

New Volvovirus Isolate From Domesticated House Cricket, *Acheta Domesticus* (Orthoptera: Gryllidae)

Somjit Homchan

Naresuan University

Yash Munnalal Gupta (✉ yashmunnalalg@nu.ac.th)

Naresuan University <https://orcid.org/0000-0003-3306-832X>

Research Article

Keywords: *Acheta domesticus*, cricket, evolution, infection, phylogeny, virus, volvovirus isolate

Posted Date: June 18th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-524451/v1>

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

[Read Full License](#)

Abstract

Background

DNA viruses have a broad variety of genetic material. Most viral DNA experiments focused on those with recognized pathogenic characteristics. Herein, we employed hybrid approach to systematically identify viral DNA from the *Acheta domesticus* genome and also curated primer library to reconfirm the infection of *Acheta domesticus* volvoxvirus (AdVVV) from *A. domesticus* samples obtained from breeding facility in Thailand.

Methods and Results

The AdVVV nucleotide sequence was anchored and examined from genome sequence of *A. domesticus*. Subsequently, we sequenced the overlapping amplified DNA to assemble the whole genome of AdVVV isolate. The genome sequence began with the putative nonanucleotide origin of replication (1-TAGTATTAC), and had four open reading frames. The circular nature of AdVVV was confirmed as typical stemloop assembled by complementary initial and ending nucleotides sequences. The newly discovered volvoxvirus isolates from Thailand is also highly homologous (97.34 -98.77 %) with previously identified volvoxvirus sequences and having identical gene organization. This correlation is particularly surprising considering that the identified volvoxvirus considerably mutated compared to previously discovered volvoxvirus isolates.

Conclusions

This research appears to be the first to find volvoxvirus isolates from Thailand with an alternative approach to identify viral isolates. These results widen our knowledge on rapidly evolving volvoxvirus, and provide insights for better identification and the phylogenetic analysis based on different segments of viral genome. Also, the alternative method used in the present study will aid in the detection of DNA viruses and the assembly of their genomes.

1. Introduction

Insect based products have received much attention over the last two decades due to their high nutritional values [1]. Among many edible insects, *Acheta domesticus* (house crickets) gradually becoming popular around Thailand due to their nutty taste and crunchy texture. Numerous breeding facilities are rearing insects due to their high economical value. Usually, these breeding facilities are having enclose vessels for rearing insects. These captive breeding environments are very susceptible to epizootic diseases [2].

A. domesticus is a commercially important species of field cricket and has been dominating insect breeding industry for over two decades in Thailand [3]. The infectious viruses often cause fatal disease triggering mortality and less vitality in the respective host population. The common problem in

commercial insect breeding is mortality or reduction in yield. Cricket rearing needs a healthy atmosphere to protect them from environmental or natural hazards. Preventing insect pathogen entering in the breeding environment is important to ensure health of rearing insect population [4]. Insects are susceptible to the several kinds of microorganism including viruses that can lead mortality to the host insect population like *A. domesticus* densovirus (AdDNV) decimated house cricket in North America in 2009 [5]. Due to the heavy losses in the *A. domesticus* production caused by AdDNV in USA and also in Japan after 2009, the Jamaican cricket (*G. assimilis*) has been introduced in breeding facilities as an alternative cricket for the sustainable production. Besides AdDNV, new viruses have been discovered which infects various species of cricket. In the report of a single-stranded circular DNA volvoxviruses, researchers hypothesized that this ssDNA viruses infects several species of crickets. One of rapidly evolving ssDNA virus is *Acheta domesticus* volvoxvirus (AdVVV). The volvoxvirus isolates were found in samples of dead house crickets from Japan and Jamaican crickets from the USA [6-8]. These rapidly evolving AdVVV isolates are not related to *cycloviruses*, *circoviruses*, *geminiviruses* or *nanoviruses* [9], could belong to new a family or genus.

Screening of pathogen such as viruses in breeding centres can be part of preventive care before it contaminates population and sub-populations. However, screening of large number of individuals from populations requires the optimized and reproducible protocol for specific virus extraction and detection. The detection of pathogen without having prior information about symptoms can be laborious with multi-disciplinary procedures. The known infectious viruses for *A. domesticus* are *A. domesticus* densovirus (AdDNV), *A. domesticus* mini ambidensovirus (AdMADV), and *A. domesticus* volvoxvirus (AdVVV) [10, 11]. These viruses are having DNA as genetic material. Known infectious viral genomes of *A. domesticus* are made of DNA, and it is apparent to find their DNA in the infected host. Recent advancement in metagenomics and bioinformatics has repositioned approaches towards discovery and examination of novel viruses. But, it becomes difficult to classify rapidly evolving viruses with large spectrum host infectivity. Previously, studies were conducted to sequence volvoxvirus isolates (AdVVV-Japan, and AdVVV-Ga) in 2013 by a group of researchers [7], and they suggested that these two volvoxvirus isolates are closely related to previously discovered single stranded DNA virus isolate (AdVVV-IAF) [6]. A recent review literature on the circular DNA virus also mentioned the novel AdVVV and validated that ssDNA viruses are widespread and infects several terrestrial insects and invertebrate species [12]. The four volvoxvirus genomes that have been reported to date, including CrACV-1 [13]. These four volvoxvirus genomes are very identical (>99%) suggesting single viral family. Past studies on volvoxviruses only proposed genome organization of this new family or genus which is rapidly evolving ssDNA virus. However, the mechanism of evolution of this virus is still unexplained and yet to be understood. For AdVVV, this mechanism is needed better understanding for prevention in case of pathogenicity that might cause mortality in breeding facility. Therefore, it is essential to sequence and compare the viral DNA.

