

**Comparative Genomic studies for
understanding the hypothetical proteins in
Pseudomonas phages for understanding their
co-evolution with *Pseudomonas* species**

By

Krishna Khairnar, Rajshree Chandekar, Swapnil Sanmukh, Waman Paunikar

International **E – Publication**

www.isca.me , www.isca.co.in



Comparative Genomic studies for understanding the hypothetical proteins in *Pseudomonas* phages for understanding their co-evolution with *Pseudomonas* species

By

Krishna Khairnar

Environmental Virology Cell, Council of Scientific and Industrial Research-National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur-440020, Maharashtra (India).

E-mail: kskhairnar@gmail.com

Rajshree Chandekar

Environmental Virology Cell, Council of Scientific and Industrial Research -National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur-440020, Maharashtra (India).

E-mail: rajshri1022@gmail.com

Swapnil Sanmukh

Environmental Virology Cell, Council of Scientific and Industrial Research -National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur-440020, Maharashtra (India).

E-mail: swamukh1985in@rediffmail.com

Waman Paunikar

Environmental Virology Cell, Council of Scientific and Industrial Research -National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur-440020, Maharashtra (India).

E-mail: wn_paunikar@neeri.res.in

2014

International E - Publication

www.isca.me , www.isca.co.in



International E - Publication

427, Palhar Nagar, RAPTC, VIP-Road, Indore-452005 (MP) INDIA

Phone: +91-731-2616100, Mobile: +91-80570-83382

E-mail: contact@isca.co.in , Website: www.isca.me , www.isca.co.in

© Copyright Reserved

2014

All rights reserved. No part of this publication may be reproduced, stored, in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, reordering or otherwise, without the prior permission of the publisher.

ISBN: 978-93-83520-86-2

Comparative Genomic studies for understanding the hypothetical proteins in *Pseudomonas* phages for understanding their co-evolution with *Pseudomonas* species

Krishna Khairnar*, Rajshree Chandekar, Swapnil Sanmukh, Waman Paunekar

Environmental Virology Cell, Council of Scientific and Industrial Research-National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur-440020, Maharashtra (India)

E-mail: Khairnar Krishna (kskhairnar@gmail.com); Rajshree Chandekar (rajshri1022@gmail.com); Swapnil Sanmukh (swamukh1985in@rediffmail.com); Waman Paunekar (wn_paunekar@neeri.res.in)

***Corresponding Author Address:** Krishna Khairnar, Environmental Virology Cell, Council of Scientific and Industrial Research - National Environmental Engineering Research Institute (CSIR-NEERI), Nehru Marg, Nagpur-440020, Maharashtra, India. Email: kskhairnar@gmail.com

Abstract

The Pseudomonas phages are the bacterial viruses infecting Pseudomonas species. The role of Pseudomonas phages in evolution of Pseudomonas species through transduction and prophagic conversion is well known but studies related to the prediction of functional proteins in Pseudomonas phages are not yet reported. Hence, there is a need to understand the genetic makeup of Pseudomonas phages imparting various characteristics to their host. The present paper is an attempt to understand some of the known Pseudomonas phages by studying their hypothetical proteins. The complete genome sequence of 35 Pseudomonas phages were screened for presence of hypothetical proteins which revealed the presence of 853 unclassified proteins. The probable function prediction of the 784 hypothetical proteins was possible among the 853 hypothetical proteins found in all 35 Pseudomonas phages by using bioinformatics web tools like CDD-Blast, Interproscan, pfam and COGs. The subcellular localization predictions of hypothetical protein were carried out by using CELLO v 2.5. The structure prediction for 290 hypothetical proteins was possible by using PS2 protein modeling server. This study revealed the presences of various functional domains in hypothetical proteins of Pseudomonas phages which are yet to be discovered. These interpreted data for hypothetical proteins can prove helpful for the understanding of structural, functional,

evolutionary and metabolic development of Pseudomonas phages and its life cycle along with their role in host co-evolution.

Keywords: Bacterial viruses; Genetic makeup; Prophagic Conversion; Bioinformatics web tools; Subcellular localization; Co-evolution.

1. Introduction

The use of bacteriophages as highly specific antimicrobial agents is widely documented in the literature ^{1,2}. For therapeutic uses, obligatory lytic bacteriophages are highly desirable, as they result in rapid killing of their target host cell, bacteriophage numbers increase rapidly and transduction is relatively rare. DNA sequencing of bacteriophage genomes is now used to confirm both identity and the absence of undesirable elements, such as functional lysogenic components or bacterial toxins. Early attempts to use bacteriophages for therapy of bacterial infections were compromised by a lack of understanding of the nature of the agents involved ³.

The genomic sequencing of bacteriophages has proved to be helpful in examination of lysogenic cassettes and bacterial toxin genes. In *Pseudomonas aeruginosa* bacteriophage FCTX toxin gene is present, it forms pores in cell membranes ^{4,5}. Several important points emerged when the predicted proteins were examined for similarity to known bacterial and bacteriophage sequences that were deposited in public databases. First, the phage proteomes are rich sources of untapped protein sequence diversity. The biological function of a large proportion of predicted proteins cannot be determined by comparison to current entries in genome databases. By providing primary sequence information from a large group of *P. aeruginosa* phages, this report not only provides a compendium of novel protein sequences, but also sets the stage for future studies aimed at better understanding virus/host relationships.

Pseudomonas aeruginosa is increasingly recognized as an emerging opportunistic pathogen. The members of the genus demonstrate a great deal of metabolic diversity and have the ability to metabolise a variety of diverse nutrients which helps them to colonize a wide range of niches ⁶. They can produce exopolysaccharides known as biofilms such as alginate, making it difficult to be phagocytosed ⁷. *Pseudomonas* sp can thrive in harsh conditions due to their cell wall that contains porins. Their resistance to most antibiotics is attributed to efflux pumps, which pump out some antibiotics before the antibiotics are able to act. *Pseudomonas* sp has low antibiotic susceptibility due

to multidrug efflux pumps with chromosomally-encoded antibiotic resistance genes e.g. mexAB-oprM, mexXY, etc ⁸. Besides intrinsic resistance, *Pseudomonas* sp easily develops acquired resistance either by mutation in chromosomally-encoded genes, or by the horizontal gene transfer of antibiotic resistance determinants ⁸.

Some members of the genus *Pseudomonas* are able to metabolize chemical pollutants toluene ^{9, 10}, cyanide ¹¹, carbazole ^{12, 13, 14}, carbon tetrachloride ^{15, 16}, pentachlorophenol ¹⁷, catechol ¹², etc in the environment, and as a result can be used for bioremediation. As a result of their ability to grow at low temperatures and ubiquitous nature due to metabolic diversity, many *Pseudomonas* spp. can cause food spoilage which includes dairy spoilage by *P. fragi* ¹⁸, *P. lundensis*, which causes spoilage of milk, cheese, meat, and fish and many more ¹⁹.

We are particularly interested to identify and understand the structure and functional characteristics attributed by hypothetical proteins present in the *Pseudomonas* phages. Bioinformatics web tools like CDD-Blast (<http://www.ncbi.nlm.nih.gov/BLAST/>)²⁰⁻²⁶, Interproscan (<http://www.abi.ac.uk/interpro>)²⁷⁻³⁰ Pfam (<http://www.pfam.sanger.ac.uk/>)^{31,32} COGs (<http://www.ncbi.nlm.nih.gov/COG/>)³³⁻³⁶ can predict the functionality of the hypothetical proteins by comparing it with biological sequence databases, whereas, PS2 protein modeling server (<http://www.ps2.life.nctu.edu.tw/>)³⁷⁻⁴³ can assist in determination of 3-D structure of the target sequence. The in-silico computational biology study of *Pseudomonas* phages hypothetical proteins (Uncharacterized proteins) will assist us to predict the functionality in these protein sequences using the different strategies of comparative proteomics.

2. Methodology

2.1 Sequence retrieval

The Complete gene sequences from whole genomes of 35 different *Pseudomonas* phages viz. *Pseudomonas* phage PaP3 ⁴³, *Pseudomonas* phage PAK_P1 ⁴⁴, *Pseudomonas* phage phiIBB-PF7A ¹⁶, *Pseudomonas* phage KPP10 ⁴⁵, *Pseudomonas* phage phi15 ⁴⁶, *Pseudomonas* phage 14-1 ⁴⁷, *Pseudomonas* phage LKA1 ⁴⁸, *Pseudomonas* phage SN ⁴⁷, *Pseudomonas* phage LMA2 ⁴⁷, *Pseudomonas* phage LBL3 ⁴⁷, *Pseudomonas* phage LIT1 ⁴⁹, *Pseudomonas* phage LUZ7 ⁴⁹, *Pseudomonas* phage phi-2 ⁵⁰, *Pseudomonas* phage YuA ⁵¹, *Pseudomonas* phage LUZ19 ⁵², *Pseudomonas* phage LUZ24 ³², *Pseudomonas* phage phikF77 ⁵³, *Pseudomonas* phage MP38 ⁵⁴, *Pseudomonas* phage MP29 ⁵⁴, *Pseudomonas* phage Pf3 ⁵⁵, *Pseudomonas* phage F8 ⁵⁶, *Pseudomonas* phage M6 ^{51, 56}, *Pseudomonas* phage PA11 ⁵⁶, *Pseudomonas* phage 119X ⁵⁶, *Pseudomonas* phage 73

⁵⁶, Pseudomonas phage F10 ⁵⁶, Pseudomonas phage F116 ⁵⁷, Pseudomonas phage gh-1 ⁵⁸, Pseudomonas phage MP22 ⁵⁹, Pseudomonas phage LKD16 ⁴⁸; Pseudomonas phage D3112 ⁶⁰, Pseudomonas phage PAJU2 ⁴⁵, Pseudomonas phage EL ⁶¹; Pseudomonas phage PaP2 ⁶² and Pseudomonas phage B3 ⁶³ were downloaded from the KEGG Database (<http://www.genome.jp/kegg/>).

2.2 Functional Annotations and Characterizations

The function annotation and characterization of hypothetical proteins in 35 different *Pseudomonas* phages were carried out by using the bioinformatics web tools like CDD-Blast (<http://www.ncbi.nlm.nih.gov/BLAST/>) ²⁰⁻²⁶; Interproscan (<http://www.abi.ac.uk/interpro>) ²⁷⁻³⁰; Pfam (<http://www.pfam.sanger.ac.uk/>) ³¹ and COGs (<http://www.ncbi.nlm.nih.gov/COG/>) ^{33-36, 64}.

2.3 Percentage of similarity and confidence level

The annotation of gene functions for the unidentified hypothetical proteins analyzed by the function prediction web tools such as CDD-BLAST, INTERPROSCAN, PFAM and COGs shows variable results with different confidence levels when searched for the conserved domains. The confidence level of the annotated gene can be determined on the basis of collective results of these web- tools ⁶⁵.

1. If the given four tools indicate the same functions then the confidence level were to be 100 percent.
2. If the given three tools indicate the same functions and one is showing different function then the confidence level were to be 75 percent.
3. If the given two tools indicate the same functions and two are showing different functions then the confidence level were to be 50 percent.
4. If only one tool indicates the function and other tools are showing different functions then the confidence level were to be 25 percent ⁶⁵.

2.4 Prediction of Subcellular Localization

The predictions of sub-cellular localization of hypothetical proteins were determined by using CELLO v.2.5 (<http://cello.life.nctu.edu.tw/>) ^{66, 67}.

2.5 Protein Structure Predictions

The 3-D protein structure predictions of the provided hypothetical protein gene were determined by using PS2 protein structure prediction server (<http://www.ps2.life.nctu.edu.tw/>) ^{20, 26-28, 37-42, 68}. The

server accepts the protein (query) sequences in FASTA format and uses the strategies of Pair-wise and multiple alignments to generate resultant proteins 3D structures.

3. Results and Discussions

Our thanks to the experts who have contributed towards development of the template. The comparative genomic studies for understanding the hypothetical proteins of 35 different *Pseudomonas* phages were carried out successfully by using different bioinformatic tools like CDD-Blast, InterProscan, Pfam and COGs. In all 853 hypothetical proteins were screened and characterized by using sequence similarity search with close orthologous family members available in various protein databases using the web tools. The function prediction in 784 hypothetical proteins was possible which are represented in the respective tables viz. *Pseudomonas* phage PaP3 Table 1, *Pseudomonas* phage PAK_P1 Table 2, *Pseudomonas* phage phiIBB-PF7A Table 3, *Pseudomonas* phage KPP10 Table 4, *Pseudomonas* phage phi15 Table 5, *Pseudomonas* phage 14-1 Table 6, *Pseudomonas* phage LKA1 Table 7, *Pseudomonas* phage SN Table 8, *Pseudomonas* phage LMA2 Table 9, *Pseudomonas* phage LBL3 Table 10, *Pseudomonas* phage LIT1 Table 11, *Pseudomonas* phage LUZ7 Table 12, *Pseudomonas* phage phi-2 Table 13, *Pseudomonas* phage YuA Table 14, *Pseudomonas* phage LUZ19 Table 15, *Pseudomonas* phage LUZ24 Table 16, *Pseudomonas* phage phikF77 Table 17, *Pseudomonas* phage MP38 Table 18, *Pseudomonas* phage MP29 Table 19, *Pseudomonas* phage Pf3 Table 20, *Pseudomonas* phage F8 Table 21, *Pseudomonas* phage M6 Table 22, *Pseudomonas* phage PA11 Table 23, *Pseudomonas* phage 119X Table 24, *Pseudomonas* phage 73 Table 25, *Pseudomonas* phage F10 Table 26, *Pseudomonas* phage F116 Table 27, *Pseudomonas* phage gh-1 Table 28, *Pseudomonas* phage MP22 Table 29, *Pseudomonas* phage LKD16 Table 30; *Pseudomonas* phage D3112 Table 31; *Pseudomonas* phage PAJU2 Table 32; *Pseudomonas* phage EL Table 33; *Pseudomonas* phage PaP2 Table 34 and *Pseudomonas* phage B3 Table 35. The 3-D structures predictions of protein were done by using online automated PS² protein structure prediction server. The structure prediction was satisfactorily carried out in 290 hypothetical proteins depending upon the sequence similarity and by using best scored orthologous template available in the submitted protein sequence through the server. The generated 3-D structures are represented in the order as Template ID, Identity, Score and E-value in structure column from Table 1 through Table 35 of each *Pseudomonas* phage analyzed.

4. Conclusions

We can conclude from the above *in-silico* studies that the generated information from the analysis of the hypothetical proteins from 35 different *Pseudomonas* phages can help us to understand the functional characteristics of the important genes which are yet to be characterized practically. Moreover, the study also signifies the importance of these proteins in modifying the host to which the phages may infect through prophagic conversion and transduction. The characterized 784 hypothetical proteins give a deeper insight about their role in *Pseudomonas* host-phage interaction and their objective by imparting novel characteristics to its host and affecting metabolic mechanism through their infection strategies. The predicted structures for 290 hypothetical proteins can be verified by cloning the specific genes in expression vectors and then we will be able to establishing their role in life cycle of *Pseudomonas* phages as well as the host they infect. However, the study also clearly resolves the unknown nature of hypothetical proteins from all the 35 different analysed *Pseudomonas* phages and helps us to understand the *Pseudomonas* phages with respect to their co-evolution with their host.

5. Acknowledgements

R.H.C and S. G. S. carried out extensive bioinformatics work and data curation for annotation of genes for identified hypothetical proteins in *Pseudomonas* phages. W.N.P., K.S.K., R.H.C. and S.S. critically edited the manuscript along with referencing and literature review. Manuscript was written by S.G.S., R.H.C., W.N.P. and K.S.K.

6. References

- [1] Harper D.R. and Kutter E., Bacteriophage: therapeutic uses, The Encyclopedia of Life Sciences. Chichester: John Wiley & Sons (2008)
- [2] Kutter E. and Sulakvelidze A., Bacteriophages: Biology and Applications, Boca Raton, FL: CRC Press (2005)
- [3] Krueger A.P. and Scribner E.J., The Bacteriophage: its nature and its therapeutic use, *JAMA* **116**, 2160-2167 (1941)
- [4] Ceysens P.J. and Lavigne R., Bacteriophages of *Pseudomonas*, *Future Microbiol* **5**, 1041-1055 (2010)

- [5] Baltch A.L., Smith R.P., Franke M., Ritz W., Michelsen P. and Bopp L., Pseudomonas aeruginosa cytotoxin as a pathogenicity factor in a systemic infection of leukopenic mice, *Toxicon* **32**, 27-34(1994)
- [6] Madigan M. and Martinko J., Brock Biology of Microorganisms, (11th ed.), Prentice Hall. ISBN 0-13-144329-1. (2005)
- [7] Ryan K.J. and Ray C.G., Sherris Medical Microbiology, (4th ed.), McGraw Hill. ISBN, 0-8385-8529-9 (2004)
- [8] Poole K., Efflux-mediated multiresistance in Gram-negative bacteria, *Clin. Microbiol. Infect.*, **10** (1), 12–26(2004)
- [9] Marques S. and Ramos J.L., Transcriptional control of the Pseudomonas putida TOL plasmid catabolic pathways, *Molecular Microbiology*, **9**(5), 923–929, doi:10.1111/j.1365-295(1993)
- [10] Yen K.M., Karl M.R., Blatt L.M., Simon M.J., Winter R.B. and Fausset P.R.. Cloning and characterization of a Pseudomonas mendocina KR1 gene cluster encoding toluene-4-monooxygenase, *J. Bacteriol.*, **173** (17), 5315–27 (1991)
- [11] Huertas M.J., Luque-Almagro V.M., Martinez-Luque M., Blasco R., Moreno-Vivián C. and Castillo F.. Cyanide metabolism of Pseudomonas pseudoalcaligenes CECT5344: role of siderophores, *Biochem. Soc. Trans.* **34** (Pt 1), 152–5(2006)
- [12] Nojiri H., Maeda K., Sekiguchi H., Urata M., Shintani M. and Yoshida T.. Organization and transcriptional characterization of catechol degradation genes involved in carbazole degradation by Pseudomonas resinovorans strain CA10, *Biosci. Biotechnol. Biochem.* **66** (4), 897–901(2002)
- [13] O'Mahony M.M., Dobson A.D., Barnes J.D. and Singleton I., The use of ozone in the remediation of polycyclic aromatic hydrocarbon contaminated soil, *Chemosphere* **63** (2), 307–314(2006)
- [14] Onaca C., Kieninger M., Engesser K.H. and Altenbuchner J., Degradation of alkyl methyl ketones by Pseudomonas veronii, *Journal of Bacteriology* **189** (10), 3759–3767(2007)
- [15] Sepúlveda-Torres L.C., Rajendran N., Dybas M.J. and Criddle C. S., Generation and initial characterization of Pseudomonas stutzeri KC mutants with impaired ability to degrade carbon tetrachloride, *Arch Microbiol* **171** (6) 424–429(1999)
- [16] Sillankorva S., Kluskens L. D., Lingohr E. J., Kropinski A. M., Neubauer P. and Azeredo J., Complete genome sequence of the lytic Pseudomonas fluorescens phage phiIBB-PF7A, *Virol. J.* **8** (1), 142(2011)

- [17] Nam I.H., Chang Y.S., Hong H.B., and Lee Y.E., A novel catabolic activity of *Pseudomonas veronii* in biotransformation of pentachlorophenol, *Applied Microbiology and Biotechnology* **62** (2–3), 284–290, doi:10.1007/s00253-003-1255-1. PMID 12883877(2003)
- [18] Pereira J.N., and Morgan M.E., Nutrition and physiology of *Pseudomonas fragi*, *J Bacteriol.*, **74** (6), 710–3. PMC 289995. PMID 13502296 (1957)
- [19] Gennari M. and Dragotto F., A study of the incidence of different fluorescent *Pseudomonas* species and biovars in the microflora of fresh and spoiled meat and fish, raw milk, cheese, soil and water, *J Appl Bacteriol.*, **72** (4), 281–8, PMID 1517169(1992)
- [20] Altschul S. F., Madden T. L., Schaffer A. A., Zhang J., Zhang Z., Miller W. and Lipman D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic Acids Res.* **25** (17), 3389-402, (1997)
- [21] Aron M. B., John B. A., Myra K. D., Carol D. S., Noreen R. G., Marc G., Luning H., Siqian H., David I. H., John D. J., Zhaoxi K., Dmitri K., Christopher J. L., Cynthia A. L., Chunlei L., Fu L., Shennan L., Gabriele H. M., Mikhail M., James S. S., Narmada T., Roxanne A. Y., Jodie J. Y., Dachuan Z. and Stephen H. B., CDD: a conserved domain database for interactive domain family analysis, *Nucleic Acids Research*, Vol. **35**, D237–D240 (2006)
- [22] Cédric N., Desmond G. H. and Jaap H., T-coffee: a novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol.*, **302**, 205-217 (2000)
- [23] Marchler-Bauer A., Anderson J.B., Derbyshire M.K., DeWeese-Scott C., Gonzales N.R. and Gwadz M. L., CDD: a conserved domain database for interactive domain family analysis, *Nucleic Acids Res.* Jan;**35** (Database issue):D237-40. Epub 2006 Nov 29 (2007)
- [24] Marchler-Bauer A., Anderson J.B., Derbyshire M.K., DeWeese-Scott C., Gonzales N.R. and Gwadz M., CDD: a Conserved Domain Database for the functional annotation of proteins, *Nucleic Acids Res.***39**(D), 225-9 (2011)
- [25] Marchler-Bauer A., and Bryant S.H., CD-Search: protein domain annotations on the fly, *Nucleic Acids Res.***32** (W) 327-331(2004)
- [26] Schaffer A. A., Aravind L., Madden T. L., Shavirin S., Spouge J. L. and Wolf Y. I., Improving the accuracy of PSI-BLAST protein database searches with composition-based statistics and other refinements, *Nucleic Acids Res.* **29**(14), 2994-3005(2001)
- [27] Burge S., Kelly E., Lonsdale D., Mutowo-Muellenet P., McAnulla C. and Mitchell A., Manual GO annotation of predictive protein signatures: the InterPro approach to GO curation, Database; doi: 10.1093/database/bar06 (2012)

