



HAL
open science

Qualitative proteomic analysis of *Tipula oleracea* nudivirus occlusion bodies

Annie Bézier, Grégoire Harichaux, Karine Musset, Valérie Labas, Elisabeth A.
Herniou

► **To cite this version:**

Annie Bézier, Grégoire Harichaux, Karine Musset, Valérie Labas, Elisabeth A. Herniou. Qualitative proteomic analysis of *Tipula oleracea* nudivirus occlusion bodies. *Journal of General Virology*, 2017, 98 (2), pp.284-295. 10.1099/jgv.0.000661 . hal-01595093

HAL Id: hal-01595093

<https://hal.science/hal-01595093>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Journal of General Virology

Qualitative proteomic analysis of *Tipula oleracea* nudivirus (ToNV) occlusion bodies --Manuscript Draft--

Manuscript Number:	JGV-D-16-00587R1
Full Title:	Qualitative proteomic analysis of <i>Tipula oleracea</i> nudivirus (ToNV) occlusion bodies
Article Type:	Standard
Section/Category:	Insect Viruses - DNA
Corresponding Author:	Anne Bézier Univ. of Tours, France FRANCE
First Author:	Anne Bézier
Order of Authors:	Anne Bézier Grégoire Harichaux Karine Musset Valérie Labas Elisabeth A. Herniou
Abstract:	<p>Nudiviruses are arthropod-specific large double-stranded circular DNA viruses, related to baculoviruses, which replicate in the nucleus of the cells they infect. Up to date six fully sequenced nudiviral genomes are available in databases and protein profile from nudivirus particles were mainly characterized by polyacrylamide gel electrophoresis. However, only few direct matches were completed between genomic and proteomic data to the exception of the major occlusion body protein from <i>Penaeus monodon</i> nudivirus (PmNV) and four nucleocapsid proteins from <i>Helicoverpa zea</i> nudivirus 2 (HzNV-2). Function of predicted nudiviral proteins is still inferred from what is known from baculoviruses or endogenous nudiviruses (i.e. bracoviruses). <i>Tipula oleracea</i> nudivirus (ToNV) is the causative agent of crane fly nucleopolyhedrosis. With PmNV, ToNV is the second fully sequenced nudivirus to be described as forming occlusion bodies. Protein profile revealed by Coomassie-stained SDS-PAGE is quite similar to those observed for other nudiviruses with five major protein bands of about 75, 48, 35, 25 and 12 kDa. Proteomic analysis using on-line nanoflow liquid chromatography tandem high resolution mass spectrometry revealed ToNV occlusion bodies are composed of 52 viral proteins, the most abundant of which are the functional homolog of baculovirus polyhedrin/granulin and the homologs of three HzNV-2 predicted proteins: the two virion structural proteins 34K (Hz2V052, the baculovirus capsid protein VP39 homolog) and 11K (Hz2V025); and the hypothetical protein Hz2V079, a newly identified nudivirus core gene product.</p>

1 Article type: Standard research paper

2 Section/Category: DNA Viruses - Insect

3

4 Title: **Qualitative proteomic analysis of *Tipula oleracea* nudivirus (ToNV) occlusion**
5 **bodies**

6 Running title: ToNV Proteomics

7

8 Annie Bézier,¹ Grégoire Harichaux,² Karine Musset,¹ Valérie Labas,² Elisabeth A. Herniou¹.

9

10 ¹Institut de Recherche sur la Biologie de l’Insecte (IRBI), UMR 7261 CNRS Université François-Rabelais, 37200 Tours, France

11 ²INRA, PRC UMR85-CNRS 7247-UFR-IFCE, Laboratoire de Spectrométrie de masse, Plateforme d’Analyse Intégrative des Biomolécules
12 et de Phénomique des Animaux d’Intérêt Bio-agronomique (PAIB2), 37380 Nouzilly, France

13

14 Correspondence: Annie Bézier; annie.bezier@univ-tours.fr

15 Phone number +33 2 47 36 73 83

16 Fax number +33 2 47 36 69 66

17

18 Keywords: Nudivirus, Proteome

19

20 Word count (Abstract): 228

21 Word count (Main text): 5272

22 Categories: Proteomics

23 Figure number: 2

24 Table number: 2

25

26 **ABSTRACT**

27 Nudiviruses are arthropod-specific large double-stranded circular DNA viruses, related to
28 baculoviruses, which replicate in the nucleus of the cells they infect. Up to date six fully
29 sequenced nudiviral genomes are available in databases and protein profile from nudivirus
30 particles were mainly characterized by polyacrylamide gel electrophoresis. However, only
31 few direct matches were completed between genomic and proteomic data to the exception of
32 the major occlusion body protein from *Penaeus monodon* nudivirus (PmNV) and four
33 nucleocapsid proteins from *Helicoverpa zea* nudivirus 2 (HzNV-2). Function of predicted
34 nudiviral proteins is still inferred from what is known from baculoviruses or endogenous
35 nudiviruses (i.e. bracoviruses). *Tipula oleracea* nudivirus (ToNV) is the causative agent of
36 crane fly nucleopolyhedrosis. With PmNV, ToNV is the second fully sequenced nudivirus to
37 be described as forming occlusion bodies. Protein profile revealed by Coomassie-stained
38 SDS-PAGE is quite similar to those observed for other nudiviruses with five major protein
39 bands of about 75, 48, 35, 25 and 12 kDa. Proteomic analysis using on-line nanoflow liquid
40 chromatography tandem high resolution mass spectrometry revealed ToNV occlusion bodies
41 are composed of 52 viral proteins, the most abundant of which are the functional homolog of
42 baculovirus polyhedrin/granulin and the homologs of three HzNV-2 predicted proteins: the
43 two virion structural proteins 34K (Hz2V052, the baculovirus capsid protein VP39 homolog)
44 and 11K (Hz2V025); and the hypothetical protein Hz2V079, a newly identified nudivirus core
45 gene product.

46

47 INTRODUCTION

48

49 Formerly described as non-occluded baculoviruses (Huger & Krieg 1991), nudiviruses (NV)
50 form a highly diverse group that shares biological, structural and genomic features with other
51 nuclear arthropod large circular dsDNA viruses (baculoviruses, hytrosaviruses, nimaviruses or
52 filamentous viruses). However, they also share lineage specific characters (Wang *et al.*, 2012;
53 Jehle *et al.*, 2013a, Bézier *et al.*, 2015; Gauthier *et al.*, 2015), which justified the creation of
54 the new *Nudiviridae* family by the International Committee on Taxonomy of Viruses (ICTV)
55 (Jehle *et al.*, 2013b).

56 Nudiviruses display wider host range than baculoviruses, hytrosaviruses or nimaviruses, as
57 they infect more insect orders, as well as crustaceans. They have been described in
58 Lepidoptera with *Heliothis zea* nudivirus-1 (HzNV-1) (Granados *et al.*, 1978; Cheng *et al.*,
59 2002), and *Helicoverpa zea* nudivirus-2 (HzNV-2) (Raina *et al.*, 2000, Burand & Rallis, 2004;
60 Burand *et al.*, 2012), in Diptera with *Drosophila innubia* NV (DiNV) (Unkless, 2011), *Tipula*
61 *paludosa* NV (TpNV) (Smith & Xeros, 1954; Bergoin & Guelpa 1977) and *Tipula oleracea*
62 NV (ToNV) (Bézier *et al.*, 2015), in Orthoptera with *Gryllus bimaculatus* NV (GbNV) (Huger
63 1985, Huger & Krieg, 1991; Wang *et al.*, 2007a), in Coleoptera with *Oryctes rhinoceros* NV
64 (OrNV) (Huger, 1966; Payne 1974; Payne *et al.*, 1977; Wang *et al.*, 2007b), and in Decapoda
65 with *Penaeus monodon* NV (PmNV) (Mari *et al.*, 1993; Yang *et al.*, 2014), *Penaeus vannamei*
66 single nucleopolyhedrovirus (PvNV) (Bonami *et al.*, 1995) and possibly *Crangon crangon*
67 intranuclear bacilliform virus (CrcrIBV) (Stentiford *et al.*, 2004). Other nudiviruses are also
68 suspected in Trichoptera, Siphonaptera, Neuroptera, Homoptera, Thysanura, Acarina and
69 Araneina (Huger & Krieg, 1991). So far, only six fully sequenced genomes are available in
70 databases: HzNV-1 (AF451898; Cheng *et al.*, 2002) and HzNV-2 (NC_004156; Burand *et al.*,

71 2012), GbNV (EF203088; Wang *et al.*, 2007a), OrNV (EU747721; Wang *et al.*, 2007b),
72 ToNV (KM610234; Bézier *et al.*, 2015) and PmNV (KJ184318; Yang *et al.*, 2014).

73 Nudiviruses are sexually or orally transmitted. All developmental stages (larvae, pupae,
74 adults) can become infected. Depending on the host, the infection can be asymptomatic in
75 both larvae and adults (HzNV-2), lethal in larvae and rather chronic in adults (GbNV and
76 OrNV), and also induce malformations and sterility (HzNV-2) (reviewed in Burand, 1998).
77 The tissue tropism is variable and concerns reproductive tissues (HzNV-2), fat body (GbNV
78 and OrNV), mid-gut (OrNV), haemocytes (TpNV and PmNV) or hepatopancreas (PmNV).
79 Nudiviruses are particularly prone to endogenization, as they are able to integrate their
80 genome in host cells (Granados *et al.*, 1978; Lin *et al.*, 1999) and to infect reproductive
81 tissues (Raina *et al.*, 2000). Recent reports highlight striking cases of nudivirus domestication
82 especially the symbiotic bracoviruses associated with braconid wasps (Bézier *et al.*, 2009)
83 since 100 MYA (Thézé *et al.*, 2011), which likely derive from a nudivirus more related to
84 ToNV (Bézier *et al.*, 2015), or the virus-like particles (VLP) from the ichneumonid wasp
85 *Venturia canescens*, derived from an *Alphanudivirus* (Pichon *et al.*, 2015). So far, there is no
86 evidence that the endogenous alphanudivirus of the brown planthopper, *Nilaparvata lugens*
87 (Cheng *et al.*, 2014), retains any functional activity.