Next-generation sequencing technology like HiSeq X Ten is capable of sequencing short DNA reads from the whole-genome. A number of DNA viruses have been identified from host genome [14]. These short reads being assembled into long DNA sequences, which can be good source of DNA sequences come from foreign DNA if host is infected. A number of DNA viruses have been identified from host

genome [14]. In our previous study, we sequenced whole genome of *A. domesticus* to develop microsatellite markers [15]. In present study, we utilised same genome database to detect whether any DNA particles from any known viruses are incorporated in genome. Preliminary *in silico* analysis provided evident that *A. domesticus* infected with AdVVV. We extended study to develop PCR based method to detect AdVVV as secondary confirmation. We developed primer library in order to amplify several DNA fragments. It is important to amplify multiple DNA fragments as AdVVV is mutating and might alter the primer binding sites that lead to unsuccessful amplification. The primer pairs were used to generate PCR products to confirm the viral infection in *A. domesticus* samples which were collected from breeding facilities. The same primer pairs were used to produce DNA fragments with overlapping sequence in for whole viral genome assembly. Unique strategy is used in present research that might serve as new mode of pathogen identification and viral genome sequencing. We have sequenced and annotated genome of newly discovered AdVVV isolate from Thailand, and developed primer library. Additionally, provided investigation also demonstrates phylogenetic analysis with previously discovered AdVVV isolates to facilitate the better classification of volvoxvirus found in domesticated *A. domesticus* from breeding facility. However, these findings might not be entirely representative of the evolutionary mechanism, but provides the backbone for AdVVV identification and the phylogenetic analysis based on different segments of DNA.

2. Material And Methods

2.1 Viral DNA sequence detection

A. domesticus genome assembly was used in present study for viral DNA sequence detection (GenBank accession number: JAAVVF000000000). Further details methodology used for DNA extraction, sequencing can be found in our previous research [15]. The local database was made using NCBI BLAST+ makeblastdb tool [16]. In present study, three viral genomes, *A. domesticus* densovirus (AdDNV), *A. domesticus* mini ambidensovirus (AdMADV), and *A. domesticus* volvoxvirus (AdVVV-Japan) were downloaded from NCBI nucleotide database with accession number NC_004290.1, NC_022564.1 and NC_021074.1, respectively. Single local database of three viral genomes was made using BLAST+ makeblastdb tool [16]. The nucleotide database of *A. domesticus* and virus genomes were compared for sequence similarity search using NCBI BLAST+ blastn [16]. Only one sequence was found to be having high similarity with isolate AdVVV-Japan. The anchored sequence GenBank accession number is JAAVVF010229171.1. BLASTn tool to inspect nucleotide database and BLASTX tool to inspect protein database from National Center for Biotechnology Information (NCBI) [17] were used to search similar sequences for secondary confirmation of nucleotide sequence extraction from whole genome of *A. domesticus* having similarity with viral genome sequence.

Originally, whole genome of *A. domesticus* was sequenced from samples collected from farm located at Nakhon Ratchasima, Thailand [15]. In order to check the viral infection in breeding facility, more samples of *A. domesticus* were collected from the same farm and their DNA extracted to check viral infection using PCR primers. Modified conventional DNA extraction method was used to extract DNA from twelve

A. domesticus individuals. PCR primers were developed using viral query sequence (GenBank accession: KC794540.1) and anchored sequence from whole genome data (GenBank accession: JAAVVF010229171.1). The PCR primers were design using Primer3 (web-based primer designing tool) [18] using default parameters. However, PCR primers were manually selected to produce overlapping DNA fragments. The primer pairs used to check infection and PCR amplification of overlapping DNA fragments is given in Table 1.

2.2 Viral DNA identification

DNA samples of twelve samples of *A. domesticus* were randomly selected from farm located at Nakhon Ratchasima, Thailand were used to inspect viral DNA particles using primer pair VV1. The amplified DNA was verified using gel electrophoresis (1.5 % Agarose gel).

2.3 Viral genome sequencing and assembly

To assemble viral genome and check its circular folding, the same DNA sample from single cricket was used in order to amplify overlapping DNA fragments using all primer pairs. These PCR products with overlapping regions were sequenced by 1st BASE DNA Sequencing Division, Malaysia. The sequenced DNA fragments were analysed and assembled in whole circular genome using CAP3 (*de novo* assembly) [19]. The draft DNA assembly was compared with reference sequence in order to validate the newly assembled viral genome organization.

2.4 Phylogenetic analysis

The sequence analysis was also followed by phylogenetic analysis for complete genome sequence with additional genome sequences from GenBank. Multiple sequence alignment was performed using MUSCLE algorithm. Evolutionary analyses were conducted in MEGA X [20].

3. Results

3.1 *in silico* sequence analysis of viral DNA

Previously published genome assembly of *A. domesticus* (GenBank accession: JAAVVF000000000) used to detect sequence resemble with any of three viral genome sequences which is targeted in present study. *A. domesticus* genome dataset is having contig length from 500 bp to 50,990 bp with an average length of 1,309 bp with zero count of ambiguous nucleotide. The genome sequence quality showed promising insights for viral sequence detection. Therefore, downloaded viral genomes, which are *A. domesticus* densovirus (AdDNV), *A. domesticus* mini ambidensovirus (AdMADV), and *A. domesticus* volvovirus (AdVVV-Japan) were used for similarity search against *A. domesticus* genome. Local nucleotide last result reveals that *A. domesticus* whole genome sequence is consisting viral DNA sequence which is highly analogous with *A. domesticus* volvovirus (AdVVV-Japan) (GenBank accession: NC_021074.1).

The anchored sequence from local nucleotide blast (GenBank accession: JAAVVF010229171.1) showed multiple hits with AdVVV-Japan genome sequence. AdVVV-Japan is single-strand DNA virus with 2517 bp sequence length [7]. Significant sequence similarity in single contig is evident the presence of AdVVV DNA in the genome of *A. domesticus* used in the present study. The length of the matched sequence is 2587 bp and nucleotide sequence that is very close to the sequence of *A. domesticus* volvovirus (AdVVV-Japan), which is 2517 bp. AdVVV-Japan genome sequence and anchored sequence produced three significant alignments with high identity and low gaps in the alignment (Table 2).

