Delft University of Technology

Document Version
Final published version

Licence
CC BY-NC

Citation (APA)
van den Berg, D. F. (2026). *Phage immunity and evasion*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:d1bec46a-a979-40bd-80db-b03f5baa4e3e

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
In case the licence states “Dutch Copyright Act (Article 25fa)”, this publication was made available Green Open Access via the TU Delft Institutional Repository pursuant to Dutch Copyright Act (Article 25fa, the Taverne amendment). This provision does not affect copyright ownership.
Unless copyright is transferred by contract or statute, it remains with the copyright holder.

Sharing and reuse
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.



PHAGE IMMUNITY & EVASION

D.F. van den Berg

Phage Immunity and Evasion

Phage Immunity and Evasion

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr. ir. H. Bijl,
chair of the Board for Doctorates,
to be defended publicly on
Thursday 5 February 2025 at 17:30 o'clock

by

Daan Frits VAN DEN BERG

Master of Science in Biomedical Sciences,
University of Amsterdam, the Netherlands,
born in Haarlem, the Netherlands.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	chairperson
Prof. dr. ir. S.J.J. Brouns	Delft University of Technology, *promotor*
Dr. D.C. Swarts	Wageningen University & Research, *copromotor*

Independent members:

Prof. dr. M.J.T. Reinders	Delft University of Technology
Prof. dr. D. Bikard	Institute Pasteur, France
Prof. dr. G.C. Atkinson	Lund University, Sweden
Prof. dr. B.D. Rowland	Delft University of Technology / Netherlands Cancer Institute
Dr. L. Loeff	Leiden University Medical Center

Reserve member:

Prof. dr. C. Joo	Delft University of Technology







This work was financially supported by the European Research Council (ERC).

Keywords: Phage defense systems, phage defense evasion, *Pseudomonas aeruginosa*, bacteriophages.

Front: The cover depicts a phage-like David holding a glowing anti-defense protein, facing a resting Goliath, which represents its bacterial host. The night sky is filled with Greek constellations, symbolizing the Greek god names of phage defense systems. Design by the author.

Printed by: Gildeprint – Enschede.

Copyright © 2026 by D.F. van den Berg

ISBN 978-94-6496-525-4

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Table of Contents

	Summary	6
	Samenvatting	8
Chapter 1	Introduction and Overview.	11
Chapter 2	Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*.	20
Chapter 3	R2-type pyocins are associated with phage defense.	45
Chapter 4	Bacterial homologs of innate eukaryotic antiviral defenses with anti-phage activity.	51
Chapter 5	Discovery of phage defense systems through component modularity networks.	81
Chapter 6	Reduced prevalence of phage defense systems in *Pseudomonas aeruginosa* strains from cystic fibrosis patients.	103
Chapter 7	Phages encode broad and specific counter-defense genes to overcome defense systems.	111
Chapter 8	Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations.	131
Chapter 9	Phage tRNAs: Decoding the Enigma.	137
Chapter 10	Summary and Discussion.	149
	Epilogue: Consciousness	161
	References	162
	Acknowledgements	186
	Curriculum Vitae	190
	About the author	191
	List of publications	192

Summary

Bacteria are under constant predation by their viruses, also known as bacteriophages or phages. To protect themselves against phages, bacteria evolved a multitude of phage defense mechanisms. Some of these phage defense mechanisms, such as CRISPR-Cas systems, have been known for more than a decade. More recently, it has become evident that the repertoire of phage defense systems of bacteria is far more diverse and numerous than previously thought. This finding not only reshaped our understanding of phage defense systems but also raised the question of how phages can still successfully infect these complex immune responses. In this dissertation, we investigate this complex interplay between phages and their hosts, specifically looking at the importance of phage defense systems in phage resistance, approaches for discovering additional phage defenses, and how phages evade these phage defense systems.

In the **first chapter**, we provide a brief overview of the current understanding of bacterial immune systems and how phages evade these systems. We begin by describing how phages infect bacteria, and how bacteria respond to this infection. We then provide an overview of the known strategies that phages use to neutralize the host response, which in response prompted bacteria to evolve new strategies. Over time, creating a complex interplay between phage defense systems and evasion strategies of the phage.

In the **second chapter,** we show that bacterial species *Pseudomonas aeruginosa* has a large and numerous repertoires of phage defense systems. We demonstrate that the number of phage defense systems per strain correlates with the broadness of its resistance against a wide range of phages.

In the **third chapter,** we aimed to uncover previously unknown phage defense systems by searching for gene clusters that are associated with a higher resistance to phage infection. To achieve this, we assessed the infection ability of our *Pseudomonas* phages across our *P. aeruginosa* collection and conducted a genome-wide-association study. We identified one gene-cluster to be significantly associated with an increased phage resistance, corresponding to a R2-type pyocin. These pyocins are remnants of ancient phages that have been domesticated by bacteria to lyse nearby cells. How these R2-type pyocin may convey phage defense remains unknown.

In the **fourth chapter,** we look for homologs of eukaryotic viral defense systems in bacteria to uncover previously unknown phage defenses. We demonstrate that these homologs provide protection against phages using *P. aeruginosa* as a model organism. These bacterial phage defense systems resemble eukaryotic viral defense mechanisms in several ways, including preventing viral attachment, R-loop-acting enzymes, the inflammasome, the ubiquitin pathway, and pathogen recognition signalling.

In the **fifth chapter,** we search for previously unknown phage defense systems by capitalizing on the observation that phage defense systems often exhibit high degrees of modularity, with sensing, signal transmission, and effector enzymes frequently being exchanged among phage defense gene clusters. By searching for gene clusters with defense-associated genes or functional domains, we uncovered several new phage defense systems.

In the **sixth chapter,** we observe that the prevalence of phage defense systems of *P. aeruginosa* strains from cystic fibrosis lung patients is reduced compared to *P. aeruginosa* strains from patients with other lung conditions, suggesting that cystic fibrosis-associated strains are more susceptible to phages. This observation provides a promising perspective for treating *P. aeruginosa* infections with phages.

In the **seventh chapter,** we set out to identify additional evasion strategies that phages use to evade phage defense systems. In this chapter, we focussed on phage genes that were located within the highly variable genomic regions of *Pbunaviruses*, a *Pseudomonas* phage family, and testing their ability to inhibit bacterial phage defense systems. Using this approach, we discovered several genes that were able to

prevent the host immune response from acting, including those both broad and specific inhibitors. We showed that these genes are prevalent among a large variety of phage taxa.

In the **eight chapter,** we provide a hypothesis that offers a new perspective on a long-standing mystery: why phages encode their own tRNAs. An observation that has intrigued the phage field since its discovery in the early 1950s. We suggest that these phage tRNAs serve as an evasion strategy against phage defense systems that deplete host tRNAs, which would otherwise inhibit the ability of the phage to translate its genes and prevent phage propagation. Supporting our hypothesis, we observe that phage tRNAs have mutations that render these insensitive to tRNA targeting phage defenses.

In the **nineth chapter,** we review the current state of the phage tRNA field by highlighting their diverse roles in phage infection. We discuss their multifunctional roles for temperate, as well as the role of phage tRNAs for virulent phages, where they primarily serve to replenish the depleted tRNA pool of the host. Additionally, we highlight currently known phage defense systems that convey phage protection by depleting the host tRNA pool. We conclude the review by discussing the multiple layers of tRNA-targeting phage defenses, not all of which act by depleting tRNAs; some instead act on tRNA maturation steps and incorporation during translation.

In the **tenth chapter,** we present a general summary of the thesis and discuss the implications of these insights. Many questions remain, including: What are the biological laws that seem to govern the composition of the phage defense repertoire? And how can such a relatively small entity overcome the defenses of a much larger host? We also discuss several conceptual considerations, such as: Are we studying phage defense systems in the appropriate biological context? The discussion concludes by providing a perspective on the future of the field.

In conclusion, this dissertation investigates phage defense systems and evasion strategies. It provides insights into the cumulative role of *Pseudomonas* phage defense systems that facilitate a broad resistance to phages, it describes several methods to further uncover the bacterial immune system and provides new perspectives on how phages circumvent these defenses.

Samenvatting

Bacteriën worden constant geïnfecteerd door hun virussen, die ook wel bacteriofagen of fagen worden genoemd. Om zichzelf tegen deze fagen te beschermen, hebben bacteriën een aantal verdedigingsmechanismen bij zich. Sommige van deze mechanismen, zoals CRISPR-Cas-systemen, zijn al langer bekend, maar recent onderzoek heeft aangetoond dat het afweersysteem van bacteriën veel diverser en talrijker is dan eerder gedacht. Dit heeft niet alleen ons begrip van deze systemen veranderd, maar roept ook de vraag op hoe fagen desondanks deze hoeveelheid verdedigingssystemen nog steeds succesvol hun gastheer kunnen infecteren. In dit proefschrift onderzoeken we deze wisselwerking om het beter te begrijpen, waarbij wij specifiek hebben gekeken of een groter aantal verdedigingsmechanismen per bacterie ook leidt tot meer resistentie tegen een brede collectie van fagen. Daarnaast hebben we ook nieuwe methoden ontwikkeld om zowel meer verdedigingsmechanismen van de bacterie als mechanismen die fagen gebruiken om de immuunreactie van de gastheer te omzeilen, in kaart te brengen.

In het **eerste hoofdstuk** geven we een beknopte introductie van de huidige kennis over bacteriële immuunsystemen en de manieren waarop fagen deze systemen omzeilen. We beginnen met een beschrijving van hoe fagen bacteriën infecteren en hoe bacteriën deze infectie proberen te stoppen. Vervolgens geven we een overzicht van strategieën die fagen gebruiken om de immuunreactie van de gastheer te omzeilen, waarop bacteriën dan weer nieuwe strategieën ontwikkelen. Over tijd heeft dit geleid tot een complexe wisselwerking tussen verdedigingssystemen en de strategieën van de fagen om deze systemen te omzeilen.

In het **tweede hoofdstuk** onderzoeken we het immuunsysteem van de bacteriesoort *P. aeruginosa*. Hieruit blijkt dat deze bacteriesoort over een groot en uitgebreid repertoire aan verdedigingssystemen beschikt. Daarna tonen we aan dat het aantal verdedigingssystemen dat de gastheer bij zich draagt correleert met hoe breed zijn weerstand is tegen onze collectie van verschillende fagen.

In het **derde hoofdstuk** zijn we op zoek gegaan naar nieuwe verdedigingsmechanismen door te zoeken naar genen die geassocieerd zijn met een hogere weerstand tegen fagen. Hiervoor hebben we al onze *Pseudomonas* fagen getest op onze *P. aeruginosa* collectie, waar we daarna een correlatieanalyse op hebben uitgevoerd op gen-niveau. Één gencluster bleek significant geassocieerd te zijn met een verhoogde weerstand tegen fagen. Dit gencluster komt overeen met een R2-type pyocine. Deze pyocinen zijn verwante aan fagen van vroeger die door bacteriën zijn "getemd" om nabijgelegen cellen te lyseren tijdens bacteriële conflicten. Hoe deze R2-type pyocinen precies bijdragen aan faag afweer is nog onbekend.

In het **vierde hoofdstuk** hebben we naar homologen van eukaryotische virale afweersystemen gezocht in bacteriën om nieuwe verdedigingsmechanismen te ontdekken. We tonen aan dat deze bacteriële homologen bescherming bieden tegen fagen. Deze bacteriële verdedigingsmechanismen vertonen op meerdere manieren overeenkomsten met hun eukaryotische virale afweersystemen verwanten, waaronder het voorkomen van het binden van het virus aan de gastheer, enzymen die op R-loops werken, componenten van het inflammasoom, het ubiquitineringscascade en pathogeen herkennings-receptor signalering.

In het **vijfde hoofdstuk** zoeken we ook naar nieuwe verdedigingsmechanismen tegen fagen door gebruik te maken van de observatie dat verdedigingsmechanismen tegen fagen vaak een hoge mate van modulariteit vertonen, waarbij sensor-, signaaloverdrachts- en effector-enzymen vaak uitgewisseld worden tussen verdedigingsmechanismen. Hierbij hebben we ons specifiek gericht op uitgewisselde componenten, zoals genen of functionele eiwitdomeinen, die voorkomen in genetische contexten die nog niet eerder waren onderzocht, bijvoorbeeld met andere flankende genen. Met deze methode hebben we verschillende nieuwe verdedigingsmechanismen ontdekt.

In het **zesde hoofdstuk** onderzochten we of er verschillen zijn tussen welke en hoeveel verdedigingssystemen gedetecteerd konden worden in *P. aeruginosa* stammen afkomstig uit de longen

van patiënten met taaislijmziekte ten opzichte van andere longaandoeningen. Bij dit onderzoek vonden we dat de stammen die afkomstig waren uit de longen van patiënten met taaislijmziekte een minder divers en talrijk verdedigingsrepertoire hadden tegen fagen dan de andere stammen. Deze bevinding suggereert dat *P. aeruginosa* stammen die een infectie veroorzaken in de longen van taaislijm patiënten waarschijnlijk gevoeliger zijn voor fagen. Dit biedt een veelbelovend perspectief voor de behandeling van *P. aeruginosa* infecties bij deze patiënten.

In het **zevende hoofdstuk** richten we ons op het identificeren van voorheen onbekende strategieën die fagen gebruiken om verdedigingssystemen te omzeilen. Hiervoor onderzochten we het immuunsysteem remmende vermogen van genen die zich bevinden in de sterk variabele delen van het genoom van *Pbunavirussen*, een familie van *Pseudomonas* fagen. Met deze methode hebben we verschillende genen ontdekt die de immuunreactie van de gastheer konden hinderen, waaronder zowel brede als specifieke remmers. Deze genen zijn terug te vinden in een groot aantal fagentaxa.

In het **achtste hoofdstuk** presenteren we een hypothese die een nieuw perspectief biedt op een lang bestaand raadsel: waarom fagen hun eigen tRNA's met zich meebrengen. Een observatie die het fagen veld sinds de ontdekking in het begin van de jaren 1950 heeft geboeid. Wij stellen dat deze faag tRNA's dienen als een omzeilingsstrategie tegen de verdedigingsmechanismen die de gastheer gebruikt om tRNA's van de gastheer afbreken in het geval van een infectie van een faag. Dit vermindert het vermogen van de faag om zijn genen om te zetten naar eiwitten en het vermogen van de faag om zich te vermenigvuldigen. Ter ondersteuning van onze hypothese laten we zien dat de tRNA's van fagen mutaties bevatten waardoor deze ongevoelig zijn voor verdedigingsmechanismen die tRNA's afbreken.

In het **negende hoofdstuk** geven we een overzicht van de huidige kennis op het gebied van faag tRNA's, waarbij we hun diverse rollen tijdens faag infectie uitlichten. We bespreken hun multifunctionele functies bij "temperate" fagen, dat zijn fagen die zichzelf integreren in het genoom, evenals de rol van faag tRNA's bij virulente fagen, waarbij ze voornamelijk dienen om de afgebroken tRNA-voorraad van de gastheer aan te vullen. Daarnaast belichten we bekende verdedigingsmechanismen die bescherming bieden door de tRNA-voorraad van de gastheer af te breken. We sluiten de review af met het bespreken van de verschillende lagen van tRNA gefocuste verdedigingsmechanismen, waarvan niet alle werken door tRNA's af te breken, maar ook andere belichten die ingrijpen op verschillende stappen van tRNA-maturatie en de incorporatie van tRNA's tijdens de translatie.

In het **tiende hoofdstuk** geven we een algemeen overzicht van het proefschrift en bespreken we de betekenis van de bevindingen. Verder bespreken we ook openstaande vragen, zoals: Welke biologische factoren de samenstelling van het verdedigingsrepertoire bepalen? En hoe een relatief kleine faag de verdediging van een veel grotere gastheer kan overwinnen? Daarnaast bespreken we enkele conceptuele vraagtekens, zoals: Bestuderen we deze verdedigingssystemen wel in de juiste biologische context? Ten slot, sluiten we de discussie af met een vooruitblik op de toekomst van het vakgebied.

Samenvattend onderzoekt dit proefschrift verdedigingsmechanismen van de gastheer ten opzichte van de fagen en de omzeilingsstrategieën van de fagen zelf. Het biedt nieuwe inzichten in de cumulatieve rol van verdedigingsmechanismen in *P. aeruginosa*-stammen die een brede weerstand tegen fagen bieden. Daarnaast beschrijft dit proefschrift verschillende methoden om het bacteriële immuunsysteem verder te ontrafelen en biedt het nieuwe perspectieven op hoe fagen deze verdedigingsmechanismen omzeilen.

Chapter 1



Introduction and Overview

1

Bacteria and their viruses

Bacteria were among the earliest organisms to roam the world (*1*). Over time, bacteria have evolved and adapted to an incredible diverse range of environments, from the ocean to even residing inside eukaryotic cells, making them the most abundant and diverse life form on earth (*1, 2*). Despite their success in adapting to different environments, up to 40% of all bacteria die each day to viral infections caused by bacteriophages (*3*). Bacteriophages, also known as phages, are viruses that infect bacteria to reproduce themselves (**Figure 1**). Phages are highly diversified, ranging from single-stranded(+) RNA icosahedral-shaped phages that only encode a few genes, to large double-stranded DNA "Jumbo" phages that encode more than 250 genes and form a sophisticated cellular structures made up of several specialized components (*4, 5*). Regardless of the type of genomic content or the structural complexity of the phage, the first step of the infection begins with the recognition and attachment to the outer membrane of the bacterial host cell (**Figure 1**). To recognize the host, phages have evolved specialized proteins (tail fibers) that bind to host specific factors on the bacterial surface. Once the phage has found its host, it ejects its genomic content, which usually resides within the capsid, into the host cell through their phage tail (**Figure 1**).

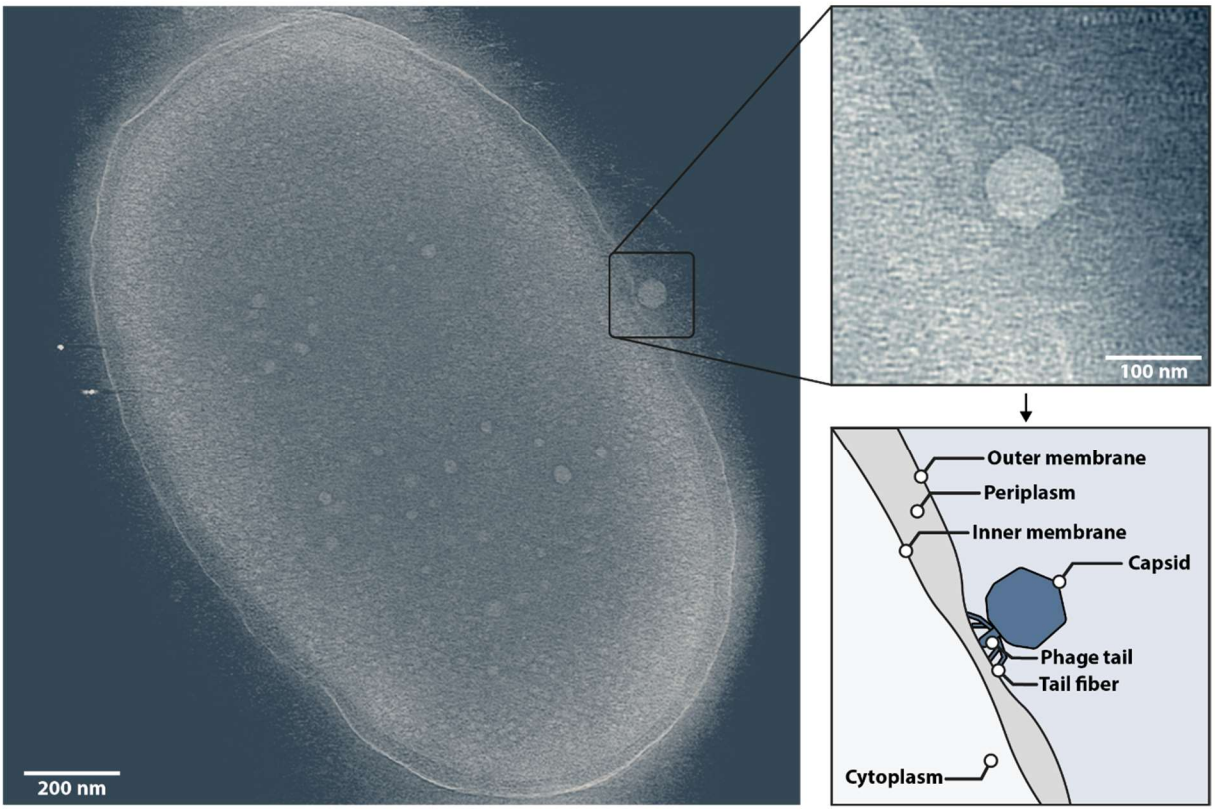

Figure 1. Phage adsorption to bacterial *Pseudomonas aeruginosa* strain PAO1. (**Left**) Shown is a false colored cryo-TEM image of the adsorption of *Pseudomonas* phage vB_PaeP_FBPa22, which has a podophage morphology, to its *Pseudomonas aeruginosa* strain PAO1 host. (**Top right**) Close-up view of the phage attachment to the host. (**Bottom right**) An illustration of the close-up view of the phage attachment, indicating the structures within the image including the host inner and outer membrane, as well as the periplasm and cytoplasm. In addition, the phage capsid, tail, and tail fibers are indicated. Scale bars: 200 nm (left) and 100 nm (top right). Cryo-TEM image was taken by dr. V.M.A. Ongenae from Claessen Lab and colorized by D.F. van den Berg.

After injection, the subsequent steps of the infection depend on the infection strategy of the phage, including lytic and lysogenic strategies (*6*) (**Figure 2**). Virulent (lytic) phages enter their host and immediately start replicating and producing phage components, which are assembled into new phage particles, followed by the lysis of the host cell to release the newly produced phages into the environment to infect their next host (*6*). While virulent phages start replicating immediately,

temperate (lysogenic) phages can integrate their genome, as a prophage, into the genome of the host instead. Here, the phage genome is passively replicated together with the host, until an environmental cue causes the excision of the phage genome from the host chromosome and the initiation of its lytic cycle (*6*). Besides the lytic and lysogenic phage lifestyles, several additional lifestyles have been discovered over the years including the chronic lifestyle (a temperate-like lifestyle in which the host is not lysed, but phage particles are still released), the pseudo lysogenic lifestyle (a temperate-like lifestyle in which the phage genome is not integrated into the host (*7*)), the helper-dependent lifestyle (undergone by satellite phage, which require a co-infecting phage to complete their infection cycle, for example by providing its phage particle (*8*)), and the plasmid-like lifestyle (undergone by phage plasmids, which typically exist as plasmids, but can spread as a plasmid through conjugation, or as a phage through cell lysis (*9*)).

Although not all the above-mentioned phage infection strategies result in the immediate death of the host, ultimately all reproducing phages redirect their host's resources to their own replication, imposing a metabolic burden and thereby a selective disadvantage to the host.

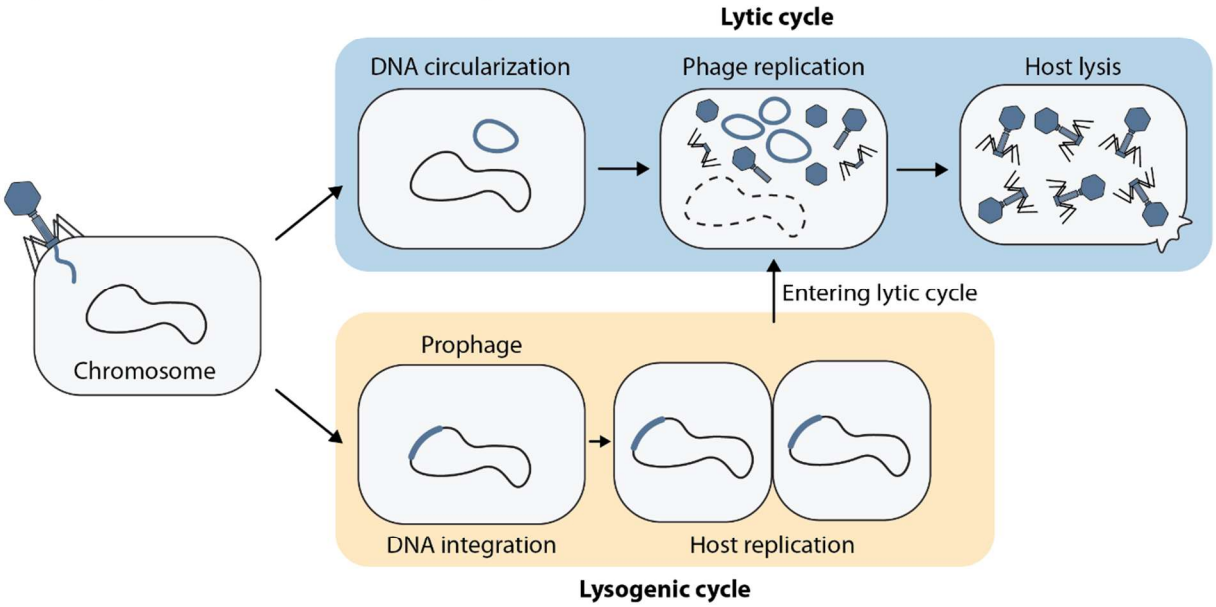

Figure 2. Phage infection strategies. The lytic (blue) and lysogenic (orange) life cycles, that phages employ during the infection of their host. Common to both cycles are the initial step of adsorbing to the host cell. In the lysogenic cycle, the phage genome resides in the host chromosome as a prophage, passively replicating together with the host, until an environmental cue causes the prophage to excise and start its lytic cycle. In the lytic cycle, phages replicate their genomes and produce new phage particles that are then released upon host cell lysis.

Phage immunity

To counteract phage infections, a variety of phage defense systems have evolved in bacteria (*10*). These bacterial defenses can be broadly categorized into extracellular and intracellular mechanisms.

Extracellular phage defenses

Extracellular phage defenses prevent phages from adsorbing to the host (*11*) (**Figure 3**). A wide variety of extracellular phage defenses are known, such as mutating, modifying or masking the molecules that phages use to recognize the host (*11*). Bacteria can also produce extracellular outer membrane vesicles that act as decoys or start producing a protective mucus-like barrier (*11-13*).

Intracellular phage defenses

When the phage overcomes the extracellular barrier of phage defense, additional intracellular phage defenses are in line to detect components of the infiltrated phage (*10*) (**Figure 3**). These phage defenses are specialized in detecting specific pathogen-associated molecular patterns (PAMPs), including phage-derived RNA, DNA, proteins, metabolites, as well as phage-induced cell envelope perturbations and metabolic changes (*14-28*). When a PAMP is detected, the phage defense system initiates its anti-viral response, which can be phage-directed or/and self-directed, depending on the phage defense system (*10*). These anti-viral responses range from the destruction of the incoming phage genome (phage-directed) to altruistic cell death to protect neighbour cells (self-directed) (*10, 29*). These defense systems often do not act in isolation, with some defenses slowing down the host metabolism to slow down phage replication, which then provides more time for other intracellular phage defenses to act (*10, 15*).

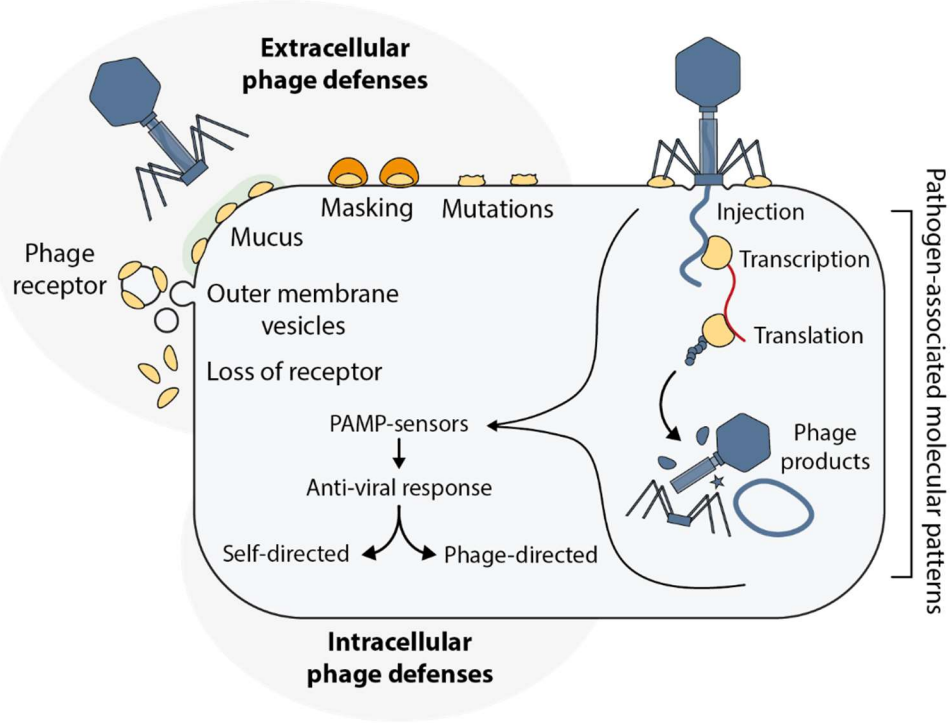

Figure 3. Phage defense. Shown are various extracellular and intracellular phage defenses of the bacterial host. The extracellular phage defenses include the loss of phage receptors, production of extra cellular vesicles, production of mucus, masking the receptor, and mutating the receptor in ways that result in the loss of recognition by the phage. The intracellular phage defenses recognize the infecting phage, based on the accompanied pathogen-associated molecular patterns (PAMPs). These PAMPs are detected by PAMP-sensors, upon which the anti-viral response is initiated. These responses can be self-directed or/and phage-directed, each providing distinct mechanisms of protection against phages.

Phage evasion

In response to the immunity of the host, phages have evolved numerous ways to evade the host response, most prominently through mutagenesis (*30*) (**Figure 4**). This mechanism is relevant for overcoming both extracellular and intracellular defenses (*11, 31*). For instance, surface modifications of the host, which affect host recognition by the phage, can be overcome by mutations in the phage tail fibers that enable the phage to identify and bind the modified surface receptor (*11, 31*). Similarly, mutagenesis of PAMPs can prevent the detection of the host immune response (*10, 32*). However, mutagenesis alone is not always sufficient to bypass the host immune response. In these instances, the phage encodes additional phage genes, such as extra tail fibers, mucus clearing enzymes, or genes that can tinker with the immune response of the host (*10, 33*) (**Figure 4a)**. For instance, phage-encoded decoy proteins revert the attention of phage defenses or prevent the alarm signalling of the immune response (*10*), and some phage-encoded proteins compensate for the effect of degraded host factors

by the activated phage defense systems. For example, by producing tRNAs to compensate for tRNAs that are depleted because of the immune response by the host (*34*) (**Figure 4b**). These outlined strategies are just a handful of examples that phages utilize to evade the host immune response. Many of these strategies have been characterized over the years with many more yet to be discovered (*10, 35*).

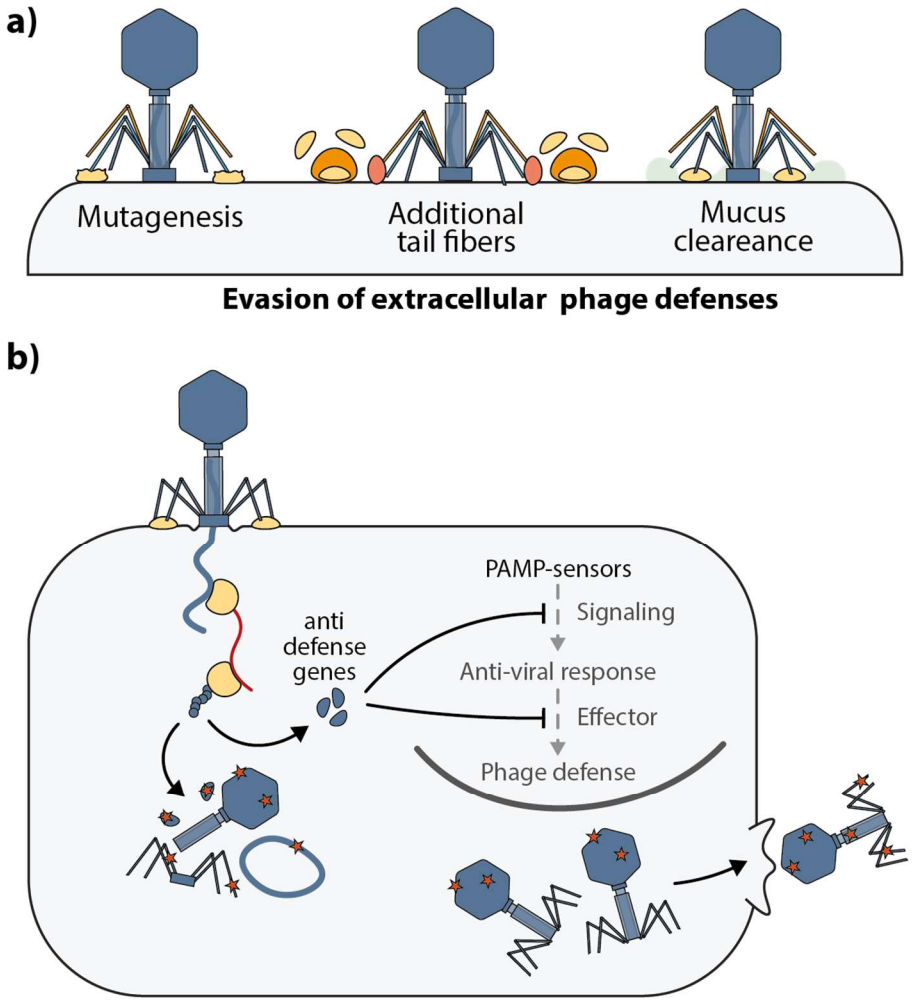

Figure 4. Phage evasion strategies. (a) Strategies that phages employ to circumvent the extracellular phage defenses of bacteria. These include the use of extra tail fibers to target additional phage receptors in cell surfaces, clearance of the mucus barrier to reach their receptors, and mutation of tail fibers to adapt to mutated receptors. **(b)** The strategies that the phage employs to circumvent the intracellular phage defenses, including anti-defense genes, that prevent the PAMP-sensors from signalling the rest of the immune response to convey the phage defense, and mutagenesis of the parts of the phage (red stars) that are detected by the PAMP-sensors.

Co-evolution of bacteria and phages

Bacteria deploy several countermeasures against these evasion strategies of the phage (*10*) (**Figure 5**). Foremost, the bacterial phage defense system mutates together with the phage-defense-evading mutations of the phage, to remain sensitive to the PAMP and survive the phage infection (*10, 36*). Moreover, some defense systems are activated upon detecting anti-defense phage genes that tinker with the immune response of the host, such as PARIS and Retron systems (*37-39*) (**Figure 5**). In response, phages also encode genes to overcome these additional defenses, that are then counteracted again by the host (*39*) (**Figure 5**). Over time, this co-evolution creates a complex ever-growing layered network of immune responses and immune evasion (**Figure 5**). As the phage evolves additional anti-

defense genes to overcome the complex layered immune response of the host, the phage is limited to the genetic capacity of the capsid, leading to a trade-off between essential and non-essential genes (*40*). Similarly, the host evolutionarily prefers to encode vital phage defenses, while getting rid of phage defenses that provide no benefit (*41, 42*). This is reflected in the genomic content of both the phage and the host, resulting in highly variable genetic regions that are critical for their survival in specific niche conditions (*41, 43*).

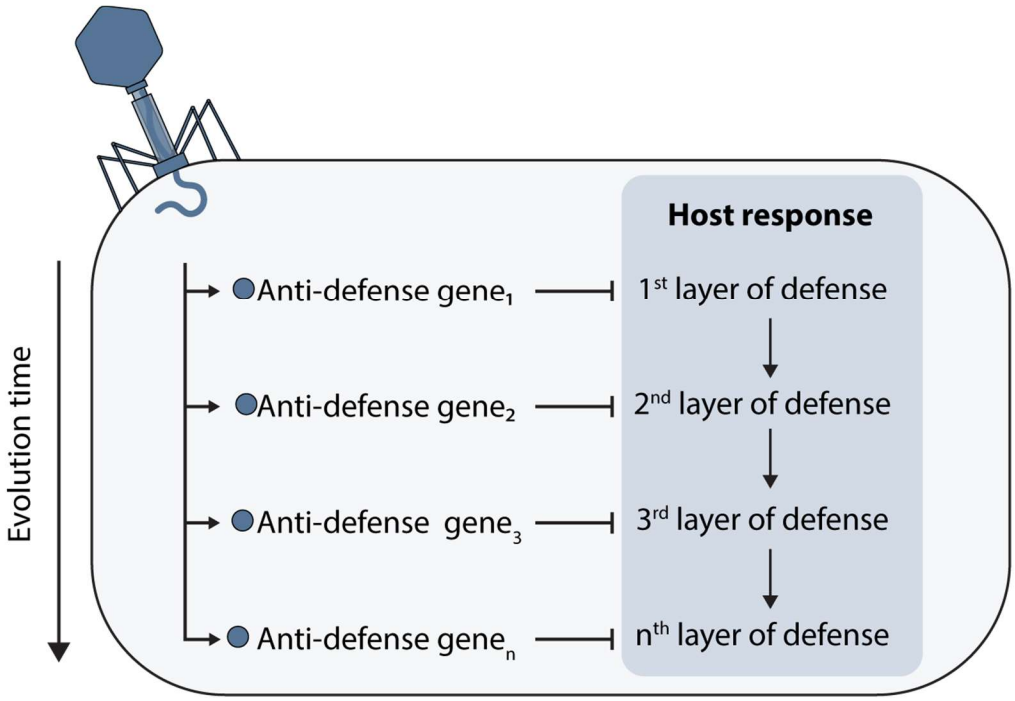

Figure 5: Co-evolution of phage and host. Shown is a depiction of the evolutionary conflict between phage and host, where every phage response is accompanied by a host response, evolving over time into a complex network of immune and anti-defense gene layers.

Scope and objectives of this thesis

The aim of this thesis is to gain a deeper understanding of phage immunity in bacteria and how phages evade this immunity. To achieve this, we chose to primarily focus on the phage-host dynamics of *P. aeruginosa*. This Gram-negative bacterial species can be found in a wide variety of environments, from crude oil to the ocean and soil (*44*). It is notorious for causing infections in burn wounds, immunocompromised patients, and the lungs of cystic fibrosis patients (*44*). These infections are often difficult to treat due to their antibiotic-resistant nature, with 15% of *P. aeruginosa* isolates being pan drug-resistant (*45*). To treat these antibiotic-resistant infections, phages could be used as an alternative to antibiotics (*46*) (**Figure 1**). However, it remains unclear if using phages is a long-lasting solution, since *P. aeruginosa* strains are also known to be highly adaptable and to have a large repertoire of phage defenses (*47, 48*). Paradoxically, *P. aeruginosa* also has a large and diverse phage population that can overcome these defenses (*48*). This unique combination makes *P. aeruginosa* an excellent model organism for gaining a deeper understanding of the interplay between phage and host, which can later be explored for the development of phage-based antibacterial therapies.

Overview of the dissertation

In this thesis, we investigated the impact of defense systems on the host-range of phages and found that encoding more phage defenses leads to a broader defense against phages (**chapter 2**). Using several bioinformatic approaches, we identified additional phage defenses in *P. aeruginosa* genomes by

searching for host genes that correlate with the loss of phage infectivity (**chapter 3**), by searching for homologs of known eukaryotic viral defenses (**chapter 4**), and by looking for components of known bacterial phage defenses in new genetic contexts (**chapter 5**). We also observed that the environment of *P. aeruginosa* can significantly impact the abundance of phage defenses. Specifically, *P. aeruginosa* encodes less phage defenses when present in the lungs of cystic fibrosis patients (**chapter 6**), likely because of the reduced air flow and thicker mucus that result in lower inflow of phages. We then investigated the mechanisms that phages deploy to circumvent phage defenses using different methodological approaches. First, we mined the highly variable gene regions of *Pseudomonas* phages to identify five new anti-defense genes (**chapter 7**). Then, we postulated that phages encode tRNAs to circumvent tRNA-targeting phage defenses, addressing the 50-year-old mystery of why phages encode their own tRNAs (**chapter 8**). This finding is placed in a broader context in a review on phage tRNAs (**chapter 9**). Finally, we discuss the knowledge gained during the research to put it in a broader perspective (**chapter 10**).

Chapter 2



Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*

Prokaryotes encode multiple distinct anti-phage defense systems in their genomes. However, the impact of carrying a multitude of defense systems on phage resistance remains unclear, especially in a clinical context. Using a collection of antibiotic-resistant clinical strains of *Pseudomonas aeruginosa* and a broad panel of phages, we demonstrate that defense systems contribute substantially to defining phage host range and that overall phage resistance scales with the number of defense systems in the bacterial genome. We show that many individual defense systems target specific phage genera and that defense systems with complementary phage specificities co-occur in *P. aeruginosa* genomes likely to provide benefits in phage-diverse environments. Overall, we show that phage-resistant phenotypes of *P. aeruginosa* with at least 19 phage defense systems exist in the populations of clinical, antibiotic-resistant *P. aeruginosa* strains.

A modified version of this chapter has been published as

Costa, A. R.*, van den Berg, D. F. *, Esser, J. Q. *, Muralidharan, A., van den Bossche, H., Bonilla, B. E., van der Steen, B.A., Haagsma, A. C., Fluit, A. C., Nobrega, F. L., Haas, P., & Brouns, S. J. J. (2024). Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*. *Science advances*, *10*(8), eadj0341.

Introduction

Bacteriophage predation imposes a strong evolutionary pressure on bacteria to evolve mechanisms to defend against phage infection (*49*). These defense mechanisms include modification of cell surface receptors (*50, 51*) and intracellular defenses (*11, 41, 52*) such as CRISPR-Cas (*53, 54*) and Restriction-Modification (RM) systems (*55, 56*).

More recently, dozens of previously unknown anti-phage immune systems have been discovered. In most instances, they were identified based on the observation that immune systems often cluster in defense islands (*57-62*). The presence and composition of these defense islands vary among individual strains (*41, 61, 63*), and strongly contribute to phage-host co-evolution in natural populations (*49*). The presence of multiple variable defense systems in bacterial genomes raises the important question what the impact is of all these immune systems on overall phage resistance of bacterial pathogens.

To address this question, we assembled a set of 32 clinical, antibiotic-resistant *Pseudomonas aeruginosa* strains and compiled a custom panel of 28 phages from 12 phylogenetic groups. We then analysed phage infectivity and adsorption of the strains across the panel. This revealed that intracellular phage defense mechanisms are an important determinant of the phage susceptibility of *P. aeruginosa*, and that strains rich in phage defense systems are inherently more resistant to phage infection. Five strains contained a large number (13-19) of anti-phage defense systems and displayed an extended phage-resistant phenotype, in addition to having an extended drug-resistant (XDR) phenotype. Our data further revealed that defense systems can be specific to certain phage families, and that the activity of these individual defense systems in model strains can often predict the resistance of clinical strains to the same phages. Additionally, we have found that some combinations of defense systems with complementary phage specificity often co-occur in *P. aeruginosa* genomes and may provide phage defense with broader phage specificity. Overall, our findings have implications for our understanding of phage defense and potentially for the development of phage-based antibacterial therapeutics, as antibiotic-resistant strains with an extended phage-resistant phenotypes are present in clinical settings.

Results

Defense systems are abundant and diverse in clinical *P. aeruginosa* strains

We selected a set of 32 antibiotic-resistant clinical strains covering the diversity of defense systems in the *P. aeruginosa* species. The *P. aeruginosa* genomes from the RefSeq database carry 71% (119/167) of the known defense system subtypes (**Figure 1a**, **Table S1**), which is in line with recent observations that *P. aeruginosa* has a diverse arsenal of anti-phage defense (*63*). The defense systems found in the RefSeq genomes resembled the defense arsenal in our clinical isolates (**Figure 1b,c, Table S2**), and the number of defense systems per genome ranged between 1 and 19 systems for both datasets (**Figure S1d**,**e**). Additionally, we assessed the phylogenetic distribution of our collection using a maximum likelihood analysis of the core genes using Parsnp (*64*), which revealed the distribution of the clinical isolates across the two main phylogroups, with 23 strains in the largest phylogeny group 1 and 9 strains in phylogeny group 2 (**Figure S1**). We found that the RefSeq genomes belonging to phylogeny group 2 carried slightly higher number of defense systems compared to those of phylogeny group 1 (median phylogenetic group 1 = 7, median phylogenetic group 2 = 9, $p = 0.0004$), but these differences are not observed in the strains present in our collection ($p = 0.612$) (**Figure S1**). Overall, we demonstrate that *P. aeruginosa* genomes encode multiple different anti-phage defense systems and establish that our collection of clinical isolates is phylogenetically diverse and covers the range of defense systems both in types and numbers per strain as observed in *P. aeruginosa*

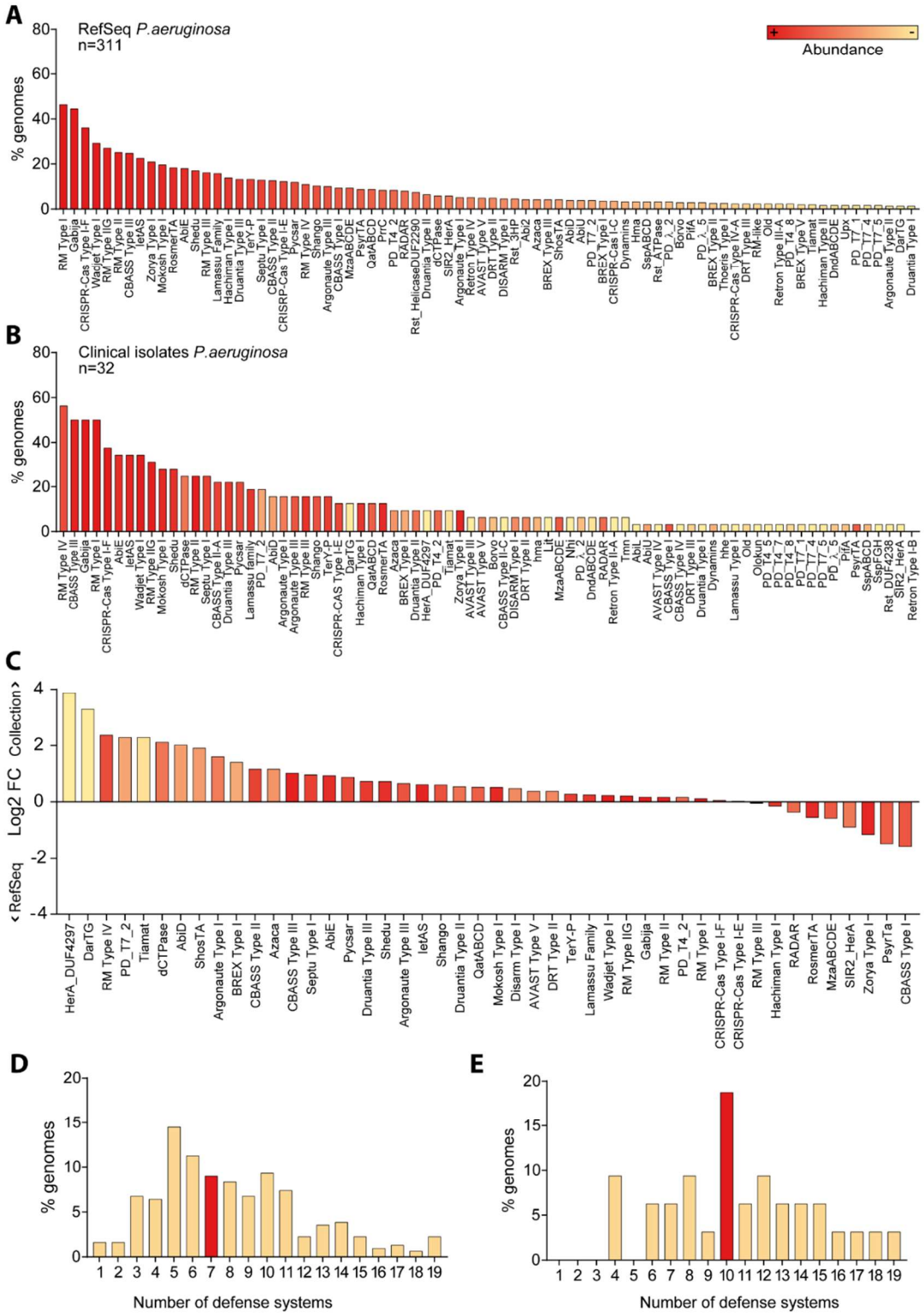

Figure 1. Defense systems are abundant and diverse in *P. aeruginosa* strains. (a) Diversity of defense systems found in the genomes of 311 *P. aeruginosa* strains from the RefSeq database, organized and colored in a gradient from most (left) to least (right) abundant. Only the most prevalent defense systems are shown (see Table S1 for the full list of defense systems). **(b)** Diversity of defense systems found in the genomes of 32 clinical isolates of *P. aeruginosa* from our collection. Systems are organized from most (left) to least (right) prevalent and colored according to the abundance in (a). All defense systems found in the clinical strains are shown. **(c)** Fold-change of defense system abundance in our collection of clinical strains relative to the RefSeq strains. **(d)** Number of defense systems per genome in *P. aeruginosa* strains from the RefSeq database. **(e)** Number of defense systems per genome in *P. aeruginosa* strains from our collection of clinical isolates. For (d) and (e) the median number of defense systems is shown in red.

Phage resistance correlates with the number of defense systems

To obtain a relevant panel of phages for these clinical isolates, we used a subset of 22 *P. aeruginosa* strains as hosts to enrich and isolate different phages from sewage water. We obtained a total of 27 phages (**Table S3**), consisting of *Caudoviricetes* (dsDNA tailed phages), including 13 podophages (5 *Autographiviridae*, 3 *Bruynoghevirus*, 1 *Schitoviridae*, 1 *Zobellviridae*, and 3 unassigned phages), 4 siphophages (1 *Samunavirus*, 1 *Casadabanvirus*, 1 *Mesyanzhinovviridae*, and 1 *Detrevirus*), and 10 myophages (9 *Pbunavirus* and 1 *Phikzvirus*, a Jumbo myophage related to nucleus-forming *Pseudomonas* phage phiKZ (*65*)). To broaden the diversity of our phage panel beyond dsDNA phages, we additionally included *Fiersviridae* PP7 (ssRNA phage). We then used vConTACT2 (*66*) to assess the taxonomic diversity of our phage panel, and found that it represents 9 out of the 16 phage clusters observed in *P. aeruginosa* phages overall, thus indicating a diverse representation of phages (**Figure S2**). A complete overview of the diversity within phage families and genera in our panel can be found in **Figure S3**.

To determine the effect of the defense systems on the susceptibility of the clinical strains to our panel of phages, we first assessed the ability of the phages to infect the strains, or to only adsorb to their cell surface without infecting. Out of a total of 924 phage-host combinations (28 phages times 33 hosts, including PAO1), 630 phage-host combinations did not result in infection (**Figure 2a**). We hypothesize the non-infected phenotype can occur in two ways: either the phage fails to adsorb to the cell surface (i.e., no receptor), or phage propagation is unsuccessful, possibly due to anti-phage defense. Based on adsorption assay data, we found that a large proportion out of the 630 non-infection cases showed adsorption. More precisely, 68% (429) of the phage-host combinations showing no infection exhibited adsorption when considering an adsorption threshold of 50% (**Figure 2a**), and 32% (201) when adopting a more conservative threshold of 90% (**Figure S4a**). These results indicate internal defense mechanisms could play a role in preventing infection in a substantial fraction of strains where no infection was observed. We ruled out the possibility of temperate phages (φPa19, φPa28, φPa42, and φPa47) integrating into the bacterial genome as the cause of instances where adsorption occurred but no infection, based on sequencing analysis of bacterial strains following phage infection, which showed no evidence of integrated phages (sequencing data available on the associated GitHub).

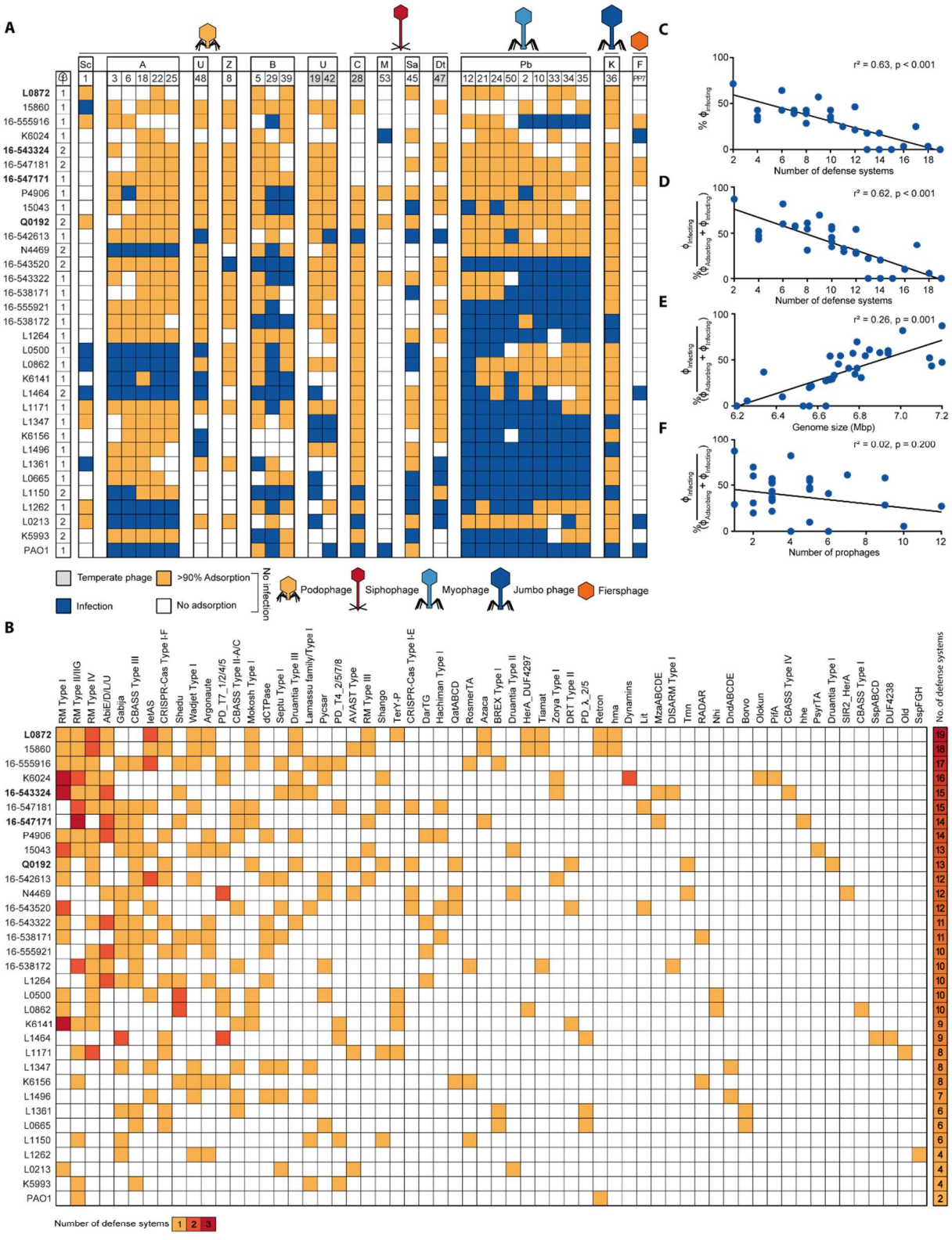

Figure 2. Innate and adaptive defense systems correlate with phage resistance. (a) Host range of phages against 32 *P. aeruginosa* clinical isolates and strain PAO1. Phages are clustered by phylogeny (**Table S3**). Phage-bacteria interactions are depicted as infection (blue), adsorption (>50%) but no infection (orange), or no interaction (white). Letters above the phage numbers indicate family or genus (for phages unassigned to a family): A, *Autographiviridae*; B, *Bruynoghevirus*; C, *Casadabanvirus*; Dt, *Detrevirus*; F, Fiersviridae; K, *Phikzvirus*; M, *Mesyanzhinovviridae*; Pb, *Pbunavirus*; Sa, *Samunavirus*; Sc, *Schitoviridae*; Z, *Zobellviridae*; U, unassigned. Strains not infected by phages in our panel are highlighted in bold. **(b)** Defense systems found in

We observed that some of the *P. aeruginosa* strains (L0872, Q0192, 16-543324, and 16-547171), which encode more defense systems than average, exhibited complete resistance to our phage panel (i.e. no phage could establish productive infection despite adsorbing to these strains in many cases) (**Figure 2a,b**). Therefore, we questioned which factor is the main driver of phage resistance among (i) the number of defense systems, (ii) genome size, and (iii) number of prophages (*49*). To assess this, we performed a multiple linear regression analysis including these factors, which showed that defense systems are the only relevant indicator of phage resistance ($r^2 = 0.63$ and $p < 0.001$ for correlation with % infecting phages; $r^2 = 0.62$ and $p < 0.001$ for correlation with % adsorbing phages that can establish a productive infection (% $\varphi_{Infecting}$ / ($\varphi_{Adsorbing}$+$\varphi_{Infecting}$)) (**Figure 2c-f, Fig. S4b-e**). The correlation between defense systems and phage resistance was reduced but remained significant when considering a conservative adsorption threshold of 90% ($r^2 = 0.35$ and $p < 0.001$) (**Figure S4f-h**), or when only one representative phage per family or genus was included (# defense systems $r^2 = 0.48$ and $p < 0.001$, genome size $r^2 = 0.11$ and $p = 0.033$, # prophage $r^2 = -0.02$ and $p = 0.6256$) (**Figure S4i-l**). Although the number of prophages did not show significant correlation with phage resistance, we did observe mechanisms of superinfection exclusion (*67*) for strains L1361 and L1496, which contain one prophage each that likely provides protection against closely related temperate phages φPa47 (100% pident, 100% coverage) and φPa42 (100% pident, 89% coverage), respectively.

Overall, our findings suggest that the number of defense systems in *P. aeruginosa* is associated with phage resistance. This observation is best exemplified by the five *P. aeruginosa* strains (L0872, 16-543324, 16-547171, 16-547181, Q0192) with 13-19 defense systems (**Figure 2a,b**) that were found to be resistant to the complete phage panel and to our attempts of phage isolation using wastewater from different sources.

In summary, we show that phage resistant strains of *P. aeruginosa* have accumulated phage defense systems in their genome, suggesting that phage defense systems could be a contributing factor of the phage sensitivity of the host.

Adaptive immunity targets temperate phages

Half (16/32) of the clinical strains contain adaptive immune systems in the form of CRISPR-Cas Type I-F (12 strains) and Type I-E (4 strains). To investigate the contribution of CRISPR-Cas to phage resistance we identified all spacers targeting our phage panel and assessed their potential effect on the phage host range. We detected 70 spacers (43 unique) across 16 strains matching our phage panel (**Figure S5a**), among which 54 are predicted to be interference-proficient (i.e., spacers with matching protospacer adjacent motif (PAM) and protospacer) and 16 priming-proficient (i.e., spacers with a ±1 slipped PAM (*68*) or up to 5 protospacer mutations (*69, 70*)) (**Table S4**). Most of the spacers (65) originate from CRISPR-Cas Type I-F systems, with only five spacers from CRISPR-Cas Type I-E systems (strains 15-547181, N4469, and 15043, all targeting φPa28). Interestingly, 66 of the 70 spacers target temperate phages (φPa19, φPa28, φPa42, and φPa47), and only four spacers match a virulent phage (φPa8). This is in line with previous findings that spacers of *P. aeruginosa* mostly match temperate phages (*71*). Our data on phage infection and adsorption reveal that in 83% (25/30) of cases with matching I-E (4/4) and I-F (21/26) spacers, the targeted phage was unable to infect (**Figure S5a**). Out of the five cases where the protective effects of matching spacers were not observed (**Figure S5a**), one was linked to the presence of an Acr (φPa42 infecting 16-542613) (**Figure S5b**, **Table S5**).

Overall, our results suggest that CRISPR-Cas Type I-E and I-F may contribute to resistance of the clinical strains against temperate phages but plays a minor role against the vast majority of virulent phages in the panel because they are not targeted. Although the catalogue of spacers may evolve through CRISPR adaptation, the current set of spacers alone does not explain the observed infection profiles.

Innate defense systems provide anti-phage activity against specific phage families

To understand the contribution of individual innate defense systems to broad-spectrum phage immunity, we inserted 14 individual defense systems (**Figure S6a**) from *P. aeruginosa* clinical isolates

into the low-copy plasmid pUCP20 (*72*), under their native promoters. The plasmids were introduced into *P. aeruginosa* strain PAO1, which is infected by 18 phages of our panel. We validated that the defense systems represent no obvious burden or toxicity to cell growth (**Figure S6b**, Kruskal-Wallis test followed by Dunn's post-hoc test), and subsequently assessed the defense-containing PAO1 strains for changes in phage susceptibility in solid and liquid media using efficiency of plating assays (EOP, **Figure 3a,b**) and infection dynamics assays to measure phage titer over time (**Figure S7**). We further monitored bacterial survival upon phage exposure using culture collapse assays (**Figure 3c**, and **Figure S8**).

The EOP (**Figure 3b**) and infection dynamics (**Figure S7**) assays show that 9 out of 14 defense systems exhibit activity against the phage panel. Of these, most (5) are active against at least two phage families or genus. For example, Zorya is active against podo *Bruynoghevirus*, sipho *Casadabanvirus*, and *Fiersviridae*. QatABCD and RADAR are both strong defenses against myo *Pbunavirus* and target members of different families of siphophages (*Casadabanvirus* and *Mesyanzhinovviridae*, respectively). CBASS Type III-C displays the broadest protection in our set, acting against podo *Autographiviridae*, sipho *Mesyanzhinovviridae* and *Casadabanvirus*, myo *Pbunavirus*, and Jumbo *PhiKZvirus*. This suggests that these defense systems employ sensing and targeting mechanisms that rely on phage features shared among different families, such as proteins (*21, 73-75*) or phage-induced changes in host metabolism (*76*) and cell integrity (*38, 77*), or that their effector is activated by other cellular responses, such as a general stress response. Interestingly, Zorya Type I is the only defense system among those tested that prevents infection of ssRNA phage PP7 (Fiersviridae), suggesting also that the effector may be activated by general cellular responses to phage infection. Additionally, CBASS Type III-C and AVAST Type V systems provide robust (>10^5-fold) protection against infection by phiKZ-like, nucleus-forming Jumbo phage φPa36 (**Figure 3b,c**). Both systems have been reported to act via altruistic cell death upon sensing of a specific phage protein (*21, 78, 79*), a strategy that (like RNA-targeting CRISPR-Cas systems (*15*)) circumvents the nuclear shell defense used by Jumbo phages to overcome DNA-targeting defense systems (*80-82*). The bacterial culture collapse assays provided additional information about the protective effect of the defense systems in liquid culture. TerY-P and Zorya Type I protect the cell population at both low and high phage multiplicity of infection (MOI) (**Figure 3c**, **Figure S8**), and most cells survive infection monitored using microscopy with propidium iodide as an indicator of membrane permeability and cell death (**Figure 3d**). For QatABCD, RADAR, Druantia Type III, AVAST Type V, CBASS Type III-C, II-A, and II-C, a culture collapse is observed when the phage is introduced at high MOI. In summary, our findings indicate that some individual defense systems display phage-targeting activity against specific phage families, while others are broader and target a few families. While most of the defense systems could provide protection against phages at low concentration, only two were efficient against phage at high densities.

Linking native infection profiles with protection patterns of individual systems

To understand the overall phage protection observed in genetically inaccessible clinical strains (**Figure 2**) in relation to the effect of individual phage defense systems (**Figure 3**), we initially assessed the phage infectivity levels in these strains using efficiency of plating assays (**Figure S9a**). We then compared the phage susceptibility profiles of the clinical strains with those of PAO1 strains equipped with a single defense system. Our findings indicate that in 84% of the cases, the phage susceptibility profile of the clinical strains aligns with the expected profile (**Figure S9b)**. This is especially evident for Zorya Type I, TerY-P, Druantia Type III, AVAST Type V, and CBASS Type II-A. The most notable disparity in the results was observed for RADAR and CBASS Type II-C, as they were providing protection against *Pbunavirus* in PAO1 and not in the clinical strains (φPa34 and φPa35 for RADAR; φPa2, φPa10, φPa33, and φPa35 for CBASS Type II-C).

To assess if known phage-encoded anti-defenses impact the phage infectivity profile of the clinical strains, we searched phage genomes for anti-defense genes including anti-RM (*83-85*), anti-CBASS (*73, 86*), anti-Pycsar (*86*), anti-TIR-STING (*87*), and anti-AVAST (*21*) proteins (**Figure S5b**, **Table S5**).

We focus here specifically on the anti-defenses against the defense systems introduced in PAO1, which include anti-CBASS and anti-AVAST. Our search identified one phage-encoded anti-defense gene, an anti-CBASS Type II (*acbII*) in phage φPa48. The *acbII* gene inhibited only the activity of CBASS type II-C in PAO1 (**Figure S5c**), but its impact on the phage host range in the clinical strains is not clear since φPa48 can only infect one (L1496) out of the two strains that carry CBASS Type II-C (**Figure 2a,b**, **Table S2**). This outcome is possibly due to other defense systems that target φPa48 in this strain.

Next, we focused on the anti-defense genes present within the bacterial genomes. We postulated that an increased quantity of these anti-defense genes within a strain might lead to increased susceptibility to our phage panel. To test this hypothesis, we performed a multiple linear regression analysis. Surprisingly, we observed that the presence of known anti-defenses does not have a detrimental effect on phage resistance of the host ($r = 0.36$, $r^2 = 0.13$, $p = 0.04$). This may be linked to the regulation of anti-defense gene expression in the host, as observed for Aca repression of Acrs (*88*). Among the anti-defense genes identified in the strains (**Figure S5b**), only *acbII* was found to target a defense system that was tested in PAO1. The *acbII* gene was found in strains L1347 and 16-547171, which carry the CBASS Type II-C and II-A systems, respectively, and here it is in line with the expected infection phenotype in four out of six cases (**Figure 2a**, **Figure S5b**). The suppression of *acbII* expression or the presence of other defense systems may be the reasons why in two cases the bacteria can resist phages despite having an anti-defense gene.

Overall, our findings underscore the intricate nature of phage susceptibility in natural settings, which is likely influenced by the interaction between different defense and anti-defense mechanisms present in both the strain and the phage.

Co-occurrence of defense systems with complementary specificities

Based on the observation that some defense systems provide distinct genera-specific anti-phage activities (**Figure 3**), we hypothesized that combinations of defense systems may be advantageous for cells by providing a wider protective range and would be a conserved feature in bacterial genomes to efficiently achieve broader antiviral specificity. To test this hypothesis, we assessed the co-occurrence (i.e. presence in the same genome) of defense systems in *P. aeruginosa* genomes in the RefSeq database ($n = 311$), while taking phylogeny into account (*89*).

We found multiple defense system co-occurrences (147 out of 1317 (11%) combinations tested, Bonferroni-corrected binomial exact test statistic with a p-value < 0.01), seven of which involved defense systems with anti-phage activity in this study (**Table S6**). Of these, six combinations have complementary phage specificity, including: i) Druantia Type III (*Pbunavirus*) and TerY-P (*Autographiviridae*), ii) Druantia Type III and Zorya Type I (*Autographiviridae*, *Bruynoghevirus*, *Casadabanvirus*, *Fiersviridae*), iii) Druantia Type III and AVAST Type V (*Phikzvirus*), iv) AVAST Type V (*Phikzvirus*) and QatABCD (*Casadabanvirus*, *Pbunavirus*), v) AVAST Type V and Zorya Type I (*Autographiviridae*, *Bruynoghevirus*, *Casadabanvirus*, *F*), and vi) TerY-P (*Autographiviridae*) with CBASS Type II (*Casadabanvirus*, *Pbunavirus*). Druantia Type III and CBASS Type II have overlapping specificity for *Pbunavirus*, with CBASS Type II adding specificity to *Casadabanvirus*. A binomial test indicates that complementary defense system combinations co-occurred significantly more often than expected (Binomial test: $n = 7$, $x = 6$, $p = 26/56$, $P(X \geq x) = 0.042$).

Overall, our analysis suggests that defense systems with complementary anti-phage activity co-occur at a probability higher than by chance in *P. aeruginosa* genomes and could provide an advantage for bacterial survival in phage diverse environments.

Discussion

Bacterial strains carry numerous distinct phage defense systems in their genomes (*63*). We found that *P. aeruginosa* strains carry at least 71% of all currently known defense systems, making this species a versatile bacterial model to study phage immunology. Using a diverse set of clinical isolates of *P. aeruginosa*, we observe that strains that have accumulated defense systems in their genome display broad and robust immunity against phage infection.

By testing the activity of 14 defense systems against our phage panel, we observed that the majority (7 out of 9 active systems) of the defense systems tested in PAO1 can protect the cell population at low phage concentration (MOI <1), but not at high concentration (MOI ≥ 1). This phenotype could be caused by the defense system being overwhelmed at high phage concentration, or by death or dormancy of the infected cell (*90*) which serves as a means of protecting the cell population through kin selection. In addition, we show that defense systems have anti-phage specificity that is often linked to a few phage families or genera, suggesting that these defense systems employ a more general sensing mechanism or that their effector is activated by other cellular responses. This is especially evident for Zorya Type I and CBASS Type III-C, which were effective against multiple distinct phage families. While Zorya remains largely uncharacterized, the current knowledge of CBASS activation suggests a variety of phage sensing strategies, including the recognition of peptides (HORMA domain) and dsDNA binding (cyclase) in *Escherichia coli* CBASS Type III-C (*79*), binding of structured phage RNA in *Staphylococcus schleiferi* CBASS Type I-B (*91*), and phage-driven depletion of folate-like molecules in *Vibrio cholerae* CBASS Type II (*92*).

Interestingly, we found that several pairs of the tested defense systems, which exhibit complementary anti-phage specificities, co-occur in *P. aeruginosa* strains. We expect that these combinations of defense systems could provide a broader range of phage protection through the complementary phage specificities of each individual defense system. The complementary activities of naturally co-occurring defense systems have also been found to enable resuscitation from defense system-induced bacterial dormancy (R-M and CRISPR-Cas) (*93*) and to prevent plasmid dissemination in *Vibrio cholerae* El Tor strains (DdmABC and DdmDE) (*94*), the latter proposing that defense system cooperation might play a role in bacterial pathogenicity.

The strong correlation found between number of defense systems and phage resistance further indicates that multiple defense system combinations are beneficial to cover the whole range of predating phages. The importance of the number of defense systems in determining phage resistance is further evidenced by the high levels of phage-resistance of five of our clinical isolates that encode between 13 to 19 defense systems. Attempts to isolate phages from diverse wastewater samples against defense-rich strains or to transform plasmid DNA proved more difficult, again pointing to the strains' inherent ability to defend well from incoming threats.

Importantly, our findings demonstrate the significance of the individual defense systems in predicting susceptibility of *P. aeruginosa* strains to phages. However, factors such as genetic context, interactions among defense systems (*93, 95-97*), the presence of unknown defense systems and anti-defense mechanisms, will affect the outcome of phage infection. Further research is required to enhance the predictive accuracy of genomic analysis, which could prove beneficial for ecological and evolutionary studies.

Altogether, our results show that while phage host range has traditionally been linked to receptor-associated factors (*98*), the number of defense systems is also a strong indicator of the susceptibility of cells to phage in *P. aeruginosa*. Naturally occurring *P. aeruginosa* clinical strains with many defense systems show an increased resistance to phages and may be selected for upon more widespread use of therapeutic phages. Therefore, monitoring the evolution and spread of phage-resistant clinical pathogens, and selecting or engineering phages with anti-defense properties may become instrumental to combat antimicrobial resistance using phage.

Materials and Methods

Bacteria

A set of 22 clinical isolates of *P. aeruginosa* provided by the University Medical Center Utrecht (UMCU) was used for phage isolation and 32 for characterization of host range (**Table S2**). The antibiotic susceptibility of the strains was established using the broth microdilution method outlined by EUCAST for determining the minimal inhibitory concentration and interpreted according to the EUCAST 2023 breakpoints (www.eucast.org). *Escherichia coli* strain Dh5α was used to clone plasmid pUCP20 with individual defense systems. *P. aeruginosa* strains containing pUCP20 with individual defense systems were constructed from *P. aeruginosa* strain PAO1. All bacterial strains were grown at 37 °C in Lysogeny Broth (LB) with 180 rpm shaking for liquid cultures, or in LB agar (LBA) plates for solid cultures. Strains containing plasmid pUCP20 were grown in media supplemented with 100 µg/ml of ampicillin (for *E. coli*) or 200 µg/ml of carbenicillin (for *P. aeruginosa*).

Bacteriophages

Phages used in this study are described in **Table S3**. All phages were isolated from sewage water. Approximately 1 ml of sewage sample was added to 20 ml of LB, inoculated with 100 µl of overnight cultures of each *P. aeruginosa* clinical isolate, and incubated overnight at 37 °C with 180 rpm shaking. Samples were centrifuged at 3,000 × *g* for 15 min and filter-sterilized (0.2 µm PES). The phage-containing supernatant was serially diluted in LB and spotted onto double-layer agar (DLA) plates of the isolation strains for the detection of phages. Single plaques with distinct morphologies were picked with sterile toothpicks and spread with sterile paper strips onto fresh bacterial lawns. The procedure was repeated until a consistent plaque morphology was obtained. Phages from purified plaques were then produced in liquid media with their respective host, centrifuged, filter-sterilized, and stored as phage lysates at 4 °C. For EOP and liquid infection assays (see below), phage stocks were obtained from lysates prepared on PAO1 and their concentration normalized to $\approx 1 \times 10^8$ pfu/ml. Additional efforts were made to isolate phages for the *P. aeruginosa* clinical isolates that exhibited the highest phage-resistance. This involved using sewage water from various sources and following the enrichment procedure outlined above, but with individual strains instead of mixtures.

Phage host range

Phages were 10-fold serially diluted in LB and spotted onto DLA plates containing each of the 32 *P. aeruginosa* clinical strain used for phage characterization (**Table S2**). The plates were incubated overnight at 37 °C and the phage plaques were observed to distinguish productive infection (lysis with individual phage plaques formed) from lysis from without (*99*) (lysis without individual phage plaques). Efficiency of plating of phages in each clinical strain was determined by comparing phage titer to that obtained in PAO1 (for phages that infect this strain) or in the clinical strain with the highest phage titer (for phages that cannot infect PAO1).

Adsorption assays

Early-exponential (optical density at 600 nm, $OD_{600} \approx 0.3$) cultures of the *P. aeruginosa* clinical isolates were added in triplicates to the wells of 96-well plates. Phages were added to these cultures at an MOI of 0.01 and incubated at 37 °C with 100 rpm shaking for 15 min. The plates were centrifuged, and a sample of the supernatant was taken, 10-fold serially diluted, and plated onto DLA plates of PAO1 to determine the titer of phages that did not adsorb to the clinical strain. A control plate in which phages were added to LB was used to determine the total phage concentration. The concentration of adsorbed phages was determined by subtracting non-adsorbed phage concentration from the total phage concentration in the suspension. The percentage of adsorbed phages was calculated as the ratio between adsorbed phages and total phages. Phages were considered to adsorb when over half of their population on average adhered to the cells.

Extraction of phage DNA and bacterial DNA

Phage DNA was extracted using phenol-chloroform. For this, 5 mL of each phage lysate at >10^9 pfu/mL were treated with 1 µg/mL of DNase I and RNase for 30 min. Ethylenediaminetetraacetic acid (EDTA), proteinase K, and sodium dodecyl sulfate (SDS) were added to the sample at final concentrations of 20 mM, 50 µg/mL, and 0.5% respectively, and the samples were incubated at 56 °C for 1 h. The samples were then mixed with an equal volume of chloroform and centrifuged at 3,000 × *g* for 10 min. The aqueous phase was recovered, and the procedure was repeated sequentially with a 1:1 mixture of phenol:chloroform, and chloroform. The resulting aqueous phase was mixed with 0.1 volume of sodium acetate 3M (pH 5) and 2.5 volumes of ice-cold absolute ethanol and incubated at -20 °C overnight. The extracted DNA was pelleted at 14,000 × *g* for 15 min and washed in ice-cold 70% ethanol, before re-suspending in ultrapure water. Bacterial genomic DNA was extracted using the GeneJET Genomic DNA Purification kit (Thermo Fisher). The quality and quantity of extracted phage and bacterial DNA were estimated using a NanoPhotometer and a Qubit fluorometer, respectively.

Phage genome sequencing

For samples sequenced at Beijing Genomics Institute (BGI) (**Table S3**), the phage genomic DNA was fragmented by Covaris 55 µL series Ultrasonicator and used to construct DNA nanoball (DNB)-based libraries by rolling circle replication. DNA was sequenced using the BGI MGISEQ-2000 platform (BGI Shenzhen, China) with a paired-end 100 nt strategy, generating 4.6-19.2 Gb sequencing data for each sample. For phage samples sequenced in-house, phage DNA was fragmented by Covaris M220 Focused-Ultrasonicator, and libraries were prepared using the NEBNext Ultra II DNA Library Prep Kit. Size distribution was checked on an Agilent D1000 Screen Tape System, and the libraries were pooled equally and spiked with approximately 5% of the PhiX control library. The pooled library was sequenced with an Illumina MiSeq using the MiSeq Reagent Nano Kit v2 (500-cycles). For samples sequenced at the Microbial Genome Sequencing Center (MiGS, Pittsburgh, PA, USA), sample libraries were prepared using the Illumina DNA Prep kit and IDT 10 bp UDI indices, and sequenced on an Illumina NextSeq 2000, producing 2x151 bp reads. Demultiplexing, quality control, and adapter trimming were performed with bcl-convert (v3.9.3). Reads obtained for all samples were assembled using Unicycler v0.5.0 (*100*). For samples sequenced in-house, the control PhiX was manually removed from the assembled contigs using Bandage (*101*).

Bacterial genome sequencing

For samples sequenced at BGI (**Table S2**), the bacterial genome was fragmented by Covaris 55 µL series Ultrasonicator and used to construct paired-end libraries with an insert size of 200-400 bp. Bacterial genomes were sequenced on the BGISEQ-500 (MGI, BGI-Shenzhen) platform, generating 1.4-2.0 Gb sequencing data for each sample with a sequencing depth >100x. Reads were checked for contamination using kraken2 (*102*) and only considered for further analysis if >90% of the reads identified as *P. aeruginosa*. Quality control of the raw data was performed using FastQC (*103*) with default parameters. For samples sequenced at MiGS, sequencing was performed as described above for phages. Reads obtained for all samples were assembled using Unicycler and the assembly quality was assessed using assembly-stats.v1.0.1 (*104*) and BUSCO.v4 (*105*) (pseudomonodales_odb10), and the GC% was calculated using bioawk (https://github.com/lh3/bioawk). The sequencing depth was calculated using minimap2 (*106*) and samtools mpileup (*107, 108*).

Bacteria genome annotation and phylogenomic

Bacterial genomes of the clinical strains were annotated using Prodigal (*109*). The genomes were used to determine the multi-locus sequence type (MLST) of the strains using the PubMLST website (https://pubmlst.org/) (*110*), and the serotype using the *P. aeruginosa* serotyper PAst (https://github.com/Sandramses/PAst). A total of 311 complete *P. aeruginosa* genomes were downloaded from RefSeq in February 2022. A phylogenetic tree of the core genome of *P. aeruginosa* was constructed using Parsnp (*111*) with default parameters using *P. aeruginosa* strain PAO1 (NC_002516.2) as the reference genome. Parsnp aligns microbial genomes to identify both structural

and point variations by searching for maximal unique matches (MUMs) to produce a core-genome alignment. SNPs in this core-genome alignment are filtered by Parsnp based on repetitive sequences, small locally collinear blocks (LCB) size, poor alignment quality, poor base quality, and possible recombination events. The final alignment was given to FastTree2 for the construction of the phylogenetic tree. Phylogeny groups were determined as previously described (*64*). The number of complete prophages present in RefSeq and clinical strains was predicted with virsorter2 v2.2.4(*112*), checkv v1.0.1 (*113*) (end_to_end with checkv-db-v1.0) and a second round of virsorter2 v2.2.4, following the protocol described in https://www.protocols.io/view/viral-sequence-identification-sop-with-virsorter2-5qpvoyqebg4o/v3. Superinfection exclusion was considered when the prophage and temperate phage shared a nucleotide similarity of pident >90% and coverage > 85%.

Phage genome annotation, taxonomy and phylogenomic

Phage genomes were annotated using the RAST server (*114*), the start of the phage genome was determined using PhageTerm (*115*), and partial genes were manually verified and removed. The phage lifestyle was predicted using PhageAI (*116*). Phages from our collection were classified taxonomically using GRAViTy (*117*). Phage diversity was evaluated using vConTACT2 (*66*) with the default settings and the ProkaryoticViralRefSeq94-Merged database, specifically selecting for *P. aeruginosa* phage genomes. The output of vConTACT2 was visualized in a circular layout using Cytoscape (*118*). Phages within the same family/genus were compared using clinker (*119*) for their similarity in gene structure, and on sequence level using blastn (*120*).

Detection of defense systems in bacterial genomes

Defense systems were detected in the *P. aeruginosa* genomes of RefSeq and clinical isolates with PADLOC-DB v1.4.0 (*121*), DefenseFinder (*63*), and the HMMs with completeness rules and thresholds as applied in Gao *et al* (2020) (*122*). In addition, the representative sequences provided by Rousset *et al* (2023) (*123*) were used to search for the defense system Detocs described in this work. Homology searches were performed via blastp (*120*) (> 0.7 subject length / query length < 1.5 ; 0.7 > query coverage < 1.3; evalue < 1e-9). Systems were considered complete when all genes were present without more than 2 genes in between. In case of discrepancies between the algorithms, we considered the output reporting the most hits. For PADLOC, we excluded defense systems of the "other" categories. For DefenseFinder, we excluded results of defense systems that were not discriminated into subtypes, e.g. BREX.

