Spring 4-26-2022

Diversity of Bacteriophage in Burkholderia Species

Abigail Price
aaprice@bgsu.edu

Diversity of Bacteriophage in *Burkholderia* species

Abigail Price

Honors Project

Submitted to the Honors College at Bowling Green
State University in partial fulfillment of the requirements for
graduation with University Honors

April 2022

Dr. Raymond Larsen, Department of Biological Sciences, Advisor
Dr. Nathan Hensley, Department of Environment and Sustainability, Advisor

Introduction

Antibiotic resistance has been a developing problem over the past few decades and will continue to create problems for a variety of fields. Antibiotic resistance occurs when bacteria evolve to survive the antibiotic medication that has been used to kill them. Bacteria are constantly evolving and are able to quickly evolve to survive antibiotics if they are not administrated appropriately. Some of the main causes of antibiotic resistance include the overuse of antibiotics when not necessary and not finishing a course of antibiotics when prescribed to a patient. The use of antibiotics in agricultural settings can contribute to resistance as some farmers use antibiotics when not necessary to improve the meat quantity without feeding as much. This can cause bacteria to evolve in agricultural animals and be possibly spread to humans through animal product consumption. Bacteria can also become resistant to antibiotics in natural ways through selective pressure and mutations, as the bacteria that survive in the natural environment when faced against organisms' immune systems have been selected for the genes and mutations that enable them to survive a more powerful immune system. Antibiotic resistance has been an impact on the medical, environmental, and microbial fields and will continue to be a problem. There are some solutions to help prevent the further impact of antibiotic resistance to these fields.

One solution to antibiotic resistance could be the use of bacteriophages. Bacteriophages are viruses that infect and replicate within bacteria and offer the potential of a therapeutic alternative to chronic infections that do not respond to antibiotic-based therapies. There are a variety of bacteriophage that can be used to combat the rising number of bacteria that are resistant to traditional antibiotics.

Burkholderia is a genus of bacteria that are similar to *Pseudomonas* and are commonly involved in chronic respiratory infections in individuals with compromised respiratory functions. *Burkholderia vietnamiensis* is one of a number of *Burkholderia* species (collectively termed the *Burkholderia cepacia* complex; "Bcc") that commonly cause chronic drug resistant infections in the lungs of individuals with compromised respiratory systems, as found in those with chronic obstructive pulmonary disease (COPD) and, most especially, are of particular significance in patients with cystic fibrosis (Jassem et al., 2016; Lauman and Dennis, 2021; Naveen et al., 2019). The bacteriophage that infect *Burkholderia* species could provide possible therapeutic options for patients with cystic fibrosis and other chronic respiratory diseases caused by *Burkholderia* species. There are a variety of bacteriophages within these species of *Burkholderia* and the genetic diversity within those bacteriophages is important when trying to identify specific bacteriophage that can combat antibiotic resistance.

The extent of the genetic diversity of *Burkholderia* phage in environmental isolates can be useful when developing bacteriophage therapy for bacteria that have become resistant to common antibiotics. It is difficult to understand the evolutionary history and relationships of bacteriophage and viruses as they evolve so quickly compared to eukaryotic creatures. For researchers to be able to track the evolution of these organisms, there must be a gene or some

segment of DNA that is identifiable and highly conserved within all species. For bacteria, this is the 16S rRNA gene that is highly conserved and when sequenced, can be used for identification of different species. This can be used to predict phylogenetic trees and a genetic history of these organisms evolving throughout time. Bacteriophage lack rRNA, thus development of phylogenetic trees that could help researchers determine the evolutionary history of these bacteriophages must rely on bacteriophage genomic DNA. The sequence of such DNA can be obtained either from DNA purified from bacteriophage, or DNA identified within the genomes of their bacterial hosts.

