Research article

Molecular evidence for the evolution of ichnoviruses from ascoviruses by symbiogenesis

Yves Bigot*[1,2,3], Sylvie Samain[4], Corinne Augé-Gouillou[1,2] and Brian A Federici[5]

Address: [1]Université François Rabelais de Tours, GICC, UFR des Sciences & Techniques, Parc de Grandmont, 37200 Tours, FRANCE, [2]CNRS, UMR 6239, UFR des Sciences & Techniques, Parc de Grandmont, 37200 Tours, France, [3]CHRU de Tours, Bd Tonnelé, 37032 Tours, France, [4]GENOSCOPE, 2 rue Gaston Crémieux, CP 5706, 91057 Evry, France and [5]Department of Entomology & Graduate Programs in Genetics, Microbiology, and Molecular Biology, University of California, Riverside, CA92521, USA

Email: Yves Bigot* - bigot@univ-tours.fr; Sylvie Samain - samain@genoscope.cns.fr; Corinne Augé-Gouillou - auge@univ-tours.fr; Brian A Federici - brian.federici@ucr.edu

* Corresponding author

Abstract

Background: Female endoparasitic ichneumonid wasps inject virus-like particles into their caterpillar hosts to suppress immunity. These particles are classified as ichnovirus virions and resemble ascovirus virions, which are also transmitted by parasitic wasps and attack caterpillars. Ascoviruses replicate DNA and produce virions. Polydnavirus DNA consists of wasp DNA replicated by the wasp from its genome, which also directs particle synthesis. Structural similarities between ascovirus and ichnovirus particles and the biology of their transmission suggest that ichnoviruses evolved from ascoviruses, although molecular evidence for this hypothesis is lacking.

Results: Here we show that a family of unique pox-D5 NTPase proteins in the *Glypta fumiferanae* ichnovirus are related to three *Diadromus pulchellus* ascovirus proteins encoded by ORFs 90, 91 and 93. A new alignment technique also shows that two proteins from a related ichnovirus are orthologs of other ascovirus virion proteins.

Conclusion: Our results provide molecular evidence supporting the origin of ichnoviruses from ascoviruses by lateral transfer of ascoviral genes into ichneumonid wasp genomes, perhaps the first example of symbiogenesis between large DNA viruses and eukaryotic organisms. We also discuss the limits of this evidence through complementary studies, which revealed that passive lateral transfer of viral genes among polydnaviral, bacterial, and wasp genomes may have occurred repeatedly through an intimate coupling of both recombination and replication of viral genomes during evolution. The impact of passive lateral transfers on evolutionary relationships between polydnaviruses and viruses with large double-stranded genomes is considered in the context of the theory of symbiogenesis.

Background

Parasitic wasps belonging to the families Ichneumonidae and Braconidae are among the most successful higher eukaryotic organisms. These two large families contain more than 37,000 known species, and estimates of the total number of species are as high as 100,000 [1].

Approximately two-thirds of these wasps are endoparasites, meaning that the larval stages develop within the body cavity of their hosts, typically other insects. Among the most successful of these endoparasitic wasps are those that use lepidopteran larvae as hosts. Owing to the economic importance of these insects and the utility of their wasp parasites as biological control agents, the ability of these parasites to develop within lepidopteran hosts without triggering an intense immune response has been the subject of numerous studies over the past forty years.

Early studies of the Mediterranean flour moth, *Ephestia kuhniella*, parasitized by the ichnemonid, *Venturia canescens*, showed that eggs of this species are coated with particles that resemble virions [2-4] and contain surface proteins that mimic host proteins, thus keeping the eggs and larvae from being recognized as foreign material by their host. These particles lack DNA, and thus are not considered virions [5]. With respect to both species number and mechanisms that lead to successful parasitism, endoparasitic wasps are known to inject secretions at oviposition, but only a few lineages use viruses or virus-like particles (VLPs) to evade or to suppress host defences. In the family Ichneumonidae, for example, four types of host defence suppression mediated by the injection of fluids or suspensions are known that lead to successful parasitism.

1) Fluid injected with eggs bypasses host defences without the aid of viruses or VLPs [6].

2) Wasps inject a virus that replicates in both the wasp and lepidopteran host. One example is the wasp *Diadromus pulchellus*, which injects an ascovirus, DpAV4 [7] into host pupae to circumvent host defence response.

3) The wasp injects VLPs capable of molecular mimicry and/or direct defence suppression.

4) The wasp injects polydnavirus particles that contain genes coding for proteins that interfere with host defence responses.

The last mechanism is by far the best-studied type of direct immune suppression by ichneumonid wasps, and occurs in many species belonging to genera *Campoletis*, *Hyposoter* and *Tranosema* (Ichneumonidae, Campopleginae), and *Glypta* (Ichneumonidae, Banchinae) [8]. In these cases, female wasps inject eggs along with ichnovirus particles into their hosts. Similarly, in certain lineages of endoparasitic braconid wasps, other types of immunosuppressive particles containing DNA occur in the fluid injected along with eggs [[9]; for a review, [10]]. Once in the host, ichneumonid and brachonid particles enter host nuclei and their DNA is transcribed, producing proteins that selectively suppress various steps in the host defence response.

As a result of this unusual biology, these particles were described as symbiotic viruses belonging to new viral family, *Polydnaviridae* [10-12]

Since the 1970's, it was assumed that the DNA in the polydnavirus particles, as with all other viruses, encoded typical enzymes and proteins for viral replication and virion assembly and structure. However, several recent genomic studies have shown that only a small number of the genes vectored into lepidopteran hosts, less than 2%, have homologs in other viruses. Most viral DNA is non-coding, except that which codes for wasp proteins involved in suppression of immune pathways, such as phenoloxidase activation and the toll pathways [8,13,14]. Even before these genomic studies, it was suggested that these particles were more similar to organelles than viruses [15]. The similarities between particle structure and virions of known types of complex DNA insect viruses are striking, and suggest these immunosuppressive particles originated by symbiogenesis between viruses and endoparasitic wasps, the same evolutionary process by which mitochondria and plastids originated from symbiotic bacteria [16]. For example, most braconid wasps produce enveloped bacilliform particles classified as bracoviruses, and these resemble baculovirus and nudivirus virions [10,15]. Similarly, ichneumonid wasps produce enveloped spindle-shaped particles classified as ichnoviruses that resemble virions of ascoviruses, viruses lethal to lepidopterans, which, interestingly, are vectored by endoparasitic wasps [15]. It must also be noted that ichnoviruses resemble other true virus particles that are structurally very similar to virions of ascoviruses, but which remain unclassified because the lack of information about their genomes [17-21]. However, ascoviruses and ichnoviruses display very different genome properties; similar genomic differences occur between bracoviruses and baculoviruses or nudiviruses, suggesting that convergent evolution led to the origin the different polydnavirus types from at least two different types of viruses.

In ascoviruses, the genome consists of a single circular DNA molecule ranging from 119- to 180-kpb in size [7]. Phylogenetic analyses of several viral genes have revealed that ascoviruses are closely related to iridoviruses [22], and likely evolved from them. In contrast, the genome of ichnoviruses is composed of multiple circular DNA molecules (25 to 105) representing a total size of 250 to 300-kbp, all of which are replicated from the wasp chromosomes. The ichnovirus proviral genome is specifically excised and amplified in several segments in the female calyx cells, the only wasp tissue in which ichnovirus virogenesis occurs. After assembly, these particles are secreted into the female genital tract. Once injected into the host, the ichnovirus genome does not replicate, and does not lead to the production of a new virus generation. The third

characteristic of ichnoviruses is that most of the genes borne by the particles are not related to viral genes. Among the 7 annotated ichnovirus gene families, there are four (rep, PRRP, N, and TrV) for which no homology with known eukaryotic (or prokaryotic) proteins has been detected and for which no function has been proposed. Among the remaining three (cys, ank and inx), cys-motif proteins have no clear homologs among eukaryotic (or prokaryotic) proteins, although the "cysteine knot" that they form is a folding domain found in many proteins, but not one that is necessarily related to eukaryotic host immune systems [10,14]. However, some protein domains and their putative functions suggest that they might be related to regulatory components of eukaryotic host defence systems that are not sufficiently elucidated.

