

Investigation of a *Phaeobacter* gene transfer agent

Bachelor project, 20 ECTS. Carried out in spring 2023 at DTU Bioengineering

Authors: Robert Friis-Møller & Frederik Reimert

Supervisors: Laura Louise Lindqvist, Morten Dencker Schostag, Shengda Zhang, & Lone Gram

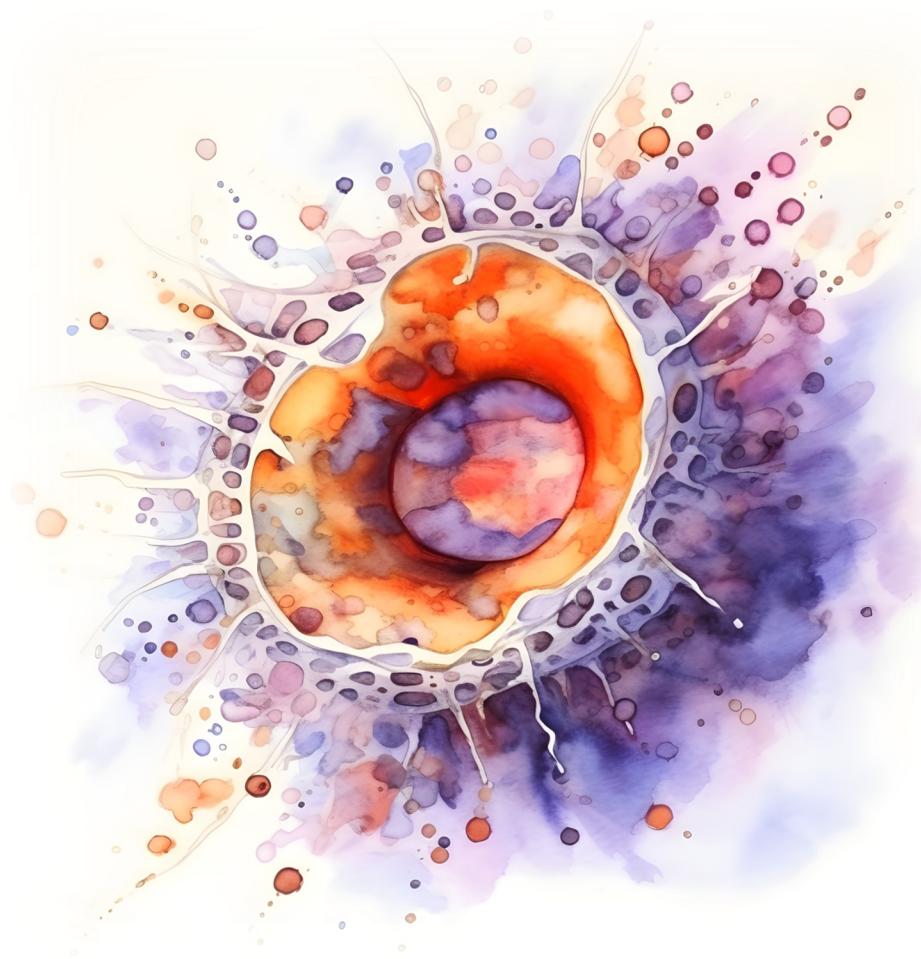


Image produced using AI image generation tool Midjourney

Table of contents

List of abbreviations.....	3
Abstract	4
Study importance.....	4
Background	5
Introduction	5
Gene transfer agents	5
Properties of RcGTAs.....	6
Mechanism of RcGTAs.....	8
Packaging and lysis	8
Delivery and uptake	8
Detection and quantification of RcGTAs	9
Regulation of GTA production and release.....	10
The GafA regulator	10
Growth phase dynamics and quorum sensing.....	11
Effects of tropodithietic acid (TDA) on GTA production	11
Evolution and ecological role of GTAs.....	12
Scope and aim of experiments	13
Materials and methods	14
Bacterial strains used in this study	14
Media and culture conditions	14
RNA extraction/purification and cDNA library construction.....	15
Analysis of GTA gene expression by RT-qPCR	15
Detection of a prophage from <i>P. piscinae</i> S26 monocultures by flow cytometry.....	17
Data analysis	18
Flow cytometry.....	18
RT-qPCR	19
Software:.....	19
Results.....	20
Detection of a potential GTA dsDNA band by electrophoresis of RNA extraction.....	20
Detection of a prophage from <i>P. piscinae</i> S26 monocultures by flow cytometry.....	21
Investigation of <i>Phaeobacter</i> sp. GTA capsid gene expression by RT-qPCR.....	24
Summary of results.....	26
Discussion	27
Acknowledgements	32
Bibliography.....	33
Appendix.....	40

List of abbreviations

Abbreviation	Definition
ANOVA	Analysis of variance
cDNA	Copy-DNA
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded DNA
EM	Electron microscopy
FCM	Flow cytometry
gDNA	Genomic DNA
GER	Gene expression ratio
GOI	Gene of interest
GTA	Gene transfer agent
HGT	Horizontal gene transfer
HKG	Housekeeping gene
MLM	Multivariate linear model
MoA	Mechanism of action
QS	Quorum sensing
RcGTA	<i>Rhodobacter capsulatus</i> gene transfer agent
RNA	Ribonucleic acid
RT-qPCR	Reverse transcription quantitative polymerase chain reaction
TDA	Tropodithietic acid
TPMC	Two-step preparative monolithic chromatography
WT	Wildtype

Abstract

A single milliliter of coastal seawater contains from 10^6 to 10^{11} viral particles, some of which are believed to be gene transfer agents (GTAs). GTAs are tailed, dsDNA containing capsid particles, akin to that of bacteriophages, which are natively encoded by some prokaryotic organisms. GTAs have been detected in several members of the *Rhodobacterales* order, such as *Phaeobacter piscinae* S26, which is also known to produce the antibiotic secondary metabolite tropodithietic acid (TDA). Transcriptome and proteome profiling demonstrate that the TDA-deficient mutant S26 Δ tdaB has increased protein abundance related to a putative GTA and a prophage. The purpose of this study was to detect and quantify the *P. piscinae* S26 GTA and verify that it was expressed at a higher quantity in the TDA-deficient mutant compared to the wildtype (WT). To attempt a direct detection of GTA particles, we sampled and fixated *P. piscinae* S26 WT and S26 Δ tdaB monocultures at 24, 48, 72 and 96 hours for flow cytometry analysis. It was not possible to quantify GTAs using flow cytometry, possibly due to their small size. However, induction of a prophage was detected, and the prophage particle count was significantly higher in S26 Δ tdaB compared to the WT at 48 ($p < 0.01$) and 72 ($p < 0.01$) hours of growth. To investigate GTA gene expression patterns, we sampled *P. piscinae* S26 WT and S26 Δ tdaB monocultures at 24, 48 and 72 hours for RT-qPCR analysis of the GTA major capsid gene. The strain S26 Δ tdaB has a significantly higher expression of the GTA major capsid gene after 48 hours ($p < 0.05$) and 72 hours ($p < 0.01$) of growth compared to S26 WT. This would suggest that in *P. piscinae* S26, the ability to produce TDA positively impacts GTA gene expression and prophage induction of the producer.

Study importance

Transcriptome and proteome profiling has demonstrated that the ability of *P. piscinae* S26 to produce the antibiotic secondary metabolite tropodithietic acid (TDA) affects the expression of GTA genes and induction of a prophage. However, little is known about the relationship between microbial secondary metabolism and horizontal gene transfer (HGT) mediated by GTAs and phage transduction in bacteria in natural microbiomes. There is currently no reliable method for direct quantification of GTAs. This project aims to detect and quantify a putative *Phaeobacter* sp. gene transfer agent (GTA) in *P. piscinae* S26 using RT-qPCR and flow cytometry (FCM). This is done to investigate how production of a microbial secondary metabolite can influence GTA-mediated HGT in natural microbiomes.

Background

Introduction

A single milliliter of coastal seawater contains anywhere from 10^6 to 10^{11} viral particles (Brusaard et al., 2016). Marine viruses shape the microbial communities of their respective ecosystems by transferring an estimated $\sim 10^{29}$ genes/day through transduction (Paul et al., 1999). However, several virus-like elements (viriforms) may also play a significant role in horizontal gene transfer (HGT) in marine microbial communities - among them the so-called gene transfer agents (GTAs). GTAs are phage-like elements natively encoded in prokaryotic genomes (Lang, et. Al 2017) and they facilitate HGT in bacterial communities (Brimacombe et al., 2015). GTAs were initially identified in 1974 in *Rhodobacter capsulatus* (RcGTA, Marrs et al., 1974), a member of the Rhodobacteraceae family. This family encompasses the globally distributed *Roseobacter* group, wherein GTAs have also been observed in multiple species, including *Phaeobacter* spp. *Roseobacter* members may represent up to 8% of surface water bacterioplankton globally (Wagner-Döbler et al., 2006; Wietz, Gram & Jørgensen et al., 2010), and are important players in the global marine biogeochemical cycling.

Several species of the *Phaeobacter* genus produce tropodithietic acid (TDA), a secondary metabolite with antibiotic, cation-chelating, and quorum sensing (QS) properties (Beyersmann et al., 2017; Henriksen et al., 2022). The antimicrobial mechanism of action (MoA) of TDA relies on an antiporter mechanism where the cytosol pH is lowered, and cations diffuse through the cell membrane (Beyersman et al., 2017). TDA production, or lack thereof, affects the expression of a putative *Phaeobacter* sp. GTA as demonstrated by transcriptomics and proteomics profiling (Lindqvist et al., 2023). Understanding the role of secondary metabolites in GTA expression is a crucial step in elucidating the mechanisms by which microbial communities interact and exchange genetic information.

Gene transfer agents

GTAs are tailed, dsDNA containing capsid particles, akin to that of bacteriophages. They are encoded in some prokaryotic genomes and mediate genetic transfer between cells (Lang et al., 2017). They were discovered in 1974, and as with many other biological phenomena, it happened by accident. While investigating means of gene transfer in purple phototrophic bacteria, Barry Marrs serendipitously discovered a unique and novel vector in *Rhodobacter capsulatus* (Marrs et al., 2002). By exhibiting resistance to DNases and lacking a requirement for cell-to-cell contact, the mechanism was determined to be most similar to that of phage transduction. However, the sedimentation coefficient of 70S-rRNA was found to be much lower than that of any known transducing phage (Marrs et al., 1974). As a result,

the mechanism was originally coined “capsduction” in honor of the species in which it was first discovered. Over the next few years, the nature of this new vector was further elucidated in the research group of Marrs (Marrs et al., 2002). The phage-like particles produced in *R. capsulatus* would go on to become the most studied GTAs and are known as *R. capsulatus* gene transfer agents (RcGTAs). This section will focus on RcGTAs, but it is important to note that other genetically distinct GTAs have been discovered, suggesting multiple instances of convergent evolution (Table 1) (Lang et al., 2017). Recently, transcriptome and proteome profiling also demonstrated that another member of the Rhodobacterales order, *Phaeobacter piscinae* S26 encodes several GTA genes (Lindqvist et al., 2023).

Table 1: List of GTA-producing organisms and their respective GTA families. Adapted from Lang, 2017.

GTA Family	Producing organism(s)	DNA packaged	Head diameter
		(kb)	(mm)
RcGTA	<i>Rhodobacter capsulatus</i> ; other members of order Rhodobacterales	4.5	30
Dd1	<i>Desulfovibrio desulfuricans</i>	13.6	43
VSH-1	<i>Brachyspira hyodysenteriae</i>	7.5	45
VTA	<i>Methanococcus voltae</i>	4.4	40
BaGTA (BLP)	<i>Bartonella</i> spp.	14	40

Properties of RcGTAs

GTAs facilitate horizontal gene transfer (HGT), which typically refers to one of three well-studied mechanisms: transduction, conjugation, and transformation. However, GTAs mediate HGT in a quite unique manner. As Barry L. Marrs discovered in 1974, the mechanism of RcGTAs seem to be analogous to transduction (Marrs et al., 1974), which is generally recognized as the most impactful mechanism of gene transfer between bacteria (Chen et al., 2018). But it also involves aspects of natural transformation, as import genes are required in the recipient cell (Brimacombe et al., 2015).

RcGTA particles are exceptionally small. Using cryo-electron microscopy (cryo-EM) the structure has been determined before and after DNA is released from the capsid. If symmetry in the structure is assumed, the reconstructions of the oblate head, portal complex, tail and baseplate have a resolution of 3.3-4.5 Å. If symmetry is not assumed, the resolution is 4.3 Å. It has been speculated that the oblate heads of RcGTAs are the smallest capsid of a tailed phage particle that can possibly be assembled (Figure 1) (Bárdy et al., 2020).

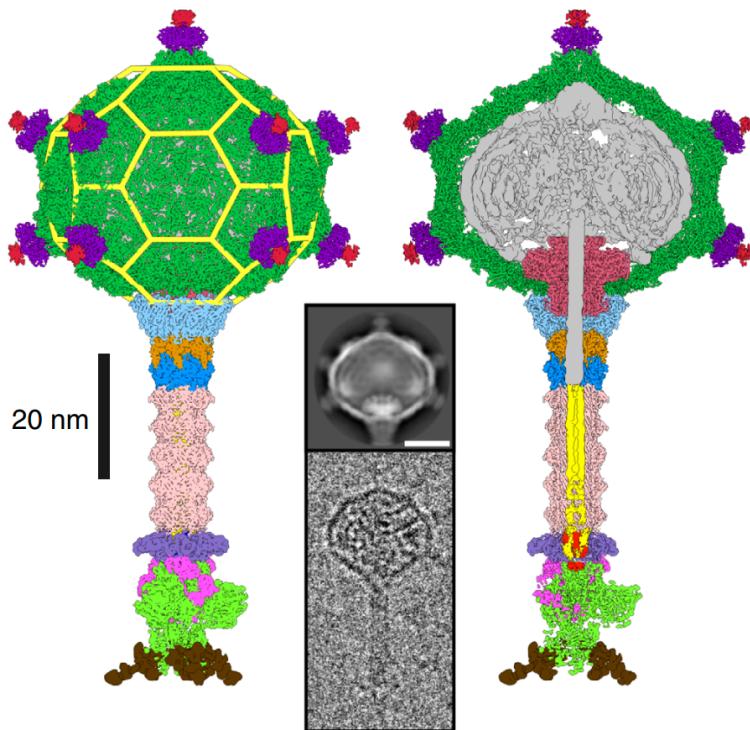


Figure 1: Cryo-EM reconstruction of the structure of the RcGTA particle from *R. capsulatus* strain DE442, calculated from 42.242 particle images in Pavol Bárdy's research group. The left side shows the complete particle, and on the right the front half has been removed to show internal DNA and proteins (Bárdy et al., 2020).

