Characterization of the Neospora caninum NcROP40 and NcROP2Fam-1 rhoptry proteins during the tachyzoite lytic cycle

Journal:	*Parasitology*
Manuscript ID	PAR-2015-0243.R2
Manuscript Type:	Research Article - Standard
Date Submitted by the Author:	n/a
Complete List of Authors:	Pastor-Fernández, Iván; Complutense University of Madrid, Animal Heatlh Regidor-Cerrillo, Javier; Complutense University of Madrid, Animal Heatlh Jiménez-Ruiz, Elena; Complutense University of Madrid, Animal Heatlh Álvarez-García, Gema; Complutense University of Madrid, Animal Health Marugán-Hernández, Virginia; Complutense University of Madrid, Animal Heatlh Hemphill, Andrew; Institute of Parasitology, University of Bern Ortega-Mora, Luis; Complutense University of Madrid, Animal Health
Key Words:	Neospora caninum, NcROP40, NcROP2Fam-1, characterization, in silico analysis, lytic cycle of tachyzoites, immunolocalization, secretion assays, mRNA expression profile, protein phosphorylation

Characterization of the *Neospora caninum* NcROP40 and NcROP2Fam-1 rhoptry proteins during the tachyzoite lytic cycle

Iván Pastor-Fernández[1], Javier Regidor-Cerrillo[1], Elena Jiménez-Ruiz[1], Gema Álvarez-García[1], Virginia Marugán-Hernández[1], Andrew Hemphill[2], Luis M. Ortega-Mora[1*].

[1] SALUVET, Animal Health Department, Faculty of Veterinary Sciences, Complutense University of Madrid, Ciudad Universitaria s/n, 28040-Madrid, Spain.

[2] Institute of Parasitology, Vetsuisse Faculty, University of Berne, Länggass-Strasse 122, CH-3012 Berne, Switzerland.

Running title: **NcROP40 and NcROP2Fam-1 characterization during the tachyzoite lytic cycle**

[*]Corresponding author: Luis Miguel Ortega Mora. Tel.: +34-913944069; fax: +34-913944098. E-mail address: luis.ortega@vet.ucm.es

SUMMARY

Virulence factors from the ROP2-family have been extensively studied in *Toxoplasma gondii*, but in the closely related *Neospora caninum* only NcROP2Fam-1 has been partially characterized to date. NcROP40 is a member of this family and was found to be more abundantly expressed in virulent isolates. Both NcROP2Fam-1 and NcROP40 were evaluated as vaccine candidates and exerted a synergistic effect in terms of protection against vertical transmission in mouse models, which suggests that they may be relevant for parasite pathogenicity. NcROP40 is localized in the rhoptry bulbs of tachyzoites and bradyzoites, but in contrast to NcROP2Fam-1, the protein does not associate with the parasitophorous vacuole membrane due to the lack of arginine-rich amphipathic helix in its sequence. Similarly to NcROP2Fam-1, Nc*ROP40* mRNA levels are highly increased during tachyzoite egress and invasion. However, NcROP40 up-regulation does not appear to be linked to the mechanisms triggering egress. In contrast to NcROP2Fam-1, phosphorylation of NcROP40 was not observed during egress. Besides, NcROP40 secretion into the host cell was not successfully detected by immunofluorescence techniques. These findings indicate that NcROP40 and NcROP2Fam-1 carry out different functions, and highlight the need to elucidate the role of NcROP40 within the lytic cycle and to explain its relative abundance in tachyzoites.

KEYWORDS

Neospora caninum, NcROP40, NcROP2Fam-1, characterization, *in silico* analysis, lytic cycle of tachyzoites, immunolocalization, secretion assays, mRNA expression profile, protein phosphorylation.

KEY FINDINGS

NcROP40 is localized in the rhoptry bulbs of tachyzoites and bradyzoites.

NcROP40 does not associate with the PVM, and its secretion could not be ruled out.

NcROP2Fam-1 secretion was detected during or following host cell invasion.

NcROP40 and NcROP2Fam-1 mRNA levels are highly increased during tachyzoite egress and invasion.

DTT-induced egress increases transcription of NcROP2Fam-1, while NcROP40 expression is not affected.

In contrast to NcROPFam-1, NcROP40 phosphorylation is not associated with egress.

INTRODUCTION

Neospora caninum is a cyst-forming parasite that causes neuromuscular disorders in dogs, and abortion, stillbirth and birth of weak offspring in bovines. This protozoan is phylogenetically related to *Toxoplasma gondii*, with which it shares the ability to cross the placenta and to infect the foetus. In cattle, asexually proliferating tachyzoites and bradyzoites are the only stages described. Tachyzoites have a high proliferative potential and are thus responsible for the dissemination of the parasite into different tissues. Bradyzoites ensure parasite persistence by forming tissue cysts located in immune-privileged organs such as the brain (Dubey and Schares 2011). Since these two stages are strictly intracellular, they have developed a number of mechanisms to actively invade their host cells and modulate their intracellular compartment to optimize intracellular survival and growth. These processes are grouped under the name of lytic cycle (Hemphill et al. 2013). Important structures exclusively found in apicomplexans, namely the apical complex and specialized secretory organelles such as micronemes, rhoptries and dense granules play important roles in the lytic cycle. Contents of these secretory organelles are sequentially released to ensure invasion, intracellular maintenance and replication of the parasite in parasitophorous vacuoles, where they mediate and influence the host cell machinery (Kemp et al. 2013). Among apicomplexan parasites, the molecular basis of the lytic cycle is highly conserved, and the underlying of mechanisms described for *T. gondii* (Carruthers and Sibley 1997) and *Plasmodium* spp. (Cowman et al. 2012) are likely to be similar in *N. caninum* (Hemphill et al. 2013).

Rhoptries have been the subject of extensive studies during the last years due to the role of their proteins in host cell invasion and cell regulation processes. Some of these proteins (RONs) are restricted to the neck, and others (ROPs) to the bulb of these

organelles. RONs are involved in the formation of the moving junction required for parasite entry into the host cells (Beck et al. 2014). The ROP2-family represents one of the largest and best-studied group of ROP proteins in *T. gondii*, and includes protein kinases and pseudokinases that are proven virulence factors (Etheridge et al. 2014; Lei et al. 2014; Reese et al. 2014; Schneider et al. 2013). To our knowledge, most of the ROP2-like proteins are secreted into the host cytosol during invasion and some of them can associate with the parasitophorous vacuole membrane (PVM), but their function is still largely unknown (Boothroyd and Dubremetz 2008; Bradley and Sibley 2007; El Hajj et al. 2006). The ROP2-family has been recently catalogued in *N. caninum* (Talevich and Kannan 2013), but only limited information is available on this protein family. Currently, only NcROP1, NcROP2Fam-1, NcROP4, NcROP5, NcROP9, NcROP30 and NcROP40 have been identified by proteomic studies (Marugán-Hernández et al. 2011; Regidor-Cerrillo et al. 2012; Sohn et al. 2011), but their function has not been described. To date, the only *N. caninum* rhoptry protein that has been partially characterized is NcROP2Fam-1 (Alaeddine et al. 2013). This protein was previously considered the orthologue of TgROP7 (Reid et al., 2012). However, it has been recently shown that TgROP7 and NcROP2Fam-1 are unlikely to be orthologues (Alaeddine et al. 2013). A fragment of NcROP2Fam-1 has been employed as a vaccine in mouse models, showing relatively high protection rates against challenge infection (Debache et al. 2008; Debache et al. 2009; Debache et al. 2010). Another rhoptry protein, NcROP40, was found to be more abundantly expressed in virulent isolates of *N. caninum* (Regidor-Cerrillo et al. 2012), thus posing the obvious question whether NcROP40 plays a potential role in parasite virulence as described for other rhoptry proteins in *T. gondii*. When applied as vaccines a combined NcROP40+NcROP2Fam-1 protein formulation had a synergistic effect and was able to induce a partial block in

transplacental transmission in a pregnant mouse model of neosporosis (Pastor-Fernández et al. 2015).

The aim of the present work was to characterize NcROP40 and compare its features with NcROP2Fam-1 during the lytic cycle of *N. caninum* development. This includes the molecular characterization of the NcROP40 through *in silico* studies, define its subcellular localization throughout the lytic cycle in comparison with NcROP2Fam-1, and to study protein dynamics, the transcript expression profile and their phosphorylation in order to predict their putative functional role in the different phases of the tachyzoite lytic cycle.

151

152

MATERIALS AND METHODS

In silico *analysis and NcROP40 sequencing*

All the sequences were obtained from ToxoDB v7.3. and v12 (www.toxodb.org) and edited using the BioEdit software v7.1.1. BLAST tools from NCBI (www.ncbi.nlm.nih.gov/BLAST/) and ToxoDB websites were used to match homologous sequences. Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) was employed to align nucleotide and protein sequences. Identity and similarity percentages were calculated with the Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/ident_sim.html). Open Reading Frames (ORFs) and introns were predicted through the ORF Finder Tool (www.ncbi.nlm.nih.gov/gorf/gorf.html, NCBI) and the Splign Tool (www.ncbi.nlm.nih.gov/sutils/splign/splign.cgi, NCBI), respectively. SignalP 4.1 server (www.cbs.dtu.dk/services/SignalP/, CBS) was used to predict signal peptides. Potential alpha helices in the arginine-rich amphipathic helix (RAH) domain were searched using Jpred3 (http://www.compbio.dundee.ac.uk/www-jpred/), PSIPRED v3.0 (http://bioinf.cs.ucl.ac.uk/psipred/) and PSSpred (http://zhanglab.ccmb.med.umich.edu/PSSpred/) tools. Trans-membrane regions were predicted with the TMPred tool (www.ch.embnet.org/software/TMPRED_form.html, ExPASy) and protein families from Pfam database (pfam.sanger.ac.uk/, Sanger). Potential phosphorylation sites were analyzed by the NetPhos v2.0 (http://www.cbs.dtu.dk/services/NetPhos/), NetPhosK v1.0 (http://www.cbs.dtu.dk/services/NetPhosK/) and the Diphos v1.3. (http://www.dabi.temple.edu/disphos/) servers.

177 The Nc*ROP40* gene (previously named Nc*ROP8*, NCLIV_012920 in ToxoDB v12) was
178 sequenced and compared among three *N. caninum* isolates of different origins. For this
179 purpose, total genomic DNA from Nc-Liv (Barber et al. 1993), Nc-Spain7 (Regidor-
180 Cerrillo et al. 2008) and Nc-Spain1H (Rojo-Montejo et al. 2009) isolates was purified
181 with the DNeasy Blood & Tissue Kit (Qiagen) following the manufacturer's
182 recommendations. The *NcROP40*-ORF (1176 bp) and the up and down-stream regions
183 (750 + 992 bp) were amplified from the three isolates using the Fw-chrV_ROP40 and
184 Rv-chrV_ROP40 primers (Additional file 1). PCR conditions were 95º C for 5 min, 35
185 cycles at 95º C for 1 min, 58º C for 1 min and 72º C for 1 min, and a final elongation at
186 72º C for 10 min. PCRs were carried out with the Platinum® Taq DNA Polymerase
187 High Fidelity (Invitrogen) and all primers were purchased from Sigma-Aldrich.
188 Amplified fragments were purified with the GENECLEAN Turbo kit (MP Biomedicals)
189 from 1% low melting agarose gels. DNA was sequenced in two directions with an ABI
190 Prism 377 DNA sequencer (Applied Biosystems) in the Genomics Unit of the Scientific
191 Park of Madrid. Six pairs of primers were employed for this purpose (Additional file 1).
192 Sequences were edited and aligned using the BioEdit software v7.1.1.
193
194 ***Parasite culture***
195 *N. caninum* (Nc-Liv isolate) tachyzoites were propagated *in vitro* by continuous passage
196 in MARC-145 cell culture using standard procedures (Pérez-Zaballos et al. 2005). For
197 transmission electron microscopy, murine epidermal keratinocyte cultures were infected
198 with the same isolate as described earlier (Vonlaufen et al. 2002). *In vitro* tachyzoite-to-
199 bradyzoite stage conversion was induced and checked by BAG1 and CC2 expression as
200 previously described (Hemphill et al. 2004). Evacuole assays were performed with the

Nc-Liv isolate in human foreskin fibroblasts (HFFs) as previously described (Dunn et al. 2008).

Generation of plasmids

Nc*ROP40* (NCLIV_012920 in ToxoDB v12) and Nc*ROP2Fam-1* (NCLIV_001970 in ToxoDB v12) were cloned in the pET45b(+) expression system (Novagen) as previously described (Pastor-Fernández et al. 2015; Regidor-Cerrillo et al. 2012). On the other hand, NcAlpha-Tubulin (TUBα) (NCLIV_058890 in ToxoDB v12) and NcSAG1 (NCLIV_033230 in ToxoDB v12) fragments were amplified from *N. caninum* cDNA and cloned within the pGEM-T-Easy vector (Promega). Primer sequences for cloning are summarized in Additional file 2. All primers were purchased from Sigma-Aldrich, and the Expand High Fidelity Plus PCR System (Roche) was used for all PCRs. Amplicons were purified with the GENECLEAN Turbo kit (MP Biomedicals) from 1% low melting agarose gels (Lonza).

Production of recombinant proteins, mass spectrometry analysis and SDS-PAGE

E. coli NovaBlue Single Competent Cells (Novagen) were transformed with construct-containing plasmids, which were isolated using the QIAprep Spin Miniprep Kit (Qiagen) and sequenced with an ABI Prism 377 DNA sequencer (Applied Biosystems) using T7 forward and reverse primers in the Genomics Unit of the Scientific Park of Madrid. All sequences were aligned with 100% consensus.

E. coli BL21(DE3) pLysS competent cells (Agilent Technologies) were transformed with the resulting expression vectors and foreign expression of rNcROP40 and rNcROP2Fam-1 as a (His)6-tagged fusion proteins was carried out following standard procedures (Álvarez-García et al. 2007). Denatured proteins were on-column refolded

and purified using HisTrapHP columns coupled to the ÄKTAprime Plus system (GE Healthcare) as previously described (Pastor-Fernández et al. 2015). Recombinant proteins included the whole NcROP40 sequence (1-392 aa) and the C-terminus domains for rNcROP2Fam-1 (238-594 aa), excluding the RAH domains. Concentration and purity of recombinant proteins was checked by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) with a standard BSA scale (Roche) and using the GS-800 densitometer coupled to the Quantity One software (Bio-Rad Laboratories) (Álvarez-García et al. 2007). Electrophoresed proteins were manually excised from prepared Coomassie-stained 1-D gels for mass spectrometry (MS) analysis (peptide mass fingerprinting) following standard procedures (Risco-Castillo et al. 2007).

Polyclonal antibody production and affinity purification

Polyclonal sera against rNcROP40 (Regidor-Cerrillo et al. 2012) and rNcROP2Fam-1 were raised in New Zealand White rabbits (Harlan Laboratories) following a procedure previously described (Risco-Castillo et al. 2007). Samples of pre-immune serum were collected to confirm the absence of antibodies against *N. caninum* by Western-Blot. All protocols followed the proceedings detailed by the current legislation at the time of the experiment (Spanish Royal Decree 1201/2005) and were approved by the Animal Research Committee of the Complutense University. Affinity purified antibodies were prepared from polyclonal antibodies (PAbs) following standard procedures (Álvarez-García et al. 2007).

Immunoblots

Detection of NcROP40 and NcROP2Fam-1 proteins in parasite extracts was carried out by Western-Blot following SDS-PAGE under reducing conditions. Unless otherwise

stated, all reagents were purchased from Bio-Rad Laboratories. 2×10^7 purified Nc-Liv tachyzoites were disrupted by bath-sonication, electrophoresed in 15% bis-acrylamide gels and transferred onto nitrocellulose membranes according to standard procedures. PAbs α-rNcROP40 and α-rNcROP2Fam-1 were diluted at 1:5,000. Goat anti-rabbit IgG antibody conjugated to peroxidase (Sigma-Aldrich) was used as secondary antibody at 1:25,000 dilution. Reactions were developed by chemiluminiscence with the Immobilon Western Chemiluminescent HRP Substrate (Millipore). For image acquisition, AGFA films (Curix/RP2 Plus) and AGFA CP1000 processor were used after 1 to 30 seconds of exposure time.

Immunogold-labeling and transmission electron microscopy (TEM)

Infected keratinocyte cultures were fixed and LR-White embedded and on-section labeled as previously described (Risco-Castillo et al. 2007). Affinity-purified rabbit α-rNcROP40 was diluted 1:2 in PBS-0.3% BSA and sections were incubated for 1 h in a moist chamber. They were then washed in three changes of PBS, 10 min each, and goat anti-rabbit conjugated to 10 nm diameter gold particles (Amersham) was applied at a dilution of 1:5 in PBS-0.3% BSA as secondary antibody. After another 3 washes, 10 min each, grids were air dried and contrasted with uranyle acetate and lead citrate (Hemphill et al. 2004). Specimens were viewed on a Phillips 600 TEM operating at 60 kV.

Immunofluorescence staining

Protein localization dynamics in *N. caninum* tachyzoites were studied on infected MARC-145 cells on rounded coverslips at different time-points after infection. A total of 5×10^4 cells were placed on sterile 13 mm-coverslips onto 24-well plates and

incubated overnight at 37ºC on a 5% CO_2 atmosphere. Then, tachyzoites were scrapped from culture flasks, passed through a 21-gauge needle and counted on a haemacytometer by trypan blue exclusion. Subsequently, MARC-145 monolayers were infected with *N. caninum* for 20 and 40 min (MOI 3), 1, 2, 6 (MOI 3), 24, 32 (MOI 2) and 48 h (MOI 1). After infection, non-adherent parasites were removed from coverslips by three PBS washes. Then, three methods of fixation were employed. Absolute methanol, 2% paraformaldehyde in PBS and 2% paraformaldehyde-0.05% glutaraldehyde in PBS were used as fixatives for 10 to 30 minutes at room temperature. All samples were immediately processed for immunofluorescence staining.

Coverslips were blocked and permeabilised with PBS containing 3% bovine serum albumin (Roche) and 0.25% Triton X-100 (Merck Chemicals) for 30 min at 37ºC. Then, cultures were labelled with the monoclonal antibody (MAb) α-NcSAG1 as a surface marker ((Bjorkman and Hemphill 1998); 1:250 dilution) and affinity purified PAbs α-NcROP40 and α-NcROP2Fam-1 (1:8 dilution) by incubation for 1 h at room temperature. Following three washes with PBS, coverslips were incubated with Alexa Fluor 488-conjugated goat anti-mouse IgG and Alexa Fluor 594-conjugated goat anti-rabbit IgG at 1:1,000 dilution (Molecular Probes) for 1 h at room temperature. Nuclei were stained with 4',6-diamidino-2-phenylindole dye (DAPI, Lonza) at 1:5,000 dilution in PBS. Finally, coverslips were mounted on glass slides with ProLong® Gold antifade reagent (Molecular Probes). Evacuoles were detected on infected HFFs in the presence of cythochalasin D following the same protocol (Additional file 5).

To phalloidin staining, coverslips fixed in 2% paraformaldehyde were blocked, permeabilised and labelled with MAb α-NcSAG1 and affinity purified PAb α-NcROP2Fam-1 as described above. After washing, they were incubated with Alexa

Fluor 647-conjugated goat anti-mouse IgG (1:1,000), Alexa Fluor 488-conjugated goat anti-rabbit IgG (1:1,000), phalloidin–TRITC (1:250) (Sigma-Aldrich) and DAPI dyed. Single 1µm slices of immunofluorescence stainings were captured with a Leica TCS-SPE confocal laser-scanning microscope (Leica Microsystems) in the Department of Biochemistry and Molecular Biology IV of the Complutense University (Madrid). Image processing was performed using the LAS AF (Leica Microsystems) and the ImageJ software (NCBI, http://rsb.info.nih.gov/ij/).

Secretion assays

Secretion assays were performed with tachyzoites obtained from cultures prior to egress. For this purpose, parasites were scrapped from culture flasks, pelleted by centrifugation (1350 × g, 10 min, 4º C), passed through a 21-gauge needle and purified by PD-10 desalting columns (GE Healthcare). Then, 1×10^8 tachyzoites were placed on 500 µl of cold phenol red-free DMEM (Life-Technologies) and stimulated with either 10 µM A23187 (Sigma-Aldrich), 1% ethanol (Merck Chemicals), or 10 mM dithiothreitol (DTT, Calbiochem) for 20 minutes at 37º C (Naguleswaran et al. 2001). Non-stimulated parasites were kept on ice during the same period of time. After the incubation, secretion supernatants were recovered by double centrifugation (1350 × g, 10 min, 4º C and 8000 × g, 10 min, 4º C), passed through 0.2 µm PVDF filters (Whatman, GE Healthcare) and supplemented with phosphatase and protease inhibitor cocktails (Sigma-Aldrich). Pelleted parasites were washed once in cold PBS supplemented with phosphatase and protease inhibitor cocktails, and recovered by centrifugation (1350 × g, 10 min, 4º C). All samples were stored at -80º C until further analyses.

