NEW COMPOUNDS INDUCING MODEL PHAGES

IDENTIFYING NEW COMPOUNDS CAPABLE OF INDUCING MODEL PHAGES

By

ANISHA NANDY

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for the Degree

Master of Science

McMaster University MASTER OF SCIENCE (2020) Hamilton, Ontario
(Department of Biochemistry and Biomedical Sciences)

TITLE: Identifying new compounds capable of inducing model phages

AUTHOR: Anisha Nandy, B.Sc. (McMaster University)

SUPERVISOR: Dr. Alexander P. Hynes

NUMBER OF PAGES: xi, 80

LAY ABSTRACT

Bacterial viruses (phages) can lie dormant as prophages in their host bacterium until a signal triggers their activation, production of viruses, and rapid killing of the host. This switch from dormant prophage to active phage is called induction. Almost all molecules that result in prophage inductions belong to a limited set of compounds which elicit a specific stress response in bacteria.

Screening 3936 compounds for their ability to inhibit the growth of bacteria carrying known prophages resulted in the identification of a small subset associated with increased phage production. For one *Escherichia coli* prophage—HK97, a model of induction—we found 49 compounds not previously known as inducers. For another model prophage—Mu, a prophage thought to be chemically uninducible—we identified seven such compounds.

These compounds will serve as tools to determine what signals prophages can respond to, and potentially identify new stress pathways of interest in bacteria.

ABSTRACT

Prophages are the genomes of bacteriophages (phages, bacterial viruses) that integrate into the chromosome of their host upon infection, lying dormant until conditions favour their reactivation. A cell harbouring a prophage is called a lysogen, as, upon exposure to certain signals, the prophage will initiate a replicative cycle ending in lysis of the host bacterium and release of phages. This process is known as induction. Canonically, induction occurs through activation of the bacterial SOS-response, a DNA-repair cascade initiated by detection of DNA damage. Studies of prophage induction have almost exclusively relied on challenges with compounds that result in the initiation of the host SOS response.

Recent studies have identified some signals that affect prophage induction independently of the SOS response, but these approaches have not been systematic. To identify non-canonical triggers of prophage induction, I screened 3,936 compounds against two model lysogens. The first, carrying phage HK97, is a model for induction. The second, carrying phage Mu—a prophage thought to be uninducible—serves as a control. Any compound which inhibited bacterial growth in only our HK97 lysogen was considered to have resulted in a phage-mediated response. The 171 compounds identified in this screen were then used to re-challenge the lysogen at a range of concentrations, and monitor the resulting release of free phages associated with induction. Increases in phage counts were seen for 86 compounds. While 38 of these were known SOS activators, 49 were novel, ‘non-canonical’

inducers. Unexpectedly, the screening also revealed seven unique chemical inducers for the supposedly un-inducible model prophage, Mu.

The 56 new phage-inducers identified by this work include compounds likely to be driving phage induction through non-canonical pathways. As prophages are thought to respond to bacterial stress, these may reflect stressors acting through new mechanisms. Using these compounds as tools opens up an avenue to probe other stress pathways in bacteria, and, as evidenced by induction of Mu, potentially help discover new phages that don't respond to canonical inducers.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Alexander P. Hynes, for his guidance throughout this project, as well as my committee member, Dr. Eric Brown who has been an invaluable source of advice and ideas. I am grateful to Dr. Jake Magolan for accepting to be my committee member for my thesis defence. I have thoroughly enjoyed working with Dr. Hynes and my committee, and appreciate each of their contributions to my work. My coworkers have been incredibly helpful and inspiring. I want to specially thank Clara for teaching me helping me with new microbiology techniques, Rabia for reading providing me edits for every written work I had to submit and Felix for helping with illustrations. My lab has been an excellent source for social well-being too from going out to dinner, to board games nights, to escape rooms each of them has had a positive impact in my life.

I would like to thank Dr. Tracey Campbell and Susan McCusker for their guidance and suggestions throughout my M.Sc. project. They trained and trusted me with all the High throughput screening equipment.

I would also specifically like to thank Dr. Jennifer Stearns for being an amazing mentor during my volunteer experience in her lab, and contributing to my positive experience in the Hynes lab that resulted in my decision to remain here for graduate school. To Hiba, who has shared the experience of being the first of the Hynes lab M.Sc. students, I appreciated our endless brainstorming sessions and life chats over drinks. I would especially like to thank her for being a friend to talk to when times got hard.

Thank you to my parents for helping me explore the world at a young age and for tolerating all my, tedious unsolicited explanations of my experiments and science that I believed were exciting.

Lastly, I would like to thank my McMaster friends, who have been sources of both knowledge and guidance. To Uzair, I would especially like to thank you for all your help and endless support—proofreading my abstracts, assignments and thesis, quizzing me before tests (even for courses you weren't taking), and helping me draft my first email to Dr. Hynes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

LB— Lysogeny Broth

STEC—Shiga-Toxin-Encoding *Escherichia coli*

CTX—Cholera Toxin

TCP—Toxin Coregulated Pilus

UV—Ultraviolet

BIA—Biochemical Induction Assay

CMCB— Center for Microbial Chemical Biology

OD—Optical Density

MIC—Minimum Inhibitory Concentration

DMSO—Dimethyl sulfoxide

VLP—Virus like particles

GLOSSARY OF TERMS

Phage: Bacterial virus, common form of full "bacteriophage"

Prophage: Quasi-dormant bacterial virus, integrated in its host genome

Lysogen: Bacterium host harbouring a prophage

Induction: Process through which a prophage switches from its quasi-dormant state to a lytic cycle and releases progeny phages

HK97: A model lambdoid phage, inducible through canonical means

Mu: A model phage that is not thought to be inducible.

DECLARATION OF ACADEMIC ACHIEVEMENT

I, Anisha Nandy, declare this thesis to be my own work.
The research work associated with this thesis was completed by me.

Clara Fikry generated the Mu lysogen and confirmed its incorporation in *ygbM* gene.

Dr. Tracey Campbell performed the analysis for the pilot study.

CHAPTER 1: INTRODUCTION

Bacteriophages (phages) are viruses that exclusively infect bacteria. Temperate phages represent a category of phage which, during infection, can integrate into the host genome and stay in a quasi-dormant state. This incorporated viral genome is called a **prophage,** and the host cell, bearing the phage DNA, is called a **lysogen** (Ptashne 2004).

1.1 Prophages in bacteria

Of the sequenced bacteria, over half are lysogens (Touchon et al. 2016) replete with prophages, which can account for up to 20% of the bacterial genome (Casjens 2003; Keen and Dantas 2018). These prophages can have large impacts on the lifestyle, fitness and virulence of the lysogens (Howard-Varona et al. 2017). The most prevalent of these is superinfection immunity where, once the phage genome integrates into the bacterial DNA, the lysogen is typically resistant to subsequent infections by closely related phages (Abedon 2015; Howard-Varona et al. 2017).

Superinfection immunity is only one example of how a prophage is not completely dormant, and can confer novel phenotypes to the lysogen. This process of altering the bacterial host is called lysogenic conversion. Key virulence factors of several prominent bacterial pathogens are prophage-encoded exotoxins e.g. Shiga-toxin-encoding *E. coli* (STEC), *Vibrio cholerae*, and *Clostridium botulinum* (Keen and Dantas 2018). In the famous case of *V. cholerae*, its toxicogenic nature is attributed to prophage CTX_{φ} which encodes

the cholera toxin (CTX). Only CTX_{φ} prophage-carrying strains of *V. cholerae* cause pandemic and epidemic cholera (Faruque & Mekalanos 2012)

Perhaps the defining property of lysogens, from which they derive their name, is the potential of the prophage to start independent replication, resulting in the eventual lysis and killing of its host to release progeny phages.

1.2 Induction of prophages

While the quasi-dormant prophages can be stably maintained in a bacterial population over evolutionary timescales, they can also switch into an active lytic cycle. This transition from quasi-dormant prophage to independently replicating phage that eventually lyses the host cell is known as **induction**.

Ecologically, the lysogenic cycle is thought to be favoured when the host density is low and phages that do not eradicate their hosts would be favoured. Conversely, when host density is high, the lytic cycle allows for and rapid propagation through the population (Nanda et al. 2015). In the lab, induction happens spontaneously, at a very low rate.

In the majority of the cases studied in the lab, an external stressor is used to trigger induction of the prophage. The most well-characterized and canonical trigger is the bacterial host's DNA-damage response, known as the SOS response (section 1.3.2).

1.3 Canonical prophage induction

1.3.1 Phage lambda, the model for induction

Lambda is a phage that has been the subject of extensive early phage research (Casjens and Hendrix 2015). Work done on phage lambda from the

mid-1950s to mid-1980s was critical to understanding gene regulation, phage particle assembly, and, more specifically, the molecular nature of lysogeny (Ptashne 2004). Lambda lysogeny is the defining model for the integration and induction of prophages. The study of this model defines the closely related lambdoid (lambda-like) phages, like **HK97** (Hendrix 2005; Lander et al. 2008) but also forms the foundation of all research on temperate phages.

The initiation of a lysogenic cycle in phage lambda can be said to start with the site-specific integration of the phage genome into the bacterial chromosome (Alberts et al. 2002; Dhillon et al. 1980). For phage HK97, as for phage lambda, the attachment site is in the vicinity of the *E. coli* gal operon (Hatfull and Hendrix 2011). Upon integration, the prophage produces the repressor protein CI. This repressor protein blocks promoters of the lytic genes in phage lambda, thus maintaining the lysogenic state. There are other classes of temperate phages that infect *E. coli* but do not follow the same mechanism of integration as phage HK97 and other lambdoid phages.

Other known temperate phages use random transposition like phage **Mu** (Bukhari 1975) or plasmid-like partitioning like phage N15 (Hatfull and Hendrix 2011). Any phage that has a lysogenic state unlike lambda are said to be the exception to the rule.

The lambdoid phages are not only the model for integration, but also the model for studying induction. Induction of a lambdoid prophage begins with the reversal of the integration event—the excision of the phage genome from the bacterial chromosome. In lambda, the excision reaction requires the product of the phage gene *xis* (Griffiths et al. 2000; Ptashne 2004). The

repressor protein in phage lambda, CI, maintains lysogeny and by extension represses the expression of *xis*. If the repressor isn't produced or is broken down, the phage will induce and start the lytic cycle. There are prophages that are exceptions to this rule—like Mu, which does not encode a CI-like repressor and, in the absence of certain specific mutations (Bukhari 1975), is thought not to be induced by bacterial stressors.

While many ecological factors are thought to influence a prophage's lysis/lysogeny decision, the decision is most often studied in the context of bacterial stress. This is often considered analogous to the prophages 'abandoning a sinking ship', leaving a close association with a stressed cell to seek out hosts that are faring better.

1.3.2 Bacterial SOS response

Irrespective of their habitat and environment, bacterial genetic material is subjected to insults from multiple sources such as UV, oxidative damage, antibiotics and mutagens. In order to avoid DNA damage, bacteria have evolved repair mechanisms that are specifically expressed only when they sense a threat to their DNA (Michel 2005; Žgur-Bertok 2013). The enhanced expression of DNA repair genes was proposed by Miroslav Radman in 1973 who termed it the SOS response (Simmons et al. 2008).

The SOS response in bacteria is regulated by two key proteins; RecA, an inducer, and LexA, a repressor. The SOS response system initiates when damaged DNA interacts with the RecA protein and activates it (Radman 1975). In the absence of this activation, LexA binds to the operator region of DNA-damage repair genes to prevent their expression (Sauer et al.

1982). However, activated RecA (RecA*) has a co-protease activity that causes the repressor LexA to self-cleave and consequently, initiates the SOS response to the DNA damage.

As much as the SOS response plays a part in the DNA repair for the host by activating the RecA protein, it also results in the induction of prophages (Rozanov et al. 1998). The bacterial LexA repressor and phage lambda's CI repressor are homologous at their carboxy-terminal domains (Janion 2008). Upon activation of RecA, other LexA homologues like the CI repressor also undergo self-cleavage (Rozanov et al. 1998; Nanda et al. 2014). This cleavage derepresses the *xis* gene and thereby allows the expression of excicionase, which initiates lambdoid prophage induction (Fig. 1). By comparison, this sequence of events does not occur in prophages like Mu, due to the lack of a similar CI-like repressor.

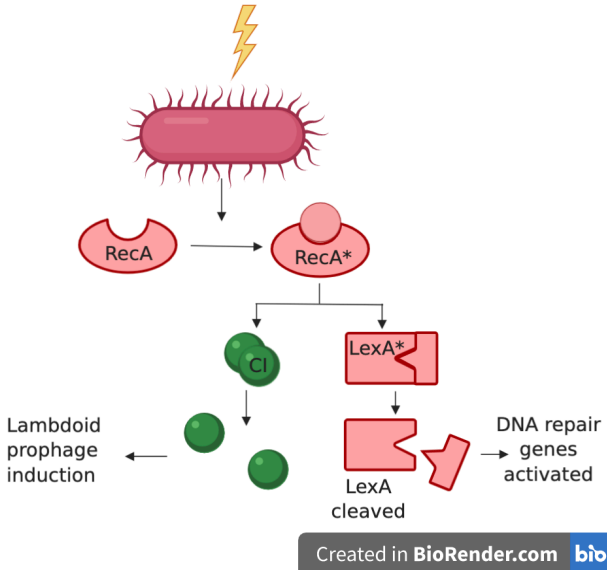


Figure 1: Host SOS-response dependent prophage induction. A schematic showing how the RecA protein is activated when the cell encounters DNA damaging agents. This activation cleaves the downstream LexA protein to unlock the DNA repair genes. In doing so it also cleaves the CI protein which causes lambdoid prophage induction.

1.3.3 Biochemical induction assays

This property of prophages responding to DNA damage has been extensively exploited to find and test new DNA-damaging compounds. Prophage induction assays are routinely used to screen antitumor agent and are an efficient method to predict the potency of chemotherapeutic drugs (Elespuru and White 1983).

Induction of phage lambda, specifically, was used as an indicator in biochemical induction assays (BIA) to test antitumor reagents (Heinemann 1971; Elespuru and Yarmolinsky 1979; Anderson et al. 1980; Akeju et al. 1998). The lambda prophage is inserted to interrupt the function of the *gal* gene, and successful prophage induction reinstates the production of β-galactosidase, allowing the cell to break down the colorimetric substrate β-galactoside (Elespuru and White 1983). The resulting colour intensity reflected the potency of the antitumor drug.

Biochemical prophage induction assays offer a framework to test potency of new compounds and antibiotics.

1.3.4 Fluoroquinolones and β-lactams

In addition to chemotherapeutics, phage induction has been widely studied using different classes of antibiotics (Majtanova et al. 1994). Sub-inhibitory concentration of ciprofloxacin, a fluoroquinolone antibiotic actively prescribed for bacterial infections of the urinary and respiratory tract, was shown to cause prophage induction and increase phage related virulence factors in *Staphylococcus aureus* (Goerke et al. 2006). This is also observed in multidrug-resistant *Salmonella*, enterohemorrhagic *E. coli* and *Streptococcus*

pneumonia (Walterspiel et al. 1992; López et al. 2014; Bearson and Brunelle 2015), among others.

Fluoroquinolones are one of the most widely prescribed class of broad-spectrum antibiotics used against gram-positive and gram-negative bacteria. (Blondeau 2004; Kimura et al. 2008). Their mode of action is through interaction with DNA gyrase, preventing it from relieving the stress from double-stranded DNA being unwound by helicases (López et al. 2014; Bearson and Brunelle 2015). This can cause single stranded breaks in the DNA and consequently trigger the bacterial SOS response. Accordingly, fluoroquinolones are excellent inducers of lambdoid phages.

Fluoroquinolones are not the only class of antibiotics known to induce prophages; β-lactams are also used in this fashion. β-lactams bind to penicillin-binding proteins in bacteria (Delhaye et al. 2019). Most of these proteins are required for cross-linking of the peptidoglycan layer that forms the cell-wall. Despite no interaction with DNA damage, damage to the peptidoglycan through β-lactams appears to indirectly activate the SOS response, leading to prophage induction (Tipper and Strominger 1965; Rodríguez-Tébar and Vázquez 1984; Miller 2004; Adamus-Białek et al. 2019; Delhaye et al. 2019; Maiques et al. 2006). This has been demonstrated by inducing prophages in the *Staphylococcus aureus* genome using the clinically prescribed β-lactam antibiotics ceftriaxone and cloxacillin (Zeng and Lin 2013). Ampicillin and penicillin, two other β-lactams, also induce phages 80α and ϕ11 in *S. aureus*. No induction of these prophages was seen in *recA* mutants unable to initiate the SOS response (Maiques et al. 2006).

Parallels between SOS repair and prophage induction were established early (Heinemann 1971). Since then, research regarding induction of prophages has almost always relied on the activation of the host SOS response, and SOS-dependent phage induction continues to be the standard for the discovery of inducible prophages (Oh et. al 2019).

1.4 Non-canonical signals influencing phage life cycles

Recent evidence shows the tendency of a temperate phage to adopt lysis or lysogeny depends on many more factors than simply the SOS response including host perception of nutrient availability, host density and even signaling from other prophages through mechanisms similar to quorum sensing (Feiner et al. 2015; Knowles et al. 2016; Silpe and Bassler 2019).

1.4.1 Phage communication

A recent study by Erez et. al. (2017) showed that phage phi3T can sense chemical signals released by the previous generation of phages to decide whether to kill upon infection, or lysogenize the host. This chemical signal was a phage-encoded, 6 amino acid polypeptide called ‘arbitrium’. Arbitrium is released out of the host bacterium every time it is lysed by phage phi3T. The higher the concentration of arbitrium in the media, the more hosts phi3T has recently killed, and consequently the lesser the availability of uninfected bacterial hosts for the next generation of phages. The study proved that in the presence of increased arbitrium in the media a phage, upon infecting the new host bacterium, favours lysogeny. Erez et al. also identified two additional phage phi3T encoded proteins which are responsible for measuring the amount of arbitrium responding to those levels.

This remarkable study showed direct phage communication between a phage and its predecessors to assess the density of viable host cells. These factors affecting the decision to influence lysogeny shed light on how lysis/lysogeny decisions are not always at the mercy of a host stress pathway, and that phage can take responsibility, as it were, for their own fate.

In a related strategy, Silpe and Bassler (2019) showed that phage detection of the host-produced autoinducers, used in quorum-sensing, can also control the phage lysis/lysogeny switch. *V. cholerae*, the host of phage VP882, secretes the signaling molecule called DPO, made of amino acids threonine and alanine. DPO is involved in the activation of genes required for biofilm formation and toxin production in situations of high host density.
Phage VP882—either as a prophage or a newly injected phage genome—detects the same signal through a phage-encoded protein, using it to gauge bacterial abundance in the environment. This directly informs the lysis/lysogeny decision, favouring a lytic cycle at high host densities by sequestering the phage repressor.

Prophages are detecting far more than just the host SOS-response. They are not only using self-encoded polypeptides (Erez et al. 2017) but also the small molecules that are used as QS signals of their bacterial host (Silpe and Bassler 2019) to decide between lysis or lysogeny.

