Genetic analyses of Chinese isolates of *Toxoplasma gondii* reveal a new genotype with high virulence to murine hosts

J.M. GAO[a], Y.T. XIE[a], Y. Z. WEN[a], Z.S. XU[a], H. CHEN[b], G. HIDE[c], T.B. Yang[a], J.L. SHEN[b], D.H. LAI [a,*], and Z.R. LUN[a,c,*]

[a] Center for Parasitic Organisms, State Key Laboratory of Biocontrol, School of Life Sciences, and Key Laboratory of Tropical Disease Control of the Ministry of Education, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510275, China

[b] The Anhui Provincial Laboratory of Pathogen Biology, Anhui Medical University, Hefei 230032, China

[c] Ecosystems and Environment Research Centre and Biomedical Research Centre, School of Environment and Life Sciences, University of Salford, Salford, M5 4WT, UK

*Corresponding authors. School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, P.R. Tel: + 86-20-84115079; fax: + 86-20-84036215; E-mail address: lsslzr@mail.sysu.edu.cn (Z.-R. Lun) or laidehua@mail.sysu.edu.cn (D.-H. Lai)

ABSTRACT

A great deal of evidence demonstrates that a strongly clonal population structure of *Toxoplasma gondii* strains exists in humans and animals in North America and Europe, while the strains from South America are genetically separate and more diverse. Potential differences in virulence between different strains mean that an understanding of strain diversity is important to human and animal health. However, to date, only one predominant genotype, ToxoDB#9 (Chinese I), and a few other genotypes, including ToxoDB#205, have been identified in China. By using DNA sequence-based phylogenetic analyses, we have re-evaluated the population structure of *T. gondii* strains collected from China and compared them with other global strains. Based on phylogenetic analysis of restriction fragment length polymorphisms, multilocus sequence typing and neutral gene sequences from *T. gondii*, we propose that the Chinese isolates described as Chinese I are divided into two groups called Chinese I and Chinese III. Our results demonstrate that significant differences were found in mouse mortality caused by some Chinese strains, and also the archetypal I, II, III strains in mice. Furthermore, a comparison of cyst loading in the brains of infected rats showed some Chinese strains to be capable of a high degree of cyst formation. Furthermore we show that genotyping using neutral genetic markers may not be a useful predictor of pathogenicity phenotypes.

Key word: Chinese; *Toxoplasma gondii*; Genotype; MLST; RFLP; Virulence

1. Introduction

Toxoplasma gondii is an important, globally distributed, intracellular parasite and provides a valuable model system to understand the evolution of intracellular pathogens. It not only infect large numbers of warm blooded animals including birds, livestock and humans, but also marine mammals (Dubey, 2010; Montoya and Liesenfeld, 2004). Animals and humans are mainly infected by ingesting food or water contaminated with *T. gondii* oocysts or consuming raw or undercooked meat containing parasite cysts (Dubey and Beattie, 1988). The life cycle of the parasite includes domestic cats and other felids as definitive hosts (Frenkel et al., 1970) while virtually all warm-blooded vertebrates can act as intermediate hosts (Dubey and Beattie, 1988). In the intermediate hosts, *T. gondii* undergoes asexual reproduction as either tachyzoites, during acute infection, or bradyzoites (cysts) during chronic infection. In the definitive host it goes through sexual reproduction to produce a high output (many millions daily) of the highly infective oocyst stage.

However, despite the sexual reproductive phase in the life cycle, initially only a few genotypes were recognized in *T. gondii* and were referred to as type I, II, and III. These archetypal types, all together, accounted for 95% of the strains isolated in North America and Europe (Ajzenberg et al., 2002; Howe and Sibley, 1995). In addition, these strains (type I, II, and III) also predominate in chickens from Africa, where a higher prevalence of type II and III are found (Velmurugan et al., 2008). Intriguingly, although the differences at the genomic level among the three main lineages are less than 1%, the virulence phenotypes in mice can differ markedly. Typically, type I strains are uniformly lethal ($LD_{100} = 1$) to mice; in contrast, types II and III strains are less virulent ($LD_{50} \geq 10^5$) (Howe and Sibley, 1995; Khan et al., 2009; Sibley and

Boothroyd, 1992). These archetypal strains are typically identified by techniques such as restriction fragment length polymorphism (RFLP) (Pena et al., 2008) or microsatellite analysis (Lehmann et al., 2006). These techniques have been widely used for genotyping a broad range of organisms (Anderson et al., 2000; Cameron et al., 1988; Hide and Tait, 2009; Widmer et al., 2004; Severson et al., 1995). In addition, use of these methods has identified other minor subpopulations of *T. gondii*, e.g. the Africa 1 and Africa 3 strains which sit amongst the major types I, II and III in Africa (Mercier et al., 2010; Rajendran et al., 2012). Using the same approaches, a wide diversity of genotypes have been observed in South America and probably reflects frequent sexual recombination (Khan et al., 2006; Lehmann et al., 2006; Pena et al., 2008; Su et al., 2012). To date, *T. gondii* strains have been categorized into at least includes six major clades and over one hundred haplotypes (Lehmann et al., 2006; Su et al., 2012). However, data from China has revealed that only a few genotypes of *T. gondii* have been isolated from human and animal hosts, including ToxoDB#205 (designated as Chinese II in this paper) and predominantly Chinese I (ToxoDB#9) (>73% based on over 60 isolates) (Chen et al., 2011; Dubey et al., 2007; Qian et al., 2012; Wang et al., 2013; Zhou et al., 2009) This dominance of Chinese I occurs across the wide geographical range of these investigations (over 640,000 km^2).

