CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Archive Ouverte en Sciences de l'Information et de la Communication

Characterization of a nuclear pore protein sheds light on the roles and composition of the Toxoplasma gondii nuclear pore complex

Flavie Courjol, Thomas Mouveaux, Kevin Lesage, Jean-Michel Saliou, Elisabeth Werkmeister, Maurine Bonabaud, Marine Rohmer, Christian Slomianny, Franck Lafont, Mathieu Gissot

To cite this version:

Flavie Courjol, Thomas Mouveaux, Kevin Lesage, Jean-Michel Saliou, Elisabeth Werkmeister, et al.. Characterization of a nuclear pore protein sheds light on the roles and composition of the Toxoplasma gondii nuclear pore complex. Cellular and Molecular Life Sciences, Springer Verlag, 2017, 74 (11), pp.2107-2125. 10.1007/s00018-017-2459-3 . hal-02106431

HAL Id: hal-02106431
https://hal.archives-ouvertes.fr/hal-02106431

Submitted on 23 Apr 2019

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

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

Characterization of a nuclear pore protein sheds light on the roles and composition of the *Toxoplasma gondii* nuclear pore complex.

Flavie Courjol[1], Thomas Mouveaux[1], Kevin Lesage[1], Jean-Michel Saliou[1], Elisabeth Werkmeister[1], Maurine Bonabaud[2], Marine Rohmer[2], Christian Slomianny[3], Franck Lafont[1] and Mathieu Gissot[1,*].

1: Univ. Lille, CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CIIL - Centre d'Infection et d'Immunité de Lille, F-59000 Lille, France.

2 : MGX-Montpellier GenomiX, c/o Institut de Génomique Fonctionnelle, 141 rue de la cardonille, 34094 Montpellier Cedex 5, France.

3 : Laboratory of Cell Physiology, INSERM U 1003, Université Lille Nord de France, Villeneuve d'Ascq, France

* Corresponding author: mathieu.gissot@pasteur-lille.fr

Acknowledgements

The authors would like to thank Dr. Valerie Doye for helpful discussions and Dr. R. Walker for critically reading the manuscript. We also thank Ludovic Huot for checking the integrity of RNA samples, Etienne Dewailly for electronic microscopy, Antonino Bongiovanni for his help with microscopy data analyses, and Quentin Deveuve for phylogenetic tree recommendations. The authors also thank the BioImaging Center Lille for access to instruments. This work was supported by Centre National de la Recherche Scientifique (CNRS), Institut National de la Santé et de la Recherche Médicale (INSERM), grants from the French National Research Agency (ANR) [grant number ANR-13-JSV3-0006-01 to MG and ANR-11-LABX-0024 to Plateforme Protéomique et Peptides Modifiés (P3M)], the Fonds Européen de Développement Economique Régionale (13003300-42405 Labex Parafrap to P3M) and the Métropole Européenne de Lille (MEL).

Abstract

The nuclear pore is a key structure in eukaryotes regulating nuclear-cytoplasmic transport as well as a wide range of cellular processes. Here, we report the characterization of the first *Toxoplasma gondii* nuclear pore protein, named TgNup302, which appears to be the orthologue of the mammalian Nup98-96 protein. We produced a conditional knock-down mutant that expresses TgNup302 under the control of an inducible tetracycline-regulated promoter. Under ATc treatment, a substantial decrease of TgNup302 protein in iKD parasites was observed, causing a delay in parasite proliferation. Moreover, the nuclear protein TgENO2 was trapped in the cytoplasm of ATc-treated mutants, suggesting that TgNup302 is involved in nuclear transport. Fluorescence *in situ* hybridization revealed that TgNup302 is essential for 18S RNA export from the nucleus to the cytoplasm, while global mRNA export remains unchanged. Using an affinity tag purification combined with mass spectrometry, we identified additional components of the nuclear pore complex, including proteins potentially interacting with chromatin. Furthermore, reverse immunoprecipitation confirmed their interaction with TgNup302, and structured illuminated microscopy confirmed the NPC localization of some of the TgNup302-interacting proteins. Intriguingly, facilitates chromatin transcription components were identified, suggesting the existence of an NPC-chromatin interaction in *T. gondii*. Identification of TgNup302-interacting proteins also provides the first glimpse at the NPC structure in Apicomplexa, suggesting a structural conservation of the NPC components between distant eukaryotes.

Keywords: *Toxoplasma gondii*, nuclear pore complex, nucleoporins, gene expression, apicomplexa

Abbreviations:

ATc: AnhydroTetraCycline

Co-IP: Co- ImmunoPrecipitation

FACT: Facilitates Chromatin Transcription

FISH: Fluorescence *in situ* Hybridization

IFA: ImmunoFluorescence Assay

iKD: conditional Knock-Down

NPC: Nuclear Pore Complex

NUP: Nucleoporin

RNA Seq: RNA Sequencing

Sg RNA: Single guide RNA

61 SIM: Structured Illumination Microscopy

62 Tg: Toxoplasma gondii

63

64

65

66

67

68

69

70

71

72

73

74 **Introduction**

75

76 *Toxoplasma gondii* is a unicellular eukaryotic pathogen. It belongs to the apicomplexan phylum, which
77 encompasses some of the deadliest pathogens of medical and veterinary importance, including *Plasmodium* (the
78 cause of malaria), *Cryptosporidium* (responsible for cryptosporidiosis) and *Eimeria* (coccidiosis). *T. gondii* is an
79 obligate intracellular parasite that leads to the development of focal central nervous system infections in patients
80 with HIV/AIDS. In addition, *Toxoplasma* is also a clinically important opportunistic pathogen that can cause birth
81 defects in the offspring of newly infected mothers. The life cycle of *T. gondii* is complex, with multiple
82 differentiation steps that are critical to parasite survival in human and feline hosts [1]. Although gene expression
83 is tightly controlled in Apicomplexa, which is particularly evident during the cell cycle [2] [3], the molecular
84 mechanisms underlying its regulation are poorly understood. Initial studies suggest that histone modifications and
85 chromatin remodeling have an important role in chromatin structure for gene regulation [4] [5]. Eukaryotic nuclei
86 are enclosed by the double-membrane nuclear envelope (NE), which is perforated by large protein structures
87 termed nuclear pore complexes (NPCs), allowing a controlled bidirectional nucleocytoplasmic transport of
88 macromolecules [6]. In eukaryotes, the organization and composition of NPCs has remained conserved throughout
89 evolution, and they play an important role in other biological processes, such as chromosomal segregation
90 mechanisms, mitotic spindle formation, transcription activation and cytokinesis [7]. Each NPC is composed of
91 multiple copies of approximately 30 different proteins known as nucleoporins (NUPs), for which the composition
92 and structure have been largely characterized in *Saccharomyces cerevisiae* [8], mammals [9] [10] and
93 *Trypanosoma brucei* [11,12]. Strikingly, the description of the putative components of the NPC in *T. brucei* led to
94 the discovery of a conserved protein arrangement that spans the eukaryotic kingdom [11]. Three conserved classes
95 of NPC proteins were previously described in yeast and humans [6] : (i) membrane–bound nucleoporins, which
96 link the nuclear pore to the NE; (ii) core-scaffold NUPs, which are restricted at the structural level; and (iii) NUPs
97 that are distributed on the cytoplasmic or nucleoplasmic face of the NE. The three distinct structural arrangements
98 found in eukaryotic core-scaffold NUPs (β-propeller fold, α-solenoid fold and a mixture of both) were found to be
99 conserved in *T. brucei*. Moreover, the sequence of some *T. brucei* core-scaffold NUPs also encompasses
100 phenylalanine-glycine (FG) repeats, which is typical of NUPs [11]. The FG NUPs are known to regulate the
101 transport of molecule through the nuclear pore by extending their FG repeats region into the channel [14] . These
102 results point to a possible ancient inheritance of the structural components of the NPCs [11]. However, it is still
103 unknown whether this apparent structural conservation between Opisthokonts (humans and yeast) and

104 Discicristates (Trypanosoma) also holds true for other distant eukaryotes. Apicomplexan parasites (Alveolates)
105 may provide another point of comparison between eukaryotic groups. However, although the dynamics of the
106 nuclear pores were described during the cell cycle of the erythrocytic forms of *P. falciparum* [15] and for an
107 individual component of the nuclear pore (PfSec13) [16], the composition of the *P. falciparum* NPC remains
108 unknown. In *T. gondii*, the NPC has not been studied, and its components remain to be identified.
109 We characterized the NPC in *T. gondii* by identifying the protein complex associated with a conserved component
110 of the core-scaffold NUPs. In this study, we show that TgNup302 is an essential protein that presents a typical
111 perinuclear staining. In a conditional TgNup302 mutant, controlled bidirectional nuclear-cytoplasmic transport
112 was severely impaired, and gene expression was also perturbed. Immunoprecipitation of proteins associated with
113 TgNup302 provide further characterization of the NPC in *T. gondii* and indicate that structural conservation among
114 eukaryotic NUPs is also true for Alveolates.
115
116 **Materials and Methods**
117
118 *Parasite tissue culture and manipulation*
119
120 *Toxoplasma gondii* strain RHΔKu80 TaTi (a strain with ATc inducible system and high homologous
121 recombination; [17]) tachyzoites were propagated *in vitro* in human foreskin fibroblasts (HFF) using Dulbeccos's
122 modified Eagles medium supplemented with 10% fetal calf serum (FCS), 2mM glutamine, and 1% penicillin-
123 streptomycin. *T. gondii* tachyzoites were grown in ventilated tissue culture flasks at 37°C and 5% CO_2. Transgenes
124 were introduced by electroporation into tachyzoites of *T. gondii* strains and stable transformants were selected by
125 culture in the presence of 2μM pyrimethamine or chloramphenicol (34 mg/ml). Clonal lines were obtained by
126 limiting dilution. Prior to RNA and protein purification, intracellular parasites were purified by sequential syringe
127 passage with 17-gauge and 26-gauge needles and filtration through a 3-μm polycarbonate membrane filter.
128
129 *Generation of transgenic T. gondii strains*
130
131 The TgNup302 iKD line was generated using RHΔku80TaTi and a plasmid containing genomic fragments
132 encompassing 2 kb upstream the gene and 2 kb from the predicted ATG [17]. To produced Myc-tagged TgNup115,
133 TgNup134, TgNup129, TgFACT140, TgNup503, TgNup407, TGGT1_311625 proteins by the knock-in strategy,

134 a DNA fragment of 2kb upstream of the stop codon from the genomic sequence was amplified from genomic DNA
135 of Δku80 RH *T. gondii* (type I strain) and cloned in the pLIC-Myc-CAT plasmid. The plasmid (25µg) was
136 linearized with NheI, ApaI, NcoI, BstBI, NcoI, NsiI, NarI, respectively, and transfected in 5.10^6 TgNup302 iKD
137 tachyzoites followed by chloramphenicol selection. The sequences of all primers used in this study are listed in
138 Supplementary Table 1.
139
140 *Plaque assays and parasite growth assays*
141
142 Plaque assays were performed using 6-well plates containing human fibroblast cells infected with 200 parasites
143 per well in media with or without 1µg/ml ATc, fixed after 7 days post infection and labelled with a crystal violet
144 solution.
145 For parasite growth assays, 8.10^5 parasites per well in a 24-well plate were incubated 4h in normal media or media
146 with ATc. After coverslips were incubated for 24, 48 or 72h, fixation and staining were carried out using
147 formaldehyde and an antibody directed against the TgENO2. The number of parasites per vacuole was counted for
148 100 vacuoles per condition.
149
150 *Antibodies*
151
152 The anti-TgENO2 rabbit [18], anti-TgChromo1 mouse [19] and anti-TgNF3 [20] mouse antibodies were used at
153 1:1000, 1:200 respectively. Anti-HA rabbit (Eurogentech) and rat (Invitrogen) antibodies were used at 1:500 in
154 IFA and in Western blots. Anti-Myc mouse (ThermoFisher) was used at 1:200 in IFA and 1:500 in Western blot.
155
156 *Immunofluorescence Assay, confocal imaging and morphology Microscopy*
157
158 Intracellular and extracellular parasites tachyzoites were fixed with 4% paraformaldehyde in PBS for 15 min,
159 followed by two PBS washes. Extracellular parasites were dried on Teflon slides. Parasites were permeabilized
160 with 0.1% Triton X-100 in PBS containing 0.1% glycine for 10min at room temperature. Samples were blocked
161 with FCS in the same buffer and the primary antibodies were added on parasites in the same buffer for 1h at room
162 temperature. Secondary antibody coupled to Alexa-488 or to Alexa-594 (Molecular probes) diluted at 1:1000 was
163 added in addition to DAPI for nucleus staining. Confocal imaging was performed with an LSM880 microscope

