Genotyping of Toxoplasma gondii strains from immunocompromised patients reveals high prevalence of type I strains

A. Khan
Washington University School of Medicine in St. Louis

C. Su
Washington University School of Medicine in St. Louis

M. German
Washington University School of Medicine in St. Louis

G. A. Storch
Washington University School of Medicine in St. Louis

D. B. Clifford
Washington University School of Medicine in St. Louis

See next page for additional authors

Follow this and additional works at: https://digitalcommons.wustl.edu/open_access_pubs

Please let us know how this document benefits you.

Recommended Citation
https://digitalcommons.wustl.edu/open_access_pubs/2523

This Open Access Publication is brought to you for free and open access by Digital Commons@Becker. It has been accepted for inclusion in Open Access Publications by an authorized administrator of Digital Commons@Becker. For more information, please contact vanam@wustl.edu.
Authors
Genotyping of *Toxoplasma gondii* Strains from Immunocompromised Patients Reveals High Prevalence of Type I Strains


Department of Molecular Microbiology, Department of Pediatrics, and Department of Neurology and Medicine, Washington University School of Medicine, St. Louis, Missouri 63110

Received 15 June 2005/Returned for modification 21 July 2005/Accepted 27 September 2005

*Toxoplasma gondii* is an important food- and waterborne opportunistic pathogen that causes severe disease in immunocompromised patients. *T. gondii* has an unusual clonal population structure consisting of three widespread lineages known as I, II, and III. To establish the genotypes of strains of *T. gondii* associated with human toxoplasmosis, we have developed a set of four highly sensitive and polymorphic nested PCR markers. Multiplex nested PCR analysis was used to genotype parasites in cerebral spinal fluid samples from 8 of 10 human immunodeficiency virus-positive patients. Remarkably, a majority of these patients had infections with type I strains or strains containing type I alleles, despite the fact that this lineage is normally uncommon in humans and animals. Multiplex analysis of these four unlinked makers was able to distinguish all three common genotypes and also detected two strains with mixed genotypes. Further analysis based on sequencing of a polymorphic intron revealed that one of these recombinant strains was an exotic lineage distinct from the archetypal clonal lineages. The multiplex nested PCR analysis described here will be useful for analyzing the contribution of parasite genotype to toxoplasmosis.

*Toxoplasma gondii* is a common cause of infection in many warm-blooded animals, including humans (6). Infection in humans generally occurs either by consuming tissue cysts in raw or undercooked meat or by accidental ingestion of oocysts shed in the feces of infected cats (20). Following acute infection, *T. gondii* differentiates into tissue cysts that reside in the muscles and brain. Most cases of human infection are mild, but devastating disease can occur in congenitally infected children (33) and in immunocompromised individuals, including patients with AIDS (15, 19). Toxoplastic encephalitis (TE) is a serious clinical complication in immunocompromised patients, especially in patients with AIDS. The majority of these cases are thought to result from recurrence of chronic infections following declining cellular immunity (18). Prophylactic antibiotics that were adopted to prevent *Pneumocystis pneumonia* also resulted in decreased incidence of TE. Additionally, the introduction of advanced highly active antiretroviral therapy (HAART) further resulted in dramatically reduced incidence compared to pre-HAART therapy (25, 26). However, in many geographic regions of the world, antibiotic or HAART therapies are not widely available, and hence, toxoplasmosis continues to be a significant problem in human immunodeficiency virus (HIV) patients, as cited, for example, in recent studies from Italy (3), Brazil (32), Malaysia (24), and Ethiopia (4). *T. gondii* has a highly unusual population structure consisting of three clonal lineages known as types I, II, and III, which are widespread in North America and Europe (5, 14). Studies of genetic polymorphism reveal that at each locus there exist only two alleles, indicating that these three lineages arose from a common source and that they have since undergone limited genetic exchange (9, 30). A small number (less than 5%) of isolates have mixtures of the two-allele patterns seen in the type strains, indicating that they are natural recombinants. Even less common are exotic strains, which have many unique polymorphisms, indicating they have a more ancient origin. The clonal lineages share the trait of direct oral infectivity to intermediate hosts, thus bypassing the need for sexual recombination (28, 30). Direct oral transmission presumably led to the recent emergence and widespread success of the clonal lineages. It has recently been suggested that a wider range of genotypes might be found in geographically remote areas where transmission may rely more often on the sexual phase, which only occurs in cats (1).

