THE ROYAL SOCIETY

Published online 18 January 2002

Genetic approaches to studying virulence and pathogenesis in *Toxoplasma gondii*

L. David Sibley*, Dana G. Mordue, Chunlei Su, Paul M. Robben and Dan K. Howe†

Department of Molecular Microbiology, Washington University School of Medicine, St Louis, MO 63110, USA

Toxoplasma gondii is a common protozoan parasite that causes disease in immunocompromised humans. Equipped with a wide array of experimental tools, *T. gondii* has rapidly developed as a model parasite for genetic studies. The population structure of *T. gondii* is highly clonal, consisting of three distinct lineages that differ dramatically in virulence. Acute virulence is probably mediated by the genetic differences that distinguish strain types. We have utilized a combination of genetic approaches to investigate the acute virulence of toxoplasmosis using the mouse model. These studies reveal the surprising finding that pathogenicity is due to the over-stimulation of normally protective immune responses. Classical genetic linkage mapping studies indicate that genes that mediate acute virulence are linked to chromosome *VII* in the parasite. To increase the resolution of genetic mapping studies, single-nucleotide polymorphisms are being developed based on an extensive database of expressed sequence tags (ESTs) from *T. gondii*. Separately, DNA microarray studies are being used to examine the expression of parasite and host genes during infection. Collectively, these approaches should improve current understanding of virulence and pathogenicity in toxoplasmosis.

Keywords: genome; pathogenesis; linkage mapping; toxoplasmosis; cytokines

1. INTRODUCTION

Toxoplasmosis is a classic zoonosis. A wide range of vertebrate animals serve as hosts and while human infections are common, serious complications occur primarily in immunocompromised hosts. Human infections are caused by accidental ingestion of oocysts, shed into the environment by cats, or tissue cysts contained in undercooked meat. Infections in healthy adults are generally benign, although toxoplasmic retinitis is frequently a cause of serious eye disease in otherwise healthy adults (Gilbert *et al.* 1999). More profound disease occurs in immunocompromised hosts (Luft *et al.* 1993) or as the result of congenital infections (McLeod *et al.* 2000). The factors that control disease severity are not well understood, although a variety of components including host genotype, nutrition, immune status, infection load and parasite genotype have been suggested as important in influencing the outcome of infection.

Following initial infection, the parasite proliferates mitotically as a rapidly growing form called the tachyzoite that is responsible for dissemination of the parasite within its host. Tachyzoites are not normally shed into the environment and are relatively sensitive to the low pH of the stomach; hence, they are not directly involved in transmission between animals and humans. The innate and adaptive immune responses that ensue following the initial infection are capable of controlling the growth of the parasite. In response, tachyzoites switch to a slow growing intracellular form. The resulting bradyzoites are capable of long-term survival in the host and are responsible for chronic infections, which may persist for the life of the host. Intracellular cysts containing bradyzoites elicit little inflammation and do not cause significant tissue damage. Instead, clinical complications caused by toxoplasmosis are due to the replication of tachyzoites, either during primary infection or reactivation.

2. CLONALITY AND VIRULENCE OF *TOXOPLASMA GONDII*

One of the most unusual features of *T. gondii* is the markedly clonal population structure of the parasite. Despite the presence of a sexual cycle that occurs exclusively in cats, natural populations of *T. gondii* are primarily clonal, existing as three distinct but closely related lineages (Dardé *et al.* 1992; Sibley & Boothroyd 1992; Howe & Sibley 1995). These three strain types are found in both animals and humans and the parasite genotype does show strong correlations with geographical boundaries or the host range. Type II strains are most common in human toxoplasmosis and as chronic infections in animals used for human consumption (Howe & Sibley 1995). Importantly, while type I strains are relatively rare in animals they occur with increasing frequency in human congenital toxoplasmosis (Fuentes *et al.* 2001) and in ocular disease (Grigg & Boothroyd 2001). These findings suggest that type I strains may be more likely to cause severe disease in humans.

The acute virulence of *T. gondii* strains is well characterized in the mouse model. Type I strains typically have

*Author for correspondence (sibley@borcim.wustl.edu).
† Present address: Gluck Equine Research Center, Veterinary Science Department, University of Kentucky, Lexington KY 40546, USA.

an LD_{100} of a single infectious organism, while types II and III typically have lethal doses of $\geqslant 10^3$ organisms, based on survival following i.p. challenge. Additionally, while type I strains are uniformly lethal, mice that survive challenge with types II and III typically remain chronically infected and seropositive (Howe *et al*. 1996). The extreme virulence of the commonly used RH strain could potentially be due to extensive laboratory passages since its original isolation (Sabin 1941). However, more than 20 recently isolated type I strains also share this highly virulent phenotype in mice, indicating that acute virulence is tightly correlated with the genotype of the parasite (Sibley & Boothroyd 1992).

3. IMMUNE RESPONSES DURING ACUTE TOXOPLASMOSIS

Non-lethal infections with type II or III strains in mice are characterized by rapid induction of protective immunity that is characterized by production of T_H1 (Hunter *et al*. 1996; Denkers & Gazzinelli 1998). The principal components of this response are the production of IL-12 by macrophages, dendritic cells and neutrophils and the subsequent production of IFN-γ by NK cells and T lymphocytes. IFN-γ, in combination with TNF-α, induces formation of reactive oxygen and reactive nitrogen intermediates and leads to nutrient deprivation, thus leading to parasite proliferation control.

Several lines of evidence also suggest that appropriate control of T_H1 cytokines is important for avoidance of immunopathology. For example, C57BL/6 mice that are orally challenged with high doses of the normally non-virulent type II strain Me49, develop severe gut necrosis that is mediated by high levels of TNF-α, IFN-γ and IL-12 (Liesenfeld *et al*. 1996, 1997, 1999). This particular model requires the specific combination of the genotype of the host with the strain and route of inoculum of the parasite. However, other studies indicate that downregulation of T_H1 cytokines during infection is also a general mechanism that is important for protection against immunopathology. For example, IL-10 KO mice challenged with normally non-lethal doses of type II strains of *T. gondii* produce higher levels of IFN-γ and IL-12, which result in a fatal cytokine shock syndrome (Gazzinelli *et al*. 1996). These results suggest that IL-10 normally performs an important role in modulating the degree of the immune response.

