JOURNAL OF VIROLOGY, Aug. 1987, p. 2589–2598
0022-538X/87/082589-10$02.00/0

Vol. 61, No. 8

Physical Analysis of the *Campoletis sonorensis* Virus Multipartite Genome and Identification of a Family of Tandemly Repeated Elements

DAVID A. THEILMANN AND MAX D. SUMMERS*

Department of Entomology, Texas A&M University, and Texas Agricultural Experiment Station, College Station, Texas 77843-2475

Received 3 February 1987/Accepted 2 May 1987

This report is an analysis of cross-hybridizing sequences found within the 28 superhelical (SH) DNAs of the multipartite genome of the polydnavirus *Campoletis sonorensis* virus (CsV). A Southern cross-blot hybridization analysis showed that the majority of CsV *Eco*RI restriction fragments cross-hybridize to multiple *Eco*RI fragments. These sequence homologies were analyzed by hybridizing recombinant clones of the CsV SH DNAs B, H, M, and O[1] to Southern blots of undigested CsV DNA, using different hybridization stringencies. The results indicated that homologous regions among the SH DNAs include closely related sequences that are detectable under stringent conditions and related but more diverged sequences which are only detectable under reduced stringencies. A sequence that hybridized to the majority of the CsV SH DNAs was identified and subcloned from the SH DNAs O[1], H, and B. Nucleotide sequence data revealed that these homologous regions contained a family of imperfectly conserved repeated elements. These repeat elements were arranged singly or in direct tandem arrays and had an average length of 540 base pairs. Within the sequenced regions that contained the repeated elements six putative open reading frames were identified. These results show that the CsV genome consists of SH DNAs with complex sequence interrelationships that may have arisen due to multiple recombinational events.

Campoletis sonorensis virus (CsV; *Polydnaviridae*) replicates in the calyx epithelium of the reproductive tissue of the female parasitic wasp *C. sonorensis* (*Ichneumonidae*) (26). CsV is secreted into the oviduct lumen and during oviposition is coinjected with the wasp egg into the hemocoel of the lepidopteran host *Heliothis virescens* (*Noctuidae*) (26). If the wasp egg is artificially injected into *H. virescens* without CsV, the egg will be encapsulated by the host hemocytes (11). In the presence of CsV, the parasite egg will not be encapsulated and the egg will hatch, with the eventual result being the death of the host and the development of a *C. sonorensis* adult (11, 26). Thus it appears that by some presently unknown mechanism, CsV protects the parasite egg and is therefore required for the successful parasitism of *H. virescens* by *C. sonorensis*.

The injection of calyx fluid or purified CsV into *H. virescens*, which mimics natural parasitization, results in dramatic effects upon the development and physiology of the host. These effects include reduced growth rate and retarded development (27–29), changes in hormone titers (9a), and alterations to the number and behavior of plasmatocytes (7a). Molecular analyses of CsV transcription in *H. virescens* show that CsV transcripts are detected during the entire development of the endoparasitic wasp larva and that at least 12 different mRNAs can be identified (2, 3, 12). During this same period, CsV persists in *H. virescens* but no viral replication is detected (27).

The genome of CsV consists of at least 28 superhelical (SH) DNAs that range in size from 6 to greater than 20 kilobase pairs (kbp) (2, 16). In the initial analysis of the CsV genome, sequence relationships among the individual SH DNAs were analyzed by hybridizing labeled probes of gel-purified SH DNAs to Southern blots of CsV DNA (16). Because only limited cross-hybridization between the SH DNAs was detected, it was concluded that the CsV genome comprised mostly unique sequences. However, more recent data showed that when recombinant clones of CsV DNA were hybridized to the CsV genome, homology to multiple SH DNAs was observed (3, 13, 27). For example, Blissard et al. (3) showed that a CsV clone (p2H-5300), which contains sequences homologous to a viral mRNA, hybridizes strongly to at least 11 different SH DNAs. These results indicate that some of the SH DNAs of the CsV genome do not contain mostly unique sequences and that some DNA sequences were repeated on more than one SH DNA. In the present investigation, we conducted a more detailed study of the physical organization of the viral genome specifically to determine if the repetition of sequences was more common throughout the CsV genome than previously thought. In addition we analyzed sequence relationships between specific cloned SH DNAs and the rest of the CsV genome.

Our Southern cross-blot hybridization analysis demonstrated that the majority of the CsV *Eco*RI fragments hybridized to at least one region of the CsV genome other than themselves and usually to more than one. Cross-hybridization of individual CsV SH DNAs was analyzed by hybridization of cloned SH DNAs to the CsV genome under various hybridization stringencies. The results showed that each of the cloned SH DNAs analyzed contain sequences that cross-hybridize to other SH DNAs. Some of the cross-hybridizing regions could only be detected under conditions of reduced stringency. Fragments that cross-hybridize to most of the SH DNAs in the CsV genome were isolated and cloned from three separate SH DNAs. The nucleotide sequences of these homologous regions were determined and shown to consist of a family of repeated elements that were arranged singly or in tandem arrays. The results of this study indicated that the CsV genome comprises interrelated SH

* Corresponding author.

DNAs that may have arisen by multiple recombinational events and sequence divergence.

MATERIALS AND METHODS

Insect cultures and virus purification. The insects *H. virescens* and *C. sonorensis* were reared at 27°C on a 14-h–10-h (light:dark) photoperiod as described by Krell et al. (16). Third-instar *H. virescens* were parasitized by mated female *C. sonorensis* wasps as described by Fleming et al. (12).

CsV was purified from dissected female wasp oviducts by the method of Krell et al. (16). Viral DNA was extracted from purified virions by using the procedure described by Fleming et al. (12).

Southern blots. DNA samples were separated by agarose gel electrophoresis and transferred to nitrocellulose membranes by the procedures of Southern (25) and Smith and Summers (24).

[^{32}P]dATP-labeled probes (22) were hybridized to Southern blots by using various hybridization stringencies by the procedures of Howley et al. (15). Hybridizations of labeled probes were performed at 43°C in 50, 40, or 30% formamide (untreated; Bethesda Research Laboratories, Inc., Gaithersburg, Md.)–150 mg of calf thymus DNA per ml–5× Denhardt solution (0.02% Ficoll [Sigma Chemical Co., St. Louis, Mo.], 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin)–20 mM sodium phosphate (pH 6.5)–5× SSC (0.75 M NaCl, 0.075 M sodium citrate). The blots were washed three times in 2× SSC–0.1% sodium dodecyl sulfate at room temperature for 10 min, followed by three 30-min washes in the low salt buffer (0.1× SSC, 0.1% sodium dodecyl sulfate) at temperatures equivalent to the stringencies of the hybridization conditions. These temperatures were determined by using the equation $T_m = 81.5 + 16.6 (\log [Na^+]) + 0.41$ (% G+C)−0.72 (% formamide) (15). A G+C content of 50% was previously determined by Krell et al. (16) and was used in the calculations. Blots hybridized in 50, 40, or 30% formamide were washed at 50, 43, or 37°C, respectively. Under these conditions the 50, 40, and 30% formamide hybridization buffers will permit hybridization when homology is greater than 77, 70, and 64%, respectively, assuming that a 1% mismatch lowers the T_m by 1.0°C (19).