In contrast to bacteriophage-mediated transduction, which primarily aims at self-preservation, RcGTAs package seemingly random DNA fragments, excised from the entire genome of the cell, and thus do not necessarily contain any genes related to its own production machinery (Hynes et al., 2012). This can be attributed to several factors. Firstly, the average size of a prokaryotic gene is around 1,000 base pairs, and 93% of *R. capsulatus* genes are shorter than 2000 base pairs. This makes the 4,000-4,500 base pairs of dsDNA packaged in the RcGTA capsid suitable to transfer genes in *Rhodobacteraceae* (Bárdy et al., 2020). However, 4,000-4,500 bp is insufficient to encode the major cluster of structural proteins required for their production, which is about 15,000 bp in size (Lang et al., 2000). Genes encoding proteins forming the RcGTA particles are even packaged less frequently compared to the other genomic regions (Hynes et al., 2012). Moreover, the genes required to regulate and produce RcGTAs are distributed over 5 distinct loci. (Hynes et al., 2016).

Mechanism of RcGTAs

As previously mentioned, the RcGTA particles require expression of a structural gene cluster of approximately 15 kbp (Lang et al., 2000). At least 5 auxiliary clusters are required to express attachment, release, adsorption, gene regulation and virion maturation genes (Hynes et al., 2016). Although the precise mechanism of RcGTAs has not been definitively established, a proposed mechanism has been described and will be outlined below.

Packaging and lysis

In dsDNA phages, the DNA packaging mechanisms are dependent upon an endonuclease complex known as terminase. The terminase complex is responsible for initiating the cutting, packaging, and translocation of DNA into the particle (Black et al., 1989). Likewise, the RcGTA structural gene cluster does indeed encode a large terminase gene *orfg2* (Lang et al., 2000), which shows homology to known T4 phage terminase which has no sequence specificity. This sequence-independent terminase would support the idea that RcGTAs package DNA essentially by random, using a headful packaging mechanism (Hynes et al., 2012). It is speculated that the small size and reduced DNA packaging capacity of the RcGTA heads may be a mechanism in place to prevent the separation of the RcGTA propagation from the producer cells (Bárdy et al., 2020). Once the RcGTA particles are assembled and packaged with dsDNA, evidence suggest they are released through cell lysis. This is supported by the detection of normally intracellular pigments and expression of endolysin/holin system. (Fogg et al., 2012; Westbye et al., 2013).

Delivery and uptake

The delivery mechanism has been proposed based on the structure, and cryo-EM images of the RcGTA particles attached to *R. capsulatus* cells in Pavol Bárdy's research group. 11 head spikes are located on the head of the oblate capsid head. These spikes contain oligosaccharide-binding sites, which mediate the initial and reversible attachment to the cell. Cryo-EM images also reveal that numerous RcGTAs will bind to some *R. capsulatus* cells, whereas others do not attract any. This suggests that the ability of the cell to bind RcGTAs is heterogenous. The tail of RcGTAs contains 21 additional oligosaccharide-binding sites. The particles will initially attach to the cell in random orientations. It is proposed that passive penetration through the 40-130 nm thick cell capsule is achieved through dynamic binding of the various receptor-binding sites on the cell capsid (Bárdy et al., 2020).

It is vital for successful DNA ejection that the RcGTA particles orient with their baseplate toward the inner membrane of the cell. This is presumed to be enabled through binding of tail fibers to an outer membrane receptor. The baseplate of the cell may be destabilized by additional receptor bindings, and

a transmembrane pore will be formed. Then, conformational changes of the baseplate may trigger release of inner tail proteins along with a tail peptidase, allowing packaged dsDNA to be ejected into the periplasm of the cell. The dsDNA can remain in the periplasm of the cell for several hours before it is finally imported into the cytoplasm (Bárdy et al., 2020). Interestingly, in the recipient cell, a natural transformation-like machinery, must be induced in order to facilitate successful DNA uptake and recombination (Brimacombe et al., 2014; Brimacombe et al., 2015). This means that the RcGTAs delivery is entirely dependent on the uptake capability of the recipient cells. This supports the idea that RcGTAs does indeed provide a selective advantage on the part of *R. capsulatus* cells and not just an accidental and selfish phenomenon. Once the RcGTAs have entered the recipient cell, the dsDNA is suggested to be incorporated into the genome by a homologous recombination, through a pathway resembling natural transformation. Additionally, the uptake process relies on the presence of DprA, which demonstrates a specific binding affinity towards incoming dsDNA, effectively shielding it from potential endonuclease-mediated breakdown (Brimacombe et al., 2014).

Detection and quantification of RcGTAs

Due to the small size of GTAs and our relatively limited understanding of the complete cycle, methods to directly detect and accurately quantify them are lacking. Intact RcGTAs have only been directly observed using electron microscopy (Bárdy et al., 2020; Westbye et al., 2016). This method, along with computational modelling and genome mining, has allowed researchers to construct complete models of the RcGTA structure (Figure 1). While this method produces extremely detailed images of the RcGTA particles, it does not reflect the expression dynamics of RcGTAs in natural systems, nor does it quantify the particles.

To this day, detection of GTAs is often implicitly elucidated by indirect methods such as metagenomics analysis, functional assays, proteomics and transcriptomics. The dsDNA content of RcGTA capsids has also been purified by Phenol:Chloroform DNA extraction (Fogg, 2019). Following agarose gel electrophoresis, a band at ~4-5 kbp could be observed, which is hypothesized to be RcGTA dsDNA. This method allows for easy and flexible confirmation of GTA content in a variety of samples and allows for purification of GTA dsDNA.

In recent years, functional RcGTAs have also been purified from monocultures of *R. capsulatus* using two-step preparative monolithic chromatography (TPMC) (Langille et al., 2022). Particles purified using this method were detectable using nanoparticle tracking analysis (NTA) and were confirmed functional based on a rifampicin resistance (Rif^R) gene transfer bioassay. According to the authors, this method is applicable to other GTA-producing species. While GTA detection and purification methods

have improved remarkably over the last years, a simple and reliable detection assay that would allow for detection and quantification of GTAs is still missing.

Regulation of GTA production and release

GTA production comes at a high cost to the individual cell, as it must lyse to release the GTA particles (Fogg et al., 2012). It follows that GTA production and release is tightly governed by multiple regulatory systems, several of which are not entirely understood. Comprehensive models of RcGTA regulation in *R. capsulatus* have only recently been constructed (Fogg, 2019), and analogues of regulatory RcGTA genes and proteins have been observed in *P. piscinae* S26 (Lindqvist et al., 2023). It can be assumed that many regulatory pathways described in *R. capsulatus* is shared with other members of the Rhodobacteraceae family, but regulatory mechanisms unique to *Phaeobacter* spp. may exist, and are yet to be elucidated. It has been hypothesized that *R. capsulatus* and *P. piscinae* share a common regulatory mechanism: The GTA Activation Factor A (GafA) regulator, which is a transcriptional activator of GTA expression in *R. capsulatus* (Fogg, 2019; Lindqvist et al., 2023).

The GafA regulator

GafA is a direct activator of RcGTA production in *R. capsulatus* (Fogg, 2019). It is encoded in the *gafA* gene, and contains a predicted DNA-binding, helix-turn-helix domain (HtH) in its N-terminal domain, as well as a C-terminal domain with similarity to various sigma factors (Fogg, 2019). The GafA regulator was originally discovered in *R. capsulatus*, but has been detected in several members of *Rhodobacterales*, among them *P. piscinae* (Hynes et al., 2016; Lindqvist et al., 2023). Deletion of *gafA* in the RcGTA hyperproducer *R. capsulatus* DE442 abolishes RcGTA gene transfer, while overexpression of *gafA* in *R. capsulatus* SB1003 increased antibiotic resistance gene transfer 57-fold (Hynes et al., 2016).

GafA is proposed to interact directly with the RNA polymerase omega subunit, altering promoter specificity of the RNA polymerase holoenzyme and inducing expression of core RcGTA genes (Sherlock & Fogg, 2022). GafA is an autoinducer of *gafA* expression, yet only the C-terminal domain is required for this mechanism (Sherlock & Fogg, 2022). Additionally, positive autoinduction of *gafA* is dependent on the global regulator protein CtrA. Deletion of *ctrA* (encoding CtrA) abolishes RcGTA gene transfer and production of RcGTA capsid protein (Lang & Beatty, 2000). It has previously been demonstrated that overexpression of *gafA* in a *ctrA* knockout of *R. capsulatus* DE442 produced increased transcription of *gafA* (198-fold), terminase genes (126-fold) and capsid genes (22-fold), but diminished transcription of endolysin genes, which would prevent release of RcGTAs (Fogg, 2019). In its phosphorylated form, CtrA~P is also required for production of secondary structural proteins, which are required for infectivity.

This indicates that both GafA and CtrA must work in concert to produce functional RcGTAs, and that RcGTA production is regulated at multiple levels: Production, maturation, and release.

Growth phase dynamics and quorum sensing

Production of RcGTAs is highest in the stationary phase, which is characterized by high cell density, reduced cell replication and limited nutrient availability. These conditions positively affect RcGTA production (Solioz & Yen et al., 1975; Hynes et al., 2012; Brimacombe et al., 2013). Quorum sensing (QS) occurs at high cell density in many organisms, and in *R. capsulatus*, production of RcGTAs is dependent on the GtaIR QS system at every level of production (Leung et al., 2012; Fogg, 2019). The GtaR protein contains a DNA-binding motif, which is proposed to bind to the *gafA* promoter, repressing expression (Leung et al., 2012; Fogg, 2019). The acylhomoserine lactone synthase, GtaI, synthesizes N-acyl homoserine lactones (AHLs), which interact with promoter-bound GtaR to release it from the target promoter. Additionally, GtaR also binds directly to GafA, repressing binding affinity to all GafA targets.

Effects of tropodithietic acid (TDA) on GTA production

TDA is a secondary metabolite with antibiotic, cation-chelating and proposed QS functions produced by several members of the *Roseobacter* group (Geng & Belas, 2010). Recent transcriptome and proteome profiling of a TDA-deficient mutant of *P. piscinae* S26 (*S26ΔtdaB*) indicates that abolishment of TDA production leads to upregulation of several systems related to horizontal gene transfer (HGT), including a putative GTA and a GafA homologue (Lindqvist et al., 2023). In surface-level marine environments inhabited by *Phaeobacter* spp., TDA has a minor by measurably effect on the microbial community composition (Bech et al., 2023). Based on these findings, it is hypothesized that TDA may play a key role in coordinating colonization and adaptation to new environments (Henriksen et al., 2022). TDA is proposed to act as a QS signal through the PgaRI system, which is homologous to the *gafA*-repressing LuxRI system in *Dinoroseobacter shibae* (Beyersmann et al., 2017; Lindqvist et al., 2023). This implies a reduction in GTA expression in TDA-producing cells. TDA may interact directly with the PgaR AHL-regulator, which in turn reduces binds to the *gafA* promoter, reducing expression of GafA (Figure 2). The connection between TDA and GTA production may shine light on the ecological roles of GTA-mediated HGT in natural microbial communities.

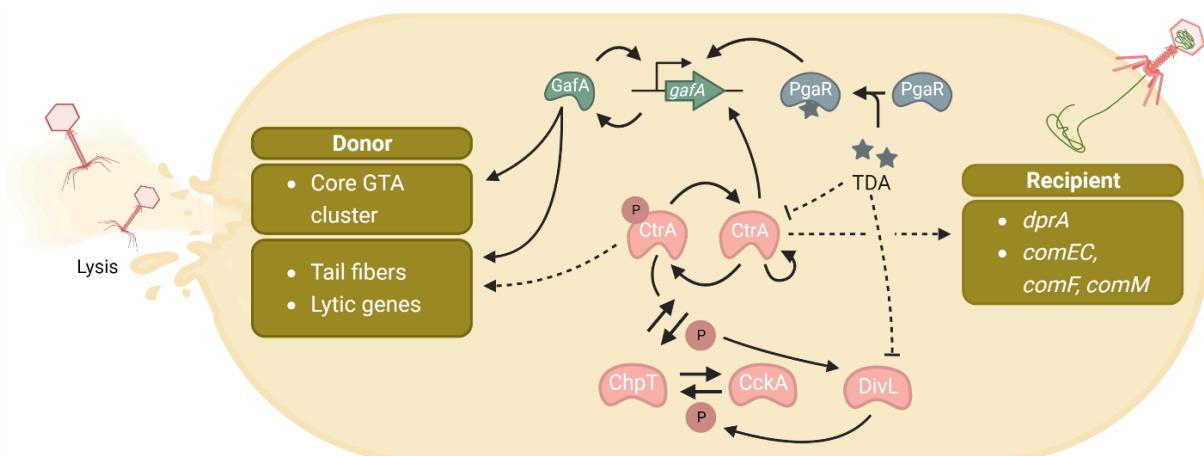


Figure 2: Proposed interaction pathway between TDA and GTA release in *P. piscinae* S26. Figure from Lindqvist et al., 2023.

Evolution and ecological role of GTAs

RcGTAs have previously been described as defective prophages (Redfield et al., 2001). However, they exhibit characteristics that go beyond being mere defective prophages and seem to provide a selective advantage to the producing organisms. Deciphering the evolutionary basis for sustaining GTA production poses a challenge, as the cells responsible for actively producing and releasing GTAs meet their demise.

A proposed theory of the origin of RcGTAs is that they have evolved from a prophage early in the alphaproteobacterial lineage (Lang et al., 2017). This is in part supported by the discovery of novel phages that contain homologs of the structural and regulatory genes of RcGTA (Zhan et al., 2016; Huang et al., 2011; Zhang, & Jiao, 2009; Bollivar et al., 2016). Furthermore, in two phages that infect members of the order Rhodobacterales, proteins homologous to the response regulator CtrA proteins of alphaproteobacteria has been found, suggesting a link between a phage ancestor and RcGTA regulation by CtrA (Brilli et al., 2010). By analyzing similarities between phages and bacterial sequences, it becomes evident that the phage sequences are not recent acquisitions. This is indicated by the fact that the phage sequences do not belong to the bacterial CtrA clade. The existing understanding of convergent evolution reveals that diverse taxa have independently developed at least 5 genetically distinct GTAs (Table 1), which further strengthens the notion that GTAs confer a selective advantage to the organisms that produce them (Lang et al., 2017).