- [28] Quevillon E., Silventoinen V., Pillai S., Harte N., Mulder N., and Apweiler R., InterProScan: protein domains identifier, *Nucleic Acids Res.* **33** (Web Server issue):W116-W120(**2005**)
- [29] Hunter S., Jones P., Mitchell A., Apweiler R., Attwood T. K. and Bateman A., InterPro in 2011: new developments in the family and domain prediction database, *Nucleic Acids Research* 2011; doi: 10.1093/nar/gkr948(**2011**)
- [30] Zdobnov E. M. and Rolf A., Interproscan- an integration platform for the signatures recognition methods in InterPro:, *Bioinformatics* **17**,847-848(**2001**)
- [31] Alex B., Lachlan C., Richard D., Robert D. F., Volker H., Sam G. J. and Ajay K., The Pfam families' database, *Nucleic Acids Research*, Vol. **32**, D138-D141(**2004**)
- [32] Ceysens P.J., Demeke M., Lavigne R., Ackermann H.W., Noben J.P. and Volckaert G., Analysis of the genome, proteome and transcriptome of *Pseudomonas aeruginosa* phage LUZ24” (Unpublished)
- [33] Galperin M.Y., Comparing microbial genomes: How the gene set determines the lifestyle. In Organization of the Prokaryotic Genome, Charlebois R.L., Ed. (*American Society of Microbiology*, Washington, DC) pp. 91-108 (**1999**)
- [34] Koonin E.V., Tatusov R.L. and Galperin M.Y., Beyond complete genomes: from sequence to structure and function, *Curr. Opin. Struct. Biol.* **8**, 355-363(**1998**)
- [35] Roman L. T., Michael Y. G., Darren A. N. and Eugene V. K., The COG database: a tool for genome –scale analysis of protein functions and evolution, *Nucleic Acid Research.* **28**, 33-36(**2000**)
- [36] R.L. Tatusov, E.V. Koonin and D.J. Lipman. “A genomic perspective on protein families”, *Science* 278: 631-637(**1997**)
- [37] Chen C.C., Hwang J.K. and Yang J.M., (PS)2: protein structure prediction server, *Nucleic Acid Research*, Vol. **34**, W152-W157, 2006.
- [38] Fiser A., Do R.K. and Sali A., Modeling of loops in protein structures, *Protein Science* **9**, 1753-1773(**2000**)
- [39] Hung L.H., Ngan S.C., Liu T. and Samudrala R., PROTINFO: New algorithms for enhanced protein structure prediction”, *Nucleic Acids Research*, **33**, W77-W80(**2005**)
- [40] Laskowski R. A., MacArthur M. W., Moss D. S. and Thornton J. M., PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Cryst.*, **26**, 283-291(**1993**)

- [41] Marti-Renom M.A., Stuart A., Fiser A., Sanchez R., Melo F. and Sali A., Comparative protein structure modeling of genes and genomes, *Annu. Rev. Biophys. Biomol. Struct.*, 29, 291-325(2000)
- [42] Notredame C., Higgins D. and Heringa J., T-Coffee: A novel method for multiple sequence alignments, *J. Mol. Biol.*, **302**, 205-217(2000)
- [43] Tan Y., Zhang K., Rao X., Jin X., Huang J. and Zhu J., Whole genome sequencing of a novel temperate bacteriophage of *P. aeruginosa*: evidence of tRNA gene mediating integration of the phage genome into the host bacterial chromosome, *Cell. Microbiol.* **9 (2)**, 479-491(2007)
- [44] L. Debarbieux, D. Leduc, D. Maura, E. Morello, A. Criscuolo and O. Grossi. "Bacteriophages can treat and prevent *Pseudomonas aeruginosa* lung Infections", *J. Infect. Dis.* 201 (7), 1096-1104, 2010.
- [45] Uchiyama J., Rashel M., Matsumoto T., Sumiyama Y., Wakiguchi H. and Matsuzaki S., Characteristics of a novel *Pseudomonas aeruginosa* bacteriophage, PAJU2, which is genetically related to bacteriophage D3, *Virus Res.*, **139 (1)**, 131-134 (2009)
- [46] Cornelissen A., Ceysens P.J., T'Syen J. and Van Praet H., The T7-Related *Pseudomonas putida* Phage phi15 Displays Virion-Associated Biofilm Eradication Properties, (Unpublished)
- [47] Ceysens P.J., Miroshnikov K., Mattheus W., Krylov V., Robben J. and Noben J.P., Comparative analysis of the widespread and conserved PB1-like viruses infecting *Pseudomonas aeruginosa*, *Environ. Microbiol.*, **11 (11)**, 2874-2883(2009)
- [48] Ceysens P. J., Lavigne R., Mattheus W., Chibeu A., Hertveldt K. and Mast J., Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: establishment of the phiKMV subgroup within the T7 supergroup, *J. Bacteriol.*, **188 (19)**, 6924-6931(2006)
- [49] Ceysens P.J., Brabban A., Rogge L., Lewis M.S., Pickard D. and Goulding D., Molecular and physiological analysis of three *Pseudomonas aeruginosa* phages belonging to the N4-like viruses, *Virology*, **405 (1)**, 26-30 (2010)
- [50] Paterson S., Vogwill T., Buckling A., Benmayor R., Spiers A.J. and Thomson N.R., Antagonistic coevolution accelerates molecular evolution, *Nature*, **464 (7286)**, 275-278(2010)
- [51] Ceysens P.J., Mesyanzhinov V., Sykilinda N., Briers Y., Roucourt B. and Lavigne R., The genome and structural proteome of YuA, a new *Pseudomonas aeruginosa* phage resembling M6, *J. Bacteriol.*, **190 (4)**, 1429-1435(2008)

- [52] Lammens E., Ceysens P.J., Voet M., Hertveldt K., Lavigne R. and Volckaert G., Representational Difference Analysis (RDA) of bacteriophage genomes, *J. Microbiol. Methods*, **77** (2), 207-213(2009)
- [53] Kulakov L.A., Ksenzenko V.N., Kochetkov V.V., Mazepa V.N. and Boronin A.M., DNA homology and adsorption specificity of *Pseudomonas aeruginosa* virulent bacteriophages, *Mol. Gen. Genet.*, **200** (1), 123-127 (1985)
- [54] Lee Y.R., Chung I.Y., Heo Y.J., and Cho Y.H., Direct Submission, (Unpublished)
- [55] Luiten R.G., Putterman D.G., Schoenmakers J.G., Konings R.N. and Day L.A., Nucleotide sequence of the genome of Pf3, an IncP-1 plasmid-specific filamentous bacteriophage of *Pseudomonas aeruginosa*. *J. Virol.*, **56** (1), 268-276 (1985)
- [56] Kwan T., Liu J., Dubow M., Gros P. and Pelletier J., Comparative Genomic Analysis of 18 *Pseudomonas aeruginosa* Bacteriophages, *J. Bacteriol.*, **188** (3), 1184-1187(2006)
- [57] Byrne M. and Kropinski A.M., The genome of the *Pseudomonas aeruginosa* generalized transducing bacteriophage F116, *Gene*, **346**, 187-194(2005)
- [58] Kovalyova I. V. and Kropinski A. M., The complete genomic sequence of lytic bacteriophage gh-1 infecting *Pseudomonas putida*--evidence for close relationship to the T7 group, *Virology*, **311** (2), 305-315(2003)
- [59] Heo Y.J., Chung I.Y., Choi K.B., Lau G.W. and Cho Y.H., Genome sequence comparison and superinfection between two related *Pseudomonas aeruginosa* phages, D3112 and MP22", *Microbiology* (Reading, Engl.) **153** (PT 9), 2885-2895 (2007)
- [60] Wang P.W., Chu L. and Guttman D.S., Complete sequence and evolutionary genomic analysis of the *Pseudomonas aeruginosa* transposable bacteriophage D3112, (Unpublished)
- [61] Hertveldt K., Lavigne R., Pleteneva E., Sernova N., Kurochkina L. and Korchevskii R., Genome comparison of *Pseudomonas aeruginosa* large phages, *J. Mol. Biol.* **354** (3), 536-545(2005)
- [62] Hu F., Huang J., Zhang K., Tan Y., Hu X., Rao X., Zhu J., and Jin X., Complete genome sequence of *Pseudomonas aeruginosa* bacteriophage PaP2, (Unpublished)
- [63] Braid M.D., Silhavy J.L., Kitts C.L., Cano R.J. and Howe M.M., Complete genomic sequence of bacteriophage B3, a Mu-like phage of *Pseudomonas aeruginosa*, *J. Bacteriol.*, **186** (19), 6560-6574(2004)
- [64] Tatusov R.L., Galperin M.Y., Natale D.A. and Koonin E.V., The COG database: a tool for genome-scale analysis of protein functions and evolution, *Nucleic Acids Res.* **28**, 33-6(2000)

- [65] Anandakumar S. and Shanmughavel P., Computational Annotation for Hypothetical Proteins of Mycobacterium Tuberculosis, *Journal of Computer Science & Systems Biology*, **1**, 50-62(2008)
- [66] Yu C.S., Chen Y.C., Lu C.H. and Hwang J.K., Prediction of protein subcellular localization”, *Proteins: Structure, Function and Bioinformatics*, **64**,643-651(2006)
- [67] Yu C. S., Lin C. J. and Hwang J. K., Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n-peptide compositions, *Protein Science*, **13**,1402-1406 (2004)
- [68] Samudrala R. and Levitt M., A comprehensive analysis of 40 blind protein structure predictions”, *BMC Structural Biology* **2**, 3-18(2002)

Table 1. *Pseudomonas* phage PaP3

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
2700534	PHA01782 super family[cl10333]	No	Firmicute fructose-1,6-bisphosphatase	No	Cytoplasmic 2.337 Periplasmic 2.120	No
2700536	No	No	Dopey, N-terminal;DnaB-helicase binding domain of primase		Cytoplasmic 2.643	No
2700545	No	No	Ubiquitin fold modifier 1 protein	No	Cytoplasmic 1.847 Periplasmic 1.782	
2700546	No	No	Gyrovirus capsid protein (VP1)	No	Cytoplasmic 2.720	No
2700547	No	No	Progressive ankylosis protein (ANKH)	No	InnerMembrane 2.512	
2700548	No	<i>Pseudomonas</i> phage PaP3, Orf56	Protein of unkNown function (DUF2566)	No	InnerMembrane 2.963	
2700549	No	No	Cellulase (glycosyl hydrolase family 5)	No	Cytoplasmic 2.726	
2700556	Phage phiEco32-like COOH.NH2 ligase-type 2;	No	Phage phiEco32-like COOH.NH2 ligase-type 2	No	Cytoplasmic 4.021	No
2700559	No	Glutamine amidotransferase, type II	Glutamine amidotransferase domain	Glucosamine 6-phosphate synthetase, contains amidotransferase and phosphosugar isomerase domains	Cytoplasmic 2.724	1ecfB 16 297 3e-81
2700561	Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase) [Coenzyme metabolism / Translation, ribosomal structure and biogenesis	ATP-grasp fold, subdomain 2	D-ala D-ala ligase C-terminus	D-alanine-D-alanine ligase and related ATP-grasp enzymes	OuterMembrane 1.413 Cytoplasmic 1.197 InnerMembrane 1.123	2q7dA- 13- 43- 1e-04
2700570	No	No	Salmonella virulence plasmid 28.1kDa A protein	No	Cytoplasmic 2.855 Periplasmic 1.076	No
2700573	No	IPR012340 Nucleic acid-binding, OB-	No	No	Cytoplasmic 2.962 Periplasmic 1.482	1je5B- 24- 114- 2e-26

		fold					
2700575	No	No	Acetyltransferase (GNAT) family	No	Cytoplasmic Periplasmic	2.531 1.339	No
2700577	CtsR super family[cl01850], Firmicute transcriptional repressor of class III stress genes (CtsR); This family consists of several Firmicute transcrFirmicute transcriptional repressor of class III stress genes (CtsR); This family consists of several Firmicute transcriptional repressor of class III stress genes (CtsR) proteins. CtsR of L. moNocytogenes negatively regulates the clpC, clpP and clpE genes belonging to the CtsR regulon.	No	No	No	Cytoplasmic	4.893	No
2700578	Uncharacterized ACR; family of bacterial proteins present mostly in environmental bacteria and metageNomes	No	Uncharacterized ACR	No	Cytoplasmic Periplasmic	2.621 1.254	No
2700581	No	No	CENP-A-nucleosome distal (CAD) centromere subunit	No	Cytoplasmic	2.929	
2700586	No	No	No	No	OuterMembrane	3.128	2pziA- 10- 30- 0.009
2700592	Phage Tail Collar Domain	No	Phage Tail Collar Domain	No	Extracellular	3.568	1pdiA- 13- 54- 2e-08
2700603	No	No	Terminase small subunit	No	Cytoplasmic	4.262	No

Table 2. *Pseudomonas* phage B3

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
3192561	Phage virion morphogenesis family	Phage virion morphogenesis protein	Phage virion morphogenesis family	No	OuterMembrane 1.524 Periplasmic 1.345 Cytoplasmic 1.302	No
3192562	Putative phage tail protein	No	Putative phage tail protein	No	Cytoplasmic 2.016 OuterMembrane 1.402	No
3192563	No	Bacteriophage B3, Orf5;Uncharacterised conserved protein, UCP028111 type	No	No	Cytoplasmic 3.809	No
3192564	No	Bacteriophage B3, Orf6	No	No	Cytoplasmic 4.496	3bt7A 21 35 0.010
3192566	No	No	No	No	Cytoplasmic 4.717	2ij2A 20 36 0.002
3192567	No	Protein of unkNown function UCP028589	No	No	Extracellular 1.614 OuterMembrane 1.537 Periplasmic 1.486	1lfwA 30 36 0.005
3192568	No	Bacteriophage Mu, Gp16	No	No	Cytoplasmic 2.516	2ijd1 15 39 0.001
3192570	HTH super family[cl00088], Helix-turn-helix domains	Transcription regulator IclR, N-terminal;Winged helix-turn-helix transcription repressor DNA-binding	IclR helix-turn-helix domain; AAA domain	Transcriptional regulator	Cytoplasmic 3.161	2ia2A 17 47 6e-07
3192572	No	No	Conserved hypothetical protein 2217	No	OuterMembrane 2.116 Cytoplasmic 1.707	2e6jA 14 33 0.002
3192574	Phage Mu protein F like protein	Phage head morphogenesis domain	Phage Mu protein F like protein	No	Cytoplasmic 2.469 Periplasmic 2.076	No
3192576	No	Bacteriophage Mu, Gp36	No	No	Cytoplasmic 3.159	2dgkA 17 35 0.009
3192583	No	No	COG (conserved oligomeric Golgi) complex component, COG2 ;Intermediate filament protein ;UV radiation resistance protein and autophagy-related subunit 14 ;Mer2	No	Cytoplasmic 4.708	No
3192588	No	No	Cortexillin I, coiled coil ;Replication initiation factor	No	Cytoplasmic 1.796 Periplasmic 1.247	No

3192590	nitrilase super family[cl11424], Nitrilase superfamily, including nitrile- or amide-hydrolyzing enzymes and amide-condensing enzymes	No	No	No	OuterMembrane Extracellular	1.863 1.677	No
3192593	Phage_BR0599 super family;phg_TIGR02218[TIGR02218], phage conserved hypothetical protein BR0599;	Bacteriophage phiJL001, Gp84;Bacteriophage phiJL001, Gp84, C-terminal;Bacteriophage phiJL001, Gp84, N-terminal	Phage conserved hypothetical protein BR0599 ;Helix-destabilising protein	No	Extracellular	3.054	No
3192595	No	No	Transposase ;Death domain	No	Periplasmic	2.383	No
3192602	No	Burkholderia phage BcepMu, Gp37	Gp37 protein	No	Cytoplasmic	3.647	
3192604	No	No	Membrane-bound lysozyme-inhibitor of c-type lysozyme	No	Periplasmic Cytoplasmic	2.255 1.673	No
3192609	No	No	Ligand-gated ion channel	No	Cytoplasmic	2.163	No
3192612	Helix-turn-helix XRE-family like proteins	No	Helix-turn-helix domain of resolvase	No	Periplasmic	2.477	No
3192613	No	No	UL49 family ;Sec23/Sec24 zinc finger ;zinc-ribbon domain	No	Cytoplasmic	2.862	No
3192614	No	No	Centrosome microtubule-binding domain of Cep57	No	Cytoplasmic Periplasmic OuterMembrane	1.366 1.323 1.119	No

Table 3. *Pseudomonas* phage PaP2

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
2846021	No	No	Growth arrest and DNA-damage-inducible proteins-interacting protein 1	No	OuterMembrane 1.483 Cytoplasmic 1.283 Periplasmic 1.129	No
2846025	No	No	Prokaryotic membrane lipoprotein lipid attachment site	No	Cytoplasmic 1.980	No
2846026	No	No	No	Kef-type K+ transport systems, membrane components	InnerMembrane 3.618	No
2846029	No	Ribonuclease H-like domain	No	No	Cytoplasmic 3.084	No
2846031	No	No	7tm Chemosensory receptor	No	InnerMembrane 2.314	No
2846035	VanY super family[cl00813], D-alanyl-D-alanine carboxypeptidase	No	D-alanyl-D-alanine carboxypeptidase	No	Periplasmic 2.784	No
2846037	No	No	Tripartite tricarboxylate transporter TctA family	No	Periplasmic 1.981	No
2846044	No	No	Helix-turn-helix domain	No	Cytoplasmic 2.254 Periplasmic 1.774	No
2846053	No	No	No	No	Periplasmic 1.903 InnerMembrane 1.383	1InzA- 19 -30- 0.009
2846055	No	No	Fibronectin-binding protein (FBP)	No	Cytoplasmic 2.605	No
2846056	No	No	Coiled-coil domain-containing protein (DUF2037)	No	Cytoplasmic 2.798	No
2846062	No	No	Phorbol esters/diacylglycerol binding domain (C1 domain) ;Spt4/RpoE2 zinc finger	No	Cytoplasmic 3.372	No
2846063	No	No	Phorbol esters/diacylglycerol binding domain (C1 domain);Spt4/RpoE2 zinc finger	No	Cytoplasmic 2.239	No

2846067	No	No	Oxidoreductase family, NAD-binding Rossmann fold	No	Periplasmic	3.126	No
2846069	No	No	MazG nucleotide pyrophosphohydrolase domain	No	Cytoplasmic	3.923	No
2846070	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases	No	AAA domain	ABC-type multidrug transport system, ATPase component	Cytoplasmic Periplasmic	1.903 1.586	1dekA- 13- 87- 2e-18
2846072	No	No	Primase C terminal 2 (PriCT-2)	No	Cytoplasmic	3.326	No
2846076	No	No	AAA domain	No	Cytoplasmic Periplasmic	2.414 1.753	No
2846078	Cas4_I-A_I-B_I-C_I-D_II-B super family[cl00641], CRISPR/Cas system-associated protein Cas4	No	PD-(D/E)XK nuclease superfamily	No	Cytoplasmic	3.745	No

Table 4. *Pseudomonas* phage EL

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5176610	No	No	Macrocin-O-methyltransferase (TyIF)	No	Cytoplasmic 4.041	No
5176613	No	No	No	No	Cytoplasmic 4.536	2q8uA -18- 49- 3e-06
5176614	No	No	DNA polymerase family B; DNA polymerase family B	No	Cytoplasmic 3.086	2p5oB 13 39 2e-05
5176616	No	No	Bacterial DNA-binding protein	No	Cytoplasmic 2.776	No
5176617	No	No	No	Orotate phosphoribosyltransferase	Cytoplasmic 3.505	1hqmD 15 145 1e-35
5176624	HNH_3 super family[cl16264], HNH endonuclease;	HNH nuclease	HNH endonuclease; Putative glycoside hydrolase xylanase; NUMOD1 domain	No	Cytoplasmic 2.216 OuterMembrane 1.607	1u3eM 25 162 9e-41
5176634	No	No	AAA domain ; Nucleopolyhedrovirus P10 protein	Chromosome segregation ATPases	Cytoplasmic 3.977	1xd4A 20 36 0.010
5176635	No	No	Biotin-requiring enzyme	DNA gyrase (topoisomerase II) A subunit	Cytoplasmic 4.513	2a6hC 20 372 1e-104
5176638	No	No	No	No	Periplasmic 1.579 OuterMembrane 1.428 Extracellular 1.190	1lzIA 20 33 0.002
5176643	No	No	AMIN domain	Chaperonin GroEL (HSP60 family)	Cytoplasmic 4.299	No
5176646	No	No	No	No	Cytoplasmic 4.204	1vzyA 14 33 0.002
5176647	No	No	Opacity-associated protein A LysM-like domain	No	Extracellular 1.801 OuterMembrane 1.457 Periplasmic 1.246	No
5176648	No	No	No	No	Cytoplasmic 2.051 OuterMembrane 1.911	1hqmD- 14- 214- 2e-56
5176650	No	No	No	No	Cytoplasmic 4.780	2pmzA 18 123 7e-29
5176652	No	No	Glycosyl hydrolases family 31 ; DNA polymerase processivity factor	No	Extracellular 2.393	No

5176653	Arc super family[cl12211], Arc-like DNA binding domain	Arc-like DNA binding; Ribbon-helix-helix; Arc-type ribbon-helix-helix	Arc-like DNA binding domain	No	Cytoplasmic	2.899	1mntA 33 77 1e-15
5176654	No	Tubulin/FtsZ, GTPase domain	No	No	Cytoplasmic OuterMembrane Periplasmic	1.736 1.198 1.052	No
5176658	No	No	T-complex protein 10 C-terminus	No	Cytoplasmic	3.904	No
5176669	No	No	No	Nucleoside diphosphate kinase	Cytoplasmic	2.444	No
5176671	No	No	No	No	Cytoplasmic	2.371	2c7yB 50 36 0.004
5176672	No	No	Glyoxal oxidase N-terminus	No	Cytoplasmic	4.110	No
5176673	No	No	Collagen triple helix repeat (20 copies); Collagen triple helix repeat (20 copies)	No	OuterMembrane Periplasmic Extracellular	1.793 1.321 1.161	No
5176676	No	No	No	No	Cytoplasmic	2.910	1n71A- 21- 32 -0.005
5176680	No	No	No	No	Cytoplasmic	4.267	1n8zC 16 36 0.009
5176681	Nudix_Hydrolase super family	No	No	No	Cytoplasmic	3.911	No
5176682	No	No	Saccharopine dehydrogenase	No	Cytoplasmic	3.409	No
5176683	Ntn hydrolases (N-terminal nucleophile) super family	No	No	No	Cytoplasmic	2.496	1iruH 21 67 2e-12
5176687	No	No	BCL7, N-terminal conserver region	No	Cytoplasmic	2.223	No
5176688	No	No	DMPK coiled coil domain like ;	No	Cytoplasmic	3.310	No
5176689	No	No	SH3 domain	No	Cytoplasmic	3.474	No
5176690	No	No	Tc5 transposase DNA-binding domain ; Apolipoprotein A1/A4/E domain	No	Cytoplasmic Periplasmic	2.047 1.546	No
5176704	No	No	Nickel-containing superoxide dismutase; RNA binding protein She2p	No	Cytoplasmic	3.248	No
5176706	No	No	No	No	Cytoplasmic	3.999	2j69A 9 33 0.002

