Genetic and Biological Characterization of *Toxoplasma gondii* from Uganda

Eva Irene Bontell

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds
Faculty of Biological Sciences
Research Institute of Integrative and Comparative Biology

April 2009

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others

This copy has been supplied on the understanding that it is copyright material and that no quotation from this thesis may be published without proper acknowledgement

Acknowledgements

First and foremost I want to thank my supervisors, Judith Smith and Johan Lindh, for introducing me to the weird and wonderful world of parasitology and for being enthusiastic and supportive throughout my PhD-work. I have really enjoyed working with you both!

Special thanks also go to....

Our collaborators in Uganda, Fred Kironde, Kaddu-Mulindwa and JD Kabasa for welcoming me in their labs, and to Erima, William and Maria, for their invaluable help with sample collection.

JP Dubey and lab members at USDA; Sundar, Oliver, Katherine and Luciana, for help with the bioassay and isolation of the Ugandan strains, and for the nice time in Beltsville.

Neil Hall and Kevin Ashelford in Liverpool deserve loads of credit for the hard work of assembling the TgCkUg2 genome, and for the preparation of user-friendly files!

Jon Boyle in Pittsburgh, for reanalyzing and sharing the SNP data from the three clonal lineages, and for helpful discussions at the Toxo100 meeting.

Various members of the Toxo community for interesting and helpful discussions, Chunlei Su for sharing unpublished data on strain diversity and the ToxoDB team (especially David Roos and Brian Brunk) for taking on the task to integrate the genome of TgCkUg2 in the database.

Marie Curie EST, NERC and SMI for funding.

In addition, I want to thank my friends and colleagues for support and friendship during these years;

Liz, for your never-ending toxo-enthusiasm and for showing me the best of UK, Yang, Aurore and Sam, I really enjoyed your company and miss you all, Beth, Zofia, Karl, Robbie, Elspeth, Hannah, James, Paul, Becky and lots of other people for help in the lab and for making my time in Leeds so memorable.

Everybody at the parasitology department at SMI, especially Johan A, Anders, Jonas, Jessica, Linda, and the big and mighty malaria group.

Helena & Jakob, Niloo and Karin for great times in Uganda (more to come I hope!)

My "old" friends for being there although I have been away, especially Malin for always providing a home in Stockholm when I needed it, Ebba (time for you to move back soon!), Jenny, Anna, Klara and all the other TB girls for being who you are!

Last but not least, my family for love and support and for always believing in me; mamma, pappa, Åsa, Lena and Josefin. Also the Bontell family, especially Anita and Göran, for all the good times in Timmervik.

Ludvig, mitt älskadeste lilla ludd, I can't even try to explain how much you mean to me. I love you, and I dedicate this thesis to you, so now you have to read it ☺

"It is written"

Slumdog Millionaire, 2008

Abstract

Toxoplasma gondii is a widely distributed protozoan parasite, estimated to infect 25-50% of the global human population. While most infections are asymptomatic, AIDS-associated toxoplasmic encephalitis causes a high level of morbidity and mortality and this is a particular problem in sub-Saharan Africa, where HIV affects a substantial part of the population. Despite the scale of the problem, virtually nothing has been known about the *T. gondii* strains present in Africa. The work presented in this thesis showed for the first time that all the three main clonal lineages of *T. gondii* are present and cause disease in African HIV-patients, and that genotype II was the most common in Uganda. Subsequent isolation of eight strains revealed that the *T. gondii* parasites in Uganda are highly similar to the strains found in Europe and North America, but over 1,200 novel mutations were identified and many of these were indicative of positive selection of genes active in the interface with the host. An important finding was the discovery of a rare natural recombinant strain, which possessed genetic elements from both genotype II and III and displayed an intermediate *in vivo* growth rate, compared with the genuine type II and III strains from the same area. Whole genome sequencing of this isolate revealed that sexual recombination may be more frequent than previously thought, but that selective pressures appear to favour a conserved genetic composition. The studies included in this thesis provide novel insight into the genotype and phenotype of *T. gondii* strains from Africa, and describe for the first time the genomics of a natural recombinant *T. gondii* strain.

Publications resulting from this research

Lindström-Bontell, I., Hall, N., Ashelford, K.E., Dubey, J.P., Boyle, J.P., Lindh, J., Smith, J.E., 2009. Whole genome sequencing of a recombinant type II/III *Toxoplasma gondii* strain from Uganda reveals chromosome sorting and local allelic variants. Genome Biology, accepted.

Lindström, I., Sundar, N., Lindh, J., Kironde, F., Kabasa, J.D., Kwok, O.C.H., Dubey, J.P., Smith, J.E., 2008. Isolation and genotyping of *Toxoplasma gondii* from Ugandan chickens reveals frequent multiple infections. Parasitology 135, 39-45.

Lindström, I., Kaddu-Mulindwa, D.H., Kironde, F., Lindh, J., 2006. Prevalence of latent and reactivated *Toxoplasma gondii* parasites in HIV-patients from Uganda. Acta Trop 100, 218-22.