4. ANALYSIS OF VIRULENCE OF *TOXOPLASMA GONDII* STRAIN TYPES

The marked virulence of type I strains of *T. gondii* is typified by the commonly used laboratory strain RH where injection of a single viable organism into a mouse leads to death in 8–12 days. While infection with type I strains is uniformly fatal in mice, the specific causes of pathology during the acute infection are poorly understood. There are several possible explanations for this remarkable virulence including:

(i) inadequate triggering of the normally protective immune response;
(ii) refractoriness on the part of the parasite to killing mechanisms of the host;
(iii) increased growth potential of the parasite resulting in greater tissue burdens; and
(iv) release of a lytic or destructive factor that causes direct tissue damage to the host.

Remarkably, none of these hypotheses has been directly addressed experimentally despite the widespread use of type I strains in experimental models.

To examine the basis for the differences in virulence between strains of *T. gondii*, we utilized the mouse model to evaluate acute virulence as monitored by survival following injection of tachyzoites. While not a natural route of infection, this model offers a precise and quantifiable inoculum and a reproducible measure of virulence, calculated as an LD_{50}. Long-term passaged isolates of *T. gondii* strains, which were cloned by limiting dilution, were chosen for comparison: the RH strain as typical of type I strains, PTG, which is a clone of the commonly used type II strain ME49, and CTG, which is a cloned strain isolated from a cat. In outbred mice, these strains have the following lethal doses: RH strain $LD_{100} = 1$, PTG strain $LD_{50} \geqslant 10^3$, CTG $LD_{50} \geqslant 10^3$.

Initially, we considered the possibility that the enhanced virulence of type I strains was simply due to a faster division time. We compared the growth rate of *T. gondii* strains *in vitro* in both murine macrophage and human fibroblast cells. A low inoculum was used (multiplicity of infection of less than 1) to allow maximum expression of differences between strains over several cycles of intracellular infection. These analyses reveal that there is only a slight growth advantage of the type I strain (RH) over type II (PTG) and type III (CTG) strains *in vitro* (figure 1). This slight growth advantage is unlikely to explain the marked differences of these strains in virulence *in vivo*.

To further examine the differences in virulence between *T. gondii* strains, we examined their growth and dissemination in mice following i.p. challenge. The virulent RH strain (type I) and non-virulent PTG strain (type II) both rapidly disseminated to a variety of organs following i.p. challenge (Mordue *et al*. 2001). A typical kinetic profile of tissue burdens from the spleen and lungs of mice infected with lethal doses of RH or PTG strain versus those given a non-lethal infection with PTG is shown in figure 2. Dissemination to the lung, liver, spleen, heart and brain was readily detected within 4 days following i.p. inoculation with as few as 100 RH strain parasites. Tissue burdens for the RH strain reached a level of $\geqslant 10^6\ g^{-1}$ by day 6 post-infection, after which mice rapidly succumbed to infection. Inoculation with lethal doses of PTG (10^5) also reached peak parasite tissue burdens at day 6 post-infection. Inoculation with a low dose (100 parasites) of PTG strain also led to rapid dissemination. Parasites reached levels of approximately 10^5 to $10^6\ g^{-1}$ of tissue by day 8, yet remarkably these animals recovered and tissue burdens began to drop after day 10. Lethal infections were accompanied by an elevated release of serum liver enzymes and pathological changes consistent with hepatocyte damage (Mordue *et al*. 2001). Significantly, this damage occurred systemically and was evident primarily in cells that were not infected with the parasite. Consequently, the markedly different LD_{50} values of *T. gondii* strain types cannot be ascribed to differences in the growth rate, tissue dissemination or the burden of parasites

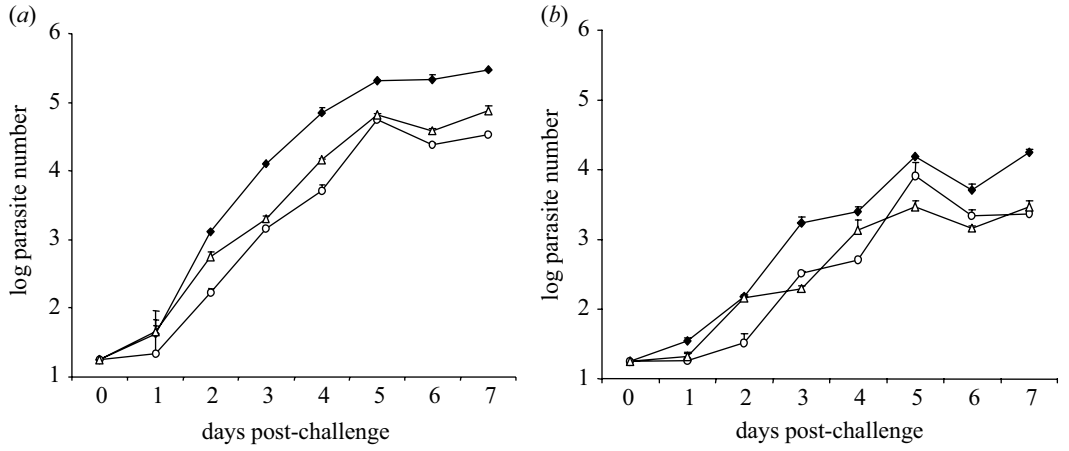

Figure 1. *In vitro* growth rate of *T. gondii* strains. (*a*) Growth rate in human fibroblasts, (*b*) growth rate in mouse macrophages. Parasite numbers were determined by enzymatic content using β-galactosidase-tagged strains as described previously (Dobrowolski & Sibley 1996). Black diamond, type I strain RH; white circle, type II strain PTG; white triangle, type III strain CTG.