Subsequently, Anchored sequence was extracted from the database for further examination using NCBI server with BLAST for nucleotide and protein sequence similarity search. The nucleotide BLAST shows the topmost similar sequences are virus nucleotide sequences, which are *A. domesticus* volvovirus isolates and having percentage identity between 82.21 to 82.45 % (Table 3). The protein BLAST reveals that (GenBank accession: JAAVVF010229171.1) query sequence is having multiple hits with *A. domesticus* volvovirus protein sequences. The protein BLAST was restricted to the protein sequences assigned for AdVVV-Japan (NC_021074.1). Result revealed that query is covering all four hypothetical proteins (GenBank accession: YP_007878130.1, YP_007878129.1, YP_007878131.1, YP_007878128.1). The hypothetical proteins alignment having accession YP_007878128.1 is divided into two segments located in the starting and in the end part of the query. This indicates the circularity of the anchored nucleotide fragment.

The identified nucleotide sequence conspicuous the presence of AdVVV in the house cricket sample which was used for whole genome sequencing. The nucleotide and protein blast also showed a strong correlation and authenticated that the viral DNA is also pooled into whole genome sequence of *A. domesticus*. The *A. domesticus* DNA sample used for whole genome sequencing was obtained from insect breeding facility, that leads to higher probability of AdVVV infection in breeding population. Therefore, further analysis was necessary to substantiate the presence of AdVVV in breeding population of *A. domesticus* from Nakhon Ratchasima farm.

3.2 *in vitro* sequence analysis of viral DNA

Initially, only one primer pair was employed to check presence of viral DNA in total DNA extracted from twelve house cricket samples using PCR technique. The primer pair VV1 successfully amplified DNA fragments with expected size, which has proven AdVVV infection in breeding facility (Fig. 1). Results reported by Pham et al. in 2013a suggests that AdVVV is rapidly evolving single stranded DNA virus and have different strains. In present study, we also encounter mismatches in sequence alignments (see Table 2). Moreover, we also found gaps in alignment that could be due to sequencing or sequencing error. Hence, we developed primer pairs targeted for viral DNA to generate overlapping fragments that can be sequenced and assembled into full genome of AdVVV isolate from Thailand. The assembled genome is deposited to NCBI genome database under GenBank accession: MW288623. The Thai isolate, as we named AdVVV-Thailand is containing 2,516 nucleotides. The newly assembled genome is having same architecture compared with other AdVVV isolates (Fig. 2a). Genome contains four open reading frame

(ORFs) of potential coding sequences (CDS). CDSs are previously assigned for four hypothetical proteins. According to reference genome sequence of AdVVV (Accession number: NC_021074 identical to KC794540), size of these four protein coding genes is the same with CDS present in newly assembled AdVVV isolate from Thailand and also coding for same length of peptides (Table 4). The hairpin structure (stemloop) producing sequence contains thirty-two nucleotides (at the end of genome 2504-2516, and at the beginning of genome 1-19).

We also designed primer pair (VV8, see Table 1) to obtain PCR product containing the single nucleotide fragment which is responsible for formation of hairpin structure with putative nonanucleotide origin of replication

(1nt, 5'-TAGTATTAC-3') (Fig. 2b). This evolutionary conserved sequence found in other viral genome as recognition site for the replication protein [21]. The primer pair VV8 successfully mapped origin of replication in PCR product of 315 bp fragment. The origin of replication and ending part of viral genome sequence adjacent to each other in sequenced fragment. The newly assembled genome sequence was utilized to conduct phylogenetic analysis with other isolates of AdVVV. The phylogram shows the scale based on genetic change in AdVVV isolates (Fig. 3). Three AdVVV isolates clustered together, AdVVV-IAF, Circular virus 1, and AdVVV-GA, respectively. AdVVV-Japan and AdVVV-Thailand are genetically more related to each other than cluster of other three isolates, However, AdVVV-Thailand isolate is distanced from all four isolates implied by higher observed sequence change in genome sequence.

Further analysis was conducted using four gene sequences of AdVVV. Each gene sequence from five AdVVV isolates were aligned to distinguish them on bases of variable sites, singleton sites, and parsimony informative sites (Table 5). Most of genetic variation are singleton type. The test revealed no significant difference in ORF4, ORF2, and ORF3 sequences compare to ORP1 sequences. 85.33 % variable sites were constrained to ORF1. As ORF1 found to be most variable gene, the phylogram was reconstructed to ensure genetic distance (Fig 4). The phylogram generated by employing ORF1 gene sequences is happened to be very identical with phylogram generated using whole genome sequences.

4. Discussion

This study appears to be the first to find AdVVV isolate from Thai *A. domesticus* sample. We used whole genome sequence of *A. domesticus* to find viral DNA [15]. In previous research of RNA viruses in human and invertebrates, researchers used bioinformatic approach to detect viral transcripts from transcriptome [22, 23]. In present research we used genome sequence and reference viral DNA sequence to detect the viral DNA. Present *In silico* approach made virus detection easier and less laborious. However, the assembling the whole virus genome of AdVVV isolate was need. The anchored sequence of viral DNA has confirmed the presence of viral DNA in genome, but it was inadequate for assembling the whole viral genome. Moreover, the *A. domesticus* sample used for genome sequencing was collected from breeding facility, systematically, we also collected samples from same breeding facility to verify the infection. Our result provided the first proof of infection in Thai breeding facility. Additionally, PCR primer used in present study will also facilitate volvoxvirus detection from cricket species as AdVVV has potency

to infect several cricket species [7]. Additionally, researchers found AdVVV isolates from dead cricket samples [6], but the pathogenicity of AdVVV isolate is not evaluated till the date due to limited knowledge of this evolving virus. AdVVVs are single stranded DNA viruses, could transmit vertically to their predators. Recently, researchers found single stranded DNA virus in a spider species, and they suggested vertical transmission of viral particle from prey to the predators [24]. Thus, vertical transmission of virus rises question of food security as crickets commonly used as food and feed. Besides vertical transmission to other species, virus can transmit within species through sexual behaviour [25] or cannibalism [26]. Furthermore, recent research proposed that breeding facilities in Thailand exchanges cricket eggs to maintain genetic variation, in such case, virus spread becomes pervasive in inbreeding population [15].

