Cell Host & Microbe
Resource

Integrative Genomic Approaches Highlight a Family of Parasite-Specific Kinases that Regulate Host Responses

Lucia Peixoto,[1] Feng Chen,[1] Omar S. Harb,[1] Paul H. Davis,[2] Daniel P. Beiting,[1] Catie Small Brownback,[1] Dinkorma Ouloguem,[1] and David S. Roos[1,*]
[1]Department of Biology and Penn Genome Frontiers Institute, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Biology, University of Nebraska at Omaha, Omaha, NE 68182, USA
*Correspondence: droos@sas.upenn.edu
DOI 10.1016/j.chom.2010.07.004

SUMMARY

Apicomplexan parasites release factors via specialized secretory organelles (rhoptries, micronemes) that are thought to control host cell responses. In order to explore parasite-mediated modulation of host cell signaling pathways, we exploited a phylogenomic approach to characterize the *Toxoplasma gondii* kinome, defining a 44 member family of coccidian-specific secreted kinases, some of which have been previously implicated in virulence. Comparative genomic analysis suggests that "ROPK" genes are under positive selection, and expression profiling demonstrates that most are differentially expressed between strains and/or during differentiation. Integrating diverse genomic-scale analyses points to ROP38 as likely to be particularly important in parasite biology. Upregulating expression of this previously uncharacterized gene in transgenic parasites dramatically suppresses transcriptional responses in the infected cell. Specifically, parasite ROP38 downregulates host genes associated with MAPK signaling and the control of apoptosis and proliferation. These results highlight the value of integrative genomic approaches in prioritizing candidates for functional validation.

INTRODUCTION

The phylum Apicomplexa includes thousands of obligate intracellular parasites, many of which are important sources of morbidity and mortality in humans and animals. *Plasmodium* parasites are responsible for malaria (World Health Organization, 2009), while *Toxoplasma* is a leading source of congenital neurological birth defects and a prominent opportunistic infection in AIDS (Hill and Dubey, 2002); *T. gondii* has also emerged as an experimentally tractable model system (Roos et al., 1994; Roos, 2005). These parasites have evolved novel mechanisms for invasion and intracellular survival, including an apical complex of specialized secretory organelles: "micronemes" are associated with host cell attachment, while secretion from "rhoptries" is associated with establishment of an intracellular "parasitophorous vacuole" (Carruthers and Sibley, 1997; Bradley and Sibley, 2007). Several rhoptry (ROP) proteins contain kinase-like domains, although many lack an obvious catalytic triad (El Hajj et al., 2006). Recent work on the active rhoptry kinases ROP16 and ROP18 (El Hajj et al., 2007a) shows that the former is secreted into the infected cell and alters STAT 3/6 phosphorylation (Saeij et al., 2007), while the latter is an important virulence determinant (Saeij et al., 2006; Taylor et al., 2006).

Eukaryote protein kinases (ePKs) are phylogenetically related (Hanks and Hunter, 1995) and typically reside in the cytoplasm, where they play key roles in signal transduction (Manning et al., 2002a). Genome sequencing has defined the complete kinome for various species (Hunter and Plowman, 1997; Plowman et al., 1999; Manning et al., 2002b), including that of *Plasmodium falciparum* (Ward et al., 2004), helping to elucidate molecular players that may be involved in signaling. We have exploited the *T. gondii* genome (Gajria et al., 2008; http://toxodb.org) to define this parasite's kinome, including 159 putative ePKs, of which 108 are predicted to be active. The largest family of *T. gondii* kinases (ROPK) contains 44 members, including ROP16 and ROP18; orthologs of most ROPK proteins are also recognizable in the *Neospora caninum* genome. ROPK proteins have not been identified in *Plasmodium*, although these parasites possess another group of secreted kinases, the FIKK kinases (Ward et al., 2004; Anamika et al., 2005; Nunes et al., 2007). This report, in conjunction with previous studies, indicates that ROPK proteins are secreted into the parasitophorous vacuole, trafficking to the intravacuolar membranous network, the vacuolar surface membrane, and/or the host cell.

Using comparative genomic approaches, we show that the ROPK family has been under positive selection since the divergence of *Neospora* and *Toxoplasma*. ROPK genes also exhibit an unusual degree of differential expression between strains and/or during differentiation. Integrating these genomic scale data sets highlights the previously uncharacterized kinase ROP38 as likely to be functionally important. Virulent RH strain *T. gondii* normally expresses virtually no ROP38, but this transcript is abundant in the relatively avirulent VEG strain parasites. Infection of mammalian cells with RH transgenics engineered to express VEG levels of ROP38 significantly alters the expression of ~1200 host genes (383 by > 2-fold), usually manifested as a suppression of host genes induced by RH infection. Functional clustering shows that parasite expression of ROP38

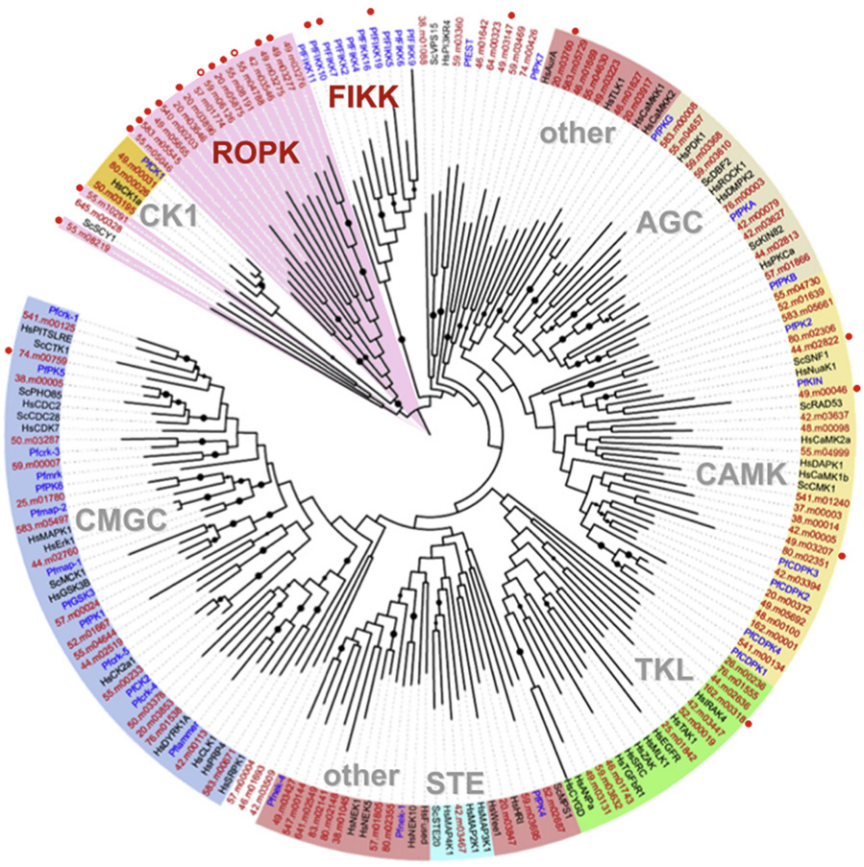

Figure 1. The *T. gondii* Kinome
Classification of 108 active kinases predicted from the *T. gondii* genome. Black, human and yeast; blue, *P. falciparum*; red, *T. gondii*. Colored arcs highlight major kinase groups: AGC, CMGC, CAMK, TKL, CK1, and STE (Hanks and Hunter, 1995; Manning et al., 2002b). Red lettering, apicomplexan-specific groups ROPK (pink) and FIKK. Red circles, kinases with predicted secretory signal sequence or signal anchor (open, newly recognized); black dots, bootstrap support > 50%.

exerts a potent effect on the expression of host transcription factors, signaling pathways, and the regulation of cell proliferation and apoptosis. Genes downregulated > 4-fold by ROP38 include c-*fos*, EGR2, and other early response genes such as CXCL1 and NAMPT, consistent with regulation of host-cell MAPK cascades (particularly ERK signaling).

RESULTS

The *T. gondii* Kinome Contains 108 Putative Kinases and 51 Pseudokinases

Analysis of the *T. gondii* genome (Experimental Procedures) identifies a total of 159 ePKs, including 108 predicted to be active based on the presence of 12 complete kinase subdomains, Pfam domain PF0069, and three conserved amino acids constituting the catalytic triad (Lys30, Asp125, Asp143; Manning et al., 2002b). Representatives of previously defined human (Manning et al., 2002b), yeast (Hunter and Plowman, 1997), and *P. falciparum* (Ward et al., 2004) ePK subfamilies were used as seeds for phylogenetic classification of all *T. gondii* kinases predicted to be active, most of which could readily be assigned to established ePK groups (Figure 1, Figure S1, Tables S1 and S2).

