Polymorphic Secreted Kinases Are Key Virulence Factors in Toxoplasmosis
J. P. J. Saeij, et al.
Science 314, 1780 (2006);
DOI: 10.1126/science.1133690

The following resources related to this article are available online at www.sciencemag.org (this information is current as of September 19, 2007):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:
http://www.sciencemag.org/cgi/content/full/314/5806/1780

Supporting Online Material can be found at:
http://www.sciencemag.org/cgi/content/full/314/5806/1780/DC1

This article cites 23 articles, 13 of which can be accessed for free:
http://www.sciencemag.org/cgi/content/full/314/5806/1780#otherarticles

This article has been cited by 1 article(s) on the ISI Web of Science.

This article has been cited by 1 articles hosted by HighWire Press; see:
http://www.sciencemag.org/cgi/content/full/314/5806/1780#otherarticles

This article appears in the following subject collections:
Microbiology
http://www.sciencemag.org/cgi/collection/microbio

Information about obtaining reprints of this article or about obtaining permission to reproduce this article in whole or in part can be found at:
http://www.sciencemag.org/about/permissions.dtl
Polymorphic Secreted Kinases Are Key Virulence Factors in Toxoplasmosis


The majority of known Toxoplasma gondii isolates from Europe and North America belong to three clonal lines that differ dramatically in their virulence, depending on the host. To identify the responsible genes, we mapped virulence in F1 progeny derived from crosses between type II and type III strains, which we introduced into mice. Five virulence (VIR) loci were thus identified, and for two of these, genetic complementation showed that a predicted protein kinase (ROP18 and ROP16, respectively) is the key molecule. Both are hypervariable rhoptry proteins that are secreted into the host cell upon invasion. These results suggest that secreted kinases unique to the Apicomplexa are crucial in the host-pathogen interaction.

Toxoplasma gondii is an obligate intracellular parasite capable of infecting a wide variety of warm-blooded animals. Infections are widespread in humans and can lead to severe disease in utero or in individuals with a suppressed immune system. The majority of European and North American isolates belong to three distinct clonal lines, referred to as types I, II, and III (I, 2). Types I and III appear to be the result of just one or two matings between an ancestral type II strain and, respectively, one or other of a pair of closely related strains that are distinct from type II (3–6). The three major Toxoplasma lines differ in a number of phenotypes (7), the best described of which is virulence in mice: type I strains are the most virulent with a lethal dose (LD 100) of one cell; type II strains are the least virulent with very low virulence were identified by injection of 100 parasites (injection of 100,000 parasites); and type III strains are intermediate in virulence (8, 9).

33. Preliminary genomic and/or cDNA sequence data generated by The Institute for Genomic Research or the Sanger Center was accessed via http://ToxoDB.org. We are grateful for the contributions of J. Ajioka, J.-F. Dubremetz, S. Håkansson, E. Groisman, I. Paulsen, D. Roos, A. Schalfer, and J. Suetterlin. Supported by the NIH (AI36629, AI059176, AI44640, and Centers of Biomedical Research Excellence (COBRE), National Centers for Research Resources P20 RR-020185), the intramural program of National Center for Biotechnology Information at NIH, the Burroughs Welcome Fund, the Swedish Research Council, and the Swedish Foundation for Strategic Research. The GenBank accession number for the ROP18-III allele is AB159598.

Supporting Online Material
www.sciencemag.org/cgi/content/full/314/5806/176/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S3
References and Notes
9 August 2006; accepted 3 November 2006
10.1126/science.1133643

References and Notes
13. Preliminary genomic and/or cDNA sequence data generated by The Institute for Genomic Research or the Sanger Center was accessed via http://ToxoDB.org. We are grateful for the contributions of J. Ajioka, J.-F. Dubremetz, S. Håkansson, E. Groisman, I. Paulsen, D. Roos, A. Schalfer, and J. Suetterlin. Supported by the NIH (AI36629, AI059176, AI44640, and Centers of Biomedical Research Excellence (COBRE), National Centers for Research Resources P20 RR-020185), the intramural program of National Center for Biotechnology Information at NIH, the Burroughs Welcome Fund, the Swedish Research Council, and the Swedish Foundation for Strategic Research. The GenBank accession number for the ROP18-III allele is AB159598.