In addition, a manual search of the neighborhood of the defense systems identified by the algorithms led to the identification of a variant of the TerY-P system that contained all three genes and corresponding functional domains of the original system (*122*). The new TerY-P sequences were used to search for this variant in the bacterial genomes using blastp with evalue < 2.34e-29 and pident > 30. Systems were considered complete when all genes were present with less than 3 genes in between.

Detection of CRISPR-Cas I-F and I-E spacers targeting phages from our collection

Spacers were detected in the bacterial genomes using CRISPRDetect (*124*), and were mapped to our phage collection using blastn (word size = 8; evalue = 1; query coverage > 90; pident > 90; no gaps; maximum of 1 mismatch allowed). The non-target strand PAM (5'-CC for I-F, 5'-AAG for I-E) was manually checked, with a +1 or -1 PAM slippage allowed for I-F (*68*). Spacers with a matching PAM and protospacer were categorized as interference-proficient, while spacers with a PAM slippage or up to five protospacer mutations (with correct PAM) were categorized as priming-proficient spacers.

Detection of anti-defense genes in bacteriophage and bacterial genomes

Acrs were detected using AcrFinder (*125*). For the detection of anti-RM (*ardA* (*83*), *klcA* (*84*), *ardB*, *ocyA*, *ocr*, *darA*, *darB* (*85*)), anti-CBASS Type I (*acb1*) (*86*), anti-CBASS Type II (*acbII*) (*73*), anti-Pycsar (*apyc*) (*86*), anti-TIR-STING (*87*), and anti-AVAST (*lidtsur-6*, *lidtsur-17*, *forsur-7*, *penshu1-7*, *usur-3*, *smaarsur-6*, and *mellemsur-6*) (*21*) genes, we first searched for *P. aeruginosa* homologs using PSI-BLAST (*126*) (maximum of 3 runs with 500 sequences; coverage > 60%, pident > 20%). Homolog functionality

was checked using HMMer (*127*) and HHpred (*128*). *P. aeruginosa* homologs were only found for anti-defense genes *acb1* and *acbII*. These homologs were searched for in our phage and bacterial genomes with the use of blastp (evalue < 10^-8; pident > 30; coverage > 60%; 2.0 < subject length / query length > 0.5). For genes with no *P. aeruginosa* homologs, we created an HMM from the multiple alignment file obtained from the PSI-BLAST search above, using hmmbuild v3.3.2 (*127*) with default settings. These HMMs were used to search for the anti-defense genes in our phage and bacterial collections (evalue < 10^-6). All hits obtained were checked for the presence of the expected functional domains by HMMer and HHpred.

Co-occurrence of defense systems

Coinfinder (*89*) was used for detecting the co-occurrence of defense systems in the *P. aeruginosa* genomes of the RefSeq database, using the Parsnp (66) phylogenetic tree as input. We calculated the percentage of overlapping (for at least 1 phage) and complementary combinations in our set of tested defense systems. For this analysis, we combined the two CBASS type II subtypes (A and C) since our co-occurrence analysis in the RefSeq database was performed on the combination of all CBASS type II subtypes. A binomial test was then performed to test if the frequency of co-occurring complementary defense systems deviated from the expected.

Cloning of defense systems in PAO1

Defense systems were amplified from *P. aeruginosa* strains using the primers indicated in **Table S7** with Q5 DNA Polymerase (New England Biolabs), in reactions that added regions of homology to plasmid pUCP20. PCR products were run on 1% agarose gels and bands of the desired size were excised and cleaned using the Zymoclean Gel DNA Recovery Kit. Plasmid pUCP20 (pEmpty, **Table S8**) was digested with BamHI and EcoRI, treated with FastAP (Thermo Scientific), and cleaned with the Zymo DNA Clean & Concentrator Kit. Each defense system was cloned into pEmpty using the NEBuilder HiFi DNA Assembly Master Mix and transformed into chemically competent NEB® 5-alpha Competent *E. coli* following the manufacturer's instructions. Plasmids were extracted using the GeneJET Plasmid Miniprep kit, confirmed by sequencing (Macrogen, primers in **Table S7**), and electroporated into PAO1 as previously described (*129*). Briefly, an overnight culture of PAO1 was centrifuged at 16,000 × *g* for 2 min at room temperature, and the pellet was washed twice and resuspended in 300 mM of sucrose. The suspension was mixed with 100-500 ng of plasmid DNA and electroporated at 2.5 kV in a 2 mm gap electroporation cuvette. Cells were recovered in LB for 1-2 h at 37 °C and plated in LBA plates supplemented with 200 µg/ml of carbenicillin.

Cloning of anti-defense gene *acbII* into PAO1

Gene *acbII* was amplified from øPa48 using the primers indicated in **Table S7** with Q5 DNA polymerase, in reactions that added regions of homology to plasmid pSTDesR (*130*). Plasmid pSTDesR was amplified with the primers indicated in **Table S7** and digested with DpnI (New England Biolabs) for 1h at 37 °C. PCR products were run on 1% agarose gels and bands of the desired size were excised and cleaned using the Zymoclean Gel DNA Recovery Kit. Gene *abcII* was cloned into the amplified pSTDesR using the NEBuilder HiFi DNA Assembly Master Mix and transformed into chemically competent NEB® 5-alpha Competent *E. coli* following the manufacturer's instructions. Plasmids were extracted using the GeneJET Plasmid Miniprep kit, confirmed by sequencing (Macrogen, primers in **Table S7**), and electroporated into PAO1.

Efficiency of plating

The 10^8 pfu/ml phage stocks were 10-fold serially diluted in LB and the dilutions were spotted onto DLA plates of PAO1, or DLA+carbenicillin plates of PAO1 with pEmpty or PAO1 with individual defense systems following the small plaque drop assay (*131*). The phage dilution that resulted in countable phage plaques was used in double-layer overlay plaque assays (*132*) with PAO1, PAO1 with pEmpty, or PAO1 with the defense systems. The anti-phage activity of the systems was determined as the fold reduction in phage plaques in comparison to the number of plaques obtained in the

PAO1:pEmpty control. The diameter of the phage plaques was measured to determine differences in plaque size caused by the defense systems.

Infection dynamics of phage-infected cultures

Bacterial cultures of PAO1 with pEmpty or with individual defense systems at an OD_{600} ≈ 0.1 were infected with phage at an MOI <1. The cultures were incubated at 37 °C with rocking, and samples were taken at 0h, 2h, 4h, and 6h to measure phage concentration. The sample was centrifuged at 3,000 × *g* for 5 min, and the phage-containing supernatant was 10-fold serially diluted and spotted onto DLA plates of PAO1 to estimate phage concentration.

Liquid culture collapse assays

Overnight grown bacteria were diluted to an OD_{600} of approximately 0.1 in LB media. The cell suspension was distributed into the wells of 96-well plates, and phages were added at MOIs of 10, 0.1, 0.01, and 0.001). Assays were performed in triplicates. The plates were incubated at 37°C in an Epoch2 microplate spectrophotometer (Biotek) for OD_{600} measurements every 10 min for 24h, with double orbital shaking. The growth rate of uninfected cells carrying the empty plasmid and each defense system was determined by comparing the OD measurement at the beginning of the log phase (OD1 at 3h) to that at the end (OD2 at 6h), using the natural log: ln((OD2-OD1) / (t2 – t1)). The growth rate of the cells carrying defense systems was compared to that of the cells carrying an empty plasmid using the Kruskal-Wallis test followed by Dunn's post-hoc test.

Sequencing of phage-infected bacterial strains

The genomic DNA of bacterial strains infected with phage for 6h was extracted using the GeneJET Genomic DNA Purification kit. The quality and quantity of extracted phage and bacterial DNA were estimated using a NanoPhotometer and a Qubit fluorometer, respectively. The bacterial genome was sequenced at Plasmidsaurus (US). 5% of the reads with the lowest quality were filtered out using Filtlong v0.2.1 (default) (*133*). Miniasm v0.3 (*134*) was used to create a first draft of the assembly using a 250 Mb subset of reads. Low quality reads were removed until a sequence depth of around 100x was achieved. Reads were assembled using Flye v2.9.1 (*135*) with parameters selected for high quality ONT reads. Polishing of the assembly was conducted using Medaka v1.8.0 (https://github.com/nanoporetech/medaka) with the reads. Blastn v2.14.1 (*120*), with query coverage > 70% and pident > 70% was used to search for integrated temperate phages. Temperate phages were considered integrated in instances where the contig was at least 30% larger than the phage itself.

Fluorescence microscopy of defense systems

Exponentially growing (OD_{600} ≈ 0.3) cultures of PAO1 strains containing pEmpty or the defense systems were infected with phage at an MOI of ≥ 3, and the phage was adsorbed for 10 min at 37 °C. Cells were centrifuged at 9,000 × *g* for 1 min, and the cell pellet was re-suspended in 5 µL of 1 µM of propidium iodide. The stained cells were spotted onto 1% agarose pads (*136*), and visualized using a Nikon Eclipse Ti2 inverted fluorescence microscope equipped with a 100× oil immersion objective (Nikon Apo TIRF; 1.49 numerical aperture). Time-lapse phase-contrast (CD Retiga R1) and fluorescence images (after excitation with a 561 nm laser 2000 609/54 bandpass filter, EM-CCD Andor iXON Ultra 897) were acquired every 5 min using Metamorph.

Statistical analysis

Unless stated otherwise, data are presented as the mean of biological triplicates ± standard deviation. All correlation analysis were determined by linear (multiple) regression models using the lm function of R, and p-values were adjusted with the Bonferroni post-hoc test. The significance of differences between phylogenetic groups was determined using the Kruskal-Wallis test with Dunn's post-hoc test, while the differences in infection dynamics were determined by two-way ANOVA followed by Sidak's multiple comparison test. For all statistical analysis, a significance level of 0.05 was used.

Supplementary material

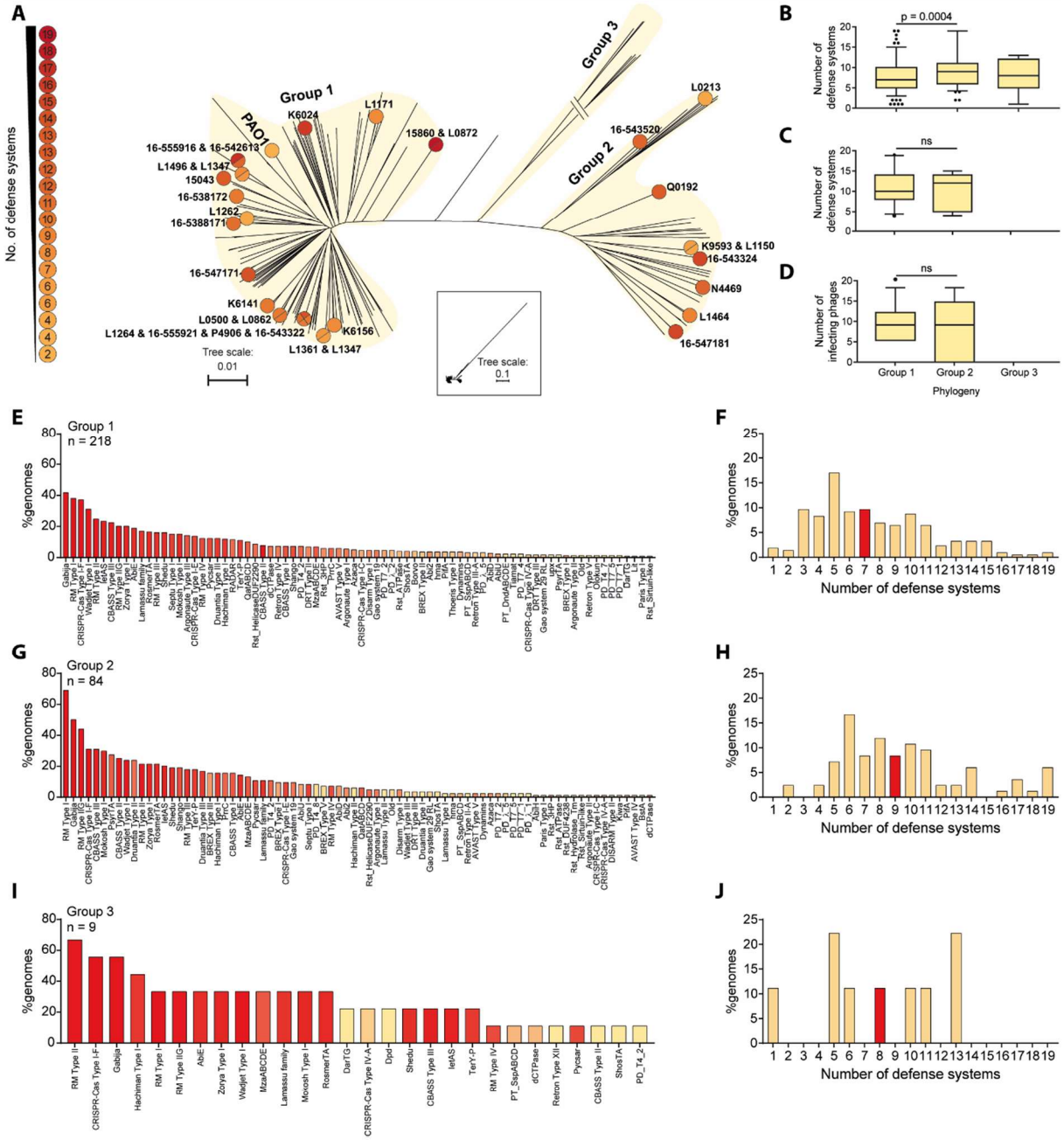

Figure S1. Phylogenomic analysis of *P. aeruginosa* strains. (a) Maximum likelihood phylogenetic tree of *P. aeruginosa* strains built using core genome SNPs mapped to the reference strain PAO1 (NC_002516.2) using Parsnp and corrected for branch length with ClonalFrameML to account for recombination. Circles and squares indicate the positions of clinical strains in phylogenetic groups 1 and 2, respectively, and are colored based on the number of defense systems present in each strain. **(b)** Comparison of the number of defense systems found in the RefSeq *P. aeruginosa* genomes among phylogenetic groups. **(c)** Comparison of defense systems in the genomes of the 32 clinical strains among phylogenetic groups. **(d)** Comparison of the number of infecting phages in the genomes of the clinical strains among phylogenetic groups. For panels (b), (c), and (d), statistical analysis was 3 determined using the Kruskal-Wallis test with Dunn's post-hoc test, with a significance value of 0.05. **(e)** Diversity of defense systems and **(f)** number of defense systems per genome in *P. aeruginosa* strains from the RefSeq database of phylogeny group 1. **(g)** Diversity of defense systems and **(h)** number of defense systems per genome in RefSeq *P. aeruginosa* of phylogeny group 2. **(i)** Diversity of defense systems and **(j)** number of defense systems per genome in RefSeq *P. aeruginosa* of phylogeny group 3. In panels (e), (g), and (i), the defense systems are organized from most (left) to least (right) prevalent and colored according to the abundance in Figure 1a. In panels (f), (h), and (j), the median number of defense systems is shown in red.

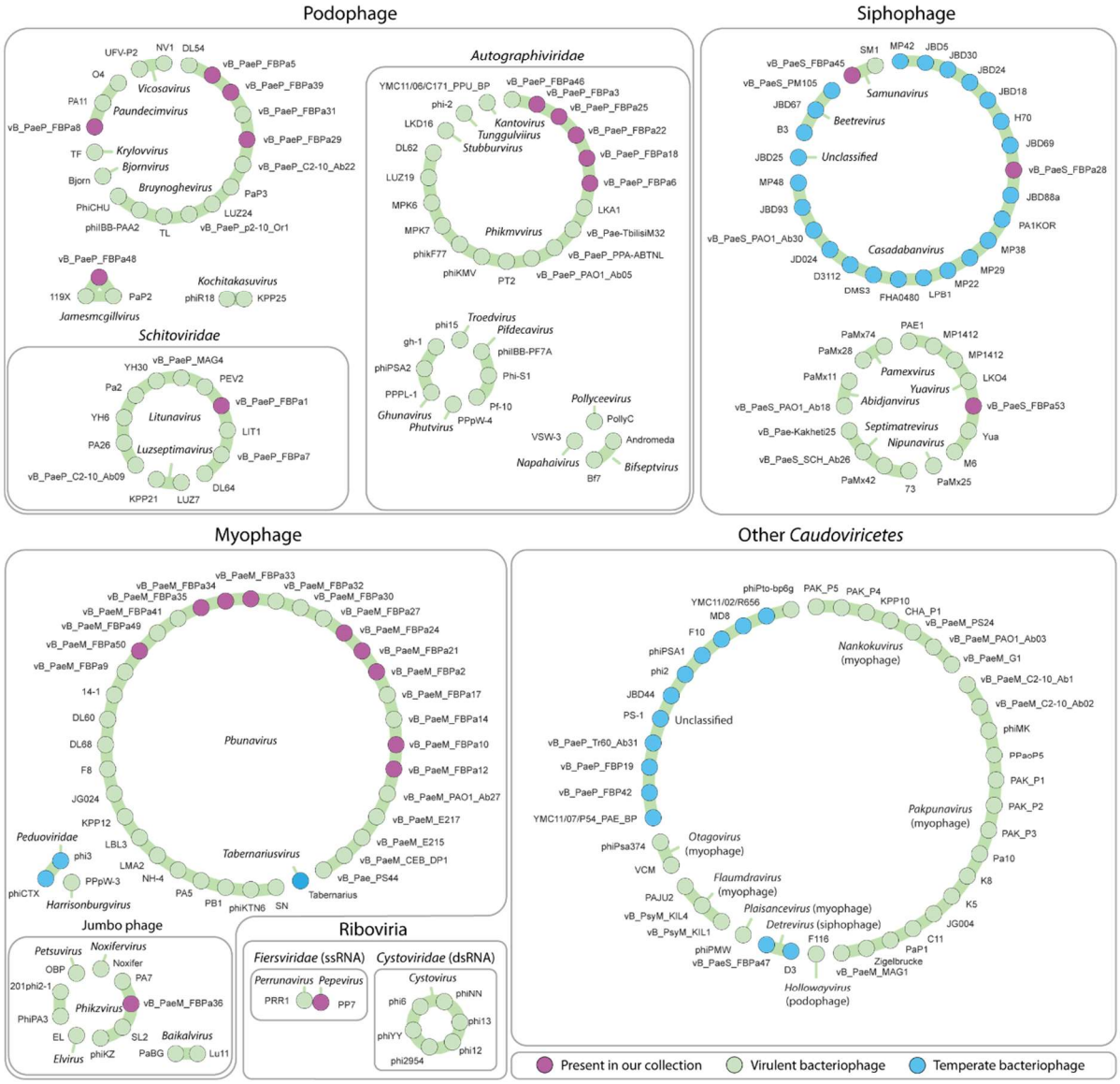

Figure S2. Diversity of the phage panel used in this study. *P. aeruginosa* phages were clustered using vConTACT2, based on guilt-by-contig-association classification of genomic sequences, and visualized in Cytoscape. The phage families and genus are indicated in italics, and the phages from the collection are highlighted in purple.

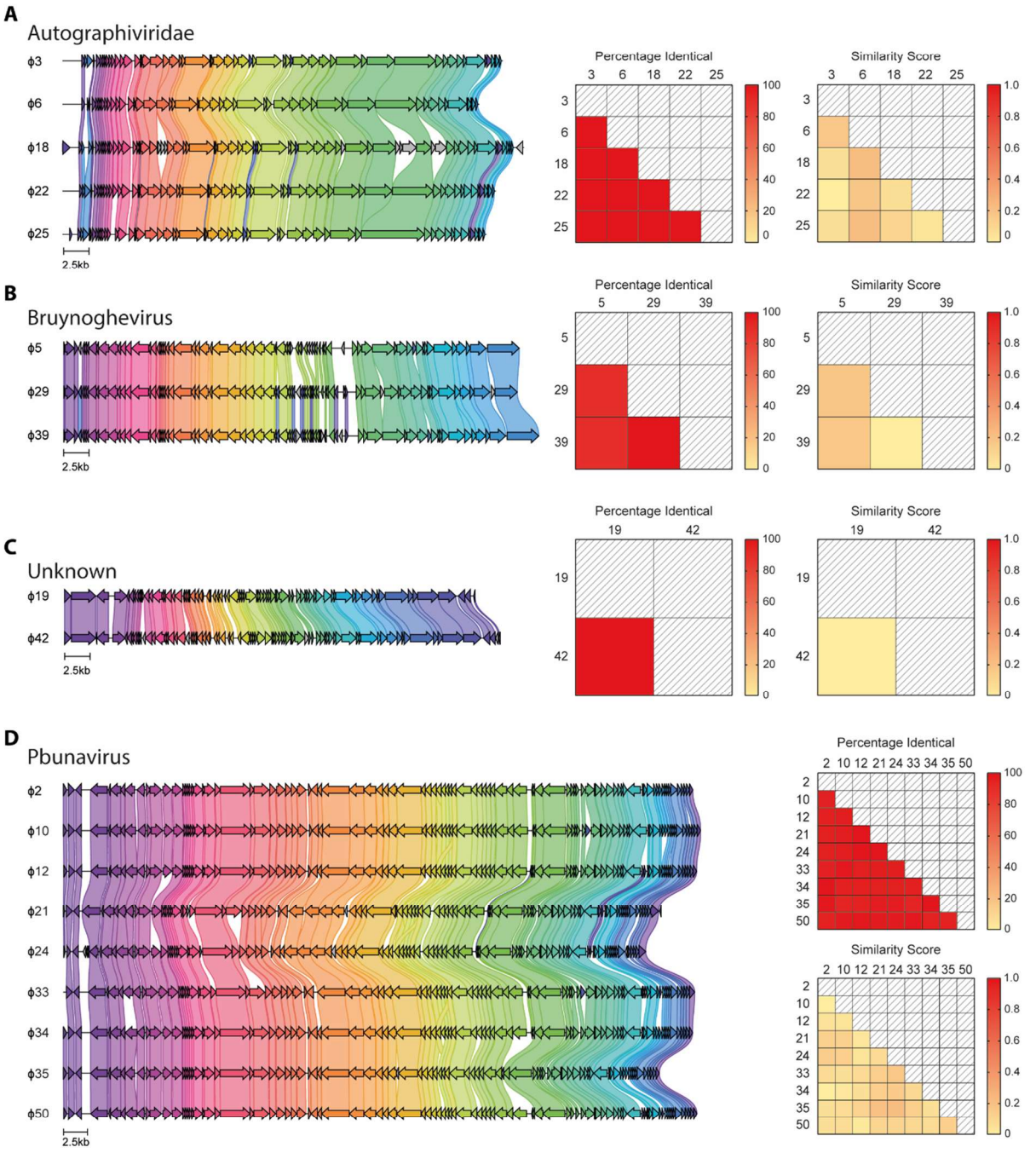

Figure S3. Gene architecture and sequence similarity among phages within the same family or genus. Gene cluster comparisons (left) and percentage identify and similarity scores (right) for **(a)** *Autographiviridae*, **(b)** *Bruynoghevirus*, **(c)** Unknown, and **(d)** *Pbunavirus*.

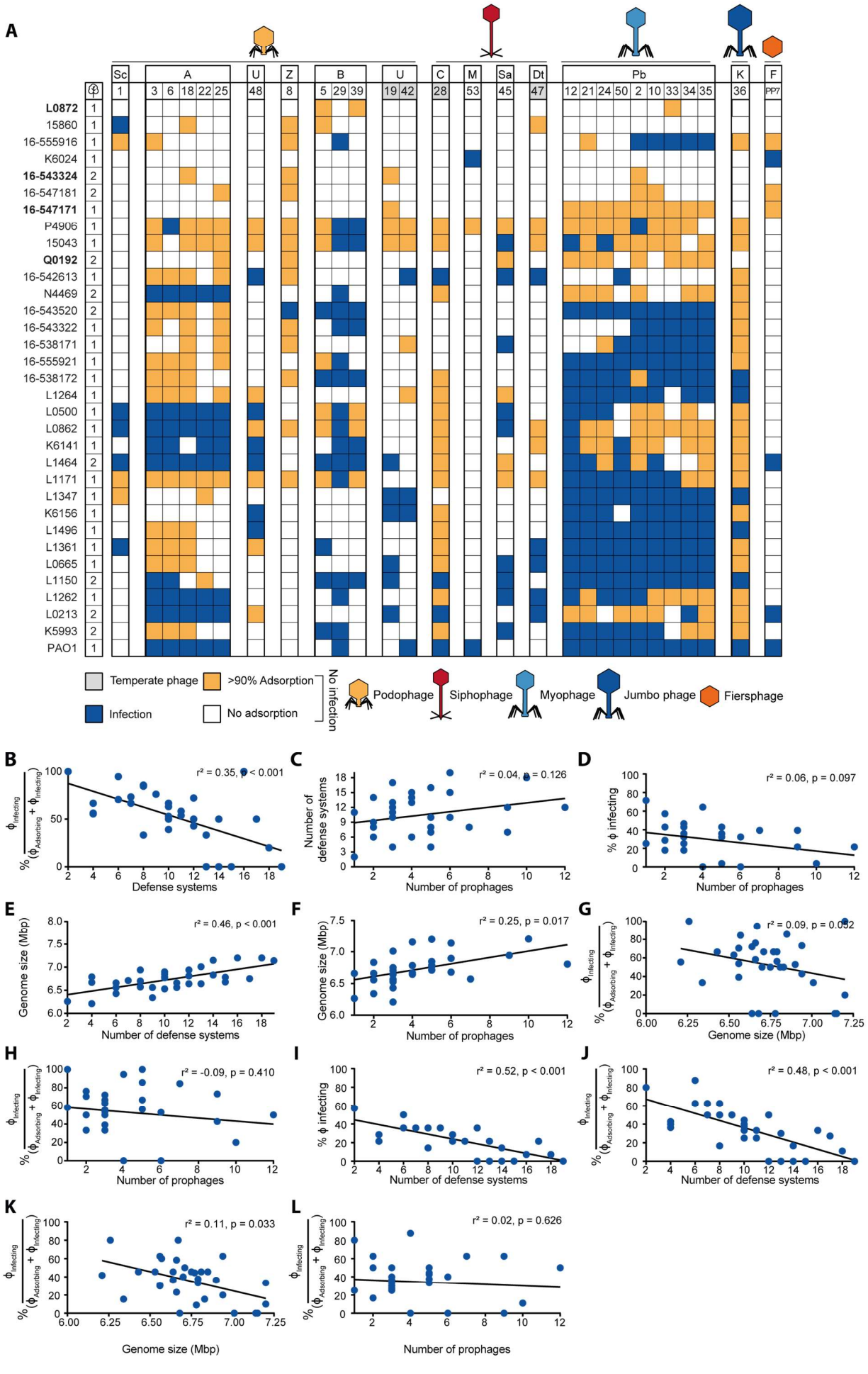
A
Sc
A
U
Z
B
U
C
M
Sa
Dt
Pb
K
F
1
3 6 18 22 25
48
8
5 29 39
19 42
28
53
45
47
12 21 24 50 2 10 33 34 35
36
PP7
L0872
15860
16-555916
K6024
16-543324
16-547181
16-547171
P4906
15043
Q0192
16-542613
N4469
16-543520
16-543322
16-538171
16-555921
16-538172
L1264
L0500
L0862
K6141
L1464
L1171
L1347
K6156
L1496
L1361
L0665
L1150
L1262
L0213
K5993
PAO1
Temperate phage
>90% Adsorption
Infection
No adsorption
No infection
Podophage
Siphophage
Myophage
Jumbo phage
Fiersphage
B
r² = 0.35, p < 0.001
Defense systems
C
r² = 0.04, p = 0.126
Number of prophages
D
r² = 0.06, p = 0.097
% ϕ infecting
Number of prophages
E
r² = 0.46, p < 0.001
Genome size (Mbp)
Number of defense systems
F
r² = 0.25, p = 0.017
Genome size (Mbp)
Number of prophages
G
r² = 0.09, p = 0.052
Genome size (Mbp)
H
r² = -0.09, p = 0.410
Number of prophages
I
r² = 0.52, p < 0.001
% ϕ infecting
Number of defense systems
J
r² = 0.48, p < 0.001
Number of defense systems
K
r² = 0.11, p = 0.033
Genome size (Mbp)
L
r² = 0.02, p = 0.626
Number of prophages

Figure S4. Linear regression analysis of correlation between different variables and phage resistance in *P. aeruginosa*. (a) Host range of phages against 32 clinical 7 isolates of *P. aeruginosa* and strain PAO1. Phages are clustered by phylogeny. Phage-bacteria interactions are depicted as infection (blue), adsorption (>90%) but no infection (orange), or no interaction (white). Letters above the phage numbers indicate family or genus (for phages unassigned to a family): A, *Autographiviridae*; B, *Bruynoghevirus*; C, *Casadabanvirus*; Dt, *Detrevirus*; F, Fiersviridae; K, *Phikzvirus*; M, *Mesyanzhinovviridae*; Pb, *Pbunavirus*; Sa, *Samunavirus*; Sc, *Schitoviridae*; Z, *Zobellviridae*; U, unassigned. Linear regression analysis considering all 28 phages of the panel for **(b)** genome size and number of prophages; **(c)** number of defense systems and number of prophages; **(d)** percentage of infecting phages and number of prophages; **(e)** genome size and number of defense systems. Linear regression analysis considering all 28 phages and adsorption at a conservative level of 90% for **(f)** percentage of adsorbing phages that can establish a productive infection (% $\varphi_{Infecting}$ / ($\varphi_{Adsorbing}$ + $\varphi_{Infecting}$)) and number of defense systems; **(g)** percentage of adsorbing phages that can establish a productive infection and genome size; (h) percentage of adsorbing phages that can establish a productive infection and number of prophages. Linear regression analysis considering a selection of representative phages for **(i)** percentage of infecting phages and number of defense systems; **(j)** percentage of adsorbing phages that can establish a productive infection (% $\varphi_{Infecting}$ / ($\varphi_{Adsorbing}$ + $\varphi_{Infecting}$)) and number of defense systems; **(k)** percentage of adsorbing phages that can establish a productive infection and genome size; **(l)** percentage of adsorbing phages that can establish a productive infection and number of prophages. r^2 represents adjusted R-squared, a goodness-of-fit measure for the linear regression models. Representative phages include φPa1, φPa2, φPa8, φPa12, φPa28, φPa34, φPa36, φPa39, φPa42, φPa45, φPa47, φPa48, φPa53, and φPP7.

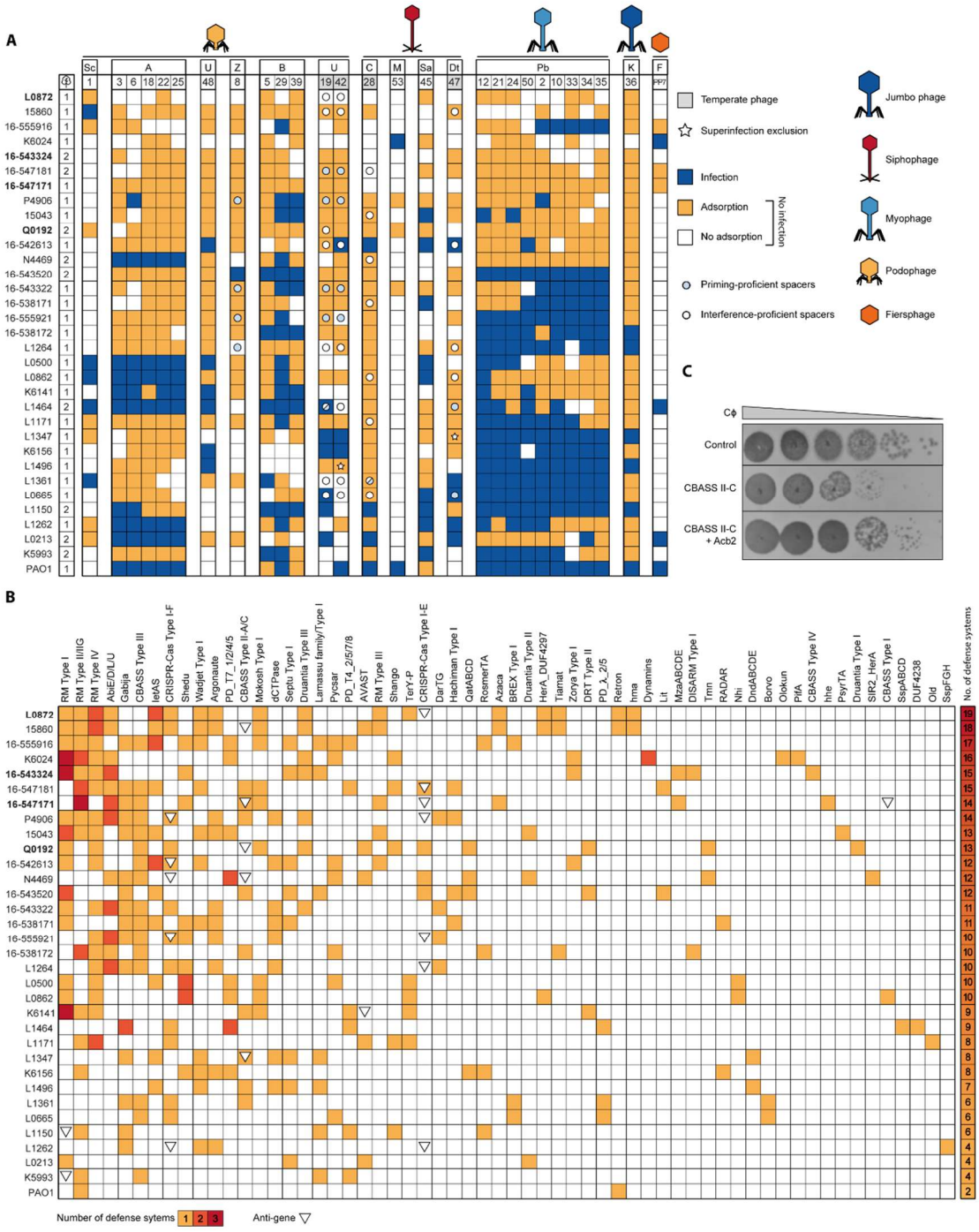

Figure S5. Role of adaptive immunity, superinfection exclusion, and anti-defense genes in defining the phage host range in *P. aeruginosa*. (a) Spacers found in the clinical isolates to match the phages in our panel. Spacers that have a matching protospacer adjacent motif (PAM) and protospacer are labelled as interference-proficient, indicated by white circles. Spacers with a ±1 slipped PAM or with up to 5 protospacer mutations are labelled as priming-proficient, indicated by blue circles. Superinfection exclusion occurrences are 9 denoted with a star symbol when a strain carries a prophage that resembles a temperate phage in the phage panel, with a pident >90% and coverage >85%. Phage-bacteria interactions are depicted as infection (blue), adsorption but no infection (orange), or no interaction (white). Letters above the phage numbers indicate family or genus (for phages unassigned to a family): A, *Autographiviridae*; B, *Bruynoghevirus*; C, *Casadabanvirus*; Dt, *Detrevirus*; F, Fiersviridae; K, *Phikzvirus*; M, *Mesyanzhinovviridae*; Pb, *Pbunavirus*; Sa, *Samunavirus*; Sc, *Schitoviridae*; Z, *Zobellviridae*; U, unassigned. **(b)** Anti-defense genes found in the genomes of the clinical isolates are indicated with an inverted triangle in the position corresponding to the defense system potentially affected by the anti-defense gene. The number of instances of each defense system type per strain is indicated in yellow, orange, or red for 1, 2, or 3 respectively. The total number of defense systems found per strain is indicated in a heatmap bar on the right. The complete list of spacers and anti-defense genes found in the clinical strains and phages can be found in **Table S4** and **Table S5**, respectively. (c) Spot test of phage φPa33 onto top agar lawns of PAO1 with empty plasmid (control), CBASS Type II-C, or CBASS Type II-C with Acb2.

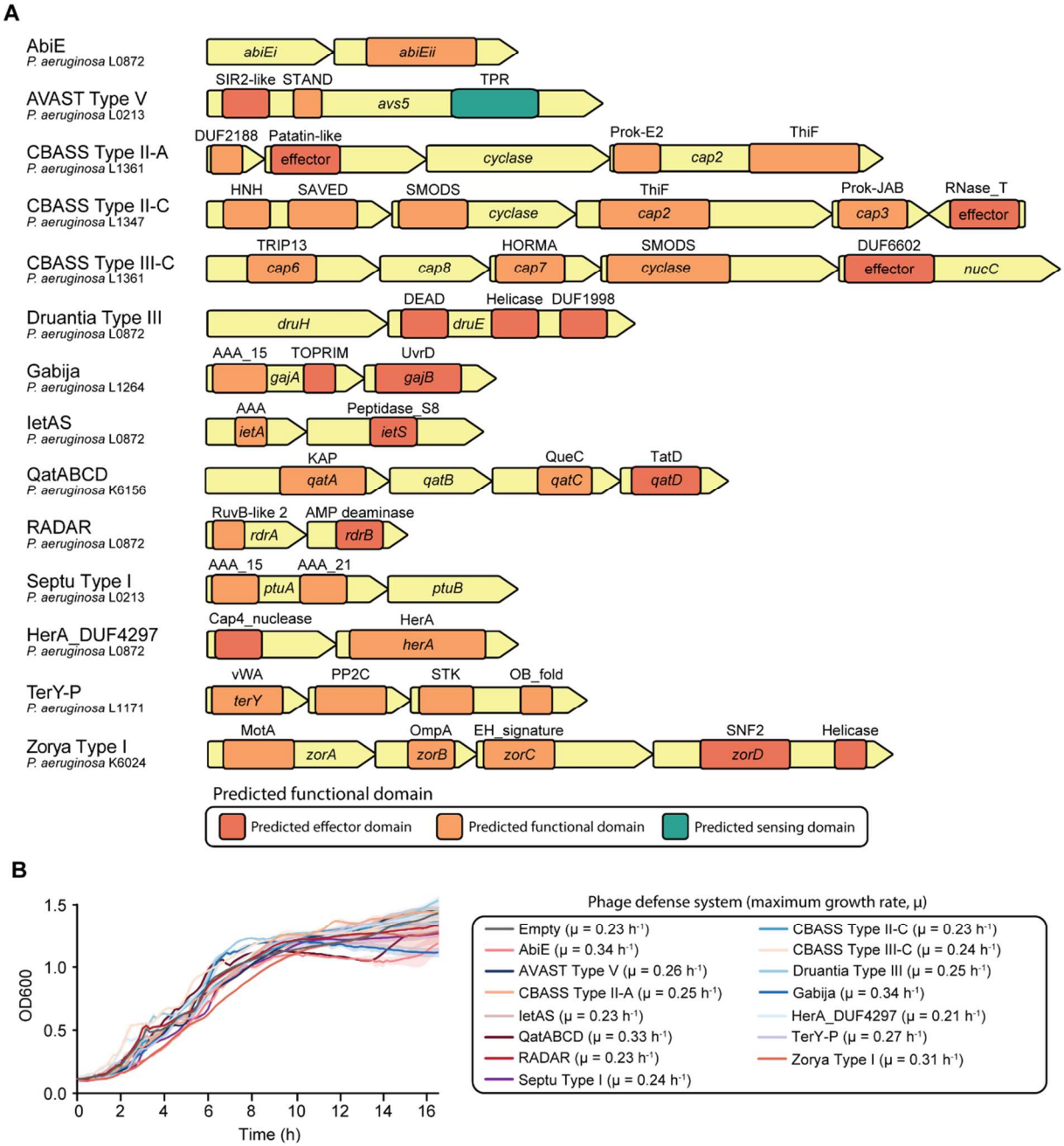

Figure S6. Individual defense systems cloned in *P. aeruginosa* strain PAO1. (a) Gene composition and functional domains of the 14 defense systems cloned into *P. aeruginosa* strain PAO1. **(b)** Growth of PAO1 strains containing individual defense systems. The maximum growth rate, μ, is shown next to the defense system name in the legend. No toxic effect is observed in cell growth, as measured by optical density at 600 nm.

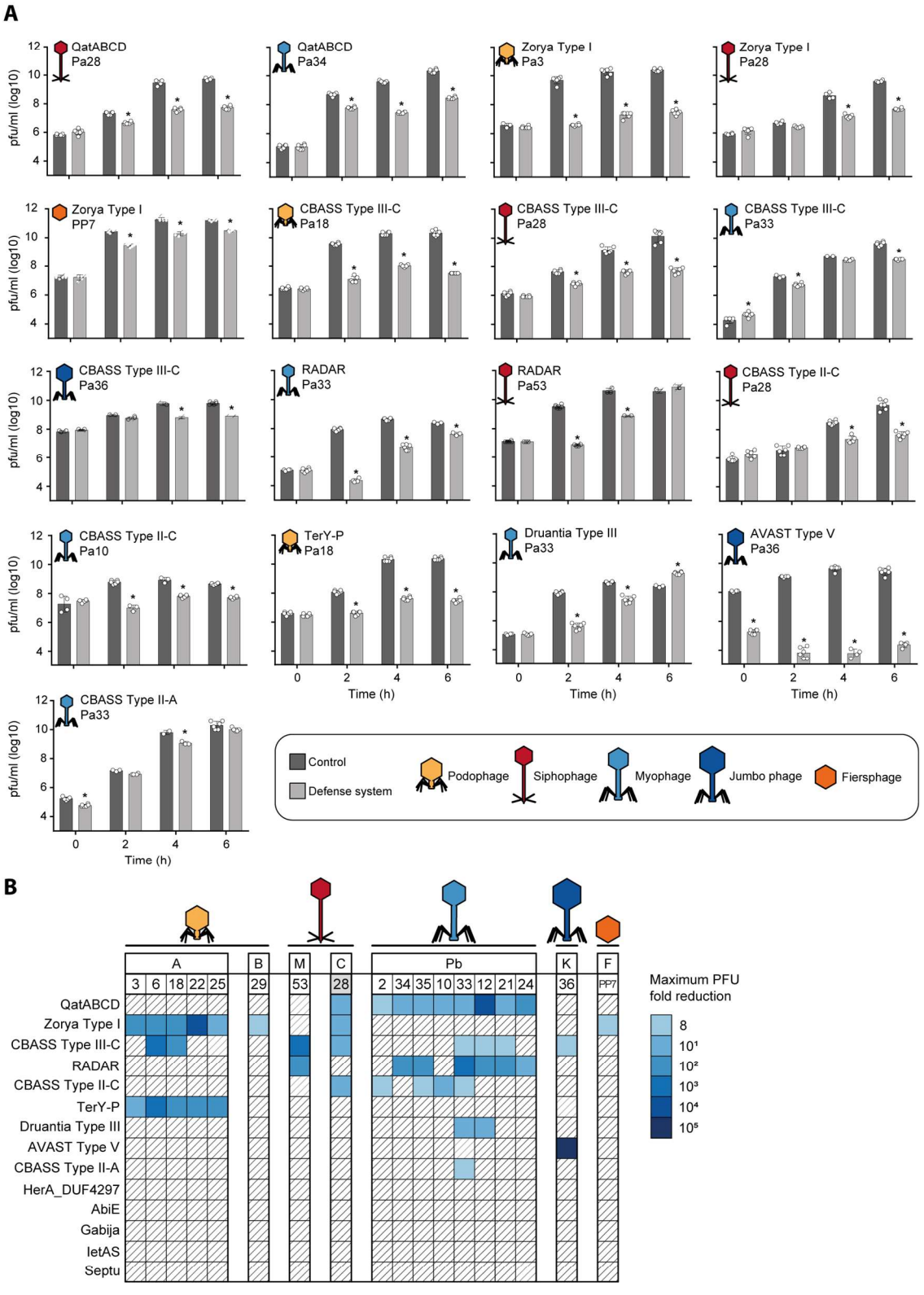

Figure S7. Effect of individual defense systems on phage infection dynamics in *P. aeruginosa* PAO1. (a) Graphical representation of the infection dynamics performed with phages representative of the families targeted by each defense system. Results are shown as the mean ± standard deviation of the phage concentration measured at 12 0-, 2-, 4-, and 6-hours post-infection of the cells containing an empty plasmid (control) or defense system. Statistically significant differences ($p < 0.01$) were determined by two-way ANOVA followed by Sidak's multiple comparison test and are indicated with *. **(b)** Heatmap summarizing the results of the infection dynamics assay. The heatmap shows the maximum fold reduction in phage concentration obtained during the infection dynamics assay. Letters above the phage numbers indicate family or genus (for phages unassigned to a family): A, *Autographiviridae*; B, *Bruynoghevirus*; C, *Casadabanvirus*; F, Fiersviridae; K, *Phikzvirus*; M, *Mesyanzhinovviridae*; Pb, *Pbunavirus*; U, unassigned.

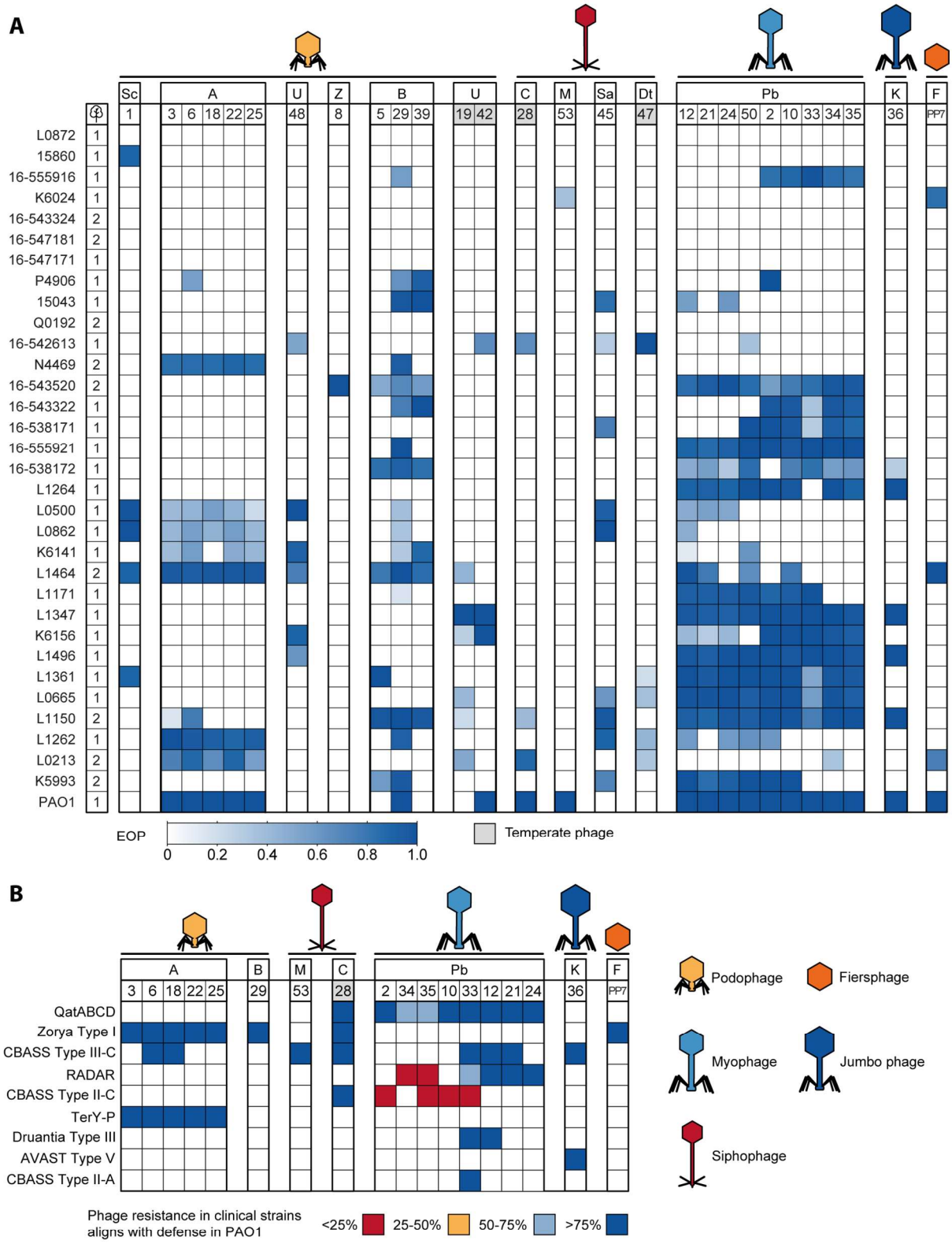

Figure S8. Effect of individual defense systems on bacterial growth during phage infection. Graphical representation of the bacterial growth upon infection with phage representatives of the families targeted by each defense system. Results are shown as the average optical density at 600 nm of cultures of strains containing a control plasmid (WT) or defense system (DS), uninfected or infected with phage at a multiplicity of infection (MOI) of 10 or 0.01.

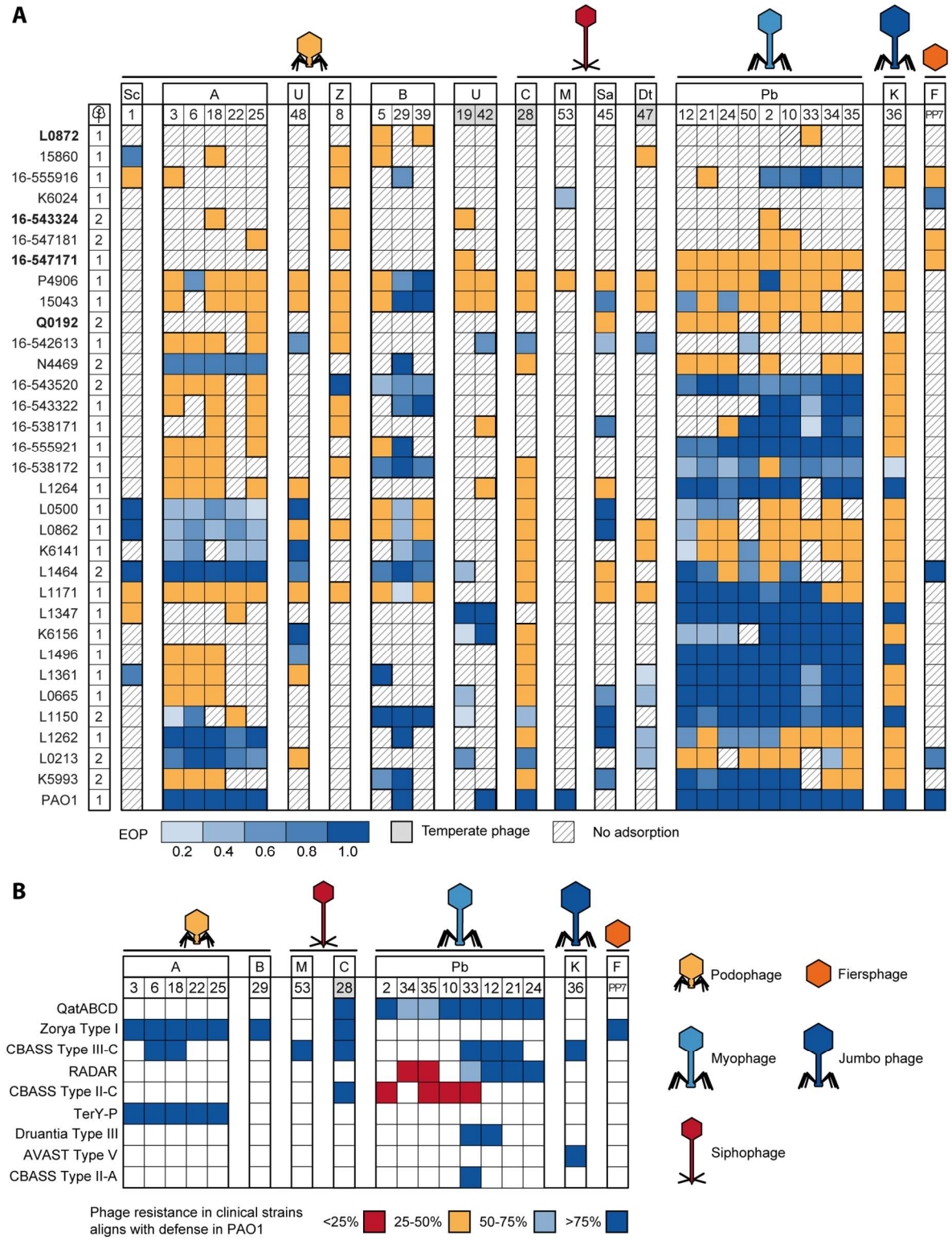

Figure S9. Predictive value of defense system activity in PAO1 to phage susceptibility of the clinical strains. **(a)** Efficiency of platting (EOP) of 28 phages against the panel of 32 clinical strains and PAO1. The EOP was determined as the ratio of the phage concentration in each clinical strain to that in either PAO1 or the strain that the phage infects most efficiently. **(b)** Heatmap representation of matches between the phage infectivity profile in the clinical 15 strains and the expected protection based on assays in PAO1 strains containing individual defense systems. The matches are categorized into four groups: 75%, which are color-coded as red, yellow, light blue, and dark blue, respectively. Letters above the phage numbers indicate family or genus (for phages unassigned to a family): A, *Autographiviridae*; B, *Bruynoghevirus*; C, *Casadabanvirus*; Dt, *Detrevirus*; F, Fiersviridae; K, *Phikzvirus*; M, *Mesyanzhinovviridae*; Pb, *Pbunavirus*; Sa, *Samunavirus*; Sc, *Schitoviridae*; Z, *Zobellviridae*; U, unassigned.

The following supplementary files are available online at doi: 10.1126/sciadv.adj0341

Table S1. Matrix of defense systems identified in the 311 RefSeq genomes of *P. aeruginosa*.

Table S2. Features of the clinical isolates of *P. aeruginosa* used in this work, including defense system presence.

Table S3. Features of the *P. aeruginosa* phages used in this study.

Table S4. List of CRISPR-Cas Type I-F and I-E interference-proficient and priming-proficient spacers found in the clinical isolates of *P. aeruginosa* to target phages from our collection.

Table S5. Anti-defense genes found in our collection of *P. aeruginosa* clinical strains and phages.

Table S6. Co-occurrence of defense systems in the *P. aeruginosa* genomes of the RefSeq database identified by Coinfinder.

Table S7. List of primers used in this work.

Table S8. List of plasmids used in this work.

Chapter 3



R2-type pyocins are associated with phage defense

To survive the constant phage predation, bacteria encode phage defense systems. Over 200 phage defenses have been identified to date, yet it remains unknown whether this is the full extent of the bacterial immune response. Identification of additional phage defense systems is important for understanding the bacterial immunity against phages. In this study, we searched for previously unknown phage defense systems through genome-wide association studies (GWAS) using the phage-sensitivity profiles of 33 *Pseudomonas aeruginosa* bacterial strains for 28 *Pseudomonas* phages. We identified only one gene cluster that significantly associates with phage resistance. This gene cluster is an R2-type pyocin, a phage tail-like bacteriocin that lyses nearby cells and is released during stress conditions and bacterial conflict. We propose the hypothesis that these pyocins may convey phage defense by lysing phage-infected kin.

This chapter has been a collaborative effort by
van den Berg, D. F., Fierlier, D., & Brouns, S. J. J. (2025).

Introduction

Bacteria are under constant predation from their viruses, phages (*49*). These phages infect bacteria to replicate themselves and lyse the bacterial host in the process (*49*). To protect themselves, bacteria evolved a multitude of phage defenses (*49*). To date, more than 200 bacterial phage defenses have been discovered, including the well-known CRISPR-Cas systems, as well as the more recently discovered phage defense systems (*59, 122, 137-142*). Despite the immense number of discovered phage defense systems, the full extent of the phage defense repertoire of bacteria remains unknown. Identification of these unknown phage defense systems is key for understanding how bacteria become resistant to phages, particularly for using phages to treat bacterial infections (*143*).

A multitude approaches have been used to search for these unknown phage defenses, including guilt-by-association studies, which leverages that genes with similar functions often co-localize within the bacterial genome (*59, 122, 137-142*). As well as functional screens such as transposon insertion assays, where transposon insertions result in the loss of phage resistance of the host (*144*). However, analytic phenotype-driven approaches such as using the phage-sensitivity profiles of a set of bacterial strains to discover novel phage defenses is yet to be applied. Where the presence of specific genes is associated with an increased resistance to a specific phage, also known as genome-wide-association studies (GWAS) (*145*).

Here, we set out to perform such a GWAS approach on the phage-sensitivity-profiles of *Pseudomonas aeruginosa* bacterial strains for 28 diverse *Pseudomonas* phages to correlate the presence of certain bacterial genes with a change in phage sensitivity (EOP) (*146*).

Results and discussion

The GWAS using linear models that were implemented by pyseer (*147*). To take the phylogeny into account, we also included a phylogenetic tree, that was creation from the core genes of our set of *P. aeruginosa* strains using panaroo, into the linear model (*148*). From this analysis, 13 genes (**Table S1**) were predicted to affect the phage sensitivity of the *P. aeruginosa* strains (**Figure 1a,b**). These genes were associated with the gain of resistance against *Pseudomonas* phage vB_PaeP_FBPa39 (ΦPa39), a phage that is part of the *Bruynoghevirus* genus with a podophage morphology (*146*). To our surprise, all these genes were found to be part of the same gene cluster that encodes of a R2-type pyocin, suggesting that R2-type pyocins are involved in the phage defense of *P. aeruginosa* isolates (**Figure 1c, S1**) (*149-153*).

R2-type pyocins are phage-like particles that have lost their ability to replicate and only consist of a phage-tail (**Figure 1d**) (*149-152*). R2-type pyocins originate from ancient P2-phages and are released during bacterial conflict to lyse competing bacterial strain (*153*). The possible involvement of R2-type pyocins in phage defense suggests that R2-type pyocins are also important during bacterial-phage conflict, where ancient phage remnants provide phage defense against new phage infections.

Apart from the phage-tail genes of the R2-type pyocin, the neighbouring genes of this gene cluster encode for essential functions for the pyocin as well, including genes that regulate its expression and translation, and the lysis machinery (**Figure S1**) (*149-153*).

The prevalence of R2-type pyocins is widespread in *P. aeruginosa* and are encoded in more than half of the strains in our collection (73%). In instances where the phage ΦPa39 adsorbs to the host, defense is observed in 14 out of 17 instances (82.4%), while R2-pyocin (-) strains, only 2 out of 7 instances show defense (28.6%), giving an increased resistance profile for R2-pyocin (+) strains by 2.7-fold (76.5/28.6) (**Table S1 & Figure 1c**). A similar trend is observed for other *Bruynoghevirus*es in the collection, including ΦPa5 and ΦPa29, which show an increase of 1.2 and 1.5-fold, respectively. In contrast, other phages with the podophage morphology in our collection are more effective at infecting R2-type pyocin (+) strains compared to R2-type pyocin (-) strains, these phages infect these strains at least 30% more often.

Discussion

How R2-type pyocins may confer phage defense remains unresolved. It is possible that the phage defense is conveyed by lysing the host before the infecting phage can replicate itself. Another possible explanation is that pyocins lyse nearby phage-infected cells and thereby terminating the phage infection. An argument against this hypothesis is that R2-type pyocins are not able to lyse other closely related bacteria under normal conditions (*154*). However, prior work has also shown that *P. aeruginosa* strains that are undergoing an SOS-response become sensitive to R2-type pyocins, even when these strains are closely related to each other (*154*). Notably, this same SOS-response is frequently triggered by phage infection, including during the infection of *Bruynoghevirus*es such as phage ΦPa39, which R2-type pyocins are predicted to act against (*155*). Furthermore, the SOS-response pathway also induces the production and release of pyocins themselves (*152*).

Together, we propose the hypothesis that R2-type pyocins are produced and released upon phage infection and provide phage defense through lysing the host and nearby bacteria that are phage-infected, while non-infected kin are unaffected (**Figure 1e**).

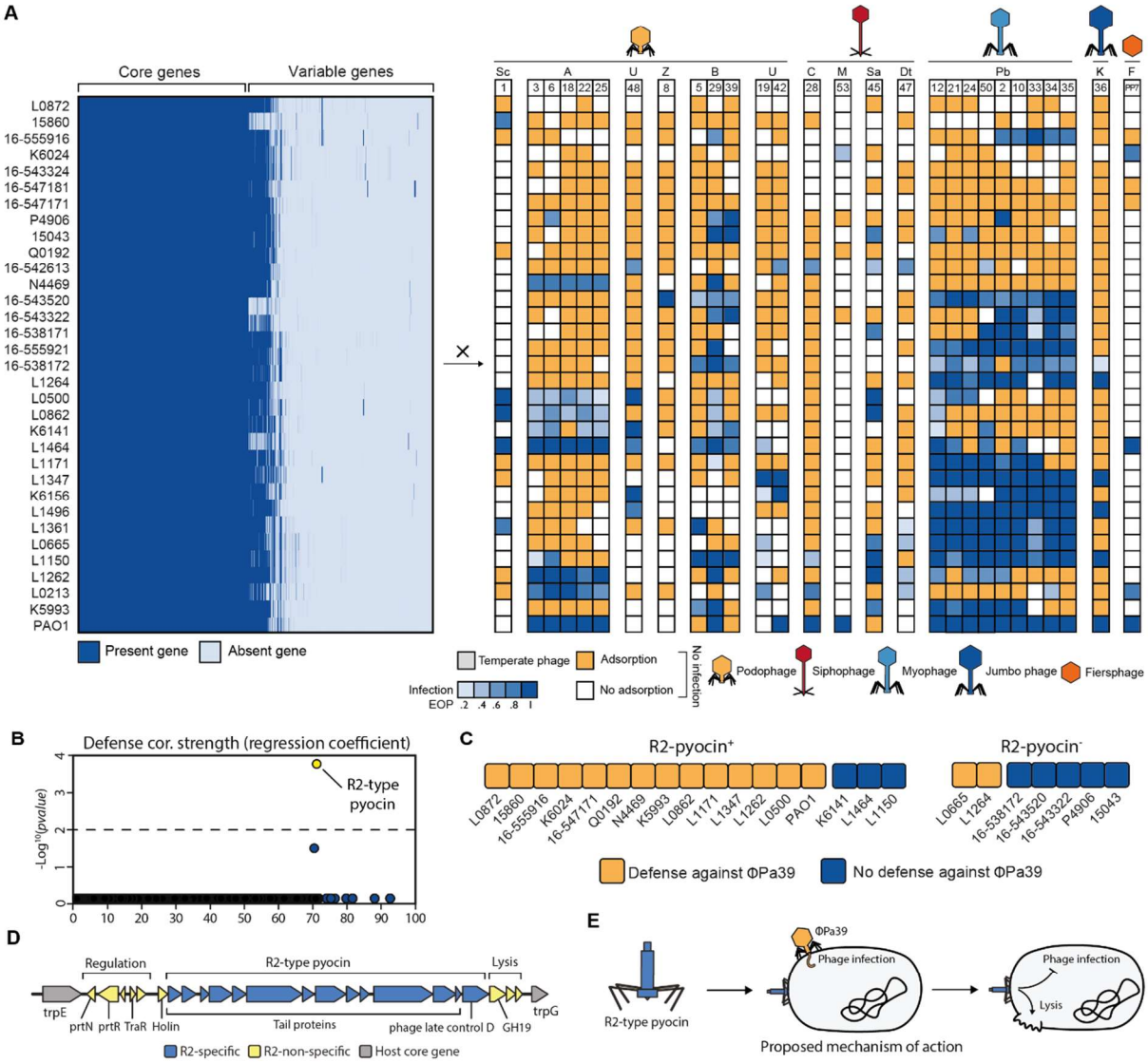

Figure 1. R2-type pyocins are associated with phage defense. (a) The gene presence and absence of 32 *P. aeruginosa* clinical isolates and strain *P. aeruginosa* strain PAO1. As well as the host range of their phages. Phage-bacteria interactions are depicted as infection (blue), adsorption (>50%) but no infection (orange), or no interaction (white). Letters above the phage numbers indicate family or genus (for phages unassigned to a family): A, *Autographiviridae*; B, *Bruynoghevirus*; C, *Casadabanvirus*; Dt, *Detrevirus*; F, Fiersviridae; K, *Phikzvirus*;

M, *Mesyanzhinovviridae*; Pb, *Pbunavirus*; Sa, *Samunavirus*; Sc, *Schitoviridae*; Z, *Zobellviridae*; U, unassigned. **(b)** Statistical results of the GWAS-analysis, depicting each variable gene within the *P. aeruginosa* collection based on its strength correlation with phage defense (Beta) with their statistical significance (-Log_{10} p-value). Significantly correlating genes are shown in yellow, while non-significant genes are shown in blue. **(c)** The phage sensitivity of R2-type pyocin (+) and (-) strains in respect to phage vB_PaeP_FBPa39. **(d)** The gene composition and neighbourhood of the R2-type pyocin in *P. aeruginosa* strain PAO1. Depicted are the R2-specific genes (blue) and its regulatory and lysis machinery (yellow), while also showing the core genes of the host (grey). **(e)** Our proposed mechanism of action that R2-type pyocins lyse phage infected cells, resulting in the termination of the phage infection.

Conclusion

We identified R2-type pyocins as a predicted phage defense system through genome-wide association studies (GWAS) using the phage-sensitivity profiles of 33 *P. aeruginosa* bacterial strains for 28 *Pseudomonas* phages. We propose the hypothesis that these R2-type pyocins are released in response to phage infection and provide phage defense to nearby bacteria by lysing phage-infected kin. Further experimental exploration is required to determine the precise involvement of R2-type pyocins in phage defense.

Materials and Methods

Genome-wide association study of phage defense-associated genes

All genome assemblies from the set of clinical isolates of *P. aeruginosa* used in the study of Costa *et al.* (2024) were downloaded (*146*). As well as the infectivity (EOP) data of each of the phages in the collection against these strains, which is present in the supplementary data.

To perform the GWAS, prokka v1.14.6 (*156*) was run on all the assemblies, followed by the creation of a presence/absence matrix and an alignment of the core genes using panaroo v1.5.1 (*148*). This alignment was then used to create a phylogenetic tree using IQ-Tree2 v2.4.0 (--alrt 1000, -B 1000 -m TEST) (*157*). To perform the association between the gene presence/absence matrix and the infectivity of the phages in the collection, pyseer v1.3.12 (*147*) was applied to estimate these effects using linear models, which took the phylogeny of the strains into account. The found associations were then adjusted for multiple testing using the Bonferroni correction. An adjusted p-value lower than 0.05 was seen as significant.

Quantification and statistical analysis

Unless stated otherwise, data are presented as the mean of biological triplicates ± standard deviation. A Bonferroni-adjusted p-value of less than 0.05 was considered significant.

Acknowledgements

This work was supported by grants from the European Research Council (ERC) CoG under the European Union's Horizon 2020 research and innovation program (grant agreement No. 101003229) to S.J.J.B. Lastly, we thank members of the Brouns lab and Dr. Véronique Ongenae for the many discussions and ideas that improved our work.

Supplementary information

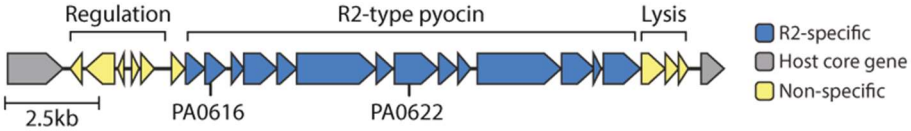

Figure S1. The R2-type pyocin genomic context. R2-type pyocin gene region. Shown are the 13 genes of R2-type pyocin that encode for the R2-type pyocin particle (blue), as well as the pyocin regulation and lysis genes (yellow). Lastly, the core genes of the host (grey) are indicated.

Gene ID	Annotation	p-value	adj. p-value	R^2	Beta	Beta-SE	MAF	Defense against
PA0615	Collar	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0616	Spike	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0617	Sheath initiator	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0618	Tri1	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0619	Tri2	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0620	Tail fibre	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0622	Sheath	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0623	Tube	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0624	Hypothetical	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0625	Tape measure	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0626	Ripcord	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0627	LysM (glue)	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39
PA0628	Hub	0,000148	0.0389	0,922	-84,4	12,5	0,5	vB_PaeP_FBPa39

Table S1. Genes that significantly correlate with phage defense. Shown are the 13 genes that that significantly (adjusted p-value <0.05) associate with phage defense, from a genome-wide association study (GWAS) that used the phage-sensitivity profiles of 33 *P. aeruginosa* bacterial strains for 28 *Pseudomonas* phages.

Chapter 4



Bacterial homologs of innate eukaryotic antiviral defenses with anti-phage activity

Prokaryotes have evolved a multitude of defense systems to protect against phage predation. Some of these resemble eukaryotic genes involved in antiviral responses. Here, we set out to systematically project the current knowledge of eukaryotic-like antiviral defense systems onto prokaryotic genomes using *Pseudomonas aeruginosa* as a model organism. Searching for phage defense systems related to innate antiviral genes from vertebrates and plants, we uncovered over 450 candidates. We validated six of these phage defense systems, including factors preventing viral attachment, R-loop-acting enzymes, the inflammasome, ubiquitin pathway, and pathogen recognition signalling. Collectively, these defense systems support the concept of deep evolutionary links and shared antiviral mechanisms between prokaryotes and eukaryotes.

A modified version of this chapter has been published as

van den Berg, D. F.*, Costa, A. R.*, Esser, J. Q.*, Stanciu, I., Geissler, J. Q., Zoumaro-Djayoon, A. D., Haas, P., & Brouns, S. J. J. (2024). Bacterial homologs of innate eukaryotic antiviral defenses with anti-phage activity highlight shared evolutionary roots of viral defenses.

Cell Host & Microbe, *32*(8), 1427-1443.

Introduction

Prokaryotes are subject to persistent predation by bacteriophages, driving the evolution of diverse host defense systems and phage-encoded countermeasures to evade these defense systems (*52*). This ongoing arms race relies on the rapid turnover of defensive and counter-defensive mechanisms (*40, 61*), and is facilitated by mobile genetic elements (MGEs) that encode the large majority of the known anti-phage defense mechanisms (*158*). These MGEs integrate into genetic hotspots within the genome, known as defense islands (*8, 61, 122, 146, 158-161*). Searching for new defense systems within these defense islands has significantly expanded our understanding of the prokaryotic immune repertoire, leading to the discovery of more than 100 defense mechanisms in the past six years (*10, 122, 141, 159, 160*).

A few of these defense mechanisms exhibit a striking resemblance to eukaryotic genes involved in antiviral response, such as Toll/interleukin-1 receptor (TIR), Argonaute, Gasdermin, cyclic GMP-AMP synthase (cGAS), and Dynamin-like (Mx) proteins (*38, 159, 160, 162-167*). These findings highlight shared evolutionary strategies in the context of the perpetual struggle against viral pathogens, bridging the gap between prokaryotic and eukaryotic immune systems. Further investigating the link between prokaryotic and eukaryotic immune systems can provide valuable insights into the evolutionary origins and underlying mechanisms of eukaryotic immune responses.

Motivated by these factors, we set out to systematically project the current knowledge of eukaryotic-like antiviral defense systems onto prokaryotic genomes using *Pseudomonas aeruginosa* as a model organism, a bacterium known for its exceptional diversity and abundance of defense systems (*63, 146, 161*). Here, we report the first instance of inositol-monophosphate phosphatase Hermes (IMPase), a plant-tolerance-like factor, which provides anti-phage defense via modifications of the cell surface that prevent viral adsorption. We also report a DNA replication helicase/nuclease 2 (DNA2)-containing anti-phage defense protein (Prometheus) that is similar to eukaryotic antiviral R-loop acting enzymes, important components of the innate immune system of various eukaryotes (*168-179*). In addition, we discovered four new eukaryotic-like anti-phage defense systems, two of which are like antiviral inflammasome components featuring a new clade of NACHT domain-containing proteins with a distinct architecture and two novel anti-phage effectors: NucS and SfsA (bNACHT Erebus and bNACHT Hypnos). Moreover, we discovered another new defense system that contains eukaryotic-like ubiquitin-related components, consisting of a fused E1-E2-JAB protein combined with a metallo-beta-lactamase fold (MBL-fold) protein (6A-MBL). This defense system is a distant homolog of the pathogen receptor signalling of Toll-like receptors and introduces a new subtype of the TIR-containing Thoeris anti-phage defense family. Thoeris type III is characterized by the presence of a ThsB-like protein with a novel NucS endonuclease effector domain and an uncharacterized SLOG domain protein from a lineage not previously linked to Thoeris systems.

These newly identified defense systems can be found in up to 11 bacterial phyla, being the most prevalent among Proteobacteria, and enhance our understanding of the complex connection between bacterial and eukaryotic immune systems.

Results

Eukaryotic-like antiviral homologs show anti-phage activity

To identify new eukaryotic-like defense systems in *P. aeruginosa*, we used the entire set of experimentally verified protein sequences associated with eukaryotic innate immunity from the InnateDB (*180*) as our starting database (n = 10,139), including *Homo sapiens* (n = 7,886), *Mus musculus* (n = 2,100), and *Bos taurus* (n = 153) (**Table S1**). Additionally, we applied DRAGO3 (*181*) to search for plant pathogen recognition proteins and a custom literature-based list of functional domains associated with plant and vertebrate innate immunity (**Table S1**). We used this starting set of proteins as a seed to identify all eukaryotic homologs included in Eukprot (*182*) with the use of MMseqs2 (easy-cluster) (*183*) and subsequently built an HMM protein signature from an average of 218 sequences for

each validated innate immune protein. By applying these signatures to all representative proteins within regions of genomic plasticity of *P. aeruginosa*, we identified 757 homologs, indicating that 11.2% of the representative genes within the overall pangenome of *P. aeruginosa* exhibit similarity to eukaryotic proteins associated with innate immunity (**Figure 1a**, **Table S2**).

To narrow down the selection of these eukaryotic homologs for testing possible antiphage activity, we focused on gene clusters that occur in the proximity of known defense systems. The resulting 457 homologs contained a large range of functional domains, including the most abundant type III restriction enzyme (RESIII) and DEAD/DEAH box helicase (DEAD). These two functional domains are also present within the most notable nucleotide-acting antiviral defenses within eukaryotes, including RIG-I and DICER (*184, 185*). To assess the likelihood of these functional domains being associated with anti-phage activity, we determined if they were enriched in defense islands compared to other genomic regions. This was achieved by computing the log-fold change (log2FC) in the prevalence of the domain within proximity to known defense systems (< 0.1 Mb) compared to its prevalence at greater distances (> 0.1 Mb) (**Figure 1a**). The 0.1 Mb threshold was determined based on the distances observed between known defense systems within defense islands of *P. aeruginosa*, which align with those previously reported in *E. coli* (*158*). Among the eukaryotic antiviral associated functional domains that are most enriched in phage defense islands were several already known to be shared by bacterial and eukaryotic antiviral strategies. These included components involved in pathogen recognition, such as signal transduction ATPases with numerous domain (STAND) proteins, and cGAS/DncV-like nucleotidyltransferases (CD-NTases) involved in the eukaryotic antiviral pathway of cGAS (*186-188*).

We then set out to experimentally test the anti-phage activity of eleven conserved gene clusters that contain at least one of the most enriched functional domains, including DEAD, RESIII, ATPase, Helicase, NACHT, Ubiquitin-activating E1, TIR, and the less common IMPase, by introducing the gene clusters with their native promoters into *P. aeruginosa* strain PAO1 on a low-copy plasmid (*146*) (**Figure 1b**, **Table 1**). These strains were subsequently challenged with a set of representative *P. aeruginosa* phages from six taxonomic groups (**Figure 1a**) (*146, 189*), revealing six new eukaryotic-like defense systems. Remarkably, all except one (Prometheus) demonstrated a complete inhibition (more than 10^6-fold) of phage propagation for at least one phage, including Hermes, bNACHT Erebus, bNACHT Hypnos, 6A-MBL, and Thoeris III (**Figure 1b**). Together, these findings confirm the discovery of six new defense systems with eukaryotic-like antiviral functional domains. We will discuss these in more detail in the subsequent sections.

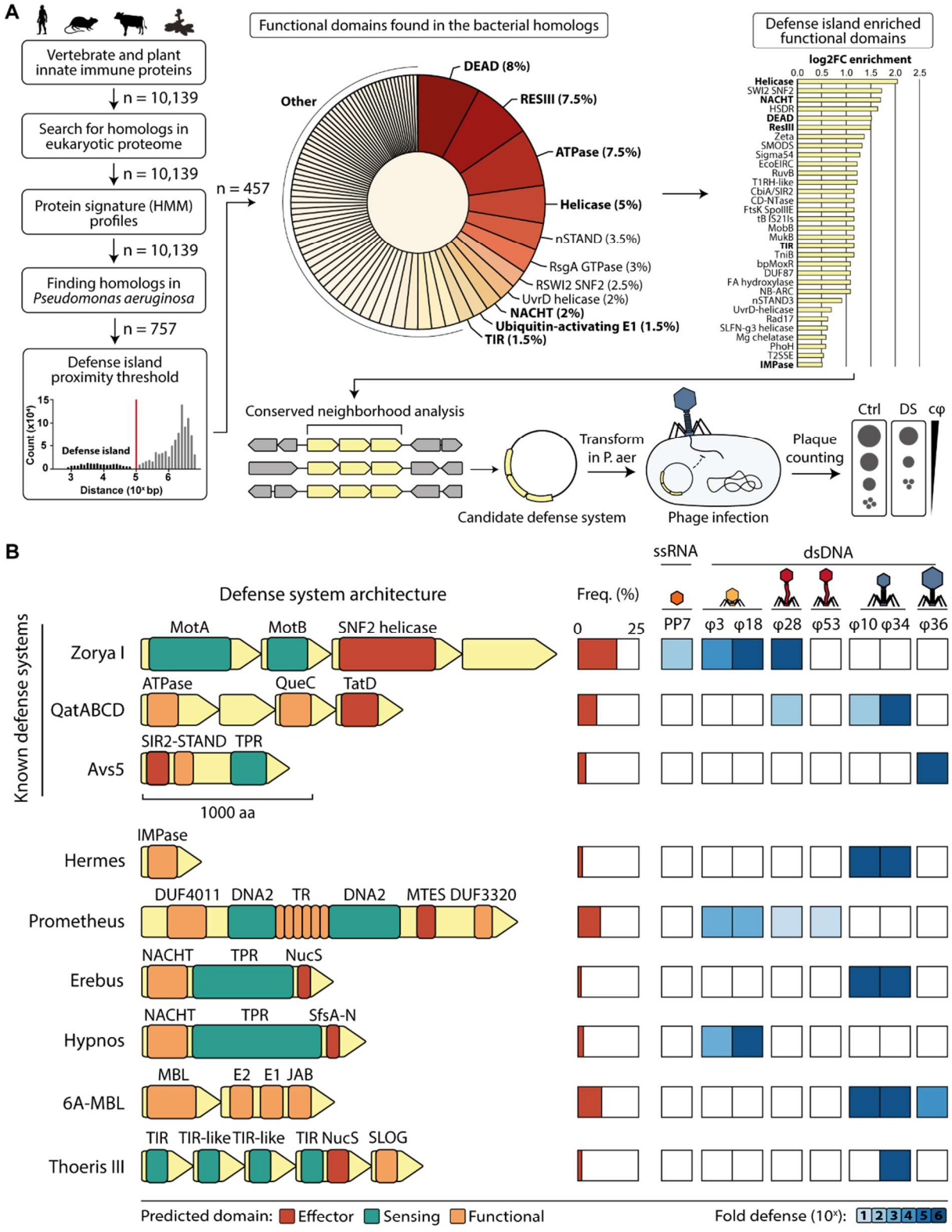

Figure 1. Anti-phage activity of defense systems identified in *P. aeruginosa* that are homologous to eukaryotic antiviral immunity. (a) Strategy used for the identification of bacterial homologs of proteins involved in eukaryotic innate immunity, from vertebrates (*Homo sapiens*, *Mus musculus*, and *Bos taurus*; InnateDB) and plants (363 species; PRGdb) that are in proximity of known defense systems. Shown are the functional domains found within these homologs, in addition to their prevalence in the *P. aeruginosa* homolog set in comparison to homologs more distant from defense systems (log2FC). A subset of conserved gene clusters containing these enriched functional domains were selected for subsequent assessment of their potential antiviral activity (in bold). The candidate defense systems were cloned with their native promoters into pUCP20, and then introduced into the *P. aeruginosa* strain PAO1. The anti-phage activity of the candidate defense systems was assessed using efficiency of plating assays. **(b)** The anti-phage activity of the defense systems against a panel of eight phages from six

distinct taxonomic groups, measured by efficiency of plating assays. Previously validated defense systems (*146*) were included as controls. The defense system architecture shows functional domains, color-coded based on their predicted function (**Table 1**). Genes are drawn to scale, with the scale bar representing 1000 amino acids. The bar graph shows the abundance of the defense systems in *P. aeruginosa* genomes of RefSeq.

Table 1. Genes and functional domains of the eukaryotic-like defense systems

Defense system	Gene	Functional domain(s)	COG/pfam/CDD	Best eukaryotic hit
Hermes	*hrsA*	IMPase[a]	PF00459 cl00289	IMPA1/2
Prometheus	*proA*	DUF4011, DNA2[a,b], tandem repeat (TR)[c], REase_MTES, DUF3320	PF13195 PF18741 COG1112 PF11784	MOV-10 RNA helicase
bNACHT Erebus	*eruA*	NACHT[a,b], TPR[a,c], NucS endonuclease[d]	cl26020 PF05729 PF13176 PF01939	NLRC4 inflammasome
bNACHT Hypnos	*hyoA*	NACHT[a,b], TPR[a,c], SfsA-N DNA-binding[d]	cl26020 PF05729 PF13176 PF17746	NLRC4 inflammasome
6A-MBL	*mblB*	ComA-like MBL-fold	cd07731	
	cap2-3	E2[a], E1[a], JAB[a]	PF14457 cl37499 PF14464	UBE2E1
Thoeris type III	*thcB1*	Cap12-like TIR[a]	PF01582	SARM1
	thcB2	TIR-like DUF1863	PF08937	
	thcB3	TIR-like DUF1863	PF08937	
	thcB4	ThsB-like TIR[a,b], NucS endonuclease[d]	PF01582	SARM1 TLR adaptor
	thcA	SLOG	PF18178	

[a] Functional domain associated with eukaryotic antiviral response.

