Genomic Nucleotide-Based Distance Analysis for Delimiting Old World Monkey Derived Herpes Simplex Viruses Species.

Aaron W Kolb (awkolb@wisc.edu)
University of Wisconsin Madison

Curtis R Brandt
University of Wisconsin Madison

Research article

Keywords:

Posted Date: October 8th, 2019

DOI: https://doi.org/10.21203/rs.2.15777/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published on June 26th, 2020. See the published version at https://doi.org/10.1186/s12864-020-06847-w.
Abstract

Herpes simplex viruses form a genus within the alphaherpesvirus subfamily, with three identified viral species isolated from Old World monkeys (OWM); Macacine alphaherpesvirus 1 (herpes B), Cercopithecine alphaherpesvirus 2 (SA8), and Papiine alphaherpesvirus 2 (PaHV-2; herpes papio). Herpes B is endemic to macaques, while PaHV-2 and SA8 appear endemic to baboons. All three viruses are genetically and antigenically similar, with SA8 and PaHV-2 thought to be avirulent in humans, while herpes B is a biosafety level 4 pathogen. Recently, next-generation sequencing (NGS) has resulted in an increased number of published OWM herpes simplex genomes, allowing an encompassing phylogenetic analysis. In this study, phylogenetic networks, in conjunction with a genome-based genetic distance cutoff method were used to examine 27 OWM monkey herpes simplex isolates. Genome-based genetic distances were calculated, resulting in distances between Lion and Pig-tailed simplex viruses themselves, and versus herpes B core strains that were higher than those between PaHV-2 and SA8 (approximately 14% and 10% respectively). The species distance cutoff was determined to be 8.94%, with the method recovering separate species status for PaHV-2 and SA8 and showed that Lion and Pig-tailed simplex viruses (vs core herpes B strains) were well over the distance species cutoff. In conclusion, we propose designating Lion and Pig-tailed simplex viruses as separate, individual viral species, and that this may be the first identification of viral cryptic species.

Background

The alphaherpesvirinae comprise a subfamily within Herpesviridae, with most of its members establishing latency in the peripheral nervous system. The five genera which comprise the alphaherpesvirinae infect birds (Iltovirus, Mardivirus), sea turtles (Scutavirus), mammals (Varicellovirus, Simplexvirus), as well as lizards (currently unassigned). Until fairly recently, simplexviruses were thought to only infect primates, however simplex viruses have been isolated from cattle, bats, rabbits, and marsupials [1-5]. Various species of macaque monkeys are the natural reservoir for the herpes B simplexvirus. Herpes B was first described in 1933, following an incident where a 29-year-old laboratory worker was bitten by an asymptomatic monkey and later died from encephalitis [6, 7]. Herpes B has been demonstrated to be highly neurovirulent with ~80% mortality and is categorized as a BSL-4 level pathogen by the CDC [8, 9]. In spite of considerable work with macaques in laboratory settings, as well as close contact between humans and macaques particularly in Asia, there have only been 46 documented cases of zoonotic transmission since 1933 [10, 11]. A recent commentary has questioned the high neurovirulence of Herpes B and has raised the possibility of higher rates of viral shedding in laboratory settings due to stress [11].

Herpes B has an approximately 156,400 bp genome, a high GC content of 74.5%, and has been shown to be closely related to papiine alphaherpesvirus 2 (HVP-2; herpes papio) and cercopithecine alphaherpesvirus 2 (SA8). With the advent of next-generation sequencing (NGS) the genomes of 19 Herpes B isolates have been sequenced [12-14]. The sequenced strains were isolated from six macaque species; Macaca (M.) fascularis (Crab-eating; Cynomologous; Cyno), M. fuscata (Japanese), M. mulatta
(Rhesus), *M. nemestrina* (Pig-tailed), *M. radiata* (Bonnet), and *M. silenus* (Lion-tailed). Macaque phylogenetic research has shown that of the macaque species featured in the current study, *M. silenus* and *M. nemestrina* are basal to the remaining species [15]. A herpes B multi-isolate analysis previously showed that herpes B strains isolated from *M. silenus* and *M. nemestrina* were distant from the remaining macaque derived sequences according to percent coding identity [12].

For several decades, the classic definition of species originating from Ernst Mayr has been “species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups” [16, 17]. This definition is problematic in virology as viruses undergo recombination [18-26], but they do not interbreed per se, so an alternative definition is required. The definition of species has not been static, with several alternative species concepts proposed based on biological, ecological, evolutionary, cohesion, phylogenetic, phenetic, and genotypic cluster properties, many of which have further subdivisions [27]. Related to challenges regarding species concepts, are cryptic species which have been described since the early 18th century [28, 29]. Cryptic species appear identical based on morphology but are on different evolutionary paths [29]. The definition of cryptic species lacks clarity, however, a recently proposed conceptual framework for identifying cryptic species involves “statistically separable and divergent genotypic clusters” [29]. To address these challenges several methods of species delimitation have been used in organisms ranging from bacteria to eukaryotes such as arbitrary distance thresholds, *in silico* DNA-DNA hybridization (isDDH) and generalized mixed Yule coalescent (GMYC) [30-33]. Previous phylogenetic studies of porcine circovirus type 2 (PCV2), H5N1 influenza, FHV-1, and the varicellovirus genus have used genomic nucleotide distance to establish clade cutoffs [34-37]. The goal of the current study was to use this genomic distance cutoff approach to determine if the herpes B strains isolated from *M. silenus* and *M. nemestrina* constituted cryptic viral species, warranting species status.