Supporting Online Material
www.sciencemag.org/cgi/content/full/314/5806/176/DC1
Materials and Methods
Figs. S1 to S6
Tables S1 to S3
References and Notes
9 August 2006; accepted 3 November 2006
10.1126/science.1133643

The parental type III clone was significantly less virulent than the parental type II clone, and among the recombinant progeny, a range of distinct phenotypes was observed (table S1). Injection of 100 parasites of some progeny (including strain S23 described above) resulted in uniform acute mortality in mice. There was also a large class of avirulent clones, unable to kill mice even when injected at the high dose (100,000) (table S1). The log-likelihood of association of virulence phenotypes with genetic markers was analyzed for three different phenotypes: (i) "high-dose survivability" (log 10 of the survival time after injection of 100,000 parasites); (ii) "avirulence," a binary trait defined as no mortality at any dose; and (iii) "low-dose survivability" (log 10 of the survival time after injection of 100 parasites). In the initial genome scan (15), two quantitative trait loci (QTLs) were identified: one that was highly significant at the left end of chromosome XII (VIR1) and one of borderline significance on chromosome X (VIR2) (Fig. 1A). A high log-arithm of odds (LOD) score in such analyses, as seen for the chromosome XII QTL, can obscure the contribution of other loci, especially when the number of progeny being analyzed is limited. When the variance associated with the QTL at the beginning of chromosome XII was fixed as a covariate, three additional QTLs emerged as significant (Fig. 1B): one each on chromosomes V/VIa (VIR3), VIIb (VIR4), and toward the right end of chromosome XII (VIR5). (See table S2 for a complete summary of all QTLs and their effects.)

Although the genomic regions spanned by the detected QTLs are large [–1 megabase (Mb)], we identified candidate genes on the basis of predicted coding function (16, 17), amount of polymorphism (5), and gene expression information (expressed sequence tags (EST) frequencies (18) and microarray data (this study)). Of par-
particular interest were polymorphic genes whose products were predicted to interact with the host (i.e., expected to be secreted into and/or beyond the parasitophorous vacuole) and with differences in expression between the type II and type III strains.

The VIR3 QTL spans a maximum of ~1.1 Mb within which there are 140 predicted proteins (16, 17). ROP18 stood out as the strongest candidate gene within this interval for a number of reasons. First, it has high homology to the ROP2 family of serine-threonine protein kinases in \textit{Toxoplasma}, which are released from specialized apical secretory organelles, called rhoptries, that are unique to the Apicomplexa. Second, ROP18 was previously identified in our laboratory as part of the rhoptry proteome [ROP2L2 of (19)] and has recently been confirmed to be a functional kinase (20) that ends up in the parasitophorous vacuole after host cell invasion (20, 21). Third, expression levels of ROP18 are very different between strains: Initial microarray results indicated that strains having the type III allele expressed significantly less mRNA than those with a type II allele, and quantitative polymerase chain reaction (QPCR) showed this difference to be on the order of 10,000-fold (Fig. 2A and fig. S3A). Analysis of the F1 progeny showed that this difference cosegregates with the VIR3 QTL (fig. S1; peak LOD score 4.5, *P* < 0.01). Sequencing the entire ROP18 gene (including promoter regions) and comparing type I, II, and III strains revealed that type III strains have a 2.1-kb sequence inserted 85 bp upstream of the ATG start codon, which is not present in type I or type II ROP18, or anywhere else in the only fully sequenced \textit{Toxoplasma} genome (from a type II strain) (16, 17). It seems likely that this insertion (relative to types I and II) in the 5'-untranslated region (UTR)—promoter of the type III ROP18 allele is involved in the major difference in expression of this locus in type III strains, although this has not been directly tested. Last, ROP18 is found in a genomic span that is generally dominated by type II–specific single-nucleotide polymorphisms (SNPs) on the basis of polymorphism maps derived from EST sequences (5, 8) and, as predicted by this, the type II ROP18 allele [(16) Gene model 20.m03896] has 11 (0.66%) SNPs in the coding region relative to genes I (GenBank accession number CA277113) and III (GenBank accession number EF092842) (Fig. 2B). Superimposed on this expected level of polymorphism, however, is a completely unexpected and extremely large number (85; 5.0%) of type III–specific SNPs (64 of which result in amino acid changes) (Fig. 2B).

To determine whether ROP18 was in fact the VIR3 QTL, we introduced a type II (ME49) allele of ROP18 into an avirulent type III background (CEP). This was done by cloning the type II gene (with 588 bp of endogenous promoter and the complete coding region) in frame with a C-terminal hemagglutinin (HA) tag. HA-specific antibody immunofluorescence on the resulting parasites revealed the presence of ROP18-HA in the rhoptries as expected (Fig. 2C), and also in vacuole-like structures within the host cell early after infection (Fig. 2D), similar to those previously observed for other rhoptry proteins (22). In BALB/c mice, these type III:ROP18 allele parasites were at least 4 logs more virulent than the wild type strain (Fig. 3). A second, genetically distinct type III:ROP18 allele clone (from a different transfection than the clone described in Fig. 3) was isolated that had similarly enhanced virulence in mice.

