

# Genetic diversity, clonality and sexuality in *Toxoplasma gondii*<sup>☆</sup>

D. Ajzenberg<sup>a</sup>, A.L. Bañuls<sup>b</sup>, C. Su<sup>c</sup>, A. Dumètre<sup>a</sup>, M. Demar<sup>d</sup>, B. Carne<sup>d</sup>, M.L. Dardé<sup>a,\*</sup>

<sup>a</sup>Laboratoire de Parasitologie-Mycologie, EA3174, Faculté de Médecine, 2 rue du Dr Marcland, 87031 Limoges, France

<sup>b</sup>GEMI, UMR CNRS-IRD 2724, Centre IRD, 34395 Montpellier, France

<sup>c</sup>Department of Molecular Microbiology, Washington University, St. Louis, MO 63110, USA

<sup>d</sup>Parasitologie-Mycologie, EA 3593 UFR de Médecine (Université Antilles Guyane) et Centre Hospitalier de Cayenne, BP 6006, F97306 Cayenne, French Guiana

Received 26 March 2004; received in revised form 28 June 2004; accepted 28 June 2004

## Abstract

The majority of *Toxoplasma gondii* strains from a variety of human and animal sources have been grouped into three highly clonal but closely related lineages. The low occurrence of nucleotide differences among the three predominant lineages and their unusual dimorphic allelic composition suggest that they have arisen from a recent common ancestry. Less than 1% of the previously studied strains contain unique genotypes and high divergence of DNA sequence, and therefore are considered 'exotic' or 'atypical' strains. The seemingly low genetic diversity in *T. gondii* may have been underestimated because most parasite strains in previous studies were collected from human patients and domestic animals in North America and Europe. To investigate the genetic diversity of *T. gondii*, we analysed parasite strains isolated from remote geographical regions by multilocus microsatellite sequencing and phylogenetic analysis. The genetic diversity indices, the molecular analysis of microsatellite genotypes and the constructed phylogram considered together suggest that the global *T. gondii* population is highly diversified and not characteristic of a clonal organism. The most parsimonious hypothesis is that *T. gondii* presents a complex population structure with a mix of clonal and sexual propagation as a function of the environmental conditions. The comparison between domestic strains data on one hand and wild strains data on the other hand is in favour of more frequent sexual recombinations in wild environment even though *Toxoplasma* subpopulation in human and domestic animals is largely clonal.

© 2004 Australian Society for Parasitology Inc. Published by Elsevier Ltd. All rights reserved.

**Keywords:** *Toxoplasma gondii*; Atypical strains; Microsatellites; Genetic diversity; Clonality; Sexuality

## 1. Introduction

*Toxoplasma gondii* is a protozoan parasite that infects virtually all warm-blooded animal species worldwide. It has a two-host life cycle with felids as definitive hosts in which the sexual replication of the parasite occurs, while mammals and birds are intermediate hosts in which the asexual replication occurs. Even though there is no geographic boundary and host specificity, molecular genotyping studies

have showed a fundamental clonal population structure with three clonal lineages, namely types I, II and III (Tibayrenc et al., 1991; Dardé et al., 1992; Sibley and Boothroyd, 1992; Howe and Sibley, 1995; Dardé, 1996; Ajzenberg et al., 2002a). Difference at DNA sequence level among the predominant clonal lineages is less than 2% (Grigg et al., 2001a). There are several explanations for the existence of clonal population structure in *T. gondii*. First, this parasite is able to transmit among intermediate hosts through carnivorousism and scavenging, bypassing sexual recombination events in definitive host cats (Howe and Sibley, 1995; Su et al., 2003). Second, many macrogametes of the parasite remain unfertilised but are capable of forming oocysts in the small intestine of cats by parthenogenesis (Ferguson, 2002). Third, cats simultaneously infected with different strains of *T. gondii* are likely to be very rare events in nature, therefore there is a temporal barrier for recombination to

<sup>☆</sup> Nucleotide sequences reported in this paper are available in the GenBank™, EMBL and DDBJ databases under the accession numbers AY572648, AY572692, AY572736, AY572780, AY572562–AY572604, AY572605–AY572647, AY572649–AY572691, AY572693–AY572735, AY572737–AY572779.

\* Corresponding author. Tel.: +33-5-55-05-6160; fax: +33-5-55-05-6177.

E-mail address: [darde@unilim.fr](mailto:darde@unilim.fr) (M.L. Dardé).

occur. Though a largely clonal population structure predominates, genetic diversity does exist in *T. gondii* (Howe and Sibley, 1995; Dardé, 1996; Bossi et al., 1998; Dardé et al., 1998; Cole et al., 2000; Lehmann et al., 2000; Grigg et al., 2001b; Carme et al., 2002b; Ghosn et al., 2003; Miller et al., 2004). Moreover, recombination phenomena have been described as occasional events (Howe and Sibley, 1995). To better estimate the genetic diversity and sexual recombination of *T. gondii*, we analysed a substantial set of strains of wild and/or tropical origin, mainly from French Guiana, by multilocus microsatellite sequencing and phylogenetic analysis.

## 2. Materials and methods

### 2.1. *Toxoplasma* and *Hammondia* strains

A total of 43 *T. gondii* strains were selected for this study (Table 1). Most strains were selected from a pool of 245 strains mainly collected by the BRC ToxoBS group and categorised based on length polymorphism of five microsatellite (MS) markers. Eighty-eight percent (216/245) of these strains belong to the three classical type I, II and III lineages, with a clear predominance of type II strains (188/245, 77%). For sequence analysis, 14 strains were arbitrarily selected to represent the three major lineages, whereas the other 29 strains were selected based on their unique MS profiles. Of these 43 strains, 28 were previously typed by the classical isoenzymatic or PCR-RFLP analysis, and the remaining 15 are new strains (see Table 1).

The 14 strains representing the major lineages are three type I strains (BK, ATIH and PIL), four type II strains (Me49, BEV, DEG and REN-2002-VAL) and seven type III strains (NED, CEP, LGE97-3, M7741, OPA, LEG-NJA and LPN-2002-SEE). Among the 29 strains with unique profiles, six strains (GPHT, RMS-2001-MAU, IPP-2002-URB, TONT, BOU and MAS) originated from France, seven strains (ARI, B41, B73, P80, SOU, T61 and WTD-1) from USA, nine strains (RUB, VAND, GUY-2001-DOS, GUY-2002-KOE, GUY-2002-MAT, IPP-2002-BAT, GUY-2003-BAS, GUY-2003-MEL and GUY-2003-ADA) from French Guiana, one strain (ENVL-2002-MAC) from Barbados, one strain (CASTELLS) from Uruguay, five strains most likely from tropical areas although isolated in France (PSP-2003-ERO from Guadeloupe island, TOU-2002-ALI from Reunion island, PSP2003-KOM from Cameroon, WIK and GANGI from Africa). Among the seven strains from USA, one was isolated from a pig (P80) whereas the six other strains, provided by L.D. Sibley as a DNA sample, were selected from the paper of Howe and Sibley (1995) because they exhibited a recombinant or non-classical PCR-RFLP pattern (SOU and ARI) and/or an unusual host origin (bears for B41 and B73, turkey for T61 and deer for WTD1).

One *Hammondia hammondi* strain (Kansas H-H 34) was kindly gifted to our laboratory by Dr J.K. Frenkel in 1989. It was used as outgroup for phylogenetic analysis.

### 2.2. Genetic markers

Five MS loci were chosen for sequencing (Table 2). *TUB2*, *TgM-A* and *W35* were previously used for strain typing by Ajzenberg et al. (2002a). Two new microsatellite markers (*B17* and *B18*) were developed by the method described previously (Ajzenberg et al., 2002a). Marker *TUB2* was mapped to chromosome IX (Sibley and Boothroyd, 1992). Markers *B17*, *B18*, *TgM-A* and *W35* were mapped to *T. gondii* chromosomes XII, VII, X and II, respectively, by genetic mapping and linkage analysis (Chunlei Su and David Sibley, personal communication).

