Reciprocal sequence exchange between non-retro viruses and hosts leading to the appearance of new host phenotypes

Eyal Maori, Edna Tanne, Ilan Sela*

Virus Laboratory, The Hebrew University of Jerusalem, Faculty of Agricultural, Food and Environmental Quality Sciences, Rehovot 76100, Israel

Received 29 September 2006; returned to author for revision 15 November 2006; accepted 22 November 2006

Abstract

Divergence among individuals of the same species may be linked to positional retrotransposition into different loci in different individuals. Here we add to recent reports indicating that individual variance occurs due to the integration of non-retroviral (potyviral) RNAs into the host genome via RNA recombination followed by retrotransposition. We report that in bees (Apis mellifera), approximately 30% of all tested populations carry a segment of a dicistrovirus in their genome and have thus become virus-resistant. Reciprocally, segments of host sequences have been found within defective-interfering-like sequences of a dicistrovirus. Similarly, host sequences were found fused to potyviral sequences, previously described integrated into their host genome. A potential, continuous RNA exchange leading to divergence is discussed.

© 2006 Elsevier Inc. All rights reserved.

Keywords: RNA recombination; Virus-host exchange; RNA heredity; evolution

Introduction

Bee mortality has recently become a more salient concern in Israeli apiculture, due to the discovery of a virus (tentatively termed Israeli acute paralysis virus: IAPV) which has been isolated, partially sequenced and characterized as a dicistrovirus (Mayo, 2002; Liljas et al., 2002). Bees are known to carry latent viruses, and epidemics can be elicited by environmental and artificial changes, causing disease outbreaks (Anderson and Gibbs, 1988). Thus, the parasitic bee mite Varroa destructor can engender disease by transmitting a virus or by simply puncturing a bee: its role in disease outbreaks has yet to be determined (Camazine and Liu, 1998; Bowen-Walker et al., 1999; Chen et al., 2004). One surprising phenomenon associated with IAPV syndrome is that only some hives in a cluster are affected, even though they are all exposed to the same environmental conditions and forage in the same fields. Since all individuals in a hive are progeny of the same queen, their genetic makeup (except for about 50% of the haploid males) is very similar (at least heterozygously), but different from that of the other hives. Therefore, genetic diversity could explain the susceptibility/resistance of one hive population relative to the next.

The Dicistroviridae are a family of insect viruses belonging to the picorna-like super family (van Regenmortel et al., 1999; Mayo, 2002). They carry positive-strand RNA of ca. 9 kb with two open reading frames (ORFs), each coding for a polyprotein that is processed to several active proteins during infection. The 5′ proximal ORF expresses non-structural proteins related to viral replication, and the ORF at the 3′ end expresses the various structural proteins (Govan et al., 2000; de Miranda et al., 2004). The two ORFs are separated by an intergenic region. Translation is not initiated by met-tRNA but rather from internal ribosomal entry sites (IRES), one located in the 5′ untranslated region and the other in the intergenic region. The structural polyprotein is processed by proteolysis to four or five capsid proteins, but (due to intermediate structures?) the number of viral proteins in infected tissue is larger (Govan et al., 2000; Wilson et al., 2000).

Recently, two independent papers have described integration into the host genome of DNA versions of non-retro RNA virus genome segments: flavivirus genome segments were found in its mosquito host (Crochu et al., 2004), and a segment of a potyviral genome was found in some (but not all) grapevine...
varieties (Tanne and Sela, 2005). In the latter case, a history of retrotransposition was evident, and RNA recombination was suggested as the step initiating integration.