It is likely that the differences in the expression levels of ROP18 are enough to account for the large difference in virulence between type III:ROP18 allele and the wild-type line (CEP), and ROP18 levels were about eight times as high in the virulent type III:ROP18 allele as in a type II strain (Fig. S3). However, the large number of polymorphisms in the type III coding region could also play a role, and future experiments using strains with allelic replacement of coding and/or promoter sequences will distinguish between these possibilities. Consistent with the results presented here, a separate study examining progeny from a cross between types I and III to determine the genetic basis for the extreme virulence of type I strains, Taylor et al. (21) found that expression of the type I ROP18 allele in an avirulent type III strain also dramatically increases virulence in mice. These data confirm that ROP18 is indeed a virulence gene and is likely the genetic basis for the virulence QTL on chromosome VIIa that was identified in both studies (9, 21).

\textit{VIR4} falls within a 0.55-Mb interval on chromosome VIIIb. Within this candidate region, one locus, ROP16, immediately stood out because of its extreme variability (the type II allele has 39 nonsynonymous SNPs compared with types I and III in ~2100 bp of coding sequence (23)), its status as a rhoptry protein kinase (and thus likely injected during invasion), and our recent results demonstrating that this rhoptry protein kinase is injected into the host cell cytosol and is involved in the strain-specific differences in induction of interleukin 12 (IL-12) secretion by mouse macrophages (23) [IL-12 is well known to be key in \textit{Toxoplasma} pathogenesis (24, 25)]. To test whether ROP16 was indeed the VIR4 QTL, we introduced either a type III or a type I allele of ROP16 into a type II strain. Both alleles were pursued because our previous studies indicated that the type II allele is recessive (loss of function) to both the type I and type III alleles, and the type I and type III alleles have only three nonsynonymous SNPs.

\textbf{Fig. 1.} Genetic mapping of virulence phenotypes of F1 progeny from II × III crosses. BALB/c and CBA/J mice were infected with 100,000 or 10 tachyzoites from 40 different F1 recombinant progeny from II × III crosses, and mortality was recorded daily for 40 days. As there were no significant differences between the two different mouse strains, results were pooled. Three phenotypes are represented: (i) “high-dose survivability,” log$_{10}$ survival time (in days) after injection of 100,000 parasites (black line); (ii) “avirulence,” a binary trait defined as no mortality at any dose (red line); (iii) “low-dose survivability,” log$_{10}$ survival time (in days) after injection of 100 parasites (blue line). Plots indicate the log-likelihood association of phenotypes with markers aligned across the genome. Marker positions (in cM) are given by tick marks. Significance levels determined by 1000 permutations are indicated by horizontal lines [upper lines are significant; lower line is suggestive (*P* = 0.1)]. Because the significance levels for all three of the phenotypes differed by less than 0.1 LOD unit, only one significance line is drawn for all three. (A) Primary genome scan (see text). (B) Secondary genome scan after the effect of the major virulence peak on chromosome XII, evident in (A) and cosegregating with the SAG3 marker, is neutralized by making it a covariate.
between them (23). The engineered strains were injected into mice, and virulence was compared with that of parental type II. The results (Fig. 4) show a substantial decrease in virulence in the strains carrying either the type I or the type III transgene. This is consistent with our QTL analyses, which indicate that the type III allele is associated with lower virulence (table S1). Obtaining two genetically distinct clones (type II: ROP16I and type II: ROP16III) with identical virulence phenotypes argues against finding the insertion site of the DNA construct to be responsible for the drop in pathogenicity.

The results presented above strongly support the idea that polymorphic rhoptry kinases provide the genetic basis for two of the virulence QTLs mapped here. The identities of the genes contributing to the other three VIR loci are not yet known: None had a candidate gene as compelling as ROP16 and ROP18 (see supplementary results for a discussion of the other three loci). In no case do any of the five VIR loci appear to mediate virulence through an enhancement in growth, because no significant growth phenotype was seen for any of the F1 progeny strains reported here, with the notable exception of CL15 which, as reported previously, has a substantial growth retardation (3). Furthermore, simple growth rate is not likely the basis of the virulence differences, because the parental type III strain has a slight growth advantage compared with the ME49 strain despite its being less virulent (26). ROP18 has recently been shown to be an active kinase.

**Fig. 2.** Expression level, polymorphism analysis, and localization of ROP18. (A) Expression level of ROP18 in the parental lines (type II, ME49, and type III, CEP) and 18 F1 progeny. Data are displayed as fold-difference relative to the type III parent. The genotype for each of the F1 progeny at the ROP18 locus is indicated by the bar color (white or black). (B) Variation in the percentage of type I, II, and III SNP polymorphisms across chromosome VIIa. The entire chromosome was divided into 10-kb windows, and the number of SNPs of each type in each window [based on EST assemblies; see (5)] was divided by the number of sites with data from all three strains to compute a polymorphism percentage. The number of each polymorphism type in the ROP18 coding region is also shown. *Accession number AM075204 (28). **Downloaded from (16) (Gene model 20m.03896). (C) Immunofluorescence assay of type III: ROP18II showing HA-specific staining (green) of HA-tagged ROP18 colocalizing with the rhoptry marker ROP2/3/4 (red) in human foreskin fibroblasts 40 hours post inoculation. (D) Immunofluorescence assay of type III: ROP18II showing HA-specific staining of tagged ROP18 in human foreskin fibroblasts 2 hours post inoculation. Punctate staining indicates the location of ROP18-HA that has been secreted into the infected host cell (22). Color scheme same as in (C).