**Methods**

**Genome sequences and Genomic Sequence Alignment**

The genomic sequences of the viral strains used in the current study were downloaded from NCBI and can be found in Table 1. Several genomic multiple sequence alignments (MSAs) were generated with MAFFT (Linux ver. 7.394) using the FFT-NS-1 strategy option [38, 39]. MSAs with and without an outgroup were generated for herpes B, PaHV-2, and all available old-world monkey (OWM) genomic sequences. The generated MSAs were manually inspected, and locally aligned for optimization using ClustalW within the MEGA 7 package [40, 41]. The alignments generated for this study can be downloaded at [https://brandt.ophth.wisc.edu/data-sets/](https://brandt.ophth.wisc.edu/data-sets/).

**Nucleotide Substitution Model Optimization and Phylogeny**

Prior to phylogenetic network construction, the optimal substitution model for each MSA, and subsequent optimal model parameters were calculated using IQ-TREE version 1.6.3 [42]. Phylogenetic
networks for each of the alignments were generated using Splitstree 4 [43] using the optimal substitution model and parameters calculated by IQ-TREE. Maximum likelihood trees were generated using RAxMLGUI (ver. 1.3) using the GTRCATI option with 1000 bootstrap replicates [44].

**Genomic Nucleotide Distance and Clade Cutoff Calculations**

To determine clade cutoff parameters, pairwise distances were first calculated using the genomic MSAs without outgroups. The genomic MSAs without outgroups were used in order to minimize alignment gaps usually created by including an outgroup sequence. A previous statistical description of establishing clades using genomic nucleotide distance has been previously described [35]. Briefly a variance analysis framework was used, where the $F$ statistic was calculated for each dataset and plotted as a curve. Maximum composite likelihood (MCL) pairwise distances were calculated with MEGA 7 rather than uncorrected $p$-distances as have been used previously [34-37]. Species distance cutoffs were established by using the Old World monkey MSA, followed by graphing the frequency of the pairwise MCL distances using the using the R software package (ver. 3.4.4) [45]. A kernel density plot also generated in R to assist in determining the clade cutoff value by finding the trough between the low and high MCL distance populations. Intraspecies clade cutoffs were established in a similar manner, using the core herpes B, and herpes papio MSAs (minus outgroup) respectively.

**Results**

**Old World Monkey Simplex Virus Phylogeny**

To investigate if the Pig and Lion-tailed macaque simplex viruses warranted separate species status, the genomes of the available Old-World monkey (OWM) derived simplex viruses were downloaded from Genbank (Table 1). The available PaHV-2 strains were included in the analysis in order set an overall species cutoff for the OWM simplex viruses. The viral genomes were first aligned, and then the terminal repeat segments were deleted from the genomic multiple sequence alignment (MSA). The optimal nucleotide substitution model for the dataset was also calculated. This MSA alignment was used to generate a phylogenetic network which illustrates phylogenetic dissonance within the dataset (Figure 1A). The phylogenetic network in Figure 1A shows a “genetic continuum” with the core herpes B strains at one end, the Pig and Lion-tailed macaque derived strains located approximately in the middle, and the baboon viruses at the opposite end of the continuum. Additionally, the herpes B strain E90-136, isolated from a cyno macaque was separated from the core herpes B strains. A maximum likelihood (ML) tree was also generated to establish phylogenetic robustness, and the subsequent tree produced highly similar results to phylogenetic network (Figure 1B). The OWM simplex virus phylogenetic network and ML tree (Figures 1A and B) show similar phylogenetic tree topology to the Old-World monkey hosts (Figure 1C).

**Establishing Species Level Cutoffs**
Genomic nucleotide distance-based cutoff values have been used in the past in an effort to define viral intraspecies clades empirically [34-37]. In the current study we applied this distance-based method to define species level cutoffs. To begin to establish species level cutoffs, the maximum composite likelihood (MCL) pairwise distances between the 28 OWM viruses was calculated, the frequencies plotted, and a kernel density graph was overlaid (Figure 2A). A genomic distance cutoff for establishing species status was derived by marking the lowest point of the kernel density plot (8.94%) and is denoted by the vertical dashed line in Figure 2A. Thus, for the current data set, genomic distances over 8.94% merit species status, and under 8.94% do not. Using this genomic nucleotide-based distance cutoff approach, the Pig and Lion-tailed macaque simplex viruses merit separate, individual species status, as the distances between each other was 10.1%. The distance of the Pig and Lion-tailed macaques from the core herpes B strains was approximately 14% (Figure 2B), suggesting they are separate species. Using this method, SA8 and PaHV-2 retained species status, however the outlying core herpes B isolate E90-136 did not merit species status (6.1% distance; Figure 2B).