Although the resemblance of the polydnavirus virions to those of conventional insect viruses suggests that the former evolved from the latter, to date no molecular evidence supports this hypothesis. In the case of ascoviruses and ichnoviruses, well-conserved genes found among the three ascoviruses sequenced so far (SfAV1a [23], TnAV2c [24], and HvAV3e [25]) are not found in ichnovirus genomes. As noted above, the principal reason for this is that the genomes of the latter viruses appear to contain mainly wasp genes, not viral genes. This highlights the need for new and alternative types of sequence data obtained from pertinent biological systems. In this regard, DpAV4 has features that could provide important insights. Indeed, it is the only ascovirus known to replicate in both its wasp and caterpillar hosts. It is transmitted vertically from wasp to caterpillars to suppress the defence response of the latter host, thereby enabling parasite development [26,27]. Moreover, in males and females of *D. pulchellus*, the DpAV4 genome resides in the nuclei of all hosts cells, providing a possible example of what may have been an intermediate stage in the symbiogenesis that led to the evolutionary origin of ichnoviruses.

We recently sequenced the DpAV4 genome, and a combination of our analysis of this genome and recent data from new types of ichnoviruses, as well as new software programs that elucidate protein relationships based on structural analysis, have enabled us to detect phylogenetic relationships between proteins encoded by open reading frames of DpAV4 and the *Glypta fumiferanae* (GfIV) and *Campolitis sonorensis* (CsIV) ichnoviruses. In support of the symbiogenesis hypothesis for the origin of ichnoviruses, data and analyses suggest two independent symbiogenic events, in agreement with what was previously proposed [28]. The first led to the ichnoviruses in Banchinae lineage. This hypothesis is based on the occurrence of a gene cluster present in GfIV and DpAV4. The second symbiogenic event led to ichnoviruses in the Campopleginae wasp lineage. This hypothesis is based on relationships of the major capsid proteins among CsIV, ascoviruses and iridoviruses. Extending our investigations to proteins encoded by open reading frames of certain ascoviruses and bracoviruses, hosts and bacteria, in the light of recent analyses about the involvement of the replication machinery of virus groups related to ascoviruses in lateral gene transfer [29], we discuss the robustness and the limits of the molecular evidence supporting an ascovirus origin for ichnovirus lineages.

Results and discussion
DpAV4 and ichnovirus related proteins
The DpAV4 genome sequenced by Genoscope (France) is 119,334-bp in length. Its organization, gene content and evolutionary characteristics will be detailed in a separate publication (manuscript in preparation; Additional file 1). However, BLAST results obtained with several ORFs in the DpAV4 genome provide evidence that certain ichnovirus ORFs have their closest relatives in an ascovirus genome. Specifically, we identified a 13-kbp region that contains a cluster of three genes (Fig. 1, ORF90, 91 and 93; Additional files 1 and 2) that have close homologs in a GfIV gene family composed of seven members [28]. All contain a domain similar to a conserved domain found in the pox-D5 family of NTPases. To date, this pox-D5 domain has been identified as a NTP binding domain of about 250 amino acid residues found only in viral proteins encoded by poxvirus, iridovirus, ascovirus and mimivirus genomes. These genes seem to be specific to GfIV, as they are absent in the three sequenced genomes of other ichnoviruses, namely CsIV, *Tranosema rostrales* ichnovirus (TrIV), and *Hyposoter fugitivus* ichnovirus (HfIV).

More specifically, in DpAV4, ORF90 encodes a protein of 925 amino acid residues that is 40% similar from position 140 to 925 to a protein of 972 amino acid residues encoded by the ORF1 contained in the segment C20 in the GfIV genome (Fig. 2). These two proteins can therefore be considered putative orthologs. The 480 C-terminal residues of this DpAV4 protein are also 42% similar to the C-terminal domain of the protein homologs encoded by the ORF1 of the D1 and D4 GfIV segments, 36% similar to the N-terminal and the C-terminal domains of the protein encoded by the ORFs 184R and 128L of the iridovirus CIV and LCDV, and 30% similar with those encoded by ORFs 119, 99 and 78 in the ascovirus genomes of HvAV3e, SfAV1a and TnAV2c, respectively. Overall, this indicates that this DpAV4 protein is more closely related to that of GfIV than to those found in other ascovirus and iridovirus genomes currently available in databases. ORF091 encodes a protein of 161 amino acid residues similar only with the C-terminal domain of three proteins encoded by the ORFs 1, 1 and 3, contained, respectively, in GfIV segments D1, D4 and D3. In contrast, ORF93 is closer to iridovirus and ascovirus genes than to GfIV genes. This
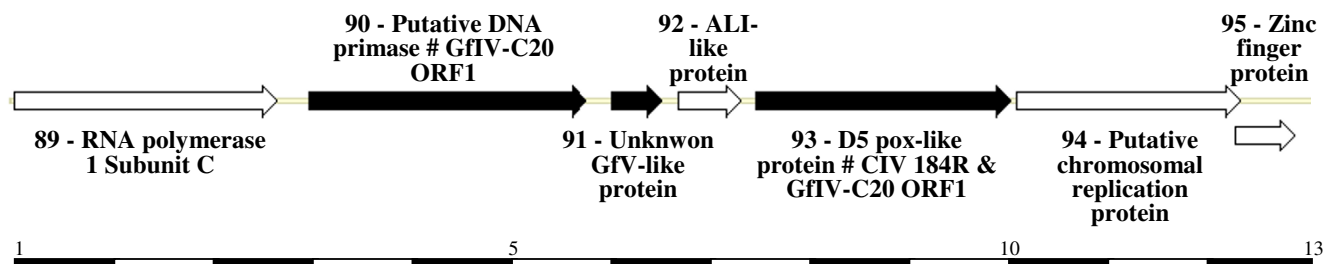
Figure 1
Map of the 13-kbp region of the DpAV4 genome (EMBL Acc. N°CU469068**and**CU467486**) that contains the gene cluster with direct homologs in the genome of the *Glypta fumiferanae* ichnovirus.** DpAV-4 ORF with well-characterized direct homologs among other ascovirus and iridovirus genomes are represented by white arrows. Homologous ORF of the GfIV genes are represented by black arrows. The sequence and the features of the protein encoded by these ORF are detailed in fig. S1. Below, the graph is scaled in kbp.

protein of 849 amino acid residues is 43% similar over all its length to CIV ORF184R orthologs in all iridoviral and ascoviral genomes and is only 36% similar over 350 amino acid residues to the C-terminal domain of the GfIV protein homologs encoded by the ORF1, 2, 1, 1, 1 and 1 in, respectively, the C20, C21, D1, D2, D3 and D4 segments of this virus.

Analysis of the genes surrounding the DpAV4 ORF-90-91-93 cluster confirms that this virus has an ascovirus origin since this region contains ORFs that are close homologs of genes in iridovirus and ascovirus genomes. Upstream from the ORF-90-91-93 cluster, an ORF encoding the DNA-dependent RNA polymerase 1 subunit C is present, which is an ortholog of the iridoviral CIV ORF176R and the ascoviral SfAV1a ORF008. Downstream from this cluster, there are two genes, absent in known ascoviral genomes, but similar to the iridoviral CIV ORF115L and CIV ORF132L. These two genes encode, respectively, a chromosomal replication initiation protein and zinc finger protein. In between them, a gene encoding a small protein is present that is similar to that encoded by the ORF069L of the iridovirus CIV, and which corresponds to the ALI-like protein also found in entomopoxviruses [30].