The size of AdVVV-Thailand is 2,516 bp. Previously found volvoxvirus isolates have two genome sizes, 2,516 bp (AdVVV-GA, Cricket associated circular virus 1) and 2,517 bp (AdVVV-IAF, AdVVV-Japan). We utilized whole genome sequences of all five isolates in phylogenetic analysis to track the evolutionary the history. Result showed that Cricket associated circular virus 1 is positioned between AdVVV-IAF and AdVVV-GA; which is also a volvoxvirus isolate. Three genome which is clustered together were isolated from crickets (samples from America). The Japanese AdVVV isolate is separated from the clustered AdVVV isolates and related Thai AdVVV isolate found in present study. Consequently, the polygenetic evaluation is also comparable with geographical location from where crickets were collected and used for AdVVV genome sequencing, where clustered isolates are from America, Thai isolates is neighbored with Japanese isolate. After conducting variation analysis on four CDS from five genomes, ORF1 found to be most variable CDS which is previously designated for hypothetical protein (GenBank: AGJ03170.1)/putative capsid protein (GenBank accession: AXL65914.1). Our study suggests that AdVVV genome (ORF2, 3 & 4) is considerably conserved compared to diversity observed in ORF1. The construction of the ssDNA virus group is well conserved, but genomes may have high genetic variation [27]. Secondly, ssDNA virus genome replication is well conserved compared to capsid protein sequence [28]. On other hand, ORF2 found to be well conserved compared to other ORFs from five isolates of AdVVV. BLASTx was used to examine the conserve nature of OFR2 with other viruses. According to BLASTx (limited to viruses) result, ORF2 showed alignments with sequences designated for replication associated protein or putative Rep protein for ssDNA viruses. The sequences coverage ranged from 98-88% and percentage identity ranged from 42.91- 28.36% (AdVVV alignments were excluded). The majority of ssDNA viruses are recognised on bases of conserved segment that codes for Rep protein, and structure proteins are not optimal due to high diversity [27]. However, highly conserved DNA segments would not be an ideal choice for evolutionary analysis within family or genera, where conserved regions are identical but structural proteins are variable. In case of AdVVV isolates, we utilised ORF1 to established phylogenetic relation among five isolates, including newly discovered AdVVV isolate in present study.

5. Conclusion

Our study suggests that, OFR2 would be optimal segment of genome to recognise AdVVV isolates, but ORF1 would be better for phylogenetic analysis among AdVVV isolates as phylogram produced using

whole genome sequence and ORF1 are nearly identical. Phylogram revealed that three isolates clustered in one clade and other branched into taxon. Accordingly, we proposed that volvovirus could be the family of rapidly evolving ssDNA virus, which reasoning is also supported by variation analysis of AdVVV isolates. Due to limited studies on AdVVV, researchers were perplexed to determine the status to separate these viruses into family or genus [6], and in our study we attempt to clear ambiguity. We intensively utilized genome and ORF segments for classification study like previously utilized to classify family of *Genomoviridae* into nine genera [29]. Thus, it is important and necessary to classify these cricket viruses for viral taxonomy and for better understanding of AdVVV evolutionary mechanism in case of virulent strain present/evolve in environment, which can cause mortality in cricket breeding facility.

Declarations

Data availability: <https://www.ncbi.nlm.nih.gov/nucleotide/MW288623>

Funding: This research was supported by Department of Biology, Faculty of Science, Naresuan University, Thailand (Project code: R2564E073).

Author contributions

All authors contributed to the study conception and design. Material preparation, insect sample collection and analysis were performed by YMG and SH. The first draft of the manuscript was written by YMG and corrections were made and approved by SH. All authors read and approved the final manuscript.

Acknowledgments

The authors hereby acknowledge Department of Biology, Faculty of Science, Naresuan University, Thailand for awarding research grant and also for their laboratory facilities.

Conflicts of interest: The best of our knowledge, the named authors have no conflict of interest, financial or otherwise.

Ethics approval

This study was approved by the Division of Research and Innovation, Naresuan University, Thailand. All of the experimental procedures involving insects were conducted in accordance with the Center for Animal Research, Naresuan University, Thailand (Project number: NU-AE610835, Certificate document number: 62 01 004).

Consent to Participate

There are no human subjects in this article and consent to participate is not applicable.

Consent for publication: All authors consent to publish this work.

