Phage stability in public transport, their diversity and self-assembly in spaceflight, and evaluation of copper surfaces for their antiviral potential

Von der Fakultät für Lebenswissenschaften
der Technischen Universität Braunschweig
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)
genehmigte
D i s s e r t a t i o n

von Bruno Pavletić
aus Zagreb

1. Referentin: Professorin Dr. Melanie Brinkmann
2. Referentin: Privatdozentin Dr. Bianca Schulte
eingereicht am: 03.12.2025
mündliche Prüfung am: 12.03.2026
Druckjahr: 2026

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Publikationen

B. Pavletić, K. Runzheimer, K. Siems, S. Koch, M. Cortesão, A. Ramos-Nascimento, R. Moeller, Spaceflight Virology: What Do We Know about Viral Threats in the Spaceflight Environment? Astrobiology. 22 (2022), doi:10.1089/ast.2021.0009.

F. X. Lehr & B. Pavletić, T. Glatter, T. Heimerl, R. Moeller, H. Niederholtmeyer, Enhanced assembly of bacteriophage T7 produced in cell-free reactions under simulated microgravity. npj Microgravity. 10, 6–12 (2024). doi: 10.1038/s41526-024-00378-4

Tagungsbeiträge

Posterbeiträge

B. Pavletić (2021). Spaceflight Synthetic Biology Pannel discussion. German Conference for Synthetic Biology. Virtual conference.

B. Pavletć, R. Moeller (2022). Stability of Bacteriophage MS2 in Spaceflight Conditions: Stability of Bacteriophage MS2 in Spaceflight Conditions. 2nd German Phage Symposium, Stuttgart, Germany.

F. Webner, and A. Kohl, and B. Pavletić, and F.S. Arndt, and Y.-T. Ly, and I. Talai, and D. Diallo, and H. A. Grübbel, and J. Soltau, (2023) DLR Graduiertenschule GANDALF: Bewusstsein und Bekämpfung pandemischer Gefahren. In: DLR GANDALF Mid-Term Evaluation Workshop. DLR GANDALF Mid-Term Evaluation Workshop, 2023-10-26 - 2023-10-27, Köln, Deutschland.

B. Pavletić (2023) Phages as tools in spaceflight virology research. 8th European Congress of Virology 2023, 2023-05-04 - 2023-05-07, Gdańsk, Poland

Acknowledgements

I was funded by a DLR-DAAD fellowship. I thank Prof. Dr. Ralf Moeller, who unexpectedly passed during this research. Still, he was always there for me, supporting me in any way he can and always focusing on science and discovery. I dedicate this thesis to him. I will never forget his enthusiasm and love for microbiology. Thank you, Ralf. This work would not have been possible without you.

I also sincerely thank Prof. Dr. Melanie Brinkmann and PD Dr. Bianca Schulte for their mentorship, help, and fruitful discussions. Additionally, I express my gratitude to the German Aerospace Center (DLR) for the support in the experiments. Finally, I thank TU Braunschweig for granting me the opportunity to conduct this PhD thesis.

Very often authors give thanks for various kinds of support to their families, partners, friends, colleagues etc. making it seem as if they passed through immense pain and suffering on their PhD journey and needed strong support of various kinds. However, many of us do not truly understand what suffering is until we experience it - breaking us from within and reforming us into new and improved people. When we experience true breakpoints, rather than mild inconveniences in life, PhD, or any other journey we are on, the support of closest to us counts. In those moments, we remove the mask we put out to the world to seem ideal and show our true selves with all our imperfections. At first, I underestimated the amount of pain I experienced on this journey, so I wanted to leave this part out, seeing it unprofessional or superficial. However, looking back, I realize that I am a different person than before, and I owe my sincerest thanks to the people who truly always supported me. Everyone I mention here deserves gratefulness from the bottom of my heart. I would not be where I am without my grandfather, Prof. Dr. Milan Sikirica who pointed me in the direction of science and showed its beauty since I was a small child. I am happy I could follow in your steps in becoming a doctor. Thank you for all your guidance and care. Thank you sincerely to my parents, Josip and Paula Pavletić, for being by my side in all the moments that I needed emotional, logistical or financial support – without you I could not have made it. Thank you dearly to my grandparents Đurđa Sikirica, Desanka and Grga Pavletić – you were giving me company and guided me with your wisdom through the journey and making me pursue biology from the early days. This thesis would not be here without you. Thank you deeply to my dear friend, Martin Matešić who is the greatest team member I know, and a creative person beyond description. I appreciate all the sacrifices you did for our team – I am sure we will create a great startup and make the world a better place, despite any setbacks. Finally, I would like to sincerely and from the heart thank my girlfriend, Aurora Britania Diaz Fernandez MD, the one who means to me more than anything in the world. She is the light that guides me through this path of life, day and night. I

cannot thank her enough for sticking with me while passing through so much hard times herself. Thank you, my love, for being there always even though I am aware that times have been hard. I know I can trust in you with my life.

Summary

Viruses are the most diverse and numerous biological entities on Earth. They are a crucial part of any ecosystem and pose a range of health-related risks but also biotechnological opportunities. This is true for bacterial viruses (bacteriophages or phages), archaeal viruses, and eukaryotic viruses. Due to their appreciable diversity, it is complex to describe how their characteristics like infectivity or diversity change in a given environment.

The research presented in this thesis was divided into four main chapters. In Chapter 1, different bacteriophages frequently used as surrogates for human/animal viruses were tested for their stability on glass surfaces after drying. The phages examined here were Enterobacteria Phage MS2, Enterobacteria Phage M13, Enterobacteria Phage PhiX174, Enterobacteria Phage T4, and Pseudomonas Phage Phi6. All the phages except PhiX174 were mixed with the host bacteria while their optical density at 400 nm (OD_{400}) was measured over time. This was performed to try to corelate the Plaque-forming unit assays (PFU assays) to the effect of the host growth curve. However the growth curves showed significant variation, so PFU assays were chosen to determine the effect of drying. PFU assays were therefore performed to determine the infectivity of the bacteriophages after drying on glass for up to 1, 3 and 7 days at room temperature and fridge (approximately 4 °C). Additionally, the molecular detection through loop-mediated isothermal amplification (LAMP) was also performed on the dried phages. This experiment served to introduce the virus cultivation and molecular detection and demonstrate the infectivity and detectivity of viral surrogates after drying on glass surfaces. The results show significant variation in stability between different surrogates. Specifically, PFU assays showed that M13 and T4 both showed high stability at room temperature and low temperature (approximately 4 °C) with minimal reductions after 7 days. On the other hand, PhiX174 showed high stability in the fridge, but not at the room temperature – with infectivity near 100% after 7 days in the fridge but near no detection at RT. At the same time, MS2 and Phi6 showed low stability both at RT and in the fridge with infectivity approaching the limit of detection after 7 days. In addition, the phages were detected with LAMP after 1, 7 and 31 days. Despite some lacking infectivity, all of them were detected using LAMP after up to 7 days, and PhiX174 and T4 were only ones not detectable after 31 days. The experiments demonstrate the variation on infectivity of virus surrogates of different structures though molecular detection can be performed even long after some viruses are rendered non-infectious. The experiments also introduced and tested safe surrogate viruses for space and public transportation research.

Chapter 2 explores the potential use of copper-based antiviral surfaces against human pathogenic viruses, Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-Cov-2) and Monkeypox Virus (MPXV) as well as one phage surrogate, Phi6. Three different variants of

SARS-CoV-2 (B.3, B.1.617.2, and XBB.1.5) were tested and all four MPXV clades (Ia, Ib, IIa, and IIb). For SARS-CoV-2, variant B.3 showed the lowest stability on copper, while XBB.1.5 showed the highest. B.3 fully inactivated after 10 minutes on copper surface and XBB.1.5 showed minimal reduction even after 15 minutes on copper. B.617.2 showed intermediate stability with almost full titer reduction after 10 minutes on copper. The Spike receptor trimer structure and electric potential distribution was determined and compared among the tested variants. It was shown that all variants possess differences, and XBB.1.5 showed the highest overall electric potential with the lowest amount of acidic amino acids. This could result in the least copper cation interaction of all other Spike variants. No genomic damage was observed through qPCR. For MPXV, it was shown that the clade Ia had the highest stability on copper, IIb and IIa intermediate stability, and Ib had the lowest stability. For each clade, the genome was assembled from Illumina sequencing data and genetic variants were determined. The variations in the surface glycosylated proteins were found for each clade. All the clades except Ia had substitution mutations in the heavily glycosylated surface protein OPG210. All the mutations could affect the protein structure or glycosylation effect. Additionally, novel 3D-printed antimicrobial surfaces containing copper microparticles and nanoparticles, respectively, were tested against SARS-CoV-2 and Phi6. The nanoparticle-containing 3D printed materials had a better antiviral effect against both SARS-CoV-2 and Phi6. Phi6 was shown to have low stability on copper compared to animal viruses, making it unsuitable surrogate when copper surface stability is considered.

Chapter 3, in the context of environmental virology, explores the diversity of integrated prophages that are found within bacterial genomes of the International Space Station (ISS) microbiota. All known full genomic sequences from the isolates from the ISS were downloaded from NCBI. Their sequences were probed for prophage gene clusters. The found genomes were then analyzed taxonomically and functionally. This revealed a significant phage diversity of families *Siphoviridae* and *Myoviridae*. Their phylogenies were determined, together with their overall diversity and the diversity within host genera. It was shown that bacterial genus *Stapylococcus* contains by far the highest diversity of prophages inside their genomes. However, this could also be because they represented the highest number of genomes of all ISS isolates. Functional diversity of phage genomes was also determined and plotted. Among all the genomes, the antimicrobial resistance virulence resistance genes were identified. All those genes are available for potential spread within the bacterial population via prophage infection. Therefore, this study showed the potential genes that could be transmitted in this environment. The project was done on genomic data, but the same approach can be done with metagenomic sequences from sequencing samples of ISS or public transportation. Therefore, a similar approach could be used in the future for public transportation, which is especially interesting because of the spread of antimicrobial resistance and virulence genes among the

population. Also, the approach could be applied in the future for space station analysis of prophages.

Finally, in Chapter 4, bacteriophage T7 virion assembly was explored in simulated microgravity through clinorotation in a transcription/translation solution. The virions were synthesized in a solution of phage genomic DNA and bacterial extract while on a rotation system often used to simulate microgravity on Earth, called clinorotation. Sampling the solution over time and performing PFU assays revealed that the endpoint-titer is higher and reached faster in the clinorotated solutions than the stationary controls. The effect on clinorotation on gene expression was explored with a GFP gene-containing control plasmid and measurement of fluorescence produced in clinorotation versus stationary control. In addition, the protein production was assessed with liquid chromatography mass spectrometry and dot blot. The replication of the viral DNA in the solution was also assessed through qPCR quantification. The protein production, gene expression and replication were not affected by clinorotation. Higher number of assembled virions was observed with transmission electron microscopy. Taken together, the results suggest a more efficient assembly of T7 virions in simulated microgravity. This is an important point for future of human space exploration.

Taken together, this work explores the viral resistance to stressors on Earth and in space. This is done by means of four different parts, divided into chapters above. Desiccation stability is an important aspect of virus transmission on Earth and was therefore explored through its effect on different bacteriophage surrogate viruses. By comparing the virus surrogate desiccation stabilities among each other, a line can be drawn between virus type and their stability. The surrogate that is most structurally similar to human respiratory viruses, Phi6, was then compared to human virus stability, specifically, SARS-CoV-2 and MPOXV. Then, the research moves to space environment where the diversity of prophages on the International Space Station is characterized. The pipeline is established to determine the gene and species diversity aboard. Finally, the specific condition of spaceflight, microgravity is explored through its effect on self-assembly of bacteriophage T7. Overall, the stress environments on Earth and in space are evaluated here through their effects on viral stability, assembly and detection.

Table of contents

I. List of abbreviations

AMR	Antimicrobial resistance
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BCA	Bicinchoninic acid
BLAST	Basic local alignment search tool
BSL	Biosafety level
cDNA	Complementary DNA
COVID-19	Coronavirus disease of 2019
Ct	Cycle threshold
CTP	Cytosine triphosphate
DMEM	Dulbecco's Modified Eagle Medium
DPBS	Dulbecco's Phosphate Buffered Saline
dsDNA	Double stranded DNA
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
dsRNA	Double stranded RNA
ELISA	Enzyme-linked immunosorbent assay
GFF8 & 14	Germ-free flying 8 and 14, antimicrobial surfaces used in the study
GFP	Green fluorescence protein
GTP	Guanosine triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
ICTV	International Committee on Taxonomy of Viruses
IgG	Immunoglobulin G
ISS	International space station
K-W test	Kruskal-Wallis Test
LAMP	Loop-mediated isothermal amplification
LB medium	Luria-Bertani medium, a nutritious medium
LC/MS^2	Liquid chromatography/mass spectrometry/mass spectrometry
LEO	Low-Earth orbit
MDA	Multiple displacement amplification (MDA)
MEM	Modified Eagle Medium
MOI	Multiplicity of infection
MPXV	Monkeypox Virus
OD	Optical density
PDB	Protein Data Bank

PFU	Plaque-forming unit
PLA	Polylactic acid
PTFE	Polytetrafluoroethylene
PVC	Polyvinyl chloride
qPCR	Quantitative PCR
RF	Replicative form of viral genome
ROS	Reactive oxygen species
RPM	Rounds per minute
RT	Room Temperature
RT-qPCR	Reverse transcription quantitative PCR
SARS-CoV-2	Severe Acute Respiratory Syndrome Coronavirus 2
SD	Standard deviation
SEM	Scanning electron microscopy
sfGFP	Superfolder Green Fluorescence Protein
SLS	Sodium lauryl sulfate
SNP	Single nucleotide polymorphism
ssDNA	Single-stranded DNA
ssRNA	Single Stranded RNA
TBST	Tris-buffer saline polysorbate
TEM	Transmission electron microscopy
TFA	Trifluoroacetic acid
tmRNA	Transfer messenger RNA
tRNA	Transfer RNA
TSA	Tryptic soy agar
TX/TL	Transcription/translation
UTP	Uracyl triphosphate
UV	Ultra violet
VIR genes	Virulence genes
xg	Unit of centrifugal force corresponding to 9.81 m/s^2

II. List of Tables

III. List of Figures

Introduction

Viral diseases are a big concern for public health. Considering their significant diversity, it is essential to develop ways to detect and reduce viruses in closed environments such as public transportation. The new and emerging viral diseases are unpredictable, and it is assumed their prevalence will grow the more humans come into contact with natural habitats where they are circulating (Epstein *et al.*, 2020; Lee-Cruz *et al.*, 2021; Zhou and Shi, 2021; Kummer and Kranz, 2022; Lawrence *et al.*, 2023). Those biological threats are the problem, both in public transportation and spaceflight. Public transportation is a space where significant amounts of infectious diseases spread, particularly airborne (Ellingjord-Dale *et al.*, 2022; Jaydarifard, Morawska and Paz, 2024; Tapiador, Gomez and Vassallo, 2024). The conditions on trains, buses, and planes allow for a high number of passengers in relatively close proximity. Therefore, public transportation can become a hub for virus transmission (Cai *et al.*, 2019; Hu *et al.*, 2021; Pung *et al.*, 2022; Mazzoli *et al.*, 2023).

Viruses can spread in different ways. Some of the ways include air transmission, fomite transmission, direct contact transmission, vector transmission via e.g. mosquitos, fecal-oral transfer, and many others (Louten, 2016; Leung, 2021). Depending on the context of public transportation or spaceflight, many of the above-mentioned ways of transmission are relevant.

Public transportation and space stations are human-built environments. Therefore, the control of environmental conditions is better than the natural environments, but control is still limited. For example, the built environments such as space stations can precisely control temperature, humidity, and similar. On the other hand, public transportation vehicles have harder to control environments due to them being open to the outer surroundings. The level of control of the environment differs in different vehicles, depending on the type and application. For example, trams, buses, and trains are widely used transportation systems, used daily and exchanging thousands of passengers each day. In aviation, the level of sanitation control is higher, but still, high passenger numbers make control of biological contamination harder. Also, while trams, buses, and trains are constantly exposed to outside conditions, planes have only limited exposure. Therefore, planes have higher environmental atmospheric control than other forms of public transportation. Even planes, however, use air from the outside environment, making necessary adaptations to the internal atmosphere to make it more sustainable for the passengers. Compared to this, space stations have their internal environment fully isolated from the Earth's conditions and have a much higher level of environmental control. Still, in all types of public transportation, as well as space stations, we can find full microbiomes, mostly originating from humans. Those microorganisms contaminate the environment and sometimes they can even be dangerous for passengers and crew members.

Conditions and viral spread in public transportation

From the point of view of the average population, public transportation is significantly more relevant than spaceflight. Trains, airplanes, buses, and other vehicles are daily transporting large numbers of people, and many depend on it. Therefore, during outbreak situations, a large proportion of the population will naturally be concerned about their health safety while using public transportation.

The detection and suppression of viral transmission in those environments is relevant from the public health perspective and enables public transportation companies to continue operating during pandemic events. Innovations in that field are therefore valuable. For example, during the COVID-19 pandemic, the aviation industry suffered significant losses. It has since recovered, even though it took three years to reach the recovery point. Figure 1 represents the average monthly number of commercial flights in the period 2019-2023 until the air flight industry showed signs of recovery. Other public transportation industries also suffered as the pandemic increased car usage, walking, and cycling, as 2021 survey-based research from Germany shows (Anke *et al.*, 2021). It was evident in the times of the lockdowns that the innovations to include in public transportation are beneficial for reducing the burden on the industries and limiting the impact on society and public health (Boon, 2020; Amankwah-Amoah, 2021). This includes, for example touchless technology, apps for contact tracking, and antimicrobial coatings introduced in public transportation systems.

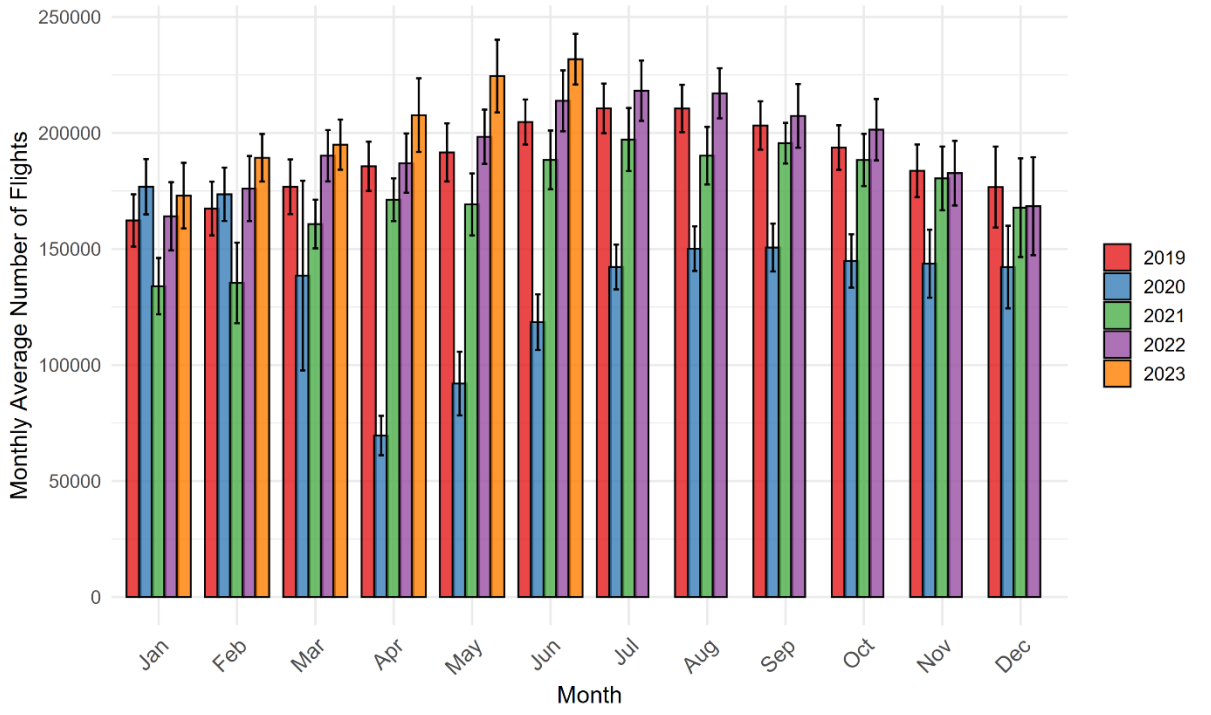


Figure 1. The average monthly numbers of flights throughout 2019-2023. The COVID-19 pandemic had a major negative impact on the air flight industry. The number of flights returned to the 2019 level only in 2023. The industry visibly recovered after that. Data source: flightradar24 (https://www.flightradar24.com/).

Ground-based public transportation such as trains and buses is susceptible to contamination brought in by passengers. The surfaces that come into direct contact with human skin are seat handles and headrests. The air is ventilated through the windows, the ventilation system, and after opening the doors at every station.

Airplane environments have specific sets of conditions. Many people exchanging microbes inside aircraft can result in a complex microbial community structure. In contrast, air turnover and air filtration reduce the chances of transmission. The low humidity, low oxygen availability, airflow, and colder temperature onboard aircraft affect human physiology. In the low humidity of the aircraft in-flight, our mucosal layer in the respiratory system gets thinner, making us more susceptible to air-borne pathogens (Guarnieri *et al.*, 2023; Seok, Lee and Yoon, 2024) . Additionally, the lower oxygen availability causes the state of hypoxia, negatively impacting the immune system activity. Commercial airplanes like public transport vehicles usually contain many people in a closed environment for a limited time. Cabins of commercial airplanes can contain up to 850 seats (two-story Airbus A380 cabin), however, it is worth noting that most flights contain far fewer travelers. In 2019, there were 4.5 billion passengers worldwide on a total of 38.9 million flights (115 passengers per flight on average) according to the International Civil Aviation Organization (ICAO). In 2020 however, 1.8 billion passengers traveled aboard

16.4 million flights (109 passengers per flight on average) (ICAO, 2023). Although the number of passengers has decreased, the number of people sharing the same cabin has remained high. Therefore, biological safety for aircraft passengers in a cabin should remain high.

Viruses possess high structural diversity, with different shapes, sizes, glycosylation profiles, and genome types. All those factors can affect viral stability in any environment. Surrogate viruses, i.e. non-infectious viruses similar to the infectious ones, are valuable tools to determine the connection between viral structure and stability in the environment.

According to International Air Transport Association (IATA), four concerning communicable diseases (the diseases spread from person to person) are COVID-19, Mpox, Acute Hepatitis, and Ebola virus disease, but also flu (IATA, 2023). They are all caused by viruses, most of which could be airborne (all except Mpox, which spreads through direct contact), but other pathogens also could pose a threat. The infected person exhales, sneezes, coughs, or vomits the droplets and aerosolized particles of saliva (99% water mixed with salts and other organic compounds) and mucous (Bagheri *et al.*, 2023). There is no clear physical distinction between “aerosol particle” and droplet, however, often 5-10 µm diameter serves as a threshold in the field of aerodynamics and often a distinction is made, based on the sedimentation time of the particle (e.g. the time the particle needs to fall down from 1 m height to the ground in still air). In literature distinctions are also based on whether the particles (>5µm) are aqueous liquid (droplets) or completely solid and dried (particles) (Heikkilä *et al.*, 2024; Vörösmarty, Hopke and Salma, 2024). In the following text we will refer only to the term “droplet” before evaporation onset or >1000 µm and “aerosol-particle” or “particle” for all other particles, since the evaporation of a 10 µm droplet under normal environmental conditions happens in fractions of a second (Shang *et al.*, 2022). Aerosol particles floating in the air for long time periods (minutes or even hours) and can be inhaled by a healthy individual. The pathogens carried inside can then cause disease. Those air-suspended particles that carry pathogens or other microbes in the air are called bioaerosols. Aerosol particles less than 5 µm in diameter can be barely seen by the human eye and pose a significant medium for the spread of diseases such as influenza, measles, tuberculosis, SARS and COVID-19 (Sze To *et al.*, 2009; Jayaweera *et al.*, 2020; Tang *et al.*, 2020; Azimi *et al.*, 2021) and are the center of respiratory disease research since many years.

Due to air conditioning and highflying altitude, the air pressure and humidity in an aircraft cabin are lower than on the ground (Grün, Trimmel and Holm, 2012). Normally, in the aircraft cabin, the relative humidity (RH, % water in the air out of the maximum amount that could be in the air at a given temperature) is 6-10% (De Ree *et al.*, 2000).This is an important environmental factor in terms of the airborne microbiome and cabin epidemiology. For instance, it is well established that humidity has a significant effect on virus infectivity inside aerosols (Wan and

C., 2012; Lin and Marr, 2020). It is generally believed that the relative humidity is a better predictor of virus stability rather than absolute humidity (AH, the total mass of water in the air per unit volume). Pathogenic respiratory viruses, which mostly have a lipid envelope, generally tend to be more stable at lower RH. A recent study by Nieto-Caballero et al. found, that the half-life of aerosolized coronavirus (murine hepatitis virus) doubled to 120 minutes, when the RH was lowered from 45-60% to 25% (Nieto-Caballero *et al.*, 2023). Also, at lower RH, the mucosal layer in the human respiratory tract becomes drier and thus more permeable. The mucosal layer provides a barrier between the airborne microbes and the human organism. Thus, at lower RH, humans have less protection from airborne pathogens (Sunwoo *et al.*, 2006). Given that the air RH is low inside an aircraft cabin in-flight, this poses an increased risk of infection. This might become more concerning the longer the flight is because an organism is exposed to such conditions for a longer time. According to the models, the infection risks differ for different airborne viruses (Aganovic *et al.*, 2022). This points out the need for research of the effect of viral physical characteristics like genome, size, structure or shape on their stability in an environment.

The pressure inside a cabin at cruising altitude is also lower than on the ground. An average commercial flight reaches 10-12 km altitude (Ahmedzai *et al.*, 2011; Campbell *et al.*, 2018). At this altitude, the pressure in the cabin can drop down to ~75% of the sea-level pressure, equivalent to the pressure at 2 km altitude. The main way that this low pressure affects microbes and humans is through the lower availability of oxygen. The oxygen availability inside the cabin in-flight is equivalent as it consisted of ~15% on the ground. This drop in oxygen availability affects the metabolism of humans and microorganisms.

The state of an organism exposed to lowered oxygen is called hypoxia. The human organism responds to hypoxia in many ways, but the longer it lasts, the more evident the effects are. There is some evidence that air flight-induced hypoxia might alter the immune response after only 30 min of exposure (Hinkelbein *et al.*, 2017). However, this association is currently not clear and the studies of the immune response in air flight conditions are lacking. Therefore, humans might be more prone to infection in-flight, though this is still controversial. From the microbiology point of view, lack of oxygen is also likely the main factor affecting the communities present in the environment. Oxygen availability is an important environmental factor because it is highly toxic to some microbes (anaerobic microbes), but essential to others (aerobic microbes). For anaerobic microbes, oxygen causes oxidation of biomolecules, which damages them and impairs their function, killing them in the process. For aerobic microbes, however, oxygen is an important molecule for their metabolism to keep them alive. It is expected that the lower oxygen content in an airplane significantly affects microbiome composition. This has yet to be confirmed or disproven.

Oxygen causes oxidation of viral membrane and proteins. This would mean that viruses can hypothetically be more stable in the low-oxygen environment. In reality, the relationship between oxygen and viruses is complex. Viruses that contain a lipid envelope are rapidly inactivated by oxidation (Käsermann and Kempf, 1998; Zeng *et al.*, 2020). Those studies confirmed that the viral lipid envelope prevents virus entry after being oxidized. This would mean the enveloped viruses, such as influenza and coronaviruses, benefit from a lower-oxygen environment in an airplane cabin. Also, while some viruses reproduce in oxygen-rich tissues, some exploit the hypoxic environment of the infected cells for reproduction. Therefore, there is a need to classify viruses according to their oxygen requirement after infection (Morinet *et al.*, 2015), the same way other microbes are characterized. This would aid in risk assessment of virus infectivity in the aviation environment because varying oxygen levels could elevate or reduce the risk of infection depending on the viruses present in the environment. Also, the infectivity of a virus and the severity of a viral disease depends on its interaction with the host immune system under hypoxia of an airplane cabin.

A portion of microbial communities from the air settle on the aircraft surfaces instead of filtering. Some of those microbes can be transmitted by direct contact with the contaminated surfaces. This is mostly true for bacteria and fungi since viruses are usually present in too low concentrations on surfaces to cause an infection (Ren *et al.*, 2020). This is due to their inability to reproduce without a host. Even though the chances are low, the possibility still exists, although for bacteria and fungi it is much higher.

With the advances of sequencing and molecular detection technologies, tracking diseases is becoming more approachable. Still, laboratory cultivation remains the only reliable method to determine the infectivity of a virus from an environmental sample. Figure 2 demonstrates the available nucleic acids-based methods for virus detection in airplanes and airports. The infectivity of infectious viral particles could be elucidated through research on surrogate viruses such as bacteriophages and other viruses that do not infect humans. This helps to introduce and standardize disinfection methods, making air travel medically safe. An understanding of the spread of exhaled aerosol particles inside the aircraft cabin will help to further understand the pathway of disease transmission aboard and improve the development of airborne pathogen reducing ventilation concepts, protecting future passengers even better.

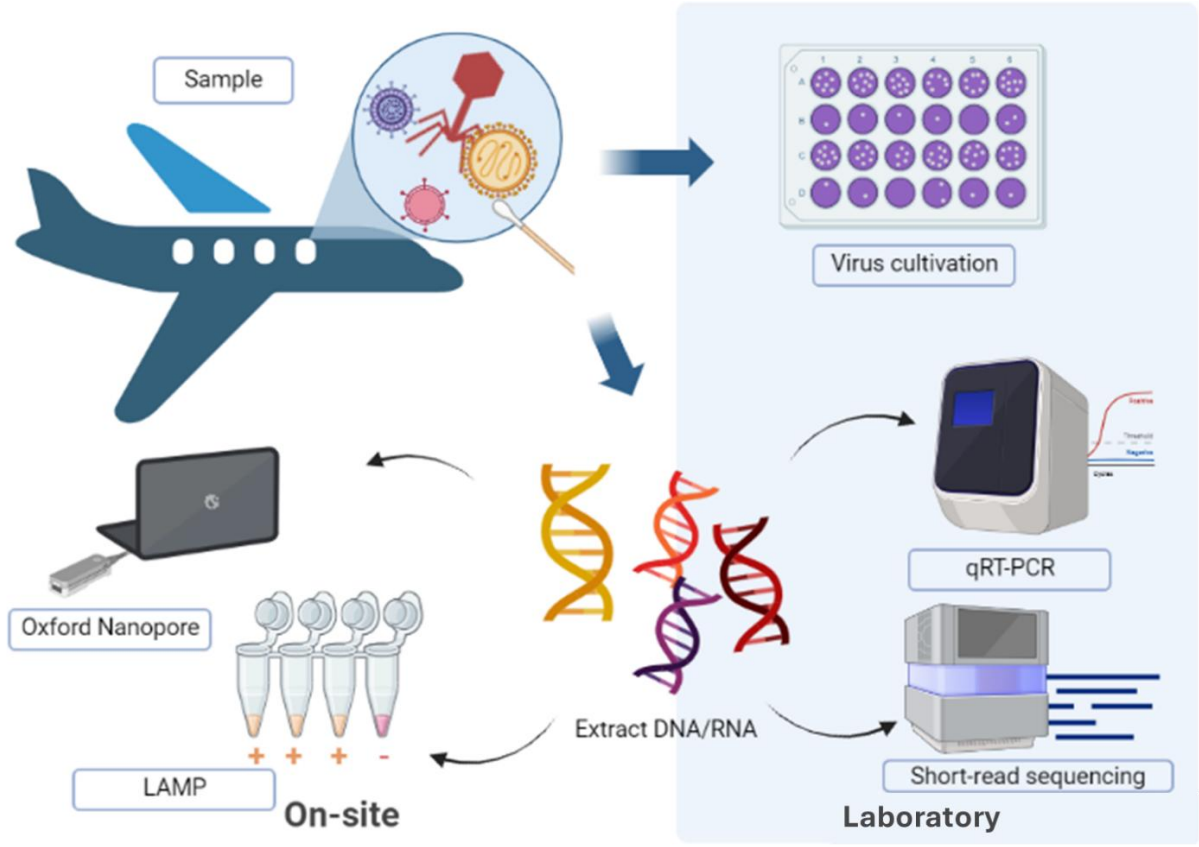


Figure 2. Nucleic acid-based methods available for virus detection in airplanes and airports. Additionally, viral cultivation is shown as the most reliable method to determine infectivity of the viruses found. Created with the help of BiorRender (BioRender.com)

Viruses are more challenging to study by metagenomics than bacteria and fungi due to the various complications that they bring into the analysis. First, they don't possess marker regions, making classification complex. The only way to assign viral sequences to their respective viruses is non-targeted sequencing of all nucleic acids in a sample, followed by the assembly of the parts or whole viral genomes, which are then queried against a database containing a collection of known virus sequences. An additional challenge is that many viral groups are underrepresented in databases. A systematic quantitative analysis of the National Center for Biotechnology Information Viral Genome Database from 2018 determined that the highest rise in the number of new sequences comes from already well-represented groups, while the under-represented groups (bacterial RNA, and archaeal viruses) remained underrepresented (Mahmoudabadi and Phillips, 2018). However, it must be noted that at the time of the analysis (August 2015), the NCBI database contained 4,378 complete viral genomes while today it is more than four times that size with 18,706 complete viral genomes (as of the end of June 2025). Therefore, the situation today might be different. Due to the under-representation of certain viral groups, and lack of the marker region, some viruses can be missed in the metagenomic analysis. However, since human pathogenic viruses are well-

represented in the NCBI database, it is unlikely that such viruses could be missed. An additional complication of virus metagenomics is that they have diverse genome types. While all cellular organisms have double-stranded DNA (dsDNA) genomes, viruses can have dsDNA, dsRNA, single-stranded DNA (ssDNA), or ssRNA genomes. This means that the metagenomic methods that imply DNA isolation are biased to detect only the viruses with DNA genomes, completely missing on the RNA viruses present in a sample. To detect those viruses together with the DNA viruses, additional steps are taken which are described by Greninger (Greninger, 2018). Normally, a reverse-transcription step first needs to be included before the sequencing step, which transcribes the RNA into a single-stranded complementary DNA (cDNA). After reverse-transcription, RNA needs to be degraded by the use of RNase enzyme which selectively breaks down RNA, leaving DNA in the reaction intact. The remaining single-stranded cDNA is then non-selectively amplified by adding adapter sequences at the ends of all DNA molecules present. Those adapters are later also used for the preparation of sequencing libraries. The advances in RNA sequencing have enabled direct sequencing of RNA instead of first transcribing it into cDNA, allowing for faster and more reliable results.

Additionally, the development of long-read sequencing technologies like Oxford Nanopore and PacBio enable sequencing of full viral genomes in metagenomic samples, allowing for better variant detection and tracking. The long-read sequencing technologies can also be adapted both for DNA and RNA sequencing. Oxford Nanopore is adaptable for in-situ sample preparation and sequencing (Zorz *et al.*, 2023), allowing for high precision tracking of outbreaks (Ji *et al.*, 2024).

Additional limitation in viral metagenomic detection is the low abundance of viral sequences in the environment. This problem can be solved by specific and non-specific DNA amplification. If the viral genomes are to be sequenced, a non-specific amplification could be used pre-sequencing to gain a qualitative and nonbiased view of the viral communities. A good method to non-specifically amplify the DNA in a sample is multiple displacement amplification (MDA). It randomly amplifies all the DNA in a sample, and has a low error-rate, allowing for higher DNA yields for sequencing. However, if the target virus is known, quantitative reverse transcription polymerase chain reaction (qRT-PCR), and loop-mediated isothermal amplification (LAMP) can be applied to amplify and detect viral DNA or RNA (after reverse transcription). While qRT-PCR requires expensive equipment, LAMP can be applied in situ and requires minimal technical expertise and on-site, making it suitable for airports.

As the result of all the challenges mentioned above, the knowledge about viromes of aircraft cabins is limited. However, some studies have explored this question. In one of the studies, the dust particles from the aircraft have been collected and analyzed metagenomically. The overall results showed a total of 7,437 identified species, out of which 389 were viral (5.23%

of species) (Sun *et al.*, 2020). The relative abundance of viral sequences was 0.3%. Most of them were bacterial viruses (bacteriophages), mainly Lactococcus, Staphylococcus and Propionibacterium phages. However, among the detected viral species, there were also pathogenic ones, mainly Molluscum contagiosum virus (relative abundance 0.025%), Human herpesvirus 1 (relative abundance 0.002%), and Polyomavirus (relative abundance 0.007%). the results of Sun et al. show that among DNA viruses, bacteriophages constitute the most significant portion, similarly to most other environments on Earth (Mann, 2005). However, it must be noted that only DNA sequences were extracted and sequenced. Therefore, important respiratory RNA viruses such as Influenza viruses and Coronaviruses could not be detected even if they were present. The epidemiological data suggests that those, and other pathogenic viruses might be transmitted aboard aircraft (Leitmeyer and Adlhoch, 2016; Rosca *et al.*, 2021).

Another study focused on the analysis of airplane wastewater as a worldwide surveillance of pathogens and antimicrobial resistance (Petersen *et al.*, 2015). In this study, two genogroups of Norovirus (GI and GII) were detected by qRT-PCR in waste of 18 commercial long-distance travel airplanes. This virus is highly transmissible via human waste and excretions, posing a significant concern in public health. Norovirus GI and GII genogroups were detected in 10, and 14 out of 18 total airplanes respectively. The qRT-PCR shows promise in tracking viral contamination in aircraft. This is particularly useful in detecting RNA viruses which can be present in low amounts in the environment due to the low stability of RNA.

The third study was also based on the analysis of aircraft toilet waste on 19 commercial long-distance flights, and, it was solely focused on virus detection by metagenomic sequencing (Hjelmsø *et al.*, 2019). Many pathogenic viruses were detected, with diverging geographical trends with respect to viral communities. In total, the authors detected 104 viral species distributed into 31 families. Human pathogens constituted 37 species, including the members of *Picornaviridae*, *Caliciviridae*, *Polyomaviridae*, *Reoviridae*, and *Picobirnaviridae* families. The majority of virus reads (90%) were mapped to bacteriophages and plant viruses which is because human waste was analyzed (bacterial viruses originating from gut bacteria and plant viruses from food). The detection of pathogenic viruses in commercial aircraft points to the fact that they might be stable in such environments, even though the aircraft viral communities are largely unknown due to limited research in the topic. There are still many unknowns in the virome dynamics in an aircraft cabin. To shed light on this question, the combination of “real-world” environmental sequencing and laboratory “simulation” research is required.

To study the stability and resistance of respiratory viruses such as Influenza or SARS CoV-2, non-pathogenic surrogates can be used as models. Ideal surrogates have similar response rates to the tested environmental stress factors such as the virus of interest.

Conditions and viral spread in spaceflight

In the spaceflight sector, there are constant concerns about the astronaut's health and well-being. The spaceflight environment has a set of stress-inducing conditions. Examples include microgravity, radiation, artificial light regulation, isolation, limited nutrition, and psychological stress due to distance from Earth. There is extensive evidence that space travel suppresses the human adaptive immune system (Guéguinou *et al.*, 2009; Akiyama *et al.*, 2020). Due to reduced activity in microgravity, bone and bone marrow mass reduces. Since the B-cells are produced in bone marrow, their numbers drop, and immune activity is reduced (Akiyama *et al.*, 2020). Additionally, spaceflight causes atrophy of thymus, which is the main source of T-cells (Benjamin *et al.*, 2016; Akiyama *et al.*, 2020). The astronauts therefore show symptoms associated with immune system suppression. Notably, the herpesviruses in their bodies reactivate, and they are more susceptible to allergies and infectious diseases. An astronaut who gets sick during a mission can hinder activity, compromising or altering a mission. Due to the reduced immune function in astronauts, infectious diseases and sepsis are considered to be among the most dangerous biological threats to long-duration space missions, meaning more than 6 months spent in space (Cowen, Zhang and Komorowski, 2024). The instances of astronauts getting sick during space missions have been recorded frequently during the 63-year history of human spaceflight. First astronauts like Yuri Gagarin (USSR) and Alan Shepard (USA) did not report any symptoms, as the durations of their respective flights were short, only 108, and 15 min respectively. Among the earliest examples of infections come from the Apollo 7 and 8 missions (NASA, 1968, 1969), where nasal infections burdened the crew. In those cases, the infections spread among the crew members rapidly and led to a range of discomforting symptoms, halting their performance. Apollo 7 and Apollo 8 missions lasted 11, and 6 days respectively. Since those incidents, astronauts are required to pass through a 14-day quarantine before and after a trip to space. Therefore, due to the imposed quarantine, it could be assumed that the conditions on space stations are sanitary and safe from the spread of disease. However, the data shows that astronauts are still not safe from on-board infectious diseases (Cowen, Zhang and Komorowski, 2024). The analysis of symptoms on 46 long-term International space station (ISS) missions, each lasting approximately 6 months, shows 3.4 incidences of infectious disease per flight year (Crucian *et al.*, 2016; Cowen, Zhang and Komorowski, 2024). The most significant infections were skin rashes (1.12 per flight year). Even though the infections currently encountered during spaceflight could be considered subclinical, it is important to know that currently, the longest duration space mission lasted for 437 from a cosmonaut Valeri Polyakov. That is more than 1 year and 2 months. Still, the missions planned for the Moon and Mars will be significantly longer. For instance, only the trip to (or from) Mars takes between 4 and 9 months, depending on Earth-Mars' relative positions

and the amount of fuel available. All things considered, infections during spaceflight should be taken seriously, despite currently yielding subclinical symptoms. The long-term immunodeficiency of astronauts in longer-duration missions might cause common infections to yield more severe clinical outcomes. Therefore, learning about the infections on the current, shorter, space missions will enable the implementation of countermeasures in long-duration spaceflight.

As there are microbial threats from the outside, they are also present on the inside of the astronauts. Most people on earth carry herpes viruses. Most of the time, they are kept inactive by the immune system preventing their activation. During immunosuppression such as in spaceflight, the herpes viruses can reactivate, being able to spread throughout the body, potentially causing disease or even infecting a healthy host. When herpesviruses reactivate, they can lead to a range of clinical symptoms. For example, Herpes simplex viruses (HSV-1 and -2) cause pain and burning on the skin, which is unpleasant, and it can require treatment, depending on the severity (WHO, 2025b). On the other hand, Herpes Zoster Virus (VZV) can cause severe neuropathic pain and widespread rash, with postherpetic neuralgia which patients feel after they experience the disease (CDC, 2025). Also, Epstein-Barr Virus (EBV) can cause fatigue fever and pharyngitis (CDC, 2024). Despite posing health concern, herpesviruses are far from the only microorganisms posing dangers to prolonged human spaceflight. Other viral and bacterial threats exist such as antibiotic-resistant bacteria and respiratory viruses. Respiratory viruses might spread more efficiently in reduced gravity such as that of Moon and Mars. That is because the large droplets carrying microorganisms stay resuspended in air for a longer time. A medical emergency in space would have much more serious consequences than on Earth. That is because of the limited resources, and special conditions different than on the ground. Resupplying missions to the moon would be costly, and to Mars even more so. Additionally, the orbital transfer window to most efficiently get to Mars opens roughly every 2.5 years. With the 4–9-month trip to Mars, any resupply would take between 3 and 6 years. Therefore, any settlers on Mars need to be self-reliant for the most part. With the limited resources on Mars compared to Earth and combined with the harsh environment of a thin atmosphere without oxygen, high radiation, and reduced gravity, it is likely that health treatment options there will be very limited but highly needed.

Space radiation is one of the factors that causes the immunosuppression of astronauts. On the ISS, the human body is exposed to a higher level of ionizing radiation than on Earth. However, the ISS is still to an extent protected by the Earth’s magnetic field, which captures the charged particles coming from the sun and outside of the Solar system (galactic cosmic radiation). The magnetic field captures the particles, which rotate around the Earth in two “belts”, the so-called Van Allen belts. The inner belt is at a height of 1,000 – 6,000 km, while

the outer belt is at a height of 13,000 – 60,000 km. The ISS orbits the Earth at roughly 370 - 460 km height. For reference, the average distance from Earth to the Moon is 384,400 km. The distance from Earth to Mars, at its closest is 56 million km. Both reach far beyond the protection of Earth's magnetic field. Neither the Moon nor Mars have a magnetic field. This means that the environmental ionizing radiation made of fast, charged particles such as protons, helium nuclei, and iron ions, on their surface is higher than the radiation onboard the ISS (Hassler *et al.*, 2014; Zhang *et al.*, 2025). It will likely cause more severe consequences for the immune system than currently observed. The current radiation shielding measures managed to protect the crew from the proton-radiation in the Van Allen belts to significantly less extent than the metal ion radiation (George *et al.*, 2024). This data was observed from radiation measurements inside the Orion capsule of the Artemis 1 mission.

After considering the external factors of the ISS that pose challenges to the astronauts, the attention should be given to the internal challenges. The microbial analysis of the environment of the International Space Station shows that the environment contains microbial communities similar to the average household (Kumar *et al.*, 2022), and even some antibiotic-resistant pathogens have been isolated from this environment (Urbaniak *et al.*, 2018; Sengupta *et al.*, 2024).

The production of pharmaceuticals such as antibiotics in space colonies such as on Mars will be crucial to sustain the health and wellbeing of the colonizers. As bacteria become more resistant to antibiotics (Padgen *et al.*, 2020; Akram, Imtiaz and Haq, 2023), they might be inconvenient to synthesize for space colonies. On the other hand, in vitro production of bacteriophages is simple compared to antibiotics and results in highly concentrated solutions. Additionally, phages co-evolve with the bacteria as they develop resistance mechanisms. All that is needed is the genomic DNA from the target phage, freeze-dried *E. coli* extract and addition of water. The solution incubated at 29 °C will produce a significant amount of phage. Therefore, the production of bacteriophages that could be used as medicine against pathogenic bacteria is an interesting concept to explore from the space mission's application point of view.

The exchange rate of astronauts on the ISS is significantly lower than in public transportation. Even with the advent of space tourism, it is unlikely that space travel will reach the level of passengers at the level of public transportation any time soon. This is due to the high costs and resources used in space missions. Still, the microbiomes of space stations are highly specific as it is impossible to prevent humans from carrying contamination onboard, and ISS is characterized by the unique environment. The microbial communities of space stations come with their own sets of risks.

Chapter 1: Establishment of PFU assays for infectious virion quantification and Loop-mediate isothermal amplification methods

1.1 Introduction

Studying the infectivity of human pathogenic viruses requires high safety standards. For many laboratories around the world, those safety standards are not achievable and therefore it is not possible to perform those types of experiments. This means that the possibilities to work with many real human pathogenic viruses are limited. Still, it is of interest to explore viral stability in an environment and also the antiviral effect of different treatments. This research helps to elucidate the threat that different viruses might pose in the environment and evaluates the novel antimicrobials for application in different settings.

1.1.1 Exploring viruses safely

To maximize safety and the possibility of working with any biological agent, including viruses, they are categorized into different biosafety levels (BSL). With microorganisms in the BSL-1 category, such as bacteriophages and plant viruses, standard safety precautions have to be applied. The BSL-2 category is used for pathogenic microorganisms not transmitted through air. Animal viruses are at a minimum in this category. Additionally, pseudoviral particles of coronaviruses, which are discussed below, also fall into this category. BSL-3 includes microorganisms that could transmit through the airborne route. This category includes human coronaviruses, influenza viruses, yellow fever virus, mpox virus, and others. BSL-4 is the highest safety level and includes human pathogens that pose a significant outbreak threat and cause severe symptoms. The higher the biosafety level, the more complex and costlier the laboratory security measures are. BSL-3 laboratories are not common, and there are in total four BSL-4 laboratories in Germany (Agency, 2017). The use of surrogates and pseudoviral particles is useful to study general viral resistance in BSL-1 or BSL-2 setting.

To allow widespread research of antimicrobials and infectivity of viruses, pseudoviral particles or viral surrogates could be used. Pseudoviral particles are virus particles that have viral surface proteins similar to the real virus but without its genetic material. Usually, in those particles, the viral genome is replaced with a marker gene such as luciferase or green fluorescent protein (GFP). When cells are infected with pseudoviral particles, the presence of the marker is evaluated inside the cells to determine the infectivity. Those particles fall into the BSL-2 category; thus, they allow working with enveloped particles resembling highly pathogenic viruses, but without being able to replicate in humans. Surrogate viruses on the other hand, fall into maximum BSL-2 group and therefore studying them is a widely available option. Experiments with surrogate viruses are a widely accepted method to study viral

response to the environment (Steinmann, 2004; Serrano-Aroca, 2022; Heffron *et al.*, 2024). Surrogate viruses structurally resemble or are related to human pathogenic viruses. Although there are notable differences between surrogate viruses and real pathogens, many surrogate viruses can even be used in BSL-1 settings, e.g. bacteriophages or plant viruses. As humans advance into natural habitats, their contact with wild animals becomes more common (Epstein *et al.*, 2020; Lee-Cruz *et al.*, 2021; Zhou and Shi, 2021; Kummer and Kranz, 2022; Lawrence *et al.*, 2023). This creates opportunities for newly emerging viral diseases, as viruses jump from animals to humans. When such events happen, surrogate viruses can, for example, help to find the most suitable methods of inactivation for newly emerging pathogenic viruses.

The main difference between animal and bacterial/plant viruses is that animal viruses in most cases possess lipid envelopes with incorporated receptor proteins. This is not a common occurrence for bacterial and plant viruses. While plant viruses generally lack an envelope, there are many examples of phages possessing an envelope, though this is less common than non-enveloped phages. Bacterial viruses have generally more widespread use since plant viruses have longer life cycles and many of them can rarely be found outside of the cell. Plant viruses spread through the plasmodesmata connecting neighboring cells, which are microchannels that allow cross-cell communication (Leslie, 2011). Plant viruses also do not lyse the host plant cells as these have a thick cellulose cell wall. Rather, they spread from plant to plant via vectors such as plant lice, which puncture the cells, collecting the viruses inside in the process. When the insect punctures another plant, it will deliver the virus into the cell (Walkey, 1991). Therefore, the structure and life cycle of plant viruses significantly differ from that of animal viruses. Two characteristics of viruses are important to look at when choosing a valuable surrogate. Depending on the condition that is explored, the surrogate virus should either have the same surface type and/or the same genome type.

Concerning the surface types, viruses can either have a protein surface capsid or a lipid envelope. Animal viruses possess both enveloped and non-enveloped species with high prevalence of the enveloped viruses (Valero-Rello and Sanjuán, 2022). This is especially true for respiratory viruses. Therefore, the surrogates for those types of viruses also need to possess a lipid envelope. Those surrogates will normally be another animal virus, which might be closely related to the virus of interest. However, those viruses are BSL-2. For the BSL-1 conditions, bacteriophages are required. Known bacteriophages mostly do not possess lipid envelopes. Only the bacteriophage family *Cystoviridae* and some archaeal viruses possess the lipid membrane, as far as it is known. For viruses without the lipid envelope, many bacteriophages can be fitting surrogates. However, a few of them stand out as the most widely used and recognized. Some of those surrogates are represented in Table 1.

Table 1. List of commonly used surrogate viruses

SURROGATE VIRUS	BSL	GENOME	LIPID ENVELOPE	SHAPE	FAMILY	DIAMETER
ESCHERICHIA PHAGE MS2	1	ssRNA	No	Icosahedral	*Fiersviridae*	30 nm
ESCHERICHIA PHAGE T4	1	dsDNA	No	Head-tail	*Myoviridae*	200 nm
ESCHERICHIA PHAGE PHIX174	1	ssDNA	No	Icosahedral	*Microviridae*	30 nm
ESCHERICHIA PHAGE M13	1	ssDNA	No	Filament	*Inoviridae*	700-2000 nm long
ESCHERICHIA PHAGE T7	1	dsDNA	No	Head-tail	*Autographiviridae*	60 nm
PSEUDOMONAS PHAGE PHI6	1	dsRNA	Yes	Spherical	*Cystoviridae*	85 nm
ESCHERICHIA PHAGE Qß	1	ssRNA	No	Icosahedral	*Fiersviridae*	30 nm
ESCHERICHIA PHAGE PRD1	1	dsDNA	No	Icosahedral	*Tectiviridae*	80 nm
ESCHERICHIA PHAGE PR772	1	dsDNA	No	Icosahedral	*Tectiviridae*	80 nm
PSEUDOALTEROMONAS PHAGE PM2	1	dsDNA	No	Icosahedral	*Corticoviridae*	56 nm
AVIAN INFECTIOUS BRONCHITIS VIRUS (IBV)	2	ssRNA	Yes	Spherical	*Coronaviridae*	120 nm
BOVINE CORONAVIRUS (BCV)	2	ssRNA	Yes	Spherical	*Coronaviridae*	120 nm
MURINE HEPATITIS VIRUS (MHV)	2	ssRNA	Yes	Spherical	*Coronaviridae*	120 nm
MURINE NOROVIRUS (MNV)	2	ssRNA	No	Spherical	*Caliciviridae*	35 nm

The spread of diseases via fomites, i.e. through contaminated objects, is a significant mode of transmission for many pathogens, including viruses. In public transportation, this is particularly challenging as various people contact the same objects. This is due to large scale of all public transport. Therefore, it is important to understand the scale of how long viruses can sustain on surfaces desiccated while they are still infectious. Also, the question arises if nucleic acid-based detection is used to detect viruses in public transportation, does it still mean that they pose a threat. To introduce the methods of virology into the laboratory, the preliminary experiments that were performed was the introduction of plaque assays for bacteriophages, as means of evaluation of virus resistance to desiccation. Those assays are crucial to understanding the infectivity of viruses in any environment. Therefore, stability and detection assays were performed on five different bacteriophages usually used as surrogates for human-related viruses. Those experiments helped to develop plaque assay methods and understand the environmental stability of different types of viruses. In the next sub-chapter, an overview of the bacteriophage surrogates used in the study is presented.

1.1.2. Overview of bacteriophages tested

In this PhD dissertation, six surrogate bacteriophages were used as models. They are represented schematically in Figure 3. All the phages differ significantly in their shapes, sizes, and genome types. At the start of the thesis, they were chosen to represent a wide variety of viruses and test their stability in various types of environments.

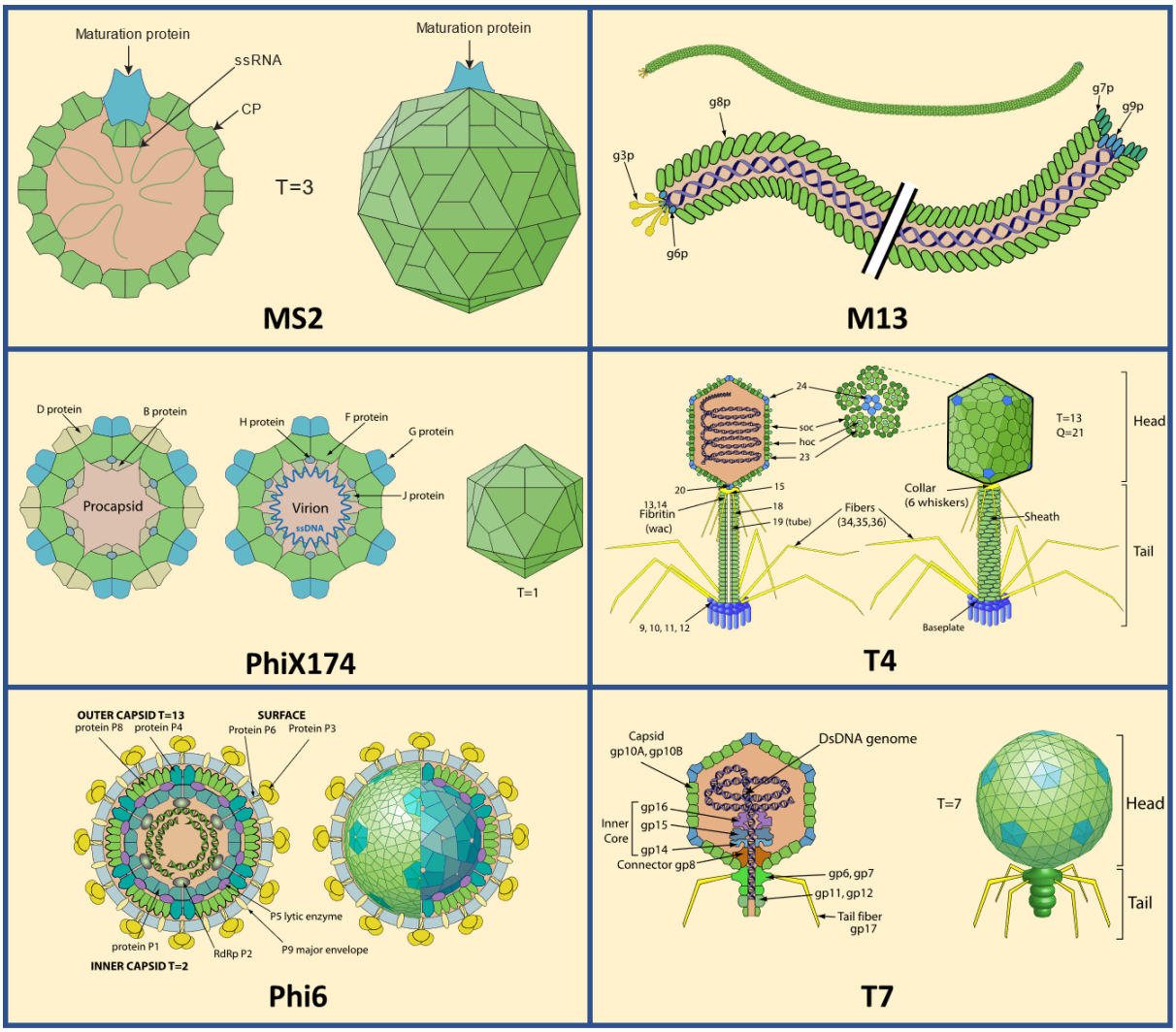


Figure 3. The bacteriophages used in the studies of PhD dissertation. They all diversify in structure, size, and life cycle. All phages are *E. coli* phages except Phi6 which is a *Pseudomonas syringae* phage. Designed and created from the schematics: Swiss Institute of Bioinformatics / ViralZone (viralzone.expasy.org)

Escherichia Phage MS2 is a widely researched small bacteriophage. It is often used to determine water treatment quality (Boudaud *et al.*, 2012; Yüzbasi *et al.*, 2022; Yang *et al.*, 2024). It is a small and relatively simple bacteriophage with a positive-sense single-stranded RNA genome of 3 569 base pairs and four genes in total. It has an icosahedral architecture and no lipid envelope. The diameter of the phage is approximately 30 nm. It has a life cycle of approximately 10 to 15 minutes, allowing for fast growth on the agar plates and in liquid. As a receptor it uses the F-pilus of *E. coli* which the bacteria usually use for conjugation. It attaches to the F-pilus through the single RecA receptor on the viral surface. It is currently unknown how the infection proceeds on the molecular level, but it is assumed that the phage inserts its genome through the F-pilus into the bacterial cell. There the positive sense ssRNA comes to the ribosomes and starts producing protein-coding genes. The RNA-dependent RNA polymerase is produced first, replicating the genome. Then the capsid protein (cp) is produced

together with RecA to encapsulate the ssRNA and produce live virions. Finally, when the cell is oversaturated with cp proteins, they bind to the translating RNA, opening the reading frame of the lysis (lys) protein which initiates cell lysis from inside the cell. The mature virions come out of the lysed cell.