1.4.2 Unusual compounds

Boling et. al studied the responses of three common gut bacteria species *Bacteroides thetaiotaomicron, Enterococcus faecalis*, *Staphylococcus aureus*, as well as an opportunistic pathogen *Pseudomonas aeruginosa*, to 117

commonly consumed foods, plant extracts, and chemical additives (Boling et al. 2020). They showed that compounds like clove, propolis (derived from bees), aspartame and stevia were prophage inducers.

To demonstrate this, the authors clustered bacterial growth curves to which these compounds had been added, identifying 28 bacteriostatic or bactericidal compounds. Through flow cytometry to count stained virus like particles (VLPs), they determined which of these compounds resulted in increased viral particle counts indicative of prophage induction. Of the 28 antimicrobial compounds, 11 resulted in higher levels of VLPs. Stevia, a plant-derived sugar substitute, was the most potent prophage inducer, increasing the number of VLPs more than 400% (Boling et al. 2020). As stevia is not known to induce the SOS response, this work suggests that there are many other potential triggers of prophage induction.

However, simply because a compound is unusual or not known to induce the SOS response does not mean it induces phages through non-canonical pathways. Oh et al. (2019), demonstrated that dietary fructose stimulated prophage induction in *Lactobacillus reuteri* 6475. They found that *L. reuteri* grows poorly on fructose-supplemented media compared to media containing glucose, galactose, or arabinose. This reduced growth is associated with increased virus production as measure from VLPs increase. In contrast, prophage induction in culture, when supplemented with fructose, was reduced in an acetate kinase (*ackA*) mutant. The authors suggested that fructose utilization promotes conversion of acetyl phosphate to acetic acid, and that the resulting accumulation of acetic acid in the bacterial cell causes DNA damage.

The DNA lesions, through activation of RecA, induce a prophage in *L. reuteri* (Oh et al. 2019).

1.5 Hypothesis

Prophage induction is canonically achieved by activating the SOS response of the bacterium. However, studies have shown that prophages can sense a variety of non-canonical signals that can drive their lysis-lysogeny decisions.

Although there has been mounting evidence for non-canonical phage induction, there have been no systematic characterizations of the novel signals capable inducing prophages. These compounds could be used as tools to probe additional bacterial stress pathways, induce phages not induced through canonical pathways, and shed light on phage lysis lysogeny decisions.

I hypothesize that screening large compound libraries against lysogens containing model prophages will lead to the identification of novel chemical signals that result in induction.

Due to a detailed mechanistic understanding of prophage HK97 induction, I selected it as a model for identifying novel inducers. A high throughput primary screen of a compound library against an HK97 lysogen will help us identify antibacterial compounds on the basis of impaired bacterial growth. Those compounds will be further investigated through a secondary high throughput screen to determine if the antibacterial effect was associated with increased phage production. Throughout the process, I will compare the HK97 lysogen to a Mu lysogen. The Mu prophage does not induce in response to the SOS response and is generally thought not to be inducible through

bacterial stress pathways. As such, it will serve as a negative control with the potential to also identify new prophage-inducing compounds.

CHAPTER 2: MATERIALS & METHODS

2.1 Media and growth conditions

E. coli was grown in Lysogeny broth (LB) with shaking at 130 rpm, at 37 °C. Bacterial stocks were prepared by freezing an overnight culture which was pelleted then resuspended in 850 µL of LB, and transferred to a tube with 150 µL sterile glycerol. LB plates were prepared with LB supplemented with 1% w/v agar, unless used for streak tests or stamp plates, in which case 1.5% w/v agar was used instead.

To generate bacterial lawns, with or without phages, bacteria were added to molten soft (0.75%) LB agar, and overlain on an LB plate—then grown overnight at 37 °C.

2.2 Phage manipulation

Filtrations were done using a 10 mL syringe and 0.45 µm syringe filter. Phage dilutions were made in LB.

Phage amplification: Frozen phage stock (-80 °C) was scraped and added to 10 mL of ~0.2 OD of host bacteria and grown for 6 h. The culture was filtered and the lysate of the first amplification was stored at 4 °C. Of this lysate, 100 µL were then added to 10 mL of freshly grown ~0.2 OD host bacterium and incubated for 12 h. The culture was filtered and the lysate of the first amplification was stored at 4 °C.

Phage Titration: Phage lysates were titred by plating 10-fold serial dilutions of the lysate on the host bacterium. Briefly, 100 µL of the lysate dilution was added to 3 mL molten overlay agar (55 °C) containing 300 µL of

an overnight bacterial culture. The plates were incubated overnight, and the plaques were counted (pfu/mL, a measure of phage particles able to form plaques per unit volume) on plates containing between 30 and 300 plaques.

2.3 Host bacterial strains and lysogen generation

Strains used in this work are listed in Table 1. Lysogens of both HK97—our model lambdoid phage, and Mu, our model transposable phage—were generated by co-incubation of phage and host, as follows: the host bacteria were grown overnight, then used to inoculate 10 mL of fresh LB at 1% v/v. Approximately $5x10^8$ pfu of phage were added to the culture, then left to incubate for 4 h. The phage-host mixture was then serially diluted in LB, and dilutions were plated using soft agar overlay to isolate surviving bacteria. This assay exploits the property of superinfection immunity, whereby a lysogen will be immune to subsequent infections by related phages.

Table 1: *E. coli* strains and phages used.

Organism	Strain	Description	Source
E. coli	K-12S (amber suppressant)	Host for HK97 and lambda vir F+, tyrT58 (AS), mel-	(HER 1382)*
	40	Host for Mu whose provenance can be traced back to AI Bukhari, exact genotype is unclear	(HER 1252)*
	HK97 lysogen	K-12 with HK97 prophage insertion near *gal* genes	Generated in this work
	Mu lysogen	*E. coli* 40 with Mu prophage in *ygbM* gene	Generated in this work
Phage	HK97	Caudovirales, Siphoviridae	(HER 382)*
	Mu	Caudovirales, Myoviridae	(HER 253)*
	Lambdavir	Caudovirales, Siphoviridae (virulent), not lysogenic	(HER 37)*

* *Félix d'Hérelle* Reference Center for Bacterial Viruses

To confirm integration of the HK97 prophage was responsible for the surviving bacterial colonies, purified colonies were streaked on a plate through phage lysates. Resistance to HK97 but susceptibility to the virulent, but closely related phage LambdaVir indicated a likely lysogen. These candidates were then confirmed by PCR of the known phage-host junction. I designed the primers attBF and HK97lysR, tested them *in silico* using geneious (Kearse et al., 2012), and had them synthesized by *Integrated DNA Technologies* (IDT, IA). Colony PCR resulting in a band across the phage-host junction confirmed the colony was an HK97 lysogen.

Due to the random integration of Mu in the host genome, the Mu lysogen was confirmed by arbitrary PCR with nested primers. This involved two nested primers (SP1 and SP2) specific to the end of the Mu genome. The first PCR amplified lysogen DNA using SP1 and an arbitrary primer (arb01 or arb06). This second PCR used primers SP2 (Mu_circle_R) and arb02—which is nested within the arb01 and arb06 sequences. These products were sequenced with SP2 to determine the site of integration. This work on the Mu lysogen generation was performed by Clara Fikry (co-op student, Winter 2018).

Table 2: Primers used.

Primer name	Sequence	Purpose
HK97lysR	GCGTGTAATTGCGGAGACTT	With attBF, Detecting HK97 lysogens, binds to prophage genome
attBF	TGAATCCGTTGAAGCCTGCT	With HK97lysR, Detecting HK97 lysogens, binds to host genome
Mu3_ygbN_F	CCGATGTCTGCGTGGAGTAA	Confirm Mu integration site in lysogen with arb02
Mu_Circle_R/ SP2	TCCAATGTCCTCCCGGTTTTT	Amplifies inward for circle PCR to locate Mu
arb01	GGCCACGCGTCGACTAGTAC NNNNNNNNNNGATAT	Arbitrary PCR with SP1
arb02	GGCCACGCGTCGACTAGTAC	Arbitrary PCR (SP2), binds tag on arb01 and arb06
arb06	GGCCACGCGTCGACTAGTAC NNNNNNNNNNACGCC	Arbitrary PCR with SP1
SP1	CTTGCAAGCCCCACCAAATC	Arbitrary PCR to find Mu integration site (arb01and arb06)

2.4 Induction tests of model lysogens with ciprofloxacin

To test inducibility of our model lysogens, a known SOS response activator, ciprofloxacin, was first used to determine the minimum inhibitory concentration for the host *E. coli* HER 1382 (Table 1). The culture was added to a 96 well plate with two-fold dilutions of ciprofloxacin prepared in DNAse-free water, ranging from 0.00625 μg/mL to 0.1 μg/mL. The ability of the

newly generated lysogens to induce was tested at 1/2MIC and 1/4MIC ciprofloxacin my measuring the change in OD_{600nm} every 30-minutes.

2.5 High-throughput screening

2.4.1 Validation tests

To determine growth and concentration parameters for our high throughput screens, I performed validation tests in 96-well plate by adding ciprofloxacin and DMSO to 150 µL HK97 lysogen culture at $OD_{600nm\ o}$ of ~0.2, ~0.4 and ~0.6. Another validation test checked any variation between the signals by testing eight different concentrations of a ciprofloxacin ranging from 1.5 µM to 100 µM, each with eight technical replicates. Both these validations were performed in biological duplicates with the HK97 lysogen.

2.4.2 Pilot study

The pilot test involved screening the conventional model of lysogeny—HK97 with a subset of the library of interest. Access to the bioactive library was generously provided by the Center of Microbial Chemical Biology at McMaster University (CMCB). I selected one of the 384-well plates of the 'bioactive compound library', for its high number of canonical prophage inducing compounds for this pilot study (Appendix II). A 10 ml overnight culture of the HK97 lysogen was used as a 1% v/v inoculant for 60 ml of fresh LB media in an 125 ml Erlenmeyer flask. The culture was then grown at 37 °C with shaking at 230 rpm until it reached an OD_{600nm} of 0.2.

The freshly grown culture was transferred independently to two 384-well plates using the Biomek® FXP integrated liquid handler (Beckman Coulter, IN) at the CMCB, adding 49.5 µL ~0.2 OD_{600nm} HK97 lysogen

culture using a Beckman-Coulter robotic arm attached with a 384-well channel head. The same device transferred 0.5 µL of a compound from the compound collection plates to the wells containing 49.5 µL of lysogen culture. The final concentration of the compound in the culture was 10 µM. An ultrasonic wash was performed between each transfer for the replicates to eliminate cross-contamination. The plates were incubated at 37 °C with no shaking and OD_{600nm} readings were taken every 30 min for 4 h.

2.5.2 Z' test for hit identification

For a well-defined signal window, Z' should be greater than 0.5; that is, the sum of the errors of the controls should not be more than 50% of the total separation between the average of the negative and positive controls (Mangat et al. 2014). The Z' value was calculated at each time point.

Screening data consists of a time zero read and an additional eight time points after compound addition. The screening data were normalized using the interquartile method (IQM) (Mangat et al. 2014) per plate basis by dividing the optical density at that time point of every well on the plate by the mean optical density of the plate. The standard cut-off to use for pilot study hit selection is the mean of the data minus three times the standard deviation (3 SD).

2.5.3 OD drop method for hit identification

I wrote a custom R-script using the tidyverse library in R-studio (Wickham et al. 2019) to look for the characteristic growth curve indicating prophage induction (section 2.3) for the compounds used in the pilot study. The OD read of the two replicates for every well at every time point was averaged. The OD drop was calculated from these averaged values. A

compound was said to be inducing the prophage if the difference between the maximum OD and the end point OD was ≥ 0.01 for this analysis method (github/anishanandy/Plate_analyses). The compounds showing an average OD drop of 0.01 or more were considered primary screen hits. The graphing to separate these curves from the non-phage inducing compounds was done using ggplot library, also developed in R (Villanueva et al. 2016).

2.6 Primary screening of the compound library

2.6.1 Screening method

OD_{600nm} measurements were taken at 30-min intervals as in the pilot study for all 3936 compounds and no compound controls. The experiment time was increased to 6 h after pilot screen analysis showed more hits were identified at the end of the screen. The assays were performed in duplicates for all the compounds in the library with both HK97 lysogen and Mu model. All these compounds were tested at a singular concentration of 10 μM.

2.6.2 Primary screen analysis

The same custom R-script was used to generate individual graphs for both replicates of the thirteen 384-well plates containing the 3936 different compounds. This R-script separated the compounds that showed a normal growth curve and the bacteriostatic compound from the ones that show the characteristic induction curve (where max OD - end point OD is ≤ 0.01) (github/anishanandy/Plate_analyses). All primary screen hits were grouped based on whether they were known, canonical prophage inducers or non-canonical prophage inducers using another R-script. (github/anishanandy/Plate_analyses/circular-plot.R) (Appendix I).

2.7 Induction test based on compound concentration

The primary screen was performed at a singular concentration. For testing dose-dependent effects, a 150 µL of ~0.2 OD HK97 lysogen culture was added to each well of a 96-well plate. Four compounds, levofloxacin, meclocycline, thiamphenicol and domperidones were chosen for their primary screen hit OD drop and different methods of action. They were added at eight different concentrations from 0.375 µM to 50 µM with two-fold steps to look for any dose dependent phage induction. I measured the OD of each of the wells over 6 h with envision multiplate reader (PerkinElmer, MA).

2.8 Secondary screening of primary hits

2.8.1 Dose-dependent induction

For the secondary screen, primary screen hits were tested at five different concentrations ranging from 1.5 µM to 25 µM with two-fold steps and a no-compound-added control. To prepare an induction plate the compounds were transferred from a source plate to a sample 384-well plate containing ~0.2 OD_{600nm} lysogen culture using the Echo 555 liquid handler (Labcyte, CA).

The induction plates were incubated as they were in the primary screen (Section 2.3). After the incubation, the dead cell debris and any bacterial remnants were separated from the free-floating phages by centrifuging the 384-well plate at 15,000 x g for 10 min. The supernatant from the plates was then subjected to 10-fold serial dilutions to obtain dilutions from 10^0 to 10^{-4}.

2.8.2 Phage count scoring post-induction

The wildtype host bacterium was grown until ~0.2 OD_{600nm}. For creating a lawn, 750 µL of this culture was added to 7.5 mL of 0.7% LB agar and then poured on 1% LB agar Nunc Omnitray plate (ThermoFisher, MA). Each of the dilutions from the plates obtained in section 2.8.1 were then pinned using 384-well disposable replicator pins (V & P Scientific, CA) on the lawn of susceptible bacteria, *E. coli* 40 for Mu and *E. coli* K12 for HK97 and incubated overnight.

The plates were observed for plaques or cleared zones on both the replicates. The compounds were scored based on the baseline induction clearing of the culture only control. The compound concentrations resulting in phage counts 10 to 100-fold higher than the baseline were considered secondary screen hits. If compound resulted plaquing at two different dilutions across the replicates then highest one was recorded.

CHAPTER 3: RESULTS & DISCUSSION

3.1 HK97 & Mu lysogens serve as models for testing induction

To identify new signals prophages can respond to, I chose to study a traditional lambdoid prophage (HK97), and compare the findings with the control prophage Mu. Both HK97 and Mu are model temperate phages and infect *E. coli*. HK97 follows the canonical SOS-dependent phage induction mechanism, whereas Mu is considered uninducible. This allows us to directly compare our lysogens, as they are largely identical except for the phage within them, and the site of insertion of that phage.

Sequencing of the amplified phage-host junction confirmed that my HK97 lysogen, as expected (Alberts et al. 2002; Dhillon et al. 1980), contained the prophage integrated in the vicinity of the *gal* operon. The host for my control phage Mu is *E. coli* 40, a strain that is very closely related to the host for HK97. In contrast with HK97, Mu can insert anywhere in the genome of its host. Sequencing of the nested PCR product revealed the location of the Mu prophage, within the *ygbM* gene of *E. coli* 40. This disrupted the *ygbM* gene, predicted to encode of a truncated hydroxypyruvate isomerase. Disruption of this gene does not affect the growth rate or survival of the bacterium (Li et al. 2014). As neither prophage is integrated in a region that would be predicted to affect how the host reacts to stressors, our two lysogens are readily comparable—the only relevant difference between them being the prophages themselves.

In order to observe how the two lysogens responded to traditional induction assays, I performed growth curves of both the lysogen models as well as the HK97 host *E. coli* in the presence of ciprofloxacin, a known SOS-response inducer (Fig. 2). The MIC of ciprofloxacin was determined to be 0.025 µg/mL for the host bacterium. At sub-MIC levels, 0.0125 µg/mL (1/2MIC) and 0.0625 µg/mL (1/4MIC), the HK97 lysogen showed a characteristic induction curve with an initial rise in OD_{600nm} followed by an eventual crash (Fig. 2C). The same induction assay test was performed with the 'uninducible' Mu lysogen, which displayed no characteristic induction, and resembled the growth profile of the phage-free host *E. coli*. (Fig. 2B). The increased sensitivity of only the HK97 lysogen to ciprofloxacin, with its growth curve characteristic of induction, confirmed that our prophage models behaved as expected with canonical prophage inducers. This allowed me to proceed to tests with unknown compounds.

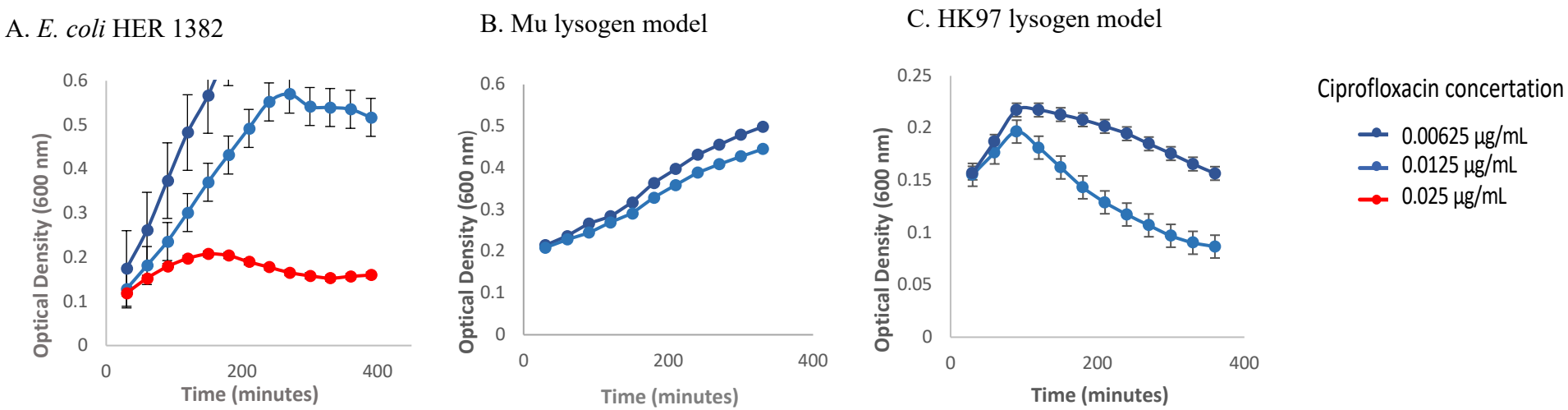


Figure 2: The canonical prophage inducer ciprofloxacin induces only the HK97 lysogen. Each panel represents a growth curve of *E. coli* over time, with each point being an average of three biological replicates. The error bars show the standard error in these triplicates. (A) The wild type was challenged with a range of concentrations of ciprofloxacin. The lowest concentration which inhibited growth was determined to be the MIC (red line). The lysogen models HK97 (C) and Mu (B) were challenged with 1/2MIC (Light Blue) or 1/4MIC (Dark Blue) ciprofloxacin. Only the HK97 lysogen (C) responded with a growth curve characteristic of phage induction.