164 (Zeiss) and a Plan Apochromat objective (Plan-Apochromat 63x:1.40 Oil DIC M27, Zeiss). Ultrastructural
165 morphology was performed using conventional microscopy [20], except that 8% paraformaldehyde containing
166 0.01% glutaraldehyde was used for cryo-IEM.
167
168 *Structured illumination microscopy (SIM)*
169
170 SIM was used to obtain high-resolution images using an ElyraPS1 microscope system (Zeiss) with a 100x oil-
171 immersion lens (alpha Plan Apochromat 100x, NA 1.46, oil immersion) and a resolution of 120 nm along the x-y
172 axis and 500 nm along the z-axis (PSF measured on 100 nm beads; Sampling voxel size: 0,050 µm*0,050
173 µm*0,150 µm). Three lasers (405, 488, and 561 nm) were used for excitation. SIM images were acquired with an
174 EMCCD camera (Andor Technology Ltd, UK) and processed with ZEN software; exposure times varied between
175 100 and 120 ms. Three-dimensional images were generated using a z-step of 150 nm (total thickness ~5 µm),
176 while reconstructions and co-distributions were determined with IMARIS software (Pearson's coefficient). We
177 determined the same A (red-HA signal) and B (green-Myc signal) thresholds for all parasite strains. The acquisition
178 was performed sequentially using Zeiss Filter Sets 43HE, 38HE and BP 420-480. Fifteen frames were acquired to
179 reconstruct one image (5 rotations x 3 phases, with a SIM grating period of 51 µm for the blue channel, 42 µm for
180 the green channel, 34 µm for the red channel). One hundred nanometer beads were imaged to measure the
181 chromatic misalignment of our system (fit procedure by the Zen software); these parameters enabled us to further
182 correct the alignment on each acquired multi-channel stack.
183
184 *RNA FISH*
185
186 Intracellular parasites were treated with or without ATc for 48 hours and purified as described earlier. The parasites
187 were fixed with 4% paraformaldehyde for 30 min and settled on RNA-treated slides as previously described [21].
188 Cy3-oligo d(T)40 and Cy3-18S RNA were hybridized on the slide at 50°C overnight and washed as previously
189 described [22][21]. After 10 min DAPI staining, the slides were mounted on coverslips.
190
191 *Cellular fractionation and Western blot for reverse immunoprecipitation*
192

193 Intracellular parasites (5x10^8 tachyzoites) of the TgNup302-HA ΔKu80RHTaTi strain were purified on a 3-mm
194 filter and washed twice with PBS. The parasite pellet was resuspended in 1 mL of NEB1 buffer (10 mM HEPES
195 pH7.9, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT, 0.1 mM EDTA, 0.65% NP40 and 0.5 mM PMSF), incubated
196 on ice for 10 min and centrifuged at 1500 g for 10 min at 4°C. The supernatant was kept as the cytoplasmic extract.
197 The pellet was then resuspended with 100 µl of buffer NEB2 (20 mM HEPES pH7.9, 1.5 mM $MgCl_2$, 420 mM
198 NaCl, 0.2 mM EDTA, 0.5 mM DTT, 25% glycerol and 0.2 mM PMSF) for 10 min on ice and centrifuged at 12000
199 g for 10 min at 4°C.
200 The supernatant was kept as the nuclear extract. The insoluble material was extracted using an SDS buffer (2%
201 SDS, 10 mM Tris and 0.2 mM PMSF) for 20 min at room temperature and centrifuged at 12000 g for 10 min. and
202 the supernatant was kept. Then, 30 µl of Pierce anti-c-Myc tag beads were washed twice with 1X TBS (50 mM
203 Tris-HCl, 150 mM NaCl and 0.5 mM PMSF) with centrifugation of 4000 g for 1 min between each wash.
204 Saturation of the beads was established with 500 µL of 1X TBS and 5 µL of BSA (1 mg/mL) under stirring for 15
205 min at 4°C, and then the beads were washed twice as previously. Nuclear extract and insoluble material are added
206 to the anti-c-Myc beads (the volume was extended to approximately 900 µl of 1X TBS to dilute the salt
207 concentration) and incubated overnight at 4°C under stirring. The next day, the beads were washed five times with
208 1X TBS-T (1X TBS, Tween 20% and 0.5 mM PMSF) and one time with 62.5 mM Tris pH 6.8 + PMSF, with
209 centrifugation at 4000 g for 1 min between each wash. Finally, the immunoprecipitated proteins from the beads
210 were eluted with 1X DTT (Tris 0.5 M pH 6.8, SDS 20%, saccharose, DTT 1 M), warmed at 95°C for 5 min and
211 centrifuged at 14000 rpm for 1 min. The supernatants were kept to perform a Western blot.
212
213 *gRNA CRISPR/Cas9 screening*
214
215 To generate guide RNAs (gRNAs) to disrupt specific TgNup115, TgNup302, TgNup593, TgSec13, TgNup530,
216 TgNup37, TgNup68, TgNup216, TgNup67, TgFACT140, TGGT1_228100, TgNup129, TgNup134, TgNup407,
217 TgNup503 and TGGT1_311625 genes, we modified the CRISPR Cas9-gRNA plasmid (pSAG1::Cas9-
218 U6::sgUPRT plasmid) as previously described [23]. To study the impact of targeted mutations at a specific locus,
219 RH Δhxgprt parasites (10.10^6) were transfected with 100 µg of CRISPR plasmids generated previously by
220 electroporation. To estimate the frequency of CRISPR/CAS9-mediated gene disruptions, 30 µL of transfection
221 reagent was added to HFF monolayers and analyzed by immunofluorescence staining for GFP 24 h post-
222 transfection.

223

224 *Mass spectrometry proteomic analysis*

225

226 After denaturation at 100°C in 5% SDS, 5% β-mercaptoethanol, 1 mM EDTA, 10% glycerol, 10 mM Tris pH 8
227 buffer for 3 min, protein samples were fractionated on a 10% acrylamide SDS-PAGE gel. The electrophoretic
228 migration was stopped as soon as the protein sample entered 1 cm into the separating gel. The gel was briefly
229 labeled with Coomassie Blue, and five bands, containing the whole sample, was cut. In gel digestion of gel slices
230 was performed as previously described [24].

231 An UltiMate 3000 RSLCnano System (Thermo Fisher Scientific) was used for separation of the protein digests.
232 Peptides were automatically fractionated onto a commercial C18 reversed phase column (75 μm×150 mm, 2 μm
233 particle, PepMap100 RSLC column, Thermo Fisher Scientific, temperature 35 °C). Trapping was performed
234 during 4 min at 5μL/min, with solvent A (98 % H2O, 2% ACN and 0.1 % FA). Elution was performed using two
235 solvents A (0,1 % FA in water) and B (0,1 % FA in ACN) at a flow rate of 300 nL/min. Gradient separation was
236 3 min at 5% B, 37 min from 5 % B to 30% B, 5 min to 80% B, and maintained for 5 min. The column was
237 equilibrated for 10 min with 5% buffer B prior to the next sample analysis.

238 The eluted peptides from the C18 column were analyzed by Q-Exactive instruments (Thermo Fisher Scientific).
239 The electrospray voltage was 1.9 kV, and the capillary temperature was 275 °C. Full MS scans were acquired in
240 the Orbitrap mass analyzer over m/z 300–1200 range with resolution 35,000 (m/z 200). The target value was
241 5.00E+05. Ten most intense peaks with charge state between 2 and 4 were fragmented in the HCD collision cell
242 with normalized collision energy of 27%, and tandem mass spectrum was acquired in the Orbitrap mass analyzer
243 with resolution 17,500 at m/z 200. The target value was 1.00E+05. The ion selection threshold was 5.0E+04
244 counts, and the maximum allowed ion accumulation times were 250 ms for full MS scans and 100 ms for tandem
245 mass spectrum. Dynamic exclusion was set to 30 s.

246

247 *Proteomic data analysis*

248

249 Raw data collected during nanoLC-MS/MS analyses were processed and converted into *.mgf peak list format
250 with Proteome Discoverer 1.4 (Thermo Fisher Scientific). MS/MS data was interpreted using search engine Mascot
251 (version 2.4.0, Matrix Science, London, UK) installed on a local server. Searches were performed with a tolerance

252 on mass measurement of 0.2 Da for precursor and 0.2 Da for fragment ions, against a composite target decoy
253 database (50620 total entries) built with 3 strains of *Toxoplasma gondii* ToxoDB.org database (strains ME49, GT1
254 and VEG, release 12.0, September 2014, 25264 entries) fused with the sequences of recombinant trypsin and a list
255 of classical contaminants (46 entries). Cysteine carbamidomethylation, methionine oxidation, protein N-terminal
256 acetylation and cysteine propionamidation were searched as variable modifications. Up to one trypsin missed
257 cleavage were allowed. For each sample, peptides were filtered out according to the cutoff set for proteins hits
258 with 2 or more peptides taller than 7 residues, ion score > 25, identity score > 0 and no false positive identification.
259

260 *Library preparation and RNA Seq*

261 RNA was extracted using Invitrogen Trizol Reagent (cat#15596018), followed by genomic DNA removal and
262 cleaning using an RNase-free DNase. An Agilent 2100 Bioanalyzer was used to assess the integrity of the RNA
263 samples. Only RNA samples having an RNA integrity score between 8 and 10 were used. Library preparation was
264 performed using the TruSeq Stranded mRNA Sample Preparation kit (Illumina) according to the manufacturer's
265 instructions. Libraries were validated using a Fragment Analyzer and quantified by qPCR (ROCHE LightCycler
266 480). Clusters were generated on a flow-cell within a cBot using the Cluster Generation Kit (Illumina), and libraries
267 were sequenced as 50-bp reads on a HiSeq 2000 using a Sequence By Synthesis (SBS) technique (Illumina).
268 Image analysis and base calling were performed using the HiSeq Control Software and Real-Time Analysis
269 component. Demultiplexing was performed using Illumina's conversion software (bcl2fastq 2.17). The quality of
270 the data were assessed using FastQC from the Babraham Institute and the Illumina software Sequence Analysis
271 Viewer (SAV). Potential contaminants were investigated with the FastQ Screen software from the Babraham
272 Institute.

273 Contamination by *Escherichia coli* sequences was evident in the sequenced samples. This contamination, due to a
274 reagent from Life Technologies (SuperScript II enzyme), has been eliminated by aligning sequences (using BWA
275 v0.7.12-r1039 [25]) to the *E. coli* genome and keeping the unmapped reads for downstream analysis.