The active *T. gondii* kinome includes 10 cyclic nucleotide regulated kinases (AGC), 20 cyclin-dependent kinases and close relatives (CMGC, including CDK, MAPK, GSK), 20 calcium/calmodulin regulated kinases (CAMK), three casein kinase-like proteins (CK1), sevevn tyrosine kinase-like proteins (TKL), and one MAP kinase kinase (MAPKK, STE). Additional kinases ("Other" in Table S1) include nine Nima/NEK, four ULK, one Aurora, two Wee, and three PIK3R4 (two display architecture distinct from their animal/fungal homologs). More than half of the "Other" kinases identified in *T. gondii* are restricted to the phylum apicomplexa. The *T. gondii* genome is also predicted to include 51 pseudokinases, defined as inactive based on the absence of a complete catalytic triad and/or extremely low HMM scores. Only four of these could be classified into major kinase groups (two CMGCs, one CAMK, one AGC), and most are specific to the apicomplexa. To facilitate cross-species comparison for functional analysis, a combined kinome for *P. falciparum* and *T. gondii* kinome (including representative human and yeast orthologs) is presented as Figure S1 (see also Table S2).

Apicomplexans Have Evolved a Unique Repertoire of Secreted Kinases

In order to understand how kinases have evolved in the protozoa, the *T. gondii* and *P. falciparum* kinomes were compared with previously published analyses of the amoeba *Dictyostelium discoideum* (Goldberg et al., 2006) and the kinetoplastid parasites *Leishmania major*, *Trypanosoma brucei*, and *T. cruzi* (Parsons et al., 2005), in addition to plants (*Arabidopsis thaliana*, *Oryza sativa*; Dardick et al., 2007), fungi (*Saccharomyces cerevisiae;* Plowman et al., 1999), and animals (*Caenorhabditis elegans*, *Drosophila melanogaster*, *Homo sapiens;* Plowman et al., 1999; Manning et al., 2002b), as shown in Table S1. While the complete proteomes of *Toxoplasma* and *Plasmodium* are estimated to differ by < 25% in size (http://eupathdb.org/), the kinome of *T. gondii* is almost double the size of the *P. falciparum* kinome. Parasite kinases exhibit similar distribution among the major groups, with many CMGC and CAM kinases and no tyrosine kinases (TK) or receptor guanylate cyclases (RGC) (Table S1). Only a single STE kinase (MAPKK) was identified, in *T. gondii*. Consistent with the secondary endosymbiotic history of the apicoplast (Foth and McFadden, 2003), three *T. gondii* kinases exhibit probable plastid origin: two CAMKs and 1 AGC (Table S2).

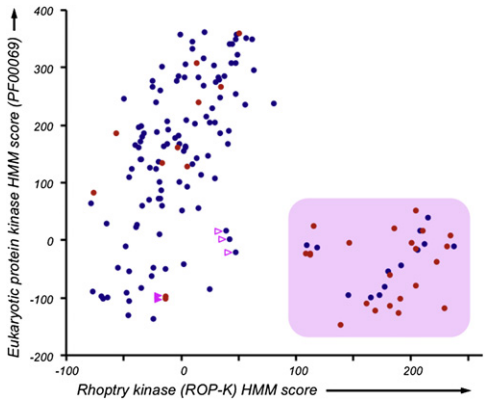

Figure 2. The *T. gondii* ROPK Family

A group-specific HMM based on the active ROPK genes (Figure 1) was applied to the entire *T. gondii* genome (see the Experimental Procedures). Pink indicates 35 genes clearly distinguishable from kinases defined by Pfam 00069, including tandem duplications ROP2A/2B and ROP4/7 (misannotated as a single gene in http://toxodb.org/). Ten additional genes were also identified as probable ROPKs: two with a truncated kinase domain (filled arrows), three with internal insertions (open arrows), and five not scored by Pfam 00069 (see Table S2). Red indicates predicted signal sequence.

The most striking feature of the apicomplexan kinome is the large fraction of kinases that do not fall within traditional groups: "Other" represents ~55% of the apicomplexan kinome, versus 30%–37% for other unicellular species and ~20% for metazoa. Most are parasite specific, usually at the species level: 11 are shared between apicomplexa (including the FIKK family), 24 are unique to *P. falciparum* (Ward et al., 2004; Nunes et al., 2007), and 51 are unique to *T. gondii,* including the virulence factor ROP18 (Saeij et al., 2006; Taylor et al., 2006; El Hajj et al., 2007a). Interestingly, while ePKs are typically cytosolic, an unusual number of apicomplexan kinases are predicted to contain secretory signal sequences (red dots in Figure 1), including the FIKK family and several *T. gondii* kinases from the ROP18 clade.

The ROPK Family: A Coccidian-Specific Family of Secreted Kinase-Related Proteins

In order to further define the *T. gondii* rhoptry kinase (ROPK) family, the monophyletic group of active kinases highlighted in Figure 1, plus ROP16 (but excluding 49.m05665 and 20.m03646, which harbor large insertions interrupting the kinase domain) was used to construct a family-specific profile HMM (see the Experimental Procedures). Applying this ROPK HMM to the entire *T. gondii* genome identifies 34 unique genes (Figure 2); application to the *T. gondii* kinome alone yielded identical results. ROPK genes were also identified in *Neospora caninum*, and assigned as orthologs based on synteny. Degenerate ROPK genes (pseudogenes and inaccurate gene models) were detected based on sequence similarity (see the Experimental Procedures), identifying ten additional family members: three with insertions > 3 kb in the kinase domain (ROP33, ROP34, ROP46), and seven with truncated kinase domains (five of which are not recognized by PF00069; Table S3). In aggregate, the *T. gondii* ROPK family contains at least 44 genes, including all previously reported kinase-like rhoptry proteins and more than doubling the number of previously described family members (Boothroyd and Dubremetz, 2008).

Twenty-four of the 33 *T. gondii* kinases and pseudokinases predicted to be secreted outside of the parasite are members of the ROPK family. Among the 34 nondegenerate ROPK family members (Figure 2), 22 contain an obvious N-terminal signal peptide (red), two (ROP26 and ROP28) are predicted to contain a signal anchor (Table S3), and experimental reanalysis of two (ROP4 and ROP7) identifies signal sequences previously missed (Carey et al., 2004). Signal sequence identification depends on accurate 5′ end prediction, which is notoriously difficult in large eukaryotic genomes (Liu et al., 2008). Four more ROPKs (ROP45, ROP29, ROP30, and ROP41) were found to contain signal peptides based on 5′ rapid amplification of cDNA ends (RACE) and/or comparison with syntenic orthologs in *N. caninum* genome. It appears likely that all ROPK family members encode a secretory signal.

In order to evaluate the accuracy of these predictions, seven novel ROPKs (ROP19, ROP20, ROP21, ROP22, ROP23, ROP25, and ROP38) and four proteins found in the rhoptries by proteomic analysis (ROP17, ROP24, ROP39, and ROP40; Bradley et al., 2005) were engineered as recombinant HA-tagged fusion proteins for expression in *T. gondii.* Transient transfections (Figure 3 and Figure S2) demonstrate that at least nine of these ROPK proteins colocalized with a rhoptry marker (Figures 3, top, and Figure S2A), and all traffic to the parasitophorous vacuolar membrane or network (a tubular membrane complex within the vacuole; Coppens et al., 2006), although many lack the predicted amphipathic helix known to facilitate membrane association of some ROPK proteins (Reese and Boothroyd, 2009). ROP21 and ROP22 did not colocalize with the rhoptry marker (Figure S2B) but were nevertheless observed in the parasitophorous vacuole; ROP21 was also observed in the host cell cytoplasm (Figure 3, bottom). It is unlikely that these patterns of distribution are attributable to overexpression using a heterologous promoter, as overexpression of secretory proteins more commonly results in staining of the parasite cytoplasm or endoplasmic reticulum, rather than promiscuous secretion (Nishi et al., 2008).

The ROPK Family Is under Diversifying Selection

Multiple sequence alignment of the ROPK family (Figure S3) shows a high degree of divergence (average 16% pairwise identity). Conservation is concentrated within the N-terminal portion of the kinase domain encompassing the activation loop and substrate-binding site. Considerable degeneracy was observed at the initial position of the "KDD" catalytic triad, accounting for the large number of pseudokinases. With the exception of ROP16, the activation loop of all active ROPKs includes the Ser/Thr whose phosphorylation is responsible for regulation in other ePKs. Recent work indicates that ROPK phosphorylation is important mechanisms for regulation (Qiu et al., 2009).