88 A point of particular interest in the *Nudiviridae* is their phenotypic diversity, as within this
89 family some viruses are transmitted as enveloped occlusion bodies (OBs), such as ToNV and
90 TpNV (Bergoin & Guelpa, 1977; Bézier *et al.*, 2015), or non-enveloped OBs, such as PmNV
91 (Yang *et al.*, 2014; Chaivisuthangkura *et al.*, 2008; Bonami *et al.*, 1997), others as non-
92 occluded virions, such as *Penaeus* NVs (Couch, 1974; Yang *et al.*, 2014), HzNV-1 (Granados
93 *et al.*, 1978) or GbNV (Granados *et al.*, 1978; Huger & Krieg, 1991), or as facultatively
94 occluded virions, such as OrNV (Huger & Krieg, 1991; Raina *et al.*, 2000). The genetic bases
95 of OrNV facultative OBs still remain to be determined (Wang *et al.*, 2011a; Bézier *et al.*,

96 2015). Protein profile obtained by polyacrylamide gel electrophoresis (Burand *et al.*, 1983;
97 Crawford & Sheehan, 1985; Bonami *et al.*, 1997) estimate nudivirus particles to contain less
98 than 30 proteins. With the exception of the major occlusion body protein (MOBP) from
99 PmNV (Chaivisuthangkura *et al.*, 2008) and four viral structural proteins (11K, 15K, 28K and
100 34K) from HzNV-2 (Kim, 2009), no direct match has ever been made between complete
101 genomic and proteomic data of nudiviruses.

102 The genome of *Tipula oleracea* nudivirus (ToNV) is 145,704 bp and contains 131 predicted
103 open reading frames (ORFs) (Bézier *et al.*, 2015). Comparative genomic analyses revealed
104 ToNV shares 32 core genes with *Nudiviridae* and 21 with *Baculoviridae* (Bézier *et al.*, 2015).
105 Based on sequence similarity with baculoviruses, these genes are involved in DNA replication
106 and processing, transcription, packaging and assembly, and *per os* infectivity (reviewed in
107 Rohrmann, 2013). Scanning and transmission electron microscopy further revealed ToNV
108 enveloped OBs contain a large number of rod-shaped virions embedded within a dense
109 protein matrix (Bézier *et al.*, 2015). In this study, we described the first comprehensive
110 qualitative proteomic analysis of an occluded nudivirus. By using on-line nanoflow liquid
111 chromatography in tandem with high-resolution mass spectrometry (nanoLC-HR-MS/MS),
112 we identified the major occlusion body proteins that compose ToNV occluded particles and
113 further characterized all other components.

114

115 **RESULTS AND DISCUSSION**

116

117 **Coomassie-stained SDS-PAGE profile revealed five major protein bands**

118 Separation of approximately $4 \cdot 10^6$ OBs on a 12.5% SDS-PAGE gel revealed five major
119 protein bands (Fig. 1). The apparent molecular weight of the major band was estimate at
120 25 kDa. Four other major bands were also identified with an apparent molecular weight of

121 75 kDa, 48 kDa, 35 kDa and 12 kDa. A low complexity profile was also observed for the
122 PvNV (Bonami *et al.*, 1995) with a unique major band corresponding to the major
123 polypeptide (polyhedrin); and for the non-occluded baculovirus-like agent (CBV), most
124 probably a nudivirus also pathogenic to penaeid shrimp, with four major bands that were
125 approximately 75, 27.5, 23.5 and 19 kDa (Nadala *et al.*, 1998). Replication studies on other
126 nudiviruses in insect cells revealed the presence of about nine major bands corresponding to
127 structural proteins from about 14 to 153 kDa in HzNV-1 (Burand *et al.*, 1983); 13.5 to
128 190 kDa in HzNV-2 (Kim, 2009) and 11.5 to 69 kDa in OrNV (Crawford & Sheehan, 1985).
129 Protein analyses of occluded circular dsDNA viruses, like baculoviruses, revealed more or
130 less complex profiles with proteins ranging from about 6 to over 200 kDa according to
131 genera, virion morphotypes, techniques used to separate proteins, and staining methods
132 (Braunagel *et al.*, 2003; Deng *et al.*, 2007; Perera *et al.*, 2007; Wang *et al.*, 2010; Wang *et al.*,
133 2011b; Hou *et al.*, 2013). On the other hand, gradient SDS-PAGE analysis of purified
134 *Glossina pallidipes* salivary gland hypertrophy virus (GpSGHV) with Coomassie blue
135 staining revealed a complex profile with at least 35 proteins ranging from 10 to over 130 kDa,
136 but only around ten distinctive bands (Kariithi *et al.*, 2010). Considering the lower sensitivity
137 of the Coomassie blue staining approach compared to radiolabeling, silver or colloidal blue
138 stainings, the ToNV protein profile appeared quite similar to those observed for nudiviruses,
139 baculoviruses or hytrosaviruses.

140

141 **ToNV occlusion bodies contain of at least 48 viral proteins**

142 To allow the identification of all proteins that composed the ToNV occlusion bodies, major
143 bands were excised individually from polyacrylamide gel and the remaining gel lane was
144 divided into twenty other gel sections (Fig. 1). After in-gel trypsin digestion, peptides were
145 extracted from each slice and subjected to bottom-up proteomic (Cottrell, 2011) using

146 nanoLC-HR-MS/MS, which is a highly sensitive proteomic method well suited for complex
147 samples, available in limited quantity and for non-sequenced organisms (Ishihama, 2005;
148 Rieux *et al.*, 2011). Protein identification with a minimal 95.0% probability against ToNV
149 predicted proteins database revealed ToNV OBs were composed of at least 48 viral proteins
150 (Table 1). Among them, 41 proteins were identified with two or more exclusive peptides and
151 seven proteins were unambiguously identified from a unique peptide with a minimal 95.0%
152 probability (Table 1). Four additional proteins could be identified with a unique peptide at the
153 same confidence level (>95%) but a lower protein threshold probability (50 to 79%). They
154 corresponded to GbNVORF19, ORF080, P47 and ORF122. Genes encoding these proteins
155 were distributed throughout the genome, contrary to what was observed in hytrosaviruses, in
156 which genes involved in virions production appears confined to a specific genomic region
157 (Garcia-Maruniak *et al.*, 2008; Kariithi *et al.*, 2010).

158 ToNV OBs thus harboured between 48 and 52 viral proteins, which is the largest number of
159 components detected in nudiviruses. In HzNV-1, HzNV-2 and OrNV, 28, 16 and 27 virus
160 structural proteins were identified respectively, after infection of insect cell lines using pulse-
161 labelling or silver staining approaches (Burand *et al.*, 1983; Lu, 1997; Crawford & Sheehan,
162 1985). This high number of viral proteins identified in ToNV OBs might reflect that ToNV is
163 an occluded nudivirus or the high sensitivity of nanoLC-HR-MS/MS enabling the detection of
164 proteins present in very small amounts. On the other hand, this number was in the range of
165 what has been observed in *Baculoviridae* with equivalent techniques (Braunagel *et al.*, 2003;
166 Perera *et al.*, 2007; Liu *et al.*, 2008; Xu *et al.*, 2011; Wang *et al.*, 2010; Wang *et al.*, 2011b;
167 Hou *et al.*, 2013; Braconi *et al.*, 2014; Zhang *et al.*, 2015). For baculoviruses, proteomic
168 analyses have led to a general understanding of virus particle structure for both virion
169 morphotypes, the occlusion-derived viruses (ODVs) that initiate primary infection and the
170 budded viruses (BVs) that spread the infection from cell to cell leading to a systemic infection

171 (van Oers & Vlak, 2007, Slack & Arif, 2007; Rohrmann, 2013). ODVs are composed of 16 to
172 73 protein products (Liu *et al.*, 2008; Zhang *et al.*, 2015) and BVs of 23 to 51 (Wang *et al.*,
173 2010; Hou *et al.*, 2013). However, a higher number of proteins, up to 86, were detected by
174 LC-MS/MS in hytrosaviruses (Kariithi *et al.*, 2010; Abd-Alla *et al.*, 2016), that are enveloped
175 but not occluded nuclear arthropod large circular dsDNA viruses.

176

177 **Comparison with other arthropod-specific large double-stranded circular DNA viruses**

178 Of the 48 identified proteins with a 95% protein threshold, 21 were already known viral
179 particle components in baculoviruses (NPV and/or GV) (reviewed in Rohrmann, 2013),
180 nudiviruses (NV) (Chaivisuthangkura *et al.*, 2008; Kim, 2009), hytrosaviruses (SGHV)
181 (Garcia-Maruniak *et al.*, 2008; Kariithi *et al.*, 2010) or endogenous nudiviruses such as
182 bracoviruses (Bézier *et al.*, 2009; Wetterwald *et al.*, 2010; Burke *et al.*, 2013) or VLPs of the
183 wasp *Venturia canescens* (VcVLP) (Pichon *et al.*, 2015); and 27 were to date so far unique
184 viral components of ToNV OBs. The previously identified polyhedrin/granulin homolog
185 (MOBP, ORF059) (Bézier *et al.*, 2015) appeared as the major protein component of OBs in
186 ToNV (Table 1), as it is for baculoviruses. This nudiviral polyhedrin/granulin homolog might
187 thus form the matrix of OBs in which ToNV virions are embedded (Bézier *et al.*, 2015) and
188 participate in long-term virion conservation in the environment, although it might confer less
189 effective protection than baculovirus polyhedra (Slack & Arif, 2007; Meynardier *et al.*, 1964).
190 VP39, 38K and VLF-1 homologs found associated with both baculovirus virion morphotypes
191 (Braunagel *et al.*, 2003; Wang *et al.*, 2010; Hou *et al.*, 2013), where they play a crucial role in
192 nucleocapsid formation and assembly, or with particles of endogenous nudiviruses (Bézier *et*
193 *al.*, 2009; Wetterwald *et al.*, 2010; Burke *et al.*, 2013; Pichon *et al.*, 2015), were identified as
194 ToNV OB components (Table 1). P33, which was identified as baculovirus ODV (Braunagel
195 *et al.*, 2003; Perera *et al.*, 2007; Wang *et al.*, 2011b; Hou *et al.*, 2013), bracovirus particles or

196 VcVLP component (Burke *et al.*, 2013; Pichon *et al.*, 2015), was also present. Additionally
197 two nudiviral core proteins, the integrase INT and the FLAP endonuclease FEN-1, most
198 probably involved in DNA metabolism (Grainge & Jayaram, 1999; Liu *et al.*, 2004;
199 Balakrishnan & Bambara, 2013), were identified. However INT is the only one that has
200 already been identified as particles associated protein in bracoviruses (Burke *et al.*, 2013) and
201 hytrosaviruses (GpSGHVORF30; Kariithi *et al.*, 2010). In combination with VLF-1, which is
202 also an integrase/recombinase, the nudiviral INT and FEN-1 might play a critical role in viral
203 DNA and/or virion maturation, but also most generally in viral replication (Vanarsdall *et al.*,
204 2004; Li *et al.*, 2005; Vanarsdall *et al.*, 2006; Burke *et al.*, 2013), they might also be involved
205 in nudiviral DNA integration and its excision from host genome.