**Core Herpes B Clade**

The core herpes B strains isolated from rhesus, bonnet, and Japanese macaques were next examined to establish intraspecies genomic distance-based clade cutoff. Similar to the method described above, MSAs comprising the 15 core herpes B strains identified in Figures 1A and B were generated with and without an outgroup (M. nemestrina isolate KQ). Next, a phylogenetic network and maximum likelihood tree were constructed (Figures 3A and B) based on the alignment with an outgroup. The tree topology patterns between the two phylogenetic methods were nearly identical, with two basic groupings, aside from an outlier strain (9400371). Next, pairwise distances between the core herpes B strains were calculated using the core herpes B MSA without an outgroup, and the frequencies were plotted (Figure 3C). The genomic distance clade cutoff derived from the kernel density trough was 0.2031% (Figure 3C). The distance between groups 1 and 2 was 0.7689% (Figure 3D), which is above the distance cutoff validating their status as clades. The distance between strain 900371 and clades 1 and 2 was 0.07246% and 0.05295% respectively, therefore 900371 warrants consideration as a single member of a third clade.

**PaHV-2 Clade Structure**

The phylogenetic structure of the seven available PaHV-2 genomic sequences was next examined. Both the phylogenetic network and maximum likelihood tree recovered three groupings (Figures 4A and B). The clade cutoffs were performed in the same manner as described above, with the cutoff value calculated at 1.9611% distance (Figure 4C). The distances between groups 1, 2 and 3 were above the cutoff (Figure 4D), thus validating their clade status.

**Discussion**

In the current study we utilized a genomic nucleotide distance-based method previously used for identifying phylogenetic clades and applied it to detect viral species. The results suggest that herpes
simplex viruses isolated from Lion and Pig-tailed macaques should be designated as separate species. To our knowledge this is the first time this technique was been applied to virus species and may be useful in detecting cryptic viral species.

**Host-Virus Co-speciation**

Herpesviruses have been shown to cospeciate with their hosts [46], however they can cross species barriers [47], especially in captivity [48-55]. These captive transmissions, especially between macaque species can complicate phylogenetic analysis. In particular, cross-species transmission appears to be fairly common among the core herpes B strains, and has been discussed previously in depth by Eberle et al [12]. In some of the herpes B strains, the original source of the virus appears to be unclear. For instance, the cynomolgus macaque derived strain E90-136 is more distant and phylogenetically separated from the core herpes B strains (Figure 1), however it was not sufficiently distant (Figure 2) to be considered a separate species. Interestingly, strain E90-136 was isolated from a Cyno macaque which died due to a disseminated infection caused by the virus [56]. Herpes B strains are generally asymptomatic within the natural host, which may suggest that Cyno macques are not the natural reservoir for this particular strain. For other OWM strains, interspecies spread is well documented. The isolate 8100812 was originally isolated from a DeBrazza monkey, however restriction digest patterns showed that the Lion-tailed macaque was the natural host [51]. Phylogenetically, this appears appropriate as strain 8100812 forms a node with the two Pig-tailed macaque isolates (Figures 1A and B), and importantly matches phylogenetic profile of the macaque species themselves (Figure 1C). The correlation between Lion and Pig-tailed virus and macaque phylogeny strongly suggests host-virus co-speciation. Additionally, while natural cross-species viral transmissions between animals does occur [47, 57-59], natural species viral transmissions between the animals and viruses in this study are fairly unlikely given that the natural host ranges of the monkeys (Figure 5). Exceptions are Lion-tailed and Bonnet macaques, as well as the Pig-tailed and Cyno macaques, where there is considerable overlap in the host ranges (Figure 5).

**Viral Species Concept**

Standard definitions of what constitutes a biological species, such a reproductively isolated population [16], are insufficient for viruses as they replicate, but do not reproduce like other organisms. Originally, viruses were simply classified according to the host that was infected, i.e. bacterial, plant or animal [60]. It wasn't until 1950 that official principles of animal virus classification were established, with categories such as morphology, chemical composition, method of transmission, tropism and symptomatology [60]. In 1963 the International Committee on Nomenclature of Viruses (ICNV) was established and in 1966 the body proposed a taxonomic framework and classification rules which included class, order, family. This organization is now known as the International Committee for Taxonomy of Viruses (ICTV) [60, 61]. In 1990 the ICTV established an official definition of viral species which was stated as “a virus species is a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche” [62], and has since evolved to state “a monophyletic group of
viruses whose properties can be distinguished from those of other species by multiple criteria....not limited to natural and experimental host range, cell and tissue tropism, pathogenicity, vector specificity, antigenicity, and the degree of relatedness of their genomes or genes [63]. While this statement recommends distinguishing properties for determining species, the process is still ambiguous.