The ROPK phylogenetic tree contains two main clades (Figure 4). One accommodates most previously identified rhoptry proteins (ROP2, ROP4, ROP5, ROP7, ROP8, and ROP18), including many recent duplications. ROP18 is the only active member of this clade, half of which have degenerated into pseudogenes in *Neospora*. A second clade contains most of the active kinases, including ROP16 and most of the novel ROPKs

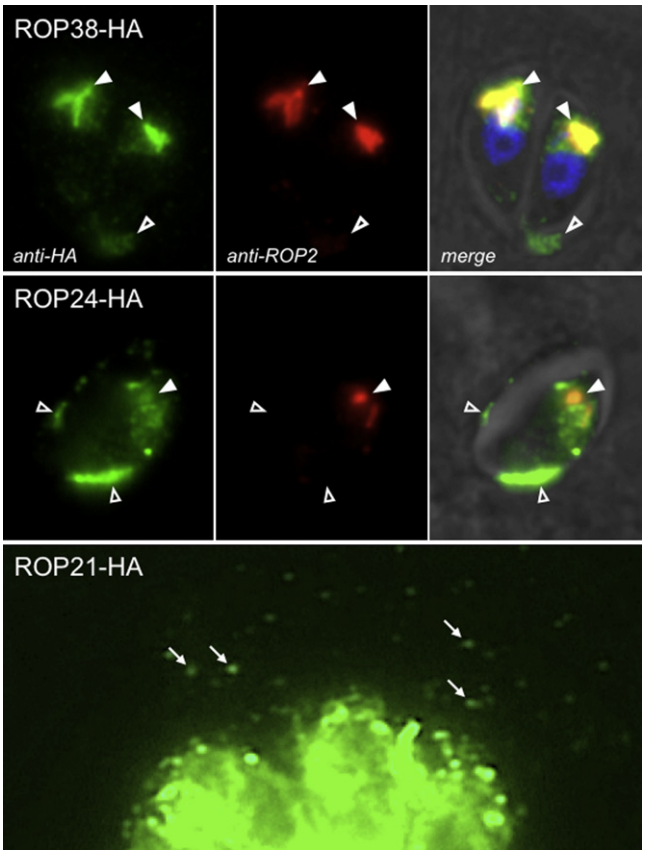

Figure 3. ROPK Localization

Most ROPK proteins localize to the rhoptries and parasitophorous vacuole, as detected by colocalization of antibodies to HA-tagged antigen (green) and native ROP2/3/4 (red) (blue, DAPI; ROP38 only). Filled arrowheads, rhoptries; open arrowheads, patches on the parasitophorous vacuole membrane and/or intravacuolar membrane network. Similar results were obtained for eight additional ROPK proteins (ROP17, ROP19, ROP20, ROP23, ROP25, ROP38, ROP39, and ROP40; Figure S2A). ROP21 and ROP22 do not colocalize with rhoptry proteins but are secreted into the parasitophorous vacuole and can be seen within the host cytoplasm late during infection (Figure S2B).

described in this report. Several derive from recent duplications, including ROP38, ROP29, ROP19, and two degenerate ROPKs on chromosome VI, and ROP42, ROP43, and ROP44 on chromosome Ib. The N-terminal portion of rhoptry kinases has been implicated in secretory targeting (Reese and Boothroyd, 2009), and this domain is highly conserved in the ROP2/ROP18 group, contributing to confidence in the monophyly of this clade. The ROPK tree retains the same basic two clade structure even when the N-terminal domain is excluded from analysis, however (data not shown).

Amplification of the ROPK family clearly preceded the divergence of *T. gondii* and *N. caninum*, as most genes are represented by orthologs in both species (Figure 4A; tick marks indicate branch points of *Neospora* orthologs). Comparing these genes indicates that all but one exhibit nucleotide sequence identity equal to or greater than the observed amino acid conservation, suggesting diversifying selection (Table S3). The ratio of nonsynonymous to synonymous substitutions (dN/dS) is commonly used as a marker of evolutionary pressure, with values > 1 indicating positive selection. Because selection is only expected to apply to a small subset of amino acids, we employed a likelihood ratio test to assess whether the observed data are better explained by a model including or excluding sites with dN/dS > 1, i.e., site-specific positive selection (PAML models M7 and M8; Yang, 1997). Sixteen ROPK proteins show signs of site-specific positive selection ($p < 0.01$), including all *T. gondii* specific duplications (Figure 4A, Table S3). None of the ROPK show signs of positive selection over the entire gene (average dN/dS), suggesting that selection occurs at a few specific sites, although the available sequence data preclude identification of specific sites, as partitioning the protein leaves too few independent sites in the alignment for reliable estimation (it will be interesting to revisit this question as sequences for other coccidian ROPK genes become available). It has long been known that expanded gene families show relaxed constraints and are likely to show rapid divergence. This may be the case for ROP38/ROP29/ROP19, where nucleotide alignments reveal independent triplication in both *T. gondii* and *N. caninum* but amino acid alignments suggest functional conservation (or convergence) between *Tg*ROP38 and *Nc*ROP19.2 (Figure 4B).

Taking advantage of the complete genome sequence available for three representative lineages (GT1, ME49, VEG; http://toxodb.org/), single nucleotide polymorphisms (SNPs) were identified in 30 ROPK genes (Table S3). The ROPK family is significantly more polymorphic than the genome, the secretome (SignalP+), or the kinome as a whole ($p = 0.09$; Figure 5A), and several ROPKs (ROP5, ROP16, ROP17, ROP18, ROP19, ROP24, ROP26, ROP39, and ROP40) also show a high ratio of nonsynonymous to synonymous polymorphisms (Table S3), although more extensive sampling will be required to determine whether any of these genes is under selection at the population level.

The ROPK Family Is Differentially Regulated among *T. gondii* Strains, and during Tachyzoite-to-Bradyzoite Transition

As previous studies on ROP18 showed that expression levels are an important for virulence (Saeij et al., 2007), we exploited genome-wide expression profiling to identify other differentially expressed ROPK genes. RNA from the rapidly growing tachyzoite stage of five *T. gondii* isolates (representing all three major lineages common in the US: type 1, RH and GT1; type 2, PRU and ME49; type 3, VEG) was hybridized to an Affymetrix microarray (http://toxodb.org/) and analyzed as described in the Experimental Procedures. Ninety percent of ROPKs are expressed in tachyzoites (Figure 5B)—a significantly higher fraction than observed for the entire genome (75%), kinome (75%), or secretome (67%); ROP4/7, ROP11, and ROP40 are among the most highly expressed genes in the genome. Others ROPK genes display dramatic differences in expression relative to the RH reference (Figure 5C), including ROP18 (>100 times lower in VEG), ROP38 (up > 64 times in VEG; > 8 times in ME49), and ROP35 (up > 16 times in VEG).

Differentiation of tachyzoites into bradyzoite "tissue cysts" is among the most biologically and clinically significant events in *T. gondii* biology (Dzierszinski et al., 2004). To explore changes in gene expression during this conversion, Prugniaud strain parasites were subjected to alkaline conditions or CO_2 starvation (Bohne et al., 1999), following known bradyzoite markers as

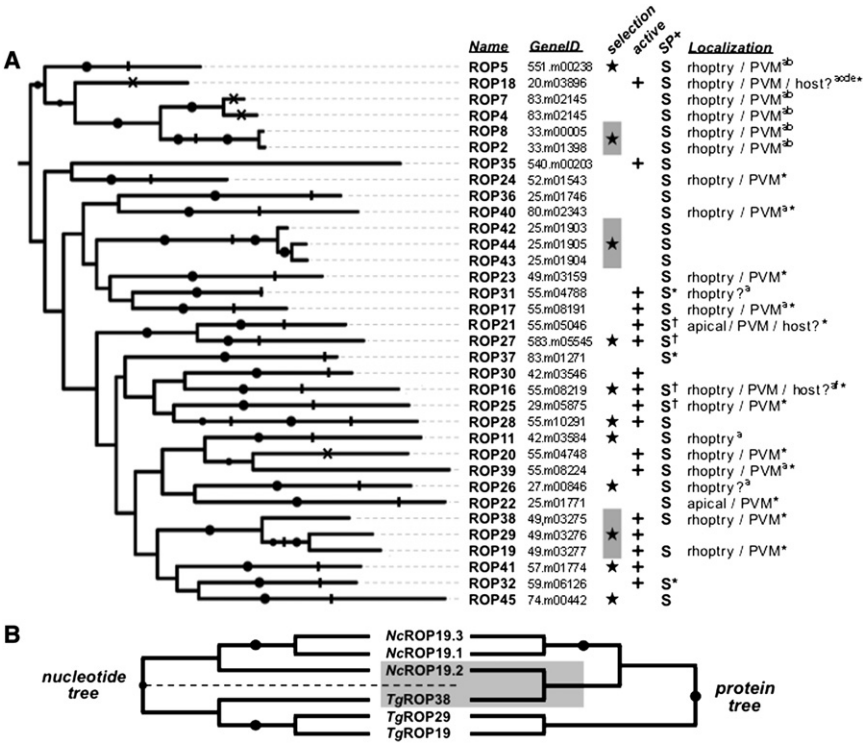

Figure 4. Evolution of the ROPK Family of Secreted Kinases

(A) Phylogenetic representation of 34 *T. gondii* ROPK proteins (excluding seven degenerate kinases, three genes with large insertions in the kinase domain and tandem duplication of ROP2; Table S2). Maximum Likelihood tree based on full-length alignments (Figure S3); black dots indicate > 70% bootstrap support (100 replicates). Tick marks show the branch points for *Neospora caninum* orthologs and pseudogenes (crosses); stars indicate positive selection (PAML codeml); shading indicates *T. gondii*-specific amplifications. +, putative catalytic triad; S, putative signal peptide or signal anchor; †, nuclear localization signal (low confidence). Superscripts indicate references for subcellular localization: [a]El Hajj et al., 2006; [b]Saeij et al., 2006; [c]Taylor et al., 2006; [d]El Hajj et al., 2007a; [e]El Hajj et al., 2007b; [f]Saeij et al., 2006; *, this report.

