Genome Sequence of a Polydnavirus: Insights into Symbiotic Virus Evolution

Eric Espagne, *et al.*

Science **306**, 286 (2004);
DOI: 10.1126/science.1103066

of tri- and tetrapeptides (Table 1, entry 10). When an excess of the oxidizing agent was used with phenylalanine thiocarbamate **2**, LC-MS established that a 63% yield of dipeptide was obtained in just 5 min, along with 13% tripeptide, 3% tetrapeptide, and traces of penta- and hexapeptide (Table 1, entry 11).

In experiments in which a mixture of L-serine (Ser, 50 mM) and the phenylalanine thiocarbamate **2** (25 mM) in CHES (400 mM, pH 9.0) were allowed to react, either in the presence of $CdCl_2$ (25 mM) or $K_3Fe(CN)_6$ (25 mM), a mixture of peptides was produced corresponding to Phe-Ser, Phe-Phe, Phe-Phe-Ser, and Phe-Phe-Phe. No homopolymers of serine were observed. In another experiment, a mixture of L-serine and L-phenylalanine was exposed to COS (Table 2, entry 4). In contrast to the previous reaction, Ser-Ser and Ser-Ser-Ser were produced, along with polymers of phenylalanine and mixed peptides (Fig. 1B). These observations strongly suggest that the activated α-aminoacyl compound derives from the thiocarbamate structure and that, once activation has occurred, peptide formation proceeds via nucleophilic attack by a second α-amino acid molecule on the in situ–formed NCA. The generality of the COS-mediated α-amino acid condensation reactions in the presence of Pb^{2+} was established with reaction mixtures containing equimolar mixtures of L-phenylalanine and either L-tyrosine, L-leucine, L-alanine, or L-serine (Table 2, fig. S5). In all reactions, efficient production of mixed dipeptides and tripeptides was observed.

Present-day levels of COS in volcanic gases have been reported up to 0.09 mol % (*14*). Because the gas hydrolyzes rapidly on a geological time scale, it is unlikely to have accumulated to a high concentration in the atmosphere. Thus, if COS was important in prebiotic chemistry, it is likely to have functioned in localized regions close to its volcanic sources. Although it may be unlikely that a substantial proportion of any amino acids present would have been converted to thiocarbamates, this would have been no obstacle to a "polymerization on the rocks" scenario (*15*, *16*) in which peptides long enough to be irreversibly adsorbed near the source of the COS were subject to slow chain elongation. The direct elongation of peptide chains using COS as a condensing agent and the condensations catalyzed by Fe^{2+} or Pb^{2+} ions seem plausible as prebiotic reactions (*17*). The very efficient polymerizations brought about by oxidizing agents are more problematic as prebiotic reactions, but $[Fe(CN)_6]^{3+}$ has been discussed as a potential prebiotic oxidizing agent (*13*).

It remains to be determined whether COS could have participated in prebiotic chemistry in other ways—for example, as an intermediate in the reduction of CO_2 (*18*, *19*) and as a condensing agent in phosphate chemistry (*20*, *21*).