We chose to focus our efforts on genomic distance in order to apply a quantitative measure to delimit viral species. Several species delimitation methods have been used in bacteria and eukaryotes. One of the most common and recent methods for species delimitation in bacteria and eukaryotes is generalized mixed Yule coalescent, where branching patterns of a single tree transition from Yule process interspecies branching to coalescent process intra-species branching [33]. Single loci can be used for this method, however more recently multiple genes and morphological characters can be used [64]. Previously, a distance method based on gene homology and sharing was used to reevaluate viral family classifications [65]. A relatively simple genomic distance cutoff method has been used to validate viral clades [34-37] and was applied to delimit species in the current study. A caveat with the distance cutoff value used in the current study is that the cutoff value is not universal, but dataset dependent. A potential issue with using the distance cutoff method to establish species boundaries is that as the genomes of additional viruses are sequenced, the species cutoff value could potentially shift, resulting in species cutoff values that could vary over time. A general complication of the method used in the current study and in other genetic data delimitation techniques is that the methods may be delimiting populations, and not necessarily species [66]. We cannot eliminate this possibility in our analysis however this is unlikely given the large distance values between species in the dataset. In our study to determine if the Lion and Pig-tailed derived simplex viruses were species separate from herpes B, we included all sequenced Old-World monkey strains in an effort not to bias the results and establish a general cutoff for the Old-World monkey group. The results of our study showed the genome-based genetic distance between Lion/Pig-tailed macaque derived viruses and the core herpes B strains were both approximately 14%, which was actually greater than the distance observed (~10%) between SA8 and herpes papio (Figure 2B), previously established viral species. The recovery of SA8 and PaHV-1 as separate species helps to validate the method. Both of these values were well above the species cutoff value (8.94%; Figure 2B). The genetic distance data, and the data supporting co-speciation of the Lion and Pig-tailed macaque viruses reinforces the idea that these should be designated as separate, individual species from herpes B, and each other.

Cryptic Viral Species

The term cryptic species is related to similar concepts such as sibling species, species complex, and superspecies, with the definitions between these concepts often blurred. Cryptic species are generally defined as species which appear virtually identical phenotypically, but belong to different taxa, and were thus “hidden”. Cryptic species were originally described three centuries ago [28, 29], and with modern molecular techniques have been increasingly identified across multiple organisms [67-71]. To our knowledge, the concept of cryptic species has not been applied to viruses, however species complex occasionally has [72, 73]. From the phylogenetic network of the Old-World monkey simplex viruses
(Figure 1A), these viruses could be described as a series of species complexes (i.e. a group closely related viruses that are difficult to separate), one comprising the macaque viruses and a second encompassing the baboon simplex viruses. The genetic distance cutoff method may be useful in establishing species boundaries in these complexes, as the method confirmed species status for the baboon derived PaHV-2 and SA8. Importantly, the method identified Lion and Pig-tailed simplex viruses as separate species (Figure 2), defining these viruses essentially cryptic species. The genetic distance cutoff method provides a quantitative threshold to determine species status and could be another tool for establishing species status among viral cryptic species complexes.

Summary

In summary genome-based phylogenetic and genetic distance cutoff techniques were applied to the available Old-World monkey simplex virus genome sequences. The results showed that Lion and Pig-tailed macaque simplex viruses were approximately 14% distant from core herpes B strains, which was more distant than between PaHV-2 strains and SA8, previously established viral sequences. The genomic distance cutoff method recovered PaHV-2 and SA8 as separate strains, and Lion and Pig-tailed macaque simplex viruses as separate species, effectively identifying these macaque viruses as cryptic species. Based on the genetic distance analysis, the fact that the OWM hosts are designated as separate species, and herpes viruses co-evolve with their hosts, we propose establishing Lion and Pig-tailed macaque simplex viruses as separate species. This may be the first identification of cryptic viral species.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The multiple sequence alignments used for this study are available for download at http://sites.ophth.wisc.edu/brandt/.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We would like to thank Dr. Cécile Ané (University of Wisconsin-Madison) for her helpful comments regarding the manuscript.
Authors contributions

AK and CRB conceived and designed the experiments. AK performed the experiments. AK and CRB analyzed the data. AK and CRB contributed to the writing of the manuscript.

Funding

This study was supported by grants from the NIH Core Grant for Vision Research (P30EY016665), NIH/HEI (R21AI137280), an unrestricted grant to the Department of Ophthalmology and Visual Sciences from Research to Prevent Blindness, Inc, and the NSF (190289).

