Genetic Analysis of Israel Acute Paralysis Virus: Distinct Clusters Are Circulating in the United States


Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, New York; Department of Entomology, Pennsylvania State University, University Park, Pennsylvania; Bee Research Laboratory, U.S. Department of Agriculture-Agricultural Research Service, Beltsville, Maryland; Bureau of Plant Industry-Apiculture, Pennsylvania Department of Agriculture, Harrisburg, Pennsylvania; Department of Entomology, Hebrew University of Jerusalem, Rehovot, Israel; and Center for Infectious Disease Dynamics, Department of Biology, Pennsylvania State University, University Park, Pennsylvania

Received 4 February 2008/Accepted 14 April 2008

Israel acute paralysis virus (IAPV) is associated with colony collapse disorder of honey bees. Nonetheless, its role in the pathogenesis of the disorder and its geographic distribution are unclear. Here, we report phylogenetic analysis of IAPV obtained from bees in the United States, Canada, Australia, and Israel and the establishment of diagnostic real-time PCR assays for IAPV detection. Our data indicate the existence of at least three distinct IAPV lineages, two of them circulating in the United States. Analysis of representatives from each proposed lineage suggested the possibility of recombination events and revealed differences in coding sequences that may have implications for virulence.

Over the winter of 2006 and 2007, an estimated 23% of all beekeeping operations in the United States experienced losses of hives attributed to colony collapse disorder (CCD) (30). More than 90 cultivated crops depend on the honey bee, Apis mellifera, for pollination; thus, CCD has profound implications for the food supply (23). We recently reported that the presence in hives of the dicistrovirus Israel acute paralysis virus (IAPV) was strongly correlated with the presence of CCD (8). IAPV was first described in 2004 in Israel (21), where infected bees presented with shivering wings, progressed to paralysis, and then died outside the hive. IAPV has features comparable to those of members of the family Dicistroviridae of the superfamily Picornaviridae. Unlike members of the Picornaviridae, which have a single open reading frame (ORF) encoding a single polyprotein, viruses in the family Dicistroviridae have two ORFs encoding two polyproteins. Dicistroviruses have two internal ribosomal entry sites (IRES), one found in the 5′ untranslated region (UTR) and the other located in the intergenic region between ORF1 and ORF2. Other viruses known to infect honey bees, and related to picornaviruses, are sacbrood virus, deformed wing virus, acute bee paralysis virus (ABPV), Kashmir bee virus (KBV), and black queen cell virus (4, 7).

Given the importance of honey bees as pollinators and the reported association between CCD and IAPV, we pursued phylogenetic analysis of geographically discrete IAPV isolates. We report the complete genome sequences of representatives from each of three lineages of IAPV and demonstrate through detailed analysis of four regions of the viral genome (the 5′ UTR and 5′ terminus of ORF1; the 3′ terminus of ORF1, the intergenic region, and the 5′ terminus of ORF2; ORF2; and the RNA-dependent RNA polymerase [RdRp]) the presence on three continents of at least three clusters of IAPV, two of them currently circulating in the United States.

MATERIALS AND METHODS

Sample collection and RNA extraction. Adult bees were collected from migratory beekeeping operations in the United States with CCD. Diseased apiaries were identified based on evidence of recent collapse of colonies within the apiary and a lack of dead bees in the collapsed colonies. From each colony, 150 adult bees were collected and stored at −80°C until they were processed. All adult bee samples were collected from two brood nest combs shaken into a plastic tub. Samples from imported Australian bees were collected directly from the shipping packages and stored at −80°C until they were used. Samples from Israel were collected in 2007 from normal or ailing colonies. Samples from New Brunswick and British Columbia, Canada, were collected in 2007 from ailing colonies. Individual bees were ground in liquid nitrogen and extracted using Tri-Reagent (Molecular Research Center Inc., Cincinnati, OH).

Consensus PCR for phylogenetic and diagnostic analyses. Primers for diagnostic purposes were designed using Greene SCPrimer software (http://scprimer.cpmc.columbia.edu/SCPrimerApp.cgi) (15). The program uses a greedy algorithm to identify the most conserved sequences, creating the minimum set of primers needed for amplification of all sequences in the alignment. Primers are selected within standard design constraints (melting temperature, 55°C to 65°C; GC content, 40% to 60%; no hairpins); degenerate positions are introduced where necessary. The primers in conserved regions of the RdRp and ORF2 regions of the IAPV genome used for diagnosis of IAPV infection are shown in Table 1.