### 2.3. DNA extraction and sequencing

*Toxoplasma gondii* DNA was extracted from haploid stages (tachyzoites or bradyzoites) present in brains or ascitic fluids of infected mice using the QIAamp<sup>®</sup> DNA MiniKit (Qiagen, Courtaboeuf, France). *Hammondia hammondi* DNA was extracted from sporulated oocysts (containing haploid sporozoites) using a method previously described for *Eimeria* species (Zhao et al., 2001). Primers (Table 2) were designed for each marker on the basis of the published sequence in GenBank using primer 3 software ([http://www.broad.mit.edu/cgi-bin/primer/primer3\\_www.cgi](http://www.broad.mit.edu/cgi-bin/primer/primer3_www.cgi)) to amplify short sequences (between 113 and 259 bp) comprising the microsatellite and its flanking regions.

The amplification reaction mixture consisted of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, 0.2 mM of dATP, dCTP and dGTP (Roche diagnostics, Meylan, France), 0.4 mM of dUTP (Roche diagnostics, Meylan, France), 15 pmol of each primer, 5% (vol/vol) dimethyl sulfoxide (except for *B17* and *TgM-A*: 6%), 1.5 U of Taq DNA polymerase (Amersham Biosciences, Orsay, France), and 6 µl of DNA in a 50 µl reaction volume. PCR conditions, carried out in a GeneAmp<sup>®</sup> PCRSystem 2700 thermocycler (Applied Biosystems, Courtaboeuf, France), were denaturation at 94 °C for 5 min, followed by 35 cycles consisting of 94 °C for 30 s, 54 °C for 30 s, and 72 °C for 30 s. The last extension step was at 72 °C for 10 min. Confirmation of DNA amplification was done on a 2% agarose gel stained with ethidium bromide. PCR products were purified using the CONCERT<sup>™</sup> Rapid PCR Purification System (GibcoBRL, Life technologies, Cergy-Pontoise, France) and then directly sequenced (i.e. without cloning) using the BigDye<sup>®</sup> Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems, Courtaboeuf, France) in a GeneAmp<sup>®</sup> PCRSystem 2700 thermocycler (Applied Biosystems, Courtaboeuf, France). Sequencing was done in both directions when there was a problem of interpretation in one direction. The fragments were purified by isopropanol precipitation. The sequencing gel was run on the ABI

Table 1

Description of the 43 *Toxoplasma gondii* strains with geographical and host origin, classical typing by PCR-RFLP or isoenzyme analysis, SNP and microsatellite (MS) types obtained by combination of SNP and MS alleles, respectively. SNP and MS alleles were defined in Fig. 1

Strain	Host	Geographical origin	Classical typing <sup>a</sup>	Reference	Alleles by SNP analysis						Alleles by MS analysis					
					Type <sup>b</sup>	TUB2	W35	TgM-A	B18	B17	Type <sup>b</sup>	TUB2	W35	TgM-A	B18	B17
BK	Human	Netherlands	I (Z1)	Dardé (1996)	I	1.2.3	1	1.2	1.2.3	1	I	1	1	1	1.3	1
ATIH	Human	Uruguay	I (Z1)	Ajzenberg et al. (2002a)	I	1.2.3	1	1.2	1.2.3	1	I	1	1	1	1.3	1
PIL	Human	France (Paris)	I (Z1)	Dardé (1996)	I	1.2.3	1	1.2	1.2.3	1	I	1	1	1	1.3	1
GPHT	Human	France (Paris)	I (Z1)	Dardé (1996)	I/III	1.2.3	1	3	1.2.3	1	I/III	1	1	3	1.3	1
WIK	Human	Africa	I (Z1)	Ajzenberg et al. (2002a)	I/III	1.2.3	1	3	1.2.3	1	I/III	1	1	3	1.3	1
PSP-2003-KOM	Human	Cameroon	ND		I/III	1.2.3	1	3	1.2.3	1	I/III*	1	<b>6</b>	3	1.3	1
PSP-2003-ERO	Human	Guadeloupe island	ND		I/III	1.2.3	3	3	1.2.3	1	I/III*	1	3	3	<b>4</b>	1
ME49	Sheep	USA	II (Z2)	Dardé et al. (1992)	II	1.2.3	2	1.2	1.2.3	2.3	II	2.3	2	2	2	2.3
BEV	Rabbit	England	II (Z2)	Dardé et al. (1992)	II	1.2.3	2	1.2	1.2.3	2.3	II	2.3	2	2	2	2.3
REN-2002-VAL	Human	France (Rennes)	ND		II	1.2.3	2	1.2	1.2.3	2.3	II	2.3	2	2	2	2.3
DEG	Human	France (Limoges)	II (Z4)	Dardé et al. (1992)	II*	1.2.3	2	1.2	<b>4</b>	2.3	II	2.3	2	2	2	2.3
BOU	Human	France (Limoges)	II (Z2)	Dardé et al. (1992)	II	1.2.3	2	1.2	1.2.3	2.3	II*	2.3	<b>5</b>	2	2	2.3
GANGI	Human	Africa	II (Z2)	Ajzenberg et al. (2002a)	II	1.2.3	2	1.2	1.2.3	2.3	II/I	2.3	2	1	2	2.3
B73	Bear	USA	II/III	Howe and Sibley (1995)	II	1.2.3	2	1.2	1.2.3	2.3	II/I or II/III	2.3	2	2	1.3	2.3
ARI	Human	USA	II	Howe and Sibley (1995)	II*	1.2.3	2	1.2	1.2.3	<b>4</b>	II/I	2.3	2	1	2	2.3
WTD-1	Deer	USA	II	Howe and Sibley (1995)	II/I	1.2.3	2	1.2	1.2.3	1/2.3	II/I	2.3	2	1	1.3	2.3
B41	Bear	USA	II	Howe and Sibley (1995)	II/I	1.2.3	2	1.2	1.2.3	1/2.3	II*	2.3	2	2	<b>4</b>	2.3
RMS-2001-MAU	Human	France (Reims)	ND		II/I	1.2.3	2	1.2	1.2.3	1/2.3	II/I* or II/III*	2.3	2	<b>5</b>	1.3	2.3
SOU	Human	USA	II/III	Howe and Sibley (1995)	II/III	1.2.3	2	3	1.2.3	2.3	II/III	2.3	2	3	2	2.3
T61	Turkey	USA	II	Howe and Sibley (1995)	II/I*	1.2.3	2	<b>8</b>	1.2.3	1/2.3	II/I* or II/III*	2.3	2	<b>5</b>	1.3	2.3
NED	Human	France (Limoges)	III (Z3)	Dardé et al. (1992)	III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3
CEP	Cat	USA	III (Z3)	Dardé et al. (1992)	III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3
LGE97-3	Human	France (Limoges)	III (Z3)	Ajzenberg et al. (2002a)	III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3

(continued on next page)

Table 1 (continued)