Escherichia Phage M13 is a filamentous ssDNA phage with a circular genome. It has a diameter of only 6 nm, but its length can be 700 – 2 000 nm. It is used as a surrogate for filamentous viruses such as the family *Filoviridae* (Yuan, Zhang and Li, 2018) though some research shows that it is not always good surrogate for those viruses (Gallandat and Lantagne, 2017). Similarly to the phage MS2, it uses the F-pilus as a target cell receptor. Many capsid proteins are assembled around the coiled-coil circular ssDNA genome of 6 700 bp. It contains 11 protein-coding genes. The capsid is composed of 2 700 copies of major capsid protein g8p. On each side of the capsid, there is a unique structure. On one side of the capsid, there is a protein complex made of minor capsid proteins g3p and g6p while on the other side complex of minor capsid proteins g7p and g9p. The former complex is the receptor that recognizes the *E. coli* F-pilus, initiating the infection. The latter protein complex binds to the encapsulation signal made by the DNA 3D structure, kickstarting the packaging process. M13 phage enters through the pore at the base of the pilus after pushing the pilus into the bacterial cell, disintegrating it. Upon contact with the base, it injects the coiled-coil (+)ssDNA into the bacterial cytoplasm. The replication of Escherichia Phage M13 proceeds by a rolling circle mechanism. First, a host DNA polymerase produces the (-) strand of the viral genome, converting its ssDNA to closed dsDNA. This viral dsDNA is the replicative form (RF) of the genome. While the (+)ssDNA strand serves as a template for replication, the (-) strand codes for all the genes. Host RNA polymerase transcribes the (-) strand of the DNA. Genes g5p and g8p are the most actively transcribed genes. The protein g2p starts the replication by nicking the dsDNA at the (+) strand. In the rolling circle fashion, the (+) strand is synthesized and displaced again and again, producing many (+)ssDNA strand copies. Those copies are then also converted to RF dsDNA by the host DNA polymerase. At the moment the cell becomes oversaturated with g5p, this conversion of (+)ssDNA into RF is inhibited but the replication proceeds, making more (+)ssDNA strands which now each bind many copies of g5p. The g5p coat guides the genome to the membrane where g8p is located and oversaturated, replacing the g5p and gradually growing a virion from a bacterial cell.

Enterobacteria Phage PhiX174 is an icosahedral, tailless ssDNA phage. It has a diameter of 27 nm and a genome size of 5 386 bp. The circular ssDNA genome codes for 11 proteins, all of which are transcribed in the same direction. The capsid is composed of 60 copies of protein F. On each of the 12 vertices, there is a spike protein complex composed of five copies of G protein and one H protein. During an infection, protein H oligomerizes to form a temporary tail

for the phage to inject its genome inside the host (Sun *et al.*, 2014). Inside the capsid, protein J binds the ssDNA and condenses it for packaging. Those four structural genes (J, F, G, and H, in order as they appear in the genome) are the only non-overlapping genes. Other genes have a high level of overlap. The protein A gene is a DNA replication protein for the phage genome. It has an internal alternative transcription start site, producing the protein A* which interferes with the host DNA replication by cleaving the short DNA strands with the target sequence. The gene for protein B is also encoded within gene A but in a different reading frame. It is incorporated into the procapsid of the phage and is a crucial element of the capsid construction. Gene K is another gene transcribed inside gene A. It is a nonessential gene, enhancing the burst size of the phage, though the exact mechanism is not known (Gillam *et al.*, 1985). Its termination codon overlaps with Gene C. The gene C encodes for a DNA packaging protein. Protein C binds to protein A and host DNA polymerase to start DNA packaging (Aoyama and Hayashi, 1986). Right behind it in the genome, the gene D is present which encodes for a provirion scaffolding protein during the genome packaging. The D protein binds together structural proteins F, G, B, and H to form a scaffold while the DNA packages. After the packaging is complete, the protein D dissociates from the complex (Mukai, Hamatake and Hayashi, 1979; Asako and A., 2005). Inside gene D, there is a gene E which codes for a protein that induces host cell lysis. Protein E inhibits the activity of the host MraY protein, which is crucial for cell wall biosynthesis (Bernhardt, Roof and Young, 2000). The infection of the host cell starts with the spike protein complex (protein G homo-pentamer and H protein) interacting with the lipopolysaccharide of Enterobacteria. A series of structural changes during the interaction leads to protein H oligomerization into a tube, which transfers the viral circular ssDNA into the cell. In the cell, the host DNA polymerase synthesizes the (-) strand, creating a dsDNA (replicative form I, RF-I). The early replicative phage genes are then transcribed by the host RNA polymerase. Protein A cleaves the dsDNA (+) strand at the origin of replication and binds covalently to the 5´-end of the DNA. The remaining 3´-end serves as a primer for the rolling circle replication by host DNA polymerase. This produces a new dsDNA which is the replicative form II (RF-II). The replication proceeds until late viral genes start expressing. Protein C binds to the replicative machinery, starting the process of packaging the circular (+)ssDNA into the procapsids. The procapsids mature into capsids, and the protein E causes the lysis of the cell, releasing the mature virions outside of the host.

Enterobacteria Phage T4 is a tailed, complex bacteriophage. It contains a linear dsDNA genome of 168 903 bp in size, containing 289 genes. It belongs to the family *Myoviridae*. T4 is one of the most complex-tailed phages. Its structure contains a head, neck, collar, contractile tail, baseplate, and tail fibers. The elongated icosahedral head is made of 24 different proteins and contains the genome. The tail is contractile as it is made of an internal base tube made of 144 copies of gp19 protein and an external cylinder wrapping around the base, made of 144

copies of gp18 protein (Yap and and Rossmann, 2014). The external cylinder retains the possibility to contract by rotating and tightening the whole structure. The internal tube therefore gets pushed into the cell during the infection by the external contracting cylinder, ejecting the T4 genomic DNA into the cell. The base plate of the phage contains the connections to the tail fibers and has a lysozyme activity. The tail fibers attach to the lipopolysaccharides or OmpC receptor of an *E. coli* cell (Washizaki, Yonesaki and Otsuka, 2016). The double-stranded DNA genome of the phage is linear, but it has terminal repeats of roughly 1.6 kbp, enabling it to circularize the genome in the case the repeats are single-stranded. As in many other viruses, the genome is structured into regions according to the functionality. Therefore, there is a genome region dedicated to the head structural genes, tail, baseplate, neck and collar, nucleotide, and DNA metabolism regions. The genome during the replication and packaging is concatemeric, meaning that multiple genome copies are linked in series one after the another. However, each head gets only a single copy of the genome as it is cut by the terminase after entry into the head. The DNA is highly condensed to the near-crystalline density inside the phage head. The phage T4 has a gene for DNA methyl transferase, providing an opportunity to study the epigenetics of host/virus interactions. As T4 is a lytic phage, it does not require insertion into the host genome, making the infection process simpler than the temperate phages.

Escherichia phage T7 is a head-tail virus with a linear dsDNA genome. It is approximately 60 nm in diameter. The genome size is 39937 bp and harbors a total of 60 genes, all protein-coding. This phage was used as a model in the TX/TL synthesis experiments in microgravity. It has an icosahedral capsid head containing genomic DNA. The contractile tail is significantly shorter than that of T4. T7 has an interesting property that after it comes into contact with the host bacterial cell, it uses its tail fibers to randomly “walk” across the cell surface in search of the receptor to inject its genome (Kiss *et al.*, 2022). It achieves this by sequential confirmation changes of the single tail fibers at a time. The phage produces its own DNA-dependent RNA polymerase to transcribe its class II genes during the infection but does not carry it in the virion. Class I genes, such as the RNA polymerase itself, are transcribed by the host RNA polymerase. Another early class I gene is gp0.3 which inhibits the host restriction enzymes, gp0.7 which phosphorylates host transcription factors, shutting down the host transcription, and gp1.2 which inhibits the host dGTPase. After the transcription of the class II genes, the replication starts. Those genes are also responsible for the degradation of host DNA. The concatemeric viral DNA builds up through replication. Simultaneously, class III genes are transcribed which are connected to virion assembly and host lysis. First, the prohead assembles around the DNA, to be joined by the tail and its fibers. Among the structural class III genes, there are also lysis genes expressed which lyse the host cell, releasing the virions.

Pseudomonas Phage Phi6 is the phage with a lipid envelope. Therefore, it is considered a suitable surrogate for many human or animal viruses (Fedorenko *et al.*, 2020; Baker, Gutierrez and Gibson Kristen, 2022; Short. *et al.*, 2023). It is a dsRNA virus, belonging to the family *Cystoviridae*. Its genome is segmented into three parts: large, medium, and small segments. Phi6 virion is 60 nm in diameter. It infects the cell after binding to the PilA protein of *Pseudomonas syringae*. PilA is a structural protein of the pili of the host. The virion has an internal protein capsid and a membrane structural protein P9 incorporated into the lipid bilayer. The protein P3 is the spike protein that binds to the pilus of the bacterium. The virion gets attracted to the bacterial surface where it infects the cell. The protein P6 mediates the membrane fusion between the virus and the host. This enables the capsid to enter the periplasmic space after which the protein P5 degrades the peptidoglycan wall, enabling the capsid to enter the cytoplasm. The dsRNA genome of the virus is never exposed to the cytoplasm, evading the host antiviral mechanisms. The transcription of the viral proteins starts inside the capsid with the protein P2 which is an RNA-dependent RNA polymerase. The new mRNA exits the capsid where the new viral proteins are made and assemble around the RNA transcripts after which the (-) strand is synthesized. As the phage proteins are synthesized, the phage creates a capsid from the inner membrane of *Pseudomonas* and lyses the cell.

1.1.3. Loop-mediated isothermal amplification for virus detection

Additionally, the viruses were detected by loop-mediated isothermal amplification (LAMP) after drying, compared with the PFU data. LAMP is a nucleic acid detection method. Like PCR, it works by amplifying the targeted DNA, and if the reverse transcriptase is included, it can also amplify cDNA to detect RNA. It normally includes a set of 4-6 primers complementary to the targeted region. The design of the LAMP primers differs significantly from the PCR primers due to the distinct amplification mechanism. The amplification mechanism of LAMP is shown in Figure 4. The obvious difference between the two methods is that LAMP is done isothermally, meaning no temperature changes are required, while PCR requires a program of temperature changes over the reaction amplification. This is because in LAMP, the double strands are displaced due to the loops generated in the amplicons which open the double strand of DNA, making it possible for a primer to invade a single strand and continue the amplification. Because of that, the reaction only requires a heater and not a thermal block which can change the temperatures in a pre-programmed manner, as it is required for PCR. The primer design for LAMP has a different set of requirements than for (q)PCR. The mechanism of LAMP is based on finding the regions within genomes that have sequence structure similar to what is represented in Figure 4 part 1**.** LAMP uses a DNA polymerase which has a strand displacement activity, rather than exonuclease activity. This is important so that the DNA can be amplified isothermally. LAMP also requires a minimum of four primers. They are usually called Forward primer (F3), backward primer (B3), forward internal primer (FIP),

and backward internal primer (BIP). The process begins by the invasion of one of the internal primers (FIP or BIP) into the targeted region. The invasion unwinds the DNA double helix, and the continued replication displaces the second DNA strand. For example, in Figure 4, the F2 region of the FIP primer (F1c-F2) binds to F2c and DNA polymerase extends the strand, producing F1, B1c, B2c, and B3c. Therefore, the other strand gets displaced without heating it. This opens the way for BIP (B1c-B2) to bind to the now single-stranded B2c region and start the replication in the opposite direction, displacing the opposite strand. In addition, the B3c region was also displaced in the previous step, making it single-stranded and allowing the external B3 primer to bind behind the growing chain originated by BIP, displacing it during replication. This is shown in Figure 4 2. Because FIP and BIP have reverse complementary sequence to F1 and B1 regions, the first amplified strand of DNA will have a self-hybridized loop at one of its ends as shown in Figure 4 3. In the second replication step, the same invasion and displacement happens. Just from the other side, starting with BIP(B1c-B2). However, now this strand on its other finishes in a loop. Therefore, the final product now makes self-complementary single-stranded loops at its ends. In the middle of those loops are F2 and B2 regions which act as perfect spot for further FIP and BIP invasion and replication, making many looped products as in Figure 4 4. Additionally, two more primers can be added to enhance the sensitivity of the amplification. Those are called loop primers, and they bind anywhere in the loop regions of the single-stranded looped DNA structure. They help in the amplification of the target.

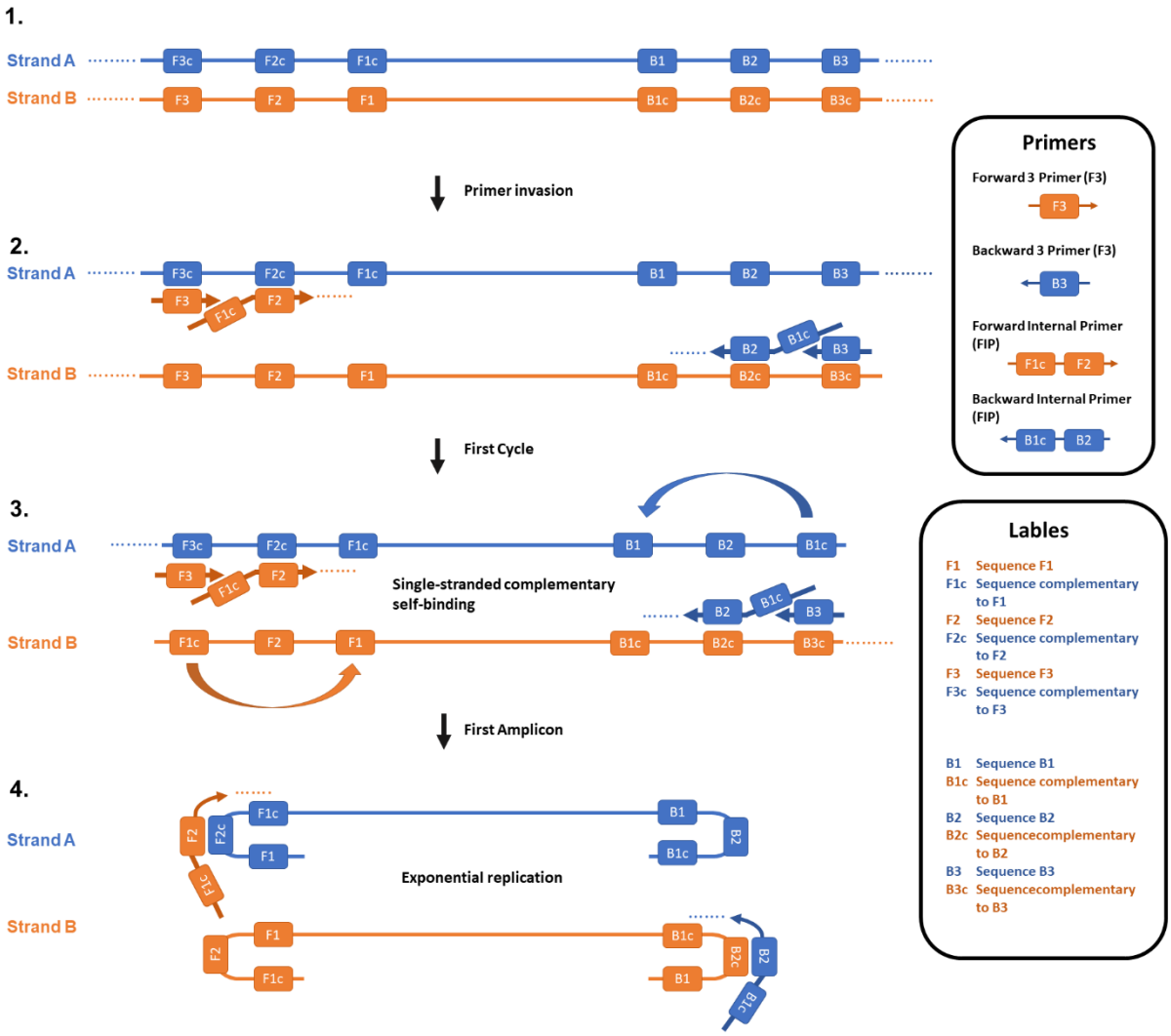


Figure 4. The mechanism of amplification in the LAMP reaction. 1 Starting point with both DNA strands. **2** Invasion and binding of primers to the template. **3** Displacement of the internal primers by the external ones and formation of loop DNA regions. **4** New binding of the primers for the next replication cycle. Created by Bruno Pavletić in MS PowerPoint.

LAMP replication works through strand displacement, rather than changing the temperatures to dissociate the double helix. The accumulating DNA can either be quantified fluorescently by adding fluorescent DNA dye such as SYBR green in the same manner as in qPCR. Alternatively, it can be detected qualitatively by observing pH changes that result from DNA replication since this reaction releases pyrophosphate ions ($P_2O_7^{4-}$). Pyrophosphate is an anhydride made of two phosphate ions bound through an oxygen atom. It is acidic in water solutions. A colorimetric pH indicator can be added to detect that change. Also, if there is extensive amount on magnesium ions present in the solution, magnesium pyrophosphate can crystalize, making the solution turbid. The increase in turbidity of the solution can therefore indicate DNA replication.

1.1.4. Experimental design

In this research, different bacteriophages usually used as surrogates for human viruses were explored. Here, all the bacteriophages and their hosts/media are described that were used throughout the thesis in the various experimental designs.

For four different bacteriophages, the relation of their multiplicity of infection and effect to bacterial growth was explored in an effort to make a simpler and more robust assay. The optical densities of liquid bacterial cultures were tracked over time to observe if there is a correlation with the phage/bacteria MOI and bacterial growth.

Second, phages were exposed on the glass slides in phosphate buffer saline (PBS) for 1, 3, and 7 days and PFU assays were performed to explore their detectability. In addition, loop-mediated isothermal amplification (LAMP) assays were performed to determine the possibility of molecular detection after drying exposure.

1.2. Materials and Methods

1.2.1. Bacteriophage propagation

The composition of the media and buffers that were used are listed in Table 2. The nutritious media used for the bacterial cultures are NZCYM (Sigma-Aldrich, N3643-1KG), Luria-Bertani Lennox (LB; Sigma-Aldrich, L3022-1KG), and Tryptic soy (TS; Millipore, 22092-500G) media. The buffers used were SM buffer and phosphate buffer saline (PBS). While the nutritious media were ordered from Merck, the buffers were prepared in the lab. After preparation, the media and the buffers were autoclaved at 120 °C for 15 minutes.

The bacteriophages and their respective hosts were ordered from the German Collection of Microorganisms and Cell Cultures (DSMZ). Upon the arrival of the host bacteria, the freeze-dried pellet was diluted in 500 µL of their respective nutritious media (Table 2). The liquid cultures were spread on solid agar plate media and incubated for 18 hours at their respective growth temperature (Table 3). The rest of the culture was frozen at –80 °C with the addition of 25% v/v glycerol as a cryoprotectant. After 18 hours of incubation, the bacterial hosts grew colonies. A colony of the bacteria was inoculated into 50 mL of the respective liquid media and left shaking at 150 rounds per minute (RPM) to grow for another 18 hours. The plates with the colonies could be held in the fridge at 4 °C for a maximum of seven days before disposing of them. From the new, pure liquid culture, 250 µL was added to a cryo-vial and 250 µL of 50% glycerol, resulting in the solution of pure bacterial cultures in nutritious media and 25% v/v glycerol. The solution was frozen at –80 °C and used as a stock culture for the host bacteria whenever they would be needed to grow. The bacterial growth curves were determined

through the measurement of optical density at 600 nm in the presence of bacteriophages at various multiplicities of infection, including in the absence of the bacteriophages.

Table 2. Recipes for the media, buffers, and chemicals prepared for bacteriophage experiments.

Media/buffer	Chemical	Amount	Concentration
NZCYM broth	Casamino Acids	1 g	0.0018 mol/L
	Magnesium sulphate ($MgSO_4$)	0.98 g	0.0082 mol/L
	Sodium chloride (NaCl)	5 g	0.0856 mol/L
	Tryptone	10 g	0.0490 mol/L
	Yeast extract	5 g	-
	Distilled water	1 L	-
Luria-Bertani broth (LB)	Sodium chloride (NaCl)	5 g	0.0856 mol/L
	Tryptone	10 g	0.0490 mol/L
	Yeast extract	5 g	-
	Distilled water	1 L	-
Tryptic soy broth (TSB)	casein peptone (pancreatic)	17 g	0.0008 mol/L
	Potassium hydrogen phosphate (K_2HPO_4)	2.5 g	0.0144 mol/L
	Glucose ($C_6H_{12}O_6$)	2.5 g	0.0139 mol/L
	Sodium chloride (NaCl)	5 g	0.0856 mol/L
	Soya peptone	3 g	-
	Distilled water	1 L	-
SM buffer	Sodium chloride (NaCl, ACS reagent, S9888-1KG)	5.8 g	0.1000 mol/L
	Magnesium sulphate heptahydrate (MgSO4•7H2O, Sigma-Aldrich, 63138)	2 g	0.0081 mol/L
	Tris HCl (1 M, pH 5.4, Sigma-Aldrich, T2194)	0.05 L	0.0500 mol/L
	Distilled water	1 L	-
Phosphate buffer saline (PBS), 10x	Sodium chloride (NaCl, ACS reagent, S9888-1KG)	80 g	1.3689 mol/L
	Potassium chloride (KCl, ACS reagent, Supelco, 1.04936)	2 g	0.0268 mol/L
	Sodium hydrogen phosphate (Na_2HPO_4, Sigma-Aldrich, S9763)	14.4 g	0.1014 mol/L
	Potassium dihydrogen phosphate (KH_2PO_4, Supelco, 1.04873)	2.4 g	0.0176 mol/L
	Distilled water	1 L	-

Table 3. The list of the bacterial hosts used and their media and growth temperatures.

Host	Virus	Media	Growth temperature
Escherichia coli F+ (DSM 5695)	MS2 (DSM 13767) M13 (DSM 13976)	NZCYM broth	37 °C
Escherichia coli (Migula 1895) Castellani and Chalmers 1919 (DSM 13127)	PhiX174 (DSM 4497)	LB broth	37 °C
Escherichia coli (Migula 1895) Castellani and Chalmers 1919 (DSM 613)	T4 (DSM 4505) T7 (synthesized in the lab)	LB broth	37 °C
Pseudomonas sp. (DSM 21482)	Phi6 (DSM 21518)	TS broth	25 °C

To cultivate bacteriophages, double-layer cultivation has been utilized. The bacterial lawns had to be produced for the bacteriophages to grow. The lawns were produced such that the plastic petri dish plates were filled with the nutritious media containing 1.5% w/v agar on the bottom and on the top, there was a layer of nutritious media with 0.75% w/v agar containing the host bacteria. First, the plates were prepared with 1.5% agar media. They were held at 4 °C, no more than 30 days. The 0.75% agar media was prepared in the 500 mL bottle and melted in a microwave oven for approximately 2 minutes at 800 W, every time the PFU assay was performed. The media was distributed into 15-mL tubes. There was 10 mL of 0.75%-agar media in each tube. When the media cooled down on air until it was possible to hold it in the hand indefinitely, 500 µL of the host bacterial culture at OD_{600}=1 was added. The solution was well mixed and poured onto the plate with 1.5% agar media. After solidification for 10 minutes, the double-layered agar plates with the bacterial lawns were ready to use. For the plaque assays, bacteriophage dilutions in PBS would be spotted on the agar plates.

All of the bacteriophages ordered arrived as freeze-dried pallets, except Phi6, which is not resistant to freeze-drying. Phi6 was received suspended in PBS. The freeze-dried pallet of each phage was diluted in 500 µL PBS. The bacterial lawn was prepared as described in the above paragraph. Then, 10 µL of the suspension was spotted on a bacterial lawn containing the dedicated bacteria for each virus. As Phi6 was delivered in the liquid suspension, it was spotted directly onto the bacterial lawn. The samples were left to dry and absorbed into the 0.75% agar for 10 minutes at room temperature. The plates were then left in the incubator for 18 hours at their respective temperatures. A clearance in the bacterial lawn developed at the place of the virus droplet, representing an area full of infectious virions. The top layer with

0.75% agar was scrapped off and put into a 15 mL tube. Then, 5 mL of PBS was added and vortexed thoroughly to release the virions. The tubes were centrifuged in a *swing-bucket* centrifuge Eppendorf 5810R for 30 minutes at 4,000xg. In this way, the agar, cells, and cell debris were sedimented while the viruses were left in the supernatant. The supernatant was filtered through a 0.2 µm filter. The solution was then distributed into 1-mL cryovials vials, each 500 µL of the virus suspension. To each vial, 500 µL of glycerol was added and the stocks were frozen at –80 °C. Those pure cultures were used as stocks for further experiments.

1.2.2. PFU assays

To determine the infectious titer of the bacteriophages, the decimal dilutions of virus samples were prepared in PBS. This was done in 96-well plates. In each well, 90 µL of the PBS was added. To the first well, 10 µL of viral sample was added. Then, 10 µL of the resulting solution was transferred to the next well with 90 µL of PBS. In this way, the various decimal dilutions were prepared for each phage. The phages were then plated by spotting and distributing the droplet on a designated area of the double-layered plate. By counting the plaques, each originating from a single virion, the infectious titer could be determined. The infectious titer is measured in plaque-forming units (PFU) per milliliter. It can be calculated from the plaque numbers based on the formula: IT = N / (V × D). Where IT is the infectious titer in PFU/mL, N is the number of plaques at the respective dilution (D, varying from 10^-1, 10^-2, ..., 10^-8). The volume added is always 10 µL. For the experiments including inactivation, the infectious titer after treatment would be divided by the original infectious titer to get the percentage of the remaining infectious virions.

1.2.3. Effect of phage infection to bacterial growth

Since PFU assays are complex and require a duration of multiple days per experiment, it was explored whether a phage infectivity could be tracked through tracking the optical density of the liquid media while bacteria grow. This would significantly reduce the experimental complexity and chances of error introduction. Therefore, bacterial hosts were added to their respective nutritious media in liquid form in addition to bacteriophages M13, MS2, Phi6 and T4 at various multiplicities of infection (MOIs). MOI is a ratio of the number of viral infectious particles and the host cells, so it can be thought of as number of infectious virions per host cell. Table 4 represents the phages used, their hosts, and MOIs explored in this research. The phage and bacterial starting titers were determined by plating them the same day of the experiment. The measurements were performed in biological and technical triplicates.

Table 4. The bacteriophages, their hosts, and multiplicities of infection (MOIs) explored in this research

Bacteriophage	Host	MOI
M13	*Escherichia coli* DSM 5695	30, 3, 0.3
MS2	*Escherichia coli* DSM 5695	15, 1.5, 0.15
Phi6	*Pseudomonas sp.* DSM 21482	3, 0.3, 0.03
T4	*Escherichia coli* DSM 613	5, 0.5, 0.05

1.2.4. Phage resistance to air drying

A test was made to determine the resistance of various bacteriophages to drying in PBS in a glass slide. On each slide, 20 µL of the virus dilution at approximately 1×10^9 PFU/mL of viruses M13, MS2, T4, and Phi6 was left on the glass slides. One batch was held at room temperature while the other was held at 4 °C for one, three, seven and thirty-one days. The glass slides were each collected into its own Eppendorf tube and diluted into 1 mL PBS. The solution was vortexed for 30 seconds. 10 µL was diluted in decimal serial dilutions in PBS and each dilution plated according to the protocol above. The graphical protocol is shown in Figure 5.

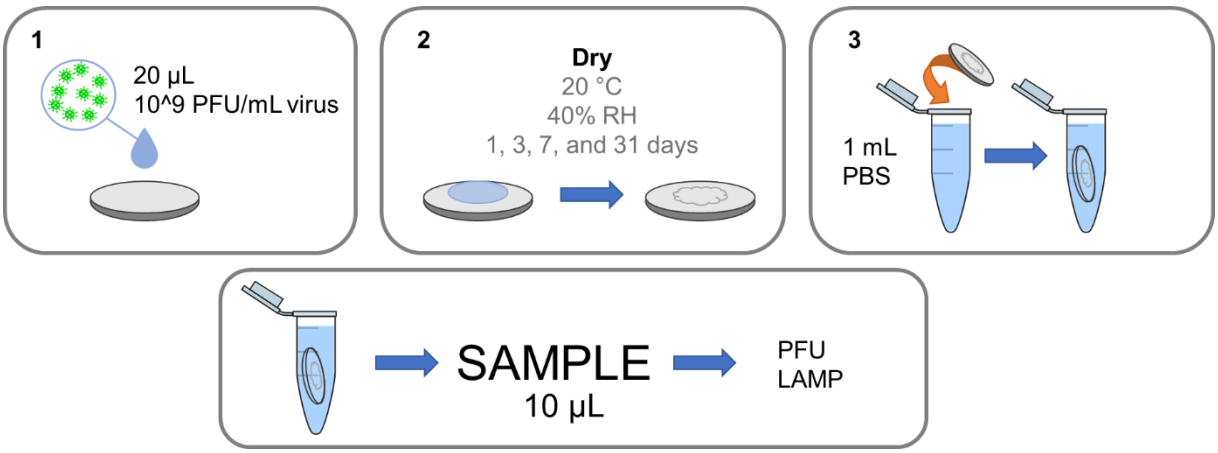


Figure 5. Graphical representation of the drying experiments and downstream analysis with different phages

LAMP detection

The primers for each of the phages were designed using the NEB LAMP Primer Design Tool. Table 5 shows the sequences of the primers that were designed and used for molecular detection of phages. The colorimetric WarmStart® LAMP Kit (New England Biolabs, #M1804S) was used to detect the viruses. To shatter the viral capsids, the solution was heated at 95 °C for 5 minutes. Since the capsids encapsulate and protect the genome, it was interesting to

explore the difference in the molecular detection when the heating (“boiling”) was applied and when it was not. Therefore, the LAMP reactions were performed with and without heating. The LAMP reactions were performed at 65 °C for 60 minutes. Normally, according to producer instructions, 30 minutes is enough to detect the target DNA or RNA, but in this case 60 minutes was chosen to eliminate uncertainties. The designed primers were ordered from Eurofins, Germany.

Table 5. The list of LAMP primers used for the detection of bacteriophages. The orientation of the primers is 5'-3'.

Primer	Primer sequence
F3_M13	AGCATTTGAGGGGGATTC
B3_M13	AACGCCAAAAGGAATTACG
FIP_M13	CCAGAGGGGGTAATAGTAAAATGTTAATGAATATTTATGACGATTCCGC
BIP_M13	CTCGCTATTTTGGTTTTTATCGTCGAGGCATAGTAAGAGCAACAC
LF_M13	AGACTGGATAGCGTCCAATAC
LB_M13	TCTGGTAAACGAGGGTTATGA
F3_MS2	AGGTGCCTTTCTCATTCGTT
B3_MS2	TACCCCTCGATGCATGGC
FIP_MS2	ATGTAGGAGCATCCCACGGGGGTCGACTGGCTCCTACCTG
BIP_MS2	AAGCGTTGACGCTCCCTACGGAGATTTGGGCCTTAGCAGT
LF_MS2	GCCCTCGAGCATGTTACCTA
LB_MS2	GGTGGACTGTGGAGAGACA
F3_Phi6	GCTGAGGGCTTATAGTGGTG
B3_Phi6	GGGATCGAGGATGTTGTCCA
FIP_Phi6	CAAAGAGACCGCGCTCATGGTTTATTCCTCCCCAGGAGTTCC
BIP_Phi6	GAGTGTCGCTCTTCTGCCTACGTGGGTTGAGGAGGTGTCC
LF_Phi6	AGTGGCCGAAATGGGAG
LB_Phi6	TCCCTCGAGTTGACGCTTCA
F3_PhiX174	GGTTTCGCTGAATCAGGTTA
B3_PhiX174	TCACCCATGCCTACAGTA
FIP_PhiX174	AGAGCAGAAGCAATACCGCCGAGATTATTTGTCTCCAGCCA
BIP_PhiX174	CCATGTCTAAATTGTTTGGAGGCGATCGGTAGCAAGCACATC
LF_PhiX174	GCACCAAACATAAATCACCTCACTT
LB_PhiX174	GTCAAAAAGCCGCCTCCG
F3_T4	CCCTTAATATAACGGATGCTATG
B3_T4	GGGCTATCTATAGAAATACCTCAT
FIP_T4	GGTCGGTAAAAAACAAGAATGGTTTTTAGGTAGAGCATCTTTGGC
BIP_T4	TTGACAAAGCGAATTCGTCCTAGGTTCTATTTATCCTTCTCTGC
LF_T4	ACAGTCGCTGAATATGAGAGT
LB_T4	TGCTCAAACAATTCTGGCC

1.3. Results

1.3.1. Effect of phage infection on bacterial growth

The growth curves of all the tested host bacteria *E. coli* and *Pseudomonas sp*. are shown in Figure 6. It can be observed that without the virus, bacteria clearly grew the fastest. The timepoints for *E coli* go only until 6 hours because 96-well plates were used for the measurements and at 37 °C, if higher timepoints would be used, the media would evaporate. For *Pseudomonas sp* the time points could go up to 20 hours because the temperature was around room temperature (25 °C).

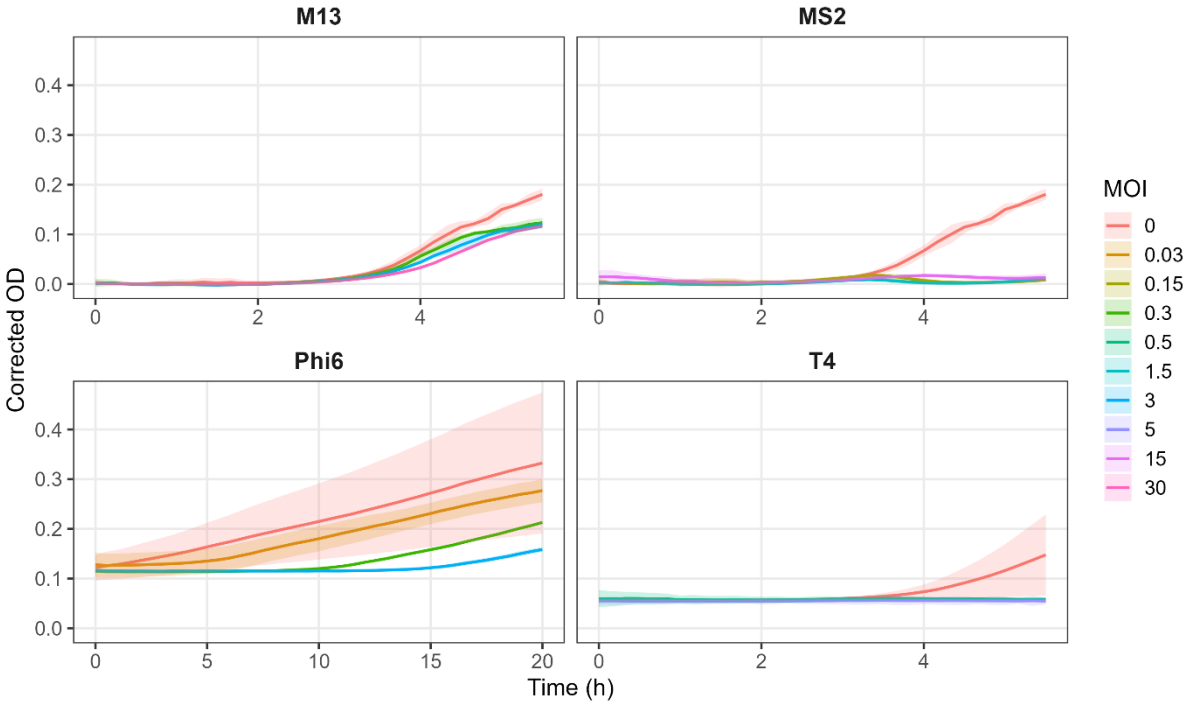


Figure 6. The growth curves of E. coli DSM 5695, DSM 613, and Pseudomonas sp. DSM 21482 in the presence of phages MS2, M13, T4, and Phi6 at different multiplicities of infection (MOI).

The MOIs differ between the viruses because of the different titers that were obtained of the viruses and bacteria. Therefore, the MOIs were calculated after the PFU and CFU assays were performed.

1.3.2. Phage resistance to air drying

The phages were also tested for their infectivity after drying on glass slides. The results are summarized in Figure 7. The infectivity of each phage was determined by PFU assays after drying.

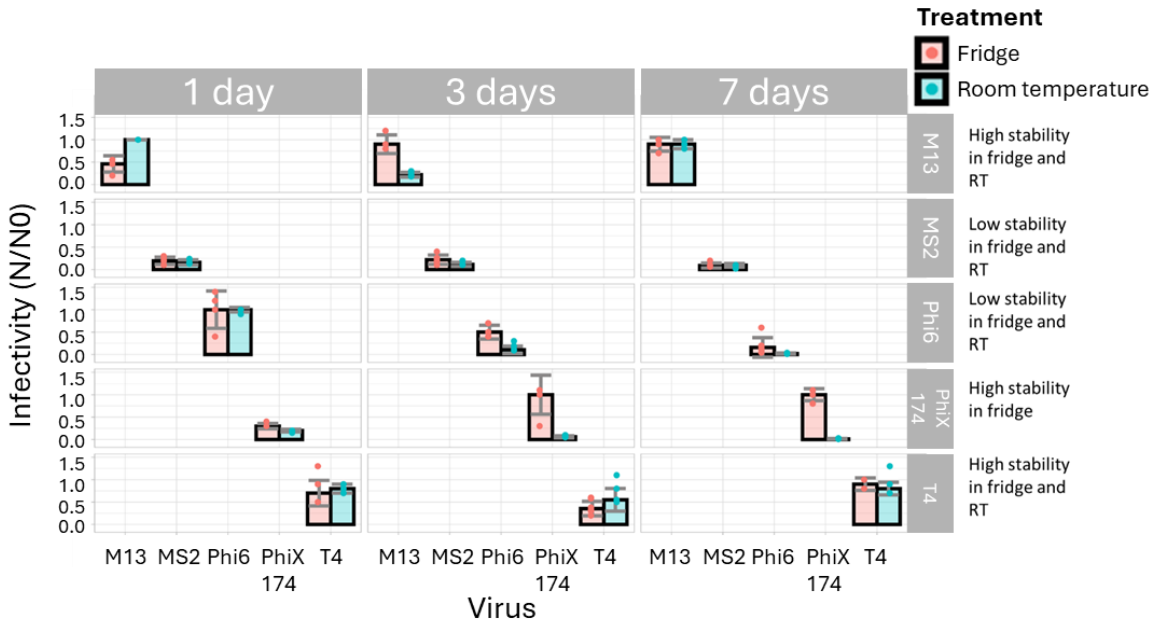


Figure 7. An overview of bacteriophage infectivity over the course of seven days dried on glass surfaces.

A significant variation can be observed between the bacteriophages and conditions. Particularly, M13 and T4 showed great resistance to drying. MS2 showed generally low resistance to drying, while PhiX174 showed higher infectivity after holding at approximately 4 °C. At the same time, the infectivity of Phi6 drops over time, and more rapidly at the room temperature.

1.3.3. Phage molecular detection via colorimetric LAMP assay

The colorimetric LAMP assays were also performed for different viruses after 1, 7, and 31 days. Figure 8 shows the results after “boiling” the viruses and without “boiling” at 95 °C for 5 minutes and after that cooling for two minutes on ice. A drawback to the colorimetric LAMP is that color intensities can vary and be relatively subjective to interpret. Additionally, the colors can be hard to see depending on the local lighting. Therefore, Table 6 represents the results as they were observed in the moment after the reaction was finished. It also represents the detection with PFU assays.

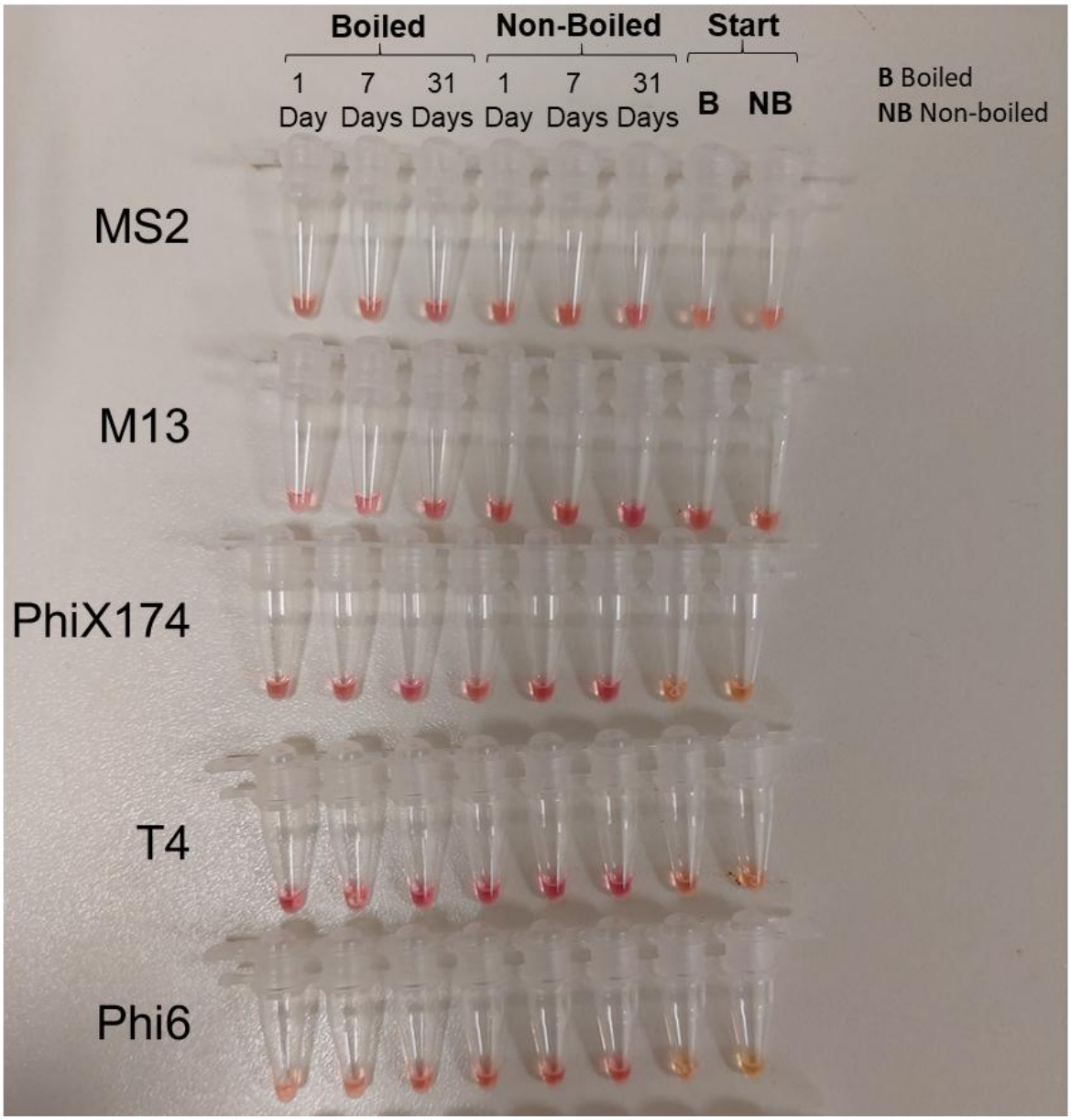


Figure 8. The results of the colorimetric LAMP assays for the bacteriophages after drying. The results are only qualitative but with varying color intensities. Therefore, the photographic results shown here need to be taken with caution. Boiling" refers to heating the phage solution at 95 °C for 5 minutes before the LAMP assay. Yellow color is a positive reaction while pink is negative.

Table 6. The summary of the results comparing bacteriophage detection with PFU assays and LAMP assays. The N/A values represent the experiment not performed.

Virus	Timepoint	PFU assay	LAMP assay
MS2	1 day	Detected	Detected
	3 days	Detected	N/A
	7 days	Detected	Detected
	31 days	N/A	Detected
M13	1 day	Detected	Detected
	3 days	Detected	N/A
	7 days	Detected	Detected
	31 days	N/A	Detected
PhiX174	1 day	Detected	Detected
	3 days	Detected	N/A
	7 days	Detected	Detected
	31 days	N/A	Not detected
T4	1 day	Detected	Detected
	3 days	Detected	N/A
	7 days	Detected	Detected
	31 days	N/A	Not detected
Phi6	1 day	Detected	Detected
	3 days	Detected	N/A
	7 days	Detected	Detected
	31 days	N/A	Detected

1.4. Discussion

Concerning the growth curves of bacterial hosts together with the phages tested (Figure 6),.it can be observed that various phages have different effects on bacterial hosts. In particular, it is visible that MS2 and T4 inactivated the bacterial growth completely at all MOIs tested. However, M13 and Phi6 showed bacterial growth for all MOIs, though suppressed. While MS2 and T4 burst a bacterial cell, effectively inactivating it, M13 follows a specific life cycle where it replicates inside the living host over time and the virions are exported outside of the live cell. This means that for some time, a cell can survive and replicate while carrying the replicating virus which continuously releases infectious virions to the environment. This could explain the fact that bacteria seem to grow, even at high MOI. For Phi6, as bacterial host grows slower, the phage life cycle is slower than for *E. coli*. Also, the host showed high variation in OD over time. So even though Phi6 showed a correlation between MOI and reduction of growth, the high variation in the host growth makes it unsuitable to correlate to PFU assay results.

From the phage drying (Figure 7) it was shown that different phages have varying stabilities after drying. Interestingly, all the bacteriophages were detected even after 7 days of being dried both at approximately 4 °C and at the room temperature. However, the infectivity varied greatly across different bacteriophages. Some phages showed high variability over time such as PhiX174 and M13. This might be due to the variability among the replicates combined with the experimental error. Here, it was shown how the bacteriophages of significantly varying structures greatly differ in their environmental resistance.

The LAMP detection assays (Figure 8; Table 6) show that phages could be detected after drying on the glass surfaces with nucleic acid assay. This shows a level of genomic integrity even after drying. However, it is important to note that it is not proof of genomic integrity. To show that the genome is truly complete, next generation sequencing should be done. Still, heating proved to rise the detectability of bacteriophages, proving that some genomic nucleic acids are still enclosed and protected within virions. Heating destroys the viral capsids and envelopes, releasing the nucleic acids to the environment. The combined results of PFU assays and LAMP assays are summarized in Table 6. The results show that LAMP can reliably detect bacteriophages even after drying, and most of the bacteriophages were detectable for up to 31 days. The only phages not detected after 31 days with colorimetric LAMP are PhiX174 and T4. Heating the samples at 95 °C for 5 minutes improved detectivity.

1.5. Conclusion

The virus detection methods and infectivity assays were successfully introduced. In an effort to correlate the effect of bacteriophages on bacterial growth, host growth curves were

compared in the presence of MS2, M13, Phi6, and T4. Correlation could not be shown with MS2 and T4 as they inactivated bacterial growth at all tested MOIs. M13 on the other hand showed close similarities between all MOIs applied due to its specific life cycle, again preventing a strong correlation. Phi6 was the only virus that showed a significant gradient of mean growth curves with varying MOI. However, the host showed high variability in growth, making the correlation between ODs and PFUs insignificant. Therefore, the PFU assays, though complex, are still an important tool to determine virus infectivity.

Therefore, for testing the resistance of the infectivity of various phages to drying at room temperature (~20 °C) and at approximately 4 °C, PFU assays were applied. Through the PFU assays, it was shown that the different phages have significantly different infectivity after drying. The filamentous phage M13 and complex head-tail phage T4 showed great infectivity after drying, even after up to 7 days at room temperature and in the cold environment. T4 is a dsDNA phage and M13 has an ssDNA genome. The small, simple, globular phage MS2 is an ssRNA phage and it showed low stability after drying up to 7 days. Another globular phage, PhiX174, carries an ssDNA genome and showed low stability at room temperature, but higher stability in the cold environment. The enveloped dsRNA phage Phi6 showed gradual reduction in infectivity over the course of 7 days.

The colorimetric LAMP assay was successfully established and utilized to detect the genomes of all bacteriophages. Heating the phage solution at 95 °C for 5 minutes improved the detection rate of all bacteriophages. Phages MS2, M13, and Phi6 were successfully detected with LAMP after up to 31 days. On the other hand, PhX174 and T4 were detectable after 7 days, but not 31 days.

In conclusion, various phage surrogate viruses were tested for their resistance to drying on glass slides. They were recovered and plated to compare their infectivity. The differences between their stabilities were obvious but despite that, LAMP was able to detect all viruses for at least up to 7 days. And most phages were detectible even up to 31 days of drying despite their infectivity dropping to undetectable level. The phage cultivation and detection was therefore successfully introduced and the differences in drying resistance and detectability were shown between carious types of viruses.

Chapter 2: Testing antiviral methods applicable to public transport and spaceflight

2.1. Introduction

As previously unknown viral diseases emerge unpredictably, it is useful to test antimicrobial methods against pathogens. They are valuable resources to slow or stop the spread of infectious diseases and therefore find a purpose in places where expected transmission might be high, such as hospitals, public transportation, and other public spaces. Most environmental disinfection methods generally target microorganisms, making them useful for a wide range of targets.

Among all the microorganisms, viruses are among the least stable in the environment, and most susceptible to disinfectants. On the other hand, spore-producing bacteria and fungi are among the most environmentally resistant microorganisms (Rheinbaben *et al.*, 2007; Geller, Varbanov and Duval, 2012; Peter, 2014). The spores contain a thick cell wall and remove all the water from their internal matrix, conserving the proteins, glycans, and nucleic acids, thereby preventing oxidative stress, chemical, and mechanical damage. In that state, microorganisms can sustain millions of years and be revived when the right conditions come, such as nutrients, water, and the absence of external stress factors. No viruses that are currently known possess any similar mechanism to spore formation. Their biological functions are dependent on the presence of the host cell. Without it, most viruses will only exist as molecular structures without biological activity in the environment. Therefore, viruses have lower resistance to environmental factors and lower stability in the environment than other microorganisms. In addition, they have smaller genomes and lower structural complexity than most cellular microorganisms. Their genomes are highly susceptible to damage from oxidative stress and radiation. Their molecular structure can be sensitive to desiccation and high temperatures.

As explained in Chapter 1, among other criteria viruses are divided according to the presence or lack of the lipid bilayer membrane at their surface, into enveloped and non-enveloped viruses. Generally, enveloped viruses are considered less stable in the environment than non-enveloped viruses. Since viruses differ in their environmental stability, it is reasonable to assume that their stability to antiviral surfaces might also be evident. Here, the stability of viruses on copper has been explored.

Antimicrobial surfaces are one of the well-established antimicrobial methods usable in public transportation and public spaces to reduce fomite transmission. Others include the use of personal sanitation products such as alcohol-based hand disinfectants, chemical surface disinfectants such as hydrogen peroxide or sodium hypochlorite, and light-based methods

such as UV or blue light treatments. The possession of such sanitation equipment is a personal choice, and hence it is difficult to regulate. On the other hand, chemical disinfectants that are applied on surfaces, as well as treatment with UV light, can have a damaging effect on the materials. Additionally, UV can be damaging to human tissues. Antimicrobial materials can potentially reduce the overall microbial burden in closed spaces, as microorganisms can stay active on conventional surfaces for between hours and months (Bäumler *et al.*, 2022). Currently, they are used e.g. in hospitals to prevent outbreaks.

Some surfaces have an antimicrobial effect by preventing the attachment of microorganisms. That is achieved through surface electrical charge, micro/nano-structuring, or hydrophobicity of the surface (Bäumler *et al.*, 2022). Also, there are antimicrobial coatings available for surfaces that do not work by repulsing microorganisms. Rather, they inactivate them when the surface comes into contact with the microorganisms through interference of metal ions with genomes or proteins on the surface. Additionally, those surfaces produce reactive oxygen species (ROS) that oxidize biological macromolecules. The efficiency of antimicrobial coating is still under debate. Some of the well-known metal-based antimicrobial surfaces are titanium, silver, zinc, and copper alloys. Metal alloys sometimes act as better antimicrobials than pure metals, as different metal ions have different minimum inhibitory concentrations to inactivate microorganisms (Zhang *et al.*, 2021).

Antimicrobial materials are widely researched in the context of medical implants. It is desirable that they are biocompatible and not prone to hold microorganisms. Titanium is a very attractive metal for this application due to its preferable biocompatibility with human tissues, but it lacks inherent antimicrobial activity. Therefore, many modifications of the titanium surface are performed to make it antimicrobial. Those examples include physical modification of its surface (Šístková *et al.*, 2024), zinc, silver, copper, cobalt or metal oxide coating, and use of titanium alloys (Rack and Qazi, 2006). The metal oxides that are used as coatings for titanium implants are usually zinc, titanium and copper oxides. They are known to be very strong antimicrobials, through the production of metal ions that penetrate the cell walls and membranes of bacteria and viruses, leaving pores in the structure and making the infectious agent burst. The antimicrobial alterations of surfaces are widely applied to implants and widely researched for bacteria, but less so for viruses. The main rationale behind this is that bacteria do not need the host to reproduce so if they find themselves on the material in question, they could be sustained there, gaining nutrients from the body and growing. A virus on the other hand, would need to find its target tissue, depending on the viral tissue tropism. Still, if an implant containing infectious viral particles is inserted into human tissue, the virions could potentially spread throughout the organism via the blood or cause inflammation.

For environmental testing, copper and zinc are the most attractive options since silver and cobalt as well as their oxides are expensive metals. Both were shown to be inactivating bacteria strongly (Poole, 2017; T. N. Nguyen *et al.*, 2022; Raja, Worthington and Martin, 2023). Crucially, copper and zinc ions are the ones that cause damage and formation of reactive oxygen species, and permeabilization of membranes and cell walls. In this project, the focus was on the antimicrobial qualities of copper on viruses. Copper is one of the most efficient antimicrobial materials known, and it has been used for that purpose since ancient times. In contact with viral surface proteins, copper ions affect the viral protein structure, making them unable to infect the host (Warnes, Summersgill and Keevil, 2015; Govind *et al.*, 2021; Rakowska *et al.*, 2021). This property is shared by copper with magnesium and zinc, all of which are used as antimicrobials (Rakowska *et al.*, 2021). Modern uses, however, require more than plain copper surfaces. In the case of the use in airplanes, copper is inconvenient as it is a heavy and costly material. Also, copper conducts electricity, which is not always beneficial for use as an antimicrobial in the real world. Another example is in hospital settings where copper materials are used, but inconvenient due to oxidation of the surface, thereby losing their antimicrobial effect.

Novel surfaces using copper as an antimicrobial agent have been developed to suit modern needs and use cases. To address the above-mentioned problems, copper particles are incorporated directly into the resin for 3D printing to maximize microbial activity. It is then printed by resin-based 3D printing, making it convenient for large-scale production. Embedding copper particles into 3D-printing filaments to combat antimicrobial resistance has also been developed in recent years. The surfaces tested in this research are code-named *Germ-free flying* (GFF) 8 and GFF14. GFF8 and GFF14 are produced by printing them from polylactic acid (PLA) resin with incorporated copper nanoparticles of copper (GFF8, 2% copper w/w) and microparticles of copper (GFF14, 2% copper w/w). Both GFF 8 and GFF 14 were produced at the DLR Institute of Lightweight Systems by Dr Fabian Kühnast and Dr. Sebastian Geier. These types of surfaces were chosen for exploratory purposes to determine the difference between nanoparticles and microparticles of copper in their antimicrobial effect. This copper amount was used in previous research, even though without comparison between the approximate particle sizes (Ekonomou et al. 2023)

Novel antimicrobial surfaces are worth researching, because they can act passively and reduce the environmental microbial burden. They inactivate infectious agents non-selectively and do not require significant maintenance. Additionally, they come at lower long-term costs than many other alternatives such as UV light or chemical disinfectants.