References

1. Sheppard M, May JT: Location and characterization of the bovine herpesvirus type 2 thymidine kinase gene. J Gen Virol 1989, 70 (Pt 11):3067-3071.
2. Babra B, Watson G, Xu W, Jeffrey BM, Xu JR, Rockey DD, Rohrmann GF, Jin L: Analysis of the genome of leporid herpesvirus 4. Virology 2012, 433(1):183-191.
3. Sasaki M, Setiyono A, Handharyani E, Kobayashi S, Rahmadani I, Taha S, Adiani S, Subangkit M, Nakamura I, Sawa H et al: Isolation and characterization of a novel alphaherpesvirus in fruit bats. J Virol 2014, 88(17):9819-9829.
4. Vaz PK, Mahony TJ, Hartley CA, Fowler EV, Ficorilli N, Lee SW, Gilkerson JR, Browning GF, Devlin JM: The first genome sequence of a metatherian herpesvirus: Macropodid herpesvirus 1. BMC Genomics 2016, 17:70.
5. Mahony TJ, Smith GA, Thomson DM: Macropodid herpesviruses 1 and 2 occupy unexpected molecular phylogenetic positions within the Alphaherpesvirinae. J Gen Virol 1999, 80 (Pt 2):433-436.
6. Holden FPGaM: The Herpes Encephalitis Problem, II. The Journal of Infectious Diseases 1933, 53(3):17.
7. Holden FPGaM: Isolation of herpes virus from several cases of epidemic encephalitis. Proc Soc Exp Biol Med 1933, 30:3.
8. Cohen JL, Davenport DS, Stewart JA, Deitchman S, Hilliard JK, Chapman LE, Group BVW: Recommendations for prevention of and therapy for exposure to B virus (cercopithecine herpesvirus 1). Clin Infect Dis 2002, 35(10):1191-1203.
9. B Virus (herpes B, monkey B virus, herpesvirus simiae, and herpesvirus B) [https://www.cdc.gov/herpesbvirus/index.html]
10. Hilliard J: Monkey B virus. In: Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Edited by Arvin A C-FG, Mocarski E, et al., 1st edn: Cambridge University Press; 2007.
11. Eberle R, Jones-Engel L: Questioning the Extreme Neurovirulence of Monkey B Virus (Macacine alphaherpesvirus 1). Adv Virol 2018, 2018:5248420.
12. Eberle R, Maxwell LK, Nicholson S, Black D, Jones-Engel L: Genome sequence variation among isolates of monkey B virus (Macacine alphaherpesvirus 1) from captive macaques. Virology 2017, 508:26-35.

13. Ohsawa K, Black D, Ohsawa M, Eberle R: Genome sequence of a pathogenic isolate of monkey B virus (species Macacine herpesvirus 1). Arch Virol 2014, 159(10):2819-2821.

14. Perelygina L, Zhu L, Zurkuhlen H, Mills R, Borodovsky M, Hilliard JK: Complete sequence and comparative analysis of the genome of herpes B virus (Cercopithecine herpesvirus 1) from a rhesus monkey. J Virol 2003, 77(11):6167-6177.

15. Li J, Han K, Xing J, Kim HS, Rogers J, Ryder OA, Disotell T, Yue B, Batzer MA: Phylogeny of the macaques (Cercopithecidae: Macaca) based on Alu elements. Gene 2009, 448(2):242-249.

16. Mayr E: Systematics and the Origin of Species. In.: New York: Columbia University Press; 1942.

17. de Queiroz K: Ernst Mayr and the modern concept of species. Proc Natl Acad Sci U S A 2005, 102 Suppl 1:6600-6607.

18. Casto AM, Roychoudhury P, Xie H, Selke S, Perchetti GA, Wofford H, Huang ML, Verjans G, Gottlieb GS, Wald A et al: Large, stable, contemporary interspecies recombination events in circulating human herpes simplex viruses. J Infect Dis 2019.

19. Burrel S, Boutolleau D, Ryu D, Agut H, Merkel K, Leendertz FH, Calvignac-Spencer S: Ancient Recombination Events between Human Herpes Simplex Viruses. Mol Biol Evol 2017, 34(7):1713-1721.

20. Koelle DM, Norberg P, Fitzgibbon MP, Russell RM, Greninger AL, Huang ML, Stensland L, Jing L, Magaret AS, Diem K et al: Worldwide circulation of HSV-2 x HSV-1 recombinant strains. Sci Rep 2017, 7:44084.

21. Pringle CR: Evidence of Genetic Recombination in Foot-and-Mouth Disease Virus. Virology 1965, 25:48-54.

22. Ledinko N: Genetic recombination with poliovirus type 1. Studies of crosses between a normal horse serum-resistant mutant and several guanidine-resistant mutants of the same strain. Virology 1963, 20:107-119.

23. Ledinko N: Temperature-sensitive mutants of type 12 adenovirus defective in a late function: protein synthesis and evidence for recombination between mutants in complementation group D. J Gen Virol 1976, 30(1):153-156.

24. Kolbourne ED: Recombination of influenza A viruses of human and animal origin. Science 1968, 160(3823):74-76.

25. Wildy P: Recombination with herpes simplex virus. J Gen Microbiol 1955, 13(2):346-360.

26. Lee K, Kolb AW, Sverchkov Y, Cuellar JA, Craven M, Brandt CR: Recombination analysis of herpes simplex virus type 1 reveals a bias towards GC content and the inverted repeat regions. J Virol 2015, 89:7214-7223.

