



**HAL**  
open science

## Phylogeography of *Toxoplasma gondii* points to a South American origin.

Emilie Bertranpetit, Thibaut Jombart, Emmanuel Paradis, Hilda Pena, Jitender Dubey, Chunlei Su, Aurélien Mercier, Sébastien Devillard, Daniel Ajzenberg

### ► To cite this version:

Emilie Bertranpetit, Thibaut Jombart, Emmanuel Paradis, Hilda Pena, Jitender Dubey, et al.. Phylogeography of *Toxoplasma gondii* points to a South American origin.. *Infection, Genetics and Evolution*, 2016, 48, pp.150-155. 10.1016/j.meegid.2016.12.020 . hal-01502248

**HAL Id: hal-01502248**

**<https://unilim.hal.science/hal-01502248v1>**

Submitted on 5 Dec 2023

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

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

1  
2 **Phylogeography of *Toxoplasma gondii* Points to a South American Origin**  
3  
4

5 Emilie Bertranpetit<sup>a</sup>, Thibaut Jombart<sup>b</sup>, Emmanuel Paradis<sup>c</sup>, Hilda Pena<sup>d</sup>, Jitender  
6 Dubey<sup>e</sup>, Chunlei Su<sup>f</sup>, Aurélien Mercier<sup>a</sup>, Sébastien Devillard<sup>g,1,2</sup>, Daniel Ajzenberg<sup>h,1,2</sup>  
7

8 <sup>a</sup>*INSERM UMR\_S 1094, Neuroépidémiologie Tropicale, Laboratoire de Parasitologie-*  
9 *Mycologie, Faculté de Médecine, Université de Limoges, Limoges, 87025, France*

10 <sup>b</sup>*MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease*  
11 *Epidemiology, School of Public Health, Imperial College London, United Kingdom*

12 <sup>c</sup>*Institut des Sciences de l'Évolution, Université Montpellier/CNRS/IRD/EPHE, Place*  
13 *Eugène Bataillon – CC 065, 34095 Montpellier cédex 05, France*

14 <sup>d</sup>*Departamento de Medicina Veterinária Preventiva e Saúde Animal, Faculdade de*  
15 *Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, Brazil*

16 <sup>e</sup>*United States Department of Agriculture, Agricultural Research Service, Beltsville*  
17 *Agricultural Research Center, Animal Parasitic Diseases Laboratory, Building 1001,*  
18 *Beltsville, Maryland, 20705-2350, USA*

19 <sup>f</sup>*Department of Microbiology, University of Tennessee, Knoxville, Tennessee, 37996-*  
20 *0845, USA*

21 <sup>g</sup>*Université de Lyon, F-69000, Lyon ; Université Lyon 1 ; CNRS, UMR 5558, Laboratoire*  
22 *de Biométrie et Biologie Evolutive, F-69622, Villeurbanne, France*

23 <sup>h</sup>*Centre National de Référence (CNR) Toxoplasmose / Toxoplasma Biological Resource*  
24 *Center (BRC), Centre Hospitalier-Universitaire Dupuytren, Limoges, 87042, France and*  
25 *INSERM UMR\_S 1094, Neuroépidémiologie Tropicale, Laboratoire de Parasitologie-*  
26 *Mycologie, Faculté de Médecine, Université de Limoges, Limoges, 87025, France*  
27

28 <sup>1</sup>Correspondence and requests for materials should be addressed to S.D (email:  
29 [sebastien.devillard@univ-lyon1.fr](mailto:sebastien.devillard@univ-lyon1.fr) Tel: +33472448111) and D.A. (email: [ajz@unilim.fr](mailto:ajz@unilim.fr)  
30 Tel: +33555056160)  
31

32 <sup>2</sup> Authors contributed equally  
33  
34  
35  
36  
37  
38  
39  
40

41 **Abstract**

42 *Toxoplasma gondii*, a protozoan found ubiquitously in mammals and birds, is the  
43 etiologic agent of toxoplasmosis, a disease causing substantial Public Health burden  
44 worldwide, including about 200,000 new cases of congenital toxoplasmosis each year.  
45 Clinical severity has been shown to vary across geographical regions, with South  
46 America exhibiting the highest burden. Unfortunately, the drivers of these  
47 heterogeneities are still poorly understood, and the geographical origin and historical  
48 spread of the pathogen worldwide are currently uncertain. A worldwide sample of 168 *T.*  
49 *gondii* isolates gathered in 13 populations was sequenced for five fragments of genes  
50 (140 single nucleotide polymorphisms from 3,153 bp per isolate). Phylogeny based on  
51 Maximum likelihood methods with estimation of the time to the most recent common  
52 ancestor (TMRCA) and geostatistical analyses were performed for inferring the putative  
53 origin of *T. gondii*. We show that extant strains of the pathogen likely evolved from a  
54 South American ancestor, around 1.5 million years ago, and reconstruct the subsequent  
55 spread of the pathogen worldwide. This emergence is much more recent than the  
56 appearance of ancestral *T. gondii*, believed to have taken place about 11 My ago, and  
57 follows the arrival of felids in this part of the world. We posit that an ancestral lineage of  
58 *T. gondii* likely arrived in South America with felids and that the evolution of oral  
59 infectivity through carnivorousism and the radiation of felids in this region enabled a new  
60 strain to outcompete the ancestral lineage and undergo a pandemic radiation.