5176709	spore_SleB; spore cortex-lytic enzyme	EF-Hand 1, calcium-binding site; EF-HAND 2	Putative peptidoglycan binding domain; Transglycosylase SLT domain; EF hand	No	OuterMembrane 1.659 Periplasmic 1.322 Extracellular 1.126	3bkhA 16 38 0.002
5176713	No	No	Lantibiotic streptin immunity protein; SEA domain	No	Cytoplasmic 4.520	2q1fA 10 38 0.005
5176716	No	DNA-directed RNA polymerase, subunit 2, domain 6	RNA polymerase Rpb2, domain 6	DNA-directed RNA polymerase beta subunit/140 kD subunit (split gene in Mjan, Mthe, Aful)	Cytoplasmic 3.170	No
5176721	No	No	MechaNosensitive ion channel; Baculovirus 11 kDa family; Putative ATP-dependent Lon protease	No	Cytoplasmic 4.171	No
5176723	No	No	Podoplanin	No	Extracellular 1.957	No
5176732	No	No	Clr5 domain	No	Cytoplasmic 2.532	No
5176734	No	No	Enterobacterial putative membrane protein (DUF943)	No	Cytoplasmic 4.798	No
5176737	No	No	No	No	OuterMembrane 2.283	1fnfA 27 37 0.006
5176740	phiKZ-like phage internal head proteins	Bacteriophage phiKZ, Orf92, internal head	phiKZ-like phage internal head proteins ; Calmodulin binding protein-like	No	OuterMembrane 2.384 Cytoplasmic 2.295	No
5176741	No	No	Bombolitin family	No	Cytoplasmic 2.732	No
5176742	phiKZ_IP[pfam12699], phiKZ-like phage internal head proteins	Bacteriophage phiKZ, Orf92, internal head	phiKZ-like phage internal head proteins; Prominin ; Spc7 kinetochore protein ; Reovirus sigma C capsid protein ; Modifier of rudimentary (Mod(r)) protein ; Autophagy protein Apg17 ; AAA domain ; Tat binding protein 1(TBP-1)-interacting protein (TBPIP) ; IncA protein ; Apolipoprotein A1/A4/E domain ; Growth-arrest specific micro-tubule binding ; Septum formation	No	Cytoplasmic 3.136	No

			initiator ; Nucleopolyhedrovirus P10 protein ; Laminin Domain II ; Tumour suppressor protein ; DivIVA protein ; COG (conserved oligomeric Golgi) complex component, COG2 ; GDP/GTP exchange factor Sec2p;				
5176743	No	No	Reovirus sigma C capsid protein ; Spectrin repeat; Dip2/Utp12 Family; Apolipoprotein III precursor (apoLp-III) ; TipAS antibiotic- recognition domain	Methyl- accepting chemotaxis protein	Cytoplasmic	3.606	
5176745	No	No	Imelysin ; AAA domain ; Inner membrane protein import complex subunit Tim54 ; Invariant surface glycoprotein	No	Cytoplasmic	3.648	No
5176746	No	No	No	No	Cytoplasmic	2.030	2eutA 36 36 0.005
5176751	No	No	No	No	Cytoplasmic	2.677	2q74A -23- 32 -0.002
5176752	No	No	Mediator complex subunit 2	No	Cytoplasmic	3.643	No
5176753	No	No	Ribbon-helix-helix protein, copG family	No	Cytoplasmic	4.665	No
5176755	TRFH super family[cl02779],Te lomic Repeat binding Factor or TTAGGG Repeat binding Factor, central (dimerization) domain Homology	No	Telomere repeat binding factor (TRF) ; Seryl-tRNA synthetase N-terminal domain; Tumour suppressor protein	Methyl- accepting chemotaxis protein	Cytoplasmic	3.507	2i1kA- 17 -42 -1e-04
5176756	No	No	Neuraminidase	No	Cytoplasmic Periplasmic Extracellular	1.623 1.183 1.008	No
5176760	No	No	Replication factor A protein 3	No	Cytoplasmic	2.268	No
5176763	No	Ribonuclease H- like domain	No	No	Cytoplasmic	2.702	1qhtA- 13- 31- 0.007
5176764	No	No	AAA domain	No	Cytoplasmic	3.386	No
5176767	No	No	ParB-like nuclease domain	No	Cytoplasmic	3.732	No

5176770	No	No	Predicted coiled-coil domain-containing protein (DUF2360); Bul1 N terminus	No	Periplasmic Cytoplasmic Extracellular	1.588 1.451 1.129	No
5176771	No	No	Initiator binding protein 39 kDa	No	Cytoplasmic Periplasmic Extracellular	1.373 1.171 1.163	1jqgA- 25 -36 -0.002
5176772	PRK09946 super family[cl14575], hypothetical protein	No	No	No	Cytoplasmic Extracellular Periplasmic	1.712 1.424 1.046	
5176773	CE4_SF super family[cl15692], Catalytic NodB homology domain of the carbohydrate esterase 4 superfamily: AAA_25[pfam13481], AAA domain;	No	AAA domain ; Nucleopolyhedrovirus P10 protein; Rio2, N-terminal	No	Cytoplasmic	4.471	1q57A- 11- 55- 3e-10
5176778	No	No	Cyclin-dependent kinase inhibitor	No	Cytoplasmic	3.365	No
5176779	SbcC[COG0419], ATPase involved in DNA repair [DNA replication, recombination, and repair]; RecF[COG1195], Recombinational DNA repair ATPase (RecF pathway)	Prokaryotic chromosome segregation/condensation protein MukB, N-terminal	AAA ATPase domain; RecF/RecN/SMC N terminal domain; Hr1 repeat	ATPase involved in DNA repair	Cytoplasmic	3.921	too long
5176780	No	No	No	ABC-type exporter of toluene and related compounds, ATPase component	Cytoplasmic OuterMembrane	3.795 1.942	1hjrA -20- 41- 2e-04
5176788	Merozoite surface antigen 2c(MSA-2c) super family	No	Eukaryotic domain of unkNown function (DUF1716)	No	OuterMembrane	2.135	1xi4A -35 -32 -0.003
5176790	No	No	Putative ATP-dependent Lon protease	No	Cytoplasmic	2.791	No
5176791	No	No	Hok/gef family	No	Cytoplasmic Periplasmic InnerMembrane	1.491 1.182 1.019	No
5176792	No	No	Vpu protein ; Phage virion morphogenesis family ; Double-strand	No	Periplasmic	2.049	

			recombination repair protein				
5176798	No	No	LisH	No	Cytoplasmic	3.582	No
5176804	No	No	No	No	Extracellular Cytoplasmic OuterMembrane	1.427 1.386 1.352	2r7eA- 20- 35- 0.010
5176806	2'-5' RNA ligase superfamily	No	2'-5' RNA ligase superfamily	No	Cytoplasmic	3.370	No

Table 5. *Pseudomonas* phage PAJU2

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
6989627	No	No	Neuraminidase ; DNA polymerase II large subunit DP2	No	Cytoplasmic 1.793 Periplasmic 1.658	No
6989628	No	No	F subunit of K+-transporting ATPase (Potass_KdpF)	No	InnerMembrane 1.545 Cytoplasmic 1.538 Periplasmic 1.102	No
6989629	No	No	YebO-like protein	No	Cytoplasmic 2.707	No
6989635	No	No	WW domain	No	Periplasmic 2.174 Cytoplasmic 1.525	No
6989638	Bro-N super family[cl10591], BRO family, N-terminal domain(This family includes the N-terminus of baculovirus BRO and ALI motif proteins); ORF6C super family[cl11181], ORF6C domain	BRO N-terminal domain	BRO family, N-terminal domain ; ORF6C domain	Prophage antirepressor	Cytoplasmic 3.396	No
6989645	No	No	YnbE-like lipoprotein	No	Periplasmic 2.643	No
6989646	No	No	Hemagglutinin esterase	No	Cytoplasmic 2.711	No
6989648	No	No	ET module	No	Cytoplasmic 3.449	No
6989650	DUF1833 super family[cl07467],	Bacteriophage D3, Orf22	No	No	Cytoplasmic 2.948	No
6989654	HTH_MerR-SF super family[cl02600], Helix-Turn-Helix DNA binding domain of transcription regulators from the MerR superfamily	No	Helix-turn-helix domain	No	Cytoplasmic 3.135	No
6989656	No	No	Putative phage tail protein	No	OuterMembrane 3.019	No
6989661	No	No	Phorbol esters/diacylglycerol binding domain (C1	No	Periplasmic 2.336 Cytoplasmic 2.163	No

			domain)				
6989662	No	No	BNR/Asp-box repeat	No	Cytoplasmic	2.512	No
6989664	No	No	No	No	Cytoplasmic	3.147	1z45A- 22 -37 -0.005
6989665	No	No	YnbE-like lipoprotein	No	Periplasmic Cytoplasmic	1.911 1.524	No
6989667	No	No	Bombolitin family	No	Cytoplasmic	2.432	No
6989674	No	No	Gyrovirus capsid protein (VP1)	No	Cytoplasmic	2.626	No
6989675	No	No	Orthopoxvirus protein of unkNown function (DUF830)	No	Cytoplasmic	3.312	2k1gA -17- 100- 7e-23
6989677	NADB_Rossmann super family[cl09931], Rossmann-fold NAD(P)(+)-binding proteins	No	Methyltransferase domain	2-polyprenyl-3-methyl-5-hydroxy-6-methoxy-1,4-benzoquinol methylase	Cytoplasmic	3.388	3cc8A- 16 -51- 2e-07
6989680	No	No	Hydrogenase expression/synthesis hypA family ; TFIIIB zinc-binding ;	No	Cytoplasmic	4.467	No
6989683	No	No	Homeodomain-like domain	No	Cytoplasmic	2.551	No
6989684	PTPc super family[cl00053], Protein tyrosine phosphatases	Dual specificity phosphatase, catalytic domain; Protein-tyrosine/Dual-specificity phosphatase; Protein-tyrosine phosphatase, active site	Dual specificity phosphatase, catalytic domain; Uroporphyrinogen decarboxylase (URO-D)	No	Cytoplasmic	3.681	1wrmA- 27- 36 0.002
6989687	gp6[cd08054], Head-Tail Connector Protein gp6 of Bacteriophage HK97 and similar proteins	Bacteriophage QLRG family, putative DNA packaging	Phage gp6-like head-tail connector protein	No	Cytoplasmic Periplasmic	2.183 1.672	No
6989689	No	No	No	Permeases of the major facilitator superfamily	Cytoplasmic Periplasmic	1.827 1.289	No
6989691	DUF551 super family[cl04523],	Domain of unkNown function DUF551	No	No	Cytoplasmic Periplasmic	2.099 2.082	No
6989694	DUF646 super family[cl12124],	Bacteriophage HK97, Gp10	Bacteriophage protein of unkNown function	No	Cytoplasmic	3.167	No

(DUF646)

6989695	No	No	Prokaryotic N-terminal methylation site; Short repeat of unknown function (DUF308)	No	Cytoplasmic	2.432	No
6989696	No	No	Putative transmembrane protein (PGPGW)	No	InnerMembrane	2.043	No
6989702	No	No	D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase	No	Cytoplasmic	2.480	No
6989704	DUF1643 super family[cl01787],	Bacteriophage D3, Orf41.6	No	No	Cytoplasmic Periplasmic	1.825 1.533	No

Table 6. *Pseudomonas* phage D3112

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
2648150	No	No	No	No	Cytoplasmic 2.480	2outA -38 -49 -2e-07
2648157	DUF2793 super family[cl12557]; wac[PHA02607], fibritin;	Fibritin/6-phosphogluconate dehydrogenase, C-terminal extension;	Fibronectin-binding repeat	No	Extracellular 2.656	No
2648161	DUF1320 super family[cl01818],	Bacteriophage Mu, Gp36	Type specific antigen		Cytoplasmic 2.528	No
2648165	nitrilase super family[cl11424],	No	No	No	OuterMembrane 1.861 Cytoplasmic 1.636	No
2648166	No	No	Heavy-metal resistance; Phasin protein	No	Periplasmic 2.161	No
2648171	Phage_BR0599 super family[cl10710],; phg_TIGR02218[TIGR02218], phage conserved hypothetical protein BR0599	Bacteriophage phiJL001, Gp84; Bacteriophage phiJL001, Gp84, C-terminal; Bacteriophage phiJL001, Gp84, N-terminal	Phage conserved hypothetical protein BR0599	No	Extracellular 2.617	No
2648173	No	Winged helix-turn-helix transcription repressor DNA-binding	Transcriptional regulator ; IclR helix-turn-helix domain		Periplasmic 1.648 Extracellular 1.270 OuterMembrane 1.246	1c3cA -15 -37 -0.002
2648177	DUF1018 super family[cl01815],	Bacteriophage Mu, Gp16	No	No	Cytoplasmic 2.302	No
2648181	No	No	TAFII55 protein conserved region ; Helix-turn-helix domain	No	Cytoplasmic 3.103	No
2648182	No	No	Helix-turn-helix	No	Periplasmic 2.314 Cytoplasmic 1.792	No
2648183	COG5362 super family[cl02216], Phage-related terminase [General function prediction only]	Archaeophage PsiM2, terminase large subunit	No	No	Cytoplasmic 4.355	2o0jA -16 -88- 4e-18
2648185	No	No	Type specific antigen	No	Cytoplasmic 2.851	No
2648194	No	No	No	No	InnerMembrane 2.170 Cytoplasmic 1.700	2iujA- 17- 31 -0.007

[2648195](#)

Phage-tail_3 super
family[cl16300],
Putative phage tail
protein

Putative phage tail
protein

OuterMembrane 3.322 2pziA- 15- 33- 0.002

Table 7. *Pseudomonas* phage LKD17

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5687413	PHA02046 super family[cl10354], hypothetical protein	No	No	No	Cytoplasmic 3.744	No
5687416	No	No	Herpesvirus pp38 phosphoprotein	No	Cytoplasmic 2.127	No
5687420	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily;	Ribonuclease H-like domain;	RNase_H superfamily ; Zinc ribbon domain	No	Cytoplasmic 4.053	1s5jA -18 -59 -2e-09
5687426	No	No	Spc7 kinetochore protein ; KR domain; CLIP, MHC2 interacting	No	Cytoplasmic 2.854	No
5687430	DNA_pol_A super family[cl02626]	No	No	No	Cytoplasmic 3.032	No
5687441	PolyA_pol super family[cl14657], Poly A polymerase head domain	No	No	No	Cytoplasmic 2.935	1ou5A- 41 -37- 0.002
5687443	ALDH-SF super family[cl11961], NAD(P)+-dependent aldehyde dehydrogenase superfamily	No	No	No	Cytoplasmic 2.309	No
5687444	No	No	No	No	Cytoplasmic 2.910	No
5687446	No	No	ABC-2 family transporter protein ; N-terminal TM domain of oligopeptide transport permease C; NfeD-like C-terminal, partner-binding ; EXS family	No	InnerMembrane 2.403	No
5687456	No	Acyl-CoA N-acyltransferase	Acetyltransferase (GNAT) domain	No	Cytoplasmic 3.912	2beiB 21- 42 -1e-04

Table 8. *Pseudomonas* phage MP22

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5648794	No	No	Fibritin C-terminal region ; Fibronectin-binding repeat	No	Extracellular 3.078	1zelB -29 -38 -0.003
5648795	No	No	C-terminal domain of 1-Cys peroxiredoxin	No	Periplasmic 2.064	No
5648798	No	No	No	No	InnerMembrane 2.382	1rrhA -23- 30 -0.009
5648803	No	No	No	No	Cytoplasmic 3.179	1u2mC- 20- 30- 0.010
5648804	nitrilase super family[cl11424], Nitrilase superfamily, including nitrile- or amide-hydrolyzing enzymes and amide-condensing enzymes	No	No	No	OuterMembrane 1.853 Extracellular 1.661	No
5648807	No	Bacteriophage Mu, Gp16	No	No	Cytoplasmic 2.330	No
5648811	Helix-turn-helix domains	Bacteriophage D3112, Orf24	ATP-grasp domain	No	Cytoplasmic 2.458	No
5648815	No	No	Putative sensor ; SICA extracellular beta domain; EcoEI R protein C-terminal	No	Periplasmic 2.029	No
5648821	No	No	Helix-turn-helix	No	Periplasmic 2.154 Cytoplasmic 1.979	No
5648825	No	No	No	No	Cytoplasmic 2.453 Periplasmic 1.809	
5648826	Phage_BR0599 super family[cl10710], Phage conserved hypothetical protein BR0599;DUF2163 super family[cl10805], Uncharacterized conserved protein (DUF2163); phg_TIGR02218[TIGR02218], phage conserved hypothetical protein BR0599;	Bacteriophage phiJL001, Gp84; Bacteriophage phiJL001, Gp84, C-terminal; Bacteriophage phiJL001, Gp84, N-terminal	No	No	Extracellular 2.177	No

5648827	Mu-like_gpT super family[cl01826], Mu-like prophage major head subunit gpT	Bacteriophage Mu, GpT	Mu-like prophage major head subunit gpT	No	Periplasmic	2.904	No
5648828	No	No	Phage major coat protein, Gp8	No	Cytoplasmic OuterMembrane	1.657 1.153	No
5648831	No	No	Type specific antigen	No	Cytoplasmic	3.166	No
5648833	n	No	Presenilin	No	Periplasmic Cytoplasmic Extracellular	1.633 1.561 1.256	No
5648834	No	Bacteriophage Mu, Gp36	Bacterial extracellular solute-binding protein, family 7	No	Cytoplasmic	4.138	No
5648836	Phage-tail_3 super family[cl16300], Putative phage tail protein	No	Putative phage tail protein	No	OuterMembrane	3.855	2pziA- 15- 33- 0.002
5648842	No	Winged helix-turn-helix transcription repressor DNA-binding	Transcriptional regulator ; IclR helix-turn-helix domain	No	Extracellular Cytoplasmic Periplasmic OuterMembrane	1.286 1.284 1.163 1.026	2o0yC -21- 38- 0.002

Table 9. *Pseudomonas* phage gh-1

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
2641707	No	No	PLC-beta C terminal ; Phage tail fibre repeat	No	Cytoplasmic 2.015 Periplasmic 1.582	No
2641710	PolyA_pol super family[cl14657], Poly A polymerase head domain	No	Poly A polymerase head domain ; Neuraxin and MAP1B repeat	No	Cytoplasmic 4.443	1miwA- 21 -151- 9e-38
2641711	No	Bacteriophage T7, Gp1.1	Bacteriophage 1.1 Protein ; Peptidase family C1 propeptide	No	Cytoplasmic 2.325	No

Table 10. *Pseudomonas* phage F116

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
3197172	No	No	Restriction endonuclease, FokI, C-terminal/endonuclease I, core	No	Periplasmic 2.842	1m0dA -28- 95- 6e-21
3197174	No	No	No	No	Cytoplasmic 2.217	1qhuA -21- 42- 2e-04
3197175	No	No	No	No	OuterMembrane 2.590	2pziA- 12 -35- 4e-04
3197178	No	No	Rad52/22 family double-strand break repair protein ; HSCB C-terminal oligomerisation domain ; Stage III sporulation protein AF (Spore_III_AF)	No	Cytoplasmic 2.463 Periplasmic 2.272	No
3197179	No	No	NADPH-dependent FMN reductase	No	Periplasmic 3.310	No
3197184	No	No	Sterile alpha motif (SAM)/Pointed domain	No	Cytoplasmic 4.361	No
3197186	No	No	3D domain	No	Cytoplasmic 2.828	No
3197187	RyR super family[cl03409], RyR domain; This domain is called RyR for RyaNodine receptor.	RyaNodine receptor Ryr	RyR domain	No	Periplasmic 2.898	No
3197188	No	Bacteriophage F116, Orf28	Transcription factor Tfb2	No	Cytoplasmic 3.243	No
3197192	No	No	No	L-lactate permease	Extracellular 3.483	No
3197194	No	No	Autophagy-related protein 11	No	Cytoplasmic 2.018 Periplasmic 1.439	No
3197201	No	No	Bacterial transcriptional repressor	No	Cytoplasmic 2.277 Periplasmic 1.643	No
3197207	No	No	Rio2, N-terminal	No	Cytoplasmic 4.397	No
3197210	No	No	No	Molecular chaperones (contain C-terminal Zn finger domain)	Cytoplasmic 2.833	No

3197215	No	No	No	No	Periplasmic Cytoplasmic	2.140 1.569	2i1kA- 19- 37- 0.005	
3197216	No	Listeria phage P100, Gp150	Transposase zinc- ribbon domain ; Restriction alleviation protein Lar	No	Cytoplasmic	2.572	No	
3197220	No	No	Neurotransmitter- gated ion-channel transmembrane region	No	Cytoplasmic	3.170	No	
3197224		PLDc_SF super family[cl15239], Catalytic domain of phospholipase D superfamily proteins;	No	Carboxypeptidase regulatory-like domain	No	Extracellular	2.602	No
3197226	No	No	tRNA_anti-like; Outer membrane protein (OmpH-like) ; Gas vesicle synthesis protein GvpL/GvpF; Bacteriophage lysis protein	No	Periplasmic OuterMembrane	2.205 1.606	No	
3197227	No	No	Mak16 protein C- terminal region	No	Cytoplasmic	2.491	No	
3197229	No	No	Uncharacterised ACR, COG2135	Uncharacteriz ed ACR	Cytoplasmic	3.471	2icuB- 41- 197 -1e-51	
3197235	No	No	Bacterial transcriptional activator domain	No	Periplasmic	2.038	2h9aA- 22- 34 -0.010	
3197240	No	No	Hydrogenase expression/synthesis hypA family ; Prokaryotic RING finger family 1 ; Putative zinc finger motif, C2HC5-type ; Double zinc ribbon	No	Cytoplasmic Periplasmic	1.967 1.455	No	
3197241	No	No	Double zinc ribbon ; zinc-finger	No	Periplasmic	2.545	No	

Table 11. *Pseudomonas* phage F10

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5228140	DUF2190 super family[cl02289], Uncharacterized conserved protein (DUF2190);	Bacteriophage VT1-Sakai, H0018	Uncharacterized conserved protein (DUF2190)	No	Periplasmic 1.556 Cytoplasmic 1.536	No
5228145	No	No	NADPH oxidase subunit p47Phox, C terminal domain	No	Periplasmic 2.340	No
5228148	No	No	Mu-like prophage FluMu protein gp41	No	Cytoplasmic 3.536	No
5228150	No	No	Bucentaur or craniofacial development	No	Cytoplasmic 1.971 Periplasmic 1.687	No
5228151	No	No	No	DNA-directed RNA polymerase sigma subunits (sigma70/sigma32)	Cytoplasmic 2.660	No
5228160	No	Arc-like DNA binding; Ribbon-helix-helix; Arc-type ribbon-helix-helix	Arc-like DNA binding domain	No	Cytoplasmic 3.238	1u9pA- 19- 43 2e-05
5228168	FimT[COG4970], Tfp pilus assembly protein FimT [Cell motility and secretion / Intracellular trafficking ...	Prokaryotic N-terminal methylation site;Prepilin-type cleavage/methylation, N-terminal;	Type IV pilin N-term methylation site GFxxxE ; Type II transport protein GspH	No	Extracellular 2.728	2hi2A -16 -74- 2e-14
5228173	No	No	Gyrovirus capsid protein (VP1) ; Haemolysin-type calcium binding protein related domain	No	Extracellular 3.349	No
5228176	No	No	6-pyruvoyl-tetrahydropterin synthase related domain; membrane protein;	No	Periplasmic 2.159	No
5228181	No	No	Putative nucleotidyltransferase substrate binding	No	Cytoplasmic 3.720	No

5228191	No	No	domain Prophage miNor tail protein Z (GPZ)	No	Cytoplasmic	1.893	No
-------------------------	----	----	--	----	-------------	-------	----