[b] Functional domain is present in the eukaryotic homolog.

[c] Predicted with HHrepID.

[d] Predicted with Foldseek or DALI.

Hermes is a homolog of eukaryotic immune IMPases that prevents phage adsorption

Hermes was named after the Greek deity that functions as the messenger of the gods and consists of one gene with an inositol monophosphatase (IMPase) functional domain (**Figure 2a, Table 1**). Hermes provides complete protection against myophages ϕPa10 and ϕPa34 from the *Pbunavirus* genus, causing a reduction in phage infectivity of at least 10^6-fold in efficiency of plating assays (**Figure 1b**), reducing phage propagation in liquid cultures by 10^2-fold (**Figure 2b**, **Figure S1a**), and allowing an almost complete recovery of bacterial growth under phage predation (**Figure 2c**). A point mutation in a conserved aspartic acid residue (D89A) in the predicted IMPase active site completely abolished phage protection (**Figure 2d**). Hermes shows strong similarity to eukaryotic IMPases (IMPA) involved in innate antiviral responses, both structurally (DALI Z-Score: 32.8) and at the sequence level (20-30% pident) (**Figure S1b,c**). Hermes also shares resemblance with the bacterial IMPase SuhB, a constituent of the core genome of *P. aeruginosa* (*190, 191*) (31% pident; DALI Z-score: 30.0) (**Figure S1b**). Similar to SuhB and IMPA, Hermes is predicted to form a dimer structure (**Figure S1d**). Noteworthy, phylogenetic analyses suggest that Hermes is the closest bacterial relative to the eukaryotic IMPase family (**Figure 2e**) (*192-197*). Eukaryotic IMPases typically function by dephosphorylating inositol monophosphate to produce myo-inositol, which serves as the precursor for key signalling molecules in several cellular processes, including apoptosis, stress tolerance, cAMP production, and cell growth (*192-194*). In addition, eukaryotic IMPases provide viral defense through affecting viral attachment and interfere with viral packaging by altering lipid composition and expression of viral receptors (*198*). In bacteria and archaea, inositol-containing molecules have been linked to osmotic balance (*199*), capsule expression (*200*), and the biosynthesis of membrane phosphatidylinositol, which can be modified to anchor proteins or complex carbohydrates to cell surfaces (*201*). Given the role of the IMPase family in surface alterations in both eukaryotes and prokaryotes, we hypothesized that Hermes-mediated phage defense might involve surface modifications preventing phage adsorption. Further investigations revealed that *Pbunavirus* ϕPa34 completely failed to adsorb to Hermes cells, while *PhiKzvirus* ϕPa36 adsorbed similarly to control and Hermes cells (**Figure 2f**).

Analysis of cell surface components, including capsule (CPS, **Figure 2g**), lipopolysaccharides (LPS, **Figure 2h**), and outer membrane proteins (**Figure S1e**), demonstrated alterations at all levels in cells expressing Hermes. Because *Pbunavirus* have been shown to use LPS as their receptor (*202, 203*), we hypothesized that, modifications in LPS could account for the observed inability of ϕPa34 to adsorb to Hermes cells. Strikingly, while ϕPa34 successfully adsorbed to LPS extracted from control and Hermes D89A cells (which also contained CPS), no adsorption was observed for LPS derived from Hermes-expressing cells (**Figure 2i**), supporting our hypothesis that Hermes provides phage defense through LPS and/or CPS alterations at the cell surface via IMPase-dependent pathways.

To gain insights into the substrate specificity of Hermes, we inspected the binding pocket and compared the conserved residues among Hermes, IMPA1, and SuhB (**Figure S1b**). We observed that Ala90 in the active site of Hermes differs from the corresponding Thr95 in IMPA1 and Thr109 in SuhB, which are critical residues for sequestering magnesium ions that catalyse the hydrolysis of inositol monophosphate into myo-inositol (*196*). This variation suggests that Hermes may target a related but distinct substrate, a hypothesis further supported by mass spectrometry analysis, which showed identical levels of inositol monophosphate and myo-inositol in both Hermes-expressing and control cells (**Figure S1f**).

Collectively, our findings reveal a conserved mechanism of antiviral defense by IMPase-relatives in both prokaryotes and eukaryotes, which prevents viral adsorption through alterations of the cell surface (**Figure 2j**).

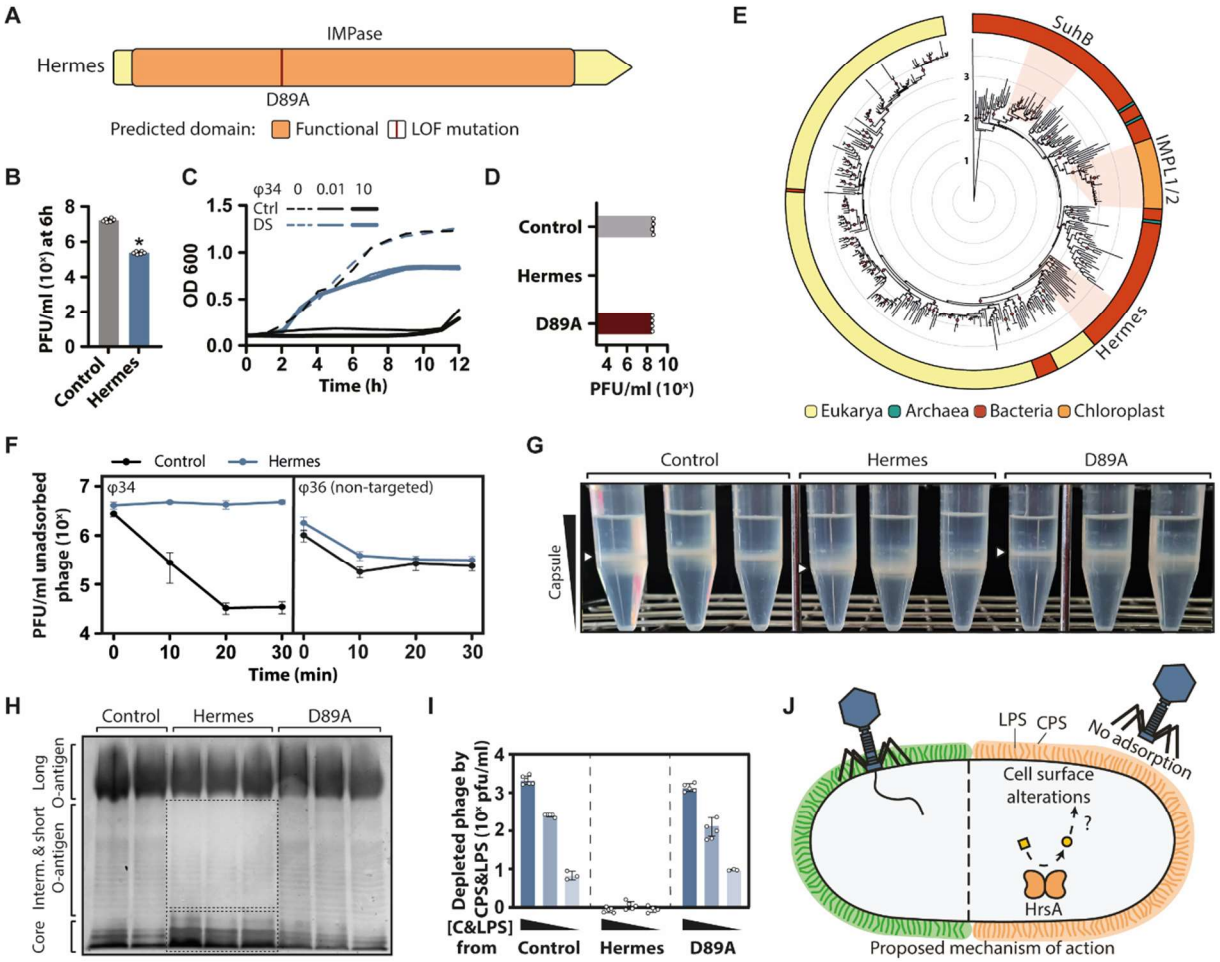

Figure 2. Hermes is an anti-phage defense system with strong similarities to eukaryotic antiviral IMPase proteins by preventing phage adsorption. (a) The functional domains of Hermes and mutation sites tested in (d). LOF, loss of function. **(b)** Impact of Hermes on phage propagation in liquid culture. The propagation of phage ϕPa34 was monitored over time and the 6h time point is shown here, with all data points displayed in Figure S1A. The control bar represents phage propagation in PAO1 containing an empty plasmid. **(c)** Effect of the defense system on bacterial growth upon phage infection. PAO1 cells containing an empty plasmid (control) or Hermes (defense system, DS) were infected with phage ϕPa34 at low (0.01) and high (10) multiplicity of infection and their growth was monitored for a period of 12h. **(d)** Effect of mutations in the functional domains of the defense system on phage protection. The infectivity of phage ϕPa34 on PAO1 cells containing an empty plasmid (control), Hermes, or Hermes with point mutations was assessed by plaque assay. **(e)** Phylogenetic tree of IMPase-containing proteins. The phylogenetic tree of 366 representative proteins was inferred and bootstrapped using IQ-Tree2, and rooted with a fructose-1,6-bisphosphatase class I protein (FBPase class I; NCBI accession: CEI80039.1) from *P. aeruginosa.* The clades of human IMPA1 and IMPA2, plant IMPL1 and IMPL2, and prokaryotic HrsA and SuhB are indicated in the tree. **(f)** Adsorption of phage ϕPa34 to cells harboring either an empty plasmid (control) or Hermes. The values represent the concentration of unbound phage during a 30-minute incubation period. Adsorption of phage ϕPa36, which is not targeted by Hermes, serves as a control. **(g)** Capsule amount produced by cells harboring empty plasmid (control), Hermes, or Hermes with a point mutation (D89A). The capsule amount was determined using Percoll density gradients (40, 60, 80%). The location of the capsule band is highlighted with a white triangle. **(h)** Lipopolysaccharide (LPS) fractions extracted from cells expressing empty plasmid (control), Hermes, or Hermes with a point mutation (D89A). LPS samples were separated on a 12% SDS-PAGE gel and stained with SYPRO Ruby. Regions exhibiting band intensity differences in Hermes cells compared to control and D89A cells are indicated by dashed rectangles. **(i)** Adsorption of phage to lipopolysaccharide (LPS) extracts containing also capsular polysaccharides (CPS) from cells carrying an empty plasmid (control), Hermes, or Hermes with an LOF point mutation (D89A). The values indicate the concentration of phage bound to decreasing amounts of C/LPS from each strain. **(j)** Proposed model for anti-phage defense by Hermes (HrsA). Hermes induces changes in the surface structures of *P. aeruginosa*, such as capsular (CPS) and lipopolysaccharides (LPS), leading to the prevention of phage adsorption

Prometheus is a bacterial homolog of human DNA2 antiviral defense genes

Prometheus, named after the Greek titan associated with the creation and protection of humanity, is part of the bacterial DNA2-like enzyme (Bad) family (*204*) which also includes anti-phage genes hhe

(*122*), Mokosh MkoA, and Mokosh MkoC (*160*). Prometheus is composed of one very large open reading frame encoding a protein of 2,221 amino acids with four predicted functional domains: DUF4011, DUF3320, DNA2 helicase/nuclease and REase MTES domains (**Figure 3a**, **Table 1**).

The DNA2 domain is shared with several human antiviral genes, including NFX1-type zinc finger-containing protein 1 (ZNFX1), Senataxin (SETX), DNA-binding protein SMUBP-2 (IGHMBP2), helicase with zinc finger 2 (HELZ2), Moloney leukemia virus 10 homolog (MOV-10), and UP Frameshift 1 helicase (Upf1) (**Figure 3a,b**, **Figure S2a**) (*168, 172-179*). These human genes share a DNA2 and (ribo)nuclease domain with Prometheus, and provide protection against various viruses, such as Epstein-Bar virus, Influenza, West Nile virus, and HIV-1 (*168, 172-179*). Interestingly, in both Prometheus and its eukaryotic homologs, the DNA2 domain is split into two parts by tandem repeats, although the function of this arrangement is unknown (**Figure 3a**). The antiviral human homologs of Prometheus regulate viral transcription through a R-loop dependent mechanism. Targeting of these RNA-DNA hybrids with a displaced single-stranded DNA strand (*205*), which are formed during transcription (*206*), often results in transcription termination or attenuation (*207*).

We hypothesized that Prometheus exerts its antiviral response by interfering with phage transcripts through action on R-loops formed during transcription. Testing this hypothesis, we observed that Prometheus provides protection against podophages ϕPa3 and ϕPa18 from the *Autographiviridae* family, siphophage ϕPa28 from the *Casadabanvirus* genus, and siphophage ϕPa53 from the *Mesyanzhinovviridae* family, as observed in efficiency of plating assays (**Figure 1b**), liquid cultures (**Figure 3c**, **Figure S2b**) and bacterial growth experiments during phage predation (**Figure 3d**). Interestingly, these phages contain 3 to 12 transcribed DNA regions prone to R-loop formation (*208*), whereas phages unaffected by Prometheus have no predicted R-loops (ϕPa36, ϕPa10, and ϕPa34) (**Figure 3e**, **Figure S2c**). We noticed that Prometheus homolog hhe protected against phage lambda with one predicted R-loop and was unable to protect against phages lacking R-loops, such as T2, T4, T5, P1, and M13(*122*) (**Table S3**), while Mokosh can protect more broadly (*160*).

We found that mutations of conserved residues in any of the four domains (DUF4011, DNA2, Rease MTES, DUF3320) abrogated protection, indicating that all domains are essential for Prometheus function (**Figure 3f**). Further investigations using RNA-seq to analyse transcription profiles of phage ϕPa18 in control versus Prometheus cells, revealed a 10-fold decrease in phage overall transcription (**Figure S2d**). This decrease was also evident in the phage genomic abundance in the presence of Prometheus, suggesting suppression of phage replication (**Figure S2d**). To understand the suppression of replication we further investigated the domain organization of Prometheus. The REase MTES domain belongs to the Swt1-like protein family found in several eukaryotic groups (*209*). In eukaryotes, Swt1 proteins act as RNA endonucleases, playing a role in the quality control of nuclear mRNA export, a crucial step in eukaryotic gene expression (*210*). However, research on a bacterial Prometheus homolog in *Geobacillus stearothermophilus* indicates that the protein functions as a single-stranded DNA (ssDNA) endonuclease (*204*). This suggests that Prometheus might act as a ssDNA endonuclease on the displaced ssDNA of the R-loop, through the actions of the DNA2 nuclease and/or REase MTES domains. The activity of the DNA2 helicase/nuclease domain may be facilitated by the putative helicase-related domains DUF4011 and DUF3320(*211*) (**Figure 3f**). Notably, DUF3320 is absent in 54% of instances of Prometheus, even though mutagenesis of this domain (as well as of DUF4011) results in inactivation of the system (**Figure 3f**). Structural predictions indicate that the DUF3320 domain in Prometheus is like a DNA-binding fork head domain (**Figure S2e**; DALI Z-score: 6.3), suggesting a potential role in facilitating the helicase or R-loop binding of DNA2 in Prometheus.

In summary, our analysis highlights Prometheus as a defense system with similarities to eukaryotic antiviral proteins containing DNA2 helicase/nuclease domains. This resemblance suggests a shared mechanism for countering viral infections by acting on viral R-loop structures formed during transcription (hypothetical model proposed in **Figure 3g**).

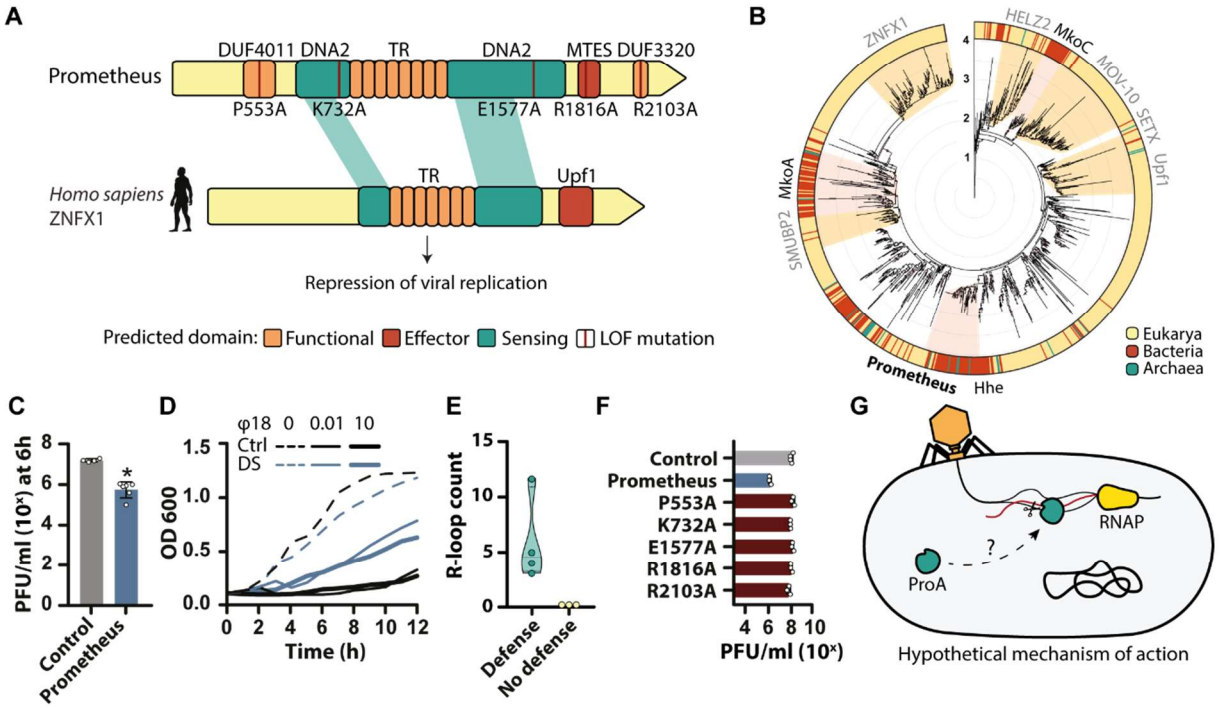

Figure 3. Prometheus is a homolog of a multitude of human DNA2 antiviral defense genes. (a) The functional domains of Prometheus, mutation sites tested in (f), and sequence similarity to the antiviral eukaryotic protein ZNFX1 from *Homo sapiens* (NP_066363.1). LOF, loss of function. **(b)** Phylogenetic tree of DNA2-containing proteins. The phylogenetic tree of 1,124 representative proteins was inferred and bootstrapped using IQ-Tree2. Branches with bootstrap confidence interval ≥90% are indicated with red dots. The phylogenetic tree was rooted with the ATP-dependent DNA helicase RecG protein (NCBI accession: NP_254032.1) from *P. aeruginosa* PAO1. The clades of human MOV-10, HELZ2, Upf1, ZNFX1, SETX, and SMUBP-2, prokaryotic Prometheus ProA, and DNA2-containing phage defense systems hhe and Mokosh type I/II (MkoA and MkoC) are indicated in the tree. **(c)** Impact of Prometheus on phage propagation in liquid culture. The propagation of phage ϕPa18 was monitored over time, and the 6h time point is shown here, with all data points displayed in **Figure S2B**. The control bar represents phage propagation in PAO1 containing an empty plasmid. **(d)** Effect of the defense system on bacterial growth upon phage infection. PAO1 cells containing an empty plasmid (control) or Prometheus (defense system, DS) were infected with phage ϕPa18 at low (0.01) and high (10) multiplicity of infection, and their growth was monitored for a period of 12h. **(e)** A violin plot comparing the number of predicted R-loops in targeted and non-targeted phages. **(f)** Effect of mutations in the functional domains of Prometheus on phage protection. The concentration of phage ϕPa18 when spotted on PAO1 cells containing an empty plasmid (control), Prometheus, or Prometheus with LOF point mutations, as assessed by plaque assay. **(g)** Prometheus is hypothesized to interfere with phage nucleic acids during transcription.

Erebus and Hypnos contain NACHT-domains

bNACHT Erebus and bNACHT Hypnos are named after the Greek god personification of darkness (Erebus) and his nephew Hypnos, the Greek god of sleep. Both anti-phage systems show similarities with eukaryotic antiviral inflammasome components, representing NACHT domain-containing NLR proteins (nucleotide-binding domain leucine-rich repeat containing, also known as nucleotide oligomerization domain (NOD)-like receptors). NLRs are part of the STAND superfamily and function in antiviral activities within almost all domains of life (*21*) (**Figure 4a**, **Table 1**). Of these eukaryotic NLRs, Erebus and Hypnos are most like intracellular-pathogen NLR family CARD domain-containing protein 4 (NLRC4). NLRC4 contains several functional domains, including NACHT, an N-terminal caspase recruitment domain (CARD), and a C-terminal leucine-rich repeat (LRR) (*212*). Erebus and Hypnos share common features with NLRC4 such as the presence of the NACHT domain, but also a protein repeat structure: the TPR domain. In addition, Erebus and Hypnos contain DNA-acting functional domains not previously associated with phage defense, instead of the CARD domain in NLRC4.

Hypnos and Erebus provide robust protection against myophages ϕPa10 and ϕPa34 from the *Pbunavirus* genus (Erebus), and podophages ϕPa3 and ϕPa18 from the *Autographiviridae* family (Hypnos), resulting

in a reduction of phage infectivity by at least 10^4-fold in efficiency of plating assays (**Figure 1b**). Both defense systems limit phage propagation (ϕPa34 by Erebus, ϕPa18 by Hypnos) in liquid culture (**Figure 4b**, **Figure S3a**). While Erebus allows full recovery of bacterial growth during phage infection (**Figure 4c**), Hypnos provides a benefit in bacterial growth approximately 2h post-infection (**Figure 4c**). Point mutations of conserved amino acids in the NACHT Walker A ATP binding pocket of both defense systems (K50A in Erebus and K52A in Hypnos) resulted in the complete loss of phage protection (**Figure 4d**).

In humans, plants, and fungi, NACHT-containing NLRs are widely involved in the recruitment of the inflammasome during the innate immune response upon sensing pathogen-derived biomolecules (*213*). Although some NACHT-containing NLRs are found to be activated upon sensing viral dsRNA (*214*), the vast majority sense pathogen-derived proteins (*215-217*). Recently, some clades of prokaryotic NACHT domain containing NLRs have been found to provide phage defense (*218*). Here, we show the first instance of anti-phage activity by NACHT proteins from clade 20 (**Figure 4e**), Erebus and Hypnos, characterized by an N-terminal NACHT domain, a central tetratricopeptide repeat (TPR) domain, and novel effector domains NucS and SfsA-N at their C-terminus (**Figure 4a**,**f**). This organization is distinct from most NACHT domain-containing NLRs, which typically display an N-terminal effector instead (*218*). Moreover, Erebus and Hypnos encode two TPR regions linked by a helical bridge (**Figure 4f**), and mutagenesis of this bridge resulted in the loss of phage defense (R740A in Erebus and R817A in Hypnos, **Figure 4d**).

In other NACHT-like defense systems, the NACHT-like domain physically associates with the effector domain to prevent its activity. This inhibition is lost upon TPR-mediated sensing of phage infection, resulting in protein multimerization (*21*), and activation of the effector domain. In line with these observations, the effector domains of both Erebus and Hypnos are structurally predicted to be near the NACHT domain, suggesting that the NACHT domain may prevent the activity of the effector domain until the target is sensed, similar to the role of NACHT in eukaryotes (*217*) (**Figure 4f**, **Figure S3b,c**). It remains unclear if multimerization is required for the activation of the effector domains (*219, 220*), since these domains and this unusual domain organization has not been characterized before. However, the effector domain of Erebus, NucS, is known for binding and cleaving ssDNA extremities of branched DNA (*221*), while the oligonucleotide/oligosaccharide-binding (OB) fold of SfsA-N in Hypnos likely provides non-specific DNA-binding or cleaving capabilities (*220, 222*).

Collectively, our findings reveal the anti-phage activity of two new NACHT-containing NLR proteins, Erebus and Hypnos, with strong similarity to eukaryotic NACHT-containing antiviral proteins. Based on this similarity, we propose that upon sensing the presence of phage proteins by the bridge-linked TPR domains, the NACHT domain undergoes a conformational change, activating the effector domain (NucS or SfsA-N) and triggering the initiation of the antiviral response, possibly through DNA cleaving or DNA repressing mechanisms (hypothetical model proposed in **Figure 4g**).

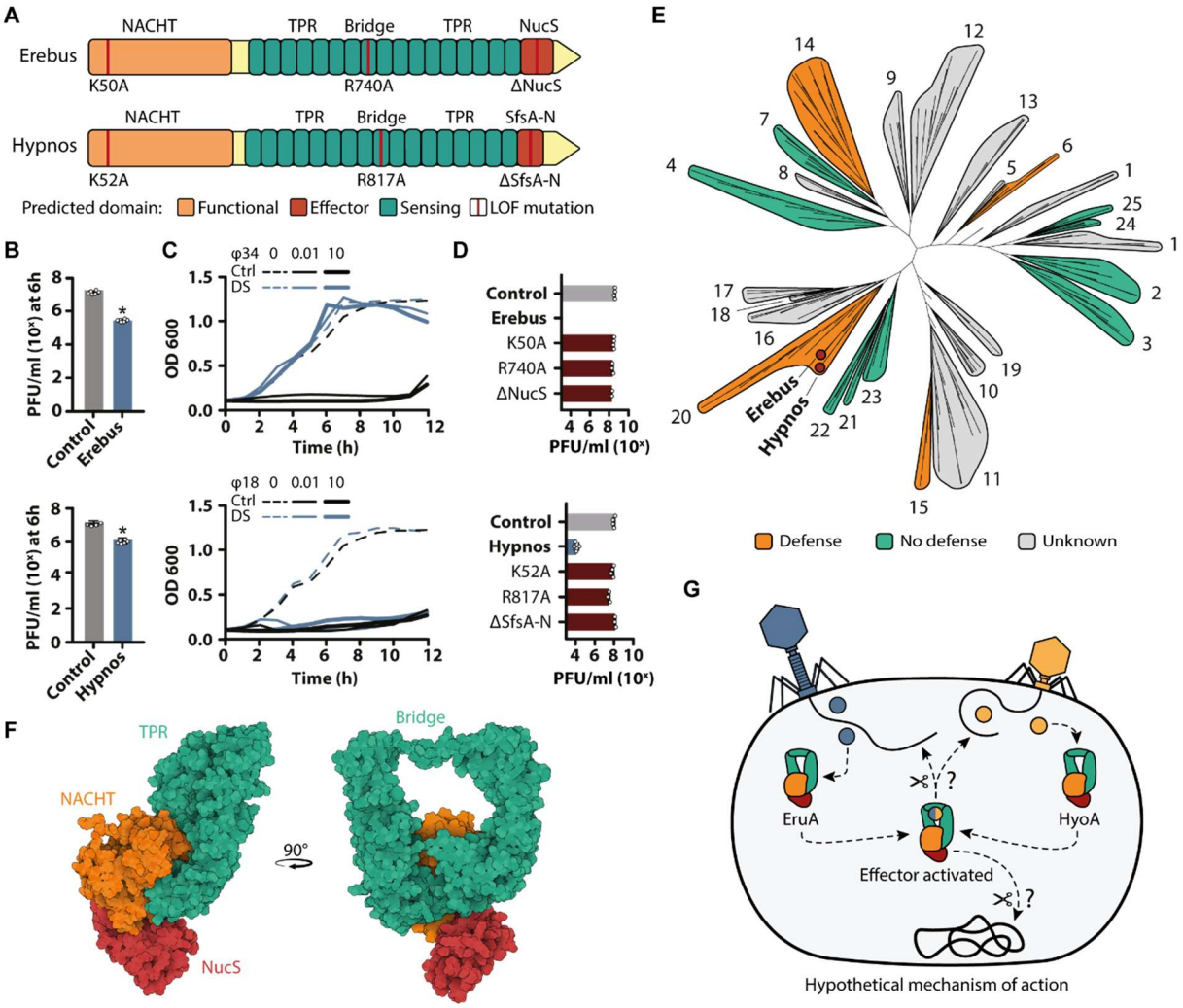

Figure 4 Erebus and Hypnos are NACHT-containing anti-phage NLRs with strong homology to eukaryotic antiviral NLRs. (a) The functional domains of Erebus and Hypnos and mutation sites tested in (d). LOF, loss of function. **(b)** Impact of Erebus and Hypnos on phage propagation in liquid culture. The propagation of phage ϕPa34 for Erebus and ϕPa18 for Hypnos was monitored over time, and the 6h time point is shown here, with all data points displayed in **Figure S3A**. The control bar represents phage propagation in PAO1 containing an empty plasmid. **(c)** Effect of Erebus (top) or Hypnos (bottom) on bacterial growth upon phage infection. PAO1 cells containing an empty plasmid (control) or the defense system (DS) were infected with phage ϕPa34 for Erebus and ϕPa18 for Hypnos at low (0.01) and high (10) multiplicity of infection and their growth was monitored for a period of 12h. **(d)** Effect of mutations in the functional domains of Erebus and Hypnos on phage protection. The infectivity of phage ϕPa34 for Erebus and ϕPa18 for Hypnos on PAO1 cells containing an empty plasmid (control), defense system, or defense system with point mutations was measured by plaque assay. **(e)** Phylogenetic tree of all bacterial NACHT-domain containing NLRs. The phylogenetic tree was built from the proteins provided by (*218*) (n = 3247) and inferred using FastTree. The clades are colored based on their antiviral properties. **(f)** Tertiary structure of Erebus predicted by Alphafold2. The functional domains are color-coded, with TPR shown in green, NACHT in orange, and the effector domain (NucS) in red. **(g)** Hypothetical model for anti-phage defense by Erebus (EruA) and Hypnos (HyoA), with only Erebus shown. Erebus and Hypnos are hypothesized to sense phage infection using the TPR domains, causing a conformational change in the NACHT domain that leads to the release of the effector domain. The effector domains will likely initiate the antiviral response via DNA acting mechanisms.

6A-MBL contains a ubiquitination-like domain and an MBL-fold protein

6A-MBL contains one gene with a fused E1-E2-JAB functional domain (Cap2-3), and a second gene with an MBL-fold functional domain (MblB) (*223*) (**Figure 5a**, **Table 1**). In eukaryotes, E1, E2, and JAB (DUB) function in the ubiquitin signalling pathway, which fine-tunes the eukaryotic innate immunity by either modulating the stability of key molecules or by regulating cytokine production (*224*). In bacteria, these functional domains provide phage defense in CBASS type II systems via a ubiquitin-like mechanism

termed cGASylation (*225*). The E1-E2 Cap2 protein of CBASS type II is responsible for cGAS conjugation of CD-NTase to a target molecule upon viral infection. JAB Cap3 functions as a regulator of cGASylation by cleaving the cGAS conjugates (*225*). Previously identified phage defense systems had E1-E2 and JAB encoded by separate genes (*160, 225*), but in 6A-MBL these are fused. Here, we show that this configuration provides strong protection against myophages ϕPa10 and ϕPa34 of the *Pbunavirus* genus, and jumbo myophage ϕPa36 of the *Phikzvirus* genus, resulting in a reduction of phage infectivity by at least 10^4-fold in efficiency of plating assays (**Figure 1b**). 6A-MBL limits the propagation of phage ϕPa34 in liquid cultures (**Figure 5b**, **Figure S4a**) and allows full recovery of bacterial growth during phage infection (**Figure 5c**). Point mutations in conserved amino acids of each predicted functional domain resulted in the complete loss of phage protection (**Figure 5d**).

Based on previous studies on E1-E2 and JAB components of antiviral responses in both eukaryotes and prokaryotes (*224, 225*), we hypothesize that upon phage infection, the E1-E2 and JAB domains of 6A-MBL prime the MBL-fold protein for activation through a ubiquitin-like mechanism. This may be aided by the predicted complex formation between the two proteins, as predicted by Alphafold. The complex consists of a dimer of Cap2-3 connected by the E1 domains. The E2 and JAB domains of Cap2-3 interact with MblB, forming a heterodimeric complex (**Figure S4b**). In addition, we observed a conserved glycine (G395) in the C-terminal region of MblB in proximity of the active site of the JAB domain of Cap2-3(*225*) (**Figure 5a**). These conserved glycine residues at the C-terminal region often serve as the conjugation site in other ubiquitin-like phage defense systems, such as CBASS type II (*225-228*). In CBASS type II, conjugation of the cyclase to its target increases signal molecule production activating downstream anti-phage effectors. Here, we suspect MblB to go through a similar activation step and this is supported by the loss of protection observed when deleting G395-A396-S397 in MblB (**Figure 5d**). However, the mechanism of this activation remains unclear. The MBL-fold was originally discovered in beta-lactamases and later found to be widely distributed in various proteins, exhibiting hydrolase activity targeting nucleic acids and small molecules (*229*). The MBL-fold protein of 6A-MBL shows strong structural similarity to the OB-fold domain of DNA internalization-related competence protein ComEC (Foldseek (*230*): E-value: 1.56e-11), a protein predicted to be a nuclease (*231*). It is therefore possible that the MBL-fold protein of 6A-MBL acts as the nucleic acid acting effector upon phage infection.

In summary, our findings demonstrate that 6A-MBL is an anti-phage defense system with similarities to eukaryotic antiviral proteins associated with the ubiquitin pathway and CBASS type II. Based on this similarity, we hypothesize that 6A-MBL senses phage infection by the fused E1-E2-JAB protein (Cap2-3), which may use a ubiquitin-like pathway to activate the MBL-fold protein (MblB), resulting in anti-phage activity (hypothetical model proposed in **Figure 5e**).

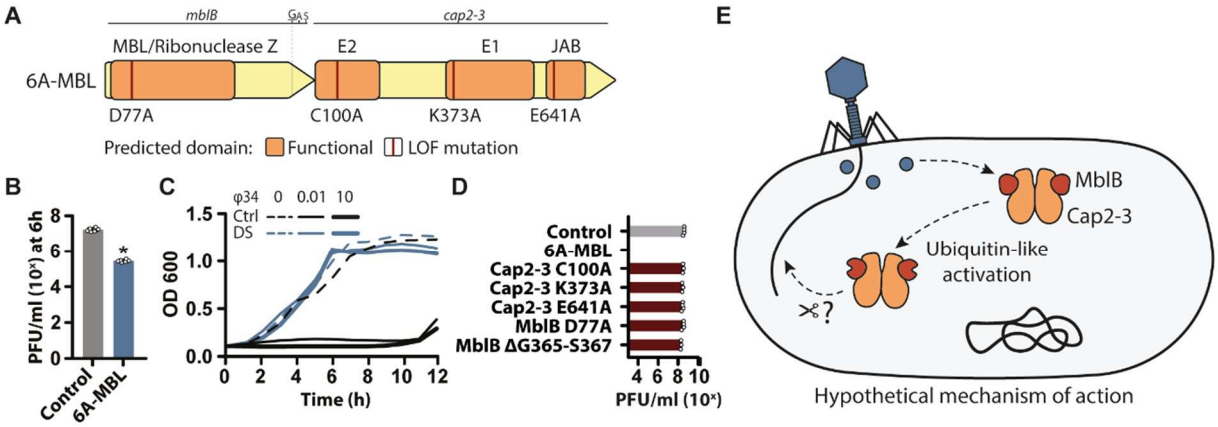

Figure 5 6A-MBL employs a ubiquitination-like mechanism to activate an MBL-fold protein. (a) The functional domains of 6A-MBL and mutation sites tested in (d). LOF, loss of function. The location of a conserved glycine (g) at the C-terminal of the MBL is indicated above the gene and represents the residue where the substrate protein is commonly linked

to its targeted via a ubiquitin-like mechanism. **(b)** Impact of 6A-MBL on phage propagation in liquid culture. The propagation of phage ϕPa34 was monitored over time and the 6h time point is shown here, with all data points displayed in **Figure S4a**. The control bar represents phage propagation in PAO1 containing an empty plasmid. **(c)** Effect of the defense system on bacterial growth upon phage infection. PAO1 cells containing an empty plasmid (control) or 6A-MBL (defense system, DS) were infected with phage ϕPa34 at low (0.01) and high (10) multiplicity of infection and their growth was monitored for a period of 12h. **(d)** Effect of mutations in the functional domains of 6A-MBL on phage protection. The infectivity of phage ϕPa34 on PAO1 cells containing an empty plasmid (control), 6A-MBL, or 6A-MBL with point mutations was assessed by plaque assay. (**e)** Hypothetical model for anti-phage defense by 6A-MBL. 6A-MBL is hypothesized to sense phage infection through the fused E1-E2-JAB protein (Cap2-3), leading to the activation of a potential nuclease activity of the MBL-fold protein (MblB).

Thoeris type III expands the TIR-domain antiviral family

Thoeris type III contains a SLOG domain protein (ThcA) accompanied by four TIR(-like) domain-containing proteins (ThcB1-B4) (**Figure 6a**, **Table 1**). ThcA and ThcB4 are consistently encoded within the Thoeris type III cluster, while the remaining genes exhibit variable presence (ThcB1: 51%; ThcB2 and ThcB3 combined: 18%; either ThcB2 or ThcB3: 64%), establishing them as core genes of this defense system. Thoeris type III exhibits robust protection against a myophage from the *Pbunavirus* genus (ϕPa34), reducing its infectivity by more than 10^6-fold in EOP assays (**Figure 1b**). It efficiently reduces phage proliferation in liquid cultures (**Figure 6b**, **Figure S5a**) and facilitates nearly complete recovery of bacterial growth during phage predation (**Figure 6c**). Mutations in conserved amino acids within core protein ThcA (SLOG, H49A) and accessory TIR-domain containing protein ThcB1 (TIR, D55A) resulted in the loss of defense, while mutagenesis of the active site of other TIR proteins (ThcB2 E82A, ThcB3 E83A, ThcB4 E203A) within the cluster did not result in the loss of defense activity (**Figure 6d**). However, deletion of any TIR domain-proteins led to a complete loss of protection, indicating that the presence of all TIR domain proteins is necessary, while all the individual TIR enzymatic activities are not (**Figure 6d**).

In eukaryotes, TIR domains serve as scaffolding adapters during inflammation, orchestrating pro-inflammatory responses through complex formation involving TIR domains of different proteins (*232*). ThcB4 is homologous to one such eukaryotic scaffolding TIR domain-containing protein, SARM1 (**Table 1**). However, unlike SARM1, ThcB4 lacks obvious multimerization (SAM) and autoinhibition (ARM) domains that regulate TIR domain activity (*233, 234*). Eukaryotic TIR domains also serve to signal pathogen recognition by protein repeats of the Toll-like receptors (*235*). In bacteria, the TIR domains of the Thoeris defense family signal defense effectors upon sensing phage infection. Unlike eukaryotic Toll-like receptors, Thoeris TIR domains function both as scaffolding proteins and phage sensors, independently of protein repeat structures that function as the main pathogen sensors in eukaryotes (*162*). In Thoeris types I and II, TIR proteins sense phage infection, generating secondary messenger molecules (cyclic ADP-ribose (cADPR) or histidine-ADPR, respectively) (*236*). These molecules are recognized by the SLOG(-like) domain, which activates the effector present on the same protein, leading to phage defense.

In Thoeris type III, the two core proteins, ThcB4 and ThcA, are predicted to form a tetrameric complex of two heterodimer subunits of ThcB4 and ThcA (**Figure 6e, Figure S5b,c**). In this configuration, the NucS endonuclease domain of ThcB4 is in direct contact with the SLOG domain of ThcA, indicating a potential regulatory role of SLOG. This observation aligns with previous characterizations of Thoeris types, where SLOG substrate binding physically activates the effector domain (*236*). Unlike previous Thoeris types, the SLOG domain of Thoeris type III belongs to the LD_cluster_3 clade of the SMF/DprA SLOG superfamily, distinct from the traditional SIR2/TIR-associated SLOG domains (STALD) (**Figure 6f**). Moreover, the SLOG domain of Thoeris type III is also predicted to establish direct contact with the TIR domain of ThcB4 (**Figure 6e**), likely facilitating TIR domain signalling to the SLOG domain upon phage infection sensing. However, mutagenesis of the TIR domain of ThcB4 (E203A) did not result in defense loss (**Figure 6d**), suggesting the production of signalling molecules may be assumed by other TIR domain proteins within Thoeris type III. For ϕPa34, this role appears to

be carried out by the TIR domain of accessory protein ThcB1, as mutation of its active site (D55A) resulted in defense loss (**Figure 6d**). Based on the requirement of ThcB2 and ThcB3 but not their TIR activity for full system functioning (**Figure 6d**), we expect these to be sensors of phages outside our panel. Mutagenesis of a conserved residue (G61A) in the Rossmann fold of ThcB4, predicted to bind NADH and to be essential for TIR activity, resulted in complete loss of protection, indicating that the activity of ThcB4 is also necessary for Thoeris defense (**Figure 6d**).

Interestingly, the TIR domain of ThcB1, which is required for the defense against ϕPa34, is phylogenetically distinct from previously described TIR domains of Thoeris and structurally resembles the dimeric Cap12 TIR domain present in effectors of the CBASS and Pycsar defense systems (*237*) (**Figure S5c,d,e**). Unlike these systems where the Cap12 TIR domain causes cellular NAD+ depletion upon phage infection, no change was observed in cellular NAD+ levels during phage infection in Thoeris type III, suggesting that ThcB1 likely acts as a phage sensor rather than an NAD+ depleting effector (**Figure S5f**).

Here, the TIR domains in Thoeris type III likely produce ADPR derivatives upon phage sensing, like Thoeris type I and II systems, to signal the SLOG domain of ThcA for activating the NucS endonuclease domain of core protein ThcB4. Mutagenesis of a conserved amino acid (D368A) in the NucS domain resulted in loss of defense, further supporting a crucial role of the NucS domain, likely as an anti-phage effector nuclease (**Figure 6d**).

In summary, Thoeris type III represents a novel addition to the antiviral TIR-domain family, which carries a novel Thoeris-associated class of SLOG domain that likely activates the anti-phage NucS nuclease effector in ThcB4 (hypothetical model proposed in **Figure 6g**).

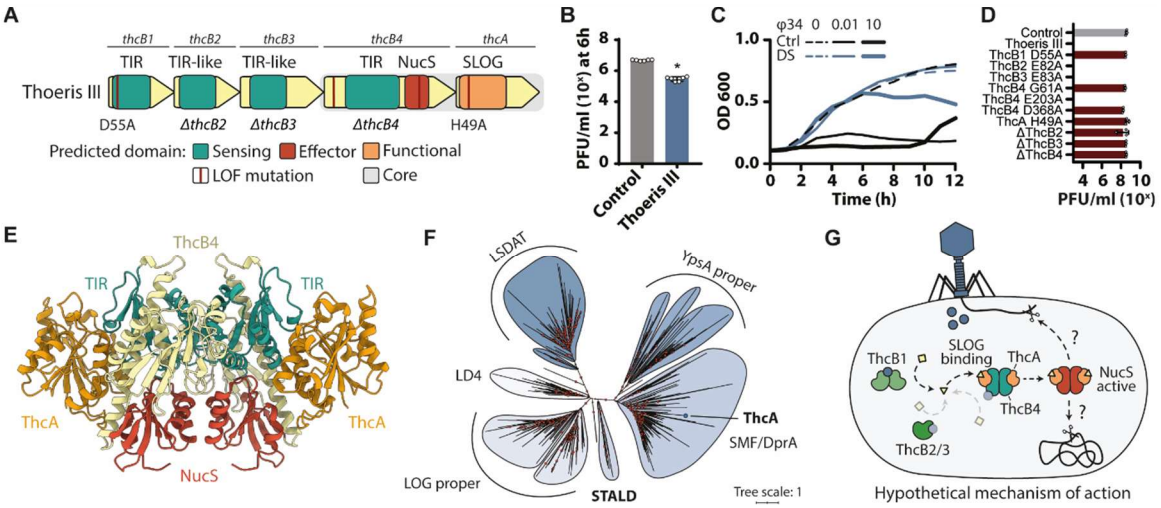

Figure 6 Thoeris type III is a phage defense system of the TIR-domain antiviral family. (a) The functional domains of Thoeris III and mutation sites tested in (d). **(b)** Impact of Thoeris type III on phage propagation in liquid culture. The propagation of phage ϕPa34 was monitored over time and the 6h time point is shown here, with all data points displayed in **Figure S5A**. The control bar represents phage propagation in PAO1 containing an empty plasmid. **(c)** Effect of Thoeris type III on bacterial growth upon phage infection. PAO1 cells containing an empty plasmid (control) or Thoeris type III (defense system, DS) were infected with phage ϕPa34 at low (0.01) and high (10) multiplicity of infection, and their growth was monitored for a period of 12h. **(d)** Effect of gene deletions and mutations in the functional domains of Thoeris type III on phage protection. The infectivity of phage ϕPa34 on PAO1 cells containing an empty plasmid (control), Thoeris type III, or Thoeris type III with point mutations was assessed by plaque assay. **(e)** Tertiary structure of the ThcA:ThcB4 tetramer predicted by Alphafold2. The SLOG-containing ThcA is shown in orange, and the TIR and NucS domains of ThcB4 are shown in green and red, respectively. **(f)** Phylogenetic tree of SLOG-containing proteins. The phylogenetic tree of 13,399 representative proteins was inferred and bootstrapped using IQ-Tree2. The representative proteins include all groups within the SLOG clan (CL0349), as indicated in the tree. Branches with bootstrap confidence interval ≥90% are indicated with red dots. **(g)** Hypothetical model for anti-phage defense by Thoeris type III. Thoeris type III is hypothesized to sense phage infection by TIR proteins (ThcB1 for ϕPa34), leading to the production of signalling molecules. These are likely detected by the SLOG domain of ThcA for the activation of the NucS endonuclease effector response by ThcB4, with which it forms a complex.

Distribution of the anti-phage repertoire across bacterial phyla

Analysis of the distribution of the new defense systems in prokaryotic genomes revealed that Prometheus and 6A-MBL are the most widely distributed systems (**Table S4**, **Figure S6**). Prometheus was found in seven bacterial phyla, exhibiting the highest abundance in Actinomycetota (**Figure S2f,g**), while 6A-MBL was found in eleven bacterial phyla, predominantly prevalent within Proteobacteria (**Figure S4c,d**). The remaining four defense systems were almost exclusively (>92%) observed in Proteobacteria, spanning various classes, except for Hermes, which is restricted to Gammaproteobacteria (**Figures S1g,h**, **S3d,e**, **S5g,h**, **Table S4**).

To assess the impact of the novel defense systems on the diverse and abundant defense repertoire in *P. aeruginosa* (*63, 146*), we determined their distribution across all complete genomes (n = 541) available from the RefSeq database. Prometheus and 6A-MBL are well distributed across the *P. aeruginosa* groups (**Figures S2h**, **S4e**, **Table S5**). Hermes, Erebus, Hypnos, and Thoeris III are less abundant in *P. aeruginosa* (**Figures S1I**, **S3f**, **S5i**, **Table S5**). Collectively, the six validated systems constitute 3% of the total defense systems found in *P. aeruginosa* and increase the average number of defense systems per strain from 9.7 to 10.0 (**Table S5**). All new defense systems can often be found within proximity of known defense systems (**Figure 1a, Figure S1j, S2i, S3g, S4f, S5j**) (*158, 161*). In summary, the defense systems described here have a prominent presence among Proteobacteria and contribute to the phage defense repertoire of *P. aeruginosa*.

Discussion

In this study, we searched for bacterial homologs of eukaryotic antiviral defense genes in *P. aeruginosa* with antiphage properties. We uncovered more than 450 candidates located in proximity to known anti-phage defenses. Experimental testing of eleven candidates revealed six eukaryotic-like novel anti-phage defenses, including Hermes, Prometheus, Erebus, Hypnos, 6A-MBL, and Thoeris III. Five candidate defense systems eluded experimental validation, possibly due to the limited panel of phages, the presence of unknown phage-encoded anti-defense genes, or incompatibility with the genetic background in the *P. aeruginosa* test strain. Of the ones that were validated, Hermes features a eukaryotic-like antiviral functional domain not previously associated with anti-phage activities in bacteria. This defense system causes alterations in surface structures crucial for phage receptor interactions, effectively preventing phage adsorption, akin to the antiviral function of its eukaryotic homologs (*198*). Likewise, Prometheus is thought to play a role analogous to its eukaryotic counterparts (*206, 207*).

We also report the eukaryotic-like defense systems bNACHT Erebus, bNACHT Hypnos, 6A-MBL, and Thoeris III, each potentially equipped with nucleic acid interfering activities. Among these, Erebus and Hypnos represent the first instances of anti-phage NACHT-containing NLR proteins from NACHT clade 20, which display a domain architecture distinct from most NACHT-containing NLRs (*218*), characterized by an N-terminal NACHT domain, central TPR domain, and C-terminal effector. This is an indication that the antiviral function of NACHT-containing NLRs is independent of domain organization and conserved across almost all domains of life (*21*). Furthermore, while the precise phage sensing mechanisms of these systems remain enigmatic, the presence of TPR domains in Erebus and Hypnos suggests a potential interaction with specific phage proteins as their targets.

Another interesting finding is the first instance of an anti-phage E1-E2-JAB fusion in the 6A-MBL system. Previously, these domains were described in unfused proteins, found both in eukaryotes and prokaryotes, where they participate in ubiquitin-like signalling pathways crucial for modulating innate immunity (*188, 224, 237*). In the prokaryotic defense systems CBASS and Pycsar, these proteins prime bacterial cyclases for increased production of signalling molecules, which subsequently activate an effector (*225*). In the context of 6A-MBL, it seems likely that the E1-E2-JAB fusion protein regulates the activity of an MBL protein through a ubiquitin-like mechanism that regulates its anti-phage activity.

Finally, Thoeris III represents a novel member of the TIR-domain antiviral family found across eukaryotes and prokaryotes. Comprising several TIR-domain-containing accessory proteins (ThcB1,

ThcB2, and ThcB3) and two core proteins (ThcB4 and ThcA), Thoeris type III is distinct from other Thoeris types by employing a unique class of SLOG, an endonuclease as a putative effector, and encoding the effector and SLOG domains in unfused proteins that form a tetrameric complex (*186*). It is noteworthy that the endonuclease NucS present in Thoeris type III is a novel endonuclease effector domain linked to anti-phage defense and is also found in bNACHT Erebus.

Four out of six of the validated defense systems in our study share domains with recently reported phage defense systems in bacteria (*188, 218, 237*). This phenomenon underscores the remarkable shared nature of anti-phage protein domains to other genetic contexts, which underlies the rapid diversification in host-pathogen interactions and evolution of innovative functions (*238*). This observation is relevant due to its potential to accelerate the mechanistic characterization of these defense systems. Currently, the pace of understanding the workings of defense mechanisms significantly lags the discovery of new systems (*76, 239-241*). Shared domains among these systems offer a promising bridge to narrow this knowledge gap. By leveraging insights gained from the study of one system to better comprehend another, we can expedite our understanding of the molecular intricacies that drive these defense mechanisms.

The discovery that multiple defensive proteins employed by human cells possess direct homologs in prokaryotes, functioning in viral defense, illuminates the cross-domain links in the evolution of the human defense system and holds implications for future mechanistic studies of human innate immunity. By expanding our knowledge of eukaryotic-like phage defenses, we can gain further insights into these perspectives. The comprehensive exploration and characterization of prokaryotic antiviral defenses could yield a plethora of new biotechnological, therapeutic and diagnostic tools, as seen previously with restriction-modification systems, CRISPR-Cas and Argonaute, which have led to widespread methods for DNA and RNA engineering (*15, 242-247*).

Altogether, we show the existence of new antiviral pathways that are shared between bacteria and eukaryotes. Detailed future studies will be required to unveil the molecular mechanisms of these new defense systems, and this will further help to understand the complex relationship between host defense mechanisms of different domains of life that are potentially forged through a shared viral past.

Materials and Methods

Identification of anti-phage defense systems

We searched for defense systems using PADLOC v1.1.0 with PADLOC-DB v1.4.0 (*121*), and DefenseFinder v1.0.9 with defense-finder-models v1.2.2 (*63*), and the HMMs with completeness rules and thresholds as applied in Gao *et al* (2020) (*122*). Additionally, Detocs was searched for using blastp (*120*) (> 0.7 subject length / query length < 1.5 ; 0.7 > query coverage < 1.3; evalue < 1e-9) and found complete if all three genes were present with a maximum of 1 gene in between (*123*). We observed *dtcB* to fail gene annotation, thus we manually checked for its presence when *dtcA* and *dtcC* were identified within proximity (1 gene in between).

Identification of conserved gene clusters within the variable regions of *P. aeruginosa*

We used PPanGGOLiN v1.2.74 (*248, 249*) to identify the conserved gene clusters within the variable regions of all 541 complete *P. aeruginosa* assemblies from the Refseq database on June 16, 2022. Functional domains were identified using HMMER v3.3.2 (*250*) in combination with the Pfam-A models v36.0.

We utilized Eukprot v3 (*182*) to find homologs of innate immune proteins from vertebrates and plants. For the vertebrate innate proteins, we built HMM profiles of the aligned InnateDB (*180*) in June 2022 using MUSCLE v5.1 (*251*) and hmmbuild (*127*). We applied the same method to find homologs of plant innate proteins, using the Plant Resistance Genes database (PRGd) v4.0 (*181*). Additionally, we applied DRAGO3 (*181*) to search for plant pathogen recognition proteins. For both vertebrate and plant innate proteins, we created a custom database of functional domains found to be associated with innate

defense in literature (**Table S1**). This list of functional domains and the created HMM profiles were then searched for in the representative genes of the *P. aeruginosa* pangenome, with an e-value lower than 0.01 considered significant. We adopted a guilt-by-association approach to categorize the eukaryotic innate-like homologs into "near" and "remote" based on their distance to known phage defense systems. This classification was determined by their proximity within 100kb of a known defense system. We established this threshold based on the bimodal distribution of distances between known defense systems in *P. aeruginosa* (**Figure 1a**). We then annotated the homologs found in *P. aeruginosa* using the Pfam-A HMM library and checked for enriched functional domains within the "near" defense island compared to "remote". A subset of conserved gene clusters with these enriched functional domains were selected for further testing of antiviral activity. The conservation of these gene clusters across species was further assessed using Webflags v1 (*252*). In addition, the signal peptide and transmembrane regions were annotated with the use of SignalP v5.0 (*253*) and Phobius v1.01 (*254*). All confirmed novel defense systems were checked for the presence of repeat regions using HHrepID (*255*). Regions of the proteins that remained uncharacterized were further analysed using DALI (*191*) and Foldseek v6 (*230*). A confidence level higher than 90% and an e-value lower than 0.001 was considered significant.

Cloning of the putative eukaryotic-like defense systems

The putative eukaryotic-like defense systems were amplified from *P. aeruginosa* strains using primers from **Table S6** with Q5 DNA Polymerase (New England Biolabs). PCR products were run on 1% agarose gels, and bands were excised and cleaned using the Zymoclean Gell DNA Recovery Kit (Zymo Research). Plasmid pUCP20 was digested with BamHI and EcoRI (NEB), treated with FastAP (Thermo Scientific), and cleaned with the Zymo DNA Clean & Concentrator Kit (Zymo Research). The defense systems were then cloned into the digested pUCP20 using the NEBuilder HiFi DNA Assembly Master Mix and transformed into chemically competent NEB® 5-alpha Competent *E. coli* following the manufacturer's instructions. Plasmids (**Table S7**) were extracted using the GeneJET Plasmid Miniprep kit and confirmed by sequencing (Macrogen). Confirmed plasmids were transformed into PAO1 by electroporation as previously described (*129*) and the cells were plated on LBA plates supplemented with 200 µg/ml of carbenicillin.

Selection for point mutations

A combination of strategies was employed to identify the amino acids to mutate in the defense system proteins. These included performing a literature search to identify known critical amino acids of the functional domains of the proteins, multiple protein alignments using PSI-BLAST (*126*) and Clustal W v2.1 (*256*) to identify conserved amino acids, and Alphafold2 (*257, 258*) 3D prediction to visually inspect active sites.

Cloning of knockout and point mutants of the defense systems

Gene knockouts and point mutations of the defense systems in pUCP20 were obtained by round-the-horn site-directed mutagenesis using phosphorylated primers (**Table S6**) and Q5 DNA Polymerase. PCR products were digested with DpnI (NEB), and run on 1% agarose gels, and the bands were extracted and cleaned with the Zymo Gel DNA Recovery Kit. The amplified plasmids were ligated with T4 DNA ligase (NEB) at room temperature for 2 hours and transformed into chemically competent NEB 5-alpha Competent *E. coli* following the manufacturer's instructions. Plasmids were extracted using the GeneJET Plasmid Miniprep kit and confirmed by sequencing (Macrogen). Confirmed plasmids were transformed into PAO1 by electroporation as previously described.

Efficiency of plating

Phage stocks were 10-fold serially diluted in LB and the dilutions were spotted onto double layer agar plates of PAO1 with empty pUCP20 or pUCP20 with defense systems following the small plaque drop assay (*131*). The anti-phage activity of the systems was calculated as the fold reduction of phage

infectivity of the PAO1 strain that contains the pUCP20-encoded defense systems, compared to the infectivity of the PAO1 strain containing the empty plasmid.

Infection dynamics of phage-infected cultures

Bacterial cultures of PAO1 strains with empty pUCP20 or pUPC20 containing a defense system at an $OD_{600} \approx 0.1$ were infected with phage at an MOI <1 and incubated at 37°C and 180 rpm. Samples were taken at 0h, 2h, 4h, and 6h, and centrifuged at 3,000 × *g* for 5 min. The phage-containing supernatant was serially diluted and spotted onto DLA plates of PAO1 to estimate phage concentration.

Liquid culture collapse assays

Overnight bacterial cultures were diluted to an $OD_{600} \approx 0.1$ in LB and distributed into 96-well plates. Phages were added at MOIs of 0.01 and 10, and the plates were incubated at 37°C in an Epoch2 microplate spectrophotometer (Biotek) for OD_{600} measurements every 10 min for 24h, with double orbital shaking.

Screening for mutant phages that escape defense

Screening for mutant phages that evade defense was performed as previously described with minor adjustments (*32*). Briefly, phages were 10-fold serially diluted and spotted onto DLA plates containing PAO1 strains expressing individual defense systems. After overnight incubation at 37°C, plates were examined for plaque formation. Single plaques were selected and transferred into LB, incubated for 1h at room temperature, and centrifuged at 8,000 × *g* for 5 min to recover the phage-containing supernatant. The recovered phage suspensions were then serially diluted and spotted onto DLA plates containing PAO1 strains expressing either an empty pUCP20 vector or individual defense systems. Original phages were also diluted and spotted for comparison. Following overnight incubation at 37°C, the fold change in phage concentration between the negative control cells and those containing defense systems was determined.

LC-MS analysis

The LC-MS analyses were performed using an Agilent LC/MS system consisting of a high-pressure liquid chromatography set-up coupled to a triple-quadrupole (QQQ) mass spectrometer (G6460C) equipped with a standard electrospray ionization (ESI) source. Both systems were operated through MassHunter v10.1. The samples were randomized prior to their analysis. 3 µl of each sample were injected into the column of the HPLC. Myo-inositol and inositol monophosphate were delivered to a CSH C_{18} guard-column and a CSH C_{18} column (Waters) (2.1 mm by 5 mm, 1.7-µm pore size) at 30°C with a flow rate of 0.3 ml/min using the following binary gradient: 0% B (acetonitrile), ramp to 5% B in 5 min followed by a 6 min ramp to 75% B then a 1 min ramp back to 100% A (5 mM ammonium acetate) and 3 min re-equilibration (A, 20 mM ammonium formate). The metabolites eluting from the column were sprayed into the mass spectrometer operated in data-dependent mode, as in dynamic multiple reaction monitoring (dMRM) mode using transitions. The MRM transitions were generated by optimizing the fragmentor voltage and the collision energy. The dMRM was acquired in negative mode with a cycle time of 500 ms. Data processing was done using Skyline (*259*). Peaks corresponding to myo-inositol and inositol monophosphate were integrated for quantification, and the area under the curves were exported for further analysis.

Phage adsorption kinetics in Hermes

Bacterial cultures of PAO1 strains containing pUCP20, pUCP20-Hermes, or pUCP20-Hermes D89A were grown until an OD_{600} of 0.3, infected with phage ϕPa34 or ϕPa36 at an MOI of 0.1, and incubated at 37°C and 180 rpm. Samples were taken at 0-, 10-, 20-, and 30-min post-infection and immediately centrifuged at 12,000 × *g* for 1 min. The supernatant was serially diluted and spotted onto DLA plates of PAO1 to estimate the concentration of non-adsorbed phages.

Extraction and visualization of outer membrane protein

Outer membrane proteins were extracted from PAO1 strains containing pUCP20, pUCP20-Hermes, or pUCP20-Hermes D89A as previously described with slight modifications (*260*). Overnight grown cultures were centrifuged at 10,000 × *g* for 20 min, and the cell pellet re-suspended in 10 mM Tris-HCl pH 8.0. Cells were lysed using B-PER™ Bacterial Protein Extraction Reagent (Thermo Fisher Scientific) following manufacturer instructions. The supernatant was centrifuged at 21,000 × *g* for 1h at 15°C, and the cell pellet containing cytoplasmic and outer membranes re-was re-suspended in 10 mM Tris-HCl pH 8 containing 2% Sarkosyl. Samples were incubated in a rotary shaker at room temperature for 30 min and centrifuged at 21,000 × *g* for 30 min at 15°C. The pellet was re-suspended in 10 mM Tris-HCl pH 8 containing 2% Sarkosyl and centrifuged again. The outer membrane pellet was then washed twice and re-suspended in 10 mM Tris-HCl pH 8. Protein concentration was determined using the Pierce BCA Protein Assay Kit (Thermo) according to manufacturer instructions and the samples run on a 4-12% SurePAGE, Bis-Tris gel (GenScript) at 200V in MOPS buffer (GenScript). The gel was stained with Bio-Safe Coomassie G-250 stain (Bio-Rad) and visualized on Bio-Rad Gel Doc XR+.

Extraction and visualization of lipopolysaccharides

LPS was extracted from *P. aeruginosa* PAO1 cells harboring pUCP20, pUCP20-Hermes, or pUCP20-Hermes D89A as previously described with minor modifications (*261*). Briefly, overnight cell cultures were adjusted to an OD_{600} of 0.5 in NaCl 0.9% (w/v). Cells were centrifuged at 10,600 × *g* for 1 min and the pellet re-suspended in NaCl 0.9% and centrifuged again. The cell pellet was re-suspended in lysis buffer (1M Tris-HCl pH 6.8, 2% SDS, 4% 2-mercaptoethanol, 10% glycerol) and incubated at 95°C for 10 min. Proteinase K solution (25 µg/ml proteinase K in lysis buffer) was added to cooled down samples and incubated at 56°C for 1h with shaking. The solution was directly loaded onto a 4-12% SurePAGE, Bis-Tris gel and run at 20 mA in MOPS buffer. The gel was stained with the SilverQuest Silver Staining Kit (Thermo Scientific) following manufacturer instructions and imaged with Bio-Rad GelDoc XR+.

Extraction and visualization of capsule density

Bacteria were separated by capsule density using a discontinuous density gradient as previously described with minor adaptations (*262*). Briefly, overnight cultures of *P. aeruginosa* PAO1 cells harboring pUCP20, pUCP20-Hermes, or pUCP20-Hermes D89A were centrifuged for 10 min at 3,200 × *g* for 10 min, and the pellet re-suspended in 1X phosphate buffered saline (PBS). After centrifugation, the pellet was washed once more with 1X PBS and re-suspended in 1 ml of 1X PBS. The cell suspension was added onto a discontinuous density gradient of 30%, 60%, 80% Percoll (w/v in 1XPBS, Sigma), and centrifuged for 30 min at 3,000 × *g* for 30 min.

Phage adsorption to capsule and lipopolysaccharides

CPS and LPS were extracted from PAO1 strains containing pUCP20, pUCP20-Hermes, or pUCP20-Hermes D89A as previously described with slight modifications (*263*). Briefly, overnight cell cultures were adjusted to an OD_{600} of 0.65 and centrifuged at 12,000 × *g* for 5 min. The cell pellet was re-suspended and vortexed in lysis buffer (60 mM Tris-HCl pH 8, 5 µg/ml of $CaCl_2$, 2 mg/ml $MgCl_2$) containing 3 mg/ml lysozyme and incubated at 37°C for 1h. Samples were then subjected to three cycles of freeze (-80°C) and thaw (37°C). DNase and RNase were added at 20 µg/ml and incubated at 37°C for 30 min, followed by additional incubation with 0.1% SDS. Extractions were boiled at 100°C for 10 min, cooled down and mixed 2 mg/ml of proteinase K in lysis buffer. After incubation at 60°C for 1h, extractions were centrifuged at 14,000 × *g* for 2 min. The supernatant was recovered, mixed with cold 75% ethanol, and incubated at -20°C overnight. Extractions were centrifuged at 14,000 × *g* for 30 min at 4°C and the CPS/LPS pellets resuspended in LB. The CPS/LPS pellets were 10-fold serially diluted and incubated with phage ϕPa34 at 10^6 pfu/ml for 1h at 37°C with shaking. The mixtures were 10-fold serially diluted and spotted onto DLA plates containing PAO1 to determine the amount of phage that did not adsorb to the CPS/LPS extracts.

Protein complex prediction

Predicted aligned error plots were generated with Alphafold2 multimer (*257, 258*) for each hetero- and homomeric co-folded protein combination. Overlapping low predicted aligned error rates between proteins were seen as an indication for putative protein complex formation.

Phage R-loop prediction

QmRLFS-finder (*264*) was applied to predict the R-loop structures within the phage collection. Predicted R-loops were considered only if they were identified within a transcript and in the same transcriptional direction.

Comparison of IMPA, SuhB, and HrsA

We compared the structure of bacterial SuhB (*M. tuberculosis*, PDB: 2Q74) and eukaryotic Human IMPA (PDB: 1AWB) with the Alphafold2 predicted structure of Hermes (WP_023087430.1) using DALI pairwise structure comparison. The alignments of these structures were further investigated using ChimeraX matchmaker.

RNA-seq experiments with Prometheus

Bacterial cultures of PAO1 strains containing pUCP20, Prometheus, and their mutants, were grown until an OD_{600} of 0.3 and infected with phage at MOI of 1. Samples were taken pre-infection and at 20 min post infection and centrifuged at 5000 × *g* for 5 min at 4°C. RNA was extracted from the cell pellets using the RNeasy Mini kit and the RNAse-free DNase set (Qiagen) following manufacturer's instructions. RNA concentration was determined using Qubit RNA High Sensitivity (HS) (Thermo), and the samples subjected to prokaryotic mRNA-seq (WOBI) by Novogene (Cambridge, UK). Adapters (5' adapter: 5'-AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGCCGTATCATT-3'; 3' adapter: 5'-GATCGGAAGAGCACACGTCTGAACTCCAGTCACGGATGACTATCTCGTATGCCGTCTTCTGCTTG-3') were removed from the reads and the reads filtered out if they contained an N content greater than 10%, or low-quality (Qscore<= 5) bases for more than 50% of the total bases. The remaining sequences were mapped on their corresponding reference genome using bwa v0.7.18 (*265*). Subsequently, the number of reads per gene were determined using Samtools v1.20 (*108*) in cohorts with bedtools intersect v2.31.1 (*266*) and the NCBI genome annotation file (gff) (*267, 268*).

Sequencing during phage infection

Bacterial cultures of PAO1 strains containing pUCP20, Prometheus, and its mutants were grown until an OD_{600} of 0.3 and infected with phage at MOI of 1. Samples were taken pre-infection and at 20 min post infection and centrifuged at 5000 × *g* for 5 min at 4°C. DNA was extracted from the cell pellets using the GeneJET Genomic DNA Purification Kit (Thermo Scientific) following manufacturer instructions. DNA was sequenced using Nanopore at Plasmidsaurus. Obtained reads were quality checked and the bottom 5% of reads with the lowest quality score reads removed via Filtlong v0.2.1 (default parameters). The filtered reads were mapped onto the corresponding reference genome of the phage and host using minimap2 v2.28 (*106*) (map-ont parameter). The alignment statistics were obtained using Samtools v1.20 and visualized using Integrative Genomics Viewer (IGV) (*269*).

NAD/NADH measurements

Bacterial cultures of PAO1 strains containing pUCP20, Thoeris III, and Thoeris III H49A SLOG were grown until an OD600 of 0.3. Cultures were divided and one part infected with phage ϕPa34 at an MOI of 1 for 30 minutes. Uninfected and infected cultures were centrifuged at 5000 × g for 5 min at 4°C, and the pellet washed once with cold 1x PBS. The cell pellets were then used for NAD/NADH quantification using the NAD/NADH quantification kit (Sigma) following manufacturer instructions.

Building HMM models of the defense system

HMM models for multi-gene defense systems were created using cblaster v1.3.18 (*270*) to obtain all clusters within *P. aeruginosa* in which the defense genes were present and near each other (20% pident). Single-gene defense systems were searched for using PSI-BLAST (*126*) search, and sequences with a pident lower than 20% and coverage lower than 70% were excluded. Muscle5 (*251*) (-super5) was used to align the obtained protein sequences, and hmmbuild v3.3.2 (*127*) was used to build the HMM models. The HMM model scoring thresholds were set based on the 100% sensitivity-point with the help of an ROC curve analysis that scored the HMM sensitivity for the defense gene compared to all other genes within *P. aeruginosa*. The HMM bitscore obtained for each defense system gene is as follows: Prometheus *proA*, 3500; Hermes *hrsA*, 325; 6A-MBL *mblB*, 130, and *cap2-3*, 230; Erebus *eruA*, 190; Hypnos *hypA*, 350; and Thoeris III *thcA*, 75, *thcB1*, 50, *thcB2*, 50, *thcB3*, 50, *thcB4*, 75, with the presence of *thcA* and *thcB4* genes being mandatory.

Search for the defense systems in archaeal and bacterial genomes

To detect all prokaryotic homologs of the novel defense systems, we applied the HMM models on the proteins of all representative genomes, which were downloaded from RefSeq in February 2023. The taxonomy of each identified defense system is based on the annotation provided by RefSeq itself.

Phylogenetic tree and annotation of the defense system genes

Phylogenetic trees of ProA (Prometheus) and HrsA (Hermes) were made by obtaining all available Uniprot Release 2024_01 (*271*) protein sequences that contained DNA2 (n = 131,097) or IMPase (n = 56,571) functional domains, respectively. ATP-dependent DNA helicase RecG protein (NCBI accession: NP_254032.1) and fructose-1,6-bisphosphatase class I (FBPase class I; NCBI accession: CEI80039.1) were used to root the ProA and the HrsA phylogenetic trees, respectively. To generate the TIR and SLOG domain phylogenetic trees, all sequences of the corresponding protein clans, STIR (CL073, n = 108k) and SLOG (CL0349, n = 155k), were downloaded from the InterPro 98.0 database (*272*) on February 6, 2024. Duplicate sequences were removed using Seqkit v2.7.0 (rmdup) (*273*), followed by a down sampling step (Seqkit sample). Representative sequences were obtained using mmseqs2 v15.6 (easy-cluster) (*274*), with the degree of down sampling based on the number of representatives obtained (between 1000 and 1500 sequences). The representative sequences were aligned using Muscle v5.3 (-super5) (*275*) and trimmed using trimAl v1.8 (-automated1) (*276*). The resulting trimmed alignment was used to build and bootstrap a phylogenetic tree using IQ-Tree2 v2.2.6(*157*) (-B 1000, --alrt 1000, -m TEST). The phylogenetic tree of Hypnos and Erebus was built with the NACHT-containing NLR protein accessions and phylogenetic group information provided by Kibby *et al.* (2023) (*218*) (n = 3247). These sequences were aligned using Muscle v5.3 (-super5), trimmed using trimAl v1.8 (*276*) (-automated1), and a phylogenetic tree was built using FastTree v2.1.11 (*277*) with default settings. All phylogenetic trees were visualized using iTOL v1.0 (*278*).

Phylogeny tree of *P. aeruginosa*

A phylogenetic tree was constructed using Parsnp v1.7.4 (*111*) on all complete *P. aeruginosa* assemblies from the Refseq database on June 16, 2022 (n = 541) and visualized using iTOL (*279*).

Quantification and statistical analysis

Unless stated otherwise, data are presented as the mean of biological triplicates ± standard deviation. A Bonferroni-adjusted p-value of less than 0.05 was considered significant.

Supplemental Information

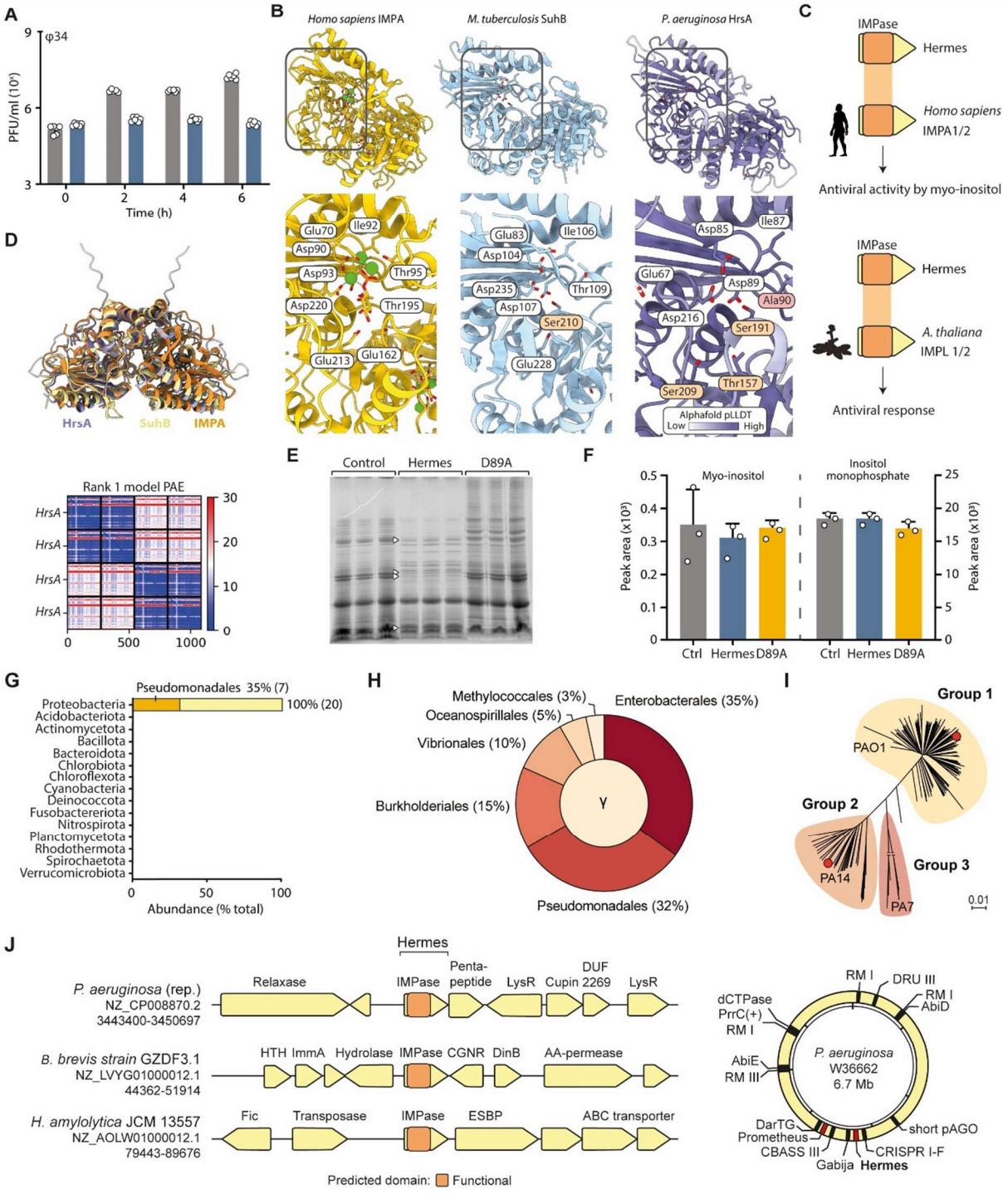

Figure S1 Features of the defense system Hermes. (a) Hermes limits propagation of phage ϕPa34 in liquid cultures. **(b)** Structure of homologous proteins *H. sapiens* IMPA (PDB: 1AWB), *M. tuberculosis* SuhB (PDB: 2Q74), and *P. aeruginosa* HrsA (predicted by Alphafold). The IMPase active site is depicted in detail with residues relevant for interaction with inositol monophosphate indicated. Residue differences between SuhB or HrsA compared to IMPA are highlighted in orange when substituted by residues of similar characteristics and in red when distinct. **(c)** Similarity of Hermes to antiviral eukaryotic IMPase proteins IMPA1/2 from Homo sapiens (NP_001138359.1) and IMP-Like 1/2 (IMPL) from Arabidopsis thaliana (NP_001118558.1). **(d)** Structure comparison of homologous proteins *H. sapiens* IMPA (PDB: 1AWB, orange), *M. tuberculosis* SuhB (PDB: 2Q74, yellow), and *P. aeruginosa* HrsA (predicted by Alphafold, purple). The PAE plot of the Alphafold multimer model for dimeric HrsA is also shown. **(e)** Outer membrane proteins extracted from cells carrying an empty plasmid (control), Hermes, or Hermes with a point mutation (D89A) and visualized by SDS-PAGE. Proteins differentially expressed in Hermes are indicated with white arrows. **(f)** Quantification of myo-inositol and inositol monophosphate via liquid chromatography–mass spectrometry (LC–MS) of extracts obtained from cells carrying an empty plasmid (control), Hermes,

or Hermes with a point mutation (D89A). **(g)** The abundance of Hermes within bacterial phyla. **(h)** The distribution of Hermes proteins within Proteobacteria. γ: Gammaproteobacteria. **(i)** The distribution of Hermes across *P. aeruginosa*. Red circles indicate strains that contain Hermes. **(j)** The gene neighborhood of Hermes within and outside *P. aeruginosa* (left), and representative genome showing the location of Hermes and other defense systems (right). Defense systems discovered in this study are shown as a red line.

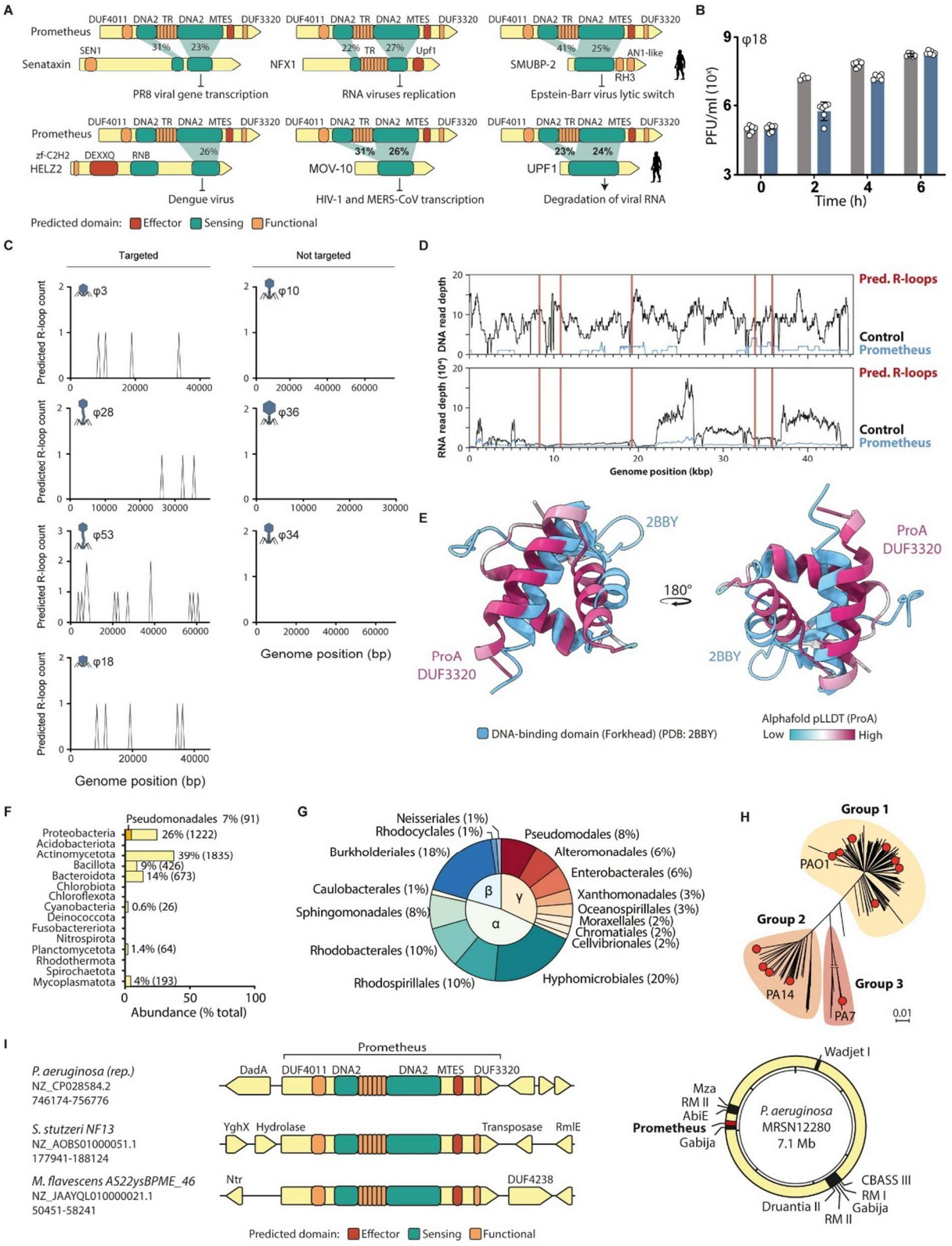

Figure S2 Features of the defense system Prometheus. (a) The human homologs of Prometheus, depicting gene architecture and regions of sequence similarity. **(b)** Prometheus limits propagation of phage ϕPa18 in liquid cultures. **(c)** R-loops identified in phage genomes from our panel. Phages with multiple R-loops are susceptible to the anti-phage activity of Prometheus (targeted), while those with a limited number of R-loops are resistant (not targeted). **(d)** DNA (top) and RNA (bottom) read depth across the ϕPa18 genome during infection of cells containing an empty plasmid (Control) or a plasmid expressing Prometheus. DNA and RNA were extracted at 20 minutes post-infection. Predicted location or R-loops in the phage genome are indicated with a red line. **(e)** Overlay of Prometheus ProA DUF3320 domain predicted by Alphafold2 with the DNA-binding Forkhead domain (PBD: 2BBY). **(f)** The abundance of Prometheus within bacterial phyla. **(g)** The distribution of Prometheus proteins within Proteobacteria. α: Alphaproteobacteria, β: Betaproteobacteria, γ: Gammaproteobacteria. **(h)** The distribution of Prometheus across *P. aeruginosa*. Red circles indicate strains that contain Prometheus. **(i)** The gene neighborhood of Prometheus within and outside *P. aeruginosa* (left), and representative genome showing the location of Prometheus and other known phage defense systems (right). Defense systems discovered in this study are shown as a red line.