Given the premise that prophages have most probably undergone evolutionary development from a prophage in the distant past: Why does the production of GTAs provide a selective advantage

for the producer? To address this question effectively, it is important to approach the problem from an evolutionary or systemic perspective, rather than solely focusing on individual bacteria. The primary focus of most theories concerning GTAs is the idea of division of labor and social behaviors observed in microbial communities. This can lead to the emergence of specialized individuals within a population who undertake resource-demanding behaviors, ultimately benefiting the population as a whole (West et al., 2016; West et al., 2006). This would be supported by the coordinated upregulation of GTA production and recipient capability by QS as well as the fact that only a small subpopulation of <3% of the community produce GTAs (Fogg, 2019; Ding et al., 2019). RcGTAs seem to transfer DNA only to closely related cells (Wall et al., 1975), and enable homologous recombination of alleles within a population as opposed to the transfer of novel sequences (Brimacombe et al 2015). This has led to the emergence of several theories, of the possible benefits of GTAs, which include, but are not limited to, repair of damaged alleles (Vos, 2009; Takeuchi et al., 2014; Gozzi K, 2016) and removal of invading mobile genetic elements (Croucher et al., 2016).

Scope and aim of experiments

Due to their small size and limited knowledge, there is currently no consistently reliable method for detection and quantifying GTAs. The purpose of this project is to detect and quantify GTAs of the *Phaeobacter* GTA in an *P. piscinae* S26 wildtype strain (S26 WT), as well as a TDA-deficient mutant (S26 Δ tdaB). This will be attempted by measuring transcription of the *Phaeobacter* sp. GTA major capsid gene through an RT-qPCR assay, as well as time-resolved flow cytometry (FCM). Comparing the S26 WT and S26 Δ tdaB strains, the experiments aim to describe the effect of host TDA production on *Phaeobacter* sp. GTA production. Based on proteomics and transcriptomics data obtained by Lindqvist et al. in 2023, the expression of the GTA major capsid gene is expected to be higher in the S26 Δ tdaB strain than the S26 WT strain. Ideally, the GTA particles should be detectable by flow cytometry after nucleic acid staining with SYBR Green H.

Materials and methods

Bacterial strains used in this study

Expression of the putative *Phaeobacter* GTA was originally inferred from proteomics and transcriptomics data obtained from *P. piscinae* S26 WT / S26 Δ *tdaB* by Lindqvist et al. The same *P. piscinae* strains were used in all experiments described in this report (Table 2). The marine genus *Phaeobacter* belongs to the *Roseobacter* subgroup within the *Rhodobacteraceae* family (Simon et al., 2017), and encompasses the strain *P. piscinae* S26. This strain was originally isolated from a Greek sea bass aquaculture unit in 2013 (Sonnen schein et al., 2017) and is, like several other members of the *Roseobacter* group, a producer of the antibiotic secondary metabolite TDA. To study the effect of host TDA production in *Phaeobacter* spp., a scarless, TDA-deficient mutant (S26 Δ *tdaB*) was produced by Laura Louise Lindqvist and colleagues in 2022 (Table 2). In this study, strains S26 and S26 Δ *tdaB* strain are compared to analyze the effect of TDA biosynthesis potential on GTA gene expression and particle quantities in monocultures.

Table 2: List of strains used in this study.

Organism	Strain	Description	Source
<i>Phaeobacter</i> <i>piscinae</i>	S26 Wildtype (WT)	Isolated from a Greek sea bass aquaculture unit in 2013 and characterized by Grotkær et al., in 2016. GenBank no. AJ536670.1.	Grotkær et al., 2016; Sonnen schein et al., 2017
<i>Phaeobacter</i> <i>piscinae</i>	S26 Δ <i>tdaB</i> (Mut)	Markerless, in-frame deletion of the <i>tdaB</i> gene in S26. TDA-deficient mutant of <i>P. piscinae</i> S26.	Lindqvist et al., 2023

Media and culture conditions

For all cultivations, -80°C cryostocks of *P. piscinae* S26 (WT and S26 Δ *tdaB*, Table 2) were cultured on Marine Agar (MA, BD Biosciences, Cat. 212185) for 48h, or until individual colonies could be observed with the naked eye. Individual colonies were picked using sterile inoculation loops and transferred to 2 mL Marine Broth (MB, BD Biosciences, Cat. 27910) in 15 mL Falcon™ Centrifuge Tubes (Thermo Fisher Scientific, Cat. 352196). Precultures were incubated at 25°C for 24h w. shaking at 200 RPM.

Precultures were diluted from 10^8 to 10^6 CFU/mL in sterile demineralized water containing 3% w/v Instant Ocean salts (IO 3%, Aquarium Systems Inc., Prod. SS15-10) by serial dilution. In TPP6-well tissue culture plates (Merck, Cat. Z707759), and cultures were inoculated to 5 mL IO medium w. Casamino Acids and glucose (IOCG, 30 g/L Instant Ocean salts (IO, Aquarium Systems Inc., Prod. SS15-10) + 3 g/L HEPES + 3 g/L Bacto Casamino Acids (Thermo Fisher Scientific, Cat. 223050) + 2 g/L glucose (Sigma-Aldrich, Prod. D9434)) to a final concentration of 10^4 CFU/mL. Cultures were placed in a humidity chamber and incubated at 25°C under stagnant conditions. Experiments were carried out using five biological replicates for flow cytometry (FCM) experiments, and four biological replicates for reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) experiments.

One mL samples for RNA extraction were taken in 2 mL Eppendorf tubes. 1 mL RNAProtect Bacteria reagent (Qiagen, Cat. 76506) was added to each tube, which were incubated at room temperature (approx. 20°C) for 5 minutes. Cells were harvested in a microcentrifuge at 10.000 x g for 5 minutes. The supernatant was removed by careful pipetting, and the pellet was stored at -80°C until RNA extraction.

RNA extraction/purification and cDNA library construction

RNA extractions were performed using the RNeasy Micro Kit (Qiagen, Cat. 74004). 2x20+1x10 µL aliquots of eluted RNA were stored at -80°C. Quality of RNA was tested by gel electrophoresis (100V, 40 min.) of 5 µL samples. Following gel electrophoresis, gels were stained in 0.0001%w/v ethidium bromide for 20 minutes. Imaging of gels was performed using the Gel Doc™ XR+ Molecular Imager (Bio-Rad, Cat. 1708195EDU). Quality of RNA was determined by the presence of strong bands from 16S-, 23S- and 5S-rRNA, as well as the amount of smearing. Size of samples were measured against a GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific, Cat. SM0311). Genomic DNA (gDNA) was removed from samples by DNase treatment with DNA-free™ kit (Thermo Fisher Scientific, Cat. AM1906). For each 20 µL sample, 0.1 volumes of 10X DNase buffer and 2 µL of DNase was used. Removal of DNA was confirmed by agarose gel electrophoresis, similar to quality testing of RNA quality. A cDNA library was constructed from 10 µL DNase treated template by reverse transcriptase treatment using the SuperScript™ IV Reverse Transcriptase kit (Invitrogen, Cat. 18090010). 1 µL of 50 µM random hexamers (Thermo-Fischer, Cat. N8080127) was used as a primer.

Analysis of GTA gene expression by RT-qPCR

Expression levels of the OL67_001823 gene, which codes for the major capsid structure of the *Phaeobacter* GTA (Lindqvist, L.L. et al., 2023), were analyzed by RT-qPCR. For future reference, the OL67_001823 gene will be referred to as the “capsid” gene. The workflow for processing of culture samples can be found in section “**RNA extraction/purification and cDNA library construction**”. All

PCR reactions were performed using the T100 Thermal Cycler (Bio-Rad, Cat. 1861096), and all qPCR reactions were performed using the CFX Opus 96 Real-Time PCR system (Bio-Rad, Cat. 12011319). Primers were designed based on the full genome sequence of *P. piscinae* S26 (Sonnenschein & Gram et al., 2021. NCBI Ref. NZ_CP080275.1) using the NCBI Primer-BLAST tool (Ye et al., 2012) (Table 3). The housekeeping gene *rpoB*, coding for the prokaryotic RNA-polymerase subunit β, was chosen as a reference housekeeping gene for expression normalization.

Table 3: List of primers used in this study.

Primer	Amplified gene	Sequence (5'-3')	Provider
Major capsid Fw	OL67_001823	CACTGTGACCGAACATCCA	Merck
Major capsid Rv	OL67_001823	GTCAAAGGCACTGTCATCCA	Merck
RNA-polymerase subunit β Fw	<i>rpoB</i>	AAACGGGCATCCAGAGCA	Merck
RNA-polymerase subunit β Rv	<i>rpoB</i>	AGATCGAGCGTGAAGAAGT	Merck

Binding specificity of primers was confirmed by temperature gradient PCR on *P. piscinae* S26 gDNA using TEMPase Hot Start 2X Mastermix (Avantor, Cat. 733-2415). Temperature gradient was set to 55 - 65°C. Gel electrophoresis (100V, 40 minutes) visualization was performed to assess binding specificity, and to determine the optimal temperature. Optimal annealing temperature for primers was determined to 62°C.

For each qPCR reaction, two master mixes were prepared – each with their own primer pair: The major capsid gene (target) and *rpoB* (housekeeping gene, HKG). Master mixes were prepared using Luna Universal qPCR Master Mix (New England Biolabs Inc., Cat. M3003S). 20 µL (19 µL master mix + 1 µL template) qPCR reactions were run in non-skirted, white 96-well plates. A standard curve to test primer binding affinity was included by using a serial dilution of 1 µL raw WT RNA extraction (10^{-1} – 10^{-4} , in nuclease free water) as template. Raw RNA extraction samples were included as positive controls in two technical replicates. DNase treated samples were included as negative controls in one technical replicate. Samples of cDNA library, the actual samples of interest, were included in 2 technical replicates. Furthermore, negative controls with nuclease free water as template were included in three technical replicates.

Detection of a prophage from *P. piscinae* S26 monocultures by flow cytometry

Cultures of *P. piscinae* S26 WT and S26 Δ tdaB were grown in accordance with section

Media and culture conditions in replicates of 5 with 1 negative medium control in each TPP6

plate. Final culture volume was 7 mL/well with an estimated starting cell concentration of 10⁴ CFU/mL. Culture samples of 1 mL were taken for flow cytometry (FCM) analysis at time points 24h, 48h, 72h and 96h with corresponding 100 μ L samples for determination of colony forming units (CFU) per mL.

FCM samples were transferred to Nunc CryoTubes (Merck, Cat. V7884). For fixating the cultures, 20 μ L 25%v/v glutaraldehyde was added to each tube to a final concentration of 0.5% v/v. Tubes were mixed by tilting approx. 10-20 times. Tubes were incubated for 15 minutes in the dark at 4°C, then flash-frozen in liquid nitrogen (LN2) and stored at -80°C. For FCM analysis, samples were transported on LN2 to the University of Copenhagen (KU), Marine Biology Section in Helsingør. All FCM experiments were performed under guidance of Prof. Matthias Middelboe (KU, Section for Marine Biology).

Samples were diluted 50x in Tris-EDTA buffer (TE buffer (10 mM Trishydroxymethyl-aminomethane, 1 mM ethylenediaminetetraacetic acid), Invitrogen, Cat. 12090015) in Falcon™ polystyrene round-bottom tubes (Fisher Scientific, Cat. 14-959-40B) by serial dilution. For staining of nucleic acids, a commercial stock of SYBR™ Green H Nucleic Acid Gel Stain (10.000x concentrate in DMSO, Thermo Fisher Scientific, Cat. S7563) was added to a 0.5*10⁻⁴ dilution of stock. Blanks were prepared from autoclaved, 0.2 μ m filtered TE-buffer and treated identically to samples. All FCM analysis was performed using the BD FACSCanto™ II flow cytometer (BD Biosciences). To determine flowrate of samples, two BD Trucount absolute counting tubes (BD Biosciences, Cat. 340334) were analyzed for each run: One before samples, and one after the sample analysis. Each sample was run for 30 seconds. The following voltage settings were used:

Table 3: Voltage settings used for flow cytometry analysis of *P. piscinae* WT and Mut monocultures.

Sample	Forward scatter (FSC) voltage [V]	Side scatter (SSC) voltage [V]	SYBR Green H voltage [V]
Trucount beads	404	390	379
<i>P. piscinae</i> S26 culture samples	404	604	485

Flow cytometer setup, calibration and data collection was performed with the BD FACSDiva™ (BD Biosciences) software. Gating and labelling of measurements was performed using the Flowing Software package (Cell Imaging and Cytometry, Turku Bioscience Centre. 2013). All samples from WT

samples were run at the predefined “High” flowrate, and the Mut samples were run at the predefined “Medium” flowrate. The gates were shaped to encompass the appropriate particles by examining the fluorescence intensity and the forward scatter (Figure 4). This was completed with the guidance and expertise of Prof. Matthias Middelboe, PDR. Morten Dencker Schostag & Ass. Prof. Mikael Lenz Strube. The “Bacteria” gates were adjusted to accommodate morphological differences between *P. piscinae* S26 WT and S26 Δ tdaB.

Data analysis

The stochastic expression patterns across the samples posed a challenge, making the data analysis particularly puzzling and rendering traditional parametric analysis tools, such as the student's T-test, invalid, as the variance could no longer be assumed to be normally distributed. As a result of this, multivariate linear models (MLM) were chosen as the analysis approach for both experiments with the help of Assoc. Prof. Mikael Lenz Strube. The data was transformed which had the advantage of stabilized variance and increased statistical power by accounting for all variables including covariates.

Flow cytometry: Gating and labelling of the flow cytometry measurements was performed using the Flowing Software package (Cell Imaging and Cytometry, Turku Bioscience Centre. 2013). This resulted in particle counts of 4 distinct groups: Virus/Phages, Cell debris, Bacteria and True count (Figure 4). Truecount beads were used to estimate the exact flowrate for each run. Subsequently, the events in each gate were counted for all samples to prepare for data analysis using R (R Core Team, 2021). Data was visualized in a combined plot to show the bacterial and virus/phage counts for each strain, replicate and timepoint (Figure 5). To stabilize the variance, a $\log_{10}(x + 1)$ transformation was used for the bacterial and viral counts. The addition of one (+1) was done as values of zero were present in the data in the data, which similarly to negative values cannot be processed directly by the logarithm function. A correlation plot / heat map was made to check for multicollinearity before making a statistical model (Figure 6). To investigate the combined effects on virus counts, we used the categorical variables of Timepoints (24, 48, 72) and Group (S26 WT, S26 Δ tdaB), and the continuous variable of bacterial counts in a multivariate linear model. We assessed the statistical effect of the variables using Analysis of Variance (ANOVA), including all interactions (Appendix fc.1). The model assumptions were assessed by examining diagnostics plots (Appendix fc.2). Significance of each variable in the model was calculated to determine their statistical significance on the viral count (Appendix fc.1). To compare virus counts between the two strains at each timepoint (pairwise comparison), the least-square means of the strain/time factor combination were calculated and compared (Appendix fc.1).