There are many literature pieces that have assisted in this research and project. Articles written by Arndt et al., Kearse et al., NCBI Resource Coordinators, and Sievers et al., are resources that explain the uses of computer software programs such as NCBI Blast, Clustal Omega, Geneious, and Phaster. These articles provide information on how to use these programs in an efficient way to access and interpret the bacteriophage information available in bacterial genetic databases. Other articles by Jassem et al., Kumar et al., Laumen et al., Mahenthiralingam et al., Rhodes et al., and Ronning et al., provide information on the *Burkholderia* species including the known genetic diversity and how phages can be possibly used for therapy. They also provide information on how certain species within the *Burkholderia* genus are multidrug resistant and often cause respiratory infections and other diseases. They discuss the pathways and mechanisms used by the species to infect humans and then evolve to become multi-drug resistant. Articles written by Bengtsson-Palme et al., Carmody et al., Fuhrmeister et al., Huemer et al., and López-Causapé et al., discuss the wider picture of antibiotic resistance within the medical field and how humans and the environment have contributed to this issue. How phage therapy can be used in patients with Cystic Fibrosis and other pulmonary infections is studied and discussed. These articles will help to relate my research into the bigger picture and interdisciplinary ways of antibiotic resistance and phage therapy techniques. They will also help to find possible solutions to human caused drug resistance and its effects on the environment.

There are a few questions that are guiding the research for this project. The main is: What is the extent of genetic diversity of *Burkholderia* phage in environmental isolates? This can be discovered by working with many online database and programs to explore the different phage within discovered isolates of *Burkholderia.* Another question that could be proposed based on the results of research is: Is there a bacteriophage protein that could serve as a signature that would help trace the evolutionary past of viruses? If this could be discovered, then the range of evolution and diversification within bacteriophages of multiple species and families of viruses could be explored. The implication of this research expands out of the microbiology field and can help within the agricultural and environmental science fields. Since some causes of antibiotic resistance is the misuse of antibiotics by humans, whether that is humans taking the antibiotics or giving the medication to agricultural animals. One possible solution to this continuously evolving issue is education. Education about the environment and how humans impact the planet and diversity of species can be critical to the prevention of

antibiotic resistance. If programs were implemented for more citizens to be educated about the rising issue of drug-resistant bacteria, then the number of variants that resist common antibiotics could lower. The issue of antibiotic resistance also can cause issues within the environment and the species that inhabit it. These resistant bacteria that are created are released back into the natural environment where they could infect other animals and cause the wild populations to decrease. The natural populations of bacteria and viruses could be impacted negatively by the fast-developing bacteria. This makes the discovery of different bacteriophage that could battle resistant bacteria ever important as it can affect more than humans.

Methods

A number of computer software programs such as NCBI Blast, Phaster, Clustal Omega, Microsoft Excel, and Geneious were used. Because very few bacteriophages from *Burkholderia* species have been isolated and sequenced; this work focused on temperate bacteriophage that could be detected in the sequenced genomes of *Burkholderia vietnamiensis.* These programs were used to manage and curate a database of diverse bacteriophage within *Burkholderia* species along with efficiently evaluating each component of these phages. First, ten isolates were searched in the National Center for Biotechnology Information (NCBI)'s database called Basic Local Alignment Search Tool (BLAST). Full sequenced genomes of these isolates were collected from this database for further examination. An online program called PHASTER (PHAge Search Tool – Enhanced Release; Arndt et al., 2016) was then used to scan the genomes of each of the ten *B. vietnamiensis* isolates for containment of bacteriophage originating sequences. Only complete regions of chromosomes were used for the rest of the research. Viral genes that encode proteins common to most or all identified bacteriophage were noted, including the polymerase, major capsid protein, tape measure protein, head protein, portal protein, and integrase. The full sequences of these ten isolates were ran through NCBI's Open Reading Frame Finder to note all of the open reading frames among the genetic code. The open reading frames (ORFs) were then compared to the possible phages found by the PHASTER program to make sure that the possible phages and proteins were true. This was done by finding ORFs and loci that matched in base pair length and checking the sequences of both. An excel sheet was made with all of the matches of ORFs to possible phages.