References

1. Belluco S, Losasso C, Maggioletti M, Alonzi CC, Paoletti MG, Ricci A (2013) Edible insects in a food safety and nutritional perspective: a critical review. *Comprehensive reviews in food science and food safety* 12(3):296-313. <https://doi.org/10.1111/1541-4337.12014>
2. Snyder NF, Derrickson SR, Beissinger SR, Wiley JW, Smith TB, Toone WD, Miller B (1996) Limitations of captive breeding in endangered species recovery. *Conservation biology* 10(2):338-348.
3. Halloran A, Roos N, Flore R, Hanboonsong Y (2016) The development of the edible cricket industry in Thailand. *Journal of Insects as Food and Feed* 2(2):91-100. <https://doi.org/10.3920/JIFF2015.0091>
4. Eilenberg J, Vlak J, Nielsen-LeRoux C, Cappellozza S, Jensen AB (2015) Diseases in insects produced for food and feed. *Journal of Insects as Food and Feed* 1(2):87-102. <https://doi.org/10.3920/JIFF2014.0022>
5. Szelei J, Woodring J, Goettel M, Duke G, Jousset F-X, Liu K, Zadori Z, Li Y, Styer E, Boucias D (2011) Susceptibility of North-American and European crickets to *Acheta domesticus* densovirus (AdDNV) and associated epizootics. *Journal of invertebrate pathology* 106(3):394-399. <https://doi.org/10.1016/j.jip.2010.12.009>
6. Pham HT, Bergoin M, Tijssen P (2013) *Acheta domesticus* volvoxvirus, a novel single-stranded circular DNA virus of the house cricket. *Genome Announcements* 1(2). <https://doi.org/10.1128/genomeA.00079-13>
7. Pham HT, Iwao H, Bergoin M, Tijssen P (2013) New volvoxvirus isolates from *Acheta domesticus* (Japan) and *Gryllus assimilis* (United States). *Genome announcements* 1(3). <https://doi.org/10.1128/genomeA.00328-13>
8. Weissman DB, Gray DA, Pham HT, Tijssen P (2012) Billions and billions sold: pet-feeder crickets (Orthoptera: Gryllidae), commercial cricket farms, an epizootic densovirus, and government regulations make for a potential disaster. *Zootaxa* 3504(1175-5326):67-88. <https://doi.org/10.11646/zootaxa.3504.1.3>
9. Tijssen P, Péntzes JJ, Yu Q, Pham HT, Bergoin M (2016) Diversity of small, single-stranded DNA viruses of invertebrates and their chaotic evolutionary past. *Journal of invertebrate pathology* 140:83-96. <https://doi.org/10.1016/j.jip.2016.09.005>
10. Maciel-Vergara G, Ros VI (2017) Viruses of insects reared for food and feed. *Journal of Invertebrate Pathology* 147:60-75. <https://doi.org/10.1016/j.jip.2017.01.013>

11. Fernandez-Cassi X, Supeanu A, Vaga M, Jansson A, Boqvist S, Vagsholm I (2019) The house cricket (*Acheta domesticus*) as a novel food: a risk profile. *Journal of Insects as Food and Feed* 5(2):137-157. <https://doi.org/10.2903/j.efsa.2018.e16082>
12. Dayaram A, Potter KA, Pailes R, Marinov M, Rosenstein DD, Varsani A (2015) Identification of diverse circular single-stranded DNA viruses in adult dragonflies and damselflies (Insecta: Odonata) of Arizona and Oklahoma, USA. *Infection, Genetics and Evolution* 30:278-287. <https://doi.org/10.1016/j.meegid.2014.12.037>
13. Rosario K, Mettel KA, Benner BE, Johnson R, Scott C, Yousseff-Vanegas SZ, Baker CC, Cassill DL, Storer C, Varsani A (2018) Virus discovery in all three major lineages of terrestrial arthropods highlights the diversity of single-stranded DNA viruses associated with invertebrates. *PeerJ* 6:e5761. <https://doi.org/10.7717/peerj.5761>
14. Amirian ES, Bondy ML, Mo Q, Bainbridge MN, Scheurer ME (2014) Presence of viral DNA in whole-genome sequencing of brain tumor tissues from the cancer genome atlas. *Journal of virology* 88(1):774-774. <https://doi.org/10.1128/jvi.02725-13>
15. Gupta YM, Tanasarnpaiboon S, Buddhachat K, Peyachoknagul S, Inthim P, Homchan S (2020) Development of microsatellite markers for the house cricket, *Acheta domesticus* (Orthoptera: Gryllidae). *Biodiversitas Journal of Biological Diversity* 21(9). <https://doi.org/10.13057/biodiv/d210921>
16. Cock PJ, Chilton JM, Grüning B, Johnson JE, Soranzo N (2015) NCBI BLAST+ integrated into Galaxy. *Gigascience* 4(1):s13742-13015-10080-13747. <https://doi.org/10.1186%2Fs13742-015-0080-7>
17. Wheeler DL, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Edgar R, Federhen S (2007) Database resources of the national center for biotechnology information. *Nucleic acids research* 36(suppl_1):D13-D21. <https://doi.org/10.1093/nar/gkq1172>
18. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3 –new capabilities and interfaces. *Nucleic acids research* 40(15):e115-e115. <https://doi.org/10.1093/nar/gks596>
19. Huang X, Madan A (1999) CAP3: A DNA sequence assembly program. *Genome research* 9(9):868-877. <https://doi.org/10.1101/gr.9.9.868>
20. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular biology and evolution* 35(6):1547-1549. <https://doi.org/10.1093/molbev/msy096>
21. Mankertz A, Persson F, Mankertz J, Blaess G, Buhk H-J (1997) Mapping and characterization of the origin of DNA replication of porcine circovirus. *Journal of virology* 71(3):2562-2566. Mankertz, A., Persson, F., Mankertz, J., Blaess, G., & Buhk, H. J. (1997). Mapping and characterization of the origin of

DNA replication of porcine circovirus. *Journal of virology*, 71(3), 2562–2566.

<https://doi.org/10.1128/JVI.71.3.2562-2566.1997>

22. Bhuvaneshwar K, Song L, Madhavan S, Gusev Y (2018) viGEN: An open source pipeline for the detection and quantification of viral RNA in human tumors. *Frontiers in microbiology* 9:1172.

