# THE EVOLUTION OF HOST MITOCHONDRIAL ASSOCIATION AND ITS IMPACT ON TOXOPLASMA GONDII INFECTION

by

# Elizabeth D. English

B.S., Molecular and Cellular Biology, University of Connecticut, 2011

Submitted to the Graduate Faculty of the

Kenneth P. Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2017

# UNIVERSITY OF PITTSBURGH

# KENNETH P. DEITRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Elizabeth D. English

It was defended on

January 11, 2017

and approved by

James M. Pipas, Ph. D., Professor, Biological Sciences

Jeffrey G. Lawrence, Ph. D., Professor, Biological Sciences

Mark Rebeiz, Ph. D., Associate Professor, Biological Sciences

Carolyn B. Coyne, Ph. D., Associate Professor, Microbiology and Molecular Genetics

Dissertation Advisor: Jon P. Boyle, Ph. D. Associate Professor, Biological Sciences

Copyright © by Elizabeth D. English

2017

# THE EVOLUTION OF HOST MITOCONDRIAL ASSOCIATION AND ITS IMPACT ON *TOXOPLASMA GONDII* INFECTION

Elizabeth D. English, Ph. D.

University of Pittsburgh, 2017

The association of intracellular pathogens with host mitochondria has been observed across taxa, from bacterial pathogens, such as Legionella pneumophila and Chlamydia trachomati, to the eukaryotic pathogen Toxoplasma gondii. However the functional impact of host mitochondrial association (HMA) remains difficult to assess in most of these species because in many cases the genes responsible for this phenomenon have not yet been identified. The recent discovery of the T. gondii gene responsible for HMA, Mitochondrial Association Factor 1 (MAFI) has provided us with the tools to begin to understand the evolution and impact of HMA. Here we use multispecies sequence analysis to determine that the MAF1 locus is tandemly duplicated and diversified in both T. gondii and its nearest extant relative Hammondia hammondi, but not another close relative *Neospora caninum*. Using cross-species complementation we find that T. gondii and H. hammondi harbor copies of MAF1 able to mediate HMA, while N. caninum does not. We have begun mutational analysis using naturally occurring HMA+ and HMA- paralogs of *MAF1* in order to determine the portions of MAF1 protein necessary for HMA. Additionally, we have identified the first *in vivo* phenotypes associated with HMA using multiple mouse models, for both acute and chronic infection. Taken together these data indicate that HMA likely evolved via neofunctionalization of a duplicated ancestral MAF1 gene, and that the neofunctionalized, HMA competent copy of *MAF1* provides a selective advantage.

# TABLE OF CONTENTS

PREFACEXV				
1.0		INTRO	DUCTION1	
	1.1	ТС	DXOPLASMA GONDII1	
		1.1.1	Life Cycle1	
		1.1.2	Host range and pathogenesis4	
		1.1.3	Comparisons to closely related species4	
		1.1.4	Non-rodent experimental models7	
		1.1	1.4.1 Cattle	
		1.1	1.4.2 Sheep	
		1.1	1.4.3 Goats9	
		1.1.5	Rodent models of infection and disease10	
	1.2	PA	ATHOGEN MANIPULATION OF HOST MITOCHONDRIA15	
		1.2.1	Viral manipulations of mitochondrial function15	
		1.2.2	Bacterial manipulations of mitochondrial function16	
		1.2.3	Association of host mitochondria with T. gondii17	
	1.3	Gl	ENE DUPLICATION AND DIVERSIFICATION17	
		1.3.1	Gene duplication18	
		1.3.2	Functional divergence19	

2.0	THE IMPACT OF TGMAF1 DUPLICATION AND DIVERSIFICATION ON
HOST M	ITOCHONDRIAL ASSOCIATION IN TOXOPLASMA GONDII AND CLOSELY
RELATE	D SPECIES
2.1	INTRODUCTION22
2.2	RESULTS23
	2.2.1 MAF1 is uniquely expanded in <i>T. gondii</i> and exhibits inter- and intra-
	lineage copy-number variation23
	2.2.2 <i>MAF1</i> paralogs are uniquely divergent and under diversifying selection
	in <i>T. gondii</i>
	2.2.3 MAF1a and MAF1b gene families have distinct inter-strain
	transcriptional profiles
	2.2.4 MAF1a and MAF1b protein expression differs between <i>T. gondii</i> strains
	2.2.5 T. gondii MAF1 paralogs differ in their ability to mediate host
	mitochondrial association in <i>T. gondii</i> and <i>N. caninum</i> 40
	2.2.6 MAF1b1 from <i>T. gondii</i> and <i>H. hammondi</i> can confer the HMA
	phenotype in type II T. gondii, while TgMAF1b0 and HhMAF1a1 cannot44
	2.2.7 Expression of TgMAF1RHb1, but not TgMAF1RHa1, in type II T.
	gondii increases competitive advantage during infection47
2.3	CONCLUSIONS
3.0	UTILIZING NATURALLY OCCURING APIMAF1 ISOFORMS AND
MUTATI	ONAL ANALYSIS TO DETERMINE TGMAF1 RESIDUES NECESSARY FOR
HOST M	ITOCHONDRIAL ASSOCIATION IN T. GONDII

	3.1	INTRODUCTION		
	3.2	RE	ESULTS	
		3.2.1	Primary sequence alignment of MAF1 isoforms reveals regions of	
		dissimil	arity56	
		3.2.2	The proline-rich region of TgMAF1RHb1 is not necessary for HMA59	
		3.2.3	Structural analysis of the C-terminal domains of TgMAF1RHa1 and	
		TgMAF	<b>TRHb1</b> shows a high degree of similarity in three dimensions60	
		3.2.4	Conserved residues selected for mutational analysis in the C-terminal	
		domain	of TgMAF1RHb1 are not required for HMA62	
	3.3	CC	ONCLUSIONS	
4.0		THE E	FFECTS OF TGMAF1 EXPRESSION THROUGHOUT THE T. GONDII	
LIF	'E CY	CLE		
	4.1	INTRODUCTION		
	4.2	2 RESULTS		
		4.2.1	Exogenous TgMAF1RHb1 expression in type II parasites increases cyst	
		burden	of chronically infected CBA/J mice 4 and 8 weeks post infection69	
		4.2.2	Serum cytokine levels differ between mice infected with wildtype type II	
		parasite	es and type II parasites expressing TgMAF1RHb1 throughout infection.73	
		4.2.3	Screen of Candidate cytokines <i>in vivo</i> 78	
	4.3	CC	DNCLUSIONS	
5.0		CONCI	LUSIONS AND FUTURE DIRECTIONS83	
6.0		MATE	RIALS AND METHODS89	
	6.1	PA	RASITE STRAINS AND MAINTENANCE	

6.2	SEQUENCE COVERAGE AND COPY NUMBER ANALYSIS
6.3	HIGH MOLECULAR WEIGHT SOUTHERN BLOTTING92
6.4	AMPLIFICATION AND CLONING OF MAF1 PARALOGS FROM T.
GONDII	, H. HAMMONDI, AND N. CANINUM93
6.5	SEQUENCE ANALYSIS94
6.6	GENERATION OF CONSTRUCTS AND TRANSGENIC PARASITES95
6.7	TGMAF1RHA1 AND TGMAF1RHB1 CLONING, PROTEIN
PRODU	CTION, AND PURIFICATION95
6.8	IMMUNOFLUORESCENCE ASSAYS AND CONFOCAL MICROSCOPY
	96
6.9	QUANTIFICATION OF VACUOLE COVERAGE
6.10	WESTERN BLOT ANALYSIS97
6.11	IN VITRO GROWTH ASSAY98
6.12	IN VITRO INTERFERON-GAMMA INHIBITION ASSAY99
6.13	ASSAY FOR GROWTH IN PRESENCE OF INTERFERON-GAMMA99
6.14	H. HAMMONDI ROP16 PROMOTER ASSAY100
6.15	IN VITRO COMPETITION ASSAY100
6.16	ANIMAL EXPERIMENTS101
6.17	GENERATION OF POLYCLONAL ANTIBODIES101
6.18	SERUM COLLECTION
6.19	ANALYSIS OF CYTOKINE EXPRESSION
6.20	PARASITE INFECTION AND BIOLUMINESCENCE IMAGING103
6.21	IN VIVO COMPETITION ASSAY

6.22	CYST PURIFICAION	104
6.23	CYST STAINING AND QUANTIFICAION	
APPENDIX	A	
APPENDIX	В	
APPENDIX	С	116
APPENDIX	D	118
BIBLIOGRA	АРНҮ	151

# LIST OF TABLES

Table 1. A summary of the known host range and virulence properties of <i>T. gondii</i> and the
closely related species <i>H. hammondi</i> and <i>N. caninum</i>
Table 2. KaKs values for T. gondii "b" paralogs including b0
Table 3. Ks values for T. gondii "b" paralogs including b0
Table 4. Conserved MAF1 residues63
Table 5. Strains used throughout dissertation
Table 6. G-CSF119
Table 7. Eotaxin   120
Table 8. GM-CSF121
Table 9. IFNg   122
Table 10. IL-1a
Table 11. IL-1b
Table 12. IL-2
Table 13. IL-3
Table 14. IL-4
Table 15. IL-5
Table 16. IL-6

Table 17. IL-7	
Table 18. IL-9	
Table 19. IL-10	
Table 20. IL-12p40	
Table 21. IL-12p70	
Table 22. IL-13	
Table 23. IL-15	
Table 24. IL-17	
Table 25. IP-10	
Table 26. KC	
Table 27. LIF	
Table 28. LIX	141
Table 29. MCP-1	
Table 30. M-CSF	
Table 31. MIG	
Table 32. MIP-1a	
Table 33. MIP-1b	146
Table 34. MIP-2	147
Table 35. RANTES	
Table 36. TNFa	149
Table 37. VEGF	

# LIST OF FIGURES

Figure 1-1. Life Cycle of <i>Toxoplasma gondii</i>
Figure 1-2. Head-to-head comparison of luciferase-tagged N. caninum and T. gondii in mice13
Figure 2-1. The MAF1 locus exhibits copy number variation across strains of T. gondii and has
comparatively low copy number in <i>H. hammondi</i> and <i>N. caninum</i> 25
Figure 2-2. The MAF1 locus exhibits copy number variation between T. gondii strains within
and between strain types
Figure 2-3. Scal sites are conserved among T. gondii strains used in Southern blot analysis27
Figure 2-4. The T. gondii and H. hammondi MAF1 loci harbor two distinct isoforms while only
one isoform is present in <i>N. caninum</i>
Figure 2-5. Alignment of select MAF1 sequences from T. gondii, H. hammondi, and N. caninum.
Figure 2-6. Transcript and protein expression analyses of MAF1a and b paralogs
Figure 2-7. Higher molecular weight band is reduced upon increased boiling of sample
Figure 2-8. <i>T. gondii</i> MAF1 paralog expression differs between lineages
Figure 2-9. Transgenic expression of TgMAF1RHb1 in ME49 detected by paralog-specific
antibodies40

Figure 2-10. Host mitochondrial association is a feature of T. gondii and H. hammondi
infections, but not N. caninum
Figure 2-11. MAF1RHa1 and MAF1RHb1 differ in their ability to complement HMA in T.
gondii and N. caninum
Figure 2-12. T. gondii and H. hammondi harbor MAF1 isoforms that differ in their ability to
mediate HMA
Figure 2-13. Complementation with TgMAF1RHb1 has does not have a significant impact on
acute virulence in vivo during infection in Balb/c mice49
Figure 2-14. Expression of TgMAF1RHb1, but not TgMAF1RHa1, in type II T. gondii increases
competitive advantage
Figure 2-15. Expression of TgMAF1RHb1 in Type II T. gondii increases competitive advantage
<i>in vitro</i> 51
Figure 3-1. Conserved and divergent regions of MAF1
Figure 3-2. The proline-rich region of TgMAF1RHb1 is not necessary for HMA60
Figure 3-3. Structure of TgMAF1RHb1 and TgMAF1RHa1 c-terminal domains61
Figure 3-4. Conserved residues in the C-terminal domain are not necessary for HMA64
Figure 4-1. TgMAF1RHb1 expression does not significantly alter parasite burden during acute
infection in CBA/J mice
infection in CBA/J mice
infection in CBA/J mice.72Figure 4-2. TgMAF1RHb1 increases cyst burden in chronically infected CBA/J mice.72Figure 4-3. Serum cytokine levels during chronic infection measured by Luminex.77
infection in CBA/J mice
infection in CBA/J mice.72Figure 4-2. TgMAF1RHb1 increases cyst burden in chronically infected CBA/J mice.72Figure 4-3. Serum cytokine levels during chronic infection measured by Luminex.77Figure 4-4. Bioluminescence monitoring of acute infection in cytokine knockout mice.79Figure 6-1. T. gondii and N. caninum growth rates in vitro.109

Figure 6-3. <i>N. caninum</i> growth is inhibited in the presence of IFNγ111
Figure 6-4. Impact of transgenic TgROP18 <sub>II</sub> and TgROP5 expression on N. caninum growth
inhibition in the presence of IFNγ112
Figure 6-5. Deletion of 16 bp upstream of ROP16 transcriptional start site disrupts expression.
Figure 6-6. Localization of 203290117

### PREFACE

This is the part where I thank all the people who helped along the arduous journey that led me to the point where I am writing the preface for my dissertation. And there are a lot of people to thank. First and foremost I would like to thank my advisor Dr. Jon Boyle. I am struggling to find the words to adequately convey my gratitude for the wonderful mentorship I have received. Thank you for supporting me when I needed support, and for giving me a push when I needed to be pushed. Thank you for fostering a supportive yet challenging lab environment, filled with awesome people at every level. I would not have made it to the end without the support of those graduate students who came before me, Drs. Yaw Adomako-Ankomah and Gregory Wier, and those who come after me (the future Drs.) Rachel Coombs, Matthew Blank, and Sarah Sokol. Especially Greg, who continues to be a wonderful and supportive (if overly cynical) friend even though we don't share a bay anymore. I have had the pleasure (when I wrote that I accidentally wrote pressure, which is also fitting) of mentoring several undergrads, including Jeffrey Danielson and Sarah Post. Thank you both for tolerating my first ventures into mentorship; I hope I served you well.

I would also like to thank the wonderful collaborators who have shared and discussed data with me over the past few years, Drs. John Boothroyd and Felice Kelly at Stanford University and Drs. Martin Boulanger and Michelle Parker at the University of Victoria. Our interactions have emphasized the importance of collaboration and communication in science; a lesson I will carry with me as I progress in my career.

Thank you to my cohort buddies, especially Christine Cucinotta and Lynley Doonan. You two kept me (mostly) sane and I definitely wouldn't have made it to the end without both of you. We were once accused of walking everywhere holding hands, and that is a wonderful description of our friendship. Thank you for all hand holding and all the support.

Finally, I would like to thank my family. Thank you to my mother, Stacy Atchley, for encouraging my reading, art, and science habits, all of which have served me well throughout my education. Thank you to my father, Curtis English, for all the groan-worthy dad jokes and random texts to "get to work" late at night on Friday or Saturday throughout my entire undergraduate and graduate career. I would also like to thank my stepmother Pam Long, for her love and encouragement. I would like to thank my grandparents, especially the woman I was named after, my grandma Elizabeth Copelman. You have always encouraged me to question everything, even authority, and to pursue my wildest dreams. Thank you for helping to make me the person I am today. Thank you to my crazy siblings, especially my sister Stephanie for feeding me regularly. And thank you to my wonderful partner in crime, Dr. Derek Horkel, for encouragement when I am down, distraction when I'm anxiously awaiting grant or paper decisions, and for ignoring me when I try to distract myself from writing by talking to you. Thank you.

# **1.0 INTRODUCTION**

The phylum Apicomplexa encompasses a variety of obligate intracellular pathogens of importance to both veterinary and human health. Notable members of the phylum include veterinary pathogens *Eimeria tenella*, *Sarcocystis neurona*, and *Neospora caninum*, as well as human pathogens such as the causative agent of malaria (*Plasmodium spp.*), the diarrheal pathogen *Cryptosporidium spp.*, and the causative agent of Toxoplasmosis (*Toxoplasma gondii*).

## 1.1 TOXOPLASMA GONDII

# 1.1.1 Life Cycle

*T. gondii* has a life-cycle typical of tissue-dwelling coccidian parasites. Sexual reproduction occurs exclusively in the definitive host, members of the family *Felidae* [1], while asexual reproduction occurs in a variety of intermediate hosts. *T. gondii* is transmitted to new hosts primarily via ingestion of either environmentally stable oocysts shed in cat feces [2] or meat from chronically infected hosts [3]. Humans can acquire *T. gondii* infection by either of these routes [4], or congenitally if a woman is infected for the first time while pregnant and infection is transmitted to the developing fetus [5].

Once an intermediate host ingests an infective form of the parasite, *T. gondii* differentiates into a form of the parasite called the tachyzoite. This life stage causes acute infection, replicating quickly, damaging tissue, and stimulating a robust immune response from the host [6, 7]. After several days, due to signals and mechanisms that are still largely unknown, the tachyzoites differentiate into a slower growing form called a bradyzoite [8]. Bradyzoites reside inside tissue-dwelling cysts and are insensitive to all currently approved drug treatments, leaving hosts chronically infected for the rest of their lives. If a chronically infected host later becomes immunocompromised, due to AIDs or immunosuppressive therapy, these bradyzoites are able to differentiate back into tachyzoites. This reactivation of acute infection can have devastating results, including brain damage or death [9]. Unlike many of its close relatives, *T. gondii* is able to transmit from one intermediate host to another via carnivory, without undergoing the sexual half of the lifecycle [10].

*T. gondii* undergoes sexual reproduction following ingestion of either tissue cysts or oocysts by the definitive host, members of the *Felidae* family. The ability of oocysts to infect the definitive host is unique to *T. gondii* compared to closely related species [10]. During sexual development in the gut of the cat, *T. gondii* differentiates into both male and female gametes. These gametes fuse to form a zygote, which develops into a premature oocyst and is then shed in the feces of the infected cat. Once in the environment the oocysts begin to mature, forming two sporocysts contained within the oocyst wall. Each sporocyst contains four sporozoites, which are an infective form of the parasite. The maturation process takes 24-48 hours [11], and mature oocysts remain environmentally stable and infective for as long as 18 months, depending on environmental conditions [12]. The entire lifecycle, including aspects unique to *T. gondii* compared to close relatives, is illustrated in Figure 1-1.



Figure 1-1. Life Cycle of *Toxoplasma gondii* 

#### **1.1.2** Host range and pathogenesis

Strikingly, *T. gondii* is capable of infecting virtually any warm-blooded animal, from birds to humans [13], and this broad intermediate host range is not only unique in comparison to closely-related species like *Hammondia hammondi* and *N. caninum*, but also with respect to most eukaryotic parasites. An estimated one third of the world's human population is currently infected with *T. gondii*, and while healthy individuals are able to control infection, those with compromised immune systems are at risk for developing life-threatening symptoms [14-16]. In addition, some *T. gondii* strains have been found to cause severe, and even fatal, disease in immunocompetent adults [17-19]. Acute infection during pregnancy often results in fetal abnormalities including blindness, hearing loss, seizures, or severe cognitive disabilities [5, 20]. In some cases this results in fetal loss or infant death [21]. *T. gondii* is also the cause of fetal loss in a number of domestic animals including sheep, goats, and pigs [22-25].

#### **1.1.3** Comparisons to closely related species

Virulence in humans has never been observed for *H. hammondi*, which is the most closely related extant relative to *T. gondii* and shares the same definitive host [10, 26]. While it is also assumed that *H. hammondi* is incapable of infecting humans, it is worth noting that given the antigenic similarity between these species [27], and that the most commonly used serological test for *T. gondii* infection is based on immunoreactivity to *T. gondii* surface antigen 1 (p30; SAG1), it is certain that if *H. hammondi* is capable of infecting humans such an infection would be misidentified as a *T. gondii* infection. Development of a serum-based diagnostic test that could

distinguish *T. gondii* from *H. hammondi* would allow for a direct test of the infectivity of *H. hammondi* in humans.

While the host range of *T. gondii* and *N. caninum* have been extensively studied, less is known about the host range in *H. hammondi*. Most *H. hammondi* isolates have been obtained from infected cats, but a wide variety of animals have been experimentally infected with this parasite, including cats, mice, rats, hamsters, and monkeys [28, 29]. Importantly, however, birds appear to be refractory to *H. hammondi* infection [30]. Another important distinction between *T. gondii* and *H. hammondi* is the inability of *H. hammondi* to be transmitted (at least experimentally) from one intermediate host to the other, and from one definitive host to another [10]. In fact this is a key diagnostic feature to distinguish isolates of these parasites in the laboratory [10, 31].

*N. caninum* does not share the cat as a definitive host, but rather utilizes canines for sexual reproduction [32]. Dogs are also intermediate hosts, as *N. caninum* infection in dogs causes a variety of neurological symptoms including encephalitis and ascending paralysis, often resulting in death [33, 34]. The known intermediate host range of *N. caninum* is more restricted than that of *T. gondii*, and consists of dogs, cattle, water buffalo, sheep, goats, and horses [34-39]. With the exception of dogs and horses, all of these intermediate hosts are members of the Bovidae family. *N. caninum* causes abortion in cattle [40, 41], much like *T. gondii* infection in sheep or goats. Unlike *T. gondii*, and similar to *H. hammondi*, there is no evidence that *N. caninum* infects humans [42], however as with *H. hammondi* the antigenic similarity between *T. gondii* and *N. caninum* makes it difficult to rule out the possibility of *N. caninum* infections in humans [43].

These overlapping, yet distinct host ranges for *T. gondii*, *H. hammondi*, and *N. caninum* have been observed for quite some time (see Table 1), yet the genes (both parasite and host) responsible for these differences remain unknown. It should also be noted that while all the host range of these parasites overlap, the pathologies and host response in overlapping intermediate hosts are not always the same.

 Table 1. A summary of the known host range and virulence properties of *T. gondii* and the closely related species *H. hammondi* and *N. caninum*

	Toxoplasma gondii			H. hammondi	N. caninum
	Type I	Type II	Type III		
Definitive host		Felids		Felids	Canids
Intermediate host range		Mammals, birds		Rodents	Bovids, horses
Disease associated with infection in intermediate hosts?		Yes		No	Yes
Transmission dynamics between hosts		Definitive Host Intermediate Host		Definitive Host	Definitive Host
Infects humans?		Yes		Unlikely	Unlikely
Disease associated with human infection?		Yes		Unlikely	Unlikely
Virulence phenotype in mice	$LD_{100} = 1$	$LD_{50} = 10^2  10^4$	$LD_{50} = 10^5 - 10^6$	Avirulent	Avirulent

#### **1.1.4** Non-rodent experimental models

Since both *T. gondii* and *N. caninum* can cause spontaneous abortion in livestock, experimental infections of sheep, goats, and cattle have been used to understand the pathology and modes of transmission of these parasites. Rodent models have also been developed to study infection, dissemination and transmission of *T. gondii*, *H. hammondi* and *N. caninum* and will be discussed in the next section.

**1.1.4.1 Cattle** The prevalence of *T. gondii* versus *N. caninum* infection in cattle is variable by region or herd. In some areas, such as Southern Vietnam and Western Thailand, T. gondii is more prevalent than N. caninum [44, 45], however in Southern China the prevalence of N. *caninum* infection is slightly higher than that of T. gondii [46]. Natural infection of cattle by T. gondii does occur [47, 48], but is not associated with abortion [49]. Experimental infection suggests that there is a low rate of abortion in cattle upon T. gondii infection [50], and that this rate increases with T. gondii strains that are typically more virulent in mice [51]. Surveys of aborted calves show a strong association with N. caninum infection, but no association with T. gondii infection [52], and herds with high abortion rates tend to have a high rate of N. caninum infection [53]. Experimental infection of pregnant cattle shows that N. caninum infection during early gestation is likely to cause abortion [54], however virulence among isolates does vary, and less virulent isolates do not appear to cause abortion [55]. While some earlier studies suggested that infection late during pregnancy facilitates vertical transmission, but does not cause abortion [56], more recent studies show that infection with N. caninum late in gestation can cause abortion [57]. There have also been conflicting studies suggesting that horizontal transfer of N. caninum infection and abortion does not occur in subsequent pregnancies after initial infection

[58], while other studies suggest that chronic infection can lead to recurrent abortions [59]. Clearly more experimental work is needed to clarify these conflicting data, and to take into account both the genetics of the parasite and the host. Experimental infections of cattle reveal that *N. caninum* disseminates to a variety of tissues including the heart, lung, kidney, skeletal muscle, and perhaps most importantly the brain [43]. In fact, in one study they detected *N. caninum* in the brain and spinal cord, but in no other surveyed location, including the gastrointestinal tract, liver, kidney, heart, lung, and skeletal muscle [60].

The genetic differences between *T. gondii* and *N. caninum* responsible for the differences in virulence in a bovine host have not yet been identified. It is interesting to note that *N. caninum* appears more virulent in cattle than *T. gondii*, whereas in most other shared intermediate hosts it appears that *T. gondii* is more virulent than *N. caninum*. This could be due to the fact that cattle are the natural, and most common, intermediate host for *N. caninum* and it has evolved specialized methods for evading the bovine immune response. Further studies are required in order to determine why *N. caninum* is so successful in the bovine host, whereas *T. gondii* is not.

**1.1.4.2 Sheep** Surveys of the prevalence of *N. caninum* and *T. gondii* in sheep herds show a significantly higher proportion of sheep infected with *T. gondii* than *N. caninum* [61-64]. A combination of serological studies and experimental infections demonstrate the ability of both *T. gondii* and *N. caninum* to cause abortion in sheep, particularly when infected during early pregnancy [65-69]. There is also experimental evidence that both *T. gondii* and *N. caninum* cause recurrent abortions in chronically infected ewes [70, 71]. Histological studies of aborted fetuses, weak lambs, congenitally infected healthy lambs, and experimentally infected ewes show that *T. gondii* and *N. caninum* dissemination patterns are quite similar [6, 65-67, 72]. Aborted *T. gondii*-infected fetuses have lesions primarily in the brain, with some specificity for regions such as the

optic tract and rostral margin of the pons [72]. Experimental infection in male sheep (rams) has shown that *T. gondii* does infect the male reproductive organs [73], and *T. gondii* infection can be sexually transmitted from infected rams to uninfected ewes [74, 75], but this has not been examined with *N. caninum* infection.

Taken together this suggests that although seroprevelence of *T. gondii* is higher than that of *N. caninum* in domestic sheep, both species are successful parasites of sheep and clearly cause similar pathology. This is in stark contrast to experimental infection in cattle, (described above), where *N. caninum* is clearly much more virulent than *T. gondii*.

**1.1.4.3 Goats** The seroprevalence of *T. gondii* infection in goats is generally much higher than that of *N. caninum* [76-78]. Surveys of aborted goat fetuses suggest that *T. gondii* infection contributes to a number of these abortions [79, 80], and experimental infection confirms that *T. gondii* is capable of causing abortions in goats [81, 82]. There are very few studies examining *N. caninu*m infection in goats, compared to the number of studies done in sheep and cattle. Experimental infection of pygmy goats during pregnancy suggests that *N. caninum* infection in goats does cause abortion when infection occurs early during gestation, and that abortion in these goats does not recur with subsequent pregnancies [83]. As with infection in sheep, *T. gondii* does infect the male reproductive organs [84], and infection can be sexually transmitted [85].

As with cattle and sheep, the genetics underlying the pathology differences between *N*. *caninum* and *T. gondii* in goats are not known. Additional studies in each of these intermediate hosts with genetically engineered parasites may help to uncover the genes responsible for both similarities, and differences in host range and host response in *T. gondii* and *N. caninum*.

#### **1.1.5** Rodent models of infection and disease

Several rodent models have been developed for studying *T. gondii*, *N. caninum*, and *H. hammondi* infections. These are particularly relevant as rodents are a natural intermediate host for both *T. gondii* and *H. hammondi*, and likely play an important role in the evolutionary history of these parasites.

Multiple mouse strains have been utilized in developing models of T. gondii infection, including both outbred (CD-1) and inbred (Balb/c; CBA/J, C57BL6) mouse strains [86, 87], and most recently, the house mouse [88]. Pregnant mouse models have also been developed to better understand why T. gondii infection causes abortion [7, 89-91]. Mice infected with T. gondii have enlarged spleens and lymph nodes, caused in part by an increase in mononuclear phagocytes and CD8+ T-cells, which produce Interferon- $\gamma$  (IFN- $\gamma$ ) [92]. Production of innate immune effectors such as Interleukin-12 (IL-12) and IFN-y increases shortly after infection, and is required for host survival and control of parasite growth [91, 93, 94]. Immune-compromised mice (lacking both B cells and T cells) that have a larger population of Natural Killer cells, and are therefore able to produce higher levels of IFN- $\gamma$ , have a lower parasite burden [89], once again providing evidence that IFN-y production is essential for mouse survival following T. gondii infection. Neutralization of IFN- $\gamma$  increases parasite burden in these mice, but decreases transmission of T. gondii infection to offspring [89]. Symptoms of acute infection by T. gondii generally decrease after several weeks, when the adaptive immune system has had time to respond and produce antibodies and effector cells to combat T. gondii [7].

The population structure of *T. gondii* isolates has been studied extensively in an effort to better understand parasite virulence and host interaction. The majority of North American and European *T. gondii* isolates can be grouped into three main linages that vary in virulence, as well

as host responses [95, 96]. *T. gondii* strains exhibit a broad range of virulence in mice, with the most virulent type I stains being capable of killing a mouse after infection with a single parasite [96, 97]. Less virulent type II and type III strains of *T. gondii* have 50% lethal doses of greater than  $10^3$  and  $10^5$  parasites, respectively [97]. Some "atypical" *T. gondii* strains from South America, which do not belong to any of the three major linages, have also been shown to be highly virulent in mice [98]. Comparisons of these strains and differences in host-response following infection have facilitated the discovery of many parasite factors responsible for virulence and/or interaction with the host, including rhoptry proteins 5, 16 and 18 (ROPs) [96, 99-105] and dense granule proteins 15, 24, 25 and MAF1 [106-109].

In general, immunocompetent mice experimentally infected by intraperitoneal injection of *N. caninum* tachyzoites exhibit no signs of disease; however, immunosuppression of mice using methylprenicolone acetate (MPA) results in a range of neurological symptoms, from a slight head tilt to paralysis and death, depending on the dose of immunosuppressant [110]. It also appears that subcutaneous injection of tachyzoites in inbred Balb/c mice results in a number of neurological symptoms without the use of MPA immunosuppression [111]. IFN- $\gamma$ -deficient mice, as well as mice lacking Toll-like receptor 4 (TLR-4) and a functional IL-12 receptor, are also susceptible to *N. caninum* infection by intraperitoneal injection of tachyzoites [112-115]. In these mice, parasites can be found in the pancreas, liver, lung, intestine, heart, and brain, while parasites are not detectable in these organs in immunocompetent mice [115]. Infection of dendritic cells is likely important for the dissemination of *N. caninum* within the host, as adoptive transfer of *N. caninum*-infected dendritic cells increases parasite load as well as vertical transmission in pregnant mice [116]. Since head-to-head comparisons between *T. gondii* and *N. caninum* have not been conducted in mice, we tagged *N. caninum* strain NC-1 [34] with luciferase and compared its proliferation *in vivo* to a highly avirulent strain of *T. gondii*, S1T. S1T is an F1 progeny clone derived from a cross between a *T. gondii* type II and type III strain, and contains avirulent alleles of all five identified *T. gondii* virulence factors [117, 118]. Mice eventually control parasite proliferation and are able to survive infection with up to  $1 \times 10^6$  tachyzoites of this parasite clone. Both species proliferate at a similar rate during the first 20 hours post-infection, but then *N. caninum* is rapidly controlled while *T. gondii* S1T continues to proliferate (Figure 1-2). This suggests that the inability of *N. caninum* to be virulent in wild type mice does not have to do with an inability to replicate within mouse cells *in vivo*, but rather an inability to disrupt host innate immune defenses that rapidly control this parasite. It will be interesting in future studies to compare host responses to these two species during the early stages of infection.

Given this attenuated phenotype in mice, several genetically altered mouse models have been developed for *N. caninum* infection with tachyzoites (described above). However even IFN- $\gamma$  knockout mice are not susceptible to oral infection by *N. caninum* oocysts [119]. Because of this, a gerbil (*Meriones unguiculatus*) model of infection has been developed, as gerbils are highly susceptible to oral infection with oocysts by *N. caninum* [120].

Similar to *T. gondii*, virulence differs among *N. caninum* isolates, which has been revealed by a number of comparisons [111, 121, 122]. The NC-Liverpool strain, isolated from the brain tissue of a young dog euthanized after presenting with severe neurological symptoms, is a more pathogenic strain than the NC-SweB1 strain, isolated from a stillborn calf [121]. NC-Nowra, isolated from a congenitally infected calf, is also less pathogenic than LC-Liverpool, but does cause some disease in a small portion of infected mice [123]. NC-1 and NC-3 were both isolated from the tissues of congenitally infected dogs, and NC-1 is much more pathogenic than

NC-3 [111]. No studies have been done to compare the pathogenicity of all *N. caninum* isolates, however these studies suggest a wide range in ability to cause neurological disease when injected subcutaneously in Balb/c mice. Genetic crosses in the definitive canine host between these strains with distinct phenotypes could potentially lead to the identification of the virulence factors responsible. However it is not known if they would be relevant to natural *N. caninum* infections since, in contrast to *T. gondii*, rodents do not appear to be a relevant host for *N. caninum* in the wild.



Figure 1-2. Head-to-head comparison of luciferase-tagged N. caninum and T. gondii in mice.

Balb/c mice (3 per strain) were intraperitoneally infected with  $1 \times 10^{6}$  tachyzoites of luciferase-tagged *N. caninum* (strain NC-1) or *T. gondii* (strain S1T) and *in vivo* bioluminescence imaging was used to quantify parasite burden over the first 96 hours of infection. (A) Average total flux (photons/s) indicating parasite burden over the course of infection. All mice survived the infection. (B) Representative images of infections quantified in (A), *N. caninum* is cleared within the first 44 hours of infection.

Relatively little work has been carried out in *H. hammondi*-infected mice, as there is currently no way to grow H. hammondi parasites in cell culture in order to perform the same types of experiments that have been done with T. gondii and N. caninum. Much of the work has been carried out in IFN- $\gamma$  knockout mice. In parenteral infections in both wild type and IFN- $\gamma$ knockout mice *H. hammondi* is benign, resulting in chronically-infected mice that show almost no symptoms of infection (based on behavioral responses to hyperinflammation or adverse neurological symptoms). However oral infections with large numbers of H. hammondi oocysts can cause severe disease [10] and even mortality in Swiss-Webster mice. It is important to note that IFN-y KO mice that are chronically infected with *H. hammondi* are infective to the definitive host, and rodents have been found to harbor H. hammondi in the wild [10]. Given that IFN- $\gamma$  is required for control of both T. gondii and N. caninum, it is intriguing that this cytokine is not required for control of *H. hammondi*. This could be due to as yet unidentified host innate immune responses, or it could be due to a hard-wired developmental program in H. hammondi that results in the spontaneous conversion from rapidly growing tachyzoites to slow-growing, encysted bradyzoites. Consistent with this latter explanation, H. hammondi-infected mice have orally infective tissue cysts in muscle and other non-CNS tissues in both wild type and IFN- $\gamma$  KO mice, and multiple groups have observed the spontaneous conversion of H. hammondi tachyzoites to infectious cysts during cultivation in vitro [10, 124]. Further analyses will be necessary to more fully characterize the differences in parasite development between H. hammondi and particularly T. gondii. Regardless of the root cause, overall the existing work on H. hammondi indicates that it is unique compared to both T. gondii and N. caninum in terms of its behavior in immune-deficient mice.

#### 1.2 PATHOGEN MANIPULATION OF HOST MITOCHONDRIA

Mitochondria are traditionally considered the powerhouses of the cell, due largely to the fact that they are a major site of ATP synthesis in many eukaryotes. However, it is becoming increasingly clear that mitochondria are also responsible for a wide variety of other cellular functions, including calcium homeostasis, regulation of apoptosis, and immune signaling [125-127]. Not surprisingly, a number of intracellular pathogens, including *T. gondii*, have been found to manipulate mitochondria localization and function.

# **1.2.1** Viral manipulations of mitochondrial function

There is a growing body of literature exploring viral manipulation of mitochondria and their many functions. Viruses have been found to manipulate various mitochondrial functions from calcium ( $Ca^{2+}$ ) homeostasis to antiviral signaling. For example herpes simplex virus 1 (HSV-1) reduces mitochondrial uptake of  $Ca^{2+}$  [128]. Several hepatitis C virus (HCV) proteins also alter  $Ca^{2+}$  homeostasis in different ways at different stages of HCV infection. The core protein increases mitochondrial uptake of  $Ca^{2+}$ , which in turn increases the production of  $Ca^{2+}$ -dependent reactive oxygen species (ROS) [129]. The HCV p7 protein localizes to the mitochondria and increases  $Ca^{2+}$  flux from the mitochondria to the cytoplasm [130].

Viruses can either inhibit or promote mitochondria-mediated apoptotic signaling, often through the use of virally encoded B cell lymphoma-2 (Bcl-2) family proteins. Epstein-Barr virus (EBV) encodes at least two Bcl-2 homologs, both of which suppress apoptosis by associating with pro-apoptotic factors in order to inhibit apoptotic signaling [131, 132]. Hepatitis B virus (HBV) encodes protein X, a pro-apoptotic Bcl-2 homolog that localizes to the mitochondria, alters mitochondrial membrane potential, and induces apoptosis [133, 134].

In addition to apoptotic signaling, viruses can also inhibit mitochondrial antiviral signaling (MAVS). Both HCV and GB virus B (GBV-B) encode NS3/4A proteins that cleave MAVS proteins from the outer mitochondrial membrane thus inhibiting the ability to induce interferon signaling [135-137]. Alterations of MAVS, as well as apoptotic signaling and other mitochondrial functions, allow viruses to increase host cell survival allowing the virus more time to replicate and increasing virus survival.

### **1.2.2** Bacterial manipulations of mitochondrial function

Intracellular bacteria also alter host mitochondrial localization, structure, and function. The causative agent of Legionnaires' disease, *Legionella pneumophila*, is a rod shaped, aerobic, intracellular bacterium that replicates inside a *Legionella*-containing vacuole (LCV), and recruits host ER-derived vesicles and mitochondria to the LCV [138]. It has also been found that *L. pneumophila* infection induces apoptosis via the mitochondrial death signaling pathway, which can be blocked by overexpression of the anti-apoptotic Bcl-2 [139]. Several proteins secreted by *L. pneumophila* via the Dot/Icm secretion system have been shown to activate caspase 3, resulting in apoptosis via the mitochondrial death signaling pathway [140]. However, whether any of these proteins are responsible for the recruitment of host mitochondria, or if recruitment of host mitochondria and apoptosis are connected remains to be determined.

Several species of *Chlamydia* also associate with host mitochondria [141, 142]. Interestingly, *Chlamydia trachomatis* and *Chlamydia pneumoniae*, both of which do not associate with host mitochondria [141], inhibit apoptosis of infected cells at a step within the mitochondrial death signaling pathway [143]. This suggests that association with host mitochondria is not necessary for manipulation of host mitochondrial function during *Chlamydia* infection. The biological impact of association with host mitochondria remains unknown for *Chlamydia* species that do associate with mitochondria, in part because the factor responsible for this association remains unknown.

### 1.2.3 Association of host mitochondria with T. gondii

Following invasion of a host cell, the parasitophorous vacuole membrane (PVM) surrounding intracellular *T. gondii* associates with the host mitochondria. While host mitochondrial association (HMA) with the PVM of intracellular *T. gondii* has been observed for decades, the parasite gene responsible for this phenotype has only recently been discovered. The identification of the parasite protein mitochondrial association factor 1 (MAF1; [106]) provides the opportunity to use comparative analyses of *T. gondii* strains differing only in their expression of MAF1, and therefore HMA, in order to determine the functional impact of HMA. Initial studies suggest that MAF1-mediated HMA may alter the host innate immune response [106], however how HMA evolved, the molecular mechanism of HMA, and the *in vivo* impact of HMA on *T. gondii* biology remain largely unknown.

### **1.3 GENE DUPLICATION AND DIVERSIFICATION**

How species evolve new and diverse traits is a long-standing question in the world of evolutionary biology. One way by which new traits arise within a population is gene duplication followed by functional divergence. Gene duplication can occur by several different mechanisms, resulting in two or more copies of the same gene, which are then considered paralogs of one another. The outcome of these duplication events depends on the evolutionary pressures on the population, as well as the cost or benefit of having an extra copy of the original gene.

### **1.3.1** Gene duplication

There are several molecular mechanisms by which a single copy of one or more genes can become duplicated within a genome. Polyploidization, or the segregation of two or more genomes into a single nucleus, results in whole genome duplication (WGD) [144]. WGD has occurred in the evolutionary history of a large variety of organisms from yeast [145] to vertebrates [146], although it is most commonly studied in flowering plants due their high incidence of polyploidy [147]. While large-scale duplication events, such as WGD, provide a large amount of genetic material for subsequent diversification or adaptation, gene duplication also occurs on smaller scales.

Smaller scale gene duplication events include retroposition and unequal crossing over. Retroposition occurs when RNA is reverse transcribed into DNA and then inserted into the genome [148]. Genes duplicated by retroposition can be found throughout a genome, and are marked by a lack of introns and remnants of flanking direct repeats. Unequal crossing over during DNA replication can result in the duplication of one or more genes in tandem [149, 150]. This can occur between homologous chromosomes during meiosis, when sections of the maternal and paternal chromosomes are exchanged unequally. This results in daughter cells with copy number variation (CNV) of one or more genes depending on the length of DNA involved in the unequal crossover event. Duplicated genes have been identified in *T. gondii*, and in fact the most potent virulence factor during infections in mice is coded for by the tandemly duplicated rhoptry protein 5 (*ROP5*) gene [102]. Perhaps even more interestingly, the subset of genes that are duplicated in *T. gondii* are distinctly different from the subset of genes duplicated in *N. caninum* and *H. hammondi* [151], suggesting that duplicated genes may account for phenotypic differences between these species.

#### **1.3.2** Functional divergence

Following duplication events, there are a number of possible outcomes for the paralogs. It is likely that the majority of gene duplication events are lost within populations, unless by some selective advantage, or perhaps random chance, the duplicated gene rises in frequency within the population. If having two or more copies of the same gene provides sufficient selective advantage due to increased expression of that gene, the paralogs can be maintained within the genome without acquiring new functions [148, 152]. Alternatively, a number of proposed outcomes following gene duplication fall within the category of functional divergence, where paralogs accumulate mutations rendering them functionally distinct from one another.

There are three proposed categories of functional divergence following a gene duplication event: nonfunctionalization, subfunctionalization, and neofunctionalization. Nonfunctionalization occurs when a paralog accumulates mutations rendering the gene nonfunctional, and is the most common outcome of gene duplication events [153]. Gene duplicates can also undergo subfunctionalization, where the two or more paralogs partition the functions of the ancestral single copy gene. This is often achieved by complementary degenerative mutations, where one copy loses function A and the other copy loses function B, so

19

that the organism must now retain both copies in order to be viable [154]. For example, the Zebrafish (*Danio rerio*) engrailed-1 paralogs, *eng1* and *eng1b*, are expressed in different tissues during embryonic development; *eng1* is expressed in the pectoral appendage bud, and *eng1b* is expressed in specific neurons in the hindbrain and spinal cord [154]. Vertebrates with only one copy of *eng1*, including the mouse (*Mus musculus*) and chicken (*Gallus gallus*), express *eng1* in both the pectoral appendage bud and specific neurons in the hindbrain and spinal cord [154]. These data suggest that following duplication of an ancestral *eng1*, *D. rerio eng1* lost neuronal expression while retaining pectoral limb bud expression, and *D. rerio eng1b* lost pectoral limb bud expression while maintaining neuronal expression. This supports the hypothesis that gene duplicates can be maintained within a genome by edegenerative mutation [154].

Neofunctionalization occurs when a paralog obtains a new function distinct from any function associated with the ancestral copy of the gene. There are an increasing number of examples of neofunctionization in the literature. A recently discovered example of neofuncitonalization is the avoidance of freezing in Antarctic notothenioid fish eggs. The eggs of Antarctic notothenioid fishes avoid freezing through the expression of an expanded repertoire of zona pellucida proteins (ZPs), which form a protective matrix around the egg [155]. Antarctic and sub-Antarctic species of notothenioid fishes have more copies of several families of ZPs compared to temperate species, which are able to lower the melting point of ice crystals [155]. This suggests that duplicated ZPs acquired the ability to protect eggs from freezing.

We have recently found evidence that the HMA phenotype observed in *T. gondii* is a result of neofunctionalization of a duplicated gene. Chapter 2 will discuss the evidence for neofunctionalization of the *MAF1* gene in *T. gondii*, as well as the selective advantage that led to the fixation of this gene in parasite populations. In Chapter 3, I will present mutational analysis
of the MAF1 protein using comparative functional and crystallographic data to select residues that may be important for association with host mitochondria. Chapter 4 will discuss the impact of MAF expression during the chronic stages of *T. gondii* infection, including alterations in the host immune response and the number of tissue cysts in the brains of chronically infected mice. Together, this dissertation provides a clear example of neofunctionalization in Apicomplexan parasites, and fills a gap in understanding why *T. gondii* associates with host mitochondria.

## 2.0 THE IMPACT OF TGMAF1 DUPLICATION AND DIVERSIFICATION ON HOST MITOCHONDRIAL ASSOCIATION IN *TOXOPLASMA GONDII* AND CLOSELY RELATED SPECIES

## 2.1 INTRODUCTION

Gene duplication and subsequent divergence is one method by which new phenotypes evolve, while still maintaining ancestral phenotypes. This is a process known as neofunctionalization, whereby a gene duplicate accumulates mutations, ultimately acquiring a new function. Here we present evidence for the occurrence of neofunctionalization in the ancestor of *Toxoplasma gondii* and *Hammondia hammondi*, resulting in the host mitochondrial association (HMA) phenotype.

The parasite protein responsible for HMA with the parasitophorous vacuole (PV) membrane in *T. gondii* has recently been identified and has been dubbed mitochondrial association factor 1 (MAF1; [106]). In this study we show that the *MAF1* gene is duplicated, and has undergone diversification in *T. gondii* and *H. hammondi* compared to another close relative *Neospora caninum*. Diverse copies of *MAF1* are differentially expressed across *T. gondii* strains, and the expression of isoforms containing a proline/serine motif correlates with the HMA phenotype. Functional studies of diverse *MAF1* isoforms cloned from *T. gondii* and *H. hammondi* reveal differences in the ability of isoforms to mediate HMA, and both *T. gondii* and *H. hammondi* harbor copies of *MAF1* able to mediate HMA. Additionally, expression of an

HMA-competent copy of *MAF1* in a *T. gondii* strain that does not normally exhibit HMA provides a competitive advantage within a mixed population compared to the wildtype HMA negative parasites *in vivo*. Taken together, these data provide evidence for neofunctionalization of an ancestral *MAF1* duplicate in the common ancestor of *T. gondii* and *H. hammondi* resulting the in the HMA phenotype, and provides evidence for the selective pressure that lead to the fixation of the new *MAF1* gene in *T. gondii* populations.

### 2.2 **RESULTS**

## 2.2.1 MAF1 is uniquely expanded in *T. gondii* and exhibits inter- and intra-lineage copynumber variation

We previously reported that MAF1 is a multicopy locus in *T. gondii*, and based on sequence read coverage exhibits strain-specific copy-number variation between representatives of the canonical *T. gondii* lineages (Types I, II and III: GT1, ME49 and VEG; [106, 151]). We have extended these copy number analyses to 5 additional *T. gondii* clonotypes outside of the 3 major lineages and also find that MAF1 is similarly expanded in these strains (Figure 2-1). Similar to Types I, II and III, there is significant copy number variation between strains at this locus, ranging from an estimated 8-10 copies for MAS to 4-6 copies for P89, FOU, VAND and RUB (Figure 2-1). While these data provide only an estimate of copy number differences between strains, they do confirm that the multi-copy state of the *MAF1* locus is conserved across highly diverse *T. gondii* isolates.