Studies with mice have shown that infections with the different clonal lineages of *T. gondii* result in very different outcomes. Type I strains are highly virulent, whereas type II and type III are relatively nonvirulent (29). The enhanced virulence of type I strains is in part due to overstimulation of a Th1 immune response that leads to pathology (23). However, it remains uncertain to what extent the genotype of the parasite directly contributes to the clinical severity of human toxoplasmosis. Comparison of the genotypes of *T. gondii* strains associated with animal versus human infections showed an increased frequency of type I strains in AIDS patients, although a majority of infections were caused by type II strains (14). Type I strains were also frequently associated with recurrent ocular toxoplasmosis observed in patients that were immunocompromised (10). These studies suggest that type I strains are more pathogenic or more likely to cause infection in immunocompromised patients. However, in most studies centered in North America and Europe, type II strains are the most prevalent cause of human toxoplasmosis both in congenital infection (2, 5, 14) and in AIDS patients (13, 14).
The continued expansion of HIV on a global level has increased the need for more sensitive and efficient diagnostic tests for opportunistic infections. Nested PCR (nPCR) has been used widely in different studies for sensitive diagnosis of toxoplasmosis (22, 27). In most cases, the sensitivity of specific detection of the assay is of primary concern. However, most of the highly repetitive markers that provide sensitive detection (i.e., B1, rRNA genes) are not useful for genotyping the parasite due to a relative absence of polymorphisms. We have previously described an nPCR assay for the SAG2 locus that allows sensitive typing of all three lineages, and this has now been widely adopted by other groups (13). However, the limitation of using a single locus for typing is that mixed strains or exotic genotypes will be misclassified as having a simple genotype. Mixed genotypes may be more common in some regions, and such strains may present different clinical complications following infection in humans. Consequently, in the present study, we developed four independent nPCR markers to provide rapid and specific diagnosis of toxoplasmosis in clinical samples.

**MATERIALS AND METHODS**

**Clinical isolates.** We analyzed 11 clinical samples (designated WU01 through WU11) collected during the period of 1990 to 1996. Analysis was performed retrospectively on cerebral spinal fluid (CSF) samples that had been collected for retrospective on cerebral spinal fluid (CSF) samples that had been collected for

**Determining the genotypes of toxoplasmosis in clinical samples.** We analyzed 11 clinical samples (designated WU01 through WU11) collected during the period of 1990 to 1996. Analysis was performed retrospectively on cerebral spinal fluid (CSF) samples that had been collected for retrospective on cerebral spinal fluid (CSF) samples that had been collected for

**Genotyping of isolates by PCR-RFLP.** To determine the genotypes of *T. gondii* in clinical samples, we developed highly sensitive nPCR markers from four different genetic markers: 5′-SAG2, 3′-SAG2, BTUB, GRA6, and SAG3. The presence of sequence polymorphism within each locus results in distinctive restriction fragment length polymorphism (RFLP) patterns that were used to assign alleles for each type strain. Multiplex PCR using external primers for four different markers was used for the initial round of amplification. PCR was carried out in a 50-μl reaction mixture consisting of 5 μl of 10× PCR buffer without MgCl2 (Sigma, St. Louis, MO), 4 μl deoxyribonucleotide triphosphates (2.5 μM each), 3 μl of 25 mM MgCl2, 1.5 μl each primer (50 μM), and 0.5 μl AmpliTaq (5 U/μl) (Sigma). Amplification was conducted for 40 cycles at an annealing temperature of 55°C. PCR-amplified products (5 μl) were then used for second-round amplification of each marker separately in a 25-μl-volume reaction mixture (all components were at the same concentrations as given above) using primers given in Table 1. Positive controls consisted of cell lysate from *T. gondii* type I (RH), type II (ME49), or type III (CTG) strains. Negative controls consisted of DNA-free water and proteinase K-treated cell lysate of noninfected HFF cells. The amplified fragments were digested with appropriate restriction enzymes for different markers, and the fragments were analyzed by 3% agarose gel electrophoresis, stained with ethidium bromide, and imaged by Alpha Imag-ger, version 5.5 (Alpha Innotech Corp., San Leandro, CA). Estimates of the sizes of fragments were based on comparison to dX 174 DNA digested with HaeIII (New England Biolabs, Beverly, MA).