**Fig. 3.** Effect of expressing the type II allele for ROP18 in a type III strain background (CEP) on virulence in mice. Mice were infected intraperitoneally with either CEP.hxgprt complemented with HXGPRT alone (CEP*) or CEP parasites complemented with the type II allele for ROP18 along with 588 bp of upstream sequence (type III: ROP18II). Infections were verified in all the survivors based on the presence of Toxoplasma-specific antibodies, and only seropositive survivors are represented on the graph. Ten mice were used for all strain and dose combinations except for type III: ROP18II-100 (eight mice) and type III: ROP18II-10 (seven mice).

**Fig. 4.** Effect of expressing the type I and III strain allele of ROP16 in the type II Prugniaud strain (PRU.hxgprt) on virulence in mice. Mice were infected intraperitoneally with 5000 tachyzoites of type II (PRU.hxgprt) or an engineered version expressing an HA-tagged copy of the ROP16 allele from the type III CEP strain (type II: ROP16III) or the type I RH strain (type II: ROP16I). For PRU.hxgprt, 24 mice were used, 10 for type II: ROP16III, and 14 for type II: ROP16I. Infections were verified in all the survivors on the basis of the presence of Toxoplasma-specific antibodies. Results from three independent experiments were pooled.
The extreme sequence divergence observed for ROP16 and ROP18 could be a result of immune selection, but it is striking that these proteins are substantially more different among the three strains than the major surface antigens, SAG1 and SAG2 (3), which are highly immunogenic (at least in terms of antibodies) and thus are likely to represent an extreme for targets of selective pressure. It could be that ROP16 and ROP18 are subject to immune pressure by T cell responses rather than by antibodies (the major T cell antigens of Toxoplasma are not known, but the presence of ROP16 and ROP18 within the infected host cell would make them readily available for antigen presentation). For ROP18, pairwise comparisons of nonsynonymous to synonymous SNPs among the three alleles suggests that selection is operating on type I and type II ROP18. All 28 SNPs between them result in amino acid changes; the ratio of nonsynonymous nucleotide substitutions to synonymous nucleotide substitutions (Ka/Ks) is 4.6 (27) (fig. S2B). It is less clear how much selection has operated on the type III allele, where 20 of the 85 SNPs specific to this strain are synonymous, which gives Ka/Ks values closer to 1.0 [a ratio expected under a neutral selection model (fig. S2B); Ka/Ks = 1.7 and 1.3 versus type I and type II alleles, respectively]. It is possible that the very low expression of ROP18 in type III strains makes type III ROP18 a “null” allele that can more readily accumulate random mutation, although the level of variation in this allele is so atypical for this strain and genomic region (Fig. 2B) (5) that both positive selection and neutral drift may have operated on this allele over time. It will be key to sequence the ROP18 gene in other T. gondii isolates to determine the range of selective pressure at this locus in the population. Finally, instead of immune pressure, another possibility is that at least some of the sequence divergence of these genes could be a result of optimization for interaction with specific proteins (e.g., their substrates) in one or more hosts that are central to their transmission. It is difficult to predict what those hosts are or were but Mus spp. has probably not played that role for all three strain types, given the very different allele-specific interactions described here. Rather, other species of rodents or birds that are central to the parasite’s transmission, perhaps in distinct parts of the world, could be the true, evolutionarily relevant hosts in which selection for different versions of these key genes occurred.

Nitrogen Fixation at 92°C by a Hydrothermal Vent Archaeon

Mausumi P. Mehta* and John A. Baross

A methanogenic archaeon isolated from deep-sea hydrothermal vent fluid was found to reduce N₂ to NH₃ at up to 92°C, which is 28°C higher than the current upper temperature limit of biological nitrogen fixation. The 16S ribosomal RNA gene of the hyperthermophilic nitrogen fixer, designated F5406-22, was 99% similar to that of non–nitrogen-fixing Methanocaldococcus jannaschii DSM 2661. At its optimal growth temperature of 90°C, F5406-22 incorporated ^15N, and expressed nifH messenger RNA. This increase in the temperature limit of nitrogen fixation could reveal a broader range of conditions for life in the subseafloor biosphere and other nitrogen-limited ecosystems than previously estimated.