206 All PIF proteins (P74, PIF-1 to PIF-6, and VP91), that are essential for oral infectivity of
207 baculoviruses (Peng *et al.*, 2010 ; Peng *et al.*, 2012 ; Zhu *et al.*, 2013), were also identified.
208 PIFs are conserved in exogenous nudiviruses (Bézier *et al.*, 2015) and to a lesser extent in
209 hytrosaviruses (Garcia-Maruniak *et al.*, 2008; Kariithi *et al.*, 2010) and filamentous viruses
210 (Gauthier *et al.*, 2015). Recent studies revealed that these proteins are also conserved particle
211 components of endogenous nudiviruses, which are not transmitted orally (Bézier *et al.*, 2009;
212 Burke *et al.*, 2013; Pichon *et al.*, 2015). As for baculoviruses, PIF proteins must therefore play
213 a crucial role in infectivity of nudivirus virions.

214 Surprisingly, we could detect in ToNV OBs neither the homologs of Ac81 or of P6.9, which
215 they were both detected in baculoviruses (Tweeten *et al.*, 1980; Braunagel *et al.*, 2003; Wang
216 *et al.*, 2010; Hou *et al.*, 2013) and in GpSGHV for Ac81 (Kariithi *et al.*, 2010; Abd-Alla *et al.*,
217 2016). The absence of P6.9 might be due to technical issues. ToNV P6.9 (ORF051) is
218 composed of 119 amino acids, a theoretical trypsin digestion using PeptideCutter ExPasy tool
219 (http://web.expasy.org/peptide_cutter) revealed 33 cleavage sites generating 22 unique
220 peptides, ranging in size from one to a maximum of eleven amino acids (not shown). Only

221 two peptides of 10 and 11 amino acids were theoretically just within the detection limits of
222 the mass spectrometry method. This could explain why this protein was not detected in ToNV
223 OBs.

224 LEF-4 and P47 homologs were present in ToNV OBs. Although the presence of LEF-4 has
225 not been previously described as viral particle component, the presence of one or more of the
226 other RNA polymerase subunits is not unusual, as it has already been observed in
227 baculoviruses (Zhang *et al.*, 2015; Perera *et al.*, 2007) and hytrosaviruses (Kariithi *et al.*,
228 2010; Abd-Alla *et al.*, 2016).

229 Twenty-one core nudiviral proteins were detected in ToNV OBs (Fig. 2, in black). Seven
230 proteins (HzNVORF128, PmNVORF47, HzNVORF143, Cc50C22.6, Cc50C22.5,
231 PmNVORF7 and PmNVORF64) were shown for the first time to be present in OBs. Five
232 nudiviral proteins (PmNVORF62, PmV, HzNVorf106, HzNVORF64 and HzNVORF9) had
233 so far exclusively been characterized as particles component of endogenous nudiviruses
234 (Bézier *et al.*, 2009; Burke *et al.*, 2013; Pichon *et al.*, 2015). The 11K protein had been
235 detected in endogenous nudiviruses as well as HzNV-2 (Kim, 2009).

236 Lastly, the four additional proteins, GbNVORF19, ORF080, P47 and ORF122, seemed
237 relevant components of ToNV OBs, although below the filter thresholds. In particular, P47
238 need to be confirmed as its detection relies on a 100% peptide sequence identity based on a
239 sequence also found in Bacteria and Eukaryota (not shown). However, P47 and GbNVORF19
240 have already been identified as viral particle component of *Clostera anachoreta* granulovirus
241 (ClanGV) (Zhang *et al.*, 2015) and VcVLP (Pichon *et al.*, 2015) respectively, which suggests
242 that they are genuine particle components. ORF080 and ORF122 both appear specific to
243 ToNV and therefore are described as OB associated proteins for the first time.

244 Overall, from all these 52 identified proteins, 21 were encoded by nudivirus core genes, 11
245 had homologs in at least one exogenous or endogenous nudivirus and 20 were to date specific
246 to ToNV (Bézier *et al.*, 2015).

247

248 **Protein Abundance**

249 On the basis of the total spectrum count obtained by Scaffold software (Proteome software
250 Inc), the relative quantity of each OB protein can be estimated (Table 1 & Fig. 2) (Searle,
251 2010; Lundgren *et al.*, 2010). Among ToNV OB proteins, the most abundant protein is
252 encoded by *orf059*, the functional homolog of the baculovirus polyhedrin (Bézier *et al.*,
253 2015). MOBP represented about 65% of all identified spectrum counts and MOBP peptides
254 were found within all gel slices subjected to nanoLC-HR-MS/MS (not shown), including all
255 major bands. The identification of single proteins in several bands is not unusual as it was
256 already observed for the NPV and GV major structural proteins (Deng *et al.*, 2007; Perera *et*
257 *al.*, 2007; Wang *et al.*, 2010; Wang *et al.*, 2011b). Moreover, protein abundance estimated by
258 calculating the emPAI using Scaffold Q+ software 4.6.1 (Ishihama *et al.*, 2005) further
259 revealed that MOBP represented ~95% of the viral proteins present in the 25 kDa major band.
260 The second component forming this band was ORF023 (~4%).

261 The other main OB components were encoded by *orf087*, *orf028* and *orf019*, that are
262 respectively homologs to three HzNV-2 predicted proteins corresponding to the virion
263 structural proteins 34K (Hz2V052, the baculovirus capsid protein VP39 homolog) and 11K
264 (Hz2V025) (Kim, 2009), and to the hypothetical protein Hz2V079, which was previously
265 determined as a nudivirus core gene (Bézier *et al.*, 2015). Of these proteins, VP39 was
266 identified as the second major protein in ToNV OBs (~10% of all identified spectrum counts)
267 and also the second main viral component of the major band estimated at 35 kDa (~38%;
268 MOBP amount to ~61% of that band). VP39 was also identified as a highly abundant

269 structural protein in baculoviruses (Person *et al.*, 1988; Thiem & Miller, 1989) or exogenous
270 (Kim, 2009) and endogenous (Bézier *et al.*, 2009; Wetterwald *et al.*, 2010; Burke *et al.*, 2013)
271 nudiviruses. 11K is a nudiviral specific core protein. This protein was found in the 12 kDa
272 band in the same proportion (~0.5%) as VP39 and two other proteins with close estimated
273 molecular weight (ORF041, 8.7 kDa and ORF053, 10.7 kDa). The HzNV-2 11K homolog
274 was suspected as a possible viral envelope component (Kim, 2009), in contrast bracovirus
275 homologs (HzNVORF124-like) were not found as particle component (Wetterwald *et al.*,
276 2010; Burke *et al.*, 2013). PmNVORF62 (ORF019) was found in low proportion in both
277 remaining major bands, mainly in combination with MOBP and ORF023 in the 48 kDa band
278 (~0.2%) and with MOBP and VP39 in the 75 kDa band (~2%).

279 All the *per os* infectivity factors were present in ToNV OBs, but only two (b and c) out of the
280 three members of the PIF-5 ToNV multigenic family were found. PIFs were not highly
281 represented with 13 to 278 total spectra, totalling only ~2% of all identified viral peptides.
282 However such low abundance was also observed in baculoviruses (Faulkner *et al.*, 1997;
283 Braunagel & Summers, 2007; Kikhno *et al.*, 2002; Peng *et al.*, 2010), hytrosaviruses (Abd-
284 Alla *et al.*, 2016) and endogenous nudiviruses (Wetterwald *et al.*, 2010).

285 Eight proteins, encoded by *orf019* (PmNVORF62), *orf027* (GbnVORF19), *orf028* (11K),
286 *orf031* (PmV), *orf049* (HzNVORF106), *orf059* (MOBP), *orf112* (HzNVORF64) and *orf117*
287 (HzNVORF9), had homologs already identified as particle components with no other known
288 function in nudiviruses (Chaivisuthangkura *et al.*, 2008; Kim, 2009), bracoviruses (Bézier *et*
289 *al.*, 2009; Wetterwald *et al.*, 2010; Burke *et al.*, 2013) and VcVLP (Pichon *et al.*, 2015)
290 (Table 1 and Fig. 2). They represented 73.23% of all identified spectrum counts. Finally, 20
291 proteins did not have any known viral homolog and therefore did not have so far any known
292 function, except putative associated or viral structural proteins in ToNV OBs. Among them,
293 ORF034, belongs to the ToNV specific ORF032 multigenic family (Bézier *et al.*, 2015), from

294 which only two (ORF032 and ORF034) out of the three members were detected as OB
295 components. These 20 proteins specific to ToNV represented 12.90% of all identified
296 spectrum counts and about a sixth of all identified spectrum counts corresponding to
297 accessory gene products (not shown). *In fine*, from the 52 identified proteins, nudivirus core
298 genes and accessory genes represented respectively 20% and 80% of all identified spectrum
299 counts (not shown).