27. De Queiroz K: Species concepts and species delimitation. Syst Biol 2007, 56(6):879-886.
28. Winker K: Sibling species were first recognized by William Derham (1718). Auk 2005, 122(2):706-707.
29. Struck TH, Feder JL, Bendiksby M, Birkeland S, Cerca J, Gusarov VI, Kistenich S, Larsson KH, Liow LH, Nowak MD et al.: Finding Evolutionary Processes Hidden in Cryptic Species. Trends Ecol Evol 2018, 33(3):153-163.
30. Schloss PD, Handelsman J: Toward a census of bacteria in soil. PLoS Comput Biol 2006, 2(7):e92.
31. Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL, Polz MF: Fine-scale phylogenetic architecture of a complex bacterial community. Nature 2004, 430(6999):551-554.
32. Meier-Kolthoff JP, Auch AF, Klenk HP, Goker M: Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013, 14:60.
33. Tang CQ, Humphreys AM, Fontaneto D, Barraclough TG, Paradis E: Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. Methods Ecol Evol 2014, 5(10):1086-1094.
34. Lewin AC, Kolb AW, McLellan GJ, Bentley E, Bernard KA, Newbury SP, Brandt CR: Genomic, recombinational and phylogenetic characterization of global feline Herpesvirus 1 isolates. Virology 2018, 518:385-397.
35. Kolb AW, Lewin AC, Moeller Trane R, McLellan GJ, Brandt CR: Phylogenetic and recombination analysis of the herpesvirus genus varicellovirus. BMC Genomics 2017, 18(1):887.
36. Segales J, Olvera A, Grau-Roma L, Charreyre C, Nauwynck H, Larsen L, Dupont K, McCullough K, Ellis J, Krakowka S et al.: PCV-2 genotype definition and nomenclature. Vet Rec 2008, 162(26):867-868.
37. Xiao CT, Halbur PG, Opriessnig T: Global molecular genetic analysis of porcine circovirus type 2 (PCV2) sequences confirms the presence of four main PCV2 genotypes and reveals a rapid increase of PCV2d. J Gen Virol 2015, 96(Pt 7):1830-1841.
38. Katoh K, Toh H: Parallelization of the MAFFT multiple sequence alignment program. Bioinformatics 2010, 26(15):1899-1900.
39. Katoh K, Standley DM: MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013, 30(4):772-780.
40. Kumar S, Stecher G, Tamura K: MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 2016, 33(7):1870-1874.
41. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentín F, Wallace IM, Wilm A, Lopez R et al.: Clustal W and Clustal X version 2.0. Bioinformatics 2007, 23(21):2947-2948.
42. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS: ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods 2017, 14(6):587-589.
43. Huson DH, Bryant D: Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 2006, 23(2):254-267.
44. Berger SA, Krompass D, Stamatakis A: Performance, accuracy, and Web server for evolutionary placement of short sequence reads under maximum likelihood. Syst Biol 2011, 60(3):291-302.
45. Team RC: R: A language and environment for statistical computing. In.: R Foundation for Statistical Computing; 2013.

46. McGeoch DJ, Dolan A, Ralph AC: Toward a comprehensive phylogeny for mammalian and avian herpesviruses. J Virol 2000, 74(22):10401-10406.

47. Pedersen K, Turnage CT, Gaston WD, Arruda P, Alls SA, Gidlewski T: Pseudorabies detected in hunting dogs in Alabama and Arkansas after close contact with feral swine (Sus scrofa). BMC Vet Res 2018, 14(1):388.

48. Malherbe H, Strickland-Cholmley M: Simian herpesvirus SA8 from a baboon. Lancet 1969, 2(7635):1427.

49. Malherbe H, Strickland-Cholmley M: Virus from baboons. Lancet 1969, 2(7633):1300.

50. Fukushi H, Tomita T, Taniguchi A, Ochiai Y, Kirisawa R, Matsumura T, Yanai T, Masegi T, Yamaguchi T, Hirai K: Gazelle herpesvirus 1: a new neurotropic herpesvirus immunologically related to equine herpesvirus 1. Virology 1997, 227(1):34-44.

51. Thompson SA, Hilliard JK, Kittel D, Lipper S, Giddens WE, Jr., Black DH, Eberle R: Retrospective analysis of an outbreak of B virus infection in a colony of DeBrazza's monkeys (Cercopithecus neglectus). Comp Med 2000, 50(6):649-657.

52. Loomis MR, O'Neill T, Bush M, Montali RJ: Fatal herpesvirus infection in patas monkeys and a black and white colobus monkey. J Am Vet Med Assoc 1981, 179(11):1236-1239.

53. Wilson RB, Holscher MA, Chang T, Hodges JR: Fatal Herpesvirus simiae (B virus) infection in a patas monkey (Erythrocebus patas). J Vet Diagn Invest 1990, 2(3):242-244.

54. Coulibaly C, Hack R, Seidl J, Chudy M, Itter G, Plesker R: A natural asymptomatic herpes B virus infection in a colony of laboratory brown capuchin monkeys (Cebus apella). Lab Anim 2004, 38(4):432-438.

55. Sekulin K, Jankova J, Kolodziejek J, Huemer HP, Gruber A, Meyer J, Nowotny N: Natural zoonotic infections of two marmosets and one domestic rabbit with herpes simplex virus type 1 did not reveal a correlation with a certain gG-, gI- or gE genotype. Clin Microbiol Infect 2010, 16(11):1669-1672.

56. Simon MA, Daniel MD, Lee-Parritz D, King NW, Ringler DJ: Disseminated B virus infection in a cynomolgus monkey. Lab Anim Sci 1993, 43(6):545-550.

57. Parrish CR, Holmes EC, Morens DM, Park EC, Burke DS, Calisher CH, Laughlin CA, Saif LJ, Daszak P: Cross-species virus transmission and the emergence of new epidemic diseases. Microbiol Mol Biol Rev 2008, 72(3):457-470.