Southern cross-blot hybridization analysis. The Southern blot cross-hybridization was performed with GeneScreen and GeneScreen-Plus membranes by the procedure of the manufacturer (New England Nuclear Corp., Boston, Mass.). Briefly, CsV DNA was completely digested with the restriction enzyme *Eco*RI. The digest was separated into two portions, and one portion was 3′ end labeled with [^{32}P]dATP by using the Klenow fragment of DNA polymerase I. The labeled (2×10^4 dpm) and unlabeled (1 μg) *Eco*RI CsV digests were separated in parallel in two 0.7% agarose gels at 55 V for 15 h by using single, wide preparatory wells. Control lanes of the labeled and unlabeled digests were electrophoresed alongside the wide sample lanes. The separated unlabeled and labeled digests were transferred to GeneScreen-Plus and GeneScreen membranes, respectively, by capillary transfer as described by the manufacturer. The DNA was fixed to the GeneScreen-Plus membrane, placed at right angles to the labeled blot, immersed in hybridization buffer (0.5 M NaCl, 10% dextran sulfate, 50% formamide, 150 mg of calf thymus DNA per ml), sandwiched between two glass plates, and hybridized overnight at 43°C. The GeneScreen-Plus membrane was washed in 2× SSC–0.1% SDS at room temperature followed by washes in 0.1× SSC–0.1% sodium dodecyl sulfate at 50°C. The blot was then air dried and exposed to X-ray film (XAR-5; Eastman Kodak Co., Rochester, N.Y.) with an intensifying screen (Cronex Lightning-Plus; E. I. du Pont de Nemours & Co., Inc., Wilmington, Del.) at −80°C.

Sequencing. The sequencing reactions were performed by using the dideoxy-nucleotide procedure of Sanger et al. (23). DNA fragments to be sequenced were cloned into the bacteriophage vectors M13 mp18 and mp19 (31). Overlapping sets of deletions were generated by using the double-stranded DNA *Exo*III-*Exo*VII procedure described by Yanisch-Perron et al. (31) and the single-stranded DNA T4-DNA polymerase technique of Dale et al. (7). The sequences of both strands of all DNA fragments were determined.

Sequence data analysis was performed with the University of Wisconsin Genetics Computer Group DNA analysis software for the VAX minicomputer (8, 18).

Cloning of CsV DNAs. Standard methods used for the isolation, cloning, and purification of DNAs are described in Maniatis et al. (19). Clones of specific CsV SH DNAs were obtained by band isolating individual SH DNAs from 0.7% agarose gels in which undigested viral genomic DNA had been electrophoretically separated. The SH DNAs are named according to the previously established nomenclature (2, 16). The isolated DNAs were digested with different restriction enzymes, and the fragments were cloned into the plasmid vectors pUC8 or pUC18. Overlapping restriction maps of the cloned fragments were used to identify complete SH DNAs. CsV SH-B had been previously cloned (27). Single insert clones of SH-O^1 and -M were isolated from an existing pBR325 *Sal*I shotgun library of the CsV genome after the SH-DNAs were mapped from band-isolated subclones. The plasmid clones of the CsV SH DNAs B, M, and O^1 will be referred to as pSH-B, pSH-M, and pSH-O^1 respectively. The 8.4-kbp *Hin*dIII clone of CsV SH-H was previously identified as p4H-8460 (2, 3). This CsV SH-H clone is believed to contain the complete sequence of the CsV SH-H DNA and will be referred to as pSH-H.

RESULTS

To examine the overall cross-hybridization throughout the CsV genome a Southern cross-blot hybridization was performed by using an *Eco*RI digest of CsV genomic DNA. This procedure is similar to individually hybridizing each CsV *Eco*RI restriction fragment to a Southern blot of an *Eco*RI digest of the entire CsV genome. An *Eco*RI digest was used because the resulting fragments were distributed over a wide range of sizes, allowing the resolution of 54 bands (Fig. 1). Analysis of the Southern cross-blot hybridization (Fig. 1) shows an obvious diagonal of dark spots from the top left-hand corner to the bottom right-hand corner which represents the hybridization of each restriction fragment to itself. Spots off the diagonal represent the cross-hybridization between restriction fragments of different sizes. The regions above and below the diagonal are mirror images; variation in spot intensity on either side is due in part to differences in the sizes and molar amounts of the end-labeled probe fragments. The large number of spots occurring off the diagonal show that numerous cross-hybridizing sequences are found throughout the CsV genome. A majority of the 54 resolvable bands (48) cross-hybridized to at least one different band and usually to more than one (Fig. 1). We expected the large fragments to cross-hybridize to more of the DNA fragments because large fragments could poten-

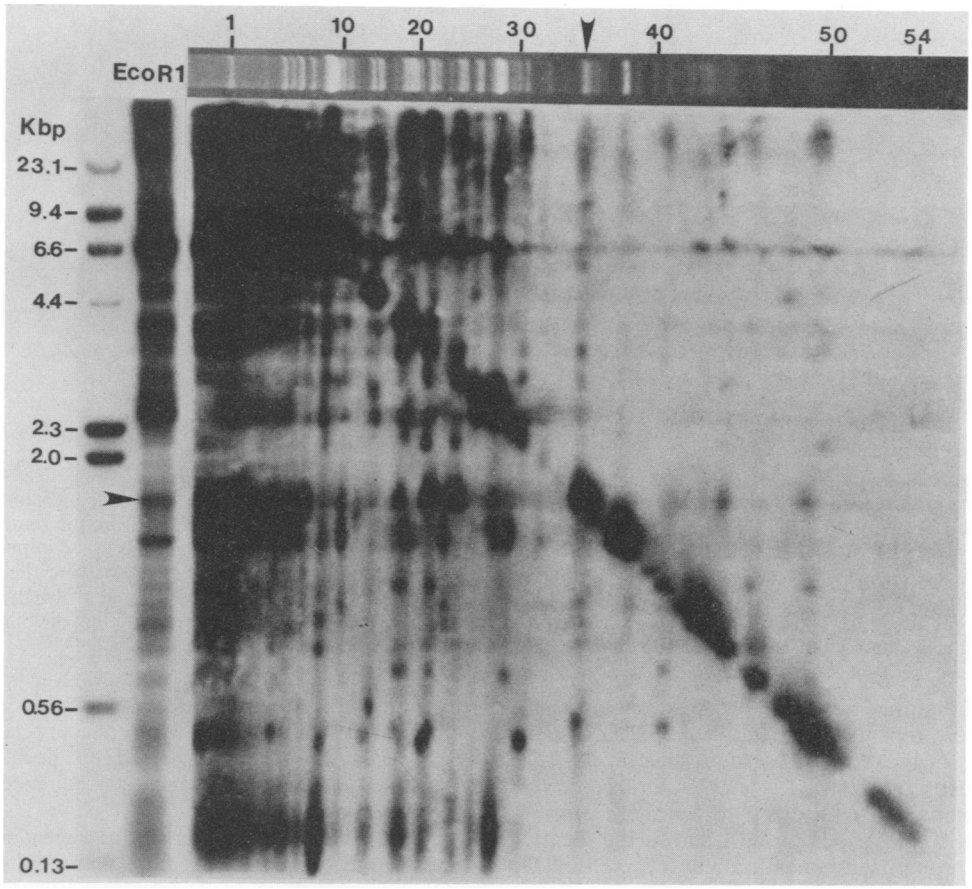

FIG. 1. Southern cross-blot hybridization of *Eco*RI-digested CsV DNA. The top panel is the ethidium bromide-stained marker lane of the unlabeled CsV digest. The numbers refer to the individual *Eco*RI fragments, with 1 being the largest and 54 the smallest detectable band. The side panel shows an autoradiogram of the [^{32}P]dATP end-labeled *Eco*RI digest of CsV DNA and of end-labeled molecular weight markers (*Hin*dIII digest of lambda phage DNA). The arrowheads at the top and side panels refer to band 34. The details of this procedure are described in the text.

tially contain a greater number of cross-hybridizing sequences. However, smaller DNAs also cross-hybridized to many restriction fragments. For example, band 34 (Fig. 1, arrows) is less than 2 kbp but hybridizes to bands 49, 46, 41, 32, 31, 30, 28, 24, 25, 22, 21, 19, 15, 11, 9, 8, 7, 4, and 1.