RNA recombination between viral RNA segments and between a viral RNA and its respective transgene has been well documented (for example, Nagy and Simon, 1997; Adams et al., 2003, Oberste et al., 2004). However, little information has been published regarding viral–host RNA recombination (Monroe and Schlesinger, 1983; Greene and Allison, 1994; Charini et al., 1996). Recently, genome segments of non-retro RNA viruses have been found integrated in their respective host genomes (Crochu et al., 2004; Tanne and Sela, 2005), and in one case, an inserted segment of Potato virus Y (PVY) was shown to reside in a grapevine genome within a retrotransposomal element and a history of RNA recombination was implicated (Tanne and Sela, 2005). It has been suggested that retrotransposition is involved in the generation of pseudogenes (Esnault et al., 2000), but no evidence of transposition leading to speciation was found (Kidwell and Lisch, 2001; Mourier, 2005). Recently, it has been shown that long interspersed nuclear element-1 (LINE-1) causes individual diversity by retrotransposition into different gene loci in different neural cells, altering different gene functions (Muorti et al., 2005). Here we report a case showing not only the integration of a viral segment into the genome of its host (honeybee) but also the reciprocal integration of a host–genome segment into a virus-derived sequence. As a result of integration of a viral sequence into the host genome, a new phenotype has emerged: individual bees harboring the integrated viral sequence were found to be virus-resistant. In addition, we show that exchange of sequences between host and virus also occurs in the case of PVY and tobacco.

Results

Occurrence of an IAPV sequence in the bee genome and its expression

We PCR-tested the DNA of 100 individual bees from 19 different, randomly selected hives (located 20 to 150 km apart) and found that 31 of them carry a segment of IAPV, residing between bases 1380 and 1808 (GenBank Accession No. AY738529; see Materials and methods) of the structural protein cistron (Fig. 1A). Sequences upstream of this segment were not detected in the bee genome. This was corroborated by Southern blot analysis (Fig. 1B), where four out of eight individuals (lanes 2, 3, 5, 6) reacted positively.
discern transcription of genome-integrated viral sequences from virus infection: (a) A sequence upstream of the integrated one (bases 1263 to 1400 in the structural protein clone) has never been found in the bee genome, and hence the presence of RNA carrying this sequence is indicative of viral (or latent viral) infection. On the other hand, the occurrence of RNA corresponding to bases 1380 to 1808 and the absence of nucleotides 1263 to 1400 (Figs. 2 and 4) implies transcription from the integrated viral segment in the absence of infection. (b) Since RNA viruses replicate via a double-stranded intermediate, RNA in both sense (viral) and antisense orientations is expected in cells with a replicating virus, whereas RNA transcribed from an integrated sequence should be in either sense or antisense orientation. This was determined by RT–PCR with primers amplifying the two aforementioned viral sequences, as well as with primers for the 1380 to 1808 segments, designed to amplify either sense or antisense RNA.

Genome-walking analyses identified five loci of integration of the viral segment in the bee genome: clone EB29 integrated into the bee chromosome LG3, clone EA19 into chromosome LG5, clone EC5 into chromosome LG6, clone EC42 into chromosome LG12 and clone EC21 into chromosome LG16.

Fig. 3. Expression of an IAPV-derived protein in bees; a western blot of bee proteins. The blot was reacted with antibodies to IAPV. Lane 1: Proteins extracted from purified IAPV virions. Lanes 2 and 3: Proteins extracted from bees carrying an IAPV segment in their genome. Lane 4: Protein extracted from bee not harboring any IAPV segment in their genome. Size markers are indicated on the left.

Fig. 4. Discerning viral infection from expression of integrated viral sequences. RT–PCR was performed with primers amplifying a region outside the one found integrated into the bee genome. The various lanes represent RNA samples extracted from various individual bees. The appearance of a 150-bp band is indicative of infection. Arrows indicate sizes in bp. Lane M: size markers. Lane 6: negative control; PCR without a template. Lane 7: negative control; PCR without RT. Lane 8: positive control; the template was viral RNA. Lanes 1–5: RNA was extracted from various bees. Size marker ladder is in 100 bp increments.

EC5 (as an example) is diagrammatically illustrated in Fig. 5A and the entire sequence is shown in Fig. 6. Chromosomal locations and positions within the chromosomes of all clones are summarized in Table 1.