**Intron sequencing.** To determine the relative divergence between *T. gondii* strains, the frequency of single-nucleotide polymorphisms (SNPs) present in the UPRT-1 intron sequence (GenBank accession no. YA134132) was analyzed (30). Nested primers were used to amplify the genomic region flanking this intron: external primers, UPRT1 (EXT) F (5′-TCCACAGGCTCTTGTTAAT-3′) and UPRT1 (EXT) R (5′-GAAGTACGGAACAGCTGCA-3′); internal primers, UPRT-F1 (5′-CCGGATATCCCGTCAAGG-3′) and UPRT-R1 (5′-GGAGTACAGCTGTTTGACTG-3′). PCR products were purified using Quick PCR purification kit (QIAGEN) according to the manufacturer’s instructions. Sequencing was done from three independent templates using BigDye cycle sequencing (Applied Biosystems, Foster City, CA) conducted by SeqWright DNA Technology Services, Houston, TX). Nucleotide sequences were analyzed using Vector NTI, version 9 (Informax Introgen Life Sciences, Carlsbad, CA). CLUSTALX (12) was used to align the sequences for comparison.

**RESULTS**

To provide for highly sensitive genotyping of *T. gondii*, we developed four independent nPCR markers, SAG2, BTUB, and...
GRA6, and SAG3, which are located on chromosomes VIII, IX, X, and XII, respectively (Table 1) (16). Because clinical samples are often available in limited quantity, a multiplex nPCR was employed using external primers for all four loci in the initial round of amplification. Internal primers were then selected to separately amplify the four loci. To test the sensitivity of the nested PCR assay, normal human CSF was spiked with freshly isolated parasites (1, 2.5, 5, and 10 parasites) and used for multiplex PCR amplification, followed by nested PCR amplifications for gene-specific markers. All of the nested PCR markers shown in Fig. 1 were readily amplified from CSF samples spiked with 10 parasites (Fig. 1). SAG3 was also detected when as few as five parasites were spiked in human CSF (Fig. 1). Negative controls remained free of amplified products (Fig. 1).

The nested PCR assay for the SAG2 locus (13) and SAG3 locus (11) have been described previously. The combination of polymorphisms in the 5' and 3' ends of the SAG2 gene allow typing of all three lineages based on a unique combination of alleles. The design of nPCR markers for GRA6 (8) was modified from a previous study that defined polymorphism in this gene. Each of these markers contains several closely spaced polymorphisms that result in three-way typing with a single restriction enzyme (Fig. 2). Digestion of the amplified product of BTUB with BsiEI distinguished genotype I (allele 1) from genotypes II and III (allele 2) and with TaqI distinguished genotype II (allele 2) from genotypes I and III (allele 1) (Fig. 2). Consequently, each of the four markers can unambiguously detect all three genotypes based on distinct RFLP patterns.

We examined 11 CSF samples collected during the period of 1990 to 1996 from immunocompromised patients (HIV positive) patients. Six of these patients were clinically diagnosed as having confirmed (WU01, WU03, and WU05) or suspected (WU04, WU07, and WU08) TE (Table 2). Classification was based on having two or more of the following criteria: (i) positive serological response to *T. gondii*, (ii) single or multiple contrast-enhancing lesions on computed tomography or magnetic resonance scan, (iii) progressive cognitive or motor impairment, and (iv) response to anti-*T. gondii* therapy. One HIV-positive patient with no history of toxoplasmosis (WU02) but with lymphoma was selected as a negative control. Multiplex nPCR was used to detect and identify the genotypes of *T. gondii* present in these clinical samples.