These full sequences were then further evaluated using additional online DNA analysis software such as Clustal Omega (Sievers et al., 2011). Clustal Omega was used to catalog bacteriophage strains evident in *B. vietnamiensis* genomes and evaluate the relatedness of different strains through phylogenetic trees. The viral genes encoding proteins common to bacteriophage including head proteins, integrase, and portal proteins were researched further. This was done by taking the genetic sequences of all of the head proteins between all of the isolates and aligned though Clustal Omega. This was done the same way for portal proteins and integrase. A *B. vietnamiensis* isolate purified and sequenced in Dr. Larsen's lab was then used to

compare and establish more relationships. This isolate's genetic sequence was put though PHASTER and the possible phage and proteins were annotated in the same excel sheet as the other isolates. The lab isolate sequence was ran through the Open Reading Frame Finder and the ORFs were compared to the possible phages. Finally, this lab isolate was compared to the isolates found on BLAST by running an alignment through Clustal Omega.

Results

The accession numbers of the eight different isolates chosen from BLAST are displayed in Table 1 showing the names of the isolates along with the chromosome and region numbers used for further analysis. Image 1 shows the results of one of the isolates being ran through Phaster. Phaster shows the whole chromosome and has regions of it that contain phages or phage-encoding viral proteins. These viral proteins are all labeled and unidentified proteins are labeled as hypothetical proteins. Image 2 shows the viral proteins and hypothetical proteins that are found in a certain region of the chromosome of the isolate ran through Phaster. Phaster included the genetic sequences of each of the individual proteins found along with the location of those proteins, the homology and E value of the homology. Some of the homologies included terminase, protease, portal protein, tail shaft, integrase, transposase, coat protein, fiber protein, and plate protein. The results of searching for the ORFs in the Open Reading Frame Finder on NCBI are organized in an excel sheet where they are easily compared to the possible phages from Phaster. Image 3 shows one of the isolates ORFs with the matching loci from the chromosomal regions of the same isolate. This shows there are many true phages or viral proteins encoding phages that are in regions of these isolates' chromosomes. This was done for each of the isolates and regions listed in Table 1.

Figure 1 shows a phylogenetic tree that is comprised of all of the isolates from BLAST. It shows that there are two main branches that split into smaller relations between isolates. There are two isolates at the bottom of the tree that are not closely related to any of the others or each other. In Clustal Omega, the integrase, head proteins, and portal proteins were focused on and it was found that head proteins have three distinct families based on the data collected. Integrase had some similarity between the three families at the end of the analysis and portal proteins were not very similar. The head proteins were labeled numerically based on the excel sheet and used to distinguish the proteins in Clustal Omega. The three distinct families were (1 & 9) (7,3,10) (2 & 11). Figure 2 shows a phylogenetic tree of the head proteins ran together and shows the similarities between the families found in the alignment by Clustal Omega. Figure 3 shows a phylogenetic tree of the BLAST isolates with the isolate from Dr. Larsen's lab. This shows that the lab isolate is actually closely related to many of the isolates online and is a good example of common environmental isolates. Now exists a database of bacteriophage and their viral genes that encode different proteins that can be used to expand the knowledge of temperate phage within the *Burkholderia* genus, especially with *B. vietnamiensis.*

Table 1: Isolate Information from BLAST

Isolate Name	Chromosome	Region	Accession Number
ASM171877v1	Chromosome 1	Region 2,3	CP013433.1
ASM171877v1	Chromosome 2	Region 1	CP013434.1
ASM171849v1	Chromosome 1	Region 1	CP013393.1
ASM171849v1	Chromosome 2	Region 1	CP013394.1
ASM171891v1	Chromosome 1	Region 2,3	CP013453.1
ASM171881v1	Chromosome 1	Region 1	CP013439.1
ASM171881v1	Chromosome 3	Region 1	CP013441.1
ASM1620v1	Chromosome 1	Region 1	CP000614.1