The clones’ sequences were deposited with GenBank (Accession nos. DQ468104–DQ468108). All viral insertions were found in intergenic regions of the various chromosomes.

Bees carrying a viral sequence in their genome are resistant to IAPV infection

Sixty-four randomly selected, healthy-looking bees were tested by DNA–PCR for the presence of integrated viral sequence, and by RT–PCR of the upstream (non-integrated) region for latent viral infection. Nineteen bees (29%) carried the viral segment in their genome, as evidenced by their positive reaction to the DNA–PCR assay. None of them reacted to the RT–PCR of the upstream sequence, and they were thus considered virus-free. The remaining 45 individuals did not carry the viral segment in their genome, but RT–PCR indicated the presence of latent virus in 22 of them. The absence of latent infection in viral–segment-carrying bees raised the question of whether bees harboring a viral segment are resistant/tolerant to viral infection. This was tested directly by inoculating bees with IAPV. The survival rate of bees (not carrying the viral segment) following injection of IAPV (1 μl containing 1 μg of virus per larva or pupa) was tested on over 300 individuals (in groups of 20–100 each) and was found to average less than 2%. We then injected 80 bee pupae with a purified virus preparation. The experiment and its controls are summarized in Table 2. Of the 80 virus-injected bees, 67 individuals died, but 13 emerged as mature, healthy-looking adults, indicating not only survival, but normal development as well. RT–PCR assays with primers amplifying the aforementioned viral sequence (bases 1380 to 1808 of the structural protein clone) in either sense or antisense
orientation revealed that (as expected) all dead bees carried RNA in both orientations, whereas all surviving individuals expressed the integrated viral sequence in only the sense orientation, indicating expression only from the integrated segment and that viral replication had not taken place. An example of these results is shown in Fig. 7.

A segment of a bee gene is embedded within an IAPV-associated defective RNA

Viral RNA recombination has hitherto been attributed to template-switching of the viral replicase. Recent reports, however, indicate a non-replicative, non-sequence-specific form of viral RNA recombination attributed to a breakage and joining mechanism (Agol, 1997; Nagy and Simon, 1997; Gallei et al., 2004). The latter is not necessarily dependent on recognition by a replicase, and may theoretically engender recombination between any two RNA species, including viral and host RNAs; a few such cases have been reported (Monroe and Schlesinger, 1983; Charini et al., 1996). While template-switching recombines one RNA species with another, breakage and joining is reminiscent of crossing over in DNA, where a segment of one DNA species may be flanked on both sides by segments of the other DNA species. Since we suggested that in the case of IAPV, the initial step was RNA recombination between a viral sequence and a bee sequence, we postulated that the recombination sites were adjacent to that viral sequence which was found integrated in the bee genome. Therefore, we performed RT–PCR on RNA extracted from CsCl-purified IAPV virions, using primers adjacent to the bee-integrated sequence. This resulted in the appearance of the expected major viral band with a few minor bands alongside it. As demonstrated diagrammatically in Fig. 5B, one of these minor bands (Fig. 8), which is reminiscent of a “defective-interfering” viral RNA (DI-RNA), was found interrupted by 41 nucleotides of a non-viral sequence homologous to rac1-activated kinase of several organisms (96% identity with the respective human segment; Fig. 9).

