JVI Accepted Manuscript Posted Online 3 June 2015
J. Virol. doi:10.1128/JVI.01112-15

Arenavirus Co-Infections are Common in Snakes with Boid Inclusion Body Disease

Hepojoki J[1*], Salmenperä P[2], Sironen T[1], Hetzel U[3,4], Korzykov Y[1], Kipar A[3,4], and Vapalahti O[1,4,5]

[1]University of Helsinki, Department of Virology, Haartman Institute, Helsinki, Finland

[2]Blueprint Genetics, Helsinki, Finland

[3]Institute of Veterinary Pathology, Vetsuisse Faculty, University of Zürich, Switzerland

[4]Department of Veterinary Biosciences, Faculty of Veterinary Medicine, University of Helsinki, Finland

[5]Department of Virology and Immunology, University of Helsinki, Hospital District of Helsinki and Uusimaa, Finland

Running title: Co-infection in BIBD

***Corresponding author**

Mailing address: **Department of Virology**
Haartman Institute
P.O. Box 21
Haartmaninkatu 3
FI-00014 University of Helsinki
Finland
***Phone:* +358-9-19126608**
***Fax:* +358-9-191 26491**
***E-mail:* jussi.hepojoki@helsinki.fi**

Word number in summary: 244

Word number in body text: 1325

ABSTRACT

Recently, novel arenaviruses were found in snakes with Boid Inclusion Body Disease (BIBD); these form a new genus *Reptarenavirus* within the family *Arenaviridae*. We used NGS and *de novo* sequence assembly to investigate reptarenavirus isolates from our previous study. Four of the six isolates and all samples from snakes with BIBD contained at least two reptarenaviruses species. The sequenced viruses comprise four novel reptarenavirus species and a representative of a new arenavirus genus.

TEXT

Until very recently, arenaviruses were known as a group of mainly rodent-borne zoonotic viruses *(1)*. The negative-sense RNA genome of arenaviruses is divided into two segments, the small (S, approximately 3.5 kb) and the large (L, approximately 7-7.5 kb) segment; both use an ambisense coding strategy *(1)*. The S segment encodes for the glycoprotein precursor (GPC) and the nucleoprotein (NP), whereas the RNA-dependent RNA polymerase (RdRp) and the Z protein (ZP) are encoded in the L segment *(1)*. The discovery of arenaviruses in snakes with boid inclusion body disease (BIBD), by three independent groups *(2-4)*, has expanded the family *Arenaviridae* by a new group of viruses. In fact, the BIBD–associated arenaviruses (BIBDAVs) have been suggested to form a new arenavirus genus called *Reptarenavirus (5)*. At the same time, the arenavirus study group of the International Committee on Taxonomy of Viruses (ICTV) has suggested that the genus *Arenavirus*, harboring the "classical" Old and New World arenaviruses, is renamed to *Mammarenavirus (6)*. According to a recent proposal, the genus *Reptarenavirus* would contain three species: Alethinophid reptarenavirus-1 (member virus: Golden Gate virus), -2 (CAS virus, CASV), and -3 (University of Helsinki virus-1 (UHV-1), Boa AV NL B3 virus). While *in vitro* evidence suggests a causal relationship between arenavirus infection and BIBD *(4)*, the *in vivo* proof is still missing. Also, the reservoir host(s) of the reptarenaviruses has not yet been confirmed; however, our recent study suggests that these viruses preferentially grow in organisms with body temperatures close to 30°C *(7)*.

To study the diversity of reptarenaviruses, we applied next-generation sequencing (NGS) to characterize certain isolates described in our previous report *(4)*. Based on phylogeny, we selected isolates originating from six *Boa constrictor* snakes, and used a continuous *B. constrictor* kidney cell line, I/1Ki *(4)*, for their propagation. While the virus isolate of snake 1 has already been almost fully sequenced (GenBank accession numbers KF297880.1 and

KF297881.1), only partial L segment sequences were available for isolates from snakes 5 (KF564801.1), 8 (KF564796.1), 9 (KF564800.1), 11, and 41 (KF564797.1). We infected clean I/1Ki cultures with tissue homogenates and collected the viruses produced during the first passage by pelleting through a sucrose cushion as described *(4)*. Viral RNAs were extracted from the pelleted viruses using the QIAmp Viral RNA Mini Kit (Qiagen) according to the manufacturer's instructions, without carrier RNA. The isolated RNA was further purified and concentrated using SPRISelect beads (Agencourt). Indexed illumina sequencing libraries were prepared using the NEBNext Ultra RNA Library Prep Kit (New England Biolabs). Pooled libraries were sequenced on the Illumina MiSeq with 161 bp paired end reads. Reads were demultiplexed, adapter sequences removed, and sample FASTQ files were produced using the MiSeq reporter. *De novo* contiguous sequence (contig) assembly was performed with MIRA version 4.0.2 (http://mira-assembler.sourceforge.net/docs/DefinitiveGuideToMIRA.html), using the CSC (IT Center for Science Ltd., Finland) supercluster Taito. Initial contigs were used to remove the cellular RNA background with Mirabait tool, after which *de novo* assemblies were run with a subset of reads extracted with Chipster v.3.1.0 *(8)*.

