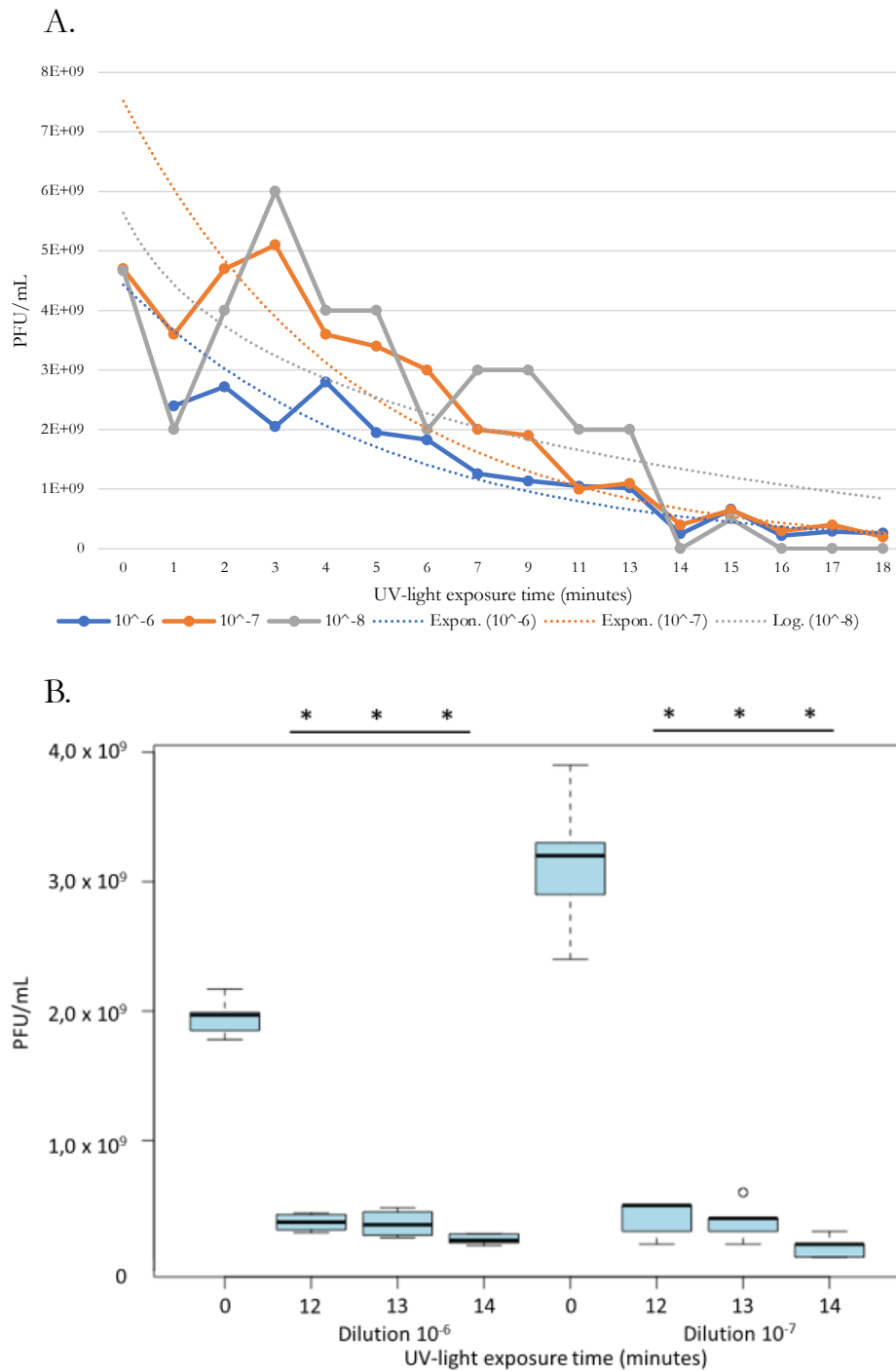
Figure 4.6. Effect of UV-B radiation on the number of ΦX174 (PFU count). **A.** The graph shows three different dilutions, which were analysed for their number of pfu as a function of UV-light exposure time. The solid lines indicate the corresponding dilution row (blue line: 10^{-6}; orange line: 10^{-7}; grey line: 10^{-8}). The dotted lines indicate the trendlines for each dilution (blue dotted line: 10^{-6}; orange dotted line: 10^{-7}; grey dotted line: 10^{-8}). **B.** The figure shows two different dilutions for five replicates, which were analysed for their number of pfu as a function of UV-light exposure

time. Each boxplot indicates the number of pfu units per phage lysate millilitre as a function of UV-light exposure time for five replicates. The asterisks indicate that the respective groups differ significantly ($p < 0.01$) from the untreated group of the respective dilution ($t = 0$). From Böhmker, C. 2020. *Unpublished internship report.*

4.3.3.2. Effect of the UV-B radiations on plaque morphology

I extended the evolution of the unsuccessful phage lineages by five more transfers and introduced a mutagenesis step. After each transfer, each phage lysate was exposed to UV-B radiation at 302 nm for 13 minutes (**Fig. 4.7**, see **Chapter II** section **2.3.6.2**). The three ΦX174 control lines from the phage evolution experiment – Standard also underwent the same protocol (**Fig. 4.7**).
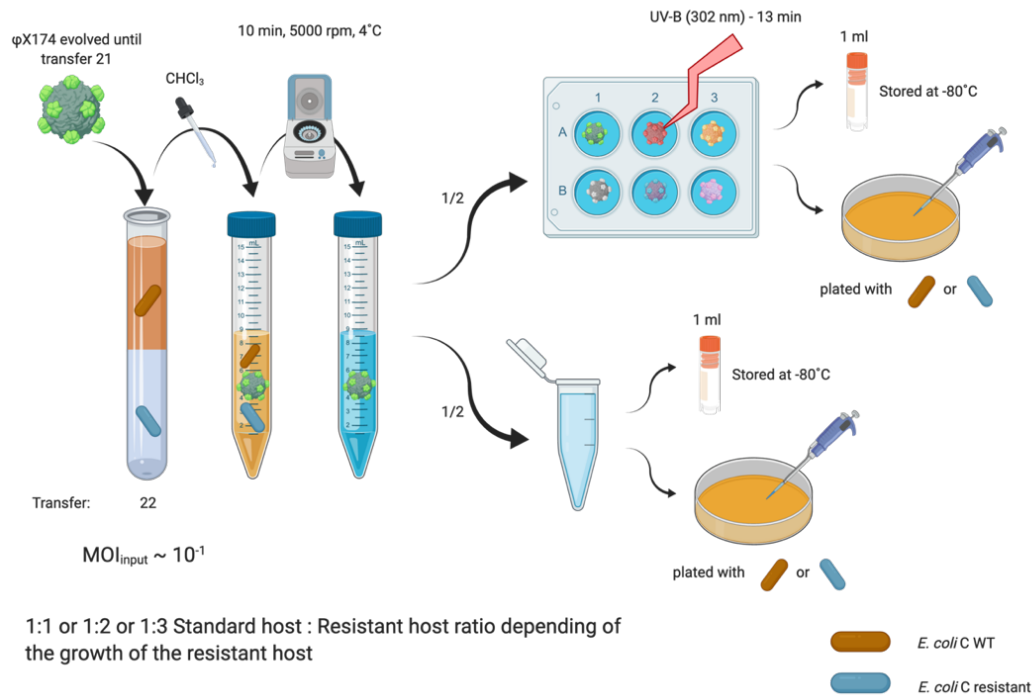
Fig. 4.7. Experimental evolution of ΦX174 under UV-B radiation. The first part of each transfer is identical to the phage evolution experiment – Standard (**Fig. 4.4**). At the end of each transfer, 2.5 ml of each purified phage lysate was exposed to UV-B light (302 nm, 100% intensity, UVP Chemstudio Plus ®) for 13 min. The three control lines were similarly exposed to UV-B light (see **Chapter II** section **2.3.6.2**).

After being exposed to UV-B light, a large proportion of phages from the evolved and control lines quickly reverted to the ancestral "bull's eye" plaque morphotype (**Fig. 4.8**). The cause for

this reversion is presumably a mutation that decreased phage adsorption to the bacterial outer membrane[247,248]. Adsorption became weaker, leading to a better diffusion of phages in the bacterial population. Phages could adsorb and kill bacteria further away from their initial location before hosts entered the stationary phase, which increased the size of their plaques[246,247,249].
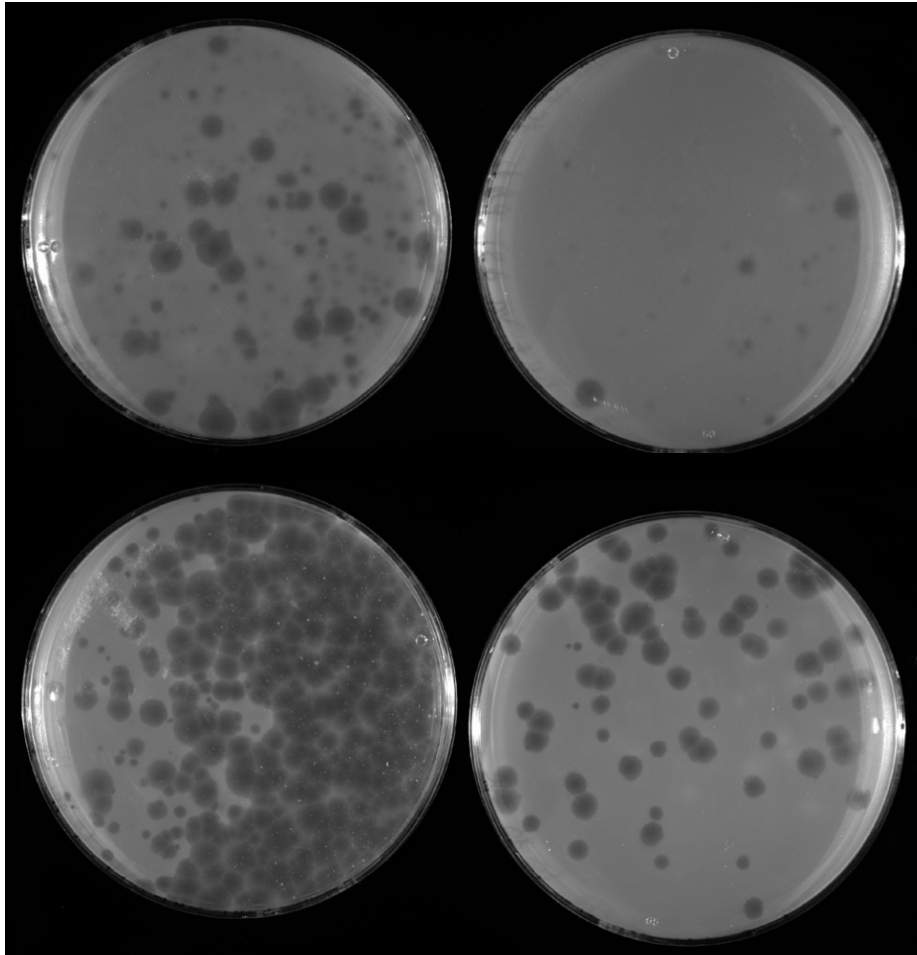
Fig 4.8. **ΦX174 populations reverted to the ancestral plaque morphotype after exposure to UV-B radiation.** Plaques produced by evolved ΦX174 after five additional transfers under UV-B light on a lawn of *E. coli* C wildtype. Top: ΦX174 R9 T26 UV; bottom ΦX174 R35 T26 UV; left: dilution 10^{-5}; right: dilution 10^{-6}. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number.

Unfortunately, no phage lineage evolved to infect their corresponding resistant strains at the end of transfer five. Thus, I tried another approach to increase the phage sequence diversity and designed two further evolution experiments to increase the chances of evolving phages capable of infecting the remaining resistant strains (**Fig. 4.9**).

4.3.4. Phage cocktail evolution experiments via serial transfers

I designed two more experiments to increase phage diversity. Diversity can increase not only through the acquisition of new mutations but also by recombining existing genotypes. In both experiments, I combined all evolved phage strains that successfully re-infected their corresponding resistant strains into a single phage cocktail[240,241]. The phage cocktail was serially transferred daily, once per day, in the phage cocktail evolution experiment – Increased diversity (**Fig 4.9A**). In the phage cocktail evolution experiment – Increased diversity and generations, four transfers per day were performed to shorten its duration (**Fig. 4.9B**). Phage diversity was maintained from one transfer to another by adding all corresponding resistant strains to the host culture (see **Chapter II** sections **2.3.8** and **2.3.9**).
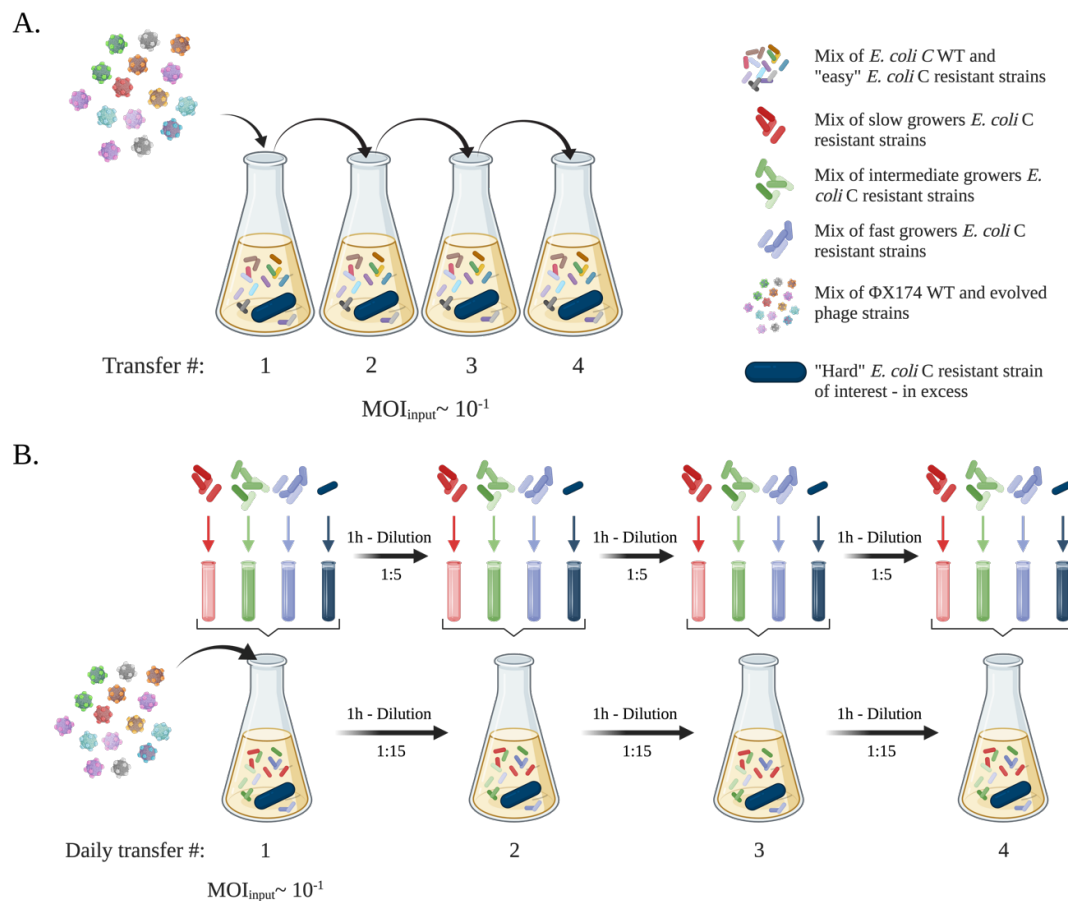
Fig. 4.9. Phage cocktail evolution experiments used to overcome *E. coli* C resistance by phage ΦX174. A. Phage cocktail evolution experiment – Increased diversity. A cocktail of ΦX174 wildtype and 14 previously evolved ΦX174 strains was generated. Despite being removed from the mutational analysis, evolved phages infecting R5 and R19, respectively, were part of the final cocktail. The phage cocktail was transferred daily for up to four days on non-evolving host cultures

containing (i) *E. coli* C wildtype, (ii) the 14 host strains for which resistance had been overcome (permissive hosts), and (iii) an excess of the still-resistant strain of interest (non-permissive host). Phages capable of infecting the resistant strain of interest were isolated from two lines. **B.** Phage cocktail evolution experiment – Increased diversity and generations. The same cocktail used for the previous phage cocktail evolution experiment (panel B) was transferred four times daily for four days (i.e., a maximum of 16 transfers) on non-evolving host cultures containing (i) *E. coli* C wildtype, (ii) the 14 strains for which resistance had been overcome (permissive hosts), and (iii) the still-resistant strain of interest (non-permissive host). Notably, all transfers completed on the same day involved transferring both phages and bacteria; supernatants were collected on every fourth transfer, and only phages were transferred. Phages capable of infecting the resistant strain of interest were isolated from the two lines. Details of the evolved phages are presented in **Table 4.2**, and further methodological details are provided in **Chapter II** sections **2.3.8** and **2.3.9**.

At the end of these two evolution experiments, all resistant strains could be infected or cross-infected by at least one of the evolved phage mutants (**Fig. 4.10**). A detailed analysis of the infection matrix is given in the next chapter (see **Chapter V** section **5.2**). My method of evolving phages with an extended infection range is simple and effective. It allowed the successful evolution of phages that, together, can infect all of the 31 resistant *E. coli* C phenotypes.
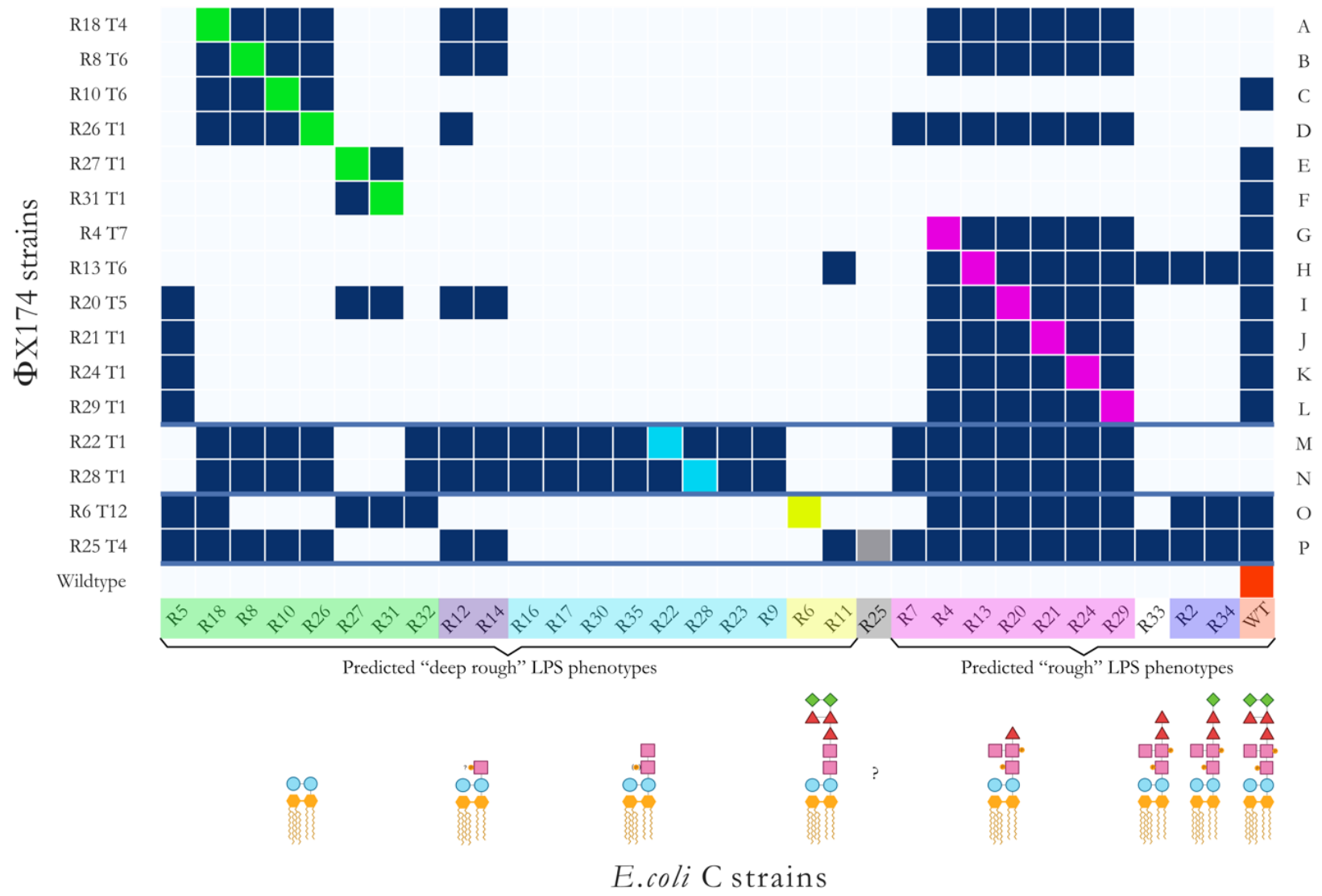
Fig. 4.10. Infection matrix of evolved ΦX174 phages on the 31 resistant *E. coli* C strains. The infection matrix was produced by combining the results of two spotting methods and plaque assays (see **Chapter II** section **2.4.1.1**). *E. coli* C strains are grouped and coloured based on their predicted core LPS structures[106,107,110,113,115,118–121,123–125,128,129,131,144,235]. ΦX174 strains are ordered based on the predicted LPS structures they infected during the evolution experiments (A-F: overcame heptoseless *waaC* and *hldE* mutants; G-L: overcame *waaO* mutants; M-N: overcame *galU* and *waaG* mutants; O: overcame *waaP/pssA* mutant; P: overcame *rfaH* mutant). The solid blue lines separate the evolved phages according to the evolution experiment in which they were isolated (A-L: Standard; M-N: Increased diversity; O-P: Increased diversity and generations). Dark blue squares = infection; light blue squares = no infection; coloured squares = control infection by a phage evolved on that host. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed. "?": core LPS structure of *E. coli* C R25 (*rfaH* mutant) could not be predicted. *E. coli* C R32 cannot be infected by any evolved phages from the phage evolution experiment – Standard. The phage infecting R5 was the only phage that could infect R32 (**Fig. 4.S1**) but was removed from the subsequent analysis due to a lack of isogeny

4.4.   Increasing host diversity accelerates ΦX174 adaptation

To identify the mutation(s) responsible for the re-infection of the resistant mutants, I performed whole genome re-sequencing on the 16 evolved phages (isogenic isolates) that successfully overcame the corresponding resistance in one of the single phage and phage cocktail evolution experiments (see **Chapter II** section **2.3.10.1**). I found mutations in *F* and/or *H* genes in all 16 (**Table 4.2**) and discussed their molecular details in depth in the subsequent section. Notably, all hard-resistant *E. coli* C strains – those whose resistance was overcome in the phage cocktail evolution experiments – are only infected by phage carrying four mutations (**Fig. 4.10** and **Table 4.2**).

The evolutionary emergence of a multi-mutation phage is expected to require a large number of transfers in a two-strain evolution experiment (**Fig. 4.11B**). However, this process is significantly accelerated by propagating a diverse mixture of phages that can overcome easy resistance, some of which may act as evolutionary intermediates (i.e., carry a subset of mutations required to overcome hard resistance). The mutational sets required to overcome hard resistance can subsequently emerge in a single mutational step, either by recombination[92] or by *de novo* mutation (**Fig. 4.11C**). For example, hard-resistant strain *E. coli* C R22 (a *galU* mutant; see **Table 3.1**) is overcome in the phage cocktail evolution experiment – Increased diversity by ΦX174 R22 T1, a phage that carries four mutations (three in *F*, one in *H*; see **Table 4.2**). Each of these mutations was present in the phage cocktail used to initiate the phage cocktail evolution experiment – Increased diversity: ΦX174 R18 T4 (which overcomes the easy resistance of *E. coli* C R18) contains three of the four mutations, and the final mutation is found in ΦX174 R10 T6 (which overcomes the easy resistance in *E. coli* C R10). Thus, adaptation to an easy resistant LPS structure (e.g., that of *E. coli* C R18) yields an evolutionary intermediate phage that can more easily evolve to infect a hard-resistant LPS structure (e.g., that of *E. coli* C R22).
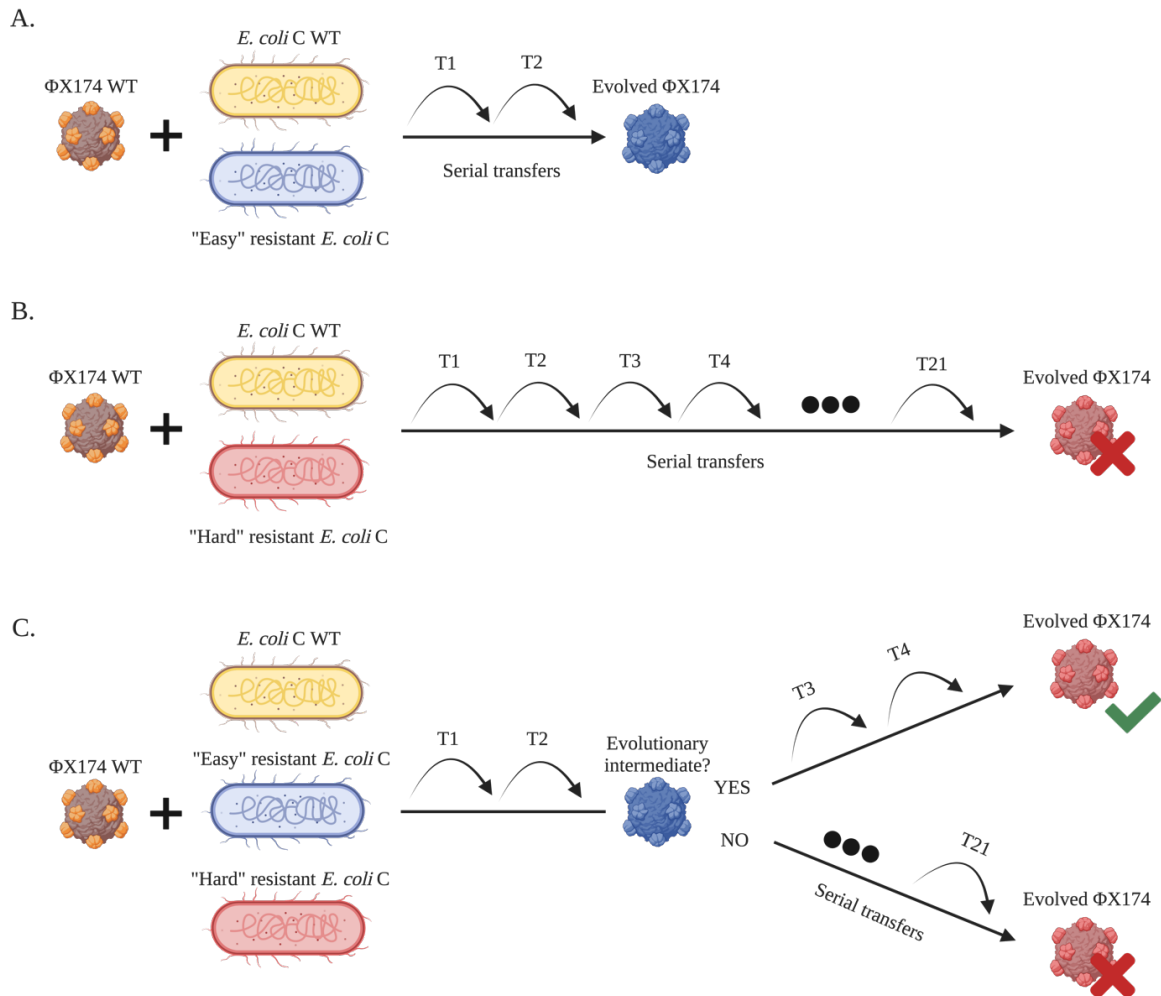
Fig. 4.11. Increased host diversity facilitates viral adaptation. A. ΦX174 wildtype evolves on *E. coli* C wildtype and an "easy" *E. coli* C resistant strain. Adaptation to the resistant mutant is fast and requires few transfers. **B.** ΦX174 wildtype evolves on *E. coli* C wildtype and a hard *E. coli* C resistant strain that it failed to infect within 21 standard serial transfers. Infecting this resistant strain may be difficult and take a long time. **C.** ΦX174 wildtype evolves in a three-host mix containing *E. coli* C wildtype and two *E. coli* C resistant strains: one easy and one hard. Phage adapts quickly to the easy-resistant phenotype. If the infection of the easy-resistant strain leads to the production of an evolutionary intermediate phage, then the hard-resistant phenotype can be infected after a few more transfers. If no evolutionary intermediate phage is produced, then no phage evolves to infect the hard-resistant *E. coli* C strain.