Table 1
Characteristics of the loci in which an IAPV viral sequence was integrated into the bee genome

<table>
<thead>
<tr>
<th>Clone</th>
<th>The bee chromosome into which a viral sequence was integrated</th>
<th>Position on the chromosome</th>
<th>Accession no. of the bee sequence and % identity with that of the clone</th>
</tr>
</thead>
<tbody>
<tr>
<td>EB29</td>
<td>LG3</td>
<td>244882 to 245322</td>
<td>NW_622759; 97%</td>
</tr>
<tr>
<td>EA19</td>
<td>LG5</td>
<td>182579 to 182748</td>
<td>NW_622804; 99%</td>
</tr>
<tr>
<td>EC5</td>
<td>LG6</td>
<td>519505 to 520633</td>
<td>NW_622830; 99%</td>
</tr>
<tr>
<td>EC42</td>
<td>LG12</td>
<td>391790 to 392141</td>
<td>NW_622565; 98%</td>
</tr>
<tr>
<td>EC21</td>
<td>LG16</td>
<td>394542 to 394849</td>
<td>NW_622631; 99%</td>
</tr>
</tbody>
</table>

Table 2
Summary of bee analysis following the injection of purified virus to pupae

<table>
<thead>
<tr>
<th>Carrying a sense viral sequence</th>
<th>Carrying an antisense viral sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Died</td>
<td>2</td>
</tr>
<tr>
<td>Died</td>
<td>7</td>
</tr>
<tr>
<td>Died</td>
<td>67</td>
</tr>
<tr>
<td>Died</td>
<td>12</td>
</tr>
<tr>
<td>Died</td>
<td>1</td>
</tr>
<tr>
<td>Died</td>
<td>62</td>
</tr>
<tr>
<td>Died</td>
<td>71</td>
</tr>
<tr>
<td>Died</td>
<td>6</td>
</tr>
<tr>
<td>Died</td>
<td>67</td>
</tr>
<tr>
<td>Died</td>
<td>9</td>
</tr>
<tr>
<td>Died</td>
<td>13</td>
</tr>
<tr>
<td>Died</td>
<td>1</td>
</tr>
</tbody>
</table>
RNA of a segment of the host DNA into the virus. This indicates that a transcript of a retroelement extended into a neighboring gene was spliced, joining together retro and host gene sequences (Derr, 1998; Mourier, 2005). Currently, we are determining the full structure of this chimeric sequence which has never been found integrated into the bee genome.

A segment of a Nicotiana gene is embedded within a PVY-associated defective RNA

Since PVY sequences were also found integrated in a plant genome (Tanne and Sela, 2005), a similar search with postulated PVY-DI-like sequences was also carried out. A PVY-DI carrying a segment of a stress-induced gene of a tobacco species (Nicotiana benthamiana; 99 bp, 86% homology) was found embedded in the PVY-DI sequence (Figs. 5C and 10). This stress-inducible gene segment carries several direct repeats (homology among the various repeats was 83–95%). Thus, the reciprocal exchange of genomic segments between virus and host at the RNA level was experimentally demonstrated.

Discussion

Evidence for reciprocal exchange of genetic material between virus and host

Previous papers have reported that DNA versions of RNA viral segments can be found in the respective host genome (Crochu et al., 2004; Tanne and Sela, 2005). This paper reports not only the integration of a DNA version of an RNA segment from a bee-affected dicistrovirus into the genome of the honeybee, but also the reciprocal integration of an RNA version of a segment of the host DNA into the virus. This indicates RNA–DNA mobility in both directions in a non-retroviral system. Based on structural features of the integrated viral sequences, we postulated an initial step of recombination between viral RNA and host mRNA, or a host RNA sequence carried on a retrotransposable element (Tanne and Sela, 2005). This is further corroborated in the present paper by the direct and inverted repeats in the RNA sequences as well as by the fact that the chimeric structure represented in Fig. 5B carries a viral sequence which has never been found integrated into the bee genome, implying that the recombination was at the RNA level. The five clones of the bee genome found to harbor IAPV sequences are positioned within intergenic regions of various bee chromosomes. This suggests that their origin may have been a transposable element, now deteriorated and non-autonomous. However, IAPV-derived RNA and protein are expressed, indicating that some IAPV-harboring loci remained as autonomous, expressible, transposable elements, or were integrated, in-frame, into the coding sequences of a host gene, or downstream of a host promoter.