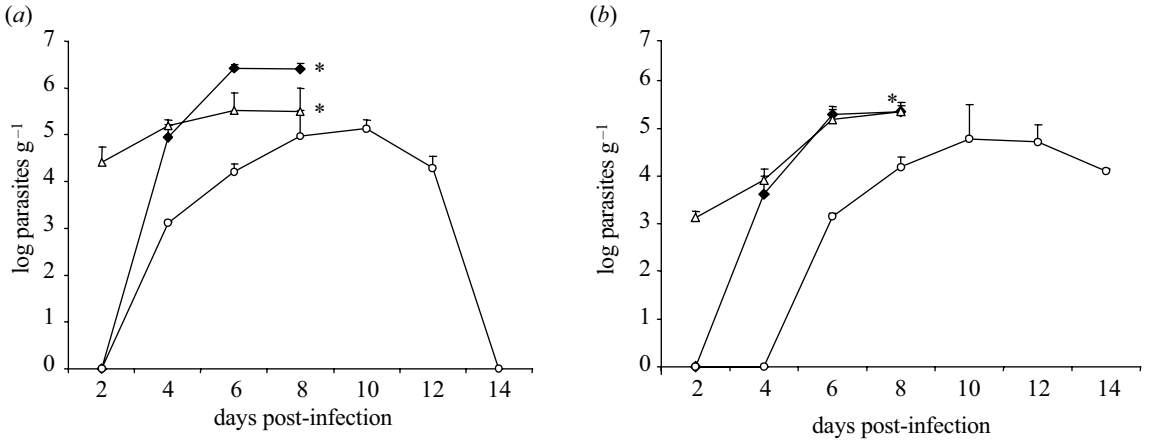

Figure 2. *In vitro* dissemination of *T. gondii* in the spleen (*a*) and lung (*b*). Parasites were injected i.p. and tissues were collected at 2 day intervals. Parasite numbers were calculated as plaque-forming assays on monolayers grown *in vitro* from dilutions of tissue homogenates. The RH strain was given at low doses (100 parasites), while PTG was given at low doses (100 parasites) or high doses (10^5 parasites). Asterisks, Lethal outcome; black diamond, RH strain; white circle, PTG strain (low dose); white triangle, PTG strain (high dose).

present in the tissues. Instead, pathogenesis appears to occur due to the presence or induction of a soluble, inflammatory mediator(s).

5. ACUTE VIRULENCE IS DUE TO OVER-STIMULATION OF T_H1 CYTOKINES

To address the possibility that RH strain infections do not induce an adequate immune response, we examined the induction of T_H1 cytokines in mice following i.p. inoculation. Significant increases were observed in the serum levels of TNF-α, IL-12 for both lethal and non-lethal infections (Mordue *et al.* 2001). By contrast, lethal infections resulted in extremely elevated levels of IFN-γ and IL-18, whereas non-lethal infections resulted in only moderate levels of these cytokines. The over-production of T_H1 cytokines was not due to an absence of production of IL-10, as levels of this cytokine were also extremely elevated during lethal infections. Similar induction of inflammatory cytokines and liver damage is seen in endotoxaemia caused by LPS released from gram-negative bacteria. However, mice that are deficient in LPS signalling (toll-like receptor 4 mutants such as C3H/J mice) are equally susceptible to toxoplasmosis (Mordue *et al.* 2001). Consequently, the lethal outcome of toxoplasmosis in mice is not due to bacterial sepsis or leakage of endotoxin from the gut.

Collectively, these findings reveal that the lethality of infections caused by RH strain is not due to a lack of induction of T_H1 cytokines. In contrast, the extremely high levels of several of these cytokines suggests that mice succumb due to cytokine-mediated pathology. As elevated levels of T_H1 cytokines have previously been implicated in causing pathology, we examined their potential role in acute toxoplasmosis using either KO mice or antibodies that selectively neutralize specific cytokines *in vivo*. In particular, we were interested in evaluating the roles of TNF-α, IL-12, IL-18 and IFN-γ.

High levels of TNF-α are associated with cerebral malaria, a clinically severe form of the disease (McGuire *et al.* 1994), and elevated TNF-α are implicated in pathogenesis of severe malaria (Odeh 2001). Elevated levels of

TNF-α are also implicated in liver damage mediated by LPS in the mouse model (Jaeschke *et al.* 1998). TNF-α plays an important role in resistance to acute and chronic toxoplasmosis; however, the extremely high levels of this cytokine produced during lethal infection could potentially contribute to pathogenesis. Consequently, we examined the effects of neutralizing TNF-α by *in vivo* injection of specific neutralizing antibodies or using TNF receptor KO mice. In both cases, mice infected with lethal doses of RH strain *T. gondii* succumbed with the same kinetics as wild-type mice, indicating that elevated TNF-α is not primarily responsible for death.

Elevated levels of IL-12 have been implicated in a variety of pathological conditions (Trinchieri 1998). In the acute toxoplasmosis model, levels of this cytokine were similarly elevated in both lethal and non-lethal infections, therefore IL-12 alone is unlikely to be responsible for the severe pathology associated with lethal infection (Mordue *et al.* 2001). However, exogenous administration of high levels of IL-12 and IL-18 have previously been implicated in causing pathology and death in mice (Carson *et al.* 2000). IL-18 is an IL-1 like cytokine that mediates its effects by potentiating the effects of IL-12 (Akira 2000), including drive production of IFN-γ. IL-18 has been implicated in protective responses to a variety of bacterial and parasitic infections (Dinarello 1999), although it appears to play a minor role in controlling infections with *T. gondii* (Cai *et al.* 2000). To examine the possibility that elevated IL-18 contributes to pathology in toxoplasmosis, neutralizing antibodies were given to mice following lethal challenge with RH strain parasites. These studies revealed that elevated levels of IL-18 are partially responsible for the lethal outcome of toxoplasmosis in mice (Mordue *et al.* 2001). Surprisingly, the effects of IL-18 in acute toxoplasmosis appear to be independent of IFN-γ, as neutralization of IFN-γ was not able to reverse the pathology or protect mice from lethal infection (Mordue *et al.* 2001).