3.2 Pilot HTS

Prior to high-throughput screens, I investigated parameters that might influence reproducibility of the assay in 384-well format. Validation tests performed with ciprofloxacin and DMSO added to HK97 lysogen culture at various culture densities revealed that an OD_{600nm} of 0.2 resulted in the most reproducible growth curves (Fig. S1). Testing eight different concentrations of ciprofloxacin in biological duplicates of HK97 model revealed high levels of reproducibility, with small standard errors lending confidence in the approach (Fig. S2).

The chosen compound library contains 3,936 compounds. To validate our high-throughput methodology, we performed a pilot study with one of the 384-well plates of compounds from the library against the HK97 lysogen. This plate of compounds was selected because of its high number of canonical prophage-inducing antibiotics.

3.2.1 IQM analysis of pilot screen

The pilot study highlighted reproducibility between replicates, as evidenced by the clustering of data points along the line y=x at all time points (Fig. 3). Based on the 3-SD cut off set up using a Z' and the interquartile method, the pilot run identified 13 out of the 320 compounds as 'hits'. As most of these are identified only near the end of the growth curve, I decided that future runs should extend observations a further 2 h. All the hits were antibiotics except one, phenylmercuric acetate. Encouragingly, this list of 'hits' included seven of the 10 fluoroquinolones tested, therefore reproducibly identifying known inducers.

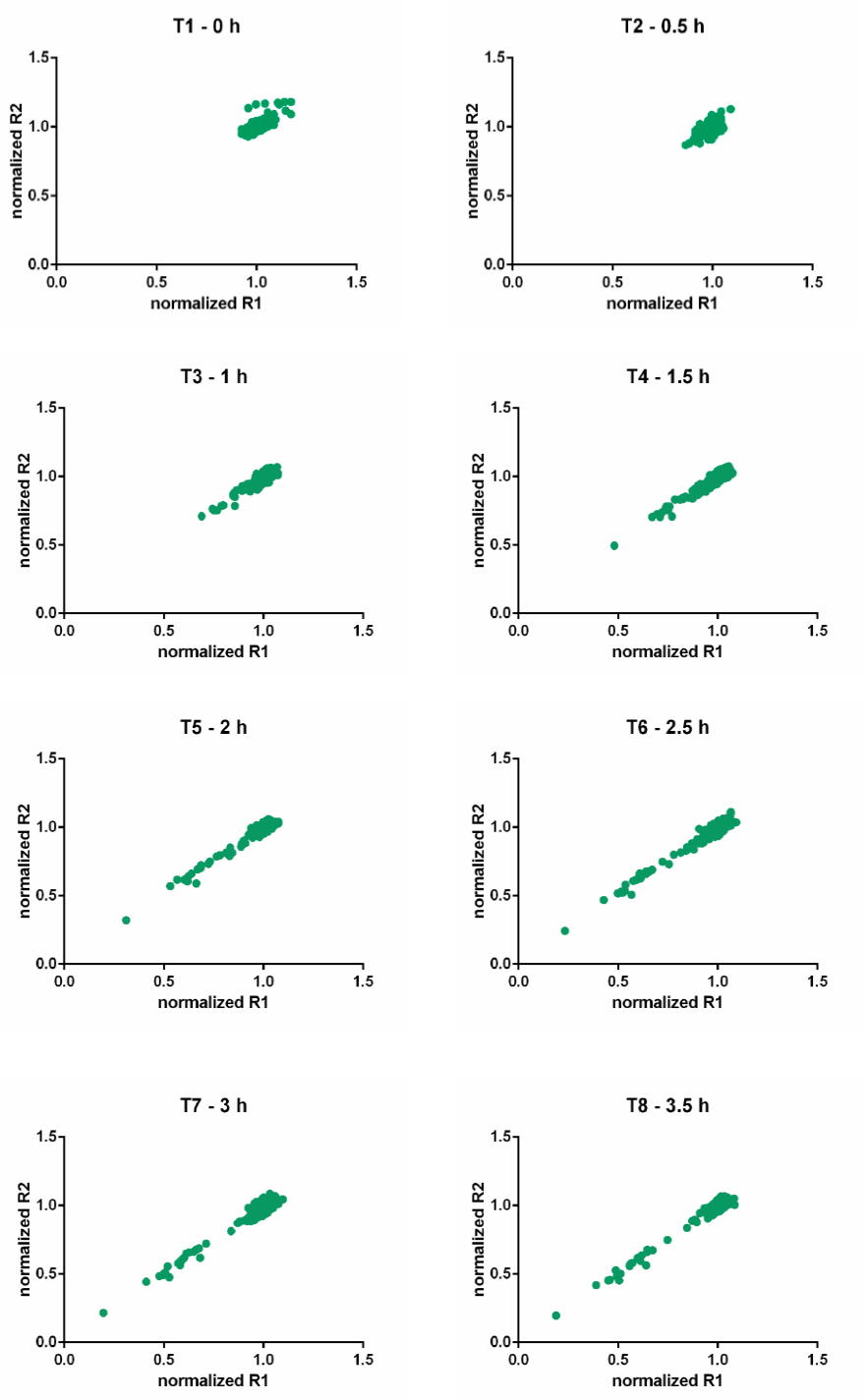


Figure 3: Pilot study shows detectable signals and reproducibility. OD_{600nm} was measured for all 384 wells, 320 of which had compounds with HK97 lysogen model for every 30 min from T1- 0 h to T8-3.5 h. Each of the replicate plots show the OD_{600nm} normalized to the plate mean at that time point, for each of the 384—wells.

My primary conceptual objection to the IQM method of identifying inducers was that it would consider bacteriostatic compounds as 'hits', because their endpoint reading would differ significantly from that of the plate mean. These 'flat lines indicating growth arrest would almost certainly be false positives, as they lack the rise-and-fall characteristic of induction. Instead I decided to track the shape of the growth curves for each well in order to identify primary hits.

3.2.2 Growth curve analysis of pilot screen

To avoid detecting bacteriostatic compounds as phage inducing compounds, the analysis was changed from endpoint OD to a continuous growth curve. The growth curves showed patterns typical of regular growth curves, bacteriostatic-growth arrest and induction-associated population crashes (Fig. 4 B). I developed an R-script to analyse the pilot screen that compares the final OD to the highest OD recorded for that curve. The threshold for a compound to be called a potential hit based on the results of the canonical compounds of the pilot study was set as follows—the difference between the maximum OD and the end point OD must be ≥ 0.01 (github/anishanandy/Plate_analyses). This threshold was selected based on the lowest OD drop from all canonical compounds that resulted as hits in the pilot screen. Any well that resulted in an endpoint OD ≥ 0.1 lower than the highest OD reached was called a 'primary hit' by this method (Fig. 4). This increased the likelihood that a compound deemed a 'hit' would be inducing the prophage.

On comparing the primary hits derived from my R-script with the hits I obtained with the IQM (Fig. 3), I determined that my result also found 14 partially overlapping with those from of the IQM analysis. The IQM analysis would have identified seven more compounds that were bacteriostatic as the hits. For a detailed listing of the hits on IQM and OD drop analysis see appendix II. This shows that analysing the entire growth curve rules out the likely false-positive bacteriostatic compounds.

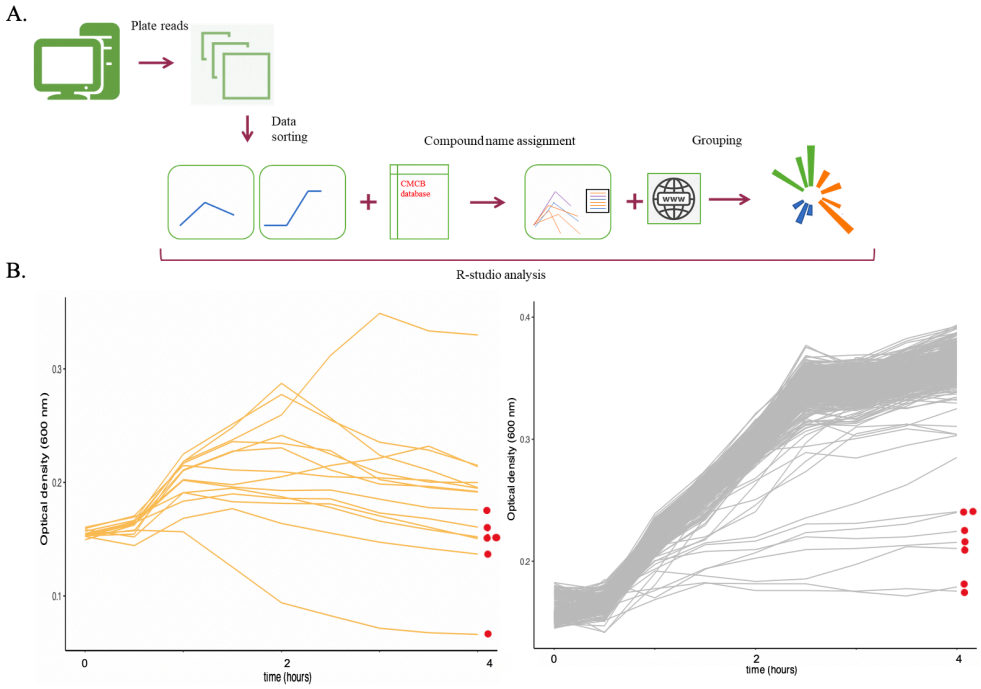


Figure 4: Growth curve analysis of the pilot screen data. A) Schematic of my separation of bacteriostatic compounds from phage inducing compounds for the pilot study. The data obtained from the plate reader were converted into .csv and edited to be read as a data frame in R-studio. The data were sorted into 'hits' and 'no hits' solely based on the drop of optical density. The CMCB database was used to assign names to the hits. The final grouping involved acquiring the functions of each bioactive from the source and graphing the OD_{600nm} drop. (B) All pilot run growth curves sorted into 'hits' (left, yellow) and 'no hits' (right, grey). The red dots indicate the 13 hits identified using the IQM method, for comparison.

3.3 Primary screen: non-canonical compounds show induction-like growth curves

To test the model phages with the full 3,936 compound library, I performed a high throughput primary screen on both lysogens, in duplicate. The primary screen results were analysed using the R-script as in the pilot screen. All compounds with an average OD drop ≥ 0.01 across replicates were also found to have had OD drops of ≥ 0.01 in each replicate.

The library contained 259 duplicate compounds. For instance, ciprofloxacin was present on two different plates in the library. Both plates tested with separately grown culture resulted in average OD drop across replicates of 0.06 and of 0.049—well above our threshold. Out of the 257 compounds repeated on two separate plates, 253 were classified as "no hits" and 4 were recognized as "hits" for each of the plates that they appeared on. This affirms the screening method and the pipeline used to analyze the data is reproducible, even across multiple days and separate plate preparations.

After generating the list of compounds showing induction-like growth curves, I classified them on the basis of canonical (fluoroquinolone and β-lactam) and non-canonical prophage inducers. The largest drops in OD were seen for compounds like cefixime and enoxacin, which are known canonical inducers of lambdoid prophages. The primary screen resulted in total of 175 compounds (171 unique) hits for the HK97 lysogen (Fig. 5A). The library also contains compounds that are known SOS activators and out of those compounds 88% of the cephalosporins and 84.7% of the fluoroquinolones in the compound collection were hits. These canonical inducers were

overrepresented in our lists, comprising only 4% of the library and 25% of our hits. However, hits with my R-script analysis were not limited to these compounds, and included 127 (126 unique) “non-canonical” compounds.

The primary screen also identified a few compounds as hits for the control, uninducible, phage Mu (Fig. 5B). This list consisted of 0.9% of the library (37 unique compounds), compared to the 4.4% (171 compounds) for the HK97 lysogen. Of these 37 primary hits, 24 (65%) were fluoroquinolones and β-lactams, suggesting that we are capturing some direct toxicity of the antibiotics. The remaining 13 non-canonical compounds identified were considered to be reflective of a false-positive rate in our screen.

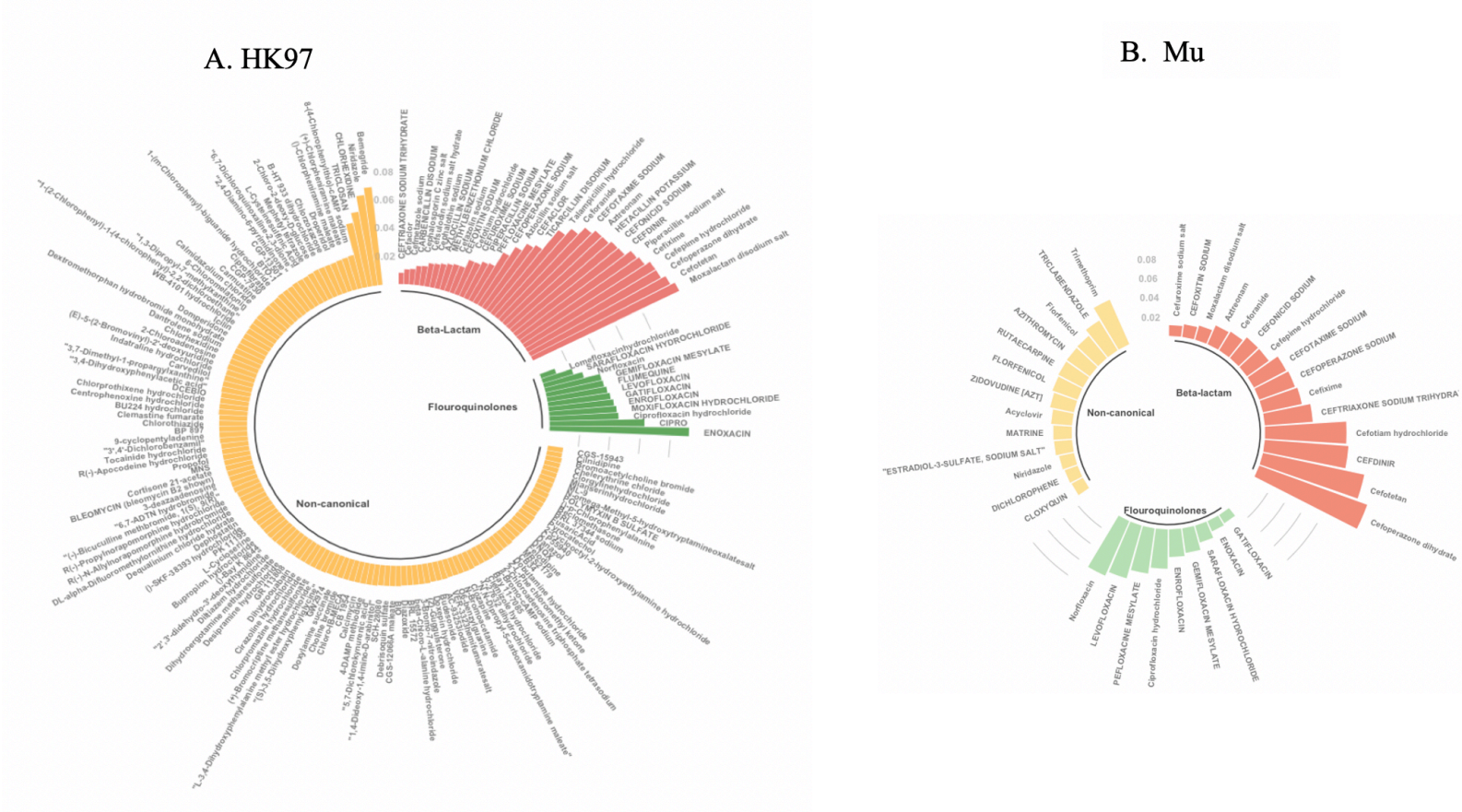


Figure 5: Primary screen hits for the two model lysogens. The height of the bar corresponds to the average magnitude of the drop in OD_{600nm} of the lysogen culture from the highest point the growth curve for that specific compound and the end point. Primary screen identified 171 unique 'hit' (175 shown here, including the duplicates) compounds for HK97 lysogen models (A) and 37 unique compounds as hits for Mu (B). Non-canonical compounds are in orange. The expected fluoroquinolones and β-lactams are shown in green and red respectively.

A comparison between the primary hits identified for both the HK97 lysogen and the control Mu lysogen revealed 23 overlapping compounds (Fig. 6). Only one non-antibiotic compound was common for both my lysogen models: niridazole. All other compounds overlapping both hitlists were either β-lactams or fluoroquinolones (Fig. 6), supporting my assumption that most of these Mu-prophage 'hits' reflect toxicity of the antibiotics independent of the prophage.

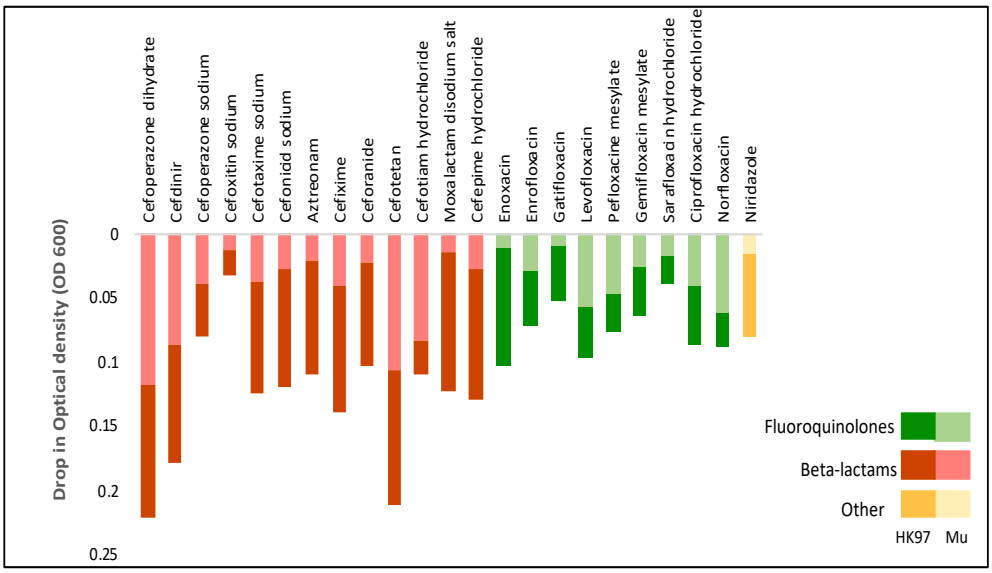


Figure 6: Overlaps in the primary hit list for the two model lysogens. The y-axis shows the average magnitude of OD drop (max OD - end point OD). The red bars show the effects of β-lactams and the green bars those of fluroquinolones. The darker shade is for HK97 and lighter one is for Mu. One single compound in yellow is the only common compound from the non-canonical group.