Diagnostic assays were standardized using synthetic DNA standards containing corresponding IAPV sequences cloned into pGEM-T Easy vectors (Promega, Fermentas, Lithuania). After quantitation by UV spectrometry, DNA standards were serially diluted in 2.5 μg/ml human placental DNA (Sigma) and subsequently used to assess the sensitivity of the generated primer sets via SYBR green real-time PCR (Applied Biosystems, Foster City, CA) (Fig. 1). Specificity was determined using a set of 22 IAPV-positive and 17 IAPV-negative adult bee samples (5 of which were also KBV positive). Five microliters of cDNA were used in a 25-μl 1× SYBR green Real-Time PCR Mix (Applied Biosystems) containing 0.3 μM (each) of the forward and reverse primers. The
Sybr green signals were measured for duplicate samples and standards for 44 cycles of 95°C for 15 s and 60°C for 1 min.

Sequence analysis. PCR products were size fractionated by agarose gel electrophoresis, purified (QIAquick Gel Extraction), and sequenced on both strands using an ABI Prism dye terminator cycle-sequencing ready-reaction kit (Applied Biosystems). Sequence data were analyzed with CHROMAS software and compared and aligned to other samples or DNA database sequences using CLUSTAL X, version 1.83 (28). Phylogenetic analysis was performed using the

Table 1. IAPV primer sets used in analysis

<table>
<thead>
<tr>
<th>IAPV primer set</th>
<th>Sequence (5’—3’)</th>
<th>Orientation</th>
<th>Region amplified on DVE31</th>
<th>Purpose or use</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF2-F8092</td>
<td>CCAGCCGTGAACCATGTTTCTTACC</td>
<td>Sense</td>
<td>8042–8267</td>
<td>SYBR green diagnostic assay</td>
</tr>
<tr>
<td>ORF2-R8318</td>
<td>ACATAGTGCACGCCAATACGGAAC</td>
<td>Antisense</td>
<td>5525–5661</td>
<td>SYBR green diagnostic assay</td>
</tr>
<tr>
<td>IAPVrDrp-F4b</td>
<td>AATGGAACCATGATTTTYTCWATGCT</td>
<td>Sense</td>
<td>7969–8082</td>
<td>SYBR green diagnostic assay</td>
</tr>
<tr>
<td>IAPVrDrp-R111</td>
<td>CGAAGGTTTTAATCATCCGACTGTC</td>
<td>Antisense</td>
<td>5477–5880</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV_B4SO427_R130</td>
<td>GCATCGAGTCCCTTCAGGAGGT</td>
<td>Antisense</td>
<td>6205–6971</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV-F-RdRp</td>
<td>GTGTCAGGCGATTGAGAAAGT</td>
<td>Sense</td>
<td>7854–8693</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV-R-RdRp</td>
<td>GTTGGCCTGGTGTGATCAT</td>
<td>Antisense</td>
<td>144–1142</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV_I1R_F</td>
<td>CGATGAACAACGGAAAGGTTT</td>
<td>Sense</td>
<td>100–180</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV_I1R_R</td>
<td>TCACCCATAGGCTGGTGG</td>
<td>Antisense</td>
<td>600–1400</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV_Capsid1_F</td>
<td>GGTCACACCCACCGAATCAAGG</td>
<td>Sense</td>
<td>100–180</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV_Capsid1_R</td>
<td>CAGTCGTCAGGCTGGTGGGT</td>
<td>Antisense</td>
<td>600–1400</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV66F</td>
<td>CGACATTAGTTAGATTTAACATTACACGZ’Z’</td>
<td>Sense</td>
<td>100–180</td>
<td>Phylogenetic analysis</td>
</tr>
<tr>
<td>IAPV1065R</td>
<td>TTCTTACATCCCTCTGGAAGGZ’Z’</td>
<td>Antisense</td>
<td>600–1400</td>
<td>Phylogenetic analysis</td>
</tr>
</tbody>
</table>

