

**Part 1:** **TITLE, AUTHORS, APPROVALS, etc**

|  |  |  |
| --- | --- | --- |
| **Code assigned:** | ***2022.018B*** |  |
| **Short title:** Create a new genus in the *Caudoviricetes* class containing *Klebsiella* phage  vB\_KpnS-Carvaje | | |
|  | | |

**Author(s) and email address(es)**

|  |  |
| --- | --- |
| Sousa JC, Sillankorva S, Faustino A, Carvalho CM | jessica.sousa@inl.int; sanna.sillankorva@inl.int; albertafaustino@gmail.com; carla.carvalho@inl.int |

**Author(s) institutional address(es) (optional)**

|  |
| --- |
| International Iberian Nanotechnology Laboratory (INL) [SJC, SS, CCM]  Instituto de Engenharia de Sistemas e Computadores – Microsistemas e Nanotecnologias (INESC MN) [SJC]  Faculdade de Engenharia da Universidade do Porto (FEUP) [SJC]  Departmento de Patologia Clínica, Hospital de Braga [FA] |

**Corresponding author**

|  |
| --- |
| carla.carvalho@inl.int [CCM] |

**List the ICTV Study Group(s) that have seen this proposal**

|  |
| --- |
| Caudoviricetes Study Group; Bacterial Viruses Subcommittee |

**ICTV Study Group comments and response of proposer**

|  |
| --- |
|  |

**ICTV Study Group votes on proposal**

|  |  |  |  |
| --- | --- | --- | --- |
| **Study Group** | **Number of members** | | |
| **Votes support** | **Votes against** | **No vote** |
|  |  |  |  |
|  |  |  |  |

**Authority to use the name of a living person**

|  |  |
| --- | --- |
| **Is any taxon name used here derived from that of a living person (Y/N)** | N |

|  |  |  |
| --- | --- | --- |
| **Taxon name** | **Person from whom the name is derived** | **Permission attached (Y/N)** |
|  |  |  |
|  |  |  |

**Submission dates**

|  |  |
| --- | --- |
| Date first submitted to SC Chair | April 2022 |
| Date of this revision (if different to above) |  |

**ICTV-EC comments and response of the proposer**

|  |
| --- |
|  |

**Part 3:** **TAXONOMIC PROPOSAL**

**Name of accompanying Excel module**

|  |
| --- |
| 2022.018B.N.v1.Carvajevirus\_ng.xls |

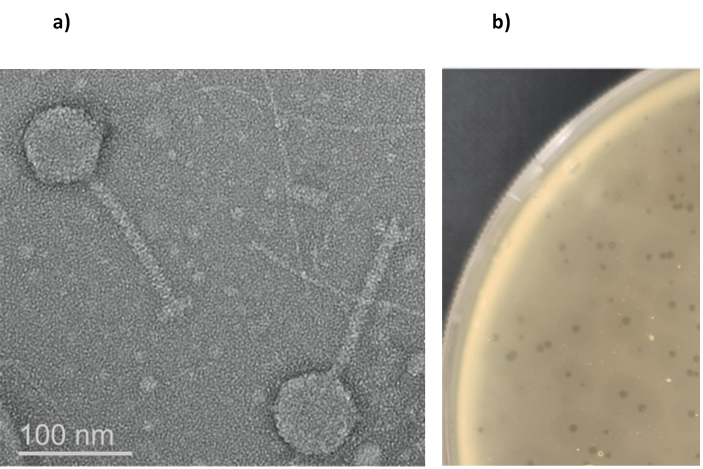
**Abstract**

|  |
| --- |
| A novel *Klebsiella pneumoniae* phage (vB\_KpnS-Carvaje) has been isolated and characterized. Its genome length is approximately 57 kb presenting 99 open reading frames (ORFs), 33 of which have assigned functions while 66 are unknown. This phage differs from other sequenced *Klebsiella* phages, showing the closest resemblance (up to 65.32%) with Salmonellaphages belonging to the *Nonanavirus* and *Sashavirus* genera*.* Comparisons at the amino acid level and phylogeny analysis among homologous genomes indicate that the Klebsiella phage Carvaje forms a novel sister taxon within the node of the Nonanaviruses and Sashaviruses cluster. Due to the unique genomic and proteomic features of Carvaje phage, we suggest the establishment of a new genus within the *Caudoviricetes* class. The genome sequence of Carvaje was deposited in GenBank under the accession number OL604152. |