To further confirm differential expansion of the *MAF1* locus in *T. gondii* and to identify differences in MAF1 copy number between them, we performed high molecular weight Southern blot analysis of the *MAF1* locus in 6 *T. gondii* strains. These strains comprised 2 each from the Type I (GT1, RH), Type II (ME49, PRU) and Type III (VEG, CTG) lineages. Genomic DNA for each strain was digested with ScaI, which cuts on either side of the entire locus but not within, allowing for locus size (and therefore copy number) to be estimated (Figure 2-2; [102, 151]). Sequence coverage analysis shows higher copy number for GT1 compared to ME49 and VEG (Figure 2-1), and the Southern blot was consistent with this observation: GT1 has the largest *MAF1* locus (~44.9 Kb), while the *MAF1* loci in ME49 and VEG were smaller (28.4 Kb; Figure 2-2B). No other bands were visible on the blot (which resolved fragments ranging in size from 4.9 Kb to 53.9 Kb), indicating that the entire locus was intact for all strains. Moreover, the ScaI sites flanking the expanded locus were conserved in all 6 strains tested (Figure 2-3), indicating these differences are not due to mutations within the flanking sequences.

Based on the size of the *MAF1* repeat unit (3612 bp) and the known size of the regions between the locus and the ScaI sites (see Figure 2-2A), we *estimate* that there are 6 copies of *MAF1* in GT1, 4 in RH, ME49, PRU and VEG, and 2 in CTG. *MAF1* exhibits copy number variation within members of the *same* clonal lineage (i.e., GT1 vs. RH and VEG vs. CTG), suggesting that expanded loci change more rapidly in size and copy number compared to the single nucleotide polymorphism rate at single-copy loci, similar to what we have observed previously at other expanded loci in *T. gondii* [151].

We also performed copy number analysis of the *MAF1* loci in both *H. hammondi* and *N. caninum*, two relatives of *T. gondii* with distinct virulence and host range phenotypes [10, 156]. For *H. hammondi*, copy number analysis suggested the presence of 2 copies of *MAF1*, which is consistent with the presence of two predicted *MAF1* paralogs in the *H. hammondi* genome (HHA\_220950 and HHA\_279100). For *N. caninum*, we predicted the existence of 1 to 2 copies of *MAF1* (Liverpool strain; Figure 2-1; <u>www.toxodb.org</u>). In version 10.0 of the *N. caninum* genome there is only a single predicted *MAF1* ortholog (NCLIV\_004730).



Figure 2-1. The *MAF1* locus exhibits copy number variation across strains of *T. gondii* and has comparatively low copy number in *H. hammondi* and *N. caninum*.

Coverage depth analysis for the *MAF1* locus in 8 *T. gondii* strain types and for the syntenic locus in *H. hammondi* and *N. caninum. T. gondii* sequences are from ToxoDB v7.3. Portions of the upper left panel of this figure were similarly represented in Pernas et al, 2014. Raw reads were plotted as described in Materials and Methods, and normalized to the coverage 20 Kb upstream of the repetitive locus. Arrowheads indicate the location of predicted

gene sequences based on ToxoDB (v7.3 for *T. gondii*; v26 for all other species). Asterisks indicate smaller repetitive sequence unrelated to *MAF1* (see Chapter 6: Materials and Methods for further explanation).



Figure 2-2. The MAF1 locus exhibits copy number variation between T. gondii strains within and between strain types.

(A) Schematic representation of the *MAF1* locus showing ScaI restrictions sites outside of the locus, the size of the regions flanking the *MAF1* locus, and the size of the repeat unit used to estimate copy number based on Southern blotting. The most relevant *T. gondii* ME49 gene name is indicated (from ToxoDB v7.3), although it does not fully match the sequenced paralogs. (B) ScaI-digested gDNA from each of 6 *T. gondii* strains was resolved by PFGE and probed with a MAF1-specific probe. The blot shows copy-number variation consistent with predictions from sequence coverage analysis for strain types GT1, ME49 and VEG. Copy-number for each strain was determined based on the schematic presented in (A).



Figure 2-3. ScaI sites are conserved among T. gondii strains used in Southern blot analysis.

A) Schematic representation of the *MAF1* locus showing ScaI restrictions sites used to determine locus size. a, b, c and d represent primers used to verify location of the ScaI restriction sites. The most relevant *T. gondii* ME49 gene name is indicated (from ToxoDB v7.3), although it does not fully match the sequenced paralogs. B) A PCR-based diagnostic digest was performed to confirm that the ScaI sites were present at the same predicted locations in all strains queried.

### 2.2.2 MAF1 paralogs are uniquely divergent and under diversifying selection in T. gondii

To further characterize the MAF1 locus across strains and species, we sequenced 6 PCR-derived MAF1 clones from each of 3 representative strains from the Type I (RH), II (ME49) and III (CTG) T. gondii lineages, 13 clones from H. hammondi HhCatGer041, and 10 clones from N. caninum strain NC-1 [157]. We found that in T. gondii the MAF1 locus harbors multiple diverse paralogs of the MAF1 gene, indicating that the locus has both amplified and diversified. Importantly, all cloned T. gondii MAF1 paralogs were distinct from those found in existing annotation datasets for the T. gondii genome, including the putative MAF1 paralogs TGME49\_020950 and TGME49\_220950 (Figure 2-4B). Therefore we have resorted to a new nomenclature that can be found in Figure 2-4B. In addition to a putative signal peptide, each paralog is predicted to encode a single transmembrane domain located in the N-terminal region (Figure 2-4A). The most significant distinguishing feature across the sequenced MAF1 paralogs is the presence or absence of a repetitive stretch of 4-7 prolines followed by a serine ( $P{4:7}S$ ), as well as the amino acids surrounding the proline motif (~20 N-terminal to the motif and ~10 Cterminal to the motif). Importantly, the MAF1 paralog that was shown previously to complement the host mitochondrial association phenotype in Type II T. gondii (TgMAF1RHb1) also harbors this P{4:7}S motif [106]. This motif is either completely missing or repeated up to 6 times depending on the paralog (Figures 2-4B and 2-5).

For RH, 5 of the 6 sequenced clones contain the P $\{4:7\}$ S motif, while all 6 CTG clones, which represent only 2 unique coding sequences, have some form of the repeat motif. Interestingly, of the 6 clones sequenced from ME49, 3 are pseudogenes with premature stop codons, and all 3 of these clones are predicted to encode *MAF1* paralogs with the P $\{4:7\}$ S motif. Of the remaining 3 clones, 2 harbored the P $\{4:7\}$ S motif while the other did not. Based on amino

acid identity of 15 non-pseudogenized genes from RH, ME49 and CTG we identified a total of 8 unique coding sequences (4 for RH, 2 for ME49 and 2 for CTG). However there is also significant variation across these sequences outside of the P{4:7}S motif. We calculated pairwise dN/dS ratios for all unique *T. gondii* paralogs, and find significantly higher dN/dS ratios in TgMAF1RHb3 when compared to other RH paralogs as well as those from ME49 and CTG (P<0.05; Figure 2-4C). While not significant (P>0.05), TgMAF1ME49b1 also shows a higher dN/dS ratio when compared to TgMAF1RHb1 and b2 (Figure 2-4C). Because the dN/dS ratio is strongly influenced by the value of dS [158], we also report the Ka/Ks values (Table 2) and the individual Ks values (Table 3) for these comparisons. Ka/Ks values were calculated pairwise using Kimuras two parameter model [159], and are similar to the dN/dS values as expected. Ks values for all comparisons ranged from 0.003 to 0.010.

When we sequenced 13 distinct clones for the *H. hammondi MAF1* locus, we identified only 2 distinct sequences. One contained a stretch of 23 prolines in the same region as the P{4:7}S motif in *T. gondii MAF1* paralogs (Figures 2-4B and 2-5; as found in HHA\_220950), and the other lacked this proline-rich region (Figures 2-4B and 2-5; similar to HHA\_279100). This suggests that, consistent with the sequence coverage analysis and genome annotations, the *MAF1* locus harbors only 2 paralogs in *H. hammondi*, but that these paralogs also differ in the presence or absence of a proline-rich motif. Finally based on the current genome assembly and direct sequencing of 10 clones, the sole *N. caninum MAF1* paralog (NCLIV\_004730) is more similar to the *T. gondii* and *H. hammondi MAF1* paralogs that do not have a proline-rich region (Figure 2-4B). A maximum likelihood tree of amino acid sequences for all unique *MAF1* paralogs from *T. gondii*, *H. hammondi* and *N. caninum* is shown in Figure 2-4B and illustrates these relationships.

Given this diversity of sequences both within and between species we have named the identified MAF1 paralogs and deposited them in Genbank. As shown in Figure 2-4B, based on sequence similarity MAF1 has two major groups, and we have dubbed these "a" and "b", and all of the "a" paralogs lack the P{4:7}S motif. For the "b" paralogs we identified 2 sequences without the P{4:7}S motif and have named these "b0", and then named all other MAF1 paralogs with the  $P{4:7}S$  motif as b1, b2, etc. (Figure 2-4B). We feel this nomenclature accurately reflects the relationships between the various sequences in terms of broad groupings as well as the presence or absence of the  $P{4:7}S$  motif. From this point onward when we use "MAF1" without further indication of paralog, it is because the exact paralog is unknown or the statement applies to all known paralogs. Paralogs belonging to the "a" subfamily are most closely related to TGME49\_279100, and the "b" subfamily is related to TGME49\_020950 (www.toxodb.org). All MAF1 paralogs are located in tandem on chromosome II as shown in Figure 2-2, and the locus was previously identified as Expanded Locus 4 (EL4;[151]). Individual paralog numbers vary by strain; identified paralogs are named in Figure 2-4B. We do not assert that this represents the full complement of MAF1 paralogs from all queried strains. Indeed, further analysis of sequences and genomic sequence reads from the strains of interest will be necessary to determine this.



Figure 2-4. The *T. gondii* and *H. hammondi MAF1* loci harbor two distinct isoforms while only one isoform is present in *N. caninum*.

(A) Schematic representation of the predicted MAF1 protein. The signal peptide (SP) was predicted using SignalP v4.0 and the putative transmembrane domain (TM) was predicted by TMHMM v2.0. The proline-rich region (Pro-Rich) stretches from AA152-164 of TgMAF1RHb1 and is not found within all *MAF1* paralogs (e.g., TgMAF1RHa1, a2). (B) Phylogram of either cloned MAF1 amino acid sequences from *T. gondii*, *H. hammondi* and *N. caninum*, or

those downloaded directly from ToxoDB (with TG Gene numbers). Cloned sequences of all of the "b" paralogs from *T. gondii* did not match any predicted gene models in ToxoDB in terms of predicted coding region length and were left out of the analysis. Paralog family is indicated at the end of each name (e.g., "a1", "b1", "b2", etc.) (C) dN/dS ratio calculations for all *T. gondii* "b" *MAF1* paralogs, including b0. (\*) indicates significant evidence for diversifying selection for that particular paralog comparison (P<0.05).

	RHb3	RHb1	RHb2	RHb0	Me49b1	Me49b0	CTGb1	CTGb2
RHb3								
RHb1	2.696							
Rhb2	2.696	infinity						
RHb0	0.909	0.908	0.908					
Me49b1	1.365	1.363	1.363	0.918				
Me48b0	0.906	0.905	0.905	0.914	0.915			
CTGb1	1.356	1.354	1.354	0.911	0.915	0.908		
CTGb2	0.609	0.910	0.910	0.919	0.920	0.916	0.913	

Table 2. KaKs values for T. gondü "b" paralogs including b0

Table 3. Ks values for *T. gondü* "b" paralogs including b0

	RHb3	RHb1	RHb2	RHb0	Me49b1	Me49b0	CTGb1	CTGb2
RHb3								
RHb1	0.003							
Rhb2	0.003	0						
RHb0	0.009	0.009	0.009					
Me49b1	0.006	0.006	0.006	0.009				
Me48b0	0.009	0.009	0.009	0.009	0.009			
CTGb1	0.006	0.006	0.006	0.006	0.006	0.010		
CTGb2	0.009	0.010	0.010	0.009	0.010	0.009	0.010	

