XIII Argentine Congress of Bioinformatics and Computational Biology

XIII International Conference of the Iberoamerican Society of Bioinformatics

III Annual Meeting of the Ibero-American Artificial Intelligence Network for Big BioData

Proteomic study of the Occlusion Bodies of Spodoptera frugiperda Granulovirus

Tomás Masson[1], Matías L. Pidre[1], Ricardo Salvador[2], Víctor Romanowski[1], María Leticia Ferrelli[1]

[1]Instituto de Biotecnología y Biología Molecular (IBBM), Facultad de Ciencias Exactas, Universidad Nacional de La Plata - CONICET, La Plata, Buenos Aires, Argentina

[2]Instituto de Microbiología y Zoología Agrícola (IMyZA), Centro de Investigaciones en Ciencias Agronómicas y Veterinarias (CICVyA), Instituto Nacional de Tecnología Agropecuaria (INTA), Hurlingham, Buenos Aires, Argentina

INTRODUCTION

Spodoptera frugiperda Granulovirus (SpfrGV) infects the fall armyworm *Spodoptera frugiperda*, an important pest for maize and other crops. SpfrGV belongs to the *Baculoviridae* family, an attractive group of insect viruses with potential applications in biological pest control. Baculoviruses produce two virion morphotypes, being the Occlusion Derived Virus (ODV) the responsible for primary infection within the larval midgut. These virions are found in the environment embedded in a proteinaceous matrix that forms occlusion bodies (OBs), a phenotype that confers protection from adverse environmental conditions. As OBs play a key role in horizontal transmission, we used mass spectrometry (MS) to investigate the protein composition of the OBs of an Argentinean isolate of SpfrGV and analyzed the detected proteins.

MATERIALS AND METHODS

SpfrGV OBs were purified from infected *S. frugiperda* larvae. Total protein mass in the sample was quantified using the Bradford assay. The subsequent analyses were performed at the Proteomics Core Facility CEQUIBIEM, at the University of Buenos Aires/CONICET. Briefly, the sample was trypsin digested and subjected to nanoHPLC coupled to a mass spectrometer (MS) with Orbitrap technology, allowing for the separation and subsequent identification of peptides obtained from trypsin digestion. The ionization of the samples was achieved through electrospray. Data analysis was conducted using the Proteome Discoverer program using a database with predicted proteins from the SpfrGV genome (Ferrelli *et al.* 2018). Relative abundance of proteins were calculated using the emPAI value and normalized their abundance using VP39 as a reference.

We subjected the detected proteins to AlphaFold prediction using ColabFold (Mirdita *et al.* 2022). In order to find similarities among the structures we used DALI all-against-all comparison. Resulting similar structures were aligned with Flexible structure AlignmenT by Chaining Aligned fragment pairs allowing Twists (FATCAT) to confirm similarities.

RESULTS

The peptides produced by MS were successfully mapped to the theoretical proteome of SpfrGV ARG. This permitted the identification of 72 proteins within the OB, constituting 47% of the theoretical proteome encoded in the virus genome. Twenty-eight proteins can be categorized within the essential 39 core gene group of *Baculoviridae* (Fig. 1). Among the 44 remaining proteins, several correspond to conserved structural and accessory proteins commonly found in baculovirus. However, there are at least 18 of them with no known associated function, including ORF 28 and ORF 40, that are unique to SpfrGV. We classified the detected proteins according to their role described in bibliography and calculated their relative abundance (Fig 2)