Strain	Host	Geographical origin	Classical typing <sup>a</sup>	Reference	Alleles by SNP analysis						Alleles by MS analysis					
					Type <sup>b</sup>	TUB2	W35	TgM-A	B18	B17	Type <sup>b</sup>	TUB2	W35	TgM-A	B18	B17
M7741	Sheep	USA	III (Z3)	Dardé (1996)	III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3
OPA	Pig	Uruguay	III (Z3)	Dardé (1996)	III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3
LEG-NJA	Human	France (Nancy)	III (Z3)	Rabaud et al. (1999)	III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3
LPN-2002-SEE	Human	France (Nice)	ND		III	1.2.3	3	3	1.2.3	2.3	III	2.3	3	3	1.3	2.3
TONT	Human	France (Lyon)	Atypical (Z8)	Dardé (1996)	III	1.2.3	3	3	1.2.3	2.3	III*	2.3	3	3	<b>4</b>	2.3
IPP-2002-BAT	Human	French Guiana	ND		III	1.2.3	3	3	1.2.3	2.3	III/I*	1	3	3	<b>4</b>	2.3
ENVL-2002-MAC	Monkey	Barbados	ND		III	1.2.3	3	3	1.2.3	2.3	III/I	1	3	3	1.3	2.3
P80	Pig	USA	Atypical (Z11)	Dardé (1996)	III*	1.2.3	3	3	1.2.3	<b>10</b>	III/II*	1	3	3	2	<b>6</b>
TOU-2002-ALI	Human	Reunion island	ND		Atypical	1.2.3	<b>5</b>	1.2	1.2.3	<b>5</b>	II/III*	2.3	2	3	1.3	<b>4</b>
RUB	Human	French Guiana	Atypical (Z6)	Dardé et al. (1998)	Atypical	1.2.3	<b>7</b>	<b>5</b>	1.2.3	<b>12</b>	Atypical	2.3	3	3	<b>7</b>	<b>9</b>
GUY-2003-BAS	Human	French Guiana	ND		Atypical	1.2.3	<b>6</b>	<b>6</b>	1.2.3	<b>9</b>	Atypical	2.3	2	<b>4</b>	<b>5</b>	<b>5</b>
GUY-2002-KOE	Human	French Guiana	ND		Atypical	1.2.3	<b>9</b>	<b>6</b>	1.2.3	<b>6</b>	Atypical	2.3	<b>4</b>	<b>4</b>	<b>5</b>	<b>4</b>
GUY-2003-MEL	Human	French Guiana	ND		Atypical	1.2.3	<b>7</b>	<b>6</b>	1.2.3	<b>7</b>	Atypical	2.3	2	<b>4</b>	2	<b>5</b>
GUY-2001-DOS	Human	French Guiana	ND		Atypical	1.2.3	<b>10</b>	<b>6</b>	1.2.3	<b>7</b>	Atypical	2.3	<b>4</b>	<b>4</b>	1.3	<b>5</b>
GUY-2002-MAT	Human	French Guiana	ND		Atypical	1.2.3	<b>6</b>	<b>6</b>	1.2.3	<b>13</b>	Atypical	1	<b>7</b>	<b>4</b>	1.3	<b>11</b>
VAND	Human	French Guiana	Atypical (Z12)	Bossi et al. (1998)	Atypical	1.2.3	<b>8</b>	<b>7</b>	1.2.3	<b>8</b>	Atypical	1	2	<b>4</b>	<b>4</b>	<b>5</b>
GUY-2003-ADA	Human	French Guiana	ND		Atypical	<b>4</b>	<b>11</b>	1.2	1.2.3	1/2.3	Atypical	1	<b>4</b>	3	<b>6</b>	<b>4</b>
MAS	Human	France (Nice)	Atypical (Z5)	Dardé et al. (1992)	Atypical	1.2.3	<b>5</b>	1.2	1.2.3	<b>11</b>	Atypical	1	2	3	<b>4</b>	<b>10</b>
IPP-2002-URB	Human	France (Paris)	ND		Atypical	1.2.3	<b>4</b>	<b>4</b>	1.2.3	<b>11</b>	Atypical	<b>4</b>	2	2	1.3	<b>7</b>
CASTELLS	Sheep	Uruguay	Atypical (Z7)	Dardé, 1996	Atypical	1.2.3	<b>4</b>	<b>4</b>	1.2.3	<b>11</b>	Atypical	<b>4</b>	2	2	2	<b>8</b>

Alleles 1, 2, 3 are reserved for clonal lineages I, II and III. Allele 1.2 means that types I and II share the allele; allele 2.3 means that types II and III share the allele; allele 1.2.3 means that types I, II and III share the allele. Allele 1/2.3 means a combination of clonal type I with type II or III allele. 'Allele 4' and above are used for atypical alleles (in bold).

<sup>a</sup> Z, zymodeme; ND, not done.

<sup>b</sup> (\*) indicates an atypical allele in the MS or SNP type.

Table 2  
Genetic markers

Marker and accession number <sup>a</sup>	Chromosome	Coding function <sup>b</sup>	Primer sequences <sup>c</sup>	Size range (bp)	Total no. alleles	No. MS alleles	No. SNP alleles (flanking regions)	GenBank accession numbers
<i>TUB2</i> (M20025)	IX	Beta-tubulin gene	(F) 5' GTCCGGGTGTTCTCACAAA 3' (R) 5' TTGGCCAAAAGACGAAGTTGT 3'	255–259	4	3	2	AY572562– AY572604
<i>W35</i> (W35487)	II	Unknown (EST)	(F) 5' GGTTCACTGGATCTTCCTCAA 3' (R) 5' AATGAACGTCGGTTGTTCC 3'	210–216	15	7	11	AY572605– AY572647
<i>TgM-A</i> (Y17507)	X	Myosin A gene	(F) 5' GCGTCGACATGAGTTTCTC 3' (R) 5' TGGGCATGATAATGTAGAGATG 3'	152–160	10	5	7	AY572649– AY572691
<i>B18</i> (BM189462)	VII	Unknown (EST)	(F) 5' TGGTCTTCAACCTTTCATCC 3' (R) 5' AGGATAAGTTTCTCACAAACA 3'	113–128	7	6	2	AY572693– AY572735
<i>B17</i> (BM175053)	XII	Unknown (EST)	(F) 5' CGACAAGTCCATGCGAACTA 3' (R) 5' GGCAACAGGAGGTAGAGGAG 3'	163–189	16	10	13	AY572737– AY572779

<sup>a</sup> GenBank accession number of RH original sequence.<sup>b</sup> EST, expressed sequence tags.<sup>c</sup> F, forward primer; R, reverse primer.

Prism<sup>®</sup> 3100 Genetic Analyzer (Applied Biosystems, Courtaboeuf, France) and sequences were evaluated using Sequencing Analysis 3.7 software.

#### 2.4. Phylogenetic and genetic diversity analyses

The nucleotide sequences of the five MS loci, comprising the microsatellite regions and its flanking regions, were aligned using Multiple Sequence Alignment Program, ClustalX (version 1.81, June 2000) (Thompson et al., 1997). Two phylogenetic methodologies were used: distance and parsimony methodologies. First, for the distance analysis, in order to consider the indel characters (insertion/deletion mutations), we have used the GapCoder program (Young and Healy, 2003). The output of this last program is a Nexus format file compatible with PAUP package. The distance calculations and tree building by the Neighbour process (Saitou and Nei, 1987) were performed by the PAUP package for windows (Swofford, D.L., 2001. PAUP: phylogenetic analysis using parsimony (and other methods), vers. 4.0b10. Sinauer Associates, Sunderland MA). The *P* (uncorrected distance often referred as *p*-distances or dissimilarity (*D*) distance) and Jukes–Cantor (Jukes and Cantor, 1969) distances were used to construct a neighbour-joining tree. Second, phylogenetic analysis was performed using the PHYLIP package (Felsenstein, 1989). We applied parsimony method with bootstrapping (100 replicates) with SEQBOOT, DNAPARS, and CONSENSE programs. *Hammondia hammondi* was used as outgroup. The construction of trees was performed by Treedyn software (Chevenet, 2000, <http://viradium.mpl.ird.fr/treedyn>). This kind of analysis must be performed without the repeated genotypes in order to study the real phylogenetic relationships and to avoid the bias produced by a high number of identical genotypes. The sequences of each locus were studied individually on one hand and grouped on the other hand.

In order to get the distribution of genetic variability in the sample, two diversity indices comprising genotypic diversity (number of different genotypes on the total number of genotypes), and mean genetic distances based on *P* distances were evaluated. The indices were calculated within the whole sample, within the French sample and the French Guiana one (see Table 3).