276 RNA-seq reads were aligned to the *Toxoplasma gondii* genome (ToxoDB-25_TgondiiGT1_Genome.fasta from
277 the ToxoDB Toxoplasma Genomics Resource, downloaded on 08/31/15) with a set of gene model annotations
278 (ToxoDB-25_TgondiiGT1.gff from the ToxoDB Toxoplasma Genomics Resource, downloaded on 08/31/15)
279 using the splice junction mapper TopHat 2.0.13 [26] (with bowtie 2.2.3 [27]). Final read alignments having more
280 than 3 mismatches were discarded. Gene counting was performed using HTSeq-count 0.6.1p1 (union mode) [28].

281 Because the data come from a strand-specific assay, the read must be mapped to the opposite strand of the gene.
282 Before statistical analysis, genes with less than 15 reads (combining all the analyzed samples) were filtered out.
283 edgeR differentially expressed genes were identified using the Bioconductor [29] package edgeR [30] 3.6.7. The
284 data were normalized using the Relative Log Expression (RLE) [31] normalization factors. Genes with adjusted
285 p-values less than 5% (according to the FDR method from Benjamini-Hochberg) were declared differentially
286 expressed.

287 *Quantitative real-time PCR*

288

289 All primers were designed online using Primer2 v.0.4.0 and are listed in Table S1; the cDNA samples were
290 synthesized from total RNA samples using the RevertAid First Strand cDNA Synthesis Kit (Fermentas). qRT-
291 PCR was carried out on an Mx3000P system (Agilent Technologies). Individual reactions were prepared with 0.5
292 µM of each primer, ~5 ng of cDNA and SYBR Green PCR Master Mix (Applied Biosystems, CA) to a final
293 volume of 20 µl. All experiments were performed twice with separate biological replicates. For each experiment,
294 reactions were performed in triplicate, and the expression of individual genes was normalized to the housekeeping
295 tubulin gene Ct values.

296

297 *Quantifications*

298 The differences of localization for TgNup115, TgNup134, TgNup407 and TgNup503 proteins with or without
299 ATc treatment were calculated with Image J Software. We designed a Macro which after manually determining
300 the parasite edge (polygon tool and manual drawing on the phase contrast image) and the nucleus edge (threshold
301 based on DAPI signal) is able to determine the quantity of Myc-tagged signal (in green) in the nucleus and the
302 cytoplasm of parasites for each strains.
303 For each parasite we measured the area in μm^2 of regions of interest (nucleus and cytoplasm) and the integrated
304 and mean intensities in gray levels in each region. With the different values, we calculated the ratio cytoplasm
305 signal versus the nuclear signal with or without ATc treatment for each strains. This ratio corresponds to the
306 integrated intensity for the cytoplasm vs integrated Intensity for the nucleus.

307

308 *Statistics*

309 A non-parametric Student's T-test was performed where statistical analysis was required.

310

311 **Results**

312

313 *TgNup302 is crucial for T. gondii growth*

314

315 Computational searching through the *Toxoplasma* genome database (toxodb.org) using the human Nup98/96
316 protein as a template revealed the existence of a protein (TGGT1_259640) that is conserved in most eukaryotes
317 from yeast to mammals as an integral component of the NPCs. We constructed a phylogenetic tree including some
318 members of the Apicomplexa, Fungi, Plantae and Eumetazoa, where there is a level of characterization of the NPC
319 proteome and some direct experimental information [10] [8] [32] [33] [11]. The Nup98/96 homologs from the
320 Apicomplexa taxon, joining representative branch species such as *Toxoplasma*, *Neospora*, *Eimeria*, *Theileria* and
321 *Babesia*, are relatively close to Eumetazoa, followed by Plantae. Surprisingly, members of the Plasmodium
322 species, which belongs to the Apicomplexa phylum as *Toxoplasma*, present a distant branching taxon separated
323 from the Apicomplexa (Fig. S1). The TgNup302 gene (TGGT1_259640) is predicted to produce a long 2894-
324 amino-acid protein containing an autocleavage domain (Pfam domain: PF04096), that may produce two distinct
325 N-terminal (1079 aa) and C-terminal (1620 aa) peptides after self-cleavage. The presence of GLFG repeats, a
326 typical variant of the FG repeats present in the human Nup98 protein, at its N-terminus and a Nup96-like domain
327 (PF12110) identified by the Pfam database at its C-terminus is typical of this protein family. This structure is
328 conserved among eukaryotes, with Nup98/96 and NUP145 being the representative proteins in humans and yeast,
329 respectively. In vertebrates, Nup98 binds directly to Nup96, the C-terminal half of the proteolytically processed
330 Nup98/Nup96 polyprotein, and is a component of the Nup107-160 complex (the Nup84 complex in yeast) [34]
331 [9]. However, the size of the unprocessed *T. gondii* protein (302 kDa) is much larger than in yeast (145 kDa) and
332 in humans (194 kDa).
333 To better characterize the biological role of this protein, we used a promoter replacement strategy to produce a
334 conditional knock-down (iKD) mutant strain. In this strain, the expression of the TgNup302 transcript is under the
335 control of anhydrotetracycline (ATc) (Fig.S2) [17]. When ATc is added to the culture media, the regulatable
336 promoter is repressed, and the *TgNup302* transcript is no longer produced. We simultaneously added an HA-tag
337 to the N-terminus of the protein in order to follow the protein expression (Fig. S2). Because the human and yeast
338 homologous proteins undergo a self-proteolytic cleavage, producing two polypeptides, we also tagged the
339 TgNup302 protein with a Myc tag at its C-terminus (Fig. S2) in the iKD strain. Therefore, we produced a strain
340 expressing TgNup302 under the control of ATc and tagged at its N- and C-termini. The correct insertion of the

341 construction at the *TgNup302* locus was validated by genomic PCR (Fig.S2). Using anti-HA (α-HA) and anti-Myc
342 (α-Myc) monoclonal antibodies by immunofluorescence assay (IFA), we observed that TgNup302 presented a
343 punctate perinuclear pattern in the parasite, consistent with nuclear pore staining (Fig. 1a). Under ATc treatment,
344 a substantial decrease in TgNup302 protein expression in iKD parasites was observed for both HA- and Myc-
345 tagged proteins. Western blots of total protein extracts from this transgenic parasite line revealed the expression
346 of the two polypeptides tagged with HA- or Myc-tags of the predicted size (150 and 170 kDa, respectively) (Fig.
347 1b) in the absence of ATc indicating the cleavage of the TgNup302 protein into two polypeptides. Notably,
348 TbNup158, the *T. brucei* homologue of TgNup302, does not perform self-cleavage and remains as one polypeptide
349 [11] and therefore is closer to the yeast Nup116 or Nup100 protein.
350 In the presence of ATc, we observed a drastic reduction in the signal from both tagged proteins, with an
351 undetectable level after 48 hours of treatment (Fig. 1b). As a first assessment of the requirement for the *TgNup302*
352 locus, we tested the ability of parasites to grow and produce plaques on host cell monolayers. After 7 days, the
353 wild-type parasites grew normally and developed equal sized plaques in the absence or presence of ATc. Similarly,
354 the growth of the iKD line gave rise to normal plaque numbers only in the absence of ATc. In the presence of ATc,
355 the TgNup302-deficient parasites were unable to proliferate and form plaques (Fig. 1c), indicating a drastic
356 impairment of growth and/or invasion of the parasites lacking the expression of TgNup302. This result was
357 confirmed by an experiment recording the number of parasites per vacuole at a given time (Fig. 1d). While the
358 parental strain with or without ATc showed similar growth to that of the iKD strain without ATc for 48 hours, we
359 noticed that the iKD-HA strain produced vacuoles with a smaller number of parasites. This indicates that the
360 growth of this parasite is impaired by the absence of TgNup302 (Fig. 1d).
361
362 *TgNup302 is involved in nuclear import and essential to 18S RNA export*
363
364 To validate the role of TgNup302 in the nuclear import and export of macromolecular complexes, we examined
365 the localization of a nuclear marker in the iKD strain. Previous work has established the predominant nuclear
366 location of the glycolytic isoenzyme enolase 2 (TgENO2) [35]. We tested the localization of the TgENO2 protein
367 by IFA using a specific antibody (α-TgENO2) in the presence and absence of ATc for 48 hours. We observed
368 predominant nuclear localization of the protein in the parental strain and in the iKD strain in the absence of ATc,
369 as expected (Fig. 2a). However, parasites of the iKD strain grown in the presence of ATc showed a marked
370 delocalization of TgENO2 into the cytoplasm of the parasite (Fig. 2a). This result was confirmed after

371 quantification of the parasites which showed the nuclear or cytoplasmic localization of TgENO2 (Fig. 2b). While
372 the iKD strain without ATc and the parental parasites exhibited predominant nuclear localization of TgENO2,
373 most of the ATc-treated iKD parasites exhibited cytoplasmic localization, indicating a defect in the transport of
374 this protein in the mutant parasites.
375 To investigate whether TgNup302 is involved in mRNA export, we examined the cellular distribution of
376 poly(A)+RNA in the parental and TgNup302 iKD strains via RNA fluorescence *in situ* hybridization (FISH) with
377 labeled oligodT after 48 hours with or without ATc treatment. In the parental and iKD strains, localization of the
378 polyA+ RNA was mostly cytoplasmic (Fig. S3). We also examined the localization of the 18S ribosomal RNA
379 using a specific probe. In the wild-type and iKD strains without 48 hours of ATc treatment, most 18S RNA was
380 found to be cytoplasmic (Fig. 2c). In contrast, the 18S RNA signal was mainly nuclear in the iKD parasites treated
381 with ATc (Fig. 2c and d). These results demonstrate that TgNup302 is essential to 18S RNA export, a typical role
382 of the NPC. To test whether TgNup302 was involved in other proteins nuclear import, we used IFA of the parental
383 and iKD-HA strains to evaluate the localization of a centromeric protein (TgChromo1) that was previously
384 described to specifically bind pericentromeric heterochromatin [19]. We also tested the localization of a nucleolar
385 marker, (TgNF3). No differences in the localization of these proteins was observed between the parental and iKD-
386 HA strains in all conditions tested (with or without ATc for 48 hours) (Fig. S4).
387 The morphology of the parasite was assessed using electron microscopy for parental and TgNup302 iKD
388 intracellular parasites following 48 h of growth with or without ATc treatment. No morphological differences were
389 observed between the parental and TgNup302 iKD strains, which present a typical nucleus. ATc treatment had
390 also no particular impact on the presence of other organelles, such as rhoptries and micronemes, for each strain,
391 suggesting that TgNup302 is not essential for nuclear morphology (Fig. S5).
392
393 *Deep sequencing suggests a role for TgNup302 in gene expression*

394 RNA sequencing (RNA Seq) was carried out to study gene expression dynamics in the TgNup302 iKD. Total
395 RNA was purified from tachyzoites of the TgNup302 iKD strain grown under normal growth conditions with or
396 without ATc treatment for 48 hr (in triplicate). RNA-seq reads were aligned to the *Toxoplasma gondii* genome
397 with a set of gene model annotations using the splice junction mapper TopHat 2.0.13. The data were normalized
398 using the relative log expression (RLE) normalization factors. Genes with adjusted p-values less than 5%
399 (according to the FDR method from Benjamini-Hochberg) were considered differentially expressed. Data analysis
400 revealed significant changes in the transcription profile of *T. gondii,* with 145 genes upregulated and 65 genes

401 downregulated (Table S2). The predominance of upregulated genes suggests a role for the NPC in the repression
402 of gene expression. We examined the differential chromosomal distribution of differentially expressed genes and
403 did not identify a cluster of differentially regulated genes (Fig. S6). Collectively, these data suggest that TgNup302
404 could play a role as a repressor and its absence modulates the transcriptional regulation of genes that are normally
405 repressed at the tachyzoite stage of *T. gondii*. Validation of the RNA Seq data was also carried out using qRT-
406 PCR in RNA samples from the iKD TgNup302 strain with and without ATc treatment for 48 hours. For that, we
407 selected 5 upregulated genes with a logFC>-2 (TGGT1_258670, TGGT1_301150, TGGT1_267160,
408 TGGT1_270273, TGGT1_360460) (Fig. 3a) and TgNup302 (TGGT1_259640) (Figure 3b) as a control. As
409 expected, Figure 3a clearly shows that ATc treatment induced an increased steady-state mRNA level of all five
410 upregulated genes, while the TgNup302 transcript was downregulated in the iKD samples treated with ATc.