(B) Phylogenetic trees and chromosomal synteny indicate that the ROP19 clades were independently triplicated in *Toxoplasma gondii* (*Tg*) and *Neospora caninum* (*Nc*). Discordance between nucleotide and protein trees suggests selection for conserved function in *Tg*TOP38 and *Nc*ROP19.2.

controls (see the Experimental Procedures). Only 6% of the entire *T. gondii* genome (8% of the secretome, 5% of the kinome) was differentially expressed, versus 48% of the ROPK family (Figure 5D; Table S3). Most of these ROPK genes were downregulated during differentiation, but ROP28 and ROP38 were induced ~5-fold.

ROP38 Dramatically Alters Host-Cell Responses to Infection

The evolutionary and functional characterization outlined above demonstrates that the ROPK family exhibits various attributes—stage- and strain-specific expression, secretion, positive selection—likely to be associated with important aspects of parasite biology. ROP18 emerges as being of interest, and this gene has previously been shown to play an important role in regulating parasite virulence (Taylor et al., 2006). ROP38 also emerges from this analysis: its ancestor gene was triplicated independently in *T. gondii* and *N. caninum*, and ROP38 exhibits hallmarks of selection (Figure 4). ROP38 is also among the most profoundly regulated genes in the parasite genome (http://toxodb.org/), and it is the only ROPK gene that is both differentially expressed between strains and induced during differentiation (Figures 5C and 5D).

In order to explore the biological significance of ROP38, RH strain parasites (which normally express this gene at very low levels) were engineered to express an HA-tagged transgene under control of the β-tubulin promoter (RH-ROP38), and a parallel mutant was engineered to overexpress HA-tagged ROP21, which lacks indicators of selection or differential expression (Figures 4 and 5) but traffics into the host cell (Figure 3, bottom). The tagged products of both transgenes were found to associate with parasitophorous vacuole membranes (PVM or PV network), and microarray analysis demonstrated upregulation by > 2^5-fold (Table S4), raising ROP38 expression to the levels typically observed in wild-type VEG strain parasites. Parallel changes were also observed in steady-state RNA abundance for several *T. gondii* genes, including TGME49_116390, a coccidian gene of unknown function that is strongly upregulated in both the RH-ROP21 and RH-ROP38 transgenics. Overall, however, these parasites are more notable for their similarities than their differences. In vitro replication of the RH-ROP38 line was comparable to wild-type RH (doubling time ~6.8 hr), as was virulence in mice (100% morbidity by 10 days after intraperitoneal inoculation of Balb/c mice with 100 tachyzoites), in contrast to VEG strain parasites, which replicate slowly and are relatively avirulent (Jerome et al., 1998; Saeij et al., 2005).

Illumina arrays were used to examine the effects of parasite ROPK expression on host cell transcript levels (Figure 6A), providing a far more detailed picture than previously available for *T. gondii* infection (Blader et al., 2001). VEG strain parasites significantly increase transcript levels of 400+ host cell genes and reduce levels of > 250, while RH parasites exert a far more dramatic effect, reproducibly upregulating > 5000 genes and downregulating > 1000. Transcription factors (c-*fos*, EGR2), cytokines (NAMPT, CXCL1), and kinases (especially those associated with MAPK signaling) are prominent among host genes upregulated by RH infection (Table S5). Modulation of MAPK signaling by *T. gondii* has been reported previously (Kim et al., 2004; Molestina et al., 2008), and increased c-*fos* and CXCL1 expression was confirmed by quantitative RT-PCR (data not shown).

Remarkably, expression of the ROP38 transgene suppressed most of the transcriptional changes induced by RH strain

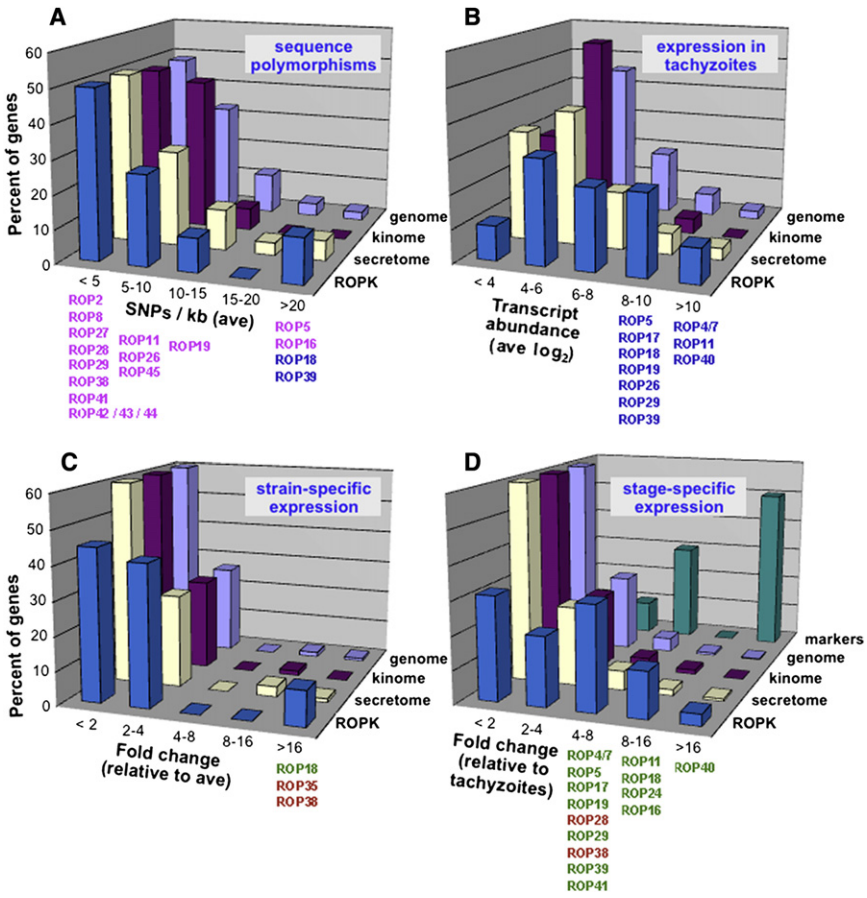

Figure 5. Functional Genomics of the ROPK Family

Shown is distribution of (A) polymorphism densities (SNPs/kb, comparison between strains GT1, ME49, VEG), (B) steady-state transcript abundance in tachyzoites (average across strains), (C) transcriptional regulation between strains (maximal $\log_2$ fold change for any strain relative to RH), and (D) transcriptional regulation during tachyzoite-bradyzoite differentiation (Pru strain; 72 hr postinduction at pH 8.2). In all panels, distribution of ROPK family members (blue) is significantly different from the entire genome (lavender, http://toxodb.org/), kinome (magenta, Table S1), or secretome (beige, Signal P+); turquoise, bradyzoite reference markers (see the Experimental Procedures). Names highlight genes of particular interest: magenta, under selection in Figure 4A; red, upregulated; green, downregulated (Table S2).

parasites (Table S5, Figure 6). Infection with RH-ROP38 parasites significantly altered the expression of only ~400 host cell genes—an effect more similar to the impact of infection with the avirulent VEG strain than that with the RH parental parasite line. For example, c-*fos* was induced > 16 times by RH infection but < 4 times by RH-ROP38 infection, and not at all by the VEG strain; similar effects were observed for CXCL1, EGR2, NAMPT, and many other genes. Quantitative RT-PCR showed a 2.8-fold lower CXCL1 levels in RH-ROP38 parasites relative to RH controls (average of two replicate experiments on each of two independent ROP38 transgenics), validating the 4.5-fold reduction observed on Illumina arrays (first line in Table S5).

Genome-wide, the impact of ROP38 on expression of host genes constitutes an ~2-fold suppression of the effects caused by infection with RH strain parasites (Figure 6B). Observed differences in c-*fos* and CXCL1 expression were confirmed by quantitative RT-PCR, and modulation of MAPK signaling pathways was confirmed by phosphoERK immunoblots (Figure S4). Increased expression of ROP21 did not significantly affect host gene expression (Figure S5B, which also serves to illustrate the reproducibility of independent biological experiments), demonstrating that the profound effect of ROP38 on host cell transcription is not simply the consequence of expressing any ROPK protein in the host cell cytoplasm (although note that ROP21 differs from ROP38 in its localization within the parasite and infected host cell; Figure 3 and Figure S2B).