61 **Keywords:** *Toxoplasma gondii*, phylogeography, Maximum likelihood phylogeny, time  
62 to the most recent common ancestor (TMRCA), genetic diversity

63

64 **1. Introduction**

65 Toxoplasmosis imposes a substantial disease burden across the world. Serological  
66 studies demonstrate its presence in virtually every country, with seroprevalence  
67 exceeding 60% in some parts of South America, Africa, and South-East Asia ([Pappas et](#)  
68 [al., 2009](#)). While asymptomatic in most patients, toxoplasmosis is a major cause of  
69 uveitis in immunocompetent patients and a potentially life-threatening illness in  
70 immunocompromised patients and fetuses ([Montoya and Liesenfeld, 2004](#)). Congenital  
71 toxoplasmosis alone has a yearly global incidence of about 200,000 cases, causing a  
72 burden exceeding 1 million disability-adjusted life year (DALYs) ([Torgerson and](#)  
73 [Mastroiacovo, 2013](#)). The overall disease burden attributable to all forms of  
74 toxoplasmosis is certainly far greater, and highest in South America where ocular  
75 toxoplasmosis is unusually frequent and severe ([Glasner et al., 1992; de-la-Torre et al.,](#)  
76 [2008; Gilbert et al., 2008; Torgerson and Mastroiacovo, 2013](#)).

77

78 The extent to which host parasite genetics, host immune status, and exposure rate  
79 contribute to the increased severity of toxoplasmosis in South America is unclear, but  
80 differences in the genetic makeup of *T. gondii* strains are likely to play a major role  
81 ([Khan et al., 2006; Gilbert et al., 2008](#)). In North America, Europe, Africa, and Asia, the  
82 population structure of *T. gondii* is dominated by a few prevalent clonal strains, whereas  
83 much greater genetic diversity is seen in tropical South America where the populations  
84 lack sign of recent genetic bottleneck and clonal structure seen in the other parts of the  
85 world ([Shwab et al., 2014; Lorenzi et al., 2016](#)).

86 Unfortunately, the drivers of the pathogen's genetic diversity are still poorly understood,  
87 and the origin of extant lineages of *T. gondii* remains controversial. Recent work  
88 suggested a potential South American origin, while another study advocated that co-  
89 migration with felids led to the divergence of South American strains from pre-existing  
90 North American ones ([Lehmann et al., 2006](#); [Khan et al., 2007](#)). The estimation of the  
91 time to the most recent common ancestor (TMRCA) of extant lineages is also disputed,  
92 with estimates ranging from 150,000 to  $10^7$  years ([Morrison, 2005](#); [Khan et al., 2007](#)).  
93 In the present study, using a large collection of genetic sequences of *T. gondii* sampled  
94 worldwide, we reconstructed the phylogeography of *T. gondii* as a basis to address the  
95 controversial questions regarding the evolution of this parasite and its geographical  
96 origin.

97

## 98 **2. Materials and Methods**

99

### 100 **2.1. Collection of *T. gondii* strains and selection of markers**

101 A total of 168 *T. gondii* strains from 13 populations collected worldwide in North  
102 America, South America, the Caribbean, Europe, Asia, and Africa, were used in this  
103 study (Supplementary information and S1 Table). Our collection was specifically  
104 designed to ensure extensive geographic coverage of strains clustered in true  
105 populations, and includes a large number of strains from Africa and China, which were  
106 so far underrepresented or absent in the previous studies ([Lehmann et al., 2006](#); [Khan  
107 et al., 2007](#)). A preliminary genetic analysis with 15 microsatellite markers (Ajzenberg et  
108 al., 2010) was performed to exclude clones of strains sampled in the same area. Each  
109 isolate was sequenced in both directions for five markers (*GRA6*, *GRA7*, *SAG3*, *UPRT1*

110 and *UPRT7*) that had the highest polymorphic rates after a preliminary analysis of  
111 genetic polymorphism of 30 fragments of genes retrieved from GenBank and ToxoDB  
112 (Supplementary information and S2 Table).