Table 12. *Pseudomonas* phage 73

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5228201	Phage_BR0599 super family[cl10710], Phage conserved hypothetical protein BR0599	Bacteriophage phiJL001, Gp84, C-terminal; Bacteriophage phiJL001, Gp84, N-terminal	Phage conserved hypothetical protein BR0599	No	Extracellular 3.580	No
5228204	No	No	HAMP domain	No	Cytoplasmic 1.741 Periplasmic 1.708	No
5228210	No	No	No	No	Cytoplasmic 1.895 OuterMembrane 1.426 InnerMembrane 1.232	2e6jA- 17- 32- 0.004
5228211	Phage-tail_3 super family[cl16300], Putative phage tail protein;	No	Putative phage tail protein	No	OuterMembrane 4.428	No
5228212	No	No	Phage related protein ; DJ-1/Pfpl family	No	Periplasmic 1.721 Cytoplasmic 1.526	No
5228221	No	No	Intra-flagellar transport protein 57 ; Z1 domain ; V-type ATPase 116kDa subunit family ; Hepatic lectin, N-terminal domain	No	Cytoplasmic 2.455 Periplasmic 1.780	2i1kA -19 -38 -0.002
5228222	No	No	HemY protein N-terminus ; HAMP domain ; Formate dehydrogenase N, transmembrane	No	Cytoplasmic 2.199	No
5228224	No	No	No	No	Periplasmic 3.397	2jesA -12- 32- 0.004
5228229	No	No	VRR-NUC domain	No	Periplasmic 4.023	No
5228231	No	No	Fusaric acid resistance protein family	No	Cytoplasmic 3.950	No
5228236	No	No	Double zinc ribbon; Sgf11 (transcriptional regulation protein) ; Zinc finger, C3HC4 type (RING finger)	No	Periplasmic 2.261	No
5228238	COG4983[COG4983], Uncharacterized conserved protein [Function unkNown]		Telomere regulation protein Stn1	No	Cytoplasmic 1.603 InnerMembrane 1.428 Periplasmic 1.091	1iq0A- 10- 32 -0.004

5228239	NLPC_P60 super family[cl11438], NlpC/P60 family;	Conserved hypothetical protein CHP02594	CHAP domain	No	Extracellular Cytoplasmic	1.829 1.397	2jrnA -21- 35- 0.009
5228240	No	No	Pedibin/Hym-346 family	No	Cytoplasmic	2.361	No
5228241	NTP-PPase super family[cl16941], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain superfamily	No	MazG nucleotide pyrophosphohydrolyase domain	No	Cytoplasmic	3.835	2yxhA- 30 -39- 4e-04
5228244	No	Streptococcus phage Sfi11, Gp151	Nucleoporin FG repeat region	No	Extracellular	2.174	No
5228245	No	No	ERF superfamily	No	Cytoplasmic	3.400	No
5228249	No	No	No	No	Periplasmic	3.502	1vpkA -12 -75 --4e-14

Table 13. *Pseudomonas* phage 119X

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5142606	No	No	No	No	Periplasmic 1.829	1lnzA- 19 -30- 0.009
-					InnerMembrane 1.434	
5142607	NTP-PPase super family[cl16941], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain superfamily; NTP-PPase super family[cl16941], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain superfamily;	No	MazG nucleotide pyrophosphohydrolase domain	No	Cytoplasmic 4.304	No
5142611	No	No	Fibronectin-binding protein (FBP)	No	Cytoplasmic 2.616	No
5142613	No	No	No	No	Cytoplasmic 2.917	2o0jA- 10 -35- 5e-04
5142615	No	No	7tm Chemosensory receptor	No	InnerMembrane 2.081	No
5142617	No	No	Saccharopepsin inhibitor I34	No	Cytoplasmic 3.101	No
5142619	n	No	Transcription factor Vhr1 ; Kinetoplastid membrane protein 11 ;	No	Cytoplasmic 3.401	No
5142620	n	No	Coiled-coil domain-containing protein (DUF2037)	No	Cytoplasmic 2.945	No
5142621	Cas4_I-A_I-B_I-C_I-D_II-B super family[cl00641], CRISPR/Cas system-associated protein Cas4; CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)	No	PD-(D/E)XK nuclease superfamily	No	Cytoplasmic 3.676	No

5142622	No	No	C4-dicarboxylate transporter/malic acid transport protein	No	No	InnerMembrane 1.588 Cytoplasmic 1.323 Periplasmic 1.148	No
5142624	No	Ribonuclease H-like domain	No	No	No	Cytoplasmic 3.478	No
5142630	No	No	Tripartite tricarboxylate transporter TctA family	No	No	Periplasmic 1.930 Cytoplasmic 1.242	No
5142634	No	No	AAA domain	No	No	Cytoplasmic 2.414 Periplasmic 1.753	No
5142637	No	No	RNA polymerases N / 8 kDa subunit; Spt4/RpoE2 zinc finger	No	No	Cytoplasmic 2.219	No
5142639	No	No	No	Dehydrogenases with different specificities (related to short-chain alcohol dehydrogenases)	No	Periplasmic 3.248	No
5142640	No	No	Primase C terminal 2 (PriCT-2)	No	No	Cytoplasmic 3.607	2coeA- 18- 34- 0.006
5142641	No	VanY super family[cl00813], D-alanyl-D-alanine carboxypeptidase;	D-alanyl-D-alanine carboxypeptidase;	No	No	Periplasmic 2.720	No
5142643	No	No	O-methyltransferase	No	No	Extracellular 2.551	No
5142646	No	No	Prokaryotic membrane lipoprotein lipid attachment site	No	No	Cytoplasmic 2.080	No
5142650	No	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;	AAA domain	ABC-type multidrug transport system, ATPase component	No	Cytoplasmic 1.888 Periplasmic 1.586	1dekA- 13 -82- 5e-17
5142653	No	No	No	Kef-type K ⁺ transport systems, membrane components	No	InnerMembrane 3.618	No
5142654	No	No	Oxidoreductase family, NAD-binding Rossmann fold	No	No	Periplasmic 3.072	No
5142656	No	HTH_XRE super family[cl15761], Helix-turn-helix	Helix-turn-helix domain	No	No	Cytoplasmic 2.535	No

XRE-family like
proteins

[5142657](#)

No

No

Septum formation
initiator ; Growth
arrest and DNA-
damage-inducible
proteins-interacting
protein 1

No

OuterMembrane 2.049
Periplasmic 1.616

No

Table 14. *Pseudomonas* phage PA11

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5236983	No	No	MEDS: MEthaNogen/methylotroph, DcmR Sensory domain; Major surface glycoprotein	No	Periplasmic 1.751 Cytoplasmic 1.385 Extracellular 1.191	No
5236984	No	No	PPP4R2	No	Cytoplasmic 3.329	No
5236985	LbetaH super family[cl00160], Left-handed parallel beta-Helix (LbetaH or LbH) domain; RimK[COG0189], Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl transferase)	ATP-grasp fold, subdomain 2	RimK-like ATP-grasp domain	S6 modification enzyme (glutaminyl transferase)	Cytoplasmic 3.648	1gsaA 19 57 4e-09
5236987	H3TH_53EXO[cd09898], H3TH domain of the 5'-3' exonuclease of Taq DNA polymerase I and homologs;PIN_SF super family[cl14812], PIN (PiIT N terminus) domain: Superfamily; PHA00439[PHA00439], exonuclease	5'-3' exonuclease, C-terminal domain	5'-3' exonuclease, N-terminal resolvase-like domain; 5'-3' exonuclease, C-terminal SAM fold	5'-3' exonuclease (including N-terminal domain of Poll)	Cytoplasmic 4.628	1bgxT- 18 -218 -8e-58
5236988	Sigma70_r2 super family[cl08419], Sigma-70 region 2;sigma70-ECF[TIGR02937], RNA polymerase sigma factor, sigma-70 family	RNA polymerase sigma factor 70, ECF, conserved site; RNA polymerase sigma-70 region 2; RNA polymerase sigma factor, region 2; RNA polymerase sigma-70;	Sigma-70 region 2	DNA-directed RNA polymerase specialized sigma subunits, sigma24 homologs	Cytoplasmic 4.426	1or7A -19 -170- 1e-43
5236989	Gn_AT_II[cd00352], Glutamine amidotransferases class-II (GATase).; GATase_6[pfam13522], Glutamine	Glutamine amidotransferase, type II	Glutamine amidotransferase domain	Glucosamine 6-phosphate synthetase, contains amidotransferase and	OuterMembrane 2.270	1ao0A -19- 286 -4e-78

	amidotransferase domain			phosphosugar isomerase domains			
5236991	Amidoligase_2 super family[cl13634], Putative amidoligase enzyme	Putative amidoligase enzyme	Putative amidoligase enzyme	No	Cytoplasmic	4.525	No
5236992	GAT_1 super family[cl00020], Type 1 glutamine amidotransferase (GATase1)-like domain; carA[CHL00197], carbamoyl-phosphate synthase arginine-specific small subunit	Peptidase C26; Glutamine amidotransferase type 1;	Peptidase C26	Predicted glutamine amidotransferases	Cytoplasmic	2.522	2a9vD -14 -121- 8e-29
5236993	No	No	No	No	Periplasmic	2.500	1egwB- 38 -31 -0.004
5236994	No	No	Copine ; EF-hand domain pair	No	Periplasmic OuterMembrane	1.957 1.490	No
5236998	COG5410 super family[cl02239], Uncharacterized protein conserved in bacteria	Terminase, large subunit	Terminase-like family	No	Cytoplasmic	3.650	2o0jA 12 37 1e-04
5236999	HNH_3 super family[cl16264], HNH endonuclease	No	HNH endonuclease ; Antitoxin SpoIIISB, type II toxin-antitoxin system ; Ta0938	No	Cytoplasmic Periplasmic	1.552 1.532	1u3eM- 19- 38- 0.001
5237001	PLN03003 super family[cl14103], Probable polygalacturonase At3g15720	Pectin lyase fold/virulence factor; Pectin lyase fold;	Pectate lyase superfamily protein; N terminal extension of bacteriophage endosialidase	No	Extracellular	3.593	2vjiA -12 -43- 2e-04
5237002	Description PssmId Multi-dom E-value; HNH_3 super family[cl16264], HNH endonuclease	No	HNH endonuclease; Beta-2-glycoprotein-1 fifth domain; FANCL C-terminal domain; Homeodomain-like domain	No	Cytoplasmic	3.427	1u3eM -15- 61- 1e-10
5237009	COOH-NH2_lig super family[cl16852], Phage phiEco32-like COOH.NH2 ligase-type 2;	No	Phage phiEco32-like COOH.NH2 ligase-type 2	No	Periplasmic	3.013	No

5237010	HNHc super family[cl00083], HNH nucleases;	No	HNH endonuclease	No	Cytoplasmic	3.228	2e8yA 30 38 5e-04
5237012	No	No	No	No	Extracellular	3.659	1t71A- 27- 37 -0.002
5237016	Head-tail_con super family[cl10251], Bacteriophage head to tail connecting protein	Head-to-tail joining protein, podovirus-type	Bacteriophage head to tail connecting protein	No	Cytoplasmic	3.425	No
5237018	No	No	No	No	Cytoplasmic	4.293	2fteA -9- 37- 0.008
5237021	No	Bacteriophage B103, Gp8, head fibre	Head fiber protein	No	Cytoplasmic	2.142	No
5237022	No	No	Translation initiation factor eIF3 subunit;	No	Extracellular	2.361	No
5237023	RNR_PFL super family[cl09939], Ribonucleotide reductase and Pyruvate formate lyase;	Ribonucleotide reductase large subunit, C-terminal; Ribonucleoside-diphosphate reductase alpha chain; Ribonucleotide reductase large subunit, N-terminal;	Ribonucleotide reductase, all-alpha domain ; Ribonucleotide reductase, barrel domain; Ribonucleotide reductase, barrel domain ; Ribonucleotide reductase, barrel domain ; Infectious salmon anaemia virus haemagglutinin	Ribonucleotide reductase alpha subunit	Cytoplasmic	2.656	1zyzA- 23- 345- 9e-98
5237024	RNRR2[cd01049], Ribonucleotide Reductase, R2/beta subunit, ferritin-like diiron-binding domain	Ribonucleotide reductase; Ferritin/ribonucleotide reductase-like; Ribonucleotide reductase-related;	Ribonucleotide reductase, small chain	Ribonucleotide reductase beta subunit	Cytoplasmic	4.752	1mxrA- 20- 247- 1e-66
5237027	No	No	Nucleopolyhedrovirus P10 protein; MutS domain III; Phage miNor structural protein GP20; Type I restriction and modification enzyme - subunit R C terminal ; Cdc37 N terminal kinase binding ; Erythromycin esterase; EcoEI R protein C-terminal; Enterocin A Immunity ; Coiled-coil domain-containing protein	No	Periplasmic	2.081	No

(DUF2037)

5237028	No	No	Protein involved in formate dehydrogenase formation ; Restriction alleviation protein Lar ; Probable zinc-binding domain ; Ogr/Delta-like zinc finger	Transcription initiation factor IIB	Cytoplasmic	2.047	No
5237030	No	No	Polysaccharide deacetylase	No	Periplasmic Cytoplasmic	2.001 1.791	No
5237031	DNA_pol_A super family[cl02626], Family A polymerase primarily fills DNA gaps that arise during DNA repair, recombination and replication; DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily;	DNA-directed DNA polymerase, family A, palm domain; Ribonuclease H-like domain;	DNA polymerase family A ; RNase_H superfamily	DNA polymerase I - 3'-5' exonuclease and polymerase domains	Cytoplasmic	4.137	1x9mA- 18 -384- 1e-107
5237033	Dcm[COG0270], Site-specific DNA methylase [DNA replication, recombination, and repair]	No	No	No	Cytoplasmic	3.635	1dctA- 18 -40- 4e-04
5237037	TRX_family[cd02947], TRX family; composed of two groups:	Thioredoxin; Thioredoxin-like fold; Thioredoxin domain;	Thioredoxin ; Type-F conjugative transfer system pilin assembly protein	Thiol-disulfide isomerase and thioredoxins	Cytoplasmic	3.568	2trxA -27- 128- 3e-31
5237038	No	No	MamL-1 domain; Spc7 kinetochore protein; TATA element modulatory factor 1 TATA binding	No	Cytoplasmic	3.661	3cnfT- 17- 42- 9e-05
5237039	No	No	PKD domain	No	Extracellular	4.464	No
5237041	No	Bacteriophage T7, Gp1.7	Cobalt chelatase (CbiK)	No	Cytoplasmic Periplasmic	1.899 1.467	No
5237042	No	No	No	Aspartyl-tRNA synthetase	OuterMembrane	3.697	

5237044	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily	Ribonuclease H-like domain	RNase_H superfamily; Adenylate kinase, active site lid ; Helix-turn-helix domain of transposase family ISL3 ; Transposase zinc-ribbon domain ; Predicted integral membrane metal-binding protein (DUF2296) ; Restriction alleviation protein Lar	DNA polymerase elongation subunit (family B)	Cytoplasmic	3.754	2jguA- 17 -48- 3e-06
5237045	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;TOPRI M_primases[cd01029], TOPRIM_primases	Toprim domain; DNA helicase, DnaB-like, C-terminal;	Toprim-like; Zinc-binding domain of primase-helicase ; AAA domain	Replicative DNA helicase	Cytoplasmic	4.738	1q57A -24 -348- 1e-96
5237046	endolysin_autolysin[cd00737]	Glycoside hydrolase, family 24; Lysozyme-like domain;Lysozyme domain	Phage lysozyme	No	Cytoplasmic	3.524	2anvB- 40 -157- 5e-40
5237047	No	No	Glycosyl hydrolase family 46; HNH endonuclease	No	Periplasmic	2.090	No
5237048	No	No	Tropomyosin like; TPR/MLP1/MLP2-like protein ; MbeD/MobD like	Pyruvate kinase	Cytoplasmic	2.638	1phpA- 31- 34- 0.005
5237049	hy1 super family[cl03630], Thymidylate synthase complementing protein	Thymidylate synthase ThyX	Thymidylate synthase complementing protein	Predicted alternative thymidylate synthase	Cytoplasmic	3.934	1o26B -33 -228 -6e-61

Table 15. *Pseudomonas* phage M6

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5237052	Radical_SAM super family[cl15377], Radical SAM superfamily; Radical_SAM[pfam04055], Radical SAM superfamily	Aldolase-type TIM barrel	Radical SAM superfamily	DNA repair photolyase	Cytoplasmic 2.847	1tv8A- 12 -82- 1e-16
5237055	No	No	Lysozyme-like	No	Cytoplasmic 1.612 OuterMembrane 1.100 Periplasmic 1.084	No
5237057	Phage_lysis super family[cl14955], Bacteriophage lysis protein	Bacteriophage lysis endopeptidase, Rz-type; Bacteriophage lysis endopeptidase, predicted	Bacteriophage lysis protein ; Phosphonate metabolism protein PhnG ; Outer membrane protein (OmpH-like) ; gp58-like protein	No	Periplasmic 1.847 OuterMembrane 1.327	2i1kA- 17 -36 -0.003
5237059	cytidine_deaminase-like super family[cl00269], Cytidine and deoxycytidylate deaminase zinc-binding region; SGNH_hydrolase super family[cl01053], SGNH_hydrolase, or GDSL_hydrolase, is a diverse family of lipases and esterases	CMP/dCMP deaminase, zinc-binding; Cytidine deaminase; Cytidine deaminase-like;	Cytidine and deoxycytidylate deaminase zinc-binding region	Deoxycytidylate deaminase	Cytoplasmic 2.325	2hvwA -24 -131 -5e-32
5237060	No	Coagulation factor 5/8 C-terminal type domain; Galactose-binding domain-like;	F5/8 type C domain	No	Extracellular 3.627	2v72A- 14 -34 -6e-04
5237061	No	No	DNA helicase IV / RNA helicase N terminal	No	Cytoplasmic 3.225	No
5237065	PRK02224[PRK0224], chromosome segregation protein	No	No	No	Cytoplasmic 3.994	1i84S -24 -38 -0.003

5237068	ape_meas_nterm super family[cl15680], tape measure domain; COG3941[COG3941], Mu-like prophage protein; BexC_CtrB_KpsE[FIGR01010], polysaccharide export inner-membrane protein, BexC/CtrB/KpsE family	Caudovirus, tape measure, N-terminal;	CheC-like family	Methyl-accepting chemotaxis protein	InnerMembrane OuterMembrane	2.168 1.728	No
5237069	No		MethaNol dehydrogenase beta subunit	No	Extracellular	2.776	No
5237070	No	No	Uracil DNA glycosylase superfamily	No	Periplasmic	2.670	1vk2A- 17 -38- 0.001
5237073	PRK13962[PRK13962], bifunctional phosphoglycerate kinase/triosephosphate isomerase	No	No	No	Cytoplasmic	2.505	2zc0A- 28 -34 -0.008
5237074	No	No	Poly(hydroxyalcalonate) granule associated protein (phasin)	No	Cytoplasmic	2.913	No
5237076	No	No	Bacteriophage related domain of unknown function	No	Cytoplasmic	1.895	No
5237077	PRK13539[PRK13539], cytochrome c biogenesis protein CcmA;	No	AAA domain	ABC-type cobalamin/Fe ³⁺ -siderophores transport systems, ATPase components	Cytoplasmic	3.493	1p5zB -18 -44- 3e-05
5237078	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily;	No	No	No	Cytoplasmic	2.689	No
5237079	No	No	Ribonulcease H	No	Cytoplasmic Periplasmic	2.235 1.910	1j9aA- 17- 35 -0.008
5237083	S1_like super family[cl09927], S1_like: Ribosomal protein S1-like	No	Double zinc ribbon; zinc-finger; S1 domain	No	Cytoplasmic	3.488	No

	RNA-binding domain							
5237084	RNR_II_dimer[cd02888], Class II ribonucleotide reductase, dimeric form; NrdJ_Z[TIGR02504], ribonucleoside-diphosphate reductase, adenocobalamin-dependent;	Ribonucleotide reductase large subunit, C-terminal; Ribonucleoside-diphosphate reductase, adenocobalamin-dependent	Ribonucleotide reductase, barrel domain	Ribonucleotide reductase alpha subunit	Cytoplasmic Periplasmic	2.541 1.974	1zyzA -19 -589 -1e-169	
5237086	HDc super family[cl00076], Metal dependent phosphohydrolases with conserved 'HD' motif	HD domain	HD domain	No	Cytoplasmic	4.611	2gz4A- 21- 124- 1e-29	
5237090	Phage-tail_3 super family[cl16300], Putative phage tail protein	No	Putative phage tail protein	No	OuterMembrane	3.837	1oh4A- 30 -31- 0.008	
5237091	NTP-PPase super family[cl16941], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain superfamily	No	HD domain; MazG nucleotide pyrophosphohydrolase domain	No	Cytoplasmic	3.149	2oieB -23 -55 -2e-08	
5237094	Trp-synth-beta_II super family[cl00342], Tryptophan synthase beta superfamily (fold type II)	Pyridoxal phosphate-dependent enzyme, beta subunit	Pyridoxal-phosphate dependent enzyme ; Phytotoxin PcF protein	No	Periplasmic	2.840	1j0aA -15 -58 -3e-09	
5237096	No	No	Apolipoprotein A1/A4/E domain ; Putative transcription regulator (DUF1323) ; Viral A-type inclusion protein repeat	No	Cytoplasmic	3.996	No	
5237098	HepA[COG0553], Superfamily II DNA/RNA helicases, SNF2 family [Transcription / DNA replication,	SNF2-related; Helicase, C-terminal;	SNF2 family N-terminal domain ; Helicase conserved C-terminal domain; ParB-like nuclease domain	Superfamily II DNA/RNA helicases, SNF2 family	Cytoplasmic	4.787	1z6aA- 20- 172 -2e-45	
5237100	TS_Pyrimidine_Hase super family[cl00358], Thymidylate	Thymidylate synthase; Thymidylate synthase/dCMP	Thymidylate synthase	Thymidylate synthase	Cytoplasmic	2.955	2h2qA- 20- 264 -1e-71	

5237103	synthase and pyrimidine hydroxymethylase : DNA_pol_A super family[cl02626], Family A polymerase primarily fills DNA gaps that arise during DNA repair, recombination and replication; DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily; PolA[COG0749], DNA polymerase I - 3'-5' exonuclease and polymerase domains	hydroxymethylase domain DNA-directed DNA polymerase, family A, palm domain; DNA polymerase A; Ribonuclease H-like domain; DNA-directed DNA polymerase, family A, conserved site	DNA polymerase family A	DNA polymerase I - 3'-5' exonuclease and polymerase domains	Cytoplasmic	4.216	2kfnA 24 636 0.0
5237104	No	No	No	No	Cytoplasmic	4.474	1gk9B- 18 -37 -0.002
5237106	Phage_BR0599 super family[cl10710], Phage conserved hypothetical protein BR0599;; DUF2163 super family[cl10805], Uncharacterized conserved protein (DUF2163);phg_TI GR02218[TIGR02218], phage conserved hypothetical protein BR0599	Bacteriophage phiJL001, Gp84; Bacteriophage phiJL001, Gp84, C-terminal; Bacteriophage phiJL001, Gp84, N-terminal	Uncharacterized conserved protein (DUF2163); Phage conserved hypothetical protein BR0599	No	Extracellular	2.399	No
5237107	UvrD_C_2 super family[cl15862], Family description; P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;	No	AAA domain ; Family description	No	Cytoplasmic	3.574	1w36D- 17 -140- 8e-34
5237108	No	Bacteriophage T7, Gp1.7	No	ATP-dependent exoDNase (exonuclease)	Cytoplasmic	4.318	No