300

301 **Associated arthropod proteins**

302 Nine host-cell derived proteins were identified to be associated with ToNV OBs with high
303 confidence (over 99%). They mainly concerned cytoskeleton proteins, proteins involved in
304 DNA or mRNA processing, and potential oxidative stress response proteins (Table 2). Such
305 host proteins might have been co-purified with ToNV OBs and consequently represent
306 contaminants but they might also be intricately associated with the OBs or the viral particles
307 as it was shown for other phylogenetically related (Wang *et al.*, 2010; Hou *et al.*, 2013;
308 Braconi *et al.*, 2014) or unrelated (Ott, 2002; Shaw *et al.*, 2008) enveloped viruses. They
309 might therefore play a crucial role in the viral life cycle. Similarly *Helicoverpa armigera* NPV
310 (HearNPV) BV and ODV were shown to be associated with 101 and 21 host proteins,
311 respectively (Hou *et al.*, 2013). Furthermore, the presence of such cellular proteins is not
312 unusual, even among nudiviruses at large, as we found that ultra-deep purified particles
313 (sucrose and cesium chloride gradients) of the endogenous nudivirus, *Cotesia congregata*
314 bracovirus, contain heat shock proteins, ATP-synthases, ATPases, tubulin and actin
315 (unpublished data).

316 In particular, actin, beta tubulin, histone and heat shock protein 70 were identified in
317 baculovirus ODV (Lanier & Volkman, 1998; Wang *et al.*, 2010; Hou *et al.*, 2013; Braconi *et*
318 *al.*, 2014; Zhang *et al.*, 2015), herpesviruses (Shaw *et al.*, 2008), human immunodeficiency

319 virus (Saphire *et al.*, 2005) and other retrovirus (Segura *et al.*, 2008). During baculovirus
320 infection, cellular actin is directly involved in nucleocapsid transit between the cytoplasm and
321 the nucleus, in nucleocapsid assembly in the virogenic stroma, in their egress from the
322 nucleus and budding through the cytoplasmic membrane (Ohkawa *et al.*, 2010; Volkman,
323 2007; Marek *et al.*, 2011).

324 Cu-Zn superoxide dismutase (SOD) is involved in detoxifying superoxide free radicals
325 released during host defense reaction (Bannister *et al.*, 1987). Such enzyme mainly occurs in
326 eukaryota and bacteria but SODs have been convergently acquired in several virus families
327 (Thézé *et al.*, 2015). SODs were described as virion components in nucleocytoplasmic large
328 DNA viruses (NCLDV) such as vaccinia virus (Almazan *et al.*, 2001; Chung *et al.*, 2006) or
329 in the Megavirus chilensis (Lartigue *et al.*, 2015). *Sod* homologs were found in the genome of
330 entomopoxviruses (Thézé *et al.*, 2013) and almost all lepidopteran baculoviruses (Tomalski *et*
331 *al.*, 1991). They were described as associated with ODV in *Chrysodeixis chalcites* NPV (Xu
332 *et al.*, 2011) and HearNPV (Hou *et al.*, 2013). In such viral context, SODs are suspected to
333 counteract the host stress response thus increasing virus survival or to inhibit or regulate
334 cellular SOD function (Almazan *et al.*, 2001; Teoh *et al.*, 2003). Concerning ToNV, the insect
335 SOD protein might have been captured by the virus by chance, while infected cells contained
336 many cellular SODs, or by oriented way so as to play a similar protective role upon virus
337 infection.

338 Hrp59 is a nuclear protein that belongs to the heterogeneous nuclear ribonucleoprotein
339 (hnRNP) M type. Such proteins are evolutionarily conserved from yeast to human. They are
340 involved in packaging nascent transcripts into ribonucleoprotein complexes; they present gene
341 target specificity depending on the considered type (Krecic & Swanson, 1999; Dreyfuss *et al.*,
342 2002). Human M type homolog is implicated in premRNA splicing regulation under stress
343 conditions (Gattoni *et al.*, 1996). HnRNPs seem to be recurring targets for many RNA viruses

344 (Jorba *et al.*, 2008; Dechtawewat *et al.*, 2015; Jagdeo *et al.*, 2015), but to our knowledge
345 nothing is known about hnRNP association with DNA virus particles or hnRNP exploitation
346 by DNA viruses.

347 Other host proteins might also be associated with ToNV OBs that could not be identified
348 because the crane fly genome is not sequenced. Finally, to be confirmed as ToNV OB
349 components, immunogold labelling analyses of these proteins should be done. Although
350 cytoskeleton and structural host proteins could be involved in virus entry and other protein in
351 preventing stress response, the real implication of these host factors in nudivirus life cycle
352 remains largely unknown.

353

354 **CONCLUSIONS**

355

356 Comprehensive qualitative proteomic analysis using on-line nanoLC-HR-MS/MS revealed for
357 the first time exhaustive composition of nudivirus OBs. The occluded nudivirus infecting
358 *Tipula oleracea* contained 52 viral proteins and 9 host proteins. About half of the viral
359 proteins were already identified as particle components in other exogenous or endogenous
360 nudiviruses as well as in baculoviruses and hytrosaviruses. These mainly concern baculovirus
361 homologs involved in oral infectivity and packaging, assembly and release. Almost all host
362 proteins were previously identified as particle component in other enveloped viruses. Thus,
363 conservation of the main protein components through various viral entities, in particular
364 nudiviruses, both demonstrated cases of nudivirus endogenization (bracoviruses and VcVLP),
365 baculoviruses and hytrosaviruses, strengthens the idea of a similar basic cell entry mechanism
366 within insect. However, each virus retains specificity, as over half identified proteins are to
367 date specific for *Tipula oleracea* nudivirus. This study does not provide any information
368 regarding whether these proteins are associated with the nucleocapsids, with the envelopes, or

369 with the protein matrix of the *Tipula oleracea* nudivirus. To discriminate further between
370 these different compartments it would be necessary to treat the OBs with, for example,
371 Nonidet P-40, to compare SDS-Page profiles between treated and non-treated occluded
372 particles and to proceed to new protein sequencing. Immunogold strategy using candidate
373 proteins and scanning transmission electron microscopy could also reveal protein localization.

374

375 **MATERIALS AND METHODS**

376

377 **Viral sample.**

378 ToNV archival sample (*Nudivirus*, Unclassified) was originally obtained from the historical
379 insect virus collection held at the Natural Environment Research Council, Centre for Ecology
380 and Hydrology (NERC-CEH, Wallingford, England). The occluded particles solution used in
381 this study was the same as those used for previous genomic analyses (Bézier *et al.*, 2015).

382

383 **SDS-Page and gel section excision.**

384 Five microliters of purified virus (approximately 3 µg total proteins quantified using Qubit
385 Protein AssayKit, Life Technologies) were denatured 30 min at 95°C in 1x PBS and SDS-
386 PAGE sample (Laemmli, 1970) buffers then placed on ice few minutes. Sample was then
387 electrophoresed during 3h at 70V on a 12.5% polyacrylamide running gel combined to a 5%
388 stacking gel, along with Kaleidoscope molecular weight marker (Biorad). After staining with
389 Coomassie blue, major bands were first individually excised from polyacrylamide gel then the
390 remaining lane was also taken and divided in twenty slices so as to include all the present
391 components in the study.

392

393 **In-gel digestion.**

394 Each band was washed in water, acetonitrile (1:1) for 5 min followed by a second wash in
395 acetonitrile for 10 min. Cystein reduction and alkylation were performed by successive
396 incubations in solutions of 10 mM dithiothreitol in 50 mM sodium bicarbonate for 30 min at
397 56°C and 55 mM iodoacetamide in 50 mM sodium bicarbonate for 20 min at room
398 temperature in the dark, respectively. Gel slices were washed by incubation in 50 mM sodium
399 bicarbonate, acetonitrile (1:1) for 10 min followed by incubation in acetonitrile for 15 min.
400 Proteins were digested overnight in 25 mM sodium bicarbonate with 12.5 ng/μl trypsin
401 (Sequencing Grade, Roche). The resulting peptides were extracted from the gel using
402 incubation in 0.1% formic acid, acetonitrile (1:1) for 10 min followed by incubation for
403 5 min, in acetonitrile. The collected extractions were pooled, dried with a Speed-Vac
404 (Thermofisher), reconstituted with 12 μL of 0.1% formic acid, 2% acetonitrile, and sonicated
405 for 10 min.

406

407 **NanoLC-HR-MS/MS approach.**

408 Peptide mixtures were analysed by on-line nanoflow liquid chromatography tandem high
409 resolution mass spectrometry (nanoLC-HR-MS/MS). All experiments were performed on a
410 LTQ Orbitrap Velos Mass Spectrometer (Thermo Fisher Scientific) coupled to an Ultimate®
411 3000 RSLC Ultra High Pressure Liquid Chromatographer (Dionex) controlled by Chromeleon
412 Software (version 6.8 SR11; Dionex). Five microliters of each sample were injected and
413 loaded on an LCPackings trap column (Acclaim PepMap 100 C₁₈, 100 μm inner diameter x
414 2 cm long, 3 μm particles, 100 Å pores). Mobile phases consisted of (A) 0.1% formic acid,
415 97.9 % water, 2% acetonitrile (v/v/v) and (B) 0.1% formic acid, 15.9 % water, 84%
416 acetonitrile (v/v/v). Peptides were desalted and preconcentrated for 10 min at 5 μL/min with
417 4% solvent B. The peptide separation was conducted using a LCPackings nano-column
418 (Acclaim PepMap C₁₈, 75 μm inner diameter x 50 cm long, 3 μm particles, 100 Å pores). The

419 gradient consisted of 4 to 55% B for 90 min, 55 to 99% B for 1 min, constant 99% B 20 min
420 and return to 4% B in 1 min. The column was re-equilibrated for 15 min at 4% B between
421 runs. The nanoflow rate was set to 300 nl/min. Standard mass spectrometric conditions for all
422 experiments were spray voltage 1.2 kV, no sheath and auxiliary gas flow; heated capillary
423 temperature, 200°C; predictive automatic gain control (AGC) enabled, and an S-lens RF level
424 of 50%.