RT-qPCR: The qPCR data was first subjected to quality control and visualization using the CFX Maestro V2.2 software package (Bio-Rad). The standard curves were then evaluated to assess the efficiency of the primers, and to later correct the Cq values. To calculate the Cq scores for gene amplification, the average Cq score from duplicate measurements from each biological replicate was obtained. The relative gene expression ratio (GER) of target gene compared to the housekeeping gene *rpoB* was calculated using the Pfaffl method (Pfaffl, M.W., 2001). This method was used to incorporate the amplification efficiency of the target gene and reference gene primers used for each plate. The Pfaffl method uses the following formula to calculate GER:

$$GER = \frac{(E_{Caps})^{\Delta Cq\ Caps}}{(E_{rpoB})^{\Delta Cq\ rpoB}}$$

E describes primer efficiency, and ΔCq describes the difference between the mean WT C_q scores (control) and the mean C_q score for the given biological replicate. These calculations were performed in Microsoft Excel, and the raw data can be found in (Supplementary material. qPCR.1). RT-qPCR was performed on two separate experiments of exact same experimental conditions. This resulted in the addition of 3 replicates at the 72-hour time point, to expand the data size. To analyze the GER data, a multivariate linear model was made using R to describe the GER, similarly to the method as described in the flow cytometry data analysis section. GER values were log10-transformed to stabilize variance. The independent variables of the model were the categorical variables of Time (24, 48, 72) and Group (S26 WT, S26Δ*tdaB*). Significance of each term in the model was calculated by ANOVA, which included all second order interactions (Appendix qPCR.1). The model assumptions were assessed using 4 different diagnostics plots (Appendix qPCR.2). To compare GER between the two strains at each timepoint (pairwise comparison), the least-square means of the strain/time factor combination were calculated and compared (Appendix qPCR.3).

Software: Most of the data analysis, including all the statistical tests were performed in the R programming language (R Core Team, 2021) in the RStudio environment (RStudio Team, 2020). Especially helpful to the data analysis and visualization was the TidyR package (Wickham et al., 2023) and ggplot2 package (Wickham et al., 2016). A helpful package to perform pairwise comparisons using the ANOVA models was the R package “emmeans” (Lenth, V.R. et al., 2023). Some data handling and the Pfaffl method calculations of the qPCR data were partly done in Microsoft Excel.

Results

Detection of a potential GTA dsDNA band by electrophoresis of RNA extraction

After incubating S26 WT and S26 Δ tdaB cultures for 72 hours and extracting RNA, agarose gel electrophoresis revealed a distinct band in S26 Δ tdaB samples at around 4-5 kbp, and this was observed over multiple replicates and experiments (Figure 4). The length of the band closely matched that of packaged DNA of an RcGTA, suggesting that it may serve as an indication of GTA dsDNA presence in the sample. It should be noted that these RNA extractions also extract DNA as an artefact. For future reference, this band will be referred to as the GTA band. The content of the band was confirmed as DNA, as the band completely disappeared following DNase treatment with the DNA-freeTM kit (Thermo Fisher Scientific, Cat. AM1906).

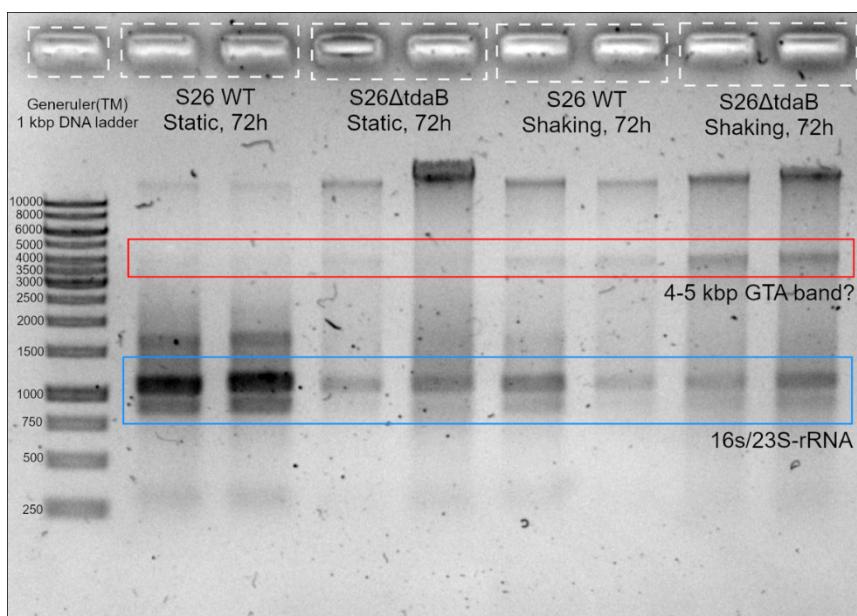


Figure 4: Agarose gel electrophoresis of RNA extractions performed on 72h cultures of *P. piscinae* S26 WT/Mut cultivated in static and shaking conditions. Hypothesized GTA band is marked in red, and 16S/23S-rRNA bands are marked in blue.

In both shaken and static cultures, the intensity of the GTA band was consistently higher in S26 Δ tdaB compared to the wildtype strain. However, regardless of the strain, shaken cultures displayed more intense GTA bands compared to static cultures. Interestingly, when genomic DNA (gDNA) extractions were performed using the NucleoSpin Tissue Kit (Macherey-Nagel, Cat. 740952.50) on the same cultures, no GTA band was observed after agarose gel electrophoresis.

Detection of a prophage from *P. piscinae* S26 monocultures by flow cytometry

Samples for FCM analysis were cultivated and sampled according to protocol “**Detection of a prophage from *P. piscinae* S26 monocultures by flow cytometry**”. The purpose of this experiment was to determine whether prophages or GTAs could be identified through flow cytometry, which could potentially serve as a reliable means of quantifying GTAs. The data for each sample were visualized and gated, and the measured particles clustered in clearly distinct groups of bacteria, virus/phages and cell debris (Figure 5). There was an abundance of virus/phage size particles in many of the samples from both wildtype and mutant strains. However, *Phaeobacter* GTA particles could not be quantified using flow cytometry. They were either indistinguishable from prophages or simply not observable by our FCM protocol, as the particle size was too small to detect. The “Virus” counts described in this section will thus refer to a separate, uncharacterized prophage.

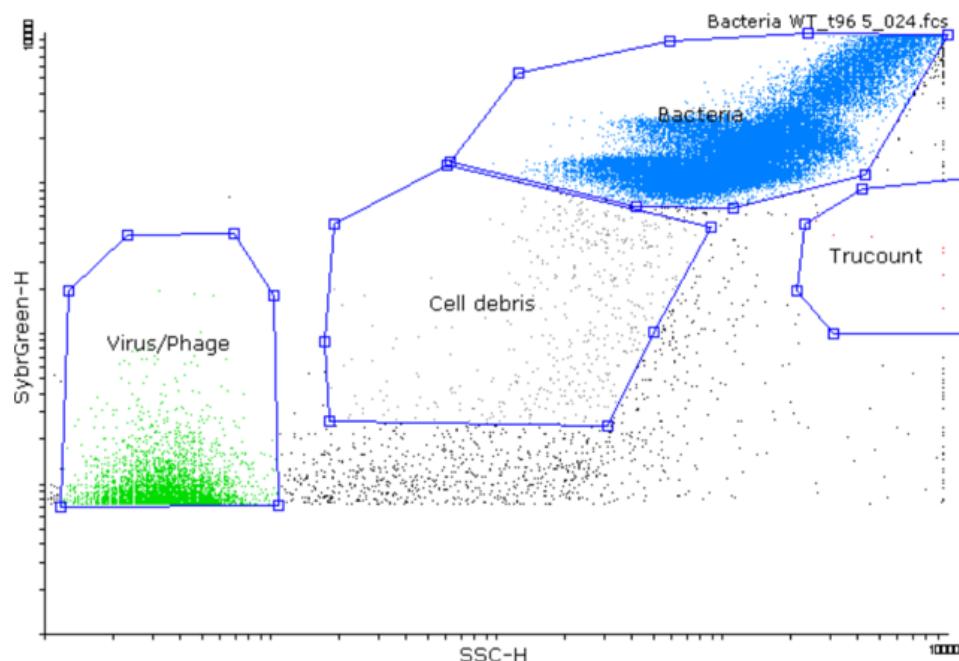


Figure 5: Cytogram showing gating of measurements from FCM analysis. Produced using Flowing Software (Cell Imaging and Cytometry, Turku Bioscience Centre. 2013). X-axis describes side scattering of light, and Y-axis describes intensity of SYBR Green H fluorescence signal. Figure shows a WT sample, $t = 96\text{h}$.

To further interpret the data for every sample at all time points, the flow cytometry date was subjected to statistical analysis (see “**Data analysis**

" section). Induction of the prophage was not observed at the 24h samples (Figure 6). Beyond this timepoint, expression of the prophage was highly stochastic, inducing at any point from 24 to 96h. A peak in virus counts were associated with a subsequent reduction in bacterial counts. In cultures where no, or few, phages were detected, bacterial counts did not decrease with time.

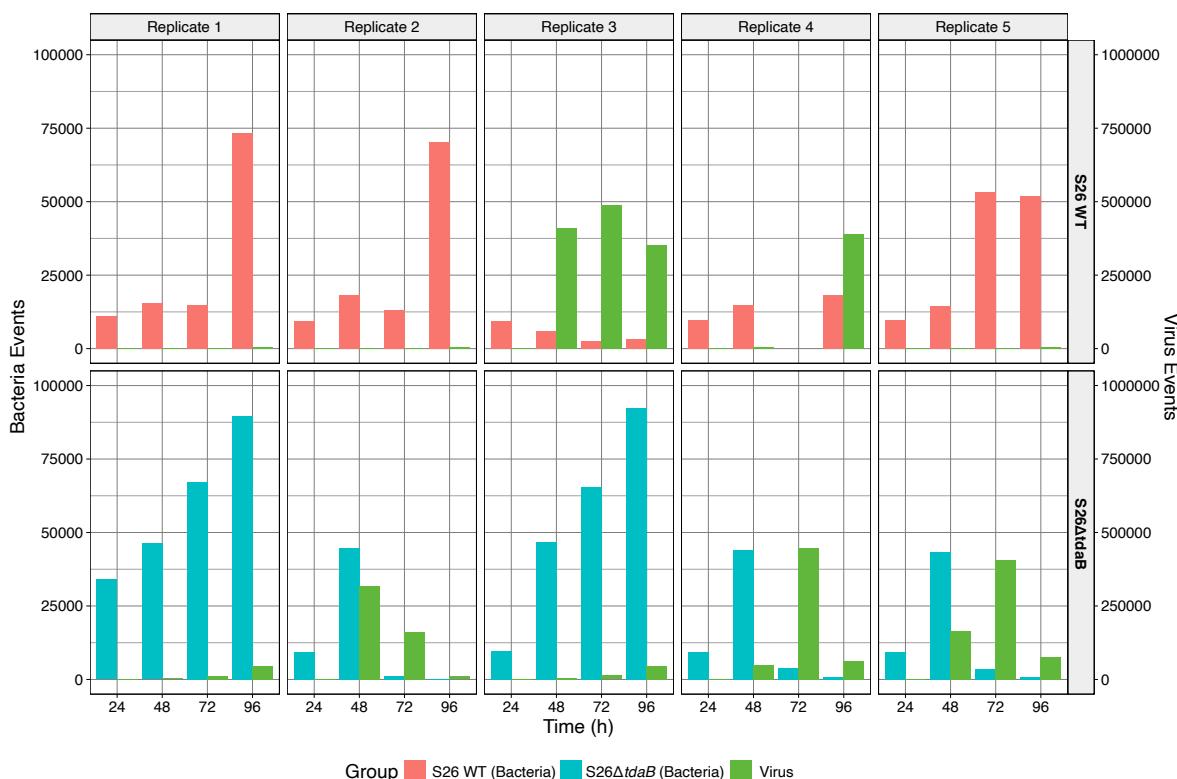


Figure 6: Counts of bacteria (left axis) and viruses (right axis) based on FCM analysis of *P. piscinae* WT (red) and Mut (cyan) monocultures. A single time-point measurement for WT Rep. 4, 72h was lost due to a technical error.

To gain insight into the dynamics of the flow cytometry data, count data from the gated variables was assessed for correlations. Interestingly, bacterial and viral counts were slightly negatively correlated (-0.18). Virus counts and cell debris were strongly positively correlated (0.86). Bacteria and cell debris counts were slightly positively correlated (0.29).

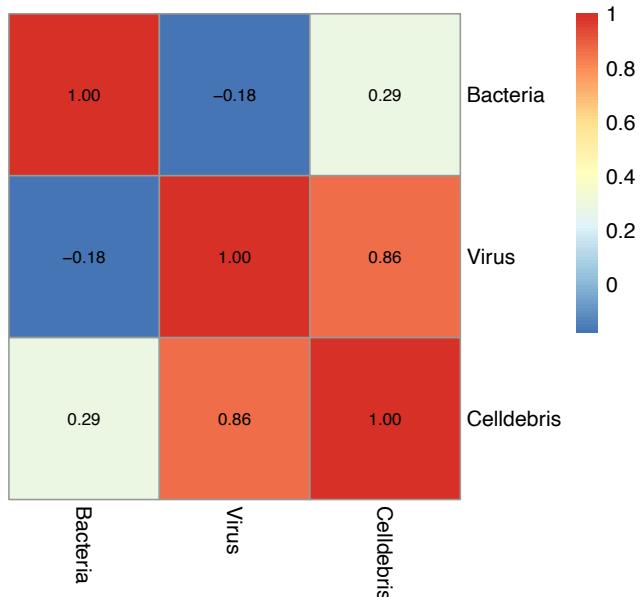


Figure 7: Correlation matrix on gated variables measured by flow cytometry. Blue indicates a negative correlation. Red indicates a positive correlation.

The Virus and Bacteria counts were $\log_{10}(x+1)$ -transformed to stabilize variance, and a multivariate linear model (MLM) was developed to examine the relationship between the viral count and the independent variables of time, group, and bacteria count. The statistical significance of the variables and their interactions in the MLM was evaluated by Analysis of Variance (ANOVA) (Appendix fc.1). The most significant variables of the model were Time ($p = 1.39 \cdot 10^{-7}$) and the interaction effect of bacterial counts and time ($p = 3.15 \cdot 10^{-4}$). The remaining statistically significant predictors of the Virus count in the model was Group (S26 WT, S26 $\Delta tdaB$), bacterial count as well as the interaction effects of Time:Group, Group:Bacteria and Time:Group:Bacteria (Appendix fc.1). At the time point of 48 and significantly higher viral counts were shown by the model ($p = 0.00576$), and at 72 hours an increase in viral counts were slightly significant ($p = 0.0504$).