Collectively, these studies reveal that acute virulence in *T. gondii* occurs by a paradoxical mechanism. Infection with virulent strains induces over-production of T_H1 cytokines, which are normally protective. The unregulated production of these factors, principally IL-18, and to a lesser extent IFN-γ, IL-12 and TNF-α, results in lethality. While the actual cause of death is unknown, it is probable that the extremely elevated levels of these cytokines cause increased vascular permeability that ultimately leads to multiple organ failure.

6. IDENTIFICATION OF VIRULENCE GENES BY LINKAGE MAPPING

The availability of a classical genetic linkage map for *T. gondii* makes it possible to perform classical genetic studies to search for unknown genes that determine specific phenotypes (Sibley *et al.* 1992). One complication of this approach is the necessity to perform genetic crosses in cats. Nonetheless, a pair of crosses between a cloned isolate of the type II strain ME49 and the type III strain CEP were utilized to examine the segregation of approximately 70 polymorphic markers consisting of RFLP (Sibley *et al.* 1992). These studies established that *T. gondii* has 11 chromosomes that segregate randomly and that the average map unit (distance over which 1% recombination occurs) is *ca.* 300 kb. The haploid genome consists of *ca.* 8×10^7 bp, while the available genetic map comprises 150 map units (the equivalent of 4. 5×10^7 bp). Whether this discrepancy is due to inaccuracy in how the genome size was calculated (relative fluorescence intensity of DNA binding fluorochromes) or an incomplete genetic map, is presently unclear. The relatively large map unit in *T. gondii* facilitates mapping of loci that control phenotypes such as drug resistance to specific chromosomes using a small number of progeny. However, it also limits the ability to precisely localize genes due to the large number of progeny that would have to be analysed to find informative crossovers.

The dramatically different virulence of *T. gondii* strain types can be measured as an LD_{100} (the minimal dose which causes 100% mortality) to provide a phenotype that has the following useful features for genetic mapping:

(i) the magnitude of the differences is more than 2 log while the variance of measurements is generally less than 0.5 log;
(ii) the trait is stable with repeated passage;
(iii) the trait is shared by multiple isolates of the same genotype, indicating it probably has a basis in the underlying genetic differences between strains.

To explore the genetic basis of acute virulence in *T. gondii*, we crossed a highly virulent type I strain called GT-1 (Dubey 1980) with the non-virulent type III strain CEP. GT-1 has a virulent phenotype like RH, where a single infectious organism is uniformly lethal while CEP has an LD_{50} of *ca.* 10^3 in outbred mice. Both strains are low passage isolates that are fully capable of completing the life cycle. The cross was performed by co-infecting cats with brain homogenates from chronically infected mice, which harbour cysts of the parasite. Analysis of the resulting progeny was facilitated using drug resistance markers that were first introduced into the parental lines. These markers were used to select doubly resistant progeny, which are the result of a sexual cross between the strains as opposed to progeny that arise by self mating (Pfefferkorn *et al.* 1977).

Analysis of the genotypes of recombinant progeny was conducted by converting RFLP markers to probes that can be readily analysed by PCR amplification. An example of such analysis is shown in figure 3 that demonstrates the two alleles at the *SAG3* locus for the parental types and several recombinant progeny. Separately, each clone was tested to establish the virulence in mice following i.p. inoculation. By comparing the segregation of markers that are located at known map positions on specific chromosomes with virulence, it is possible to analyse the genetic contribution of acute virulence. These preliminary studies have revealed several important findings. First, acute virulence is clearly a heritable trait, that is clones exhibit stable virulence phenotypes that reflect one or another parental line. Second, in addition to parental phenotypes (virulent or non-virulent), some clones demonstrated intermediate phenotypes. This result indicates that acute virulence is multigenic.

Previously, we have examined natural recombinants of *T. gondii* and described an association between a locus that is linked to *SAG1* on chromosome *VIII* and acute virulence (Howe *et al.* 1996). However, it was not clear from

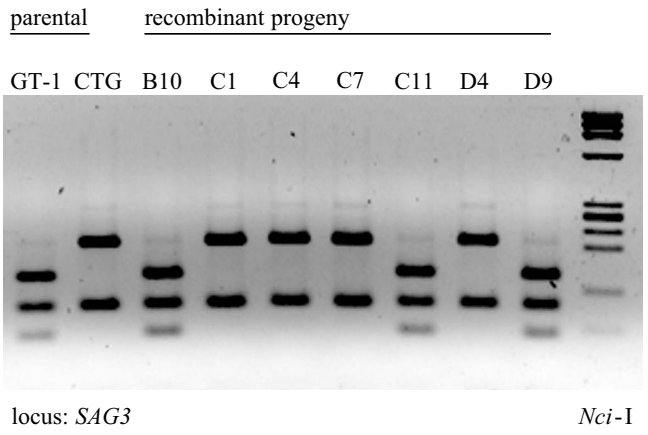

Figure 3. RFLP mapping of recombinant progeny. The *SAGr3* locus was analysed by PCR amplification followed by digestion with *Nci*-I. Markers correspond to ϕX 174 digested with *Hae* III.

these studies if this linkage was fortuitous or due to the presence of a gene(s) in this region that contributed directly to acute virulence. Analysis of progeny from the type I × III cross revealed that there is no significant linkage with markers on chromosome *VIII*. Instead a statistically significant linkage was identified with one end of chromosome *VII* (χ^2, $p = 0.0008$). Since acute virulence can be quantified, it is well suited for quantitative trait mapping, a process that is able to assign a relative contribution to different genetic loci that control complex traits (Lander & Botstein 1989). However, one of the main limitations to genetic mapping in *T. gondii* is the low resolution of the present genetic map.