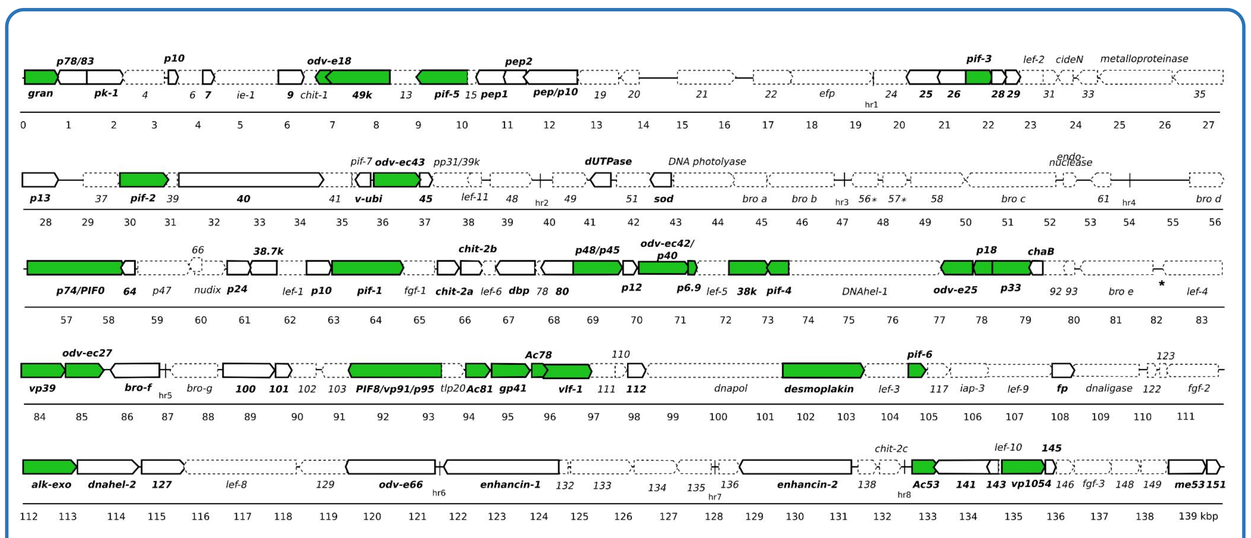

Figure 1. Distribution of 72 detected proteins along SpfrGV genome. Genes that code for proteins detected in this proteomic study are highlighted with solid lines. Baculovirus core genes are in green.

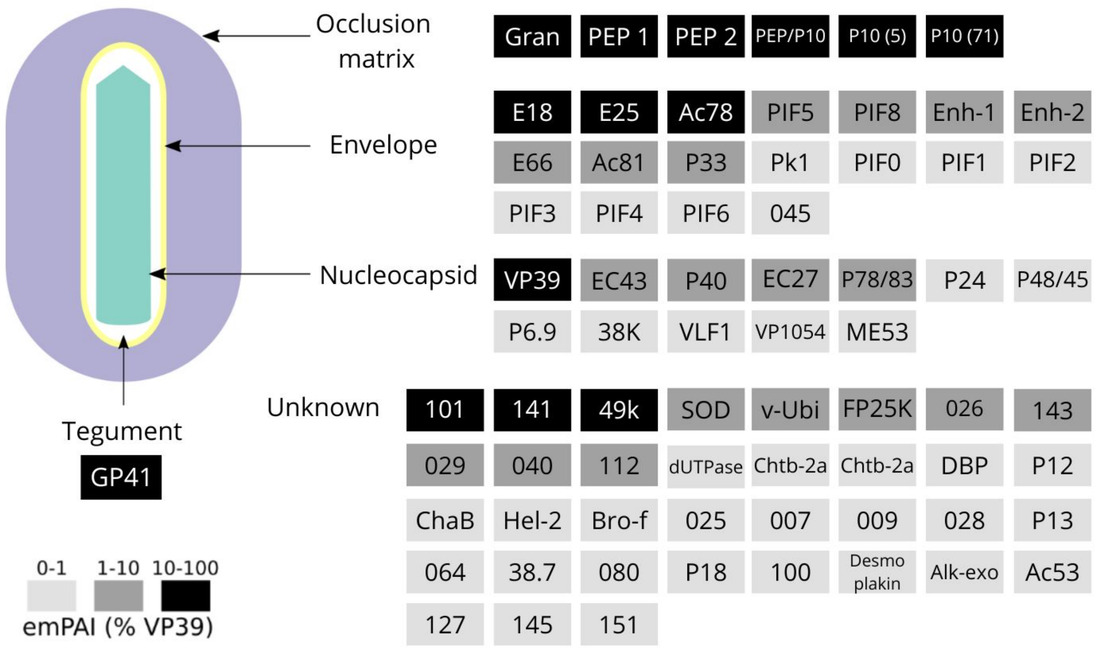

Figure 2. Localization and protein abundance of 72 detected proteins in the SpfrGV occlusion body. Proteins localizations are classified according bibliography of their homologues.Protein abundance was classified in three categories according to their empPAI values relative to the capsid protein VP39.