To further explore the effects of *T. gondii* ROP38 on human host cells, we examined the enrichment of functional annotation (GO terms, KEGG pathways, Interpro domains, UniProt keywords) associated with parasite-induced or -repressed genes (Dennis et al., 2003; Huang et al., 2009). Terms associated with host genes upregulated by RH infection include transcriptional control, signaling, proliferation/apoptosis, sterol biosynthesis, and cell-cycle regulation (Table S6, Figure 7). Organellar functions, including mitochondrial metabolism and vesicular transport, were enriched among transcripts downregulated by RH infection (Table S6). VEG parasites also induced genes related to cell-cycle control and sterol biosynthesis, but not those associated with transcriptional control, signaling, or negative regulation of apoptosis, and did not downregulate mitochondrial metabolism. As noted above, ROP38 expression counteracts many of the effects of RH infection, specifically by downregulating RH-induced transcription factors and signaling molecules associated with proliferation/repression of apoptosis, and upregulating mitochondrial function (Figure 7; Table S6).

DISCUSSION

The *T. gondii* Kinome and the ROPK Family of Secreted Kinases

The 108 predicted active *T. gondii* kinases in Figure 1 (20% more than *Plasmodium*, but < 25% the size of the human kinome) constitute the largest apicomplexan kinome described to date (Table S1). This classification confirms previous experimental evidence, while expanding our knowledge of signaling cascades likely to be present in apicomplexans. Receptor kinases, including tyrosine kinases and receptor guanylate cyclases, are completely absent (although tyrosine kinases have been found in some unicellular species; Shiu and Li, 2004). The

A

	VEG vs HFF	RH vs HFF	RH-ROP38 vs HFF	RH-ROP38 vs RH
> 2X:	204	546	121	158
Induced:	439	5239	221	559
No change:	5950	6011	6913	11107
Repressed:	260	1078	193	704
> 2X:	196	700	179	225
undetectable:	8906	3227	8228	3185

B

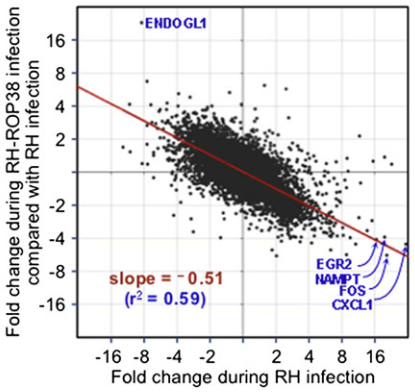

Figure 6. The Effect of ROP38 on Host Cell Gene Expression

(A) Host cell gene expression changes induced by infection with VEG, RH, or RH-ROP38 *T. gondii*.

(B) ROP38 causes a 2-fold global reduction of host response to infection. The transcriptional impact of RH-ROP38 infection (*y* axis) is negatively correlated with the impact of RH infection (x axis), $r^2 = 0.59$.

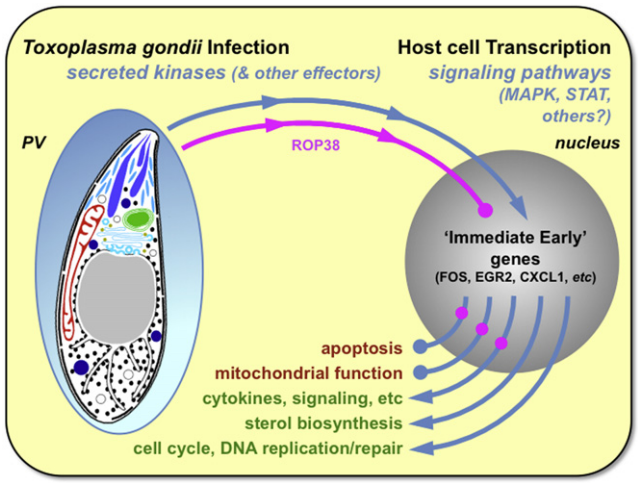

Figure 7. Effects of *T. gondii* Infection (±ROP38) on Host Cell Transcripts

Host cell transcripts affected by infection with RH strain parasites (blue) are enriched in functional terms associated with transcriptional control, signaling, and metabolic pathways, as noted (arrows, upregulated; dots, downregulated). Increased expression of ROP38 in the parasite selectively reverses the effects of RH infection (magenta; Table S5).

T. gondii genome is predicted to encode a single STE kinase, and *P. falciparum* appears to lack MAPKK orthologs entirely, suggesting limited ability to exploit traditional MAPK cascades. A large fraction of the *T. gondii* kinome lacks the canonical catalytic triad, although we note that mounting evidence suggests that at least some such ''pseudokinases'' are able to phosphorylate subtrates (Kannan and Taylor, 2008; Kornev and Taylor, 2009).

Even where parasite enzymes can be classified into one of the major kinase groups, they are often highly divergent; 78 of the 108 ''active'' *T. gondii* kinases lack an obvious ortholog in human or yeast (Table S2). Far more apicomplexan kinases share orthologs between *Toxoplasma* and *Plasmodium* (Figure S1, Table S2). While secreted protein kinases are rare in eukaryotes, ~15% of apicomplexan kinases are predicted to be secreted outside of the parasite (Figure 1), suggesting that they may affect host-pathogen interactions. Most of these kinases belong to parasite-specific families, including the *P. falciparum* FIKKs and the *T. gondii* ROPKs. Two ROPK proteins (ROP16 and ROP18) have been reported as the only active members of a 9–12 member family dominated by kinase-like proteins lacking catalytic activity (El Hajj et al., 2006; Sinai, 2007). Our analysis reveals that the ROPK family contains at least 44 members, including 16 predicted to be active (Figure 2). Several ROPK genes are tandemly duplicated in the *T. gondii* genome (ROP2/ROP8, ROP4/ROP7, ROP42/ROP43/ROP44, ROP19/ROP29/ROP38), and comparison with the genome assembly (http://toxodb.org/) reveals that many lie at contig breaks, suggesting the presence of additional tandemly duplicated ROPK genes.

ROP proteins were originally defined by their association with the rhoptries—part of the distinctive apical complex of secretory organelles defining the phylum apicomplexa. Rhoptries facilitate the interaction of *T. gondii* with host cells, including establishment of the parasitophorous vacuole within which these obligate intracellular parasites survive and replicate (Boothroyd and Dubremetz, 2008). Subcellular localization using tagged transgenes demonstrates that most ROPK proteins target to the rhoptries and are secreted into the parasitophorous vacuole, where they associate with the vacuolar membrane and/or tubulovesicular network (Figures 3 and 4, Figure S2). A recent report (Reese and Boothroyd, 2009) described an N-terminal amphipathic α helix associated with some ROPK genes (particularly the ROP2/ROP18 clade; Figure 4A) that facilitates trafficking to the parasitophorous vacuole membrane when expressed in the cytoplasm of infected mammalian cells. We find that ROPK proteins lacking this domain (e.g., ROP38, formerly known as ROP2L5; Figure 3, top) are also able to associate with the tubulovesicular network, via unknown mechanisms.

Leveraging Genomic-Scale Data Sets to Prioritize ROPK Genes for Further Analysis

Because advantageous traits frequently emerge through gene duplication and functional divergence (Ohno et al., 1968), expanded gene families can provide useful insights into organismal biology. Several members of the ROPK family have

previously been shown to regulate *T. gondii* virulence and host-pathogen interactions: ROP18 was identified as a virulence factor by genetic mapping (Saeij et al., 2006; Taylor et al., 2006), and ROP16 alters phosphorylation of host STAT3/6 (Saeij et al., 2007). Having defined the full spectrum of ROPK genes in the *T. gondii* genome, can we identify those that are most likely to play important roles in parasite biology?

Although the ROPK family forms a single clade distinct from previously characterized kinase families (Figure 1), it is quite diverse: the subtree defined by ROP16 (Figure 4) exhibits greater protein sequence diversity than the entire human AGC family (PKA, PKC, etc.), and the subtrees defined by ROP16 and ROP18 are as different from each other as PKA versus CAM kinases. Many ROPK proteins lack the complete catalytic triad required for kinase activity, but even inactive pseudokinases may play important roles in substrate binding and/or allosteric interactions (Boudeau et al., 2006).

