Highly diverged novel subunit composition of apicomplexan F-type ATP synthase identified from *Toxoplasma gondii*

Rahul Salunke[1], Tobias Mourier[2], Manidipa Banerjee[3], Arnab Pain[2] and Dhanasekaran Shanmugam[#1]

[1]Biochemical Sciences Division, CSIR-National Chemical Laboratory, Pune, Maharashtra, India.

[2]Pathogen Genomics Laboratory, BESE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia.

[3]Kusuma School of Biological Sciences, Indian Institute of Technology-Delhi, Hauz Khas, New Delhi, India.

[#]Correspondence to be addressed to d.shanmugam@ncl.res.in

Short title: Novel subunits of *T. gondii* F-type ATP synthase

Key words: *Toxoplasma gondii*, Apicomplexa, Mitochondrion, F-type ATP synthase, Oxidative Phosphorylation

Abstract

The mitochondrial F-type ATP synthase, a multi-subunit nanomotor, is critical for maintaining cellular ATP levels. In *Toxoplasma gondii* and other apicomplexan parasites, many subunit components, necessary for proper assembly and functioning of this enzyme, appear to be missing. Here, we report the identification of 20 novel subunits of *T. gondii* F-type ATP synthase from mass spectrometry analysis of partially purified monomeric (~600 kDa) and dimeric (>1 MDa) forms of the enzyme. Despite extreme sequence diversification, key F_O subunits, a, b and d, can be identified from conserved structural features. Orthologs for these proteins are restricted to apicomplexan, chromerid and dinoflagellate species. Interestingly, their absence in ciliates indicates a major diversion, with respect to subunit composition of this enzyme, within the alveolate clade. Discovery of these highly diversified novel components of the apicomplexan F-type ATP synthase complex could facilitate the development of novel anti-parasitic agents. Structural and functional characterization of this unusual enzyme complex will advance our fundamental understanding of energy metabolism in apicomplexan species.

48 **Introduction**

49 The F-type ATP synthase is a ubiquitous nanomotor present on the inner membrane
50 of mitochondria, chloroplast and bacterial plasma membrane that catalyzes ATP formation
51 from ADP and Pi [1,2]. In mitochondria, the energy required to drive the nanomotor is
52 harnessed from the proton motive force and associated membrane potential, which are
53 generated by the mitochondrial respiratory complexes [3,4]. The mitochondrial F-type ATP
54 synthase complex consists of two sub-complexes called the F_1 and F_O sectors. The
55 hydrophilic globular F_1 sector includes three each of α and β subunits, and a central stalk
56 containing one each of γ, δ and ε subunits. The site of ATP formation resides in the catalytic
57 center located in the β subunit [1,2,4,5]. Unlike the F_1 sector subunit components, which are
58 well conserved across various eukaryotic phyla, many F_O sector subunit components are
59 highly diverse [6-9] and are not readily identified based on sequence similarity. The complete
60 set of F_O sector subunits have been mapped only in a few species [10,11]. For example, in the
61 yeast *Saccharomyces cerevisiae*, the F_O sector is composed of the hydrophobic membrane
62 spanning oligomeric subunit c (10-12 monomers), along with subunits a, b, d, f, 8, h and
63 OSCP (Oligomycin sensitivity confirming protein) [10]. In addition, other accessory subunits
64 such as e, g, i, j and k are also associated with the enzyme. Orthologs of all yeast F_O sector
65 subunits, except subunits j and k, are present in the mammalian counterpart (bovine enzyme),
66 which additionally has subunits AGP and MLQ [11].

67 Numerous biochemical and structural studies have addressed the assembly and
68 interaction between the various subunit components of F-type ATP synthases [1,12]. These
69 studies have provided in-depth insights on the functional coupling of the proton motive force
70 to ATP synthesis. The F_O sector subunits c and a interact to form the proton channel, and
71 proton translocation, which occurs along the interface of these two subunits, is coupled to the
72 rotary motion of the c-ring and central stalk of the enzyme [13-16]. The peripheral stalk

73 structure of the enzyme, comprised of F_O sector subunits b, d, h (f6 in bovine) and OSCP, has
74 the critical role of holding the $\alpha_3\beta_3$ catalytic domain in place during the clockwise rotation of
75 the central stalk. These events elicit conformational changes in the $\alpha_3\beta_3$ catalytic domain,
76 which then results in ATP synthesis [2,4,17,18]. F_O sector subunits e and g are known to be
77 involved in dimer assembly of the enzyme complex in yeast [19,20], which appears to
78 facilitate cristae formation by the mitochondrial inner membrane [21-24].

79 However, despite the importance of the various F_O sector subunits, only subunit c and
80 OSCP can be detected from sequence conservation in the vast majority of unicellular
81 eukaryotes, including in the alveolate infrakingdom, which includes the phylum *Apicomplexa*
82 [25-27]. It is unlikely that the missing subunits are totally absent in these species. Plausible
83 scenarios are that they have either diverged in sequence beyond recognition or that a
84 completely new set of proteins have substituted for the missing subunits in these species. This
85 is indeed the case in organisms such as *Euglena gracilis* [9], *Trypanosoma brucei* [7],
86 *Tetrahymena thermophile* [8] and *Chlorophyceae* algae [6]. Hence, our interest was to
87 characterize the complete subunit composition of the F-type ATP synthase from the
88 apicomplexan parasite *Toxoplasma gondii*.

89 *T. gondii* is an obligate intracellular protozoan parasite, responsible for toxoplasmosis
90 in humans and animals [28]. The parasite completes its life cycle in two host species; the
91 definitive hosts are naïve feline species which support the sexual development of the parasite,
92 and all warm blooded animals are intermediate hosts and support its asexual development
93 [29]. During asexual development, the parasites can reversibly differentiate between fast-
94 growing virulent tachyzoites and slow growing latent bradyzoites [30]. By virtue of their fast-
95 growing nature, tachyzoites are metabolically more active and glycolysis is the primary
96 source for carbon and energy for optimal growth of this parasite [31-33]. However, recent
97 findings have revealed that glucose is not an essential nutrient [34], and parasites survive in

98 the presence of glutamine as an alternative nutrient source [35,36]. In the absence of glucose
99 oxidation *via* glycolysis, oxidative phosphorylation is the only source for bulk cellular ATP.
100 Evidence for an active oxidative phosphorylation in *T. gondii* exists [37-39], and treatment
101 with atovaquone, a potent inhibitor of mitochondrial electron transport chain (mtETC),
102 results in inhibition of ATP synthesis and parasite growth (EC_{50} < 10 nM) [38].

103 Very little is known about the structure and function of the F-type ATP synthase from
104 *T. gondii* and other apicomplexan parasites. Comparative genomics has revealed that, while
105 the five canonical F_1 sector subunits - α, β, γ, δ, and ε - are readily identified in all
106 apicomplexan parasites based on amino acid sequence conservation, only two F_O sector
107 subunits - c and OSCP - can be identified by sequence [25-27]. Characterization of F-type
108 ATP synthase from solubilized mitochondria lysates of the human malaria parasite
109 *Plasmodium falciparum* revealed that the enzyme assembles into monomer and dimer forms
110 [26]. This suggests that a full complement of subunits, typical for the eukaryotic enzyme, is
111 present in *P. falciparum* and likely in other apicomplexan parasites as well.

112 Here, we have investigated the subunit composition of the enzyme from *T. gondii.* We
113 were able to identify and partially purify the F-type ATP synthase enzyme complex from
114 solubilized mitochondrial preparation of *T. gondii* using blue native PAGE separation, by
115 immuno-precipitation and by chromatographic enrichment. LC-MS analysis of the enzyme
116 preparations revealed the identity of the proteins associated with the *T. gondii* F-type ATP
117 synthase. We have identified 20 novel proteins (of unknown function) as being *bona fide*
118 subunit constituents of *T. gondii* F-type ATP synthase based on consensus from multiple and
119 independent experiments. Phyletic profiling revealed that orthologs for many of these
120 proteins are present in most apicomplexan species, except in those *Cryptosporidium sp* which
121 possess a degenerate mitochondrion devoid of all proteins involved in oxidative
122 phosphorylation [40]. Further, the phyletic profiles of these novel proteins reveal their origin

123 to be ancient, probably even ancestral to alveolate radiation. The identification of these novel
124 protein components of the apicomplexan F-type ATP synthase will facilitate further
125 structural, functional and inhibitor discovery studies on this important enzyme complex.

126

127 **Results**

128 *Identification of F-type ATP synthase subunits from T. gondii proteome* – Comprehensive *in*
129 *silico* analysis on the presence and absence of key enzymes involved in mitochondrial
130 metabolism in apicomplexan species has been previously carried out [25,27,41]. These
131 studies have highlighted the missing subunits of the F-type ATP synthase enzyme from these
132 organisms. We have further confirmed these findings from our efforts to identify *T. gondii*
133 orthologs for yeast and bovine F-type ATP synthase subunits. Orthologs for F_1 sector
134 subunits α (TGME49_204400), β (TGME49_261950), δ (TGME49_226000),
135 ε (TGME49_314820), and γ (TGME49_231910) and F_O sector subunits c
136 (TGME49_249720) and OSCP (TGME49_284540) were readily identified from *T. gondii*.
137 This suggested that all the subunits necessary for assembling the catalytic core (α3β3),
138 central stalk (δ, ε and γ) and the subunit c oligomer portions of the F-type ATP synthase (Fig
139 1A) are encoded by *T. gondii*. However, as reported previously [25,27], we were unable to
140 identify the orthologs for critical F_O sector subunits involved in proton translocation and
141 formation of the peripheral stalk structure of the enzyme in *T. gondii* through routine
142 bioinformatic analysis. This appears to be the case in all apicomplexan species and in other
143 taxa within the alveolate infrakingdom, such as the dinoflagellates, ciliates and chromerids
144 (Fig 1B). It is likely that the missing F_O subunits are either highly divergent or completely
145 novel. Interestingly, this situation is not exclusive to alveolates, since many other unicellular
146 eukaryotes also appear not to possess the corresponding orthologs for many of the yeast or

147 bovine F_O subunits. In fact, novel F-type ATP synthase associated proteins have been
148 previously identified from *Tetrahymena* [8], *Chlamydomonas* [6] and *Trypanosoma* [7].
149 Therefore, we attempted to purify the F-type ATP synthase enzyme from isolated *T. gondii*
150 mitochondria in native form and identify its protein composition by mass spectrometry.

151 *Generating transgenic parasites expressing YFP-HA tagged TgATP–β and TgATP-OSCP*
152 *proteins* – To facilitate the purification of the *T. gondii* F-type ATP synthase enzyme in
153 native form, we engineered in-frame genomic YFP-HA tags in the 3' end of the respective
154 genes encoding the F_1 β and F_O OSCP subunits (Fig 2A). The modified genes continued to be
155 expressed under the control of their endogenous promoters. The correct insertion of the tags
156 was confirmed by genomic PCRs (S1 Data) from clonal isolates of the respective transgenic
157 parasites. Expression of the desired YFP-HA tagged proteins, and their mitochondrial
158 localization, were confirmed by Western blotting (Fig 2B) and microscopy (Fig 2C),
159 respectively. We then compared the ability of the parental and transgenic strains of the
160 parasites to maintain total cellular ATP levels in the presence/absence of glucose. Although
161 *T. gondii* tachyzoites are known to prefer glucose as the primary nutrient source, in the
162 absence of glucose they are known to switch to glutaminolysis for carbon and energy supply.
163 In the later case, ATP is obtained by the parasites *via* oxidative phosphorylation, and this can
164 be inhibited using atovaquone [25,42,43]. Like the parental parasites, transgenic parasites
165 expressing YFP-HA tagged F_1 β and F_O OSCP subunits were capable of maintaining cellular
166 ATP levels *via* oxidative phosphorylation in the absence of glucose, which was inhibited by
167 atovaquone (Fig 2D). These results confirm that the modification of either F_1 β or F_O OSCP
168 proteins with YFP-HA tag had no detrimental effect on the function of the enzyme,
169 implicating that the structure of the tagged enzyme remained intact.

170 *Identification of monomer and dimer forms of T. gondii F-type ATP synthase by Blue Native*
171 *Page (BNP) analysis* – The fully assembled F-type ATP synthase is known to exist as dimers
172 on the inner mitochondrial membrane and this dimerization is known to influence the
173 characteristic cristae formation by the membrane [21,24]. In case of the human malaria
174 parasite *P. falciparum* F-type ATP synthase, dimer formation has been observed previously
175 [26]. To find out whether the *T. gondii* enzyme can assemble into dimers, we carried out BNP
176 analysis of detergent solubilized mitochondrial preparations. After resolving the samples on a
177 3% - 12% gradient BNP gel capable of resolving a molecular weight range between 20 kDa
178 to ~1200 kDa (Fig 3A lane M), Coomassie staining revealed prominent bands at the sizes
179 corresponding to dimeric and monomeric forms of the enzyme (Fig 3A lane A). This was
180 further confirmed by Western blotting after BNP separation using α-HA antibodies (Fig 3A
181 lane B), which indicated that the dimer form of the enzyme is more abundant than the
182 monomer form.