7. EST SEQUENCING FOR GENE DISCOVERY

Random sequencing of cDNAs to generate ESTs provides a rapid, inexpensive method of gene discovery (Adams *et al.* 1991). EST sequencing is particularly useful in systems with complex genomes, which are often interrupted by introns or non-coding regions. Parasites are well suited for EST analysis as they fit these criteria and are often understudied, hence few specific genes are known. An example of the power of EST sequencing is provided by the efforts to sequence approximately 10 000 randomly selected cDNAs from *T. gondii*, a process that successfully identified more than 2000 putative genes (based on BLASTX similarities of $e \leqslant 10^{-6}$) (Ajioka *et al.* 1998; Manger *et al.* 1998*b*). By comparison, prior to this effort there were less than 30 genes for *T. gondii* in the public gene databases. Thus, EST sequencing rapidly identified a large number of candidate genes for further study, including those related to conserved functions and an important class of genes that are unique to the Apicomplexa (Ajioka *et al.* 1998).

EST analysis also provides useful data on the expression level of genes that are moderately or abundantly expressed. The most highly expressed genes in *T. gondii* tend to be parasite specific, or those related to the unique biology of the Apicomplexa. Among these are secretory proteins such as the *GRA* genes and the surface antigens such as the *SAG*-related genes. Additionally, a number of unknown genes are highly expressed (no homology to any other entry in the databases) and these presumably play important roles in the biology of the parasite. For highly expressed genes, ESTs analysis offers the possibility of identifying the complete ORF by constructing an overlapping consensus sequence from multiple ESTs. These reconstructed ORFs are valuable for their ability to rapidly identify the complete coding sequence of a new gene. They can also be used to identify alternatively spliced messages and gene families consisting of related genes. Additionally, ORFs that are validated from consensus EST sequences are useful for training the algorithms that are currently used for predicting gene boundaries from genomic DNA sequences (i.e. GLIMMER, GENEFINDER, HEXAMER). Application of these ORF- predicting programs to the malarial genome sequence currently results in numerous different predictions for a given region, highlighting the need for additional EST sequences from *Plasmodium falciparum*.

Sequencing cDNAs present in libraries made from different life cycle stages allows identification of genes that are specifically expressed in one stage versus those which are constitutively expressed. For example, among the approximately 16 000 *T. gondii* ESTs currently in GenBank (as of April 2001), approximately 12 700 are from the tachyzoite stage and 3300 are from the bradyzoite stage of the parasite. Analysis of the abundant ESTs reveals that a significant number of these genes are expressed in a stage-specific manner. Interestingly, among this set of genes are members of the surface antigen genes related to *SAG1*. This family consists of genes with overall low-sequence identity but which contain a conserved spacing and number of cysteine residues (Manger *et al.* 1998*a*,*b*). The major surface antigens of tachyzoites and bradyzoites consist of different members of this gene family, which may be involved in immune evasion. Current efforts are underway to expand the stage-specific EST analysis to include the sporozoite stage of the life cycle using recently constructed sporozoite or oocyst libraries (M. White, unpublished data).

While EST sequencing provides a rapid means of gene discovery and comparative data on expression levels, it is unable to identify genes that are expressed at low levels or during limited times in the life cycle. These genes can only be identified by a full genome sequencing effort; however, for many parasitic organisms, their limited economic and health importance may not currently warrant such expensive and comprehensive efforts. By contrast, EST sequencing offers a wealth of data for a relatively modest investment. For this reason, we have recently embarked on comparative EST projects for several apicomplexan parasites including *Eimeria tenella*, a pathogen of poultry, *Sarcocystis neurona*, which causes paralysis in horses, and *Neospora caninum*, which causes abortion in cattle. Additionally, EST projects have been undertaken for *Plasmodium* (Chakrabarti *et al.* 1994) and *Cryptosporidium*. Comparison of the most highly expressed genes from these taxa reveals several interesting trends. First, among the highly expressed genes of this group are actin and actin-regulatory proteins, which attests to the importance of these components in the active motility and invasion mechanism employed by these parasites. Second, secretory proteins found in micronemes, which typically function in adhesion, are highly conserved, such as the TRAP

family. Finally, surface antigen encoding genes are more restricted. The *SAG*-related genes are found in the closely related parasites *T. gondii*, *N. caninum* and *S. neurona*, but so far have not been identified in *Plasmodium* or *Eimeria*. Further comparative EST analyses promises to provide many important clues about genes that are important for members of this phylum including those which control specialized features such as host range.

8. ANALYSIS OF STRAIN-SPECIFIC DIFFERENCES IN GENE EXPRESSION AND ALLELIC DIFFERENCES

The *T. gondii* EST database also allows for comparisons between strain types as the sequences are distributed between the prototypic strains RH (type I), Me49/PTG (type II) and VEG (type III). By comparing the frequencies of moderately abundant ESTs between the different strain types, it is possible to identify differences in gene expression. Such genes might be expected to underlie the important differences in virulence between strain types. Remarkably, only one strain-specific gene was identified by this means. The clone TgESTzy64e04 (AF323780) represents a cluster of 29 related ESTs in PTG (type II) but only four ESTs each in RH (type I) or VEG (type III) strains. Preliminary analysis of this gene does not reveal any similarity with other genes in the public databases, therefore its potential role is uncertain. Among the remaining ESTs, there do not appear to be major differences between strains in the compliment or expression of genes. However more sensitive techniques such as SAGE (Velculescu *et al.* 1995), or microarray analysis may identify other candidate genes that might be differentially expressed between *T. gondii* strains.

Comparison of sequences for a given gene among different strains *of T. gondii* has also revealed an abundance of SNPs. On average, the genomes of the three lineages diverge by 1–2% for coding regions such as *SAG1*. Sequencing of approximately 34 different loci (a total of 24 365 bp) has revealed a frequency of a SNP every 150–250 nt (C. Su, unpublished data). Surprisingly, these markers are dimorphic, that is there are only two alleles at every locus, despite the presence of three strain types. This pattern has previously been pointed out using RFLP markers (Sibley *et al.* 1992), and it implies that the three clonal types are recently derived from two distinct parental types.

For genes that are moderately to highly abundant, it is possible to establish a consensus allele for each strain from the EST database. An example of this analysis is shown in figure 4 for the highly expressed gene *GRA1*. Within this region of the sequence, the presence of an A versus G at a single position changes a *Nla*III restriction site to a *Mae*II site. Thus, by amplifying this region using flanking primers and testing for the presence of a restriction site, it is possible to type the allele at this locus based on this single nucleotide substitution.