113  
114 **2.2. Phylogenetic analysis.**

115 *Hammondia hammondi* is the most closely parasite related to *T. gondii* and was used  
116 as an outgroup in phylogenetic analyses. Sequences of the *H. hammondi* strain H.H.34  
117 corresponding to *T. gondii* *GRA6*, *GRA7*, *SAG3*, *UPRT1* and *UPRT7* sequences were  
118 retrieved from GenBank and ToxoDB, and aligned with MUSCLE ([Edgar, 2004](#)). We  
119 used *ape* and *pegas* R packages to extract haplotypes and build phylogenetic trees with  
120 three distance-based methods: NJ, BIONJ, and FastME which were used as starting  
121 trees for the ML analyses. Maximum likelihood phylogenetic analyses were performed  
122 with the R package *phangorn* using four partitions of the sequence data by crossing two  
123 criteria: exons vs. introns on one side, and *GRA6*, *GRA7* and *SAG3* vs. *UPRT1* and  
124 *UPRT7* on the other (Supplementary information). A GTR +  $\Gamma$  + I model was used with  
125 parameters that could vary among data partitions. The different model fits performed  
126 with *phangorn* were compared with AIC. The three trees obtained with the distance-  
127 based methods were all tested as initial trees.

128  
129 **2.3. Geostatistical analyses.** We used the same geostatistics approach for inferring  
130 the putative origin of *T. gondii* as previously used for uncovering the origins of *P.*

131 *falciparum* ([Tanabe et al., 2010, 2013a, 2013b](#); [Mita and Jombart, 2015](#)). We implemented  
132 this approach in the R package *geoGraph* (<http://thibautjombart/geoGraph>), in which we  
133 provided extensive documentation replicating the analyses described below using  
134 publicly available data ([Cann et al., 2002](#)). The method implemented in *geoGraph* relies  
135 on the idea that migration events result in successive bottlenecks which reduce the  
136 genetic diversity within populations as they are located further away from the origin  
137 ([Tanabe et al., 2010](#)). Accordingly, we expect to observe a negative correlation between  
138 within-population diversity and the distance from the origin. While in practice the true  
139 origin is often unknown, one can infer the most plausible origin by assessing this  
140 relationship for a number of candidate origins, and retaining the origin yielding the  
141 strongest negative correlation. This method requires two types of distances, genetic and  
142 spatial, to be computed. Here, the genetic diversity was mostly structured by varying  
143 frequencies of a small number of haplotypes within populations (S3 Table and Fig 1).  
144 Therefore, we used haplotype richness (i.e. number of distinct haplotypes) as a  
145 measure of diversity within populations. Spatial distances through landmasses were  
146 computed using *geoGraph*. The package models movements on the surface of the  
147 Earth using a spherical, pseudo-regular grid with approximately 40,000 nodes. Each  
148 node possesses an 'habitat' attribute, here used to distinguish landmasses from seas.  
149 Shortest path between locations were computed using the dijkstra algorithm ([Jungnickel,](#)  
150 [2013](#)) implemented in the R package RBGL ([Edmonds et al., 2006](#); [Carey et al., 2011](#)). To

151 define candidate origins, 1,800 combinations of regularly spaced longitudes and  
152 latitudes were used to cover the globe, which resulted in 433 non-redundant locations  
153 on landmasses on the grid used by *geoGraph*. For each location, the shortest path  
154 through each sampled population was identified, and the corresponding distance  
155 computed in kilometers. These distances were then used to assess patterns of  
156 decrease of genetic diversity from the putative origin using simple linear regression. The  
157 most likely origin was inferred as the location which yielded the most negative  
158 correlation between geographic distances and haplotype diversity within populations.

159

160 **2.4. Time to the most recent common ancestor (TMRCA).** In order to estimate the  
161 time to the most recent common ancestor (TMRCA) of *T. gondii*, we used two different  
162 approaches: a simple molecular dating method based on the divergence with *H.*  
163 *hammondi* which is estimated to be around 11 My, and a coalescent approach using the  
164 expectation of TMRCA which is equal to twice the effective population size ( $N_e$ ). Both  
165 approaches need an estimate of the mutation rate ( $\mu$ ), and the second one also needs  
166 an estimate of the population parameter  $\theta$  ( $= 2 \mu N_e$ ). We estimated  $\mu$  for the non-  
167 coding introns of *UPRT1* and *UPRT7*. We did two neutrality tests: D's Tajima and the  
168 R2 test. We calculated a standard-error of this estimate using the variance of ratio of  
169 two random variables, here the number of mutations per site and twice the time of  
170 divergence. The former was estimated with a Tamura-Nei distance with its associated



171 variance, and the latter was 11 My with an arbitrary  $sd = 1$  My. The population  
172 parameter  $\theta$  was estimated in two ways: with a Markov chain Monte Carlo (MCMC)  
173 approach as implemented in the R package *coalescentMCMC*, and with the nucleotide  
174 diversity ( $\pi$ ) calculated with *pegas*. Both ways calculate the standard-error of the  
175 estimate of  $\theta$ . In the end, three estimates of TMRCA were obtained with their respective  
176 95% confidence interval (CI).