5237112	HTH_XRE super family[cl15761], Helix-turn-helix XRE-family like proteins; HTH_XRE super family[cl15761], Helix-turn-helix XRE-family like proteins	Lambda repressor-like, DNA-binding	Helix-turn-helix ; Helix-turn-helix domain	No	Cytoplasmic	3.653	No		V), alpha subunit - helicase superfamily I member
5237113	primase_Cterm super family[cl15909], phage/plasmid primase, P4 family, C-terminal domain	No	Virulence-associated protein E ; Poxvirus D5 protein-like	No	Cytoplasmic	3.710	No		
5237114	No	No	No	No	Cytoplasmic	2.855	No		
5237115	Terminase_6[pfa m03237], Terminase-like family; This family represents a group of terminase proteins	Terminase, large subunit	Terminase-like family	No	Cytoplasmic	3.046	2o0JA 12 35 3e-04		
5237116	No	No	Prolyl-tRNA synthetase, C-terminal	No	Cytoplasmic	3.087	No		
5237117	No	No	Apolipoprotein A1/A4/E domain	No	Periplasmic	2.888	No		
5237118	Antirestrict super family[cl03947], Antirestriction protein	Antirestriction protein	Antirestriction protein	No	Cytoplasmic	3.708	No		
5237120	No	No	No	No	Extracellular	2.800	2fteA- 18- 46- 1e-05		
5237124	No	No	DnaJ central domain	No	Cytoplasmic	3.288	No		
5237126	No	Zinc finger, CCHC-type	Predicted membrane protein (DUF2318); Zinc knuckle	No	Cytoplasmic	3.809	No		
5237127	Phage_Mu_F super family[cl10072], Phage Mu protein F like protein	Phage head morphogenesis domain; Bacteriophage K1H, Orf3	Phage Mu protein F like protein	No	Cytoplasmic	3.086	No		

5237128	Predicted soluble lytic transglycosylase fused to an ABC-type amino acid-binding protein [Cell envelope biogenesis, outer membrane]	Lysozyme-like domain	No	Soluble lytic murein transglycosylase and related regulatory proteins (some contain LysM/invasin domains)	Periplasmic OuterMembrane	1.804 1.329	1qsaA- 9 -49- 8e-07
5237130	No	No	Kelch motif	No	Cytoplasmic	4.481	No
5237131	No	No	Phage portal protein	No	Cytoplasmic	2.321	No
5237133	No	No	Plastid and cyanobacterial ribosomal protein (PSRP-3 / Ycf65)	No	Cytoplasmic	2.574	No
5237134	No	No	No	Carbamoylphosphate synthase large subunit (split gene in MJ)	InnerMembrane	4.155	No
5237135	No	No	No	No	OuterMembrane	2.523	1zunB 21 37 0.009
5237136	No	No	Conserved hypothetical protein 2217 (DUF2460)	No	Cytoplasmic	2.995	No

Table 16. *Pseudomonas* phage F8

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5237144	No	No	SigmaW regulon antibacterial ; GAD domain; Ribosomal protein S27a	No	Cytoplasmic 3.047	No
5237147	No	No	S-layer like family, N-terminal region	No	Cytoplasmic 2.295	No
5237148	No	No	Aspartyl protease	No	Cytoplasmic 3.467	No
5237149	DUF2213 super family[cl15403], Uncharacterized protein conserved in bacteria (DUF2213);	Uncharacterised conserved protein UCP029215	Uncharacterized protein conserved in bacteria (DUF2213)	Uncharacterized BCR	Cytoplasmic 3.501	No
5237150	No	No	Flagellar P-ring protein	No	Periplasmic 2.477	No
5237152	No	No	Stress responsive A/B Barrel Domain	No	Cytoplasmic 3.675	No
5237154	No	No	Double zinc ribbon	No	Cytoplasmic 2.081 Periplasmic 1.702	No
5237158	No	No	Alphaherpesvirus glycoprotein E	No	InnerMembrane 3.820	No
5237165	No	No	HrpJ-like domain; Bacteriophage protein of unkNown function (DUF646)	No	Cytoplasmic 3.391	2qzvA 15 37 0.003
5237168	VirE super family[cl14928], Virulence-associated protein E	Virulence-associated E	Virulence-associated protein E ; Phage Mu protein F like protein; Dodecin	No	Cytoplasmic 4.461	2v9pD -18 -34 0.001
5237169	No	No	Ca2+ regulator and membrane fusion protein Fig1 ;	No	InnerMembrane 1.569 Cytoplasmic 1.326	No
5237172	SMC_prok_B[TIGR 02168], chromosome segregation protein SMC, common bacterial type	Heat shock protein DnaJ, cysteine-rich domain	DnaJ central domain; RIM-binding protein of the cytomatrix active zone	No	Cytoplasmic 2.478 Periplasmic 2.260	2f9yB- 30-43- 8e-05
5237174	No	No	Eukaryotic and archaeal DNA primase small subunit	No	Cytoplasmic 2.392	No

5237175	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;UvrD_C_2 super family[cl15862], Family description	No	AAA domain; Family description	ATP-dependent exoDNAse (exonuclease V), alpha subunit - helicase superfamily I member	Cytoplasmic	4.781	1w36D- 17 -174 2e-44
5237176	No	No	Daxx Family	No	Cytoplasmic	2.883	No
5237177	No	No	No	FAD binding domain	Cytoplasmic	3.652	2b9wA -40 -38 -0.005
5237184	No	No	Nitrile hydratase beta subunit	No	Cytoplasmic	3.434	No
5237185	Thy1 super family[cl03630], Thymidylate synthase complementing protein	Thymidylate synthase ThyX	Thymidylate synthase complementing protein	Predicted alternative thymidylate synthase	Cytoplasmic	3.806	2af6G -21- 213- 2e-56
5237186	Terminase_3 super family[cl12054],	Bacteriophage terminase, large subunit;Caudovirales, terminase large subunit	Phage terminase large subunit	Phage terminase large subunit	Cytoplasmic	4.137	2o0jA- 12- 55 -5e-08
5237187	COG3567 super family[cl12082], Uncharacterized protein conserved in bacteria ;	No	Inorganic pyrophosphatase	Uncharacterized BCR	Periplasmic	3.741	No
5237190	pseT[PHA02530], polynucleotide kinase; Provisional	HAD-like domain	HAD superfamily, subfamily IIIB (Acid phosphatase); Ketopantoate reductase PanE/ApbA C terminal	No	Cytoplasmic	3.223	1ltqA- 21- 163- 3e-41
5237192	DEDDh[cd06127], DEDDh 3'-5' exonuclease domain family;	Exonuclease; Ribonuclease H-like domain; Exonuclease, RNase T/DNA polymerase III;	Exonuclease	DNA polymerase III epsilon subunit and related 3'-5' exonucleases	Cytoplasmic	4.625	2p1jB -27- 179 -2e-46
5237193	DNA_pol3_alpha super family[cl06665], Bacterial DNA polymerase III ; PRK07373[PRK07373], DNA polymerase III subunit	Bacterial DNA polymerase III, alpha subunit	Bacterial DNA polymerase III alpha subunit ; PHP domain	DNA polymerase III alpha subunit	Cytoplasmic	3.540	No

5237195	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; HepA[COG0553], Superfamily II DNA/RNA helicases, SNF2 family	SNF2-related; Helicase, C-terminal;	SNF2 family N-terminal domain ; Helicase conserved C-terminal domain	Superfamily II DNA/RNA helicases, SNF2 family	Cytoplasmic	4.521	1z6aA- 19-481 -1e-136
5237197	No	No	Vitamin-D-receptor interacting Mediator subunit 4 ; TSC-22/dip/bun family	No	Cytoplasmic	4.129	No
5237199	No	No	Cytochrome b5-like Heme/Steroid binding domain	No	Periplasmic	3.390	No
5237200	No	No	Outer membrane protein (OmpH-like); TipAS antibiotic-recognition domain ;	No	Periplasmic InnerMembrane OuterMembrane	1.564 1.286 1.030	No
5237201	NTP-PPase_u3[cd11540], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8)	Phosphoribosyl-ATP pyrophosphohydrolase-like	Phosphoribosyl-ATP pyrophosphohydrolase	No	Cytoplasmic	4.729	1vmgA- 21 -51 -2e-07
5237202	Adenylation_DNA_ligase_like super family[cl12015]; OBF_DNA_ligase_family super family[cl08424],	DNA ligase, ATP-dependent, central; Nucleic acid-binding, OB-fold; Nucleic acid-binding, OB-fold-like	ATP dependent DNA ligase domain	No	Cytoplasmic	4.692	1fviA -25- 108 -3e-24
5237203	No	No	V-ATPase subunit H	No	Cytoplasmic	2.973	No
5237204	lysozyme_like super family[cl00222], lysozyme_like domain	Glycoside hydrolase, family 19, catalytic; Lysozyme-like domain;	Chitinase class I	Predicted chitinase	Periplasmic Extracellular	2.103 1.683	1wvva 17 135 4e-33
5237206	No	No	Bacterial protein of unkNown function (DUF905)	No	Extracellular	3.585	No
5237207	No	No	No	No	Extracellular	2.088	1qzfA- 17- 31- 0.007

5237208	Phage_Mu_F super family[cl10072], Phage Mu protein F like protein;	Phage head morphogenesis domain	Phage Mu protein F like protein	Uncharacterized protein, homolog of phage Mu protein gp30	Cytoplasmic	2.383	No
5237210	No	Uncharacterized protein conserved in bacteria (DUF2345);Bacteriophage Mu Gp45 protein	No	No	Periplasmic Extracellular	1.769 1.391	No
5237214	LT_GEWL[cd00254], Lytic Transglycosylase (LT) and Goose Egg White Lysozyme (GEWL) domain	Prokaryotic transglycosylase, active site; Lytic transglycosylase-like, catalytic; Lysozyme-like domain	Transglycosylase SLT domain ; Colicin pore forming domain	Soluble lytic murein transglycosylase and related regulatory proteins (some contain LysM/invasin domains)	Periplasmic Cytoplasmic	1.467 1.205	No
5237217	No	No	Mycobacterium membrane protein ; Tetratricopeptide repeat	No	Cytoplasmic	2.194	No
5237218	No	No	No	No	Extracellular OuterMembrane	1.583 1.582	1n8yC- 20 -36- 0.006
5237220	No	No	Met-10+ like-protein	No	Cytoplasmic	3.578	No
5237222	No	No	No	No	Periplasmic Cytoplasmic Extracellular	1.496 1.483 1.141	1a6zA- 19 -37 -0.002
5237223	No	Bacteriophage rv5, Orf53	No	No	Extracellular	2.847	No
5237224	No	No	Bacteriophage protein of unkNown function (DUF646)	No	Cytoplasmic InnerMembrane	2.120 1.518	No
5237228	No	No	Beta propeller domain	No	Periplasmic Cytoplasmic Extracellular	1.418 1.261 1.257	No

Table 17. *Pseudomonas* phage Pf3

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
1260909	Helitron_like_N super family[cl16715], Helitron helicase- like domain at N- terminus	No	Helitron helicase-like domain at N-terminus	No	Cytoplasmic 2.869	No
1260910	P-loop_NTPase super family[cl09099], P- loop containing Nucleoside Triphosphate Hydrolases	Zona occludens toxin	Zonular occludens toxin (Zot)	No	Periplasmic 1.927 Cytoplasmic 1.539	2r2aA- 29- 170 -3e-43

Table 18. *Pseudomonas* phage MP29

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
7056597	Aconitase super family[cl00285], Aconitase catalytic domain	No	No	3-isopropylmalate dehydratase large subunit	Cytoplasmic 3.180	No
7056600	No	No	Antitoxin Phd_YefM, type II toxin-antitoxin system	No	Cytoplasmic 3.167	No
7056604	No	Winged helix-turn-helix transcription repressor DNA-binding	Transcriptional regulator; MarR family	No	Extracellular 1.406 Periplasmic 1.244 Cytoplasmic 1.138 OuterMembrane 1.004	2o0yC -21- 37- 0.004
7056606	No	No	Helix-turn-helix	No	Periplasmic 2.290 Cytoplasmic 1.866	No
7056607	No	No	Type specific antigen	No	Cytoplasmic 3.166	No
7056613	No	Bacteriophage Mu, Gp16	No	No	Cytoplasmic 2.523	No
7056615	No	No	C-terminal domain of 1-Cys peroxiredoxin	No	Periplasmic 1.840 Cytoplasmic 1.272	No
7056618	No	No	Putative sensor ; SICA extracellular beta domain; EcoEI R protein C-terminal	No	Periplasmic 1.965	No
7056620	HTH super family[cl00088], Helix-turn-helix domains;	Bacteriophage D3112, Orf24	Type II secretion system (T2SS), protein F	No	Cytoplasmic 2.310	No
7056621	No	No	No	No	Cytoplasmic 2.980	1u2mC-20 -30- 0.008
7056629	Mu-like_gpT super family[cl01826], Mu-like prophage major head subunit gpT; Mu-like_gpT super family[cl01826], Mu-like prophage major head subunit gpT; Mu-like_gpT super family[cl01826], Mu-like prophage major head subunit gpT	Bacteriophage Mu, GpT	Mu-like prophage major head subunit gpT ; Mu-like prophage major head subunit gpT ; Mu-like prophage major head subunit gpT	No	Periplasmic 3.383	No

7056630	odpA[CHL00149], pyruvate dehydrogenase E1 component alpha subunit;	No	No	Topoisomeras e IA	Cytoplasmic Periplasmic	2.203 1.940	No
7056632	No	Bacteriophage Mu, Gp36	Type specific antigen	No	Cytoplasmic	2.528	No
7056638	nitrilase super family[cl11424], Nitrilase superfamily, including nitrile- or amide- hydrolyzing enzymes and amide-condensing enzymes	No	No	No	Extracellular OuterMembrane	1.824 1.709	No
7056640	No	No	No	No	InnerMembrane	2.347	1rrhA -23 -30- 0.009
7056641	Phage_BR0599 super family[cl10710], Phage conserved hypothetical protein BR0599; DUF2163 super family[cl10805], Uncharacterized conserved protein (DUF2163); phg_TIGR02218[TI GR02218], phage conserved hypothetical protein BR0599;	Bacteriophage phiJL001, Gp84;Bacterioph ge phiJL001, Gp84, C-terminal; Bacteriophage phiJL001, Gp84, N- terminal;	Uncharacterized conserved protein (DUF2163) ; Phage conserved hypothetical protein BR0599	No	Extracellular	2.828	No
7056643	Phage-tail_3 super family[cl16300], Putative phage tail protein;	No	Putative phage tail protein	No	OuterMembrane	3.619	2pziA 15 33 0.001
7056644	No	No	No	No	Extracellular	3.192	1zelB 29 38 0.003

Table 19. *Pseudomonas* phage MP38

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
7056546	No	Winged helix-turn-helix transcription repressor DNA-binding Method Identifier Description	Transcriptional regulator ; MarR family	No	Cytoplasmic 1.549	2g7uA -17- 36 -0.008
7056549	No	No	Helix-turn-helix	No	Periplasmic 2.290 Cytoplasmic 1.866	No
7056552	No	No	Type specific antigen	No	Cytoplasmic 2.579	No
7056553	No	No	UBA/TS-N domain	No	Cytoplasmic 2.836	No
7056556	No	No	Bacteriophage Mu, Gp16	No	Cytoplasmic 2.523	No
7056558	No	No	C-terminal domain of 1-Cys peroxiredoxin	No	Periplasmic 2.010 Cytoplasmic 1.737 Cytoplasmic 1.361	No
7056562	No	No	Putative sensor ; SICA extracellular beta domain ; EcoEI R protein C-terminal	No	Periplasmic 2.029	No
7056563	HTH super family[cl00088], Helix-turn-helix domains;	Bacteriophage D3112, Orf24	Retrograde transport protein Dsl1 N terminal ; Type II secretion system (T2SS), protein F	No	Cytoplasmic 2.400	No
7056565	No	No	Human Cytomegalovirus UL139 protein	No	Cytoplasmic 2.872	No
7056566	COG5362 super family[cl02216], Phage-related terminase [General function prediction only]	Archaeophage PsiM2, terminase large subunit	No	No	Cytoplasmic 4.468	1sxjC- 10 -38- 0.003
7056571	Mu-like_gpT super family[cl01826], Mu-like prophage major head subunit gpT	Bacteriophage Mu, GpT	Mu-like prophage major head subunit gpT	No	Periplasmic 3.018	No
7056573	No	No	Tumour suppressor, Mitostatin ; CP12 domain	No	Cytoplasmic 2.971	No
7056574	No	Bacteriophage Mu, Gp36	Bacterial extracellular solute-binding protein,	No	Cytoplasmic 4.138	No

			family 7				
7056575	No	No	Phage major coat protein, Gp8	No	Cytoplasmic	1.851	No
7056582	No	No	No	No	InnerMembrane	2.363	1giyQ- 20 -30 -0.010
7056583	Phage_BR0599 super family[cl10710], Phage conserved hypothetical protein BR0599;DUF2163 super family[cl10805], Uncharacterized conserved protein (DUF2163); phg_TIGR02218[TIGR02218], phage conserved hypothetical protein BR0599	Bacteriophage phiJL001, Gp84; Bacteriophage phiJL001, Gp84, C-terminal; Bacteriophage phiJL001, Gp84, N-terminal	Uncharacterized conserved protein (DUF2163); Phage conserved hypothetical protein BR0599	No	Extracellular	2.620	No
7056585	Aconitase super family[cl00285], Aconitase catalytic domain;	No	3-isopropylmalate dehydratase large subunit	No	Cytoplasmic	3.180	No
7056586	Phage-tail_3 super family[cl16300], Putative phage tail protein	No	Putative phage tail protein	No	OuterMembrane	3.467	2pziA -15- 33 -0.001
7056587	No	No	Fibritin C-terminal region	No	Extracellular	3.558	1zelB-31- 38- 0.003
7056588	No	No	Bacillus PapR protein	No	Cytoplasmic	2.990	No
7056593	No	No	Helix-turn-helix domain	No	Cytoplasmic	2.999	No

Table 20. *Pseudomonas* phage phikF77

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
7670361	No	No	Antenna complex alpha/beta subunit ; Photosystem II reaction centre M protein (PsbM)	No	Cytoplasmic 2.172	No
7670369	Spc7[pfam08317], Spc7 kinetochore protein;		Glycoprotein GP40 of Cryptosporidium; SF-assemblin/beta giardin ; Spc7 kinetochore protein ; Mnd1 family ; Phage late control gene D protein (GPD); BCL7, N-terminal conserved region ; CLIP, MHC2 interacting	No	Cytoplasmic 2.909	No
7670370	DNA_pol_A super family[cl02626], Family A polymerase primarily fills DNA gaps that arise during DNA repair, recombination and replication	No	DNA polymerase family A ; Divergent polysaccharide deacetylase ; Mediator complex subunit 13 N-terminal	DNA polymerase I - 3'-5' exonuclease and polymerase domains	Cytoplasmic 2.021 Periplasmic 1.545	1bgxT- 47- 43 -7e-05
7670372	Rel-Spo_like super family[cl11966], RelA- and SpoT-like ppGpp Synthetases and Hydrolases, catalytic domain	No	Poly A polymerase head domain	tRNA nucleotidyltransferase/poly(A) polymerase	Periplasmic 2.488	1miwA- 25- 38- 0.003
7670378	No	No	No	Helicase subunit of the DNA excision repair complex	Cytoplasmic 3.188	No
7670384	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily	Ribonuclease H-like domain	RNase_H superfamily ; InsA N-terminal domain	No	Cytoplasmic 3.910	2kfnA- 18 -38- 0.005
7670386	No	No	No	No	Cytoplasmic 4.088	1v5eA- 32 -36 0.001
7670389	No	Acyl-CoA N-	Acetyltransferase	No	Cytoplasmic 3.988	2beiB 21 42 1e-04

Table 21. *Pseudomonas* phage LUZ24

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5896679	No	No	CENP-A-nucleosome distal (CAD) centromere subunit ; Tropomyosin like ; Peptidase of plants and bacteria	No	Cytoplasmic 3.154	No
5896684	HNH_3 super family[cl16264], HNH endonuclease	No	HNH endonuclease; Esterase PHB depolymerase	No	Periplasmic 1.693 Cytoplasmic 1.351 Extracellular 1.255	1u3eM -15- 63- 4e-11
5896692	No	No	Putative amidoligase enzyme	No	Cytoplasmic 4.429	No
5896700	GGCT_like[cd06661], GGCT-like domains, also called AIG2-like family. Gamma-glutamyl cyclotransferase	AIG2-like;Butirosin biosynthesis, BtrG-like;	AIG2-like family	Uncharacterized ACR	Cytoplasmic 3.939	2qikA -25- 91 -8e-20
5896703	Phage_endo_I super family[cl11622], Phage endonuclease I;	No	No	No	Cytoplasmic 2.730	1m0dA- 21- 34-0.008
5896705	No	No	Dynamitin	No	Cytoplasmic 3.490	No
5896707	No	No	TATA element modulatory factor 1 DNA binding ; YjbD family (DUF3811)	No	Cytoplasmic 2.508	No
5896708	PHA01782 super family[cl10333], hypothetical protein	No	PCI domain	No	Periplasmic 3.969	No
5896714	No	No	DASH complex subunit Dad1 ; AAA domain ; Rabaptin	No	Cytoplasmic 1.988	No
5896715	PRK14571[PRK14571], D-alanyl-alanine synthetase A;RimK[COG0189] , Glutathione synthase/Ribosomal protein S6 modification enzyme (glutaminyl	ATP-grasp fold, subdomain 2	Calcineurin-like phosphoesterase superfamily domain; Beta-ketoacyl synthase, C-terminal domain; D-ala D-ala ligase C-terminus	D-alanine-D-alanine ligase and related ATP-grasp enzymes	Cytoplasmic 4.382	2p0aA -12 -53 -1e-07

transferase)

5896716	No	No	Cytochrome oxidase c subunit VIII	No	Cytoplasmic Periplasmic	1.757 1.249	No
5896717	HNH_3 super family[cl16264], HNH endonuclease	No	HNH endonuclease ; Gas vesicle synthesis protein GvpO	No	Periplasmic	2.778	1u3eM -15 71- 1e-13
5896721	No	n	DmpG-like communication domain	No	Cytoplasmic	2.719	No
5896729	COOH-NH2_lig super family[cl16852], Phage phiEco32-like COOH.NH2 ligase-type 2	No	Phage phiEco32-like COOH.NH2 ligase-type 2	No	Cytoplasmic	2.986	No
5896730	No	No	Phage tail protein	No	Cytoplasmic	3.149	No
5896741	Collar super family[cl10081], Phage Tail Collar Domain; MdpB[COG4675], Microcystin-dependent protein [Function unkNown]	No	Phage Tail Collar Domain	No	Extracellular	4.115	1pdiA 13 50 4e-07