Intriguingly, previous studies have demonstrated that the Chinese I isolates vary in their virulence to mice (Chen et al., 2011; Qian et al., 2012; Wang et al., 2013). Some isolates consistently display high virulence, while others show low virulence. This suggests that they could be virulence mutants or possibly distinct strains that have not yet been resolved using traditional approaches. In order to better understand the correlation between virulence and genotype of these Chinese *T. gondii* isolates, a more sensitive and accurate method is needed. Multilocus Sequence Typing (MLST) is an

approach which could create a better dataset by revealing a greater range of possible DNA sequence polymorphisms including SNPs, insertions and deletions among strains (Su et al., 2012). Therefore, here, we re-examined the genotyping of an expanded set of *T. gondii* isolates using classical RFLP as well as MLST and performed further phylogenetic and population genetic analyses to re-evaluate the population structure of isolates in China.

2. Materials and Methods

2.1. Animals

Female Swiss Webster mice and female Sprague-Dawley (SD) rats were purchased from the Experimental Animal Center of Sun Yat-Sen University. Animals at age 8-10 weeks (mice, weighing 20-25 g) and 4 weeks (rats, weighing approximately 60 g) were used for experiments. They were routinely maintained in a special pathogen free room with free access to food and water. Protocols for the use of animals were approved by the Institutional Review Board for Animal Care at Sun Yat-Sen University.

2.2. Toxoplasma gondii isolates

A total of 18 Chinese *T. gondii* isolates were included for analysis in this study alongside a bank of reference strains (Table 1). The Chinese *T. gondii* isolates have been genotyped previously (Chen et al., 2011; Qian et al., 2012; Wang et al., 2013). Tachyzoites of *T. gondii* isolates were maintained in mice or cryopreserved in liquid nitrogen. At the late stage of animal infection, animals were euthanized. Ascites were collected by injection of ice cold D-Hanks. Two steps of differential centrifugation at

4 ºC (50 × g for 8 min and 1500 × g for 10 min) were applied to remove host macrophages and pelleted tachyzoites (Li et al., 2012; Zhao et al., 2013). Additionally a brief trypsin digestion with 0.25% trysin and 0.02% EDTA were used to digest residual macrophages, as modified from Derouin et al (1987) and Wu et al (2012). *T. gondii* tissue cysts were obtained from the brains of orally infected mice and prepared as previously described (Brinkmann et al., 1987; Letscher-Bru et al., 2003).

2.3. DNA extraction, amplification and sequencing

Genomic DNA was extracted from haploid stages (tachyzoites or bradyzoites within cysts) of *T. gondii* isolates using the AXYGEN DNA Kit (AXYGEN, USA). Primers (see Table 2) were designed for each MLST marker based on the published sequences in ToxoDB (http://www.toxodb.org/toxo/) and using Primer 3 software (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and primers for each RFLP marker were based on previous publications (Grigg et al., 2001; Khan et al., 2005; Su et al., 2006).

PCR amplification was carried out on eight introns using 5 house-keeping genes including the *UPRT*, *MIC*, *BTUB*, *HP* and *EF* genes, exons from rhoptry protein 18 (*ROP 18*) and dense granule proteins (*GRA6* and *GRA7*), and RFLP markers *SAG1*, *SAG3*, *BTUB*, *GRA6*, *c22-8*, *c29-2*, *L358*, *PK1*, *N-SAG2* and *Apico* as described previously (Khan et al., 2007; 2009; 2011). In brief, the target DNA sequences were used for the first amplification with PCR primers for each marker separately and PCR amplified products (1 µl) were then independently used for nested PCR amplification with sequencing primers. Amplicons were purified using the UNIQ-10 Spi Column PCR Product Purification Kit (Sangon, China) and subjected to commercial DNA sequencing (BGI, China) with the sequencing primers.

153 Nucleotide sequence data reported in this paper are available in the GenBank
154 databases under the accession numbers: KY618681-KY618707;
155 KY628060-KY628212.

156

157 2.4. *DNA polymorphism analysis*

158 Sequences of classical RFLP typing fragments from representive *T. gondii* strains
159 were downloaded from ToxoDB (http://www.toxodb.org/toxo/). The restriction sites
160 and resultant fragments obtained after digestion were predicted by NEB Cutter
161 (http://nc2.neb.com/NEBcutter2/index.php). The actual electrophoresis patterns were
162 obtained from the literature (Su et al., 2010) where fragments of similar molecular
163 weight or small fragments (< 40 bp) could not be distinguished.

164 The similarity coefficient (SC) between samples was calculated using the formula
165 $SC_{XY} = 2n_{XY}/(n_X + n_Y)$ (Nei and Li, 1979), where SC_{XY} represents the similarity
166 coefficient of taxa X and Y, n_{XY} is the number of common fragments for taxa X and Y,
167 and n_X or n_Y is the number of specific fragments for taxon X or Y, respectively. The
168 genetic distance (GD_{XY}) was generated using the formula $GD_{XY} = 1 - SC_{XY}$. SC and
169 GD were calculated using data from each RFLP gene and combined later both as a
170 simple average and as a weighted mean of all genes (by number of fragments). These
171 were denoted $SC_{average}$ and $GD_{average}$ and $SC_{weighted}$ and $GD_{weighted}$ respectively.

172 Nucleotide sequences were obtained from PCR-RFLP or MLST from our results or
173 taken from ToxoDB (http://www.toxodb.org/toxo/) and were compiled and aligned
174 with Clustal X 1.83 (Thompson et al., 1997) using default parameters and further
175 manual verification. A substitution model and the gamma distribution shape parameter
176 for the rate of heterogeneity among sites were determined using Modeltest 3.07
177 (Posada and Crandall, 1998) based on Hierarchical Likelihood Ratio Tests (hLRTs).