NCLIV 004730	1 MHVKRERVSRCRVVV CAAVCCLLGV APGLGVETDDRIAETM, APADIDEVPVOERRNNVEEOSEGIRIOTORSAPH	ISGR 80
HIMAF1HIGer04101	1 MWRYGRSRLTFLFAA-GULLGALTAGLGSUTSGSAGRNYQAPAGYADA - PQEAGDYVEERRERNEQQTFAPGPP	RGH 75
TgMAF1RHb0	0 1 MWR I WR C R L S F L F V T - G C L L G A L T A G L G S G M S D S V G R N V Q A P A G V A D A - S Q E A G D V V E E R T E R T E E Q V F A P G P P	RRH 75
TgMAF1ME49b0	0 1 MWR I WRCRL SFL FAT - GCLL GAL TAGL GSQMSD SVG RNVQ APA GVADA - SQEAGD V VEERT ERTEEH VFAPGP	RRH 75
TaMAF1RHb1	1 1 MWR I WRCRLSFLFAT - GCLLGAL TAGLGSOMSDSVGRNVO APAGVADA - SQEAGDVVEERTERTERQAFALGPP	RRH 75
TCMAE1DUb2	2 1 NUMPERING CHISCHEAT TACLOSOMS DSVC DNVC ADACVADA SOCACDVVCCDTED TED TECOVEAD CH	
		75
IGNAFICIODI	T T MWRTWRCRL SFLFAT-GULLGALTAGLGSUMSDSVGRNVQAPAGVADA - SUEAGDVVEERTERTEEQAFALGPP	<b>RRH</b> 75
TgMAF1ME49b1	1 1 MWR I WR CRL SFL FAT - G CL L GAL TAGL G SQMSD SV G RN VQ APAG VADA - SQ EAGD V V E ERTERT E EH I FALGPP	RRH 75
TaMAF1CTGb2	2 1 MWR I WR CRL SFL FAT - GCL L GAL TAGL G SOM SDS V GRN V QAPAG V ADA - SQEAGD V V E ERTERTE EN I FALGPP	RRH 75
HbMAE1HbGer041a1	1 1 WWRVGKSRUXELEAA - COLLOALTAGUGSOUSGSACRNVOARAGVADA - ROEAGDVVEERRERNEOOLEARGRE	RGH 75
	IMWRVGRSRETTETAA-GULLOALTAGLOSUTSGSAGRNVQAFAGVADA-FQEAGDVVEERRERREQUTFAFGFF	KGH 75
IGVEG_279100 (a1)	) 1 MWRTWRCRLSFLFVT - GCLLGALTAGLGSQMSDSVGRNVQAPAGVADA - SQEAGDVVEERTERTEEQAFALGPP	RRH 75
TGME49_279100 (a1)	) 1 MWR I WR CRL SFL FAT - GCL L GAL TAGL G SQMSD SVG RNVQ APAG VADA - SQE AGD V VEERTERTEQQ V FAL GPP	RRH 75
ToMAF1RHa1	1 1 MWR LWRCRL SEL FAT - GOLL GAL TAGL G SOMSDS VG RN VO A PAG VADA - SO FAG DV V F F RT F FO V F A PG P	RRH 75
TCCT1 270100 (a1)		0.04 76
10011_2/9100 (a1)	) I MWRIWRCRESFERVI-GCELGAE IAGEGSUMSDSVGRNVGAFAGVADA - SGEAGDVVEERIERIEEGVFAFGF	<b>KKH</b> 75
NCLIV 004730	0 81 SRGSFVSRRTSALTSKLRRNRKA I VMGVVGVA AV LAALYVARRWWTQKPREPGDSPPPEEPGDSPPRAGK	150
HbMAE1HbGar041b1	1 76 SSESLEDRSDSVTARDERNER LALVATAVGVAVILAALVAL BREWARD CEPEDRAPDSMEDDEVLEKEDESSEDE	PPP 156
		111 100
IGMAF1RH60	0 76 SSESLEPENPSVTARRERNERTTLTATAVGVAVTLAALYVLERRERAGPPGEPEPPTELETPEPPTELETPEPPTEL	145
TgMAF1ME49b0	0 76 SSE <mark>SLFPRNPSVTARRRRRRRRIALIATAVGVAVILAALYVLRR</mark> RWAQPPQEPEPPTRLRTPRPRAPSGQQ	145
TaMAF1RHb1	1 76 SSESLEPRNASVTARRRRNRRIALIATAVGVAVILAAVYVLRRRAQRPQDPEPPAPRSVEDPEVLPEEDEASSSLPP	PPP 156
TCMAE1DUb2		DDD 160
TOWAF INHD2	2 70 SEELFFANFOUTSRAAMATALTATAVGVAVILAALTVLAALTVLAARAGF GEFEFFAFRSVEDFEVLAEEDEASSSLFF	FFF 150
TgMAF1C⊺Gb1	1 76 SSESLFPRNPSVTARRRRNRIALIATAVGVAVILAAVYVLRRRWAQPPQEPEPPAPRSVEDPEVLPEEDEASSSLPP	PPP 156
TgMAF1ME49b1	1 76 SSESLFPRNPSVTARRRRNRRIALMATAVGVAVILAAVYVLRRRWAQPPQEPEPPAPPPVEDPEVLPEEDEASSSLPP	PPP 156
TaMAF1CTGb2	2 76 SSESLEPRNPSVTARRRRNR I ALMATAVGVAVILAAVYVLRRRWAOPPOAPEPPAPRSVEDPEVLPEEDEASSSLPP	PPP 156
HbMAE1HbCor041a1		146
	TO SEELFF AF OVIARA AND A LAVOYAVE AALTVE AR WARFFEEFEFFT AF AT A REFT F SOUL	145
TGVEG_279100 (a1)	) 76 SSESLFPRNPSVTARRRRNRIALIATAVGVAVILAALYVLRRRRAQPPQEPEPPTRPRTPRPRAPSGQQ	145
TGME49 279100 (a1)	) 76 SSESLFPRNASVTARRRRRRRIALIATAVGVAVILAALYVLRRRRAQPPQEPEPPTRLRTPRPRAPSGQQ	145
TaMAF1RHa1	1 76 SSESLEPRNASVTARRRRNRR LAPLATAVGVAVILAAL VVI RRRAOPPOEPEPPTRI RTPRPRAPSEOO	145
TCCT1 270100 (a1)		145
IGG11_279100 (a1)	) 76 SSESLFPRNASVTARRRRNR I AL I ATAVGVAVI LAALYVLRRRRAQPPQEPEPPTRLRTPRPRAPSEQQ · · · · · · ·	145
NCLIV 004730	0 151PKAPSQPPTERRDVPTGVRQVGARDL PAPPASHTALLACLENRQIEFFGPSNSPHGFTPLYDPEPOKRVAN	VDA 224
HEMAE1UbCor041b1		VDA 225
HIMAF HIGeld4101	I ST FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	223
TgMAF1RHb0	0 146	<b>VDA</b> 204
TgMAF1ME49b0	0 146PTCL_RDTK_VTFFGPSGRQH <mark>GFTPLYDP</mark> SPNKRVAT	VDA 204
TaMAF1RHb1	1 157 P SPPPPPPVEDPLSPESOTVDLSCLSGTTVRFFGPSHHFGGFTPLYDPAPDKRVAT	VDA 215
TCMAE1DU62		VDA 220
I GWAF IKHD2	2 IS/ FSFFFFSFFFFFFVEUFLSFESQTVULSGESGTTVKFFGFSHIFGGFTFLTUPAFUKKVAT	VUA 220
TgMAF1CTGb1	1 157 PSPPPPSPPPPSPPPPPVEV <mark>P</mark> LS <b>P</b> ESQTVDLS <mark>CL</mark> TGTT <mark>VRFFGPS</mark> HHSG <mark>GFTPLYDP</mark> APD <mark>KRVA</mark> T	VDA 225
TaMAF1ME49b1	1 157 PSPPPP	VDA 223
TaMAE1CTGb2		VDA 235
		2000
HhmAF1HhGer041a1	1 146 ··································	VNA 204
TGVEG_279100 (a1)	)146	VDA 204
TGME49 279100 (a1)	) 146	VDA 204
ToMAE1PHo1		VDA 204
IgiMAFIKHAT	140	VDA 204
TGGT1_279100 (a1)	)146	VDA 204
		204
NCLIV 004730	0 225 GTNAL FIGGAGVNGEVARL I TEEARRHEVR TPEOLSEHSKRVOEDL RLAVOHPRT LIELDTGARSPVFARSVGFVS	VVP 305
NCLIV_004730	0 225 GTNALFIGGAGVNGEVARL TEEARRHEVRLTPEQLSEHSKRVQEDLLRLAVQHPRTLIELDTGARSPVFARSYGFVS 4 226 GTNALFIGGGQNAGEEAKTILEEAOKHGNULTSSGLSAHSOOLOEIM NDAVKKDGKIVEVDTGAGSDVEADSEAFUG	VVP 305
NCLIV_004730 HhMAF1HhGer041b1	0 225 GTNALFIGGAGVNGEVARLLTEEARRHEVRLTPEQLSEHSKRVQEDLLRLAVQHPRTLIELDTGARSPVFARSYGFVS 1 226 GTNDLFIGGGGMDGEFAKTLLEEAQKHGMPLTSSGLSAHSQQIQEIMLNRAVKKPGKLVEVDTGAGSPVFARSFAFVP	VVP 305 VVP 306
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0	0 225 GTNALFIGGAGVNGEVARL TEEARRHEVRLTPEQLSEHSKRVQEDLLRLAVQHPRTLIELDTGARSPVFARSYGFVS 1 226 GTNDLFIGGGGMDGEFAKTLLEEAOKHGMPLTSSGLSAHSOOIOEIMLNRAVKKPGKLVEVDTGAGSPVFARSFAFVP 0 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLTSVALSEHSORIQOSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP	VVP 305 VVP 306 VVP 285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1ME49b0	0 225 G T NALFIGGAGVNG EVARL I TEEARRHEVR LTPEQLSEHSKRVQEDLLRLAVQHPRT LIELDTGARSPVFARSYGFVS 1 226 G T NALFIGGGGMDGEFAKT LLEEAQKHGMPLTSSGLSAHSQQIQEIMLNRAVKKPGKLVEVDTGAGSPVFARSFAFVP 0 205 GANALFIGGGGLNGQFAKT LLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP 2 205 GANALFIGGGGLNGQFAKT LLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP	VVP 305 VVP 306 VVP 285 VVP 285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1ME49b0 TgMAF1MEHb1	0 225 G T NAL FIGGAG VNG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 1 226 G T NDLFIGGGGMDGEFAKTLLEEAOKHGMPLTSSGLSAHSOOIOEIM. NRAVKKPGKLVEVDTGAGSPVFARSFAFVP 0 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP 2 206 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP 2 216 GANALFIGGGGLNGOFAKTLLEEAEKHGIR TPEELSOHSORIOQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP 2 216 GANALFIGGGGLNGOFAKTLEEAEKHGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1ME49b0 TgMAF1MEHb1 TGMAF1RHb1	0 225 G T NAL FIGGAGVING EVARL TEEARRHEVR TPEQLSEHSKRVQEDLLRLAVQHPRT LIELDTGARSPVFARSYGFUS 1 226 G T NAL FIGGGGMDGEFAKT LLEEAQKHGMPLTSSGLSAHSQQIQEIMLNRAVKKPGKLVEVDTGAGSPVFARSFAFV 0 205 GANAL FIGGGGLNGOFAKT LLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFV 0 205 GANAL FIGGGGLNGOFAKT LLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFV 1 216 GANAL FIGGGGLNGOFAKT LLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFV 2 201 GANAL FIGGGGLNGOFAKT LLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFV 2 201 GANAL FIGGGGLNGOFAKT LLEEAEKNGIR TPEELSQHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFV 2 201 GANAL FIGGGGLNGOFAKT LLEEAEKNGIR TPEELSQHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVF 2 201 GANAL FIGGGGLNGOFAKT LLEEAEKNGIR TPEELSQHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVF	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 296
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1ME49b0 TgMAF1RHb1 TGMAF1RHb1	0 225 G T NAL FIGGAGVNG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 1 226 G T NDLFIGGGGMOGEFAKTLLEEAGKHGMP TSSGLSAHSOOIOEIM NRAVKKPGKLVEVDTGAGSPVFARSFAFVP 0 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 126 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 296 VVP 301
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1ME49b0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1	0 225 G T NAL FIGGAGVNG EVARL TEEARRHEVRL TPEQLSEHSKRVQEDLL RLAVQHPRTLIELDTGARSPVFARSYGFVS 1 226 G T NDLFIGGGGMDGEFAKTLLEEAGKHGMPLTSSGLSAHSQQIQEIMLNRAVKKPGKLVEVDTGAGSPVFARSFAFVP 2 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLTSVALSEHSQRIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP 0 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLTSVALSEHSQRIQQSLRRAVKSPGKLVELDTGVASPVFARSFGFVP 1 216 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLTPEELSQHSQRIQQSLRRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLTPEELSQHSQRIQQSLRRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLSFVALSEHSQRIQQSLRRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIRLSFVALSEHSQRIQQSLRRAVKSPGKLVELDTGVASPVFARSFGFVP	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 296 VVP 301 VVP 306
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1	0 225 G T NAL FIGGAGVNG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 1 226 G T NDLFIGGGGMOGEFAKTLLEEAGKHGMP TSSGLSAHSOOIOEIM NRAVKKPGKLVEVDTGAGSPVFARSFAFVP 0 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 216 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TYPEOISOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TYPEOISOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TYPEOISOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 296 VVP 301 VVP 306 VVP 304
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2	0 225 G T NAL FIGGAG V NG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 1 226 G T NDL FIGGGGLNG G FAKTLLEEARKHGIR TSSGLSAHSO OLO EIMLNRAVKKPGKLVEVDTGAG SPVFARSFAFVP 0 205 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TSVALSEHSORIQQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVP 1 216 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TSVALSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEOLSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEOLSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 221 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEOLSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEOLSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEOLSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 1 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEOLSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEELSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEELSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEELSGHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGGLNG G FAKTLLEEAEKNGIR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG G FAKTLLEEAEKNGIR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG G FAKTLLEEAEKNGIR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG G FAKTLLFEVFYPENTANG FOR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG G FAKTLLFYPENTANG FOR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG FAKTLLFYPENTANG FOR TPEELSEHSORIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG FAKTLLFYPENTANG FOR TPEELSEHSORIQUSL RRAVKSPGKLVELTGVASPVFARSFGFVP 2 226 G ANAL FIGGGLNG FAKTLLFYPENTANG FOR TPEELSENG	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 306 VVP 301 VVP 306 VVP 304 VVP 304
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2	0 225 G T NAL F I GG AG V NG EVARL I TE EARRHE V R LTPEQL SEHS KRYQEDL RLAVQHPRT LIELDTG AR SPVFARSYGF V S 1 226 G T NDLF I GG GG M OG EFAKTLLE EA OKHOMP I TS SGLSAHSOO I OE I MLNRAVKK PGKLVELDTG VAS PVFARSFAF V P 2 25 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTSVALSEHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTSVALSEHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTSVALSEHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 24 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 24 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 24 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 24 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 24 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 24 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 23 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 23 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OQSLLRRAVKSPGKLVELDTG VAS PVFARSFGF V P 2 23 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OSL RAVKSPGKLVELDTG VAS PVFARSFGF V P 2 23 G ANAL F I GG GGLNG OFAKTLLE EA EKNG I RLTPEELSOHS OR I OSL RAVKSPGKLVELDTG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE FAR KNG I RLTPEELSOHS OR I OSL RAVKSPGKLVELD TG VAS PVFARSFGF V P 2 26 G ANAL F I GG GGLNG OFAKTLLE F F V F F F F F F F F F F F F F F F F	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 296 VVP 306 VVP 306 VVP 304 VVP 316
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1	0 225 G T NAL F I GGAG V NG EVARLL T EEARRHEVRL TPEQL SEHSKRVQE DLL RL AVQHP RT L I EL D T GARS P VFARS YG F VS 226 G T NDL F I GGGGM D G F AKTLL E EAOKHOM PL TSSGL SAH SOO I O E I ML NRAVKK PGKL VEL D T G VAS P VFARS F AF VP 0 205 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TSVAL SEHSOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 1 216 GANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TSVAL SEHSOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 21 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TSVAL SEHSOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 21 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPECI SOH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 221 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPECI SOH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 1 226 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPECI SOH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 1 226 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPECI SOH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 236 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPEEL SH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 246 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPEEL SH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 236 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPEEL SH SOR I QQSLL RRAVKSPGKL VEL D T G VAS P VFARS F G F VP 2 236 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPEEL SH SOR I QQSLL RRAVKSPGKL VUL D T G I AS P VFARS F G F VP 2 236 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPEEL SH SOR I QQSLL RRAVKSPGKL VUL D T G I AS P VFARS F G F VP 2 236 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I RL TPEEL SH SOR I QQSLL RRAVKSPGKL VUL D T G I AS P VFARS F AF VA 2 236 G ANAL F I GGGGL NG O F AKTLL EEAEKNG I SL TAAELSAH SQEL Q E I L RAVKSPGKL VUL D T G I AS P VFARS F AF VA 2 236 G ANAL F I GGGGL NG G F AKTLL EEAEKNG I SL TAAELSAH SQEL Q E RL L RAVKSPGKL VUL D T G I AS P VFARS F AF VA 2 2 36 G ANAL F I GGGSL NG G F AKTL I F A K S F AF VA	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 301 VVP 306 VVP 304 VVP 316 VVP 316 VVP 285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1)	0 225 G T NAL F I GG AG V NG E V ARLL T EE ARRHE V RL T PEQL SEHS K RVQ E DLL RL AVQHP RT L I EL D T G AR S P V F AR S Y G F V S 226 G T N DL F I GG GG M O G F A KT LL EE A G K HOMP L T S S G L S H S G O I O E I ML NR AV K K P G K L V E V D T G A G S P V F AR S F G F V P 0 205 G A NAL F I GG GG L NG O F A KT LL EE A E K NG I RL T S V AL S EH S G R I O Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 0 205 G A NAL F I GG GG L NG O F A KT LL EE A E K NG I RL T S V AL S EH S G R I O Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 0 205 G A NAL F I GG GG L NG O F A KT LL EE A E K NG I RL T S V AL S EH S G R I O Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 1 216 G A NAL F I GG GG L NG O F A KT LL EE A E K HG I RL T P E EL S G H S G R I O Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 2 221 G A NAL F I GG GG L NG O F A KT LL EE A E K HG I RL T P E EL S G H S G R I O Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 1 226 G A NAL F I GG G G L NG O F A KT LL EE A E K NG I R L T P E EL S H S G R I Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 1 224 G A NAL F I GG G G L NG O F A KT LL EE A E K NG I R L T P E EL S G H S G R I Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 2 236 G A NAL F I GG G G L NG O F A KT LL EE A E K NG I R L T P E EL S G H S G R I Q S LL RR AV K S P G K L V EL D T G V A S P V F AR S F G F V P 2 236 G A NAL F I GG G G L NG O F A KT LL EE A E K NG I R L T P E EL S G H S G R I Q S LL RR AV K S P G K L V L D T G I A S P V F AR S F G F V P 2 236 G A NAL F I GG G G M NG F A KT LL EE A E K NG I R L T P E EL S G H S G R I Q S LL RR AV K S P G K L V L D T G V A S P V F AR S F G F V P 2 236 G A N L F I GG G G M NG E F A KT I I E E A R R NH I S T A A E L S A H S Q E I Q E R L R D A E R P G S L V I D S G F S P V F AR S F A F V A 2 205 G T N L F I GG G G M M G E F A KT I I E E A R R NH I S L T A A E L S A H S Q E I Q E R L R D A E R P G S L V I D S G R S S P V F AR S F A F V A 2 205 G T N L	VVP 305 VVP 306 VVP 285 VVP 285 VVP 286 VVP 296 VVP 301 VVP 306 VVP 304 VVP 316 VVP 316 VVP 316 VVP 316
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGME49_279100 (a1)	0 225 G T NAL FIGGAG VNG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 226 G T NDLFIGGGGMOG FAKTLLEEACKHGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 205 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 216 GANALFIGGGGLNGOFAKTLLEEAEKHGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 221 20 ANALFIGGGGLNGOFAKTLLEEAEKHGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2221 20 ANALFIGGGGLNGOFAKTLLEEAEKHGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2226 CANALFIGGGGLNGOFAKTLLEEAEKHGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 223 CANALFIGGGGLNGOFAKTLLEEAEKHGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 224 CANALFIGGGGLNGOFAKTLLEEAEKNGIR LTPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 225 CANALFIGGGGLNGOFAKTLLEEAEKNGIR LTPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 CANALFIGGGGLNGFAKTLLEAAEKNGIR TTPELSOHSORIOQSL RRAVKSPGKLVELDTGVSPKFARSFAVA	VVP         305           VVP         306           VVP         285           VVP         285           VVP         285           VVP         296           VVP         301           VVP         306           VVP         285           I         P           285         I
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1MHC49b1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1)	0 225 G T NAL F I GG AG V NG E V ARLL T EE ARRHE V RL T PEQL SEHS K RVQ EDLL RL AVQH P RT L I EL D TG ARS P VF ARS YG F VS 226 G T N DL F I GG GG M OG E F AKTLL EE AE KHGI RL TS YGL SEHS (R 10 QSLL RRAVK PG KL VE V D TG AG S P VF ARS F G F VF 0 205 G ANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TS YGL SEHS QR I QQSLL RRAVK SPG KL VEL D TG YG S P VF ARS F G F VF 2 205 GANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TS YGL SEHS QR I QQSLL RRAVK SPG KL VEL D TG YG S P VF ARS F G F VF 2 205 GANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS QH SQR I QQSLL RRAVK SPG KL VEL D TG YG S P VF ARS F G F VF 2 216 GANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS QH SQR I QQSLL RRAVK SPG KL VEL D TG YG S P VF ARS F G F VF 2 221 GANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS GH SQR I QQSLL RRAVK SPG KL VEL D TG YG S P VF ARS F G F VF 2 226 G ANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS BHS QR I QQSLL RRAVK SPG KL VEL D TG YA S P VF ARS F G F VF 2 226 G ANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS BHS QR I QQSLL RRAVK SPG KL VEL D TG YA S P VF ARS F G F VF 2 226 G ANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS BHS QR I QQSLL RRAVK SPG KL VEL D TG YA S P VF ARS F G F VF 2 226 G ANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS GH SQR I QQSLL RRAVK SPG KL VUL D TG YA S P VF ARS F G F VF 2 236 G ANAL F I GG GG L NG O F AKTLL EE AE KHGI RL TP EELS QH SQR I QQSLL RRAVK SPG KL VUL D TG YA S P VF ARS F G F VF 2 236 G ANAL F I GG GG M NG E F AKTI I I EE ARRNH I S TAAEL SAH SQE I Q E RL RDA E RROPG SL VE I D TG YA S P VF ARS F G F VF 2 205 G T NRL F I GG GG M NG E F AKTI I I EE ARRNH I SL TATEL SAE SQE I Q E RL RDA E RROPG SL VE I D SG RF S R VF ARS F A F V 2 205 G T NRL F I GG GG M NG E F ANTI I I EE ARRNH I PL TATL SAE SQE I Q E RL RDA E RROPG SL VE I D SG RF S R VF ARS F A F V 2 2 05 G T NRL F I GG GG M NG E F ANTI I EE ARRNH I PL TATL SAE SQE I Q E RL RDA E RROPG SL VE I D SG RF S R VF ARS F A Y VA 2 2 05 G T NRL F I GG GG M NG E F ANTI I E E ARRNH I PL TATL SAE SQE I Q E RL RDA E RROPG	VVP 305 VVP 306 VVP 285 VVP 285 VVP 285 VVP 296 VVP 306 VVP 306 VVP 306 VVP 306 VVP 306 VVP 316 VVP 316 VVP 285 I VP 285 I VP 285 VVP 285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1)	0 225 G T NAL FIGGAG VNG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 226 G T NDLFIGGGGMOEFAKTLLEEARKHGIR TSVALSEHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 205 GANALFIGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 206 GANALFIGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 216 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 221 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2226 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2227 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2228 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2236 GANALFIGGGGMNGEFAKTILEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2236 GANALFIGGGGMNGEFAKTLLEEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 2236 GANALFIGGGGMNGEFAKTILEEARRNFISTAAELSAHSQELQERL RDAERRPGLVEIDSGSFSPVFARSFAFVA 2045 GTNLFIGGGGMNGEFAKTILEEARRNFISTAAESAHSQELQERL RDAERRPGLVEIDSGSFSPVFARSFAFVA 2055 GTNLFIGGGGMNGEFANTIIEEARRNFISTAAESAHSQELQERL RDAERRPGTVEIDSGRFSRVFARSFAVA 2055 GTNLFIGGGGMNGEFANTIIEEARRNFISTAAESAHSQELQERL RDAERRPGTVEIDSGRFSRVFARSFAVA	VVP 305 VVP 306 VVP 285 VVP 285 VVP 286 VVP 306 VVP 306 VVP 306 VVP 304 VVP 316 VVP 316 VVP 316 VVP 316 VVP 285 I VP 285 I VP 285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TgMAF1CTGb2 TgMAF1CTGb2 TgMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGMAF4RHa1 TGGT1_279100 (a1)	0 225 G T NAL F I GGAG V NG E VARL T E E ARRHE V R T P E QL SE H SKRV QE DLL RLAVQH P RT LIEL D T G AR S P V F AR S Y G F V S 226 G T N DL F I GG GG M D G F A KT LLE E A C KH G M P L T S G L S A H S Q I Q E I M L NRAV K K P G K V E V D T G A G S P V F AR S F G F V P 0 205 G A N AL F I G G G L NG O F A KT LLE E A E KH G I R L T S V AL S E H S Q R I Q S LL RRAV K S P G K V E L D T G V A S P V F AR S F G F V P 1 216 G A N AL F I G G G L NG O F A KT LLE E A E KH G I R L T S V AL S E H S Q R I Q S LL RRAV K S P G K V E L D T G V A S P V F AR S F G F V P 2 216 G A N AL F I G G G L NG O F A KT LLE E A E KH G I R L T P E L S Q H S Q R I Q S LL RRAV K S P G K L V E L D T G V A S P V F AR S F G F V P 2 216 G A N AL F I G G G G L NG O F A KT LLE E A E KH G I R L T P E L S Q H S Q R I Q S LL RRAV K S P G K L V E L D T G V A S P V F AR S F G F V P 2 216 G A N AL F I G G G G L NG O F A KT LLE E A E KH G I R L T P E D S G H S Q R I Q S LL RRAV K S P G K L V E L D T G V A S P V F AR S F G F V P 2 226 G A N AL F I G G G G L NG O F A KT LLE E A E KH G I R L S F V A L S E H S Q R I Q S LL RRAV K S P G K L V E L D T G V A S P V F AR S F G F V P 2 246 G A N AL F I G G G G L NG O F A KT LLE E A E KH G I R L T P E E L S H S Q R I Q S LL RRAV K S P G K L V E L D T G V A S P V F AR S F G F V P 2 246 G A N AL F I G G G G L NG O F A KT LLE E A E KH G I R L T P E E L S H S Q R I Q S LL RRAV K S P G K L V L D T G I A S P V F AR S F G F V P 2 236 G A N AL F I G G G G L NG O F A KT LLE E A E KH G I R L T P E E L S H S Q R I Q S LL RRAV K S P G K L V L D T G I A S P V F AR S F G F V P 2 236 G A N AL F I G G G G M NG E F A KT I I E E A R R N I S L A A E L S A H S Q E L Q E R L R A V K S P G K L V L D T G I A S P V F AR S F G F V P 2 236 G T N L F I G G G G M NG E F A NT I I E E A R R N I P L T A T Q S A E S Q E I Q E R L R D A E R P G S L V L D S G F S R V F AR S F A Y A A 2 2 05 G T N R L F I G G G G M NG E F A NT I I E E A R R N I P L T A T Q S A E S Q E I Q E R L R D A E R P G T L V I D S G R F S R V F A R S F A Y A A 2 2 05 G	VVP         305           VVP         306           VVP         285           VVP         285           VVP         296           VVP         285           VVP         304           VVP         285           VP         285           VP         285           VP         285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1)	0 225 G T NAL FIGGAG VNG EVARL TEEARRHEVR TPEQLSEHSKRVQEDL RLAVQHPRT LIELDTGARSPVFARSYGFVS 226 G T NDLFIGGGGMOEFAKTLLEEARKHGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 0 205 GANALFIGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 216 GANALFIGGGLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 216 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 216 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 221 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GANALFIGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GANALFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GINNLFIGGGGLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSL RRAVKSPGKLVELDTGVASPVFARSFGFVP 226 GINNLFIGGGGNNGEFAKTLIEEARRNNIS TAAELSAHSGELOERL RDAERRGKLVELDTGVASPVFARSFGFVP 226 GINNLFIGGGGNNGEFAKTLIEEARRNNIS TAAELSAHSGELOERL RDAERRGTVEIDSGSFSPVFARSFAVA 205 GINNLFIGGGGNNGEFANTIIEEARRNNIS TAAELSAHSGELOERL RDAERRGTVEIDSGRFSRVFARSFAVA 205 GINNLFIGGGGNNGEFANTIIEEARRNNIS TAAELSAHSGELOERL RDAERRGTVEIDSGRFSRVFARSFAVA 205 GINNLFIGGGGNNGEFANTIIEEARRNNIS TAAELSAHSGELOERL RDAERRGTVEIDSGRFSRVFARSFAVA 205 GINNLFIGGGGNNGEFANTIIEEARRNNIS TAAELSAHSGELOERL RDAERGAGTVEIDSGRFSRVFARSFAVAA 205 GINNLFIGGGGNNGEFANTIIEEARRNNIS TAAELSAHSGELOERL RDAERGGTVEIDSGRFSRVFARSFAVAA	VVP         305           VVP         306           VVP         285           VVP         285           VVP         285           VVP         296           VVP         301           VVP         304           VVP         305           VVP         306           VVP         306           VVP         306           VVP         306           VVP         304           VVP         304           VVP         285           VP         285           VP         285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1)	0 225 G T NAL F I GGAG VNG E VARLL T EE ARRHE VR LTPEQL SEH SKRVQE DLL RLAVQHP RT LIEL DT GARS PVFARS YG F VS 226 G T NDLF I GGGGMOE F AKTLLE EA OKHOM PLTSSGLSAH SOOIO E IML NRAVKKPGKLVEL DT GVAS PVFARS FAF VR 0 205 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTSVALSEH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 216 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTSVALSEH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 216 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTSVALSEH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 216 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 221 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2226 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2236 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2246 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2256 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2266 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2266 GANALF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2266 GT NTLF I GGGGLNGO F AKTLLEE AE KHGI RLTPEEL SOH SORIOQSLL RRAVKSPGKLVEL DT GVAS PVFARS FGF VR 2266 GT NTLF I GGGGMNGE FANT I I EE ARRNH I SLTAAEL SAH SQELO ERLL RDAE RRPGSLVEL DT SG SF S PVFARS FAF VA 2056 GT NRLF I GGGMNGE FANT I I EE ARRNH I PLTATEL SAE SOEIO ERLL RDAE RRPGSLVE I D SG SF S RVFARS FAF VA 2056 GT NGLF I GGVGMNGE FADT I I EE ARRNR I PLTATEL SAE SOEIO ERLL RDAE ROPGT I VE I D SG RF SRVFARS FAF VA 2056 GT YGLF I GG VGMNGE FADT I I EE ARRNR I PLTATEL SAE SOEIO ERLL HDAE ROPGT I VE I D SG RF SRVFARS FAF VA 2056 GT YGLF I GG VGMNGE FADT I I EE ARRNR I PLTATEL SAE SOEIO ERLL HDAE ROPGT I VE I D SG RF SRVFARS FAF VA 2056 TYGLF I GG VGMNGE FADT I I EE ARRNR I PLTATEL SAE SOEIO ERLL HDAE ROPGT I VE I D SG RF SRVFARS FAF VA	VVP         305           VVP         306           VVP         285           VVP         285           VVP         296           VVP         301           VVP         306           VVP         306           VVP         316           VP         316           VP         316           VP         316           VP         285
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMF17700 (a1) NCLIV_004730 HhMAF1HhGer041b1	0 225 G T NAL F I GG AG VNG EVARL T E E ARRHE VR T T PEQL S EH SKRVQ E DL RL AVQHP RT L I EL D TG AR S P VF ARS YG F VS 226 G T N DL F I GG GG M G E F AKT L L E E A GK HGMP T S SG L S AH S GO I O E I M. NR AVKK P G K L VE V D TG AG S P VF AR S F AF V P 0 205 G ANAL F I GG GG L NG O F AKT L L E E A EK NG I R L T S VAL S EH S GR I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 226 G ANAL F I GG GG L NG O F AKT L L E E A EK NG I R L T S VAL S EH S GR I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 205 G ANAL F I GG GG L NG O F AKT L L E E A EK NG I R L T S VAL S EH S GR I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 216 G ANAL F I GG GG L NG O F AKT L L E E A EK HG I R L T PEEL S G H S G R I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 216 G ANAL F I GG GG L NG O F AKT L L E E A EK HG I R L T PEEL S G H S G R I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 221 G ANAL F I GG GG L NG O F AKT L L E E A EK NG I R L T P E L S H S G R I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 222 G ANAL F I GG GG L NG O F AKT L L E E A EK NG I R L T P E L S H S G R I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 223 G ANAL F I GG G G L NG O F AKT L L E E A EK NG I R L T P E E L S H S G R I QO S L RR AVKS P G K L VE L D TG VA S P VF AR S F G F V P 224 G ANAL F I GG G G L NG O F AKT L L E E A EK NG I R L T P E E L S H S G R I QO S L RR AVKS P G K L VL D TG I A S P VF AR S F G F V P 225 G T N L F I GG G G M NG E F AKT L L E E A EK NG I R L T P E E L S G H S G R I QO S L RR AVKS P G K L VL D TG I A S P VF AR S F G F V P 226 G T N L F I GG G G M NG E F AKT L I E E A RR NR I P L T A T L S A S G E I Q E R L R D A ER P G S L V E I D S G R F S R VF AR S F A Y M 205 G T N L F I GG G G M NG E F ANT I I E E A RR NR I P L T A T E L S A E S G E I Q E R L R D A ER P G T L V E I D S G R F S R VF AR S F A Y M 205 G T N L F I G G G G M NG E F ANT I I E E A RR NR I P L T A T E L S A E S G E I Q E R L R D A ER P G T L V E I D S G R F S R VF AR S F A Y M	VVP         305           VVP         285           VVP         285           VVP         285           VVP         285           VVP         301           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VVP         316           VVP         285           IVP
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGME49_279100 (a1) TGMAF1279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF12Hb0	0 225 G T NAL F I GGAG VNG EVARL T EEARRHEVR T PEQL SEHSKRVQEDL RLAVQHPRT LIELD TGARSPVFARSYGFVS 226 G T NDLF I GGGGMOG EFAKTLLE EAQKHGMP TSSGLSAHSQQIQ EIM. NRAVKKPGK VEVD TGAG SPVFARSFAFVP 0 205 GANALF I GGGGLNGOFAKTLLEEAEKNGI RLTSVALSEHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 216 GANALF I GGGGLNGOFAKTLLEEAEKNGI RLTSVALSEHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 2212 GANALF I GGGGLNGOFAKTLLEEAEKNGI RLTPEELSQHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 2225 GANALF I GGGGLNGOFAKTLLEEAEKNGI RLTPEELSQHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 2236 GANALF I GGGGLNGOFAKTLLEEAEKNGI RLTPEELSQHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 2305 GTNLFI GGGGMNGEFAKTI I EEARRNH I TATELSAESOEI QERLL RDAERRPGS VEI DSG SFSPVFARSFAVA 2405 GTNLFI GGGGMNGEFAKTI I EEARRNH I TATELSAESOEI QERLL RDAERRPGS VEI DSG SFSPVFARSFAVA 2405 GTNLFI GGGGMNGEFANTI I EEARRNH I PLTATELSAESOEI QERLL RDAERRPGT VEI DSG RFSRVFARSFAVA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERLL RDAERRPGT VEI DSG RFSRVFARSFAVAA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERLL RDAERRPGT VEI DSG RFSRVFARSFAVAA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERLL RDAERRPGT VEI DSG RFSRVFARSFAVAA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERLLHDAEROPGT VEI DSG RFSRVFARSFAVAA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERLLHDAEROPGT VEI DSG RFSRVFARSFAVAA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERLLHDAEROPGT VEI DSG RFSRVFARSFAVAA 2405 GTYGLFI GGVGMNGEFADTI I EEARRNRI PLTATELSAESOEI QERL HDAEROPGT VE	VVP         305           VVP         285           VVP         285           VVP         285           VVP         286           VVP         200           VVP         200           VVP         200           VVP         200           VVP         301           VVP         306           VVP         306           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VP         304           VP         285           VP <td< td=""></td<>
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TGMF1RHb0	0 225 G T NAL F I GG AG VNG EVARL T EE ARRHEVR T PEQL SEHS KRYQEDL RLAVQHPRT LIELDTG ARSPVFARSYGFVS 226 G T NDLFI GG GG MOE FAKTLLE EA RKHEVR T SYALSEHS OR IO SLARAVKKPGKLVELDTG VASPVFARSFAFVF 0 205 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T SYALSEHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 206 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T SYALSEHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 207 208 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T SYALSEHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 208 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T SYALSEHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 209 209 CANALFI GG GG LNG OFAKTLLE EA EKNGIR T SYALSEHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 201 201 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T PEELS OHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 202 202 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T PEELS EHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 202 6 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T PEELS EHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 202 6 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T PEELS EHS OR IO SLARAVKSPGKLVELDTG VASPVFARSFGFVF 203 6 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T PEELS GHS OR IO SLARAVKSPG KLVELDTG VASPVFARSFGFVF 204 GANALFI GG GG LNG OFAKTLLE EA EKNGIR T PEELS GHS OR IO SLARAVKSPG KLVELDTG VASPVFARSFGFVF 205 G T NLFI GG GG MNG EFAKTLI E EA RRNNI SLAAELS AHS OEL DE RLLR DA ERPORT VEIDS GSFS PVFARSFGFVF 205 G T NLFI GG GG MNG EFAKTLI I E A RRNNI PLATTELSAES OEI DE RLLR DA ERPORT VEIDS GSFS PVFARSFA VA 205 G T NRLFI GG GG MNG EFANTI I E A RRNNI PLATTELSAES OEI DE RLLR DA ERPORT VEIDS GRFS RVFARSFA VA 205 G T NGLFI GG VG MNG EFADTI I E EA RRNNI PLATTELSAES OEI DE RLLR DA ERPORT VEIDS GRFS RVFARSFA VA 205 G T YGLFI GG VG MNG EFADTI I E A RRNNI PLATTELSAES OEI DE RLLRDAE ROPGT T VEIDS GRFS RVFARSFA VA 205 G T YGLFI GG VG MNG EFADTI I E A RRNNI PLATTELSAES OEI DE RLLRDAE ROPGT T VEIDS GRFS RVFARSFA VA 206 G T YGLFI GG VG MNG EFADTI I E A RRNNI PLATTELSAES OEI DE RLLRDAE ROPGT T VEIDS GRFS RVFARSFA VA 206 G T YGLFI GG VG MNG EFADTI I E A RRNNI PLATTELSAES OEI DE RLLHDAE ROPGT VEIDS GRFS	VVP         305           VVP         285           VVP         285           VVP         285           VVP         285           VVP         286           VVP         285           VVP         306           VVP         285           VP         285           PIE         386           PLQ         387           PVQ         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb0	0 225 G T NAL F I GGAGVNG EVARL TEEARRHEVR TPEQL SEHSKRVQEDL RLAVQHPRT I IELD TGARSPVFARSYGFVS 226 G T NDLF I GGGGMD E FAKTLLE EAQKHGMP TSSGLSAHSQOIQEIM. NRAVKKPGK VEVD TGAGSPVFARSFAFVP 0 205 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 216 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 216 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TSVALSEHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 216 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TYPEOLSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 217 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TYPEOLSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 228 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TYPEOLSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 229 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GANALF I GGGGLNG O FAKTLLE EAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GANALF I GGGGMNG E FAKTLI EEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GTNLF I GGGGMNG E FAKTLI EEAEKNGIR TPEELSOHSQRIQQSL RRAVKSPGK VELD TGVASPVFARSFGFVP 226 GTNLF I GGGGMNG E FAKTLI EEAEKNGIR TATELSAHSQELQERL RDAERROGT VEID SGRFSRVFARSFAVA 205 GTNLF I GGGGMNG E FAKTLI EEAERNRIP TATELSAESQEIQERL RDAERROGT VEID SGRFSRVFARSFAVAA 205 GTYGLF GGVGMNG E FADTII EEARRNRIP TATELSAESQEIQERL RDAERPGT VEID SGRFSRVFARSFAVAA 205 GTYGLF GGVGMNG E FADTII EEARRNRIP TATELSAESQEIQERL RDAEROPGT VEID SGRFSRVFARSFAVAA 205 GTYGLF GGVGMNG E FADTII EEARRNRIP TATELSAESQEIQERL RDAEROPGT VEID SGRFSRVFARSFAVAA 205 GTYGLF GGVGMNG E FADTII EEARRNRIP TATELSAESQEIQERL RDAEROPGT VEID SGRFSRVFARSFAVAA 205 GTYGLF GGVGMNG E FADTII EEARRNRIP TATELSAESQEIQERL RDAEROPGT VEID SGRFSRVFARSFAVAA 206 GTGWDESQTGRNVGATFIHLKVPVYGDVNNNWYVVAR SGAPADASYSAAVATIKAT I SGVI	VVP         305           VVP         305           VVP         205           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VP         305           VP         306           VP         306           VP         285           PIE         386           PVQ         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1	0 225 G T NAL F I GG GG VNG EVARL T EE ARRHEVR T PEQL SEHS KRYQEDL RLAVQHPRT LIELDTG ARSPVFARSYGFVS 226 G T NDLFI GG GG NG G F AKTLLE EA EKNGIR T SYALSEHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 0 205 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T SYALSEHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 206 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T SYALSEHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 216 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSORIO SLARAVKSPGK VELDTG VASPVFARSFGFVF 216 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSORIO SLARAVKSPGK VELDTG VASPVFARSFGFVF 216 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSORIO SLARAVKSPGK VELDTG VASPVFARSFGFVF 216 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSORIO SLARAVKSPGK VELDTG VASPVFARSFGFVF 221 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSOHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 222 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSOHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 223 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSOHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 224 G ANALFI GG GG LNG OF AKTLLE EA EKNGIR T PEELSOHS OR I O SLARAVKSPGK VELDTG VASPVFARSFGFVF 225 G T NLFI GG GG MNG EFANTI I E EA RNN I PLTATELSA ESO EI O ERLLRDA ER PGSLVEI DTG VASPVFARSFGFVF 205 G T NLFI GG GG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 205 G T NLFI GG GG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 205 G T NLFI GG GG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 205 G T YGLFI GG VG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 205 G T YGLFI GG VG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 206 G T YGLFI GG VG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 206 G T YGLFI GG VG MNG EFANTI I E EA RNN I PLTATELSAESO EI O ERLLRDA ER PGT VEI D SG RFSRVFARSFA VA 206 G T YGLFI GG VG MNG EFANTI I E EA RNN I PLATELSAESO EI O ERLLRDAE ROPGT VEI D SG RFSRVFARSFA V	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         301           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         301           VP         304           VVP         301           VP         304           VP         305           VP         305           VP         301           VP         301           VP         301           VP         301           VP         285           PIE         386           PVQ         366           PVQ         366           PVQ         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb1	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR LTPEQL SEHS KRYQEDLL RLAVQHPRT LIELD TG ARSP VF ARSYGF VS 226 G T NDLF I GG GG MO G F AKTLLE EA OKHGMP I TSSGLSAHSOO I O E IM. NRAVKKPGK VEV DT G AG SP VF ARSF AF VP 0 205 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 206 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 207 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 208 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 209 C ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 201 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 202 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 203 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 204 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR LTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 205 G T NLF I GG GG MNG E F AKTLI E EA RRNNIS LTAAELSAHSOE I QE RL RAVKSPGK VELD TG VA SP VF ARSF GF VP 205 G T NLF I GG GG MNG E F ANTI I E E ARRNNIS LTAAELSAHSOE I QE RL RDA ERP GSL VE I D SG RF SR VF ARSF AT VA 205 G T NLF I GG GG MNG E F ANTI I E E ARRNNIS LTAAELSAHSOE I QE RL RDA ERP GT VE I D SG RF SR VF ARSF AT VA 205 G T NLF I GG GG MNG E F ANTI I E E ARRNNIS LTAAELSAHSOE I QE RL RDA ERP GT VE I D SG RF SR VF ARSF AT VA 205 G T YG LF I GG VG MNG E F ADTI I E E ARRNNIP L TATELSAESOE I QE RL RDA ERP GT VE I D SG RF SR VF ARSF AT VA 205 G T YG LF I GG VG MNG E F ADTI I E E ARRNNIP L TATELSAESOE I QE RL RDA ERP GT VE I D SG RF SR VF ARSF AT VA 205 G T YG LF I GG VG MNG E F ADTI I E E ARRNNIP L TATELSAESOE I QE RL RDA ERP GT VE I D SG RF SR VF ARSF AT VA 205 G T YG LF I GG VG MNG E F ADTI I E E ARRNNIP L TATELSAESOE I QE RL RDA ERP GT VE I D SG RF SR VF ARSF AT VA 2	VVP         305           VVP         305           VVP         205           VVP         285           VVP         285           VVP         296           VVP         306           VVP         306           VVP         306           VVP         306           VP         305           VP         305           VP         285           VP         307           PVO         366           PVO         377
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1	0 225 G T NAL F I GG GG VNG EVARL T EE ARRHEVR T PEQL SEHS KRYQEDL RLAVQHPRT LIELD TG ARS PVF ARS YG F VS 226 G T NDLF I GG GG MO G F AKTLLE EA GKHGMP T SSGL SAHSOO I O E IM NRAVKKPGK VELD TG VAS PVF ARS F AF VP 0 205 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T SVALSEHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 206 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T SVALSEHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 207 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T SVALSEHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 208 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELSOHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 210 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELSOHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 221 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELSOHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 221 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELS BHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 223 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELS BHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 236 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELS BHSOR I OQSL RRAVKSPGK VELD TG VAS PVF ARS F G F VP 2 236 G ANALF I GG GG LNG O F AKTLLE EA EKNGIR T PEELS G HSOR I OQSL RRAVKSPGK VUED TG VAS PVF ARS F G F VP 2 236 G ANALF I GG GG MNG E F AKTLLE EA EKNGIR T PEELS OHSOR I OQSL RRAVKSPGK VUED TG VAS PVF ARS F G F VP 2 236 G TNLF I GG GG MNG E F AKTLLE EA EKNGIR T PEELS OHSOR I OQSL RRAVKSPGK VUED TG VAS PVF ARS F AF VA 2 056 G TNLF I GG GG MNG E F ANTTI I E ARRNRIP T TATELSAESOE I O ERLL RDA ERP GT VE I D SG RF S RVF ARS F AF VA 2 056 G TNLF I GG GG MNG E F ANTTI I E ARRNRIP T ATELSAESOE I O ERLL RDA ERP GT VE I D SG RF S RVF ARS F AF VA 2 056 G TYGLF I GG VG MNG E F ANTTI I E ARRNRIP T TATELSAESOE I O ERLL RDA ERP GT VE I D SG RF S RVF ARS F AF VA 2 056 G TYGLF I GG VG MNG E F ANTTI I E ARRNRIP T ATELSAESOE I O ERLL RDA ERP GT VE I D SG RF S RVF ARS F AF VA 2 056 TYGLF I GG VG MNG E F ANTTI I E ARRNRIP T ATELSAESOE I O ERLL RDA ERP GT VE I D SG RF	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         301           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         301           VP         304           VVP         301           VP         304           VP         285           IVP         285           IVP         285           PIC         387           PVQ         366           PVQ         366           PVQ         382           PVQ         382
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HHGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR LTPEQL SEHS KRYQEDLL RLAVQHPRT LIELD TG ARSP VF ARSYGF VS 226 G T NDLF I GG GG MO G F AKTLLE EA QKHGMP I TSSGLSAHSOO I QE I MLNRAVKKPGK VEV D TG AG SP VF ARSF AF VP 0 205 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 205 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 206 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 207 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTSVALSEHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 208 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 209 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 224 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 225 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 226 G ANALF I GG GG LNG Q F AKTLLE EA EKNG I RLTPEELSOHSOR I QQSLLRRAVKSPGK VELD TG VA SP VF ARSF GF VP 226 G T NLF I GG GG MNG E F AKTLI E EA RRNN I SLTAAELSAHSOE I Q ERLLRDA ERRPGS LVELD TG VA SP VF ARSF GF VP 205 G T NLF I GG GG MNG E F AKTLI E EA RRNN I SLTAAELSAHSOE I Q ERLLRDA ERRPGS LVE I D SG RF SRVF ARSF A V A 205 G T NLF I GG GG MNG E F ANT I I E A RRNN I PLTATELSAESOE I Q ERLLRDA ERRPGT LVE I D SG RF SRVF ARSF A V A 205 G T NLF I GG GG MNG E F ANT I I E A RRNN I PLTATELSAESOE I Q ERLLRDA ERRPGT LVE I D SG RF SRVF ARSF A V A 205 G T NGL F GG GG MNG E F ANT I I E A RRNN I PLTATELSAESOE I Q ERLLRDA ERRPGT LVE I D SG RF SRVF ARSF A V A 205 G T NGL F GG GG MNG E F ANT I I E A RRNN I PLTATELSAESOE I Q ERLLRDA ERRPGT LVE I D SG RF SRVF ARSF A V A 205 G T NGL F GG GG MNG E F ANT I I E A RRNN I PLTATELSAESOE I Q ERLLRDA ERRPGT LVE I D SG RF SRVF ARSF A V A 205 G T NGL F GG GG MNG E F ANT I I E A RRNN I PLTATELSAESOE I Q ERLLRDA ERQ PGT LVE I D SG RF SRVF ARSF A V A 205 G T NGL F	VVP         305           VVP         305           VVP         205           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VP         255           VP         255           VP         285           PVQ         366           PVQ         367           PVQ         377           PVQ         377           PVQ         377
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR L TPEQL S EH S KRYQ EDL RL AVQHP RT L I EL D TG AR S P VF ARS YG F VS 226 G T NDL F I GG GG M G E F AKTLL E E A EKNG I R L TS VAL S EH S GRI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 0 205 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TS VAL S EH S GRI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 205 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TS VAL S EH S GRI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 205 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TS VAL S EH S GRI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 201 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TP E L S GH S G RI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 221 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TP E L S EH S G RI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 221 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TP E L S EH S G RI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 226 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TP E L S EH S G RI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 236 G ANAL F I GG GG L NG O F AKTLL E E A EKNG I R L TP E L S EH S G RI QO SLL RRAVKS P G K VEL D TG VA S P VF ARS F G F VF 2 236 G ANAL F I GG GG M NG E F AKTLL E E A EKNG I R L TP E L S HS G RI QO SLL RRAVKS P G K VUL D TG VA S P VF ARS F G F VF 2 236 G T N L F I GG GG M NG E F AKTLL E E A EKNG I R L TP E L S AH S G LI Q E RL R D A ERR P G L VE I D S G F S R VF ARS F A F V A 2 205 G T N L F I GG GG M NG E F AKTLL E E A RRNR I P L TA T L S A E S Q E I Q E RL R D A E R P G T VE I D S G R F S R F A RS F A T V A 2 205 G T Y G L F I G G VG M NG E F ANTI I L E A RRNR I P L TA T L S A E S Q E I Q E RL R D A E R P G T L VE I D S G R F S R VF ARS F A T V A 2 205 G T Y G L F I G G VG M NG E F ANTI I L E A RRNR I P L TA T L S A E S Q E I Q E R L R D A E R P G T L VE I D S G R F S R VF A RS F A T V A 2 205 G T Y G L F I G G VG M NG E F ANTI I L E A RRNR I P L TA T L S A E S Q E I Q E R L R D A E R P G T L VE I D S G R	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         301           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         301           VP         304           VVP         301           VP         285           IVP         285           IVP         285           IVP         285           PIC         387           PVQ         366           PVQ         386           PVQ         387           PVQ         387           PVQ         387
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HHGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR LTPEQL SEHS KRYQ EDL RL AVQHP RT LIELD TG ARS P VF ARS YG F VS 226 G T NDL F I GG GG MO G F AKTLL EE AQKHGMP I TSSGL SAHSQO I Q E I M. NRAVKKPGK V VE VD TG AG S P VF ARS FG F VF 0 205 G ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTS VAL SEHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 206 G ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTS VAL SEHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 207 G ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTS VAL SEHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 208 G ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 209 C ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 200 C ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 200 C ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 200 C ANAL F I GG GG L NG Q F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VE LD TG VAS P VF ARS FG F VF 209 C T NLF I GG GG MNG E F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VL D TG I AS P VF ARS FG F VF 200 S G T NLF I GG GG MNG E F AKTLL EE AEKNG I RLTPEEL SOHS QR I Q SL RRAVKSPGK V VL D TG I AS P VF ARS FG F VF 200 S G T NLF I GG GG MNG E F ANT I I EE ARRNR I P L TATEL SAES QE I Q ERL RDAERRPG SL VE I D SG RF S RVF ARS FA V A 205 G T NLF I GG GG MNG E F ANT I I EE ARRNR I P L TATEL SAES QE I Q ERL RDAERRPG T VE I D SG RF S RVF ARS FA V A 205 G T NGL F I GG GG MNG E F ANT I I EE ARRNR I P L TATEL SAES QE I Q ERL RDAERRPG T VE I D SG RF S RVF ARS FA V A 205 G T NGL F I GG GG MNG E F ANT I I EE ARRNR I P L TATEL SAES QE I Q ERL RDAERRPG T VE I D SG RF S RVF ARS FA V A 205 G T NGL F I GG GG MNG E F ANT I I EE ARRNR I P L TATEL SAES QE I Q ERL RDAERRPG T VE I D SG RF S RVF ARS FA V A 205 G T NGL F I GG GG MNG E F ANT I I EE ARRNR I P L TATEL SAES QE I Q ERL RDAERDPG T VE I D SG RF S RVF ARS FA V A	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         304           VVP         304           VVP         304           VP         285           PVQ         366           PVQ         387           PVQ         387           PVQ         387           PVQ         387           PVQ         387           PVQ         387
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb1 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR L TPEQL SEHS KRYQ EDL RL AVQHPRT L I EL D TG ARS PVF ARS YG F VS 226 G T NDL F I GG GG MOG E F AKTLL E E AG KHOMP I TS SG L SAHS OO I O E I ML NRAVKK PG K VEL D TG VAS PVF ARS F AF VP 0 205 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TS VAL SEHS OR I O QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 205 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TS VAL SEHS OR I O QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 205 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TY PE E SOHS OR I O QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 21 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TP E EL SOHS OR I O QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 221 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TP F E L SOHS OR I O QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 221 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L SF VAL SEHS OR I O QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 224 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TP E EL SHO R I Q QSLL RRAVKS PG K VEL D TG VAS PVF ARS F G F VP 2 236 G ANAL F I GG GG L NG O F AKTLL E E AE KHO I R L TP E EL SOHS OR I Q QSLL RRAVKS PG K VUL D TG VAS PVF ARS F G F VP 2 236 G ANAL F I GG GG MNG E F AKTLI E E AE KHO I R L TP E EL SOHS OR I Q QSLL RRAVKS PG K VUL D TG VAS PVF ARS F G F VP 2 205 G T NL F I GG GG MNG E F ANTI I E E ARRNN I P L TA TE L SAES OE I Q E RL R DA E RRPGS L VE I D SG RF S RVF ARS F AF V A 2 205 G T NL F I GG GG MNG E F ANTI I E E ARRNN I P L TA TE SAES OE I Q E RL R DA E RRPG SL VE I D SG RF S RVF ARS F AF V A 2 205 G TYGL F I GG VG MNG E F AD TI I E E ARRNN I P L TA TE SAES OE I Q E RL R DA E RRPG SL VE I D SG RF S RVF ARS F AF V A 2 205 G TYGL F I GG VG MNG E F AD TI I E E ARRNN I P L TA TE SAES OE I Q E RL R DA E RRPG SL VE I D SG RF S RVF ARS F AF V A 2 205 G TYGL F I GG VG MNG E F AD TI I E E ARRNN I P L TA TE SAES OE I Q E RL R DA E RRPG SL VE I D SG RF S RVF ARS F AF V A 2 205 G TYGL F I GG VG MNG E F AD TI I E E ARRNN I P L A TE L SAES OE I Q E RL R DA E RRPG SL	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         280           VVP         301           VVP         304           VVP         301           VVP         304           VVP         301           VVP         304           VVP         301           VP         355           VP         285           IVP         285           PLQ         366           PVQ         366           PVQ         387           PVQ         387           PVQ         387           PVQ         387           PVQ         397           PVQ         397
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1CTGb2 HhMAF1HHGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041a1 TGMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR LTPEQL SEHS KRYQEDLL RLAVQHP RT LIELD TG ARS P VF ARS YG F VS 226 G T NDLF I GG GG MO G F AKTLL EE AQKHGMP ITS SG SAHSOO I QE IM NRAVKKPGKLVELD TG VAS P VF ARS FG F VP 0 205 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTS VAL SEHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 206 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTS VAL SEHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 207 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTS VAL SEHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 208 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTS VAL SEHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 209 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTPEELS QHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 201 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTPEELS EHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 2026 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTPEELS EHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 2026 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTPEELS EHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 2026 G ANALF I GG GG LNG Q F AKTLL EE AEKNG I RLTPEELS QHSOR I QQSLL RRAVKSPGKLVELD TG VAS P VF ARS FG F VP 2026 G T NLF I GG GG MNG E F AKTLI EE ARRNR I P LTATQL SAESOE I Q ERLL RDA ERP GSLVE I DS GSF S P VF ARS FA V A 2020 ST NRLF I GG GG MNG E F ANTI I E E ARRNR I P LTATELS AESOE I Q ERLL RDA ERP G T VE I DS G FF S R VF ARS FA Y A 2036 G T NG L F GG GG MNG E F ANTI I E E ARRNR I P LTATELS AESOE I Q ERLL RDA ERP G T VE I DS G FF S R VF ARS FA Y A 2046 G T YG LF I GG YG MNG E F ANTI I I E E ARRNR I P LTATELS AESOE I Q ERLL RDA ERP G T VE I DS G R FS R VF ARS FA Y A 2056 G T NG L F AG YG MNG E F ANTI I E E ARRNR I P LTATELS AESOE I Q ERLL RDA ERP G T VE I DS G R FS R VF ARS FA Y A 2056 G T YG L F I GG YG MNG E F ANTI I I E E ARRNR I P LTATELS AESOE I Q ERLL RDA ERP G T VE I DS G R FS R VF ARS FA Y A 2056 G T YG L W K S S VG ANG Y F I H I LK P E YT P YG NL NN YMMY T YAP S G AAPD AS YS Q AY Q AT I KG YYG A AA YN KS P YG A AS YN D T P AG Q Y 2056 G T Y	VVP         305           VVP         305           VVP         206           VVP         285           VVP         301           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VP         285           IVP         366           VVQ         366           VV         366           VV         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 TgMAF1CTGb4 TgMAF1 TgMAF1CTGb4 TgMAF1CTGb4 TgMAF1 T	0 225 G T NAL F I GG AG VNG E V ARLL T EE ARRHE VR L TPEQL S EH S KRYQ EDL RL AVQHPRT L I EL D TG AR S P VF AR S YG F VS 226 G T NAL F I GG GG L NG G F AKTLL EE AE KNGI RL TS YAL S EH S GRI QO SLL RRAVK PG KL VEL D TG VAS P VF AR S F G F VF 0 205 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TS YAL S EH S GRI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 1 216 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TS YAL S EH S GRI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 2 221 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TY PE EL S G H S GRI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 2 221 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TY PE EL S G H S GRI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 2 221 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TY PE EL S G H S GRI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 2 224 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TY PE EL S HS G RI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 2 243 G ANAL F I GG GG L NG G F AKTLL EE AE KNGI RL TP E EL S HS G RI QO SLL RRAVK S PG KL VEL D TG VAS P VF AR S F G F VF 2 245 G T N L F I GG GG L NG G F AKTLL EE AE KNGI RL TP E L S HS G RI QO SLL RRAVK S PG KL VUL D TG VAS P VF AR S F G F VF 2 256 G T N L F I GG GG M NG E F AKTL I EE AE KNGI RL TP E L S G H S G R I Q SLL RRAVK S PG KL VUL D TG VAS P VF AR S F G F VF 2 256 G T N L F I GG GG M NG E F ANTI I E E AR RNR I PL TA T L S AE SOCI Q E R LL RD A E R R PG SL VE I D SG R F S R VF AR S F A F V A 2 056 G T N L F I GG GG M NG E F ANTI I E E AR RNR I PL TA T L S AE SOCI Q E R LL RD A E R R PG SL VE I D SG R F S R VF AR S F A Y A 2 056 G T N L F I GG GG M NG E F ANTI I E E AR RNR I PL TA T L S AE SOCI Q E R LL RD A E R R G T L VE I D SG R F S R VF AR S F A Y A 2 056 T Y G L F G G Y G M NG E F AD T I I E E AR RNR I PL TA T L S AE SOCI Q E R LL RD A E R R G T L VE I D SG R F S R VF AR S F A Y A 2 056 T Y G L F G G Y G N M G E F AD T I I E E AR RNR I PL TA T L S AE SOCI Q E R LL RD A E R R G T L Y L Y D SG R F S R YF A R S F A	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         280           VVP         301           VVP         204           VVP         301           VVP         304           VVP         301           VVP         304           VVP         301           VVP         301           VP         355           PIE         386           PVQ         366           PVQ         366           PVQ         366           PVQ         367           PVQ         385           PVQ         385           PVQ         385           PVQ         385           PVQ         385           PVQ         385           PVQ         366           PVQ         385           PVQ         385           PVQ         366           PVQ         365           PVQ         366           PVQ         365           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 HhMAF1HHGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGMF1HhGer041b1 TGMAF1HhGer041b1 TgMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TgMAF1RHb1 TgMAF1RHb1 TgMAF1RHb1 TgMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TgMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGME49_279100 (a1) TGME49_279100 (a1)	0 225 G T NAL F I GG AG VNG E V ARL T TE E ARRHE VR LTPEQL SEHS KRYQ E DL RL AVQHP RT LIEL DTG ARS P VF ARS YG F VS 226 G T NDL F I GG GG MO G F AKTLL E E AQKHGMP LTSSGL SAHSOO I O E IM. NRAVKKPGKL VEL DTG VAS P VF ARS F G F VP 0 205 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTSVAL SEHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 216 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTSVAL SEHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 216 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTSVAL SEHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 216 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTSVAL SEHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 216 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTPEEL SAHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 221 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTPEEL SHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 224 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTPEEL SHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 224 G ANAL F I GG GG LNG O F AKTLL E E AEKING I RLTPEEL SHSOR I QOSL RRAVKSPGKL VEL DTG VAS P VF ARS F G F VP 225 G T NLF I GG GG MNG E F AKTLL E E AEKING I RLTPEEL SQHSOR I QOSL RRAVKSPGKL VU D TG I AS P VF ARS F G F VP 226 G T NLF I GG GG MNG E F AKTLL E E AEKING I RLTPEEL SQHSOR I QOSL RRAVKSPGKL VU D TG I AS P VF ARS F A F V A 205 G T NLF I GG GG MNG E F ANTI I E E ARRNRI P L TATG L SAESOE I Q ERLL RDA ERP G T VE I D SG RF S R VF ARS F A Y VA 205 G T NG L M G GG MNG E F ANTI I E E ARRNRI P L TATE L SAESOE I Q E RLL RDA ERP G T L VE I D SG RF S R VF ARS F A Y VA 205 G T YG L F I GG VG MNG E F ANTI I I E E ARRNRI P L TATE L SAESOE I Q E RLL RDA ERP G T L VE I D SG RF S R VF ARS F A Y VA 205 G T YG L F I GG VG MNG E F ANTI I I E E ARRNRI P L TATE L SAESOE I Q E RLL RDA ERP G T L VE I D SG RF S R VF ARS F A Y VA 205 G T YG L F I GG VG MNG E F ANTI I I E E ARRNRI P L TATE L SAESOE I Q E RLL RDA ERP G T L VE I D SG RF S R VF ARS F A Y VA 205 G T YG L I W K S S V G S N YG N T I H I L K P Y T YG D NNNVMY T VAP SG	VVP         305           VVP         305           VVP         206           VVP         285           VVP         285           VVP         301           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         304           VVP         304           VVP         304           VP         285           IVP         366           PVQ         367           PVQ         367           PVQ         367           PVE         366           PVE         366           PVE         366           PVE         366           PVE
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHC041a1 TGVEG_279100 (a1) TGVEG_279100 (a1) TGVEG_279100 (a1)	0 225 G T NAL F I GG AG VNG E V ARLL T E E ARRHE VR L TPEQL S EHS KRYQ E DL RL AVQHPRT L I EL D TG ARS P VF ARS YG F VS 226 G T NAL F I GG GG L NG G F AKTLL E E A EKNGI RL TS VAL S EHS QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 205 G ANAL F I GG GG L NG G F AKTLL E E A EKNGI RL TS VAL S EHS QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 205 G ANAL F I GG GG L NG G F AKTLL E E A EKNGI RL TS VAL S EHS QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 206 G ANAL F I GG GG L NG G F AKTLL E E A EKNGI RL TP E LS QH S QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 207 G ANAL F I GG GG L NG G F AKTLL E E A EKNGI RL TP E LS QH S QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 208 G ANAL F I GG GG L NG G F AKTLL E E A EKNGI RL TP E LS GH S QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 204 G ANAL F I GG GG L NG G F AKTLL E E A EKNGI RL TP E LS EHS QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 205 G T N RL F I GG GG L NG G F AKTLL E E A EKNGI RL TP E LS EHS QR I Q SLL RRAVKS P GK L VEL D TG VAS P VF ARS F G F VP 205 G T N RL F I GG GG L NG G F AKTLL E E A EKNGI RL TP E LS G H S QR I Q SLL RRAVKS P GK L VUL D TG VAS P VF ARS F G F VP 205 G T N RL F I GG GG M NG E F ANT I I E E A RNNH I S L TA A ELS A H S Q E I Q E RL RD A ERR P G SL VE I D S G R F S R VF ARS F A F V A 205 G T N RL F I GG GG M NG E F ANT I I E E A RNNH I S L TA TE LS A E S Q E I Q E RL RD A ERR P G SL VE I D S G R F S R VF ARS F A F V A 205 G T N RL F I GG GG M NG E F ANT I I E E A RNNH I P L A T L S A E S Q E I Q E RL RD A ERR P G T L VE I D S G R F S R VF A RS F A T VA 205 G T N RL F I GG GG M NG E F ANT I I E E A RNNH I P L A T L S A E S Q E I Q E RL RD A ERR P G T L VE I D S G R F S R VF A RS F A T VA 205 G T N RL F I GG GG M NG E F A D T I I E A RNNH I P L A T L S A E S Q E I Q E RL RD A ERR P G T L VE I D S G R F S R VF A RS F A T VA 205 G T N G L W R S S VG N V T F I H I L K P V T P Y G D P K N N M M Y T VA P S G A P P T T Y S L A Y K T I I G Y ( G A A A Y N D T P A G Q Y	VVP         305           VVP         305           VVP         206           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VP         305           VP         306           VP         306           PVQ         366           PVQ         382           PVQ         382           PVQ         397           PVQ         366           PVQ         366           PVQ         366           PVE         366           PVE         366           PVE         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TGMAF1CTGb2 HhMAF1HCG2279100 (a1) TGMAF1RHa1 TGMAF1 TGMAF1RHa1 TGMAF1 T	0 225 G T NAL F I GG AG VNG E V ARL T E E ARRHE VR LTPEQL S EH S KRYQ E DL RL AVQHP RT LIEL D TG ARS P VF ARS YG F VS 226 G T ND LF I GG GG MO G F AKT LL E E AQKHGMP ITS SG LS AH SOO I O E IM. NRAVKK P GK LVEL D TG VAS P VF ARS F G F VF 0 205 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTS VAL S EH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 20 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTS VAL S EH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 16 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTS VAL S EH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 16 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SOH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 21 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SOH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 226 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SHSOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 226 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SHSOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 226 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SHSOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 236 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SOH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 236 G ANAL F I GG GG LNG O F AKT LL E E AEKING I R LTPEEL SOH SOR I QOSL RRAVKS P GK LVEL D TG VAS P VF ARS F G F VF 2 236 G ANAL F I GG GG MNG E F ANT I I E E ARRNRI P L TA TO LS A E SOC I Q E RL RD AER P G T LVE I D SG RF S RVF ARS F A V A 2 050 G T NL F I GG GG MNG E F ANT I I E E ARRNRI P L TA TE L SA E SOC I Q E RL RD AER P G T LVE I D SG RF S RVF ARS F A V A 2 050 G T YG LF I GG VG MNG E F ADT I I E E ARRNRI P L TA TE L SA E SOC I Q E RL RD AER P G T LVE I D SG RF S RVF ARS F A V A 2 050 G T YG LF I GG VG MNG E F ADT I I E E ARRNRI P L TA TE L SA E SOC I Q E RL RD AER P G T LVE I D SG RF S RVF ARS F A V A 2 050 G T YG L W K S E VG SNN Q Y T I H I L K P V T P YG D NNN M Y T VAP S G AP P T Y S LA YKT I I KG V Y	VVP         305           VVP         305           VVP         206           VVP         285           VVP         285           VVP         301           VVP         304           VVP         304           VVP         301           VVP         304           VVP         304           VVP         304           VVP         304           VVP         301           VP         304           VP         285           VP         285           VP         285           VP         285           VP         285           VP         285           VP         366           PVQ         366           PVQ         382           PVQ         382           PVQ         397           PVE         366           PVE         <
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RH60 TGMAF1 TGMAF1RH60 TGMAF1	0 225 GTNAL FIGGA VNG EVA RL LTEEARRHEYRLTPEQLSEHSKRYQEDLL RLAVQHPRT LIELDTGARSPVFARSYGFYS 226 GTNAL FIGGA MOEFAKTLLEEAEKNGIRLTSYALSEHSARVQEDLL RLAVKSPGKLVELDTG VASPVFARSFAFYS 205 GANALFIGGA LNG GFAKTLLEEAEKNGIRLTSYALSEHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 216 GANALFIGGA LNG GFAKTLLEEAEKNGIRLTSYALSEHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GANALFIGGA LNG GFAKTLLEEAEKNGIRLSFVALSEHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GANALFIGGA LNG GFAKTLLEEAEKNGIRLSFVALSEHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GANALFIGGA LNG GFAKTLLEEAEKNGIRLSFVALSEHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GANALFIGGA LNG GFAKTLLEEAEKNGIRLTSELSGHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GANALFIGGA LNG GFAKTLLEEAEKNGIRLTSELSGHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GANALFIGGA LNG GFAKTLLEEAEKNGIRLTSELSGHSARI OQSLLRRAVKSPGKLVELDTG VASPVFARSFAFYS 226 GTNRLFIGGA LNG GFAKTLLEEAEKNGIRLTSELSGHSARI OLSLLRAVSSPGKLVELDTG VASPVFARSFAFYS 226 GTNRLFIGGA LNG GFAKTLIEEARRNNIPLTATELSGHSARI OLSLLRAVSSPGKLVELDTG VASPVFARSFAFYS 205 GTNRLFIGGA LNG GFAKTLIEEARRNNIPLTATELSAESOEI O ERLLRD AERRPG LVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG EFANTI I EEARRNNIPLTATELSAESOEI O ERLLRD AERRPG LVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG EFANTI I EEARRNNIPLTATELSAESOEI O ERLLHD AEROPG TLVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG EFANTI I EEARRNNIPLTATELSAESOEI O ERLLHD AEROPG TLVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG EFANTI I EEARRNNIPLTATELSAESOEI O ERLLHD AEROPG TLVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG FFANTI I EEARRNNIPLTATELSAESOEI O ERLLHD AEROPG TLVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG FFANTI I EEARRNNIPLTATELSAESOEI O ERLLHD AEROPG TLVEIDSG RFSRVFARSFAFYS 205 GTYGLFIGG VGNNG FFANTI I EEARRNNIPLTATELSAESOEI O ERLLHD AEROPG TLVEIDSG RFSRVFARSFAFYS 205	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VP         305           VP         285           PVQ         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) MCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TGMAF1CCGC TGMAF1CTGb2	0 225 GTNAL FIGGAOVNGEVARL LTEEARRHEVRLTPEQLSEHSKRVQEDL RLAVQHPRT IELDTGARSPVFARSYGFVS 226 GTNDLFIGGGMDGEFAKTLLEEAEKNGIRLTSVALSEHSORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 0 205 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTSVALSEHSORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 20 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTSVALSEHSORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 21 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSANGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 22 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSANGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 22 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSANGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 22 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSENGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 22 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSENGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 22 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSENGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 23 GANALFIGGGLNGGFAKTLLEEAEKNGIRLTPELSENGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 24 GANALFIGGGGLNGGFAKTLLEEAEKNGIRLTPELSENGORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 25 GTNTFFIGGGGMNGEFAKTIIEEARRNNIPLTATGESABORIOQSLLRAVKSPGKLVELDTGVASPVFARSFGFVF 2 25 GTNFFIGGGGMNGEFAKTIIEEARRNNIPLTATGLSABORIOQSLLRAVKSPGKLVELDTGVASPVFARSFAVA 2 05 GTNFFIGGGMNGEFAKTIIEEARRNNIPLTATGLSABORIOQSLLRAVKSPGKLVELDTGVGAVSPVFARSFAVA 2 05 GTYGFIGGVGMNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLNDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGMNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGMNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGNNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGNNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGNNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGNNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHDAEROPGTLVEIDSGRFSRVFARSFAVA 2 05 GTYGFIGGVGNNGEFAKTIIEEARRNNIPLTATELSAESOEIOERLLHAAEAANNGTVGAAAANNDTPAGOOV 2 2 000 GTYGFIGGVGNNGEFAKTIIEEARNNIPLTATELSAESOEIOERLLHAAEAANNOTAGAAAANNDTPAGOOV 2 2 0000 GTYGFIGGVGNNGEFAKTIIEEARNNIPLTATELSAESOEIOERLHAAAANNGT	VVP         305           VVP         305           VVP         206           VVP         285           VVP         285           VVP         301           VVP         306           VVP         301           VVP         301           VVP         301           VVP         304           VVP         301           VVP         301           VP         304           VVP         301           VP         285           VP         366           PVQ         366           PVQ         382           PVQ         382           PVQ         366           PVQ <t< th=""></t<>
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHc3b1 TGMAF1RHc3b1 TGMAF1RHc3b1 TGMAF1RHc3b2 TGMAF1CTGb1 TGMAF1RHc3b1 TGMAF1 TGMAF1RHc3b1 TGMAF1RHc3b1 TGMAF1	0 225 GTNALFIGGA SVNGEVARLL TEEARRHEVRLTPEQLSEHSKRVOEDLRLAVQHPRTLIELDTGARSPVFARSYGFVS 226 GTNOLFIGGGOLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLLRRAVKSPGKLVELDTGVASPVFARSFGFVF 205 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 206 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 207 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 208 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TSVALSEHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 208 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 209 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 204 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 205 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSLRRAVKSPGKLVELDTGVASPVFARSFGFVF 206 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSLRRAVKSPGKLVULDTGVASPVFARSFGFVF 208 GANALFIGGGOLNGOFAKTLLEEAEKNGIR TPEELSOHSORIOQSLRRAVKSPGKLVULDTGVASPVFARSFGFVF 205 GTNRLFIGGGOMNGEFANTIIEEARRNRIP TATOLSAESOEIOERLROAERPGSLVEIDSGRFSRVFARSFAFVA 205 GTNRLFIGGGOMNGEFANTIIEEARRNRIP TATOLSAESOEIOERLROAERPGSLVEIDSGRFSRVFARSFAFVA 205 GTNGLFIGGGOMNGEFANTIIEEARRNRIP TATELSAESOEIOERLLROAERPGTLVEIDSGRFSRVFARSFAFVA 205 GTVGLFIGGVOMNGEFADTIIEEARRNRIP TATELSAESOEIOERLLROAERPGTLVEIDSGRFSRVFARSFAFVA 205 GTVGLFIGGVOMNGEFADTIIEEARRNRIP TATELSAESOEIOERLLROAERPGTLVEIDSGRFSRVFARSFAFVA 205 GTVGLFIGGVOMNGEFADTIIEEARRNRIP TATELSAESOEIOERLLROAERVGVIVATISGVIGAAAAYNDTPAGOO 206 GLMWESKVGANGVTFIHILKPEVTPYGDPKNNAMVTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOO 206 GLMWESKVGANGVTFIHILKPEVTPYGNLNNVMMYTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOO 207 GLMWESEVGNNGVTFIHILKPEVTPYGNLNNVMMYTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOO 206 GLMWESEVGNNGVTFIHILKPEVTPYGNLNNVMYTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOO 207 GLMWESEVGNNGVTFIHILKPEVTPYGNLNNNVMYTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOO 207 GLMWESEVGNNGVTFIHILKPEVTPYGNLNNVMYTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOO 207 GLMWESEVGNNGVTFIHILKPEVTPYGNLNNVMYTVAPSGAAPDTYSLAVKTTIAGVIGAAAAYNDTPAGO	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VP         255           VP         255           VP         265           VP         285           PVQ         386           PVQ         386           PVQ         366           PVE         366           PVE         366           PVE         366           PVE         366           PVE         366           PVE         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) MCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RHb3 TGMAF1RHb1 TGMAF1RH0200 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1RH1 TGMAF1 TGMA	0 225 GTNALFIGGAGVNGEVARLL TEEARRHEVRLTPEQLSEHSKRVQEDLRLAVQHPRTLIELDTGARSPVFARSYGFVS 226 GTNDLFIGGGMDEFAKTLLEEACKHGIRLTSVALSEHSQTIQGSLLRAVKKPGKLVELDTGVASPVFARSFAFVA 205 GANALFIGGGLNQOFAKTLLEEAEKNGIRLTSVALSEHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 205 GANALFIGGGLNQOFAKTLLEEAEKNGIRLTSVALSEHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 206 GANALFIGGGLNQOFAKTLLEEAEKNGIRLTSVALSEHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 207 GANALFIGGGLNQOFAKTLLEEAEKNGIRLTSVALSEHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 208 GANALFIGGGGLNQOFAKTLLEEAEKNGIRLTPEELSQHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 208 GANALFIGGGLNQOFAKTLLEEAEKNGIRLTPEELSQHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 208 GANALFIGGGGLNQOFAKTLLEEAEKNGIRLTPEELSQHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 208 GANALFIGGGGLNQOFAKTLLEEAEKNGIRLTPEELSQHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 205 GTNRLFIGGGGNNGEFANTIIEEARRNIRPTATOLSAESQHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 205 GTNRLFIGGGGNNGEFANTIIEEARRNIPTTATELSAESQHSQTIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFVP 205 GTNRLFIGGGGNNGEFANTIIEEARRNRIPTTATOLSAESQETGERLLRDAERRPGSLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTTATELSAESQETGERLLRDAERRPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTTATELSAESGETGERLLRDAERRPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTTATELSAESGETGERLLRDAERRPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTTATELSAESGETGERLLRDAERRPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGERLHDAEROPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGERLLRDAERRPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGERLHDAEROPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGERLHDAEROPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGERLHDAEROPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGERLHDAEROPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGENLHDAEROPGTLVEIDSGRFSRVFARSFAYVA 205 GTYGLFIGGGNNGEFANTIIEEARRNRIPTATELSAESGETGENLAAAAXNDTPAGOY 20	VVP         305           VVP         305           VVP         206           VVP         285           VVP         285           VVP         301           VVP         206           VVP         301           VVP         301           VVP         304           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         301           VVP         301           VVP         305           VVP         305           VVP         305           VVP         306           PVQ         366           PVQ         367           PVQ         366           PVQ         366           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGGT1_279100 (a1) TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb2 TgMAF1RHb2 TgMAF1RH6041a1 TGVEG_279100 (a1) TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHB1 TGMAF1 TG	0 225 GTNALFIGGAGVNGEVARLLTEEARRHEVRLTPEQLSEHSKRVQEDLRLAVQHPRTLIELDTGARSPVFARSYGFVS 226 GTNOLFIGGAGNDEFAKTLLEEARNGIRLTSVALSEHSQIQELRAVKKPGKLVELDTGVASPVFARSFAFVA 226 GTNOLFIGGGLNQGFAKTLLEEARKNGIRLTSVALSEHSQIQGSLLRAVKSPGKLVELDTGVASPVFARSFGFV 216 GANALFIGGGGLNQGFAKTLLEEARKNGIRLTSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 216 GANALFIGGGGLNQGFAKTLLEEARKNGIRLTSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 216 GANALFIGGGGLNQGFAKTLLEEARKNGIRLTSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGLNQGFAKTLLEEARKNGIRLSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGLNQGFAKTLLEEARKNGIRLSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGLNQGFAKTLLEEARKNGIRLSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGLNQGFAKTLLEEARKNGIRLSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGLNQGFAKTLLEEARKNGIRLSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGLNQGFAKTLLEEARKNGIRLSVALSEHSQRIQQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGNNGFFATTIEEARRNNISLAAELSAHSQEDGENGRIQSLLRAVKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGNNGFFATTIEEARRNNISLAAELSAHSQEDGENLRDARKSPGKLVELDTGVASPVFARSFGFV 226 GANALFIGGGGNNGFFATTIEEARRNNISLAAELSAHSQEDGENLLRDAERRPGSLVEIDSGRFSRVFARSFAFV 205 GTNTLFIGGGGNNGFFATTIEEARRNNIPLTATGLSAESQEIGERLLRDAERRPGSLVEIDSGRFSRVFARSFAFV 205 GTNTLFIGGGVMNGFFATTIEEARRNNIPLTATGLSAESQEIGERLLRDAERRPGSLVEIDSGRFSRVFARSFAFV 205 GTYGLFIGGVGMNGFFATTIEEARRNRIPLTATELSAESQEIGERLLRDAERRPGTLVEIDSGRFSRVFARSFAFV 205 GTYGLFIGGVGMNGFFATTIEEARRNRIPLTATELSAESQEIGERLLRDAERRPGTLVEIDSGRFSRVFARSFAFV 205 GTYGLFIGGVGMNGFFATTIEEARRNRIPLTATELSAESQEIGERLLRDAERAGTIVEIDSGRFSRVFARSFAFV 205 GTYGLFIGGVGMNGFFATTIEEARRNRIPLTATELSAESQEIGERLLRDAERAGTIVEIDSGRFSRVFARSFAFV 205 GTYGLFIGGVGMNGFFATTIEEARRNRIPLTATELSAESQEIGERLLRDAERAGTIVEIDSGRFSRVFARSFAFV 205 GTVGLFIGGVGMNGFFATTIEEARRNRIPLTATELSAESQEIGERLLRDAERAGTIVEIDSGAAAAYNOTPAGQQY 206 GLMWESSVGSNVGVTFIHLLKPEVTPYGNLNNNMMYTVAPSGAAPDTTVSLAXKTTIAGVIGAAAAYNOTPAGQQY 207 GLMWESSVGSNVGVTFIHLLKPEVTPYGNLNNNMMYTVAPSGAAPDTTVSLAXKTTIAGVIGAAAAYNOTPAGQQY 208 GLMM	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VVP         306           VP         285           PVQ         366           PVQ         386           PVQ         386           PVE         366           PVE         366           PVE         366           PVE         366           PVE         366
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) MCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb0 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME40_279100 (a1)	0 225 GT NALF I GG GA VNG EVA RLLT EE ARRHEVR TPEQL SEHSKRV0 EDLLRLAVQHPRT I ELDTG ARSPVFARSYGF VS 226 GT NDLF I GG GG NG GF AKT LLEE ARRHEVR TPEQL SEHSKRV0 EDLLRLAVK PG KLVE DTG AG BPVFARSFAF VF 226 GT NDLF I GG GG NG GF AKT LLEE AE KNG I R TSVALSEH GO I OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TSVALSEH GO I OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 221 GANALF I GG GG NG GF AKT LLEE AE KNG I R TSVALSEH GO I OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 222 GANALF I GG GG NG GF AKT LLEE AE KNG I R TYPE I SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TYPE I SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TYPE I SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TYPE I SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TYPE I SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TPEEL SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TPEEL SCH SORI OG SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 226 GANALF I GG GG NG GF AKT LLEE AE KNG I R TPEEL SCH SOLO SLLRRAVK SPG KLVELDTG VASPVFARSFGF VF 236 GT NT LF I GG GG NG GF AKT LLEE AE KNG I R TPEEL SCH SOLO SLLRRAVK SPG KLVELDTG VASPVFARSFAF VA 240 S GT NT LF I GG GG NG GF AKT LLEE AE KNG I R TPEEL SCH SOLO SLLRRAVK SPG KLVELDTG VASPVFARSFAF VA 250 ST NG LF I GG VG MNG EF ANT I I EE ARRNRIP TATOL SAE SOEI O C RLLROA ERRPG TLVE I DSG RF SRVFARSFAF VA 250 ST NG LF I GG VG MNG EF ANT I I EE ARRNRIP TATEL SAE SOEI O C RLLROA ERRPG TLVE I DSG RF SRVFARSFAF VA 250 ST YG LF I GG VG MNG EF ADT II EE ARRNRIP TATEL SAE SOEI O C RLLROA ERRPG TLVE I DSG RF SRVFARSFAF VA 250 ST YG LF I GG VG MNG EF ADT II EE ARRNRIP TATEL SAE SOEI O C RLLROA ERRPG TLVE I DSG RF SRVFARSFAF VA 250 ST YG LF I GG VG MNG EF ADT II EE ARRNRIP TATEL SAE SOEI O C RLLROA ERRPG TLVE I DSG RF SRVFARSFAF VA 25	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         301           VVP         285           VVP         301           VVP         301           VVP         304           VVP         304           VVP         301           VVP         302           VVP         303           PVQ         360           PVQ         361           PVQ         362           PVQ         362           PVQ         362           PVQ         362           PVQ         366           PVQ         366           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HHGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb0 TgMAF1RHb0 TgMAF1RHb2 TgMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0 TgMAF1RHb0	0 225 GTNALF I GGGAVNGEVA RLLT EE ARRHEVR T PEOL SEHSKRVD EDLL RLAVQH PRT LIELD TG ARS PVFARSYGF VS 226 GTNALF I GGGGUNGEFAKT LLEE ACKHGMP T SSGL SAHSOO I O E IML NRAVKK PGKLVELD TG VAS PVFARSFAF VF 226 GTNALF I GGGGL NGGFAKT LLEE AEKNG I R. TSVALSEHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGL NGGFAKT LLEE AEKNG I R. TSVALSEHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGL NGGFAKT LLEE AEKNG I R. TSVALSEHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGL NGGFAKT LLEE AEKNG I R. TYPEL SQHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGL NGGFAKT LLEE AEKNG I R. TYPEL SQHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGL NGGFAKT LLEE AEKNG I R. TYPEL SQHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGL NGGFAKT LLEE AEKNG I R. TYPEL SQHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGN GGFAKT LLEE AEKNG I R. TPEEL SGHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGN MGEFAKT LLEE AEKNG I R. TPEEL SGHSOR I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGN MGEFAKT I I EE ARRNH I P TATOLSAE SQH I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 226 GANALF I GGGGN MGEFAKT I I EE ARRNH I P TATOLSAE SQH I QOSLLRRAVKSPGKLVELD TG VAS PVFARSFGFV 205 GTNLF I GGGGN MGEFAKT I I EE ARRNH I P TATOLSAE SQH I QERLLRAPARD TVS PG KLVELD SGR FSRVFARSFAYVA 205 GTVGLF I GGVGN MGEFA TI I I EE ARRNH I P TATEL SAE SQH I QERLLRAPARD TVS I DSGR FSRVFARSFAYVA 205 GTVGLF I GGVGN MGEFA TI I I EE ARRNR I P TATEL SAE SQH I QERLLRAPARD TVG TI I SGR FSRVFARSFAYVA 205 GTVGLF I GGVGN MGEFA TI I I EE ARRNR I P TATEL SAE SQH I QERLL HDAERO FGT LVE I DSGR FSRVFARSFAYVA 205 GTVGLF I GGVGN MGEFA TI I I EE ARRNR I P TATEL SAE SQH I QERLL HDAERO FGT LVE I DSGR FSRVFARSFAYVA 205 GTVGLF I GGVGN MGEFA TI I I EE ARRNR I P TATEL SAE SQH I QERLL HDAERO FGT LVE I DSGR FSRVFARSFAYVA 205 GTVGLF I GGVGN MGEFA TI I I E PVTY YGD NNNWMYTVAP SGAAPD TY SLAYKT I I AGVI GAAAA NNDT PAGQOY 206 GLWK E SEVG NG VTF I H I LKPEYTPYGN NNNWMYTVAP SGAAPD TY SLAYK	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         306           PVQ         366           PVE         366           A55         456           4355         456
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1 TGMAF1RHb1 TGMAF1	0 225 GT NALF I GG GA V NG EVA RLL T EE ARRHE VR. T PEOL SEH SK R VO EDLL RLAVQH PRT. I ELD TG AR S P VF AR S YG F VS 226 GT ND LF I GG GG NG GF AKT LLE EA ANHGMP. T S SGL SAH BOO I O E IM. NRAVKK PG KL VE D TG AG P VF AR S FAF VF 226 GT ND LF I GG GG NG GF AKT LLE EA EKNG I R. T S VAL SEH BOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR S FG F VF 226 GANALF I GG GGL NG GF AKT LLE EA EKNG I R. T S VAL SEH BOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR S FG F VF 226 GANALF I GG GGL NG GF AKT LLE EA EKNG I R. T YAL SEH BOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR S FG F VF 226 GANALF I GG GGL NG GF AKT LLE EA EKNG I R. T YEEL SOH BOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR S FG F VF 226 GANALF I GG GGL NG GF AKT LLE EA EKNG I R. T YEEL SOH BOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR SFG F VF 226 GANALF I GG GGL NG GF AKT LLE EA EKNG I R. T P EEL S EH SOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR SFG F VF 226 GANALF I GG GGL NG GF AKT LLE EA EKNG I R. T P EEL S EH SOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR SFG F VF 226 GANALF I GG GG NG GF AKT LLE EA EKNG I R. T P EEL S EH SOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR SFG F VF 226 GANALF I GG GG NG GF AKT LLE EA EKNG I R. T P EEL S EH SOR I QO SLL RRAVK SP GKL VE LD TG VAS P VF AR SFG F VF 226 GANALF I GG GG NG GF AKT LLE EA EKNG I R. T P EEL S EH SOR I QO SLL RRAVK SP GKL VU D TG I AS P VF AR SFG F VF 226 GT NLF I GG GG NM GE F AKT I I EE ARRNH I S. T AAEL SAH SOE LO ERLL RAD ER RPG T VE I D SG R F S RVF AR SFA Y M 205 GT NLF I GG GG NM GE F AKT I I EE ARRNR I P. T A T LS AS ES OE I O ERLL RAD ER RPG T VE I D SG R F S RVF AR SFA Y M 205 GT YG F I GG YG MM GE F ADT I I EE ARRNR I P. T A T LL S AE SOE I O ERLL RAD ER RPG T VE I D SG R F S RVF AR SFA Y M 205 GT YG F I GG YG MM GE F ADT I I EE ARRNR I P. T A T LL S AE SOE I O ERLL RAD ER RPG T VE I D SG R F S RVF AR SFA Y M 205 GT YG F I GG YG MM GE F ADT I I EE ARRNR I P. T A T L S AE SOE I O ERLL RAD ER RPG T VE I D SG R F S RVF AR SFA Y M 205 GT YG L G G YG M G F A T I I L K P YT PYG N	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         301           VVP         285           VVP         301           VVP         301           VVP         304           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         285           VP         366           PVQ         387           PVQ         386           PVQ         386           PVQ         386           PVQ         386           PVE         366           PVE         366           PVE         366           PVE         366           PVE         366           PVE
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HHGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2	0 225 GT NALEF IGG GAG V NG EV ARL T EEA RRHEV R. TPEQL SEHSKRVG EDL RLAVGHPRT LIELDTG ARSP VFARSYGF VS 1 226 GT ND LF IGG GG ND G FAKT LLEEA EKNG IR TSYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 05 GANALF IGG GG LNG OFAKT LLEEA EKNG IR TSYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 05 GANALF IGG GG LNG OFAKT LLEEA EKNG IR TSYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 16 GANALF IGG GG LNG OFAKT LLEEA EKNG IR STYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 216 GANALF IGG GG LNG OFAKT LLEEA EKNG IR STYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 216 GANALF IGG GG LNG OFAKT LLEEA EKNG IR STYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 216 GANALF IGG GG LNG OFAKT LLEEA EKNG IR STYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 216 GANALF IGG GG LNG OFAKT LLEEA EKNG IR STYALSEHS (NO GS LL RRAVKSPGKLVELDTG VASP VFARSFGFVP 2 205 GANALF IGG GG LNG OFAKT LLEEA EKNG IR STYALSEHS (NO GS LL RRAVKSPG KLVELDTG VASP VFARSFGFVP 2 205 GT NT LF IGG GG NNG FA KT LLEEA EKNG IR TPEELS (SHS (NO GS LL RRAVKSPG KLVELDTG VASP VFARSFGFVP 2 205 GT NT LF IGG GG NNG FFANT I EEA RRNH IS TA AFE SAHSOLD ER LRAD ER NO ROSS (L RRAVKSPG KLVELDTG VASP VFARSFGFVP 2 05 GT NT LF IGG GG NNG FFANT I EEA RRNH IS TA AFE SAHSOLD ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NT LF IGG GG NNG FFANT I I EEA RRNH IP TA TE LSAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NG LF IGG GV NNG FFADT I I EEA RRNH IP TA TE LSAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NG LF IGG GV NNG FFADT I I EEA RRNH IP TA TE LSAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NG LF IGG GV NNG FFADT I I EEA RRNH IP TA TE LSAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NG LF IGG GV NNG FFADT I I EEA RRNH IP TA TE LSAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NG LF IGG GV NNG FFADT I I EEA RRNH IP TA TA ELSA SAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05 GT NG LF IGG GG NNG FFANT I I EEA RRNH IP TA TE LSAESOLO ER LLRAD ER NOFT VE IDSG RFS RVFARSFAYVA 2 05	VVP         305           VVP         305           VVP         306           VVP         285           VVP         301           VVP         304           VVP         285           IVP         366           PVE         366           PVE         366           PVE         366           PVE         366           PVE         366           IVE         366           IVE         366           IVE
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 TgMAF1RHb1 TGGT1_279100 (a1) TGMAF1HhGer041b1 TGMAF1AHhGer041b1 TGMAF1CTGb1 TGMAF1RHb2 TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1 TGMAF1 TGMAF1 TGMAF1 TGMAF1 TG	0 225 GT NAL F I GG AG V NG EV ARL T TE EARRHE VR. T PEQ. SEH SKRVG E DL. RLAVQH PRT I I EL D TG AR S P VF AR SYGY S 1 226 GT ND L F I GG GG MD G F AKT L LE EA EKNG IR. T SYAL SEH SCH OG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 05 GANA L F I GG GG L NG G F AKT L EE A EKNG IR T SYAL SEH SCH OG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 05 GANA L F I GG GG L NG G F AKT L EE A EKNG IR. T SYAL SEH SCH OG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 16 GANA L F I GG GG L NG G F AKT L EE A EKNG IR. T SYAL SEH SCH OG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 16 GANA L F I GG GG L NG G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLT RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 16 GANA L F I GG GG L NG G F AKT L EE A EKNG IR. SFYAL SEH SCH TG SLT RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 24 GANA L F I GG GG L NG G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLT RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 25 GANA L F I GG GG L NG G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLT RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 26 GANA L F I GG GG M G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 26 GANA L F I GG GG M G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 26 GANA L F I GG GG M G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 26 GT NT L F I GG GG M G F AKT L EE A EKNG IR. T PEEL SGH SCH TG SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 26 GT NT L F I GG GG M G F AKT L EE A EKNG IR. T PEEL SGH SCH D C SLL RRAVKS PGK VEL D TG VA S P VF AR SFG F VF 2 26 GT YG L F I GG VG M G F ANT I I EE A RRNN IP T A TE SA ESOE I D E R LL RD A ERRPG T VEI D SG RF S R VF AR SFG Y VF 2 26 GT YG L F I GG VG M G F ANT I I EE A RRNN IP T A TE SA ESOE I D E R LL RD A ERRPG T VEI D SG RF S R YF AR SFG Y VF 2 05 GT YG L F I GG VG M M G F A DT I I EE A RRNN IP T A TE SA ESOE I D E R LL RD A ERRPG T VEI D SG RF S R YF AR SFG Y VF 2 05 GT YG L F I GG VG M G F ANT I I EE A RRNN IP T A TE SA ESOE I	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         285           VVP         285           VVP         301           VVP         301           VVP         304           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VP         285           VP         285           VP         285           VP         285           VP         285           VQ         366           PVQ         387           PVQ         386           PVQ         387           PVQ         387           PVQ         387           PVQ         387           PVQ         387           PVQ         366           PVQ         366           PVQ         367           PVQ         366           PVQ         366           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1CTGb2 HhMAF1HH62001a1 TGVE6279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMF1CTGb2 HhMAF1HhGer041a1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041a1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HHGer041a1 TGMAF1CTGb2 HhMAF1HHGer041a1 TGMAF1CTGb2 HhMAF1HHGer041a1 TGMAF1RHb1 TGMAF1RHb0 TGMAF1RHb0 TGMAF1RHb0 TGMAF1RHb0 TGMAF1RHb1 TGMAF1RHb0 TGMAF1RHb1	225 GT NALEFIGG GAG VNG EVARL TE EAARRHEVR TPEOLSEHSKRVDEDL RLAVGHPRT LIELDTG ARSPVFARSYGFVS 226 GT ND LFIGG GG MD GEFAKT LLEEA EKNOTR TSYALSEHSOR 10 GEI MN NRAVKRAGK VELDTG VASPVFARSFGFVP 205 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TSYALSEHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 216 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TSYALSEHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2216 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TSYALSEHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2216 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TSYALSEHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2216 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TPEELSOHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2216 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TPEELSOHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2226 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TPEELSOHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2236 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TPEELSOHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2236 GANALFIGG GG LNG OFAKT LLEEA EKNOTR TPEELSOHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2236 GANALFIGG GG LNG OFAKT LEEA EKNOTR TPEELSOHSOR 10 GS LL RRAVKSPGK VELDTG VASPVFARSFGFVP 2305 GT NT LFIG GG GMN GEFANT 11 EEA RRNN IS TA AFE SAHSOELO ER LLRD AERROGT VELDSG SSPVFARSFAYVA 2405 GT NT LFIG GG GMN GEFANT 11 EEA RRNN IP TA TA LSAE SOH 10 GS LL RRAVKSPGK VELDTG VASPVFARSFAYVA 2505 GT NT LFIG GG GMN GEFANT 11 EEA RRNN IP TA TA LSAE SOH 10 ER LLRD AERROGT VEIDSG RFSRVFARSFAYVA 2505 GT NG LFIG GG VGMN GEFANT 11 EEA RRNN IP TA TA LSAE SOH 10 ER LLRD AERROGT VEIDSG RFSRVFARSFAYVA 2505 GT NG LFIG GG VGMN GEFANT 11 EEA RRNN IP TA THE SAE SOH 10 ER LLRD AERROGT VEIDSG RFSRVFARSFAYVA 2505 GT NG LFIG GG VGMN GEFANT 11 EEA RRNN IP TA THE SAE SOA 10 ER LLRD AERROGT VEIDSG RFSRVFARSFAYVA 2505 GT NG LFIG GG VGMN GEFANT 11 EEA RRNN IP TA TA LSAE SOA 10 ER LLRD AERROGT VEIDSG RFSRVFARSFAYVA 2505 GT VG LFIG GG VG VG TFIHLLK PEVT PY GD PKN NNW YT VAP SG AAP DT TY SLAVKT 11 AGV 16 AAAA NN DT PAGO Y 2505 GT VG LFIG GG VG VG TFIHLLK PEVT PY GD PKN NNW YT VAP SG AAP DT TY SLAVKT 11 AGV 16 AAAA NN DT PAGO	VVP         305           VVP         305           VVP         306           VVP         285           VVP         285           VVP         304           VVP         285           IVP         285           PVQ         366           PVQ         395           PVQ         366           PVE         366           PVE         366           PVE         366           4355         4456           435         4461           455         4456           455         4456
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 TgMAF1CTGb1 TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1RHb2 TGMAF1CTGb1 TGMAF1RHb2 TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1 TGMAF1RHb2 TGMAF1 TGMAF	225 GT NALF I GG GA VING EV ARLL TE EAR RHEVRL TPEQL SEH SKRV GEDL RLAVOH RT LIELD TG AR EPVEARS YG F VS 226 GT NDLF I GG GG MOG F AKTLL EE ACKHOMPL TSSGL SAH BOQ I O EI ML NRAVKK PG KLVE VD TG AG SPVF ARSF AF V 226 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTS YAL SEH GOR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 205 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTS YAL SEH GOR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 216 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTS YAL SEH GOR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 216 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTS YAL SEH GOR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 216 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTS YAL SEH GOR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 217 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTS FVAL SEH GOR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 218 GA NALF I GG GG LNOG F AKTLL EE ACKNO I RLTPEL SEH SGR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 229 GA NALF I GG GG MOG F AKTLL EE ACKNO I RLTPEL SEH SGR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 230 GA NALF I GG GG MOG F AKTLL EE ACKNO I RLTPEL SEH SGR I O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 230 GA NALF I GG GG MOG F AKTLL EE ACKNO I RLTPEL SANS GUI O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 230 GT NTLF I GG GG MOG F AKTLL EE ACKNO I RLTPEL SANS GUI O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF GF V 240 GT TYGL F I GG GG MOG F AKTLL EE ARKNN I PLTA TOL SANS GUI O GSLL RRAVKS PG KLVE D TG VAS PVF ARSF AFY M 240 GT YGL F I GG YG MOG F AKTLL EE ARKNN I PLTA TE SAE SGE I O ER L RD AE RR GSLVE I D SG RF SRVF ARSF AFY M 240 GT YGL F I GG YG MMG EF ADTT I EE ARRNR I PLTA TE SAE SGE I O ER L RD AE RR GSLVE I D SG RF SRVF ARSF AFY M 240 GT YGL F I GG YG MMG EF ADTT I I E ARRNR I PLTA TE SAE SGE I O ER L RD AE RR GSLVE I D SG RF SRVF ARSF AFY M 240 GT YGL F I GG YG MMG EF ADTT I I E ARRNR I PLTA TE SAE SGE I O ER L RD AE RR GSLVE I D SG RF SRVF ARSF AFY M 240 GT YGL F I GG YG MMG EF ADTT I I E ARRNR I PLTA TE SAE SGE I O ER L RD AE RR GSLVE I D SG RF SRVF AR	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         285           VVP         285           VVP         301           VVP         301           VVP         304           VVP         301           VVP         302           PVQ         366           PVQ         366           PVQ         367           PVQ         367           PVQ         366           PVQ         367           PVQ         366           PVQ         366           PVQ         366           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGVE6_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMF1RHa7 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041a1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1CTGb2 HhMAF1HhGer041a1 TGGT1_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb0 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb1 TGMAF1CTGb2	225 G T NAL T I GGAG VING EV RIL I T E ARRHEVRL TPEQ. SEH KRY V EDI. RLAVOH HAT LIELD TGAR BY VIARS YGF V 226 G T NOL F I GGGG NOG F AKTLL EE ACKNOR RL SYAL SEHSOR I OQSL LRAVKSPGKL VE DT GAG SPVF ARSFGF V 205 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL SYAL SEHSOR I OQSL LRAVKSPGKL VE DT GVAS PVF ARSFGF V 205 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL SYAL SEHSOR I OQSL LRAVKSPGKL VE DT GVAS PVF ARSFGF V 216 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL SYAL SEHSOR I OQSL LRAVKSPGKL VE DT GVAS PVF ARSFGF V 217 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL TY VAL SGHSOR I OQSL LRAVKSPGKL VE DT GVAS PVF ARSFGF V 218 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL SYVAL SEHSOR I OQSL LRAVKSPGKL VE DT GVAS PVF ARSFGF V 224 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL SYVAL SEH SOR I OQSL LRAVKSPGKL VE DT GVAS PVF ARSFGF V 224 GANAL F I GGGGL NOG F AKTLL EE AKKNOF RL SYVAL SEH SOR I OQSL LRAVKSPGKL VU DT G I AS PVF ARSFGF V 225 G T N LF I GG GGN GG F AKTLL EE AKKNOF RL SPE SH SOR I OQSL LRAVKSPGKL VU DT G I AS PVF ARSFGF V 226 G T N LF I GG GGN GG F AKTLL EE AKKNOF RL SPE SH SOR I OQSL LRAVKSPGKL VU DT G I AS PVF ARSFGF V 226 G T N LF I GG GGN GF A KTLL EE AKKNOF NUT T P L T AT L SAESOE I OER L RDAERNGS V U DT G I AS PVF ARSFGF V 205 GT Y N LF I GG GGNNG E F ANT I I EE ARRNNF I P LT AT L SAESOE I OER L RDAERNGS V U DT G I AS PVF ARSFGF V 205 GT Y GL F I GG Y GMNG E F ADT I I EE ARRNNF I P LT AT L SAESOE I OER L RDAERNGS I VU D SGF S RYF ARSFA Y W 205 GT Y GL FI GGY VMNG E F ADT I I EE ARRNNF I P LT AT L SAESOE I OER L RDAERNGS I VU E D SGRF SRYF ARSFA Y W 205 GT Y GL FI GGY VMNG E F ADT I I EE ARRNNF I P LT AT L SAESOE I OER L RDAERNGS I VU E D SGRF SRYF ARSFA Y W 205 GT Y GL FI GGY VMNG E F ADT I I EE ARRNNF I P LT AT L SAESOE I OER L HDAERNG ST VU E D SGRF SRYF ARSFA Y W 205 GT Y GL FI GGY VMNG E F ADT I I EE ARRNNF I P LT AT L SAESOE I OER L HDAERNG ST VU E D SGRF SRYF ARSFA Y W 205 GT Y G L R W K SK VG ANVG Y T I H L K P VT P Y ON L NNNWY Y VY P S GAAP D T Y SL Y KTI I ASV Y GAAAAN ND T A SO Y 206 GT Y G L R W SK Y G	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         301           VVP         285           VVP         301           VVP         301           VVP         301           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         285           VVP         285           VVP         285           VVP         285           VVP         366           PVQ         366           PVQ         366           PVE         366           PVE         366           425         446           455         446           456         455           4466         451
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 TgMAF1RHb1 TGGT1_279100 (a1) TGME49_279100 (a1) TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHC30 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHC30 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHC30 TGMAF1RHb1 TGMAF1RHC30 TGMAF1	225 G T NAL F I GG G V NG E Y M RL I T E FAR HE Y RL T PEOL SEH KR V O E DI L RL AVOH R TL I EL D T G AR B Y Y A RS Y G F M S 226 G T NOL F I GG G G N O G F AK TL L E FARKNG I RL T S Y AL SEH SQ I O G L I RL N RA Y KK B G K L Y E D T G Y A S P Y A RS F G F M Y 206 G ANA F I G G G L NO G F AK TL L E FARKNG I RL T S Y AL SEH SQ I O G L R RA Y KS P G K L Y E D T G Y A S P Y A RS F G F M Y 207 G ANA F I G G G L NO G F AK TL L E FARKNG I RL T S Y AL SEH SQ I O G L R RA Y KS P G K L Y E D T G Y A S P Y A RS F G F M Y 207 G ANA F I G G G G L NO G F AK TL L E FARKNG I RL T S Y AL SEH SQ I O G L R RA Y KS P G K L Y E D T G Y A S P Y A RS F G F M Y 207 G ANA F I G G G G L NO G F AK TL L E FARKNG I RL T PEEL S G H SQ RI O G S L R RA Y KS P G K L Y E D T G Y A S P Y A RS F G F M Y 207 G ANA F I G G G G L NO G F AK TL L E FARKNG I R L F Y AL S E H S Q I O G S L R RA Y KS P G K L Y E D T G Y S P Y A RS F G F M Y 208 G ANA F I G G G G L NO G F AK TL L E FARKNG I R L F Y AL S E H S Q I O G S L R RA Y KS P G K L Y E D T G Y S P Y A RS F G F M Y 206 G ANA F I G G G G L NO G F AK TL L E FARKNG I R L T P E E S G H O G R I O G S L R RA Y KS P G K L Y U D T G I A S P Y A RS F G F M Y 208 G T N R F I G G G G M O G F AK TL L E FARKNG I S L T FAR L S A H O G L O E R L R D A E R M G S L V E I D S G F S P Y A RS F G F M Y 208 G T N R F I G G G G M N G F AN T I I E FARRNH I P L T A T L S A E S Q E I O E R L R D A E R M G S L V E I D S G F S RY F A RS F A Y A 208 G T N R F I G G G G M N G F A NT I I E FARRNH I P L T A T L S A E S Q E I O E R L R D A E R M G T L U E J S G R S R Y A RS F A Y A 209 G T Y G L T G G V G M N G F A D T I I E FARRNH I P L T A T L S A E S Q E I O E R L R D A E R M G T L U E S G R S R Y A RS F A Y A 200 G T N R F I G G G G M N G F A D T I I E FARRNH I P L T A T L S A E S Q E I O E R L H D A E R O G T L Y E I S G R S R Y A RS F A Y A 200 G T N R F I G G G G M N G F A D T I I E A R R N N I P L T A T L S A E S Q E I O E R L H D A E R O G T L Y E I S G R S R Y A RS F A Y A 200 G T N R F I G G G G M N G F	VVP         305           VVP         305           VVP         206           VVP         285           VVP         285           VVP         285           VVP         206           VVP         301           VVP         301           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         301           VP         285           PVQ         366           PVQ         366           PVQ         367           PVQ         367           PVQ         366           PVQ         367           PVQ         367           PVQ         367           PVQ         366           PVQ         366           PVQ         366           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGMAF1HhGer041a1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TgMAF1CTGb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1RHb2 TGMAF1CTGb2 TGMAF1CTGb2 HhMAF1HhGer041a1 TGMAF1CTGb2 TGMAF1 TGMA	225 G T NA F I GG AG YNG E Y AR LLT EE ARRHE YR T PEO. S EH K RYDE DL R LYAOH PRT I I EL DTG AR B Y YF AR S YCF W 226 G T ND F I GG GG MO G F AKT LLE E ACKNG IR LT SYAL S EH G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 206 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LT SYAL S EH G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 216 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LT YAL S EH G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 212 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LTYPE I S G H G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 212 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LTYPE I S G H G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 212 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LTYPE I S G H G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 212 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LTYPE I S G H G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 213 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LTPE LS GH G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 214 G ANA F I GG GG LNG G A KT LLE E ACKNG IR LTPE LS GH G RI DOG LL RRAYKS PGK. VEL DTG YA B YYF AR SFG F W 215 G T NT F I G G G G M G F AKT I LE E ACKNG IR LTPE LS GH G RU L RD A KNS PGK. VYL DTG I A PYF AR SFG F W 216 G T NT F I G G G G M G F AKT I LE E ARKNH IS LTA AE S AN G EL D EN L RD A KNS PGK. VYL DTG I A B YYF AR SFG F W 216 G T NT F I G G G G M G F AKT I LE E ARKNH IS LTA AE S A SO EL D EN L RD A KNS PGK. VYL D TG I A B YYF AR SFG F W 216 G T NT F I G G G G M G F AKT I LE E ARKNH IS LTA AE S A SO EL D EN L RD A KNS PGK. VYL D TG I A B YYF AR SFG F W 2105 G T NT F I G G G G M G F AKT I LE E ARKNH IS LTA AE SO EL D EN L RD A KNS PGK. VYL D TG I A B YF AR SFG F W 2105 G T NG F I G G Y G M G F A TT I LE A RRNN IP L TA TE SA E SO EL D EN L RD A KNS PGK. VYL D TG I A B YF A R SFG F W 2105 G T NG L W K S KYG A NYG T F I H L K P YT PYG N NAWY Y W A T SO A PD T Y SL A X XT I A SY I I A SY I A	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         301           VVP         285           VVP         301           VP         285           VP         285           VP         285           VP         285           VP         285           VP         285           VP         366           PVQ         366           PVQ         397           PVQ         366           PVE         366           PVE         366           PVE         366           PVE         366           4355         4466           456         4454           456         4454           456         4455           4456
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb2 TgMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGMAF1AHhGer041b1 TGMAF1HhGer041b1 TGMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1 TGMAF	225 GTNALFIGGAG VIGEVARLETEEARRHEVRETPEO. SEH KRVDEDLERLVAV VOHPRTEIED TGARBPVFARSYGFV 226 GTNDLFIGGGOMDDEFAKTLLEEARRNGIRLTSVALSENSOR JOGSLERAVVSPGKLVEDTGVABPVFARSFGFV 226 GANALFIGGGGLNGOFAKTLLEEAKNGIRLTSVALSENSOR JOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 226 GANALFIGGGGLNGOFAKTLLEAKNGIRLTSVALSENSOR JOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 226 GANALFIGGGGLNGOFAKTLLEAKNGIRLTSVALSENSOR JOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 227 GANAFFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 228 GANALFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 229 GANALFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 220 GANAFFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 220 GANAFFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 220 GANAFFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 226 GANAFFIGGGGLNGOFAKTLLEAKNGIRLTYEELSGHBGRIOGSLERAVVSPGKLVELDTGVABPVFARSFGFV 226 GANAFFIGGGGUNGFFANTIIEEARNNIPLTATELSAESGEIDERLLFDAERRPGSLVELDSGFSPVFARSFAVA 206 GTNRFFIGGGGUNGFFANTIIEEARNNIPLTATELSAESGEIDERLLHDAERRPGSLVELDSGRFSRVFARSFAVA 206 GTNRFFIGGGGUNGFFANTIIEEARNNIPLTATELSAESGEIDERLLHDAERRPGSLVEIDSGRFSRVFARSFAVA 207 GTVKFFYGGVGVNGFFADTTIEEARRNNIPLTATELSAESGEIDERLLHDAEROFGTLVEIDSGRFSRVFARSFAVA 208 GTNRFFIGGVGVNGFFADTTIEEARRNNIPLTATELSAESGEIDERLLHDAEROFGTLVEIDSGRFSRVFARSFAVA 209 GTGVDESOTGRVGATFIHILKPEVTPYONLNNVMYTVAPSGNAPDASVSVYGATIKGVGAVSFFRRVFARSFAVA 209 GTGVDESOTGRVGATVINLENPEVTPYONLNNVMYTVAPSGNAPDASVSVYGATIKGVGAVSFFNRTFEEGO 208 GLMWKSSKVGANQVTFIHILKPEVTPYONLNNVMYTVAPSGNAPDASVSVYGATIKGVGAVSFFNRTFEEGO 208 GLMWKSSKVGANVGVTFIHILKPEVTPYONLNNVMYTVAPSGNAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOOY 208 GLMWKSSKVGANVGVTFIHILKPEVTPYONLNNVMYTVAPSGNAPDTYSLAVKTTIAGVIGAAAAYNDTPAGOOY 207 GTWKESEVGSNVGTFVHILKPEVTPYONLNNVMYTVAPSGNAPTTYSLAVKTTIAGVIGAAAAYNDTPAGOOY 208 GTVGDESTGRVVGATGVTFIHILKPEVTPYONLNNVMYTVAPSGNAPTTYSLAVKTTIAGVIGAAAAYNDTPAGOOY 207 GTWKESEVGSNVGTFVHILKPEVTPYONLNNVMYTVAPSGNAPTTYSLAVKTTIAGVIGAAAAYNDTPAGOOY 208 GTVGVESVGNVGTFVHI	VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         285           VVP         285           VVP         285           VVP         301           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         301           VP         285           PVQ         366           PVQ         367           PVQ         366           PVQ         366           PVQ
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) TGME49_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1CTGb2 TGMAF1 TGMF	225 GTNALF IGG AG VNG EVARLET E ARRHEVRET PEOLS EH KRV0EDLERLING VAPRTEI ELDTGARSPVEARSYGFV 236 GTND_FIGG GUNDGEFAKTLLEE ARRHEVRET PEOLS EH KRV0EDLERLING VKKPGKLVED TGARSPVEARSYGFV 236 GANALFIGG GUNDGEKTLLEE AKKNORTSVALSENSOR (QGSLERAVKSPGKLVED TGVASPVEARSYGFVE 236 GANALFIGG GUNDGEKTLLEE AKKNORTSVALSENSOR (QGSLERAVKSPGKLVELDTGVASPVEARSYGFVE 236 GANALFIGG GUNDGEKTLLEE AKKNORTSVALSENSOR (QGSLERAVKSPGKLVELDTGVASPVEARSYGFVE 237 GANAFFIGG GUNDGEKTLLEE AKKNORTSVALSENSOR (QGSLERAVKSPGKLVELDTGVASPVEARSYGFVE 238 GANALFIGG GUNDGEKTLLEE AKKNORTSVALSENSOR (QGSLERAVKSPGKLVELDTGVASPVEARSYGFVE 239 GANALFIGG GUNDGEKTLLEE AKKNORTFFE SON (QGSLERAVKSPGKLVELDTGVASPVEARSYGFVEARSYGFVE 240 GANAFFIGG GUNDGEKTLLEE AKKNORTFFE SON (QGSLERAVKSPGKLVELDTGVASPVEARSYGFVEA	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         301           VVP         285           VVP         301           VVP         301           VVP         301           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         285           VP         285           VP         285           VP         285           VP         285           VP         366           PVQ         366           PVQ         367           PVQ         367           PVQ         366           PVE         366           4355         4466           455
NCLIV_004730 HhMAF1HhGer041b1 TgMAF1RHb0 TgMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TgMAF1CTGb1 TgMAF1CTGb2 HhMAF1HhGer041a1 TGVEG_279100 (a1) TGMAF1RHa2 TGMAF1RHa1 TGGT1_279100 (a1) NCLIV_004730 HhMAF1HhGer041b1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb1 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1RHb2 TGMAF1CTGb1 TGMAF1CTGb1 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CTGb1 TGMAF1CTGb2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CCG2 TGMAF1CC	225 GTNALF IGG AG VNG EVARLLTEE ARRHEVRIT PEOLS EH KRVGEDL RLAVADHPRT LIELDTGAR PVFARSYGFV 226 GTND FIGG GGUND GFAKTLLEE ARRHGIRLTSVALSENSOR I QSLLRAVKSPGK. VED TGAG PVFARSYGFV 226 GANALF I GG GGUND GFAKTLLEE AEKKNGIRLTSVALSENSOR I QSLLRAVKSPGK. VED TGAG PVFARSYGFV 226 GANALF I GG GGUND GFAKTLLEE AEKKNGIRLTSVALSENSOR I QSLLRAVKSPGK. VED TGVA PVFARSYGFV 227 GANALF I GG GGUND GFAKTLLEE AEKKNGIRLTSVALSENSOR I QSLLRAVKSPGK. VED TGVA PVFARSYGFV 228 GANALF I GG GGUND GFAKTLLEE AEKKNGIRLTSVALSENSOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 229 GANALF I GG GGUND GFAKTLLEE AEKKNGIRLTSVALSENSOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 220 GANALF I GG GGUND GFAKTLLEE AEKKNGIRLTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 220 GANALF I GG GGUND GFAKTLLEE AEKKNGIR LTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 221 GANALF I GG GGUND GFAKTLLEE AEKKNGIR LTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 222 GANALF I GG GGUND GFAKTLLEE AEKKNGIR LTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 223 GANA FI GG GGUND GFAKTLLEE AEKKNGIR LTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 236 GTNR FI GG GGUND GFAKTLLEE AEKKNGIR LTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 236 GTNR FI GG GGUND GFAKTLLEE AEKKNGIR LTPEELSGNBOR I QG LLRAVKSPGK. VED TGVA PVFARSYGFV 236 GTNR FI GG GGUND GFAKTLLEE AEKKNGIR LTALESSA 236 GTNR FI GG GGUND GFAKT I LEARNNRIP LTATELSA 236 GTNR FI GG GGUND GFAKT I LEARNRIP LTATELSA 237 GTNR FI GG GGUND GFAKT I LEARNRIP	VVP         305           VVP         305           VVP         305           VVP         285           VVP         285           VVP         285           VVP         285           VVP         301           VVP         301           VVP         304           VVP         304           VVP         301           VVP         301           VVP         301           VVP         301           VVP         285           PVQ         366           PVQ         385           PVQ         385           PVQ         385           PVQ         385           PVQ         366           PVQ         366           PVQ         366           PVQ         365           4355