<https://doi.org/10.3389/fmicb.2018.01172>

23. Cheng R-L, Li X-F, Zhang C-X (2021) Novel Dicistroviruses in an Unexpected Wide Range of Invertebrates. *Food and Environmental Virology*:1-9. <https://doi.org/10.1007/s12560-021-09472-2>

24. Rosario K, Mettel KA, Greco AM, Breitbart M (2019) Prevalence of a vertically transmitted single-stranded DNA virus in spinybacked orbweavers (*Gasteracantha cancriformis*) from Florida, USA. *Journal of General Virology* 100(8):1253-1265. <https://doi.org/10.1099/jgv.0.001293>

25. Adamo SA, Kovalko I, Easy RH, Stoltz D (2014) A viral aphrodisiac in the cricket *Gryllus texensis*. *Journal of Experimental Biology* 217(11):1970-1976. <https://doi.org/10.1242/jeb.103408>

26. Adamo SA, Bartlett A, Le J, Spencer N, Sullivan K (2010) Illness-induced anorexia may reduce trade-offs between digestion and immune function. *Animal Behaviour* 79(1):3-10.

<http://doi.org/10.1016/j.anbehav.2009.10.012>

27. Rosario K, Duffy S, Breitbart M (2012) A field guide to eukaryotic circular single-stranded DNA viruses: insights gained from metagenomics. *Archives of virology* 157(10):1851-1871.

<https://doi.org/10.1007/s00705-012-1391-y>

28. Kazlauskas D, Dayaram A, Kraberger S, Goldstien S, Varsani A, Krupovic M (2017) Evolutionary history of ssDNA bacilladnaviruses features horizontal acquisition of the capsid gene from ssRNA nodaviruses. *Virology* 504:114-121. <https://doi.org/10.1016/j.virol.2017.02.001>

29. Varsani A, Krupovic M (2017) Sequence-based taxonomic framework for the classification of uncultured single-stranded DNA viruses of the family *Genomoviridae*. *Virus evolution* 3(1):vew037.

<https://doi.org/10.1093/ve/vew037>

Tables

Table 1 Primer pairs used to detect viral DNA and whole viral genome sequencing and assembly

Primer pair	Primer sequence	Â°C	Fragment size (bp)
VV1-F	ATTACCAGTGCTGTGCGTCT	55.4	368
VV1-R	CGGAGGAAACGGATCAAAAGT		
VV8-F	AAGTCTGTTTGTTCGGGGC	58.3	343
VV8-R	CTGTAGCACCCCAACCAATAG		
LVV1-F	TGGACTGTTGGTAAAGCTGGA	58.4	737
LVV1-R	AGGCTCAGGATAACGTGAC		
LVV2-F	CGTGGGGCTAGAGAGTTTAGT	57.3	827
LVV2-R	TGCTCTAAAACCCCAACAGG		
LVV3-F	ACTCCCGATGCTGTTGAGAT	57.5	728
LVV3-R	AGCGGGGACAGGAAAATCAT		
LVV4-F	ATTGAGACGTGGATGGGGTT	58.5	721
LVV4-R	TCATGTTTGTATACCCGGTGTC		

F, forward; R, reverse; Â°C, annealing temperature.

Table 2 Significant alignment description of AdVVV-Japan (NC_021074.1) against anchored sequence (JAAVVF010229171.1)

Significant pairwise alignments	Identity	Length	Mismatch	Gap	E-value	Table 3 Top four sequence alignment with a query sequence with
1	82.449	1282	167	21	0.0	
2	72.396	797	170	13	4.74e-124	
3	84.774	243	35	1	3.42e-75	

zero error values

Accession	Max Score	Total Score	Query Cover	E value	Per. Ident
KC794539.1	1261	1965	88%	0	82.28%
MH545526.1	1257	1960	88%	0	82.21%
KC794540.1*	1256	1959	87%	0	82.45%
KC543331.1	1255	1963	88%	0	82.22%

Note*: KC794540.1 is the accession for AdVVV-Japan, the reference sequence accession for AdVVV-Japan is NC_021074.1

Table 4 Open reading frame organization in genome of AdVVV-Thailand

ORFs	Strand	Start	Stop	Length (nt aa)
ORF4	+	70	462	393 130
ORF1	+	447	1532	1086 361
ORF2	-	2444	1632	813 270
ORF3	-	2392	1769	624 207

Note: Start codon is 'ATG'

Table 5 Variable site in nucleotide sequence of four CDS from five genomes of AdVVV isolates

CDS	ORFs	V	Pi	S	Alignment length
Hypothetical protein	ORF4	5	1	4	393
Hypothetical protein/Putative capsid protein	ORF1	64	6	58	1086
Hypothetical protein/Replication-associated protein	ORF2	3	-	3	813
Hypothetical protein	ORF3	3	-	3	624

CDS, coding sequence; ORF, open reading frame; V, variable site; Pi, parsimony informative site; S, singleton site

Figures

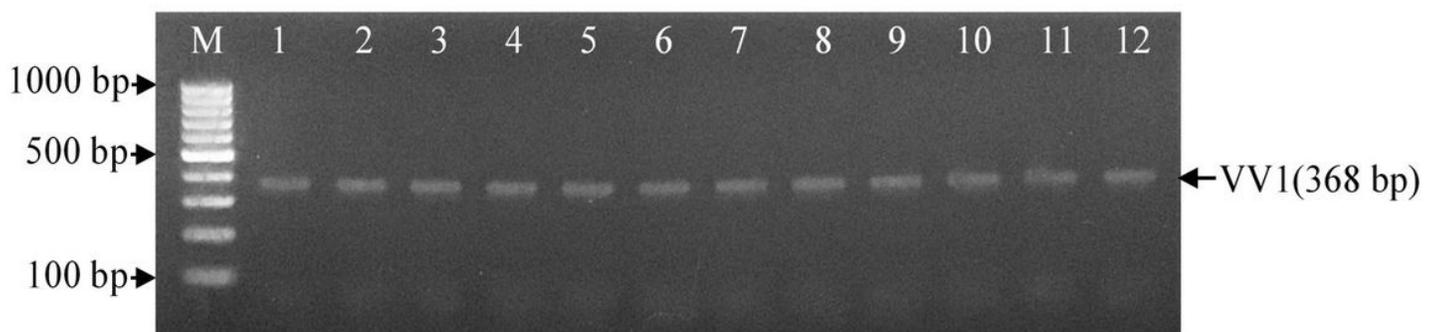


Figure 1

Agarose gel electrophoresis of amplified DNA fragments from twelve different *A. domesticus* DNA samples using primer pair VV1 (Lane:1-12)