177

### 178 **3. Results**

#### 179 **3.1. Genetic diversity.**

180 Sequences of the five markers represented a total of 3,153 bp per isolate, including 140  
181 variable sites. Without taking into account sites with gaps, 26, 30, 27, 32, and 25 SNPs  
182 were identified in the *GRA6* (607 bp), *GRA7* (677 bp), *SAG3* (638 bp), *UPRT1* (574 bp),  
183 and *UPRT7* (657 bp) genes, respectively (S4 Table and S1 Fig). Strains from the  
184 Caribbean, Guiana shield, Northeast Brazil, and Southeast Brazil showed more  
185 sequence polymorphism with 65, 90, 59, and 65 SNPs, respectively (S3 Table).  
186 Polymorphism was lower in the African and European populations with a number of  
187 SNPs ranging from one to 39. Polymorphism was intermediate in the Asian and North  
188 American populations with 41, 60, 47, and 51 SNPs in Turkey, China, Minnesota, and  
189 Pennsylvania, respectively.

190 Data concatenation revealed 60 haplotypes. The number of haplotypes ranged from  
191 four in Europe to 32 in South America. Of the 32 genotypes in the 44 strains from South  
192 America, 29 were endemic in South America whereas only three were common in other  
193 populations (one in the Caribbean, one in Africa, and one both in Asia and Africa). In  
194 contrast, of the four genotypes in the 34 strains from Europe, three were common in  
195 other populations (one in Asia and Africa, one in North America, Asia and Africa, and  
196 one in the Caribbean, North America, Asia and Africa) and the unique genotype differed  
197 by only one SNP from the one common to North America, Asia, and Africa (S5 Table).  
198 The higher values of haplotype diversity, estimated from concatenation of the five  
199 markers by the number of haplotypes divided by the number of isolates, were observed  
200 in South America and the Caribbean, whereas the lower values were observed in  
201 Europe, Asia, and Africa (S3 Table). Haplotype diversity was intermediate in North  
202 America. Overall the highest genetic diversity was found in South-America.

203

### 204 **3.2. Phylogeny of *T. gondii* strains**

205 Preliminary analyses of genetic diversity (Supplementary information, S6 Table and S2  
206 and S3 Figs.) revealed different mutation patterns and rates of evolution between  
207 coding and non-coding segments, and for the two *UPRT* genes compared to the others.  
208 Accordingly, we defined four partitions of the sequence data crossing these two  
209 categories, and reconstructed separate phylogenies by maximum likelihood (ML) to  
210 investigate potential phylogenetic incongruence ([Som, 2015](#)). Statistical tests and

211 examination of model selection criterion (AIC) revealed the existence of distinct  
212 topologies (Fig. 1), suggesting that these sequence partitions have undergone different  
213 evolutionary histories and selective pressures. Interestingly, only South American  
214 strains were consistently placed at a basal position (close to the root) in all topologies.  
215 To investigate this pattern further and identify the common evolutionary history of these  
216 genes, a consensus topology was inferred from the four ML-partitioned topologies (Fig.  
217 2). This new tree supported the more ancestral status of South American isolates, with  
218 35 out of 44 samples located at the root of the tree. However, as expected in the  
219 presence of conflicting phylogenetic signal, this tree was only partially resolved, and  
220 strains from other locations (China: 12 samples; Africa: 7 samples) also belonged to the  
221 large basal multifurcation.

222

### 223 **3.3. On the geographic origin of *T. gondii***

224 As a complementary analysis, we used a geostatistical approach previously employed  
225 for identifying the origin of *Plasmodium falciparum*, the main etiologic agent of malaria  
226 ([Tanabe et al., 2010, 2013b; Mita and Jombart, 2015](#)). This method identifies likely  
227 geographic origins as the locations from which patterns of decrease in genetic diversity,  
228 expected to be observed due to repeated migration and founder effects, are most  
229 consistent ([Tanabe et al., 2010](#)). Because of the low level of polymorphism observed in  
230 the sequenced genes and the highly clonal nature of *T. gondii*, haplotype richness was  
231 used as a measure of genetic diversity within populations (S3 Table). Testing a large  
232 number of hypothetical origins across the world, this approach identified South America,  
233 and more specifically Colombia as the most likely origin ( $r=-0.81$ ,  $p=0.9 \times 10^{-4}$ , Fig. 3).  
234 While substantial uncertainty remains about the exact location, this analysis brings

235 strong support to a South American origin for *T. gondii* suggested by the phylogenetic  
236 approach. Our results further suggest that *T. gondii* initially spread through the  
237 Americas and then colonized Asia and Europe via the Bering Strait, before entering  
238 Africa through two different migration routes (Fig. 3).