Abbreviations

AIDS	acquired immunodeficiency syndrome
AMA	apical membrane antigen
BLAST	basic local alignment search tool
bp	base pairs
CD	cluster of differentiation
CDS	coding sequence
CNS	central nervous system
CT	computer tomography
DA	direct agglutination
DMEM	Dulbecco's modified Eagle's medium
DMSO	dimethyl sulfoxide
EDTA	ethylene diamine tetra-acetic acid
ELISA	enzyme linked immunosorbent assay
EST	expressed sequence tag
FBS	foetal bovine serum
GRA	dense granule protein
HAART	highly active antiretroviral therapy
HFF	human foreskin fibroblast
HIV	human immunodeficiency virus
IFN	interferon
IgG	immunoglobulin G
IL	interleukin
kbp	kilo base pairs
MAT	modified agglutination test
Mbp	mega base pairs
MIC	microneme protein
MJ	moving junction
MRCA	most recent common ancestor
nt	nucleotides
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEST	penicillin - streptomycin
PV	parasitophorous vacuole
Q-PCR	quantitative real-time PCR
RAPD	random amplification of polymorphic DNA
RFLP	restriction fragment length polymorphism
RON	rhoptry neck protein

ROP	rhoptry protein
rpm	revolutions per minute
SAG	surface antigen
SMI	Swedish Institute for Infectious Disease Control
SNP	single nucleotide polymorphism
SRS	SAG1-related sequence
TE	toxoplasmic encephalitis
TEM	transmission electron microscopy
USDA	United States Department of Agriculture

Table of contents

Acknowledgements I

Abstract III

Publications resulting from this research IV

Abbreviations V

Table of contents VII

List of figures XI

List of tables XIII

Chapter 1. Introduction 1

1.1 The parasite, *Toxoplasma gondii* 1
1.1.1 History and classification 1
1.1.2 Life cycle 3
1.1.3 Cellular and molecular biology 7
1.1.4 Invasion, stage conversion and egress 12

1.2 The disease, toxoplasmosis 16
1.2.1 Toxoplasmosis in immunocompetent adults 16
1.2.2 Congenital toxoplasmosis 17
1.2.3 Toxoplasmosis in immunocompromised hosts 18
1.2.4 Ocular toxoplasmosis 19
1.2.5 Outbreaks 19

1.3 Diversity of *Toxoplasma gondii* strains 21
1.3.1 Phenotypic and genetic differences between strains 21
1.3.2 Genotyping markers 22
1.3.3 Genotype distribution 23

1.4 Aims of this thesis 26

Chapter 2. Infection with *Toxoplasma gondii* is highly prevalent in Ugandan HIV-patients and all three lineages cause disease 27

2.1 Introduction 27
2.1.1 HIV and opportunistic diseases in Africa 27
2.1.2 Epidemiology of *T. gondii* and transmission in Africa 30
2.1.3 Toxoplasmosis in HIV-patients 33
2.1.4 Diagnosis and treatment 34
2.1.5 Aims 36

2.2 Materials and methods 37
2.2.1 Study site 37
2.2.2 Study populations and sample collection 37
2.2.3 *T. gondii* serology (Toxo-Screen DA) 39
2.2.4 DNA-extraction 40
2.2.5 Diagnostic and genotyping PCRs 40

2.2.6 Quantitative real-time PCR 42
2.2.7 Statistical analysis 43
2.2.8 Isolation in cell culture 43
2.2.9 Isolation in mice 44

2.3 Results 45
2.3.1 Patient demographics, CD4-counts and symptoms 45
2.3.2 *T. gondii* seroprevalence 48
2.3.3 Diagnostic PCR 49
2.3.4 Genotyping 52

2.4 Discussion 53
2.4.1 High seroprevalence and slow increase with age 53
2.4.2 Reactivation is common and may occur in seronegative patients 55
2.4.3 Diagnosis of latent and active toxoplasmosis in Uganda 57
2.4.4 Three lineages are present and cause reactivation 58

2.5 Conclusions 59

Chapter 3. Multiple *Toxoplasma gondii* infections are common in Ugandan chickens, providing opportunities for sexual recombination 60

3.1 Introduction 60
3.1.1 Chicken as indicator species 60
3.1.2 *T. gondii* in chickens from different regions 61
3.1.3 *T. gondii* in rats and mice 63
3.1.4 Aims 64

3.2 Materials and methods 65
3.2.1 Free-range chickens 65
3.2.2 Wild-caught mice and rats 65
3.2.3 Transport 66
3.2.4 *T. gondii* serology (MAT) 66
3.2.5 Bioassay and isolation in mice 67
3.2.6 Cryopreservation of *T. gondii* from mouse brains 68
3.2.7 Bioassay in cats 70
3.2.8 DNA-extraction 70
3.2.9 Genotyping PCRs 71
3.2.10 Sequencing and sequence analysis 73

3.3 Results 74
3.3.1 Serological screening of *T. gondii* infection in chickens and rodents 74
3.3.2 Direct genotyping of *T. gondii* from Ugandan chickens and rodents 77
3.3.3 Bioassay results and genotyping of isolates 80

3.4 Discussion 84
3.4.1 Serotitre was related to amplification and isolation success 84
3.4.2 Genotype related bias in isolation success 85
3.4.3 Similar genotype distribution in chickens and HIV-patients 86
3.4.4 High prevalence of multiple infections provides opportunities for recombination 88

3.5 Conclusions 89

Chapter 4. Characterization of a natural recombinant *Toxoplasma gondii* isolate from Uganda reveals an intermediate phenotype *in vivo* 90

4.1 Introduction 90
4.1.1 Sexual recombination in *T. gondii* 90
4.1.2 Origin of three clonal lineages I, II and III 93
4.1.3 Global population structure 96
4.1.4 Aims 97

4.2 Materials and methods 98
4.2.1 Toxoplasma gondii strains 98
4.2.2 Cell culture methods 98
4.2.3 Subpassage in mice 99
4.2.4 Isolation of *T. gondii* in cell culture from infected mouse brains 100
4.2.5 *In vivo* growth assay 101
4.2.6 Separation of multiple strains 101
4.2.7 *In vitro* growth assay 101
4.2.8 Transmission Electron Microscopy 102
4.2.9 Sequencing and sequence analysis 103
4.2.10 Random Amplification of Polymorphic DNA 105