Supernatants and pellets were analysed by immunoblotting, and secretion was estimated by comparing equal amounts of secretion supernatants and tachyzoite lysates. Monoclonal antibodies directed against NcTUBα (α-TUBα MAb, Sigma-Aldrich) were used on immunoblots of secreted supernatant fractions to monitor tachyzoite lysis, and antibodies directed against NcMIC2 were used as a positive control of secretion (Lovett et al. 2000). The α-TUBα MAb specifically recognized NcTUBα protein in tachyzoite extracts. PVDF membranes were incubated with rabbit α-rNcROP40, α-rNcROP2Fam-1 and α-rNcMIC2 at 1:5,000 dilutions, whereas α-TUBα MAb was employed at 1:10,000 dilution. Secondary antibodies were employed at 1:25,000 (goat anti-rabbit IgG antibody conjugated to peroxidase) and at 1:80,000 dilutions (goat anti-mouse IgG antibody conjugated to peroxidase) (Sigma-Aldrich). Reactions were developed by chemiluminiscence with the Immobilon Western Chemiluminescent HRP Substrate as describe above.

Evaluation of NcROP40 and NcROP2Fam-1 mRNA expression levels

The mRNA expression levels of NcROP40 and NcROP2Fam-1 were assessed by real-time reverse transcription PCR throughout the lytic cycle of tachyzoites at four representative points which illustrate the recent invasion, PV formation and maturation, exponential growth of parasites and tachyzoite egress. For this purpose, MARC-145 cultures were infected with the Nc-Liv isolate at MOI 3 for 6, 24, 48 and 56 h. Infected cultures were synchronised by washing the monolayer twice with pre-warmed PBS and replacing the culture media at 6 hours post-infection (hpi), to remove non-adherent parasites. Cells were harvested with a cell scraper and recovered by centrifugation at 1,350 × *g* for 15 minutes at 4º C. Pelleted parasites were conserved at -80º C until RNA extraction. The experiment was carried out in triplicate. For each experiment, three

different flasks were analysed at each time-point. The effect of induced egress of tachyzoites on expression levels of NcROP40 and NcROP2Fam-1 was also studied in parallel. For this purpose, five flasks from three different experiments containing cells that were infected for 48 h were treated with 10 mM DTT for 1 h, after which tachyzoites had undergone egress from approximately 80% of parasitophorous vacuoles. Tachyzoites were then recovered as described above.

Total RNA was extracted using the Maxwell® 16 LEV simplyRNA Purification Kit (Promega), that includes a DNAse treatment, following the manufacturer's recommendations. RNA concentrations were determined by spectrophotometry (Nanophotometer, Implen), and RNA integrity was checked by the visualization of the 18S and 28S ribosomal fragments after electrophoresis on 1% agarose gels. Reverse transcription was carried out by the master mix SuperScript® VILO™ cDNA Synthesis Kit (Invitrogen) in a 20 µl reaction using 2.5 µg of total RNA. Resulting cDNA was diluted 1:20 and analysed by real-time PCR.

Real-time PCR reactions were performed using the Power SYBR® Green PCR Master Mix in the ABI 7300 Real Time PCR System (Applied Biosystems) following standard conditions. Primers used for amplification of NcROP40, NcROP2Fam-1 and the housekeeping genes NcTUBα and NcSAG1 are shown in Table 1. A seven-point duplicate standard curve based on 10-fold serial dilutions was included on each run. pET45b(+)-NcROP40, pET45b(+)-NcROP2Fam-1, pGEM-T-NcTUBα and pGEM-T-NcSAG1 plasmids were used as standards.

mRNA expression levels for each target were normalized by the -ΔCt method (Schmittgen and Livak 2008). -ΔCt values were calculated by subtracting the Ct value of the normalizer genes from the Ct value of each sample. Relative fold increases or decreases were assessed by the $2^{-\Delta\Delta Ct}$ method (Schmittgen and Livak 2008). Since

374 expression levels at 24 hpi were the lowest for both proteins, the -ΔΔCt value was
375 calculated by subtracting the mean -ΔCt values for each protein at 24 hpi as baseline
376 samples as indicated in this formula: -ΔΔCt = -[(Ct NcROP*x* - Ct NcTUBα) - (mean Ct
377 NcROP*x* at 24 h - mean Ct NcTUBα at 24 h)]. Raw RNA samples were included in
378 each batch of amplifications to confirm the absence of *N. caninum* genomic DNA. Data
379 analyses of mRNA expression levels were carried out by Kruskal-Wallis and Dunn's
380 tests using GraphPad Prism v6.01 software.
381
382 ***Phosphorylation assays***
383 Phosphorylation assays were performed on denatured lysates generated from infected
384 MARC-145 cell cultures at 56 hpi, when tachyzoite had escaped from parasitophorous
385 vacuoles and invaded neighboring cells. Freshly pelleted cell monolayers were
386 resuspended on alkaline phosphatase-compatible buffer (100 mM sodium chloride
387 [Panreac], 50 mM Tris-HCl [Panreac], 10 mM magnesium chloride [Merck Chemicals],
388 1 mM DTT [Calbiochem], 0.2 % Triton X-100 [Merck Chemicals] and protease
389 inhibitor cocktail [Sigma-Aldrich], pH 7.9) or in phosphatase inhibitor buffer (50 mM
390 HEPES [Sigma-Aldrich], 100 mM sodium fluoride [Sigma-Aldrich], 2 mM sodium
391 orthovanadate [Sigma-Aldrich], 2 mM EDTA [Sigma-Aldrich], 1 mM DTT, 0.2 %
392 Triton X-100 and protease inhibitor cocktail). Extracts were disrupted on ice for 15
393 minutes by bath-sonication (Ultrasons, Selecta) and shaked by vortexing during an
394 additional 45 minutes. Alkaline phosphatase treatment (20 U CIP/2×10^7 tachyzoites,
395 New England Biolabs) was only applied on extracts resuspended in alkaline
396 phosphatase-compatible buffer for 90 minutes at 37º C. Resulting extracts were stored
397 at -80ºC until further analysis.

Phosphorylated proteins experience a mobility shift on Phos-tag SDS-PAGE electrophoresis (Kinoshita et al. 2006). To determine if NcROP40 and NcROP2Fam-1 are phosphorylated, tachyzoite extracts resuspended in alkaline phosphatase-compatible buffer (CIP) and phosphatase inhibitor buffer (PI) were electrophoresed in 15% bis-acrylamide gels supplemented with 25 μM Phos-Tag (Wako Pure Chemicals Industries) and 50 μM manganese (II) chloride (Merck Chemicals). After electrophoresis, gels were washed once in 0.1 M EDTA in transfer buffer and once in transfer buffer without EDTA to remove metal complexes. Then, gels were transferred onto nitrocellulose membranes according to standard procedures. Membranes were incubated with α-rNcROP40 and α-rNcROP2Fam-1 at 1:1,000 dilution, and then incubated with goat anti-rabbit IgG antibody conjugated to peroxidase at 1:1,000 dilution. Reactions were developed using 4-chloro-1-naphtol as substrate until signal visualization.

423

424

425

426

427

RESULTS

In silico *analysis and NcROP40 sequencing*

NcROP40 (NCLIV_012920, chromosome V) is currently classified as an unspecified product, but in previous releases of ToxoDB (v7.3) the protein was named as rhoptry kinase family protein ROP40 and considered as orthologous gene of TgROP40. No introns are predicted in the Nc*ROP40* sequence, which contains 1176 bp and codes for a product of a predicted molecular weight of 43 kDa. In contrast, the Tg*ROP40* sequence (TGME49_291960, chromosome IX) contains two introns, a coding sequence of 1578 bp and a predicted molecular weight of 57.9 kDa. However, according to previous releases of ToxoDB (v7.3), the Tg*ROP40* gene (TGME49_091960) has no introns. The predicted peptide sequences of NcROP40 and TgROP40 proteins (v7.3) share 32.3% identity and 48.7% similarity. In addition, a Pfam database search identified a protein kinase-like domain in NcROP40, but catalytic activity is lacking (PANDIT: PF14531).

Due to the observed inconsistencies between Nc*ROP40* and Tg*ROP40*, the chromosome V-sequence of the *N. caninum* Nc-Liv genome was analyzed in detail. First, the N-terminus of the Nc*ROP40* sequence was strikingly shorter than that annotated for Tg*ROP40* (~400 bp). In order to elucidate these differences, up and down-stream Nc*ROP40* sequences (from positions 662772 to 665947, chromosome V) were submitted to the ORF Finder tool, which displayed a unique 1578 bp ORF, which corresponds to a putative protein with a calculated molecular weight of 57.8 kDa. This finding is consistent with the Tg*ROP40* sequence. The presence of the additional N-terminal fragment in the Nc*ROP40* ORF was confirmed by RT-PCR using cDNA from tachyzoites at two different time points of infection (Additional file 3). The newly identified ORF (now termed Nc*ROP40-long*) contains the Nc*ROP40* sequence as listed

453 in ToxoDB and an additional 402 bp at its N-terminus (Additional file 4) (GenBank:
454 KP731805, KP731806 and KP731807). According to the previous Tg*ROP40* gene
455 (TGME49_091960 in ToxoDB v7.3), no introns were predicted when the Nc*ROP40-*
456 *long* DNA sequence was submitted to the Splign tool. The percentage of amino acid
457 sequence identity between TgROP40 and NcROP40-long increased from 32.3 to 42.9,
458 whilst similarity increased from 46.3% to 61.4% (Fig. 1A).
459 Protein trans-membrane regions were predicted for the NcROP40-long protein between
460 the positions 5 and 25, but according to SignalP predictions the signal peptide is cleaved
461 between amino acids (aa) 19 and 20. Sequence comparison among the most
462 representative members of the ROP2-family (Fig. 1B), as well as alpha helices
463 prediction within the structure of the NcROP40-long protein (Fig. 1C) suggest that the
464 protein lacks RAH domains. Phosphorylation sites both in NcROP40-long and
465 NcROP2Fam-1 were subjected to three different prediction programs (NetPhos v2.0,
466 NetPhosK v1.0, and Diphos v1.3.), and were only considered when detected by at least
467 two of them. In this sense, NcROP40 showed two putative phosphorylation sites at
468 position S-75 and S-78, whilst phosphorylation of NcROP2Fam-1 was predicted to
469 occur at positions S-82 and S-129.
470 The Nc*ROP40-long* ORF and its up and down-stream sequences were amplified by
471 PCR from DNA of three different isolates, which have shown manifest differences in
472 virulence: Nc-Liv, Nc-Spain7 and Nc-Spain1H. Primers were designed to amplify a
473 fragment of 2918 bp, containing the ORF. After PCR, amplicons were sequenced and
474 analyzed in detail. For all isolates a 2362 bp consensus fragment was sequenced in two
475 directions. This fragment comprised the Nc*ROP40-long* ORF (1578 bp) and an
476 additional 148 and 636 bp in its flanking regions. Thus, comparative analyses did not
477 show differences in the amplified sequences among the three isolates (Additional file 4).

Protein sequence and immunodetection

The identities of rNcROP40 and rNcROP2Fam-1 were confirmed by mass spectrometric analyses. rNcROP40 matched with the NCLIV_012920 sequence (score: 175; 18/60 matched values; 64% of sequence coverage), whereas rNcROP2Fam-1 matched with the NCLIV_001970 annotation (score: 345; 33/65 matched values; 69% of sequence coverage). These analyses corroborated the composition of both proteins, with a molecular weight of 43.9 and 43.2 kDa respectively, according to the predicted molecular weights of rNcROP40 and rNcROP2Fam-1, which exclude part of their N-terminal domains. Hence, rNcROP40 and rNcROP2Fam-1 were used to develop PAbs in rabbits. *N. caninum* tachyzoite crude extracts were separated by SDS-PAGE under reducing conditions. Western blots revealed that α-NcROP40 reacted with five distinct bands of approximately 53, 44, 38, 32 and 28 kDa. The polyclonal α-NcROP2Fam-1 antiserum detected six different bands of approximately 58, 48, 40, 36, 34 and 26 kDa (Fig. 2).

Subcellular localization of NcROP40 by TEM

In order to confirm the subcellular localization of NcROP40, immunogold-TEM was carried out on sections of keratinocytes infected with *N. caninum* tachyzoites, and of cultures infected with bradyzoites generated *in vitro* by sodium nitroprusside treatment. In both, *N. caninum* tachyzoites (Fig. 3, A-B) and *in vitro* induced-bradyzoites (Fig. 3, C-D), affinity-purified anti-NcROP40 antibodies localized to rhoptry bulbs.

NcROP40 and NcROP2Fam-1 tracing throughout the lytic cycle

Immunofluorescence staining of NcROP40 on *N. caninum*-infected cultures showed a rhoptry-like pattern in tachyzoites throughout the lytic cycle, from 20 minutes to 56 hpi.

503 Methanol fixation showed the clearest results in terms of NcROP40
504 immunolocalization. In contrast, fixation with paraformaldehyde and glutaraldehyde
505 mixtures resulted in a lower staining intensity (Fig. 4). The rhoptry-like pattern was
506 clearly associated with the apical end of tachyzoites in all the micrographs, and
507 disappeared as the captured slices intersected the parasites in more external areas (Fig.
508 4, 24 and 32 hpi, MeOH fixation). Interestingly, the presence of NcROP40 was not
509 detected, neither in evacuoles during the invasion phases nor in the PVM during the
510 development and establishment of the PV, with similar results obtained using three
511 different fixation protocols. Hence, no secretion of NcROP40 protein could be detected
512 under the tested conditions.
513 Concerning NcROP2Fam-1, our PAbs specifically recognized rhoptry-like structures at
514 the parasite apex at all time-points as described for NcROP40. In contrast, secretion of
515 the protein was detected from 20 min to 24 hpi using all three fixation protocols (Fig.
516 5). Specifically, evacuoles were detected from 20 minutes to 6 hpi. These rhoptry-
517 derived secretory vesicles were localized intracellularly and surrounded the host cell
518 nucleus, as shown by the phalloidin stainings in those coverslips fixed with
519 paraformaldehyde (Fig. 5, 6 h, PFA fixation). At later time points, during the
520 establishment of the PV, NcROP2Fam-1 was detected on the PVM (Fig. 5, 6 hpi) and in
521 the PV matrix (Fig. 5, 24 hpi) under all the fixation methods. Thereafter, NcROP2Fam-
522 1 was restricted to rhoptries (Fig. 5, 32 and 48 hpi). Interestingly, the protein was
523 released again during egress, where it appeared to localize on the surface of the
524 extracellular tachyzoites (Fig. 5, 56 hpi).
525 Identical results were obtained for NcROP40 and NcROP2Fam-1 proteins by specific
526 evacuole assays carried out in human foreskin fibroblasts (Additional file 5).
527

528 ***Induced secretion of NcROP40 and NcROP2Fam-1***

529 Freshly purified tachyzoites were treated with A23187, ethanol or DTT in order to
530 induce the calcium-related protein secretion from apical organelles. However,
531 NcROP40 and NcROP2Fam-1 proteins were not detected on secretome supernatants by
532 immunoblotting. In contrast, a manifest secretion of the NcMIC2 protein was observed
533 under the same conditions, according to previous findings (Lovett et al. 2000).
534 Moreover, inadvertent lysis of tachyzoites during these secretion assays could be
535 discarded since no NcTUBα could be detected in any of the secreted fractions
536 (Additional file 6).

537

538 ***NcROP40 and NcROP2Fam-1 mRNA expression during the lytic cycle***

539 NcROP40 and NcROP2Fam-1 transcription levels were monitored at four
540 representative time points during the lytic cycle. Similar results in the mRNA pattern
541 were observed using NcTUBα and NcSAG1 as normalizer genes (data not shown). The
542 results presented here were processed using NcTUBα as normalizer.

543 The lowest NcROP40 and NcROP2Fam-1 mRNA levels were found at 24 hpi. In
544 contrast, mRNA levels were the highest at 6 hpi (during the invasion phase) and at 56
545 hpi (egress phase) for both proteins (P<0.005; Kruskal-Wallis test) (Fig. 6, A).
546 Differences in fold increases of mRNA transcription were calculated by the $2^{-\Delta\Delta Ct}$
547 method. Since the lowest normalized values for both NcROP40 and NcROP2Fam-1
548 were observed at 24 hpi, this time point was used as baseline to calculate the mRNA
549 transcription fold increases during egress and invasion. NcROP40 showed a 4-fold
550 increase in mRNA levels at 6 and 56 hpi, and NcROP2Fam-1 exhibited a 3-fold
551 increase at the same time points. At 48 hpi, mRNA levels displayed a 2-fold increase for
552 both NcROP40 and NcROP2Fam-1.

553 Egress of *N. caninum* and *T. gondii* tachyzoites can be artificially induced *in vitro* by
554 the addition of DTT into the culture medium. Thus, the effect of DTT supplementation
555 on the expression of NcROP40 and NcROP2Fam-1 mRNA was studied. Different
556 responses were observed: while NcROP40 did not exhibit significant increases in its
557 mRNA levels upon DTT treatment, NcROP2Fam-1 mRNA transcription was
558 significantly increased (P<0.005; Kruskal-Wallis test) (Fig. 6, B). Moreover, mean
559 values for NcROP2Fam-1 mRNA remained above those observed at 56 hpi, while the
560 corresponding values for NcROP40 remained below (Fig. 6, B).
561
562 ***Phosphorylation of NcROP40 and NcROP2Fam-1 at the egress***
563 The phosphorylation status of NcROP40 and NcROP2Fam-1 was studied at 56 hpi, as
564 this was the time point when the mRNA levels for both proteins within the lytic cycle
565 were the highest, simultaneously to tachyzoite egress and early invasion. Tachyzoites
566 were harvested, processed under conditions that preserve the phosphorylation status,
567 and extracts were separated by Phos-Tag SDS-PAGE electrophoresis. The
568 electrophoretic mobility of NcROP40 on Phos-Tag gels was similar in both, alkaline
569 phosphatase-treated and phosphatase inhibitor-treated extracts. In contrast,
570 NcROP2Fam-1 showed a mobility shift in those extracts treated with phosphatase
571 inhibitors, which suggests that NcROP2Fam-1 is phosphorylated at 56 hpi (Fig. 7).
572
573
574
575
576
577

578

579

DISCUSSION

Considerable efforts have been undertaken to increase the understanding on how apicomplexan parasites interact with their host cells and how they maintain and optimize their intracellular life style. It is widely known that components of distinct secretory organelles, namely rhoptries, micronemes and dense granules, play a crucial role in defining the host-parasite relationship (Carruthers and Sibley 1997) and therefore corresponding antigens are being extensively studied as vaccine targets to prevent infections by apicomplexan parasites.

TgROP proteins have shown to be important virulence factors (Lim et al. 2012). In contrast, little is known about the rhoptry proteins in *N. caninum*. Several NcROP and NcRON proteins have been identified by different proteomic approaches and monoclonal antibodies (Marugán-Hernández et al. 2011; Regidor-Cerrillo et al. 2012; Sohn et al. 2011; Straub et al. 2009), but only NcROP2Fam-1, which has been associated with the tachyzoites invasion process, has been partially characterized to date (Alaeddine et al. 2013).

NcROP40 was shown to be more abundantin virulent isolates of *N. caninum* (Regidor-Cerrillo et al. 2012). In *T. gondii*, limited information is available for the orthologous gene product Tg*ROP40* (initially named as TgROP2L6). Tg*ROP40* is highly expressed in tachyzoites (Peixoto et al. 2010) and increased expression levels were observed during acute infections in mice (Pittman et al. 2014). A number of studies suggested that NcROP40 is one of the major rhoptry components, since it has been detected by three different proteomic approaches (Marugán-Hernández et al. 2011; Pollo-Oliveira et al. 2013; Regidor-Cerrillo et al. 2012). Hence, NcROP40 expression could be an important element related to parasite virulence. Indeed, a vaccine formulation combining NcROP40+NcROP2Fam-1 recombinant proteins was recently assessed in a

pregnant mouse model of neosporosis and conferred partial protection against congenital transmission of *N. caninum*, with NcROP40 and NcROP2Fam-1 acting synergistically (Pastor-Fernández et al. 2015).