239

### 240 **3.4. Time to the most recent common ancestor ( $T_{\text{MRCA}}$ ) of *T. gondii*.**

241 To understand the processes which may have led to a pandemic radiation of *T. gondii*  
242 from South America, the emergence of this ancestral, highly successful lineage has to  
243 be dated. To this end, we derived estimates of time to the most recent common  
244 ancestor (MRCA) of the extant strains using standard molecular approaches. While  
245 confidence intervals indicated substantial uncertainty, overall results suggest that the  
246 MRCA of *T. gondii* appeared around 1.5 My ago (Table 1). This emergence is much  
247 more recent than the existence of *T. gondii* itself, estimated to have diverged from its  
248 closest ancestor *Hammondia hammondi* some 11 My ago ([Morrison et al., 2004](#)).

249

## 250 **4. Discussion**

251

252 A major event occurred in the evolutionary history of *T. gondii* which led to a selective  
253 sweep about 1.5 My ago. We hypothesized that an ancestral form of *T. gondii* was  
254 introduced in South America through the migration of Felidae after the emergence of the  
255 Isthmus of Panama about 2 to 3 My ago, at the end of Pliocene ([O'Brien et al., 2008](#)). It is  
256 believed that Felidae species quickly expanded after their arrival and diversified into the  
257 "ocelot" lineage in South America. Interestingly the Muridae, potential intermediate  
258 hosts for *T. gondii*, also showed extensive diversification in South America with the

259 appearance of several genera ([Webb, 2006](#)) around the same time. As previously  
260 suggested ([Webb, 2006](#)), we posit that this expansion, diversification and mixing of host  
261 populations certainly resulted in similar processes in their parasites and favoured the  
262 accumulation of genetic diversity in *T. gondii*, which eventually led to a selective sweep  
263 by a highly successful mutant lineage.

264

265 The selective pressures underlying this selective sweep can be debated. Previous work  
266 attributed this radiation to the emergence of transmission through carnivorism (i.e., oral  
267 infectivity of tissue cysts) between intermediate hosts in clonal strains 10,000 years ago  
268 ([Su et al., 2003](#)). However, oral infectivity was shown to be also a trait of many South  
269 American strains ([Came et al., 2002](#); [Khan et al., 2007](#)). Because South American strains  
270 were the first to diverge from the MRCA, it is likely that transmission by carnivorism  
271 evolved earlier than the apparition of clonal lineages. This trait conferred a better  
272 transmission of the current form of *T. gondii* which likely outcompeted the ancestral form  
273 arriving in South America with Felidae. The transmission of *T. gondii* between its  
274 different hosts would allow some genotypes to migrate to North America, then to go  
275 through the Bering Strait to colonize Asia, Europe and Africa. The current population  
276 structure of *T. gondii* with a predominance of a few successful clonal strains in Africa,  
277 Asia and Europe, is likely to be the consequence of the recent expansion of the  
278 domestic cat, an Old World species until the sixteenth century, that tremendously  
279 amplified a specific subset of pre-adapted genotypes ([Müller and Howard, 2016](#)).

280

281 In addition to being the likely origin of modern *T. gondii* strains, South America also  
282 suffers from the highest burden of toxoplasmosis. Prevalence, incidence, and severity of

283 acquired and congenital ocular toxoplasmosis (OT) in some areas of Brazil, Colombia,  
284 and Argentina are considerably higher than anywhere else, which makes OT a genuine  
285 public health issue in South America ([Glasner et al., 1992](#); [de-la-Torre et al., 2008](#); [Gilbert et](#)  
286 [al., 2008](#); [Rudzinski et al., 2016](#)). Because South America is also the hotspot of *T. gondii*'s  
287 genetic diversity, it has been hypothesized that severe forms of toxoplasmosis may be  
288 the consequence of poor adaptation of the human host to the unusual diversity of  
289 strains in this part of the world, resulting in impaired immune response and, thus, a  
290 more aggressive disease ([Khan et al., 2006](#); [Gilbert et al., 2008](#); [Demar et al., 2012](#);  
291 [de-la-Torre et al., 2013](#); [Rudzinski et al., 2016](#)). The societal and economic costs of  
292 care for symptomatic cases of congenital toxoplasmosis can be considerable but the  
293 cost-effectiveness of national routine prenatal screening and treatment program are still  
294 debated ([Wallon et al., 1999](#); [Jones et al., 2014](#)). There is a need for randomized placebo-  
295 controlled trials to help determine the effectiveness of these interventions.

296

## 297 **5. Conclusion**

298 Our reconstruction of *T. gondii*'s phylogeography provides a new framework for  
299 understanding patterns of genetic diversity in sampled populations of the parasite, and  
300 for predicting diversity in unsampled locations. Because genetic diversity seems to  
301 impact directly the severity of the disease, our results can be used as a basis for  
302 explaining geographic heterogeneities in disease burden, and identifying priority targets  
303 for potential future interventions.

