Apidologie 31 (2000) 17–23

Original article

Nucleotide sequence variations in Kashmir bee virus isolated from *Apis mellifera* L. and *Varroa jacobsoni* Oud.

Akey C.F. HUNG[a]*, Christine Y.S. PENG[b], H. SHIMANUKI[a]

[a] Bee Research Laboratory, USDA-ARS, Beltsville, MD 20705, USA
[b] Department of Entomology, University of California, Davis, CA 95616, USA

(Received 14 June 1999; accepted 19 August 1999)

Abstract – The 374 bp amplicon of 20 Kashmir bee virus (KBV) isolates from 14 bee and 6 mite samples from California was sequenced. The sequence data and those of homologous KBV amplicons of one bee each from Australia, Canada, Maine, and Maryland were analyzed. Pairwise comparisons of nucleotide sequences show the mean distance between Californian isolates varying from 0% to 3.2%. The mean distance between Californian isolates and the two U.S. or the Canadian isolates are equal or less than 3.2%. The Australian isolate is different from all North American isolates with mean distances of 11.5 to 12.3%. This 374 bp sequence contains a single open reading frame, the 124 amino acid sequence matched closely with the sequences of many virus polyproteins.

Kasmir bee virus / direct RT-PCR / pairwise comparison / phylogenetic tree

1. INTRODUCTION

At least 16 viruses are known to infect honey bees [2]. Immunodiffusion has been routinely used for bee virus identification. This technique, however, appears to be effective only when the virus has multiplied sufficiently to kill the individual larvae, pupae and adult bees [2]. Stoltz et al. [13] used reverse transcription-polymerase chain reaction (RT-PCR) to detect Kashmir bee virus (KBV). This technique requires time-consuming virus purification and RNA extraction steps. Hung and Shimanuki [9] developed a direct RT-PCR method of KBV detection that circumvents these steps.

KBV is one of the picorna-like viruses found in honey bees. It was first isolated from a diseased adult bee of the Asian honey bee, *Apis cerana* [4]. KBV has also been

* Correspondence and reprints
E-mail: ahung@asrr.arsusda.gov

found in adult bees of the Western honey bee, *Apis mellifera*, in Canada, Spain, India, Australia and New Zealand [1], Fiji [3] and the U.S. [5,7]. Partial nucleotide sequence data have been used to characterize KBV genome [9,10]. Nucleotide sequence variations from KBV isolated from honey bees and the parasitic mite, *Varroa jacobsoni*, are presented here to further characterize this bee virus.

2. MATERIALS AND METHODS

2.1. Sources of samples

Fifty five adult bees and 82 mites from a mite-infested colony and 32 adult bees from a mite-free colony were used for KBV detection by RT-PCR. Bees from the mite-free colony were inoculated with mites from the mite-infested colony. Both bee colonies were located in apiaries at University of California, Davis about four miles apart.

Individual bees were homogenized in 0.45 ml of 0.01M potassium phosphate buffer (pH 6.7) with a tissue grinder, but only 2 μl/mite of the same buffer was used to homogenize two pooled mite samples (23 and 30 mites each). We later used 15 μl of buffer to homogenize individual mites. All homogenates were centrifuged at ca. 4 000·g for 15 min and stored at –20 °C without removing the supernatant.

2.2. RT-PCR analyses and DNA sequencing

The detection of KBV was carried out with an aliquot of 2 μl supernatant of the homogenate from each sample, using the direct reverse transcription-polymerase chain reaction method (direct RT-PCR) as described by Hung and Shimanuki [9]. The primer pair 5'-GATGAACGTCGACC-TATTGA-3' and 5'-TGTGGGTTGGC-TATGAGTCA-3' of Stoltz et al. [13] was used in PCR, each at a final concentration of 0.4 μM.