425 Data were acquired using Xcalibur software (version 2.1; Thermo Fisher Scientific). The LTQ
426 Orbitrap Velos instrument was operated in positive ion mode in data-dependent mode to
427 switch between high resolution full scan MS spectra collected in profile mode and low-
428 resolution MS/MS spectra in centroid mode. Resolution in the Orbitrap was set to $R = 60,000$.
429 In the scan range of m/z 300-1800, the 20 most intense peptide ions with charge states ≥ 2
430 were sequentially isolated (isolation width, 2 m/z ; 1 microscan) and fragmented in the high-
431 pressure linear ion trap by CID (Collision Induced Dissociation) with normalized collision
432 energy of 35% and wideband-activation enabled. Ion selection threshold was 500 counts for
433 MS/MS, and the maximum allowed ion accumulation times were 200 ms for full scans and
434 50 ms for CID-MS/MS measurements. An activation $q = 0.25$ and activation time of 10 ms
435 were used. Dynamic exclusion was activated during 30 s with a repeat count of 1. The lock
436 mass was enabled for accurate mass measurements. Polydimethylcyclsiloxane (m/z ,
437 445.1200025, $(\text{Si}(\text{CH}_3)_2\text{O})_6$) ions were used for internal recalibration of the mass spectra.

438

439 **Protein identification and data validation.**

440 Raw data files were converted to MGF (Mascot Generic Format) with Proteome Discoverer
441 software (version 1.3; Thermo Fischer Scientific). Precursor mass range of 350-5000 Da and
442 signal to noise ratio of 1.5 were the criteria used for generation of peak lists. In order to
443 identify the proteins, the peptide and fragment masses obtained were matched automatically

444 against a local database comprising all the ToNV predicted proteins, such a comparison with a
445 specific database was all the more necessary that overall nudiviral homologous sequences are
446 relatively distant from a virus to another. MS/MS ion searches were performed using
447 MASCOT Daemon and search engine (version 2.3; Matrix Science). The parameters used for
448 database searches include trypsin as a protease with allowed two missed cleavage,
449 carbamidomethylcysteine (+57 Da), oxidation of methionine (+16) and N-terminal protein
450 acetylation (+42) as variable modifications. The tolerance of the ions was set to 5 ppm for
451 parent and 0.8 Da for fragment ion matches. Mascot results obtained from the target and
452 decoy databases searches were incorporated to Scaffold software (version 3.6.4, Proteome
453 Software), a bioinformatic tool that integrates algorithms to validate identified peptides and
454 proteins by search engine and thus reduces number of false positives (Searle, 2010). Peptide
455 identifications were accepted if they could be established at greater than 95.0% probability as
456 specified by the Peptide Prophet algorithm (Keller *et al.*, 2002). Protein identifications were
457 accepted if they could be established at greater than 95.0% probability as specified by the
458 Protein Prophet algorithm (Nesvizhskii *et al.*, 2003). Supplementary peptides with a 95.0%
459 peptide probability and a 50 to 79% protein probability were also examined and validated
460 manually against the non-redundant protein sequences using blastp (NCBI).

461

462 **ACKNOWLEDGEMENTS**

463

464 This study was funded by the European Research Council starting grant GENOVIR (205206).
465 The high-resolution mass spectrometer was financed (SMHART project) by the European
466 Regional Development Fund (ERDF), the Conseil Régional du Centre, the French National
467 Institute for Agricultural Research (INRA) and the French National Institute of Health and
468 Medical Research (Inserm). Lucie Combes-Soia is thanked for helping with data analysis.

469

470 **AUTHORS' CONTRIBUTIONS**

471

472 AB performed and supervised the experiments and wrote the manuscript. GH and VL
473 performed the proteomic analyses and contributed to the writing of the manuscript. KM
474 realized the SDS-Page gel. EAH directed the study and contributed to the writing of the
475 manuscript. All authors have read and approved the final manuscript prior submission.

476

477 **REFERENCES**

478

- 479 **Abd-Alla, A. M., Kariithi H. M., Cousserans, F., Parker, N. J., Ince, I. A., Scully, E. D.,**
480 **Boeren, S., Geib, S. M., Mekonnen, S., Vlak, J. M., Parker, A. G., Vreysen, M. J. &**
481 **Bergoin, M. (2016).** Comprehensive Annotation of the *Glossina pallidipes* Salivary Gland
482 Hypertrophy Virus from Ethiopian Tsetse Flies: A proteogenomics Approach. *J Gen Virol* **97**,
483 1010-1031.
- 484 **Almazan, F., Tschärke, D. C., & Smith, G. L. (2001).** The vaccinia virus superoxide
485 dismutase-like protein (A45R) is a virion component that is nonessential for virus replication.
486 *J Virol* **75**, 7018-7029.
- 487 **Balakrishnan L. & Bambara, R. A. (2013).** Flap Endonuclease 1. *Annu Rev Biochem* **82**,
488 119-138.
- 489 **Bannister, J. V., Bannister, W. H. & Rotilio, G. (1987).** Aspects of the structure, function,
490 and applications of superoxide dismutase. *CRC Crit Rev Biochem* **22**, 111-180.
- 491 **Bergoin, M. & Guelpa, B. (1977).** Dissolution des inclusions du virus de la polyhedrose
492 nucléaire du diptère *Tipula paludosa* MEIG. Etude ultrastructurale du virion. *Arch Virol* **53**,
493 243-254.

494 **Bézier, A., Annaheim, M., Herbinière, J., Weraintterwald, C., Gyapay, G., Bernard-**
495 **Samain, S., Wincker, P., Roditi, I., Heller, M., Belghazi, M., Pfister-Wilhem, R.,**
496 **Periquet, G., Dupuy, C., Huguet, E., Volkoff, A. N., Lanzrein, B. & Drezen J-M. (2009).**
497 Polydnviruses of braconid wasps derive from an ancestral nudivirus. *Science* **323**, 926-930.
498 **Bézier, A., Thézé, J., Gavory, F., Gaillard, J., Poulain, J., Drezen, J-M. & Herniou, E. A.**
499 **(2015).** The genome of the nucleopolyhedrosis-causing virus from *Tipula oleracea* sheds new
500 light on the *Nudiviridae* family. *J Virol* **89**, 3008-3025.
501 **Bonami, J-R., Bruce, L. D., Poulos, B. T., Mari, J. & Lightner, D. V. (1995).** Partial
502 characterization and cloning of the genome of PvSNPV (= BP-type virus) pathogenic for
503 *Penaeus vannamei*. *Dis Aqua Org* **23**, 59-66.
504 **Bonami, J-R., Aubert, H., Mari, J., Poulos, B. T. & Lightner D. V. (1997).** The polyhedral
505 of the occluded baculoviruses of marine decapod crustacean: a unique structure, crystal
506 organization, and proposed model. *J Struct Biol* **120**, 134-145.
507 **Braconi, C. T., Ardisson-Araujo, D. M., Leme, A. F., Oliveira, J. V., Pauletti, B. A.,**
508 **Garcia-Maruniak, A., Ribeiro, B. M., Maruniak, J. E. & Zanotto, P. M. (2014).**
509 Proteomic analyses of baculovirus *Anticarsia gemmatalis* multiple nucleopolyhedrovirus
510 budded and occluded virus. *J Gen Virol* **95**, 980-989.
511 **Braunagel, S. C., Russel, W. K., Rosas-Acosta, G., Russel, D. H. & Summers, M. D.**
512 **(2003).** Determination of the protein composition of the occlusion-derived virus *Autographa*
513 *californica* nucleopolyhedrovirus. *PNAS* **100**, 9797-9802.
514 **Braunagel, S. C. & Summers, M. D. (2007).** Molecular biology of the baculovirus
515 occlusion-derived virus envelope. *Curr Drug Targets* **8**, 1084-1095.
516 **Burand, J. P., Stiles, B. & Wood, H. A. (1983).** Structural and intracellular proteins of the
517 nonoccluded baculovirus Hz-1. *J Virol* **46**, 137-142.

518 **Burand, J. P. (1998).** Nudiviruses. In *The Insect Viruses*, pp. 69-90. Edited by L. K. Miller &
519 L. A. Ball. New York and London, FL: Plenum Press.

520 **Burand, J. P. & Rallis, C. P. (2004).** *In vivo* dose-response of insects to Hz-2V infection.
521 *Virology* **15**, 15.

522 **Burand, J. P., Kim, W., Afonso, C. L., Tulman, E. R., Kutish, G. F., Lu, Z. & Rock, D.
523 L. (2012).** Analysis of the genome of the sexually transmitted insect virus *Helicoverpa zea*
524 nudivirus 2. *Viruses* **4**, 28-61.

525 **Burke, G. R., Thomas, S. A., Eum, J. H. & Strand, M. R. (2013).** Mutualistic
526 polydnviruses share essential replication gene functions with pathogenic ancestors. *Plos*
527 *Pathog* **9**, e1003348.

528 **Chaivisuthangkura, P., Tawilert, C., Tejangkura, T., Rukpratanporn, S., Longyant, S.,
529 Sithigorngul, W. & Sithigorngul, P. (2008).** Molecular isolation and characterization of a
530 novel occlusion body protein gene from *Penaeus monodon* nucleopolyhedrovirus. *Virology*
531 **381**, 261-267.

532 **Cheng, C. H., Liu, S. M., Chow, T. Y., Hsiao, Y. Y., Wang, D. P., Huang, J. J. & Chen H.
533 H. (2002).** Analysis of the complete genome sequence of the Hz-1 virus suggests that it is
534 related to members of the *Baculoviridae*. *J Virol* **76**, 9024-9034.

535 **Cheng, R-L., Lou, Y-H., Gu, L-Z., Wang, Z., Xu, J-Y., Xu, H-J. & Zhang, C-X. (2014).**
536 The brown planthopper nudivirus DNA integrated in its host genome. *J Virol* **88**, 5310-5318.

537 **Chung, C-S., Chen, C-H., Ho, M-Y., Huang, C-Y., Liao, C-L. & Chang W. (2006).**
538 Vaccinia virus proteome: identification of proteins in vaccinia virus intracellular mature virion
539 particles. *J Virol* **80**, 2127-2140.

540 **Cottrell, J. S. (2011).** Protein identifications using MS/MS data. *J Proteomics* **74**, 1842-1851.

541 **Couch, J. A. (1974).** Free and occluded virus, similar to *Baculovirus*, in hepatopancreas of
542 pink shrimp. *Nature* **247**, 229-231.