Figure 2-5. Alignment of select MAF1 sequences from *T. gondii*, *H. hammondi*, and *N. caninum*.

Alignments were performed using CLUSTAL-Omega, and visualized using JalView. Residues are colored by percent identity. Sequences in bold were used in complementation experiments in the present study. Portions of proteins used as antigens for antibody production are indicated by red (TgMAF1RHb1) and black (TgMAF1RHa1) boxes.

#### 2.2.3 *MAF1a* and *MAF1b* gene families have distinct inter-strain transcriptional profiles

We reported previously that MAF1 transcript levels were of lower abundance in a type II T. gondii strain (ME49) compared to Types I and III (RH and CTG, respectively;[106, 160]). The previously reported data were derived from spotted cDNA microarray experiments which would not distinguish transcripts for MAF1a and b paralogs. Therefore we used the sequence alignments shown in Figure 2-4 to determine if probes for the MAF1a and b paralog families could be found on the T. gondii Affymetrix array [161] and therefore could be used to assess paralog-specific transcript levels. MAF1a was most similar to TGME49\_279100 and the MAF1b family was most similar to TGME49\_220950, respectively (Figure 2-4B; www.toxodb.org). Similar to what was reported previously ([106] and Figure 2-6A), we found that MAF1b transcript levels were lower in Type II strains (ME49 and PRU; www.toxodb.org and Figure 2-6A,B) compared to Type I strains (GT1 and RH) and Type III strains (CTG and VEG; Figure 2-6A,B). In contrast, we found that transcript levels for TGME49\_279100 (MAF1a) were of comparatively high abundance (>90<sup>th</sup> Percentile; data not shown) across all 6 queried T. gondii strain types (Figure 2-6B). These data indicate that MAF1a and MAF1b have significantly diverged in terms of their transcript abundance in the Type II T. gondii lineage.





A) Spotted cDNA microarray (which would not distinguish MAF1a from MAF1b transcript) data illustrating reduced MAF1a/b transcript levels in ME49 compared to RH and CTG. Two replicates per strain type are shown.
B) Affymetrix microarray data (derived from probes that were chosen to be unique to each gene, which therefore would distinguish between MAF1a and b transcript levels) illustrating similar expression of MAF1a across multiple *T. gondii* strains, and reduced MAF1b transcript levels in Type II *T. gondii* (Pru, ME49). Data downloaded from.

C) Levels of MAF1b protein were compared among 2 strains each from the 3 predominant lineages of *T. gondii* using polyclonal antibodies against the C-terminus of TgMAF1RHb1. Expression polymorphism of MAF1 correlates with strain-specificity of the host mitochondrial association phenotype. SAG1 is used as a loading control. D) Densitometric analysis of relative levels of MAF1 in the six strains examined based on the MAF1/SAG1 intensity ratio. GT1 was set to 1. E) Isolation of paralog-specific polyclonal mouse antisera against MAF1a1 and b1. Three mice were exposed to purified MAF1RHb1 and 2 to purified MAF1RHa1 and used to probe blots containing the immunizing antigen. All but mouse 2 had antibodies that were fully specific to the input antigen.

## 2.2.4 MAF1a and MAF1b protein expression differs between T. gondii strains

We previously demonstrated a lack of MAF1 protein expression in TgME49 compared to TgRH and TgCTG using a polyclonal anti-MAF1 antibody raised against the C-terminus of TgMAF1RHb1 [106]. Using this same antibody, we compared MAF1 protein expression in the 6 T. gondii strains examined in the Southern blot analysis and observed MAF1b expression in the Type I and III strains, but did not detect any MAF1b protein in either of the Type II strains (Figure 2-6C,D). Additionally, we observed that MAF1b protein from both Type III strains had a slightly higher apparent molecular weight compared to those in the Type I strains (Figures 2-7B) and 2-6C). This is consistent with the observation that the 2 clones of MAF1b sequenced from the CTG strain encode either 4 or 6 P{4:7}S repeat motifs, while the highest number of P{4:7}S repeat motifs in RH was 3 (Figures 2-4B and 2-5). To determine if Types I, II, and III all express a MAF1a isoform, we generated new polyclonal antibodies against the C-terminus of TgMAF1RHa1 (Ser173 to Ser443) or TgMAF1RHb1 (Thr159 to Asp435) (indicated in Figure 2-5). We exposed 2 mice to the TgMAF1RHa1 and 3 mice to TgMAF1RHb1. Polyclonal serum from 4 of the 5 mice was specific for the input antigen, while 1 mouse exposed to TgMAF1RHb1 harbored antibodies that bound to both MAF1 paralogs (Figure 2-6E and data not

shown). Given the amount of similarity between the two antigens, it is likely that the epitopes recognized by sera from most of the mice were derived from the dissimilar regions. Additionally it is likely that each polyclonal serum is capable of recognizing multiple "a" or "b" paralogs. In Western blots against the input antigen, a higher molecular weight band is detected by all of these antisera in addition to the major species at the expected molecular weight (Figure 2-6E). The antisera recognize the higher MW band with similar specificity as the purified protein. This higher MW band may be a dimer of the purified protein, as it is approximately twice the size of the major species, can be seen upon Coomassie staining, and its quantity is reduced after longer boiling times of the purified protein (Figure 2-7). Using antibodies from mouse 5 for immunofluorescence, we detected MAF1a protein in all three strains, while once again we did not detect MAF1b in type II when using antibodies from mouse 1 (Figure 2-8A). We also saw a similar pattern of expression by Western blot (Figure 2-8B). The specificity of the MAF1b antiserum to MAF1b and not MAF1a was further confirmed by the fact that the MAF1b antiserum bound to Type II T. gondii when expressing an ectopic copy of TgMAF1RHb1 (Figure 2-9). These data indicate significant strain-specific variation between major clonotypes in both MAF1 protein level and in the qualitative nature of the paralogs that are expressed.





A) Coomassie staining of purified TgMAF1RHa1 C-terminus (Ser173 to Ser443) or TgMAF1RHb1 (Thr159 to Asp435) where the sample was boiled for 10 minutes (lanes 1 and 2) or 30 minutes (lanes 3 and 4) before being run on a gel. The higher MW band is reduced (TgMAF1RHa) or eliminated (TgMAF1RHb1) after 30' of boiling compared to 10'. B) Western blot using mouse 5 antiserum against purified TgMAF1Rha1 C-terminus boiled for 10 or 30 minutes before loading the gel. There was no higher MW band observed in this blot, which was run on the same gel as C. C) Western blot using mouse 1 antiserum against purified TgMAF1RHb1 C-terminus boiled for 10 or 30 minutes. The higher MW band is reduced after 30' compared to 10. Additional break down of the sample is also observed (lower MW band).



Figure 2-8. T. gondii MAF1 paralog expression differs between lineages.

Polyclonal antibodies were generated specifically against the C-termini of TgMAF1RHa1 or TgMAF1RHb1. (A) Protein expression was compared by immunofluorescence across 3 strains representing clonotypes I, II, and III. Antibodies against TgMAF1RHa1 detected protein in all 3 strains, while antibodies against TgMAF1RHb1 detected protein on in RH and CTG (not ME4). (B) Expression of paralogs confirmed by western blot, blots were stripped and probed with antibodies against ROP5 as a loading control.



**Figure 2-9. Transgenic expression of TgMAF1RHb1 in ME49 detected by paralog-specific antibodies.** Transgenic TgMAF1RHb1 expression in a TgME49 background detected using the TgMAF1RHb1-specific antibody compared to expression of empty-vector transfected TgME49. Expression was detected only in parasites transfected with a plasmid containing *TgMAF1RHb1*.

# 2.2.5 *T. gondii MAF1* paralogs differ in their ability to mediate host mitochondrial association in *T. gondii* and *N. caninum*.

Host mitochondrial association (HMA) is a strain-specific trait in *T. gondii* (lacking in Type II stains; Figure 2-10A), and this trait is consistent with reduced MAF1b transcript and protein levels in members of the Type II lineage (Figures 2-8, 2-6A-D). In contrast the *MAF1a* gene is highly expressed at the transcript and protein level equally well across multiple *T. gondii* strains. To determine if the *MAF1a* and *b* genes differed in their ability to confer HMA in HMA<sup>-</sup> parasites, we generated N-terminally HA-tagged clones of the 2 paralogs that differed in the absence or presence of the P{4:7}S motif (*TgMAF1RHa1* and *TgMAF1RHb1*, respectively). To do this we cloned *TgMAF1RHa1* in place of *TgMAF1RHb1*, while retaining the *TgMAF1RHb1* promoter in the construct to ensure equal expression between paralogs. We expressed these

genes in both a Type II strain (TgME49) and in *N. caninum* (NC-1;[157]) and used confocal microscopy and mitochondrial staining to determine the impact on HMA. TgMAF1RHb1 expression was sufficient to mediate HMA in *T. gondii* strain ME49 (Figure 2-11A) and also in *N. caninum* (Figure 2-11B) 18 h post-infection. In contrast, TgMAF1RHa1 was unable to mediate HMA in either TgME49 or *N. caninum* although its protein localization profile was similar to that of MAF1RHb1 (Figure 2-11A,B, bottom panels). We also generated clones of TgME49:MAF1RHb1 and NC-1:MAF1RHb1 for electron microscopy. Both wildtype TgME49 and NC-1 have little, if any, HMA (Figure 2-11C, left panels). However, when these strains express MAF1RHb1 they become HMA<sup>+</sup> and there is an increase in host mitochondria directly adjacent to the PVM (Figure 2-11C, right panels).



Figure 2-10. Host mitochondrial association is a feature of *T. gondii* and *H. hammondi* infections, but not *N. caninum*.

(A) NRK-mitoRFP cells were infected with GFP-expressing Type I, II, and III (RH, PRU, CTG) parasites. Type II parasites are HMA<sup>-</sup>, while Types I and III are HMA<sup>+</sup>. (B) NRK-mitoRFP cells were infected with *N. caninum* strain NC-1. Cells were fixed and counterstained with Hoechst stain. Wild type *N. caninum* are HMA<sup>-</sup>. (C) HFFs were

infected with *H. hammondi* sporozoites for 8 days before fixation. Host-mitochondria were visualized using an antibody to human MTCO2. *H. hammondi* is HMA<sup>+</sup>.



Figure 2-11. MAF1RHa1 and MAF1RHb1 differ in their ability to complement HMA in *T. gondü* and *N. caninum*.

(A) HFFs were labeled with MitoTracker and infected with parasites transiently transfected with either HA-MAF1RHa1 or HA-MAF1RHb1. MAF1RHb1 but not MAF1RHa1 is able to confer the HMA phenotype in TgME49. (B) Identical results were obtained for *N. caninum*. Scale bar, 5.0 μm. (C) HA-MAF1RHb1 was transfected into either TgME49 (top panels) or *N. caninum* (bottom panels), and HA-positive clones were isolated by limiting dilution. Wild type (WT, left panels) and TgMAF1RHb1-complemented (right panels) were grown for 18 h in HFFs and processed for electron microscopy. Asterisks indicate host mitochondria. Scale bar: 500 nm.

## 2.2.6 MAF1b1 from *T. gondii* and *H. hammondi* can confer the HMA phenotype in type II *T. gondii*, while TgMAF1b0 and HhMAF1a1 cannot

HMA is greatly reduced in the closely related *N. caninum* (Figure 2-10B;[162]), but the HMA phenotype of the nearest extant relative of *T. gondii*, *H. hammondi*, is unknown. To test this we assessed HMA in sporozoite-derived tachyzoites of *Hammondia hammondi* (strain HhCatEth1; [163]), and found clear evidence for HMA in this species (Figure 2-10C). Therefore we hypothesized that *T. gondii* and *H. hammondi* would harbor *MAF1* paralogs that could complement the HMA defect in Type II *T. gondii*, while *N. caninum* would not. To test this hypothesis we cloned N-terminally tagged *MAF1* paralogs from *T. gondii*, *H. hammondi* and *N. caninum*. For *T. gondii*, the coding sequences for TgMAF1RHb0 and TgMAF1RHb1 with the endogenous promoters were cloned directly from RH strain genomic DNA. Similar constructs were made for *H. hammondi* and *N. caninum*. Each construct was transfected into the HMA<sup>-</sup> TgME49 strain and the ability of each isoform to mediate HMA was assessed by immunofluorescence. Similar to our results with TgMAF1RHa1, TgMAF1RHb0 was unable to mediate HMA (Figure 2-12A), indicating that not all "b" paralogs are capable of mediating this

phenotype. Importantly, the same was true for HhMAF1a1: when transfected into Type II *T. gondii* this protein did not confer the HMA phenotype, although it did have an localization profile that was distinct from other MAF1 paralogs (Figure 2-12C). In contrast, both TgMAF1RHb1 (as shown previously) and HhMAF1b1 could confer the HMA phenotype when ectopically expressed in Type II *T. gondii* (Figure 2-12B,D). We quantified percent vacuole coverage for 20 vacuoles for each MAF1 paralog using confocal microscopy and found that parasites expressing TgMAF1RHb1 or HhMAF1b1 had significantly more vacuole membrane associated with host mitochondria than wildtype Type II parasites (Figure 2-12E). The localization of TgMAF1RHb0, TgMAF1RHb1 and HhMAF1b are all similar. We also performed the same experiment with NcMAF1 (based on NCLIV\_004730), but we were unable to detect any protein following multiple (>3) transfections. Whether this is due to upstream regulatory sequences or some other species-specific factor is unknown.