183 Next, we carried out in-gel activity assays, to detect the ATPase activity associated with the
184 isolated enzyme. Results from these assays confirmed that the ATPase activity was intact, for
185 both dimer and monomer forms of the enzyme, after BNP separation (Fig 3B). We found
186 that, although the dimer form of the enzyme was more abundant, its ATPase activity was less
187 than that of the monomer form. This is supported by the fact that F_1 ATPase activity is
188 inhibited *in vivo* to minimize the risk of ATP hydrolysis and favor ATP synthesis [44]. Since
189 the in-gel activity assay is an ATP hydrolysis assay, the dimer has less of this activity. Based
190 on BNP mobility, we deduce the size of the dimer form of the enzyme to be between 1 – 1.2
191 MDa and that of the monomer form as ~600 kDa. This is in agreement with what has been
192 reported previously for *P. falciparum* [26], *S. cerevisiae* [19], and bovine [45] enzymes, and
193 suggests that a full complement of F_O subunits is present in *T. gondii* F-type ATP synthase.
194 Therefore, we excised the regions in the BNP corresponding to the dimer and monomer

195 forms of the enzyme and proceeded to identify the proteins in the gel band by LC-MS
196 analysis (Fig 3A).

197 *Identification of T. gondii F-type ATP synthase associated proteins by LC-MS/MS proteomics*
198 – A standard protocol (see methods section) was followed for in-gel LC MS/MS analysis of
199 the gel bands corresponding to the dimer and monomer forms of the enzyme identified by
200 BNP. This was done on samples prepared from both *Tgatp–β–YFP-HA* and *Tgatp-*
201 *OSCP–YFP-HA* transgenic parasites. Since the dimer form of the enzyme was more abundant
202 than the monomer form after BNP separation, we had better success with LC-MS/MS
203 analysis of the gel band corresponding to the dimer form of the enzyme. A total of 96
204 proteins were identified with high confidence from consensus data obtained from multiple in-
205 gel LC MS/MS analysis (Fig 5). Proteins corresponding to all F_1 subunits and F_O OSCP were
206 detected, while F_O subunit c was not detected in any of the experiments. In addition, many
207 proteins of unknown function were also detected. However, due to co-migration of several
208 non-specific proteins, such as myosin, in the area corresponding to the gel band processed for
209 LC MS/MS analysis, it is likely that some of these proteins of unknown function are not
210 *bona-fide* subunits of the F-type ATP synthase

211 In order to specifically identify the protein subunits of F-type ATP synthase, we resorted to
212 partially purify the enzyme before LC MS/MS analysis by chromatographic separation (Fig
213 4A, B) and immunoprecipitation (IP; using the α-HA antibody). In contrast to BNP analysis,
214 after size exclusion chromatography we observed that the dimer form of the enzyme was less
215 abundant than the monomer form (Fig 4B). This is likely due to stability issues with the
216 dimer, which might progressively fall apart as monomers during the process of
217 chromatographic separation. The fractions corresponding to the dimer and monomer forms of
218 the enzyme were pooled and concentrated separately before processing for LC-MS/MS
219 analysis. A total of 64 proteins were detected (Fig 5A) from the fractions corresponding to

220 monomer form of the enzyme and we did not get reliable data from the dimer fractions owing
221 to very low protein concentration.

222 LC-MS/MS analysis on the enzyme enriched by IP resulted in the identification of a total of
223 29 proteins, of which 28 were also detected from either BNP or chromatography samples (Fig
224 5A). It is notable that 19 proteins, including most F_1 subunits and F_O OSCP, were detected in
225 samples prepared by all three methods. We detected very few non-specific proteins from the
226 IP samples, i.e., functionally not related to F-type ATP synthase (3 out of 29). Out of the
227 remaining 26 proteins, 5 were known F_1 subunits (α, β, δ, ε, γ), one was F_O OSCP, and the
228 remaining 20 were proteins of unknown function. It should be noted that we were unable to
229 identify F_O subunit c in any of our LC-MS/MS analysis, probably owing to its highly
230 hydrophobic nature. Based on the consensus of proteins identified by the three different
231 approaches, the F-type ATP synthase from *T. gondii* appears to be comprised of at least 27
232 protein subunits. We have coined the termed 'ATP synthase Associated Proteins' (ASAP) to
233 refer to the 20 novel subunit components of the enzyme identified in this study. A complete
234 list of all *T. gondii* F-type ATP synthase subunits identified by mass spectrometry, along with
235 the experiments from which they were detected, is given in the table in Fig 5B.

236 *Identification of F_O subunits a, b and d* – As shown in Fig 1, in addition to subunits c and
237 OSCP, the F_O sector consists of at least 6 other subunits in yeast, mammalian and plant
238 enzymes. Subunit a is essential for proton conductance, which it facilitates along with subunit
239 c [13-15]. Subunit b forms the core of the stator structure, which is essential for holding the
240 catalytic $\alpha_3\beta_3$ structure in place during the rotary motion of the central stalk [18]. Subunit d is
241 part of the stator structure. We attempted to find out which of the ASAPs correspond to these
242 three subunits in *T. gondii* F-type ATP synthase. Since none of the ASAPs had any sequence
243 identity to the yeast F_O a, b and d subunits, we resorted to conserved structure based
244 identification using pairwise comparison of profile hidden Markov models as implemented in

245 HHPred [46,47]. For this analysis we used the corresponding tool available from web based
246 MPI bioinformatics toolkit (*toolkit.tuebingen.mpg.de*). We analyzed each of the ASAPs using
247 this tool and successfully identified hits for F_O subunits a, b and d based on previously known
248 structures for these proteins from other species. The ASAP-1 was identified as F_O subunits a
249 (Fig 6A), and we were also able to generate sequence alignments for the C-terminal domain
250 of this protein which showed the conservation of key amino acid residues arginine (important
251 for proton translocation) and glutamine (yellow highlight in Fig 6B). Similarly, searches with
252 ASAP-2 and ASAP-3 came up with hits for F_O subunits b and d, respectively (S2 Data and
253 S3 Data).

254 *Mitochondrial localization of ASAPs* – Confirming the mitochondrial localization for the
255 ASAPs is essential for validating their subunit membership in *T. gondii* F-type ATP synthase.
256 We first predicted the presence of signal sequence for mitochondrial targeting using the
257 Mitoprot tool [48] in twelve ASAPs (Fig 7). Five ASAPs, including ASAP-2 (F_O subunit b
258 like protein), were previously shown to localize in the mitochondrion [49], and we
259 experimentally confirmed the mitochondrial localization for another four (S4 Data).
260 Moreover, unpublished results from Hyper-LOPIT analysis [50] of *T. gondii* tachyzoite
261 proteome in an independent study (personal communication, Dr. Ross Waller, Cambridge
262 University, UK), confirmed that all ASAP proteins (except ASAP-1 and TGME49_249720,
263 for which no data were available) are localized to the mitochondria (Fig 7).

264 *Phylogenetic analysis reveals conserved F-type ATP synthase subunit composition in three*
265 *major alveolate taxons* – From our ortholog identification protocol, and from previous studies
266 [27,41], it was apparent that the F-type ATP synthase subunits missing from *T. gondii* were
267 also missing from all other species grouped within the *Alveolata* infrakingdom. In fact, the
268 novel subunit components identified from *Tetrahymena* (*Alveolata*; *Ciliophora*) enzyme were
269 found to be unique to ciliates and not conserved in other alveolate organisms [8]. Therefore,

270 we were interested in finding out whether the novel ASAPs identified in this study are unique
271 to *T. gondii* F-type ATP synthase. Interestingly, we were able to identify orthologs for many
272 of the ASAPs in three major alveolate taxons – *Apicomplexa*, *Chromerida* and *Dinoflagellata*
273 (Fig 7). Out of the 20 ASAPs, 15 were conserved in all apicomplexan clades, except in case
274 of *Cryptosporidium*, where only *C. muris* contained orthologs for 10 of these proteins. A few
275 ASAPs were unique to the *Coccidian* clade. More importantly, all ASAPs, except one, were
276 conserved in *Chromerida*, and at least 9 and 12 ASAPs were also conserved in *Symbiodinium*
277 and *Perkinsus* respectively.

278 To obtain further insights on the evolutionary origin of these proteins, we generated
279 neighbor-joining phylogenetic trees for ortholog sequences of all F_1 and F_O subunits from
280 representative species of *Haemosporida*, *Piroplasmida*, *Coccidia*, *Cryptosporidiidae*,
281 *Chromerida, Dinoflagellata* and *Ciliophora* (Fig 7 and S5 Data). *Stramenopile* (*Ectocarpus*
282 *siliculosis*, *Phaeodactylum tricornumtum*, *Pythium ultimum*, *Thalassiosira pseudonana*) and
283 *Plantae* (*Chlamydomonas reinhardtii* and *Arabidopsis thaliana*) species were included as
284 outgroups while constructing the phylogenetic trees for the highly conserved F_1 subunits.
285 Even though the topology of most trees did not reflect the expected evolutionary relationship
286 between the included species, monophyletic grouping was observed in general at the taxon
287 level, and importantly, this was evident for the conserved F_1 / F_O subunits, as well as the
288 novel ASAPs (S5 Data). Thus, the evolutionary origin of the newly identified highly
289 divergent ASAPs in *Apicomplexa*, *Chromerida* and *Dinoflagellata* clades appears to be
290 ancient.

291 *Essentiality and gene co-expression analysis of ASAPs from T. gondii and P. falciparum* –
292 Due to the functional importance of the F-type ATP synthase enzyme, it is reasonable to
293 expect that the enzyme would be essential in *T. gondii*. This is indeed the case, and all known
294 subunits and ASAPs (except ASAP-19 and 20; Fig 7) of the enzyme were found to be

295 essential in a previous study [49]. In order to obtain further independent evidence for ASAPs
296 as *bona fide* subunits of F-type ATP synthase, we carried out transcriptome co-expression
297 correlation analysis using publicly available gene expression datasets for *T. gondii* and *P.*
298 *falciparum* [51,52]. We assumed that the ASAPs interacting with one other would show
299 significant pair-wise correlation profiles in their expression levels. Strikingly, we found very
300 high correlation of gene co-expression profiles among the F-type ATP synthase subunits, in
301 comparison to co-expression with other unrelated gene pairs, in both *T. gondii* and *P.*
302 *falciparum* transcriptome datasets (Fig 8). This finding further supports the fact that the novel
303 ASAPs are indeed co-expressed together and are likely *bona fide* subunits of the unusual F-
304 type ATP synthase enzyme from *T. gondii*.

305

306 **Discussion**

307 Mitochondrial oxidative phosphorylation is an important source of ATP in most
308 eukaryotic organisms, and is facilitated by the multimeric F-type ATP synthase enzyme
309 complex [1,2,5]. The enzyme consists of two functionally distinct F_1 and F_O sectors, which
310 act in concert to convert the electrochemical energy into mechanical energy to facilitate ATP
311 synthesis. F_O sector subunits a and c come together to form the proton translocation channel
312 across the inner mitochondrial membrane. Subunit c, which forms a membrane embedded
313 oligomeric ring, converts the proton gradient energy into mechanical energy *via* subunit
314 rotation. The F_1 sector $\alpha_3\beta_3$ catalytic core is connected to the rotary motor formed by F_O
315 subunit c ring *via* the rotating central stalk structure comprised of the γ, δ and ε F_1 subunits
316 [1,2,5]. The stator structure, comprised primarily of the F_O subunits b, d, h and OSCP, along
317 with other accessory subunits helps in holding the $\alpha_3\beta_3$ catalytic core in place while the
318 central stalk rotates [17]. The asymmetric structure of the γ subunit of the rotating central

319 stalk induces conformational changes in the $\alpha_3\beta_3$ catalytic core, which facilitates ATP
320 synthesis. Thus all subunit components and structural features of this unique enzyme are
321 critical for efficient ATP synthesis.

322 In addition to the core subunit composition described, other accessory proteins
323 participate in assembling the V shaped dimer form of the enzyme complex [19]. The paired
324 arrangement of the F_1 complex (spaced by 12 nm) was observed in *Paramecium* inner
325 mitochondrial membrane by freeze-fracture and deep-etching studies [53]. Subsequent
326 studies on yeast, bovine and other species confirmed the dimeric arrangement of the F-type
327 ATP synthase complex by BNP, atomic force microscopy and cryo-electron microscopy
328 [19,54,55]. In fact, the dimer form of the enzyme appears to facilitate cristae formation by the
329 inner mitochondrial membrane [21,22,24] and is implicated in maintenance of the
330 mitochondrial membrane potential [56]. However, the identity of the proteins responsible for
331 dimer formation has not been ascertained in most species. In case of the yeast enzyme,
332 subunits e, g, k were identified as dimer specific components, and studies with genetic
333 mutants revealed that subunits e and g are essential in dimer formation [19,20]. Orthologs for
334 these proteins are present in mammalian species as well.