543 **Crawford, A. M. & Sheehan, C. (1985).** Replication of *Oryctes baculovirus* in cell culture:
544 Viral morphogenesis, infectivity and protein synthesis. *J Gen Virol* **66**, 529-539.

545 **Dechtawewat, T., Songprakhon, P., Limjindaporn, T., Puttikhunt, C., Kasinrerak, W.,**
546 **Saitornuang, S., Yenchitsomanus P-T. & Noisakran, S. (2015).** Role of human
547 heterogeneous nuclear ribonucleoprotein C1/C2 in dengue virus replication. *Virology* **6**, 12-14.

548 **Deng, F., Wang, R., Fang, M., Jiang, Y., Xu, X., Wang, H., Chen, X., Arif, B. M., Guo,**
549 **L., Wang, H. & Hu, Z. (2007).** Proteomics analysis of *Helicoverpa armigera* single
550 nucleocapsid nucleopolyhedrovirus identified two new occlusion-derived virus-associated
551 proteins, HA44 and HA100. *J Virol* **81**, 9377-9385.

552 **Dreyfuss, G., Kim, V. N. & Kataoka, N. (2002).** Messenger-RNA-binding proteins and the
553 messages they carry. *Nat Rev Mol Cell Biol* **3**, 195-205.

554 **Faulkner, P., Kuzio, J. G., Williams, V. & Wilson, J. A. (1997).** Analysis of p74, a PDV
555 envelope protein of *Autographa californica* nucleopolyhedrovirus required for occlusion body
556 infectivity *in vivo*. *J Gen Virol* **78**, 3091-3100.

557 **Garcia-Maruniak, A., Maruniak, J. E., Farmerie, W. & Boucias, D. G. (2008).** Sequence
558 analysis of a non-classified, non-occluded DNA virus that causes salivary gland hypertrophy
559 of *Musa domestica*, MdSGHV. *Virology* **377**, 184-196.

560 **Gattoni, R., Mahé, D., Mähl, P., Fischer, N., Mattei, M-G., Stévenin, J. & Fuchs, J-P.**
561 **(1996).** The human hnRNP-M proteins: structure and relation with early heat shock-induced
562 splicing arrest and chromosome mapping. *Nucleic Acids Res* **24**, 2535-2542.

563 **Gauthier, L., Cornman, S, Hartmann, U, Cousserans, F, Evans, J. E., de Miranda, J. R.**
564 **& Neumann P. (2015).** The *Apis mellifera* filamentous virus genome. *Viruses* **7**, 3798-3815.

565 **Grainge, I. & Jayaram, M. (1999).** The Integrase family of recombinase: organization and
566 function of the active site. *Mol Microbiol* **33**, 449-456.

567 **Granados, R. R., Nguyen, T. & Cato, B. (1978).** An insect cell line persistently infected
568 with baculovirus-like particle. *Intervirology* **10**, 309-317.

569 **Hou, D., Zhang, L., Deng, F., Fang, W., Wang, R., Liu, X., Guo, L., Rayner, S., Chen, X.,**
570 **Wang, H. & Hu Z. (2013).** Comparative proteomics reveal fundamental structural and
571 functional differences between the two progeny phenotypes of a baculovirus. *J Virol* **87**, 829-
572 839.

573 **Huger, A. M. (1966).** A virus disease of the Indian rhinoceros beetle, *Oryctes rhinoceros*
574 (Linnaeus), caused by a new type of insect virus, *Rhabdionvirus oryctes* gen. n., sp. n. *J*
575 *Invertebr Pathol* **8**, 38-51.

576 **Huger, A. M. (1985).** A new virus disease of crickets (Orthoptera: Gryllidae) causing
577 macronucleosis of fatbody. *J Invertebr Pathol* **45**, 108-111.

578 **Huger, A. M. & Krieg, A. (1991).** *Baculoviridae*. Nonoccluded baculoviruses. In *Atlas of*
579 *Invertebrate Viruses*, pp. 287-319. Edited by J. R. Adams & J.-R. Bonami. Boca Raton, FL:
580 CRC Press.

581 **Ishihama, Y. (2005).** Proteomic LC-MS systems using nanoscale liquid chromatography with
582 tandem mass spectrometry. *J Chromatogr A* **1067**, 73-83.

583 **Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J. & Mann M.**
584 **(2005).** Exponentially modified protein abundance index (emPAI) for estimation of absolute
585 protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell*
586 *Proteomics* **4**, 1265-1272.

587 **Jagdeo, J. M., Dufour, A., Fung, G., Luo, H., Kleifeld, O., Overall, C. M. & Jan, E.**
588 **(2015).** Heterogenous nuclear ribonucleoprotein M facilitates enterovirus infection. *J Virol*
589 **89**, 7064-7078.

590 **Jehle, J. A., Abd-Alla, A. M. M. & Wang, Y. (2013a).** Phylogeny and evolution of
591 *Hytrosaviridae*. *J Invertebr Pathol* **112**, S62-S67.

592 **Jehle, J. A., Burand, J., Herniou, E. A., Harrison, R., Arif, B., Thielmann, D., van Oers,**
593 **M. & Becnel, J. (2013b).** Creation of a new Family *Nudiviridae* including two new genera
594 and three species. Taxonomy Proposals
595 [talk.ictvonline.org/files/proposals/taxonomy_proposals_invertebrate1/m/inv04/4770.aspx].
596 **Jorba, N., Juarez, S., Torreira, E., Gastaminza, P., Zammerreño, N., Albar, J. P. & Ortin,**
597 **J. (2008).** Analysis of the interaction of influenza virus polymerase complex with human cell
598 factors. *Proteomics* **8**, 2077-2088.
599 **Kariithi, H. M., Ince, I. A., Boeren, S., Vervoort, J., Bergoin, M., van Oers, M. M., Abd-**
600 **Alla A. M. M. & Vlak, J. M. (2010).** Proteomic analysis of *Glossina pallidipes* salivary
601 gland hypertrophy virus virions for immune intervention in tsetse fly colonies. *J Gen Virol* **91**,
602 3065-3074.
603 **Keller, A., Nesvizhskii, A. I., Kolker, E. & Aebersold, R. (2002).** Empirical statistical
604 model to estimate the accuracy of peptide identifications made by MS/MS and database
605 search. *Anal Chem* **74**, 5383-5392.
606 **Kikhno, I., Gutierrez, S., Croizier, L., Croizier, G. & Lopez-Ferber, M. (2002).**
607 Characterization of *pif*, a gene required for the per os infectivity of *Spodoptera littoralis*
608 nucleopolyhedrovirus. *J Gen Virol.* **83**, 3013-3022.
609 **Kim, W. (2009).** Characterization of genome and structural proteins of the sexually
610 transmitted insect virus, Hz-2V. Doctoral Dissertations. Paper AAI3379978.
611 **Krecic, A. M. & Swanson, M. S. (1999).** hnRNP complexes: composition, structure, and
612 function. *Curr Opin Cell Biol* **11**, 363-371.
613 **Laemmli, U. K. (1970).** Cleavage of structural proteins during the assembly of the head of
614 bacteriophage T4. *Nature* **227**, 680-685.
615 **Lanier, L. M. & Volkman, L. E. (1998).** Actin binding and nucleation by *Autographa*
616 *california* M nucleopolyhedrovirus. *Virology* **243**, 167-177.

617 **Lartigue, A., Burlat, B., Coutard, B., Chaspoul, F., Claverie, J-M. & Abergel, C. (2015).**
618 The megavirus chilensis Cu,Zn-superoxide dismutase : the first viral structure of a typical
619 cellular copper chaperone-independent hyperstable dimeric enzyme. *J Virol* **89**, 824-832.

620 **Li, Y., Wang, J., Deng, R., Zhang, Q., Yang, K. & Wang, X. (2005).** *Vlf-1* deletion brought
621 AcMNPV to defect in nucleocapsid formation. *Virus Genes* **31**, 275-284.

622 **Lin, C., Lee, J. C., Chen, S. S., Wood, H. A., Li, M. L. & Chao, Y. C. (1999).** Persistent
623 Hz-1 virus infection in insect cells: evidence for insertion of viral DNA into host
624 chromosomes and viral infection in a latent status. *J Virol* **73**, 128-139.

625 **Liu, Y., Kao H. I. & Bambara, R. A. (2004).** Flap endonuclease 1: a central component of
626 DNA metabolism. *Annu Rev Biochem* **73**, 589-615.

627 **Liu, X., Chen, K., Cai, C. & Yao, Q. (2008).** Determination of protein composition and
628 host-derived proteins of *Bombyx mori* nucleopolyhedrovirus by 2-dimensional electrophoresis
629 and mass spectrometry. *Intervirology* **51**, 369-376.

630 **Lu, H. (1997).** Characterization of a novel baculovirus, gonad-specific virus, GSV. Thesis
631 dissertation. Department of Microbiology, University of Massachusetts, Amherst, MA.

632 **Lundgren, D. H., Hwang, S. I., Wu, L. & Han, D. K. (2010).** Role of spectral counting in
633 quantitative proteomics. *Expert Rev Proteomics* **7**, 39-53.

634 **Mari, J., Bonami, J-R., Poulos, B. & Lightner, D. (1993).** Preliminary characterization and
635 partial cloning of the genome of a baculovirus from *Penaeus monodon* (PmSNPV = MBV).
636 *Dis Aqua Org* **16**, 207-215.

637 **Marek, M., Merten, O. W., Galibert, L., Vlak, J. M. & van Oers, M. M. (2011).**
638 Baculovirus VP80 protein and the F-actin cytoskeleton interact and connect the viral
639 replication factory with the nuclear periphery. *J Virol* **85**, 5350-5362.

640 **Meynardier, G., Ricou, G. & Bergoin, M. (1964).** Virose à corps d'inclusion chez *Tipula*
641 *paludosa* (Diptera) en France. *Re Pathol Vég Entomo Agric Fr* **43**, 113-118.

642 **Nadala, E. C. B., Tapay, L. M. & Loh, P. C. (1998).** Characterization of a non-occluded
643 baculovirus-like agent pathogenic to penaeid shrimp. *Dis Aqua Org* **33**, 221-229.