Pseudomonas Phage Phi6, described in Chapter 1, has long been used as a model virus for testing antimicrobials against viruses, without the need for high biosafety level laboratories

(Fedorenko *et al.*, 2020; Baker, Gutierrez and Gibson Kristen, 2022; String *et al.*, 2022; K. *et al.*, 2023). Despite general structural similarities of surrogates to human-pathogenic viruses, Phi6 has several differences compared to human viruses. The proteins and glycoproteins of the viral surface, as well as the genomes, differ among viruses and Phi6. Hence, for every antimicrobial tested, on any surrogate virus, caution must be taken when extrapolating the conclusions onto the human pathogens.

SARS-CoV-2 is a virus causing respiratory infection, which spread rapidly from 2020 onwards, causing a global pandemic. It is an enveloped ssRNA virus, mainly targeting lung tissue, though it is capable of infecting other tissues as well (Zeng *et al.*, 2022; Behboudi *et al.*, 2024). This broad tropism partially causes the wide variety of symptoms observed in various patients (Behboudi *et al.*, 2024). The virus has diverged into many different variants since its discovery, some of which spread more efficiently than others, and some supplanting entire viral populations, rapidly transforming their genotypic structure. Early in the course of the pandemic (February 2020), one of the diverged viral lineages, pangolin lineage B.3, was the first one that appeared in Germany, together with B.1 (Korencak *et al.*, 2022). Therefore, those lineages are closely related to the Wuhan strain that started the pandemic. Eventually, B.1 variant took over in the population and diverged further into more sub-lineages. It gave rise to other lineages such as B.1.1.7 (“alpha”), B.1.351 (“beta”), B.1.167.2 (“delta”), and B.1.1.529 (“omicron”). Each of those variants was more infectious than the previous, spreading through the entire population and giving rise to more daughter lineages from 2020 through 2023. The variant B.1.1.529 gave rise to new distinct lineages labeled BA.1 to BA.5. Again, each of those subvariants gave rise to more lineages. The recombination of BA.2.10.1.1 and BA.2.75.3.1.1.1 resulted in the completely new sub-lineage XBB.1. This lineage brought about XBB.1.5, which quickly became the dominant strain around the world, being highly transmissible and causing significant problems to health sectors throughout the world. SARS-CoV-2 left a significant mark in human society and history as patients were overflooding hospitals and more than six million people died worldwide. Furthermore, this period was marked by significant lockdowns and serious financial burdens to entire economies and healthcare systems.

SARS-CoV-2 efficiently spreads through direct contact and through air from person to person. Even though those are the main routes of spread for this virus, fomite transmission also occurs more rarely (Short and Cowling, 2023). The fact that the virus spreads through direct contact and air means that wearing masks and isolation are the most efficient ways to prevent further infection. However, the use of antimicrobial surfaces or other disinfection methods should significantly help in limiting the spread. As an example of a microorganism with well-studied ancestry and high significance, SARS-CoV-2 presents a significant opportunity to study the adaptation to an antimicrobial overtime. Therefore, it was selected for this study as one of the

viruses to evaluate the efficiency of copper-based antimicrobial surfaces. It is well-known that copper and copper-blends inactivate SARS-CoV-2 (Govind *et al.*, 2021; Purniawan *et al.*, 2022; Glass *et al.*, 2023). However, the resistance of different variants to copper has not been explored. Especially, the knowledge of the phylogeny and the timeline of SARS-CoV-2 variants, gives a unique opportunity to assess the potential of viruses co-evolving to adapt to different antiviral agents commonly used in hospital settings. Additionally, developing novel antiviral surfaces might be a good way to combat microbial adaptation to antimicrobials.

MPXV is an enveloped dsDNA virus of the family Poxviridae and the genus Orthopoxvirus, which mainly spread through direct contact. The virus causes skin lesions which are highly infectious upon direct contact. In addition to lesions, it causes fever, headache, muscle- and joint-pain, swollen glands and shivering. Poxviruses have a wide range of hosts, including vertebrates and insects. MPXV specifically has four clades: Ia, Ib, IIa, and IIb. The clade I members are more closely related to each other than to clade II and vice versa. Clade I members generally show higher virulence and mortality than clade II (~10% mortality compared to 1% of clade II). It is believed that clade Ia spreads from deceased and live animals, through households, and in hospitals. It mostly affects children up to the age of 15. On the other hand, clade Ib seems to be spreading mainly through sexual contact. The virus used to be endemic to Africa, but recently it started to spread to other parts of the world as well. In 2022, the mpox virus of the clade IIb caused a global epidemic and concern, spreading to more than 100 countries. After that, in 2024, the clade Ib started to spread outside Africa. Considering higher virulence and disease severity, this last outbreak caused a bigger concern. According to WHO, the clade IIa is currently present in West Africa as well as Nigeria, while Ia remains endemic to regions of DR Congo. WHO classifies the MPXV clade Ib as high risk. As of April 2025, there are still viruses of clade II circulating around the world, mostly travel-associated cases. Clade II caused more than 100,000 cases in 122 countries around the world where no previous mpox cases were reported. Even though there is currently no major concern regarding mpox reaching unpredictable levels of spread, the fact that it is circulating around the world, especially with clade Ib showing higher than usual virulence and severity, means that caution and care must be taken while tracking its spread.

Being a dsDNA virus, MPXV has significantly lower mutation rates than SARS-CoV-2. It also has a much more complex structure. Its genome, with the size of roughly 200 kbp, encodes around 200 proteins. In comparison, SARS-CoV-2 has genome length of 29.9 kbp and encoding 13 proteins. Like other Orthopoxviruses, mpox virus comes in two different infectious forms – an enveloped virion and mature virion. The enveloped virions are double-enveloped and exit the cell through exocytosis. The outer membrane originates from the surface of the cell. The mature virions exit by cell bursts and only have a single lipid membrane on their

surface. The mature virion is better suited for environmental stability and surface spread while the enveloped virion spreads better inside the host. Though the main form of transmission is considered to be direct contact, fomite transmission is also known to occur (Li *et al.*, 2023). This means that exploring the antimicrobial effect of copper surfaces might be beneficial in reducing the spread. Considering the nature of the spread of mpox, combined with the excessive knowledge of its phylogeny, this virus presents a valuable candidate to explore resistance to copper surfaces between various clades. Knowledge about the effectiveness of copper surfaces in inactivation of these viruses can help to reduce or prevent the spread of this disease.

In this study, we tested the efficacy of copper surfaces to inactivate strains of SARS-CoV-2, MPXV, and bacteriophage Phi6. The strains of SARS-CoV-2 used for assays were B.3 (a variant closely related to the Wuhan variant), B.1.617.2 (the “delta” variant), and XBB.1.5 (the “omicron” variant). The MPXV subclades that were tested were Ia (DR Congo endemic clade), Ib (cause of the 2024 epidemic), IIa (West and Central Africa endemic), and IIb (causing 2022 epidemic). The mature virion form of those viruses was studied. SARS-CoV-2 variants B.3 and XBB.1.5 were tested in artificial saliva to determine whether there is an effect of the surrounding matrix on the retention of infectivity. We compare the efficacy of the surfaces to inactivate virions to pure copper. As a reference with no antiviral quality, we used stainless steel. In addition, we tested novel 3D-printed antimicrobial materials made of polylactic acid (PLA) containing 1.5% copper nanoparticles (*GFF8,* or Germ-free-flying 8) and 1.5% copper microparticles (*GFF14*, or Germ-free-flying 14) to assess the resistance of SARS-CoV-2 B.3 variant and Phi6 to them.

2.2. Materials and Methods

2.2.1. Preparation of artificial saliva

The artificial saliva was prepared according to Woo et al (Woo *et al.*, 2010). All the components described there were mixed and diluted in 999 mL of deionized water and autoclaved. Then, 1 mL of Dulbecco's Modified Eagle Medium (DMEM, Thermo Fischer Scientific #11965092) was added to the mix. Table 7 shows the composition of artificial saliva solution.

Table 7. The composition of the artificial saliva tested in this research

Chemical component	Concentration
$MgCl_2$	0.18 mM
$CaCl_2$	1.00 mM
$NaHCO_3$	5.00 mM
KH_2PO_4	1.54 mM
K_2HPO_4	2.46 mM
NH_4Cl	2.05 mM
KSCN	1.95 mM
$(NH_2)_2CO$	2.00 mM
NaCl	15.06 mM
KCl	13.95 mM
Mucin	0.54 mM
DMEM	0.1% (v/v)

2.2.2. Preparation and titration of Pseudomonas Phage Phi6

Both the host bacterium (*Pseudomonas sp.*, DSM 21482) and Pseudomonas phage Phi6 (DSM 21518) were ordered from German Collection of Microorganisms and Cell Cultures (DSMZ). The host bacterium was inoculated with an inoculation loop into 20 mL of Tryptic Soy Broth (TSB) medium (Millipore, 51228) and incubated for 18 hours at 25 °C shaking at 150 RPM. The culture was spread on a Tryptic Soy Agar (TSA) Petri dish using a sterile loop. The bacteria were left to form colonies for 18 hours at 25 °C. A single colony was taken and diluted in 10 mL of TSB and left for 3 h to grow to OD_{600} around 1. Bacterial lawns were created by adding 200 µL of the solution into 10 mL of TSA with 0.75% agar (soft agar) which was first melted in the microwave oven and cooled to 48 °C in a heating bath. The soft agar TSA containing *Pseudomonas sp*. was poured immediately onto the Petri dish containing TSA with 1.5% agar. The petri dishes were left for 1 minute at room temperature for the soft agar to solidify.

Each sample of the virus was diluted in serial 10x dilutions by adding 10 µL of the viral solution to 90 µL of PBS. The dilutions were done until 10^{-4}x dilution. A droplet of 10 µL of each viral dilution was added to different locations on an agar plate with a bacterial lawn. The droplets were left to dry in a sterile environment for 15 min and left at 25 °C. The next day, plaques were counted where they were present and countable. The formula to calculate the titer was t(PFU/mL) = N/[V(mL)×D], where t is the titer, N is the number of plaques, V is volume of the virus solution (0.01 mL) and D is a dilution (10^{-1}-10^{-4}).

2.2.3. Preparation and titration of SARS-CoV-2 and MPXV

SARS-CoV-2 and MPXV strains were isolated from patients at the University Clinic of Bonn from the patient mucus and skin swabs respectively by Dr. Bianca Schulte. The strains were confirmed by Illumina sequencing. For the PFU assays, different cells were chosen for SARS-CoV-2 and mpox virus. For SARS-CoV-2, Vero E6 cells were chosen to be the host cell line, as they are widely used hosts for this virus (Jang and Ross, 2020; Ogando *et al.*, 2020; Essaidi-Laziosi *et al.*, 2021; Pires De Souza *et al.*, 2022; Melis *et al.*, 2024). For mpox, BSC-40 cells were used as they were giving a higher yield during virus propagation (Mucker *et al.*, 2024) (oral communication with Dr. Bianca Schulte), DMEM (Gibco, 12491015), Dulbecco's phosphate buffer saline (DPBS, ThermoFisher Scientific, J67670.AP), and Trypsin (Gibco, 25200072) were preheated in a water bath at 37 °C before the addition of Vero E6 to prevent stress caused by temperature changes. For cell line maintenance, Vero E6 cells were cultivated in T175 cell culture flasks with DMEM and 10% fetal calf serum (FCS). Cells were split one time per week by removing the media and washing the cell lawn with 10 mL of DPBS. The whole amount of DPBS was removed and replaced with 6 mL Trypsin. The trypsinization process took place for 2 minutes in the incubator at 37 °C and 5% CO_2, which speeds up the process. The cells were observed under the microscope to check whether they were detached. The cells were shaken, and the culture was hit gently against the surface to help with cell detachment. The cell suspension was then pipetted up and down gently to separate all cells from one another. Then, 2 mL of cell suspension was transferred into a growth medium with a volume of approximately 43 mL in a new culturing flask. Then, the culturing flasks were incubated at 37 °C and 5% CO_2.

For the contact inactivation assays, cell lines were plated in 24-well plates. Thus, after trypsinization of the cells, 9 mL of DMEM medium was added to the cell suspension to dilute the trypsin. Then the solution of 15 mL was transferred into a 50 mL falcon tube. A set number of cells (Vero E6 1.25×10^5; BSC-40 1×10^5) needs to be plated in each well of the 24-well plates. Therefore, the number of cells in the suspension were determined. In an Eppendorf tube, 10 µL of Trypan blue stain and 10 µL of cell suspension were added. Then, 10 µL of the cell-stain suspension was injected into an EVE™ Cell counting slide. The EVE™ Cell Counter

automatically counted all living cells. To calculate how much volume of the cell suspension should be used to transfer 1.25×10^5 of Vero E6 or 1×10^5 BSC-40 cells per well, the following formula was used: $V_{initial} = c_{final} \times V_{final} / c_{initial}$. To achieve the desired cell density, the necessary volume of medium was added to dilute the cells, the suspension was pipetted up and down gently to mix, then 1 mL of the cell, suspension was added to each of the 24 wells, and the 24-well plates were incubated at 37 °C and 5% CO_2 for up to 24 h.

Before the contact inactivation experiment, SARS-CoV-2 B.3 was diluted by a factor of 10, by adding 90 µL into 810 µL of the desired medium, either PBS or artificial saliva. In the case of variant XBB.1.5, it was diluted by a factor of 1.16, meaning 258.6 µL of virus was added to 41.4 µL medium. The variant B.1.617.2 was diluted 3 times by adding 100 µL into 200 µL DPBS. For MPXV, Ia was diluted 2.3x with DPBS by adding 130 µL of virus into 170 µL of DPBS, Ib was diluted 19.23x by adding 15.6 µL of virus suspension into 284.4 µL DPBS. The clade IIa was diluted 10x by adding 30 µL of the virus into 270 µL DPBS and IIb was diluted 6.66x by adding 45 µL into 255 µL of DPBS. Then, all the virus variants were exposed to the surfaces to be tested for specified time points. After exposure to the surfaces, the samples were diluted 300x except MPXV Ia which had a too low titer and was diluted only 65x. The medium covering the cells in each well was completely removed using a vacuum pump. Then, 200 µL of the diluted virus was added to each well containing Vero E6 or BSC-40 cells. After the whole 24-well plate was processed, it was incubated at 37 °C, and 5% CO_2 for one hour. Subsequently, the supernatant DPBS was removed with the help of a vacuum pump, and very quickly, 1 mL of an overlay medium, containing 1.5% (w/v) carboxymethylcellulose and 2x Modified Eagle Medium (MEM) in 1:1 v/v ratio was added to each well, and the plates incubated at 37 °C, and 5% CO_2 for three days.

After three days of incubation, the infected cells were fixed so that they could be analyzed and kept for a longer time. The overlay medium was removed from each well quickly. Then, the 24-well plates were fully submerged in a container filled with 6% formaldehyde. The plates stayed in the fixing solution for at least 30 minutes, after which the plates were washed with plenty of deionized water to remove formaldehyde residues and stained with 1% (w/v) crystal violet. One mL of the staining solution was added into each well and incubated at room temperature for 1 hour. Any excess staining solution was washed away with tap water. The 24-well plates were air-dried completely before plaques were counted.

2.2.4. Contact inactivation

Virions of the tested viruses, namely SARS-CoV-2 variants B.3, B.1.617.2, and XBB1.5, MPXV clades aI, bI, aII, and bII as well as their surrogate virus, Pseudomonas Phage Phi6, were prepared in Dulbecco's phosphate-buffered saline (DPBS). The starting titers of viruses are shown in Table 8. Each virus was treated in one set of experiments. However, both SARS-

CoV-2 and MPOX were compared to surrogate Phi6 which is therefore shown in the results for both viruses. The viruses were first diluted in DPBS and therefore corrected to approximately 25,000-30,000 PFU/mL because of the limited volume of available SARS-CoV-2 and MPOX virus solutions and to represent more realistic environmental count than the starting titers listed in Table 8. An untreated virus sample was taken at point zero (just before the experiment) in three 10 µL-technical replicates. For the contact inactivation, 40 µL of virus solutions were then spiked onto copper surfaces in three biological replicates for each virus. The copper/steel plate dimensions were 10 mm × 25 mm. The virus suspension was added to a defined, round area with a diameter of 6 mm. After each time point, 10 µL of each solution would be sampled. Those 10 µL-samples were diluted 300-fold in DPBS and then added to Vero cells (for SARS-CoV-2) or Bsc-40 cells (for MPXV) in 24-well plates to determine infectivity. The inactivation was determined by dividing the determined titer at every time point by the titers at the timepoint 0. In this way, the percentage of the virions that were left infectious was determined. In a similar way, the contact inactivation of Phi6 and SARS-CoV-2 B.3 was performed on surfaces GFF14 (with copper microparticles) and GFF8 (with copper nanoparticles).

Table 8. The list of viruses used and their mean starting titers with their respective standard deviations. The list of viruses used and their mean starting titers with their respective standard deviations. The titers were done on three biological and 3 technical replicates for each virus.

Virus	Solution	Mean starting titer	Standard deviation
SARS-CoV-2 B.3	DPBS	1.11×10^6	1.43×10^5
SARS-CoV-2 B.3	Artificial saliva	1.19×10^6	2.41×10^5
SARS-CoV-2 B.1.617.2	DPBS	1.72×10^5	9.20×10^3
SARS-CoV-2 XBB.1.5	DPBS	1.23×10^6	2.50×10^5
SARS-CoV-2 XBB.1.5	Artificial saliva	1.33×10^6	2.23×10^5
Mpox Ia	DPBS	7.43×10^5	1.52×10^5
Mpox IIa	DPBS	9.80×10^6	1.91×10^6
Mpox Ib	DPBS	3.78×10^6	4.88×10^5
Mpox IIb	DPBS	1.36×10^7	9.14×10^5
Pseudomonas Phage Phi6	DPBS	1.61×10^7	8.95×10^6

2.2.5. qPCR assays

For SARS-CoV-2, RNA was extracted using QIAamp Viral RNA Kit according to the kit manual. Each sample volume was 40 µL. Therefore, 40 µL sample of virus was diluted in 160 µL of DPBS. The RNA was therefore extracted from 200 µL of the sample according to the protocol.

For SARS-CoV-2, RealStar SARS-CoV-2 RT-PCR Kit 1.0 was used to perform the qRT-PCR. Passive reference dye from the kit was added to the master mix, and internal control was included.

The DNA of MPXV was extracted using QIAamp® DNA Mini Kit according to the protocol. Similarly, 40 µL of the samples were diluted in 160 µL DPBS, making 200 µL sample volume.

For MPXV, the NZYtech Monkeypox virus qPCR Kit was used for qPCR quantification of mpox according to the kit protocol.

All the qPCR runs were done on BIO-RAD CFX96 Real-Time PCR Detection System w/ C1000 Touch Thermal Cycler. The default amplification program was used for the PCR, being composed of 3 min initial denaturation at 95 °C, and then forty-nine PCR cycles with 10 s denaturation at 95 °C and 60 s amplification at 55 °C. The channel for Cy5 was used for amplification of Spike gene, and JOE™ for internal control.

The qPCR curves together with mean Ct values and standard deviation error bars was plotted using Python v3.12.7, with pandas v2.2.3 and matplotlib v3.10.1 libraries.

2.2.6. Glycoprotein electrostatic potential glycosylation profile analysis

The protein sequences and the PDB files from the Protein Data Bank (PDB) for all three tested strains of SARS-CoV-2 (B.3., B.1.617.2, XBB.1.5). The PDB IDs are 6vyb (for B.3), 7tov (for B.1.617.2) and 8vkk (for XBB.1.5). The PDB structures contain the spike protein in the open state (Receptor Binding Domain in the UP conformation). All the files also contain the glycans added to the structure.

The sequences upon which the structures were based were analyzed for glycosylation sites, and the distribution of electrostatic potentials was modeled with apbs v3.4.1 (Jurrus *et al.*, 2018) and pdb2pqr v3.7.1 (Dolinsky *et al.*, 2004). Program apbs was used to solve the Poisson-Boltzman electrostatic equations for each protein. Program pdb2pqr was used to prepare the PDB structures for next step of the pipeline by adding missing atoms and atomic radii. CHARMM36m forcefield model (Huang and MacKerell Jr, 2013) was used to model electrostatic potentials the pH 7.5. When the electrostatic potentials were calculated, they were plotted over the protein sequence with Python v3.12.7 with its packages (numpy, matplotlib, csv, and os) and visualized in PyMOL v3.1.0 Open-Source.

2.2.7. MPXV genome assembly and glycosylation comparisons

The genome of every MPXV strain was sequenced with Illumina sequencing (thanks to Dr. Bianca Schulte). The sequence quality was analyzed with falco v1.2.5 (de Sena Brandine and Smith, 2021). All the genomes were assembled based on the reference genome (RefSeq: NC_063383.1) as described in the further text. The adapters and polyA tails were trimmed

from the Illumina reads using Trim Galore v0.6.10. The reads were aligned to the reference genome through alignment by Burrows-Wheeler transformation (bwa v0.7.19-r1273) (Li and Durbin, 2009). The files were indexed with samtools v1.21 (Li *et al.*, 2009) to sort the reads by coordinates, index them and determine the alignment statistics. Finally, bcftools v1.21 (Danecek *et al.*, 2021) was used to make the consensus sequence from the aligned reads and determine the variants. The mutations in the surface proteins were considered. Among those, mutations were screened for any glycosylation site SNPs. To determine the SNPs, the consensus sequence was compared to the reference DNA sequence using NUCmer v3.1 (Kurtz *et al.*, 2004). This tool aligns two genomic sequences and identifies any SNPs or indels. When the nucleotide differences were known, the genes inside which they are found were determined through a custom python script analyzing the annotated genbank file of the reference mpox genome (RefSeq: NC_063383). Out of the affected proteins, the highest interest was in the viral surface proteins. Therefore, the protein sequences encoded by the affected genes were isolated into a separate fasta file. The transmembrane regions were identified, if present, using DeepTMHMM v1.0 (Hallgren *et al.*, 2022). This program uses deep learning to determine whether a given protein is transmembrane or not. Additionally, it identifies the potential transmembrane regions. Then, the proteins that were determined to be transmembrane were scanned for the potential glycosylation sites using NetNGlyc v1.0 (for N-glycosylation sites) and NetOGlyc v4.0.0.13 (Gupta and Brunak, 2002). Those programs identify the N- and O-glycosylation motifs and amino acids. They assign a score between 0 and 1 for the certainty that the site might be glycosylated. The threshold score to be assumed as a glycosylation site was 0.75. The protein sequences were then cross-referenced with UniProt database to determine whether the 3D structure or glycosylation sites were available.

2.2.8. Scanning electron microscopy

The copper, steel, GFF8, and GFF14 surfaces were imaged with a scanning electron microscope (SEM) (DSM Ultra 55, Carl Zeiss NTS, Wetzlar, Germany). To achieve high surface contrast, the secondary electron detector was used with an acceleration voltage of 5 kV. Additionally, the surface roughness was determined by a profilometer (Hommel Tester 1000, Hommelwerke GmbH, Schwenningen, Germany).

2.3. Results

2.3.1. SARS-CoV-2 infectivity on copper and steel surfaces

The copper inactivation experiments were performed to assess the stability of various strains of SARS-CoV-2 to copper. The strains that emerged in various times were tested. The results are summarized in Figure 9. While Phi6 shows by far the lowest stability on copper surfaces, it is followed by B.3, B.1.617.2 and finally XBB.1.5. Particularly notable is the significant difference between the strains of SARS-CoV-2.

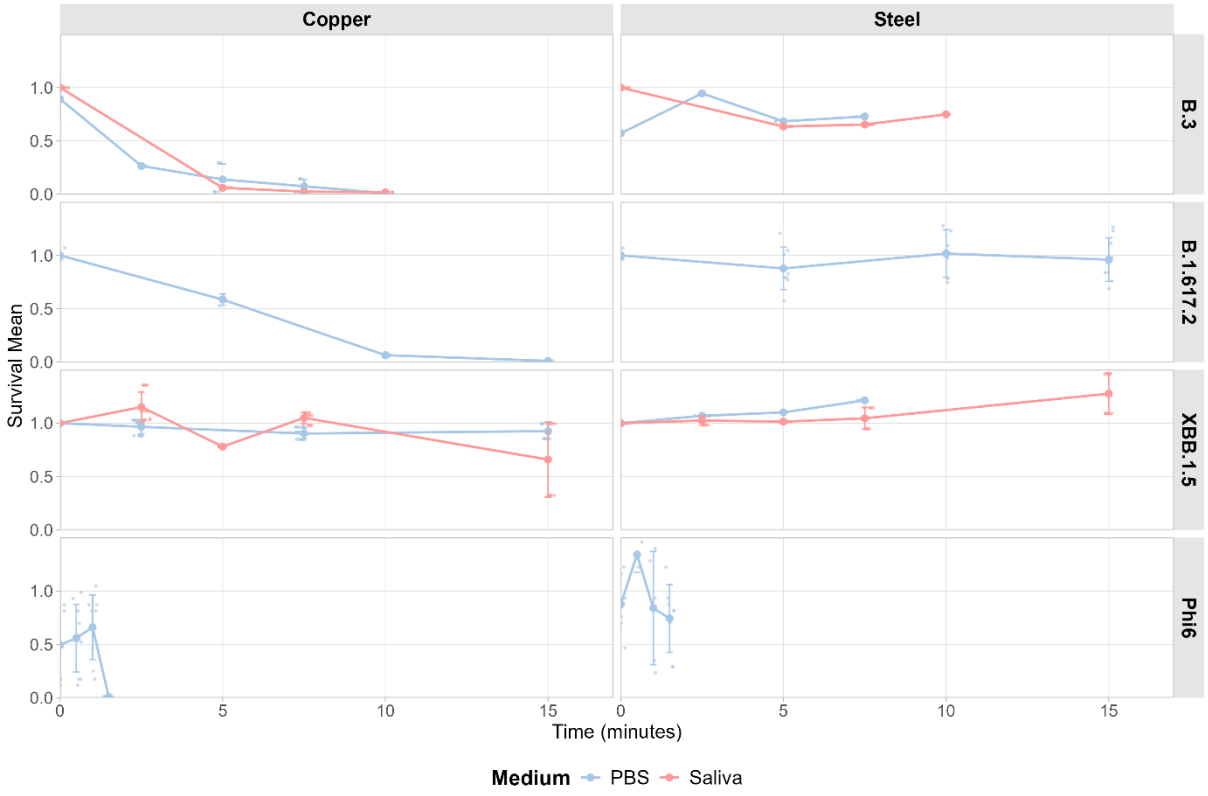


Figure 9. Inactivation of SARS-CoV-2 variants and Phi6 over time on polished copper and steel surfaces. Phi6 were tested for up to 2.5 minutes, as in the early preliminary experiments, they shown to be inactivated quickly. At steel surface, they would be inactivated after up to 7.5 minutes, as shown by the early experiments. Due to the limited number of surfaces available, the experiment with replicates was performed up to 2.5 minutes for Phi6. The experiments performed in PBS are shown in blue while those in artificial saliva are shown in red (SARS-CoV-2 B.3 and XBB.1.5). Each of the first three rows represents a strain of SARS-CoV-2 and the bottom row Phi6. The left column represents viruses treated on copper and the right on steel. The time is shown on the x axis and mean survival (N/N_0) on the y axis.

2.3.2. Spike protein glycosylation and electrostatic charge

The surface representation of Spike surface glycoprotein receptor (Figure 10). The protein models of spike protein revealed potential copper cation (Cu^{2+}) binding sites within the structure of each variant. The main assumption was made that those ions bind predominantly to free thiol groups coming from cysteines, histidine clusters that can coordinate a copper cation, or anionic aspartates and glutamates. Figure 10 shows the binding sites for the Spike protein of each SARS-CoV-2 strain tested here. Firstly, there were no free thiol groups. All of

those available at the surfaces of Spike were forming disulfide bonds among each other. Secondly, histidine clusters were not found in the structure. However, negatively charged surface Asp and Glu were available in each structure, though in different numbers. Those acidic amino acids and the glycan sites are represented on the protein 3D models.

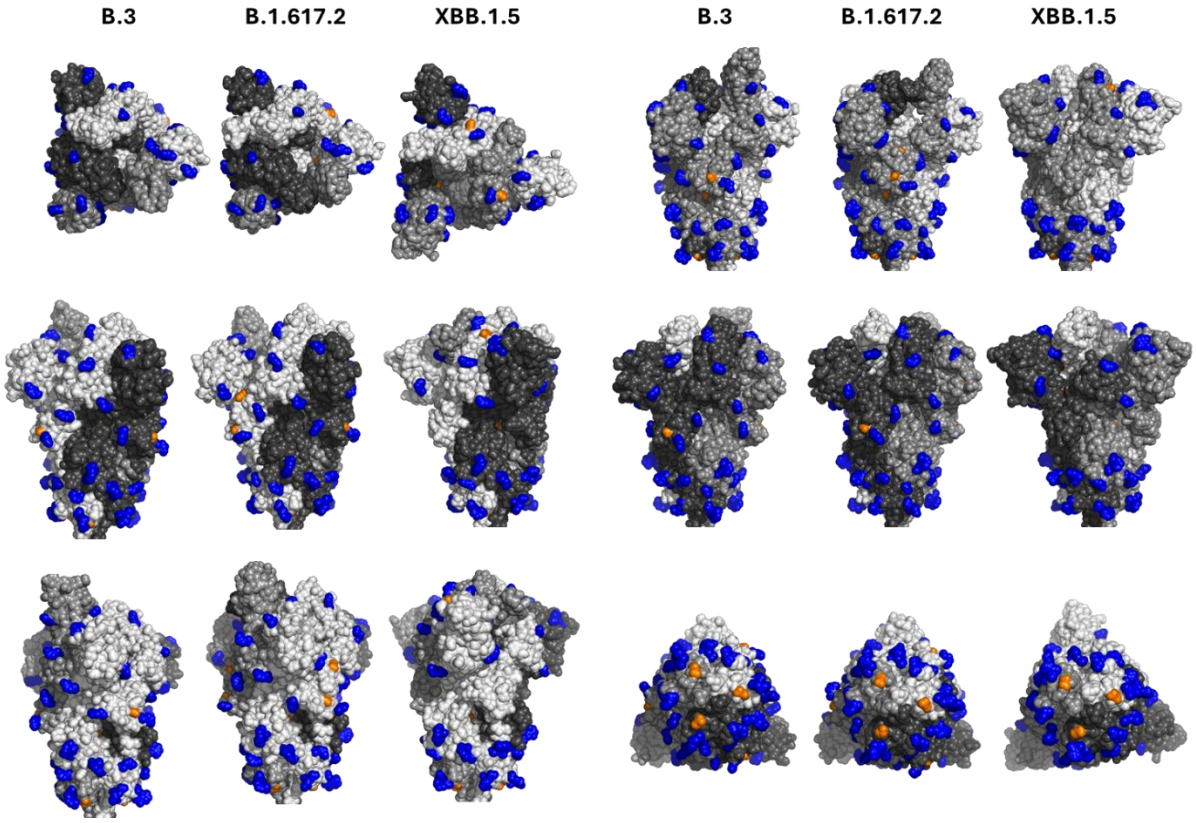


Figure 10. Comparison of SARS-CoV-2 spike protein points of all the tested strains from different angles. The PDB IDs are 6vyb (B.3), 7tov (B.1.617.2), and 8vkk (XBB.1.5). The glycan positions are highlighted in blue and aspartate/glutamate amino acids which can bind copper cations in orange. Each of the three Spike chains is colored in a different shade of gray (Visualized in PyMOL v3.10).

Additionally, electrostatic modeling was performed on the spike receptor for all tested SARS-CoV-2 strains. The histogram of calculated electrostatic potential distributions was plotted (Figure 11 A). Despite being useful to see the general charge distribution in a protein molecule, a histogram only shows the numbers of points over the whole structure which have similar electrostatic potential. Still, the local distribution of potential is not visible from the histogram. To determine whether glycosylation or structure of the surface spike receptor might play a role in this result, the computational analysis of the protein structure, as well as glycosylation sites, has been explored between the SARS-CoV-2 strains. The plot of the surface charge distribution over each amino acid in a chain are shown in Figure 11 B. It can be observed that the potential varies between chains significantly. Still, glycosylation sites are not necessarily associated with high or low electrostatic potentials by the position of amino acid. It seems glycosylation sites in general are, more often than not, neutral.

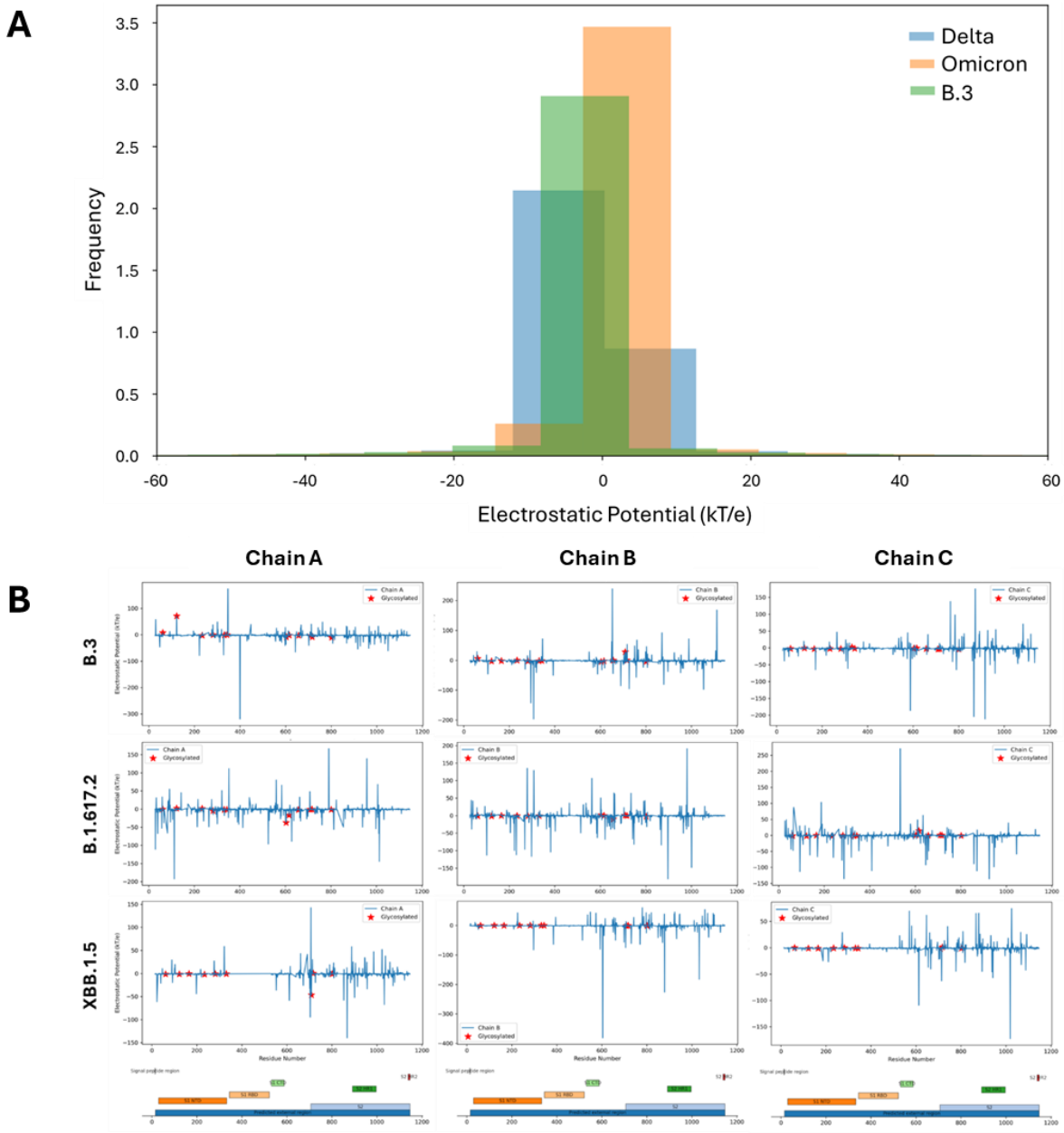


Figure 11. Electric potential distribution over the surface of spike glycoprotein chains. A Histogram of the electrostatic potentials present in the Spike protein of all three tested strains. Generally, XBB.1.5 (Omicron) has the most positive potential distribution of all the strains. At the same time, B.1.617.2 (Delta) variant has a divided distribution of potential. B Distribution of electrostatic potential across the peptide chains of Spike protein. Each column represents one chain (left to right: **A**, **B**, and **C**). The top row are the chains of B.3, the middle one is from B.1.617.2, and the bottom row is from XBB.1.5. Stars represent the glycosylated sites on each chain. The bottom bars represent various protein domains of each Spike chain.

Finally, to visualize the electrostatic potential distribution over the molecular surface, the Spike receptor 3D models with the overlayed electrostatic potentials are shown in Figure 12.

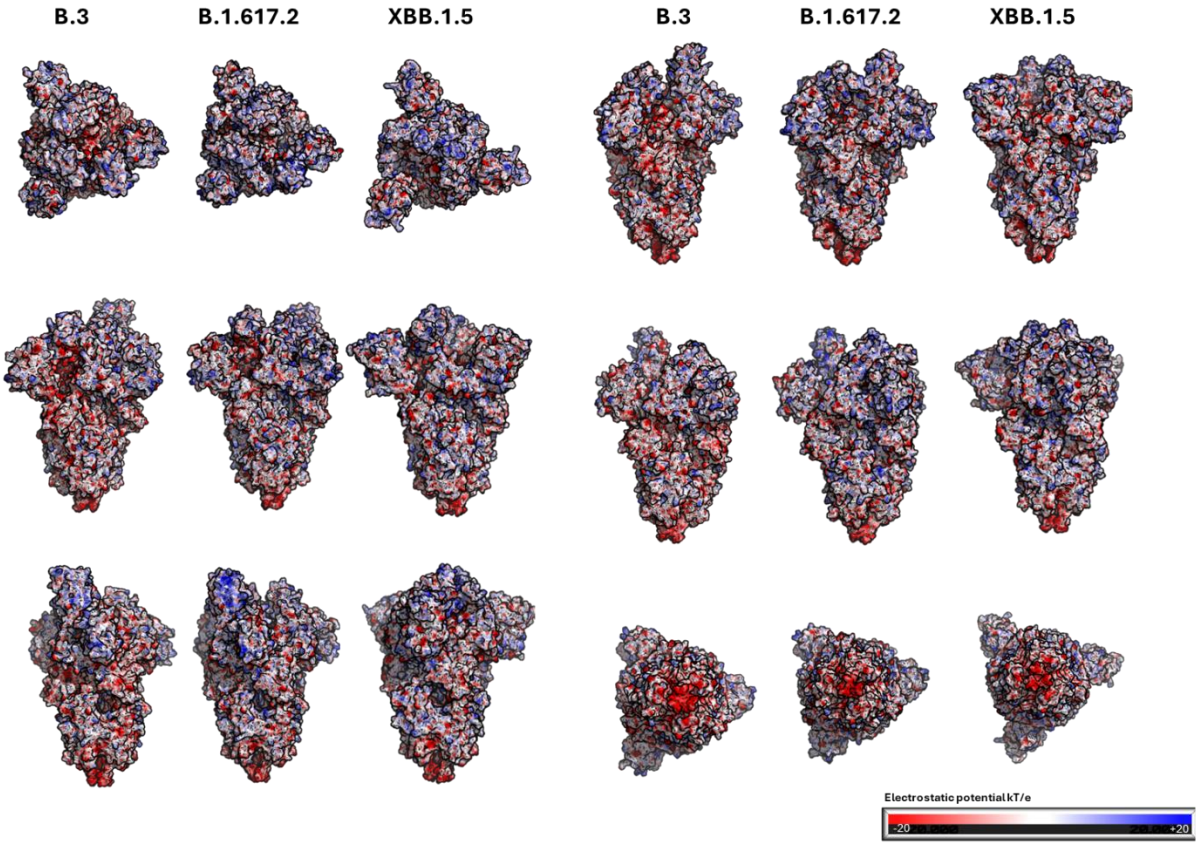


Figure 12. The distribution of electrostatic potentials over the external fragment of Spike protein surface at pH 7.5 of all three tested strains from various angles. The PDB IDs are 6vyb (B.3), 7tov (B.1.617.2), and 8vkk (XBB.1.5). Even though some spots show similar charge distribution, such as the bottom part of the protein, there are many differences in charges, as also observed in other analyses.

2.3.3. SARS-CoV-2 genome damage

One-step RT-qPCR was performed on all the samples to test if genome damage could be present. The mean amplification curves are shown on the upper panel of Figure 13. There are no observable differences between the curves. However, all the samples showed similar Ct values as could be observed from the curves. Even though the amplification curves are informative, they serve as a qualitative assessment. To quantitatively determine the differences in Ct values across the groups, their means and SDs were calculated and plotted. Also, they were compared through ANOVA and Tukey's post hoc test. This plot is shown in the bottom panel of Figure 13. This confirmed no significant statistical differences between any experimental group Ct values. However, they all only differed from the positive control's mean Ct.

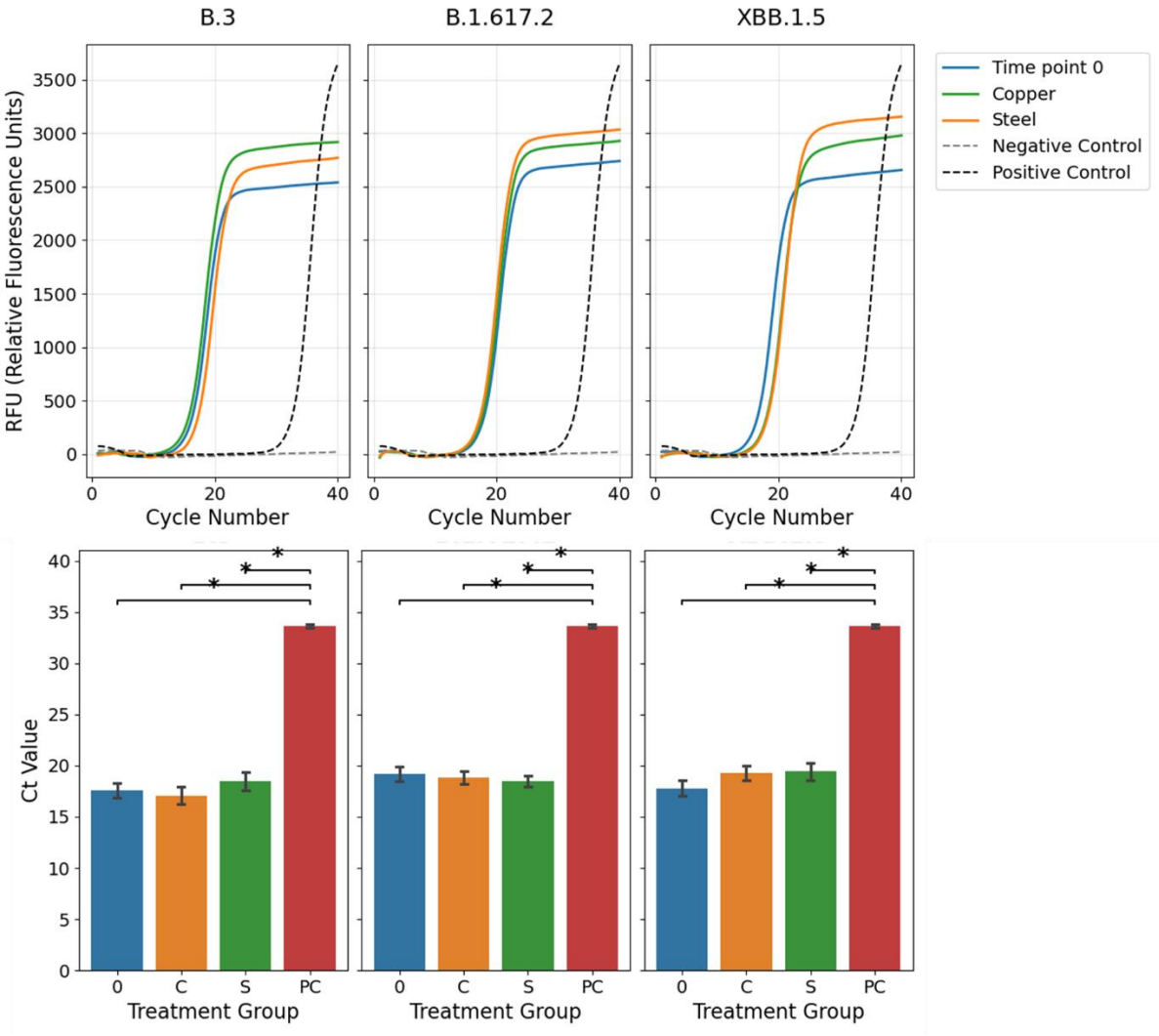


Figure 13. qPCR results of detection of SARS-CoV-2 variants after treatment on copper and steel. The upper row shows the mean amplification curves. Three biological and three technical replicates were performed. Due to high variation in the rapid amplification stage, 95% CI was omitted, but all the raw curves can be found in the Supplementary Figure 4. The bottom row shows Ct-values of qPCR of the amplification curves. 0 represents negative control, C is copper-inactivated, S is steel control, and PC is positive control. The non-template control was not detected at the threshold, so it is not shown.

MPXV infectivity after copper treatment

The results for MPXV contact inactivation assays are visible in Figure 14, which shows that the highest resistance to copper was exhibited by clade Ia. Clades IIa and IIb showed medium-level resistance, while Ib showed the lowest level. Ia is the only clade which didn't show a significant reduction in titer after 15 minutes of treatment. Again, Phi6 appeared not to be a good surrogate for mpox virus.

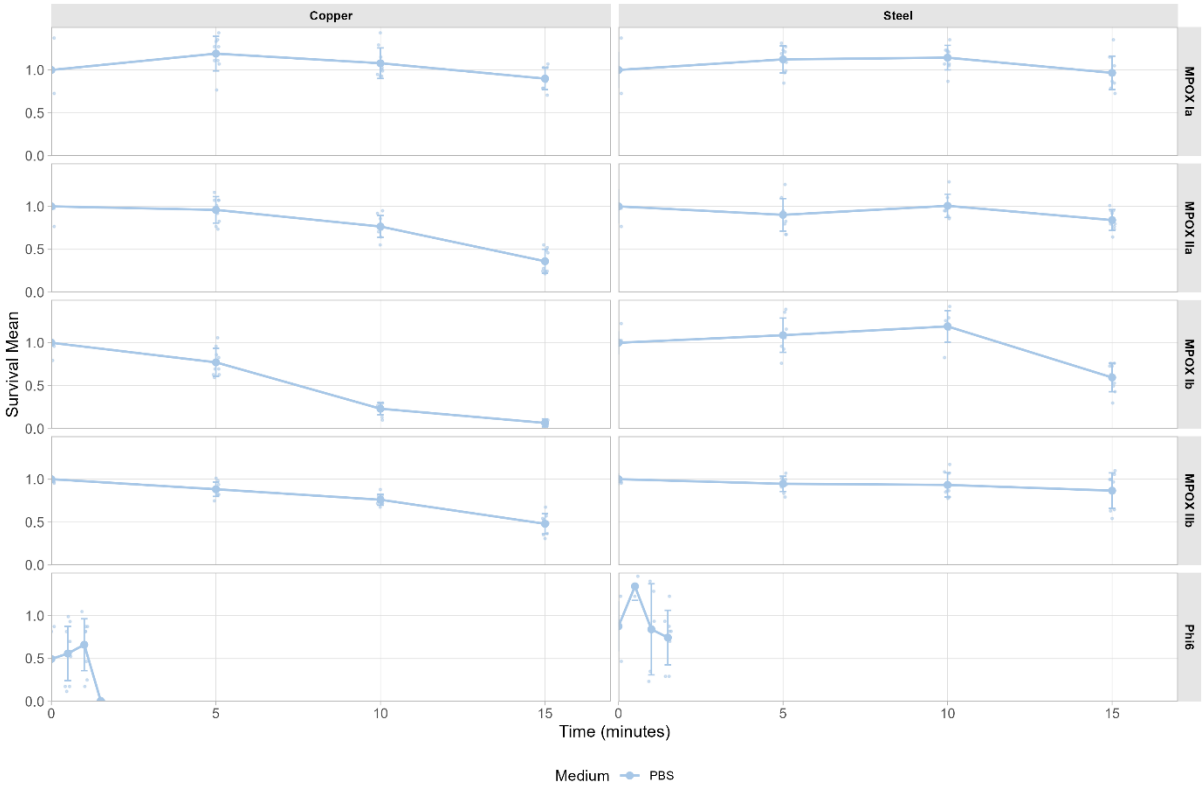


Figure 14. Inactivation of MPXV strains and Phi6 over time on polished copper and steel surfaces. Phi6 were tested for up to 2.5 minutes, as in the early preliminary experiments, they shown to be inactivated quickly. At steel surface, they would be inactivated after up to 7.5 minutes, as shown by the early experiments. Due to the limited number of surfaces available, the experiment with replicates was performed up to 2.5 minutes for Phi6. Y axis shows the survival rate (N/N0) while the x axis shows time points. Each row is one MPOXV clade and they are compared to the survival of Phi6 over time. The left column shows the curve for viruses inactivated on copper surface while the left column shows the ones on steel surface.

2.3.5. MPXV genome assemblies

The mechanism of the varying resistance to copper between various clades might be connected to the mutations in the surface proteins or their glycan profile. Therefore, the genome sequences of all the MPXV strains tested were analyzed for single nucleotide polymorphisms (SNPs) and other mutations. Figure 15 shows the SNPs found in the proteins determined to be transmembrane. MPXV IIa did not show any SNPs in transmembrane proteins.

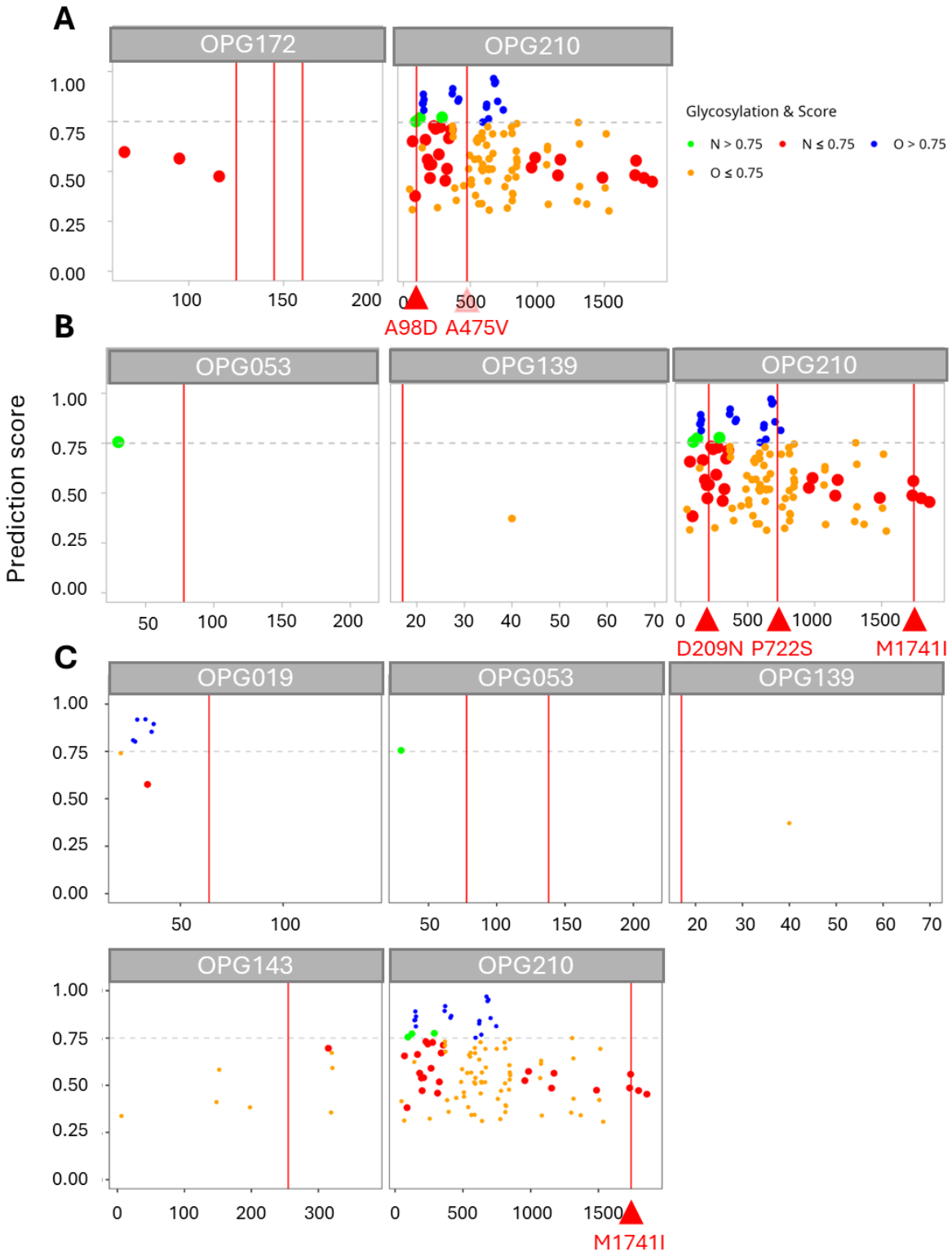


Figure 15. Transmembrane proteins with single nucleotide polymorphisms (SNPs) in different MPXV strains sequenced. A MPXV Ia, **B** Ib, **C** IIb. The horizontal dashed line is 0.75 score threshold from NetNGlyc and NetOGlyc. The smaller dots are the O-glycosylation sites and bigger dots N-glycosylation sites. The dots surpassing the threshold score 0.75 for N-glycosylation sites are green and for O-glycosylation sites are blue. The dots below the threshold score 0.75 for N-glycosylation sites are red and for O-glycosylation sites are orange. The SNP positions are labeled with vertical red lines. The mutation M1741 in the surface glycoprotein OPG210 in the potential glycosylation region is labeled.

2.3.6. Mpox qPCR detection

To determine detectability and potential genome damage of MPXV exposed to copper surface, qPCR was performed. The amplification curves of the qPCR reactions together with the Ct values of the qPCR analysis are shown in Figure 16. It can be seen that all the samples differ significantly from both the positive and the negative controls. Interestingly, among the samples themselves, there is no significant difference except in the clade IIb where both steel and copper have slightly, but significantly higher Ct values than the non-treated control (Figure 16). The single curves are represented in Supplementary Figure 5.

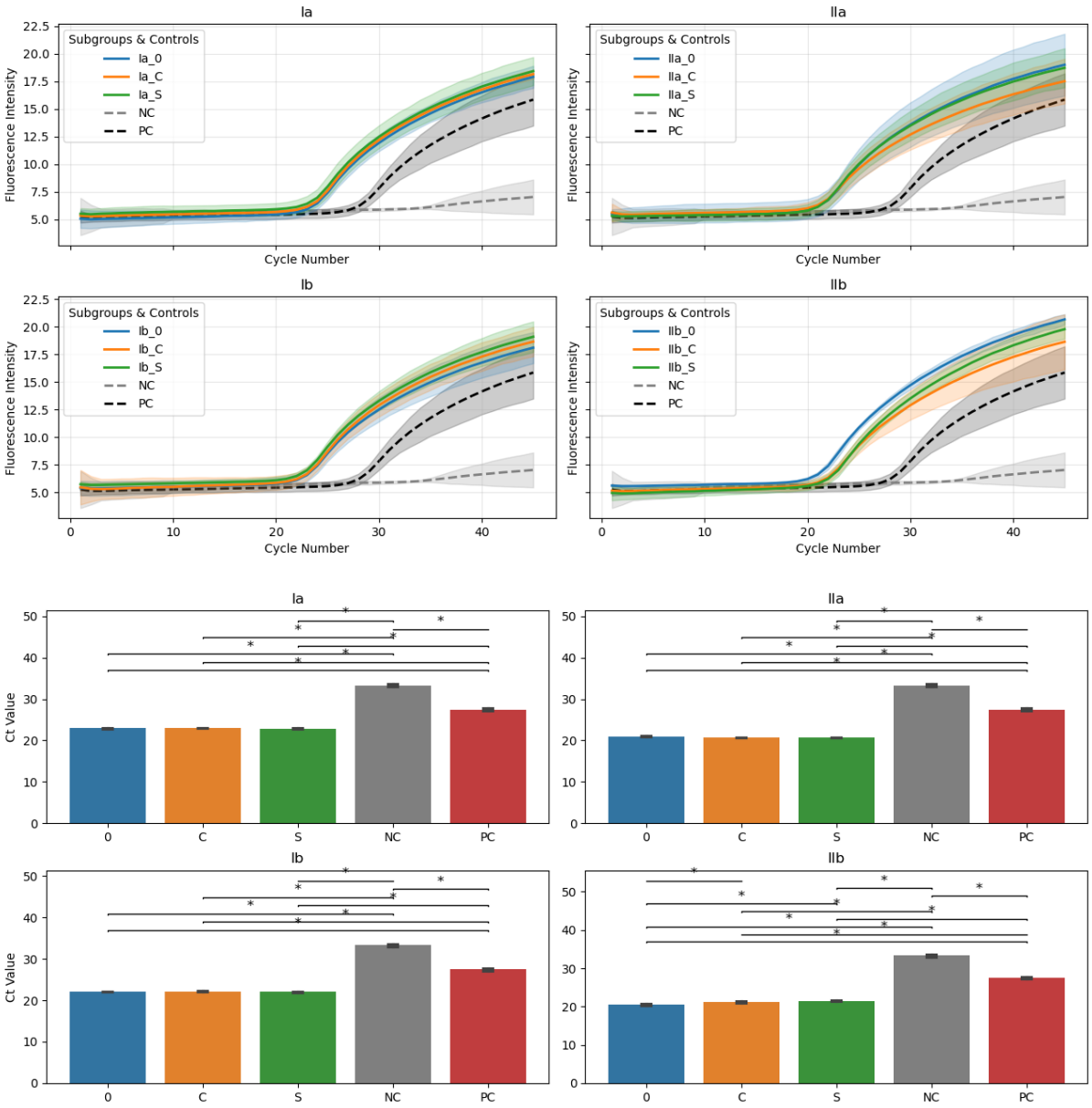


Figure 16. qPCR results of detection of MPXV variants after treatment on copper and steel. Above, the qPCR amplification curves of MPXV DNA for each clade. The average values of biological and technical triplicates are plotted together with 95% confidence intervals. 0 is the starting time-point, C is copper-inactivated, S is steel control, NC is negative control and PC is positive control. Below, the average qPCR Ct values of MPXV clades after copper-based inactivation. 0 represents the negative control, which was not treated, C is copper-treated, S is steel-treated, NC is negative control without DNA, and PC is positive control. The stars show the statistical significance of the difference between any two of the groups.