There are presently almost 200 SNPs that can readily be identified in the existing EST database for *T. gondii*; however, only about 45% of these distinguish RFLPs. Previous studies have established that SNPs can be analysed by gel-based sequencing or oligoinucleotide hybridization, thus providing a sensitive means of typing human SNPs

accession no.	sequence
1164790	CAATGTTAAC**A**TGGAGGA
487944	CAATGTTAAC**G**TGGAGGA
574095	CAATGTTAAC**A**TGGAGGA
571465	CAATGTTAAC**A**TGGAGGA
607858	CAATGTTAAC**A**TGGAGGA
1165586	CAATGTTAAC**A**TGGAGGA
1164743	CAATGTTAAC**A**CGGAGGA
1165494	CAATGTTAAC**A**TGGAGGA
915340	CAATGTTAAC**A**TGGAGGA
487556	CAATGTTAAC**G**TGGAGGA
488019	CAATGTTAAC**G**TGGAGGA
487601	CAATGTTAAC**G**TGGAGGA
*Nla*III	C**A**TG
*Mae*II	AC**G**T

Figure 4. SNP analysis. The presence of a single base change in the GRA1 locus provides a convenient RFLP for typing. Alignment of GRA1 SNPs detected in EST clones form RH (type I) and ME49 (type II) strains of *T. gondii*.

(Wang *et al.* 1998). This method has been further refined to provide for allele-specific primer extension that can be used to specifically amplify the alleles present in a complex sample (Pastinen *et al.* 2000). Hybridization of these labelled probes to allele-specific DNA oligos spotted onto glass slides can then be used to discriminate between the alleles present in the sample (Hirschhorn *et al.* 2000). This method has the advantage of allowing rapid analysis of multiple loci from a single sample. We are presently adapting this technique for typing *T. gondii* strains based on SNPs that have been identified in the EST database. These markers should provide a convenient and powerful approach for genetic mapping and epidemiological studies. For example, expanding the genetic linkage map by incorporation of additional markers will greatly increase the resolution of mapping studies.

9. MICROARRAY EXPRESSION STUDIES

The ability to examine global changes in gene expression has been revolutionized by techniques for generating microarrays of DNAs on glass slides (Shalon *et al.* 1996). The resulting 'DNA chips' contain thousands of spots of individual DNAs that can then be used for simultaneous hybridization with complex mRNAs to evaluate changes in gene expression, for example in human T cells following heat shock or phorbol ester treatments (Schena *et al.* 1996). Two-colour fluorescence detection also allows simultaneous comparison of mRNAs from two samples, for example mRNAs from different tissues of *Arabidopsis* (Schena *et al.* 1996). Application of this technology to yeast reveals global changes in gene expression in response to heat shock–cold shock (Lashkari *et al.* 1997) and following transfer from fermentation to respiration (DeRisi *et al.* 1997). More recently, this comparative analysis has been used to detect genes that differ between yeast strains and that encode a multi-drug resistance locus (Winzeler *et al.* 1998).

Microarray technology is ideally suited for global comparisons of gene expression during the interaction between microbes and their hosts. Often gene expression levels change when the microbe encounters a new niche within

the host and in response, many signalling and response pathways are triggered in the host. Current knowledge of these pathways is presently limited by preconceived ideas about which genes may be of interest combined with the difficulty of analysing particular candidates by traditional means. Microarray studies provide a broad survey without prior assumptions about which genes may be of interest. The use of clustering programs can also highlight genes that change in patterns that are correlated, thus perhaps revealing new insights into the interaction of different pathways. Whole genome arrays have been generated for a variety of pathogens including *Mycobacterium, Escherichia coli, Helicobacter* and others (Cummings & Relman 2000). These arrays have been useful for probing questions about metabolism, virulence and drug resistance (Cummings & Relman 2000). Additionally, comparative hybridization studies have been used to examine global changes in mycobacterial genomes that are related to attenuation of virulence and the maintenance of vaccine stocks (Behr *et al.* 1999).

A number of studies have examined the response of host cells to intracellular pathogens such as *Salmonella* and *Listeria* (Cohen *et al.* 2000; Eckmann *et al.* 2000; Rosenberger *et al.* 2000) using either large filter arrays or DNA arrays printed onto glass slides. Predictably, these studies reinforce the importance of induction of inflammatory mediators as well as identifying unanticipated genes involved in cell-cycle regulation and transcription. Human fibroblasts infected with *T. gondii* also display profound changes in gene expression, albeit in a small number of genes (Blader *et al.* 2001). Among those genes that are upregulated are a variety of inflammatory mediators. More surprisingly, genes involved in glycolysis and in lipid synthesis are also upregulated, implying that intracellular infection alters the basic metabolism of the host, perhaps in a manner to support intracellular growth of the parasite.

While these profiling studies are an important first step, they have thus far not extended beyond the relatively artificial environment of *in vitro* culture models. Using a combination of *T. gondii* ESTs and host genes corresponding to inflammatory mediators, cytokines and signalling pathways, we have recently shown that it is feasible to simultaneously monitor changes in parasite and host genes during infection (P. Robben, unpublished data). This approach will enable comprehensive analyses of changes in gene expression that occur during infection *in vivo* in the mouse model. These studies will allow us to address two key points:

(i) differences in gene expression in the host that occur upon infection with different strains of the parasite; and
(ii) differences in gene expression between different strains of the parasite.

Such global expression studies are designed to be the first step towards formulating hypotheses for further testing. In this regard, both the host, the laboratory mouse, and the parasite, *T. gondii*, provide robust models for genetic, biochemical and cellular analysis of function. These findings will provide useful leads for exploring the roles of parasite virulence and pathogenesis in toxoplasmosis.

The authors are grateful to Tovi Lehman and Michael Grigg for sharing sequence polymorphism data and to Buddy Brownstein for assistance with microarrays. L.D.S. is the recipient of a Burroughs Wellcome Award in Molecular Parasitology. The work described here was partially supported by grants from the National Institutes of Health.