Using the model, pairwise comparison was conducted to examine the disparities in virus observations between S26 $\Delta tdaB$ and S26 WT samples at different timepoints. The comparison involved analyzing the predicted viral counts generated by the model for S26 WT and S26 $\Delta tdaB$. This showed a significant increase of virus counts in S26 $\Delta tdaB$ compared to S26 WT at 48 and 72 hours ($p_{48h} = 0.00525$, $p_{72h} = 0.00340$) (Figure 8). No statistical difference in virus counts was estimated at 24 and 96 hours.

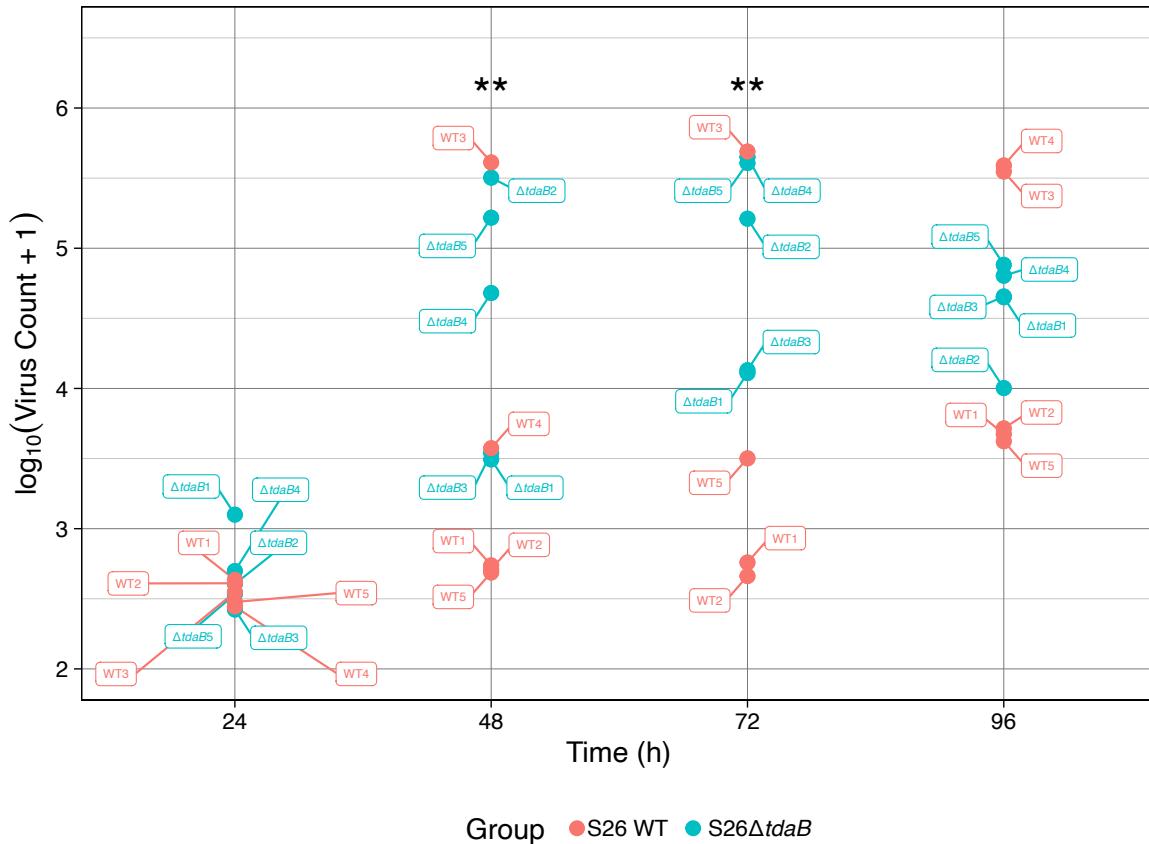


Figure 8: $\log_{10}(x+1)$ -transformed virus counts obtained from FCM analysis of *P. piscinae* S26 WT and S26 Δ tdaB monocultures. In a pairwise comparison using the MLM model, cultures of S26 Δ tdaB show significantly higher virus counts at 48h and 72h (Pairwise comparison, Bonferroni adjusted. $p_{48h} = 0.00524$, $p_{72h} = 0.00340$). Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Investigation of *Phaeobacter* sp. GTA capsid gene expression by RT-qPCR

To investigate the GTA expression levels in S26 WT and S26 Δ tdaB, RT-qPCR experiments were conducted to measure expression the capsid gene (OL67_001823) at 3 different time points: 24, 48 and 72 hours. Assessing the melting curves, negative controls and DNase treated wells showed no signal before 35 cycles, indicating a successful DNase treatment and no contamination of the samples. Positive controls consisting of raw RNA extraction samples produced a strong signal, indicating the presence of DNA in the original samples. With a visual inspection, the GERs revealed little to no increase of capsid GER at 24 hours in both strains, with one exception (S26 WT Replicate 3) (Figure 9). At the 48- and 72-hour time points a stochastic pattern of increased GER was observed, resulting in high degree of variance within the groups.

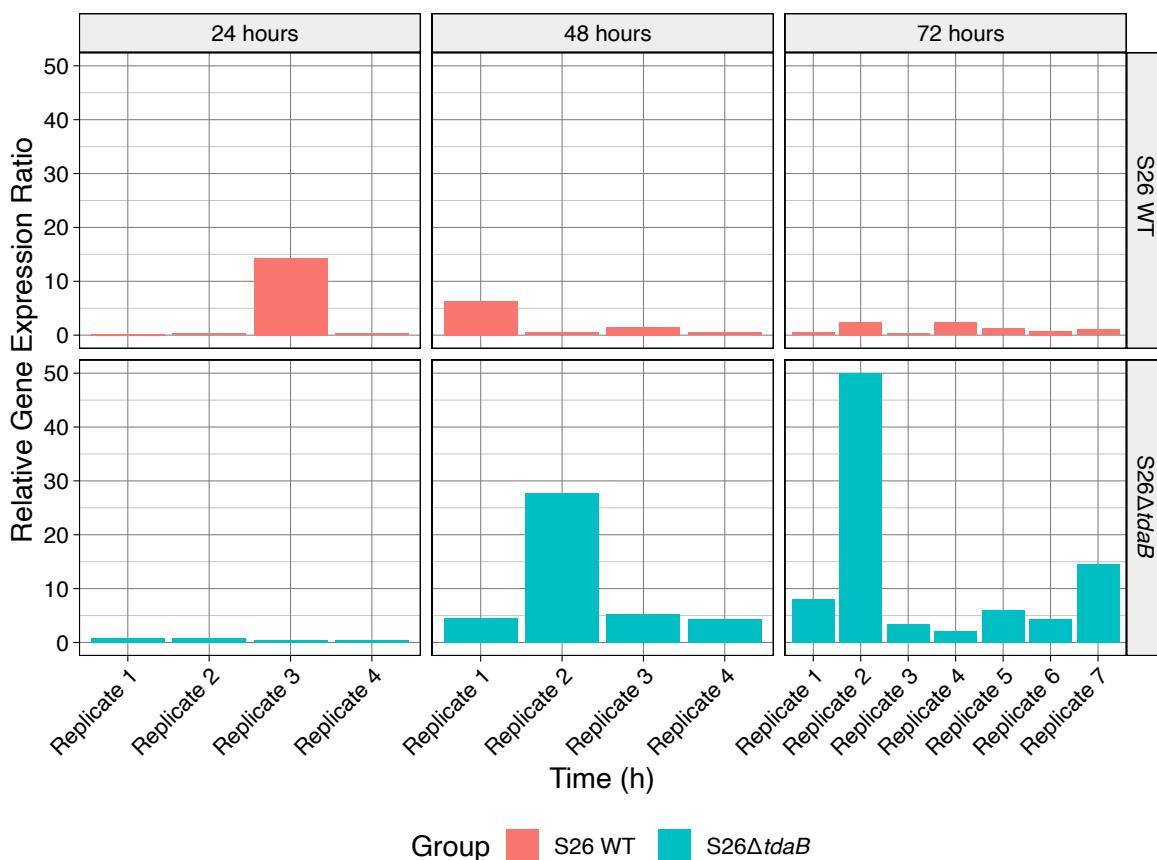


Figure 9: Gene expression ratio (GER) of Capsid/rpoB genes in *P. piscinae* S26 WT and S26 Δ tdaB. Three additional replicates were included at 72h, which were produced under identical circumstances.

To further analyze the data, GER values were $\log_{10}(x)$ -transformed to stabilize the variance. A multivariate linear model (MLM) was created to examine the relationship between the GER and the independent variables time and group (strain). Assessing the model using ANOVA, time and group were estimated to have a significant effect on the GER ($p = 0.0167$, $p = 0.00234$, respectively). The most significant positive effect on GER was found for S26 Δ tdaB at 72 hours ($p = 0.03$). According to the MLM model, S26 Δ tdaB was estimated to express significantly higher capsid gene expression levels at timepoints 48h and 72h ($p_{48h} = 0.0262$, $p_{72h} = 0.0023$) (Figure 10).

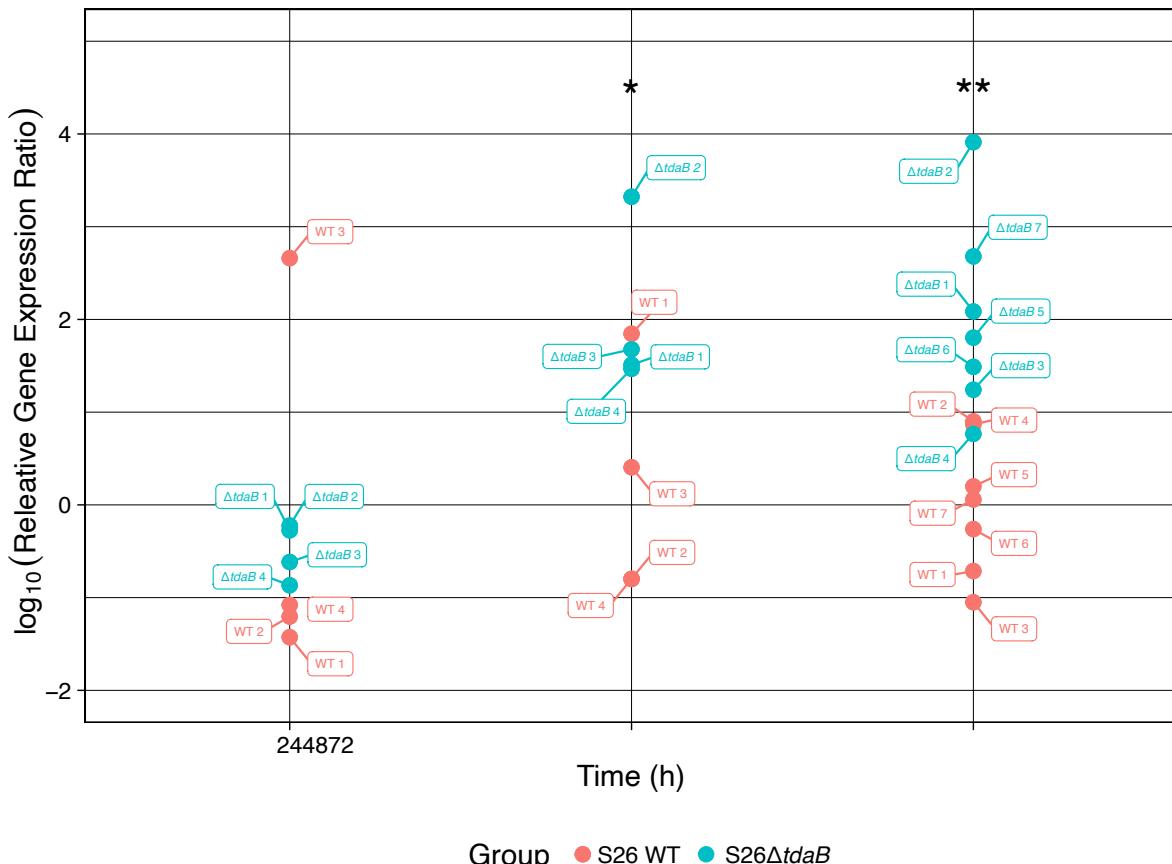


Figure 10: log₁₀(Relative Gene Expression Ratio) values for *P. piscinae* S26 WT and S26ΔtdaB. Three additional replicates were included at 72h, which were produced under identical circumstances. The mutant strain, S26ΔtdaB, showed significantly higher expression of GTA capsid gene at 48h and 72h (LLM pairwise comparison, Bonferroni adjusted. $p_{48h} = 0.0262$, $p_{72h} = 0.0023$) (Appendix fc.1). Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Summary of results

Gel electrophoresis of RNA extractions revealed a distinct band in S26ΔtdaB cultures, closely resembling the size of packaged DNA of an RcGTA. This band, referred to as the GTA band, was consistently more intense in S26ΔtdaB compared to the wildtype strain. Flow cytometry analysis of S26 monocultures failed as a tool to quantify GTA particles, but instead detected a likely prophage. The experiment revealed a stochastic induction pattern of the prophage after 24 hours. Using a multivariate linear model (MLM) it was shown that the prophage was expressed significantly more in S26ΔtdaB compared S26 WT at the 48- and 72-hour timepoints. Quantification of capsid gene expression (OL67_001823) using qPCR revealed a similar stochastic pattern of increased gene expression after

24 hours. Comparative analysis using a MLM demonstrated significantly higher capsid gene expression in S26 Δ tdaB compared to S26 WT at the 48- and 72-hour time points.

Discussion

We demonstrate that GTAs and prophages were generally produced once cultures entered the stationary phase (>24h), and significantly more in S26 Δ tdaB compared to S26 WT at 48 and 72 hours. This finding is supported by previous papers on *P. piscinae* S26 WT and S26 Δ tdaB, as well as studies on RcGTAs (Lindqvist et al., 2023; Solioz et al., 1975; Lang & Beatty, 2000). In FCM analysis of S26 WT and S26 Δ tdaB monocultures, GTA particles were either undetectable or indistinguishable from an uncharacterized prophage (Figure 5). Side scattering of light (SSC) was measured to assess the size and complexity of the particles, while fluorescence intensity was employed to detect the content of double-stranded DNA (dsDNA). Assuming that the size and complexity of S26 GTAs are comparable to that of other known GTAs and considering the 4-5 kb band observed in the gel electrophoresis is representative of the dsDNA content of S26 GTAs (Figure 4), it appears that the size, complexity, and dsDNA content of S26 GTA particles are likely insufficient for detection by FCM. Even if the GTA particles were detected, they would likely be indistinguishable from prophage particles (Brussaard et al., 2010).