Table 3  
Values of genetic diversity indices in the whole sample and two subgroups

	Number of strains	Genotypic diversity	Mean <i>P</i> distance (standard deviation)
Whole sample	43	0.74	0.016 (0.007)
French Guiana sample	9	1	0.021 (0.007)
France sample	13	0.69	0.013 (0.007)



### 3. Results

#### 3.1. Molecular analysis

All 43 *T. gondii* strains were successfully sequenced at the five MS loci and the sequences were submitted to GenBank (see Table 2 for accession numbers). Based on the polymorphisms of dinucleotide repeats in microsatellite regions and the single nucleotide polymorphisms (SNPs) in flanking regions, a number of total alleles were identified for each marker including four alleles at *TUB2*, 15 at *W35*, 10 at *TgM-A*, seven at *B18* and 16 at *B17* (Fig. 1 and Table 2).

The number of SNP alleles varies depending on the markers. There are 2, 11, 7, 2 and 13 different SNP alleles in *TUB2*, *W35*, *TgM-A*, *B18* and *B17*, respectively (Fig. 1 and Table 2). Atypical SNPs that are not in the three clonal lineages are mainly observed in French Guiana strains at markers *W35*, *TgM-A* and *B17* (Fig. 1 and Table 1). In comparison to type II reference strain ME49, the most polymorphic strains were all originated from French Guiana. RUB and GUY-2003-MEL are the most polymorphic strains at *W35* locus (2.6% polymorphism), VAND at *TgM-A* (2.14% polymorphism) and GUY-2002-MAT at *B17* (3.9% polymorphism due not only to SNPs but also to insertion of three nucleotides). Both *TUB2* and *B18* are less polymorphic, with only one SNP in flanking regions of *TUB2* in a French Guiana strain GUY-2003-ADA, and one SNP associated with a deletion of three nucleotides in flanking regions of *B18* for a type II strain DEG.

A number of MS alleles also are identified for each marker. There are 3, 7, 5, 6 and 10 different alleles in *TUB2*, *W35*, *TgM-A*, *B18* and *B17*, respectively (Fig. 1 and Table 2). For marker *B18*, there are four atypical MS alleles even though no SNPs were found in their flanking regions. The MS polymorphism consists in size variation of dinucleotide tandem repeats but also in rare mutations (only for French Guiana strains and located at *W35*, *B18*, and *B17*) (Fig. 1). As in SNP alleles, the majority of atypical MS alleles are found in French Guiana strains and some MS alleles are encountered only in French

Guiana strains such as alleles 4 at *TgM-A* and *W35* or alleles 5 at *B18* and *B17* (Table 1).

Combination of the five MS or SNP markers allows classifying the majority of strains in the three classical types although some type I or II strains as determined by classical typing exhibited a mixed MS or SNP type, and/or, in some cases, one atypical allele (Table 1). These latter strains and also undetermined ones can be related to one of the main types when SNP alleles and/or the majority of their MS alleles were characteristic of this type. For instance, in Table 1, from BK to PSP-2003-KOM, some classical type I or previously undetermined genotypes that exhibit one type III MS allele (*TgM-A*) and sometimes an atypical allele (*W35* for PSP-2003-KOM) can be considered as type I related strains with mixed MS and SNP type I/III. Mixed types, with or without atypical alleles, can also be described for strains related to types II and III.

Three strains (PSP-2003-ERO, T61, and P80) are difficult to relate to one of the main types by molecular analysis because they exhibit a clearly mixed genotype with type I, II, and III alleles associated to one atypical MS or SNP allele.

From TOU-2002-ALI to CASTELLS (Table 1), the strains can be considered as atypical, not related to one of the main types because they presented more than two atypical MS and/or SNP alleles in their genotype.

Primers used for *T. gondii* allowed the amplification of only four sequences in *H. hammondi* genome (*TUB2* was not amplified). A microsatellite sequence similar to *Toxoplasma* was detected only for *W35* and *B18* sequences. The corresponding sequences were submitted to GenBank (accession nos. AY572648, AY572692, AY572736 and AY572780).

#### 3.2. Genetic diversity analysis

In order to better understand the distribution of the diversity in the sample, the following subgroups were studied (Table 3): (i) the whole sample; (ii) strains from French Guiana; and (iii) strains from France. The two genetic indices, genotypic diversity and mean genetic

Fig. 1. SNP and MS alleles defined by polymorphic sites in flanking region and microsatellite sequences, respectively. In each marker, sites demarcated by an asterisk (\*) indicate deletion sites in comparison with the longest microsatellite sequence. Among 14 strains arbitrarily selected to represent the three major lineages, BK, Me49 and CEP are reference type I, II and III strains, respectively. Alleles 1, 2, 3 are reserved for clonal lineages I, II and III. Allele 1.2 means that types I and II share the allele; allele 2.3 means that types II and III share the allele; allele 1.2.3 means that types I, II and III share the allele. Allele 1/2.3 means a combination of clonal type I with type II or III allele. Allele 4 and above are used for atypical alleles. These data were used to generate data for Table 1. <sup>a</sup>ATIH, PIL, GPHT, WIK, PSP-2003-KOM, PSP-2003-ERO, IPP-2002-BAT, ENVL-2002-MAC, P80, GUY-2002-MAT, VAND, MAS. <sup>b</sup>BEV, DEG, REN-2002-VAL, B41, BOU, GANGI, ARI, WTD-1, T61, RMS-2001-MAU, B73, SOU, TOU-2002-ALI, NED, LGE97-3, M7741, OPA, LEG-NJA, LPN-2002-SEE, TONT, RUB, GUY-2003-BAS, GUY-2002-KOE, GUY-2003-MEL, GUY-2001-DOS. <sup>c</sup>ATIH, PIL, GPHT, WIK. <sup>d</sup>BEV, DEG, REN-2002-VAL, B41, GANGI, ARI, WTD-1, T61, RMS-2001-MAU, B73, SOU. <sup>e</sup>NED, LGE97-3, M7741, OPA, LEG-NJA, LPN-2002-SEE, PSP-2003-ERO, TONT, IPP-2002-BAT, ENVL-2002-MAC, P80. <sup>f</sup>ATIH, PIL, GANGI, ARI, WTD-1. <sup>g</sup>BEV, DEG, REN-2002-VAL, B41, BOU, B73. <sup>h</sup>NED, LGE97-3, M7741, OPA, LEG-NJA, LPN-2002-SEE, GPHT, WIK, PSP-2003-KOM, PSP-2003-ERO, SOU, TONT, IPP-2002-BAT, ENVL-2002-MAC, P80. <sup>i</sup>GUY-2003-BAS, GUY-2002-KOE, GUY-2003-MEL, GUY-2002-MAT. <sup>j</sup>ATIH, PIL, GPHT, WIK, PSP-2003-KOM, WTD-1, T61, RMS-2001-MAU, B73, TOU-2002-ALI, NED, LGE97-3, M7741, OPA, LEG-NJA, LPN-2002-SEE, ENVL-2002-MAC, GUY-2001-DOS, GUY-2002-MAT, IPP-2002-URB. <sup>k</sup>BEV, REN-2002-VAL, BOU, GANGI, ARI, SOU, P80, GUY-2003-MEL, CASTELLS. <sup>l</sup>PSP-2003-ERO, TONT, IPP-2002-BAT, VAND, MAS. <sup>m</sup>ATIH, PIL, GPHT, WIK, PSP-2003-KOM, PSP-2003-ERO. <sup>n</sup>BEV, DEG, REN-2002-VAL, BOU, GANGI, B73, SOU, NED, LGE97-3, M7741, OPA, LEG-NJA, LPN-2002-SEE, TONT, IPP-2002-BAT, ENVL-2002-MAC.