To verify that the exposure to these compounds actually induced the prophage in the model, we proceeded to a secondary screen testing our primary hits for increases in phage counts.

3.4 Secondary screen: model prophages are induced by non-canonical compounds

The results from the primary screen growth curve analysis use host cell lysis, detected through a drop in OD, as a proxy for prophage induction. If induction is occurring, the number of phages in the compound-exposed cultures should be higher than the baseline counts seen resulting from spontaneous induction. I chose to screen all primary 'hits' for their ability to produce increases in phage titres. Additionally, in a small-scale test of our different primary hits over a range of concentrations, HK97 prophage induction was dose specific (Fig S3). Phage induction is known to be a dose-dependent response (Goerke et al. 2006) (Fig. 2). I wished to determine if the effect of our compounds could be increased by optimizing the concentration of the suspected inducer. With only 208 compounds across both phages, I was also able to test a range of compound concentrations for their induction effects.

From the secondary screen stamp plates, I compared the plaques resulting from filtrates of culture only (e.g. baseline induction) with the plaques resulting from filtrates of cultures exposed to enoxacin (a fluoroquinolone) at five different concentrations. Enoxacin had the highest OD difference in the primary screen and resulted in a 10-fold to 100-fold increase in phage-mediated clearing of the susceptible bacterial lawn from that of the baseline induction of culture only control (Fig. 7). This measurement of increase in phage presence at lower dilutions of lysate was used to score all the primary hit compounds.

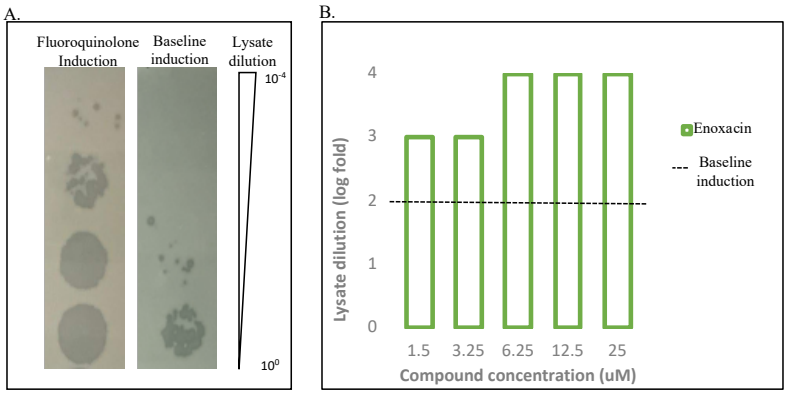


Figure 7: Increased phage production in fluoroquinolone induced cultures compared to baseline induction. A) Plaques counted for the culture-only tube that was grown for 4 h without any compound added (baseline induction) compared with the plaques of the lysates with fluoroquinolone added culture. B) Summarizing the spot test from A) the number of phages culture only control is represented by a dashed line where no compound was added to the HK97 lysogen. The green bars represent the phage counts in the culture post induction with enoxacin (fluoroquinolone) and removal of the host cell remnants. The difference in the counts is measured over five different doses of the compound. The highest difference is seen at 6.25 µM, 12.5 µM and 25 µM of enoxacin.

Phages plaquing on the susceptible bacterial lawn at much lower dilutions (10^{-4}) compared to of the baseline induction (10^{-2} for HK97, 10^{-1} for Mu) shows that the compound was able to initiate the lytic cycle of the phage and release more phage particles in the media. With HK97 lysogens, a total of 86 compounds showed an increased phage production, resulting in a hit rate from our primary screen to secondary screen of 50%. Of these, 37 were canonical phage inducers—β-lactams and fluoroquinolones — representing 84% of the primary-screen 'canonical' compounds (Fig S4). The remaining 49 compounds were non-canonical inducers, representing 38.5% of the primary screen hits in that category (Fig. 8). The historical reliance on DNA-damaging compounds to induce phages has clearly constrained the field — I show here

that there are many other compounds that can drive the prophage to induce out of its cell.

The hits for Mu were identified using the same scoring method as the for the HK97 model as from Figure 7. However, unlike HK97, there is no known inducer of Mu prophage, and I did not have an enoxacin-like compound to serve as a positive control. Unexpectedly, for my 'uninducible' control prophage Mu, the secondary screen showed increased phage production with seven different compounds (Fig. 7B). Although this represents only 18.9% of primary screen hits, Mu is not known to be induced by any external agent, not even by SOS-activating antibiotics. One of the hits was Cefoperazone sodium—a β-lactam which curiously did not induce HK97 prophage in my model, while the rest were non-canonical compounds.

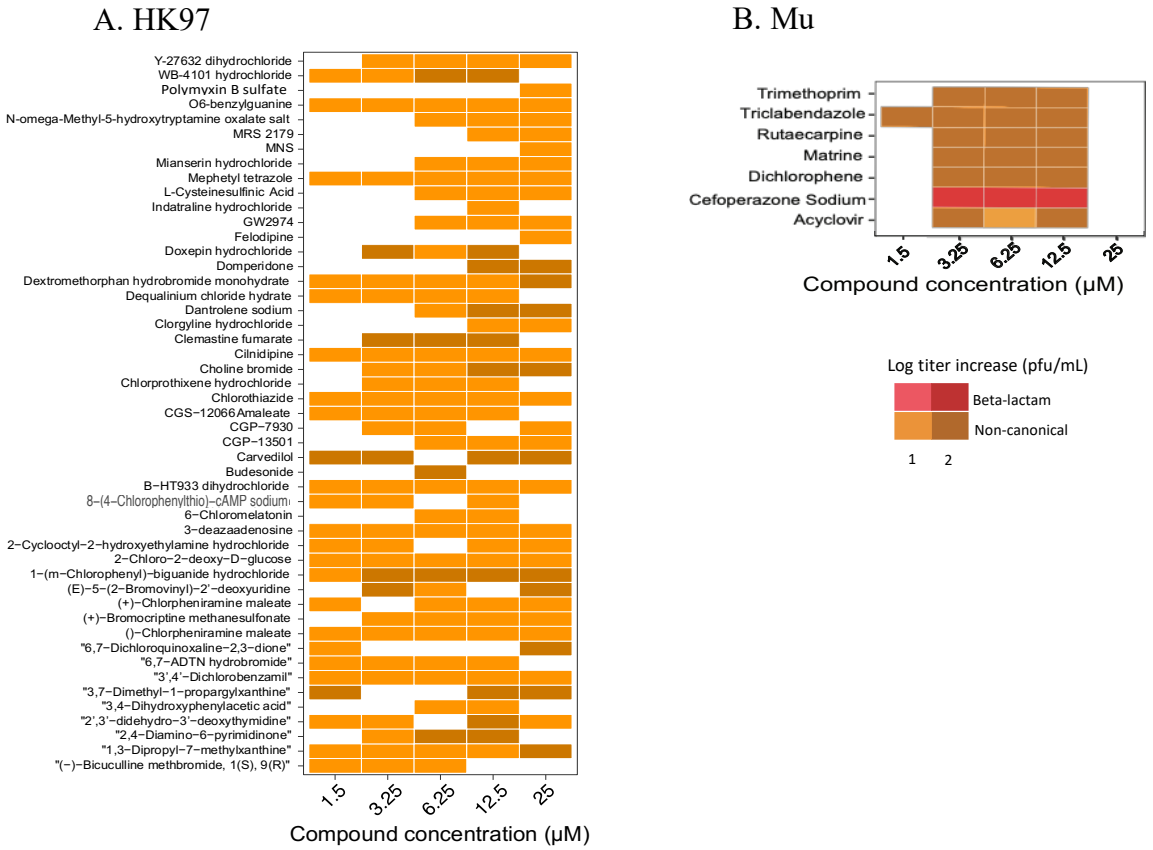


Figure 8: Non-canonical compounds increased phage production in the two model lysogens. The heatmap lists the compounds that showed increased phage production at any one of the concentrations from the range tested concentrations shown on the x-axis, in HK97 lysogen model (A) and Mu lysogen model (B) in the lysates after 6 h of incubation with the compounds. Each of the compounds showed high phage counts over a range of concentrations shown on the x-axis. The compound classes are highlighted in separate colours. Non-canonical compounds are in orange and the canonical β-lactam is shown in red. The increase in phage count is depicted by the intensity of colour.

This work is the first to record induction of Mu prophage as a response to a compound — and, as no compound was found to induce both HK97 and Mu, this further validates our choice of models. Furthermore, since host genome is almost identical, this supports a phage-specific mode of induction.

For the secondary screen with the Mu model, none of the compounds causes increased phage production at the highest concentration tested, 25 μM. This may reflect toxicity of the compounds to the bacterium, preventing successful induction dependent on functional host machinery. The contrast with HK97 may suggest that Mu is responding to compounds closer to their threshold of toxicity for the host. Additionally, one hit for Mu, Trimethoprim, is not from the β-lactam or fluoroquinolone class of compounds, but it has been found to activate the host SOS response (Lewin and Amyes 1991; Goerke et al. 2006). These compounds and the concentration stringency can also support the existence of different pathway inducing Mu prophage compared to that of HK97.

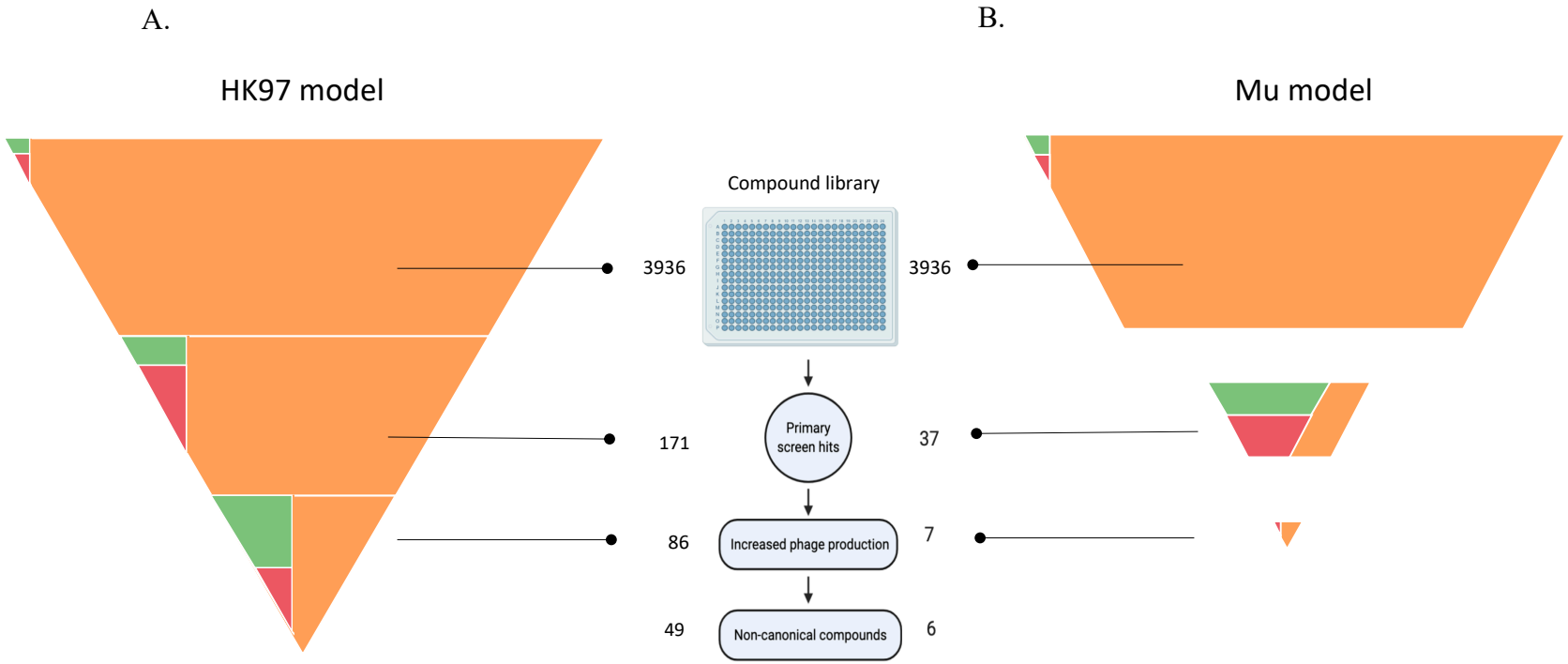


Figure 9: Summary of sequential screens across the two model lysogens. (A) HK97 model has higher number of hits in primary and secondary screens compared to that of (B) Mu lysogen model. Non-canonical inducing compounds are in orange. The fluoroquinolones and β-lactams are shown in green and red respectively. The area of a coloured shape within a horizontal tranche of the pyramid is proportional to the number of compounds represented by that category.

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

In summary (Fig. 9), I identified different 185 compounds that, at a singular concentration of 10 μM, caused crash in growth in our *E. coli* lysogen models host through the primary screen. A small overlapping subset of these (23, 22 of which were antibiotics) affected the host regardless of the prophage it carried, while the remaining 162 yielded phage-specific effects. Through a secondary screen across a wider range of concentrations, we confirmed that many of these—half for HK97 and a fifth for Mu—resulted in increased phage titres reflective of induction. None induced both model prophages. Of these compounds, 49 of the 86 inducing HK97 were not canonical phage inducers, and are not known to have phage-related activity (Fig. 8). All seven compounds resulting in Mu induction are of considerable interest, despite one being a canonical phage inducer, as Mu is traditionally considered uninducible. These 56 compounds with newfound phage-related activities will serve as tools to probe bacterial stress pathways, the prophage induction response, and potentially be used to identify new phages not inducible through traditional, DNA-damaging agents.

Although the compounds are not associated with phage induction, they may still be acting through canonical, SOS-induction-based pathways. This was the case in Oh et. al. 2019. To rule this out, we would repeat our screens in mutants unable to activate the SOS response, such as a *recA*$^{-}$ *E. coli*. Because such mutants are pleiotropic (Bianco 2001), we may be better by high-throughput phenomics (e.g. (French et al. 2018), screening a library of *E. coli* reporter-gene fusions upon exposure to the compounds to see whether the

compounds are resulting in activation of known SOS-signatures, or those of other stress pathways.

Many of our compounds are actively prescribed drugs, and our targeted bacterium, *E. coli* is native to the human gut. These compounds could easily already be causing phage driven microbial changes in the human gut microbiome (Maurice et al. 2013; Boling et al. 2020). Interestingly, six of our non-canonical HK97-inducing compounds were categorized as neuro-active — either dopaminergic or have serotonin-related functions (Table 3). Given that recent studies have shown that bacteria—commensal, probiotic, or pathogenic, in the gastrointestinal tract can activate neural pathways and central nervous signaling systems (Foster and McVey Neufeld 2013), I am curious to see if phage-mediated responses to neuroactive compounds might be contributing to their efficacy, modulating the microbiome through the gut-brain axis (Foster and McVey Neufeld 2013).

Table 3: Neuroactive compounds identified as inducers.

Compounds	**Effect**
Bromocriptine methane sulfonate	Dopaminergic
CGS-12066A maleate	Serotonin
Chlorprothixene hydrochloride	Dopaminergic
Mianserin hydrochloride	Serotonin
N-omega-Methyl-5-hydroxytryptamine oxalate	Serotonin
Domperidone	Dopaminergic

Finally, while most bacteria carry prophages (Touchon et al. 2016), many are thought to be defective remnants of past infections. This

categorization is largely based on the failure of these prophages to respond to canonical inducing agents such as ciprofloxacin or mitomycin C (Campbell 1998). Non-canonical triggers of induction — especially those seven able to induce the traditionally uninducible prophage Mu could lead to the discovery of previously undetectable classes of phages.

REFERENCES

Abedon, S.T., 2015. Bacteriophage secondary infection. Virologica Sinica, 30(1), pp.3-10.

Adamus-Białek, W., Wawszczak, M., Arabski, M., Majchrzak, M., Gulba, M., Jarych, D., Parniewski, P., and Głuszek, S. 2019. Ciprofloxacin, amoxicillin, and aminoglycosides stimulate genetic and phenotypic changes in uropathogenic *Escherichia coli* strains. Virulence 10(1): 260–276.

Akeju, M.O., McChesney, J.D., and Williamson, J.S. 1998. Effect of polyphenolic compounds on a λ-prophage induction assay. Nat. Prod. Lett. 11(3): 201–205.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. 2002. Studying gene expression and function. In Molecular Biology of the Cell. 4th edition. Garland Science

Anderson, W.A., Moreau, P.L., Devoret, R., and Maral, R. 1980. Induction of prophage λ by daunorubicin and derivatives correlation with antineoplastic activity. Mutat. Res. Toxicol. 77(3): 197–208.

Bearson, B.L., and Brunelle, B.W. 2015. Fluoroquinolone induction of phage-mediated gene transfer in multidrug-resistant *Salmonella*. Int. J. Antimicrob. Agents 46(2): 201–204.

Bianco, P. R. (2001). RecA protein. *eLS*, 1-12.

Blondeau, J. M. 2004. Fluoroquinolones: mechanism of action, classification, and development of resistance. Surv. ophthalmol., 49(2), S73-S78.

Boling, L., Cuevas, D.A., Grasis, J.A., Kang, H.S., Knowles, B., Levi, K., Maughan, H., McNair, K., Rojas, M.I., Sanchez, S.E., Smurthwaite, C., and Rohwer, F. 2020. Dietary prophage inducers and antimicrobials: toward landscaping the human gut microbiome. Gut Microbes: 1–14.

Bukhari, A.I. 1975. Reversal of mutator phage mu integration. J. Mol. Biol. 96(1): 87–99.

Campbell, A. M. (1996). Cryptic prophages. *Escherichia coli* and *Salmonella:* cellular and molecular biology. ASM Press, Washington*, DC*, 2041-204

Casjens, S. 2003. Prophages and bacterial genomics: what have we learned so far? Mol. Microbiol. 49(2): 277–300.

Casjens, S.R., and Hendrix, R.W. 2015. Bacteriophage lambda: early pioneer and still relevant. Virology 479–480: 310–30.

Delhaye, A., Collet, J.0F., and Laloux, G. 2019. A fly on the wall: how stress response systems can sense and respond to damage to peptidoglycan. Front. Cell. Infect. Microbiol. 9: 380.

Dhillon, E.K.S., Dhillon, T.S., Lai, A.N.C., and Linn, S. 1980. Host range, immunity and antigenic properties of lambdoid coliphage HK97. J. Gen. Virol. 50(1): 217–220.

Elespuru, R.K., and Yarmolinsky, M.B. 1979. A colorimetric assay of lysogenic induction designed for screening potential carcinogenic and carcinostatic agents. Environ. Mutagen. 1(1): 65–78.