For most virus preparations, contigs of full length or almost full length S and L segments of arenaviruses were obtained (Table 1). The contig coverages, determined using Bowtie2 *(9)* in Unipro UGENE 1.14.2 *(10)*, are presented graphically in Table 1. The virus isolate from snake 1 served as a positive control for the study, since it represents the UHV-1 that we have characterized in detail *(4)*. Bowtie2 alignment of the NGS data in Unipro UGENE 1.14.2 to the UHV-1 reference L segment (KF297881.1) revealed that such sequence did not exist in the sample. Instead, two nearly complete L segment sequences were recovered from the purified UHV preparation by *de novo* sequence assembly. The database UHV sequence KF297881.1 was assembled from NGS data and by Sanger sequencing of a PCR cloned fragment from pGEM-T vector. At the time we did not expect co-infection with multiple arenaviruses and

accidentally generated an "*in silico*" recombinant of the two viruses in the preparation. The KF297881.1 combines the first ~2,200 nt from the 5´-end of the first L segment and ~5,000 nt from the 3´end of the second L segment found in the preparation. The region where the mismatch occurred contains 24/26 identical nucleotides in both sequences. The fact that an S segment nearly identical (3367/3393, >99% nt identity) to the UHV-1 reference S segment (KF297881.1) was recovered from the NGS data provides further evidence of the reliability of the chosen *de novo* assembly approach. Interestingly, three of the other virus preparations studied contained more than one L and S segment, indicating that several snakes were co- or superinfected with two or more reptarenaviruses. For most (3/4) samples with several reptarenaviruses, two L segment sequences were constructed, however, one of the isolates (from snake 8) contained six different L segments. Curiously, at maximum two full length (or nearly full length) S segments were constructed per virus preparation, which may reflect stronger selection pressure or an unknown method-induced bias. The GenBank accession numbers for the assembled S and L segments are provided in Table 1.

We then used PAirwise Sequence Comparison (PASC, *(11)*, available at http://www.ncbi.nlm.nih.gov/sutils/pasc/viridty.cgi, to compare the obtained sequences to each other and to the arenavirus S and L segments in the databases (Table 1). According to PASC analysis and ICTV criteria several of the viruses that we sequenced showed <76% nt identity to known arenaviruses and thus represent putative new Alethinophid reptarenavirus species, as depicted by phylogeny (Fig. 1). For Alethinophid reptarenavirus 4 (University of Giessen virus (UGV)-1 and -2) we recovered both segments, since isolates from snakes 5 and 9 only contained a single virus each. However, for the putative representatives of Alethinophid reptarenavirus 5 to 7 (Tavallinen suomalainen mies virus, TSMV; Hans Kompis virus, HKV; and Suri Vanera virus, SVaV), we only obtained the sequence for the L segment. Remarkably, according to the PASC analysis, one of our isolates, Haartman Institute snake virus (HISV),

would be sufficiently distant from the known arenaviruses to represent a novel arenavirus genus. The S segment of HISV encodes GPC and NP by ambisense coding strategy, but despite of our attempts we were unable to obtain the full-length L segment (the ZP gene and part of RdRp missing). To roughly estimate the relative quantities of the viruses in each preparation, we selected primers targeting the L and S segments specific for each identified virus (primer sequences available upon request). We performed quantitative reverse transcription polymerase chain reaction (qRT-PCR) from purified virus preparations with the respective primers and confirmed the presence of viruses and primer specificity by Sanger sequencing. The prevalence of L segments per purified virus preparation is presented in table 1.