Image 1: Results of Phaster showing a chromosome from an isolate with the regions available for viewing of phages

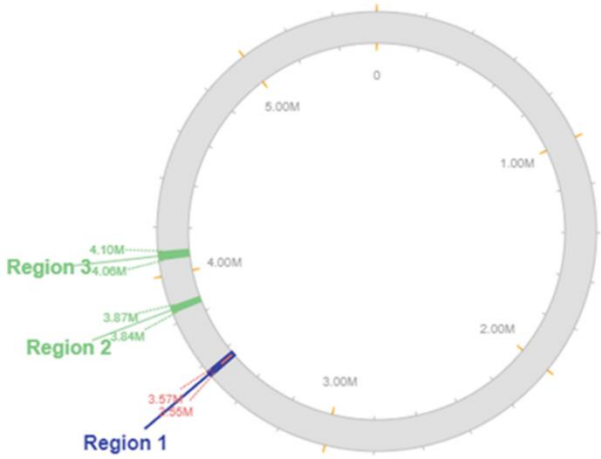


Image 2: Results of Phaster showing all possible phages and phage-encoding viral proteins for one region of a chromosome of an isolate

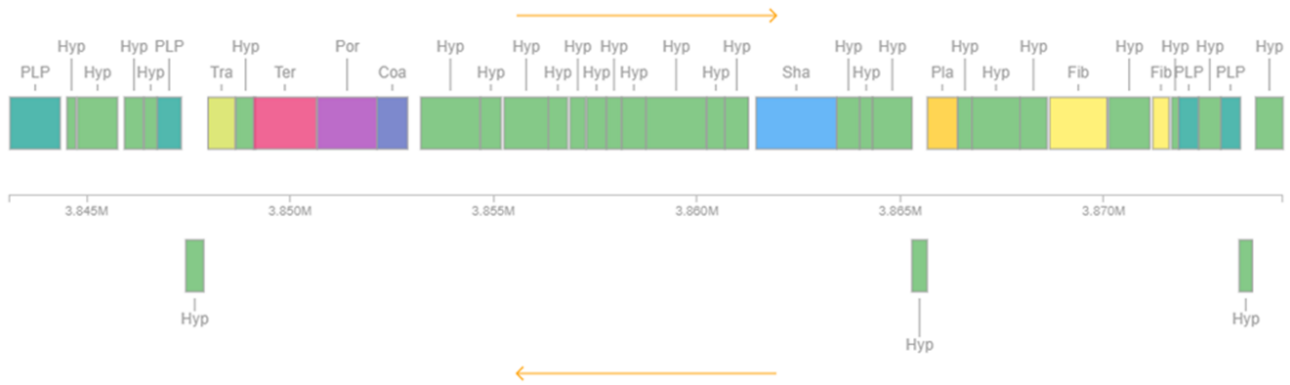


Image 3: Excel sheet sample of ORF and Phaster results comparison for single isolate