Elespuru, R. K., and White, R. J. 1983. Biochemical prophage induction assay: a rapid test for antitumor agents that interact with DNA. Cancer research, 43(6), 2819-2830.

Erez, Z., Steinberger-Levy, I., Shamir, M., Doron, S., Stokar-Avihail, A., Peleg, Y., Melamed, S., Leavitt, A., Savidor, A., Albeck, S. and Amitai, G., 2017. Communication between viruses guides lysis–lysogeny decisions. *Nature*, *541*(7638), pp.488-493.

Faruque, S. M., & Mekalanos, J. J. (2012). Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence*, *3*(7), 556-565.

Feiner, R., Argov, T., Rabinovich, L., Sigal, N., Borovok, I., and Herskovits, A.A. 2015. A new perspective on lysogeny: prophages as active regulatory switches of bacteria. Nat. Rev. Microbiol. 13(10): 641–650.

Foster, J.A., and McVey Neufeld, K.A. 2013. Gut-brain axis: how the microbiome influences anxiety and depression. Trends neurosci, 36(5), 305-312.

French, S., Coutts, B.E., and Brown, E.D. 2018. Open-source high-throughput phenomics of bacterial promoter-reporter strains. Cell Syst. 7(3): 339-346.e3.

Goerke, C., Koller, J., and Wolz, C. 2006. Ciprofloxacin and trimethoprim cause phage induction and virulence modulation in *Staphylococcus aureus*. Antimicrob. Agents Chemother. 50(1): 171–177.

Griffiths, A. J., Wessler, S. R., Lewontin, R. C., Gelbart, W. M., Suzuki, D. T., & Miller, J. H. (2005). An introduction to genetic analysis. Macmillan.

Hatfull, G.F., and Hendrix, R.W. 2011. Bacteriophages and their genomes. Curr. Opin. Virol. 1(4): 298–303.

Heinemann, B. 1971. Prophage induction in lysogenic bacteria as a method of detecting potential mutagenic, carcinogenic, carcinostatic, and teratogenic agents. In Chemical Mutagens pp. 235-266.

Hendrix, R. W. (2005). Bacteriophage HK97: assembly of the capsid and evolutionary connections. Advances in virus research, 64, 1-14

Howard-Varona, C., Hargreaves, K.R., Abedon, S.T., and Sullivan, M.B. 2017. Lysogeny in nature: mechanisms, impact and ecology of temperate phages. ISME J. 11(7): 1511–1520.

Janion, C. 2008. Inducible SOS response system of DNS repair and mutagenesis in *Escherichia coli*. Int. J. Biol. Sci: 4.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C. and Thierer, T., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), pp.1647-1649.

Keen, E.C., and Dantas, G. 2018. Close encounters of three kinds: bacteriophages, commensal bacteria, and host immunity. Trends Microbiol. 26(11), 943-954.

Kimura, S., Sonoda, K., Yamane, S., Maeda, H., Matsumura, K., and Hatakeyama, M. 2008. Function approximation approach to the inference of reduced NGnet models of genetic networks. BMC Bioinformatics 9(1): 23.

Knowles, B., Silveira, C.B., Bailey, B.A., Barott, K., Cantu, V.A., Cobián-Güemes, A.G., Coutinho, F.H., Dinsdale, E.A., Felts, B., Furby, K.A.,

George, E.E., Green, K.T., Gregoracci, G.B., Haas, A.F., Haggerty, J.M., Hester, E.R., Hisakawa, N., Kelly, L.W., Lim, Y.W., Little, M., Luque, A., McDole-Somera, T., McNair, K., de Oliveira, L.S., Quistad, S.D., Robinett, N.L., Sala, E., Salamon, P., Sanchez, S.E., Sandin, S., Silva, G.G.Z., Smith, J., Sullivan, C., Thompson, C., Vermeij, M.J.A., Youle, M., Young, C., Zgliczynski, B., Brainard, R., Edwards, R.A., Nulton, J., Thompson, F., and Rohwer, F. 2016. Lytic to temperate switching of viral communities. Nature 531(7595): 466–470.

Lander, G.C., Evilevitch, A., Jeembaeva, M., Potter, C.S., Carragher, B., and Johnson, J.E. 2008. Bacteriophage lambda stabilization by auxiliary protein GPD: timing, location, and mechanism of attachment determined by cryo-em. Structure 16(9): 1399–1406.

Lewin, C.S., and Amyes, S.G.B. 1991. The role of the SOS response in bacteria exposed to zidovudine or trimethoprim. J. Med. Microbiol. 34(6): 329–332.

Li, J.J., Lee, C.S., Sheng, J.F., and Doi, Y. 2014. Complete sequence of a conjugative incn plasmid harboring blaKPC-2, blaSHV-12, and qnrS1 from an *Escherichia coli* sequence type 648 strain. Antimicrob. Agents Chemother. 58(11): 6974–6977.

López, E., Domenech, A., Ferrándiz, M.J., Frias, M.J., Ardanuy, C., Ramirez, M., García, E., Liñares, J., and de la Campa, A.G. 2014. Induction of prophages by fluoroquinolones in *Streptococcus pneumoniae*: implications for emergence of resistance in genetically-related clones. PLoS One 9(4): e94358.

Maiques, E., Úbeda, C., Campoy, S., Salvador, N., Lasa, Í., Novick, R.P., Barbé, J., and Penadés, J.R. 2006. β-lactam antibiotics induce the SOS response and horizontal transfer of virulence factors in *Staphylococcus aureus*. J. Bacteriol. 188(7): 2726–2729.

Majtanova, L., Hostacka, A., & Majtan, V. (1994). Effects of subinhibitory concentrations of antibiotics on biological properties of *Salmonella typhimurium*. Folia microbiologica, *39*(2), 141-146

Mangat, C.S., Bharat, A., Gehrke, S.S., and Brown, E.D. 2014. Rank ordering plate data facilitates data visualization and normalization in high-throughput screening. J. Biomol. Screen. 19(9): 1314–1320.

Maurice, C.F., Haiser, H.J., and Turnbaugh, P.J. 2013. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. Cell 152(1–2): 39–50.

Michel, B. 2005. After 30 years of study, the bacterial SOS response still surprises us. PLoS Biol. 3(7): e255.

Miller, C., Thomsen, L.E., Gaggero, C., Mosseri, R., Ingmer, H. and Cohen, S.N., 2004. SOS response induction by ß-lactams and bacterial defense against antibiotic lethality. Science, 305(5690), pp.1629-1631.

Nanda, A.M., Heyer, A., Krämer, C., Grünberger, A., Kohlheyer, D., and Frunzke, J. 2014. Analysis of SOS-induced spontaneous prophage induction in *Corynebacterium glutamicum* at the single-cell level. J. Bacteriol. 196(1): 180–8.

Nanda, A. M., Thormann, K., and Frunzke, J. 2015. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. J. Bacteriol., 197(3), 410-419

Oh, J.H., Alexander, L.M., Pan, M., Schueler, K.L., Keller, M.P., Attie, A.D., Walter, J., and van Pijkeren, J.P. 2019. Dietary fructose and microbiota-derived short-chain fatty acids promote bacteriophage production in the gut symbiont *Lactobacillus reuteri*. Cell Host Microbe 25(2): 273-284.e6.

Ptashne, M. 2004. A genetic switch : phage lambda revisited. CSHL press.

Radman, M. 1975. SOS repair hypothesis: phenomenology of an inducible DNA repair which is accompanied by mutagenesis. Basic Life Sci. 5A: 355–67.

Rodríguez-Tébar, A., and Vázquez, D. 1984. Penicillin-binding proteins and peptidoglycan peptide-interacting proteins. Microbiol. Sci. 1(9): 211–4.

Rozanov, D. V, D'Ari, R., and Sineoky, S.P. 1998. RecA-independent pathways of lambdoid prophage induction in *Escherichia coli*. J. Bacteriol. 180(23): 6306–6315.

Sauer, R.T., Ross, M.J., and Ptashne, M. 1982. Cleavage of the lambda and p22 repressors by RecA protein. J. Biol. Chem. 257(8): 4458–4462.

Silpe, J.E., and Bassler, B.L. 2019. A host-produced quorum-sensing autoinducer controls a phage lysis-lysogeny decision. Cell 176(1–2): 268-280.e13.

Simmons, L.A., Foti, J.J., Cohen, S.E., and Walker, G.C. 2008. The SOS regulatory network. EcoSal Plus 2008.

Tipper, D.J., and Strominger, J.L. 1965. Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. Proc. Natl. Acad. Sci. U.S.A. 54(4): 1133–1141.

Touchon, M., Bernheim, A., and Rocha, E.P.C. 2016. Genetic and life-history traits associated with the distribution of prophages in bacteria. ISME J. 10(11): 2744–2754.

Villanueva, R.A.M., Chen, Z.J., and Wickham, H. 2016. Ggplot2: elegant graphics for data analysis using the grammar of graphics. Springer

Walterspiel, J.N., Morrow, A.L., Cleary, T.G., and Ashkenazi, S. 1992. Effect of subinhibitory concentrations of antibiotics on extracellular shiga-like toxin I. Infection 20(1): 25–29.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. 2019. Welcome to the tidyverse. J. Open Source Softw. 4(43): 1686.

Zeng, X., and Lin, J. 2013. Beta-lactamase induction and cell wall metabolism in gram-negative bacteria. Front microbiol, 4, 128

Žgur-Bertok, D. 2013. DNA damage repair and bacterial pathogens. PLoS Pathog. 9(11): e1003711.

SUPPLEMENTARY FIGURES

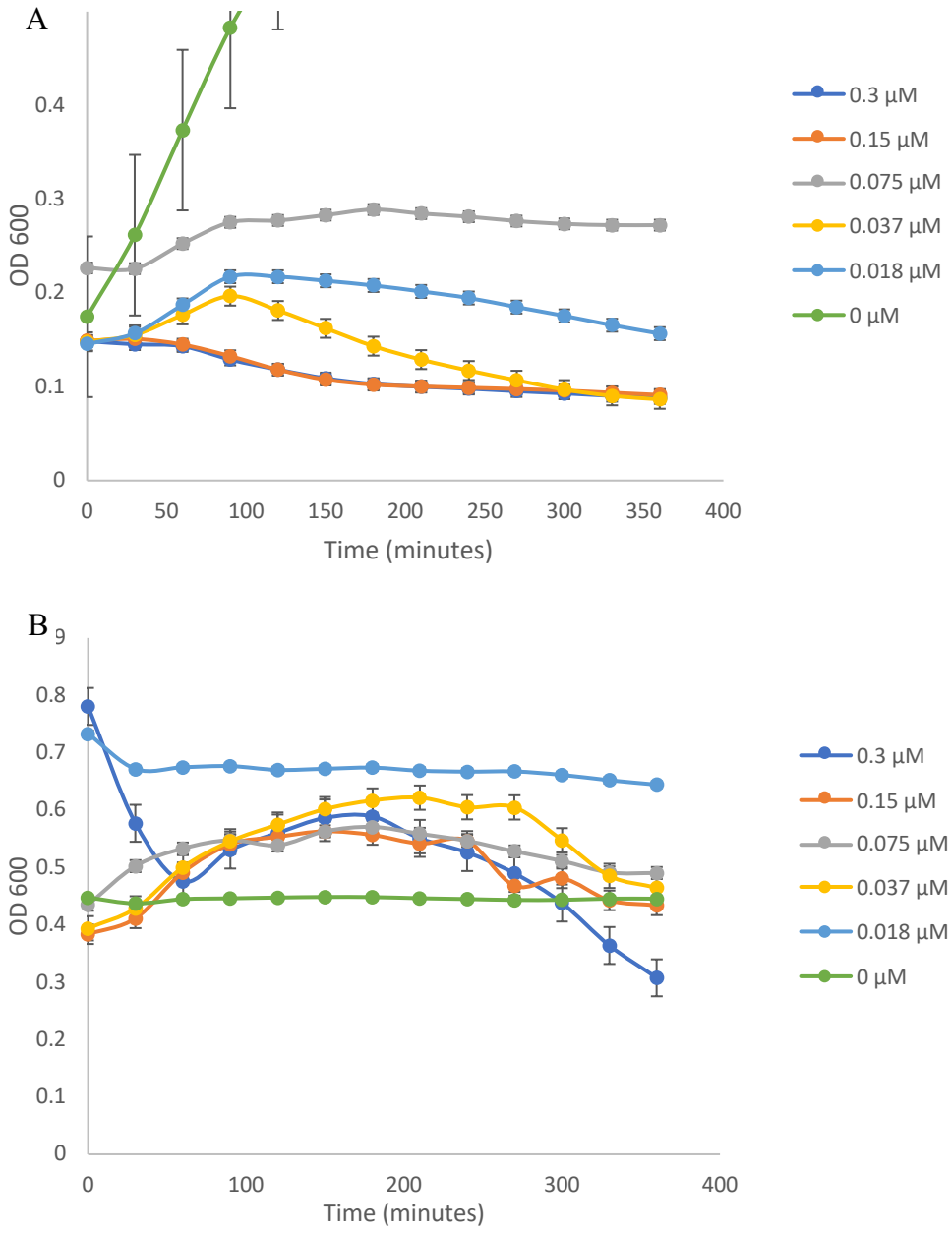


Figure S1: OD_{600nm} of 0.2 is optimal for the addition of a compound to study induction. A) HK97 lysogen at OD 0.2 challenged with sub-MIC concentration of ciprofloxacin. The concentrations were 2-fold increments starting from 0.00625 µg/mL to 0.1 µg/mL B) Same as A except HK97 lysogen culture was at OD 0.4 and the two lines starting at a higher OD may be because of bubbles. Each line is an average of biological triplicate. The error bars represent standard error. 0 µg/mL was no compound added culture only control.

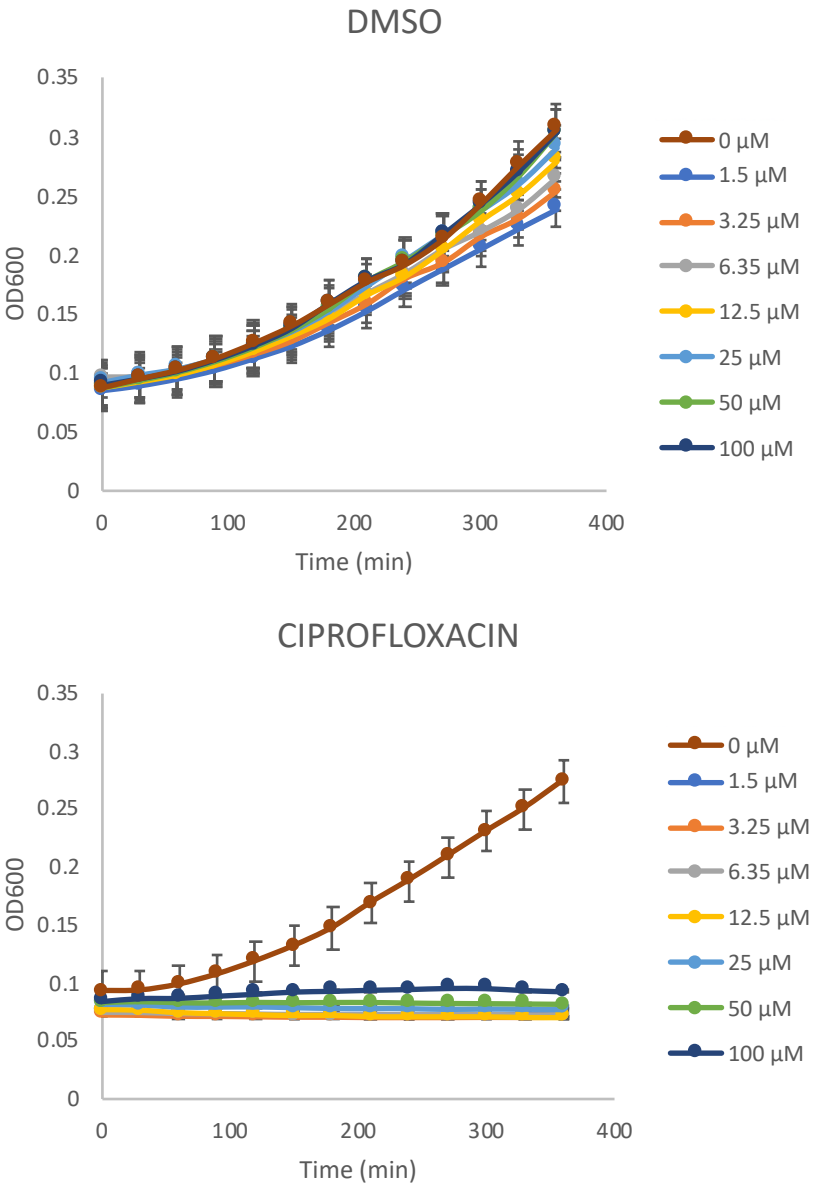


Figure S2: Reproducibility in 384-well analysis of lysogen culture. The ciprofloxacin data for lower concentration is shown in Figure 3c). HK97 lysogen at OD 0.1 challenged eight different concentration of ciprofloxacin (A) and DMSO (B). The concentrations were two-fold increments starting from 1.5 µM to 100 µM Each line is an average of biological triplicate. The error bars represent standard error. 0 µg/mL was no compound added culture only control.