Comparative genomics shows that the ROPK family is restricted to coccidia: the ROPK HMM identifies family members in *Eimeria*, *Neospora*, and *Toxoplasma*, but not *Plasmodium*, *Babesia*, or *Theileria.* The apicomplexa also possess another family of secreted kinases (FIKK; Nunes et al., 2007), which is expanded in *P. falciparum* only. *Neospora* and *Toxoplasma* share a recent common ancestor (Frenkel and Smith, 2003), and comparison of their ROPK families is particularly informative (Figure 4). Most ROPK genes display a 1:1 correspondence between species, but there is evidence for rapid diversification. ROP4, ROP7, ROP18, and ROP20 correspond to *N. caninum* pseudogenes, and *Nc*ROP46 corresponds to a pseudogene in *T. gondii* (Table S3). Independent tandem duplications are also evident (Figure 4). It is interesting to note that DNA sequence is more highly conserved than protein sequence for most ROPK genes, and there is evidence for positive selection in 16 family members (Figure 4A, Table S3).

Functional genomics data may also be exploited to identify genes of likely biological interest. Transfection studies have shown that the importance of ROP18 in parasite virulence is mediated through the regulation of transcript abundance. (Saeij et al., 2006). Analysis of expression in the parasite kinome, secretome, and genome shows that a disproportionate number of ROPK genes are highly expressed in parasite tachyzoites (Figure 5B), differentially expressed in different strains (Figure 5C), and/or transcriptionally regulated during differentiation (most are downregulated in bradyzoites; Figure 5D). Considerable difference is observed even between recently-duplicated ROPK genes (compare ROP 19 versus ROP38; Table S3).

Integrating these comparative and functional genomics analyses, ROP18 and ROP38 display the most striking indicators of biological significance (Table S3): both show evidence of evolutionary selection at the population or species level (Khan et al., 2009; Figure 4), both display > 16-fold differences in expression level between strains (Figure 5C), and both are differentially regulated during bradyzoite differentiation (Figure 5D). ROP38 is particularly intriguing, as this clade has been independently expanded in both *Toxoplasma* and *Neospora* (Figure 4B), with *Nc*ROP19.2 and *Tg*ROP38 retaining (or converging upon) similar sequence while the other paralogs have diverged. Preliminary efforts to delete this gene have been unsuccessful, perhaps arguing for essentiality, as suggested by its high degree of conservation among *T. gondii* strains and between coccidian species.

ROP38 Is a Potent Regulator of Host-Cell Transcription

Previous studies showed that infection with *T. gondii* induces transcriptional changes in the host cell (Blader et al., 2001; Saeij et al., 2007), but the extent of change has not been accurately defined. We identified ~700 genes whose expression is significantly up- or downregulated 24 hr after infection with VEG strain parasites (~400 > 2-fold). RH strain parasites exhibit a strikingly different pattern, significantly up- or downregulating the expression of > 6000 host cell genes (~1200 > 2-fold), including all previously described examples of induction by RH infection (CXCL1, EGR2, HIF1α, etc. [Blader et al., 2001; Spear et al., 2006; Phelps et al., 2008]; Figure 6, Table S5). Functional analysis reveals that infection with any *T. gondii* strain induces transcription of host genes associated with cell-cycle control, DNA replication/repair, RNA processing, and sterol biosynthesis (Table S6), but RH infection specifically induces transcriptional control and signaling pathways (including inhibitors of apoptosis, especially MAPKs), and represses organellar pathways (Table S6). The striking changes in host cell transcription induced by this strain may be responsible for some of its unusual biological characteristics (Saeij et al., 2005).

As noted above, several lines of evidence suggest that ROP38 is functionally important for *T. gondii* biology, and indeed, transgenic expression in RH strain parasites dramatically diminishes the impact of infection on host expression profiles. Increasing ROP38 expression to levels normally observed in VEG strain parasites significantly alters transcript levels of > 1200 host genes (Figure 6B, Table S5). Although there is no evidence that significant quantities of ROP38 leave the parasitophorous vacuole (Figure 3, top), overexpression specifically downregulates host transcription factors and genes associated with regulation of signaling and apoptosis/proliferation (all strongly induced by RH infection, but not altered by infection with VEG strain parasites). These genes include the transcription factors c-*fos* and EGR2 (known to be induced in a rhoptry-dependent manner; Phelps et al., 2008). Functional clustering implicates host MAPK cascades, and preliminary results show that the kinetics of ERK phosphorylation in RH-ROP38-infected cells is distinct from the response to the parental RH line (Figure S4). We note, however, that the dramatic changes of parasite ROP38 expression on host transcript levels do not appear to affect parasite replication in vitro or virulence in vivo.

This report highlights the potential of integrating multiple, diverse genomic-scale data sets to aid in the discovery of biologically important molecules. Comparative genomic approaches defined the parasite kinome, revealing the full diversity of rhoptry kinases, and evolutionary genomic analysis of positive selection indicates the importance of this family. Functional genomics data sets facilitated the prioritization of ROPK family members for further exploration, and transcriptional profiling of the host cell, coupled with functional clustering, highlights pathways likely to be regulated by parasite infection and a role for ROP38 in the regulation of host transcription. In aggregate, integrating phylogenetics with functional genomic analysis and experimental manipulation of transgenic parasites has expanded our understanding of secreted kinases and their role

as effector molecules during host cell infection. It will be interesting to determine the targets of these kinases, and whether they act directly or indirectly on host cell signaling pathways.

EXPERIMENTAL PROCEDURES

T. gondii Protein Kinase Phylogeny and Classification

The *T. gondii* proteome (http://toxodb.org/) was searched for protein kinase domains using Pfam hidden Markov model (HMM) PF00069 (http://pfam.sanger.ac.uk/) and the HMMer package (Eddy, 1998; cutoff −150, E < 1). Matches were expanded to include orthologs identified by OrthoMCLv1 (http://orthomcl.org/). Identical results were obtained using the SMART HMMs SM00219 and SM00220 and Interpro IPR017442. Active kinases were defined based on a putative KDD catalytic triad and HMM score >−100; proteins scoring from −100 to −150 or lacking a complete triad were designated as pseudokinases.

Experimentally validated representatives from all major kinase groups were selected from the published kinomes of *Homo sapiens* (Manning et al., 2002b), *Saccharomyces cerevisiae* (Hunter and Plowman, 1997), and *Plasmodium falciparum* (Ward et al., 2004), and aligned with active *T. gondii* kinase domains using HMMer, with PF00069 as a reference. Kinase subdomains were assessed by manual inspection, removing other regions from the alignment. PHYML 3.0 (Guindon and Gascuel, 2003; Guindon et al., 2005) was used for phylogeny reconstruction; 100 and 1000 bootstrap replicates yielded comparable results. Each pseudokinase was classified independently by constructing a new alignment and ML tree (100 bootstrap replicates), using the set of active kinases noted above.

Identification of the ROPK Family and Analysis of Divergence

A ROPK-specific HMM was constructed based on kinase domain alignment of all active ROPKs without insertions: ROP16, ROP17, ROP18, ROP19, ROP21, ROP25, ROP27, ROP29, ROP30, ROP31, ROP32, ROP35, ROP38, and ROP41. Truncated genes were identified based on higher sequence identity to the ROPK HMM than any other sequence in the kinome. *Neospora* orthologs were identified by conceptual translation of syntenic regions from http://toxodb.org/. MUSCLE (Edgar, 2004) was used for multiple sequence alignment of full-length ROPK proteins (excluding columns with gaps in > 90% of sequences), and PHYML 3.0 was used for phylogeny reconstruction. For analysis of site-specific positive selection, full-length orthologous protein sequence pairs from *T. gondii* and *N. caninum* (along with recent paralogs) were aligned with MUSCLE, and the underlying nucleic acid sequence alignments used as input for PAML codeml (Yang, 1997), using nested models M7 and M8 to generate a likelihood ratio test; sequences under positive selection were determined based on a standard chi-square probability distribution.

Parasite Transfection and Immunolabeling

ROP16, ROP17, ROP18, ROP19, ROP20, ROP21, ROP22, ROP23, ROP24, ROP25, ROP38, ROP39, and ROP40 were amplified from an RH strain *T. gondii* cDNA library using gene specific primers, subcloned into the NheI/BglII cloning sites in *p*tub-HA/sagCAT (Nishi et al., 2008), and sequenced to confirm fidelity. All gene-specific primers amplified a single gene product. Transfections were performed as previously described (Roos et al., 1994) and examined by immunofluorescence 24 hr posttransfection; stable transgenic lines expressing either 55.m05046 (RH-ROP21) or 49.m03275 (RH-ROP38) were isolated by chloranphenicol selection and cloned by limiting dilution.

For subcellular localization, 4×10^5 parasites were inoculated into confluent host cell monolayers on 22 mm glass coverslips, incubated 24 hr at 37°C, fixed in 3.7% paraformaldehyde, permeabilized with 0.25% Triton X-100, and stained with (1) mouse monoclonal anti-HA conjugated to Alexa 488 (Roche; 1:1,000), (2) mouse monoclonal anti-*Tg*ROP2/3/4 (1:50,000, kindly provided by Jean François Dubremetz) followed by Alexa 594-conjugated goat anti-mouse (Molecular Probes; 1:5,000), and (3) 2.8 μM 4′,6-diamidino-2-phenylindole dihydrochloride (DAPI, Invitrogen). This antibody recognizes proteins in the rhoptries, with PVM localization observed only immediately after infection, presumably due to inaccessibility of the epitope (Jean François Dubremetz, personal communication). Samples were visualized on a Leica DM IRBE inverted microscope equipped with a motorized filter wheel, 100W Hg-vapor lamp and Orca-ER digital camera (Hamamatsu). Openlab software (Improvision) was used for all image acquisition and manipulation.