2.3.7. 3D-printed antimicrobial surface testing

To test the antiviral activity of 3D-printed materials containing copper, SARS-CoV-2 and Phi6 were exposed to them over the course of up to 90 minutes. The surfaces tested are GFF14

containing copper microparticles and GFF8 containing copper nanoparticles. Figure 17 shows the results of the inactivation experiments on those surfaces.

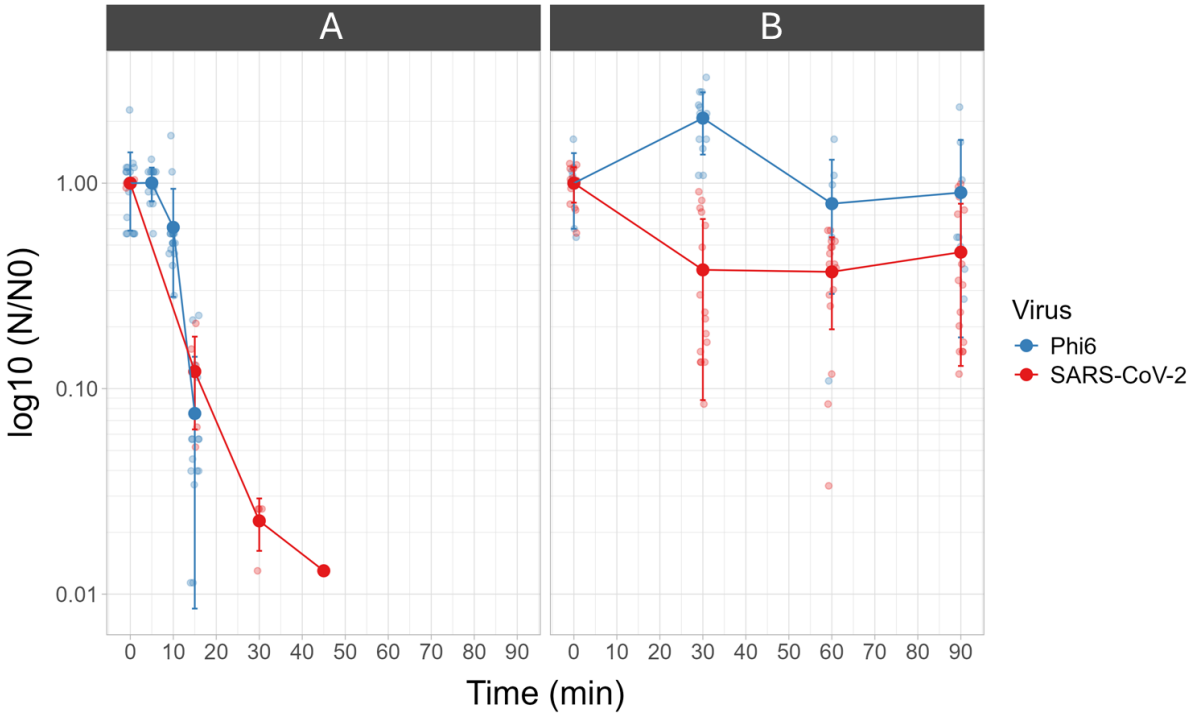


Figure 17. The effect of the 3D printed surfaces on Phi6 and SARS-CoV-2 B.3 over the course of 90 min. **A** The results of viral inactivation on material GFF14 (PLA with copper microparticles). **B** The results of viral inactivation on material GFF8 (PLA with copper nanoparticles). The infectivity is calculated as the titer at each time point (N) divided by the starting titer (N0). For clarity, the data is represented at a $\log_{10}$ scale. The enlarged solid points represent mean values, small dots the raw data, while error bars represent standard error over three biological and three technical replicates. GFF14 showed a much greater reduction in the infectivity effect for both Phi6 and SARS-CoV-2.

Photos of samples of the surfaces GFF8 and GFF14 are shown in Figure 18 together with the SEM images of their surfaces together with steel and copper surface SEM micrographs are shown. As shown, they are mostly plain with minor aberrations. GFF14 and GFF8 showed high roughness as visible from the SEM images. The exact roughness measurements are shown in Table 9. It is visible that the surfaces differ in their roughness significantly.

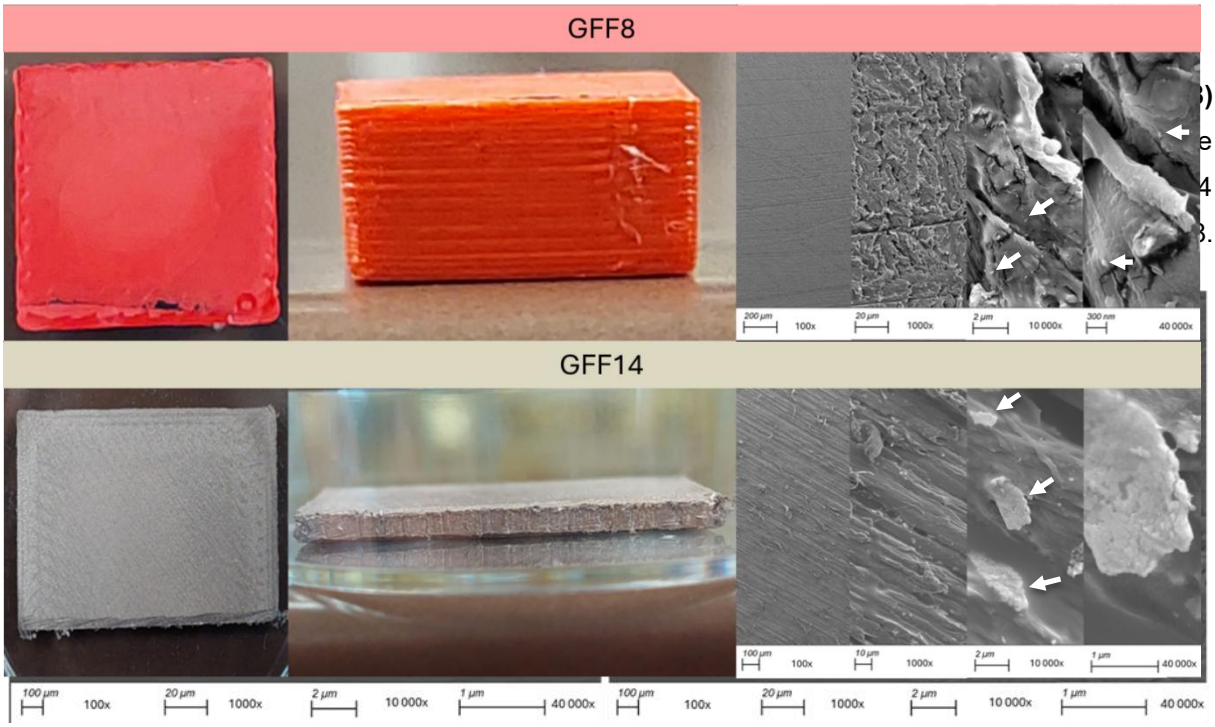


Figure 18. The Scanning electron microscopy images of the polished copper and steel surfaces. Both show smooth features with roughness visible only at higher magnifications. The arrows indicate copper nanoparticles and microparticles respectively.

Table 9. The roughness of the surfaces tested in this research. Two pieces of each GFF surface were measured. The mean values are shown out of 3 measurements. The distance over which the measurement was performed was 2.3 mm in length. $\mathbf{R_{max}}$ is the highest difference between a hill and a valley. $\mathbf{R_z}$ is a mean roughness depth, meaning the average maxima between hills and valleys over the scanned area. $\mathbf{R_a}$ is a mean distance of the surface profile from the mean line.

Material	*Roughness (R_{max}, µm)*	*R_z, µm*	*R_a, µm*
GFF14	25.24	13.41	2.030
GFF14	14.56	9.17	1.554
GFF8	3.25	2.35	0.373
GFF8	3.25	2.62	0.399
Copper	0.22	0.14	0.027
Steel	0.63	0.31	0.129

2.4. Discussion

2.4.1. SARS-CoV-2 infectivity

Compared to all the variants of SARS-CoV-2, Phi6 showed a greater reduction of infectivity after the treatment on copper. The experiments performed in this study show that when it comes to the efficiency of copper as an antiviral material, Phi6 does not seem to be a suitable surrogate for SARS-CoV-2. This is potentially due to the molecular differences between human/animal viruses such as SARS-CoV-2 or MPXV and a bacteriophage Phi6. There is some evidence that the starting titter of the virus is important to consider in experiments. Particularly for Phi6, Bangiyev et al. showed that the higher titer of Phi6 protects the virus from environmental decay (Ronald *et al.*, 2021). Phage Phi6 can achieve significantly higher titers during its growth in culture (in the range of 10^9 PFU/mL) than SARS-CoV-2 or mpox (in the range of 10^6 PFU/mL). Therefore, we ensured that the starting titers of the viruses were comparable by diluting Phi6 to a similar range as SARS-CoV-2 and MPXV. Additionally, the surface area of the virions might play a role in the case of surface saturation. Because larger virions will saturate the surface at lower titers than smaller ones. Therefore, it could be expected that when a copper surface is saturated, viruses with larger virions might demonstrate a slower drop in infectivity. Additionally, a significant aspect that distinguishes Phi6 from the other tested viruses is glycosylation of surface proteins. Since bacteria do not possess the ER or Golgi apparatus where glycosylation of proteins occurs, glycosylation in general differs significantly between phages and animal viruses. Glycosylation in bacteria is rare and if present, it is restricted to outer membrane proteins. *Pseudomonas syringae*, the host of Phi6, indeed does glycosylate its proteins, but it is limited and most crucial for outer-membrane-bound flagellin (Kasumi *et al.*, 2003). The probability that Phi6 contains glycans on any of its proteins during life cycle is low, especially since it takes its envelope from the host inner (cytoplasmic) membrane. While host surface glycoproteins might be important for Phi6 to target its host, there is currently no evidence that the virus itself has glycoproteins on its surface.

The infectivity of Phi6 was reduced significantly faster on copper than on steel, showing that steel is good control, even for a phage like Phi6 which showed lower stability than other viruses.

The XBB1.5 variant showed surprisingly high resistance to copper compared to the other tested variants of SARS-CoV-2. The infectivity of Phi6 was reduced rather quickly on both steel and copper surfaces. Phi6 and SARS-CoV-2 have been shown to have similar diameters of 80-100 nm (Fedorenko *et al.*, 2020). However, there are significant genomic differences between Phi6 and SARS-CoV-2. While Phi6 is a double-stranded RNA virus, SARS-CoV-2 is

a single-stranded RNA virus. The main mechanism of copper´s antiviral activity is to produce reactive oxygen species, thereby damaging the viral surface and genome. In addition, copper can bind directly to proteins. Glycans on the protein surface, which Phi6 lacks, can help protect the protein itself from oxidative damage.

In nature, the virus will normally not find itself in PBS, but rather in respiratory aerosols. Those aerosols are made of saliva and lung fluid (Woo *et al.*, 2010). They are composed mainly of protein, water, and salts with traces of cell debris and lipids. In this research, we used a formula for artificial aerosol fluid developed by Woo et al. for research on viral spread. It is important to note that here the main protein is porcine mucin, while in human saliva the main proteins are degrading enzymes, among which are peptidases. Due to their activity, saliva contains a high content of low-molecular-weight proteins (Loo *et al.*, 2010). Those peptidases could potentially affect viral stability in saliva and aerosols. Even though the mucin allows for realistic viscosity of the saliva, it does not possess a degrading activity. Additionally, mucin is a high-molecular-weight protein. This means that the artificial saliva used here contains a low diversity of high molecular weight proteins, while the real human saliva contains highly diverse low molecular weight proteins. This is a detail that could be explored further in studies with saliva of human or animal origin. Nevertheless, it is unlikely that the chemical composition plays a key role as saliva is composed of 99.5% water and all other components make 0.5%. Results consistent with that claim were observed here. Two different SARS-CoV-2 variants were tested in artificial saliva and PBS. The reduction in infectivity was similar for both strains of SARS-CoV-2 in PBS and the artificial saliva. This indicates no effect of the saliva on the infectivity of virions. However, it will be of interest to experiment on human or animal saliva.

2.4.2. Spike protein glycosylation and electrostatic charge

Glycans might play a role in changing the surface electrostatic potential. Since copper cations (Cu^{2+}) interacting with negatively charged protein structures and nucleic acids are among the main ways that copper inactivates biological agents, electrostatic potential distribution of the phage surface is worth considering the effect of copper on the virus. Research suggests that copper binds in the proteins to histidine clusters, thiol groups (from cysteine), and acidic amino acid side chains (from aspartate and glutamate). The structure of the P3 phage receptor protein homotrimer was not available in PDB. For SARS-CoV-2, all three tested strains had available spike protein structures in PDB. Figure 10 shows the spike protein in all three strains with highlighted potential Cu^{2+} binding sites. In the 3D structure analysis, no free thiol groups were found as all of them formed disulfide bonds. Also, histidine clusters were not found in any of the structures. However, aspartate and glutamate sites were found in all the strains in different numbers.

One possible explanation is that the slower reduction of infectivity of XBB.1.5 compared to B.3 is due to the difference in glycosylation profile and possibly copper binding sites. The spike viral receptor is heavily glycosylated. Spike protein of SARS-CoV-2 has 22 potential N- and 17 potential O-glycosylation sites (Aloor *et al.*, 2022). Recent evidence shows that even though all variants have conserved glycosylation sites in their Spike proteins, the N- and O-glycosylation profiles change between variants (Shajahan *et al.*, 2023). Notably, the more recent variants such as omicron showed a reduced abundance of complex-type glycans and an overall increase in oligomannosidic-type glycans. Those differences stem from the variations in primary, secondary, and tertiary structure of the spike monomers between variants. It is unknown how this could affect environmental stability, but it is worth noting that glycosylation represents a difference between variants on the surface of virions which might play a role in environmental stability. Further experiments are needed to confirm this hypothesis.

The distribution of the potential is informative while plotted on the graph, but still, it is useful to compare it on the proteins in 3D to observe which sites might electrically attract copper cations. While the acidic amino acids pointed out in Figure 10 can act as spots for copper binding, additionally, the spikes with more negative electrostatic potential at the given pH (here 7.5), have a higher probability of interaction with copper cations through electrostatic interactions. Figure 11 compares all the Spike proteins electrostatic potentials at pH 7.5, as calculated by Adaptive Poisson-Boltzmann Solver (APBS) and modeled with CHARMM36.

XBB.1.5 has exposed lower amount of acidic amino acids to the surface. This could mean less possibility of copper binding to the protein. Those findings also go in line with the results displayed in the histogram of charges over the whole structure. The XBB.1.5 variant has the most positive net charge, meaning it would have the least affinity towards copper cations. The Spike protein of each strain showed slight differences.

2.4.3. SARS-CoV-2 genome damage

The result of no significant differences in qPCR shows there was no significand genomic RNA damage in the region that was amplified (S gene). However, it does not exclude the possibility of genomic damage outside that region. Therefore, the limitation of the use of qPCR to assess the genomic damage is that it is limited to a specific genomic region. Long-range RT-qPCR could be used to cover the wider genomic regions but still cannot detect the damage in all regions. Mass spectrometry could also be used to assess the damage of the genome, however it requires high amounts of DNA/RNA which is not always achievable with viral samples. Still, even though genome damage cannot be completely excluded as a possibility, the results of qPCR assays show that at least in the S gene region of the protein, the damage is not significant.

2.4.4. Mpox infectivity after copper treatment

MPXV has a more complex structure than SARS-CoV-2. Due to the higher surface protein diversity and complexity, the MPXV might show less variation between the different clades tested. Still, differences could be observed. Clade Ia showed the highest resistance to copper surface. It generally shows lower transmissibility, mortality and severity than others, and according to WHO, as of April 2025, only three reported cases were recorded in 2025 outside Africa – two in China and one in Ireland (WHO, 2025a). Ia regularly circulates inside central Africa and shows spillovers from animals to humans. Nevertheless, it is important to know that this clade of MPXV might be more resistant to copper. Of the four tested clades, clade Ib showed the lowest resistance, exhibiting almost complete inactivation after 15 minutes. IIa and IIb showed similar and intermediate resistance to copper.

2.4.5. MPXV genome assemblies

The explanation for different resistances to copper might stem from the structures and electrostatic potentials of the surface glycoproteins, similar to SARS-CoV-2. Even though the 15-minute time point for clade IIa was not enough to cause inactivation, it was enough to observe the differences between various clades. Depending on the application, those time points can or don't have to be realistic. It depends on how long an individual can stay in potential contact with a contaminated surface. For example, in the public health settings where people spend long times, the time points of hours or more should be considered. For the public transportation settings, minutes to hours are considered enough because humans don't stay long in those settings. To cover a wide range of applications, the time points beyond 15 minutes can be tested in the future to determine whether this difference is significant when longer time is considered. However, antimicrobial surfaces are of most use in the areas that are touched frequently. Therefore, exploring the stability of MPXV strains up to 15 minutes has more relevance than the longer time points. The variation in the glycan profile among the clades could also play a role in resistance to copper inactivation.

While MPXV IIa did not show any sequence-altering (i.e. missense/nonsense) mutations in the transmembrane proteins, all other strains showed them in the gene OPG210, with the product named gp173. It is a big protein of 1880 amino acids with multiple transmembrane domains and many glycosylation signals. N-glycosylation signal is N-X-S/T (Lowenthal *et al.*, 2016). Among other transmembrane genes that contained such mutations, none was on a potential glycosylation site. On the other hand, MPXV Ia showed A98D (alanine-98 to aspartic acid) that is near a strong candidate N-glycosylation site (Asparagine-94). Therefore, the sequence of the N-glycosylation signal **NYT**IA which becomes **NYT**ID. Since aspartic acid is a negatively charged amino acid at pH 7.5 while alanine is neutral, this mutation changes the local electric potential. In MPXV Ib, there were three mutations in the gene OPG210, of which D209N was

directly next to an N-glycosylated asparagine-210. Therefore, the sequence changes from D**NNT** to N**NNT**. Again, this changes the electrostatic local environment, rising the local potential. Even though this mutation is not directly part of the glycosylation signal, it could have an effect on the glycosylation efficacy. Another mutation present in MPXV Ib and IIb is M1741I which is near a potential N-glycosylation site of lower certainty (asparagine-1737). Interestingly, this same mutation appears in both sub-clades of clade b but in none of the clade a. It changes the sequence from **NDT**YM to **NDT**YI.

Overall, no SNPs were found that would affect a glycosylation motif sequence. However, there were three missense mutations immediately close to the motif. Out of those three, two change the local electrostatic potential. Interestingly, even though Ib showed the lowest copper resistance, it also has the highest number of non-silent mutations in OPG210. The main interesting mutation might be D209N which raises the local electrostatic potential. Therefore, it could be that it affects the glycosylation process, potentially reducing the prevenance of N-glycan there. However, this would still require experimental confirmation. OPG210 is a surface glycoprotein of MPXV and is known to evolve rapidly (Wang *et al.*, 2023). Therefore, it is not surprising to see it mutated in three out of four strains. This is a gene that is believed to disrupt T-cell response through yet unknown mechanism and some mutations show human host preference (Brayden *et al.*, 2024). Still, as a surface membrane glycoprotein, it is probable it affects the resistance and stability of the virus in environment.

No missense/nonsense mutations were found near or within O-glycosylation sites. However, N-glycosylation is much more complex though less prevalent. Still, N-glycosylation could play a major role in the protein stability and interaction with other proteins.

2.4.6. Mpox qPCR detection

To determine whether there is any DNA damage that occurred as the result of copper treatment, qPCR was performed on all the viruses tested. As visible in Figure 16, in most of the clades, the Ct values do not differ among the starting timepoint, copper, or steel. The only exception to this is in clade IIb where both copper-treated and steel-treated samples show significantly higher Ct value than those at the starting time point. However, steel and copper do not differ among each other. This means that some DNA damage could be present after in this clade, but then it would not be unique to copper. This is surprising since the inactivation results show the viruses were inactivated but to lesser extent than Ia. The argument that there might be other genomic damage in the places not amplified by qPCR stands, as is the case for SARS-CoV-2. Therefore, in the highly inactivated clade such as Ia, it might be the case that the genomic damage is present mostly in other genomic regions while in the case of IIb the damage was found in that region. However, to confirm this, it is required to make a wider genome damage assessment.

2.4.7. 3D-printed antimicrobial surface testing

Copper in its plain form is not suitable for use as an antimicrobial in aircraft, as it is a heavy and costly metal. Here we explored the antiviral activity of lower-cost and lighter 3D-printed materials incorporating copper particles. These materials could be antimicrobial materials suitable for use in aircraft. Thus, we compared the stability of Phi6 and the ancestral variant of SARS-CoV-2 (B.3.) against such surfaces containing copper particles (GFF8 and GFF14). The 3D-printed GFF14, which had incorporated copper micro-particles, had a significant antimicrobial effect against an ancestral variant of SARS-CoV-2 and even higher on Phi6, even in the scale of hours (Figure 17). This makes it a suitable candidate for potential use in aircraft, when considering the antimicrobial activity, as it acts quickly. This could not be shown for GFF8, which contains copper nanoparticles, where both Phi6 and SARS-CoV-2 B.3 stably remained for 90 min without significant reduction in titer. Even though the precise explanation has not been experimentally explored here, we propose potential explanations, which remain to be tested in the future.

As seen in Table 9, GFF14 showed the highest degree of roughness, as also visible in the scanning electron microscopy images (Figure 18). This affirms the notion that in addition to the antiviral effect of copper on the surfaces, an additional effect could be the trapping of the virus particles in the topography of the surfaces. This would make some of the viruses unavailable for sampling. However, it is assumed that a virus will move together with the water matrix, as its surface proteins form hydrogen bonds with water. Therefore, it is not likely that a virus would get immobilized as a consequence of the surface topography. This means that among many viruses that were present on the surface, a few would remain non-sampled. Even if the number of viruses that are non-sampled would be significant, the fact of not being sampled makes them unavailable for infection in the real world. Therefore, the roughness of the surface could reduce the spread of a virus, adding to its antimicrobial effect. Additionally, the higher roughness of the surface increases the contact surface area between virions and the surface. This allows for a rougher surface to be a better antimicrobial agent than a smoother surface (Georgakopoulos-Soares *et al.*, 2023). This would add to the explanation the higher antiviral activity of GFF14 compared to GFF8.

2.5. Conclusion

Different MPXV and SARS-CoV-2 strains, together with phage Phi6, were treated on copper, and steel. Also, variant B.3 of SARS-CoV-2 and Phi6 were also treated on 3D-printed copper nano- and micro-particle containing surfaces for up to 15 minutes . They were then plated for PFU assays to determine their infectivity. Additionally, since the 3D structures of the SARS-CoV-2 were available, they were compared against each other, considering their electric

potential distribution and potential copper-binding sites. For MPXV, no such structures were available so genomic DNA was sequenced and genome assembled, and variations were screened among the surface glycoproteins.

The three SARS-CoV-2 variants tested showed varying resistance to copper. Specifically, the variant closely related to the Wuhan ancestor, B.3, showed the lowest resistance to copper, while variant B.1.617.2 (“Delta”) showed a higher copper resistance and XBB.1.5 was the most resistant variant to copper. Even though only three variants were tested, this trend suggests that more resistance was developed over SARS-CoV-2 evolution. The Spike receptor protein allows the recognition of the target cell. It was hypothesized that the exposure of the copper-binding sites, the electrostatic potential distribution on the surface of the Spike receptor protein, and it’s glycosylation sites might play a role in this resistance. Indeed, the XBB.1.5 variant had the highest average Spike protein electrostatic potential and hidden copper-binding sites (acidic amino acids). The glycosylation sites of the Spike protein were not bound to significant changes in electrostatic potential of the Spike protein among SARS-CoV-2 variants. However, the surface glycans generally can shield the protein from the external ions. Additionally, the viruses were detected with qRT-PCR after treatment on copper and steel surfaces. The qRT-PCR quantification showed no differences between the copper-treated samples and steel or the non-treated control which had its genome directly isolated and qRT-PCR performed. Therefore, the detectability on copper and steel both was the same even though infectivity was not. This points to the limited genome damage in SARS-CoV-2 at last in the region amplified by PCR (Spike gene).

For MPXV, there was also variation in the resistance between different clades. While clade Ia showed high resistance to copper, clades IIa and IIb showed intermediate copper resistance and Ib the lowest resistance. The differences in all the clades were explored after assembly of their genomes and searching for variations. The variations in the membrane protein genes were isolated and scanned for glycosylation sites. There were several transmembrane protein-coding genes with mutations (Ia: OPG172, OPG210; Ib: OPG053, OPG139, OPG210; IIb: OPG019, OPG053, OPG139, OPG143, OPG210). Significantly, three mutations were found near an N-glycosylation site, but no mutations near O-glycosylation sites. The qRT-PCR quantification showed for almost all clades that the Ct-value did not differ between copper and steel, even not from the non-treated control. However, clade IIb showed slight but statistically significant differences in the Ct-values, which were higher in steel- and copper-treated samples than the non-treated control. This is in line with the infectivity results as IIb had the lowest resistance to copper and reduction was even visible on steel after 15 minutes. Phi6 had significantly lower resistance than SARS-CoV-2, showing it is not a suitable surrogate when it comes to copper resistance.

Finally, SARS-CoV-2 B.3 and Phi6 were tested on 3D-printed copper-based antimicrobial surfaces, which are suitable for use in aircraft due to their light weight and inexpensiveness. Overall, the surface with copper microparticles showed better efficacy in inactivation of the SARS-CoV-2 and Phi6 than the surface with copper nanoparticles. The surfaces containing copper microparticles showed themselves as a good antimicrobial, inactivating roughly 1 $\log_{10}$ (90%) of SARS-CoV-2 virions in less than one hour. This brings it a step closer to the use in aircraft.

Chapter 3: Prophage diversity in spaceflight

3.1. Introduction

Microorganisms are traveling on humans and reaching all environments that humans reach, and even further. The microorganisms in our habitats therefore form complex communities, made of viruses, bacteria, archaea, and eukaryotic microorganisms. As the most numerous biological entities on Earth, bacteriophage diversity is interesting to explore in spaceflight environment. In addition to controlling and shaping bacterial populations, bacteriophages carry and transfer bacterial genes between genomes. Some of those genes can cause antimicrobial resistance (AMR). Therefore, bacteriophages can spread various genes, including AMR through bacterial populations. That is especially important when the phages jump between species and a new bacterial species, with a different set of genes, now possesses genes coming from another species. For bacteria to be able to utilize those genes, a phage should be integrated into its genome so that the characteristic passes onto new generations. Those phages, integrated into the host genome, are called prophages. Their genomes are therefore sequenced together with the bacterial genomes. Space stations are isolated environments with specific set of conditions such as radiation and microgravity. Also, the welfare of the crew is of high importance during missions. Therefore, it is of interest to understand the viral prophage diversity of International Space Station as an example to learn for the future space stations.

For the detection of microorganisms, many methods exist, though not many are suitable for direct detection on-site. The methods can be divided into nucleic acid-based, protein-based, and whole organism-based. Of all of those methods, nucleic acid-based ones are among the most sensitive. Nucleic acid-based methods include qPCR, sequencing, and loop-mediated isothermal amplification (LAMP). Normally, all those methods require laboratory conditions, and highly purified DNA to be performed. As sensitive and specific as they are, their weakness in environmental detection is that it is impossible to say if the microorganism is infectious or not, as only its genes are detected. Still, this can also be a strength of the nucleic acid-based tests, as they do not necessarily need to detect live microorganisms, but also can be used to detect genes, like antibiotic resistance genes. Most of the nucleic acid-based methods are targeted, therefore they cannot detect novel pathogens. However, DNA sequencing is non-targeted and therefore can detect unknown microorganisms. At the same time, qPCR and LAMP are targeted and can only detect pre-determined microorganisms. The protein-based detection methods are Enzyme-linked immunosorbent assay (ELISA) and mass spectrometry. They can detect proteins belonging to specific microorganisms. ELISA is based on the highly specific reaction of an antibody to the target protein antigen. In the case of detection of the binding of the antibody, fluorescence is observed, as the fluorescently labeled antibody will not be washed away. This signifies the presence of the target antigen, and therefore, the pathogen.

Another method of detection is cultivation, which takes time and resources. Another possibility is matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS). There the pathogens are directly subjected to mass spectrometry, where the pathogens are determined based on the ratio of the mass and charge. Each pathogen will have its unique pattern and can therefore be determined from the databases. As a mass spectrometry method, it is highly costly and not easily available to research/diagnostic labs. Therefore, it is not a convenient method in most cases. It is also not convenient for *in situ* detection of microbes as the mass spectrometer is not portable, especially not in public transport. Additionally, it requires a fully-grown bacterial culture which also requires time.

The methods that apply to public transportation should be portable, quick, and simple to use. LAMP may be a method of choice for viral detection on-site. Another possibility would be a portable qPCR system has been recently developed for in situ testing (Xun *et al.*, 2021; Zowawi *et al.*, 2021; P. Q. M. Nguyen *et al.*, 2022; Song *et al.*, 2023). However, those types of systems are used for diagnostics and food control and usually require sample purification. Compared to environmental samples, diagnostic samples have significantly more biomass, and therefore, the results can be more conclusive. With the high sensitivity of LAMP and qPCR, it is feasible to use them for environmental detection of viruses (Klymus *et al.*, 2020).

In spaceflight, it is challenging to track microbial and viral contaminations due to all the restrictions associated with space travel. The samples need time to get returned to Earth which changes them biologically. Also, the resources and time available for experiments on space stations are limited. Cultivation of microorganisms is available in spaceflight, but time and resources are limited for such assays. Molecular detection is a promising way to detect and track different microbial contaminations. Currently, the qPCR system has been developed to work in microgravity (Parra *et al.*, 2017). Though qPCR is a sensitive and specific method for microbial detection, it requires highly pure samples, and it is targeted. This means that the list of the contaminants to be tracked needs to be pre-determined, allowing the unknown or unaccounted-for contaminants to slip undetected. Nanopore sequencing has been used in spaceflight to sequence full genomes of mouse, *E. coli*, and bacteriophage λ (Castro-Wallace *et al.*, 2017). However, in that research, the DNA was pre-prepared in the laboratory, just to be sent to ISS for sequencing. Microbial monitoring is still being done on the ground after the samples have been returned to Earth. Since the ISS resupply and return happens on average 5-8 times per year, depending on the agencies and re-suppliers included, the microbial monitoring of the ISS is far from real-time. The sample return can happen on average every several months, which is a lot of time for microbial contamination to cause disease. There is research that shows the bacterial microbiome of the ISS surfaces shows similarities to an average household, with some significant differences, such as heavily being comprised of

human-associated bacteria (Lang *et al.*, 2017). The limitation of the study was that it was based on 16S sequencing data, missing the eukaryotic and viral parts of the microbiome. The viral part of the ISS microbiome is not heavily studied.

Still, the virome of the space stations remains to be explored. Here we are trying to close the gap by utilizing public databases of nucleic acids such as the NCBI Database. The NCBI GenBank offers a digital collection of nucleic acid sequences of hundreds of thousands of projects, among which are the sequencing projects of the isolates of the ISS. Among the bacterial genomes from the ISS, many of them have incorporated temperate bacteriophages. Those viruses can carry genes inside them that, once the phage reactivates, can be carried between the host bacteria. This is especially inconvenient in the case of antimicrobial resistance genes. This is how the temperate phages spread resistance in microbial communities. Therefore, it is interesting to understand the viruses available in the genomes of the bacteria isolated from the ISS. This information is available in public databases such as GenBank, but it is not a representation of all the viruses there, as only the isolates are considered. Still, provided that the viral genes inside the bacterial genomes can be recognized, and assessed whether the virus genome is complete, there could be an assessment of the diversity of potentially active phages and the assessment of the antimicrobial resistance genes that could be transmitted.

In this research, the genomic sequences of all the bacteria isolated from the International Space Station were collected and analyzed for the presence of viral genes. The prophages inside the bacterial genomes of the space stations could potentially be activated to a higher level than on Earth, due to various stress conditions such as ionizing radiation and microgravity. Normally, prophages reactivate inside the bacteria that are in a stressful environment. The research from Irby and Broddrick 2024 suggests that bacterial adaptation to spaceflight is mediated through bacteriophages inside the bacterial genomes (Irby and Broddrick, 2024). This shows the importance of understanding the diversity of prophages in the spaceflight environment.

The prophages inside the bacterial genomes were collected and analyzed together. The search was performed for any antibiotic resistance genes within the identified prophage regions.

3.2. Materials and Methods

3.2.1. Downloading the bacterial isolate genomes and prophage detection

The GenBank and RefSeq databases were searched with the keywords: “International space station” AND “bacteria” [organism]. Only full genomes were selected. The list of obtained

genomes is included in the supplementary information. The full genomes were analyzed for the presence of bacteriophage genes using the PHASTER API tool utilized through a UNIX-based command line. The UNIX-based command line was accessed on Windows through the software MobaXterm Personal Edition v20 Build 4286. Briefly, the genome sequences were downloaded in the same directory and extracted to get the FATSA files (text format) containing sequences. The sequences were then split into two folders based on the number of entries in each file. They were divided into single-entry genomes and multi-entry genomes. Single-entry FASTA files contain only one sequence which is a full-length genome. The multi-entry FASTA files contain contig-level assembled genomes and genomes including plasmid sequences. This categorization was performed because the PHASTER algorithm needs to note whether it is analyzing a single entry or multiple entries. The logic of this categorization is represented in Figure 19.

PHASTER algorithm scans the provided FASTA nucleotide sequence file and identifies the protein coding genes present in it. It then searches the database of bacteriophage protein sequences consisting of 420,000 entries and a bacterial protein database made of 16,000,000 entries. It maps the phage genes onto the bacterial genome and searches for the regions containing clusters of the mapped phage genes. Then it scores the identified regions based on the phage gene content. Briefly, if 100% of the genes in the cluster region can be assigned to the same bacteriophage, the maximum score is assigned of 150 and the region is termed “intact”. If a single phage can be assigned more than 50%, but less than 100% of the genes in the identified cluster region, the percentage of the genes that identify with that phage are multiplied by 100 and the total length percentage of the phage in relation to the total region length is multiplied by 50 and added to the total score. That value is noted as the first score. Then, PHASTER searches for the keywords related to the phage genes: 'capsid', 'head', 'integrase', 'plate', 'tail', 'fiber', 'coat', 'transposase', 'portal', 'terminase', 'protease' or 'lysin'. For each keyword found, the score is increased by 10. Also, in the cases of regions above 30 kb in length, the score is increased by 10. If there are at least 40 phage proteins in the region, the score is increased by 10. Also, if all phage proteins consist more than 70% of the region´s genes, the score will be increased by 10. That score is noted as the second score. The larger of the two scores is then evaluated. Scores below 70 are termed incomplete, between 70 and 90 is questionable, and above 90 are complete phage genomes. The full genomes are also searched against a bacterial genome database to assign the bacterial genes that might also find themselves inside phage regions. Since PHASTER is primarily an online tool for detection of prophages, it is best suited to small amount of genome submissions, as they need to be submitted manually. However, PHASTER has an Application Programming Interface (API) through which big batches of genomic sequences could be submitted. Since this process is

not obvious, I describe it here together with the bash script code provided which could be reproduced on any machine.

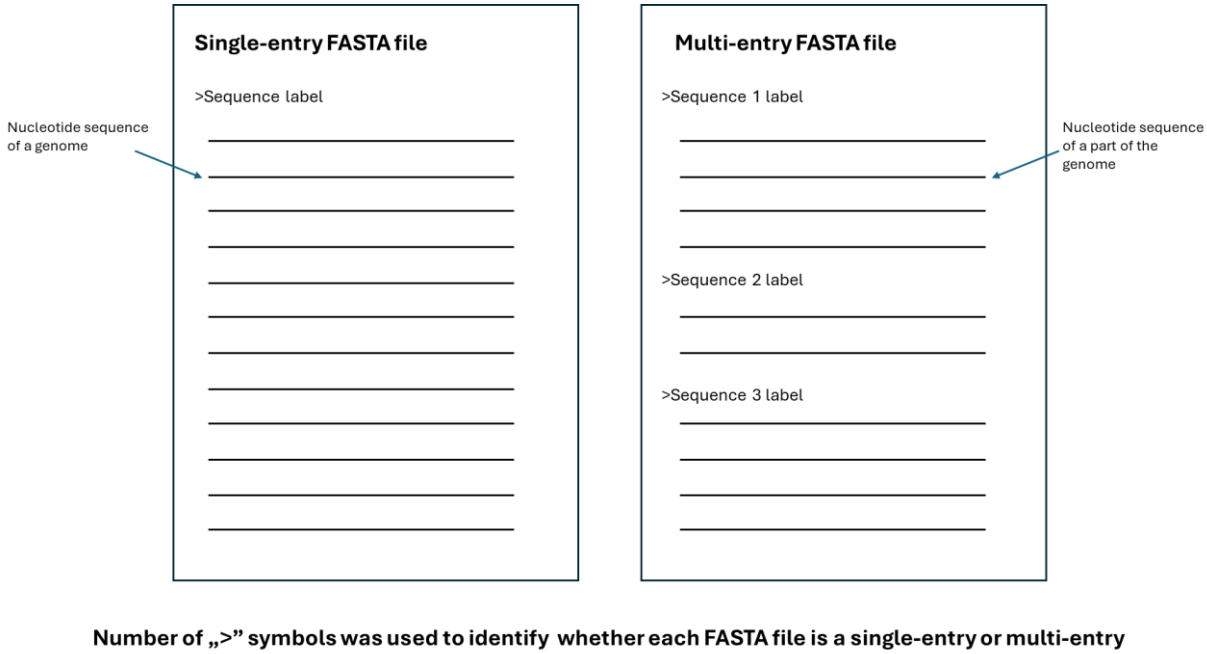


Figure 19. Schematic representation of the logic behind the code for categorization of the FASTA files. All contig-level sequences are represented as multi-entry FASTA files. However, some complete genomes containing plasmid(s) are also represented as multi-entry FASTA files.

The categorization of the FASTA files was performed by executing a bash script containing the following code:

```
#!/bin/bash

# Create the directories
mkdir -p Single-entry Multi-entry

# Loop through all .fna files in the current directory
for file in *.fna; do
    if [ -f "$file" ]; then # Check if the file exists
        # Count the number of '>' symbols in the file
        count=$(grep -o ">" "$file" | wc -l)

        # Move the file based on the count
        if [ "$count" -eq 1 ]; then
            mv "$file" Single-entry/
        elif [ "$count" -gt 1 ]; then
```

```
            mv "$file" Multi-entry/
        fi
    fi
done

echo "FASTA categorization done."
```

This code determined whether a FASTA file, a text file containing the genomic sequence, contains one or multiple “>” symbols, signifying the label of a sequence. The files with multiple “>” symbols are multi-entry genomes and those with one “>” are the full genomes. This division was necessary because the PHASTER algorithm requires specifying the level of genome assembly before the analysis.

Once the sequences were divided by the assembly level, in each folder, all the sequences were submitted to the PHASTER server. The single-entry sequences were submitted by executing a script containing the following bash code, from the directory containing the single-entry FASTA files:

```
# Loop through all .fna files in the current directory
for FILE in *.fna; do
    echo "Processing file: $FILE"

    # Send the file to the API and append the response along with
the filename to Status.txt
    wget --post-file="$FILE" "http://phaster.ca/phaster_api" -O -
>> Status.txt
    echo " -- File: $FILE" >> Status.txt

    echo "File $FILE processed. Waiting for 10 seconds before the
next request."
    sleep 10
done

echo "All files processed. Results saved in Status.txt."
```

For the multi-entry files, the same script was executed but from the multi-entry FASTA file directory, and with the difference in the URL to submit to the PHASTER API: http://phaster.ca/phaster_api?contig=1. With the addition of “?contig=1” to the URL, it was

specified that the file would contain multiple entries. It is to be noted that despite the label “contig” used by PHASTER, this does not necessarily mean the genomes are assembled to the contig level. Some genomes contain plasmids which can also be shown as additional sequences in the file together with the nucleoid genome. Other genomes are indeed assembled to the contig level. All of them are categorized as a multi-entry FASTA file. In the script, 10 seconds of “sleep” time was included to enable *wget* full establishment of the connection to the server before proceeding with the command. Otherwise, the script would result in an error.

After each job was submitted, the specific ID was assigned automatically by the system to each submitted sequence. The IDs are in the format ZZ_ followed by ten alphanumerical characters (for example ZZ_a1b2c3d4ef). The analysis was based on the server so the waiting time was prolonged.

To later fetch the results of the analysis of each genome, this ID needed to be used as a reference. The output of the code above created a Status.txt file with the message in the format "*job_id": ZZ_a1b2c3d4ef,"status":"You're next!."* for each sequence submitted. The code above also appends the name of the file submitted beneath the message. Therefore, the job IDs could be connected to the sequence that was analyzed. In the end, the directory contained the collection of FASTA files with genomic sequences of bacteria, and a Status file containing the success messages with the job IDs, and under each job ID, the FASTA file name was appended to later identify each job ID. Therefore, the Status.txt file was used as the reference to connect each species with its respective job ID. Figure 20. schematically represents the Status.txt file.

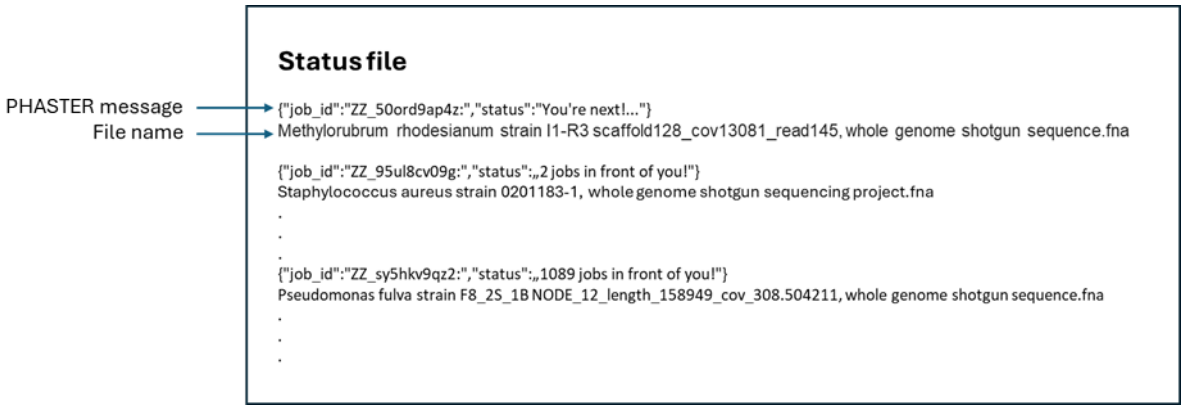


Figure 20. Schematic representation of the Status file that incorporated the message from PHASTER and each FASTA file name underneath it.

The list of all IDs was created by including the IDs and the species names. The following bash script was executed in the UNIX command line:

```
#!/bin/bash

# Output CSV file
output_file="ID_names.csv"

# Initialize the CSV file with headers
echo "jobID,name" > "$output_file"

# Read the Status.txt file and process it
while IFS= read -r line; do
    # Check if the line contains a valid jobID (ZZ_ followed by 10
alphanumeric characters)
    if [[ $line =~ ^(ZZ_[a-z0-9]{10})$ ]]; then
        jobID="${BASH_REMATCH[1]}"

        # Read the next line for the "name"
        IFS= read -r name_line

        # Append to the CSV file
        echo "$jobID,$name_line" >> "$output_file"
    fi
done < Status.txt

echo "CSV file '$output_file' has been created."
```

The code above extracts all the job IDs and makes a table with a job ID and the file name.

After approximately seven days, the analysis was complete. The results of the analysis are fetched from the PHASTER server by downloading them from the link phaster.ca/submissions/jobID, where jobID is the job ID assigned to each genome with the submission. To fetch the results associated with each ID, the following command was used in the UNIX command line:

```
#!/bin/bash

# Input CSV file
input_file="ID_names.csv"
```

```
# Ensure the input file exists
if [[ ! -f "$input_file" ]]; then
    echo "Error: $input_file not found!"
    exit 1
fi

# Read the CSV file, skipping the header
tail -n +2 "$input_file" | while IFS=',' read -r jobID name; do
    # Clean up the name to avoid invalid characters in filenames
    sanitized_name=$(echo "$name" | tr -d '[:space:]' | tr -cd 'A-
Za-z0-9_-')

    # Construct the download URL and output file name
    url="http://phaster.ca/submissions/${jobID}.zip"
    output_file="Output_${sanitized_name}.zip"

    # Download the file
    echo "Downloading $url as $output_file..."
    wget -q "$url" -O "$output_file"

    # Check if the download was successful
    if [[ $? -eq 0 ]]; then
        echo "Downloaded: $output_file"
    else
        echo "Failed to download: $url"
        rm -f "$output_file" # Remove partially downloaded file if
any
    fi
done

echo "Download complete."
```

The script above queries the URL in the format http://phaster.ca/submissions/${jobID}.zip where ${jobID} is replaced by the job ID for each genome queried. For each ID, the output was named according to the genome FASTA file so that the results are saved as FASTAname.zip. The IDs and names are read from the CSV table generated in the previous script. The results of each analysis are contained within a .zip file with all the phage DNA sequences identified,

their location within the genome with the scores of certainties if the phage is active, inactive, or questionable, and the list of genes identified.

Descriptive statistics were performed on phage genomes found within the bacterial genomic sequences. The distribution of active, questionable, and inactive phages was determined.

3.2.2. Annotation of prophage genomes

The phage genome sequences identified inside the bacterial isolates were uploaded to the *Galaxy* public server (Boekel *et al.*, 2015) and annotated using Pharokka (v1.3.2) (Bouras *et al.*, 2023). Pharokka identifies the coding sequences (CDS) inside any provided DNA sequence using PHANOTATE (McNair *et al.*, 2019). It translates the CDS and compares to the reference bacteriophage proteomic database – PHROG (Terzian *et al.*, 2021) which contains 868,340 phage protein sequences. It performs the search with HMMER v3.4 algorithm (Potter *et al.*, 2018; Larralde, 2022). The threshold e-value used to identify a gene as a homologous was 10^{-5}. Based on the homology to the genes in the database, Pharokka categorizes the CDS according to function. It additionally determines any genes associated with antimicrobial resistance or virulence by searching against Comprehensive Antimicrobial Resistance Database (CARD) (Alcock *et al.*, 2020) and Virulence Factor Database (VFDB) (Liu *et al.*, 2019). It also uses tRNAscan-SE 2.0 (Chan *et al.*, 2021) to find tRNA genes. Additionally, to find transfer messenger RNA (tmRNA) genes, it uses Aragorn (Laslett and Canback, 2004). Therefore, the genetic diversity of the prophage sequences could be determined. The global genetic variation was visualized by plotting a heatmap of phages and their predicted gene functions.

A heatmap was produced, all phages with tRNA genes were categorized according to their host genus. To avoid the bias resulting from copies, phages were also categorized according to their genomic 95%-identity group. In this way, if the phages of any group appeared more than once, they would be considered copies and only one phage would be considered for the analysis. In this way, it was ensured that only the tRNA genes from distinct phages are shown.

Moreover, to determine whether the distribution of the phage groups among the bacterial hosts, the number of clusters determined by CD-HIT v4.8.1(Li and Godzik, 2006; Fu *et al.*, 2012) at 95% identity threshold were plotted on a per-species and per-genus level. As some of those species and genera were showing significantly higher number of clusters than the others, the next question that was explored is the distribution of those clusters among individual genomes within a species and also among individual species within genera.

3.2.3. Phage classification

After the CDS, tRNA, and tmRNA genes were determined for each phage sequence, the phages could be classified based on their gene content. This was performed using vContact2

v0.9.19 (Bin Jang *et al.*, 2019) on the KBase server v18 (Arkin *et al.*, 2018). The genetic content of each phage is compared to bacteriophage genomes in the reference database - Prokaryotic RefSeq with ICTV and NCBI Taxonomy v201, containing 3,500 curated phage genomes. The homology is determined using BLAST Protein (BLASTP) algorithm. The clustering is then performed based on the percentage of shared genes between any two phage genomes. That is performed among the query genomes as well as the reference genomes. Phages that share higher portions of genes are clustered closer together and those sharing lower portions are further away. Based on the grouping to the reference genomes, many of the identified phages could be classified into families and some even down to the genus level. The pie chart was created showing the distribution of the identified families. It must be noted that because the clustering algorithm was solely based on the shared gene content, and not the sequences themselves, many phage sequences still could not be identified because some of the phage genomes were fragmented, and some possessed a high level of genetic diversity.

To improve the classification of the identified phages, they were clustered according to 95% sequence identity using CD-HIT. This means that the sequences that had a minimum 95% identity to each other were clustered together. Since the nucleotide sequences were compared, the word-size used was 10. The advantage of CD-HIT is that it will compare even the sequences of different sizes and adjusts the identity to the shorter sequence. All the clusters that contained at least one bacteriophage sequence previously successfully classified with vContact2 using reference genomes were identified. Those clusters that additionally contained non-classified sequences were assigned to the family that any other sequence in the same cluster was assigned to. Due to the limited length of some sequences, the genus-level classification could not be resolved for the unclassified sequences, but the family-level classification was better resolved overall.

When the family classification was as accurate as possible, and genomes annotated, the large terminase subunit (terL) genes were used to construct phylogenetic relationships between prophages that contain at least one homolog of terL. The protein sequences of all terL genes identified were filtered from the whole dataset. Then, the terL sequences were separated into distinct FASTA files according to their bacteriophage family. For each family, terL sequences were aligned using MAFFT v7 (Kuraku *et al.*, 2013; Katoh, Rozewicki and Yamada, 2019). The strategy for the alignment used was G-INS-1, which is slower but more accurate while suitable for less than 2,000 sequences. The alignments were trimmed to remove the poorly aligned and non-informative regions using trimAl v1.5.rev0 (Capella-Gutiérrez, Silla-Martínez and Gabaldón, 2009). Option *automated1* was used to decide between three trimming models (strict, gappyout, and strictplus) based on the multiple alignment output from MAFFT. Then, based on the high-quality aligned regions, the phylogenies were constructed using IQ-TREE

v2.4.0 (Minh *et al.*, 2020). The program was used to find the best-fitting evolutionary model for each dataset, and then based on the selected model, the highest likelihood tree was generated with 1000 bootstrap reiterations.

3.2.4. Identifying inter- and intra-species prophage jumps

To determine if there was any horizontal transfer of phages between the host genera and species, the clusters previously generated by CD-HIT were utilized. Every phage sequence was part of a bacterial genomic sequence. The 95% identity clusters of phage sequences were used to determine the contigs that were going to be compared together. Those phages that were known to be similar or even identical were all found in different bacterial genomes. The question was asked whether the surrounding genetic environment is the same or different in those phages. If two phages have the same genetic environment, it is assumed they originated from the same dormant phage. If the surrounding sequences were significantly different, the phage was assumed to have jumped from one genome to another and integrated itself into a different locus in that other genome.

Therefore, the genomic contigs inside of which the phages were found, were compared using fastANI v1.34 (Jain *et al.*, 2018). Only the contigs with the phages within the same cluster were compared. This ensured that the contigs with highly similar bacteriophages are compared to determine their mutual identity. The contigs with low mutual identity but highly similar phages were determined to be potential signs of phage transfer between genomes. A python script running fastANI on all the sequences, comparing only within the same clusters and outputting an identity matrix of contigs for each cluster was prepared. Normally, fastANI is used to compare whole bacterial genomes by dividing the query sequence into chunks (called *words*) and comparing their sequences to the subject sequence. Based on the contig length distribution, the default settings were of fastANI were adapted to comparison of shorter sequences. Word size was set to 250 bp (compared to normal 3000) and the number of words to match to be considered a hit was set to 15 (instead of default 50).

3.2.5. Phage gene diversity determination

The analysis above confirmed that there were no inter-genera transfers of prophage sequences in the data (see *Results*). This means that the prophage diversity could be analyzed independently for each host genus. To explore the prophage gene diversity across the genera of bacterial ISS isolates, Roary v3.13.0 (Page *et al.*, 2015) was utilized. Similar to vContact2, Roary uses BLASTP with e-value threshold 10^{-5} to identify homologous proteins between the sequences. It builds a phylogenetic tree based on the shared gene contents of the input sequences and makes a tree based on the shared gene content. It also determines the percentage of genomes sharing each gene. Roary requires multiple General feature Format (gff) files containing protein sequences. The previous Pharokka annotation provided a

combined gff file containing all annotated protein and nucleotide sequences. Using Python v3.12.7 and Biopython v1.85 (Cock *et al.*, 2009), the big .gff file was divided into individual phage genomes which were distributed according to the host bacterial genus. Therefore, the output was a set of directories named according to host bacterial genus, and containing the .gff files with the sequences coming from each individual identified phage of that bacterial genus. Another python script was used to loop Roary across each genus directory and compare the genes between all the phages found in that bacterial genus. The output directory was generated for each genus.

To visually inspect the shared genes between phages, Phandango v1.1.0 (Hadfield *et al.*, 2018) was used. However, Phandango does not offer color-coding genes in any way. Color-coding the represented genes by function creates better visual information on the gene distribution, mosaicism, and functional evolution. Therefore, the Java Script source code of Phandango was modified to allow custom color-coding of the genes represented. All that was required was adding a “color” or “colour” column into the input csv table generated by Roary. That column should be filled with a HEX color code of user’s choice for each gene. Using Python, colors were assigned based on the gene function deduced by the similarity to the genes in PHAROG database. The phages were categorized based on the host species, previously determined family, and the level of completeness.

3.2.6. Analysis of antibiotic resistance and virulence genes

The antibiotic resistance and virulence genes were categorized by the host species. Basic descriptive statistics were performed for all the genes found. As some highly similar AMR and VIR genes were found in different bacteriophages, it was interesting to determine if all the similar genes stem from the same of from different, independent phages. To do that, again CD-HIT clusters were utilized. The phages belonging to the same cluster with the same AMR or VIR genes were concluded to come from the same origin and therefore, there was no horizontal gene transfer between the phages. The phages that were found with the similar AMR or VIR genes and belonging to different clusters were examined further as they presented a potential for horizontal gene transfer between prophages.

3.3. Results

3.3.1. Description of the prophage population in the context with the hosts

First, the basic descriptive statistics of the population were performed. The distribution of the viral genome completeness found among the bacterial genomes isolated from the ISS is shown in Figure 21. Of all the temperate phages found, only 23.9% were active. There was a total of 2,217 phage genomes identified in 602 bacterial genomes, which means an average of 3.68 bacteriophages per bacterial genome. The exact distribution of the amount of the bacteria

according to the number of phages identified in a genome is shown in Figure 22. In total, 99.17% of all bacterial genomes had at least one prophage in the genome. Of course, it must be noted that the bacteria analyzed here do not represent the whole ISS microbiome, but only a part of it. Specifically, the bacterial strains that were isolated and grown to sufficient amount that their full genomes can be sequenced.