Following the initial multiplex PCR using external primers, no amplification products were observed (data not shown), indicating that the amount of parasite DNA present in these clinical samples was likely very low. In the second round of amplification, 8 of 11 samples gave positive amplicons for one or more of the markers (data not shown) using separate sets of internal primers for the four nPCR markers. All of the confirmed and suspected cases of TE were positive for at least one marker. WU02, which was a negative control, was consistently negative for all markers. Two of the patients for whom TE was not a suspected diagnosis (WU10 and WU11) also remained negative in all analyses. No products were detected from negative controls (water and proteinase K-treated cell lysate of noninfected fibroblast cultures) used to check for contamination.

Following restriction digestion and gel electrophoresis of the products, the strains were classified based on the alleles shown in Fig. 2. Six (WU01, WU03, WU04, WU07, WU08, and WU09) of 8 positive clinical samples possessed type I alleles at some or all of the markers (Table 2). Four of these samples were supported by two or more independent markers, whereas the remaining 2 samples were only supported by the marker SAG3. SAG3 is the most sensitive marker within this set (Fig. 1), thus it is typically the easiest to detect when parasite DNA is limiting. Due to limited sample amounts, we were not able to successfully amplify and type the remaining loci for these samples. WU06 showed a type III genotype in all four nPCR markers. Interestingly, sample WU05 showed an unusual genotype and possessed type I (at SAG2 and BTUB loci), type II (at the GRA6 locus), and type III (at the SAG3 locus) alleles at different loci (Fig. 2; Table 2). These results gave the first indication that the *T. gondii* strain present in WU05 carried a novel nonarchetypal genotype.

The true extent of polymorphism between alleles cannot be fully assessed by RFLP analysis, thus we determined the sequence of a polymorphic intron in the UPRT gene as described previously (30). The frequency of SNPs present in the UPRT-1 sequence of WU05 was compared to the archetypal clonal genotypes, and it was found to contain the normal biallelic pattern and to have four novel SNPs (Table 3). This finding
indicates that WU05 contains sufficient novel SNPs in its genome to classify this strain as exotic. By comparison, intron analysis of WU08 revealed a type III sequence without evidence of any additional mutations (Table 3).

**DISCUSSION**

We have developed a sensitive nPCR assay for multiplex analysis of the genotypes of *T. gondii* in clinical samples. Application of this method to CSF samples from patients with confirmed or presumptive TE revealed a striking percentage of strains with a type I genotype. In addition, the combination of 4 independent markers was useful for identifying a mixed strain, which upon further analysis turned out to have an exotic genotype. The multiplex nPCR method is highly sensitive and specific and provides comprehensive genotyping from small amounts of clinical material. Application of this methodology to a broader set of samples should be useful in classifying strains of *T. gondii* that cause toxoplasmosis.

Genotyping of *T. gondii* samples from clinical infections is complicated by the chronic nature of the infection, which is characterized by semidormant tissue cysts and an absence of circulating parasites. For these reasons, we lack good baseline data on the strains of parasite that cause clinically unapparent infections in humans (the majority of cases). While recent attempts to develop serological tests for genotyping infections show promise (17), they lack the ability to unambiguously identify all three lineages and they have not yet been widely applied. Furthermore, the kinetics of parasite dissemination during infection is poorly understood, and there are no systematic protocols for reliably detecting parasites in patients with toxoplasmosis. One of the advantages of the method described here is that it can be conducted on small samples of CSF, amniotic fluid, or Buffy coats, typically using materials that are collected in the course of performing other clinical tests. The sensitivity of the multiplex nPCR method described here is estimated to be between 5 to 10 parasites, which makes it well suited to clinical