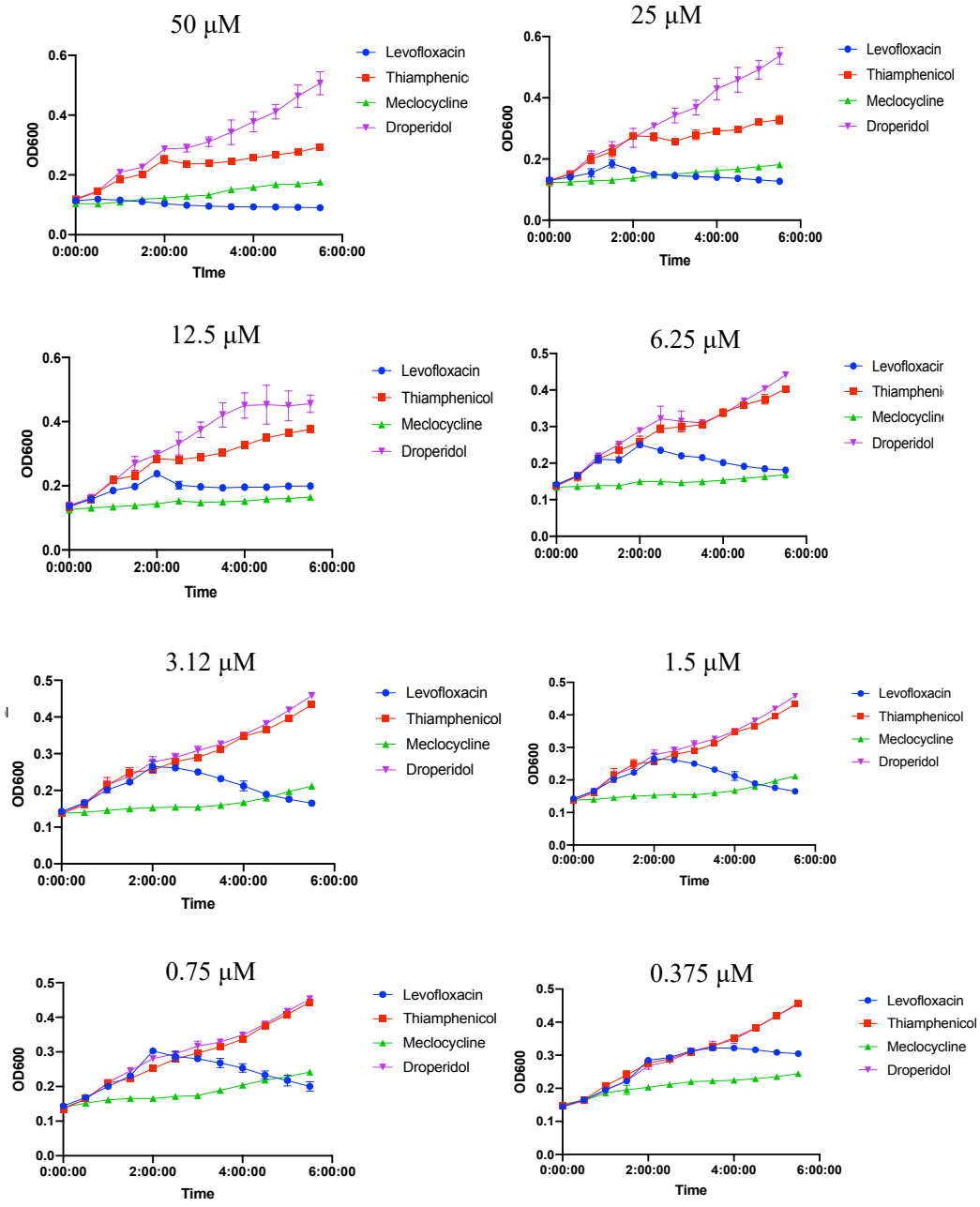
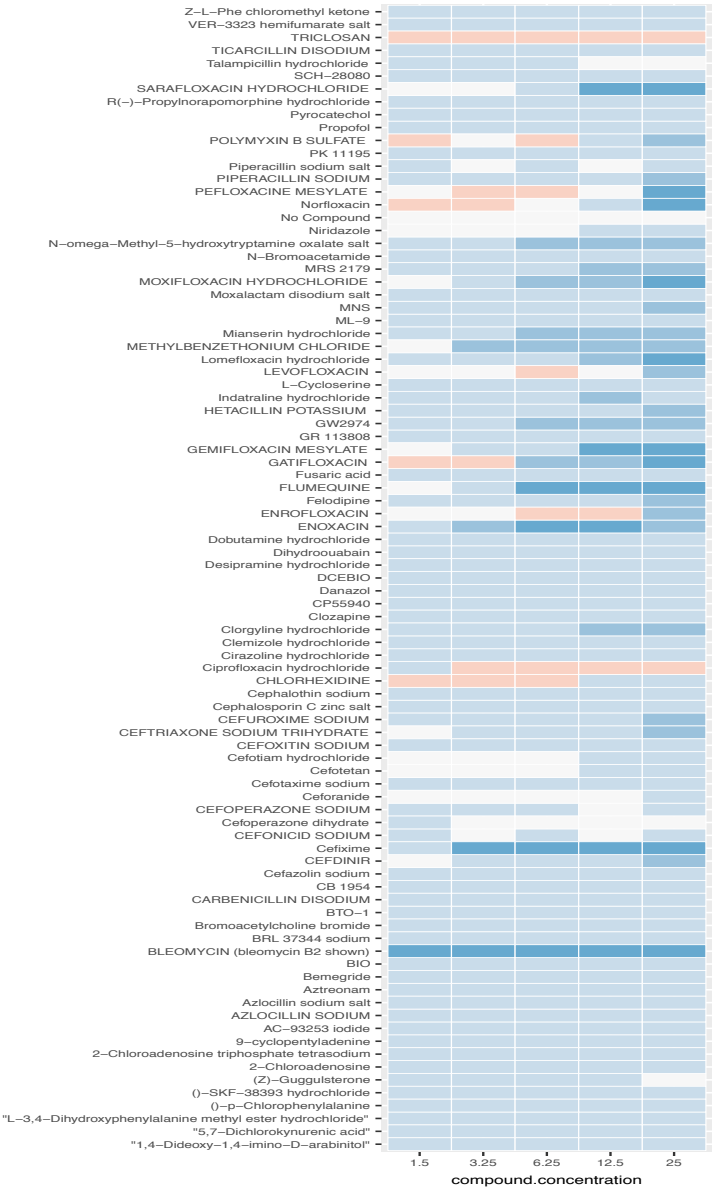


Figure S3: Four selected compounds show dose-dependent effect on HK97 lysogen. Compounds were added to the culture after it was grown it ~ 0.15 and the optical density for each of these compounds was measured over 6 h after the compound additions. Eight different concentrations were tested with two-fold increments. All these lines are averaged of biological triplicates and the standard error is represented by error bars.



(continued on next page)

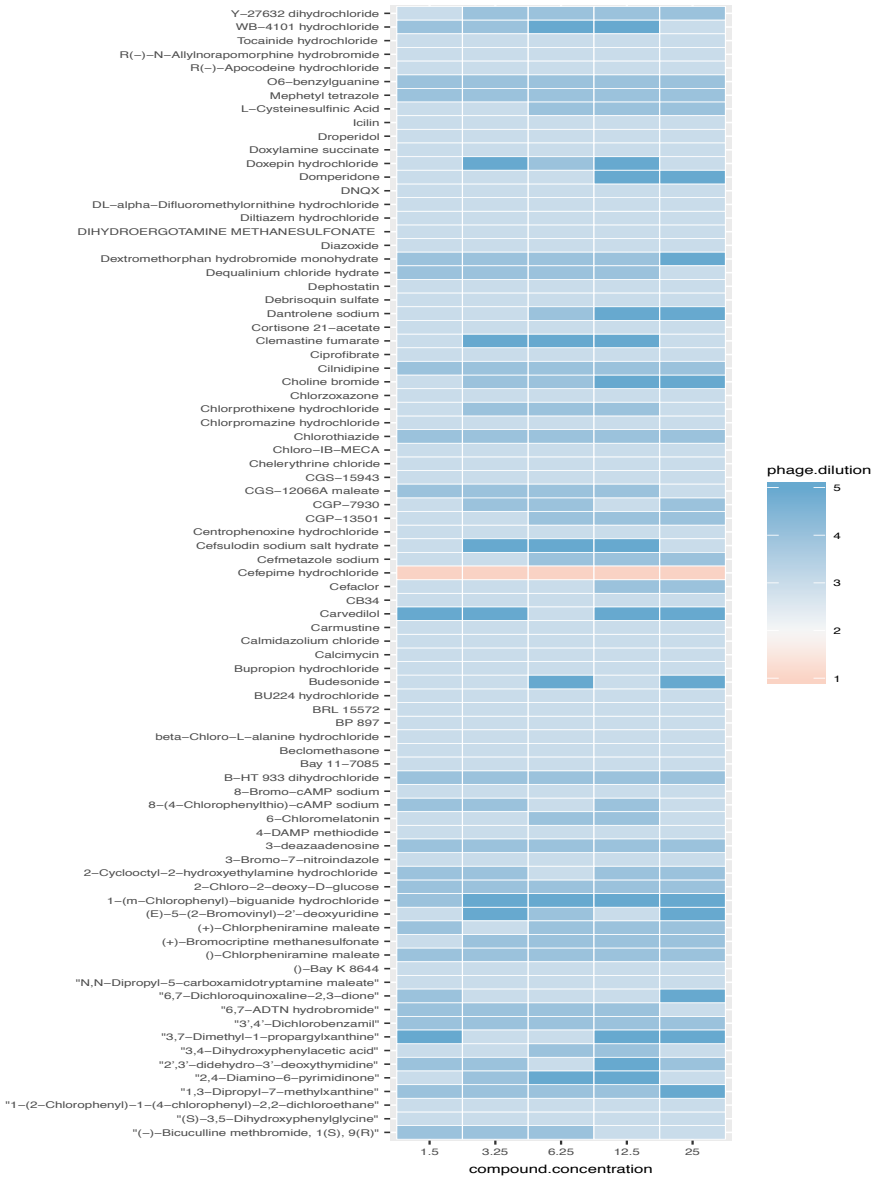


Figure S4: HK97 secondary screen raw data including the SOS-activating compounds. The heatmap lists all the compounds tested over the range of concentrations shown on the x-axis with HK97 lysogen model, in the lysates after 6 h of incubation with the compounds. Pink represent the ones that showed less phage production than the baseline induction. classes are highlighted in separate colours. The phage count is depicted by the intensity of colour.

APPENDIX I: SCRIPTS USED

All the scripts for this study were written in R and can be found on

https://github/anishanandy/Plate_analyses.

Separating hits vs no hits

```
#####
Library
Calls
#####

library("tidyverse")
library("reshape")



##### Initial input section #####

#The original datafile in .csv format
df <—read.csv("data/HTSplate_3.csv", header = FALSE)

#The number of time points taken
timepoints = 9

#The number of technical replicates
replicate = 2

#The number of readings per hour
read.hour = 1

#The Treshold for significance for variations
treshold = 0.01

#The path to the csv file that contains the list of all variation >
than the threshold
Non0_Ouput <—"NonZero_ODVariationttestalex.csv"


##### Data wrangling section #####

#Removes the empty lines in the .csv files
df %>%
 select(-1) %>%
 filter(!is.na(V2))—> df2
```

```
#Make a list of nested dataframes for each time point
chunk <- 17
n <—nrow(df2)
r <—rep(1:ceiling(n/chunk),each=chunk)[1:n]
list_df <- split(df2,r)
remove(n)
remove(r)
remove(chunk)
remove(df2)

#Rename the nested dataframes with the time point and replicate
names.1 <—sprintf("t%d",rep(seq(1:timepoints),
each=replicate))
names.2 <—rep(sprintf(".%d", seq(1:replicate)),timepoints)
fill <—paste0(names.1, names.2)
names(list_df) <—fill
remove(names.1)
remove(names.2)

#Remove first row from each dataframes
list_df <—lapply(list_df, function(x) {
 slice(x, 2:17)
})

#Make dataframes into single column
list_df_single <—lapply(list_df, function(x) {
 x <—data.frame(unlist(x))
})

#Creating the vertical ID column
ID.V <— c("A01", "B01", "C01", "D01", "E01", "F01", "G01",
"H01", "I01", "J01", "K01", "L01", "M01",
 "N01", "O01", "P01", "A02", "B02", "C02", "D02", "E02",
"F02", "G02", "H02", "I02", "J02",
 "K02", "L02", "M02", "N02", "O02", "P02", "A03", "B03",
"C03", "D03", "E03", "F03", "G03",
 "H03", "I03", "J03", "K03", "L03", "M03", "N03", "O03",
"P03", "A04", "B04", "C04", "D04",
 "E04", "F04", "G04", "H04", "I04", "J04", "K04", "L04",
"M04", "N04", "O04", "P04", "A05",
 "B05", "C05", "D05", "E05", "F05", "G05", "H05", "I05",
"J05", "K05", "L05", "M05", "N05",
```

"O05", "P05", "A06", "B06", "C06", "D06", "E06", "F06",
"G06", "H06", "I06", "J06", "K06",
"L06", "M06", "N06", "O06", "P06", "A07", "B07", "C07",
"D07", "E07", "F07", "G07", "H07",
"I07", "J07", "K07", "L07", "M07", "N07", "O07", "P07",
"A08", "B08", "C08", "D08", "E08",
"F08", "G08", "H08", "I08", "J08", "K08", "L08", "M08",
"N08", "O08", "P08", "A09", "B09",
"C09", "D09", "E09", "F09", "G09", "H09", "I09", "J09",
"K09", "L09", "M09", "N09", "O09",
"P09", "A10", "B10", "C10", "D10", "E10", "F10", "G10",
"H10", "I10", "J10", "K10", "L10",
"M10", "N10", "O10", "P10", "A11", "B11", "C11", "D11",
"E11", "F11", "G11", "H11", "I11",
"J11", "K11", "L11", "M11", "N11", "O11", "P11", "A12",
"B12", "C12", "D12", "E12", "F12",
"G12", "H12", "I12", "J12", "K12", "L12", "M12", "N12",
"O12", "P12", "A13", "B13", "C13",
"D13", "E13", "F13", "G13", "H13", "I13", "J13", "K13",
"L13", "M13", "N13", "O13", "P13",
"A14", "B14", "C14", "D14", "E14", "F14", "G14", "H14",
"I14", "J14", "K14", "L14", "M14",
"N14", "O14", "P14", "A15", "B15", "C15", "D15", "E15",
"F15", "G15", "H15", "I15", "J15",
"K15", "L15", "M15", "N15", "O15", "P15", "A16", "B16",
"C16", "D16", "E16", "F16", "G16",
"H16", "I16", "J16", "K16", "L16", "M16", "N16", "O16",
"P16", "A17", "B17", "C17", "D17",
"E17", "F17", "G17", "H17", "I17", "J17", "K17", "L17",
"M17", "N17", "O17", "P17", "A18",
"B18", "C18", "D18", "E18", "F18", "G18", "H18", "I18",
"J18", "K18", "L18", "M18", "N18",
"O18", "P18", "A19", "B19", "C19", "D19", "E19", "F19",
"G19", "H19", "I19", "J19", "K19",
"L19", "M19", "N19", "O19", "P19", "A20", "B20", "C20",
"D20", "E20", "F20", "G20", "H20",
"I20", "J20", "K20", "L20", "M20", "N20", "O20", "P20",
"A21", "B21", "C21", "D21", "E21",
"F21", "G21", "H21", "I21", "J21", "K21", "L21", "M21",
"N21", "O21", "P21", "A22", "B22",
"C22", "D22", "E22", "F22", "G22", "H22", "I22", "J22",
"K22", "L22", "M22", "N22", "O22",
"P22", "A23", "B23", "C23", "D23", "E23", "F23", "G23",
"H23", "I23", "J23", "K23", "L23",

```
 "M23", "N23", "O23", "P23", "A24", "B24", "C24", "D24",
"E24", "F24", "G24", "H24", "I24",
 "J24", "K24", "L24", "M24", "N24", "O24", "P24")
tAll.raw <—as.data.frame(ID.V)

#Merge dataframes into 1
list_ID <—lapply(list_df_single, function(x){
 as.data.frame(append(tAll.raw, x))
})

merge.all <—function(x, y) {
 merge(x, y, all=TRUE, by="ID.V")
}

tAll.raw <- Reduce(merge.all, list_ID)
colnames(tAll.raw)[1] <- "ID.H"
ID.H <— as.vector(tAll.raw$ID.H)
names.1 <—sprintf("t%d",rep(seq(1:timepoints),
each=replicate))
names.2 <- rep(sprintf(".%d", seq(1:replicate)),timepoints)
dataset.vector <—paste0(names.1, names.2)
fill <—append("ID.H", dataset.vector)
colnames(tAll.raw) <—fill
remove(merge.all)
remove(fill)
remove(list_ID)
remove(list_df)



##### Data analysis section #####

#Obtain the dataframe of the averages
as.data.frame(within(tAll.raw, {
 pair.colmeans <—sapply(seq(2, ncol(tAll.raw), 2), function(i) {
 rowMeans(tAll.raw[, c(i, i+1)], na.rm=TRUE)
 })
 }))—> intermediate
as.data.frame(intermediate$pair.colmeans, drop = F)—> tAll
colnames(tAll) <—(sprintf("t%d",rep(seq(1:timepoints))))
row.names(tAll) <—tAll.raw$ID.H
remove(intermediate)

#Transpose and melt the dataframe for easier graphing
```

```
time.vector <—seq(from = 0, to = (timepoints/read.hour)—
(1/read.hour), by = (1/read.hour))
tAll.t <—as.data.frame(t(tAll))
tAll.t %>%
 mutate(time = time.vector) %>%
 select(time, A01:P24)—> tAll.t
tAll.long <—melt(tAll.t, id="time")

#Compute the difference between the highest and last point
tAll.t %>%
 select(A01:P24)—> tAll_diff

colMax <—function(data) sapply(data, max, na.rm = TRUE)

tAll_max <—as.data.frame(colMax(tAll_diff))
colnames(tAll_max) <—c("max")
tAll_last <—as.data.frame(t(slice(tAll_diff, timepoints)))
colnames(tAll_last) <—c("last")

tAll_diff <—bind_cols(tAll_max, tAll_last)
tAll_diff %>%
 mutate(diff = max-last) %>%
 mutate(ID = ID.H) %>%
 select(ID, everything())—> tAll_diff
remove(tAll_last)
remove(tAll_max)

#compute relative difference
tAll_diff %>%
 mutate(rel.diff = 1—(last/max))—> tAll_diff

#Extract all differences above the treshold
tAll_diff %>%
 filter(diff>treshold)—> tAll_diff_hits

#Separates the curve data into hits and non hits
hit.ID <—as.vector(tAll_diff_hits$ID)
tAll.long %>%
 filter(variable %in% hit.ID) %>%
 mutate(hit = "hit")—> tAll.hits

tAll.long %>%
 filter(!(variable %in% hit.ID)) %>%
 mutate(hit = "No.hit")—> tAll.NotHits
```

```
tAll.long <—union(tAll.hits, tAll.NotHits)
remove(tAll.hits)
remove(tAll.NotHits)

##### Output section #####

#Extract csv of hits
write.csv(tAll_diff_hits, file = Non0_Ouput, row.names =
FALSE)



##### Plotting Section #####

#Example of a plot with low number of variables
ggplot() +
 geom_line(data = tAll.t,
  aes(x = time, y = O22, color = "O22"),
  size = 2) +
 geom_line(data = tAll.t,
  aes(x =time, y = F22, color = "F22"),
  size = 2) +
 geom_line(data = tAll.t,
  aes(x =time, y = N18, color = "N18"),
  size = 2) +
 scale_color_manual(values = c(
 "O22" = "#7A003C",
 "F22" = "#FDBF57",
 "N18" = "#5E6A71")) +
 labs(x = "Time (H)", y = expression(OD[600]), color = "Well")

#Plot all variables together with hits highlited and none hits in
grey
ggplot(subset(tAll.long, hit %in% c("hit"))) +
 geom_line(aes(x = time,
  y = value,
  group = variable,
  colour = hit)) +
 scale_color_manual(values = c("#7A003C")) +
 expand_limits(x=c(0,4), y=c(0.15, 0.4))

ggplot(subset(tAll.long, hit %in% c("No.hit"))) +
 geom_line(aes(x = time,
```

```
 y = value,
 group = variable,
 colour = hit)) +
scale_color_manual(values = c("#5E6A71")) +
expand_limits(x=c(0,4), y=c(0.15, 0.4))


#Plot the Histogram of difference between max and last
(removes the 0s)
ggplot(data=tAll_diff_hits,
 aes(tAll_diff_hits$diff)) +
 geom_histogram(binwidth = 0.001,
  color = "#5E6A71",
  fill = "#7A003C") +
 theme(legend.position="none") +
 ylim(0,5) +
 labs(x = expression(OD[600]~Variation), y = "Count")
```