ASM171877v1			Chromosome 1		Region 2								
NCBI	**location**	**length (nt)**						**Phaster**	**location**	**length**	**homology**	**Homolgy E value**	
ORF 1	<1-1239	1239						Locus 1: WK23_16875	3843114-3844352	1238 bps	phage-like protein	0	
ORF2	1402-1620	219						Locus 2: WK23_16880	3844515-3844733	218	hypothetical protien	n/a	
ORF36	1658-2647	990						Locus 3: WK23_16885	3844771-3845760	989	hypothetical protein	1.82E-25	
ORF5	3619-4212	594						Locus 6: WK23_16900	3846732-3847325	593	phage-like protein	5.42E-102	
ORF219	4763-4320	444						Locus 7: WK23_16905	3847433-3847876	443	hypothetical protein	n/a	
ORF78	4869-5534	666						Locus 8: WK23_16910	3847982-3848647	665	transposase	1.97E-78	
ORF79	5553-6032	480						Locus 9: WK23_16915	3848666-3849145	479	hypothetical protein	1.99E-43	
ORF45	9020-9775	756						Locus 12: WK23_16930	3852133-3852888	755	head protein	1.52E-56	
ORF46	10097-11557	1461						Locus 13: WK23_16935	3853210-3854670	1460	hypothetical protein	6.15E-56	
ORF47	11567-12073	507						Locus 14: WK23_16940	3854680-3855186	506	hypothetical protein	1.47E-12	
ORF86	13236-13703	468						Locus 16: WK23_16950	3856349-3856816	467	hypothetical protein	n/a	
ORF87	14178-14660	483						Locus 18: WK23_16960	3857291-3857773	482	hypothetical protein	3.20E-32	
ORF17	15040-15630	591						Locus 20: WK23_16970	3858153-3858743	590	hypothetical protein	1.84E-35	
ORF18	15640-17115	1476						Locus 21: WK23_16975	3858753-3860228	1475	hypothetical protein	2.52E-65	
ORF19	17131-17571	441						Locus 22: WK23_16980	3860244-3860684	440	hypothetical protein	4.16E-52	
ORF93	18339-20318	1980						Locus 24: WK23_16990	3861452-3863431	1979	tail protein	2.85E-44	
ORF57	20327-20890	564						Locus 25: WK23_16995	3863440-3864003	563	hypothetical protein	5.07E-26	
ORF24	20890-21207	318						Locus 26: WK23_17000	3864003-3864320	317	hypothetical protein	4.80E-17	
ORF116	25546-22169	378						Locus 28: WK23_17010	3865282-3865659	377	hypothetical protein	n/a	
ORF59	22550-23290	741						Locus 29: WK23_17015	3865663-3866403	740	plate protein	1.60E-53	
ORF98	23298-23651	354						Locus 30: WK23_17020	3866411-3866764	353	hypothetical protein	3.42E-29	
ORF27	25555-26967	1413						Locus 33: WK23_17035	3868668-3870080	1412	fiber protein	8.56E-15	
ORF28	27010-28020	1011						Locus 34: WK23_17040	3870123-3871133	1010	hypothetical protein	n/a	
ORF65	28094-28492	399						Locus 35: WK23_17045	3871207-3871605	398	fiber protein	3.79E-09	
ORF66	28727-29224	498						Locus 37: WK23_17055	3871840-3872337	497	phage-like protein	1.56E-104	
ORF105	29766-30245	480						Locus 39: WK23_17065	3872879-3873358	479	phage-like protein	1.29E-56	