Figure 2-12. *T. gondii* and *H. hammondi* harbor *MAF1* isoforms that differ in their ability to mediate HMA. (A) N-terminally HA-tagged MAF1 isoforms were expressed in TgME49 parasites and HMA was assessed using MitoTracker or immunofluorescence assay using antibodies against the mitochondrial marker MTCO2.

TgMAF1RHb0 and HhMAF1a1 did not mediate HMA, while TgMAF1RHb1 and HhMAF1b1 are both able to mediate HMA. (B) Quantification of percent vacuole coverage, determined by confocal microscopy. Twenty vacuoles were quantified for each of the MAF1 paralog indicated, as well as wild-type TgME49.  $\chi^2$  *P*-values: \*0.0144; \*\*0.0005; \*\*\*<0.0001.

# 2.2.7 Expression of TgMAF1RHb1, but not TgMAF1RHa1, in type II *T. gondü* increases competitive advantage during infection

In order to directly examine the impact of MAF1 in an *in vivo* infection system, we infected Balb/c mice with TgME49 wild-type or a TgMAF1RHb1-complemented line and measured their rates of proliferation and dissemination *in vivo* using bioluminescence imaging [164]. We used a sub-lethal dose for a Type II strain (100 tachyzoites; [165]) to allow the mice to survive the full course of the infection and enable us to detect any subtle differences in parasite dissemination. We observed marginally higher, but statistically insignificant, parasite burdens in infection with TgME49:TgMAF1RHb1 (Figure 2-13). Given the marginally higher parasite burden observed in mice infected with TgME49:TgMAF1RHb1 compared to wildtype TgME49, we hypothesized that this increase in parasite growth, although marginal, could provide a competitive advantage during an infection with a mixed population. To test this hypothesis we created mixed populations of TgME49 and TgME49:TgMAF1RHb1, and infected female Balb/c mice with these populations of known proportions. Mice were infected with 10<sup>5</sup> tachyzoites of 1:4 or 4:1 proportions of TgME49 to TgME49:TgMAF1RHb1 parasites. Initial population proportions were quantified by IFA. Parasite burden was monitored by bioluminescence imaging for 5 days, after which mice were euthanized and parasites were collected by peritoneal lavage and the population proportions were determined by IFA. After 5 days the proportion of TgMAF1RHb1-expressing

parasites increased significantly, regardless of initial proportion (Figure 2-14A). This competitive advantage was not observed when mice were infected with a mixed population of TgME49 and TgME49:*TgMAF1RHa1* parasites (Figure 2-14A). While we did not notice any differences in growth rate *in vitro* between these strains, we also constructed mixed populations of TgME49 and TgME49:*TgMAF1RHb1* and maintained them in HFFs by serial syringe lysis and passage *in vitro* for 8 weeks. Population proportions were determined by IFA at days 0, 28, and 54-59. In 3 of 4 populations, TgMAF1RHb1-expressing parasites significantly increased in proportion compared to their wildtype counterparts (Figure 2-15). However the competitive advantage of TgMAF1RHb1 expression appears much greater *in vivo* than *in vitro*, as evidenced by the 15-fold greater percent change per day of TgMAF1RHb1 expressing parasites within the population during *in vivo* compared to *in vitro* (Figure 2-14B).



Figure 2-13. Complementation with TgMAF1RHb1 has does not have a significant impact on acute virulence *in vivo* during infection in Balb/c mice.

Balb/C mice were infected intraperitoneally with 100 tachyzoites of luciferase-expressing TgME49 complemented with TgMAF1RHb (n=5) or empty vector (TgME49WT, n=4). Parasite burden was measured using *in vivo* bioluminescence imaging (BLI) at 24-hour time-points. (A) Quantitation of parasite burden by *in vivo* BLI. (B) Representative images of parasite burden in mice at days 6, 8 and 10 post-infection. Scale: photons/sec/cm<sup>2</sup>/sr x10<sup>5</sup>.



Figure 2-14. Expression of TgMAF1RHb1, but not TgMAF1RHa1, in type II *T. gondü* increases competitive advantage.

(A) Mice were infection with mixed populations of TgME49:EV and TgME49:MAF1 with the indicated isoforms and ratios. Infection was allowed to progress for 5 days and population proportions before and after infection were quantified by IFA. Both HMA+ and HMA- MAF1 isoforms were assessed. TgME49:TgMAF1RHb1 significantly increases in proportion to TgME49:EV. \* $\chi$ 2 P-value <0.05. The proportion of TgMAF1RHa1-expressing parasites did not increase during infection. (B) Percent change per day was calculated for the populations that started with 4:1 TgME49:EV to TgME49:TgMAF1RHb1 both in vitro and in vivo by dividing the total percent increases of TgMAF1RHb1-expressing parasites within the populations by the number of days of infection. The first bar of both the in vitro and in vivo infections represent one clone set, while the second bar for each represents a second clone set. (C) Representative images of a mixed population from A before and after a 5-day in vivo infection. HA staining indicates TgMAF1RHb1-positive vacuoles.



Figure 2-15. Expression of TgMAF1RHb1 in Type II *T. gondii* increases competitive advantage *in vitro*. Mixed populations at the indicated ratios were passed and assayed after 0, 4 and 8 weeks of serial passage using IF imaging. TgME49:TgMAF1RHb1 consistently outcompeted TgME49:EV (WT) during *in vitro* co-infections in HFF cells. TgME49:TgMAF1RHb1 increased an average of 1.8-fold over TgME49:EV at 8 weeks post-infection. \* $\chi^2$  p-value < .05X 10<sup>5</sup>.

## 2.3 CONCLUSIONS

The identification of MAF1 as the parasite protein responsible for HMA of intracellular *T. gondii* provides the opportunity to examine the impact of this host-pathogen interaction on parasite biology [106]. Here we have traced the evolutionary history of the *MAF1* locus with respect to

HMA, and for the first time provide evidence of the selective pressure maintaining HMA within parasite populations.

Computational and functional examination of the MAF1 locus in T. gondii, H. hammondi, and N. caninum reveals that the MAF1 gene is duplicated, and individual copies of the MAF1 gene are under diversifying selection. Not all copies of MAF1 mediate HMA, and expressing a *MAF1* copy capable of mediating HMA provides a selective advantage during acute infection in a mouse, while expressing an extra copy of MAF1 that does not mediate HMA does not provide a selective advantage during acute infection. These data support our proposed model wherein the duplication of an ancestral MAF1 gene (MAF1a) in the common ancestor of T. gondii and H. hammondi was followed by diversification and eventual neofunctionalization of a copy of MAF1 into the paralog that is able to mediate HMA (MAF1b). This model is consistent with the observation that both T. gondii and H. hammondi exhibit the HMA phenotype, while the next closest relative, N. caninum, does not exhibit HMA. All three-parasite species have at least one copy of MAF1, however only T. gondii and H. hammondi have copies of MAF1 that fall into the "b" category, and these "b" isoforms are able to mediate HMA while the "a" isoforms are not (Figures 2-11, 2-12). The function of the ancestral *MAF1a* copy remains unknown, however it is not essential *in vitro*[106], and does not mediate HMA (Figure 2-11).

It is also possible that the observed lack of HMA in *N. caninum* parasites is due to a secondary loss of an HMA<sup>+</sup> copy of *MAF1*, or that the ancestral copy of *MAF1* is HMA<sup>+</sup> and parasite populations are losing the ability to mediate HMA. However, these explanations seem less likely for several reasons. Electron micrographs of another *Neospora* species, *Neospora hughesi*, demonstrate a clear lack of HMA in these parasites [166], providing further evidence that HMA evolved in *T. gondii* and *H. hammondi* after the split from the *Neospora* lineage.

Further information regarding the presence or absence of a *MAF1* gene and the HMA phenotype in additional closely related species would help to determine the likelihood that a secondary loss of an HMA<sup>+</sup> *MAF1b* copy occurred in *N. caninum* parasites. It also seems unlikely that the HMA phenotype is ancestral, and *MAF1* copies are losing the ability to mediate HMA, as there is evidence that HMA is selectively advantageous for parasites (discussed below).

The natural variation in MAF1 protein sequence and ability to mediate HMA provides the foundation for future studies regarding the structural aspects and molecular mechanism of MAF1-mediated HMA. Broadly, isoforms that are able to mediate HMA fall into the "b" family which are distinguished from isoforms in the "a" family by several regions of variability, most notably a region containing a proline/serine motif. All isoforms tested that lack this proline-rich region were unable to mediate HMA, making it an ideal candidate for mutational analysis, which will be discussed in Chapter 3.

While tracing the evolutionary history of the *MAF1* locus with regards to HMA, we were also able to find evidence for the selective advantage that likely led to the fixation of the HMA<sup>+</sup> copy of *MAF1* in parasite populations. During initial studies comparing type II parasites (TgME49; HMA<sup>-</sup>) to type II parasites complemented with an HMA<sup>+</sup> copy of MAF1 (TgME49:TgMAF1RHb1; HMA<sup>+</sup>) we did not observe any significant difference in parasite burden or mouse morbidity (Figure 2-13). However, when we competed these HMA<sup>+</sup> and HMA<sup>-</sup> parasite strains head-to-head during an acute mouse infection, we observed a significant growth advantage for the TgMAF1RHb1-expressing parasites. Over the course of five days, the TgMAF1RHb1-expressing parasites outcompeted their HMA<sup>-</sup> counterparts by 7-10% per day (Figure 2-14). This competitive advantage was not observed when we conducted a similar

experiment using type II parasite expressing an extra copy of the HMA<sup>-</sup> isoform, TgMAF1RHa1. These data provide evidence that HMA itself provides a selective advantage.

Despite this clear advantage during acute infection, there is no detectable MAF1b protein expression in type II parasites (Figures 2-6C, 2-8). This lack of MAF1b protein correlates with the lack of HMA in type II parasites, and demonstrates that MAF1b expression (and therefore HMA) is not essential for parasite survival. However, type II parasites do harbor a copy of MAF1b that, if expressed would likely mediate HMA based on comparisons to MAF1 isoforms of known HMA function. It will be interesting for future studies to examine how and why MAF1b protein expression suppressed in type II parasites.

## 3.0 UTILIZING NATURALLY OCCURING APIMAF1 ISOFORMS AND MUTATIONAL ANALYSIS TO DETERMINE TGMAF1 RESIDUES NECESSARY FOR HOST MITOCHONDRIAL ASSOCIATION IN *T. GONDII*

## 3.1 INTRODUCTION

The association of the parasitophorous vacuole membrane (PVM) of intracellular *Toxoplasma gondii* with host organelles is an intimate interaction, on the order of 12-18nm [167]. In fact, electron micrographs reveal that ribosomes are excluded from the space between the rough endoplasmic reticulum and the PVM [167]. The identification of mitochondrial association factor 1 (MAF1) as the *T. gondii* protein responsible for the association of the host mitochondria with the PVM has provided the opportunity to further study this interaction on both a phenotypic and molecular level [106]. This chapter will focus on the further elucidation of the molecular mechanism behind host mitochondrial association (HMA) by using natural variation in the MAF1 protein across Apicomplexan parasites to determine the portions of MAF1 necessary for HMA.

The primary sequence of the MAF1 protein reveals very little about the mechanism behind the HMA phenotype. Aside from a predicted N-terminal signal peptide, a putative transmembrane helix, and several phosphorylation sites there are no identifiable domains of known or predictable function [106]. To begin to understand the mechanism of MAF1 mediated HMA, it is important to identify regions of the MAF1 protein necessary for HMA. The natural variation in MAF1 isoforms and their ability to mediate HMA provide the opportunity to use sequence and structural comparisons to select candidate residues for mutational analysis.

Functional studies discussed in Chapter 2, as well as additional unpublished data from the lab, have broadly categorized MAF1 isoforms into HMA-competent (HMA<sup>+</sup>) and HMA-incompetent (HMA<sup>-</sup>) isoforms. Comparisons of the primary sequences of HMA-competent and HMA-incompetent MAF1 paralogs revealed several regions of dissimilarity, specifically the previously observed proline-rich region, as well as several other small regions of sequence divergence. Here I used site directed mutagenesis to determine that neither the proline-rich region, nor several other residues conserved among HMA-competent MAF1 paralogs are necessary for HMA.

## 3.2 **RESULTS**

#### **3.2.1** Primary sequence alignment of MAF1 isoforms reveals regions of dissimilarity

In Chapter 2, we determined that MAF1 isoforms across *T. gondii*, *H. hammondi*, and *N. caninum* exhibit sequence diversity (Figures 2-4, 2-5), and that not all isoforms mediate HMA (Figures 2-11, 2-12). Further examination of the sequence alignment in Figure 2-5 reveals several regions of dissimilarity among MAF1 isoforms. Here I have simplified the alignment to include only the isoforms for which we have HMA competency data (Figure 3-1). Colored boxes have been drawn around the most diverse regions among MAF1 isoforms. The proline-rich
region (magenta box) and the residues immediately following (green box) are particularly interesting as the MAF1 isoforms that are HMA<sup>+</sup> (TgMAF1RHb1 and HhMAF1b1) appear more similar to one another and divergent from the HMA<sup>-</sup> isoforms (TgMAF1RHb0, TgMAF1RHa1, and HhMAF1a1). This particular grouping is not evident for the rest of the regions of dissimilarity (indicated by the blue, yellow, black, orange, and grey boxes). For these regions TgMAF1RHa1 and HhMAF1a1 are more similar to one another, and divergent from TgMAF1RHb0, TgMAF1RHb1, and HhMAF1b1.

HhMAF1a1 TgMAF1RHa1 TgMAF1RHb0 TgMAF1RHb1 HhMAF1b1	1 1 1 1	MWRVGKSRLYFLFAAGO MWRIWRCRLSFLFATGO MWRIWRCRLSFLFVTGO MWRIWRCRLSFLFATGO MWRVGKSRLYFLFAAGO	DELEGALTAGE DELEGALTAGE DELEGALTAGE DELEGALTAGE DELEGALTAGE	GSQ ISGSAGR GSQMSDSVGR GSQMSDSVGR GSQMSDSVGR GSQ ISGSAGR	NVQAPAGVAE NVQAPAGVAE NVQAPAGVAE NVQAPAGVAE NVQAPAGVAE	APQEAGDVV ASQEAGDVV ASQEAGDVV ASQEAGDVV ASQEAGDVV	55 55 55 55 55
HhMAF1a1 TgMAF1RHa1 TgMAF1RHb0 TgMAF1RHb1 HhMAF1b1	56 56 56 56	EERRERNEQQ I FAP GPF EERTERTEEQVFAP GPF EERTERTEEQVFAP GPF EERTERTEEQAFAL GPF EERRERNEQQ I FAP GPF	PRGHSSESLF PRHSSESLF PRHSSESLF PRHSSESLF PRGHSSESLF	PRSPSVTARR PRNASVTARR PRNPSVTARR PRNASVTARR PRSPSVTARR	RRNRR I ALVA RRNRR I AP I A RRNRR I TL I A RRNRR I AL I A RRNRR I ALVA	TAVGVAV <mark>V</mark> L TAVGVAVIL TAVGVAVIL TAVGVAVIL TAVGVAVIL	110 110 110 110 110
HhMAFlal TgMAF1RHal TgMAF1RHb0 TgMAF1RHb1 HhMAF1b1	111 111 111 111 111 111	AALYVLRRRWARPPEEF AALYVLRRRRAQPPQEF AALYVLRRRRAQPPQEF AAVYVLRRRRAQPPQEF AAVYVLRRRRAQRPQDF AALYALRRRWARPPGEF	PEPP TRPRTR PEPP TRLRTP PEPP TRLRTP PEPP APRSVE PEPP APPSME	RFTTPSGQQQ RPRAPSEQQQ RPRAPSGQQQ DPEVLPEEDE DPEVLPKEDE	P  ASSSLPPPPP PSPSFPPPPP	P	147 147 147 157 165
HhMAFla1	148	SPSGRPAGVPAVPO	BELTLRLTCL	RDTEVKFFGP	SSHRHGFTPL	YDPVANKRV	199
TgMAF1RHa1	148	SESEPPAEVPMTPD	DPLTLRFTCL	GDRNVIFFGP	SGRQDGFTPL	YDPSPSKRV	199
TgMAF1RHb0	148	SESEPPAGVPMKPO	BELTLPFTCL	GDTKVTFFGP	SGRQHGFTPL	YDPSPSKRV	199
TgMAF1RHb1	158	SPPPPPPVEDPLSPE	ESQTVDLSCL	SGTTVRFFGP	SHHFGGFTPL	YDPAPDKRV	210
HhMAF1b1	166	PPPPPPPPVEEPLSPE	ESQAIELSCL	SGTTVRFFGP	SNHPEGFTPL	YDPAPDKRV	220
HhMAF1a1	200	ATVNAGTNTLFIGGSGN	ANGEFAKTII	EEARRNHISL	TAAELSAHSO	ELQERLLRD	254
TgMAF1RHa1	200	ATVDAGTYGLFIGGVGN	ANGEFADTII	EEARRNRIPL	TATELSAESO	EIQERLLHD	254
TgMAF1RHb0	200	ATVDAGANALFIGGGGL	NGQFAKTLL	EEAEKNGIRL	TSVALSEHSO	RIQQSLLRR	254
TgMAF1RHb1	211	ATVDAGANALFIGGGGL	NGQFAKTLL	EEAEKHGIRL	TPEELSOHSO	RIQQSLLRR	265
HhMAF1b1	221	ATVDAGTNDLFIGGGGN	AGQFAKTLL	EEAQKHGMPL	TSSGLSAHSO	QIQEIMLNR	275
HhMAF1a1	255	AERRPGSLVEIDSGSFS	SPVFARSFAF	VAVPSNLFWD	ESETGKNVGA	AFLHILKPE	309
TgMAF1RHa1	255	AEROPGTLVEIDSGRFS	RVFARSFAY	VAIVPNTVWD	ESETGKNVGA	TFLHILKPE	309
TgMAF1RHb0	255	AVKSPGKLVELDTGVAS	SPVFARSFGF	VPVVPGLMWK	ESKVGANVGV	TFIHILKPE	309
TgMAF1RHb1	266	AVKSPGKLVELDTGVAS	SPVFARSFGF	VPVVPGLMWE	ESEVGPNVGV	TFVHILKPE	320
HhMAF1b1	276	AVKKPGKLVEVDTGAGS	SPVFARSFAF	VPVVPGLLWK	ESEVGSNVGV	TFVHILKPE	330
HhMAF1a1	310	VTPHGNPMNDVMIYTVA	APLGNASDDA	YAMAYKATML	SIVGAVSKYN	KTPLGEEKP	364
TgMAF1RHa1	310	VTPHGNEMNDVMLYTVA	APFGNASDSA	YNMAYKATML	GIVGAVSEYN	KTPWGEVKP	364
TgMAF1RHb0	310	VTPYGNLNNNVMMYTVA	APCGAPPDTT	YSLAYKTTIA	GVIRAAAAYN	DTPAGQQYP	364
TgMAF1RHb1	321	VTPYGNLNNNVMMYTVA	APSGAAPDKT	YSLAYKTTIA	GVIGAAAAYN	DTPAGQQYP	375
HhMAF1b1	331	VTPYGNLNNNVMMYTVA	APSGDAPDKA	YSVAYKATIS	GVIGAAAAYN	KSPVGQQYP	385
HhMAF1a1	365	VDA I RLPLLGAGHFRGH	HRSLESIGRV	NATAVKTAIS	QFAPSVELQY	MYDSSDAAF	419
TgMAF1RHa1	365	VEA I RLPLLGAGHFRGH	RRGLHSIGRA	NAVAVEAAIT	RFDPRVELQF	MYEPSDTAL	419
TgMAF1RHb0	365	VQGLRLPLLRGG I FRRN	NRSLESIGRA	NAEGTSLAIT	QYGPNFELQY	MYDPSNAAL	419
TgMAF1RHb1	376	VQGLRLPLLGGG I FRRN	NRSLESIGRA	NAEGTSLAIT	RYGPNFELQY	MYDPSNAAL	430
HhMAF1b1	386	LQGLRLPLLGGG I FRGN	NRSFESIGTA	NAEGTSLAIT	RYGPGFELQY	MYDPSGAAL	440
HhMAFlal TgMAF1RHa1 TgMAF1RHb0 TgMAF1RHb1 HhMAF1b1	420 420 420 431 441	RGLLETEEKFKSHQRD RGLMESERKYKFPQGD HGLQEAESTYLASMLD HGLQEAESTYLASMLD HGLQEAESKYLASMVD					435 435 435 446 456

#### Figure 3-1. Conserved and divergent regions of MAF1.

Alignment was made using Clustal Omega and visualized with Jalview. Residues are colored by conservation (% identity). Colored boxes indicate several different areas with lower levels of conservation. The magenta box indicates the previously identified proline-rich region.

### 3.2.2 The proline-rich region of TgMAF1RHb1 is not necessary for HMA

Previous work determined that amino acids in the proline-rich region do not need to be prolines in order to mediate HMA (data not shown). However, it is possible that this stretch acts as some sort of linker, or that the amino acids immediately surrounding the proline-rich region are important for HMA. In order to determine if the proline-rich region is necessary for HMA, I mutated residues 131-174 of TgMAF1RHb1 to the corresponding residues of TgMAF1RHb0 (131-163), this region encompasses the entire magenta box and several residues into the green box indicated in the alignment (Figure 3-1). I then expressed the mutant MAF1 protein (TgMAF1RHb1∆pro) in type II parasites (Me49), and determined by immunofluorescence assay (IFA) that TgMAF1RHb1∆pro was still able to mediate HMA (Figure 3-2).



Figure 3-2. The proline-rich region of TgMAF1RHb1 is not necessary for HMA.

Type II (Me49) parasites were transfected with a plasmid containing an N-terminally HA-tagged copy of TgMAF1RHb1 where the proline-rich region of the protein was mutated to the corresponding residues of TgMAF1RHb0, thereby eliminating the proline-rich region. Mutant TgMAF1RHb0∆pro region was visualized using an antibody against the HA-tag and host mitochondria were visualized using an antibody against the host cell mitochondrial marker MTCO2. The mutant lacking the proline-rich region is still able to mediate HMA.

## **3.2.3** Structural analysis of the C-terminal domains of TgMAF1RHa1 and TgMAF1RHb1 shows a high degree of similarity in three dimensions

There are no domains of known function predicted by the primary sequence of MAF1, however crystal structures of the C-termini of TgMAF1RHb1 (HMA<sup>+</sup>) and TgMAF1RHa1 (HMA<sup>-</sup>) reveal a large amount of structural similarity between the HMA-competent and HMA-incompetent isoforms (Parker and Boulanger, unpublished; Figure 3-3). Purified C-termini used for these crystallographic studies were the same as those used to generate antibodies in Chapter 2, and are indicated by red (TgMAF1RHb1) and black (TgMAF1RHa1) boxes in Figure 2-5.



Figure 3-3. Structure of TgMAF1RHb1 and TgMAF1RHa1 c-terminal domains.

(A,B) Ribbon diagrams of the three-dimensional structure of the C-termini of TgMAF1RHa1 (A) and TgMAF1RHb1 (B) determined by x-ray crystallography (Parker and Boulanger, unpublished). (C) An alignment of the ribbon diagrams from A and B, illustrating the structural similarity of TgMAF1RHa1 (blue) and TgMAF1RHb1 (red) C-termini.

## 3.2.4 Conserved residues selected for mutational analysis in the C-terminal domain of TgMAF1RHb1 are not required for HMA

Using the crystal structures of the C-termini of TgMAF1RHa1 and TgMAF1RHb1, as well as primary sequence and HMA functionality data discussed in Chapter 2, candidate residues were selected for mutational analysis. To be selected for mutational analysis residues must 1) be different between TgMAF1RHb0 (HMA<sup>-</sup>) and TgMAF1RHb1 (HMA<sup>+</sup>), 2) be conserved among TgMAF1RHb1 and all isoforms containing the proline/serine-rich region, and 3) must not be conserved among TgMAF1RHb0 and other isoforms lacking the proline/serine-rich region. A total of 10 residues fit all three criteria and are surface exposed when mapped to the crystal structure (Table 4). Of these 10 residues, seven are located within two beta strands located at the N-terminus of the crystalized portion of the MAF1 protein.

To determine if the identity of these residues is important for HMA, constructs containing a mutant copy of TgMAF1RHb1 were expressed in type II parasites (Me49), and ability for these mutants to mediate HMA was assessed by IFA. This analysis included a mutant where the entire two beta strands located at the N-terminus of the crystalized portion of TgMAF1RHb1 (Q174-R187) were mutated to the analogous residues of TgMAF1RHb0 (L163-T176) (Figure 3-4A), which encompasses all of the seven conserved residues found within that region. An additional mutant was made where the serine at position 339 in TgMAF1RHb1 was mutated to a phenylalanine, the residue observed at that position in TgMAF1RHb0 (Figure 3-4B). Both mutants were able to mediate HMA. Table 4. Conserved MAF1 residues.

Conserved TgMAF1RHb1	Alternate TgMAF1RHb0 or
residue	TgMAF1RHa1 amino acid
Glutamine 174	Leucine
Aspartate177	Proline
Serine 179	Threonine
Serine 182	Glycine
Glycine 183	Aspartate
Threonine 185	Lysine
Arginine 187	Threonine
Alanine 205	Serine
Aspartate 207	Serine
Serine 339	Phenylalanine



Figure 3-4. Conserved residues in the C-terminal domain are not necessary for HMA.

Type II parasites (Me49) were transfected with a plasmid containing N-terminally HA-tagged mutant versions of TgMAF1RHb1. (A) Ribbon diagram of the C-terminus of TgMAF1RHb1 (red) and the residues mutated to TgMAF1RHb0 (blue). This ribbon diagram is rotated along (B) Mutant MAF1 protein was visualized using an antibody against the HA-tag, and host mitochondria were visualized using an antibody against the mitochondrial marker MTCO2. This mutant MAF1 is able to mediate HMA. (C) Ribbon diagram of the C-terminus of TgMAF1RHb1 with the mutated residue at position 339 highlighted in blue. (D) The conserved serine at position 339 of TgMAF1RHb1 was mutated to the corresponding residue of TgMAF1RHa1, which is a phenylalanine. This mutant is also able to mediate HMA.

### 3.3 CONCLUSIONS

The natural variation in MAF1 protein sequence provides the opportunity to use comparisons between HMA-competent and HMA-incompetent isoforms in order to identify the regions of MAF1 necessary for association with host mitochondria. Initial attempts to identify regions of interest revealed several portions throughout the MAF1 protein with variation between isoforms. Of particular interest were the proline-rich region and amino acids immediately following this stretch of prolines, as TgMAF1RHb1 and HhMAF1b1 (both HMA<sup>+</sup>) are more similar to one another than the HMA<sup>-</sup> isoforms TgMAFRHb0, TgMAF1RHa1, and HhMAF1a1. Mutational analysis was conducted using comparisons of the primary sequence, as well as information gleaned from crystallographic studies. These studies determined that the proline-rich region, as well as several conserved residues in the beta strands found at the N-terminus of the crystalized portion of MAF1 are not necessary for MAF1-mediated HMA (Figures 3-2, 3-4A). Additionally, a surface exposed serine in the C-terminal domain, which is conserved among HMA<sup>+</sup> MAF1 isoforms, is not necessary for HMA (Figure 3-4B). Further mutational analysis must be conducted to see if any of the other identified regions of MAF1 diversity are necessary for HMA.

The remaining regions of sequence divergence indicated in Figure 3-1 do not cluster into HMA<sup>+</sup> and HMA<sup>-</sup> isoforms in the same manner as the proline-rich region and surrounding amino acids. In fact, for these regions the sequences cluster by "a" and "b" paralog, where TgMAF1RHb0 (HMA<sup>-</sup>), TgMAF1RHb1 (HMA<sup>+</sup>) and HhMAF1b1 (HMA<sup>+</sup>) are more similar to one another, and TgMAF1RHa1 (HMA-) and HhMAF1a1 (HMA-) are more similar. Interestingly, both HhMAF1a1 and TgMAF1RHb0 appear to have altered localization patterns compared to TgMAF1RHa1, TgMAF1RHb1, and HhMAF1b1 (discussed briefly in Chapter 2; illustrated in Figure 2-12). This suggests that the ability of MAF1 to mediate HMA is dependent

on two things 1) proper localization and 2) the correct folding or sequence of the region that interacts with the host mitochondria. Previous attempts to C-terminally tag MAF1 constructs have resulted in a lack of HMA (data not shown), suggesting the importance of the C-terminus in MAF1-mediate HMA. Given these data, the most C-terminal region of should be given high priority for future mutational studies.

The recent identification of the interaction partner on the host mitochondria (SAMM50; Kelly and Boothroyd, unpublished) provides the opportunity for additional mechanistic studies. Future mutational analysis will be able to include not only HMA functionality assays through IFA analysis, but also both *in vivo* and *in vitro* assays to determine if these mutations specifically interrupt the ability of MAF1 to interact with SAMM50. Additionally, co-crystallographic studies with SAMM50 and HMA<sup>+</sup> and HMA<sup>-</sup> isoforms of MAF1 may provide additional information on how MAF1 mediates HMA. Further elucidation of how MAF1 mediates HMA will provide the foundation for further studies on the functional consequences of HMA on parasite biology.

### 4.0 THE EFFECTS OF TGMAF1 EXPRESSION THROUGHOUT THE *T. GONDII* LIFE CYCLE

### 4.1 INTRODUCTION

Chronic *T. gondii* infection is marked by the presence of slow growing parasites (called bradyzoites) within cysts that reside primarily in the muscle and central nervous system tissue, including the brain [168, 169]. These tissue cysts are refractory to current drug treatments and persist for the life of the infected host [170].

Cysts can be found in the brains of infected mice as early as two weeks post infection, and the number of cysts in the brains of infected mice reaches a plateau four weeks post infection [171]. The majority of cysts in the brain are spheroidal and have a diameter between 5 and 70  $\mu$ m [169]. However, large cysts with a diameter greater than 70  $\mu$ m are observed, and the number of observed large cysts is greater 5-8 weeks post infection compared to 3-4 week post infection [172]. Cysts remain intracellular and are surrounded by a glycosylated cyst wall [169, 173].

A number of cyst wall components and other parasite proteins necessary establish or maintain chronic infection have been identified. CST1 is a glycoprotein found within the cyst wall, and is responsible for *Dolichos biflorus* binding to *T. gondii* cysts [173, 174]. Deletion of *CST1* results in a thin and fragile cyst wall, as well as reduced cyst numbers in the brains of infected mice [175]. The disruption of the bradyzoite pseudokinase 1 (*BPK1*) gene results in

significantly smaller cysts 8 weeks post infection, and these cysts have a reduced ability to cause oral infection in mice [176]. There have also been several studies demonstrating the importance of secreted rhoptry kinases during chronic infection [177, 178].

Studies focusing on the host response during chronic *T. gondii* infection have identified components of the host immune response necessary for controlling chronic infection. Several cytokines have been implicated in resistance to toxoplasmic encephalitis (TE), specifically Cxcl9, Cxcl10 (a.k.a. IP-10), and Ccl5 (a.k.a. RANTES), as they are expressed during chronic infection in mice that are naturally resistant to TE [179]. Interferon-gamma (IFN $\gamma$ ), which is necessary for control of acute T. gondii infection [93], is also necessary for control of chronic infection. When IFN $\gamma$  is ablated by antibody injection 10 weeks post infection, there is an increase in brain inflammation compared to control mice, and large areas of necrotic brain tissue can be observed [180].

Previous work has established that TgMAF1RHb1 expression, and therefore host mitochondrial association (HMA), alters the host cytokine response [106] and provides a competitive advantage during the acute phase of infection in a mouse model (Chapter 2; [181]). However, it is possible that MAF1-mediated HMA has additional impacts on parasite biology during other phases of infection. Here we provide evidence that MAF1-expression increases cyst burden in chronically infected mice and alters host cytokine response throughout chronic infection.

### 4.2 **RESULTS**

# 4.2.1 Exogenous TgMAF1RHb1 expression in type II parasites increases cyst burden of chronically infected CBA/J mice 4 and 8 weeks post infection

Previous work used a Balb/c mouse model to determine that TgMAF1RHb1 expression and subsequent HMA provides a competitive advantage during acute infection (Chapter 2). It is difficult to assess chronic infection in a Balb/c mouse, as chronic infection in Balb/c mice results in low numbers of brain cysts [182]. In order to assess chronic infection, we used a CBA/J mouse model, as CBA/J mice are more susceptible to chronic infection and have higher numbers of brain cysts during chronic infection [183, 184]. Since CBA/J mice are more resistant to acute infection, I conducted a dose curve with type II (TgME49:EV) and TgMAF1RHb1 complemented type II parasites (TgME49:TgMAF1RHb1) in CBA/J mice to determine an appropriate sub-lethal dose. CBA/J mice were infected with 10, 100, or 1,000 parasites of either TgME49:EV or TgME49:TgMAF1RHb1 and acute infection was assessed (3 mice per strain per dose). Dose was confirmed by plaque assay (TgME49:EV, 18/100; TgME49:TgMAF1RHb1, 23/100), and parasite burden was monitored by bioluminescence. As previously observed in Balb/c mice (Figure 2-13), there was no significant difference in parasite burden between TgME49:EV and TgME49:TgMAF1RHb1 infected mice during acute infection at any dose (Figure 4-1), with the exception of a single time point, 7 dpi at a dose of 1,000 parasites, for which mice infected with TgME49:EV had significantly higher bioluminescence than mice infected with TgME49:TgMAF1RHb1 (p = 0.0004).

To assess the impact of MAF1 expression on chronic infection, CBA/J mice were infected with 1,000 parasites of either TgME49:EV or TgME49:TgMAF1RHb1 and infection

was allowed to progress to the chronic stage of infection (10 mice per strain). Mice were then euthanized at the indicated times, and brains were removed in order to quantify cyst burden. Cyst burden was quantified for half the mice 28 days post infection, and at 56 (experiment 1) or 60 (experiment 2) days post infection. Each experiment was conducted with independently generated TgME49:TgMAF1RHb1 and TgME49:EV clones to control for variation due to TgMAF1RHb1 insertion location. For both experiments there was a higher cyst burden in mice infected with TgME49:TgMAF1RHb1 compared to mice infected with TgME49:EV (Figure 4-2). This difference in cyst burden was not significant at 28 days post infection, however the observed difference in cyst burden was statistically significant at 56 (p = 0.0082) or 60 (p =0.0006) days post infection. For TgME49:TgMAF1RHb1 infected mice, cyst burden was significantly higher at 60 days post infection compared to 28 days post infection in experiment 2 (p = 0.0019); however, only five mice infected with TgME49:TgMAF1RHb1 in experiment 1 survived acute infection, rendering the sample size too small to achieve statistical significance between TgME49:MAF1RHb1 infected mice 28 and 56 days post infection.



## Figure 4-1. TgMAF1RHb1 expression does not significantly alter parasite burden during acute infection in CBA/J mice.

Female CBA/J mice were infected with 10, 100, or 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1 and acute infection was monitored by bioluminescence imaging. (A) Quantification of bioluminescence in mice infected with 10 or 100 parasites. There was no significant difference in bioluminescence between TgME49:EV- and TgME49:TgMAF1RHb1- infected mice at any time at either dose. (B) Quantification of bioluminescence in mice infected with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1. Parasite burden is similar between mice infected with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1. Parasite burden is similar between mice infected with TgME49:EV and TgME49:TgMAF1RHb1, with the exception of the time point 7 days post infection, where mice infected with TgME49:EV had significantly higher bioluminescence compared to mice infected with TgME49:TgMAF1RHb1 (\* p = 0.0004). (C) Images of infected mice quantified in B on days 3, 6, 9, and 12 post infection.



Figure 4-2. TgMAF1RHb1 increases cyst burden in chronically infected CBA/J mice.

Female CBA/J mice were infected with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1. Dose was confirmed by plaque assay (Exp 1: TgMe49:EV 382/1000, TgMe49:MAF1RHb 404/1000; Exp 2: TgMe49:EV 271/1000, TgMe49:MAF1RHb 364/1000). Half of each group was sacrificed 28 dpi, and the remainder were

sacrificed 56 (experiment 1) or 60 (experiment 2) dpi. Brains were removed from sacrificed mice and divided sagitally. The left half of each brain was homogenized and cysts were isolated by Percoll gradient before staining with rhodamine-conjugated *Dolichos biflorus* agglutinin. Cysts were then quantified using and inverted fluorescent microscope. Mice infected with TgME49:TgMAF1RHb1 had a higher cyst burden than mice infected with TgME49:EV. This difference was significant 56 or 60 dpi (\*p = 0.0082; \*\*\*p = 0.0006). There were also significantly more cysts 60 dpi compared to 28 dpi in TgME49:TgMAF1RHb1 infected mice in experiment 2 (\*\* p= 0.0019).

## 4.2.2 Serum cytokine levels differ between mice infected with wildtype type II parasites and type II parasites expressing TgMAF1RHb1 throughout infection

Previous work has shown that the host cytokine response is elevated during the acute stage of *T*. *gondii* infection when TgMAF1RHb1 is expressed [106]. To determine if the host response is altered throughout infection, I infected C57BL6/J mice with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1 and collected blood samples during chronic infection (10 mice per strain). For this experiment, C57BL6/J mice were used as they are susceptible to chronic *T*. *gondii* infection [183, 184], and C57BL6/J is the background strain for the vast majority of knockout mouse strains. Serum cytokine levels were measured by Luminex 0, 21, 28, and 57 days post infection. Mice infected with TgME49:TgMAF1RHb1 generally had elevated serum cytokine concentrations compared to mice infected with TgME49:EV (Figure 4-3; Appendix D). Two mice appeared to be uninfected, as there was no observed bioluminescence during imaging and there was no interferon-gamma (IFN $\gamma$ ) response, so these mice were eliminated from graphical and statistical analyses in Figure 4-3 (mice numbered 2-2 and 3-2; values for these mice available in Appendix D). To determine if the observed differences in cytokine expression

were significant, I conducted a two-way ANOVA for each individual cytokine followed by multiple comparisons with a Sidak correction to compare the means of TgME49:EV infected mice to TgME49:TgMAF1RHb1 infected mice at each time point. There were 13 cytokines with at least one significant time point, including two cytokines with two significant time points (IP-10 and RANTES). At 21 days post infection mice infected with TgME49:TgMAF1RHb1 had significantly higher levels of Eotaxin and IP-10 compared to TgME49:EV infected mice. At 28 days post infection TgME49:TgMAF1RHb1 infected mice had significantly higher levels of G-CSF, IL-6, IL-10, IL-12p40, IP-10, MCP-1, and RANTES, while at 57 days post infection these mice had higher IL-13, LIF, MIG, MIP-1b, RANTES, and VEGF compared to TgME49:EV infected mice. If we apply an additional Bonferroni by multiplying each p-value by 32 (the number of cytokines analyzed), only three points of comparison are significant. These points are IL-10 at 28 days post infection, LIF at 57 days post infection, and VEGF at 57 days post infection.





















\*\*

v

.

5





Female C57BL6/J mice were infected with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1. Blood samples were taken prior to infection, as well as 21, 28, and 57 days post infection and processed to obtain serum samples. Serum samples were then analyzed by Luminex for the presence of 32 mouse cytokines. For each cytokine, serum levels are given as fluorescence intensity (FI) minus background. Each point represents a single mouse, blue circle indicated mice infected with TgME49:EV and red squares indicate mice infected with TgME49:MAF1RHb1. A two way ANOVA followed by multiple comparisons of the means of TgME49:EV versus TgME49:TgMAF1RHb1 infected mice was used to determine significant differences in cytokine level for each cytokine at each time point. A Sidak correction for multiple comparisons was used. Significant differences indicated by: \* p < 0.05; \*\* p < 0.01; \*\*\*\* p < 0.001.

### 4.2.3 Screen of Candidate cytokines in vivo

Candidate cytokines were selected based on Luminex data and presence in the literature indicating a possible role during *T. gondii* infection, with an emphasis on cytokines implicated in a possible role during chronic infection. Female RANTES knockout (KO; a.k.a. Ccl5 -/-), IP-10 KO (a.k.a. Cxcl10 -/-), and NOS2 KO, or wildtype C57BL6/J mice were infected with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1 and acute infections were monitored by bioluminescence imaging (Figure 4-4). Parasite viability was confirmed by plaque assay (TgME49:EV, 40/100; TgME49:TgMAF1RHb1, 31/100).

For the wildtype mice (C57BL6/J) there was a slightly, but not significantly, higher parasite burden in TgME49:TgMAF1RHb1 infected mice compared to TgME49:EV infected mice, which is consistent with previous data collected from both Balb/c and CBA/J mice (Figures 2-13, 4-1, and 4-4A). Ccl5 knockout mice infected with TgME49:TgMAF1RHb1 had a slightly higher parasite burden compared to Ccl5 knockout mice infected with TgME49:EV, but again this difference was not significant (Figure 4-4B). All Ccl5 knockout mice infected with either parasite strain became morbid by 12 DPI, and there was no significant difference in survival. For the Cxcl10 knockout mice, there was no significant difference in parasite burden for TgME49:EV or TgME49:TgMAF1RHb1 infected mice. In fact for the first 6 days post infection, in contrast to what has been observed for multiple wildtype strains of mice, there was not even a modestly higher parasite burden for TgME49:EV. NOS2 knockout mice infected with TgME49:TgMAF1RHb1 had a slightly higher parasite burden than the NOS2 knockout mice infected with TgME49:TgMAF1RHb1 had a slightly higher parasite burden than the NOS2 knockout mice infected with TgME49:TgMAF1RHb1 had a slightly higher parasite burden than the NOS2 knockout mice infected with TgME49:TgMAF1RHb1 had a slightly higher parasite burden than the NOS2 knockout mice infected with TgME49:EV. However, only one NOS2 knockout mices was successfully

infected with TGME49:EV, therefore it was not possible to perform statistics on the NOS2 knockout mice. All NOS2 knockout mice infected with either parasite strain became morbid by 9 DPI, however once again the low sample size prevented statistical analysis.





Female cytokine KO or wildtype mice were infected with 1,000 parasites of TgME49:EV or TgME49:TgMAF1RHb1 and acute infection was monitored by bioluminescence imaging. Parasite viability was confirmed by plaque assay (TgME49:EV 40/100; TgME49:TgMAF1RHb1 31/100). (A) For wildtype C57BL6/J, luminescence is higher in TgME49:TgMAF1RHb1 infected mice, however there is no significant difference between these mice and the TgME49:EV infected controls. (B) Ccl5 (RANTES) KO mice infected with TgME49:TgMAF1RHb1 had a slightly higher, but not statistically significant, parasite burden compared to

TgME49:EV infected mice. (C) Cxcl10 (IP-10) KO mice infected with TgME49:EV and TgME49:TgMAF1RHb1 had similar parasite burdens until 7 DPI, when mice infected with TgME49:TgMAF1RHb1 have a slightly, but not statically significant, higher parasite burden until 12 DPI. (D) NOS2 KO mice infected with TgME49:TgMAF1RHb1 had a higher parasite burden compared to TgME49:EV infected mice. No statistics were performed for these mice due to n=1 for TgME49:EV infected NOS2 KO mice.

### 4.3 CONCLUSIONS

Here we provide evidence that TgMAF1RHb1 expression may play a role in the establishment and maintenance of chronic *T. gondii* infection. Type II parasites, such as the TgME49 strain, do not normally express a copy of TgMAF1 capable of mediating HMA. Despite this, TgME49 parasites are quite capable of establishing chronic infection and forming tissue cysts in the brains of infected mice. However, the exogenous expression of an HMA-competent copy of TgMAF1 (TgMAF1RHb1) leads to a higher number of cysts in the brains of chronically infected mice when compared to mice infected with wildtype type II parasites. This increase in cyst burden is observed at 4 weeks post infection, but is even more pronounced and statistically significant at 8 weeks post infection. There is also a significant increase in cyst burden from 4 to 8 weeks post infection in mice infected with TgME49:TgMAF1RHb1, but this increase over time is not observed in mice infected with TgME49:EV.

It has traditionally been thought that bradyzoites within tissue cysts remain relatively quiescent, and are limited in their replicative capabilities. However, a recent study looking directly at bradyzoite division within cysts *in vivo* has shown that bradyzoites replicate both asynchronously and synchronously within mature cysts [172]. Our data showing an increase in cyst burden of mice infected with TgME49:TgMAF1RHb1 over time also support the idea that

bradyzoites, and perhaps the cysts themselves, grow and divide. However, the mechanism by which the cyst burden of TgME49:TgMAF1RHb1 infected mice increases remains unknown. Further investigations into the mechanism of increased cyst burden over time in these mice will be limited by the lack of technology available to track chronic infection *in vivo*. The newest technology examining replication of bradyzoites within cysts may be useful in determining if TgMAF1RHb1 expressing parasites have increased replication rates *in vivo* compared to wildtype type II bradyzoites.

In addition to a higher number of brain cysts, mice chronically infected with TgME49:TgMAF1RHb1 have generally elevated cytokine levels compared to the TgME49:EV infected controls. There were 13 cytokines that had a significant difference between TgME49:TgMAF1RHb1 and TgME49:EV infected mice during at least one time point. Interestingly, IP-10 (Cxc110) and RANTES (Cc15) each had two time points with significant differences, and both of these cytokines have been implicated in resistance to TE. Despite the implication that both Cxc110 and Cc15 promote resistance to TE, we see that mice infected with TgME49:TgMAF1RHb1 and therefore elevated Cxc110 and Cc15, actually have more brain cysts, suggesting that the role of these cytokines is perhaps more complicated.

In the absence of Ccl5. mice infected with either TgME49:EV or TgME49:TgMAF1RHb1 become morbid 9-12 days post infection, with no significant difference in morbidity between mice infected with either parasite strain. Similar to wildtype mice, we see a slightly, but not significantly, higher parasite burden in TgME49:TgMAF1RHb1 infected mice compared to mice infected with TgME49:EV. Interestingly, in the absence of Cxcl10, we did not observe this slightly higher parasite burden in mice infected with TgME49:TgMAF1RHb1 compared to TgME49:EV infected mice until 7 days post infection.

81

Mice lacking the inducible nitric oxide synthase (NOS2) were also used for this small scale *in vivo* screen. NOS2 KO mice were included because nitric oxide (NO) has been implicated in the control of chronic [185], *T. gondii* infection. In this study NOS2 KO mice infected with both TgME49:EV and TgME49:TgMAF1RHb1 became morbid 8-9 days post infection, with no significant differences between mice infected with either strain. Similarly to the Ccl5 KO and wildtype C57BL6/J mice, there was a slightly higher parasite burden in TgME4:TgMAF1RHb1 infected mice, however only one NOS2 mouse was infected with TgME49:EV and we were therefore unable to determine if this difference was significant.

Together these data suggest that TgMAF1RHb1 plays a role in the persistence of chronic *T. gondii* infection, however it does not determine the role of any of the candidate cytokines during this process. Many cytokines are likely to have a role in controlling acute infection and the loss of those cytokines during acute infection will prevent mice from surviving until chronic infection. To avoid this problem, future studies regarding the role of TgMAF1RHb1 expression during chronic infection should utilize conditional knockouts or antibody ablation of cytokines during chronic infection.

### 5.0 CONCLUSIONS AND FUTURE DIRECTIONS

Gene duplication, and subsequent diversification, is one mechanism by which new phenotypes evolve. The process by which a duplicated gene accumulates mutations and ultimately acquires a new function distinct from any ancestral gene functions is called neofunctionalization. Chapter 2 presents compelling evidence that the HMA phenotype during *T. gondii* infection evolved via neofunctionalization of a duplicated gene.

The *MAF1* gene, which is responsible for the HMA phenotype, is uniquely duplicated and diversified in *T. gondii*, compared to its closest relatives *H. hammondi* and *N. caninum*. We have shown that both type I and III *T. gondii* strains and *H. hammondi* are able to mediate HMA, while type II *T. gondii* strains and *N. caninum* are not. The presence or absence of the HMA phenotype correlates with the expression of "b" type MAF1 paralogs containing a proline-rich region. And indeed, "b" paralogs from both *T. gondii* and *H. hammondi* (TgMAF1RHb1 and HhMAF1b1) are able to mediate HMA when expressed in parasites that do not normally associate with host mitochondria, while "a" paralogs from *T. gondii* and *H. hammondi* (TgMAF1RHa1 and HhMAF1a1) do not mediate HMA. Perhaps most importantly, we have determined that expression of an HMA-competent MAF1 paralog provides a competitive advantage to parasites that do not normally associate with host mitochondria. This advantage is linked to expression of the HMA<sup>+</sup> paralog, and exogenous expression of an HMA<sup>-</sup> paralogs does not provide the same advantage. It is therefore likely, that the competitive advantage provided to parasites expressing and HMA<sup>+</sup> MAF1 paralog led to the fixation of the HMA phenotype, and the neofunctionalized *MAF1* gene, in parasite populations.