**Alleles**

**TUB2 (255-259 bp): 4 alleles**

	SNP	MS	46	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	
BK and 12 identical <sup>a</sup>	1.2.3	1	C	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C
ME49, CEP and 25 identical <sup>b</sup>	1.2.3	2.3	C	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C
CASTELLS, IPP-2002-URB	1.2.3	4	C	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C
GUY-2003-ADA	4	1	T	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C

**W35 (210-216 bp): 15 alleles**

			17	35	36	38	40	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	105	124	125	128	132	142	187						
BK (I) and 4 identical <sup>a</sup>	1	1	G	C	T	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	G	T	G	*	*	G	C	G	T	T	C	T		
ME49 (II) and 11 identical <sup>a</sup>	2	2	G	C	G	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
CEP (III) and 11 identical <sup>a</sup>	3	3	G	C	G	A	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
CASTELLS, IPP-2002-URB	4	2	G	T	G	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
MAS, TOU-2002-ALI	5	2	G	C	G	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
GUY-2003-BAS	6	2	G	C	G	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
GUY-2003-MEL	7	2	G	C	G	A	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
VAND	8	2	C	C	G	G	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
RUB	7	3	G	C	G	A	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
GUY-2002-KOE	9	4	G	C	G	G	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
GUY-2001-DOS	10	4	G	C	G	G	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
GUY-2003-ADA	11	4	G	C	G	G	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C
BOU	2	5	G	C	G	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
PSP-2003-KOM	1	6	G	C	T	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T
GUY-2002-MAT	6	7	G	C	G	G	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T	C	T

**TgM-A (152-160 pb): 10 alleles**

			1	16	22	30	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	98	123																		
BK (I) and 5 identical <sup>a</sup>	1.2	1	T	A	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
ME49 (II) and 6 identical <sup>a</sup>	1.2	2	T	A	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
CEP (III) and 15 identical <sup>a</sup>	3	3	T	A	G	T	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
CASTELLS, IPP-2002-URB	4	2	T	G	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
MAS, TOU-2002-ALI, GUY-2003-ADA	1.2	3	T	A	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
RUB	5	3	T	A	G	T	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
GUY-2001-DOS and 4 identical <sup>a</sup>	6	4	G	G	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
VAND	7	4	G	G	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
RMS-2001-MAU	1.2	5	T	A	G	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G
T61	8	5	T	A	A	C	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G

**B18 (113-128 bp): 7 alleles**

			33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	68	88	90	91	92													
BK, CEP and 20 identical <sup>a</sup>	1.2.3	1.3	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A
ME49 (II) and 9 identical <sup>a</sup>	1.2.3	2	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A
B41 and 5 identical <sup>a</sup>	1.2.3	4	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A
GUY-2003-BAS, GUY-2003-KOE	1.2.3	5	C	A	C	A	C	G	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A
GUY-2003-ADA	1.2.3	6	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A
RUB	1.2.3	7	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A
DEG	4	2	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A

**B17 (163-189 bp): 16 alleles**

			15	21	33	34	43	44	54	77	78	79	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	12
--	--	--	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

distances were calculated in each group (Table 3). The *H. hammondi* strain was not included in the calculations.

We could observe that even if the genotypic diversity is high in some subgroups, the mean *P* genetic distance remains low (Table 3). These results showed the higher diversity in the French Guiana sample compared to the France subgroup.

3.3. Phylogenetic analysis

As the results of parsimony and distance analyses are similar, here we report the result of Wagner analysis due to its ease of visual presentation. Fig. 2 shows the result based on the sequence data of all 43 *T. gondii* strains. The three clonal lineages (the column ‘classical typing’ in Fig 2) were individualised on the phylogram.

All the strains belonging to type I were clustered together with a bootstrap value equal to 63.4. PSP-2003-KOM and PSP-2003-ERO strains, which were not identified before, clustered more closely with type I than the other strains. As the strains pertaining to the classical type III revealed only one genotype, one branch represented this lineage in the tree. Four strains were phylogenetically linked or similar to type III: IPP-2002-BAT, ENVL-2002-MAC and LPN-2002-SEE which were also non-characterised and TONT previously identified as an atypical strain. All the strains pertaining to type II were clustered in a node sustained by a weak bootstrap value of 47.8. Four strains were phylogenetically included in the type II cluster: RMS-2001-MAU and REN-2002-VAL which were uncharacterised strains and the two recombinant strains (SOU and B73) previously identified by Howe and Sibley (1995).

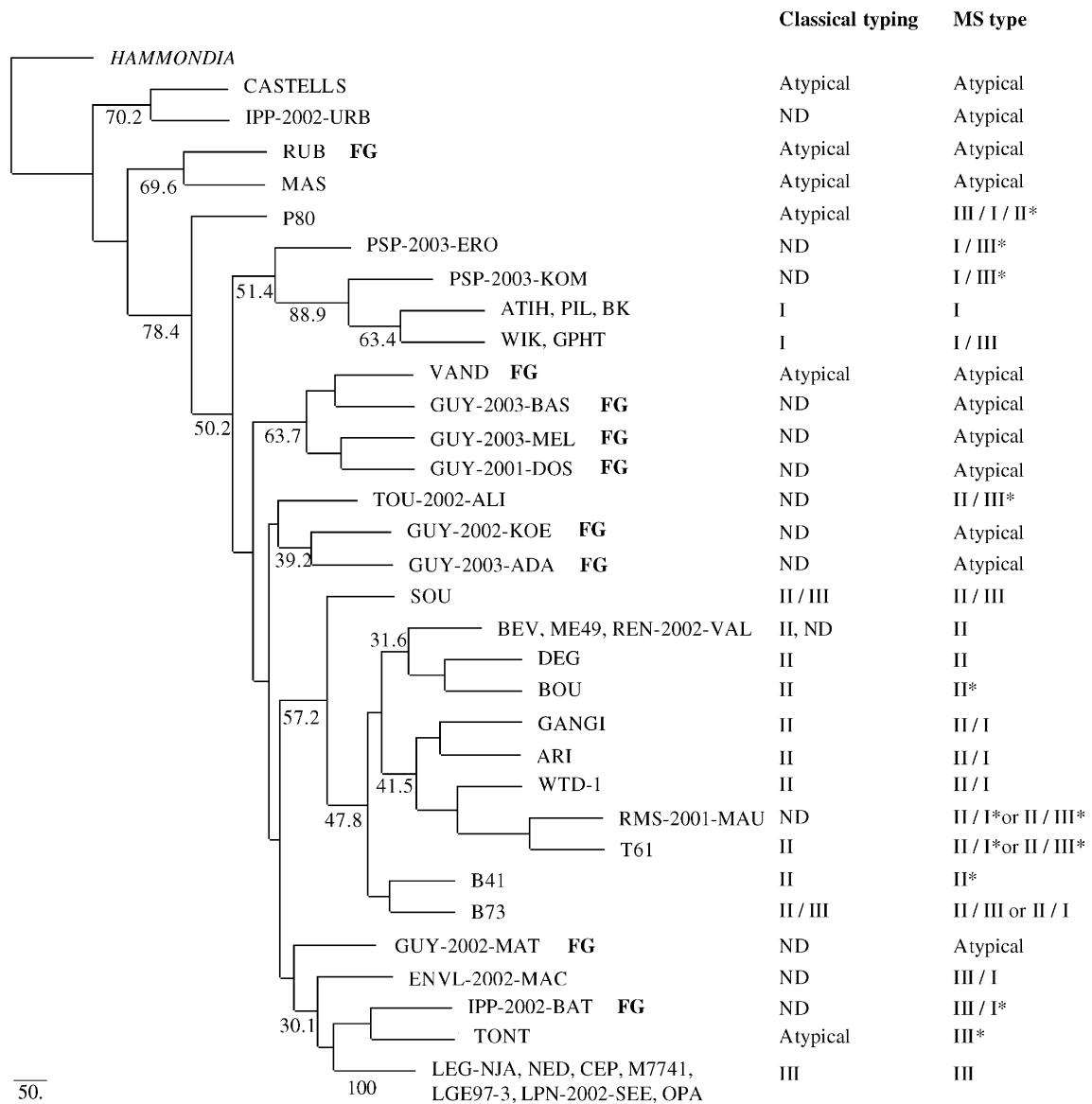


Fig. 2. Phylogram of 43 *Toxoplasma gondii* strains as determined by analysis of the entire sequences of the five markers. *Hammondia hammondi* was used as the outgroup. The tree was built by Wagner analysis after bootstrapping with 100 repetitions. (FG) indicates a French Guiana strain. The classical typing and MS genotype are detailed for each strain. Asterisk indicates an atypical allele in the MS genotype (see Table 1).