Increasing both phage and host diversities in the phage cocktail evolution experiments accelerated the adaptation of ΦX174 to resistant *E. coli* C strains. This result seemingly contradicts recent research by Sant *et al.*, which elegantly shows that higher host diversity decreases the rate at which

phages evolve[250]. By adapting phage ΦJB01 (a phage that likely uses the LPS to infect its *E. coli* host[250]) to a set of different *E. coli* strains, the authors demonstrate that adaptation to resistant (non-permissive) strains is faster in evolution experiments where ΦJB01 is exposed to only one permissive and one non-permissive host than when two non-permissive and one permissive strain are used. Similar to the two-strains experiment of Sant *et al.*[250], I evolved phage ΦX174 in the presence of one permissive and one non-permissive *E. coli* C strain in the phage evolution experiment – Standard (**Fig. 4.4**). While the experiment quickly gave rise to phages able to infect most non-permissive strains (**Figs. 4.10** and **4.11A**), some phage-resistant hosts could only be infected when both host and phage diversities were increased (**Figs. 4.10, 4.11B** and **4.11C**).

While adaptation to re-infect easy resistant host strains can be rapid under low host diversity (as shown in previous studies[137,250]), the results in this section demonstrate that it can be nearly impossible to infect hard-resistant *E. coli* C strains in the absence of additional hosts that allow the evolution of intermediate phages.

4.5. Mutational causes of ΦX174 adaptation to the resistant strains

4.5.1. Exclusion of two evolved ΦX174 from the final analyses

I found that the evolved phage infecting R5 was non-isogenic (**Tables 4.3** and **4.4**). Sanger sequencing showed that the glycerol stocks of phages isolated at the end of the phage evolution experiment – Standard are composed of at least two populations (c1 strains, **Table 4.4**). Re-isolating phages from the c1 strains corrected the lack of isogeny (see **Chapter II** section **2.3.10.2**), and the subsequent analyses were performed only with the re-isolated phage strains (c2 strains, **Table 4.4**). However, the lack of isogeny persisted for the re-isolated phage infecting R5 (ΦX174 R5 T2 c2; **Tables 4.4**). Therefore, the phage infecting R5 was removed from the final mutational and phenotypical analyses.

The evolved phage infecting R19 (ΦX174 R19 T1, **Table 4.3**) was also removed from the mutational and phenotypical analyses. Its corresponding resistant strain was the only one remaining sensitive to ΦX174 wildtype infection (see **Fig. 3.6**) and is predicted to have no LPS modification (**Table 3.2**). ΦX174 R19 T1 can only infect *E. coli* C wildtype and R19 in a semi-solid environment. Therefore, its substitution S284Y in the H protein is probably not involved in the host LPS recognition.

4.5.2. ΦX174 overcomes *E. coli* C resistance by mutations in the F capsid and H minor spike proteins

The remaining 16 evolved phages collectively carry 40 mutations, 15 of which are unique (**Fig. 4.12A and Table 4.2**). All are non-synonymous and occur in either *F* (encoding the viral capsid, with mutations occurring specifically in regions encoding the E-F and H-I loops[237]; the loops are involved in the irreversible adsorption of the capsid to the bacterium's cell surface[238]) or *H* (encoding the minor spike protein involved in viral DNA injection) genes (**Figs. 4.12B** and **4.12C**). Almost all evolved phages (13 of 16, or ~81%) carry multiple (between two and four) mutations (**Fig. 4.12D**). These results are consistent with previous studies demonstrating that mutations in *F* and *H* are relevant for ΦX174 adaptation to novel hosts[76,80,251].

Parallel evolution was common in my evolved phages, particularly in the control lines and the lineages (populations) that failed to infect their corresponding resistant *E. coli* C strains during the phage evolution experiment – Standard (**Table 4.5**). Of the three controls and 19 resistant lines that remained uninfected by their own phages after the phage evolution experiment – Standard, a total of 43 mutations were detected, but only 10 were unique. All these mutations are located in the *F* gene, except for one, which is located in the *A/A** genes (at genomic position 4817). The mutation in the *F* gene at the genomic position at position 2280 (causing amino acid change S427L in the F protein) arose in 16 lineages. Among them, eleven (69 %) carry between one to three additional mutations. High degrees of parallel evolution have also previously been observed in ΦX174 wildtype strains and other viruses[66,77,82,252,253] and is an indicator of either (i) a low number of mutational targets with high fitness gains, or (ii) mutational bias towards certain positions in the genome[147,253].
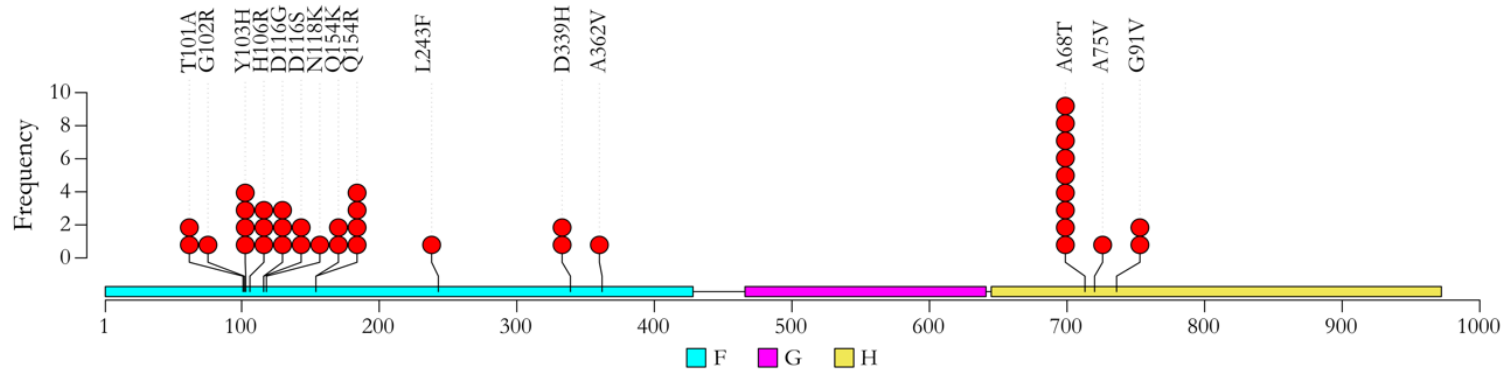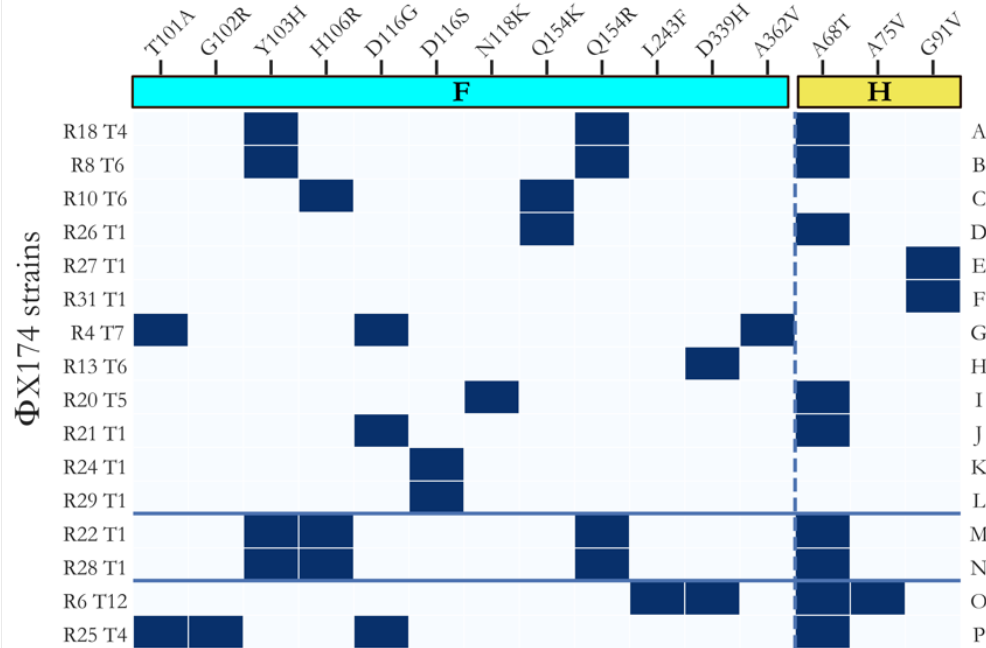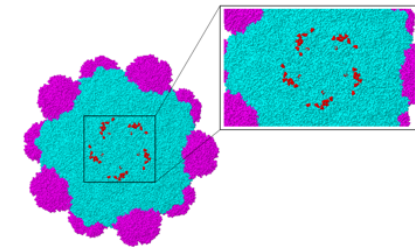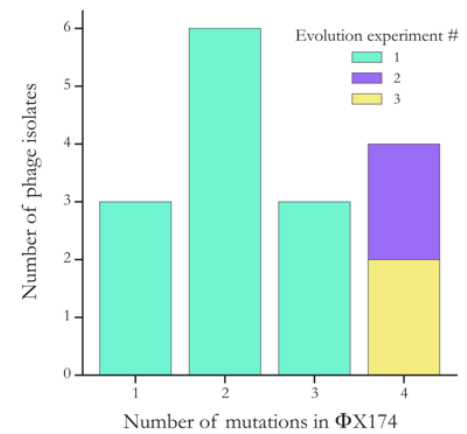
Fig. 4.12. ΦX174 adaptation to resistant strains of *E. coli* C. A. Positions of all predicted amino acid changes in the F and H proteins observed in the 16 evolved ΦX174 strains. **B.** Substitution matrix of all amino acid changes observed in ΦX174. ΦX174 strains are ordered based on the predicted LPS structures they infected during the evolution experiments (A-F: overcame heptoseless *waaC* and *hldE* mutants; G-L: overcame *waaO* mutants; M-N: overcame *galU* and *waaG* mutants; O: overcame *waaP/pssA* mutant; P: overcame *rfaH* mutant). The solid blue lines separate the evolved phages according to the evolution experiment from which they were isolated (A-L: Standard; M-N: Increased diversity; O-P: Increased diversity and generations). The x-axis shows amino acid substitutions in the F and H proteins (separated by the dashed line), R# indicates the resistant bacterial strain that the phage evolved on, and T# is the number of transfers required to overcome host resistance. **C.** Location of protein changes on the ΦX174 three-dimensional structure. The F capsid proteins are coloured in cyan, and the G major spike proteins are in magenta. Residues affected by the mutations in this study are in red and shown for one F capsid unit only. The G protein spike (above that pictured here) has been removed to improve visualisation. The image was generated using Geneious prime (version 2020.1.2) and Protein Data Bank accession code 2BPA[237]. Changes in the H protein cannot be displayed as they all occur outside the currently determined crystallographic structure[254]. **D.** Barplot of the number of evolved phage isolates carrying 1, 2, 3, 4 mutations. Most phage isolates – particularly those isolated from the phage cocktail evolution experiments – Increased diversity and – Increased diversity and generations – carry multiple mutations.

I also observed high levels of parallel evolution in the isogenic phage strains that evolved to infect *E. coli* C mutants with the same predicted LPS phenotype (**Figs. 4.12A** and **4.12B**). Parallelism was especially remarkable among phages infecting the four bacterial strains with the shortest (deep rough heptoseless) predicted LPS structure (see **Fig. 3.7**). The four phages infecting *E. coli* C R18, R8, R10, and R26 all carry an amino acid substitution at position 154 in the F protein (either Q154K or Q154R). In addition, the phage strains adapted to the *E. coli* C mutants R8 (*hldE* mutant) and R18 (*waaC* mutant) also carry the mutations Y103H in the F protein and A68T in the H protein. Finally, while *E. coli* C R27 and R31 both became resistant via the same amino acid substitution in HldE (G27A), their corresponding phages evolved similarly, each acquiring a different mutation that led to the same change at the protein level (G91V in the H protein).

Parallel evolution was also observed in phages infecting bacteria within several other classes of predicted LPS structure. For example, the evolved phages infecting *galU* mutant *E. coli* C R22 and *waaG* mutant *E. coli* C R28 carry an identical set of four mutations, one in the *H* gene and three in the *F* gene (**Table 4.2 and Fig. 4.12B**). Another example is the mutation at position 116 in the F protein (either D116G or D116S) that occurred in four different phages all infecting different *waaO* mutants (**Table 4.2 and Fig. 4.12B**).

ΦX174 strains	Evolution experiment #	Gene mutated in the corresponding *E. coli* C resistant strains	Mutants sensitive to the evolved ΦX174 strain infection	Gene mutated in ΦX174 strains	Descriptions	Positions	Nucleotide changes	Amino acid changes	Are mutations present in a control lineage?
R21 T1	1	*waaO*	R5, R4, R13, R20, R21, R24, R29, WT	*F*	Capsid protein	1,347	a→g	D116G	No
				H	Minor spike protein	3,132	g→a	A68T	No
R24 T1	1	*waaO*	R5, R4, R13, R20, R21, R24, R29, WT	*F*	Capsid protein	1,346	g→a	D116S	No
				F	Capsid protein	1,347	a→g	D116S	No
R26 T1	1	*hldE*	R18, R8, R10, R26, R12, R7, R4, R13, R20, R21, R24, R29	*F*	Capsid protein	1,460	c→a	Q154K	No
				H	Minor spike protein	3,132	g→a	A68T	No
R27 T1	1	*hldE*	R27, R31, WT	*H*	Minor spike protein	3,202	g→t	G91V	No
R29 T1	1	*waaO*	R5, R4, R13, R20, R21, R24, R29, WT	*F*	Capsid protein	1,346	g→a	D116S	No
				F	Capsid protein	1,347	a→g	D116S	No
R31 T1	1	*hldE*	R27, R31, WT	*H*	Minor spike protein	3,202	g→t	G91V	No
R18 T4	1	*waaC*	R18, R8, R10, R26, R12, R14, R4, R13, R20, R21, R24, R29	*F*	Capsid protein	1,307	t→c	Y103H	No
				F	Capsid protein	1,461	a→g	Q154R	No
				H	Minor spike protein	3,132	g→a	A68T	No
R20 T5	1	*waaO*	R5, R27, R31, R12, R14, R4, R13, R20, R21, R24, R29, WT	*F*	Capsid protein	1,354	t→a	N118K	No
				H	Minor spike protein	3,132	g→a	A68T	No
R8 T6	1	*hldE*	R18, R8, R10, R26, R12, R14, R4, R13, R20, R21, R24, R29	*F*	Capsid protein	1,307	t→c	Y103H	No
				F	Capsid protein	1,461	a→g	Q154R	No

				F	Capsid protein	3,132	g→a	A68T	No
R10 T6	1	*hldE*	R18, R8, R10, R26, WT	F	Capsid protein	1,317	a→g	H106R	No
				F	Capsid protein	1,460	c→a	Q154K	No
R13 T6	1	*waaO*	R11, R4, R13, R20, R21, R24, R29, R33, R2, R34, WT	F	Capsid protein	2,015	g→c	D339H	No
R4 T7	1	*waaO*	R4, R13, R20, R21, R24, R29, WT	F	Capsid protein	1,301	a→g	T101A	Yes (in C1 T21)
				F	Capsid protein	1,347	a→g	D116G	No
				F	Capsid protein	2,085	c→t	A362V	No
R22 T1	2	*galU*	R18, R8, R10, R26, R32, R12, R14, R16, R17, R30, R35, R22, R28, R23, R9, R7, R4, R13, R20, R21, R24, R29	F	Capsid protein	1,307	t→c	Y103H	No
				F	Capsid protein	1,317	a→g	H106R	No
				F	Capsid protein	1,461	a→g	Q154R	No
				H	Minor spike protein	3,132	g→a	A68T	No
R28 T1	2	*waaG*	R18, R8, R10, R26, R32, R12, R14, R16, R17, R30, R35, R22, R28, R23, R9, R7, R4, R13, R20, R21, R24, R29	F	Capsid protein	1,307	t→c	Y103H	No
				F	Capsid protein	1,317	a→g	H106R	No
				F	Capsid protein	1,461	a→g	Q154R	No
				H	Minor spike protein	3,132	g→a	A68T	No
R25 T4	3	*rfaH*	R5, R18, R8, R10, R26, R12, R14, R25, R11, R7, R4, R13, R20, R21, R24, R29, R33, R2, R34, WT	F	Capsid protein	1,301	a→g	T101A	Yes (in C1 T21)
				F	Capsid protein	1,304	g→c	G102R	No
				F	Capsid protein	1,347	a→g	D116G	No
				H	Minor spike protein	3,132	g→a	A68T	No

R6 T12	3	*waaP*	R5, R18, R27, R31, R32, R6, R4, R13, R20, R21, R24, R29, R2, R34, WT	F	Capsid protein	1,727	c→t	L243F	No
				F	Capsid protein	2,015	g→c	D339H	No
				H	Minor spike protein	3,132	g→a	A68T	No
				H	Minor spike protein	3,154	c→t	A75V	No

Table 4.2. List of mutations identified in ΦX174 evolved isolates. ΦX174 strains: R# is the corresponding resistant strain number, and T# is the transfer number where plaques were observed for the first time on the given resistant strain. Evolution experiment #: experiment number from which the evolved phages were isolated. Mutants sensitive to the evolved ΦX174 strains infection: list of all resistant strains that a specific phage strain can infect (See **Figs. 3.7 and 4.10**). Descriptions: names of proteins encoded by the listed genes. Positions: genomic coordinates of mutation (according to GenBank accession number AF176034.1). Nucleotide changes: observed nucleotide change. Amino acid changes: resulting change in amino acid sequence. "→": indicates 5'➔3' direction. Are mutations present in a control lineage?: indicates mutations observed in at least one control lineage (see **Chapter II** section **2.3.10.1**).

ΦX174 strains	Evolution experiment #	Gene mutated in the corresponding *E. coli* C resistant strains	Mutants sensitive to the evolved ΦX174 strain infection	Gene mutated in ΦX174 strains	Descriptions	Positions	Nucleotide changes	Amino acid changes	Are mutations present in a control lineage?
R19 T1	1	*yajC*	(R19), WT	*H*	Minor spike protein	3,781	c→a	S284Y	No
R5 T2	1	*gmhB*	R5, R18, R8, R10, R26, R27, R31, R32, WT	*H*	Minor spike protein	3,169	c→t	G80V	No

8

9 **Table 4.3. List of all evolved ΦX174 isolates excluded from the analysis.** ΦX174 strains: R# is the corresponding resistant strain number, and T#

10 is the transfer number where plaques were observed for the first time on the given resistant strain. Evolution experiment #: experiment number where

11 the evolved phages were obtained. Descriptions: names of proteins encoded by the listed genes. Positions: genomic coordinates of mutation (according

12 to GenBank accession number AF176034.1). Nucleotide changes: observed nucleotide change. "→": indicates 5'➔3' direction. Amino acid changes:

13 resulting change in amino acid sequence. Are mutations present in a control lineage?: indicates mutations observed in at least one control lineage (see

14 **Chapter II** section **2.3.10.1**).

15

16

ΦX174 strains clones	Transfer #	Gene mutated in the corresponding *E. coli* C	Gene mutated in ΦX174 strains	Descriptions	Positions	Nucleotide changes	Amino acid changes
R5 T2 c1	2	*gmhB*	*H*	Minor Spike protein	3,169	g→g or t	G80V
R5 T2 c2	2	*gmhB*	*H*	Minor Spike protein	3,169	g→g or t	G80V
			H	Minor Spike protein	3,202	g→g or t	G91V
R5 T2 c2B	2	*gmhB*	*H*	Minor Spike protein	3,169	g→t	G80V
R8 T6 c1	6	*hldE*	*F*	Capsid protein	1,307	t→c	Y103H
			F	Capsid protein	1,461	a→g or a	Q154R
R8 T6 c2	6	*hldE*	*F*	Capsid protein	1,307	t→c	Y103H
			F	Capsid protein	1,461	a→g	Q154R
			H	Minor Spike protein	3,132	g→a	A68T
R8 T6 c2B	6	*hldE*	*F*	Capsid protein	1,307	t→c	Y103H
			F	Capsid protein	1,461	a→g	Q154R
R10 T6 c1	6	*hldE*	*F*	Capsid protein	1,307	t→c	Y103H
			F	Capsid protein	1,626	g→a	R229H
			F	Capsid protein	2,016	a→g	D339G

R10 T6 c2	6	*hldE*	F	Capsid protein	1,317	a→g	H106R
			F	Capsid protein	1,460	c→a	Q154K
R10 T6 c2B	6	*hldE*	F	Capsid protein	1,317	a→g	H106R
			F	Capsid protein	1,460	c→a	Q154K
R13 T6 c1	6	*waaO*	F	Capsid protein	2,015	g→c	D339H
R13 T6 c2	6	*waaO*	F	Capsid protein	2,015	g→c	D339H
R13 T6 c2B	6	*waaO*	F	Capsid protein	2,015	g→c	D339H
R4 T7 c1	7	*waaO*	F	Capsid protein	1,301	a→g	T101A
			F	Capsid protein	1,347	a→g	D116G
			H	Minor Spike protein	3,111	g→a	V61I
R4 T7 c2	7	*waaO*	F	Capsid protein	1,301	a→g	T101A
			F	Capsid protein	1,347	a→g	D116G
			F	Capsid protein	2,085	c→t	A362V
R4 T7 c2B	7	*waaO*	F	Capsid protein	1,301	a→g	T101A
			F	Capsid protein	1,347	a→g	D116G

F	Minor Spike protein	3,111	g→a	V61I

17

Table 4.4. Genome troubleshooting of evolved ΦX174 isolates from the phage evolution experiment – Standard via Sanger Sequencing

ΦX174 strains: R# is the corresponding resistant strain number, and T# is the transfer number where plaques were observed for the first time on the given resistant strain. Evolution experiment #: experiment number where the evolved phages were obtained. c1: plaques isolated immediately during the evolution experiment on a lawn of the corresponding *E. coli* C resistant strains; c2: clones isolated from single plaques made by c1 growing on a lawn of the corresponding *E. coli* C resistant strains; c2B: second clones isolated from single plaques made by c1 growing on a lawn of the corresponding *E. coli* C resistant strains. Descriptions: names of proteins encoded by the listed genes. Positions: genomic coordinates of mutation (according to GenBank accession number AF176034.1). Nucleotide changes: observed nucleotide change. "→": indicates 5'➔3' direction. Amino acid changes: resulting change in amino acid sequence (see **Chapter II** section **2.3.10.2**).