178 The p-distance model was selected for the phylogenetic analysis. Genetic
179 relationships among haplotypes were reconstructed using the neighbour-joining (NJ)
180 and maximum parsimony (MP) methods implemented in MEGA 4.0 (Saitou and Nei,
181 1987; Tamura et al., 2007) and bootstrap values were estimated with 1,000 replicate
182 searches to evaluate support for the NJ/MP trees. Pairwise genetic differentiation
183 between *T. gondii* genetic types was estimated using the fixation index (F_{ST}), and
184 statistical significances were tested with 10,000 permutations and F_{ST} calculations
185 was performed in Arlequin 3.5 (Excoffier & Lischer, 2010.).
186 In addition, genealogical relationships were examined by constructing haplotype
187 networks in Network 5.0.0.0 using the median – joining network approach (Bandelt et
188 al., 1999) with Maximum Parsimony (MP) calculation (Polzin and Daneshmand,
189 2003).
190
191 2.5. ***Pathogenicity of representative T. gondii isolates***
192 Tachyzoites of TgCtsd1, TgCtsx1, TgCtwh6, and TgCtxz3 isolates were maintained
193 separately in Swiss Webster mice. To determine the virulence phenotypes of *T. gondii*
194 isolate, Swiss Webster mice (6-8 animals per group) were inoculated (i.p.) with
195 1000,100 or 10 tachyzoites; the RH strain was used as a reference. Mortality in the
196 mice was recorded daily after inoculation.
197 Furthermore, cyst forming ability was also tested. Each rat of the Sprague-Dawley
198 (SD) strain ($n \geq 5$) was intraperitoneally infected with 10^6 tachyzoites of each isolate;
199 the Prugniaud (PRU) strain was used as a reference. The mortality of the rats was
200 recorded, and the *T. gondii* cyst numbers were counted (Brinkmann et al., 1987;
201 Letscher-Bru et al., 2003) in the brains of all surviving rats 60 days later and were
202 given as Mean ± Standard Error of Mean (SEM). The Student´s t-test was used to

compare the difference between two unpaired samples and statistical differences were designated by # ($p<0.05$) and *** ($p<0.001$), using GraphPad Prism version 5.

3. Results

3.1. Phylogenetic analyses on representative strains using RFLP and MLST methods

We first reviewed the classical RFLP method to construct a phylogenetic tree, based on the restriction enzyme fragments of both predicted and measured electrophoretic patterns (Table S1).

Not surprisingly, the two trees, containing Chinese (TgCtsx1, TgCtxz3 and TgCtsd1) representative strains and 15 reference strains (Table 1), are similar (Fig. 1A and B) in shape, containing two major clades. Clade I, the larger one could be subdivided into three obvious groups, including archetype group I (ToxoDB#10), Brazil group (ToxoDB#19 & 42), and archetype III group (ToxoDB#2). Clade II could be subdivided into two branches, including the Chinese I group (ToxoDB#9) and archetype II (ToxoDB#3) (Fig. 1A and 1B). The high degree of consistency, of inclusion of the representative Chinese strains in clade II, between the trees constructed using different methods, gives confidence that these genuinely cluster alongside the archetype II strains.

We also investigated these strains using the MLST method with the same set of genes as RFLP above. A total of 150 sequences were obtained which covered 4843 bp of the 10 genes in each isolate/strain. The phylogenetic tree presents a similar topology as the ones generated using the RFLP method, indicating the consistency of both methods (Fig 1C). However, a minor but significant difference between the two results is that the lone group of Chinese I isolates identified by RFLP is split into two

groups by MLST, one in Clade I, the other in Clade II. This raises the question as to whether at least two subgroups are present in Chinese I strains.

To better understand the relationship among these isolates/strains, we also constructed an MLST-based phylogenetic tree with a published set of gene sequences (Table 2), including eight introns and three exons. A total of 198 sequences, covering 2930 bp of the eight introns and 2401 bp of the three exons were obtained for each isolate/strain . It is interesting to see that the constructed tree (Fig.1D) has a distinct topology and a higher resolution of groups, as Clade II only contains archetype III strains and most of the branch nodes are well supported (bootstrap values > 50), than that based on the other gene set (Fig.1C). As shown in Fig 1D, the Chinese I and II isolates are clearly distinct. Furthermore, the Chinese I isolates are split into two further subgroups, one of these subgroups is closely related to archetype I, while the other is not. Therefore, we define the Chinese ToxoDB#205 isolate as Chinese II, and rename some of the Chinese I isolates (represented by TgCtsd1) that are closely related to archetype I as Chinese III.

3.2. Phylogenetic relationships and population structure of T. gondii isolates

To confirm the existence of two Chinese I subgroups, the genetic diversity of more Chinese *T. gondii* isolates were accessed, including an additional 13 isolates from the definitive host (cats) and one from chicken (TgCkjs) (Table 1). Another 19 reference isolates available from the ToxoDB database, including one Chinese reference strain (TgCatPRC2), were also included in analysis. Altogether, phylogenetic trees were generated using a total of 396 sequences from 36 strains, covering the same 2930 bp and 2401 bp regions as above. The combined tree (Fig. S1) was largely consistent with the one from a smaller data set (Fig. 1D), revealing four Chinese II isolates

(TgCtxz3,-5,-7,-8), two Chinese III isolates (TgCtsd1,TgCkjs) and the remaining 12 Chinese I isolates.

The phylogeny-based on concatenated sequencing of 8 introns (Fig. 2A) showed that the Chinese III group is closely related to the RH and GT1 strains (both archetype I, ToxoDB#10), sharing the same alleles in sequences from the *URPT*, *BTUB*, *EF1* and *HP* loci (except for a minor difference in one SNP at *MIC2*). Furthermore, we found 4 Chinese II isolates were closely related to the ME49 and PRU strains (both are type II, ToxoDB#1) and share the same alleles at the *BTUB*, *HP* and *URPT* loci. Network analysis was carried out in these sequences and this also revealed a similar haplotype topology where the Chinese I and Chinese II strains broadly clustered, while the Chinese III strains (TgCtsd1 and TgCkjs) closely clustered with the archetype I strains (RH and GT1) (Fig. 2C).