When we considered the different type lineages with the related strains, we still observed a phylogenetic individualisation, but with changes in the branch robustness. Indeed the type I and related strains group was supported by a bootstrap value of 51.4, type II and related strains by a bootstrap value of 57.2, and type III and related strains by a bootstrap value of 30.1.

The strains originating from French Guiana, France and USA were distributed along the phylogram without geographical structuration. Moreover, the atypical strains did not represent a phylogenetically individualised group, as they were also distributed all over the tree (Fig. 2).

#### 4. Discussion

The clonal theory of *T. gondii* population structure, with three predominant lineages, is based on *T. gondii* samples collected mainly from the European or the North American domestic cycle either from clinical cases of human toxoplasmosis or from meat producing animals such as pigs, sheep and chicken (Dardé et al., 1992; Howe and Sibley, 1995; Ajzenberg et al., 2002a,b). This collection may not reflect the true status of *T. gondii* in remote geographical areas or in tropical regions where the ecological system is very different from that of the civilised world. Several studies have examined the distribution of genotypes in chickens from other countries such as Egypt, Argentina, India and Brazil (Dubey et al., 2002, 2003a–d; Sreekumar et al., 2003; Lehmann et al., 2004), but the genetic studies were limited by the use of only one marker (SAG2) (except Sreekumar et al., 2003; Lehmann et al., 2004), and because chickens in farms are indicators of strain prevalence in a domestic or peridomestic environment. The rare atypical strains were sampled mainly in tropical regions such as French Guiana (Bossi et al., 1998; Dardé et al., 1998; Carme et al., 2002b), but also in France (Dardé, 1996), or in unusual host species such as deer, bear, cougar or sea otter (Howe and Sibley, 1995; Cole et al., 2000; Lehmann et al., 2000; Miller et al., 2004).

In this study, we performed multi-locus sequencing of five microsatellite markers for 43 *T. gondii* strains from different geographical and host origins to reconsider the population structure and develop a more realistic picture of diversity between *Toxoplasma* strains by phylogenetic and molecular analysis.

The use of microsatellite markers for such a study was justified by their high discriminatory power and their usefulness to assess population genetic structure and molecular epidemiology (Goldstein and Schlötterer, 1999) and to infer phylogenetic relationships at intraspecific level or for recently diverged species (Goldstein et al., 1995; Paetkau et al., 1997; Harr et al., 1998; Fisher et al., 2000; Metzgar et al., 2001; Richard and Thorpe, 2001; Bulle et al., 2002). Strelman et al. (1998) showed that sequences flanking a microsatellite were also very

informative for inferring phylogeny. Concerning *T. gondii*, MS length polymorphisms have shown their usefulness for detecting the genetic polymorphism (Ajzenberg et al., 2002a; Blackston et al., 2001; Lehmann et al., 2004). The most polymorphic of these loci revealed substantial within-lineage variation usually explained by rapid accumulation of mutations by genetic drift. However, some of them have been demonstrated to be stable enough to type *T. gondii* and to detect recombination events (Ajzenberg et al., 2002a). The sequencing of the microsatellite region presented in this study allowed a better allelic definition than GeneScan analysis of length polymorphism, and had a higher discriminative power than classical typing by isoenzymatic and PCR-RFLP analysis which defined the types I, II and III and atypical. Alleles were defined by SNP and MS data, but MS alleles were more helpful in detecting recombination events (Table 1).

This study showed that, while the majority of isolates fall into the three clonal lineages as expected, the nine strains from French Guiana were clearly atypical, very different from the three major lineages, but also from atypical strains from France and Uruguay. Each French Guiana strain has a unique multi-locus genotype, and some of its alleles are not observed in strains from other areas. They concentrate the majority of gene pool of this study. This situation is markedly contrasted with our domestic sample from France where 95% of strains belonged to the three main types, and where 88% of strains had the same MS genotype II (BRC ToxoBS group, personal data). This is confirmed by the highest values of genetic diversity indices obtained in the French Guiana sample compared with the France subgroup. However, the global observation of the genetic diversity indices shows that even if we obtained a high allelic diversity, the real genetic polymorphism of sequences is very low which is in agreement with all the genetic data previously published (Grigg et al., 2001a; Su et al., 2003)..

The existence of sexual recombination in *T. gondii* is thought to occur as occasional events in a largely clonal population structure (Sibley and Howe, 1996; Lehmann et al., 2000, 2004; Grigg et al., 2001a; Grigg and Suzuki, 2003). Considered together for the data in this study, the phylogram hierarchy and the sequencing data are not in agreement with a strictly clonal model. On the one hand, the detailed analysis of MS genotypes shows different allele associations, not only in the atypical strains from French Guiana but also in the atypical strains of other countries, and in the three type lineages and their related strains. This is not in agreement with a theoretical clonal structure characterised by a global association between alleles due to linkage disequilibrium (Tibayrenc et al., 1990). On the other hand, from a phylogenetic point of view, the constructed phylogram does not present a structure typical of a clonal organism as occurs for *Trypanosoma cruzi*, *Entamoeba histolytica* or *Leishmania* (Tibayrenc et al., 1990; Bañuls et al., 1999). These last microorganisms are subdivided into discrete phylogenetic lineages supported by strong

bootstrap values. In *T. gondii*, the phylogenetic clusters are not supported by strong bootstrap values if we consider the three types with their related strains (see Fig. 2). This suggests that phylogenetic divergence of *T. gondii* is clouded by recombination. From these genetic and phylogenetic analyses, the most parsimonious hypothesis is that *T. gondii* presents a complex population structure with a mix of clonal and sexual propagation.

The proportion of typical strains compared to the strains with recombined genotypes in the different regions suggests that the highest proportion of genetic exchanges would occur in the wild cycle in countries where domestic breeding has a brief history. Thus, the under-representation of the wild cycle strains might explain the predominance of typical strains and the description of a global clonal structure in the previous studies. More frequent sexual recombination in *T. gondii* was found in tropical areas, not only in the wild life cycle in French Guiana, but also in areas where breeding is not intensive (e.g. Africa, Reunion Island or Caribbean Islands, although our sample from these countries is not representative). Interestingly, five out of six strains from African countries (including three strains not submitted to sequencing in this study), showed the same mixed MS genotype I/III by microsatellite length polymorphism suggesting that an emerging genotype may be present in this continent. In non-tropical areas, analysis of genotypes in a few wild animals in North America (WTD1, B41 and B73) showed genotypes with different combinations of types I, II, III, and atypical alleles, in agreement with previous typing of these strains (Howe and Sibley, 1995). Other papers also described atypical genotypes among wild animals such as cougar (Lehmann et al., 2000), and sea otters (Lehmann et al., 2000; Miller et al., 2004). So, this work suggests that a variable population structure exists in a range from clonal to sexual propagation and that *T. gondii* uses different strategies for transmission and survival in different hosts as a function of the environmental conditions. As recently proposed by Lehmann et al. (2004), after comparing microsatellite analysis of *T. gondii* from Brazil and America, *T. gondii* has evolved in order to optimise its transmission in its specific environment.