335 BNP and size exclusion chromatography studies on purified yeast F-type ATP
336 synthase revealed that the average molecular size of the dimer and monomer form of the
337 enzyme is around 1 MDa and between 500 – 600 kDa, respectively [19]. Given the complex
338 structure and assembly of the enzyme, the expected number of protein subunits in an intact
339 dimer is around twenty, based on yeast and mammalian enzyme compositions [10,11].
340 Interestingly, while the F_1 subunits are well conserved across various taxa, the F_O subunits,
341 except c and OSCP, appear to be poorly conserved and in many instances are not readily
342 identified from sequence. This is especially true for a variety of unicellular eukaryotes,
343 including many free living and parasitic protists, as evident from KEGG data (*kegg.jp*).

344 Comparative genomics studies revealed that orthologs for yeast and mammalian F_O subunits,
345 except c and OSCP, are missing in the entire apicomplexan phylum [25-27]. In fact, along
346 with *Apicomplexa*, this appears to be the case for all other taxa belonging to the alveolata
347 infrakingdom. A study on the subunit composition for F-type ATP synthase from the ciliate
348 *Tetrahymena thermophila*, an alveolate species, resulted in the identification of 13 novel
349 proteins [8]. Surprisingly, these proteins were unique to ciliates, indicating the possible
350 existence of unique set of F_O components in other alveolate species as well.

351 The *P. falciparum* enzyme was identified in dimer form with a molecular size of > 1
352 MDa, confirming the presence of novel F_O subunits [26]. This study also revealed that the
353 enzyme is essential in *P. falciparum*. However, in case of *P. berghei*, F_1 subunit β was found
354 to be essential for development in mosquito, but not for survival of blood stage parasites in
355 mammalian host [57]. Studies with metabolic mutants in *T. gondii* have revealed that
356 mitochondrial oxidative phosphorylation is an important source of ATP, especially during
357 glucose deprivation and in the absence of a glycolytic flux [34,35]. Mitochondrial ATP
358 synthesis is also likely essential for formation and maintenance of tissue cyst forms of *T.*
359 *gondii* in infected hosts. Despite the importance of F-type ATP synthase in apicomplexan
360 parasites, not much is known about its subunit composition, structure and function in these
361 parasites.

362 In this study, we have successfully identified the subunit constituents of F-type ATP
363 synthase from *T. gondii*, a model apicomplexan parasite. Intact monomer and dimer forms of
364 the enzyme can be identified by BNP analysis from detergent solubilized parasite
365 mitochondria preparations. We have identified 20 novel proteins (ASAPs) as subunit
366 constituents of the *T. gondii* F-type ATP synthase by LC-MS/MS analysis of partially
367 purified enzyme using BNP, IP and chromatographic techniques. While some of these
368 proteins could be counterparts of the missing F_O subunits, others are most likely accessory

369 proteins which may be involved in dimer formation. Importantly, we were able to identify
370 putative F_O subunits a (ASAP-1), b (ASAP-2) and d (ASAP-3), based on conserved structural
371 features. Although the F_O subunit a from yeast is known to have six transmembrane domains,
372 the *T. gondii* protein has only three. Nevertheless, the critical arginine residue, which is
373 important for proton translocation [14], appears to be conserved (Fig 6). The hallmark
374 feature of F_O subunit b from yeast and other species is the presence of an extended helix that
375 spans the distance between the membrane and the $\alpha_3\beta_3$ catalytic core. Although there was
376 very poor sequence conservation, very high structural similarity (>97% probability in
377 HHPred assignment) can be found between yeast and *T. gondii* F_O subunit b proteins. F_O
378 subunit d has no transmembrane domain and is known to interact with F_O subunit b *via*
379 parallel/anti-parallel coiled coil helical domains [2]. The putative *T. gondii* F_O subunit d was
380 highly similar in secondary structure features over the conserved regions with the bovine
381 enzyme. Further follow up work on the 3D structure of the enzyme is needed to verify these
382 findings.

383 Interestingly, all F_1 subunits, and F_O c and OSCP subunits, along with 18 ASAPs,
384 were found to be essential in *T. gondii*, in a genome wide CRISPR/Cas9 mediated gene
385 knockout screen [49] (Fig 7). In the same study, some ASAPs were shown to be localized in
386 the mitochondrion. We interrogated the mitochondrial localization for all subunits of *T.*
387 *gondii* F-type ATP synthase by first considering evidence from *in silico* prediction of
388 mitochondrial localization signals, followed by experimental localization of selected ASAPs
389 (Fig 7). We also obtained confirmation for mitochondrial localization of all but two ASAPs
390 from Hyper-LOPIT analysis of tachyzoite stage *T. gondii* (unpublished data from an
391 independent study, Dr. Ross Waller, University of Cambridge, UK; personal
392 communication). Further, we found that the transcript co-expression for the subunits of F-
393 type ATP synthase was highly correlated, independently in both *T. gondii* and *P. falciparum*

394 transcriptome datasets (Fig 8), thus providing additional evidence for functional interaction
395 between these proteins across apicomplexan parasite taxa. In summary, evidence from gene
396 essentiality, subcellular localization and transcript co-expression together provide
397 independent levels of support for our identification of ASAPs as *bona fide* subunits of *T.*
398 *gondii* F-type ATP synthase.

399 As can be expected for such an important protein complex playing a fundamental
400 biological role, phylogenetic analysis revealed that most ASAPs identified from *T. gondii* are
401 conserved across the phylum *Apicomplexa*. This suggests that the origin of a divergent F-type
402 ATP synthase may predate the origin of apicomplexan parasites. In support of this
403 conclusion, we find that many of the ASAPs are also conserved in two other alveolate
404 phylums, *Chromerida* and *Dinoflagellata*. This opens up the opportunity for studying the
405 structure and function of this unusual enzyme from free-living alveolate species, from which
406 the enzyme can be isolated in native conditions at higher yields and purity. Moreover, a
407 deeper understanding of the structure and function of this enzyme will facilitate the discovery
408 of anti-parasitic compounds with pan-apicomplexan effect.

409

410 **Methods**

411 *Sequence analysis and ortholog identification* -

412 *Molecular reagents and methods* – Genomic DNA and total RNA from tachyzoite stage *T.*
413 *gondii*, and plasmid DNA from recombinant *E. coli* (DH5α) were isolated using Qiagen kits
414 (Germany); cDNA was prepared from total RNA using the reverse transcription kit from
415 Thermo Fisher Scientific (USA). The manufacturer's protocol was followed while using the
416 various kits. Polymerase chain reactions (PCR) was done using a proof reading DNA
417 polymerase obtained from Takara Bio (Japan), and all primers used in this study (S1 Table)

418 were obtained from Integrated DNA Technologies (USA). Ligase and various restriction
419 endonucleases were purchased from New England BioLabs (USA) or Promega (USA).
420
421 *Plasmid constructs* – Various PCR products were cloned into the Topo 2.1 plasmid backbone
422 (Thermo Fisher Scientific, USA) and sequenced before sub-cloning into other plasmids used
423 for either modification of target genomic loci or ectopic expression of full-length cDNA for
424 gene of interest in *T. gondii*. The plasmid construct used for tagging the 3' end of F-type ATP
425 synthase subunits β (*Tg*atpβ) and OSCP (*Tg*atpOSCP) with yellow fluorescent protein plus
426 hemagglutinin (YFP-HA) tag had the following features: 2606 bp (BglII and AvrII) and 1831
427 bp (BamHI and AvrII) PCR fragments amplified from the 3' end of *Tg*atpβ and *Tg*atpOSCP
428 loci respectively, excluding the stop codon, were cloned into the Topo 2.1 plasmid, upstream
429 of a YFP-HA tag coding sequence, followed by the 3' UTR of the *T. gondii* dihydrofolate
430 reductase-thymidilate synthase (*Tg*dhfr-ts) gene [58]. A modified pBluescript plasmid
431 backbone was used for ectopic expression of full-length cDNA of selected *T. gondii* genes
432 encoding the novel F-type ATP synthase subunits identified from this work. The cDNAs
433 were cloned as BamHI and XbaI/NheI fragments in-frame to a HA tag, for constitutive
434 expression using the *T. gondii* β-tubulin gene promoter and *Tg*dhfr-ts gene 3' UTR. For
435 isolating clonal lines of transgenic parasites, the respective plasmid constructs include either
436 the DHFR cassette (for endogenous tagging; expresses a mutant version of the *Tg*dhfr cDNA
437 that confers pyrimethamine resistance [59] or the CAT cassette (for ectopic expression;
438 expresses the chloramphenicol acetyl transferase gene, which confers resistance to
439 chloramphenicol [60].

440 *T. gondii culture and genetic manipulation* – Tachyzoite stage of *T. gondii* parasites (Type-I
441 RH (RH) and Type-I RH ΔKu80 [61] (RHΔKu80)) strains were propagated following
442 published protocols [62]. Briefly, Human foreskin fibroblast (HFF), which were used as host

443 cells for parasite infection were grown in Dulbecco's modified Eagle medium (DMEM High
444 glucose) supplemented with 10% heat-inactivated fetal bovine serum, 2 mM GlutaMAX, 25
445 mM HEPES and 50 µg/ml Gentamicin. All cell culture grade reagents were procured from
446 Thermo Fisher Scientific, USA. The HFF cells were maintained at 37°C in a humidified
447 atmosphere containing 5% CO_2. *T. gondii* tachyzoites were propagated by inoculating ~10^5
448 freshly isolated tachyzoites into T25 flasks containing two-weeks old confluent HFF cell
449 monolayer. The culture medium used for parasite growth is similar to that used for host cell
450 growth except that it lacks serum. Parasites are harvested 48 h after infection by scrapping the
451 host cell monolayer, physically disrupting the HFF cell suspension by passing it repeatedly
452 through a 22-gauge needle, and filtering it through a 3 µM nucleopore membrane (Whatman,
453 GE Healthcare ,USA (to obtain a suspension of parasites devoid of HFF cell debris .The
454 number of parasites present in 1 ml of suspension is estimated from cell counts obtained
455 using a hemocytometer.
456
457 For generating transgenic parasites expressing the *Tg*ATPβ and *Tg*ATPOSCP proteins from
458 the endogenous loci, the RHΔKu80 parasites were transfected [61] with the respective
459 tagging plasmids. Transfection was achieved by electroporating 10^7 freshly harvested
460 tachyzoite stage parasites resuspended in 400 µl parasite culture medium containing 50 µg of
461 linearized sterile plasmid DNA, using a BioRad (USA) Gene Pulser system using 10µF
462 capacitance, ∞ ohms resistance and 1.5 kV voltage settings. Transfected parasites were
463 immediately inoculated into a T25 flask containing HFF monolayer and allowed to invade
464 and replicated for 12-15 hours before beginning drug selection with 1 µM pyrimethamine.
465 For ectopic expression of cDNA with HA tags for selected genes, RH parasites were
466 transfected with the respective plasmid constructs and stable transformants were selected

467 using 20 μM chloramphenicol. Clonal lines of stable transgenic parasites were isolated using
468 the limiting dilution technique [62].
469
470 *Measurement of intracellular ATP* - For this assay, freshly isolated extracellular tachyzoite
471 stage parasites were incubated in culture media which either contained 5.5 mM glucose and 4
472 mM glutamine or 0 mM glucose and 4 mM glutamine, in order to evaluate the ability of the
473 parasites to maintain ATP homeostasis. The anti-parasitic drug atovaquone [43], an inhibitor
474 of mtETC, was used to specifically inhibit mitochondrial ATP synthesis. After 2 h of
475 incubation, total cellular ATP was measured using the ViaLight Plus Cell Proliferation and
476 Cytotoxicity Bioassay Kit (Lonza, Switzerland) as per the manufacture protocol. The
477 luminescence readout was quantified using the Varioskan Flash plate reader (Thermo Fisher
478 Scientific, USA).
479
480 *Microscopy and Western blotting* – For visualizing the expression and subcellular
481 localization of selected ASAPs fused to a reporter tag, replicating intracellular tachyzoite
482 stage parasites stably expressing these proteins were imaged by microscopy. Before imaging
483 the parasite infected HFF monolayers cultured on glass coverslips were washed with 1X
484 phosphate buffered saline (PBS), and fixed using 4% paraformaldehyde for 20 min, before
485 again washing and mounting on glass slides using the fluoroshield reagent (Sigma, USA). To
486 counterstain the parasite mitochondrion, the cells were treated with 250 nM Mitotracker Red
487 (Thermo Fisher Scientific, USA) for 20 minutes before fixing. The slides were imaged using
488 the 63X oil immersion objective fitted to the Axio Observer inverted fluorescent microscope
489 (Carl Zeiss, Germany), and images were processed using the Zen software (Carl Zeiss,
490 Germany). Intrinsic fluorescence from YFP and Mitotracker Red were visualized using the
491 excitation/emission filter combination of 493/520 and 578/599 respectively. HA tagged