644 **Nesvizhskii A. I., Keller A., Kolker E. & Aebersold R. A. (2003).** Statistical model for
645 identifying proteins by tandem mass spectrometry. *Anal Chem.* **75**, 4646-4658.

646 **Ohkawa, T., Volkman, L. E. & Welch, M. D. (2010).** Actin-based motility drives
647 baculovirus transit to the nucleus and cell surface. *J Cell Biol* **190**, 187-195.

648 **Ott, D. E. (2002).** Potential roles of cellular proteins in HIV-1. *Rev Med Virol* **12**, 359-374.

649 **Payne, C. C. (1974).** The isolation and characterization of a virus from *Oryctes rhinoceros*. *J*
650 *Gen Virol* **25**, 105-116.

651 **Payne, C. C., Compson, D. & de Looze S. M. (1977).** Properties of the nucleocapsids of a
652 virus isolated from *Oryctes rhinoceros*. *Virology* **77**, 269-280.

653 **Pearson, M. N., Russell, R. L. Q., Rohrmann, G. F. & Beaudreau, G. S. (1988).** P39, a
654 major baculovirus structural protein: Immunocytochemical characterization and genetic
655 location. *Virology* **167**, 407-413.

656 **Peng, K., van Oers, M. M., Hu, Z., van Lent J. W. M. & Vlak J. M. (2010).** Baculovirus
657 *per os* infectivity factors form a complex on the surface of occlusion-derived virus. *J Virol* **84**,
658 9497-9504.

659 **Peng K., van Lent, J. W. M., Boeren, S., Fang, M., Theilman, D. A., Erlandson M. A.,**
660 **Vlak, J. M. & van Oears, M. M. (2012).** Characterization of novel components of the
661 baculovirus *per os* infectivity factor complex. *J Virol* **86**, 4981-4988.

662 **Perera, O. P., Green, T. B., Stevens, S. M., White, S. & Becnel, J. J. (2007).** Proteins
663 associated with *Culex nigripalpus* nucleopolyhedrovirus occluded virions. *J Virol* **81**, 4585-
664 4590.

665 **Pichon, A., Bézier, A., Urbach, S., Aury J-M., Jouan, V., Ravvalet, M., Guy, J.,**
666 **Cousserans, F., Thézé, J., Gauthier, J., Demettre, E., Schmieder, S., Wurmser, F., Sibut,**

667 **V., Poirié, M., Colinet, D., de Silva, C., Couloux, A., Barbe, V., Drezen, J-M. & Volkoff,**
668 **A-N. (2015).** Recurrent DNA virus domestication leading to different parasite virulence
669 strategies. *Sci Adv* **1**, e1501150.

670 **Raina, A. K., Adams, J. R., Lupiani, B., Lynn, D. E., Kim, W., Burand, J. P. &**
671 **Dougherty, E. M. (2000).** Further characterization of the Gonad-Specific Virus of corn
672 eatworm, *Helicoverpa zea*. *J Inverteb Pathol* **76**, 6-12.

673 **Rieux, L., Sneekes, E-J. & Swart, R. (2011).** Nano LC: principles, evolution, and state-of-
674 the-art of the technique. *LCGC NA* **29**, 926–934, [Online]
675 [http://www.chromatographyonline.com/lcgc/Column%3A+Innovations+in+HPLC/Nano-LC-](http://www.chromatographyonline.com/lcgc/Column%3A+Innovations+in+HPLC/Nano-LC-Principles-Evolution-and-State-of-the-Art-/ArticleStandard/Article/detail/745381)
676 [Principles-Evolution-and-State-of-the-Art-/ArticleStandard/Article/detail/745381.](http://www.chromatographyonline.com/lcgc/Column%3A+Innovations+in+HPLC/Nano-LC-Principles-Evolution-and-State-of-the-Art-/ArticleStandard/Article/detail/745381)

677 **Rohrmann, G. F. (2013).** *Baculovirus Molecular Biology*, 3rd edn. Bethesda (MD): National
678 Library of Medicine (US), National Center for Biotechnology Information.
679 [<http://www.ncbi.nlm.nih.gov/books/NBK114593/>].

680 **Saphire, A. C. S., Gallay, P. A. & Bark, S. J. (2005).** Proteomic analysis of human
681 immunodeficiency virus using liquid chromatography/tandem mass spectrometry effectively
682 distinguishes specific incorporated host proteins. *J Proteome Res* **5**, 530-538.

683 **Searle, B. C. (2010).** Scaffold: A bioinformatic tool for validating MS/MS-based proteomic
684 studies. *Proteomics* **10**, 1265-1269.

685 **Segura, M. M., Garnier, A., Di Falco, M. R., Whissell, G., Meneses-Acosta, A. Arcand,**
686 **N. & Kamen, A. (2008).** Identification of host proteins associated with retroviral vector
687 particles by proteomic analysis of highly purified vector preparations. *J Virol* **82**, 1107-1117.

688 **Shaw, M. L., Stone, K. L., Colangelo, C. M., Gulcicek, E. E. & Palese, P. (2008).** Cellular
689 proteins in influenza virus particles. *PLoS Pathog* **4**, e1000085.

690 **Slack, J. & Arif, B. M. (2007).** The Baculoviruses Occlusion-Derived Virus: Virion
691 Structure and Function. *Adv Virus Res* **69**, 99-165.

692 **Smith, K. M. & Xeros, N. (1954).** An unusual virus disease of a dipterous larva. *Nature* **173**,
693 866-867.

694 **Stentiford, G. D., Bateman, K. & Feist, S. W. (2004).** Pathology and ultrastructure of an
695 intranuclear bacilliform virus (IBV) infecting brown shrimp *Crangon crangon* (Decapoda:
696 Crangonidae). *Dis Aqua Org* **58**, 89-97.

697 **Teoh, M. L., Walasek, P. J. & Evans, D. H. (2003).** *Leporipoxvirus* Cu,Zn-superoxide
698 dismutase (SOD) homologs are catalytically inert decoy proteins that bind copper chaperone
699 for SOD. *J Biol Chem* **278**, 33175-33184.

700 **Thézé, J., Bézier, A., Periquet, G., Drezen, J-M. & Herniou, E. A. (2011).** Paleozoic origin
701 of insect large dsDNA viruses. *PNAS* **108**, 15931-15935.

702 **Thézé, J., Tajatsuka, J., Li, Z., Gallais, J., Doucet, D., Arif, B., Nakai, M. & Herniou E.**
703 **A. (2013).** New insights into the evolution of Entomopoxvirinae from the complete genome
704 sequences of four entomopoxviruses infecting *Adoxophyes honmai*, *Choristoneura biennis*,
705 *Choristoneura rosaceana*, and *Mythimna separata*. *J Virol* **87**, 7992-8003.

706 **Thézé, J., Takatsuka, J., Nakai, M., Arif, B. & Herniou, E. A. (2015).** Gene Acquisition
707 Convergence between Entomopoxviruses and Baculoviruses. *Viruses* **7**, 1960-1974.

708 **Thiem, S. M. & Miller, L. K. (1989).** Identification, sequence, and transcriptional mapping
709 of the major capsid protein gene of the baculovirus *Autographa californica* nuclear
710 polyhedrosis virus. *J Virol* **63**, 2008-2018.

711 **Tomalski, M. D., Eldridge, R. & Miller, L. K. (1991).** A baculovirus homolog of a Cu/Zn
712 superoxide dismutase gene. *Virology* **184**, 149-161.

713 **Tweeten, K. A., L. A. Bulla., J. R. & Consigli, R. A. (1980).** Characterization of an
714 extremely basic protein derived from granulosis virus nucleocapsids. *J Virol* **33**, 866-876.

715 **Unkless, R. L. (2011).** A DNA virus of *Drosophila*. *PloS One* **6**, e26564.

716 **van Oers, M. M. & Vlak, J. M. (2007).** Baculovirus Genomics. *Curr Drug Targets* **8**, 1051-
717 1068.

718 **Vanarsdall, A. L., Okano, K. & Rohrmann, G. F. (2004).** Characterization of a baculovirus
719 with a deletion of *vlf-1*. *Virology* **326**, 191-201.

720 **Vanarsdall, A. L., Okano, K. & Rohrmann, G. F. (2006).** Characterization of the role of
721 VLF-1 in baculovirus capsid structure and DNA processing. *J Virol* **80**, 1724-1733.

722 **Volkman, L. E. (2007).** Baculovirus infectivity and the actin cytoskeleton. *Curr. Drug*
723 *Targets* **8**, 1075-1083.

724 **Wang, R., Deng, F., Hou, D., Zhao, Y., Guo, L., Wang, H. & Hu, Z. (2010).** Proteomics of
725 the *Autographa californica* nucleopolyhedrovirus budded virions. *J Virol* **84**, 7233-7242.

726 **Wang, X-F., Zhang, B-Q., Xu, H-J. Cui, Y-J. Xu, Y-P., Zhang, M-J., Han, Y. S., Lee, Y.**
727 **S., Bao, Y-Y. & Zhang, C-X. (2011b).** ODV-associated proteins of the *Pieris rapae*
728 granulovirus. *J Proteome Res* **10**, 2817-2827.

729 **Wang, Y., Kleespies, R. G., Huger, A. M. & Jehle, J. A. (2007a).** The genome of *Gryllus*
730 *bimaculatus* nudivirus indicates an ancient diversification of baculovirus-related non-
731 occluded nudiviruses of insects. *J Virol* **81**, 5395-5406.

732 **Wang, Y., van Oers, M. M., Crawford, A. M., Vlak, J. M. & Jehle, J. A. (2007b).**
733 Genomic analysis of *Oryctes rhinoceros* virus reveals genetic relatedness to *Heliothis zea*
734 virus 1. *Arch Virol* **152**, 519-531.

735 **Wang, Y., Bininda-Emonds, O. R. P., van Oers, M. M., Vlak, J. M., Jehle & J. A.**
736 **(2011a).** The genome of *Oryctes rhinoceros* nudivirus provides novel insight into the
737 evolution of nuclear arthropod-specific large circular double-stranded DNA viruses. *Virus*
738 *Genes* **42**, 444-456.