In the domestic environment, *Toxoplasma* seroprevalence in cats is high (Gauss et al., 2003), suggesting possible ingestion of preys containing more than one strain. This explains the persistence of a few recombined MS genotypes in the domestic cycle. However, in Europe or North America, intensive raising of a narrow range of domestic meat producing animals such as pigs, sheep or chickens offers a major niche to *Toxoplasma*. Among many genotypes, the three clonal lineages seem to be the most successfully adapted to these domestic hosts (Lehmann et al., 2003). High levels of transmission have been demonstrated in a pig farm (Lehmann et al., 2003) associated with an increase of prevalence in the surrounding wild environment with proximity to the pig sties. The authors defined the pig farm as a reservoir,

explaining the preponderance of *T. gondii* clonal propagation in this area. To support this, in a recent paper (Su et al., 2003) the authors estimated that the three major lineages would have diverged about 10,000 years ago which coincides with the domestication of companion and agricultural animals.

In a wild environment such as French Guiana, there is also a high seroprevalence in animals, ranging between 42 and 62% for non-carnivorous mammals like paca, armadillo, collared peccary, anteater or tapir (Carne et al., 2002a). Prevalence of the few wild felids tested soon after their capture in Amazonia suggests high levels of transmission for wild definitive hosts in the Amazonian rainforest: for felids in French Guiana three out of three were positive (Carne et al., 2002a) and in Brazilian Amazonia five out of eight (Ferraroni and Marzochi, 1980). The greater genetic diversity of *Toxoplasma* in a wild and tropical environment in comparison to the domestic one is likely to be due not only to more frequent genetic exchanges, but also to the presence of SNP polymorphisms and rare alleles. This greater *Toxoplasma* diversity is probably the result of the host diversity which is considerably higher in the wild host populations. Thus, in French Guiana, where the fauna, like the flora, is abundant and highly diverse with 179 mammalian wild species identified in 1998 (Carne et al., 2002a), a plurality of alleles is needed for *Toxoplasma* to colonise the available ecological niches in such a diverse environment.

The French Guiana strains were isolated from human patients with severe disseminated toxoplasmosis, except one strain IPP-2002-BAT that was isolated from a case of asymptomatic congenital toxoplasmosis. Since the human infections occurred in the rainforest by eating undercooked game or drinking untreated river water (Carne et al., 2002b; and unpublished data), these strains can probably be considered as an indirect sampling of the wild cycle of *T. gondii* in the Amazonian rainforest. Though this sample could not represent all the strains circulating in this wild ecosystem, it did allow us to compare the genetic data of strains obtained in a relatively restrictive wild environment with data of strains obtained in an allopatric domestic environment. One must keep in mind that *T. gondii* strains from ecosystems such as remote tropical regions still are largely underrepresented. To better understand the population structure in *T. gondii*, more strains need to be sampled, on one hand directly from wild animals in Amazonian rain forest to avoid the human disease sampling bias, and on the other hand, from an area in which wild and domestic cycles are sympatric.

#### Acknowledgements

This work was supported by French Ministry of Research (support no. 02 g 0418 for the development of Biological Resource Centers). We wish to thank for strain providing

the BRC ToxoBS group (French parasitologist network for *Toxoplasma* strain collection): J.M. Pinon, I. Villena (Reims, RMS), P. Thulliez (Paris Institut de Puériculture, IPP), P. Marty (Nice, LPN), M.H. Bessières (Toulouse, TOU), J.P. Gangneux (Rennes, REN), L. Paris (Paris Pitié-Salpêtrière, PSP), J. Berthonneau (Poitiers), A. Bonnin (Dijon), E. Candolfi (Strasbourg), D. Chabasse (Angers), B. Couprie (Bordeaux), L. Delhaes (Lille), C. Duhamel (Caen), T.H. Duong (Tours), L. Favennec (Rouen), P. Flori (Saint Etienne), B. Fortier (Nancy), J. Franck (Marseille), J.F. Garin (Paris Saint Louis), O. Morin (Nantes), G. Nevez (Amiens), H. Pelloux (Grenoble), F. Peyron (Lyon), R. Piarroux (Besançon), F. Pratlong (Montpellier), D. Quinio (Brest), H. Yera (Paris Cochon). We also acknowledge D. Sibley for DNA samples, J.K. Frenkel for giving us the *Hammondia hammondi* strain, C. Chevillon, T. De Meeus and P. Durand for their useful discussions on population genetics and microsatellites.