ΦX174 populations	Gene mutated in the corresponding *E. coli* C resistant strains	Gene mutated in ΦX174 strains	Descriptions	Positions	Nucleotide changes	Amino acid changes	Frequencies (%)	Are mutations present in a control lineage?
C1 T21	None (WT)	*F*	Capsid protein	1,301	a→g	T101A	69,8	-
		F	Capsid protein	1,965	g→a	G322D	70,9	-
C2 T21	None (WT)	*F*	Capsid protein	1,968	a→g	N323S	17,8	-
		F	Capsid protein	2,277	c→t	T426I	56,7	-
C3 T21	None (WT)	*F*	Capsid protein	2,277	c→t	T426I	5,7	-
		F	Capsid protein	2,280	c→t	S247L	80,7	-
R2 T21	*waaW*	*F*	Capsid protein	1,965	g→a	G322D	19,0	Yes (in C1 T21)
		F	Capsid protein	2,276	a→g	T426A	30,4	No
		F	Capsid protein	2,279	t→c	S427P	10,9	No
		F	Capsid protein	2,280	c→t	S247L	30,9	Yes (in C3 T21)
R5 T21	*gmhB*	NA	NA	NA	NA	NA	NA	NA
R6 T21	*waaP*	*F*	Capsid protein	1,965	g→a	G322D	11,7	Yes (in C1 T21)
		F	Capsid protein	2,276	a→g	T426A	40,4	No
		F	Capsid protein	2,279	t→c	S427P	29,2	No
R7 T21	*waaO*	*F*	Capsid protein	2,279	t→c	S427P	5,4	
		F	Capsid protein	2,280	c→t	S247L	93,2	Yes (in C3 T21)
R9 T21	*waaG*	*F*	Capsid protein	2,280	c→t	S247L	100	Yes (in C3 T21)
R11 T21	*waaP*	NA	NA	NA	NA	NA	NA	NA

R12 T21	*waaF*	F	Capsid protein	2,280	c→t	S247L	100	Yes (in C3 T21)
R14 T21	*waaF*	F	Capsid protein	2,280	c→t	S247L	93,5	Yes (in C3 T21)
R16 T21	*galU*	F	Capsid protein	2,280	c→t	S247L	100	Yes (in C3 T21)
R17 T21	*galU*	F	Capsid protein	2,279	t→c	S427P	6,7	No
		F	Capsid protein	2,280	c→t	S247L	93,1	Yes (in C3 T21)
R22 T21	*galU*	F	Capsid protein	1,301	a→g	T101A	24,7	Yes (in C1 T21)
		F	Capsid protein	2,179	g→t	Q393H	19,7	No
		F	Capsid protein	2,279	t→c	S427P	25,4	No
		F	Capsid protein	2,280	c→t	S247L	54,5	Yes (in C3 T21)
R23 T21	*waaG*	F	Capsid protein	1,301	a→g	T101A	5,6	Yes (in C1 T21)
		F	Capsid protein	2,280	c→t	S247L	91,6	Yes (in C3 T21)
R25 T21	*rfaH*	F	Capsid protein	2,280	c→t	S247L	100	Yes (in C3 T21)
R28 T21	*waaG*	F	Capsid protein	2,279	t→c	S427P	6,9	No
		F	Capsid protein	2,280	c→t	S247L	93,0	Yes (in C3 T21)
R30 T21	*galU*	F	Capsid protein	2,279	t→c	S427P	7,9	No
		F	Capsid protein	2,280	c→t	S247L	90,0	Yes (in C3 T21)
R32 T21	*hldE*	F	Capsid protein	2,276	a→g	T426A	29,6	No
		F	Capsid protein	2,277	c→t	T426I	6,8	Yes (in C2 T21 and C3 T21)
		F	Capsid protein	2,280	c→t	S247L	61,9	Yes (in C3 T21)

R33 T21	*waaT*	F	Capsid protein	2,276	a→g	T426A	42,9	No
		F	Capsid protein	2,277	c→t	T426I	14,8	Yes (in C2 T21 and C3 T21)
		F	Capsid protein	2,280	c→t	S247L	41,7	Yes (in C3 T21)
		A/A*	Replication-associated protein/protein A*	4,817	c→t	D279D/D107D	12,8	No
R34 T21	*waaW*	NA	NA	NA	NA	NA	NA	NA
R35 T21	*galU*	F	Capsid protein	1,301	a→g	T101A	19,1	Yes (in C1 T21)
			Capsid protein	1,649	c→t	R217C	8,0	No
			Capsid protein	2,179	g→t	Q393H	6,8	No
			Capsid protein	2,280	c→t	S247L	82,9	Yes (in C3 T21)

27

28 **Table 4.5. List of mutations found in the evolved ΦX174 populations that did not infect their corresponding *E. coli* C resistant strains during**

29 **the phage evolution experiment – Standard.** ΦX174 populations: R# is the corresponding resistant strain number, and T# is the transfer number

30 where plaques were observed for the first time on the given resistant strain. Descriptions: names of proteins encoded by the listed genes. Positions:

31 genomic coordinates of mutation (according to GenBank accession number AF176034.1). Nucleotide changes: observed nucleotide change. Amino

32 acid changes: resulting change in amino acid sequence. "→": indicates 5'→3' direction. Frequencies (%): frequency of a given mutation in the phage

33 population. Are mutations present in a control lineage?: indicates mutations observed in at least one control lineage. NA: no data available (see **Chapter**

34 **II** section **2.3.10.1**).

35

4.6.   Discussion

4.6.1.   Increasing phage and host diversity allows the infection of hard LPS-resistant phenotypes

Another major finding of my study is that increasing phage and host diversity allows the infection of hard-resistant phenotypes (see **Fig. 4.11C**). Increased phage sequence diversity was obtained by pooling all evolved phages obtained during the phage evolution experiment – Standard in one phage cocktail. By knowing the mutational content of this phage cocktail used to start the phage cocktail evolution experiments (see **Chapter II** sections **2.3.8** and **2.3.9**), I was able to determine that some of these mutations (especially the ones obtained during the first adaption to the easy LPS mutants *waaC*, *hldE*, and *waaO*) were stepping stone mutations. These stepping stone mutations are required for the fixation of additional mutations (four in total), making accessible the re-infection of the *waaP/pssA*, *waaG*, *galU*, and *rfaH* resistant strains (**Table 4.2**, **Fig. 4.12B**).

This phenomenon – the successful infection of the hard-resistant phenotypes that became accessible by increasing phage and host diversity – is similar to another phenomenon observed in a co-evolution experiment between a different host-phage pair[255]. Phage λ also requires at least four mutations to infect the novel OmpF receptor in *E. coli* B, making it almost impossible to evolve from the wildtype phage in a single step[256]. Evolving all four mutations in one phage λ genome first required the presence of λ phage evolutionary intermediates that already carried specific mutations (see **Fig. 4.11C**). However, Meyer and colleagues did not speculate on how these λ phage evolutionary intermediates first arose in the phage population.

Similar to what I observed during ΦX174 evolution to the hard-resistant strains, the λ phage evolutionary intermediates may arise from the infection of easy-resistant *E. coli* B mutants (the *malT* loss-of-function mutants) that could still spontaneously induce low levels of the traditional host receptor, LamB. By increasing the host diversity, the phage λ eventually evolved the additional mutations (four in total) to overcome the hard-resistant *E. coli* B mutants (*lamB* loss-of-function mutants)[255,257]. My phage cocktail evolution methods could be used to quickly identify the mutations necessary for the later re-infection of various resistant bacterial strains and speed up the phage adaption process (see **Fig. 4.11C**), not only for the phage λ-*E. coli* B model presented above, but also for other host-phage pairs.

4.6.2. Specific substitutions are linked to the adaptation of ΦX174 to specific modified LPS structures

Through my evolution experiments, I showed that substitutions occurred at specific amino acid sites during ΦX174 adaptation to specific predicted LPS phenotypes (**Fig. 4.12B**, **Tables 4.2**, **4.3** and **4.5**). For instance, both substitutions Y103H and Q154K/R in the F capsid were found associated in multiple ΦX174 strains that re-infected the predicted heptoseless (shortest LPS) strains (**Fig. 4.12B**) and happened to be stepping stones in overcoming the *galU/waaG* hard-resistant phenotypes (**Fig. 4.11C**).

Similarly, the substitution at D116 (an amino acid at the immediate vicinity of known protein-protein interaction sites[74,258]) often arose in ΦX174 evolved strains that re-infected specifically *waaO* bacterial mutants (**Figs. 4.12B** and **4.12C**). Interestingly, this amino acid was either found to be mutated alone or was associated with other substitutions (especially D116S). While this amino acid site could be directly involved in the host recognition of the predicted *waaO* LPS structure, it is possible that, in some cases, this mutation was compensatory and helped stabilize the modified F protein structure[259].

I also found that substitutions T426A/I and S427L/P in the F protein were massively enriched in ΦX174 lineages that failed to adapt to their corresponding *E. coli* C resistant strains (**Table 4.5**). While these substitutions may not participate in infecting previously resistant *E. coli* C strains, they are more likely an effect of adaptation to growing in liquid *E. coli* C wildtype cultures and may be under selection for the wildtype host recognition and adsorption to the native LPS structure.

A large proportion of the substitutions occurred in the minor spike protein H of almost every evolved ΦX174 strains, regardless of the predicted modified LPS structure that has been overcome (11 out of 15, ~73%, **Fig 4.12B**). Out of these 11 phage strains with at least one substitution in the H protein, nine (~82%) carry the substitution A68T (**Fig 4.12B**). These substitutions should be localized in one of the predicted transmembrane domains, which anchor the assembled tube-like structure into the inner and outer cell membranes during the DNA ejection process[254]. Previous work has demonstrated that both the hexose residues of the outer-core LPS and the lipid A part of the LPS play a role in triggering conformational changes of the H proteins to initiate the DNA ejection[218,219,234]. While the observed substitutions in the H protein could be compensatory, the lack of specificity for one particular LPS phenotype could also suggest that ΦX174 adapt to

the new balance between hydrophobic and hydrophilic residues and the better-exposed lipid A, rather than the core LPS sequence[219].

To clearly assess the role of each phage mutation in host recognition, future experiments should aim to engineer phage mutants with all possible single and multiple combinations of mutations via site-directed mutagenesis and test their infectivity toward each *E. coli* C. In this manner, the contingency and epistasis effects between mutations could help better predict ΦX174 evolutionary pathways to overcome resistance.

For example, Y103H, Q154K and Q154R phage mutants can test the role of these substitutions in the specific infection of the heptoseless LPS strains (strains with the shortest LPS structure, see **Fig. 3.7**). The combination of Q154K/R (in the F protein) and A68T (in the H protein) can test if only these two substitutions mediate the infection of the *waaG* and *galU* mutants or if all four are needed (**Fig. 4.12B**). Similarly, the D116G/S phage mutants can be generated to (i) specify the role of these substitutions in the recognition of the *waaO* mutants and (i) see if they are involved in the acquisition of the additional mutations found in the evolved phages infecting the *waaP/pssA* and *rfaH* mutants. To ensure that an absence of lysis is not due to viruses' inactivation, phage infection will be tested against *E. coli* C wildtype. However, this approach will be difficult to use if phage mutants lose their ability to infect *E. coli* C wildtype and/or other resistant strains.

4.6.3. Conclusion

In a series of evolution experiments, I successfully bred the highly LPS-specific phage ΦX174 to infect a wide variety of resistant mutants with modified LPS structures. I showed that ΦX174 overcomes *E. coli* C resistance via specific mutations in the F capsid and H minor spike proteins and that increasing host and phage diversities could accelerate the adaptation of phage ΦX174 under certain circumstances. Finally, parallel evolution in phages that evolved to infect *E. coli* C strains with the same predicted LPS structure generally supported the predicted classes of LPS structure. However, analyses in the following chapter (see **Chapter V**) demonstrate that LPS structure predictions generally underestimate LPS diversity.

In the next chapter, I challenge the current LPS model based on the knowledge acquired from single gene deletions and complementation approaches. Using the high specificity of my evolved ΦX174 strains for their host LPS structures, I turned them into efficient "biosensors" to differentiate between LPS phenotypes.

4.7. Supplementary Figure
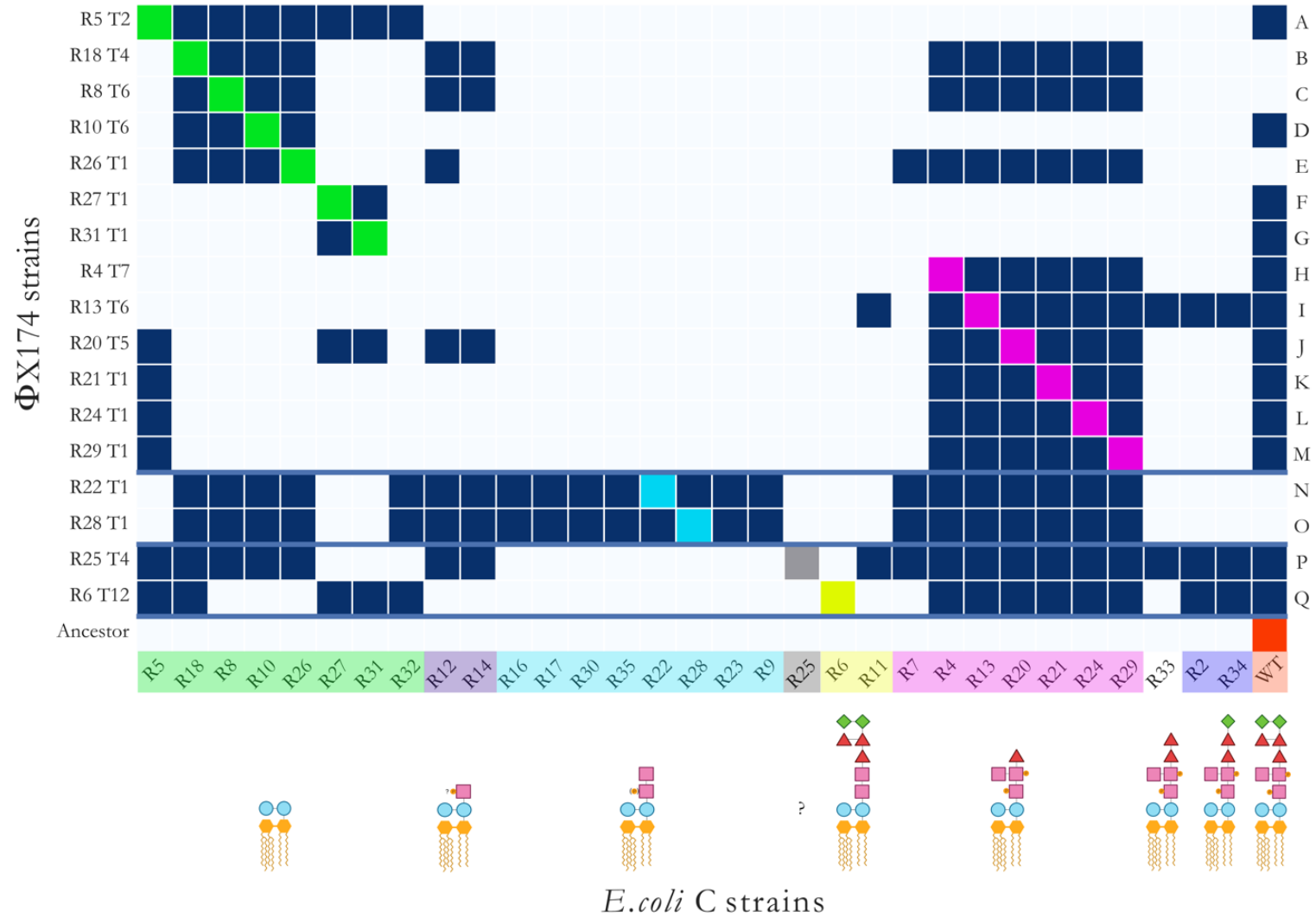
Fig. 4.S1. Infection matrix of evolved ΦX174 phages (including the evolved phage infecting R5) on the 31 resistant *E. coli* C strains. The infection matrix was produced by combining the results of two spotting methods and plaque assays (see **Chapter II** section **2.4.1.1**). *E. coli* C strains are grouped and coloured based on their predicted core LPS structures[106,107,110,113,115,118–121,123–125,128,129,131,144,235]. ΦX174 strains are ordered based on the predicted LPS structures they infected during the evolution experiments (A-G: overcame heptoseless *waaC* and *hldE* mutants; H-M: overcame *waaO* mutants; N-O: overcame *galU* and *waaG* mutants; P: overcame *rfaH* mutant Q: overcame *waaP/pssA* mutant). The solid blue lines separate the evolved phages according to the evolution experiment in which they were isolated (A-M: Standard; N-O: Increased diversity; P-Q: Increased diversity and generations). Dark blue squares = infection; light blue squares = no infection; coloured squares = control infection by a phage evolved on that host. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed. "?": core LPS structure of *E. coli* C R25 (*rfaH* mutant) could not be predicted.

Chapter V – Challenging the predicted LPS genotype – phenotype maps with phages

5.1. Evolved ΦX174 as biosensor tools to test the current understanding of LPS biology

My third and final objective was to challenge the "LEGO® block" principle of LPS biosynthesis and assembly. The "LEGO® block" principle is a simplified approach I used to predict the LPS genotype to phenotype maps of each *E. coli* C resistant strain. LPS structures are predicted solely based on what genes are mutated. I assumed that mutations that likely have an effect comparable to a gene deletion (i.e., frame shifts and non-sense mutations) and mutations that simply cause the substitution of an amino acid lead to the complete absence of the mutated LPS gene and result in the production of a single LPS phenotype. Therefore, there is only a very limited variety of possible LPS structures which can be generated.

Following this principle, I predicted that my 31 *E. coli* C mutants collectively produce eight distinct LPS structures based on the position of each mutation (see **Fig 3.7**). If my predictions are correct, an evolved phage infecting a resistant strain with a particular LPS phenotype should be able to cross-infect all other resistant strains sharing the same predicted LPS structure regardless of the underlying mutation (see **Fig 3.7**). Hence, only eight distinct infectivity profiles should be observed after testing the infection of each *E. coli* C resistant strain with each of the 16 evolved phages. Otherwise, it would indicate that the underlying mutation leads to a different LPS structure.

The resistance/sensitivity of the 31 *E. coli* C resistant strains against the evolved phage strains was tested via cross-infection assays in a semi-solid environment and revealed 14 distinct profiles instead of eight. Phages that evolved to infect the same predicted LPS structure were not always able to cross-infect each other's host, suggesting that the diversity of LPS structures generated by LPS-pathway mutations is higher than previously anticipated.

5.2. Evolved phages can be used to discriminate between LPS phenotypes

To determine precisely which of the 16 evolved phages can infect which of the 31 initially resistant *E. coli* C strains (i.e., the infection pattern of each phage), I first performed a cross-infection assay (**Fig. 4.10**). It involved measuring the infectivity of all possible host-phage combinations by two distinct methods: (i) phage spotting on host lawns and (ii) host spotting on phage lawns. In cases where a mismatch between the two results was observed, standard plaque assays were used as a tie-breaker (see **Chapter II** section **2.4.1.1**). If my predictions – based solely on what genes are mutated – of seven distinct LPS structures among the 31 initially resistant *E. coli* C strains are accurate (the final structure of R25 could not be predicted, see **Fig 3.7**), then all phages that evolved to infect a mutant with a given predicted LPS structure are expected to cross-infect all other host strains with the same predicted LPS structure. That is, the infection patterns in **Fig. 4.10** would be identical for each class (colour) of *E. coli* C mutant.

To visually determine whether infection patterns cluster according to the predicted LPS structures, a hierarchical agglomerative clustering analysis was performed on the infection matrix in **Fig. 4.10**. Of the six predicted LPS structures that occur in at least two *E. coli* C mutants, three form unbroken clusters with respect to infectivity patterns (highlighted in cyan, light purple, and dark purple; **Fig. 5.1**), and three do not (highlighted in magenta, yellow, and green; **Fig. 5.1**). I discussed each group in more detail below.

LPS predictions that cluster by infection pattern. The isogenic ΦX174 mutants R22 T1 and R28 T1 that evolved to infect resistant strains with mutations in *galU* (*E. coli* C R22) and *waaG* (*E. coli* C R28) can cross-infect all eight *E. coli* C mutants predicted to exhibit the same LPS structure (R9, R16, R17, R22, R23, R28, R30, R35; highlighted in cyan in **Figs. 4.10** and **5.1**). Furthermore, both *E. coli* C *waaW* mutants (R2 and R34; light purple) are cross-infected by the same three evolved phage genotypes (R6 T12, R13 T6, R25 T4). Also, both *E. coli* C *waaF* mutants (R12 and R14; dark purple) are infected by a near-identical set of seven (R12) or six (R14) phages. In each of these cases, phage infectivity patterns appear to be a good indicator of host LPS phenotype.
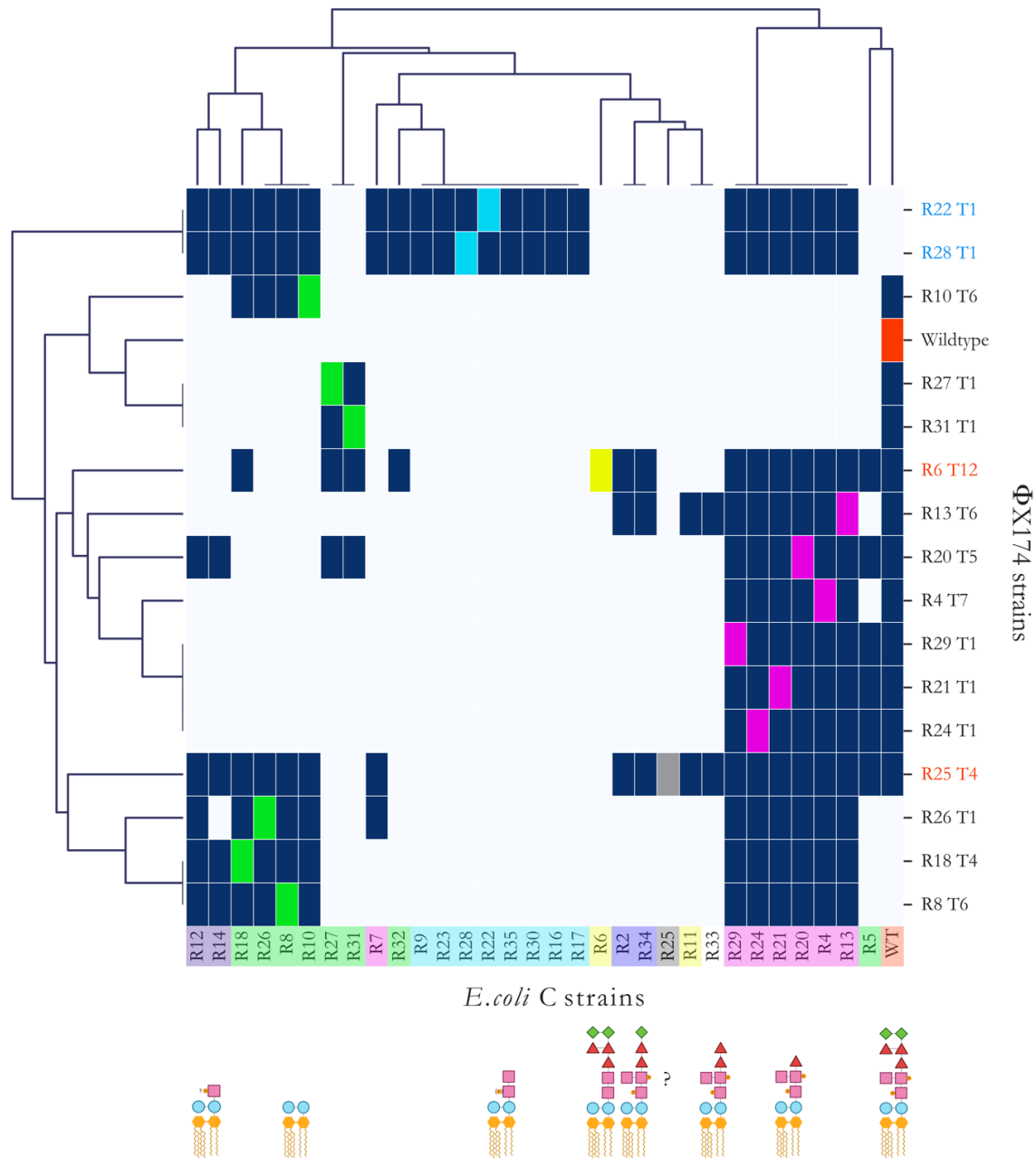
Fig. 5.1. Hierarchical agglomerative clustering analysis of the host ranges of evolved phage. To unravel the structures in the heatmap infection matrix (**Fig. 4.10**), I performed a clustering analysis using the default settings of the Seaborn library (version 0.11.2) for Python (version 3.7.4). Briefly, each data point is considered as a cluster. The linkage method computes the distance between them, then repeatedly combines the two nearest clusters into larger clusters, until a single cluster is left. *E. coli* C strains are coloured based on their predicted core LPS structure (**Fig. 3.7**). Phage names in black were obtained during the first 21 serial transfers (phage evolution experiment – Standard). Phage strains in blue were obtained in the phage cocktail evolution experiment – Increased diversity (one transfer per day). Phage strains in red were obtained in the

phage cocktail evolution experiment – Increased diversity and generations (four transfers per day). Dark blue squares = infection; light blue squares = no infection; coloured squares = control infection by a phage evolved on that host. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed. "?": core LPS structure of *E. coli* C R25 (*rfaH* mutant) could not be predicted.

LPS predictions that do not cluster by infection pattern. Six of the seven *E. coli* C *waaO* mutants (highlighted in magenta in **Figs. 4.10** and **5.1**) can be cross-infected by the same set of 13 phages. The seventh *waaO* mutant (*E. coli* C R7) was the only *waaO* mutant for which the corresponding phage lineage did not evolve to infect the resistant mutant in the phage evolution experiment – Standard. It can only be infected by the phage infecting R26 and phages that evolved in the phage cocktail evolution experiments – Increased diversity and – Increased diversity and generations (**Fig. 5.1**). Hence, it is likely that the LPS structure of R7 differs from that of the other *waaO* mutants. The difference could conceivably result from residual (or altered) WaaO function; *E. coli* C R7 carries an in-frame deletion of a 6-bp repeat in *waaO*, which alters two WaaO residues, but leaves the remainder of the protein intact (see **Table 3.1**). Because R7 is close to the *galU*/*waaG* mutants (**Fig. 5.1**), it may be possible that its LPS structure resembles the ones of the *galU*/*waaG* mutants. However, more detailed LPS studies are needed to elucidate the LPS structure of R7.

The infection patterns of the two *waaP* mutants *E. coli* C R6 and R11 (highlighted in yellow in **Figs. 4.10** and **5.1**) also suggest that, in contrast to the prediction, these strains possess different LPS structures. R6 is infected by a single phage that evolved directly on host R6 (phage R6 T12 from the phage cocktail evolution experiment – Increased diversity and generations), while R11 is cross-infected by two different phages (phage R13 T6 from the phage evolution experiment – Standard, and R25 T4 from the phage cocktail evolution experiment – Increased diversity and generations). The cluster map analysis indicates that R11 displays an LPS phenotype comparable to the rough LPS phenotype of R33, a *waaT* mutant (**Fig. 5.1**). The R6 mutation in *waaP* affects the LPS structure differently, possibly by affecting the number of phosphorylated hexoses in the inner core LPS (**Fig. 3.7**). The exact nature of the R6 LPS phenotype is difficult to infer since only a single phage evolved to infect R6 at the very end of the phage cocktail evolution experiment – Increased diversity and generations.

The combination of phages that can infect each of the eight deep rough *E. coli* C mutants predicted to have the shortest LPS structure (green highlighting in **Figs. 4.10 and 5.1**) is remarkably

different. These *E. coli* C strains cluster into four different phenotypic sub-groups (**Fig. 5.1**). The first sub-group consists of *E. coli* C R8, R10, R18, and R26. Each of these strains can be cross-infected by their corresponding evolved phages, regardless of their genotypes (a mixture of *waaC* or *hldE* mutants). However, none of the four evolved phages can cross-infect *E. coli* C R27 or R31 (*hldE* mutants). *E. coli* C R27 and R31 comprise the second sub-group. Both *E. coli* C R27 and R31 and their corresponding evolved phages acquired identical genotypes and, not surprisingly, they also share phenotypes. *E. coli* C R5 (*gmhB* mutant) and *E. coli* C R32 (*hldE* mutant) fall into a third and a fourth sub-groups. *E. coli* C R32 is the only deep-rough mutant for which the corresponding phage lineage failed to infect during the phage evolution experiment – Standard (see **Figs. 4.10** and **4.S1**). Only a very narrow but different set of evolved phages can infect R5 and R32 (**Figs. 5.1** and **5.S1**).

A total loss of function in *hldE* or *waaC* genes always leads to a complete truncation of the inner core LPS (**Fig. 3.2A**). Thus, the involvement of these two genes in the production of the deep rough phenotypes (highlighted in green in **Figs. 4.10** and **5.1**) cannot explain the vast phenotypic diversity in LPS structures. For example, *E. coli* C R8, R10, R18, and R26 belong to the same phenotypic group despite having both *hldE* or *waaC* mutated, while R27, R31, and R32 (all *hldE* mutants) do not belong to this group (**Table 3.1**, **Fig. 5.1**). Instead, the location and type of mutation are more likely to cause the observed LPS structure diversity.