**FIG. 2.** RFLP analysis of type strains and a mixed clinical isolate of *T. gondii* using 4 nested PCR markers. Agarose gel electrophoresis of undigested and restriction-digested products for type strains (type I RH, type II Me49, and type III CTG; lanes I, II, and III, respectively) and a mixed-genotype isolate (lanes WU05) is shown. Products were resolved on 3% agarose gels stained with ethidium bromide. Lanes: Neg, negative PCR control; M, molecular mass markers corresponding to /H9278 digested with HaeIII. SAG2 typing was based on a previously reported method (13). GRA6 typing was adapted from a previously reported method (8). SAG3 typing was adapted from a previously reported method (11).
samples containing small numbers of organisms. Thus, wider application of this methodology may yield a broader data set for drawing conclusions about the occurrence of toxoplasmosis associated with particular genotypes of the parasite.

Previous studies have shown that the type II lineage is the most prevalent in animals and humans (5, 13, 14). Since toxoplasmosis is acquired as a food-borne or waterborne pathogen (20), infections in agricultural animals and cats provide the reservoir for human infection. Type II strains are the most abundant, while type I strains are relatively rare in agricultural and wild animals from North America (7, 14, 21). However, an elevated frequency of type I strains is observed in some studies of immunocompromised patients. For example, Howe and Sibley reported a slight increase in the frequency of type I strains in HIV patients with advanced TE (14). A study of recurrent retinal toxoplasmosis in immunocompromised patients also reported an elevated frequency of type I strains or strains bearing type I alleles (11). A larger study of immunocompromised patients may reveal if this pattern is really due to increased pathogenicity of type I strains or simply due to small sample sizes. Type I strains are capable of causing lethal infection in mice (23, 29), and this trait has been linked to a parasite gene on chromosome Vila that is conserved among all type I strains examined (31).

The majority of samples examined in the present study were of the type I genotype or at least contained the type I alleles. Six of 8 positive samples contained type I alleles (75%), while the frequency of type I strains in animals is in the range of 5 to 10% (13, 21). For several of the samples, only a single marker (SAG3) was positive in the PCR analysis; consequently, it is possible that these strains have a mixed genotype consisting of some type I alleles. While we cannot be sure of the reason for the high frequency of type I alleles and strains, it is quite likely that the samples studied here contain a selection bias, since the samples were obtained due to progressive neurological deterioration.

All of the patients in the present study were receiving antiviral therapy, although many of the patients had advanced AIDS and likely concomitant reduced immune function. Due to the retrospective nature of the study, it was not possible to access patient records for all of the samples. However, from the available data, the range of clinical severity of the patients was substantial, as some patients had severe TE characterized by multifocal lesions (WU01, WU03, and WU05). Others had a single lesion (WU04 and WU08) and, in some cases, were not suspected of having TE initially (patient WU08, for example, had a negative Toxoplasma serology). It is unlikely that the positive result obtained here by PCR is due to a false positive, as we never observed contamination in any of the control reactions. However, sequencing the UPRT-1 intron for this isolate revealed a type III sequence. These data are consistent with this strain being a natural recombinant between the type I and III lineages. Due to the sensitivity of the nPCR assay described here, it may also provide a rapid means of establishing the presence of T. gondii DNA in clinical samples.

### TABLE 2. Genotypes of human toxoplasmosis samples tested by multiplex nested PCR

<table>
<thead>
<tr>
<th>Sample</th>
<th>HIV status</th>
<th>Diagnosis</th>
<th>Toxoplasma serology</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>WU01</td>
<td>+</td>
<td>TE, multiple lesions</td>
<td>IgG(^a) positive</td>
<td>I—III</td>
</tr>
<tr>
<td>WU03</td>
<td>+</td>
<td>TE, large, multiple lesions</td>
<td>IgG positive, IgM negative</td>
<td>I—III</td>
</tr>
<tr>
<td>WU05</td>
<td>+</td>
<td>TE, multiple lesions, ring enhancing</td>
<td>IgG positive</td>
<td>I—III</td>
</tr>
<tr>
<td>WU04</td>
<td>+</td>
<td>TE, single lesion</td>
<td>NA(^a)</td>
<td>I—III</td>
</tr>
<tr>
<td>WU07</td>
<td>+</td>
<td>Small lesions, white matter</td>
<td>NA</td>
<td>I—III</td>
</tr>
<tr>
<td>WU08</td>
<td>+</td>
<td>Single lesion, nonenhancing</td>
<td>Negative</td>
<td>I—III</td>
</tr>
<tr>
<td>WU06</td>
<td>+</td>
<td>NA</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WU09</td>
<td>+</td>
<td>NA</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WU10</td>
<td>+</td>
<td>NA</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WU11</td>
<td>+</td>
<td>NA</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>WU02</td>
<td>+</td>
<td>Lymphoma</td>
<td>Negative</td>
<td>I—III</td>
</tr>
</tbody>
</table>