Table 22. *Pseudomonas* phage LUZ19

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5896746	PLN02225[PLN02225], 1-deoxy-D-xylulose-5-phosphate synthase	No	No	No	InnerMembrane 2.516	No
5896749	PolyA_pol super family[cl14657], Poly A polymerase head domain	No	Poly A polymerase head domain	No	Periplasmic 2.474 Cytoplasmic 1.778	1ou5A- 61- 36- 0.004
5896751	No	No	Spc7 kinetochore protein; KR domain ; CLIP, MHC2 interacting	No	Cytoplasmic 3.095	No
5896752	DNA_pol_A super family[cl02626], Family A polymerase primarily fills DNA gaps that arise during DNA repair	No	No	DNA polymerase I - 3'-5' exonuclease and polymerase domains	Cytoplasmic 2.346	1bgxT- 47- 43- 8e-05
5896759	No	No	Fungal fruit body lectin	No	Cytoplasmic 1.679	No
5896764	PHA01782 super family[cl10333], hypothetical protein	No	No	Thiamine biosynthesis protein ThiC	Cytoplasmic 2.415 Periplasmic 2.020	No
5896777	No	No	No	Helicase subunit of the DNA excision repair complex	Cytoplasmic 3.921	No
5896778	No	No	Laminin Domain I ; EspA-like secreted protein	No	Cytoplasmic 2.220	No
5896782	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily;	Ribonuclease H-like domain	RNase_H superfamily ; InsA N-terminal domain	No	Cytoplasmic 3.974	1q8iA- 14- 62- 1e-10
5896783	, putative structural protein	Acyl-CoA N-acyltransferase	Acetyltransferase (GNAT) domain	No	Cytoplasmic 3.974	2beiB- 15- 39- 7e-04

5896784	No	No	No	No	Cytoplasmic	4.041	1v5eA- 32- 37-8e-04
5896785	PHA01808 super family[cl10339], putative structural protein		EAP30/Vps36 family ; Protein gp23 (Bacteriophage A118)	No	Periplasmic Cytoplasmic	1.948 1.289	No
5896788	No	No	Herpesvirus pp38 phosphoprotein	No	Cytoplasmic	2.127	No

Table 23. *Pseudomonas* phage YuA

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5797678	No	No	No	No	Cytoplasmic 3.934	3c5eA- 21- 37- 0.004
5797679	HepA[COG0553], Superfamily II DNA/RNA helicases, SNF2 family [Transcription / DNA replication, recombination, and repair]	SNF2-related; Helicase, C- terminal;	SNF2 family N- terminal domain ; Helicase conserved C- terminal domain	No	Cytoplasmic 4.833	1z6aA-19-171-3e-45
5797684	UvrD_C_2 super family[cl15862], Family description; P- loop_NTPase super family[cl09099], P- loop containing Nucleoside Triphosphate Hydrolases		AAA domain; Family description	ATP- dependent exoDNase (exonuclease V), alpha subunit - helicase superfamily I member	Cytoplasmic 3.654	1w36D- 17-139- 1e-33
5797688	HDc super family[cl00076], Metal dependent phosphohydrolase s with conserved 'HD' motif	HD domain	HD domain	No	Cytoplasmic 4.454	2gz4A -19 -150-1e-37
5797689	No	No	Kelch motif	No	Cytoplasmic 2.564 Cytoplasmic 3.594	No
5797694	No	No	Poly(hydroxyalcaNoat e) granule associated protein (phasin)	No	Cytoplasmic 2.880	No
5797695	No	No	No	No	Cytoplasmic 2.574	2iakA- 37 -35- 0.007
5797697	Trp-synth-beta_II super family[cl00342], Tryptophan synthase beta superfamily (fold type II)	Pyridoxal phosphate- dependent enzyme, beta subunit	Pyridoxal-phosphate dependent enzyme ;Phytotoxin PcF protein	No	Periplasmic 3.008	1f2dA- 18- 50- 8e-07
5797703	No	No	Lysozyme-like	No	Cytoplasmic 1.612 OuterMembrane 1.100 Periplasmic 1.084	No

5797704	GGDEF[cd01949], Diguanylate- cyclase (DGC) or GGDEF domain; Diguanylate- cyclase (DGC) or GGDEF domain	Diguanylate cyclase, predicted; Adenylyl cyclase class-3/4/guanylyl cyclase	GGDEF domain ; Glycosyl transferase family 2	GGDEF domain	Cytoplasmic	2.632	1w25A- 24 -128- 4e-31
5797706	HTH_XRE super family[cl15761], Helix-turn-helix XRE-family like proteins. Prokaryotic DNA binding proteins belonging to .the xNobiotic response element family of transcriptional regulators.	Lambda repressor- like, DNA-binding domain	Helix-turn-helix domain	No	Cytoplasmic	3.749	No
5797709	PRK13962[PRK139 62], bifunctional phosphoglycerate kinase/triosephos phate isomerase	No	No	No	Cytoplasmic	2.505	2zc0A -28- 34- 0.008
5797710	No	No	Influenza RNA- dependent RNA polymerase subunit PB2	Carbamoylpho sphate synthase large subunit (split gene in MJ)	Cytoplasmic	2.329	No
5797712	Antirestrict super family[cl03947], Antirestriction protein	Antirestriction protein	Antirestriction protein	No	Cytoplasmic	4.297	No
5797716	No	No	Outer membrane protein (OmpH-like)	No	Cytoplasmic Cytoplasmic	4.697 2.780	No
5797719	No	No	Defensin propeptide	No	Periplasmic	2.668	No
5797721	No	No	FUN14 family; Viral A- type inclusion protein repeat	No	Cytoplasmic	4.089	No
5797722	RNR_II_dimer[cd0 2888], Class II ribonucleotide reductase, dimeric form; NrdJ_Z[TIGR02504 , ribonucleoside- diphosphate reductase, adeNosylcobalami n-dependent	Ribonucleotide reductase large subunit, C- terminal; Ribonucleoside- diphosphate reductase, adeNosylcobalami n-dependent	Ribonucleotide reductase, barrel domain	Ribonucleotid e reductase alpha subunit	Cytoplasmic Periplasmic Cytoplasmic	2.476 1.984 4.739	1zyzA-18-394 -1e-112
5797724	PRK13539[PRK135 39], cytochrome c biogenesis protein	No	AAA domain	ABC-type cobalamin/Fe 3+-	Cytoplasmic	3.693	1ki9B -11- 37 -0.003

	CcmA			siderophores transport systems, ATPase components			
5797726	No	No	No	No	Cytoplasmic	2.972	1xjA- 28 -262 -8e-71
5797728	NTP-PPase super family[cl16941], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain superfamily	No	HD domain;MazG nucleotide pyrophosphohydrolase domain	No	Cytoplasmic	2.548	2oieB- 22- 53 -4e-08
5797730	Radical_SAM[cd01335], Radical SAM superfamily.; Radical_SAM[pfam04055], Radical SAM superfamily	Aldolase-type TIM barrel	Radical SAM superfamily	DNA repair photolyase	Cytoplasmic Periplasmic	2.349 2.204	1tv8A- 13- 76- 1e-14
5797732	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily	Ribonuclease H-like domain	Peptidase dimerisation domain	No	Cytoplasmic Periplasmic	2.311 1.875	2p1jB- 14- 49- 8e-07
5797737	No	No	Prolyl 4-Hydroxylase alpha-subunit, N-terminal region ; Plastid and cyanobacterial ribosomal protein (PSRP-3 / Ycf65)	No	Cytoplasmic	2.741	No
5797743	TS_Pyrimidine_HMase super family[cl00358], Thymidylate synthase and pyrimidine hydroxymethylase	Thymidylate synthase; Thymidylate synthase/dCMP hydroxymethylase domain	Thymidylate synthase	Thymidylate synthase	Cytoplasmic	2.750	1j3kC- 16 -241-1e-64
5797744	No	Zinc finger, CCHC-type	Predicted membrane protein (DUF2318) ; Zinc knuckle	No	Cytoplasmic	3.377	No
5797746	Terminase_6[pfam03237], Terminase-like family	Terminase, large subunit	Terminase-like family	No	Cytoplasmic	3.128	2o0jA -12- 36- 2e-04

5797747	DNA_pol_A super family[cl02626], Family A polymerase primarily fills DNA gaps that arise during DNA repair, recombination and replication; DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamilyPolA[COG0749], DNA polymerase I - 3'-5' exonuclease and polymerase domains [DNA replicationrecombination, and repair]	DNA-directed DNA polymerase, family A, palm domain; DNA polymerase A; Ribonuclease H-like domain; DNA-directed DNA polymerase, family A, conserved site	DNA polymerase family A	DNA polymerase I - 3'-5' exonuclease and polymerase domains	Cytoplasmic	4.064	2kfnA- 24- 635- 0.0
5797749	cytidine_deaminase-like super family[cl00269], Cytidine and deoxycytidylate deaminase zinc-binding region; SGNH_hydrolase super family[cl01053], SGNH_hydrolase, or GDSL_hydrolase, is a diverse family of lipases and esterases	CMP/dCMP deaminase, zinc-binding; Cytidine deaminase-like; dCMP deaminase	Cytidine and deoxycytidylate deaminase zinc-binding region	Deoxycytidylate deaminase	Cytoplasmic	2.507	2hvwA -26-101-1e-22
5797750	primase_Cterm super family[cl15909], phage/plasmid primase, P4 family, C-terminal domain	No	Virulence-associated protein E ; Poxvirus D5 protein-like	No	Cytoplasmic	3.547	No

Table 24. *Pseudomonas* phage phi-2

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
8673566	No	No	Ecdysteroid kinase	No	Periplasmic 4.164	No
8673572	2OG-Fell_Oxy super family[cl15773], 2OG-Fe(II) oxygenase superfamily	Oxoglutarate/iron-dependent dioxygenase	2OG-Fe(II) oxygenase superfamily	No	Extracellular 1.964 Periplasmic 1.833	2hbtA- 18-41-9e-05
8673574	PolyA_pol super family[cl14657], Poly A polymerase head domain	No	Poly A polymerase head domain	tRNA nucleotidyltransferase/poly(A) polymerase	Cytoplasmic 3.898	1ou5A- 21 -39 7e-04
8673575	No	No	Leucine-rich repeats of kinetochore protein Cenp-F/LEK1 ; TATA element modulatory factor 1 DNA binding	No	Cytoplasmic 2.333	No
8673578	No	No	Part of AAA domain	No	Cytoplasmic 3.584	No
8673580	No	No	Fibronectin-binding protein A N-terminus (FbpA)	No	Cytoplasmic 3.773	No
8673586	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily	Ribonuclease H-like domain	RNase_H superfamily	No	Cytoplasmic 4.130	1q8iA- 14- 61- 3e-10
8673589	No	Acyl-CoA N-acyltransferase	Acetyltransferase (GNAT) family	No	Cytoplasmic 4.229	1i12D -11- 37- 0.002
8673606	No	No	Pre-toxin domain with VENN motif ; Ilarvirus coat protein	Seryl-tRNA synthetase	Extracellular 1.777 Cytoplasmic 1.734	No

Table 25. *Pseudomonas* phage LUZ7

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
8684362	No	No	No	No	Cytoplasmic 1.770 Periplasmic 1.133	1ciiA- 25- 44- 8e-06
8684364	No	No	ThiC-associated domain ; ThiC-associated domain	Superoxide dismutase	InnerMembrane 1.674 Cytoplasmic 1.394 Extracellular 1.041	No
8684365	No	No	YesK-like protein ; 7tm Chemosensory receptor ;	No	InnerMembrane 1.792 Cytoplasmic 1.171	No
8684370	Mito_fiss_Elm1 super family[cl09404], Mitochondrial fission ELM1	No	PhnA protein	No	Cytoplasmic 4.079	No
8684377	No	No	Endoribonuclease XendoU ; ATP synthase D chain, mitochondrial (ATP5H)	No	Cytoplasmic 2.847	No
8684381	No	No	PMP-22/EMP/MP20/Claudin tight junction; Cytochrome oxidase assembly protein ; Positive regulator of sigma(E), RseC/MucC ; Enoyl-CoA hydratase/isomerase family	No	InnerMembrane 2.215	No
8684382	SMC_prok_B[TIGR02168], chromosome segregation protein SMC, common bacterial type	No	Ribonuclease 2-5A ; LXG domain of WXG superfamily ; Syntaxin 6, N-terminal ; Heat shock factor binding protein 1 ; Putative cell-wall binding lipoprotein ; Laminin Domain I ; SH3 domain-binding protein 5 (SH3BP5); IncA protein ; Calcium binding and coiled-coil domain (CALCOCO1) like; Gammaherpesvirus protein of unknown function (DUF848) ; GDP/GTP exchange factor Sec2p ; MGS-like domain ; Fibrinogen alpha/beta	ATPase involved in DNA repair	Cytoplasmic 3.421	2dfsA- 13 41 1e-04

			chain family ; DNA packaging protein FI				
8684384	No	No	V-type ATPase 116kDa subunit family ; SIS domain; PDZ domain of MCC-2 bdg protein for Usher syndrome	No	Cytoplasmic	4.006	No
8684385	No	No	L-arabinose isomerase C-terminal domain	No	Cytoplasmic	2.696	No
8684388	AAA[cd00009], The AAA+ (ATPases Associated with a wide variety of cellular Activities) superfamily represents an ancient group of ATPases belonging to the ASCE (for additional strand, catalytic E) division of the P-loop NTPase fold; ClpX[COG1219], ATP-dependent protease Clp, ATPase subunit [Posttranslational modification, protein turnover, chaperones]	AAA+ ATPase domain; ATPase, AAA-2	AAA domain (Cdc48 subfamily) ;	ATP-dependent protease Clp, ATPase subunit	Cytoplasmic	3.651	1g4aE- 16- 268- 8e-7
8684395	cytidine_deaminase-like super family[cl00269], Cytidine and deoxycytidylate deaminase zinc-binding region	CMP/dCMP deaminase, zinc-binding; Cytidine deaminase; APOBEC/CMP deaminase, zinc-binding; Cytidine deaminase-like	Cytidine and deoxycytidylate deaminase zinc-binding region	Deoxycytidylate deaminase	Periplasmic	2.880	No
8684400	No	No	Mediator of CRAC channel activity	No	Cytoplasmic InnerMembrane	1.908 1.517	No
8684401	No	No	Inhibitor of apoptosis-promoting Bax1 ; 5TMR of 5TMR-LYT ; Lipid A Biosynthesis N-terminal domain	No	InnerMembrane Cytoplasmic	2.039 1.534	No
8684404	No	No	MbeB-like, N-term conserved region	No	Periplasmic OuterMembrane Cytoplasmic	1.475 1.382 1.250	No

8684408	SGNH_hydrolase super family[cl01053], SGNH_hydrolase, or GDSL_hydrolase, is a diverse family of lipases and esterases; ATP-synt_B super family[cl07975], ATP synthase B/B' CF(0)	Esterase, SGNH hydrolase-type; Esterase, SGNH hydrolase-type, subgroup	Potato leaf roll virus readthrough protein	No	Extracellular OuterMembrane	2.331 1.849	No
8684411	COG4733[COG4733], Phage-related protein, tail component [Function unknown]	Immunoglobulin-like fold	Reovirus sigma C capsid protein ; Fibrin C-terminal region ; Histone acetyltransferase subunit NuA4	No	Extracellular OuterMembrane	2.272 2.131	2edeA- 20- 38- 0.004
8684413	No	No	Bacteriophage Mu transposase	No	Cytoplasmic Periplasmic	1.854 1.523	No
8684422	No	No	Carbamoyl-phosphate synthase small chain, CPSase domain	No	Cytoplasmic	2.728	No
8684424	No	No	MFS/sugar transport protein ; Colicin V production protein	No	InnerMembrane Cytoplasmic	1.830 1.476	No
8684426	No	No	No	No	Periplasmic	2.253	2o5vA- 16 -34- 0.001
8684433	No	No	Flagellar FljJ protein	No	Cytoplasmic	3.936	No
8684435	trimeric_dUTPase[cd07557], Trimeric dUTP diphosphatases	DeoxyUTP pyrophosphatase; DeoxyUTP pyrophosphatase, dUTPase subfamily	dUTPase	dUTPase	Cytoplasmic	2.988	1q5hC-34-179-2e-46
8684445	No	No	MetRS-N binding domain	No	Cytoplasmic	2.058	No
8684449	No	No	Phosphatidylinositol 3- and 4-kinase	No	Cytoplasmic Periplasmic	1.861 1.426	No
8684467	No	No	Tim10/DDP family zinc finger	No	Cytoplasmic Periplasmic Extracellular	1.612 1.311 1.243	No
8684468	No	No	Proteins of 100 residues with WXG ; Uncharacterized conserved protein H4 (DUF2046)	No	Periplasmic Cytoplasmic	2.003 1.395	No

Table 26. *Pseudomonas* phage LIT1

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
8684263	No	No	YhzD-like protein ; Mannan-binding protein	Succinate dehydrogenas e/fumarate reductase Fe- S protein	Cytoplasmic 2.952	No
8684267	No	No	MA3 domain	No	Cytoplasmic 3.186	No
8684271	AP_MHD_Cterm super family[cl10970], C- terminal domain of adaptor protein (AP) complexes medium mu subunits and its homologs (MHD);	No	No	No	InnerMembrane 2.558	No
8684279	No	No	YHS domain	No	Cytoplasmic 2.395	No
8684285	No	No	IncA protein ; UbiA prenyltransferase family ; MMPL family	No	InnerMembrane 2.421	No
8684287	No	No	IncA protein; Growth- arrest specific micro- tubule binding; Phage miNor structural protein GP20 ; Autophagy protein 16 (ATG16)	No	Cytoplasmic 3.529	No
8684289	No	No	Afadin- and alpha - actinin-Binding ; MechaNosensitive ion channel porin domain; Plant protein of unkNown function (DUF827) ; ATP synthase subunit D ; bZIP transcription factor	No	Cytoplasmic 3.702	No
8684292	P-loop_NTPase super family[cl09099], P- loop containing Nucleoside Triphosphate Hydrolases	ATPase, AAA-2	AAA domain (Cdc48 subfamily)	ATP- dependent protease Clp, ATPase subunit	Cytoplasmic 2.270	1um8A- 20- 228-1e-60

8684300	cytidine_deaminase-like super family[cl00269], Cytidine and deoxycytidylate deaminase zinc-binding region	CMP/dCMP deaminase, zinc-binding; APOBEC/CMP deaminase, zinc-binding; Cytidine deaminase-like	Cytidine and deoxycytidylate deaminase zinc-binding region	Deoxycytidylate deaminase	Cytoplasmic	3.030	2hwvA- 24- 159- 2e-40
8684306	No	No	Phosphotransferase system, EIIC	No	InnerMembrane Periplasmic Cytoplasmic	1.543 1.482 1.372	No
8684307	No	No	Flagellar protein (FlbD) ; Bacteriophage lambda Kil protein	Glucose-6-phosphate 1-dehydrogenase	Cytoplasmic	3.301	No
8684309	No	No	CDK-activating kinase assembly factor MAT1 ; IncA protein ; DivIVA protein; JNK_SAPK-associated protein-1 ; UV radiation resistance protein and autophagy-related subunit 14 ;	No	Cytoplasmic	2.746	1g8xA- 5- 36 -0.006
8684311	PRK07710[PRK07710], acetolactate synthase catalytic subunit	No	Pilus assembly protein, PilP	Thiamine pyrophosphate-requiring enzymes [acetolactate synthase, pyruvate dehydrogenase (cytochrome), glyoxylate carboxylase, phosphoenolpyruvate decarboxylase]	Cytoplasmic	3.020	No
8684312	nirA[PRK09567], ferredoxin-nitrite reductase	No	No	No	Cytoplasmic	3.852	No
8684314	SGNH_hydrolase[cd00229], SGNH_hydrolase, or GDSL_hydrolase, is a diverse family of lipases and esterases	Esterase, SGNH hydrolase-type; Esterase, SGNH hydrolase-type, subgroup	GDSL-like Lipase/Acylhydrolase family ; FixH ; AF-4 proto-oncoprotein	No	OuterMembrane Extracellular	2.139 1.463	1fxwF- 18- 61- 8e-12
8684317	COG4733[COG4733], Phage-related protein, tail component [Function	Immunoglobulin-like fold	Reovirus sigma C capsid protein; Mbed/MobD like	Methyl-accepting chemotaxis protein	Extracellular OuterMembrane	2.239 1.693	2e7sD- 20- 41- 6e-04

unkNown]

8684324	No	No	RecO N terminal ; Centrosome localisation domain of PPC89	No	Cytoplasmic	2.764	No
8684327	No	No	Levivirus coat protein	No	Cytoplasmic Periplasmic	1.943 1.552	No
8684330	No	No	Uncharacterised Sugar-binding Domain	No	Periplasmic Cytoplasmic	2.170 1.770	No
8684331	No	No	NfeD-like C-terminal, partner-binding ; Major Facilitator Superfamily ; Sulfite exporter TauE/SafE;	No	InnerMembrane Cytoplasmic Periplasmic	1.741 1.227 1.068	No
8684333	No	No	Lantibiotic streptin immunity protein	No	OuterMembrane Extracellular	1.798 1.287	No
8684342	trimeric_dUTPase[cd07557], Trimeric dUTP diphosphatases	DeoxyUTP pyrophosphatase; DeoxyUTP pyrophosphatase, dUTPase subfamily;	dUTPase	dUTPase	Cytoplasmic Extracellular Periplasmic	1.571 1.564 1.174	1sixA- 36- 142- 2e-35
8684343	No	No	Uncharacterized protein, homolog of nitrogen regulatory protein PII	No	Cytoplasmic	2.790	No
8684348	No	No	Potyviriidae polyprotein	No	Cytoplasmic	2.087	No
8684350	Description Pssmld Multi- dom E-value, COesterase[pfam0 0135], Carboxylesterase family;	No	No	No	Cytoplasmic	2.978	No
8684352	No	No	No	Methyl- accepting chemotaxis protein	Cytoplasmic	2.545	No

Table 27. *Pseudomonas* phage LBL3

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
6803696	No	No	Fibronectin-binding repeat	No	Cytoplasmic 1.806 Periplasmic 1.499	No
6803697	No	No	Acetyltransferase (GNAT) domain	No	Cytoplasmic 2.254	No
6803702	No	No	Ribosomal L18ae/LX protein domain ; Archaeal flagella protein	No	Cytoplasmic 3.847	No
6803707	No	No	VRR-NUC domain	No	Cytoplasmic 3.978	No
6803708	No	No	V-ATPase subunit H	No	Cytoplasmic 3.086	No
6803713	No	No	V-type ATPase 116kDa subunit family ; Vitamin-D-receptor interacting Mediator subunit 4 ; Tektin family ; TSC-22/dip/bun family	No	Cytoplasmic 4.243	No
6803715	No	No	Outer membrane protein (OmpH-like) ; TipAS antibiotic-recognition domain ;	No	Periplasmic 1.564 InnerMembrane 1.286 OuterMembrane 1.030	No
6803716	NTP-PPase_u3[cd11540], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain found in a group of uncharacterized proteins from bacteria and archaea	Phosphoribosyl-ATP pyrophosphohydrolase-like	Phosphoribosyl-ATP pyrophosphohydrolase	No	Cytoplasmic 4.548	1vmgA- 22- 50- 3e-07
6803718	No	No	HrpJ-like domain	No	Cytoplasmic 3.345	2qzva- 15- 38 -0.003
6803721	DnaJ_CXXCXGXG[pfam00684], DnaJ central domain	No	PADR1 (NUC008) domain; DnaJ central domain ; Phage DNA packaging protein Nu1	No	Periplasmic 2.604 Cytoplasmic 2.211	No
6803723	No	No	Neuronal voltage-dependent calcium channel alpha 2acd	No	Cytoplasmic 4.638	No