For immunoblot analysis of phosphorylated ERK, *T. gondii*-infected HFF host cells were passed through a 26 ga needle to ensure efficient egress, filtered through a 3 μm filter, and resuspended in CO_2 equilibrated reduced serum medium (Opti-MEM; GIBCO) prior to inoculation into serum-starved confluent HFF cells in 25 cm^2 T flasks (MOI 5:1). At various times postinfection, samples were lysed in ice-cold RIPA buffer containing protease and phosphatase inhibitors (Sigma #P5726; Roche Complete Mini Protease Inhibitor Tablets) and frozen. Lysates were resuspended in Invitrogen NuPage LDS buffer containing 5 mM β-mercaptoethanol, incubated 5 min at 85°C, and 4 μg protein (determined by Bradford assay) loaded on a NuPage 10% BT SDS gel, followed by semidry transfer to nitrocellulose membrane, and sequential 1 hr incubations in 10% nonfat milk in PBS, monoclonal antidiphosphorylated ERK 1&2 (Sigma, 1:4000) and monoclonal anti-α-tubulin (Sigma, 1:7500), and HRP conjugated goat-anti-mouse (Bio-Rad, 1:3000), proteins were visualized by 5 min incubation with Immobilon Western Chemiluminescent HRP substrate (Millipore) and 2 min exposure to BioMaxMR Film (Kodak).

Toxoplasma gondii Expression Profiling

T. gondii parasites were cultivated in vitro in HFF cells (moi 10:1) using standard methods (Roos et al., 1994), and RNA harvested from purified parasites (RNeasy; QIAGEN); quality was ascertained using a spectrophotometer (NanoDrop) and confirmed with an Agilent Bioanalyzer. The Affymetrix Expression 3′ One-Cycle amplification kit was used to prepare labeled cRNA, which was hybridized to a custom *T. gondii* Affymetrix microarray (http://toxodb.org/), and the fluorescent signal collected by excitation at 570 nm and confocal scanning at 3 μm resolution; low-affinity probes and sequences redundant in the parasite genome were excluded (not a concern for ROP38, which harbors few SNPs). *T. gondii* microarray data has been deposited with GEO (GSE22315, GSE22258) and is also available for querying and downloading at http://toxodb.org/.

Two sequential scans were averaged for each microarray feature, and robust multiarray analysis (RMA) normalization used to calculate relative RNA abundance for each gene (log_2 values). Three replicates of each experiment were used to assign a p value and average log fold change between strains or time points during differentiation using the limma R package (Bioconductor). Tachyzoite expression was determined for strains RH, GT1, Prugniaud, ME49, VEG, RH-ROP38, and RH-ROP21, and bradyzoite differentiation data were obtained at various time points after alkaline induction of the Prugniaud strain (Dzierszinski et al., 2004). The following known bradyzoite markers were used as controls: 9.m03411, 72.m00004, 80.m00010, 55.m00009, 641.m01562, 72.m00003, 59.m03410, 44.m00006, 59.m00008, 44.m00009, and 641.m01563. A power analysis was carried out for all pairwise strain comparisons to establish statistical significance of ≥ 2-fold differential expression, based on a calculated false discovery rate of 8% at p = 0.05 (1% at p = 0.01). Maximum log fold change between strains was defined as the maximum difference between any two strains.

Host-Cell Expression Profiling and Cluster Analysis of Functional Enrichment

Freshly harvested VEG, RH, RH-ROP21, and RH-ROP38 parasites were inoculated and incubated 24 hr before harvest and RNA isolation as described above (moi 20:1). At least three experimental replicates were conducted for each assay. Only samples with comparable ratios of parasite:host RNA (determined by Agilent Bioanalyzer) were used for hybridization. Illumina HumanRef8_V2 and V3 microarrays were hybridized according to the manufacturer's instructions, scanned on a BeadScan unit, analyzed using the gene expression module of GenomeStudio software (Illumina), and deposited in the GEO database (GSE22402). Due to incompatibilities between V2 and V3 probes, only 15,554 human genes were analyzed in comparisons across all samples. Genes were defined as expressed if they displayed a detection p value < 0.05 in at least one sample. Comparison with uninfected or RH-infected HFF cells was used to determine differential expression based on an Illumina differential expression score of > |30| (nominally equivalent to p < 0.001).

Data sets for functional annotation were defined as the set of genes exhibiting statistically significant up- or downregulation > 2 × . Enrichment of functional annotation (GO Biological process, GO Molecular Function,

KEGG pathways, Biocarta pathways, InterPro and PFAM domains, SwissProt and Protein Information Resource keywords) was assessed using the DAVID package (Dennis et al., 2003; Huang et al., 2009). Enrichment relative to the 15,554 Illumina probes common to V2 and V3 arrays was defined as a $p < 0.05$ with at least three genes per term per data set. Fuzzy heuristical clustering was performed using kappa similarity > 0.3–0.35 and requiring an enrichment p value geometric mean >0.05.

ACCESSION NUMBERS

T. gondii microarray data have been deposited with GEO under accession numbers GSE22315 and GSE22258 and are also available for querying and downloading at http://toxodb.org/. Illumina HumanRef8_V2 and V3 microarrays were hybridized according to the manufacturer's instructions, scanned on a BeadScan unit, analyzed using the gene expression module of GenomeStudio software (Illumina), and deposited in the GEO database under accession number GSE22402.

SUPPLEMENTAL INFORMATION

Supplemental Information includes five figures and six tables and can be found with this article online at doi:10.1016/j.chom.2010.07.004.

ACKNOWLEDGMENTS

We thank the Roos laboratory and *T. gondii* and ISCB research communities for helpful discussions. This work was supported by research grants AI28724, AI077268, AI075846, and RR016469 from the National Institutes of Health.

Received: December 23, 2009
Revised: April 21, 2010
Accepted: July 14, 2010
Published: August 18, 2010

REFERENCES

Anamika, Srinivasan, N., and Krupa, A. (2005). A genomic perspective of protein kinases in *Plasmodium falciparum*. Proteins *58*, 180–189.

Blader, I.J., Manger, I.D., and Boothroyd, J.C. (2001). Microarray analysis reveals previously unknown changes in *Toxoplasma gondii*-infected human cells. J. Biol. Chem. *276*, 24223–24231.

Bohne, W., Holpert, M., and Gross, U. (1999). Stage differentiation of the protozoan parasite *Toxoplasma gondii*. Immunobiology *201*, 248–254.

Boothroyd, J.C., and Dubremetz, J.F. (2008). Kiss and spit: the dual roles of *Toxoplasma* rhoptries. Nat. Rev. Microbiol. *6*, 79–88.

Boudeau, J., Miranda-Saavedra, D., Barton, G.J., and Alessi, D.R. (2006). Emerging roles of pseudokinases. Trends Cell Biol. *16*, 443–452.

Bradley, P.J., Ward, C., Cheng, S.J., Alexander, D.L., Coller, S., Coombs, G.H., Dunn, J.D., Ferguson, D.J., Sanderson, S.J., Wastling, J.M., and Boothroyd, J.C. (2005). Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in *Toxoplasma gondii*. J. Biol. Chem. *280*, 34245–34258.

Bradley, P.J., and Sibley, L.D. (2007). Rhoptries: an arsenal of secreted virulence factors. Curr. Opin. Microbiol. *10*, 582–587.

Carey, K.L., Jongco, A.M., Kim, K., and Ward, G.E. (2004). The *Toxoplasma gondii* rhoptry protein ROP4 is secreted into the parasitophorous vacuole and becomes phosphorylated in infected cells. Eukaryot. Cell *3*, 1320–1330.

Carruthers, V.B., and Sibley, L.D. (1997). Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts. Eur. J. Cell Biol. *73*, 114–123.

Coppens, I., Dunn, J.D., Romano, J.D., Pypaert, M., Zhang, H., Boothroyd, J.C., and Joiner, K.A. (2006). *Toxoplasma gondii* sequesters lysosomes from mammalian hosts in the vacuolar space. Cell *125*, 261–274.

Dardick, C., Chen, J., Richter, T., Ouyang, S., and Ronald, P. (2007). The rice kinase database. A phylogenomic database for the rice kinome. Plant Physiol. *143*, 579–586.

Dennis, G., Jr., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., and Lempicki, R.A. (2003). DAVID: database for annotation, visualization, and integrated discovery. Genome Biol. *4*, 3.