References and Notes

1. R. S. Dewey *et al.*, *J. Org. Chem.* **36**, 49 (1971).
2. C. Huber, G. Wachtershauser, *Science* **281**, 670 (1998).
3. C. Huber, W. Eisenreich, S. Hecht, G. Wachtershauser, *Science* **301**, 938 (2003).
4. Materials and methods are available as supporting material on *Science* Online.
5. COS is reported to dissolve in water at room temperature to give up to 20 to 30 mM solutions (*6*, *22*).
6. R. J. Ferm, *Chem. Rev.* **57**, 621 (1957).
7. During the course of the reaction substantial quantities of H_2S are generated, for example, through the hydrolysis of COS. Attack of HS^- on the NCA would generate α-amino thioacids that can participate in the formation of peptides and side products (*23*).
8. The observed half-life of phenylalanine thiocarbamate **2** (25 mM in D_2O, pD 8.6) formed in situ from the amino acid and COS was 10 hours. In a separate NMR study using an authentic sample of **2** (50 mM in D_2O, pD 9.0), a hydrolysis half-life of ~20 hours was observed.
9. Condensations of NCAs with free amino acids (100 mM each in borate buffer pH ≈ 10) at 4°C are typically complete in less than 2 min (*1*, *24*).
10. Metal ions that might be present as impurities in the sample are not required for condensation, as demonstrated by formation of product in the presence of the metal chelator EDTA (Table 1, entry 4).
11. R. Liu, L. E. Orgel, *Nature* **389**, 52 (1997).
12. J. F. Kasting, L. L. Brown, in *The Molecular Origins of Life*, A. Brack, Ed. (Cambridge Univ. Press, New York, 1998), pp. 35–56.
13. A. D. Keefe, S. L. Miller, *Origins Life Evol. Biosphere* **2**, 111 (1996).
14. R. B. Symonds, W. I. Rose, G. J. S. Bluth, T. M. Gerlach, *Rev. Mineral.* **30**, 1 (1994).
15. L. E. Orgel, *Origins Life Evol. Biosphere* **28**, 227 (1998).
16. A. R. Hill Jr., C. Bohler, L. E. Orgel, *Origins Life Evol. Biosphere* **28**, 235 (1998).
17. Alternative potentially prebiotic condensing agents with relatively high efficiency are inorganic polyphosphates (*25*, *26*).
18. W. Heinen, A. M. Lauwers, *Origins Life Evol. Biosphere* **2**, 131 (1996).
19. D. R. Herrington, P. L. Kuch, U.S. Patent 4,618,723 (1986).
20. W. C. Buningh, U.S. Patent 3,507,613 (1970).
21. J.-P. Biron, R. Pascal, *J. Am. Chem. Soc.* **126**, 9189 (2004).
22. U.S. Environmental Protection Agency, *Chemical Summary for Carbonyl Sulfide* (Publication 749-F-94-009a, Environmental Protection Agency, Washington, DC, 1994; www.epa.gov/chemfact/s_carbns.txt).
23. T. Wieland, K. E. Euler, *Chem. Ber.* **91**, 2305 (1958).
24. R. Hirschmann *et al.*, *J. Org. Chem.* **32**, 3415 (1967).
25. Y. Yamagata, H. Watanabe, M. Saitoh, T. Namba, *Nature* **352**, 516 (1991).
26. J. Rabinowitz, J. Flores, R. Kresbach, G. Rogers, *Nature* **224**, 795 (1969).

27. We thank NASA Astrobiology Institute and NASA Exobiology (NAG5-12160) for financial support. L.L. is the recipient of an NSF Predoctoral Fellowship.

Supporting Online Material
www.sciencemag.org/cgi/content/full/306/5694/283/DC1
Materials and Methods
Figs. S1 to S5
Reference

13 July 2004; accepted 25 August 2004

Genome Sequence of a Polydnavirus: Insights into Symbiotic Virus Evolution

Eric Espagne,[1]*† Catherine Dupuy,[1]†‡ Elisabeth Huguet,[1] Laurence Cattolico,[2] Bertille Provost,[1] Nathalie Martins,[2] Marylène Poirié,[1] Georges Periquet,[1] Jean Michel Drezen[1]

Little is known of the fate of viruses involved in long-term obligatory associations with eukaryotes. For example, many species of parasitoid wasps have symbiotic viruses to manipulate host defenses and to allow development of parasitoid larvae. The complete nucleotide sequence of the DNA enclosed in the virus particles injected by a parasitoid wasp revealed a complex organization, resembling a eukaryote genomic region more than a viral genome. Although endocellular symbiont genomes have undergone a dramatic loss of genes, the evolution of symbiotic viruses appears to be characterized by extensive duplication of virulence genes coding for truncated versions of cellular proteins.

Once regarded as a rare biological event, symbiosis is now known to be central to the origin of eukaryotic cellular organelles. The genomes of mitochondria and plastids are known to be dramatically reduced compared with those of their ancestors—free-living bacteria (*1*). There are also examples of viral symbionts, but almost nothing is known about the genome rearrangements these have undergone during their evolution.