Since the three DpAV4 genes have relatives in all ascovirus and iridovirus genomes sequenced so far, their presence in the DpAV4 genome cannot result from a lateral transfer that occurred from an ichnovirus genome related GfIV to DpAV4. Thus, as these DpAV4 genes are the closest relatives of the pox-D5 gene family present in GfIV identified so far, they could be considered a landmark of the symbiogenic ascovirus origin of the ichnovirus lineage to which this polydnavirus belongs. An alternative explanation is that the presence of DpAV4-like genes in the genome of GfIV resulted from a lateral transfer from viral genomes closely related to those of GfIV and DpAV4.

Indeed, this might have happened when a *Glypta* wasp was infected by an ancestral virus related to DpAV4. Nevertheless, the symbiogenic origin of GfIV from ascoviruses is also supported by morphological features of its virions [28], which, aside from similarities in shape, also show reticulations on their surface in negatively stained preparations, a characteristic of the virions of all ascovirus species examined to date [7].

Relationships between ascovirus and ichnovirus virion proteins
Because ascovirus virions and ichnovirus particles display structural similarities, we developed an approach to search for homologs of virion structural proteins in ichnoviruses. These approaches were initiated in 2000 and recently finalized, but some of the conclusions have been published [14]. To date, only two virion proteins from the *Campoletis sonorensis* ichnovirus (CsIV) have been characterized [31,32]. The first is the P44 (Acc N° AAD01199), a structural protein that appears to be located as a layer between the out envelope and nucleocapsid, and the second, P12, a capsid protein (Acc N° AF004367). Presently, there are more than one hundred ascoviral or iridoviral MCP sequences in databases. BLAST searches using these sequences failed to detect any similarities between CsIV virion proteins and ascoviral or iridoviral MCPs, or any other proteins [33]. To evaluate the possibility that homology between ichnovirus and ascovirus virion proteins may simply not be detectable by conventional Blastp searches, we used a different method, WAPAM (weighted automata pattern matching; [34]). The models were designed on the basis of a previous study [22] demonstrating that MCP encoded by ascovirus, iridovirus, phycodnavirus and asfarvirus genomes are related, and all contain 7 conserved domains separated by hinges of very variable size. We investigated these conserved domains further using hydrophobic cluster analysis (HCA, [35]). This

```
>ref|YP_001029423.1| Gene info GfV-C20-ORF1 [Glypta fumiferanae ichnovirus]
Length=972

 Score =  432 bits (1112),  Expect = 5e-119, Method: Composition-based stats.
 Identities = 192/813 (23%), Positives = 326/813 (40%), Gaps = 74/813 (9%)

Query  140  VCPPTPGYAWLHNSQYVSPPPMPRELLDLLFPEPADVDRARALCTTFRNVDKATLAYVVE  199
            V P     + +  ++  S PP          +    +      +     K +L  +V
Sbjct  198  VQVPETSESLIDINESNSEPP----------SKKMKTNENPGNISYN----KKSLERLVM  243

Query  200  GIPARRADVRDEWMRVVWAVADTAAKNGYDALDVADAFSRRSHKYAGPEDVEAMYRQSNG  259
            + +R ++  D W V+WA+      +NGYD LD+AD FS+RS KY   EDV+   ++ S+G
Sbjct  244  DLESRHSEKYDLWRNVLWALNTCGLENGYDTLDIADKFSKRSAKYKNIEDVKKTFQSSDG  303

Query  260  AVTFGTLVHFSE---RGLPRDEGSVT----LKNTDLLTSDLKAFDDKVERIDDFEVNGNV  312
            +    L+ S+   GL  E  ++         L  + +   D  R+       + +
Sbjct  304  TMALQYLIDISKIYEEGLHDMEEYLSTIVPPIYLPLTLTQMDLRDRSKLRLTVKCSDASY  363

Query  313  INLVCGDLKIRIDINNLAIYVNGQYRGHLKTGLEFEND-LSFIHKHFSEGGGSIFHMNEN  371
            + +   L + D N + +         LT+    + L +    S+G   F +
Sbjct  364  VLFIIPHLGVLTDHNGIRL-------NFLTTKIHSSSSTLRSLCAEMSDG----FTVTFQ  412

Query  372  RTDLVSRGQYGENRIKFMRN----GKQKHYILMENGRNVGTRSKKADINHLEQLIVDTNQ  427
             D V R +   + ++F+ N      G+          + R   T   +   L +
Sbjct  413  EDDYVFRSR--RHSLEFILNNPLDGENASITFPNHKRGRKTCDLQRIHKALRKETMKLLS  470

Query  428  RQMAGYNILNQNNVNVFNITNNFYFDEDNLRSDDQIIDALVRSHPSIVDHFKFCSRDYGS  487
             +          + N +  FN  N+      D  R D +++  AL   + P+I  + +
Sbjct  471  IKSGKIFAKDDNCIVYFNNKNS-----DATRDDIKLVQALYAAKPNIKTRYCYS------  519

Query  488  FNGIYLFGPHTGIWRKEHNATVNKMLTTRIKKYVP-MLTEKETKYCETQKNKAGLRMAFI  546
             N I    P T +W    +        + L T + +      LT+ E K+  T+   A +R  +
Sbjct  520  GNTISCCDPKTNVWEIIGEKKFDFHLRTYLTEIAEIKLTKAELKHMNTKNAIANIRNIIL  579

Query  547  EHIVDNDFESKIDENLDVFATANCVYDVGE--GALRKARPHDFVYTNTGWDYERAADEH  604
             +     +F ++++  L +F T N   D        +R    D   T  GW Y+   A ++
Sbjct  580  GGVEITNFANRLNSKLHLFITDNKTIDTSSIPPIIRHIEIEDLAKTTCGWSYNSELAAKY  639

Query  605  LPAVREFFERVFPVAQERDVVVTYLASLIHGYRTDKKLLAFTDKRNGNNGKTTLMTLLRT  664
             V+ +F  +FP   ERD  +++A  ++G R D+   L  TD+  G  GK TL  LL+
Sbjct  640  KEPVQSYFNTLFPEQSERDWFLSFIARFLNGRRMDEPFLILTDEHEGKTGKKTLAKLLKA  699

Query  665  FFGDYTKTNNNFFLKGAFAKDKDAHDGGTATLKGKRLLICDELKKSMRLDEGTVKNITGG  724
            FG Y +N+     G   ++ +  GG        KRLL+ D L+K+   L+ G +K ITG
Sbjct  700  VFGSYYLSNSTIVRAGGTRENNE-FAGGLYGFNEKRLLLADGLQKTDTLNCGFIKAITGD  758

Query  725  ANQELQGRRMGKEDVFRLTVQCGVILIFNEGDCPKFDATDNAFMERLVIVPFRSKFVGGE  784
             N  ++               F   VQ G+++I NE + PKFD TD  F++++V+ P RS+FV  E
Sbjct  759  CNYYMKHINKFFRTQFEFVVQAGLVVIANEKNMPKFDKTDENFLKKMVVCPLRSRFVTQE  818

Query  785  ADPDTYTFQRDC------NISDKF--KLWRSALLKHFISHCRKNGIADMEMPSSMVDWKN  836
               +     +        N S K       WRSA L    +            +P SM  WK
Sbjct  819  EFQEMRRSKNKMTNILIANDSSKLPFHEWRSATLDLLTEY---YNKPISTIPDSMKAWKK  875

Query  837  EILEENNVVAEW--LWDAVRPQEGSFVQLAELRERYKKDHPHERAFKNKDIER----MIN  890
            + +    +W  L        E  FV  E+ R KK    E+    ++++R     M
Sbjct  876  RVCDMFFDYTDWMNLHLCKSQNENEFVSATEIMIRIKKP---EKYMDPREMQRLQIAMKE  932

Query  891  SAFNAKGAYVKDHHNYYIDGARKQGRRVYLNFA  923
            A   +  Y   H +       R + R V +N
Sbjct  933  WAIKNECKYKSRHMHPDDKRKRVETRSVIMNAT  965
```

Figure 2
Amino acid sequence comparison resulting from a BLAST search done with the DpAV4 ORF90 as a query, and the best hit corresponding to the protein encoded by the ORF1 of the ichnovirus segment GfV-C20 (Subject; Genbank Acc. N° YP_001029423).

analysis revealed that most conservation occurred at the level of hydrophobic residues, as expected for structural proteins (Additional file 3a and 3b). The size variability of the hinges between conserved domains and the conservation of hydrophobic residues might explain why BLAST searches using iridoviral and ascoviral MCP sequences have limited ability to detect MCP orthologs in phycodnavirus and asfarvirus genomes. We designed two syntactic models (see Materials and Methods), which together were able to specifically align all MCP sequences of the four

virus families. Importantly, WAPAM aligned the CsIV ichnovirus P44 structural protein with both models. Complementary structural and HCA confirmed the presence of the seven conserved domains in this CsIV structural protein (Fig. 3a and Additional file 3c).