Comparison of the Nc*ROP40* ORF and its potential regulatory expression sequences among three different *N. caninum* isolates with differing virulence and *in vitro* behavior (Pereira García-Melo et al. 2010; Regidor-Cerrillo et al. 2010; Regidor-Cerrillo et al. 2011) did not reveal any polymorphism that could explain differences in virulence as described for TgROP18 (Steinfeldt et al. 2010). However, dissimilarities in protein abundance among isolates might be due to regulatory mechanisms such as epigenetics, which have been shown to be involved in genome reprogramming during tachyzoite to bradyzoite conversion in *T. gondii* (Dixon et al. 2010). These analyses allowed a detailed dissection of the Nc*ROP40* gene, including flanking regions, and lead to the description of the Nc*ROP40-long* sequence (Nc*ROP40* sequence with an additional 402 bp in its N-terminus), whose presence was confirmed by RT-PCR (Additional file 3). In addition, an improved transcriptome annotation for *NcROP40* has recently been submitted (GenBank: CEL65449.1), confirming our results (Ramaprasad et al. 2015).

Inconsistencies in the measured (53 kDa) and theoretical (58 kDa) molecular weight of NcROP40-long, as well as the presence of different bands on immunoblots, could reflect the maturation process described for all ROP2-family rhoptry proteins, which are synthesized as pro-proteins (Hajagos et al. 2012). In fact, Alaeddine and colleagues described the processing of the NcROP2Fam-1 protein by Western-blot through different affinity purified antibodies directed against peptides located at the C-terminal end of the protein (Alaeddine et al. 2013). In *T. gondii* TgSUB2 protease is in charge to remove the N-terminal domains that are involved in rhoptry targeting at a highly conserved SΦx(E/D) site (Hajagos et al. 2012). This sequence was also found in the N-

terminal domain of NcROP2Fam-1 (Alaeddine et al. 2013), but is absent in NcROP40. In any case, the polyclonal antibody recognizes a main band of 53 kDa and a number of additional bands by Western blotting, which may reflect a protein maturation process. Nevertheless, further studies must be carried out to define more accurately the implication of these changes on protein function.

In silico analyses are useful to predict certain protein features, and were employed in this study to further characterize the NcROP40 protein. The presence of a signal peptide is an important pre-requisite for a protein to enter the secretory pathway in eukaryotes, and putative signal peptides are present in both TgROP40 (El Hajj et al. 2006) and NcROP40 (this work). This is in contrast with our observations, since NcROP40 secretion is not detected by immunofluorescence microscopy, even when employing three different fixation protocols or specific evacuole assays. Moreover, NcROP40 does not interact with the PVM, albeit this finding is consistent with the predicted lack of RAH domains in its sequence. RAH domains are the regions displaying the highest similarities between each of the members of the ROP2-family, and are required for PVM association. These domains are also absent in the TgROP40 protein, and consequently it does not associate with the PVM (El Hajj et al. 2006; Reese and Boothroyd 2009). Nevertheless, previous studies carried out with the toxofilin protein of *T. gondii* showed that secretion of low abundance proteins may be undetectable by immunofluorescence approaches, which is especially relevant for those proteins that are not concentrated on a membrane or in an intracellular compartment (Lodoen et al. 2009). This may be the case for the NcROP40 protein that could be secreted into the host cell cytosol. Interestingly, the TgROP40 protein shows some nuclear localization when is heterologously expressed in infected HFF, suggesting that the protein may be translocated into the host cell nucleus after its secretion in the cytosol. Nevertheless, the

immunodetection of the NcROP40 into the host cell nucleus was not achieved in this study (Reese and Boothroyd 2009). In contrast to NcROP40, NcROP2Fam-1 was extensively secreted under the tested conditions (Alaeddine et al. 2013). The protein was easily detected in evacuoles, and then surrounding invasive tachyzoites from 1 to 24 hpi. Similar findings have been described for the TgROP2 protein, which may participate in the PVM formation (Beckers et al. 1994; Dunn et al. 2008; Nakaar et al.,2003; Sinai and Joiner 2001).Consistent with our findings, most of the rhoptry proteins described to date in *T. gondii* are secreted and participate in host cell invasion, PV formation and maturation, and/or are involved in hijacking the host cell machinery (Kemp et al. 2013).

On the other hand, we could not detect any NcROP40 and NcROP2Fam-1 protein in the secretory fractions after induction of tachyzoite secretion using A23187, ethanol or DTT stimulation. This indicates that rhoptry discharge is not affected by elevated intracellular calcium levels as previously stated for *T. gondii* rhoptry proteins (Carruthers and Sibley 1999), and that rhoptry secretion can be only induced upon host cell contact.

The mRNA levels of NcROP40 and NcROP2Fam-1 transcripts were quantified during defined time points of the lytic cycle of tachyzoites grown in MARC-145 cells (Regidor-Cerrillo et al. 2011). Both proteins displayed higher mRNA levels at 6 hpi (which largely represents recently invaded tachyzoites) and at 56 hpi (representing tachyzoites shortly prior to or already undergoing egress). Lower mRNA levels were measured at 24 hpi (a time point representing early exponential replication). Subsequently, once exponential growth of parasites was almost completed (48 hpi), mRNA levels of NcROP40 and NcROP2Fam-1 gradually increased to reach again their highest value. According to our findings, developmental transitions in *Plasmodium falciparum* and *T. gondii* have shown to be strongly influenced by changes in mRNA

levels (Le Roch et al. 2004; Radke et al. 2005). Indeed, a modal switch from expression of proteins involved in invasion and motility has been also described in extracellular tachyzoites of *T. gondii* (Gaji et al. 2011; Lescault et al. 2010). This could suggest that NcROP40 and NcROP2Fam-1 proteins are required for the subsequent phases of the lytic cycle in which both are highly transcribed. This phenomenon is consistent with the "just-in-time" concept stated for *P. falciparum* and *T. gondii*, whereby gene expression is only activated as their biological function becomes necessary to the parasite (Behnke et al. 2010; Llinas and DeRisi 2004; Radke et al. 2005).

In addition to monitoring mRNA levels during egress under normal culture conditions, the same was done by inducing egress artificially employing DTT at 48 hpi (Esposito et al. 2007). DTT treatment induced a dramatic increase in NcROP2Fam-1 expression after DTT supplementation, to levels similar to naturally occurring egress. Strikingly, and in contrast to NcROP2Fam-1, NcROP40 mRNA levels were not substantially increased by DTT addition. To date, the mechanisms governing egress are not fully understood, but mounting evidence shows that intracellular calcium levels trigger the abrupt exit of parasites from PV, which is accompanied with a rapid decrease in host cell ATP (Blackman and Carruthers 2013). NcROP40 mRNA levels were unresponsive to the artificially-induced egress, suggesting that its up-regulation does not rely on the mechanisms triggering egress in contrast to NcROP2Fam-1.

Phosphorylation has a prominent key role in cellular regulatory processes and influences the functional activity of a plethora of enzymes and structural proteins. At 56 hpi, when the mRNA expression levels for NcROP40 and NcROP2Fam-1 reached a peak and tachyzoites were undergoing egress to infect another host cell, phosphorylation was evident in NcROP2Fam-1, but not in NcROP40. However, we cannot exclude that NcROP40 is phosphorylated at another phase of the lytic cycle.

Predicted phosphorylation sites were found in NcROP40 and in NcROP2Fam-1. However, it is important to note that the phosphorylation prediction algorithms are optimized for mammalian cells or other cell types, and that rhoptry proteins are unique among eukaryotes and are only found in apicomplexan parasites. Thus, potential phosphorylation sites might not be accurately predicted. Previous works have shown that TgROP2 and TgROP4 are also phosphorylated, but only in intracellular parasites, indicating that phosphorylation is associated with protein regulation and its potential participation within the lytic cycle (Carey et al. 2004; Dunn et al. 2008). For NcROP2Fam-1, phosphorylation coincides with high mRNA levels, and since NcROP2Fam-1 was shown to be involved in host cell invasion (Alaeddine et al. 2013), this could indicate that the protein is being activated to prepare tachyzoites for egress and/or invasion. To date, there is no information about the relevance of phosphorylation in NcROP proteins, and phosphorylation of all the known TgROP proteins has not been studied. Previous works suggested that phosphorylation of dense granule proteins has an influence on PVM association (Labruyere et al. 1999; Mercier et al. 2005). Thus, phosphorylation of NcROP2Fam-1 could be important for secretion and its subsequent association to the PVM, and this could be also applied to TgROP2 and TgROP4, both of which exhibit similar properties. However, further studies must be carried out in order to determine the role of rhoptry protein phosphorylation during the lytic cycle of *N. caninum* tachyzoites.

Pseudokinases are emerging as key regulators of cellular signaling (Reese et al. 2014). Several studied rhoptry proteins in *T. gondii* have been described as kinases or pseudokinases, and some of them have shown the ability to remodel cellular transduction and the transcriptome of the host cell through phosphorylation events (Jacot and Soldati-Favre 2012; Lim et al. 2012). Specifically, the TgROP18 and

730 TgROP17 kinases and the TgROP5 pseudokinase form complexes and by that inactivate
731 host immune responses and inflammation (Du et al. 2014; Etheridge et al. 2014).
732 Moreover, TgROP16 regulates host innate immunity through STAT3 and STAT6
733 phosphorylation (Jensen et al. 2013) and TgROP38 modulates MAPK signaling to
734 control apoptosis and cell proliferation (Peixoto et al. 2010). In our case, NcROP40 has
735 been described as a predicted member of the rhoptry kinase family (ROPK) lacking the
736 key kinase sequence motifs (Talevich and Kannan 2013). The protein contains a
737 structurally conserved N-terminal extension to the kinase domain that displays high
738 sequence similarity to the NcROP5 and TgROP5 pseudokinases, among others.
739 TgROP5 also lacks kinase activity (Reese and Boothroyd 2011), but in contrast to
740 NcROP40, is clearly secreted during invasion and associates with the PVM (El Hajj et
741 al. 2007). Therefore, the role of NcROP40 as pseudokinase remains unclear.
742 Nevertheless, the protein could be implicated in the regulation of still unknown
743 virulence factors. Unfortunately, little is known about the existence of rhoptry virulence
744 factors that could alter the host transcriptome after the infection with *N. caninum*. To
745 date, the information about the orthologues of TgROP5, TgROP16, TgROP18 and
746 TgROP38 in *N. caninum* is limited, and the only study in which they have been
747 described is restricted to genomic and transcriptomic information that highlights the
748 divergence of rhoptry proteins between *T. gondii* and *N. caninum* (Reid et al. 2012).
749 Hence, despite the common features of *N. caninum* and *T. gondii,* these distinct
750 differences in their secreted virulence factors make it difficult to make direct
751 extrapolations from one species to the other. However, the description of common
752 mechanisms of the ROP2-family members required for the success of the lytic cycle and
753 parasite proliferation could represent a valuable source for the development of novel
754 vaccine candidates.

In summary, this study describes highly interesting features of the NcROP40 protein, and another member of the ROP2-family, NcROP2Fam-1, during the lytic cycle of *N. caninum* tachyzoites. Immunogold TEM clearly localized NcROP40 in the rhoptry bulbs of *N. caninum* tachyzoites but, in contrast to NcROP2Fam-1, we were unable to detect NcROP40 secretion into the host cell, which is likely an effect of the protein dilution within the host cytosol. mRNA quantification showed that NcROP40 is highly expressed during egress and invasion, although its mRNA levels were not affected when egress was induced by DTT supplementation. These findings suggest differences in the transcriptional regulation and functional role of NcROP40 and NcROP2Fam-1. In addition, no evidence was found for NcROP40 phosphorylation at the time point of egress, in contrast to NcROP2Fam-1. NcROP40, together with NcROP2Fam-1, is a promising vaccine candidate, thus further studies will be carried out in order to elucidate its functionality. Epitope-tag assays and generation of Δ*rop40* knockout parasites would be useful to confirm more accurately whether NcROP40 is secreted or not, and to establish the role of the NcROP40 protein within the lytic cycle of *N. caninum*.

COMPETING INTERESTS

The authors declare that they have no competing interests. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

JRC, GAG and LMOM conceived and designed the experiments. IPF, EJR, VMH and AH performed the experiments. IPF, JRC, GAG and LMOM analyzed the data. IPF, JRC, EJR, GAG, VMH, AH and LMOM wrote the paper.

ACKNOWLEDGEMENTS

This work was supported by the AGL2010-22191/GAN and the AGL2013-44694-R grants from the Spanish Ministry of Economy and Competitiveness (M.I.N.E.C.O.) and the S2013/ABI2906 grant from the Community of Madrid. Iván Pastor-Fernández was supported by a fellowship from the Spanish Ministry of Education, Culture and Sports (M.E.C.D.), as part of the Program of Training of University Staff (F.P.U., grant number AP2009-0354). Andrew Hemphill was supported by the Swiss National Science Foundation (grant No. 310030 146162). We gratefully acknowledge Prof. Mª Teresa Miras Portugal's group from the Department of Biochemistry and Molecular Biology IV of the Complutense University (Madrid) for their confocal microscopy system. We also thank Dr. Diana Williams from the Liverpool School of Tropical Medicine (Liverpool, UK) for the *N. caninum* Nc-Liv isolate, and Dr. David Sibley from the Washington University School of Medicine (St. Louis, MO, USA) for the NcMIC2 antibody.

805

806

807 **FIGURES**

808

809 Fig. 1. A: Sequence alignment of the ROP40 protein, both in *N. caninum*
810 (NCLIV_012920) and *T. gondii* (TGME49_091960). NcROP40-long is referred to the
811 NcROP40 sequence incorporating and additional 134 aminoacids in the N-terminus. SP:
812 signal peptide. Empty arrow head: potential phosphorylation sites. Filled arrow head:
813 origin of the NcROP40 protein as shown in ToxoDB. Boxes and roman numerals:
814 conserved motifs of likely inactive rhoptry kinase regions as described for Talevich and
815 colleagues in 2013. . B: Comparison of the RAH domains among rhoptry proteins from
816 the ROP2-family. Boxes designate the three domains described for El Hajj and
817 colleagues in 2006. Sequences were obtained from ToxoDB with the following
818 accession numbers: TGME49_005250 (TgROP18), TGME49_108080 (TgROP5),
819 TGME49_095110 (TgROP7), TGME49_015780 (TgROP2A), TGME49_015770
820 (TgROP8), TGME49_091960 (TgROP40) and NCLIV_012920 (NcROP40). For A and
821 B asterisks (*) indicate fully conserved residues, whilst colons (:) and periods (·)
822 indicate conservation between groups of strongly or weakly similar properties,
823 respectively. C: Secondary structure predictions of the NcROP40-long RAH domains
824 by PSSpred, PSIPRED and Jpred3 servers. H: helix. C: coil. Dashes: undefined.

825

826 Fig. 2: *N. caninum*-based Western-blot showing the immuno-reactivity of α-rNcROP40
827 and α-rNcROP2Fam-1 antibodies. Five bands of approximately 53, 44, 38, 32 and 28
828 kDa were detected with α-NcROP40 antibodies, whilst six different bands of
829 approximately 58, 48, 40, 36, 34 and 26 kDa were detected with α-NcROP2Fam-1
830 antibodies. B: Phosphorylation detection of NcROP40 and NcROP2Fam-1 by Phos-Tag
831 SDS-PAGE. Tachyzoite extracts obtained at 56 hpi were processed with alkaline

phosphatase (CIP) and phosphatase inhibitors (PI), electrophoresed on SDS-PAGE and Phos-Tag SDS-PAGE gels, and blotted to nitrocellulose membranes. Then, both proteins were detected by their respective antibodies in order to detect a mobility shift of the proteins treated with PI.

Fig. 3. NcROP40 is a rhoptry protein associated with rhoptry bulbs. Transmission electron microscopy and immunogold staining in tachyzoites (A-B) and bradyzoites (C-D). Rhoptries (rh), dense granules (dg) and micronemes (mic) are indicated on the pictures. Bars represent 1 μm.

Fig. 4. Confocal laser scanning microscopy of NcROP40 along the lytic cycle of tachyzoites. Infected cultures were fixed with metanol (MeOH), paraformaldehyde (PFA) and paraformaldehyde combined with glutaraldehyde (PFA+GA) and double labelled with affinity purified antibodies against NcROP40 (red) and monoclonal antibodies against NcSAG1 (green). Nuclei were stained with DAPI (blue). All the images show a single 1 μm slice. Bars represent 4 μm.

Fig. 5. Confocal laser scanning microscopy of NcROP2Fam-1 along the lytic cycle of tachyzoites. Infected cultures were fixed with metanol (MeOH), paraformaldehyde (PFA) and paraformaldehyde combined with glutaraldehyde (PFA+GA) and double labelled with affinity purified antibodies against NcROP2Fam-1 (red) and monoclonal antibodies against NcSAG1 (green). Nuclei were stained with DAPI (blue). PFA-fixed cultures were also labelled with phalloidin to delimitate host-cell surface (white). All the images show a single 1 μm slice. Bars represent 4 μm.

Fig. 6. mRNA expression of NcROP40 and NcROP2Fam-1. Real time-PCR was employed to assess the mRNA expression of both proteins along the lytic cycle. Top panel: photomicrographs showing the infection dynamics of the Nc-Liv isolate on MARC-145 cultures at recent invasion (6 h), PV formation and maturation (24 h), exponential growth of parasites (48 h) and tachyzoite egress (56 h and 48 h + DTT). A: mRNA expression levels of NcROP40 and NcROP2Fam-1 during the lytic cycle. B: Effect of DTT supplementation to artificially induce egress at 48 h on mRNA expression for both proteins. For A and B, each point represents a single sample and bars represent the mean value. *a*, *b* and *c* indicate significant differences ($P<0.005$; Kruskal-Wallis test).

Fig. 7: Phosphorylation detection of NcROP40 and NcROP2Fam-1 by Phos-Tag SDS-PAGE. Tachyzoite extracts obtained at 56 hpi were processed with alkaline phosphatase (CIP) and phosphatase inhibitors (PI), electrophoresed on SDS-PAGE and Phos-Tag SDS-PAGE gels, and blotted to nitrocellulose membranes. Then, both proteins were detected by their respective antibodies in order to detect a mobility shift of the proteins treated with PI.

882

883 **TABLES**

884

885 **Table 1:** Primers used to amplify Nc*ROP40*, Nc*ROP2Fam-1*, Nc*SAG1* and Nc*Tubilin alpha* sequences by real time-PCR.

Protein	ToxoDB accession number	Primer sequences	Reference	Length	Introns [a]	Slope [b]	R^2 [b]
NcROP40	NCLIV_012920	Fw-CATCAAGCAGCCCAGAATCA Rv-TGGTGACTGCGACCAACTTA	This study	94 bp. 1021-1114	No	-3.71	0.996
NcROP2Fam-1	NCLIV_001970	Fw-TTCTTCCTCTCCAAGCGACA Rv-TTGAGTCGTTCCCGAAGTTG	Alaeddine et al., 2013	140 bp. 1604-1743	No	-3.68	0.996
NcSAG1	NCLIV_033230	Fw-CGGTGTCGCAATGTGCTCTT Rv-ACGGTCGTCCCAGAACAAAC	Fernández-García et al., 2006	150 bp. 504-653	No	-3.24	0.997
NcTUBα	NCLIV_058890	Fw-GGTAACGCCTGCTGGGAG Rv-GCTCCAAATCCAAGAAGACGCA	Alaeddine et al., 2013	166 bp. 49-214	Yes*	-3.24	0.994

886

887 [a] Primers for intron-containing sequences were designed using cDNA as template. * Forward primer for NcTUBα amplification annealed at

888 intron splice junction to prevent amplification of genomic DNA.

889 [b] Descriptive values of real time-PCR from standard curves for each pair of primers are shown.

890

891

892 **REFERENCES**

893

894 **Alaeddine F., Hemphill A., Debache K. and Guionaud C.** (2013). Molecular cloning and
895 characterization of NcROP2Fam-1, a member of the ROP2 family of rhoptry proteins in
896 *Neospora caninum* that is targeted by antibodies neutralizing host cell invasion *in vitro*.
897 *Parasitology* **140**, 1033-1050.

898 **Álvarez-García G., Pitarch A., Zaballos A., Fernández-García A., Gil C., Gómez-**
899 **Bautista M., Aguado-Martínez A. and Ortega-Mora L. M.** (2007). The NcGRA7 gene
900 encodes the immunodominant 17 kDa antigen of *Neospora caninum*. *Parasitology* **134**, 41-
901 50.