6803725	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;UvrD_C_2 super family[cl15862], Family description	No	AAA domain; Family description	ATP-dependent exonuclease (exonuclease V), alpha subunit - helicase superfamily I member	Cytoplasmic	4.742	1w36D- 18- 141- 4e-34
6803726	No	No	Ca2+ regulator and membrane fusion protein Fig1 ;	No	InnerMembrane Cytoplasmic	1.569 1.326	No
6803736	No	No	FAD binding domain	No	Cytoplasmic	3.233	2b9wA- 40 -38- 0.004
6803737	No	No	Uncharacterized protein conserved in bacteria (DUF2345); Bacteriophage Mu Gp45 protein	No	Periplasmic Extracellular	1.769 1.391	No
6803746	No	No	YebG protein ; Alpha helical Porin B	No	Cytoplasmic	3.078	No
6803752	No	No	Colicin pore forming domain ; Photosystem II reaction centre N protein (psbN)	No	Cytoplasmic	2.410	1ciiA- 25- 40-2e-04
6803753	No	No	Beta propeller domain	No	Cytoplasmic Extracellular	1.792 1.387	No
6803758	No	No	DNA polymerase family B exonuclease domain, N-terminal	No	Cytoplasmic	2.962	No
6803762	No	No	Nitrile hydratase beta subunit	No	Cytoplasmic	3.563	No
6803763	No	No	NIPSNAP ; SigmaW regulon antibacterial ; Ribosomal protein S27a	No	Cytoplasmic	3.148	No
6803764	No	No	Flagellar P-ring protein	No	Cytoplasmic Periplasmic	2.215 1.745	No
6803765	No	No	Mediator complex subunit 30 ; Tc3 transposase ; Nuclear pore complex subunit Nro1	No	Cytoplasmic	3.796	No
6803766	No	No	DnaJ central domain	No	Cytoplasmic	3.377	No
6803770	No	No	Iron Transport-associated domain	No	Cytoplasmic	2.614	No
6803771	pseT[PHA02530], polynucleotide kinase	HAD-like domain	HAD superfamily, subfamily IIIB (Acid phosphatase) ; Ketopantoate	No	Cytoplasmic	3.419	1yj5B-17- 37- 0.004

6803774	No	No	reductase PanE/ApbA C terminal					
			Bacteriophage protein of unkNown function (DUF646)	No	Cytoplasmic InnerMembrane	2.120 1.518	No	

Table 28. *Pseudomonas* phage LMA2

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
6803782	No	No	Formin Homology 2 Domain ; NIPSNAP	No	Cytoplasmic 3.226	No
6803783	No		Biofilm development protein YmgB/AriR	No	Cytoplasmic 2.025	No
6803788	No	No	No	No	Cytoplasmic 2.528	1ciiA- 22- 35- 0.003
6803792	No	No	Iron Transport-associated domain	No	Cytoplasmic 2.628	No
6803796	No	No	YebG protein	No	Cytoplasmic 3.129	No
6803797	pseT[PHA02530], polynucleotide kinase	HAD-like domain	HAD superfamily, subfamily IIIB (Acid phosphatase) ; Glyoxalase-like domain ; Ketopantoate reductase PanE/ApbA C terminal	No	Cytoplasmic 3.584	1ltqA- 20- 163- 4e-41
6803799	NTP-PPase_u3[cd11540], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain found in a group of uncharacterized proteins from bacteria and archaea	Phosphoribosyl-ATP pyrophosphohydrolase-like	Phosphoribosyl-ATP pyrophosphohydrolase	No	Cytoplasmic 4.743	3beoA- 23- 35- 0.005
6803801	No	No	DnaJ central domain	No	Cytoplasmic 3.280	No
6803804	No	No	V-ATPase subunit H	No	Cytoplasmic 3.112	No
6803808	No	No	T4 recombination endonuclease VII, dimerisation ;	No	Cytoplasmic 2.331	No
6803809	Hs1pro-1_N superfamily[cl06319], Hs1pro-1 N-terminus	No	No	No	Cytoplasmic 2.852	No
6803813	No	No	Archaeal flagella protein	No	Cytoplasmic 3.918	No
6803816	No	No	Mediator complex subunit 30; Tc3 transposase ; Nuclear	No	Cytoplasmic 3.834	No

			pore complex subunit Nro1				
6803823	No	No	PQQ enzyme repeat ; Ribosomal protein S27a	No	Cytoplasmic	2.499	No
6803824	No	No	Tumour protein D52 family ; Ribosomal L29 protein	No	Cytoplasmic	2.572	No
6803828	No	No	V-type ATPase 116kDa subunit family ; Vitamin-D-receptor interacting Mediator subunit 4 ; Tektin family ; TSC- 22/dip/bun family	No	Cytoplasmic	4.165	No
6803834	No	No	Fibronectin-binding repeat	No	Cytoplasmic Periplasmic	1.806 1.499	No
6803835	No	No	No	No	Cytoplasmic	3.466	1jxhA 31 36 0.004
6803839	No	No	Bacteriophage Mu Gp45 protein	No	Periplasmic Extracellular	1.775 1.220	No
6803847	No	No	Non-structural protein C	No	Cytoplasmic	3.566	No
6803856	No	No	Outer membrane protein (OmpH-like) ; TipAS antibiotic- recognition domain;	No	Periplasmic InnerMembrane OuterMembrane	1.564 1.286 1.030	No
6803858	No	No	FAD binding domain	No	Cytoplasmic	3.694	2b9wA -40- 38- 0.005
6803860	No	No	Eukaryotic and archaeal DNA primase small subunit	No	Cytoplasmic	2.392	No
6803861	No	No	VRR-NUC domain	No	Cytoplasmic	3.562	No
6803862	No	No	HrpJ-like domain	No	Cytoplasmic	3.261	2qzvA -15- 38- 0.003
6803865	No	No	Beta propeller domain	No	Cytoplasmic Extracellular	1.755 1.326	No
6803866	No	No	Photosystem P840 reaction-centre cytochrome c-551 ; Membrane MotB of proton-channel complex MotA/MotB	Flagellar motor protein	InnerMembrane Cytoplasmic	1.783 1.270	No
6803868	SMC_prok_B[TIGR 02168], chromosome segregation protein SMC, common bacterial type	No	INositol 1,4,5- trisphosphate/ryaNodi ne receptor ; Cortexillin I, coiled coil ; SecD export protein N-terminal TM region	Chromosome segregation ATPases	Cytoplasmic	3.869	1qvrA- 26- 39- 7e-04

6803873	RNase_H super family[cl14782], RNase H is an endonuclease that cleaves the RNA strand of an RNA/DNA hybrid in a sequence Non-specific manne	No	No	No	Cytoplasmic	4.577	No
6803874	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; UvrD_C_2 super family[cl15862], Family description	No	AAA domain ; Family description	ATP-dependent exoDNase (exonuclease V), alpha subunit - helicase superfamily I member	Cytoplasmic	4.760	1w36D- 18- 141- 3e-34

Table 29. *Pseudomonas* phage SN

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
7113303	No	No	Helicobacter pylori protein of unkNown function (DUF874)	No	Cytoplasmic 3.314	No
7113304	AdoMet_MTases super family[cl16911], S-adeNosylmethionine-dependent methyltransferases (SAM or AdoMet-MTase), class I	No	Methyltransferase domain	No	Cytoplasmic 4.358	No
7113305	No	No	TATA element modulatory factor 1 DNA binding ; Mediator complex subunit 30 ; Tc3 transposase ; Nuclear pore complex subunit Nro1	No	Cytoplasmic 3.851	No
7113306	No	No	3-Oxoacyl-[acyl-carrier-protein (ACP)] synthase III C terminal	No	Cytoplasmic 2.726	No
7113316	No	No	Beta propeller domain	No	Cytoplasmic 1.728 Extracellular 1.365	No
7113325	No	No	No	No	Periplasmic 2.172	1korB- 30- 35- 0.005
7113339	No	No	V-ATPase subunit H	No	Cytoplasmic 3.007	No
7113341	NTP-PPase_u3[cd11540], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8) MazG-like domain found in a group of uncharacterized proteins from bacteria and archaea	Phosphoribosyl-ATP pyrophosphohydrolase-like	Phosphoribosyl-ATP pyrophosphohydrolase	No	Cytoplasmic 3.558	1vmgA- 21- 52- 1e-07
7113342	No	No	Outer membrane protein (OmpH-like) ; TipAS antibiotic-recognition domain	No	Periplasmic 1.564 InnerMembrane 1.286 OuterMembrane 1.030	No
7113344	No	No	V-type ATPase 116kDa subunit family; Vitamin-D-receptor	No	Cytoplasmic 4.153	No

interacting Mediator
subunit 4 ; TSC-
22/dip/bun family

7113345	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases;HepA[COG0553], Superfamily II DNA/RNA helicases, SNF2 family [Transcription / DNA replication, recombination, and repair]	SNF2-related; Helicase, C-terminal;	SNF2 family N-terminal domain ; Helicase conserved C-terminal domain	Superfamily II DNA/RNA helicases, SNF2 family	Cytoplasmic	4.363	1z6aA- 20 485 1e-138
7113349	pseT[PHA02530], polynucleotide kinase	HAD-like domain	HAD superfamily, subfamily IIIB (Acid phosphatase) ; Glyoxalase-like domain ; Ketopantoate reductase PanE/ApbA C terminal	No	Cytoplasmic	3.538	1ltqA 19 164 2e-41
7113350	No	No	Iron Transport-associated domain	No	Cytoplasmic	2.628	No
7113351	Thy1 super family[cl03630], Thymidylate synthase complementing protein	Thymidylate synthase ThyX	Thymidylate synthase complementing protein	Predicted alternative thymidylate synthase	Cytoplasmic	3.873	2af6G 20 212 5e-56
7113352	No	No	YebG protein	No	Cytoplasmic	3.108	No
7113357	No	No	FAD binding domain	No	Cytoplasmic	3.535	2b9wA 40 38 0.005
7113360	No	No	Eukaryotic and archaeal DNA primase small subunit	No	Cytoplasmic Periplasmic	2.155 1.578	No
7113362	PADR1 super family[cl06892], PADR1 (NUC008) domain; PRK03918[PRK03918], chromosome segregation	No	PADR1 (NUC008) domain ; DnaJ central domain ; RIM-binding protein of the cytomatrix active zone ; Seryl-tRNA synthetase N-terminal	No	Periplasmic	3.440	No

	protein		domain ; Centrosome localisation domain of PPC89					
7113365	No	No	Photosystem P840 reaction-centre cytochrome c-551 ; Membrane MotB of proton-channel complex MotA/MotB	Flagellar motor protein	InnerMembrane Cytoplasmic	1.700 1.194	No	
7113368	No	No	HrpJ-like domain ;	No	Cytoplasmic	3.372	2qzvA 13 37 0.003	
7113372	No	No	Glycerophosphoryl diester phosphodiesterase	No	Cytoplasmic	3.677	No	
7113376	No	No	Double zinc ribbon	No	Cytoplasmic Periplasmic	2.192 1.518	No	
7113380	No	No	Coxiella burnetii protein of unkNown function (DUF762)	No	Cytoplasmic	3.926	No	
7113382	SMC_prok_B[TIGR 02168], chromosome segregation protein SMC, common bacterial type	No	INositol 1,4,5-trisphosphate/ryaNodine receptor ; Cortexillin I, coiled coil ; IncA proteinSecD export protein N-terminal TM region ; DNA packaging protein FI ; Interferon-induced 35 kDa protein (IFP 35) N-terminus	Chromosome segregation ATPases	Cytoplasmic	4.001	1qvrA -25- 41- 2e-04	
7113389	No	No	SigmaW regulon antibacterial; Ribosomal protein S27a	No	Cytoplasmic	3.138	No	

Table 30. *Pseudomonas* phage LKA1

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
5687472	DnaQ_like_exo super family[cl10012], DnaQ-like (or DEDD) 3'-5' exonuclease domain superfamily	Ribonuclease H-like domain	RNase_H superfamily ; Helix-turn-helix domain of transposase family ISL3 ; Restriction alleviation protein Lar ; Transcription factor zinc-finger	DNA polymerase elongation subunit (family B)	Cytoplasmic 2.111 Extracellular 1.424	2kfnA- 19- 46- 2e-05
5687475	No	No	Vesicle transport v-SNARE protein N-terminus	No	Cytoplasmic 2.644	No
5687477	SMC_prok_B[TIGR 02168], chromosome segregation protein SMC, common bacterial type	No	Proteins of 100 residues with WXG ; Autophagy protein 16 (ATG16) ; Fez1 ; AAA domain ; Ezrin/radixin/moesin family ; Syntaxin ; 4-hydroxyphenylacetate 3-hydroxylase N terminal ; Herpesvirus UL6 like ; Phosducin; Centromere protein H (CENP-H) ; Phage miNor structural protein GP20	No	Cytoplasmic 4.428	2dfsA- 26- 35- 0.006
5687480	No	No	Biogenesis of lysosome-related organelles complex-1 subunit 2 ; Tat binding protein 1(TBP-1)-interacting protein (TBPIP) ; SlyX ; TATA element modulatory factor 1 DNA binding ; V-type ATPase 116kDa subunit family ; Septum formation initiator ; bZIP transcription factor ;	No	Cytoplasmic 3.717	No
5687496	No	Acyl-CoA N-acyltransferase	Acetyltransferase (GNAT) domain	No	Cytoplasmic 3.202	1wwzA- 12- 37- 0.002
5687501	No	Glycoside hydrolase, family 22, conserved site	Flavodoxin	No	Cytoplasmic 3.311	No
5687507	No	No	SPW repeat	No	Cytoplasmic 1.451 InnerMembrane 1.390 Periplasmic 1.180	No

5687508	No	No	No	No	Periplasmic	3.637	2pziA- 7- 38- 0.004
5687517	No	No	NifZ domain	No	Cytoplasmic	2.876	No

Table 31. *Pseudomonas* phage 14-1

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
7095618	Phage_Mu_F super family[cl10072], Phage Mu protein F like protein; COG2369[COG2369], Uncharacterized protein, homolog of phage Mu protein gp30 [Function unkNown]	Phage head morphogenesis domain	Phage Mu protein F like protein	Uncharacterized protein, homolog of phage Mu protein gp30	Cytoplasmic 2.694	No
7095620	No	No	Colicin pore forming domain ; Photosystem II reaction centre N protein (psbN)	No	Cytoplasmic 2.410	1ciiA- 25- 40- 2e-04
7095621	DUF2213 super family[cl15403], Uncharacterized protein conserved in bacteria (DUF2213);	Uncharacterised conserved protein UCP029215	Uncharacterized protein conserved in bacteria (DUF2213) ; BUD22	Uncharacterized BCR	Cytoplasmic 2.914	No
7095623	No	No	AAA ATPase domain	No	OuterMembrane 2.268	No
7095624	No	No	Beta propeller domain	No	Cytoplasmic 1.792 Extracellular 1.387	No
7095628	No	No	Bacteriophage protein of unkNown function (DUF646)	No	Cytoplasmic 2.120 InnerMembrane 1.518	No
7095629	No	Bacteriophage rv5, Orf53	Pre-pro-megakaryocyte potentiating factor precursor (Mesothelin)	No	Extracellular 2.929	No
7095630	No	No	No	No	Cytoplasmic 1.510 Extracellular 1.283 Periplasmic 1.274	1a6zA- 19- 36- 0.002
7095631	No	No	Matrixin	No	Cytoplasmic 1.483 Extracellular 1.257 Periplasmic 1.199	No
7095632	No	No	Met-10+ like-protein	No	Cytoplasmic 3.453	No
7095634	No	No	No	No	Extracellular 1.583 OuterMembrane 1.582	1n8yC- 20- 36- 0.006
7095635	No	No	Mycobacterium membrane protein ; Tetratricopeptide	No	Cytoplasmic 2.194	No

			repeat					
7095638	LT_GEWL[cd00254], Lytic Transglycosylase (LT) and Goose Egg White Lysozyme (GEWL) domain.	Prokaryotic transglycosylase, active site; Lytic transglycosylase-like, catalytic Lysozyme-like domain	Transglycosylase SLT domain ;Colicin pore forming domain ; RTX C-terminal domain	Soluble lytic murein transglycosylase and related regulatory proteins (some contain LysM/invasin domains)	Periplasmic	1.430	No	
7095641	No	No	Bacteriophage Mu Gp45 protein	No	Periplasmic Extracellular	1.849 1.312	No	
7095642	Baseplate_J super family[cl01294], Baseplate J-like protein	No	No	No	Extracellular Periplasmic OuterMembrane	1.784 1.159 1.134	No	
7095643	No	No	No	No	Extracellular OuterMembrane	2.048 1.610	2cndA- 15- 31- 0.007	
7095644	No	No	Beta-lactamase inhibitor (BLIP)	No	Extracellular	3.425	No	
7095646	lysozyme_like super family[cl00222], lysozyme_like domain	Glycoside hydrolase, family 19, catalytic; Lysozyme-like domain;	Chitinase class I	Predicted chitinase	Extracellular Periplasmic	1.904 1.899	2cjlA -19- 137- 2e-33	
7095647	No	No	V-ATPase subunit H	No	Cytoplasmic	3.007	No	
7095648	Adenylation_DNA_ligase_like super family[cl12015], Adenylation domain of proteins similar to ATP-dependent polynucleotide ligases;OBF_DNA_ligase_family super family[cl08424],The Oligonucleotide/oligosaccharide binding (OB)-fold domain is a DNA-binding module that is part of the catalytic core unit of ATP dependent DNA ligases.	DNA ligase, ATP-dependent, central; Nucleic acid-binding, OB-fold;Nucleic acid-binding, OB-fold-like	ATP dependent DNA ligase domain	No	Cytoplasmic	4.686	1fviA- 26- 260- 2e-70	
7095649	NTP-PPase_u3[cd11540], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8)	Phosphoribosyl-ATP pyrophosphohydrolase-like	Phosphoribosyl-ATP pyrophosphohydrolase	No	Cytoplasmic	4.409	1vmgA- 21- 51- 2e-07	

	MazG-like domain found in a group of uncharacterized proteins from bacteria and archaea							
7095650	No	No	TniQ ; Outer membrane protein (OmpH-like); TipAS antibiotic-recognition domain ;	No	Periplasmic	1.540	InnerMembrane 1.210	No
7095651	No	No	No	No	Periplasmic	2.994		1fviA -26 -260- 2e-70
7095652	No	No	V-type ATPase 116kDa subunit family ; Vitamin-D-receptor interacting Mediator subunit 4 ; TSC-22/dip/bun family	No	Cytoplasmic	4.153		No
7095653	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; HepA[COG0553], Superfamily II DNA/RNA helicases, SNF2 family [Transcription / DNA replication, recombination, and repair]	SNF2-related; Helicase, C-terminal;	SNF2 family N-terminal domain ; Helicase conserved C-terminal domain ; Glyoxalase-like domain	Superfamily II DNA/RNA helicases, SNF2 family	Cytoplasmic	4.221		1z6aA- 20- 483- 1e-137
7095654	No	No	VRR-NUC domain	No	Cytoplasmic	3.978		No
7095655	DNA_pol3_alpha super family[cl06665], Bacterial DNA polymerase III alpha subunit;PRK07373 [PRK07373], DNA polymerase III subunit alpha	Bacterial DNA polymerase III, alpha subunit	Bacterial DNA polymerase III alpha subunit ; PHP domain	DNA polymerase III alpha subunit	Cytoplasmic	3.608		No
7095656	DEDDh[cd06127], DEDDh 3'-5' exonuclease domain family;	Exonuclease; Ribonuclease H-like domain; Exonuclease,	Exonuclease	DNA polymerase III epsilon subunit and	Cytoplasmic	4.331		2p1jB- 27- 177- 8e-46

	DEDDh exonucleases, part of the DnaQ-like (or DEDD) exonuclease superfamily	RNase T/DNA polymerase III		related 3'-5' exonucleases			
7095657	Description Pssmid Multi-domain E-value; pseT[PHA02530], polynucleotide kinase	HAD-like domain	HAD superfamily, subfamily IIIB (Acid phosphatase) ; Glyoxalase-like domain ; Ketopantoate reductase PanE/ApbA C terminal	No	Cytoplasmic	3.538	1ltqA- 19- 164- 2e-41
7095658	No	No	Iron Transport-associated domain	No	Cytoplasmic	2.697	No
7095659	Thy1 super family[cl03630], Thymidylate synthase complementing protein	Thymidylate synthase ThyX	Thymidylate synthase complementing protein	Predicted alternative thymidylate synthase	Cytoplasmic	3.873	2af6G- 20- 212- 5e-56
7095660	No	No	YebG protein	No	Cytoplasmic	3.108	No
7095665	No	No	FAD binding domain	No	Cytoplasmic	3.324	2pyxA- 13- 37- 0.009
7095666	No	No	No	No	Cytoplasmic	2.945	1c1gA- 19- 40- 4e-04
7095667	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; UvrD_C_2 super family[cl15862], Family description	No	AAA domain ; Family description	ATP-dependent exoDNase (exonuclease V), alpha subunit - helicase superfamily I member	Cytoplasmic	4.720	1w36D- 16- 177- 2e-45
7095668	No	No	Eukaryotic and archaeal DNA primase small subunit	No	Cytoplasmic	2.451	No
7095670	PADR1 super family[cl06892], PADR1 (NUC008) domain	No	PADR1 (NUC008) domain ; DnaJ central domain; Phage DNA packaging protein Nu1	No	Periplasmic Cytoplasmic	2.588 2.226	No
7095673	No	No	Ca2+ regulator and membrane fusion protein Fig1 ;	No	InnerMembrane Cytoplasmic	1.569 1.326	No
7095674	VirE super family[cl14928], Virulence-associated protein E	Virulence-associated E	Virulence-associated protein E ; Dodecin	No	Cytoplasmic	4.368	1qvrA- 13- 60- 1e-09
7095676	No	No	HrpJ-like domain	No	Cytoplasmic	3.398	No

7095680	No	No	Glycerophosphoryl diester phosphodiesterase	No	Cytoplasmic	3.677	No	
7095687	No	No	Coxiella burnetii protein of unkNown function (DUF762)	No	Cytoplasmic	3.926	No	
7095689	No	No	Hydantoinase B/oxoprolinase ; Shugoshin N-terminal coiled-coil region ; CHD5-like protein ; CHDCT2 (NUC038) domain ; Shugoshin N-terminal coiled-coil region ;	Chromosome segregation ATPases	Cytoplasmic	3.695	1g8xA- 13- 38- 0.002	
7095691	No	No	HemN C-terminal domain; Ribosomal L29 protein	No	Cytoplasmic	2.661	No	
7095695	No	No	NIPSNAP ; SigmaW regulon antibacterial ; Ribosomal protein S27a	No	Cytoplasmic	3.148	No	
7095697	No	No	Transketolase, thiamine diphosphate binding domain	No	Cytoplasmic	3.623	No	
7095701	No	No	DnaJ central domain ;	No	Cytoplasmic	2.785	No	
7095702	No	No	AdoMet_MTases super family[cl16911], S-adenosylmethionine-dependent methyltransferases (SAM or AdoMet-MTase), class I	Methyltransferase domain	No	Cytoplasmic	4.220	3cc8A- 13- 47- 2e-06
7095703	No	No	Mediator complex subunit 30 ; Tc3 transposase ; Nuclear pore complex subunit Nro1	No	Cytoplasmic	3.794	No	
7095704	No	No	Flagellar P-ring protein	No	Cytoplasmic	2.373	No	