The selective advantage of HMA<sup>+</sup> parasites compared to HMA<sup>-</sup> parasites provides evidence against the alternative hypothesis that HMA<sup>+</sup> *MAF1b* is the ancestral phenotype and that *MAF1* paralogs are losing the ability to mediate HMA. This advantage also argues against the possibility that *N. caninum* lacks HMA due to secondary loss of an HMA<sup>+</sup> copy of *MAF1*. The apparent lack of HMA in *N. hughesi* also supports the proposed neofunctionalization model where HMA evolved in the ancestor of *T. gondii* and *H. hammondi* following the split from the *Neospora* lineage.

Type II parasite genomes contain multiple copies of MAF1, including copies that are predicted to mediate HMA if expressed. However, these potentially HMA<sup>+</sup> isoforms are never expressed on the protein level. How and why type II parasites silence the expression of potentially HMA-competent MAF1 isoforms remains unknown. There are many known mechanisms of regulating protein expression before, during, and after transcription, as well as during or after translation.

Given that there is evidence of an antisense transcript across the 3' end of the *MAF1* gene, it seems likely that MAF1 protein expression is regulated at the transcriptional level. The apparent length and location of the antisense transcript suggests a possible long noncoding RNA (lncRNA) might be responsible for the lack of MAF1b protein expression in type II parasites. It will be interesting for future studies to focus on this potential lncRNA, and the mechanism of silencing a potentially HMA<sup>+</sup> MAF1 isoforms in the type II parasite linage. It will also be interesting to determine what evolutionary pressures or events led to the silencing of these MAF1 isoforms only in type II parasites.

Chapter 3 uses the naturally occurring sequence variation of MAF1 paralogs to identify regions of diversity among paralogs that may be important for mediating HMA. Candidate residues for mutational analysis were determined by conservation among HMA<sup>+</sup> MAF1 isoforms compared to HMA<sup>-</sup> MAF1 isoforms, and informed by structural information obtained from crystallographic studies. The regions of diversity tested in Chapter 3 included the proline-rich region, a stretch of amino acids comprising two beta sheets immediately C-terminal to the proline-rich region, and a serine residue conserved among HMA<sup>+</sup> MAF1 isoforms. Surprisingly, none of these regions were necessary for HMA, as type II parasites expressing TgMAF1RHb1 with amino acid substitutions in these regions were able to associate with host mitochondria. There are still a number of regions of MAF1 diversity that may prove to be necessary for HMA. Future studies should focus on these regions, particularly the region at the C-terminus of the protein. The C-terminus is a particularly attractive candidate for mutational analysis, as Cterminally HA-tagging TgMAF1RHb1 disrupts its ability to mediate HMA (data not shown), suggesting that this region may be involved in mediating HMA. It will also be interesting to see if the same portions of MAF1 that are necessary for HMA are the portions of MAF1 that interact with the host protein SAMM50.

In Chapter 2 we determined that TgMAF1RHb1 expression, and subsequent HMA, provided a competitive advantage during acute infection with a mixed population of both HMA<sup>+</sup> and HMA<sup>-</sup> parasites. In Chapter 4 we explore the impact of TgMAF1RHb1 expression during chronic infection. We find that TgMAF1RHb1 expression in a type II *T. gondii* background increases the number of cysts during chronic infection 4 and 8 weeks post infection. Perhaps most interestingly, the number of cysts in the brains of mice chronically infected with TgME49:TgMAF1RHb1 was significantly higher 8 weeks post infection compared to 4 weeks

post infection. These data suggest that cysts, or the bradyzoite parasites within the cysts, are capable of dividing even after chronic infection has been established. This is consistent with recently published work showing that bradyzoites do in fact replicate, even within mature tissue cysts [172].

We have also determined that the host cytokine response during chronic infection is significantly different between TgME49:EV and TgME49:TgMAF1RHb1 infected mice. Several of the cytokines that are significantly higher in the serum of TgME49:TgMAF1RHb1 infected mice have previously been implicated in control of chronic infection, and were therefore used in a small scale *in vivo* screen. Mice lacking either Ccl5 (RANTES) or NOS2 became morbid during acute infection, however there was no difference in morbidity between TgME49:EV and TgME49:TgMAF1RHb1 infected mice. Similar to wildtype C57BL6/J mice, there was a slightly higher parasite burden for mice infected with TgME49:TgMAF1RHb1 compared to those infected with TgME49:EV, but this difference was not significant. With Cxcl10 KO mice, this difference was not observed until after 7 days post infection. This suggests some sort of interaction between TgMAF1RHb1 expression and Cxcl10, however exactly how the presence or absence of Cxcl10 impacts parasite growth with or without TgMAF1RHb1 expression remains to be determined.

Cyst burden increases over time in mice chronically infected with type II parasites expressing an exogenous copy of TgMAF1RHb1, and therefore exhibiting HMA. Exactly how TgMAF1RHb1 expression leads to more cysts over time remains unknown. Further analysis of the differences in cytokine expression during chronic infection between HMA<sup>-</sup> and HMA<sup>+</sup> strains may reveal additional candidate cytokines or host pathways manipulated by TgMAF1Rhb1 expression. The use of conditional knockout mice, or antibody ablation of specific cytokines during chronic infection, will help us to understand the mechanism of increased cyst burden in mice infected with TgMAFRHb1-expressing strains.

While the exact impact of HMA on mitochondrial function remains unclear, known mechanisms of mitochondrial manipulation by viral and intracellular bacterial pathogens may provide insight into what is happening to host cells when their mitochondria associate with *T. gondii*. For instance, several viruses manipulate  $Ca^{2+}$  homeostasis, altering  $Ca^{2+}$  flux into and out of the mitochondria [129, 130]. It is possible that *T. gondii* mediated HMA disrupts  $Ca^{2+}$  homeostasis in host cells. A variety events in the *T. gondii* life cycle are controlled by  $Ca^{2+}$ , including invasion, motility, secretion of parasite proteins into the host cell, and egress [186]. If MAF1-mediated HMA alters  $Ca^{2+}$  homeostasis in host cells, it might positively impact the ability of parasites to perform any of these critical tasks, leading to increased parasite virulence, and the observed selective advantage.

Alternatively, HMA could alter mitochondrial signaling pathways, including apoptotic or immune signaling. Viruses manipulate host apoptotic signaling to prevent apoptosis, allowing more time for viral replication [131, 132], or to induce apoptosis and facilitate host cell lysis [133, 134]. It has been shown that *T. gondii* infected host cells are resistant to apoptosis triggered by several different inducers of apoptosis [187]. It is possible that the association of host mitochondria with the vacuole surrounding intracellular *T. gondii* inhibits the mitochondrial apoptotic signaling pathway. This could result in increased parasite survival and therefore increased parasite virulence. Interestingly, this hypothesis is supported by a study showing that cells infected with a type II strain (ME49) had higher rates of apoptosis compared to a type I strain (RH) [188].

As for immune signaling, our studies indicate that expression of TgMAF1RHb1 in type II parasites results in higher serum cytokine levels of chronically infected mice. It has also been shown that mice infected with type II parasites expressing TgMAF1RHb1 have higher serum cytokine levels during acute infection compared to mice infected with wildtype type II parasites [106]. It is therefore possible that HMA increases mitochondria-mediated immune signaling. However, it is important to note that mice infected with type II parasites expressing TgMAF1RHb1 have an increased parasite burden, which could explain the increased cytokine expression in these mice. Further studies examining the impact of HMA on host cell signaling will determine exactly what this parasite phenotype is doing to host cells and how these changes to the host cell benefit the parasite.

### 6.0 MATERIALS AND METHODS

### 6.1 PARASITE STRAINS AND MAINTENANCE

All *T. gondii* and *N. caninum* parasite strains were maintained by passage in human foreskin fibroblasts (HFFs) at 37° C and 5% CO<sub>2</sub>. HFFs were grown in Dulbecco's modified Eagles Medium (DMEM) supplemented with 10% fetal bovine serum, 2 mM glutamine, and 50  $\mu$ g/mL each of penicillin and streptomycin (CDMEM). Mouse embryonic fibroblasts (MEFs) were similarly maintained at 37° C and 5% CO<sub>2</sub> in CDMEM, however MEFs were passed at a 1:5 or 1:10 ratio once confluent. MEFs were allowed to remain confluent for a maximum of 5 days before passage. RAW 264.7 cells were maintained at 37° C and 5% CO<sub>2</sub> in CDMEM buffered with 10 mM HEPES (Gibco).

To produce *H. hammondi* oocysts, interferon- $\gamma$  KO mice were fed 10<sup>4</sup> *H. hammondi* oocysts and killed ~60 days postinfection (pi). Muscles from infected mice were then fed to 10-to 20-week-old cats, and feces were collected during days 5–11 postinfection. Unsporulated oocysts were isolated by sucrose flotation, and the resulting oocysts were allowed to sporulate at ambient temperature in 2% H<sub>2</sub>SO<sub>4</sub> [10]. Oocyst preparations were stored at 4° for no longer than 6 months. Sporulated oocysts (40–80 million) were washed four times in Hank's Buffered Saline Solution (HBSS) and treated with 10% bleach in PBS for 30 min. Pellets were resuspended in 4 ml HBSS and vortexed at maximum speed along with 1 g of sterile glass beads (710–1180  $\mu$ M,

Sigma) for 30 sec, allowed to cool for 30 sec, and then vortexed for 30 sec again. DNA was isolated directly from the pellet of cracked oocysts (containing sporocysts released from the oocysts and debris) using the DNAzol reagent (Invitrogen; Carlsbad, CA).

In other cases, we used the sporocyst preparation to generate *in vitro* cultures of *H*. *hammondi*. To do this, we exposed the cracked oocyst preparation to PBS containing trypsin (Sigma T4799; 12.5 mg/ml) and taurocholic acid (Sigma T4009; 50 mg/ml) at  $37^{\circ}$  for 30 min. The reaction was quenched by the addition of cDMEM (containing 10% FBS) and we removed debris from the preparation by filtration through 5-µm syringe filters (Millipore). The resulting sporozoites were used to infect confluent monolayers of HFFs seeded on 12-mm circle glass coverslips. Samples were fixed and processed for immunofluorescence (IF) as described below. Parasite strains used throughout this dissertation have been included in Table 3.

Strain name	Species	Notes
ME49:TgMAF1RHb1	T. gondii	Transgenic type II expressing MAF1 isoform
U		TgMAF1RHb1, parent strain ME49∆HPT:Luc
ME49:EV	T. gondii	Empty vector (EV) control, parent strain
		ME49AHPT:Luc complemented with HPT
ME49:TgMAF1RHa1	T. gondii	Transgenic type II expressing MAF1 isoform
		TgMAF1RHa1, parent strain ME49∆HPT:Luc
ME49:TgMAF1RHb0	T. gondii	Transgenic type II expressing MAF1 isoform
6		TgMAF1RHb0, parent strain ME49∆HPT:Luc
ME49:HhMAF1a1	T. gondii	Transgenic type II expressing MAF1 isoform
		HhMAF1a1, parent strain ME49∆HPT:Luc
ME49:HhMAF1b1	T. gondii	Transgenic type II expressing MAF1 isoform
		HhMAF1b1, parent strain ME49△HPT:Luc

Table 5. Strains used throughout dissertation.

ME49:TgMAF1RHb1Apro	T. gondii	Type II strain expressing a mutant TgMAF1RHb1
		isoform lacking the proline-rich region
ME49:TgMAF1RHb1:174-187b0	T. gondii	Type II strain expressing a mutant TgMAF1RHb1
U		isoform where residues 174-187 were substituted
		for corresponding residues found in TgMAF1Rhb0
ME49:TgMAF1RHb1:S339F	T. gondii	Type II strain expressing a mutant TgMAF1RHb1
Ŭ		isoform where serine 339 is mutated to a
		phenylalanine
ME49ΔHPT:Luc	T. gondii	Type II strain lacking HPT and expressing
		Luciferase, used as parent strain for transgenic type
		II strains
CTG ΔHPT:Luc:GFP	T. gondii	Type III strain lacking HPT and expressing
		Luciferase and GFP, used as parent strain for
		transgenic type III strains
S1T:dsRed:Luc	T. gondii	F1 progeny of a II x III cross, lack all known
		virulence factors, expresses dsRed and Luciferase
S26:dsRed:Luc	T. gondii	F1 progeny of a II x III cross, lack all known
		virulence factors, expresses dsRed and Luciferase
S22:dsRed:Luc	T. gondii	F1 progeny of a II x III cross, lack all known
		virulence factors, expresses dsRed and Luciferase
S23:dsRed:Luc	T. gondii	F1 progeny of a II x III cross, has all known
		virulence factors, expresses dsRed and Luciferase
NC-1∆HPT:dsRed:Luc	N. caninum	NC-1 isolate, lacks HPT, expresses both dsRed and
		Luciferase, used as parent train for transgenic
		Neospora strains
H. hammondi	H. hammondi	H. hammondi parasites derived from oocysts

### 6.2 SEQUENCE COVERAGE AND COPY NUMBER ANALYSIS

Copy number analysis was performed as described previously [106, 151]. Briefly, raw sequence reads from multiple *T. gondii* strains and *N. caninum* (Liverpool strain)[189] were downloaded from the NCBI trace archive in fasta format (strains GT1, ME49, and VEG were derived from Sanger-based shotgun sequencing; MAS, P89, FOU, VAND, and RUB were generated using Roche 454 technology). *T. gondii* and *N. caninum* reads were aligned to the *T. gondii* ME49 genome (ToxoDB version 7.3; www.toxodb.org) using BLAT (parameters: -fastMap – minIdentity = 95 –minScore = 90) [190] and coverage was calculated in each 500-bp window using coverageBed (from the Bedtools suite) [191]. *H. hammondi* reads (strain HH34) [192] were aligned using Bowtie2 (using default parameters plus –end-to-end) [193],and sequence coverage calculations were made using the integrated genome browser (IGB) [194]. All coverage and annotation data were then plotted using custom scripts in R statistical software. To do this, start and end coordinates of regions of the *MAF1* locus were noted and data were normalized to the average coverage of ~20 Kb upstream of the locus [192].

### 6.3 HIGH MOLECULAR WEIGHT SOUTHERN BLOTTING

Southern blotting was performed as previously described [151]. The six strains of *T. gondii* used were GT1 and RH (type I), ME49 and PRU (type II), and VEG and CTG (type III). Genomic DNA from each strain was digested with *Sca*I restriction enzyme in a 100-µl reaction volume for  $\sim$ 12 hr and resolved by pulsed field gel electrophoresis (Bio-Rad CHEF-DR III system).
Resolved fragments were probed with DIG-labeled (Roche) *MAF1*-specific probes followed by chromogenic detection as per manufacturer's protocol.

# 6.4 AMPLIFICATION AND CLONING OF *MAF1* PARALOGS FROM *T. GONDII*, *H. HAMMONDI*, AND *N. CANINUM*

Due to the fact that the *MAF1* locus exhibits significant copy number variation across species and strains, we used long-extension PCR and cloning to identify *MAF1* paralogs in *T. gondii* (strains RH, ME49, and CTG), *H. hammondi* (strain HhCatGer041), and *N. caninum* (strain NC-1). Long extension PCR was used to minimize the potential for chimera formation between different *MAF1* paralogs (as described in [106]). For cloned sequences, all polymorphisms were validated by querying a local copy of the sequence read database for the presence of that polymorphism along with at least 40 bp of flanking sequence (RH was compared to GT1; ME49 was compared to ME49; CTG was compared to VEG) [192]. This served three purposes: validation of SNPs specific to a given clonal lineage, elimination of PCR-derived SNPs, and controlled for the possibility of generating interparalog chimeric sequences during PCR amplification when the polymophisms were  $\leq$ 40 bp apart. Since our SNP curation method relied on comparing between members of the same clonal lineage (*e.g.*, RH *vs.* GT1), it is possible that some isolate-specific SNPs were artificially eliminated during curation. We did not identify any evidence for chimerism in our sequences although this outcome cannot be completely ruled out.

## 6.5 SEQUENCE ANALYSIS

Coding sequences of *MAF1* paralogs from multiple *T. gondii* strains and other species were analyzed using algorithms implemented in MEGA6 [195] as follows: Specifically, coding sequences were translated into protein and aligned using Muscle (default settings). Phylogenetic trees were constructed using maximum parsimony and the subtree-pruning-regrafting algorithm [196]. Search level was 1 and the initial trees were obtained by the random addition of sequences (10 replicates were performed). Branch lengths were calculated using the average pathway method. All positions containing gaps and missing data were eliminated, and there were a total of 359 useful positions in the final dataset.

We calculated pairwise dN/dS ratios for all "b" paralogs (including the HMAincompetent b0 paralogs) to determine if they had been under either positive or purifying selection. To do this, we used the modified Nei-Gojobori method with the assumed transition/transversion bias of 2 [197], and as above, all positions containing gaps were eliminated. All analyses were conducted in MEGA6 [195] and pairwise *P*-values for dN/dS ratios were deemed significant at P < 0.05.

To calculate pairwise Ka/Ks values for the same sequences used in the dN/dS analysis, sequences were aligned using ClustalOmega[198], and all positions containing gaps were removed by hand. Sequences were analyzed pairwise using an online KaKs calculation tool [159], values reported are Ka/Ks and Ks calculated using Kimuras two parameter model.

94

## 6.6 GENERATION OF CONSTRUCTS AND TRANSGENIC PARASITES

Generation of pMAF1RHb1 (N-terminally hemagglutinin (HA)-tagged MAF1b) expression construct has been described previously [106]. For pMAF1RHa, the coding sequence for *MAF1RHa* was amplified from RH cDNA, cloned, and then used in a splicing by overlap extension (SOE) PCR reaction to fuse the N-terminal portion of *MAF1RHb1* gene and the Cterminal portion of the *MAF1RHa1* gene. The specific construct contained the *MAF1RHb1* promoter, start codon, signal sequence, an HA tag (as in [106]), and this was followed by the remainder of the C terminus encoding portion of the MAF1RHa1 gene. The TgMAF1RHb0, TgMAF1RHb1, HhMAF1a1, HhMAF1b1, and NcMAF1 constructs were made using SOE PCR to introduce an HA-tag following the predicted signal peptide for each isoform. Plasmid templates for the first round of PCR were generated from genomic DNA, which included 1116 bp upstream of the start site to include the putative promoter. Transgenic parasite lines were generated by transfecting TGME49Δ*hpt* (MΔLuc) and NC-1Δ*hpt* parental strains with 50  $\mu$ g of *Hind*III-linearized plasmid. Stable expression lines were isolated by selection in mycophenolic acid (MPA)/xanthine followed by limiting dilution in 96-well plates.

# 6.7 TGMAF1RHA1 AND TGMAF1RHB1 CLONING, PROTEIN PRODUCTION, AND PURIFICATION

A construct encoding the predicted C-terminal domain of TgMAF1RHb1 (Thr159 to Asp435) was codon optimized for *Escherichia coli* and synthesized by GenScript. A construct of *TgMAF1RHa1* containing the analogous C-terminal domain (TgMAF1RHa1; Ser173 to Ser443)

was amplified from *T. gondii* cDNA. Each construct was subcloned into a modified pET28a vector encoding an N-terminal hexa-histidine tag separated from the sequence of interest by a tobacco etch virus (TEV) protease cleavage site. Constructs were produced recombinantly in *E. coli* BL21 cells. Following 4 hr of growth at 310 K and 12 hr at 303 K, the cells were harvested by centrifugation, resuspended, and lysed using a French press. TgMAF1 proteins were purified from the soluble fraction by Ni-affinity chromatography, the His tag was removed by TEV protease, and TgMAF1 proteins were further purified by size exclusion chromatography on a Superdex 75 16/60 HiLoad column in HBS (20 mM Hepes pH 7.5, 150–300 mM NaCl) with 1% glycerol and 1 mM dithiothreitol.

## 6.8 IMMUNOFLUORESCENCE ASSAYS AND CONFOCAL MICROSCOPY

HFFs or NRK-mitoRFP cells were seeded on 12-mm coverslips in 24-well plates and grown to ~80% confluency. NRK-mitoRFP cells were infected with *N. caninum* or *T. gondii* strains expressing GFP and incubated for 8 hr. For HFFs, MitoTracker staining was performed as follows: Growth medium on the HFF monolayer was replaced with DMEM containing MitoTracker (Red CMXRos, Invitrogen) at a 30-nM concentration and incubated for 30 min at 37°. Cells were then washed with PBS, infected with parasites in prewarmed DMEM, and incubated for 4 hr at 37°. After incubation, the infected cells were washed with PBS, fixed with 3% paraformaldehyde in PBS for 15 min, and blocked/permeabilized in PBS containing 5% BSA and 0.2% Triton X-100. Alternatively, NRK-mitoRFP infected cells were fixed with 3% PFA and either mounted directly or Hoechst stained prior to mounting followed by visualization. Fixed cells were then immunostained with rat monoclonal anti-HA (3F10 clone, Roche) at

1:1000, mouse anti-MAF1a/b polyclonal antibodies at 1:1000, or mouse monoclonal anti-MTCO2 (ab110258, Abcam) at 1:500.

## 6.9 QUANTIFICATION OF VACUOLE COVERAGE

Percent vacuole coverage was determined using confocal microscopy and ImageJ. Populations transfected with HA-tagged TgMAF1RHb1 or HhMAF1b1 were fixed and stained with anti-HA and anti-MTCO2 primary antibodies. Confocal images were taken in three channels; 594 (anti-MTCO2), 488 (anti-HA), and DIC. All three images were converted to 8-bit images and merged using ImageJ. Vacuoles were traced while only the DIC and green channels were visible, and then pixel intensity along the vacuole was measured in the red channel. Pixel intensities >20 were considered to be mitochondria. Percent vacuole coverage was calculated by measuring the length of the vacuole trace with pixel intensity >20 and dividing it by total vacuole trace length. Twenty HA-positive vacuoles were measured for both the TgMAF1RHb1 and HhMAF1b1 populations. Ten HA-negative vacuoles were measured from each population (20 total) as a WT control.

## 6.10 WESTERN BLOT ANALYSIS

Parasites were filtered away from host cell debris and lysed in  $1 \times$  SDS lysis buffer. Proteins were resolved by SDS-PAGE, transferred onto nitrocellulose membrane, and blocked for 1 hr in 5% (w/v) milk in TBS-Tween20 (TBS-T). Primary antibody incubation was performed in blocking

buffer for 45–120 min followed by three washes in TBS-T. Anti-HA and anti-MAF1 [106] antibodies were used at 1:1000 while anti-SAG1 was used at 1:2000 and rabbit anti-ROP5 [99] was used at 1:40,000. Anti-TgMAF1RHa1 and anti-TgMAFRHb1 antibodies generated for this study were used at a 1:10,000 dilution. Secondary antibody incubation was performed with horseradish peroxidase-conjugated secondary antibodies to the respective primary antibodies in blocking buffer for 45 min. Bands were visualized with SuperSignal West Pico chemiluminescent substrate (Thermo Scientific). Densitometric analysis was performed using ImageJ.

## 6.11 IN VITRO GROWTH ASSAY

96-well plates of confluent HFFs were infected with  $10^4$  parasites per well of luciferase expressing *T. gondii* (TgS1T, TgS22, TgS23) or *N. caninum* (NcNC1). Outermost wells were excluded to avoid edge effects due to media evaporation. Media was aspirated and cells were lysed using 50-100 µL of 1x cell lysis buffer (Promega) at 4, 12, 28, and 48 hours post infection. Cell lysate was transferred to a new 96-well plate and samples were stored at -80 degrees until all samples were collected. Once all time points were collected, samples were thawed at room temperature and centrifuged at 800 xg for 10 minutes to remove bubbles. 50 µL of the luciferase substrate, LARII (Promega), was added carefully to avoid introducing bubbles, and relative light units were measured using MicroWin2000 software.

#### 6.12 IN VITRO INTERFERON-GAMMA INHIBITION ASSAY

96-wells of confluent host cells (HFFs, MEFs, or RAW 264.7) were treated with increasing doses of human (HFFs) or mouse (MEFs or RAW 264.7) recombinant IFN- $\gamma$  for 24 hours before infection with *T. gondii* (TgS22, TgS23) or *N. caninum* (NcNC1). Parasites were collected 48 hours post infection and samples were analyzed as described above for the *in vitro* growth assay.

#### 6.13 ASSAY FOR GROWTH IN PRESENCE OF INTERFERON-GAMMA

96-wells were seeded with MEFs and grown under normal culture conditions (37° C, 5% CO<sub>2</sub>, CDMEM) until confluent. Once confluent cells were allowed to grow an additional 24-48 hours maximum prior to infection. Wells were infected with 10,000 parasites of *T. gondii* (TgS23) or *N. caninum* (NcNC1, Nc1:TgROP18<sub>II</sub>, NC1:TgROP5, or Nc1:TgROP18<sub>II</sub>:ROP5) in 100  $\mu$ L of pre-warmed CDMEM. After 4 hours of infection an additional 100  $\mu$ L of pre-warmed media was added to each well, with or without IFN $\gamma$  (200 U/mL, for a final concentration of 100 U/mL). Media was aspirated at the indicated time points cells were lysed using 100  $\mu$ L of 1x cell lysis buffer (Promega). Lysis buffer was pipetted up and down 3 times before being transferred to a fresh 96-well plate so that samples could be stored at -80° C until all time points were collected. Samples were analyzed as described above for the *in vitro* growth assay.

#### 6.14 H. HAMMONDI ROP16 PROMOTER ASSAY

Putative ROP16 promoters for *T. gondii* (881 bp upstream of the start codon) and *H. hammondi* (865 bp upstream of the start codon) were cloned into vectors upstream of a firefly luciferase coding region followed by a T. gondii dihydrofolate reductase 3' UTR. Mutant *T. gondii* and *H. hammondi* ROP16 promoter constructs were made using splicing by overlap extension PCR. We deleted 16 bp immediately preceding the putative transcriptional start site of TgROP16, and similarly inserted 16 bp from the TgROP16 promoter into the HhROP16 promoter. All promoter constructs were tested for activity by a dual-luciferase reporter assay.

15-20 million TgME49 parasites were transfected with 50-70  $\mu$ g of the ROP16 promoter construct along with 20  $\mu$ g of a vector containing a *Renilla* luciferase gene under a *T. gondii* tubulin promoter (to normalize for transfection efficiency between experiments). Parasites were then lysed using 1 x cell lysis buffer (Promega), and both Firefly and *Renilla* luciferase activity was measured using the Promega dual-luciferase reporter assay system.

#### 6.15 IN VITRO COMPETITION ASSAY

An ME49 strain engineered to express an N-terminal HA-tagged Type I (RH) MAF1 (ME49:TgMAF1RHb1) was mixed with ME49:WT in ratios 4:1 and 1:4. These two mixed populations were used to infect human foreskin fibroblast cells at an MOI of 3. Flasks were passed via syringe lysis every 3 days. At the 0, 4, and 8 week time mark HFFs grown on 12 mm glass coverslips were infected at an MOI of 3 and the proportion of HA+ and HA- parasites was calculated using immunofluorescence using rat  $\alpha$ -HA (as above) and serum from a mouse

chronically infected with *Toxoplasma gondii* at 1:1000 dilution. The ratio of HA+ to HA- was determined by counting at least 200 vacuoles. The entire experiment was repeated two times, each time with a genetically distinct clone set (WT and complemented).

#### 6.16 ANIMAL EXPERIMENTS

Mouse experiments were performed with 4-8wk old female Balb/c, CBA/J, C57BL6/J, or cytokine knockout mice in a C57Bl6/j background, obtained from Jackson Laboratories. All animal procedures in this study meet the standards of the American Veterinary Association and were approved locally under IACUC protocols #12010130 and #15015428.

#### 6.17 GENERATION OF POLYCLONAL ANTIBODIES

Female Balb/c mice were injected intraperitoneally (IP) with 100  $\mu$ g of either TgMAF1RHa1 antigen or TgMAF1RHb1 antigen suspended in 100  $\mu$ L phosphate buffered saline (PBS) and mixed 1:1 with Sigma Adjuvant (Sigma) to a final volume of 200  $\mu$ L. Additional injections of 50  $\mu$ g of the appropriate antigen mixed 1:1 with Sigma Adjuvant to a final volume of 200  $\mu$ L were administered 14, 35 and 56 days after the initial injection. Sera were collected prior to initial injection, as well as day 31, 81, and 88. All sera were tested for reactivity against both TgMAF1RHa1 and TgMAF1RHb1 by Western blot prior to use in western blots or immunofluorescence assays.

#### 6.18 SERUM COLLECTION

Blood samples were collected by tail or submandibular bleeding at specified time points. For each sample 20-50  $\mu$ L of whole blood was collected, incubated on ice for 1 hour, and centrifuged at 2500 x g for 10 minutes. Alternatively, blood was collected into serum separator tubes, allowed to clot for 5-10 minutes, and then centrifuged at 10,000 xg for 10 minutes. In all cases, serum supernatant was transferred to a new tube for further analysis, and the pelleted clot was discarded. For final time points, blood was collected after euthanasia by cardiac puncture for a total volume of 200-500  $\mu$ L, incubated for 1 hour on ice, and serum was collected as described above.

## 6.19 ANALYSIS OF CYTOKINE EXPRESSION

Serum was collected as described above and stored at -80° C until sent for Luminex. Samples were analyzed by The UPCI Cancer Biomarkers Facility: Luminex Core Laboratory (supported in part by award P30CA047904) for the presence of 32 mouse cytokines. The serum level of each cytokine was reported as fluorescence intensity minus background, and cytokine levels were analyzed using GraphPad PRISM software. Mice were grouped by parasite strain, and each mouse had between one and four time points used in the analysis. A two-way ANOVA was performed, followed by multiple comparisons using a Sidak correction. Comparisons were done between the means of mice infected with TgME49:EV or TgME49:TgMAF1RHb1 to determine significant differences between serum cytokine levels.

#### 6.20 PARASITE INFECTION AND BIOLUMINESCENCE IMAGING

Unless otherwise stated, all infections were achieved by intraperitoneal (IP) injection of quantified tachyzoites suspended in 200  $\mu$ L of PBS. Briefly, well infected flasks were washed with fresh CDMEM, then monolayers were scraped from the bottom of the flask and syringe lysed by serial passage through a 27 and then 25 gauge needle. Lysed parasites were pelleted by centrifugation at 800 x g for 10 minutes, CDMEM was aspirated, and the pellet was suspended in 1-3 mL PBS. Parasites were then quantified by counting on a hemocytometer and diluted with PBS to the appropriate number of parasites per 200  $\mu$ L. For parasite strains expressing luciferase, infections were then monitored by bioluminescence.

## 6.21 IN VIVO COMPETITION ASSAY

Using the same genetically engineered ME49 clone sets, we again created mixed populations at ratios of 1:4, 1:1, 4:1, 100% ME49:TgMAF1RHb1, and 100% ME49:WT. We injected 100,000 tachyzoites intraperitoneally of these 5 populations into Balb/c mice in 200  $\mu$ L of PBS (3-5 mice per population). In a separate experiment we transfected ME49 with the same pTgMAF1RHa1, and grew the population under MPA/Xanthine selection for 2 weeks, and then injected 100,000 tachyzoites of this mixed population into Balb/c mice as above. On the day of injection, we used the same parasite preparation to infect HFFs seeded on glass coverslips to quantify the exact input proportions using IF imaging as described above. Parasite burden and location was assessed daily for the next five days using *in vivo* bioluminescence imaging [164] since the parental ME49 strain expressed click beetle luciferase off of a DHFR promoter [164]. On day 5

pi, all mice were sacrificed and an intraperitoneal lavage was performed in order to harvest peritoneal cells and associated parasites. Samples were spun down and resuspended in cDMEM and used to infect HFFs. After 1 passage parasites were used to infect HFFs seeded onto glass coverslips coverslips at an MOI of 3 to quantify proportions using IF imaging as above.

## 6.22 CYST PURIFICAION

Mice were euthanized and whole brains were removed. Either the entire brain or the left half was stored in PBS with 3% fetal bovine serum (3% FBS) on ice and then at 4° C until processing. Brain tissue was homogenized by passage through a 100 µm nylon cell strainer and suspended in 3% FBS in a 50 mL conical tube. Brain homogenate was pelleted by centrifugation at 1500 x g for 15 minutes and 4° C. Excess 3% FBS was removed, leaving ~5 mL of 3% FBS and brain homogenate. Cysts were then purified using a Percoll gradient of 5 mL each of 90%/40%/20% Percoll diluted in PBS layered into a new 50 mL conical tube. Brains were resuspended in PBS/FBS to a total volume of 5 mL, passaged twice through an 18 gauge needle, and then layered dropwise on top of the gradient. Gradients were then subjected to centrifugation at 1500 x g for 15 minutes at 4° C. Following centrifugation the top layer of brain tissue and the 90% Percoll layers were removed. Additional 3% FBS was added to the remaining Percoll and 3% FBS layers, disrupting the gradient and bringing the total volume to 45 mL. The tubes were inverted 2-3 times to insure complete disruption of the gradient before centrifugation at 1500 x g for 15 minutes at 4° C. The top 40 mL was then removed, leaving 5 mL of 3% FBS and cysts. This 5 mL was mixed and transferred to a new 15 mL conical tube, and 10 mL of 3% FBS was added before inverting 2-3 times to mix before centrifugation at 1500 x g for 15 minutes at 4° C.

The top 14.5 mL was then removed and the remaining 500  $\mu$ L was mixed before transferring to a 1.5 mL Eppendorf tube. An additional 1 mL of 3% FBS was added and tubes were inverted 2-3 times to mix before centrifugation at 2000 x g for 15 minutes. 3% FBS was carefully removed by pipette leaving ~200  $\mu$ L of 3% FBS and cysts at the bottom of the tube. Purified cysts were then stored at 4° C until staining or oral infection.

# 6.23 CYST STAINING AND QUANTIFICAION

A portion of whole brain homogenate or purified cysts were pelleted before resuspension in 100-250 uL 3% /FBS and incubated at room temperation for 1 hour. Following 1 hour incubation, rhodamine conjugated *Dolichos biflorus* Agglutinin was added (1:500 dilution) and samples were incubated for 1 hour at room temperature or 4° C overnight. Samples were then washed 2x by adding 1 mL 3% FBS, inverting 6-10 times, and then pelleting at 2,000 xg for 10 minutes. Following the second wash, stained cysts were suspended in 200 uL and 50 uL aliquots were placed into wells of a 96-well dish. Cysts were counted using an inverted fluorescent microscope, and estimated cysts per brain were calculated based on portion of the whole brain used for quantification.

# **APPENDIX** A

# A COMPARISON OF INTERFERON-GAMMA INHIBITION OF *N. CANINUM* AND *T. GONDII* GROWTH IN CELL CULTURE

Despite approximately 85% genomic syntenny [189], *T. gondii* and *N. caninum* have dramatically different host ranges and disease outcomes in various hosts [199]. Disease progression in a mouse model exemplifies this disparity, as even the least virulent strains of *T. gondii* (TgS22 or TgS1T; both lack all known virulence factors [96]) cause disease and morbidity when a mouse is infected with a high enough dose (>10<sup>6</sup> parasites), while *N. caninum* does not cause disease or morbidity, even when mice are infected with >10<sup>7</sup> parasites [199]. In cell culture, *T. gondii* and *N. caninum* are both able to invade host cells, grow and divide, and eventually lyse their host cells in order to then infect new host cells. *In vitro* growth rates of these parasites are quite similar, despite dramatic differences *in vivo*. The molecular and genetic causes underlying the differences between *T. gondii* and *N. caninum in vivo* remain elusive, in part due to the lack of an *in vitro* screening process in order to identify candidate genes.

In order to develop an *in vitr*o screen to identify candidate genes responsible for the differences between *T. gondii* and *N. caninum* disease progression in a mouse model, I measured the growth of several *T. gondii* strains (avirulent strain TgS22; virulent strain TgS23) in a variety

of cell culture conditions and compared it to the growth of *N. caninum* under the same conditions. Initial comparisons of growth in human foreskin fibroblasts (HFFs) at 37°C in 5%  $CO_2$  confirmed that under normal cell culture conditions in HFFs, *T. gondii* and *N. caninum* grow at similar rates *in vitro* (Figure 6-1).

However, growth conditions in HFFs are dramatically different from growth conditions in a mouse, which has many different cell and tissue types, as well as both an innate and adaptive immune response. In order to better simulate an immune response *in vitro*, I stimulated host cells with recombinant purified Interferon- $\gamma$  (IFN $\gamma$ ) before or after infecting with *T. gondii* or *N. caninum*. I did this using HFFs, as well as murine embryonic fibroblasts (MEFs) and a cell line derived from murine macrophage isolated from blood (RAW 264.7). In all three cell types, both *T. gondii* and *N. caninum* growth was inhibited by IFN $\gamma$  pre-treatment. Regardless of cell type or IFN $\gamma$  dose, growth was similarly inhibited for all parasites (Figure 6-2). Inhibition of parasite growth was greater with higher doses of IFN $\gamma$ , except in the case of RAW cells where even the smallest dose of IFN $\gamma$  inhibited parasite growth as well as the highest dose (Figure 6-2E). Treating the cells with IFN $\gamma$  post-infection yielded varying results. Overall, parasite growth was inhibited less by post-treatment than pre-treatment. Interestingly, *N. caninum* growth was the most inhibited, compared to both *T. gondii* strains, in post-treated HFFs, but the least inhibited in post-treated RAW cells (Figure 6-2B,F)

To determine the impact of IFN $\gamma$  treatment on growth rate, I conducted a growth assay in HFFs, where parasites were allowed to invade host cells for 4 hours prior to IFN $\gamma$  treatment. Consistent with the previous data, *T. gondii* growth was minimally inhibited by IFN $\gamma$  treatment of host cells 4 hours post infection, while *N. caninum* growth was greatly inhibited (Figure 6-3). These data are consistent with the recent identification of a *T. gondii* protein (TgIST) responsible

for the suppression of interferon-stimulated transcription that is absent in *N. caninum* [200, 201]. Future studies will include the transgenic expression of TgIST in *N. caninum* parasites to determine if TgIST expression is responsible for the observed difference in *T. gondii* and *N. caninum* growth in the presence of IFN<sub>γ</sub>.

Another recently published study suggested that TgROP18, which is a pseudogene in *N. caninum*, is responsible for virulence differences *in vivo* [202]. While other members of the lab are conducting experiments to determine if TgROP18 expression in the *N. caninum* strain NC1 increases *N. caninum* virulence *in vivo*, I have been testing if TgROP18 expression impacts *in vitro* growth in the presence of IFN $\gamma$ . To do this I have conducted growth assays of transgenic *N. caninum* expressing TgROP18 cloned from type II *T. gondii* (TgROP18<sub>II</sub>) in the presence and absence of IFN $\gamma$ . Transgenic expression of TgROP18<sub>II</sub> in *N. caninum* does not alter parasite growth inhibition by IFN $\gamma$  (Figure 6-4A). Because TgROP18 works in concert with TgROP5 in *T. gondii* [100, 203], I also conducted growth assays with transgenic *N. caninum* expressing TgROP5, as well as transgenic *N. caninum* expressing both TgROP18<sub>II</sub> and TgROP5. Interestingly, when TgROP5 was expressed in *N. caninum*, there was a lack of IFN $\gamma$  mediated growth inhibition (Figure 6-4B), however when both TgROP18<sub>II</sub> and TgROP5 were expressed together, I once again observed *N. caninum* growth inhibition in the presence of IFN $\gamma$  (Figure 6-4C). This experiment will need to be repeated in order to confirm these results.



## Figure 6-1. T. gondii and N. caninum growth rates in vitro.

*In vitro* growth assay of TgS22, TgS23, and NcNC1. Parasites expressing luciferase were grown in human foreskin fibroblasts for the indicated amount of time, lysed and exposed to the luciferin substrate. Relative light units were used as a readout for parasite growth.



Figure 6-2. T. gondii and N. caninum growth inhibition by IFNy in vitro.

Host cells were infected with a virulent strain of *T. gondii* (S23), an avirulent strain of *T. gondii* (S22) or *N. caninum* (NC1) in the presence of increasing IFN<sub>Y</sub> concentrations. Host cells were treated 12 hours before infection (A, C, E)

or 12 hours after infection (B, D, F). Parasite growth was measured 48 hours post infection by bioluminescence and normalized to an untreated control. Pretreatment of all three host cell types (HFFs, MEFs, and RAW 264.7 cells) with IFN<sub>γ</sub> inhibited growth of all three parasite strains.



Figure 6-3. *N. caninum* growth is inhibited in the presence of IFN<sub>γ</sub>.

96-well plates of confluent HFFs were infected with 10,000 parasites per well of either *T. gondii* (S23) or *N. caninum* (NC1). Parasites were allowed to infect host cells for 4 hours before the addition of IFN $\gamma$  (dose = 100 units), or additional pre-warmed media for control wells. Parasites were quantified by bioluminescence for each treatment at the indicated time points (n=3). Exponential growth curves were fit to the data using GraphPad Prism. Using an extra sum-of-squares F test, the k values for the curves fitting *T. gondii* growth in the presence or absence of IFN $\gamma$  were not significantly different (p = 0.578), however the k values for the curves fitting *N. caninum* in the presence or absence of IFN $\gamma$  were significantly different (p = 0.014)



Figure 6-4. Impact of transgenic TgROP18<sub>II</sub> and TgROP5 expression on *N. caninum* growth inhibition in the presence of IFN $\gamma$ .

Growth assays were conducted and analyzed in the same manner as in Figure 6-3, this time with transgenic *N*. *caninum* expressing TgROP18<sub>II</sub> (Nc1:TgROP18<sub>II</sub>), TgROP5 (Nc1:TgROP5) or both together (Nc1:TgROP18<sub>II</sub>:TgROP5). The k values for the curves fit to Nc1:TgROP18<sub>II</sub> and Nc1:ROP18<sub>II</sub>:ROP5 in the presence or absence of IFN $\gamma$  were significantly different (p = 0.005 and 0.0007, respectively), however the k values for the curves fir to Nc1:TgROP5 in the presence or absence of IFN $\gamma$  were not significantly different (p = 0.740).

## **APPENDIX B**

# H. HAMMONDI ROP16 PROMOTER CHARACTERIZAION

Despite remarkable similarity at the genomic level, *Toxoplasma gondii* and *Hammondia hammondi* have vastly different host ranges and disease outcomes in infected intermediate hosts. The factors, both parasite and host, contributing to these differences remain largely unknown. Previous studies utilizing genetic and phenotypic variation between strains of *T. gondii* have identified a number of secreted parasite proteins that interact with and modify the host cell, thus contributing to parasite virulence [96, 102, 204]. To determine if these identified factors also contribute to phenotypic differences between *T. gondii* and *H. hammondi*; studies have assessed the ability of *H. hammondi* copies of each of the identified virulence factors ROP5, ROP16, ROP18, and GRA15 to complement *T. gondii* strains lacking each factor individually. Initial studies indicate that *H. hammondi* copies of both ROP5 and ROP18 are functionally conserved when expressed in *T. gondii* [164]. This appendix contains work that was done to aid in determining if *H. hammondi* harbors a functional copy of ROP16, and was published along with work identifying a functional *H. hammondi* GRA15 [205].

Comparisons of the putative promoter region of ROP16 (881 bp upstream of the start codon) in *T. gondii* types I, II, and III to the putative promoter region of the *H. hammondi* 

ROP16 (865 bp upstream of the start codon) revealed a 16 base pair deletion just upstream of the predicted transcriptional start site. To determine if this deletion altered the ability of this putative promoter region to drive gene expression, upstream regions from *H. hammondi* and *T. gondii* were cloned in front of a firefly luciferase reporter gene and tested for the ability to drive luciferase expression. Constructs were co-transfected with a plasmid containing a tubulin promoter driven *Renilla* luciferase to control for transfection efficiency, and promoter activity is represented as Firefly over *Renilla* activity.

The wildtype promoter region cloned from *T. gondii* (type II strain, ME49) was able to drive significantly higher expression of Firefly luciferase compared to the wildtype promoter region from *H. hammondi* (Figure 6-5; P = 0.008). Additionally, the deletion of those 16 bp from the *T. gondii* promoter significantly reduced promoter activity compared to the wildtype *T. gondii* promoter (P = 0.01). The addition of the 16 bp region to the *H. hammondi* promoter did increase promoter activity, however this increase was not significant when compared to the wildtype *H. hammondi* ROP16 promoter activity (P = 0.3).

These data suggest that the *H. hammondi* ROP16 promoter is not functional in a *T. gondii* background, and that the 16 bp deletion just upstream of the transcriptional start site is at least partially responsible for the lack of expression. Therefore, further studies assessing the functionality of *H. hammondi* ROP16 were done using a construct where *H. hammondi* ROP16 was expressed under the GRA1 promoter [205].



Figure 6-5. Deletion of 16 bp upstream of ROP16 transcriptional start site disrupts expression.

(A) Alignment of sequence upstream and including the ROP16 transcriptional start site in *T. gondii* and *H. hammondi*, illustrating the 16 bp deletion in the *H. hammondi* sequence. Arrow indicates predicted transcription start site. (B) Putative ROP16 promoters from type II T. gondii and H. hammondi were cloned upstream of Firefly luciferase and co-transfected with a plasmid containing *Renilla* luciferase under a tubulin promoter into wildtype ME49 (T. gondii; type II). Firefly and *Renilla* luciferase expression levels were measured using a dual luciferase reporter assay system. Promoter activity is represented by the ratio of Firefly over *Renilla* activity. Means with the same letter are not significantly different (P > 0.05).

# **APPENDIX C**

## LOCALIZATION OF TGME49\_203290

While in some model organisms, such as the budding yeast *Saccharomyces cerevisiae*, the majority of protein coding genes within the genome have known or predicted functions [206], most putative protein coding genes within the genome of *T. gondii* have no known or predicted function. Some of these putative protein coding genes have signal peptides or motifs that indicate they may be expressed and secreted by the parasite in order to interact with and modulate various host cell signaling pathways. One such predicted protein coding gene is annotated on ToxoDB.org as TGME49\_203290 (203290), which is differentially expressed across *T. gondii* strains, with higher expression in type II strains compared to types I and III. To begin studies on 203290, I cloned and C-terminally HA-tagged 03290 from both a type II strain (ME49) and a type III strain (CTG) and expressed them in opposite backgrounds (the type II version expressed in a type III strain and vice versa) to determine localization.

Interestingly, despite evidence that the type III version of 203290 may not be expressed, both the type II 203290 expressed in a type III background and type III 203290 expressed in a type II background were expressed and visualized by HA-staining (Figure 6-6). Both versions of 203290 localize to the vacuole surrounding the parasites and partially co-localize with the dense granule protein GRA7, which is consistent with the recent identification of 203290 as a dense granule protein (now designated GRA34; [207]).



Figure 6-6. Localization of 203290.

Type II (ME49) or type III (CTG) parasites were transfected with a C-terminally HA-tagged 203290 cloned from type III (CTG) or type II (ME49) genomic DNA, respectively. Transfected parasites were then stained by immunofluorescence using an antibody against HA. Parasites were also stained for the dense granule protein GRA7.

#### APPENDIX D

## **RAW LUMINEX DATA FROM CHAPTER 4**

This appendix contains data obtained from the Luminex 32-plex mouse cytokine panel of TgME49Lev and TgME49:TgMAF1RHb1 infected C57BL6/J mice pre-infection as well as 21, 28 and 57 DPI. Cytokine measurement is given in fluorescent intensity minus background. The cytokines in this panel include: Granulocyte colony-stimulating factor (G-CSF), Eotaxin, granulocyte macrophage colony-stimulating factor (GM-CSF), Interferon-gamma (IFNg), Interleukin- (IL) 1 alpha (IL-1a), IL-1beta (IL-1b), IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-12 beta (IL-12p40), IL-12 heterodimer (IL-12p70), IL-13, IL-15, IL-17, Interferon-gamma induced protein 10 (IP-10), keratinocyte chemoattractant (KC; a.k.a. Cxcl1), leukemia inhibitory factor (LIF), lipopolysaccharide-induced CXC cytokine (LIX; a.k.a. Cxcl5), monocyte chemoattractant cytokine 1 (MCP-1), macrophage colony-stimulating factor (M-CSF), monokine induced by gamma interferon (MIG), macrophage inflammatory protein- (MIP) 1 alpha (MIP-1a), MIP-1 beta (MIP-1b), MIP-2, RANTES, tumor necrosis factor alpha (TNFa), and vascular endothelial growth factor (VEGF). "†" denotes data not obtained as mouse became morbid during acute infection. "---" denotes data not obtained as mouse was euthanized 28 DPI. Mouse 2-2 and 3-2 appeared uninfected by bioluminescence and lack of IFNg (\*).