The involvement of the *gmhB* gene in *E. coli* C R5, however, might be involved in the production of a different deep rough phenotype. Previous research demonstrated that the deletion of *gmhB* only causes a partial defect in the synthesis of the LPS core, resulting in the formation of a heptoseless and a heptose-rich form[124]. Because the defect caused by a deletion of *gmhB* is only partial and can potentially be compensated, it is unclear whether any mutation in *gmhB* leads to the same mix of a unique deep rough and the wildtype phenotypes or if multiple deep rough phenotypes can appear and persist.

The "LEGO® block" principle is a useful model to predict LPS structures when no further information on all structural alterations and how they can be regulated are available[109]. However, LPS structures predicted based only on what genes are mutated are not always accurate. Not all mutations have an effect comparable to a knock-out mutation as previously assumed but can have much more subtle effects on LPS structure that can be demonstrated an infection matrix (**Figs. 4.10** and **5.1**).

5.3. Evolved bacteria can be used to link phage genotypes to infection phenotypes

Phages' infectivity patterns (host-ranges) can be used to construct phage genotype-phenotype maps. For example, every phage strain that loses the ability to infect *E. coli* C wildtype has a combination of two mutations: A68T in the H protein and Q154K/R in the F protein (**Figs. 4.12 & 5.1**). In isolation, A68T confers the ability to infect *waaO* mutants and Q154K/R to infect [R8, R18, R10, R26]-heptoseless mutants. While the two mutations combined allow the phage to infect both heptoseless and *waaO* mutants, the phage loses the ability to infect the *E. coli* C wildtype (**Figs. 4.12 & 5.1**). ΦX174 R25 T4 is the only phage that can infect all three genotypes: *waaO*, the [R8, R18, R10, R26]-heptoseless mutants, and *E. coli* C wildtype. Presumably, this is because this phage does not contain the Q154K/R mutation and instead uses another set of mutations that allows R25 to infect [R8, R18, R10, R26]-heptoseless mutants (**Table 4.2**).

There are also other ways to infect bacteria with mutations in the *waaO* gene. For example, evolved phages adapted to R4, R21, R24 and R29 carry a D116G/S mutation in the F protein and can infect *waaO* mutants. Similarly, phages adapted to R13 and R20 evolved, respectively, a D339H and N118K mutation in the F protein (**Fig. 4.12B**) and can cross-infect *waaO* mutants (except R7). Hence, D116G/S, N118K or D339H mutations in the F protein alone or associated with A68T in the H protein are probably sufficient to infect *waaO* mutants (except R7). These results show that overcoming *waaO* phenotypes can be achieved through multiple evolutionary pathways.

Unlike bacteria, phage phenotypes are usually underpinned by more than one mutation (~81% of evolved phages carry two or more mutations, see **Fig. 4.12B**), which makes it very hard to understand the role of the mutation (or combination mutations) in the underlying infection phenotype. To gain insights, many more phage mutants have to be evolved, for example, by repeating the different evolution experiments multiple times. If, however, an infection phenotype is always caused by a combination of different mutations, it will be almost impossible to pin down the function of each mutation (due to combinatorial effects).

5.4. ΦX174 adaptation to a hard-resistant *E. coli* C mutant expanded its host range

Mutations acquired during ΦX174 adaption to the different *E. coli* C resistant strains may allow the infection of other *E. coli* strains. To identify the mutation or combinations of mutations responsible for phage infection, I tested whether any evolved phages can infect any of the four chosen *E. coli* strains (*E. coli* B REL606 and the K-12 strains MG1655, AQ4591, BZB1011), following the *method 1* and *standard plaque assay*'s protocols described in **Chapter II** section **2.4.1.1**. All *E. coli* strains harbour a rough LPS-type structure that cannot be used by ΦX174 wildtype (**Fig. 5.2**). However, the phage can use the different *E. coli* strains' cell machinery to initiate its replication[107,110,260,261]. Consequently, bacterial lysis will indicate that the evolved ΦX174 strain can use the novel LPS receptor.
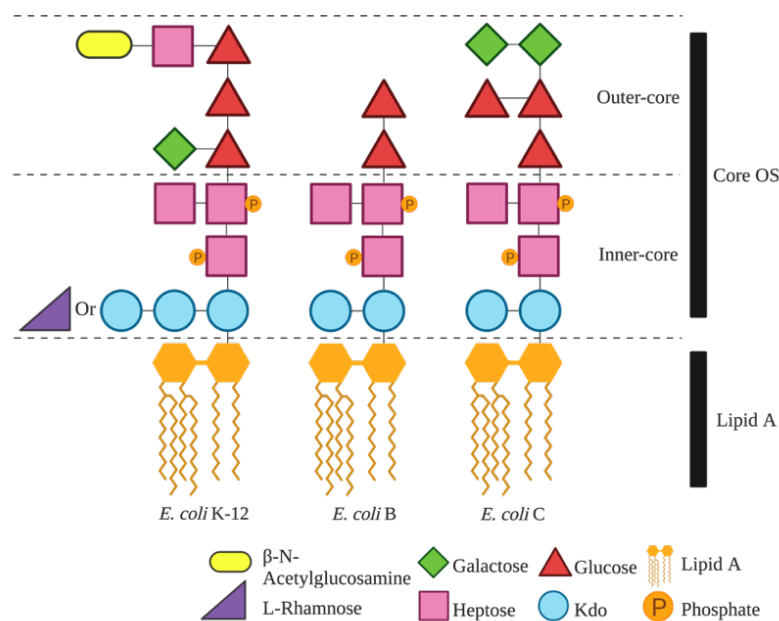
Fig 5.2. LPS structures of three *E. coli* strains. *E. coli* K-12, *E. coli* B, and *E. coli* C display a rough LPS-type structure composed of two parts: lipid A (composed of an acetylated and 1,4'-diphosporylated β(1→6)-linked glucosamine (GlcN) disaccharide), and the core oligosaccharide[107]. The core oligosaccharide is subdivided into a structurally conserved inner core and an outer core that variate between strains. Kdo: 3-deoxy-D-manno-octulosonic acid; Heptose: L-glycero-D-manno-heptose; Glucose: D-glucose; Galactose: D-galactose

All *E. coli* K-12 strains remained uninfected by any evolved phage, meaning that none of the mutations present in the evolved phage genomes is involved in the recognition of *E. coli* K-12's LPS structure. The outer core LPS structure of *E. coli* K-12 belongs to a completely different type than the one of *E. coli* C wildtype[110,113,119] and may require completely different (or additional) mutations to be used by the phage[251,262].

E. coli B REL606 could be infected by the evolved phage infecting R25 (**Fig. 5.3**), meaning that one or more mutations in the evolved phage genome are involved in the recognition and adsorption to *E. coli* B REL606's LPS. Compared to *E. coli* C wildtype, *E. coli* B REL606's outer core part LPS is truncated (**Fig. 5.2**). *E. coli* B REL606 *waaT* gene is interrupted by an IS1 insertion[261], resulting in a truncation of the outer core LPS similar to an *E. coli* C Δ*waaT* (**Fig. 3.7**). The evolved phage infecting R25 can infect R33, a spontaneous *E. coli* C *waaT* mutant (**Figs. 4.10** and **5.1**). Therefore, the mutations that accumulated during ΦX174 adaption to R25 not only allow the phage to re-infect several *E. coli* C LPS mutants (**Figs. 4.10** and **5.1**), but also increase its host range[263]. The phage is now able to infect an *E. coli* strain that shared its LPS structure with one of the *E. coli* C LPS mutants (**Fig. 5.3**).
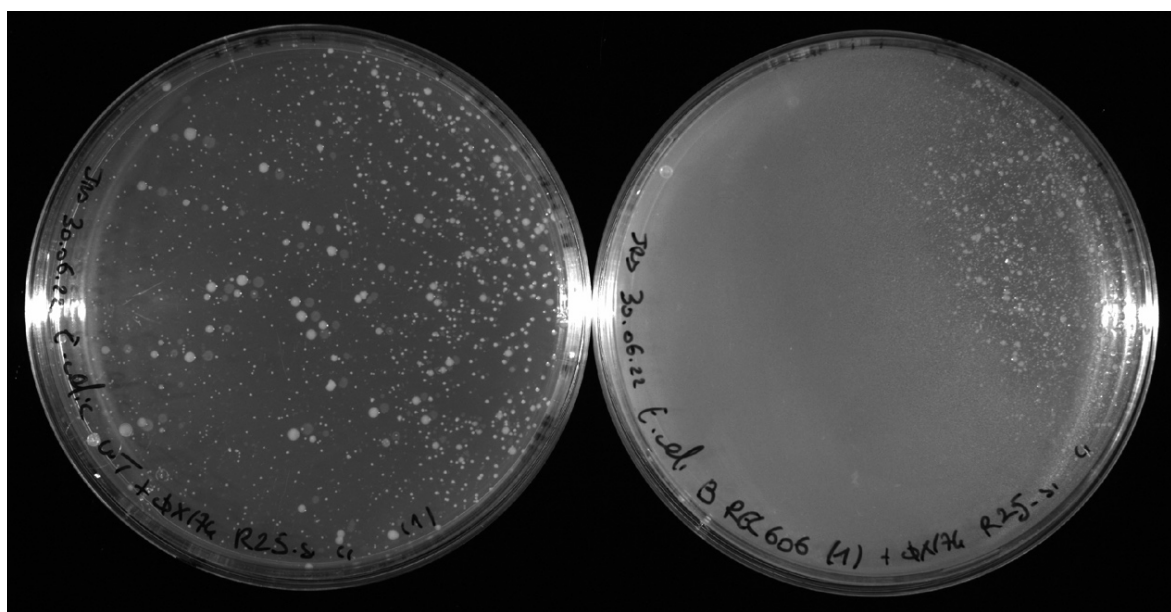
Fig. 5.3. *E. coli* B REL606 is susceptible to the ΦX174 R25 T4. Top agar overlays were prepared by mixing 100μl of ΦX174 R25 T4 phage lysate (~10^8pfu/ml) with 200μl of *E. coli* C wildtype (left) or 200μl of *E. coli* B REL606 (right) in 4ml SSA with CaCl$_2$ and MgCl$_2$ (at a final concentration of 5mM and 10mM, respectively). ΦX174 R25 T4 completely erased the bacterial lawn made by *E. coli* C wildtype after ~17 hours (at 37°C). Only spontaneous resistant colonies

survived the phage infection. ΦX174 R25 T4 can infect *E. coli* B REL606 but was unable to kill all bacteria.

I believe that my two phage cocktail evolution experiments (see **Chapter II** sections **2.3.8** and **2.3.9**) could be exploited in the context of human health. By increasing ΦX174 sequence diversity previous to the evolution experiment, evolving ΦX174 strains that can infect any new hard-resistant *E. coli* strains (bacterial strains that cannot be infected by phages in a few mutational steps, see **Figs. 4.11B** and **4.11C**) would become easy and fast.

5.5.   Preliminary investigations and results

5.5.1.   Host ranges of the evolved ΦX174 strains increase in a spatially structured environment

These preliminary experiments aimed to determine whether the evolved phages' infection patterns (host ranges) observed in a semi-solid environment (**Figs. 4.10** and **5.1**) are consistent across different environments. It is important to note that the non-isogenic phage strains c1 clones (see **Table 4.4**) were used for these experiments. Controls on the phage strains' isogeneity were not performed at this moment. I tested the infection phenotype of the evolved phages in two environments: one with a spatial structure and one without. For the environment with a spatial structure, I spotted drops of high titer phage lysates on the top of bacterial lawns on solid LB plates (**Chapter II** section **2.4.1.2**). For the environment without a spatial structure, I infected well-mixed liquid cultures of bacteria in 96 well plates (**Chapter II** section **2.4.1.3**).

Compared to the evolved phages' host ranges observed in the semi-solid environment (**Figs. 4.10** and **5.1**), host ranges are drastically increased when spotting assays were performed on solid LB plates (**Fig. 5.S2**) and reduced when performed in the liquid, well-mixed environment (**Fig. 5.S3**). The increase in phage infectivity is especially striking with ΦX174 wildtype (ancestor). While its host range remains unchanged in the liquid cultures, ΦX174 wildtype can re-infect 14 out of 31 *E. coli* C resistant strains when the bacteria are grown on solid LB plates (**Fig. 5.S3**).

Except for the *gmhB*, *galU* and *waaG* mutants (**Table 3.1**), all resistant bacteria are susceptible to infection by at least three evolved phages in the spatially structured environment (**Fig. 5.S2**). Moreover, some bacterial strains that are usually infected by a narrow set of evolved phages in a semi-solid environment suddenly became susceptible to almost all phages. For example, *E. coli* C

R7 can be infected by all evolved phages (**Figs. 4.10** and **5.S2**), and *E. coli* C R32 can be infected by 11 evolved phages (out of 13, **Figs. 4.10** and **5.S2**). Even the hard-resistant strains R6 (*waaP/pssA* mutant) and R25 (*rfaH* mutant) which required specific experimental conditions to be overcome (**Chapter II** section **2.3.9**) can now be infected by ten and three evolved phages, respectively. The bacterial outer membrane of the resistant strains may be more exposed and susceptible to phage adsorption and infection when the bacteria grow on solid LB plates due to the LPS truncations impairing cell-cell and cell-surface interactions[264]. Alternatively, ΦX174 may be better adapted and more efficient at infecting bacteria in spatially structured environments (**Figs. 4.10** and **5.S2**) rather than in a liquid environment (**Fig. 5.S3**).

While these results must be taken with the highest precaution and confirmed via further experiments (lack of replicates, phage strains missing or bacterial and phage strains found later to be non-isogenic), they suggest that the infectivity pattern of a phage is bound to a specific environment and cannot be transposed elsewhere.

5.5.2. The phage cocktail does not prevent the emergence of LPS-based phage resistance

This preliminary experiment aimed to determine if a cocktail made of evolved ΦX174 can prevent the emergence of LPS-based phage-resistant bacteria. Since each of the 31 *E. coli* C resistant strains can be infected by at least one evolved phage (**Figs 4.10** and **5.1**), I assumed that pooling all evolved phage strains retrieved from my different evolution experiments should be able to prevent the emergence of spontaneous phage-resistant *E. coli* C strains with the same predicted LPS structures (**Fig. 3.7**).

Using the same approach described in **Chapter II** section **2.2.1**, I generated and isolated eleven spontaneous *E. coli* C strains resistant to a cocktail of phages constituted of (i) ΦX174 wildtype and (ii) 21 evolved ΦX174 strains (see **Chapter II** section **2.4.2**). Whole-genome sequencing was performed on each resistant mutant to detect the mutations responsible for the resistant phenotypes (**Table 5.1**).

E. coli C resistant	Genes	Locus tags	Synonyms	Descriptions	Positions	Nucleotide changes	Amino Acid changes	Predicted LPS structure phenotypes
22ΦR1	*gmhB*	*B6N50_18960* ←	-	D-glycero-beta-D-manno-heptose 1,7-bisphosphate 7-phosphatase	3,693,608:1	+a	Y91fsTer#99[a]	Deep rough
22ΦR2	*waaP*	*B6N50_00410* →	*rfaP*	Lipopolysaccharide core heptose(I) kinase	75,978	c→t	R190W	Deep rough
	pgm	*B6N50_16200* ←	-	Phosphoglucomutase	3,172,605	Δ1 bp	K347fsTer#354	
22ΦR3	*rfaH*	*B6N50_22865* →	-	Transcription/translation regulatory transformer protein	4,485,917	c→t	Q18stop	Unknown
22ΦR4	*rfaH*	*B6N50_22865* →	-	Transcription/translation regulatory transformer protein	4,485,917	c→t	Q18stop	Unknown
22ΦR5	[*galU*] - [*rssB*] - [*rssA*] - [*ychJ*] - [*purU*]	[*B6N50_13325* ←] - [*B6N50_13330* ←] - [*B6N50_13335* ←] - [*B6N50_13340* →] - [*B6N50_13345* →]	-	[UTP--glucose-1-phosphate uridylyltransferase] - [Regulator of rpoS] - [Putative patatin-like phospholipase] - [NTF2-like domain-containing protein] - [Formyltetrahydrofolate deformylase]	2,611,466 - 2,615,824	Δ4,359 bp	Δ[*galU*_A213_Intergenic region][b]	Deep rough
22ΦR6	-	*B6N50_00035* ← / → *B6N50_00040*		-	8,469	t→c	Intergenic region	Deep rough
	waaV	*B6N50_00435* →	-	UDP-glucose:(Glucosyl) LPS β1,3-Glucosyltransferase	80,534	$(t)_{7 \to 6}$	F149fsTer#158	
	gmhD	*B6N50_00455* ←	*hldD*	ADP-L-glycero-D-mannoheptose-6-epimerase	85,299	c→t	G36D	

22ΦR7	*waaG*	B6N50_00405 →	*rfaG*	UDP-glucose:(heptosyl) LPS α1,3-glucosyltransferase (glucosyltransferase I)	74,345	c→a	R18S	Deep rough
22ΦR8	*waaG*	B6N50_00405 →	*rfaG*	UDP-glucose:(heptosyl) LPS α1,3-glucosyltransferase (glucosyltransferase I)	74,946 - 74,952	Δ7 bp	L218fsTer#265	Deep rough
22ΦR9	*waaV*	B6N50_00435 →	-	UDP-glucose:(Glucosyl) LPS β1,3-Glucosyltransferase	81,038:1	+t	V317fsTer#322	Deep rough
	galU	B6N50_13325 ←	-	UTP--glucose-1-phosphate uridylyltransferase	2,611,225	c→t	W293stop	
22ΦR10	*waaT*	B6N50_00420 →	-	UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase	77,462	c→g	S69stop	Rough
	macB	B6N50_15175 ←	-	Macrolide ABC transporter permease/ATP-binding protein	2,966,384	c→a	W247L	
22ΦR11	*waaT*	B6N50_00420 →	-	UDP-galactose:(glucosyl) LPS α-1,2-galactosyltransferase	77,563	g→a	A103T	Rough

Table 5.1. List of mutations found in eleven *E. coli* C strains that are resistant to a cocktail of 22 evolved ΦX174. *E. coli* C resistant strains: 22ΦR# is the strain number. Genes: name of the gene(s) in which mutations have been identified (as compared with wildtype *E. coli* C). Locus tags: identifier of each listed gene. Synonyms: alternative gene names. Descriptions: protein product encoded by each listed gene. Positions: genomic coordinates of mutation. Nucleotide changes: observed nucleotide change. Amino acid changes: resulting change in amino acid sequence. Predicted LPS structure phenotypes: based on predicted LPS structure (see **Fig. 3.7**). "→": indicates 5'➜3' direction. "INS": insertion. "stop": stop codon. "ᵃ": Y91fsTer#99 indicates a frameshift (fs) leading to a premature codon stop (Ter); #: the position of the premature stop codon. "ᵇ": Δ[*galU*_A213-Intergenic region] indicated a long deletion starting from *galU* [B6N50_13325 ←] to the intergenic region [B6N50_13145 →/ B6N50_13150].

Mutations in phage-resistant mutants occurred in previously affected LPS genes (*gmhB*, *waaP*, *waaG*, *galU*, *waaT*, and *rfaH*; **Table 5.1**). Seven resistant strains are predicted to carry a single mutation, three carry two mutations, and one carries three mutations. A total of 16 mutations are identified, 15 of which are unique (including nine nucleotide substitutions, four deletions, and two single nucleotide insertions). Nine (out of 15, or 60%) mutations introduce premature stop codons or lead to frameshifts and thus are highly likely to disrupt gene function (see **Table 5.1**). The bacterial mutants are predicted to collectively produce five different LPS structures, including both "rough" and "deep rough" phenotypes.

One phage-resistant strain mutation, however, carries a mutation in a previously unaffected LPS gene. 22ΦR6 carry one mutation in *gmhD* and the deletion (t)$_{7 \rightarrow 6}$ in *waaV* at position 80,534 (**Table 5.1**). This deletion is identical to the one identified previously in the excluded *E. coli* C R1 strain (**Table 3.2**). All mutations occurring in *waaV* are systematically associated with at least another mutation in a gene linked to the LPS biosynthesis (*waaO*) or assembly (*gmhD*, *galE*, and *galU*), suggesting that a single mutation in *waaV* may not be sufficient to resist the phage infection. Furthermore, the *gmhD* mutant was found in the phage cocktail treatment. Thus, the absence of *gmhD* mutants in the first 31 resistant strains was likely due to chance during the isolation step (see **Chapter II** section **2.2.1**).

Even if these mutations are located in the same LPS genes, they may cause different phage-resistant phenotypes that are resistant to all phages present in the cocktail (**Fig. 5.1**). In the following preliminary experiment, I assessed the resistance of the novel mutants to a set of evolved phages. The evolved phages were chosen based on which corresponding resistance they overcame during the different evolution experiments and/or their infection phenotypes (**Figs. 4.10** and **5.1**). It is also important to note that not all phage-bacteria combinations were tested at this time. On the tested pairs, I found that nine phage-cocktail resistant strains (out of 11, ~73%) remain sensitive to the infection of at least a single evolved phage strain (**Fig. 5.4**).
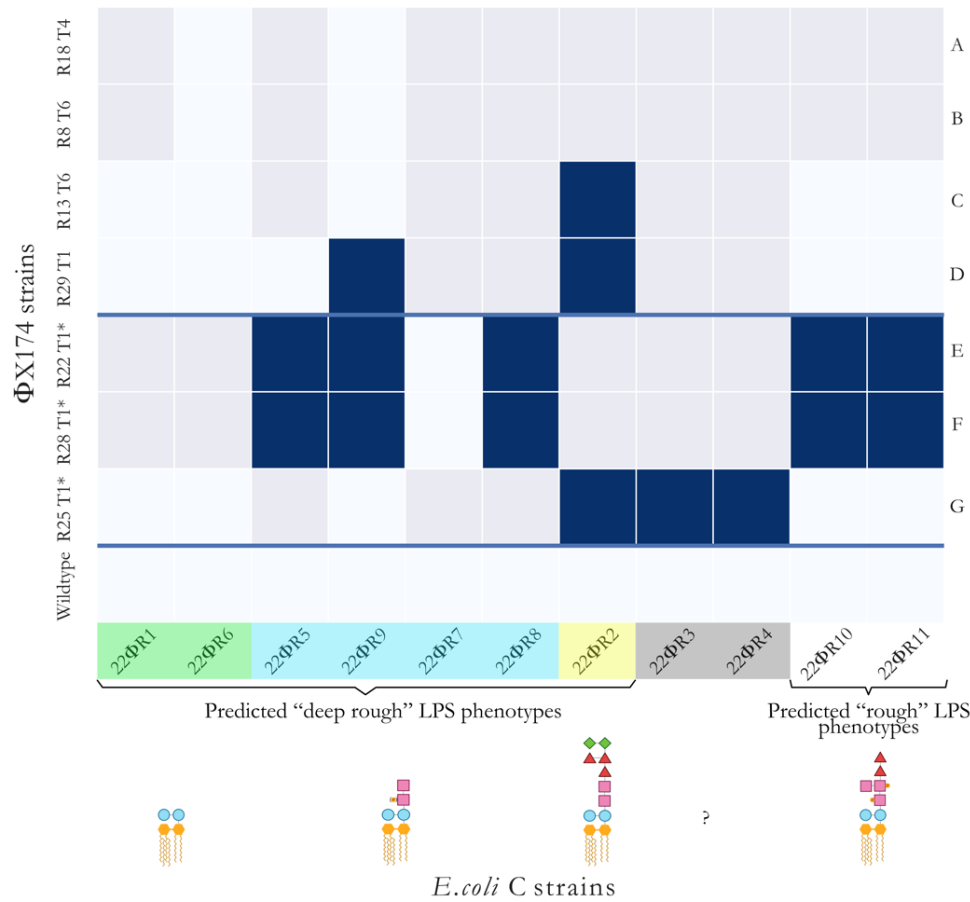
Fig. 5.4. Infection matrix of evolved ΦX174 phages on the eleven phage-cocktail resistant *E. coli* C strains. *E. coli* C strains (22ΦR1 to 22ΦR11) are grouped and coloured based on their predicted core LPS structures[106,107,110,113,115,118–121,123–125,128,129,131,144,235]. ΦX174 strains are ordered based on the predicted LPS structures they infected during the evolution experiments (A-B: overcame heptoseless *waaC* and *hldE* mutants; C-D: overcame *waaO* mutants; E-F: overcame *galU* and *waaG* mutants; G: overcame *rfaH* mutant). ΦX174 R22 T1*, ΦX174 R25 T1*, and ΦX174 R28 T1* were obtained from the preliminary phage cocktail evolution experiment (see **Chapter II** section **2.3.7**). Details on their genotypes and infection phenotypes can be found in **Table 5.S1** and **Fig. 5.S4**. The solid blue lines separate the evolved phages according to the evolution experiment in which they were isolated (A-D: Standard; E-G: Preliminary). Dark blue squares = infection; light blue squares = no infection; grey squares = the phage-bacterium combination was not tested. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed. "?": core LPS structure of *E. coli* C 22ΦR3 and 22ΦR4 (*rfaH* mutants) could not be predicted.

A reason why sensitive bacterial mutants survived the phage-cocktail infection may be that a spatially structured environment (such as a solid LB plate) often offers spatial refuges free of phage for slow-growing phage-sensitive mutants to persist through the phage infection[199,265–267]. These refuges can emerge from an unequal distribution of phages on the plate, a slow diffusion of phages through the bacterial population or a physical barrier. For example, all phage-cocktail-resistant strains harbour structural modifications on their LPS that can impair cell-cell and cell-surface interactions. These impairments can promote the formation of aggregates and biofilms that, in turn, protect the trapped phage-sensitive strains from phages[264,268,269].

Hence, I focused my investigations on preventing the emergence of LPS-based resistance in a liquid, well-mixed environment where spatial refuges are absent (see **Chapter II** section **2.4.3**). I also reduced the complexity of the experiment by using a phage cocktail composed of the minimum number of evolved phage strains necessary to infect each LPS phenotype (**Figs. 4.10** and **5.S4**).

5.5.3. The phage cocktail delays the emergence of phage-resistant bacteria in a liquid environment

In this preliminary experiment, I tested whether a phage cocktail consisting of ΦX174 R22 T1*, ΦX174 R25 T1*, and ΦX174 R29 T1 was more efficient at (i) killing *E. coli* C wildtype and (ii) better preventing the emergence of phage-resistant bacteria than using ΦX174 wildtype (ancestor) alone in a liquid environment (see **Chapter II** section **2.4.3**).