411

412 *New components of the T. gondii nuclear pore*

413

414 NPC components are unknown in *T. gondii*. To uncover the components of the NPC in *T. gondii*, we performed
415 co-immunoprecipitation (co-IP) experiments on the iKD Nup302-HA parasites using α-HA antibody followed by
416 mass spectrometry identification of the co-IP proteins (from two experiments). As a control, we performed an α-
417 HA immunoprecipitation on protein extracts from the parental strain, RHΔKu80TaTi. We selected the proteins
418 identified in both experiments but not identified in the control and those present in the control with a single peptide
419 (Table S3 and Table 1). Among the proteins identified and listed in Table 1, we confirmed the presence of
420 TgNup302 (TGGT1_259640) with peptides spanning both the N-terminus and the C-terminus (Fig. S7), indicating
421 that both products of TgNup302 autocleavage are present in the IP fraction. As shown for the human Nup98/Nup96
422 protein [36], both polypeptides may also interact in *T. gondii.*
423 Most of the identified proteins did not share primary sequence conservation with known nucleoporins using
424 BLAST searches. However, a conserved structure was identified between distant eukaryotic NUPs [11], we
425 therefore examined the secondary structure prediction of the selected proteins (Table 1, Fig. S8). We also searched
426 the proteins for other motifs using the Pfam database. Among the co-immunoprecipitated proteins, five putative
427 NUPs encompassing FG-repeat motifs (TgNup593, TgNup37, TgNup68, TgNup216 and TgNup67) were
428 identified. We also found proteins with a β-propeller fold (TgSec13, TGGT1_311625 and TGGT1_228100),
429 including a homolog of the yeast protein Sec13, shown to be part of the Y complex through its interaction with
430 Nup96 [37]. Interestingly, six of the TgNup302-associated proteins (TgNup530, TgNup407, TgNup129,

431 TgNup503, TgNup115 and TgNup134) had no recognizable Pfam motif, with the exception of TgNup115, which
432 contains a Zinc-finger domain (CCCH). However, these proteins do contain an α-solenoid fold, a structure also
433 found in other eukaryotic core-scaffold NUPs (Fig. S8). Moreover, two of these proteins exhibited large segments
434 of coiled-coil structure (TgNup530, TgNup407) (Fig. S8). We further examined the sequence of these proteins
435 using a secondary structure prediction software [38]. This allowed the identification of a domain in TgNup68 with
436 a strong homology to the yeast Nsp1 crystal structure (Table 1). Similarly, a domain in TgNup129 was homologous
437 to the yeast Nic96 structure (Table 1). These data indicate that the proteins composing the *T. gondii* NPC may
438 have retained the same level of structural conservation as *T. brucei* NUPs. Notably, the size of the identified
439 proteins substantially exceeded the size of their putative homologs. Therefore, the assigned homologs listed in
440 Table 1 are based on the structural properties of the proteins rather than their expected size.
441 As expected, TgNup302 is also associated with the putative homolog of the ribonucleic acid export 1 protein
442 (TgRae1, TGGT1_272350), which is implicated in the mRNA export pathway in other eukaryotes, which also
443 interact with the GLFG Nup98 protein [39]. Surprisingly, we also found the two proteins (TgFACT140 and
444 TgFACT80) corresponding to the homologs of the proteins composing the FACT complex, which is categorized
445 as a histone chaperone critical for nucleosome reorganization during replication and transcription [26] [27] [28].
446 This observation indicates a potential interaction between the FACT complex and the NPC in *T. gondii*.
447
448 *Identified proteins interact with TgNup302*
449
450 To confirm the mass-spectrometry results, we performed a reverse IP of TgNup302 using the identified partner
451 proteins. Protein extracts from parasite strains expressing each Myc-tagged protein in the TgNup302-HA iKD
452 background were immunoprecipitated using anti-Myc tag beads. The TgNup302-HA iKD strain was used as a
453 negative control. We then performed a Western blot using an anti-HA antibody to detect TgNup302 in the eluates
454 of the Myc-tagged immunoprecipitates (Fig. 4). As shown in Figure 4a, the anti-Myc antibody failed to
455 immunoprecipitate the HA-tagged TgNup302 in the absence of Myc-tagged proteins, as expected. In contrast, we
456 were able to confirm the interaction between TgNup302 and the TgFACT140 protein, indicating a link between
457 the NPC and the FACT complex in *T. gondii* (Fig. 4a, line 2, top panel). We tested whether some of the unknown
458 proteins with an α-solenoid fold listed in Table 1 were also able to co-immunoprecipitate TgNup302. We
459 confirmed the presence of the TgNup302 protein in the TgNup129 co-immunoprecipitated proteins (Fig. 4a, line
460 3). Similarly, TgNup302 was detected after immunoprecipitation of TgNup134 and TgNup407, indicating that
461 these proteins of unknown function may participate in the NPC (Fig. 4a, lines 4 and 5, top panel). Western blot

462 membranes were probed with an anti-Myc antibody to detect specific Myc-tagged proteins (Fig. 4a, bottom panel).
463 Using the anti-Myc antibody, the immunoprecipitated proteins extracted from the TgNup302 iKD parasite line
464 showed no signal, as expected (Fig. 4a, line 1, bottom panel). We confirmed the presence of each Myc-tagged
465 protein at the expected molecular weights for TgFACT140 (135 kDa), for TgNup129 and TgNup134 (130 kDa)
466 and for TgNup407 (115 kDa). We verified that the level of the TgNup302 protein was similar in the starting
467 material of each strain (Figure S9a). Similarly, the presence of the myc-tagged proteins was checked for the same
468 input samples (Figure S9b). As a control, we repeated this experiment using a strain expressing TgNup302-HA
469 and a myc-tagged protein (TgChromo1) that was not identified by mass-spectrometry (Figure 4b). After
470 immunoprecipitation with anti-myc coated beads, we performed a Western blot using an anti-HA antibody to
471 detect TgNup302 in the input and eluates (Fig. 4b left panel). We were able to confirm the interaction between
472 TgNup302 and the TgNup134 protein (Fig. 4b, line 2, left panel). As expected, TgNup302 is not detected in the
473 negative control as well as in the TgChromo1 co-immunoprecipitated proteins (Fig. 4b, lines 4 and 6 left panel)
474 but is present in the input (Fig. 4b, lines 1, 2, 3, left panel). Western blot membranes were probed with an anti-
475 Myc antibody to detect specific Myc-tagged proteins in the input and eluates (Fig. 4b, right panel). Using the anti-
476 Myc antibody, the immunoprecipitated proteins extracted from the TgNup302 iKD parasite line showed no signal,
477 as expected (Fig. 4b, line 1 and 4, right panel). We confirmed the presence of each Myc-tagged protein in the
478 eluates for TgChromo1 (Fig. 4b, line 6, right panel) and TgNup134 (Fig. 4b, line 5, right panel) confirming that
479 the myc-tagged proteins were indeed immunoprecipitated.
480 These results confirm the interaction between the NPC protein TgNup302 and the selected binding partners
481 identified by mass-spectrometry.
482
483 *Identified proteins localize to the T. gondii nuclear pore*
484
485 To gain more information about these unknown proteins presenting an α-solenoid fold, we tagged several of them
486 using a Myc-tag in the iKD TgNup302-HA strain. As a control, we tagged a component of the FACT complex
487 (TgFACT140) whose localization was anticipated to be nuclear. Using IFA, we observed that the five potential
488 partners exhibited a TgNup302-like pattern (Figs. 5a, 6a and S9) in the absence of ATc (after 48 hours of growth),
489 indicating that they may indeed be components of the NPC. Therefore, we propose to rename them as TgNup
490 proteins, as suggested in Table 1. As anticipated, the TgFACT140 protein showed strong nuclear localization that

491 extended to the perinuclear periphery (Fig. S9). Notably, the TgNup503 protein showed a more spread pattern that
492 extended toward the cytoplasm of the cell (Fig. 5a) and might not be connected to the nuclear envelope.
493 We next examined whether the localization of these proteins was perturbed in the TgNup302-deficient parasites.
494 We performed IFA and measured the quantity of signal in the cytoplasm and in the nucleus (as defined by DAPI
495 staining) in the absence and presence of ATc (Fig. 5b) in multiple vacuoles. We identified two proteins
496 (TgNup115, TgNup134) with a greater proportion of the fluorescence in the cytosol in the presence of ATc (Fig.
497 5b). In contrast, TgNup503 presented a distribution that remained unchanged after ATc treatment (Fig. 5b).
498 Notably, we observed that TgNup407 localization was perinuclear in the absence of ATc and changed in the
499 presence of ATc, with a more pronounced localization inside the nucleus (Fig. 6a). We measured the quantity of
500 signal present at the nuclear periphery and inside the nucleus in the presence and absence of ATc (after 48 hours)
501 (Fig. 6b). We found that there was much more signal in the nucleus than at the nuclear periphery in the presence
502 of ATc compared to in the absence of ATc (Fig. 6b). We examined the same ratio for TgNup129 protein
503 localization, which exhibited a punctate perinuclear pattern that remained unchanged by ATc treatment (Fig. S10).
504 Similarly, TgFACT140 exhibited a nuclear localization that was not perturbed by ATc treatment (Fig. S10).
505
506 *Structured illuminated microscopy imaging of TgNup302 and its binding partners.*
507
508 To better characterize the structure of the NPC in *T. gondii*, we performed structured illuminated microscopy
509 (SIM) on selected parasite strains. For each strain, colocalization between TgNup302-HA and its Myc-tagged
510 protein partner was determined by Imaris software (Pearson's coefficient). Hence, we determined a Pearson's
511 coefficient average of 0.65 for the Myc-tagged TgNup302 C-terminus, 0.46 for TgNup67, 0.63 for TgNup129,
512 0.28 for TgNup503, 0.23 for TgNup115 and 0.56 for TgNup134 with the HA-tagged TgNup302 protein (Fig. 7a).
513 Using SIM microscopy, we observed for four proteins (TgNup302 C-terminus, TgNup67, TgNup129 and
514 TgNup134) between 4 and 8 perinuclear co-distribution foci (indicated with white arrows) per parasite,
515 representing co-distribution between HA (red) and Myc (green) staining (Fig. 7b). This may indicate that the
516 parasite has 4 or 8 complete NPC structures per nuclei, a number that is comparable to what was observed in *P.*
517 *falciparum* (Fig. 7c) [15] [40]. In contrast, TgNup115 and TgNup503 had lower Pearson's coefficients, displaying
518 a lower number of foci per parasite and indicating that these proteins may only be transiently interacting with
519 NPCs (Figs. 7b and c).
520