Dzierszinski, F., Nishi, M., Ouko, L., and Roos, D.S. (2004). Dynamics of *Toxoplasma gondii* differentiation. Eukaryot. Cell *3*, 992–1003.

Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755–763.

Edgar, R.C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics *5*, 113.

El Hajj, H., Demey, E., Poncet, J., Lebrun, M., Wu, B., Galéotti, N., Fourmaux, M.N., Mercereau-Puijalon, O., Vial, H., Labesse, G., and Dubremetz, J.F. (2006). The ROP2 family of *Toxoplasma gondii* rhoptry proteins: proteomic and genomic characterization and molecular modeling. Proteomics *6*, 5773–5784.

El Hajj, H., Lebrun, M., Arold, S.T., Vial, H., Labesse, G., and Dubremetz, J.F. (2007a). ROP18 is a rhoptry kinase controlling the intracellular proliferation of *Toxoplasma gondii*. PLoS Pathog. *3*, e14.. 10.1371/journal.ppat.0030014.

El Hajj, H., Lebrun, M., Fourmaux, M.N., Vial, H., and Dubremetz, J.F. (2007b). Inverted topology of the *Toxoplasma gondii* ROP5 rhoptry protein provides new insights into the association of the ROP2 protein family with the parasitophorous vacuole membrane. Cell. Microbiol. *9*, 54–64.

Foth, B.J., and McFadden, G.I. (2003). The apicoplast: a plastid in *Plasmodium falciparum* and other Apicomplexan parasites. Int. Rev. Cytol. *224*, 57–110.

Frenkel, J.K., and Smith, D.D. (2003). Determination of the genera of cyst-forming coccidia. Parasitol. Res. *91*, 384–389.

Gajria, B., Bahl, A., Brestelli, J., Dommer, J., Fischer, S., Gao, X., Heiges, M., Iodice, J., Kissinger, J.C., Mackey, A.J., et al. (2008). ToxoDB: an integrated *Toxoplasma gondii* genome database. Nucleic Acids Res. *36*, D553–D556.

Goldberg, J.M., Manning, G., Liu, A., Fey, P., Pilcher, K.E., Xu, Y., and Smith, J.L. (2006). The *Dictyostelium* kinome. Analysis of the protein kinases from a simple model organism. PLoS Genet. *2*, e38.. 10.1371/journal.pgen.0020038.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. *52*, 696–704.

Guindon, S., Lethiec, F., Duroux, P., and Gascuel, O. (2005). PHYML Online. A web server for fast maximum likelihood-based phylogenetic inference. Nucleic Acids Res. *33*, W557–W559.

Hanks, S.K., and Hunter, T. (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. *9*, 576–596.

Hill, D., and Dubey, J.P. (2002). *Toxoplasma gondii:* transmission, diagnosis and prevention. Clin. Microbiol. Infect. *8*, 634–640.

Huang, D.W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Hunter, T., and Plowman, G.D. (1997). The protein kinases of budding yeast: six score and more. Trends Biochem. Sci. *22*, 18–22.

Jerome, M.E., Radke, J.R., Bohne, W., Roos, D.S., and White, M.W. (1998). *Toxoplasma gondii* bradyzoites form spontaneously during sporozoite-initiated development. Infect. Immun. *66*, 4838–4844.

Kannan, N., and Taylor, S.S. (2008). Rethinking pseudokinases. Cell *133*, 204–205.

Khan, A., Taylor, S., Ajioka, J.W., Rosenthal, B.M., and Sibley, L.D. (2009). Selection at a single locus leads to widespread expansion of *Toxoplasma gondii* lineages that are virulent in mice. PLoS Genet. *5*, e1000404.. 10.1371/journal.pgen.1000404.

Kim, L., Butcher, B.A., Lee, C.W., Uematsu, S., Akira, S., and Denkers, E.Y. (2004). *Toxoplasma gondii* interferes with lipopolysaccharide-induced mitogen-activated protein kinase activation by mechanisms distinct from endotoxin tolerance. J. Immunol. *172*, 3003–3010.

Kornev, A.P., and Taylor, S.S. (2009). Pseudokinases: functional insights gleaned from structure. Structure *17*, 5–7.

Liu, Q., Crammer, K., Pereira, F.C., and Roos, D.S. (2008). Reranking candidate gene models with cross-species comparison for improved gene prediction. BMC Bioinformatics *9*, 433.

Manning, G., Plowman, G.D., Hunter, T., and Sudarsanam, S. (2002a). Evolution of protein kinase signaling from yeast to man. Trends Biochem. Sci. *27*, 514–520.

Manning, G., Whyte, D.B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002b). The protein kinase complement of the human genome. Science *298*, 1912–1934.

Molestina, R.E., El-Guendy, N., and Sinai, A.P. (2008). Infection with *Toxoplasma gondii* results in dysregulation of the host cell cycle. Cell. Microbiol. *10*, 1153–1165.

Nishi, M., Hu, K., Murray, J.M., and Roos, D.S. (2008). Organellar dynamics during the cell cycle of *Toxoplasma gondii*. J. Cell Sci. *121*, 1559–1568.

Nunes, M.C., Goldring, J.P., Doerig, C., and Scherf, A. (2007). A novel protein kinase family in *Plasmodium falciparum* is differentially transcribed and secreted to various cellular compartments of the host cell. Mol. Microbiol. *63*, 391–403.

Ohno, S., Wolf, U., and Atkin, N.B. (1968). Evolution from fish to mammals by gene duplication. Hereditas *59*, 169–187.

Parsons, M., Worthey, E.A., Ward, P.N., and Mottram, J.C. (2005). Comparative analysis of the kinomes of three pathogenic trypanosomatids: *Leishmania major, Trypanosoma brucei* and *Trypanosoma cruzi*. BMC Genomics *6*, 127.

Phelps, E.D., Sweeney, K.R., and Blader, I.J. (2008). *Toxoplasma gondii* rhoptry discharge correlates with activation of the early growth response 2 host cell transcription factor. Infect. Immun. *76*, 4703–4712.

Plowman, G.D., Sudarsanam, S., Bingham, J., Whyte, D., and Hunter, T. (1999). The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms. Proc. Natl. Acad. Sci. USA *96*, 13603–13610.

Qiu, W., Wernimont, A., Tang, K., Taylor, S., Lunin, V., Schapira, M., Fentress, S., Hui, R., and Sibley, L.D. (2009). Novel structural and regulatory features of rhoptry secretory kinases in *Toxoplasma gondii*. EMBO J. *28*, 969–979.

Reese, M.L., and Boothroyd, J.C. (2009). A helical membrane-binding domain targets the *Toxoplasma* ROP2 family to the parasitophorous vacuole. Traffic *10*, 1458–1470.

Roos, D.S. (2005). Themes and variations in apicomplexan parasite biology. Science *309*, 72–73.

Roos, D.S., Donald, R.G., Morrissette, N.S., and Moulton, A.L. (1994). Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii*. Methods Cell Biol. *45*, 27–63.

Saeij, J.P., Boyle, J.P., and Boothroyd, J.C. (2005). Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host. Trends Parasitol. *21*, 476–481.

Saeij, J.P., Boyle, J.P., Coller, S., Taylor, S., Sibley, L.D., Brooke-Powell, E.T., Ajioka, J.W., and Boothroyd, J.C. (2006). Polymorphic secreted kinases are key virulence factors in toxoplasmosis. Science *314*, 1780–1783.

Saeij, J.P., Coller, S., Boyle, J.P., Jerome, M.E., White, M.W., and Boothroyd, J.C. (2007). *Toxoplasma* co-opts host gene expression by injection of a polymorphic kinase homologue. Nature *445*, 324–327.

Shiu, S.H., and Li, W.H. (2004). Origins, lineage-specific expansions, and multiple losses of tyrosine kinases in eukaryotes. Mol. Biol. Evol. *21*, 828–840.

Sinai, A.P. (2007). The toxoplasma kinase ROP18: an active member of a degenerate family. PLoS Pathog. *3*, e16.. 10.1371/journal.ppat.0030016.

Spear, W., Chan, D., Coppens, I., Johnson, R.S., Giaccia, A., and Blader, I.J. (2006). The host cell transcription factor hypoxia-inducible factor 1 is required for *Toxoplasma gondii* growth and survival at physiological oxygen levels. Cell. Microbiol. *8*, 339–352.

Taylor, S., Barragan, A., Su, C., Fux, B., Fentress, S.J., Tang, K., Beatty, W.L., Hajj, H.E., Jerome, M., Behnke, M.S., et al. (2006). A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen *Toxoplasma gondii*. Science *314*, 1776–1780.

Ward, P., Equinet, L., Packer, J., and Doerig, C. (2004). Protein kinases of the human malaria parasite *Plasmodium falciparum:* the kinome of a divergent eukaryote. BMC Genomics *5*, 79.

World Health Organization. (2009). World Malaria Report (http://who.int/malaria/world_malaria_report_2009).

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. *13*, 555–556.