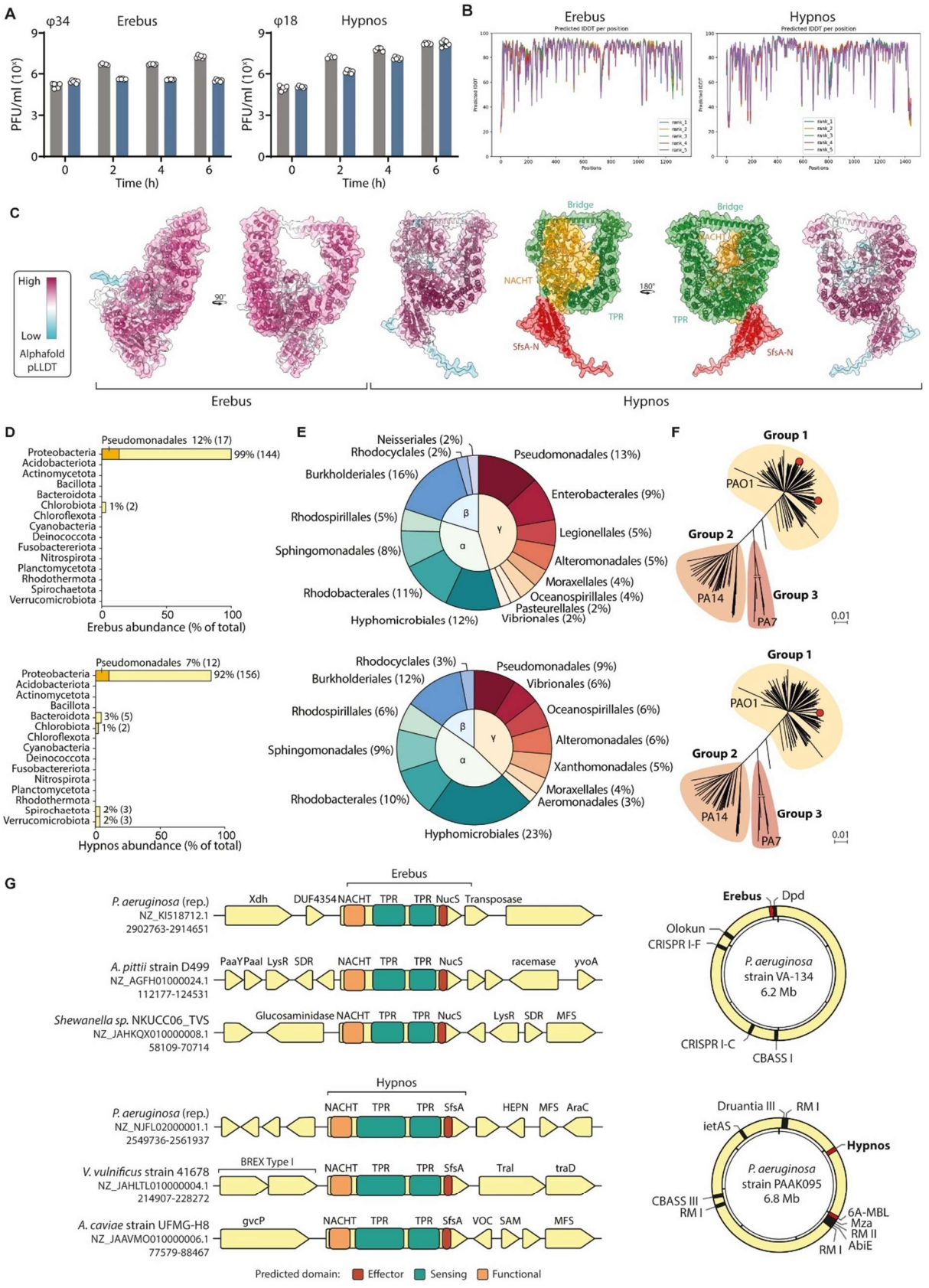

Figure S3 Features of the defense systems Erebus and Hypnos. (a) Erebus and Hypnos limit the propagation of phages ϕPa34 and ϕPa18, respectively, in liquid cultures. **(b)** Predicted lDDT per position for Erebus and Hypnos Alphafold2 structures. **(c)** Tertiary structure of Erebus and Hypnos predicted by Alphafold2. On the left, the structure of Erebus is color-coded by pLLDT score. On the left, the structure of Hypnos is shown both color-coded by pLLDT score and by functional domains, with TPR shown in green, NACHT in orange, and the effector domain (SfsA-N) in red. **(d)** The abundance of Erebus

(top) and Hypnos (bottom) within bacteria phyla. **(e)** The distribution of Erebus (top) and Hypnos (bottom) proteins within Proteobacteria. α: Alphaproteobacteria, β: Betaproteobacteria, γ: Gammaproteobacteria. **(f)** The distribution of Erebus (top) and Hypnos (bottom) across *P. aeruginosa*. Red circles indicate strains that contain the defense system. **(g)** The gene neighborhood of Erebus (top) and Hypnos (bottom) within and outside *P. aeruginosa*, and representative genome showing the location of Erebus (top) and Hypnos (bottom) and other defense systems. The defense systems discovered in this study are shown as a red line.

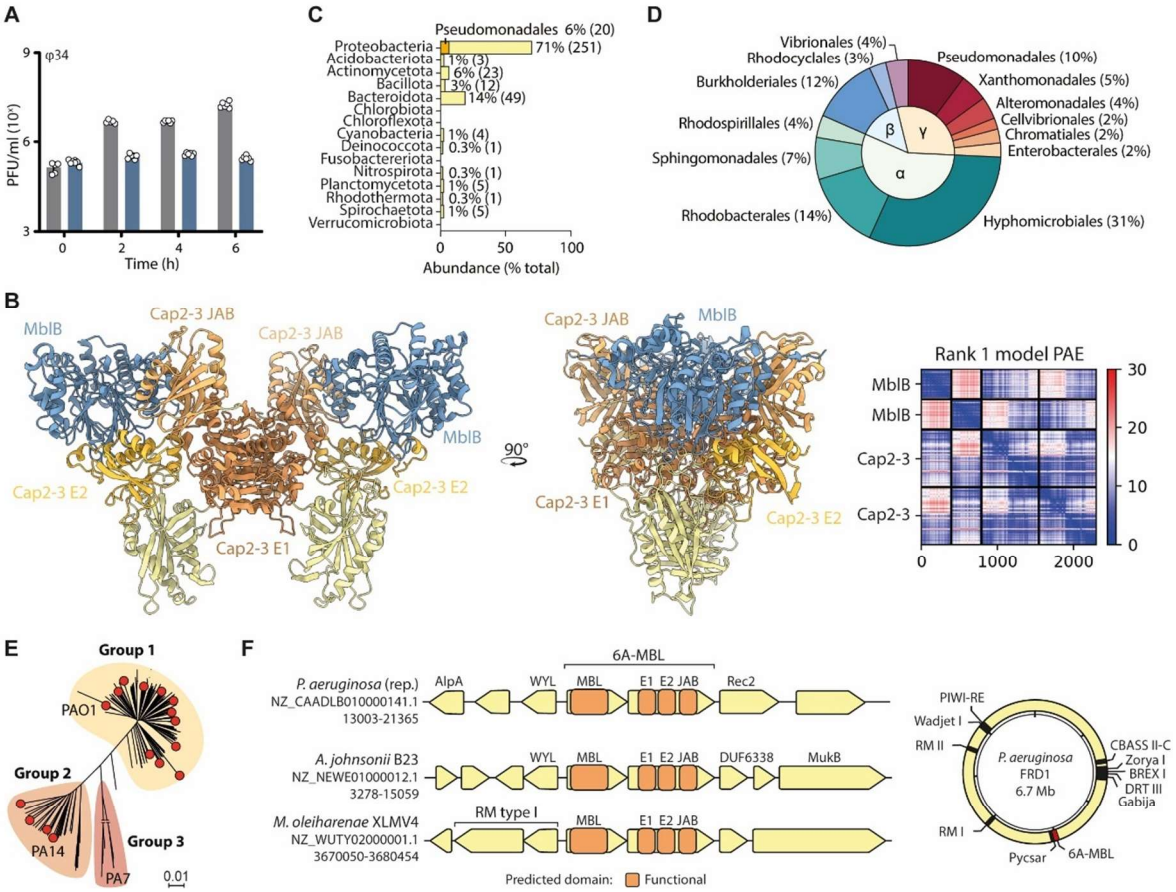

Figure S4 Features of the defense system 6A-MBL. (a) 6A-MBL limits propagation of phage ϕPa34 in liquid cultures. **(b)** Tertiary structure of the complex formed between MblB and Cap2-3 proteins of 6A-MBL, predicted by Alphafold 2. On the left the structure of the complex is color-coded by functional domains with MblB shown in blue, Cap2-3 E1 in dark orange, Cap2-3 E2 in yellow, and Cap2-3 JAB in light orange. **(c)** The abundance of 6A-MBL within bacteria phyla. **(d)** The distribution of 6A-MBL proteins within Proteobacteria. α: Alphaproteobacteria, β: Betaproteobacteria, γ: Gammaproteobacteria. **(e)** The distribution of 6A-MBL across *P. aeruginosa*. Red circles indicate strains that contain 6A-MBL. **(f)** The gene neighborhood of 6A-MBL within and outside *P. aeruginosa* and representative genome showing the location of 6A-MBL and other defense systems. The defense systems discovered in this study are shown as a red line.

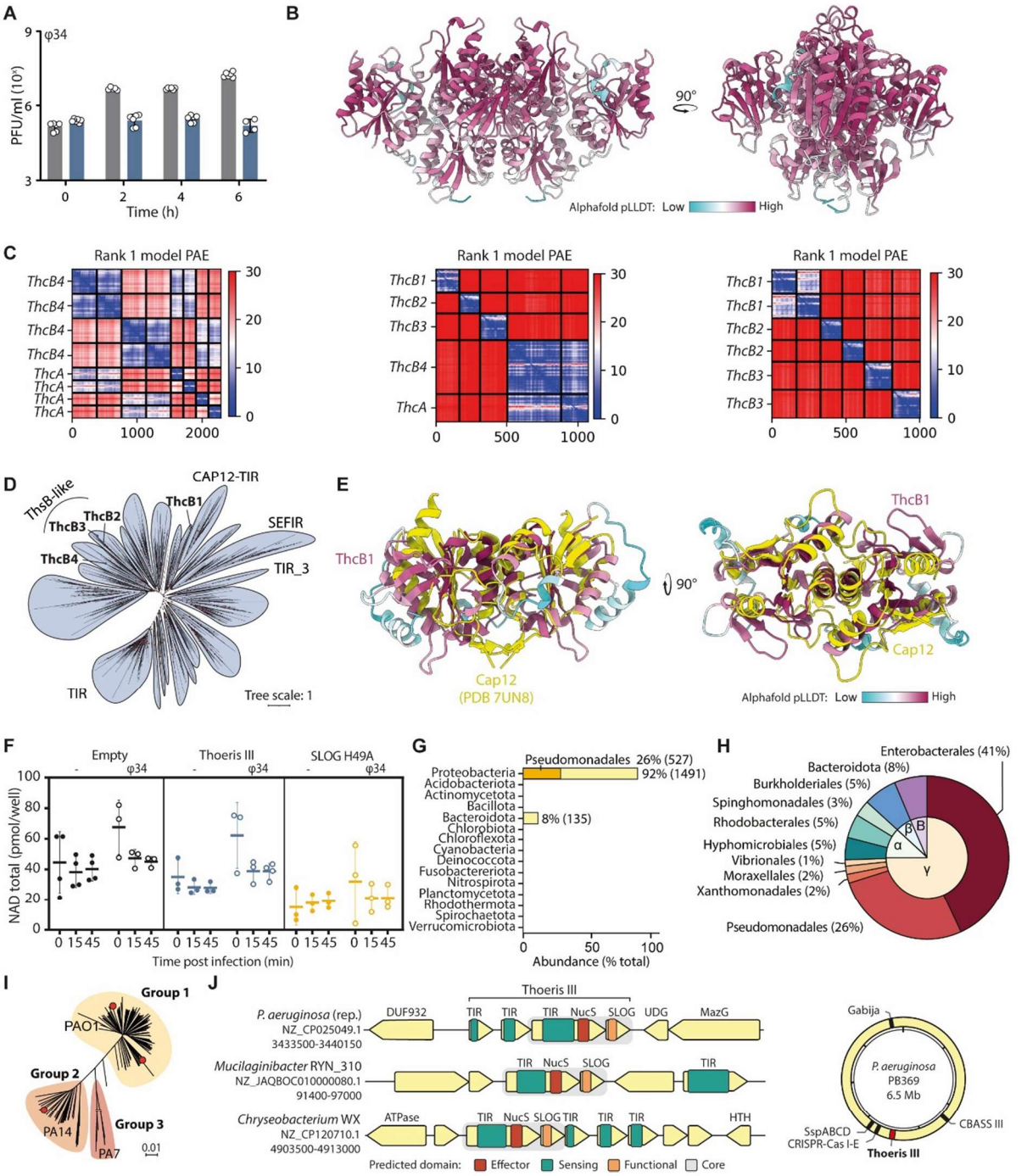

Figure S5 Features of the defense system Thoeris type III. (a) Thoeris type III limits propagation of phage ϕPa34 in liquid culture. **(b)** Tertiary structure of the ThcA:ThcB4 tetramer predicted by Alphafold2, color-coded by pLLDT score. **(c)** PAE plot of AlphaFold multimer models for complexes formed between ThcA and ThsB1-B4 proteins and among ThsB1-B4 proteins of Thoeris III. In each square, the shade in blue indicates the expected distance error in Angström (Å) for a pair of residues. **(d)** Phylogenetic tree of TIR-containing proteins. The phylogenetic tree of 5,201 representative proteins was inferred and bootstrapped using IQ-Tree2. The representative proteins include all groups within the STIR clan (CL0173), as indicated in the tree. Branches with bootstrap confidence interval ≥90% are indicated with red dots. **(e)** Tertiary structure of ThcB1 predicted by Alphafold 2 and overlaid with the TIR domain of Cap12 (PDB 7UN8). Cap12 is shown in yellow and ThcB1 is color-coded by pLLDT score. **(f)** Total intracellular NAD measured in cells harboring an empty plasmid, Thoeris III, or mutation H49A in the SLOG domain, during infection with phage ϕPa34. **(g)** The abundance of Thoeris type III within bacteria phyla. **(h)** The distribution of Thoeris type III within Proteobacteria. α: Alphaproteobacteria, β: Betaproteobacteria, γ: Gammaproteobacteria, B: Bacteroidota. **(i)** The distribution of Thoeris type III across *P. aeruginosa*. Red circles indicate strains that contain Thoeris type III. **(j)** The gene neighborhood of Thoeris III within and outside *P. aeruginosa*, and representative genome showing the location of Thoeris type III and other defense systems. The defense systems discovered in this study are shown as a red line.

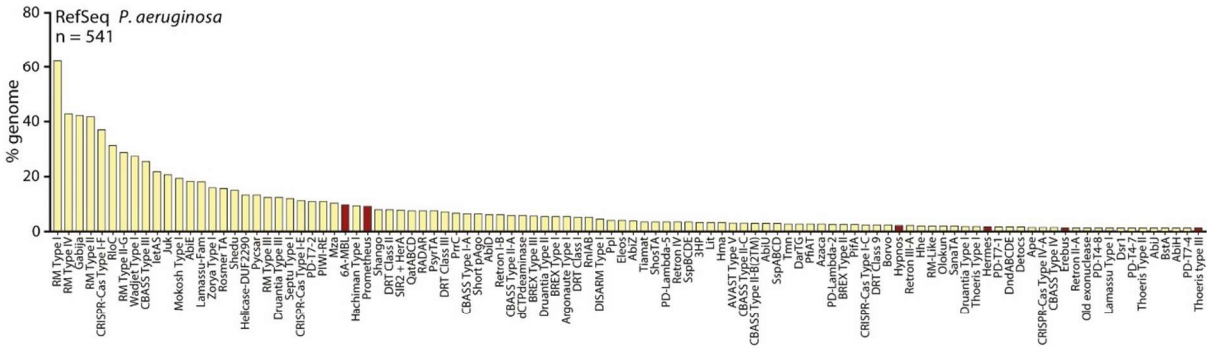

Figure S6 Impact of the new defense systems on the *P. aeruginosa* anti-phage repertoire. Diversity of defense systems found in the genomes of 541 *P. aeruginosa* strains from the RefSeq database, organized from most (left) to least (right) abundant. The complete list of defense systems is shown on **Table S5**. Defense systems discovered in this study are depicted in red.

The following supplementary files are available online at doi: 10.1016/j.chom.2024.07.007

Table S1 List of eukaryotic innate proteins retrieved from the literature.

Table S2 Representative genes of the variable regions within the pangenome of 542 RefSeq *P. aeruginosa* genomes, identified by PPanGGoLiN.

Table S3 R-loops predicted in phage genomes from Costa *et al.* (2024), Gao *et al.* (2020), and Millman *et al.* (2022).

Table S4 Presence of the defense systems in RefSeq representative genomes of prokaryotes.

Table S5 Defense systems identified in the 542 RefSeq genomes of *P. aeruginosa*.

Table S6 List of primers used in this study.

Table S7 List of plasmids used in this work.

Chapter 5



Discovery of phage defense systems through component modularity networks

Phage defense systems in bacteria exhibit high degrees of modularity, with sensing, signal transmission, and effector enzymes frequently being exchanged among phage defense gene clusters. In this study, we capitalized on this modularity to discover phage defense systems by searching for defense-associated modules in new gene contexts. This approach revealed a large and interconnected network of modular components distributed across diverse gene clusters. From over 500 candidate defense systems, we selected nine for experimental testing and validated three: Dionysus, a TerB-encoding system that disrupts early phage infection vesicle formation by Jumbo phages; Ophion, a Radical SAM-containing system that prevents the formation of the Jumbo phage nucleus; and Ambrosia, a tightly regulated RM-like system. Collectively, we demonstrate that leveraging the modular architecture of phage defense systems is an effective approach to their discovery.

A modified version of this chapter has been published as

van den Berg, D. F.*, Costa, A. R.*, Esser, J. Q.*, Muralidharan, A., van den Bossche, H., & Brouns, S. J. J. (2025). Discovery of phage defense systems through component modularity networks. bioRxiv, doi: 10.1101/2025.09.30.679545

Introduction

Bacteriophages (phages) are viruses that infect and replicate within a bacterial host (*280*). To defend themselves against these viral threats, bacteria evolved a diverse repertoire of phage defense systems. These systems detect the infecting phage and trigger immune responses that range from cleaving the invading phage DNA to disrupting key metabolic processes, such as ATP and NAD^+ metabolism, thereby preventing phage propagation (*10*). In response, phages evolved counter-defenses, prompting further bacterial adaptations (*281*). Over time, this ongoing evolutionary race has generated remarkable diversity among bacterial anti-phage strategies (*10*).

It has become increasingly evident that many defense systems are built from modular components. This modularity is exemplified by phage defense families which comprise multiple subtypes that are unified by conserved core elements, such as the cyclase of CBASS, the ATPase NsnB of Menschen, LmuB of Lamassu, YprA family helicase of ARMADA, and the protease-nuclease pair of Canu (*160, 188, 282-284*). Apart from shared genes, many systems have also interchanged functional domains, creating novel architectures from a common pool of modular parts. For example, sirtuin (SIR2) domains appear in several phage defense families, including Thoeris, prokaryotic argonautes (pAgos), defense-associated sirtuins (DSRs), and Avs (*122, 159, 160, 186, 285, 286*). Other frequently repurposed domains include superfamily helicase (SF2), HNH endonuclease, CD-NTase associated protein 4 (Cap4), modified DNA rejection and restriction (Mrr), restriction endonuclease (REase), Toll/interleukin-1 receptor/resistance protein (TIR), Trypsin-like, Caspase, Metallo-β-lactamase (MBL), and purine nucleoside phosphorylase (PNP) domains (*238, 283, 284, 287*).

The modular exchange of these genes and domains provides a genetic breeding ground for the genesis and diversification of phage defense systems. Understanding these modular components that are shared in evolution is key to uncovering the full spectrum of bacterial antiviral mechanisms. In this study, we use this principle to identify candidate defenses by analysing gene clusters that combine at least one known defense-associated gene with different domain contexts. We focused on *P. aeruginosa*, a species with a highly diverse phage defense repertoire (*161, 288, 289*). This approach uncovered more than 500 candidate systems, nine of which were selected for experimental testing. Of these, we validated three: Dionysus, a TerB-containing system specifically targeting jumbo phages; Ophion, a Radical SAM (rSAM)-containing system also targeting jumbo phages; and Ambrosia, a restriction-modification (RM)-like system targeting *Pbunavirus* and *Mesyanzhinovviridae* phages. Our findings demonstrate the power of modular domain-based discovery approaches to expand the known diversity of phage defense mechanisms and offer new insights into the architecture of bacterial immunity.

Results

Modularity of phage defense system components

More than 95% of known phage defense systems in *P. aeruginosa* reside within regions of genomic plasticity (RGPs) (*248*). To identify previously uncharacterized phage defense systems, we performed an extensive search for gene clusters residing in these RGPs across 541 *P. aeruginosa* genomes. This analysis identified 1040 conserved gene clusters (**Figure 1a**) of which 515 (40%) contained functional domains associated with known defense systems. The most prevalent defense-associated domains were P-loop-containing ATPases (10.6%) and methyltransferases (MTases, 3.8%) (**Table S1**). The modularity of gene clusters that contained phage defense-associated functional domains was further analysed using a gene network approach to investigate shared genes, resulting in a large, interconnected network of phage defense-related clusters (**Figure 1b-d, Supplementary File 1**).

We then classified the identified clusters into six distinct categories: 1. Complete defense systems (12.2%); 2. Complete systems with additional defense-associated genes (4.2%); 3. Complete systems with additional genes not previously associated with defense (5.6%); 4. Complete systems with additional genes both defense-associated and non-defense-associated (7.0%); 5. Clusters containing

defense-associated genes in new combinations (57%); and 6. Clusters with both defense and non-defense genes (14%). Importantly, nearly half of the gene clusters (247/515; 48%) shared at least one gene with another cluster, allowing us to combine these into gene networks (**Figure 1b; Table S1**). Most of these networks are small, but three gene networks (networks 3, 5, and 8) each contain more than 15 genes (**Figure 1c; Table S1**). These larger networks consist of several gene clusters and may include novel or accessory elements relevant to phage defense. For instance, gene network 8, which is the largest network, spans 95 genes from 33 gene clusters and is largely composed of genes not previously associated with phage defense, aside from a few ATPases and STAND proteins. On the other hand, networks 3 and 5 are strongly enriched in known defense elements. Network 3 contains several type I and II restriction-modification (RM) system components, while network 5 includes genes that are part of BREX, DISARM, and Druantia systems. Although we did not further investigate these networks in this study, they offer starting points for future exploration of phage defense diversity and modularity.

To investigate the key contributors to defense network formation, we evaluated the gene connectivity within defense system families (degree centrality score: number of connections) and their role in linking distinct subnetworks (betweenness centrality score: bridge formation between subnetworks) (**Figure 1e,f; Table S1**). Interestingly, most genes with high number of connections within phage defense families (14 out of the top 20) have yet to be associated with phage defense, while known defense genes predominantly scored high in bridging phage defense subnetworks (15 out of the top 20) (**Figure 1e,f**).

For example, NAD(P)-binding protein (gene 233, **Table S1**) from gene network 3 is ranked the highest in both degree and betweenness centrality but has not previously been associated with phage defense. In contrast, other genes with high scores in both degree centrality and betweenness have been previously associated with phage defense, including MTases and REases of RM systems. Similar to what was observed for the largest gene networks 3 and 5, we found these to be commonly shared among RM subtypes as well as other RM-like systems (*159, 241, 290*). Similarly, ATPases, such as SMC-like ATPase (gene 3542), PARIS ATPase AriA (gene 1225), and the ATP-Binding Cassette (ABC)-2 type transporter (gene 4514), scored highly in both categories. These ATPase domains are often phage-sensing components of defense systems, suggesting they might be shared among defense systems for phage specificity(*291, 292*).

These insights into the fluidity of the phage defense repertoire of *P. aeruginosa* suggested that we could utilize this gene network to discover additional phage defense systems. To test this, we selected nine gene clusters with shared defense components (clusters type 2-5) for experimental validation of phage defense and plasmid conjugation in *P. aeruginosa* strain PAO1 (*288*). Three gene clusters (33%), which we named Dionysus, Ophion, and Ambrosia (**Figure 1b, Table 1**), exhibited robust anti-phage activity and were further characterized, as discussed in detail below.

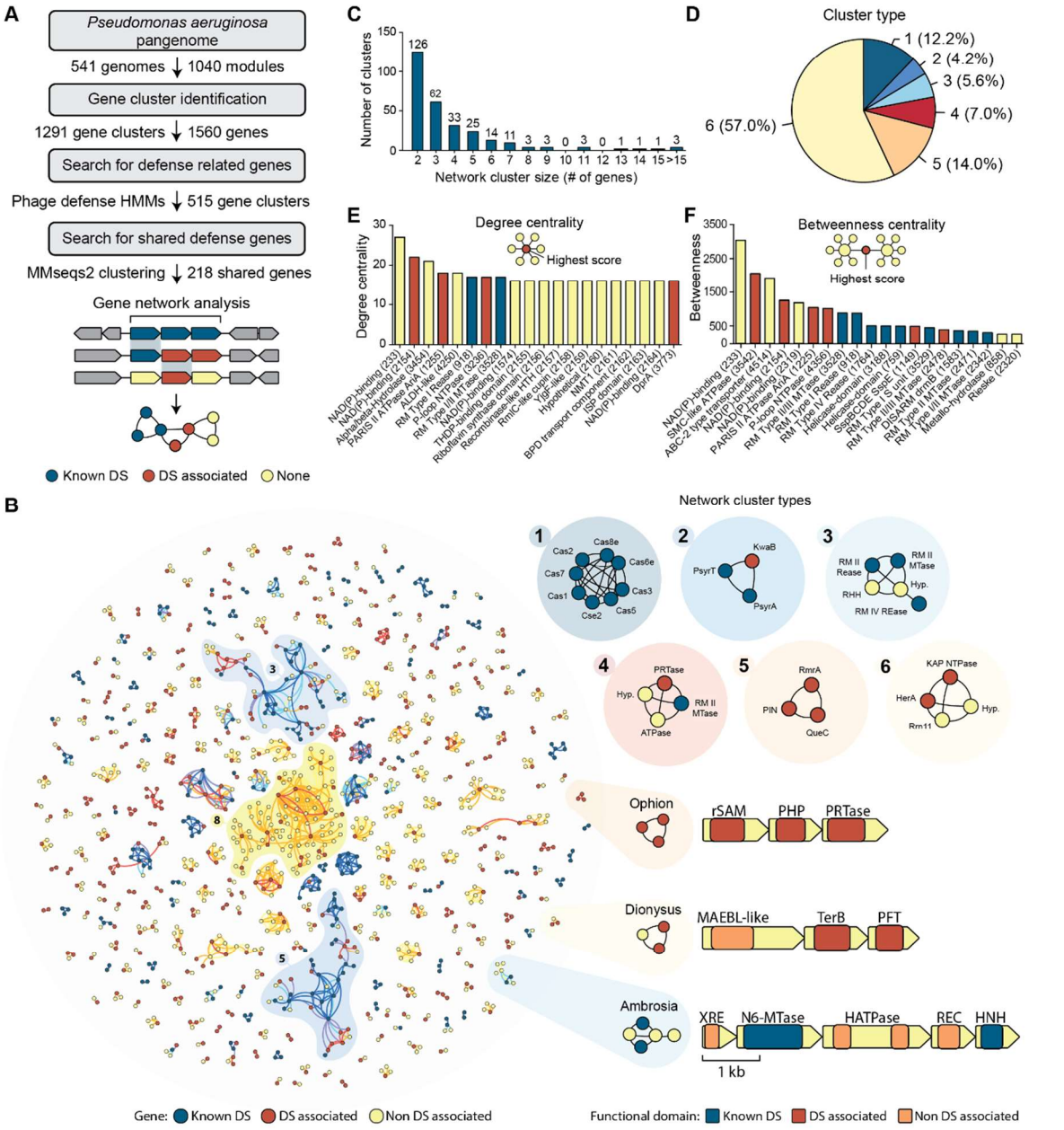

Figure 1. The gene network of phage defense system components in *P. aeruginosa*. (a) Identification of gene clusters located in regions of genomic plasticity that contain functional domains previously associated with anti-phage activity. Gene clusters containing anti-phage functional domains were searched for shared genes, followed by generating a network of common components among these gene clusters. **(b)** Network of shared components between gene clusters containing anti-phage functional domains. Six types of network clusters were observed, including: 1. Complete defense systems; 2. Complete defense systems with additional defense-associated genes; 3. Complete defense systems with additional genes not previously associated with defense; 4. Complete defense systems with additional genes both defense-associated and non-defense-associated; 5. Clusters containing defense-associated genes in new combinations; and 6. Clusters with both defense-associated and non-defense-associated genes. The defense system clusters on the bottom right show the architecture and functional domains of the validated defense systems in this study. Genes are drawn to scale, with the scale bar representing 1 kilo base pairs (kb). **(c)** The size distribution of each cluster within the network. **(d)** Pie chart depicts the frequency of each cluster type in the network. **(e)** Number of connections that a gene establishes within the network of gene clusters containing anti-phage components (degree centrality). A higher score corresponds to a gene that is found together with more genes, either by being present in a large gene cluster or by being shared among multiple gene clusters. **(f)** Number of bridges that a gene establishes between subnetworks (betweenness centrality). A higher score corresponds to a gene more frequently shared among different families of phage defense systems.

Table 1. Genes and functional domains of the phage defense systems identified in this study

Defense system	Gene	Functional domain(s)	Identifier
Dionysus	*dinA*	Merozoite adhesive erythrocytic binding protein (MAEBL)	cl34696, cl31754
		Transmembrane domain	-
		Domain with unknown function	CATH: No match
	dinB	Tellurite resistance protein TerB-like[a]	IPR029024
	dinC	α-helical pore-forming toxin (α-PFT)[a]	-
		Domain with unknown function	CATH: No match
Ophion	*opnA*	Radical S-adenosyl-L-methionine (rSAM)[a]	PF04055
	opnB	Polymerase and Histidinol Phosphatase (PHP)[a]	IPR004013
	opnC	Phosphoribosyl transferase domain (PRTase)[a]	PF00156
		Domain with unknown function (2x)	CATH: No match
Ambrosia	*abrR*	Xenobiotic Response Element (XRE) transcriptional regulator	cd02209
	abrA	DNA N6- adenine methyltransferase (MTase)[a]	PF05869
	abrB	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase (HATPase)	PF13589, COG0642
	abrC	Phosphoacceptor receiver (REC) domain	cd00156
		Winged helix-turn-helix (wHTH) domain	IPR036390
	abrD	HNH endonuclease[a]	PF13395

[a] Functional domain previously associated with phage defense

Dionysus is a TerB-containing defense system with activity against *Phikzvirus* jumbo phages

Dionysus is named after the god associated with winemaking and the celebration of life. This defense system consists of three genes encoding: DinA, featuring a merozoite adhesive erythrocyte binding protein (MAEBL) and a transmembrane (TM) domain; DinB, a TerB-like protein; and DinC, a pore-forming toxin (PFT) with four TM domains (**Figure 2a**). Dionysus can be found in both Gamma- (312 instances) and Betaproteobacteria (49 instances) (**Figure S1a, Table S2**).

Dionysus provides robust protection against jumbo *Chimallivirus* phiKZ and ϕPa36, both in efficiency of plating (>10^6-fold reduction) and liquid cultures (**Figure 2a,b**). The one-step growth curve of ϕPa36 in control and Dionysus cells revealed that the life cycle of the phage is severely compromised by the defense system, with a similar latency period but a significantly reduced burst size (82 ± 24 PFU/cell in

control cells versus 9 ± 2 PFU/cell in Dionysus-expressing cells) (**Figure 2c,d**). All three proteins of Dionysus are essential for its protective capacity, as mutating a conserved amino acid within the MAEBL-like domain of DinA (H326A), and deletion of DinB or DinC, resulted in the partial (DinA H326A) or complete loss of anti-phage activity (**Figure S1b**). Lastly, *in silico* co-folding with AlphaFold 3 and further analysis using AlphaBridge predicts DinC to interact with the C-terminus of DinA to form a complex (1:1 stoichiometry; piCSi: 0.77), and DinB to form a homodimer (piCSi: 0.83) (**Figure S1c**).

Dionysus genes *dinB* and *dinC* encode domains that have previously been associated with phage defense. DinB has a TerB domain, which was originally identified in the tellurite resistance gene cassette and believed to be involved in the bacterial stress response (*293, 294*). More recently, TerB domains were also found in phage defense systems such as Bunzi (BnzA) and Shango (SngA)(*160, 161*) (**Figure S2d**). It is still unclear how TerB domain-containing phage defense proteins convey phage protection. It has been proposed that this could be related to TerB proteins having amphitropic characteristics (*295*), meaning they can reversibly switch between the cytoplasm and membrane, depending on environmental signals or ligand binding. These previous observations suggest that DinB may be involved in membrane surveillance.

DinC contains a PFT domain, which has also been reported to act as the anti-phage effector in various defense systems, including CBASS, Pycsar, CRISPR-Cas, Retrons, and bGSDM (*10, 38, 161, 188, 296*). PFTs often form multi-subunit pores with a hydrophilic interior in the cellular membrane to depolarize the bacterium and block phage propagation and can be broadly classified into two categories based on their secondary structure: α-PFTs, which form α-helical pores, and β-PFTs, which form β-barrel pores (*297*). Inspection of the AlphaFold3 modelling of DinC revealed that it is an α-pore-forming toxin (α-PFT) consisting of multiple amphipathic α-helices, structurally similar to enterotoxins such as haemolysin BL (Hbl), non-haemolytic enterotoxin (Nhe), and cytolysin A (ClyA) (*298*) (**Figure S1e**).

Dionysus gene *dinA* encodes an MAEBL domain, not previously linked to phage defense. This domain is poorly understood, apart from its role in *Plasmodium* species, where MAEBL is a transmembrane protein that binds erythrocyte receptor(*299*). Therefore, DinA may act as a membrane-associated protein involved in detecting phage components during phage infection.

Based on the predicted function of DinA, DinB, and DinC, we hypothesized that Dionysus may convey phage protection by detecting membrane-associated processes of the jumbo phage through DinA and DinB, with the result of DinC forming pores in the cellular membrane to cause cell death and abort phage infection. To investigate this hypothesis, we assessed membrane permeabilization using propidium iodide (PI) staining. We observed that at 90 min post infection by jumbo phage ϕPa36, the vast majority of Dionysus cells were not lysed, suggesting that Dionysus does not act through damaging of the cellular membrane (**Figure S1f**). This is further supported by quantification of cell survival at both one hour and two hours post infection, where no control cells survived after 2 hours, while Dionysus cells remained viable (**Figure S1g**). These results show that DinC does not form pores in the host membrane. We therefore hypothesized that DinC might act on the early phage infection (EPI) vesicles that jumbo phages produce to protect their DNA from host defense systems before transitioning into their nucleus-like compartment (*300*). To explore this possibility, we performed confocal microscopy using 4′,6-diamidino-2-phenylindole (DAPI) to stain phage and host DNA, expecting that DNA foci appear at EPI vesicle and nucleus compartments. These compartments shield the phage DNA from DNA-targeting host defenses (*82*). This experiment revealed no EPI vesicles or nuclei in cells expressing Dionysus, while both compartments could be observed in control (empty plasmid) and ΔDinC mutant cells (**Figure 2e-g; S1h**). At 50 minutes post-infection, EPI vesicles were present in almost half of the control and ΔDinC-expressing cells (48% and 39%, respectively), while none could be detected for Dionysus-expressing cells (**Figure 2f**). Moreover, no nuclei were detected in Dionysus cells, while 24% control and 28% ΔDinC cells contained nuclei (**Figure 2g**). These results were not affected by differences in adsorption, which was found to be identical between control and Dionysus-expressing cells (**Figure S1i**). To further investigate the potential interaction of Dionysus

with the EPI vesicles of the Jumbo phage, we labelled either DinA or DinC with mNeonGreen and observed that, 15 min post infection, both proteins form foci where the EPI vesicles are formed, suggesting possible interactions between these proteins and further supporting the hypothesis of early activity of Dionysus on the EPI vesicle (**Figure 2h,i; Figure S1j)**. Together, these results show that Dionysus specifically acts on the phage-derived EPI vesicle upon phage detection, preventing progression of jumbo phage infection at the first stage in its infection cycle (**Figure 2j**).

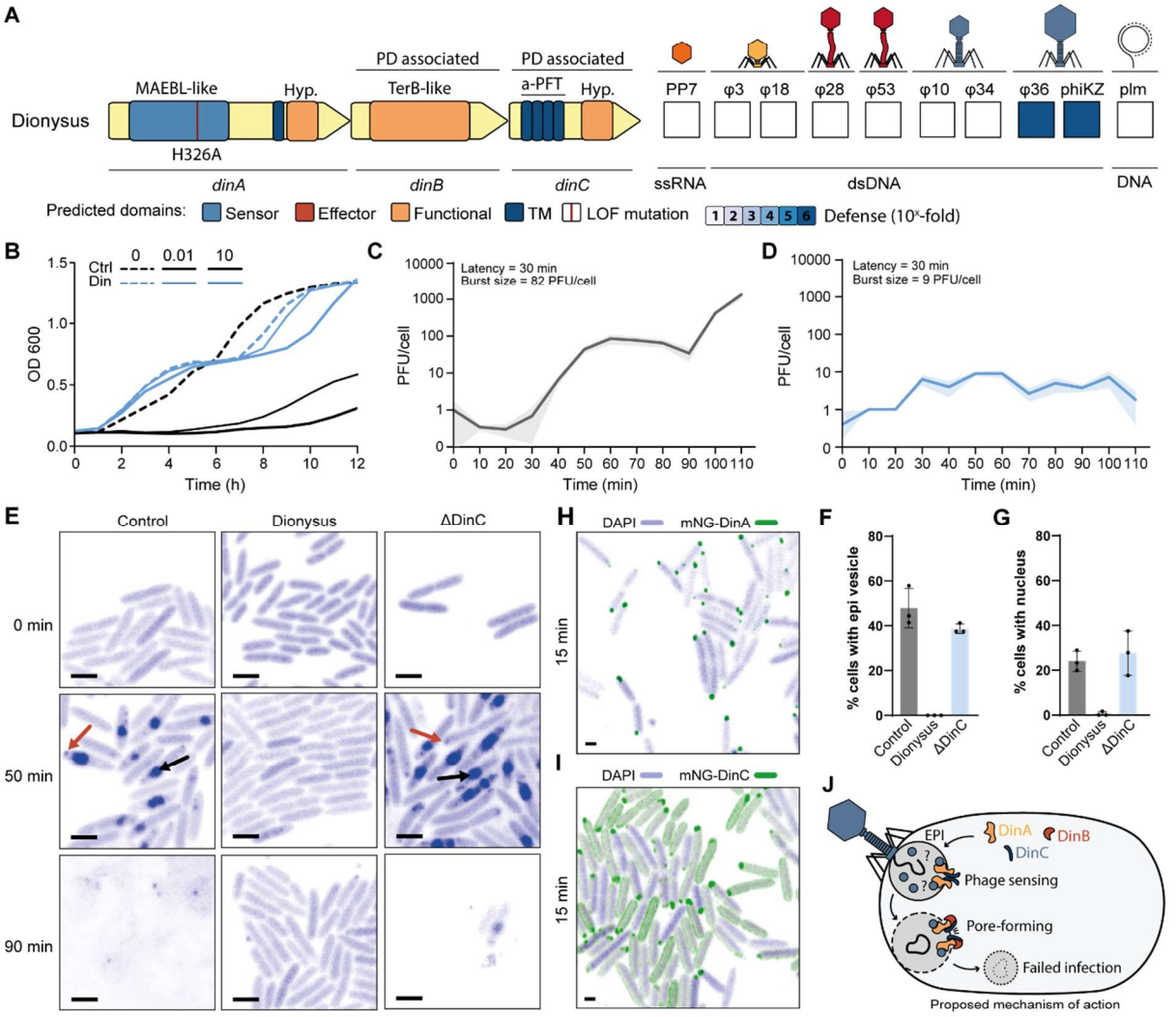

Figure 2. Dionysus is a jumbo phage specific defense system. (a) The functional domains of Dionysus, with loss of function mutation sites indicated as red lines. Genes are drawn to scale and indicated for their association with phage defense (PD associated). Dionysus was cloned with its native promoter into pUCP20 and introduced into *P. aeruginosa* strain PAO1, followed by assessing anti-plasmid activity (plm) by conjugation assays and anti-phage activity by efficiency of plating assays. The phage panel included *Fiersviridae* PP7, podophages *Autographiviridae* ϕPa3 and ϕPa18, siphophages *Casadabanvirus* ϕPa28 and *Mesyanzhinoviridae* ϕPa53, myophages *Pbunavirus* ϕPa10 and ϕPa34, and jumbo phages *Phikzvirus* ϕPa36 and phiKZ. **(b)** Effect of Dionysus on bacterial growth upon phage infection. *P. aeruginosa* strain PAO1 cells containing an empty plasmid (Ctrl) or Dionysus (Din) were infected with phage ϕPa36 at low (0.01) and high (10) multiplicity of infection, and their growth was monitored for a period of 12h. (**c,d**) One-step phage growth curve of ϕPa36 infecting PAO1 with (c) empty plasmid and (d) Dionysus. (e) Confocal microscopy images during ϕPa36 infection (0, 50, and 90 min) of PAO1 with empty plasmid, Dionysus, and a knock-out of DinC (ΔDinC). 4′,6-diamidino-2-phenylindole (DAPI) was used to stain phage and host DNA. The scale bar is 1 μm in length. (**f,g**) The percentage of observed (f) EPI and (g) nuclei during ϕPa36 infection of control, Dionysus-expressing, and ΔDinC cells. (**h,i**) Confocal microscopy images during ϕPa36 infection (15 min) of Dionysus cells with (h) DinA-mNeonGreen and (i) DinC-mNeonGreen. The scale bar is 1 μm in length. (j) Working model for phage defense by Dionysus, which is hypothesized to form pores in the EPI vesicle upon phage infection, protecting against jumbo phages.

Ophion is a radical SAM-containing anti-jumbo phage defense system proposed to act via nucleic-acid interfering mechanisms

Ophion, named after a primordial serpent from Greek mythology, is a three gene system encoding a radical S-adenosyl-L-methionine (rSAM) enzyme (OpnA), a polymerase- and histidinol-phosphatase (PHP, OpnB), and a phosphoribosyl transferase (PRTase, OpnC) (**Figure 3a**). Distributed across diverse Gammaproteobacteria (**Figure S2a; Table S2**), Ophion shows strong protection against jumbo phages ϕPa36 and PhiKZ, reducing phage infectivity by at least 10^6-fold in EOP assays and blocking phage propagation in liquid cultures (**Figure 3a,b**). The one step growth curve of ϕPa36 in Ophion cells showed no detectable burst, indicating that the phage life cycle is completely disrupted (**Figure 3c**). Mutating conserved residues of the rSAM and PRTase domains revealed these to be essential for protection, with partial dependence on the PHP domain (**Figure S2b**).

Each Ophion protein contains a domain associated with anti-phage mechanisms. For instance, the rSAM domain of OpnA is related to, but phylogenetically distinct from, those in prokaryotic Viperins (pViperins) (**Figure S2c**), which inhibit phage propagation by generating 3'-deoxy-3',4'-didehydro-nucleotides (ddh-NTPs). These nucleotide analogs terminate RNA synthesis when incorporated into phage transcripts (*301, 302*). The PHP domain of OpnB is found in the ppl phage defense system and in family X DNA polymerases where it is involved in DNA proofreading as a metal-dependent exonuclease of the phosphate backbone, removing misincorporated nucleotides (*122, 303, 304*). The PRTase domain of OpnC is known to play crucial roles in the biosynthesis of purine, pyrimidine, and pyridine nucleotides, as well as in purine and pyrimidine salvage (*305*). PRTases were previously identified as the phage defense effector of the phage defense type III Retron systems and as the toxin in the toxin-antitoxin phage defense systems ShosTA and PsyrTA (*160, 306*). These phage defense associated PRTases are phylogenetically related and located within the same clade of the PRTase family, suggesting a shared function (**Figure S2d**). Of these phage defense-associated PRTases, only the function of the ShosTA PRTase has been characterized, which disrupts the purine metabolism by producing nucleoside monophosphate, eventually disrupting DNA duplication of both the host and the phage (*307*). The Ophion proteins are not predicted to interact with each other (**Figure S2e**).

Together, these observations indicate that each protein encoded by Ophion contains functional domains that are known to interfere with nucleotide modification and salvage pathways. To gain a broader understanding of Ophion's potential mechanism of action, we set out to identify the stage of phage infection at which Ophion acts, using confocal microscopy to monitor ϕPa36 jumbo phage infection. In control cells and in cells expressing an OpnC mutant (K98A), the phage progressed normally through both early (EPI vesicle) and late (nucleus) stages. In contrast, in Ophion-expressing cells, ϕPa36 failed to develop nuclei (**Figure 3d-f**), indicating that infection is arrested early, before the translation of nucleus-forming components (*308*). Given the nucleotide-modifying potential of the rSAM domain of OpnA and PRTase of OpnC (*301*), it is possible that early phage transcription within the EPI vesicle is disrupted by incorporating transcription-halting nucleotides. Notably, Ophion does not seem to affect the host and does not trigger abortive infection (**Figure S2f**), implying a phage-specific mechanism.

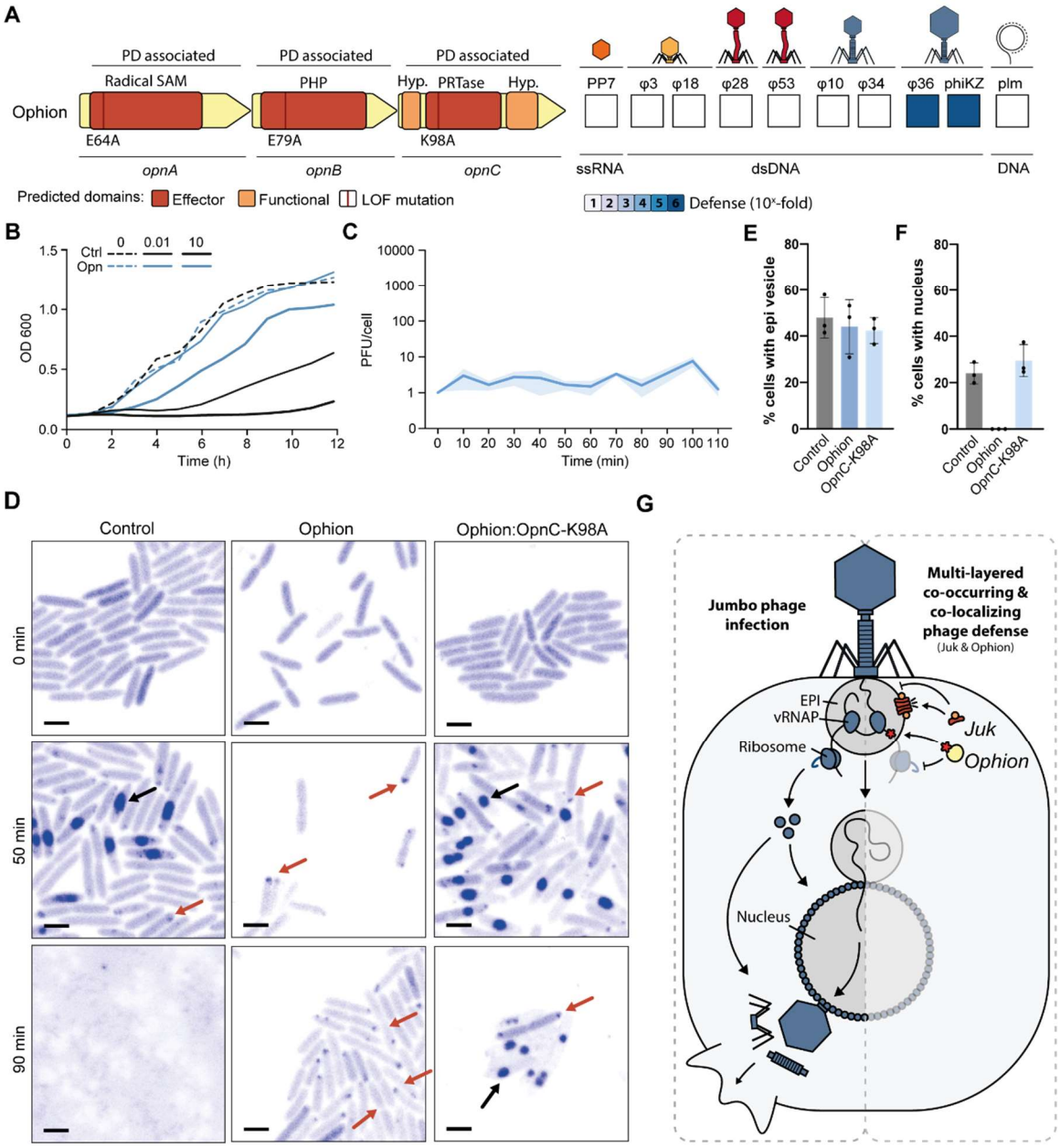

Figure 3. Ophion is a radical SAM-containing phage defense system proposed to act via nucleic-acid interfering mechanisms. (a) The functional domains of Ophion and the mutation sites (shown as red lines) that were tested. Genes are drawn to scale and indicated for their association with phage defense (PD associated). Ophion was assessed for anti-plasmid activity (plm) by conjugation assays and for anti-phage activity by efficiency of plating assays. The phage panel includes *Fiersviridae* PP7, podophage *Autographiviridae* ϕPa3 and ϕPa18, siphophage *Casadabanvirus* ϕPa28, siphophage *Mesyanzhinoviridae* ϕPa53, myophage *Pbunavirus* ϕPa10 and ϕPa34, and jumbo phages *Phikzvirus* ϕPa36 and phiKZ. **(b)** Effect of the defense system on bacterial growth upon phage infection. *P. aeruginosa* strain PAO1 cells containing an empty plasmid (Ctrl) or Ophion (Opn) were infected with phage ϕPa36 at low (0.01) and high (10) multiplicity of infection, and their growth was monitored for a period of 12h. **(c)** One-step growth curve of ϕPa36 infecting PAO1 with Ophion. (**d)** Confocal microscopy images during ϕPa36 infection (0, 50, and 90 min) of control (empty plasmid), Ophion, and OpnC mutant (K98A) cells. Both the bacterial and phage DNA were labelled using 4′,6-diamidino-2-phenylindole (DAPI). (**e,f**) The percentage of observed (e) EPI and (f) nuclei formed during ϕPa36 infection of control (empty plasmid), Ophion, and OpnC-K98A cells. **(g)** Proposed multilayered defense against jumbo phages. On the left side, jumbo phage progresses through its infection cycle, from formation of the EPI vesicle (early stage) to the transition into the nucleus stage, and the final assembly of the phage particles before lysing the host. On the right, jumbo phage infection is stalled at early stages when co-occurring and co-localizing jumbo phage defense systems Juk (targeting epi vesicle) and Ophion (likely targeting early phage transcripts) are present.

Notably, we observed that Ophion strongly co-localizes and co-occurs with Jumbo phage defense system Juk (*309*) in 98.2% of the cases (108 out of the 110 times Ophion is found in *P. aeruginosa*; Observed = 108; Expected = 23; Z-score = 18.25; Adjusted p-value = 0; Mean distance = 0 bp). Interestingly, Ophion and Juk convey protection against jumbo phages at different stages of their infection cycle. Specifically, Juk disrupts EPI vesicles which form early in the infection, while Ophion affects the phage in a later stage by blocking the progression from phage EPI vesicle to the nucleus stage. Combined, these systems seem to provide a multilayered defense to limit the escape possibilities of jumbo phages **(Figure 3g)**.

In summary, Ophion provides potent and specific defense against jumbo phages by blocking progression from the EPI vesicle to the nucleus stage, possibly by halting the transcription of the nucleus forming genes. The co-occurrence of Ophion with the EPI vesicle-targeting system Juk strongly suggests beneficial effects of targeting distinct features and processes of early jumbo phage infection.

Ambrosia combines features of RM and two-component regulatory systems

Ambrosia, named after the food and drinks that provide immortality to the Greek gods, consists of five genes and provides protection against siphophages from the *Mesyanzhinovviridae* family and myophages from the *Pbunavirus* genus, limiting phage propagation in both solid and liquid cultures (**Figure 4a,b**). One step growth curve of ϕPa34 in control and Ambrosia cells revealed that the life cycle of the phage is severely compromised by the defense system, with a burst size significantly reduced from 165 ± 18 PFU/cell in control cells to undetectable in Ambrosia-expressing cells (**Figure 4c,d)**. Ambrosia is found across Gammaproteobacteria, with most of its occurrences (76%) in Pseudomonadales (**Figure S3a; Table S2**).

The five genes of Ambrosia encode a HipB-like XRE transcriptional regulator (AbrR), a RM type II-associated N6 adenine-specific DNA methylase (MTase, AbrA), a sensor histidine kinase/heat-shock-protein(HSP)90-like ATPase (HATPase, AbrB), a response regulator (REC, AbrC), and an RM type IV HNH endonuclease (AbrD) (**Figure 4a; Figure S3b,c**). Each protein is required for anti-phage activity, since mutating conserved residues in any of the predicted functional domains completely abolishes Ambrosia defense activity (**Figure S3d**).

Ambrosia contains both elements from an RM system (MTase and HNH endonuclease (*310*)), which are known to be involved in phage defense, and a two-component regulatory system (HATPase and REC, a stimulus-response coupling mechanism (*311*)). Previous studies have reported that two-component systems often localize near RM systems, particularly RM types II and IV (*312, 313*), but their integration in a single gene cluster was not previously described. It has been proposed that the two-component system regulates the transcription of the RM systems, with the sensor kinase phosphorylating the response regulator to initiate transcription in response to specific environmental stimuli (*314*). We hypothesized that this could also be the case for Ambrosia, especially since the REC-encoding AbrC contains a winged-helix-turn-helix (wHTH) domain, which is associated with these transcription initiating RECs (*315*). In the context of Ambrosia, this suggests that AbrB detects phage infection and then phosphorylates AbrC, triggering a downstream transcriptional response to phage infection. The importance of AbrC phosphorylation is supported by the loss of phage protection when the predicted phosphorylated site (D60) was mutated to alanine (D60A) (**Figure S3d**). In addition, we quantified the expression of mNeonGreen inserted into the Ambrosia cluster in place of AbrA or AbrD and observed that these regions were barely expressed in normal conditions but highly expressed during phage infection (**Figure S3e**). These results suggest that the expression of Ambrosia proteins is regulated at the transcription level.

In addition, Ambrosia expression may also be controlled by the HipB-like XRE transcriptional regulator AbrR, as these transcriptional regulators are known to inhibit transcription, especially of their own gene clusters (*316*). In the context of RM systems, these XRE transcriptional repressors are known as control (c) proteins, which prevent the expression of the nuclease before the methylase has had time to methylate the host DNA (*317*).

The apparent need for multiple transcriptional regulators to control Ambrosia gene expression suggests that its RM-like components could be toxic to the host if expressed inappropriately and are therefore only expressed upon phage infection. Consistent with this idea, AbrA is predicted to be an N6 DNA methyltransferase and AbrD belongs to the RM type IV HNH endonuclease family, which typically cleaves methylated DNA. This raises the possibility that AbrD could act on modifications introduced by AbrA and, if expressed outside infection, damage host DNA. In classical type II RM systems, by contrast, the methylase protects host DNA from nuclease activity (*318*).

To test whether AbrA methylation might provide protection from the nuclease of Ambrosia, we propagated phages in *P. aeruginosa* PAO1 expressing only AbrA from an inducible promoter and then challenged these phages against PAO1 carrying the complete Ambrosia system. In line with the hypothesis, phages remained sensitive, indicating that AbrA methylation does not protect against AbrD activity (**Figure S3f,g**). Interestingly, we did not observe toxicity of Ambrosia to the host during phage defense, which would have been expected if AbrD efficiently targeted the host chromosome (**Figure 4e**). One possibility is that AbrD activity is more damaging to phage DNA than to host DNA, as nicking nucleases may disproportionately affect linear genomes compared with circular chromosomes (*319*).

Overall, our results indicate that the RM-like components of Ambrosia do not operate as a conventional type II RM system. Instead, Ambrosia appears to represent a distinct, tightly regulated phage defense system in which a HipB-like XRE protein and a two-component regulatory module control the expression of RM-like enzymes in response to infection (**Figure 4F**).

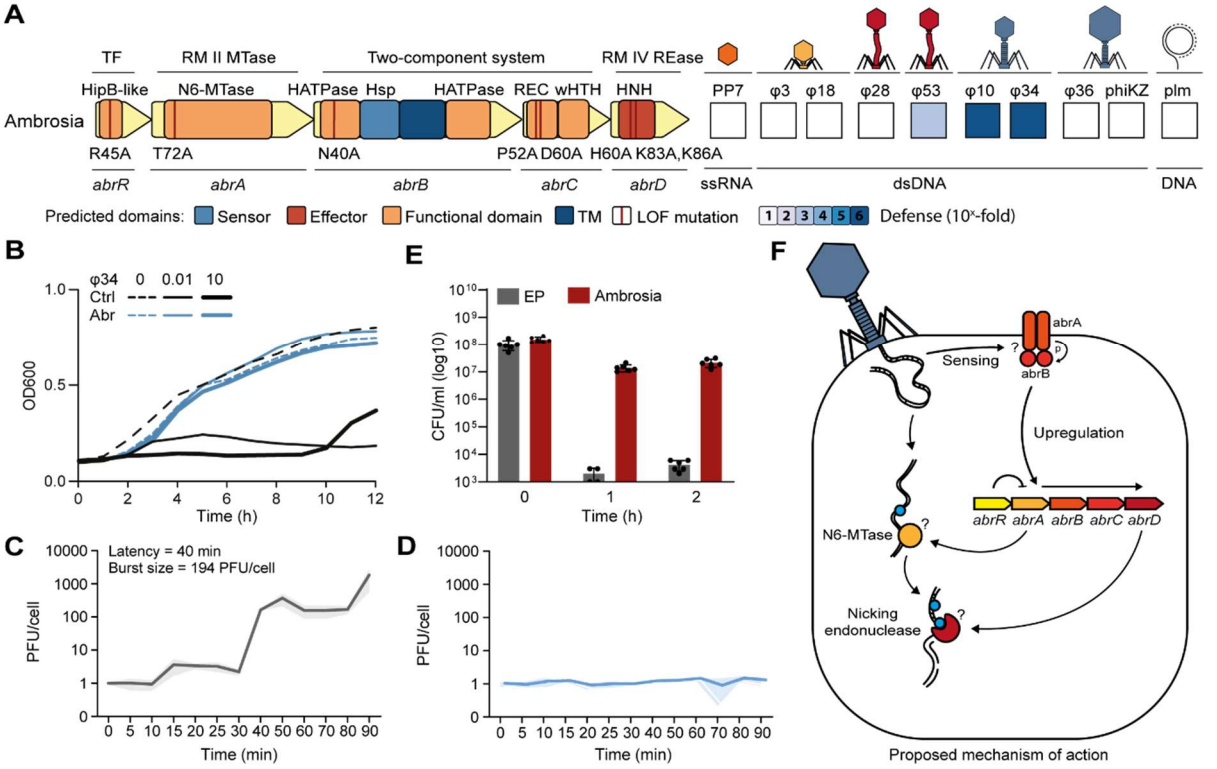

Figure 4. Ambrosia is a phage defense system that combines components of a restriction-modification system and a two-component system. (a) The functional domains of Ambrosia with the mutation sites tested indicated as red lines. Genes are drawn to scale and indicated for their association with phage defense (PD associated). Ambrosia was assessed for anti-plasmid activity (plm) by conjugation assays and for anti-phage activity by efficiency of plating assays. **(b)** Effect of Ambrosia on bacterial growth upon phage infection. P. aeruginosa strain PAO1 cells containing an empty plasmid (Ctrl), or Ambrosia (Abr) were infected with phage ϕPa34 at low (0.01) and high (10) multiplicity of infection, and their growth was monitored for a period of 12 h. (c,d) One-step growth curve of ϕPa34 infecting **(c)** control (empty plasmid) and **(d)** Ambrosia-expressing cells. **(e)** Assessment of cell concentration (CFU/ml) in phage-infected cultures of control (empty plasmid, EP) or Ambrosia-expressing cells. Bars represent average values of at least three biological replicates with individual points overlaid. **(f)** Working model for anti-phage defense by Ambrosia. Ambrosia is proposed to represent a defense system where a HipB-like XRE gene together with a two-component regulatory mechanism appears to control the expression of an RM-like system that targets its own DNA modification in response to phage infection.

Discussion

Evolutionary adaptation often proceeds from evolutionary tinkering processes, where existing genetic elements are repurposed or recombined (*320*). Key to this process is the modular organization of genes and their constituent domains, which enables rapid exchange of distinct functional units among unrelated pathways in a "plug-and-play" fashion (*320*). This is especially well exemplified by the core glycolytic genes, which are recruited across diverse metabolic pathways, including amino acid biosynthesis and aerobic respiration (*321, 322*).

Likewise, phage defense systems of bacteria exhibit high degrees of modularity, with sensing, signal transmission, and effector enzymes exchanged between phage defense systems (*238*). This modularity may allow rapid adaptation to phage countermeasures (*323-325*), for instance by replacing the sensory or effector domain that are inhibited by phage encoded anti-genes. In turn, the phage overcomes this adaptation, resulting in a continuous evolutionary race between the phage and the phage defense systems of the host.

In this study, we capitalized on the modularity of phage defense systems to discover three previously unknown defense systems in *P. aeruginosa* that incorporated genes or functional domains previously associated with phage defense in new architectures. Our discovery approach offers several powerful features. First, by searching for components already associated with phage defense, we increase the likelihood of finding positive candidate defense systems, independently of their abundance across bacterial genomes. Second, previously unknown anti-phage genes and functional domains can be integrated into subsequent searches, creating a feedback loop that continually expands our understanding of the phage defense repertoire. Lastly, this growing catalogue enables deeper insight into the evolutionary relationships between defense systems and the modular components they share. For example, many of the most frequently shared elements are RM components, and one of the systems uncovered by our approach, Ambrosia, combines an unusual RM-like module with regulatory genes.

The defense systems that we discovered using this approach include also Dionysus and Ophion. Dionysus contains a TerB-like and a pore-forming domain shared with several other defense systems but also contains an additional gene with a MAEBL functional domain that was not associated with phage defense before. Ophion features three genes with phage defense associated functionalities that have not previously been found in this combination. This system includes a rSAM, which is related to bacterial Viperins that convey phage defense (*166, 301*), along with PHP and PRTase domains. Both Dionysus and Ophion specifically block infection by jumbo phages, which have evolved specialized compartments to protect their DNA from DNA-targeting host defenses. After entry, the phage genome is enclosed by a membranous EPI vesicle, where early transcription occurs. At a later stage of the infection, the content of the EPI vesicle is transferred into a nucleus-like compartment where DNA replication and late transcription (*300, 326-329*). While these compartments provide jumbo phages with an advantage against some host defenses, our findings show that they can also be exploited as a vulnerability targeted by phage defenses such as Dionysus and Ophion.

Specifically, we show that Dionysus forms pores in the EPI vesicle of the jumbo phage to disrupt the infection cycle of the phage at an early stage, similar to the mechanism described for the Juk phage defense system(*309*). In addition, we show that Ophion specifically prevents the progression of nucleus formation, potentially by halting phage transcription in the EPI vesicle stage. Together with jumbo phage specific Avs5 defense, which recognizes an early-expressed jumbo phage protein (*330*), evidence is growing of the existence of a jumbo phage specific defense system repertoire. We also observed that Ophion almost exclusively co-localizes and co-occurs with the Juk phage defense system, suggesting evolutionary advantages of co-targeting.

Altogether, we explored the modularity of phage defense components across defense systems, expanding the known network of phage defense-associated domains employed by bacteria, and highlighting the diverse molecular functions recruited for phage defense.

Limitations of the study

This study demonstrates the utility of gene modularity networks for uncovering new phage defense systems, but several limitations remain. First, while we show that Dionysus and Ophion arrest jumbo phage infection at distinct stages, the exact molecular processes underlying these effects remain to be elucidated, and future biochemical and structural work will be required to define the effector activities in detail. Second, our analysis of Ambrosia suggests regulation of an RM-like module by a two-component system, but it is not yet clear whether the effector activity primarily targets phage DNA, host DNA, or both, and how host toxicity is avoided. Third, our validation was limited to *P. aeruginosa* and a subset of phages. The broader distribution, mechanistic diversity, and ecological roles of these systems remain to be determined. Finally, while our modularity-based search expands discovery potential, it inevitably misses defense systems that lack known components or occur outside the networks captured here.

Materials and Methods

Bacterial strains and bacteriophages

The candidate defense systems were amplified from *P. aeruginosa* clinical isolates obtained from the University Medical Centre Utrecht (*146*). pUCP20-based plasmids containing the defense systems were cloned in *Escherichia coli* strain Dh5α and subsequently in *P. aeruginosa* strain PAO1. *E. coli* strain ET12567/pUZ8002 was used in conjugation assays. All bacterial strains were grown in Lysogeny Broth (LB) at 37 °C and 180 rpm, or in LB agar (LBA, 1.5 % agar (w/v)) plates at 37 °C, unless stated otherwise. *E. coli* ET12567/pUZ8002 was grown in media supplemented with 50 µg/ml of kanamycin and 25 µg/ml of chloramphenicol. Strains containing plasmid pUCP20 were grown in media supplemented with 100 µg/ml of ampicillin (for *E. coli*) or 200 µg/ml of carbenicillin (for *P. aeruginosa*). Strains containing plasmid pSTDesR (*130*) were grown in media supplemented with streptomycin at 25 µg/ml. *Pseudomonas* phage PP7 was acquired from LGC Standards, and all the remaining phages were obtained from the Fagenbank (*146*). Phages were propagated in liquid media using *P. aeruginosa* strain PAO1 as the bacterial host. The resulting bacterial lysate was centrifuged at 3,000 × *g* for 15 min, filter-sterilized (0.2 µm PES), and stored at 4 °C until further use.

Identification of putative anti-phage gene clusters in variable regions of *P. aeruginosa*

We used PPanGGOLiN v1.2.74 (*248, 249*) to identify conserved gene clusters in the variable regions of all 541 complete *P. aeruginosa* assemblies from the RefSeq database as of June 16, 2022. Genes in these conserved gene clusters were analysed for their proximity to one another using cblaster v1.3.19 (*270*), applying a threshold of 100-base-pair distance. Gene clusters encoding known and complete defense systems were identified using DefenseFinder v1.0.9 with defense-finder-models v1.2.2 (*63*). Additionally, the defense-finder models were used to identify incomplete phage defense systems. Lastly, anti-phage-related functional domains were identified in the sub-selected gene clusters using the Pfam-A (*331*) and Superfamily HMM models (*332*) using HMMer (*127*) and Interproscan v5.60-92.0 (*333*) respectively, from a list of anti-phage-associated functional domains compiled from a literature search (**Table S3**). R package visNetwork (https://datastorm-open.github.io/visNetwork/) and Gephi (*334*) were used to perform a network analysis of the multi-gene clusters that contained anti-phage associated functional domains in addition to other functional domains previously not associated with phage defense. Confirmed anti-phage gene clusters were further annotated using HHpred, InterProScan (*333*), and CDD (*128, 335, 336*). Moreover, each protein was also predicted using AlphaFold3 (*257, 258*), searched for homologs with Foldseek (*230*), and further analyzed for functional domains using CATH-AlphaFlow 2024 (*337*) domain predictions. Signal peptides and transmembrane regions were annotated using Phobius v1.01 and SignalP v6.0 (*253, 254, 257, 258, 338*).

Cloning of candidate defense systems

The candidate defense systems were amplified from *P. aeruginosa* strains using Q5 DNA Polymerase (New England Biolabs) with primers from **Table S4**. The PCR products were run on 1% agarose gels, and bands were excised and cleaned using the Zymoclean Gel DNA Recovery Kit (Zymo Research). Plasmid pUCP20 was digested with BamHI and EcoRI (NEB), dephosphorylated with FastAP (Thermo Scientific), and cleaned with the Zymo DNA Clean & Concentrator Kit (Zymo Research). The defense systems were cloned into the digested pUCP20 using the NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs) and transformed into chemically competent NEB® 5-alpha Competent *E. coli* (New England Biolabs) following the manufacturer's instructions. Plasmids (**Table S5**) were extracted using the GeneJET Plasmid Miniprep kit (Thermo Scientific) and confirmed by Sanger sequencing at Macrogen. Confirmed plasmids were transformed into *P. aeruginosa* strain PAO1 by electroporation as previously described (*129*) and the cells were plated on LBA plates supplemented with 200 µg/ml of carbenicillin.

Selection and cloning of point mutations of the defense systems

Loss of function mutagenesis candidates were determined using a combination of literature search, multiple protein alignments using PSI-BLAST and ClustalW v2.1 (*256, 339*), and Alphafold3 structural prediction (*257, 258*). Gene knockouts and point mutations of the defense systems in pUCP20 were obtained by round-the-horn site-directed mutagenesis using Q5 polymerase with the phosphorylated primers indicated in **Table S4**. The PCR products were digested with DpnI (New England Biolabs) to remove methylated template DNA and run on 1% agarose gels. The bands of the expected size were excised and cleaned with the Zymo Gel DNA Recovery Kit and ligated with T4 DNA ligase (New England Biolabs) at room temperature for 2 hours. The ligated products were transformed into chemically competent NEB® 5-alpha Competent *E. coli* following the manufacturer's instructions. Plasmids (**Table S5**) were extracted and sequenced as indicated above and transformed into *P. aeruginosa* strain PAO1 by electroporation.

Efficiency of plating

Ten-fold serial dilutions of phage stocks were spotted onto double-layer agar (DLA) plates of *P. aeruginosa* strain PAO1 with empty pUCP20 or pUCP20 with the candidate defense systems following the small plaque drop assay (*131*). The anti-phage activity of the systems was calculated as the fold reduction in phage infectivity of *P. aeruginosa* strain PAO1 that contains the defense system, compared to the phage infectivity of *P. aeruginosa* strain PAO1 containing the empty plasmid.

Infection dynamics of phage-infected cultures

Overnight bacterial cultures of *P. aeruginosa* strain PAO1 with either empty or defense-containing pUCP20 were adjusted to an OD_{600} of 0.1 in LB and subjected to phage infection at an MOI lower than 1. The cultures were incubated at 37 °C and 180 rpm. Samples were collected at 0h, 2h, 4h, and 6h, followed by centrifugation at 3,000 × *g* for 5 min. The resulting phage-containing supernatant was serially diluted and spotted onto DLA plates of *P. aeruginosa* strain PAO1 to estimate phage concentration.

Liquid culture collapse assays

Bacterial cultures grown overnight were diluted to an OD_{600} of 0.1 in LB and dispensed into 96-well plates. Phages were introduced at MOIs of 0.01 and 10, and the plates were incubated at 37 °C in an Epoch2 microplate spectrophotometer (Biotek). OD_{600} measurements were taken every 10 min over a 24h period, with double orbital shaking.

Plasmid conjugation

Plasmid pSTDesR was conjugated into *P. aeruginosa* strain PAO1 containing empty pUCP20 or pUCP20 with individual defense systems using the puddle matting approach as previously described (*340*). Briefly, overnight cultures of the strains were diluted 1:2 in LB and incubated at 42 °C for a minimum of 3h. Then, 0.5 ml of the *P. aeruginosa* cultures were mixed with 1.5 ml of an exponentially grown

culture of *E. coli* ET12567/pUZ8002 containing pSTDesR. The cells were collected by centrifugation at 10,000 × *g* for 5 min, and resuspended in 50 µl of LB. The cell mixture was transferred onto the middle of a pre-warmed LBA plate and allowed to dry. The plates were incubated overnight at 30 °C. The bacterial puddle was then scaped off the surface of the LBA plate and resuspended in 1 ml of sterile phosphate buffer saline (1x PBS). Two-fold serial dilutions of this cell suspension were then platted on LBA plates containing either 100 µg/ml of carbenicillin (for total cell count) or additional 12.5 µg/ml streptomycin (for conjugants quantification). The plates were incubated for 48h at 37 °C.

Adsorption assays

P. aeruginosa strain PAO1 carrying either an empty plasmid or a plasmid expressing a defense system was grown to early exponential phase (optical density at 600 nm [OD_{600}] ≈ 0.3). Phages were added at a multiplicity of infection (MOI) of 0.01 and incubated at 37°C with shaking at 180 rpm for 10 minutes. Samples were collected at 0- and 10-minutes post-infection, centrifuged at 9,000 × *g*, and the supernatant, containing non-adsorbed phages, was serially diluted 10-fold. The dilutions were plated on PAO1 double-layer agar plates to quantify non-adsorbed phage titers.

One step growth curve

P. aeruginosa strain PAO1 carrying either an empty plasmid or a plasmid expressing a defense system was grown to an OD_{600} of 0.3–0.4. Cultures were centrifuged at 3,200 × *g* for 10 minutes, and the cell pellet was resuspended in LB to half the original culture volume. Phages (ϕPa34 or ϕPa36) were added at a multiplicity of infection (MOI) of 0.01 and allowed to adsorb for 10 minutes at 37°C with shaking at 180 rpm. After adsorption, cultures were centrifuged, and the cell pellet was resuspended in LB to the original culture volume. Cultures were incubated at 37°C with shaking, for 90 minutes (Ambrosia) or 110 minutes (Ophion and Dionysus). For Ambrosia, samples were taken at time 0, every 5 minutes for the first 30 minutes, and then every 10 minutes thereafter. For Ophion and Dionysus, samples were collected at time 0 and at 10-minute intervals. Samples were immediately serially diluted 10-fold and spotted on PAO1 double-layer agar plates for phage quantification.

Confocal microscopy

Confocal microscopy was performed as previously described (*330*). *P. aeruginosa* strain PAO1 carrying either an empty plasmid or a plasmid expressing Dionysus or Ophion was grown to an OD_{600} of 0.3–0.4. Cells were infected with phage ϕPa36 at an MOI >3, and adsorption was allowed for 10 min at 37°C, 180 rpm. Cells were then centrifuged at 12,000 × *g* for 1 min to remove non-adsorbed phages, resuspended in LB, and incubated at 37°C, 180 rpm. Following incubation at different time points, cells were pelleted again and resuspended in 100 µl of LB. DAPI (to stain DNA) was added to final concentrations of 0.1 mg/ml. For cell viability staining, propidium iodide was added at a final concentration of 1 µM. Stained cells were spotted onto 1% agarose pads (*136*) and imaged using a Nikon A1R/SIM laser scanning confocal microscope (inverted Nikon Ti Eclipse body) equipped with a 100× oil immersion objective (SR Apo TIRF; numerical aperture 1.49). DAPI was excited at 405 nm (emission filter: 450/50 nm), with signals passed through a 405/488/543/640 excitation dichroic mirror. Z-stacks were acquired using a Nikon A1 Piezo Z Drive at intervals of 0.1 or 0.2 µm (10–20 slices), capturing different focal planes of all bacteria in the field of view (512 × 512 pixels, corresponding to 36.79 × 36.79 µm, satisfying Nyquist criteria). A pinhole size corresponding to 1.2 Airy Units (AU), referenced to the shortest wavelength, was used. Images were acquired at 12-bit depth using a Galvano scanner with Nikon NIS-Elements software. Image analysis was performed using Fiji.

Presence of the novel phage defense systems in bacteria

To detect all instances of Dionysus, Ophion, and Ambrosia in bacteria, we applied cblaster v1.3.18 (*270*) to obtain all clusters in the RefSeq database (minimum identity 20%). The taxonomy of all Dionysus, Ophion, and Ambrosia instances by cblaster was based on the annotation provided by RefSeq. Next, we built HMM profiles for the individual proteins of these defense systems using MUSCLE (-super5) v5.1 (*251*) to align the obtained protein sequences, and hmmbuild v3.3.2 (*127*) to build the HMM models. HMM models were searched against protein sequence databases using hmmsearch (*127*). The HMM model scoring thresholds were set based on the 100% sensitivity point with the help of an ROC curve analysis that scored the HMM sensitivity for the defense protein compared to all other proteins within *P. aeruginosa*. The HMM bitscore obtained for each defense system protein is as follows: Dionysus: DinA, 400, DinB, 100 and DinC, 200; Ophion: OpnA, 500, OpnB, 350, and OpnC, 450; and Ambrosia: AbrA, 100, AbrB, 500, AbrC, 150, and AbrD, 200.

Co-occurrence analysis

A pairwise co-occurrence analysis was performed using the presence and absence of all known jumbo phage defense systems, including Juk, Dionysus, CBASS type III, Avs5, 6A-MBL, and Ophion in the set of 541 complete *P. aeruginosa* assemblies from the RefSeq database as of June 16, 2022. The Z-scores and associated p-values were calculated based on the expected co-occurrence of these phage defense systems with the assumption of independence(*341*). These p-values were then adjusted using the Bonferroni correction (*342*).

Phylogenetic analysis of shared phage defense genes

Phylogenetic trees of AbrA (N6-Methylase), AbrD (HNH), DinB (TerB), OpnB (PHP), and OpnC (PRTase), were made by obtaining all available Uniprot Release 2024_04 (*343*) proteins containing the corresponding functional domain. The phylogenetic tree of OpnA (rSAM) was built using the radical SAM containing protein list from Bernheim *et al.* (2021) (*301*). Duplicate sequences were removed, and the left-over proteins were subsequently down sampled, aligned, and trimmed as previously described (*138*). The resulting alignments were used to build a phylogenetic tree using FastTree v2.1.11 (*277*) with default settings and visualized, without branch lengths, using the interactive Tree of Life (iTol) v6 (*344*). All previously known defense genes were then searched using the HMM profiles of Defense Finder Models v1.3.0 and PADLOC-DB v2.0.0 (*121, 345*). A bitscore higher than 100 was seen as significant.

Statistical analysis

Unless stated otherwise, data are presented as the means of biological triplicates ± standard deviation. A Bonferroni-adjusted p-value of less than 0.05 was considered significant.