Phylogenetic analyses of the same set of 36 *T. gondii* isolates/strains were carried out based on DNA sequences from the antigen-encoding genes including *GRA6*, *GRA7* and *ROP18* (Fig. 2B). This analysis was unable to distinguish the Chinese I and II strains when mixed, in with the archetype II strains (ME49 and PRU). With one exception, TgCtbj was clustered quite separately from all other Chinese strains, due to sharing the same allele at the *ROP18* locus with CTG and VEG (type III, ToxoDB#1). Again, the suggested new group of Chinese III strains (TgCtsd1 and TgCkjs) were phylogenetically distinct from the other Chinese strains and remained closely clustered with the RH and GT1 strains (archetype I)(Fig. 2B). Network analysis using these antigen coding genes (Fig. 2D) showed, the majority of Chinese I and Chinese II strains formed a single node with only TgCtbj being represented in a neighboring node and related to the archetype III strains (CTG and VEG). On the bottom left of Fig. 2D, the newly suggested Chinese III strains (TgCtsd1 and TgCkjs) clustered with

278 the archetype I (RH and GT1) strains, far removed from the other Chinese strains.

279 In addition, genetic differentiation among the genotypes was assessed using
280 population genetic F_{ST} pairwise comparisons, using sequences of a total 26 strains that
281 fall into 7 groups, as archetype I, II, III, Brazil and Chinese I, II and III (Table 1).

282 As shown in Table 3, the F_{ST} values are generally higher when comparisons are
283 done among the known types (>0.57, archetype I, II, III and Brazil). F_{ST} values of
284 Chinese I with these reference groups were around 0.718-0.906, with statistical
285 significance (P<0.05). The F_{ST} values between the three Chinese groups were
286 0.669-0.936, which falls into the required range (P=0.001-0.070). Interestingly, the
287 F_{ST} value between Chinese III and RH was zero.

288 ***3.3 Pathogenicity of representative T. gondii isolates in Swiss Webster mice and***
289 ***Sprague-Dawley (SD) rats***

290 The virulence of Chinese isolates was also investigated. As shown in Figure 3A,
291 RH and most representative Chinese isolates showed strong virulence to mice (i.p.
292 inoculated 1000 tachyzoites) with 100% mortality within 6-11 days post-infection
293 (dpi), except TgCtwh6 which had zero mortality. A small inoculum of 100 tachyzoites
294 (i.p.) could kill mice within 6-12 days in most tested isolates except for TgCtwh6.
295 Even 10 tachyzoites from TgCtsd1 and RH, could kill mice within 7-12 days.

296 We also tested the ability of the same Chinese isolates mentioned above to form
297 cysts in a more resistant host, SD rat, with the PRU strain as a reference. All infected
298 SD rats survived until sacrifice after one and half months post-infection. The Chinese
299 III (TgCtsd1) and Chinese II (TgCtxz3) isolates formed significantly more cysts in rat
300 brain than the two Chinese I isolates (TgCtsx1 and TgCtwh6) and PRU strain (Fig. 3B:
301 TgCtsd1 *vs* Chinese I and PRU, $p < 0.001$; TgCtxz3 *vs* Chinese I, $p < 0.05$).

302

4. Discussion

With the exception of the predominant ToxoDB#9 (Chinese I), only a small number of *T. gondii* isolates, including ToxoDB#10 (Type I), ToxoDB#1(Type II), ToxoDB#204 and ToxoDB#205 (designated as Chinese II in this paper) have been reported in China (Shen & Wang, 2014) and, furthermore, showed very limited genetic diversity. Fewer genotype differences were found in these isolates despite being from different provinces of China (Fig.2) (Chen et al., 2011; Wang et al., 2013). This suggests that there might be a common origin and recent expansion through this country, possibly due to limited recombination. It has been demonstrated that *T. gondii* is a haploid protozoan and genetic recombination can only occur when the cat is simultaneously infected with at least two different strains (Alan and Pitt, 2012). Therefore, this greatly restricts the occurrence of genetic recombination in *T. gondii*. In China, the infected cat population is less than 20% (Chen et al., 2011; Qian et al., 2012; Wang et al., 2013), to some extent, limiting the opportunities for genetic recombination during sexual reproduction within (Boothroyd and Grigg, 2002). Therefore, this might be the reason why we observed limited genotypes in the Chinese strains.

To our surprise, inconsistent with previous results based on RFLP (Chen et al., 2011), MLST analysis found that TgCtsd1 and TgCkjs comprise an entirely new and unique lineage of genotypes which we have designated here as Chinese III (Fig. 2). Interestingly, these two isolates were found to be genetically closely related to archetypal type I strains (such as RH and GT1), with only a few base pair differences. This indicates that Chinese III might be a mutant or a hybrid of archetype I.

Phylogenetic analyses of *T. gondii* isolates based on introns which are selectively neutral loci are thought to provide an unbiased estimation, while secretory antigens,