902 **Barber J., Trees A. J., Owen M. and Tennant B.** (1993). Isolation of *Neospora caninum*
903 from a British dog. *The Veterinary Record* **133**, 531-532.

904 **Beck J. R., Chen A. L., Kim E. W. and Bradley P. J.** (2014). RON5 is critical for
905 organization and function of the *Toxoplasma* moving junction complex. *PLoS Pathogens* **10**,
906 e1004025.

907 **Beckers C. J., Dubremetz J. F., Mercereau-Puijalon O. and Joiner K. A.** (1994). The
908 *Toxoplasma gondii* rhoptry protein ROP2 is inserted into the parasitophorous vacuole
909 membrane, surrounding the intracellular parasite, and is exposed to the host cell cytoplasm.
910 *The Journal of Cell Biology* **127**, 947-961.

911 **Behnke M. S., Wootton J. C., Lehmann M. M., Radke J. B., Lucas O., Nawas J., Sibley**
912 **L. D. and White M. W.** (2010). Coordinated progression through two subtranscriptomes
913 underlies the tachyzoite cycle of *Toxoplasma gondii*. *PloS One* **5**, e12354.

914 **Björkman C. and Hemphill A.** (1998). Characterization of *Neospora caninum* iscom
915 antigens using monoclonal antibodies. *Parasite Immunology* **20**, 73-80.

916 **Blackman M. J. and Carruthers V. B.** (2013). Recent insights into apicomplexan parasite
917 egress provide new views to a kill. *Current Opinion in Microbiology* **16**, 459-464.

918 **Boothroyd J. C. and Dubremetz J. F.** (2008). Kiss and spit: the dual roles of *Toxoplasma*
919 rhoptries. *Nature Reviews.Microbiology* **6**, 79-88.

920 **Bradley P. J. and Sibley L. D.** (2007). Rhoptries: an arsenal of secreted virulence factors.
921 *Current Opinion in Microbiology* **10**, 582-587.

922 **Carey K. L., Jongco A. M., Kim K. and Ward G. E.** (2004). The *Toxoplasma gondii*
923 rhoptry protein ROP4 is secreted into the parasitophorous vacuole and becomes
924 phosphorylated in infected cells. *Eukaryotic Cell* **3**, 1320-1330.

925 **Carruthers V. B. and Sibley L. D.** (1997). Sequential protein secretion from three distinct
926 organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. *European*
927 *Journal of Cell Biology* **73**, 114-123.

928 **Carruthers V. B. and Sibley L. D.** (1999). Mobilization of intracellular calcium stimulates
929 microneme discharge in *Toxoplasma gondii*. *Molecular Microbiology* **31**, 421-428.

930 **Cowman A. F., Berry D. and Baum J.** (2012). The cellular and molecular basis for malaria
931 parasite invasion of the human red blood cell. *The Journal of Cell Biology* **198**, 961-971.

932 **Debache K., Guionaud C., Alaeddine F., Mevissen M. and Hemphill A.** (2008).
933 Vaccination of mice with recombinant NcROP2 antigen reduces mortality and cerebral

infection in mice infected with *Neospora caninum* tachyzoites. *International Journal for Parasitology* **38**, 1455-1463.

Debache K., Alaeddine F., Guionaud C., Monney T., Muller J., Strohbusch M., Leib S. L., Grandgirard D. and Hemphill A. (2009). Vaccination with recombinant NcROP2 combined with recombinant NcMIC1 and NcMIC3 reduces cerebral infection and vertical transmission in mice experimentally infected with *Neospora caninum* tachyzoites. *International Journal for Parasitology* **39**, 1373-1384.

Debache K., Guionaud C., Alaeddine F. and Hemphill A. (2010). Intraperitoneal and intra-nasal vaccination of mice with three distinct recombinant *Neospora caninum* antigens results in differential effects with regard to protection against experimental challenge with *Neospora caninum* tachyzoites. *Parasitology* **137**, 229-240.

Dixon S. E., Stilger K. L., Elias E. V., Naguleswaran A. and Sullivan W. J.,Jr. (2010). A decade of epigenetic research in *Toxoplasma gondii*. *Molecular and Biochemical Parasitology* **173**, 1-9.

Du J., An R., Chen L., Shen Y., Chen Y., Cheng L., Jiang Z., Zhang A., Yu L., Chu D., Shen Y., Luo Q., Chen H., Wan L., Li M., Xu X. and Shen J. (2014). *Toxoplasma gondii* virulence factor ROP18 inhibits the host NF-kappaB pathway by promoting p65 degradation. *The Journal of Biological Chemistry* **289**, 12578-12592.

Dubey J. P. and Schares G. (2011). Neosporosis in animals-The last five years. *Veterinary Parasitology* **180**, 90-108.

Dunn J. D., Ravindran S., Kim S. K. and Boothroyd J. C. (2008). The *Toxoplasma gondii* dense granule protein GRA7 is phosphorylated upon invasion and forms an unexpected

association with the rhoptry proteins ROP2 and ROP4. *Infection and Immunity* **76**, 5853-5861.

El Hajj H., Demey E., Poncet J., Lebrun M., Wu B., Galeotti N., Fourmaux M. N., Mercereau-Puijalon O., Vial H., Labesse G. and Dubremetz J. F. (2006). The ROP2 family of *Toxoplasma gondii* rhoptry proteins: proteomic and genomic characterization and molecular modeling. *Proteomics* **6**, 5773-5784.

El Hajj H., Lebrun M., Fourmaux M. N., Vial H. and Dubremetz J. F. (2007). Inverted topology of the *Toxoplasma gondii* ROP5 rhoptry protein provides new insights into the association of the ROP2 protein family with the parasitophorous vacuole membrane. *Cellular Microbiology* **9**, 54-64.

Esposito M., Moores S., Naguleswaran A., Muller J. and Hemphill A. (2007). Induction of tachyzoite egress from cells infected with the protozoan *Neospora caninum* by nitro- and bromo-thiazolides, a class of broad-spectrum anti-parasitic drugs 9. *International Journal for Parasitology* **37**, 1143-1152.

Etheridge R. D., Alaganan A., Tang K., Lou H. J., Turk B. E. and Sibley L. D. (2014). The *Toxoplasma* pseudokinase ROP5 forms complexes with ROP18 and ROP17 kinases that synergize to control acute virulence in mice. *Cell Host & Microbe* **15**, 537-550.

Gaji R. Y., Behnke M. S., Lehmann M. M., White M. W., Carruthers V. B. (2011). Cell cycle-dependent, intercellular transmission of *Toxoplasma gondii* is accompanied by marked changes in parasite gene expression. *Molecular Microbiology* **79**, 192-204.

Hajagos B. E., Turetzky J. M., Peng E. D., Cheng S. J., Ryan C. M., Souda P., Whitelegge J. P., Lebrun M., Dubremetz J. F. and Bradley P. J. (2012). Molecular

dissection of novel trafficking and processing of the *Toxoplasma gondii* rhoptry metalloprotease toxolysin-1. *Traffic* **13**, 292-304.

Hemphill A., Vonlaufen N., Naguleswaran A., Keller N., Riesen M., Guetg N., Srinivasan S. and Alaeddine F. (2004). Tissue culture and explant approaches to studying and visualizing *Neospora caninum* and its interactions with the host cell. *Microscopy and Microanalysis* **10**, 602-620.

Hemphill A., Debache K., Monney T., Schorer M., Guionaud C., Alaeddine F., Mueller N. and Mueller J. (2013). Proteins mediating the *Neospora caninum*-host cell interaction as targets for vaccination. *Frontiers in Bioscience (Elite edition)* **5**, 23-36.

Jacot D. and Soldati-Favre D. (2012). Does protein phosphorylation govern host cell entry and egress by the Apicomplexa? *International Journal of Medical Microbiology* **302**, 195-202.

Jensen K. D., Hu K., Whitmarsh R. J., Hassan M. A., Julien L., Lu D., Chen L., Hunter C. A. and Saeij J. P. (2013). *Toxoplasma gondii* rhoptry 16 kinase promotes host resistance to oral infection and intestinal inflammation only in the context of the dense granule protein GRA15. *Infection and Immunity* **81**, 2156-2167.

Kemp L. E., Yamamoto M. and Soldati-Favre D. (2013). Subversion of host cellular functions by the apicomplexan parasites. *FEMS Microbiology Reviews* **37**, 607-631.

Kinoshita E., Kinoshita-Kikuta E., Takiyama K. and Koike T. (2006). Phosphate-binding tag, a new tool to visualize phosphorylated proteins. *Molecular & Cellular Proteomics* **5**, 749-757.

Labruyere E., Lingnau M., Mercier C. and Sibley L. D. (1999). Differential membrane targeting of the secretory proteins GRA4 and GRA6 within the parasitophorous vacuole formed by *Toxoplasma gondii*. *Molecular and Biochemical Parasitology* **102**, 311-324.

Le Roch K. G., Johnson J. R., Florens L., Zhou Y., Santrosyan A., Grainger M., Yan S. F., Williamson K. C., Holder A. A., Carucci D. J., Yates J. R. and Winzeler E. A. (2004). Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Research* **14**, 2308-2318.

Lei T., Wang H., Liu J., Nan H. and Liu Q. (2014). ROP18 is a key factor responsible for virulence difference between *Toxoplasma gondii* and *Neospora caninum*. *PloS One* **9**, e99744.

Lescault P. J., Thompson A. B., Patil V., Lirussi D., Burton A., Margarit J., Bond J., Matrajt M. (2010). Genomic data reveal *Toxoplasma gondii* differentiation mutants are also impaired with respect to switching into a novel extracellular tachyzoite state. *PLoS One* **12,** e14463.

Lim D. C., Cooke B. M., Doerig C. and Saeij J. P. (2012). *Toxoplasma* and *Plasmodium* protein kinases: roles in invasion and host cell remodelling. *International Journal for Parasitology* **42**, 21-32.

Llinas M. and DeRisi J. L. (2004). Pernicious plans revealed: *Plasmodium falciparum* genome wide expression analysis. *Current Opinion in Microbiology* **7**, 382-387.

Lodoen M. B., Gerke C., Boothroyd J. C. (2009). A highly sensitive FRET-based approach reveals secretion of the actin-binding protein toxofilin during *Toxoplasma gondii* infection. *Cellular Microbiology* **12**, 55-66.

Lovett J. L., Howe D. K. and Sibley L. D. (2000). Molecular characterization of a thrombospondin-related anonymous protein homologue in *Neospora caninum*. *Molecular and Biochemical Parasitology* **107**, 33-43.

Marugán-Hernández V., Álvarez-García G., Tomley F., Hemphill A., Regidor-Cerrillo J. and Ortega-Mora L. M. (2011). Identification of novel rhoptry proteins in *Neospora caninum* by LC/MS-MS analysis of subcellular fractions. *Journal of Proteomics* **74**, 629-642.

Mercier C., Adjogble K. D., Daubener W. and Delauw M. F. (2005). Dense granules: are they key organelles to help understand the parasitophorous vacuole of all apicomplexa parasites? *International Journal for Parasitology* **35**, 829-849.

Naguleswaran A., Cannas A., Keller N., Vonlaufen N., Schares G., Conraths F. J., Bjorkman C. and Hemphill A. (2001). *Neospora caninum* microneme protein NcMIC3: secretion, subcellular localization, and functional involvement in host cell interaction. *Infection and Immunity* **69**, 6483-6494.

Nakaar V., Ngô H. M., Aaronson E. P., Coppens I., Stedman T. T., Joiner K. A. (2003). Pleiotropic effect due to targeted depletion of secretory rhoptry protein ROP2 in *Toxoplasma gondii*. *Journal of Cell Science* **116**, 2311-2320.

Pastor-Fernández I., Arranz-Solís D., Regidor-Cerrillo J., Álvarez-García G., Hemphill A., García-Culebras A., Cuevas-Martín C. and Ortega-Mora L. M. (2015). A vaccine formulation combining rhoptry proteins NcROP40 and NcROP2 improves pup survival in a pregnant mouse model of neosporosis. *Veterinary Parasitology* **207**, 203-215.

Peixoto L., Chen F., Harb O. S., Davis P. H., Beiting D. P., Brownback C. S., Ouloguem D. and Roos D. S. (2010). Integrative genomic approaches highlight a family of parasite-specific kinases that regulate host responses. *Cell Host & Microbe* **8**, 208-218.

Pereira García-Melo D., Regidor-Cerrillo J., Collantes-Fernández E., Aguado-Martínez A., Del Pozo I., Minguijon E., Gómez-Bautista M., Aduriz G. and Ortega-Mora L. M. (2010). Pathogenic characterization in mice of *Neospora caninum* isolates obtained from asymptomatic calves. *Parasitology* **137**, 1057-1068.

Pérez-Zaballos F. J., Ortega-Mora L. M., Álvarez-García G., Collantes-Fernández E., Navarro-Lozano V., García-Villada L. and Costas E. (2005). Adaptation of *Neospora caninum* isolates to cell-culture changes: an argument in favor of its clonal population structure. *The Journal of Parasitology* **91**, 507-510.

Pittman K. J., Aliota M. T. and Knoll L. J. (2014). Dual transcriptional profiling of mice and *Toxoplasma gondii* during acute and chronic infection. *BMC Genomics* **15**, 806-2164-15-806.

Pollo-Oliveira L., Post H., Acencio M. L., Lemke N., van den Toorn H., Tragante V., Heck A. J., Altelaar A. F. and Yatsuda A. P. (2013). Unravelling the *Neospora caninum* secretome through the secreted fraction (ESA) and quantification of the discharged tachyzoite using high-resolution mass spectrometry-based proteomics. *Parasites & Vectors* **6**, 335.

Radke J. R., Behnke M. S., Mackey A. J., Radke J. B., Roos D. S. and White M. W. (2005). The transcriptome of *Toxoplasma gondii*. *BMC Biology* **3**, 26.

Ramaprasad A., Mourier T., Naeem R., Malas T. B., Moussa E., Panigrahi A., Vermont S. J., Otto T. D., Wastling J. and Pain A. (2015). Comprehensive evaluation of *Toxoplasma*

gondii VEG and *Neospora caninum* LIV genomes with tachyzoite stage transcriptome and proteome defines novel transcript features. *PloS One* **10**, e0124473.

Reese M. L. and Boothroyd J. C. (2009). A helical membrane-binding domain targets the *Toxoplasma* ROP2 family to the parasitophorous vacuole. *Traffic* **10**, 1458-1470.

Reese M. L. and Boothroyd J. C. (2011). A conserved non-canonical motif in the pseudoactive site of the ROP5 pseudokinase domain mediates its effect on *Toxoplasma* virulence. *The Journal of Biological Chemistry* **286**, 29366-29375.

Reese M. L., Shah N. and Boothroyd J. C. (2014). The *Toxoplasma* pseudokinase ROP5 is an allosteric inhibitor of the immunity-related GTPases. *The Journal of Biological Chemistry* **289**, 27849-27858.

Regidor-Cerrillo J., Gómez-Bautista M., Pereira-Bueno J., Adúriz G., Navarro-Lozano V., Risco-Castillo V., Fernández-García A., Pedraza-Díaz S. and Ortega-Mora L. M. (2008). Isolation and genetic characterization of *Neospora caninum* from asymptomatic calves in Spain. *Parasitology* **135**, 1651-1659.

Regidor-Cerrillo J., Gómez-Bautista M., Del Pozo I., Jiménez-Ruiz E., Aduriz G. and Ortega-Mora L. M. (2010). Influence of *Neospora caninum* intra-specific variability in the outcome of infection in a pregnant BALB/c mouse model. *Veterinary Research* **41**, 52.

Regidor-Cerrillo J., Gómez-Bautista M., Sodupe I., Aduriz G., Álvarez-García G., Del Pozo I. and Ortega-Mora L. M. (2011). *In vitro* invasion efficiency and intracellular proliferation rate comprise virulence-related phenotypic traits of *Neospora caninum*. *Veterinary Research* **42**, 41.

Regidor-Cerrillo J., Álvarez-García G., Pastor-Fernández I., Marugán-Hernández V., Gómez-Bautista M. and Ortega-Mora L. M. (2012). Proteome expression changes among virulent and attenuated *Neospora caninum* isolates. *Journal of Proteomics* **75**, 2306-2318.

Reid A. J., Vermont S. J., Cotton J. A., Harris D., Hill-Cawthorne G. A., Konen-Waisman S., Latham S. M., Mourier T., Norton R., Quail M. A., Sanders M., Shanmugam D., Sohal A., Wasmuth J. D., Brunk B., Grigg M. E., Howard J. C., Parkinson J., Roos D. S., Trees A. J., Berriman M., Pain A. and Wastling J. M. (2012). Comparative genomics of the apicomplexan parasites *Toxoplasma gondii* and *Neospora caninum*: Coccidia differing in host range and transmission strategy. *PLoS Pathogens* **8**, e1002567.

Risco-Castillo V., Fernández-García A., Zaballos A., Aguado-Martínez A., Hemphill A., Rodríguez-Bertos A., Álvarez-García G. and Ortega-Mora L. M. (2007). Molecular characterisation of BSR4, a novel bradyzoite-specific gene from *Neospora caninum*. *International Journal for Parasitology* **37**, 887-896.

Rojo-Montejo S., Collantes-Fernández E., Blanco-Murcia J., Rodríguez-Bertos A., Risco-Castillo V. and Ortega-Mora L. M. (2009). Experimental infection with a low virulence isolate of *Neospora caninum* at 70 days gestation in cattle did not result in foetopathy. *Veterinary Research* **40**, 49. **Schmittgen T. D. and Livak K. J.** (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nature Protocols* **3**, 1101-1108.

Schneider A. G., Abi Abdallah D. S., Butcher B. A. and Denkers E. Y. (2013). *Toxoplasma gondii* triggers phosphorylation and nuclear translocation of dendritic cell

STAT1 while simultaneously blocking IFNgamma-induced STAT1 transcriptional activity. *PloS One* **8**, e60215.

Sinai A. P. and Joiner K. A. (2001). The *Toxoplasma gondii* protein ROP2 mediates host organelle association with the parasitophorous vacuole membrane. *The Journal of Cell Biology* **154**, 95-108.

Sohn C. S., Cheng T. T., Drummond M. L., Peng E. D., Vermont S. J., Xia D., Cheng S. J., Wastling J. M. and Bradley P. J. (2011). Identification of novel proteins in *Neospora caninum* using an organelle purification and monoclonal antibody approach. *PloS One* **6**, e18383.

Steinfeldt T., Konen-Waisman S., Tong L., Pawlowski N., Lamkemeyer T., Sibley L. D., Hunn J. P. and Howard J. C. (2010). Phosphorylation of mouse immunity-related GTPase (IRG) resistance proteins is an evasion strategy for virulent *Toxoplasma gondii*. *PLoS Biology* **8**, e1000576.

Straub K. W., Cheng S. J., Sohn C. S. and Bradley P. J. (2009). Novel components of the Apicomplexan moving junction reveal conserved and coccidia-restricted elements. *Cellular Microbiology* **11**doi: 10.1111/j.1462-5822.2008.01276.x.

Talevich E. and Kannan N. (2013). Structural and evolutionary adaptation of rhoptry kinases and pseudokinases, a family of coccidian virulence factors. *BMC Evolutionary Biology* **13**, 117.

Vonlaufen N., Muller N., Keller N., Naguleswaran A., Bohne W., McAllister M. M., Bjorkman C., Muller E., Caldelari R. and Hemphill A. (2002). Exogenous nitric oxide

triggers *Neospora caninum* tachyzoite-to-bradyzoite stage conversion in murine epidermal keratinocyte cell cultures. *International Journal for Parasitology* **32**, 1253-1265.