Primary hit arranged according to groups
(adapted from—https://www.r-graph-gallery.com/297-circular-barplot-with-groups/)

```
 #
 library
            library(tidyverse)

            list = read.csv("Final_HKhitlist(1000).csv", header = TRUE)
            sortedList <—list[order(list$ODdip),]
            #sortedList <—list[order('ODdip')]
            # Create dataset
            data=data.frame(
             individual=(sortedList$Hit_compound),
```

```
 group=(sortedList$group),
 value=(sortedList$ODdip)
)

# Set a number of 'empty bar' to add at the end of each group
empty_bar=3
to_add = data.frame( matrix(NA, empty_bar*nlevels(data$group),
ncol(data)) )
colnames(to_add) = colnames(data)
to_add$group=rep(levels(data$group), each=empty_bar)
data=rbind(data, to_add)
data=data %>% arrange(group)
data$id=seq(1, nrow(data))

# Get the name and the y position of each label
label_data=data
number_of_bar=nrow(label_data)
angle= 90 - 360 * (label_data$id-0.5) /number_of_bar # I
substract 0.5 because the letter must have the angle of the center
of the bars. Not extreme right(1) or extreme left (0)
label_data$hjust<-ifelse( angle < -90, 1, 0)
label_data$angle<-ifelse(angle < -90, angle+180, angle)

# prepare a data frame for base lines
base_data=data %>%
 group_by(group) %>%
 summarize(start=min(id), end=max(id)—empty_bar) %>%
 rowwise() %>%
 mutate(title=mean(c(start, end)))

# prepare a data frame for grid (scales)
grid_data = base_data
grid_data$end = grid_data$end[ c( nrow(grid_data),
1:nrow(grid_data)-1)] + 1
grid_data$start = grid_data$start - 1
grid_data=grid_data[-1,]

# Make the plot
p = ggplot(data, aes(x=as.factor(id), y=value, fill=group)) + #
Note that id is a factor. If x is numeric, there is some space
between the first bar

 geom_bar(aes(x=as.factor(id), y=value, fill=group),
stat="identity", alpha=0.5) +
```

```
 # Add a val=100/75/50/25 lines. I do it at the beginning to make
sur barplots are OVER it.
 geom_segment(data=grid_data, aes(x = end, y = 80, xend = start,
yend = 80), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +
 geom_segment(data=grid_data, aes(x = end, y = 60, xend = start,
yend = 60), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +
 geom_segment(data=grid_data, aes(x = end, y = 40, xend = start,
yend = 40), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +
 geom_segment(data=grid_data, aes(x = end, y = 20, xend = start,
yend = 20), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +

 # Add text showing the value of each 100/75/50/25 lines
 annotate("text", x = rep(max(data$id),4), y = c(20, 40, 60, 80),
label = c("0.02", "0.04", "0.06", "0.08") , color="grey", size=3 ,
angle=0, fontface="bold", hjust=1) +

 geom_bar(aes(x=as.factor(id), y=value, fill=group),
stat="identity", alpha=0.5) +
 ylim(-100,120) +
 theme_minimal() +
 scale_fill_brewer(palette="Set1") +
 theme(
 legend.position = "none",
 axis.text = element_blank(),
 axis.title = element_blank(),
 panel.grid = element_blank(),
 plot.margin = unit(rep(-1,4), "cm")
 ) +
 coord_polar() +
 geom_text(data=label_data, aes(x=id, y=value+10,
label=individual, hjust=hjust), color="black",
fontface="bold",alpha=0.6, size=2.5, angle= label_data$angle,
inherit.aes = FALSE ) +

 # Add base line information
 geom_segment(data=base_data, aes(x = start, y = -5, xend = end,
yend = -5), colour = "black", alpha=0.8, size=0.6 , inherit.aes =
FALSE ) +
```

```
 geom_text(data=base_data, aes(x = title, y = -18, label=group),
hjust=c(1,1,0.9,0.9,0), colour = "black", alpha=0.8, size=3,
fontface="bold", inherit.aes = FALSE)

p

##https://www.r-graph-gallery.com/297-circular-barplot-with-
groups/

list = read.csv("Final_Muhitlist(1000)V.csv", header = TRUE)
sortedList <—list[order(list$ODdip),]
#sortedList <—list[order('ODdip')]
# Create dataset
data=data.frame(
 individual=(sortedList$Hit_compound),
 group=(sortedList$group),
 value=(sortedList$ODdip)
)

# Set a number of 'empty bar' to add at the end of each group
empty_bar=3
to_add = data.frame( matrix(NA, empty_bar*nlevels(data$group),
ncol(data)) )
colnames(to_add) = colnames(data)
to_add$group=rep(levels(data$group), each=empty_bar)
data=rbind(data, to_add)
data=data %>% arrange(group)
data$id=seq(1, nrow(data))

# Get the name and the y position of each label
label_data=data
number_of_bar=nrow(label_data)
angle= 90 - 360 * (label_data$id-0.5) /number_of_bar # I
substract 0.5 because the letter must have the angle of the center
of the bars. Not extreme right(1) or extreme left (0)
label_data$hjust<-ifelse( angle < -90, 1, 0)
label_data$angle<-ifelse(angle < -90, angle+180, angle)

# prepare a data frame for base lines
base_data=data %>%
 group_by(group) %>%
 summarize(start=min(id), end=max(id)—empty_bar) %>%
 rowwise() %>%
 mutate(title=mean(c(start, end)))
```

```
# prepare a data frame for grid (scales)
grid_data = base_data
grid_data$end = grid_data$end[ c( nrow(grid_data),
1:nrow(grid_data)-1)] + 1
grid_data$start = grid_data$start - 1
grid_data=grid_data[-1,]

# Make the plot
q = ggplot(data, aes(x=as.factor(id), y=value, fill=group)) + #
Note that id is a factor. If x is numeric, there is some space
between the first bar
geom_bar(aes(x=as.factor(id), y=value, fill=group),
stat="identity", alpha=0.5) +

 # Add a val=100/75/50/25 lines. I do it at the beginning to make
sur barplots are OVER it.
 geom_segment(data=grid_data, aes(x = end, y = 80, xend = start,
yend = 80), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +
 geom_segment(data=grid_data, aes(x = end, y = 60, xend = start,
yend = 60), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +
 geom_segment(data=grid_data, aes(x = end, y = 40, xend = start,
yend = 40), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +
 geom_segment(data=grid_data, aes(x = end, y = 20, xend = start,
yend = 20), colour = "grey", alpha=1, size=0.3 , inherit.aes =
FALSE ) +

 # Add text showing the value of each 100/75/50/25 lines
 annotate("text", x = rep(max(data$id),4), y = c(20, 40, 60, 80),
label = c("0.02", "0.04", "0.06", "0.08") , color="grey", size=3 ,
angle=0, fontface="bold", hjust=1) +

 geom_bar(aes(x=as.factor(id), y=value, fill=group),
stat="identity", alpha=1.0) +
 ylim(-100,140) +
 theme_minimal() +
 scale_fill_manual(values=c("#b0cc6d", "#61af8e", "#83d7fb")) +
 theme(
 legend.position = "none",
 axis.text = element_blank(),
 axis.title = element_blank(),
```

```
 panel.grid = element_blank(),
 plot.margin = unit(rep(-1,4), "cm")
 ) +
 coord_polar() +
 geom_text(data=label_data, aes(x=id, y=value+10,
label=individual, hjust=hjust), color="black",
fontface="bold",alpha=0.6, size=2.5, angle= label_data$angle,
inherit.aes = FALSE ) +

 # Add base line information
 geom_segment(data=base_data, aes(x = start, y = -5, xend = end,
yend = -5), colour = "black", alpha=0.8, size=0.6 , inherit.aes =
FALSE ) +
 geom_text(data=base_data, aes(x = title, y = -20, label=group),
hjust=c(1.0,0.5,0.1), colour = "black", alpha=0.8, size=3,
fontface="bold", inherit.aes = FALSE)

 q
```

Secondary hits graphing

```
install.packages("vir
dis")
                    install.packages("ggplot2")
                    install.packages("hrbrthemes")
                    # library
                     library(ggplot2)
                     library(viridis)
                     library(hrbrthemes)

                     # create a dataset
                     list = read.csv("actives_to_hits(nonSOS).csv",
                    header = TRUE)
                     list2 = read.csv("Dopaminehits.csv", header =
                    TRUE)
                     list3 = read.csv("Serotoninhits.csv", header =
                    TRUE)
                     # # Small multiple
                     # ggplot(data, aes(fill=condition, y=value,
                    x=specie)) +
                     # geom_bar(position="stack", stat="identity") +
                     # scale_fill_viridis(discrete = T) +
                     # ggtitle("Studying 4 species..") +
                     # xlab("")
```

```
#
# # Small multiple
# ggplot(list, aes(fill=name, y=Total,
x=Compound.name)) +
# geom_bar(position="stack", stat="identity") +
# scale_fill_viridis(discrete = FALSE) +
# xlab("")
#
# ggplot() +
# geom_bar(data=list, aes(y = Total, x =
Compound.name), width = 0.3, stat="identity",
#  position='stack') +
# coord_flip() +
# facet_grid( name ~ .)
# scale_fill_viridis()
#
# ggplot() +
# geom_bar(data=list, aes(y = Total, x =
Compound.name), width = 0.3, stat="identity",
#  position='stack') +
# facet_grid( name ~ .)
# scale_fill_viridis()
#
# list2 = list[list$name == 'one',]
# theme_set(theme_gray(base_size = 7))
# ggplot() +
# geom_bar(data=list2, aes(y = Total, x =
Compound.name), width = 0.3, stat="identity",
#  position='stack') +
# coord_flip() +
# scale_fill_viridis()
# Dummy data
# x <—LETTERS[1:20]
# y <—paste0("var", seq(1,20))
# data <—expand.grid(X=x, Y=y)
# data$Z <- runif(400, 0, 5)
ind <—seq(190)
ind2 <- seq(416,860)
list1 <—list[ind, ]
list2 <—list[ind2,]
# Heatmap
ggplot(list, aes(compound.concentration,
Compound.name, fill= phage.dilution)) +
geom_tile(
```

```
 colour = "white",position="identity") +
 theme(text = element_text(size=5))+
 scale_fill_gradient2(low="#ef8a62",
high="#67a9cf", limits=c(4,5)) +

scale_x_discrete(labels=c("1.5","3.25","6.25","12.
5","25")) +
 coord_fixed(ratio = 0.3)


 #edited heatmap for hits
 p <—ggplot(list, aes(compound.concentration,
Compound.name, fill= phage.dilution)) +
 geom_tile(
 colour = "white",position="identity") +
 theme(text = element_text(size=5))+
 #scale_fill_manual(drop=FALSE,
values=colorRampPalette(c("white","red"))(5),
na.value="#EEEEEE", name="phage.dilution") +
 scale_fill_gradient( low="orange",
high="orange3" , limits = c(3,5)) +

scale_x_discrete(labels=c("1.5","3.25","6.25","12.
5","25")) +
 scale_y_discrete(labels=NULL) +
 #coord_fixed(ratio = 0.3) +

 theme_test() +
 theme(axis.ticks.x = element_blank(), axis.title.x
= element_blank())

 p + coord_flip() #flips co-ordinates
 # theme(axis.text.x = element_text(angle = 0,
hjust = 1)) #changes the angle of the axis label


q <—ggplot(list2, aes(compound.concentration,
Compound.name, fill= phage.dilution)) +
 geom_tile(
 colour = "white",position="identity") +
 theme(text = element_text(size=5))+
 #scale_fill_manual(drop=FALSE,
values=colorRampPalette(c("white","red"))(5),
na.value="#EEEEEE", name="phage.dilution") +
```

```
 scale_fill_gradient( low="skyblue",
high="deepskyblue3" , limits = c(3,5)) +

scale_x_discrete(labels=c("1.5","3.25","6.25","12.
5","25")) +
 scale_y_discrete(labels=NULL) +
 #coord_fixed(ratio = 0.3) +

 theme_test() +
 theme(axis.ticks.x = element_blank(), axis.title.x
= element_blank())

 q + coord_flip() #flips co-ordinates
 # theme(axis.text.x = element_text(angle = 0,
hjust = 1)) #changes the angle of the axis label


 r <—ggplot(list3, aes(compound.concentration,
Compound.name, fill= phage.dilution)) +
 geom_tile(
 colour = "white",position="identity") +
 theme(text = element_text(size=5))+
 #scale_fill_manual(drop=FALSE,
values=colorRampPalette(c("white","red"))(5),
na.value="#EEEEEE", name="phage.dilution") +
 scale_fill_gradient( low="mediumpurple3",
high="mediumpurple4" , limits = c(3,5)) +

scale_x_discrete(labels=c("1.5","3.25","6.25","12.
5","25")) +
 scale_y_discrete(labels=NULL) +
 #coord_fixed(ratio = 0.3) +

 theme_test() +
 theme(axis.ticks.x = element_blank(), axis.title.x
= element_blank())

 r + coord_flip() #flips co-ordinates
 # theme(axis.text.x = element_text(angle = 0,
hjust = 1)) #changes the angle of the axis label

 #theme_ipsum()
```

```
 base_size <- 9
 p + theme_grey(base_size = base_size) +
 labs(x = "",
 y = "") + scale_x_discrete(expand = c(0, 0)) +
 scale_y_discrete(expand = c(0, 0)) +
opts(legend.position = "none",
  axis.ticks = theme_blank(), axis.text.x =
theme_text(size = base_size *
  0.8, angle = 330, hjust = 0, colour = "grey50"))

 # Heatmap
 ggplot(list2, aes(compound.concentration,
Compound.name, fill= phage.dilution)) +
 geom_tile(
 colour = "white",position="identity") +
 theme(text = element_text(size=5))+
 scale_fill_gradient2(low="#ef8a62",
mid="#f7f7f7", high="#67a9cf", midpoint=2,
limits=c(1,5)) +

scale_x_discrete(labels=c("1.5","3.25","6.25","12.
5","25")) +
 coord_fixed(ratio = 0.3)
```

APPENDIX II: PILOT SCREEN DATA

Table A1: Subset of compounds used for pilot study. The compounds highlighted in grey were hits by the OD drop method and the ones highlighted in red are IQM hits. The values for compound showing OD drop less than 0.01 are not shown.

Compound Name	OD max R1	OD final R1	OD max R2	OD final R2	Average OD max-OD end point
TOLAZOLINE HYDROCHLORIDE	0.344	0.344	0.346	0.346	
SULFACHLORPYRIDAZINE	0.344	0.344	0.35	0.35	
TRIAMCINOLONE DIACETATE	0.339	0.339	0.344	0.344	
RANITIDINE	0.342	0.341	0.345	0.345	
SPARTEINE SULFATE	0.343	0.339	0.335	0.335	
FLURANDRENOLIDE	0.349	0.349	0.34	0.34	
VIDARABINE	0.356	0.354	0.337	0.337	
MERBROMIN	0.348	0.348	0.337	0.337	
PROCHLORPERAZINE EDISYLATE	0.346	0.343	0.334	0.334	
AMINACRINE	0.329	0.329	0.321	0.321	
PSEUDOEPHEDRINE HYDROCHLORIDE	0.341	0.341	0.328	0.328	
DROPERIDOL	0.349	0.349	0.337	0.337	
RESORCINOL	0.351	0.349	0.333	0.328	
FENBENDAZOLE	0.357	0.356	0.341	0.341	
SULFABENZAMIDE	0.351	0.349	0.337	0.331	
MEBEVERINE HYDROCHLORIDE	0.355	0.354	0.343	0.337	
SULFASALAZINE	0.346	0.346	0.347	0.344	
NICERGOLINE	0.353	0.353	0.349	0.346	
TRIPELENNAMINE CITRATE	0.342	0.342	0.338	0.336	
SULFADIMETHOXINE	0.35	0.35	0.351	0.348	
EDOXUDINE	0.347	0.347	0.343	0.343	
LITHIUM CITRATE	0.355	0.355	0.355	0.355	
LISINOPRIL	0.362	0.362	0.353	0.353	
DONEPEZIL HYDROCHLORIDE	0.374	0.374	0.358	0.358	
PHTHALYLSULFATHIAZOLE	0.374	0.374	0.359	0.359	
CLORSULON	0.369	0.369	0.347	0.347	
CEFUROXIME SODIUM	0.205	0.186	0.199	0.166	0.026

BENAZEPRIL HYDROCHLORIDE	0.366	0.366	0.349	0.349	
BEZAFIBRATE	0.373	0.373	0.351	0.351	
SIBUTRAMINE HYDROCHLORIDE	0.366	0.366	0.347	0.347	
TINIDAZOLE	0.369	0.369	0.351	0.351	
TELMISARTAN	0.367	0.367	0.345	0.345	
BENZYL BENZOATE	0.368	0.368	0.348	0.348	
MEPIVACAINE HYDROCHLORIDE	0.371	0.371	0.347	0.345	
QUINAPRIL HYDROCHLORIDE	0.369	0.369	0.354	0.351	
TRIFLURIDINE	0.37	0.37	0.342	0.34	
BROMHEXINE HYDROCHLORIDE	0.363	0.363	0.361	0.357	
ACRISORCIN	0.355	0.355	0.348	0.348	
TENOXICAM	0.367	0.361	0.359	0.355	
GATIFLOXACIN	0.245	0.207	0.238	0.193	0.0415
WARFARIN	0.36	0.36	0.35	0.35	
SPIPERONE	0.356	0.356	0.356	0.356	
PROCYCLIDINE HYDROCHLORIDE	0.357	0.357	0.353	0.353	
ERYTHROMYCIN ESTOLATE	0.363	0.363	0.357	0.357	
PYRANTEL PAMOATE	0.372	0.372	0.352	0.352	
PHENTOLAMINE HYDROCHLORIDE	0.383	0.383	0.349	0.349	
RIFAMPIN	0.366	0.366	0.345	0.345	
BEKANAMYCIN SULFATE	0.357	0.357	0.343	0.343	
SULFACETAMIDE	0.363	0.363	0.347	0.347	
FAMOTIDINE	0.366	0.366	0.347	0.347	
SULFATHIAZOLE	0.359	0.359	0.342	0.342	
MECLOCYCLINE SULFOSALICYLATE	0.188	0.184	0.179	0.174	
THIABENDAZOLE	0.362	0.362	0.342	0.342	
SULFAQUINOXALINE SODIUM	0.359	0.359	0.347	0.347	
TOLMETIN SODIUM	0.365	0.365	0.349	0.347	
SULOCTIDIL	0.366	0.366	0.346	0.344	
TRICHLORMETHIAZIDE	0.366	0.362	0.37	0.366	
ESTRADIOL PROPIONATE	0.361	0.359	0.365	0.364	
PROGESTERONE	0.359	0.353	0.358	0.354	
BETAHISTINE HYDROCHLORIDE	0.36	0.355	0.364	0.358	
ENOXACIN	0.292	0.187	0.283	0.204	0.092