The band containing the RT-PCR reaction product was excised from a 1.5% agarose gel after electrophoresis, purified and sequenced on an ABI DNA sequencer (models 373a and 377, PE Applied Biosystems, Foster City, CA, USA). Both DNA strands were sequenced three times in each direction and the sequences of each strand were compared to each other to confirm the accuracy of the sequence data. Sequence data were edited using DNASIS™ for Windows® Version 2.5 (Hitachi Software, South San Francisco, CA, USA). Sequences were entered in a NEXUS file and analyzed with parsimony using PAUP 3.0 (Phylogenetic Analysis Using Parsimony, by D. Swofford, Illinois Natural History Survey, Champaign, Illinois). DNASIS™ as also used for protein translation in preparation for a BLASTP search of sequence similarity.

3. RESULTS

The PCR primers generated an amplicon of about 400 bp in 40 of the 55 adult bees from the mite-infested colony, 16 of the 29 mites, 10 of the 31 inoculated adult bees, and the two pooled mite samples. This amplicon is about the same size as the KBV amplicon reported by Stoltz et al. [13] and Hung and Shimanuki [8] using isolated KBV and extracted RNA. This KBV amplicon from 11 adult bees (CAbees 1–11) of the mite-infested colony, three inoculated bees (CAbees 12–14), the two pooled mite samples (CAmites 1–2) and four individual mites (CAmites 3–6) were sequenced.

Based on the actual nucleotide count (including the two primer sites), the size of this amplicon was 414 bp. The nucleotide sequences of a homologous KBV amplicon in three isolates from Australia, Canada and Maryland were reported by Hung and Shimanuki [8,10]. Sequence data of another homologous amplicon of a KBV isolate from Maine are also available from GenBank (Hung, unpublished data). These four sets of sequence data were included for

comparison. Thus a total of 24 sets of sequence data were used in the analysis using PAUP. To minimize sequencing errors caused by missing or mismatched bases at both ends, the analysis was carried out without the 40 bases describing the sense and antisense primer sites.

The results of analyses using this 374 bp fragment are shown in Figure 1 and Table I. As shown in Figure 1, there are three branches with eight Californian isolates (i.e. CAbees 7–10 and CAmites 3–6) forming a homogeneous group with no variation. The three isolates from Maryland, Maine, Canada and the remaining Californian isolates formed a heterogenous group. As shown in Figure 1, isolates from bees and mites of the mite-infested colony (CAbees 1–11 and CAmites 1–6) and isolates from bees of the mite-free colony (CAbees 12–14) do not each form a cluster.

Pairwise distances between isolates are shown in Table I. Mean distances between the Californian isolates vary from 0% (e.g. CAbee6 and CAbee12) to 3.2% (e.g. CAbee1 and CAbee). Mean distances between the Canadian isolate and the Californian isolates vary from 2.1% (e.g. CANbee and CAbee12) to 2.9% (e.g. CANbee and CAbee7). The mean distances between Maine or Maryland isolate and the California isolates, however, vary from 2.7% (e.g. MDbee and CAbee12) to 3.2% (e.g. MDbee and CAbee1). The Australian isolate is distinctly different from all North American isolates, with mean differences between 11.5% (e.g. AUSbee and CAbee12) to 12.3% (e.g. AUSbee and CANbee).

This 374 bp sequence contains a single open reading frame (ORF1) of 124 amino acids with codon starting at base 2 (deduced amino acid sequence not shown). Amino