\(^a\) Genotypes I and II are the same.  
\(^b\) Genotypes I and III are the same.  
\(^c\) –, negative amplification product.  
\(^d\) This strain has a recombinant genotype based on intron sequence data shown in Table 3.  
\(^e\) NA, data not available.  
\(^f\) IgG, immunoglobulin G.

### TABLE 3. Summary of polymorphisms in UPRT-1 intron 1 from T. gondii strains

<table>
<thead>
<tr>
<th>Allele type</th>
<th>Strain</th>
<th>Source</th>
<th>5’ SAG2(^a)</th>
<th>3’ SAG2(^b)</th>
<th>BTUB</th>
<th>SAG3</th>
<th>GRA6</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>RH</td>
<td>Human</td>
<td>176</td>
<td>197</td>
<td>207</td>
<td>335</td>
<td>385</td>
<td>389</td>
</tr>
<tr>
<td>III</td>
<td>CTG</td>
<td>Cat</td>
<td>C</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>T</td>
</tr>
</tbody>
</table>

\(^a\) Nucleotides are shown at the numbered positions that contained one or more SNPs. The total intron size is 443 bp, and positions not indicated were identical for all strains. Size is based on the noncoding region including the GT/AC splice site. Biallelic polymorphisms are found in the type strains and WU08, which has a type III genotype. WU05 contains biallelic polymorphisms shared with the clonal types and also contains 4 unique SNPs (bold).
We previously described an nPCR method for genotyping the three predominant lineages of *T. gondii* based on the SAG2 locus (13). While this marker provides accurate genotyping of the majority of strains, it is not capable of detecting recombinant or exotic strains. The advantage of using the combination of 4 independent markers in the present study is that they are much more likely to detect recombinant genotypes. Recombination between the clonal lineages occurs rarely in the wild and is seen in 1 to 5% of strains sampled. These strains contain the archetypal biallelic pattern, but they have been shuffled due to genetic recombination that occurs via meiosis following co-infection in the cat. A second class of strains is typified by having a significantly greater level of polymorphism, and these stains are thought to have a recombinant genotype based on the combination of alleles at 4 different makers. However, when the intron from the UPRT-1 gene was sequenced, this strain was supported by NIH grants AI38858, NS32228, and NIMH-22005.  

ACKNOWLEDGMENTS  

We thank Daniel K. Howe (Veterinary Science Department, University of Kentucky), who performed assisted studies; Bill Campbell and Bill Mulford, who assisted in sample acquisition; and Julie Sueterlin, who provided technical assistance.  

This work was supported by an NIH grant (AI36629) to L.D.S. B.D.C. was supported by NIH grants AI38858, NS32228, and NIMH-22005.  

REFERENCES  


This work was supported by an NIH grant (AI36629) to L.D.S. D.B.C. was supported by NIH grants AI38858, NS32228, and NIMH-22005.  


ACKNOWLEDGMENTS  

We thank Daniel K. Howe (Veterinary Science Department, University of Kentucky), who performed assisted studies; Bill Campbell and Bill Mulford, who assisted in sample acquisition; and Julie Sueterlin, who provided technical assistance.  

This work was supported by an NIH grant (AI36629) to L.D.S. B.D.C. was supported by NIH grants AI38858, NS32228, and NIMH-22005.