4.3 Results 106
4.3.1 Isolation in mice 106
4.3.2 Quantification of parasites in mouse organs 109
4.3.3 Isolation and growth in cell culture 112
4.3.4 Morphology 115
4.3.5 Microsatellites and sequence analysis 116
4.3.6 RAPD analysis 120
4.3.7 Composite genotype of the recombinant strain 121

4.4 Discussion 123
4.4.1 Phenotypic differences and similarities between isolates 123
4.4.2 Limited genetic divergence among *T. gondii* from Uganda 125
4.4.3 Isolation of a natural recombinant strain 127

4.5 Conclusions 129

Chapter 5. Whole genome sequencing of a recombinant *Toxoplasma gondii* strain from Uganda reveals chromosome sorting and genes under selection 130

5.1 Introduction 130
5.1.1 The *T. gondii* genome and ToxoDB 130
5.1.2 Genetic differences between strains 133
5.1.3 Massively parallel genome sequencing technologies 135
3.1.4 Aims 138

5.2 Materials and methods 139
5.2.1 Parasite strain 139
5.2.2 DNA extraction and quality control 139
5.2.3 454-sequencing, creation of output files and bioinformatic analysis 140
5.2.4 Divergence time estimate 143

List of tables

Table 1-1 Global overview of HIV-infections in 2007 28
Table 1-2 Overview of common opportunistic infections in HIV-patients 30
Table 1-3 Overview of the five patient groups and the corresponding data sets 38
Table 1-4 Primers and restriction enzymes used for detection and genotyping of *T. gondii* 41
Table 2-5 Seroprevalence and reactivation rate of *T. gondii* in groups G1-G3 49
Table 3-1 Summary of chicken genotyping studies across the world 63
Table 3-2 Primers and enzymes used for detection and genotyping 71
Table 3-3 Direct genotyping results from chicken and rodent tissues 79
Table 3-4 Genotyping results from mice inoculated with tissue from infected chickens 81
Table 4-1 Summary of recombination patterns in progeny from experimental crosses 92
Table 4-2 Genomic locations and primers used for sequencing 104
Table 4-3 Origin and designation of retrieved isolates 107
Table 4-4 Detection and quantification of parasites and isolation success in mice 108
Table 4-5 Parasite density per gram of infected mouse organs 110
Table 4-6 Polymorphic sites in the UPRT-1 intron 118
Table 4-7 Summary of divergences between strains over >5 kb of sequence 119
Table 4-8 Summary of RAPD results 119
Table 5-1 Nucleotide differences between types I, II and III 134
Table 5-2 SNP colour scheme 141
Table 5-3 Primers used for complementary sequencing of additional Ugandan isolates 145
Table 5-4 Contig and coverage summary for all chromosomes and the apicoplast 146
Table 5-5 SNP summary for all chromosomes 147
Table 5-6 Details of SNPs in the Apicoplast genome 153
Table 5-7 Divergence time estimates based on the *T. gondii* intron mutation rate 154
Table 5-8 Divergence time estimate based on the relation between major and minor SNPs 155
Table 5-9 Genes with ≥2 novel SNPs in the coding region of TgCkUg2 157
Table 5-10 Amino acids at three polymorphic loci for Ugandan and reference strains 159
Table 5-11 Sequence similarity between Ugandan type II strains and the reference type I 159
Table 5-12 Different SNP types on TgCkUg2 chromosome VIII suggests recombination 160
Table 5-13 TgCkUg2 compared with circulating non-archetypal lineages and African strains 164

5.2.5 Targeted sequencing and phylogeny....144

5.3 Results....146

5.3.1 Coverage and contig metrics....146

5.3.2 SNP distribution....147

5.3.3 The apicoplast genome....152

5.3.4 Divergence time estimates....153

5.3.5 Unique SNPs and genes under selection....156

5.3.6 Local allelic variants and relationship between strains....158

5.4 Discussion....162

5.4.1 Recombination through chromosome sorting....162

5.4.2 Apicoplast genome copy number....164

5.4.3 Level of divergence and relatedness to other strains....164

5.4.4 Genes under selection....167

5.5 Conclusions....169

Chapter 6. General discussion....170

References....176

Appendices....202

List of figures

Figure 1-1 Cladogram placing *Toxoplasma* in the tree of life 2
Figure 1-2 Life cycle of *Toxoplasma gondii* 4
Figure 1-3 Morphology and major organelles of *T. gondii* 7
Figure 1-4 Distribution of major protein families over the 14 chromosomes (Ia-XII) 10
Figure 1-5 Morphology of the endysymbiotic organelles 11
Figure 1-6 Overview of the invasion process and mediating proteins 13
Figure 1-7 *T. gondii* bradyzoites in tissue cyst 15
Figure 1-8 *T. gondii* genotype distribution in different hosts and geographic regions 24
Figure 2-1 *T. gondii* seroprevalence in Africa 32
Figure 2-2 *T. gondii* direct agglutination test (Toxo-Screen DA) 39
Figure 2-3 Sensitivity testing of PCR assays used for *T. gondii* detection 42
Figure 2-4 Cumulative age distribution by gender and patient category 45
Figure 2-5 CD4-count in different patient categories 46
Figure 2-6 Prevalence of clinical symptoms in HIV-patients at different CD4-levels 47
Figure 2-7 Prevalence of clinical symptoms in seropositive and seronegative patients 48
Figure 2-8 Gel picture of nested B1-PCR 49
Figure 2-9 Box-plots showing the distribution of CD4-counts in different patient groups 50
Figure 2-10 Prevalence of clinical symptoms in PCR-positive and PCR-negative patients 51
Figure 2-11 Overview of serology, PCR-results and SAG2 genotyping for G1 52
Figure 2-12 Seroprevalence in different age groups in symptomatic and HIV-test patients. 54
Figure 3-1 Sample processing overview for Ugandan chickens 69
Figure 3-2 Sensitivity testing of the genotyping PCRs for BTUB and SAG3 72
Figure 3-3 Map of sampling areas 74
Figure 3-4 Chicken serology results, by place of origin 75
Figure 3-5 Chicken age distribution, by place of origin 76
Figure 3-6 Chicken seroprevalence in different age groups 76
Figure 3-7 Nested SAG2-3' screening results for Ugandan chicken and rodent samples 78
Figure 3-8 Double restriction enzyme digests of SAG1 amplicons 78
Figure 3-9 Chromatogram with double peaks indicative of co-infection 82
Figure 3-10 Amplification success in relation to serological titre in chicken 84
Figure 3-11 Genotype related bias in mouse bioassay 85
Figure 3-12 *T. gondii* strain distribution in chickens and HIV-patients from Uganda 87
Figure 4-1 Prevalence of recombinant chromosomes in experimental *T. gondii* crosses 93
Figure 4-2 Schematic of the cross that gave rise to the canonical three lineages I, II and III 94
Figure 4-3 Overview of *T. gondii* divergence over time. 95
Figure 4-4 Polymorphic loci selected for sequencing 103
Figure 4-5 Overview of the isolate retrieval procedure 106