Polydnaviruses (PDVs) are used by parasitoid wasps to facilitate development of their progeny within the body of immunocompetent insect hosts, which are typically lepidopteran larvae (*2*). Viral particles are produced in the

[1]Institut de Recherche sur la Biologie de l'Insecte, CNRS UMR 6035, UFR Sciences et Techniques, Parc de Grandmont, 37200 Tours, France. [2]Genoscope, Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 Evry, France.

*Present address: Institut de Génétique et Microbiologie, Université Paris Sud, Bat. 400, 91405 Orsay cedex, France.
†These authors contributed equally to this work.
‡To whom correspondence should be addressed. E-mail: catherine.dupuy@univ-tours.fr

wasp ovaries and are injected via the wasp ovipositor into the insect host along with the parasitoid eggs (*2*). Viral gene products act by manipulating host immune defenses and development, thereby ensuring the emergence of adult parasitoid wasps (*3*). Unlike most viruses, polydnaviruses are not transmitted by infection, because no virus replication occurs in parasitized host tissues. They are exclusively inherited as an endogenous "provirus" integrated in the wasp genome (*4*–*6*).

The Polydnaviridae are a unique insect virus family on the basis of the molecular features of their genome and of their obligate association with endoparasitoid wasps (*7*, *8*). They are composed of two genera, bracoviruses and ichnoviruses, associated with braconid and ichneumonid wasps, respectively, with distinct evolutionary origins (*2*). Bracovirus-bearing species have a common ancestor (*9*). The classical hypothesis is that bracoviruses originate from an "ancestor virus" initially integrated into the genome of the ancestor wasp species that lived 73.7 ± 10 million years ago (*10*).

Several PDV genes expressed in parasitized host tissues have been isolated from various wasp species but the organization and content of PDV genomes are largely unknown (*11*). Here, we present the complete nucleotide sequence of the bracovirus (CcBV) injected by the wasp *Cotesia congregata* into its lepidopteran host *Manduca sexta*.

With a full length of 567,670 base pairs (bp), the CcBV genome (Table 1) is one of the largest viral genomes sequenced so far (*11*). The segmented genome is composed of 30 DNA circles ranging from 5 to 40 kb and contains 156 coding DNA sequences (CDSs) (Fig. 1). The overall sequence displays a strong bias toward A-T content (66%), and more than 70% of the sequence corresponds to noncoding DNA. The circles encode at least one gene (with the exception of a single noncoding circle), and the percentage of potential coding sequences varies from 7.4 to 53.9% depending on the circle, a gene density that is markedly different from the highly compact structure of a "classical" virus genome. Unlike most viral genes, many CcBV genes contain introns (69%), and 42.3% of putative CDSs have no similarity to previously described genes (Fig. 2). Another unique feature of the CcBV genome, compared with classical viruses, is the abundance of gene families: 66 genes (42.5%) are organized in nine families (Table 2). It is noteworthy that the proteins encoded by four of these gene families contain highly conserved domains previously described in virulence factors used by bacterial pathogens or parasitic nematodes.

Table 1. Genomic features of CcBV (*Cotesia congregata* bracovirus).

Genomic features	Complete genome
Length (bp)	567670
A+T ratio (%)	66.05
Percent coding sequence	26.9
tRNA coding genes	7
Predicted genes encoding proteins	156
Genes with functional assignments	42
LTR and transposons	10

The largest CcBV gene family comprises 27 genes encoding protein tyrosine phosphatases (CcBV PTP). PTPs are known to play a key role in the control of signal transduction pathways by dephosphorylating tyrosine residues on regulatory proteins (*12*). We recently identified PTPs in bracoviruses of two distantly related braconid subfamilies (*13*) (Table 2), which suggests that they constitute a common component of bracovirus genomes. Bracovirus PTPs share significant similarity with cellular PTPs, but they are not homologous to baculovirus or poxviruses PTPs, which counters the hypothesis that bracoviruses originated from baculoviruses as initially suspected (*14*). Note hat some bacterial pathogens, such as *Yersinia pestis*, inhibit host macrophage phagocytosis by injecting PTPs that interfere with the signal transduction pathways controlling actin cytoskeleton dynamics (*15*). In response to the injection of a foreign body, insect hosts enclose it in a cellular sheath of hemocytes in an encapsulation process that requires adherence, spreading, and attachment of immune cells. Like pathogenic bacteria, parasitoid wasps may inhibit the cytoskeleton dynamics of immune cells using viral PTPs and thus may prevent encapsulation of parasitoid eggs.