521 *CRISPR/Cas9 screening reveals the importance of the TgNup503 and TgNup134 proteins for nuclear-cytoplasmic*
522 *shuttling.*
523
524 Previous studies have shown that combining CRISPR/CAS9 with sgRNAs can be used to generate site-specific
525 double-strand breaks in the target DNA that are repaired by non-homologous-end joining, leading to short
526 insertions and deletions that inactivate the gene [41]. Recently, the CRISPR/CAS9 system was adapted to *T. gondii*
527 with a plasmid expressing a nuclear-localized CAS9 fused to green fluorescent protein (CAS9-NLS-GFP) driven
528 by the SAG1 promoter and a single guide RNA (sgRNA) driven by the *T. gondii* U6 (TgU6) promoter [23]. We
529 used CRISPR/Cas9 and a single gRNA placed downstream of the ATG to target the fifteen proteins previously
530 identified by co-immunoprecipitation with the TgNup302 protein (Table 1). As a control, we used a sgRNA
531 targeting a non-essential gene (TgAlba1) [21]. For each construct, GFP expression was monitored to determine
532 the transfection efficiency at 24 h after electroporation, revealing that ~30 to 70% of cells received the plasmid
533 (Fig. S11). We used IFA to determine the localization of the TgENO2 protein, a nuclear marker. We determined
534 the percentage of vacuoles that were GFP-negative and had cytoplasmic TgENO2 staining, as a control. We also
535 scored the percentage of vacuoles that were GFP-positive (CAS9-expressing) and had cytoplasmic TgENO2
536 staining, illustrating the impact of the potentially mutated gene on the TgENO2 localization (Fig. 8). As expected,
537 no differences were observed between these two scores for the parasite expressing CAS9 and the sgRNA targeting
538 TgAlba1. Similarly, most of the targeted genes did not result in an increase in the number of vacuoles that had
539 cytoplasmic TgENO2 and were GFP-positive. In contrast, the parasite expressing Cas9 and an sgRNA directed
540 against the TgNup302 gene presented a drastic increase of the cytoplasmic TgENO2 signal (Fig. 8), in good
541 concordance with the phenotype observed with the iKD strain (Fig. 2b). Interestingly, for sgRNA targeting the
542 TgNup503 and TgNup134 genes, we observed a higher percentage (93 and 57%, respectively) of GFP-positive
543 vacuoles with cytoplasmic TgENO2 staining than that of GFP-negative vacuoles with cytoplasmic TgENO2
544 staining (14 and 16%, respectively) (Fig. 8). Collectively, these results may indicate that in addition to TgNup302,
545 the TgNup503 and TgNup134 proteins are essential for the nuclear localization of TgENO2.
546
547
548 **Discussion**
549
550 Nuclear pore complex components have been widely studied in yeast and humans. More recently, NUP proteins
551 were uncovered in a distant eukaryote, *T. brucei* [42,43]. Strikingly, structural conservation is a key feature of *T.*

552 *brucei* core-scaffold NUPs compared to those in yeast and humans. Using *T. gondii* as a model Apicomplexa, we
553 investigated the role and the composition of the nuclear pore from another distant eukaryotic branch, the
554 Alveolates. Using a conserved nuclear pore protein (TgNup302) homolog of the NUP98/96 and NUP145 proteins
555 in humans and yeast, respectively, we identified potential interacting proteins. Interestingly, TgNup302 has all the
556 features of its human and yeast homologs, including the family-specific GLFG repeats. *T. brucei* presents only
557 one GLFG repeat protein that does not contain an autocleavage domain, as observed for the yeast Nup100 protein.
558 This indicates that in distant eukaryotes, the GLFG repeat motif is a key feature that was retained during evolution.
559 Because of the stringency of our protocol and as shown by the characterization of the identified proteins, we
560 believe that a significant number of *T. gondii* NUPs were identified during the course of this study. Among these
561 proteins, we purified proteins encompassing FG (phenylalanine-glycine)-repeat motifs, which are distinctive of
562 NPC components. In fact, it was shown that a proportion of NUPs are composed of FG repeats, which function is
563 to mediate the passage of transport receptors and their cargos with selectively gated transport [7].
564 Moreover, other proteins, bearing no recognizable motifs, presented structural features that are shared with
565 eukaryotic NUPs. We further showed that their localization corresponds to known *T. gondii* NUPs, indicating that
566 they may participate in the parasite NPC. This correlates with earlier studies identifying *T. brucei* NUPs [42,43].
567 Indeed, structural conservation led to the hypothesis of an ancient inheritance of the core-scaffold NUPs. This
568 hypothesis is confirmed by our study identifying NUPs in Alveolates. We also noted that most of the proteins that
569 co-immunoprecipitated with TgNup302 were very large compared with their eukaryotic counterparts. This is also
570 true for the TgNup302 protein itself. This may indicate that fewer proteins are needed to produce a functional
571 nuclear pore in *T. gondii*.
572 During the course of this study, we also characterized the biological function of TgNup302, demonstrating its role
573 in the nuclear transport of proteins and rRNA. When TgNup302 was depleted, these vital functions were impaired,
574 and as a consequence, the parasite rapidly dies. It is worth noting that the nuclear pore structure is still present in
575 the TgNup302-depleted parasites because some of the nuclear pore proteins exhibited normal localization
576 (TgNup67 and TgNup134). The remaining nuclear pore structure may also still be sufficient for the active transport
577 of the large ribonucleoprotein complexes, insuring the nuclear export of polyA+ mRNAs. This indicates that
578 TgNup302 performs specific roles in the complex, as shown for NUP98/96 and NUP145, the representative human
579 and yeast homologs, respectively. In fact, deletion of the essential *Nup145* gene results in a defect not in protein
580 import but in polyA+ RNA export [44] [45]. Moreover, NUP98 plays a specific role in RNA export from the
581 nucleus, and it appears to be an essential component of multiple RNA export pathways [46] [47]. NUP98 functions

582 as a docking protein for the cytosol-mediated docking of a model import substrate. The docking function was
583 localized to the N-terminal half of NUP98, which contains all its GLFG repeats [48]. However, we cannot exclude
584 the possibility that an undetectable amount of TgNup302 is still present at the nuclear pore after 48 hours of ATc
585 treatment, therefore allowing the passage of mRNAs but not rRNA. We also noted that the nuclear marker
586 TgENO2 was mainly cytoplasmic in the TgNup302-depleted parasite. However, the localization of a nucleolar
587 marker, TgNF3 or that of a pericentromeric marker (TgChromo1), was not affected after 48h. This may be
588 explained by the presence of multiple nuclear import pathways in the parasite, as previously suggested by a
589 genomic identification of potential import factors [49]. While the expected molecular weight of TgNF3 (34 kDa)
590 would allow this protein to diffuse freely in the nucleus through the nuclear pore, TgChromo1 (98 kDa) exceeds
591 the size cut-off (approximately 40 kDa) for passive diffusion of molecules. Alternatively, this might reflect the
592 differences in the dynamics of nuclear import of these proteins, especially in parasites for which growth is altered
593 by TgNup302 depletion.
594 Using SIM imaging, we identified two proteins (TgNup115 and TgNup503) that associate with TgNup302 to a
595 much lesser extent. These proteins may represent transporters that transiently associate with the NPC and should
596 be investigated further, as this class of proteins is under-represented in Apicomplexa genomes [49]. In particular,
597 TgNup115 presents a CCCH zinc finger that may bind RNA and therefore play a role in RNA transport. Overall,
598 we estimated that the number of nuclear pores in *T. gondii* tachyzoites to be between 4 and 8, a number that is
599 close to what was observed in *P. falciparum* [15]. However, epitope-tag availability and the microscopy limitations
600 may have hampered the identification of proteins and therefore may have led to an under-estimation of the pore
601 number.
602 To further investigate the potential components of the *T. gondii* nuclear pore, we performed a CRISPR/Cas9
603 screen. We used the ability of the *T. gondii* RH strain to perform non-homologous recombination after Cas9
604 cleavage, leading to mutations in the target gene that may impair its expression. Using TgENO2 as a nuclear
605 marker, we identified two other proteins (TgNup503 and TgNup134) that showed a marked defect in TgENO2
606 localization in addition to TgNup302. Given the TgNup503 localization and its association with TgNup302, we
607 believe that these data reinforce the hypothesis that this protein play a major role in the TgENO2 import and
608 potentially other proteins. TgNup503 localization is not dependent on the presence of TgNup302, and therefore it
609 may associate with other proteins of the NPC.
610 The TgNup134 protein strongly associates with TgNup302 and is delocalized in the absence of TgNup302.
611 Therefore, it may associate with the NPC through its interaction with TgNup302 and may represent a peripheral

612 NUP. TgNup134 may also have a role in nuclear import. Further studies should be performed to validate this
613 hypothesis.
614 In *T. gondii*, previous studies hypothesized a model of centromere sequestration, in which centromere attachment
615 to the centrocone, a subcompartment of the nucleus, serves as an organizer of chromosome apical region location
616 throughout the parasite cell cycle [19,50]. We hypothesized that centromeric heterochromatin could be anchored
617 through protein components of the nuclear membranes via nucleoporins. In the TgNup302 iKD mutant, in the
618 presence of ATc, we showed that a marker of the pericentromeric heterochromatin is still sequestered at all times
619 of the cell cycle, indicating that TgNup302 depletion is not sufficient to uncouple the potential link between nuclear
620 pore and centromeric heterochromatin. However, as previously shown, TgNup302 depletion is not sufficient to
621 fully abrogate the NPC presence at the nuclear membrane. Therefore, other proteins in the remaining NPC may
622 still interact with centromeric heterochromatin. Moreover, TgNup302, a FG nucleoporin, may not have a direct
623 access to chromatin as nuclear basket nucleoporin would [51]. Alternatively, proteins directly interacting with the
624 nuclear membrane may be involved in centromere sequestration at the nuclear periphery.
625 Finally, RNA-seq revealed that most of the transcripts that are impacted in the iKD mutant when TgNup302 is
626 depleted are upregulated. This indicates a potential role of the TgNup302 protein in regulating gene expression. .
627 Further experiments are needed to explore the potential role of the NPC in *T. gondii* gene expression. Because the
628 parasites lacking TgNup302 are severely affected, a kinetic experiment was not performed and would be useful to
629 identify the genes that are impacted early after the depletion of TgNup302. In higher eukaryotes, an alternative
630 splice variant of the Nup98/96 transcript produces a shortened version of the Nup98/96 polypeptide corresponding
631 to the Nup98 protein with a stop codon inserted after the auto-proteolytic domain. This polypeptide is found in
632 the intranuclear pool and can interact with chromatin [52]. Such splicing variants were not detected by in the RNA-
633 seq data available on ToxoDB. Interestingly, TgNup129 was found to be a structural homolog of the Nup93 (or
634 Nic96 in yeast) protein. This protein was shown to interact at multiple loci on chromatin [53]. Further experiments
635 should be performed to validate this hypothesis.
636 The presence of the HDAC3 protein in the IP (Table S3), may indicate that *T. gondii* NPCs are recruiting
637 repression-specific complexes at the nuclear periphery. TgHDAC3 has been implicated in the control of the
638 expression of bradyzoite-specific genes during differentiation [54]. Therefore, the NPC may have a role in
639 maintaining the repression state for those loci. However, these genes were scattered in the genome, which may
640 indicate that some genomic regions may contact the nuclear periphery, resulting in the repression of gene