Both phage treatments decrease the OD600 values quickly after inoculation (**Fig. 5.5**). However, the decrease is stronger and faster when the cocktail treatment is applied. With the phage cocktail, this decrease only stopped when resistant strains regrew after ~13h of incubation. With the ΦX174 wildtype treatment, the OD600 value reaches a plateau (~0.48) after ~7h30 of incubation, and the resistant strains regrew earlier after ~10h of incubation (**Fig. 5.5**). After 24h of incubation, the OD600 value is lower with the phage cocktail than with the wildtype ΦX174 (ancestor). Thus, the cocktail of three phages is more efficient at killing *E. coli* C wildtype and delays the emergence of phage-resistant bacteria better than ΦX174 wildtype alone.
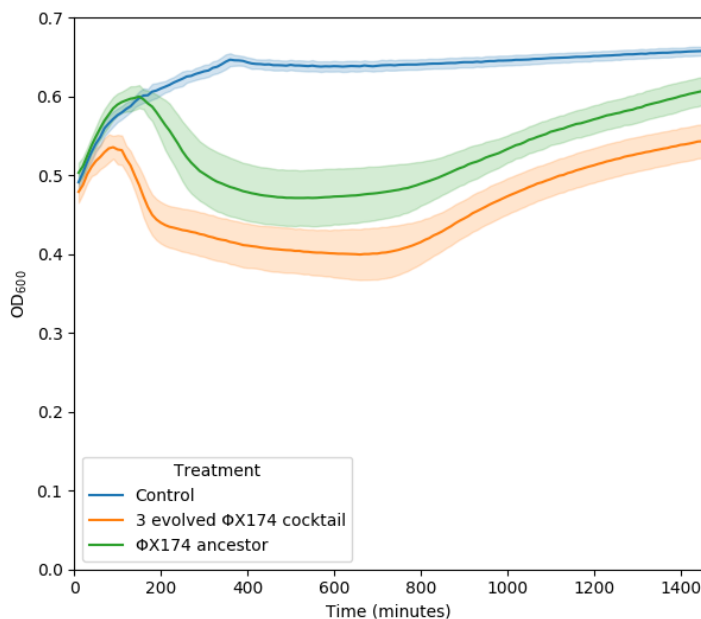
Fig. 5.5. Infectious dynamics of a phage cocktail against *E. coli* C wildtype. Bacteria were infected with ΦX174 wildtype (ancestor) alone or with the cocktail of three phages in 96 well plates (see **Chapter II** section **2.4.3**). *E. coli* C wildtype was also grown without phage as a control. Thick lines represent the average value of the OD600 from five replicates. Transparent bands correspond to the upper and lower values of the OD600. Blue line: *E. coli* C wildtype control without phage; green line: *E. coli* C wildtype + ΦX174 wildtype (ancestor); orange line: *E. coli* C wildtype + ΦX174 R22 T1* + ΦX174 R25 T1* + ΦX174 R29 T1. Evolved phages ΦX174 R22 T1* and ΦX174 R25 T1* were obtained from the preliminary phage cocktail evolution experiment (see **Chapter II** section **2.3.7**). Details on their genotypes and infection phenotypes can be found in **Table 5.S1** and **Fig. 5.S4**.

Although the phage cocktail is more efficient than the single ΦX174 wildtype treatment, resistant bacteria re-occurred in both cases within a day (**Fig. 5.5**). One explanation could be that the phage resistance selected in the liquid media may be completely different from the one selected in a solid media (novel LPS phenotypes or type of resistance), making the cocktail unable to re-infect the bacterial mutants. Unfortunately, I do not have the information regarding the type of resistance that arose at the end of the experiment. Future works should therefore aim to isolate and characterise phage-resistance bacteria from a liquid environment. If phage resistances are indeed different from the ones observed in a solid environment (**Table 3.1**), new ΦX174 strains should be evolved against them to generate a more efficient phage cocktail.

Alternatively, the adsorption rate of the evolved phages may be too low. A low adsorption rate ($\sim 10^{-11}$ or less) allows the coexistence between phage and bacteria (even sensitive ones) at high densities in a liquid environment[270]. Therefore, the MOI_{actual} - the number of phages that have *actually* adsorbed/infected bacteria and have contributed to a multiplicity of infection[271] (to be distinguished for the MOI_{input}, the ratio of phages *added* to bacteria[271]) - may not be sufficient to effectively kill the LPS mutants before they reach their stationary phase. Further investigations could also focus on measuring and, if necessary, improving the evolved ΦX174 strains' adsorption rates by resuming their evolution only against their corresponding *E. coli* C resistant strains.

5.6.   Discussion

5.6.1.   The success of a phage cocktail at preventing the emergence of phage resistance may be influenced by the environment in which the phages have primarily evolved

In my preliminary investigations, I observed that phage infection[199,247] (and evolution) was very different from one environment to another and suggested that the infection phenotype associated with each phage may only be relevant for the environment in which it has been determined first. Thus, the fact that the evolved phages can collectively infect all resistant strains in semi-solid does not necessarily imply that they can do the same in liquid or solid. If this is true, then phage cocktails generated based on host ranges determined only in a semi-solid environment may not infect bacteria in a liquid or solid environment (**Figs. 5.S2** and **5.S3**).

To test this hypothesis, lineages of ΦX174 wildtype will be adapted against each *E. coli* C resistant strain in a solid environment. For this purpose, the phage evolution experiment – Standard will be modified for the solid and semi-solid environments. ΦX174 wildtype will be mixed with both *E. coli* C wildtype and one of the resistant strains in a 96-well plate then plated with sterile beads in a solid LB plate or inoculated in a top-agar overlay. After 6 hours of incubation, the agar of each plate will be scraped and dissolved. Drops of chloroform will be added to retrieve and purify the phage lysates. To start the next transfer, the purified phage lysates from the previous day will be mixed with fresh bacterial hosts before being plated and incubated again. Serial transfers will be performed daily for a total of 21 transfers. Isogenic evolved phage strains that successfully reinfected their corresponding resistant strains will be isolated.

If some resistant strains remain uninfected, the phage cocktail evolution experiment – Increased diversity will be modified for the solid environment. Finally, a phage cocktail will be generated for phages that evolved in liquid or solid environments. The different phage cocktails will be made with the minimum set of evolved phage strains that cover all *E. coli* C resistant strains in the relevant environment. Phage cocktails' efficiencies at killing *E. coli* C wildtype and preventing the emergence of LPS-based phage resistance will also be compared in the other environments to determine whether or not the results are consistent across the environments.

I would expect each phage cocktail to be more efficient at killing *E. coli* C wildtype and preventing the emergence of LPS-based phage resistance in the media in which the phages have been primarily adapted. If this is true, then experimentalists would have to consider for the design of their phage cocktails: (i) the type of environment (structured or unstructured) in which the focal bacterial strain settle and grow, (ii) in what environment a wildtype phage is most fit to infer the environmental conditions of the evolution experiment, and (iii) evolve the wildtype phage to infect the strain in an environment which is the closest to the real infection conditions[272].

However, not being able to prevent all LPS resistance at once might not be a bad thing. While most phage cocktails are designed to (i) reduce the growth of the focal strain and (ii) prevent or delay as much as possible the emergence of phage resistance[273–275], they could instead be designed as an evolutionary trap. In this scenario, a phage cocktail primarily aims to drive the evolution of the bacteria toward predictable and often cost-heavy outcomes for the bacterium (such as the deleterious deep rough phenotypes which are associated with costly pleiotropic effect, especially in the absence of threat[96,97,121,149,151,276]) that could lead to a lower capacity to evolve new resistance types[277]. Then, phages specifically evolved to infect the new resistant strains can be used alone or in association with antibiotics that synergize with the phage treatment to kill the remaining mutants (sequential treatment)[158,276,278].

5.6.2. Conclusion

While the predictions of the current LPS model were sometimes correct, I also observed that specific mutations in the **same** gene could cause **different** phage-resistance phenotypes, indicating that LPS phenotypes are much more diverse than anticipated from the current one-gene-one-phenotype model of LPS biology. This greater LPS diversity often leads to a larger range of phage resistance phenotypes.

My work demonstrates that ΦX174 can be effectively bred to infect a wide variety of resistant mutants and open the possibility to evolve ΦX174 to infect any *E. coli* strain, including pathogenic strains. Given that ΦX174 is already approved for use in humans[98,99], it may prove a viable alternative for phage therapy to uncharacterized, potentially unsafe, environmental phage isolates.

5.7. Supplementary Figures and Tables
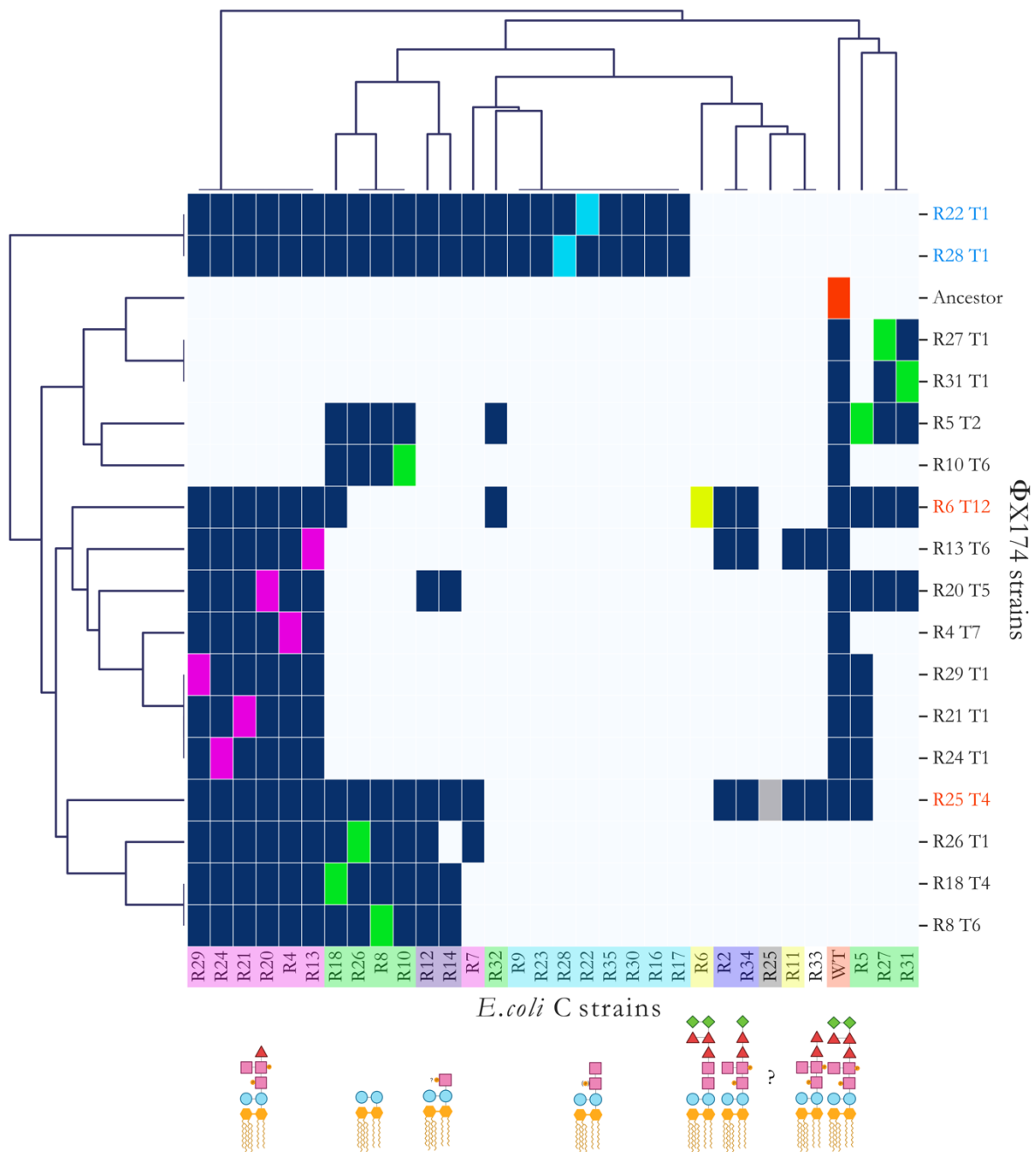
Fig. 5.S1. Hierarchical agglomerative clustering analysis of the host ranges of evolved phage (including the evolved phage infecting R5). To unravel the structures in the heatmap infection matrix (**Fig. 4.10**), I performed a clustering analysis using the default settings of the Seaborn library (version 0.11.2) for Python (version 3.7.4). Briefly, each data point is considered as a cluster. The linkage method computes the distance between them, then repeatedly combines the two nearest clusters into larger clusters, until a single cluster is left. *E. coli* C strains are coloured based on their predicted core LPS structure (**Fig. 3.7**). Phage names in black were obtained during the first 21 serial transfers (phage evolution experiment – Standard). Phage strains in blue were obtained in the phage cocktail evolution experiment – Increased diversity (one transfer per day). Phage strains in red were obtained in the phage cocktail evolution experiment – Increased diversity and generations (four transfers per day). Dark blue squares = infection; light blue squares = no infection; coloured squares = control infection by a phage evolved on that host. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed. "?": core LPS structure of *E. coli* C R25 (*rfaH* mutant) could not be predicted.

Fig. 5.S2. Infection matrix of evolved ΦX174 phages on the 31 resistant *E. coli* C strains in a solid environment – Preliminary results. The infection matrix was produced by combining the results from the solid environment assays (see **Chapter II** section **2.4.1.2**). *E. coli* C strains are ordered based on their predicted core LPS structures[106,107,110,113,115,118–121,123–125,128,129,131,144,235]. Controls on the isogeneity of the phage strains (c1 clones) were not performed at this time. No evolved phage infecting the *waaG* and *galU* mutants was found yet. Dark blue squares = infection; light blue squares = no infection. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed.

Fig. 5.S3. Infection matrix of evolved ΦX174 phages on the 31 resistant *E. coli* C strains in a liquid environment – Preliminary results. The infection matrix was produced by combining the results from the liquid environment assays (see **Chapter II** section **2.4.1.3**). *E. coli* C strains are ordered based on their predicted core LPS structures[106,107,110,113,115,118–121,123–125,128,129,131,144,235]. Controls on the isogeneity of the phage strains (c1 clones) were not performed at this time. No evolved phage infecting the *waaG* and *galU* mutants was found yet. Dark blue squares = infection; light blue squares = no infection. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed.
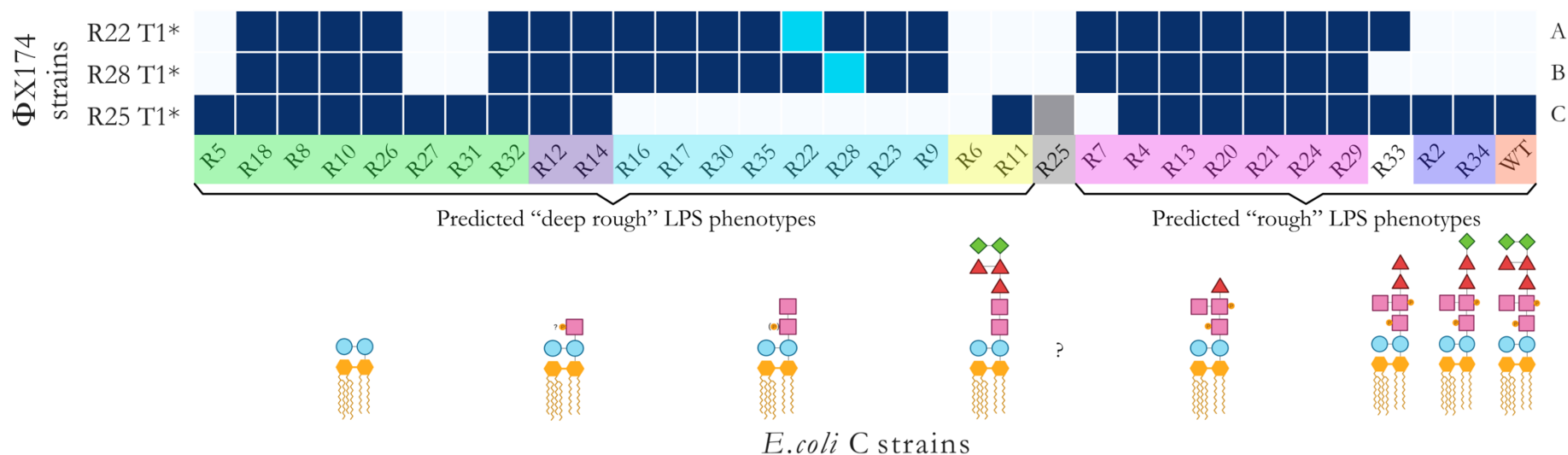
Fig. 5.S4. Infection matrix of evolved ΦX174 phages retrieved from the preliminary phage cocktail evolution experiment. *E. coli* C strains are grouped and coloured based on their predicted core LPS structures[106,107,110,113,115,118–121,123–125,128,129,131,144,235]. ΦX174 strains are ordered based on the predicted LPS structures they infected during the evolution experiments (A-B: overcame *galU* and *waaG* mutants; C: overcame *rfaH* mutant). Dark blue squares = infection; light blue squares = no infection; coloured squares = control infection by a phage evolved on that host. R# indicates the number of the resistant strain the phage evolved on, and T# is the transfer number where plaques were first observed. "*": indicates that these evolved phage strains were retrieved from the preliminary phage cocktail evolution experiment (see **Chapter II** section **2.3.7**). "?": core LPS structure of *E. coli* C R25 (*rfaH* mutant) could not be predicted.

ΦX174 strains	Gene mutated in the corresponding *E. coli* C resistant	Mutants sensitive to the evolved ΦX174 strain infection	Gene mutated in ΦX174 strains	Descriptions	Positions	Nucleotide changes	Amino acid changes	Are mutations present in a control lineage?
R22 T1*	*galU*	R18, R8, R10, R26, R32, R12, R14, R16, R17, R30, R35, R22, R28, R23, R9, R7, R4, R13, R20, R21, R24, R29	F	Capsid protein	1,307	t→c	Y103H	No
			F	Capsid protein	1,317	a→g	H106R	No
			F	Capsid protein	1,460	c→a	Q154K	No
			H	Minor spike protein	3,132	g→a	A68T	No
R25 T1*	*rfaH*	R5, R18, R8, R10, R26, R27, R31, R32, R12, R14, R25, R11, R4, R13, R20, R21, R24, R29, R33, R2, R34, WT	F	Capsid protein	1,304	g→c	G102R	No
			F	Capsid protein	1,614	c→a	T205N	No
R28 T1*	*waaG*	R18, R8, R10, R26, R32, R12, R14, R16, R17, R30, R35, R22, R28, R23, R9, R7, R4, R13, R20, R21, R24, R29	F	Capsid protein	1,307	t→c	Y103H	No
			F	Capsid protein	1,317	a→g	H106R	No
			F	Capsid protein	1,461	a→g	Q154R	No
			H	Minor spike protein	3,132	g→a	A68T	No

Table 5.S1. List of mutations identified in ΦX174 evolved isolates retrieved from the preliminary phage cocktail evolution experiment. ΦX174 strains: R# is the corresponding resistant strain number, and T# is the transfer number where plaques were observed for the first time on the given resistant strain. "*": indicates that these evolved phage strains were retrieved from the preliminary phage cocktail evolution experiment (see

Chapter II section **2.3.7**). Mutants sensitive to the evolved ΦX174 strains infection: list of all resistant strains that a specific phage strain can infect (See **Fig. 5.S4**). Descriptions: names of proteins encoded by the listed genes. Positions: genomic coordinates of mutation (according to GenBank accession number AF176034.1). Nucleotide changes: observed nucleotide change. Amino acid changes: resulting change in amino acid sequence. "→": indicates 5'→3' direction. Are mutations present in a control lineage?: indicates mutations observed in at least one control lineage.

Chapter VI – Discussion

6.1.   Overview of the thesis

6.1.1.   A wide variety of LPS structures leads to ΦX174 resistance

The evolution of *E. coli* C strains resistant to ΦX174 wildtype infection and their subsequent mutational analyses marked the first step of my study (see **Chapter III**). I showed that *E. coli* C becomes resistant to ΦX174 wildtype infection through mutations in genes involved in the LPS biosynthesis, assembly, or regulation. These mutations often introduce premature stop codons or lead to frameshifts that shut down the gene function. For the vast majority of the mutant strains (27 out of 31, ~87%), only one single mutation was enough to confer resistance to the phage infection (**Table 3.1**).

Surprisingly, a few LPS mutants carry additional mutations in genes without an apparent link with LPS biosynthesis, assembly, or regulation (**Table 3.1**). A closer look revealed that these mutations impact genes involved in metabolic pathways that transform toxic products into non-toxic ones or in the stabilisation of outer membrane proteins[173,174,179,191]. Since only a single mutation in one of the LPS genes is presumably enough to avoid phage infection, the presence of mutants with additional mutations was unexpected. Excesses of mutational jackpot events have already been reported elsewhere[198,200,201] and are considered to be key factors in the emergence of complex phenotypes, such as drug resistance or cancer[198]. Implications in the acquisition and spread of phage resistance may also be possible yet remain unknown and unexplored.

From the current model of LPS biosynthesis and assembly[119,120,113,121,123,110], I predicted a final core LPS structure for each resistant strain based on which LPS genes are mutated (**Figs. 3.1**, **3.2**, and **3.7**). I determined that the bacterial strains collectively produced seven predicted LPS phenotypes: four were categorized as deep-rough and three as rough (**Fig 3.7**). The LPS structure of *E. coli* C R25 remains unpredictable since any LPS biosynthetic gene could potentially be affected by a mutation in the *rfaH* gene[109,133].

Changes in the LPS structure, especially the ones leading to deep rough phenotypes, are often associated with growth impairments[156–159]. I distinguished three growth categories from OD600 measurements in a well-mixed liquid media and the absence of phage: fast, intermediate, and slow (**Fig. 3.10**). However, I could not accurately discriminate the seven predicted LPS structures with

growth analyses alone since not all LPS mutants show growth deficiency compared to *E. coli* C wildtype (fast category). Hence, I evolved ΦX174 wildtype to reinfect each LPS mutant and used their high-LPS specificity to test my LPS genotype-phenotype maps.

6.1.2. Infection of resistant bacteria by ΦX174 requires evolutionary intermediates

Evolving ΦX174 wildtype to reinfect each resistant strain marked the second step of my thesis (see **Chapter IV**). I developed and performed an evolution experiment where I grew ΦX174 in a culture consisting of *E. coli* C wildtype and one of the resistant strains of interest for a total of 21 daily transfers (see **Chapter II** section **2.3.5** and **Fig. 4.4**). The twelve isogenic phage strains were isolated from phage lineages that reinfected their corresponding resistant strains. They all carry mutations in only two genes. These genes are involved in host recognition (F capsid protein) and DNA injection (H minor spike protein) (**Table 4.2**, **Fig 4.12A** and **4.12C**). Parallel evolution was often observed among phage lineages that adapted to bacterial strains with the same predicted core LPS structures (**Fig 4. 12B**), suggesting that my predictions based on the current LPS model are accurate.

To reinfect the remaining resistance strains that could not be cross-infected by any of the evolved phage strains (considered to display a hard-resistant phenotype), I designed additional evolution experiments (see **Chapter II** sections **2.3.6 to 2.3.9**). I increased the phage sequence diversity before starting the serial transfers by pooling ΦX174 wildtype with all evolved phage strains that successfully reinfected their corresponding resistant strains in one phage cocktail. Then, I serially transferred it daily, up to four times a day, in a mixed culture of fresh permissive and non-permissive hosts (see **Chapter II** sections **2.3.8** and **2.3.9**, **Fig. 4.9**). Increasing both phage and host diversity and increasing the number of phage generations sped up the adaptation process of ΦX174 significantly. Four phage strains evolved that were able to infect *waaP/pssA*, *waaG*, *galU* and *rfaH* mutants (the hard-resistant strains). They all carry four mutations in only the *F* and *H* genes (**Table 4.2**, **Fig. 4. 12B**).

Interestingly, almost all mutations of these four evolved phage strains are also found in the phage genomes present in the initial cocktail (**Table 4.2**, **Fig. 4.12B**). Both additional mutations and recombination events created hybrid phage strains able to reinfect hard-resistant bacteria. ΦX174 adaptation depends upon the presence of specific mutations[68,279] and, therefore, may not be possible to achieve in an evolution experiment starting from only the wildtype genotype.

6.1.3.   The current model of LPS biology can be challenged using evolved phages

Testing the predicted LPS phenotypes marked the last step of my thesis (see **Chapter V**). I predicted LPS structures by assuming that all LPS mutations cause loss of function phenotypes. If my LPS structure predictions are correct, then all bacteria predicted to carry mutations in the same LPS gene should be susceptible to the same set of phages.

Phage infectivity patterns (host range) showed that phages that evolved to infect the same predicted LPS structure are not always able to cross-infect each other's host. Of the six predicted LPS structures that occur in at least two *E. coli* C mutants, three fitted my LPS predictions. They constituted groups in which bacterial strains are infected by the same set of evolved phages (highlighted in cyan, light purple, and dark purple; **Fig. 5.1**). However, three predicted LPS structures did not fit my predictions. The bacterial strains predicted to belong to the same group were infected by a different set of evolved phages (highlighted in pink, green, and yellow; **Fig. 5.1**). This result demonstrates that the LPS structural diversity is greater than anticipated and leads to a wider range of resistance.

Phage genotype-phenotype maps can also be constructed from the phage infectivity patterns. I demonstrated that a combination of the two phage substitutions Q154K/R in the F capsid protein and A68T in the H minor spike protein leads to an infectivity trade-off and the loss of the ability to infect *E. coli* C wildtype in favour of *waaG* and *galU* mutants' infection.

6.2.   Implications of the results for the field of phage evolution

6.2.1.   Stepwise evolution of phage and its host can be used to explore the evolution of many resistance and counter-resistance strategies

Antagonistic co-evolution between phages and bacteria – the evolutionary processes by which phages impose selection for resistant hosts, while the newly resistant hosts impose selection for phages to evolve to reinfect them – have been extensively studied via "top-down" approaches[136,256,277,280–285]. A "top-down" approach typically aims to gain an overview of the system by first analysing it as a whole, then breaking it down and further analysing the individual parts. In the laboratory, the experimentalist sets up the experimental conditions (e.g., media, temperature, pH, growth regime) and allows both phages and bacteria to evolve over many generations[141,142,286].

These pioneering studies revealed that parallel evolution – the repeated evolution of the same phenotype or genotype in different populations – is commonly observed during micro-organisms' adaptation to similar environmental conditions[65,66,82,287–289] (e.g., different bacterial populations repeatedly evolve the same resistance strategies against antibiotics or phages[154,290]). The main advantage of observing repeatable evolutionary outcomes is that these regularities can be used to make predictions on how microorganisms will evolve in the future[136,256,277,280–285,291–293].