Figure 4-6 Parasite density per gram of infected mouse organs 109
Figure 4-7 Growth rate *in vitro*, direct inoculation versus trypsinized parasites 112
Figure 4-8 Growth during the first passage in culture (trypsinized samples) 113
Figure 4-9 Growth during the eighth round of cell culture 114
Figure 4-10 TEM pictures of Ugandan isolates and the Me49 reference strain 115
Figure 4-11 Phylogram for microsatellite markers B17 and W35 117
Figure 4-12 Amino acid substitutions in Toxofilin 119
Figure 4-13 Amplification with random primers 12.2+12.4 121
Figure 4-14 Mapping of genotyped loci in TgCkUg2 122
Figure 4-15 Genetic subgroups among the Ugandan type II strains 126
Figure 5-1 Comparison of the genome browser in ToxoDB versions 4.3 (A) and 5.0 (B) 132
Figure 5-2 Shared and unique predicted genes among the three annotated strains 134
Figure 5-3 Chemistry of 454 sequencing 136
Figure 5-4 Reads from a single bead in 454 sequencing 136
Figure 5-5 Artemis screenshot showing SNP features 142
Figure 5-6 Schematics of the cross yielding types I, II, III, and the MRCA time estimates 143
Figure 5-7 SNP density graphs of chromosomes II, VIII and XI 149
Figure 5-8 Overview of SNP distribution in TgCkUg2 150
Figure 5-9 Overview of dominant SNP types in different chromosomal regions 150
Figure 5-10 SNP density graph overlaid with SNP type graph (chromosome IX) 151
Figure 5-11 Distribution of novel SNPs in TgCkUg2 156
Figure 5-12 Phylogenetic trees of the relationship between Ugandan and reference strains 161
Figure 5-13 SNP density graph of chromosome VIII, with polymorphic loci suggestive of recombination 163
Figure 5-14 Average SNP density in type II and type III derived chromosomes 166

Chapter 1 Introduction

1.1 The parasite, *Toxoplasma gondii*

1.1.1 History and classification

Toxoplasma gondii is an obligate intracellular protozoan parasite that infects a wide range of warm-blooded animals, including humans (Dubey and Beattie, 1988). It was first identified in 1908, by two separate research groups. Alfonso Splendore at the Bacteriological Institute in Sao Paolo found the parasite in a rabbit , but misidentified the protozoan as a species of *Leishmania* (Splendore, 1908). Charles Nicolle and Louis Mancaeux at the Pasteur Institute in Tunis discovered the parasite in the North African rodent *Ctenodactylus gundi* which, although misspelt, gave the parasite its species name (Nicolle and Manceaux, 1908). Initially they named the parasite *Leishmania gondii*, but in 1909 the genus *Toxoplasma* was introduced (Nicolle and Manceaux, 1909). The name refers to the crescent morphology of the organism; *toxon* means crescent-shaped in Greek and *plasma* means form (Hill et al., 2005). In the following decades, several species of *Toxoplasma* were named (Levine, 1977), but all of these were later reclassified as *T. gondii* or moved to other genera including *Hammondia*, *Atoxoplasma* and *Isospora*. *T. gondii* was first recognised as a human pathogen in the 1930s, when it was shown to cause encephalitis in neonates (Wolf et al., 1939), but the life cycle of the organism as well as transmission routes remained largely unknown until cats were identified as the definitive hosts in the late 1960's (see section 1.1.2).

Toxoplasma belongs to the phylum *Apicomplexa*, an ancient phylum that may date back as long as 1 billion years, based on analysis of small subunit rRNA (Sogin and Silberman, 1998). The phylum consists of several thousands of species, nearly all of which are obligate intracellular parasites (Levine, 1988), the main exception being the early branching extracellular Gregarines (Leander, 2008). Many important human and veterinary pathogens are found within the *Apicomplexa*, including

Plasmodium, the causative agent of malaria; *Cryptosporidium,* which causes gastrointestinal disease; *Eimeria,* an important pathogen in poultry; and *Theileria* and *Babesia,* which can cause fatal disease in cattle. An overview of the relationship between *Toxoplasma* and other *Apicomplexa* and their taxonomic placement in relation to other protozoan parasites and major organism groups is shown in a cladistic tree in Figure 1-1.