The Southern blot cross-hybridization analysis is a qualitative assessment of cross-hybridizing sequences throughout the CsV genome. Since a restriction digest of the CsV genome is used in this procedure and the origin of most of the CsV *Eco*RI fragments is unknown, the results do not specifically illustrate the nature of these sequence relationships. To analyze the sequence relationships between individual SH DNAs, recombinant clones of four CsV SH DNAs of known origin were used for analysis of cross-hybridizing sequences. They were pSH-B (6.6 kbp), pSH-H (8.4 kbp), pSH-M (10.8 kbp), and pSH-O^1 (11.2 kbp). These DNAs were chosen because they represent a wide range of the SH DNA sizes present in the CsV genome (Fig. 2A).

Figure 2B shows the hybridization of cloned SH DNAs to Southern blots of CsV genomic DNA under various conditions of hybridization stringency. Each of the cloned SH DNAs hybridized to multiple SH DNAs other than itself. The number of SH DNAs to which clones pSH-B, -H, and -O^1 cross-hybridized increased as the stringency of hybridization was decreased. SH-B cross-hybridized intensely to SH-Q and -A^1 under the lower stringencies of 40 and 30% formamide; minor cross-hybridization to five other SH DNAs was also observed (Fig. 2B). Under stringent conditions, pSH-H hybridized mainly to itself and showed only minor hybridization to SH-L^1 and -Q. However, as the stringency was reduced hybridization to SH-O^1, -R, and -T was also easily detected. A similar hybridization pattern has been previously reported for pSH-H (3). Unlike the other three SH DNAs tested, hybridization of pSH-M under conditions of reduced stringencies did not identify additional SH DNAs that contained homologous sequences.

Of the four cloned SH DNAs analyzed, pSH-O^1 cross-hybridized to the most DNAs. Under stringent conditions SH-O^1 hybridizes with an almost equal or greater intensity to SH-Q, -R and -U as it does to itself (Fig. 2B). These results indicated that the four SH DNAs O^1, Q, R, and U contain some sequences that are very closely related. Under low-stringency conditions (40 and 30% formamide), it was observed that SH-O^1 also hybridized at lower intensities to nearly every SH DNA. This includes DNAs that migrate below SH-A^1, at which no DNAs are visually detectable on the ethidium bromide-stained control lane (Fig. 2A). Low levels of hybridization to the majority of the SH DNAs could also be observed if the blot hybridized in 50% formamide was exposed to X-ray film for long periods.

Identification of repeated sequences. The low-level hybridization of SH-O^1 to a majority of the SH DNAs in the CsV

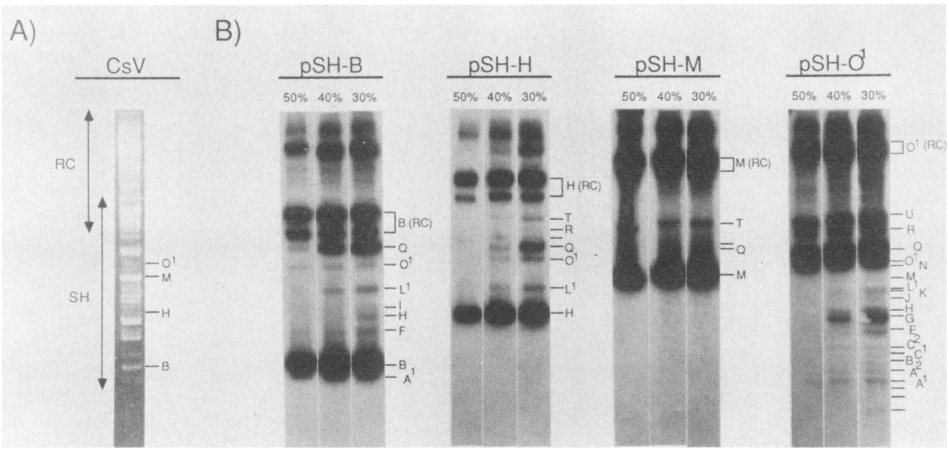

FIG. 2. Hybridization analysis of cloned CsV SH DNAs. (A) Ethidium bromide-stained 0.7% agarose gel of electrophoretically separated undigested CsV genomic DNA (1 μg). The locations of the cloned CsV SH DNAs are indicated on the right. The lines on the left indicate the regions of the relaxed circular (RC) and SH forms of the CsV DNAs. (B) Southern blots of duplicate lanes from the same gel shown in panel A were hybridized with pSH-B, -H, -M, and -O^1 as indicated. The stringency of the hybridization used, expressed in percent formamide, is shown above each lane. The probe concentration and specific activity and autoradiograph exposures times were the same for the three stringencies used for each of the individual cloned CsV SH DNAs. The SH DNAs to which each of the cloned probes hybridized are shown on the right side of each section. The RC forms of each probe are shown. Markers with no letter designation indicate bands for which there is no corresponding previously identified band in the ethidium bromide-stained control lane.

genome indicated that this molecule contained a sequence or sequences that were repeated on most of the SH DNAs. To map the locations of these repeated sequences on pSH-O^1, 17 overlapping restriction fragments of this molecule were gel purified, labeled, and hybridized to Southern blots of CsV genomic DNA (Fig. 3). Figure 3A shows the location and size of each probe used beneath a map of pSH-O^1. On the basis of the intensity of hybridization, one of two hybridization profiles was observed for all the probes tested. Of the 17 probes, 12 hybridized intensely to SH-O^1, -Q, -R, and -U and less intensely to the majority of the other SH DNAs and therefore were scored as containing the repeat (Fig. 3B, Repeat Positive). This was the same hybridization profile as that seen for the complete clone pSH-O^1 (Fig. 2). The other hybridization profile observed is also shown in Fig. 3B. Of the 17 probes, 5 hybridized intensely to SH-O^1, -Q, -R, and -U but did not have the less intense hybridization to the majority of the SH DNAs in the CsV genome. The fragments that gave this hybridization profile were scored as not containing the repeated sequences (Fig. 3B, Repeat Negative). The probes containing the repeated sequences mapped to two locations on SH-O^1 that were 1.1 kbp (Fig. 3A, region II) and 5.6 kbp (regions I and III) in size. The cumulative size of the repeat-positive fragments is 6.7 kbp, which accounts for approximately 60% of pSH-O^1.