Figure 1: Phylogenetic tree of the isolates found in BLAST

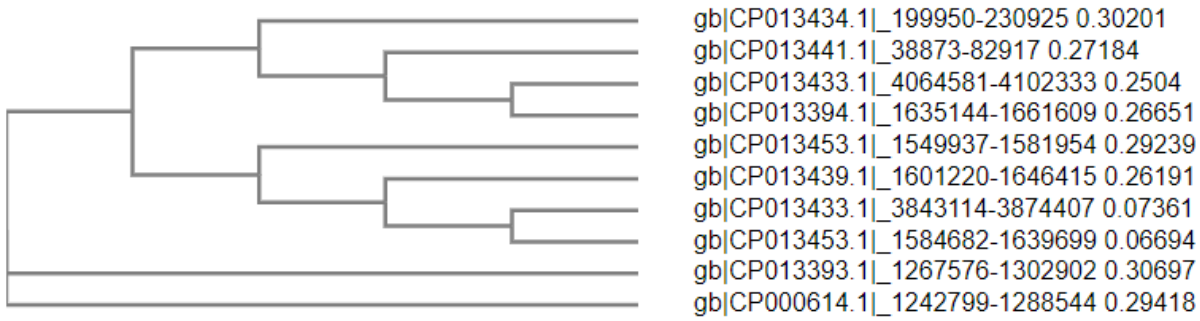


Figure 2: Phylogenetic tree of the head proteins from the BLAST isolates

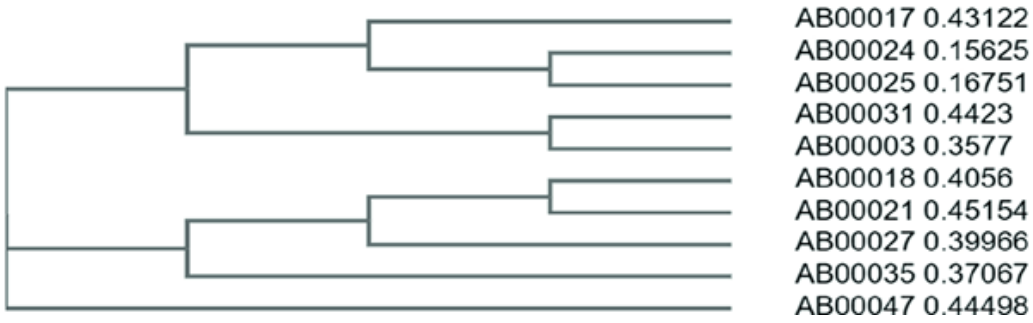
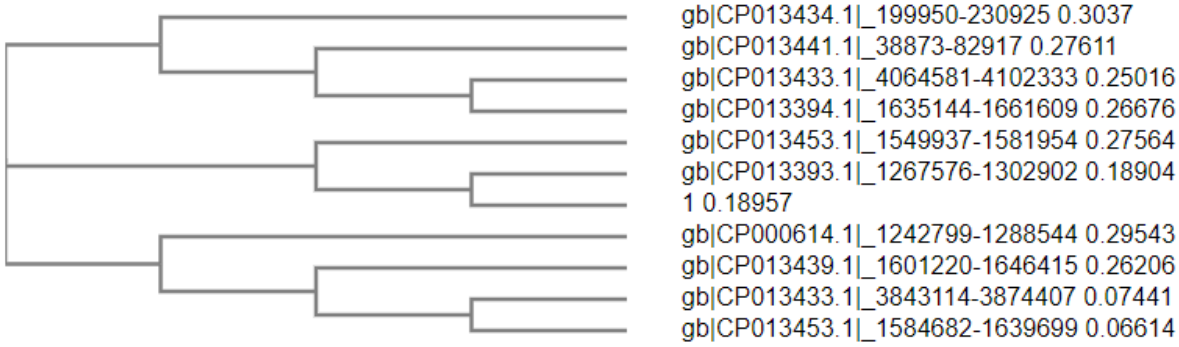


Figure 3: Phylogenetic tree of all of the isolates from BLAST and the lab isolate



Discussion and Future Implications

The results of this project can be useful for future research and will help with the overall fight against antibiotic resistance. The results indicate that there is not much genetic diversity between the isolates collected from BLAST and the lab isolate. This shows that there is has not been lots of genetic exchange between wild isolates of Burkholderia and there are not signs of significant evolution among the variants. The families formed between the head proteins of the isolates indicates that there is not much genetic diversity among head proteins and they are conserved well among stains of isolates. There is a large range of possible phage among the isolates from BLAST and the lab isolate. This means that there is potential for some of these

proteins to be used in the future for bacteriophage therapy to combat drug-resistant bacteria. The collection of data from this project will help future researchers look into the relationships and evolutionary history between *B. vietnamiensis* and possibly predict which variants arise and could cause respiratory infections in people. There is much research that can be done with the data that has been collected in this project and it will serve as a guide to researching bacteriophage relations and evolution more.