Supplemental information

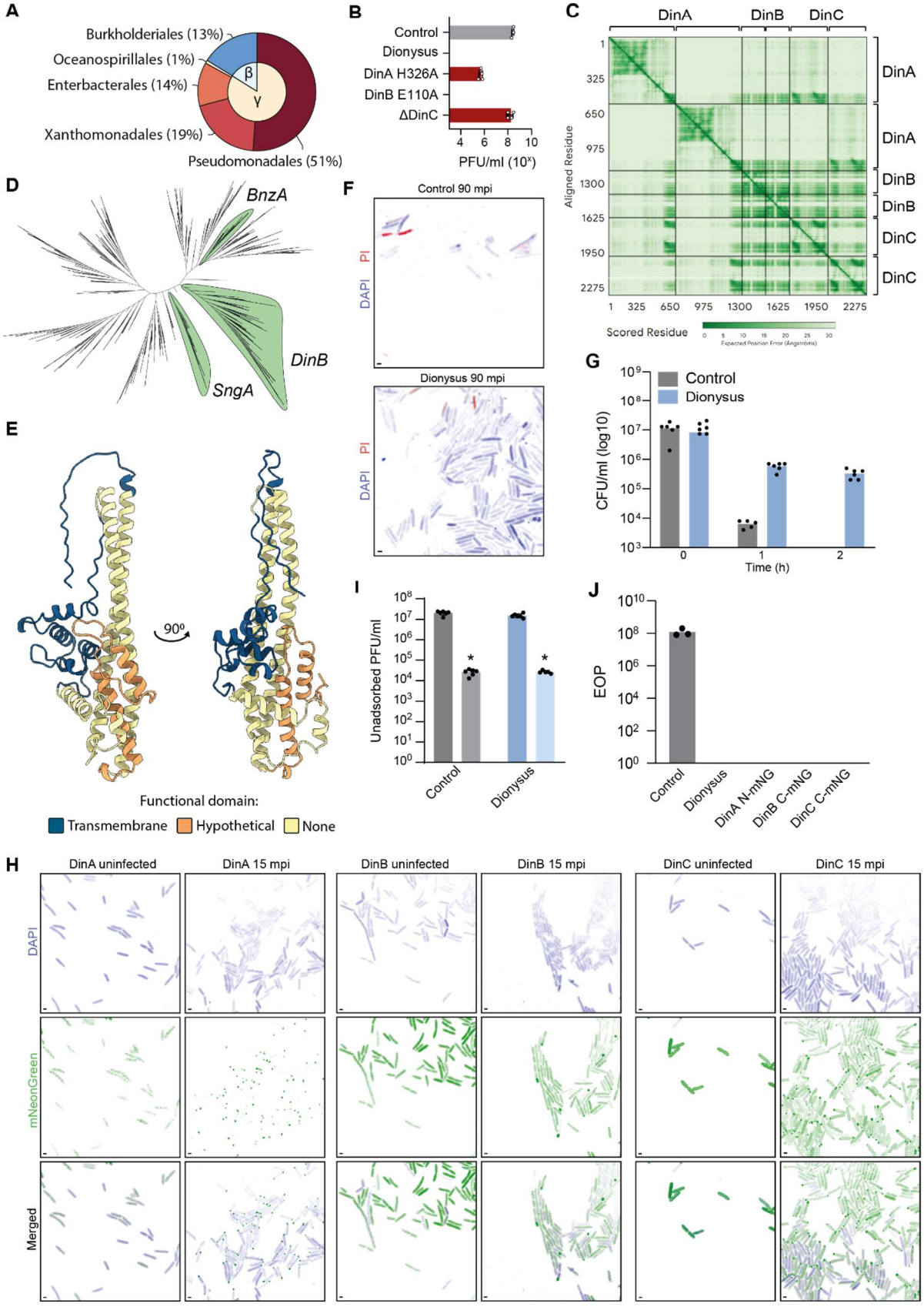

Figure S1. Characteristics of the phage defense system Dionysus, Related to Figure 2. (a) The prevalence of Dionysus within Prokaryotes. Dionysus was found in a total of 713 instances, and its distribution (in percentage of total) in

different orders is shown in the outer circles of the graphic, while the inner circle indicates the classes. γ: Gammaproteobacteria. β: Betaproteobacteria. **(b)** Effect of gene deletions and mutations in the functional domains of Dionysus on phage protection. The infectivity of phage ϕPa36 on *P. aeruginosa* strain PAO1 cells containing an empty plasmid (control), Dionysus, or Dionysus with point mutations or deleted genes was assessed by plaque assay. Bars represent average values of at least three biological replicates with individual points overlaid. **(c)** PAE plot obtained by co-folding the three Dionysus proteins (two copies of each) with AlphaFold3. **(d)** Phylogenetic tree of TerB containing proteins. The phylogenetic tree of 1406 representative proteins were inferred using Fasttree2. The representative proteins include all groups within the TerB-like family (IPR029024), as indicated in the tree. **(e)** Predicted Alphafold3 structure of alpha-PFT DinC. **(f)** The membrane permeability of control and Dionysus-expressing cells 90 min post infection with ϕPa36, assessed with propidium iodide (PI) staining. **(g)** The cell concentration (CFU/ml) of PAO1 containing an empty plasmid (control) or Dionysus, during ϕPa36 infection at an MOI of 10. **(h)** Confocal microscopy images during ϕPa36 infection (0 and 15 min) of PAO1 expressing Dionysus with fluorescently labelled mNeonGreen-DinA, mNeonGreen-DinB, or mNeonGreen-DinC, stained with DAPI. Shown are the separate captured images for DAPI, and mNeonGreen, as well as the merged images. **(i)** Adsorption assay of ϕPa36 on PAO1 containing an empty plasmid (control) or Dionysus. **(j)** EOP assay of ϕPa36 in PAO1 containing an empty plasmid (control), Dionysus, or Dionysus with mNeonGreen (mNG) fused to N- or C-terminus of each Dionysus protein (DinA N-mNG, DinB C-mNG, and DinC C-mNG).

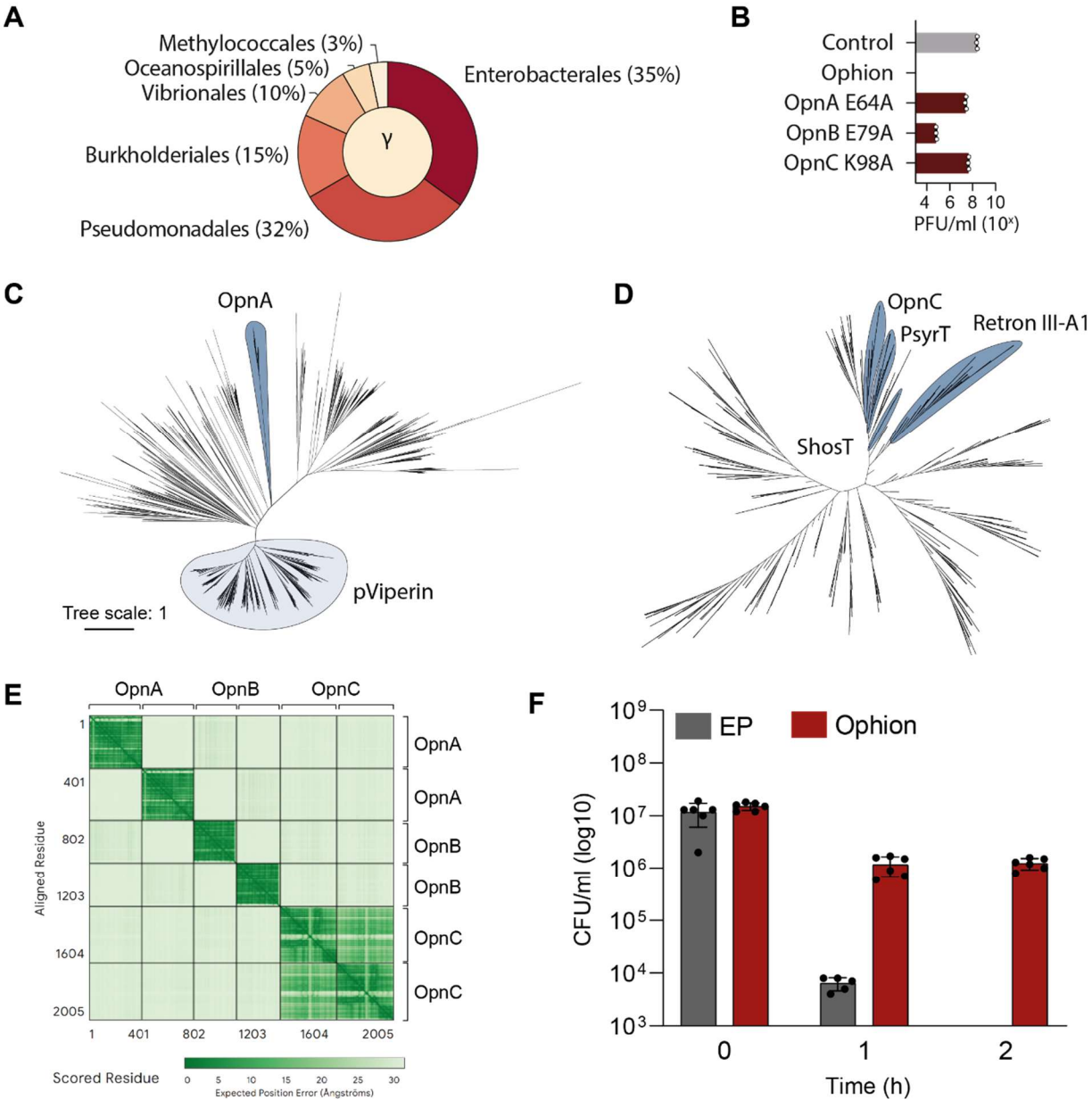

Figure S2. Characteristics of the phage defense system Ophion, Related to Figure 3. (a) The prevalence of Ophion in prokaryotes. Ophion was found in a total of 4103 instances, and its distribution (in percentage of total) in different orders is shown in the outer circles of the graphic. γ: Gammaproteobacteria. **(b)** Effect of mutations in the functional domains of Ophion on phage protection. The infectivity of phage ϕPa36 on *P. aeruginosa* strain PAO1 containing an empty plasmid (control), Ophion, or Ophion with point mutations was assessed by plaque assay. Bars represent average values of at least three biological replicates with individual points overlaid. **(c)** Phylogenetic tree of radical SAM containing proteins. The phylogenetic tree of 1976 representative proteins was inferred using Fasttree2. The representative proteins include all groups within the radical SAM superfamily (PF04055), as indicated in the tree. **(d)** Phylogenetic tree of PRTase-containing proteins. The phylogenetic tree of 1174 representative proteins was inferred using Fasttree2. The representative proteins include all groups within PFAM clan0533, as indicated in the tree. **(e)** PAE plot obtained by co-folding all Ophion proteins (two copies each) using AlphaFold3. (f) The cell concentration (CFU/ml) of PAO1 containing an empty plasmid (control) or Ophion, during infection of ϕPa36 at an MOI of 10.

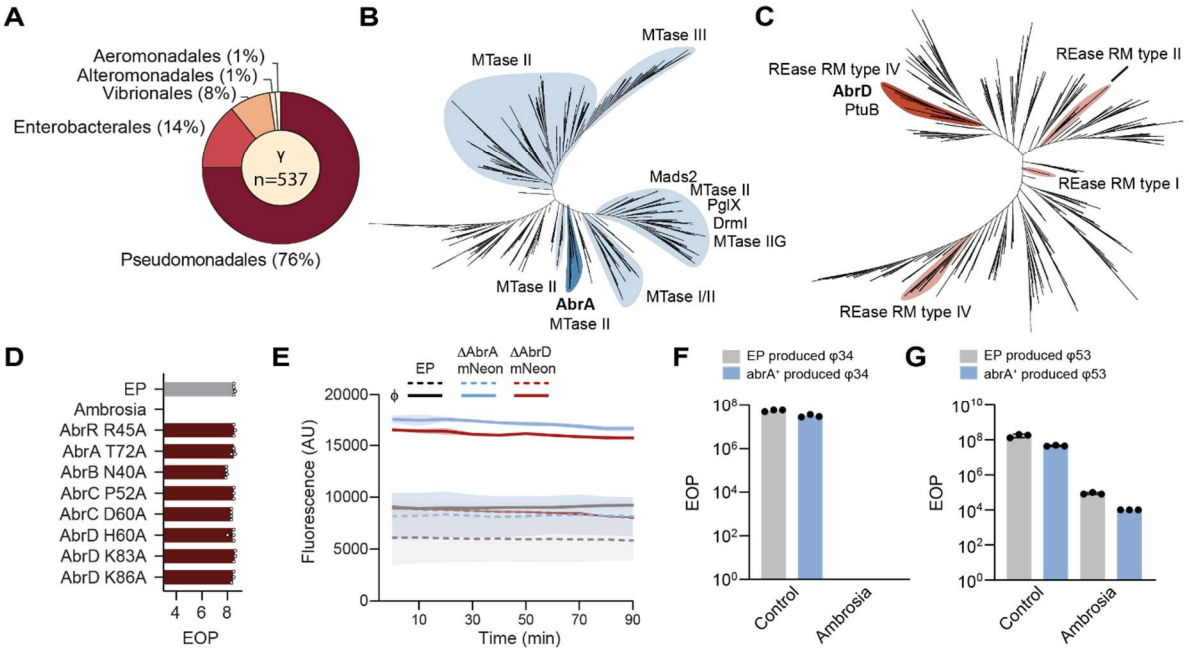

Figure S3. Characteristics of the phage defense system Ambrosia, Related to Figure 4. (a) The prevalence of Ambrosia in prokaryotes. Ambrosia was found in a total of 537 instances, and its distribution (in percentage of total) in different orders is shown in the outer circles of the graphic. γ: Gammaproteobacteria. **(b)** Phylogenetic tree of N6 MTase-containing proteins. The phylogenetic tree of 2066 representative proteins was inferred using Fasttree2. The representative proteins include all groups in N6-specific DNA methylases (IPR002052), as indicated in the tree. **(c)** Phylogenetic tree of HNH containing proteins. The phylogenetic tree of 1681 representative proteins was inferred using Fasttree2. The representative proteins include all groups in HNH (IPR002711), as indicated in the tree. **(d)** Effect of mutations in the functional domains of Ambrosia on phage protection. The infectivity of phage ϕPa34 on *P. aeruginosa* strain PAO1 cells containing an empty plasmid (control), Ambrosia, or Ambrosia with point mutations was assessed by plaque assay. Bars represent average values of at least three biological replicates with individual points overlaid. **(e)** Fluorescence measurements (AU) over 90 minutes of infection by ϕPa34 on *P. aeruginosa* strain PAO1 cells containing an empty plasmid (EP), Ambrosia with abrA replaced by mNeonGreen, or Ambrosia with abrD replaced by mNeonGreen. **(f)** EOP assay to assess the infectivity of ϕPa34 produced in PAO1 or PAO1:abrA$^+$, on PAO1 encoding an empty plasmid (control) or Ambrosia. **(g)** EOP assay to assess the infectivity of ϕPa53 produced in PAO1 or PAO1 abrA$^+$, on PAO1 encoding an empty plasmid (control) or Ambrosia.

The following supplementary files are available online at doi: 10.1101/2025.09.30.679545

Supplementary file 1. Geography markup language (.gml) format vector file of the modular gene network of phage defense systems, Related to Figure 1.

Table S1. Information on the conserved gene clusters found by PPanGGoLiN in the set of 541 RefSeq *P. aeruginosa* genomes.

Table S2. Identified instances of Dionysus, Ophion, and Ambrosia in bacteria.

Table S3. Functional domains associated with phage defense.

Table S4. List of primers used in this study.

Table S5. List of plasmids used in this study.

Chapter 6



Reduced prevalence of phage defense systems in *Pseudomonas aeruginosa* strains from Cystic Fibrosis patients

Cystic fibrosis is a genetic disorder that affects mucus clearance, particularly of the lungs. As a result, cystic fibrosis patients often experience infections from bacteria, which contribute to the disease progression. *Pseudomonas aeruginosa* is one of the most common opportunistic pathogens associated with cystic fibrosis. The presence of *P. aeruginosa* complicates the treatment due to its high antibiotic resistance. Thus, research is ongoing to treat these infections with bacterial viruses instead, known as bacteriophages. Notably, *P. aeruginosa* clinical strains possess a variety of phage defense mechanisms that may limit the effectiveness of phage therapy. In this study, we compared the defense system repertoire of *P. aeruginosa* strains isolated from cystic fibrosis patients with those from non-cystic fibrosis patients. Our findings reveal that *P. aeruginosa* strains isolated from cystic fibrosis patients have fewer phage defense mechanisms per strain than from non-cystic fibrosis patients, suggesting altered phage selection pressures in strains colonizing cystic fibrosis patient lungs.

A modified version of this chapter has been published as

van den Berg, D. F., & Brouns, S. J. J. (2025). Reduced prevalence of phage defense systems in *Pseudomonas aeruginosa* strains from cystic fibrosis patients. *mBio, 16(4),* doi: 10.1128/mbio.03548-24.

Introduction

Cystic fibrosis (CF) is a genetic disorder that affects mucus clearance, especially of the lungs (*346*). The decrease in mucus clearance creates a breeding ground for opportunistic pathogens, such as bacteria, to colonize the lungs *(346)*. These infections are often chronic, accumulate resistance to antibiotic treatment, and cause major lung damage (*346, 347*). Over time, this damage results in respiratory failure, the most common cause of death for CF patients (*346*). Lung infections in CF patients are often caused by strains of *Pseudomonas aeruginosa* (*346*). Although antibiotic treatments can alleviate patient symptoms and prolong life, they generally fail to fully eradicate *P. aeruginosa* from the lungs (*348*). Consequently, research has turned to alternative treatment options, including the use of bacterial viruses (phages) (*349-353*). However, the efficacy of phage-based treatments might also be limited since *P. aeruginosa* strains often possess a variety of phage defense mechanisms that may limit the sensitivity of strains to phages (*146*). Particularly, a recent study on the abundance of defense systems in clinical isolates of *P. aeruginosa* found that these strains accumulate a large number of defense systems and are less susceptible to phage infection (*146*). However, it is unknown if the distribution of phage defense systems varies with the pathogen environment or patient conditions. In this study, we investigated the defense system repertoire of clinical *P. aeruginosa* isolates from mucus-rich lungs of CF patients by comparing them to strains isolated from non-CF lung patients.

Results and discussion

We utilized the *Pseudomonas* Genome Database to investigate the defense system repertoire of *P. aeruginosa* isolates from the respiratory system of CF and non-CF (e.g. pneumonia) patients (*354*). We first compared the prevalence of the phage defense systems and found isolates from CF patients to encode significantly fewer phage defense systems per strain (non-CF: 9.8 vs CF: 6.8; a 30% reduction; Welch Two Sample t-test, t = -9.878, df = 601.37, p-value < 2.2e-16) (**Figure 1a; Figure S1a**) (*146*). This possibly renders the isolates from CF lungs more susceptible to phages, as we previously observed that *P. aeruginosa* strains including CF strains, encoding fewer phage defense systems, exhibit greater susceptibility to phages (*146*). This finding remains valid when only considering lung isolates (n = 11, adjusted R-squared = 0.37, p-value < 0.05). Interestingly, this reduction in the number of defense systems per strain did not affect the overall diversity of the defense repertoire, where the Shannon index (Both: 3.86) and evenness (Both: 0.80) remained identical (**Table S1**). Therefore, we hypothesized that the decrease in defense systems per strain is not the result of a broad, non-specific reduction. Instead, we suspected that the overall reduction is possibly caused by differences in the abundance of a few prevalent defense systems. To assess this, we compared the abundance of each defense system in CF and non-CF lung isolates (**Table S1**) using a chi-square test followed by a post-hoc test with false positive rate (FDR) adjusted p-values, considering adjusted p-values lower than 0.01 to be significant (**Table S2**). Additionally, we considered changes greater than 2-fold (log2FC > 1 or log2FC < -1) as relevant (**Table S2**).

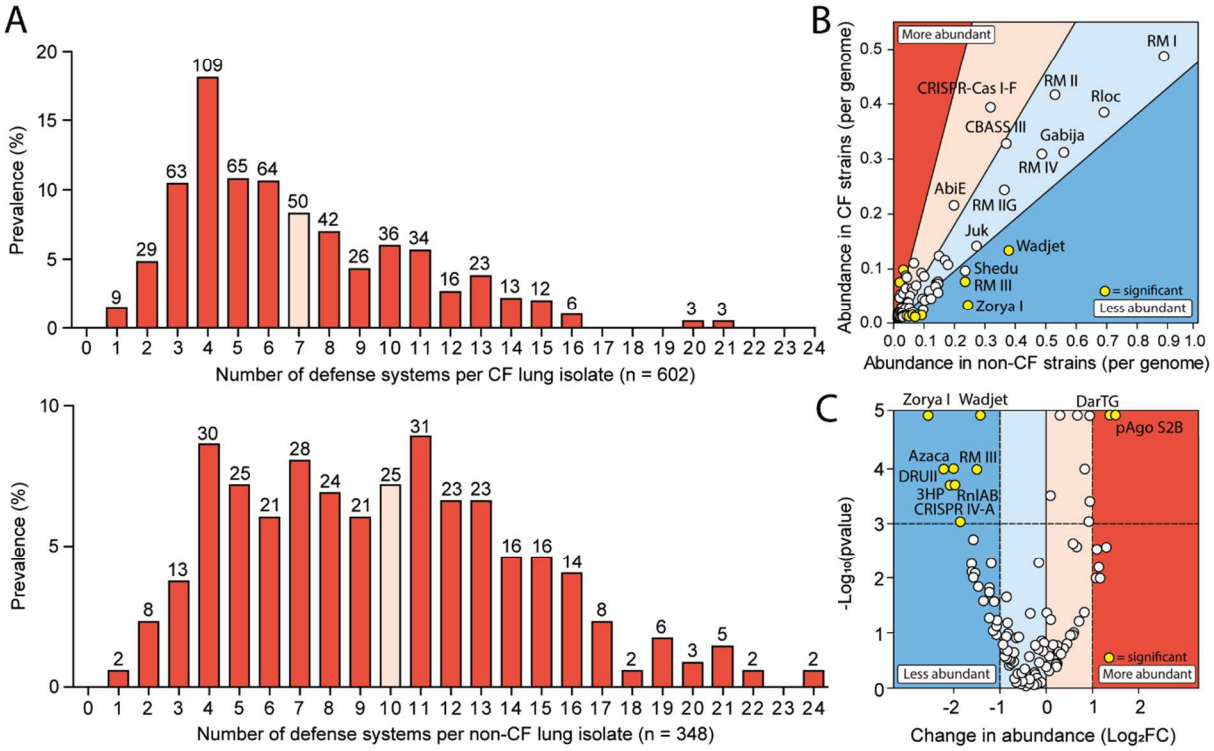

Figure 1. *P. aeruginosa* strains isolated from cystic fibrosis (CF) patient lungs have fewer defense systems. (a) The number of defense systems per *P. aeruginosa* strain isolated from CF lungs (top) and non-CF lungs (bottom). The median number of defense systems per strain is indicated with a lighter red colored bar. **(b)** The comparative abundance of individual defense systems in *P. aeruginosa* strains isolated from CF and non-CF lungs. The graph is divided into four sections: Red represents a log2FC greater than 1, light red indicates a log2FC between 0 and 1, light blue corresponds to a log2FC between 0 and -1, and blue signifies a log2FC less than -1. The significance of the change in abundance was calculated using a Chi-Square test with false positive rate (FDR) adjusted p-values. Significant defense systems (adjusted p-value < 0.01) are indicated as yellow dots. **(c)** A volcano plot depicting the change in abundance (Log2FC) of defense systems between CF and non-CF isolates as well as the significance of this change.

We identified ten defense systems with considerable changes in abundance (two were more abundant and eight were less abundant in CF strains; **Figure 1b-c, Sb-f**), but only three of these systems were prevalent enough to potentially affect the overall number of defense systems per strain (present in more than 10%) in either group, all three were more depleted in CF isolates. These defense systems include Wadjet type I (*355*) (non-CF: 37.8% vs CF:13.4%; log_2FC = -1.4), RM type III (non-CF: 23.5% vs CF: 7.7%; log_2FC = -1.5) and Zorya type I (27) (non-CF: 24.4% vs CF: 3.4%; log_2FC = -2.5). The reduced prevalence of just these three defense systems accounts for one-fifth (344 out of 1786 fewer defense systems than expected: 19.3% of total) of the total reduction in defense systems observed in CF compared to non-CF isolates. Besides these defense systems, CRISPR-Cas type IV-A is also noteworthy for being completely absent in the CF isolates, while present in the non-CF isolates (non-CF: 2.6% vs CF: 0%; log_2FC = -1.8). Interestingly, all of the above-mentioned defense systems act on foreign DNA, and notably, two of these defense systems, Wadjet type I and CRISPR-Cas type IV-A, are known to act upon plasmids specifically (*159, 355, 356*). The reduced presence of plasmid-restrictive mechanisms of these strains may highlight the role of plasmids in CF-infecting strains, potentially facilitating the acquisition of plasmids to gain resistance against antibiotics used for treating CF infections (*357, 358*). Notably, the reduction in the phage defense repertoire of CF lung isolates is not due to the reduced presence of plasmid-restrictive mechanisms alone, such as Wadjet, RM systems, CRISPR-Cas systems, DISARM, and MADS. Rather the observed reduction remains unaffected by the exclusion of these plasmid-restrictive defenses from the analysis (CF mean: 4.2; non-CF mean: 5.8; Welch Two Sample t-test: t= 9.1, df = 680, p-value < $2.2e^{-16}$).

It remains unclear whether the reduced prevalence of defense systems in CF isolates was caused by strain selection before, or during lung infection in CF patients. As for antibiotic resistance, strains infecting CF lungs adapt over time (*346*), and could therefore also adapt in the context of phages and other mobile genetic elements such as plasmids. In the CF lung environment, this adaptation seems to involve the loss of phage defense mechanisms or reduction of their expression. Possibly because CF lungs are less penetrable to phages due to factors such as reduced air circulation, a thicker and dehydrated mucus layer, and a higher prevalence of biofilms (*346, 359*). Supporting this hypothesis, we observed fewer prophages in the genomes of CF lung isolates, suggesting reduced exposure to phage predation (CF: 3.3 vs non-CF: 5.3 prophages per genome; 38% decrease; Welch Two Sample t-test: t = -12.379, df = 508.88, p-value < $2.2e^{-16}$) (**Figure S1g**). Besides the reduced phage defenses of *P. aeruginosa* isolates from CF lung isolates we also observed these strains to have a significantly smaller genome (CF: 6.5 Mb vs non-CF: 6.8 Mb; a 4% decrease; Welch Two Sample t-test: t = -13.811, df = 594.47, p-value < $2.2e^{-16}$) and to encode fewer genes (CF: 6053 vs non-CF: 6298 genes; a 4% decrease; Welch Two Sample t-test: t = -11.715, df = 625.07, p-value < $2.2e^{-16}$) (**Figure S1h,i**). These characteristics reflect a more adapted and specialized genome of *P. aeruginosa* strains in the CF lung environment (*360-362*).

Conclusion

We observed that *P. aeruginosa* strains isolated from the lungs of CF patients encode a more limited number of phage defense systems compared to strains isolated from non-CF patient lungs. This can be attributed to a reduced abundance of several key phage defense systems including Wadjet, Zorya type I, and RM type III, which likely results from adaptation to the CF lung environment. This provides a promising perspective that these bacterial strains are more susceptible to phage therapeutic options.

Materials and Methods

All complete *P. aeruginosa* genomes were downloaded from the *Pseudomonas* Genome Database on the 9th of April 2023 (n = 14,230). The metadata was utilized to select the strains for: Species = *P. aeruginosa*, Assembly version status = latest, Host taxonomic name = Homo sapiens. A further selection was then made between isolates from the respiratory system of CF (n = 602) and those of non-CF (n = 348) (**Table S1**). The non-CF patients primarily were affected by unspecified respiratory tract infection (73.0%) and ventilator-associated pneumonia (20.1%). Chronic bronchitis and chronic obstructive pulmonary disease (COPD) were less frequently prevalent, 4.9% and 2.0% respectively. Defense systems were detected in *P. aeruginosa* genomes with DefenseFinder v1.3.0 (*363*) using DefenseFinderModels v1.3.0. Prophages were identified using PhiSpy v4.2.21 (default settings) (*364*). Differentially abundant defense systems among CF strains compared to non-CF strains were detected using a simulated p-value Chi-Square test followed by a post hoc analysis (R-package: chisq.posthoc.test v0.1.2) (*365*) with false discovery rate (FDR) adjusted p-values. An adjusted p-value lower than 0.01 was considered significant. Moreover, a $\log_2$FC of at least (-)1 was used as an additional threshold. Visualization of this analysis was done using a volcano plot created by the R-package EnhancedVolcano v1.24.0 (*268*).

Supplementary information

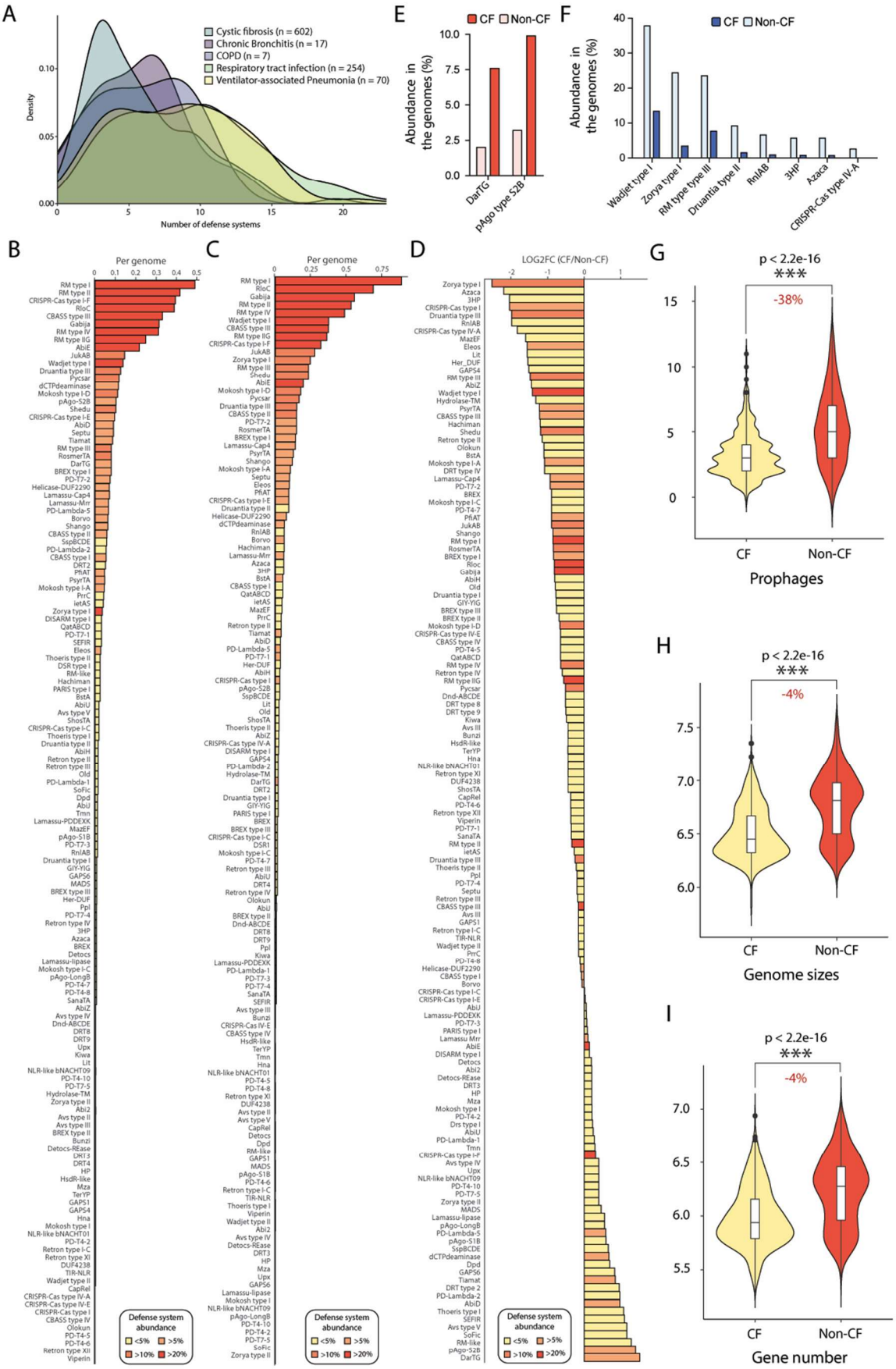

Figure S1. Comparison of the phage defense system repertoire in *P. aeruginosa* strains isolated from cystic and non-cystic fibrosis lungs. (a) A density plot of the number of defense systems per *P. aeruginosa* strain isolated from CF lungs and the other non-CF lung sources. **(b)** The prevalence of individual defense systems per *P. aeruginosa* strain isolated from CF lungs. **(c)** The prevalence of individual defense systems per *P. aeruginosa* strain isolated from non-CF lungs. **(d)** The relative log2FC change in prevalence of individual defense systems between *P. aeruginosa* strains isolated from CF and non-CF patients. **(e)** The subset of significantly differentially abundant (adjusted p-value < 0.01) defense systems that are found to be overly abundant (log2FC > 1) and **(f)** less abundant (log2FC < -1) in CF isolates. **(g)** The number of detectable prophages in *P. aeruginosa* strains isolated from CF and non-CF patients. **(h)** The genome size of *P. aeruginosa* strains isolated from CF and non-CF patients. **(i)** The number of genes present in *P. aeruginosa* strains isolated from CF and non-CF patients.

The following supplementary files are available online at doi: 10.1128/mbio.03548-24.

Table S1. Summary table of the *P. aeruginosa* strains and additional information, including the defense system repertoire.

Table S2. The relative fold changes in defense system prevalence in *P. aeruginosa* strains isolated from CF lungs compared to those from non-CF lungs.

Chapter 7



Phages encode broad and specific counter-defense genes to overcome defense systems

Bacteria and bacteriophages have been engaged in a relentless evolutionary arms race, driving a rapid evolution of bacterial defense mechanisms and leading to their scattered distribution across genomes. We hypothesized that the variability in defense systems presence in bacterial genomes leads to equally variable counter-defense repertoires in phage genomes. To test this, we analysed the variable regions in *Pseudomonas* model phages of the *Pbunavirus* genus, uncovering five anti-defense genes inhibiting Zorya type I, RADAR, Hypnos, Druantia type I and III, and Thoeris type III. Remarkably, a typical *Pbunavirus* encodes up to five known anti-defense genes, some broadly inhibiting four unrelated defense systems with distinct nucleic acid-targeting mechanisms. Structural searches revealed that these broad-acting inhibitors are encoded across diverse phage taxa infecting multiple bacterial hosts. The broad activity and prevalence of these inhibitors indicate strong selective pressures exerted by defense systems and offer opportunities to improve phage-based therapeutics.

A modified version of this chapter has been published as

Costa, A. R.*, van den Berg, D. F.*, Esser, J. Q.*, van den Bossche, H., Pozhydaieva, N., Kalogeropoulos, K., & Brouns, S. J. J. (2025). Bacteriophage genomes encode both broad and specific counter-defense repertoires to overcome bacterial defense systems. *Cell Host & Microbe*, *33*(7), 1161-1172.

Introduction

Prokaryotes have evolved an extensive array of defense mechanisms to protect themselves from bacteriophage predation (*10, 49*). In response, phages have developed countermeasures to overcome these defenses, including mutations that enable them to escape recognition, and anti-defense genes that disrupt phage defense systems (*32, 52, 61*). Well-known examples of the latter include anti-restriction and anti-CRISPR proteins, which disable restriction-modification and CRISPR-Cas systems, respectively (*310, 366, 367*). With the increasing discovery of new phage defense systems (*10, 122, 141, 159, 160, 368*), novel anti-defense genes targeting these mechanisms are beginning to be uncovered as well (*281, 369*).

The ongoing phage-bacteria arms race is highlighted by instances where defense systems detect interference by specific anti-defense genes. For example, systems like Retrons (*38, 370*), Ronin, and PARIS (*62, 371*) can sense tampering of other defense systems by phages and trigger additional defensive measures, prompting phages to evolve further countermeasures (*37, 323*).

Much like bacteria do not encode every possible defense system (*41*), it is likely that phages do not carry all anti-defense genes. In phages, the number of anti-defense genes may be limited by genomic constraints, as the size of a phage genome is restricted by the capacity of the capsid (*40*). In addition, some anti-defense proteins such as Ocr (i.e., overcome classical restriction) are sensed by specific defense systems, presenting additional vulnerabilities (*62*). To balance these trade-offs, it is possible that phages carry only a subset of anti-defense genes, acquiring others through recombination with related phages (*372*). These recombination events, driven by high evolutionary pressure from phage defenses, are main drivers of phage differentiation and variability (*30*). Based on this, we hypothesized that anti-defense genes are likely encoded in regions of high genomic variability within phage genomes. To test this hypothesis, we used phages of the *Pbunavirus* genus as a model, which infect *Pseudomonas* species (*373*) and have been used in pre-clinical studies and clinical trials for phage therapy (*374, 375*). By analysing regions of high genomic variability in these phages, we selected 43 candidate genes and tested them for anti-defense activity against a panel of 12 individual defense systems in *Pseudomonas aeruginosa* strain PAO1. Our screen identified five novel anti-defense genes: one inhibiting Druantia type III (DadIII-1), another targeting Thoeris type III (TadIII-1), one inhibiting Zorya type I (ZadI-1), and two related broad defense inhibitors (Bdi1 and Bdi2) targeting four defense systems including Zorya type I, RADAR, Hypnos, and Druantia type I. This approach demonstrates the potential for identifying new anti-defense genes by exploring variable regions of phage genomes and sets the stage for broader investigations across other phage taxonomic groups. Understanding the full repertoire of phage anti-defense mechanisms will not only shed light on phage evolution but also advance the development of phage-based therapeutics.

Results

Novel anti-defense genes in highly variable genomic regions of *Pbunavirus*

To identify new anti-defense genes in highly variable regions of phage genomes, we analysed the pangenome of 162 *Pbunavirus* genomes available in RefSeq using PPanGGOLiN (*249*). This analysis categorized the genes into three distinct groups: 30 core genes (18%, present in at least 95% of genomes), 71 shell genes (39%, present in more than 1.5% and less than 95% of genomes) and 66 cloud genes (42%, present in less than 1.5% of genomes) (**Table S1**). Notably, the highest genomic variability was observed in the 50-65 kb region, which is predominantly composed of small hypothetical proteins (**Figure 1a**).

To test our hypothesis that proteins with anti-defense activity are encoded in these highly variable regions, we mapped known anti-defense genes onto the *Pbunavirus* genome. This analysis identified only the gene for the anti-restriction protein Lar (*376*) within these variable regions (**Figure 1a**, **Table S1**). To investigate the anti-defense activity of other genes in this region, we selected 43 representative genes from the variable regions of the *Pbunavirus* pangenome (**Table S2**) and tested them against a

collection of 12 defense systems expressed in *P. aeruginosa* strain PAO1 (*138, 146*) (**Figure 1b**). These candidate anti-defense genes encode proteins ranging in size from 37 to 475 amino acids, with the majority (75%) being smaller than 150 amino acids (**Figure S1a**).

Each candidate anti-defense gene was expressed individually from a plasmid in host cells, which were then challenged with a *Pbunavirus* that does not encode the candidate anti-defense gene. However, some defense systems, such as Druantia type I, Hypnos, and Zorya type I, failed to protect from *Pbunavirus* phages in our collection. In these cases, we used a surrogate phage from the *Autographiviridae* family, which is vulnerable to these systems, to test the effect of the candidate anti-defense genes. The combinations of defense systems and candidate anti-defense genes for testing were based on the known infectivity of *Pbunavirus* against PAO1 expressing each defense system (*146*). If a defense system protected against a phage specifically encoding the candidate anti-defense gene as well, that combination was excluded from testing (**Table S2**).

Our search revealed five anti-defense genes targeting various defense systems, including Zorya type I, RADAR, Hypnos, Druantia types I and III, and Thoeris type III (**Figure 1c**). Interestingly, the validated anti-defense proteins are significantly smaller (average size: 90 amino acids) than other proteins in the *Pbunavirus* core (average: 559 amino acids), shell (average: 319 amino acids) or cloud (average: 314 amino acids) pangenomes.

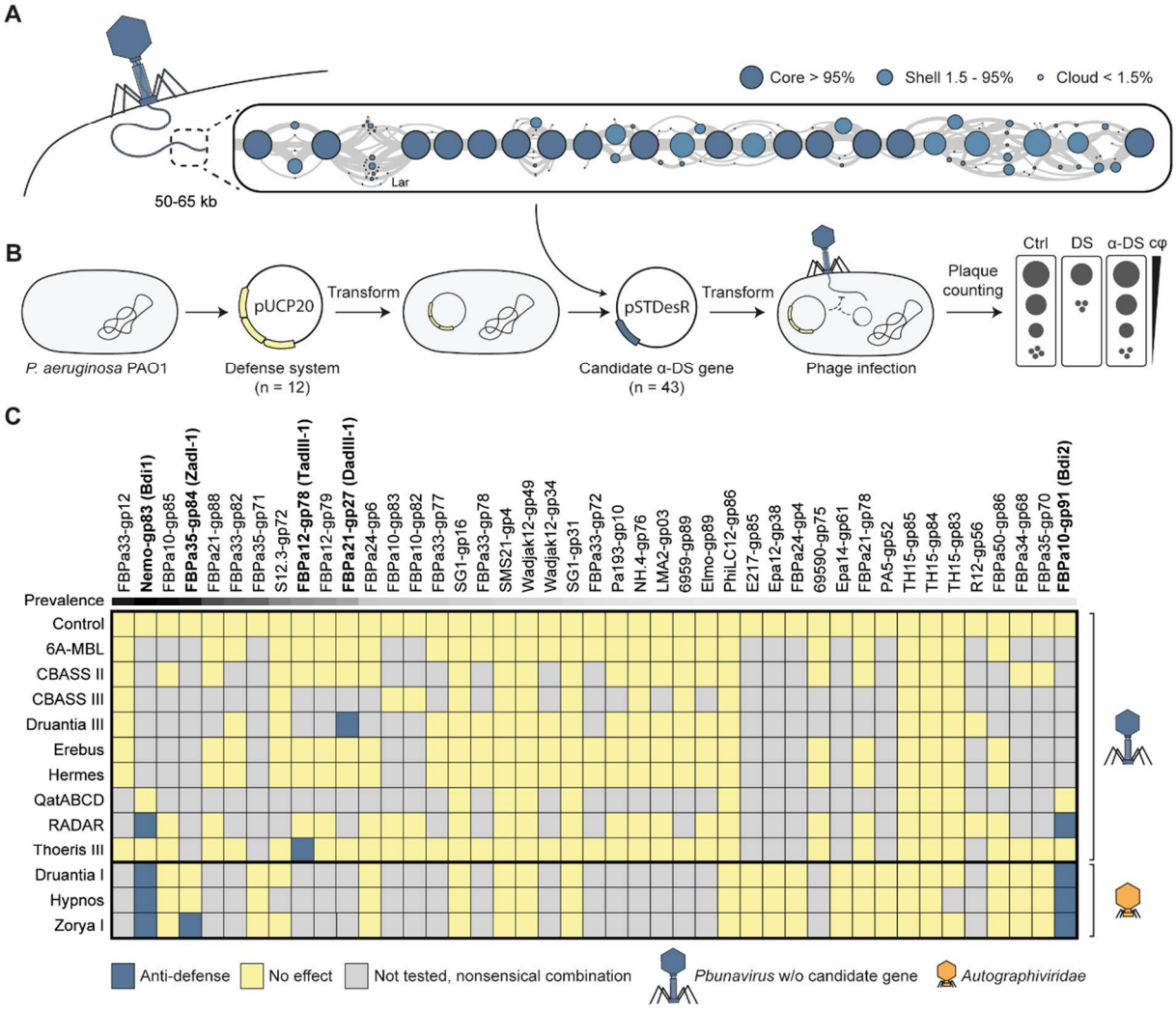

Figure 1. The pangenome of *Pbunavirus* reveals highly variable regions containing genes with anti-defense properties. (a) Representation of the pangenome of the 50-65 kb region of *Pseudomonas Pbunavirus* (n = 162) generated by PPanGGOLiN. Genes are classified as core when present in >95% of the genomes, as shell when found in >1.5% and <95% of the genomes, and as cloud when present in <1.5% of the genomes. The location of known anti-defense gene Lar is indicated.

(b) Individual defense systems were cloned with their native promoters into pUCP20 and then introduced into the *P. aeruginosa* strain PAO1, followed by the introduction of the candidate anti-defense gene on a second plasmid. The anti-defense activity of the candidate genes was assessed using efficiency of plating assays with a phage that does not encode the candidate anti-defense gene. **(c)** Tricolor heatmap showing the combinations of defense system and candidate anti-defense gene tested. The combinations tested are shown in yellow, with those showing anti-defense properties marked in dark blue. Combinations not tested (grey) were excluded because phages carrying the candidate gene are targeted by the defense system. The anti-defense activity of the candidate gene was tested using a *Pbunavirus* where this gene is absent or tested using a podophage of the *Autographiviridae* family. The candidate genes are organized from most (left, 87%) to least (right, 0.62%) prevalent in *Pbunavirus*.

Bdi1 and Bdi2 are inhibitors of a broad range of nucleic acid acting defense systems.

Among our candidate genes, Bdi1 (broad defense inhibitor 1, ORF83 from phage Nemo, Genbank: KT372694.1) and Bdi2 (ORF91 from vB_PaeM_FBPa10, Genbank: ON857929.1) surprisingly demonstrated broad-spectrum inhibitory activity against the defense systems Zorya type I, RADAR, Hypnos, and Druantia type I in both efficiency of plating and culture collapse assays (**Figure 2a**, **Figure S2a**).

To understand the broad inhibitory activity of Bdi1 and Bdi2, we first searched for functional domains at the sequence level but found none. We then analysed the proteins at the structural level and found that, despite sharing only 33.8% identity at amino acid level (**Figure S2a**), Bdi1 and Bdi2 have strong structural similarity with each other (AlphaFold 3, DALI (*191*) Z-score: 12.1), suggesting a similar mechanism for phage defense system inhibition (**Figure 2b**, **Figure S2c**). Multimeric predictions using AlphaFold 3 suggest that Bdi1, but not Bdi2, forms multimers, with the most confident prediction being a hexamer (**Figure 2c**, **Figure S2d**). A structural similarity search using Foldseek with the multimeric structure did not reveal any significant structural homologs, but analysis of the monomeric structure indicated significant similarity (prob 1.00, %ident 16.2, E-value 7.62e-3) to elongation factor P, specifically to its N-terminal region containing a Kyprides-Ouzounis-Woese (KOW)-like domain (**Figure S2e**). The KOW domain is a subclass of the Src homology-3 (SH3) domain, an ancient fold that is involved in signalling pathways through protein-protein interactions (*377*). The KOW-like domain is also known to function through protein-protein interactions and to bind ribonucleoproteins (*378, 379*). Since the four defense systems inhibited by Bdi1 and Bdi2 have no shared proteins or functional domains that could serve as common targets, but have been shown to, or postulated to, convey defense via targeting of nucleic acids (DNA for Hypnos (*138*) and Druantia type I (*122*), RNA for RADAR (*239, 240*), DNA and RNA for Zorya type I (*380*), **Figure S2f**), we hypothesized that Bdi1 and Bdi2 interact with protein-nucleic-acid complexes of these defense systems, inhibiting their activity.

To investigate this hypothesis, we used AlphaFold 3 to co-fold Bdi1 (predicted with higher confidence than Bdi2, **Figure S2c**) with each protein of the defense systems it inhibits. We did not observe highly confident co-folding of Bdi1 with any of the defense system proteins (**Figure S2g**). Since Bdi1 is inhibiting multiple DNA- and RNA-targeting defense systems, we hypothesized that it would act as a nucleic acid mimic, but analysis of the electrostatic charges of the monomeric or multimeric structures are not suggestive of this hypothesis (**Figure 2d**).

To further explore the multimeric structure of Bdi1, we used Predictomes (*381*) to analyse AlphaFold 3 structures and identify residue pairs mediating interactions among monomers (**Table S3**). Based on this information and amino acid conservation in Bdi1 and Bdi2 homologues (**Figure 2e**, **Figure S2h**), we selected three amino acids in Bdi1 that may contribute to multimer formation for mutational analysis (**Figure 2c,f**). Mutating each of these amino acids (D14A, L33E, and W54A) completely abolished the inhibitory activity of Bdi1 against Zorya type I (**Figure 2f**). However, none of these mutations affected the anti-defense activity of Bdi1 against RADAR, and only D14A and L33A abolished its activity against Hypnos and Druantia type I (**Figure 2f**). The selective effect of these mutations on

each defense system suggests that Bdi1 employs distinct mechanisms or protein surfaces to inhibit different phage defense systems.

In summary, Bdi1 and Bdi2 function as multipurpose anti-defense proteins capable of inhibiting four defense systems involved in nucleic acid binding activities.

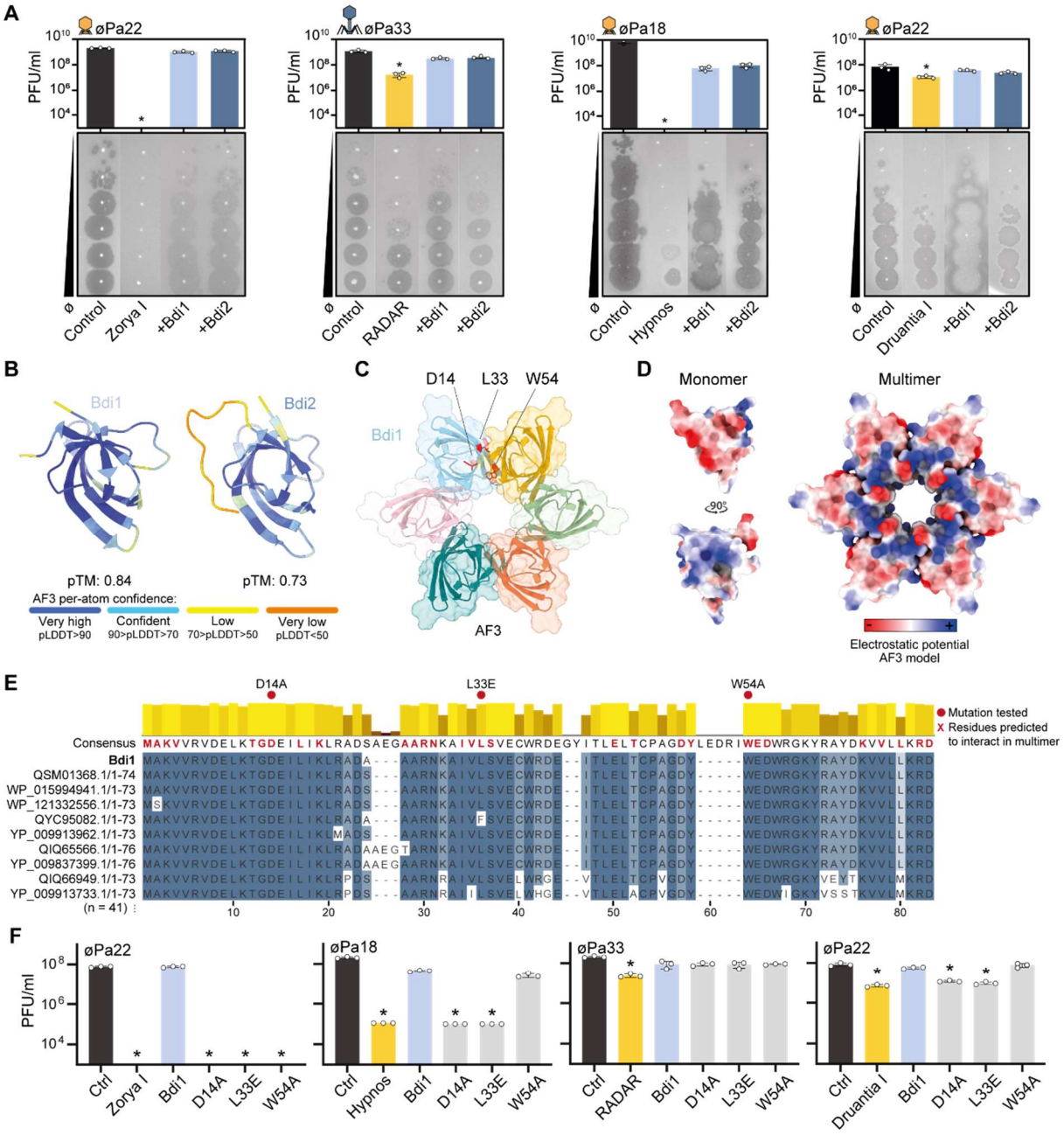

Figure 2. The inhibitory activity of Bdi1 and Bdi2 on Zorya type I, RADAR, Hypnos, and Druantia type I defense systems. (a) Effect of anti-defense proteins Bdi1 and Bdi2 on Zorya type I, RADAR, Hypnos, and Druantia type I protection against phage infection, evaluated by measuring phage titers (PFU/ml). Phage titers were determined, with an example plate shown alongside a bar graph representing the average titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). Control is *P. aeruginosa* strain PAO1 containing empty pUCP20 and pSTDesR plasmids. **(b)** Monomeric structure of Bdi1 and Bdi2, as predicted by AlphaFold 3 (AF3) and colored by pLDDT. **(c)** Hexameric structure of Bdi1, as predicted by AF3 and colored by monomer. Three amino acids of Bdi1 predicted to be important for multimer formation, and mutated in panel F, are shown in red. **(d)** Monomeric (left) and hexameric (right) AF3 structures of Bdi1 colored by electrostatic potential. **(e)** Multiple sequence alignment of Bdi1 with 40 homologous sequences. Ten representative sequences are shown, with consensus indicated above. All residues predicted to be involved in multimer formation are marked in red in the consensus sequence,

and mutations tested in F are indicated with a red dot. **(f)** Effect of point mutations in selected amino acids of Bdi1 on anti-defense activity against Zorya type I, RADAR, Hypnos, and Druantia type I. The bar graph displays the average phage titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). DS, phage defense system, indicated above each graphic.

ZadI-1 is an inhibitor of Zorya type I

Our search uncovered an additional anti-defense gene, ZadI-1 (ORF84 of vB_PaeM_FBPa35, Genbank: ON857938.1) that fully inhibits Zorya type I defense, as shown by efficiency of plating and culture collapse assays (**Figure 3a,b**).

ZadI-1 does not contain any known functional domains that could suggest its mechanism of action. To explore potential functions, we used AlphaFold 3 to predict its structure. ZadI-1 shows no sequence or structural similarity to the anti-Zorya proteins Bdi1 and Bdi2. AlphaFold 3 predictions indicate that ZadI-1 likely functions as a monomer (**Figure 3c**) and suggests a weak physical interaction with the ZorC protein (ipTM 0.31, pTM 0.54, **Figure S3a,b**). ZadI-1 appears to be buried in the EH signature domain of ZorC, interacting also with the globular C-terminal domain of ZorC, both shown to be required for ZorC DNA binding activity (*380*) (**Figure S3b**). Thus, we hypothesize that ZadI-1 could inhibit ZorC by blocking its DNA-binding activity. However, given the weakly predicted interaction between ZadI-1 and ZorC, we were unable to confidently identify specific amino acids involved in this interaction for further mutational analysis. Instead, we mutated two conserved lysine residues (K64 and K68) based on homologues of ZadI-1 (**Figure 3d**) and tested their impact on anti-defense activity. Both K64A and K68A mutations abolished the anti-defense function of ZadI-1 against Zorya type I, indicating that these residues are critical for its activity (**Figure 3e**).

In summary, ZadI-1 inhibits Zorya type I through a mechanism that remains to be characterized, potentially by interfering with the DNA binding activity of ZorC.

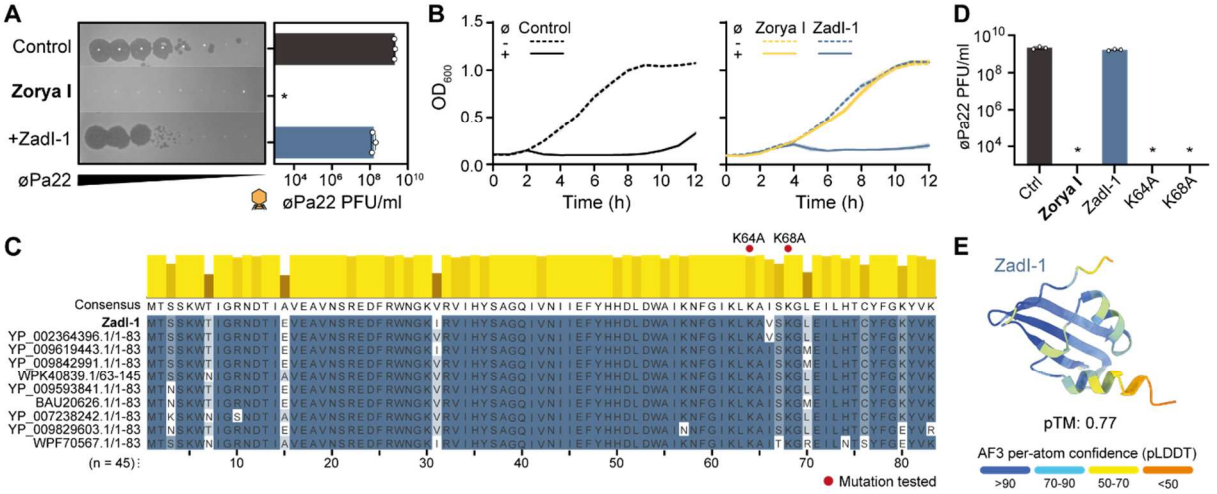

Figure 3. The inhibitory activity of ZadI-1 on Zorya type I defense. (a) Effect of anti-defense protein ZadI-1 on Zorya type I protection against phage vB_PaeP_FBPa22 (øPa22), evaluated by measuring phage titers (PFU/ml). An example plate is shown alongside a bar graph representing the average phage titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). **(b)** Effect of anti-defense protein ZadI-1 on Zorya type I protection against phage vB_PaeP_FBPa22 (øPa22), evaluated by measuring bacterial growth via optical density (OD 600 nm). The graph displays the average optical density readings and standard deviations from three biological replicates. **(c)** Structure of ZadI-1 predicted by AlphaFold 3 (AF3) and colored by pLDDT. **(d)** Multiple sequence alignment of Zad-1 with 43 homologous sequences. Ten representative sequences are shown, with consensus indicated above. Mutations tested in D are indicated with a red dot. **(e)** Effect of point mutations in selected amino acids of ZadI-1 on anti-defense activity against Zorya type I. The bar graph displays the average phage titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*).

TadIII-1 is an inhibitor of Thoeris type III

We identified TadIII-1 (ORF78 of vB_PaeM_FBPa12, GenBank: ON857930.1), a protein that inhibits Thoeris type III defense, as demonstrated by both efficiency of plating and culture collapse assays (**Figure 4a,b**). In these assays, TadIII-1 fully restored phage infectivity in the presence of the defense system. Despite extensive analysis, we found no known functional domains in TadIII-1, nor any significant structural similarity to previously characterized proteins (**Figure 4c**).

To identify potential interactions between TadIII-1 and the Thoeris type III components, we used co-folding predictions. TadIII-1 was predicted to interact with the phage-sensing TIR proteins ThcB1 and ThcB3, as well as SLOG domain-containing protein ThcA (**Figure S4a**). Since ThcB1 is predicted to form a dimer and ThcA is predicted to interact with ThcB4 (*138*), we also co-folded TadIII-1 with these complexes. The results showed a highly confident interaction between two TadIII-1 molecules and the ThcB1 dimer (ipTM = 0.73, pTM = 0.80, **Figure S4b,c**).

The TIR protein ThcB1 is responsible for producing signal molecules that activate the Thoeris type III defense response upon phage infection (*138*). TadIII-1 appears to bind to this protein, leading us to hypothesize that it interferes with signal production. Supporting this hypothesis, TadIII-1 was found to bind near the catalytic aspartate residue (D55) of the ThcB1 TIR domain (**Figure 4d**), a residue critical for the TIR-domain's signalling function (*138*).

Using Predictomes, we identified two conserved amino acids in TadIII-1 – T47 and S49 – that are predicted to strongly interact with the D55 residue of ThcB1 (**Table S3**, **Figure 4d,e**). Mutating these amino acids individually abolished TadIII-1 anti-defense activity (**Figure 4f**), confirming their importance for its function. These findings suggest that TadIII-1 inhibits Thoeris type III by physically blocking the active site of the ThcB1 TIR domain, thereby interfering with signal production. This mechanism differs from that of other known anti-Thoeris proteins, such as Tad1 and Tad2, which function by sequestering immune signalling molecules produced by TIR-domain proteins from other Thoeris types (*382, 383*).

In summary, TadIII-1 inhibits the anti-phage activity of Thoeris type III potentially by blocking the active site of the TIR domain and inhibiting signal molecule production.

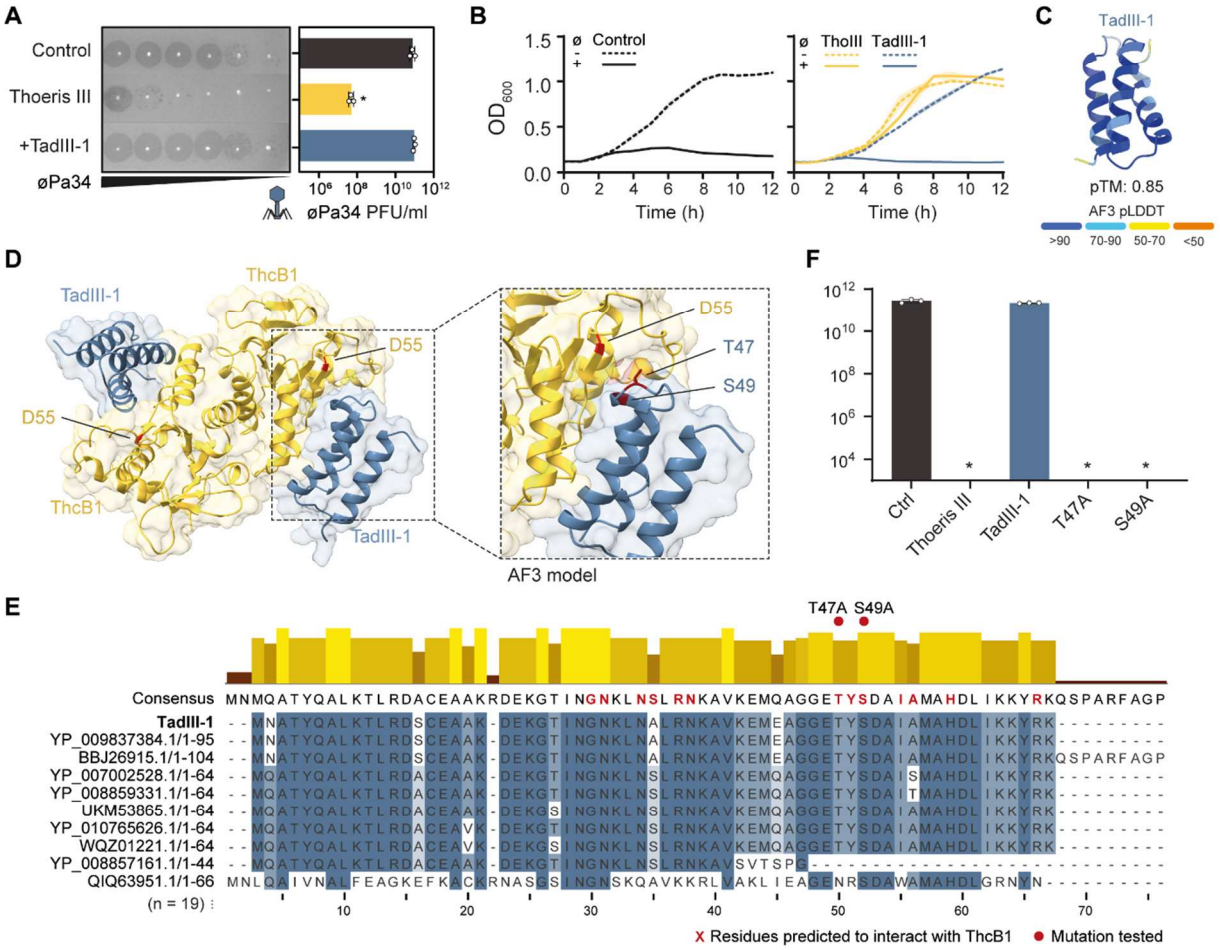

Figure 4. The inhibitory activity of TadIII-1 on Thoeris type III defense. (a) Effect of anti-defense protein TadIII-1 on Thoeris type III protection against phage vB_PaeM_FBPa34 (øPa34), evaluated by measuring phage titers (PFU/ml). An example plate is shown alongside a bar graph representing the average titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). **(b)** Effect of anti-defense protein TadIII-1 on Thoeris type III protection against phage vB_PaeM_FBPa34 (øPa34), evaluated by measuring bacterial growth via optical density (OD 600 nm). The graph displays the average optical density readings and standard deviations from three biological replicates. **(c)** Structure of TadIII-1 predicted by AlphaFold 3 (AF3) and colored by pLDDT. **(d)** Structure of the complex formed between two TadIII-1 monomers (blue) and a ThcB1 dimer (yellow), as predicted by AlphaFold 3. The catalytic aspartate D55 of ThcB1 is indicated in red. The inlet shows two amino acids (red) of TadIII-1 predicted to be important for interaction with D55 of ThcB1 and mutated in panel F. **(e)** Multiple sequence alignment of TadIII-1 with 18 homologous sequences. Ten representative sequences are shown, with consensus indicated above. All residues predicted to interact with ThcB1 are marked in blue in the consensus sequence, and mutations tested in F are indicated with a red dot. **(f)** Effect of point mutations in selected amino acids of TadIII-1 on anti-defense activity against Thoeris type III. The bar graph displays the average phage titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*).

DadIII-1 is an inhibitor of Druantia type III

We identified DadIII-1 (ORF27 of vB_PaeM_FBPa21, Genbank: ON857942.1) as an anti-defense gene effective against the defense system Druantia type III, as demonstrated in efficiency of plating and culture collapse assays (**Figure 5a,b**). Druantia III comprises the proteins DruE and DruH, with DruE being conserved across all Druantia types (*159*). To determine whether DadIII-1 interferes with DruH or DruE, we tested its effect on Druantia type I, hypothesizing that if DadIII-1 targeted DruE, it would inhibit all Druantia types. We observed that DadIII-1 did not affect Druantia type I defense (**Figure 5c**), indicating that DadIII-1 likely targets DruH rather than DruE.

DruH lacks predicted functional domains, complicating efforts to elucidate the molecular basis for DadIII-1-mediated inhibition. Structural analysis using AlphaFold 3 did not suggest any complex

formation between DadIII-1 and the Druantia type III proteins (**Figure S5a,b**), implying that inhibition may not rely on direct protein-protein interactions. However, the structural prediction of DadIII-1 was of low confidence, which limits our ability to draw conclusions about its mechanism of inhibition. This poor prediction suggests that DadIII-1 likely has a novel structure, hinting at potentially new functions. Additionally, mutating the conserved lysine residue (K85A) in DadIII-1 did not alter its anti-defense activity (**Figure 5d,e**).

In summary, the phage protein DadIII-1 inhibits Druantia type III, most likely through a mechanism specifically targeting the activity of DruH, a protein unique to this Druantia type.

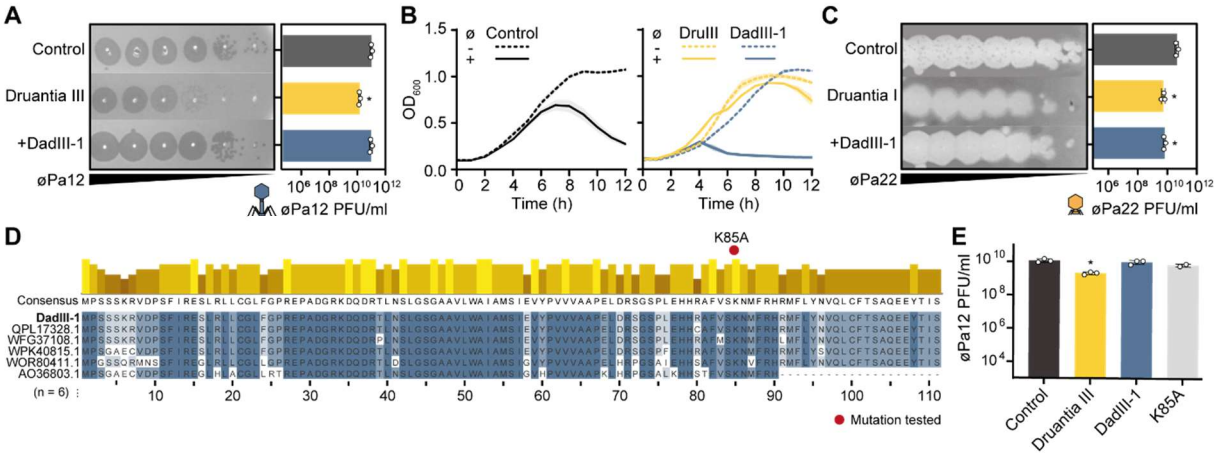

Figure 5. The inhibitory activity of DadIII-1 on Druantia type III defense. (a) Effect of anti-defense protein DadIII-1 on Druantia type III protection against phage vB_PaeM_FBPa12 (øPa12), evaluated by measuring phage titers (PFU/ml). An example plate is shown alongside a bar graph representing the average titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). **(b)** Effect of anti-defense protein DadIII-1 on Druantia type III protection against phage øPa12, evaluated by measuring bacterial growth via optical density (OD 600 nm). The graph displays the average optical density readings and standard deviations from three biological replicates. **(c)** Effect of anti-defense protein DadIII-1 on Druantia type I protection against phage vB_PaeP_FBPa22 (øPa22), evaluated by measuring phage titers (PFU/ml). An example plate is shown alongside a bar graph representing the average titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). **(d)** Multiple sequence alignment of DadIII-1 with its five homologous sequences. The consensus sequence is indicated on top. Mutations tested in E are indicated with a red dot. **(e)** Effect of point mutations in a selected amino acid of DadIII-1 on anti-defense activity against Druantia type III. The bar graph displays the average phage titers from three independent biological replicates. Error bars indicate standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*).

Widespread occurrence of the anti-defense genes in phage infecting diverse bacterial taxa

Using a sequence-based homology search, we found that several of the anti-defense genes from our study are highly prevalent in *Pbunavirus*, particularly Bdi1 (99%) and Zadl-1 (78%) (**Figure 6a**). These genes are often co-encoded by the same phage (**Figure 6a, Figure S6**, **Table S4**), but we did not detect significant pairs of anti-defense genes that are significantly co-occurring, even when taking the phylogeny of the large phage terminase into account.

On average, phages of the *Pbunavirus* group encode 2.9 anti-defense genes per phage (**Figure 6b**), but no phage encodes all six known anti-defense genes (**Figure 6a, Figure S6**). This may be due to constraints such as capsid size, the metabolic burden of maintaining additional genes, or the need to balance anti-defense gene expression with evading their detection by some phage defenses. These factors suggest that the phage anti-defense pangenome acts as a reservoir from which individual phages selectively acquire anti-defense genes as needed. Supporting this, our analysis showed that anti-defense genes are unevenly distributed across different phylogenetic branches of *Pbunavirus*. For example, Zadl-1 is nearly absent in the PB1 branch, while Bdi2 is almost absent in the *Pseudomonas* phage Epa7, TH15,

and PB1 branches (**Figure 6a**). These genes are typically located within clusters of small genes in highly variable regions of the phage genome (**Figure S6**). Interestingly, despite anti-defense activity across different phage groups (*Pbunavirus* and *Autographiviridae*), sequence similarity searches failed to detect the anti-defense genes in *Pseudomonas* phages outside the *Pbunavirus* genus, suggesting very little exchange of these genes between phage groups. The exception is ZadI-1, which was also found in the *Arenbergviridae* phage family (**Figure 6c**, **Table S4**). Furthermore, sequence similarity searches also failed to detect the anti-defense genes in phages infecting other bacterial species.

We hypothesized that the insufficient sensitivity of sequence-based approaches may limit the detection of anti-defense homologs. To overcome this, we employed a structural-based strategy, searching for structural homologs of the identified anti-defense proteins. Using AlphaFold3 structures, we queried the Big Fantastic Virus Database (BFVD) (*384*) using Foldseek (*230*). The BFVD consists of 351,242 viral structures, with all sequences sharing less than 30% amino acid identity. This analysis identified one structural homolog for ZadI-1 and two for TadIII-1 (**Table S5**). Bdi1 and Bdi2 yielded the most hits, with a total of 54 unique structural homologs distributed across more than 32 phage taxa infecting more than 17 bacteria genera. These span Gram-negative and Gram-positive hosts of the bacterial phyla Actinobacteria, Actinomycetota, Bacteroidota, and Pseudomonadota (**Figure 6d**, **Table S5**). Remarkably, the structural homologs of Bdi1 and Bdi2 exhibited very low sequence identity (7.4-29.7%), highlighting the limitations of sequence-based methods for detecting these relationships (**Figure 6d**, **Table S5**).

In summary, the anti-defense genes identified in this study represent a major step in uncovering the anti-defense repertoire of phage genomes and underline the importance of phage defense systems as a barrier to phage infection. The ability of these phages to encode multiple, but not all anti-defense genes, suggests they exploit a dynamic anti-defense pangenome, selectively drawing from it to adapt their strategies against bacterial immune systems while balancing other genomic constraints.

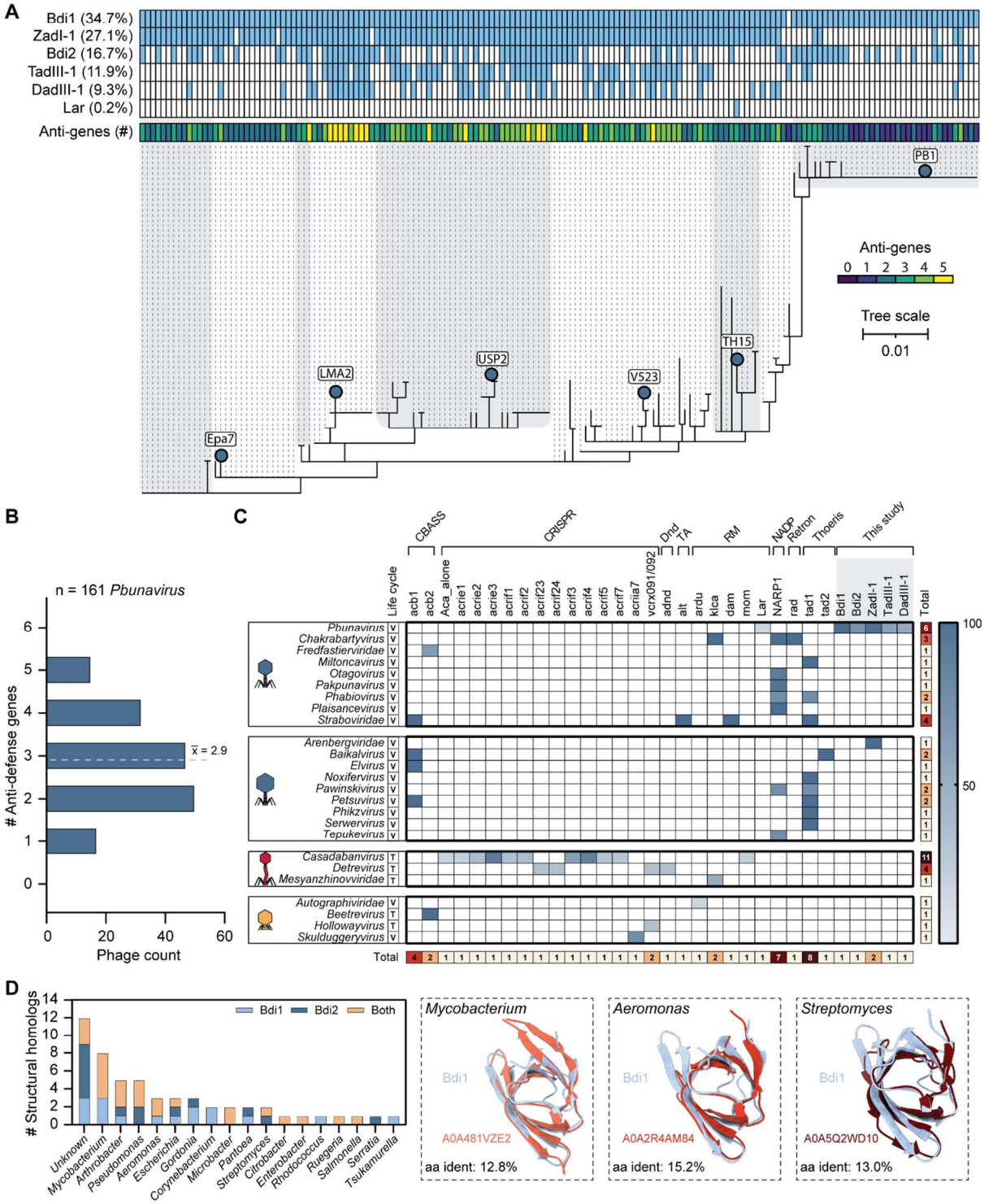

Figure 6. Distribution of anti-defense genes in *Pseudomonas* phage genomes. (a) Phylogenetic tree of *Pbunavirus*, showing the presence of anti-defense genes across the tree. Phages highlighted in Figure S6 are indicated in the tree. **(b)** Prevalence of anti-defense genes across 162 *Pbunavirus* genomes. $\bar{x}$ is the average number of anti-phage defense genes per phage. **(c)** Distribution of anti-defense genes across different phage taxonomic groups infecting *Pseudomonas*, with their life cycle denoted by a "V" for virulent and a "T" for temperate phages. **(d)** Diversity of host genera of phages encoding Bdi1 and Bdi2 structural homologs found in the BFVD (*384*), a database of 351,242 viral structures sharing less than 30% amino acid identity. The structural alignment of Bdi1 with example homologs is depicted on the right, with the amino acid identity indicated at the bottom.

Discussion

In this study, we explored the highly variable genomic regions of a common myophage infecting *P. aeruginosa* to identify novel anti-defense genes, which hold potential for improving phage therapy applications. Using *Pbunavirus* as a model phage, we identified and validated 5 out of 43 candidate genes. Collectively, these genes showed activity against six of the 12 defense systems tested. Among these anti-defenses, the structural homologs Bdi1 and Bdi2 individually inhibited four distinct defense systems: Zorya type I, RADAR, Hypnos, and Druantia type I. We postulate that these anti-defense genes interfere with DNA- and RNA-interacting proteins of these defense systems. Interestingly, mutational data on Bdi1 suggest distinct specificities towards different targets. Some other proteins such as Ocr, AdrA, NARP1, and JSS1_004 have been shown to inhibit more than one phage defense system (*34, 385-387*). Ocr and AdrA achieve this by mimicking DNA and inhibiting defense systems that target specific DNA forms, while NARP1 overcomes NAD+ depleting systems by reconstituting NAD^+ from its degradation products. Here, Bdi1 and Bdi2 do not seem to target conserved effector domains among defense systems, akin to JSS1_004, which inhibits several distinct defense systems via a phosphorylation-based mechanism (*387*). These findings suggest that broadly acting anti-defenses are common among phages, and may provide evolutionary advantages, given the genomic constraints imposed by capsid size, restricting the amount of genetic material they can encode (*388*). This versatility likely arises from the functional similarities between different defense systems, which share common domains, sense similar phage-associated molecular patterns, and employ identical effector mechanisms (*389*). For Bdi1 and Bdi2, it is unclear if there is a common feature among the nucleic acid interacting defense systems targeted by these inhibitors, and whether there are more defense systems inhibited by Bdi proteins.

In addition to Bdi1 and Bdi2, we identified ZadI-1 as another inhibitor of Zorya type I. ZadI-1 likely binds to ZorC, highlighting the diverse mechanisms phages can use to inhibit the same phage defense system. This diversity may increase the evolutionary challenge for bacterial hosts, complicating the development of phage resistance. This diversity in inhibitory strategies is further demonstrated by TadIII-1, which inhibits Thoeris type III by physically blocking its TIR domain. Unlike previously identified anti-Thoeris proteins (Tad1 and Tad2), which sequester signalling molecules produced by TIR domains (*382, 383, 390*), TadIII-1 likely directly interacts with the TIR-domain proteins, thereby preventing effector activation. The inhibitory mechanism of DadIII-1 against Druantia type III remains elusive, as neither functional domains nor interactions could be predicted.

Our analysis revealed that most of the identified anti-defense genes are likely encoded on the leading strand of *Pbunavirus* genomes. This aligns with previous findings showing anti-CRISPR proteins and other anti-defense genes also tend to be encoded on the leading strand of conjugative plasmids (*391*) to facilitate early expression in the host. It further predicts that the exploration of leading strands in other phage families and genera could yield further discoveries of anti-defense mechanisms.

Interestingly, while amino acid similarity searches suggest that the anti-defense genes identified in this study are restricted to *Pseudomonas* phages and largely specific to the *Pbunavirus* genus, structural analysis reveals their broader distribution across phages of multiple taxa infecting diverse bacterial hosts from all major clades. This discrepancy between sequence- and structure-similarity approaches likely arises because anti-defense genes can undergo extensive sequence divergence while retaining their structure and function, making them undetectable by sequence-based methods. In such cases, structure-based approaches can uncover relationships between genes that have diverged too far at the sequence level. These findings highlight the value of structural approaches in assessing the prevalence of anti-defense genes across diverse phages, although their success ultimately depends on the accuracy of structure predictions.

While our study shows that *Pbunavirus* typically encode an average of three anti-defense genes, and up to five in some cases leading to the inhibition of up to seven defense systems, anti-defense genes targeting newly discovered defense systems remain massively underrepresented (*392*). It is likely that

the actual number of anti-defense genes per phage as well as the phage defense systems they inhibit, is higher than current estimates suggest. The discovery of potent, broadly acting anti-defense genes holds great potential to improve phage-based therapeutics, particularly by informing the engineering or selection of phages that can overcome abundant phage defense systems in phage-resistant pathogenic species (*146*). Our study contributes valuable anti-defense genes and presents a strategy for discovering others by exploring highly variable regions in genomes of other phage groups.