FIG. 1. IAPV Sybr green real-time PCR assay. (A) IAPV standards assayed at concentrations ranging from 500,000 to 5 genome copies for both the ORF2 and RdRp regions. (B) Standard curve of the IAPV real-time SYBR green PCR assay performed with 10-fold dilutions of IAPV DNA standards (indicated on the x axis); the corresponding threshold cycles (Ct) are presented on the y axis. Duplicate amplifications of each dilution were tested. The correlation coefficient (R²), y intercept, and slope were calculated and are shown for both assays. Rn, differential normalized reporter.
FIG. 2. Phylogenetic analysis of IAPV sequences in the following areas: RdRp (nucleotides [nt] 5448 to 5790) (A); ORF2 (nt 7817 to 8588) (B); 3′ terminus of ORF1-intergenic region-5′ terminus of ORF2 (nt 6140 to 6879) (C); 5′ UTR-5′ terminus of ORF1 (nt 121 to 1018) (D). Sequence data were aligned with those of other samples or DNA database sequences using CLUSTAL X, version 1.83. Phylogenetic analysis was performed using the Kimura model of nucleotide substitution. Programs from the MEGA package were used to produce phylogenetic trees, reconstructed by the neighbor-joining method. The statistical significance of a particular tree topology was evaluated by bootstrap resampling of the sequences 1,000 times. Strains are annotated by number, country of isolation, and, if known, year of isolation. Lineages (groups 1, 2, and 3) and virus names are given on the right.

Kimura model of nucleotide substitution. Programs from the MEGA package (18) were used to produce phylogenetic trees and reconstructed by the neighbor-joining method. The statistical significance of a particular tree topology was evaluated by bootstrapping of the sequences 1,000 times. Similar topologies were obtained with alternative substitution and analysis models.

Complete genome sequencing. Primers were designed to amplify 1-kb regions of the IAPV genome using Primer3Plus (http://www.bioinformatics.nl/primer3plus; 29) (the primers are available on request). A 500-bp staggered-overlap strategy was employed to improve the fidelity of the sequences. One microtiter of viral cDNA was added to 47 μL PCR master mixture consisting of 1× OptiBuffer, 2.5 mM MgCl₂, 0.5 mM of each deoxynucleoside triphosphate, and 5 units of Short Bio-X-Act polymerase (Bioline) containing 2 μM of each primer. Samples were incubated at 95°C for 10 min, followed by 45 cycles of incubation at 95°C for 1 min, 55°C for 1 min, and 72°C for 1.5 min, with a final extension at 72°C for 7 min.

Detection of recombination events. Potential recombination patterns were screened using the programs GARD, Simplot, RDP, Bootscan, and LARD (likelihood analysis of recombination) (11, 17, 19, 22, 24). GARD is designed to search for evidence of segment-specific phylogenies. The method searches for breakpoints in the alignment, inferring phylogenies for each putative nonrecombinant fragment. Putative recombinant sequences were also identified with Simpleplot using a sliding window of 200 bp and a step size of 20 bp. Bootscanning was carried out as implemented in Simplot using the putative recombinant sequence as the query. LARD, as implemented in the RDP program, runs the informative sites test for recombination detection.

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been deposited under GenBank accession numbers EU436423 to EU436537.

RESULTS

Diagnostic PCR for IAPV. Two PCR assays targeting the RdRp region and part of ORF2 were optimized to a sensitivity threshold of 160 genome copies per bee (10 copies per assay) using cloned IAPV standards in a background of non-CCD bee RNA (Fig. 1). Both diagnostic assays correctly identified all IAPV-positive samples with titers ranging from 1,120 to 320,000 copies/bee. Assay specificity was confirmed by the absence of signal in extracts from the IAPV-negative samples,
including five that contained KBV, the closest known relative of IAPV.

Phylogenetic analysis of RdRp sequence. The RdRp region is both highly conserved and well represented in sequence databases. It is useful for establishing broad taxonomic relationships but lacks the variability needed for detailed phylogenetic analysis. The phylogenetic tree obtained with a data set of related dicistrovirus RdRp sequences is shown in Fig. 2A. This analysis suggested that some virus isolates from Australia, France, and Russia previously reported as KBV may represent IAPV. Dicistrovirus nomenclature remains obscure pending formal classification by the International Committee on Taxonomy of Viruses. Indeed, IAPV may ultimately be designated a subtype of KBV. Nonetheless, these data indicate that current isolates from Israel, Canada, Australia, and the United States, as well as older isolates from Australia, Russia, and France, represent a distinct evolutionary lineage.

Molecular epidemiology of IAPV. Dicistrovirus sequence information is scarce in genome regions other than the RdRp. Indeed, only one complete genome sequence of IAPV has been reported. Genetic analysis using areas around the intergenic region (IGR) (739 bp), the 5’ UTR and helicase (897 bp), and the capsid sequence (771 bp) demonstrated that isolates cluster into three groups (Fig. 2B, C, and D). Group 1 comprises isolates from Australian packages and isolates from the western United States (Texas and California); group 2 comprises the initial isolate (EF219380) and currently circulating variants from Israel; group 3 comprises isolates circulating in the eastern United States (Maryland, Wisconsin, Pennsylvania, and Florida) and Canada.