This topic of diversity of bacteriophage within the *Burkholderia* genus can be related to many other fields and disciplines. Antibiotic resistance is a wicked problem that is causing issues within the environment and human health and needs researchers from many fields within the sciences to come together to handle the issue. Phage therapy is a possible solution to certain drug resistant bacteria, so my project could possibly assist future researchers in developing phage therapy for these multi-drug resistant bacteria. Antibiotic resistance is caused by a variety of things, but many are the result of human development. This includes the overuse of antibiotics in humans and agricultural animals, unsanitary living conditions, ethically unsound practices, and the transmission of bacteria across the globe. One possible solution to this continuously evolving issue is education. Education about the environment and how humans impact the planet and diversity of species can be critical to the prevention of antibiotic resistance. The issue of antibiotic resistance also affects the health of the environment as bacteria that has become drug-resistant thought the misuse of humans can go back into the environment and infect other organisms and possibly affecting the population of some. This could expand and affect the food chains of some organisms and cause changes in the ecosystem. It is important that research is continued on this topic to ensure the safety and health of all organisms on the planet.

In the future, Dr. Larsen plans to use the data found in this project to further investigate the relationships of bacteriophage in Burkholderia species. From this research, a next step would be to search for a signature protein that could serve as a guide for tracking the evolutionary past of phages. The discovery of this would be useful when comparing isolates, as looking at one protein could show the direct relationship between isolates. This identification of a wide set of bacteriophages will also support the development of such a resource to give future researchers and physicians data that could help in the development and use of bacteriophage within medicinal therapy. These outcomes would be beneficial to the disciplines of microbiology and pathology as this research would help other researchers and scholars looking for data within the Burkholderia genus and how bacteriophage in this genus can be possibly used in the future to develop bacteriophage therapy to treat bacterial infections that are multidrug-resistant, such as Cystic Fibrosis and other pulmonary infections. The research on antibiotic resistance needs to be continues as this is an issue that affects the entire ecosystem and can prevent the recovery of many people with immunocompromised health and severe health issues. Education of this issues is important as one of the biggest causes of resistance is the misuse of antibiotics among humans, whether that is on humans themselves or

on agricultural animals. Antibiotic resistance is an increasingly large issue that needs to be researched further to see if bacteriophage therapy can be a possible and realistic solution.

References

David Arndt, Jason R. Grant, Ana Marcu, Tanvir Sajed, Allison Pon, Yongjie Liang, and David S. Wishart (2016). PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44(Web Server issue): W16–W21. doi: 10.1093/nar/gkw387

Fabian Sievers, Andreas Wilm, David Dineen, Toby J Gibson, Kevin Karplus, Weizhong Li, Rodrigo Lopez, Hamish McWilliam, Michael Remmert, Johannes Söding, Julie D Thompson, and Desmond G Higgins (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7: 539. Published online 2011 Oct 11. doi: 10.1038/msb.2011.75

Jassem, A. N., Zlosnik, J. E., Henry, D. A., Hancock, R. E., Ernst, R. K., & Speert, D. P. (2011). In vitro susceptibility of Burkholderia vietnamiensis to aminoglycosides. Antimicrobial agents and chemotherapy, 55(5), 2256–2264. https://doi.org/10.1128/AAC.01434-10

Naveen Kumar, Devanga Ragupathi, Balaji Veeraraghavan (2019). Accurate identification and epidemiological characterization of Burkholderia cepacia complex: an update.Ann Clin Microbiol Antimicrob. 18: 7. Published online 2019 Feb 4. doi: 10.1186/s12941-019-0306-0

NCBI Resource Coordinators (2016). Database resources of the National Center for Biotechnology Information, Nucleic Acids Res. 44 (Database issue): D7–D19. Published online 2015 Nov 28. doi: 10.1093/nar/gkv1290

Philip Lauman, Jonathan J. Dennis (2021) Advances in Phage Therapy: Targeting the Burkholderia cepacia Complex. Viruses. 2021 Jul; 13(7): 1331. Published online 2021 Jul 9. doi: 10.3390/v13071331