Postulated mechanisms for RNA-based exchange of genomic information

Once a viral sequence has been mobilized into the host genome, it may serve, due to its inherent repeats, as a hot spot for further DNA transpositions within the DNA genome, resulting in gradual deterioration of the inserted sequence (for example, Vitte and Panaud, 2005). The grapevine locus into which a viral segment has been integrated (Tanne and Sela, 2005) and the aforementioned DI-like RNA sequences carry footprints of both transposition and retrotransposition. Recombined RNA, suggestive of a mechanism resembling crossing over (such as that of the DI-RNAs in Figs. 8 and 9), indicates the potential for recombination of any two RNAs (probably at a very low rate, depending on their proximity and abundance). Retrotransposition of heterologous sequences has been reported; however, it was assumed that the mechanism of fusion of two sequences had been DNA recombination, or that a transcript of a retroelement extended into a neighboring gene was spliced, joining together retro and host gene sequences (Derr, 1998; Mourier, 2005). Currently, we are developing a model system which may confirm the mobilization of a non-retro RNA into the genome, the frequency of these events, and the frequency of integration events leading to expression.

![Fig. 7. Example of the PCR analyses carried out as summarized in Table 2. RT–PCR of RNA extracted from bees. M: Size markers. S: Reverse transcription was performed with a primer adhering to sense-oriented RNA. −RT: A control reaction to ensure that no DNA is reacting; the reverse transcription step was performed in the absence of reverse transcriptase. In the second, third and fourth lanes (from left) RNA was extracted from a dead bee. In the fifth, sixth and seventh lanes (from left) RNA was extracted from a surviving bee. +: Positive control with viral RNA; reverse transcription was performed with a sense-adhering primer. Size marker ladder is in 100 bp increments.](image)

![Fig. 8. Electrophoretic pattern of the products of RT–PCR performed on RNA extracted from purified IAPV virions. PCR was conducted with primers corresponding to IAPV sequences immediately adjacent to the viral-integrated sequence. The sequence presented in Fig. 9 is of the minor band indicated by the arrow. Size marker ladder is in 100 bp increments.](image)
Possible impact of RNA recombination and reciprocal mobilization on genome dynamics and evolution

The potential impact of viral RNA recombination on viral as well as host evolution, disease severity and avoidance of the host immune system has been discussed at length (Bujarski and Kaesberg, 1986; Strauss and Strauss, 1988; Jarvis and Kirkegaard, 1992; Lai, 1992; Gallei et al., 2004). A case showing transposition-mediated divergence among individuals has also been demonstrated (Muorti et al., 2005). The present report experimentally supports a hypothesis of perpetual emergence of variance within a population leading to the appearance of new phenotypes within the same species or sub-species (strains, races, varieties, etc.). Thus, another factor may be a presumed cause of variation in both viruses and hosts. Moreover, we suggest that events leading to genome mosaicism and perpetual genome dynamics also take place at the RNA level. Stress, including viral infections, increases the rate of retrotransposition (Wessler, 1996) and the tobacco Tnt1 retrotransposon shares transcriptional activation regions with stress-inducible genes (Mhiri et al., 1997). The abundance of viral RNA and transcripts of some host genes, and the template-switching capacity of both reverse transcriptases and the viral replicases may enable RNA recombination and insertion into the host as well as the viral genomes.

Bi-directional RNA mobilization may lead to dual co-evolution of host and virus. Expression of viral sequences from the host genome is probably an infrequent event prone to deterioration. However, when expressible, it may lead to the appearance of a new host phenotype. The engendering of resistance might be a result of an RNAi mechanism now directed towards the viral sequence, or due to the removal of the pertinent host gene segment (Esnault et al., 2000). However, the reciprocal mobilization of a host sequence into the virus leads to a more stable, possibly replicative, virus-derived form, affecting virus divergence and virus–host interactions. This is due to the postulated ability of the native virus to support replication of the chimeric virus–host entities. Furthermore, if