To identify the regions on other SH molecules which hybridize to the repeated sequences of SH-O^1, the smaller 1.1-kbp region of SH-O^1 (Fig. 3A, region II) was subcloned into pUC8 and this clone, pO^1-HC1185, was used as a probe to map homologous sequences on the cloned SH DNAs B, H, and M. Clone pO^1-HC1185 was hybridized to Southern blots of restriction enzyme digests of clones pSH-O^1, pSH-H, pSH-B, and pSH-M. The regions on these molecules to which pO^1-HC1185 hybridized are shown in Fig. 4. On pSH-H a region homologous to pO^1-HC1185 was mapped to a 2.7-kbp region. On pSH-B, pO^1-HC1185 hybridized to a 1.0-kbp region but at very low levels. No hybridization of pO^1-HC1185 to pSH-M was detected. When pO^1-HC1185 was hybridized to restriction digests of the parent molecule pSH-O^1 hybridization was not detected to repeat region III, even under reduced (40% formamide) stringencies. This suggested that at least two types or classes of repeated sequences may be present on SH-O^1 because region III had been shown to contain repeated sequences by our hybridization assay (Fig. 3).

Nucleotide sequence analysis of repeat regions. Comparisons of the homologous regions identified on SH-B, -H, and -O^1 did not reveal any common restriction sites (Fig. 4). Therefore, the nucleotide sequence of pO^1-HC1185 and the homologous regions on pSH-B, and pSH-H were determined. The regions sequenced and the strategy used are shown in Fig. 4.

The regions sequenced on pSH-O^1, -H, and -B were 1,303, 2,743, and 1,153 bp in length, respectively. Dot matrix homology comparisons were used initially to analyze the three regions to identify homologous sequences (Fig. 5). The diagonal lines of a dot-plot comparison represent regions of homology, spaces in the lines indicate regions of sequence mismatch relative to the window size and stringencies used. The five major broken diagonal lines of the dot-plot comparison in Fig. 5A indicate that sequences on pSH-B between nucleotides (nt) 44 and 587 (BR1) are imperfectly repeated approximately 4.5 times (HR1 through HR5) in a direct tandem arrangement on the pSH-H molecule. A similar comparison between the pSH-O^1 and pSH-B sequences (Fig. 5C) indicates that the same region of pSH-B that is repeated on pSH-H is also repeated on pSH-O^1 approximately 2.5 times (OR1 through OR3) and that the sequences are also arranged in a direct tandem array. Comparison of the repeated sequences of pSH-H and pSH-O^1 (Fig. 5B) results in a complex pattern of five long overlapping diagonal lines that are due to the homology between the 4.5 repeats of pSH-H and the 2.5 repeats of pSH-O^1 which are organized in tandem arrays.

The dot plots showed that the sequences from the three SH DNAs share homology due to a repeated element present in 1, 4.5, and 2.5 copies on the sequenced regions of pSH-B, pSH-H, and pSH-O^1, respectively. The broken diagonal lines show that the homology between the individual repeated elements is not perfect. To analyze the relative

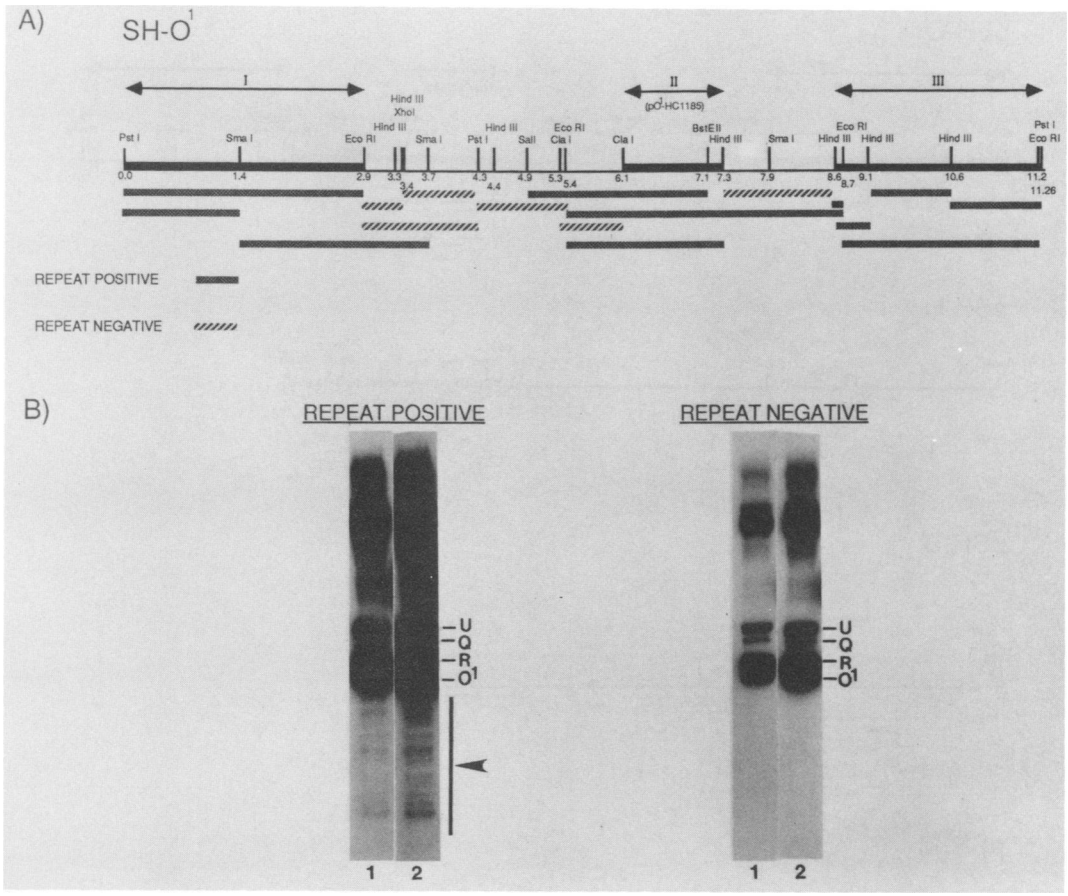

FIG. 3. Mapping of repeat sequences on the cloned SH-O[1]. (A) The locations of the 17 gel-purified restriction fragments used as hybridization probes to map the repeat sequences are shown beneath the restriction map of pSH-O[1]. The consensus locations of the repeat sequences determined by the probes are indicated by the solid black lanes on the restriction map (regions I, II, and III). Regions I and III would be joined in the circular molecule. (B) Hybridization patterns of the probes shown in panel A that either contained the repeated sequences (REPEAT POSITIVE) or did not contain the repeated sequences (REPEAT NEGATIVE). The labeled gel-purified probes were hybridized to Southern blots of undigested CsV genomic DNA (1 μg) under stringent conditions (50% formamide). The minor hybridization to multiple SH DNAs that was specific to the repeat-positive probes is indicated by the arrow. For both the repeat positive and repeat negative panels, lanes 1 and 2 represent 12- and 30-h autoradiograph exposures, respectively.

sequence composition of all the partial or complete repeat sequences from pSH-B, -H, and -O[1], all were optimally aligned by computer and a consensus sequence was determined (Fig. 6).