Materials and Methods

Bacteria and phages

A set of *Pbunavirus* from the Fagenbank were used to amplify candidate anti-defense genes. Phages ϕPa18, ϕPa22, ϕPa33, and ϕPa34 were used to test the effect of the candidate genes. *Escherichia coli* strain Dh5α was used for cloning plasmid pUCP20 with individual defense systems and plasmid pSTDesR with individual candidate anti-defense genes. *P. aeruginosa* strain PAO1 was used to clone both plasmids. All bacterial strains were grown in Lysogeny Broth (LB) at 37 °C and 180 rpm, or in LB agar (LBA, 1.5 % agar (w/v)) plates at 37 °C. *E. coli* strains containing pUCP20 or pSTDesR were grown in media supplemented with 50 µg/ml of streptomycin or 100 µg/ml of ampicillin, respectively. *P. aeruginosa* strains containing both plasmids were grown in media supplemented with 200 µg/ml of carbenicillin and 25 µg/ml of streptomycin. Phages were amplified in liquid media with PAO1, centrifuged at 3,000 × *g* for 15 min, filter-sterilized (0.2 µm PES), and stored at 4 °C.

Identification of variable genes within the pangenome of *Pbunavirus*

We used PPanGGOLiN v1.2.74 (*248, 249*) to identify the variable genes of 162 complete *Pbunavirus* from the Refseq database on August 1, 2023. A phylogenetic tree was constructed by aligning the large terminase protein sequence using MAFFT v7.525 (default settings)(*393*), and inferring by maximum likelihood phylogenetics with the use of IQ-tree v2.3.6 (-B 1000 –alrt 1000 –m TEST) (*394*) (**File S1**). Core genes (genes present in more than 95% of the genomes) were excluded from further analyses. The variable genes were clustered based on their sequence similarity using MMseqs2 (*183*) (default settings) and subsequently annotated (July 2024) using HMMER v3.3.2 (*127*), CD-SearchD v3.20 (*395*), InterProScan v5.60-92.0 (*333*), Foldseek v6 (*230*), Phobius v1.01 (*254*), MobiDB-lite v3.8.4 (*396*), Gene3D (*397*), PANTHER (*398*), and PROSITE (*399*). Previously identified anti-defense genes were searched for using dbAPIS (*400*) and Anti-DefenseFinder v1.0.0 (*392*) with default settings. Representative variable genes within genetic hotspots were selected and tested for their anti-defense activity (**Table S2**).

Cloning of the candidate anti-defense genes

The selected candidate anti-defense genes were amplified from *Pbunavirus* of the Fagenbank using primers (Integrated DNA Technologies) from **Table S6** with Q5 DNA Polymerase (New England Biolabs) or synthesized as gene fragments (gblocks, IDT). Plasmid pSTDesR was amplified by PCR to insert an RBS using primers from **Table S6**. PCR products were run on 1% agarose gels, and bands were excised and cleaned using the Zymoclean Gell DNA Recovery Kit (Zymo Research). The defense systems were cloned into pSTDesR using the NEBuilder HiFi DNA Assembly Master Mix and transformed into chemically competent NEB® 5-alpha Competent *E. coli* following the manufacturer's instructions. Plasmids were extracted using the GeneJET Plasmid Miniprep kit and confirmed by sequencing at Macrogen. Confirmed plasmids (**Table S7**) were transformed into PAO1 containing individual defense systems (*146*) by electroporation, as previously described (*129*). The recovered cells were plated on LBA plates supplemented with 200 µg/ml of carbenicillin and 25 µg/ml of streptomycin.

Protein complex prediction

Protein complex predictions were performed using AlphaFold 3 (*401*) for confirmed anti-defense and defense system proteins. Co-folding was carried out for each heteromeric and homomeric protein combination to model potential interactions. Predicted aligned error (PAE) plots were generated for

these co-folded proteins to assess the likelihood of complex formation. Low PAE values, particularly in regions of overlap between the two proteins, were interpreted as indicative of potential protein-protein interactions.

Selection for point mutations

Point mutations in the anti-defense proteins were selected using two criteria: 1) Multiple protein alignments from PSI-BLAST (*126*) and Clustal W v2.1 (*256*) to identify conserved amino acids (**File S2**); and 2) AlphaFold 3 co-folding predictions, followed by Predictomes (*381*) analysis to identify amino acids likely involved in protein-protein interactions (**Table S3**).

Cloning of point mutants of the anti-defense genes

Point mutations of the anti-defense genes in pSTDesR were obtained by round-the-horn site-directed mutagenesis using phosphorylated primers (**Table S6**) and Q5 DNA Polymerase. PCR products were digested with DpnI (NEB) and ran on 1% agarose gels. The bands of correct size were extracted and cleaned with the Zymo Gel DNA Recovery Kit. The amplified plasmids were ligated with T4 DNA ligase (NEB) at room temperature for 2 hours and transformed into chemically competent NEB 5-alpha Competent *E. coli*. Plasmids were extracted using the GeneJET Plasmid Miniprep kit and confirmed by sequencing (Macrogen). Confirmed plasmids (**Table S7**) were transformed into PAO1 by electroporation.

Efficiency of plating

Phage stocks were 10-fold serially diluted in LB and the dilutions were spotted onto double-layer agar plates of PAO1 with 1) empty pUCP20 and empty pSTDesR, 2) pUCP20 with a defense system and empty pSTDesR, 3) empty pUCP20 and pSTDesR with the candidate anti-defense gene, or 4) pUCP20 with a defense system and pSTDesR with the candidate gene, following the small plaque drop assay (*131*). The assays were performed using top agar (0.5% agar, w/v) supplemented with 200 µg/ml of carbenicillin and 25 µg/ml of streptomycin for plasmid maintenance, and 10 mM Rhamnose (*130*) for induction of expression of candidate anti-defense proteins.

Liquid culture collapse assays

Overnight bacterial cultures were diluted to an OD_{600} of approximately 0.1 in LB and transferred to 96-well plates. Phages were added at a multiplicity of infection of less than 1. The plates were then incubated at 37°C, with OD_{600} readings taken every 10 minutes over a 12-hour period using an Epoch2 microplate spectrophotometer (Biotek), with double orbital shaking (**Table S8**).

Building HMM models of the anti-defense proteins

HMM models of the anti-defense proteins were created using the PSI-BLAST-derived clusters of the *Pbunavirus* protein sequences. MAFFT was used to align the obtained protein sequences, and hmmbuild v3.4 (*127*) was used to build the HMM models. The HMM models were applied to all viral proteins available from NCBI using hmmsearch v3.4 (*127*), and scoring thresholds were set based on the bimodal distribution of the HMM-hit bitscore. The HMM bit score obtained for each anti-defense gene is as follows: Bdi1, 125; Bdi2, 135; ZadI-1, 100; TadIII-1, 135; and DadIII-1, 100. The HMM profiles of the anti-defense proteins can be found in **File S3.**

Search for the anti-defense proteins in all phages

The HMM models were used to search for homologues of known anti-defense proteins and those found in this study in all phage proteins from RefSeq (downloaded in September 2024). Phages containing anti-defense homologues were filtered for *Pseudomonas* phages and their taxonomy recovered from the RefSeq database (**Table S4**).

Structural-based homology detection

Foldseek easy-search (version 9.427df8a) was used in exhaustive search mode (--exhaustive-search 1) to identify structural homologs of the identified anti-defense genes in the BFVD (*384*), which contains 351,242 viral genes. The structural matches were filtered using a tmalnscore > 0.7, qcov > 0.8 and tcov

> 0.8. The code used to run the Foldseek algorithms and analyse results is available in **File S4**. The Foldseek 3Di sequences of the anti-defense genes identified in this study are available in **File S5**.

Comparison of the variable region of *Pbunavirus* genomes

Clinker v1.33 (*119*) was used to visualize the genomic context of the anti-defense genes among phages.

Quantification and statistical analysis

Unless stated otherwise, data are presented as the mean of biological triplicates ± standard deviation. A Bonferroni-adjusted p-value of less than 0.05 was considered significant.

Supplementary information

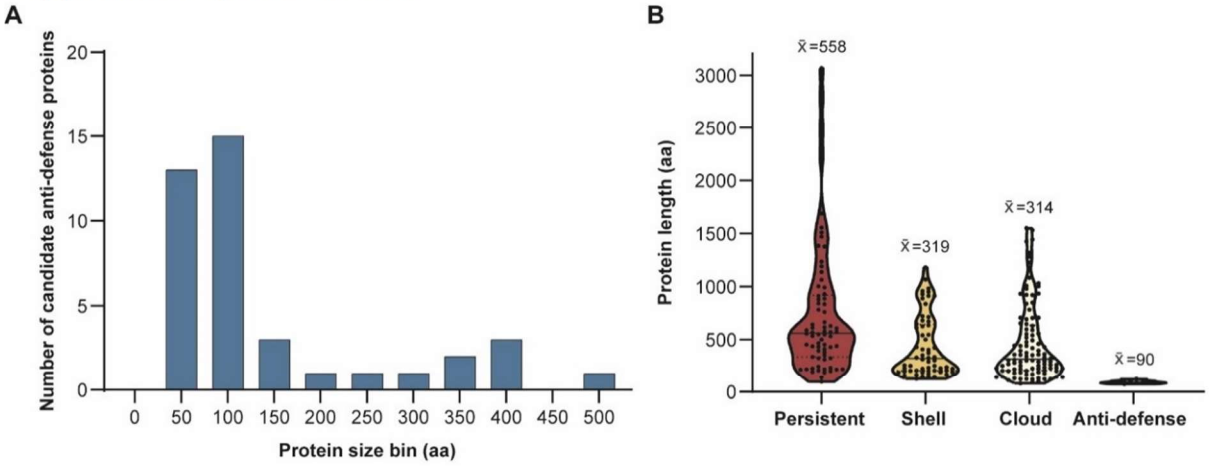

Figure S1. Protein length of candidate and confirmed anti-defense proteins. (a) Distribution of the protein lengths (in amino acids, aa) for the candidate anti-defense proteins. **(b)** Comparison of protein lengths for persistent, shell, cloud, and found anti-defense proteins in *Pbunavirus*. Each point represents the length of an individual protein, with the median length shown by a bold line and the interquartile range represented by dotted lines. The average protein length ($\bar{x}$) is show at the top of each violin plot.

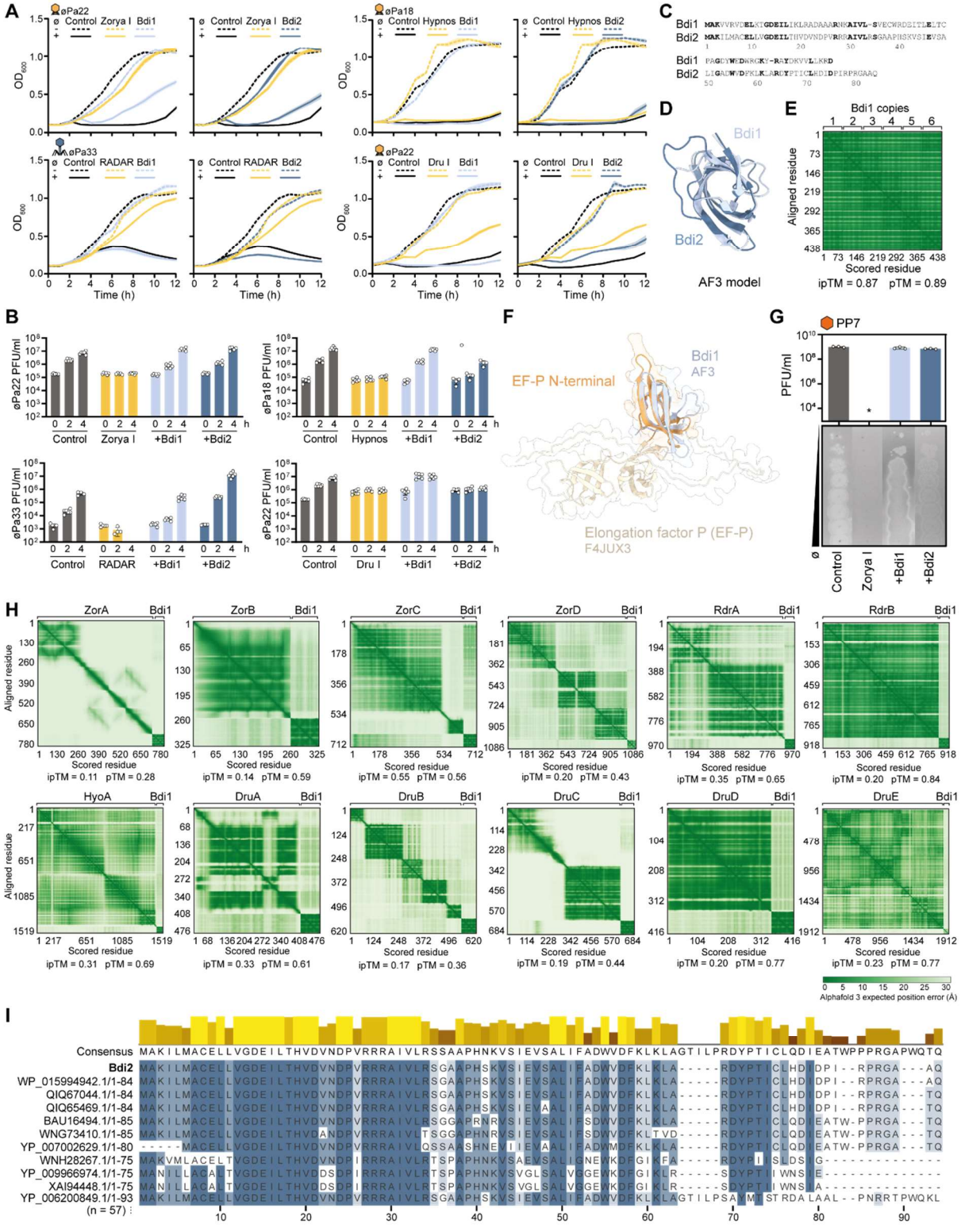

Figure S2 Characteristics of anti-defense proteins Bdi1 and Bdi2. (a) Effect of anti-defense proteins Bdi1 and Bdi2 on Zorya type I, RADAR, Hypnos, and Druantia type I protection against phage infection, evaluated by measuring bacterial growth (OD_{600}). Control is *P. aeruginosa* strain PAO1 having empty pUCP20 and pSTDesR plasmids. **(b)** Effect of anti-defense proteins Bdi1 and Bdi2 on Zorya type I, RADAR, Hypnos, and Druantia type I protection against phage infection, evaluated by measuring phage propagation (PFU/ml) over a period of 4h. Control is *P. aeruginosa* strain PAO1 containing empty pUCP20 and pSTDesR plasmids. **(c)** Amino acid similarity between Bdi1 and Bdi2. Identical amino acids are shown in bold. **(d)** Superimposition of Bdi1 and Bdi2 structures predicted with AlphaFold 3 (AF3). **(e)** Predicted aligned error (PAE) plot, pTM and ipTM values obtained by AlphaFold 3 co-folding of six Bdi1 proteins. **(f)** Superimposition of Bdi1 (predicted with AF3) and Elongation factor P (EF-P, Uniprot F4JUX3) showing that Bdi1 has significantly structural similarity to the N-terminal region of EF-P having a KOW-like domain. **(g)** Zorya type I protection against RNA *Leviviridae* PP7 and inhibition of the

protection by proteins Bdi1 and Bdi2. Phage titers were determined, with an example plate shown alongside a bar graph representing the average titers from three independent biological replicates. Error bars show standard deviation, and statistically significant differences ($p<0.05$) are marked with an asterisk (*). (h) Predicted aligned error (PAE) plots, pTM and ipTM values obtained by AlphaFold 3 co-folding Bdi1 with each protein of the phage defense systems it inhibits. (i) Multiple sequence alignment of Bdi2 with 56 homologous sequences. Ten representative sequences are shown, with consensus indicated above.

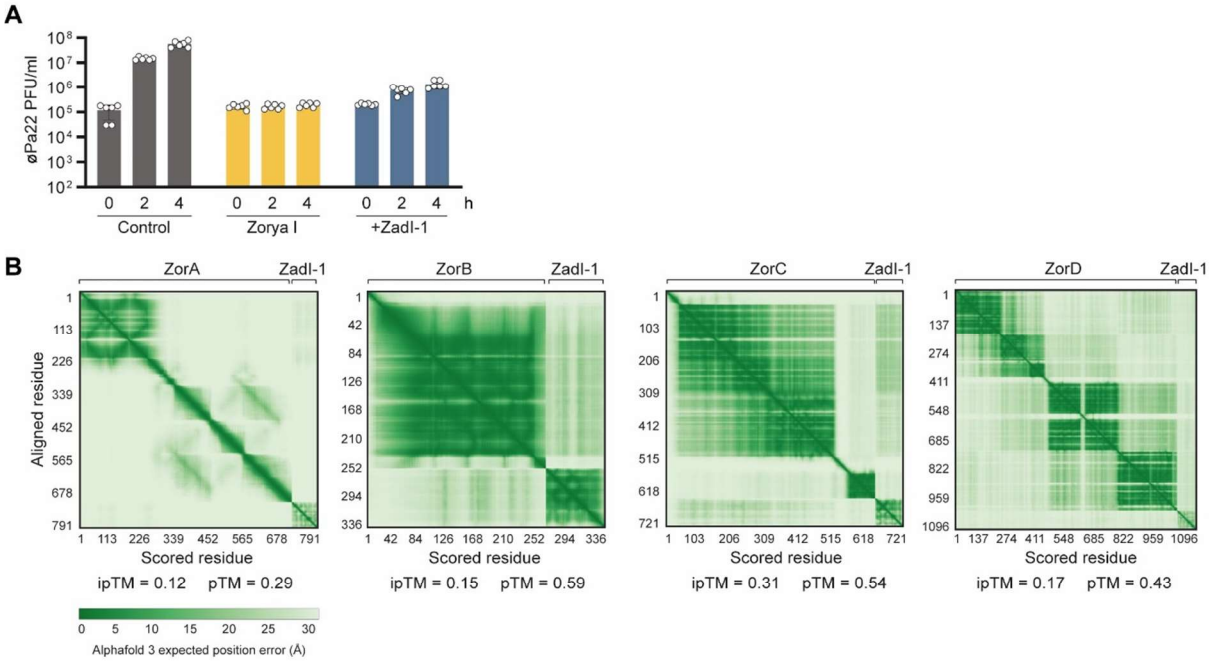

Figure S3. Characteristics of anti-defense protein ZadI-1. (a) Effect of anti-defense protein ZadI-1 on Zorya type I protection against phage infection, evaluated by measuring phage propagation (PFU/ml) over a period of 4h. Control is *P. aeruginosa* strain PAO1 having empty pUCP20 and pSTDesR plasmids. **(b)** Predicted aligned error (PAE) plots, pTM and ipTM values obtained by AlphaFold 3 co-folding ZadI-1 with each protein of Zorya type I.

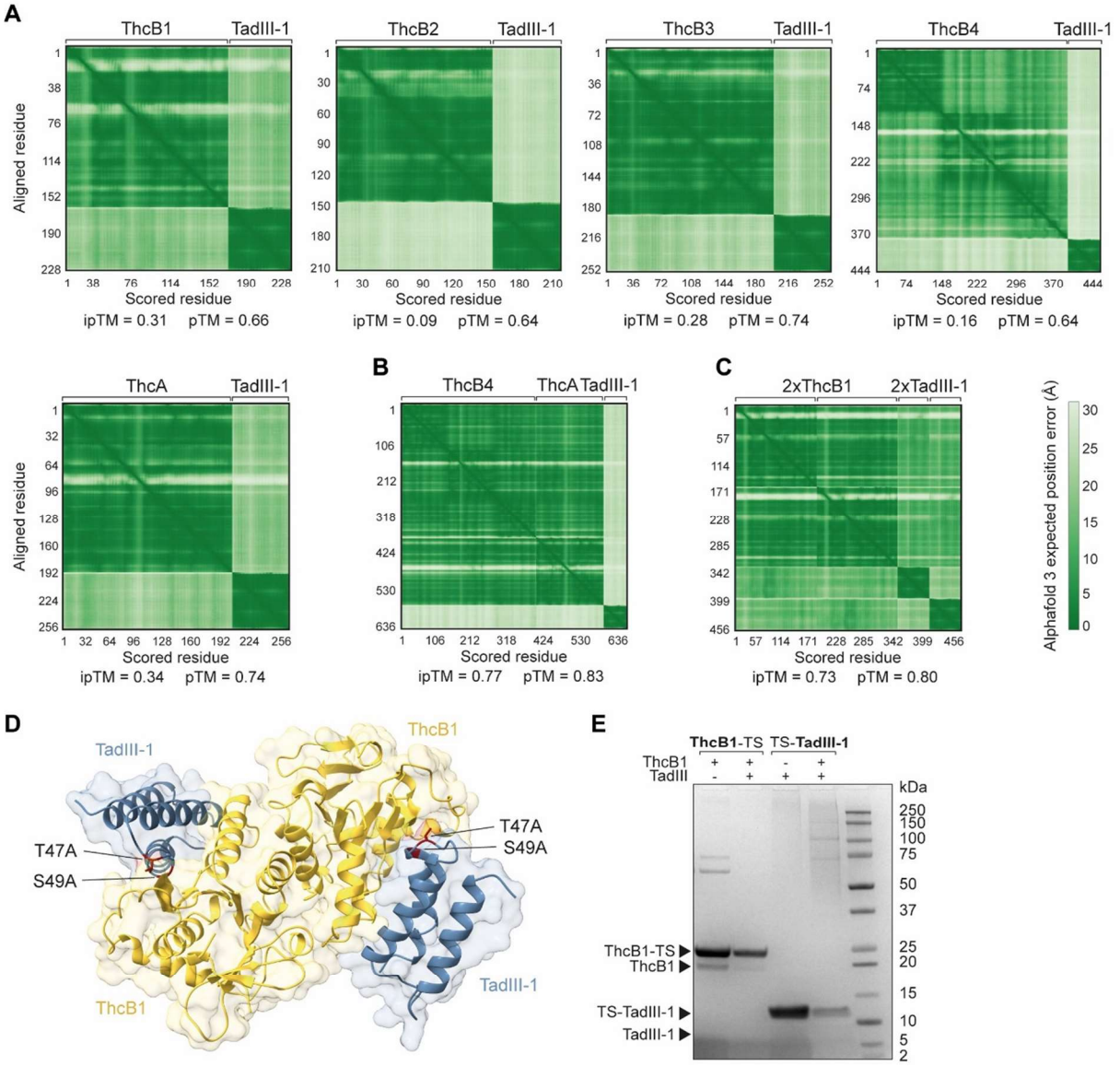

Figure S4. Predicted aligned error (PAE) plots, pTM and ipTM values obtained by AlphaFold 3 co-folding of TadIII-1 with Thoeris type III proteins, and experimental validation attempts. (a) Co-folding of TadIII-1 with each individual protein of Thoeris type III. **(b)** Co-folding of TadIII-1 with the predicted TIR-containing ThcB4 and SLOG-containing ThcA protein complex of Thoeris type III. **(c)** Co-folding of two TadIII-1 monomers with the predicted TIR-containing ThcB1 dimer of Thoeris type III. **(d)** Structure of the complex formed between two TadIII-1 monomers (blue) and a ThcB1 dimer (yellow), as predicted by AlphaFold 3. Two amino acids of TadIII-1, which are predicted to interact with ThcB1 and mutated in Figure 4E, are shown in red. **(e)** Co-purification assay testing interactions between TadIII-1 and ThcB1. Tagged and untagged combinations of the proteins were co-expressed in E. coli BL21-AI, subjected to affinity purification via Strep-tag, and analyzed by SDS-PAGE.

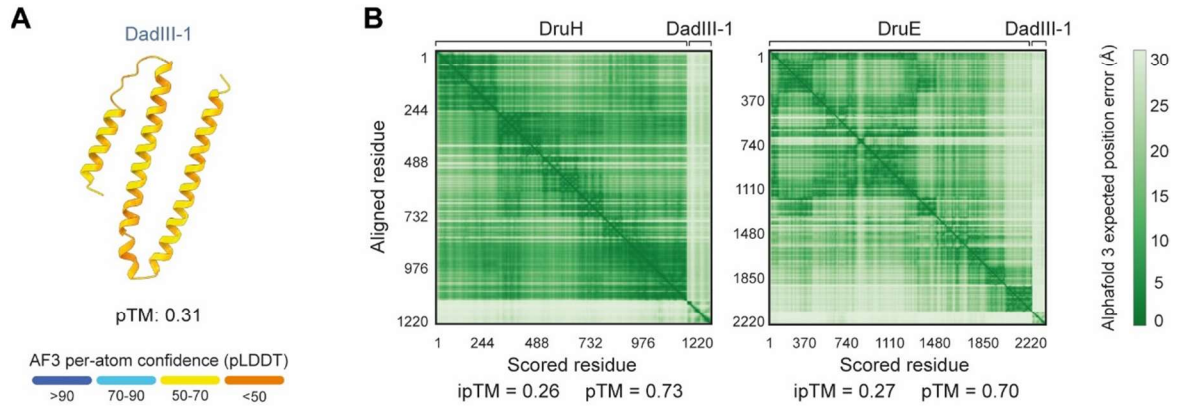

Figure S5. Characteristics of anti-defense protein DadIII-1. (a) Structure of TadIII-1 predicted by AlphaFold 3 (AF3) and colored by pLDDT. **(b)** Predicted aligned error (PAE) plots, pTM and ipTM values obtained by AlphaFold 3 co-folding of DadIII-1 with each protein of Druantia type III.

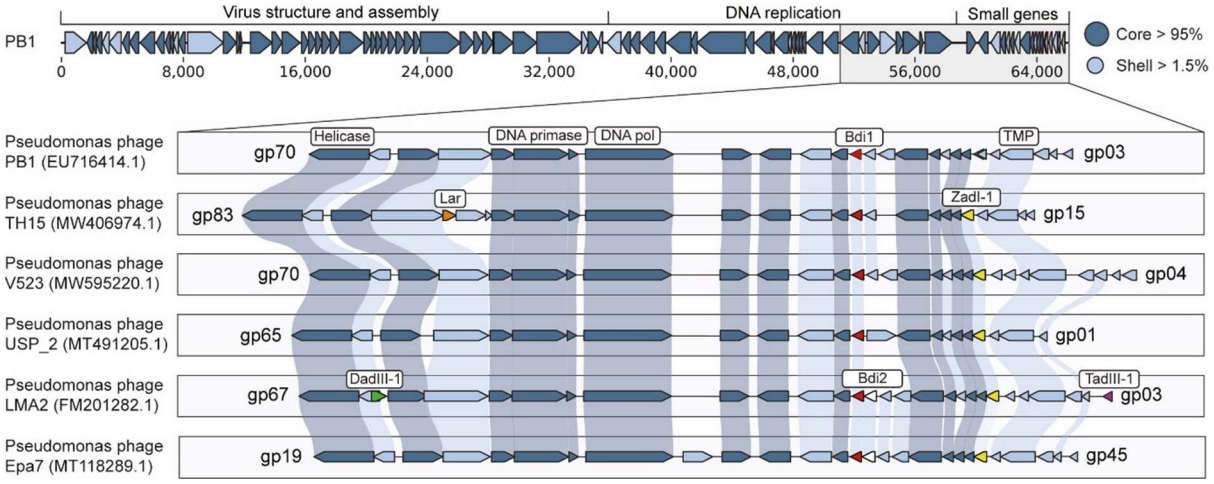

Figure S6. Genomic structure of the representative *Pbunavirus* PB1, with a zoomed-in view of the highly variable region having anti-defense genes. The gene composition of this variable region is shown for six distinct *Pbunavirus* genomes.

The following supplementary files are available online at doi: 10.1016/j.chom.2025.06.010

Table S1 Summary of PPanGGoLiN analysis of 162 *Pbunavirus* genomes.
Table S2 List of candidate genes tested for anti-defense activity.
Table S3 Predictomes analysis of the AlphaFold 3-predicted complex 2ThcB1:2TadIII-1.
Table S4 Prevalence of anti-defense genes in *Pseudomonas* phages of different taxonomies.
Table S5 Structural homologs of the anti-defense genes found in the BFVD.
Table S6 List of primers and gene fragments used in this study.
Table S7 List of plasmids used in this study.
Table S8. Raw data of liquid culture assays.
File S1. Phylogenetic tree of 162 *Pbunavirus*.
File S2. Multiple sequence alignments of the individual anti-defense proteins identified in this study.
File S3. HMM profiles of the individual anti-defense proteins identified in this study.
File S4. Code for searching structural homologs of the anti-defense proteins.
File S5. Foldseek 3Di sequences for the anti-defense proteins identified in this study.

Chapter 8



Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations

tRNAs in bacteriophage genomes are widespread across bacterial host genera, but their exact function has remained unclear for more than 50 years. Several hypotheses have been proposed, and the most widely accepted one is codon compensation, which suggests that phages encode tRNAs that supplement codons that are less frequently used by the host. Here, we combine several observations and propose a new hypothesis that phage-encoded tRNAs counteract the tRNA-depleting strategies of the host using enzymes such as VapC, PrrC, Colicin D, and Colicin E5 to defend from viral infection. Based on mutational patterns of anticodon loops of tRNAs encoded by phages, we predict that these tRNAs are insensitive to host tRNAses. For phage-encoded tRNAs targeted in the anticodon itself, we observe that phages typically avoid encoding these tRNAs. Further supporting the hypothesis that phage tRNAs are selected to be insensitive to host anticodon nucleases. Altogether our results support the hypothesis that phage-encoded tRNAs have evolved to be insensitive to host anticodon nucleases.

A modified version of this chapter has been published as

van den Berg, D. F., van der Steen, B. A., Costa, A. R., & Brouns, S. J. J. (2023). Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations. *Elife*, *12*, e85183.

Introduction

Transfer RNAs (tRNAs) were first discovered in the 1950s (*402*) and have since been found to play a vital role in the central dogma of molecular biology in all living systems (*403*). During the 1960s, tRNAs were also reported in viruses of bacteria (phages) (*404*). We now know that phage-encoded tRNAs are widespread, especially among virulent phages (*405*). Multiple hypotheses have been proposed for the role of these phage-encoded tRNAs. The most established being codon compensation, where codons rarely used by the host but necessary to the phage are supplemented by the tRNAs carried by the phage (*405*). Why phages are pushed towards these alternative codons is generally believed to be a side effect of differences in the GC content of phage and host (*405-407*). A recent study by Yang *et al.* (2021) may have hinted at an additional factor: phage tRNAs represent a means to counteract the depletion of host tRNAs that occurs as an early response to phage infection (*408*). The host uses several mechanisms to deplete its tRNA pool, such as downregulating the expression of its tRNAs, modifying tRNAs to make them unusable for translation, and most notably cleaving the tRNAs using anticodon nucleases (*408-412*). Exactly what activates host tRNA cleavage is often unknown; an exception is anticodon nuclease PrrC, which cleaves tRNA-Lys(ttt) when triggered after sensing phage-encoded protein Stp (*413*). In response, phage T4 encodes a tRNA ligase that repairs the cleaved tRNA-Lys (*413*). Recently, a phage encoding a tRNA was found to replenish the host tRNA depleted by the Retron Ec78 phage defense system, thereby preventing the inhibition of phage propagation (*39*). However, it remains unclear how phage tRNAs avoid being degraded by the same mechanism that results in the depletion of host tRNAs during phage infection. Here, we have analysed phage-encoded tRNAs and hypothesize that the tRNAs encoded by phages are insensitive to tRNA anticodon nuclease activity, preventing depletion of the tRNA pool and translation stalling during phage infection.

Results and discussion

To investigate our hypothesis that phage tRNAs are insensitive to tRNA anticodon nucleases, we analysed the tRNAs encoded by a large and well-characterized dataset of tRNA-rich bacteriophages (33 tRNAs per phage on average) that infect mycobacteria: mycobacteriophage cluster C1 (*414*) (**Figure 1a,b**). We specifically selected this mycobacteriophage dataset because the bacterial host encodes a range of well-characterized tRNA nucleases (tRNAses), such as VapC, MazF, and RelE (*415-420*). A subset of these tRNAses target the tRNA anticodon loop and are activated upon various stress responses, including phage infection (*421*). Upon activation, these anticodon nucleases cleave specific tRNAs in conserved regions within the anticodon loop to inactivate the tRNAs and thereby regulate protein translation of the host (*415*). The cleavage region within the tRNA anticodon loop is sequence-dependent and highly specific to the type of tRNA. Mutations in the recognition and cleavage sequence within the anticodon loop have been found to result in insensitivity to these anticodon nucleases (*415, 417*). To check phage tRNAs for mutations that are known to cause insensitivity to anticodon nucleases (*415-417*), we compared them to those of their host. We found that all 10 mycobacteriophage-encoded tRNAs that are targeted by anticodon nucleases contained anticodon loop mutations known to affect cleavage by VapC (**Figure 1c, Supplementary File 1a,1b**). These findings support the idea that phage-encoded tRNAs are insensitive to cleavage by anticodon nucleases (**Figure 1c**). We propose that these phage tRNAs serve as a mechanism to counteract the depletion of tRNAs by anticodon nucleases during phage infection, thereby allowing the phage to translate its proteins and successfully complete the infection cycle (**Figure 1d**).

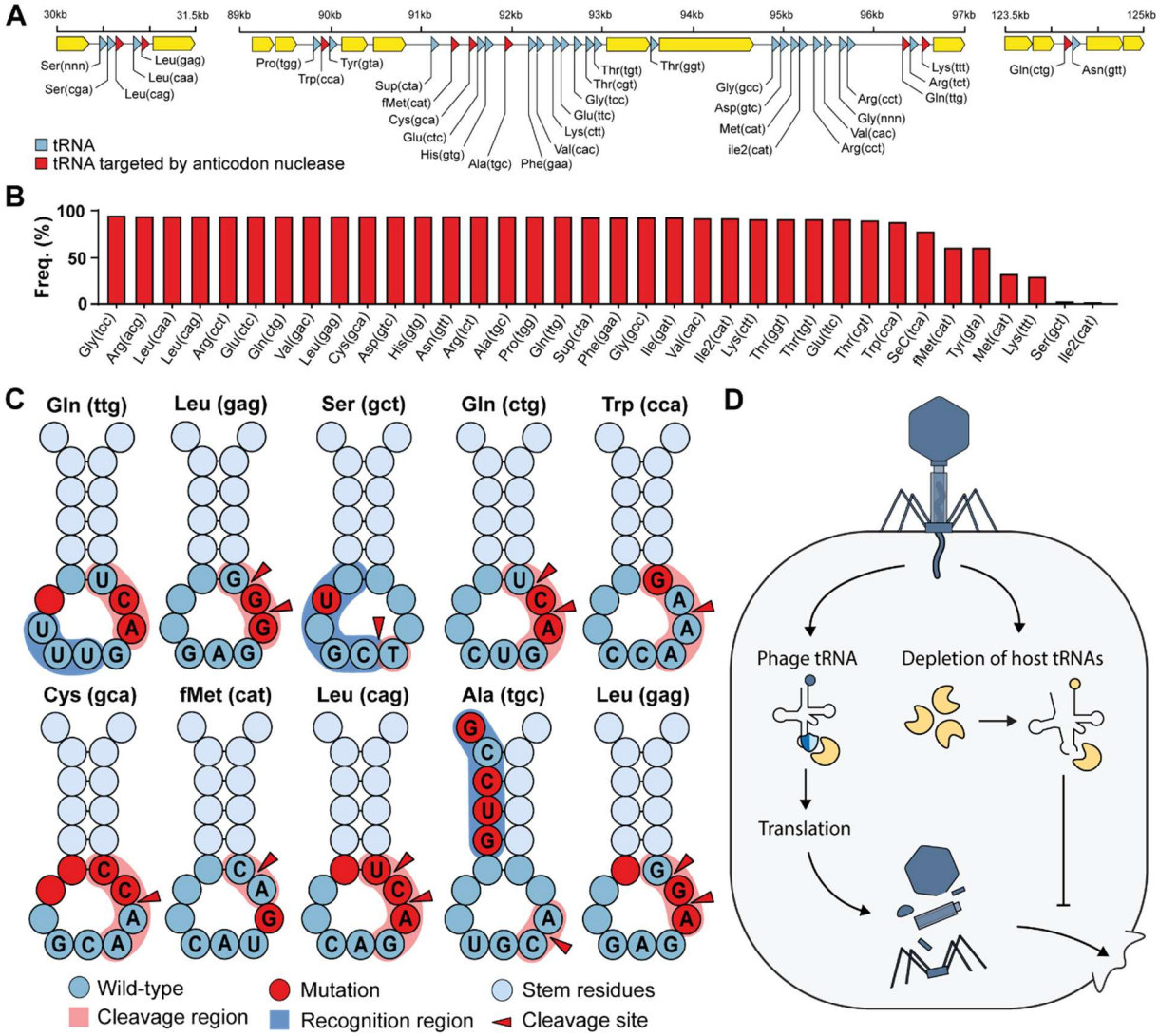

Figure 1 Phage tRNAs are predicted to be anticodon nuclease resistant. (a) The genomic context of the tRNA clusters containing 36 tRNAs present in C1 mycobacteriophage Rizal (*414*). **(b)** Prevalence of individual phage-encoded tRNAs in the C1 mycobacteriophage cluster, composed of 161 phages. **(c)** Mutations in the anticodon-loop of phage tRNAs in comparison to host tRNAs, located in the cleavage site of anticodon nucleases. **(d)** Proposed mechanism of action of phage tRNAs. During phage infection, tRNAses are activated and deplete the host tRNA pool via tRNA cleavage to prevent phage propagation. Phage tRNAs are insensitive to cleavage allowing the phage to propagate.

In addition, we observe that mycobacteriophages rarely encode tRNAs that are cleaved within the anticodon itself (**Supplementary File 1b**), suggesting that these anticodons are avoided by the phage and that no tRNAs evolved to be insensitive to cleavage. Specifically, this avoidance is seen for the serine-coding tRNAs that are cleaved at the GA site within the anticodon (tRNA-Ser(gga), tRNA-Ser(tga), tRNA-Ser(cga), and tRNA-Ser(aga)) (*415*) (**Supplementary File 1b**). To compensate for this, the phage encodes a serine tRNA (tRNA-Ser(gct)) that is not targeted by nucleases. We observed the same avoidance for known targets of RelE (*420*), including stop codon (cta), tRNA-Leu(tta), and tRNA-Gln(cga) (**Supplementary File 1b**). Interestingly, when comparing the codon frequency of phage genes, we observed no differences in the codon frequency between codons for which the phage encodes a tRNA and those for which it does not (Welch Two Sample t-test, t = 1.0471, df = 41.591, p-value = 0.3011) (**Supplementary File 1b**). Moreover, we did not observe a difference in the codon frequency between phage and host genes (Paired t-test, p-value = 0.999) (**Supplementary File 1b**). We also found that in only 2 out of 23 instances, the preferred codon (i.e. the most frequently encoded codon per amino acid) of the phage did not match that of the host (tRNA-Val(cac) and tRNA-Ala(cgc))

(**Supplementary File 1b**). Together, these observations suggest that the phage-encoded tRNAs were likely not selected for codon compensation. Overall, our findings support the hypothesis that phage tRNAs in Mycobacteria evolved to resist anticodon nucleases to overcome host tRNA-depletion strategies that limit phage propagation. To investigate if this hypothesis could extend more generally outside Mycobacteria, we examined other species with anticodon nucleases, including *Shigella flexneri* (VapC-LT2) (*422*), *Escherichia coli* (VapC, PrrC, Colicin D, and Colicin E5) (*413, 422-424*), and *Salmonella enterica* (VapC/MvpT) (*422*). It is important to point out that the following analyses are based solely on the known cleavage site, and that the recognition sequence of these anticodon nucleases is unknown, thus limiting the ability to identify possible insensitivity-causing mutations in the recognition site of the anticodon loop. Despite this limitation, we observed mutations in the anticodon loop near or at the cleavage site of anticodon loop-targeted phage-encoded tRNAs in nearly all instances (9 out of 11) (**Figure S1; Supplementary File 1c**). When the anticodon itself was the target of the anticodon nuclease, as is the case for VapC-LT2 in *S. flexneri*, we found that the phage generally avoided encoding the target tRNA (tRNA-fMet), except for two instances. One of these is the phage-encoded tRNA-Lys(ttt) in coliphages. In this instance, the phage encodes a tRNA-Lys ligase (rnl1 & rnl2) that repairs tRNA-Lys(ttt) after it has been cleaved by PrrC (*413*). The other instance involves phage-encoded tRNAs targeted by *E. coli* Colicin E5. The cleavage activity of Colicin E5 depends on modifying the wobble position (*424*), which might be absent in these phage tRNAs possibly caused by mutations in the anticodon loop, which we observed for three out of the four targeted tRNAs (**Figure S1c**). Alternatively, these mutations might affect the recognition by Colicin E5. In summary, our findings in Mycobacteria are consistent with almost all (9 out of 11) currently known instances of targeted phage tRNAs in Enterobacteria (**Supplementary File 1c**). The discrepancies in these cases can be accounted for by the presence of tRNA ligases, reliance on tRNA modifications of the cleavage site, or mutations within the uncharacterized recognition site. We speculate that our hypothesis may be extended beyond Mycobacteria and Enterobacteria, given the ubiquity of virus-encoded tRNAs and host tRNAses (*424-426*).

Conclusion

We propose that phage-encoded tRNAs escape targeting by host tRNAses via insensitivity causing mutations within the tRNA cleavage and recognition site. This proposed hypothesis can be helpful in selecting or engineering bacteriophages capable of infecting hosts containing anticodon nucleases.

Materials and Methods

tRNA analysis in mycobacteria and mycobacteriophages

Mycobacterium smegmatis MC²-155 (CP000480.1) and Mycobacterium tuberculosis H37Rv (NC_000962.3) were used as references for obtaining the host tRNA sequences. All C1 cluster mycobacteriophage genomes were downloaded from PhagesDB (*414*) the 1st of September 2022. tRNAs were annotated using Aragorn (*427*) and tRNAscan-SE (*428*). Codon frequency and fraction were determined using cusp (EMBOSS v6.6.0.0) (*429*).

tRNA analysis in non-mycobacteria species and phages

We performed a literature search for all anticodon nucleases with known specificities on the 1st of March 2023. These were found in *Leptospira interrogans (430)*, *S. flexneri* (*422*), *E. coli* (*413, 422-424*), *S. enterica* (*422*), *Deinococcus radiodurans (431)*, and *Geobacillus kaustophilus (432)*. We excluded *L. interrogans*, *D. radiodurans*, and *G. kaustophilus* due to the absence of known tRNA-encoding phages. The species-specific tRNA sequence of the host for each anticodon in question was obtained using tRNAviz (*433*). Species-specific phages were obtained from the public PhageAI database (*116*) on the 1st of March 2023. tRNAs were annotated and analyzed as indicated above for mycobacteria and mycobacteriophages.

Supplementary information

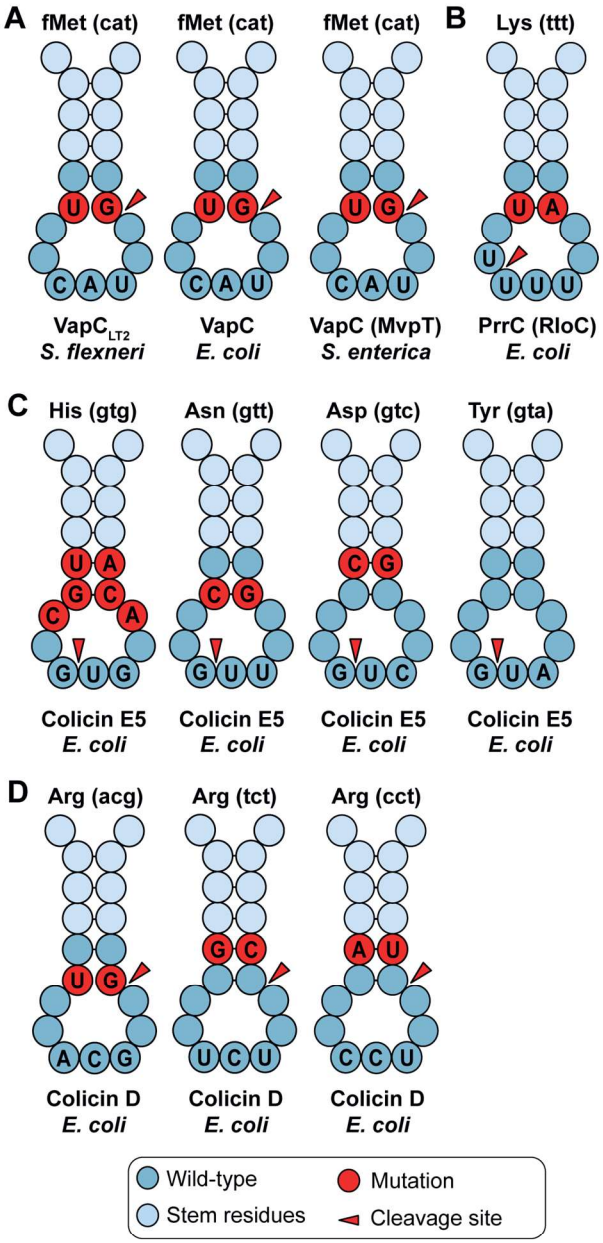

Figure 1S Phage tRNAs from Enterobacteria. (a) Mutations in the anticodon-loop of phage tRNAs in comparison to host tRNAs, located in the cleavage site of anticodon nucleases of tRNA fMet (cat) that are targeted by VapCs. **(b)** Mutations in the anticodon-loop of phage tRNAs in comparison to host tRNAs, located in the anticodon loop of tRNA Lys (ttt) that is targeted by PrrC (Rloc), which is repaired by the tRNA-Lys ligase. **(c)** Mutations in the anticodon-loop of phage tRNAs in comparison to host tRNAs, located in the anticodon loops of tRNAs that are targeted by Colicin E5 in the anticodon itself **(d)** Mutations in the anticodon-loop of phage tRNAs in comparison to host tRNAs, located in the anticodon loops of tRNAs that are targeted by Colicin D from E. coli.

The following supplementary files are available online at doi: 10.7554/eLife.85183

Supplementary file 1. Phage versus host-encoded tRNA comparisons.

Chapter 9



Phage tRNAs: Decoding the Enigma

The presence of tRNAs in bacteriophage genomes has intrigued scientists since their discovery in the early 1960s, as phages were believed to rely on the host tRNAs for their translation. Over the years, a multitude of hypotheses have been explored, providing evidence that phages with different lifestyles utilize tRNAs in distinct ways. In recent years, several studies have provided evidence that phage tRNAs play a crucial role in evading phage defense systems. In this review we summarize the current state of the field of phage tRNAs, highlighting their diverse roles in phage infection. We also discuss the host response to phage tRNAs and the application of this knowledge to improve phage-based therapeutics to combat bacterial infections.

A modified version of this chapter has been published as
Berg, D. F., & Brouns, S. J. J. (2025). Phage tRNAs: decoding the enigma. *Trends in Microbiology*.

Viruses of Bacteria encode transfer RNAs for Translation

Transfer RNAs (tRNAs) were first discovered in the 1950s (*402*) and have since been recognized to be vital in the central dogma of molecular biology in all living systems (Box 1) (*403*). Surprisingly, during the 1960s, tRNAs were also reported in viruses of bacteria (phages) (*404*). This finding challenged the view of viruses at that time, which were believed to hijack the host translation machinery for their replication, including the host tRNAs instead of encoding their own (*434*). Follow-up studies observed that the tRNAs from *Escherichia coli* phage T4 were expressed early in the infection and that removing these tRNAs from the genome resulted in a 20-fold reduced infectivity (*435, 436*). Intriguingly, recent efforts found that phages more often encode tRNAs than previously thought, observing tRNAs in more than 30% of phage genomes, ranging from 1 to 62 tRNAs (*437*). Yet, despite their widespread occurrence, the precise function of phage tRNAs during the **infection cycle** (see Glossary) remains unclear. Recent breakthroughs have uncovered unexpected roles for these tRNAs, from circumventing the immune response of the host to regulating the phage replication cycle. We summarize these recent advances in this review.

Box 1: The Cellular Function of tRNAs

tRNAs make up around 15% of the total ribonucleic acid (RNA) in the bacterial cell, making them the second most abundant RNA after ribosomal RNA (rRNA) (*438*).

After transcription, tRNAs go through several maturation steps before becoming functional, including the removal of the 5' and 3' precursor sequences by a variety of RNases, and the addition of a CCA sequence on its 3' end by the nucleotidyl transferases (NTases) (*439, 440*). This CCA sequence is charged with its corresponding amino acid by the aminoacyl-tRNA synthetase (aaRS) based on the anticodon of the tRNA (*35, 441, 442*). Additionally, tRNAs undergo various post-transcriptional modifications to stabilize the tertiary structure, regulate its translation efficiency, and tune the precise base-pairing of its anticodon (*35, 441*).

After the tRNA is mature and charged with their amino acid, elongation factor thermo-unstable (EF-Tu) guides the tRNA to correct **A-site** of the ribosome during translation (*443*). During translation, the amino acid that is attached to the tRNA is incorporated into the nascent polypeptide chain by the ribosome, forming a peptide bond between the amino acids, in accordance with the codon on the mRNA template (*441, 442, 444, 445*).

To ensure that each of these steps is performed with high accuracy, tRNAs consist of several key features (**Figure I**) (*442*):

- Acceptor stem: Serves as the amino acid attachment site acid at its conserved CCA site at the 3' end during aminoacylation (addition of the amino acid to the tRNA).
- Dihydrouridine (d)-arm: Serves to stabilize the L-shaped structure of the tRNA that is required to fit into the ribosome during translation, as well as being crucial for the recognition of the aminoacyl tRNA synthase (aaRS) during aminoacylation.
- Variable loop: Differs in length among different tRNAs, and together with the D-arm aids the aaRS with distinguishing between tRNAs for the correct aminoacylation.
- TΨC-arm: Assists in stabilizing the L-shaped structure of the tRNA together with the D-arm, as well as aiding the binding of the mature tRNA to the correct position within the ribosome during translation.
- Anticodon loop: part of the anticodon stem which is responsible for the base pairing of the tRNA with the mRNA codon.
- Post-transcriptional modifications: Can serve a multitude of functions, such as stabilizing the tertiary structure, regulating its translation efficiency, and tuning the base-pairing of its anticodon. For example, through **7-cyano-7-deazaguanine** and **archaeaosine** modifications of specific tRNA residues (*446, 447*).

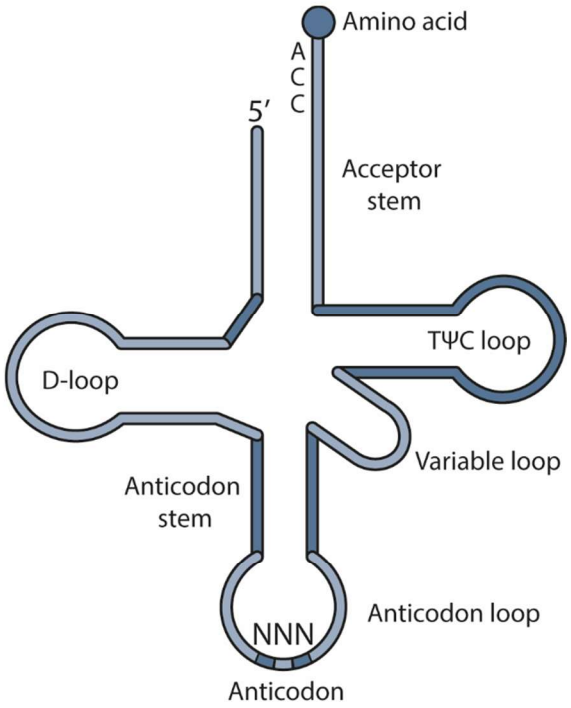

Figure I. Overview of the structural composition of the transfer RNA. Shown are the features of the transfer RNA (tRNA), consisting of the amino acid, acceptor stem, D-arm, anticodon stem, anticodon loop, variable loop and TΨC loop, as well as the CCA-end of the acceptor stem.

The disruption of any of the tRNA maturation steps leads to a variety of detrimental effects, including non-functional tRNAs, mischarging, and misincorporation of the tRNA during translation (*448*). To resolve these tRNAs from negatively affecting translation, cells have evolved several quality control checks during each of these steps that degrade or repair faulty tRNAs (*448*).

The Diverse Roles of Phage tRNAs

The role of tRNA genes in phage genomes has been extensively discussed and explored, with numerous studies identifying a diverse set of potential roles that can be attributed to phage tRNAs (*405, 449*). One important factor in this discussion is that the lifestyle of the tRNA-encoding phage appears to determine the specific function of the tRNA. These lifestyles can be grouped into two main types: **virulent and temperate phages** (*6*). Virulent phages follow the **lytic cycle**, where phage replication and formation of new particles start upon genome ejection into the host cell, while temperate phages can follow the **lysogenic cycle** where they integrate into the genome of the host. Here, the temperate phage can remain integrated, passively replicating itself with the host chromosome. Under stressful conditions for the host, temperate phages excise from the host genome and transition to the lytic cycle, like virulent phages, to abandon ship. Some types of temperate phages, such as plasmid-phages, deviate from the lysogenic cycle by not integrating in the chromosome, but instead, these phages replicate as plasmids (*450*). Interestingly, temperate phages typically encode just a few phage tRNAs, while virulent phages are found to encode a multitude, sometimes up to 61 different tRNAs (*323, 405*). This suggests that phages with different lifecycles may have different needs for tRNAs. Over the years, several hypotheses and roles have been proposed for phage tRNAs.

Diverse Functions of tRNA from Temperate Phages

As mentioned above, for temperate phages that integrate into the tRNA gene of their hosts, the encoded phage tRNAs may replace the function of the dysfunctional host tRNA that the temperate phage integrates into during its lysogenic cycle. In the majority of observed cases, host tRNA genes are a prime integration target of temperate phages, since host tRNAs are prevalent and conserved across a large diversity of hosts, making it a reliable integration site (*451, 452*). A disadvantage of integrating into the host tRNAs is that the function of this host tRNA is disrupted, leading to a disadvantage for the host, and consequently, for the integrated phage that relies on host survival for its own existence. To resolve this downside, some phages encode the same tRNA in their genome to replace the function of the integrated host tRNA (**Figure 1a**) (*405, 453*). Alternatively, some temperate phages, such as Mlo38S, are flanked by tRNA halves at their 3' and 5' end of their genome, which forms two complete functional tRNA copies when it integrates into the tRNA of the host (**Figure 1b**) (*454, 455*). Notably, temperate phages often encode multiple tRNAs, suggesting that the function of phage tRNAs in temperate phages is not limited to compensating for the disrupted tRNA of the host.

Several hypotheses could explain these additional tRNAs in temperate phage genomes. One hypothesis is that tRNAs are acquired by random chance during **transduction events**, however, no study has found evidence for this hypothesis (*405*). Moreover, tRNAs of temperate phages may function as an integration site for beneficial **mobile elements**, such as those encoding phage defense systems. This benefits the temperate phage in inter-viral conflicts, as observed for *E. coli* P2-like phages and their P4-like **satellites**, which encode phage defense systems in a specific integration site in their genome (*62*). Lastly, besides possibly providing an advantage in inter-viral conflicts, temperate phage tRNAs may provide an advantage in their conflict with the host defenses. A recent study discovered that temperate Mu phages use a charged $tRNA^{gly}$ from the host as a substrate to hypermodify the adenines of the phage DNA, preventing host DNases from acting (*456*). It is possible that other phages utilize a similar mechanism with their own charged tRNAs.

In addition to regular tRNAs, temperate phages have also been found to encode suppressor tRNAs ($tRNA^{sup}$), also known as **nonsense suppressors** (*457-459*). These tRNAs match the stop codon on the mRNA, and instead of terminating the translation process, the tRNA is charged with an amino acid (*460*). This incorporation results in readthrough of the stop codon, preventing translation termination, resulting in translation of dysfunctional protein. The role of these suppressor tRNAs in temperate phages has recently been proposed as preventing the correct translation of lytic genes while being in the lysogenic phase, especially in crAss-like phages where the TGA stop codon is re-assigned to glutamine (**Figure 1c**) (*457-459, 461*).

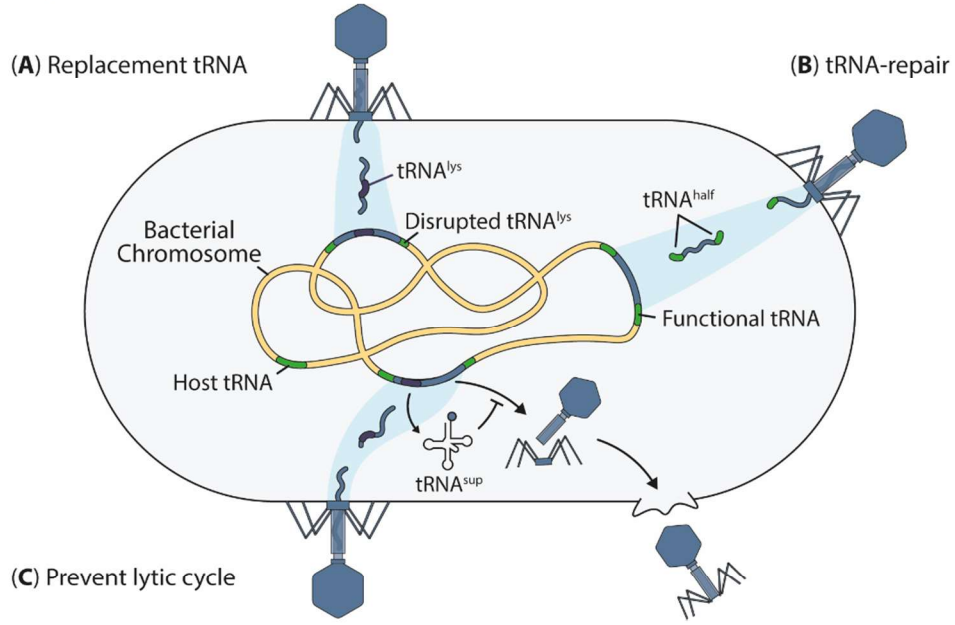

Figure 1: Overview of the tRNA conflict between temperate phages and their hosts. (a) Phage tRNAs are encoded by temperate phages to compensate for the disrupted host tRNA which is caused by the integration of the temperate phage into the host tRNA during its lysogenic cycle. **(b)** Some temperate phage genomes are flanked by two tRNA halves that restore the function of the disrupted host tRNA after integration. **(c)** Suppressor tRNAs are encoded to prevent the premature expression of the lytic genes when the temperate phage is integrated into the host genome.

Phage tRNAs in Phage Replication and Codon Usage.

While temperate phages only encode a few phage tRNAs, virulent phages can encode up to 61 different tRNAs. Similar to temperate phages, some virulent phages, such as *E. coli* phage T4, also encode suppressor tRNAs (*460, 462*). These suppressor tRNAs may function to alter the codon table of the phage, possibly interfering with the translation of the proteins involved in the host response during phage infection (*461, 463*). Additionally, phage tRNAs of virulent phages may facilitate phage infection through codon compensation, where codons rarely used by the host but necessary to the phage are supplemented by the tRNAs encoded by the phage (**Figure 2**) (*405*). Why phages evolved to differ in codon usage from their hosts remains largely unknown. One factor may be that the translation efficiency of certain codons is affected by phage infection, driving the phage to encode more of the efficiently translated codons, while avoiding those that are negatively affected (*464, 465*). Meanwhile, the host evolved to encode the codons that are well translated during non-phage infection (*466*). Another relevant factor for encoding phage tRNAs is to bridge the discrepancies in the **GC-content** that are often observed between the phage genome and its host, resulting in a mismatch between the frequently used codons of the host with those of the phage, as well as opening up the **host range** from GC-rich hosts to those that are those with a lower GC-content (*405, 407, 467-469*). This is in line with virulent phages, that do not encode tRNAs, translate their highly expressed genes using the most abundant tRNAs in the host pool, while tRNA-encoding virulent phages use their own tRNAs, especially for their late-stage genes (**Figure 2**) (*405, 406, 449, 453*). This discrepancy may be explained by the **tRNA pool** dynamics during phage infection. Early in the infection, the host tRNA pool is still abundant enough for efficient translation (**Figure 2**). However, at a later stage, phage tRNAs take over (**Figure 2**) (*408*). This shift shows the dynamic nature of the tRNA pool during phage infection.

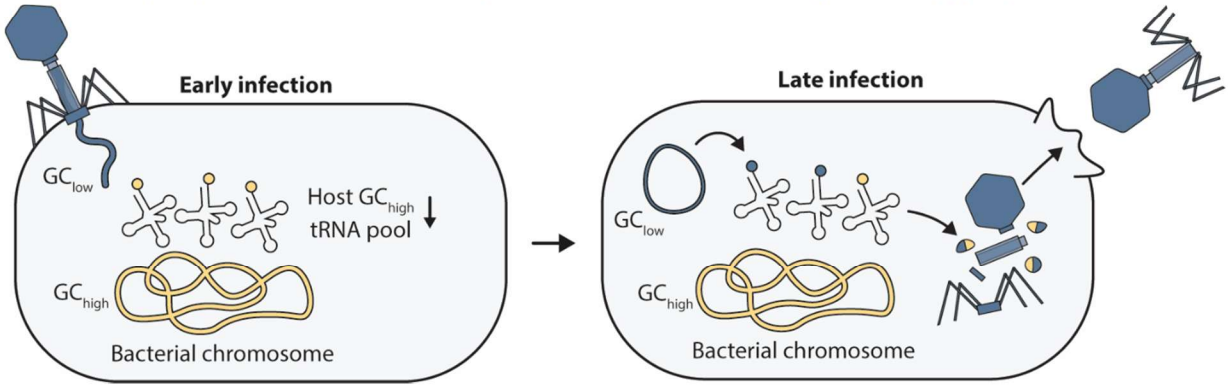

Figure 2. Phage tRNAs in phage replication and codon usage. Phage tRNAs may compensate for the differences in GC-content between the phage and host. Early phage translated genes are translated using the most abundant host tRNAs. However, at later stages, phage genes are translated using the tRNAs of the phage.

Phage tRNAs Replenish the Depleted Host tRNA Pool

The depletion of the host tRNA pool during phage infection was first observed several decades ago (*470*). Since then, various causes for this phenomenon have been identified, including the downregulation of the host tRNAs during the early stages of the infection (*471, 472*) and the active cleavage of the tRNA pool (**Figure 3a**) requiring phages to evolve countermeasures against these depletion strategies (**Figure 3b**) (*408, 413, 472-474*). For example, *E. coli* phage T4 utilizes a tRNA

ligase to repair the host lysine tRNAs that are cleaved by PrrC. This cleavage occurs upon detecting the interference by the Stp phage T4 protein of the restriction-modification (RM) system of the host (*475*).

An alternative and recently identified strategy involves phages encoding tRNAs that are resistant to the host **tRNA nucleases** (tRNases) (**Figure 3a**) (*37, 39, 323, 408*). This strategy was only recently proposed when investigating differences in phage tRNAs compared to host tRNAs (*323*). These mutations were found to overlap with known resistance-gaining mutations in the cleavage site of a broad range of tRNA nucleases, including VapC, Colicin E5, Colicin D, and PrrC (**Figure 3a**) (*323*). Several studies have subsequently expanded the relevance of these cleavage-resistant phage tRNAs as an evasion mechanism to other phage defense systems that cleave host tRNAs upon detecting the phage infection, including phage defense systems RM-PrrC, PARIS, Retron-Eco7 (type I-A Retron), Cas13, and RemAIN (**Figure 3a**) (*37, 39, 476, 477*). While the impact of phage tRNAs on most of these defense mechanisms is unexplored, their involvement in the evasion of PARIS and Retron-Eco7 is well established (*37, 39, 476, 477*). Although their mechanism of action is distinct from each other. Because PARIS (and RM-PrrC) detect the presence of phage proteins that mimic DNA, including Ocr and Stp, both resulting in the cleavage of $tRNA^{lys}$ (*37, 478, 479*); while tRNA-targeting phage defense system Retron-Eco7 is activated by a phage exonuclease that cleaves the multicopy single-stranded DNA (msDNA) of the Retron-Septu type II complex, causing the complex to fall apart. This releases the HNH-domain containing effector (PtuB), which starts cleaving the host $tRNA^{tyr}$ or $tRNA^{ser}$, depending on the specificity of the system (*39, 476*).

Moreover, CRISPR derived Cas13a complexes, which do not recognize phage proteins, but instead recognize the phage through crRNA matching phage transcripts with one of the spacers from its CRISPR-array, upon which the collateral RNase activity of Cas13a cleaves several host tRNAs including $tRNA^{lys}$, $tRNA^{glu}$, $tRNA^{gln}$, and $tRNA^{thr}$ (**Figure 3a**) (*15, 410*).

Lastly, RemAIN, a prophage encoded-phage defense system, protects the host from lytic activity of other (pro)phages through cleaving several essential host tRNAs, including $tRNA^{leu}$, $tRNA^{pro}$, $tRNA^{fmet}$, and $tRNA^{ser}$, upon detecting an unknown trigger (**Figure 3a**) (*477*).

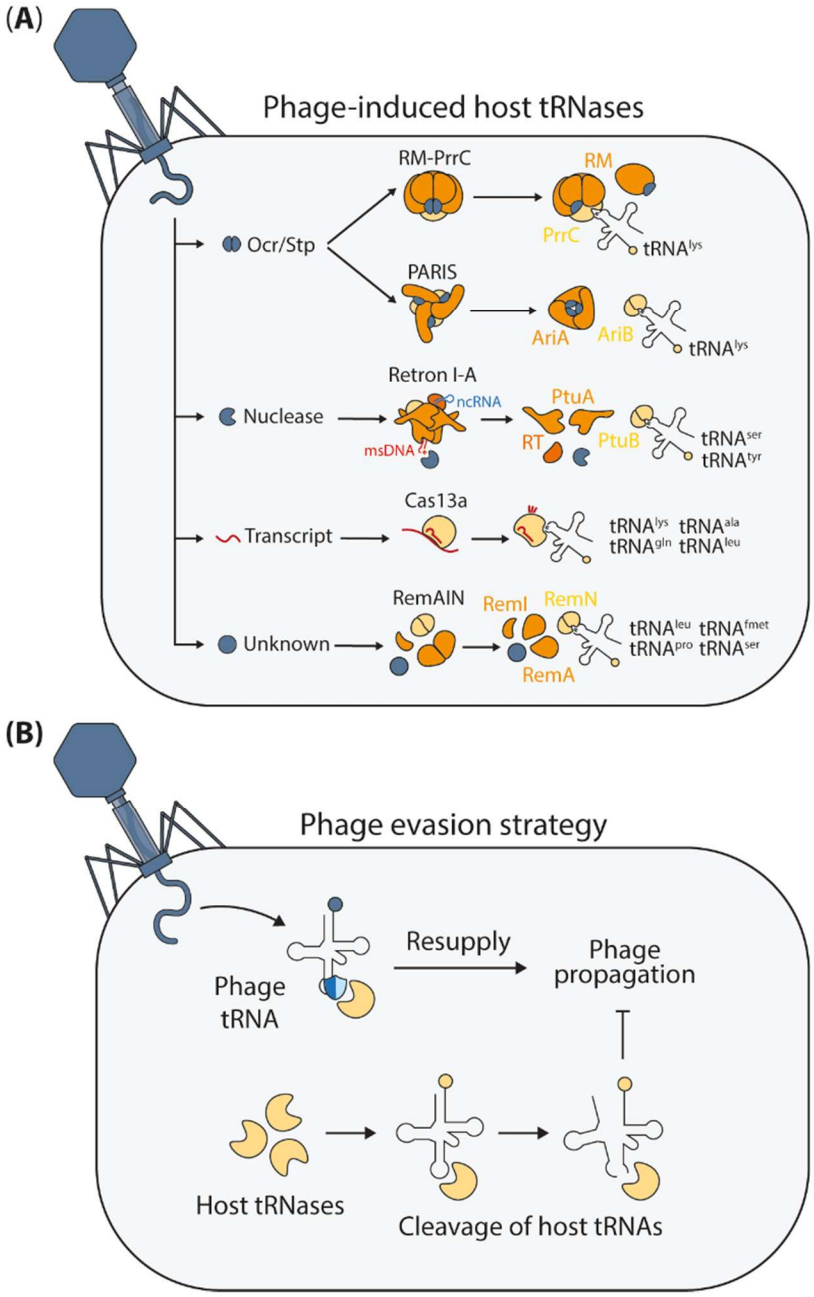

Figure 3: Phage tRNAs replenish the depleted host tRNAs. (a) During phage infection, several phage factors are known to activate the tRNases of the host, such as DNA-mimics Stp and Ocr, which interfere with the defense response of the host, and induce the activity of RM-PrrC and PARIS to cleave the $tRNA^{lys}$ of the host. In addition, Retron-Eco7 (type I-A Retron) is activated by an exonuclease of the phage, resulting the cleavage of the msDNA of the Retron. This causes the Retron-Eco7 complex to disassemble and start cleaving the $tRNA^{tyr}$ or $tRNA^{ser}$ of the host. Moreover, Cas13a is activated by phage transcripts that match its spacer, leading to the cleavage of several host tRNAs, including $tRNA^{lys}$, $tRNA^{gln}$, $tRNA^{ala}$, and $tRNA^{leu}$. Lastly, RemAIN, a prophage encoded-phage defense system, protects the host from lytic activity of other (pro)phages through cleaving several essential host tRNAs, including $tRNA^{leu}$, $tRNA^{pro}$, $tRNA^{fmet}$, and $tRNA^{ser}$, upon detecting an unknown trigger. **(b)** The cleavage of tRNAs by these phage defenses result in the inhibition of translation and phage propagation. In response, phages encode tRNAs, that are resistant to these host tRNases, to replenish the tRNA pool and propagate.

Although the involvement of phage tRNAs in evading each of these phage defense systems has not yet been demonstrated, current evidence strongly suggests their potential neutralize these phage defenses. It is also likely that the number of tRNA-targeting defense systems exceeds the above-mentioned defenses, since the functional protein domain that cleaves tRNAs for each of these phage defense

systems is shared across various phage defense systems. PARIS contains the Old/TOPRIM functional domain, which is responsible for their tRNA cleaving capacity. This domain is also present in at least eleven other phage defense systems, including Menshen and Gabija (*59, 141, 142, 159, 480-483*). Moreover, Retron-Eco7 (type I-A Retron) cleaves tRNAs using its HNH nuclease domain, a commonly shared domain among phage defense systems, including Septu type I, Zorya type II, RM type II systems, and CRISPR-Cas systems (*159, 484, 485*). Additionally, Cas13a, RemAIN, and PrrC share their tRNA cleaving HEPN domain with at least 17 other phage defenses, including CoCoNuTs, AbiF, and Ape (*122, 141, 142, 481, 486-491*). Lastly, VapC shares its PIN domain with phage defenses HEC-03 and HEC-09 (*142, 492*). While these domains are not necessarily always involved in tRNA cleavage and may have different nucleic acid targets, it is likely that at least some of these phage defenses target tRNAs, potentially even the tRNAs of the phage. Promising candidates may be found in the VapC family, as many VapCs have unknown tRNA targets, while extensively investigated for their activity against the tRNAs of the host (*415*).

Proposed Phage tRNA Functions

In addition to the above-mentioned roles and proposed hypotheses regarding the role of tRNAs in phages, several additional novel hypotheses may be proposed. One of these hypotheses is that some phage tRNAs are not resistant to cleavage and instead act as tRNA-derived small RNAs (tsRNAs) and tRNA-derived fragments (tRFs), which are known to modulate the transcription and translation of the host (*493*). Moreover, phage tRNAs might act to stall the ribosomes of the host during infection, preventing a host defense response (*494*), or interfere with other tRNA-binding proteins (*495*). Alternatively, integrated temperate phages might encode tRNAs to increase the available tRNA pool of the host, which may accelerate their replication, resulting in an increased number of integrated phage copies. Lastly, eukaryotic viruses, such as human cytomegalovirus, have been found to use tRNAs as scaffolds to assemble their capsids. It is possible that phages use tRNAs in a similar fashion (*496*).

Taken together, phage tRNAs have diverse functions. In temperate phages, tRNAs may restore the detrimental effects of their integration into the tRNA of the host and prevent the premature expression of their lytic genes. In virulent phages, tRNAs are primarily encoded to benefit the phage during its replication cycle, to counteract the phage defense response of the host and overcome discrepancies in codon usage between phage and host genes. Notably, these lytic-related functions could also apply to the functionality of the tRNAs of temperate phages that have entered the lytic cycle. More speculative functions of phage tRNAs, include their involvement in the **hypermodification** of phage DNA, to provide an integration site for beneficial mobile elements, increase the tRNA pool of the host for improved replication, function as capsid scaffolds, and regulate host translation and transcription through **RNA-interference** and ribosomal stalling.

Layers of tRNA-targeted Phage Defense

Cleavage of host tRNAs is not the only tRNA-targeted layer of phage defense that the host immune response initiates upon phage infection (**Figure 4**). Another significant mechanism involves the toxic small alarmone synthetase (toxSAS) family, including toxSAS and phage defense system CapRel, which pyrophosphorylate the highly conserved CCA-ends in the acceptor stem of tRNAs to inhibit translation and prevent phage propagation (**Figure 4b**) (*74, 497*). Other host mechanisms that act on tRNAs, though not yet linked to phage defense, are part of the nucleotidyltransferase (NTase) family (*498*). The most notable member of the NTase family is the MenT toxin of *Mycobacterium tuberculosis* (*498*). MenT extends the tRNA at the acceptor stem with additional cytosines, preventing the aminoacylation process during maturation and inhibiting translation (**Figure 4b**) (*498*). Other members of the NTase family are less understood but are known to convey phage defense, including SanaTA, AbiE, and AbiG (*59, 499*). Lastly, TacT, a member of the acetyltransferase (GNAT) family, disrupts translation through acetylation of the glycine aminoacyl-tRNAs (**Figure 4b**).

In addition to these toxins, it is conceivable that hosts distinguish self-tRNAs from non-self-tRNAs through several specific modifications to the tRNA during maturation (*442*). To bypass these host defenses, phages may act to become independent of the host and encode their own tRNA-maturing enzymes to steer the phage tRNA pool away from these host regulatory pathways (**Figure 4a**) (*500*). Supporting this hypothesis, a recent study has found that some phages encode not only one, but several steps of tRNA maturation, including their own CCA-adding tRNA NTases, aminoacyl-tRNA synthetases (aaRS), and elongation factors (EF-Tu), as well as tRNA modifying enzymes such as QueC and tRNA-guanine transglycosylase (TGT)-like proteins (*446, 447, 500*), which are known to modify the guanine (g) of tRNAs into 7-cyano-7-deazaguanine ($preQ_0$) and archaeaosine (G^+) through their involvement in the deazaguanine metabolism (**Figure 4a**) (*446, 447*). The exact benefit for phages to modify the guanine (g) into $preQ_0$ or G^+ remains to be speculated. A similar **tRNA modification**, deazaguanine derivative queuosine (Q), is present at the first position of the anticodon and improves the codon-anticodon interaction, making the translation faster and more accurate (*501, 502*). However, unlike Q, $preQ_0$ and G^+ modifications are present in the D-arm and act to improve the stability of the tRNA during stressful conditions, which might be beneficial for the phage during phage infection (*503*). Moreover, $preQ_0$ and G^+ may render phage tRNAs more resistant to the tRNA cleaving enzymes of the host, since the G^+ modification in DNA is known to render DNA resistant to a multitude of endonucleases, including EcoRV, HaeIII, DraI, EcoRI, and MboI (*504*). In addition, $preQ_0$ or G^+ modification might also be a way to distinguish phage tRNAs from host tRNAs (*442*).

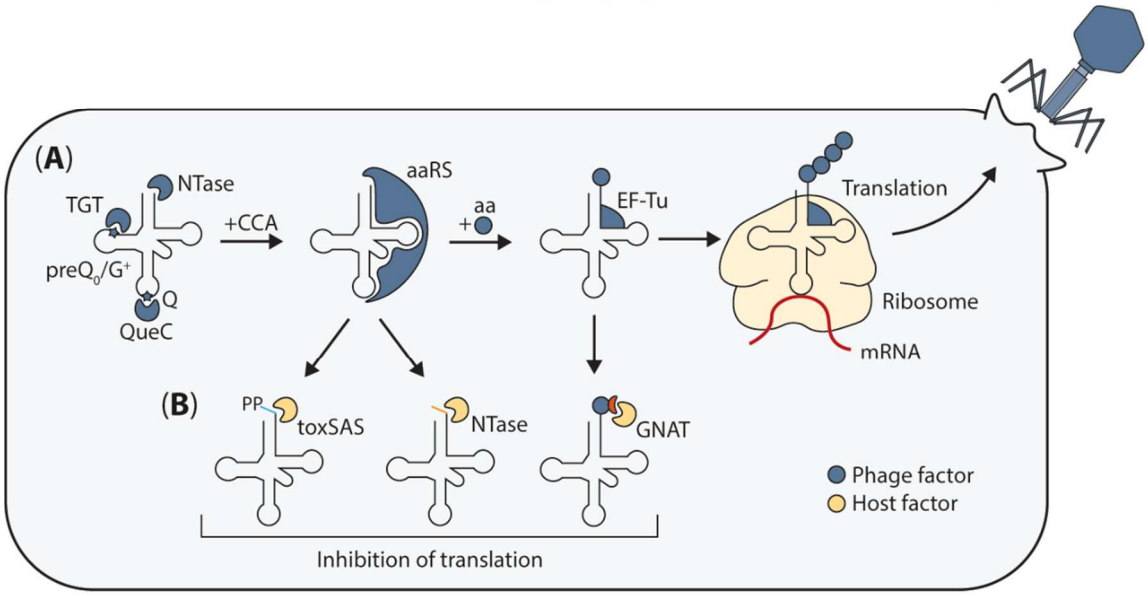

Figure 4. The layers of tRNA-targeted phage defense in the maturation steps of phage tRNA. (a) The stages of phage tRNA maturation are shown, from transcription to tRNA modifications, processing, amino-acetylation, translocation, and translation. Phage factors are shown in blue, while bacterial factors are shown in yellow. **(b)** The proposed host defenses that may interfere with the maturation of the phage tRNA are also illustrated, including toxSAS, NTase, and GNAT. preQ0, 7-cyano-7-deazaguanine; G^+, archaeaosine; TGT, tRNA-guanine transglycosylase; NTase, nucleotidyltransferase; Q, queuosine; aaRS, aminoacyl-tRNA synthetase; aa, amino-acid; EF-Tu, Elongation factor thermo unstable; toxSAS, toxic small alarmone synthetase; GNAT, Gcn5-related N-acetyltransferases.

These complex dynamics highlight the arms race between the host and phage, with numerous of unresolved questions to be explored.

Concluding Remarks and Future Perspectives

Although phage tRNAs have intrigued scientists since their discovery in the 1960s, it is only recently that researchers have begun to understand their complexity, uncovering an array of unexpected roles for phage tRNAs during phage infection. These diverse roles make predicting the function of individual

phage tRNAs challenging and are likely affected by multiple factors. However, certain overarching functions appear to be conserved, such as virulent phages are known to encode a multitude of tRNAs to overcome the defenses of the host and improve translation efficiency. For temperate phages, tRNAs are likely encoded to prevent a detrimental effect of integrating into the tRNA of the host, as well as controlling the host response during the lytic and lysogenic stages. These novel insights provide promising potential for phage therapeutic approaches. For instance, natural or genetically engineered phages, that encode additional phage tRNAs, could significantly enhance the efficiency of the phage infection to overcome phage resistant bacteria that utilize tRNA-depleting defense strategies. Furthermore, engineering phages might also be a way to expand their host range by overcoming codon usage discrepancies between the host and phage. A comprehensive understanding of phage tRNAs will be crucial for these applications. Intriguingly, several unexplored hypotheses remain that may be validated in the years to come (see Outstanding Questions). These questions might be answered with a recent burst of novel methods to investigate tRNAs *in vivo*, such as the improved detection of tRNA modifications and high-throughput sequencing of tRNAs (*505-507*). Understanding the full complexity and role of phage tRNAs holds the potential to benefit phage-based therapeutics through empowering phages to overcome tRNA-targeting phage defenses (*508*).

Outstanding Questions

1. Are there phage defense systems that specifically target phage tRNAs?
2. Do phage tRNAs function as mobile element integration hotspots?
3. Which roles do the other phage encoded non-coding RNAs have, such as transfer-messenger RNAs?
4. Can we extrapolate the roles of phage tRNAs into tRNAs that are encoded by human viruses?
5. Can we improve phage-therapeutics with phage tRNAs?

Chapter 10



Summary and discussion

Bacteria and phages have engaged in an evolutionary dance since the dawn of life. Only recently have we begun to understand the depth of this dance, as a growing number of phage defenses and countermeasures continue to be uncovered. Yet, it remains largely unclear what the specific roles and impact of each of these mechanisms is, and whether we can unearth additional mechanisms. This dissertation has advanced our understanding of this phage-host conflict, finding novel insights into both phage defense systems and how phages circumvent them. Here, we discuss these advances and put these in the context of the discoveries that were made before, while also reflecting on the current paradigms in the field and its future.

Introduction

Since their emergence, bacteria have been under predation from their viruses, known as bacteriophages (phages) (*3*). To defend themselves, immune mechanisms have evolved in bacteria that recognize and act on phage associated molecular patterns (*10*). In response, phages have evolved countermeasures against each of these mechanisms, prompting bacteria to adapt, resulting in a complex conflict between the phage and host (*10*). Over the last decade, the field has made tremendous progress in our understanding of this conflict (*10*). In this discussion, we place these discoveries in the context of the findings presented in this thesis and discuss the current and future direction of the field.

Insights into phage immunity and evasion strategies

The accumulation of phage defense systems correlates with a broader phage resistance

The complexity of the bacterial immune system is exemplified by the sheer number of distinct mechanisms that have been discovered so far, from preventing adsorption to degrading host factors upon detecting phage infection (*11-28*). Bacteria almost always encode several phage defenses that have been accumulated over time (**Figure 1**) (*63*). However, the impact of encoding multiple defense systems remained unclear. Although long hypothesized, we demonstrate that the greater the number of phage defense systems the bacterium encodes, the broader its resistance to a collection of phages (**Chapter 2**) (*146*). Later studies from the Filippovs group confirmed our findings when also investigating the phage defense resistance of *Pseudomonas aeruginosa* (*509*).