We further tested the original tissues (stored frozen at -80°C) from the diseased constrictor snakes that had served to generate the isolates (liver from snakes 5, 8, 9, and 11) by using the L segment primers in RT-PCR, and confirmed the presence of these viruses in the original tissues by Sanger sequencing the products. To our surprise, also the liver of snakes 5 and 9, unlike the respective isolates, were found to have multiple reptarenaviruses (UGV-2, HKV, and ABV-3 in snake 5; HKV and UGV-3 in snake 9) and snake 11 was found positive for all except ABV-2 and HISV. All studied snakes were from Europe. PASC analysis of the sequences revealed that the virus preparations of co-infected snakes always contained at least two different reptarenavirus species and thus co- or superinfection might be essential for the development of BIBD. *In vitro* studies with mammarenaviruses demonstrate that a persistent arenavirus infection restricts the replication of a serologically similar, but not more distant arenavirus *(12)*. Some strains of lymphocytic choriomeningitis virus (LCMV), the prototype arenavirus, induce an acquired immunosuppression in their natural host *(14)*, as also suggested for BIBD *(13)*. We hypothesize that reptarenaviruses may establish a chronic infection in snakes. Co- or superinfection of a chronically infected snake with another reptarenavirus

139 species might then result in amplified replication and, eventually, in BIBD, due to
140 immunosuppression induced by the chronically infecting virus.

141 The data presented herein demonstrates that the recently established genus *Reptarenavirus*
142 is likely to expand in the near future, since we identified four novel representatives of this
143 genus from only six *B. constrictors*. Furthermore, our results suggest that co- or superinfection
144 with reptarenaviruses is common and might be relevant for the pathogenesis of BIBD. Isolation
145 of the individual viruses is required to confirm the hypothesis and to study the potential re-
146 assortation between reptarenaviruses.

ACKNOWLEDGEMENTS

This work was supported by the following grant awarding bodies: Academy of Finland (251836 “Arboviruses in Northern Europe” to OV), and the Finnish Foundation of Veterinary Research (to UH).

152 **REFERENCES**

153 1. **Moraz, M. L., and S. Kunz.** 2011. Pathogenesis of arenavirus hemorrhagic fevers. Expert
154 Rev. Anti Infect. Ther. **9:**49-59. doi: 10.1586/eri.10.142.

155 2. **Stenglein, M. D., C. Sanders, A. L. Kistler, J. G. Ruby, J. Y. Franco, D. R. Reavill, F.**
156 **Dunker, and J. L. Derisi.** 2012. Identification, characterization, and in vitro culture of highly
157 divergent arenaviruses from boa constrictors and annulated tree boas: candidate etiological
158 agents for snake inclusion body disease. MBio. **3:**e00180-12. doi: 10.1128/mBio.00180-12;
159 10.1128/mBio.00180-12.

160 3. **Bodewes, R., M. J. Kik, V. S. Raj, C. M. Schapendonk, B. L. Haagmans, S. L. Smits,**
161 **and A. D. Osterhaus.** 2013. Detection of novel divergent arenaviruses in boid snakes with
162 inclusion body disease in The Netherlands. J. Gen. Virol. **94:**1206-1210. doi:
163 10.1099/vir.0.051995-0; 10.1099/vir.0.051995-0.

164 4. **Hetzel, U., T. Sironen, P. Laurinmaki, L. Liljeroos, A. Patjas, H. Henttonen, A. Vaheri,**
165 **A. Artelt, A. Kipar, S. J. Butcher, O. Vapalahti, and J. Hepojoki.** 2013. Isolation,
166 identification, and characterization of novel arenaviruses, the etiological agents of boid
167 inclusion body disease. J. Virol. **87:**10918-10935. doi: 10.1128/JVI.01123-13;
168 10.1128/JVI.01123-13.

169 5. **EC Approved taxonomic proposal "2014.011a-dV.A.v2.reptarenavirus.pdf" awaiting**
170 **ICTV ratification**, downloaded 16th of February 2015 from
171 http://talk.ictvonline.org/files/proposals/taxonomy_proposals_vertebrate1/m/vert04/5108.aspx.

172 6. **EC Approved taxonomic proposal "2014.012aV.A.v3.Arenavirus_ren.pdf" awaiting**
173 **ICTV ratification**, downloaded 16th of February 2015 from
174 http://talk.ictvonline.org/files/proposals/taxonomy_proposals_vertebrate1/m/vert04/5109.aspx.

175 7. **Hepojoki, J., A. Kipar, Y. Korzyukov, L. Bell-Sakyi, O. Vapalahti, and U. Hetzel.** 2015.
176 Replication of Boid Inclusion Body Disease-Associated Arenaviruses Is Temperature Sensitive
177 in both Boid and Mammalian Cells. J. Virol. **89:**1119-1128. doi: 10.1128/JVI.03119-14 [doi].