The second largest CcBV gene family (CcBV *ank*) comprises six genes encoding proteins with ankyrin repeat motifs. These proteins belong to the IκB family (*16*), whose members are inhibitors of nuclear factor κB (NF-κB)/Rel transcriptional factors, implicated in vertebrate and *Drosophila* immune responses (*17*). As reported recently for other PDVs, CcBV Ank proteins lack the regulatory elements associated with the basal degradation of IκB proteins. Normally, proteolysis of the inhibitor of nuclear factor κB (IκB) releases NF-κB/Rel, sequestered in the cytoplasm by IκB, to translocate to the nucleus and to initiate transcription of immune response genes (*17*). A similarly truncated IκB-like protein is used by a poxvirus (the African swine fever virus) to inhibit the vertebrate immune response (*18*). The truncated forms of the six CcBV Ank proteins may play the same role in lepidopteran hosts.

The third gene family encodes for four predicted cysteine-rich proteins (CcBV *crp*) containing a particular cysteine knot motif (*19*). A similar protein—teratocyte secreted protein 14 (TSP 14)—is encoded by a cellular gene of a braconid wasp species (*20*). The TSP14 protein is secreted by teratocytes (i.e., wasp cells circulating within the host's hemolymph) and, notably, inhibits storage protein synthesis. CcBV Crp proteins may also inhibit translation of storage proteins, such as arylphorin, the level of

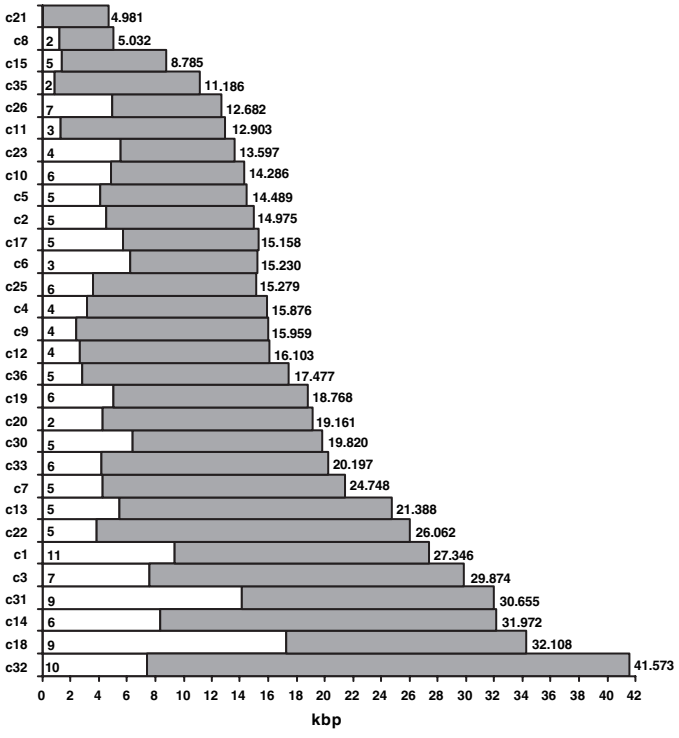

Fig. 1. Graphical representation of the gene distribution for each CcBV circle. Each circle is represented by a bar. Areas in white represent the length of the coding sequence, with the number of coding sequences indicated in black. Areas in gray represent noncoding sequences. The total length of each circle (bp) is indicated in black.

which is dramatically decreased in the hemolymph of parasitized *Manduca sexta* (*21*). Selective disruption of host protein translation is thought to redirect host metabolism to support endoparasite growth and development.

The fourth gene family encodes three cysteine protease inhibitors (CcBV *cyst*) of the cystatin superfamily. Cystatins have been described in a variety of organisms (*22*) but have apparently not previously been found in viruses (*23*). Interestingly, cystatins are also secreted by parasitic filarial nematodes and account for a major part of their immunosuppressive activity (*24*).