In addition to the above analysis, ten syntactic models were developed using proteins conserved in the three sequenced ascovirus species (SfAV1a, TnAV2c, and HvAV3a) and twelve iridoviruses [36]. None of these

Figure 3
Sequence (lanes 1 to 3) and secondary structure (lanes 4 to 6) comparisons among (a) MCP and (b) SfAV1a ORF061 orthologs from CsIV (lanes 1 and 4, typed in black), DpAV4 (lanes 2 and 5, typed in blue) and SfAV1a (lanes 3 and 6, typed in purple). Conserved positions among the amino acid sequence of CsIV and those of DpAV4 and SfAV1a are highlighted in grey. Secondary structures in the three SfAV1a ORF061 orthologs were calculated with the Network Protein Sequence Analysis at http://npsa-pbil.ibcp.fr/ and the statistical relevance of the secondary structures were evaluated with Psipred at http://bioinf.cs.ucl.ac.uk/psipred/. C, E and H in lanes 4 to 6 respectively indicated for each amino acid that it is involved in a coiled, b sheet or a helix structure. Using default parameters of Psipred, upper case letters indicate that the predicted secondary structure is statically significant in Psipred results. Significant secondary structures are highlighted in yellow. In (a), the comparisons were limited to three of the seven conserved domains (Additional file 3a, b and 3c), the 2, 5 and 7. Indeed, classical *in silico* methods appeared to be inappropriate to predict statistically significant secondary structures in conserved structural protein rich in b strand such as iridovirus and ascovirus MCP. In contrast, a complete and coherent domain comparison was obtained by HCA profiles (fig. S3b, c).

models detected homologs among ichnovirus proteins available in databases, except for one (see Materials and Methods section), developed from small proteins encoded by the DpAV4 ORF041, SfAV1a ORF061, HvAV3a ORF74, and TnAV2c ORF118 in the ascovirus genomes, and iridovirus CIV ORF347L and mimivirus MIV ORF096R genomes, respectively. Importantly, these proteins have orthologs in vertebrate iridoviruses, phycodnaviruses, and asfarvirus. In SfAV1a, the peptide encoded by ORF061 is one of the virion components. In ascoviruses, iridoviruses, phycodnaviruses, and the asfarvirus, they have been annotated as thioredoxines, proteins that play a role in initiating viral infection [37-39]. Database mining with our model revealed four hits with CsIV sequences (Acc N°. M80623, S47226, AF236017, AF362508) each a homolog ORF of SfAV1a ORF061. In fact, these sequences correspond to several variants of a single region contained in the B segment of the CsIV genome. To date, these have not been annotated in the final CsIV genome, probably because they overlap a recombination site. HCA analyses confirmed that the hydrophobic cores were conserved (Fig. 3b and Additional file 3d and 3e).

The chromosomal locations of genes encoding these two CsIV proteins, i.e., P44 and P12, were also consistent with the symbiogenesis hypothesis. In fact, the ORF encoding P44 is not found in proviral DNA. It is notable that no ORFs encoding orthologs of P44 or other structural proteins such as MCPs are found in any of the other three ichnovirus genomes sequenced – TrIV, GfIV, HfIV [8,14]. Therefore, this indicates that the orthologs of ichnovirus MCPs and other virion structural proteins are also probably located in the genomes of these wasps, i.e., not in proviral DNA. In contrast to this, we found that the gene encoding the CsIV ortholog of SfAV1a ORF061 is located within the proviral DNA. Whether ortholog proteins are similarly involved in the TrIV, GfIV and HfIV biology, their genes are not found in proviral DNA, since no matches were detected in their viral genomes. The phylogenetic analysis performed previously on P44 and the SfAV1a ORF061 orthologs [15] indicated that they have an ancestor close to that of the ascoviruses and iridoviruses.

As in the case of genes encoding pox-D5 family of NTPases in all ascoviruses, iridoviruses, and GfIV, genes encoding virion proteins cannot result from a horizontal transfer from a Campoplegine or Banchine ichnovirus genome to all ascovirus, iridovirus, phycodnaviruses and asfarvirus genomes. As the ascovirus genes encoding the two virion proteins investigated here are the closest relatives of virion proteins in CsIV, they can be considered a landmark reflecting the symbiogenic origin of the two ichnovirus lineages from ascoviruses closely related to

DpAV4. In fact, the difficulty encountered in elucidating their sequence relationships can be explained by a combination of the marked transition from ascovirus to ichnovirus, and the significant selection constraints that resulted as the latter virus type evolved from the former.

Impact of the intimate environment on lateral gene transfer

Analysis of available ascovirus, iridovirus and ichnovirus genomes provides some of the first molecular support for the hypothesis that ichnoviruses evolved from ascoviruses by symbiogenesis. However, examining genes shared only by ascovirus, iridovirus and ichnovirus genomes likely limits the sources of genes that contributed to the evolution and complexity of these viruses, especially of the role of lateral gene transfer. Relevant to this is the recent finding that an important part of the mimivirus and phycodnavirus genomes had a bacterial origin [28]. Obviously, this did not lead to the conclusion that these viruses had a bacterial origin. The cytoplasmic environment in which these viruses replicate is rich in bacterial DNA because their amobae and unicellular algae hosts feed on bacteria that they digest in their cytoplasm. Thus, it has been proposed [28] that lateral transfers of bacterial DNA within these viral genomes were driven by intimate coupling of recombination and viral genome replication. Indeed, replication of these viruses is similar to that of bacteriophage T4. This mode of replication has been called recombination-primed replication. It permits integration of DNA molecules with sequence homology as short as 12-bp [28,40]. The replication machinery used by ascoviruses, iridoviruses, mimiviruses, phycodnaviruses, and other nucleocytoplasmic large DNA viruses (NCLDV) [41,42] is common to all of them, despite differences in the specifics of replication in each virus family. It can therefore be expected that recombination-primed replication occurred repeatedly during evolution of both these viruses and the genome of their eukaryotic hosts. In an eukaryotic cellular environment in which bacteria, chromosomes, NCLDV viruses and non-NCLDVs (such as baculoviruses) intimately cohabit temporarily or permanently, recombination-primed replication is able to allow reciprocal passive lateral transfers between viral genomes, host chromosomes, and bacterial DNA. Under these conditions, lateral transfers are considered passive since they just result from the intimate environment and not from an active mechanism dedicated to genetic exchanges. In ascoviruses and iridoviruses, the occurrence of such lateral transfers is supported by BLASTp searches that detected the presence of ORFs whose closest relatives have their origin within eukaryotic genomes (e.g., for DpAV4, in Additional data 1, ORFs 029, 049, 077, 080, 083, 118), bacterial genomes (e.g., for DpAV4, in Additional data 1, ORFs 056, 057, 059, 112, 115 and119) or viruses belonging to other

NCLDV and non-NCLDV families (e.g., for DpAV4, in Additional data 1, ORFs 007, 037, 062, 068).