178 8. **Kallio, M. A., J. T. Tuimala, T. Hupponen, P. Klemela, M. Gentile, I. Scheinin, M.**
179 **Koski, J. Kaki, and E. I. Korpelainen.** 2011. Chipster: user-friendly analysis software for
180 microarray and other high-throughput data. BMC Genomics. **12:**507-2164-12-507. doi:
181 10.1186/1471-2164-12-507 [doi].

182 9. **Langmead, B., and S. L. Salzberg.** 2012. Fast gapped-read alignment with Bowtie 2. Nat.
183 Methods. **9:**357-359. doi: 10.1038/nmeth.1923 [doi].

184 10. **Okonechnikov, K., O. Golosova, M. Fursov, and UGENE team.** 2012. Unipro UGENE:
185 a unified bioinformatics toolkit. Bioinformatics. **28:**1166-1167. doi:
186 10.1093/bioinformatics/bts091 [doi].

187 11. **Bao, Y., V. Chetvernin, and T. Tatusova.** 2014. Improvements to pairwise sequence
188 comparison (PASC): a genome-based web tool for virus classification. Arch. Virol. **159:**3293-
189 3304. doi: 10.1007/s00705-014-2197-x [doi].

190 12. **Damonte, E. B., S. E. Mersich, and C. E. Coto.** 1983. Response of cells persistently
191 infected with arenaviruses to superinfection with homotypic and heterotypic viruses. Virology.
192 **129:**474-478.

193 13. **Schumacher, J., E. R. Jacobson, B. L. Homer, and J. M. Gaskin.** 1994. Inclusion Body
194 Disease in Boid Snakes. Journal of Zoo and Wildlife Medicine. **25:**pp. 511-524.

195 14. **Botten, J., J. Sidney, B. R. Mothe, B. Peters, A. Sette, and M. F. Kotturi.** 2010.
196 Coverage of related pathogenic species by multivalent and cross-protective vaccine design:

197 arenaviruses as a model system. Microbiol. Mol. Biol. Rev. **74:**157-170. doi:
198 10.1128/MMBR.00045-09 [doi].

199 15. **Abascal, F., R. Zardoya, and M. J. Telford.** 2010. TranslatorX: multiple alignment of
200 nucleotide sequences guided by amino acid translations. Nucleic Acids Res. **38:**W7-13. doi:
201 10.1093/nar/gkq291 [doi].

202

FIGURE LEGENDS

FIG 1. Phylogenetic relationships of the newly sequenced BIBDAV. (A) A phylogenetic tree of aligned core polymerase domains of the RNA dependent RNA polymerases (RdRps) of segmented negative-strand viruses. The tree was reconstructed using the maximum likelihood method in Mega 6.06 with 1000 bootstrap replicates. The amino acid substitution model WAG was used as suggested by Mega as the best fitting model. Bootstrap support values >70 are given at the nodes. (B) Maximum likelihood trees were built based on the amino acid sequences of the RdRp, and on the (C) nucleotide sequences of the nucleoprotein. The phylogeny was reconstructed using the MEGA 5.05 software with 1000 bootstraps. The sequence dataset (including representatives of Old and New World arenaviruses, and the previously reported reptarenaviruses) was compiled from the Virus Pathogen Resource (ViPR) database. The nucleotide sequence alignment was guided by amino acid translations using Translator X *(15)* with the MAFFT algorithm. The viruses and their abbreviations: CAS virus (CASV), University of Helsinki virus-1 to -4 (UHV-1 to -4), Boa Av NL B3, Aurora borealis virus-1 to -3 (ABV-1 to -3), Golden Gate virus (GGV), Tavallinen suomalainen mies virus (TSMV), Hans Kompis virus (HKV), Suri Vanera virus (SVaV), University of Giessen virus-1 to -4 (UGV-1 to -4), Haartman Institute snake virus (HISV). The newly sequenced viruses are highlighted in bold.

Table 1. Length, coverage, and sequence identities of the *de novo* constructed BIBDAV S and L segments. The sequences highlighted in yellow and orange are based on PASC analysis and sufficiently distant to represent new species and genus, respectively. The viruses and their abbreviations: CAS virus (CASV), University of Helsinki virus-1 to -4 (UHV-1 to -4), Boa Av NL B3, Aurora borealis virus-1 to -3 (ABV-1 to -3), Golden Gate virus (GGV), Tavallinen suomalainen mies virus (TSMV), Hans Kompis virus (HKV), Suri Vanera virus (SVaV), University of Giessen virus-1 to -4 (UGV-1 to -4), Haartman Institute snake virus (HISV), Amapari virus (AMPV), Bear Canyon virus (BCV).