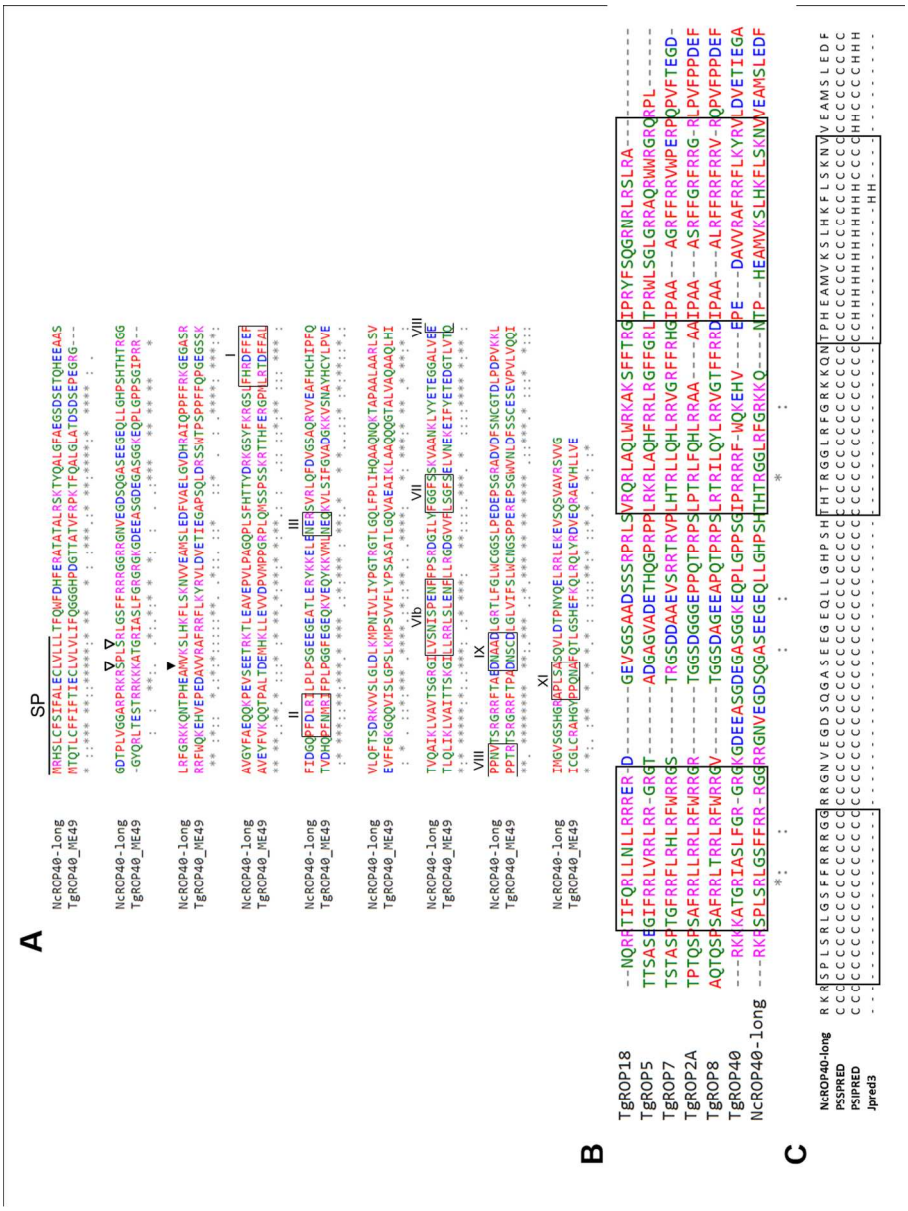


Fig. 1. A: Sequence alignment of the ROP40 protein, both in N. caninum (NCLIV_012920) and T. gondii (TGME49_091960). NcROP40-long is referred to the NcROP40 sequence incorporating and additional 134 aminoacids in the N-terminus. SP: signal peptide. Empty arrow head: potential phosphorylation sites. Filled arrow head: origin of the NcROP40 protein as shown in ToxoDB. Boxes and roman numerals: conserved motifs of likely inactive rhoptry kinase regions as described for Talevich and colleagues in 2013. . B: Comparison of the RAH domains among rhoptry proteins from the ROP2-family. Boxes designate the three domains described for El Hajj and colleagues in 2006. Sequences were obtained from ToxoDB with the following accession numbers: TGME49_005250 (TgROP18), TGME49_108080 (TgROP5), TGME49_095110 (TgROP7), TGME49_015780 (TgROP2A), TGME49_015770 (TgROP8), TGME49_091960 (TgROP40) and NCLIV_012920 (NcROP40). For A and B asterisks (*) indicate fully conserved residues, whilst colons (:) and periods (·) indicate conservation between groups of strongly or weakly similar properties, respectively. C: Secondary structure predictions of the NcROP40-long RAH domains by PSSpred, PSIPRED and Jpred3 servers. H: helix. C: coil. Dashes: undefined.

190x254mm (300 x 300 DPI)

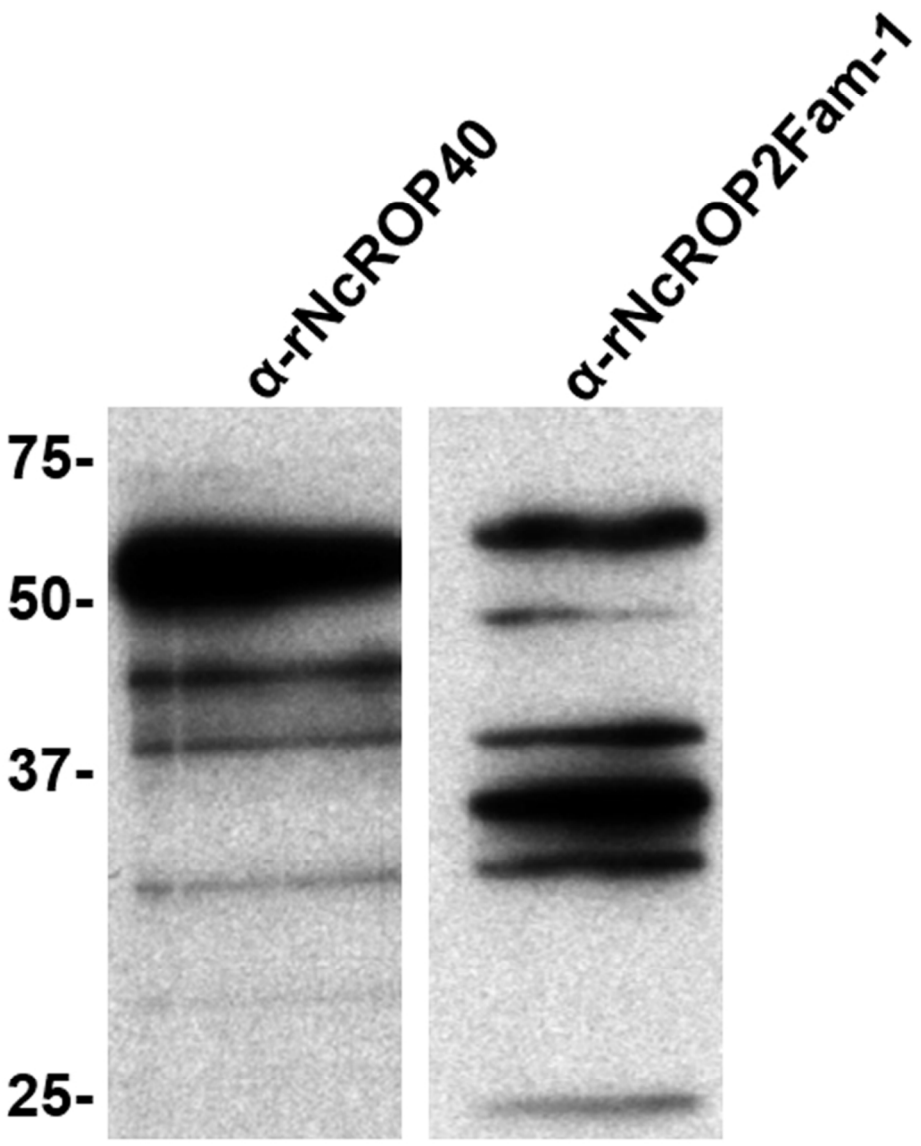


Fig. 2: N. caninum-based Western-blot showing the immuno-reactivity of α-rNcROP40 and α-rNcROP2Fam-1 antibodies. Five bands of approximately 53, 44, 38, 32 and 28 kDa were detected with α-NcROP40 antibodies, whilst six different bands of approximately 58, 48, 40, 36, 34 and 26 kDa were detected with α-NcROP2Fam-1 antibodies. B: Phosphorylation detection of NcROP40 and NcROP2Fam-1 by Phos-Tag SDS-PAGE. Tachyzoite extracts obtained at 56 hpi were processed with alkaline phosphatase (CIP) and phosphatase inhibitors (PI), electrophoresed on SDS-PAGE and Phos-Tag SDS-PAGE gels, and blotted to nitrocellulose membranes. Then, both proteins were detected by their respective antibodies in order to detect a mobility shift of the proteins treated with PI.
79x98mm (300 x 300 DPI)

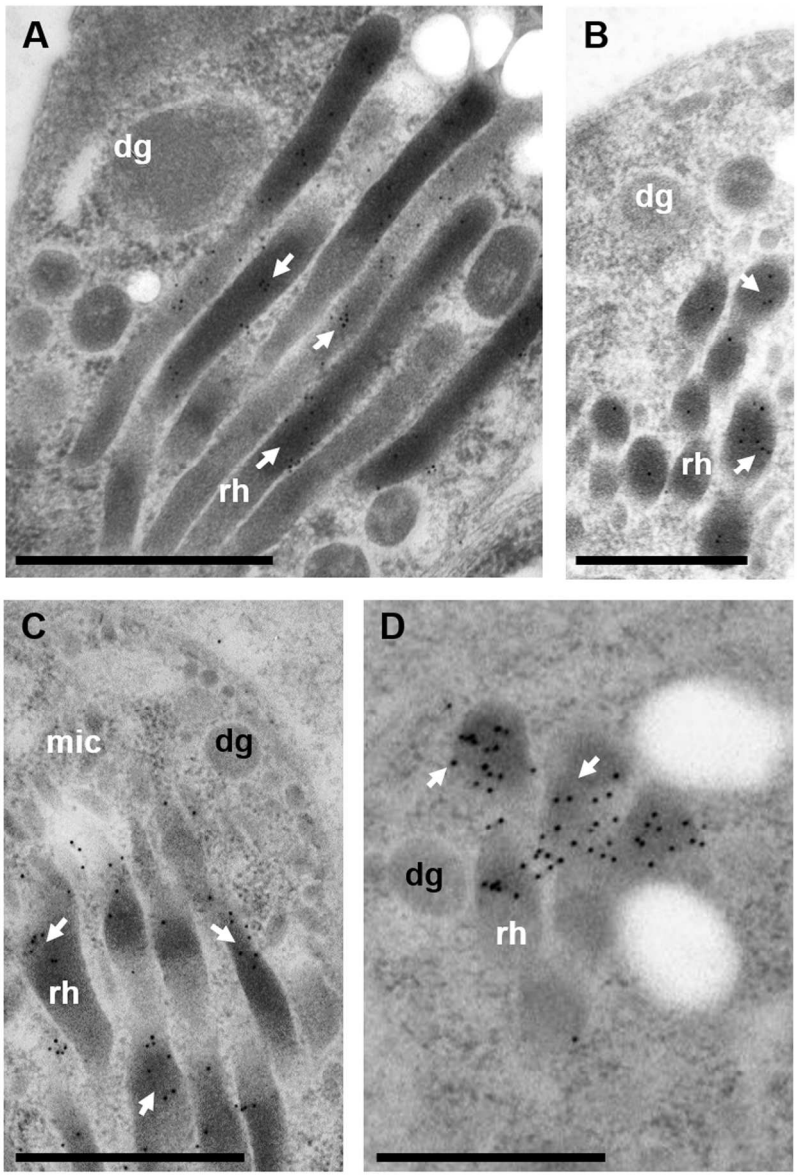


Fig. 3. NcROP40 is a rhoptry protein associated with rhoptry bulbs. Transmission electron microscopy and immunogold staining in tachyzoites (A-B) and bradyzoites (C-D). Rhoptries (rh), dense granules (dg) and micronemes (mic) are indicated on the pictures. Bars represent 1 µm.
145x209mm (300 x 300 DPI)

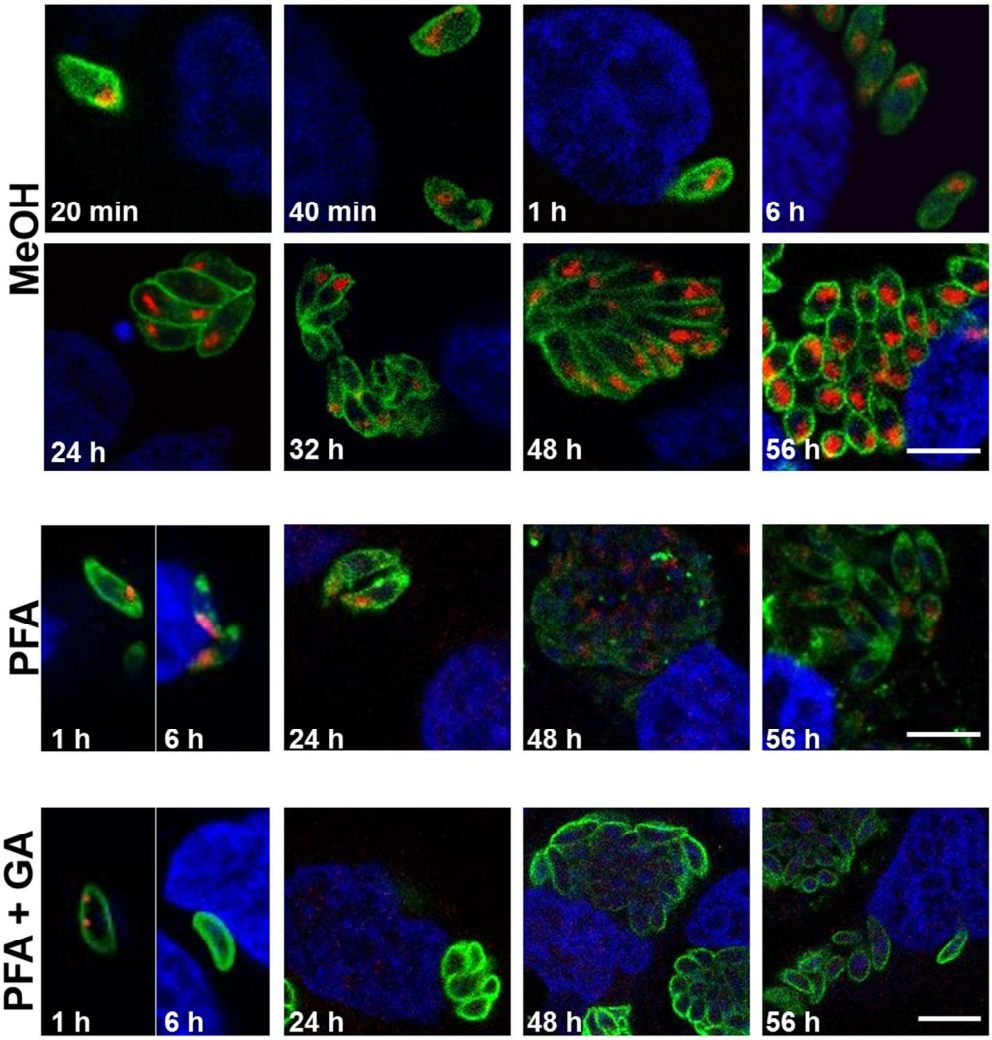


Fig. 4. Confocal laser scanning microscopy of NcROP40 along the lytic cycle of tachyzoites. Infected cultures were fixed with metanol (MeOH), paraformaldehyde (PFA) and paraformaldehyde combined with glutaraldehyde (PFA+GA) and double labelled with affinity purified antibodies against NcROP40 (red) and monoclonal antibodies against NcSAG1 (green). Nuclei were stained with DAPI (blue). All the images show a single 1 µm slice. Bars represent 4 µm.
176x185mm (300 x 300 DPI)

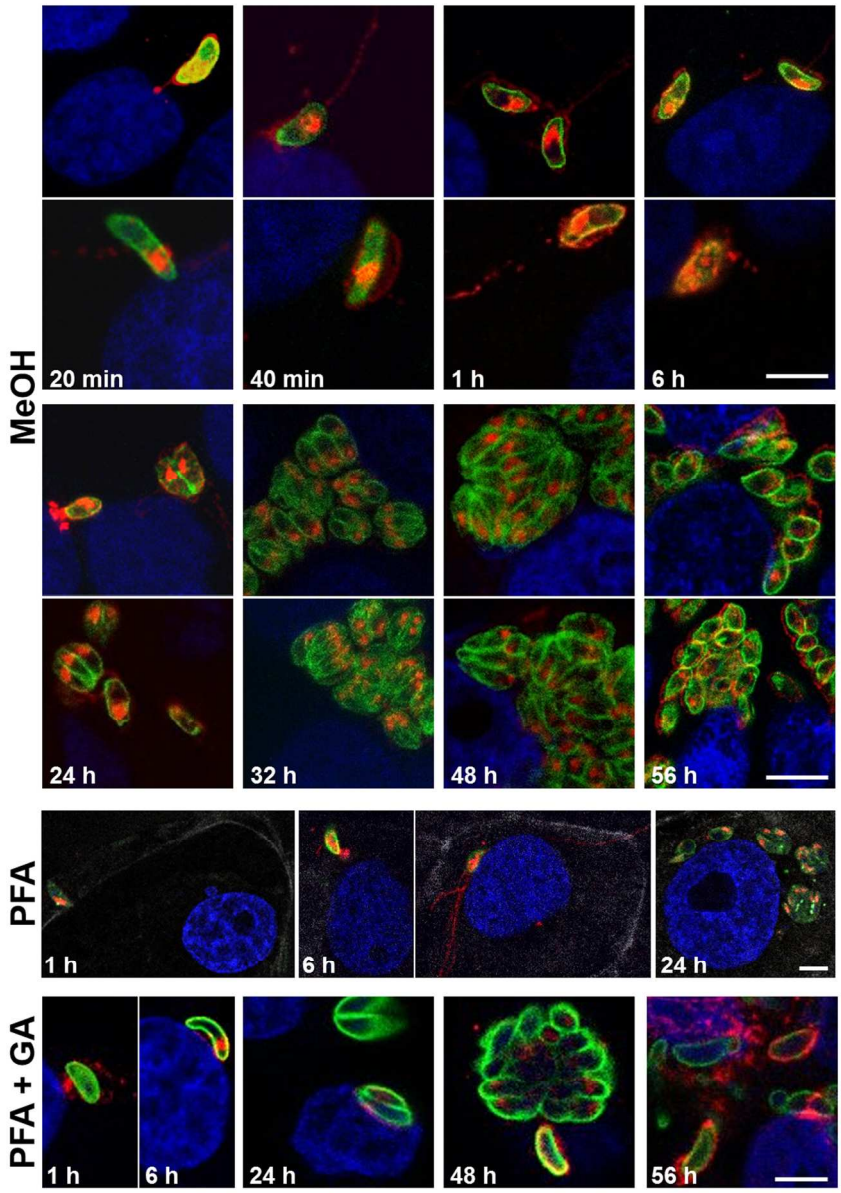


Fig. 5. Confocal laser scanning microscopy of NcROP2Fam-1 along the lytic cycle of tachyzoites. Infected cultures were fixed with metanol (MeOH), paraformaldehyde (PFA) and paraformaldehyde combined with glutaraldehyde (PFA+GA) and double labelled with affinity purified antibodies against NcROP2Fam-1 (red) and monoclonal antibodies against NcSAG1 (green). Nuclei were stained with DAPI (blue). PFA-fixed cultures were also labelled with phalloidin to delimitate host-cell surface (white). All the images show a single 1 μm slice. Bars represent 4 μm.
176x251mm (300 x 300 DPI)