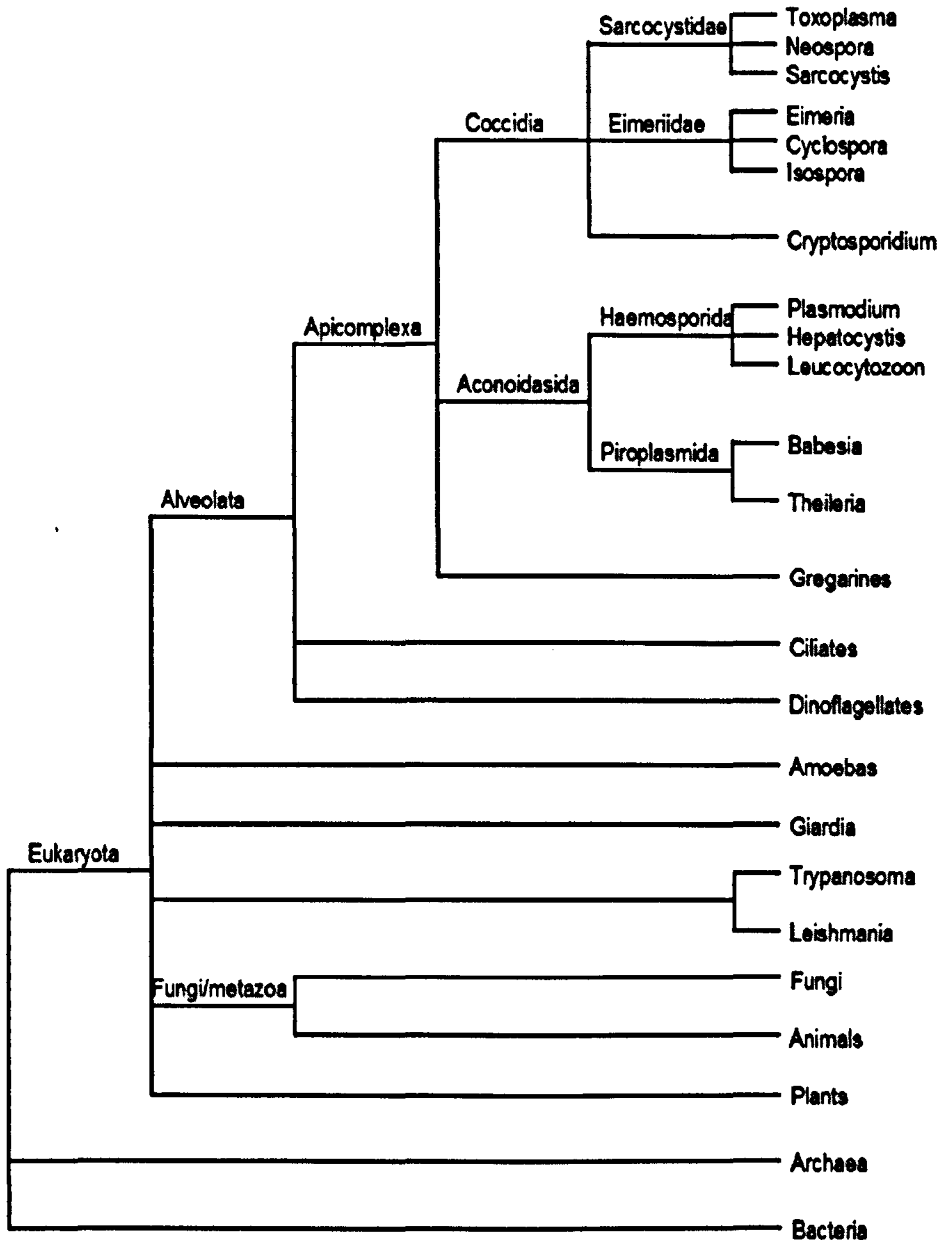


Figure 1-1. Cladogram placing *Toxoplasma* in the tree of life. *Toxoplasma* belongs to the phylum *Apicomplexa* and is a coccidian parasite closely related to *Neospora* and *Sarcocystis spp.* The tree is abstracted from the NCBI taxonomy database (http://www.ncbi.nlm. nih.gov/Taxonomy/).

The type characteristics of the phylum *Apicomplexa* are seen in organelles adapted to allow intracellular invasion, including the apical complex and a range of secretory organelles (Dubremetz et al., 1998). However, apicomplexan parasites differ in many other important aspects of their biology, including host and tissue specificity, intracellular location, organelle content, transmission route and both sexual and asexual development (Plattner and Soldati-Favre, 2008). Two remarkable traits that set *Toxoplasma* apart from the majority of apicomplexan parasites are the capacity to invade and propagate in almost any cell type in a wide range of host species as well as its ability of direct transmission between intermediate hosts, thus circumventing the sexual cycle. Most parasites are highly specific, for example there are dozens of species of *Plasmodium* with a definite intermediate host range (Rich and Ayala, 2003) and among the hundreds of *Eimeria* species, seven are exclusively specialized to chickens, each targeting a discrete area of the gut (Morris and Gasser, 2006). *T. gondii* on the other hand, has been isolated from hundreds of different host species all over the world (Tenter et al., 2000), and has even been shown to invade, but not persist in fish cells (Omata et al., 2005), even though cold-blooded animals are not within the known host range. The ease with which *T. gondii* is cultivated both *in vivo* and *in vitro*, and the possibility of genetic manipulation, has made it a favourite model organism for investigation of the cellular and molecular biology of the phylum (see 1.1.3).