The products of the five other gene families do not contain any conserved domains that would allow prediction of their function (Fig. 2). Two genes are only known from *Cotesia congregata* bracovirus (CcBV *hypothetical1* and CcBV *hypothetical2* families), and the other three families contain genes described in viruses associated with other *Cotesia* species (*25*) (CcBV *EP1-like*, CcBV *family1*, CcBV *family2*). Most of these genes are expressed in the host tissues—the EP1 protein, for example, can account for 10% of the hemolymph protein content in parasitized hosts (*26*)—and presumably are required for successful parasitism.

The complex genome of CcBV devotes at least 26% of its CDS to potential virulence factors. Several genes probably originate from duplication events, resulting in multiple multigenic families consisting of up to 27 genes and constituting almost half the CDS. Such gene diversification may have facilitated the radiation of the bracovirus-bearing wasp complex, which now consists of 17,500 species (*9*). Strikingly, CcBV *ank* and CcBV *PTP* resemble truncated versions of cellular genes. Cysteine-knot motif genes have not only been described in PDV genomes, but also in the genome of a braconid wasp (*Microplitis croceipes*) (*20*). Finally, some of the CcBV genes, such as cystatin and histone H4 genes, have apparently not yet been described previously in viral genomes, which suggests that some of the PDV genes have been acquired from the cellular genome. Gene transfer may have occurred into the chromosomally integrated form of the virus, after recombination or transposition events.

Apart from the abundance of virulence factors, the CcBV genome lacks CDS with significant similarity to other virus genes. There are remnants of genes from retrovirus-like elements, but only three genes share significant similarities with sequences from free replicating viruses. Two putative proteins have a significant similarity with a baculovirus protein (48% similarity with *Autographa californica* M nuclear polyhedrosis virus gp94) nonessential for infectivity (*27*). A third protein shows significant similarity (39.9%) to a hypothetical protein from *Spodoptera frugiperda* ascovirus 1 (SfAV1), a member of a family of lepidopteran-infecting viruses (*28*).

Unexpectedly, the bracovirus genome does not contain any set of genes that offers a hallmark for a known virus family. The paucity of "virus-like" genes may be partly explained by the selection pressures acting on PDVs. The genes involved in the production of virus particles do not have to be present on the DNA injected into insect hosts, because virus particles' replication is restricted to wasp ovaries. The demonstration that the p44 gene encoding a structural protein of the *Campoletis sonorensis* ichnovirus is amplified in female wasps undergoing virus

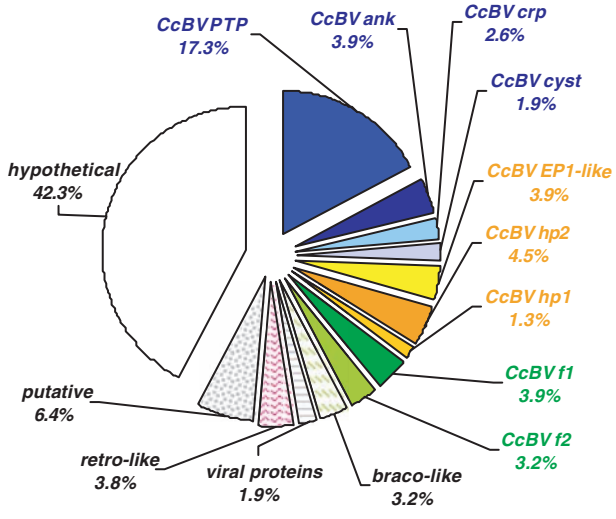

Fig. 2. Classification of the 156 genes identified in the CcBV genome: 42.3% of the genes encode proteins showing no similarity to proteins in databanks (in white); 42.5% of the genes are organized in nine multigenic families (indicated with different colors). In blue are shown genes encoding proteins with well-known conserved domains (PTPs, protein tyrosine phosphatases; ank, ankyrin; crp, cysteine-rich proteins; cyst, cystatins). In orange are shown gene families specific of CcBV (hp1 and hp2: hypothetical 1 and 2). In green are shown gene families common to other species of the *Cotesia* genus. Of the genes, 3.2% are single genes encoding proteins that are homologous to "bracovirus proteins" (hatched green); 1.9% (hatched gray) correspond to the three genes encoding proteins with viral structural domains and 3.8% to the genes that resemble retrovirus-like elements (hatched pink). In dotted-line gray are shown 6.4% of the genes encoding proteins that have similarity with proteins in hypothetical databanks.