RT-qPCR has previously been used to analyze expression of the *gafA* gene in *R. capsulatus* (Sherlock & Fogg, 2022), but to the knowledge of the authors, never for quantifying gene transcripts encoding GTA particle components. The use of RT-qPCR proved to be an effective method for GTA capsid gene transcript quantification and can likely be used as a proxy for GTA production. This is supported by a recent study demonstrating that the same capsid gene, among other GTA-related genes as a prophage, were upregulated in S26 Δ tdaB compared to S26 WT by transcriptome and proteome analysis (Lindqvist et al., 2023). A similar stochastic induction pattern of GTAs and prophages was observed in both FCM and RT-qPCR data. Signals would spike in some biological replicates, where others showed little to no signal. Similar stochastic induction patterns have been observed in *R. capsulatus* RcGTA gene expression (Ding et al., 2019). In *R. capsulatus*, the Rc280 protein acts as an extracellular repressor of RcGTA gene expression. This mechanism is believed to limit the number of RcGTA-producers in a community to a very small subpopulation (< 3%) and is characterized as a stochastic process. By genetic stabilization and autoinduction mechanisms, stochastic effects may result in phenotypic “bifurcations” in microbial communities – even in isogenic cultures under homogenous conditions (Ryall et al., 2012; Veening et al., 2008). Despite similar induction patterns and the upregulation of GTA-related genes and prophage-related genes in S26 Δ tdaB compared to S26 WT (Lindqvist et al., 2023), it is important to note that there is no known association between the GTAs and

the prophage, and they should not be treated as analogous entities. Nevertheless, the induction patterns and possible connection between prophages and GTAs are certainly an interesting observation worth further investigating in the future.

While this study provides insights into the induction patterns of GTAs and prophages in *P. piscinae* S26 WT and S26 Δ tdaB, there are certain limitations and disadvantages to the experiments and data handling that should be acknowledged. Firstly, it would have been advantageous to take samples from the same cultures for RT-qPCR and flow cytometry. This would result in the establishment of a unified model with enhanced statistical power, potentially enabling an exploration of the relationship between GTA and phage induction. Furthermore, given the stochastic induction patterns observed for prophage and GTAs, additional samples would be beneficial to reduce the variance and increase the statistical power even further.

The experiments described in this report do not accurately reflect colony heterogeneity. For *R. capsulatus*, only a small subpopulation (< 3% of cells) will produce RcGTAs in an induced, stationary-phase culture (Ding et al., 2019). Culture samples of *P. piscinae* S26 taken for RT-qPCR and FCM experiments contain both GTA(+) and GTA(-) phenotypes in an unknown ratio, and the single-cell level expression dynamics of the *Phaeobacter* sp. GTA remain unknown. A GTA gene reporter strain of *P. piscinae* S26 could be constructed to further explore GTA production at the single-cell level. A reporter strain could be used for real-time tracking of GTA gene expression, as well as fluorescence-activated cell sorting (FACS) of GTA-producers, both in monocultures and synthetic communities. The overarching goal of this project was to enable further understanding of GTA-mediated HGT in natural microbiomes, yet all experiments described in this report were performed using monocultures of *P. piscinae* S26. This does not accurately reflect the complex interactions found in natural microbial communities, and significantly higher rates of GTA-mediated HGT has been observed in natural environments compared to laboratory conditions (McDaniel et al., 2010).

We observed a fluctuation in color intensity during the incubation period for both S26 WT and S26 Δ tdaB. Surprisingly, some biological replicate cultures displayed a decrease in turbidity and brown TDA pigmentation after 24-48 hours, despite being exposed identical culture conditions. Our initial assumption was that reduced TDA production caused this phenomenon, but we observed a similar pattern in S26 Δ tdaB, indicating that it might be a decline in cell count instead. Although we did not further explore this phenomenon, it would be interesting to examine the factors responsible for this reduction in color and determine if there is correlation with phage/GTA production.

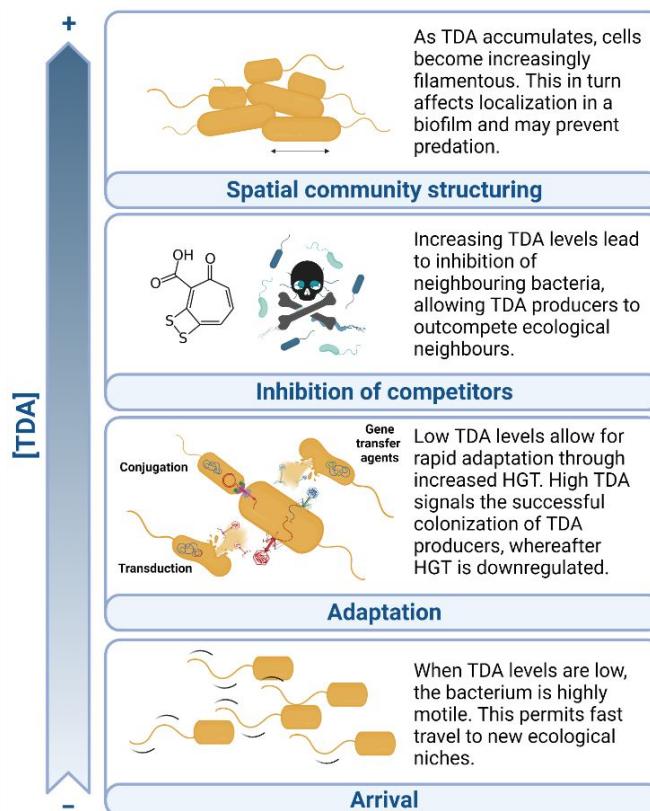


Figure 11: Proposed role of TDA in colonization of new ecological niches, showing 4 levels of increasing TDA concentrations by *Phaeobacter* spp. Figure by Lindqvist, L. L. et al., 2023.

The expression of approximately 10% of *P. inhibens*' total genes are significantly influenced by TDA (Beyermann et al., 2017). In *P. piscinae* S26, the ability to produce TDA affects the motility of the producing organism, with higher levels of TDA leading to decreased motility (Lindqvist et al., 2023). In combination with the findings that abolishment of TDA biosynthesis leads to upregulation of three HGT-related systems (type IV secretion system, prophage induction and GTA expression), this points to TDA playing a role in the colonization of novel niches by the producing organism (Figure 11) (Lindqvist et al., 2023). At low TDA concentrations, higher rates of HGT, including GTA production, would allow for colonizing organisms to rapidly share templates for homologous recombination, allowing for diversification of genetic traits within the community. For the TDA-deficient strain S26 $\Delta tdaB$, the absence of TDA production would prevent successful colonization signaling, and HGT would remain upregulated. This could possibly explain the increase in GTA production observed in S26 $\Delta tdaB$.

The host's ability to produce TDA also seem to affect induction of a prophage, as observed in this report. In a study of *P. inhibens* DSM 17395, the 262-kb chromid encoding the biosynthesis of TDA (among other functions) was deleted. With genomic resequencing, it was found that deletion of the 262-kb chromid resulted in 1.9 to 3.7-fold higher DNA coverage of a *Myoviridae* prophage ($\Delta 262$. Wünsch et al., 2020). It was unclear whether the 262-kb chromid was responsible for repression of the prophage, or the prophage was induced at higher cell densities achieved in the $\Delta 262$ strain. Building upon these results with the findings of this report and Lindqvist's recent work, it appears that the abolishment of TDA biosynthesis leads to an upregulation of not only a GTA, but also prophage induction in *Phaeobacter* spp. The prophage encoded in *P. piscinae* S26 is, however, still uncharacterized and may respond differently to TDA biosynthesis compared to the $\Delta 262$ -encoded *Myoviridae*.

While *Phaeobacter* sp. GTA gene expression and prophage induction are both affected by the ability to produce TDA, it remains unknown whether these mechanisms are directly co-regulated. One proposed link between GTAs and phages is the role of GTAs as a CRISPR-based "vaccination" system (Redfield, 2018). When a phage enters its lytic cycle, the relative content of phage dsDNA in the cell will be higher than in homeostasis. GTAs could serve as a vaccination system, transferring parts of phage dsDNA to recipients and triggering development of CRISPR immunity (Figure 12). Testing this hypothesis would require a controlled system of phage induction, and ideally a purification protocol yielding intact GTAs or GTA dsDNA for further study (Langille & Bottaro et al., 2022).

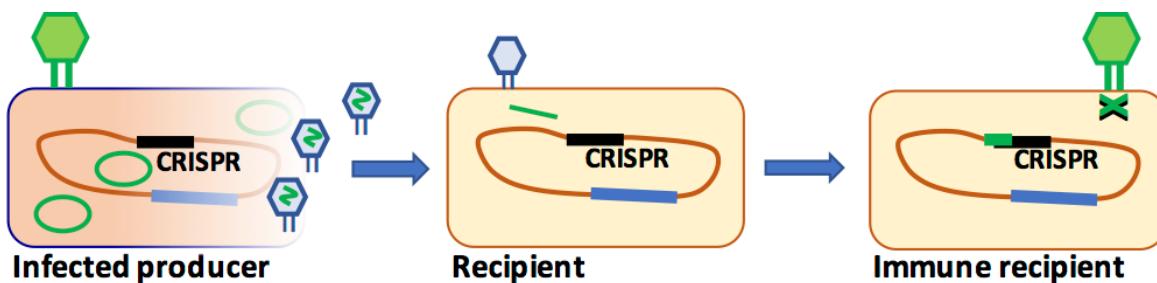


Figure 12: Proposed mechanism for GTA-mediated CRISPR immunity development. Figure by Redfield, R., 2018.

Assuming the validity of the vaccination hypothesis, a phage outbreak would lead to a greater relative abundance of phage dsDNA in the released GTA particles, which could be analyzed. In this project, the hypothesized GTA gel band was not repurified and analyzed, meaning no conclusions can be made about the quality or composition of the fragment. Following GTA dsDNA purification, a restriction enzyme-based assay could be employed to check for patterns in the GTA dsDNA content.

More detailed information about the contents of GTA particles could be obtained by sequencing, which could reveal any packaging bias.

This would require adequate quantities of GTA DNA, which may pose a challenge to obtain. Investigation of GTAs faces a constraint due to the small size of particles and limited quantity in samples. To facilitate detection and analysis of GTAs, it would be beneficial to develop a GTA-overproducing strain of *Phaeobacter* sp. akin to the RcGTA hyperproducer *R. capsulatus* DE442 (Ding et al., 2019). A *Phaeobacter* sp. GTA overproducer could possibly be constructed by upregulating production of the GafA homologue, which is present in *P. piscinae* S26 (Lindqvist et al., 2023).

Due to their small size and the limitations of existing tools, there is presently no established approach for detecting and measuring GTAs. The primary goal of this report was to explore means for reliable detection and quantification of GTAs. We demonstrated that while GTA particles were either undetectable or indistinguishable from bacteriophages by flow cytometry, RT-qPCR provided a reliable quantification of gene expression ratios for GTA genes. Our results indicate that GTAs and prophages are expressed significantly more in S26 Δ tdaB as compared to S26 at 48 and 72 hours of incubation in IOCG medium. Additionally, an interesting resemblance between induction of GTAs and prophages in *P. piscinae* S26 emerged, revealing stochastic production patterns with striking similarities. The understanding of GTAs has advanced significantly in recent years, but much still remains unknown about the role and regulation of GTAs and their function in natural microbiomes. Thus, the field of GTAs has an exciting future ahead, for further exploration and discoveries to come.

Acknowledgements

We would first and foremost like to thank every member of the Bacterial Ecophysiology & Biotechnology group at DTU Bioengineering for their lasting support, guidance, and friendship. The past year has been filled with lots of challenges and unforgettable experiences, which we cherish greatly.

We express our heartfelt gratitude to all our 4 supervisors. We extend a thanks to Shengda Zhang for generously sharing her expertise in molecular biology for the past year, and consistently being a joy to work with. We are grateful to Morten Dencker Schostag for his unwavering patience during the countless hours he dedicated to guiding us through the intricacies of working with genomic material in the laboratory. Tons of appreciation goes to Laura Louise Lindqvist for sharing her remarkable knowledge in the field of gene transfer agents. Laura's own research constitutes the foundation for our project, for which we are incredibly grateful. Lastly, we would like to thank Lone Gram for her invaluable microbiological wisdom and for providing us with the opportunity to conduct our Bachelor thesis in her exceptional research group, where we have spent the majority of our study time over the past year. Being part of this group has been a meaningful part of our journey, and we look forward to the possibility of rejoining for another project down the line!

We would also like to thank Prof. Matthias Middelboe (KU, Section for Marine Biology) for his invaluable assistance with our flow cytometry experiments. Without his expertise and technical know-how, much of this project could not have been realized. Much appreciated was the statistical guidance of Mikael Lenz Strube, helping us with the data analysis and the statistical models. I, Robert, would also like to thank every member of DTU ScienceShow for their companionship. Celebrating and performing with you has kept me sane and joyful.

Lastly, our gratitude goes out to the Danish National Research foundation for funding the Center for Microbial Secondary Metabolites (CeMiSt, DNRF 137), which has provided us with the necessary aid to realize our experiments.