641 expression, a mechanism widely recognized in other eukaryotes [51] that was never investigated in *T. gondii* until
642 now.
643 We also showed that TgNup302 is able to interact with two members of the FACT complex. Indeed, the FACT
644 complex is involved in the reorganization of the nucleosome structure during crucial processes such as DNA
645 replication and transcription elongation [55]. Proteins involved in transcription elongation have already been
646 associated with the NPC [56], but the specific role of the FACT complex at the nuclear periphery was not
647 investigated. Notably, a genetic interaction was identified in yeast between the Pob3 gene (a homolog of
648 TgFACT80) and the Nup100 gene, coding for a GLFG NUP (homologous to TgNup302) [57]. The FACT complex
649 is also known to be required for proper export of mRNA either by favoring an efficient mRNA splicing [58] or
650 through a direct interaction with mRNA export factors [59]. Therefore, we hypothesize that *T. gondii*, and possibly
651 other apicomplexan parasites, may have evolved a strong link between gene transcription, elongation, splicing and
652 export of mRNA. The FACT complex might therefore facilitate the export of mRNPs through the NPC in *T. gondii*.
653 This would allow, as shown for other eukaryotes, the rapid expression of genes in *T. gondii.*
654 In conclusion, we uncovered novel members of the nuclear pore in *T. gondii* and also identified novel proteins that
655 may have a role in the nuclear-cytoplasmic transport of proteins. TgNup302 is an essential protein that may be
656 used as a tool to investigate the role of the *T. gondii* nuclear pore in regulating gene expression.
657

658 ***Bibliography***

659 [1] K. Kim, L.M. Weiss, Toxoplasma gondii: the model apicomplexan, Int J Parasitol. 34 (2004) 423–32.
660 [2] M.S. Behnke, J.B. Radke, A.T. Smith, W.J. Sullivan, M.W. White, The transcription of bradyzoite genes
661 in Toxoplasma gondii is controlled by autonomous promoter elements, Mol. Microbiol. 68 (2008) 1502–
662 1518. doi:10.1111/j.1365-2958.2008.06249.x.
663 [3] M.S. Behnke, J.C. Wootton, M.M. Lehmann, J.B. Radke, O. Lucas, J. Nawas, L.D. Sibley, M.W. White,
664 Coordinated progression through two subtranscriptomes underlies the tachyzoite cycle of Toxoplasma
665 gondii, PLoS ONE. 5 (2010) e12354. doi:10.1371/journal.pone.0012354.
666 [4] W.J. Sullivan, M.A. Hakimi, Histone mediated gene activation in Toxoplasma gondii, Mol Biochem
667 Parasitol. 148 (2006) 109–16.
668 [5] M. Gissot, K.A. Kelly, J.W. Ajioka, J.M. Greally, K. Kim, Epigenomic Modifications Predict Active
669 Promoters and Gene Structure in Toxoplasma gondii, PLoS Pathog. 3 (2007) e77.
670 [6] F. Alber, S. Dokudovskaya, L.M. Veenhoff, W. Zhang, J. Kipper, D. Devos, A. Suprapto, O. Karni-
671 Schmidt, R. Williams, B.T. Chait, A. Sali, M.P. Rout, The molecular architecture of the nuclear pore
672 complex, Nature. 450 (2007) 695–701. doi:10.1038/nature06405.
673 [7] M. Capelson, M.W. Hetzer, The role of nuclear pores in gene regulation, development and disease, EMBO
674 Rep. 10 (2009) 697–705. doi:10.1038/embor.2009.147.
675 [8] M.P. Rout, J.D. Aitchison, A. Suprapto, K. Hjertaas, Y. Zhao, B.T. Chait, The yeast nuclear pore complex:
676 composition, architecture, and transport mechanism, J. Cell Biol. 148 (2000) 635–651.
677 [9] S.K. Vasu, D.J. Forbes, Nuclear pores and nuclear assembly, Curr. Opin. Cell Biol. 13 (2001) 363–375.
678 [10] J.M. Cronshaw, A.N. Krutchinsky, W. Zhang, B.T. Chait, M.J. Matunis, Proteomic analysis of the
679 mammalian nuclear pore complex, J. Cell Biol. 158 (2002) 915–927. doi:10.1083/jcb.200206106.
680 [11] J.A. DeGrasse, K.N. DuBois, D. Devos, T.N. Siegel, A. Sali, M.C. Field, M.P. Rout, B.T. Chait, Evidence
681 for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic
682 ancestor, Mol. Cell Proteomics. 8 (2009) 2119–2130. doi:10.1074/mcp.M900038-MCP200.

683 [12] S.O. Obado, M. Brillantes, K. Uryu, W. Zhang, N.E. Ketaren, B.T. Chait, M.C. Field, M.P. Rout,
684 Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex, PLoS Biol. 14
685 (2016) e1002365. doi:10.1371/journal.pbio.1002365.
686 [13] J.A. DeGrasse, B.T. Chait, M.C. Field, M.P. Rout, High-yield isolation and subcellular proteomic
687 characterization of nuclear and subnuclear structures from trypanosomes, Methods Mol. Biol. 463 (2008)
688 77–92. doi:10.1007/978-1-59745-406-3_6.
689 [14] R.Y.H. Lim, B. Fahrenkrog, J. Köser, K. Schwarz-Herion, J. Deng, U. Aebi, Nanomechanical Basis of
690 Selective Gating by the Nuclear Pore Complex, Science. 318 (2007) 640–643.
691 doi:10.1126/science.1145980.
692 [15] A. Weiner, N. Dahan-Pasternak, E. Shimoni, V. Shinder, P. von Huth, M. Elbaum, R. Dzikowski, 3D
693 nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria
694 parasite Plasmodium falciparum, Cell. Microbiol. 13 (2011) 967–977. doi:10.1111/j.1462-
695 5822.2011.01592.x.
696 [16] N. Dahan-Pasternak, A. Nasereddin, N. Kolevzon, M. Pe'er, W. Wong, V. Shinder, L. Turnbull, C.B.
697 Whitchurch, M. Elbaum, T.W. Gilberger, E. Yavin, J. Baum, R. Dzikowski, PfSec13 is an unusual
698 chromatin-associated nucleoporin of Plasmodium falciparum that is essential for parasite proliferation in
699 human erythrocytes, J. Cell. Sci. 126 (2013) 3055–3069. doi:10.1242/jcs.122119.
700 [17] L. Sheiner, J.L. Demerly, N. Poulsen, W.L. Beatty, O. Lucas, M.S. Behnke, M.W. White, B. Striepen, A
701 systematic screen to discover and analyze apicoplast proteins identifies a conserved and essential protein
702 import factor, PLoS Pathog. 7 (2011) e1002392. doi:10.1371/journal.ppat.1002392.
703 [18] F. Dzierszinski, M. Mortuaire, N. Dendouga, O. Popescu, S. Tomavo, Differential expression of two plant-
704 like enolases with distinct enzymatic and antigenic properties during stage conversion of the protozoan
705 parasite Toxoplasma gondii, J. Mol. Biol. 309 (2001) 1017–1027. doi:10.1006/jmbi.2001.4730.
706 [19] M. Gissot, R. Walker, S. Delhaye, L. Huot, D. Hot, S. Tomavo, Toxoplasma gondii chromodomain protein
707 1 binds to heterochromatin and colocalises with centromeres and telomeres at the nuclear periphery, PLoS
708 ONE. 7 (2012) e32671. doi:10.1371/journal.pone.0032671.
709 [20] A. Olguin-Lamas, E. Madec, A. Hovasse, E. Werkmeister, I. Callebaut, C. Slomianny, S. Delhaye, T.
710 Mouveaux, C. Schaeffer-Reiss, A. Van Dorsselaer, S. Tomavo, A novel Toxoplasma gondii nuclear factor
711 TgNF3 is a dynamic chromatin-associated component, modulator of nucleolar architecture and parasite
712 virulence, PLoS Pathog. 7 (2011) e1001328. doi:10.1371/journal.ppat.1001328.
713 [21] M. Gissot, R. Walker, S. Delhaye, T.D. Alayi, L. Huot, D. Hot, I. Callebaut, C. Schaeffer-Reiss, A.V.
714 Dorsselaer, S. Tomavo, Toxoplasma gondii Alba proteins are involved in translational control of gene
715 expression, J. Mol. Biol. 425 (2013) 1287–1301. doi:10.1016/j.jmb.2013.01.039.
716 [22] J. Thompson, In situ detection of RNA in blood- and mosquito-stage malaria parasites, Methods Mol.
717 Med. 72 (2002) 225–233. doi:10.1385/1-59259-271-6:225.
718 [23] B. Shen, K.M. Brown, T.D. Lee, L.D. Sibley, Efficient gene disruption in diverse strains of Toxoplasma
719 gondii using CRISPR/CAS9, MBio. 5 (2014) e01114-01114. doi:10.1128/mBio.01114-14.
720 [24] L. Miguet, G. Béchade, L. Fornecker, E. Zink, C. Felden, C. Gervais, R. Herbrecht, A. Van Dorsselaer, A.
721 van Dorsselaer, L. Mauvieux, S. Sanglier-Cianferani, Proteomic analysis of malignant B-cell derived
722 microparticles reveals CD148 as a potentially useful antigenic biomarker for mantle cell lymphoma
723 diagnosis, J. Proteome Res. 8 (2009) 3346–3354. doi:10.1021/pr801102c.
724 [25] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform,
725 Bioinformatics. 25 (2009) 1754–1760. doi:10.1093/bioinformatics/btp324.
726 [26] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, TopHat2: accurate alignment of
727 transcriptomes in the presence of insertions, deletions and gene fusions, Genome Biol. 14 (2013) R36.
728 doi:10.1186/gb-2013-14-4-r36.
729 [27] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2, Nat. Methods. 9 (2012) 357–359.
730 doi:10.1038/nmeth.1923.
731 [28] S. Anders, P.T. Pyl, W. Huber, HTSeq--a Python framework to work with high-throughput sequencing
732 data, Bioinformatics. 31 (2015) 166–169. doi:10.1093/bioinformatics/btu638.
733 [29] R.C. Gentleman, V.J. Carey, D.M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J.
734 Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A.J.
735 Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J.Y.H. Yang, J. Zhang, Bioconductor: open
736 software development for computational biology and bioinformatics, Genome Biol. 5 (2004) R80.
737 doi:10.1186/gb-2004-5-10-r80.
738 [30] M.D. Robinson, D.J. McCarthy, G.K. Smyth, edgeR: a Bioconductor package for differential expression
739 analysis of digital gene expression data, Bioinformatics. 26 (2010) 139–140.
740 doi:10.1093/bioinformatics/btp616.
741 [31] S. Anders, W. Huber, Differential expression analysis for sequence count data, Genome Biol. 11 (2010)
742 R106. doi:10.1186/gb-2010-11-10-r106.