492 proteins were visualized by immunofluorescence staining, using rabbit α-HA primary
493 antibodies (1:1000) followed by Alexa 488 conjugated goat anti rabbit secondary antibody
494 (1:1000), both purchased from Thermo Fisher Scientific, USA. After fixing, the cells were
495 permeabilized for 5-10 minutes using 0.25% Triton X-100 in 1X PBS. They were then treated
496 with 2% fetal bovine serum in 1X PBS for 30 minutes, followed by primary antibody for 1 h,
497 3X wash with 1X PBS containing 0.25% Triton X-100, and secondary antibody for 1 h.
498 Finally, the cells were washed 3X with 1X PBS containing 0.25% Triton X-100, followed by
499 a water wash, before mounting the cover slip on glass slides for imagining.
500
501 For Western blotting, purified mitochondria were solubilized in 1X Laemmli buffer (120 mM
502 Tris-HCL pH 6.8, 2% SDS, 10% glycerol and 0.01% w/v bromophenol blue) and resolved by
503 SDS-PAGE, followed by electro-transfer to PVDF membrane (using Tris-glycine buffer pH 8
504 containing 20% methanol) at 65 mA for ~1 h 20 min. The blots were kept in blocking buffer
505 (5% skimmed milk in 1X PBS) overnight at 4°C and then treated with primary antibody
506 (1:5000; rabbit α-HA monoclonal antibody, Thermo Fisher Scientific, USA) for 1 h at
507 ambient temperature. This was followed by 3X wash with 1X PBS containing 0.1% Tween
508 20, before incubating with horse radish peroxidase coupled secondary antibody (1:5000;
509 donkey anti-rabbit antibody, Nif 824 GE Healthcare, USA) for 1 h at ambient temperature,
510 and 3X wash with 1X PBS containing 0.1% Tween 20. The blot was then developed using
511 either the 3,3'-diaminobenzidine substrate (Sigma, USA) or the chemiluminescent ECL
512 Western blotting kit (GE Healthcare, USA).
513
514 *Preparation of mitochondria from T. gondii* – Freshly harvested tachyzoite stage transgenic
515 parasites (~10^9), expressing either *Tg*ATPβ-YFP-HA or *Tg*ATPOSCP-YFP-HAHA, were
516 washed with PBS and resuspended in 1-2 ml of hypotonic lysis buffer (15 mM phosphate

517 buffer pH 7.5 and 2 mM Glucose). The cells were lysed by sonication in an iced water bath
518 for 30 minutes, and samples were centrifuged at 2000 g for 2 minutes at 4°C to remove
519 unbroken cells and large cell debris. The supernatant was recovered into a new tube and
520 centrifuged at 21,000 g for 15 minutes at 4°C to pellet the mitochondria, which was then re-
521 suspended in 500 µl of mitochondria storage buffer (320 mM sucrose, 1 mM EDTA, 10 mM
522 Tris pH 7.4). Total protein in the mitochondria suspension was estimated by Bradford method
523 [63] and the lysate was stored in -80°C until further use. When required, the mitochondria
524 were recovered from storage buffer by centrifuging at 21,000 g for 15 minutes at 4°C and
525 lysed using mitochondria solubilization buffer A (MSB-A; 50 mM NaCl, 50 mM Imidazole,
526 2 mM 6-Aminohexoanic acid, 1 mM EDTA, pH 7.0) containing β-dodecyl maltoside (DDM)
527 detergent in total protein to detergent ratio of 1:5. Solubilization was allowed to continue
528 overnight with rocking at 4°C, and the solubilized fraction was separated by centrifuging at
529 100,000 g for 20 minutes at 4°C.
530
531 *Blue Native Page separation, in-gel ATPase activity assay and Western blotting* – In order to
532 identify the monomeric and dimeric forms of the *T. gondii* F-type ATP synthase, the
533 solubilized mitochondrial lysate was subject to one-dimensional Blue Native PAGE (BNP)
534 analysis [64,65]. Samples were prepared by adding glycerol to a final concentration of 5%
535 and Coomassie Blue G 250 5% (w/v) dye was added such that DDM detergent to dye ratio
536 was 8:1. Samples were separated on a 3-12 % native PAGE (Thermo Fisher Scientific, USA)
537 at 150 V setting for 30 minutes in cathode buffer A, which was then switched to cathode
538 buffer B and separation was continued till the blue dye exited the gel. The gel was visualized
539 under white light to identify the various dye stained bands. 25 mM imidazole pH 7.0 was
540 used as the anode buffer, while the composition of cathode buffers A and B were 50 mM

541 Tricine, 7.5 mM imidazole pH 7.0, 0.02% Coomassie Blue G 250 dye and 50 mM Tricine,
542 7.5 mM imidazole pH 7.0, 0.002% Coomassie Blue G 250 dye, respectively.
543
544 For in-gel ATPase assay, after sample separation by BNP, the gel was incubated overnight in
545 assay buffer (35 mM Tris-HCL pH 7.8, 270 mM glycine, 14 mM MgSO4, 0.2% $Pb(NO_3)_2$,
546 and 8 mM ATP). The gel was then washed with water and the ATPase activity of F-type ATP
547 synthase was observed by the formation of a milky white precipitate, visible against a black
548 background, at the region corresponding to the expected monomeric and dimeric protein
549 bands. For Western blotting after BNP separation, proteins were electro-transferred onto a
550 PVDF membrane using Tris-glycine buffer pH 8.0 without methanol at 65 mAh for ~1hour
551 20 minutes. The membrane was then fixed in 8 % acetic acid for 15 minutes, air dried at
552 room temperature for 30 minutes, and washed with methanol several times to remove the
553 Coomassie dye stain [26]. Further steps were similar to that described above for SDS PAGE
554 Western blotting.
555
556 *Immunoprecipitation of F-type ATP synthase complex* – The supernatant from solubilized
557 mitochondria preparation from tachyzoite stage transgenic *T. gondii* expressing YFP-HA
558 tagged *Tg*ATPOSCP subunit was used for immunoprecipitating the F-type ATP synthase
559 using α-HA antibodies (Thermo Fisher Scientific, USA). The antibodies were first cross-
560 linked to Protein A/G Plus agarose beads (Santa Cruz Biotechnology, USA) using the
561 following protocol. ~50 µl of Protein A/G Plus agarose beads were washed twice in 500 µl
562 1X PBS (pH 7.4) at 4°C and mixed with 1 ml of 1X PBS containing 5-7 µg of rabbit α-HA
563 antibody and incubate overnight at 4°C with gentle mixing on rocker. The beads were then
564 separated by centrifugation (1000g for 2 minute at 4°C), equilibrate with 0.2 M
565 Triethanolamine (pH 8.2) for 2 minutes and then washed twice with the same buffer. The

566 bound antibodies were then cross-linked to the beads using 20 mM DMP (Sigma, USA) in 1
567 ml of 0.2 M Triethanolamine at room temperature for 45 minutes, followed by washing with
568 1ml of 50 mM Tris pH 7.5 twice for 15 minutes to quench the cross-linking reaction.
569 Unbound antibodies are removed using three quick washes with 0.2 M glycine buffer pH 2.3.
570 The antibody conjugated beads were then washed twice with 1X PBS and equilibrate with
571 MSB containing the detergent DDM at critical micelle concentration. The supernatant from
572 solubilized mitochondria was added to the antibody coupled beads and incubated overnight at
573 4°C with gentle mixing on rocker. The beads were then washed three times with 1X PBS
574 containing DDM at critical micelle concentration. Proteins captured by the antibodies were
575 eluted in 100 µl of 0.2 M glycine buffer (pH 2.3) in three rounds, and the pooled eluate was
576 neutralized with 1M Tris (pH 8.0) and concentrated using a 3 KDa cut-off concentrator
577 (Amicon Millipore, Germany). The samples were equilibrated with 0.1% RapiGest (Waters,
578 USA) in 50 mM ammonium bicarbonate buffer in preparation for LC-MS analysis.
579
580 *Purification of native F-type ATP synthase by chromatography* – We followed a protocol that
581 was previously reported for the purification of the F-type ATP synthase from the *Polytomella*
582 *sp*. [66]. About 13 mg (wet weight) of mitochondrial preparation from transgenic *T. gondii*
583 expressing YFP-HA tagged *Tg*atpβ subunit was solubilized overnight in 3 ml of
584 mitochondrial solubilization buffer B (MSB-B; 50 mM Tris HCl (pH 8.0), 1 mM $MgCl_2$, and
585 DDM at protein to detergent ration 1:5). The supernatant was loaded onto a HiTrap DEAE
586 sepharose column (GE Healthcare Life sciences, USA) having a bed volume of 1 ml and
587 equilibrated with buffer A (50 mM Tris HCl (pH 8.0) $MgCl_2$ 1mM, 0.05 % (w/v) DDM).
588 After completion of loading, the column was first washed with 10 column volumes of buffer
589 B_{20} (50 mM Tris HCl (pH 8.0), 1 mM $MgCl_2$, 20 mM NaCl, and 0.05 % (w/v) DDM) and
590 eluted with a 10X column volume linear gradient using buffer B_{20} and buffer B_{500} (50 mM

591 Tris HCl (pH 8.0), 1 mM $MgCl_2$, 500 mM NaCl, and 0.05 % (w/v) DDM) at a flow rate of
592 0.5 ml per minute. 500 µl fractions were collected and checked by Western blotting using α-
593 HA antibodies. Peak positive fractions were identified, pooled and concentrated to a final
594 volume of 200 µl using the Vivaspin concentrator (100 kDa cut-off; GE Healthcare Life
595 sciences, USA) and separated by size exclusion chromatography using the Superose 6
596 Increase 3.2/300 column (GE Healthcare Life sciences, USA) having a bed volume of 2.4 ml.
597 50 µl sample was loaded on the column previously equilibrated with buffer B_{20}, and then
598 eluted with 1.5 column volumes of the same buffer at 50 µl flow rate. 100 µl fractions were
599 collected and positive fractions, identified by Western blotting using α-HA antibodies, were
600 pooled and concentrated using a 3 KDa cut-off concentrator (Amicon Millipore, Germany).
601 Samples were then equilibrated with 0.1% RapiGest in 50 mM ammonium bicarbonate buffer
602 in preparation for LC-MS analysis. All chromatography steps were carried out using the
603 AKTApure system (GE Healthcare Life sciences, USA).
604
605 *LC-MS proteomics* – In order to identify the subunit components of the F-type ATP synthase
606 from *T. gondii*, we performed LC-MS analysis on the partially purified enzyme obtained
607 from BNP, immunoprecipitation and chromatographic purification. Sample preparation for
608 LC-MS analysis was based on a previous report [67].
609 *Sample preparation after BNP separation*: The bands corresponding to the dimer and
610 monomer forms of the protein were excised out from BNP gel, cut to small pieces, destained
611 with 50% acetonitrile in 50 mM ammonium bicarbonate, and dehydrate with 100 %
612 acetonitrile. The gel pieces were then treated with 10 mM DTT in 50 mM ammonium
613 bicarbonate for 45-60 minutes at 56ºC to reduce the proteins, followed by alkylation in dark
614 with 55 mM Iodoacetamide in 50 mM ammonium bicarbonate at ambient temperature for 45
615 minutes. Then, trypsinization of protein was carried out overnight at 37°C, and peptides were

616 extracted using 50% acetonitrile in 2% formic acid. The extracts were dried using a speedvac
617 and the peptides were reconstituted in 50 µl of 50 mM ammonium bicarbonate, acidified with
618 HCl and desalted using the C_{18} Zip tip columns (Millipore, Germany). The samples were then
619 dried in a speedvac and stored at -80°C until further use.

620 *Sample preparation after Immunoprecipitation and size exclusion chromatography*:
621 Samples obtained from immunoprecipitation and size exclusion chromatography were
622 immediately equilibrated with 0.1 % RapiGest as stated above. The mixtures were then
623 heated to 80°C and maintained at this temperature for 15 minutes. The denatured proteins
624 were reduced by 100 mM DTT at 60°C for 15 minutes, followed by alkylation with 200 mM
625 iodoacetamide at ambient temperature and in dark for 30 minutes. Samples were then treated
626 overnight with trypsin (Sigma Aldrich, USA) at 20:1 substrate to enzyme ratio at 37°C.
627 Trypsinization was stopped by addition of 2 µl of 1N HCL and incubating at 37°C for 20
628 minutes, after which they were vortexed and centrifuged. The peptide samples were desalted
629 using the C_{18} Zip tip (Millipore, Germany), dried and stored at -80°C until further use.