In order to find more information on the proteins associated with SpfrGV OB we used ColabFold to predict all 72 protein structures. Upon visual examination, we observed some proteins with apparent similarity. In order to assess structural similarity among SpfrGV OB proteins we used DALI all-against-all comparison and found that 6 proteins shared domain similarities (Fig 3). Interestingly, 3 of this group are well characterized baculovirus proteins: ORFs 16, 17 and 18 code for the Polyhedron Envelope Protein N-terminus domain (Baculo_PEP_N, PFAM 04512) that plays a crucial role in the formation of the polyhedron envelope, which surrounds the occlusion bodies (OBs) providing stability and resistance to adverse external environment. On the other hand, ORFs 7, 25 and 98 are uncharacterized proteins.

To confirm the structural similarity among them we performed pairwise structural alignments with FATCAT (Fig 4). ORF98 contains three globular domains that were splitted in order to align them with the other proteins and among them. All the alignements, showed significant similarity. PEP-P10 only shares similarity in its N-ter domain.

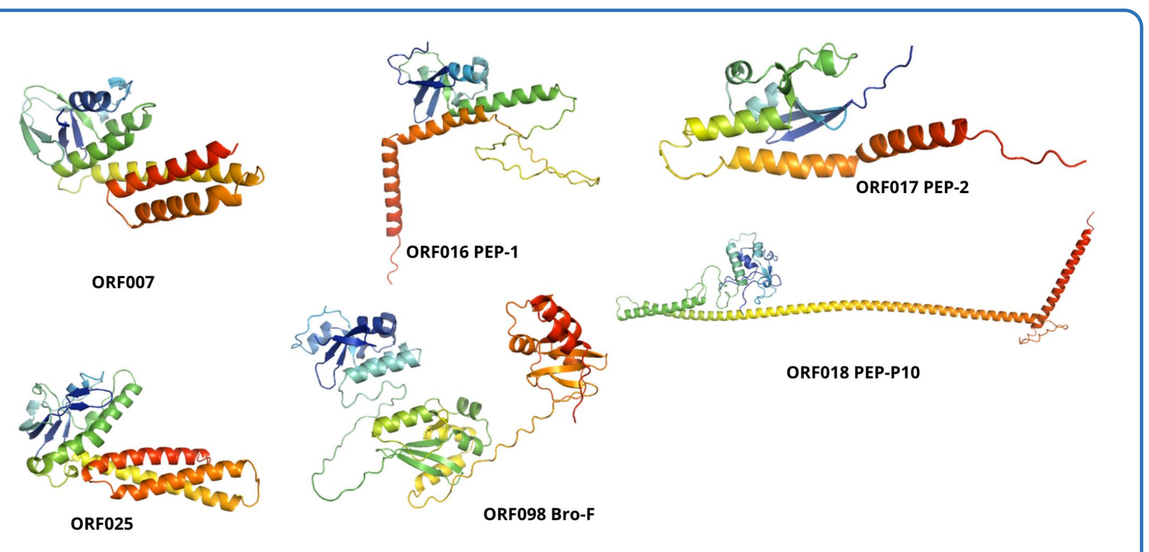

Figure 3. Alphafold structures of 6 detected proteins in SpfrGV OB that share structural similarities. Structure predictions were obtained with ColabFold and the images with rainbow colouring were performed with pymol