Table 2. Features of the CcBV gene families. The features of each gene family are detailed with the circle (C) localization of each gene and the number of related genes on each circle. The average % of similarity of the related proteins are indicated for each gene family. Other PDVs containing such families are indicated. GiBV, *Glyptapanteles indiensis* bracovirus; CsIV, *Campoletis sonorensis* ichnovirus; MdBV, *Microplitis demolitor* bracovirus; HfIV, *Hyposoter fugitidus* ichnovirus; TnBV, *Toxoneuron nigriceps* bracovirus; CkBV, *Cotesia kariyai* bracovirus; CgBV, *Cotesia glomerata* bracovirus.

Parameter	CcBV families								
	PTP	ank	crp	cyst	EP1-like	hp1	hp2	f1	f2
Number of related genes	27	6	4	3	6	2	7	6	5
Circle no.: no. of related genes	C1:8	C11:1	C18:2	C19:3	C1:3	C30:1	C3:1	C9:2	C19:3
	C4:2	C14:2	C32:1		C5:1	C18:1	C6:1	C23:1	C25:1
	C7:1	C15:1	C35:1		C7:1		C9:1	C25:1	C30:1
	C10:5	C26:2			C8:1		C20:1	C33:2	
	C14:3						C23:1		
	C17:5						C25:1		
	C26:3						C33:1		
Percent similarity	<5	19.49	13.79	75	16.34	63.28	33	41.48	75.14
PDVs in which similar gene families are found	GiBV	CsIV	CsIV	None	CkBV	None	None	CkBV	CkBV
	TnBV	HfIV	CgBV					GiBV	
		TnBV	MdBV						
		MdBV							

replication, but is not encapsidated, lends support to this hypothesis (*29*).

The idea that all the genes involved in viral DNA replication and virion production have been transferred to the wasp genome is nevertheless difficult to sustain. A more parsimonious hypothesis would be that bracoviruses do not originate from any of the large genome viruses characterized to date (*30*). They may have been built up from a simple system producing circular DNA intermediates, such as mobile elements, within the wasp genome. The acquisition of a capsid protein, possibly of viral origin, around the circular DNA intermediates would have allowed infection of lepidopteran cells. Finally, virulence genes could have been acquired from the wasp genome at different times during evolution of bracovirus-bearing wasp lineages, thus explaining why CcBV genes encoding proteins with a predicted function resemble cellular genes.

From their genome content, bracoviruses can be discerned as biological weapons directed by the wasps against their hosts. The wasp strategy for delivery of bracovirus genes could inspire medical applications for gene therapy, whereas PDV virulence factors are of interest in agriculture. Currently, a parasitoid gene is already in use in pest-control studies: TSP 14–producing transgenic plants significantly reduce *Manduca sexta* larvae growth and development (*31*). Cystatins also have pesticide activity, because when expressed in transgenic plants, they can reduce the growth of nematodes (*32*). Other potential virulence factors encoded by PDV genomes may also serve as a source of natural molecules with insecticide activity of high specificity (*33*).