630 *LC-MS analysis*: The peptide samples were analyzed using LC-MS^E workflow on the
631 nano ACQUITY (UPLC) Synapt system (Waters Corporation, USA). 4 µl of the digested
632 peptide sample was injected into a 5 µm Symmetry C_{18} trapping column (180 µm x 20 mm)
633 at a flow rate of 5 nl/min and peptides were eluted by using the following protocol: 3-40 % B
634 (B - 100% acetonitrile with 0.1% formic acid) for 90 minutes, 40-85% B from 90-105
635 minutes and 97% A (A - 0.1% formic acid in water) and 3% B from 105-120 minutes, on a
636 bridge ethyl hybrid C_{18} column (75 µm x 250 mm, 1.7 µm) at flow rate of 250 nl/min. The
637 system was coupled to the Synapt High Definition Mass Spectrometer (Waters Corporation,
638 USA) with a nonoLock spray as the ion source. Standard glu-fibrinopeptide B (Sigma
639 Aldrich, USA) as the lockmass calibrate peptide was infused into ion source having 500
640 nl/min flow rate and sampled every 30 second. Acquisition of LC-MS^E data was performed in

641 positive V mode with mass range m/z 50-1900 having scan time of 0.75 second, a constant
642 low energy of 4 V for MS mode, and 20-40 V of collision energy during high energy MS^E
643 mode scan. A capillary voltage of 3.5 kV and cone voltage 38 V were maintained during
644 analysis. LC-MS^E data were processed using Protein Lynx Global Server (PLGS Version
645 2.5.3, Waters Corporation, USA) for identifying the proteins, with reference to the *T. gondii*
646 proteome taken from *ToxoDB.org* release 36.
647
648 *Gene co-expression analysis and phylogenetic studies* – Normalized signal intensity values
649 from microarray hybridizations performed at 13 different time points in the tachyzoite stage
650 [51] were obtained directly from *EupathDB.org*. Normalized microarray intensities from time
651 series data for intraerythrocytic stage *P. falciparum* were retrieved from reference [52].
652 Pearson correlation coefficients were calculated for all gene pairs using R (*R-project.org*). P-
653 values were calculated using the Mann-Whitney U test [68] for co-expression between ATP
654 synthase genes vs. co-expression between ATP synthase genes and non- ATP synthase genes,
655 and for co-expression between ATP synthase genes vs. co-expression between non- ATP
656 synthase genes. Since minute differences in very large samples may lead to very small p-
657 values, the effect size was therefore tested using Cohen's d test [69], where values above 0.8
658 denote large effect size. Orthologs for the novel subunits of *T. gondii* ATP synthase were
659 identified as reported previously [41]. Sequences were aligned with mafft (v7.222) [70] and
660 Neighbor-Joining trees were constructed using Clustalw (2.1) [71] using 1000 bootstrap
661 replicates. The expected 'true' phylogeny was adopted from a previous study [41].
662
663
664

Acknowledgements

The project is supported by funding from Science & Engineering Research Board (SERB), India (EMR/2016/003588), awarded jointly to DS and MB and This publication is also based upon work supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No. OCRF-2014-CRG3-2267 and the faculty baseline funding (BAS/1/1020-01-01) awarded to AP. RS received fellowship from Council of Scientific and Industrial Research, India. Authors would like to thank Dr. Mahesh J. Kulkarni and Ms. Rajeshwari Rathore, for LC-MS^E data acquisition and analysis, for Protein chromatography - Dr. Subashchandrabose Chinnathambi and Ms. Shweta Sonawane, CSIR-National Chemical Laboratory, Pune, and Prof. David. S. Roos, University of Pennsylvania, Philadelphia, USA, for providing *T. gondii* strains and other biological materials as gifts.

Author contributions

DS conceived the project and designed the experiments along with RS, MB and AP. RS and TM executed the various experiments. DS, RS, AP and TM were involved in manuscript preparation.

Competing interests

The authors declare no competeing interests, financial or otherwise, related to this manuscript.

689 **References**

690 1. Capaldi RA, Aggeler R (2002) Mechanism of the F(1)F(0)-type ATP synthase, a biological rotary
691 motor. Trends Biochem Sci 27: 154-160.
692 2. Devenish RJ, Prescott M, Rodgers AJ (2008) The structure and function of mitochondrial F1F0-ATP
693 synthases. Int Rev Cell Mol Biol 267: 1-58.
694 3. Mitchell P (1966) Chemiosmotic coupling in oxidative and photosynthetic phosphorylation. Biol
695 Rev Camb Philos Soc 41: 445-502.
696 4. Stock D, Leslie AG, Walker JE (1999) Molecular architecture of the rotary motor in ATP synthase.
697 Science 286: 1700-1705.
698 5. Boyer PD (1997) The ATP synthase--a splendid molecular machine. Annu Rev Biochem 66: 717-
699 749.
700 6. Vazquez-Acevedo M, Cardol P, Cano-Estrada A, Lapaille M, Remacle C, et al. (2006) The
701 mitochondrial ATP synthase of chlorophycean algae contains eight subunits of unknown
702 origin involved in the formation of an atypical stator-stalk and in the dimerization of the
703 complex. J Bioenerg Biomembr 38: 271-282.
704 7. Zikova A, Schnaufer A, Dalley RA, Panigrahi AK, Stuart KD (2009) The F(0)F(1)-ATP synthase
705 complex contains novel subunits and is essential for procyclic Trypanosoma brucei. PLoS
706 Pathog 5: e1000436.
707 8. Balabaskaran Nina P, Dudkina NV, Kane LA, van Eyk JE, Boekema EJ, et al. (2010) Highly divergent
708 mitochondrial ATP synthase complexes in Tetrahymena thermophila. PLoS Biol 8: e1000418.
709 9. Yadav KNS, Miranda-Astudillo HV, Colina-Tenorio L, Bouillenne F, Degand H, et al. (2017) Atypical
710 composition and structure of the mitochondrial dimeric ATP synthase from Euglena gracilis.
711 Biochim Biophys Acta 1858: 267-275.
712 10. Velours J, Arselin G (2000) The Saccharomyces cerevisiae ATP synthase. J Bioenerg Biomembr 32:
713 383-390.
714 11. Meyer B, Wittig I, Trifilieff E, Karas M, Schagger H (2007) Identification of two proteins associated
715 with mammalian ATP synthase. Mol Cell Proteomics 6: 1690-1699.
716 12. Ruhle T, Leister D (2015) Assembly of F1F0-ATP synthases. Biochim Biophys Acta 1847: 849-860.
717 13. Cain BD, Simoni RD (1989) Proton translocation by the F1F0ATPase of Escherichia coli. Mutagenic
718 analysis of the a subunit. J Biol Chem 264: 3292-3300.
719 14. Valiyaveetil FI, Fillingame RH (1997) On the role of Arg-210 and Glu-219 of subunit a in proton
720 translocation by the Escherichia coli F0F1-ATP synthase. J Biol Chem 272: 32635-32641.
721 15. Rastogi VK, Girvin ME (1999) Structural changes linked to proton translocation by subunit c of the
722 ATP synthase. Nature 402: 263-268.
723 16. Diez M, Zimmermann B, Borsch M, Konig M, Schweinberger E, et al. (2004) Proton-powered
724 subunit rotation in single membrane-bound F0F1-ATP synthase. Nat Struct Mol Biol 11: 135-
725 141.
726 17. Weber J (2006) ATP synthase: subunit-subunit interactions in the stator stalk. Biochim Biophys
727 Acta 1757: 1162-1170.
728 18. Walker JE, Dickson VK (2006) The peripheral stalk of the mitochondrial ATP synthase. Biochim
729 Biophys Acta 1757: 286-296.
730 19. Arnold I, Pfeiffer K, Neupert W, Stuart RA, Schagger H (1998) Yeast mitochondrial F1F0-ATP
731 synthase exists as a dimer: identification of three dimer-specific subunits. EMBO J 17: 7170-
732 7178.
733 20. Brunner S, Everard-Gigot V, Stuart RA (2002) Su e of the yeast F1Fo-ATP synthase forms
734 homodimers. J Biol Chem 277: 48484-48489.
735 21. Paumard P, Vaillier J, Coulary B, Schaeffer J, Soubannier V, et al. (2002) The ATP synthase is
736 involved in generating mitochondrial cristae morphology. EMBO J 21: 221-230.

737 22. Dudkina NV, Heinemeyer J, Keegstra W, Boekema EJ, Braun HP (2005) Structure of dimeric ATP
738 synthase from mitochondria: an angular association of monomers induces the strong
739 curvature of the inner membrane. FEBS Lett 579: 5769-5772.
740 23. Dudkina NV, Sunderhaus S, Braun HP, Boekema EJ (2006) Characterization of dimeric ATP
741 synthase and cristae membrane ultrastructure from Saccharomyces and Polytomella
742 mitochondria. FEBS Lett 580: 3427-3432.
743 24. Davies KM, Anselmi C, Wittig I, Faraldo-Gomez JD, Kuhlbrandt W (2012) Structure of the yeast
744 F1Fo-ATP synthase dimer and its role in shaping the mitochondrial cristae. Proc Natl Acad Sci
745 U S A 109: 13602-13607.
746 25. Mather MW, Henry KW, Vaidya AB (2007) Mitochondrial drug targets in apicomplexan parasites.
747 Curr Drug Targets 8: 49-60.
748 26. Balabaskaran Nina P, Morrisey JM, Ganesan SM, Ke H, Pershing AM, et al. (2011) ATP synthase
749 complex of Plasmodium falciparum: dimeric assembly in mitochondrial membranes and
750 resistance to genetic disruption. J Biol Chem 286: 41312-41322.
751 27. Danne JC, Gornik SG, Macrae JI, McConville MJ, Waller RF (2013) Alveolate mitochondrial
752 metabolic evolution: dinoflagellates force reassessment of the role of parasitism as a driver
753 of change in apicomplexans. Mol Biol Evol 30: 123-139.
754 28. Weiss LM, Dubey JP (2009) Toxoplasmosis: A history of clinical observations. Int J Parasitol 39:
755 895-901.
756 29. Dubey JP (1998) Advances in the life cycle of Toxoplasma gondii. Int J Parasitol 28: 1019-1024.
757 30. Dubey JP, Lindsay DS, Speer CA (1998) Structures of Toxoplasma gondii tachyzoites, bradyzoites,
758 and sporozoites and biology and development of tissue cysts. Clin Microbiol Rev 11: 267-
759 299.
760 31. Polonais V, Soldati-Favre D (2010) Versatility in the acquisition of energy and carbon sources by
761 the Apicomplexa. Biol Cell 102: 435-445.
762 32. Pomel S, Luk FC, Beckers CJ (2008) Host cell egress and invasion induce marked relocations of
763 glycolytic enzymes in Toxoplasma gondii tachyzoites. PLoS Pathog 4: e1000188.
764 33. Al-Anouti F, Tomavo S, Parmley S, Ananvoranich S (2004) The expression of lactate
765 dehydrogenase is important for the cell cycle of Toxoplasma gondii. J Biol Chem 279: 52300-
766 52311.
767 34. Blume M, Rodriguez-Contreras D, Landfear S, Fleige T, Soldati-Favre D, et al. (2009) Host-derived
768 glucose and its transporter in the obligate intracellular pathogen Toxoplasma gondii are
769 dispensable by glutaminolysis. Proc Natl Acad Sci U S A 106: 12998-13003.
770 35. MacRae JI, Sheiner L, Nahid A, Tonkin C, Striepen B, et al. (2012) Mitochondrial metabolism of
771 glucose and glutamine is required for intracellular growth of Toxoplasma gondii. Cell Host
772 Microbe 12: 682-692.
773 36. Nitzsche R, Zagoriy V, Lucius R, Gupta N (2016) Metabolic Cooperation of Glucose and Glutamine
774 Is Essential for the Lytic Cycle of Obligate Intracellular Parasite Toxoplasma gondii. J Biol
775 Chem 291: 126-141.
776 37. Vercesi AE, Rodrigues CO, Uyemura SA, Zhong L, Moreno SN (1998) Respiration and oxidative
777 phosphorylation in the apicomplexan parasite Toxoplasma gondii. J Biol Chem 273: 31040-
778 31047.
779 38. Saleh A, Friesen J, Baumeister S, Gross U, Bohne W (2007) Growth inhibition of Toxoplasma
780 gondii and Plasmodium falciparum by nanomolar concentrations of 1-hydroxy-2-dodecyl-
781 4(1H)quinolone, a high-affinity inhibitor of alternative (type II) NADH dehydrogenases.
782 Antimicrob Agents Chemother 51: 1217-1222.
783 39. Lin SS, Gross U, Bohne W (2009) Type II NADH dehydrogenase inhibitor 1-hydroxy-2-dodecyl-
784 4(1H)quinolone leads to collapse of mitochondrial inner-membrane potential and ATP
785 depletion in Toxoplasma gondii. Eukaryot Cell 8: 877-887.
786 40. Mogi T, Kita K (2010) Diversity in mitochondrial metabolic pathways in parasitic protists
787 Plasmodium and Cryptosporidium. Parasitol Int 59: 305-312.