304

305

306 **ACKNOWLEDGEMENTS**

307 We are thankful to github (<http://github.com/>) and travis (<http://travis-ci.org/>) for  
308 providing great resources for the development of the R package *geoGraph*. We are  
309 indebted to Endrias Zewdu Gebremedhin, Hüseyin Can, Mert Döşkaya, Yüksel Gürüz,  
310 Susana Sousa, Lénaïg Halos, Min Li, François Peyron, and the members of the French  
311 National Reference Center for toxoplasmosis (Nicole Desbois, Saadia Azi, Hélène Yera,  
312 Magalie Demar, Bernard Carme, Stéphane Simon, Denis Blanchet, Rachida Boukhari,  
313 Isabelle Villena, Dominique Aubert) for their contribution to the collection of 168 strains  
314 that were included in the analysis.

315

316 This research did not receive any specific grant from funding agencies in the public,  
317 commercial, or not-for-profit sectors.

318

319 The sequences reported in this study have been deposited in the GenBank database  
320 (accession n° KU598987-KU599154 for *GRA6*, accession n°KU599155-KU599322 for  
321 *GRA7*, accession n°KU599323-KU599490 for *SAG3*, accession n°KU599491-  
322 KU599658 for *UPRT1* and accession n°KU599659-KU599826 for *UPRT7*).

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337 **References**

- 338 [Ajzenberg, D., Daniel, A., Frédéric, C., Aurélien, M., Philippe, V., Marie-Laure, D., 2010. Genotyping of \*Toxoplasma gondii\* Isolates with 15 Microsatellite Markers in a Single Multiplex PCR Assay. J. Clin. Microbiol. 48, 4641–4645.](#)
- 339
- 340
- 341
- 342 [Cann, H.M., de Toma, C., Cazes, L., Legrand, M.-F., Morel, V., Piouffre, L., Bodmer, J.,](#)
- 343 [Bodmer, W.F., Bonne-Tamir, B., Cambon-Thomsen, A., Chen, Z., Chu, J., Carcassi, C.,](#)
- 344 [Contu, L., Du, R., Excoffier, L., Ferrara, G.B., Friedlaender, J.S., Groot, H., Gurwitz, D.,](#)
- 345 [Jenkins, T., Herrera, R.J., Huang, X., Kidd, J., Kidd, K.K., Langaney, A., Lin, A.A., Mehdi,](#)
- 346 [S.Q., Parham, P., Piazza, A., Pistillo, M.P., Qian, Y., Shu, Q., Xu, J., Zhu, S., Weber, J.L.,](#)
- 347 [Greely, H.T., Feldman, M.W., Thomas, G., Dausset, J., Cavalli-Sforza, L.L., 2002. A human](#)
- 348 [genome diversity cell line panel. Science 296, 261–262.](#)
- 349
- 350 [Carey, V., Long, L., Gentleman, R., 2011. RBGL: An interface to the BOOST graph library. R](#)
- 351 [package version 1, 124.](#)
- 352
- 353 [Carme, B., Bissuel, F., Ajzenberg, D., Bouyne, R., Aznar, C., Demar, M., Bichat, S., Louvel, D.,](#)
- 354 [Bourbigot, A.M., Peneau, C., Neron, P., Darde, M.L., 2002. Severe Acquired](#)
- 355 [Toxoplasmosis in Immunocompetent Adult Patients in French Guiana. J. Clin. Microbiol. 40,](#)
- 356 [4037–4044.](#)
- 357
- 358 [de-la-Torre, A., López-Castillo, C.A., Gómez-Marín, J.E., 2008. Incidence and clinical](#)
- 359 [characteristics in a Colombian cohort of ocular toxoplasmosis. Eye 23, 1090–1093.](#)
- 360
- 361 [de-la-Torre, A., Sauer, A., Pfaff, A.W., Bourcier, T., Brunet, J., Speeg-Schatz, C., Ballonzoli, L.,](#)
- 362 [Villard, O., Ajzenberg, D., Sundar, N., Grigg, M.E., Gomez-Marin, J.E., Candolfi, E., 2013.](#)
- 363 [Severe South American ocular toxoplasmosis is associated with decreased Ifn- \$\gamma\$ /Il-17a and](#)
- 364 [increased Il-6/Il-13 intraocular levels. PLoS Negl. Trop. Dis. 7, e2541.](#)
- 365
- 366 [Demar, M., Hommel, D., Djossou, F., Peneau, C., Boukhari, R., Louvel, D., A.-M., B., Nasser,](#)
- 367 [V., Ajzenberg, D., M.-L., D., Carme, B., 2012. Acute toxoplasmoses in immunocompetent](#)
- 368 [patients hospitalized in an intensive care unit in French Guiana. Clin. Microbiol. Infect. 18,](#)
- 369 [E221–E231.](#)
- 370
- 371 [Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high](#)
- 372 [throughput. Nucleic Acids Res. 32, 1792–1797.](#)
- 373
- 374 [Edmonds, N., Breuer, A., Gregor, D., Lumsdaine, A., 2006. Single-source shortest paths with](#)
- 375 [the parallel boost graph library. The Ninth DIMACS Implementation Challenge: The](#)
- 376 [Shortest Path Problem, Piscataway, NJ 219–248.](#)
- 377
- 378 [Gilbert, R.E., Freeman, K., Lago, E.G., Bahia-Oliveira, L.M.G., Tan, H.K., Wallon, M., Buffolano,](#)
- 379 [W., Stanford, M.R., Petersen, E., European Multicentre Study on Congenital](#)
- 380 [Toxoplasmosis \(EMSCOT\), 2008. Ocular sequelae of congenital toxoplasmosis in Brazil](#)
- 381 [compared with Europe. PLoS Negl. Trop. Dis. 2, e277.](#)
- 382
- 383 [Glasner, P.D., Silveira, C., Kruszon-Moran, D., Martins, M.C., Burnier Júnior, M., Silveira, S.,](#)
- 384 [Camargo, M.E., Nussenblatt, R.B., Kaslow, R.A., Belfort Júnior, R., 1992. An unusually](#)
- 385 [high prevalence of ocular toxoplasmosis in southern Brazil. Am. J. Ophthalmol. 114, 136–](#)