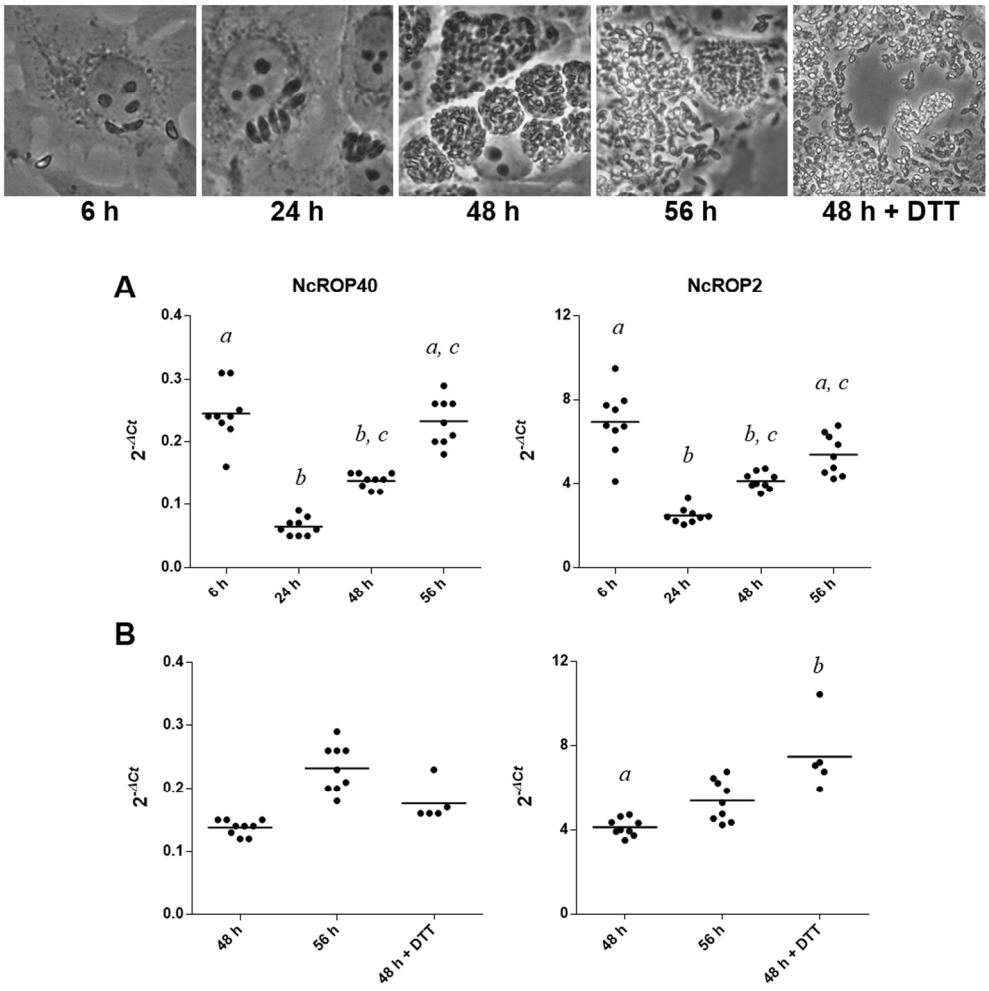


Fig. 6. mRNA expression of NcROP40 and NcROP2Fam-1. Real time-PCR was employed to assess the mRNA expression of both proteins along the lytic cycle. Top panel: photomicrographs showing the infection dynamics of the Nc-Liv isolate on MARC-145 cultures at recent invasion (6 h), PV formation and maturation (24 h), exponential growth of parasites (48 h) and tachyzoite egress (56 h and 48 h + DTT). A: mRNA expression levels of NcROP40 and NcROP2Fam-1 during the lytic cycle. B: Effect of DTT supplementation to artificially induce egress at 48 h on mRNA expression for both proteins. For A and B, each point represents a single sample and bars represent the mean value. a, b and c indicate significant differences ($P<0.005$; Kruskal-Wallis test).
182x182mm (300 x 300 DPI)

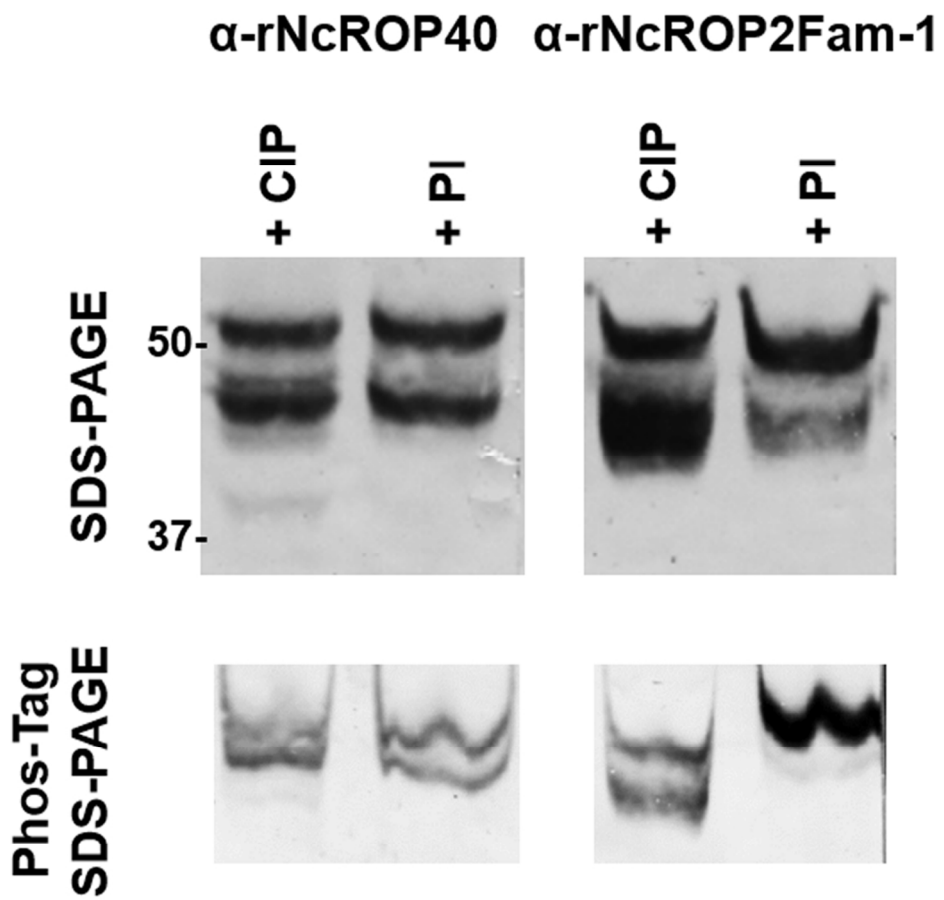


Fig. 7: Phosphorylation detection of NcROP40 and NcROP2Fam-1 by Phos-Tag SDS-PAGE. Tachyzoite extracts obtained at 56 hpi were processed with alkaline phosphatase (CIP) and phosphatase inhibitors (PI), electrophoresed on SDS-PAGE and Phos-Tag SDS-PAGE gels, and blotted to nitrocellulose membranes. Then, both proteins were detected by their respective antibodies in order to detect a mobility shift of the proteins treated with PI.
100x97mm (300 x 300 DPI)

Additional files

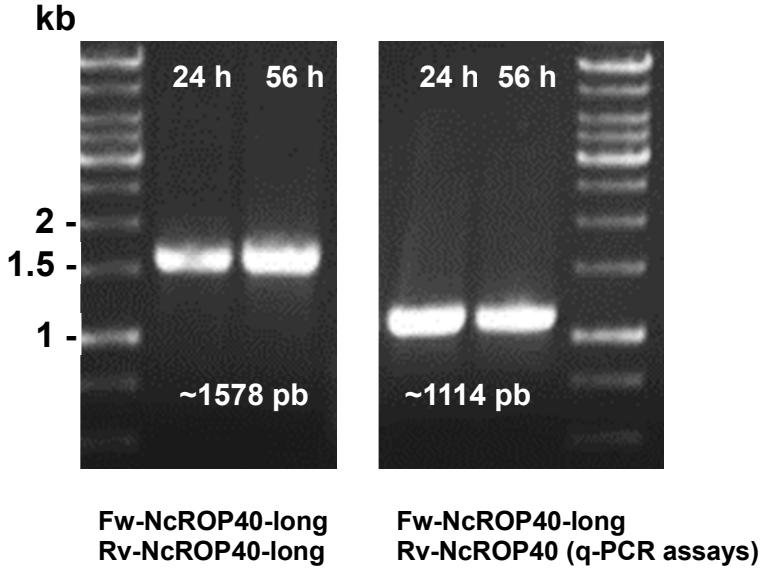
Additional file 1: Primers used for NcROP40 sequencing among three *N. caninum* isolates of different origins (Nc-Liv, Nc-Spain7 and Nc-Spain1H).

Primer	Sequences	Reference
Fw-chrV_ROP40	5'-TAAGAACGCATGGCTGACTG-3'	This study
Rv-chrV_ROP40	5'-CTGACAACGGCTCCTCTTTC-3'	This study
Fw-ROP40	5'-CGAGCTCATGGTGAAATCCCTGCACAAG-3'	Regidor-Cerrillo et al., 2012
Rv-ROP40	5'-CCTTAATTAATCACCCCACCACTGAACGC-3'	Regidor-Cerrillo et al., 2012
Rv-ROP40-int	5'-CTTCTGGCTTTTGCTGCTCC-3'	This study
Fw-ROP40-int	5'-ACGCACTCTCTTTGGCTTGT-3'	This study

Additional file 2: Primers used for NcROP40, NcROP2Fam-1, NcSAG1 and NcTUBα cloning.

Protein	ToxoDB accession number	Primer sequences	Reference	Length
NcROP40	NCLIV_012920	Fw-*Sac*I-CGAGCTCATGGTGAAATCCCTGCACAAG Rv-*Pac*I-CCTTAATTAATCACCCCACCACTGAACGC	Regidor-Cerrillo et al., 2012	1176 bp. 1-392 aa.
NcROP2Fam-1	NCLIV_001970	Fw-*Sac*I-GAGCTCTTGTGGCGTAATCAGAAGCAC Rv-*Hind*III-AAGCTTTTATAGCCTCGTGTCCTCCGT	Pastor-Fernández et al., 2015	1076 bp. 236-595 aa.
NcSAG1	NCLIV_033230	Fw-CACTGGTGGCGTTCTTTGAC Rv-GCTATCGAGCCTACGAGTCC	This study	889 bp. 56-944 aa.
NcTUBα	NCLIV_058890	Fw-GGTAACGCCTGCTGGGAG Rv-CTTCCTCTTCACCTTCGCCC	Alaeddine et al., 2013	1294 bp. 49-1342 aa.

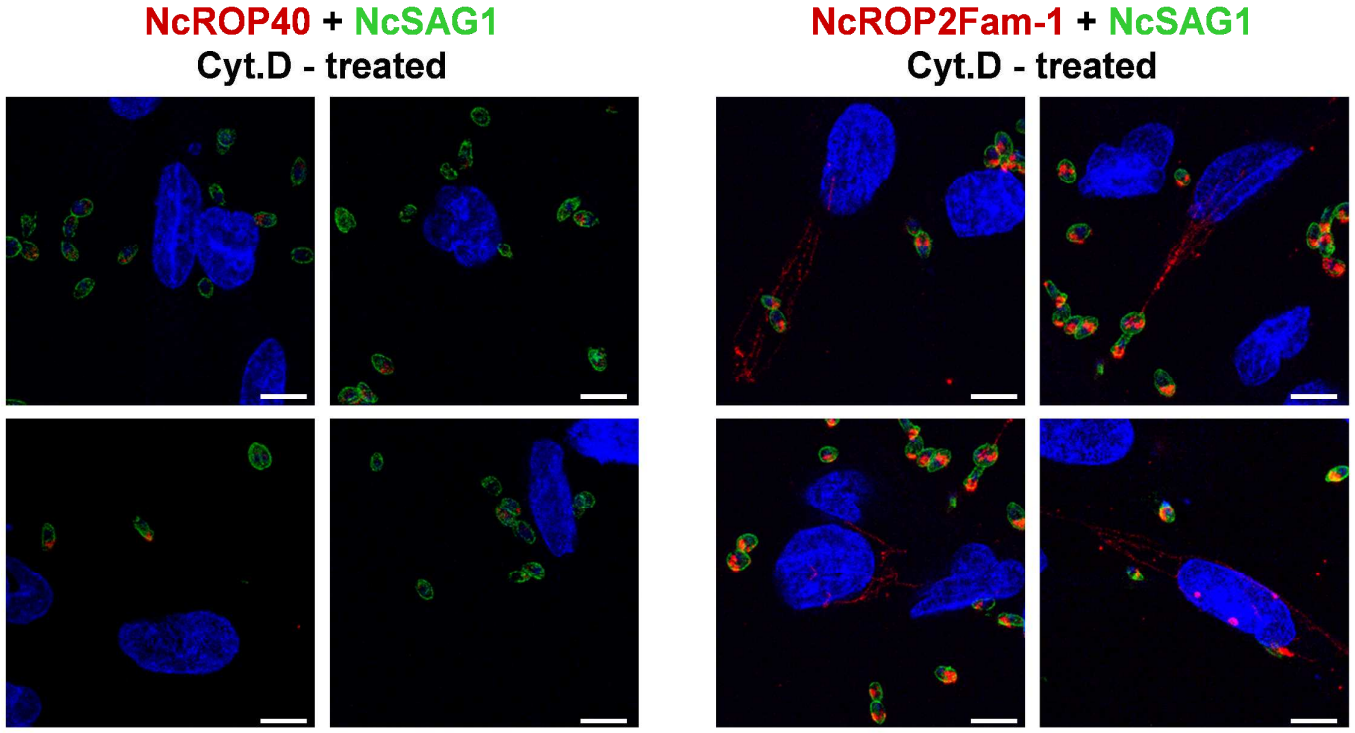
Additional file 3: The NcROP40-long gene model was corroborated by reverse transcription PCR. mRNA was obtained from tachyzoites at 24 and 56 hours post-infection as described in the methods section (*Evaluation of NcROP40 and NcROP2Fam-1 mRNA expression levels*). The whole NcROP40-long ORF was amplified from cDNA with the following primers: Fw-NcROP40-long (5'-ATGAGACACTCCTTGTGCTTTTC-3') and Rv-NcROP40-long (5'-TCACCCCACCACTGAACG-3'). In addition, the same forward primer was used with the reverse internal primer employed for the q-PCR assays (5'-TGGTGACTGCGACCAACTTA-3', from Table 1). In all the cases, PCR amplification yielded a single fragment with the expected molecular weight (see figure below).



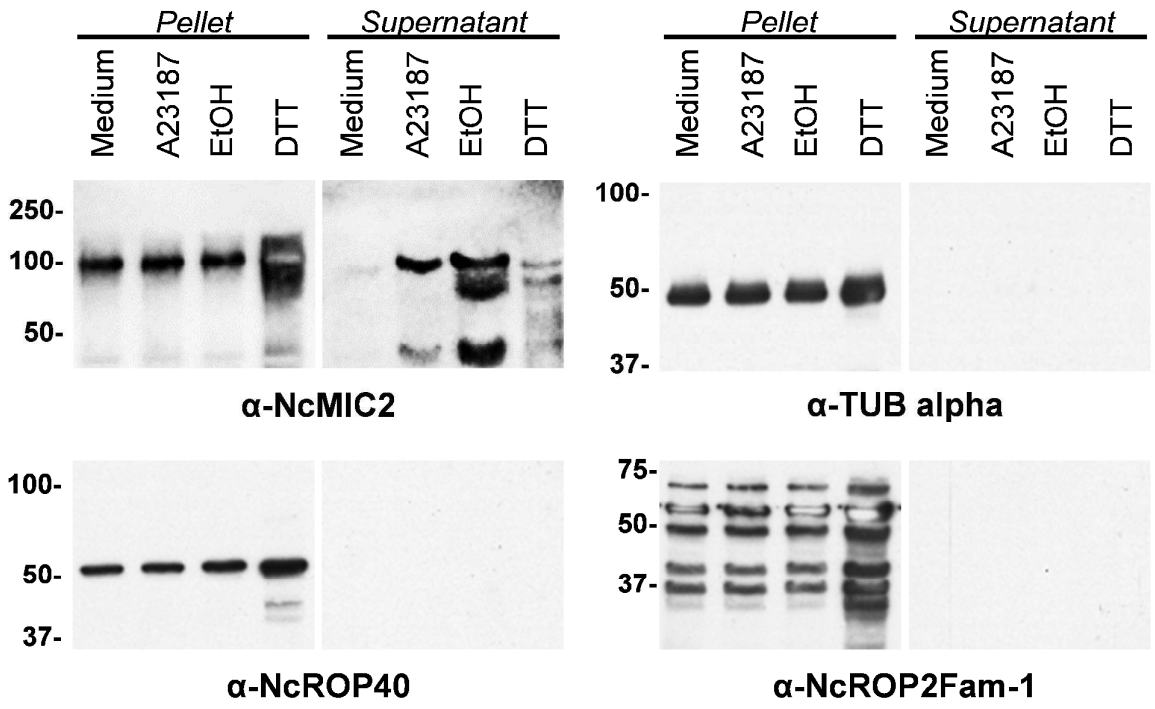
Additional file 4: Sequence alignment of the coding region for the NcROP40 protein and its up and down-stream regions within the chromosome V of the Nc-Liv genome. NCLIV_chrV (662772-665947 position), NcROP40 and NcROP40-long are displayed as templates and were obtained from the ToxoDB source as described in methods section. Consensus sequences among Nc-Liv, Nc-Spain1H and Nc-Spain7 isolates were obtained by DNA sequencing and aligned based on the template sequences. Predicted aminoacidic sequence of the NcROP40-long protein is also displayed.

See attachaed PDF

Additional file 5: Determination of NcROP40 and NcROP2Fam-1 secretion by evacuole assays. These experiments were carried out as described previously by Dunn et al., (2008), by incubating cytochalasin D-treated tachyzoites with human foreskin fibroblasts (HFFs) for 1 h prior to fixation. Evacuoles were detected by confocal laser-scanning microscopy using the affinity purified PAbs α-NcROP40 (red, left) and α-NcROP2Fam-1 (red, right). The MAb α-NcSAG1 (green) was employed as a surface marker. Only NcROP2Fam-1 was detected in evacuoles, whilst NcROP40 release was not observed. Images represent three merged stacks of 1 µm each. Bars represent 1 µm.