**Text of proposal**

|  |  |
| --- | --- |
| |  | | --- | | **Species demarcation criteria:** Two phages are assigned to the same species if their genomes are more than 95% identical over their genome length for isolates.  These values can be calculated by a number of tools, such as BLASTn – usually calculated using intergenomic distance calculator VIRIDIC [1].  **Genus demarcation criteria:** In search for criteria that create cohesive and distinct genera that are reproducible and monophyletic, the Bacterial Viruses Subcommittee has established 70% nucleotide identity of the genome length as the cut-off for genera. Genus-level groupings should always be monophyletic in the signature genes, as tested with a phylogenetic tree [2].  Phage Carvaje has a siphovirus-like morphology, according to the former ICTV classification with a 64 nm × 65 nm icosahedral capsid, a 146 nm contractile tail, and a brushy baseplate at the distal tip (Figure 1.a). Moreover, as for plaque morphology, upon overnight incubation, it forms uniform plaques with diameters of approximately 0.7 mm and a surrounding larger halo of about 2.7 mm (Figure 1.b).  Bioinformatic analysis of the Carvaje phage revealed that it presents a linear, double-stranded DNA genome with 56858 bp. Its genomic analysis allowed the identification of genes’ putative functions. In total, 99 putative ORFs were predicted, only 33 have assigned putative functions, and 66 ORFs were classified as hypothetical proteins of unknown function. Also, one tRNA was retrieved using myRAST annotation pipeline (Figure 2) [3].  BLASTn analysis [4] of the Carvaje genome revealed a higher nucleotide identity to the complete genomes of *Salmonella* phages 9NA (accession number: KJ802832.1), vB\_SenS\_Sasha (accession number: NC\_047786.1), Solent (accession number: MH586730.1), and vB\_SenS\_Sergei (accession number: KY002061.1). These Salmonella phages belong to two genera: the Nonanaviruses comprising the complete genome of Salmonella phage 9NA and the partial genome sequences of phages FLS SP-069 (SP069) and SP-062 (SP062); and the Sashaviruses including the Sasha, Sergei, and Solent phages.  The whole-genome alignment of Carvaje’s genome sequence with the phages that present higher homology (9NA, Sasha, Sergei, and Solent) using EasyFig [5] shows that Carvaje has a higher pairwise identity with the 9NA Salmonella phage (Figure 3). Also, the whole-genome alignments using MAUVE [4] displayed homologies ranging from 62.10% (Solent phage) up to 65.32% (9NA phage).  A phylogenomic tree was inferred using the 22 phage genomes with the highest homology to Carvaje phage retrieved from the BLASTn analysis. Phylogenomic tree was visualized on iTOL software and inferred using the formula d0 on VICTOR [7, 8]. The analysis revealed that Carvaje was included in the same node with the six previously mentioned phages (9NA, Sasha, Sergei, Solent, SP069 and SP062) (Figure 4).  OrthoVenn [9] comparative analysis of Carvaje and its four closest homologs, using an E-value of 1e-5 and the default inflation value of 1.5, resulted in 99 clusters, of which 49 were orthologous clusters between the five phage genomes, and 46 single-copy gene clusters, with Carvaje yielding 37 of these singletons. A total of 62 orthologous gene clusters were identified among Carvaje and the *Salmonella* most similar phages (9NA, Sasha, Sergei, and Solent). As pointed above, 9NA belongs to a different genus than Sasha, Solent and Sergei but still shares 60 clusters with these viruses, whereas Carvaje shares only 51 clusters with the latter and 58 clusters in common with 9NA (Figure 5).  Despite homology with Salmonella phages, Carvaje showed to be specific for K. pneumoniae, not infecting any of the other bacterial species tested (including Salmonellastrains). Thus, Carvaje is a novel virulent phage that, to the best of our knowledge, is the first K. pneumoniae phage with similarity with the *Salmonella* Nonanaviruses and Sashaviruses, suggesting that it has a common ancestor to the phages belonging to these genera. Nevertheless, despite this homology, the similarity of Carvaje with any other known phages isolated to date is still low, and, as so, we suggest that Carvaje should belong to another sister genus.  These results are presented in the manuscript [10], which has been accepted at Current Genetics, Springer (Q1). | |