Instead of a "top-down" approach, I chose a "bottom-up" approach to explore and predict the evolutionary responses of *E. coli* C, to the infecting coliphage ΦX174. The objectives of a "bottom-up" approach are to specify and characterise in detail the individual parts of the system and link them together to understand emerging properties of the system[286]. From this perspective, I could predict and test LPS genotype-phenotype maps of each *E. coli* C mutant. This approach was possible because of the wealth of information on LPS biosynthesis, assembly, and regulation[119,120,113,121,123,110].

The stepwise evolution of *E. coli* C and ΦX174 allows me to avoid the common pitfalls of short-term co-evolution experiments[281,294–296]. Usually, the arms race – the escalation between defence and counter-defence mechanisms generated by both species – is asymmetric and often ends with the victory of the host over the phage after only a few rounds of resistance and overcoming resistance strategies[136,294,295,297]. In the case of a simple co-evolution between *E. coli* C and ΦX174, the emergence of hard-resistant *E. coli* C strains (such as *waaP/pssA*, *waaG*, *galU*, and *rfaH* mutants) that cannot be infected in a few mutational steps (**Fig. 4.11B**) will put an end to the arms race. ΦX174 will be diluted and disappear from the host cultures after a few serial transfers. However, the fact that ΦX174 can be easily bred to reinfect even the hard-resistant strains opens the possibility of pursuing the arms race between *E. coli* C and ΦX174 for more than a few rounds. In the following paragraphs, I describe a possible evolution experiment to perform to expand the stepwise evolution of *E. coli* C and ΦX174.

First, the experimentalist will generate the second generation of spontaneous phage-resistant *E. coli* C strains by mixing each of the 16 evolved phage lysates with their corresponding *E. coli* C mutants' stationary-phase cultures on solid agar plates (fluctuation experiments, see **Chapter II** section **2.2.1**). One single colony from each plate will be isolated, and whole-genome sequencing will be performed to unravel the mutational cause behind the new phage resistance. This first step

will allow the experimentalist to see which strains of *E. coli* C can further evolve resistance and which ones cannot. Especially, the LPS mutants that got rid of their entire LPS structure (the [R8, R10, R18, and R26]-heptoseless mutants) may not be able to further evolve phage resistance through LPS modifications and could be quickly eradicated (LPS-based modifications may become an evolutionary trap). Alternatively, they may evolve different types of phage resistance that either confer a complete or partial immunity and persistence[202] to phage infection (for example, via mutations in the *yajC* gene[146]).

Then, the experimentalist will evolve the second generation of evolved ΦX174 strains by adapting each of the 16 evolved phages (first generation of evolved phage) to their corresponding new *E. coli* C resistant strain (second generation of *E. coli* C mutants) using the phage evolution experiment – Standard (see **Chapter II** section **2.3.5**). Because it is expected that *E. coli* C will become more and more resistant over the round of stepwise evolution, the phage cocktail evolution experiment – Increased diversity and generations (see **Chapter II** section **2.3.9**) would also have to be used to quickly reinfect bacterial mutants that could not be infected or cross-infected. The experimentalist will first pool all evolved phages from the first generation with all the newly evolved phages that successfully overcame their corresponding new resistant *E. coli* C strains (second generation of evolved ΦX174). Then, the experimentalist will serially transfer this cocktail daily, up to four transfers a day, in a mixing environment consisting of (i) each corresponding *E. coli* C resistant strain from the first and second generation of phage-resistant strains that have been overcome (permissive hosts) and (ii) one of the remaining resistant strains of interest from the second generation of phage-resistant mutants (non-permissive host). Whole genome re-sequencing of the evolved phages will be performed to understand how ΦX174 adapts to overcome new phage resistance. The turnover of resistance and counter-resistance rounds will be carried out for an arbitrary number of cycles or until no further evolution from either the phage or the bacterium is possible.

This simple combination of fluctuation experiments and serial transfers can be used to investigate the evolution of all possible phage resistance and overcoming resistance strategies existing between ΦX174 and *E. coli* C beyond the scope of LPS mutations. Based on this idea, my "bottom-up" approach could also be exploited to systematically test the resilience of engineered phage-resistant *E. coli* strains manufactured for different industrial fermentation processes. Phage contamination in bacterial bioprocess is responsible for devastating loss of income due to fermentation failures and inconsistent production[298]. Thus, strategies are developed to generate resistant bacterial strains

with a broad range of phage resistance[299,300]. For a bacterial strain to be considered for industrial fermentation processes, the phage resistance must be laborious to overcome, even by high-titer phages or phage cocktails. My two phage cocktail evolution experiments (see **Chapter II** sections **2.3.8** and **2.3.9**) are designed to rapidly adapt a phage to reinfect these hard-resistant bacterial phenotypes. Evolving phages to overcome these engineered phage-resistant *E. coli* strains could contribute to the development of novel strategies to generate better phage-resistant *E. coli* strains. Conversely, investigating how phages are still capable to generate novelty to overcome diverse phage resistance (despite their small and constrained genomes) could help the development of systematic human and animal phage therapy approaches.

6.2.2. Stepwise evolution experiments can be used to quickly breed phages to infect a large number of phage-resistant bacteria

While bacteria often become broadly resistant to phages via only a single mutation (**Table 3.1**), phages require the stepwise acquisition of multiple mutations to evolve broad infectivity[255,256,280] (**Table 4.2**). ΦX174 and phage λ both reinfect phage-resistant strains via the fixation of stepping stone mutations in the virus' host-recognition (and DNA infection) protein. These stepping stone mutations are required later for a last mutation to arise, allowing the phage to overcome the resistance[255]. These experiments pinpoint the importance of contingency and epistasis between mutations in the acquisition of "key innovations" and demonstrate that the phage-bacteria arms race can be extended despite their mutational asymmetry[14,283,284,301–304].

Evolution experiments that aim to breed viruses with increased host ranges can take advantage of my experimental design. In the case of ΦX174, infecting the resistant hosts could be achieved by coculturing (i) *E. coli* C wildtype, (ii) one of the "easy-resistant" LPS mutants that produce a phage evolutionary intermediate, and (iii) one of the hard-resistant strains of interest. It is important to note that the choice of the easy-resistant bacterium is not trivial and depends on the hard-resistant strain that must be reinfected. For example, adapting ΦX174 to the *hldE* mutants R8 and R10 or the *waaC* mutant R18 leads to the production of an evolved phage with the mutations required for the later infection of the *waaG* and *galU* mutants (**Fig. 4.12B**). If the *rfaH* mutant R25 or the *waaP/pssA* mutant R6 have to be reinfected, then ΦX174 will have to be adapted to the *waaO* mutants R4, R13, R21, R24, or R29 first (**Fig. 4.12B**). However, without precise knowledge of hard and easy-resistant strains, co-culturing two resistant strains together with the wildtype might lead to faster phage evolution than coculturing only a single resistant strain with *E. coli* C wildtype (**Fig. 4.11B**).

6.3   Possible experiments to improve the study

6.3.1.   Generating and using more evolved phages to distinguish additional LPS phenotypes

Specific *E. coli* C resistant strains could not be discriminated from each other despite having predicted LPS structures resulting from mutations introducing loss of functions in distinct LPS genes (**Fig. 5.1**). For example, *E. coli* C R11 (*waaP* mutant) and R33 (*waaT* mutant) both carry a mutation introducing a premature codon stop in their respective LPS genes that should modify their LPS structures in two different ways. *E. coli* C R11 is predicted to conserve the backbone of its LPS intact, while the outer core LPS of R33 is predicted to be truncated (**Fig. 3.7**). However, these two strains cannot be distinguished based on their infection patterns (**Fig. 5.1**). An additional example is the three *waaG* and four *galU* mutants that are infected by an identical set of two evolved phages and cannot be separated from each other (highlighted in blue in **Fig 5.1**). While this result suggests that these strains share the same LPS phenotypes, there may be differences in their LPS structures that cannot be detected on the sole base of the two evolved phages infection.

Similarly, R11 and R33 are sensitive to the same two evolved phages (**Fig. 5.1**). To detect potential differences in their LPS structures, a set of highly LPS specific ΦX174 strains could be evolved for each *E. coli* C resistant mutant, using the setup of the phage cocktail evolution experiments (see **Chapter II** sections **2.3.8** and **2.3.9**). The sensitivity and resistance of R11 and R33 to infection by the newly evolved phages may either provide more confidence in the LPS structures being identical or show that the LPS structures are indeed different as suggested by the mutations present in two different LPS genes.

My three phage evolution experiments (Standard, Increased diversity, and Increased diversity and generations) could also be repeated for the other *E. coli* C resistant strains that could only be infected indirectly but not by the corresponding phage population, increasing the total number of evolved ΦX174 strains from 16 to 31. Several new phage genotypes may arise that could strengthen our ability to distinguish LPS structures.

6.3.2.   Investigating the effects of the LPS mutations on other traits

In bacteria, phage-resistance mutations in genes involved in the LPS biosynthesis, assembly, or regulation often have various pleiotropic effects. Not only the mutations will result in the

modification of the LPS structures, but very often will also affect other traits simultaneously. For example, LPS mutants often show a reduction of the bacterial growth rate or competitive ability[184,282,305], different sensitivity to antibiotic[20,276] (Parab *et al.,* in prep), an unstable outer membrane and a modification of the turgor pressure[121,306], a tendency to form biofilms[268,307], etc. Many phages under consideration for therapeutic depend on the LPS to infect their hosts[276,308]. Since pleiotropic effects resulting from LPS-synthesis mutations can be clinically relevant, understanding the impacts of these mutations on other traits is, therefore, crucial for the long-term success of phages as therapeutic agents for human and animal health.

In my first investigations, the abundance of nutrients and spiking of the growth media with divalent cations such as Ca^{2+} and Mg^{2+} may have reduced the cost of phage resistance (**Fig. 3.10**)[214,216,295]. Thus, a deeper analysis of each LPS mutant's growth phenotype in resource-limited conditions[214,309] may unravel potential fitness defects induced by specific LPS-synthesis mutations. However, the presence of aggregate formation during the growth of the LPS mutants often results in wrong OD600 values and the impossibility of correctly replicating the experiment. Alternatively, the microbial cell numbers could be estimated using DNA quantification[310] or biomass calculation methods[311,312].

Aggregation phenotypes (e.g., speed of sedimentation and size of the aggregates) of each *E. coli* C resistant strain can also be characterised using real-time optical technology (oCelloScope, BioSense Solution)[313]. Because the truncation of the core LPS impairs cell-cell and cell-surface interactions (less motility[264], biofilm and aggregates formation[199,268,307]), I would expect that a bacterial strain with a mutation resulting in a loss of LPS gene function (truncated LPS) will aggregate and sediment faster than a strain that carried a mutation that modulates the activity of the LPS gene (intact LPS structure but potentially different composition). A different composition of the LPS structure may be reflected directly on the aggregation phenotype and show phenotypic variations ranging between *E. coli* C wildtype and the fully truncated mutants. Chromatography and mass-spectrometry could also help determine the final LPS composition of the *E. coli* C resistant strains[217,236,314].

6.4.   Future directions

6.4.1.   Role of environmental spatial structure in the evolution of phage resistance

Aggregates or single colonies of growing bacteria (in a spatially structured environment) are considered to be founded by isogenic clones. However, it has been shown recently that bacterial growth in a spatially structured environment is characterised by an excess of mutational jackpot events, increasing the overall genetic diversity of the bacterial populations in the aggregates or colonies[198,315]. Excesses of jackpot events contribute to the emergence of bacterial sub-populations carrying a bigger proportion of high-frequency mutations compared to bacterial growth in a well-mixed environment[198,315] with complex phenotypes, such as antibiotic resistance[198]. Bacterial mutants are often trapped as "bubbles" or can expand via surfing on the edges of the colony[198], but can be released in the environment when antibiotics are added. It has direct implications for human health. For example, antibiotic therapy used to treat a pathogenic strain that grows in biofilm could release and disseminate antibiotic-resistant bacterial mutants, compromising the long-term success of antibiotic therapy.

It is not unlikely that an excess of jackpot events could also contribute to the emergence of bacterial sub-populations with complex phage-resistance phenotypes in pathogenic biofilms. Similarly, phage treatments could also release and disseminate these phage-resistant mutants in the patient. Future experiments could investigate whether an excess of jackpot events impacts the evolution, the type, and the diversity of phage-resistance strategies. For example, wildtype ΦX174-resistant *E. coli* C can be generated by infecting the same wildtype host cultures in a solid or liquid environment. Both the number and type of phage mutants could be compared between the two environments to see whether the presence of spatial structure has a different impact on the evolution of phage resistance compared to an absence of spatial structure.

The acquisition of the first mutations in the bacterial genome may also impact the *de novo* acquisition of further adaptive mutations[316–318]. For example, mutations in genes resulting in slow growth could allow more time for the bacteria to acquire CRISPR-Cas spacers[192]. In my dataset, I also observed that some *E. coli* C mutants carry more than one mutation in genes without apparent links to phage resistance (**Tables 3.1** to **3.3**). Here, engineered *E. coli* C mutants carrying all possible combinations of mutations could be further evolved in the presence of wildtype ΦX174 to determine their impact on the evolution of phage resistance. ΦX174 could be adapted to *E. coli* C wildtype in the presence or absence of spatial structure to determine whether an excess of

jackpot mutation events is a determining factor in phage diversification and, therefore, should be taken into account during the design of phage therapies.

6.4.2.   Role of complex microbial communities on phage-bacteria co-evolution

Studies of phage-bacteria co-evolution in the laboratory usually study a one-to-one relationship between phages and their hosts. However, in a more ecological context (like *in vivo*), both phages and bacteria not only have to face each other, but also a multitude of other microorganisms plus the host-immune system. In such conditions, remodelling the LPS has direct consequences on the virulence of bacteria[152,319], its competitiveness against other foes and predators[184,282,305,320], its propensity to form biofilm[199,268,307], antibiotic resistance[276] (Parab *et al.*, in prep) and phage resistance[184]. For example, it has been recently demonstrated that a diverse bacterial community amplifies the fitness trade-offs associated with the mutations of phage receptors[320] and drives bacterial evolution toward CRISPR-based resistance[184]. Similarly, future investigations could study how bacteria evolve phage resistance in complex environments or *in vivo* conditions to better predict and control the emergence of phage resistance during phage therapy[305]. Evolution experiments could introduce another competitor during the co-evolution of *E. coli* C and ΦX174. While *E. coli* C could not shift toward CRISPR-based resistance[115], other phage resistance mechanisms may evolve (i.e., abortive infection mechanisms, blocking phage adsorption or DNA entry via the production of extracellular matrices or competitive inhibitors, or restriction-modification systems)[154].

6.5.   Final comment

Well-established model systems such as *E. coli* C and ΦX174 can be used to easily breed phages that can re-infect a wide range of resistant strains (**Fig. 4.10**, **Table 4.2**). But currently, the preferred approach to phage therapy is the isolation of phages from environmental samples that can infect specific bacterial pathogens[19]. However, the isolation, characterization, and adaptation of new phages from the environment can drastically slow their suitability for therapy. Finding new phages is often time-consuming, especially because their safety should be assessed since phages are known to carry dangerous toxins and antibiotic-resistance genes[55,56,321]. The considerable knowledge acquired from decades of research on ΦX174 structure, life cycle, and evolution showed that this phage does not carry virulence genes. ΦX174 is also highly host specific[79], meaning that it will be harmless to the patient's microbiota, in contrast to phages with broader infectivity or antibiotics known to disrupt the microbiota and lead to adverse health outcomes[322].

Furthermore, its use is already approved *in vivo* by the FDA as a marker of immune responses in patients[98,99]. ΦX174 can easily be manipulated in a laboratory, and evolution experiments can extend its host range of formerly resistant strains (**Figs. 4.4, 4.9** and **4.10**), unlocking its potential as a powerful therapeutic agent. For all these reasons, I believe that the use of bacteriophage ΦX174 could constitute a promising avenue for human and animal therapeutic (either in cocktails or in complement with antibiotics).

References

1. Twort, F.W. An investigation on the nature of ultra-microscopic viruses. *Lancet* **186**, 1241–1243 (1915).

2. D'Hérelle, F. Sur un microbe invisible antagoniste des bacilles dysentériques. *C. R. Acad. Sci. Paris* **165**, 373–375 (1917).

3. Abedon, S. T., Thomas-Abedon, C., Thomas, A. & Mazure, H. Bacteriophage prehistory. *Bacteriophage* **1**, 174–178 (2011).

4. M.E. Hankin. L'action bactéricide des eaux de la Jumna et du Gange sur le microbe du choléra. *Ann. Inst. Pasteur* 511–523 (1896).

5. Smith, H. O. & Wilcox, K. W. A restriction enzyme from *Hemophilus influenzae*. I. Purification and general properties. *J. Mol. Biol.* **51**, 379–391 (1970).

6. Weiss, B. & Richardson, C. C. Enzymatic breakage and joining of deoxyribonucleic acid, I. Repair of single-strand breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. *Proc. Natl. Acad. Sci. U.S.A* **57**, 1021–1028 (1967).

7. Zhu, B. Bacteriophage T7 DNA polymerase – sequenase. *Front. Microbiol.* **5** (2014).

8. Sanger, F. *et al.* Nucleotide sequence of bacteriophage. *Nature* **265**, 687–695 (1977).

9. Sanger, F *et al.* The nucleotide sequence of bacteriophage ΦX174. *J. Mol. Biol.* **125**, 225–246 (1978).

10. Fiers, W. *et al.* Complete nucleotide sequence of bacteriophage MS2 RNA: primary and secondary structure of the replicase gene. *Nature* **260**, 500–507 (1976).

11. Salmond, G. P. C. & Fineran, P. C. A century of the phage: past, present and future. *Nat. Rev. Microbiol.* **13**, 777–786 (2015).

12. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).

13. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).

14. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).

15. Summers, W. C. The strange history of phage therapy. *Bacteriophage* **2**, 130–133 (2012).

16. Murray, C. J. *et al.* Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629–655 (2022).

17. World Health Organization. Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance (2021).

18. Centers for Disease Control and Prevention (U.S.). Antibiotic resistance threats in the United States, 2019. https://stacks.cdc.gov/view/cdc/82532 (2019). doi:10.15620/cdc:82532.

19. Chan, B. K. *et al.* Phage treatment of an aortic graft infected with *Pseudomonas aeruginosa*. *Evol. Med. Public Health* **2018**, 60–66 (2018).

20. Chan, B. K. *et al.* Phage selection restores antibiotic sensitivity in MDR *Pseudomonas aeruginosa*. *Sci. Rep.* **6**, 26717 (2016).

21. Dion, M. B., Oechslin, F. & Moineau, S. Phage diversity, genomics and phylogeny. *Nat. Rev. Microbiol.* **18**, 125–138 (2020).

22. Ackermann, H.-W. 5500 Phages examined in the electron microscope. *Arch. Virol.* **152**, 227–243 (2007).

23. Ackermann, H.-W. & Prangishvili, D. Prokaryote viruses studied by electron microscopy. *Arch. Virol.* **157**, 1843–1849 (2012).

24. Adams, M. J. *et al.* 50 years of the International Committee on Taxonomy of Viruses: progress and prospects. *Arch. Virol.* **162**, 1441–1446 (2017).

25. Wang, F., Wang, F., Li, Q. & Xiao, X. A Novel filamentous phage from the deep-sea bacterium *Shewanella piezotolerans* WP3 is induced at low temperature. *J. Bacteriol.* **189**, 7151–7153 (2007).

26. Casjens, S. & Lenk, E. *https://commons.wikimedia.org/wiki/File:Bacteriophage_P22_Casjens_Lenk.png.* (1998).

27. Stedman, K. M., Schleper, C., Rumpf, E. & Zillig, W. Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. *Genetics* **152**, 1397–1405 (1999).

28. Poddar, S. K., Cadden, S. P., Das, J. & Maniloff, J. Heterogeneous progeny viruses are produced by a budding enveloped phage. *Intervirology* **23**, 208–221 (1985).

29. Schuch, R. & Fischetti, V. A. The secret life of the anthrax agent *Bacillus anthracis*: bacteriophage-mediated ecological adaptations. *PLoS One* **4**, e6532 (2009).

30. Kutter, E. & Sulakvelidze, A. *Bacteriophages: Biology and Applications.* (CRC Press, Boca Raton 2005).

31. Xu, J. & Xiang, Y. Membrane penetration by bacterial viruses. *J. Virol.* **91**, e00162-17 (2017).

32. Calendar, R. L. *The Bacteriophages.* (Oxford University Press, New York, 2006).

33. Lawrence, D., Baldridge, M. T. & Handley, S. A. Phages and human health: more than idle hitchhikers. *Viruses* **11**, 587 (2019).

34. Loh, B., Kuhn, A. & Leptihn, S. The fascinating biology behind phage display: filamentous phage assembly. *Molecular Microbiology* **111**, 1132–1138 (2019).

35. Shapiro, J. W., Williams, E. S. C. P. & Turner, P. E. Evolution of parasitism and mutualism between filamentous phage M13 and *Escherichia coli*. *PeerJ* **4**, e2060 (2016).

36. Hay, I. D. & Lithgow, T. Filamentous phages: masters of a microbial sharing economy. *EMBO reports* **20**, e47427 (2019).

37. Abedon, S. T. Disambiguating bacteriophage pseudolysogeny: an historical analysis of lysogeny, pseudolysogeny, and the phage carrier state. *Contemporary Trends in Bacteriophage Research,* 285-307 (Nova Science Publishers, Inc. New York 2009)

38. Mäntynen, S., Laanto, E., Oksanen, H. M., Poranen, M. M. & Díaz-Muñoz, S. L. Black box of phage–bacterium interactions: exploring alternative phage infection strategies. *Open Biology* **11**, 210188.

39. Łoś, M. & Węgrzyn, G. Pseudolysogeny. *Adv. Virus. Res.* **82**, 339–349 (2012).

40. Suttle, C. A. Viruses in the sea. *Nature* **437**, 356–361 (2005).

41. Al-Shayeb, B. *et al.* Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425–431 (2020).

42. Batinovic, S. *et al.* Bacteriophages in natural and artificial environments. *Pathogens* **8**, 100 (2019).

43. Sime-Ngando, T. Environmental bacteriophages: viruses of microbes in aquatic ecosystems. *Front. Microbiol.* **5** (2014).

44. Breitbart, M., Bonnain, C., Malki, K. & Sawaya, N. A. Phage puppet masters of the marine microbial realm. *Nat. Microbiol.* **3**, 754–766 (2018).

45. Braga, L. P. P. *et al.* Impact of phages on soil bacterial communities and nitrogen availability under different assembly scenarios. *Microbiome* **8**, 52 (2020).

46. Navarro, F. & Muniesa, M. Phages in the human body. *Front. Microbiol.* **8** (2017).

47. Zuppi, M., Hendrickson, H. L., O'Sullivan, J. M. & Vatanen, T. Phages in the gut ecosystem. *Front. Cell. Infect. Microbiol.* **11** (2022).

48. Edwards, R. A. *et al.* Global phylogeography and ancient evolution of the widespread human gut virus crAssphage. *Nat. Microbiol.* **4**, 1727–1736 (2019).

49. Angly, F. E. *et al.* The GAAS metagenomic tool and its estimations of viral and microbial average genome size in four major biomes. *PLoS Comput. Biol.* **5**, e1000593 (2009).

50. Lindell, D., Jaffe, J. D., Johnson, Z. I., Church, G. M. & Chisholm, S. W. Photosynthesis genes in marine viruses yield proteins during host infection. *Nature* **438**, 86–89 (2005).

51. Thompson, L. R. *et al.* Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. *Proc. Natl. Acad. Sci. U.S.A* **108**, 757–764 (2011).

52. De Smet, J. *et al.* High coverage metabolomics analysis reveals phage-specific alterations to *Pseudomonas aeruginosa* physiology during infection. *ISME J* **10**, 1823–1835 (2016).

53. Colavecchio, A., Cadieux, B., Lo, A. & Goodridge, L. D. Bacteriophages contribute to the Spread of antibiotic resistance genes among foodborne pathogens of the *Enterobacteriaceae* family – A review. *Front. Microbiol.* **8** (2017).

54. Colomer-Lluch, M., Imamovic, L., Jofre, J. & Muniesa, M. Bacteriophages carrying antibiotic resistance genes in fecal waste from cattle, pigs, and poultry. *Antimicrob Agents Chemother* **55**, 4908–4911 (2011).

55. Krüger, A. & Lucchesi, P. M. A. Shiga toxins and stx phages: highly diverse entities. *Microbiology (Reading, Engl.)* **161**, 451–462 (2015).

56. Jamet, A. *et al.* A widespread family of polymorphic toxins encoded by temperate phages. *BMC Biol.* **15**, 75 (2017).

57. Forde, S. E. & Jessup, C. M. Understanding evolution through the phages. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments.* 391-418 (University of California Press, 2009).

58. Futuyma, D. J. & Bennett, A. F. The importance of experimental studies in evolutionary biology. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments.* 15–30 (University of California Press, 2009).

59. Hershey, A. D. & Chase, M. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *J. Gen. Physiol.* **36**, 39–56 (1952).

60. Benzer, S. Fine structure of a genetic region in bacteriophage. *Proc. Natl. Acad. Sci. U.S.A* **41**, 344–354 (1955).

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

62. Luria, S.E & Delbrück, M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* **28**, 491–511 (1943).

63. Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. & Belshaw, R. Viral mutation rates. *J. Virol.* **84**, 9733–9748 (2010).

64. Sanjuán, R. & Domingo-Calap, P. Mechanisms of viral mutation. *Cell. Mol. Life Sci.* **73**, 4433–4448 (2016).

65. Barrick, J. E. *et al.* Genome evolution and adaptation in a long-term experiment with *Escherichia coli. Nature* **461**, 1243–1247 (2009).

66. Bertels, F., Leemann, C., Metzner, K. J. & Regoes, R. R. Parallel evolution of HIV-1 in a long-term experiment. *Mol. Biol. Evol.* **36**, 2400–2414 (2019).

67. Gould, S. J. *Wonderful Life: The Burgess Shale and the Nature of History.* (W.W. Norton, New York, London, 1989)

68. Blount, Z. D., Lenski, R. E. & Losos, J. B. Contingency and determinism in evolution: replaying life's tape. *Science* **362**, eaam5979 (2018).

69. Lobkovsky, A. & Koonin, E. Replaying the Tape of Life: Quantification of the predictability of evolution. *Front. Genet.* **3** (2012).

70. Brentlinger, K. L. *et al. Microviridae*, a family divided: isolation, characterization, and genome sequence of MH2K, a bacteriophage of the obligate intracellular parasitic bacterium *Bdellovibrio bacteriovorus. J. Bacteriol.* **184**, 1089–1094 (2002).

71. Roux, S., Krupovic, M., Poulet, A., Debroas, D. & Enault, F. Evolution and diversity of the *Microviridae* viral family through a collection of 81 new complete genomes assembled from virome reads. *PLoS One* **7**, e40418 (2012).

72. Doore, S. M. & Fane, B. A. The *Microviridae*: diversity, assembly, and experimental evolution. *Virology* **491**, 45–55 (2016).