REFERENCES

Adams, M. D., Kelly, J. M., Gocayne, J. D., Dubnick, M., Polymeropolous, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B. & Moreno, R. F. 1991 Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**, 1651–1656.

Ajioka, J. A., Boothroyd, J. C., Brunk, B. P., Hehl, A., Hillier, L., Manger, I. D., Overton, G. C., Marra, M., Roos, D. & Wan, K. L. 1998 Gene discovery by EST sequencing in *Toxoplasma gondii* reveals sequences restricted to the Apicomplexa. *Genome Res.* **8**, 18–28.

Akira, S. 2000 The role of IL-18 in innate immunity. *Curr. Opin. Immunol.* **12**, 59–63.

Behr, M. A., Wilson, M. A., Gill, W. P., Salamon, H., Schoolnik, G. K., Rane, S. & Small, P. M. 1999 Comparative genomics of BCG vaccines by whole-genome DNA microarray. *Science* **284**, 1520–1522.

Blader, I., Manger I. D., Boothroyd, J. C., 2001 Microarray analysis reveals previously unknown changes in *Toxoplasma gondii* infected human cells, *J. Biol. Chem.* **276**, 22 223–22 231.

Cai, G., Kastelein, R. & Hunter, C. A. 2000 Interleukin-18 (IL-18) enhances innate IL-12-mediated resistance to *Toxoplasma gondii*. *Infect. Immun.* **68**, 6932–6938.

Carson, W. E., Dierksheide, J. E., Jabbour, S., Anghelina, M., Bourchard, P., Ku, G., Yu, H., Baumann, H., Shah, M. H. & Cooper, M. A. 2000 Coadministration of interleukin-18 and interleukin-12 induces a fatal inflammatory response in mice: critical role of natural killer cell interferon-γ production and STAT-mediated signal transduction. *Blood* **96**, 1465–1473.

Chakrabarti, D., Reddy, G. R., Dame, J. B., Almira, E. C., Lapis, P. J., Ferl, R. J., Yang, T. P., Rowe, T. C. & Schuster, S. M. 1994 Analysis of expressed sequence tags from *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* **66**, 97–104.

Cohen, P., Bouaboula, M., Bellis, M., Baron, V., Jbilo, O., Poinot-Chazel, C., Galiegue, S., Hadibi, E. H. & Casellas, P. 2000 Monitoring cellular responses to *Listeria monocytogenes* with oligonucleotide arrays. *J. Biol. Chem.* **275**, 11 181–11 190.

Cummings, C. A. & Relman, D. A. 2000 Using DNA microarrays to study host–microbe interactions. *Emerg. Infect. Dis.* **6**, 513–525.

Dardé, M. L., Bouteille, B. & Pestre-Alexandre, M. 1992 Isoenzyme analysis of 35 *Toxoplasma gondii* isolates: biological and epidemiological implications. *J. Parasitol.* **78**, 786–794.

Denkers, E. Y. & Gazzinelli, R. T. 1998 Regulation and function of T-cell mediated immunity during *Toxoplasma gondii* infection. *Clin. Microbiol. Rev.* **11**, 569–588.

DeRisi, J. L., Iyer, V. R. & Brown, P. O. 1997 Exploring the metabolic and genetic control of gene expression on a genetic scale. *Science* **278**, 680–686.

Dinarello, C. A. 1999 IL-18: a T_H1-inducing proinflamatory cytokine and new member of the IL-1 family. *J. Allergy Clin. Immunol.* **103**, 11–24.

Dobrowolski, J. M. & Sibley, L. D. 1996 Toxoplasma invasion of mammalian cells is powered by the actin cytoskeleton of the parasite. *Cell* **84**, 933–939.

Dubey, J. 1980 Mouse pathogenicity of *Toxoplasma gondii* isolated from a goat. *Am. J. Vet. Res.* **41**, 427–429.

Eckmann, L., Smith, J. R., Housley, M. P., Dwinell, M. B. & Kagnoff, M. F. 2000 Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells

in response to infection with the invasive enteric bacteria *Salmonella*. *J. Biol. Chem.* **275**, 14 084–14 094.

Fuentes, I., Rubio, J. M., Ramirez, C. & Alvar, J. 2001 Genotypic characterization of *Toxoplasma gondii* strains associated with human toxoplasmosis in Spain: direct analysis from clinical samples. *J. Clin. Microbiol.* **39**, 1566–1570.

Gazzinelli, R. T., Wysocka, M., Hieny, S., Scharton-Kersten, T., Cheever, A., Kuhn, R., Muller, W., Trinchieri, G. & Sher, A. 1996 In the absence of endogenous IL-10, mice acutely infected with *Toxoplasma gondii* succumb to a lethal immune response dependent on $CD4^+$ T cells and accompanied by overproduction of IL-12, IFN-γ, and TNF-α. *J. Immunol.* **157**, 798–805.

Gilbert, R. E., Dunn, D. T., Lightman, S., Murray, P. I., Pavesio, C. E., Gormley, P. D., Masters, J., Parker, S. P. & Stanford, M. R. 1999 Incidence of symptomatic Toxoplasma eye disease: aetiology and public health implications. *Epidemiol. Infect.* **123**, 283–289.

Grigg, M. E. & Boothroyd, J. C. 2001 Rapid identification of virulent type I strains of the protozoan pathogen *Toxoplasma gondii* by PCR-restriction fragment length polymorphism analysis of the *B1* gene. *J. Clin. Microbiol.* **39**, 398–400.

Hirschhorn, J. N., Sklar, P., Lindblad-Toh, K., Lim, Y. M., Ruiz-Gutierrez, M., Bolk, S., Langhorst, B., Schaffner, S., Winchester, E. & Lander, E. S. 2000 SBE-TAGS: an array-based method for efficient single-nucleotide polymorphism genotyping. *Proc. Natl Acad. Sci. USA* **97**, 12 164–12 169.

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.