1.1.2 Life cycle

T. gondii has a complex life cycle. Asexual replication can take place in a wide range of mammals and birds, while the sexual cycle, which results in shedding of environmentally resistant oocysts, only takes place in the epithelial cells of the small intestine of feline hosts (for a life cycle overview, see Figure 1-2). The understanding of the *T. gondii* life cycle began to unfold when J.K.A. Beverley proved successive congenital transmission in mice (Beverley, 1959) and G. Desmonts showed that consumption of raw meat was an important route of infection in humans (Desmonts et al., 1965). W.M. Hutchison provided the first indication that cats were important for transmission in 1965 (Hutchison, 1965), and two years later he

proposed that *T. gondii* was transmitted inside the eggs of the cat nematode *Toxocara cati* (Hutchison, 1967). In the years to follow, several different groups showed that *T. gondii* transmission in cats was independent of the worm (summarized in Frenkel, 1973), and the *T. gondii* oocyst was finally described in 1970 by J.P. Dubey, N.L. Miller and J.K. Frenkel (Dubey et al., 1970). Subsequently, a number of other species were investigated, but only felids including bobcats, ocelots and cougars, were found to shed oocysts (Dubey and Beattie, 1988).

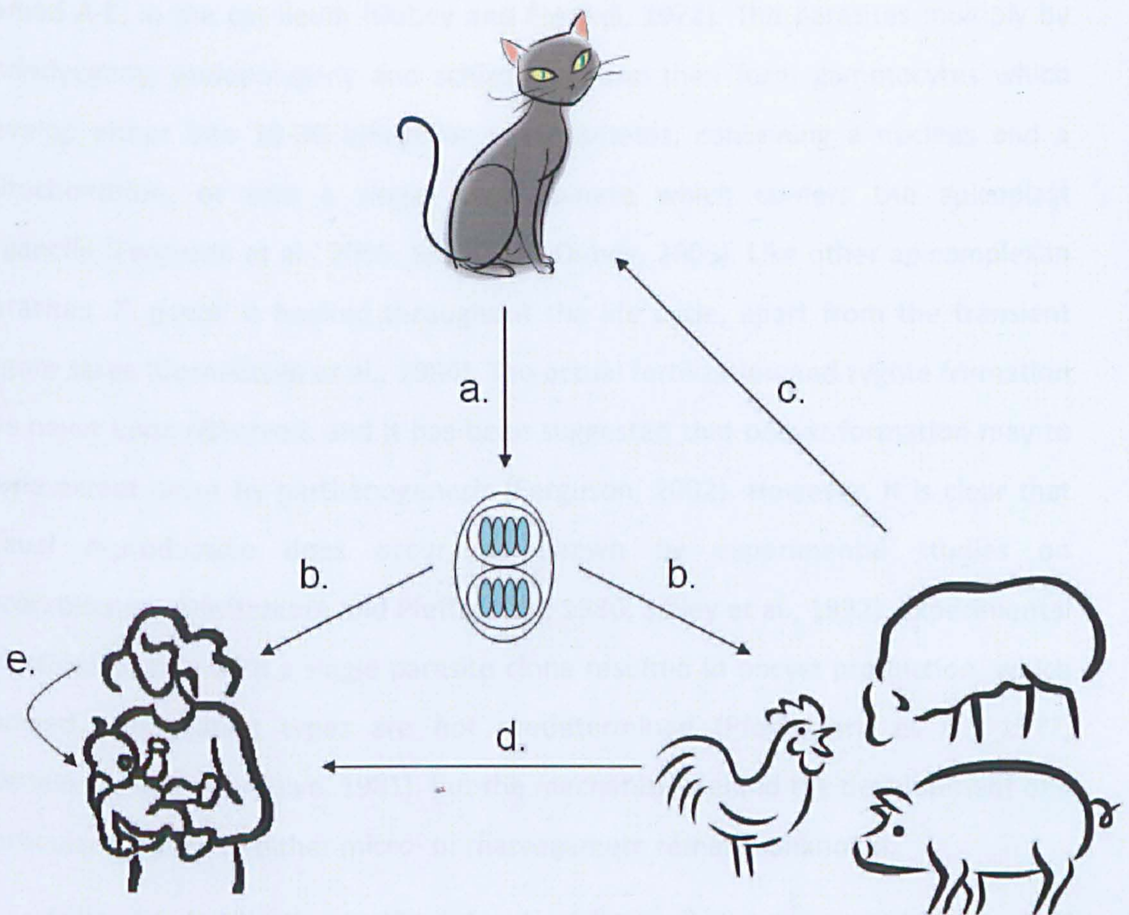


Figure 1-2. Life cycle of *Toxoplasma gondii*. (a) Felids are the definitive hosts and excrete *T. gondii* oocysts, which contaminate soil, water and vegetation. (b) Parasites ingested by intermediate hosts undergo initial rapid division and dissemination as tachyzoites, before transformation into bradyzoites in tissue cysts, which can persist for the lifetime of the host. (c) In order to return to the sexual cycle the parasites inside tissue cysts need to be eaten by a member of the cat family. (d) Intermediate hosts can become infected through bradyzoite-infected meat. (e) The tachyzoite stage can pass through the placenta and cause abortion or infected progeny; this is known as vertical transmission.