73. Bryson, S. J., Thurber, A. R., Correa, A. M. S., Orphan, V. J. & Vega Thurber, R. A novel sister clade to the enterobacteria microviruses (family *Microviridae*) identified in methane seep sediments: DNA phages associated with methane seeps. *Environ. Microbiol.* **17**, 3708–3721 (2015).

74. Wichman, H. A. & Brown, C. J. Experimental evolution of viruses: *Microviridae* as a model system. *Philos. Trans. R. Soc. B: Biol. Sci.* **365**, 2495–2501 (2010).

75. Wichman, H. A., Scott, L. A., Yarber, C. D. & Bull, J. J. Experimental evolution recapitulates natural evolution. *Phil. Trans. R. Soc. Lond. B* **355**, 1677–1684 (2000).

76. Crill, W. D., Wichman, H. A. & Bull, J. J. Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **154**, 27–37 (2000).

77. Bull, J. J. *et al.* Exceptional convergent evolution in a virus. *Genetics* **147**, 1497–1507 (1997).

78. Holder, K. K. & Bull, J. J. Profiles of adaptation in two similar viruses. *Genetics* **159**, 1393–1404 (2001).

79. Michel, A., Clermont, O., Denamur, E. & Tenaillon, O. Bacteriophage ΦX174's ecological niche and the flexibility of its *Escherichia coli* lipopolysaccharide receptor. *Appl. Environ. Microbiol.* **76**, 7310–7313 (2010).

80. Weisbeek, P. J., Van de Pol, J. H. & Van Arkel, G. A. Mapping of host range mutants of bacteriophage ΦX174. *Virology* **52**, 408–416 (1973).

81. Pepin, K. M., Domsic, J. & McKenna, R. Genomic evolution in a virus under specific selection for host recognition. *Infect. Genet. Evol.* **8**, 825–834 (2008).

82. Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. Different trajectories of parallel evolution during viral adaptation. *Science* **285**, 422–424 (1999).

83. Wichman, H. A., Millstein, J. & Bull, J. J. Adaptive molecular evolution for 13,000 phage generations: a possible arms race. *Genetics* **170**, 19–31 (2005).

84. Bull, J. J., Millstein, J., Orcutt, J. & Wichman, H. A. Evolutionary feedback mediated through population density, illustrated with viruses in chemostats. *Am. Nat.* **167**, 39–51 (2006).

85. Bull, J. J., Badgett, M. R., Springman, R. & Molineux, I. J. Genome properties and the limits of adaptation in bacteriophages. *Evolution* **58**, 692–701 (2004).

86. Bull, J. J., Badgett, M. R., Rokyta, D. & Molineux, I. J. Experimental evolution yields hundreds of mutations in a functional viral genome. *J. Mol. Evol.* **57**, 241–248 (2003).

87. Vale, P. F., Choisy, M., Froissart, R., Sanjuán, R. & Gandon, S. The distribution of mutational fitness effects of phage ΦX174 on different hosts. *Evolution* **66**, 3495–3507 (2012).

88. Rokyta, D. R., Joyce, P., Caudle, S. B. & Wichman, H. A. An empirical test of the mutational landscape model of adaptation using a single-stranded DNA virus. *Nat. Genet.* **37**, 441–444 (2005).

89. Poon, A. & Chao, L. The rate of compensatory mutation in the DNA bacteriophage ΦX174. *Genetics* **170**, 989–999 (2005).

90. Poon, A. F. Y. & Chao, L. Functional origins of fitness effect-sizes of compensatory mutations in the DNA of bacteriophage ΦX174. *Evolution* **60**, 2032–2043 (2006).

91. Pepin, K. M. & Wichman, H. A. Variable epistatic effects between mutations at host recognition sites in ΦX174 bacteriophage. *Evolution* **61**, 1710–1724 (2007).

92. Rokyta, D. R., Burch, C. L., Caudle, S. B. & Wichman, H. A. Horizontal gene transfer and the evolution of microvirid coliphage genomes. *J. Bacteriol.* **188**, 1134–1142 (2006).

93. La Rosa, R., Rossi, E., Feist, A. M., Johansen, H. K. & Molin, S. Compensatory evolution of *Pseudomonas aeruginosa*'s slow growth phenotype suggests mechanisms of adaptation in cystic fibrosis. *Nat. Commun.* **12**, 3186 (2021).

94. Zhang, G., Meredith, T. C. & Kahne, D. On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr. Opin. Microbiol.* **16**, 779–785 (2013).

95. Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593–656 (2003).

96. Mutalik, V. K. *et al.* High-throughput mapping of the phage resistance landscape in *E. coli*. *PLoS Biol.* **18**, e3000877 (2020).

97. Simpson, B. W. & Trent, M. S. Pushing the envelope: LPS modifications and their consequences. *Nat. Rev. Microbiol.* **17**, 403–416 (2019).

98. Bearden, C. M. *et al.* Rituximab inhibits the *in vivo* primary and secondary antibody response to a neoantigen, bacteriophage ΦX174. *Am. J. Transplant.* **5**, 50–57 (2005).

99. Rubinstein, A. *et al.* Progressive specific immune attrition after primary, secondary and tertiary immunizations with bacteriophage ΦX174 in asymptomatic HIV-1 infected patients. *AIDS* **14**, F55-62 (2000).

100. Picelli, S. *et al.* Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).

101. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using *breseq. Methods Mol. Biol.* **1151**, 165–188 (2014).

102. Barrick, J. E. *et al.* Identifying structural variation in haploid microbial genomes from short-read resequencing data using *breseq. BMC Genomics* **15**, 1039 (2014).

103. Deatherage, D. E., Traverse, C. C., Wolf, L. N. & Barrick, J. E. Detecting rare structural variation in evolving microbial populations from new sequence junctions using *breseq. Front. Genet.* **5** (2015).

104. Van Charante, F., Holtappels, D., Blasdel, B. & Burrowes, B. Isolation of bacteriophages. *Bacteriophages: Biology, Technology, Therapy.* 1–32 (Springer International Publishing, 2019).

105. Bono, L. M., Gensel, C. L., Pfennig, D. W. & Burch, C. L. Competition and the origins of novelty: experimental evolution of niche-width expansion in a virus. *Biol. Lett.* **9**, 20120616–20120616 (2012).

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

107. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700 (2002).

108. Qian, J., Garrett, T. A. & Raetz, C. R. H. *In vitro* assembly of the outer core of the lipopolysaccharide from *Escherichia coli* K-12 and *Salmonella typhimurium. Biochemistry* **53**, 1250–1262 (2014).

109. Klein, G. & Raina, S. Regulated assembly of LPS, its structural alterations and cellular response to LPS defects. *IJMS* **20**, 356 (2019).

110. Amor, K. *et al.* Distribution of core oligosaccharide types in lipopolysaccharides from *Escherichia coli. Infect. Immun.* **68**, 1116–1124 (2000).

111. Klein, G. *et al.* Molecular and structural basis of inner core lipopolysaccharide alterations in *Escherichia coli*: Incorporation of glucuronic acid and phosphoethanolamine in the heptose region. *J. Biol. Chem.* **288**, 8111–8127 (2013).

112. Yethon, J. A., Heinrichs, D. E., Monteiro, M. A., Perry, M. B. & Whitfield, C. Involvement of *waaY*, *waaQ*, and *waaP* in the modification of *Escherichia coli* lipopolysaccharide and their role in the formation of a stable outer membrane. *J. Biol. Chem.* **273**, 26310–26316 (1998).

113. Whitfield, C. *et al.* Assembly of the R1-type core oligosaccharide of *Escherichia coli* lipopolysaccharide. *J. Endotoxin Res.* **5**, 151–156 (1999).

114. Van der Ley, P., de Graaff, P. & Tommassen, J. Shielding of *Escherichia coli* outer membrane proteins as receptors for bacteriophages and colicins by O-antigenic chains of lipopolysaccharide. *J. Bacteriol.* **168**, 449–451 (1986).

115. Król, J. E. *et al.* Genome rearrangements induce biofilm formation in *Escherichia coli* C – an old model organism with a new application in biofilm research. *BMC Genomics* **20**, 767 (2019).

116. Belunis, C. J., Clementz, T., Carty, S. M. & Raetz, C. R. H. Inhibition of lipopolysaccharide biosynthesis and cell growth following inactivation of the *kdtA* gene in *Escherichia coli. J. Biol. Chem.* **270**, 27646–27652 (1995).

117. Klein, G., Lindner, B., Brabetz, W., Brade, H. & Raina, S. *Escherichia coli* K-12 Suppressor-free mutants lacking early glycosyltransferases and late acyltransferases: Minimal lipopolysaccharide structure and induction of envelope stress response. *J. Biol. Chem.* **284**, 15369–15389 (2009).

118. Jansson, P.-E., Lindberg, B., Lindberg, A. A. & Wollin, R. Structural studies on the hexose region of the core in lipopolysaccharides from *Enterobacteriaceae*. *Eur. J. Biochem.* **115**, 571–577 (1981).

119. Heinrichs, D. E., Yethon, J. A. & Whitfield, C. Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. *Mol. Microbiol.* **30**, 221–232 (1998).

120. Vinogradov, E. V. *et al.* The structures of the carbohydrate backbones of the lipopolysaccharides from *Escherichia coli* rough mutants F470 (R1 core type) and F576 (R2 core type): LPS from *E. coli* R1 and R2 core types. *Eur. J. Biochem.* **261**, 629–639 (1999).

121. Yethon, J. A., Vinogradov, E., Perry, M. B. & Whitfield, C. Mutation of the lipopolysaccharide core glycosyltransferase encoded by *waaG* destabilizes the outer membrane of *Escherichia coli* by interfering with core phosphorylation. *J. Bacteriol.* **182**, 5620–5623 (2000).

122. Heinrichs, D. E., Yethon, J. A., Amor, P. A. & Whitfield, C. The assembly system for the outer core portion of R1- and R4-type lipopolysaccharides of *Escherichia coli*: the R1 core-specific β-glucosyltransferase provides a novel attachment site for O-polysaccharides. *J. Biol. Chem.* **273**, 29497–29505 (1998).

123. Leipold, M. D., Vinogradov, E. & Whitfield, C. Glycosyltransferases involved in biosynthesis of the outer core region of *Escherichia coli* lipopolysaccharides exhibit broader substrate specificities than is predicted from lipopolysaccharide structures. *J. Biol. Chem.* **282**, 26786–26792 (2007).

124. Kneidinger, B. *et al.* Biosynthesis Pathway of ADP-L-glycero-β-D-manno-Heptose in *Escherichia coli*. *J. Bacteriol.* **184**, 363-369 (2002).

125. McArthur, F., Andersson, C. E., Loutet, S., Mowbray, S. L. & Valvano, M. A. Functional analysis of the glycero-manno-heptose 7-phosphate kinase domain from the bifunctional HldE protein, which is involved in ADP-L-glycero-D-manno-heptose biosynthesis. *J. Bacteriol.* **187**, 5292–5300 (2005).

126. Frey, P. A. The Leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose. *FASEB j.* **10**, 461–470 (1996).

127. Holden, H. M., Rayment, I. & Thoden, J. B. Structure and function of enzymes of the Leloir Pathway for galactose metabolism. *J. Biol. Chem.* **278**, 43885–43888 (2003).

128. Weissborn, A. C., Liu, Q., Rumley, M. K. & Kennedy, E. P. UTP: α-D-glucose-L -phosphate uridylyltransferase of *Escherichia coli*: isolation and DNA sequence of the *galU* gene and purification of the enzyme. *J. Bacteriol.* .**176**, 2611-2618 (1994).

129. Genevaux, P., Bauda, P., DuBow, M. S. & Oudega, B. Identification of Tn 10 insertions in the *rfaG*, *rfaP*, and *galU* genes involved in lipopolysaccharide core biosynthesis that affect *Escherichia coli* adhesion. *Arch. Microbiol.* **172**, 1–8 (1999).

130. Pierson, D. E. & Carlson, S. Identification of the *galE* gene and a *galE* homolog and characterization of their roles in the biosynthesis of lipopolysaccharide in a serotype O:8 strain of *Yersinia enterocolitica*. *J. Bacteriol.* **178**, 5916–5924 (1996).

131. Schnaitman, C. A. & Austin, E. A. Efficient incorporation of galactose into lipopolysaccharide by *Escherichia coli* K-12 strains with polar *galE* mutations. *J. Bacteriol.* **172**, 5511–5513 (1990).

132. Karp, P. D. *et al.* The BioCyc collection of microbial genomes and metabolic pathways. *Brief Bioinform* **20**, 1085–1093 (2019).

133. Bailey, M. J. A., Hughes, C. & Koronakis, V. RfaH and the *ops* element, components of a novel system controlling bacterial transcription elongation. *Mol. Microbiol.* **26**, 845–851 (1997).

134. Bedford, E., Tabor, S. & Richardson, C. C. The thioredoxin binding domain of bacteriophage T7 DNA polymerase confers processivity on *Escherichia coli* DNA polymerase I. *Proc. Natl. Acad. Sci. U.S.A* **94**, 479–484 (1997).

135. Nagy, G. *et al.* Down-regulation of key virulence factors makes the *Salmonella enterica* Serovar *Typhimurium rfaH* mutant a promising live-attenuated vaccine candidate. *IAI* **74**, 5914–5925 (2006).

136. Chao, L., Levin, B. R. & Stewart, F. M. A Complex community in a simple habitat: an experimental study with bacteria and phage. *Ecology* **58**, 369–378 (1977).

137. Betts, A., Gray, C., Zelek, M., MacLean, R. C. & King, K. C. High parasite diversity accelerates host adaptation and diversification. *Science* **360**, 907–911 (2018).

138. Papkou, A. *et al.* The genomic basis of Red Queen dynamics during rapid reciprocal host–pathogen coevolution. *Proc. Natl. Acad. Sci. U.S.A* **116**, 923–928 (2019).

139. Wandro, S. *et al.* Predictable molecular adaptation of coevolving *Enterococcus faecium* and Lytic Phage EfV12-Φ1. *Front. Microbiol.* **9**, 3192 (2019).

140. Perry, E. B., Barrick, J. E. & Bohannan, B. J. M. The molecular and genetic basis of repeatable coevolution between *Escherichia coli* and bacteriophage T3 in a laboratory microcosm. *PLoS One* **10**, e0130639 (2015).

141. McDonald, M. J. Microbial experimental evolution – a proving ground for evolutionary theory and a tool for discovery. *EMBO Rep* **20**, (2019).

142. Van den Bergh, B., Swings, T., Fauvart, M. & Michiels, J. Experimental design, population dynamics, and diversity in microbial experimental evolution. *Microbiol. Mol. Biol. Rev.* **82**, e00008-18 (2018).

143. La Rosa, R., Rossi, E., Feist, A. M., Johansen, H. K. & Molin, S. Compensatory evolution of Pseudomonas aeruginosa's slow growth phenotype suggests mechanisms of adaptation in cystic fibrosis. *Nat Commun* **12**, 3186 (2021).

144. Fang, J. & Wei, Y. Expression, purification and characterization of the *Escherichia coli* integral membrane protein YajC. *Protein Pept. Lett.* **18**, 601–608 (2011).

145. Schulze, R. J. *et al.* Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. *Proc. Natl. Acad. Sci. U.S.A* **111**, 4844–4849 (2014).

146. Bohm, K. *et al.* Genes affecting progression of bacteriophage P22 infection in *Salmonella* identified by transposon and single gene deletion screens: host genes affecting phage P22 infection. *Mol. Microbiol.* **108**, 288–305 (2018).

147. Lind, P. A., Libby, E., Herzog, J. & Rainey, P. B. Predicting mutational routes to new adaptive phenotypes. *eLife* **8**, e38822 (2019).

148. Moxon, E. R., Rainey, P. B., Nowak, M. A. & Lenski, R. E. Adaptive evolution of highly mutable loci in pathogenic bacteria. *Curr. Biol.* **4**, 24–33 (1994).

149. Pagnout, C. *et al.* Pleiotropic effects of *rfa*-gene mutations on *Escherichia coli* envelope properties. *Sci Rep* **9**, 9696 (2019).

150. Seregina, T. A. *et al.* The Inactivation of LPS biosynthesis genes in E. *coli* Cells leads to oxidative stress. *Cells* **11**, 2667 (2022).

151. Matsuura, M. Structural modifications of bacterial lipopolysaccharide that facilitate Gram-Negative bacteria evasion of host innate immunity. *Front. Immunol.* **4** (2013).

152. Maldonado, R. F., Sá-Correia, I. & Valvano, M. A. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol. Rev.* **40**, 480–493 (2016).

153. Hancock, R. E. & Reeves, P. Lipopolysaccharide-deficient, bacteriophage-resistant mutants of *Escherichia coli* K-12. *J. Bacteriol.* **127**, 98–108 (1976).

154. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8**, 317–327 (2010).

155. Kulikov, E. E., Golomidova, A. K., Prokhorov, N. S., Ivanov, P. A. & Letarov, A. V. High-throughput LPS profiling as a tool for revealing of bacteriophage infection strategies. *Sci. Rep.* **9**, 2958 (2019).

156. Burmeister, A. R., Sullivan, R. M. & Lenski, R. E. Fitness costs and benefits of resistance to phage lambda in experimentally evolved *Escherichia coli*. *Evolution in Action: Past, Present and Future: A Festschrift in Honor of Erik D. Goodman.* 123–143 (Springer International Publishing, 2020).

157. Zulk, J. J. *et al.* Phage resistance accompanies reduced fitness of uropathogenic *Escherichia coli* in the urinary environment. *mSphere* **7**, e00345-22 (2022).

158. Mangalea, M. R. & Duerkop, B. A. Fitness trade-offs resulting from bacteriophage resistance potentiate synergistic antibacterial strategies. *Infect. Immun.* **88**, e00926-19 (2020).

159. Wright, R. C. T., Friman, V.-P., Smith, M. C. M. & Brockhurst, M. A. Resistance evolution against phage combinations depends on the timing and order of exposure. *mBio* **10**, e01652-19 (2019).

160. Monod, J. The growth of bacterial cultures. *Annu. Rev. Microbiol.* **3**, 371–394 (1949).

161. Haaber, J., Cohn, M. T., Petersen, A. & Ingmer, H. Simple method for correct enumeration of *Staphylococcus aureus*. *J. Microbiol. Methods* **125**, 58–63 (2016).

162. Ou, F., McGoverin, C., Swift, S. & Vanholsbeeck, F. Absolute bacterial cell enumeration using flow cytometry. *J. Appl. Microbiol.* **123**, 464–477 (2017).

163. Butterfield, C. T. Comparison of the enumeration of bacteria by means of solid and liquid media. *Public Health Rep. (1896-1970)* **48**, 1292–1297 (1933).

164. Sanz, R., Battu, S., Puignou, L., Galceran, M. T. & Cardot, P. J. P. Sonication effect on cellular material in sedimentation and gravitational field flow fractionation. *J. Chromatogr. A.* **1002**, 145–154 (2003).

165. Joyce, E., Phull, S. S., Lorimer, J. P. & Mason, T. J. The development and evaluation of ultrasound for the treatment of bacterial suspensions. A study of frequency, power and sonication time on cultured *Bacillus* species. *Ultrasonics Sonochemistry* **10**, 315–318 (2003).

166. Belunis, C. J. & Raetz, C. R. Biosynthesis of endotoxins. Purification and catalytic properties of 3-deoxy-D-manno-octulosonic acid transferase from *Escherichia coli*. *J. Biol. Chem.* **267**, 9988–9997 (1992).

167. Reynolds, C. M. & Raetz, C. R. H. Replacement of lipopolysaccharide with free lipid A molecules in *Escherichia coli* mutants lacking all core sugars. *Biochem.* **48**, 9627–9640 (2009).

168. Morrison, J. P. & Tanner, M. E. A Two-base mechanism for *Escherichia coli* ADP-L-glycero-D-manno-heptose 6-epimerase. *Biochem.* **46**, 3916–3924 (2007).

169. Igarashi, K. & Kashiwagi, K. Polyamine transport in bacteria and yeast. Biochem. J. **344**. 633-642 (1999).

170. Pistocchi, R., Kashiwagi, K., Kobayashi, H. & Igarashi, K. Characteristics of the operon for a putrescine transport system that maps at 19 minutes on the *Escherichia coli* chromosome. *J. Biol. Chem.* **1**, 146-152 (1993).

171. Djoko, K. Y., Xiao, Z. & Wedd, A. G. Copper resistance in *E. coli*: the multicopper oxidase PcoA catalyzes oxidation of copper(I) in Cu(I)Cu(II)-PcoC. *ChembioChem* **9**, 1579–1582 (2008).

172. Lee, S. M. *et al.* The Pco proteins are involved in periplasmic copper handling in *Escherichia coli. Biochem. Biophys. Res. Commun.* **295**, 616–620 (2002).

173. Huffman, D. L. *et al.* Spectroscopy of Cu(II)-PcoC and the multicopper oxidase function of PcoA, two essential components of *Escherichia coli pco* copper resistance operon. *Biochem.* **41**, 10046–10055 (2002).

174. Rensing, C. & Grass, G. *Escherichia coli* mechanisms of copper homeostasis in a changing environment. *FEMS Microbiol. Rev.* **27**, 197–213 (2003).

175. Kaval, K. G. & Garsin, D. A. Ethanolamine utilization in bacteria. *mBio* **9**, e00066-18 (2018).

176. Kofoid, E., Rappleye, C., Stojiljkovic, I. & Roth, J. The 17-Gene ethanolamine (*eut*) operon of *Salmonella typhimurium* encodes five homologues of carboxysome shell proteins. *J. Bacteriol.* **181**, 5317–5329 (1999).

177. Tanaka, S., Sawaya, M. R. & Yeates, T. O. Structure and mechanisms of a protein-based organelle in *Escherichia coli. Science* **327**, 81–84 (2010).

178. Roof, D. M. & Roth, J. R. Functions required for vitamin B12-dependent ethanolamine utilization in *Salmonella typhimurium. J. Bacteriol.* **171**, 3316–3323 (1989).

179. Penrod, J. T., Mace, C. C. & Roth, J. R. A pH-sensitive function and phenotype: evidence that EutH facilitates diffusion of uncharged ethanolamine in *Salmonella enterica. J. Bacteriol.* **186**, 6885–6890 (2004).

180. Goodall, E. C. A. *et al.* The essential genome of *Escherichia coli* K-12. *mBio* **9**, e02096-17 (2018).

181. Raetz, C. R. Phosphatidylserine synthetase mutants of *Escherichia coli*. Genetic mapping and membrane phospholipid composition. *J. Biol. Chem.* **251**, 3242–3249 (1976).

182. Shi, W., Bogdanov, M., Dowhan, W. & Zusman, D. R. The *pss* and *psd* genes are required for motility and chemotaxis in *Escherichia coli. J. Bacteriol.* **175**, 7711–7714 (1993).

183. Raetz, C. R. & Dowhan, W. Biosynthesis and function of phospholipids in *Escherichia coli. J. Biol. Chem.* **265**, 1235–1238 (1990).

184. Alseth, E. O. *et al.* Bacterial biodiversity drives the evolution of CRISPR-based phage resistance. *Nature* **574**, 549–552 (2019).

185. Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60**, 174–182 (2005).

186. Couvin, D. *et al.* CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* **46**, 246–251 (2018).

187. Zhang, Q. & Ye, Y. Not all predicted CRISPR–Cas systems are equal: isolated cas genes and classes of CRISPR like elements. *BMC Bioinform.* **18**, 92 (2017).

188. Hille, F. & Charpentier, E. CRISPR-Cas: biology, mechanisms and relevance. *Philos. Trans. of R. Soc. B: Biol. Sci.* **371**, 20150496 (2016).

189. Garrett, S. C. Pruning and tending immune memories: spacer dynamics in the CRISPR array. *Front. Microbiol.* **12** (2021).

190. Lovett, S. T. Encoded errors: mutations and rearrangements mediated by misalignment at repetitive DNA sequences. *Mol. Microbiol.* **52**, 1243–1253 (2004).

191. Penrod, J. T. & Roth, J. R. Conserving a volatile metabolite: a role for carboxysome-like organelles in salmonella enterica. *J. bacteriol.* **188**, 2865–2874 (2006).

192. Dimitriu, T. *et al.* Bacteriostatic antibiotics promote CRISPR-Cas adaptive immunity by enabling increased spacer acquisition. *Cell Host & Microbe* **30**, 31-40.e5 (2022).

193. Golec, P., Karczewska-Golec, J., Łoś, M. & Węgrzyn, G. Bacteriophage T4 can produce progeny virions in extremely slowly growing *Escherichia coli* host: comparison of a mathematical model with the experimental data. *FEMS Microbiol. Lett.* **351**, 156–161 (2014).

194. Nabergoj, D., Modic, P. & Podgornik, A. Effect of bacterial growth rate on bacteriophage population growth rate. *MicrobiologyOpen* **7**, e00558 (2018).

195. Rabinovitch, A., Fishov, I., Hadas, H., Einav, M. & Zaritsky, A. Bacteriophage T4 development in *Escherichia coli* is growth rate dependent. *J. Theor. Biol.* **216**, 1–4 (2002).

196. Bull, J. J. *et al.* Phage-bacterial dynamics with spatial structure: self-organization around phage sinks can promote increased cell densities. *Antibiotics (Basel)* **7** (2018).

197. Schenk, H. & Sieber, M. Bacteriophage can promote the emergence of physiologically sub-optimal host phenotypes. *bioRxiv* (2019). http://biorxiv.org/lookup/doi/10.1101/621524 (2019) doi:10.1101/621524.

198. Fusco, D., Gralka, M., Kayser, J., Anderson, A. & Hallatschek, O. Excess of mutational jackpot events in expanding populations revealed by spatial Luria–Delbrück experiments. *Nat. Commun* **7**, 12760 (2016).

199. Testa, S. *et al.* Spatial structure affects phage efficacy in infecting dual-strain biofilms of *Pseudomonas aeruginosa. Commun. Biol.* **2**, 405 (2019).

200. Drake, J. W. Too many mutants with multiple mutations. *Crit. Rev. Biochem. Mol. Biol.* **42**, 247–258 (2007).

201. Drake, J. W., Bebenek, A., Kissling, G. E. & Peddada, S. Clusters of mutations from transient hypermutability. *Proc. Natl. Acad. Sci. U.S.A* **102**, 12849–12854 (2005).

202. Balaban, N. Q. Definitions and guidelines for research on antibiotic persistence. *Nat. Rev. Microbiol.* **17**, 441-448 (2019).

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

204. Jamasbi, R. J. & Paulissen, L. J. Influence of bacteriological media constituents on the reproduction of *Salmonella enteritidis* bacteriophages. *Antonie van Leeuwenhoek* **44**, 49–57 (1978).

205. Ramesh, N., Archana, L., Madurantakam Royam, M., Manohar, P. & Eniyan, K. 2019. Effect of various bacteriological media on the plaque morphology of *Staphylococcus* and *Vibrio* phages. *Access Microbiology* **1**, e000036.