58. Jin MJ, Hui H, Robertson DL, Muller MC, Barre-Sinoussi F, Hirsch VM, Allan JS, Shaw GM, Sharp PM, Hahn BH: Mosaic genome structure of simian immunodeficiency virus from west African green monkeys. EMBO J 1994, 13(12):2935-2947.

59. Faria NR, Suchard MA, Rambaut A, Streicker DG, Lemey P: Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. Philos Trans R Soc Lond B Biol Sci 2013, 368(1614):20120196.
60. Matthews REF: The History of Virus Taxonomy. In: A Critical Appraisal of Viral Taxonomy. Edited by Matthews REF: CRC Press; 1983: 256.
61. Lwoff A: The new provisional committee on nomenclature of viruses. International Bulletin of Bacterial Nomenclature and Taxonomy 1964, 14:53.
62. Regenmortel MHV: Virus Species. In: Species: The units of biodiversity. Edited by M.F. Claridge HADaMRW: Chapman and Hall; 1997: 549.
63. ICTV Information [https://talk.ictvonline.org/information/w/ictv-information/383/ictv-code]
64. Solis-Lemus C, Knowles LL, Ane C: Bayesian species delimitation combining multiple genes and traits in a unified framework. Evolution 2015, 69(2):492-507.
65. Aiewsakun P, Adriaenssens EM, Lavigne R, Kropinski AM, Simmonds P: Evaluation of the genomic diversity of viruses infecting bacteria, archaea and eukaryotes using a common bioinformatic platform: steps towards a unified taxonomy. J Gen Virol 2018, 99(9):1331-1343.
66. Sukumaran J, Knowles LL: Multispecies coalescent delimits structure, not species. Proc Natl Acad Sci U S A 2017, 114(7):1607-1612.
67. Surveswaran S, Gowda V, Sun M: Using an integrated approach to identify cryptic species, divergence patterns and hybrid species in Asian ladies' tresses orchids (Spiranthes, Orchidaceae). Mol Phylogenet Evol 2018, 124:106-121.
68. Saitoh T, Sugita N, Someya S, lwami Y, Kobayashi S, Kamigaichi H, Higuchi A, Asai S, Yamamoto Y, Nishiumi I: DNA barcoding reveals 24 distinct lineages as cryptic bird species candidates in and around the Japanese Archipelago. Mol Ecol Resour 2015, 15(1):177-186.
69. Hanelt B, Schmidt-Rhaesa A, Bolek MG: Cryptic species of hairworm parasites revealed by molecular data and crowdsourcing of specimen collections. Mol Phylogenet Evol 2015, 82 Pt A:211-218.
70. Crespo A, Lumbsch HT: Cryptic species in lichen-forming fungi. IMA Fungus 2010, 1(2):167-170.
71. Hahn MW, Huymann LR, Koll U, Schmidt J, Lang E, Hoetzinger M: Polynucleobacter wuianus sp. nov., a free-living freshwater bacterium affiliated with the cryptic species complex PnecC. Int J Syst Evol Microbiol 2017, 67(2):379-385.
72. Palacios G, Savji N, Travassos da Rosa A, Desai A, Sanchez-Seco MP, Guzman H, Lipkin WI, Tesh R: Characterization of the Salehabad virus species complex of the genus Phlebovirus (Bunyaviridae). J Gen Virol 2013, 94(Pt 4):837-842.
73. Gundacker ND, Carrera JP, Castillo M, Diaz Y, Valenzuela J, Tamhane A, Moreno B, Pascale JM, Tesh RB, Lopez-Verges S: Clinical Manifestations of Punta Toro Virus Species Complex Infections, Panama, 2009. Emerg Infect Dis 2017, 23(5):872-874.
74. Tyler SD, Peters GA, Severini A: Complete genome sequence of cercopithecine herpesvirus 2 (SA8) and comparison with other simplexviruses. Virology 2005, 331(2):429-440.