In addition to the highly infectious sporozoites inside oocysts, *T. gondii* appears in two other infectious stages: rapidly dividing tachyzoites (tachy=fast) and slowly multiplying bradyzoites (brady=slow) in tissue cysts (Frenkel, 1973). Felids normally become infected upon ingestion of bradyzoites in the tissue of their prey, where the parasite containing tissue cysts are most commonly found in muscle and neural tissue (Dubey, 1997a). The cysts are ruptured as they pass through the digestive tract and enter the epithelial cells of the cat small intestine, where asexual proliferation is followed by a sudden switch to sexual stages. The sexual cycle was first described in 1972 by Dubey and Frenkel, who discovered five successive stages, named A-E, in the cat ileum (Dubey and Frenkel, 1972). The parasites multiply by endodyogeny, endopolygeny and schizogony, and then form gametocytes which develop either into 10-30 biflagellar microgametes, containing a nucleus and a mitochondrion, or into a single macrogamete which confers the apicoplast organelle (Ferguson et al., 2005; Speer and Dubey, 2005). Like other apicomplexan parasites, *T. gondii* is haploid throughout the life cycle, apart from the transient zygote stage (Cornelissen et al., 1984). The actual fertilization and zygote formation has never been observed, and it has been suggested that oocyst formation may to some extent occur by parthenogenesis (Ferguson, 2002). However, it is clear that sexual reproduction does occur, as shown by experimental studies on recombination (Pfefferkorn and Pfefferkorn, 1980; Sibley et al., 1992). Experimental infections of cats with a single parasite clone resulted in oocyst production, which showed that mating types are not predetermined (Pfefferkorn et al., 1977; Cornelissen and Overdulve, 1981), but the mechanism behind the development of a particular parasite to either micro- or macrogamete remains unknown.

Following fertilization in the cat gut, millions of immature oocysts are shed with the faeces. Unsporulated oocysts contain a single primary sporoblast, which undergoes development outside of the host into two sporocysts, each containing four sporozoites (Ferguson and Dubremetz, 2007). The sporulated oocyst can survive in soil and water for several months, even years under cool and moist conditions (Dubey, 1998b). The oocyst wall protects the highly infectious sporozoites from environmental stresses, but enzymes in the gastrointestinal tract

destroy the oocyst and sporocyst walls when it enters the digestive system of an intermediate host (Dubey et al., 1998). Studies of the route of infection in mice show that sporozoites invade epithelial cells and then differentiate into rapidly dividing tachyzoites, which spread throughout the body via the blood-and lymph system (Barragan and Sibley, 2003). In tissue culture, this initial differentiation to the tachyzoite stage takes place within 24 hours post-invasion (Speer et al., 1995) and in the presence of a functioning immune system the tachyzoites convert into bradyzoites within 1-2 weeks (Dubey and Frenkel, 1976; Di Cristina et al., 2008). The parasites can remain latent in intracellular tissue cysts for extended periods, possibly for the lifetime of the host, although cyst rupture and reinvasion is likely to occur from time to time (Weiss and Kim, 2000), and if an infected animal is eaten by a cat the sexual cycle starts all over again.

Horizontal transmission between intermediate hosts occurs when an animal or human eats raw or undercooked meat containing tissue cysts and the course of events is essentially the same as upon oocyst ingestion; the cyst wall is digested by the gastric enzymes, the bradyzoites invade epithelial cells and transform into the tachyzoite stage (Dubey, 1997c). Although tachyzoites are often used for subcutaneous and intraperitoneal infection in laboratory mice, this stage is highly sensitive to both pepsin and trypsin digestion and the oral infectivity rate is low (Dubey, 1998a). Tachyzoites are, however, important in vertical transmission, where they can pass from mother to foetus through the placenta (Dubey, 1987), either following a recent infection or after reactivation of a latent infection, with conversion from bradyzoites to tachyzoites. Only primary infections are considered important in immunocompetent pregnant women (Foulon et al., 2000), but sporadic cases of congenital transmission have been reported in immunocompromised women with previously documented *T. gondii* infection (Bachmeyer et al., 2006; Cruz et al., 2007) and also as a result of superinfection with an atypical strain (Elbez-Rubinstein et al., 2009). In host species that are more sensitive to *T. gondii* infections, congenital transmission may be highly prevalent and even occur over successive generations as shown in experimental infection and

natural populations of mice (Beverley, 1959; Marshall et al., 2004) and sheep (Duncanson et al., 2001; Morley et al., 2008).

1.1.3 Cellular and molecular biology

The biology of *T. gondii* has been well studied, partly as a pathogen in its own right, but also as a model for other *Apicomplexa* since it is easily cultured and amenable to genetic manipulation (Kim and Weiss, 2004). Tachyzoites and bradyzoites can be obtained *in vivo* as well as *in vitro*, while sporozoites can only be acquired from cats. Ultrastructurally these stages are similar: approximately 7 by 2 µm, crescent-shaped or nearly elliptic, without cilia or flagella and with a pointed conoid tip, the apical complex (Figure 1-3), which contains a polar ring from which the microtubular cytoskeleton extends and through which secretory proteins are released during invasion (Dubey et al., 1998).