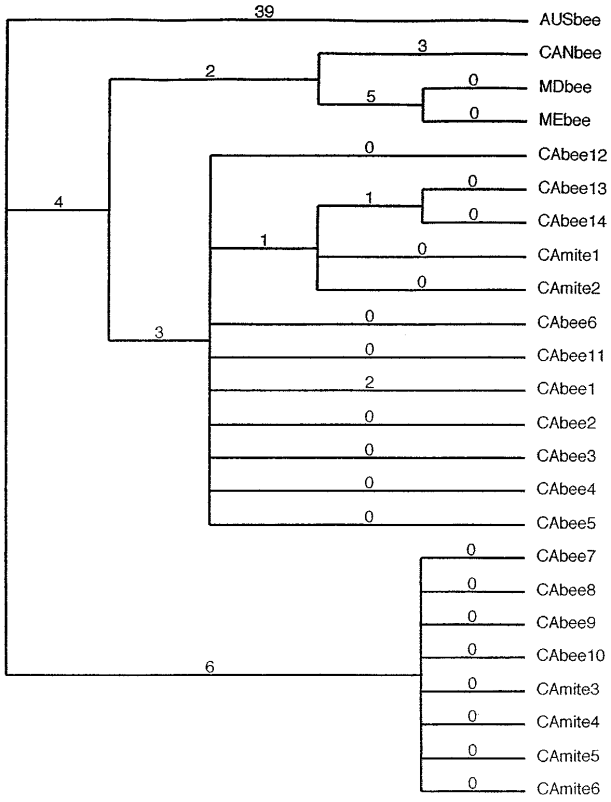

Figure 1. Phylogenetic tree showing the relationship between KBV isolates from California (CAbees 1-14 and CAmite 1-6), eastern U.S. (MDbee and MEbee), Canada (CANbee), and Australia (AUSbee). Number on each branch represents branch length.

Table I. Pairwise distances between KBV isolates. Below diagonal: absolute distances; above diagonal: mean distances.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	AUSbee		0.123	0.115	0.12	0.12	0.115	0.12	0.12	0.12	0.12	0.115	0.12	0.115	0.115	0.115	0.115	0.118	0.118	0.12	0.12	0.12	0.12	0.12	0.12
2	CANbee	46		0.021	0.021	0.021	0.021	0.029	0.029	0.029	0.029	0.021	0.021	0.021	0.021	0.021	0.021	0.024	0.024	0.029	0.029	0.029	0.029	0.021	0.021
3	CAbee12	43	8		0.005	0.005	0	0.032	0.032	0.032	0.032	0	0.005	0	0	0	0	0.003	0.003	0.032	0.032	0.032	0.032	0.027	0.027
4	CAbee13	45	8	2		0	0.005	0.032	0.032	0.032	0.032	0.005	0.011	0.005	0.005	0.005	0.005	0.003	0.003	0.032	0.032	0.032	0.032	0.032	0.032
5	CAbee14	45	8	2	0		0.005	0.032	0.032	0.032	0.032	0.005	0.011	0.005	0.005	0.005	0.005	0.003	0.003	0.032	0.032	0.032	0.032	0.032	0.032
6	CAbee6	43	8	0	2	2		0.032	0.032	0.032	0.032	0	0.005	0	0	0	0	0.003	0.003	0.032	0.032	0.032	0.032	0.027	0.027
7	CAbee7	45	11	12	12	12	12		0	0	0	0.032	0.032	0.032	0.032	0.032	0.032	0.029	0.029	0	0	0	0	0.029	0.029
8	CAbee8	45	11	12	12	12	12	0		0	0	0.032	0.032	0.032	0.032	0.032	0.032	0.029	0.029	0	0	0	0	0.029	0.029
9	CAbee9	45	11	12	12	12	12	0	0		0	0.032	0.032	0.032	0.032	0.032	0.032	0.029	0.029	0	0	0	0	0.029	0.029
10	CAbee10	45	11	12	12	12	12	0	0	0		0.032	0.032	0.032	0.032	0.032	0.032	0.029	0.029	0	0	0	0	0.029	0.029
11	CAbee11	43	8	0	2	2	0	12	12	12	12		0.005	0	0	0	0	0.003	0.003	0.003	0.032	0.032	0.032	0.027	0.027
12	CAbee1	45	8	2	4	4	2	12	12	12	12	2		0.005	0.005	0.005	0.005	0.008	0.008	0.032	0.032	0.032	0.032	0.032	0.032
13	CAbee2	43	8	0	2	2	0	12	12	12	12	0	2		0	0	0	0.003	0.003	0.032	0.032	0.032	0.032	0.027	0.027
14	CAbee3	43	8	0	2	2	0	12	12	12	12	0	2	0		0	0	0.003	0.003	0.032	0.032	0.032	0.032	0.027	0.027
15	CAbee4	43	8	0	2	2	0	12	12	12	12	0	2	0	0		0	0.003	0.003	0.032	0.032	0.032	0.032	0.027	0.027
16	CAbee5	43	8	0	2	2	0	12	12	12	12	0	2	0	0	0		0.003	0.003	0.032	0.032	0.032	0.032	0.027	0.027
17	CAmite1	44	9	1	1	1	1	11	11	11	11	1	3	1	1	1	1		0	0.029	0.029	0.029	0.029	0.029	0.029
18	CAmite2	44	9	1	1	1	1	11	11	11	11	1	3	1	1	1	1	0		0.029	0.029	0.029	0.029	0.029	0.029
19	CAmite3	45	11	12	12	12	12	0	0	0	0	12	12	12	12	12	12	11	11		0	0	0	0.029	0.029
20	CAmite4	45	11	12	12	12	12	0	0	0	0	12	12	12	12	12	12	11	11	0		0	0	0.029	0.029
21	CAmite5	45	11	12	12	12	12	0	0	0	0	12	12	12	12	12	12	11	11	0	0	0	0	0.029	0.029
22	CAmite6	45	11	12	12	12	12	0	0	0	0	12	12	12	12	12	12	11	11	0	0	0		0.029	0.029
23	MDbee	45	8	10	12	12	10	11	11	11	11	10	12	10	10	10	10	11	11	11	11	11	11		0
24	MEbee	45	8	10	12	12	10	11	11	11	11	10	12	10	10	10	10	11	11	11	11	11	11	0	