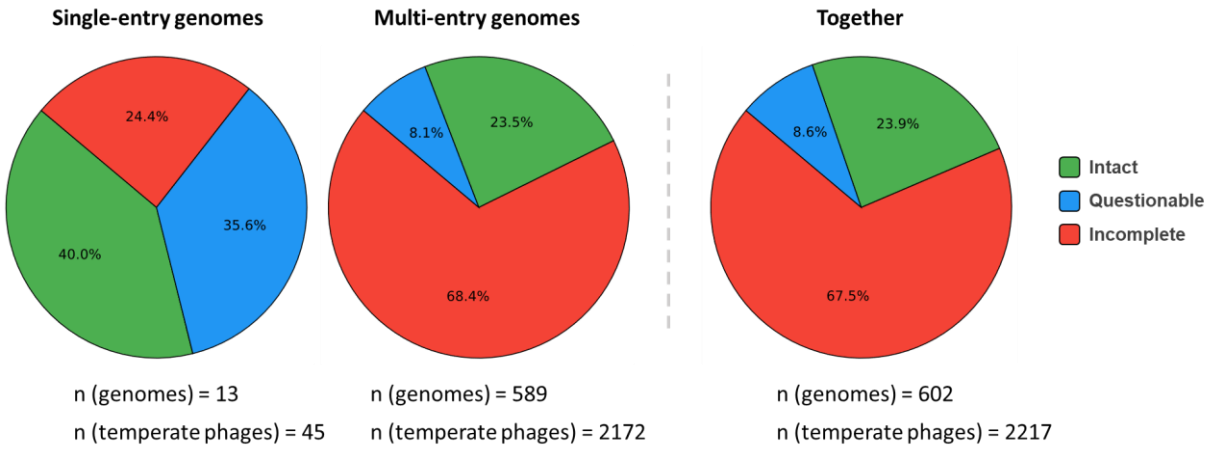


Figure 21. The basic distribution of the bacteriophages found in the bacterial genomes isolated from the International Space Station (ISS) categorized according to the determined bacteriophage activity. Intact phages are complete and potentially could be reactivated, questionable are the ones where the algorithm could not determine if the phage could be active or not, and the incomplete bacteriophages are those where the genome is incomplete but there are still some bacteriophage genes.

The medians, maxima and minima of all groups found per bacterial genome are shown in Table 10. Taken all together, the bacteria isolated from the ISS had a median of 2 (Q1 = 1; Q3 = 3) bacteriophages per genome. The maximum overall number of phages found in a genome is 7, while the minimum 0.

Table 10. Descriptive statistics of the number of prophages in bacterial isolates from the ISS.

	Median	***Q1-Q3***	***Minimum***	***Maximum***
All together	2	1-3	0	7
Intact	2	1-2	0	5
Questionable	1	1-1	0	3
Incomplete	2	1-4	0	7

The descriptive statistics were performed for all bacteriophages found. The results are plotted in Figure 22. The left column plots represent the distribution of prophages per bacterial genome, grouped by the activity of a prophage as determined by PHASTER. The top row represents all the genomes grouped together, the middle row represents only the single-entry

genomes (which are known to be complete) and the bottom row represents only the multiple entry genomes (which contain contig-level assemblies and plasmids).

To find out whether the groups are statistically significantly different from each other, a series of tests was used. The results of all statistical tests performed on the data represented in Figure 22 are shown in Supplementary Table 2.

First, the normality of the data distribution was tested for each group with the Shapiro-Wilk test. Notably, neither together, single-entry genomes nor multiple-entry genomes had all the data normally distributed, as can also be observed visually in Figure 22. Therefore, all the groups were tested for the differences among each other with Kruskal-Wallis Test. Being non-parametric test, it does not make assumptions about the distribution of the data. Notably, among the single-entry genomes, K-W test showed no difference among the number of intact, questionable, or intact. However, when all the bacteria are put together, there are significant differences in the groups. Therefore, to determine which groups are different from each other, pairwise Mann-Whitney U tests were used between each pair of the three groups (intact, questionable, and incomplete). All the groups have highly significant differences between each other. Despite the equal medians between intact and incomplete phages, the difference in their distribution is significant (Figure 22 A).

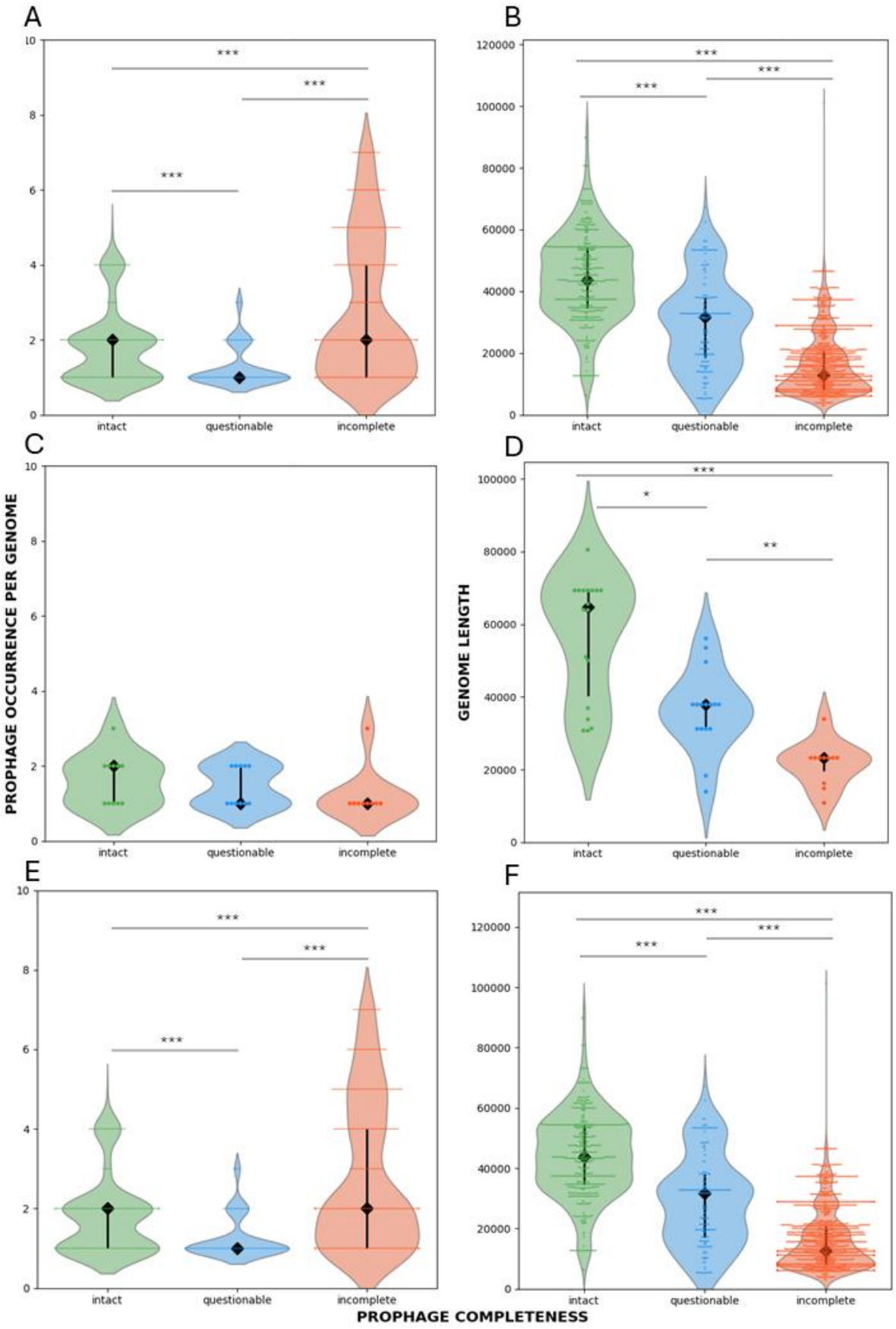


Figure 22. The descriptive statistics of the temperate phage genomic sequences found inside genomes of the ISS bacterial isolates. Median, Q1, and Q3 range are represented in black. Left

column represents the number of phages determined per bacterial genome, while the right column represents the genome length A and B, the distribution of the bacterial genomes according to the number of bacteriophages identified per genome. While significantly more inactive or incomplete bacteriophages were identified in bacteria per genome, they tend to have clearly shorter genomes, with the longest genomes being those of the active phages. C and D show the descriptive statistics of phages found in single-entry genomes, meaning they come from complete, circularized, bacterial genomes. They are a minority of the data. E and F represent the statistics of the multi-entry genomes, mostly consisting of contig-level assembled genomes and plasmids.

Most of the temperate bacteriophage sequences were inactive (67.5%). This is inferred by PHASTER based on the gene content of the identified phage sequences. That shows that even if the phage has a long genome, if it is lacking essential genes, it will be labeled as incomplete. Interestingly, bacteria tend to have one or two active phages in their genome.

To further understand the distribution of the phages across the bacterial species, the bacterial host genome lengths were plotted against phage genome lengths and phage counts found. This analysis was performed for two main reasons: First, it would show the basic view of the phage population among the hosts, and second, if those values were strongly correlated, it would mean a random distribution and high level of randomness in finding bacteriophage sequences. This is because the chances of finding a phage would depend on the size of a genome that is being analyzed and therefore, high level of phage exchange in a population. The representation of the correlations can be found in Figure 23. The strength of the trends has been determined with Pearson and Spearman correlations since there was no assumption about whether the trend would be linear (better with Pearson correlation) or non-linear (better with Spearman correlation). To retain the statistical strength of the analysis, only the genera with more than 20 bacteriophage genomes found were plotted. Additionally, the question was asked whether bigger genomes contain higher amounts of prophages. Figure 24 shows the plot of phage count plotted against the host genome length. For the statistical significance, only the genera with more than 20 bacterial genomes were plotted. Still, no significant correlation was found between the host genome size and number of phage genomes found, representing again that larger genomes do not necessarily have higher phage counts.

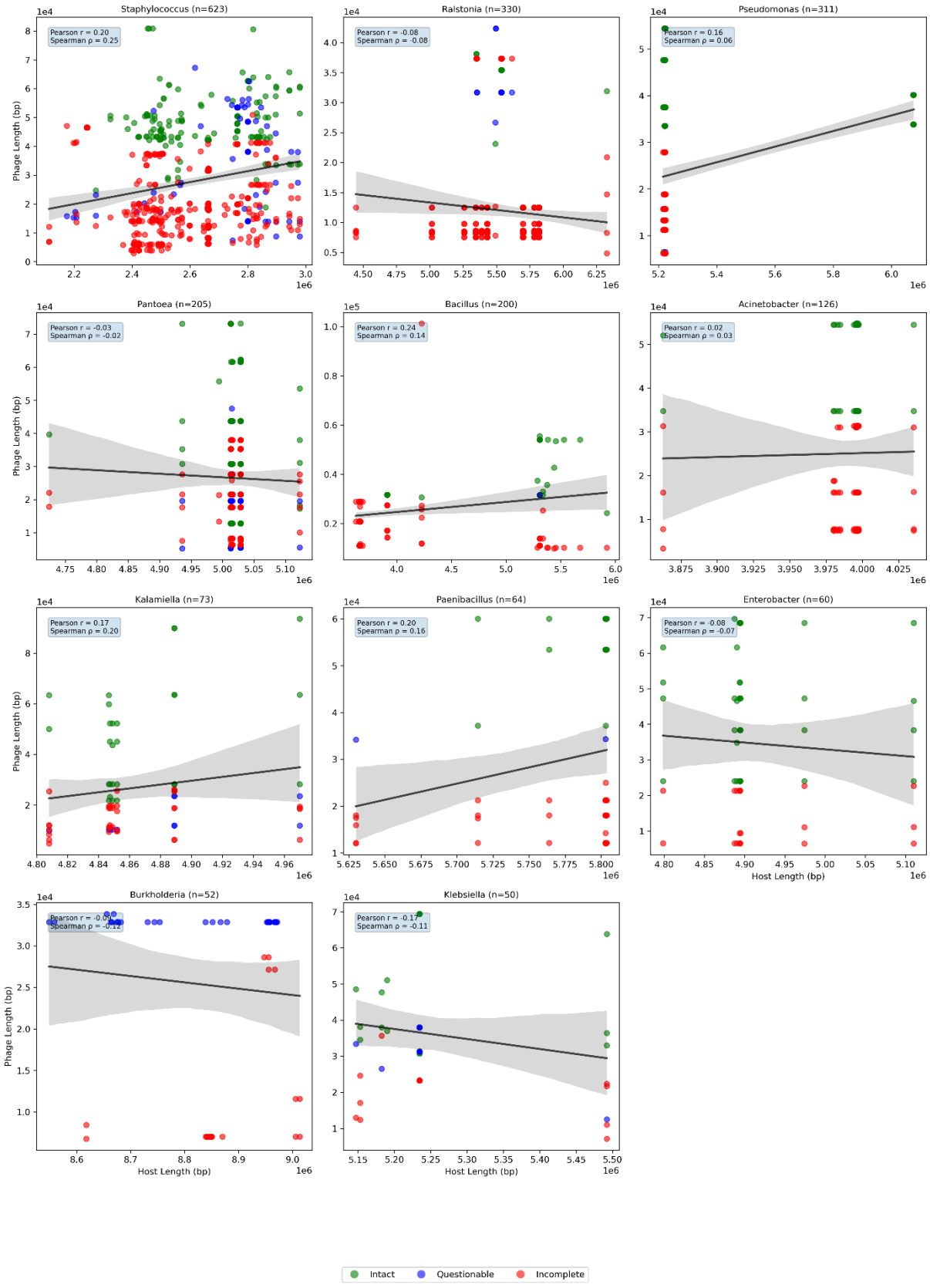


Figure 23. Correlation of the phage genome lengths with the lengths of the host genome, grouped by bacterial genus. All genera show very weak trends, both with Pearson correlation and Spearman correlation, as all absolute |r| and |ρ| values are smaller than 0.3. Around each model, a 95%

confidence interval is plotted. The points are colored according to the completeness of the genomes according to PHASTER.

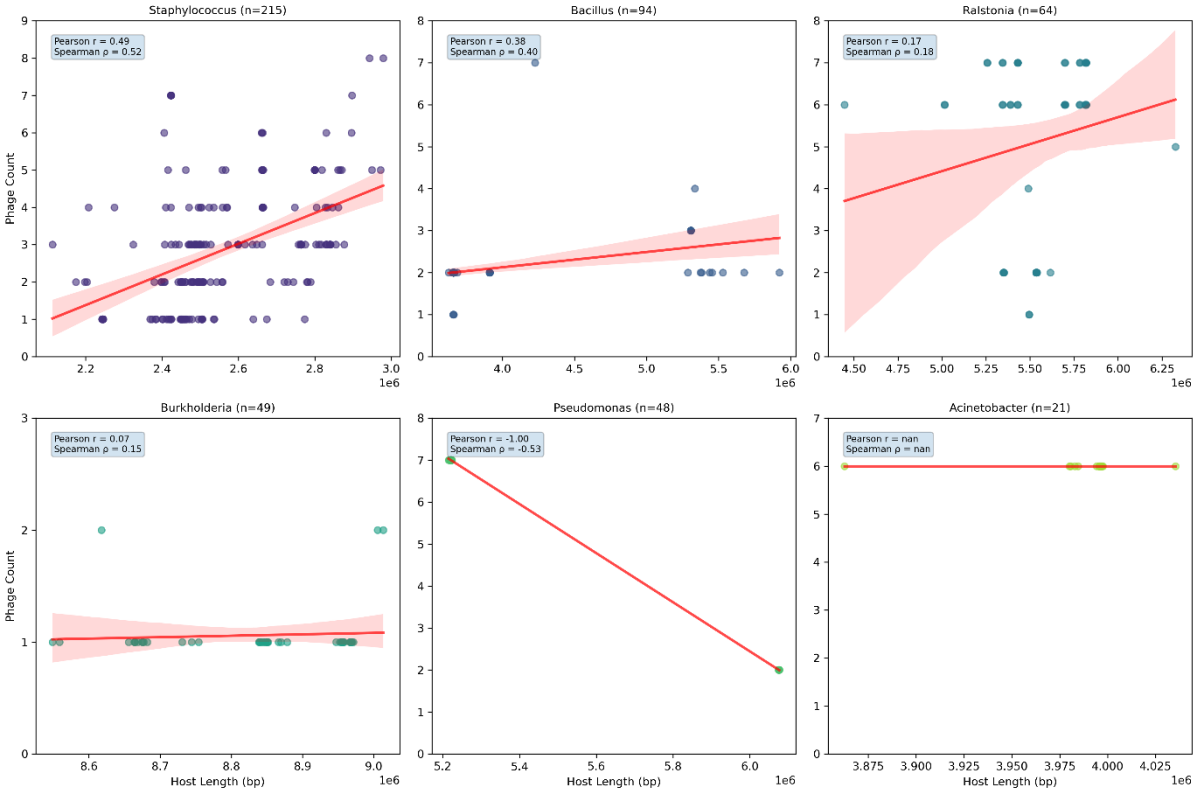


Figure 24. Regression analysis of the possible correlation between the phage number per genome and genome length. Only genera with more than 20 genomes were plotted for statistical significance of the regression. No strong correlation was found in most genera. Interesting observations come from *Pseudomonas* and *Acinetobacter*. The former has low genomic variation, meaning that despite the high total number of genomes, only a total of two different genome types were found, as seen from only two genome length points. *Acinetobacter* on the other hand, has more distributed genomic variation, but the number of phages per genome stays at 6 for all the genomes found, resulting in no ability to calculate Pearson r and Spearman ρ.

3.3.2. Analysis of the clusters of related prophages

To understand the distribution of unique phages globally, the clusters found by CD-HIT based on 95% and 100% identities were constructed. The 100% identity-based clusters represent highly strict clustering, where no mutations are allowed for the phage sequences to fall into the same cluster. On the other hand, 95% identity-based clustering allows for up to 5% global difference between phage genomes. Despite not being a perfect definition of species or genus, 95% identity is generally considered a threshold identity for bacteriophages and archaeal viruses to be considered a different species (Adriaenssens and Brister, 2017). The limitation of this approach is using pure nucleotide identities. While nucleotide sequences can show high

level of identity, the protein sequences could differ significantly (Grose and Casjens, 2014). With those limitations in mind, we can plot the numbers of clusters with mutual minimum 100% and 95% average nucleotide identity (Figure 25).

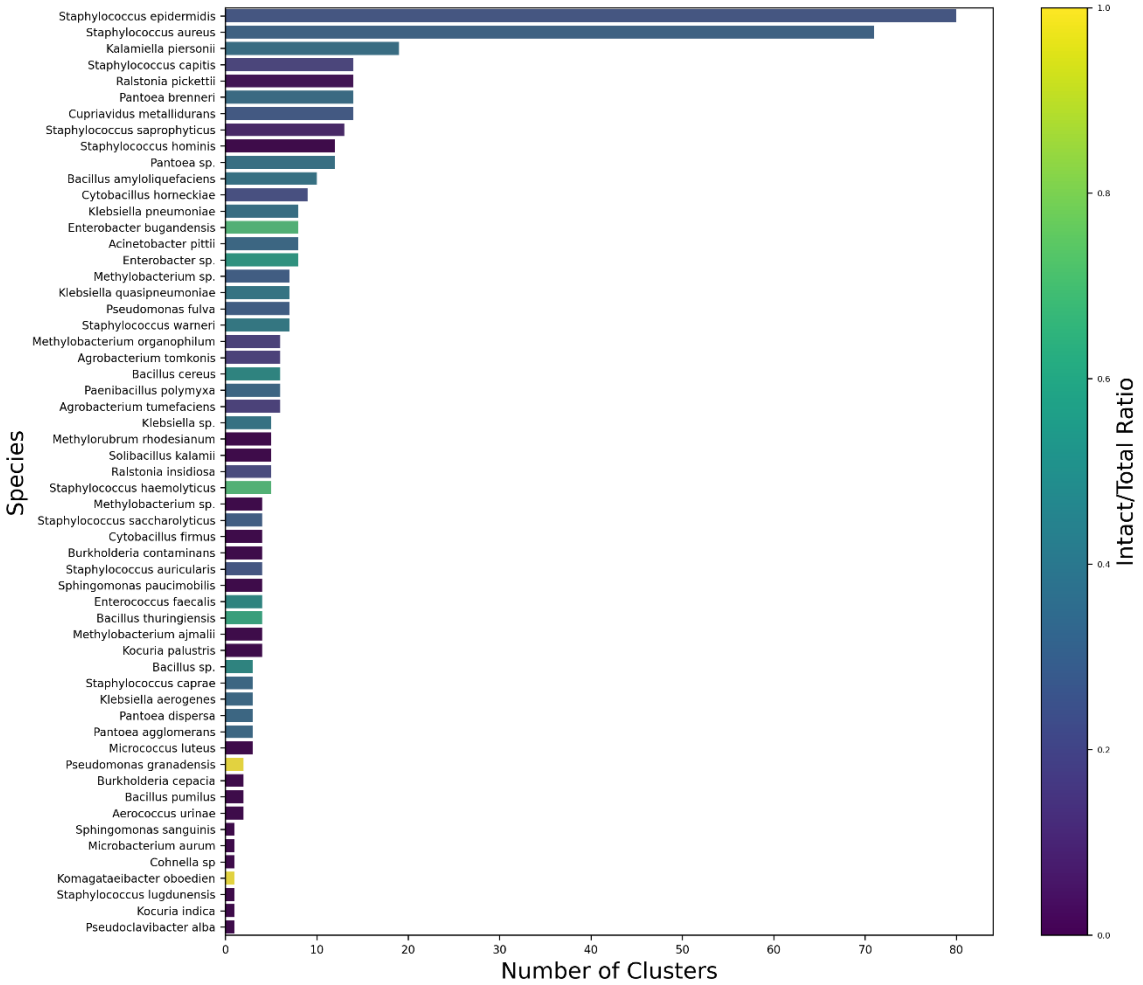


Figure 25. The distribution of phage clusters among different bacterial host species based on 95% identity. The color scale shows the overall portion of the intact phages for each species.

Next, it was explored how does this distribution looks at the host genus level. Therefore, the same plot was created for the host genus level instead of species (Figure 26). At the genus level, we observe the diversity being dominated by *Staphylococci* with 209 clusters in total. The next host genus containing high phage diversity is *Klebsiella* with 20 total clusters. This is less than 10% of the *Staphylococcus*, indicating significant dominance in phage diversity in *Staphylococci* over other bacterial genera. Though there are genera with higher proportion of intact prophages, such as *Enterobacter*, 66.67% (40/60), most of the genera show this portion to be below 50%, as it was expected based on the overall distribution (pie chart in Figure 21). Still, it is insightful to see the distribution of phage activity among the microbiome of the ISS.

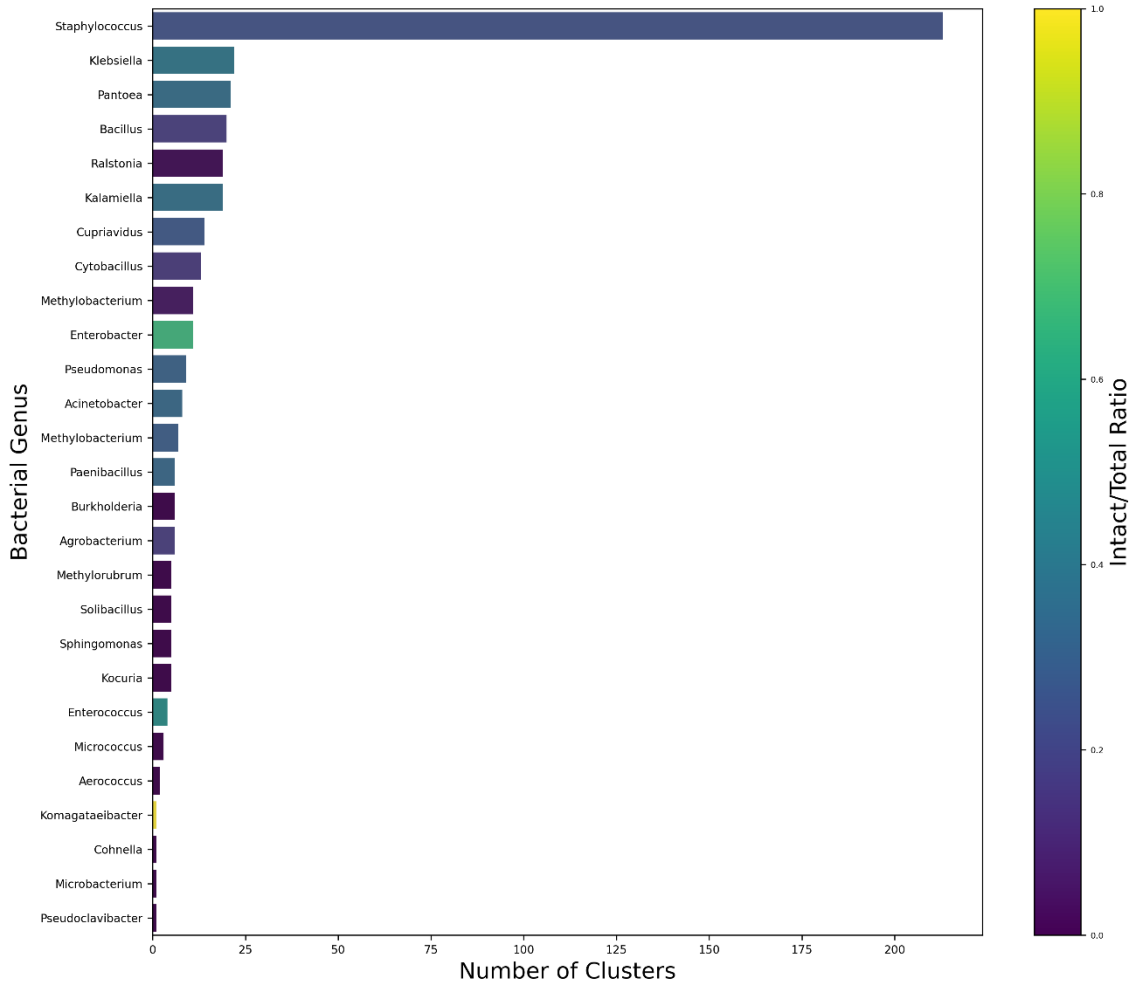


Figure 26. Distribution of the 95%-identity clusters of bacteriophages across bacterial genera. Color-coded by the portion of intact phages across the genus. Most of the clusters come from *Staphylococci*, though they do not show the highest portion of intact phages. Rather, the phage activity is distributed independently across bacterial genera.

Next, it was interesting to compare the per-genome density of bacteriophages among the genomes. This allows us to determine the species and genera that are denser in phages then the others despite the number of phage clusters. Figure 27 shows the average amount of bacteriophage clusters per genome.

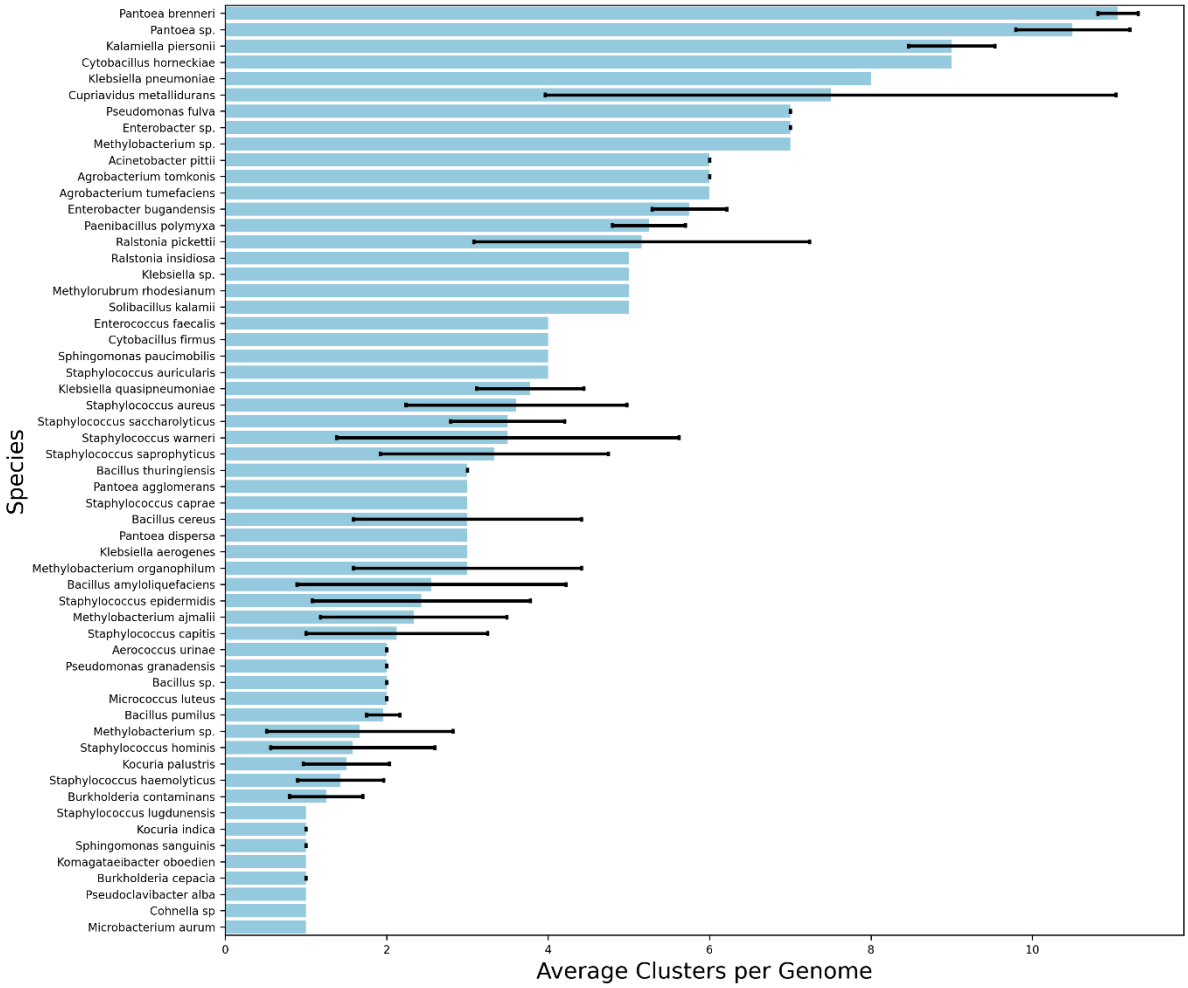


Figure 27. Diversity of bacteriophage sequences per genome of each bacterial species. The highest density of phage diversity comes from genus *Pantonea* where both *Pantonea brenneri* and *Pantonea* sp. have more than 10 clusters per genome on average. They also have low variation in that number, meaning that even though the bacterial species are low in diversity, they act as hotspots of diversity for prophages.

From the average numbers of clusters per genome, it is visible that some species like the genus *Pantonea* have higher density of diversity than others such as *Microbacterium*. To show whether there are significant differences between the species, a statistical test was used. To determine whether to use a parametric or non-parametric statistical test, Shapiro-Wilk test was utilized (Shapiro and Wilk, 1965), examining how much each species data distribution deviates from normality, for all species with more than 3 data points. Significance threshold α of 0.05 was chosen. Since the data for most species deviated from normality ($p<0.05$), a nonparametric test was used - Kruskal-Wallis test (Kruskal and Wallis, 1952), again with the significance threshold $\alpha=0.05$. The test showed $p=0$, meaning that there were significant differences found between the species. To determine the groups that differ significantly, Mann-Whitney U-test was performed (Mann and Whitney, 1947). The p-values between each pair of

species were corrected for the number of comparisons with Bonferroni correction, by multiplying the p-value with the number of comparisons performed. The bottom picture shows a heatmap showing the significance of the differences after comparing all the groups among each other and after the correction (Figure 28).

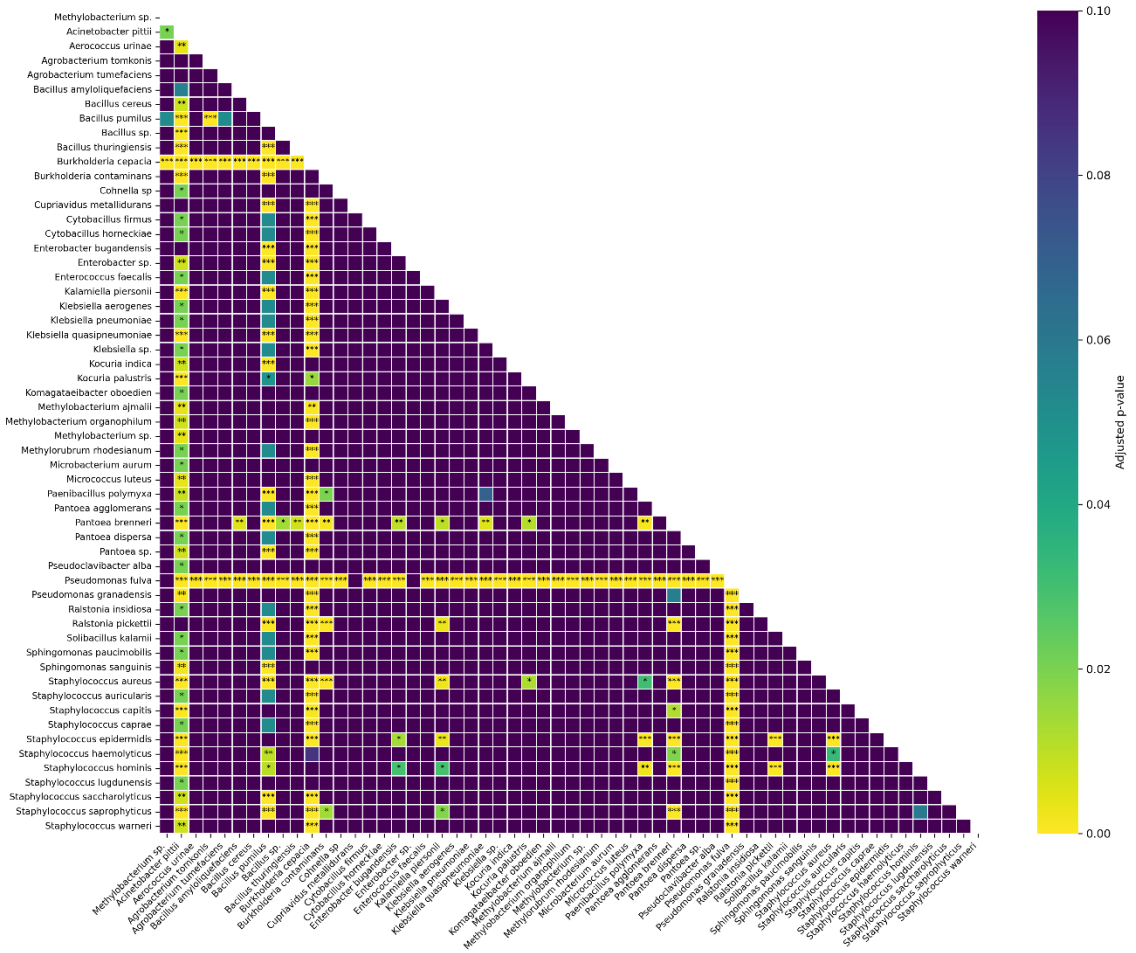


Figure 28. The heatmap of the Bonferroni-adjusted p-values comparing the average number of prophages per bacterial genome with Mann-Whitney U-test. Combined with the graph above, it represents that some species have significantly different phage diversity densities then others, showing they represent more isolated and stricter ecological niches for prophages. The $p<0.05$ are labeled with *, $p<0.01$ with **, and $p<0.001$ with ***.

Interestingly, there are multiple bacterial species that show a significant difference in phage density to almost all others. Those are *Burkholderia cepacia*, *Acinetobacter pitii*, and *Pseudomonas fulva*. *Bacillus pumilus* also shows the difference to many other species, though not as many as the former three. Those species present dense reservoirs of phage diversity. At the same time, *Staphylococcus* species, which showed highest total phage diversity, have no more than 4 phage clusters per genome. This might be due to high abundance of this genus, giving space for prophages to distribute over the population. To determine the phage density distribution on the genus level, the same plot was presented for the genera (Figure 29). In

agreement with the species-based distribution above, *Pantonea* and *Kalamiella* genera remain the densest centers of phage diversity. Interestingly, even on the genus level, the variation of some genera was low, with standard deviation of 0 – such as e.g. *Microbacterium* or *Methylobacterium*. The map of the density distribution of phage variation among genera shows insights into the diversity of the prophages on the International Space Station.

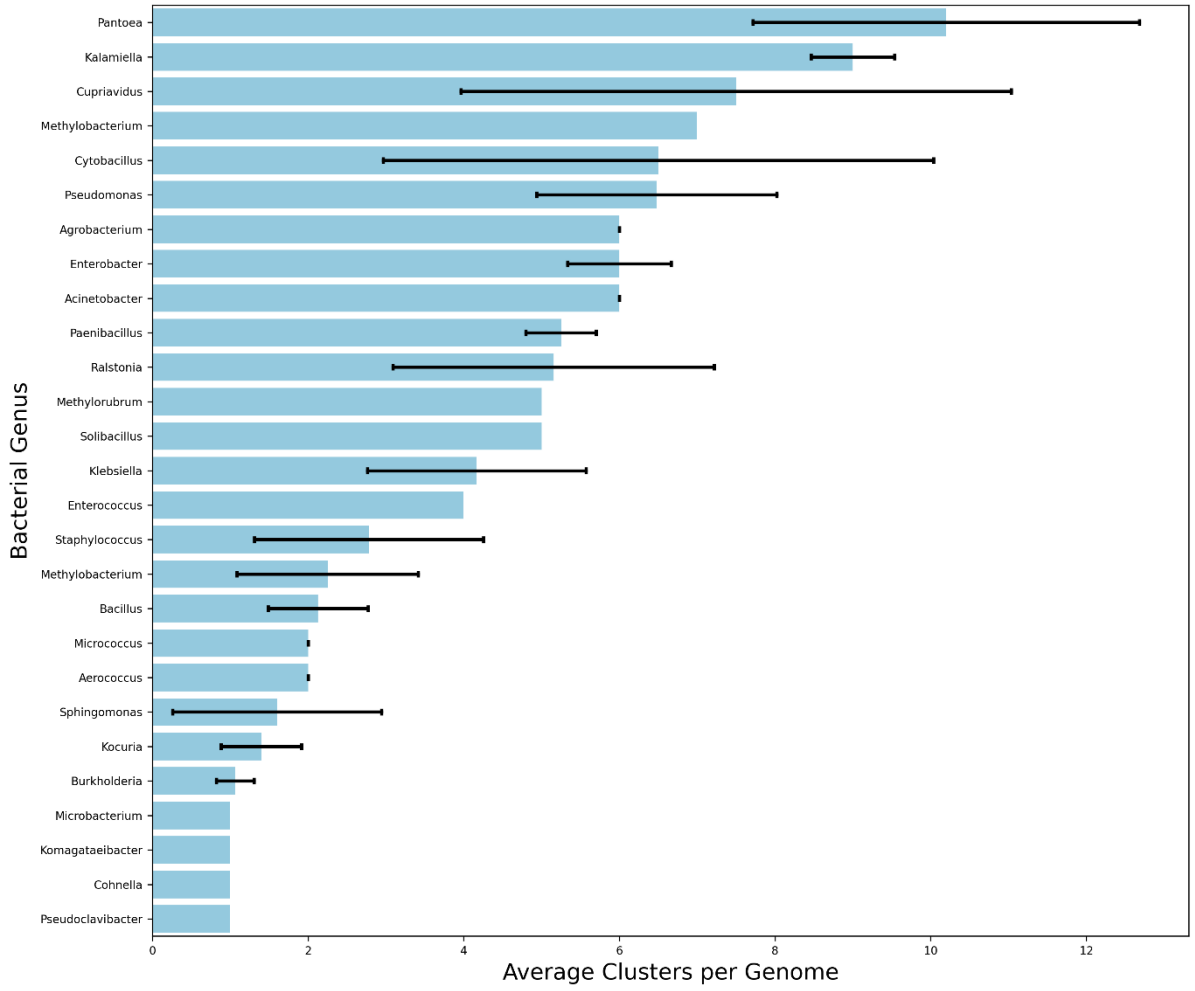


Figure 29. Diversity of bacteriophage sequences per genome by genus.

The heatmap below shows the significance of the differences between the bacterial genera in phage diversity is visible in Figure 30. There, the genera with the significantly different diversity from other genera are determined. The genera that differ significantly have a unique set of phage clusters per genome. This helps to identify the clusters of unique prophage dynamics.

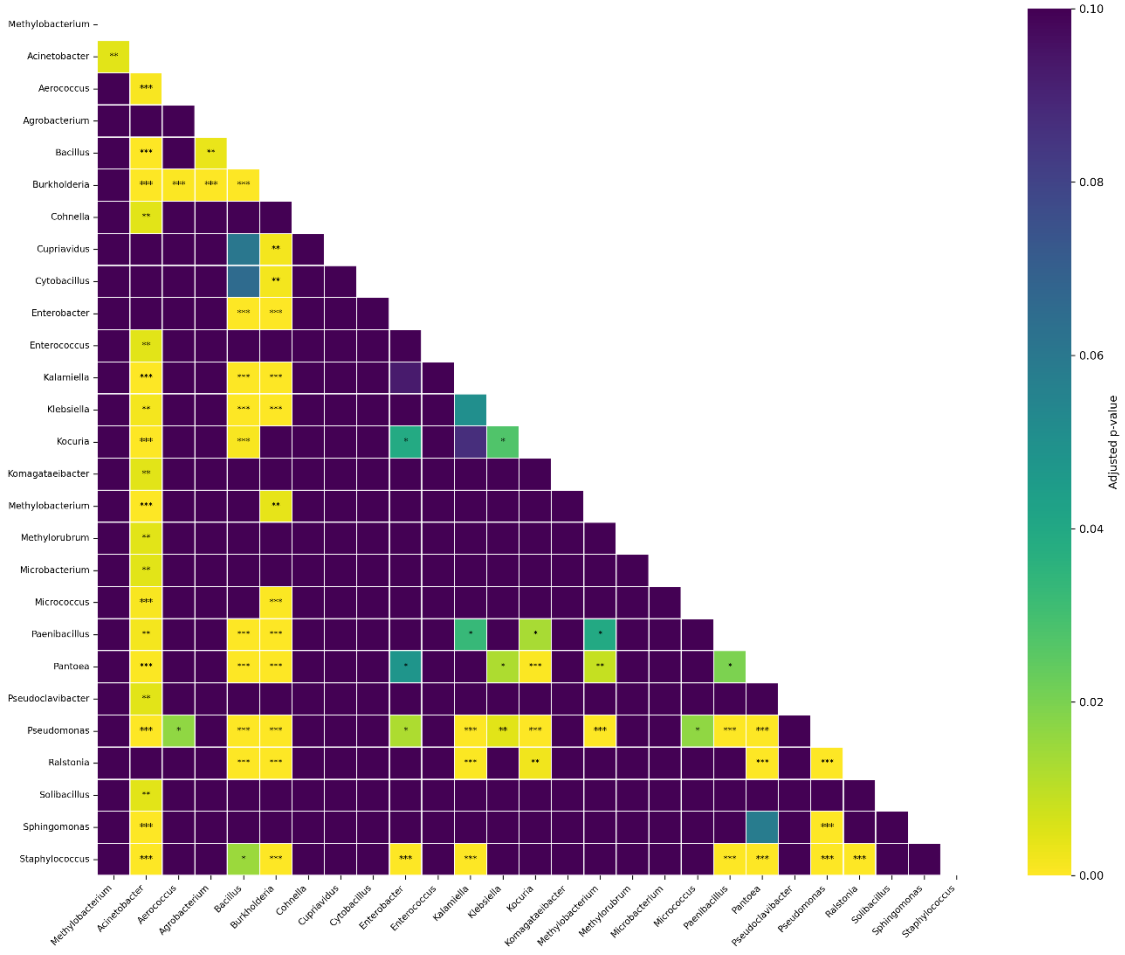


Figure 30. Heatmap of the significance of the differences between different bacterial hosts regarding their phage diversity densities. Based on the phage cluster density and comparisons using Kruskal-Wallis test and Mann-Whitney U-test as well as Bonferroni p-value correction. The significant differences in prophage diversity densities point to the species that serve as phage clusters, and those that have lower density. Therefore, it points to the distribution of prophage diversities across bacterial host genera.

All the prophage genomic sequences identified are included in the Supplementary Text 1. To access the complete report for each individual bacterial genome with all the phages found inside, it is enough to access the website https://phaster.ca/submissions/ID, where ID should be replaced by the PHASTER-assigned genome ID (ZZ_ followed by 10 alphanumerical characters) included in the sequence name.

3.3.3. Prophage genome annotation and classification

After the basic description of the phage genomes across the bacterial species and genera, the genomes were annotated with Pharokka. With the knowledge of the genes and their functions, the functional gene distribution could be explored among host genera. Therefore, after annotating all 2,217 genomes, a heatmap was created with the distribution of gene groups and their functions. If the heatmap would represent the raw gene counts, there would be significant

problems with the representation of gene diversity because many of the phages are identical, biasing the raw count towards the groups with higher gene counts. Therefore, the protein sequences were clustered into groups according to 75% identity using CD-HIT. Then, instead of raw gene counts, the heatmap was plotted to show the number of gene groups that share a minimum of 75% identity. The data was in this way corrected for the number of sequences and could reliably show the distribution of gene groups in phages across the bacterial genera. Figure 31 represents those results. The heatmap reveals the distribution of gene diversity among genera isolated from the ISS.

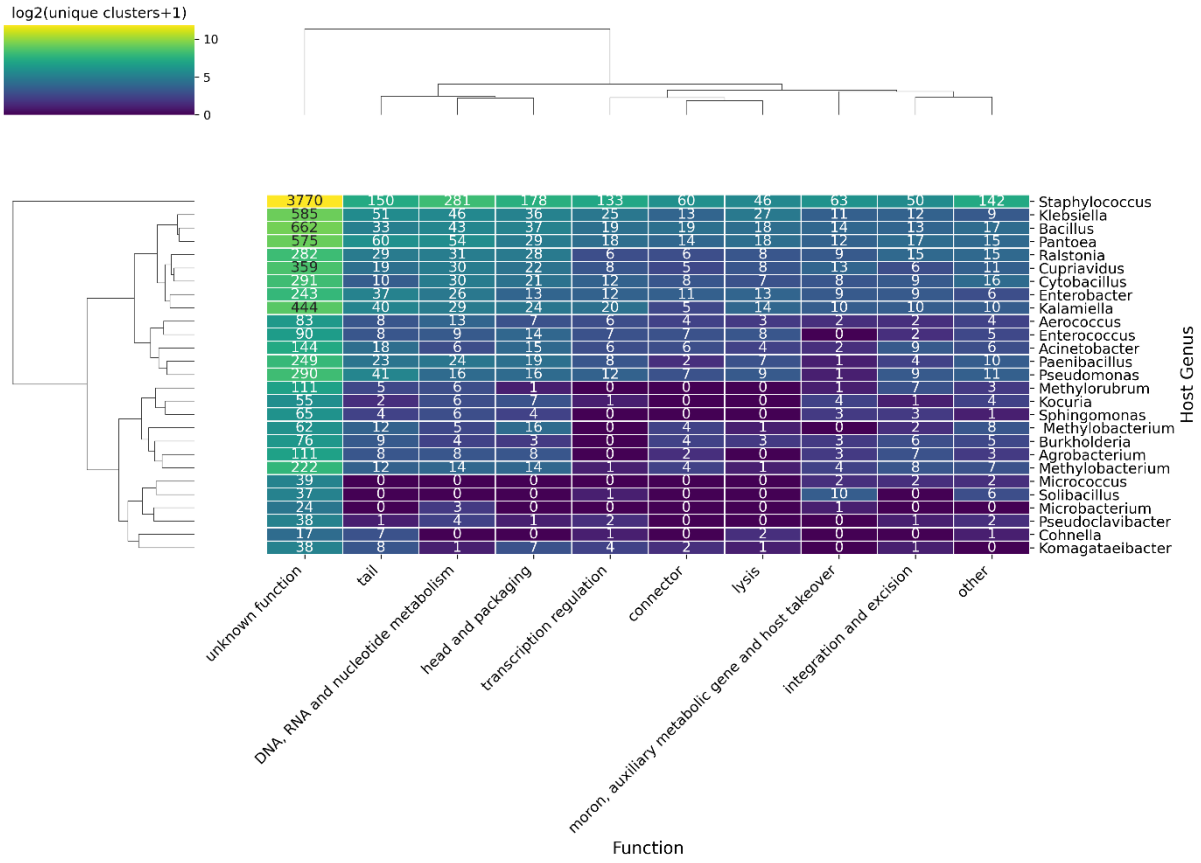


Figure 31. The heatmap of functional gene diversity among prophages across the bacterial genera. The heatmap color scale is log2-transformed for better visibility, while the raw numbers of gene clusters are written in each field. The genera and functions are ordered hierarchically according to similarity.

It is worth noting that the heatmap above shows the protein-coding genes. However, there were also 79 tRNA genes found in total. Figure 32 shows the heatmap of the distinct tRNA genes found in the prophage population. Not all genera contained tRNA genes inside their prophages, and only those that did are represented. Here again, *Staphylococcus* phages stand out with high tRNA gene counts compared to other bacterial host genera.

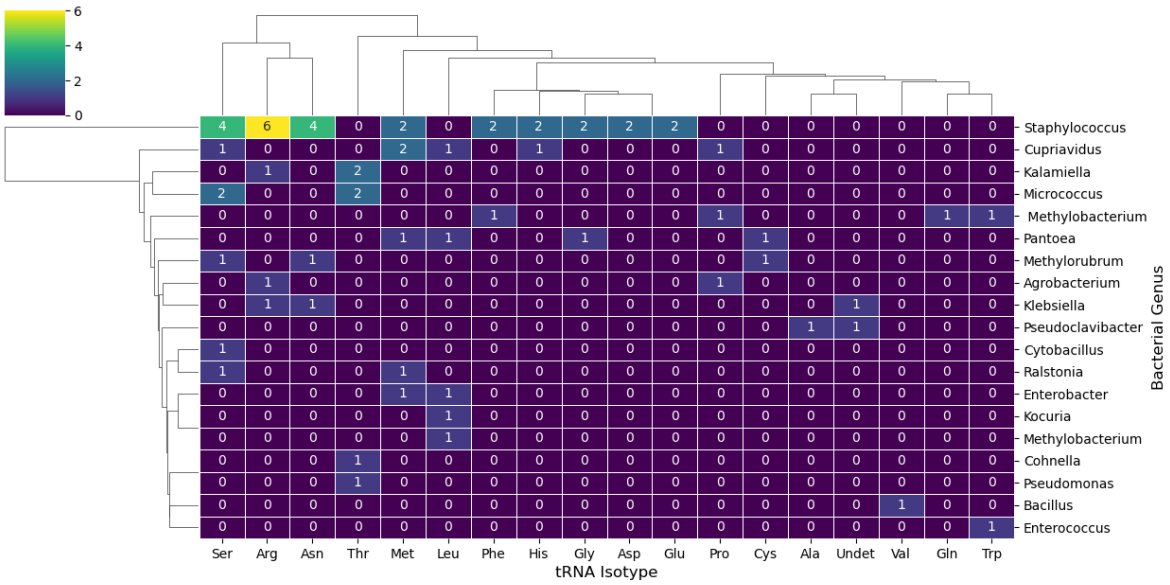


Figure 32. The tRNA genes found within prophages of the ISS categorized by host genus and amino acid. *Undet* is a label for genes which were determined to encode tRNA but not determined for which amino acid.

The classification of the bacteriophage genomes is complex, as they do not possess a conserved marker gene that evolves at an approximately constant rate. This complicates their phylogeny. One approach uses the whole proteome clustering, which aims to determine all the homologous genes shared between a set of bacteriophages. Therefore, once the phage genomes were annotated, the attempt could be made for their classification.

Based on the shared gene content, the phage genomes are clustered together. A phylogenetic tree of 45 phage sequences identified from the 13 full, single-entry bacterial genomes was constructed in this way, using VIPTree (Nishimura *et al.*, 2017). It is shown in Figure 33.

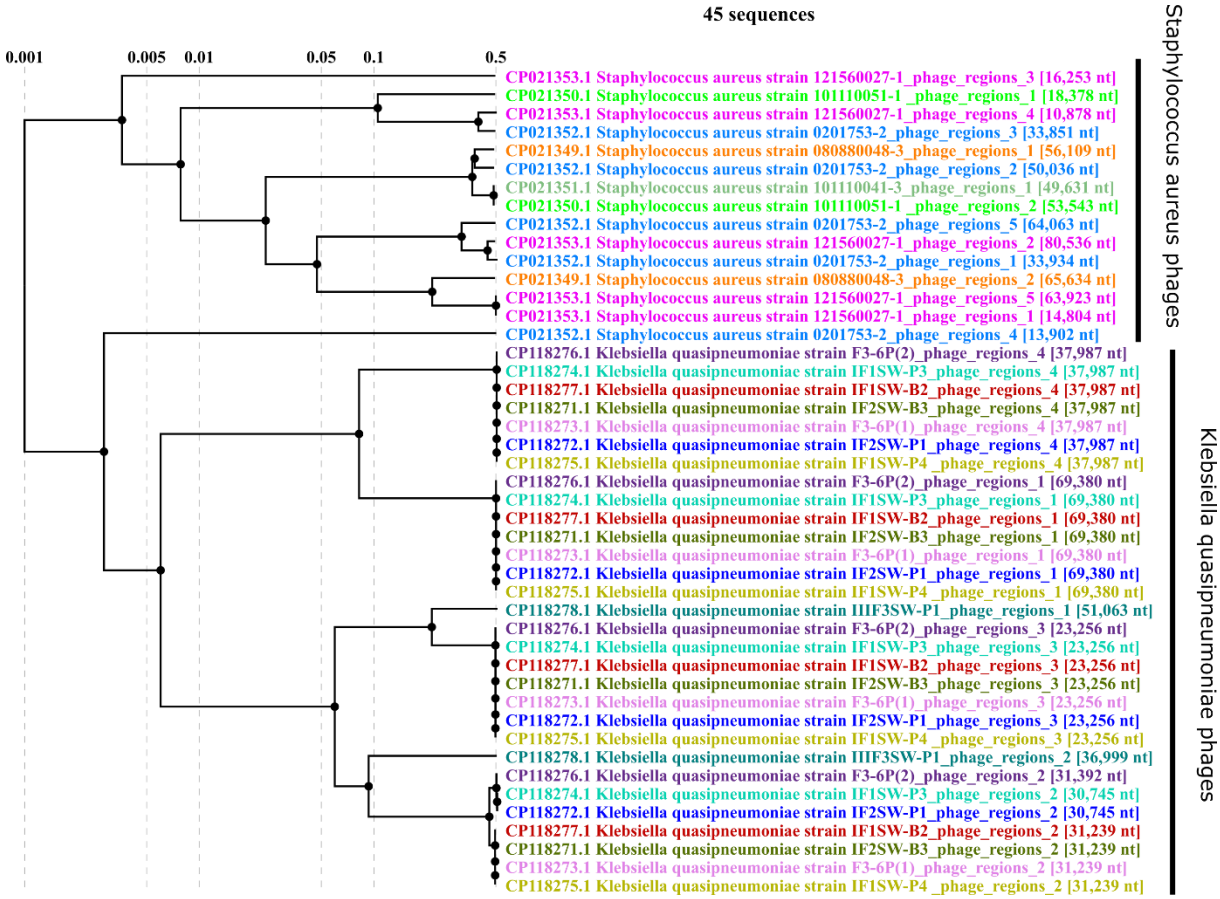


Figure 33. Relationships of the 45 phages identified in all of the single-entry genomes. The phage genomes cluster together depending on the shared gene content. The phages found within the same bacterial genome are labeled in the same color.

When the unknown phage sequences are compared to the reference genomes with the solved phylogeny, the classification of the phages can be resolved. To determine the total classification of all temperate phages found in the bacterial isolates of the ISS, the shared gene content was compared to the reference database of bacteriophage genomes and the phages found were clustered to the ones to which the gene content is the most similar. This was done using vContact2 as described in the methods. The clustering of the temperate phages found in the bacterial isolates of the ISS, are shown in Figure 34. The phages are color-coded according to the predicted activity. Prophages with more shared genes cluster closer together. Supplementary Figures 1 and 2 show the same map but color-coded by genus and pathogenicity respectively. The ntw file containing the whole network and the cdv table with the metadata are also provided in the Supplementary Text 2. The network and the data can be inspected using Cytoscape v3.10.3 (Paul Shannon *et al.*, 2003). As expected, the phages that are determined to be intact cluster overall better with the reference genomes, though this is not a strict rule.

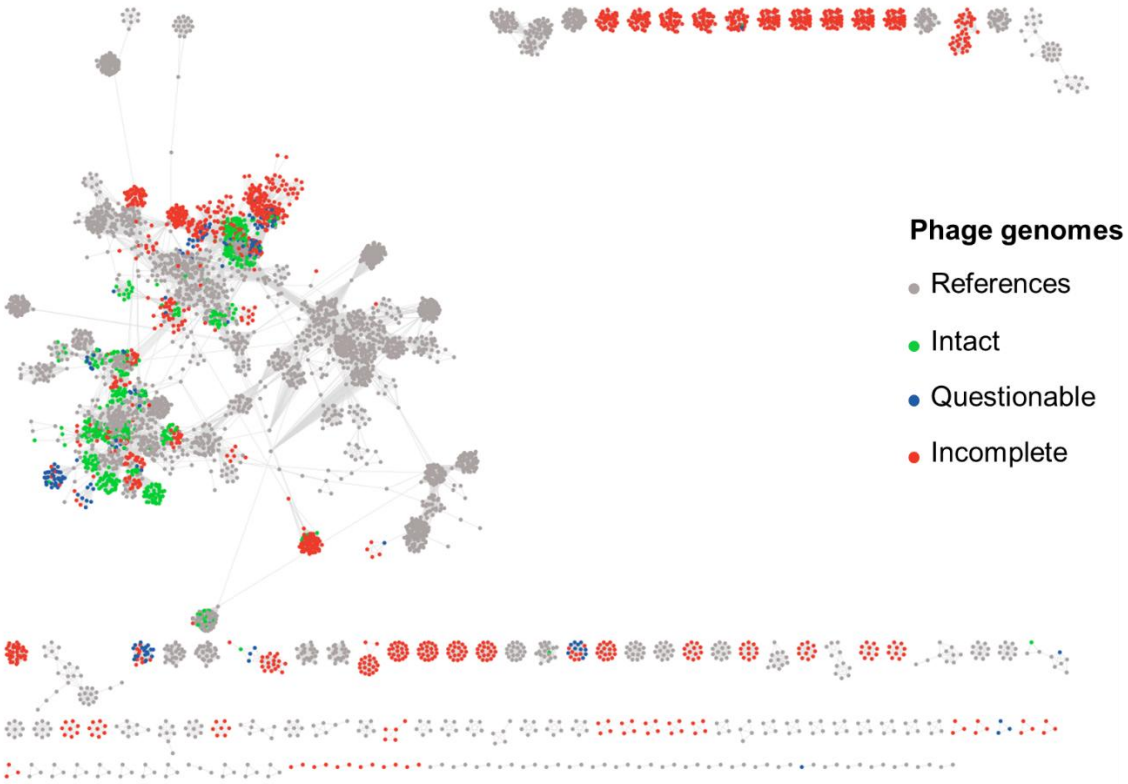


Figure 34. Clustering of the identified prophages in comparison to the reference phage genomes based on vContact2. The visualization was done in Cytoscape. The clusters away from the main tree could not be assigned to a group. Significantly, the groups of identified phages are clearly distinguishable. Multiple groups of bacteriophages were found. Most of the phages that are determined as intact and questionable cluster well with reference genomes, as well as some inactive phages. However, most of the phages determined inactive did not cluster with references, but rather among each other. Based on this clustering, the phage classification was performed.