**Supporting evidence**



**Figure 1 –** a) TEM micrograph of Carvaje phage particle (scale bar 100 nm); b) Plaque morphology of Carvaje. Retrieved from: [10]

C:\Users\jsousa50398\OneDrive - INL\Documents\PhD Thesis\paper Phage E\Figure 4.tif

**Figure 2 –** Map of the genome organization of phage Carvaje created by BRIG platform and CGView. The ORFs with predicted annotations are indicated with blue arrows (CDS – Coding Sequence). The inner rings show genome location, GC skew + (green), GC skew − (purple) and GC content (black). Encoded tRNA is represented in yellow. Hypothetical proteins are represented by hp. Retrieved from: [10]

C:\Users\jsousa50398\OneDrive - INL\Documents\PhD Thesis\paper Phage E\figures high reolution\finals\Figure 4.tif

**Figure 3 –** Schematic genomic alignment at the nucleotide level of phage Carvaje with most similar phages, generated by EasyFig program. Arrows represent ORFs (pink arrows are associated with hypothetical proteins of unknown function, while orange arrows are ORFs that encode proteins with specific functions). Gradient of grey bars indicate the identity percentage between connected homologous regions, darker grey means a higher similarity. Retrieved from: [10]

C:\Users\jsousa50398\OneDrive - INL\Documents\PhD Thesis\paper Phage E\figures high reolution\finals\Figure 5.tif

**Figure 4 –** Phylogenomic tree visualized on iTOL software and inferred using the formula d0 on VICTOR. Branches’ lengths reflect pseudo-bootstrap support values from 100 replications and are scaled in terms of the respective distance formula used. The names of the phages and respective host are given at the tip of each branch. Color ranges represent phages’ genera and colored circles are regarded as respective family. Retrieved from: [10]

C:\Users\jsousa50398\OneDrive - INL\Documents\PhD Thesis\paper Phage E\figures high reolution\finals\Figure 6_1200 dpi.tif

**Figure 5 –** OrthoVenn comparison of Carvaje gene products with those of its close related phage genomes deposited at the NCBI. Venn diagram showing the number of orthologous shared clusters. Below the Venn diagram is the number of clusters of each phage and the number of clusters shared by 5, 4, 3, and 2 phages. Retrieved from: [10]

**References**

1. Moraru C, Varsani A, Kropinski AM. VIRIDIC-A Novel Tool to Calculate the Intergenomic Similarities of Prokaryote-Infecting Viruses. Viruses. 2020 Nov 6;12(11):1268. doi: 10.3390/v12111268. PMID: 33172115.
2. Turner D, Kropinski AM, Adriaenssens EM. A Roadmap for Genome-Based Phage Taxonomy. Viruses. 2021 Mar 18;13(3):506. doi: 10.3390/v13030506. PMID: 33803862.
3. Aziz RK, Bartels D, Best A, et al (2008) The RAST Server: Rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75
4. Altschup SF, Gish W, Miller W, et al (1990) Basic Local Alignment Search Tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2
5. Sullivan MJ, Petty NK, Beatson SA (2011) Easyfig: a genome comparison visualizer. Bioinformatics 27:1009–1010. https://doi.org/10.1093/bioinformatics/btr039
6. Darling ACE, Mau B, Blattner FR, Perna NT (2004) Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. Genome Res 14:1394–1403. https://doi.org/10.1101/gr.2289704.tion
7. Ciccarelli FD, Doerks T, Creevey CJ, et al (2006) Toward Automatic Reconstruction of a Highly Resolved Tree of Life. Science (80- ) 311:1283–1287. https://doi.org/10.1126/science.1121745
8. Meier-Kolthoff JP, Göker M (2017) VICTOR: genome-based phylogeny and classification of prokaryotic viruses. Bioinformatics 33:3396–3404. https://doi.org/10.1093/bioinformatics/btx440
9. Xu L, Dong Z, Fang L, et al (2019) OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. Nucleic Acids Res 47:W52–W58. https://doi.org/10.1093/nar/gkz333
10. Sousa JC, Sillankorva S, Faustinho A, Carvalho CM (2022) Suggestion for a new bacteriophage genus for a Klebsiella pneumoniae phage. Curr Genet. https://doi.org/10.1007/s00294-022-01242-2