328 known to be functionally related to virulence and infectivity in the host (GRA6,
329 GRA7 and ROP18) (Shwab et al., 2016; Zhang et al., 2016), are also considered
330 capable of revealing great genetic diversity because they are frequently under
331 selective pressure (Miller et al., 2004; Pena et al., 2008; Sundar et al., 2008). However,
332 on the basis of these genes, phylogenetic analysis found that the avirulent (TgCtwh6)
333 and virulent Chinese I (e.g. TgCtsx1) strains were clustered together (Fig. 2A & B).
334 Additionally, the analysis could not distinguish Chinese I from Chinese II strains (Fig.
335 2B). Additionally some Chinese I strains (as defined by neutral markers, RFLP and
336 MLST) may be eventually placed in a novel clade (e.g. TgCtbj in Fig. 2B). We
337 propose that neutral markers (RFLP, MLST, introns) accurately predict phylogenetic
338 relationships between strains but may not be useful for prediction of host virulence or
339 other such functional attributes. Conversely, we propose that phylogenetic
340 relationships as generated by functional genes, which are probably under selection,
341 are not useful for strain prediction of related strains but probably reflect functional
342 mutants. Such genes may be useful in predicting other functional characteristics, such
343 as virulence, but the phylogenetic identity of the cyst forming Chinese III and the
344 non-cyst forming RH suggests that such relationships may be complex. Further
345 research is required to investigate the relationship between genotype, phylogenetic
346 relationships and virulence phenotypes. As we found that the unusual strain (TgCtbj in
347 Fig. 2B) was fixed in the network plot (Fig. 2D), we consider that the network plot
348 could be more consistent, more fault-tolerant and more useful for future investigations
349 on *T. gondii*.
350 The mouse model is a classical assay for presenting the cyst-forming capability of *T.*
351 *gondii* strains (Wagner et al., 2016; Mahmoudv et al., 2016). In consideration of the
352 great differences between mouse models and human clinical symptoms, the rat model

has been suggested as a better analogy (Darcy and Zenner, 1993; Santoro et al., 1987). In our study, most Chinese strains are virulent and kill mice within two weeks before the formation of cysts can occur (Fig. 3A). Therefore, for analysis of cyst-forming ability, the rat model system, based on the innate resistance of rats to *T. gondii* (Gao et al., 2015; Li et al., 2012), is much more useful. In comparison to the archetype II PRU strain, which itself is capable of cyst formation in certain rat strains, the Chinese strains of TgCtsd1 and TgCtxz3 produced over 200 times more cysts. This suggests that the Chinese strains are not only virulent to mice but also to rats. Although, we do not have any evidence for the status of cyst distribution and prevalence in wild rats in China, our results support the notion that rats could be significant reservoirs for *T. gondii* in nature (Dubey and Frenkel, 1998).

In conclusion, the Chinese strains of *T. gondii* isolated from different geographic regions show limited genetic diversity based on RFLP and sequencing methods. A novel genotype in China which we called Chinese III was identified with classical RFLP, MLST and neutral markers. However, these markers cannot be used to predict virulence or ability to form cysts, while the functional genes are not efficient markers to analyze the genetic phenotype. Therefore, more isolates, especially those from humans, are needed to enable a clearer understanding of the population structure and pathogenicity of *T. gondii* in China.

Acknowledgments

The authors' laboratories are supported by the grants from the National Natural Science Foundation of China (#31472058, #31401975, #31402029), Guangdong Natural Science Funds for Distinguished Young Scholar (2016A030306048), Guangzhou Science Technology and Innovation Commission (201506010011), the

Natural Science Foundation of Guangdong Province (2014A030313164). We thank Dr Yan-Zi Wen for her critical comments and other members in our laboratories who provided significant help during the work.

References

Anderson, T.J., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, L., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J.,Velez, I.D., Brockman, A.H., Nosten, F., Ferreira, M.U., Day, K.P., 2000. Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. Mol. Biol. Evol. 17, 1467-1482.

Ajzenberg, D., Cogne, N., Paris, L., Bessieres, M.H., Thulliez, P., Filisetti, D., Pelloux, H., Marty, P., Darde, M.L., 2002. Genotype of 86 *Toxoplasma gondii* isolates associated with human congenital toxoplasmosis, and correlation with clinical findings. J. Infect. Dis. 186, 684-689.

Alan Gunn, Pitt, S.J. (eds.), 2012. Parasitology. Wiley-Blackwell, UK.

Bandelt, H. J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48.

Baldursson, S., Karanis, P., 2011. Waterborne transmission of protozoan parasites: review of worldwide outbreaks - an update 2004-2010. Water. Res. 45, 6603-6614.

Boothroyd, J.C., Grigg, M.E., 2002. Population biology of *Toxoplasma gondii* and its relevance to human infection: do different strains cause different disease? Curr. Opin. Microbiol. 5, 438-442.

Brinkmann, H., Martinez, P., Quigley, F., Martin, W., Cerff, R., 1987. Endosymbiotic origin and codon bias of the nuclear gene for chloroplast

403 glyceraldehyde-3-phosphate dehydrogenase from maize. J. MOI. Evol. 26,
404 320-328.
405 Cameron, M.L., Levy, P., Nutman, T., Vanamala, C.R., Narayanan, P.R., Rajan, T.V.,
406 1988. Use of restriction fragment length polymorphisms (RFLPs) to distinguish
407 between nematodes of pathogenic significance. Parasitology. 96, 381-390.
408 Chen, Z.W., Gao, J.M., Huo, X.X., Wang, L., Yu, L., Halm-Lai, F., Xu, Y.H., Song,
409 W.J., Hide, G., Shen, J.L., Lun, Z.R., 2011. Genotyping of *Toxoplasma gondii*
410 isolates from cats in different geographic regions of China. Vet. Parasitol. 183,
411 166-170.
412 Darcy, F., Zenner, L., 1993. Experimental models of toxoplasmosis. Res. Immunol.
413 144, 16–23.
414 Darde, M.L., Bouteille, B., Pestre-Alexandre, M., 1992. Isoenzyme analysis of 35
415 *Toxoplasma gondii* isolates and the biological and epidemiological implications.
416 J. Parasitol. 78, 786-794.
417 Dubey, J.P., Beattie, C. P., 1988. Toxoplasmosis of Animals and Man. In: Beattie,
418 C.P. (Ed)., CRC Press, Boca Raton, Florida, pp. 220.
419 Dubey, J.P., Frenkel, J.K., 1998. Toxoplasmosis of rats: a review, with considerations
420 of their value as an animal model and their possible role in epidemiology. Vet.
421 Parasitol. 77, 1-32.
422 Dubey, J.P., Graham, D.H., Blackston, C.R., Lehmann, T., Gennari, S.M., Ragozo,
423 A.M., Nishi, S.M., Shen, S.K., Kwok, O.C., Hill, D.E., Thulliez, P., 2002.
424 Biological and genetic characterisation of *Toxoplasma gondii* isolates from
425 chickens (Gallus domesticus) from Sao Paulo, Brazil: unexpected findings. Int. J.
426 Parasitol. 32, 99-105.
427 Dubey, J.P., Zhu, X.Q., Sundar, N., Zhang, H., Kwok, O.C., Su, C., 2007. Genetic and