Bibliography

1. Alexander I. Culley et al., 2006. Metagenomic Analysis of Coastal RNA Virus Communities. *Science* 312, 1795–1798. DOI: [10.1126/science.1127404](https://doi.org/10.1126/science.1127404)
2. Bárdy, P., Füzik, T., Hrebík, D., Pantůček, R., Thomas Beatty, J., & Plevka, P. (2020). Structure and mechanism of DNA delivery of a gene transfer agent. *Nature Communications*, 11(1). <https://doi.org/10.1038/s41467-020-16669-9>
3. Bech, P. K., Zhang, S.-D., Henriksen, N. N. S. E., Bentzon-Tilia, M., Strube, M. L., & Gram, L. (2023). The potential to produce tropodithietic acid by *Phaeobacter inhibens* affects the assembly of microbial biofilm communities in natural seawater. *Npj Biofilms and Microbiomes*, 9(1), 12. <https://doi.org/10.1038/s41522-023-00379-3>
4. Beyersmann, P. G., Tomasch, J., Son, K., Stocker, R., Göker, M., Wagner-Döbler, I., Simon, M., & Brinkhoff, T. 2017. Dual function of tropodithietic acid as antibiotic and signaling molecule in global gene regulation of the probiotic bacterium *Phaeobacter inhibens*. *Scientific Reports*, 7(1). <https://doi.org/10.1038/s41598-017-00784-7>
5. Black, L. W. (1989). DNA packaging in dsDNA bacteriophages. *Annual Review of Microbiology*, 43, 267–292. <https://doi.org/10.1146/annurev.mi.43.100189.001411>
6. Bollivar, D. W., Bernardoni, B., Bockman, M. R., Miller, B. M., Russell, D. A., Delesalle, V. A., Krukonis, G. P., Hatfull, G. F., Cross, M. R., Szewczyk, M. M., & Eppurath, A. (2016). Complete genome sequences of five bacteriophages that infect *Rhodobacter capsulatus*. *Genome Announcements*, 4(3). <https://doi.org/10.1128/genomeA.00051-16>
7. Brilli, M., Fondi, M., Fani, R., Mengoni, A., Ferri, L., Bazzicalupo, M., & Biondi, E. G. (2010). The diversity and evolution of cell cycle regulation in alpha-proteobacteria: a comparative genomic analysis. In *BMC Systems Biology* (Vol. 4). <http://www.biomedcentral.com/1752-0509/4/52>
8. Brimacombe, C. A., Ding, H., & Beatty, J. T. (2014). *Rhodobacter capsulatus*DprA is essential for RecA-mediated gene transfer agent (RcGTA) recipient capability regulated by quorum-sensing and the CtrA response regulator. *Molecular Microbiology*, 92(6), 1260–1278. <https://doi.org/10.1111/mmi.12628>
9. Brimacombe, C. A., Ding, H., Johnson, J. A., & Thomas Beatty, J. (2015). Homologues of genetic transformation DNA import genes are required for *Rhodobacter capsulatus* gene transfer agent recipient capability regulated by the response regulator CtrA. *Journal of Bacteriology*, 197(16), 2653–2663. <https://doi.org/10.1128/JB.00332-15>

10. Brimacombe, C. A., Ding, H., Johnson, J. A., & Thomas Beatty, J. (2015). Homologues of genetic transformation DNA import genes are required for Rhodobacter capsulatus gene transfer agent recipient capability regulated by the response regulator CtrA. *Journal of Bacteriology*, 197(16), 2653–2663. <https://doi.org/10.1128/JB.00332-15>
11. Brimacombe, C. A., Stevens, A., Jun, D., Mercer, R., Lang, A. S., & Beatty, J. T. (2013). Quorum-sensing regulation of a capsular polysaccharide receptor for the Rhodobacter capsulatus gene transfer agent (RcGTA). *Molecular Microbiology*, 87(4), 802–817. <https://doi.org/10.1111/mmi.12132>
12. Brussaard, C. P. D., Payet, J. P., Winter, C., & Weinbauer, M. G. (2010). Quantification of aquatic viruses by flow cytometry. In *Manual of Aquatic Viral Ecology* (pp. 102–109). American Society of Limnology and Oceanography. <https://doi.org/10.4319/mave.2010.978-0-9845591-0-7.102>
13. Brussaard, C.P., Baudoux, AC., Rodríguez-Valera, F. 2016. Marine Viruses. In: Stal, L., Cretoiu, M. (eds) The Marine Microbiome. Springer, Cham. https://doi.org/10.1007/978-3-319-33000-6_5
14. Chen, J., Quiles-Puchalt, N., Chiang, N., Bacigalupe, R., Fillol-Salom, A., Su, M., Chee, J., Fitzgerald, J. R., & Penadés, J. R. (2018). Genome hypermobility by lateral transduction Downloaded from. In *Science* (Vol. 362). <http://science.sciencemag.org/>
15. Chen, Y. E., Tsokos, C. G., Biondi, E. G., Perchuk, B. S., & Laub, M. T. (2009). Dynamics of two phosphorelays controlling cell cycle progression in Caulobacter crescentus. *Journal of Bacteriology*, 191(24), 7417–7429. <https://doi.org/10.1128/JB.00992-09>
16. Croucher NJ, Mostowy R, Wymant C, Turner P, Bentley SD, Fraser C. Horizontal DNA Transfer Mechanisms of Bacteria as Weapons of Intragenomic Conflict. *PLoS Biol.* 2016 Mar 2;14(3):e1002394. doi: 10.1371/journal.pbio.1002394. PMID: 26934590; PMCID: PMC4774983.
17. Daniel Wünsch, Annemieke Strijkstra, Lars Wöhlbrand, Heike M. Freese, Sabine Scheve, Christina Hinrichs, Kathleen Trautwein, Michael Maczka, Jörn Petersen, Stefan Schulz, Jörg Overmann, Ralf Rabus; Global Response of *Phaeobacter inhibens* DSM 17395 to Deletion of Its 262-kb Chromid Encoding Antibiotic Synthesis. *Microb Physiol* 1 December 2020; 30 (1-6): 9–24. <https://doi.org/10.1159/000508591>
18. Ding, H., Grüll, M. P., Mulligan, M. E., Lang, A. S., & Beatty, J. T. (2019). *Induction of Rhodobacter capsulatus Gene Transfer Agent Gene Expression Is a Bistable Stochastic Process Repressed by an Extracellular Calcium-Binding RTX Protein Homologue*. <https://doi.org/10.1128/JB>

19. Fogg PCM, Westbye AB, Beatty JT. 2012. One for All or All for One: Heterogeneous Expression and Host Cell Lysis Are Key to Gene Transfer Agent Activity in *Rhodobacter capsulatus*. *PLoS One* 7:e43772.
20. Fogg, P. C. M. (2019). Identification and characterization of a direct activator of a gene transfer agent. *Nature Communications*, 10(1). <https://doi.org/10.1038/s41467-019-08526-1>
21. Fogg, P. C. M., Westbye, A. B., & Beatty, J. T. (2012). One for all or all for one: Heterogeneous expression and host cell lysis are key to gene transfer agent activity in *Rhodobacter capsulatus*. *PLoS ONE*, 7(8). <https://doi.org/10.1371/journal.pone.0043772>
22. Geng, H., & Belas, R. (2010). Expression of tropodithietic acid biosynthesis is controlled by a novel autoinducer. *Journal of Bacteriology*, 192(17), 4377–4387. <https://doi.org/10.1128/JB.00410-10>
23. Gozzi K, Tran NT, Modell JW, Le TBK, Laub MT. Prophage-like gene transfer agents promote Caulobacter crescentus survival and DNA repair during stationary phase. *PLoS Biol*. 2022 Nov 3;20(11):e3001790. doi: 10.1371/journal.pbio.3001790. PMID: 36327213; PMCID: PMC9632790.
24. Grotkjær, T., Bentzon-Tilia, M., D'Alvise, P., Dourala, N., Nielsen, K. F., & Gram, L. (2016). Isolation of TDA-producing *Phaeobacter* strains from sea bass larval rearing units and their probiotic effect against pathogenic *Vibrio* spp. in *Artemia* cultures. *Systematic and Applied Microbiology*, 39(3), 180–188. <https://doi.org/10.1016/j.syapm.2016.01.005>
25. Henriksen, N. N. S. E., Lindqvist, L. L., Wibowo, M., Sonnenschein, E. C., Bentzon-Tilia, M., & Gram, L. (2022). Role is in the eye of the beholder-the multiple functions of the antibacterial compound tropodithietic acid produced by marine Rhodobacteraceae. In *FEMS Microbiology Reviews* (Vol. 46, Issue 3). Oxford University Press. <https://doi.org/10.1093/femsre/fuac007>
26. Huang, S., Zhang, Y., Chen, F., & Jiao, N. (2011). Complete genome sequence of a marine roseophage provides evidence into the evolution of gene transfer agents in alphaproteobacteria. *Virology Journal*, 8, 1–6. <https://doi.org/10.1186/1743-422X-8-124>
27. Hynes, A. P., Mercer, R. G., Watton, D. E., Buckley, C. B., & Lang, A. S. (2012). DNA packaging bias and differential expression of gene transfer agent genes within a population during production and release of the *Rhodobacter capsulatus* gene transfer agent, RcGTA. *Molecular Microbiology*, 85(2), 314–325. <https://doi.org/10.1111/j.1365-2958.2012.08113.x>
28. Hynes, A. P., Shakya, M., Mercer, R. G., Grüll, M. P., Bown, L., Davidson, F., Steffen, E., Matchem, H., Peach, M. E., Berger, T., Grebe, K., Zhaxybayeva, O., & Lang, A. S. (2016). Functional and Evolutionary Characterization of a Gene Transfer Agent's Multilocus "genome." *Molecular Biology and Evolution*, 33(10), 2530–2543. <https://doi.org/10.1093/molbev/msw125>

29. Lang, A. S., & Beatty, J. T. (2000). Genetic analysis of a bacterial genetic exchange element: The gene transfer agent of *Rhodobacter capsulatus*. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2), 859–864.
<https://doi.org/10.1073/pnas.97.2.859>
30. Lang, A. S., & Beatty, J. T. (2007). Importance of widespread gene transfer agent genes in α-proteobacteria. *Trends in Microbiology*, 15(2), 54–62. <https://doi.org/10.1016/j.tim.2006.12.001>
31. Lang, A. S., Westbye, A. B., & Beatty, J. T. (2017). *Annual Review of Virology. The Distribution, Evolution, and Roles of Gene Transfer Agents in Prokaryotic Genetic Exchange*. <https://doi.org/10.1146/annurev-virology>
32. Langille, E., Bottaro, C. S., & Lang, A. S. (2022). Purification of Functional Gene Transfer Agents Using Two-Step Preparative Monolithic Chromatography. *PHAGE: Therapy, Applications, and Research*, 3(4), 194–203. <https://doi.org/10.1089/phage.2022.0035>
33. Laub, M. T., Chen, S. L., Shapiro, L., & Mcadams, H. H. (2002). Genes directly controlled by CtrA, a master regulator of the Caulobacter cell cycle. *Proceedings of the National Academy of Sciences*, 99(7), 4632-4637. <https://doi.org/10.1073/pnas.062065699>
34. Leung, M. M., Brimacombe, C. A., Spiegelman, G. B. & Beatty, J. T. The GtaR protein negatively regulates transcription of the gtaRI operon and modulates gene transfer agent (RcGTA) expression in *Rhodobacter capsulatus*. *Mol. Microbiol.* 83, 759–774 (2012).
35. M. Solioz, H. C. Yen, and B. Marrs, “Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*,” *Journal of Bacteriology*, vol. 123, no. 2, pp. 651–657, Aug. 1975, doi: <https://doi.org/10.1128/jb.123.2.651-657.1975>.
36. Marrs B. 1974. Genetic recombination in *Rhodopseudomonas capsulata*. PNAS, USA 71:971–973.
37. Marrs, B. L. (2002). The early history of the genetics of photosynthetic bacteria: a personal account. In *Photosynthesis Research* (Vol. 73).
38. Mascolo, E., Adhikari, S., Caruso, S. M., deCarvalho, T., Folch Salvador, A., Serra-Sagristà, J., Young, R., Erill, I., & Curtis, P. D. (2022). The transcriptional regulator CtrA controls gene expression in Alphaproteobacteria phages: Evidence for a lytic deferment pathway. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.918015>
39. McDaniel, L. D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K. B., & Paul, J. H. (2010). High frequency of horizontal gene transfer in the oceans. In *Science* (Vol. 330, Issue 6000, p. 50). American Association for the Advancement of Science.
<https://doi.org/10.1126/science.1192243>

40. Mercer, R. G., Callister, S. J., Lipton, M. S., Pasa-Tolic, L., Strnad, H., Paces, V., Beatty, J. T., & Lang, A. S. (2010). Loss of the response regulator CtrA causes pleiotropic effects on gene expression but does not affect growth phase regulation in *Rhodobacter capsulatus*. *Journal of Bacteriology*, 192(11), 2701–2710. <https://doi.org/10.1128/JB.00160-10>
41. Paul, J. H. 1999. Microbial gene transfer: An ecological perspective. *Journal of Molecular Microbiology and Biotechnology*, 1(1), 45–50.
42. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001 May 1;29(9):e45. doi: 10.1093/nar/29.9.e45. PMID: 11328886; PMCID: PMC55695.
43. Redfield, R. J. (2001). Do bacteria have sex? *Nature Reviews Genetics*, 2(8), 634–639. doi:10.1038/35084593
44. Redfield, R. R. & University of British Columbia. (2019, January 11). *Might GTA be a vaccination system for infecting phages?* RRResearch. Retrieved June 5, 2023, from <http://rrresearch.fieldofscience.com/2018/01/might-gta-be-vaccination-system-for.html>
45. Ryall, B., Eydallin, G., & Ferenci, T. (2012). Culture History and Population Heterogeneity as Determinants of Bacterial Adaptation: the Adaptonics of a Single Environmental Transition. *Microbiology and Molecular Biology Reviews*, 76(3), 597–625. <https://doi.org/10.1128/mmbr.05028-11>
46. Sherlock D, Fogg PCM. 2022. The archetypal gene transfer agent RcGTA is regulated via direct interaction with the enigmatic RNA polymerase omega subunit. *Cell Rep* 40:111183.
47. Simon, M., Scheuner, C., Meier-Kolthoff, J. P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., Klenk, H. P., Schomburg, D., Petersen, J., & Göker, M. (2017). Phylogenomics of Rhodobacteraceae reveals evolutionary adaptation to marine and non-marine habitats. *ISME Journal*, 11(6), 1483–1499. <https://doi.org/10.1038/ismej.2016.198>
48. Solioz, M., Yen, H. C., & Marrs, B. (1975). Release and uptake of gene transfer agent by *Rhodopseudomonas capsulata*. *Journal of Bacteriology*, 123(2), 651–657. <https://doi.org/10.1128/jb.123.2.651-657.1975>
49. Sonnenschein, E. C., Phippen, C. B. W., Nielsen, K. F., Mateiu, R. V., Melchiorse, J., Gram, L., Overmann, J., & Freese, H. M. (2017). *Phaeobacter piscinae* sp. nov., a species of the Roseobacter group and potential aquaculture probiont. *International Journal of Systematic and Evolutionary Microbiology*, 67(11), 4559–4564. <https://doi.org/10.1099/ijsem.0.002331>
50. Veening, J. W., Smits, W. K., & Kuipers, O. P. (2008). Bistability, epigenetics, and bet-hedging in bacteria. In *Annual Review of Microbiology* (Vol. 62, pp. 193–210).