386 [144.](#)  
387 [Jones, J.L., Parise, M.E., Fiore, A.E., 2014. Neglected parasitic infections in the United States:](#)  
388 [toxoplasmosis. Am. J. Trop. Med. Hyg. 90, 794–799.](#)  
389  
390 [Jungnickel, D., 2013. Graphs, Networks and Algorithms: Algorithms and Computation in](#)  
391 [Mathematics . Springer Berlin Heidelberg.](#)  
392  
393 [Khan, A., Fux, B., Su, C., Dubey, J.P., Darde, M.L., Ajioka, J.W., Rosenthal, B.M., Sibley, L.D.,](#)  
394 [2007. Recent transcontinental sweep of \*Toxoplasma gondii\* driven by a single](#)  
395 [monomorphic chromosome. Proc. Natl. Acad. Sci. U. S. A. 104, 14872–14877.](#)  
396  
397 [Khan, A., Jordan, C., Muccioli, C., Vallochi, A.L., Rizzo, L.V., Belfort, R., Jr, Vitor, R.W.A.,](#)  
398 [Silveira, C., Sibley, L.D., 2006. Genetic divergence of \*Toxoplasma gondii\* strains associated](#)  
399 [with ocular toxoplasmosis, Brazil. Emerg. Infect. Dis. 12, 942–949.](#)  
400  
401 [Lehmann, T., Marcet, P.L., Graham, D.H., Dahl, E.R., Dubey, J.P., 2006. Globalization and the](#)  
402 [population structure of \*Toxoplasma gondii\*. Proc. Natl. Acad. Sci. U. S. A. 103, 11423–](#)  
403 [11428.](#)  
404  
405 [Lorenzi, H., Khan, A., Behnke, M.S., Namasivayam, S., Swapna, L.S., Hadjithomas, M.,](#)  
406 [Karamycheva, S., Pinney, D., Brunk, B.P., Ajioka, J.W., Ajzenberg, D., Boothroyd, J.C.,](#)  
407 [Boyle, J.P., Dardé, M.L., Diaz-Miranda, M.A., Dubey, J.P., Fritz, H.M., Gennari, S.M.,](#)  
408 [Gregory, B.D., Kim, K., Saeij, J.P.J., Su, C., White, M.W., Zhu, X.-Q., Howe, D.K.,](#)  
409 [Rosenthal, B.M., Grigg, M.E., Parkinson, J., Liu, L., Kissinger, J.C., Roos, D.S., Sibley,](#)  
410 [L.D., 2016. Local admixture of amplified and diversified secreted pathogenesis](#)  
411 [determinants shapes mosaic \*Toxoplasma gondii\* genomes. Nat. Commun. 7, 10147.](#)  
412  
413 [Mita, T., Jombart, T., 2015. Patterns and dynamics of genetic diversity in Plasmodium](#)  
414 [falciparum: what past human migrations tell us about malaria. Parasitol. Int. 64, 238–243.](#)  
415  
416 [Montoya, J.G., Liesenfeld, O., 2004. Toxoplasmosis. Lancet 363, 1965–1976.](#)  
417  
418 [Morrison, D.A., 2005. How old are the extant lineages of \*Toxoplasma gondii\*? Parassitologia 47,](#)  
419 [205–214.](#)  
420  
421 [Morrison, D.A., Set, B., Per, T., Ulrich, W., Jörg, K., Mattsson, J.G., 2004. The current status of](#)  
422 [the small subunit rRNA phylogeny of the coccidia \(Sporozoa\). Int. J. Parasitol. 34, 501–514.](#)  
423  
424 [Müller, U.B., Howard, J.C., 2016. The impact of \*Toxoplasma gondii\* on the mammalian genome.](#)  
425 [Curr. Opin. Microbiol. 32, 19–25.](#)  
426  
427 [O'Brien, S.J., Johnson, W., Driscoll, C., Pontius, J., Pecon-Slattery, J., Menotti-Raymond, M.,](#)  
428 [2008. State of cat genomics. Trends Genet. 24, 268–279.](#)  
429  
430 [Pappas, G., Roussos, N., Falagas, M.E., 2009. Toxoplasmosis snapshots: global status of](#)  
431 [\*Toxoplasma gondii\* seroprevalence and implications for pregnancy and congenital](#)  
432 [toxoplasmosis. Int. J. Parasitol. 39, 1385–1394.](#)  
433  
434 [Rudzinski, M., Khoury, M., Couto, C., Ajzenberg, D., 2016. Reactivation of Ocular](#)  
435 [Toxoplasmosis in Non-Hispanic Persons, Misiones Province, Argentina. Emerg. Infect. Dis.](#)  
436 [22, 912–913.](#)