206. Walakira, J. k. *et al.* Identification and characterization of bacteriophages specific to the catfish pathogen, *Edwardsiella ictaluri. J. Appl. Microbiol.* **105**, 2133–2142 (2008).

207. Cvirkaitė-Krupovič, V., Krupovič, M., Daugelavičius, R. & Bamford, D. H. Calcium ion-dependent entry of the membrane-containing bacteriophage PM2 into its *Pseudoalteromonas* host. *Virol.* **405**, 120–128 (2010).

208. Bandara, N., Jo, J., Ryu, S. & Kim, K.-P. Bacteriophages BCP1-1 and BCP8-2 require divalent cations for efficient control of *Bacillus cereus* in fermented foods. *Food Microbiol.* **31**, 9–16 (2012).

209. Rowatt, E. The role of bivalent ions in the inactivation of bacteriophage ΦX174 by lipopolysaccharide from *Escherichia coli* C. *Biochem. J.* **223**, 7 (1984).

210. Rowatt, E. The effect of multivalent ions on the inactivation of bacteriophage ΦX174 by lipopolysaccharide from *Escherichia coli* C. *Biochem. J.* **231**, 4 (1985).

211. Ács, N., Gambino, M. & Brøndsted, L. Bacteriophage enumeration and detection methods. *Front. Microbiol.* **11** (2020).

212. Clokie, M. R. J., Kropinski, A. M., & Lavigne, R. *Bacteriophages: Methods and Protocols, Volume 3.* (Springer, Humana New York, 2018).

213. Serwer, P., Hayes, S. J., Thomas, J. A. & Hardies, S. C. Propagating the missing bacteriophages: a large bacteriophage in a new class. *Virol. J.* **4**, 21 (2007).

214. Clifton, L. A. *et al.* Effect of divalent cation removal on the structure of Gram-negative bacterial outer membrane models. *Langmuir* **31**, 404–412 (2015).

215. Nixdorff, K., Gmeiner, J. & Martin, H. H. Interaction of lipopolysaccharide with detergents and its possible role in the detergent resistance of the outer membrane of Gram-negative bacterua. *Biochim. Biophys. Acta. Biomembr.* **510**, 87–98 (1978).

216. Jeworrek, C. *et al.* Effects of specific versus nonspecific ionic interactions on the structure and lateral organization of lipopolysaccharides. *Biophys. J.* **100**, 2169–2177 (2011).

217. Banoub, J. H., El Aneed, A., Cohen, A. M. & Joly, N. Structural investigation of bacterial lipopolysaccharides by mass spectrometry and tandem mass spectrometry. *Mass. Spectrom. Rev.* **29**, 606–650 (2010).

218. Inagaki, M., Wakashima, H., Kato, M., Kaitani, K. & Nishikawa, S. Crucial role of the lipid part of lipopolysaccharide for conformational change of minor spike H protein of bacteriophage ΦX174. *FEMS Microbiol. Lett.* **251**, 305–311 (2005).

219. Inagaki, M. *et al.* Different contributions of the outer and inner R-core residues of lipopolysaccharide to the recognition by spike H and G proteins of bacteriophage ΦX174. *FEMS Microbiol. Lett.* **226**, 221–227 (2003).

220. Jiménez, N. *et al.* Effects of lipopolysaccharide biosynthesis mutations on K1 polysaccharide association with the *Escherichia coli* cell Surface. *J. Bacteriol.* **194**, 3356–3367 (2012).

221. Caspar, D. L. D. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1–24 (1962).

222. Sinsheimer, R. L. A single-stranded deoxyribonucleic acid from bacteriophage ΦX174. *J. Mol. Biol.* **1**, 43–53 (1959).

223. Barrell, B. G., Air, G. M. & Hutchison, C. A. Overlapping genes in bacteriophage ΦX174. *Nature* **264**, 34–41 (1976).

224. Eisenberg, S. & Ascarelli, R. The A* protein of ΦX174 is an inhibitor of DNA replication. *Nucleic Acids Res.* **9**, 1991–2002 (1981).

225. Gillam, S., Atkinson, T., Markham, A. & Smith, M. Gene *K* of bacteriophage ΦX174 codes for a protein which affects the burst size of phage production. *J. Virol.* **53**, 708–709 (1985).

226. Peck, K. M. & Lauring, A. S. Complexities of viral mutation rates. *J. Virol.* **92**, e01031-17 (2018).

227. Sanjuán, R. & Domingo-Calap, P. Mechanisms of viral mutation. *Cell. Mol. Life Sci.* **73**, 4433–4448 (2016).

228. Cuevas, J. M., Pereira-Gómez, M. & Sanjuán, R. Mutation rate of bacteriophage ΦX174 modified through changes in GATC sequence context. *Infect. Genet. Evol.* **11**, 1820–1822 (2011).

229. Pereira-Gómez, M. & Sanjuán, R. Effect of mismatch repair on the mutation rate of bacteriophage ΦX174. *Virus Evol.* **1**, vev010 (2015).

230. Wyrzykowski, J. & Volkert, M. R. The *Escherichia coli* methyl-directed mismatch repair system repairs base pairs containing oxidative lesions. *J. Bacteriol.* **185**, 1701–1704 (2003).

231. Jiricny, J. Postreplicative mismatch repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012633 (2013).

232. Król, J. E. *et al.* Invasion of *E. coli* biofilms by antibiotic resistance plasmids. *Plasmid* **70**, 110–119 (2013).

233. Feige, Ulrich & Stirm, Stephan. On the structure of *Escherichia coli* C cell wall lipopolysaccharide core and its ΦX174 receptor region. *Biochem. Biophys. Res. Commun.* **71**, 8 (1976).

234. Inagaki, M. *et al.* Characterization of the binding of spike H protein of bacteriophage ΦX174 with receptor lipopolysaccharides. *J. Biochem.* **127**, 577–583 (2000).

235. Kawaura, T. *et al.* Recognition of receptor lipopolysaccharides by spike G protein of Bacteriophage ΦX174. *Biosci. Biotech. Biochem.* **64**, 1993–1997 (2000).

236. Kojima, H., Inagaki, M., Tomita, T. & Watanabe, T. Diversity of non-stoichiometric substitutions on the lipopolysaccharide of *E. coli* C demonstrated by electrospray ionization single quadrupole mass spectrometry. *Rapid Commun. Mass Spectrom.* **24**, 43–48 (2010).

237. McKenna, R. *et al.* Atomic structure of single-stranded DNA bacteriophage ΦX174 and its functional implications. *Nature* **355**, 137–143 (1992).

238. Sun, Y. *et al.* Structural changes of tailless bacteriophage ΦX174 during penetration of bacterial cell walls. *Proc. Natl. Acad. Sci. U.S.A* **114**, 13708–13713 (2017).

239. Bläsi, U., Henrich, B. & Lubitz, W. Lysis of *Escherichia coli* by cloned ΦX174 gene *E* depends on its expression. *J. Gen. Microbiol.* **131**, 1107–1114 (1985).

240. Hermisson, J. & Pennings, P. S. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**, 2335–2352 (2005).

241. Barrett, R. D. H. & Schluter, D. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008).

242. Bleichrodt, J. F. & Verheij, W. S. D. Mutagenesis by ultraviolet radiation in bacteriophage ΦX174: on the mutation stimulating processes induced by ultraviolet radiation in the host bacterium. *Molec. Gen. Genet.* **135**, 19–27 (1974).

243. Gallet, R., Kannoly, S. & Wang, I.-N. Effects of bacteriophage traits on plaque formation. *BMC Microbiol.* **11**, 181 (2011).

244. Nojima, T., Minamimoto, S. & Fukumi, H. A consideration on the factors influencing the plaque size in a bacteriophage (MT-bacteriophage). *Jpn. J. Med. Sci. Biol.* **8**, 25–31 (1955).

245. Abedon, S. T. & Yin, J. Bacteriophage plaques: theory and analysis. *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions.* 161–174 (Springer, Humana Press, 2009).

246. Abedon, S. T. & Culler, R. R. Bacteriophage evolution given spatial constraint. *J. Theor. Biol.* **248**, 111–119 (2007).

247. Roychoudhury, P., Shrestha, N., Wiss, V. R. & Krone, S. M. Fitness benefits of low infectivity in a spatially structured population of bacteriophages. *Proc. R. Soc. B.* **281**, 20132563 (2014).

248. Tom, E. F., Molineux, I. J., Paff, M. L. & Bull, J. J. Experimental evolution of UV resistance in a phage. *PeerJ* **6**, e5190 (2018).

249. Ally, D. *et al.* The impact of spatial structure on viral genomic diversity generated during adaptation to thermal Stress. *PLoS One* **9**, e88702 (2014).

250. Sant, D. G., Woods, L. C., Barr, J. J. & McDonald, M. J. Host diversity slows bacteriophage adaptation by selecting generalists over specialists. *Nat. Ecol. Evol.* **5**, 350–359 (2021).

251. Cox, J. & Putonti, C. Mechanisms responsible for a ΦX174 mutant's ability to infect *Escherichia coli* by phosphorylation. *JVI* **84**, 4860–4863 (2010).

252. Remold, S. K., Rambaut, A. & Turner, P. E. Evolutionary genomics of host adaptation in vesicular stomatitis virus. *Mol. Biol. Evol.* **25**, 1138–1147 (2008).

253. Bertels, F., Metzner, K. J. & Regoes, R. Convergent evolution as an indicator for selection during acute HIV-1 infection. *Peer Community J.* **1**, e4 (2021).

254. Sun, L. *et al.* Icosahedral bacteriophage ΦX174 forms a tail for DNA transport during infection. *Nature* **505**, 432–435 (2014).

255. Meyer, J. R. *et al.* Repeatability and contingency in the evolution of a key innovation in phage Lambda. *Science* **335**, 428–432 (2012).

256. Burmeister, A. R., Lenski, R. E. & Meyer, J. R. Host coevolution alters the adaptive landscape of a virus. *Proc. R. Soc. B* **283**, 20161528 (2016).

257. Gupta, A. *et al.* Host-parasite coevolution promotes innovation through deformations in fitness landscapes. *eLife* **11**, e76162 (2022).

258. McKenna, R., Ilag, L. L. & Rossmann, M. G. Analysis of the single-stranded DNA bacteriophage ΦX174, refined at a resolution of 3.0 Å. *J. Mol. Biol.* **237**, 517–543 (1994).

259. DePristo, M. A., Weinreich, D. M. & Hartl, D. L. Missense meanderings in sequence space: a biophysical view of protein evolution. *Nat. Rev. Genet.* **6**, 678–687 (2005).

260. Washizaki, A., Yonesaki, T. & Otsuka, Y. Characterization of the interactions between *Escherichia coli* receptors, LPS and OmpC, and bacteriophage T4 long tail fibers. *MicrobiologyOpen* **5**, 1003–1015 (2016).

261. Jeong, H. *et al.* Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J. Mol. Biol.* **394**, 644–652 (2009).

262. Bone, D. R. & Dowell, C. E. A mutant of bacteriophage ΦX174 which infects *E. coli* K12 strains: I. Isolation and partial characterization of ΦXtB. *Virol.* **52**, 319–329 (1973).

263. Ford, B. E. *et al.* Frequency and fitness consequences of bacteriophage Φ6 host range mutations. *PLoS One* **9**, e113078 (2014).

264. Lindhout, T., Lau, P. C. Y., Brewer, D. & Lam, J. S. Truncation in the core oligosaccharide of lipopolysaccharide affects flagella-mediated motility in *Pseudomonas aeruginosa* PAO1 via modulation of cell surface attachment. *Microbiol.* **155**, 3449–3460 (2009).

265. Eriksen, R. S., Mitarai, N. & Sneppen, K. Sustainability of spatially distributed bacteria-phage systems. *Sci. Rep.* **10**, 3154 (2020).

266. Lourenço, M. *et al.* The spatial heterogeneity of the gut limits predation and fosters coexistence of bacteria and bacteriophages. *Cell Host & Microbe* **28**, 390-401.e5 (2020).

267. Schrag, S. J. & Mittler, J. E. Host-parasite coexistence: the role of spatial refuges in stabilizing bacteria-phage interactions. *Am. Nat.* **148**, 348–377 (1996).

268. Nakao, R., Ramstedt, M., Wai, S. N. & Uhlin, B. E. Enhanced biofilm formation by *Escherichia coli* LPS mutants defective in Hep biosynthesis. *PLoS One* **7**, e51241 (2012).

269. Simmons, E. L. *et al.* Biofilm structure promotes coexistence of phage-resistant and phage-susceptible bacteria. *mSystems* **5**, e00877-19 (2020).

270. Bull, J. J., Vegge, C. S., Schmerer, M., Chaudhry, W. N. & Levin, B. R. Phenotypic resistance and the dynamics of bacterial escape from phage control. *PLoS One* **9**, e94690 (2014).

271. Abedon, S. T. Phage therapy dosing: the problem(s) with multiplicity of infection (MOI). *Bacteriophage* **6**, e1220348 (2016).

272. Saussereau, E. *et al.* Effectiveness of bacteriophages in the sputum of cystic fibrosis patients. *Clin. Microbiol. Infect.* **20**, O983-990 (2014).

273. Yehl, K. *et al.* Engineering phage host-range and suppressing bacterial resistance through phage tail fiber mutagenesis. *Cell* **179**, 459-469.e9 (2019).

274. Yang, Y. *et al.* Development of a bacteriophage cocktail to constrain the emergence of phage-resistant *Pseudomonas aeruginosa. Front. Microbiol.* **11**, 327 (2020).

275. Nale, J. Y. *et al.* An optimized bacteriophage cocktail can effectively control *Salmonella in vitro* and in *Galleria mellonella. Front. Microbiol.* **11** (2021).

276. Burmeister, A. R. *et al.* Pleiotropy complicates a trade-off between phage resistance and antibiotic resistance. *Proc. Natl. Acad. Sci. U.S.A* **117**, 11207–11216 (2020).

277. Borin, J. M., Avrani, S., Barrick, J. E., Petrie, K. L. & Meyer, J. R. Coevolutionary phage training leads to greater bacterial suppression and delays the evolution of phage resistance. *Proc. Natl. Acad. Sci U.S.A* **118**, e2104592118 (2021).

278. Hall, A. R., De Vos, D., Friman, V.-P., Pirnay, J.-P. & Buckling, A. Effects of sequential and simultaneous applications of bacteriophages on populations of *Pseudomonas aeruginosa in vitro* and in wax moth larvae. *Appl. Environ. Microbiol.* **78**, 5646–5652 (2012).

279. Blount, Z. D., Borland, C. Z. & Lenski, R. E. Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli. Proc. Natl. Acad. Sci. U.S.A* **105**, 7899–7906 (2008).

280. Koskella, B. & Brockhurst, M. A. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol. Rev.* **38**, 916–931 (2014).

281. Buckling, A. & Rainey, P. B. . Antagonistic coevolution between a bacterium and a bacteriophage. *Proc. R. Soc. B: Biol. Sci.* **269**, 931–936 (2002).

282. Gomez, P. & Buckling, A. Bacteria-phage antagonistic coevolution in soil. *Science* **332**, 106–109 (2011).

283. Hall, A. R., Scanlan, P. D. & Buckling, A. Bacteria-phage coevolution and the emergence of generalist pathogens. *Am. Nat.* **177**, 44–53 (2011).

284. Hall, A. R., Scanlan, P. D., Morgan, A. D. & Buckling, A. Host-parasite coevolutionary arms races give way to fluctuating selection: bacteria-phage coevolutionary dynamics. *Ecol. Lett.* **14**, 635–642 (2011).

285. Betts, A., Kaltz, O. & Hochberg, M. E. Contrasted coevolutionary dynamics between a bacterial pathogen and its bacteriophages. *Proc. Natl. Acad. Sci. U.S.A* **111**, 11109–11114 (2014).

286. Garland, T., & Rose, M. R. *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments.* (University of California press, 2009).

287. Sackman, A. M. *et al.* Mutation-driven parallel evolution during viral adaptation. *Mol. Biol. Evol.* **34**, 3243–3253 (2017).

288. Escalera-Zamudio, M. *et al.* Parallel evolution in the emergence of highly pathogenic avian influenza A viruses. *Nat. Commun.* **11**, 5511 (2020).

289. Lang, G. I. & Desai, M. M. The spectrum of adaptive mutations in experimental evolution. *Genomics* **104**, 412–416 (2014).

290. Munita, J. M. & Arias, C. A. Mechanisms of antibiotic resistance. *Microbiol. Spectr.* **4** (2016).

291. Bull, J. J. & Molineux, I. J. Predicting evolution from genomics: experimental evolution of bacteriophage T7. *Heredity* **100**, 453–463 (2008).

292. Bailey, S. F., Blanquart, F., Bataillon, T. & Kassen, R. What drives parallel evolution?: How population size and mutational variation contribute to repeated evolution. *BioEssays* **39**, e201600176 (2017).

293. Lässig, M., Mustonen, V. & Walczak, A. M. Predicting evolution. *Nat. Ecol. Evol.* **1**, 1–9 (2017).

294. Lenski, R. E. & Levin, B. R. Constraints on the coevolution of bacteria and virulent phage: a model, some experiments, and predictions for natural communities. *Am. Nat.* **125**, 585–602 (1985).

295. Bohannan, B. J. M. & Lenski, R. E. Linking genetic change to community evolution: insights from studies of bacteria and bacteriophage. *Ecol. Lett.* **3**, 362–377 (2000).

296. Poullain, V., Gandon, S., Brockhurst, M. A., Buckling, A. & Hochberg, M. E. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. *Evolution* **62**, 1–11 (2008).

297. Spanakis, E. & Horne, M. T. Co-adaptation of *Escherichia coli* and coliphage lambda vir in continuous culture. *J. Gen. Microbiol.* **133**, 353–360 (1987).

298. Los, M. Minimization and prevention of phage infections in bioprocesses. *Methods Mol. Biol.* **834**, 305–315 (2012).

299. Zou, X. *et al.* Systematic strategies for developing phage resistant *Escherichia coli* strains. *Nat. Commun.* **13**, 4491 (2022).

300. Los, M. Strategies of phage contamination prevention in industry. *Open J. Bacteriol.* **4**, 020–023 (2020).

301. Gandon, S. & Michalakis, Y. Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *J. Evol. Biol.* **15**, 451–462 (2002).

302. Sun, C. L. *et al.* Phage mutations in response to CRISPR diversification in a bacterial population. *Environ. Microbiol.* **15**, 463–470 (2013).

303. Azam, A. H. & Tanji, Y. Bacteriophage-host arm race: an update on the mechanism of phage resistance in bacteria and revenge of the phage with the perspective for phage therapy. *Appl. Microbiol. Biotechnol.* **103**, 2121–2131 (2019).

304. Samson, J. E., Magadán, A. H., Sabri, M. & Moineau, S. Revenge of the phages: defeating bacterial defences. *Nat. Rev. Microbiol.* **11**, 675–687 (2013).

305. Scanlan, P. D., Buckling, A. & Hall, A. R. Experimental evolution and bacterial resistance: (co)evolutionary costs and trade-offs as opportunities in phage therapy research. *Bacteriophage* **5**, e1050153 (2015).

306. Pagnout, C. *et al.* Pleiotropic effects of *rfa*-gene mutations on *Escherichia coli* envelope properties. *Sci. Rep.* **9**, 9696 (2019).

307. Abdel-Rhman, S. H. Role of *Pseudomonas aeruginosa* lipopolysaccharides in modulation of biofilm and virulence factors of Enterobacteriaceae. *Ann Microbiol* **69**, 299–305 (2019).

308. Bertozzi Silva, J., Storms, Z. & Sauvageau, D. Host receptors for bacteriophage adsorption. *FEMS Microbiol. Lett.* **363**, fnw002 (2016).

309. Bohannan, B. J. M. & Lenski, R. E. Effect of resource enrichment on a chemostat community of bacteria and bacteriophage. *Ecology* **78**, 2303–2315 (1997).

310. Ishii, M., Matsumoto, Y. & Sekimizu, K. Estimation of lactic acid bacterial cell number by DNA quantification. *Drug Discov. Ther.* **12**, 88–91 (2018).

311. Solera, R., Romero, L. I. & Sales, D. Measurement of microbial numbers and biomass contained in thermophilic anaerobic reactors. *Water Environ Res.* **73**, 684–690 (2001).

312. Bratbak, G. Bacterial biovolume and biomass estimations. *Appl. Environ. Microbiol.* **49**, 1488–1493 (1985).

313. Fredborg, M. *et al.* Real-time optical antimicrobial susceptibility testing. *J. Clin. Microbiol.* **51**, 2047–2053 (2013).

314. Oyler, B. L. *et al.* Top-down tandem mass spectrometric analysis of a chemically modified rough-type lipopolysaccharide vaccine candidate. *J. Am. Soc. Mass Spectrom.* **29**, 1221–1229 (2018).

315. Schwartzman, J. A. *et al.* Bacterial growth in multicellular aggregates leads to the emergence of complex life cycles. *Curr. Biol.* **32**, 3059-3069.e7 (2022).

316. Andam, C. P. Clonal yet Different: Understanding the causes of genomic heterogeneity in microbial species and impacts on public health. *mSystems* **4**, e00097-19 (2019).

317. Blundell, J. R. *et al.* The dynamics of adaptive genetic diversity during the early stages of clonal evolution. *Nat. Ecol. Evol.* **3**, 293-301 (2018)

318. Vázquez-García, I. *et al.* Clonal Heterogeneity influences the fate of new adaptive mutations. *Cell Rep.* **21**, 732–744 (2017).

319. Paciello, I. *et al.* Intracellular Shigella remodels its LPS to dampen the innate immune recognition and evade inflammasome activation. *Proc. Natl. Acad. Sci. U.S.A* **110**, 4345–4354 (2013).

320. Westra, E. R. *et al.* Parasite exposure drives selective evolution of constitutive versus inducible defense. *Curr. Biol.* **25**, 1043–1049 (2015).

321. Colavecchio, A., Cadieux, B., Lo, A. & Goodridge, L. D. Bacteriophages contribute to the spread of antibiotic resistance genes among foodborne pathogens of the *Enterobacteriaceae* family – A Review. *Front. Microbiol.* **8** (2017).

322. Ramirez, J. *et al.* Antibiotics as major disruptors of gut microbiota. *Front. Cell. Infect. Microbiol.*
10 (2020).

Appendices

```
#!/bin/bash
#
#   submit by sbatch full_pipe_insert_size.sh
#
#   specify the job name
#SBATCH --job-name=Trim
#   how many cpus are requested
#SBATCH --ntasks=4
#   run on one node, important if you have more than 1 ntasks
#SBATCH --nodes=1
#   maximum walltime, here 10min
#SBATCH --time=30:00:00
#   maximum requested memory
#SBATCH --mem=20G
#   write std out and std error to these files
#SBATCH --error=insert.%J.err
#SBATCH --output=insert.%J.out
#   send a mail for job start, end, fail, etc.
#   which partition?
#   there are global, testing, highmem, standard, fast
#SBATCH --partition=standard



#   add your code here:
#############################################################
#############################################################
#############################################################
#####
## Developed by Loukas Theodosiou and Jordan Romeyer Dherbey:
last update 11.01.2022
##
```

```
## USE THIS PIPELINE TO:
## 1. Check quality of the sequences
## 2. Trim them
## 3. Check the quality of the trimmed files
#############################################################
#############################################################
#############################################################
#####

#path from working directory to directory that you wish the
job to be done
path= # add here the path to your files

mkdir ${path}fastQC_1

##### here you test the quality of the raw data before
trimming

fastqc
fastQC_raw=${path}fastQC_1/

#FastQC

cd ${path}
for each in *fastq.gz
do
${fastqc}fastqc ${each} ${fastQC_raw}{each%fastq}fQC
done

# Here you trim your data

# Create necessary directories
mkdir ${path}Trimmomatic
mkdir ${path}Trimmomatic/fastQC_2
```

```
mkdir ${path}Trimmomatic/unpaired


# Trimmomatic

trimmo=/data/biosoftware/Trimmomatic/Trimmomatic-0.38/
out_dir=${path}Trimmomatic/


#Trim for adaptors, min quality = 30, min length = 45
#raw read directory
cd ${path}
for each in *_R1_001.fastq.gz
do
echo ${each}
java -jar ${trimmo}trimmomatic-0.38.jar PE -threads 5 \
                    $each
${each%R1_001.fastq.gz}R2_001.fastq.gz \

    ${out_dir}${each%_R1_001.fastq.gz}P_1.fastq
${out_dir}${each%_R1_001.fastq.gz}U_1.fastq \

    ${out_dir}${each%_R1_001.fastq.gz}P_2.fastq
${out_dir}${each%_R1_001.fastq.gz}U_2.fastq \
                    ILLUMINACLIP:${trimmo}adapters/TruSeq3-PE-
2.fa:2:30:10 SLIDINGWINDOW:3:30 MINLEN:45
done

## Trim for adapters, quality, and length

fastqc
fastQC_trim=${path}Trimmomatic/fastQC_2/
unpaired=${path}Trimmomatic/unpaired/
```

```
#FastQC


cd ${out_dir}
for each in *fastq
do
${fastqc}fastqc ${each} ${fastQC_trim}{each%fastq}fQC
done


mv *U_1.fastq ${unpaired}
mv *U_2.fastq ${unpaired}
```

Appendix. A1. Script used to control the quality of the sequencing output using *FastQC* version 0.11.8 and trim the reads using *Trimmomatic*.

```
#!/bin/bash
#
#submit by sbatch 092021_WGS_colony_Breseq.sh
#  specify the job name
#SBATCH --job-name=Breseq
#  how many cpus are requested
#SBATCH --ntasks=4
#  run on one node, important if you have more than 1 ntasks
#SBATCH --nodes=1
#  maximum walltime, here 10min
#SBATCH --time=100:00:00
#  maximum requested memory
#SBATCH --mem=120G
#  write std out and std error to these files
#SBATCH --error=insert.%J.err
#SBATCH --output=insert.%J.out
#  send a mail for job start, end, fail, etc.
#  which partition?
#  there are global, testing, highmem, standard, fast
#SBATCH --partition=global


## Developed by Loukas Theodosiou and Jordan Romeyer Dherbey:
last update 11.01.2022
##


module load R/3.4.3
module load python/3.7.1
#path from working directory to directory that you wish the job
to be done


reference_file_1 = # add here the path to your reference file
fastq_dir= # add here the path to your /Trimmomatic/ files


# make the breseq_output directory
```

```
mkdir ${fastq_dir}breseq_output
out_dir=${fastq_dir}breseq_output/

cd ${fastq_dir}
for each in *_1.fastq
do
   breseq   -j   8   -p   -r   ${reference_file_1}   ${each}
${each%_1.fastq}_2.fastq \
   -o ${out_dir}${each}

Done
```

Appendix. A2. Script used to assemble and analyse the trimmed reads using the *breseq* pipeline version 0.33.2[101–103].