788 41. Woo YH, Ansari H, Otto TD, Klinger CM, Kolisko M, et al. (2015) Chromerid genomes reveal the
789 evolutionary path from photosynthetic algae to obligate intracellular parasites. Elife 4:
790 e06974.
791 42. McFadden DC, Tomavo S, Berry EA, Boothroyd JC (2000) Characterization of cytochrome b from
792 Toxoplasma gondii and Q(o) domain mutations as a mechanism of atovaquone-resistance.
793 Mol Biochem Parasitol 108: 1-12.
794 43. Baggish AL, Hill DR (2002) Antiparasitic agent atovaquone. Antimicrob Agents Chemother 46:
795 1163-1173.
796 44. Pullman ME, Monroy GC (1963) A Naturally Occurring Inhibitor of Mitochondrial Adenosine
797 Triphosphatase. J Biol Chem 238: 3762-3769.
798 45. Krause F, Reifschneider NH, Goto S, Dencher NA (2005) Active oligomeric ATP synthases in
799 mammalian mitochondria. Biochem Biophys Res Commun 329: 583-590.
800 46. Zimmermann L, Stephens A, Nam SZ, Rau D, Kubler J, et al. (2017) A Completely Reimplemented
801 MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol.
802 47. Soding J (2005) Protein homology detection by HMM-HMM comparison. Bioinformatics 21: 951-
803 960.
804 48. Claros MG, Vincens P (1996) Computational method to predict mitochondrially imported
805 proteins and their targeting sequences. Eur J Biochem 241: 779-786.
806 49. Sidik SM, Huet D, Ganesan SM, Huynh MH, Wang T, et al. (2016) A Genome-wide CRISPR Screen
807 in Toxoplasma Identifies Essential Apicomplexan Genes. Cell 166: 1423-1435 e1412.
808 50. Dunkley TP, Watson R, Griffin JL, Dupree P, Lilley KS (2004) Localization of organelle proteins by
809 isotope tagging (LOPIT). Mol Cell Proteomics 3: 1128-1134.
810 51. Behnke MS, Wootton JC, Lehmann MM, Radke JB, Lucas O, et al. (2010) Coordinated progression
811 through two subtranscriptomes underlies the tachyzoite cycle of Toxoplasma gondii. PLoS
812 One 5: e12354.
813 52. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, et al. (2003) The transcriptome of the
814 intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol 1: E5.
815 53. Allen RD, Schroeder CC, Fok AK (1989) An investigation of mitochondrial inner membranes by
816 rapid-freeze deep-etch techniques. J Cell Biol 108: 2233-2240.
817 54. Buzhynskyy N, Sens P, Prima V, Sturgis JN, Scheuring S (2007) Rows of ATP synthase dimers in
818 native mitochondrial inner membranes. Biophys J 93: 2870-2876.
819 55. Davies KM, Strauss M, Daum B, Kief JH, Osiewacz HD, et al. (2011) Macromolecular organization
820 of ATP synthase and complex I in whole mitochondria. Proc Natl Acad Sci U S A 108: 14121-
821 14126.
822 56. Bornhovd C, Vogel F, Neupert W, Reichert AS (2006) Mitochondrial membrane potential is
823 dependent on the oligomeric state of F1F0-ATP synthase supracomplexes. J Biol Chem 281:
824 13990-13998.
825 57. Sturm A, Mollard V, Cozijnsen A, Goodman CD, McFadden GI (2015) Mitochondrial ATP synthase
826 is dispensable in blood-stage Plasmodium berghei rodent malaria but essential in the
827 mosquito phase. Proc Natl Acad Sci U S A 112: 10216-10223.
828 58. Roos DS (1993) Primary structure of the dihydrofolate reductase-thymidylate synthase gene from
829 Toxoplasma gondii. J Biol Chem 268: 6269-6280.
830 59. Donald RG, Roos DS (1993) Stable molecular transformation of Toxoplasma gondii: a selectable
831 dihydrofolate reductase-thymidylate synthase marker based on drug-resistance mutations in
832 malaria. Proc Natl Acad Sci U S A 90: 11703-11707.
833 60. Kim K, Soldati D, Boothroyd JC (1993) Gene replacement in Toxoplasma gondii with
834 chloramphenicol acetyltransferase as selectable marker. Science 262: 911-914.
835 61. Huynh MH, Carruthers VB (2009) Tagging of endogenous genes in a Toxoplasma gondii strain
836 lacking Ku80. Eukaryot Cell 8: 530-539.
837 62. Roos DS, Donald RG, Morrissette NS, Moulton AL (1994) Molecular tools for genetic dissection of
838 the protozoan parasite Toxoplasma gondii. Methods Cell Biol 45: 27-63.

63. Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248-254.
64. Wittig I, Braun HP, Schagger H (2006) Blue native PAGE. Nat Protoc 1: 418-428.
65. Nijtmans LG, Henderson NS, Holt IJ (2002) Blue Native electrophoresis to study mitochondrial and other protein complexes. Methods 26: 327-334.
66. Allegretti M, Klusch N, Mills DJ, Vonck J, Kuhlbrandt W, et al. (2015) Horizontal membrane-intrinsic alpha-helices in the stator a-subunit of an F-type ATP synthase. Nature 521: 237-240.
67. Korwar AM, Vannuruswamy G, Jagadeeshaprasad MG, Jayaramaiah RH, Bhat S, et al. (2015) Development of Diagnostic Fragment Ion Library for Glycated Peptides of Human Serum Albumin: Targeted Quantification in Prediabetic, Diabetic, and Microalbuminuria Plasma by Parallel Reaction Monitoring, SWATH, and MSE. Mol Cell Proteomics 14: 2150-2159.
68. Myles Hollander DAW, Eric Chicken (2013) Nonparametric Statistical Methods, 3rd Edition: Wiley. 848 p.
69. Cohen J (1988) Statistical power analysis for the behavioral sciences. Hillsdale, N.J.: L. Erlbaum Associates. xxi, 567 p. p.
70. Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30: 772-780.
71. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947-2948.
72. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7: 539.
73. Robert X, Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res 42: W320-324.

873 **Main Figure Legends**

874 **Fig 1. Detail of F-type ATP synthase subunits missing in *T. gondii* and other**
875 **apicomplexan parasites.**

876 (A) Schematic representation of core subunit composition of F-type ATP synthase. The
877 different subunits are color coded and annotated. Middle, the complete set of subunits of the
878 core enzyme; left and right, subunits identified and not identified by *in silico* methods in
879 apicomplexan genome. (B) Table showing the presence/absence of orthologs of yeast and
880 bovine F-type ATP synthase in *Apicomplexa* and other alveolate species. +/-,
881 presence/absence of ortholog; blue, diverged (no clear orthologs) functional equivalent; light
882 blue, experimentally identified species-specific novel subunits (numeric values indicate the
883 total number of such novel subunits). Species names: *Pfal*, *P. falciparum*; *Pviv*, *P. vivax*;
884 *Tpar*, *T. parva*; *Bbov*, *B. bovis*; *Tgon*, *T. gondii* ; *Ncan*, *N. caninum*; *Cmur*, *C. muris*; *Cpar*, *C.*
885 *parvum*; *Chom*, *C. hominis*; *Cvel*, *C. velia*; *Pmar*, *P. marinus*; *Tthe*, *T. thermophila*; *Scer*, *S.*
886 *cerevisiae*; *Btau*, *B. taurus*; *Atha*, *A. thaliana*; *Crei*, *C. reinhardtii*; *Tbru*, *T. brucei*.

887 **Fig 2. Endogenous tagging of *Tgatpβ* and *Tgatposcp* genes with YFP-HA tags.**

888 (A) Schematic representation of the *Tgatpβ* (TGME49_261950) gene locus showing coding
889 region in light blue and 3' UTR in grey. The hatched region shows the homology region
890 (HR), which was cloned into a donor plasmid (dark blue) upstream of an in-frame coding
891 region for YFP-HA tag. The donor plasmid also includes the DHFR cassette, which confers
892 pyrimethamine resistance. Asterisk denotes the position of stop codon. The donor plasmid
893 was linearized with the restriction enzyme BstZ17I to facilitate single crossover homologous
894 recombination and the resulting modification of the gene locus was confirmed by PCR using
895 the primers ConF and YFP_R. A similar strategy was used to tag the *Tgatposcp* gene locus.
896 (B) and (C) Confirming the expression of the *Tg*ATPβ−YFP-HA and *Tg*ATPOSCP-YFP-HA

897 by SDS-PAGE Western blotting (B) and microscopy (C) on respective clonal transgenic
898 lines. Gel lanes in (B): 1, Cell lysate from *Tg*ATPb-YFP-HA transgenic parasites; 2, Cell lysate
899 from *Tg*ATPOSCP-YFP-HA transgenic parasites; 3, Cell lysate from parental parasites; M,
900 Molecular weight size markers. Mitotracker was used to visualize the mitochondrion in (C). (D)
901 Functional validation of mitochondrial ATP synthesis in *wt* parental strain, and *Tgatpβ-YFP-*
902 *HA* and *Tgatposcp-YFP-HA* expressing transgenic parasites. All three strains exhibited
903 similar response to atovaquone treatment in the presence and absence of glucose, thereby
904 confirming that the YFP-HA tag had no effect on mitochondrial ATP synthesis.

905 **Fig 3. Identifying the dimer and monomer forms of *T. gondii* F-type ATP synthase by**
906 **BNP analysis.**

907 (A) Mitochondria lysates were prepared from tachyzoites stage *Tgatpβ-YFP-HA* transgenic
908 parasites were separated by BNP. Lane M, native molecular weight markers; lane A,
909 Coomassie blue staining of BNP gel; lane B, Western blotting of BNP gel using α-HA
910 antibodies. The dimer and monomer forms are indicated by arrows. The boxed region in lane
911 A corresponds to the excised gel piece, which was processed for LC-MS/MS analysis. (B)
912 In-gel ATPase activity assays following BNP separation confirms that the dimer and
913 monomer forms of F-type ATP synthase are functionally intact.

914

915 **Fig 4. Partial purification of dimer and monomer forms of *T. gondii* F-type ATP**
916 **synthase by chromatography.**

917 (A) Ion exchange (DEAE sepharose) separation of mitochondrial lysates prepared from
918 *Tgatpβ-YFP-HA* transgenic parasites. Absorbance at 280 nm (filled circles) and NaCl
919 concentration (open circles) are plotted for each fraction. Fractions 5 and 18 are marked with
920 arrows. Bottom panel shows SDS-PAGE Western blotting for fractions 5-18 to find out the

921 elution profile of *Tg*ATPβ-YFP-HA. (B) Size exclusion profile of the pooled fractions from
922 ion exchange chromatography. The Absorbance at 280 nm is plotted for each fraction. The
923 size exclusion column was calibrated with the following native markers; thyroglobulin (T,
924 660 kDa), ferritin (F, 440 kDa), conalbumin (C, 75 kDa), Ovalbumin (O, 45 kDa). Peak
925 elution volume for each marker is indicated by arrow. Fractions 1-3, 4-6 and 7-9 were pooled,
926 concentrated and subject to SDS-PAGE Western blotting to detect *Tg*ATP-β-YFP-HA as
927 shown in bottom panel.

928 **Fig 5. Identification of novel *T. gondii* F-type ATP synthase subunits from LC-MS/MS**
929 **analysis.**

930 (A) The Venn diagram shows shared identification of proteins following blue native page
931 (BNP), size exclusion chromatography (SEC) and immunoprecipitation (IP) sample
932 preparation. Total proteins identified in each technique is given within brackets outside the
933 Venn diagram. The numbers shown in white font and within brackets are the final set of
934 proteins assigned as subunit components of *T. gondii* F-type ATP synthase. (B) List of all
935 ASAPs identified in this work by mass spectrometry analysis. The entries shown in bold
936 indicate that the protein was detect with high confidence from only BNP and IP samples.
937 Asterisk indicates that the corresponding peptides were detected in only one of the replicate
938 runs for SEC sample. BNP data is a combination of experiments done with both *Tg*ATPβ-
939 YFP-HA and *Tg*ATPoscp-YFP-HA expressing transgenic parasites. SEC and IP data are
940 from *Tgatpβ-YFP-HA* and *Tgatposcp-YFP-HA* transgenic parasites respectively.

941 **Fig 6. HHPred identification of novel F_O subunit a based on conserved structural**
942 **features.**

943 (A) Representation of the pairwise sequence alignments generated by HHPred [46,47] for the
944 putative F_O subunit a from *T. gondii* and four other F_O subunit a proteins for which the

945 structure is known. The table provides details of the amino acid length and a probability score
946 for the prediction from the hit alignments. The red lines indicated the 3 transmembrane
947 domains present in the *T. gondii* protein. (B) Protein sequence alignments and secondary
948 structure information were made using the Clustal Omega [72] and ESPript [73] softwares,
949 and only the C-terminal portion of respective proteins is shown. The names of alveolate
950 species are highlighted in blue. For the non-alveolate species included in the alignment, the
951 F_O subunit a is either readily identified from sequence (*Scer, Hsap, Atha, Ecol* and *Pten*), or
952 has been experimentally determined (*Crei* and *Poly*). Positions with similar amino acids are
953 highlighted in red in the alignment. The arginine and glutamine residues, highlighted in
954 yellow, are conserved in all species and important for function. The helices (α9, α10 and
955 α12) shown are from the structure of *Poly* F-type ATP synthase. Species names: *Tgon*, *T.*
956 *gondii*; *Hham*, *H. hammondi*; *Pfal*, *P. falciparum*; *Pviv*, *P. vivax*; *Tann*, *T. annulata*; *Bbov*, *B.*
957 *bovis*; *Cmur*, *C. muris*; *Cvel*, *C. velia*; *Crei*, *C. reinhardtii*; *Poly*, *Polytomella sp.*; *Scer*, *S.*
958 *cerevisiae*; *Hsap*, *H. sapiens*; *Atha*, *A. thaliana*; *Ecol*, *E. coli*; *Pden*, *P. denitrificans*.