Hydrothermal fluids venting from undetected deep-sea mid-ocean ridges are low in nitrate and ammonia (7), indicating that the microbial community inhabiting the subseafloor (2) may be limited by fixed nitrogen. Dissolved N₂ is abundant in hydrothermal vent fluids (3), and biological nitrogen fixation has been proposed to explain the depleted ^15N/^14N ratios of low–trophic level animals living around hydrothermal vents (4). The nitrogenase enzyme complex, encoded by the nifHDK genes, catalyzes nitrogen fixation, i.e., the reduction of N₂ to NH₃. Although nifH genes have been detected in hydrothermal vent fluid (5), no microorganism isolated from deep-sea vents has, to our knowledge, been reported to fix nitrogen. A wide

References and Notes
15. Materials and methods are available as supporting material on Science Online.
18. L. Li et al., Genome Res. 13, 443 (2003).
29. This work was supported by grants to J.C.B. from the NIH (AI22423, AI03203, and AI01014) and the Ellison Medical Foundation (a Senior Scholar Award); to L.D.S. from the NIH (AI36629, AI059176); to J.W.A. from the U. K. Biotechnology and Biological Sciences Research Council and the Wellcome Trust; to J.P.J.S from the California Universitywide AIDS Research Program (F04-ST-216); to S.C. from the California Universitywide AIDS Research Program (FT-207-ST); and to J.P.B. from the NIH (F32AI63063). We thank J. D. Dunn for construction of the pGRA-HA-HPT vector; A. Fouts for help with QPCR; K. W. Bowman for his many helpful suggestions on using the Rqti package; P. Bradley and J.-F. Dubremetz for exchange of unpublished data; E. Pfefferkorn for performing the original crosses as part of a collaboration on drug resistance; the U. K. Medical Research Council MRC HGPMP for printing the Toxoplasma microarrays; and M. White, J. Wootton, and J.F. Dubremetz for helpful comments on the manuscript. Preliminary genomic and/or cDNA sequence data were accessed from ToxoDB.org and/or www.tigr.org/tdb/t bons/tdb1.go. Genomic data were provided by The Institute for Genomic Research (supported by the NIH grant AI05909) and by the Sanger Center (Wellcome Trust). EST sequences were generated by Washington University (NIH grant 1R01AI045806-01A1). The GenBank accession number for the ROP18 sequence from the type III strain CTG is EF092842.
30. Supplementary Online Material
www.sciencemag.org/cgi/content/full/314/5806/1780/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Tables S1 and S2
References
9 August 2006; accepted 3 November 2006
10.1126/science.1133690

Supporting Online Material
www.sciencemag.org/cgi/content/full/314/5806/1780/DC1
Materials and Methods
SOM Text
Figs. S1 to S3
Tables S1 and S2
References
9 August 2006; accepted 3 November 2006
10.1126/science.1133690

www.sciencemag.org  SCIENCE  VOL 314  15 DECEMBER 2006  1783
Supporting Online Material for

Polymorphic Secreted Kinases Are Key Virulence Factors in Toxoplasmosis
J. W. Ajioka, J. C. Boothroyd*

*To whom correspondence should be addressed. E-mail: john.boothroyd@stanford.edu

DOI: 10.1126/science.1133690

This PDF file includes

Materials and Methods
Figs. S1 to S6
Tables S1 to S3
References and Notes
Materials and Methods

Genetic crosses
Recombinant F1 progeny derived from three different crosses (named S, CL and c96) between the ME49 strain (Type II) and the CEP strain (type III) were used in the analysis of virulence. The genotypes of these progeny and the exact details of the crosses have been described previously (S1) and can also be accessed at the Toxoplasma Genome Map web-page (http://www.toxomap.wustl.edu/).

Parasite maintenance
Strains were maintained in vitro by serial passage on monolayers of human foreskin fibroblasts (HFFs) at 37°C in the presence of 5% CO2 as previously described (S2). HFFs were grown in Dulbecco modified Eagle medium (GIBCO BRL) supplemented with 10% NuSerum (Collaborative Biomedical Products), 2 mM glutamine, 50 µg/ml penicillin, 50 µg/ml streptomycin, and 20 µg/ml gentamicin.

Determination of virulence in mice
Female BALB/c or CBA/Caj mice that were 7- to 11-weeks-old (Jackson Laboratories, Bar Harbor, Maine) were used in all experiments. For intraperitoneal (i.p.) infection, tachyzoites were grown in vitro and extracted from host cells by passage through a 27-gauge needle, washed two times in phosphate-buffered saline (PBS) and quantified with a hemocytometer. Parasites were diluted in PBS, and mice were inoculated intraperitoneally with tachyzoites of each strain (in 200 µl) by using a 27-gauge needle. The viability of parasites in each inoculum was checked by in vitro plating of a sample of the inoculum and counting plaques 5-7 days later. For the F1 recombinant progeny from the S and CL crosses, two BALB/c mice were injected with 100,000 tachyzoites. For the 21 c96 cross F1 recombinant progeny, BALB/c and CBA/Caj mice were injected with 100,000 (one mouse of each strain) or 100 tachyzoites (two mice of each strain). Mortality was recorded daily for 40 days after infection. Blood samples were collected from surviving mice and tested for the presence of antibodies against T. gondii by immunofluorescence. For the F1 recombinant progeny S1T, S2T and S28, another three CBA/Caj mice were injected with 100 parasites and another two with 100,000 parasites. For F1 recombinant progeny S29, S30, c96-C12, c96-H6 and STH10, another three CBA/Caj mice were injected with 100 parasites. The institutional APLAC committee approved all protocols.