The transmission of ascoviruses is unusual in that they are poorly infectious *per os* and appear to be transmitted among lepidopteran hosts by parasite wasp vectors at oviposition [7,43]. The genome of the ascoviruses can be replicated in presence of polydnavirus DNA either within the reproductive tissues of female wasps or within the body of the parasitized hosts infected by both polydnavirus and ascovirus. Consequently, integrated sequences of ascovirus origin can be expected within wasp and polydnavirus genomes. Reciprocally, sequences of polydnavirus origin may have been integrated in ascovirus genomes, whatever the wasp origin, ichneumonid or braconid. One gene family related to a bacterial family of N-acetyl-L-glutamate 5-phosphotransferase (Acc. N° of the closest bacterial relatives YP_001354925, CAM32558, ZP_00944224, ZP_02006449), identified only within the SfAV1a, HvAV3e and TnAV2c genomes, supports this conclusion. It has been found in the genome of a bracovirus, *Cotesia congregata* BracoVirus (CcBV [13]; Fig. 4). Since this gene is absent in the genome of *Microplitis demolitor* BV, a related bracovirus [8], it is difficult to infer the direction of the lateral transfer between the common ancestors of the three ascoviruses and of the wasp *C. congregata*. However, they unambiguously indicate that there was at least one lateral transfer for this gene between the common ancestor of ascoviruses and the parasitic wasp.

Since iridoviruses, like ascoviruses and other virus species [44,45], are, in some cases, vectored by parasitic wasps, databases were mined using all the available ichnovirus virus proteins as queries. We found no significant relationships between CsIV, HfIV and TrIV genomes and genomes of their putative closest relatives NCLDV and non-NCLDV relatives. This indicates that passive lateral gene transfers from virus to eukaryotes that are successfully spread and maintained in ichnovirus genomes remain rare events. One case of such lateral transfer was described in the CcBV genome. In this genome, aside from the presence of cardinal endogenous eukaryotic retrotransposon and Polintons that transposed in the chromosomal DNA of the proviral form of CcBV [46-48], two genes encoding AcMNPV P94-related proteins, which have their closest relatives among granuloviruses (XcGV), were found. This suggests that CcBV contained at least two cases of lateral transfers between non-NCLDV and a bracovirus.

Concluding remarks

Our results provide another source of evidence that passive lateral gene transfers have occurred regularly during evolution from bacteria to viruses and eukaryotes, and between viruses and eukaryotes [49-52]. Even if the pox-D5 NTPase genes in the GfIV genome, and the MCP and SfAV061-like genes in the CsIV genome, indicate that they have an ascovirus origin, they provide only limited evidence supporting an ascovirus origin of ichnoviruses. Indeed, their sequence conservation and biological characteristics suggest that there were repeated lateral transfers during evolution between ascoviruses and wasp genomes, including the proviral ichnovirus loci. This raises an important issue about the role of lateral transfers during co-evolution of the NCLDVs and non-NCLDVs, ichnovirus, wasp and parasitized host. Indeed, genetic materials of various origins have been exchanged and maintained during co-evolution. This therefore suggests that ichnoviruses might be chimeric entities partly resulting from sev-
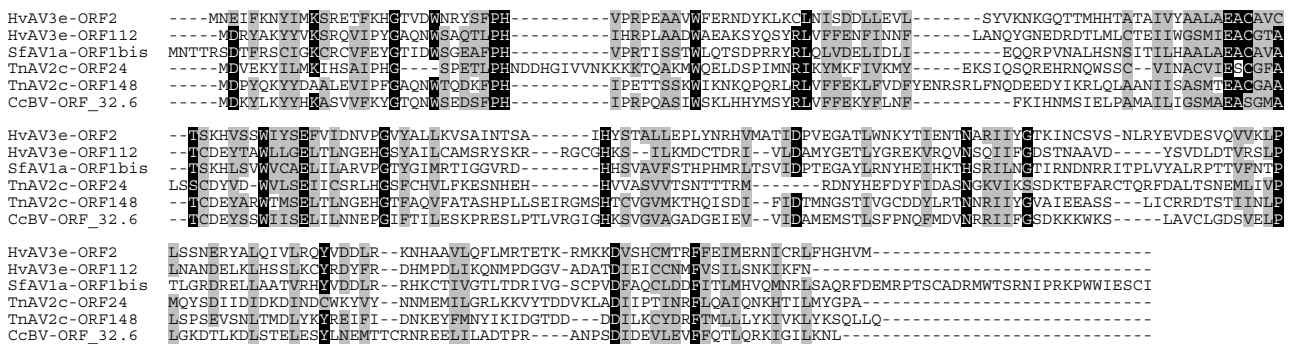

Figure 4
Amino acid sequence analysis of the N-acetyl-L-glutamate 5-phosphotransferase-like proteins encoded by one gene in SfAV1a (ORF1bis), two genes in HvAV3e (ORFs 002 and 112) and TnAV2c (ORFs 024 and 148), and one gene in CcBV. Identical residues between sequences are highlighted in black typed in white, those that are similar are highlighted in grey.

eral lateral transfer events of DNA fragments with viral origins.

Symbiogenesis was first proposed as an evolutionary mechanism when it became widely recognized that mitochondria and plastids originated from free-living prokaryotes [7]. The genomes of the endosymbiotic cyanobacteria and proteobacteria, respectively, at the origin of chloroplasts and mirochondria have evolved by reduction of several orders of magnitude to the approximate size of plasmids. Concurrently, nuclear genomes have been the recipients of plastid genomes. This relocation of the genes encoding most proteins of the endosymbiotic bacteria to the host nucleus is the ultimate step of this evolutionary process, so-called endosymbiogenesis [7,53]. Recent studies of plants have revealed a constant deluge of DNA from organelles to the nucleus since the origin of organelles [54]. This allows the host cell to have the genetic control on its organelles, in a relationship that is closer to enslavement or domestication than to a symbiosis or a mutualism in which the organelles would recover benefits from their contribution to the eukaryotic cell well-being. To date, this deluge of DNA is considered to correspond to passive lateral transfers that result from the interactions between the life cycle of the organelle and nuclear replication.

Numerous cases of symbiogenesis between endocellular bacteria and a wide variety of eukaryotic hosts have been characterized. However, recent work has demonstrated that this evolutionary process was not restricted to bacteria. It also occurred between endocellular eukaryotes such as unicellular algae and fungal endophyte in plants [55,56]. Endosymbiogenesis was also proposed as the evolutionary mechanism that allowed some invertebrate viruses with a large double-stranded DNA genome related to the nudiviruses and the ascoviruses [22], to have led, respectively, to the origin of bracoviruses and ichnoviruses, which are currently recognized as forming two genera within the family Polydnaviridae. Although presently there is no definitive evidence ruling out the hypothesis that the resemblance between ichnovirus and ascovirus virions is only an evolutionary convergence, the genomic differences between ascovirus and ichnoviruses are in good agreement with the symbiogenetic hypothesis. Indeed, they match an evolutionary scenario of endosymbiogenesis during which, from a single integration event of symbiotic virus genome, viral genes were lost and/or translocated from the provirus to other chromosomal regions (Fig. 5). In parallel, host genes of interest for the wasp parasitoid were integrated and diversified by selection and gene duplication in the proviral DNA. In this scenario, the more ancient symbiogenesis, the rarer the traces of genes from viral origin in the ichnovirus genome would be. This constitutes a constraint that dramatically limits

the possibility to investigate the evolutionary links between ascovirus and ichnovirus. Results of our analyses demonstrate that the situation is also complicated by the fact that lateral gene transfers unrelated to the origin of ichnoviruses cause important misleading background noise. Moreover, the scenario in Figure 5 is close to a previously proposed version [57], but is not consistent with results presented here, nor with recently accumulated knowledge on DNA transfer from organelles into the nucleus. Since endocellular environments favour lateral transfers between virus and wasp nucleus, it can be proposed that genes of virus origin that are involved in the ichnovirus biology were passively integrated in one or several loci, step by step over time, alone or through transfers of gene clusters, or even the entire viral genome. Since parasitoid wasps are able to vector different viruses [44,45], this second scenario opens the exciting possibility that virus genes involved in the ichnovirus biology might correspond to a gene patchwork resulting from transfers from viruses belonging to different NCLDV and non-NCLVD families. Because of the background noise due to lateral gene transfers found in these systems, elucidating the origins of ichnoviruses will be very time-consuming, requiring new accurate experimental approaches to generate more robust evidence. Sequencing wasp genomes to identify proteins of viral origin that are components of virions and involved in the assembly of these may well contribute to our understanding of how ichnoviruses and bracoviruses evolved from other insect DNA viruses.