428 biologic characterization of *Toxoplasma gondii* isolates of cats from China. Vet.
429 Parasitol. 145, 352-356.
430 Dubey, J.P., 2010. Toxoplasmosis of Animals and Humans. In, Boca. Raton. Florida:
431 CRC Press, pp. 313.
432 Excoffier, L., Lischer, H.E.L., 2010. Arlequin suite ver 3.5: a new series of programs
433 to perform population genetics analyses under Linux and Windows. Mol. Ecol.
434 Resources. 10, 564-567.
435 Frenkel, J.K., Dubey, J.P., Miller, N.L., 1970. *Toxoplasma gondii* in cats: fecal stages
436 identified as coccidian oocysts. Science. 167, 893-896.
437 Gallas-Lindemann, C., Sotiriadou, I., Mahmoodi, M.R., Karanis, P., 2013. Detection
438 of *Toxoplasma gondii* oocysts in different water resources by Loop Mediated
439 Isothermal Amplification (LAMP). Acta. Trop. 125, 231-236.
440 Gao, J.M., Yi, S.Q., Wu, M.S., Geng, G.Q., Shen, J.L., Lu, F.L., Hide, G., Lai, D.H.,
441 Lun, Z.R., 2015. Investigation of infectivity of neonates and adults from different
442 rat strains to *Toxoplasma gondii* Prugniaud shows both variation which
443 correlates with iNOS and Arginase-1 activity and increased susceptibility of
444 neonates to infection. Exp. Parasitol. 149, 47-53.
445 Grigg M.E., Ganatra J., Boothroyd J.C., Margolis T.P., 2001. Unusual abundance of
446 atypical strains associated with human ocular toxoplasmosis. J. Infect. Dis. 184,
447 633-639.
448 Hide. G., Tait. A., 2009. Molecular epidemiology of African sleeping sickness.
449 Parasitology. 136(12), 1491-500.
450 Howe, D.K., Sibley, L.D., 1995. *Toxoplasma gondii* comprises three clonal lineages:
451 correlation of parasite genotype with human disease. J. Infect. Dis. 172,
452 1561-1566.

453 Karanis, P., Kourenti, C., Smith, H., 2007. Waterborne transmission of protozoan
454 parasites: a worldwide review of outbreaks and lessons learnt. J. Water. Health. 5,
455 1-38.
456 Karanis, P., Aldeyarbi, H.M., Mirhashemi, M.E., Khalil, K.M., 2013. The impact of
457 the waterborne transmission of *Toxoplasma gondii* and analysis efforts for water
458 detection: an overview and update. Environ. Sci. Pollut. Res. Int. 20, 86-99.
459 Khan, A., Jordan, C., Muccioli, C., Vallochi, A.L., Rizzo, L.V., Belfort, R.J., Vitor,
460 R.W., Silveira, C., Sibley, L.D., 2006. Genetic divergence of *Toxoplasma gondii*
461 strains associated with ocular toxoplasmosis, Brazil. Emerg. Infect. Dis. 12,
462 942-949.
463 Khan, A., Dubey, J.P., Su, C., Ajioka, J.W., Rosenthal, B.M., Sibley, L.D., 2011.
464 Genetic analyses of atypical *Toxoplasma gondii* strains reveal a fourth clonal
465 lineage in North America. Int. J. Parasitol. 41, 645-655.
466 Khan, A., Fux, B., Su, C., Dubey, J.P., Darde, M.L., Ajioka, J.W., Rosenthal, B.M.,
467 Sibley, L.D., 2007. Recent transcontinental sweep of *Toxoplasma gondii* driven
468 by a single monomorphic chromosome. Proc. Natl. Acad. Sci. (USA) 104,
469 14872-14877.
470 Khan, A., Taylor, S., Ajioka, J.W., Rosenthal, B.M., Sibley, L.D., 2009. Selection at a
471 single locus leads to widespread expansion of *Toxoplasma gondii* lineages that
472 are virulent in mice. PLoS. Genet. 5, e1000404.
473 Khan, A., Taylor, S., Su C., Mackey, A.J., Boyle, J., Cole, R., Glover, D., Tang,
474 K., Paulsen, I.T., Berriman, M., Boothroyd, J.C., Pfefferkorn, E.R.,Dubey,
475 J.P., Ajioka, J.W., Roos, D.S., Wootton, J.C., Sibley, L.D., 2005. Composite
476 genome map and recombination parameters derived from three archetypal
477 lineages of *Toxoplasma gondii*. Nucleic. Acids. Res. 33, 2980-2992.