the viral part of the chimera carries signals for assembly, the newly formed RNA will be encapsidated with the viral coat protein(s), as demonstrated in the present report. This may be a basis for viral divergence and may explain, for example, the origin of a host-homologous heat-shock protein in all closteroviruses (Karasev, 2000). The resultant virus–host chimeras, stabilized, and possibly replicative as “parasites” of the native virus, or even encapsidated within viral particles, can now be transmitted horizontally to other host individuals, to other host species, or even acquire new host range. Viruses are transmission-competent entities, and may be better suited than transposons for horizontal transmission of genetic material.

Transgenic resistance and environmental risks

The aforementioned integration of a viral segment into the bee genome results in “naturally transgenic” bees with a pronounced new phenotype: resistance to subsequent viral infection. Man has now successfully created transgenic organisms: in plants, transformation with viral sequences has brought about (in many cases) viral resistance (Sela, 1996). In some cases, this resistance is RNA-mediated and does not require expression of a protein from the transgene (Silva-Rosales et al., 1994; Smith et al., 1994). Apparently, man has unknowingly been mimicking a natural process.

Man has been concerned with the possible hazards of virus–host RNA recombination. The authors tend to agree with Falk and Bruening (1994) that man-engendered RNA recombination is statistically insignificant relative to the seemingly frequent natural process.

Materials and methods

Virus isolation

The homogenate of a single dead bee was injected into healthy bees and the virus was isolated from them. Bee-ho-
genates (in 0.01 M phosphate buffer, pH 7.6) contained a considerable amount of bacteria which were lysed with 0.3% sodium deoxycholate and 3Brij 58 (Sigma) followed by centrifugation at 10,000 × g for 10 min. The pellet was resuspended in the aforementioned buffer and subjected to two cycles of centrifugation (10,000 × g alternating with 100,000 × g). The final pellet was resuspended in a solution of CsCl (1.33 g/ml) and centrifuged at 100,000 × g for 24 h. The resultant single band was collected, diluted with buffer and re-centrifuged at 100,000 × g for 1 h.

**Molecular procedures**

Nucleic-acid extraction, PCR, Southern, Northern and Western blot analyses, DNA sequencing and other molecular procedures were performed according to established protocol (Sambrook and Russell, 2001). Genome-walking was carried out according to the protocol of the kit provider (Clontech).

To pinpoint the ends of the viral sequence integrated into the bee genome, we first performed PCR with randomly selected primers to various parts of the viral non-structural and structural protein ORFs. Only one set of primers (of the structural proteins’ ORF) reacted positively. We then conducted PCR assays, changing one primer at a time (in increments of 10–20 bp) upstream and downstream of the reacting region, until the PCR no longer reacted. The position of the insert’s end was further specified when the junction of the bee genome/IAPV sequences was determined by genome walking. The primers used for the various (reacting) PCR assays are specified in **Table 3**.

**Genome walking**

Genome walking from the IAPV sequence into the flanking bee sequences was carried out using Clontech’s Universal Genome Walker Kit according to the manufacturer’s protocol. The DNA for the genome-walking experiments was RNase-treated prior to cleavage with restriction enzymes. All genome-walking primers were with $T_m$ values of about 70 °C. Primers used for genome walking are listed in **Table 3**. Primers #935 and AP2 were used to amplify IAPV-harboring genomic fragments of the bee. Primers #935 and AP2 were used for the nested PCR stage of this amplification. The resultant genome-walking PCR products were cloned into the vector pDrive (Qiagen) and sequenced therefrom.

**Acknowledgments**

This research was supported in part by Grant No. US-3205-01R from BARD, The United States–Israel Binational Agricultural Research and Development Fund. This study was also supported by Minerva’s Otto Warburg Center for Plant Biotechnology. We thank the Wolfson Foundation for the use of facilities contributed to the Plant Science Institute.

**References**