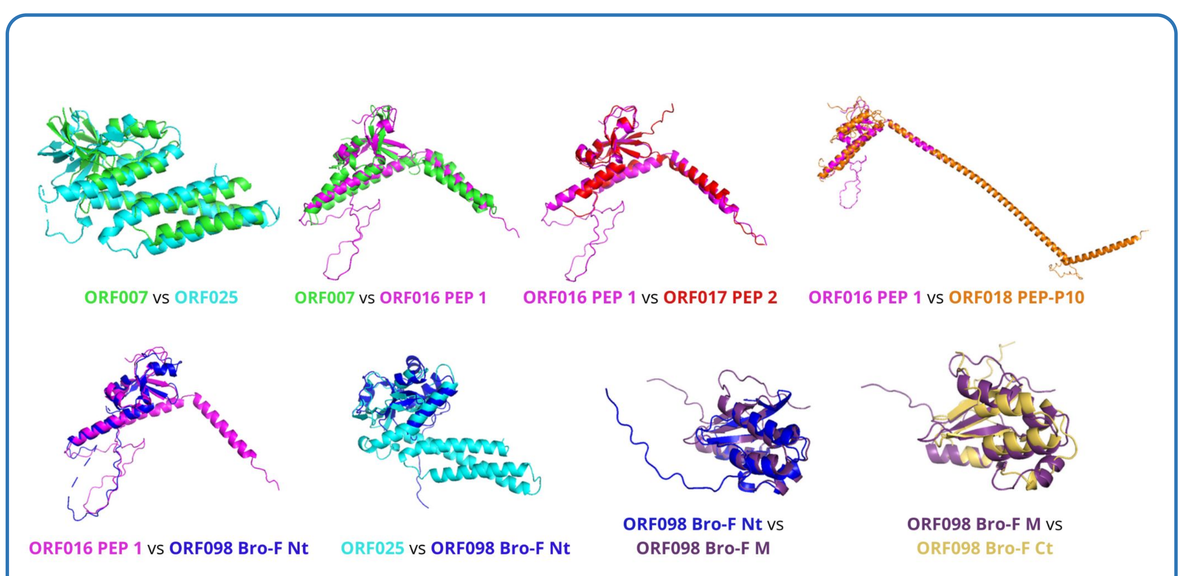

Figure 4. Structural alignments of similar proteins detected with DALI. Representative alignments performed with FATCAT of all against all alignments. ORF98 was chopped in three domains (Nt, M and Ct). All alignments resulted significantly similar according to FATCAT.

CONCLUSIONS

The SpfrGV OB is a complex structure comprising almost half of the proteins coded in its genome, and these are not confined to a particular location. Also, most of the baculovirus core genes end up in this structure. In particular, those known to belong to the virion or matrix structure were found highly abundant. However, we found other highly abundant proteins with no role assigned, like ORF101 and ORF141. Structural comparison among all the proteins detected, allowed finding shared domains between PEPs and three uncharacterized proteins. This suggests they could have a similar role to PEP, which needs to be explored in future assays. The identification of 72 proteins within SpfrGV OBs, along with novel insights from structural comparisons, enhances our understanding of these OBs and associated proteins. This knowledge has implications for virus transmission dynamics and for the development of biological pest control strategies.

REFERENCES

- Ferrelli, M. L., Pidre, M. L., Ghiringhelli, P. D., Torres, S., Fabre, M. L., Masson, T., Cédola, M.T., Sciocco-Cap, A. & Romanowski, V. (2018). Genomic analysis of an Argentinean isolate of Spodoptera frugiperda granulovirus reveals that various baculoviruses code for Lef-7 proteins with three F-box domains. Plos one, 13(8), e0202598.
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nature methods, 19(6), 679-682.
- Holm, L., Kääriäinen, S., Wilton, C., & Plewczynski, D. (2006). Using Dali for structural comparison of proteins. Current protocols in bioinformatics, 14(1), 5-5.
- Ye, Y., & Godzik, A. (2004). FATCAT: a web server for flexible structure comparison and structure similarity searching. Nucleic acids research, 32(suppl_2), W582-W585.