acid sequence analysis (data not shown) shows a similar branching pattern to that inferred from the nucleotide sequence analysis. The majority of Californian isolates, however, have identical amino acid sequences, since the few miss-matched bases did not result in amino acid coding differences. A BLASTP search showed that this ORF1 matched closely with the sequences of many virus polyproteins, e.g. (CAbee1): replicase polyprotein of *Drosophila* C virus (P = 4.8e–026); polyprotein B of red clover mottle virus (P = 2.0e–019); nonstructural polyprotein of *Rhopalosiphum padi* virus (P = 2.7e–019); polyprotein B of cowpea aphid-borne mosaic virus (P = 2.6e–017).

4. DISCUSSION

The avian myeloblastosis virus (AMV) used in the reverse transcription is known to have an error rate of 5.9×10^{-5}/base [12]. For this 374 bp amplicon an error rate of 0.022 would be expected. The mean distance of 3.2% between some Californian isolates could not be the result of transcription error by AMV. Furthermore, there are eight isolates (Fig. 1 lower branch) with identical sequences suggesting the high fidelity of the reverse transcription.

Most RNA virus infections consist of heterogeneous virus populations often referred to as quasi-species [6]. A 475 nt region of the lentivirus *pol* gene in eight isolates from naturally infected sheep from southern France was found to have nucleotide variation of 2.3–8.1% [11]. The KBV sequences of 23 North American isolates reported here do not seem to be as heterogeneous as in the lentivirus group.

Allen and Ball [1] divided KBV into two serological groups: the Canadian and Spanish strains formed one group and the Indian, South Australian and New Zealand strains formed another group. Our sequence analyses also show that the Australian isolate is distinctly different from North American isolates. According to Bruce et al. [5] there are three "serological" strains of KBV in each of the seven states they surveyed. Figure 1 in their paper, however, demonstrates only the distinction between one Californian isolate and KBV from Australia and New Zealand, it does not show that "serologically distinct strains of the virus exist in the U.S.". Our pairwise distance analysis also shows that the difference between the two eastern U.S. isolates and the Californian isolates is within the range of differences found among isolates of KBV from the same bee colony. The results of sequence analyses presented here with multiple isolates from California and single isolates from the eastern U.S., Canada, and Australia will serve as a baseline for the expected genetic variation in KBV. With this baseline information, future analyses of multiple isolates from different regions can determine if several distinct strains of KBV are present in the U.S. as Bruce et al. [5] claimed.