Table 6. G-CSF

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	208	58	176.5	
1-2	TgME49:TgMAF1RHb1	295.5	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	166.5	137.5	564	
1-4	TgME49:TgMAF1RHb1	150.5	184	98	
1-5	TgME49:TgMAF1RHb1	134.5	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	91.5	202.5	398.5	167.5
2-2*	TgME49:TgMAF1RHb1	232.5	94.5	106.5	164
2-3	TgME49:TgMAF1RHb1	71	114.5	140.5	260.5
2-4	TgME49:TgMAF1RHb1	105	106.5	97.5	95.5
2-5	TgME49:TgMAF1RHb1	77.5	Ť	Ŧ	Ť
3-1	TgME49:EV	350	122	232.5	
3-2*	TgME49:EV	157.5	101	133	
3-3	TgME49:EV	72.5	120	22	
3-4	TgME49:EV	184.5	74.5	63.5	
3-5	TgME49:EV	34.5	30.5	83.5	
4-1	TgME49:EV	266.5	70.5	115.5	138
4-2	TgME49:EV	99	51.5	68	221
4-3	TgME49:EV	89.5	21.5	65.5	46
4-4	TgME49:EV	206.5	150.5	291.5	288.5
4-5	TgME49:EV	228.5	89.5	141.5	149.5

Table 7. Eotaxin

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	563	5966	7355	
1-2	TgME49:TgMAF1RHb1	200	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	267.5	5423	6738	
1-4	TgME49:TgMAF1RHb1	132	5459.5	8090.5	
1-5	TgME49:TgMAF1RHb1	171	Ť	ŧ	ŧ
2-1	TgME49:TgMAF1RHb1	149	2330	2372	8840
2-2*	TgME49:TgMAF1RHb1	212	1837.5	641	9127.5
2-3	TgME49:TgMAF1RHb1	53	541	1121	7800.5
2-4	TgME49:TgMAF1RHb1	770	3541.5	2597	6440
2-5	TgME49:TgMAF1RHb1	478.5	Ť	Ť	Ť
3-1	TgME49:EV	664	2872.5	6697	
3-2*	TgME49:EV	407.5	326	2828.5	
3-3	TgME49:EV	123	1986.5	4293.5	
3-4	TgME49:EV	284	3006.5	4248.5	
3-5	TgME49:EV	205	1191	2359.5	
4-1	TgME49:EV	179.5	591.5	1375	7617
4-2	TgME49:EV	201	1595	1850	5198
4-3	TgME49:EV	192.5	1049	1846.5	4030
4-4	TgME49:EV	429	680	1827	9534
4-5	TgME49:EV	564.5	693	1151	3412

Table 8. GM-CSF

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	3	0	-2	
1-2	TgME49:TgMAF1RHb1	2	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	-2	-2	1	
1-4	TgME49:TgMAF1RHb1	0	-2	-5	
1-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	1	-3	-5	3
2-2*	TgME49:TgMAF1RHb1	-3	-5	-6	5
2-3	TgME49:TgMAF1RHb1	0	-6	-8	6
2-4	TgME49:TgMAF1RHb1	-2	-0.5	-1	-2
2-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
3-1	TgME49:EV	4	2	5	
3-2*	TgME49:EV	1	-2	-4	
3-3	TgME49:EV	-4	10	1	
3-4	TgME49:EV	0	-5.5	-11	
3-5	TgME49:EV	-2	-2	-5	
4-1	TgME49:EV	-2	-4	-7	1
4-2	TgME49:EV	-6	-8	-9	2
4-3	TgME49:EV	2	-4	-4	1
4-4	TgME49:EV	1	-3	-4	5
4-5	TgME49:EV	2	-2	-4	2

Table 9. IFNg

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	<b>57 DPI</b>
1-1	TgME49:TgMAF1RHb1	4	265.5	938.5	
1-2	TgME49:TgMAF1RHb1	0	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	1	135.5	676.5	
1-4	TgME49:TgMAF1RHb1	1	119.5	32.5	
1-5	TgME49:TgMAF1RHb1	0	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	1	220	110	301
2-2*	TgME49:TgMAF1RHb1	1.5	-1	0	10
2-3	TgME49:TgMAF1RHb1	3	25	18	237.5
2-4	TgME49:TgMAF1RHb1	0	164	86	27.5
2-5	TgME49:TgMAF1RHb1	3	Ť	Ŧ	Ť
3-1	TgME49:EV	2	138.5	545.5	
3-2*	TgME49:EV	0	2.5	3.5	
3-3	TgME49:EV	2	273.5	134.5	
3-4	TgME49:EV	1	155	95	
3-5	TgME49:EV	-2	223	271	
4-1	TgME49:EV	0	193	40	405.5
4-2	TgME49:EV	2	130.5	43	378
4-3	TgME49:EV	-5	65	57	6
4-4	TgME49:EV	4	183.5	141	124
4-5	TgME49:EV	4	64	96	310

Table 10. IL-1a

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	<b>57 DPI</b>
1-1	TgME49:TgMAF1RHb1	15	15	61	
1-2	TgME49:TgMAF1RHb1	25	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	104	129	134.5	
1-4	TgME49:TgMAF1RHb1	30	160	61	
1-5	TgME49:TgMAF1RHb1	18	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	14	33	37	65
2-2*	TgME49:TgMAF1RHb1	39.5	311	24	148.5
2-3	TgME49:TgMAF1RHb1	30	8	31	116.5
2-4	TgME49:TgMAF1RHb1	86	72	286	50.5
2-5	TgME49:TgMAF1RHb1	77	Ť	Ť	Ť
3-1	TgME49:EV	73	68	33	
3-2*	TgME49:EV	71	43	23	
3-3	TgME49:EV	84	57	20	
3-4	TgME49:EV	16	54	46	
3-5	TgME49:EV	6	33	24	
4-1	TgME49:EV	23	18.5	66	81
4-2	TgME49:EV	37	44	134	126
4-3	TgME49:EV	19	12.5	16	69.5
4-4	TgME49:EV	109.5	26	56	129
4-5	TgME49:EV	94	23	84	110.5

Table 11. IL-1b

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	0	2	1	
1-2	TgME49:TgMAF1RHb1	1.5	Ť	Ŧ	Ť
1-3	TgME49:TgMAF1RHb1	4	3	3	
1-4	TgME49:TgMAF1RHb1	0	4	3	
1-5	TgME49:TgMAF1RHb1	-1	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	0	0	1	4
2-2*	TgME49:TgMAF1RHb1	1	0	-1	3
2-3	TgME49:TgMAF1RHb1	1	2	1	3
2-4	TgME49:TgMAF1RHb1	1	2	0	-2
2-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
3-1	TgME49:EV	2	9	9	
3-2*	TgME49:EV	1	1	1	
3-3	TgME49:EV	0	3	2	
3-4	TgME49:EV	-1	0	2	
3-5	TgME49:EV	0	0	0	
4-1	TgME49:EV	0	1	0	1
4-2	TgME49:EV	1	1	1	2
4-3	TgME49:EV	0	0	0	3
4-4	TgME49:EV	14	0	1	1
4-5	TgME49:EV	2	-1	1	1

Table 12. IL-2

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	1	6	13	
1-2	TgME49:TgMAF1RHb1	3	Ť	Ť	†
1-3	TgME49:TgMAF1RHb1	1	4	42	
1-4	TgME49:TgMAF1RHb1	-1	5	16	
1-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	-1	2	0	34
2-2*	TgME49:TgMAF1RHb1	0	0	-1	17
2-3	TgME49:TgMAF1RHb1	0	-1	-1	49
2-4	TgME49:TgMAF1RHb1	1	3	1	12
2-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
3-1	TgME49:EV	3	5	5	
3-2*	TgME49:EV	0	2	3	
3-3	TgME49:EV	-1	4	5	
3-4	TgME49:EV	-1	-3	0	
3-5	TgME49:EV	-2	-2	-3	
4-1	TgME49:EV	1	1	0	10
4-2	TgME49:EV	-0.5	0	1	9
4-3	TgME49:EV	0	0	-2	3
4-4	TgME49:EV	2	45	15	36
4-5	TgME49:EV	1	-1	0	5

Table 13. IL-3

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	2	-1	-0.5	
1-2	TgME49:TgMAF1RHb1	2	Ŧ	Ť	Ť
1-3	TgME49:TgMAF1RHb1	0	-1	0	
1-4	TgME49:TgMAF1RHb1	0	-2	0	
1-5	TgME49:TgMAF1RHb1	-1	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	0	2	0	2
2-2*	TgME49:TgMAF1RHb1	0	-4	-4	1
2-3	TgME49:TgMAF1RHb1	-1	-3	-5	8
2-4	TgME49:TgMAF1RHb1	-1	-2	-1	-4
2-5	TgME49:TgMAF1RHb1	0	t	Ť	Ť
3-1	TgME49:EV	0	-1	-1	
3-2*	TgME49:EV	2	-3	-3	
3-3	TgME49:EV	-2	1	-2	
3-4	TgME49:EV	-3	-3	-4	
3-5	TgME49:EV	-2	-1	-2	
4-1	TgME49:EV	-4	-2	-3	0
4-2	TgME49:EV	-3	-4	-3	-3
4-3	TgME49:EV	1	-4	0	2.5
4-4	TgME49:EV	1	0	-1	3
4-5	TgME49:EV	2	-3	-3	2

Table 14. IL-4

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	<b>57 DPI</b>
1-1	TgME49:TgMAF1RHb1	2	-1	-1	
1-2	TgME49:TgMAF1RHb1	8.5	Ť	Ť	Ŧ
1-3	TgME49:TgMAF1RHb1	7.5	1.5	7	
1-4	TgME49:TgMAF1RHb1	6.5	21	19	
1-5	TgME49:TgMAF1RHb1	4.5	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	4.5	0.5	0.5	5.5
2-2*	TgME49:TgMAF1RHb1	1.5	-1.5	-3.5	7.5
2-3	TgME49:TgMAF1RHb1	4.5	-9.5	-7.5	7.5
2-4	TgME49:TgMAF1RHb1	0.5	3.5	2.5	0.5
2-5	TgME49:TgMAF1RHb1	3.5	Ť	Ť	Ť
3-1	TgME49:EV	0	25	32	
3-2*	TgME49:EV	3.5	1	2	
3-3	TgME49:EV	2	0	3	
3-4	TgME49:EV	2	5.5	3.5	
3-5	TgME49:EV	1.5	-3.5	-6.5	
4-1	TgME49:EV	1.5	1.5	1.5	3.5
4-2	TgME49:EV	6.5	0.5	-2.5	7.5
4-3	TgME49:EV	4.5	-0.5	-6.5	0.5
4-4	TgME49:EV	142.5	0.5	0.5	8.5
4-5	TgME49:EV	3.5	0.5	-2.5	6.5

Table 15. IL-5

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	13	9	17	
1-2	TgME49:TgMAF1RHb1	13	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	6	22.5	18	
1-4	TgME49:TgMAF1RHb1	3	31.5	8	
1-5	TgME49:TgMAF1RHb1	7	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	8	20	10	15
2-2*	TgME49:TgMAF1RHb1	12	4	2	16
2-3	TgME49:TgMAF1RHb1	4	9	52	16
2-4	TgME49:TgMAF1RHb1	10	24	10	2
2-5	TgME49:TgMAF1RHb1	41	Ť	t	†
3-1	TgME49:EV	24	45	11	
3-2*	TgME49:EV	4	16	14	
3-3	TgME49:EV	5	21	22	
3-4	TgME49:EV	7	9	23	
3-5	TgME49:EV	12	9	7	
4-1	TgME49:EV	6	4.5	0	12
4-2	TgME49:EV	23	21	17	134.5
4-3	TgME49:EV	8	19	13	66
4-4	TgME49:EV	151	14	9.5	22
4-5	TgME49:EV	14	26	7	31
Table 16. IL-6

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	6.5	31	144	
1-2	TgME49:TgMAF1RHb1	16	Ŧ	Ŧ	Ť
1-3	TgME49:TgMAF1RHb1	8	27	145	
1-4	TgME49:TgMAF1RHb1	2.5	39	27	
1-5	TgME49:TgMAF1RHb1	3	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	4	34	27	44
2-2*	TgME49:TgMAF1RHb1	5	2	0	12
2-3	TgME49:TgMAF1RHb1	4	20	6	82.5
2-4	TgME49:TgMAF1RHb1	1	18	3	32
2-5	TgME49:TgMAF1RHb1	-1	Ť	Ť	†
3-1	TgME49:EV	8	34	36	
3-2*	TgME49:EV	6	0	4	
3-3	TgME49:EV	2	44	20.5	
3-4	TgME49:EV	2	11	21	
3-5	TgME49:EV	0	20	11.5	
4-1	TgME49:EV	17	15	5	52
4-2	TgME49:EV	0	8	1	29
4-3	TgME49:EV	5	21	23	56.5
4-4	TgME49:EV	42	23	13	86
4-5	TgME49:EV	3	6	9	30

Table 17. IL-7

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	-1	1	8	
1-2	TgME49:TgMAF1RHb1	-3.5	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	-3.5	10.5	13	
1-4	TgME49:TgMAF1RHb1	-2.5	-5	8.5	
1-5	TgME49:TgMAF1RHb1	-7.5	Ť	Ť	ŧ
2-1	TgME49:TgMAF1RHb1	-13.5	-5.5	-3.5	13.5
2-2*	TgME49:TgMAF1RHb1	-5.5	-11.5	-11.5	14.5
2-3	TgME49:TgMAF1RHb1	-6.5	-15.5	-2.5	45.5
2-4	TgME49:TgMAF1RHb1	-9.5	-1	-6.5	10.5
2-5	TgME49:TgMAF1RHb1	-4.5	Ť	Ť	Ť
3-1	TgME49:EV	1	1	40	
3-2*	TgME49:EV	-5.5	0	2	
3-3	TgME49:EV	-8.5	-1.5	93.5	
3-4	TgME49:EV	-8.5	11.5	6.5	
3-5	TgME49:EV	9.5	28.5	11.5	
4-1	TgME49:EV	-9.5	-8.5	-4.5	13.5
4-2	TgME49:EV	-8.5	10.5	-9.5	6.5
4-3	TgME49:EV	-6.5	-12.5	-6.5	7.5
4-4	TgME49:EV	-3.5	5.5	-5.5	56.5
4-5	TgME49:EV	-2.5	-8.5	-7.5	31.5

Table 18. IL-9

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	3	5	10	
1-2	TgME49:TgMAF1RHb1	-12	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	4	-10	46.5	
1-4	TgME49:TgMAF1RHb1	-1	-4	60	
1-5	TgME49:TgMAF1RHb1	-13.5	Ť	Ť	ŧ
2-1	TgME49:TgMAF1RHb1	-15	-12	-10	43
2-2*	TgME49:TgMAF1RHb1	-4	-6.5	-10	39
2-3	TgME49:TgMAF1RHb1	5	-9.5	4	136
2-4	TgME49:TgMAF1RHb1	-4.5	-15	-14	4
2-5	TgME49:TgMAF1RHb1	-4	Ť	Ť	Ť
3-1	TgME49:EV	10	-2	22	
3-2*	TgME49:EV	7	3	9	
3-3	TgME49:EV	-6	6.5	29.5	
3-4	TgME49:EV	-5	-15	-4	
3-5	TgME49:EV	-16	-3	13	
4-1	TgME49:EV	-10.5	-7	-14	7
4-2	TgME49:EV	-11	-13	-14.5	31
4-3	TgME49:EV	-15	-10	-1	15
4-4	TgME49:EV	-13	-3.5	-9	6
4-5	TgME49:EV	2	-12	-6	13

**Table 19. IL-10** 

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	27.5	26.5	19.5	
1-2	TgME49:TgMAF1RHb1	31	Ŧ	Ŧ	Ť
1-3	TgME49:TgMAF1RHb1	37.5	46	53.5	
1-4	TgME49:TgMAF1RHb1	43	28.5	24.5	
1-5	TgME49:TgMAF1RHb1	53	Ŧ	Ť	Ť
2-1	TgME49:TgMAF1RHb1	44	46	43	8
2-2*	TgME49:TgMAF1RHb1	27	20	22	7
2-3	TgME49:TgMAF1RHb1	59	21	27	14
2-4	TgME49:TgMAF1RHb1	40	50	54.5	1
2-5	TgME49:TgMAF1RHb1	48	Ŧ	Ť	Ŧ
3-1	TgME49:EV	32.5	30.5	30.5	
3-2*	TgME49:EV	45	26.5	9	
3-3	TgME49:EV	46.5	39.5	29.5	
3-4	TgME49:EV	38	32.5	-4	
3-5	TgME49:EV	31	39	12	
4-1	TgME49:EV	41	14	-14	3
4-2	TgME49:EV	40	24	-14.5	3
4-3	TgME49:EV	45	23	-1	6
4-4	TgME49:EV	68	48	-9	7
4-5	TgME49:EV	42	19	-6	6

Table 20. IL-12p40

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	-4	9	35	
1-2	TgME49:TgMAF1RHb1	12	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	15	16	52	
1-4	TgME49:TgMAF1RHb1	12	5	35	
1-5	TgME49:TgMAF1RHb1	11	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	2	3	-8	5
2-2*	TgME49:TgMAF1RHb1	10	4.5	5	14
2-3	TgME49:TgMAF1RHb1	19	6	10	25
2-4	TgME49:TgMAF1RHb1	4	16	6.5	3
2-5	TgME49:TgMAF1RHb1	2	Ť	Ť	Ť
3-1	TgME49:EV	-4	7	3	
3-2*	TgME49:EV	11	2	6	
3-3	TgME49:EV	13	-1	27	
3-4	TgME49:EV	11	-2	2	
3-5	TgME49:EV	-5.5	4	6	
4-1	TgME49:EV	7	-1	5	9
4-2	TgME49:EV	15	-4	-2.5	12
4-3	TgME49:EV	11	1	8	12
4-4	TgME49:EV	13	11	22	35.5
4-5	TgME49:EV	16	-3	1	9

Table 21. IL-12p70

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	1	1	3	
1-2	TgME49:TgMAF1RHb1	3	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	1	-1	3	
1-4	TgME49:TgMAF1RHb1	-0.5	17	5	
1-5	TgME49:TgMAF1RHb1	0	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	0	0	-1	6
2-2*	TgME49:TgMAF1RHb1	0	-2	-2	16
2-3	TgME49:TgMAF1RHb1	0	-2	-3	6
2-4	TgME49:TgMAF1RHb1	-1	-1	-1	5
2-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
3-1	TgME49:EV	2	5	9	
3-2*	TgME49:EV	0	-1	0	
3-3	TgME49:EV	-1	1	1	
3-4	TgME49:EV	-1	1	1	
3-5	TgME49:EV	-1	-1	-2	
4-1	TgME49:EV	-1	0	-1	2
4-2	TgME49:EV	-1	0	0	3
4-3	TgME49:EV	-1	-2	-2	3
4-4	TgME49:EV	24	2	-2	9
4-5	TgME49:EV	1	-1	-1	5

**Table 22. IL-13** 

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	26	26	39	
1-2	TgME49:TgMAF1RHb1	49	Ŧ	Ť	Ť
1-3	TgME49:TgMAF1RHb1	30	34	55	
1-4	TgME49:TgMAF1RHb1	33	28	53.5	
1-5	TgME49:TgMAF1RHb1	30	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	28	30	30	49
2-2*	TgME49:TgMAF1RHb1	29	23	24	32.5
2-3	TgME49:TgMAF1RHb1	33	18	16	284
2-4	TgME49:TgMAF1RHb1	28	30	21	15
2-5	TgME49:TgMAF1RHb1	28	Ŧ	Ť	Ť
3-1	TgME49:EV	29	18	37	
3-2*	TgME49:EV	28	19	24	
3-3	TgME49:EV	24	27	35	
3-4	TgME49:EV	30.5	37	26	
3-5	TgME49:EV	22	27	19	
4-1	TgME49:EV	26	32	39	28
4-2	TgME49:EV	24	28	21	36
4-3	TgME49:EV	28	27.5	22	22
4-4	TgME49:EV	31	43	35	82.5
4-5	TgME49:EV	32	20	16	25

**Table 23. IL-15** 

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	16	17	16	
1-2	TgME49:TgMAF1RHb1	17	Ŧ	Ŧ	Ť
1-3	TgME49:TgMAF1RHb1	19	19.5	27	
1-4	TgME49:TgMAF1RHb1	20	3	29	
1-5	TgME49:TgMAF1RHb1	26.5	Ŧ	Ť	Ť
2-1	TgME49:TgMAF1RHb1	20	16	21	19
2-2*	TgME49:TgMAF1RHb1	16	10	8	14
2-3	TgME49:TgMAF1RHb1	27	8	26	93
2-4	TgME49:TgMAF1RHb1	15	26.5	23	9
2-5	TgME49:TgMAF1RHb1	25	Ť	Ť	Ť
3-1	TgME49:EV	17	16	16	
3-2*	TgME49:EV	20	14	17	
3-3	TgME49:EV	22	12	18	
3-4	TgME49:EV	19	15	21	
3-5	TgME49:EV	32	29	25	
4-1	TgME49:EV	21	5	5	8
4-2	TgME49:EV	27	13	14	9
4-3	TgME49:EV	27.5	14	19	25
4-4	TgME49:EV	24	23	31	47.5
4-5	TgME49:EV	22	10	9	6

**Table 24. IL-17** 

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	<b>57 DPI</b>
1-1	TgME49:TgMAF1RHb1	1	0	1	
1-2	TgME49:TgMAF1RHb1	3	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	2	4	4	
1-4	TgME49:TgMAF1RHb1	0	13	5.5	
1-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	1	3	3	12
2-2*	TgME49:TgMAF1RHb1	0	-3	-1	11
2-3	TgME49:TgMAF1RHb1	1	-1	-2	17
2-4	TgME49:TgMAF1RHb1	1	3	2	1
2-5	TgME49:TgMAF1RHb1	4	Ť	Ť	Ť
3-1	TgME49:EV	4	7	9	
3-2*	TgME49:EV	0	0	1	
3-3	TgME49:EV	1	4	2	
3-4	TgME49:EV	0	3	0	
3-5	TgME49:EV	-1	2	0.5	
4-1	TgME49:EV	1.5	1	-1	6
4-2	TgME49:EV	1	2	0	7
4-3	TgME49:EV	0	1.5	2	7
4-4	TgME49:EV	21	6	1	5
4-5	TgME49:EV	2	-2	-1	6

**Table 25. IP-10** 

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	309	4658	4662	
1-2	TgME49:TgMAF1RHb1	140	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	11	3440	4584.5	
1-4	TgME49:TgMAF1RHb1	59	4887	4195.5	
1-5	TgME49:TgMAF1RHb1	93	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	76	3460	2537	3543.5
2-2*	TgME49:TgMAF1RHb1	113	274	156	567.5
2-3	TgME49:TgMAF1RHb1	35.5	690.5	1610	5262.5
2-4	TgME49:TgMAF1RHb1	167	2925	2050	3399.5
2-5	TgME49:TgMAF1RHb1	184	Ť	Ť	Ť
3-1	TgME49:EV	231	2450	3018	
3-2*	TgME49:EV	119	289	1128.5	
3-3	TgME49:EV	79.5	3180	2904.5	
3-4	TgME49:EV	105	2107	2812	
3-5	TgME49:EV	54	2217	1987	
4-1	TgME49:EV	144	1135	1123	3441
4-2	TgME49:EV	144.5	2070.5	1181	3012.5
4-3	TgME49:EV	176	2480	1537	2782.5
4-4	TgME49:EV	161	2015	1799	4367
4-5	TgME49:EV	205.5	1073	1067.5	3018.5

Table 26. KC

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	<b>57 DPI</b>
1-1	TgME49:TgMAF1RHb1	25.5	140	469	
1-2	TgME49:TgMAF1RHb1	41	Ť	†	Ť
1-3	TgME49:TgMAF1RHb1	27.5	11.5	494.5	
1-4	TgME49:TgMAF1RHb1	7.5	130	72	
1-5	TgME49:TgMAF1RHb1	19.5	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	24.5	140.5	182.5	269.5
2-2*	TgME49:TgMAF1RHb1	23	103.5	47.5	240
2-3	TgME49:TgMAF1RHb1	8.5	85.5	113.5	245
2-4	TgME49:TgMAF1RHb1	59.5	192.5	352.5	56
2-5	TgME49:TgMAF1RHb1	44.5	Ť	Ť	Ť
3-1	TgME49:EV	58	257	1147.5	
3-2*	TgME49:EV	36.5	68	274	
3-3	TgME49:EV	25.5	144	197.5	
3-4	TgME49:EV	23.5	143.5	368	
3-5	TgME49:EV	12.5	78.5	228.5	
4-1	TgME49:EV	43.5	59.5	151.5	502
4-2	TgME49:EV	31.5	81.5	164.5	133
4-3	TgME49:EV	27	57.5	175.5	176
4-4	TgME49:EV	34	232.5	401	362
4-5	TgME49:EV	52.5	59.5	136.5	247

Table 27. LIF

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	-2	2	8	
1-2	TgME49:TgMAF1RHb1	-1	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	-2	-2	4	
1-4	TgME49:TgMAF1RHb1	-2	-2	2	
1-5	TgME49:TgMAF1RHb1	-3	Ť	Ť	ŧ
2-1	TgME49:TgMAF1RHb1	-3	-3	-2.5	22.5
2-2*	TgME49:TgMAF1RHb1	-3	-2	-3	11
2-3	TgME49:TgMAF1RHb1	-3	-1.5	10	170
2-4	TgME49:TgMAF1RHb1	-3	-2.5	-2	6
2-5	TgME49:TgMAF1RHb1	-2	Ť	Ť	Ť
3-1	TgME49:EV	-1	11	8	
3-2*	TgME49:EV	-3	-4	-4	
3-3	TgME49:EV	-3	0	0	
3-4	TgME49:EV	-4	-1	-2.5	
3-5	TgME49:EV	11	11	3	
4-1	TgME49:EV	-3	-1	-1	7
4-2	TgME49:EV	-1.5	-3	-2	11
4-3	TgME49:EV	-2	-2	-2	2
4-4	TgME49:EV	-1	-1	-2	26.5
4-5	TgME49:EV	1	-3	-1	4

Table 28. LIX

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	1940	7391.5	3694.5	
1-2	TgME49:TgMAF1RHb1	545.5	Ť	Ť	Ŧ
1-3	TgME49:TgMAF1RHb1	5048	130	3659	
1-4	TgME49:TgMAF1RHb1	1605.5	1799.5	4593.5	
1-5	TgME49:TgMAF1RHb1	227	Ť	Ť	ţ
2-1	TgME49:TgMAF1RHb1	278	1589	937.5	6425.5
2-2*	TgME49:TgMAF1RHb1	286	7078.5	967	20249.5
2-3	TgME49:TgMAF1RHb1	912.5	511	712	5696
2-4	TgME49:TgMAF1RHb1	3274.5	967	1784	6705.5
2-5	TgME49:TgMAF1RHb1	1480	Ť	Ť	ţ
3-1	TgME49:EV	1907.5	3858	889.5	
3-2*	TgME49:EV	2971	17010	2009	
3-3	TgME49:EV	1487.5	5128	3552.5	
3-4	TgME49:EV	109	2623	1301	
3-5	TgME49:EV	1529	3602	2802	
4-1	TgME49:EV	277	268	1458	9439.5
4-2	TgME49:EV	475	633	937.5	7199.5
4-3	TgME49:EV	1402	769	1606	6331.5
4-4	TgME49:EV	3939.5	525	896	6850.5
4-5	TgME49:EV	1810.5	622.5	8120.5	5871

Table 29. MCP-1

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	2	9	116	
1-2	TgME49:TgMAF1RHb1	3	Ŧ	Ť	Ť
1-3	TgME49:TgMAF1RHb1	7	16	166.5	
1-4	TgME49:TgMAF1RHb1	2	9	34	
1-5	TgME49:TgMAF1RHb1	2	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	1	7	6	18
2-2*	TgME49:TgMAF1RHb1	2	2	1	13
2-3	TgME49:TgMAF1RHb1	2	6	3	40
2-4	TgME49:TgMAF1RHb1	2	10	9	11.5
2-5	TgME49:TgMAF1RHb1	3	Ť	Ť	Ť
3-1	TgME49:EV	4	17	43	
3-2*	TgME49:EV	1	2	4	
3-3	TgME49:EV	1	19	19	
3-4	TgME49:EV	0.5	7	9	
3-5	TgME49:EV	1	5	4	
4-1	TgME49:EV	2	3	2	21
4-2	TgME49:EV	1	3	2	35
4-3	TgME49:EV	2	2	5	12
4-4	TgME49:EV	4.5	3	5	16
4-5	TgME49:EV	2	2	2	20

Table 30. M-CSF

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	2	2	6	
1-2	TgME49:TgMAF1RHb1	3	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	3	2	7	
1-4	TgME49:TgMAF1RHb1	2	1	9	
1-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	1	2	2	6
2-2*	TgME49:TgMAF1RHb1	1.5	2	0	7
2-3	TgME49:TgMAF1RHb1	2	3	1	9
2-4	TgME49:TgMAF1RHb1	2	3	2	2
2-5	TgME49:TgMAF1RHb1	2	Ť	Ť	Ŧ
3-1	TgME49:EV	2	2	1	
3-2*	TgME49:EV	0	2	2	
3-3	TgME49:EV	2	0	3	
3-4	TgME49:EV	1	2	3	
3-5	TgME49:EV	0	1	2.5	
4-1	TgME49:EV	2	2	2	4
4-2	TgME49:EV	3	1.5	0	5
4-3	TgME49:EV	1	5	7	12
4-4	TgME49:EV	1	8	10	23
4-5	TgME49:EV	2	0	1	3

Table 31. MIG

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	32	2940.5	7041	
1-2	TgME49:TgMAF1RHb1	21	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	10	1723	9075	
1-4	TgME49:TgMAF1RHb1	11	2404	8182	
1-5	TgME49:TgMAF1RHb1	9	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	9	2810	1138	5341
2-2*	TgME49:TgMAF1RHb1	13	62.5	44	416.5
2-3	TgME49:TgMAF1RHb1	0	459	741	11467.5
2-4	TgME49:TgMAF1RHb1	57	1399	1147.5	5756
2-5	TgME49:TgMAF1RHb1	20	Ť	Ť	Ť
3-1	TgME49:EV	53	2205.5	4308	
3-2*	TgME49:EV	39	75	64	
3-3	TgME49:EV	3	2822.5	5082.5	
3-4	TgME49:EV	12	1033	4673	
3-5	TgME49:EV	11	1362	3147	
4-1	TgME49:EV	14	533	917	3621
4-2	TgME49:EV	36	1241	664	4450.5
4-3	TgME49:EV	17.5	635	1960	993.5
4-4	TgME49:EV	58	698	964	8207
4-5	TgME49:EV	22	506.5	715	2531

Table 32. MIP-1a

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	<b>57 DPI</b>
1-1	TgME49:TgMAF1RHb1	6	11	9	
1-2	TgME49:TgMAF1RHb1	3	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	3	7	11	
1-4	TgME49:TgMAF1RHb1	-1	12	17	
1-5	TgME49:TgMAF1RHb1	0	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	-1	1.5	2	31.5
2-2*	TgME49:TgMAF1RHb1	1	-2	-1	10
2-3	TgME49:TgMAF1RHb1	3	-3	-2	11
2-4	TgME49:TgMAF1RHb1	4	8	4	4
2-5	TgME49:TgMAF1RHb1	6	Ť	Ť	Ť
3-1	TgME49:EV	4	13	8	
3-2*	TgME49:EV	1	8	5	
3-3	TgME49:EV	3	13	5.5	
3-4	TgME49:EV	-1.5	0	0	
3-5	TgME49:EV	1	-3	-5	
4-1	TgME49:EV	1	0	-1	9
4-2	TgME49:EV	0	0	0	12
4-3	TgME49:EV	0	0	2	5
4-4	TgME49:EV	5	28	7	13
4-5	TgME49:EV	4	0	-1	4

Table 33. MIP-1b

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	16	30	4.5	
1-2	TgME49:TgMAF1RHb1	13.5	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	5.5	7.5	29	
1-4	TgME49:TgMAF1RHb1	8	27	35	
1-5	TgME49:TgMAF1RHb1	9.5	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	6.5	-4.5	-1.5	42.5
2-2*	TgME49:TgMAF1RHb1	5.5	-1.5	-0.5	19.5
2-3	TgME49:TgMAF1RHb1	9.5	3.5	-4.5	56.5
2-4	TgME49:TgMAF1RHb1	7.5	13.5	4.5	42.5
2-5	TgME49:TgMAF1RHb1	10.5	Ŧ	Ť	Ť
3-1	TgME49:EV	-7	22	19	
3-2*	TgME49:EV	8.5	3	3	
3-3	TgME49:EV	5.5	31	5	
3-4	TgME49:EV	7.5	0.5	-8.5	
3-5	TgME49:EV	1.5	4.5	0.5	
4-1	TgME49:EV	3.5	15.5	4.5	15.5
4-2	TgME49:EV	-3.5	-5.5	-9.5	11.5
4-3	TgME49:EV	10.5	4.5	0.5	16.5
4-4	TgME49:EV	8.5	19.5	-4.5	49.5
4-5	TgME49:EV	12.5	-13.5	-12	24.5

Table 34. MIP-2

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	4	3	2	
1-2	TgME49:TgMAF1RHb1	3	Ť	Ŧ	Ť
1-3	TgME49:TgMAF1RHb1	5	2	4	
1-4	TgME49:TgMAF1RHb1	2	1	6	
1-5	TgME49:TgMAF1RHb1	2	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	1	1	0	19
2-2*	TgME49:TgMAF1RHb1	1	2	0	23
2-3	TgME49:TgMAF1RHb1	2	3	0	23
2-4	TgME49:TgMAF1RHb1	2	3	2	7
2-5	TgME49:TgMAF1RHb1	4	Ť	Ŧ	Ť
3-1	TgME49:EV	3	2	3	
3-2*	TgME49:EV	7	3	3	
3-3	TgME49:EV	3	3	1	
3-4	TgME49:EV	-1	2	0	
3-5	TgME49:EV	-1	2	1	
4-1	TgME49:EV	1	0	0	11
4-2	TgME49:EV	2	1	0	11
4-3	TgME49:EV	2	0	-1	11
4-4	TgME49:EV	4	9	5	26
4-5	TgME49:EV	3	-2	1	7

Table 35. RANTES

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	13	99	1576	
1-2	TgME49:TgMAF1RHb1	11	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	9	58.5	1069.5	
1-4	TgME49:TgMAF1RHb1	9	96	1540.5	
1-5	TgME49:TgMAF1RHb1	6	Ť	Ť	ŧ
2-1	TgME49:TgMAF1RHb1	9	15	38.5	1452
2-2*	TgME49:TgMAF1RHb1	7	14	7	154
2-3	TgME49:TgMAF1RHb1	6	10	-3	614
2-4	TgME49:TgMAF1RHb1	10	52	23.5	1262
2-5	TgME49:TgMAF1RHb1	12	Ť	Ť	Ť
3-1	TgME49:EV	14	137	143	
3-2*	TgME49:EV	7	13	20	
3-3	TgME49:EV	10	31	696	
3-4	TgME49:EV	6	24	116	
3-5	TgME49:EV	4	24	417	
4-1	TgME49:EV	7	15	22.5	297
4-2	TgME49:EV	7	16	16	236
4-3	TgME49:EV	12	6	22	129
4-4	TgME49:EV	33.5	7	9	1287
4-5	TgME49:EV	11	16	17	141

Table 36. TNFa

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	-1	18	14	
1-2	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	6	15	24	
1-4	TgME49:TgMAF1RHb1	-1	16	12	
1-5	TgME49:TgMAF1RHb1	1	Ť	Ť	ŧ
2-1	TgME49:TgMAF1RHb1	1	11	5	18
2-2*	TgME49:TgMAF1RHb1	1	0	0	6
2-3	TgME49:TgMAF1RHb1	0	9	2	23
2-4	TgME49:TgMAF1RHb1	0	12	5	16
2-5	TgME49:TgMAF1RHb1	1	Ť	Ť	Ť
3-1	TgME49:EV	1	7	9	
3-2*	TgME49:EV	1	5	1	
3-3	TgME49:EV	1	9	10.5	
3-4	TgME49:EV	-0.5	2	4	
3-5	TgME49:EV	0	3	6	
4-1	TgME49:EV	0	3	2	7
4-2	TgME49:EV	0	10	1	10
4-3	TgME49:EV	0	3	4	11
4-4	TgME49:EV	5	47	5	7
4-5	TgME49:EV	2	6	1	7

Table 37. VEGF

Mouse	Parasite Strain	Pre-bleed	21 DPI	28 DPI	57 DPI
1-1	TgME49:TgMAF1RHb1	7	8	22	
1-2	TgME49:TgMAF1RHb1	6	Ť	Ť	Ť
1-3	TgME49:TgMAF1RHb1	10	8	66	
1-4	TgME49:TgMAF1RHb1	6	7	139	
1-5	TgME49:TgMAF1RHb1	4	Ť	Ť	Ť
2-1	TgME49:TgMAF1RHb1	6	6	5	146
2-2*	TgME49:TgMAF1RHb1	5	4	3	109
2-3	TgME49:TgMAF1RHb1	8	0	0	335
2-4	TgME49:TgMAF1RHb1	11	12	4	290
2-5	TgME49:TgMAF1RHb1	24.5	Ť	Ť	Ť
3-1	TgME49:EV	12.5	7	3	
3-2*	TgME49:EV	8	5	69	
3-3	TgME49:EV	4	9	5	
3-4	TgME49:EV	7	2	19	
3-5	TgME49:EV	1	3	8	
4-1	TgME49:EV	4	3	2	64
4-2	TgME49:EV	9	10	3.5	32
4-3	TgME49:EV	1.5	3	1	3
4-4	TgME49:EV	17	47	15	248.5
4-5	TgME49:EV	20	6	2	2

## BIBLIOGRAPHY

1. Frenkel JK, Dubey JP, Miller NL. Toxoplasma gondii in cats: fecal stages identified as coccidian oocysts. Science. 1970;167(3919):893-6. PubMed PMID: 4903651.

2. Dubey JP, Frenkel JK. Experimental toxoplasma infection in mice with strains producing oocysts. J Parasitol. 1973;59(3):505-12. PubMed PMID: 4576142.

3. Dubey JP. Bradyzoite-induced murine toxoplasmosis: stage conversion, pathogenesis, and tissue cyst formation in mice fed bradyzoites of different strains of Toxoplasma gondii. J Eukaryot Microbiol. 1997;44(6):592-602. PubMed PMID: 9435131.

4. Tenter AM, Heckeroth AR, Weiss LM. Toxoplasma gondii: from animals to humans. Int J Parasitol. 2000;30(12-13):1217-58. PubMed PMID: 11113252; PubMed Central PMCID: PMCPMC3109627.

5. Di Carlo P, Romano A, Schimmenti MG, Mazzola A, Titone L. Materno-fetal Toxoplasma gondii infection: critical review of available diagnostic methods. Infez Med. 2008;16(1):28-32. PubMed PMID: 18367880.

6. Benavides J, Maley S, Pang Y, Palarea J, Eaton S, Katzer F, et al. Development of lesions and tissue distribution of parasite in lambs orally infected with sporulated oocysts of Toxoplasma gondii. Vet Parasitol. 2011;179(1-3):209-15. doi: 10.1016/j.vetpar.2011.03.001. PubMed PMID: 21440372.

7. McLeod R, Estes RG, Mack DG, Cohen H. Immune response of mice to ingested Toxoplasma gondii: a model of toxoplasma infection acquired by ingestion. J Infect Dis. 1984;149(2):234-44. PubMed PMID: 6699433.

8. Dubey JP, Frenkel JK. Feline toxoplasmosis from acutely infected mice and the development of Toxoplasma cysts. J Protozool. 1976;23(4):537-46. PubMed PMID: 1003342.

9. Porter SB, Sande MA. Toxoplasmosis of the central nervous system in the acquired immunodeficiency syndrome. N Engl J Med. 1992;327(23):1643-8. doi: 10.1056/NEJM199212033272306. PubMed PMID: 1359410.

10. Dubey JP, Sreekumar C. Redescription of Hammondia hammondi and its differentiation from Toxoplasma gondii. Int J Parasitol. 2003;33(13):1437-53. PubMed PMID: 14572507.

11. Dubey JP, Miller NL, Frenkel JK. The Toxoplasma gondii oocyst from cat feces. J Exp Med. 1970;132(4):636-62. PubMed PMID: 4927658; PubMed Central PMCID: PMCPMC2138867.

12. Frenkel JK, Ruiz A, Chinchilla M. Soil survival of toxoplasma oocysts in Kansas and Costa Rica. Am J Trop Med Hyg. 1975;24(3):439-43. PubMed PMID: 1098494.

13. Black MW, Boothroyd JC. Lytic cycle of Toxoplasma gondii. Microbiol Mol Biol Rev. 2000;64(3):607-23. PubMed PMID: 10974128; PubMed Central PMCID: PMCPMC99006.

14. Luft BJ, Conley F, Remington JS, Laverdiere M, Wagner KF, Levine JF, et al. Outbreak of central-nervous-system toxoplasmosis in western Europe and North America. Lancet. 1983;1(8328):781-4. PubMed PMID: 6132129.

15. Luft BJ, Naot Y, Araujo FG, Stinson EB, Remington JS. Primary and reactivated toxoplasma infection in patients with cardiac transplants. Clinical spectrum and problems in diagnosis in a defined population. Ann Intern Med. 1983;99(1):27-31. PubMed PMID: 6344718.

16. Mastrobuoni S, Dell'aquila AM, Herreros J. Fatal Toxoplasma gondii Dissemination in a Heart Transplant Recipient: Description of a Case. Case Rep Transplant. 2012;2012:524279. doi: 10.1155/2012/524279. PubMed PMID: 23259134; PubMed Central PMCID: PMCPMC3504277.

17. Carme B, Demar M, Ajzenberg D, Darde ML. Severe acquired toxoplasmosis caused by wild cycle of Toxoplasma gondii, French Guiana. Emerg Infect Dis. 2009;15(4):656-8. doi: 10.3201/eid1504.081306. PubMed PMID: 19331765; PubMed Central PMCID: PMCPMC2671434.

18. Demar M, Ajzenberg D, Maubon D, Djossou F, Panchoe D, Punwasi W, et al. Fatal outbreak of human toxoplasmosis along the Maroni River: epidemiological, clinical, and parasitological aspects. Clin Infect Dis. 2007;45(7):e88-95. doi: 10.1086/521246. PubMed PMID: 17806043.

19. Carme B, Bissuel F, Ajzenberg D, Bouyne R, Aznar C, Demar M, et al. Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. J Clin Microbiol. 2002;40(11):4037-44. PubMed PMID: 12409371; PubMed Central PMCID: PMCPMC139686.

20. Austeng ME, Eskild A, Jacobsen M, Jenum PA, Whitelaw A, Engdahl B. Maternal infection with toxoplasma gondii in pregnancy and the risk of hearing loss in the offspring. Int J Audiol. 2010;49(1):65-8. doi: 10.3109/14992020903214053. PubMed PMID: 20053157.

21. Wolf A, Cowen D, Paige B. Human Toxoplasmosis: Occurrence in Infants as an Encephalomyelitis Verification by Transmission to Animals. Science. 1939;89(2306):226-7. doi: 10.1126/science.89.2306.226. PubMed PMID: 17737029.

22. Dubey JP, Schmitz JA. Abortion associated with toxoplasmosis in sheep in Oregon. J Am Vet Med Assoc. 1981;178(7):675-8. PubMed PMID: 7204245.

23. Dubey JP. Toxoplasma-induced abortion in dairy goats. J Am Vet Med Assoc. 1981;178(7):671-4. PubMed PMID: 7204244.

24. Hunter B. Isolated, spontaneous Toxoplasma abortion in a young sow. Can Vet J. 1979;20(4):116. PubMed PMID: 427708; PubMed Central PMCID: PMCPMC1789533.

25. Hill DE, Dubey JP. Toxoplasma gondii prevalence in farm animals in the United States. Int J Parasitol. 2013;43(2):107-13. doi: 10.1016/j.ijpara.2012.09.012. PubMed PMID: 23201235.

26. Frenkel JK, Dubey JP. Hammondia hammondi: A new coccidium of cats producing cysts in muscle of other mammals. Science. 1975;189(4198):222-4. PubMed PMID: 806116.

27. Riahi H, Bouteille B, Darde ML. Antigenic similarity between Hammondia hammondi and Toxoplasma gondii tachyzoites. J Parasitol. 1998;84(3):651-3. PubMed PMID: 9645881.

28. Christie E, Dubey JP. Cross-immunity between Hammondia and Toxoplasma infections in mice and hamsters. Infect Immun. 1977;18(2):412-5. PubMed PMID: 411757; PubMed Central PMCID: PMCPMC421248.

29. Dubey JP, Wong M. Experimental Hammondia hammondi infection in monkeys. J Parasitol. 1978;64(3):551-2. PubMed PMID: 96246.

30. Frenkel JK, Dubey JP. Hammondia hammondi gen. nov., sp.nov., from domestic cats, a new coccidian related to Toxoplasma and Sarcocystis. Z Parasitenkd. 1975;46(1):3-12. PubMed PMID: 807048.

31. Dubey JP, Tilahun G, Boyle JP, Schares G, Verma SK, Ferreira LR, et al. Molecular and biological characterization of first isolates of Hammondia hammondi from cats from Ethiopia. J Parasitol. 2013;99(4):614-8. doi: 10.1645/12-51.1. PubMed PMID: 23517380.

32. McAllister MM, Dubey JP, Lindsay DS, Jolley WR, Wills RA, McGuire AM. Dogs are definitive hosts of Neospora caninum. Int J Parasitol. 1998;28(9):1473-8. PubMed PMID: 9770635.

33. Dubey JP, Carpenter JL, Speer CA, Topper MJ, Uggla A. Newly recognized fatal protozoan disease of dogs. J Am Vet Med Assoc. 1988;192(9):1269-85. PubMed PMID: 3391851.

34. Dubey JP, Hattel AL, Lindsay DS, Topper MJ. Neonatal Neospora caninum infection in dogs: isolation of the causative agent and experimental transmission. J Am Vet Med Assoc. 1988;193(10):1259-63. PubMed PMID: 3144521.

35. Dubey JP, Leathers CW, Lindsay DS. Neospora caninum-like protozoon associated with fatal myelitis in newborn calves. J Parasitol. 1989;75(1):146-8. PubMed PMID: 2493086.

36. Dubey JP, Romand S, Hilali M, Kwok OC, Thulliez P. Seroprevalence of antibodies to Neospora caninum and Toxoplasma gondii in water buffaloes (Bubalus bubalis) from Egypt. Int J Parasitol. 1998;28(3):527-9. PubMed PMID: 9559371.

37. Dubey JP, Hartley WJ, Lindsay DS, Topper MJ. Fatal congenital Neospora caninum infection in a lamb. J Parasitol. 1990;76(1):127-30. PubMed PMID: 2299518.

38. Dubey JP, Acland HM, Hamir AN. Neospora caninum (Apicomplexa) in a stillborn goat. J Parasitol. 1992;78(3):532-4. PubMed PMID: 1597802.

39. Dubey JP, Porterfield ML. Neospora caninum (Apicomplexa) in an aborted equine fetus. J Parasitol. 1990;76(5):732-4. PubMed PMID: 2213418.

40. Thilsted JP, Dubey JP. Neosporosis-like abortions in a herd of dairy cattle. J Vet Diagn Invest. 1989;1(3):205-9. PubMed PMID: 2488345.

41. Shivaprasad HL, Ely R, Dubey JP. A Neospora-like protozoon found in an aborted bovine placenta. Vet Parasitol. 1989;34(1-2):145-8. PubMed PMID: 2588466.

42. Barratt JL, Harkness J, Marriott D, Ellis JT, Stark D. Importance of nonenteric protozoan infections in immunocompromised people. Clin Microbiol Rev. 2010;23(4):795-836. doi: 10.1128/CMR.00001-10. PubMed PMID: 20930074; PubMed Central PMCID: PMCPMC2952979.

43. Ho MS, Barr BC, Rowe JD, Anderson ML, Sverlow KW, Packham A, et al. Detection of Neospora sp. from infected bovine tissues by PCR and probe hybridization. J Parasitol. 1997;83(3):508-14. PubMed PMID: 9194835.

44. Huong LT, Ljungstrom BL, Uggla A, Bjorkman C. Prevalence of antibodies to Neospora caninum and Toxoplasma gondii in cattle and water buffaloes in southern Vietnam. Vet Parasitol. 1998;75(1):53-7. PubMed PMID: 9566094.