The classification of the phages was performed based on their genomic content and homology to the reference phage genomes. Most of the prophages remain unassigned, forming their own clusters, far from reference sequences, as visible in the figure above. Those phages are labeled *Separate Cluster*. The full list of phages with the reference genomes is available in Supplementary Table 3 and without the reference genomes in Supplementary Table 4. Those tables can be loaded into Cytoscape together with the provided ntw file (Supplementary Text 2) to recreate the graphs produced here. Some unassigned phages clustered close to a reference phage genome but remained unassigned, because the reference phage itself was also of an undetermined family. Such prophages are labeled *Unassigned*. The original classification according to families was determined with graphanalyzer v1.6.0 (Mattia *et al.*, 2022). Since the phage genomes were also clustered according to 95% average nucleotide identity, this clustering was used to improve vContact2 classification based on shared genes. Therefore, the improved analysis includes the clustering based on gene content as well as

nucleotide identity. The family distributions of both kinds of analysis are represented in Figure 35. The left chart represents only the use of vContact2 while the right chart adds CD-HIT analysis. In total, 5.4% of the phages (120 out of 2,217) were reassigned from *Separate cluster* to one of the known families. The phage family that mostly profited from the improved analysis is *Siphoviridae*, gaining 118 of the phages while one came to *Podoviridae* and one to *Autographiviridae*. At the same time, the amount of *Myoviridae*, *Inoviridae,* or *Unassigned* prophages did not change after the improved analysis. Despite the improvement being small, it still adds to the accuracy of the analysis.

The prophages of the ISS seem to be dominated by phages which were not successfully classified. This again points out the fact of how little is known about overall phage diversity and phylogeny. Of the known families, *Siphoviridae* seems to be a heavily dominant family of prophages.

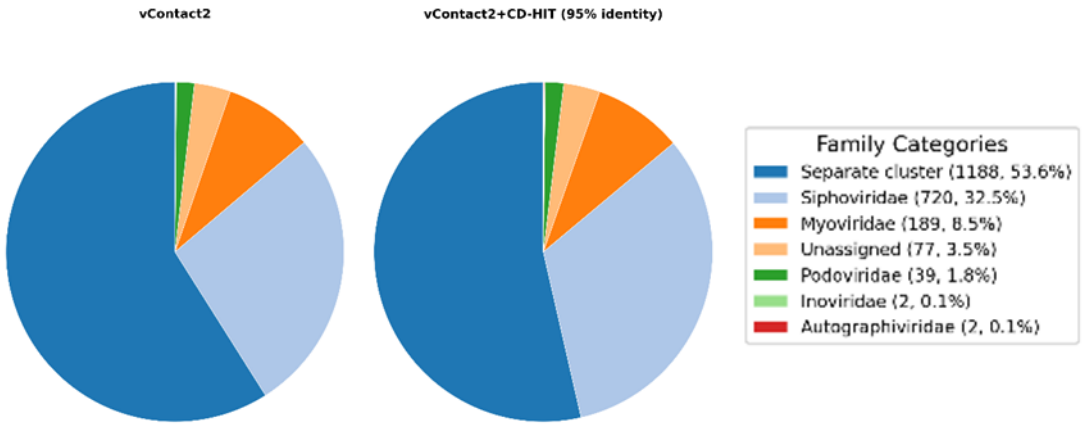


Figure 35. The diversity of temperate bacteriophage families found in the bacterial isolates of the ISS. The left pie chart represents phage families baser solely on vContact2 clustering classification, while the right pie chart represents improved classification based on vContact2 and CD-hit after clustering the genomic sequences based on minimum 95% similarity which is well within the definition of a phage family.

Once the phage families were determined, a new distribution chart could be made representing the phage family diversity among bacterial genera. This plot is similar to the one in Figure 26 but instead of representing phage completeness, this one is stacked and filled by the prophage families found within each host genus. Figure 36 shows that chart. Since the variation between genera is significant with a bias towards *Staphylococci*, a log2-scaled version of the graph is added in Supplementary Figure 3 for better visibility of the genera with low phage cluster counts. It is important to keep in mind that log2-scaled stacked graph doesn't intuitively represent well the distribution of prophage families but shows better visibility for the genera with the lower counts of 95%-identity clusters.

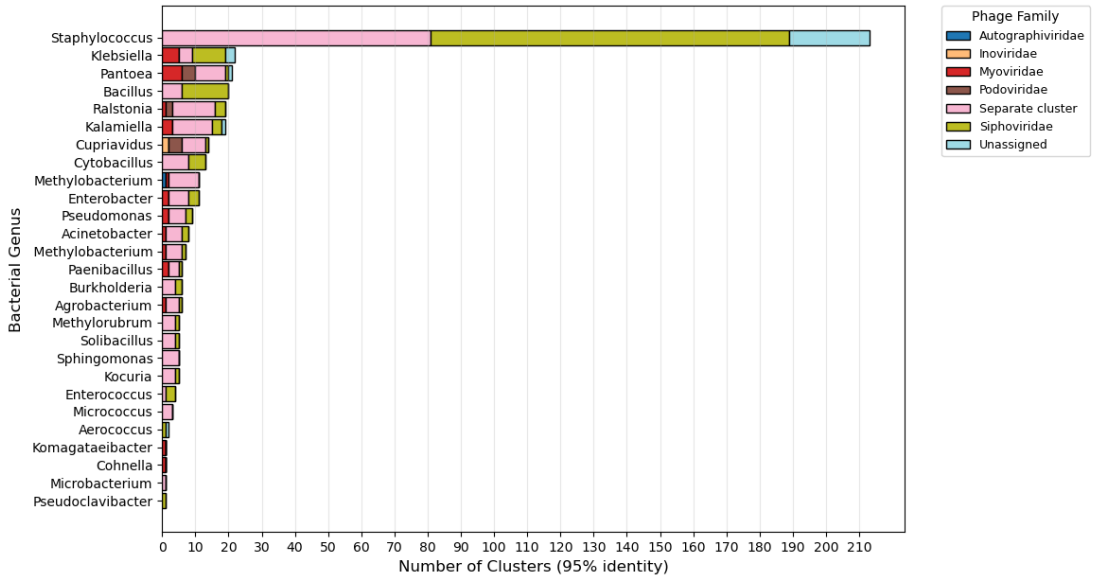


Figure 36. The diversity of identified families with vContact2 and CD-HIT throughout the bacterial genera. The counts of identified 95%-identity clusters are represented.

To further explore the gene-content similarities among the prophage families, a heatmap of the genes by function was created for all 2,217 prophage genomes found (Figure 37). All the functions were categorized by PHANOTATE and they all reference functionality according to the phage. However, if there are any bacterial genes found within prophage genomes, they are categorized among *other*. Since the genes of highest importance in horizontal transfer are antibiotic resistance genes as well as virulence factors, they were categorized into their own category to examine the load of such genes within prophages of the ISS. Table 11 shows the distribution of the total numbers of genes among the whole population.

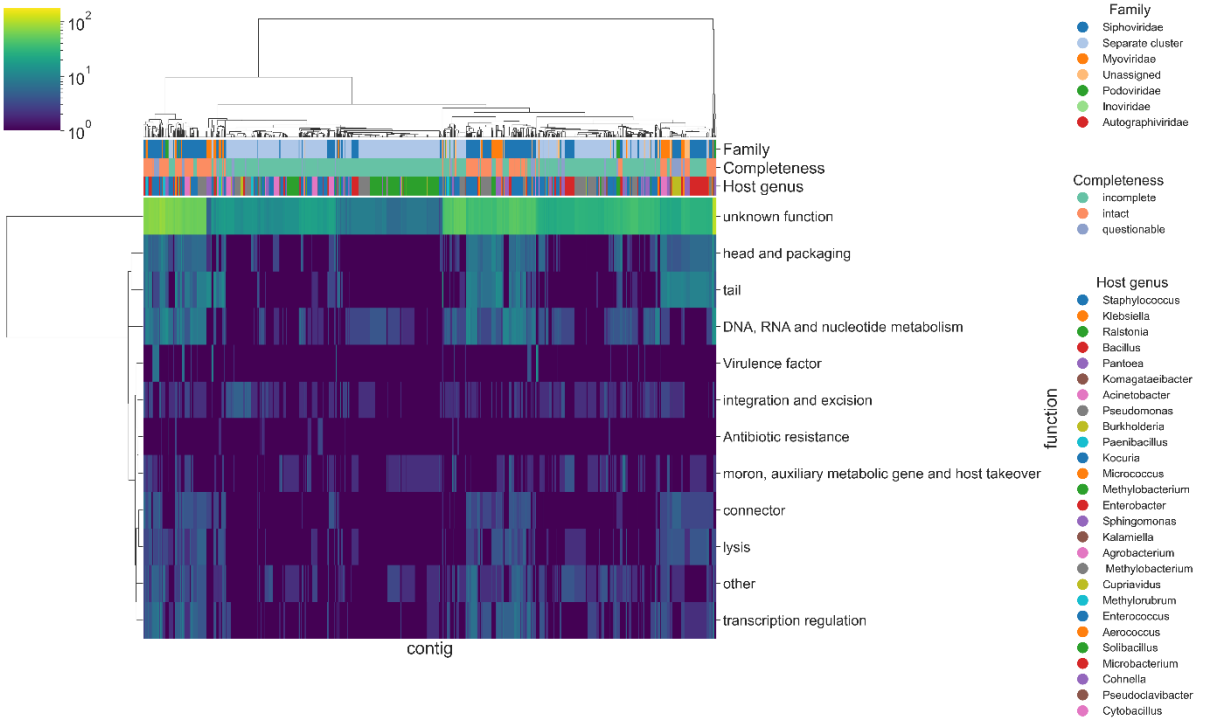


Figure 37. Heatmap of the number of genes categorized according to their function in each identified prophage. The upper three annotation rows depict from bottom-up: Host genus, prophage completeness, and prophage family. The *Separate cluster* prophages are those that could not be assigned to a group as they share no genes with other groups. The Unassigned prophages are those that share genes with reference phages which are themselves not classified into a family.

Table 11. Total number of genes by function found in the whole population of prophages, ordered by descending number of genes.

Function	Total number of genes
Unknown function	61,083
Tail	5,758
DNA, RNA and nucleotide metabolism	5,726
Head and packaging	4,756
Transcription regulation	2,836
Other	2,766
Lysis	2041
Connector	1,913
Integration and excision	1,775
Moron, auxiliary metabolic gene and host takeover	1,412
Virulence factor	517
Antibiotic resistance	128

To observe the distribution of the genes and phage interrelationships within each bacterial genus, a separate gene map including the phylogenetic tree was constructed for each bacterial genus separately. The construction was based on gene homology which was determined by BLAST (e-value 10^{-5}, 50% minimum coverage). The heatmap represented here categorizes the genes according to homology based on BLAST and compares the shared homologous genes across phages. There was a total of 23 total becaterial genera in the dataset. Using Roary and Phandango, for each genus, a pangenomic map was created of shared genes among all the phages of that genus. All the images can be found in Supplementary Collection 1, but Figure 38 shows three bacterial host genera with the most diverse prophage communities – *Staphylococcus*, *Klebsiella*, *Pantonea*.

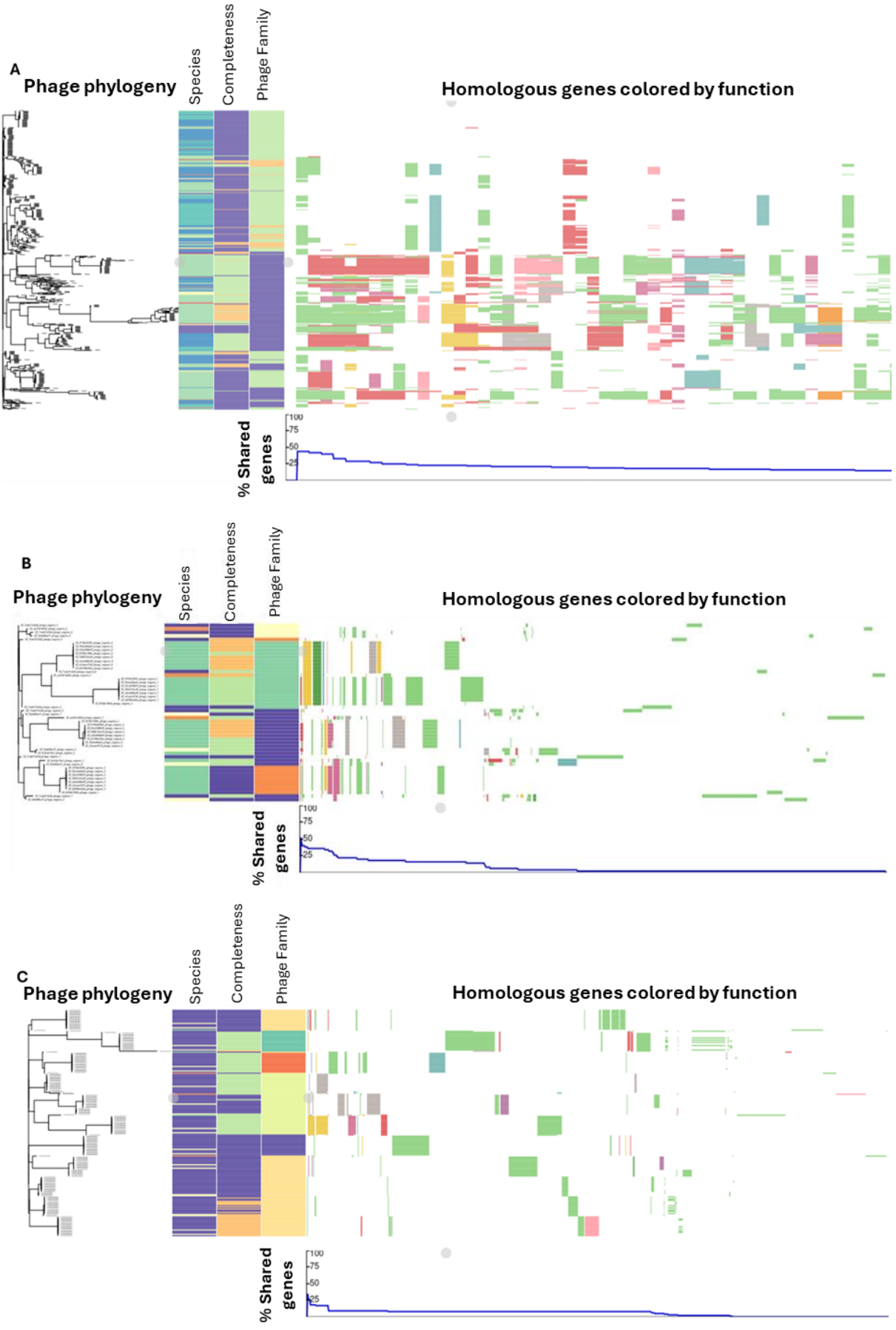

A
Phage phylogeny
Species
Completeness
Phage Family
Homologous genes colored by function
% Shared genes
100
75
50
25
B
Phage phylogeny
Species
Completeness
Phage Family
Homologous genes colored by function
% Shared genes
100
75
50
25
C
Phage phylogeny
Species
Completeness
Phage Family
Homologous genes colored by function
% Shared genes
100
75
50
25

Figure 38. Homologous gene distributions among bacteriophages colored by function of the three most prophage-diverse bacterial genera. A *Staphylococcus*, **B** *Klebsiella,* **C** *Pantonea*. On the far left the phylogeny of the phages was constructed based on the shared homologous genes. The annotation bars are colored by the host species, prophage completeness, and prophage family. The tiles that come after the bars represent individual genes. Each column contains homologous genes while each row contains all the genes from a phage. The tiles are colored by function. On the bottom, a graph represents % of phages that share homologs of each gene. The genes are ordered from the most shared to the most unique. The figures for all other genera are available in the Supplementary Collection 1. The plots were created with modified js code of Phandango v1.1.0 to allow for color-based functional annotation.

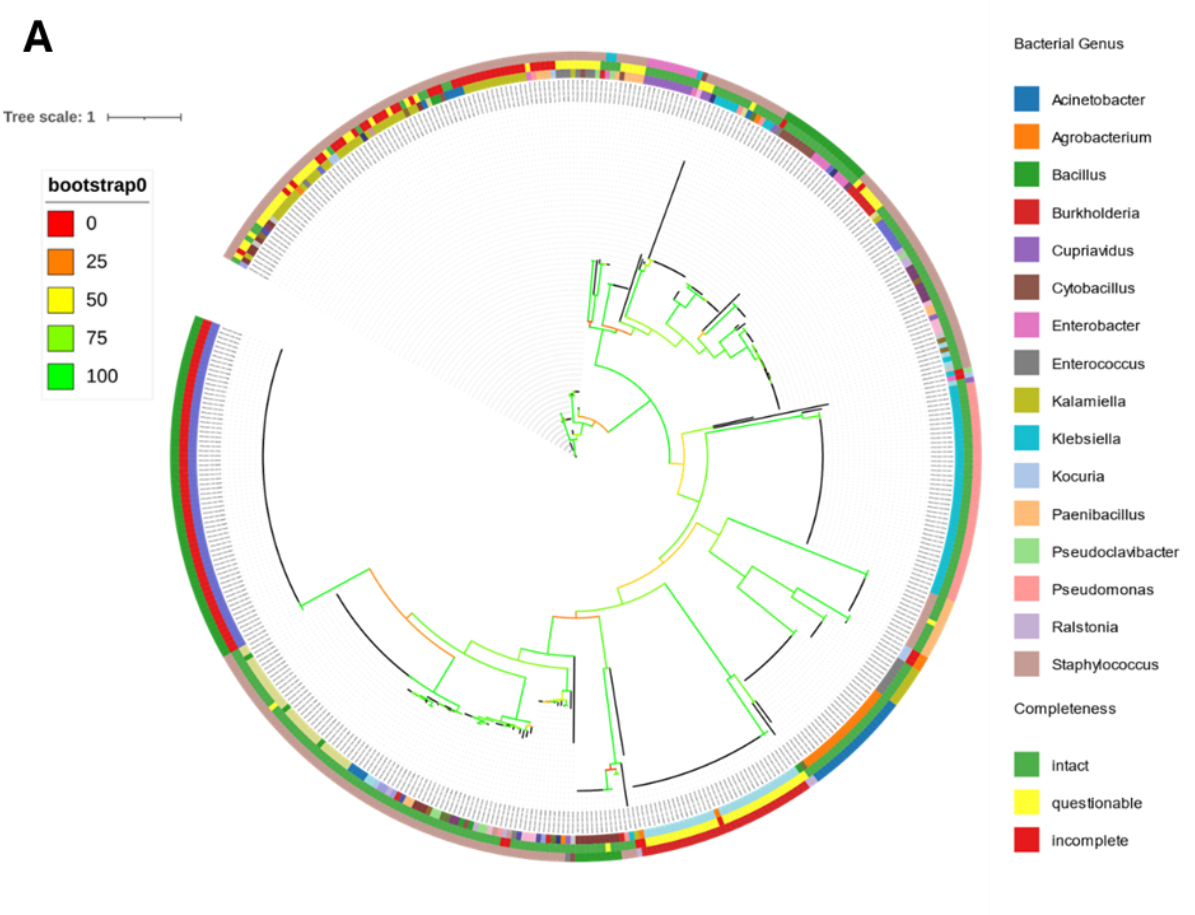
As visible from the results, two major identified families based on genomic content were Siphoviridae and Myoviridae, which are both tailed phage families. Those prophages possess a gene essential for their life cycles, large terminase subunit (terL). This protein packages phage DNA into the capsid. As an essential gene, it is used often in the phylogeny of bacteriophages (Al-Shayeb *et al.*, 2020). As such, orthologs of this gene are present in both families but divergent between the lower phylogenetic groups. Large terminase subunit sequence was used to construct the phylogenies. Within bacteriophage families where those genes were available. Not all members of each family contained the gene. The distribution of terL gene homologs across bacteriophage families is shown in Table 12. Since Autographviridae and Inoviridae contain 1 and 0 terL genes respectively, and in addition only contain 2 prophage genomes, they were removed from phylogenic analysis. Furthermore, after the multiple alignment with MAFFT, it was observed that 18 genes found within Podoviridae and 10 genes from the unassigned phages respectively were identical among each other so the tree could not be constructed (see Supplementary Collection 2 multiple alignments). This left only Siphoviridae, Myoviridae, and the phages forming separate clusters. The MAFFT alignments of terL genes by family, both unprocessed and trimmed, are added to the Supplementary Collection 2. Interestingly, there were in total 5 prophages that contained multiple terL genes. They are listed in Table 13. Both of their genes were used in the analysis. The phylogenetic trees based on terL are shown graphically in Figure 39. Even though terL genes are represented in the trees, the full list of all terL genes and their assigned phage are listed in Supplementary Collection 3.

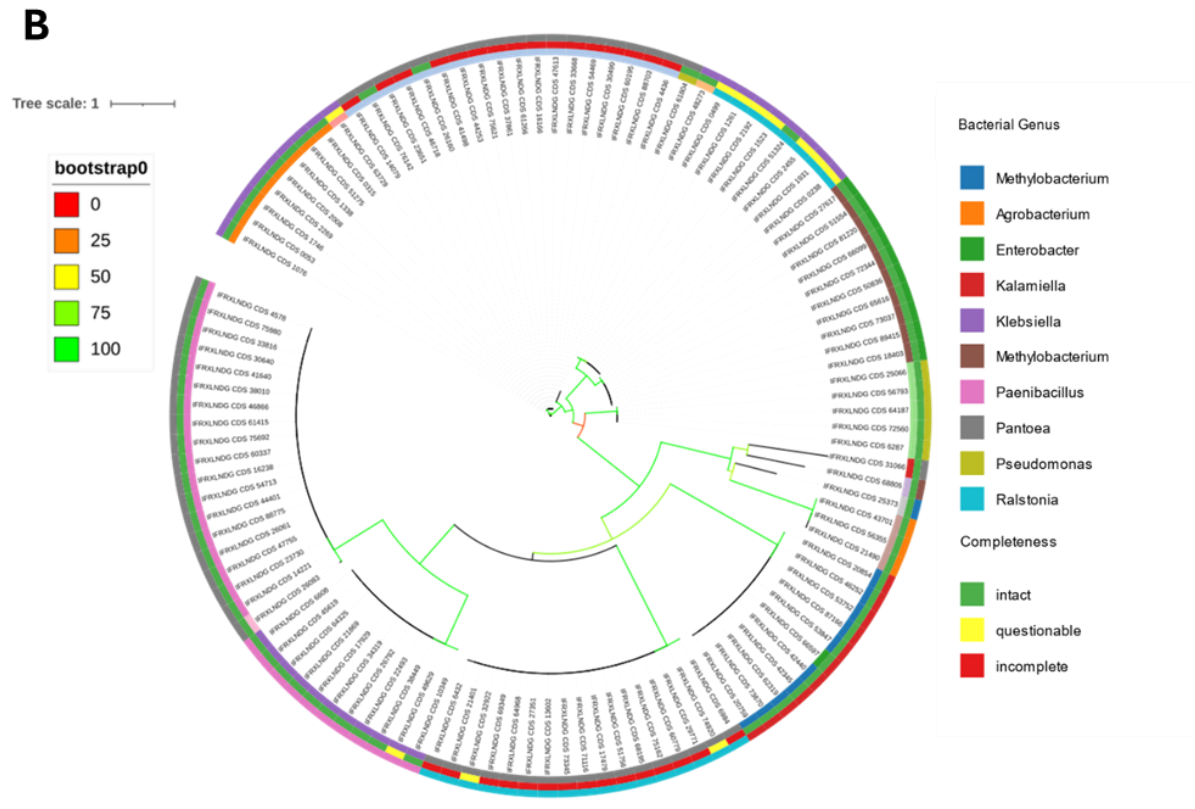
Table 12. The number of terL-containing prophages. This gene was used to construct the mutual phylogenetic relationship between phages of each family.

Bacteriophage family	Prophages containing terL/Total prophages
Siphoviridae	0.6625 (477/720)
Myoviridae	0.6137 (116/189)
Podoviridae	0.4615 (18/39)
Autographviridae	0.5000 (1/2)
Inoviridae	0.0000 (0/2)
Unassigned	0.1230 (10/77)
Separate cluster	0.0581 (69/1188)

Table 13. List of prophages with more than one large terminase subunit homolog.

Prophage label	Number of terL genes	Family	completeness	Host
ZZ_2e2108f013_phage_regions_2	2	Siphoviridae	intact	*Staphylococcus epidermidis*
ZZ_39571e11ef_phage_regions_1	2	Myoviridae	intact	*Kalamiella piersonii*
ZZ_3bf3447a40_phage_regions_2	2	Siphoviridae	questionable	*Staphylococcus epidermidis*
ZZ_6e5a162a73_phage_regions_7	2	Myoviridae	intact	*Kalamiella piersonii*
ZZ_90746a112f_phage_regions_1	2	Myoviridae	intact	*Kalamiella piersonii*


A
Tree scale: 1
bootstrap0
0
25
50
75
100
Bacterial Genus
Acinetobacter
Agrobacterium
Bacillus
Burkholderia
Cupriavidus
Cytobacillus
Enterobacter
Enterococcus
Kalamiella
Klebsiella
Kocuria
Paenibacillus
Pseudoclavibacter
Pseudomonas
Ralstonia
Staphylococcus
Completeness
intact
questionable
incomplete


B
Tree scale: 1
bootstrap0
0
25
50
75
100
Bacterial Genus
Methylobacterium
Agrobacterium
Enterobacter
Kalamiella
Klebsiella
Methylobacterium
Paenibacillus
Pantoea
Pseudomonas
Ralstonia
Completeness
intact
questionable
incomplete

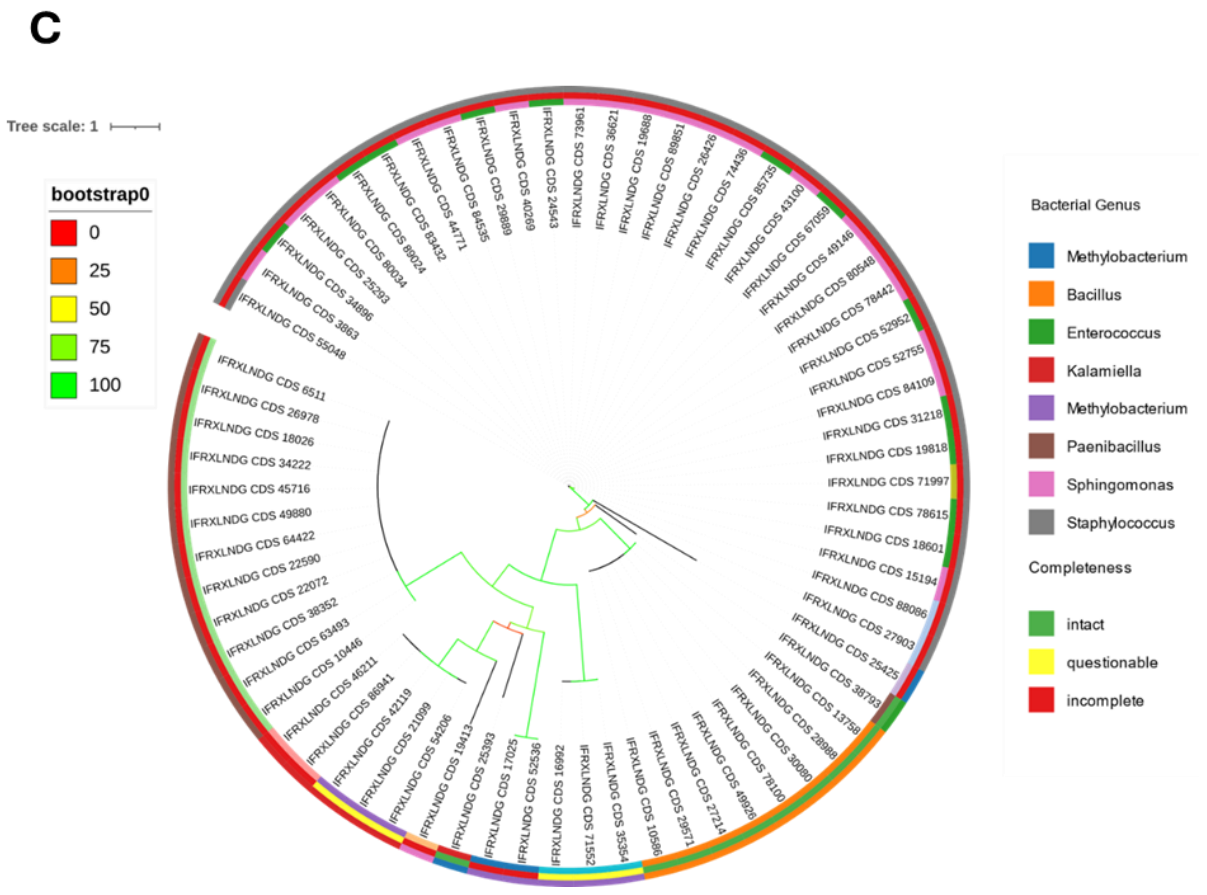


Figure 39. Phylogeny of the terL genes among groups. A Siphoviridae, **B** Myoviridae, **C** Independently clustered phages. Each node is color-coded based on its bootstrap support values after 1000 iterations. The innermost ring represents the cluster of the phages based on 95% genomic average nucleotide identity so those terL genes with the same innermost color likely come from the same phage species. The middle ring represents the completeness of the prophage from which terL comes from. The outermost ring represents the host bacterial genus from which the phage comes from. The 95%-cluster color labels were not included in the legend, as their names do not have any biological meaning and there are too many of them. Still, the list of all the terL genes, their prophages and their cluster labels are provided in Supplementary Collection 2.

Additionally, some antimicrobial resistance (AMR) and virulence genes were found inside the phage genomes. In total, there were 128 AMR genes identified inside temperate phages of 602 bacterial strains isolated from the ISS. Among those AMR genes, there are 16 unique ones, since most of those bacteriophages were closely related (due to the bacterial strains being closely related, such as demonstrated in Figure 33). The table below shows all the unique AMR genes that were identified (Table 14). This is just a descriptive table containing unique genes, and the full table with more details such as top hit proteins, E-values, and known bacteria with the said AMR genes are found in the Supplementary Table 5. The list of virulence genes can be found in Supplementary Table 6.

Table 14. The list of unique AMR genes found in the genomes of temperate bacteriophages of the ISS. The genes that are present in at least one prophage recognized as intact are shown in green, those that have no intact phages but have at least one questionable are shown in blue and those who are all inside incomplete phages are shown in white. In total 16 different genes were found, some in multiple species and some in multiple variants within one species.

CARD short name	AMR Gene Family	Resistance Mechanism	Bacteriophage Family	Bacterial species
oqxB	resistance-nodulation-cell division (RND) antibiotic efflux pump	antibiotic efflux	Myoviridae	Klebsiella quasipneumoniae
oqxA	resistance-nodulation-cell division (RND) antibiotic efflux pump	antibiotic efflux	Myoviridae	Klebsiella quasipneumoniae
Kpne_KpnH	major facilitator superfamily (MFS) antibiotic efflux pump	antibiotic efflux	Separate cluster	Pantoea brenneri
tet(K)	major facilitator superfamily (MFS) antibiotic efflux pump	antibiotic efflux	Separate cluster	Staphylococcus aureus
arlR	major facilitator superfamily (MFS) antibiotic efflux pump	antibiotic efflux	Separate cluster	Staphylococcus saprophyticus Staphylococcus hominis
dfrC	trimethoprim resistant dihydrofolate reductase dfr	antibiotic target replacement	Separate cluster	Staphylococcus saprophyticus Staphylococcus epidermidis Staphylococcus hominis Staphylococcus capitis
FosD	fosfomycin thiol transferase	antibiotic inactivation	Separate cluster	Staphylococcus saprophyticus
mphC	macrolide phosphotransferase (MPH)	antibiotic inactivation	Separate cluster	Staphylococcus hominis **Staphylococcus aureus**
aadK	ANT(6)	antibiotic inactivation	Myoviridae	Paenibacillus polymyxa
AAC6_Ie_APH2_Ia	AAC(6');APH(2'')	antibiotic inactivation	Siphoviridae	Staphylococcus epidermidis
SAT-4	streptothricin acetyltransferase (SAT)	antibiotic inactivation	Separate cluster	Staphylococcus aureus
aad(6)	ANT(6)	antibiotic inactivation	Separate cluster	Staphylococcus aureus
APH(3')-IIIa	APH(3')	antibiotic inactivation	Separate cluster	Staphylococcus aureus
Erm(44)v	Erm 23S ribosomal RNA methyltransferase	antibiotic target alteration	Separate cluster	Staphylococcus saprophyticus
msrA	msr-type ABC-F protein	antibiotic target protection	Separate cluster	Staphylococcus hominis **Staphylococcus aureus**
fusB	Target protecting FusB-type protein conferring resistance to Fusidic acid	antibiotic target protection	Siphoviridae	Staphylococcus haemolyticus

Among the virulence factor genes, a total of 517 genes were found. Of all those virulence factors, sixty-one unique ones were found within the prophages of the ISS. They are listed in Table 15. Majority of virulence factors come inside inactive phage sequences (35/61; 57%).

Table 15. List of unique virulence factor genes found in the prophages of the ISS. The genes that are present in at least one prophage recognized as intact are shown in green, those that have no intact

phages but have at least one questionable are shown in blue and those that are all inside incomplete phages are shown in white.

VFDB gene	Bacterial species	Bacteriophage Family
lukH VFG044014(gb|WP_000791407)	Staphylococcus aureus	Siphoviridae
splD VFG004681(gb|WP_001038766)	Staphylococcus aureus	Separate cluster
selo VFG004978(gb|WP_010922839)	Staphylococcus aureus	Separate cluster
selo VFG004979(gb|WP_147612242)	Staphylococcus aureus	Separate cluster
selo VFG004977(gb|WP_148141455)	Staphylococcus aureus	Separate cluster
hlyIII VFG016234(gb|NP_390062)	Bacillus amyloliquefaciens	Siphoviridae
lukF-PV VFG001276(gb|WP_024937002)	Staphylococcus aureus	Siphoviridae
seb VFG001802(gb|AAA88550)	Staphylococcus aureus	Siphoviridae
scn VFG049876(gb|WP_000702263)	Staphylococcus aureus	Siphoviridae
splC VFG004689(gb|WP_001039022)	Staphylococcus aureus	Separate cluster
selm VFG004986(gb|WP_000821649)	Staphylococcus aureus	Separate cluster
sak VFG004790(gb|WP_000920038)	Staphylococcus aureus	Siphoviridae
lukG VFG044013(gb|WP_000595324)	Staphylococcus aureus	Siphoviridae
rpoS VFG043648(gb|NP_252312)	Pseudomonas granadensis	Myoviridae
scn VFG049879(gb|WP_000702262)	Staphylococcus aureus	Siphoviridae
seh VFG001803(gb|WP_000608674)	Staphylococcus aureus	Separate cluster
lukS-PV VFG001277(gb|WP_000239545)	Staphylococcus aureus	Siphoviridae
ebp VFG004576(gb|WP_000069298)	Staphylococcus aureus	Siphoviridae
sei VFG004998(gb|WP_000721567)	Staphylococcus aureus	Separate cluster
sei VFG004996(gb|WP_000713850)	Staphylococcus aureus	Separate cluster
chp VFG002422(gb|WP_000727649)	Staphylococcus aureus	Siphoviridae
splF VFG004665(gb|WP_001038721)	Staphylococcus aureus	Separate cluster
eap/map VFG004564(gb|WP_000669729)	Staphylococcus aureus	Siphoviridae
eap/map VFG004565(gb|WP_001557458)	Staphylococcus aureus	Siphoviridae
lukE VFG004853(gb|WP_000473596)	Staphylococcus aureus	Separate cluster
hlb VFG004834(gb|WP_000239604)	Staphylococcus aureus	Siphoviridae
chp VFG049868(gb|YP_500656)	Staphylococcus aureus	Siphoviridae
selu VFG004970(gb|WP_000764692)	Staphylococcus aureus	Separate cluster
rpoS VFG000477(gb|NP_461845)	Kalamiella piersonii	Myoviridae
selu VFG004971(gb|WP_000764684)	Staphylococcus aureus	Separate cluster

splC VFG004685(gb|WP_001038860)	Staphylococcus aureus	Separate cluster
splF VFG004666(gb|YP_500436)	Staphylococcus aureus	Separate cluster
lukD VFG004864(gb|WP_000782463)	Staphylococcus aureus	Separate cluster
seg VFG005003(gb|WP_000736712)	Staphylococcus aureus	Separate cluster
seln VFG004981(gb|WP_001236366)	Staphylococcus aureus	Separate cluster
sak VFG004788(gb|YP_500658)	Staphylococcus aureus	Siphoviridae
seln VFG004983(gb|WP_001235656)	Staphylococcus aureus	Separate cluster
splB VFG004694(gb|WP_001039452)	Staphylococcus aureus	Separate cluster
ebp VFG004572(gb|WP_000069282)	Staphylococcus aureus	Siphoviridae
splE VFG004673(gb|WP_001038759)	Staphylococcus aureus	Separate cluster
sak VFG004787(gb|WP_000919350)	Staphylococcus aureus	Siphoviridae
selk VFG004995(gb|WP_000733775)	Staphylococcus aureus	Unassigned
seln VFG004982(gb|WP_001236362)	Staphylococcus aureus	Separate cluster
splA VFG004701(gb|WP_001039435)	Staphylococcus aureus	Separate cluster
seg VFG005004(gb|WP_000736707)	Staphylococcus aureus	Separate cluster
selq VFG004975(gb|WP_001033316)	Staphylococcus aureus	Unassigned
splD VFG004678(gb|WP_001038704)	Staphylococcus aureus	Separate cluster
ABK27166 VFG043786(gb|ABK27166)	Staphylococcus aureus	Separate cluster
A79E_RS10215 VFG049313(gb|WP_002910453)	Kalamiella piersonii	Separate cluster
sea VFG001325(gb|WP_000750412)	Staphylococcus aureus	Siphoviridae
splC VFG004684(gb|WP_001038867)	Staphylococcus aureus	Separate cluster
sei VFG004997(gb|WP_000713847)	Staphylococcus aureus	Separate cluster
splB VFG004693(gb|WP_001039447)	Staphylococcus aureus	Separate cluster
eap/map VFG004566(gb|WP_001793419)	Staphylococcus aureus	Siphoviridae
selm VFG004985(gb|WP_000821658)	Staphylococcus aureus	Separate cluster
splA VFG004703(gb|WP_001039427)	Staphylococcus aureus	Separate cluster
tsst-1 VFG001809(gb|WP_001035599)	Staphylococcus aureus	Unassigned
nuc VFG004731(gb|WP_002470235)	Staphylococcus epidermidis	Siphoviridae
adsA VFG049854(gb|WP_000645753)	Staphylococcus aureus	Separate cluster
selp VFG004976(gb|WP_025174055)	Staphylococcus aureus	Siphoviridae
sea VFG005012(gb|WP_000750406)	Staphylococcus aureus	Siphoviridae

Finally, with the uncovered diversity of AMR and virulence factor genes present in the bacteria onboard ISS, a natural question remained how likely the phage transfers are. Particularly,

phage transfers between genera, species, or inside species. Therefore, to confirm if there were any observable transfers available, phages were again clustered according to 95% and 100% identity. Those phages belonging to the same cluster were considered to originate from the same phage. Under the assumption that a phage that transferred from one bacterial cell to another, will integrate its genome to a random new locus once inside a new host, the local DNA sequences around the similar phages were compared. As shown earlier, most bacterial sequences came in contigs. The table with identity values across contigs and genomes containing the phages of the same cluster was produced, one for 95% identity clusters and the other for 100% identity clusters. Additionally, to confirm the results, the table was created based on the genbank files, presenting 10 upward and 10 downward coding genes (CDS) from the phage locus. The proteins coded by all those genes were compared with DIAMOND algorithm with threshold e-value 10^{-5}. If less than 5 genes among 20 matched between phages in the same cluster, those phages were considered to be exchanged. The tables with comparisons of the contigs within each cluster are available in Supplementary Tables 7 (95% identity clusters) and 8 (100% identity clusters). The results of fastANI comparison showed that when 95% identity clusters are considered, the minimal identity in the regions where similar phages are is 89.58%. With 100% identity clusters, the minimal identity between the contigs is 97.57% (Supplementary tables 7 and 8). Those identities are too high to be considered distinct loci. Still, the genes in each contig are provided in the Supplementary Table 9. The results of DIAMOND gene comparisons are included in Supplementary Table 10.

3.4. Discussion

3.4.1. Description of the prophage population in the context with the bacterial hosts

A high number of prophages were found in the bacterial community of the ISS based on the available open-source data. This is not surprising as it is widely considered that most bacteria in the biosphere contain multiple prophages, so they are called polylysogens (Silpe *et al.*, 2023). In this study, their diversity, gene content, and relationships were explored, taking into context the host species.

From the distribution of the determined phage activity in Figure 21, it is notable that the higher part of phages was active in single-entry genomes. This might be because many of the multiple-entry genomes are assembled up to the contig level, leaving potential phages in the genomes incomplete. However, it must be noted that the number of single-entry genomes was significantly lower - only 13 compared to 589 of multiple entry genomes. Still, in those 13 genomes, 45 phages were discovered. This raises a question: What is the distribution of the prophages among the bacterial genomes? Only 5 out of 602 genomes (0.83%) had no determined prophages inside them.

The analysis of variance between the groups showed that there are significant differences between the number of active, inactive and questionable prophages in the microbial community of the International Space Station (Figure 22 A). The majority of the sequences were determined to be inactive prophage sequences with incomplete genomes. Also, the analysis shows that, as would be expected, complete prophages have significantly longer genomes than questionable and incomplete ones. Additionally, even though bacteria in general tend to have a similar number of prophages in their genomes on average, the distribution of active, questionable and inactive phages differ significantly. This hints towards distinct sub-populations of bacteria with different distributions of intact, questionable and inactive prophages.

The correlation analysis between the host genome length and phage genome length in Figure 23 shows that no significant correlations were found. This shows that viruses with bigger, more complex genomes are not necessarily found in larger bacterial genomes and vice versa. This is the effect of viral specificity towards a host, as it is obvious that prophages are not randomly distributed across species. The correlation analysis between the bacterial host genome length, and the number of phages found per bacterial genome is found in Figure 24. No significant correlation was found between the phage length and the host genome length. This shows that in the microbiome the bigger phages do not necessarily infect bigger genomes. In this way, it was shown that the phages of any genome length can be found in bacteria of any genome size. Also, no significant correlation between phage counts and the host genome length was found, showing that bigger genomes do not necessarily host more prophages. As expected, this hints towards biological interactions within the community. The two interesting results from this analysis were genus *Pseudomonas* and *Acinetobacter*. *Pseudomonas* shows low genomic variation between species. There are two species present – *Pseudomonas fulva* and *Pseudomonas granadensis*. *P. fulva* has a smaller genome, a mean of 5,220,950.186 (±1,434.402) bp, yet hosts 7 phages per genome. *P. granadensis* on the other hand, has a mean genome length of 6,075,650.000 (±602.449) bp, yet contains only 2 phages per genome. From *Acinetobacter*, we can observe only a single species, and that is *A. pitii* where all the members found have exactly 6 phages. Interestingly, it is not the case that the same 6 phages are shared among all 21 genomes. This would be expected if all the genomes are highly related or there is no activity between them regarding the bacteriophage exchange. Rather, even though every one of the 21 *Acinetobacter* genomes has 6 phages, they classify into 8 distinct clusters with a minimum of 95% identity (Figure 25). This means that the lack of variation in the phage number is not a result of faulty data, like strong relatedness of the genomes, but rather shows an interesting phenomenon.

3.4.2. Analysis of the clusters of related prophages

The heatmap of phage gene cluster numbers categorized according to biological function within the ISS bacterial community is represented in Figure 31. As with the number of phage clusters from the previous analysis, *Staphylococci* phages show by far the highest gene diversity among all the genera. This is so notable that they cluster as a separate group completely away from any other genus (the tree on the left in Figure 31). It is worth pointing out a staggering number of gene groups of unknown function among only *Staphylococci*. Only in that genus, there are 3,770 gene groups of unknown function, demonstrating just how little it is known about phage gene function, especially in complex communities such as the *Stapylococci* of the ISS. Among other genera, *Klebsiella*, *Bacillus*, and *Pantonea* phages show significant diversity with high number of gene clusters that fall into each functional category. At the same time, we can see the bottom six genera (*Micrococcus, Solibacillus, Microbacterium, Pseudoclavibacter, Cohnella,* and *Komagataeibacter*) being very poor in diversity compared to others. The genes of the unknown function have the highest prevalence of all in the population. This is followed by the tail genes, DNA/RNA with nucleotide metabolism, and then head and packaging genes. Combined with the previously presented distribution of phage genomic clusters, the addition of gene functional diversity paints a clearer picture of the prophages at the ISS.

Looking at the distribution of 95%-identity clusters among species (Figure 25), it is visible that *Staphylococcus epidermidis* and *Staphylococcus aureus* vastly dominate the diversity of prophages within their genomes out of all bacteria isolated from the ISS. The former has 80 different clusters while the latter has 67. The next highest-diversity species is *Kalamiella piersonii* with only 19 clusters. However, only a portion of their prophages are determined to be active by PHASTER algorithm. It can also be observed, as would be expected, that the portion of intact prophages does not depend on the diversity, as there are species with low diversity but high portion of active or intact prophages, such as e.g. *Pseudomonas granadensis* (2 clusters, 100% intact), while there are species with high diversity and low portion of intact prophages such as *Ralstonia pickettii* with 14 different clusters and 2.77% intact phages (9/325). Still, the number of clusters is a strong indicator of prophage diversity, as even the inactive phages carry genes that can serve as a gene reservoir for similar phages to recombine and acquire new genes.

This is interesting as phages usually have their host range, rarely spanning multiple genera. Therefore, each bacterial genus could be observed as its own viral ecological niche. There were some interesting differences observed when comparing the species- and genus-level distribution of prophages. For example, despite *Kalamiella piersonii* having some of the highest diversity of phages (first after *S. epidermidis* and *S. aureus*), the genus *Kalamiella* is only 6th

by the phage diversity among bacterial genera. This is because different genera have varying numbers of species and, interestingly, each genome has its own phage content. When the plots are compared based on the total number of prophage clusters (Figure 25 and 26) and the number of prophage clusters per bacterial genome (Figure 27 and 29) striking differences can be observed. Even though *Staphylococcus* species such as *S. epidermidis* and *S. aureus* strikingly dominate with the total number of prophage clusters, their per-genome diversity is average to low. Specifically, while *S. epidermidis* and *S. aureus* show 80 and 71 clusters respectively, they show only 2.43 and 3.61 average clusters per genome. Compared to *Pantoea brenneri* and *Pantoea sp.*, which have only 14 and 12 total unique clusters while having 11.06 and 10.5 per-genome clusters, the difference is significant. This points to a bias in phage discovery due to the number of genomes available from the ISS. Since the genome data is heavily biased towards *Staphylococcus* species, the highest number of discovered phages come from that genus. However, when accounted for the number of genomes, *Staphylococci* do not show a higher diversity of phages. Therefore, it is a strong sign that with more genomes being sequenced, more prophages and their genes will be discovered in more species.

When the phage diversity is observed on the genus level (Figure 29), many genera show 0 standard deviation. This further confirms the assumption that the genome data possesses a bias towards some bacterial species over others. Only the samples of isolated and sequenced bacterial species and strains were analyzed. It is likely that some yet non-isolated bacterial species could be present in the ISS environment but since their genomic sequence is unknown, they were not included in the analysis. This means that some rare genera or species are present in the dataset in low sample numbers. Still, the assumption is that all the species included make up the majority of the ISS population. As more and more genomes are added to the database, the dataset will be able to be refined. It is visible in the heatmap of Figure 30, the genus *Pseudomonas* does not show the difference to so many other genera as did *Pseudomonas fulva* against all the other species. This points to a significant importance of *P. fulva* as an independent niche for prophage diversity, compared to other *Pseudomonas* species.

3.4.3. Prophage genome annotation and classification

Heatmap of the tRNA gene counts found in the phage genomes of the ISS are represented in Figure 32. The genes encoding tRNAs are interesting for virus evolution because once they acquire those genes, it sets their protein-coding potential on a separate evolutionary trajectory (Carbone, 2008). This is because those genes are free to mutate and evolve as long as the host Is providing a functional tRNA substitute. Additionally, the spare tRNA genes can be used as a defense mechanism against tRNA-targeting antiviral defense from the host (van den Berg

et al., 2023). Nevertheless, those new tRNA genes are free to horizontally transfer between bacterial hosts, adding new diversity and coding potential into the pool of tRNA genes. *Staphylococcus* is by far the richest genus considering the tRNA genes. Interestingly, despite being among the richer genera considering phage diversity, *Bacillus* genus only has one distinct tRNA gene. The tRNA genes for most biological amino acids are found inside prophage genomes across this microbiome. The only missing tRNA genes that carry amino acids within phages are tyrosine, lysin, and isoleucine. It seems that serine, arginine and asparagine-carrying tRNA isotypes are prevalent in high diversity through the population.

Figure 33 shows a phylogenetic tree of the identified bacteriophages from the hosts with complete genomes. As expected, the phages of the same host are more related to each other, suggesting no cross-species infection. However, looking at *Klebsiella quasipneumoniae*, it is observable that all the individual full genomes share phages, suggesting their close relatedness, and even reflecting the host phylogeny. Each phage on the tree is labeled and color-coded based on the genome from which it comes from. The identical phages are represented on the tree as groups of leaves at the edge without branching. With *S. aureus* phages, a different result is observed. The strains have limited shared phage sequences and therefore show less relatedness among *S. aureus* strains and more prophage activity. This tree demonstrates the difference in the prophage dynamics between the represented two bacterial genera. *Staphylococcus* genus has a more active set of phages which might be driven by their own genomic diversity and ecological interactions. The result is in line with what was shown previously that *Staphylococci* possess a rich diversity of prophages onboard the ISS.

The clusters of prophages according to the shared gene content are represented in Figure 34. Notably, the prophages of the ISS tend to form clusters, rather than being dispersed across the network. This points to the limited overall diversity and high level of relatedness of the temperate phages on the ISS. Arguably the highest number of clusters came from *Staphylococcus* and *Bacillus* bacteriophages as shown in the Supplementary Figure 1. With more reference genomes added to the database, the resolution of the clustering will be higher and the classification better. In this classification, most of the prophages that were classified into known families fall within *Siphoviridae* or *Myoviridae*. Other families found were *Podoviridae*, *Inoviridae* (filamentous), and *Autographviridae*. The original classification according to families comes from vContact2, after analyzing the results with graphanalyzer v1.6.0 (Mattia *et al.*, 2022) (Figure 35).

The family-level phage distribution among bacterial genera is represented in Figure 36. From the results, we can see that even though *Staphylococcus* possesses highest diversity on the lower levels such as genus/species, at the family level, they are among the ones with lower diversities. Their structure is dominated by *Siphoviridae*, and the highest portion of all

Siphoviridae falls among *Staphylococcus* hosts. Other genera contain mixes of mostly unidentified prophages or *Siphoviridae* and *Myoviridae*.

The heatmap of the gene content by phage is shown in Figure 37. It could be observed that the prophages of the ISS possess high genetic diversity, even inside families, as the families are scattered across the heatmap. As observed previously, the dominant genes in number and diversity are those of the unknown function, which is not surprising, considering that this is a very generalized category, and for many of the genes, function could not be determined based on gene homology. When looking at the family annotation bar, it is clear that the phages that could not be assigned to any family (*Separate cluster*) are significantly lacking genes in comparison to those that are classified successfully. This is also not surprising considering that classification relies on the shared gene content. Still, it shows that those prophages passed unclassified at large part due to the lack of genes, rather than representing unknown families. Considering the classification of the functions, all the functions were categorized by PHANOTATE and they all reference functionality according to the phage. However, if there are any bacterial genes found within prophage genomes, they are categorized among *other*. It is important to keep in mind that, even though it is informative to visualize the gene content on per-phage level, the raw numbers of genes get biased based on the distribution of bacteriophages in the population. For example, a high number of genes of a certain function can be a result not of real diversity, but rather because there might be many copies of the same phage in a dataset, which could result from biased sampling. For that reason, the better view of gene diversity is provided in the heatmap in Figure 31 where proteins are clustered based on their 75% identity, ensuring the copies are not considered in the analysis. Nevertheless, the per-phage heatmap shows a remarkable functional diversity of the prophage genomes of the ISS. Even though the complete heatmap such as the one in Figure 37 presents the global state of gene diversity, some finer level details are hard to see. Clustering phages in that heatmap does not indicate true evolutionary relationships because the genomes are clustered based on the content of genes falling into either of the 12 gene functional categories. Since those categories are very general, they clearly do not represent the homology between genes.

Figure 38 represents a pangenomic map of three bacterial genera with the most diverse prophage groups. The first thing that can be noticed by looking at the trees based on the shared genes is that many groups of phages have completely identical gene sets. This is visible at the leaves of the tree that have no branching. These phages are considered copies, and they bias the diversity results from the heatmap above. This is the reason why homology-based classification is better than just functional classification. However, homology-based classification makes more sense among the more closely related bacteriophages since

unrelated phages might share 0 homologs together. For that reason, it was decided to focus the homology-based plots on each host genus separately, each representing an ecological niche. Comparing the plots, the mosaicism of the prophage genomes, particularly in *Staphylococcus* phages is clearly visible. Nevertheless, this mosaicism can be seen in all represented genera. A strong sign of it is a gap between genes in a column. This means that the gene is shared between two distinct groups of phages. Another interesting point that can be observed is gene re-functionalization. It is observed when genes in the same columns have different colors, meaning that a homologous gene in different species has distinct functions. This can be strongly observed in *Staphylococcus* phages. Additional curiosity that can be seen from the graphs is that there is no gene among all the phage species that is conserved among the whole phage population. There are genes that are shared by 50% or more of the prophages in the whole population, but none is shared by all the prophages. This makes distinguishing their phylogeny more complex. Still, it is clear that *Klebsiella* and *Pantonea* phages have significantly more ordered genomes among prophages than *Staphylococci* when function and homology are considered. This is a sign of less mutual evolutionary interactions and less recombination between the prophages. This still goes to show that *Staphylococci* are dominating the diversity of prophages in numbers, gene content, and the interactions among prophages in general. On the other hand, in the species column it could be seen that many phages that share most of the genes among each other come from different host species. This is also a strong indicator of interchange between prophages in the population of the international space station, though it is not the absolute evidence of current or recent phage jumps. Rather, it shows that during time and evolution, prophages were jumping between species. Additionally, again, it is visible that many of the phages labeled as incomplete truly lack genes, but still, they share some homologs with other phages. This shows that most of the incomplete phages have originated from those that were found but due to faulty excision, recombination, or incomplete sequencing data, they lost some genes that are essential for their function.

When observing bacteriophage pangenome within each host genus (Figure 38), it is clear that the major functional diversity is present in *Staphylococcus* genus compared to other genera. The plot elucidates the striking genetic mosaicism among phages. This is not surprising, considering that phages are known for their mosaic genomes (Hendrix *et al.*, 1999; Pedulla *et al.*, 2003; Cresawn *et al.*, 2011). However, it is surprising to observe the functional divergence between the homologous genes. Those changes are present at higher phylogenetic levels. This is visible because the phages that show homologous gene re-functionalization are not closely related. This gives evolutionary time for mutations to affect the protein structure and function. Still, the fact that we observe high functional diversity is a sign of ongoing adaptations and evolution. Additionally, the wide interrelationships of the prophages present in the ISS

microbiome are seen on the line plots at the bottom of the graphs. They represent % of all phages that share each gene. No gene is present across all the prophage genomes in any of the bacterial genera tested. The genes that are most shared by any genera are up to 75%, and they are present among *Enterococcus* and *Sphingomonas* phage communities. Overall, this representation shows more accurately the high phage diversity present on the ISS.