The average length of each repeat element is approximately 540 bp, with each repeat approximately 70% homologous to the consensus repeat. The homology between individual repeats varies from 60 to 70%, with short segments (less than 50 bp) having 80 to 100% homology. The repeat element HR5 contains a sequence of 78 bp that does not share homology with any of the other repeats (Fig. 6). The reduced or partial repeat elements on the ends of the sequenced regions of pSH-O[1] (OR1 and OR3) and pSH-H (HR1) suggest that these repeat sequences extend beyond the regions analyzed. Analysis of the aligned sequences indicated that there are no specific regions within the repeat element that appear to be highly conserved among all the repeats.

Analysis of the sequence data from the three homologous regions on SH-H, -O[1], and -B has also revealed the presence of six open reading frames (ORFs) which are shown schematically in Fig. 7. The ORFs range in size from 291 to 996 nt. On pSH-H, three ORFs have been identified, SH-O[1] contains one ORF that starts at nt 326 and continues through to the end of the sequenced region (853 nt), and the pSH-B repeat sequence contains two small ORFs. Comparison of the computer-generated translation products of these ORFs shows homologies among the predicted proteins. Alignment of the predicted protein sequences of H ORF 1 and H ORF 2 identified a region of 70 amino acids that was 69% homologous (data not shown). Similar results were obtained when each of the other predicted proteins from the six ORFs were compared with each other.

DISCUSSION

The polydnavirus family is the only group of double-stranded DNA viruses that has segmented genomes (4). The CsV genome has at least 28 circular SH DNAs that range from approximately 6.0 to greater than 20 kbp, with a composite size greater than 270 kbp (2, 16). The large genome size and its segmented organization makes CsV one of the largest and most structurally complex of the known animal viruses. Preliminary analysis of the CsV genome suggested that all of the SH DNAs were for the most part unique in sequence content (16). Contrary to this, recent

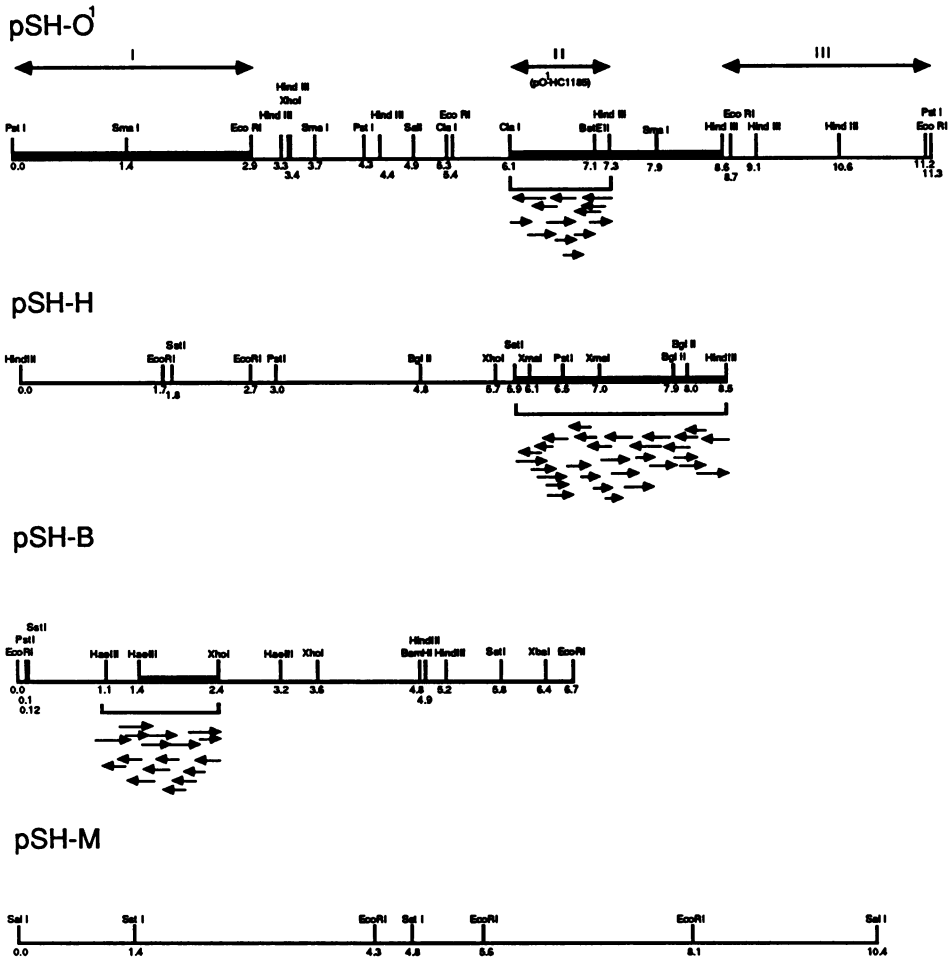

FIG. 4. Mapping and sequencing strategies of the regions homologous to the repeat plasmid pO[1]-HC1185. Restriction maps of the four cloned SH DNAs (pSH-O[1], -H, -B, and -M) are shown, with the regions homologous to pO[1]-HC1185 (determined under stringent conditions [50% formamide]) indicated by solid black lines. Sequenced regions of pSH-B, -H, and -O[1] and the sequencing strategies used, as indicated by the arrows, are shown below each map. The regions on SH-O[1] shown to contain repeated sequences (Fig. 3) are indicated (I, II, and III).

data have shown that cloned fragments of CsV DNA will cross-hybridize to as many as 11 different SH DNAs (3, 13, 27). In the present study we performed a more detailed analysis of the physical relationships among the multiple SH DNAs that constitute the CsV genome.

To assess the degree of cross-hybridizing sequences throughout the CsV genome we performed a Southern cross-blot hybridization (Fig. 1) which indicated that sequence homology exists between the majority of the SH DNAs and that most of the SH DNAs are not composed of unique sequences as previously suggested. This result reveals that the CsV genome is composed of interrelated SH DNAs.

Hybridization of DNAs under different stringencies has been used to analyze various degrees of sequence homologies and evolutionary relatedness among viral genomes, gene families, or specific DNA sequences (1, 6, 14, 15). In this study we used variation in hybridization conditions to analyze the sequence relationships among several SH DNAs of the CsV genome. Our hybridization data with four cloned SH DNAs (Fig. 2) revealed that relationships among different DNAs of the CsV genome include both closely related DNA sequences for which homology is detected under stringent conditions and related but diverged sequences which are more easily detected under reduced stringencies.

Hybridization of 17 restriction fragments of SH-O[1] indicated that every region tested hybridized intensely to the molecules O[1], Q, R, and U under stringent conditions (Fig. 3), which suggests that the SH-O[1] molecule is duplicated within these larger SH DNAs. It is possible that the larger SH DNAs (Q, R, and U) have undergone intramolecular recombination to generate smaller circular molecules that will be duplications of portions of the larger molecules. Palmer and Shields (20) reported that the 218-kbp plant mitochondrial genome of *Brassica campestris* appears to recombine via intramolecular recombination between direct repeats to generate smaller circular DNAs which are duplications of part of the parent mitochondrial molecule (20). This type of recombination may explain the cross-hybridization results we observed among SH-O[1], -Q, -R, and -U. Cloning and mapping of the larger CsV SH DNAs Q, R, and U will be required to determine whether they do indeed contain all the sequences of SH-O[1]. A 7.2-kbp CsV clone from the CsV-Q region has already been cloned and analyzed (2, 3, 13), but the restriction map of this clone shows no similarity to that of SH-O[1], which suggests that there may be comigrating bands in the Q region of the CsV genome.