Table 32. *Pseudomonas* phage phi15

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
10322096	No	No	Meiotic cell cortex C-terminal pleckstrin homology	No	Cytoplasmic 1.591 Extracellular 1.525 Periplasmic 1.072	No
10322098	No	No	Beta-galactosidase jelly roll domain ; Yip1 domain	No	Cytoplasmic 1.695 Periplasmic 1.125 Extracellular 1.122	No
10322105	No	No	SlyX ; WHEP-TRS domain	No	Cytoplasmic 3.196	No
10322120	No	No	Mrr N-terminal domain	No	OuterMembrane 1.549 Extracellular 1.368	No
10322122	PIN_SF super family[cl14812], PIN (PiIT N terminus) domain: Superfamily;PHA00439[PHA00439], exonuclease	5'-3' exonuclease, alpha-helical arch, N-terminal	5'-3' exonuclease, N-terminal resolvase-like domain ; 5'-3' exonuclease, C-terminal SAM fold	5'-3' exonuclease (including N-terminal domain of Poll)	Cytoplasmic 4.596	1bgxT- 16 -213- 2e-56
10322123	No	Bacteriophage T7-like gene 6.5	No	No	Cytoplasmic 2.454 Periplasmic 1.989	No

Table 33. *Pseudomonas* phage KPP10

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
10358671	No	Bacteriophage D3, Orf41.6	No	No	Periplasmic 2.075 Cytoplasmic 1.911	1pkyC- 18-36- 0.008
10358675	No	No	Bacterial protein of unkNown function (DUF883) ; T4 recombination endonuclease VII, dimerisation Uncharacterised protein family (UPF0150)	No	Cytoplasmic 2.150 Periplasmic 1.708	No
10358688	No	No	Tropomyosin like	No	Cytoplasmic 3.531	No
10358689	No	No	Centrosome microtubule-binding domain of Cep57	No	Cytoplasmic 3.936	1vh1D 16 38 0.004
10358692	Ead_Ea22 super family[cl16487], Ead/Ea22-like protein;	Ead/Ea22-like protein	Ead/Ea22-like protein ; bZIP transcription factor ; Homeobox associated leucine zipper	No	Cytoplasmic 4.584	No
10358693	No	Bacteriophage D3, Orf41.6	No	No	Periplasmic 2.075 Cytoplasmic 1.911	1pkyC 18 36 0.008
10358699	N_methyl_2 super family[cl06830], Type IV pilin N-term methylation site GFxxxE; PulG[COG2165], Type II secretory pathway, pseudopilin PulG [Cell motility and secretion / Intracellular trafficking and secretion]	Prepilin-type cleavage/methylat ion, N-terminal	Type IV pilin N-term methylation site GFxxxE ; Type II secretory pathway pseudopilin	General secretory pathway proteins G and H and related periplasmic/se creted proteins	Cytoplasmic 2.240	1oqwA- 66- 45- 4e-06
10358703	aroK[PRK00131], shikimate kinase	No	Proteasome subunit	No	Cytoplasmic 3.116	1e94A- 23- 48-1e-06
10358708	DNA_processg_A[pfam02481], DNA recombination-mediator protein A;	No	DNA recombination-mediator protein A ; Glycosyl transferase family, a/b domain	No	Periplasmic 2.053 Cytoplasmic 1.728	1rcuA- 25- 36- 0.010
10358709	No	No	DinB superfamily ; Small subunit of acetolactate synthase	No	Cytoplasmic 3.284	No

			; Low molecular weight S layer protein N terminal				
10358717	No	No	Geminivirus V1 protein	No	Cytoplasmic Periplasmic	2.559 1.929	No
10358719	No	No	Defensin propeptide	No	Periplasmic	3.275	No
10358728	HDPD super family[cl03784], Bacteriophage lambda head decoration protein D	No	Bacteriophage lambda head decoration protein D ; AAA domain (Cdc48 subfamily)	No	Cytoplasmic	3.259	No
10358730	No	No	InsA C-terminal domain	No	Cytoplasmic	2.853	No
10358739	No	No	Glutaredoxin ; CorA-like Mg2+ transporter protein ; Occludin homology domain ;	No	Cytoplasmic	3.156	No
10358741	Ku_bact[TIGR0272], Ku protein, prokaryotic	No	No	No	Cytoplasmic Periplasmic OuterMembrane	1.788 1.430 1.137	No
10358747	No	No	Arb2 domain	No	Cytoplasmic OuterMembrane	2.220 1.530	No
10358749	No	No	Phosphoprotein P region PCT disordered	No	Cytoplasmic	2.888	No
10358752	No	No	holin, BlyA family	No	InnerMembrane Cytoplasmic	2.261 1.605	No
10358754	No	No	Baculovirus polyhedron envelope protein, PEP, C terminus ; Transposase	No	Cytoplasmic Periplasmic	2.053 1.441	No
10358756	No	No	No	No	Periplasmic	2.104	No
10358757	No	No	BlaR1 peptidase M56 ; WW domain-binding protein 1	No	Cytoplasmic	2.257	No
10358758	No	No	Herpes virus U44 protein	No	Cytoplasmic	2.824	No
10358759	No	No	Acetyltransferase (GNAT) domain	No	Cytoplasmic	4.100	No
10358760	No	No	Tripartite ATP-independent periplasmic transporters, DctQ component	No	Cytoplasmic Periplasmic	1.959 1.583	No
10358761	Band_7 super family[cl02525], The band 7	Band 7 protein	SPFH domain / Band 7 family	Membrane protease subunits,	Cytoplasmic	2.807	3bk6A- 19- 119- 4e-28

	domain of flotillin (reggie) like proteins			stomatin/proh ibitin homologs			
10358767	No	No	Acyl-CoA thioesterase	No	Cytoplasmic	4.121	No
10358772	No	No	Sporulation related domain	No	Cytoplasmic	2.738	No
10358773	RPA_2b- aaRSs_OBF_like super family[cl09930], Replication protein A, class 2b amiNoacyl-tRNA synthetases, and related proteins with oligonucleotide/ol igosaccharide (OB) fold	No	BAG domain	No	Periplasmic	3.156	No
10358774	No	No	Catalase-related immune-responsive ; Toluene tolerance, Ttg2	No	Cytoplasmic	3.053	No
10358776	No	No	N-terminal domain of NEFA-interacting nuclear protein NIP30	No	Cytoplasmic	2.894	No
10358780	No	No	HNH endonuclease ; Platelet-derived growth factor, N terminal region ; Rho termination factor, N- terminal domain	No	Cytoplasmic	4.528	No
10358783	No	No	Tash protein PEST motif	No	Cytoplasmic	3.042	No
10358791	No	No	RNA polymerase III transcription factor (TF)IIIC subunit	No	Cytoplasmic	3.042	No
10358792	No	No	Molydopterin dinucleotide binding domain	No	Cytoplasmic	3.418	No
10358797	No	No	Huntingtin protein region	No	Cytoplasmic	2.707	No
10358800	No	No	Stc1 domain ; Hydrogenase expression/synthesis hypA family ; Double zinc ribbon	No	Cytoplasmic	2.814	No
10358802	No	No	No	Phosphoglycer omutase	Cytoplasmic	2.652	No
10358804	No	No	Herpesvirus pp38	No	Periplasmic	3.053	No

			phosphoprotein				
10358805	No	No	NTF2-like N-terminal transpeptidase domain	No	Cytoplasmic	3.983	No
10358806	No	No	Autophagy protein Apg5	No	Cytoplasmic Periplasmic	2.278 1.815	No
10358808	No	No	Repeat of unkNown function (DUF1079)	No	Cytoplasmic	2.956	No
10358811	No	No	Methyl-accepting chemotaxis protein	No	Cytoplasmic	3.060	No

Table 34. *Pseudomonas* phage phiIBB-PF7A

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
10323784	No	No	Glycosyl hydrolase family 67 middle domain	No	Periplasmic 2.411	No
10323786	P4Hc[smart00702], Prolyl 4-hydroxylase alpha subunit homologues	Oxoglutarate/iron-dependent dioxygenase; Prolyl 4-hydroxylase, alpha subunit	2OG-Fe(II) oxygenase superfamily	No	Cytoplasmic 3.373	2hbtA- 13- 42- 1e-04
10323794	No	No	Integrin, beta chain ; Influenza C hemagglutinin stalk ; SlyX	No	Cytoplasmic 2.691	No
10323795	No	No	No	No	Cytoplasmic 4.301	1dekA- 17 -82- 1e-16
10323801	PolyA_pol super family[cl14657], Poly A polymerase head domain	No	Poly A polymerase head domain	tRNA nucleotidyltransferase/poly(A) polymerase	Cytoplasmic 4.182	1vfgB- 17 -85- 7e-18
10323803	No	Small-subunit processome, Utp21	Utp21 specific WD40 associated putative domain ; ssRNA positive strand viral 18kD cysteine rich protein ; Uncharacterized protein conserved in archaea (DUF2240)	No	Cytoplasmic 2.615	No
10323806	No	No	CIDE-N domain	No	Cytoplasmic 3.786	No
10323807	No	No	FadR C-terminal domain	No	Cytoplasmic 2.315	No
10323809	No	Bacteriophage T7-like gene 6.5	No	No	Cytoplasmic 3.237	No
10323810	No	Bacteriophage T7-like, gene 6.7	No	No	Periplasmic 2.290	No
10323828	No	No	Putative transmembrane protein 170	No	Cytoplasmic 1.880 InnerMembrane 1.269	No
10323829	No	No	FACT complex subunit (SPT16/CDC68)	No	Periplasmic 2.557	No
10323830	No	No	SHNi-TPR	No	Cytoplasmic 2.787	No
10323833	No	No	RimM N-terminal domain	No	Cytoplasmic 2.337 Periplasmic 1.693	No

10323834	No	No	Glycosyl hydrolase family 67 middle domain	No	Periplasmic	2.411	No
10323835	No	No	RimM N-terminal domain	No	Cytoplasmic Periplasmic	2.337 1.693	No

Table 35. *Pseudomonas* phage PAK_P1

NCBI GENE ID	CDD-Blast	Interproscan	pfam	COGs	CELLO v 2.5	PS2 structure
10351230	No	No	Borna disease virus P10 protein	No	Cytoplasmic 4.354	No
10351233	No	No	NifZ domain	No	Cytoplasmic 2.510	No
10351234	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases	No	No	No	Cytoplasmic 2.888	No
10351237	DNA_processg_A[pfam02481], DNA recombination-mediator protein A	No	DNA recombination-mediator protein A	Predicted Rossmann fold nucleotide-binding protein involved in DNA uptake	Cytoplasmic 2.224	1rcuA- 27- 39- 0.001
10351239	No	No	Valyl tRNA synthetase tRNA binding arm	No	Cytoplasmic 2.215	No
10351240	Macro_Poa1p_like [cd02901], Macro domain, Poa1p_like family	Appr-1-p processing	Macro domain; Tetratricopeptide repeat	Uncharacterized ACR related to the C-terminal domain of histone macroH2A1	Cytoplasmic 2.909	2fg1A -28- 134- 9e-33
10351243	PLN03130[PLN03130], ABC transporter C family member	No	No	No	Cytoplasmic 2.926	No
10351245	P-loop_NTPase super family[cl09099], P-loop containing Nucleoside Triphosphate Hydrolases; AAA_33[pfam13671], AAA domain	No	AAA domain	No	Cytoplasmic 4.357	1yj5B- 16- 55- 6e-09
10351247	RNA_lig_T4_1 super family[cl09743], RNA ligase	RNA ligase, T4 RnIA-like	RNA ligase	No	Cytoplasmic 4.775	2c5uA- 20- 213- 4e-56
10351248	No	No	IncA protein	No	InnerMembrane 2.978	No

10351249	HTH_XRE super family[cl15761], Helix-turn-helix XRE-family like proteins	Lambda repressor-like, DNA-binding domain	Helix-turn-helix domain ; Hepatocyte nuclear factor 1 (HNF-1), N terminus ; Kinesin protein	No	Cytoplasmic	3.930	2a6cA- 36- 80- 1e-16
10351251	MPP_superfamily super family[cl13995], metallophosphatase superfamily, metallophosphatase domain	Calcineurin-like phosphoesterase superfamily domain	Calcineurin-like phosphoesterase superfamily domain	No	Cytoplasmic	3.485	1xm7A- 22- 120- 2e-28
10351253	No	HD/PDEase domain	HD domain	No	Cytoplasmic	3.871	1vj7A- 19 -159- 3e-40
10351254	Hydrolase_2 super family[cl06494], Cell Wall Hydrolase	Cell wall hydrolase, SleB	Cell Wall Hydrolase	No	Cytoplasmic	2.608	No
10351259	cytidine_deaminase-like super family[cl00269], Cytidine and deoxycytidylate deaminase zinc-binding region.	CMP/dCMP deaminase, zinc-binding; Cytidine deaminase; Cytidine deaminase-like;	Cytidine and deoxycytidylate deaminase zinc-binding region	Deoxycytidylate deaminase	Cytoplasmic	4.477	2hvwA- 29- 146- 1e-36
10351260	No	No	Mitochondrial ribosomal subunit S27	No	Cytoplasmic	3.004	No
10351262	No	No	Histone lysine methyltransferase SET associated	No	Periplasmic	2.590	No
10351263	No	No	No	No	Cytoplasmic	2.759	2cfmA- 44 -35- 0.004
10351265	No	No	P-11 zinc finger	No	Cytoplasmic	2.668	No
10351266	No	No	Interferon-related protein conserved region ; Major surface glycoprotein 2 C terminal	No	Cytoplasmic	2.734	No
10351268	Clp_protease_like super family[cl15763], Caseinolytic protease (ClpP) is an ATP-dependent protease	Peptidase S14, ClpP	Clp protease	Protease subunit of ATP-dependent Clp proteases	Cytoplasmic	4.326	1y7oB- 18- 147- 7e-37
10351277	No	No	AAA domain	No	Cytoplasmic	4.161	No
10351278	No	No	Bacteriophage lambda head decoration protein D ; Pyridoxal-phosphate dependent enzyme	No	Cytoplasmic	3.245	No

10351284	No	No	Bacteriophage rv5, Orf53	No	Cytoplasmic	2.228	No
10351286	No	No	No	Phosphoenolpyruvate synthase/pyruvate phosphate dikinase	Cytoplasmic InnerMembrane	1.476 1.472	No
10351290	tape_meas_nterm super family[cl15680], tape measure domain;COG3941[COG3941], Mu-like prophage protein [General function prediction only]	Caudovirus, tape measure, N-terminal	Phage related protein	No	InnerMembrane	2.367	No
10351291	No	No No	Sporulation inhibitor A	No	OuterMembrane	2.407	No
10351293	No	No	No	Polyphosphate kinase	Cytoplasmic	3.567	No
10351294	No	No	A short region before toxin nuclease domain	No	OuterMembrane	4.163	No
10351295	gpV[COG4540], Phage P2 baseplate assembly protein gpV [General function prediction only]	No	Bacteriophage Mu Gp45 protein	No	Extracellular	2.577	No
10351299	No	No	No	No	Extracellular	3.647	1k6dB- 21- 32- 0.003
10351301	34[PHA02584], long tail fiber, proximal subunit	No	Nucleic acid binding protein NABP	No	Extracellular	3.502	1v4sA- 19- 30- 0.010
10351302	No	No	Transglycosylase SLT domain; ManNosylglycoprotein endo-beta-N-acetylglucosaminidase	No	Periplasmic	3.198	No
10351303	No	No	Cation efflux family	Gamma-aminobutyrate permease and related permeases	Cytoplasmic	2.582	No
10351304	Filament[pfam00038], Intermediate filament protein	No	HemX; Intermediate filament protein	No	Periplasmic	3.042	1gk4A- 40- 37 -0.001
10351305	No	No	Type-F conjugative transfer system protein (TrbI_Ftype); Flagella accessory	No	Cytoplasmic InnerMembrane	1.571 1.505	No

			protein C (FlaC)				
10351306	No	No	No	No	Cytoplasmic	3.414	No
	NTP-PPase_u5[cd11542], Nucleoside Triphosphate Pyrophosphohydrolase (EC 3.6.1.8)						
10351307		No	No	No	Cytoplasmic	4.948	2q73B- 20- 44- 6e-05
	MazG-like domain found in a group of uncharacterized proteins from bacteria and archaea						
10351308	No	Zinc finger, LIM-type	MYM-type Zinc finger with FCS sequence motif ; Transposase zinc-ribbon domain ; LIM domain	Uncharacterized BCR	Cytoplasmic	2.650	No
	Adenylation_DNA_ligase_like super family[cl12015], Adenylation domain of proteins similar to ATP-dependent polynucleotide ligases; PHA02142[PHA02142], putative RNA ligase						
10351309		RNA ligase, DRB0094; RNA ligase domain, REL/Rln2	RNA ligase	No	Cytoplasmic	3.297	2hvqA- 13- 51- 4e-07
10351315	No	No	No	No	Cytoplasmic Periplasmic	2.252 1.707	1jjeB -20 -35- 0.009
	PhdYefM_antitox super family[cl09153], Antitoxin Phd_YefM, type II toxin-antitoxin system						
10351316		No	No	No	Cytoplasmic	4.645	2a6qA- 48- 37- 0.003
10351321	No	No	KOW motif	No	Cytoplasmic	3.566	No
10351328	No	No	Chorion family 3	No	Cytoplasmic	2.750	No
	Tetraspanin family ; Transmembrane Fragile-X-F protein ; NADH dehydrogenase subunit 5 C-terminus ; PepSY-associated TM helix ; African swine fever virus J13L protein						
10351329	No	No		No	InnerMembrane	3.112	No
10351331	No	No	Poxvirus A28 family	No	Cytoplasmic	4.503	1bgxT -15- 111- 1e-25

10351332	HNHc super family[cl00083], HNH nucleases	No	SLA1 homology domain 1, SHD1 ; prokaryotic dksA/traR C4-type zinc finger; HNH endonuclease ; Anti-repressor SinI	No	Cytoplasmic	4.172	No
10351336	No	No	Tetratricopeptide repeat ; DASH complex subunit Dam1	No	Cytoplasmic	2.989	No
10351337	No	Metallophosphoesterase domain	Calcineurin-like phosphoesterase	No	Cytoplasmic	3.497	1uteA- 13- 36 -0.010
10351341	No	No	No	No	Cytoplasmic	3.415	1h4vB -30- 37 -0.006
10351342	Thy1 super family[cl03630], Thymidylate synthase complementing protein	Thymidylate synthase ThyX	Thymidylate complementing protein	Predicted alternative thymidylate synthase	Cytoplasmic	3.444	2af6G- 20 -103- 3e-23
10351343	No	No	Transposase-associated domain; Putative zinc-finger		Cytoplasmic	4.154	No
10351349	GAT_1 super family[cl00020], Type 1 glutamine amidotransferase (GATase1)-like domain	No	No	No	Cytoplasmic	3.662	No
10351350	No	No	Double zinc ribbon	No	Cytoplasmic	2.552	No
10351351	No	No	Cyanate lyase C-terminal domain	No	Cytoplasmic	3.955	No
10351353	No	No	No	Membrane GTPase LepA	Cytoplasmic Periplasmic	1.937 1.849	No
10351356	No	No	Acyl-CoA reductase (LuxC)	No	Periplasmic	2.222	No
10351357	No	No	Flavocytochrome c sulphide dehydrogenase, flavin-binding	No	Cytoplasmic	3.178	No
10351358	SprT super family[cl01182], SprT homologues.	No	SprT-like family	No	Cytoplasmic	3.023	No
10351359	No	No	Coronavirus NS2A protein	No	Cytoplasmic	2.365	No
10351367	No	No	AP endonuclease family 2 C terminus	No	Periplasmic Cytoplasmic Extracellular	1.652 1.333 1.189	No
10351370	No	No	Clr5 domain	No	Cytoplasmic	3.039	No

10351379	No	No	RecX family	No	Cytoplasmic	3.543	No
10351382	No	No	Variant SH3 domain	Transketolase	Cytoplasmic OuterMembrane	1.871 1.173	No
10351383	No	No	T4 recombination endonuclease VII, dimerisation ; Uncharacterised protein family (UPF0150)	No	Cytoplasmic Periplasmic	2.150 1.708	No

ABOUT AUTHORS



Krishna Khairnar: Dr. Krishna Khairnar is Scientist at Environmental Virology Cell, Council of Scientific and Industrial Research - National Environmental Engineering Research Institute (CSIR-NEERI). He received his PhD in Medical Microbiology from Jawaharlal Institute of Postgraduate Medical Education & Research, Pondicherry and completed his Post doctorate in Molecular Medical Microbiology from University of Toronto, Toronto, Canada. Dr. Khairnar was a research scientist and an ID consultant at Ministry of Health and Long Term Care, Toronto, Canada and specializes in advanced infectious diseases research and molecular diagnostics of parasitic diseases. He is one of the leading research experts in the role of molecular diagnostics research development and infectious diseases. Dr. Khairnar has excellent research publications in peer-reviewed international journal in the related field with outstanding impact factor of ~82, H-index of 10, i10-index of 10 and citations of 283. As a result of outstanding research contribution, publications and expertise in the field, at the age of only 34 years Dr. Khairnar is bestowed with the honour of being selected as a chief editor of a reputed SCI journal Biomed Research International which has a Thomson Reuters Impact factor of 3. Dr. Khairnar also serves on the reviewer board for a number of SCI journals including BMC series and Experimental Parasitology.

Rajshree Chandekar: Ms. Rajshree Chandekar pursuing PhD. at Environmental Virology Cell, Council of Scientific and Industrial Research - National Environmental Engineering Research Institute (CSIR-NEERI) under the guidance of Dr. Krishna Khairnar. She has completed her Master degree in Biotechnology in 2010 from Rastrasant Tukadoji Maharaj Nagpur University, Nagpur. Her field of interest is Virology, Microbiology, Bioinformatics and Molecular biology. She has good research publications in peer-reviewed SCI journal in the related field.



Swapnil Sanmukh: Mr. Swapnil Sanmukh pursuing Ph. D. in Biotechnology at Council of Scientific and Industrial Research - National

Comparative Genomic studies for understanding the



Environmental Engineering Research Institute (CSIR-NEERI). He has completed his Master degree in Environmental Biotechnology in 2009 from Shivaji University, Kolhapur. His field of interest is Virology, Microbiology, Bioinformatics and Molecular biology. He has good research publications in peer-reviewed SCI journal in the related field.



Waman Paunikar: Dr. Waman Paunikar is Head and Principal Scientist at Environmental Virology Cell, Council of Scientific and Industrial Research - National Environmental Engineering Research Institute (CSIR-NEERI). He received his PhD in Zoology Rastrasant Tukadoji Maharaj Nagpur University, Nagpur. He is presently working in Zoology (Cell biology), Environmental Virology, Environmental Bacteriology, Environmental Microbiology, Environmental Virology, Environmental Biotechnology area. Dr. Paunikar has excellent research publications in peer-reviewed international journal in the related field with outstanding impact factor.