959 **Fig 7. Phylogenetic profile of the alveolate infrakingdom for all subunits mapped to *T.***
960 ***gondii* F-type ATP synthase.**

961 Alveolate clades are highlighted with a grey background and their expected phylogenetic
962 relationship is indicated by a tree structure above. Grey and white boxes indicate the presence
963 and absence of the corresponding ortholog, respectively. The hatched boxes represent the
964 presence of the ortholog in *C. muris* only and absence in *C. parvum* and *C. hominis*. The table
965 on the left lists the gene ID for all ASAPs, along with the annotation, essentiality phenotypes
966 (phenotype score from CRISPR/Cas9 knockout [49]), and protein localization experiments
967 from an independent unpublished study (see text in Materials and Methods for details). The
968 key for localization annotation is given below the table.

969 **Fig 8. Gene co-expression analysis for *T. gondii* and *P. falciparum* F-type ATP synthase**
970 **subunits.**

971 The distribution of co-expression values - as measured by the Pearson correlation coefficients
972 – are plotted for the three gene pairs categories, as shown for *T. gondii* (A) and *P. falciparum*
973 (B). The statistical support from Mann-Whitney U p-values and Cohen's d values are shown
974 in the table within each plot. Blue (1), co-expression correlation between ATP synthase
975 genes; Orange (2), co-expression correlation between ATP synthase genes and non- ATP
976 synthase genes; Green (3), co-expression correlation between non- ATP synthase genes.
977 MWU, Mann-Whitney U p-values; Cohen's d, Cohen's d values.

978

979

980

981

982

983

984

985

986

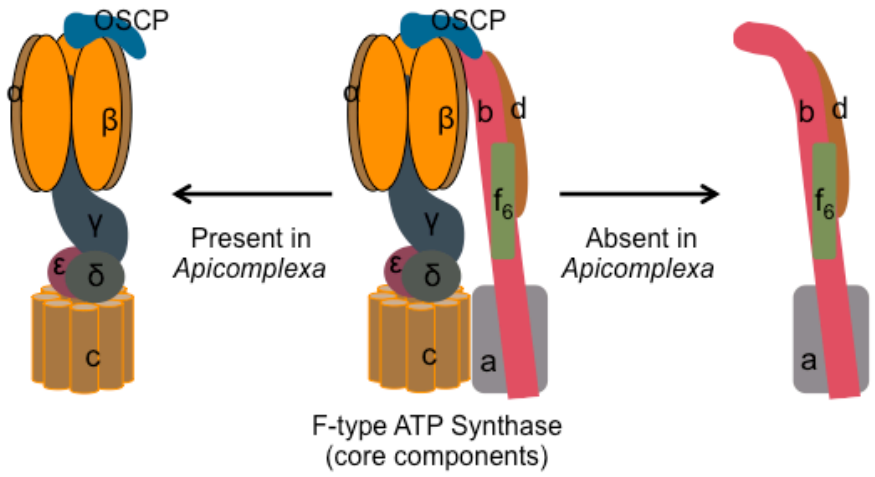
987

988

989

Salunke et. al., Fig 1

A



B

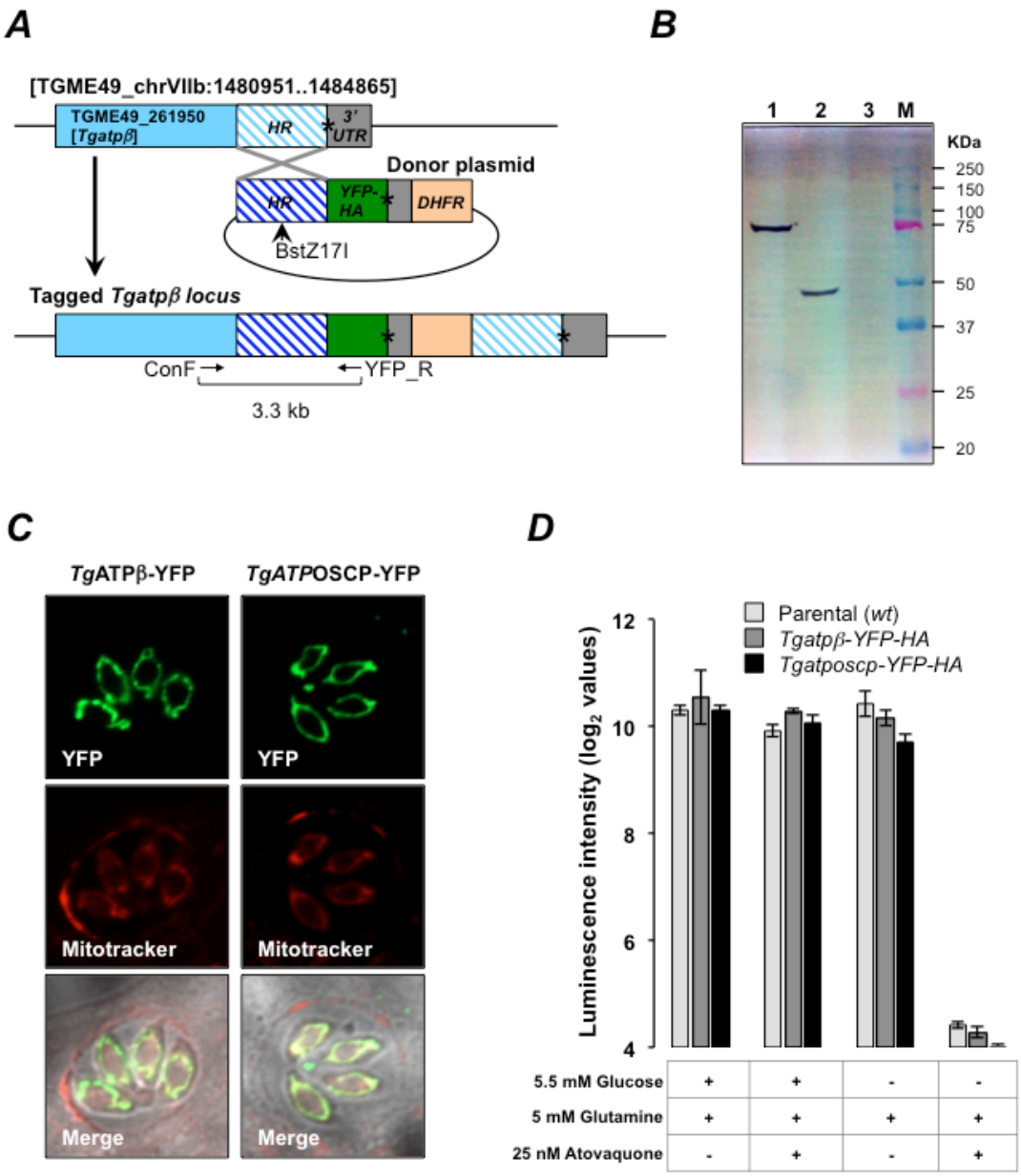
F-type ATP synthase module	Sub-unit	*Pfal*	*Pviv*	*Tpar*	*Bbov*	*Tgon*	*Ncan*	*Cmur*	*Cpar*	*Chom*	*Cvel*	*Pmar*	*Tthe*	*Scer*	*Btau*	*Atha*	*Crei*	*Tbru*
F_1 Sector	α	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	β	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
	γ	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+
	ε	+	+	+	+	+	+	+	-	-	+	+	-	+	+	+	+	+
	δ	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+
F_O Sector	c	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+
	OSCP	+	+	+	+	+	+	+	-	-	+	+	+	+	+	+	+	+
	a	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+
	b	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	+
	d	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	-	-
	f	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
	8/A6L	-	-	-	-	-	-	-	-	-	-	-	-	*	+	+	-	-
	h/f6	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
Other associated proteins	g	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
	IF1	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
	e	-	-	-	-	-	-	-	-	-	-	-	-	+	+	+	-	-
	j	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
	k	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-
	AGP	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
	MLQ	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-
Novel subunits													13				9	14