ACKNOWLEDGEMENTS

We thank Jay Evans for reviewing the manuscript and technical advice, Michele Logan and Sun Trihn for their technical assistance.

Résumé – Variations dans la séquence nucléotidique chez le virus Cachemire de l'abeille isolé chez *Apis mellifera* L. et *Varroa jacobsoni*. La détection du virus du Cachemire de l'abeille (KBV) a été faite avec un aliquot de 2 µl du surnageant du broyat individuel d'abeilles mellifères et d'acariens *V. jacobsoni* en utilisant le protocole de RT-PCR directe, tel que Hung et Shimanuki l'ont décrit [9]. La première paire d'amorces de Stoltz et al. [13] 5'-GATGAACGTCGACCTATTGA-3' et 5'-TGTGGGTTGGCTATGAGTCA-3' a été utilisée dans la PCR. La bande contenant le produit de RT-PCR de 374 pb a été excisé d'un gel de agarose à 1,5 % après

électrophorèse, puis purifié. On a séquencé le produit d'amplification de 20 isolats de KBV provenant de 14 échantillons d'abeilles et de 6 échantillons d'acariens venant de Californie et analysé ces données de séquence et celles de produits d'amplification analogues provenant d'une abeille de chacune des régions suivantes : Australie, Canada, états du Maine et du Maryland. Les comparaisons deux à deux des séquences nucléotidiques montrent que la distance moyenne entre les isolats de Californie varie de 0 à 3,2 %. La distance moyenne entre les isolats de Californie et ceux des USA ou du Canada est égale ou inférieure à 3,2 %. L'isolat australien diffère de tous les isolats américains par des distances moyennes comprises entre 11,5 et 12,3 %. Cette séquence de 374 pb contient un seul cadre de lecture ouvert, la séquence de 124 acides aminés est fortement homologue à des séquences de nombreuses polyprotéines de virus.

Virus Cachemire de l'abeille / RT-PCR directe / comparaison deux à deux / arbre phylogénétique

Zusamenfassung – Variation der Nukleotidsequenz des Kashmir Bienenvirus aus *Apis mellifera* L. und *Varroa jacobsoni* Oud. Die Bestimmung des Kashmir Bienenvirus (KBV) wurde an einem Aliqot von 2 µl Überstand des Homogenats einzelner Honigbienen und *Varroa jacobsoni* Milben unter Verwendung des von Hung und Shimanuki [9] beschriebenen Protokolls der direkten reversen Transkriptions-Polymerasen Kettenreaktion durchgeführt (direkte RT-PCR). In der PCR wurde das Primerpaar 5' – GATGAACGTCGACCTATTGA – 3' und 5' – TGTGGGTTGGCTATGAGTCA – 3' von Stoltz et al. [13] verwendet. Die Bande, die das 374 bp lange Reaktionsprodukt der RT-PCR enthielt, wurde aus einem 1.5% Agarosegel einer Elektrophorese ausgeschnitten und gereinigt. Das Amplifizierungsprodukt von 20 Isolaten des Kashmir Bienenvirus von 14 kalifornischen Bienenproben und 6 Milbenproben wurde sequenziert. Die Sequenzdaten und die Daten homologer Amplifizierungsprodukte aus jeweils einer Biene aus Australien, Kanada, Maine und Maryland wurden analysiert. In paarweisen Vergleichen variierte der mittlere Abstand zwischen den kalifornischen Isolaten zwischen 0 und 3,2%. Der mittlere Abstand der kalifornischen Isolate zu den zwei Isolaten aus den USA und Canada war gleich oder geringer als 3,2%. Das australische Isolat unterschied sich mit Abständen zwischen 11,5 und 12,3% von allen nordamerikanischen Isolaten. Die verwendete 374 Basenpaare lange Sequenz enthielt eine einzige codierende DNS - Sequenz, die Reihenfolge von 124 Aminosäuren stimmte fast mit der vieler von Viren bekannter Polyproteine überein.