Methods
Database mining
Searches for similarities were mainly developed using facilities of BLAST programs at two websites http://www.ncbi.nlm.nih.gov/blast/Blast.cgi and http:genoweb.univ-rennes1.fr/Serveur-GPO/outils.php3?id_rubrique=47. For DpAV4 genes having their origin within eukaryotic, bacterial or virus genomes belonging to NCLDV and non-NCLDV families, the closest gene was located using the distance trees supplied with each BLAST search at the NCBI website.

Defining syntactic model
Construction of syntactic models: Conserved amino acid blocks and positions described previously [15,22] and with new data sets were verified or determined using MEME at http://meme.sdsc.edu/meme/meme.html. In the first step, we used motifs resulting from MEME to make MAST minings in databases at http://meme.sdsc.edu/meme/mast.html. Since MEME motifs depend significantly on the data set use to calculate them, this approach did not enable an exhaustive detection of homologs among ascoviruses, iridoviruses, phycodnaviruses, mimiviruses and asfarviruses, and the detection sensitivity was ultimately very similar to that obtained
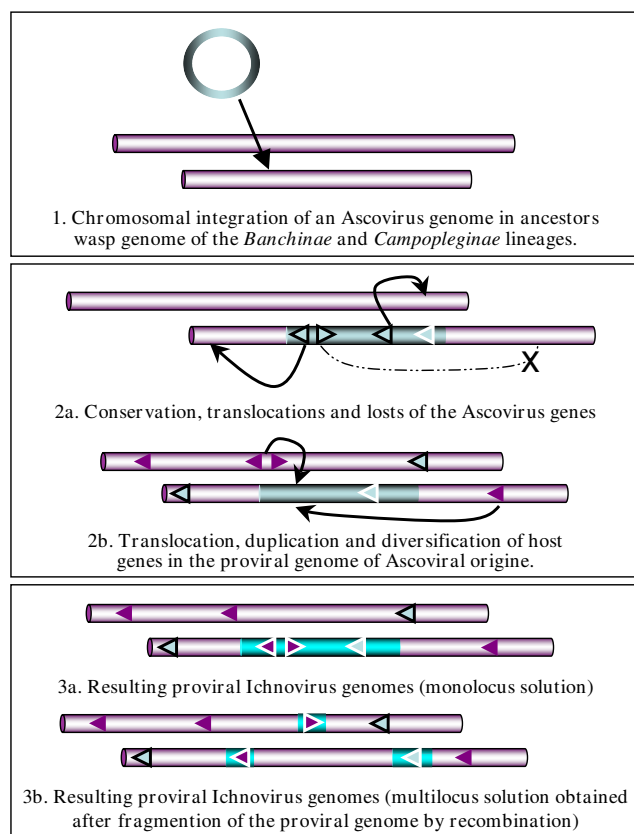
1. Chromosomal integration of an Ascovirus genome in ancestors wasp genome of the *Banchinae* and *Campopleginae* lineages.

2a. Conservation, translocations and losts of the Ascovirus genes

2b. Translocation, duplication and diversification of host genes in the proviral genome of Ascoviral origine.

3a. Resulting proviral Ichnovirus genomes (monolocus solution)

3b. Resulting proviral Ichnovirus genomes (multilocus solution obtained after fragmention of the proviral genome by recombination)

Figure 5
Hypothetical mechanism for the integration and evolution of ascovirus genomes in endoparasitic wasps.
Schematic representation of the three-step process of symbiogenesis, and DNA rearrangements that putatively occurred in the germ line of the wasp ancestors in the *Banchinae* and *Campopleginae* lineages, from the integration of an ascoviral genome to the proviral ichnoviral genome. Sequences that originate from the ascovirus are in blue, those of the wasp host and its chromosomes are in pink. Genes of ascoviral origin are surrounded by a thin black or white line, depending on their final chromosomal location. Two solutions can account for the final chromosomal organisation of the proviral ichnovirus genome, monolocus or multilocus, since this question is not fully understood in either wasp lineage. More complex alternatives to this three-step process might also be proposed and would involve, for example, the complete *de novo* creation of a mono or multi locus proviral genome from the recruitment by recombination or transposition of ascoviral and host genes located elsewhere in the wasp chromosomes. This model for the chromosomal organization of proviral DNA in polydnaviruses is consistent with data recently published [58].

with BLAST. To reach our detection objectives, we therefore constructed syntactic models that only included the most conserved positions and their variable spacing using WAPAM at the website. http://genoweb.univ-rennes1.fr/ Serveur-GPO/ outils_acces.php3?id_syndic=185&lang=en. Defining these models was obtained empirically until they allowed an exhaustive detection in refseq-protein and Genbank databases of the homologs among ascoviruses, iridoviruses, phycodnaviruses, mimiviruses and asfarviruses. The procedures were done until we were only able to detect exact match with the syntactic model. Whatever obtained with WAPAM, they required a confirmation with other approaches. Here, we used Psipred result comparison for regions with scores over 7 and HCA analyses for regions having scores lower than 7 with Psipred. This simplified the statistical treatment of the result obtained with WAPAM, since all exact matches have significance or a score of 100%.

Syntactic model for MCP orthologs in ascovirus and iridovirus genomes: G-[DE]-x(1)-ILMVFYW]-x(20,26)-[ILMV-FYW]-x(5)-[ILMVFYW]-x(3)-[ILMVFYW]-[ILMVFYW]-x(4)-[ILMVFYW]-x(17)-[ILMVFYW]-x(10,16)-[ILMV-FYW]-x(5,30)-[ILMVFYW]-x(1)-[ILMVFYW]-P-x(2)-[ILM-VFYW]-[ILMVFYW]-[ILMVFYW]-x(3,6)-[AST]-x(14)-[ILMVFYW]-x(7)-[ILMVFYW]-[ILMVFYW]-[ILMVFYW]-x(18,70)-[ILMVFYW]-x(15)-P. The syntactic model for MCP orthologs in ascovirus, iridovirus, phycodnavirus and asfarvirus genomes: [GE]-[DE]-x(1)-[ILMVFYW]-x(20,80)-[ILMVFYW]-x(9)-[ILMVFYW]-[ILMVFYW]-x(4)-[ILMVFYW]-x(12,17)-[ILMVFYW]-x(10,16)-[ILMVFYW]-x(5,105)-[ILMVFYW]-x(1)-[ILMVFYW]-x(3)-[ILMVFYW]-[ILMVFYW]-[ILMVFYW]-x(3,6)-[AST]-x(14)-[ILMVFYW]-x(7)-[ILMVFYW]-[ILMVFYW]-x(18,90)-P.

Syntactic model for orthologs of the SfAV1a ORF061 in ascovirus, iridovirus, phycodnavirus and asfarvirus genomes: [FWY]-x(3)-[ILMVFYW]-[ILMVFYW]-x(3)-[RHK]-x(15,30)-[ILMVFYW]-x(2)-[ILMVFYW]-x(3)-LPC-x(2,4)-C-x(28,38)-[RHK]-x(2)-[ILMVFYW].

Abbreviations
Acc N°: accession number; AcMNPV: *Autographa californica* multiple nucleopolyhedrovirus; BV: Bracovirus; CcBV: *Cotesia congregata* bracovirus; CIV: chilo iridescent virus; CsIV: *Campoletis sonorensis* Ichnovirus; DNA: deoxyribonucleic acid; DpAV4a: *Diadromus pulchellus* ascovirus 4a; GfIV: *Glypta fumiferanae* Ichnovirus; HfIV: *Hyposoter fugitivus* ichnovirus; HvAV3e: *Heliothis virescens* ascovirus 3e; IV: Ichnovirus; MCP: Major capsid protein; NCLDV: nucleocytoplasmic large DNA viruses; NTP: nucleotide triphosphate; ORF: Open reading frame; P: protein; RNA: ribonucleic acid; SfAV1a: *Spodoptera frugiperda* ascovirus 1a; TnAV2c: *Trichoplusia ni* ascovirus 2c; TrIV: *Tranosema rostrales* ichnovirus; VLP: virus-like particle; WAPAM: weighted automata pattern matching; XcGV: *Xestia c-nigrum* granulovirus.