References and Notes

1. S. D. Dyall, M. T. Brown, P. T. Johnson, *Science* **304**, 253 (2004).
2. M. Turnbull, B. A. Webb, *Adv. Virus Res.* **58**, 203 (2002).
3. N. E. Beckage, *Parasitology* **116** (Suppl.), S57 (1998).
4. D. B. Stoltz, *J. Gen. Virol.* **71**, 1051 (1990).
5. E. Belle *et al.*, *J. Virol.* **76**, 5793 (2002).
6. J.-M. Drezen *et al.*, *J. Insect Physiol.* **49**, 407 (2003).
7. D. B. Stoltz, P. Krell, M. D. Summers, S. B. Vinson, *Intervirology* **21**, 1 (1984).
8. B. A. Webb *et al.*, in *Virus Taxonomy*, M. H. V. Van Regenmortel *et al.*, Eds. (Academic Press, San Diego, 2002), pp. 253–260.
9. J. B. Whitfield, *Naturwissenschaften* **84**, 502 (1997).
10. J. B. Whitfield, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 7508 (2002).
11. J. A. Kroemer, B. A. Webb, *Annu. Rev. Entomol.* **49**, 431 (2004).
12. J. N. Andersen *et al.*, *Mol. Cell. Biol.* **21**, 7117 (2001).
13. B. Provost *et al.*, *J. Virol.*, in press.
14. J. B. Whitfield, *Parasitol. Today* **6**, 381 (1990).
15. F. Deleuil, L. Mogemark, M. S. Francis, H. Wolf-Watz, M. Fallman, *Cell. Microbiol.* **5**, 53 (2003).
16. S. Ghosh, M. J. May, E. B. Kopp, *Annu. Rev. Immunol.* **16**, 225 (1998).
17. M. S. Dushay, B. Asling, D. Hultmark, *Proc. Natl. Acad. Sci. U.S.A.* **93**, 10343 (1996).
18. Y. Revilla *et al.*, *J. Biol. Chem.* **273**, 5405 (1998).
19. J. Einerwold, M. Jaseja, K. Hapner, B. A. Webb, V. Copie, *Biochemistry* **40**, 14404 (2001).
20. D. L. Dahlman *et al.*, *Insect Mol. Biol.* **12**, 527 (2003).
21. N. E. Beckage, M. R. Kanost, *Insect Biochem. Mol. Biol.* **23**, 643 (1993).
22. M. Abrahamson, M. Alvarez-Fernandez, C. M. Nathanson, *Biochem. Soc. Symp.* **70**, 179 (2003).
23. E. Espagne *et al.*, in preparation.
24. P. Schierack, R. Lucius, B. Sonnenburg, K. Schilling, S. Hartmann, *Infect. Immun.* **71**, 2422 (2003).
25. T. Teramato, T. Tanake, *J. Insect Physiol.* **49**, 463 (2003).
26. S. H. Harwood, A. J. Grosovsky, E. A. Cowles, J. W. Davis, N. E. Beckage, *Virology* **205**, 381 (1994).
27. R. J. Clem, M. Robson, L. K. Miller, *J. Virol.* **68**, 6759 (1994).
28. K. Staziak, M. V. Demattei, B. A. Federici, Y. Bigot, *J. Gen. Virol.* **81**, 3059 (2000).
29. L. Deng, D. B. Stoltz, B. A. Webb, *Virology* **269**, 440 (2000).
30. L. M. Iyer, L. Aravind, E. V. Koonin, *J. Virol.* **75**, 11720 (2001).
31. I. B. Maiti *et al.*, *Plant Biotechnol. J.* **1**, 209 (2003).
32. P. E. Urwin, M. J. McPherson, H. J. Atkinson, *Planta* **204**, 472 (1998).
33. N. E. Beckage, D. B. Gelman, *Annu. Rev. Entomol.* **49**, 299 (2004).

34. This work was supported by the European Community program "Bioinsecticides from Insect Parasitoids" (QLK3-CT-2001-01586). The authors thank A. Bézier and F. Hericourt for useful suggestions; C. Ménoret and J. Dérisson for insect rearing; and N. Beckage for early contribution to the project. Genome circle sequences have been deposited in the EMBL Nucleotide Sequence Database under accession numbers (to circle 1 to 36 respectively): AJ632304; AJ632305; AJ632306; AJ632307; AJ632308; AJ632309; AJ632310; AJ632311; AJ632312; AJ632313; AJ632314; AJ632315; AJ632316; AJ632317; AJ632318; AJ632319; AJ632320; AJ632321; AJ632322; AJ632323; AJ632324; AJ632325; AJ632326; AJ632327; AJ632328; AJ632329; AJ632330; AJ632331; AJ632332; AJ632333.

Supporting Online Material

www.sciencemag.org/cgi/content/full/306/5694/286/DC1

Materials and Methods

References and Notes

21 July 2004; accepted 26 August 2004