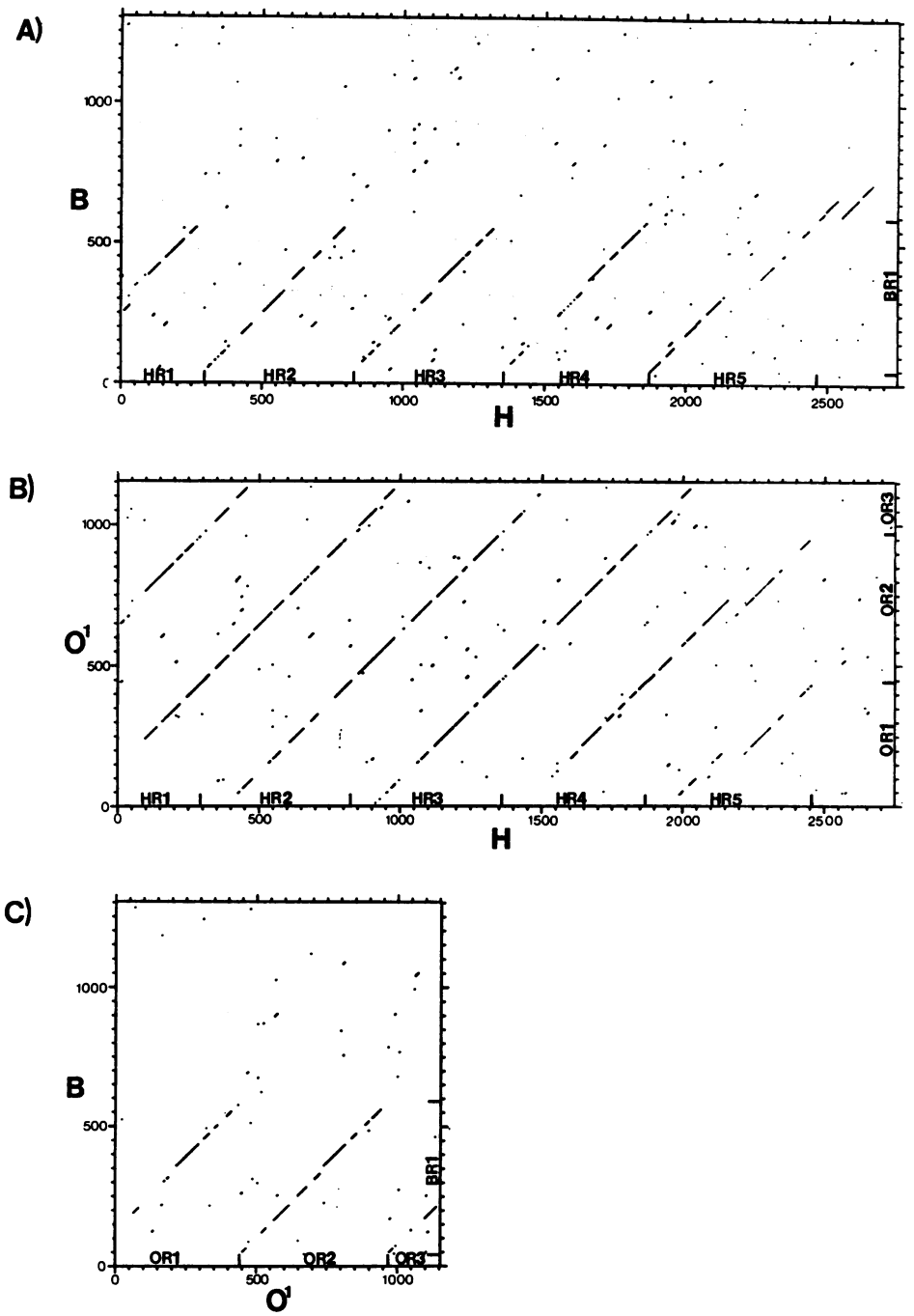

FIG. 5. Dot matrix homology plot comparisons of the repeat sequences from (A) pSH-H and pSH-B, (B) pSH-H and pSH-O[1], and (C) pSH-O[1] and pSH-B. The stringency of the comparison was a minimum 14-bp match in a window of 21 bp. The numbers on the axes refer to nucleotide base pairs. The locations of the repeat elements pSH-H repeat 1 (HR1) to pSH-H repeat 5 (HR5), pSH-O[1] repeat 1 (OR1) to pSH-O[1] repeat 3 (OR3), and pSH-B repeat 1 (BR1) are shown on the axes.

In this study we also identified, cloned, and sequenced a region of pSH-O[1] that cross-hybridizes under low stringency to the majority of the SH DNAs in the CsV genome. The sequence data from pSH-O[1] and homologous sequences from pSH-B and -H revealed that these three regions consisted of an imperfectly conserved repeated element with an average size of 540 bp. On pSH-O[1] and pSH-H the repeat elements were organized in tandem arrays, but only a single element was identified on pSH-B (Fig. 4). Comparisons of all the sequenced repeat elements suggest that no specific sequences of the repeat elements are highly conserved (Fig. 6). The sequence data showed that the homology among the repeated elements was variable, with an average homology of 60 to 70%, but that the repeats also contained short regions that were greater than 90% homologous. This explains in part why we observed increased cross-hybridization when low-stringency conditions were used (Fig. 2, pSH-O[1]). Analysis of the consensus sequence that was