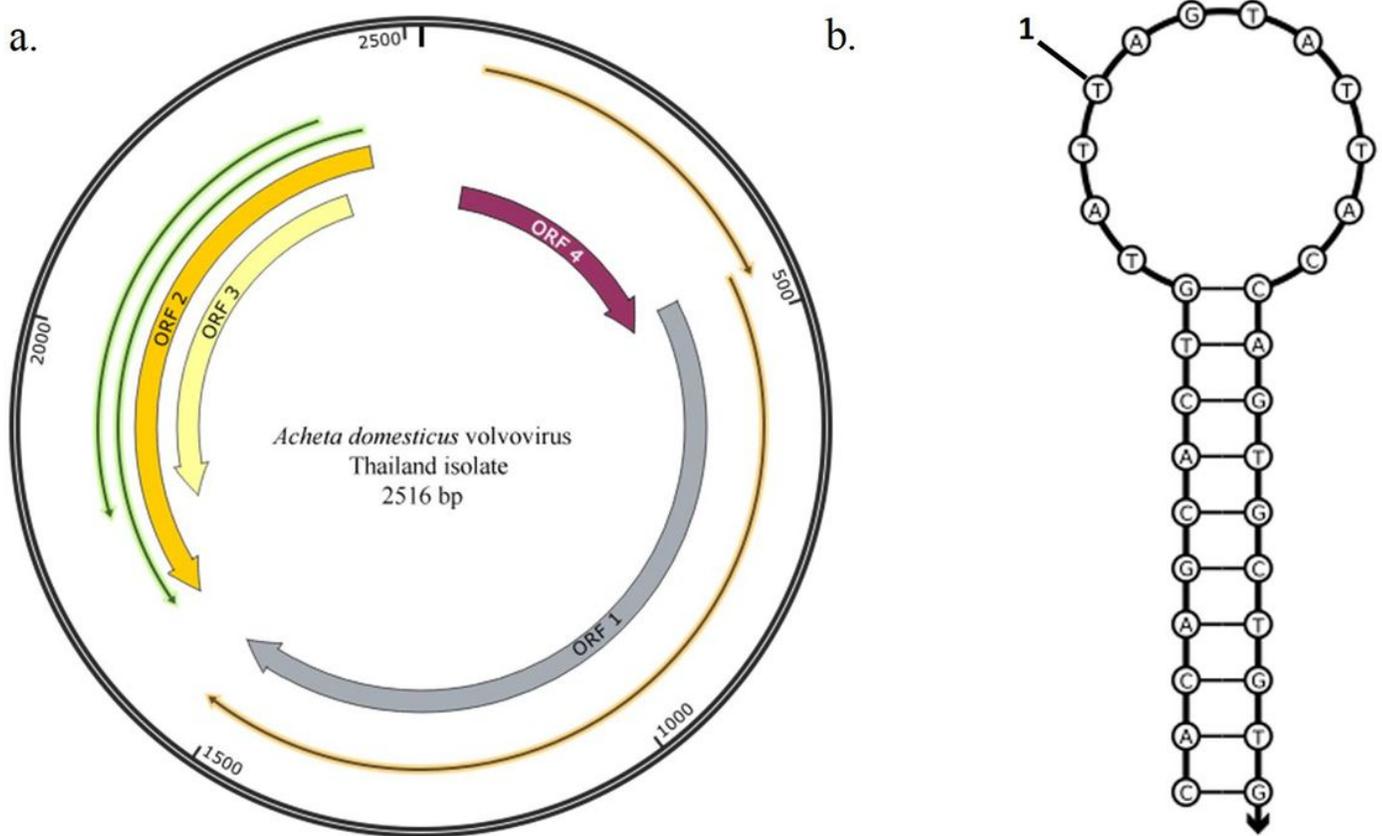


Figure 2

Genome map of newly sequenced AdVVV-Thailand isolate. (a.) AdVVV-Thailand isolate is 2,516 nucleotides in size and contains four ORFs. ORF4 and ORF1 are in the sense direction, starting at nucleotide 70 and 447, respectively. ORF2 and ORF3 are in antisense direction, starting at nucleotide 2444 and 2392, respectively. (b.) Putative nonanucleotide origin of replication (5'-TAGTATTAC-3') is located between ORF2 and ORF4, at the beginning of the genome. The typical hairpin loop structure consists (5'-TAGTATTAC-3'), which starts at 1st nucleotide of the genome sequence.