Genome-wide scans
The linkage between virulence, based on mortality/survival time in mice, and parasite genotype was performed using the statistical package 'R' (www.cran.org) with the library ‘R/qtl’ (S3), which uses computational approaches described previously (S4). Because T. gondii is haploid, only two alleles exist for each locus in a given cross, and therefore the computational approach is equivalent to an analysis of recombinant inbred lines. The method of marker regression was used because departure from Mendelian segregation was detected for a number of markers partly due to the fact that the majority of the F1 progeny were isolated by drug selection (S5). Significant thresholds for all genome scans were calculated by performing 1,000 permutations (S6).
**Toxoplasma microarrays**

HFF monolayers were infected with tachyzoites (MOI=10) from a type II strain (ME49), a type III strain (CEP) and 18 II X III F1 progeny (S and CL clones; \(S1\)) and total RNA was harvested 24 h post-infection using the Trizol reagent. Parasite microarrays were custom printed cDNA arrays from an RH tachyzoite cDNA library (S7): approximately 12,000 cDNA clones were PCR amplified, resuspended in 3X SSC, 1.5M betaine and printed onto GAPS II slides (Corning, Netherlands), using a Lucidea Pro array printer (GE Healthcare Bio-Sciences, Amersham, UK) with a 48-pin tool. The slides were baked for 18 hours at 80°C and stored dessicated at room temperature. For 7,488 of the cDNAs, end-sequence data was obtained previously and is available in Genbank (S7). Labeling and hybridization were carried out as described previously (S8). For each array spot, the log2 ratios of the normalized data from the 18 F1 progeny were tested for significant association with each genetic marker using R/qtl (S3). Genome-wide significance (P<0.05) was assessed using 1000 permutations of the genotype data (S3).

**Complementation**

For ROP18, primers were constructed to amplify from 588 bp upstream of the ATG start codon to the last amino acid from the type II strain, and contained restriction sites for ligation into a modified version of pMINI-HXGPRT (S9) that allows for insertion of a gene of interest in frame with an HA tag and subsequent selection for the presence of HXGPRT in transfectants. Constructs were used to engineer the type III parental line CEP\(\Delta\)hxgprt to express an HA-tagged version of ROP18 from the type II strain using standard culture methods (S2). For virulence studies with ROP18, a clone that inserted only the HXGPRT minigene from the vector (and not the HA-tagged ROP18) was used as a control. The type I and type III constructs for ROP16 were constructed in a similar fashion, except that 2000 bp of upstream sequence was used (as well as the relevant coding region) from the two respective donor strains and the resulting constructs were transfected into the type II strain PRU\(\Delta\)hxgprt.

**Real-time quantitative PCR**

RNA was harvested from RH (type I), ME49 (type II), CEP (type III) and Type III-ROP18II-infected HFF cell monolayers, converted to cDNA using oligo-dT, and diluted 1:10 for real-time quantitative PCR (qPCR) analysis. ROP18 primers were chosen to target regions free from polymorphisms in all 3 strains, and AMA1 (Genbank Acc. No. AF010264) was used as a control since both microarray (Figure S3) and western blot (S10) analyses show that AMA1 is similarly expressed in type II and type III strains. qPCR was performed on an iCycler® system (BioRad) using SYBR® green fluorescence to quantify PCR product accumulation. Cycle threshold (Ct) values were determined with the same threshold settings for each primer set. One-way analysis of variance was performed on the \(\Delta\)Ct values (Ct value for AMA1 minus Ct value for ROP18, which is a numerical representation of the differences in ROP18 expression) from each strain, followed by Dunnett's multiple comparison post-test to compare the ROP18 expression level for all strains to ME49. For display purposes, estimated fold-differences in ROP18 transcript level between each strain and ME49 were calculated using the \(2^{(\Delta\Delta\text{Ct})}\) method (S11).
SOM Text: Supplementary Results

QTL’s on chromosome XII (VIR1 and VIR5) and chromosome X (VIR2)