Additional file 6: Effect of A23187, ethanol and DTT on secretion of NcROP40 and NcROP2Fam-1 proteins as shown by Western-blot using respective antibodies. The same protein samples were also probed by immunoblotting with α-NcMIC2 and α-TUBα antibodies to (i) confirm induced secretion and (ii) exclude inadvertent tachyzoite lysis, respectively. Rhoptry discharge was not observed in culture supernatants upon any of these treatments, whilst NcMIC2 secretion was evident after A23187, ethanol and DTT supplementation. Tachyzoite lysis was not detected. All the antibodies specifically reacted against their respective protein on tachyzoite extracts.



```
                          10        20        30        40        50        60
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         GTTTAACAGTCTCCAAATACTGCTTGATATGAAGGAAGATCTTCAGTTATGCTGCATTTC
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     ------------------------------------------------------------
Consensus NcSp7     ------------------------------------------------------------
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------

                          70        80        90       100       110       120
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         AATTGGCGGCGGTAGAGCCCAGGGGGAGCCTGCCTCATTTTAGGCCGGATGGCCGCGGTC
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     ------------------------------------------------------------
Consensus NcSp7     ------------------------------------------------------------
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------

                         130       140       150       160       170       180
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         AGACATTCTGTGCTAGTGCACCTCTTTCCGTATCTTGACTCCGCCAATACTCCCCAGCCC
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     ------------------------------------------------------------
Consensus NcSp7     ------------------------------------------------------------
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------

                         190       200       210       220       230       240
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         GCCACCTTGGAAAACTCCCCACCGATCTACAGACTGACACCAGCCACAGCAAAAGCACGT
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     ------------------------------------------------------------
Consensus NcSp7     ------------------------------------------------------------
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------

                         250       260       270       280       290       300
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         GCCTCTTTTTTAAGAACGCATGGCTGACTGCTAGCGCTGTTCAATCCGTTGACGTCGCGC
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     ------------------------------------------------------------
Consensus NcSp7     ------------------------------------------------------------GC
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------

                         310       320       330       340       350       360
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         AAGGTTGGGCTCTGGACAGATTCCGCCGCGCAGAATTGGCGCTAGTCTAGATCACAGACA
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     ---------------------------------AATTGGCGCTAGTCTAGATCACAGACA
Consensus NcSp7     AAGGTTGGGCTCTGGACAGATTCCGCCGCGCAGAATTGGCGCTAGTCTAGATCACAGACA
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------
```

```
                        370       380       390       400       410       420
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         CGATCGTTTGAGCCGGTGTGGTTCCGAAAGTTCTTGAGTACAGGGTCTGGCTCCTACCCG
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     CGATCGTTTGAGCCGGTGTGGTTCCGAAAGTTCTTGAGTACAGGGTCTGGCTCCTACCCG
Consensus NcSp7     CGATCGTTTGAGCCGGTGTGGTTCCGAAAGTTCTTGAGTACAGGGTCTGGCTCCTACCCG
Consensus NcSp1H    ------------------------------------------------------------
Aa.sequence         ------------------------------------------------------------

                        430       440       450       460       470       480
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         TACTTTGGTGCATAAATCTGTCGGATGTATTATTCTTCTTTTGTGAAGTTTTGTTTCATA
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     TACTTTGGTGCATAAATCTGTCGGATGTATTATTCTTCTTTTGTGAAGTTTTGTTTCATA
Consensus NcSp7     TACTTTGGTGCATAAATCTGTCGGATGTATTATTCTTCTTTTGTGAAGTTTTGTTTCATA
Consensus NcSp1H    ------------------------------TATTCTTCTTTTGTGAAGTTTTGTTTCATA
Aa.sequence         ------------------------------------------------------------

                        490       500       510       520       530       540
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         GTGTTAAGCTTGTAAGCAGCTGACGTCAACTGATTATGTCTTGCGAGATCCCGCCTAGAC
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------
Consensus NcLiv     GTGTTAAGCTTGTAAGCAGCTGACGTCAACTGATTATGTCTTGCGAGATCCCGCCTAGAC
Consensus NcSp7     GTGTTAAGCTTGTAAGCAGCTGACGTCAACTGATTATGTCTTGCGAGATCCCGCCTAGAC
Consensus NcSp1H    GTGTTAAGCTTGTAAGCAGCTGACGTCAACTGATTATGTCTTGCGAGATCCCGCCTAGAC
Aa.sequence         ------------------------------------------------------------

                        550       560       570       580       590       600
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         AGACACAAACCAACCGAACGTCTGGGATTCCAGTGTGTAGCCGAATAGATTTAGCAAAAT
>NcROP40            ------------------------------------------------------------
>NcROP40-long       ------------------------------------------------------------AT
Consensus NcLiv     AGACACAAACCAACCGAACGTCTGGGATTCCAGTGTGTAGCCGAATAGATTTAGCAAAAT
Consensus NcSp7     AGACACAAACCAACCGAACGTCTGGGATTCCAGTGTGTAGCCGAATAGATTTAGCAAAAT
Consensus NcSp1H    AGACACAAACCAACCGAACGTCTGGGATTCCAGTGTGTAGCCGAATAGATTTAGCAAAAT
Aa.sequence         ------------------------------------------------------------Me

                        610       620       630       640       650       660
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         GAGACACTCCTTGTGCTTTTCGATATTTGCACTCGAATGCTTGGTGCTGCTTCTGACTTT
>NcROP40            ------------------------------------------------------------
>NcROP40-long       GAGACACTCCTTGTGCTTTTCGATATTTGCACTCGAATGCTTGGTGCTGCTTCTGACTTT
Consensus NcLiv     GAGACACTCCTTGTGCTTTTCGATATTTGCACTCGAATGCTTGGTGCTGCTTCTGACTTT
Consensus NcSp7     GAGACACTCCTTGTGCTTTTCGATATTTGCACTCGAATGCTTGGTGCTGCTTCTGACTTT
Consensus NcSp1H    GAGACACTCCTTGTGCTTTTCGATATTTGCACTCGAATGCTTGGTGCTGCTTCTGACTTT
Aa.sequence         tArgHisSerLeuCysPheSerIlePheAlaLeuGluCysLeuValLeuLeuLeuThrPh

                        670       680       690       700       710       720
                    ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV         TCAGTGGTTCGATCACTTTGAACGGGCCACCGCTACGGCACTCCGATCGAAGACGTACCA
>NcROP40            ------------------------------------------------------------
>NcROP40-long       TCAGTGGTTCGATCACTTTGAACGGGCCACCGCTACGGCACTCCGATCGAAGACGTACCA
Consensus NcLiv     TCAGTGGTTCGATCACTTTGAACGGGCCACCGCTACGGCACTCCGATCGAAGACGTACCA
Consensus NcSp7     TCAGTGGTTCGATCACTTTGAACGGGCCACCGCTACGGCACTCCGATCGAAGACGTACCA
Consensus NcSp1H    TCAGTGGTTCGATCACTTTGAACGGGCCACCGCTACGGCACTCCGATCGAAGACGTACCA
Aa.sequence         eGlnTrpPheAspHisPheGluArgAlaThrAlaThrAlaLeuArgSerLysThrTyrGl
```

```
                        730       740       750       760       770       780
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GGCATTGGGATTCGCCGAAGGTTCTGACAGCGAGACACAGCACGAGGAGGCAGCGAGCGG
>NcROP40           ------------------------------------------------------------
>NcROP40-long      GGCATTGGGATTCGCCGAAGGTTCTGACAGCGAGACACAGCACGAGGAGGCAGCGAGCGG
Consensus NcLiv    GGCATTGGGATTCGCCGAAGGTTCTGACAGCGAGACACAGCACGAGGAGGCAGCGAGCGG
Consensus NcSp7    GGCATTGGGATTCGCCGAAGGTTCTGACAGCGAGACACAGCACGAGGAGGCAGCGAGCGG
Consensus NcSp1H   GGCATTGGGATTCGCCGAAGGTTCTGACAGCGAGACACAGCACGAGGAGGCAGCGAGCGG
Aa.sequence        nAlaLeuGlyPheAlaGluGlySerAspSerGluThrGlnHisGluGluAlaAlaSerGl

                        790       800       810       820       830       840
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CGACACACCATTGGTAGGGGGGGCACGACCACGGAAGAGAAGTCCGCTTTCTCGCCTTGG
>NcROP40           ------------------------------------------------------------
>NcROP40-long      CGACACACCATTGGTAGGGGGGGCACGACCACGGAAGAGAAGTCCGCTTTCTCGCCTTGG
Consensus NcLiv    CGACACACCATTGGTAGGGGGGGCACGACCACGGAAGAGAAGTCCGCTTTCTCGCCTTGG
Consensus NcSp7    CGACACACCATTGGTAGGGGGGGCACGACCACGGAAGAGAAGTCCGCTTTCTCGCCTTGG
Consensus NcSp1H   CGACACACCATTGGTAGGGGGGGCACGACCACGGAAGAGAAGTCCGCTTTCTCGCCTTGG
Aa.sequence        yAspThrProLeuValGlyGlyAlaArgProArgLysArgSerProLeuSerArgLeuGl

                        850       860       870       880       890       900
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CTCTTTCTTTCGCAGACGCGGAGGCAGACGAGGAAATGTGGAAGGAGACTCTCAAGGCGC
>NcROP40           ------------------------------------------------------------
>NcROP40-long      CTCTTTCTTTCGCAGACGCGGAGGCAGACGAGGAAATGTGGAAGGAGACTCTCAAGGCGC
Consensus NcLiv    CTCTTTCTTTCGCAGACGCGGAGGCAGACGAGGAAATGTGGAAGGAGACTCTCAAGGCGC
Consensus NcSp7    CTCTTTCTTTCGCAGACGCGGAGGCAGACGAGGAAATGTGGAAGGAGACTCTCAAGGCGC
Consensus NcSp1H   CTCTTTCTTTCGCAGACGCGGAGGCAGACGAGGAAATGTGGAAGGAGACTCTCAAGGCGC
Aa.sequence        ySerPhePheArgArgArgGlyGlyArgArgGlyAsnValGluGlyAspSerGlnGlyAl

                        910       920       930       940       950       960
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CAGTGAAGAAGGAGAACAGCTGTTAGGTCACCCCAGCCACACACACACACGGGGGGGCCT
>NcROP40           ------------------------------------------------------------
>NcROP40-long      CAGTGAAGAAGGAGAACAGCTGTTAGGTCACCCCAGCCACACACACACACGGGGGGGCCT
Consensus NcLiv    CAGTGAAGAAGGAGAACAGCTGTTAGGTCACCCCAGCCACACACACACACGGGGGGGCCT
Consensus NcSp7    CAGTGAAGAAGGAGAACAGCTGTTAGGTCACCCCAGCCACACACACACACGGGGGGGCCT
Consensus NcSp1H   CAGTGAAGAAGGAGAACAGCTGTTAGGTCACCCCAGCCACACACACACACGGGGGGGCCT
Aa.sequence        aSerGluGluGlyGluGlnLeuLeuGlyHisProSerHisThrHisThrArgGlyGlyLe

                        970       980       990       1000      1010      1020
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GCGGTTCGGGCGCAAAAAACAGAATACACCCCACGAGGCAATGGTGAAATCCCTGCACAA
>NcROP40           ----------------------------------------ATGGTGAAATCCCTGCACAA
>NcROP40-long      GCGGTTCGGGCGCAAAAAACAGAATACACCCCACGAGGCAATGGTGAAATCCCTGCACAA
Consensus NcLiv    GCGGTTCGGGCGCAAAAAACAGAATACACCCCACGAGGCAATGGTGAAATCCCTGCACAA
Consensus NcSp7    GCGGTTCGGGCGCAAAAAACAGAATACACCCCACGAGGCAATGGTGAAATCCCTGCACAA
Consensus NcSp1H   GCGGTTCGGGCGCAAAAAACAGAATACACCCCACGAGGCAATGGTGAAATCCCTGCACAA
Aa.sequence        uArgPheGlyArgLysLysGlnAsnThrProHisGluAlaMetValLysSerLeuHisLy

                        1030      1040      1050      1060      1070      1080
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GTTCTTGAGCAAGAACGTAGTGGAGGCAATGTCCCTAGAAGATTTTGTAGCAGAACTCGG
>NcROP40           GTTCTTGAGCAAGAACGTAGTGGAGGCAATGTCCCTAGAAGATTTTGTAGCAGAACTCGG
>NcROP40-long      GTTCTTGAGCAAGAACGTAGTGGAGGCAATGTCCCTAGAAGATTTTGTAGCAGAACTCGG
Consensus NcLiv    GTTCTTGAGCAAGAACGTAGTGGAGGCAATGTCCCTAGAAGATTTTGTAGCAGAACTCGG
Consensus NcSp7    GTTCTTGAGCAAGAACGTAGTGGAGGCAATGTCCCTAGAAGATTTTGTAGCAGAACTCGG
Consensus NcSp1H   GTTCTTGAGCAAGAACGTAGTGGAGGCAATGTCCCTAGAAGATTTTGTAGCAGAACTCGG
Aa.sequence        sPheLeuSerLysAsnValValGluAlaMetSerLeuGluAspPheValAlaGluLeuGl
```

```
                        1090      1100      1110      1120      1130      1140
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TGTCGACCATCGTGCAATTCAACCTCCATTCTTTCGCAAGGGCGAGGGCGCGTCGAGAGC
>NcROP40           TGTCGACCATCGTGCAATTCAACCTCCATTCTTTCGCAAGGGCGAGGGCGCGTCGAGAGC
>NcROP40-long      TGTCGACCATCGTGCAATTCAACCTCCATTCTTTCGCAAGGGCGAGGGCGCGTCGAGAGC
Consensus NcLiv    TGTCGACCATCGTGCAATTCAACCTCCATTCTTTCGCAAGGGCGAGGGCGCGTCGAGAGC
Consensus NcSp7    TGTCGACCATCGTGCAATTCAACCTCCATTCTTTCGCAAGGGCGAGGGCGCGTCGAGAGC
Consensus NcSp1H   TGTCGACCATCGTGCAATTCAACCTCCATTCTTTCGCAAGGGCGAGGGCGCGTCGAGAGC
Aa.sequence        yValAspHisArgAlaIleGlnProProPhePheArgLysGlyGluGlyAlaSerArgAl

                        1150      1160      1170      1180      1190      1200
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TGTCGGGTATTTCGCGGAGCAGCAAAAGCCAGAAGTATCTGAAGAGACGCGGAAAACTTT
>NcROP40           TGTCGGGTATTTCGCGGAGCAGCAAAAGCCAGAAGTATCTGAAGAGACGCGGAAAACTTT
>NcROP40-long      TGTCGGGTATTTCGCGGAGCAGCAAAAGCCAGAAGTATCTGAAGAGACGCGGAAAACTTT
Consensus NcLiv    TGTCGGGTATTTCGCGGAGCAGCAAAAGCCAGAAGTATCTGAAGAGACGCGGAAAACTTT
Consensus NcSp7    TGTCGGGTATTTCGCGGAGCAGCAAAAGCCAGAAGTATCTGAAGAGACGCGGAAAACTTT
Consensus NcSp1H   TGTCGGGTATTTCGCGGAGCAGCAAAAGCCAGAAGTATCTGAAGAGACGCGGAAAACTTT
Aa.sequence        aValGlyTyrPheAlaGluGlnGlnLysProGluValSerGluGluThrArgLysThrLe

                        1210      1220      1230      1240      1250      1260
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GGAAGCGGTTGAACCAGTTCTGCCTGCCGGACAGCCACTGTCTTTTCACACTACATACGA
>NcROP40           GGAAGCGGTTGAACCAGTTCTGCCTGCCGGACAGCCACTGTCTTTTCACACTACATACGA
>NcROP40-long      GGAAGCGGTTGAACCAGTTCTGCCTGCCGGACAGCCACTGTCTTTTCACACTACATACGA
Consensus NcLiv    GGAAGCGGTTGAACCAGTTCTGCCTGCCGGACAGCCACTGTCTTTTCACACTACATACGA
Consensus NcSp7    GGAAGCGGTTGAACCAGTTCTGCCTGCCGGACAGCCACTGTCTTTTCACACTACATACGA
Consensus NcSp1H   GGAAGCGGTTGAACCAGTTCTGCCTGCCGGACAGCCACTGTCTTTTCACACTACATACGA
Aa.sequence        uGluAlaValGluProValLeuProAlaGlyGlnProLeuSerPheHisThrThrTyrAs

                        1270      1280      1290      1300      1310      1320
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CAGAAAAGGCTCTTACTTCAAAAGAGGCAGCTTGTTTCACCGAGATTTCTTCGAGTTCTT
>NcROP40           CAGAAAAGGCTCTTACTTCAAAAGAGGCAGCTTGTTTCACCGAGATTTCTTCGAGTTCTT
>NcROP40-long      CAGAAAAGGCTCTTACTTCAAAAGAGGCAGCTTGTTTCACCGAGATTTCTTCGAGTTCTT
Consensus NcLiv    CAGAAAAGGCTCTTACTTCAAAAGAGGCAGCTTGTTTCACCGAGATTTCTTCGAGTTCTT
Consensus NcSp7    CAGAAAAGGCTCTTACTTCAAAAGAGGCAGCTTGTTTCACCGAGATTTCTTCGAGTTCTT
Consensus NcSp1H   CAGAAAAGGCTCTTACTTCAAAAGAGGCAGCTTGTTTCACCGAGATTTCTTCGAGTTCTT
Aa.sequence        pArgLysGlySerTyrPheLysArgGlySerLeuPheHisArgAspPhePheGluPhePh

                        1330      1340      1350      1360      1370      1380
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CATTGATGGGCAACCTTTTGATCTGAGGATACTCCCACTACCGAGCGGTGAAGAGGGGGA
>NcROP40           CATTGATGGGCAACCTTTTGATCTGAGGATACTCCCACTACCGAGCGGTGAAGAGGGGGA
>NcROP40-long      CATTGATGGGCAACCTTTTGATCTGAGGATACTCCCACTACCGAGCGGTGAAGAGGGGGA
Consensus NcLiv    CATTGATGGGCAACCTTTTGATCTGAGGATACTCCCACTACCGAGCGGTGAAGAGGGGGA
Consensus NcSp7    CATTGATGGGCAACCTTTTGATCTGAGGATACTCCCACTACCGAGCGGTGAAGAGGGGGA
Consensus NcSp1H   CATTGATGGGCAACCTTTTGATCTGAGGATACTCCCACTACCGAGCGGTGAAGAGGGGGA
Aa.sequence        eIleAspGlyGlnProPheAspLeuArgIleLeuProLeuProSerGlyGluGluGlyGl

                        1390      1400      1410      1420      1430      1440
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        AGCTACGCTGGAACGATACAAAAAGGAGCTGGAAAATGAGCGGAGTGTTCGACTTCAATT
>NcROP40           AGCTACGCTGGAACGATACAAAAAGGAGCTGGAAAATGAGCGGAGTGTTCGACTTCAATT
>NcROP40-long      AGCTACGCTGGAACGATACAAAAAGGAGCTGGAAAATGAGCGGAGTGTTCGACTTCAATT
Consensus NcLiv    AGCTACGCTGGAACGATACAAAAAGGAGCTGGAAAATGAGCGGAGTGTTCGACTTCAATT
Consensus NcSp7    AGCTACGCTGGAACGATACAAAAAGGAGCTGGAAAATGAGCGGAGTGTTCGACTTCAATT
Consensus NcSp1H   AGCTACGCTGGAACGATACAAAAAGGAGCTGGAAAATGAGCGGAGTGTTCGACTTCAATT
Aa.sequence        uAlaThrLeuGluArgTyrLysLysGluLeuGluAsnGluArgSerValArgLeuGlnPh
```

```
                        1450      1460      1470      1480      1490      1500
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TGATGTGGGTTCTGCTCAACGTGTCGTGGAGGCCTTTCACTGTCACATTCCATTTCAAGT
>NcROP40           TGATGTGGGTTCTGCTCAACGTGTCGTGGAGGCCTTTCACTGTCACATTCCATTTCAAGT
>NcROP40-long      TGATGTGGGTTCTGCTCAACGTGTCGTGGAGGCCTTTCACTGTCACATTCCATTTCAAGT
Consensus NcLiv    TGATGTGGGTTCTGCTCAACGTGTCGTGGAGGCCTTTCACTGTCACATTCCATTTCAAGT
Consensus NcSp7    TGATGTGGGTTCTGCTCAACGTGTCGTGGAGGCCTTTCACTGTCACATTCCATTTCAAGT
Consensus NcSp1H   TGATGTGGGTTCTGCTCAACGTGTCGTGGAGGCCTTTCACTGTCACATTCCATTTCAAGT
Aa.sequence        eAspValGlySerAlaGlnArgValValGluAlaPheHisCysHisIleProPheGlnVa

                        1510      1520      1530      1540      1550      1560
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GCTGCAGTTTACAAGCGACAGAAAGGTCGTCTCACTTGGGTTAGACCTCAAGATGCCCAA
>NcROP40           GCTGCAGTTTACAAGCGACAGAAAGGTCGTCTCACTTGGGTTAGACCTCAAGATGCCCAA
>NcROP40-long      GCTGCAGTTTACAAGCGACAGAAAGGTCGTCTCACTTGGGTTAGACCTCAAGATGCCCAA
Consensus NcLiv    GCTGCAGTTTACAAGCGACAGAAAGGTCGTCTCACTTGGGTTAGACCTCAAGATGCCCAA
Consensus NcSp7    GCTGCAGTTTACAAGCGACAGAAAGGTCGTCTCACTTGGGTTAGACCTCAAGATGCCCAA
Consensus NcSp1H   GCTGCAGTTTACAAGCGACAGAAAGGTCGTCTCACTTGGGTTAGACCTCAAGATGCCCAA