Origin	Genome (segment)	Length (in nt)	Coverage (min-max)	Reads aligned	PASC identity% to other contigs (BLAST vs. Global alignment)	PASC identity% to database sequences (BLAST vs. Global alignment)	L-segment prevalence in preparation (%)	GenBank accession number
Snake 1	UHV-1 (S)	3393	7071-35430	523677	ABV-2 (74.5% vs. 66.4%)	UHV-1 (99.0% vs. 99.0)		KR870011
	ABV-1 (S)	2725	746-30609	381847	UHV-3 71.8% vs. UHV-2 72.6%	GGV (71.7% vs. 64.7%)		KR870010
	UHV-1 (L)	6834	40-20629	647002	UHV-4 (96.0% vs. 95.4%)	Boa Av NL B3 (78.7% vs. 79.6%)	2.6%	KR870020
	ABV-1 (L)	6892	688-17159	413735	ABV-2 (85.1% vs. 85.0%)	GGV (82.2% vs. 82.4%)	97.4%	KR870021
Snake 5	UGV-1 (S)	3433	1055-79695	1096065	UGV-3 (96.7% vs. 95.7%)	GGV (74.9 vs. 75.6%)		KR870012
	UGV-1 (L)	6787	1310-32480	698732	UGV-3 (81.5% vs. 81.3%)	GGV (68.1% vs. 70.0%)		KR870022
Snake 8	UGV-3 (S)	3455	68-41472	612439	UGV-2 (99.0% vs. 98.9%)	GGV (73.0% vs. 75.8%)		KR870013
	UGV-4 (S)	2774	167-41725	510392	UGV-2 (86.9% vs. 77.9%)	GGV (66.5% vs. 60.8%)		KR870014
	UGV-3 (L)	6830	91-3228	85261	UGV-2 (99.2% vs. 54.0%)	GGV 67.8% vs. Cupixi 54.0%	14.9%	KR870023
	SVaV (L)	6506	289-2547	78412	UGV-1 (69.7% vs. 69.6%)	GGV (65.3% vs. 66.2%)	0.3%	KR870024
	ABV-3 (L)	6641	73-2048	50159	ABV-1 (79.4% vs. 78.3%)	GGV (79.1% vs. 78.1%)	59.4%	KR870025
	TSMV (L)	6928	11-2439	73643	ABV-2 (70.8% vs. 72.7%)	GGV (70.5% vs. 72.3%)	20.2%	KR870026
	UHV-4 (L)	6882	127-3210	69701	UHV-3 (98.3% vs. 98.3%)	Boa Av NL B3 (79.5% vs. 81.5%)	0.6%	KR870027
	HKV (L)	6906	3-1655	37354	ABV-1 (70.3% vs. 72.8%)	GGV (69.5% vs. 73.0%)	4.6%	KR870028
Snake 9	UGV-2 (S)	3465	194-31376	468578	UGV-3 (99.0% vs. 98.9%)	GGV (73.5% vs. 76.0%)		KR870015
	UGV-2 (L)	6903	123-14415	401021	UGV-3 99.2% vs. Sn5 81.3%	GGV (68.7% vs. 71.6%)		KR870029
Snake 11	UHV-2 (S)	3383	37-4780	62970	UHV-3 (93.8% vs. 93.0%)	Boa Av NL B3 (78.5% vs. 80.0%)		KR870016
	HISV (S)	3376	11-27092	350355	UHV-1 15.2% vs. ABV-1 52.1%	Lassa (16.5% vs. 54.0%)		KR870017
	UHV-2 (L)	6894	132-2077	63741	UHV-4 (79.7% vs. 80.0%)	Boa Av NL B3 (70.4% vs. 72.8%)	93.4%	KR870030
	HISV (L)	5913	357-4493	115633	TSMV 15.3% vs. ABV-3-3 50.0%	AMPV 18.8% vs. BCV 48.5%	6.6%	KR870031
Snake 41	ABV-2 (S)	3413	28-3460	42023	UHV-1 74.6% vs. Sn5 76.1%	GGV (79.2% vs. 79.1)		KR870018
	UHV-3 (S)	3339	10-1055	13563	UHV-2 (93.8% vs. 93.0%)	Boa Av NL B3 (80.4% vs. 83.0%)		KR870019
	UHV-3 (L)	6925	3-4572	36844	UHV-4 (98.3% vs. 98.3%)	Boa Av NL B3 (80.4% vs. 81.9%)	99.7%	KR870032
	ABV-2 (L)	6938	111-2966	49692	ABV-1 (85.0% vs. 85.0%)	GGV (82.3% vs. 82.8%)	0.3%	KR870033

A