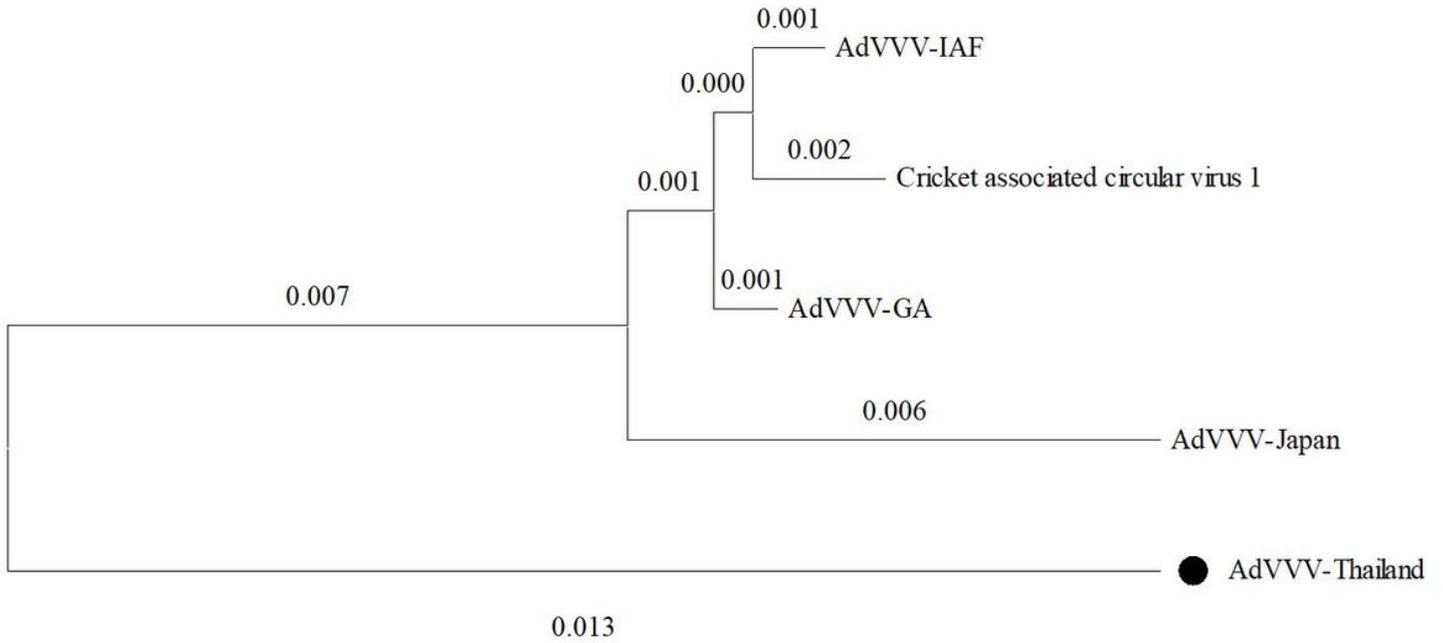


Figure 3

Evolutionary relationships of AdVVV isolates. The evolutionary history was inferred using the Neighbor-Joining method. The tree is drawn to scale, with branch lengths (next to the branches) in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Jukes-Cantor method and are in the units of the number of base differences per site. This analysis involved 5 nucleotide sequences AdVVV isolates. There were a total of 2517 positions in the final dataset. Newly sequenced genome of AdVVV isolate Thailand indicated with black dot.

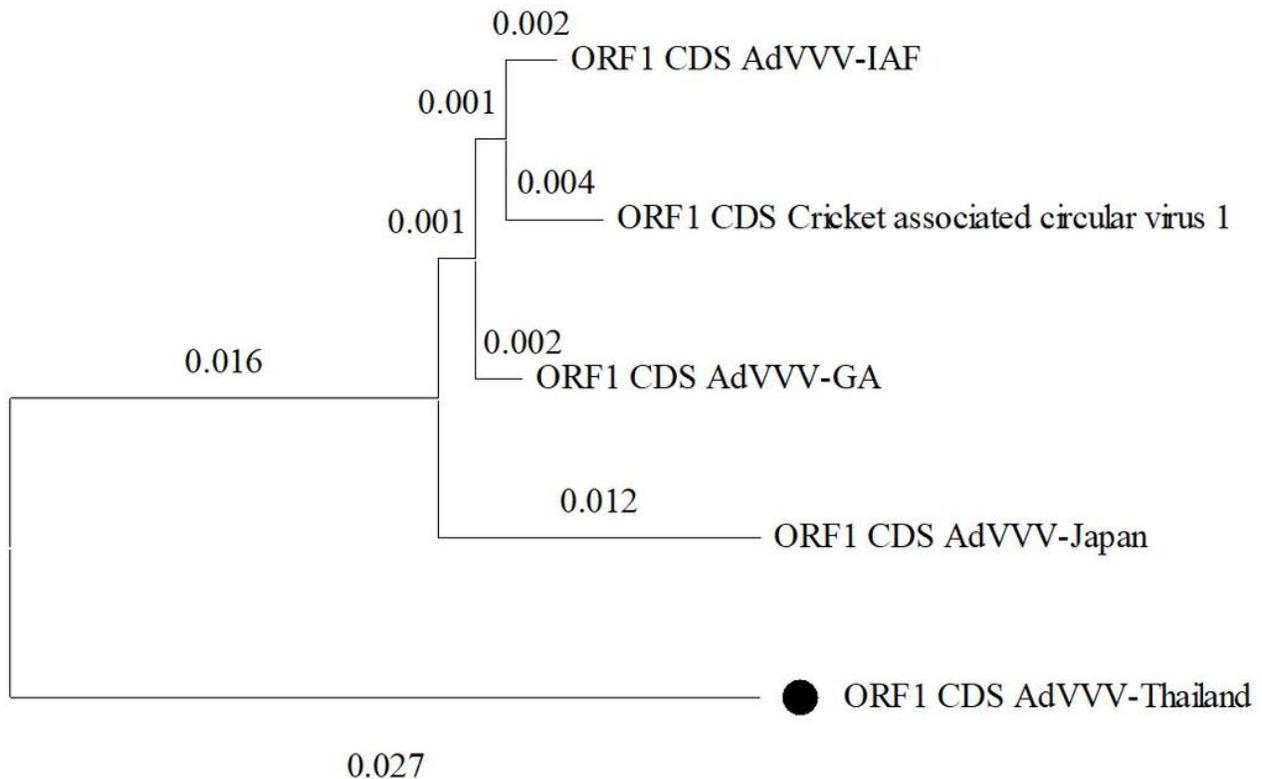


Figure 4

The evolutionary history was inferred using the Neighbor-Joining. Phylogram generated using 5 nucleotide sequences of AdVVV isolates (ORF1, CDS of Hypothetical protein/Putative capsid protein). There were a total of 1086 positions in the final dataset.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Visualabstractworkflow2.pdf](#)