Aa.sequence        lLeuGlnPheThrSerAspArgLysValValSerLeuGlyLeuAspLeuLysMetProAs

                        1570      1580      1590      1600      1610      1620
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CATTGTTCTCATCTACCCGGGCACACGTGGGACGCTCGGCCAACTCTTTCCGTTGATACA
>NcROP40           CATTGTTCTCATCTACCCGGGCACACGTGGGACGCTCGGCCAACTCTTTCCGTTGATACA
>NcROP40-long      CATTGTTCTCATCTACCCGGGCACACGTGGGACGCTCGGCCAACTCTTTCCGTTGATACA
Consensus NcLiv    CATTGTTCTCATCTACCCGGGCACACGTGGGACGCTCGGCCAACTCTTTCCGTTGATACA
Consensus NcSp7    CATTGTTCTCATCTACCCGGGCACACGTGGGACGCTCGGCCAACTCTTTCCGTTGATACA
Consensus NcSp1H   CATTGTTCTCATCTACCCGGGCACACGTGGGACGCTCGGCCAACTCTTTCCGTTGATACA
Aa.sequence        nIleValLeuIleTyrProGlyThrArgGlyThrLeuGlyGlnLeuPheProLeuIleHi

                        1630      1640      1650      1660      1670      1680
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TCAAGCAGCCCAGAATCAGAAAACCGCCCCCGCTGCTCTAGCTGCCCGGCTGAGCGTGAC
>NcROP40           TCAAGCAGCCCAGAATCAGAAAACCGCCCCCGCTGCTCTAGCTGCCCGGCTGAGCGTGAC
>NcROP40-long      TCAAGCAGCCCAGAATCAGAAAACCGCCCCCGCTGCTCTAGCTGCCCGGCTGAGCGTGAC
Consensus NcLiv    TCAAGCAGCCCAGAATCAGAAAACCGCCCCCGCTGCTCTAGCTGCCCGGCTGAGCGTGAC
Consensus NcSp7    TCAAGCAGCCCAGAATCAGAAAACCGCCCCCGCTGCTCTAGCTGCCCGGCTGAGCGTGAC
Consensus NcSp1H   TCAAGCAGCCCAGAATCAGAAAACCGCCCCCGCTGCTCTAGCTGCCCGGCTGAGCGTGAC
Aa.sequence        sGlnAlaAlaGlnAsnGlnLysThrAlaProAlaAlaLeuAlaAlaArgLeuSerValTh

                        1690      1700      1710      1720      1730      1740
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GGTGCAGGCCATTAAGTTGGTCGCAGTCACCAGTGGAAGAGGGATCTTGGTGAGTAACAT
>NcROP40           GGTGCAGGCCATTAAGTTGGTCGCAGTCACCAGTGGAAGAGGGATCTTGGTGAGTAACAT
>NcROP40-long      GGTGCAGGCCATTAAGTTGGTCGCAGTCACCAGTGGAAGAGGGATCTTGGTGAGTAACAT
Consensus NcLiv    GGTGCAGGCCATTAAGTTGGTCGCAGTCACCAGTGGAAGAGGGATCTTGGTGAGTAACAT
Consensus NcSp7    GGTGCAGGCCATTAAGTTGGTCGCAGTCACCAGTGGAAGAGGGATCTTGGTGAGTAACAT
Consensus NcSp1H   GGTGCAGGCCATTAAGTTGGTCGCAGTCACCAGTGGAAGAGGGATCTTGGTGAGTAACAT
Aa.sequence        rValGlnAlaIleLysLeuValAlaValThrSerGlyArgGlyIleLeuValSerAsnIl

                        1750      1760      1770      1780      1790      1800
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CTCGCCGGAAAATTTCTTCCCCAGTAGAGATGGAATTCTTTATTTTGGTGGCTTCTCCTC
>NcROP40           CTCGCCGGAAAATTTCTTCCCCAGTAGAGATGGAATTCTTTATTTTGGTGGCTTCTCCTC
>NcROP40-long      CTCGCCGGAAAATTTCTTCCCCAGTAGAGATGGAATTCTTTATTTTGGTGGCTTCTCCTC
Consensus NcLiv    CTCGCCGGAAAATTTCTTCCCCAGTAGAGATGGAATTCTTTATTTTGGTGGCTTCTCCTC
Consensus NcSp7    CTCGCCGGAAAATTTCTTCCCCAGTAGAGATGGAATTCTTTATTTTGGTGGCTTCTCCTC
Consensus NcSp1H   CTCGCCGGAAAATTTCTTCCCCAGTAGAGATGGAATTCTTTATTTTGGTGGCTTCTCCTC
Aa.sequence        eSerProGluAsnPhePheProSerArgAspGlyIleLeuTyrPheGlyGlyPheSerSe
```

```
                        1810      1820      1830      1840      1850      1860
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        AAAAGTAGCGGCAAACAAGCTGTACTACGAGACCGAGGGGGGCGCCCTGGTCGAGGAGCC
>NcROP40           AAAAGTAGCGGCAAACAAGCTGTACTACGAGACCGAGGGGGGCGCCCTGGTCGAGGAGCC
>NcROP40-long      AAAAGTAGCGGCAAACAAGCTGTACTACGAGACCGAGGGGGGCGCCCTGGTCGAGGAGCC
Consensus NcLiv    AAAAGTAGCGGCAAACAAGCTGTACTACGAGACCGAGGGGGGCGCCCTGGTCGAGGAGCC
Consensus NcSp7    AAAAGTAGCGGCAAACAAGCTGTACTACGAGACCGAGGGGGGCGCCCTGGTCGAGGAGCC
Consensus NcSp1H   AAAAGTAGCGGCAAACAAGCTGTACTACGAGACCGAGGGGGGCGCCCTGGTCGAGGAGCC
Aa.sequence        rLysValAlaAlaAsnLysLeuTyrTyrGluThrGluGlyGlyAlaLeuValGluGluPr

                        1870      1880      1890      1900      1910      1920
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GCCGAACGTGACTTCCAGGGGGAGACGGTTCACCGCTGAAGACAACGCAGCAGACTTAGG
>NcROP40           GCCGAACGTGACTTCCAGGGGGAGACGGTTCACCGCTGAAGACAACGCAGCAGACTTAGG
>NcROP40-long      GCCGAACGTGACTTCCAGGGGGAGACGGTTCACCGCTGAAGACAACGCAGCAGACTTAGG
Consensus NcLiv    GCCGAACGTGACTTCCAGGGGGAGACGGTTCACCGCTGAAGACAACGCAGCAGACTTAGG
Consensus NcSp7    GCCGAACGTGACTTCCAGGGGGAGACGGTTCACCGCTGAAGACAACGCAGCAGACTTAGG
Consensus NcSp1H   GCCGAACGTGACTTCCAGGGGGAGACGGTTCACCGCTGAAGACAACGCAGCAGACTTAGG
Aa.sequence        oProAsnValThrSerArgGlyArgArgPheThrAlaGluAspAsnAlaAlaAspLeuGl

                        1930      1940      1950      1960      1970      1980
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        ACGCACTCTCTTTGGCTTGTGGTGCGGCGGTTCGCTTCCTGAGGATGAACCGAGTGGGCG
>NcROP40           ACGCACTCTCTTTGGCTTGTGGTGCGGCGGTTCGCTTCCTGAGGATGAACCGAGTGGGCG
>NcROP40-long      ACGCACTCTCTTTGGCTTGTGGTGCGGCGGTTCGCTTCCTGAGGATGAACCGAGTGGGCG
Consensus NcLiv    ACGCACTCTCTTTGGCTTGTGGTGCGGCGGTTCGCTTCCTGAGGATGAACCGAGTGGGCG
Consensus NcSp7    ACGCACTCTCTTTGGCTTGTGGTGCGGCGGTTCGCTTCCTGAGGATGAACCGAGTGGGCG
Consensus NcSp1H   ACGCACTCTCTTTGGCTTGTGGTGCGGCGGTTCGCTTCCTGAGGATGAACCGAGTGGGCG
Aa.sequence        yArgThrLeuPheGlyLeuTrpCysGlyGlySerLeuProGluAspGluProSerGlyAr

                        1990      2000      2010      2020      2030      2040
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GGCCGACGTGGATTTCTCCAATTGCGGGACAGACCTGCCGGATCCTGTCAAGAAGTTAAT
>NcROP40           GGCCGACGTGGATTTCTCCAATTGCGGGACAGACCTGCCGGATCCTGTCAAGAAGTTAAT
>NcROP40-long      GGCCGACGTGGATTTCTCCAATTGCGGGACAGACCTGCCGGATCCTGTCAAGAAGTTAAT
Consensus NcLiv    GGCCGACGTGGATTTCTCCAATTGCGGGACAGACCTGCCGGATCCTGTCAAGAAGTTAAT
Consensus NcSp7    GGCCGACGTGGATTTCTCCAATTGCGGGACAGACCTGCCGGATCCTGTCAAGAAGTTAAT
Consensus NcSp1H   GGCCGACGTGGATTTCTCCAATTGCGGGACAGACCTGCCGGATCCTGTCAAGAAGTTAAT
Aa.sequence        gAlaAspValAspPheSerAsnCysGlyThrAspLeuProAspProValLysLysLeuIl

                        2050      2060      2070      2080      2090      2100
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TATGGGAGTGTCCGGCTCCCACGGACGCGCCCCTCTCAGCGCCTCCCAGGTCCTCGATAC
>NcROP40           TATGGGAGTGTCCGGCTCCCACGGACGCGCCCCTCTCAGCGCCTCCCAGGTCCTCGATAC
>NcROP40-long      TATGGGAGTGTCCGGCTCCCACGGACGCGCCCCTCTCAGCGCCTCCCAGGTCCTCGATAC
Consensus NcLiv    TATGGGAGTGTCCGGCTCCCACGGACGCGCCCCTCTCAGCGCCTCCCAGGTCCTCGATAC
Consensus NcSp7    TATGGGAGTGTCCGGCTCCCACGGACGCGCCCCTCTCAGCGCCTCCCAGGTCCTCGATAC
Consensus NcSp1H   TATGGGAGTGTCCGGCTCCCACGGACGCGCCCCTCTCAGCGCCTCCCAGGTCCTCGATAC
Aa.sequence        eMetGlyValSerGlySerHisGlyArgAlaProLeuSerAlaSerGlnValLeuAspTh

                        2110      2120      2130      2140      2150      2160
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TCCAAACTATCAGGAGCTGCGCAGATTAGAAAAAGAAGTGTCTCAGAGTGTCGCGGTGCG
>NcROP40           TCCAAACTATCAGGAGCTGCGCAGATTAGAAAAAGAAGTGTCTCAGAGTGTCGCGGTGCG
>NcROP40-long      TCCAAACTATCAGGAGCTGCGCAGATTAGAAAAAGAAGTGTCTCAGAGTGTCGCGGTGCG
Consensus NcLiv    TCCAAACTATCAGGAGCTGCGCAGATTAGAAAAAGAAGTGTCTCAGAGTGTCGCGGTGCG
Consensus NcSp7    TCCAAACTATCAGGAGCTGCGCAGATTAGAAAAAGAAGTGTCTCAGAGTGTCGCGGTGCG
Consensus NcSp1H   TCCAAACTATCAGGAGCTGCGCAGATTAGAAAAAGAAGTGTCTCAGAGTGTCGCGGTGCG
Aa.sequence        rProAsnTyrGlnGluLeuArgArgLeuGluLysGluValSerGlnSerValAlaValAr
```

```
                        2170      2180      2190      2200      2210      2220
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TTCAGTGGTGGGGTGACGTCTTTTGAGTTTCTGGACCGTGGACTGGTATCTTAAGAGTGG
>NcROP40           TTCAGTGGTGGGGTGA----------------------------------------------
>NcROP40-long      TTCAGTGGTGGGGTGA----------------------------------------------
Consensus NcLiv    TTCAGTGGTGGGGTGACGTCTTTTGAGTTTCTGGACCGTGGACTGGTATCTTAAGAGTGG
Consensus NcSp7    TTCAGTGGTGGGGTGACGTCTTTTGAGTTTCTGGACCGTGGACTGGTATCTTAAGAGTGG
Consensus NcSp1H   TTCAGTGGTGGGGTGACGTCTTTTGAGTTTCTGGACCGTGGACTGGTATCTTAAGAGTGG
Aa.sequence        gSerValValGlyEnd----------------------------------------------

                        2230      2240      2250      2260      2270      2280
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TTTTCCCGGATGGTCAGGAAGACTCTGAATGACCATTTAATGGGCCGTGTGGATGCGGGA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    TTTTCCCGGATGGTCAGGAAGACTCTGAATGACCATTTAATGGGCCGTGTGGATGCGGGA
Consensus NcSp7    TTTTCCCGGATGGTCAGGAAGACTCTGAATGACCATTTAATGGGCCGTGTGGATGCGGGA
Consensus NcSp1H   TTTTCCCGGATGGTCAGGAAGACTCTGAATGACCATTTAATGGGCCGTGTGGATGCGGGA
Aa.sequence        ------------------------------------------------------------

                        2290      2300      2310      2320      2330      2340
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GTTGCTCTAAGGCGATTGCTTTTTGCTCTTTACCCCCGATATCGCTGTGACTGCATAAGC
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    GTTGCTCTAAGGCGATTGCTTTTTGCTCTTTACCCCCGATATCGCTGTGACTGCATAAGC
Consensus NcSp7    GTTGCTCTAAGGCGATTGCTTTTTGCTCTTTACCCCCGATATCGCTGTGACTGCATAAGC
Consensus NcSp1H   GTTGCTCTAAGGCGATTGCTTTTTGCTCTTTACCCCCGATATCGCTGTGACTGCATAAGC
Aa.sequence        ------------------------------------------------------------

                        2350      2360      2370      2380      2390      2400
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GCTGTCTGGGGTCTAATAAACAGAAGAGGACAACACAGCGGTTCCACAAGCTGCACACAG
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    GCTGTCTGGGGTCTAATAAACAGAAGAGGACAACACAGCGGTTCCACAAGCTGCACACAG
Consensus NcSp7    GCTGTCTGGGGTCTAATAAACAGAAGAGGACAACACAGCGGTTCCACAAGCTGCACACAG
Consensus NcSp1H   GCTGTCTGGGGTCTAATAAACAGAAGAGGACAACACAGCGGTTCCACAAGCTGCACACAG
Aa.sequence        ------------------------------------------------------------

                        2410      2420      2430      2440      2450      2460
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GCGCGTACCTCTTGCCTTTGTAGAACGAGTTGTGAATGCTGTCACCAATGAATGTGTTTT
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    GCGCGTACCTCTTGCCTTTGTAGAACGAGTTGTGAATGCTGTCACCAATGAATGTGTTTT
Consensus NcSp7    GCGCGTACCTCTTGCCTTTGTAGAACGAGTTGTGAATGCTGTCACCAATGAATGTGTTTT
Consensus NcSp1H   GCGCGTACCTCTTGCCTTTGTAGAACGAGTTGTGAATGCTGTCACCAATGAATGTGTTTT
Aa.sequence        ------------------------------------------------------------

                        2470      2480      2490      2500      2510      2520
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TGTGAAGTATACGTTTTGTCCGAACCGTTCCGTGGATCATGCCGCACCGGCGATTTCAGT
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    TGTGAAGTATACGTTTTGTCCGAACCGTTCCGTGGATCATGCCGCACCGGCGATTTCAGT
Consensus NcSp7    TGTGAAGTATACGTTTTGTCCGAACCGTTCCGTGGATCATGCCGCACCGGCGATTTCAGT
Consensus NcSp1H   TGTGAAGTATACGTTTTGTCCGAACCGTTCCGTGGATCATGCCGCACCGGCGATTTCAGT
Aa.sequence        ------------------------------------------------------------
```

```
                        2530      2540      2550      2560      2570      2580
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        AGTCCAAGTGCGGATGGTTTTGCTGCCATCCATGCTCATGGTGGCCCTATTGCTCCCAAA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    AGTCCAAGTGCGGATGGTTTTGCTGCCATCCATGCTCATGGTGGCCCTATTGCTCCCAAA
Consensus NcSp7    AGTCCAAGTGCGGATGGTTTTGCTGCCATCCATGCTCATGGTGGCCCTATTGCTCCCAAA
Consensus NcSp1H   AGTCCAAGTGCGGATGGTTTTGCTGCCATCCATGCTCATGGTGGCCCTATTGCTCCCAAA
Aa.sequence        ------------------------------------------------------------

                        2590      2600      2610      2620      2630      2640
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CGTTGAGACAACGAGTCTGTTTAGGGGATCGTACATCCGTTGACCTCTTCAAGACATGAA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    CGTTGAGACAACGAGTCTGTTTAGGGGATCGTACATCCGTTGACCTCTTCAAGACATGAA
Consensus NcSp7    CGTTGAGACAACGAGTCTGTTTAGGGGATCGTACATCCGTTGACCTCTTCAAGACATGAA
Consensus NcSp1H   CGTTGAGACAACGAGTCTGTTTAGGGGATCGTACATCCGTTGACCTCTTCAAGACATGAA
Aa.sequence        ------------------------------------------------------------

                        2650      2660      2670      2680      2690      2700
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        ATTGATGATAGGGAACAGCAGTATGTTGGTGACCGCACCTTTTCGGTGACGCGTTAGTGA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    ATTGATGATAGGGAACAGCAGTATGTTGGTGACCGCACCTTTTCGGTGACGCGTTAGTGA
Consensus NcSp7    ATTGATGATAGGGAACAGCAGTATGTTGGTGACCGCACCTTTTCGGTGACGCGTTAGTGA
Consensus NcSp1H   ATTGATGATAGGGAACAGCAGTATGTTGGTGACCGCACCTTTTCGGTGACGCGTTAGTGA
Aa.sequence        ------------------------------------------------------------

                        2710      2720      2730      2740      2750      2760
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CCGTTTCTGGTTTGTCGTCGCCGATACGGGGCGGTGACATAAGCAGTGACCAGAAGATAT
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    CCGTTTCTGGTTTGTCGTCGCCGATACGGGGCGGTGACATAAGCAGTGACCAGAAGATAT
Consensus NcSp7    CCGTTTCTGGTTTGTCGTCGCCGATACGGGGCGGTGACATAAGCAGTGACCAGAAGATAT
Consensus NcSp1H   CCGTTTCTGGTTTGTCGTCGCCGATACGGGGCGGTGACATAAGCAGTGACCAGAAGATAT
Aa.sequence        ------------------------------------------------------------

                        2770      2780      2790      2800      2810      2820
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        AAAATAGTGAGGAAATAAAGGTTGCTGGTAATGCAGCTTGATCAAGGACTAGGCAGCACA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    AAAATAGTGAGGAAATAAAGGTTGCTGGTAATGCAGCTTGATCAAGGACTAG--------
Consensus NcSp7    AAAATAGTGAGGAAATAAAGGTTGCTGGTAATGCAGCTTGATCAAGGACTAGGCAGCACA
Consensus NcSp1H   AAAATAGTGAGGAAATAAAGGTTGCTGGTAATGCAGCTTGATCAAGGACTAGGCAGCACA
Aa.sequence        ------------------------------------------------------------

                        2830      2840      2850      2860      2870      2880
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        GAAGCGACATTTCCCGAATGTAACGATGCACGGTTCCGGAGCGAGATAGCTCGTAGCTTA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    ------------------------------------------------------------
Consensus NcSp7    GAAGCGACATTTCCCGAATGTAACGATGCACGGTTCCGGAGCGAGATAGCTCGTAGCTTA
Consensus NcSp1H   GAAGCGACATTTCCCGAATGTAACGATGCACGGTTCCGGAGCGAGATAGCTCGTAGCTTA
Aa.sequence        ------------------------------------------------------------
```

```
                        2890      2900      2910      2920      2930      2940
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CCCCGGTTCTTATGATGGTCCACTGATAAATCGTGCCTTCGAAGAAGCATCTTTGATACA
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    ------------------------------------------------------------
Consensus NcSp7    CCCCGGTTCTTATGATGGTCCACTGATAAATCGTGCCTTCGAAGAAGCATCTTTGATACA
Consensus NcSp1H   CCCCGGTTCTTATGATGGTCCACTGATAAATCGTGCCTTCGAAGAAGCATCTTTGATACA
Aa.sequence        ------------------------------------------------------------

                        2950      2960      2970      2980      2990      3000
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        CTGCAACCTGCTACACCAAAACTGCCATAACTGAATTGCTGAATGCAATTGTACCTGGAG
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    ------------------------------------------------------------
Consensus NcSp7    CTGCAACCTGCTACACCAAAACTGCCATAACTGAATTGCTGAATGCAATTGTACCTGGA-
Consensus NcSp1H   CTGCAACCTGCTACACCAAAACTGCCATAACTGAATTGCTGAATGCAATTGTACCTGGAG
Aa.sequence        ------------------------------------------------------------

                        3010      3020      3030      3040      3050      3060
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TCGCAACTGCTTCTCTCACGCGTACGAAGGATGCACTGGCAGGCTATCGTCTTAGCTACC
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    ------------------------------------------------------------
Consensus NcSp7    ------------------------------------------------------------
Consensus NcSp1H   TCGCAACTGCTTCTCTCA------------------------------------------
Aa.sequence        ------------------------------------------------------------

                        3070      3080      3090      3100      3110      3120
                   ....|....|....|....|....|....|....|....|....|....|....|....|
>NCLIV_chrV        TGCTTTCGGTTGTCTCCATGAAGTGCTCTCAGGAAGCCACCGACGGAAGGACAGGAAGGT
>NcROP40           ------------------------------------------------------------
>NcROP40-long      ------------------------------------------------------------
Consensus NcLiv    ------------------------------------------------------------
Consensus NcSp7    ------------------------------------------------------------
Consensus NcSp1H   ------------------------------------------------------------
Aa.sequence        ------------------------------------------------------------

                        3130      3140      3150      3160      3170
                   ....|....|....|....|....|....|....|....|....|....|.
>NCLIV_chrV        AAAAAGTGGTTGGTCGATTTGTCGAAGGGAAAGAGGAGCCGTTGTCAGCAGGCTGA
>NcROP40           ---------------------------------------------------------
>NcROP40-long      ---------------------------------------------------------
Consensus NcLiv    ---------------------------------------------------------
Consensus NcSp7    ---------------------------------------------------------
Consensus NcSp1H   ---------------------------------------------------------
Aa.sequence        --------
```