478 Letscher-Bru, V., Herbrecht, R., 2003. Caspofungin: the first representative of a new
479 antifungal class. J. Animicrob. Chemother. 51, 513-521.
480 Lehmann, T., Marcet, P.L., Graham, D.H., Dahl, E.R., Dubey, J.P., 2006.
481 Globalization and the population structure of *Toxoplasma gondii*. Proc. Natl.
482 Acad. Sci. (USA) 103, 11423-11428.
483 Li, Z., Zhao, Z.J., Zhu, X.Q., Ren, Q.S., Nie, F.F., Gao, J.M., Gao, X.J., Yang, T.B.,
484 Zhou, W.L., Shen, J.L., Wang. Y., Lu, F.L., Chen, X.G., Hide, G., Ayala, F.J.,
485 Lun, Z.R., 2012. Differences in iNOS and arginase expression and activity in the
486 macrophages of rats are responsible for the resistance against *T. gondii* infection.
487 PLoS One. 7(4), e35834.
488 Lunde, M.N., Jacobs, L., 1983. Antigenic differences between endozoites and
489 cystozoites of *Toxoplasma gondii*. J. Parasitol. 69, 806-808.
490 Mahmoudv, H., Ziaali, N., Ghazvini, H., Shojaee, S., Keshavarz, H., Esmaeilpour,
491 K., Sheibani, V., 2016. *Toxoplasma gondii* infection promotes
492 neuroinflammation through cytokine networks and induced hyperalgesia in
493 BALB/c mice. Inflammation. 39, 405-412.
494 Mercier, A., Devillard, S., Ngoubangoye, B., Bonnabau, H., Bañuls, A.-L., Durand, P.,
495 Salle, B., Ajzenberg, D., Dardé, M.-L. 2010. Additional haplogroups of
496 *Toxoplasma gondii* out of Africa: population structure and mouse-virulence of
497 strains from Gabon. PLoS. Neglected. Tropical. Diseases. 4, e876.
498 Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R.,
499 Jessup, D.A., Boothroyd, J.C., Brownstein, D., Conrad, P.A., 2004. An unusual
500 genotype of *Toxoplasma gondii* is common in California sea otters (Enhydra
501 lutris nereis) and is a cause of mortality. Int. J. Parasitol. 34, 275–284.
502 Montoya, J.G., Liesenfeld, O., 2004. Toxoplasmosis. Lancet. 363, 1965-1976.

503 Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms
504 of restriction endonucleases. Proc. Natl. Acad. Sci. (USA) 76, 5269–5273.
505 Pena, H.F., Gennari, S.M., Dubey, J.P., Su, C., 2008. Population structure and
506 mouse-virulence of *Toxoplasma gondii* in Brazil. Int. J. Parasitol. 38, 561-569.
507 Polzin, T., Daneshmand, S. V., 2003. On Steiner trees and minimum spanning trees in
508 hypergraphs. Operations Research Letters. 31,12-20.
509 Posada, D, Crandall, KA., 1998. Modeltest: testing the model of DNA substitution.
510 Bioinformatics, 14, 817-818.
511 Qian, W., Wang, H., Su, C., Shan, D., Cui, X., Yang, N., Lv, C., Liu, Q., 2012.
512 Isolation and characterization of *Toxoplasma gondii* strains from stray cats
513 revealed a single genotype in Beijing, China. Vet. Parasitol. 187, 408-413.
514 Rajendran, C., Su, C. and Dubey, J. P., 2012. Molecular genotyping of *Toxoplasma*
515 *gondii* from Central and South America revealed high diversity within and
516 between populations. Infect. Genet. Evol. 12, 359–368.
517 Saitou, N., NEI, M., 1987. The neighbor-joining method: a new method for
518 reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.
519 Santoro, F., Auriault, C., Leite, P., Darcy, F., Capron, A., 1987. Infection of the
520 athymic rat by *Toxoplasma gondii*. C. R. Acad. Sci. III 304, 297–300.
521 Severson, D.W., Thathy, V., Mori, A., Zhang, Y., Christensen, B.M., 1995.
522 Restriction fragment length polymorphism mapping of quantitative trait loci for
523 malaria parasite susceptibility in the mosquito Aedes aegypti. Genetics. 139,
524 1711-1717.
525 Sibley, L.D., Boothroyd, J.C., 1992. Virulent strains of *Toxoplasma gondii* comprise a
526 single clonal lineage. Nature. 359, 82-85.
527 Sibley, L.D., Khan, A., Ajioka, J.W., Rosenthal, B.M., 2009. Genetic diversity of

Toxoplasma gondii in animals and humans. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 364, 2749-2761.

Shen, J.L., Wang, L.L., 2014. Genotypes and main effectors of *Toxoplasma gondii* and their pathogenic mechanisms. Chin. J. Parasitol. Parasit. Dis. 33(6), 429-435.

Shwab, E.K., Jiang, T., Pena, H.F., Gennari, S.M., Dubey, J.P., Su, C., 2016. The ROP18 and ROP5 gene allele types are highly predictive of virulence in mice across globally distributed strains of *Toxoplasma gondii*. Int. J. Parasitol. 46(2), 141-146.

Su, C., Shwab, E.K., Zhou, P., Zhu, X.Q., Dubey, J.P., 2010. Moving towards an integrated approach to molecular detection and identification of *Toxoplasma gondii*. Parasitology. 137, 1–11.

Su, C., Khan, A., Zhou, P., Majumdar, D., Ajzenberg, D., Darde, M.L., Zhu, X.Q., Ajioka, J.W., Rosenthal, B.M., Dubey, J.P., Sibley, L.D., 2012. Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. Proc. Natl. Acad. Sci. (USA) 109, 5844-5849.

Su C., Zhang X., Dubey J.P., 2006. Genotyping of *Toxoplasma gondii* by multilocus PCR-RFLP markers: A high resolution and simple method for identification of parasites. Int. J. Parasitol. 36, 841-848.

Sundar, N., Cole, R.A., Thomas, N.J., Majumdar, D., Dubey, J.P., Su, C., 2008. Genetic diversity among sea otter isolates of *Toxoplasma gondii*. Vet. Parasitol. 151, 125–132.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic. Acids. Res. 25, 4876–4882.