Authors' contributions

YB is the leader of all aspects of the research on the biology, genomics, and evolution of DpAV4. SS coordinated the sequencing, assembly, and sequence quality control of the DpAV4 genome. CAG participated in the bioinformatics analysis of the DpAV4 genome development of the manuscript. BAF contributed original concepts regarding the evolutionary origins and role of polydnaviruses in endoparasitoid biology, provided virological expertise to optimize data interpretation, and participated in writing the manuscript.

Additional material

Additional File 1

Predicted ORFs in DpAV4 genome. Table in which is indicated the ORF number, its location in the nucleic acid sequence, the size of the peptide encoded by each ORF and its Molecular mass, and its putative function.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-253-S1.doc]

Additional File 2

Properties of the ORFs within the DpAV4 13 kbp regions containing homologs of the GfIV genes with a pox-D5 domain (Acc. N° CU467486). Table in which is described the properties of seven ORFs: ORF number, location within the 13-kbp of the Fig. 1., the size of the peptide encoded by each ORF and its Molecular mass, its pI, its putative function and its closest viral or cellular orhologs in databases.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-253-S2.doc]

Additional File 3

Analysis of ascovirus and ichnovirus virion proteins. 3a: Amino acid sequence alignment of the major capsid protein from selected large double-stranded DNA viruses. The seven conserved domains in MCP sequences of SfAV-1a, TnAV-2a, HvAV-3c, DpAV-4a, CIV, IV29, IV1, IV22, IV16, FV3, LCDV, CHV and ASFV were aligned. Identical and 100% similar positions are highlighted. 3b: HCA analysis of the conserved domain in MCP sequences. For each of the seven conserved domain in MCP, clusters that are conserved in sequences and in HCA graphs are highlighted in yellow. 3c: HCA analysis of the conserved domain in the MCP sequence of CsIV. Each of the seven conserved domain in MCP of ascoviruses, iridoviruses, phycodnaviruese and asfarvirus was located in the MCP sequence of CsIV. Clusters that are conserved in sequences and in HCA graphs are highlighted in yellow. 3d: Amino acid sequence alignment of the protein encoded by orthologous genes of the SfAV1a ORF061. Sequences of SfAV1a ORF061 orthologs extracted from in DpAV4, HvAV3e, SfAV1a, TnAV2c, CIV, MIV, FV3, LCDV, SGV, PBCV, ESV and ASFV genome were aligned. Identical and 100% similar positions are highlighted. 3e: HCA analysis of the conserved domain in SfAV1a ORF061. Clusters that are conserved in sequences and in HCA graphs of SfAV1a ORF061 orthologs are highlighted in yellow.
Click here for file
[http://www.biomedcentral.com/content/supplementary/1471-2148-8-253-S3.doc]

References

1. LaSalle J, Gauld ID: **Hymenoptera and Biodiversity.** Wallingford: CAB International, United Kingdom; 1993.
2. Rotheram S: **Immune surface of eggs of a parasitic insect.** *Nature* 1967, **214:**700.
3. Salt G: **The resistance of insect parasitoids to the defense reactions of their hosts.** *Biol Rev* 1968, **43:**200-232.
4. Bedwin O: **An Insect Glycoprotein: A Study of the Particles Responsible for the Resistance of a Parasitoid's Egg to the Defence Reactions of Its Insect Host.** *Proc R Soc Lond B Biol Sci* 1979, **205(1159):**271-86.
5. Schmidt O, Schuchmann-Feddersen I: **Role of virus-like particles in parasitoid-host interaction of insects.** *Subcell Biochem* 1989, **15:**91-119.
6. Richards EH, Parkinson NM: **Venom from the endoparasitic wasp Pimpla hypochondriaca adversely affects the morphology, viability, and immune function of hemocytes from larvae of the tomato moth, Lacanobia oleracea.** *J Invertebr Pathol* 2000, **76:**33-42.
7. Stasiak K, Renault S, Federici BA, Bigot Y: **Characteristics of pathogenic and mutalistic relationships of ascoviruses in field populations of parasitoid wasps.** *J Insect Physiol* 2005, **51:**103-115.
8. Webb BA, Strand MR, Dickey SE, Beck MH, Hilgarth RS, Barney WE, Kadash K, Kroemer JA, Lindstrom KG, Rattanadechakul W, Shelby KS, Thoetkiattikul H, Turnbull MW, Witherell RA: **Polydnavirus genomes reflect their dual roles as mutualists and pathogens.** *Virology* 2006, **347:**160-174.
9. Vinson B, Scott JR: **Particles containing DNA associated with the oocyte of an insect parasitoid.** *J Invertebr Pathol* 1975, **25:**375-378.
10. Webb BA, Beckage NE, Harayakawa Y, Krell PJ, Lanzrein B, Stoltz DB, Strand MR, Summers MD: **Family *Polydnaviridae*.** In *Virus Taxonomy. Eighth report of the International Commitee on Taxonomy of Viruses* Edited by: Fauquet CM, Mayot MA, Maniloff J, Dessselberger U, Ball LA. Academic Press, London, United Kingdom; 2005:253-260.
11. Edson KM, Vinson SB, Stoltz DB, Summers MD: **Virus in aparasitoid wasp: suppression of the cellular immune response in the parasitoid's host.** *Science* 1981, **211:**582-583.
12. Stoltz DB, Krell P, Summers MD, Vinson SB: **Polydnaviridae – a proposed family of insect viruses with segmented, double-stranded, circular DNA genomes.** *Intervialogy* 1984, **21:**1-4.
13. Espagne E, Dupuy C, Huguet E, Cattolico L, Provost B, Martin N, Poirié M, Periquet G, Drezen JM: **Genome sequence of a polydnavirus: insights into symbiotic virus evolution.** *Science* 2004, **306:**386-289.
14. Tanaka K, Lapointe R, Barney WE, Makkay AM, Stoltz D, Cusson M, Webb BA: **Shared and species-specific features among ichnovirus genomes.** *Virology* 2007, **363:**26-35.
15. Federici BA, Bigot Y: **Origin and evolution of polydnaviruses by symbiogenesis of insect DNA viruses in endoparasitic wasps.** *J Insect Physiol* 2003, **49:**419-432.
16. Margulis L: **Symbiosis in cell evolution.** Freeman: New York; 1993.
17. Avery SW, Jouvenaz DP, Banks WA, Anthony DW: **Virus-like particles in a fire ant, *Solenopsis* sp., (Hyenoptera: Formicidae) from Brazil.** *Florida Entomol* 1977, **60:**17-20.

18. Hess RT, Benham GS, Poinar GO: **The ultrastructure of micro-organisms in the tissues of *Casenaria infesta* (Hymenoptera: Ichneumonidae).** *J Invert Pathol* 1975, **26**:181-191.

19. Stoltz DB, Faulkner G: **Apparent replication of an unusual virus-like particle in both parasitoid wasp and its host.** *Can J Microbiol* 1978, **24**:1509-1514.

20. Stoltz DB, Krell P, Cook D, MacKinnon ZA, Lucarotti CJ: **An unu-sual virus from the parasitic wasp *Cotesia melanoscela*.** *Virology* 1988, **162**:311-320.

21. Barratt BIP, Evans AA, Stoltz DB, Vinson SB, Easingwood R: **Virus-like particles in the ovaries of *Microctonus aethiopoides* loan (Hymenoptera: Braconidae), a parasitoid of adult weevils (Coleoptera: Curculionidae).** *J Invert Pathol* 1999, **73**:182-188.