743 [32] K. Tamura, Y. Fukao, M. Iwamoto, T. Haraguchi, I. Hara-Nishimura, Identification and characterization
744 of nuclear pore complex components in Arabidopsis thaliana, Plant Cell. 22 (2010) 4084–4097.
745 doi:10.1105/tpc.110.079947.
746 [33] K. Tamura, I. Hara-Nishimura, The molecular architecture of the plant nuclear pore complex, J. Exp. Bot.
747 64 (2013) 823–832. doi:10.1093/jxb/ers258.
748 [34] B.M. Fontoura, G. Blobel, M.J. Matunis, A conserved biogenesis pathway for nucleoporins: proteolytic
749 processing of a 186-kilodalton precursor generates Nup98 and the novel nucleoporin, Nup96, J. Cell Biol.
750 144 (1999) 1097–1112.
751 [35] D.J.P. Ferguson, S.F. Parmley, S. Tomavo, Evidence for nuclear localisation of two stage-specific
752 isoenzymes of enolase in Toxoplasma gondii correlates with active parasite replication, Int. J. Parasitol. 32
753 (2002) 1399–1410.
754 [36] E.R. Griffis, S. Xu, M.A. Powers, Nup98 Localizes to Both Nuclear and Cytoplasmic Sides of the Nuclear
755 Pore and Binds to Two Distinct Nucleoporin Subcomplexes, Mol. Biol. Cell. 14 (2003) 600–610.
756 doi:10.1091/mbc.E02-09-0582.
757 [37] J. Enninga, A. Levay, B.M.A. Fontoura, Sec13 shuttles between the nucleus and the cytoplasm and stably
758 interacts with Nup96 at the nuclear pore complex, Mol. Cell. Biol. 23 (2003) 7271–7284.
759 [38] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein
760 modeling, prediction and analysis, Nat. Protocols. 10 (2015) 845–858. doi:10.1038/nprot.2015.053.
761 [39] C.E. Pritchard, M. Fornerod, L.H. Kasper, J.M. van Deursen, RAE1 is a shuttling mRNA export factor
762 that binds to a GLEBS-like NUP98 motif at the nuclear pore complex through multiple domains, J. Cell
763 Biol. 145 (1999) 237–254.
764 [40] J. Guizetti, R.M. Martins, S. Guadagnini, A. Claes, A. Scherf, Nuclear pores and perinuclear expression
765 sites of var and ribosomal DNA genes correspond to physically distinct regions in Plasmodium
766 falciparum, Eukaryotic Cell. 12 (2013) 697–702. doi:10.1128/EC.00023-13.
767 [41] H. Wang, H. Yang, C.S. Shivalila, M.M. Dawlaty, A.W. Cheng, F. Zhang, R. Jaenisch, One-step
768 generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering,
769 Cell. 153 (2013) 910–918. doi:10.1016/j.cell.2013.04.025.
770 [42] J.A. DeGrasse, K.N. DuBois, D. Devos, T.N. Siegel, A. Sali, M.C. Field, M.P. Rout, B.T. Chait, Evidence
771 for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic
772 ancestor, Mol. Cell Proteomics. 8 (2009) 2119–2130. doi:10.1074/mcp.M900038-MCP200.
773 [43] S.O. Obado, M. Brillantes, K. Uryu, W. Zhang, N.E. Ketaren, B.T. Chait, M.C. Field, M.P. Rout,
774 Interactome Mapping Reveals the Evolutionary History of the Nuclear Pore Complex, PLOS Biol. 14
775 (2016) e1002365. doi:10.1371/journal.pbio.1002365.
776 [44] E. Fabre, W.C. Boelens, C. Wimmer, I.W. Mattaj, E.C. Hurt, Nup145p is required for nuclear export of
777 mRNA and binds homopolymeric RNA in vitro via a novel conserved motif, Cell. 78 (1994) 275–289.
778 [45] T.C. Dockendorff, C.V. Heath, A.L. Goldstein, C.A. Snay, C.N. Cole, C-terminal truncations of the yeast
779 nucleoporin Nup145p produce a rapid temperature-conditional mRNA export defect and alterations to
780 nuclear structure, Mol. Cell. Biol. 17 (1997) 906–920.
781 [46] M.A. Powers, D.J. Forbes, J.E. Dahlberg, E. Lund, The vertebrate GLFG nucleoporin, Nup98, is an
782 essential component of multiple RNA export pathways, J. Cell Biol. 136 (1997) 241–250.
783 [47] A. Radu, G. Blobel, M.S. Moore, Identification of a protein complex that is required for nuclear protein
784 import and mediates docking of import substrate to distinct nucleoporins, Proc. Natl. Acad. Sci. U.S.A. 92
785 (1995) 1769–1773.
786 [48] A. Radu, M.S. Moore, G. Blobel, The peptide repeat domain of nucleoporin Nup98 functions as a docking
787 site in transport across the nuclear pore complex, Cell. 81 (1995) 215–222.
788 [49] M.B. Frankel, L.J. Knoll, The ins and outs of nuclear trafficking: unusual aspects in apicomplexan
789 parasites, DNA Cell Biol. 28 (2009) 277–284. doi:10.1089/dna.2009.0853.
790 [50] C.F. Brooks, M.E. Francia, M. Gissot, M.M. Croken, K. Kim, B. Striepen, Toxoplasma gondii sequesters
791 centromeres to a specific nuclear region throughout the cell cycle, Proc. Natl. Acad. Sci. U.S.A. 108
792 (2011) 3767–3772. doi:10.1073/pnas.1006741108.
793 [51] C. Lemaître, W.A. Bickmore, Chromatin at the nuclear periphery and the regulation of genome functions,
794 Histochem. Cell Biol. 144 (2015) 111–122. doi:10.1007/s00418-015-1346-y.
795 [52] Y. Liang, M.W. Hetzer, Functional interactions between nucleoporins and chromatin, Curr. Opin. Cell
796 Biol. 23 (2011) 65–70. doi:10.1016/j.ceb.2010.09.008.
797 [53] C.R. Brown, C.J. Kennedy, V.A. Delmar, D.J. Forbes, P.A. Silver, Global histone acetylation induces
798 functional genomic reorganization at mammalian nuclear pore complexes, Genes Dev. 22 (2008) 627–639.
799 doi:10.1101/gad.1632708.
800 [54] A. Bougdour, D. Maubon, P. Baldacci, P. Ortet, O. Bastien, A. Bouillon, J.-C. Barale, H. Pelloux, R.
801 Ménard, M.-A. Hakimi, Drug inhibition of HDAC3 and epigenetic control of differentiation in
802 Apicomplexa parasites, J. Exp. Med. 206 (2009) 953–966. doi:10.1084/jem.20082826.

803 [55] D. Reinberg, R.J. Sims, de FACTo nucleosome dynamics, J. Biol. Chem. 281 (2006) 23297–23301.
804 doi:10.1074/jbc.R600007200.
805 [56] C. Tous, A.G. Rondón, M. García-Rubio, C. González-Aguilera, R. Luna, A. Aguilera, A novel assay
806 identifies transcript elongation roles for the Nup84 complex and RNA processing factors, EMBO J. 30
807 (2011) 1953–1964. doi:10.1038/emboj.2011.109.
808 [57] M. Costanzo, A. Baryshnikova, J. Bellay, Y. Kim, E.D. Spear, C.S. Sevier, H. Ding, J.L.Y. Koh, K.
809 Toufighi, S. Mostafavi, J. Prinz, R.P. St Onge, B. VanderSluis, T. Makhnevych, F.J. Vizeacoumar, S.
810 Alizadeh, S. Bahr, R.L. Brost, Y. Chen, M. Cokol, R. Deshpande, Z. Li, Z.-Y. Lin, W. Liang, M.
811 Marback, J. Paw, B.-J. San Luis, E. Shuteriqi, A.H.Y. Tong, N. van Dyk, I.M. Wallace, J.A. Whitney,
812 M.T. Weirauch, G. Zhong, H. Zhu, W.A. Houry, M. Brudno, S. Ragibizadeh, B. Papp, C. Pál, F.P. Roth,
813 G. Giaever, C. Nislow, O.G. Troyanskaya, H. Bussey, G.D. Bader, A.-C. Gingras, Q.D. Morris, P.M. Kim,
814 C.A. Kaiser, C.L. Myers, B.J. Andrews, C. Boone, The genetic landscape of a cell, Science. 327 (2010)
815 425–431. doi:10.1126/science.1180823.
816 [58] T. Burckin, R. Nagel, Y. Mandel-Gutfreund, L. Shiue, T.A. Clark, J.-L. Chong, T.-H. Chang, S. Squazzo,
817 G. Hartzog, M. Ares, Exploring functional relationships between components of the gene expression
818 machinery, Nat Struct Mol Biol. 12 (2005) 175–182. doi:10.1038/nsmb891.
819 [59] G.M. Hautbergue, M.-L. Hung, M.J. Walsh, A.P.L. Snijders, C.-T. Chang, R. Jones, C.P. Ponting, M.J.
820 Dickman, S.A. Wilson, UIF, a New mRNA Export Adaptor that Works Together with REF/ALY,
821 Requires FACT for Recruitment to mRNA, Current Biology. 19 (2009) 1918–1924.
822 doi:10.1016/j.cub.2009.09.041.
823

824 **Captions**
825

826 **Fig. 1 Conditional knock-down of TgNup302 and phenotypic analysis.** (a) Localization of the TgNup302
827 protein in the conditional knock-down strain (tagged at its N- and C-termini) by confocal imaging using anti-HA
828 (N-ter) and anti-Myc (C-ter) monoclonal antibodies under normal conditions (top) and after 48-hr ATc treatment
829 (bottom). (b) TgNup302 double-tagged protein expression decreases under ATc treatment in the conditional
830 knock-down strain. Immunoblots are shown for total protein extracts from the wild-type and TgNup302 iKD strain
831 in the presence or absence of ATc for 24, 48, and 72 hrs. Western blots were probed with anti-HA (N-ter) and anti-
832 Myc (C-ter) antibodies. TgENO2 was probed as a loading control. (c) The appropriate level of TgNup302
833 expression is essential for parasite growth. Plaque assays were performed with 200 parasites of wild-type and
834 TgNup302 iKD strains with or without ATc treatment for 7 days. The wild-type strain growth was not affected by
835 ATc treatment (top). In contrast, the growth of the conditional TgNup302 knock-down strain was dramatically
836 impaired in the presence of ATc. (d) Growth assay. Parasites from the parental (blue bars) or iKD TgNup302 iKD
837 (red bars) strains were incubated with (crossed bars) or without ATc (dotted bars) for 48 hrs. The number of
838 parasites per vacuole was scored for a minimum of 100 vacuoles. The average number of parasite is represented
839 in this graph. The results shown are from three independent experiments. *, $P<0.05$; **, $P<0.001$
840

841 **Fig. 2 TgNup302 is involved in nuclear import and is essential for 18S RNA export**. (a) TgNup302 is involved
842 in the nuclear import of the TgENO2 protein. IFA was performed on paraformaldehyde-fixed intracellular parasites
843 of the wild-type and TgNup302 iKD strains using a TgENO2 antibody (α-TgENO2) and staining of the nuclear
844 DNA with DAPI. In the presence of ATc for 48 h, the TgENO2 protein is mislocalized in the cytoplasm of
845 TgNup302 iKD, whereas normal nuclear localization was observed in the absence of ATc for the parental and
846 TgNup302 iKD strains, as expected. (b) TgENO2 localization in intracellular parasites. The percentage of parasites
847 displaying nuclear and cytoplasmic TgENO2 signal was determined for 100 vacuoles in the parental and
848 TgNup302 iKD strains with or without ATc for 48 h after IFA using a α-TgENO2 staining. Predominant nuclear
849 localization of TgENO2 was observed for the iKD strain without ATc and the parental parasites, whereas
850 predominant cytoplasmic localization was observed for ATc-treated iKD parasites. The results shown are from
851 three independent experiments. **, $P<0.05$; ***, $P<0.0001$. (c) TgNup302 is essential to 18S RNA export. RNA
852 FISH was performed on intracellular parasites treated with or without ATc for 48 h. Parasites of the parental and
853 TgNup302 iKD strains were hybridized with Cy3-labeled 18S oligonucleotide primers (red), and the nuclear DNA
854 was labelled with DAPI (blue). (d) 18S RNA localization in extracellular parasites. The quantity of nuclear and
855 cytoplasmic Cy3-labeled 18S signal for 100 parasites was determined for the parental and TgNup302 iKD strains
856 with or without ATc for 48 h. In the parental strain and iKD strains without ATc treatment, most 18S RNA is
857 cytoplasmic, while it is mainly nuclear in the iKD parasites treated with ATc. The results shown are from three
858 independent experiments. ***, $P<0.0001$
859

860 **Fig. 3 TgNup302 depletion alters gene expression.** (a) Total RNA purified from iKD TgNup302 parasites under
861 either control conditions or ATc treatment for 48 hr were analyzed by quantitative RT-PCR (dark bars). Genes

coding for the transcripts identified to be upregulated with ATc treatment in RNA Seq were analyzed. Values are presented as log2 ratios of the signal given by the sample extracted from the ATc-treated parasites relative to those under control conditions (minus ATc). (b) Total RNA purified from iKD TgNup302 parasites under either control conditions or ATc treatment for 48 hr were analyzed by quantitative RT-PCR. The TgNup302 gene was analyzed alongside the TgTubulin gene, which was not affected by ATc treatment. Values are presented as the Log2 ratio of ATc-treated parasites relative to those under control conditions (minus ATc).