553 Tamura, K., Dudley, J., Nei, M., Kumar, S., 2007. MEGA4: Molecular evolutionary
554 genetics analysis (MEGA) software version 4.0. Mol. Biol. Evol. 24, 1596–1599.
555 Velmurugan, B., Selvanayagam, M., Cengiz, E.I., Uysal, E. 2008. Levels of
556 transaminases, alkaline phosphatase, and protein in tissues of Clarias gariepienus
557 fingerlings exposed to sublethal concentrations of cadmium chloride. Environ.
558 Toxicol. 23, 672–678.
559 Wagner A., Schabussova I., Drinic M., Akgün J., Loupal G., Kundi M., Joachim
560 A., Wiedermann U., 2016. Oocyst-derived extract of *Toxoplasma gondii* serves
561 as potent immunomodulator in a mouse model of birch pollen allergy. PLoS.
562 One. 11(5), e0155081.
563 Wang, L., Chen, H., Liu, D., Huo, X., Gao, J., Song, X.R., Xu, X., Huang, K., Liu, W.,
564 Wang, Y., Lu, F., Lun, Z., Luo, Q., Wang, X., Shen, J., 2013. Genot ypes and
565 Mouse Virulence of *Toxoplasma gondii* Isolates from Animals and Humans in
566 China. PLoS. One. e8.
567 Widmer, G., Feng, X., Tanriverdi, S., 2004. Genotyping of *Cryptosporidium parvum*
568 with microsatellite markers. Methods. Mol. Biol. 268, 177-187.
569 Zhang, W., Li, L., Xia, N., Zhou, Y., Fang, R., He, L., Hu, M., Shen, B., Zhao, J.
570 2016. Analysis of the virulence determination mechanisms in a local *Toxoplasma*
571 strain (T.gHB1) isolated from central China. Parasitol. Res. 115(10), 3807-15.
572 Zhao, G.W., Shen, B., Xie, Q., Xu, L.X., Yan, R.F., Song, X.K., Hassan, I.A., Li,
573 X.R., 2012. Detection of *Toxoplasma gondii* in free-range chickens in China
574 based on circulating antigens and antibodies. Vet. Parasitol. 185, 72-77.
575 Zhao, Z.J., Zhang, J., Wei, J., Li, Z., Wang, T., Yi, S.Q., Shen, J.L., Yang, T.B., Hide,
576 G., Lun, Z.R. 2013. Lower expression of inducible nitric oxide synthase and

577 higher expression of arginase in rat alveolar macrophages are linked to their
578 susceptibility to *Toxoplasma gondii* infection. PLoS One. 8(5), e63650.
579 Zhou, P., Zhang, H., Lin, R.Q., Zhang, D.L., Song, H.Q., Su, C., Zhu, X.Q., 2009.
580 Genetic characterization of *Toxoplasma gondii* isolates from China. Parasitol. Int.
581 58, 193-195.
582

Figure legends:

Fig. 1. Phylogenetic trees of 15/18 representative isolates constructed using different methods. (A) Phylogenetic analysis based on the number of fragments of a weighted mean of all restriction fragment length Polymorphism (RFLP) sequences. (B) Phylogenetic analysis based on the number of fragments of a simple mean of all RFLP sequences. (C) Phylogenetic analysis based on the MLST method using the same set of genes used for RFLP. Due to the lack of information from PRU. ARI and CTG strains, these strains were not included. (D) Phylogenetic tree of these representative isolates based on the classical MLST method. Bootstrap values were estimated with 1,000 replicate searches to evaluate support for the neighbour-joining and maximum parsimony trees. *T. gondii* strains: RH and GT1, ToxoDB#10 (Type I), from USA; PRU and ME49, ToxoDB#1 (Type II), from France and USA, respectively; CTG and VEG, ToxoDB#2 (Type III), from USA; TgCtsx1, TgCtsd1 (ToxoDB#9) and TgCtxz3 (ToxoDB#205), from China; CAST (ToxoDB#28) and ARI (ToxoDB#5), from USA; TgCatBr5 and TgCatBr9, ToxoDB#19 and ToxoDB#42, from Brazil; MAS (ToxoDB#17), FOU, VAND (ToxoDB#60) and RUB (ToxoDB#98), from France; TgCtCo5, ToxoDB#61, from Colombia.

Fig. 2. Phylogenetic analysis based on intron (house-keeping genes) or exon (antigen-encoding genes) sequences of a greater range of Chinese *T. gondii* isolates with MLST. (A) Genetic diversity was estimated by sequencing eight introns from five unlinked genes that encode house-keeping functions. (B) Combined phylogenetic tree for dense granule proteins 6 and 7 (GRA6 and GRA7) and rhoptry protein 18 (ROP 18). Median-joining network for *Toxoplasma gondii* haplotypes of concatenated eight intron genes (C) and concatenated GRA6, GRA7 and ROP18 genes (D). The sizes of the circles are proportional to haplotype frequency and the colours represent the group to which the samples belong; small hollow circles represent missing haplotypes. *T. gondii* strains: RH and GT1, ToxoDB#10 (Type I), from USA; PRU

and ME49, ToxoDB#1 (Type II), from France and USA, respectively; CTG and VEG, ToxoDB#2 (Type III), from USA; TgCtgd1/-2, TgCtsx1/-2, TgCtsd1/-3,-4,-5, TgCtwh3/-5,-6, TgCtBj, TgCatPRC2 (ToxoDB#9), TgCkjs and TgCtxz3/-5,-7,-8 (ToxoDB#205), from China; CAST (ToxoDB#28), ARI (ToxoDB#5) and P89, from USA; TgCatBr5 (ToxoDB#19) and TgCatBr9 (ToxoDB#42), from Brazil; MAS (ToxoDB#17), FOU, VAND (ToxoDB#60) and RUB (ToxoDB#98), from France; TgCtCo5, ToxoDB#61, from Colombia; COUG from Canada; and GAB2-2007-GAL-DOM2, unknown.

Fig. 3. Survival rates of Swiss Webster mice infected with representative Chinese *Toxoplasma gondii* isolates (TgCtsd1, TgCtsx1, TgCtwh6, and TgCtxz3) (A) and the cyst-forming capability of these strains in Sprague-Dawley rats (B).