The precise genes responsible for the VIR1, VIR2 and VIR5 QTLs have yet to be identified. Candidate genes do exist, however, and these will form the foundation for future studies aimed at identifying the entire complement of significant virulence genes in the II x III genetic cross. The QTL located at the left end of chromosome XII (VIR1) spans a maximum of ~0.98 Mb (marker AK37 to M144; (S1)), encompassing 154 predicted proteins, and three candidate genes were identified for this locus using the criteria listed above: these are a surface antigen (SAG3), a secreted protein kinase (ROP5), and a hypothetical protein 145.m00583 (Table S2). SAG3 has previously been shown to be crucial for infection: when it was knocked out in a type I strain there was a decrease in virulence (S12) although there is no evidence that polymorphisms between surface antigens can impact disease outcome. The predicted kinase encoded by ROP5 is secreted from the parasite’s apical rhoptries and into the parasitophorous vacuole where it could interface with the cytosol of the infected host cell (S13). Given that VIR3 and VIR4 encode ROP18 and ROP16, respectively, it is notable that another rhoptry kinase is among the list of candidate genes for VIR1. VIR2 falls within a ~1.2 Mb interval, and contains 139 predicted genes. Two candidate genes were identified for this locus (gene models 42.m03493, 42.m03409), both of which are predicted to have signal peptides and at least one transmembrane domain, but for which there are no other similarities to other known proteins or domains. Finally, the genomic segment spanned by the VIR5 QTL is the largest of all identified loci, spanning 5.7 MB and containing 719 predicted genes. Most of this QTL is found in a region where there are very few differences between type II and type III strains (S14). It should also be noted that this QTL is linked to resistance to adenosine arabinoside (AraA) which is due to a loss of function mutation in the adenosine kinase (AK) gene (S15). It has been reported that a lack of functional AK can result in fitness defects (S16). Since the type III parent carries the AraA marker and this drug was used to select 23 of the 41 recombinant progeny, it is possible that AK is the gene responsible for VIR5. No other compelling candidate genes are apparent in this region of chromosome XII.
**Figure S1:**

(A and B) Comparison of LOD scores resulting from genome-wide scans of virulence phenotypes 1 and 2 (black and red lines, respectively) and ROP18 expression (orange line). Data for virulence phenotypes are exactly as presented in Figure 1 (A and B). Only the significance level (P<0.01) for ROP18 expression is shown and was determined using 1000 permutations of the genotype data. Significance levels for virulence phenotypes can be found in Figure 1A. (A) Primary genome scan. (B) Secondary scan, fixing marker SAG3 on chromosome XII as a covariate (for virulence data only). For details, see Materials and Methods.
Fig. S2. (A and B) Alignment of the predicted coding region for ROP18 in type I (RH), II (ME49) and type III (CEP) strains. Type I sequence (Genbank Acc. No. CAJ27113), type II sequence (www.toxodb.org; Gene Model 20.m03896) and type III sequence (Genbank Acc. No. EF092842) were aligned using ClustalX. The predicted signal peptide, protein kinase ATP-binding domain and serine/threonine protein kinase active site signature sequences are highlighted in red, blue and green, respectively. Note: It is not known if the first (shown) or second methionine (context MGLATLL...) is the true start codon for ROP18. (B) Pairwise Ks and Ka values and their ratios determined with DIVERGE software (Genetics Computer Group, Accelerys) implementing the method in (S17). *Ka: number of non-synonymous mutations per 100 non-synonymous sites. **Ks: number of synonymous mutations per 100 synonymous sites.
Figure S3:

**Fig. S3.** (A and B) Microarray (A) and real-time quantitative PCR (B) analyses of transcript levels of apical membrane antigen-1 (*AMA1*) and *ROP18* in canonical strain types and the complemented strains used in this study. Error bars represent standard error of the mean. (A) Transcript levels of *AMA1* and *ROP18* in type II (ME49) and type III (CEP) strains of *Toxoplasma* as determined by microarray analysis. Data are represented as the log$_2$ ratio of the experimental sample (strain) versus T3/T7 reference RNA (see Methods). *AMA1* data are the average of 3 spots on the microarray, while those for *ROP18* are derived from a single spot. N=2 for ME49 and 3 for CEP. (B) Real-time quantitative PCR data for *ROP18* transcript levels in 3 canonical *T. gondii* strains (RH, ME49 and CEP) as well as type III:*ROP18*$_{II}$. Data are represented as estimated fold difference in transcript level versus ME49, and asterisks indicate where transcript level was significantly different ($P<0.05$) from ME49 as determined by one-way ANOVA and Dunnett's multiple comparison post-test on the ΔCt values (Ct value for *AMA1* minus that for *ROP18*). N=4 for each strain. ***: $P<0.001$. 
Supplementary Table Legends

**Table S1.** Genotype and phenotype information for all recombinant F1 progeny derived from three genetic crosses (S, CL and c96) between type II and type III strains. For each F1 recombinant (column one), three phenotypes are represented: (1) “high-dose survivability”: survival time (in days) after injection of 100,000 parasites; (2) “avirulence”: a binary trait defined as no mortalities (“0” vs. any mortalities is “1”) at any dose; (3) “low-dose survivability”: survival time (in days) after injection of 100 parasites. For each F1 progeny, their genotype is shown at the genetic markers (“3” = type III allele, “2” = type II allele) on chromosomes VIIa, VIIb, X, and XII that are most tightly linked to the five virulence QTLs (i.e., at the peak of each LOD score) as well as the most proximal marker. Their position (in cM) on each respective chromosome is also given. The allele positively contributing to heightened virulence is shaded. F1 recombinant progeny with similar genotypes for the genetic markers most tightly linked to the five virulence QTLs are grouped together.