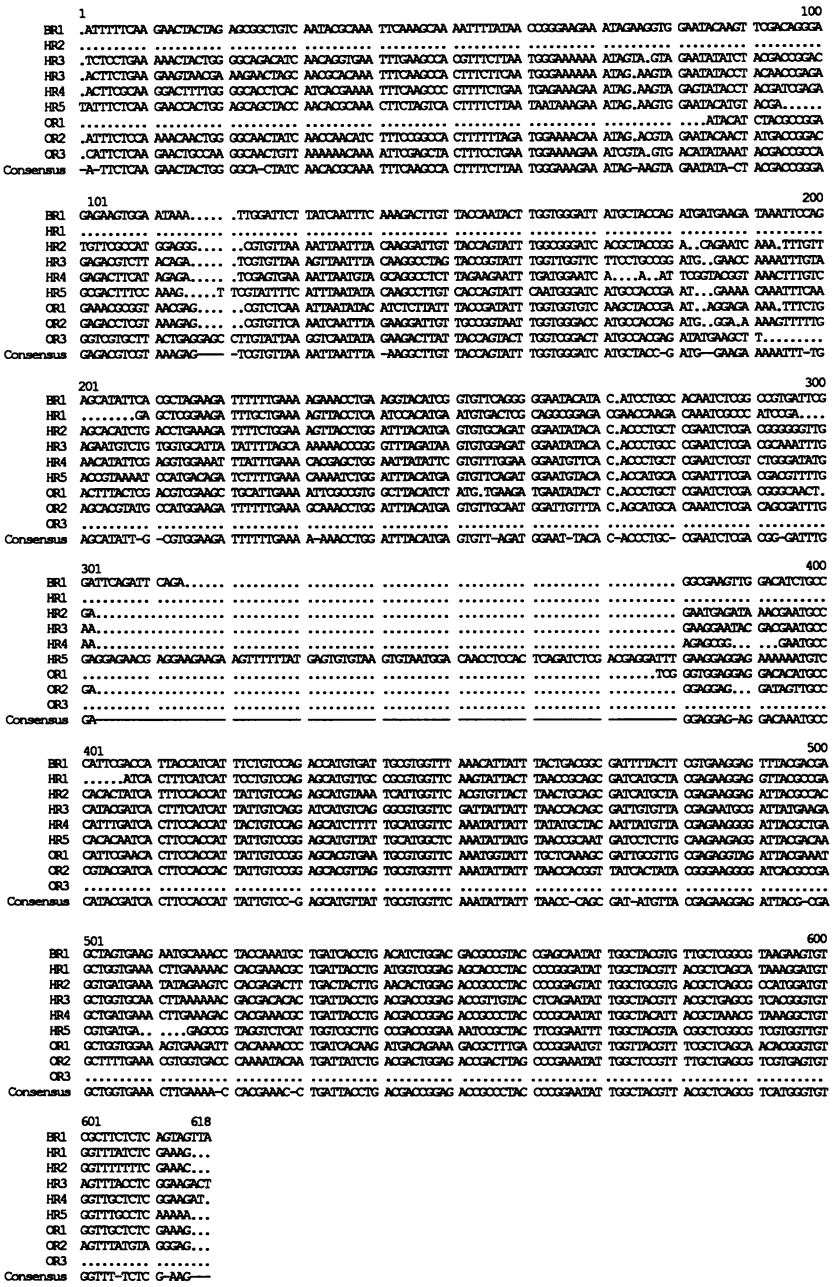

FIG. 6. Optimal alignment of the nine sequenced repeat elements (partial and complete) from pSH-B, -H, and -O[1]. The locations of the five repeats from pSH-H (HR1, nt 1 through 297; HR2, nt 298 through 822; HR3, nt 823 through 1351; HR4, nt 1352 through 1869; HR5, nt 1870 through 2462), the three repeats from pSH-O[1] (OR1, nt 1 through 442; OR2, nt 443 through 964; OR3, nt 965 through 1153), and the single repeat of pSH-B (BR1, nt 44 through 587) were determined from the dot matrix homology plots in Fig. 5. Gaps (periods) were inserted into the sequences to obtain a best fit. Dashes indicate that no consensus nucleotide was found for that specific location. The plurality used for the consensus sequence was 3.

determined by a comparison of the nine sequenced repeat elements (Fig. 6) did not reveal any distinguishing features such as the internal inverted or direct repeats which have been shown to exist in other eucaryotic repeated elements (9). The consensus sequence may be similar to an ancestral DNA that was the progenitor of the family of repeats we identified in this study.

The sequence data from pSH-O[1], -H, and -B also revealed six ORFs within the sequences of the repeated elements (Fig. 7). Comparison of the computer-generated predicted translation products of each of the six ORFs showed amino acid homology. Recently, G. W. Blissard, O. P. Smith, and M. D. Summers (Virology, in press) identified two CsV genes expressed in parasitized *H. virescens* that share significant nucleotide and amino acid sequence homology and appear to represent two members of a larger, related viral gene family. The nucleotide sequences of these two related genes show no homology to the repeated elements identified in this study. The ORFs identified in the 540-bp repeat elements of pSH-B, -H, and -O[1] may represent another

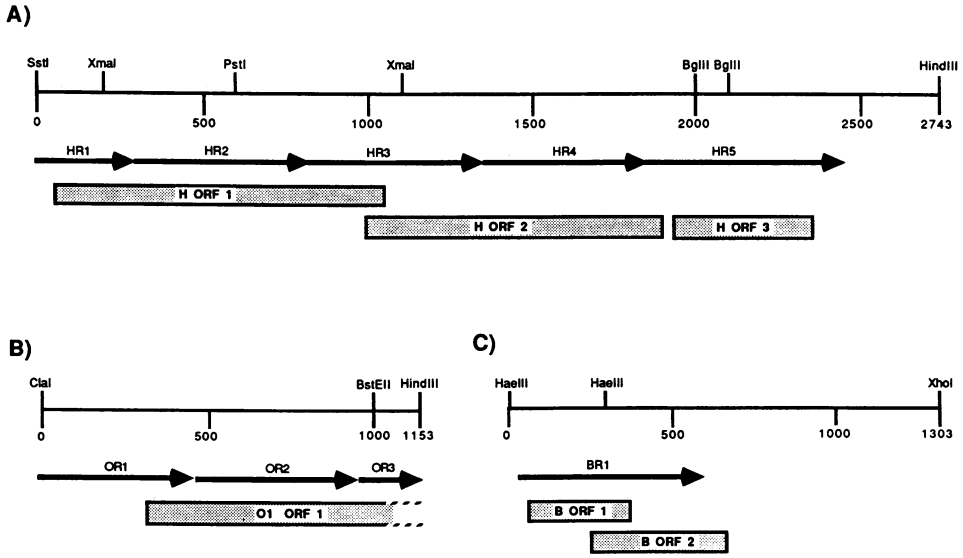

FIG. 7. Summary schematic showing the location of the 540-bp repeat elements and the ORFs of the sequenced regions of pSH-H (A), pSH-O[1] (B), and pSH-B (C). The locations of the direct repeat elements are shown by the arrows. The sizes of each ORF are as follows: 903 nt (H ORF 1), 996 nt (H ORF 2), 440 nt (H ORF 3), >853 nt (O ORF 1), 291 nt (B ORF 1), and 375 nt (B ORF 2). The open end of O[1] ORF 1 indicates that the end of the ORF is beyond the sequenced region.

example of a family of related CsV genes. Studies are currently in progress to investigate which, if any, of the six ORFs identified on the cloned SH DNAs B, H, or O[1] are actively transcribed into CsV mRNAs in either *C. sonorensis* or parasitized *H. virescens*. A 3.2-kb mRNA expressed in parasitized *H. virescens* which is homologous to and is believed to be transcribed from pSH-H has been previously reported (2, 3). This mRNA would be large enough to code for any or all of the three ORFs we identified on pSH-H.

The organization of repeat elements in tandem arrays of various lengths on the multiple SH DNAs of the CsV genome could provide a substrate for a large number of intra- and intermolecular homologous recombinational events, including equal and unequal crossover (10, 17). Multiple recombinational events combined with sequence divergence could result in a set of DNAs that would have very complex sequence relationships similar to what we presently observe for the CsV genome. It has been suggested that recombinational events between tandem arrays of repeats may provide a method of gene duplication and mutation of eucaryotic genes (10, 30). In the *C. sonorensis-H. virescens* parasite-host system, multiple recombinational events in CsV could be of evolutionary advantage for the parasitic wasp. Such changes may be related to the ability of the parasitic wasp to successfully parasitize its lepidopteran hosts that may be evolving new or altered defensive mechanisms.