Table1 Identification of proteins associated with NUP1 protein. Identification of proteins that were copurified from nuclear extract from the parental and TgNup1 iKD N-ter-HA Tag (in duplicate (1) and (2)) strains using nano LC-MS/MS. The affinity purification was performed under high stringency conditions and protein identity was determined using a composite target decoy database (50620 total entries) built with 3 strains of *Toxoplasma gondii* ToxoDB.org database (strains ME49, GT1 and VEG, release 12.0, September 2014, 25264 entries). Potential TgNup1 interacting proteins were identified and sorted by their secondary structure. FG repeats containing proteins are highlighted in green, alpha-solenoid protein in orange, beta-propeller in blue and coiled-coiled domains in yellow.

Fig. 4 Identified proteins interact with TgNup302. (a) Nuclear extract of ~500x10^6 parasites for the TgFACT140 (line 2), TgNup129 (line 3), TgNup134 (line 4) and TgNup407 (line 5) C-terminally myc tagged proteins in the TgNup302-HA iKD strain were immunoprecipitated using anti-myc antibodies. Extracts from the TgNup302 HA iKD strain were immunoprecipitated using anti-myc antibodies and used as negative controls. The immunoblot was probed with anti-HA antibody (upper panel) to detect interactions between the N-terminal domain of TgNup302 and the five different proteins. The same blots were reprobed with anti-myc antibodies to verify that the myc-tagged proteins were indeed immunoprcecipitated (bottom panel). The TgFACT140, TgNup129, TgNup134, TgNup407 C-terminally myc tagged proteins in the TgNup302-HA iKD strain and the TgNup302 HA iKD strain related inputs (produced from the same experiment) are shown in figure S9. (b) Nuclear extract of ~300x10^6 parasites for the TgNup134 (positive control) and TgChromo1 (negative control) C-terminally myc tagged proteins in the TgNup302-HA iKD strain were immunoprecipitated using anti-Myc antibodies. Extracts from the TgNup302 HA iKD strain were immunoprecipitated using anti-Myc antibodies and used as negative controls. The immunoblot was probed with anti-HA (left panel) antibody to detect the TgNup302-HA protein in the input and eluates. The immunoblot was re-probed with an anti-myc antibody (right panel) to detect the myc-tagged proteins in the input and eluates.

Fig. 5 Identified proteins localize to the *T. gondii* nuclear pore. (a) Each potential partner was tagged using a Myc-tag in the TgNup302-HA iKD strain. Endogenous TgNup302 iKD was labelled with rabbit monoclonal anti-HA (in red), and endogenous TgNup503, TgNup134, and TgNup115 were labelled with the mouse monoclonal anti-Myc (in green) antibody with or without ATc after 48 hr of growth. (b) Measurement of cytoplasmic and nuclear Myc signals were determined for endogenous TgNup115 (vertical striped bars), TgNup134 (dotted bars) and TgNup503 (horizontal striped bars) potential partners without ATc treatment (blue) and after 48-hr ATc treatment (red). The ratio of the cytoplasmic signal versus nuclear signal was determined. The results shown are from three independent experiments. *, $P<0.05$

Fig. 6 TgNup407 localization is perturbed in the TgNup302-deficient parasites. (a) TgNup407 protein was tagged using a Myc-tag in the TgNup302-HA iKD strain. Endogenous TgNup302 iKD was labelled with rabbit monoclonal anti-HA (in red) and endogenous TgNup407 was labelled with mouse monoclonal anti-Myc (in green) antibody with or without ATc after 48 hr of growth. (b) Measurement of cytoplasmic and nuclear Myc signals were determined for endogenous TgNup503 (horizontal striped bars), TgNup407 (empty bars) potential partners without ATc treatment (blue) and after 48-hr ATc treatment (red). The ratio of the cytoplasmic signal versus nuclear signal was determined. The results shown are from three independent experiments. ***, $P<0.0001$

Fig. 7 Codistribution of new nuclear pore proteins and TgNup302. (a) Codistribution between TgNup302-N-ter HA-tag staining (in red) and Myc-tag staining (in green) for TgNup302 (C-terminus), TgNup67, TgNup129, TgNup503, TgNup115 and TgNup134 knock-ins. Values are measured with Imaris software from Structured Illumination Microscopy (SIM) (Pearson's coefficient in ROI volume). ***, $P<0.0001$. (b) SIM images representing the codistribution of TgNup302 (C-ter), TgNup67, TgNup129, TgNup503, TgNup115 and TgNup134 proteins tagged using a Myc-tag (in green) in the TgNup302-HA (in red) iKD strain (highlighted by white arrows). (c) Number of pores per nucleus in TgNup302 (C-ter), TgNup67, TgNup129, TgNup503, TgNup115 and TgNup134 knock-ins. Values are measured with Imaris software from Structured Illumination Microscopy (SIM). ***, $P<0.0001$

921 **Fig. 8 CRISPR-Cas9 screening reveals the importance of the TgNup503 and TgNup134 proteins for the**
922 **nuclear localization of TgENO2.** The TgNup503 and TgNup134 proteins are important for the nuclear
923 localization of the TgENO2 marker. IFA was performed on paraformaldehyde-fixed intracellular parasites of
924 TgNup115, TgNup593, TgNup302, TgNup134, TgNup503, TgNup530, TgFACT140, TgNup67, TgNup37,
925 TgNup68, TgNup216, TGGT1_228100, TgNup407, TgSec13, TgNup129 strains and TgAlba1 as a control using
926 the TgENO2 antibody (α-TgENO2) and staining of the nuclear DNA with DAPI. The percentage of GFP-negative
927 vacuoles with cytoplasmic TgENO2 and the percentage of GFP-positive vacuoles with cytoplasmic TgENO2 was
928 determined. The results shown are from three independent experiments. **, $P<0.05$
929

930 **Supporting information captions**

931

932 **Fig. S1 *T. gondii* TgNup302 is evolutionarily conserved among Eukaryota.** Phylogenetic tree of *T. gondii*
933 TgNup302 homologues based on ClustalW alignment of sequences identified by BLASTp searches using the
934 entire sequence of TgTgNup302 gene against 15 apicomplexan parasites (brown), 5 Fungi (green), 3 Eumetazoa
935 (blue) and 3 Plantae (purple). The tree was reconstructed by maximum likehood (ML) analyses with MEGA6
936 software. Five hundred bootstrap pseudo-replicates were used to give statistical support to the clades of the
937 maximum likehood topology. Scale bar reflects number of substitutions per site. Numbers in the nodes of the tree
938 reflect the percentage of bootstrap replicates supporting each node

939

940 **Fig. S2 Construction of the TgNup302 iKD strain.** (A) Schematic of the genetic approach used to produce the
941 conditional knock-down strain by a promoter replacement strategy. After promoter replacement, the expression of
942 TgNup302 is under the control of anhydrotetracyline (ATc). The *TgNup302* gene was HA-tagged at its 5' (Top
943 panel). (B) PCR was used to confirm the correct integration of the plasmid and the creation of the recombinant
944 locus using primer pairs i and ii (sequences are in TableS1 of supplementary data). (C) The double HA-tagged (N-
945 terminal) / Myc-tagged (C-terminal) strain for the TgNup302 gene was produced by introducing a Myc tag at the
946 3' end of the gene

947

948 **Fig. S3 PolyA+ RNA were mostly cytoplasmic in the parental and iKD strains**. RNA FISH was performed on
949 intracellular parasites. Parasites of the parental and TgNup302 iKD strains were hybridized with Cy3-labeled
950 polyA+ oligonucleotides primers (red), and the nuclear DNA was labelled with DAPI (blue)

951

952 **Fig. S4 The conditional expression of TgNup302 has no impact on TgChromo1 and TgNF3 localization.**
953 Endogenous TgChromo1 and TgNF3 were labelled with the mouse monoclonal anti-Myc and 488nm Alexa goat
954 secondary antibody in parental and TgNup302 iKD strains with or without ATc treatment for 48hr

955

956 **Fig. S5 Ultrastructure of *Toxoplasma gondii*.** Intracellular *T. gondii* tachyzoite showing the nucleus (N) for the
957 parental RHΔKu80 TaTi and iKD TgNup302 strains with or without ATc treatment. Bar=500nm

958

959 **Fig. S6 Chromosomal distribution of genes differentially regulated in the TgNUP1 iKD strain.** The
960 chromosomal position and distribution of genes that were identified as differentially expressed, upregulated (top
961 panel) and downregulated (bottom panel), in TgNup302 iKD after an ATc treatment for 48hr

962

963 **Fig. S7 Peptides identified for TgNup302.** Peptides recovered from the immunoprecipitation of TgNup302 and
964 mapped onto the TgNup302 sequence are highlighted in yellow. The autocatalytic domain sequence is underlined

965

966 **Fig. S8 Predicted secondary structure features, fold and location for validated TgNups.** The horizontal black
967 line represents the polypeptide length of the proteins. The y axis indicates the confidence score of the predicted
968 secondary structure element. Predicted α-helices are indicated in blue, predicted β-sheets in orange, and predicted
969 coiled coil regions are in red arrows. The green arrows indicate FG repeats.

970

971 **Fig. S9 Identified proteins interact with TgNup302.** (a) 5μl of nuclear extract (from ~$500x10^6$ parasites)
972 conserved before the immunoprecipitation (inputs) for the TgFACT140 (line 2), TgNup134 (line 3), TgNup129
973 (line 4) and TgNup407 (line 5) C-terminally myc tagged proteins in the TgNup302-HA iKD strain were analyzed
974 by Western blot. The immunoblot was probed with anti-HA antibody to detect the presence of the TgNup302-HA
975 protein in each strains before immunoprecipitation. (b) The immunoblot was reprobed with an anti-myc antibody
976 to detect the myc tagged proteins: TgFACT140 (134kDa), TgNup129 (129kDa), TgNup134 (134kDa), TgNup407
977 (115kDa).

978

979 **Fig. S10 New components of the *T. gondii* nuclear pore.** Each potential partners were tagged using a Myc-tag
980 in the TgNup302-HA iKD strain. Endogenous TgNup302 iKD was labelled with the rabbit monoclonal anti-HA
981 (in red) and endogenous TgFACT140, TgNup129 were labelled with the mouse monoclonal anti-Myc (in green)
982 antibody with or without ATc after 48hr of growth

983

984 **Fig. S11 Transfection efficiency for the Crisp-Cas9 screening.** For each construction, the percentage of vacuole
985 with a positive GFP expression was monitored to determine the transfection efficiency at 24 hr after
986 electroporation, revealing that ~30 to 70% of cells received the plasmid. TgAlba1 is a negative control
987
988 **Table S1: Oligonucleotides used in this study**
989 **Table S2: RNA-seq results.** Presents all the genes up or down regulated with a FDR of 0.05
990 **Table S3: Mass spectrometry results.** The proteins highlighted in orange met the following criteria: less than 1
991 peptide in the control experiment and present in the two IP experiment
992