990

991

992

Salunke et. al., Fig 2

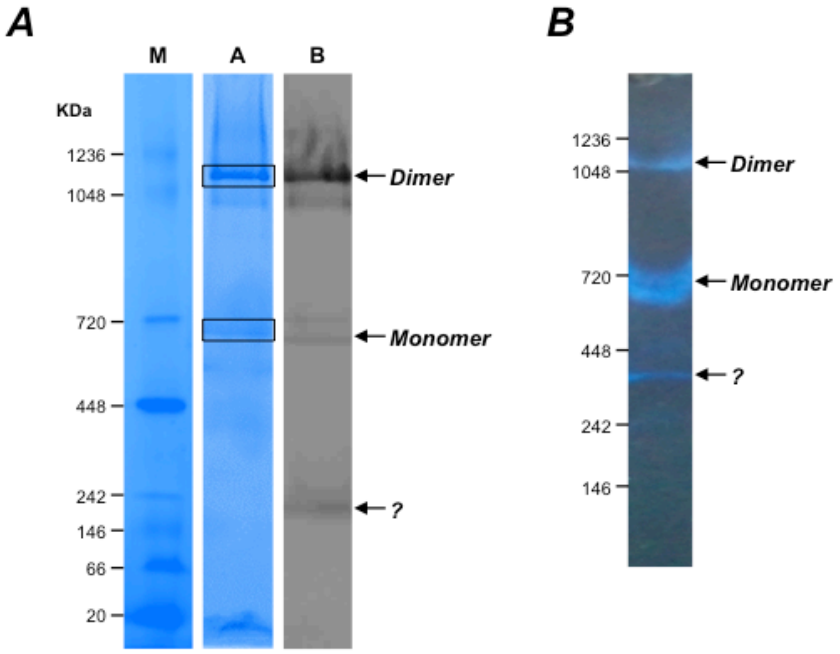


993

994

995

Salunke et. al., Fig 3

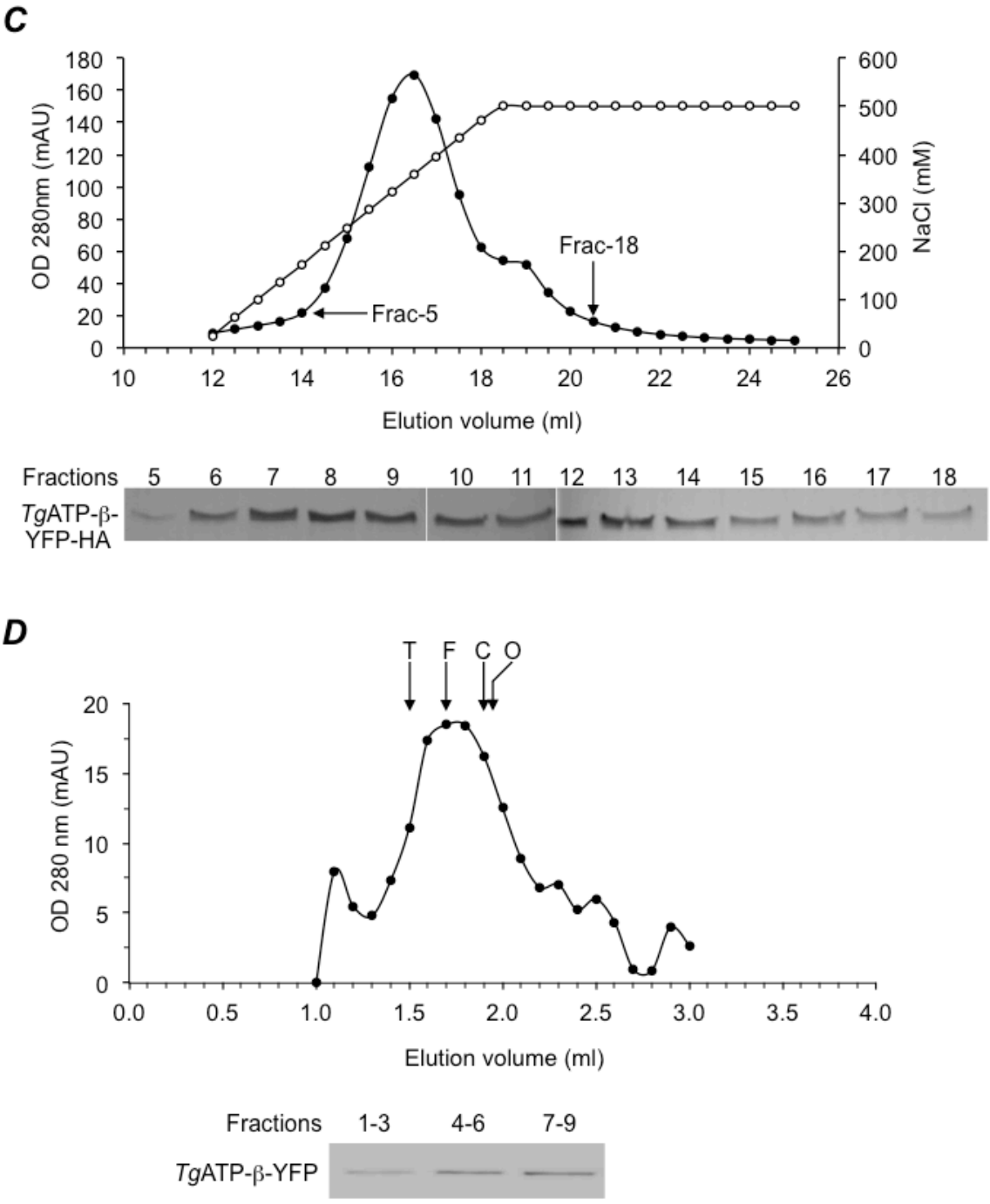


996

997

998

Salunke et. al., Fig 4



999

1000

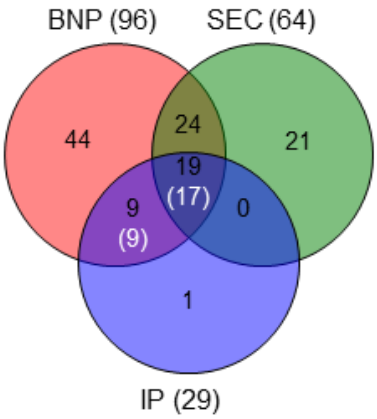
1001

1002

1003

Salunke et. al., Fig 5

A



B

S. No.	List of proteins	Product description from ToxoDB
1	TGME49_204400	F_1alpha subunit
2	TGME49_261950	F_1beta subunit
3	TGME49_226000	F_1delta subunit
4	TGME49_314820	F_1epsilon subunit
5	**TGME49_231910**	**F_1 gamma subunit**
6	TGME49_284540	F_0 OSCP subunit
7	**ASAP-1**	**hypothetical protein**
8	ASAP-2	hypothetical protein
9	ASAP-3	hypothetical protein
10	ASAP-4	hypothetical protein
11	ASAP-5	hypothetical protein
12	**ASAP-6**	**hypothetical protein***
13	**ASAP-7**	**hypothetical protein**
14	ASAP-8	hypothetical protein
15	ASAP-9	hypothetical protein
16	ASAP-10	hypothetical protein
17	**ASAP-11**	**hypothetical protein**
18	**ASAP-12**	**hypothetical protein**
19	**ASAP-13**	**hypothetical protein**
20	ASAP-14	hypothetical protein
21	**ASAP-15**	**hypothetical protein**
22	ASAP-16	hypothetical protein
23	ASAP-17	membrane protein
24	ASAP-18	hypothetical protein
25	ASAP-19	hypothetical protein
26	**ASAP-20**	**hypothetical protein**

1004

Salunke et. al., Fig 6

A

HHPred prediction for *ASAP 1*

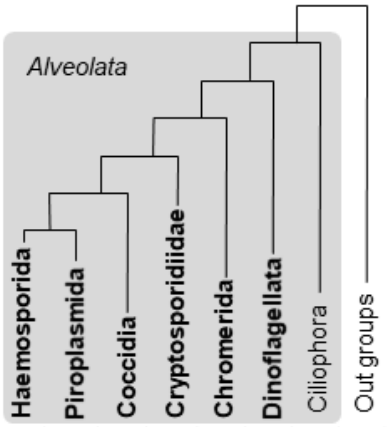
Hit	AA No.	Probability
5T40_K	196..265	64.3
6F36_M	215..308	37.7
5DN6_X	193..276	24.5
6B8H_a	154..243	23.3

B

```
                                α9                  α10
                          ooooooooo  oooooooooooooooooooooooooooooo
                     .         .         .         .         .         .
Tgon  133 FLRPTK----MASFAATFLLNFRFYFMYMARTTFQAVRPLLAFSVFGEVMKLVLATMSSG  188
Hham  133 FLRPAK----MASFAATFLLNFRFYFMYMARTTFQAVRPLLAFSVFGEVMKLVLATMSSG  188
Pfal  88  YVNPLD----MTKFAIKYILSYRFFFIYMARTTFQAVRPLMAFCVFGELMKLILATMTSG  143
Pviv  86  YVNPID----MTKFTIKYILSYRFFFIYMARTTFQAVRPLMAFCVFGELMKLILATMTSG  141
Tann  113 YVNPVS----IVKFTAKYLLSYRFFFIYMARTTFQAARPLLAFCVFGEIMKLVLANLSGG  168
Bbov  82  YVDPLS----MVKFASKYVLSLRFFFIYMARTTFQAVRPLLAFCVFGEIMKLILANISGG  137
Cmur  110 FIYFTD----ILNFTSSIFTNLREYITLLARMSFQAIRPLLAFSVLGEAVKIALALITGS  165
Cvel  304 KFSLNQ----FLSFGYGYVFSLRDIFIFAFRLIFQTGRPMMAVYIGGTLMKTVFSITLSY  359
Crei  217 F--PGHFIPGGTPWPMAFIFVPLETISYTFRAVSLGVRLWVNMLAGHTLLHILTGMALAL  274
Poly  204 F--PGHFIPGGTPWPMAFIFVPLETISYTFRAVSLGVRLWVNMLAGHTLLHILTGMALAL  261
Scer  151 F--FSLFVPAGTPLPLVPLLVIIETLSYFARAISLGLRLGSNILAGHLLMVILAGLTFNF  208
Hsap  124 A--LAHFLPQGTPTPLIPMLVIIETISLLIQPMALAVRLTANITAGHLLMHLIGSATLAM  181
Atha  276 F--FSFLLPAGVPLPLAPFLVLLELISYCFRALSLGIRLFANMMAGHSLVKILSGFAWTM  333
Ecol  173 GFTKELTLQPFNHWAFIPVNLILEGVSLLSKPVSLGLRLFGNMYAGELIFILIAGLLPWW  232
Pden  147 F--LGLFWVSSAPLALRPVLAVIELISYFVRPVSHSIRLAGNIMAGHAVIKVFAAFAAVA  204
                          .      .    :      *    .      .    .

                                α12
                          oooooooooooooooooooooooooooooooooooo
                     .         .         .         .         .         .
Tgon  189 ----------------LFSFLFSFVLAFEVFYFFLQCYISY----TFLTMFFTVLF----  224
Hham  189 ----------------LFSFLFSFVLAFEVFYFFLQCYISY----TFLTMFFTVLF----  224
Pfal  143 ----------------VFAFFFSFVLAFEVLYFFLQCYISY----TFLTMFFDVMF----  179
Pviv  142 ----------------VFAFFFSFVLAFEVLYFFLQCYISY----TFLTMFFDVMF----  177
Tann  168 ----------------VPAYLFSFVLAFEVFYFFLQLYISY----TFLMMFFTVMF----  204
Bbov  138 ----------------VPAFLFSFVLAFEVLYFFLQCYISY----TFLTMFFTVMF----  173
Cmur  166 ----------------SIAFIFSFILAFEVFYFFLQCFISF----VFLSMFFKAIV----  201
Cvel  360 ----------------SFLLFFSLIAGFEVFYFGLQCYIGF----VFLSFFSEGALM---  396
Crei  275 P--FSLSFFAMVPATFAVACLLSALVGLEYLVAVLQSGVFSILSTVYVGEFNSVKLAGPL  332
Poly  262 P--FSLGFFSMVPATFGVCCLLSALVGLEYLVAVLQSGVFSILSTVYVGEFNHDKFIGPA  319
Scer  209 M--LINL-FTLVF-GFVPLAMILAIMMLEFAIGIIQGYVWAILTASYLKDAVYLH-----  259
Hsap  182 S--TINLPSTL-----IIFTILILLTILEIAVALIQAYVFTLLVSLYLHDNT--------  226
Atha  334 L--CMNDIFYFIG-ALGPLFIVLALTGLELGVAILQAYVFTILICIYLNDAINLH-----  385
Ecol  233 SQWILNV----------------PWAIFHILIITLQAFIFMVLTIVYLSMASEEH-----  271
Pden  205 A--I----------APVSVVAITAMYGLEVLVCLIQAYVFTILTCVYLKDALHPAH----  248
                                     :.     :*  :       ::
```

Salunke et. al., Fig 7



Subunit Annotation	CRISPR Pheno Score*	Mito. location	Haemosporida	Piroplasmida	Coccidia	Cryptosporidiidae	Chromerida	Dinoflagellata	Ciliophora	Out groups
F_1 alpha	-3.84	Mito#	■	■	■	■	■	■	■	■
F_1 beta	-4.84	**Mito#**	■	■	■	■	■	■	■	■
F_1 delta	-4.57	Mito#	■	■	■	▨	■	■	□	■
F_1 epsilon	-3.21	Mito#	■	■	■	▨	■	■	□	■
F_1 gamma	-3.94	Mito#	■	■	■	▨	■	■	■	■
F_0 c	-2.98	Mito	■	■	■	▨	■	■	■	■
OSCP	-3.94	**Mito#**	■	■	■	▨	■	■	■	■
ASAP-1 (F_0 a?)	-4.49	Mito	■	■	■	▨	■	□	□	□
ASAP-2 (F_0 b?)	-5.37	Mito*#	■	■	■	▨	■	■	□	□
ASAP-3 (F_0 d?)	-3.10	Mito#	■	■	■	▨	■	■	□	□
ASAP-4	-4.49	**N#**	■	■	■	▨	■	■	□	□
ASAP-5	-3.90	N*#	■	■	■	▨	■	■	□	□
ASAP-6	-4.07	Mito*#	■	■	■	▨	■	■	□	□
ASAP-7	-3.92	N*#	■	■	■	▨	■	■	□	□
ASAP-8	-2.86	**Mito#**	■	■	■	▨	■	■	□	□
ASAP-9	-1.87	N#	■	■	■	▨	■	■	□	□
ASAP-10	-3.55	Mito*#	■	■	■	▨	■	□	□	□
ASAP-11	-3.88	**Mito#**	■	■	■	□	■	■	□	□
ASAP-12	-3.40	N#	■	■	■	□	■	□	□	□
ASAP-13	-1.37	N#	■	■	■	□	■	■	□	□
ASAP-14	-2.95	**Mito#**	■	■	■	□	■	■	□	□
ASAP-15	-3.54	Mito#	■	■	■	▨	□	□	□	□
ASAP-16	-4.01	Mito#	□	■	■	□	■	□	□	□
ASAP-17	-3.65	Mito#	■	□	■	□	■	■	□	□
ASAP-18	-3.10	Mito#	□	□	■	▨	■	■	□	□
ASAP-19	0.22	N#	□	□	■	□	■	□	□	□
ASAP-20	0.32	N#	□	□	■	□	■	□	□	□

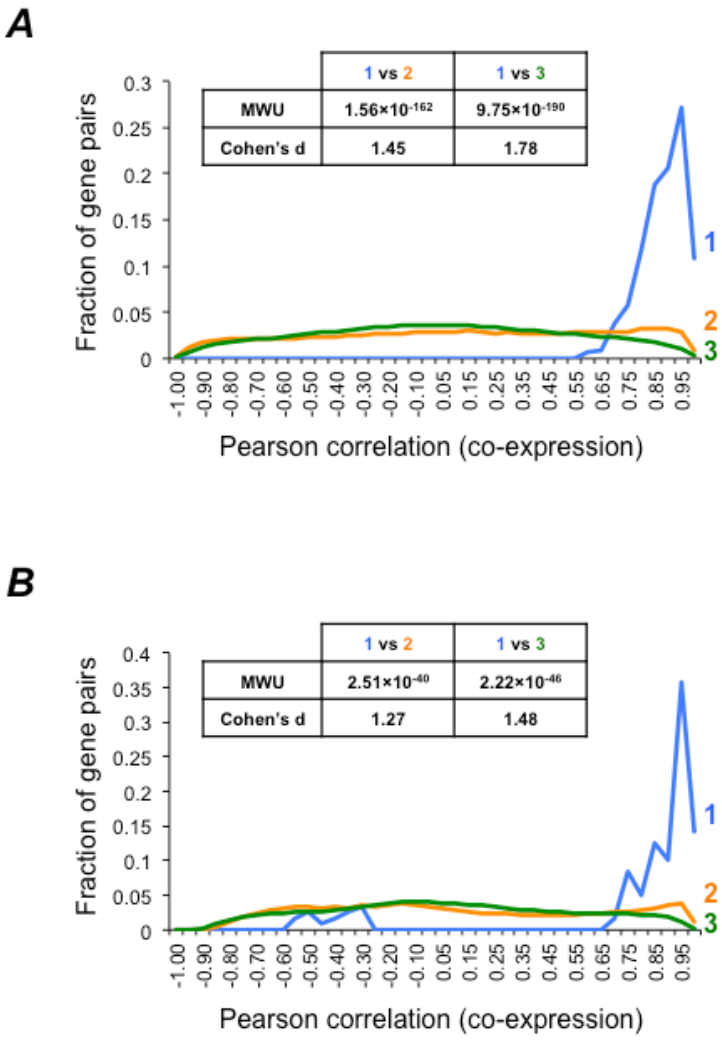
Key for location data
Mito = mitochondria localization signal present
N = no mitochondria localization signal
Bold = Localized in this study
* = Localized previously [50]
\# = unpublished evidence from Hyper-LOPIT experiment

1008

1009

1010

Salunke et. al., Fig 8



1011

1012