Kashmir Bienenvirus / direkte RT-PCR / Paarweiser Vergleich / Phylogenetischer Baum

REFERENCES

[1] Allen M.F., Ball B.V., Characterisation and serological relationships of strains of Kashmir bee virus, Ann. Appl. Biol. 126 (1995) 471–484.

[2] Allen M.F., Ball B.V., The incidence and world distribution of honey bee viruses, Bee World 77 (1996) 141–162.

[3] Anderson D.L., Pests and pathogens of the honeybee (*Apis mellifera*) in Fiji, J. Apic. Res. 29 (1990) 53–59.

[4] Bailey L., Woods R.D., Two more small RNA viruses from honey bees and further observations on sacbrood and acute bee-paralysis virus, J. Gen. Virol. 37 (1977) 175–182.

[5] Bruce W.A., Anderson D.L., Calderone N.W., Shimanuki H., A survey for Kashmir bee virus in honeybee colonies in the United States, Am. Bee J. 135 (1995) 352–354.

[6] Domingo E., Holland J., Biebricher C., Eigen M., Quasi-species: the concept and the word, in: Gibbs A.J., Calisher C., Garcia-Arenal F. (Eds.), Molecular basis of virus evolution, Cambridge University Press, Cambridge, 1995, pp. 181–191.

[7] Hung A.C.F., Adams J.R., Shimanuki H., Bee parasitic mite syndrome (II): The role of *Varroa* and viruses, Am. Bee J. 135 (1995) 702–704.

[8] Hung A.C.F., Shimanuki H., A scientific note on the partial nucleotide sequence of a U.S. strain of Kashmir bee virus isolated from *Apis mellifera* L., Apidologie 30 (1999) 355–356.

[9] Hung A.C.F., Shimanuki H., A scientific note on the detection of Kashmir bee virus in individual honeybees and *Varroa* mites, Apidologie 30 (1999) 353–354.

[10] Hung A.C.F., Shimanuki H., Nucleotide sequence and restriction site analyses in three isolates of Kashmir bee virus from *Apis mellifera* L. (Hymenoptera: Apidae), Proc. Wash. State Entomol. Soc. (in press).

[11] Leroux C., Vuillermoz S., Mornex J.-F., Greenland T., Genomic heterogeneity in the *pol* region of ovine lentiruses obtained from bronchoalveolar cells of infected sheep from France, J. Gen. Virol. 76 (1995) 1533–1537.

[12] Roberts J.D., Bebenek K., Kunkel T.A., The accuracy of reverse transcriptase from HIV-1, Science 242 (1988) 1171–1173.

[13] Stoltz D., Shen X.R., Boggis C., Sisson G**.,** Molecular diagnosis of Kashmir bee virus infection, J. Apic. Res. 34 (1995) 153–160.

APPENDIX

The nucleotide sequence data used here have been registered with GenBank under the following accession numbers (in parenthesis):

CAbee1 (AF083239); CAbee2 (AF135852); CAbee3 (AF117953); CAbee4 (AF135853); CAbee5 (AF135854); CAbee6 (AF135856); CAbee7 (AF135857); CAbee8 (AF135858); CAbee9 (AF135859); CAbee10 (AF135860); CAbee11 (AF135861); CAbee12 (AF085478); CAbee13 (AF052566); CAbee14 (AF085479); CAmite1 (AF052567); CAmite2 (AF135855); CAmite3 (AF093457); CAmite4 (AF093458); CAmite5 (AF093459); CAmite6 (AF093460); AUSbee (AF034541); CANbee (AF034542); MDbee (AF027125); MEbee (AF035359).