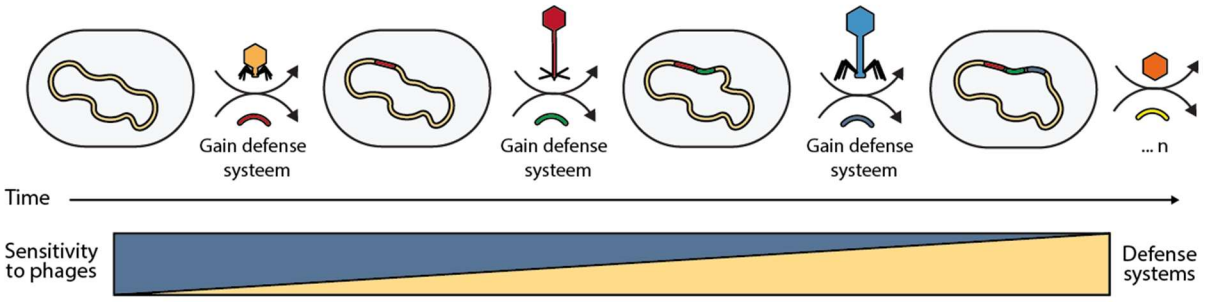

Figure 1: Accumulation of phage defense systems correlates with phage resistance. Over time, phage defense systems are accumulated by the bacterial strain to protect itself from phage predation. This increase in the number of encoded phage defense systems renders the strain less sensitive to phages.

Other studies on other genera and species, including *Vibrio* and *Escherichia* (*49, 510*), have also found that phage defense systems are indeed a significant factor for phage resistance. However, they also emphasize that adsorption factors are the large contributor to phage resistance, especially when phages infect bacterial strains that are more distantly related from their original host strain.

Why phage defense systems seem to be a larger factor in *P. aeruginosa* compared to *Vibrio* and *Escherichia* may relate to the broad adsorption range of most *Pseudomonas* phages, while phages from *Vibrio* and *Escherichia* seem to be more often adsorbing to a smaller fraction of strains (*49, 510*). One explanation could be that *Pseudomonas* only has 20 LPS/O-antigen variants, one of the most common phage receptors, while *Vibrio* and *Escherichia* have more than 180 variants (*511-514*). I hypothesize that this may give *Vibrio* and *Escherichia* species more ways to remodel their outer membrane to the environmental phages to prevent their adsorption, instead of relying on intracellular phage defenses. Future studies could further explore the possible relationship between the variability of the outer membrane phage receptors and the prevalence of intracellular phage defenses.

Mapping the phage defense repertoire of P. aeruginosa

To obtain a more comprehensive understanding of the effect of phage defense systems on the host range of *Pseudomonas* phages, we set out to map the complete phage defense repertoire of our *P.*

aeruginosa bacterial collection (**Chapters 3 to 5**) (*138, 515*). To date, several approaches have been used for such purposes, including both bioinformatic and wet-lab methods (**Figure 2**). These approaches can be further categorized in two main types: gain of function assays and loss of function approaches.

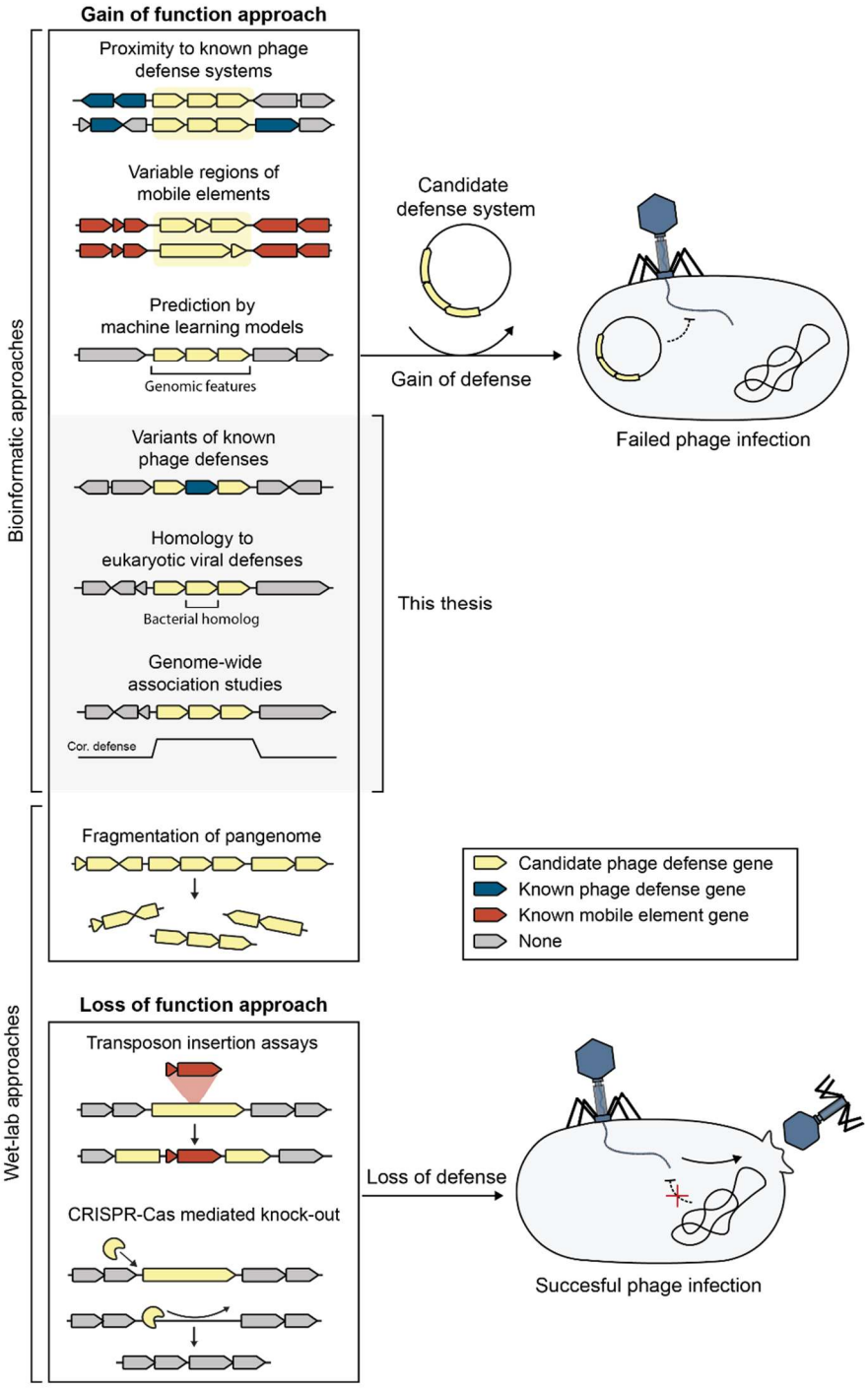

Figure 2. Overview of approaches to find novel phage defenses. A multitude of approaches have been applied to date. These can be categorized into both bioinformatic and wet-lab approaches, as well as gain of function and loss of function assays. Candidate phage defense systems are identified through several ways, including proximity to known phage defense systems, variable regions of mobile elements, machine learning predictions based on genomic features, variants of known phage defenses, homology to eukaryotic viral-defenses, genome-wide association studies, and fragmentation of the pangenome. In gain of function assays, candidate phage defense systems are cloned and transformed into a host bacterium, followed by phage resistance assays. In loss of function assays, bacterial genes are disrupted or removed by using transposons or CRISPR-Cas directed DNA degrading enzymes.

In the loss of function approaches, bacterial genes are inactivated, followed by subjecting these strains to a collection of phages to check for altered phage resistance. Thus far, two of such approaches have been applied, including transposon insertion assays and CRISPR-Cas mediated knockouts (*144, 516*). During transposon insertion assays, transposons integrate at pseudorandom positions in the genome, causing the disruption of bacterial genes (*517*). A loss of phage defense provides evidence that this disrupted gene is linked to phage resistance (*144*). Similarly, CRISPR-Cas mediated knockouts, i.e. CRISPR-Cas3 and CRISPR-Cas9, of which CRISPR-Cas3 disrupts bacterial genes by searching for the position in the genome, for which a spacer was provided, followed by the genomic degradation from that position onwards (*516, 518*). Loss of phage defense leads to follow-up analyses to provide evidence that these genes are indeed linked to phage resistance.

In the gain of function approaches, candidate phage defense systems are transplanted from their original strain into another bacterial host that is used to test whether the candidate phage defense system provides phage defense, such as *P. aeruginosa* or *Escherichia coli* (**Figure 2**). This is typically achieved by using of a cloning plasmid, that is modified to encode the candidate phage defense system. When introduced into the new host, the plasmid either persists or integrates the candidate phage defense system into the genome of its new host. To date, the most fruitful approach to successfully identify candidate phage defense has been through guilt-by-association studies which leverages that genes with similar functions often co-localize within the bacterial genome (*121, 122, 142, 159, 160*). Other approaches include mining variable genetic regions of mobile elements, such as prophages and integrons, the fragmentation of the bacterial genome, and most recently the development of machine learning algorithms that search for novel phage defenses based on the genomic and protein features of the known phage defense repertoire (*62, 139, 141, 519, 520*) (**Figure 2**).

In this thesis, we provide three additional methods for identifying phage defense systems (**Chapters 3 to 5**). First, we search for bacterial genes that, when present in the clinical strains, correlate strongly with the resistance to one of our *Pseudomonas* phages (**Chapter 3**). This method provided us with one strong candidate: pyocins. Pyocins are small phage-like particles that consist of only the tail of a phage and are believed to originate from an ancient phage that has been domesticated by the bacterium to combat competing bacteria through pyocin-mediated lysis (*149, 153*). Our approach suggests that pyocins can also act in the bacterium-phage conflict, where the release of pyocins can lyse during early stages of phage infection, preventing formation of mature phage particles. In a way, ancient phage components are used against phages that are around today. It should be noted that further investigation is required to support the involvement of pyocins in phage defense.

Secondly, we identified phage defenses by analysing conserved gene clusters that contained components of known phage defenses in new gene contexts (**Chapter 4**). This approach to find new phage defense systems capitalizes on the observation that phage defense systems in bacteria exhibit high degrees of modularity, with sensing, signal transmission, and effector enzymes frequently being exchanged among phage defense gene clusters. An approach that had already been successfully applied to the earliest discovered phage defense systems, including RM and CRISPR-Cas (*484, 521, 522*). Here, we apply the same methodology additionally using more recently uncovered phage defenses, leading to the discovery and functional characterization of three novel phage defense systems. These phage defense systems are RM systems, TerB-family proteins, and Radical SAMs. Beyond these three phage defense systems, we also uncovered over 500 candidate phage defense systems, a much larger number than the currently known phage defense repertoire, providing a broad indication of the putative phage defense repertoire of *P. aeruginosa*. We observed that of these candidate phage defense systems, restriction enzymes, methylases, and ATPases are the most frequently exchanged phage defense genes. Earlier work has also identified restriction and methylation enzymes to be commonly exchanged among phage defense systems such as RM, DISARM, and Druantia (*159, 241, 523, 524*). The exchange of these RM components likely provides the bacterial immune system with additional options for protecting against additional phages. Moreover, this heterogeneity could also be a way to limit a successful phage from infecting neighbouring bacteria in the colony (*523*). Since this phage inherits the epigenetic

markers of the previous host, rendering the RM system of the neighbouring host useless in instances where the RM systems are identical. The exchange of the restriction and methylation genes could be a way to counteract this evasion strategy and maintain phage defense at a population level (*523*). Similar to restriction and methylation enzymes, ATPases have also been described in a multitude of phage defenses, such as Gabija, Septu, PrrC, and PARIS (*62, 122, 142, 159, 525*). Of these shared ATPases, we found the ATPase of the PARIS systems to be one of the most shared genes. In PARIS, this ATPase is responsible for detecting the presence of specific phage-proteins (*37, 479*). This observation suggests that ATPases may be exchanged among phage defense systems to gain or lose sensitivity to a specific phage protein, adding an additional layer to the phage defense repertoire. Notably, the co-operative nature of ATPase does not only seem to occur at the genomic level, but also at the functional protein level, since a recent study by Wu *et al.* (2024) showed that the phage sensor ATPase of Gabija can operate as a substitute sensory ATPase for phage defense system Tmn, which then results in activating the downstream effector of the Tmn-mediated phage defense (*526*). Overall, we show that searching for conserved gene clusters that contain components of known phage defense systems is an effective approach for uncovering the full scope of the phage defense system repertoire, as well as providing new insights into the defense system repertoire dynamics.

Lastly, we leveraged the similarity of several phage defenses with viral defenses from eukaryotes, to search for additional eukaryotic viral defense-like genes in *P. aeruginosa* to uncover novel phage defenses (**Chapter 5**) (*138*). Previous studies have uncovered several phage defenses with strong similarities to the viral defenses of eukaryotes, including bacterial Viperin, caspase recruitment domains (CARD), Gasdermin, Sirtuins, prokaryotic Argonautes (pAgos), Toll-interleukin-1 receptors (TIR), and STANDs (*21, 163, 238, 301, 527-529*). For this reason, we created a catalogue of protein signatures of known viral defense genes from eukaryotes. Applying these protein signatures to the pangenome of *P. aeruginosa*, we uncovered more than 450 candidate gene clusters that may provide phage protection (*138*). We provide evidence that at least four of these candidates convey phage defense, including IMPAse Hermes, DNA2-containing Prometheus, TIR-containing Thoeris III, and ubiquitin-like containing 6A-MBL. This provides additional insight into the evolutionary link between prokaryotic and eukaryotic viral defenses. This link is reinforced by a separate study that searched for eukaryotic viral defense genes through homology to known bacterial phage defense systems and discovered seven additional anti-viral genes in the human genome, including MOV10L1, PLD6, GIMAPs, FHAD1, EFHD2, and CTRC (*530*). Each of these six genes was found to be homologous to at least one out of three bacterial phage defenses: Mokosh, Eleos, and Lamassu. To explore the origin of these shared defenses, other groups turned to Asgard Archaea, the closest living prokaryotic relative to eukaryotes (*531, 532*). Their findings reveal that Asgard Archaea harbor a multitude of eukaryotic-like viral defenses, including Mokosh, Lamassu, Viperin, Argonaute, CBASS, and NLRs. These results support the hypothesis that the eukaryotic viral defense repertoire originates from Asgard Archaea. It is not unthinkable that besides the contribution of Asgard Archaea, the bacterial phage defense repertoire also contributed to the complex viral defenses of eukaryotes today.

Overall, this thesis provides several additional approaches for discovering phage defense systems to improve our broad understanding of the phage defense repertoire of bacteria.

Fewer phage defense systems in strains isolated from the lungs of cystic fibrosis patients

Besides our curiosity in mapping the complete phage defense repertoire of *P. aeruginosa*, we were curious about its relevance in phage therapeutics. Previous studies aimed at development of therapeutic applications for treating *P. aeruginosa* infections with phages (*533*). While treating superficial bacterial infections showed promising results, these were yet to be reported for treating infections residing in the lungs of cystic fibrosis patients (*351, 533, 534*). To our surprise, we found that the strains that cause these infections have significantly fewer phage defense systems than their counterparts in other lung conditions, which should in theory render these strains more sensitive to phages (**Chapter 6**) (*535*). We postulate that this decrease in the number of phage defenses hints at the difficulty for phages

to reach these strains in the lungs of cystic fibrosis patients, likely because of the thicker dehydrated mucus and/or the reduced air flow, which lowers the need for the additional phage defense systems (*346*). The question remains whether these strains are selected for, or whether these strains lose irrelevant phage defense systems over time (**Figure 3**). Either way, the discrepancy highlights that the environments can greatly impact the phage defense repertoire, as well as providing a promising perspective for phage-based therapies for these CF patients, as was recently demonstrated by the first randomized human phage-therapy trial that successfully treated chronic *P. aeruginosa* infections in CF patients (*536*). Our findings suggests that phage treatments could be combined in the future with mucus-clearing approaches to potentially achieve even greater efficacy.

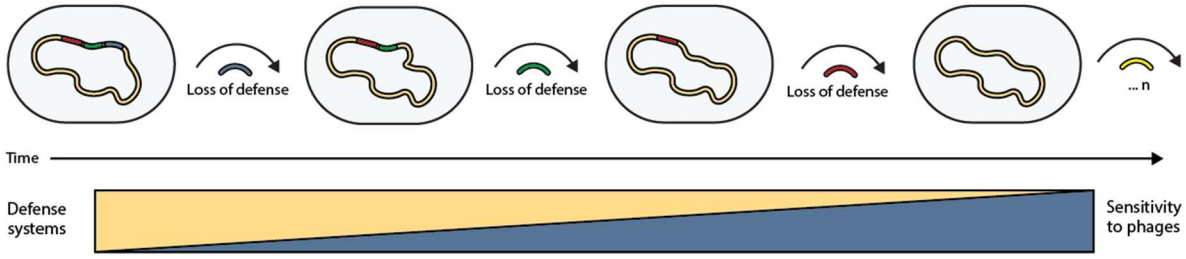

Figure 3. Loss of phage defense systems in the lungs of cystic fibrosis patients. Over time, the *P. aeruginosa* strains residing within the lungs of cystic fibrosis patients likely lose their phage defense systems, resulting encoding far fewer phage defense systems than their counterparts in other lung conditions. This is likely the result of the reduced phage predation that occurs in these lungs, due to the reduced air flow and thicker dehydrated mucus, preventing the phage from reaching the strain.

Phage evasion strategies of Pseudomonas phages

Beyond our investigation into the application of phages to combat *P. aeruginosa* strains that infect cystic fibrosis patients, we were also interested in improving and understanding how these phages infect these strains. For this reason, we set out to dissect the mechanisms that phages utilize to circumvent the bacterial immune system (**Chapter 7 to 9**).

To date, a number of studies have already set out to uncover a multitude of these mechanisms through a combination of wet-lab and bioinformatic approaches (*43, 281, 371, 537-540*) (**Figure 4**). Several of these approaches are like those applied to finding phage defense systems. For example, instead of searching for co-localizing phage defense systems, these studies search for co-localizing anti-phage defense genes within phage genomes and predicting anti-phage defense activity through machine learning algorithms (*43, 371, 540*). Other approaches are distinct. For example, one study observed defense neutralizing genes to be encoded in the leading region of plasmids. These genes are likely to be the first ones to be translated when entering the host to counteract its immune response (*391*). They used this gained knowledge to identify novel anti-defense genes in this region of other plasmids. Recently, efforts have also been made to identify phage-defense neutralizing proteins through computationally predicting a physical interaction between all phage genes and phage defense gene, after which predicted strong interacting proteins were tested for their neutralizing activity (*537*). It is not unlikely that a similar approach that leverages protein hallucinations will produce artificial proteins that able to neutralize a wide range of phage defense systems (*541*).

Besides these above-mentioned bioinformatic approaches, several wet-lab approaches have been explored to identify phage evasion strategies. This includes loss of function approaches, such as synthetic antisense oligomers (ASOs), which inhibit specific phage transcripts to investigate their role in phage defense evasion, as well as CRISPR-Cas13-mediated counter selection for phages that contain a gene of interest (*542-544*). This CRISPR-Cas13 is loaded with a spacer against the gene of interest, if present, CRISPR-Cas13 becomes active, causing cell death. Alternatively, smaller phages can be produced using PCR and ligated together using Gibson assembly and electroporating the assembled phage genome in an *E. coli* competent strain, such as HST08 (*545*). By deliberately omitting PCR products of the region of interest, the impact of these regions on the ability of the phage to overcome

the host phage defenses can be assessed. Both approaches strife to show the effect of the absence of these genes on the ability of the phage to overcome the immune system of the host.

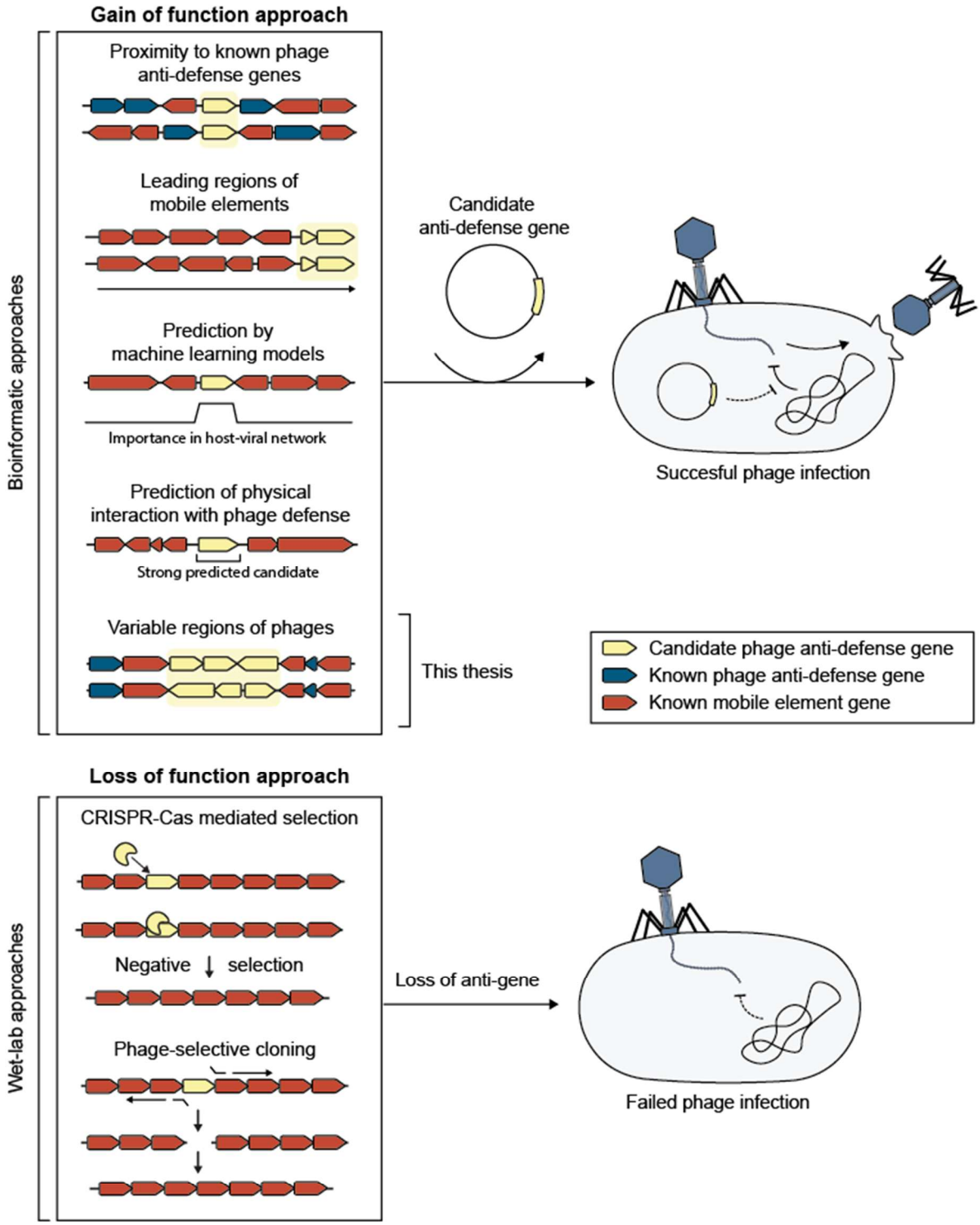

Figure 4. Overview of approaches to find novel anti-phage defense genes. A multitude of approaches have been applied to date. These can be categorized into both bioinformatic and wet-lab approaches, as well as gain of function and loss of function assays. In loss of function assays, phage genes are negatively selected by CRISPR-Cas mediated selection, or selectively not synthesized during phage-cloning. In gain of function assays, candidate anti-phage defense genes are cloned and transformed into a host bacterium with a certain phage defense system, followed by phage resistance assays. Candidate anti-phage defense genes are identified through several ways, including proximity to known anti-phage defense genes, leading regions of mobile elements, machine learning predictions based on host-viral network importance, co-folding predictions of phage genes to phage defense genes, and presence in highly variable regions of the phage genome.

In this thesis, we set out to uncover more anti-phage defense genes by looking in highly variable regions of phage genomes. We suspected that, like the defense systems of the host, phages encode their counter mechanisms in highly variable regions of their genome (**Chapter 7**). To investigate this proposition, we identified highly variable phage genes within our set of *Pseudomonas*-infecting *Pbunaviruses*. From the 43 identified candidate anti-phage defense genes, we validated five to be functional. Two of these anti-phage defense genes, bdi1 and bdi2, turned out to have a broad neutralizing activity against our set of phage defense systems, including Druantia type I, Hypnos, RADAR, and Zorya type I. Others, such as ZadI-1, TadIII-1, and DadIII-1, appear to specifically act against Zorya type I, Thoeris type III and Druantia type III, respectively. How these phage genes convey their neutralizing activity remains an open question. Previous studies have found that anti-phage defense genes can act in a variety of ways, including physical blocking, degradation or sequestering of the signalling molecule, preventing or reverting the effector activity, and inactivating phage defense genes (*34, 39, 86, 546-551*).

Phages encode tRNAs to resupply the depleted host tRNA pool during phage infection

Beyond our search for anti-phage defense genes in our *Pseudomonas* phage collection, we were also curious about the possibility that phage tRNAs may act as anti-phage defense genes. Phage tRNAs were discovered not long after the discovery of tRNAs themselves, however, its function remained an enigma until recently (*325*). We provided a novel insight into this enigma when we compared the tRNA sequence of the phage to the tRNA sequence of its host and observed conserved mutations in the phage tRNA that render them resistant to phage defenses that target tRNAs (**Figure 5**, **Chapter 8**) (*323*). Later studies confirmed our suspicion by providing evidence that phage tRNAs circumvent phage defenses PARIS and Retron type I-A (*479, 552*). We summarized these breakthroughs in a review on phage tRNAs, where we also discuss its role in the infection cycle of temperate phages (**Chapter 9**) (*325*). For temperate phages, tRNAs may act to also benefit the phage during its lysogenic cycle, especially for phages that integrate into the tRNA of the host. In these instances, phage tRNAs are utilized to compensate the detrimental effects of disrupting the host tRNA (**Figure 5**) (*325*).

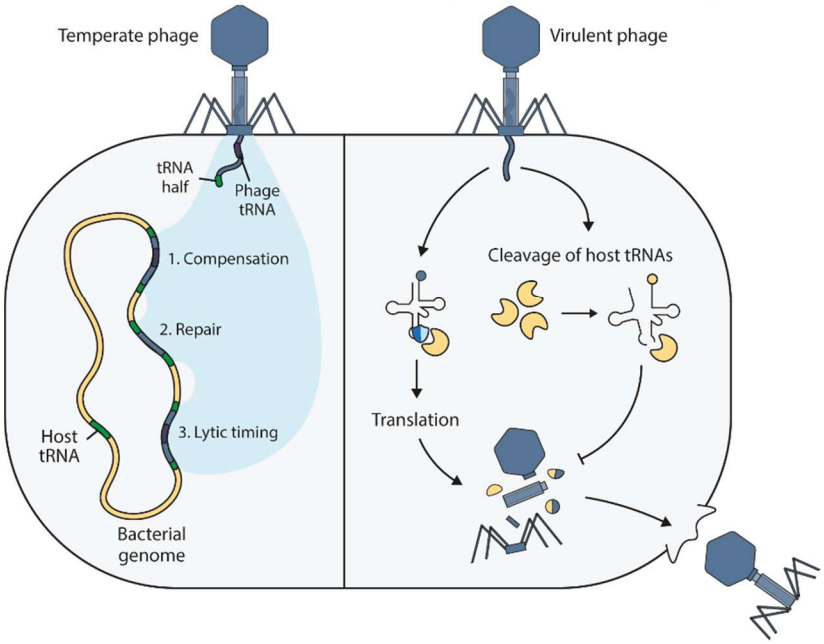

Figure 5. Overview of the roles of phage tRNAs in the lysogenic and lytic cycle. Phage tRNAs may act to compensate for the detrimental effect of phages that integrate into tRNAs. In these instances, the tRNA compensates for the same tRNA that is disrupted or repairs the tRNA. Some phages also encode suppressor tRNAs that regulate the timing of the lytic cycle. In the lytic cycle, phage tRNAs serve different functions, most notably to replenish the depleted tRNA pool of the host, as well as providing tRNAs that are rarely produced by the host, but the phage utilizes extensively.

Together, this thesis has contributed to obtaining deeper understanding of the phage-host conflict. We showed that *P. aeruginosa* strains can encode upwards of 20 distinct phage defense systems and their accumulation leads to a broader phage defense phenotype (**Chapter 2**). We developed several additional approaches to further uncover the expanding phage defense repertoire (**Chapter 3 to 5**). In addition, we show that the environment of the host affects its phage defense repertoire (**Chapter 6**). On the side of phages, we find that phages encode genes in their highly variable genomic regions that neutralize the immune system of the host (**Chapter 7**). We also shed a new light on a 50-year-old mystery on why phages encode their own tRNAs (**Chapter 8 and 9**). Although, this thesis expands our understanding of phage defense systems, it also shows that our comprehensive understanding of the host-phage conflict is incomplete, with many unanswered questions remaining.

The Quest(ions): Searching for Clarity

The broader context

First and foremost, do the insights described in this thesis hold when viewed through the lens of the natural environment? In laboratory settings, we often focus on one phage infecting one strain, a necessary simplification. However, this simplification overlooks several important factors that likely affect the host/phage interaction in the natural environment, including nutrient levels, environmental changes, time, and presence of other phages and bacteria, resulting in the likely under- and overestimating the features of the phage-host interaction (*553*). This is well exemplified by the many environmental factors that affect the decision of temperate phages to enter the lytic or lysogenic life cycle, such as salinity, temperature, UV, host density, and the presence of other phages (*554*). Similarly, the immune system of the host is also affected by external factors. For example, CRISPR-Cas mediated resistance is more prominent in low nutrient conditions compared to high nutrient conditions (*555*). Let alone the effect of other organisms within the environment, for instance the production of antibiotic products (e.g. aminoglycosides) and secondary metabolites (e.g. coelichelin), which modulate the phage sensitivity of nearby bacteria (*556, 557*). Last but not least, lab settings typically do not last longer than a few days, while the phage-host interaction in natural conditions likely unfolds over a longer timescale (*558*).

All included several components of the phage-host interactome are likely poorly reflected in lab conditions. This does not invalidate results obtained in laboratory conditions, but it does suggest that the relevance of certain factors in the phage-host interaction might differ between the lab and natural setting. Therefore, adopting a broader ecological perspective could be fruitful for fully understanding the phage-host interaction.

Constraints on the phage defense repertoire

Beyond the broader natural perspective, there are fundamental conceptual questions that have yet remained unresolved. For example, bacterial strains only encode a small fraction of the phage defense systems that are present within its species. This suggests that there is a maximum number of phage defense systems that a single strain can encode, especially when considering that every strain would strive to encode as many phage defense systems as possible to protect itself against external threats. However, we observe a wide range in the number of encoded phage defense systems per strain (**Chapter 2**) (*146*). In the natural perspective, this could be explained that bacteria in environments with a limited variety of phages, encode fewer necessary phage defense systems, and those in high phage diverse settings encoding a multitude of phage defenses. In line, this phenomenon is observed when looking across a variety of metagenomic samples (*559*). However, it does not explain why strains do not always encode as many phage defense systems as possible. A possible explanation is that bacterial strains do indeed encode as many phage defense systems as possible, while also striving to encode as little genes as possible for efficient replication to outgrow other competing bacterial strains from the same colony: the minimal phage defense conjecture. Driven by stochastic processes, these

strains lose as many phage defense systems as possible, while the overarching colony still encoding each of the individual defense systems that the colony started out with. During the next phage outbreak, these divergent strains within the colony exchange their phage defense repertoire to give birth to a new phage resistant hybrid strain. Over time, creating a multitude of closely related strains with different arrangements of phage defense systems that each provide enough phage defense to outside threats, while reproducing as efficiently as much as possible. This explanation is in line with the observation that closely related strains can encode a vast variety of phage defense system repertoire architectures (**Chapter 2**) (*146*).

Together, these observations suggests that the phage defense repertoire is thus not solely shaped by the accumulation of phage defense, but also by being able to rapidly lose these phage defense systems upon (temporal) redundancy or high energetic costs. This postulation hints at a possible reason for why phage defense systems are so commonly found on mobile elements and rarely encoded in core-regions of the bacterial genome (*146*). Beyond the growth rate and energy-efficiency benefits as major drivers for the minimal phage defense conjecture, several other factors might also be relevant. For instance, it can be postulated that the more phage defense systems a strain encodes, the fewer external sources of genetic material it is able to accept, and thus limiting its ability to genetically adapt to environmental changes compared to those with fewer phage defense systems. As well as, preventing other detrimental effects of encoding a multitude of phage defense systems, such as the increased chance of inter-phage defense conflicts and autoimmune responses. Several example of inter-phage defense conflicts include instances where the restriction and modification systems prevent the transcription of another or acts on the modification of the methylates of the other system (*97*). It is thus not unlikely that other instances of conflicting phage defense systems exist, especially those acting on the same cellular pathways. Regarding the increased chance of autoimmunity, the hypothesis is that the more self-acting phage defense systems a strain encodes, the more chance of being triggered in non-phage conditions, leading to cell death. As is seen for CRISPR-Cas systems which often seem to incorporate self-targeting spacers that are toxic to the host (*560, 561*). However, for other phage defense systems it is unclear whether these self-inflicting fitness costs are the large driver in limiting the number of encoded phage defenses. Thus far, one preliminary study has been conducted and observed that overexpressing self-acting phage defense systems, i.e. Septu and Gabija, can cause growth arrest under sub-optimal growth conditions. The activation of these systems is likely caused by the ATP depletion that occurs in these conditions. In contrast, most studies show that the activity of self-acting phage defense systems is highly regulated and only present upon detecting a specific phage-associated molecular pattern (*21, 22, 32, 74, 479*). It is likely that even more regulative processes are present to prevent the self-harming activity of these phage defense systems during non-phage infection conditions, as is seen for eukaryotic viral immune systems (*562*).

Overall, several factors seem to drive bacterial strains to encode a minimalistic phage defense repertoire, including energy-efficiency, growth rate benefits, genomic adaptability, avoiding inter-phage defense conflicts, and auto-immune fitness costs. All while remaining resistant (enough) to the next phage outbreak at a colony level.

On the rarity of phage defense systems

While some phage defense systems are commonly encoded, such as RM and CRISPR-Cas systems, others are rarer (*146*). To some extent, this rarity could be explained by their relevance being limited to a handful of niches. Yet, some phage defense systems, such as Ambrosia, which is only present once in all sequenced *P. aeruginosa* strains, are so rare that it is hard to believe that such narrow and infrequent niches exist. This raises several questions, such as: Could these rare phage defense systems be remnants of niches that used to be more prevalent? Or are they recent adapting to novel niches? Alternatively, it is also possible that these rare phage defenses are so rare because they are not native to the species in which we found them but instead are sometimes acquired from genetic interactions

with other organisms in the environment that more frequently encode these phage defenses (*563*). Many fascinating possibilities remain to be explored. Determining the origin of these rare phage defense systems could offer new insights into the dynamics of the phage defense repertoire.

David and Goliath

Thus far, we discussed various unanswered questions regarding the phage-host conflict from the host side. However, just as many questions remain on the phage side of the conflict. The main question is how such comparatively small viruses can overcome the defenses of their giant hosts. Even the smallest of viruses, such as *Pseudomonas* phage PP7 which only comprises four genes, can infect their bacterial host which encodes up to 6000 genes (*564, 565*). These ultra-small viruses possibly evade the bacterial host immune system through their simplicity or by replicating faster than the response time of the host. However, this hypothesis does not seem to hold for larger phages that are several magnitudes larger, since, unlike ultra small viruses, larger phages evidently encode genes that are linked to counteracting the host immune system. This shows that the host can convey immunity to these phages when the chance is there (**This thesis**). Each defense mechanism is counteracted by an additional phage evasion strategy, which drives the phage to encode more genes to evade all these phage defense systems. This ultimately drives the phage to encode additional genes until its genomic limit is reached, which is determined by the capacity of its genome-encapsulating capsid (*40*).

In response, phages can overcome this genomic limit by making their genes multifunctional, requiring less genes for the same outcome, or to increase their genomic limit to encode more genes. We observed the multifunctionality of phage genes when investigating the anti-phage genes in our set of *Pseudomonas* phages, where we find that several of these anti-phage genes show a neutralizing activity against a broad range of phage defenses systems (**Chapter 7**). On the other hand, there is also evidence that phages increase their genomic limit by increasing the genomic capacity of their capsid. For instance, the jumbo phage morphology is believed to have evolved from the smaller myophage morphology, increasing its maximum genomic capacity from around 150kb to 500kb (*566*). Besides increasing its genomic capacity, phages may have also expanded their genome by teaming up with other (helper) phages. This is well exemplified by several instances where researchers have found phages to be accompanied by helper satellite phages, which encode additional genes to help the phage infect its host (*567*).

Phages may also overcome their genomic limit by becoming genetically highly variable, both through mutations as well as gene-composition, to evolve faster than the host is able to adapt to. Furthermore, phages may have evolved their lytic and lysogenic lifestyles to evade the host immune response. For instance, the phage lytic cycle might have evolved to outpace the host immune response, while the lysogenic cycle evolved to allow the phage to insert into the genome, and wait until a suitable moment, for example when the host immune response is hampered, before replicating.

But even when considering all these phage evasion strategies, it seems only a matter of time before the bacterial host has found ways to counteract each of these strategies. So, why have phages not gone extinct yet? One possibility is that phages are indeed somehow able to outpace the host immune response indefinitely (Red Queen hypothesis). Alternatively, the existence of phages hints at a deeper co-dependency between phages and their hosts. In this co-dependency, bacteria cannot drive phages to extinction without negative consequences, whereas phages cannot drive bacteria extinct without going extinct themselves. In a way, it seems that bacteria and phages are in a not-so voluntary co-dependent symbiosis (mutualism).

On the origin of phage defenses

Over the eons, the viral-host conflict has evolved into an extraordinarily complex interaction, with roots tracing back to the emergence of life itself during the RNA world, when life is thought to have been characterized by self-replicating ribonucleic acid (RNA) entities (< 200 nucleotides)(*568*). A viral-

host-like phenomenon might already have existed, where the constraints of thermodynamics and entropy make replication inherently error-prone, producing non-functional copies that utilize resources without contributing to them (*568, 569*). It could be speculated that selective pressures may have driven the existence of ribozymes (RNA enzymes) that cleared these RNA entities, giving rise to one of the earliest forms of viral defense. Over time, as life became more complex, so did its viruses. Requiring the host to rapidly evolve additional phage defense systems, favoring phage defenses with a modular nature to readily create novel combinations (**Chapter 4**). The origin of what makes up these phage defense systems is quite mysterious, but it may have involved the repurposing of phage components, such as phage tails to bind phage particles, or bacterial outer membrane proteins, that the phage uses to bind to the host, to detect phage tail production from within the host cytoplasm. Moreover, a portion of phage defense systems were likely also adopted from the ongoing viral-viral conflict, in which competing phages encode defenses against each other (*8*). For instance, by domesticating defense system-encoding prophages, ultimately streamlining these prophage-remnants to retain only its beneficial genes, including the phage defense system itself. This process could explain why so many known phage defense systems are found on small mobile elements (**Chapter 1**).

In summary, I hypothesize that the origin of phage defense systems coincided with the emergence of life, or perhaps even earlier. Over time, phage defense systems drew inspiration from diverse sources, including the repurposing of phage-interacting components.

Pause for reflection

It begs the question whether we should see the phage-host conflict so deterministically as the current paradigm of the field suggests. It calls for a moment for reflection of how narrow and, at the same time, broad our view is on what constitutes as a phage defense system. The main question being: what qualifies and what does not qualify as a phage defense system? The current paradigm suggests any gene that inhibits the replication of any phage can be considered a phage defense system. But should we not be asking how relevant this inhibition is and how it came to be? Does a marginal inhibition count as true phage defense? And what about "phage defense systems" that have a primary function that is not related to phage defense? For example, most phage defense conveying toxin-antitoxins do not act on phages but instead are activated by the metabolic changes caused by the phage infection that affect the presence of the anti-toxin. Their primary function is to prevent their own removal from the host genome, so do these accidental phage defense systems count as true phage defense systems? With the current view, it is easily possible that more than half, if not all, of the host genome qualifies as a phage defense system. And at that point, you might as well refer to them as regular genes again.

"If one has a hammer one tends to look for nails"- Silvan Tomkins in 1963

Conclusions and Outlook

Advances in understanding the conflict between phage and host has significantly increased over the years, from knowing about just a few phage defenses to hundreds (*122, 138, 141, 159, 160*). For many years to come, these phage defenses will give enough food for thought for the phage-defense field. In the future, I propose that it will be important to view these phages defenses and evasion strategies in their full context, not only in the context of phage infection, but also in an ecological context and beyond. This is especially important if our goal is to use phages to combat bacterial infections, since phages' nature prevents the eradication of all its hosts. Overcoming this innate phage instinct likely requires us to engineer these phages, a delicate process that demands an immense understanding of the phage-host interaction, as well as requiring breakthroughs in biotechnology, such as the recent advances in *in silico* protein prediction. These breakthroughs will not only give us new insight into the phage-host interactions but also opens the possibility to start designing our own phage evasion strategies, enabling us to become better at designing a virus, than nature itself. The question, however, remains whether we should.

Epilogue: Consciousness

In a musky room, dimly lit by the morning light, a student wakes up after a long night of thinking. He rubs his eyes and blindly feels around for his glasses, fumbling past crumpled notes and left open books. His fingers find his glasses besides a mug filled with cold coffee from last night. He puts on his glasses. Although it wasn't much of an improvement, he was now able to perceive a better overview of the mess around him. "Right, last night" he says out loud, while he tries to remember where he left his thoughts. It comes to him, while sipping on his cold coffee that he found besides his couch.

"It still doesn't make any sense." He mumbles to himself.
"Why is it always the same?"

He gazes at his blackboard, which has completely turned white from the smeared-out scribbles of obsessive chalk drawings. "There is no reason for it" he says calmly while clearing the blackboard with his sleeve.

"Why do all electrons have the same charge and mass?".

Then it comes to him: "Maybe… they are all one and the same…"

For a long time, scientists have been curious on why all electrons have the same charge and mass. There is nothing preventing electrons from differing, yet they remain identical and indistinguishable, which has led to the one-electron theory. A theory that hypothesizes that all electrons are, in fact, one and the same. One single electron that pulsates through all possible electron positions one by one in space and time. From the electron's own perspective, the electron blinks almost instantaneously between all available positions, from one perceived electron to another across the whole universe. From the perspective of the observer, this is observed as several electrons circling around multiple nuclei, akin to how we perceive motion from film, which is inferred from rapidly changing frames.

This analogy not only holds for perception of motion from film but may also be extended to perception itself. Where one true consciousness blinks through all possible perspectives. To us, the observer, it is perceived as a continuous presence, while consciousness enters and leaves us all in every passing moment…

In a way, I is we and we is I.

References

1. M. W. Gray, Lynn Margulis and the endosymbiont hypothesis: 50 years later. *Molecular biology of the cell* **28**, 1285-1287 (2017).
2. H.-C. Flemming, S. Wuertz, Bacteria and archaea on Earth and their abundance in biofilms. *Nature Reviews Microbiology* **17**, 247-260 (2019).
3. K. E. Wommack, R. R. Colwell, Virioplankton: Viruses in Aquatic Ecosystems. *Microbiology and Molecular Biology Reviews* **64**, 69-114 (2000).
4. M. B. Dion, F. Oechslin, S. Moineau, Phage diversity, genomics and phylogeny. *Nature Reviews Microbiology* **18**, 125-138 (2020).
5. A. M. Korn, A. E. Hillhouse, L. Sun, J. J. Gill, Comparative Genomics of Three Novel Jumbo Bacteriophages Infecting Staphylococcus aureus. *Journal of Virology* **95**, (2021).
6. A. Chevallereau, B. J. Pons, S. Van Houte, E. R. Westra, Interactions between bacterial and phage communities in natural environments. *Nature Reviews Microbiology* **20**, 49-62 (2022).
7. S. Mäntynen, E. Laanto, H. M. Oksanen, M. M. Poranen, S. L. Díaz-Muñoz, Black box of phage–bacterium interactions: exploring alternative phage infection strategies. *Open Biology* **11**, 210188 (2021).
8. E. P. C. Rocha, D. Bikard, Microbial defenses against mobile genetic elements and viruses: Who defends whom from what? *PLoS Biol* **20**, e3001514 (2022).
9. R. Sayid *et al.*, Characteristics of phage-plasmids and their impact on microbial communities. *Essays in Biochemistry* **68**, 583-592 (2024).
10. H. Georjon, A. Bernheim, The highly diverse antiphage defence systems of bacteria. *Nature Reviews Microbiology*, (2023).
11. J. E. Egido, A. R. Costa, C. Aparicio-Maldonado, P.-J. Haas, S. J. J. Brouns, Mechanisms and clinical importance of bacteriophage resistance. *FEMS Microbiol Rev* **46**, fuab048 (2022).
12. V. Ongenae *et al.*, Reversible bacteriophage resistance by shedding the bacterial cell wall. *Open Biology* **12**, (2022).
13. T. Reyes-Robles *et al.*, Vibrio cholerae outer membrane vesicles inhibit bacteriophage infection. *J Bacteriol* **200**, 10.1128/jb. 00792-00717 (2018).
14. E. Huiting, J. Bondy-Denomy, Defining the expanding mechanisms of phage-mediated activation of bacterial immunity. *Current Opinion in Microbiology* **74**, 102325 (2023).
15. S. P. B. van Beljouw, J. Sanders, A. Rodríguez-Molina, S. J. J. Brouns, RNA-targeting CRISPR–Cas systems. *Nature Reviews Microbiology* **21**, 21-34 (2023).
16. D. V. Banh *et al.*, Bacterial cGAS senses a viral RNA to initiate immunity. *Nature* **623**, 1001-1008 (2023).
17. S. E. Luria, M. L. Human, A nonhereditary, host-induced variation of bacterial viruses. *J Bacteriol* **64**, 557-569 (1952).
18. K. S. Makarova *et al.*, Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nature Reviews Microbiology* **18**, 67-83 (2020).
19. O. T. Tuck *et al.*, Genome integrity sensing by the broad-spectrum Hachiman antiphage defense complex. *Cell* **187**, 6914-6928.e6920 (2024).
20. L. Loeff, A. Walter, G. T. Rosalen, M. Jinek, DNA end sensing and cleavage by the Shedu anti-phage defense system. *Cell* **188**, 721-733.e717 (2025).
21. L. A. Gao *et al.*, Prokaryotic innate immunity through pattern recognition of conserved viral proteins. *Science* **377**, eabm4096 (2022).
22. T. Zhang *et al.*, A bacterial immunity protein directly senses two disparate phage proteins. *Nature*, (2024).
23. C. G. Roberts, C. B. Fishman, D. V. Banh, L. A. Marraffini, A bacterial TIR-based immune system senses viral capsids to initiate defense. *bioRxiv*, (2024).
24. C. Bernard, Y. Li, P. Lopez, E. Bapteste, Beyond arbitrium: identification of a second communication system in Bacillus phage phi3T that may regulate host defense mechanisms. *ISME J* **15**, 545-549 (2021).

25. A. Kelly, T. J. Arrowsmith, S. C. Went, T. R. Blower, Toxin–antitoxin systems as mediators of phage defence and the implications for abortive infection. *Current Opinion in Microbiology* **73**, 102293 (2023).
26. R. Cheng *et al.*, Prokaryotic Gabija complex senses and executes nucleotide depletion and DNA cleavage for antiviral defense. *Cell Host & Microbe* **31**, 1331-1344.e1335 (2023).
27. H. Hu *et al.*, Structure and mechanism of the Zorya anti-phage defense system. *Nature*, (2024).
28. Z. Zhang *et al.*, Kiwa is a bacterial membrane-embedded defence supercomplex activated by phage-induced membrane changes. *bioRxiv*, (2023).
29. D. Mayo-Muñoz *et al.*, Type III CRISPR-Cas provides resistance against nucleus-forming jumbo phages via abortive infection. *Molecular Cell* **82**, 4471-4486.e4479 (2022).
30. A. Kupczok *et al.*, Rates of Mutation and Recombination in Siphoviridae Phage Genome Evolution over Three Decades. *Mol Biol Evol* **35**, 1147-1159 (2018).
31. H. Guo, L. Arambula, P. Ghosh, J. F. Miller. (ASM Press, 2015), pp. 1237-1252.
32. A. Stokar-Avihail *et al.*, Discovery of phage determinants that confer sensitivity to bacterial immune systems. *Cell* **186**, 1863-1876.e1816 (2023).
33. R. Ouyang, V. Ongenae, A. Muok, D. Claessen, A. Briegel, Phage fibers and spikes: a nanoscale Swiss army knife for host infection. *Current Opinion in Microbiology* **77**, 102429 (2024).
34. I. Osterman *et al.*, Phages reconstitute NAD(+) to counter bacterial immunity. *Nature* **634**, 1160-1167 (2024).
35. R. W. Holley *et al.*, Structure of a Ribonucleic Acid. *Science* **147**, 1462-1465 (1965).
36. T. Zhang *et al.*, Direct activation of a bacterial innate immune system by a viral capsid protein. *Nature* **612**, 132-140 (2022).
37. N. Burman *et al.*, A virally-encoded tRNA neutralizes the PARIS antiviral defence system. *Nature*, (2024).
38. A. Millman *et al.*, Bacterial Retrons Function In Anti-Phage Defense. *Cell* **183**, 1551-1561.e1512 (2020).
39. A. H. Azam *et al.*, Evasion of antiviral bacterial immunity by phage tRNAs. *Nat Commun* **15**, 9586 (2024).
40. S. Srikant, C. K. Guegler, M. T. Laub, The evolution of a counter-defense mechanism in a virus constrains its host range. *Elife* **11**, e79549 (2022).
41. A. Bernheim, R. Sorek, The pan-immune system of bacteria: antiviral defence as a community resource. *Nature Reviews Microbiology* **18**, 113-119 (2020).
42. E. Bolotin, R. Hershberg, Bacterial intra-species gene loss occurs in a largely clocklike manner mostly within a pool of less conserved and constrained genes. *Scientific Reports* **6**, 35168 (2016).
43. R. Pinilla-Redondo *et al.*, Discovery of multiple anti-CRISPRs highlights anti-defense gene clustering in mobile genetic elements. *Nat Commun* **11**, 5652 (2020).
44. S. Qin *et al.*, *Pseudomonas aeruginosa*: pathogenesis, virulence factors, antibiotic resistance, interaction with host, technology advances and emerging therapeutics. *Signal Transduction and Targeted Therapy* **7**, (2022).
45. S. Kamal *et al.*, Risk Factors and Clinical Characteristics of Pandrug-Resistant *Pseudomonas aeruginosa*. *Cureus*, (2024).
46. Z. Chegini *et al.*, Bacteriophage therapy against *Pseudomonas aeruginosa* biofilms: a review. *Annals of Clinical Microbiology and Antimicrobials* **19**, (2020).
47. F. Tesson *et al.*, Systematic and quantitative view of the antiviral arsenal of prokaryotes. *bioRxiv*, 2021.2009.2002.458658 (2021).
48. J. De Smet, H. Hendrix, B. G. Blasdel, K. Danis-Wlodarczyk, R. Lavigne, *Pseudomonas* predators: understanding and exploiting phage–host interactions. *Nature Reviews Microbiology* **15**, 517-530 (2017).
49. D. Piel *et al.*, Phage–host coevolution in natural populations. *Nature Microbiology*, (2022).
50. A. R. Burmeister *et al.*, Pleiotropy complicates a trade-off between phage resistance and antibiotic resistance. *Proceedings of the National Academy of Sciences* **117**, 11207-11216 (2020).
51. S. J. Labrie, J. E. Samson, S. Moineau, Bacteriophage resistance mechanisms. *Nature Reviews Microbiology* **8**, 317-327 (2010).

52. H. G. Hampton, B. N. J. Watson, P. C. Fineran, The arms race between bacteria and their phage foes. *Nature* **577**, 327-336 (2020).
53. R. Barrangou *et al.*, CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315**, 1709-1712 (2007).
54. P. Mohanraju *et al.*, Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353**, aad5147 (2016).
55. T. A. Bickle, Restricting restriction. *Molecular Microbiology* **51**, 3-5 (2004).
56. D. Dussoix, W. Arber, Host specificity of DNA produced by Escherichia coli: II. Control over acceptance of DNA from infecting phage λ. *Journal of Molecular Biology* **5**, 37-49 (1962).
57. E. V. Koonin, K. S. Makarova, Y. I. Wolf, Evolutionary Genomics of Defense Systems in Archaea and Bacteria. *Annu Rev Microbiol* **71**, 233-261 (2017).
58. N. Tal, R. Sorek, SnapShot: Bacterial immunity. *Cell* **185**, 578-578.e571 (2022).
59. A. Millman *et al.*, An expanding arsenal of immune systems that protect bacteria from phages. *bioRxiv*, 2022.2005.2011.491447 (2022).
60. C. Vassallo, C. Doering, M. L. Littlehale, G. Teodoro, M. T. Laub, Mapping the landscape of anti-phage defense mechanisms in the E. coli pangenome. *bioRxiv*, 2022.2005.2012.491691 (2022).
61. F. A. Hussain *et al.*, Rapid evolutionary turnover of mobile genetic elements drives bacterial resistance to phages. *Science* **374**, 488-492 (2021).
62. F. Rousset *et al.*, Phages and their satellites encode hotspots of antiviral systems. *Cell Host Microbe* **30**, 740-753.e745 (2022).
63. F. Tesson *et al.*, Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat Commun* **13**, 2561 (2022).
64. D. Subedi, A. K. Vijay, G. S. Kohli, S. A. Rice, M. Willcox, Comparative genomics of clinical strains of *Pseudomonas aeruginosa* strains isolated from different geographic sites. *Scientific Reports* **8**, 15668 (2018).
65. V. Krylov *et al.*, Phage phiKZ-The First of Giants. *Viruses* **13**, 149 (2021).
66. H. Bin Jang *et al.*, Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nature Biotechnology* **37**, 632-639 (2019).
67. M. Hunter, D. Fusco, Superinfection exclusion: A viral strategy with short-term benefits and long-term drawbacks. *PLoS Comput Biol* **18**, e1010125 (2022).
68. S. A. Jackson, N. Birkholz, L. M. Malone, P. C. Fineran, Imprecise Spacer Acquisition Generates CRISPR-Cas Immune Diversity through Primed Adaptation. *Cell Host Microbe* **25**, 250-260.e254 (2019).
69. P. C. Fineran *et al.*, Degenerate target sites mediate rapid primed CRISPR adaptation. *Proceedings of the National Academy of Sciences* **111**, E1629-E1638 (2014).
70. R. H. J. Staals *et al.*, Interference-driven spacer acquisition is dominant over naive and primed adaptation in a native CRISPR–Cas system. *Nat Commun* **7**, 12853 (2016).
71. K. C. Cady *et al.*, Prevalence, conservation and functional analysis of Yersinia and Escherichia CRISPR regions in clinical *Pseudomonas aeruginosa* isolates. *Microbiology (Reading)* **157**, 430-437 (2011).
72. H. P. Schweizer, Escherichia-*Pseudomonas* shuttle vectors derived from pUC18/19. *Gene* **97**, 109-112 (1991).
73. E. Huiting *et al.*, Bacteriophages antagonize cGAS-like immunity in bacteria. *bioRxiv*, 2022.2003.2030.486325 (2022).
74. T. Zhang *et al.*, Direct activation of an innate immune system in bacteria by a viral capsid protein. *bioRxiv*, 2022.2005.2030.493996 (2022).
75. A. Stokar-Avihail *et al.*, Discovery of phage determinants that confer sensitivity to bacterial immune systems. *bioRxiv*, 2022.2008.2027.505566 (2022).
76. R. Cheng *et al.*, Prokaryotic Gabija complex senses and executes nucleotide depletion and DNA cleavage for antiviral defense. *Cell Host Microbe*, (2023).
77. C. K. Guegler, M. T. Laub, Shutoff of host transcription triggers a toxin-antitoxin system to cleave phage RNA and abort infection. *Molecular Cell* **81**, 2361-2373.e2369 (2021).

78. E. Huiting *et al.*, Bacteriophages inhibit and evade cGAS-like immune function in bacteria. *Cell*, (2023).
79. Q. Ye *et al.*, HORMA Domain Proteins and a Trip13-like ATPase Regulate Bacterial cGAS-like Enzymes to Mediate Bacteriophage Immunity. *Molecular Cell* **77**, 709-722.e707 (2020).
80. V. Chaikeeratisak *et al.*, Assembly of a nucleus-like structure during viral replication in bacteria. *Science* **355**, 194-197 (2017).
81. L. M. Malone *et al.*, A jumbo phage that forms a nucleus-like structure evades CRISPR–Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nature Microbiology* **5**, 48-55 (2020).
82. S. D. Mendoza *et al.*, A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. *Nature* **577**, 244-248 (2020).
83. S. A. McMahon *et al.*, Extensive DNA mimicry by the ArdA anti-restriction protein and its role in the spread of antibiotic resistance. *Nucleic acids research* **37**, 4887-4897 (2009).
84. W. Liang *et al.*, Anti-Restriction Protein, KlcAHS, Promotes Dissemination of Carbapenem Resistance. *Frontiers in Cellular and Infection Microbiology* **7**, (2017).
85. D. Serfiotis-Mitsa *et al.*, The structure of the KlcA and ArdB proteins reveals a novel fold and antirestriction activity against Type I DNA restriction systems in vivo but not in vitro. *Nucleic acids research* **38**, 1723-1737 (2010).
86. S. J. Hobbs *et al.*, Phage anti-CBASS and anti-Pycsar nucleases subvert bacterial immunity. *Nature* **605**, 522-526 (2022).
87. P. Ho, Y. Chen, S. Biswas, E. Canfield, D. E. Feldman, Bacteriophage anti-defense genes that neutralize TIR and STING immune responses. *bioRxiv*, 2022.2006.2009.495361 (2022).
88. S. Shehreen, N. Birkholz, Peter C. Fineran, Chris M. Brown, Widespread repression of anti-CRISPR production by anti-CRISPR-associated proteins. *Nucleic Acids Research* **50**, 8615-8625 (2022).
89. F. J. Whelan, M. Rusilowicz, J. O. McInerney, Coinfinder: detecting significant associations and dissociations in pangenomes. *Microbial Genomics* **6**, (2020).
90. A. Lopatina, N. Tal, R. Sorek, Abortive Infection: Bacterial Suicide as an Antiviral Immune Strategy. *Annual Review of Virology* **7**, 371-384 (2020).
91. D. V. Banh, C. G. Roberts, A. M. Amador, S. F. Brady, L. A. Marraffini, Bacterial cGAS senses a viral RNA to initiate immunity. *bioRxiv*, 2023.2003.2007.531596 (2023).
92. A. T. Whiteley *et al.*, Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. *Nature* **567**, 194-199 (2019).
93. M. C. Williams *et al.*, Restriction endonuclease cleavage of phage DNA enables resuscitation from Cas13-induced bacterial dormancy. *Nature Microbiology*, (2023).
94. M. Jaskólska, D. W. Adams, M. Blokesch, Two defence systems eliminate plasmids from seventh pandemic Vibrio cholerae. *Nature* **604**, 323-329 (2022).
95. Y. Wu *et al.*, Defence systems provide synergistic anti-phage activity in E. coli. *bioRxiv*, 2022.2008.2021.504612 (2022).
96. M.-È. Dupuis, M. Villion, A. H. Magadán, S. Moineau, CRISPR-Cas and restriction–modification systems are compatible and increase phage resistance. *Nat Commun* **4**, 2087 (2013).
97. N. Birkholz, S. A. Jackson, R. D. Fagerlund, Peter C. Fineran, A mobile restriction–modification system provides phage defence and resolves an epigenetic conflict with an antagonistic endonuclease. *Nucleic Acids Research* **50**, 3348-3361 (2022).
98. F. L. Nobrega *et al.*, Targeting mechanisms of tailed bacteriophages. *Nature Reviews Microbiology* **16**, 760-773 (2018).
99. S. T. Abedon, Lysis from without. *Bacteriophage* **1**, 46-49 (2011).
100. R. R. Wick, L. M. Judd, C. L. Gorrie, K. E. Holt, Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* **13**, e1005595 (2017).
101. R. R. Wick, M. B. Schultz, J. Zobel, K. E. Holt, Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* **31**, 3350-3352 (2015).
102. D. E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2. *Genome Biology* **20**, 257 (2019).
103. S. Andrews. FastQC: a quality control tool for high throughput sequence data. (2010).

104. R. Challis, Assembly-stats: get assembly statistics from FASTA and FASTQ files. (2017).
105. F. A. Simão, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210-3212 (2015).
106. H. Li, Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094-3100 (2018).
107. P. Danecek *et al.*, Twelve years of SAMtools and BCFtools. *GigaScience* **10**, (2021).
108. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
109. D. Hyatt *et al.*, Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics* **11**, 119-119 (2010).
110. K. A. Jolley, M. C. J. Maiden, BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* **11**, 595 (2010).
111. T. J. Treangen, B. D. Ondov, S. Koren, A. M. Phillippy, The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biology* **15**, 524 (2014).
112. J. Guo *et al.*, VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. *Microbiome* **9**, 37 (2021).
113. S. Nayfach *et al.*, CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nature Biotechnology* **39**, 578-585 (2021).
114. R. K. Aziz *et al.*, The RAST Server: rapid annotations using subsystems technology. *BMC genomics* **9**, 75-75 (2008).
115. J. R. Garneau, F. Depardieu, L.-C. Fortier, D. Bikard, M. Monot, PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Scientific Reports* **7**, 8292 (2017).
116. P. Tynecki *et al.*, PhageAI - Bacteriophage Life Cycle Recognition with Machine Learning and Natural Language Processing. *bioRxiv*, 2020.2007.2011.198606 (2020).
117. P. Aiewsakun, E. M. Adriaenssens, R. Lavigne, A. M. Kropinski, P. Simmonds, Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. *J Gen Virol* **99**, 1331-1343 (2018).
118. P. Shannon *et al.*, Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* **13**, 2498-2504 (2003).
119. C. L. M. Gilchrist, Y.-H. Chooi, clinker & clustermap.js: automatic generation of gene cluster comparison figures. *Bioinformatics* **37**, 2473-2475 (2021).
120. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *Journal of Molecular Biology* **215**, 403-410 (1990).
121. L. J. Payne *et al.*, Identification and classification of antiviral defence systems in bacteria and archaea with PADLOC reveals new system types. *Nucleic Acids Research* **49**, 10868-10878 (2021).
122. L. Gao *et al.*, Diverse enzymatic activities mediate antiviral immunity in prokaryotes. *Science* **369**, 1077-1084 (2020).
123. F. Rousset *et al.*, A conserved family of immune effectors cleaves cellular ATP upon viral infection. *bioRxiv*, 2023.2001.2024.525353 (2023).
124. A. Biswas, R. H. J. Staals, S. E. Morales, P. C. Fineran, C. M. Brown, CRISPRDetect: A flexible algorithm to define CRISPR arrays. *BMC genomics* **17**, 356-356 (2016).
125. H. Yi *et al.*, AcrFinder: genome mining anti-CRISPR operons in prokaryotes and their viruses. *Nucleic Acids Research* **48**, W358-W365 (2020).
126. S. F. Altschul *et al.*, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402 (1997).
127. R. D. Finn, J. Clements, S. R. Eddy, HMMER web server: interactive sequence similarity searching. *Nucleic acids research* **39**, W29-W37 (2011).
128. J. Söding, A. Biegert, A. N. Lupas, The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research* **33**, W244-W248 (2005).

129. K.-H. Choi, A. Kumar, H. P. Schweizer, A 10-min method for preparation of highly electrocompetent *Pseudomonas aeruginosa* cells: Application for DNA fragment transfer between chromosomes and plasmid transformation. *Journal of Microbiological Methods* **64**, 391-397 (2006).
130. E.-M. Lammens, M. Boon, D. Grimon, Y. Briers, R. Lavigne, SEVAtile: a standardised DNA assembly method optimised for *Pseudomonas*. *Microbial Biotechnology* **15**, 370-386 (2022).
131. A. Mazzocco, T. E. Waddell, E. Lingohr, R. P. Johnson, in *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions,* M. R. J. Clokie, A. M. Kropinski, Eds. (Humana Press, Totowa, NJ, 2009), pp. 81-85.
132. A. M. Kropinski, A. Mazzocco, T. E. Waddell, E. Lingohr, R. P. Johnson, in *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions,* M. R. J. Clokie, A. M. Kropinski, Eds. (Humana Press, Totowa, NJ, 2009), pp. 69-76.
133. R. R. Wick, P. Menzel, Filtlong. *Available online: github. com/rrwick/Filtlong (accessed on 15 August 2021)*, (2017).
134. H. Li, Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* **32**, 2103-2110 (2016).
135. M. Kolmogorov, J. Yuan, Y. Lin, P. A. Pevzner, Assembly of long, error-prone reads using repeat graphs. *Nature Biotechnology* **37**, 540-546 (2019).
136. J. N. A. Vink, S. J. J. Brouns, J. Hohlbein, Extracting Transition Rates in Particle Tracking Using Analytical Diffusion Distribution Analysis. *Biophys J* **119**, 1970-1983 (2020).
137. S. Doron *et al.*, Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
138. D. F. van den Berg *et al.*, Bacterial homologs of innate eukaryotic antiviral defenses with anti-phage activity highlight shared evolutionary roots of viral defenses. *Cell Host Microbe*, (2024).
139. P. C. Deweirdt, E. M. Mahoney, M. T. Laub, DefensePredictor: A Machine Learning Model to Discover Novel Prokaryotic Immune Systems. *bioRxiv*, (2025).
140. F. Rousset, J. Dowding, A. Bernheim, E. P. C. Rocha, D. Bikard, Prophage-encoded hotspots of bacterial immune systems. *bioRxiv*, 2021.2001.2021.427644 (2021).
141. C. N. Vassallo, C. R. Doering, M. L. Littlehale, G. I. C. Teodoro, M. T. Laub, A functional selection reveals previously undetected anti-phage defence systems in the E. coli pangenome. *Nature Microbiology* **7**, 1568-1579 (2022).
142. L. J. Payne, T. C. D. Hughes, P. C. Fineran, S. A. Jackson, New antiviral defences are genetically embedded within prokaryotic immune systems. *bioRxiv*, (2024).
143. I. Bleriot *et al.*, Improving phage therapy by evasion of phage resistance mechanisms. *JAC-Antimicrobial Resistance* **6**, (2023).
144. L. Yuping *et al.*, Jumbo phage killer immune system targets early infection of nucleus-forming phages. *Cell*, (2025).
145. A. G. Moller *et al.*, Genes Influencing Phage Host Range in Staphylococcus aureus on a Species-Wide Scale. *mSphere* **6**, (2021).
146. A. R. Costa *et al.*, Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*. *Sci Adv* **10**, eadj0341 (2024).
147. J. A. Lees, M. Galardini, S. D. Bentley, J. N. Weiser, J. Corander, pyseer: a comprehensive tool for microbial pangenome-wide association studies. *Bioinformatics* **34**, 4310-4312 (2018).
148. G. Tonkin-Hill *et al.*, Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biology* **21**, (2020).
149. P. Ge *et al.*, Action of a minimal contractile bactericidal nanomachine. *Nature* **580**, 658-662 (2020).
150. L. Blasco *et al.*, Study of 32 new phage tail-like bacteriocins (pyocins) from a clinical collection of *Pseudomonas aeruginosa* and of their potential use as typing markers and antimicrobial agents. *Scientific Reports* **13**, (2023).
151. S. R. Williams, D. Gebhart, D. W. Martin, D. Scholl, Retargeting R-Type Pyocins To Generate Novel Bactericidal Protein Complexes. *Applied and Environmental Microbiology* **74**, 3868-3876 (2008).

152. M. Mei, I. Estrada, S. P. Diggle, J. B. Goldberg, R-pyocins as targeted antimicrobials against *Pseudomonas aeruginosa*. *npj Antimicrobials and Resistance* **3**, (2025).
153. K. Nakayama *et al.*, The R-type pyocin of *Pseudomonas aeruginosa* is related to P2 phage, and the F-type is related to lambda phage. *Molecular Microbiology* **38**, 213-231 (2000).
154. J. Penterman, P. K. Singh, G. C. Walker, Biological Cost of Pyocin Production during the SOS Response in *Pseudomonas aeruginosa*. *J Bacteriol* **196**, 3351-3359 (2014).
155. J. De Smet *et al.*, The bacteriophage LUZ24 "Igy" peptide inhibits the *Pseudomonas* DNA gyrase. *Cell Rep* **36**, 109567 (2021).
156. T. Seemann, Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068-2069 (2014).
157. B. Q. Minh *et al.*, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Molecular Biology and Evolution* **37**, 1530-1534 (2020).
158. D. Hochhauser, A. Millman, R. Sorek, The defense island repertoire of the Escherichia coli pan-genome. *PLOS Genetics* **19**, e1010694 (2023).
159. S. Doron *et al.*, Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).
160. A. Millman *et al.*, An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* **30**, 1556-1569.e1555 (2022).
161. M. C. Johnson *et al.*, Core defense hotspots within *Pseudomonas aeruginosa* are a consistent and rich source of anti-phage defense systems. *Nucleic Acids Research* **51**, 4995-5005 (2023).
162. D. Ka, H. Oh, E. Park, J.-H. Kim, E. Bae, Structural and functional evidence of bacterial antiphage protection by Thoeris defense system via NAD+ degradation. *Nat Commun* **11**, 2816 (2020).
163. A. G. Johnson *et al.*, Bacterial gasdermins reveal an ancient mechanism of cell death. *Science* **375**, 221-225 (2022).
164. K. S. Makarova, Y. I. Wolf, J. van der Oost, E. V. Koonin, Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. *Biol Direct* **4**, 29-29 (2009).
165. B. Koopal, A. J. Kruis, N. J. Claassens, F. L. Nobrega, J. van der Oost, Incorporation of a Synthetic Amino Acid into dCas9 Improves Control of Gene Silencing. *ACS Synthetic Biology* **8**, 216-222 (2019).
166. H. Shomar *et al.*, Viperin immunity evolved across the tree of life through serial innovations on a conserved scaffold. *bioRxiv*, 2023.2009.2013.557418 (2023).
167. P. Leão *et al.*, Asgard archaea defense systems and their roles in the origin of immunity in eukaryotes. *bioRxiv*, 2023.2009.2013.557551 (2023).
168. M. S. Miller *et al.*, Senataxin suppresses the antiviral transcriptional response and controls viral biogenesis. *Nature Immunology* **16**, 485-494 (2015).
169. A. M. Murphy, B. Otto, C. A. Brearley, J. P. Carr, D. E. Hanke, A role for inositol hexakisphosphate in the maintenance of basal resistance to plant pathogens. *The Plant Journal* **56**, 638-652 (2008).
170. M. S. Bezerra Espinola *et al.*, Inositol and vitamin D may naturally protect human reproduction and women undergoing assisted reproduction from Covid-19 risk. *Journal of Reproductive Immunology* **144**, 103271 (2021).
171. M. L. Bizzarri, A. S.; Aragona, D.; Unfer, V., Inositol and pulmonary function. Could myo-inositol treatment downregulate inflammation and cytokine release syndrome in SARS-CoV-2? *Eur Rev Med Pharmacol Sci* **24**, 3426-3432 (2020).
172. Q. Zhang, Y.-C. J. Wang, E. A. Montalvo, Sμ bp-2 Represses the Epstein–Barr Virus Lytic Switch Promoter. *Virology* **255**, 160-170 (1999).
173. D. N. Fusco *et al.*, HELZ2 Is an IFN Effector Mediating Suppression of Dengue Virus. *Front Microbiol* **8**, (2017).
174. S. Vavassori *et al.*, Multisystem inflammation and susceptibility to viral infections in human ZNFX1 deficiency. *Journal of Allergy and Clinical Immunology* **148**, 381-393 (2021).
175. L. Wang, I. Sola, L. Enjuanes, S. Zuñiga, MOV10 Helicase Interacts with Coronavirus Nucleocapsid Protein and Has Antiviral Activity. *mBio* **12**, 10.1128/mbio.01316-01321 (2021).

176. R. A. Cuevas *et al.*, MOV10 Provides Antiviral Activity against RNA Viruses by Enhancing RIG-I–MAVS-Independent IFN Induction. *The Journal of Immunology* **196**, 3877-3886 (2016).
177. A. K. P. Serquiña *et al.*, UPF1 Is Crucial for the Infectivity of Human Immunodeficiency Virus Type 1 Progeny Virions. *Journal of Virology* **87**, 8853-8861 (2013).
178. J. P. May, A. E. Simon, Targeting of viral RNAs by Upf1-mediated RNA decay pathways. *Current Opinion in Virology* **47**, 1-8 (2021).
179. A. Richardson *et al.*, S58 UPF1 is a novel modulator of antiviral responses against rhinovirus and is deficient in patients with severe asthma. *Thorax* **77**, A37-A38 (2022).
180. K. Breuer *et al.*, InnateDB: systems biology of innate immunity and beyond—recent updates and continuing curation. *Nucleic Acids Research* **41**, D1228-D1233 (2012).
181. J. Calle García *et al.*, PRGdb 4.0: an updated database dedicated to genes involved in plant disease resistance process. *Nucleic Acids Research* **50**, D1483-D1490 (2021).
182. D. J. Richter *et al.*, EukProt: A database of genome-scale predicted proteins across the diversity of eukaryotes. *Peer Community Journal* **2**, (2022).
183. M. Steinegger, J. Söding, MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nature Biotechnology* **35**, 1026-1028 (2017).
184. J. Zou, M. Chang, P. Nie, C. J. Secombes, Origin and evolution of the RIG-I like RNA helicase gene family. *BMC Evolutionary Biology* **9**, 85 (2009).
185. O. Andrisani *et al.*, Biological functions of DEAD/DEAH-box RNA helicases in health and disease. *Nature Immunology* **23**, 354-357 (2022).
186. G. Ofir *et al.*, Antiviral activity of bacterial TIR domains via signaling molecules that trigger cell death. *bioRxiv*, 2021.2001.2006.425286 (2021).
187. D. D. Leipe, E. V. Koonin, L. Aravind, STAND, a Class of P-Loop NTPases Including Animal and Plant Regulators of Programmed Cell Death: Multiple, Complex Domain Architectures, Unusual Phyletic Patterns, and Evolution by Horizontal Gene Transfer. *Journal of Molecular Biology* **343**, 1-28 (2004).
188. A. Millman, S. Melamed, G. Amitai, R. Sorek, Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. *Nature Microbiology* **5**, 1608-1615 (2020).
189. F. M. Zerbini *et al.*, Changes to virus taxonomy and the ICTV Statutes ratified by the International Committee on Taxonomy of Viruses (2023). *Archives of Virology* **168**, 175 (2023).
190. K. Li *et al.*, SuhB Is a Regulator of Multiple Virulence Genes and Essential for Pathogenesis of *Pseudomonas aeruginosa*. *mBio* **4**, 10.1128/mbio.00419-00413 (2013).
191. L. Holm, L. M. Laakso, Dali server update. *Nucleic Acids Research* **44**, W351-W355 (2016).
192. K. Tariq, B. W. Luikart, Striking a balance: PIP2 and PIP3 signaling in neuronal health and disease. *Exploration of Neuroprotective Therapy* **1**, 86-110 (2021).
193. L. Hu *et al.*, Myo-inositol mediates reactive oxygen species-induced programmed cell death via salicylic acid-dependent and ethylene-dependent pathways in apple. *Horticulture Research* **7**, 138 (2020).
194. R. Goswami *et al.*, Inositol Monophosphatase: A Bifunctional Enzyme in Mycobacterium smegmatis. *ACS Omega* **3**, 13876-13881 (2018).
195. S. Lu *et al.*, Insights into the Role of Magnesium Triad in myo-Inositol Monophosphatase: Metal Mechanism, Substrate Binding, and Lithium Therapy. *Journal of Chemical Information and Modeling* **52**, 2398-2409 (2012).
196. N. Ferruz, G. Tresadern, A. Pineda-Lucena, G. De Fabritiis, Multibody cofactor and substrate molecular recognition in the myo-inositol monophosphatase enzyme. *Scientific Reports* **6**, 30275 (2016).
197. K. Faisal Tarique, S. Arif Abdul Rehman, S. Gourinath, Structural elucidation of a dual-activity PAP phosphatase-1 from Entamoeba histolytica capable of hydrolysing both 3'-phosphoadenosine 5'-phosphate and inositol 1,4-bisphosphate. *Acta Crystallogr D Biol Crystallogr* **70**, 2019-2031 (2014).
198. S. Rattay, M. Hufbauer, P. Hoboth, M. Sztacho, B. Akgül, Viruses and phospholipids: Friends and foes during infection. *Journal of Medical Virology* **95**, e28658 (2023).
199. D. D. Martin, R. A. Ciulla, M. F. Roberts, Osmoadaptation in Archaea. *Applied and Environmental Microbiology* **65**, 1815-1825 (1999).

200. S. L. Heaver *et al.*, Characterization of inositol lipid metabolism in gut-associated Bacteroidetes. *Nature Microbiology* **7**, 986-1000 (2022).
201. T. B. Reynolds, Strategies for acquiring the phospholipid metabolite inositol in pathogenic bacteria, fungi and protozoa: making it and taking it. *Microbiology* **155**, 1386-1396 (2009).
202. F. Li *et al.*, High-resolution cryo-EM structure of the *Pseudomonas* bacteriophage E217. *Nat Commun* **14**, 4052 (2023).
203. J. Garbe *et al.*, Characterization of JG024, a *Pseudomonas aeruginosa* PB1-like broad host range phage under simulated infection conditions. *BMC Microbiol* **10**, 301 (2010).
204. O. J. Wilkinson, C. Carrasco, C. Aicart-Ramos, F. Moreno-Herrero, M. S. Dillingham, Bulk and single-molecule analysis of a bacterial DNA2-like helicase–nuclease reveals a single-stranded DNA looping motor. *Nucleic Acids Research* **48**, 7991-8005 (2020).
205. B. P. Belotserkovskii, S. Tornaletti, A. D. D'Souza, P. C. Hanawalt, R-loop generation during transcription: Formation, processing and cellular outcomes. *DNA Repair* **71**, 69-81 (2018).
206. Z. Hasanova, V. Klapstova, O. Porrua, R. Stefl, M. Sebesta, Human senataxin is a bona fide R-loop resolving enzyme and transcription termination factor. *Nucleic Acids Research* **51**, 2818-2837 (2023).
207. M. Groh, L. O. Albulescu, A. Cristini, N. Gromak, Senataxin: Genome Guardian at the Interface of Transcription and Neurodegeneration. *Journal of Molecular Biology* **429**, 3181-3195 (2017).
208. V. Brázda, J. Havlík, J. Kolomazník, O. Trenz, J. Šťastný, R-Loop Tracker: Web Access-Based Tool for R-Loop Detection and Analysis in Genomic DNA Sequences. *International Journal of Molecular Sciences* **22**, 12857 (2021).
209. V. Anantharaman, K. S. Makarova, A. M. Burroughs, E. V. Koonin, L. Aravind, Comprehensive analysis of the HEPN superfamily: identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol Direct* **8**, 15 (2013).
210. M. Skružný *et al.*, An Endoribonuclease Functionally Linked to Perinuclear mRNP Quality Control Associates with the Nuclear Pore Complexes. *PLOS Biology* **7**, e1000008 (2009).
211. J. Wang *et al.*, The conserved domain database in 2023. *Nucleic Acids Research* **51**, D384-D388 (2022).
212. J. A. Duncan, S. W. Canna, The NLRC4 Inflammasome. *Immunological Reviews* **281**, 115-123 (2018).
213. P. Rosenstiel, A. Till, S. Schreiber, NOD-like receptors and human diseases. *Microbes and Infection* **9**, 648-657 (2007).
214. S. Bauernfried, M. J. Scherr, A. Pichlmair, K. E. Duderstadt, V. Hornung, Human NLRP1 is a sensor for double-stranded RNA. *Science* **371**, eabd0811 (2021).
215. N. Maruta *et al.*, Structural basis of NLR activation and innate immune signalling in plants. *Immunogenetics* **74**, 5-26 (2022).
216. Y. Gao, M. L. Skowyra, P. Feng, T. A. Rapoport, Protein import into peroxisomes occurs through a nuclear pore–like phase. *Science* **378**, eadf3971 (2022).
217. E. Baggs, G. Dagdas, K. V. Krasileva, NLR diversity, helpers and integrated domains: making sense of the NLR IDentity. *Current Opinion in Plant Biology* **38**, 59-67 (2017).
218. E. M. Kibby *et al.*, Bacterial NLR-related proteins protect against phage. *Cell* **186**, 2410-2424.e2418 (2023).
219. E. Cebrián-Sastre, I. Martín-Blecua, S. Gullón, J. Blázquez, A. Castañeda-García, Control of Genome Stability by EndoMS/NucS-Mediated Non-Canonical Mismatch Repair. *Cells* **10**, 1314 (2021).
220. K. Takeda, C. Akimoto, M. Kawamukai, Effects of the Escherichia coli sfsA Gene on mal Genes Expression and a DNA Binding Activity of SfsA. *Bioscience, Biotechnology, and Biochemistry* **65**, 213-217 (2001).
221. B. Ren *et al.*, Structure and function of a novel endonuclease acting on branched DNA substrates. *EMBO J* **28**, 2479-2489 (2009).
222. C. Creze *et al.*, Modulation of the Pyrococcus abyssi NucS Endonuclease Activity by Replication Clamp at Functional and Structural Levels*. *Journal of Biological Chemistry* **287**, 15648-15660 (2012).