Howe, D. K., Summers, B. C. & Sibley, L. D. 1996 Acute virulence in mice is associated with markers on chromosome VIII in *Toxoplasma gondii*. *Infect. Immun.* **64**, 5193–5198.

Hunter, C. A., Suzuki, Y., Subaste, C. S. & Remington, J. S. 1996 Cells and cytokines in resistance to *Toxoplasma gondii*. *Curr. Topics Microbiol. Immunol.* **219**, 113–125.

Jaeschke, H., Fisher, M. A., Lawson, J. A., Simmons, C. A., Farhood, A. & Jones, D. A. 1998 Activation of caspase 3 (CPP32)-like proteases is essential for TNF-α-induced hepatic parenchymal cell apoptosis and neutrophil-mediated necrosis in a murine endotoxin shock model. *J. Immunol.* **160**, 3480–3486.

Lander, E. S. & Botstein, D. 1989 Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185–199.

Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., Brown, P. O. & Davis, R. W. 1997 Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl Acad. Sci. USA* **94**, 13 057–13 062.

Liesenfeld, O., Kang, H., Park, D., Nguyen, T., Parkhe, C., Watanabe, H., Abo, T., Sher, A., Remington, J. & Suzuki, Y. 1999 TNF-α, nitric oxide and IFN-γ are all critical for development of necrosis in the small intestine and early mortality in genetically susceptible mice infected perorally with *Toxoplasma gondii*. *Parasite Immunol.* **21**, 365–376.

Liesenfeld, O., Kosek, J., Remington, J. S. & Suzuki, Y. 1996 Association of $CD4^+$ T cell-dependent, interferon-γ-mediated necrosis of the small intestine with genetic susceptibility of mice to peroral infection with *Toxoplasma gondii*. *J. Exp. Med.* **184**, 597–607.

Liesenfeld, O., Kosek, J. C. & Suzuki, Y. 1997 Gamma interferon induces Fas-dependent apoptosis of Peyer's patch T cells in mice following peroral infection with *Toxoplasma gondii*. *Infect. Immun.* **65**, 4682–4689.

Luft, B. J., Hafner, R., Korzun, A. H., Leport, C., Antoniskis, D., Bosler, E. M., Bourland, D. D., Uttamchandani, R., Fuhrer, J. & Jacobson, J. 1993 Toxoplasmic encephalitis in patients with the acquired immunodeficiency syndrome. *New Engl. J. Med.* **329**, 995–1000.

Manger, I., Hehl, A. B. & Boothroyd, J. C. 1998*a* The surface of *Toxoplasma* tachyzoites is dominated by a family of glycosylphosphatidylinositol-anchored antigens related to SAG1. *Infect. Immun.* **66**, 2237–2244.

Manger, I. D., Adrian, H., Parmley, S., Sibley, L. D., Marra, M., Hillier, L., Waterston, R. & Boothroyd, J. C. 1998*b* Expressed sequence tag analysis of the bradyzoite stage of *Toxoplasma gondii*: identification of developmentally regulated genes. *Infect. Immun.* **66**, 1632–1637.

McGuire, W., Hill, A. V., Allsopp, C. E., Greenwood, B. M. & Kwiatkowski, D. 1994 Variation in the TNF-α promoter region associated with susceptibility to cerebral malaria. *Nature* **371**, 508–510.

McLeod, R., Boyer, K., Roizen, N., Stein, L., Swisher, C., Holfels, E., Hopkins, J., Mack, D., Karrison, T. & Patel, D. 2000 The child with congenital toxoplasmosis. *Curr. Clin. Topics Infect. Dis.* **20**, 189–208.

Mordue, D. G., Monroy, F., La Regina, M., Dinarello, C. A. & Sibley, L. D. 2001 Acute toxoplasmosis leads to lethal overproduction of Th1 cytokines. *J. Immunol.* **167**, 4574–4584.

Odeh, M. 2001 The role of tumor necrosis factor-alpha in the pathogenesis of complicated falciparum malaria. *Cytokine* **14**, 11–18.

Pastinen, T., Raitio, M., Lindroos, K., Tainola, P., Peltonen, L. & Syvånen, A. C. 2000 A system for specific, high throughput genotyping by allele-specific primer extension on microarrays. *Genome Res.* **10**, 1031–1042.

Pfefferkorn, E. R., Pfefferkorn, L. C. & Colby, E. D. 1977 Development of gametes and oocysts in cats fed cysts derived from cloned trophozoites of *Toxoplasma gondii*. *J. Parasitol.* **63**, 158–159.

Rosenberger, C. M., Scott, M. G., Gold, M. R., Hancock, R. E. W. & Finlay, B. B. 2000 *Salmonella typhimurium* infection and lipopolysaccharide stimulation induce similar changes in macrophage gene expression. *J. Immunol.* **164**, 5894–5904.

Sabin, A. B. 1941 Toxoplasmic encephalitis in children. *J. Am. Med. Assoc.* **116**, 801–807.

Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P. O. & Davis, R. W. 1996 Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc. Natl Acad. Sci. USA* **93**, 10 614–10 619.

Shalon, D., Smith, S. J. & Brown, P. O. 1996 A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* **6**, 639–645.

Sibley, L. D. & Boothroyd, J. C. 1992 Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. *Nature* **359**, 82–85.

Sibley, L. D., LeBlanc, A. J., Pfefferkorn, E. R. & Boothroyd, J. C. 1992 Generation of a restriction fragment length polymorphism linkage map for *Toxoplasma gondii*. *Genetics* **132**, 1003–1015.

Trinchieri, G. 1998 Interleukin 12: a cytokine at the interface of inflammation and immunity. *Adv. Immunol.* **70**, 83–243.

Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. 1995 Serial analysis of gene expression. *Science* **270**, 484–487.

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E. & Spencer, J. 1998 Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science* **280**, 1077–1082.

Winzeler, E. A., Richards, D. R., Conway, A. R., Goldstein, A. L., Kalman, S., McCullough, M. J., McCusker, J. H., Stevens, D. A., Wodicka, L. & Lockart, D. J. 1998 Direct allelic variation scanning of the yeast genome. *Science* **281**, 1194–1197.