45. Wiengcharoen J, Nakthong C, Mitchaothai J, Udonsom R, Sukthana Y. Toxoplasmosis and neosporosis among beef cattle slaughtered for food in Western Thailand. Southeast Asian J Trop Med Public Health. 2012;43(5):1087-93. PubMed PMID: 23431814.

46. Xu MJ, Liu QY, Fu JH, Nisbet AJ, Shi DS, He XH, et al. Seroprevalence of Toxoplasma gondii and Neospora caninum infection in dairy cows in subtropical southern China.

Parasitology. 2012;139(11):1425-8. doi: 10.1017/S0031182012000728. PubMed PMID: 22717118.

47. Ferguson HW, Ellis WA. Toxoplasmosis in a calf. Vet Rec. 1979;104(17):392-3. PubMed PMID: 473542.

48. Dubey JP. Isolation of Toxoplasma gondii from a naturally infected beef cow. J Parasitol. 1992;78(1):151-3. PubMed PMID: 1738059.

49. Costa GH, da Costa AJ, Lopes WD, Bresciani KD, dos Santos TR, Esper CR, et al. Toxoplasma gondii: infection natural congenital in cattle and an experimental inoculation of gestating cows with oocysts. Exp Parasitol. 2011;127(1):277-81. doi: 10.1016/j.exppara.2010.08.005. PubMed PMID: 20736009.

50. Stalheim OH, Hubbert WT, Boothe AD, Zimmermann WJ, Hughes DE, Barnett D, et al. Experimental toxoplasmosis in calves and pregnant cows. Am J Vet Res. 1980;41(1):10-3. PubMed PMID: 7362114.

51. Wiengcharoen J, Thompson RC, Nakthong C, Rattanakorn P, Sukthana Y. Transplacental transmission in cattle: is Toxoplasma gondii less potent than Neospora caninum? Parasitol Res. 2011;108(5):1235-41. doi: 10.1007/s00436-010-2172-8. PubMed PMID: 21203773.

52. Otter A, Jeffrey M, Griffiths IB, Dubey JP. A survey of the incidence of Neospora caninum infection in aborted and stillborn bovine fetuses in England and Wales. Vet Rec. 1995;136(24):602-6. PubMed PMID: 7571263.

53. Ghalmi F, China B, Kaidi R, Losson B. Neospora caninum is associated with abortion in Algerian cattle. J Parasitol. 2011;97(6):1121-4. doi: 10.1645/GE-2861.1. PubMed PMID: 21728878.

54. Gibney EH, Kipar A, Rosbottom A, Guy CS, Smith RF, Hetzel U, et al. The extent of parasite-associated necrosis in the placenta and foetal tissues of cattle following Neospora caninum infection in early and late gestation correlates with foetal death. Int J Parasitol. 2008;38(5):579-88. doi: 10.1016/j.ijpara.2007.09.015. PubMed PMID: 18021783.

55. Rojo-Montejo S, Collantes-Fernandez E, Blanco-Murcia J, Rodriguez-Bertos A, Risco-Castillo V, Ortega-Mora LM. Experimental infection with a low virulence isolate of Neospora caninum at 70 days gestation in cattle did not result in foetopathy. Vet Res. 2009;40(5):49. doi: 10.1051/vetres/2009032. PubMed PMID: 19497257; PubMed Central PMCID: PMCPMC2704335.

56. Benavides J, Katzer F, Maley SW, Bartley PM, Canton G, Palarea-Albaladejo J, et al. High rate of transplacental infection and transmission of Neospora caninum following experimental challenge of cattle at day 210 of gestation. Vet Res. 2012;43:83. doi: 10.1186/1297-9716-43-83. PubMed PMID: 23228067; PubMed Central PMCID: PMCPMC3567967.

57. Almeria S, Araujo R, Tuo W, Lopez-Gatius F, Dubey JP, Gasbarre LC. Fetal death in cows experimentally infected with Neospora caninum at 110 days of gestation. Vet Parasitol. 2010;169(3-4):304-11. doi: 10.1016/j.vetpar.2009.12.044. PubMed PMID: 20089361.

58. McCann CM, McAllister MM, Gondim LF, Smith RF, Cripps PJ, Kipar A, et al. Neospora caninum in cattle: experimental infection with oocysts can result in exogenous transplacental infection, but not endogenous transplacental infection in the subsequent pregnancy. Int J Parasitol. 2007;37(14):1631-9. doi: 10.1016/j.ijpara.2007.05.012. PubMed PMID: 17624353.

59. Pabon M, Lopez-Gatius F, Garcia-Ispierto I, Bech-Sabat G, Nogareda C, Almeria S. Chronic Neospora caninum infection and repeat abortion in dairy cows: a 3-year study. Vet Parasitol. 2007;147(1-2):40-6. doi: 10.1016/j.vetpar.2007.03.017. PubMed PMID: 17467905.

60. Nishimura M, Kohara J, Hiasa J, Muroi Y, Yokoyama N, Kida K, et al. Tissue distribution of Neospora caninum in experimentally infected cattle. Clin Vaccine Immunol. 2013;20(2):309-12. doi: 10.1128/CVI.00556-12. PubMed PMID: 23239805; PubMed Central PMCID: PMCPMC3571275.

61. Ueno TE, Goncalves VS, Heinemann MB, Dilli TL, Akimoto BM, de Souza SL, et al. Prevalence of Toxoplasma gondii and Neospora caninum infections in sheep from Federal District, central region of Brazil. Trop Anim Health Prod. 2009;41(4):547-52. doi: 10.1007/s11250-008-9220-8. PubMed PMID: 18726165.

62. Romanelli PR, Freire RL, Vidotto O, Marana ER, Ogawa L, De Paula VS, et al. Prevalence of Neospora caninum and Toxoplasma gondii in sheep and dogs from Guarapuava farms, Parana State, Brazil. Res Vet Sci. 2007;82(2):202-7. doi: 10.1016/j.rvsc.2006.04.001. PubMed PMID: 17266999.

63. Figliuolo LP, Kasai N, Ragozo AM, de Paula VS, Dias RA, Souza SL, et al. Prevalence of anti-Toxoplasma gondii and anti-Neospora caninum antibodies in ovine from Sao Paulo State, Brazil. Vet Parasitol. 2004;123(3-4):161-6. doi: 10.1016/j.vetpar.2004.06.006. PubMed PMID: 15325042.

64. Bartova E, Sedlak K, Literak I. Toxoplasma gondii and Neospora caninum antibodies in sheep in the Czech Republic. Vet Parasitol. 2009;161(1-2):131-2. doi: 10.1016/j.vetpar.2008.12.022. PubMed PMID: 19181450.

65. McAllister MM, McGuire AM, Jolley WR, Lindsay DS, Trees AJ, Stobart RH. Experimental neosporosis in pregnant ewes and their offspring. Vet Pathol. 1996;33(6):647-55. PubMed PMID: 8952023.

66. Buxton D, Maley SW, Thomson KM, Trees AJ, Innes EA. Experimental infection of non-pregnant and pregnant sheep with Neospora caninum. J Comp Pathol. 1997;117(1):1-16. PubMed PMID: 9263840.

67. Buxton D, Maley SW, Wright S, Thomson KM, Rae AG, Innes EA. The pathogenesis of experimental neosporosis in pregnant sheep. J Comp Pathol. 1998;118(4):267-79. PubMed PMID: 9651804.

68. Owen MR, Clarkson MJ, Trees AJ. Acute phase toxoplasma abortions in sheep. Vet Rec. 1998;142(18):480-2. PubMed PMID: 9612913.

69. Dubey JP, Welcome FL. Toxoplasma gondii-induced abortion in sheep. J Am Vet Med Assoc. 1988;193(6):697-700. PubMed PMID: 3192448.

70. Morley EK, Williams RH, Hughes JM, Thomasson D, Terry RS, Duncanson P, et al. Evidence that primary infection of Charollais sheep with Toxoplasma gondii may not prevent foetal infection and abortion in subsequent lambings. Parasitology. 2008;135(2):169-73. doi: 10.1017/S0031182007003721. PubMed PMID: 17922930.

71. Jolley WR, McAllister MM, McGuire AM, Wills RA. Repetitive abortion in Neosporainfected ewes. Vet Parasitol. 1999;82(3):251-7. PubMed PMID: 10348105.

72. O'Donovan J, Proctor A, Gutierrez J, Worrell S, Nally J, Marques P, et al. Distribution of lesions in fetal brains following experimental infection of pregnant sheep with Toxoplasma gondii. Vet Pathol. 2012;49(3):462-9. doi: 10.1177/0300985811424732. PubMed PMID: 22002976.

73. Lopes WD, Santos TR, Luvizotto MC, Sakamoto CA, Oliveira GP, Costa AJ. Histopathology of the reproductive system of male sheep experimentally infected with Toxoplasma gondii. Parasitol Res. 2011;109(2):405-9. doi: 10.1007/s00436-011-2268-9. PubMed PMID: 21286752.

74. Lopes WD, Rodriguez JD, Souza FA, dos Santos TR, dos Santos RS, Rosanese WM, et al. Sexual transmission of Toxoplasma gondii in sheep. Vet Parasitol. 2013;195(1-2):47-56. doi: 10.1016/j.vetpar.2012.12.056. PubMed PMID: 23384578.

75. de Moraes EP, Batista AM, Faria EB, Freire RL, Freitas AC, Silva MA, et al. Experimental infection by Toxoplasma gondii using contaminated semen containing different doses of tachyzoites in sheep. Vet Parasitol. 2010;170(3-4):318-22. doi: 10.1016/j.vetpar.2010.02.017. PubMed PMID: 20227187.

76. Czopowicz M, Kaba J, Szalus-Jordanow O, Nowicki M, Witkowski L, Frymus T. Seroprevalence of Toxoplasma gondii and Neospora caninum infections in goats in Poland. Vet Parasitol. 2011;178(3-4):339-41. doi: 10.1016/j.vetpar.2011.01.039. PubMed PMID: 21324599.

77. Faria EB, Gennari SM, Pena HF, Athayde AC, Silva ML, Azevedo SS. Prevalence of anti-Toxoplasma gondii and anti-Neospora caninum antibodies in goats slaughtered in the public slaughterhouse of Patos city, Paraiba State, Northeast region of Brazil. Vet Parasitol. 2007;149(1-2):126-9. doi: 10.1016/j.vetpar.2007.07.009. PubMed PMID: 17706359.

78. Iovu A, Gyorke A, Mircean V, Gavrea R, Cozma V. Seroprevalence of Toxoplasma gondii and Neospora caninum in dairy goats from Romania. Vet Parasitol. 2012;186(3-4):470-4. doi: 10.1016/j.vetpar.2011.11.062. PubMed PMID: 22177331.

79. Dubey JP, Miller S, Desmonts G, Thulliez P, Anderson WR. Toxoplasma gondii-induced abortion in dairy goats. J Am Vet Med Assoc. 1986;188(2):159-62. PubMed PMID: 3700211.

80. Moreno B, Collantes-Fernandez E, Villa A, Navarro A, Regidor-Cerrillo J, Ortega-Mora LM. Occurrence of Neospora caninum and Toxoplasma gondii infections in ovine and caprine abortions. Vet Parasitol. 2012;187(1-2):312-8. doi: 10.1016/j.vetpar.2011.12.034. PubMed PMID: 22260901.

81. Dubey JP, Desmonts G, Antunes F, McDonald C. Serologic diagnosis of toxoplasmosis in experimentally infected pregnant goats and transplacentally infected kids. Am J Vet Res. 1985;46(5):1137-40. PubMed PMID: 4003888.

82. Dubey JP. Repeat transplacental transfer of Toxoplasma gondii in dairy goats. J Am Vet Med Assoc. 1982;180(10):1220-1. PubMed PMID: 7085441.

83. Lindsay DS, Rippey NS, Powe TA, Sartin EA, Dubey JP, Blagburn BL. Abortions, fetal death, and stillbirths in pregnant pygmy goats inoculated with tachyzoites of Neospora caninum. Am J Vet Res. 1995;56(9):1176-80. PubMed PMID: 7486395.

84. Santana LF, da Costa AJ, Pieroni J, Lopes WD, Santos RS, de Oliveira GP, et al. Detection of Toxoplasma gondii in the reproductive system of male goats. Rev Bras Parasitol Vet. 2010;19(3):179-82. PubMed PMID: 20943023.

85. Wanderley FS, Porto WJ, Camara DR, da Cruz NL, Feitosa BC, Freire RL, et al. Experimental vaginal infection of goats with semen contaminated with the "CPG" strain of Toxoplasma gondii. J Parasitol. 2013;99(4):610-3. doi: 10.1645/12-126.1. PubMed PMID: 23391103.

86. Rachinel N, Buzoni-Gatel D, Dutta C, Mennechet FJ, Luangsay S, Minns LA, et al. The induction of acute ileitis by a single microbial antigen of Toxoplasma gondii. J Immunol. 2004;173(4):2725-35. PubMed PMID: 15294991.

87. Buzoni-Gatel D, Debbabi H, Mennechet FJ, Martin V, Lepage AC, Schwartzman JD, et al. Murine ileitis after intracellular parasite infection is controlled by TGF-beta-producing intraepithelial lymphocytes. Gastroenterology. 2001;120(4):914-24. PubMed PMID: 11231945.

88. Lilue J, Muller UB, Steinfeldt T, Howard JC. Reciprocal virulence and resistance polymorphism in the relationship between Toxoplasma gondii and the house mouse. Elife. 2013;2:e01298. doi: 10.7554/eLife.01298. PubMed PMID: 24175088; PubMed Central PMCID: PMCPMC3810784.

89. Abou-Bacar A, Pfaff AW, Georges S, Letscher-Bru V, Filisetti D, Villard O, et al. Role of NK cells and gamma interferon in transplacental passage of Toxoplasma gondii in a mouse model of primary infection. Infect Immun. 2004;72(3):1397-401. PubMed PMID: 14977944; PubMed Central PMCID: PMCPMC356035.

90. Dubey JP, Ferreira LR, Martins J, McLeod R. Oral oocyst-induced mouse model of toxoplasmosis: effect of infection with Toxoplasma gondii strains of different genotypes, dose, and mouse strains (transgenic, out-bred, in-bred) on pathogenesis and mortality. Parasitology. 2012;139(1):1-13. doi: 10.1017/S0031182011001673. PubMed PMID: 22078010; PubMed Central PMCID: PMCPMC3683600.

91. Hunter CA, Candolfi E, Subauste C, Van Cleave V, Remington JS. Studies on the role of interleukin-12 in acute murine toxoplasmosis. Immunology. 1995;84(1):16-20. PubMed PMID: 7890300; PubMed Central PMCID: PMCPMC1415199.

92. Jones TC, Alkan S, Erb P. Spleen and lymph node cell populations, in vitro cell proliferation and interferon-gamma production during the primary immune response to Toxoplasma gondii. Parasite Immunol. 1986;8(6):619-29. PubMed PMID: 3101032.

93. Suzuki Y, Orellana MA, Schreiber RD, Remington JS. Interferon-gamma: the major mediator of resistance against Toxoplasma gondii. Science. 1988;240(4851):516-8. PubMed PMID: 3128869.

94. Gazzinelli R, Xu Y, Hieny S, Cheever A, Sher A. Simultaneous depletion of CD4+ and CD8+ T lymphocytes is required to reactivate chronic infection with Toxoplasma gondii. J Immunol. 1992;149(1):175-80. PubMed PMID: 1351500.

95. Howe DK, Sibley LD. Toxoplasma gondii comprises three clonal lineages: correlation of parasite genotype with human disease. J Infect Dis. 1995;172(6):1561-6. PubMed PMID: 7594717.

96. Saeij JP, Boyle JP, Coller S, Taylor S, Sibley LD, Brooke-Powell ET, et al. Polymorphic secreted kinases are key virulence factors in toxoplasmosis. Science. 2006;314(5806):1780-3. doi: 10.1126/science.1133690. PubMed PMID: 17170306; PubMed Central PMCID: PMCPMC2646183.

97. Sibley LD, Boothroyd JC. Virulent strains of Toxoplasma gondii comprise a single clonal lineage. Nature. 1992;359(6390):82-5. doi: 10.1038/359082a0. PubMed PMID: 1355855.

98. Darde ML. Toxoplasma gondii, "new" genotypes and virulence. Parasite. 2008;15(3):366-71. PubMed PMID: 18814708.

99. Behnke MS, Khan A, Wootton JC, Dubey JP, Tang K, Sibley LD. Virulence differences in Toxoplasma mediated by amplification of a family of polymorphic pseudokinases. Proc Natl Acad Sci U S A. 2011;108(23):9631-6. doi: 10.1073/pnas.1015338108. PubMed PMID: 21586633; PubMed Central PMCID: PMCPMC3111276.

100. Fleckenstein MC, Reese ML, Konen-Waisman S, Boothroyd JC, Howard JC, Steinfeldt T. A Toxoplasma gondii pseudokinase inhibits host IRG resistance proteins. PLoS Biol.

2012;10(7):e1001358. doi: 10.1371/journal.pbio.1001358. PubMed PMID: 22802726; PubMed Central PMCID: PMCPMC3393671.

101. Reese ML, Boothroyd JC. A conserved non-canonical motif in the pseudoactive site of the ROP5 pseudokinase domain mediates its effect on Toxoplasma virulence. J Biol Chem. 2011;286(33):29366-75. doi: 10.1074/jbc.M111.253435. PubMed PMID: 21708941; PubMed Central PMCID: PMCPMC3190742.

102. Reese ML, Zeiner GM, Saeij JP, Boothroyd JC, Boyle JP. Polymorphic family of injected pseudokinases is paramount in Toxoplasma virulence. Proc Natl Acad Sci U S A. 2011;108(23):9625-30. doi: 10.1073/pnas.1015980108. PubMed PMID: 21436047; PubMed Central PMCID: PMCPMC3111280.

103. Taylor S, Barragan A, Su C, Fux B, Fentress SJ, Tang K, et al. A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen Toxoplasma gondii. Science. 2006;314(5806):1776-80. doi: 10.1126/science.1133643. PubMed PMID: 17170305.

104. El Hajj H, Lebrun M, Arold ST, Vial H, Labesse G, Dubremetz JF. ROP18 is a rhoptry kinase controlling the intracellular proliferation of Toxoplasma gondii. PLoS Pathog. 2007;3(2):e14. doi: 10.1371/journal.ppat.0030014. PubMed PMID: 17305424; PubMed Central PMCID: PMCPMC1797617.

105. Saeij JP, Coller S, Boyle JP, Jerome ME, White MW, Boothroyd JC. Toxoplasma co-opts host gene expression by injection of a polymorphic kinase homologue. Nature. 2007;445(7125):324-7. doi: 10.1038/nature05395. PubMed PMID: 17183270; PubMed Central PMCID: PMCPMC2637441.

106. Pernas L, Adomako-Ankomah Y, Shastri AJ, Ewald SE, Treeck M, Boyle JP, et al. Toxoplasma effector MAF1 mediates recruitment of host mitochondria and impacts the host response. PLoS Biol. 2014;12(4):e1001845. doi: 10.1371/journal.pbio.1001845. PubMed PMID: 24781109; PubMed Central PMCID: PMCPMC4004538.

107. Rosowski EE, Lu D, Julien L, Rodda L, Gaiser RA, Jensen KD, et al. Strain-specific activation of the NF-kappaB pathway by GRA15, a novel Toxoplasma gondii dense granule protein. J Exp Med. 2011;208(1):195-212. doi: 10.1084/jem.20100717. PubMed PMID: 21199955; PubMed Central PMCID: PMCPMC3023140.

108. Braun L, Brenier-Pinchart MP, Yogavel M, Curt-Varesano A, Curt-Bertini RL, Hussain T, et al. A Toxoplasma dense granule protein, GRA24, modulates the early immune response to infection by promoting a direct and sustained host p38 MAPK activation. J Exp Med. 2013;210(10):2071-86. doi: 10.1084/jem.20130103. PubMed PMID: 24043761; PubMed Central PMCID: PMCPMC3782045.

109. Shastri AJ, Marino ND, Franco M, Lodoen MB, Boothroyd JC. GRA25 is a novel virulence factor of Toxoplasma gondii and influences the host immune response. Infect Immun. 2014;82(6):2595-605. doi: 10.1128/IAI.01339-13. PubMed PMID: 24711568; PubMed Central PMCID: PMCPMC4019154.

110. Lindsay DS, Dubey JP. Neospora caninum (Protozoa: apicomplexa) infections in mice. J Parasitol. 1989;75(5):772-9. PubMed PMID: 2795380.

111. Lindsay DS, Lenz SD, Cole RA, Dubey JP, Blagburn BL. Mouse model for central nervous system Neospora caninum infections. J Parasitol. 1995;81(2):313-5. PubMed PMID: 7707216.

112. Khan IA, Schwartzman JD, Fonseka S, Kasper LH. Neospora caninum: role for immune cytokines in host immunity. Exp Parasitol. 1997;85(1):24-34. doi: 10.1006/expr.1996.4110. PubMed PMID: 9024199.

113. Baszler TV, Long MT, McElwain TF, Mathison BA. Interferon-gamma and interleukin-12 mediate protection to acute Neospora caninum infection in BALB/c mice. Int J Parasitol. 1999;29(10):1635-46. PubMed PMID: 10608450.

114. Nishikawa Y, Tragoolpua K, Inoue N, Makala L, Nagasawa H, Otsuka H, et al. In the absence of endogenous gamma interferon, mice acutely infected with Neospora caninum succumb to a lethal immune response characterized by inactivation of peritoneal macrophages. Clin Diagn Lab Immunol. 2001;8(4):811-6. doi: 10.1128/CDLI.8.4.811-817.2001. PubMed PMID: 11427432; PubMed Central PMCID: PMCPMC96148.

115. Botelho AS, Teixeira L, Correia-da-Costa JM, Faustino AM, Castro AG, Vilanova M. Neospora caninum: high susceptibility to the parasite in C57BL/10ScCr mice. Exp Parasitol. 2007;115(1):68-75. doi: 10.1016/j.exppara.2006.06.004. PubMed PMID: 16889774.

116. Collantes-Fernandez E, Arrighi RB, Alvarez-Garcia G, Weidner JM, Regidor-Cerrillo J, Boothroyd JC, et al. Infected dendritic cells facilitate systemic dissemination and transplacental passage of the obligate intracellular parasite Neospora caninum in mice. PLoS One. 2012;7(3):e32123. doi: 10.1371/journal.pone.0032123. PubMed PMID: 22403627; PubMed Central PMCID: PMCPMC3293873.

117. Sibley LD, Boothroyd JC. Construction of a molecular karyotype for Toxoplasma gondii. Mol Biochem Parasitol. 1992;51(2):291-300. PubMed PMID: 1574087.

118. Sibley LD, LeBlanc AJ, Pfefferkorn ER, Boothroyd JC. Generation of a restriction fragment length polymorphism linkage map for Toxoplasma gondii. Genetics. 1992;132(4):1003-15. PubMed PMID: 1360931; PubMed Central PMCID: PMCPMC1205223.

119. Lindsay DS, Dubey JP, Duncan RB. Confirmation that the dog is a definitive host for Neospora caninum. Vet Parasitol. 1999;82(4):327-33. PubMed PMID: 10384909.

120. Dubey JP, Lindsay DS. Gerbils (Meriones unguiculatus) are highly susceptible to oral infection with Neospora caninum oocysts. Parasitol Res. 2000;86(2):165-8. PubMed PMID: 10685849.

121. Atkinson R, Harper PA, Ryce C, Morrison DA, Ellis JT. Comparison of the biological characteristics of two isolates of Neospora caninum. Parasitology. 1999;118 (Pt 4):363-70. PubMed PMID: 10340326.

122. Pereira Garcia-Melo D, Regidor-Cerrillo J, Collantes-Fernandez E, Aguado-Martinez A, Del Pozo I, Minguijon E, et al. Pathogenic characterization in mice of Neospora caninum isolates obtained from asymptomatic calves. Parasitology. 2010;137(7):1057-68. doi: 10.1017/S0031182009991855. PubMed PMID: 20233488.

123. Miller CM, Quinn HE, Windsor PA, Ellis JT. Characterisation of the first Australian isolate of Neospora caninum from cattle. Aust Vet J. 2002;80(10):620-5. PubMed PMID: 12465814.

124. Riahi H, Darde ML, Bouteille B, Leboutet MJ, Pestre-Alexandre M. Hammondia hammondi cysts in cell cultures. J Parasitol. 1995;81(5):821-4. PubMed PMID: 7472890.

125. Zavodnik IB. [Mitochondria, calcium homeostasis and calcium signaling]. Biomed Khim. 2016;62(3):311-7. doi: 10.18097/PBMC20166203311. PubMed PMID: 27420625.

126. Landes T, Martinou JC. Mitochondrial outer membrane permeabilization during apoptosis: the role of mitochondrial fission. Biochim Biophys Acta. 2011;1813(4):540-5. doi: 10.1016/j.bbamcr.2011.01.021. PubMed PMID: 21277336.

127. Sandhir R, Halder A, Sunkaria A. Mitochondria as a centrally positioned hub in the innate immune response. Biochim Biophys Acta. 2016. doi: 10.1016/j.bbadis.2016.10.020. PubMed PMID: 27794419.

128. Lund K, Ziola B. Cell sonicates used in the analysis of how measles and herpes simplex type 1 virus infections influence Vero cell mitochondrial calcium uptake. Can J Biochem Cell Biol. 1985;63(11):1194-7. PubMed PMID: 4084857.

129. Li Y, Boehning DF, Qian T, Popov VL, Weinman SA. Hepatitis C virus core protein increases mitochondrial ROS production by stimulation of Ca2+ uniporter activity. FASEB J. 2007;21(10):2474-85. doi: 10.1096/fj.06-7345com. PubMed PMID: 17392480.

130. Griffin SD, Harvey R, Clarke DS, Barclay WS, Harris M, Rowlands DJ. A conserved basic loop in hepatitis C virus p7 protein is required for amantadine-sensitive ion channel activity in mammalian cells but is dispensable for localization to mitochondria. J Gen Virol. 2004;85(Pt 2):451-61. doi: 10.1099/vir.0.19634-0. PubMed PMID: 14769903.

131. Marshall WL, Yim C, Gustafson E, Graf T, Sage DR, Hanify K, et al. Epstein-Barr virus encodes a novel homolog of the bcl-2 oncogene that inhibits apoptosis and associates with Bax and Bak. J Virol. 1999;73(6):5181-5. PubMed PMID: 10233985; PubMed Central PMCID: PMCPMC112567.

132. Henderson S, Huen D, Rowe M, Dawson C, Johnson G, Rickinson A. Epstein-Barr viruscoded BHRF1 protein, a viral homologue of Bcl-2, protects human B cells from programmed cell death. Proc Natl Acad Sci U S A. 1993;90(18):8479-83. PubMed PMID: 8397406; PubMed Central PMCID: PMCPMC47380.

133. Rahmani Z, Huh KW, Lasher R, Siddiqui A. Hepatitis B virus X protein colocalizes to mitochondria with a human voltage-dependent anion channel, HVDAC3, and alters its transmembrane potential. J Virol. 2000;74(6):2840-6. PubMed PMID: 10684300; PubMed Central PMCID: PMCPMC111774.

134. Lu YW, Chen WN. Human hepatitis B virus X protein induces apoptosis in HepG2 cells: role of BH3 domain. Biochem Biophys Res Commun. 2005;338(3):1551-6. doi: 10.1016/j.bbrc.2005.10.117. PubMed PMID: 16274670.

135. Li XD, Sun L, Seth RB, Pineda G, Chen ZJ. Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. Proc Natl Acad Sci U S A. 2005;102(49):17717-22. doi: 10.1073/pnas.0508531102. PubMed PMID: 16301520; PubMed Central PMCID: PMCPMC1308909.

136. Meylan E, Curran J, Hofmann K, Moradpour D, Binder M, Bartenschlager R, et al. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature. 2005;437(7062):1167-72. doi: 10.1038/nature04193. PubMed PMID: 16177806.

137. Chen Z, Benureau Y, Rijnbrand R, Yi J, Wang T, Warter L, et al. GB virus B disrupts RIG-I signaling by NS3/4A-mediated cleavage of the adaptor protein MAVS. J Virol. 2007;81(2):964-76. doi: 10.1128/JVI.02076-06. PubMed PMID: 17093192; PubMed Central PMCID: PMCPMC1797450.

138. Horwitz MA. Formation of a novel phagosome by the Legionnaires' disease bacterium (Legionella pneumophila) in human monocytes. J Exp Med. 1983;158(4):1319-31. PubMed PMID: 6619736; PubMed Central PMCID: PMC2187375.

139. Neumeister B, Faigle M, Lauber K, Northoff H, Wesselborg S. Legionella pneumophila induces apoptosis via the mitochondrial death pathway. Microbiology. 2002;148(Pt 11):3639-50. doi: 10.1099/00221287-148-11-3639. PubMed PMID: 12427954.

140. Zhu W, Hammad LA, Hsu F, Mao Y, Luo ZQ. Induction of caspase 3 activation by multiple Legionella pneumophila Dot/Icm substrates. Cell Microbiol. 2013;15(11):1783-95. doi: 10.1111/cmi.12157. PubMed PMID: 23773455; PubMed Central PMCID: PMC3797225.

141. Matsumoto A, Bessho H, Uehira K, Suda T. Morphological studies of the association of mitochondria with chlamydial inclusions and the fusion of chlamydial inclusions. J Electron Microsc (Tokyo). 1991;40(5):356-63. PubMed PMID: 1666645.

142. Suwa T, Itakura C. Ultrastructural studies of chlamydia-infected air sacs of chicks. Avian Pathol. 1992;21(3):443-52. doi: 10.1080/03079459208418862. PubMed PMID: 18670959.

143. Fischer SF, Harlander T, Vier J, Hacker G. Protection against CD95-induced apoptosis by chlamydial infection at a mitochondrial step. Infect Immun. 2004;72(2):1107-15. PubMed PMID: 14742558; PubMed Central PMCID: PMCPMC321620.

144. Otto SP, Whitton J. Polyploid incidence and evolution. Annu Rev Genet. 2000;34:401-37. doi: 10.1146/annurev.genet.34.1.401. PubMed PMID: 11092833.

145. Wolfe KH. Origin of the Yeast Whole-Genome Duplication. PLoS Biol. 2015;13(8):e1002221. doi: 10.1371/journal.pbio.1002221. PubMed PMID: 26252643; PubMed Central PMCID: PMC4529243.

146. Dehal P, Boore JL. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 2005;3(10):e314. doi: 10.1371/journal.pbio.0030314. PubMed PMID: 16128622; PubMed Central PMCID: PMC1197285.

147. Masterson J. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science. 1994;264(5157):421-4. doi: 10.1126/science.264.5157.421. PubMed PMID: 17836906.

148. Brosius J. Retroposons--seeds of evolution. Science. 1991;251(4995):753. PubMed PMID: 1990437.

149. Zhang J. Evolution by gene duplication: an update. Trends in Ecology and Evolution. 2003;18(6):7.

150. Sturtevant AH. The Effects of Unequal Crossing over at the Bar Locus in Drosophila. Genetics. 1925;10(2):117-47. PubMed PMID: 17246266; PubMed Central PMCID: PMC1200852.

151. Adomako-Ankomah Y, Wier GM, Borges AL, Wand HE, Boyle JP. Differential locus expansion distinguishes Toxoplasmatinae species and closely related strains of Toxoplasma gondii. MBio. 2014;5(1):e01003-13. doi: 10.1128/mBio.01003-13. PubMed PMID: 24496792; PubMed Central PMCID: PMCPMC3950507.

152. Qian W, Zhang J. Gene dosage and gene duplicability. Genetics. 2008;179(4):2319-24. doi: 10.1534/genetics.108.090936. PubMed PMID: 18689880; PubMed Central PMCID: PMC2516101.

153. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. Science. 2000;290(5494):1151-5. PubMed PMID: 11073452.

154. Force A, Lynch M, Pickett FB, Amores A, Yan YL, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 1999;151(4):1531-45. PubMed PMID: 10101175; PubMed Central PMCID: PMC1460548.

155. Cao L, Huang Q, Wu Z, Cao DD, Ma Z, Xu Q, et al. Neofunctionalization of zona pellucida proteins enhances freeze-prevention in the eggs of Antarctic notothenioids. Nat Commun. 2016;7:12987. doi: 10.1038/ncomms12987. PubMed PMID: 27698404; PubMed Central PMCID: PMC5059455.

156. Dubey JP, Barr BC, Barta JR, Bjerkas I, Bjorkman C, Blagburn BL, et al. Redescription of Neospora caninum and its differentiation from related coccidia. Int J Parasitol. 2002;32(8):929-46. PubMed PMID: 12076623.

157. Rettigner C, Leclipteux T, De Meerschman F, Focant C, Losson B. Survival, immune responses and tissue cyst production in outbred (Swiss white) and inbred (CBA/Ca) strains of mice experimentally infected with Neospora caninum tachyzoites. Veterinary research. 2004;35(2):225-32. Epub 2004/04/22. doi: 10.1051/vetres:2004005

V4205 [pii]. PubMed PMID: 15099498.

158. Wolf JB, Kunstner A, Nam K, Jakobsson M, Ellegren H. Nonlinear dynamics of nonsynonymous (dN) and synonymous (dS) substitution rates affects inference of selection. Genome Biol Evol. 2009;1:308-19. doi: 10.1093/gbe/evp030. PubMed PMID: 20333200; PubMed Central PMCID: PMCPMC2817425.

159. Zhang Z, Li J, Zhao XQ, Wang J, Wong GK, Yu J. KaKs\_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics. 2006;4(4):259-63. doi: 10.1016/S1672-0229(07)60007-2. PubMed PMID: 17531802; PubMed Central PMCID: PMCPMC5054075.

160. Boyle JP, Saeij JP, Harada SY, Ajioka JW, Boothroyd JC. Expression quantitative trait locus mapping of Toxoplasma genes reveals multiple mechanisms for strain-specific differences in gene expression. Eukaryot Cell. 2008;7(8):1403-14. Epub 2008/06/17. doi: 10.1128/EC.00073-08. PubMed PMID: 18552283; PubMed Central PMCID: PMC2519772.

161. Bahl A, Davis PH, Behnke M, Dzierszinski F, Jagalur M, Chen F, et al. A novel multifunctional oligonucleotide microarray for Toxoplasma gondii. BMC Genomics. 2010;11(1):603. Epub 2010/10/27. doi: 1471-2164-11-603 [pii]

10.1186/1471-2164-11-603. PubMed PMID: 20974003.

162. Pernas L, Boothroyd JC. Association of host mitochondria with the parasitophorous vacuole during Toxoplasma infection is not dependent on rhoptry proteins ROP2/8. Int J Parasitol. 2010;40(12):1367-71. Epub 2010/07/20. doi: S0020-7519(10)00249-3 [pii]

10.1016/j.ijpara.2010.07.002. PubMed PMID: 20637758; PubMed Central PMCID: PMC2939271.

163. Dubey JP, Tilahun G, Boyle JP, Schares G, Verma SK, Ferreira LR, et al. Molecular and Biological Characterization of First Isolates of Hammondia hammondi from Cats from Ethiopia. J Parasitol. 2013;99(4):614-8. Epub 2013/03/23. doi: 10.1645/12-51.1. PubMed PMID: 23517380.

164. Walzer KA, Adomako-Ankomah Y, Dam RA, Herrmann DC, Schares G, Dubey JP, et al. Hammondia hammondi, an avirulent relative of Toxoplasma gondii, has functional orthologs of known T. gondii virulence genes. Proc Natl Acad Sci U S A. 2013;110(18):7446-51. doi: 10.1073/pnas.1304322110. PubMed PMID: 23589877; PubMed Central PMCID: PMCPMC3645575.

165. Saeij JP, Boyle JP, Boothroyd JC. Differences among the three major strains of Toxoplasma gondii and their specific interactions with the infected host. Trends Parasitol. 2005;21(10):476-81. Epub 2005/08/16. doi: 10.1016/j.pt.2005.08.001. PubMed PMID: 16098810.

166. Dubey JP, Liddell S, Mattson D, Speert CA, Howe DK, Jenkins MC. Characterization of the Oregon isolate of Neospora hughesi from a horse. J Parasitol. 2001;87(2):345-53. doi: 10.1645/0022-3395(2001)087[0345:COTOIO]2.0.CO;2. PubMed PMID: 11318565.

167. Sinai AP, Webster P, Joiner KA. Association of host cell endoplasmic reticulum and mitochondria with the Toxoplasma gondii parasitophorous vacuole membrane: a high affinity interaction. J Cell Sci. 1997;110 (Pt 17):2117-28. PubMed PMID: 9378762.

168. Dubey JP. Tissue cyst tropism in Toxoplasma gondii: a comparison of tissue cyst formation in organs of cats, and rodents fed oocysts. Parasitology. 1997;115 (Pt 1):15-20. PubMed PMID: 9226953.

169. Dubey JP, Lindsay DS, Speer CA. Structures of Toxoplasma gondii tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. Clin Microbiol Rev. 1998;11(2):267-99. PubMed PMID: 9564564; PubMed Central PMCID: PMC106833.

170. Garin JP, Eyles DE. [Spiramycin therapy of experimental toxoplasmosis in mice]. Presse Med. 1958;66(42):957-8. PubMed PMID: 13567489.

171. Chew WK, Wah MJ, Ambu S, Segarra I. Toxoplasma gondii: determination of the onset of chronic infection in mice and the in vitro reactivation of brain cysts. Exp Parasitol. 2012;130(1):22-5. doi: 10.1016/j.exppara.2011.10.004. PubMed PMID: 22027550.

172. Watts E, Zhao Y, Dhara A, Eller B, Patwardhan A, Sinai AP. Novel Approaches Reveal that Toxoplasma gondii Bradyzoites within Tissue Cysts Are Dynamic and Replicating Entities In Vivo. MBio. 2015;6(5):e01155-15. doi: 10.1128/mBio.01155-15. PubMed PMID: 26350965; PubMed Central PMCID: PMC4600105.

173. Boothroyd JC, Black M, Bonnefoy S, Hehl A, Knoll LJ, Manger ID, et al. Genetic and biochemical analysis of development in Toxoplasma gondii. Philos Trans R Soc Lond B Biol Sci. 1997;352(1359):1347-54. doi: 10.1098/rstb.1997.0119. PubMed PMID: 9355126; PubMed Central PMCID: PMC1692023.

174. Zhang YW, Halonen SK, Ma YF, Wittner M, Weiss LM. Initial characterization of CST1, a Toxoplasma gondii cyst wall glycoprotein. Infect Immun. 2001;69(1):501-7. doi: 10.1128/IAI.69.1.501-507.2001. PubMed PMID: 11119543; PubMed Central PMCID: PMC97909.

175. Tomita T, Bzik DJ, Ma YF, Fox BA, Markillie LM, Taylor RC, et al. The Toxoplasma gondii cyst wall protein CST1 is critical for cyst wall integrity and promotes bradyzoite persistence. PLoS Pathog. 2013;9(12):e1003823. doi: 10.1371/journal.ppat.1003823. PubMed PMID: 24385904; PubMed Central PMCID: PMC3873430.

176. Buchholz KR, Bowyer PW, Boothroyd JC. Bradyzoite pseudokinase 1 is crucial for efficient oral infectivity of the Toxoplasma gondii tissue cyst. Eukaryot Cell. 2013;12(3):399-410. doi: 10.1128/EC.00343-12. PubMed PMID: 23291621; PubMed Central PMCID: PMC3629768.

177. Jones NG, Wang Q, Sibley LD. Secreted protein kinases regulate cyst burden during chronic toxoplasmosis. Cell Microbiol. 2016. doi: 10.1111/cmi.12651. PubMed PMID: 27450947.

178. Fox BA, Rommereim LM, Guevara RB, Falla A, Hortua Triana MA, Sun Y, et al. The Toxoplasma gondii Rhoptry Kinome Is Essential for Chronic Infection. MBio. 2016;7(3). doi: 10.1128/mBio.00193-16. PubMed PMID: 27165797; PubMed Central PMCID: PMC4959664.

179. Wen X, Kudo T, Payne L, Wang X, Rodgers L, Suzuki Y. Predominant interferongamma-mediated expression of CXCL9, CXCL10, and CCL5 proteins in the brain during chronic infection with Toxoplasma gondii in BALB/c mice resistant to development of toxoplasmic encephalitis. J Interferon Cytokine Res. 2010;30(9):653-60. doi: 10.1089/jir.2009.0119. PubMed PMID: 20626297; PubMed Central PMCID: PMC2963637.

180. Suzuki Y, Joh K. Effect of the strain of Toxoplasma gondii on the development of toxoplasmic encephalitis in mice treated with antibody to interferon-gamma. Parasitol Res. 1994;80(2):125-30. PubMed PMID: 8202451.

181. Adomako-Ankomah Y, English ED, Danielson JJ, Pernas LF, Parker ML, Boulanger MJ, et al. Host Mitochondrial Association Evolved in the Human Parasite Toxoplasma gondii via Neofunctionalization of a Gene Duplicate. Genetics. 2016. doi: 10.1534/genetics.115.186270. PubMed PMID: 26920761.

182. Suzuki Y, Joh K, Kwon OC, Yang Q, Conley FK, Remington JS. MHC class I gene(s) in the D/L region but not the TNF-alpha gene determines development of toxoplasmic encephalitis in mice. J Immunol. 1994;153(10):4649-54. PubMed PMID: 7963536.

183. Blanchard N, Dunay IR, Schluter D. Persistence of Toxoplasma gondii in the central nervous system: a fine-tuned balance between the parasite, the brain and the immune system. Parasite Immunol. 2015;37(3):150-8. doi: 10.1111/pim.12173. PubMed PMID: 25573476.

184. Suzuki Y, Joh K, Orellana MA, Conley FK, Remington JS. A gene(s) within the H-2D region determines the development of toxoplasmic encephalitis in mice. Immunology. 1991;74(4):732-9. PubMed PMID: 1783431; PubMed Central PMCID: PMCPMC1384788.

185. Dincel GC, Atmaca HT. Nitric oxide production increases during Toxoplasma gondii encephalitis in mice. Exp Parasitol. 2015;156:104-12. doi: 10.1016/j.exppara.2015.06.009. PubMed PMID: 26115941.

186. Moreno SN, Docampo R. Calcium regulation in protozoan parasites. Curr Opin Microbiol. 2003;6(4):359-64. PubMed PMID: 12941405.

187. Nash PB, Purner MB, Leon RP, Clarke P, Duke RC, Curiel TJ. Toxoplasma gondiiinfected cells are resistant to multiple inducers of apoptosis. J Immunol. 1998;160(4):1824-30. PubMed PMID: 9469443.

188. Angeloni MB, Guirelli PM, Franco PS, Barbosa BF, Gomes AO, Castro AS, et al. Differential apoptosis in BeWo cells after infection with highly (RH) or moderately (ME49) virulent strains of Toxoplasma gondii is related to the cytokine profile secreted, the death receptor Fas expression and phosphorylated ERK1/2 expression. Placenta. 2013;34(11):973-82. doi: 10.1016/j.placenta.2013.09.005. PubMed PMID: 24074900.

189. Reid AJ, Vermont SJ, Cotton JA, Harris D, Hill-Cawthorne GA, Konen-Waisman S, et al. Comparative genomics of the apicomplexan parasites Toxoplasma gondii and Neospora caninum: Coccidia differing in host range and transmission strategy. PLoS Pathog. 2012;8(3):e1002567. doi: 10.1371/journal.ppat.1002567. PubMed PMID: 22457617; PubMed Central PMCID: PMCPMC3310773.

190. Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002;12(4):656-64. doi: 10.1101/gr.229202. Article published online before March 2002. PubMed PMID: 11932250; PubMed Central PMCID: PMC187518.

191. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841-2. doi: 10.1093/bioinformatics/btq033. PubMed PMID: 20110278; PubMed Central PMCID: PMC2832824.

192. Lorenzi H, Khan A, Behnke MS, Namasivayam S, Swapna LS, Hadjithomas M, et al. Local admixture of amplified and diversified secreted pathogenesis determinants shapes mosaic Toxoplasma gondii genomes. Nat Commun. 2016;7:10147. doi: 10.1038/ncomms10147. PubMed PMID: 26738725; PubMed Central PMCID: PMC4729833.

193. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9. doi: 10.1038/nmeth.1923. PubMed PMID: 22388286; PubMed Central PMCID: PMC3322381.

194. Nicol JW, Helt GA, Blanchard SG, Jr., Raja A, Loraine AE. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. Bioinformatics.
2009;25(20):2730-1. doi: 10.1093/bioinformatics/btp472. PubMed PMID: 19654113; PubMed Central PMCID: PMC2759552.

195. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 2013;30(12):2725-9. doi: 10.1093/molbev/mst197. PubMed PMID: 24132122; PubMed Central PMCID: PMC3840312.

196. Nei M, Kumar, S. Molecular Evolution and Phylogenetics. Oxford: Oxford University Press; 2000.

197. Zhang J, Rosenberg HF, Nei M. Positive Darwinian selection after gene duplication in primate ribonuclease genes. Proc Natl Acad Sci U S A. 1998;95(7):3708-13. PubMed PMID: 9520431; PubMed Central PMCID: PMC19901.

198. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539. doi: 10.1038/msb.2011.75. PubMed PMID: 21988835; PubMed Central PMCID: PMCPMC3261699.

199. English ED, Adomako-Ankomah Y, Boyle JP. Secreted effectors in Toxoplasma gondii and related species: determinants of host range and pathogenesis? Parasite Immunol. 2015;37(3):127-40. doi: 10.1111/pim.12166. PubMed PMID: 25655311; PubMed Central PMCID: PMCPMC4359005.

200. Olias P, Etheridge RD, Zhang Y, Holtzman MJ, Sibley LD. Toxoplasma Effector Recruits the Mi-2/NuRD Complex to Repress STAT1 Transcription and Block IFN-gamma-Dependent Gene Expression. Cell Host Microbe. 2016;20(1):72-82. doi: 10.1016/j.chom.2016.06.006. PubMed PMID: 27414498; PubMed Central PMCID: PMCPMC4947229.

201. Gay G, Braun L, Brenier-Pinchart MP, Vollaire J, Josserand V, Bertini RL, et al. Toxoplasma gondii TgIST co-opts host chromatin repressors dampening STAT1-dependent gene regulation and IFN-gamma-mediated host defenses. J Exp Med. 2016;213(9):1779-98. doi: 10.1084/jem.20160340. PubMed PMID: 27503074; PubMed Central PMCID: PMCPMC4995087.

202. Lei T, Wang H, Liu J, Nan H, Liu Q. ROP18 is a key factor responsible for virulence difference between Toxoplasma gondii and Neospora caninum. PLoS One. 2014;9(6):e99744. doi: 10.1371/journal.pone.0099744. PubMed PMID: 24927100; PubMed Central PMCID: PMCPMC4057265.

203. Behnke MS, Fentress SJ, Mashayekhi M, Li LX, Taylor GA, Sibley LD. The polymorphic pseudokinase ROP5 controls virulence in Toxoplasma gondii by regulating the active kinase ROP18. PLoS Pathog. 2012;8(11):e1002992. doi: 10.1371/journal.ppat.1002992. PubMed PMID: 23144612; PubMed Central PMCID: PMCPMC3493473.

204. Boyle JP, Saeij JP, Harada SY, Ajioka JW, Boothroyd JC. Expression quantitative trait locus mapping of toxoplasma genes reveals multiple mechanisms for strain-specific differences in gene expression. Eukaryot Cell. 2008;7(8):1403-14. doi: 10.1128/EC.00073-08. PubMed PMID: 18552283; PubMed Central PMCID: PMCPMC2519772.

205. Walzer KA, Wier GM, Dam RA, Srinivasan AR, Borges AL, English ED, et al. Hammondia hammondi harbors functional orthologs of the host-modulating effectors GRA15 and ROP16 but is distinguished from Toxoplasma gondii by a unique transcriptional profile. Eukaryot Cell. 2014;13(12):1507-18. doi: 10.1128/EC.00215-14. PubMed PMID: 25280815; PubMed Central PMCID: PMCPMC4248688. 206. Pena-Castillo L, Hughes TR. Why are there still over 1000 uncharacterized yeast genes? Genetics. 2007;176(1):7-14. doi: 10.1534/genetics.107.074468. PubMed PMID: 17435240; PubMed Central PMCID: PMCPMC1893027.

207. Nadipuram SM, Kim EW, Vashisht AA, Lin AH, Bell HN, Coppens I, et al. In Vivo Biotinylation of the Toxoplasma Parasitophorous Vacuole Reveals Novel Dense Granule Proteins Important for Parasite Growth and Pathogenesis. MBio. 2016;7(4). doi: 10.1128/mBio.00808-16. PubMed PMID: 27486190; PubMed Central PMCID: PMCPMC4981711.