437  
438 [Shwab, E.K., Zhu, X.-Q., Majumdar, D., Pena, H.F.J., Gennari, S.M., Dubey, J.P., Su, C., 2014. Geographical patterns of \*Toxoplasma gondii\* genetic diversity revealed by multilocus PCR-RFLP genotyping. \*Parasitology\* 141, 453–461.](#)  
439  
440  
441  
442 [Som, A., 2015. Causes, consequences and solutions of phylogenetic incongruence. \*Brief. Bioinform.\* 16, 536–548.](#)  
443  
444  
445 [Su, C., Evans, D., Cole, R.H., Kissinger, J.C., Ajioka, J.W., Sibley, L.D., 2003. Recent Expansion of \*Toxoplasma\* Through Enhanced Oral Transmission. \*Science\* 299, 414–416.](#)  
446  
447  
448 [Tanabe, K., Jombart, T., Horibe, S., Palacpac, N.M.Q., Honma, H., Tachibana, S.-I., Nakamura, M., Horii, T., Kishino, H., Mita, T., 2013a. \*Plasmodium falciparum\* mitochondrial genetic diversity exhibits isolation-by-distance patterns supporting a sub-Saharan African origin. \*Mitochondrion\* 13, 630–636.](#)  
449  
450  
451  
452  
453 [Tanabe, K., Mita, T., Jombart, T., Eriksson, A., Horibe, S., Palacpac, N., Ranford-Cartwright, L., Sawai, H., Sakihama, N., Ohmae, H., Nakamura, M., Ferreira, M.U., Escalante, A.A., Prugnolle, F., Björkman, A., Fämert, A., Kaneko, A., Horii, T., Manica, A., Kishino, H., Balloux, F., 2010. \*Plasmodium falciparum\* accompanied the human expansion out of Africa. \*Curr. Biol.\* 20, 1283–1289.](#)  
454  
455  
456  
457  
458  
459 [Tanabe, K., Mita, T., Palacpac, N.M.Q., Arisue, N., Tougan, T., Kawai, S., Jombart, T., Kobayashi, F., Horii, T., 2013b. Within-population genetic diversity of \*Plasmodium falciparum\* vaccine candidate antigens reveals geographic distance from a Central sub-Saharan African origin. \*Vaccine\* 31, 1334–1339.](#)  
460  
461  
462  
463  
464 [Torgerson, P.R., Mastroiacovo, P., 2013. The global burden of congenital toxoplasmosis: a systematic review. \*Bull. World Health Organ\* 91, 501–508..](#)  
465  
466  
467 [Wallon, M., Liou, C., Garner, P., Peyron, F., 1999. Congenital toxoplasmosis: systematic review of evidence of efficacy of treatment in pregnancy. \*BMJ\* 318, 1511–1514.](#)  
468  
469  
470 [Webb, S.D., 2006. The Great American Biotic Interchange: patterns and processes 1. \*Ann. Mo. Bot. Gard.\* 93, 245–257.](#)  
471  
472  
473  
474  
475  
476  
477  
478

479

480 **Table 1. Estimates of the time to the most recent common ancestor (TMRCA) of**  
481 ***Toxoplasma gondii* with three different methods on the introns of *UPRT* genes.**

482

Method	TMRCA (Ma)	[95% CI]
Molecular dating	1.59	[0.00–3.46]
Coalescent	1.26	[0.94–1.57]
Nuc. div. ( $\pi$ )	1.20	[0.00–2.44]

483

484 Nuc. div.: nucleotide diversity.

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500