In tissues of male and female *C. sonorensis*, sequences homologous to CsV DNA are found in nonviral forms which may be indicative of CsV integration in the wasp genome (13). Repeat elements in other viral and eucaryotic systems have been shown to be involved in genomic integration events (5, 21). If the replicative cycle of CsV does involve integrated viral DNAs, the 540-bp repeated elements identified in this study may also play a similar role.

In summary, the data in this study show that the multipartite CsV genome is composed of interrelated SH DNAs, indicative of a complex evolutionary history. Shared sequences between SH DNAs include both closely related and diverged sequences; an example of the latter being the imperfectly conserved repeated elements that are homologous to most of the SH DNAs in the CsV genome. The presence of ORFs within the sequenced repeated elements suggests that they may represent a group of related viral genes. Alternatively, the repeated elements may be sites of homologous recombination between the SH DNA molecules or the putative integrated forms of viral DNA. Transcriptional analysis of the repeat sequences in either *C. sonorensis* or parasitized *H. virescens* may reveal some of the functional aspects of this family of repeated elements.

ACKNOWLEDGMENTS

The authors thank Gary W. Blissard for very helpful discussions and critical review of this manuscript and Burt Beames for editorial comments. We also thank Nelida Angulo for excellent technical assistance in the rearing of insects.

This work was supported by National Science Foundation grant PCM-8319002 and by Texas Agricultural Experiment Station Project 6313. D.A.T. is supported by the Research Scientist Trainee Program of Agriculture Canada (Vancouver).

LITERATURE CITED

1. **Beltz, G. A., K. A. Jacobs, T. H. Eickbush, P. T. Cherbas, and F. C. Kafatos.** 1983. Isolation of multigene families and determination of homologies by filter hybridization methods. Methods Enzymol. **100**:266–285.
2. **Blissard, G. W., J. G. W. Fleming, S. B. Vinson, and M. D. Summers.** 1986. *Campoletis sonorensis* virus: expression in *Heliothis virescens* and identification of expressed sequences. J. Insect Physiol. **32**:351–359.
3. **Blissard, G. W., S. B. Vinson, and M. D. Summers.** 1986. Identification, mapping, and in vitro translation of *Campoletis sonorensis* virus mRNAs from parasitized *Heliothis virescens* larvae. J. Virol. **57**:318–327.
4. **Brown, F.** 1986. The classification and nomenclature of viruses: summary of results of meetings of the International Committee on Taxonomy of Viruses in Sendai, September 1984. Intervirology **25**:141–143.
5. **Carlos, P. M., and J. H. Miller.** 1980. Transposable elements.

Cell **20**:589–595.

6. **Cochran, M. A., and P. Faulkner.** 1983. Location of homologous sequences interspersed at five regions in the baculovirus AcMNPV genome. J. Virol. **45**:961–970.
7. **Dale, R. M., B. A. McClure, and J. P. Houchins.** 1985. A rapid single stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18 S rDNA. Plasmid **13**:31–40.

7a. **Davies, D. H., M. R. Strand, and S. B. Vinson.** 1987. Changes in differential haemocyte count and in vitro behaviour of plasmatocytes from host *Heliothis virescens* caused by *Campoletis sonorensis* polydnavirus. J. Insect Physiol. **33**:143–153.

8. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12**:387–395.
9. **Doolittle, W. F.** 1985. The evolutionary significance of middle repetitive DNAs, p. 443–487. *In* T. Cavalier-Smith (ed.), The evolution of genome size. John Wiley & Sons, Inc., New York.

9a. **Dover, B. A., D. H. Davies, M. R. Strand, R. S. Gray, L. L. Keeley, and S. B. Vinson.** 1987. Ecdysteroid-titre reduction and developmental arrest of last-instar *Heliothis virescens* larvae by calyx fluid from the parasitoid *Campoletis sonorensis*. J. Insect Physiol. **33**:333–338.

10. **Dyson, P., and D. Sherratt.** 1985. Molecular mechanisms of duplication, deletion, and transposition, p. 353–395. *In* T. Cavalier-Smith (ed.), The evolution of genome size. John Wiley & Sons, Inc., New York.
11. **Edson, K. M., S. B. Vinson, D. B. Stoltz, and M. D. Summers.** 1981. Virus in a parasitoid wasp: suppression of the cellular immune response in the parasitoid's host. Science **211**:582–583.
12. **Fleming, J. G. W., G. W. Blissard, M. D. Summers, and S. B. Vinson.** 1983. Expression of *Campoletis sonorensis* virus in the parasitized host, *Heliothis virescens*. J. Virol. **48**:74–78.
13. **Fleming, J. G. W., and M. D. Summers.** 1986. *Campoletis sonorensis* endoparasitic wasps contain forms of *C. sonorensis* virus DNA suggestive of integrated and extrachromosomal polydnavirus DNAs. J. Virol. **57**:552–562.
14. **Galinski, M. S., V. H. Stanik, G. F. Rohrmann, and G. S. Beaudreau.** 1983. Comparison of sequence diversity in several cytoplasmic polyhedrosis viruses. Virology **130**:372–380.
15. **Howley, P. M., M. A. Isarael, M. Law, and M. A. Martin.** 1979. A rapid method for detecting and mapping homology between heterologous DNAs. J. Biol. Chem. **254**:4876–4883.
16. **Krell, P. J., M. D. Summers, and S. B. Vinson.** 1982. Virus with a multipartite superhelical DNA genome from the ichneumonid parasitoid *Campoletis sonorensis* J. Virol. **43**:859–870.
17. **Lewin, B.** 1980. Gene expression, vol. 2, p. 503–569. John Wiley & Sons, Inc., New York.
18. **Maizel, J. V., and R. P. Lenk.** 1981. Enhanced matrix analysis of nucleic acid and protein sequences. Proc. Natl. Acad. Sci. USA **78**:7665–7669.
19. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Palmer, J. D., and C. R. Shields.** 1984. Tripartite structure of the *Brassica campestris* mitochondrial genome. Nature (London) **307**:437–440.
21. **Panganiban, A. T.** 1985. Retroviral DNA integration. Cell **42**: 5–6.
22. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
23. **Sanger, F., S. Nicklem, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
24. **Smith, G. E., and M. D. Summers.** 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. Anal. Biochem. **109**:123–129.
25. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
26. **Stoltz, D. B., and S. B. Vinson.** 1979. Viruses and parasitism in insects. Adv. Virus Res. **24**:125–171.
27. **Theilmann, D. A., and M. D. Summers.** 1986. Molecular analysis of *Campoletis sonorensis* virus DNA in the lepidopteran host, *Heliothis virescens*. J. Gen. Virol. **67**:1961–1969.
28. **Vinson, S. B.** 1972. Effect of the parasitoid, *Campoletis sonorensis*, on the growth of its host, *Heliothis virescens*. J. Insect Physiol. **18**:1509–1514.
29. **Vinson, S. B., K. M. Edson, and D. B. Stoltz.** 1979. Effect of a virus associated with the reproductive system of the parasitoid wasp, *Campoletis sonorensis*, on host weight gain. J. Invert. Pathol. **34**:133–137.
30. **Wilson, A. C., S. S. Carlson, and T. White.** 1977. Biochemical evolution. Annu. Rev. Biochem. **46**:573–639.
31. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**: 103–119.