Recombination analysis. Analysis of the original IAPV isolate (EF219380) from Israel revealed that, whereas the 5’ end of the genome clusters with current Israeli IAPV isolates, the 3’ end of the genome clusters with Australian and U.S. isolates. This observation could be explained by a recombination event. Recombination has been reported in picornaviruses and in other insect or vector-borne viruses (11, 12, 20). Since recombination events may occur at various positions in the genome, phylogenetic trees based on partial sequence data may be misleading. To address this concern, we searched for evidence of recombination by recovering and sequencing three complete genomes, one from each proposed lineage: IS1-ISRAEL-2007 (EU436455, from a bee collected from an ailing colony in Israel in 2007), OZ6-AUSTRALIA-2007 (EU436456, from a bee in a package of imported Australian honey bees), and DVE31-OP3-PA-USA-2006 (EU436423, from a bee obtained at the original site of detection of CCD in Pennsylvania). Plaque-purified viral template could not be employed for these studies because IAPV has not yet been isolated in culture. Thus, to eliminate the potential for creation of chimeras by amplification of sequences from two different coinfected viruses, the ~9,500-bp genomes were sequenced from single bees through amplification and assembly of 1-kb fragments overlapping each other by 500 bp.

Evidence for a recombination event was found between the original IAPV, the OZ6 isolate, and the IS1 isolate (Fig. 3). Searches for segment-specific phylogenies using the computer algorithm GARD identified the major breakpoint of recombination between the strains at position 6,146, the 3’ end of the first nonstructural ORF. Similar results were obtained with Simplot, Booscan, and LARD.

Characterization of the IAPV genome. The complete genome sequences (accession numbers EU436423, EU436455, and EU436456) obtained with the original IAPV and with other cripaviruses were analyzed with the objective of discovering differences that might be implicated in pathogenesis. OZ6 and DVE31 had one amino acid change (italic) in one potential 3C-pro cleavage site (274DIVKQGAR282); however, this motif is still recognized as a potential cleavage site (21). More profound differences came to light with analysis of the initiation sites of the ORFs. Both ORFs are preceded by IRES that initiate translation when the canonical cap-dependent cellular mechanism is disabled. The second IRES, also known as the IGR-IRES, has been described as a methionine-independent starting site for translation (25). The predicted initiation sites of ORF2 are similar in all IAPV and KBV sequences examined (Fig. 4). In contrast, isolate differences were observed with respect to ORF1. In sequence databases, the ORF1 start site of dicistroviruses has been assigned in various ways. ORF1 in triatoma virus (AF178440) opens with the triplet that immediately follows an in-frame stop codon, assuming an AUG-independent initiation. In both drosophila C virus (AF014388) and Taura syndrome virus (AF277675), the second in-frame AUG triplet has been assigned as the ORF1 start site. In other dicistroviruses, the first in-frame AUG triplet has been assigned as the initiation site. The original sequence reported by Maori et al. (21) proposed an AUG initiation site at position 696 of the genome. In the three complete genome sequences reported here, as well as in all IAPV-infected bees studied (more than 40 samples, including 5 collected in Israel), multiple variations were observed around the ORF1 initiation area. An AUG initiation site, in a location similar to that of the proposed KBV and ABPV ORF1 initiation sites, was detected at position 545, 151 bp upstream of the originally proposed AUG (Fig. 5). Termination sites for both ORFs and initiation sites for ORF2 were conserved among the three honey bee viruses.

DISCUSSION

IAPV was recently reported to be present in honey bees diagnosed with CCD in the United States, in bee samples
imported from Australia, and in royal jelly from China (8). In a subsequent study, IAPV was found in historical samples collected as early as 2002, preceding the documented importation of bees from Australia (6). Although the role of IAPV in the pathogenesis of CCD remains unknown, its presence in a hive has high positive predictive value for CCD (8), and the capacity of IAPV to cause paralytic disease has been demonstrated experimentally (21). Given the significance of IAPV as a marker for CCD and its candidacy as an important pathogen of pollinating species, we pursued phylogenetic analysis of geographically discrete IAPV isolates, as well as the establishment of assays to facilitate IAPV surveillance.

IAPV was first noted in 2004, when its complete genome sequence was shown to be distinct from those of KBV and ABPV (21). Analysis of RdRp sequences from IAPV isolates from the United States, Canada, Australia, and Israel revealed similarities to sequences reported as KBV from France, Australia, and Russia. Indeed, the use of consensus RdRp primers during our initial study showed cross-reactivity between IAPV and KBV that required sequencing for resolution (Fig. 2A). The highly conserved nature of the RdRp region highlighted the need to study other, more variable regions. Nonetheless, the RdRp sequence results were sufficient to demonstrate that IAPV is likely present in France, Russia, the United States, Australia, Israel, and China. It is curious that IAPV was not recognized in the United States prior to the study by Cox-Foster et al. in 2007, despite surveys using a primer pair (27) now known to amplify both IAPV and the targeted KBV.