**Table S2.** *P*-values, percent variance estimates, candidate genes and physical sizes of the five virulence QTL’s. A model using five QTLs was used to fit the data for all 3 phenotypes. The LOD score and percentage of variance explained by each QTL was estimated by dropping one QTL at a time from the model and comparing the full model to the sub-model with the QTL dropped. *P*-values were estimated from the LOD scores based on 1000 permutations. Results shaded in grey indicate locus significance (genome-wide *P*<0.05).
Supplementary Table 1: Survival data of F1 progeny and corresponding genotypes at 5 identified virulence loci

| F1 progeny | Phenotype | M95 | CS4 | CS2 | AK43 | Tgubs1 | AK103 | L339 | AK105 | cASRay | GRA2 | AK65 | SRS4 | AK37 | AK145 | SAG3 | M144 | AK185 | AK | M183 |
|------------|-----------|-----|-----|-----|------|--------|-------|------|-------|--------|------|------|------|------|-------|-----|------|-------|-----|------|-----|------|
| S23        | 7.0       | 0   | 7.8 |     |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 E7     | 8.0       | 0   | 10.0|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STC8   | 7.0       | 0   | 10.5|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STG2   | 7.0       | 0   | 11.3|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STH1   | 7.0       | 0   | 12.0|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL16       | 7.5       | 0   | 40.0|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL19       | 7.5       | 0   | 9.8 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 A5     | 13.0      | 0   | 11.8|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STF3   | 7.0       | 0   | 25.8|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL11       | 7.5       | 0   | 28.8|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S29        | 11.0      | 0   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL12       | 7.5       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL17       | 10.5      | 0   | 21.5|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STG4   | 8.5       | 0   | 27.0|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STG10  | 13.5      | 0   | 28.6|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STD10  | 20.5      | 0   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STH10  | 9.0       | 0   | 11.0|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S30        | 12.0      | 0   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 H6     | 17.5      | 0   | 22.0|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 C12    | 21.5      | 0   | 33.3|      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STC7   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STE7   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STE10  | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 B4     | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STE1   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STG2   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STD2   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STH11  | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S22        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S26        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL15       | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STH5   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S27        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL18       | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S28        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| c96 STD3   | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CL13       | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S25        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S27        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S21        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| S17        | >40       | 1   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| ME49*      | 8.5       | 0   | 10  |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |
| CEP*       | 38.5      | 0   | >40 |      |      |        |       |      |       |        |      |      |      |      |       |     |      |       |     |      |     |      |

*Parental lines

- Genotype that positively contributes to virulence
### Supplementary Table 2: P-values, percent variance estimates and candidate genes in the 5 virulence QTLs

<table>
<thead>
<tr>
<th>QTL</th>
<th>Chr, position (cM)</th>
<th>P-value</th>
<th>Variance explained (%)</th>
<th>Size (MB)</th>
<th>Genes</th>
<th>Candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Phenol 1 Phenol 2 Phenol 3</td>
<td>Phenol 1 Phenol 2 Phenol 3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIR1</td>
<td>XII, 10</td>
<td>1.6E-14 5.9E-12 3.4E-06</td>
<td>51.6 42.8 26.9</td>
<td>0.98</td>
<td>154</td>
<td>ROP5, SAG3</td>
</tr>
<tr>
<td>VIR2</td>
<td>X, 13.7</td>
<td>3.9E-03 1.6E-03 0.33</td>
<td>3.5 5.3 0.9</td>
<td>1.2</td>
<td>139</td>
<td>42.m03493, 42.m03409</td>
</tr>
<tr>
<td>VIR3</td>
<td>VIIa, 5.8</td>
<td>1.2E-04 1.9E-06 0.02</td>
<td>6.9 14.4 5.4</td>
<td>1.1</td>
<td>140</td>
<td>ROP18*</td>
</tr>
<tr>
<td>VIR4</td>
<td>VIIb, 5</td>
<td>0.02 0.36 4.5E-03</td>
<td>2.1 0.4 8.4</td>
<td>0.55</td>
<td>78</td>
<td>ROP16*</td>
</tr>
<tr>
<td>VIR5</td>
<td>XII, 60</td>
<td>0.03 0.46 3.1E-03</td>
<td>1.8 0.3 9.1</td>
<td>5.7</td>
<td>719</td>
<td>Adenosine kinase</td>
</tr>
</tbody>
</table>

| Variance explained (full model; %) | 84.6 | 81.2 | 62.4 |

*Evaluated in the present study*
**Supplementary References**