223. A. M. Burroughs, L. M. Iyer, L. Aravind, Natural history of the E1-like superfamily: Implication for adenylation, sulfur transfer, and ubiquitin conjugation. *Proteins: Structure, Function, and Bioinformatics* **75**, 895-910 (2009).
224. M. K. Isaacson, H. L. Ploegh, Ubiquitination, Ubiquitin-like Modifiers, and Deubiquitination in Viral Infection. *Cell Host Microbe* **5**, 559-570 (2009).
225. H. E. Ledvina *et al.*, An E1–E2 fusion protein primes antiviral immune signalling in bacteria. *Nature* **616**, 319-325 (2023).
226. J. M. Jenson, T. Li, F. Du, C.-K. Ea, Z. J. Chen, Ubiquitin-like conjugation by bacterial cGAS enhances anti-phage defence. *Nature* **616**, 326-331 (2023).
227. K. Andryka-Cegielski, S. Soler, E. Bartok, Unexpected bonds: ubiquitin-like conjugation of cGAS/CD-NTases supports their enzymatic activity and antiphage defense. *Signal Transduction and Targeted Therapy* **8**, 308 (2023).
228. Y. Yan *et al.*, Ubiquitin-like cGAS chain formation by a super enzyme activates anti-phage response. *bioRxiv*, 2022.2005.2025.493364 (2023).
229. I. Pettinati, J. Brem, S. Y. Lee, P. J. McHugh, C. J. Schofield, The Chemical Biology of Human Metallo- β -Lactamase Fold Proteins. *Trends in Biochemical Sciences* **41**, 338-355 (2016).
230. M. van Kempen *et al.*, Fast and accurate protein structure search with Foldseek. *Nature Biotechnology*, (2023).
231. J. A. Baker, F. Simkovic, H. M. C. Taylor, D. J. Rigden, Potential DNA binding and nuclease functions of ComEC domains characterized in silico. *Proteins: Structure, Function, and Bioinformatics* **84**, 1431-1442 (2016).
232. L. A. J. O'Neill, A. G. Bowie, The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nature Reviews Immunology* **7**, 353-364 (2007).
233. L. Luo, R. M. Lucas, L. Liu, J. L. Stow, Signalling, sorting and scaffolding adaptors for Toll-like receptors. *Journal of Cell Science* **133**, (2019).
234. M. Ademi, X. Yang, M. P. Coleman, J. Gilley, Natural variants of human SARM1 cause both intrinsic and dominant loss-of-function influencing axon survival. *Scientific Reports* **12**, 13846 (2022).
235. S. Janssens, R. Beyaert, Role of Toll-Like Receptors in Pathogen Recognition. *Clin Microbiol Rev* **16**, 637-646 (2003).
236. D. Sabonis *et al.*, TIR domains produce histidine-ADPR conjugates as immune signaling molecules in bacteria. *bioRxiv*, 2024.2001.2003.573942 (2024).
237. N. Tal *et al.*, Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. *Cell* **184**, 5728-5739.e5716 (2021).
238. F. Rousset, R. Sorek, The evolutionary success of regulated cell death in bacterial immunity. *Current Opinion in Microbiology* **74**, 102312 (2023).
239. B. Duncan-Lowey *et al.*, Cryo-EM structure of the RADAR supramolecular anti-phage defense complex. *Cell*, (2023).
240. Y. Gao *et al.*, Molecular basis of RADAR anti-phage supramolecular assemblies. *Cell*, (2023).
241. G. Ofir *et al.*, DISARM is a widespread bacterial defence system with broad anti-phage activities. *Nature Microbiology* **3**, 90-98 (2018).
242. F. A. Ran *et al.*, Genome engineering using the CRISPR-Cas9 system. *Nature Protocols* **8**, 2281-2308 (2013).
243. C.-H. Huang, K.-C. Lee, J. A. Doudna, Applications of CRISPR-Cas Enzymes in Cancer Therapeutics and Detection. *Trends in Cancer* **4**, 499-512 (2018).
244. B. Guo *et al.*, Pyrococcus furiosus Argonaute with isothermal amplification for fast and ultra-sensitive diagnosis of acute hepatopancreatic necrosis disease in shrimp. *Aquaculture* **575**, 739821 (2023).
245. R. J. Roberts, How restriction enzymes became the workhorses of molecular biology. *Proceedings of the National Academy of Sciences* **102**, 5905-5908 (2005).
246. L. Fu *et al.*, The prokaryotic Argonaute proteins enhance homology sequence-directed recombination in bacteria. *Nucleic acids research* **47**, 3568-3579 (2019).
247. D. A. Jackson, R. H. Symons, P. Berg, Biochemical Method for Inserting New Genetic Information into DNA of Simian Virus 40: Circular SV40 DNA Molecules Containing Lambda

Phage Genes and the Galactose Operon of Escherichia coli. *Proceedings of the National Academy of Sciences* **69**, 2904-2909 (1972).
248. A. Bazin, G. Gautreau, C. Médigue, D. Vallenet, A. Calteau, panRGP: a pangenome-based method to predict genomic islands and explore their diversity. *Bioinformatics* **36**, i651-i658 (2020).
249. G. Gautreau *et al.*, PPanGGOLiN: Depicting microbial diversity via a partitioned pangenome graph. *PLoS Comput Biol* **16**, e1007732 (2020).
250. J. Mistry, R. D. Finn, S. R. Eddy, A. Bateman, M. Punta, Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Research* **41**, e121-e121 (2013).
251. R. C. Edgar, High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *bioRxiv*, 2021.2006.2020.449169 (2022).
252. C. K. Saha, R. Sanches Pires, H. Brolin, M. Delannoy, G. C. Atkinson, FlaGs and webFlaGs: discovering novel biology through the analysis of gene neighbourhood conservation. *Bioinformatics* **37**, 1312-1314 (2020).
253. J. J. Almagro Armenteros *et al.*, SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology* **37**, 420-423 (2019).
254. L. Käll, A. Krogh, E. L. L. Sonnhammer, A Combined Transmembrane Topology and Signal Peptide Prediction Method. *Journal of Molecular Biology* **338**, 1027-1036 (2004).
255. A. Biegert, J. Söding, De novo identification of highly diverged protein repeats by probabilistic consistency. *Bioinformatics* **24**, 807-814 (2008).
256. J. D. Thompson, D. G. Higgins, T. J. Gibson, CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680 (1994).
257. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021).
258. M. Mirdita *et al.*, ColabFold: making protein folding accessible to all. *Nature Methods* **19**, 679-682 (2022).
259. B. MacLean *et al.*, Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966-968 (2010).
260. E. Frirdich, C. Bouwman, E. Vinogradov, C. Whitfield, The Role of Galacturonic Acid in Outer Membrane Stability in Klebsiella pneumoniae. *Journal of Biological Chemistry* **280**, 27604-27612 (2005).
261. E. E. Kulikov, A. K. Golomidova, N. S. Prokhorov, P. A. Ivanov, A. V. Letarov, High-throughput LPS profiling as a tool for revealing of bacteriophage infection strategies. *Scientific Reports* **9**, 2958 (2019).
262. T. Au - Feltwell, M. J. Au - Dorman, D. A. Au - Goulding, J. Au - Parkhill, F. L. Au - Short, Separating Bacteria by Capsule Amount Using a Discontinuous Density Gradient. *JoVE*, e58679 (2019).
263. K. A. Tipton, P. N. Rather, in *Acinetobacter baumannii: Methods and Protocols*, I. Biswas, P. N. Rather, Eds. (Springer New York, New York, NY, 2019), pp. 227-231.
264. P. Jenjaroenpun, T. Wongsurawat, S. P. Yenamandra, V. A. Kuznetsov, QmRLFS-finder: a model, web server and stand-alone tool for prediction and analysis of R-loop forming sequences. *Nucleic Acids Research* **43**, W527-W534 (2015).
265. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
266. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).
267. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2009).
268. K. Blighe, S. Rana, M. Lewis. (2023).
269. H. Thorvaldsdóttir, J. T. Robinson, J. P. Mesirov, Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**, 178-192 (2012).

270. C. L. M. Gilchrist *et al.*, cblaster: a remote search tool for rapid identification and visualization of homologous gene clusters. *Bioinformatics Advances* **1**, (2021).
271. T. U. Consortium, UniProt: a hub for protein information. *Nucleic Acids Research* **43**, D204-D212 (2014).
272. S. Hunter *et al.*, InterPro: the integrative protein signature database. *Nucleic Acids Research* **37**, D211-D215 (2008).
273. W. Shen, S. Le, Y. Li, F. Hu, SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. *PLoS One* **11**, e0163962 (2016).
274. C. Dai *et al.*, QSP: An open sequence database for quorum sensing related gene analysis with an automatic annotation pipeline. *Water Research* **235**, 119814 (2023).
275. R. C. Edgar, Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat Commun* **13**, 6968 (2022).
276. S. Capella-Gutiérrez, J. M. Silla-Martínez, T. Gabaldón, trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**, 1972-1973 (2009).
277. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PLoS One* **5**, e9490 (2010).
278. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* **49**, W293-W296 (2021).
279. I. Letunic, P. Bork, Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research* **47**, W256-W259 (2019).
280. F. d'Hérelle, Sur un microbe invisible antagoniste des bacilles dysentérique. *Acad Sci Paris* **165**, 373-375 (1917).
281. K. Murtazalieva, A. Mu, A. Petrovskaya, R. D. Finn, The growing repertoire of phage anti-defence systems. *Trends Microbiol* **32**, 1212-1228 (2024).
282. H. Li, Y. Tan, D. Basu, Kevin D. Corbett, D. Zhang, Unveiling the multifaceted domain polymorphism of the Menshen antiphage system. *Nucleic Acids Research* **53**, (2025).
283. O. T. Tuck *et al.*, Recurrent acquisition of nuclease-protease pairs in antiviral immunity. *bioRxiv*, 2025.2007.2028.667249 (2025).
284. R. T. Bell *et al.*, YprA family helicases provide the missing link between diverse prokaryotic immune systems. *bioRxiv*, 2025.2009.2015.676423 (2025).
285. J. Garb *et al.*, Multiple phage resistance systems inhibit infection via SIR2-dependent NAD+ depletion. *Nature Microbiology* **7**, 1849-1856 (2022).
286. B. Koopal, S. K. Mutte, D. C. Swarts, A long look at short prokaryotic Argonautes. *Trends in Cell Biology* **33**, 605-618 (2023).
287. S. Wang *et al.*, The role of TIR domain-containing proteins in bacterial defense against phages. *Nat Commun* **15**, 7384 (2024).
288. A. R. Costa *et al.*, Accumulation of defense systems in phage resistant strains of *Pseudomonas aeruginosa*. *bioRxiv*, 2022.2008.2012.503731 (2023).
289. D. F. v. d. Berg, S. J. J. Brouns, Reduced prevalence of phage defense systems in *Pseudomonas aeruginosa* strains from cystic fibrosis patients. *mBio* **16**, e03548-03524 (2025).
290. J. P. K. Bravo, C. Aparicio-Maldonado, F. L. Nobrega, S. J. J. Brouns, D. W. Taylor, Structural basis for broad anti-phage immunity by DISARM. *Nat Commun* **13**, 2987 (2022).
291. Y. Wu *et al.*, Bacterial defense systems exhibit synergistic anti-phage activity. *Cell Host Microbe*, (2024).
292. L. J. Payne, T. C. D. Hughes, P. C. Fineran, S. A. Jackson, New antiviral defences are genetically embedded within prokaryotic immune systems. *bioRxiv*, 2024.2001.2029.577857 (2024).
293. V. Anantharaman, L. M. Iyer, L. Aravind, Ter-dependent stress response systems: novel pathways related to metal sensing, production of a nucleoside-like metabolite, and DNA-processing. *Molecular BioSystems* **8**, 3142-3165 (2012).
294. D. E. Taylor, Bacterial tellurite resistance. *Trends in Microbiology* **7**, 111-115 (1999).
295. O. Alekhina, L. Valkovicova, J. Turna, Study of membrane attachment and in vivo co-localization of TerB protein from uropathogenic Escherichia coli KL53. *General Physiology and Biophysics* **30**, 286–292 (2011).

296. A. R. VanderWal *et al.*, Csx28 is a membrane pore that enhances CRISPR-Cas13b–dependent antiphage defense. *Science* **380**, 410-415 (2023).
297. M. W. Parker, S. C. Feil, Pore-forming protein toxins: from structure to function. *Progress in Biophysics and Molecular Biology* **88**, 91-142 (2005).
298. A. Fagerlund, T. Lindbäck, A. K. Storset, P. E. Granum, S. P. Hardy, Bacillus cereus Nhe is a pore-forming toxin with structural and functional properties similar to the ClyA (HlyE, SheA) family of haemolysins, able to induce osmotic lysis in epithelia. *Microbiology* **154**, 693-704 (2008).
299. M. Ghai, S. Dutta, T. Hall, D. Freilich, C. F. Ockenhouse, Identification, expression, and functional characterization of MAEBL, a sporozoite and asexual blood stage chimeric erythrocyte-binding protein of Plasmodium falciparum. *Molecular and Biochemical Parasitology* **123**, 35-45 (2002).
300. E. G. Armbruster *et al.*, Sequential membrane- and protein-bound organelles compartmentalize genomes during phage infection. *Cell Host Microbe* **33**, 484-497.e486 (2025).
301. A. Bernheim *et al.*, Prokaryotic viperins produce diverse antiviral molecules. *Nature* **589**, 120-124 (2021).
302. E. E. Rivera-Serrano *et al.*, Viperin Reveals Its True Function. *Annual Review of Virology* **7**, 421-446 (2020).
303. T. Barros *et al.*, A structural role for the PHP domain in E. coli DNA polymerase III. *BMC Structural Biology* **13**, 8 (2013).
304. S. Nagpal, D. T. Nair, The PHP domain of PolX from Staphylococcus aureus aids high fidelity DNA synthesis through the removal of misincorporated deoxyribo-, ribo- and oxidized nucleotides. *Scientific Reports* **11**, 4178 (2021).
305. S. C. Sinha, J. L. Smith, The PRT protein family. *Current Opinion in Structural Biology* **11**, 733-739 (2001).
306. F. M. García-Rodríguez, F. Martínez-Abarca, M. E. Wilkinson, N. Toro, Similar mechanisms of retron-mediated anti-phage defense for different families of tailed phages. *bioRxiv*, 2024.2002.2009.579579 (2024).
307. H. Pu *et al.*, A toxin-antitoxin system provides phage defense via DNA damage and repair. *Nat Commun* **16**, 3141 (2025).
308. J. Guan, J. Bondy-Denomy, Intracellular Organization by Jumbo Bacteriophages. *J Bacteriol* **203**, 10.1128/jb.00362-00320 (2020).
309. L. Yuping *et al.*, Jumbo phage killer immune system targets early infection of nucleus-forming phages. *Cell* **188**, 2127-2140.e2121 (2025).
310. M. R. Tock, D. T. F. Dryden, The biology of restriction and anti-restriction. *Current Opinion in Microbiology* **8**, 466-472 (2005).
311. A. F. Alvarez, D. Georgellis, Environmental adaptation and diversification of bacterial two-component systems. *Current Opinion in Microbiology* **76**, 102399 (2023).
312. D. Bellieny-Rabelo, W. J. S. Pretorius, L. N. Moleleki, Novel Two-Component System-Like Elements Reveal Functional Domains Associated with Restriction–Modification Systems and paraMORC ATPases in Bacteria. *Genome Biology and Evolution* **13**, (2021).
313. R. T. Bell *et al.*, Astonishing diversity and multifaceted biological connections of Type IV restriction-modification systems. *bioRxiv*, 2023.2007.2031.551357 (2023).
314. T. Mascher, J. D. Helmann, G. Unden, Stimulus Perception in Bacterial Signal-Transducing Histidine Kinases. *Microbiology and Molecular Biology Reviews* **70**, 910-938 (2006).
315. Y.-C. Lou *et al.*, Structural basis of transcriptional activation by the OmpR/PhoB-family response regulator PmrA. *Nucleic Acids Research* **51**, 10049-10058 (2023).
316. C.-Y. Lin, N. Awano, H. Masuda, J.-H. Park, M. Inouye, Transcriptional Repressor HipB Regulates the Multiple Promoters in Escherichia coli. *Journal of Molecular Microbiology and Biotechnology* **23**, 440-447 (2013).
317. Y. Furuta, I. Kobayashi, in *Madame Curie Bioscience Database*. (Landes Bioscience, Austin (TX), 2011), vol. 2000-2013.
318. G. G. Wilson, N. E. Murray, Restriction and modification systems. *Annual Review of Genetics* **25**, 585-627 (1991).

319. S. A. McMahon *et al.*, Structure and mechanism of a Type III CRISPR defence DNA nuclease activated by cyclic oligoadenylate. *Nat Commun* **11**, 500 (2020).
320. F. Jacob, Evolution and Tinkering. *Science* **196**, 1161-1166 (1977).
321. N. G. A. Kuijpers *et al.*, Pathway swapping: Toward modular engineering of essential cellular processes. *Proceedings of the National Academy of Sciences* **113**, 15060-15065 (2016).
322. S. J. Kierans, C. T. Taylor, Glycolysis: A multifaceted metabolic pathway and signaling hub. *Journal of Biological Chemistry* **300**, 107906 (2024).
323. D. F. van den Berg, B. A. van der Steen, A. R. Costa, S. J. J. Brouns, Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations. *Elife* **12**, e85183 (2023).
324. A. R. Costa *et al.*, Bacteriophage genomes encode both broad and specific counter-defense repertoires to overcome bacterial defense systems. *Cell Host Microbe* **33**, 1161-1172.e1165 (2025).
325. D. F. Van Den Berg, S. J. J. Brouns, Phage tRNAs: decoding the enigma. *Trends in Microbiology*, (2025).
326. D. Mozumdar *et al.*, Characterization of a lipid-based jumbo phage compartment as a hub for early phage infection. *Cell Host Microbe* **32**, 1050-1058.e1057 (2024).
327. T. G. Laughlin *et al.*, Architecture and self-assembly of the jumbo bacteriophage nuclear shell. *Nature* **608**, 429-435 (2022).
328. E. A. Birkholz *et al.*, A cytoskeletal vortex drives phage nucleus rotation during jumbo phage replication in E. coli. *Cell Rep* **40**, 111179 (2022).
329. K. R. Harding, N. Kyte, P. C. Fineran, Jumbo phages. *Current Biology* **33**, R750-R751 (2023).
330. A. Muralidharan *et al.*, Molecular basis for anti-jumbo phage immunity by AVAST Type 5. *bioRxiv*, 2025.2007.2008.663546 (2025).
331. E. L. L. Sonnhammer, S. R. Eddy, R. Durbin, Pfam: A comprehensive database of protein domain families based on seed alignments. *Proteins: Structure, Function, and Bioinformatics* **28**, 405-420 (1997).
332. J. Gough, C. Chothia, SUPERFAMILY: HMMs representing all proteins of known structure. SCOP sequence searches, alignments and genome assignments. *Nucleic Acids Research* **30**, 268-272 (2002).
333. P. Jones *et al.*, InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236-1240 (2014).
334. M. Bastian, S. Heymann, M. Jacomy, paper presented at the International AAAI Conference on Weblogs and Social Media., 2009.
335. M. Yang, M. K. Derbyshire, R. A. Yamashita, A. Marchler-Bauer, NCBI's Conserved Domain Database and Tools for Protein Domain Analysis. *Current Protocols in Bioinformatics* **69**, e90 (2020).
336. S. Lu *et al.*, CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Research* **48**, D265-D268 (2019).
337. V. P. Waman *et al.*, CATH 2024: CATH-AlphaFlow Doubles the Number of Structures in CATH and Reveals Nearly 200 New Folds. *Journal of Molecular Biology* **436**, 168551 (2024).
338. F. Teufel *et al.*, SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology* **40**, 1023-1025 (2022).
339. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
340. L. R. Hmelo *et al.*, Precision-engineering the *Pseudomonas aeruginosa* genome with two-step allelic exchange. *Nature Protocols* **10**, 1820-1841 (2015).
341. J. A. Veech, A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography* **22**, 252-260 (2013).
342. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**, 289-300 (1995).
343. T. U. Consortium, UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Research* **51**, D523-D531 (2022).
344. I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research* **52**, W78-W82 (2024).

345. L. J. Payne *et al.*, PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. *Nucleic Acids Research* **50**, W541-W550 (2022).
346. S. Malhotra, D. Hayes, D. J. Wozniak, Cystic Fibrosis and *Pseudomonas aeruginosa*: the Host-Microbe Interface. *Clin Microbiol Rev* **32**, (2019).
347. P. Bonyadi, N. T. Saleh, M. Dehghani, M. Yamini, K. Amini, Prevalence of antibiotic resistance of *Pseudomonas aeruginosa* in cystic fibrosis infection: A systematic review and meta-analysis. *Microbial Pathogenesis* **165**, 105461 (2022).
348. L. Jackson, V. Waters, Factors influencing the acquisition and eradication of early *Pseudomonas aeruginosa* infection in cystic fibrosis. *Journal of Cystic Fibrosis* **20**, 8-16 (2021).
349. M. Cafora *et al.*, Phage therapy against *Pseudomonas aeruginosa* infections in a cystic fibrosis zebrafish model. *Scientific Reports* **9**, (2019).
350. A. Hahn *et al.*, Bacteriophage Therapy for Pan-Drug-Resistant *Pseudomonas aeruginosa* in Two Persons With Cystic Fibrosis. *Journal of Investigative Medicine High Impact Case Reports* **11**, (2023).
351. J. Singh, E. Yeoh, D. A. Fitzgerald, H. Selvadurai, A systematic review on the use of bacteriophage in treating Staphylococcus aureus and *Pseudomonas aeruginosa* infections in cystic fibrosis. *Paediatric Respiratory Reviews* **48**, 3-9 (2023).
352. M. Bernabéu-Gimeno *et al.*, Neutralizing antibodies after nebulized phage therapy in cystic fibrosis patients. *Med* **5**, 1096-1111.e1096 (2024).
353. J.-P. Pirnay *et al.*, Personalized bacteriophage therapy outcomes for 100 consecutive cases: a multicentre, multinational, retrospective observational study. *Nature Microbiology* **9**, 1434-1453 (2024).
354. G. L. Winsor *et al.*, *Pseudomonas* Genome Database: improved comparative analysis and population genomics capability for *Pseudomonas* genomes. *Nucleic Acids Research* **39**, D596-D600 (2011).
355. A. Deep *et al.*, The SMC-family Wadjet complex protects bacteria from plasmid transformation by recognition and cleavage of closed-circular DNA. *Molecular Cell* **82**, 4145-4159.e4147 (2022).
356. R. Pinilla-Redondo *et al.*, Type IV CRISPR–Cas systems are highly diverse and involved in competition between plasmids. *Nucleic Acids Research* **48**, 2000-2012 (2020).
357. S. L. Taylor *et al.*, The cystic fibrosis gut as a potential source of multidrug resistant pathogens. *Journal of Cystic Fibrosis* **20**, 413-420 (2021).
358. S. Castañeda-Barba, E. M. Top, T. Stalder, Plasmids, a molecular cornerstone of antimicrobial resistance in the One Health era. *Nature Reviews Microbiology* **22**, 18-32 (2024).
359. K.-M. Ling, S. M. Stick, A. Kicic, Pulmonary bacteriophage and cystic fibrosis airway mucus: friends or foes? *Frontiers in Medicine* **10**, (2023).
360. P. Bentkowski, C. Van Oosterhout, T. Mock, A Model of Genome Size Evolution for Prokaryotes in Stable and Fluctuating Environments. *Genome Biology and Evolution* **7**, 2344-2351 (2015).
361. V. Merhej, M. Royer-Carenzi, P. Pontarotti, D. Raoult, Massive comparative genomic analysis reveals convergent evolution of specialized bacteria. *Biol Direct* **4**, 13 (2009).
362. E. Rossi *et al.*, *Pseudomonas aeruginosa* adaptation and evolution in patients with cystic fibrosis. *Nature Reviews Microbiology* **19**, 331-342 (2021).
363. F. Tesson *et al.*, A Comprehensive Resource for Exploring Antiphage Defense: DefenseFinder Webservice,Wiki and Databases. *Peer Community Journal* **4**, (2024).
364. S. Akhter, R. K. Aziz, R. A. Edwards, PhiSpy: a novel algorithm for finding prophages in bacterial genomes that combines similarity- and composition-based strategies. *Nucleic acids research* **40**, e126-e126 (2012).
365. D. Ebbert. (2019).
366. J. Bondy-Denomy, A. Pawluk, K. L. Maxwell, A. R. Davidson, Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429-432 (2013).
367. Y. Li, J. Bondy-Denomy, Anti-CRISPRs go viral: The infection biology of CRISPR-Cas inhibitors. *Cell Host Microbe* **29**, 704-714 (2021).

368. D. Mayo-Muñoz, R. Pinilla-Redondo, N. Birkholz, P. C. Fineran, A host of armor: Prokaryotic immune strategies against mobile genetic elements. *Cell Rep* **42**, 112672 (2023).
369. D. Mayo-Muñoz, R. Pinilla-Redondo, S. Camara-Wilpert, N. Birkholz, P. C. Fineran, Inhibitors of bacterial immune systems: discovery, mechanisms and applications. *Nature Reviews Genetics*, (2024).
370. J. Bobonis *et al.*, Bacterial retrons encode phage-defending tripartite toxin–antitoxin systems. *Nature* **609**, 144-150 (2022).
371. S. Silas *et al.*, Activation of programmed cell death and counter-defense functions of phage accessory genes. *bioRxiv*, 2023.2004.2006.535777 (2023).
372. M. Mahler, A. R. Costa, S. P. B. van Beljouw, P. C. Fineran, S. J. J. Brouns, Approaches for bacteriophage genome engineering. *Trends in Biotechnology* **41**, 669-685 (2023).
373. P.-J. Ceyssens *et al.*, Survey of *Pseudomonas aeruginosa* and its phages: de novo peptide sequencing as a novel tool to assess the diversity of worldwide collected viruses. *Environmental Microbiology* **11**, 1303-1313 (2009).
374. M. Castledine *et al.*, Parallel evolution of *Pseudomonas aeruginosa* phage resistance and virulence loss in response to phage treatment in vivo and in vitro. *Elife* **11**, e73679 (2022).
375. E. A. Ashworth *et al.*, Exploiting lung adaptation and phage steering to clear pan-resistant *Pseudomonas aeruginosa* infections in vivo. *Nat Commun* **15**, 1547 (2024).
376. G. King, N. E. Murray, Restriction alleviation and modification enhancement by the Rac prophage of Escherichia coli K-12. *Molecular Microbiology* **16**, 769-777 (1995).
377. N. Kurochkina, U. Guha, SH3 domains: modules of protein–protein interactions. *Biophysical Reviews* **5**, 29-39 (2013).
378. S. Falk *et al.*, Structural insights into the interaction of the nuclear exosome helicase Mtr4 with the preribosomal protein Nop53. *RNA* **23**, 1780-1787 (2017).
379. P. K. Zuber *et al.*, Structure and nucleic acid binding properties of KOW domains 4 and 6–7 of human transcription elongation factor DSIF. *Scientific Reports* **8**, 11660 (2018).
380. H. Hu *et al.*, Structure and mechanism of Zorya anti-phage defense system. *bioRxiv*, 2023.2012.2018.572097 (2023).
381. E. W. Schmid, J. C. Walter, Predictomes: A classifier-curated database of AlphaFold-modeled protein-protein interactions. *bioRxiv*, 2024.2004.2009.588596 (2024).
382. A. Leavitt *et al.*, Viruses inhibit TIR gcADPR signalling to overcome bacterial defence. *Nature* **611**, 326-331 (2022).
383. E. Yirmiya *et al.*, Phages overcome bacterial immunity via diverse anti-defence proteins. *Nature* **625**, 352-359 (2024).
384. R. S. Kim, E. Levy Karin, M. Mirdita, R. Chikhi, M. Steinegger, BFVD—a large repository of predicted viral protein structures. *Nucleic Acids Research*, (2024).
385. A. A. Kudryavtseva *et al.*, Broadness and specificity: ArdB, ArdA, and Ocr against various restriction-modification systems. *Front Microbiol* **14**, (2023).
386. A. Isaev *et al.*, Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence. *Nucleic Acids Research* **48**, 5397-5406 (2020).
387. S. Jiang *et al.*, A widespread phage-encoded kinase enables evasion of multiple host antiphage defence systems. *Nature Microbiology*, (2024).
388. E. Nurmemmedov, M. Castelnovo, C. E. Catalano, A. Evilevitch, Biophysics of viral infectivity: matching genome length with capsid size. *Quarterly Reviews of Biophysics* **40**, 327-356 (2007).
389. G. Mariano, T. R. Blower, Conserved domains can be found across distinct phage defence systems. *Molecular Microbiology* **120**, 45-53 (2023).
390. D. Li *et al.*, Single phage proteins sequester TIR- and cGAS-generated signaling molecules. *bioRxiv*, 2023.2011.2015.567273 (2023).
391. B. Samuel, K. Mittelman, S. Y. Croitoru, M. Ben Haim, D. Burstein, Diverse anti-defence systems are encoded in the leading region of plasmids. *Nature*, (2024).
392. F. Tesson *et al.*, Exploring the diversity of anti-defense systems across prokaryotes, phages, and mobile genetic elements. *bioRxiv*, 2024.2008.2021.608784 (2024).
393. K. Katoh, K. Misawa, K. i. Kuma, T. Miyata, MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* **30**, 3059-3066 (2002).

394. J. Trifinopoulos, L.-T. Nguyen, A. von Haeseler, B. Q. Minh, W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Research* **44**, W232-W235 (2016).
395. A. Marchler-Bauer, S. H. Bryant, CD-Search: protein domain annotations on the fly. *Nucleic Acids Research* **32**, W327-W331 (2004).
396. M. Necci, D. Piovesan, Z. Dosztányi, S. C. E. Tosatto, MobiDB-lite: fast and highly specific consensus prediction of intrinsic disorder in proteins. *Bioinformatics* **33**, 1402-1404 (2017).
397. J. Lees *et al.*, Gene3D: a domain-based resource for comparative genomics, functional annotation and protein network analysis. *Nucleic Acids Research* **40**, D465-D471 (2011).
398. H. Mi, A. Muruganujan, P. D. Thomas, PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Research* **41**, D377-D386 (2012).
399. C. J. A. Sigrist *et al.*, New and continuing developments at PROSITE. *Nucleic Acids Research* **41**, D344-D347 (2012).
400. Y. Yan, J. Zheng, X. Zhang, Y. Yin, dbAPIS: a database of anti-prokaryotic immune system genes. *Nucleic Acids Research* **52**, D419-D425 (2023).
401. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024).
402. N. Kresge, R. D. Simoni, R. L. Hill, The discovery of tRNA by Paul C. Zamecnik. *Journal of Biological Chemistry* **280**, e37-e39 (2005).
403. F. Crick, Central Dogma of Molecular Biology. *Nature* **227**, 561-563 (1970).
404. S. B. Weiss, W. T. Hsu, J. W. Foft, N. H. Scherberg, Transfer RNA coded by the T4 bacteriophage genome. *Proceedings of the National Academy of Sciences* **61**, 114-121 (1968).
405. M. Bailly-Bechet, M. Vergassola, E. Rocha, Causes for the intriguing presence of tRNAs in phages. *Genome Res* **17**, 1486-1495 (2007).
406. J. B. Lucks, D. R. Nelson, G. R. Kudla, J. B. Plotkin, Genome Landscapes and Bacteriophage Codon Usage. *PLoS Comput Biol* **4**, e1000001 (2008).
407. K. Limor-Waisberg, A. Carmi, A. Scherz, Y. Pilpel, I. Furman, Specialization versus adaptation: two strategies employed by cyanophages to enhance their translation efficiencies. *Nucleic Acids Res* **39**, 6016-6028 (2011).
408. J. Y. Yang *et al.*, Degradation of host translational machinery drives tRNA acquisition in viruses. *Cell Syst* **12**, 771-779 e775 (2021).
409. D. M. Thompson, R. Parker, Stressing Out over tRNA Cleavage. *Cell* **138**, 215-219 (2009).
410. I. Jain *et al.*, tRNA anticodon cleavage by target-activated CRISPR-Cas13a effector. *Science Advances* **10**, (2024).
411. M. Amitsur, I. Morad, G. Kaufmann, In vitro reconstitution of anticodon nuclease from components encoded by phage T4 and Escherichia coli CTr5X. *EMBO J* **8**, 2411-2415 (1989).
412. M. Wolfram-Schauerte, N. Pozhydaieva, M. Viering, T. Glatter, K. Hofer, Integrated Omics Reveal Time-Resolved Insights into T4 Phage Infection of E. coli on Proteome and Transcriptome Levels. *Viruses* **14**, 2502 (2022).
413. G. Kaufmann *et al.*, in *DNA Repair*. (IntechOpen, 2011).
414. D. A. Russell, G. F. Hatfull, PhagesDB: the actinobacteriophage database. *Bioinformatics* **33**, 784-786 (2017).
415. K. Winther, J. J. Tree, D. Tollervey, K. Gerdes, VapCs of Mycobacterium tuberculosis cleave RNAs essential for translation. *Nucleic Acids Research* **44**, 9860-9871 (2016).
416. U. Chauhan, V. C. Barth, N. A. Woychik, tRNA fMet Inactivating Mycobacterium tuberculosis VapBC Toxin-Antitoxin Systems as Therapeutic Targets. *Antimicrob Agents Chemother* **66**, (2022).
417. J. W. Cruz *et al.*, Growth-regulating Mycobacterium tuberculosis VapC-mt4 toxin is an isoacceptor-specific tRNase. *Nat Commun* **6**, 7480 (2015).
418. M. Cintrón *et al.*, Accurate target identification for Mycobacterium tuberculosis endoribonuclease toxins requires expression in their native host. *Scientific Reports* **9**, (2019).

419. V. C. Barth *et al.*, Mycobacterium tuberculosis VapC4 toxin engages small ORFs to initiate an integrated oxidative and copper stress response. *Proceedings of the National Academy of Sciences* **118**, e2022136118 (2021).
420. K. Pedersen *et al.*, The Bacterial Toxin RelE Displays Codon-Specific Cleavage of mRNAs in the Ribosomal A Site. *Cell* **112**, 131-140 (2003).
421. F. Calcuttawala *et al.*, Apoptosis like symptoms associated with abortive infection of Mycobacterium smegmatis by mycobacteriophage D29. *PLoS One* **17**, e0259480 (2022).
422. K. S. Winther, K. Gerdes, Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. *Proceedings of the National Academy of Sciences* **108**, 7403-7407 (2011).
423. M. Amitsur, R. Levitz, G. Kaufmann, Bacteriophage T4 anticodon nuclease, polynucleotide kinase and RNA ligase reprocess the host lysine tRNA. *EMBO J* **6**, 2499-2503 (1987).
424. T. Ogawa, S. Inoue, S. Yajima, M. Hidaka, H. Masaki, Sequence-specific recognition of colicin E5, a tRNA-targeting ribonuclease. *Nucleic Acids Research* **34**, 6065-6073 (2006).
425. D. Cavard, C. Lazdunski, Interaction of Colicin E4 with Specific Receptor Sites Mediates Its Cleavage into Two Fragments Inactive Towards Whole Cells. *European Journal of Biochemistry* **96**, 525-533 (1979).
426. A. M. Jones, F. Garza-Sánchez, J. So, C. S. Hayes, D. A. Low, Activation of contact-dependent antibacterial tRNase toxins by translation elongation factors. *Proceedings of the National Academy of Sciences* **114**, E1951-E1957 (2017).
427. D. Laslett, ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Research* **32**, 11-16 (2004).
428. Patricia P. Chan, Brian Y. Lin, Allysia J. Mak, Todd M. Lowe, tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Research* **49**, 9077-9096 (2021).
429. P. Rice, I. Longden, A. Bleasby, EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics* **16**, 276-277 (2000).
430. A. P. Y. Lopes *et al.*, VapC from the Leptospiral VapBC Toxin-Antitoxin Module Displays Ribonuclease Activity on the Initiator tRNA. *PLoS One* **9**, e101678 (2014).
431. T. Miyamoto *et al.*, Characterization of a Deinococcus radiodurans MazF: A UACA-specific RNA endoribonuclease. *MicrobiologyOpen* **6**, e00501 (2017).
432. E. Davidov, G. Kaufmann, RloC: a wobble nucleotide-excising and zinc-responsive bacterial tRNase. *Molecular Microbiology* **69**, 1560-1574 (2008).
433. B. Y. Lin, P. P. Chan, T. M. Lowe, tRNAviz: explore and visualize tRNA sequence features. *Nucleic Acids Research* **47**, W542-W547 (2019).
434. A. D. Hershey, M. Chase, Independent functions of viral protein and nucleic acid in growth of bacteriophage. *Journal of General Physiology* **36**, 39-56 (1952).
435. N. H. Scherberg, A. Guha, W. T. Hsu, S. B. Weiss, Evidence for the early synthesis of T4 bacteriophage-coded transfer RNA. *Biochemical and Biophysical Research Communications* **40**, 919-924 (1970).
436. J. H. Wilson, Function of the bacteriophage T4 transfer RNA's. *J Mol Biol* **74**, 753-757 (1973).
437. R. B. Cook, Nathan; Redgwell, Tamsin; Rihtman, Branko; Barnes, Megan; Clokie, Martha; Stekel, Dov J.; Hobman, Jon ; Jones, Michael A.; Millard, Andrew, INfrastructure for a PHAge REference Database: Identification of Large-Scale Biases in the Current Collection of Cultured Phage Genomes. *PHAGE* **2**, 214-223 (2021).
438. H. Dong, L. Nilsson, C. G. Kurland, Co-variation of tRNA Abundance and Codon Usage inEscherichia coliat Different Growth Rates. *Journal of Molecular Biology* **260**, 649-663 (1996).
439. Z. Li, M. P. Deutscher, Maturation Pathways for E. coli tRNA Precursors: A Random Multienzyme Process In Vivo. *Cell* **86**, 503-512 (1996).
440. K. Tomita, S. Yamashita, Molecular mechanisms of template-independent RNA polymerization by tRNA nucleotidyltransferases. *Front Genet* **5**, (2014).
441. F. H. C. Crick, The origin of the genetic code. *Journal of Molecular Biology* **38**, 367-379 (1968).
442. T. Suzuki, The expanding world of tRNA modifications and their disease relevance. *Nature Reviews Molecular Cell Biology* **22**, 375-392 (2021).

443. K. L. Harvey, V. M. Jarocki, I. G. Charles, S. P. Djordjevic, The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Front Microbiol* **10**, (2019).
444. M. W. Nirenberg, J. H. Matthaei, The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. *Proceedings of the National Academy of Sciences* **47**, 1588-1602 (1961).
445. C. A. Hutchison *et al.*, Design and synthesis of a minimal bacterial genome. *Science* **351**, aad6253 (2016).
446. G. Hutinet *et al.*, 7-Deazaguanine modifications protect phage DNA from host restriction systems. *Nat Commun* **10**, (2019).
447. G. Hutinet, M. A. Swarjo, V. de Crécy-Lagard, Deazaguanine derivatives, examples of crosstalk between RNA and DNA modification pathways. *RNA Biol* **14**, 1175-1184 (2017).
448. M. Ibba, SöLl, Dieter, Quality Control Mechanisms During Translation. *Science* **286**, 1893-1897 (1999).
449. A. Carbone, Codon Bias is a Major Factor Explaining Phage Evolution in Translationally Biased Hosts. *Journal of Molecular Evolution* **66**, 210-223 (2008).
450. E. Pfeifer, J. A. Moura De Sousa, M. Touchon, E. P. C. Rocha, Bacteria have numerous distinctive groups of phage–plasmids with conserved phage and variable plasmid gene repertoires. *Nucleic Acids Research* **49**, 2655-2673 (2021).
451. T. T. T. Tran, H. Belahbib, V. Bonnefoy, E. Talla, A Comprehensive tRNA Genomic Survey Unravels the Evolutionary History of tRNA Arrays in Prokaryotes. *Genome Biology and Evolution* **8**, 282-295 (2016).
452. T. Hackl *et al.*, Novel integrative elements and genomic plasticity in ocean ecosystems. *Cell* **186**, 47-62 e16 (2023).
453. V. A. Delesalle, N. T. Tanke, A. C. Vill, G. P. Krukonis, Testing hypotheses for the presence of tRNA genes in mycobacteriophage genomes. *Bacteriophage* **6**, e1219441 (2016).
454. S. Zhao, K. P. Williams, Integrative Genetic Element That Reverses the Usual Target Gene Orientation. *J Bacteriol* **184**, 859-860 (2002).
455. K. P. Williams, Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. *Nucleic Acids Research* **30**, 866-875 (2002).
456. R. M. B. Silva *et al.*, Modification of DNA by a viral enzyme and charged tRNA. *bioRxiv*, (2023).
457. K. Kienbeck *et al.*, Identification of HDV-like theta ribozymes involved in tRNA-based recoding of gut bacteriophages. *Nat Commun* **15**, (2024).
458. A. L. Borges *et al.*, Widespread stop-codon recoding in bacteriophages may regulate translation of lytic genes. *Nature Microbiology* **7**, 918-927 (2022).
459. N. Yutin *et al.*, Analysis of metagenome-assembled viral genomes from the human gut reveals diverse putative CrAss-like phages with unique genomic features. *Nat Commun* **12**, (2021).
460. F. W. Stahl, The amber mutants of phage T4. *Genetics* **141**, 439-442 (1995).
461. S. L. Peters *et al.*, Experimental validation that human microbiome phages use alternative genetic coding. *Nature Communications* **13**, (2022).
462. K. Foss, S.-H. Kao, W. H. McClain, Three suppressor forms of bacteriophage T4 leucine transfer RNA. *Journal of Molecular Biology* **135**, 1013-1021 (1979).
463. A. Pfennig, A. Lomsadze, M. Borodovsky, MgCod: Gene Prediction in Phage Genomes with Multiple Genetic Codes. *Journal of Molecular Biology* **435**, 168159 (2023).
464. J. Elf, D. Nilsson, T. Tenson, M. N. Ehrenberg, Selective Charging of tRNA Isoacceptors Explains Patterns of Codon Usage. *Science* **300**, 1718-1722 (2003).
465. L. E. Leiva *et al.*, Oxidative stress strongly restricts the effect of codon choice on the efficiency of protein synthesis in Escherichia coli. *Front Microbiol* **13**, (2022).
466. Y. Wei, J. R. Silke, X. Xia, An improved estimation of tRNA expression to better elucidate the coevolution between tRNA abundance and codon usage in bacteria. *Scientific Reports* **9**, (2019).
467. H. Enav, O. Béjà, Y. Mandel-Gutfreund, Cyanophage tRNAs may have a role in cross-infectivity of oceanic Prochlorococcus and Synechococcus hosts. *ISME Journal* **6**, 619-628 (2012).
468. T. Kunisawa, Synonymous codon preferences in bacteriophage T4: A distinctive use of transfer RNAs from T4 and from its host Escherichia coli. *Journal of Theoretical Biology* **159**, 287-298 (1992).

469. E. S. Miller *et al.*, Complete genome sequence of the broad-host-range vibriophage KVP40: Comparative genomics of a T4-related bacteriophage. *J Bacteriol* **185**, 5220-5233 (2003).
470. T. W. Tillack, D. W. Smith, The effect of bacteriophage T2 infection of the synthesis of transfer RNA in Escherichia coli. *Virology* **36**, 212-222 (1968).
471. B. Son, J. Patterson-West, C. O. Thompson, J. R. Iben, D. M. Hinton, Setting Up a Better Infection: Overexpression of the Early Bacteriophage T4 Gene motB during Infection Results in a More Favorable tRNA Pool for the Phage. *PHAGE: Therapy, Applications, and Research* **3**, 141-152 (2022).
472. S. Zborowsky, R. Tahan, D. Lindell, Adaptive loss of tRNA gene expression leads to phage resistance in a marine Synechococcus cyanobacterium. *Nature Microbiology* **10**, 66-76 (2025).
473. M. Penner, I. Morad, L. Snyder, G. Kaufmann, Phage T4-coded Stp: Double-Edged Effector of Coupled DNA and tRNA-Restriction Systems. *Journal of Molecular Biology* **249**, 857-868 (1995).
474. S. Uppalapati *et al.*, Prophage terminase with tRNase activity sensitizes Salmonella enterica to oxidative stress. *Science* **384**, 100-105 (2024).
475. M. David, G. D. Borasio, G. Kaufmann, T4 bacteriophage-coded polynucleotide kinase and RNA ligase are involved in host tRNA alteration or repair. *Virology* **123**, 480-483 (1982).
476. J. T. George *et al.*, Structural basis of antiphage defense by an ATPase-associated reverse transcriptase. *bioRxiv*, (2025).
477. M. R. Sargen, S. Helaine, A prophage competition element protects Salmonella from lysis. *Cell Host Microbe* **32**, 2063-2079. e2068 (2024).
478. D. Chapman *et al.*, Nucleotide and deduced amino acid sequence of stp: The bacteriophage T4 anticodon nuclease gene. *Journal of Molecular Biology* **199**, 373-377 (1988).
479. A. Deep, Q. Liang, E. Enustun, J. Pogliano, K. D. Corbett, Architecture and activation mechanism of the bacterial PARIS defence system. *Nature* **634**, 432-439 (2024).
480. F. Rousset *et al.*, A conserved family of immune effectors cleaves cellular ATP upon viral infection. *Cell* **186**, 3619-3631.e3613 (2023).
481. M.-C. Chopin, A. Chopin, E. Bidnenko, Phage abortive infection in lactococci: variations on a theme. *Current Opinion in Microbiology* **8**, 473-479 (2005).
482. H. Myung, R. Calendar, The old exonuclease of bacteriophage P2. *J Bacteriol* **177**, 497-501 (1995).
483. A. Maestri *et al.*, The bacterial defense system MADS interacts with CRISPR-Cas to limit phage infection and escape. *Cell Host & Microbe* **32**, 1412-1426.e1411 (2024).
484. H. Altae-Tran *et al.*, Uncovering the functional diversity of rare CRISPR-Cas systems with deep terascale clustering. *Science* **382**, (2023).
485. M. Saravanan, J. M. Bujnicki, I. A. Cymerman, D. N. Rao, V. Nagaraja, Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. *Nucleic Acids Res* **32**, 6129-6135 (2004).
486. R. T. Bell, H. Sahakyan, K. S. Makarova, Y. I. Wolf, E. V. Koonin, CoCoNuTs are a diverse subclass of Type IV restriction systems predicted to target RNA. *Elife* **13**, (2024).
487. M. Tangney, G. F. Fitzgerald, AbiA, a Lactococcal Abortive Infection Mechanism Functioning in Streptococcus thermophilus. *Applied and Environmental Microbiology* **68**, 6388-6391 (2002).
488. M. Gapińska *et al.*, Structure-functional characterization of Lactococcus AbiA phage defense system. *Nucleic Acids Research* **52**, 4723-4738 (2024).
489. J. Anba, E. Bidnenko, A. Hillier, D. Ehrlich, M. C. Chopin, Characterization of the lactococcal abiD1 gene coding for phage abortive infection. *J Bacteriol* **177**, 3818-3823 (1995).
490. J. Haaber, S. Moineau, L.-C. Fortier, K. Hammer, AbiV, a Novel Antiphage Abortive Infection Mechanism on the Chromosome of Lactococcus lactis subsp. cremoris MG1363. *Applied and Environmental Microbiology* **74**, 6528-6537 (2008).
491. M. Koga, Y. Otsuka, S. b. Lemire, T. Yonesaki, Escherichia coli rnlA and rnlB Compose a Novel Toxin–Antitoxin System. *Genetics* **187**, 123-130 (2011).
492. Charlotte, Amanda, S. Tachdjian, Paul, Robert, Role of vapBC toxin–antitoxin loci in the thermal stress response of Sulfolobus solfataricus. *Biochemical Society Transactions* **37**, 123-126 (2009).

493. B. Liu *et al.*, Deciphering the tRNA-derived small RNAs: origin, development, and future. *Cell Death & Disease* **13**, (2021).
494. G. S. Kushwaha, G. Bange, N. S. Bhavesh, Interaction studies on bacterial stringent response protein RelA with uncharged tRNA provide evidence for its prerequisite complex for ribosome binding. *Current Genetics* **65**, 1173-1184 (2019).
495. A. Biela *et al.*, The diverse structural modes of tRNA binding and recognition. *Journal of Biological Chemistry* **299**, 104966 (2023).
496. Y.-T. Liu, D. Strugatsky, W. Liu, Z. H. Zhou, Structure of human cytomegalovirus virion reveals host tRNA binding to capsid-associated tegument protein pp150. *Nature Communications* **12**, (2021).
497. T. Kurata *et al.*, Toxic small alarmone synthetase FaRel2 inhibits translation by pyrophosphorylating tRNAGly and tRNAThr. *Science Advances* **10**, (2024).
498. X. Xu *et al.*, MenT nucleotidyltransferase toxins extend tRNA acceptor stems and can be inhibited by asymmetrical antitoxin binding. *Nat Commun* **14**, (2023).
499. R. L. Dy, R. Przybilski, K. Semeijn, G. P. C. Salmond, P. C. Fineran, A widespread bacteriophage abortive infection system functions through a Type IV toxin-antitoxin mechanism. *Nucleic acids research* **42**, 4590-4605 (2014).
500. B. Al-Shayeb *et al.*, Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425-431 (2020).
501. F. Harada, S. Nishimura, Possible anticodon sequences of tRNAHis, tRNAAsn, and tRNAAsp from Escherichia coli. Universal presence of nucleoside O in the first position of the anticodons of these transfer ribonucleic acid. *Biochemistry* **11**, 301-308 (1972).
502. J. Díaz-Rullo, J. E. González-Pastor, tRNA queuosine modification is involved in biofilm formation and virulence in bacteria. *Nucleic Acids Research* **51**, 9821-9837 (2023).
503. T. Kawamura *et al.*, Multisite-specific archaeosine tRNA-guanine transglycosylase (ArcTGT) from Thermoplasma acidophilum, a thermo-acidophilic archaeon. *Nucleic Acids Research* **44**, 1894-1908 (2016).
504. E. Kulikov *et al.*, Genomic Sequencing and Biological Characteristics of a Novel Escherichia Coli Bacteriophage 9g, a Putative Representative of a New Siphoviridae Genus. *Viruses* **6**, 5077-5092 (2014).
505. M. C. Lucas *et al.*, Quantitative analysis of tRNA abundance and modifications by nanopore RNA sequencing. *Nature Biotechnology* **42**, 72-86 (2024).
506. N. H. Padhiar, U. Katneni, A. A. Komar, Y. Motorin, C. Kimchi-Sarfaty, Advances in methods for tRNA sequencing and quantification. *Trends in Genetics* **40**, 276-290 (2024).
507. Y. Sun, M. Piechotta, Isabel, C. Dieterich, Ann, Detection of queuosine and queuosine precursors in tRNAs by direct RNA sequencing. *Nucleic Acids Research* **51**, 11197-11212 (2023).
508. M. Skurnik *et al.*, Phage therapy. *Nature Reviews Methods Primers* **5**, (2025).
509. K. A. Burke, C. D. Urick, N. Mzhavia, M. P. Nikolich, A. A. Filippov, Correlation of *Pseudomonas aeruginosa* Phage Resistance with the Numbers and Types of Antiphage Systems. *International Journal of Molecular Sciences* **25**, 1424 (2024).
510. B. Gaborieau *et al.*, Prediction of strain level phage–host interactions across the Escherichia genus using only genomic information. *Nature Microbiology* **9**, 2847-2861 (2024).
511. S. Nasrin *et al.*, Distribution of serotypes and antibiotic resistance of invasive *Pseudomonas aeruginosa* in a multi-country collection. *BMC Microbiol* **22**, (2022).
512. J. Xu *et al.*, O Antigen Is the Receptor of Vibrio cholerae Serogroup O1 El Tor Typing Phage VP4. *J Bacteriol* **195**, 798-806 (2013).
513. B. Liu *et al.*, Structure and genetics ofEscherichia coliO antigens. *FEMS Microbiol Rev* **44**, 655-683 (2020).
514. N. K. Broeker, S. Barbirz, Not a barrier but a key: How bacteriophages exploit host's O-antigen as an essential receptor to initiate infection. *Molecular Microbiology* **105**, 353-357 (2017).
515. D. F. van den Berg *et al.*, Discovery of phage defense systems through component modularity networks. *bioRxiv*, (2025).

516. W.-X. Yee *et al.*, END nucleases: Antiphage defense systems targeting multiple hypermodified phage genomes. *bioRxiv*, (2025).
517. A. K. Cain *et al.*, A decade of advances in transposon-insertion sequencing. *Nature Reviews Genetics* **21**, 526-540 (2020).
518. H. Morisaka *et al.*, CRISPR-Cas3 induces broad and unidirectional genome editing in human cells. *Nat Commun* **10**, (2019).
519. B. Darracq *et al.*, Sedentary chromosomal integrons as biobanks of bacterial anti-phage defence systems. *bioRxiv*, (2024).
520. E. Mordret *et al.*, Protein and genomic language models chart a vast landscape of antiphage defenses. *bioRxiv*, (2025).
521. K. S. Makarova *et al.*, An updated evolutionary classification of CRISPR–Cas systems. *Nature Reviews Microbiology* **13**, 722-736 (2015).
522. R. J. Roberts, T. Vincze, J. Posfai, D. Macelis, REBASE: a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Research* **51**, D629-D630 (2023).
523. R. S. Eriksen, N. Malhotra, A. S. N. Seshasayee, K. Sneppen, S. Krishna, Emergence of networks of shared restriction-modification systems in phage–bacteria ecosystems. *Journal of Biosciences* **47**, (2022).
524. Edward *et al.*, A model for the evolution of prokaryotic DNA restriction-modification systems based upon the structural malleability of Type I restriction-modification enzymes. *Nucleic Acids Research* **46**, 9067-9080 (2018).
525. S. Blanga-Kanfi, M. Amitsur, A. Azem, G. Kaufmann, PrrC-anticodon nuclease: functional organization of a prototypical bacterial restriction RNase. *Nucleic Acids Research* **34**, 3209-3219 (2006).
526. Y. Wu *et al.*, Bacterial defense systems exhibit synergistic anti-phage activity. *Cell Host & Microbe* **32**, 557-572.e556 (2024).
527. T. Wein *et al.*, CARD domains mediate anti-phage defence in bacterial gasdermin systems. *Nature* **639**, 727-734 (2025).
528. D. Bonhomme *et al.*, A human homolog of SIR2 antiphage proteins mediates immunity via the TLR pathway. *bioRxiv*, (2024).
529. A. Bernheim, J. Cury, E. Z. Poirier, The immune modules conserved across the tree of life: Towards a definition of ancestral immunity. *PLOS Biology* **22**, e3002717 (2024).
530. J. Cury *et al.*, Conservation of antiviral systems across domains of life reveals immune genes in humans. *Cell Host & Microbe* **32**, 1594-1607.e1595 (2024).
531. P. Leão *et al.*, Asgard archaea defense systems and their roles in the origin of eukaryotic immunity. *Nat Commun* **15**, (2024).
532. C. Bastiaanssen *et al.*, RNA-guided RNA silencing by an Asgard archaeal Argonaute. *Nat Commun* **15**, (2024).
533. H. Onallah *et al.*, Refractory *Pseudomonas aeruginosa* infections treated with phage PASA16: A compassionate use case series. *Med* **4**, 600-611.e604 (2023).
534. C. S. McVay, M. VeláSquez, J. A. Fralick, Phage Therapy of *Pseudomonas aeruginosa* Infection in a Mouse Burn Wound Model. *Antimicrob Agents Chemother* **51**, 1934-1938 (2007).
535. D. F. Van Den Berg, S. J. J. Brouns, Reduced prevalence of phage defense systems in *Pseudomonas aeruginosa* strains from cystic fibrosis patients. *mBio* **16**, (2025).
536. I. Weiner *et al.*, Phage therapy with nebulized cocktail BX004-A for chronic *Pseudomonas aeruginosa* infections in cystic fibrosis: a randomized first-in-human trial. *Nat Commun* **16**, 5579 (2025).
537. E. Yirmiya *et al.*, Structure-guided discovery of viral proteins that inhibit host immunity. *Cell* **188**, 1681-1692.e1617 (2025).
538. E. Yirmiya *et al.*, Phages overcome bacterial immunity via diverse anti-defence proteins. *Nature*, (2023).
539. B. Samuel, K. Mittelman, S. Y. Croitoru, M. Ben Haim, D. Burstein, Diverse anti-defence systems are encoded in the leading region of plasmids. *Nature* **635**, 186-192 (2024).
540. A. Lopatina *et al.*, Interpretable machine learning reveals a diverse arsenal of anti-defenses in bacterial viruses. *bioRxiv*, 2024.2006.2014.598830 (2024).

541. M. Pacesa *et al.*, One-shot design of functional protein binders with BindCraft. *Nature*, (2025).
542. B. A. Adler *et al.*, Broad-spectrum CRISPR-Cas13a enables efficient phage genome editing. *Nature Microbiology* **7**, 1967-1979 (2022).
543. B. A. Adler *et al.*, Genome-wide Characterization of Diverse Bacteriophages Enabled by RNA-Binding CRISPRi. *bioRxiv*, (2023).
544. M. Gerovac *et al.*, A grad-seq view of RNA and protein complexes in *Pseudomonas aeruginosa* under standard and bacteriophage predation conditions. *mBio* **12**, 1-26 (2021).
545. S. Ojima *et al.*, Systematic Discovery of Phage Genes that Inactivate Bacterial Immune Systems. *bioRxiv*, 2024.2004.2014.589459 (2024).
546. S. Jiang *et al.*, A widespread phage-encoded kinase enables evasion of multiple host antiphage defence systems. *Nature Microbiology* **9**, 3226-3239 (2024).
547. T. Bartolec *et al.*, Pervasive phosphorylation by phage T7 kinase disarms bacterial defenses. *bioRxiv*, (2024).
548. S. P. Antine *et al.*, Structural basis of Gabija anti-phage defense and viral immune evasion. *bioRxiv*, 2023.2005.2001.538945 (2023).
549. R. B. Chang *et al.*, A widespread family of viral sponge proteins reveals specific inhibition of nucleotide signals in anti-phage defense. *bioRxiv*, (2024).
550. C. L. Loyo, A. D. Grossman, A phage-encoded counter-defense inhibits an NAD-degrading anti-phage defense system. *PLOS Genetics* **21**, e1011551 (2025).
551. D. F. van den Berg, B. A. van der Steen, A. R. Costa, S. J. J. Brouns, Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations. *Elife* **12**, (2023).
552. A. H. Azam *et al.*, Evasion of antiviral bacterial immunity by phage tRNAs. *Nat Commun* **15**, (2024).
553. M. Blazanin, P. E. Turner, Community context matters for bacteria-phage ecology and evolution. *ISME J* **15**, 3119-3128 (2021).
554. M. Zhang, T. Zhang, M. Yu, Y.-L. Chen, M. Jin, The Life Cycle Transitions of Temperate Phages: Regulating Factors and Potential Ecological Implications. *Viruses* **14**, 1904 (2022).
555. Edze *et al.*, Parasite Exposure Drives Selective Evolution of Constitutive versus Inducible Defense. *Current Biology* **25**, 1043-1049 (2015).
556. L. Kever *et al.*, Aminoglycoside Antibiotics Inhibit Phage Infection by Blocking an Early Step of the Infection Cycle. *mBio* **13**, e0078322 (2022).
557. Z. Zang *et al.*, Streptomyces secretes a siderophore that sensitizes competitor bacteria to phage infection. *Nature Microbiology* **10**, 362-373 (2025).
558. M. Middelboe *et al.*, Effects of Bacteriophages on the Population Dynamics of Four Strains of Pelagic Marine Bacteria. *Microbial Ecology* **42**, 395-406 (2001).
559. S. Meaden, E. Westra, P. Fineran, Phage defence system abundances vary across environments and increase with viral density. *bioRxiv*, (2025).
560. A. Stern, L. Keren, O. Wurtzel, G. Amitai, R. Sorek, Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet* **26**, 335-340 (2010).
561. N. C. Keith, R. A. Snyder, C. W. Euler, J. W. Modell, Bacteria exploit viral dormancy to establish CRISPR-Cas immunity. *Cell Host & Microbe* **33**, 330-340.e336 (2025).
562. S. Carpenter, L. A. J. O'Neill, From periphery to center stage: 50 years of advancements in innate immunity. *Cell* **187**, 2030-2051 (2024).
563. K. Bartke, D. L. Huseby, G. Brandis, D. Hughes, Evolution of bacterial interspecies hybrids with enlarged chromosomes. *Genome Biology and Evolution*, (2022).
564. R. C. L. Olsthoorn, G. Garde, T. Dayhuff, J. F. Atkins, J. Van Duin, Nucleotide sequence of a single-stranded RNA phage from *Pseudomonas aeruginosa*:Kinship to coliphages and conservation of regulatory RNA structures. *Virology* **206**, 611-625 (1995).
565. C. K. Stover *et al.*, Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen. *Nature* **406**, 959-964 (2000).
566. Y. Yuan, M. Gao, Jumbo Bacteriophages: An Overview. *Front Microbiol* **8**, (2017).
567. J. R. Penadés, K. D. Seed, J. Chen, D. Bikard, E. P. C. Rocha, Genetics, ecology and evolution of phage satellites. *Nature Reviews Microbiology*, (2025).

568. Y. I. Wolf, E. V. Koonin, On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. *Biol Direct* **2**, 14 (2007).
569. H. S. Bernhardt, The RNA world hypothesis: the worst theory of the early evolution of life (except for all the others)a. *Biol Direct* **7**, 23 (2012).

Acknowledgements

Before it is time to take some rest, it is time to reflect a bit on this chapter of life. What a journey it has been. A journey from a know-it-all youngster to a clueless dad. And what was I lucky to have so many of you there with me during this time.

Stan, dankjewel voor de kans die je mij destijds hebt gegeven om mezelf te kunnen bewijzen. Ik weet nog dat ik het vertikte om een PhD te beginnen, maar van deze kans kon ik echt niet vanaf zien, en daar heb ik geen moment spijt van gehad. Jij hebt vast zelf vaak gedacht: "Waar ben ik aan begonnen?" Maar dat waren voor mij juist de momenten waar ik het meest van heb geleerd. Vooral je altijd-aandachtig-luisterende oor zonder te veel vooropgestelde oordelen maakte het fijn om mijn verhaal te vertellen, van klein tot groot. Niet alleen over wetenschappelijke onderwerpen zoals hoe leuk of saai CRISPR-Cas wel niet is, maar ook over de grote lijnen van het leven en wat nou echt belangrijk is. Ik ben blij dat ik nog iets langer mag blijven en hoop op nog veel intrigerende kleurrijke discussies.

Daan (de lange), vanaf het congres in Isräel wist ik al dat het goed zou komen met mijn zoektocht naar niet alleen een co-promotor, maar ook iemand om nieuwe ideeën mee te bespreken. Een wandelende Argonaute encyclopedie met een vrijwel oneindige creativiteit. Ik was altijd al onder de indruk van jouw ontdekkingen in het veld en het opzetten van je eigen lab. Tijdens onze meetings slaagde je er altijd in om de meest kritische vragen te stellen over precies de gedeeltes waar ik onzeker over was. Alhoewel dat op dat moment erg lastig vond, heeft dat achteraf enorm geholpen. Dankjewel dat jij mijn co-promotor wilde zijn en hoop nog veel van je te kunnen afkijken. Tot slot wil ik je ook bedanken dat ik door jou nog kleiner uitzie dan ik al ben.

Rita, you have no idea. I have thought a lot about how to thank you, but I find it difficult to find the words, especially without becoming too emotional. During this time, you have been there walking this journey with me, all the time, in all ways. What a mentor and more importantly what a friend you have been. I could not have imagined how it would have been without you, for sure less enjoyable and less meaningful. I cannot thank you enough and hope that one day I could inspire someone just as much as you have done to me. Always thinking about how to help others, just promise me that you won't forget to also think about yourself, at least sometimes.

Jelger, mijn pizza en retreat buddy, en de enige normale persoon met een privé leven. Je bent altijd degene geweest die ons liet zien dat waar je ook mee bezig bent in het lab, dat het net zo belangrijk (zo niet belangrijker) is om plezier te hebben en te vinden buiten je werk om. En bovendien, om jezelf elke dag te vragen of het al vrijdag is. Wat ben jij een legende. Ik hoop jou voorlopig niet uit het oog te verliezen, makker maatje.

Rita and Jelger, what a team we were. To the truest sense: don't cry because it is over, be happy that it happened. I could spend many pages on how special my time with you guys was. I am excited to see what the next chapters of our lives will bring, but I am very certain that I will miss this one deeply. *Hakuna Matata, out.* You guys are amazing.

Baltus, jij laat mij altijd zien wat een echte vriend is. Iets waar ik jaloers op ben en elke dag naar streef. Het is een begin dat jouw verjaardag de enige verjaardag is die ik uit mijn hoofd weet. Beste verjaardagsgenoot, wie had dat gedacht bij onze eerste ontmoeting, en wat hebben we gelachen. Laat me weten wanneer je je eigen lab begint, hopelijk een beetje in de buurt van het oosten. Wat je ook gaat doen, ik ben nu al trots op waar je nu staat. **Halewijn**, frisbee partner en perfecte persoon om mijn vlaamse accent mee te oefenen. En natuurlijk het erin wrijven dat brak zijn voor jou geen begrip is. Ook liet je mij zien dat wat mensen een snelle leerling noemen, jij vier keer zo snel kan. Wat een bevoorrecht om jou te hebben mogen begeleiden en nu mee samen te werken. Volg je nieuwsgierigheid en je zult je nooit vervelen.

Ilinka and Jasper, the small-people club, whenever I joked that you guys should go back to Martijs if you guys stopped enjoying the internship, Ilinka always sharply replied: "I don't think the warranty covers damaged goods". Honestly, I couldn't agree more, since it seems to be a good summary of your combined experience with me as your supervisor. Something that **Carmen** would probably also agree on. Regardless, that we both have a connection to a small village in Spain still blows my mind, what are the odds? All three of you worked your butts off and the results showed! It does me good to see you guys do so well.

Sam, they say that there is a very tiny line between genius and madness, somehow you seem to have found a unique balance. Just promise me that you will keep in touch with your genius side. Dankjewel dat je altijd laat zien dat normaal zijn best saai is, en klaar stond om te filosoferen over van alles, hoe gek de ideeën ook waren. **Reza**, for being my ping pong and gym partner. I will remember our time together as one of the most enjoyable parts of this whole journey. Our completely different perspectives and experiences made our discussions unique and plentiful. Although we disagree on many things, somehow, we both strongly agree that life is better in the east. **Anna**, dankjewel dat je meteen klaar stond op een van mijn meest stressvolle dagen van mijn leven. Daarnaast voor de inspiratie dat kinderen en wetenschap zeker niet het einde van de wereld zijn. Nogmaals, dankjewel. **Marina**, the most quietly brilliant person I have ever met, also thank you for being my spy-next-door, sheep enthusiast, and world-sandwich expert. Every time I get to speak to you, I am reminded of the "good old" days. It is always a joy to see and talk to you, although not as often as I would have liked. **Desi en Pippi**, Jelger en ik (en de rest van het lab) zijn sinds de dag dat jullie zijn vertrokken, jullie nooit vergeten. Het is fijn om te weten dat we elkaar niet uit het oog zijn verloren.

Nadiia and Costas, new blood does the lab good. I am looking forward to the insights you guys will bring with all these new ideas! **Aswin, Alicia**, all the students and former members of Brouns lab, thank you for not killing me. Jokes aside, it's been enjoyable to share the lab with you all.

Moritz, Ran, Nicole, and the rest of the Grussmiauwer lab. Thank you for being such a warm and joyful (neighbouring) lab group, and for sharing so many moments with us. From the early days when you were just two, to now where you guys almost outnumber us. It has been a joy to share our corner with you guys. Special thanks to **Moritz** to always being up for some fussball to blow off some steam and keeping me updated with the latest changes in life, the lab, and our shared journey of the PhD. I wouldn't have wanted it any other way. It's good to see that even though you guys enjoy looking at very small things, the big things don't go unnoticed. Danke schön.

Franklin, you crazy bastard. How much life has changed since the day we met. I still remember that day vividly. I owe many of the opportunities I've had to you. Still patiently waiting on that invite to come to the UK to play chess with a bottle of good whiskey. **Martijs en Han**, dank jullie wel dat jullie je zo inzetten voor de huidige generatie studenten en wat gaat de tijd snel! Ik kan mij nog heel goed de herinneren dat ik jullie voor het eerste ontmoette, en het gezamenlijke enthousiasme over het opzetten van vakken die dieper op de huidige biomedische technieken ingaan. Veel van ons hebben heel veel aan jullie te danken. Wat kan een klein idee zo'n grote impact hebben. **Anje**, dankjewel voor het verruimen van mijn idee van wetenschap, en mij uit te dagen om verder te kijken dan boeken je kunnen brengen. **Jerry Wells and Jessica Teeling**, thank you for giving me a place to "overwinter" the COVID times.

Conferences. Thank you for the privileged experience of flying around to world and being paid for it. I won't miss the sweaty, smelly conference halls, but I will miss the drinks afterward. You showed me that some scientists are actually fun to talk to.

My parents, for everything. The more I experience the parent life, the deeper my respect grows for all you did for me.

Véronique, mijn voor-altijd partner. Wat hebben wij al veel meegemaakt in zo'n korte tijd. En ja, jij had gelijk over dat het misschien tailocins zouden kunnen zijn, maar gelijk dat heb je wel vaker. Ik ben misschien super eigenwijs, maar dat komt omdat als ik ergens zeker van ben, ik het ook écht zeker weet. En één van die dingen is dat mijn leven oneindig veel leuker is samen met jou. **Mika**, het heeft zo moeten zijn. Het leven is een stuk betekenisvoller sinds jij erbij bent. Ze zeggen wel eens dat niets je kan voorbereiden op het worden van een ouder, nou ze hebben gelijk. Zelfs een PhD is er niks bij. Denk vooral ook nu niet dat jij ook een PhD moet gaan doen, alleen maar omdat wij dat hebben gedaan. Volg vooral je hart en wees niet bang, "het loopt wel los", dat zeggen ze hier in het wijze oosten.

Joëlle, thank you for sunshine, thank you for rain, thank you for joy, thank you for pain, it's a beautiful day-ay-ay-ay-ay. **De rest van mijn nieuwe familie** voor jullie warme welkom en mij mee te nemen in het gezellige Twentse leven.

Mijn Delft vrienden: Roel, Daniëlle, Gianni, Jeroen, Menno, John en Vera. Dankjewel dat jullie mij uit mijn donkere studenten kamertje hebben weten te trekken! Wat maakten jullie zomers om naar uit te kijken, altijd een thuisgevoel en om even niet te praten over wetenschappelijk ge-nerd. Ik mis jullie ontiegelijk.

Mijn grotere broer in vele aspecten en mijn inspiratiebron, **Odmar**. Ik heb dit al vaker gezegd, maar kan het niet vaak genoeg zeggen. De manier hoe jij in het leven staat en achter je dromen aangaat is aanstekelijk. Ik weet zeker dat grootse dingen jou te wachten staan.

Last but for sure not least: **Tor-Andre, Paul, Wouter, Merlin, Femke, Baltus, Storm, Maurice, Mike, Aron, Jeroen, Ivo, Kursad, and many more.** What a bunch of people you are. Endless conversations, endless memories. Seeing all of you guys succeed brings me joy.

Although my stubbornness might not show it, I have learned something from each of you.

Lastly, I would like to thank **the unpredictable nature of life**, for keeping daily life interesting and worth waking up to. Looking back, daily life felt ordinary, but so much has changed along the way. As a frog in slowly heating water, hardly noticing the changes. I can't help to wonder what life will be cooking next.

Curriculum Vitae

Daan Frits van den Berg

11-08-1996	Born in Haarlem, the Netherlands.

Education

2021-2026	Doctoral researcher in phage defenses and evasion strategies. Department of Bionanoscience, Delft University of Technology, the Netherlands. *Advisor:* Prof. dr. Stan J.J. Brouns
2018 – 2021	MSc Biomedical Sciences: Infection and Immunity, *Cum Laude*. University of Amsterdam, Amsterdam
2015 – 2018	BSc Biomedical Sciences. University of Amsterdam, Amsterdam
2008 – 2014	VWO: Nature & Health and Nature & Techniques Kennemer Lyceum, Overveen

Internships

2021	Intern in the effect of the gut microenvironment on the pathology of Alzheimer's disease, Department of Experimental Neuroimmunology at the University of Southampton and the Host-Microbe Interactomics Group of Wageningen University and Research. *Supervisors:* Prof. dr. J.M. Wells and Prof. dr. J.L. Teeling
2020	Intern in understanding phage defenses. Department of Bionanoscience, Delft University of Technology, the Netherlands. *Supervisors:* Prof. dr. Stan J.J. Brouns and dr. Franklin L. Nobrega
2019	Intern in understanding immunosenescence of the $CD8^+$ T cell population. Faculty of Medical Sciences at the UMC Groningen. *Supervisor:* prof. Jon D. Laman
2017	Intern in understanding the role of HMGB1 in the pathology of Crohn's disease. Tytgat Institute for Liver and Intestinal research, Amsterdam UMC. *Supervisor:* Dr. Anje A. te Velde
2017	Intern in the effect of pro-resolving mediators on treating Crohn's disease. Tytgat Institute for Liver and Intestinal research, Amsterdam UMC. *Supervisor:* Dr. Anje A. te Velde

Work experience

2025 - current	Founder of BioLumens: Bright insight in Biological Data.
2025 - current	Postdoctoral researcher at Brouns Lab. Delft University of Technology, Delft.
2023	Teaching assistant: Genomics University of Amsterdam, Amsterdam.
2022	Teaching assistant: Genomics University of Amsterdam, Amsterdam.
2021	Teaching assistant: Genomics University of Amsterdam, Amsterdam.
2021	Teaching assistant: Bioinformatics Delft University of Technology, Delft.
2020	Teaching assistant: Transcriptomics University of Amsterdam, Amsterdam.
2020	Teaching assistant: Genomics University of Amsterdam, Amsterdam.
2020	Teaching assistant: Bioinformatics Delft University of Technology, Delft.

Awards

2023	Best poster award. awarded at the international SPP2330 Symposium in Berlin (Germany).

About the author



(Daan on the left)

Daan Frits van den Berg was born on August 11, 1996, in Haarlem, the Netherlands. He completed secondary school (VWO) at the Kennemer Lyceum in Overveen, following the 'Science and Health' and 'Science and Technology'. During this time, Daan was 'among one of the best' in the preliminary round of the Dutch Biology Olympiad. Following his passion, he attended the Biomedical Bachelor at the University of Amsterdam. During his bachelor's degree he chose to focus on the innerworkings of human brain in his last year. Besides this interest, he put his attention to the causes and possible treatments of Crohn's disease, a disease affecting the gut, primarily studying the human immune response, under the supervision of dr. Anje te Velde. Next, he decided to continue this path by also following the Biomedical sciences master's program at the University of Amsterdam to study the human immune response (Infection and Immunity) in more detail. Earning a *Cum Laude* and taking additional courses to be able to work with next generation sequencing techniques and statistics. In his first internship at the Brouns lab at the Delft University of Technology, he expanded beyond the human immune response by investigating the immune response of bacteria, under the supervision of dr. Franklin Nobrega and prof. dr. Stan Brouns. Specifically, the immune response of bacteria against their viruses, bacteriophages. Finding that a broader immune response repertoire correlates to a broader resistance against bacteriophages. Following this internship, he decided to do an internship that combined both interests of his bachelor's degree: the brain and the gut, by investigating the role of the gut in causing Alzheimer's disease, under the supervision of prof. dr. Jerry Wells and prof. dr. Jessica Teeling. Although the results remained preliminary, he gained additional lab experience and investigation skills.

In 2021, Daan returned to the Brouns lab at the Delft University of Technology to start his PhD under the supervision of prof. dr. Stan Brouns and continued his work from where he left off during his Masters' internship. Focussing on the immune response of bacteria against bacteriophages and how bacteriophages circumvent the immune response of the host, resulting in this dissertation: Phage Immunity and Evasion. Describing several previously unknown mechanisms that bacteria use to defend themselves from bacteriophages and how bacteriophages counteract these systems. During this PhD, he also moved to the east of the country, Enschede, and started a family.

List of publications

Journal articles related to this dissertation

7. **van den Berg, D. F.***, Costa, A. R.*, Esser, J. Q.*, van den Bossche, H., & Brouns, S. J. J. (2025). Discovery of phage defense systems through component modularity networks. *bioRxiv*, 2025-09. Under peer review at *Host Cell & Microbe*.
6. **van den Berg, D. F.**, & Brouns, S. J. J. (2025). Phage tRNAs: decoding the enigma. *Trends in Microbiology*.
5. Costa, A. R.*, **van den Berg, D. F.***, Esser, J. Q.*, van den Bossche, H., Pozhydaieva, N., Kalogeropoulos, K., & Brouns, S. J. J. (2025). Bacteriophage genomes encode both broad and specific counter-defense repertoires to overcome bacterial defense systems. *Cell Host & Microbe*, *33*(7), 1161-1172.
4. **van den Berg, D. F.**, & Brouns, S. J. J. (2025). Reduced prevalence of phage defense systems in *Pseudomonas aeruginosa* strains from cystic fibrosis patients. *mBio*, e03548-24.
3. **van den Berg, D. F.***, Costa, A. R.*, Esser, J. Q.*, Stanciu, I., Geissler, J. Q., Zoumaro-Djayoon, A. D., Haas, P. & Brouns, S. J. J. (2024). Bacterial homologs of innate eukaryotic antiviral defenses with anti-phage activity highlight shared evolutionary roots of viral defenses. *Cell Host & Microbe*, *32*(8), 1427-1443.
2. Costa, A. R.*, **van den Berg, D. F.***, Esser, J. Q.*, Muralidharan, A., van den Bossche, H., Bonilla, B. E., van der Steen, B. A., Haagsma, A. C., Fluit, A. C., Nobrega, F. L., Haas, P., & Brouns, S. J. J. (2024). Accumulation of defense systems in phage-resistant strains of *Pseudomonas aeruginosa*. *Science advances*, *10*(8), eadj0341.
1. **van den Berg, D. F.**, van der Steen, B. A., Costa, A. R., & Brouns, S. J. J. (2023). Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations. *Elife*, *12*, e85183.

Other journal articles

7. Kalogeropoulos, K.*, van Beljouw, S. P. B.*, Feldmann, D., **van den Berg, D. F.**, & Brouns, S. J. J. Proteases in prokaryotic antiviral immunity and their potential in bioengineering. Manuscript accepted at *Trends in Biochemical Sciences*.
6. Muralidharan, A., Costa, A. R., Fierlier, D., **van den Berg, D. F.**, van den Bossche, H., Zoumaro-Djayoon, A. D., Pabst, M., Pacesa, M., Correia, B.E., & Brouns, S. J. J. (2025). Molecular basis for anti-jumbo phage immunity by AVAST Type 5. *bioRxiv*, 2025-07. Manuscript accepted at *Molecular Cell*.
5. Mahler, M., Malone, L. M., **van den Berg, D. F.**, Smith, L. M., Brouns, S. J. J., & Fineran, P. C. (2023). An OmpW-dependent T4-like phage infects *Serratia sp.* ATCC 39006. *Microbial genomics*, *9*(3), 000968.
4. Japaridze, A., van Wee, R., Gogou, C., Kerssemakers, J. W., **van den Berg, D. F.**, & Dekker, C. (2023). MukBEF-dependent chromosomal organization in widened *Escherichia coli*. *Frontiers in microbiology*, *14*, 1107093.
3. Estrada Bonilla, B., Costa, A. R., **van den Berg, D. F.**, Van Rossum, T., Hagedoorn, S., Walinga, H., Xiao, M., Song, W., Haas, P., & Brouns, S. J. J. (2021). Genomic characterization of four novel bacteriophages infecting the clinical pathogen *Klebsiella pneumoniae*. *DNA Research*, *28*(4), dsab013.
2. van Beljouw, S. P. B., Haagsma, A. C., Rodríguez-Molina, A., **van den Berg, D. F.**, Vink, J. N., & Brouns, S. J. J. (2021). The gRAMP CRISPR-Cas effector is an RNA endonuclease complexed with a caspase-like peptidase. *Science*, *373*(6561), 1349-1353.
1. **van den Berg, D. F.**, & te Velde, A. A. (2020). Severe COVID-19: NLRP3 inflammasome dysregulated. *Frontiers in immunology*, *11*, 1580.

Delivered conference presentations

4. The 2025 Symposium on the immune system of bacteria, Paris, France, Poster presentation, 2025
3. The 2024 Symposium on the immune system of bacteria, Boston, USA, Poster presentation, 2024
2. International SPP2330 Symposium, Berlin, Germany, Poster presentation, 2023
1. Belgian Society for Viruses of Microbes conference, Leuven, Belgium, Poster presentation, 2022

Equal contribution indicated with an asterisk (*)



Bacteria are under constant predation by phages. To protect themselves, bacteria evolved a multitude of defense mechanisms. Some of these phage defense mechanisms, such as CRISPR-Cas systems, have been known for more than a decade.

More recently, it has become evident that the repertoire of phage defense systems of bacteria is far more diverse and numerous. This finding not only reshaped our understanding of phage defense systems, but also raised the question of how phages can still successfully infect despite this complex immune responses of its host.

In this dissertation, we investigate this question by specifically looking at the importance of phage defense systems, and discovering additional defenses and counter strategies.