FIG. 4. Alignment of class II IGRs of dicistroviruses (A) and diagram illustrating the structure of IAPV-DVE31 IGR-IRES (B). (A) The predicted base-pairing of the area encompassed by nucleotides 6413 and 6623 of IAPV EF219380 is shown. The areas involved in the base pairing are coded (a to h) and marked with arrows. Those areas forming pseudoknots are labeled PKI to -III. Conserved nucleotide areas and bulges are boxed. The accession numbers of the dicistrovirus IGRs are shown. (Figure adapted from reference 16). (B) IAPV-DVE31 IGR class II IRES structure predicted using the SINV-1 IGR IRES model (16). The sequence of IAPV IGR IRES fits well into the type II IGR IRES model. Predicted base-pairing and stem-loop structures are marked according to panel A. The predicted initiation codon is indicated (Gly).
While one subsequent survey (6) found a moderate frequency of IAPV in two populations surveyed since 2002, it is possible that the overall prevalence of IAPV was low until recently. Alternatively, since prior surveys rarely sequenced amplified viral products, IAPV-derived amplicons may have been mistakenly classified as KBV (5, 9, 10, 13, 14, 26). The Australian KBV isolate corresponds to a strain isolated in 1979 in South Australia (1, 3). This strain is antigenically different from other isolates of KBV and shows capsid proteins of distinct sizes with respect to other KBV isolates (1). Intriguingly, Australia has reported a syndrome similar to CCD, termed “disappearing disorder” (2). In 1999 and 2000, heavy losses of honey bees were observed in apiaries from Cunningham’s Gap and Gympie sites (2). The report concluded that no microbiological agents were associated with the disorder. However, it also noted that healthy pupae that were injected with larval extracts from the affected hives plus antisera against the Canberra strain of KBV became infected with another KBV strain. It is conceivable that the serologically different strain of KBV detected in the affected pupae was IAPV. 

IAPV sequence analysis in three genomic domains suggested the existence of three different groups: group 1 includes samples from operations in the western United States, as well as from bee packages imported from Australia; group 2 includes sequences from Israel; group 3 includes sequences from three operations in the eastern United States and one operation in Canada. We interpret these findings as consistent with multiple introductions of IAPV into the United States.

The IAPV strain originally reported by Maori et al. (21) did not cluster consistently with other isolates. Despite similarity to other Israeli strains in the 5′ end of the genome, differences were apparent in the 3′ end of the genome. This finding is compatible with either bona fide recombination in vivo or an artifact representing joined sequences of coinfecting viruses obtained during PCR cloning. The classical approach to eliminating the potential for artifacts is to plaque purify viruses prior to sequence analysis. However, because IAPV has not been isolated in pure culture, we adopted an alternative strategy in which 500-bp staggered overlapping products were sequenced to ensure fidelity. The original IAPV reference sequence was not obtained using either our strategy or plaque-purified virus; thus, we cannot determine whether it represents a bona fide recombination event.

Perhaps the most interesting observation concerning the sequencing of the different IAPV strains was the identification of deletions in the vicinity of the ORF1 start site identified in the original IAPV report (21). Assuming that translation from the 5′ IRES is AUG dependent and considering the changes in all IAPV sequences obtained in this study, we predict that ORF1 starts at position 545 in our isolates, as opposed to position 696, predicted to be operative in the IAPV reference strain. It is conceivable that the IAPV isolates from Israel and those identified in this study employ different ORF1 start sites. A question that cannot be resolved at this time is whether differences in translation profiles account for the apparent differences in pathogenesis between the phenotype originally described in association with IAPV infection (shivering wings, paralysis, and death in close proximity to the hive) and CCD.

Key to IAPV surveillance will be the establishment of sensitive, specific PCR assays for detection of viral nucleic acid. At the time we initiated our metagenomic study of CCD, only one sequence of IAPV was available. The use of published primers did not allow discrimination of KBV and IAPV without product sequence. The subsequent accumulation of additional IAPV and KBV sequence data has enabled the design of new assays that appear to be specific; however, deposited sequence data for IAPV and KBV remain scarce, and we anticipate that the assays we report here will require further modification as investigators focus on the distribution and biology of this intriguing agent.

REFERENCES