In the phylogenetic trees made based on terL gene sequence (Figure 39) it is visible that most of the trees were supported by high bootstrap values of more than 75%. This is good support for the presented phylogeny. It is also visible on all three represented trees that the phages infecting the same bacterial genera are more likely to be closely related (i.e. cluster together). This is visible from the colors in the outer circle. This confirms the assumption that there is no observable transfer of species between host genera. Potentially, with the discovery of new bacterial isolates from the ISS and novel prophages, this phenomenon could be observed. However, on the data presented here, no inter-host-genus transfers were found. All three trees showed a group of identical terL genes, visible on the tightly connected leaves. This points to close relatedness between many groups. However, among all, *Siphoviridae* showed the most complex tree which is expected as they are the most numerous phages.

The screening of antibiotic resistance genes among the prophage genomes of the ISS showed fascinating results (Table 14). Some gene homologs are present in multiple different variants such as dfrC or fosD. Aligning with the results of phage diversity, most of the AMR genes come from *Staphylococci* phages. Also, most of the unique AMR genes (9/16; 56%) come inside the phages deemed intact, meaning that they can still potentially spread through the population. The functionalities of the genes bringing the antimicrobial resistance vary between antibiotic efflux, target replacement, antibiotic inactivation, target alteration, and target protection. Even though not all of the genes are present in the active phages, it is important to note that all the bacteria detected with AMR genes are found in the pathogen database (Bartlett *et al.*, 2022), showing that the hosts can cause illness. This is not confirmation that those specific strains would in reality cause illness. Nevertheless, the astronaut immune system is suppressed in spaceflight so there is still a possibility of infection, even if on Earth those bacteria might not cause it. This means that an AMR gene does not need to spread through the population through phage-assisted transduction. It is enough that the circumstances align in a way that even a bacterium with an inactive phage and AMR gene could cause an illness. Additionally, those genes could still spread through the population by means of other phages and recombination. If a similar phage infects a bacterium already containing an inactive phage with the AMR gene, a recombination could occur, and an active phage could acquire it. Therefore, the fact that a gene is inside an inactive phage does not imply it could not spread. It is worth noting that this analysis was only performed on bacteriophage sequences inside the bacteria

isolated from ISS. No analysis was done outside of the phage genomes. This means that those antibiotic resistance genes found here likely do not represent the full spectrum of AMR genes found onboard ISS, but only those that are available for phage-assisted transduction.

Looking at the virulence genes present in the prophage genomes, most of them come from *Staphylococci* phages, and in particular *S. aureus.* This just confirms the magnitude of the diversity of this genus present on the ISS. Again, all the species showing the presence of virulence factors are present in the human pathogen database (Bartlett *et al.*, 2022). Therefore, similarly to AMR genes, the fact that they are not present in the active phages does not imply they cannot spread through the population. Still, knowing the spectrum of virulence factors onboard ISS adds to the knowledge and expectations of what bacteria we could expect to become pathogenic or even resistant.

Taking the results of phage jumps together, no evidence was observed of horizontal bacteriophage transfer among the bacteria of the ISS. That said, an important limitation is that the approach used here relied on the assumption that a phage will integrate into a random locus in a new genome. For every bacteriophage 100% and 95% identity cluster, nucleotide sequences of the contigs that phages were part of were compared with fastANI. If the contigs would differ significantly from each other, while the phages are the same, it would be concluded that the phage changed location, and therefore, the host. The schematic explanation is found in Figure 40. All the comparisons between the phage contigs belonging to the same cluster showed also high degree of similarity, pointing out that the phages were not exchanged. However, if a phage finds a specific homologous region to integrate, those phage jumps would have been missed because the sequence identity and gene homology would be similar between loci. Still, the fact that no phage exchange was detected is encouraging but needs to be taken with caution. Bearing in mind that in the case of phages targeting similar sites, the phage jumps would be missed. To detect such cases in the future for any microbiome, a more thorough analysis needs to be done.

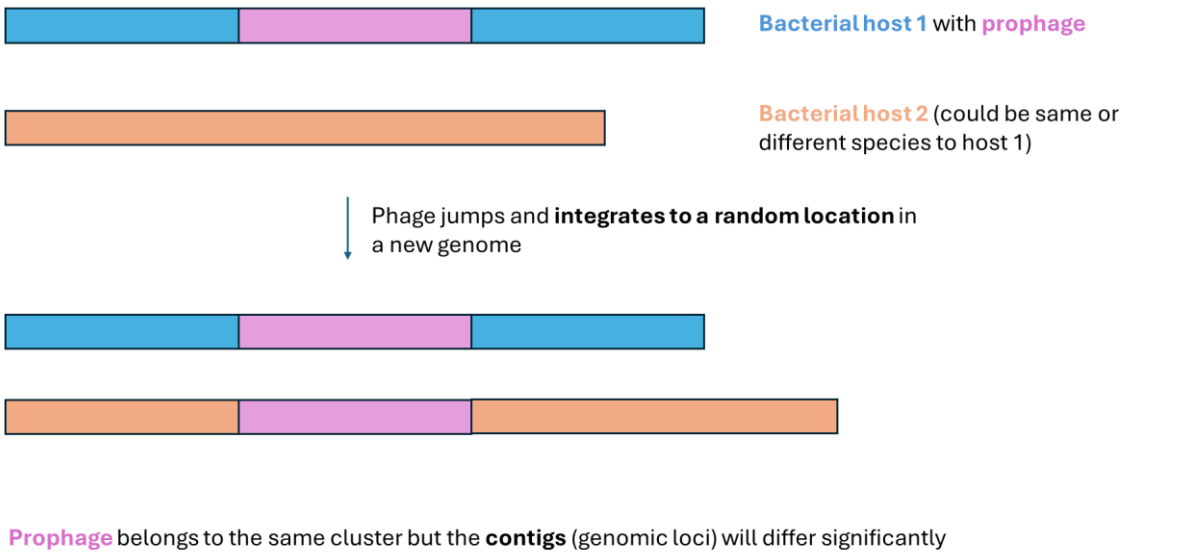


Figure 40. The rationale behind the search for prophage exchange among the population of bacterial genomes. First the clustering of phages according to their identity is performed, and then the contigs or the local genetic context is compared. Those phages that belong to the same cluster, but their locations differ, are considered to be exchanged.

The exploration of prophage diversity in an environment is not a commonly used type of analysis for environmental microbiome. The genomes of the identified bacteria are usually analyzed for their function with limited or no attention to bacteriophage sequences. However, the large-scale analysis by Touchon et al. reveals that on average bacteria have 1-2 prophages per genome with high distribution, ranging from 0 to more than 8 (Touchon, Bernheim and Rocha, 2016). Here an average of 3-4 prophages per genome was found. A similar type of study would be interesting to perform in public transport, to determine the amount and distribution of antibiotic resistance genes that can potentially spread. Also, the knowledge of prophages inside bacterial genomes of environments is interesting as phages are potent regulators of bacterial populations.

3.5. Conclusion

Through the set of sequence analysis steps, the prophage content among the bacteria isolated from the ISS was explored. Most of the identified phages could not be tracked to any known closely related phage. That is not unusual for phages as new species are discovered frequently.

Nevertheless, the phage diversity was shown to be very high within the community of the ISS. The gene diversity and interrelationships of bacteriophages were determined based on the shared gene content and the TerL protein sequence among the phages that contained it. Overall, the phage diversity was shown to be dominated within a few bacterial genera. The

highest number of unique genomes was identified within *Staphylococcus* genus, which are bacteria inhabiting human skin. This shows the importance of humans as agents that introduce new diversity to microbiomes. Pangenomic analysis among the phages within each genus also showed very diverse functions of related genes, indicating re-functionalization. The tRNA genes encoding most of the amino acids were found among the phages, enabling the divergent evolution of the genetic code.

With virulence and antibiotic resistance genes, it was shown that most phages containing those genes are within bacterial species that are on the human pathogen list. This does not mean that they would cause a disease, but it implies that in the right conditions, they could. This is especially concerning the ISS where occupants are expected to have reduced immune function.

No evidence was found of phage jumps between the host species or even intraspecies. This however does not confirm that the jumps are absent, in the same way as this analysis takes only the part of the microbiome that has their genomes sequenced.

A big limitation of the study is that no isolated samples were acquired to induce the activation and production of the discovered bacteriophages and without a question confirm the findings experimentally. This question remains open to future research. In the future, fluoroquiol or mitomycin C could be used to induce prophages from the strains of interest. If those phages were induced in the laboratory environment, it could lead to potential discoveries of new bacteriophage species, genera, or even families.

Chapter 4: Bacteriophage assembly under simulated microgravity conditions

4.1. Introduction

It is of great benefit to understand the virus behavior in microgravity. For one, it can lead to a better understanding of virus behavior in space travel, allowing better and more informed risk assessment concerning viral diseases. Additionally, viruses and their self-assembling capsids are heavily used in biotechnology and pharmacology. For example, bacteriophage therapy against bacterial infections is becoming more relevant as antibiotic resistance in bacteria becomes more prevalent (Stacey, De Soir and Jones, 2022; 'The promise of phages', 2023; Strathdee *et al.*, 2023; Cui *et al.*, 2024). Additionally, viral capsids are heavily used in gene therapies and smart drug delivery (Wang, Tai and Gao, 2019).

Biotechnology and pharmacology are becoming more and more relevant for application in human space missions. In the deep space, there will be limited supplies, as well as resupply opportunities. Therefore, being able to synthesize the phages on missions can be useful. They are used as vectors or therapies on Earth and could be one day in space. Therefore, understanding the synthesis of viruses in a microgravity environment will show the applicability of the synthesis of viral capsids in microgravity.

In this research, a virus, bacteriophage T7, was synthesized in vitro under simulated microgravity conditions by clinorotation. To determine if the virus was synthesized to a greater extent, lower, or there is no difference in the assembly. This research supports basic science by showing the phage behavior in microgravity. Additionally, it will be beneficial to determine whether clinorotation influences phage production. Therefore, despite being space-focused research, it finds great application on the ground.

Bacteriophage T7 is a complex phage that can self-assemble. It has been long used as a phage model. It is a dsDNA phage with a tail and head. The order this phage belongs to, *Caudovirales*, has a conserved prohead assembly and DNA packaging mechanism which is shared with herpes viruses. As discussed in the introduction, herpesviruses are a major problem for astronauts. Understanding the assembly of those viruses in microgravity would lead to possible explanations for the reactivation of those viruses in spaceflight, especially in zero gravity. T7 infects *Escherichia coli* and produces plaques within 2-3 hours after infection. Therefore, this phage was a suitable model to evaluate the effect of simulated microgravity on the production of complete phage virions.

To synthesize the phages in vitro, the transcription/translation solution (TX/TL) was used. It contains the extract from the host bacteria (*E. coli*), including the ribosomes, amino acids, RNA

polymerase, and other enzymes. This gives it the capacity to produce proteins based on the DNA template. A self-assembling bacteriophage such as Enterobacteria Phage T7 can be completely synthesized in vitro, starting from genomic DNA. The assembly of the phages in vitro is therefore equivalent to the assembly inside the cells. Viral production could also be observed inside the bacterial cells. However, bacteria have also been known to react to simulated and real microgravity (Klaus, Todd and Schatz, 1998; Brown, Klaus and Todd, 2002; Baker, Meyer and Leff, 2004; Benoit and Klaus, 2007). Therefore, any effects observed on the viral production could be due to the reaction of the bacteria. By using in vitro phage synthesis, the bacterial host factor is removed while studying the effect of microgravity on phage assembly. Therefore, the TX/TL solution is a good testbed for assessing the bacteriophage assembly in simulated microgravity (Figure 41).

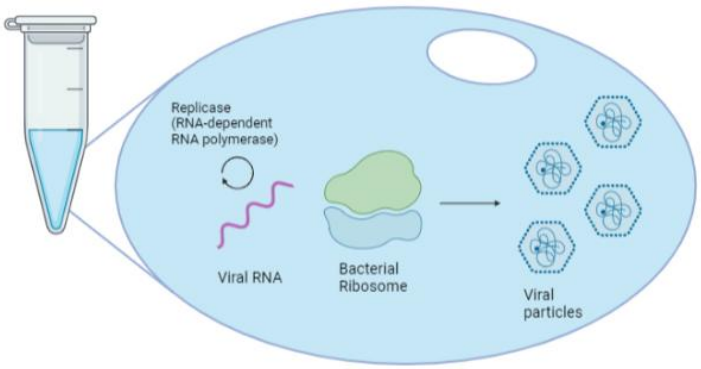


Figure 41. Picture of the concept of the assembly of the phage T7 in the TX/TL reaction. The viral genomic dsDNA is transcribed into RNA and the translation is performed on the ribosomes from *E. coli*. The phage proteins are assembled around the genomic DNA. Created with BioRender (https://www.biorender.com/)

Simulation of microgravity, or weightlessness, on the ground, is a field that requires a lot of thought and design. There are various ways of simulating weightlessness without going to space. One way is to go into a free fall. In this method, the weightless state is achieved by dropping the experiment from a certain height. While in freefall, the experiment is in a weightless state. Therefore, by dropping an experiment from a high altitude, an experiment can experience pure weightlessness for a very limited amount of time, measuring in seconds. Those types of experiments are performed by drop towers, which drop an experiment from a high altitude. Since the timespan of exposure to microgravity is too low in those experiments for many biological systems to make work, they are not very suitable for most biological experiments. Despite that, drop towers are known as some of the lowest-cost solutions to achieve weightlessness. Additionally, they achieve the “purest” weightlessness of all methods, with the g-forces going down to 10^{-8} x g. In the case of phage assembly researched here,

phage particle expression and assembly take hours to be achieved. For that reason, the drop tower is unsuitable for phage production studies.

Another method to achieve weightlessness which can take a longer time, but is higher cost and not as continuous, is parabolic flight. This method consists of flying an experiment in an airplane at a high altitude after which the plane dives and goes up. This method achieves multiple consecutive moments of weightlessness, but also hyper-gravity. Both range in time around 15-20 seconds. Through consecutive dives, the total microgravity exposure can be up to two minutes. However, an experiment in those moments passes through multiple exposures to hyper-gravity, meaning that the exposure is not completely pure. Still, the total weightlessness exposure is higher than the drop towers, which also comes at a higher cost.

Another option for getting microgravity response on Earth is the use of a sounding rocket. That method takes the experiment up to 100 km into the sky, reaching the upper atmosphere and releasing it to free-fall until the parachutes are activated. The experiment is consequently recovered, and the results are analyzed. This method achieves effective weightlessness in the range of around 10 minutes. Since those rockets do not achieve orbit, they can be theoretically launched from anywhere in the world, making them flexible and lower-cost solutions to achieve microgravity than sending an experiment to orbit. Despite still low time of exposure to microgravity, some biological experiments were performed using a sounding rocket (Demets, 2011; Hilbig *et al.*, 2011; Fajardo-Cavazos and Nicholson, 2021). Also, during the ascent, the samples go through a period of hyper-gravity, making the microgravity phase after the ascent. Also, since the samples are brought to Earth by a parachute, they still feel the drag throughout the descent.

The final option for doing genuine microgravity research is sending a sample to space. It can be done on a recoverable satellite or in an orbital laboratory (a space station). Those exposures can take months, or even years, which is highly fitting for biological studies. However, they require a high degree of experimental planning and design. Those studies are also extremely high in costs. Therefore, they are not widely accessible to most laboratories around the world. For that reason, the reproducibility of such studies is low.

Clinorotation offers an elegant and inexpensive way of simulating microgravity in liquid systems. Its principle of operation is through the randomization of the liquid movement, where gravity affects the liquid from all sides. There are different types of clinostats (Kiss *et al.*, 2019). The 2D-clinostats are well suited for fully liquid systems with micro-components such as bacterial solutions or movement of liquid inside plant tissues. They place the liquid system horizontally and rotate uniformly around their axis. This makes the effect of gravity on the system negligible, similarly as it would behave in a free fall. On the other hand, 3D-clinostats

are used for the microsystems which cannot be dispersed in liquid, such as adherent cells. Those clinostats are also called random positioning machines. They work in a way that they randomize the distribution of gravitational force direction by rotating the cells in random directions. In this way, they uniformly apply the gravitational force in all directions around the cells. In this way, the cells experience similar effect as if they were experiencing microgravity. In this study, a 2D-clinostat was used to assess the phage T7 self-assembly in TX/TL solution.

4.2. Materials and Methods

4.2.1. Experimental setup

The transcription/translation (TX/TL) solution was utilized for phage production in ground control and simulated microgravity.

The TX/TL ingredients include the cell lysate from *E. coli*, nucleotide triphosphate mix, and the phage genomic dsDNA. The final TX/TL reaction mixture is composed of the following reagents: 33% v/v of *E. coli* extract, 10 mM ammonium glutamate; 1.2 mM ATP; 0.850 mM each of GTP, UTP, and CTP; 0.034 mg/mL folinic acid; 0.175 mg/mL yeast tRNA; 2 mM amino acids; 30 mM 3-PGA; 0.33 mM NAD; 0.27 mM CoA; 1 mM putrescine; 1.5 mM spermidine; 4 mM oxalic acid; 57 mM HEPES, 3.5% PEG 8000, 5 µM of Chi6 linear DNA (Marshall *et al.*, 2017). Mg-glutamate and K-glutamate were optimized for phage production and set respectively to 7 mM and 170 mM. T7 Phage genomic DNA (Bioron) was set to 2 nM (1.204×10^{12} T7 genomic DNA molecules per mL).

The 1 mL pipette clinostat developed at the Gravitational Biology Department of DLR was used for the clinorotation of the samples (Figure 42).

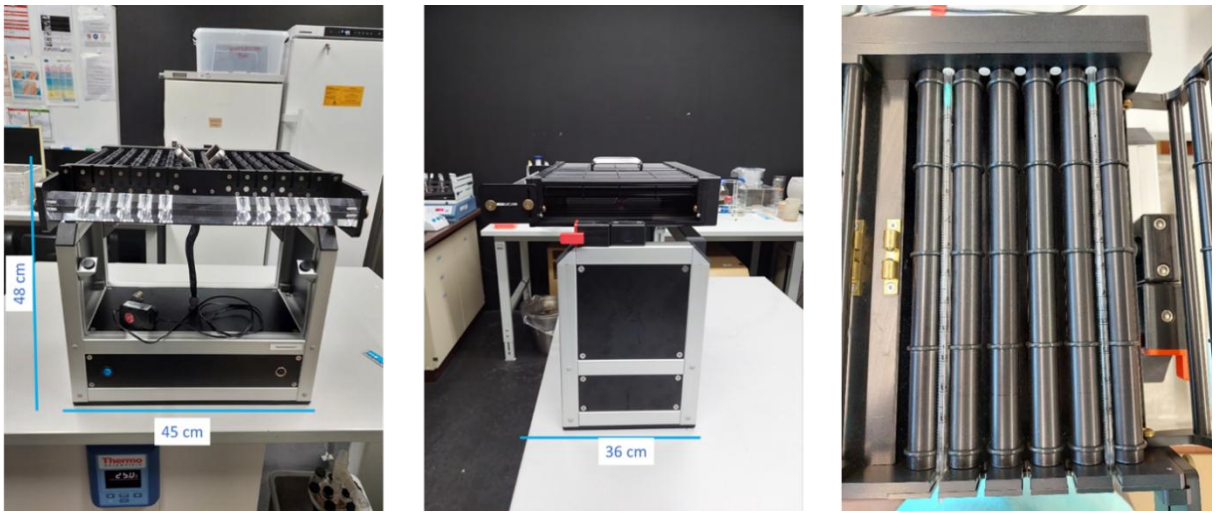


Figure 42. The clinostat device is used for the simulated microgravity experiments. Left, the front view of the clinostat; Center, the side view of the clinostat; Right, the top view with the 1 mL pipette inside.

Since the sample volumes achievable with TX/TL reaction are small, in the range of 50 µL, a piece of special equipment was prepared to rotate the samples of small volumes. A hollow tube of the exact thickness of a 1 mL pipette was 3D-printed from polyvinylchloride (PVC). This tube served as a holder. Inside, a smaller polytetrafluorethylene (PTFE) tube was inserted, containing a 50 µL sample (Figure 43). The setup aligned the liquid to rotate horizontally to the surface.

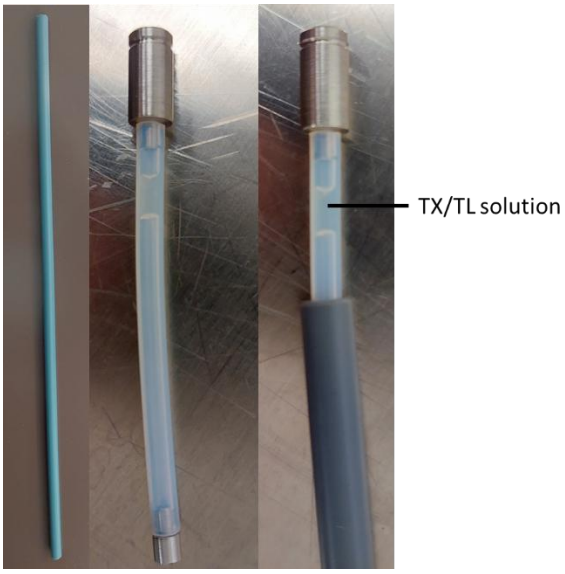


Figure 43. The costume setup for the clinorotation of small volumes of liquid. Specifically prepared for TX/TL reactions in clinostat. The big tube serves as a holder and is made of polyvinyl chloride (PVC). Inside, there is a small polytetrafluorethylene (PTFE) tube in which the reaction is running. The tube is capped from both sides with steel lids.

The setup was performed to treat triplicates of the TX/TL solution, together with the ground control, clinorotatated at 60 RPM. Before preparing the reactions, a master mix was created with all the TX/TL components, except the DNA. In the end, the DNA was added to the master mix and the mix was inoculated into the PTFE tube. The samples were held on ice until all the samples were ready and then all together put for clinorotation.

The reactions were performed at 29 °C in an incubator. The samples were taken after 1 hour, 2 h, 3 h, 4 h, 6 h, and 21 h. Out of the 20 µL, 10 µL were taken for serial dilutions and plating, and 10 µL were frozen at -80 °C with 25% v/v glycerol. Special samples were prepared for mass spectrometry and dot blot analysis.

4.2.2. Tracking synthesis of infectious virions through PFU assays

Plating was performed by overlay agar method (Kropinski *et al.*, 2009). The host bacterium used for plaque assays was *E. coli* B (DSM 613). The 50 mL liquid culture of the bacterium was grown for 18 hours in the Luria-Bertani broth (LB) medium. After that, the culture was

streaked, and a colony was isolated to grow in a new liquid culture flask containing LB. After 18 hours of incubation, the bacteria were diluted to the optical density at 600 nm (OD_{600}) between 0.8 and 1. Then, 5 μL of each serially diluted sample was mixed with 130 μL of *E. coli* B and incubated for 3 min at 37 °C, to allow phage attachment. The solution was added to 1.75 mL of the LB containing 0.75% agar. The solution was mixed well and poured onto an agar plate containing LB medium with 1.5% agar. The cultures were left for 10 min to harden and then replaced into an incubator at 37 °C for 3 hours until the plaques were visible and countable. The plaques were counted at the dilutions where the plaque density allowed counting.

4.2.3. GFP gene expression control

In addition to the plaque assays, a control was used to assess the gene expression. For this control, the reaction conditions were the same, except in the place of the T7 genomic dsDNA, there was a plasmid containing the super folder green fluorescence protein (sfGFP) gene. The plasmid is p70-sfGFP. The same amount of the plasmid was added as it would be of T7 DNA into two dedicated TX/TL samples. They were sampled after 5, and 21 hours, respectively. The green fluorescence was measured using a *Tecan Infinite M200 Pro* plate reader. For each readout, 10 μL of a sample was used. The excitation wavelength used was 485 nm and the detection wavelength was at 515 nm. The fluorescence of the clinorotated samples was compared to the stationary control.

4.2.4. Protein content determination through liquid chromatography tandem mass spectrometry

Additionally, the samples prepared for the liquid chromatography tandem mass spectrometry (LC/MS^2). The experiment was performed by Dr. Timo Glatter at Max Planck Institute for Terrestrial Microbiology. Each sample contained 50 μL of the TX/TL reaction. They were taken after 21 hours. The mass spectrometry was performed to evaluate the proteome of the phages produced, quantifying the gene expression. The MS was performed at the Max Planck Institute for Terrestrial Microbiology. Proteins contained in TX/TL samples (50 µL) were reduced by adding 5 mM Tris(2-carboxyethyl) phosphine at 90 °C for 15 min, followed by alkylation (10 mM iodoacetamide, 30 min at 25 °C). The amount of extracted proteins was measured using BCA protein assay (Thermo Fisher Scientific). 50 µg total protein was then digested with 1 µg trypsin (Promega) overnight at 30 °C in the presence of 0.5% SLS. Following digestion, SLS was precipitated with trifluoroacetic acid (TFA, 1.5% final concentration) and peptides were purified using Chromabond C18 micro spin columns (Macherey-Nagel). Acidified peptides were loaded on spin columns equilibrated with 400 µL acetonitrile and then 400 µL 0.15% TFA. After peptide loading, a washing step with 0.15% TFA was performed, followed by elution using 400 µL 50% acetonitrile. Eluted peptides were then dried by a vacuum concentrator and reconstituted in

0.15% TFA. The peptide mixtures were analyzed using liquid chromatography-mass spectrometry carried out on an Exploris 480 instrument connected to an Ultimate 3000 RSLC nano with a Prowflow upgrade and a nanospray flex ion source (all Thermo Scientific). Peptide separation was performed on a reverse phase HPLC column (75 µm × 42 cm) packed in-house with C18 resin (2.4 µm, Dr. Maisch). The following separating gradient was used: 94% solvent A (0.15% formic acid) and 6% solvent B (99.85% acetonitrile, 0.15% formic acid) to 25% solvent B over 40 min and 35% B for an additional 20 min at a flow rate of 300 mL per min. The DIA-MS acquisition method was adapted from Bekker-Jensen et al (Bekker-Jensen *et al.*, 2020). In short, the Spray voltage was set to 2.0 kV, funnel RF level at 55, and heated capillary temperature at 275 °C. For DIA experiments full MS resolutions were set to 120.000 at m/z 200 and full MS AGC target was 300% with an IT of 50 ms. The mass range was set to 350–1400. AGC target value for fragment spectra was set at 3000%. 45 windows of 15 Da were used with an overlap of 1 Da. Resolution was set to 15,000 and IT to 22 ms. Stepped HCD collision energy of 25, 27.5, and 30% was used. MS1 data was acquired in profile, and MS2 DIA data in centroid mode. Analysis of DIA data was performed using DIA-NN version 1.8 (Demichev *et al.*, 2020) using Uniprot protein databases from E. coli and bacteriophage T7. Full tryptic digest was allowed with two missed cleavage sites, and oxidized methionines and carbamidomethylated cysteines. Match between runs and remove likely interferences were enabled. The neural network classifier was set to the single-pass mode, and protein inference was based on genes. The quantification strategy was set to any LC (high accuracy). Cross-run normalization was set to RT-dependent. Library generation was set to smart profiling. DIA-NN exports were further statistically evaluated using a modified SafeQuant script made compatible with processing DIA-NN data.

From the rest of the TX/TL reactions which were not used for plaque assays, qPCR, dot-blot, and transmission electron microscopy were performed.

4.2.5. Tracking phage DNA copy number through qRT-PCR

To determine the bacteriophage T7 DNA copy number, qRT-PCR was performed. Each TX/TL sample was diluted 100x in nuclease-free water, and 5 µL of the diluted TX/TL was used for qRT-PCR, as described earlier for the T7 phage produced in TX/TL (Vogele *et al.*, 2021). The assays were performed in technical duplicates for each of the time points from all three experimental replicates. The primers and the program were adapted from Peng et al (Peng *et al.*, 2018). The kit used for the qRT-PCR reaction was NEB Luna® Universal qPCR & qRT-PCR (New England Biolabs). As a standard, 5 µL of T7 Phage DNA (Bioron) was used. The standard was diluted in nuclease-free water. Different dilutions were used: 0.2 nM (610,000,000 T7 DNA molecules), 0.2×10^{-2} nM (6,100,000 T7 DNA molecules), 0.2×10^{-4} nM (61,000 T7 DNA molecules), and 0.2×10^{-6} nM (610 T7 DNA molecules). The determined DNA

copy number in each sample was corrected for the original TX/TL reaction concentration and volume (100× concentrated, 20 µL). The reaction composition is shown in Table 16. The PCR protocol is shown in Table 17. The results were analyzed by the two-way analysis of variance (ANOVA) method.

Table 16. The components of the RT-qPCR reactions performed in this study.

Component	Volume (µL)
Luna® Universal qPCR Master Mix (2x	10
Forward primer (5'CCTCTTGGGAGGAAGAGATTTG3', 10 nM)	2.5
Reverse primer (5'TACGGGTCTCGTAGGACTTAAT3', 10 nM)	2.5
TX/TL sample (100x diluted in nuclease-free water)	5

Table 17. The PCR protocol for the PCR reactions performed in this study.

Temperature (°C)	Time (s)	Number of cycles
50	120	1x
95	120	1x
95 60	15 60	40x
95	15	1x
60	60	1x
95	15	1x

4.2.6. Evaluating expression of a single phage gene through dot blot

The dot blot analysis was performed to further evaluate the differences in gene expression. The dot blot was performed by Dr. François-Xavier Lehr at the Max Planck Institute for Terrestrial Microbiology. For that, 2 µL of the diluted TX/TL samples were blotted on a nitrocellulose membrane and let dry for 15 minutes. The membrane was incubated for 1 h with blocking buffer (Tris-buffer saline polysorbate 20, TBST 1×, 5% w/v non-fat dry milk) on a shaker at RT. The membrane was rinsed 2 times with water and incubated for 1 hour on a shaker at room temperature (RT) with the primary antibody (T7 tag, PA1-32386, Thermofisher) diluted in the blocking buffer (1:10,000). The membrane was washed 3 times with the wash buffer (TBST 1×) and incubated for 1 h on a shaker at RT with the secondary antibody (IRDye® 800CW Goat anti-Rabbit IgG, LI-COR) diluted (1:20,000) in the blocking buffer. Finally, the

membrane was washed 3 times with the wash buffer and scanned for fluorescence with an LI-COR Odyssey DLx. The calibration curve showed that the concentration of synthesized phages from the experiment was in the linear range of detection. For each experiment day (n = 3), the time points for the clinorotated and the stationary samples were blotted on the same membrane along with an internal standard (bacteriophage T7 synthesized in TX/TL at 1 nM DNA concentration). Dot blots of the three experiment days were performed on three consecutive days and analyzed in parallel by normalization to the internal standard.

4.2.7. Visualization of the synthesized phages through transmission electron microscopy (TEM)

To explore whether there are any differences in the virion structures, transmission electron microscopy (TEM) was performed by Dr. Thomas Heimerl from Max Planck Institute for Terrestrial Microbiology, to visualize the Enterobacteria Phage T7 viral particles produced in the TX/TL reactions. To do that, 30 µL of the diluted TX/TL samples from the 21 h time point was further diluted to 200 µL with Tris-buffer saline (TBS; Tris 20 mM, NaCl 150 mM, pH 7.6). The samples were extracted by adding 10 µL of chloroform and centrifuged for 10 min at 13,000 × g. The upper aqueous phase was transferred to a new tube. The bacteriophages were precipitated by adding 50 µL of polyethylene glycol (PEG)/NaCl 5x buffer (PEG-8000 20%, NaCl 2.5 M). The tubes were incubated at 4 °C overnight. The bacteriophages were pelleted by centrifugation for 10 minutes at 13,000 × g. The bulks of supernatants were removed. The samples were centrifuged again for 10 minutes at 13,000 × g and the supernatant was completely removed. The pellets were resuspended in 20 µL TBS. Carbon-coated copper grids (400 mesh) were hydrophilized by glow discharging (PELCO easiGlow, Ted Pella, USA). 5 µL of the primary antibody solution (T7 tag, PA1-32386, Thermofisher) was applied onto the hydrophilized grids. After 2 min of incubation, the solutions were briefly blotted and 5 µL of the samples were applied to the grids. Alternatively, “antibody coated” grids were incubated in diluted sample solution (1:40) for 1 h under shaking. After two short washing steps in droplets of double distilled water, samples were stained with 2% uranyl acetate. Samples were analyzed with a JEOL JEM-2100 transmission electron microscope using an acceleration voltage of 120 kV. Images were acquired with an F214 FastScan CCD camera (TVIPS, Gauteng).

4.2.8. Experimental setup visualization

The schematic of the whole experimental setup with the number of samples is shown in Figure 44.

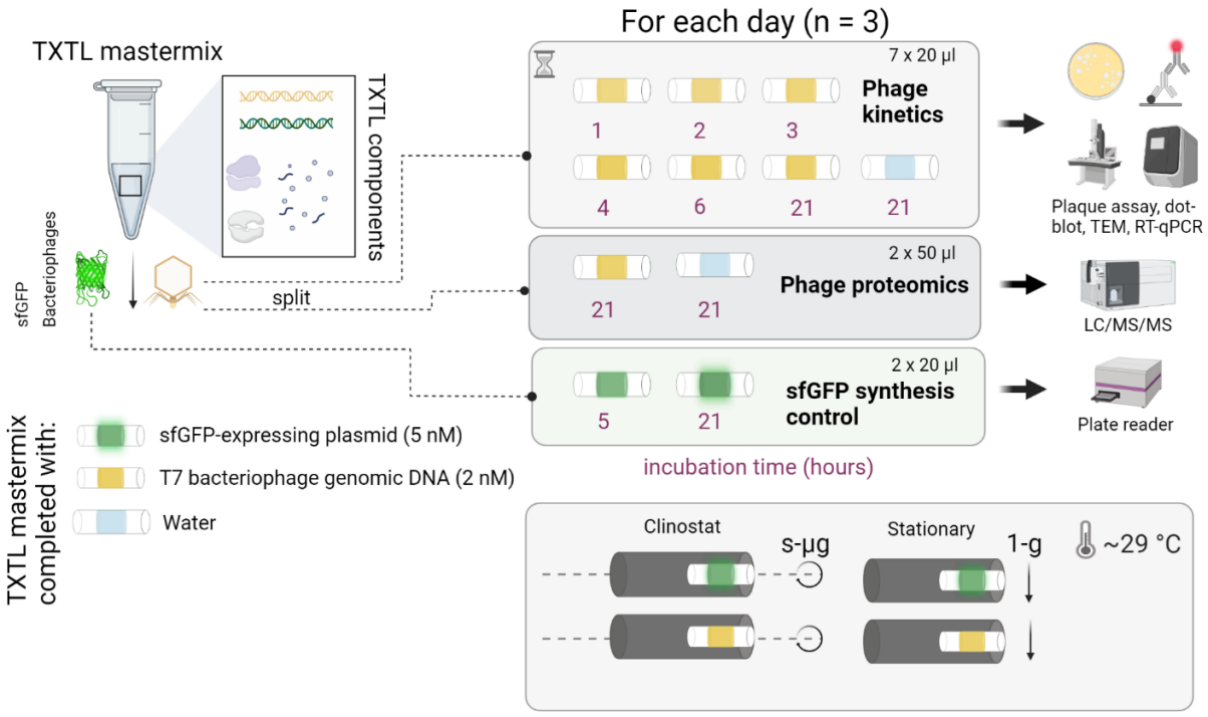


Figure 44. The schematic of the experiments performed in this study. Each day, out of three days in total, the TX/TL mastermix was prepared containing all the components as described in the materials and methods 4.2.1. The mix was separated for three experiments, particularly, phage kinreics, phage proteomics, and sfGFP synthesis control. For Phage kinetics, 7 aliquotes were taken from the master mix, one of which was a negative control with added water instead of DNA. Another aliquote was used for phage proteomics together with a negative control which was sent for Mass Spectrometry. Another two aliquotes were used as a gene expression control where a plasmid with a GFP gene was used to assess in the same solution the production of GFP. Each synthesis was performed at 29 °C and there was a sample group which was exposed to clinorotation (s-µg) and the control group which was stationary (1-g).

4.3. Results

4.3.1. Phage synthesis kinetics

The phage synthesis kinetics was determined through cultivation, i.e. plaque assays over time. The results are shown in Figure 45. Overall, the samples subjected to clinorotation showed a faster rate of synthesis and higher endpoint titer. It is visible that the efficacy of the Phage T7 production by PFU assays was higher under clinorotation as the titer grows faster and reaches a higher endpoint titer.

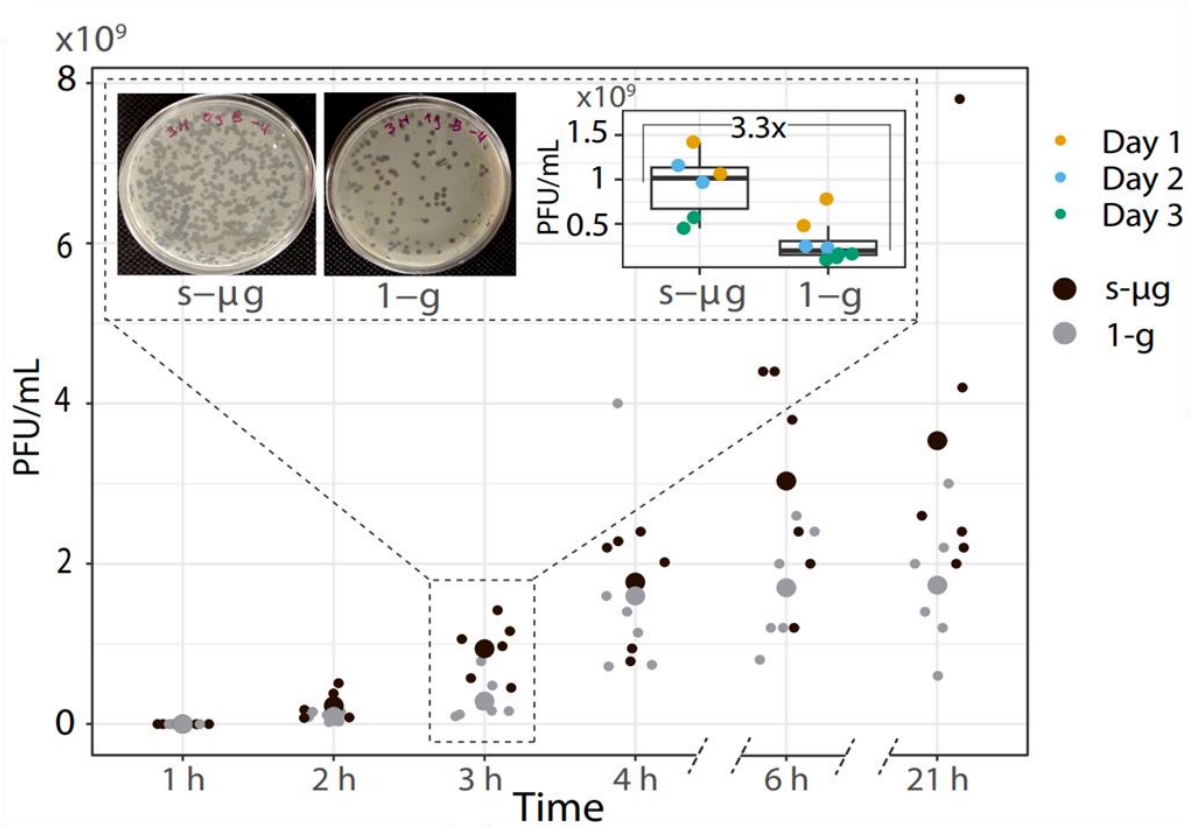


Figure 45. The kinetics of Enterobacteria Phage T7 production in TX/TL solution under clinorotation (s-µg) and in the stationary control (1-g). An example of clinorotated and stationary plates after 3 h is shown in the upper left corner, as well as their difference in the determined titers.

4.3.2. GFP control gene expression

Figure 46 shows the GFP fluorescence intensity in the clinorotated (s-µg) and stationary control (1-g) samples after 5 and 21 hours of reaction. As seen from the colored points, after 5 hours the fluorescence is lower than after 21 hours, which is to be expected. Also, the 1-g control had a slightly higher fluorescence overall, but not statistically significantly.

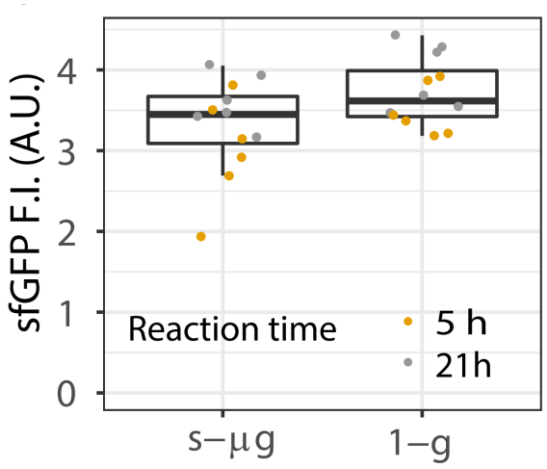


Figure 46. The fluorescence values observed from sfGFP for the TX/TL reactions containing a plasmid with the sfGFP gene. The time points are after 5 hours and 21 hours. The values in between the treatments for both time points are not significantly different.

4.3.3. Dot blot gene expression evaluation

To explore the gene expression at the protein-level, dot blot analysis was performed. The polyclonal antibody used, targeted T7 major capsid protein, encoded by gene 10. The fluorescence intensity after washing and drying is plotted in Figure 47.

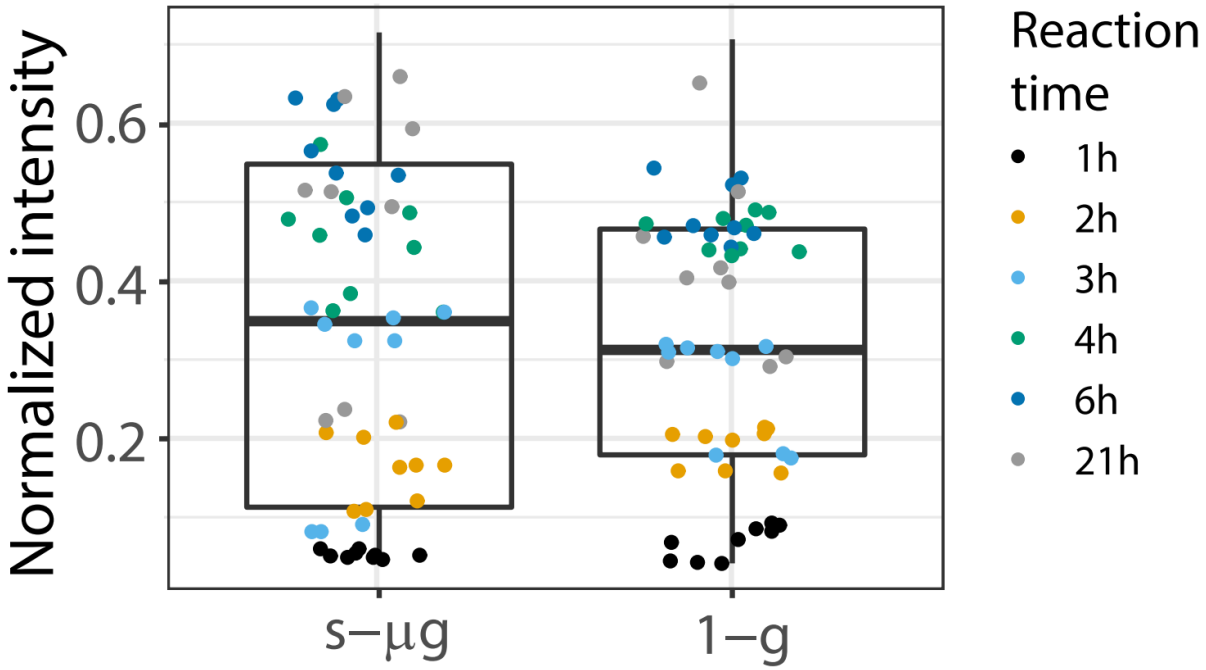


Figure 47. A plot of the normalized dot blot fluorescence intensities over time points for the phage T7 major capsid protein. Each column represents a treatment, s-µg simulated microgravity and 1-g stationary control. The time points are represented in different colors.

4.3.4. Proteome analysis by LC/MS^2

To determine the full proteome differences, the mass spectrometry was performed. Figure 48 shows the volcano plot resulting from the results. No phage protein falls beyond the statistical significance q-value of 0.05 as seen on y axis of the plot (the axis is log-scaled and -log10(0.05) = 1.3010). Also, no phage protein goes in an amount above 2x change as visible on the x axis of the plot.

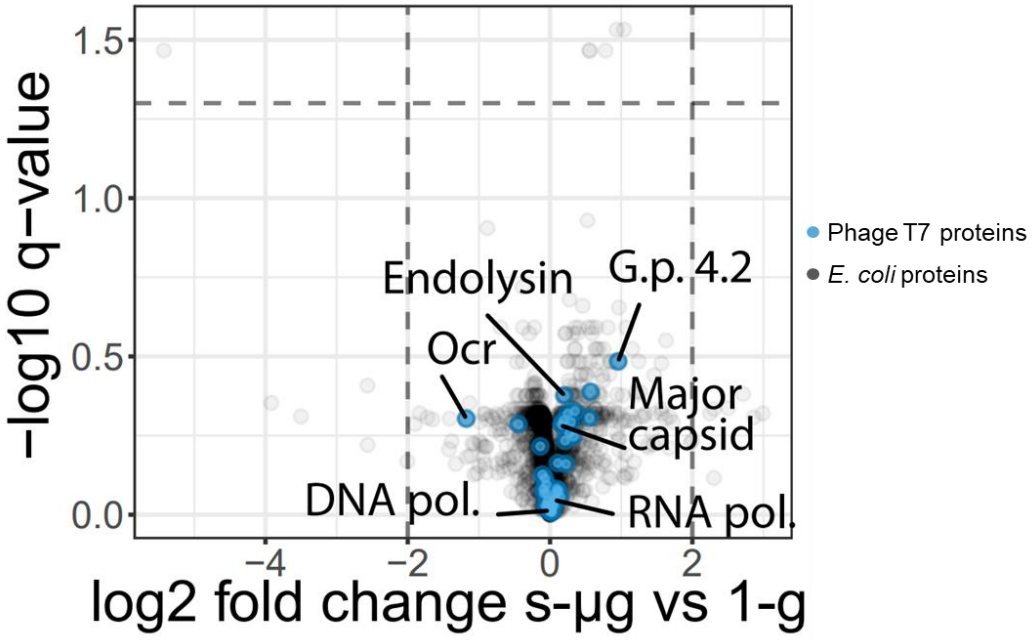


Figure 48. Vulcano plot of the mass spectrometry data after comparing the amount of proteins from clinorotated sample (s-µg) and stationary control (1-g). The blue points indicate phage T7 protein signals with some of the proteins specifically labeled. The transparent black points are remnant *E coli* proteins and are not of big interest in this study.

4.3.5. Tracking T7 DNA copy numbers

Figure 49 shows the results of qPCR amplification, quantifying the DNA copies. It is observable that for both stationary control and clinorotated sample the DNA copy number stays the same, and only after 21 hours it drops which is to be expected because DNA degrades over time. Figure 50 represents the standard curve for the samples. Table 18 represents the statistical significance between the samples over time after ANOVA analysis.

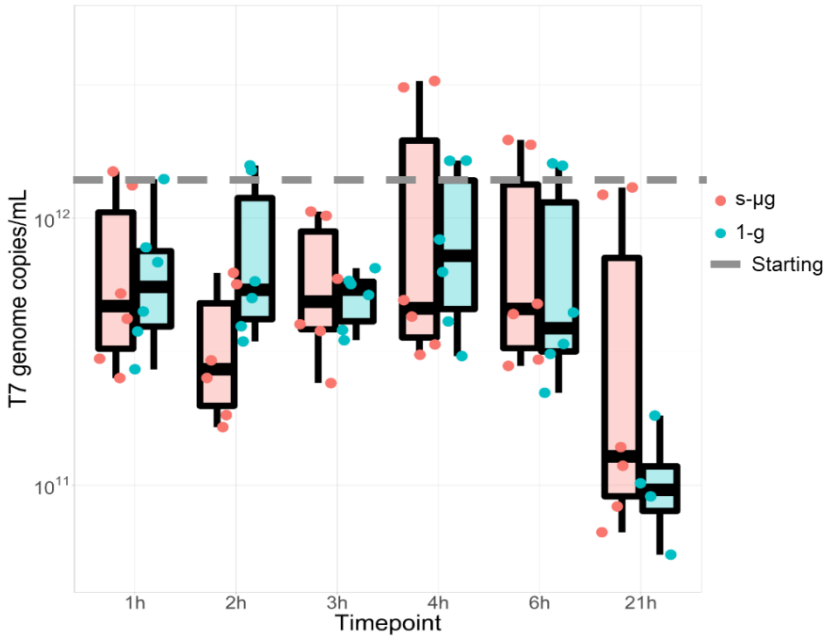


Figure 49. The results of the qPCR assay on the samples after the TX/TL reaction. The genome copy numbers were determined based on the standard curve. Throughout the experiment, the amount of T7 phage DNA stayed the same with a slight reduction after 21 hours, likely due to degradation. There were no signs of replication observed.

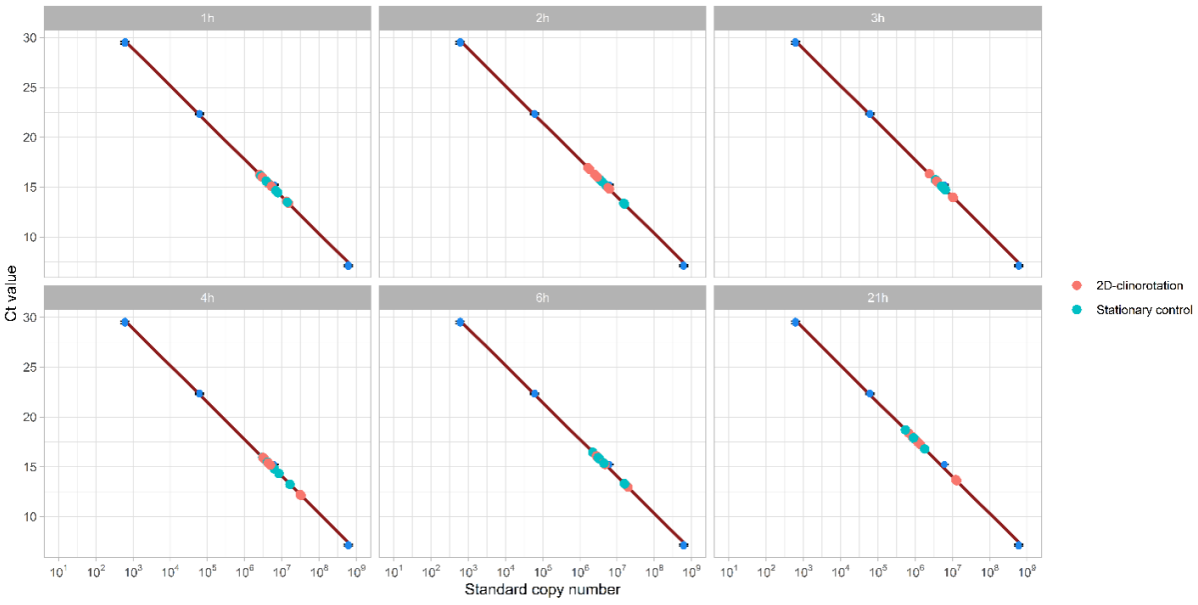


Figure 50. The standard curve of the RT-qPCR, with the plotted sample values onto it. The equation of the curve: y = -1.61000 ln(x) + 39.99968; R^2 = 0.9988.

Table 18. The results of the two-way ANOVA test on the copy numbers of T7 DNA genomes determined by qPCR. There is no significant difference between the time points and samples, hinting that the DNA is relatively stable over time in the TX/TL reaction.

Test of dependency	F	p
Copy number: Time point	1.942	0.0998
Copy number: Clinorotation	0.380	0.5349
Timepoint: Clinorotation	0.704	0.6230

4.3.6. Transmission electron microscopy

The transmission electron microscopy did not show any observable differences between the virions of the stationary samples and simulated microgravity. However, it is worth to note that the virions were more plentiful from the clinorotated sample than from the stationary samples. Figure 51 shows example TEM image of T7 phages synthesized under simulated microgravity and stationary conditions. However, this is not a quantitative method to prove a higher count of virions in any sample and PFU assays are more reliable way, especially to determine the number of infectious virions.

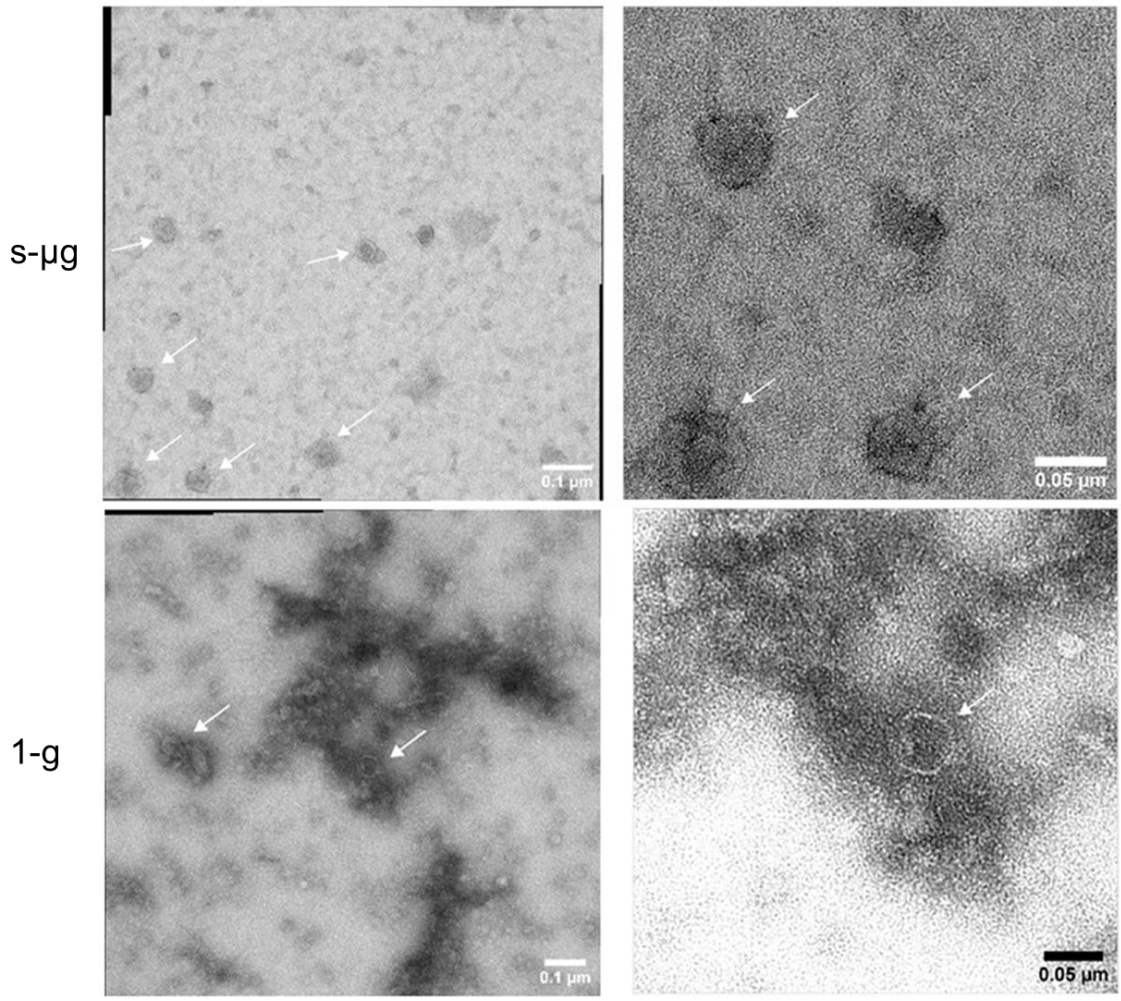


Figure 51. T7 bacteriophages synthesized in s-µg and 1-g visualized by transmission electron microscopy. White arrows indicate presumed fully assembled phage capsids.

Taking the results together, as the gene expression stays the same, as well as the DNA content while the number of infectious virions rises while clinorotation. This indicates that while the total content of the proteins and capsids is the same, the assembly of those is more efficient or less faulty in the samples that were subject to clinorotation.

4.4. Discussion

4.4.1. Phage synthesis kinetics

As visible from Figure 45, the titer of phages grows over time very quickly. First plaques can already be seen after 2 hours of synthesis, particularly in the clinorotated samples. After 3 hours, the difference is particularly striking between the clinorotated and stationary samples. The end-point titer is also higher in clinorotated samples approximately two times. This result goes in line with the assumption that mixing is higher during clinorotation. Similarly, in microgravity, the water buoyancy force does not result in downward motion. Rather, the buoyancy acts in all directions. This is similar to mixing a fluid, as it was done in clinorotation.

This is the reason why in microgravity crystals form more uniformly (McPherson and DeLucas, 2015). Because of the lack of gravity, the molecules are moving more randomly in water and therefore interact more uniformly. This effect could be present in this experiment. Even though the experiment was not performed in real microgravity, it still shows that phage synthesis could be enhanced on Earth through clinorotation.