<https://doi.org/10.1146/annurev.micro.62.081307.163002>

51. Vos M. Why do bacteria engage in homologous recombination? *Trends Microbiol.* 2009 Jun;17(6):226-32. doi: 10.1016/j.tim.2009.03.001. Epub 2009 May 20. PMID: 19464181.
52. Wagner-Döbler I, Biebl H. 2006. Environmental biology of the marine *Roseobacter* lineage. *Annu Rev Microbiol*; 60: 255–280.
53. Wall, J. D., Weaver, P. F., & Gest, H. (1975). Gene transfer agents, bacteriophages, and bacteriocins of *Rhodopseudomonas capsulata*. *Archives of Microbiology*, 105(1), 217–224. <https://doi.org/10.1007/BF00447140>
54. Wang, H., Ziesche, L., Frank, O., Michael, V., Martin, M., Petersen, J., Schulz, S., Wagner-Döbler, I., & Tomasch, J. (2014). The CtrA phosphorelay integrates differentiation and communication in the marine alphaproteobacterium *Dinoroseobacter shibae*. *BMC Genomics*, 15(1). <https://doi.org/10.1186/1471-2164-15-130>
55. West, S. A., & Cooper, G. A. (2016). Division of labour in microorganisms: An evolutionary perspective. *Nature Reviews Microbiology*, 14(11), 716–723. <https://doi.org/10.1038/nrmicro.2016.111>
56. West, S. A., Griffin, A. S., Gardner, A., & Diggle, S. P. (2006). Social evolution theory for microorganisms. *Nature Reviews Microbiology*, 4(8), 597–607. <https://doi.org/10.1038/nrmicro1461>
57. Westbye, A. B., Kater, L., Wiesmann, C., Ding, H., Yip, C. K., & Beatty, J. T. (2018). The protease ClpXP and the PAS domain protein DivL regulate CtrA and gene transfer agent production in *Rhodobacter capsulatus*. *Applied and Environmental Microbiology*, 84(11). <https://doi.org/10.1128/AEM.00275-18>
58. Westbye, A. B., Kuchinski, K., Yip, C. K., & Beatty, J. T. (2016). The Gene Transfer Agent RcGTA Contains Head Spikes Needed for Binding to the *Rhodobacter capsulatus* Polysaccharide Cell Capsule. *Journal of Molecular Biology*, 428(2), 477–491. <https://doi.org/10.1016/j.jmb.2015.12.010>
59. Westbye, A. B., Leung, M. M., Florizone, S. M., Taylor, T. A., Johnson, J. A., Fogg, P. C., & Beatty, J. T. (2013). Phosphate concentration and the putative sensor kinase protein CckA modulate cell lysis and release of the *Rhodobacter capsulatus* gene transfer agent. *Journal of Bacteriology*, 195(22), 5025–5040. <https://doi.org/10.1128/JB.00669-13>
60. Wietz, M., Gram, L., Jørgensen, B., & Schramm, A. (2010). Latitudinal patterns in the abundance of major marine bacterioplankton groups. *Aquatic Microbial Ecology*, 61(2), 179–189. <https://doi.org/10.3354/ame01443>

61. Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). *Primer-BLAST: A tool to design target-specific primers for polymerase chain reaction.* <http://www.biomedcentral.com/1471-2105/13/134>
62. Zhan, Y., Huang, S., Voget, S., Simon, M., & Chen, F. (2016). A novel roseobacter phage possesses features of podoviruses, siphoviruses, prophages and gene transfer agents. *Scientific Reports*, 6. <https://doi.org/10.1038/srep30372>
63. Zhang, Y., & Jiao, N. (2009). Roseophage RDJLØI, Infecting the aerobic anoxygenic phototrophic bacterium Roseobacter denitrificans OChlI4. *Applied and Environmental Microbiology*, 75(6), 1745–1749. <https://doi.org/10.1128/AEM.02131-08>

Appendix

fc.1 Statistical models of flow cytometry data

Detection of GTAs by Flow Cytometry

Anova (lm(formula = Virus ~ Time * Group * Bacteria))					
	df	Sum Sq	Mean Sq	F-value	p-value
Time	3	22.438	7.479	26.077	1.39E-07
Group	1	3.912	3.912	13.639	1.20E-03
Bacteria	1	1.460	1.460	5.089	3.39E-02
Time:Group	3	3.899	1.300	4.531	1.23E-02
Time:Bacteria	3	8.045	2.682	9.349	3.15E-04
Group:Bacteria	1	3.409	3.409	11.887	2.19E-03
Time:Group:Bacteria	3	2.344	0.781	2.724	6.76E-02 *
Residuals	23	6.597	0.287		

Pairwise S26 WT- S26ΔtdaB					
Timepoint	df	estimate	std.error	statistic	p-value
	24	23	-0.042	1.245	-0.034 0.97307
	48	23	32.092	10.408	3.084 0.00525 **
	72	23	1.184	0.363	3.266 0.00340 **
	96	23	-0.293	0.388	-0.755 0.45816

Summary				
term	estimate	std.error	statistic	p-value
(Intercept)	-1.223	4.398	-0.278	0.78345
Time48	273.848	89.953	3.044	0.00576 **
Time72	9.490	4.596	2.065	0.05040 .
Time96	5.464	4.464	1.224	0.23330
GroupWT	-2.662	39.060	-0.068	0.94625
Bacteria	0.953	1.075	0.887	0.38433
Time48:GroupWT	-240.145	98.145	-2.447	0.02247 *
Time72:GroupWT	5.458	39.153	0.139	0.89035
Time96:GroupWT	9.886	39.123	0.253	0.80276
Time48:Bacteria	-58.578	19.338	-3.029	0.00597 **
Time72:Bacteria	-1.791	1.125	-1.593	0.12483
Time96:Bacteria	-0.853	1.094	-0.780	0.44361
GroupWT:Bacteria	0.657	9.783	0.067	0.94700
Time48:GroupWT:Bacteria	50.556	21.693	2.331	0.02891 *
Time72:GroupWT:Bacteria	-1.625	9.805	-0.166	0.86984
Time96:GroupWT:Bacteria	-2.342	9.796	-0.239	0.81319

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

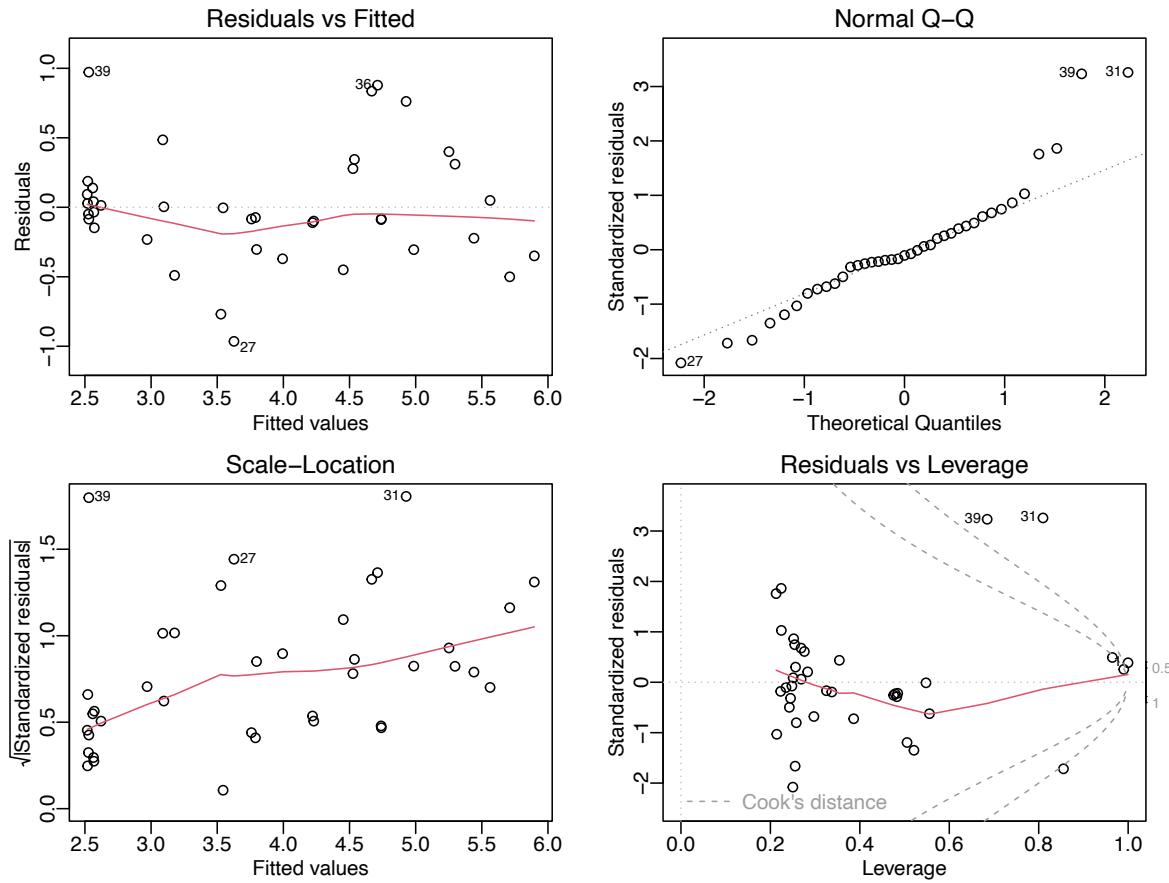
Residual standard error: 0.5356 on 23 degrees of freedom

Multiple R-squared: 0.8734, Adjusted R-squared: 0.7908

F-statistic: 10.58 on 15 and 23 DF, p-value: 5.342e-07

Table 4 Flow Cytometry:

fc.2 Model assumptions for multivariate linear model of flow cytometry data



fc.2 Model assumptions for multivariate linear model of flow cytometry data. 1. Residuals vs fitted, 2. QQ plot, 3. Scale-location, 4 Residuals vs Leverage

qPCR.1 Statistical model of RT-qPCR data

Effect of host TDA production on expression of GTAs by qPCR

Anova					
term	df	sumsq	meansq	statistic	p-value
Time	2	11.637	5.819	4.873	0.01674 *
Group	1	13.829	13.829	11.582	0.00234 **
Time:Group	2	6.938	3.469	2.905	0.07414 .
Residuals	24	28.6543121	1.19392967		

Pairwise S26 WT- S26ΔtdaB					
Timepoint	df	estimate	std.error	statistic	p-value
24	24	0.233	0.773	0.302	0.76550125
48	24	-1.831	0.773	-2.370	0.02616482 *
72	24	-1.997	0.584	-3.418	0.00225354 **

Summary				
term	estimate	std.error	statistic	p-value
(Intercept)	-0.262	0.546	-0.480	0.63558
Time48	0.426	0.773	0.551	0.58677
Time72	0.262	0.685	0.383	0.70519
GroupΔtdaB	-0.233	0.773	-0.302	0.76550
Time48:GroupΔtdaB	2.064	1.093	1.889	0.07100 .
Time72:GroupΔtdaB	2.230	0.969	2.302	0.03032 *

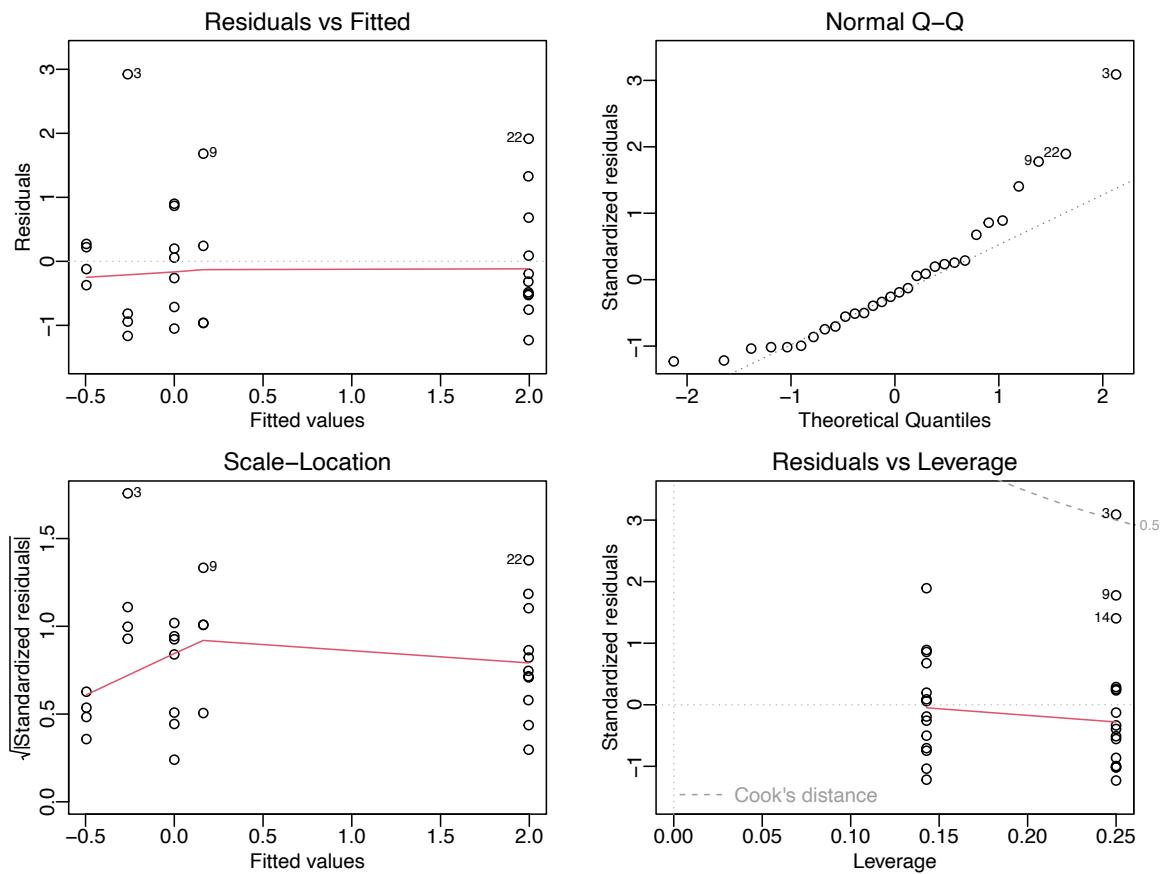
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.093 on 24 degrees of freedom

Multiple R-squared: 0.5307, Adjusted R-squared: 0.4329

F-statistic: 5.428 on 5 and 24 DF, p-value: 0.001768

qPCR.2 Model assumptions for RT-qPCR multivariate linear model



qPCR.2 Model assumptions for multivariate linear model of flow cytometry data. 1. Residuals vs fitted, 2. QQ plot, 3. Scale-location, 4 Residuals vs Leverage