HOMOSALATE	0.361	0.361	0.358	0.358	
ETHISTERONE	0.368	0.368	0.354	0.354	
ESTROPIPATE	0.371	0.371	0.358	0.358	
ISOXICAM	0.399	0.399	0.376	0.376	
FAMCICLOVIR	0.375	0.375	0.355	0.355	
AMSACRINE	0.377	0.377	0.354	0.354	
PERINDOPRIL ERBUMINE	0.376	0.376	0.351	0.351	
HYDROCORTISONE BUTYRATE	0.379	0.379	0.353	0.353	
AMINOLEVULINIC ACID HYDROCHLORIDE	0.375	0.375	0.349	0.349	
THIOTEPA	0.373	0.373	0.349	0.349	
NATEGLINIDE	0.374	0.374	0.355	0.355	
RESORCINOL MONOACETATE	0.372	0.372	0.354	0.354	
MIGLITOL	0.385	0.385	0.352	0.352	
CELECOXIB	0.377	0.377	0.359	0.359	
OLANZAPINE	0.384	0.384	0.362	0.362	
NAPROXOL	0.366	0.364	0.365	0.365	
CLAVULANATE LITHIUM	0.368	0.368	0.379	0.379	
IOPANIC ACID	0.364	0.364	0.38	0.38	
AMLODIPINE BESYLATE	0.363	0.358	0.364	0.358	
ROXARSONE	0.372	0.372	0.367	0.367	
BUDESONIDE	0.361	0.361	0.359	0.359	
SULFADIAZINE	0.35	0.35	0.35	0.35	
MEFENAMIC ACID	0.361	0.361	0.351	0.351	
SULFINPYRAZONE	0.372	0.372	0.348	0.348	
SULFAGUANIDINE	0.371	0.371	0.351	0.351	
THIMEROSAL	0.25	0.25	0.23	0.23	
RONIDAZOLE	0.38	0.38	0.355	0.355	
TOLNAFTATE	0.36	0.36	0.339	0.339	
ESTRADIOL BENZOATE	0.378	0.378	0.36	0.36	
ZOMEPIRAC SODIUM	0.355	0.355	0.339	0.339	
BUTAMBEN	0.367	0.367	0.35	0.35	
SALICYL ALCOHOL	0.361	0.361	0.342	0.341	
"CANRENOIC ACID, POTASSIUM SALT"	0.368	0.368	0.351	0.351	
SULFAMERAZINE	0.359	0.359	0.342	0.342	
METHACYCLINE HYDROCHLORIDE	0.213	0.213	0.212	0.208	
SULFISOXAZOLE	0.357	0.357	0.355	0.351	
MINAPRINE HYDROCHLORIDE	0.361	0.361	0.36	0.357	

THIOGUANINE	0.352	0.351	0.352	0.35	
NEFOPAM	0.362	0.362	0.361	0.356	
PARAROSANILINE PAMOATE	0.38	0.38	0.374	0.374	
ROSUVASTATIN CALCIUM	0.368	0.368	0.362	0.362	
MIDODRINE HYDROCHLORIDE	0.362	0.362	0.359	0.359	
TELITHROMYCIN	0.354	0.354	0.344	0.344	
ROXITHROMYCIN	0.382	0.382	0.36	0.36	
VENLAFAXINE	0.374	0.374	0.35	0.35	
TETROQUINONE	0.379	0.379	0.361	0.361	
METAXALONE	0.373	0.373	0.347	0.347	
AZITHROMYCIN	0.234	0.232	0.217	0.217	
IRBESARTAN	0.371	0.371	0.348	0.348	
ISOSORBIDE MONONITRATE	0.375	0.375	0.353	0.353	
VALDECOXIB	0.37	0.37	0.349	0.349	
HYDROXYCHLOROQUINE SULFATE	0.375	0.375	0.358	0.357	
ALCLOMETAZONE DIPROPIONATE	0.369	0.369	0.349	0.349	
PERHEXILINE MALEATE	0.372	0.372	0.359	0.359	
EZETIMIBE	0.378	0.378	0.35	0.348	
NADOLOL	0.37	0.37	0.372	0.371	
RAMIPRIL	0.359	0.359	0.355	0.353	
PIPOBROMAN	0.362	0.362	0.358	0.355	
CITALOPRAM	0.357	0.357	0.356	0.354	
TRANYLCYPROMINE SULFATE	0.355	0.355	0.352	0.352	
SULFAMONOMETHOXINE	0.365	0.365	0.355	0.355	
TUAMINOHEPTANE SULFATE	0.357	0.357	0.352	0.352	
SULFAMETER	0.369	0.369	0.354	0.354	
ACETARSOL	0.369	0.369	0.346	0.346	
ESTRADIOL ACETATE	0.373	0.373	0.357	0.357	
PYRIMETHAMINE	0.318	0.317	0.306	0.304	
ECONAZOLE NITRATE	0.354	0.353	0.342	0.342	
SALICYLAMIDE	0.365	0.365	0.345	0.345	
SUCCINYLSULFATHIAZOLE	0.377	0.377	0.355	0.355	
SULFAMETHAZINE	0.368	0.368	0.349	0.349	
ALRESTATIN	0.366	0.366	0.351	0.351	
TRIACETIN	0.368	0.368	0.348	0.348	
CARBIDOPA	0.393	0.393	0.369	0.369	
TRIMEPRAZINE TARTRATE	0.368	0.368	0.347	0.347	
BENZOYL PEROXIDE	0.388	0.388	0.352	0.352	

TYROTHRICIN	0.345	0.331	0.331	0.328	
ABAMECTIN	0.382	0.382	0.377	0.377	
ACRIFLAVINIUM HYDROCHLORIDE	0.377	0.377	0.364	0.361	
CEFTRIAXONE SODIUM TRIHYDRATE	0.352	0.338	0.332	0.326	
AMCINONIDE	0.372	0.372	0.368	0.368	
AVOBENZONE	0.363	0.363	0.362	0.362	
MEPHENTERMINE SULFATE	0.372	0.372	0.363	0.363	
CLARITHROMYCIN	0.36	0.36	0.346	0.346	
NALTREXONE HYDROCHLORIDE	0.384	0.384	0.356	0.356	
LEVOFLOXACIN	0.185	0.144	0.169	0.13	0.04
OXETHAZAINE	0.382	0.382	0.365	0.365	
MOXIFLOXACIN HYDROCHLORIDE	0.245	0.195	0.229	0.188	0.0445
ETANIDAZOLE	0.381	0.381	0.362	0.362	
THIOSTREPTON	0.372	0.372	0.348	0.348	
BENZALKONIUM CHLORIDE	0.37	0.366	0.356	0.356	
ALENDRONATE SODIUM	0.383	0.383	0.364	0.364	
CLOPIDOGREL SULFATE	0.376	0.376	0.363	0.363	
OLMESARTAN MEDOXOMIL	0.374	0.374	0.349	0.349	
BUPIVACAINE HYDROCHLORIDE	0.378	0.378	0.355	0.355	
DERACOXIB	0.375	0.375	0.348	0.348	
BETA-PROPIOLACTONE	0.378	0.378	0.362	0.362	
ATOVAQUONE	0.365	0.365	0.355	0.355	
METARAMINOL BITARTRATE	0.371	0.371	0.362	0.362	
ROFECOXIB	0.37	0.37	0.357	0.356	
PYRVINIUM PAMOATE	0.364	0.364	0.361	0.361	
CETRIMONIUM BROMIDE	0.33	0.33	0.336	0.336	
SODIUM SALICYLATE	0.361	0.361	0.361	0.361	
FLUNISOLIDE	0.372	0.372	0.366	0.366	
SULFAMETHIZOLE	0.369	0.369	0.349	0.349	
CARBENOXOLONE SODIUM	0.374	0.374	0.356	0.356	
THIOTHIXENE	0.362	0.362	0.344	0.344	
TRANEXAMIC ACID	0.367	0.367	0.356	0.356	
TRIAMCINOLONE	0.369	0.369	0.352	0.352	
CEFAMANDOLE SODIUM	0.352	0.352	0.343	0.343	
TRIMETHOBENZAMIDE HYDROCHLORIDE	0.365	0.365	0.353	0.353	

AKLOMIDE	0.375	0.375	0.361	0.361	
UREA	0.367	0.367	0.355	0.355	
AMIPRILOSE	0.371	0.371	0.356	0.356	
AMINOPYRINE	0.372	0.372	0.351	0.351	
CHLOROXINE	0.225	0.225	0.207	0.206	
SPIRONOLACTONE	0.369	0.369	0.361	0.361	
XYLAZINE	0.372	0.371	0.36	0.36	
SULFAMETHOXAZOLE	0.365	0.365	0.353	0.352	
CARPROFEN	0.382	0.382	0.367	0.367	
DIPYRONE	0.37	0.37	0.371	0.371	
CANDESARTAN CILEXTIL	0.371	0.371	0.373	0.373	
NAFRONYL OXALATE	0.371	0.371	0.373	0.373	
PIOGLITAZONE HYDROCHLORIDE	0.369	0.369	0.359	0.359	
ALBENDAZOLE	0.374	0.374	0.358	0.358	
TILMICOSIN	0.38	0.38	0.358	0.358	
MORANTEL CITRATE	0.381	0.381	0.354	0.354	
ROPINIROLE	0.38	0.38	0.353	0.353	
ETHYLNOREPINEPHRINE HYDROCHLORIDE	0.383	0.383	0.363	0.363	
CEFTIBUTEN	0.364	0.364	0.354	0.354	
HALCINONIDE	0.377	0.377	0.365	0.365	
CROTAMITON	0.377	0.377	0.359	0.359	
METHAZOLAMIDE	0.379	0.379	0.364	0.364	
CHLOROGUANIDE HYDROCHLORIDE	0.375	0.375	0.354	0.354	
SULFANILATE ZINC	0.381	0.381	0.356	0.356	
SIMVASTATIN	0.386	0.386	0.365	0.365	
QUIPAZINE MALEATE	0.379	0.379	0.365	0.365	
ROSIGLITAZONE	0.367	0.367	0.354	0.354	
CIPROFLOXACIN	0.284	0.22	0.271	0.21	0.0625
TOLTERODINE TARTRATE	0.38	0.38	0.372	0.372	
TERBUTALINE HEMISULFATE	0.363	0.363	0.353	0.353	
CEFMETAZOLE SODIUM	0.364	0.364	0.363	0.363	
TIMOLOL MALEATE	0.354	0.354	0.352	0.352	
CARBOPLATIN	0.37	0.37	0.362	0.362	
TRIAMCINOLONE ACETONIDE	0.363	0.363	0.347	0.347	
NICOTINYL ALCOHOL TARTRATE	0.371	0.371	0.351	0.351	
PHENETHICILLIN POTASSIUM	0.358	0.358	0.337	0.337	

CHLORPROTHIXENE HYDROCHLORIDE	0.364	0.364	0.353	0.353	
PHENACETIN	0.368	0.368	0.343	0.343	
CEFSULODIN SODIUM	0.373	0.373	0.356	0.356	
IODIPAMIDE	0.361	0.361	0.348	0.348	
MESNA	0.376	0.376	0.354	0.354	
ETOPOSIDE	0.367	0.367	0.351	0.351	
CISPLATIN	0.378	0.378	0.359	0.359	
FENBUFEN	0.367	0.367	0.345	0.345	
PIPERIDOLATE HYDROCHLORIDE	0.394	0.394	0.366	0.366	
MEFEXAMIDE	0.375	0.375	0.35	0.35	
GLUCONOLACTONE	0.377	0.377	0.366	0.366	
SULCONAZOLE NITRATE	0.355	0.355	0.345	0.345	
CINNARAZINE	0.375	0.375	0.359	0.359	
SELAMECTIN	0.382	0.382	0.377	0.377	
FLUNIXIN MEGLUMINE	0.363	0.363	0.365	0.365	
BUTACAINE	0.37	0.37	0.371	0.371	
QUETIAPINE	0.374	0.374	0.355	0.355	
OXYPHENCYCLIMINE HYDROCHLORIDE	0.382	0.382	0.36	0.36	
CEFDINIR	0.16	0.068	0.157	0.064	0.092
TENIPOSIDE	0.382	0.382	0.362	0.362	
CITICOLINE	0.377	0.377	0.354	0.354	
HYCANTHONE	0.398	0.398	0.374	0.374	
ZOLMITRIPTAN	0.371	0.371	0.353	0.353	
METHYLBENZETHONIUM CHLORIDE	0.353	0.329	0.345	0.331	0.019
OXFENDAZOLE	0.372	0.372	0.35	0.35	
URETHANE	0.377	0.377	0.358	0.358	
RIFAXIMIN	0.379	0.374	0.355	0.355	
PRACTOLOL	0.378	0.378	0.357	0.357	
MODAFINIL	0.378	0.378	0.352	0.352	
ENROFLOXACIN	0.2	0.156	0.19	0.148	0.043
METHYLDOPATE HYDROCHLORIDE	0.374	0.374	0.356	0.356	
CLOBETASOL PROPIONATE	0.371	0.371	0.358	0.358	
PHENYLETHYL ALCOHOL	0.366	0.366	0.357	0.357	
SULFAMETHOXYPYRIDAZINE	0.355	0.355	0.355	0.355	
GALANTHAMINE HYDROBROMIDE	0.366	0.366	0.364	0.364	

PHENYLMERCURIC ACETATE	0.19	0.177	0.177	0.174	
CEFAMANDOLE NAFATE	0.367	0.367	0.366	0.366	
LEVOTHYROXINE	0.379	0.379	0.352	0.352	
ZIDOVUDINE [AZT]	0.315	0.311	0.302	0.297	
FLUMEQUINE	0.237	0.196	0.224	0.188	0.0385
ANISINDIONE	0.381	0.381	0.357	0.357	
PROBUCOL	0.373	0.373	0.353	0.353	
AMINOHIPPURIC ACID	0.376	0.376	0.353	0.353	
PRAMOXINE HYDROCHLORIDE	0.358	0.358	0.345	0.345	
AZLOCILLIN SODIUM	0.24	0.223	0.224	0.205	0.018
RITODRINE HYDROCHLORIDE	0.365	0.365	0.348	0.348	
ENALAPRIL MALEATE	0.374	0.374	0.357	0.357	
SACCHARIN	0.377	0.377	0.353	0.353	
FOSCARNET SODIUM	0.385	0.385	0.356	0.356	
ACONITINE	0.373	0.373	0.356	0.356	
CEFOXITIN SODIUM	0.218	0.2	0.212	0.19	0.02
CIMETIDINE	0.368	0.368	0.357	0.357	
CEFOPERAZONE SODIUM	0.194	0.152	0.188	0.149	0.0405
SILDENAFIL	0.377	0.373	0.37	0.37	
IFOSFAMIDE	0.368	0.368	0.36	0.36	
METHYLPREDNISOLONE SODIUM SUCCINATE	0.374	0.374	0.371	0.371	
NATAMYCIN	0.367	0.367	0.356	0.356	
THIRAM	0.292	0.292	0.278	0.278	
AMITRAZ	0.376	0.376	0.354	0.354	
NABUMETONE	0.383	0.383	0.361	0.361	
CHLORMADINONE ACETATE	0.373	0.373	0.348	0.348	
ATORVASTATIN CALCIUM	0.387	0.387	0.371	0.371	
CEFPROZIL	0.372	0.372	0.354	0.354	
CLOPAMIDE	0.376	0.376	0.357	0.357	
ATOMOXETINE HYDROCHLORIDE	0.371	0.371	0.356	0.356	
ALAPROCLATE	0.375	0.375	0.353	0.353	
ETHOXZOLAMIDE	0.371	0.371	0.352	0.352	
ACETRIAZOIC ACID	0.383	0.383	0.364	0.364	
PREDNISOLONE HEMISUCCINATE	0.382	0.382	0.355	0.355	
CEFUROXIME AXETIL	0.363	0.363	0.352	0.352	
DOXORUBICIN	0.372	0.372	0.362	0.362	
TRIMETOZINE	0.376	0.376	0.36	0.36	

NORGESTIMATE	0.36	0.36	0.357	0.357	
LIOTHYRONINE	0.365	0.365	0.351	0.351	
TROLEANDOMYCIN	0.367	0.367	0.37	0.37	
FENOPROFEN	0.359	0.359	0.361	0.361	
MEFLOQUINE	0.368	0.368	0.365	0.365	
MEBENDAZOLE	0.37	0.37	0.35	0.35	
BACAMPICILLIN HYDROCHLORIDE	0.362	0.362	0.356	0.356	
MEPHENESIN	0.36	0.36	0.356	0.356	
THONZYLAMINE HYDROCHLORIDE	0.368	0.368	0.367	0.367	
SULPIRIDE	0.366	0.366	0.356	0.356	
"BETAMETHASONE 17,21-DIPROPIONATE"	0.372	0.372	0.355	0.355	
ACETANILIDE	0.354	0.354	0.343	0.343	
FOLIC ACID	0.369	0.369	0.355	0.355	
ACEBUTOLOL HYDROCHLORIDE	0.366	0.366	0.346	0.346	
BENZOIC ACID	0.373	0.373	0.357	0.357	
ALTHIAZIDE	0.369	0.369	0.347	0.347	
ADIPHENINE HYDROCHLORIDE	0.375	0.375	0.357	0.357	
FLUFENAMIC ACID	0.36	0.36	0.355	0.355	
BENDROFLUMETHIAZIDE	0.378	0.378	0.372	0.372	
NALBUPHINE HYDROCHLORIDE	0.37	0.37	0.348	0.348	
THIAMPHENICOL	0.318	0.311	0.307	0.297	
LOSARTAN	0.352	0.352	0.361	0.361	
PEFLOXACINE MESYLATE	0.186	0.156	0.194	0.165	0.0295
PIZOTYLINE MALATE	0.347	0.347	0.345	0.345	
AZAPERONE	0.339	0.339	0.338	0.338	
DPCLIOQUINOL	0.247	0.247	0.234	0.234	
RANOLAZINE	0.353	0.353	0.345	0.345	
RIZATRIPTAN BENZOATE	0.363	0.363	0.354	0.354	
DULOXETINE HYDROCHLORIDE	0.356	0.356	0.351	0.351	
VALSARTAN	0.36	0.36	0.351	0.351	
BENURESTAT	0.363	0.363	0.355	0.355	
CANRENONE	0.356	0.356	0.348	0.348	
SULFISOXAZOLE ACETYL	0.348	0.348	0.342	0.342	
DIRITHROMYCIN	0.362	0.362	0.348	0.348	
ACEPROMAZINE MALEATE	0.35	0.35	0.344	0.344	

FEXOFENADINE HYDROCHLORIDE	0.308	0.308	0.298	0.298	
CEFPODOXIME PROXETIL	0.353	0.353	0.336	0.336	
ZOXAZOLAMINE	0.345	0.345	0.341	0.341	
ASPARTAME	0.342	0.342	0.345	0.345	
OXCARBAZEPINE	0.351	0.351	0.348	0.348	
AZELASTINE HYDROCHLORIDE	0.347	0.347	0.343	0.343	