4.4.2. Gene Expression

The fluorescence of the samples with the sfGFP-producing plasmid was similar between the clinorotated and stationary samples. However, after 21 hours, the fluorescence became stronger for both samples than after 5 hours (Figure 46). This suggests that the gene expression is continuous over 21 hours, and it is not affected by the clinorotation. As the sfGFP was produced to the same extent in both clinorotated sample and stationary control sample, the green fluorescence produced is not significantly different. As GFP is a simple small protein without complex assembly, we can assume that the tertiary structure was achieved with high efficiency. Therefore, the fluorescence measurements from both treatments do not differ significantly, though the fluorescence values grow in both conditions, signifying that the transcription goes at a similar rate for both conditions.

Figure 47 represents the results of the dot-blot. It shows that T7 major capsid protein was present in similar amounts for all the time points in clinorotated and stationary samples. This is in line with the results of GFP protein being expressed to a similar extent as shown in Figure 46, without significant differences.

Figure 48 shows the LC/MS^2 results plotted on a volcano plot. The mass spectrometry analysis showed similar amount of phage proteins in both clinorotated and stationary TX/TL reactions, further supporting the prospect that the gene expression does not change. Since no phage protein showed significantly different amounts of protein, this points that protein-coding gene expression does not change significantly between the treatments. Considering all results, with dot-blot and mass spectrometry data, it is reasonable to assume that differences in gene expression are lacking and therefore cannot explain higher rate of phage assembly in clinorotated samples.

4.4.3. DNA copy numbers

The qPCR analysis shows that there was no significant difference in the amount of DNA between the clinorotated and stationary samples (Figure 49). The standard curve representing the sample Ct values is also shown in Figure 50. Comparing the values between each other with two-way ANOVA, testing if the variation in DNA copy number is explained by clinorotation or timepoint, we conclude that there is no significant difference comparing any of the groups (Table 18). This shows that the phage DNA is not replicating in TX/TL, contradictory to the

previous research (Shin, Jardine and Noireaux, 2012). The possible explanation for this is that the starting solution of TX/TL was highly concentrated with phage DNA, thermodynamically preventing the production of more genome copies. ANOVA results show that there are no significant differences between the clinorotated and stationary samples, and there are no significant differences in any of the timepoints.

4.4.4. Transmission electron microscopy

Figure 51 shows the T7 virions synthesized in TX/TL reaction. There is no particular difference in the virions themselves. Still, the electron micrographs confirm the presence of synthesized T7 virions. The T7 synthesis has been shown in the past in TX/TL solutions (Shin, Jardine and Noireaux, 2012; Garamella *et al.*, 2016)

4.5. Conclusion

Infectious bacteriophage T7 virions were successfully synthesized in vitro through a TX/TL reaction from its genomic DNA. The reaction was repeated while stationary and while in clinorotation at 29 °C. Under the clinorotation conditions, the bacteriophages got synthesized faster and reached a higher endpoint titer.

The effect was explored through a range of methods including a GFP plasmid control, measuring fluorescence, dot blot fluorescence measurements through the antibody specific for major capsid protein, LC/MS^2, and qPCR. The results indicate a more efficient infectious virion assembly under simulated microgravity.

Concluding Remarks

In this thesis, a vast range of questions were explored in the context of space and environmental virology. The aim was to kickstart virus-related research in spaceflight and public transport environment. Therefore, a diverse set of experiments was performed, demonstrating the possibilities of space and environmental virus research.

First, the bacteriophages tested here showed a wide range of resistance to drying and detectability on glass surfaces. It was shown in a standardized manner that viruses of different types show varying environmental stability. The differences between various surrogate phages were determined in the stability concerning survival after drying. Those bacteriophages could be used as surrogates in space virology studies. Study of viruses in spaceflight is important because it allows us to understand their stability and spread potential during human space missions and also, they act as models for the origin of life due to their simple structure. Therefore, they are invaluable assets for astrobiology. Here, the phages were introduced as potential surrogates and models for such studies.

Second, the diversity of temperate bacteriophages in the ISS environment was explored. A high level of gene transfer and re-functionalization among phages within different bacterial genera was observed. A group with a particularly wide range of bacteriophages was the genus *Staphylococcus.* The phylogeny between all the bacteriophage groups was also determined.

SARS-CoV-2, MPXV, and Phi6 were tested for their resistance against copper to determine its efficacy as an antiviral agent. Interestingly, a variation within virus variants was visible. Phi6 was shown to not be a suitable surrogate for those viruses where copper resistance is concerned. As the experiments showed, Phi6 demonstrated faster inactivation rate than either SARS-CoV-2 or MPXV. Also, 3D-printed antimicrobial surface containing copper microparticles was shown to be a promising antiviral material. It effectively inactivated Phi6 and SARS-CoV-2 within minutes. This material therefore shows promise for use in future space missions for construction of habitats or special high-touch-rate surfaces. This is due to its light weight and high antimicrobial potential. At the same time, the 3D-printed material with nanoparticles of copper showed lower effectiveness of virus inactivation.

Additionally, a bacteriophage T7 which was synthesized in the clinorotation device showed higher effectiveness of self-assembly than the stationary control. The molecular detection methods confirmed the improved assembly efficiency, rather than replication or gene expression, is the reason for that difference. This shows that in spaceflight, TX/TL synthesis could be used for the production of phages or complex biomolecules for potential therapies. Additionally, it shows that the application of rotation of the bioreactors for phage could improve the yield of the production. It was also shown that phages could potentially assemble better in

microgravity though this remains to be tested in real microgravity conditions. Given that the phage virion assembly was shown to be more efficient in simulated microgravity environment, this research gives an insight into the dynamics of phage propagation in spaceflight.

All the findings presented here add a piece of a grand puzzle about viral stability in spaceflight and on environmental surfaces. Bacteriophages were introduced and tested as valuable viral surrogates for spaceflight. The 3D-printed materials well-adapted for use in space and aviation were tested against Phi6 (one of the phage surrogates) and SARS-CoV-2, showing a promising antiviral potential. Additionally, the diversity of temperate phages of the ISS isolates was assessed as well as the self-assembly of a model phage in simulated microgravity. This work is only the beginning to better understand the factors and opportunities existing in the yet unexplored world of aerospace, public transport and aviation virology

List of References

Adriaenssens, E. and Brister, J. R. (2017) 'How to Name and Classify Your Phage: An Informal Guide', *Viruses*. doi: 10.3390/v9040070.

Aganovic, A. *et al.* (2022) 'Modeling the impact of indoor relative humidity on the infection risk of five respiratory airborne viruses', *Scientific Reports 2022 12:1*, 12(1), pp. 1–10. doi: 10.1038/s41598-022-15703-8.

Agency, I. (2017) 'WHO Consultative Meeting on High / Maximum Containment (Biosafety Level 4) Laboratories Networking', (December), pp. 13–15.

Ahmedzai, S. *et al.* (2011) 'Managing passengers with stable respiratory disease planning air travel: British Thoracic Society recommendations', *Thorax*. BMJ Publishing Group Ltd, p. 20. doi: 10.1136/thoraxjnl-2011-200295.

Akiyama, T. *et al.* (2020) 'How does spaceflight affect the acquired immune system?', *npj Microgravity*, 6(1), p. 14. doi: 10.1038/s41526-020-0104-1.

Akram, F., Imtiaz, M. and Haq, I. ul (2023) 'Emergent crisis of antibiotic resistance: A silent pandemic threat to 21st century', *Microbial Pathogenesis*, 174, p. 105923. doi: https://doi.org/10.1016/j.micpath.2022.105923.

Al-Shayeb, B. *et al.* (2020) 'Clades of huge phages from across Earth's ecosystems', *Nature*, 578(7795), pp. 425–431. doi: 10.1038/s41586-020-2007-4.

Alcock, B. P. *et al.* (2020) 'CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database', *Nucleic Acids Research*, 48(D1), pp. D517–D525. doi: 10.1093/nar/gkz935.

Aloor, A. *et al.* (2022) 'Glycosylation in SARS-CoV-2 variants: A path to infection and recovery', *Biochemical Pharmacology*, 206, p. 115335. doi: https://doi.org/10.1016/j.bcp.2022.115335.

Amankwah-Amoah, J. (2021) 'COVID-19 pandemic and innovation activities in the global airline industry: A review', *Environment International*, 156, p. 106719. doi: https://doi.org/10.1016/j.envint.2021.106719.

Anke, J. *et al.* (2021) 'Impact of SARS-CoV-2 on the mobility behaviour in Germany', *European Transport Research Review*, 13(1). doi: 10.1186/s12544-021-00469-3.

Aoyama, A. and Hayashi, M. (1986) 'Synthesis of bacteriophage ϕX174 in vitro: Mechanism of switch from DNA replication to DNA packaging', *Cell*, 47(1), pp. 99–106. doi: 10.1016/0092-8674(86)90370-3.

Arkin, A. P. *et al.* (2018) 'KBase: The United States Department of Energy Systems Biology Knowledgebase', *Nature Biotechnology*, 36(7), pp. 566–569. doi: 10.1038/nbt.4163.

Asako, U. and A., F. B. (2005) 'Identification of an Interacting Coat-External Scaffolding Protein Domain Required for both the Initiation of φX174 Procapsid Morphogenesis and the Completion of DNA Packaging', *Journal of Virology*, 79(11), pp. 6751–6756. doi: 10.1128/jvi.79.11.6751-6756.2005.

Azimi, P. *et al.* (2021) 'Mechanistic transmission modeling of COVID-19 on the Diamond Princess cruise ship demonstrates the importance of aerosol transmission', *Proceedings of the National Academy of Sciences of the United States of America*, 118(8). doi: 10.1073/pnas.2015482118.

B., F. A., Claire, B. A. and Sapna, C. D. (2022) 'Murine Hepatitis Virus, a Biosafety Level 2 Model for SARS-CoV-2, Can Remain Viable on Meat and Meat Packaging Materials for at Least 48 Hours', *Microbiology Spectrum*, 10(5), pp. e01862-22. doi: 10.1128/spectrum.01862-22.

Bagheri, G. *et al.* (2023) 'Size, concentration, and origin of human exhaled particles and their dependence on human factors with implications on infection transmission', *Journal of Aerosol Science*, 168(August 2022), p. 106102. doi: 10.1016/j.jaerosci.2022.106102.

Baker, A. C., Gutierrez, A. and Gibson Kristen, E. (2022) 'Factors Impacting Persistence of Phi6 Bacteriophage, an Enveloped Virus Surrogate, on Fomite Surfaces', *Applied and Environmental Microbiology*, 88(7), pp. e02552-21. doi: 10.1128/aem.02552-21.

Baker, P. W., Meyer, M. L. and Leff, L. G. (2004) 'Escherichia coli growth under modeled reduced gravity', *Microgravity - Science and Technology*, 15(4), pp. 39–44. doi: 10.1007/BF02870967.

Bartlett, A. *et al.* (2022) 'A comprehensive list of bacterial pathogens infecting humans', *Microbiology*, 168(12). doi: https://doi.org/10.1099/mic.0.001269.

Bäumler, W. *et al.* (2022) 'Antimicrobial coatings for environmental surfaces in hospitals: a potential new pillar for prevention strategies in hygiene', *Critical Reviews in Microbiology*, 48(5), pp. 531–564. doi: 10.1080/1040841X.2021.1991271.

Behboudi, E. *et al.* (2024) 'SARS-CoV-2 mechanisms of cell tropism in various organs considering host factors', *Heliyon*, 10(4), p. e26577. doi: https://doi.org/10.1016/j.heliyon.2024.e26577.

Bekker-Jensen, D. B. *et al.* (2020) 'A Compact Quadrupole-Orbitrap Mass Spectrometer with

FAIMS Interface Improves Proteome Coverage in Short LC Gradients*', *Molecular & Cellular Proteomics*, 19(4), pp. 716–729. doi: https://doi.org/10.1074/mcp.TIR119.001906.

Benjamin, C. L. *et al.* (2016) 'Decreases in thymopoiesis of astronauts returning from space flight', *JCI insight*, 1(12), p. e88787.

Benoit, M. R. and Klaus, D. M. (2007) 'Microgravity, bacteria, and the influence of motility', *Advances in Space Research*, 39(7), pp. 1225–1232. doi: https://doi.org/10.1016/j.asr.2006.10.009.

van den Berg, D. F. *et al.* (2023) 'Phage tRNAs evade tRNA-targeting host defenses through anticodon loop mutations', *eLife*. Edited by R. L. Gonzalez, J. L. Manley, and B. Wiedenheft, 12, p. e85183. doi: 10.7554/eLife.85183.

Bernhardt, T. G., Roof, W. D. and Young, R. (2000) 'Genetic evidence that the bacteriophage φX174 lysis protein inhibits cell wall synthesis', *Proceedings of the National Academy of Sciences*, 97(8), pp. 4297–4302. doi: 10.1073/pnas.97.8.4297.

Boekel, J. *et al.* (2015) 'Multi-omic data analysis using Galaxy', *Nature Biotechnology*, 33(2), pp. 137–139. doi: 10.1038/nbt.3134.

Boon, T. (2020) 'Qatar Airways Begins Disinfecting Aircraft Cabins With UV', *Simple Flying*.

Boudaud, N. *et al.* (2012) 'Removal of MS2, Qβ and GA bacteriophages during drinking water treatment at pilot scale', *Water Research*, 46(8), pp. 2651–2664. doi: https://doi.org/10.1016/j.watres.2012.02.020.

Bouras, G. *et al.* (2023) 'Pharokka: a fast scalable bacteriophage annotation tool', *Bioinformatics*, 39(1), p. btac776. doi: 10.1093/bioinformatics/btac776.

Brayden, Y. *et al.* (2024) 'Exploring the genomic basis of Mpox virus-host transmission and pathogenesis', *mSphere*, 9(12), pp. e00576-24. doi: 10.1128/msphere.00576-24.

Brown, R. B., Klaus, D. and Todd, P. (2002) 'Effects of space flight, clinorotation, and centrifugation on the substrate utilization efficiency ofE. coli', *Microgravity Science and Technology*, 13(4), p. 24. doi: 10.1007/BF02881678.

Cai, J. *et al.* (2019) 'Roles of Different Transport Modes in the Spatial Spread of the 2009 Influenza A(H1N1) Pandemic in Mainland China', *International Journal of Environmental Research and Public Health*. doi: 10.3390/ijerph16020222.

Campbell, C. D. *et al.* (2018) 'Air Travel for Subjects Receiving Long-Term Oxygen Therapy', *Respiratory Care*, 63(3), pp. 326–331. doi: 10.4187/respcare.05522.

Capella-Gutiérrez, S., Silla-Martínez, J. M. and Gabaldón, T. (2009) 'trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses', *Bioinformatics*, 25(15), pp. 1972–1973. doi: 10.1093/bioinformatics/btp348.

Carbone, A. (2008) 'Codon Bias is a Major Factor Explaining Phage Evolution in Translationally Biased Hosts', *Journal of Molecular Evolution*, 66(3), pp. 210–223. doi: 10.1007/s00239-008-9068-6.

Castro-Wallace, S. L. *et al.* (2017) 'Nanopore DNA Sequencing and Genome Assembly on the International Space Station', *Scientific Reports*, 7(1), pp. 1–12. doi: 10.1038/s41598-017-18364-0.

CDC (2024) *About Epstein-Barr Virus (EBV)*. Available at: https://www.cdc.gov/epstein-barr/about/index.html (Accessed: 4 August 2025).

CDC (2025) *Clinical Overview of Chickenpox (Varicella)*. Available at: https://www.cdc.gov/chickenpox/hcp/clinical-overview/index.html (Accessed: 4 August 2025).

Chan, P. P. *et al.* (2021) 'tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes', *Nucleic Acids Research*, 49(16), pp. 9077–9096. doi: 10.1093/nar/gkab688.

Cock, P. J. A. *et al.* (2009) 'Biopython: freely available Python tools for computational molecular biology and bioinformatics', *Bioinformatics*, 25(11), pp. 1422–1423. doi: 10.1093/bioinformatics/btp163.

Cowen, D., Zhang, R. and Komorowski, M. (2024) 'Infections in long-duration space missions', *The Lancet Microbe*, 5(9), p. 100875. doi: https://doi.org/10.1016/S2666-5247(24)00098-3.

Cresawn, S. G. *et al.* (2011) 'Phamerator: a bioinformatic tool for comparative bacteriophage genomics', *BMC Bioinformatics*, 12(1), p. 395. doi: 10.1186/1471-2105-12-395.

Crucian, B. *et al.* (2016) 'Incidence of clinical symptoms during long-duration orbital spaceflight', *International Journal of General Medicine*, 9, pp. 383–391. doi: 10.2147/IJGM.S114188.

Cui, L. *et al.* (2024) 'Current and future directions in bacteriophage research for developing therapeutic innovations', *Scientific Reports*, 14(1), p. 24404. doi: 10.1038/s41598-024-76427-5.

Danecek, P. *et al.* (2021) 'Twelve years of SAMtools and BCFtools', *GigaScience*, 10(2), p. giab008. doi: 10.1093/gigascience/giab008.

Demets, R. (2011) ‘Biology on Sounding Rockets: History, Requirements, Results and Scientific Interpretation’, in. Available at: https://api.semanticscholar.org/CorpusID:53597339.

Demichev, V. *et al.* (2020) ‘DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput’, *Nature Methods*, 17(1), pp. 41–44. doi: 10.1038/s41592-019-0638-x.

Dolinsky, T. J. *et al.* (2004) ‘PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations’, *Nucleic Acids Research*, 32(suppl_2), pp. W665–W667. doi: 10.1093/nar/gkh381.

Ekonomou, S. I. *et al*. (2023) ‘An explorative study on the antimicrobial effects and mechanical properties of 3D printed PLA and TPU surfaces loaded with Ag and Cu against nosocomial and foodborne pathogens’, *Journal of the Mechanical Behavior of Biomedical Materials*, 137, 105536. doi: 10.1016/j.jmbbm.2022.105536

Ellingjord-Dale, M. *et al.* (2022) ‘The use of public transport and contraction of SARS-CoV-2 in a large prospective cohort in Norway’, *BMC Infectious Diseases*, 22(1), p. 252. doi: 10.1186/s12879-022-07233-5.

Epstein, J. H. *et al.* (2020) ‘Nipah virus dynamics in bats and implications for spillover to humans’, *Proceedings of the National Academy of Sciences*, 117(46), pp. 29190–29201. doi: 10.1073/pnas.2000429117.

Essaidi-Laziosi, M. *et al.* (2021) ‘Estimating clinical SARS-CoV-2 infectiousness in Vero E6 and primary airway epithelial cells’, *The Lancet Microbe*, 2(11), p. e571. doi: 10.1016/S2666-5247(21)00216-0.

Fajardo-Cavazos, P. and Nicholson, W. L. (2021) ‘Mechanotransduction in Prokaryotes: A Possible Mechanism of Spaceflight Adaptation’, *Life*, 11. Available at: https://api.semanticscholar.org/CorpusID:231587939.

Fedorenko, A. *et al.* (2020) ‘Survival of the enveloped bacteriophage Phi6 (a surrogate for SARS-CoV-2) in evaporated saliva microdroplets deposited on glass surfaces’, *Scientific Reports*, 10(1), p. 22419. doi: 10.1038/s41598-020-79625-z.

Fu, L. *et al.* (2012) ‘CD-HIT: accelerated for clustering the next-generation sequencing data’, *Bioinformatics*, 28(23), pp. 3150–3152. doi: 10.1093/bioinformatics/bts565.

Gallandat, K. and Lantagne, D. (2017) ‘Selection of a Biosafety Level 1 (BSL-1) surrogate to evaluate surface disinfection efficacy in Ebola outbreaks: Comparison of four bacteriophages’, *PLOS ONE*, 12(5), p. e0177943. Available at: https://doi.org/10.1371/journal.pone.0177943.

Garamella, J. *et al.* (2016) 'The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology', *ACS Synthetic Biology*, 5(4), pp. 344–355. doi: 10.1021/acssynbio.5b00296.

Geller, C., Varbanov, M. and Duval, R. E. (2012) 'Human Coronaviruses: Insights into Environmental Resistance and Its Influence on the Development of New Antiseptic Strategies', *Viruses*, pp. 3044–3068. doi: 10.3390/v4113044.

Georgakopoulos-Soares, I. *et al.* (2023) 'Surface antibacterial properties enhanced through engineered textures and surface roughness: A review', *Colloids and Surfaces B: Biointerfaces*, 231, p. 113584. doi: https://doi.org/10.1016/j.colsurfb.2023.113584.

George, S. P. *et al.* (2024) 'Space radiation measurements during the Artemis I lunar mission', *Nature*, 634(8032), pp. 48–52. doi: 10.1038/s41586-024-07927-7.

Gillam, S. *et al.* (1985) 'Gene K of bacteriophage phi X174 codes for a protein which affects the burst size of phage production', *Journal of Virology*, 53(2), pp. 708–709. doi: 10.1128/jvi.53.2.708-709.1985.

Glass, A. *et al.* (2023) 'Efficacy of copper blend coatings in reducing SARS-CoV-2 contamination', *BioMetals*, 36(1), pp. 217–225. doi: 10.1007/s10534-022-00473-7.

Govind, V. *et al.* (2021) 'Antiviral properties of copper and its alloys to inactivate covid-19 virus: a review', *BioMetals*, 34(6), pp. 1217–1235. doi: 10.1007/s10534-021-00339-4.

Greninger, A. L. (2018) 'A decade of RNA virus metagenomics is (not) enough', *Virus Research*. Elsevier, pp. 218–229. doi: 10.1016/j.virusres.2017.10.014.

Grose, J. H. and Casjens, S. R. (2014) 'Understanding the enormous diversity of bacteriophages: The tailed phages that infect the bacterial family Enterobacteriaceae', *Virology*, 468–470, pp. 421–443. doi: https://doi.org/10.1016/j.virol.2014.08.024.

Grün, G., Trimmel, M. and Holm, A. (2012) 'Low humidity in the aircraft cabin environment and its impact on well-being - Results from a laboratory study', *Building and Environment*, 47(1), pp. 23–31. doi: 10.1016/j.buildenv.2011.05.004.

Guarnieri, G. *et al.* (2023) 'Relative Humidity and Its Impact on the Immune System and Infections', *International Journal of Molecular Sciences*. doi: 10.3390/ijms24119456.

Guéguinou, N. *et al.* (2009) 'Could spaceflight-associated immune system weakening preclude the expansion of human presence beyond Earth's orbit?', *Journal of Leukocyte Biology*, 86(5), pp. 1027–1038. doi: 10.1189/jlb.0309167.

Gupta, R. and Brunak, S. (2002) 'Prediction of glycosylation across the human proteome and

the correlation to protein function', in *Pacific Symposium on Biocomputing*. Center for Biological Sequence Analysis, Bldg-208, Bio-Centrum Technical University of Denmark, DK-2800 Lyngby, Denmark, pp. 310–322. Available at: https://psb.stanford.edu/psb-online/proceedings/psb02/abstracts/p310.html.

Hadfield, J. *et al.* (2018) 'Phandango: an interactive viewer for bacterial population genomics', *Bioinformatics*, 34(2), pp. 292–293. doi: 10.1093/bioinformatics/btx610.

Hallgren, J. *et al.* (2022) 'DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks', *bioRxiv*, p. 2022.04.08.487609. doi: 10.1101/2022.04.08.487609.

Hassler, D. M. *et al.* (2014) 'Mars' Surface Radiation Environment Measured with the Mars Science Laboratory's Curiosity Rover', *Science*, 343(6169), p. 1244797. doi: 10.1126/science.1244797.

Heffron, J. *et al.* (2024) 'Mitigation of viruses of concern and bacteriophage surrogates via common unit processes for water reuse: A meta-analysis', *Water Research*, 252, p. 121242. doi: https://doi.org/10.1016/j.watres.2024.121242.

Heikkilä, P. *et al.* (2024) 'Analysis and classification of individual ambient aerosol particles with field-deployable laser-induced breakdown spectroscopy platform', *Aerosol Science and Technology*, 58(9), pp. 1063–1078. doi: 10.1080/02786826.2024.2350022.

Hendrix, R. W. *et al.* (1999) 'Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage', *Proceedings of the National Academy of Sciences*, 96(5), pp. 2192–2197. doi: 10.1073/pnas.96.5.2192.

Hilbig, R. D. I. *et al.* (2011) 'Analysis of Behaviour and Habituation of Fish Exposed to Diminished Gravity in Correlation to Inner Ear Stone Formation - A Sounding Rocket Experiment (TEXUS 45)', in. Available at: https://api.semanticscholar.org/CorpusID:7573635.

Hinkelbein, J. *et al.* (2017) 'Thirty minutes of hypobaric hypoxia provokes alterations of immune response, haemostasis, and metabolism proteins in human serum', *International Journal of Molecular Sciences*, 18(9), p. 1882. doi: 10.3390/ijms18091882.

Hjelmsø, M. H. *et al.* (2019) 'Metagenomic analysis of viruses in toilet waste from long distance flights—A new procedure for global infectious disease surveillance', *PLoS ONE*, 14(1), p. e0210368. doi: 10.1371/journal.pone.0210368.

Hu, M. *et al.* (2021) 'Risk of Coronavirus Disease 2019 Transmission in Train Passengers: an Epidemiological and Modeling Study', *Clinical Infectious Diseases*, 72(4), pp. 604–610. doi: 10.1093/cid/ciaa1057.

Huang, J. and MacKerell Jr, A. D. (2013) ‘CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data’, *Journal of Computational Chemistry*, 34(25), pp. 2135–2145. doi: https://doi.org/10.1002/jcc.23354.

IATA (2023) *Air Transport & Communicable Diseases*. Available at: https://www.iata.org/en/programs/safety/health/diseases/#tab-1.

ICAO (2023) ‘Effects of Novel Coronavirus (COVID-19) on Civil Aviation: Economic Impact Analysis Economic Development of Air Transport Air Transport Bureau International Civil Aviation Organization (ICAO)’.

Irby, I. and Broddrick, J. T. (2024) ‘Microbial adaptation to spaceflight is correlated with bacteriophage-encoded functions’, *Nature Communications*, 15(1), p. 3474. doi: 10.1038/s41467-023-42104-w.

Jain, C. *et al.* (2018) ‘High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries’, *Nature Communications*, 9(1), p. 5114. doi: 10.1038/s41467-018-07641-9.

Bin Jang, H. *et al.* (2019) ‘Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks’, *Nature Biotechnology*, 37(6), pp. 632–639. doi: 10.1038/s41587-019-0100-8.

Jang, H. and Ross, T. M. (2020) ‘Dried SARS-CoV-2 virus maintains infectivity to Vero E6 cells for up to 48 h’, *Veterinary Microbiology*, 251, p. 108907. doi: https://doi.org/10.1016/j.vetmic.2020.108907.

Jayaweera, M. *et al.* (2020) ‘Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy’, *Environmental Research*. Academic Press Inc., p. 109819. doi: 10.1016/j.envres.2020.109819.

Jaydarifard, S., Morawska, L. and Paz, A. (2024) ‘Mitigating airborne infection risks in public transportation: A systematic review’, *Transport Policy*, 155, pp. 309–320. doi: https://doi.org/10.1016/j.tranpol.2024.07.009.

Ji, C.-M. *et al.* (2024) ‘The Applications of Nanopore Sequencing Technology in Animal and Human Virus Research’, *Viruses*. doi: 10.3390/v16050798.

Jurrus, E. *et al.* (2018) ‘Improvements to the APBS biomolecular solvation software suite’, *Protein Science*, 27(1), pp. 112–128. doi: https://doi.org/10.1002/pro.3280.

K., P. A. *et al.* (2023) ‘Persistence of SARS-CoV-2 and its surrogate, bacteriophage Phi6, on

surfaces and in water', *Applied and Environmental Microbiology*, 89(11), pp. e01219-23. doi: 10.1128/aem.01219-23.

Käsermann, F. and Kempf, C. (1998) 'Inactivation of enveloped viruses by singlet oxygen thermally generated from a polymeric naphthalene derivative', *Antiviral Research*, 38(1), pp. 55–62. doi: 10.1016/S0166-3542(98)00007-2.

Kasumi, T. *et al.* (2003) 'Flagellin Glycosylation Island in Pseudomonas syringae pv. glycinea and Its Role in Host Specificity', *Journal of Bacteriology*, 185(22), pp. 6658–6665. doi: 10.1128/jb.185.22.6658-6665.2003.

Katoh, K., Rozewicki, J. and Yamada, K. D. (2019) 'MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization', *Briefings in Bioinformatics*, 20(4), pp. 1160–1166. doi: 10.1093/bib/bbx108.

Kiss, B. *et al.* (2022) 'Imaging the Infection Cycle of T7 at the Single Virion Level', *International Journal of Molecular Sciences*. doi: 10.3390/ijms231911252.

Kiss, J. Z. *et al.* (2019) 'Comparison of Microgravity Analogs to Spaceflight in Studies of Plant Growth and Development', *Frontiers in Plant Science*, 10, p. 1577. doi: 10.3389/FPLS.2019.01577.

Klaus, D. M., Todd, P. and Schatz, A. (1998) 'Functional weightlessness during clinorotation of cell suspensions', *Advances in Space Research*, 21(8), pp. 1315–1318. doi: https://doi.org/10.1016/S0273-1177(97)00404-3.

Klymus, K. E. *et al.* (2020) 'Reporting the limits of detection and quantification for environmental DNA assays', *Environmental DNA*, 2(3), pp. 271–282. doi: https://doi.org/10.1002/edn3.29.

Korencak, M. *et al.* (2022) 'Reconstruction of the origin of the first major SARS-CoV-2 outbreak in Germany', *Computational and Structural Biotechnology Journal*, 20, pp. 2292–2296. doi: https://doi.org/10.1016/j.csbj.2022.05.011.

Kropinski, A. M. *et al.* (2009) 'Enumeration of Bacteriophages by Double Agar Overlay Plaque Assay BT - Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions', in Clokie, M. R. J. and Kropinski, A. M. (eds). Totowa, NJ: Humana Press, pp. 69–76. doi: 10.1007/978-1-60327-164-6_7.

Kruskal, W. H. and Wallis, W. A. (1952) 'Use of Ranks in One-Criterion Variance Analysis', *Journal of the American Statistical Association*, 47(260), pp. 583–621. doi: 10.2307/2280779.

Kumar, R. K. *et al.* (2022) ‘Metabolic modeling of the International Space Station microbiome reveals key microbial interactions’, *Microbiome*, 10(1), p. 102. doi: 10.1186/s40168-022-01279-y.

Kummer, S. and Kranz, D.-C. (2022) ‘Henipaviruses—A constant threat to livestock and humans’, *PLOS Neglected Tropical Diseases*, 16(2), p. e0010157. Available at: https://doi.org/10.1371/journal.pntd.0010157.

Kuraku, S. *et al.* (2013) ‘aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity’, *Nucleic Acids Research*, 41(W1), pp. W22–W28. doi: 10.1093/nar/gkt389.

Kurtz, S. *et al.* (2004) ‘Versatile and open software for comparing large genomes’, *Genome Biology*, 5(2), p. R12. doi: 10.1186/gb-2004-5-2-r12.

Lang, J. M. *et al.* (2017) ‘A microbial survey of the International Space Station (ISS)’, *PeerJ*. Edited by J. Mossong, 5, p. e4029. doi: 10.7717/peerj.4029.

Larralde, M. (2022) ‘Pyrodigal: Python bindings and interface to Prodigal, an efficient method for gene prediction in prokaryotes’, *Journal of Open Source Software*, 7(72), p. 4296. doi: 10.21105/joss.04296.

Laslett, D. and Canback, B. (2004) ‘ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences’, *Nucleic Acids Research*, 32(1), pp. 11–16. doi: 10.1093/nar/gkh152.

Lawrence, P. *et al.* (2023) ‘The natural virome and pandemic potential: Disease X’, *Current Opinion in Virology*, 63, p. 101377. doi: https://doi.org/10.1016/j.coviro.2023.101377.

Lee-Cruz, L. *et al.* (2021) ‘Mapping of Ebola virus spillover: Suitability and seasonal variability at the landscape scale’, *PLOS Neglected Tropical Diseases*, 15(8), p. e0009683. Available at: https://doi.org/10.1371/journal.pntd.0009683.

Leitmeyer, K. and Adlhoch, C. (2016) ‘Review Article: Influenza Transmission on Aircraft: A Systematic Literature Review’, *Epidemiology (Cambridge, Mass.)*, 27(5), pp. 743–751. doi: 10.1097/EDE.0000000000000438.

Leslie, M. (2011) ‘Plant virus gets the bends’, *Journal of Cell Biology*, 193(3), p. 426. doi: 10.1083/jcb.1933iti3.

Leung, M. H. Y. *et al.* (2021) ‘Characterization of the public transit air microbiome and resistome reveals geographical specificity’, *Microbiome*, 9(1), p. 112. doi: 10.1186/s40168-

021-01044-7.

Leung, N. H. L. (2021) ‘Transmissibility and transmission of respiratory viruses’, *Nature Reviews Microbiology*, 19(8), pp. 528–545. doi: 10.1038/s41579-021-00535-6.

Li, F. *et al.* (2023) ‘Stability of mpox virus on different commonly contacted surfaces’, *Journal of Medical Virology*, 95(12), p. e29296. doi: https://doi.org/10.1002/jmv.29296.

Li, H. *et al.* (2009) ‘The Sequence Alignment/Map format and SAMtools’, *Bioinformatics*, 25(16), pp. 2078–2079. doi: 10.1093/bioinformatics/btp352.

Li, H. and Durbin, R. (2009) ‘Fast and accurate short read alignment with Burrows–Wheeler transform’, *Bioinformatics*, 25(14), pp. 1754–1760. doi: 10.1093/bioinformatics/btp324.

Li, W. and Godzik, A. (2006) ‘Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences’, *Bioinformatics*, 22(13), pp. 1658–1659. doi: 10.1093/bioinformatics/btl158.

Lin, K. and Marr, L. C. (2020) ‘Humidity-Dependent Decay of Viruses, but Not Bacteria, in Aerosols and Droplets Follows Disinfection Kinetics’, *Environmental Science & Technology*, 54(2), pp. 1024–1032. doi: 10.1021/acs.est.9b04959.

Liu, B. *et al.* (2019) ‘VFDB 2019: a comparative pathogenomic platform with an interactive web interface’, *Nucleic Acids Research*, 47(D1), pp. D687–D692. doi: 10.1093/nar/gky1080.

Loo, J.A. *et al.* (2010) ‘Comparative Human Salivary and Plasma Proteomes’, *Journal of Dental Research*, 89(10), pp. 1016–1023. doi: 10.1177/0022034510380414.

Louten, J. (2016) ‘Chapter 5 - Virus Transmission and Epidemiology’, in Louten, J. B. T.-E. H. V. (ed.). Boston: Academic Press, pp. 71–92. doi: https://doi.org/10.1016/B978-0-12-800947-5.00005-3.

Lowenthal, M. S. *et al.* (2016) ‘Identification of Novel N-Glycosylation Sites at Noncanonical Protein Consensus Motifs’, *Journal of Proteome Research*, 15(7), pp. 2087–2101. doi: 10.1021/acs.jproteome.5b00733.

Mahmoudabadi, G. and Phillips, R. (2018) ‘A comprehensive and quantitative exploration of thousands of viral genomes’, *eLife*, 7. doi: 10.7554/eLife.31955.

Mann, H. B. and Whitney, D. R. (1947) ‘On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other’, *The Annals of Mathematical Statistics*, 18(1), pp. 50–60. Available at: http://www.jstor.org/stable/2236101.

Mann, N. H. (2005) ‘The third age of phage’, *PLoS biology*, 3(5), pp. 0753–0755. doi: 10.1371/JOURNAL.PBIO.0030182.

Marshall, R. *et al.* (2017) ‘Short DNA containing χ sites enhances DNA stability and gene expression in E. coli cell-free transcription–translation systems’, *Biotechnology and Bioengineering*, 114(9), pp. 2137–2141. doi: https://doi.org/10.1002/bit.26333.

Mattia, P. *et al.* (2022) ‘MetaPhage: an Automated Pipeline for Analyzing, Annotating, and Classifying Bacteriophages in Metagenomics Sequencing Data’, *mSystems*, 7(5), pp. e00741-22. doi: 10.1128/msystems.00741-22.

Mazzoli, M. *et al.* (2023) ‘Spatial immunization to abate disease spreading in transportation hubs’, *Nature Communications*, 14(1), p. 1448. doi: 10.1038/s41467-023-36985-0.

McNair, K. *et al.* (2019) ‘PHANOTATE: a novel approach to gene identification in phage genomes’, *Bioinformatics*, 35(22), pp. 4537–4542. doi: 10.1093/bioinformatics/btz265.

McPherson, A. and DeLucas, L. J. (2015) ‘Microgravity protein crystallization’, *npj Microgravity*, 1(1), p. 15010. doi: 10.1038/npjmgrav.2015.10.

Melis, R. *et al.* (2024) ‘The metabolic footprint of Vero E6 cells highlights the key metabolic routes associated with SARS-CoV-2 infection and response to drug combinations’, *Scientific Reports*, 14(1), p. 7950. doi: 10.1038/s41598-024-57726-3.

Minh, B. Q. *et al.* (2020) ‘IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era’, *Molecular Biology and Evolution*, 37(5), pp. 1530–1534. doi: 10.1093/molbev/msaa015.

Morinet, F. *et al.* (2015) ‘Oxygen and viruses: a breathing story’, *Journal of General Virology*, 96(8), pp. 1979–1982. doi: https://doi.org/10.1099/vir.0.000172.

Mucker, E. M. *et al.* (2024) ‘Comparison of protection against mpox following mRNA or modified vaccinia Ankara vaccination in nonhuman primates’, *Cell*, 187(20), pp. 5540-5553.e10. doi: https://doi.org/10.1016/j.cell.2024.08.043.

Mukai, R., Hamatake, R. K. and Hayashi, M. (1979) ‘Isolation and identification of bacteriophage phi X174 prohead.’, *Proceedings of the National Academy of Sciences*, 76(10), pp. 4877–4881. doi: 10.1073/pnas.76.10.4877.

NASA (1968) *APOLLO 7 MISSION REPORT DECEMBER 1968*. Houston, TX. Available at: https://www.hq.nasa.gov/alsj/a410/A07_MissionReport.pdf (Accessed: 13 June 2021).

NASA (1969) *Apollo 8 Mission Report February 1969*.

Nguyen, P. Q. M. *et al.* (2022) 'Modular micro-PCR system for the onsite rapid diagnosis of COVID-19', *Microsystems & Nanoengineering*, 8(1), p. 82. doi: 10.1038/s41378-022-00400-3.

Nguyen, T. N. *et al.* (2022) 'Enhancement of antibacterial activity by a copper(II) and zinc(II) in chelation with ethylenediaminetetra-acetic acid and urea complex', *Chemical Papers*, 76(11), pp. 7163–7176. doi: 10.1007/s11696-022-02361-3.

Nieto-Caballero, M. *et al.* (2023) 'Carbohydrate vitrification in aerosolized saliva is associated with the humidity-dependent infectious potential of airborne coronavirus', *PNAS Nexus*, 2(2), p. pgac301. doi: 10.1093/pnasnexus/pgac301.

Nishimura, Y. *et al.* (2017) 'ViPTree: the viral proteomic tree server', *Bioinformatics*, 33(15), pp. 2379–2380. doi: 10.1093/bioinformatics/btx157.

Ogando, N. S. *et al.* (2020) 'SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology', *Journal of General Virology*, 101(9), pp. 925–940. doi: https://doi.org/10.1099/jgv.0.001453.

Padgen, M. R. *et al.* (2020) 'EcAMSat spaceflight measurements of the role of σs in antibiotic resistance of stationary phase Escherichia coli in microgravity', *Life Sciences in Space Research*, 24, pp. 18–24. doi: https://doi.org/10.1016/j.lssr.2019.10.007.

Page, A. J. *et al.* (2015) 'Roary: rapid large-scale prokaryote pan genome analysis', *Bioinformatics*, 31(22), pp. 3691–3693. doi: 10.1093/bioinformatics/btv421.

Parra, M. *et al.* (2017) 'Microgravity validation of a novel system for RNA isolation and multiplex quantitative real time PCR analysis of gene expression on the International Space Station', *PLOS ONE*, 12(9), p. e0183480. Available at: https://doi.org/10.1371/journal.pone.0183480.

Paul Shannon, 1 *et al.* (2003) 'Cytoscape: A Software Environment for Integrated Models', *Genome Research*, 13(11), p. 426. doi: 10.1101/gr.1239303.metabolite.

Pedulla, M. L. *et al.* (2003) 'Origins of Highly Mosaic Mycobacteriophage Genomes', *Cell*, 113(2), pp. 171–182. doi: 10.1016/S0092-8674(03)00233-2.

Peng, X. *et al.* (2018) 'Quantitative PCR of T7 Bacteriophage from Biopanning', *JoVE*, (139), p. e58165. doi: doi:10.3791/58165.

Peter, S. (2014) 'Spore Resistance Properties', *Microbiology Spectrum*, 2(5), pp. 10.1128/microbiolspec.tbs-0003–2012. doi: 10.1128/microbiolspec.tbs-0003-2012.

Petersen, T. N. *et al.* (2015) 'Meta-genomic analysis of toilet waste from long distance flights; A step towards global surveillance of infectious diseases and antimicrobial resistance',

Scientific Reports, 5. doi: 10.1038/srep11444.

Pires De Souza, G. A. *et al.* (2022) 'Choosing a cellular model to study SARS-CoV-2', *Frontiers in Cellular and Infection Microbiology*, Volume 12. Available at: https://www.frontiersin.org/journals/cellular-and-infection-microbiology/articles/10.3389/fcimb.2022.1003608.

Poole, K. (2017) 'At the Nexus of Antibiotics and Metals: The Impact of Cu and Zn on Antibiotic Activity and Resistance', *Trends in Microbiology*, 25(10), pp. 820–832. doi: https://doi.org/10.1016/j.tim.2017.04.010.

Potter, S. C. *et al.* (2018) 'HMMER web server: 2018 update', *Nucleic Acids Research*, 46(W1), pp. W200–W204. doi: 10.1093/nar/gky448.

Pung, R. *et al.* (2022) 'Using high-resolution contact networks to evaluate SARS-CoV-2 transmission and control in large-scale multi-day events', *Nature Communications*, 13(1), p. 1956. doi: 10.1038/s41467-022-29522-y.

Purniawan, A. *et al.* (2022) 'Synthesis and assessment of copper-based nanoparticles as a surface coating agent for antiviral properties against SARS-CoV-2', *Scientific Reports*, 12(1), p. 4835. doi: 10.1038/s41598-022-08766-0.

Rack, H. J. and Qazi, J. I. (2006) 'Titanium alloys for biomedical applications', *Materials Science and Engineering: C*, 26(8), pp. 1269–1277. doi: https://doi.org/10.1016/j.msec.2005.08.032.

Raja, F. N. S., Worthington, T. and Martin, R. A. (2023) 'The antimicrobial efficacy of copper, cobalt, zinc and silver nanoparticles: alone and in combination', *Biomedical Materials*, 18(4), p. 45003. doi: 10.1088/1748-605X/acd03f.

Rakowska, P. D. *et al.* (2021) 'Antiviral surfaces and coatings and their mechanisms of action', *Communications Materials*, 2(1), p. 53. doi: 10.1038/s43246-021-00153-y.

De Ree, H. *et al.* (2000) 'Ozone and relative humidity in airline cabins on polar routes: Measurements and physical symptoms', *ASTM Special Technical Publication*. doi: 10.1520/stp14497s.

Reissner, J. *et al.* (2023) 'Stability of Feline Coronavirus in aerosols and dried in organic matrices on surfaces at various environmental conditions', *Scientific Reports*, 13(1), p. 22012. doi: 10.1038/s41598-023-49361-1.

Ren, S. Y. *et al.* (2020) 'Stability and infectivity of coronaviruses in inanimate environments',

World Journal of Clinical Cases. Baishideng Publishing Group Co, pp. 1391–1399. doi: 10.12998/WJCC.V8.I8.1391.

Rheinbaben, F. v. *et al.* (2007) 'Environmental resistance, disinfection, and sterilization of poxviruses BT - Poxviruses', in Mercer, A. A., Schmidt, A., and Weber, O. (eds). Basel: Birkhäuser Basel, pp. 397–405. doi: 10.1007/978-3-7643-7557-7_19.

Ronald, B. *et al.* (2021) 'Higher Concentrations of Bacterial Enveloped Virus Phi6 Can Protect the Virus from Environmental Decay', *Applied and Environmental Microbiology*, 87(21), pp. e01371-21. doi: 10.1128/AEM.01371-21.

Rosca, E. C. *et al.* (2021) 'Transmission of SARS-CoV-2 associated with aircraft travel: a systematic review', *Journal of Travel Medicine*, 28(7), pp. 1–14. doi: 10.1093/JTM/TAAB133.

de Sena Brandine, G. and Smith, A. D. (2021) 'Falco: high-speed FastQC emulation for quality control of sequencing data [version 2; peer review: 2 approved]', *F1000Research*, 8(1874). doi: 10.12688/f1000research.21142.2.

Sengupta, P. *et al.* (2024) 'Genomic, functional, and metabolic enhancements in multidrug-resistant Enterobacter bugandensis facilitating its persistence and succession in the International Space Station', *Microbiome*, 12(1), p. 62. doi: 10.1186/s40168-024-01777-1.

Seok, J., Lee, B. Y. and Yoon, H.-Y. (2024) 'Association between humidity and respiratory health: the 2016–2018 Korea National Health and Nutrition Examination Survey', *Respiratory Research*, 25(1), p. 424. doi: 10.1186/s12931-024-03054-z.

Serrano-Aroca, Á. (2022) 'Antiviral Characterization of Advanced Materials: Use of Bacteriophage Phi 6 as Surrogate of Enveloped Viruses Such as SARS-CoV-2', *International Journal of Molecular Sciences*. doi: 10.3390/ijms23105335.

Shajahan, A. *et al.* (2023) 'Site specific N- and O-glycosylation mapping of the spike proteins of SARS-CoV-2 variants of concern', *Scientific Reports*, 13(1), p. 10053. doi: 10.1038/s41598-023-33088-0.

Shang, Y. *et al.* (2022) 'An improved numerical model for epidemic transmission and infection risks assessment in indoor environment', *Journal of Aerosol Science*, 162, p. 105943. doi: 10.1016/J.JAEROSCI.2021.105943.

Shapiro, S. S. and Wilk, M. B. (1965) 'An Analysis of Variance Test for Normality (Complete Samples)', *Biometrika*, 52(3/4), pp. 591–611. doi: 10.2307/2333709.

Shin, J., Jardine, P. and Noireaux, V. (2012) 'Genome Replication, Synthesis, and Assembly

of the Bacteriophage T7 in a Single Cell-Free Reaction', *ACS Synthetic Biology*, 1(9), pp. 408–413. doi: 10.1021/sb300049p.

Short, K. R. and Cowling, B. J. (2023) 'Assessing the potential for fomite transmission of SARS-CoV-2', *The Lancet Microbe*, 4(6), pp. e380–e381. doi: 10.1016/S2666-5247(23)00099-X.

Silpe, J. E. *et al.* (2023) 'Small protein modules dictate prophage fates during polylysogeny', *Nature*, 620(7974), pp. 625–633. doi: 10.1038/s41586-023-06376-y.

Šístková, J. *et al.* (2024) 'Insight into antibacterial effect of titanium nanotubular surfaces with focus on Staphylococcus aureus and Pseudomonas aeruginosa', *Scientific Reports*, 14(1), p. 17303. doi: 10.1038/s41598-024-68266-1.

Song, W. *et al.* (2023) 'Portable rotary PCR system for real-time detection of Pseudomonas aeruginosa in milk', *Lab on a Chip*, 23(20), pp. 4592–4599. doi: 10.1039/D3LC00401E.

Stacey, H. J., De Soir, S. and Jones, J. D. (2022) 'The Safety and Efficacy of Phage Therapy: A Systematic Review of Clinical and Safety Trials', *Antibiotics*. doi: 10.3390/antibiotics11101340.

Steinmann, J. (2004) 'Surrogate viruses for testing virucidal efficacy of chemical disinfectants', *Journal of Hospital Infection*, 56, pp. 49–54. doi: 10.1016/j.jhin.2003.12.030.

Strathdee, S. A. *et al.* (2023) 'Phage therapy: From biological mechanisms to future directions', *Cell*, 186(1), pp. 17–31. doi: 10.1016/j.cell.2022.11.017.

String, G. M. *et al.* (2022) 'Chlorine efficacy against bacteriophage Phi6, a surrogate for enveloped human viruses, on porous and non-porous surfaces at varying temperatures and humidity', *Journal of Environmental Science and Health, Part A*, 57(8), pp. 685–693. doi: 10.1080/10934529.2022.2101845.

Sun, L. *et al.* (2014) 'Icosahedral bacteriophage ΦX174 forms a tail for DNA transport during infection', *Nature*, 505(7483), pp. 432–435. doi: 10.1038/nature12816.

Sun, Y. *et al.* (2020) 'Shotgun metagenomics of dust microbiome from flight deck and cabin in civil aviation aircraft', *Indoor Air*, 30(6), pp. 1199–1212. doi: https://doi.org/10.1111/ina.12707.

Sunwoo, Y. *et al.* (2006) 'Physiological and subjective responses to low relative humidity', *Journal of physiological anthropology*, 25(1), pp. 7–14. doi: 10.2114/JPA2.25.7.

Sze To, G. N. *et al.* (2009) 'Experimental study of dispersion and deposition of expiratory aerosols in aircraft cabins and impact on infectious disease transmission', *Aerosol Science and Technology*, 43(5), pp. 466–485. doi: 10.1080/02786820902736658.

Tang, S. *et al.* (2020) 'Aerosol transmission of SARS-CoV-2? Evidence, prevention and control.', *Environment International*, 144, pp. 106039–106039. doi: 10.1016/J.ENVINT.2020.106039.

Tapiador, L., Gomez, J. and Vassallo, J. M. (2024) 'Exploring the relationship between public transport use and COVID-19 infection: A survey data analysis in Madrid Region', *Sustainable Cities and Society*, 104, p. 105279. doi: https://doi.org/10.1016/j.scs.2024.105279.

Terzian, P. *et al.* (2021) 'PHROG: families of prokaryotic virus proteins clustered using remote homology', *NAR Genomics and Bioinformatics*, 3(3), p. lqab067. doi: 10.1093/nargab/lqab067.

'The promise of phages' (2023) *Nature Biotechnology*, 41(5), p. 583. doi: 10.1038/s41587-023-01807-7.

Touchon, M., Bernheim, A. and Rocha, E. P. C. (2016) 'Genetic and life-history traits associated with the distribution of prophages in bacteria', *ISME Journal*, 10(11), pp. 2744–2754. doi: 10.1038/ISMEJ.2016.47;TECHMETA=45;SUBJMETA=1321,181,326,631,735;KWRD=BACTERIOPHAGES,MOLECULAR+EVOLUTION.

Urbaniak, C. *et al.* (2018) 'Detection of antimicrobial resistance genes associated with the International Space Station environmental surfaces', *Scientific Reports*, 8(1), pp. 1–13. doi: 10.1038/s41598-017-18506-4.

Valero-Rello, A. and Sanjuán, R. (2022) 'Enveloped viruses show increased propensity to cross-species transmission and zoonosis', *Proceedings of the National Academy of Sciences*, 119(50), p. e2215600119. doi: 10.1073/pnas.2215600119.

Vogele, K. *et al.* (2021) 'Small Antisense DNA-Based Gene Silencing Enables Cell-Free Bacteriophage Manipulation and Genome Replication', *ACS Synthetic Biology*, 10(3), pp. 459–465. doi: 10.1021/acssynbio.0c00402.

Vörösmarty, M., Hopke, P. K. and Salma, I. (2024) 'Attribution of aerosol particle number size distributions to main sources using an 11-year urban dataset', *Atmos. Chem. Phys.*, 24(9), pp. 5695–5712. doi: 10.5194/acp-24-5695-2024.

Walkey, D. G. A. (1991) 'Mechanical Transmission and Virus Isolation BT - Applied Plant Virology', in Walkey, D. G. A. (ed.). Dordrecht: Springer Netherlands, pp. 103–120. doi: 10.1007/978-94-011-3090-5_4.

Wan, Y. and C., M. L. (2012) 'Mechanisms by Which Ambient Humidity May Affect Viruses in Aerosols', *Applied and Environmental Microbiology*, 78(19), pp. 6781–6788. doi:

10.1128/AEM.01658-12.

Wang, D., Tai, P. W. L. and Gao, G. (2019) ‘Adeno-associated virus vector as a platform for gene therapy delivery’, *Nature Reviews Drug Discovery*, 18(5), pp. 358–378. doi: 10.1038/s41573-019-0012-9.

Wang, L. *et al.* (2023) ‘Genomic annotation and molecular evolution of monkeypox virus outbreak in 2022’, *Journal of Medical Virology*, 95(1), p. e28036. doi: https://doi.org/10.1002/jmv.28036.

Warnes, S. L., Summersgill, E. N. and Keevil, C. W. (2015) ‘Inactivation of murine norovirus on a range of copper alloy surfaces is accompanied by loss of capsid integrity’, *Applied and Environmental Microbiology*. doi: 10.1128/AEM.03280-14.

Washizaki, A., Yonesaki, T. and Otsuka, Y. (2016) ‘Characterization of the interactions between Escherichia coli receptors, LPS and OmpC, and bacteriophage T4 long tail fibers’, *MicrobiologyOpen*, 5(6), pp. 1003–1015. doi: https://doi.org/10.1002/mbo3.384.

WHO (2025a) *Global Mpox Trends*. Available at: https://worldhealthorg.shinyapps.io/mpx_global/.

WHO (2025b) *Herpes simplex virus*. Available at: https://www.who.int/news-room/fact-sheets/detail/herpes-simplex-virus (Accessed: 4 August 2025).

Woo, M.-H. *et al.* (2010) ‘Method for contamination of filtering facepiece respirators by deposition of MS2 viral aerosols’, *Journal of Aerosol Science*, 41(10), pp. 944–952. doi: https://doi.org/10.1016/j.jaerosci.2010.07.003.

Xun, G. *et al.* (2021) ‘A rapid, accurate, scalable, and portable testing system for COVID-19 diagnosis’, *Nature Communications*, 12(1), p. 2905. doi: 10.1038/s41467-021-23185-x.

Yang, Z.-C. *et al.* (2024) ‘Ozone, hydrogen peroxide, and peroxymonosulfate disinfection of MS2 coliphage in water’, *Environmental Science: Processes & Impacts*, 26(5), pp. 824–831. doi: 10.1039/D3EM00527E.

Yap, M. L. and and Rossmann, M. G. (2014) ‘Structure and Function of Bacteriophage T4’, *Future Microbiology*, 9(12), pp. 1319–1327. doi: 10.2217/fmb.14.91.

Yuan, S., Zhang, Z.-W. and Li, Z.-L. (2018) ‘Bacteriophage M13 May Be Used for the Assessment of Viral Transfer during Doffing of Ebola-Level Personal Protective Equipment’, *Infection Control & Hospital Epidemiology*. 2018/04/15, 39(6), pp. 762–763. doi: DOI: 10.1017/ice.2018.74.

Yüzbasi, N. S. *et al.* (2022) ‘Removal of MS2 and fr Bacteriophages Using MgAl2O4-Modified, Al2O3-Stabilized Porous Ceramic Granules for Drinking Water Treatment’, *Membranes*. doi: 10.3390/membranes12050471.

Zeng, C. *et al.* (2022) ‘SARS-CoV-2 spreads through cell-to-cell transmission’, *Proceedings of the National Academy of Sciences*, 119(1), p. e2111400119. doi: 10.1073/pnas.2111400119.

Zeng, L. *et al.* (2020) ‘An effective inactivant based on singlet oxygen-mediated lipid oxidation implicates a new paradigm for broad-spectrum antivirals’, *Redox Biology*, 36, p. 101601. doi: 10.1016/j.redox.2020.101601.

Zhang, E. *et al.* (2021) ‘Antibacterial metals and alloys for potential biomedical implants’, *Bioactive Materials*, 6(8), pp. 2569–2612. doi: https://doi.org/10.1016/j.bioactmat.2021.01.030.

Zhang, S. *et al.* (2025) ‘First measurements of the radiation dose on the lunar surface’, *Science Advances*, 6(39), p. eaaz1334. doi: 10.1126/sciadv.aaz1334.

Zhou, P. and Shi, Z.-L. (2021) ‘SARS-CoV-2 spillover events’, *Science*, 371(6525), pp. 120–122. doi: 10.1126/science.abf6097.

Zorz, J. *et al.* (2023) ‘SituSeq: an offline protocol for rapid and remote Nanopore 16S rRNA amplicon sequence analysis’, *ISME Communications*, 3(1), p. 33. doi: 10.1038/s43705-023-00239-3.

Zowawi, H. M. *et al.* (2021) ‘Portable RT-PCR system: A rapid and scalable diagnostic tool for COVID-19 testing’, *Journal of Clinical Microbiology*, 59(5). doi: 10.1128/JCM.03004-20.

Contributors and funding sources

This work was funded by DLR Deutscher Akademischer Austauschdienst (DAAD) felliowship, financed from GANDALF DLR Graduate program. The research was performed at the German Aerospace Center (DLR) Institute of Aerospace Medicine, Cologne. The author is grateful to Dr. Francois-Xavier Lehr, Prof. Dr. Henrike Niederholtmeyer, Dr. Timo Glatter, and Dr. Thomas Heimerl for support, discussions and help on the experiments with phage T7 assembly. Additionally, the author thanks Dr Daniel Moeller and Dr Aisha Ahmed for providing the copper and steel surfaces for testing. The author also thanks Dr. Ronja Anton for assistance in scanning electron microscopy imaging.