22. Stasiak K, Renault S, Demattei MV, Bigot Y, Federici BA: **Evidence for the evolution of ascoviruses from iridoviruses.** *J Gen Virol* 2003, **84**:2999-3009.

23. Bideshi DK, Demattei MV, Rouleux-Bonnin F, Stasiak K, Tan Y, Bigot S, Bigot Y, Federici BA: **Genomic sequence of *Spodoptera fru-giperda Ascovirus 1a*, an enveloped, double-stranded DNA insect virus that manipulates apoptosis for viral reproduc-tion.** *J Virol* 2006, **80**:11791-11805.

24. Wang L, Xue J, Seaborn CP, Arif BM, Cheng XW: **Sequence and organization of the *Trichoplusia ni ascovirus* 2c (Ascoviridae) genome.** *Virology* 2006, **10**: 354:167-177.

25. Asgari S, Davis J, Wood D, Wilson P, McGrath A: **Sequenceand organization of the *Heliothis virescens ascovirus* genome.** *J Gen Virol* 2007, **88**:1120-1132.

26. Bigot Y, Rabouille A, Doury G, Sizaret PY, Delbost F, Hamelin MH, Periquet G: **Biological and molecular features of the relation-ships between *Diadromus pulchellus* ascovirus, a parasitoid hymenopteran wasp (*Diadromus pulchellus*) and its lepidop-teran host, *Acrolepiopsis assectella*.** *J Gen Virol* 1997, **78**:1149-1163.

27. Renault S, Petit A, Bénédet F, Bigot S, Bigot Y: **Effectsof the *Diadromus pulchellus* ascovirus, DpAV-4, on thehemocytic encapsulation response and capsule melanization of the leek-moth pupa, *Acrolepiopsis assectella*.** *J Insect Physiol* 2002, **48**:297-302.

28. Lapointe R, Tanaka K, Barney WE, Whitfield JB, Banks JC, Beliveau C, Stoltz D, Webb BA, Cusson M: **Genomic and morphological fea-tures of a banchine polydnavirus: comparison with bracovi-ruses and ichnoviruses.** *J Virol* 2007, **81**:6491-6501.

29. Filée J, Siguier P, Chandler M: **I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses.** *Trends Genet* 2007, **23**:10-15.

30. Afonso CL, Tulman ER, Lu Z, Oma E, Kutish GF, Rock DL: **The genome of *Melanoplus sanguinipes* entomopoxvirus.** *J Virol* 1999, **73**:533-552.

31. Deng L, Webb BA: **Cloning and expression of a gene encoding a *Campoletis sonorensis* polydnavirus structural protein.** *Arch Insect Biochem Physiol* 1999, **40**:30-40.

32. Deng L, Stoltz DB, Webb BA: **A gene encoding a polydnavirus structural polypeptide is not encapsidated.** *Virology* 2000, **269**:440-450.

33. Bamford DH, Grimes JM, Stuart DI: **What does structure tell us about virus evolution?** *Curr Opi Struct Biol* 2005, **15**:655-663.

34. Guyetant S, Giraud M, L'Hours L, Derrien S, Rubini S, Lavenier D, Raimbault F: **Cluster of re-configurable nodes for scanning large genomic banks.** *Parallel Computing* 2005, **31**:73-96.

35. Callebaut I, Labesse G, Durand P, Poupon A, Canard L, Chomilier J, Henrissat B, Mornon JP: **Deciphering protein sequence informa-tion through hydrophobic cluster analysis (HCA): current status and perspectives.** *Cell Mol Life Sci* 1997, **53**:621-645.

36. Eaton HE, Metcalf J, Penny E, Tcherepanov V, Upton C, Brunetti CR: **Comparative genomic analysis of the family *Iridoviridae*: re-annotating and defining the core set of iridovirus genes.** *Virol J* 2007, **4**:1-17.

37. Masutani H, Ueda S, Yodoi J: **The thioredoxin system in retrovi-ral infection and apoptosis.** *Cell Death Differ* 2005, **12**:991-998.

38. Renesto P, Abergel C, Decloquement P, Moinier D, Azza S, Ogata H, Fourquet P, Gorvel JP, Claverie JM: **Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products.** *J Virol* 2006, **80**:11678-11685.

39. Fenouillet E, Barbouche R, Jones IM: **Cell entry by enveloped viruses: redox considerations for HIV and SARS-coronavi-rus.** *Antioxid Redox Signal* 2007, **9**:1009-1034.

40. Tessman I: **Genetic recombination of the DNA plant virus PBCV-1 in a chlorella alga.** *Virology* 1985, **145**:319-322.

41. Iyer LM, Aravind L, Koonin EV: **Common origin of four diverse families of large eukaryotic DNA viruses.** *J Virol* 2001, **75**:11720-11734.

42. Iyer LM, Balaji S, Koonin EV, Aravind L: **Evolutionary genomics of nucleo-cytoplasmic large DNA viruses.** *Virus Res* 2006, **117**:156-184.

43. Hamm JJ, Nordlund DA, Marti OG: **Effects of the nonoccluded virus of *Spodoptera frugiperda* (Lepidoptera: Noctuidae) on the development of a parasitoid, *Cotesia marginventris* (Hymenoptera : Brachonidae).** *Environ Entomol* 1985, **14**:258-261.

44. López M, Rojas JC, Vandame R, Williams T: **Parasitoid-mediated transmission of an iridescent virus.** *J Invertebr Pathol* 2002, **80**:160-170.

45. Lawrence P: **Non-poly-DNA viruses, their parasitic wasp, and hosts.** *J Insect Physiol* 2005, **51**:99-101.

46. Drezen JM, Bézier A, Lesobre J, Huguet E, Cattolico L, Périquet G, Dupuy C: **The few virus-like genes of *Cotesia congregata*.** *Arch Insect Biochem Phys* 2006, **61**:110-122.

47. Kapitonov VV, Jurka J: **Self-synthesizing DNA transposons in eukaryotes.** *Proc Natl Acad Sci USA* 2006, **103**:4540-4545.

48. Pritham EJ, Putliwala T, Feschotte C: *Mavericks*, **a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses.** *Gene* 2006, **390**:3-17.

49. Becker Y: **Evolution of viruses by acquisition of cellular RNA or DNA nucleotide sequences and genes: an introduction.** *Virus Genes* 2000, **21**:7-12.

50. Salzberg SL, White O, Peterson J, Eisen JA: **Microbialgenes in the human genome: lateral transfer or gene loss?** *Science* 2001, **292**:1903-1906.

51. Andersson JO, Doolittle WF, Nesbo CL: **Are there bugs in our genome?** *Science express* 2001:online-1062241.

52. Hughes AL, Friedman RF: **Genome-wide survey for genes hori-zontally transferred from cellular organisms to baculovi-ruses.** *Mol Biol Evol* 2003, **20**:979-987.

53. Chapman MJ, Margulis L: **Morphogenesis by symbiogenesis.** *Int Microbiol* 1998, **1**:319-326.

54. Timmis JN, Ayliffe MA, Huang CY, Martin W: **Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes.** *Nat Rev Genet* 2004, **5**:123-135.

55. Trémouillaux-Guiller J, Huss VA: **A cryptic intracellular green alga in Ginkgo biloba: ribosomal DNA markers reveal world-wide distribution.** *Planta* 2007, **226**:553-557.

56. Rudgers JA, Holah J, Orr SP, Clay K: **Forest succession sup-pressed by an introduced plant-fungal symbiosis.** *Ecology* 2007, **88**:18-25.

57. Dupuy C, Huguet E, Drezen JM: **Unfolding the evolutionary story of polydnaviruses.** *Virus Res* 2006, **117**:81-89.

58. Desjardins CA, Gundersen-Rindal DE, Hostetler JB, Tallon LJ, Fuester RW, Schatz MC, Pedroni MJ, Fadrosh DW, Haas BJ, Toms BS, Chen D, Nene V: **Structure and evolution of a proviral locus of *Glyp-tapanteles indiensis* bracovirus.** *BMC Microbiol* 2007, **7**:61.