## References

- Ajzenberg, D., Bañuls, A.L., Tibayrenc, M., Dardé, M.L., 2002a. Microsatellite analysis of *Toxoplasma gondii* population shows considerable polymorphism structured into two main clonal groups. *Int. J. Parasitol.* 32, 27–38.
- Ajzenberg, D., Cogné, N., Paris, L., Bessières, M.H., Thulliez, P., Filisetti, P., Pelloux, H., Marty, P., Dardé, M.L., 2002b. Genotype of 86 *Toxoplasma gondii* isolates associated with human congenital toxoplasmosis and correlation with clinical findings. *J. Infect. Dis.* 186, 684–689.
- Bañuls, A.L., Hide, M., Tibayrenc, M., 1999. Molecular epidemiology and evolutionary genetics of *Leishmania* parasites. *Int. J. Parasitol.* 29, 1137–1147.
- Blackston, C.R., Dubey, J.P., Dotson, E., Su, C., Thulliez, P., Sibley, D., Lehmann, T., 2001. High-resolution typing of *Toxoplasma gondii* using microsatellite loci. *J. Parasitol.* 87, 1472–1475.
- Bossi, P., Caumes, E., Paris, L., Dardé, M.L., Bricaire, F., 1998. *Toxoplasma gondii*-associated Guillain-Barré syndrome in an immunocompetent patient. *J. Clin. Microbiol.* 36, 3724–3725.
- Bulle, B., Millon, L., Bart, J.M., Gallego, M., Gambarelli, F., Portus, M., Schnur, L., Jaffe, C.L., Fernandez-Barredo, S., Alunda, J.M., Piarroux, R., 2002. Practical approach for typing strains of *Leishmania infantum* by microsatellite analysis. *J. Clin. Microbiol.* 40, 3391–3397.
- Carne, B., Aznar, C., Motard, A., Demar, M., De Thoisy, B., 2002a. Serologic survey of *Toxoplasma gondii* in noncarnivorous free-ranging neotropical mammals in French Guiana. *Vector Borne Zoonotic Dis.* 2, 11–17.
- Carne, B., Bissuel, F., Ajzenberg, D., Bouyine, R., Aznar, C., Demar, M., Bichat, S., Louvel, D., Bourbigot, A.M., Peneau, C., Neron, P., Dardé, M.L., 2002b. Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. *J. Clin. Microbiol.* 40, 4037–4044.
- Cole, R.A., Lindsay, D.S., Howe, D.K., Roderick, C.L., Dubey, J.P., Thomas, N.J., Baeten, L.A., 2000. Biological and molecular characterizations of *Toxoplasma gondii* strains obtained from southern sea otters (*Enhydra lutris nereis*). *J. Parasitol.* 86, 526–530.
- Dardé, M.L., 1996. Biodiversity in *Toxoplasma gondii*. *Curr. Top. Microbiol. Immunol.* 219, 27–41.
- Dardé, M.L., Bouteille, B., Pestre-Alexandre, M., 1992. Isoenzyme analysis of 35 *Toxoplasma gondii* isolates and the biological and epidemiological implications. *J. Parasitol.* 78, 786–794.
- Dardé, M.L., Villena, I., Pinon, J.M., Beguinot, I., 1998. Severe toxoplasmosis caused by a *Toxoplasma gondii* strain with a new isoenzyme type acquired in French Guiana. *J. Clin. Microbiol.* 36, 324.
- Dubey, J.P., Graham, D.H., Blackston, C.R., Lehmann, T., Gennari, S.M., Ragozo, A.M., Nishi, S.M., Shen, S.K., Kwok, O.C., Hill, D.E., Thulliez, P., 2002. Biological and genetic characterisation of *Toxoplasma gondii* isolates from chickens (*Gallus domesticus*) from Sao Paulo, Brazil: unexpected findings. *Int. J. Parasitol.* 32, 99–105.
- Dubey, J.P., Graham, D.H., Dahl, E., Hilali, M., El-Ghaysh, A., Sreekumar, C., Kwok, O.C., Shen, S.K., Lehmann, T., 2003a. Isolation and molecular characterization of *Toxoplasma gondii* from chickens and ducks from Egypt. *Vet. Parasitol.* 30, 89–95.
- Dubey, J.P., Graham, D.H., Da Silva, D.S., Lehmann, T., Bahia-Oliveira, L.M.G., 2003b. *Toxoplasma gondii* isolates of free-ranging chickens from Rio de Janeiro, Brazil: mouse mortality, genotype and oocyst shedding by cats. *J. Parasitol.* 89, 851–853.
- Dubey, J.P., Navarro, I.T., Graham, D.H., Dahl, E., Freire, R.L., Prudencio, L.B., Sreekumar, C., Vianna, M.C., Lehmann, T., 2003c. Characterization of *Toxoplasma gondii* isolates from free range chickens from Parana, Brazil. *Vet. Parasitol.* 117, 229–234.
- Dubey, J.P., Venturini, M.C., Venturini, L., Piscopo, M., Graham, D.H., Dahl, E., Sreekumar, C., Vianna, M.C., Lehmann, T., 2003d. Isolation and genotyping of *Toxoplasma gondii* from free-ranging chickens from Argentina. *J. Parasitol.* 89, 1063–1064.
- Felsenstein, J., 1989. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* 5, 164–166.
- Ferguson, D.J., 2002. *Toxoplasma gondii* and sex: essential or optional extra?. *Trends Parasitol.* 18, 355–359.
- Ferraroni, J.J., Marzochi, M.C., 1980. Prevalence of *Toxoplasma gondii* infection in domestic and wild animals, and human groups of the Amazonas region. *Mem. Inst. Oswaldo Cruz* 75, 99–109.
- Fisher, M.C., Koenig, G., White, T.J., Taylor, J.W., 2000. A test for concordance between the multilocus genealogies of genes and microsatellites in the pathogenic fungus *Coccidioides immitis*. *Mol. Biol. Evol.* 17, 1164–1174.
- Gauss, C.B., Almeria, S., Ortuno, A., Garcia, F., Dubey, J.P., 2003. Seroprevalence of *Toxoplasma gondii* antibodies in domestic cats from Barcelona, Spain. *J. Parasitol.* 89, 1067–1068.
- Ghosh, J., Paris, L., Ajzenberg, D., Carcelain, G., Dardé, M.L., Tubiana, R., Bossi, P., Bricaire, F., Katlama, C., 2003. Atypical toxoplasmic manifestation after discontinuation of maintenance therapy in a human immunodeficiency virus type 1-infected patient with immune recovery. *Clin. Infect. Dis.* 37, 112–114.
- Goldstein, D.B., Schlötterer, C., 1999. Microsatellites. Evolution and Applications. University Press, Oxford.
- Goldstein, D.B., Ruiz Linares, A., Cavalli-Sforza, L.L., Feldman, M.W., 1995. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* 92, 6723–6727.
- Grigg, M.E., Suzuki, Y., 2003. Sexual recombination and clonal evolution of virulence in *Toxoplasma*. *Microbes Infect.* 5, 685–690.
- Grigg, M.E., Bonnefoy, S., Hehl, A.B., Suzuki, Y., Boothroyd, J.C., 2001a. Success and virulence in *Toxoplasma* as the result of sexual recombination between two distinct ancestries. *Science* 294, 161–165.
- Grigg, M.E., Ganatra, J., Boothroyd, J.C., Margolis, T.P., 2001b. Unusual abundance of atypical strains associated with human ocular toxoplasmosis. *J. Infect. Dis.* 184, 633–639.
- Harr, B., Weiss, S., David, J.R., Brem, G., Schlotterer, C., 1998. A microsatellite-based multilocus phylogeny of the *Drosophila melanogaster* species complex. *Curr. Biol.* 8, 1183–1186.
- Howe, D.K., Sibley, L.D., 1995. *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *J. Infect. Dis.* 172, 1561–1566.
- Jukes, T.H., Cantor, C.R., 1969. Evolution of protein molecules, in: Munro, H.N. (Ed.), *Mammalian Protein Metabolism*. Academic Press, London, pp. 21–132.

- Lehmann, T., Blackston, C.R., Parmley, S.F., Remington, J.S., Dubey, J.P., 2000. Strain typing of *Toxoplasma gondii*: comparison of antigen-coding and housekeeping genes. *J. Parasitol.* 86, 960–971.
- Lehmann, T., Graham, D.H., Dahl, E., Sreekumar, C., Launer, F., Corn, J.L., Gamble, H.R., Dubey, J.P., 2003. Transmission dynamics of *Toxoplasma gondii* on a pig farm. *Infect. Genet. Evol.* 3, 135–141.
- Lehmann, T., Graham, D.H., Dahl, E.R., Bahia-Oliveira, L.M.G., Gennari, S.M., Dubey, J.P., 2004. Variation in the structure of *Toxoplasma gondii* and the roles of selfing, drift, and epistatic selection in maintaining linkage disequilibria. *Infect. Genet. Evol.* 4, 107–114.
- Metzgar, D., Thomas, E., Davis, C., Field, D., Wills, C., 2001. The microsatellites of *Escherichia coli*: rapidly evolving repetitive DNAs in a non-pathogenic prokaryote. *Mol. Microbiol.* 39, 183–190.
- Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R., Jessup, D.A., Boothroyd, J.C., Brownstein, D., Conrad, P.A., 2004. An unusual genotype of *Toxoplasma gondii* is common in California sea otters (*Enhydra lutris nereis*) and is a cause of mortality. *Int. J. Parasitol.* 34, 275–284.
- Paetkau, D., Waits, L.P., Clarkson, P.L., Craighead, L., Strobeck, C., 1997. An empirical evaluation of genetic distance statistics using microsatellite data from bear (*Ursidae*) populations. *Genetics* 147, 1943–1957.
- Rabaud, C., Dardé, M.L., Biava, M.F., Doco-Lecompte, T., May, T., Canton, P., 1999. Deux cas de toxoplasmose pulmonaire concomitants au sein d'un couple d'hétérosexuels infectés par le VIH.4 *Med. Mal. Inf.* 29, 650–652.
- Richard, M., Thorpe, R.S., 2001. Can microsatellites be used to infer phylogenies? Evidence from population affinities of the Western Canary Island lizard (*Gallotia galloti*). *Mol. Phylogenet. Evol.* 20, 351–360.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.
- Sibley, L.D., Boothroyd, J.C., 1992. Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. *Nature* 359, 82–85.
- Sibley, L.D., Howe, D.K., 1996. Genetic basis of pathogenicity in toxoplasmosis. *Curr. Top. Microbiol. Immunol.* 219, 3–15.
- Sreekumar, C., Graham, D.H., Dahl, E., Lehmann, T., Raman, M., Bhalerao, D.P., Vianna, M.C., Dubey, J.P., 2003. Genotyping of *Toxoplasma gondii* isolates from chickens from India. *Vet. Parasitol.* 118, 187–194.
- Streelman, J.T., Zardoya, R., Meyer, A., Karl, S.A., 1998. Multilocus phylogeny of cichlid fishes (Pisces: Perciformes): evolutionary comparison of microsatellite and single-copy nuclear loci. *Mol. Biol. Evol.* 15, 798–808.
- 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.
- 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.
- Tibayrenc, M., Kjellberg, F., Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. *Proc. Natl. Acad. Sci. USA* 87, 2414–2418.
- Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Brenière, S.F., Dardé, M.L., Ayala, F.J., 1991. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. *Proc. Natl. Acad. Sci. USA* 88, 5129–5133.
- Young, N.D., Healy, J., 2003. GapCoder automates the use of indel characters in phylogenetic analysis. *BMC Bioinformatics* 19, 6.
- Zhao, X., Duszynski, D.W., Loker, E.S., 2001. A simple method of DNA extraction for *Eimeria* species. *J. Microbiol. Methods* 44, 131–137.