Tables

Table 1.
| Abbreviation | Synonym | Strain | Host                  | Genome Length | Accession Number |
|--------------|---------|--------|-----------------------|---------------|------------------|
| HSV-1        | Herpes virus type 1 | 17     | Homo sapiens          |               | NC_001806.2      |
| CeHV-2       | SA8     | B264   | Cercopithecus aethiops* | 150,715       | NC_006560.1      |
| HVP-2        | Herpes papio | X313   | Papio cynocephalus    | 156,487       | NC_007653.1      |
| HVP-2        | Herpes papio | OU4-2  | Papio ursinus         | 138,963       | KF908244.1       |
| HVP-2        | Herpes papio | OU4-8  | Papio ursinus         | 139,193       | KF908243.1       |
| HVP-2        | Herpes papio | A951   | na                    | 138,559       | KF908242.1       |
| HVP-2        | Herpes papio | OU2-5  | Papio cynocephalus    | 138,807       | KF908241.1       |
| HVP-2        | Herpes papio | OU1-76 | Papio cynocephalus    | 148,944       | KF908240.1       |
| HVP-2        | Herpes papio | A189164 | na                  | 139,366       | KF908239.1       |
| CeHV-1       | Herpes B  | E2490  | Macaca mulatta       | 156,789       | NC_004812.1      |
| CeHV-1       | Herpes B  | M12-O  | Macaca radiata       | 155,404       | KY628985.1       |
| CeHV-1       | Herpes B  | 9400371 | Macaca mulatta or fascicularis* | 155,143 | KY628983.1 |
| CeHV-1       | Herpes B  | 7709642 | Macaca fuscata       | 155,141       | KY628982.1       |
| CeHV-1       | Herpes B  | 32425-G | Macaca mulatta       | 155,528       | KY628981.1       |
| CeHV-1       | Herpes B  | 32188-O | Macaca mulatta       | 155,099       | KY628980.1       |
| CeHV-1       | Herpes B  | 32157-G | Macaca mulatta       | 155,777       | KY628979.1       |
| CeHV-1       | Herpes B  | 31618-G | Macaca mulatta       | 155,425       | KY628978.1       |
| CeHV-1       | Herpes B  | 31612-G | Macaca mulatta       | 155,321       | KY628977.1       |
| CeHV-1       | Herpes B  | 26896-O | Macaca mulatta       | 155,583       | KY628976.1       |
| CeHV-1       | Herpes B  | 26896-G | Macaca mulatta       | 155,609       | KY628975.1       |
| CeHV-1       | Herpes B  | 24105-G | Macaca mulatta       | 155,021       | KY628974.1       |
| CeHV-1       | Herpes B  | 20620  | Macaca mulatta       | 155,323       | KY628973.1       |
| CeHV-1       | Herpes B  | 16293  | Macaca mulatta       | 155,180       | KY628972.1       |
| CeHV-1       | Herpes B  | 12930  | Macaca mulatta       | 155,462       | KY628971.1       |
| CeHV-1       | Herpes B  | KQ     | Macaca nemestrina    | 157,321       | KY628970.1       |
| CeHV-1       | Herpes B  | 1504-11 | Macaca nemestrina    | 156,905       | KY628969.1       |
CeHV-1 | Herpes B | 8100812 | Macaca silenus$ | 157,447 | KY628968.1
---|---|---|---|---|---
CeHV-1 | Herpes B | E90-136 | Macaca fascicularis | 155,157 | KJ566591.2

*Subsequent studies following isolation show that the natural reservoir for SA8 is baboons [48, 49, 74].

*Host species differs between the Genbank annotation and the corresponding publication [12].

$Strain was originally isolated from *C. neglectus*, however subsequent work showed the natural reservoir is *M. silenus* [51].

**Figures**
Figure 1

Phylogenetic analysis of Old-World monkey (OWM) derived simplexviruses. OWM viral genomic sequences (Table 1) were aligned with MAFFT ver. 7.394 and the optimal substitution model was calculated by IQ-Tree [39, 42]. A) Phylogenetic network generated from the alignment using Splitstree ver. 4.14 and the HKY+G+I substitution model (gaps deleted; p-inv = 0.469; gamma = 1.138) [43]. B) Maximum Likelihood tree generated from an alignment using HSV-1 as an outgroup using RAxMLGUI
Figure C shows a macaque monkey phylogenetic tree based on data presented by Li et al [15].

![Graph showing a distribution of pairwise distances with a cutoff value of 8.94%]

**Figure 2**

Establishing viral species cutoff value. Pairwise distances in the Old World monkey virus alignment were calculated using Mega 7 [40], and the frequencies plotted using the R package. A kernel density plot was also generated and combined with the distance frequencies (A). A distance cutoff value was established
by determining the trough of the kernel plot, which is depicted by a vertical dotted line (8.94%). Mega 7 was used to calculate between group distances which is shown in Figure B.

**Figure 3**

Core herpes B phylogeny and clades. A genome sequence alignment was generated with the core herpes B strains identified in Figure 1. A phylogenetic network using the HKY+G+I substitution model (gaps deleted; p-inv = 0.686; gamma = 0.927) (A) and maximum likelihood tree (B) were then produced, finding provisionally three clades. Pairwise distances between the strains were plotted (shown in Figure C) and a clade cutoff value (vertical dotted line) was calculated (0.203%). Figure D contains a table showing the between group genetic distances.
Figure 4

PaHV-2 phylogeny and clades. A genome sequence alignment was generated with the available PaHV-2 strains (Table 1). A phylogenetic network (Figure A) was generated using the HKY+G+I substitution model recommended by IQ-Tree (gaps deleted; $p$-inv = 0.572; gamma = 0.739). Figure B shows a maximum likelihood tree which shows three clades. Pairwise distances between the strains were plotted (Figure C) and a cladecutoff value calculated (1.96%). Figure D includes a table showing the between group genetic distances.
Figure 5

Maps depicting macaque species ranges. The figure shows the natural ranges for the Pig-tailed, Lion-Tailed, Bonnet, Crab Eating/Cynomologous, Japanese and Rhesus macaques. The maps were generated in R (version 3.4.2, “maps” package), and edited using Adobe Illustrator. The ranges were based on those presented by IUCN Red List of Threatened Species (https://www.iucnredlist.org/).