739 **Wang, Y., Bininda-Emond, O. R. P. & Jehle, J. A. (2012).** Nudivirus genomics and
740 phylogeny. In *Viral Genomes - Molecular Structure, Diversity, Gene Expression Mechanisms*

741 *and Host-Virus Interactions*, pp. 33-52. Edited by M. L. Garcia & V. Romanowski.
742 [[http://www.intechopen.com/books/viral-genomesmolecular-structure-diversity-gene-](http://www.intechopen.com/books/viral-genomesmolecular-structure-diversity-gene-expression-mechanisms-and-host-virusinteractions/nudivirus-genomics-and-phylogeny)
743 [expression-mechanisms-and-host-virusinteractions/nudivirus-genomics-and-phylogeny](http://www.intechopen.com/books/viral-genomesmolecular-structure-diversity-gene-expression-mechanisms-and-host-virusinteractions/nudivirus-genomics-and-phylogeny)].
744 **Wetterwald, C., Roth, T., Kaeslin, M., Annaheim, M., Wespi, G., Heller, M., Mäser, P.,**
745 **Roditi, I., Pfister-Wilhelm, R., Bézier, A., Gyapay, G., Drezen, J-M. & Lanzrein, B.**
746 **(2010)**. Identification of bracovirus particle proteins and analysis of their transcript levels at
747 the stage of virion formation. *J Gen Virol* **91**, 2610-2619.
748 **Yang, Y. T., Lee, D. Y., Wang, Y., Hu, J. M., Li, W. H., Leu, J. H., Chang, G. D., Ke, H.**
749 **M., Kang, S. T., Lin, S. S., Kou, G. H. & Lo, C. F. (2014)**. The genome and occlusion
750 bodies of marine *Penaeus monodon* nudivirus (PmNV, also known as MBV and PemoNPV)
751 suggest that it should be assigned to a new nudivirus genus that is distinct from the terrestrial
752 nudiviruses. *BMC Genomics* **15**, 628.
753 **Xu, F., Ince I. A., Boeren, S., Vlack, J. M. & van Oers M. M. (2011)**. Protein composition
754 of the occlusion derived virus of *Chrysodeixis chalcites* nucleopolyhedrovirus. *Virus Res* **158**,
755 1-7.
756 **Zhang, X., Liang, Z., Yin, X. & Shao, X. (2015)**. Proteomic analysis of the occluded-
757 derived virus *Chlostera anachoreta* granulovirus. *J Gen Virol* **96**, 2394-2404.
758 **Zhu, S., Wang, W., Wang, Y., Yuan, M. & Yang, K. (2013)**. The baculovirus core gene
759 *ac83* is required for nucleocapsid assembly and *per os* infectivity of *Autographa californica*
760 nucleopolyhedrovirus. *J Virol* **87**, 10573-10586.
761

762 **TABLES**

763 **Table 1. The ToNV major and associated OB proteins identified by nanoLC-HR-MS/MS ranked by decreasing predicted molecular**
 764 **weight.** The four main major components, with known homologs in other nudiviruses, are highlighted in gray, with gray intensity based on
 765 normalized total spectra count. Homologous proteins were those present (+) or identified (●) as particle component for at least one
 766 representative of the referenced viral families.

ORF #	Protein designation	Length (aa)	MW (kDa)	pI	#unique peptide	Protein coverage (%)	Total spectrum count	Present (+) or identified (●) as particle component in				
								NPV - GV	NV	SGHV	BV	VcVLP
125	ORF125	808	94.99	5.02	46	49	857					
130	ORF130	765	88.28	4.92	36	50	415					
45	P74	763	87.20	9.91	8	13	40	●	+	●	●	●
16	VP91/95	742	86.95	5.33	15	25	111	●	+		●	●
57	Cc50C22.5	690	81.88	5.13	15	26	107		+		+	
108	ORF108	704	81.16	5.89	47	45	904					
19	PmNVORF62	597	68.85	9.13	42	66	2442		+			●
4	HzNVORF128	561	66.18	8.55	3	6	8		+		+	
69	PIF-1	527	60.04	8.01	10	21	101	●	+		●	●
25	LEF-4*	494	58.96	4.86	1	2	8	+	+	+	+	+
120	ORF120	497	57.75	6.00	3	7	5					
1	FEN-1	479	56.85	8.43	17	38	92		+	+	+	
102	PIF-5c	450	51.94	4.22	17	34	262	●	+	●	●	●
112	HzNVORF64	408	47.81	4.64	5	15	12		+		●	
84	ORF084*	421	47.74	5.60	1	3	4					
80	ORF080**	388	45.23	4.10	1	3	2					
65	VLF-1	385	44.81	5.93	6	22	15	●	+		●	
15	PmNVORF47	364	43.40	4.89	10	26	285		+			
96	PIF-5b	398	43.27	8.11	15	30	278	●	+	●	●	●
7	PIF-2	362	40.96	6.34	14	47	148	●	+	●	●	●

107	ORF107	336	39.58	7.76	16	45	606					
43	INT*	331	39.15	8.31	1	4	3		+	●	●	
115	P47**	333	39.09	7.75	1	2	20	●	+		+	+
119	PIF-4	327	38.31	5.76	14	48	88	●	+		●	●
87	VP39	324	37.77	6.19	47	78	4999	●	●		●	
49	HzNVORF106	307	36.30	4.88	13	32	74		+		●	
63	38K	291	35.02	9.03	7	19	25	●	+		●	
99	P33	289	34.70	8.85	3	11	30	●	+		●	●
27	GbNVORF19**	291	34.12	8.35	1	3	2		+		+	●
31	PmV	272	31.78	5.75	13	44	200		+		●	
54	Cc50C22.6	254	30.124	8.51	8	26	169				+	
59	MOBP	241	27.36	6.31	44	87	32188	●	●			
117	HzNVORF9	237	27.31	4.82	8	34	40		+		●	
26	ORF026*	226	27.00	4.97	1	7	15					
62	ORF062*	224	26.31	6.43	1	7	5					
23	ORF023	225	26.20	6.19	19	75	894					
106	ORF106	218	25.29	9.82	13	59	651					
89	ORF089	203	23.37	4.88	10	52	36					
32	ORF032	205	22.47	6.14	10	38	566					
13	PIF-3	193	21.86	4.83	6	25	43	●	+	●	●	●
124	ORF124	180	21.27	8.24	7	34	138					
34	ORF034	188	20.64	7.16	5	24	1139					
56	PIF-6	143	16.81	5.60	2	15	13	●	+		+	●
30	ORF030*	144	16.77	7.25	1	9	3					
114	PmNVORF64*	139	15.76	4.36	1	7	1		+			
122	ORF122**	133	15.68	9.92	1	7	1					
42	ORF042*	105	12.65	9.70	1	21	14					
28	11K	115	12.39	10.13	9	44	1325		●		+	●
100	PmNVORF7	92	10.73	6.41	3	23	35		+			
53	ORF053	93	10.70	9.38	2	41	11					
41	ORF041	74	8.70	9.37	5	49	107					

52 HzNVORF143 53 6.53 7.76 3 51 11 + +767

768

769 NPV, nucleopolyhedrovirus; GV, granulovirus; NV, nudivirus; SGHV, salivary gland hypertrophy virus; BV, bracovirus; VcVLP, *Venturia*
 770 *canescens* virus-like particles. MW and pI, predicted molecular weight and Isoelectric point, respectively. *, proteins unambiguously identified
 771 with a unique peptide and a minimal protein threshold at 95%; **, proteins unambiguously identified with a unique peptide, a protein threshold
 772 comprised between 50 and 94% but a peptide threshold above or equal to 95%.

773

774

775

776 **Table 2. Cellular proteins identified by nanoLC-HR-MS/MS.**

Identified Protein	GO – Biological process	Total spectra	Protein coverage (%)	GI no.	Species	Order
Cu,Zn superoxide dismutase	Response to oxidative stress	14	8	1019906	<i>Drosophila saltans</i>	Diptera
Histone H4	Nucleosome assembly	7	13	194772468	<i>Drosophila ananassae</i>	Diptera
Muscle myosin heavy chain	Myosin filament organization	5	1	183979376	<i>Papilio xuthus</i>	Lepidoptera
Actin	Cytoskeleton process	4	9	425893652	<i>Psoroptes cuniculi</i>	Astigmata
ADP, ATP carrier protein	Response to oxidative stress	4	11	195438665	<i>Drosophila willistoni</i>	Diptera
ADP, ATP carrier protein 2	Response to oxidative stress	3	11	751238593	<i>Solenopsis invicta</i>	Hymenoptera
Hrp59 protein	mRNA splicing	4	2	50880296	<i>Chironomus tentans</i>	Diptera
Tubulin beta chain	Microtubule-based process	3	2	157108656	<i>Aedes aegypti</i>	Diptera
Hsp 70 BD1	Stress response	3	4	311797659	<i>Bactrocera dorsalis</i>	Diptera

777

778 **FIGURE LEGENDS**

779

780 **Fig. 1. Coomassie-stained SDS-PAGE of ToNV occlusion bodies analysed by LC-**
781 **MS/MS.** Five microliters of purified OB (lane 2) were denatured then run on a 12.5% SDS-
782 PAGE along with Kaleidoscope molecular weight (Biorad, lane 1) at 70 V during 3 h. Major
783 bands and their apparent molecular weight are indicated by arrows.

784

785 **Fig. 2. Qualitative proteomic analyses of ToNV major and associated OB proteins**
786 **identified by nanoLC-HR-MS/MS.** Relative quantity of all identified major or associated
787 OB proteins was presented as the total spectrum count defined by the Scaffold software
788 (Proteome software Inc). Nudiviral core and accessory proteins were represented by black and
789 gray bars, respectively. In both bar graph and pie chart synthetic graph, proteins were
790 addressed according to their potential function in other viruses. *, proteins unambiguously
791 identified with a unique peptide and a minimal protein threshold at 95%; **, proteins
792 unambiguously identified with a unique peptide, a protein threshold comprised between 50
793 and 94% but a peptide threshold above or equal to 95%. All members of the ToNV ORF032,
794 PIF-5 and VP91/95 multigenic families were displayed in the figure although for some of
795 them (#) no peptide was detected from ToNV OBs (ORF033, PIF-5a and ORF083,
796 respectively).