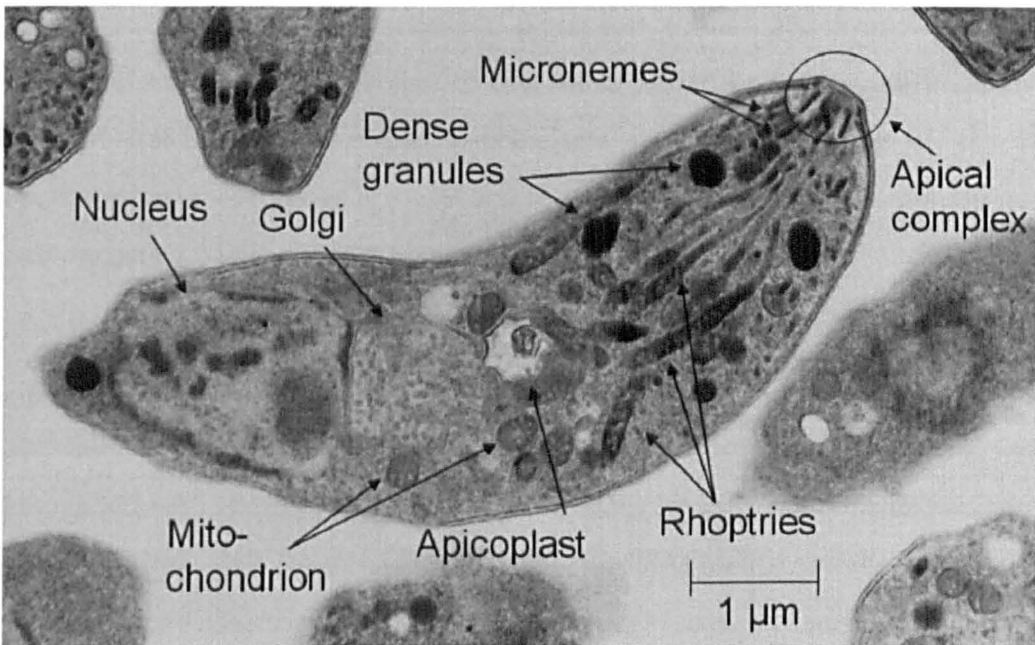


Figure 1-3. Morphology and major organelles of *T. gondii*. At the top right is the apical complex and nearby several small micronemes can be seen. The rhoptries extend from the apical end and their bulb appears partly hollow, which is typical of tachyzoites. Dense granules are seen both towards the anterior and posterior end. Nucleus, Golgi apparatus and Endoplasmic reticulum (not shown) are universal eukaryotic organelles present in *T. gondii*. Transmission electron microscopy (TEM) picture of strain Me49, for methodology see 4.2.8.

Bradyzoites can often be visually distinguished from tachyzoites by the posterior location of the nucleus, high number of amylopectin granules and electron-dense rhoptry organelles (Dubey et al., 1998). In addition, the life stages can be differentiated by the detection of stage-specific proteins, notably the surface antigens (Smith, 1995). This group of proteins includes the major tachyzoite surface antigens SAG1 and SAG2, which are highly immunogenic (Grimwood and Smith, 1996) as well as specific bradyzoite antigens including BAG1 and SRS9, against which antibodies are not found in natural infections (Kim and Boothroyd, 2005). SAG1 was early identified as important in host cell invasion (Grimwood and Smith, 1992) and allelic variants of SAG1 and 2 have been used extensively to differentiate between mouse-virulent and avirulent *T. gondii* strains (Howe and Sibley, 1995; Howe et al., 1996). In total, more than 160 members of this family (called SRS genes, which is short for SAG1-related sequences) have been identified through bioinformatic analysis and their short, intronless genes are found in tandem arrays on most chromosomes (Jung et al., 2004 and Figure 1-4). SRS proteins are important in the initial host cell attachment (Robinson et al., 2004) and it has been hypothesized that the variability of the SRS superfamily render the invasion of a wide range of cell types possible (Lekutis et al., 2001; Jung et al., 2004). Furthermore, these genes are believed to be highly important for immune evasion and persistence in the host (Kim et al., 2007).

Other important protein families in *T. gondii* originate from three kinds of secretory organelles: the club-shaped rhoptries, the small rod-shaped micronemes and the electron-dense, globular dense granules (Figure 1-3). The protein families associated with these organelles are called ROPs (rhoptry proteins), RONs (rhoptry neck proteins), MICs (microneme proteins) and GRAs (dense granule proteins), from the common nomenclature proposed by Sibley, Pfefferkorn and Boothroyd (Sibley et al., 1991). Surface and secretory proteins are essential in the invasion process, which is further discussed in section 1.1.4. Several members of these protein families have attracted a lot of interest in recent years, since they appear to be crucial for the extreme differences in mouse-virulence seen among different strains of *T. gondii*.

Rhoptry proteins (both ROPs and RONs) are specific for apicomplexan parasites, and many have enzymatic activity (Dubremetz, 2007). They are released during invasion and several ROPs migrate to the parasitophorous vacuole membrane or into the host cell cytosol and even nucleus, where they interfere with the host cell gene expression (Boothroyd and Dubremetz, 2007). ROP18 is a highly polymorphic serine-threonine kinase, which causes phosphorylation of an unknown parasite protein (Sinai, 2007), and has been shown to be a major determinant of mouse-virulence through bioinformatic analysis and transfection of recombinant proteins, (Saeij et al., 2006; Taylor et al., 2006). Other ROP-proteins that have been implicated in virulence include ROP2, ROP5 and ROP16 (El Hajj et al., 2006; Saeij et al., 2006; Bradley and Sibley, 2007) and further structural characterization of these proteins and the mechanisms behind their role in pathogenesis is in progress (Qiu et al., 2009).

Several of the GRA proteins, derived from the dense granule organelles, are highly polymorphic and variation in the GRA6 gene sequence has been used extensively for differentiation between genotypes at the DNA level (Fazaeli et al., 2000; Su et al., 2006; Zia-Ali et al., 2007), while peptides from GRA3, 6 and 7 have been used in serotyping assays (Kong et al., 2003; Sousa et al., 2008). The GRA proteins are released post-invasion and have been shown to interact with a number of host proteins, and while the exact function of each one remains to be elucidated, they appear to be important for the persistence of infection and inhibition of host cell apoptosis (Ahn et al., 2006) and nutrient acquisition (Coppens et al., 2006). In addition, the dense granule isozymes NTPase I and II are essential in the purine salvage pathway (Sibley et al., 1994; Carruthers, 1999).

MIC proteins are key molecules in the invasion process by recognition of sialic acids on the host cell surface, which establishes the first attachment and initiates the penetration (Blumenschein et al., 2007; Hager and Carruthers, 2008). Knock-out experiments have also shown that some MIC proteins, including MIC1, 2 and 3, are major determinants of virulence (Huynh and Carruthers, 2006; Ismael et al., 2006). The distribution of SRS, ROP, GRA and MIC genes across the 14 chromosomes of the haploid *T. gondii* genome is shown in Figure 1-4.

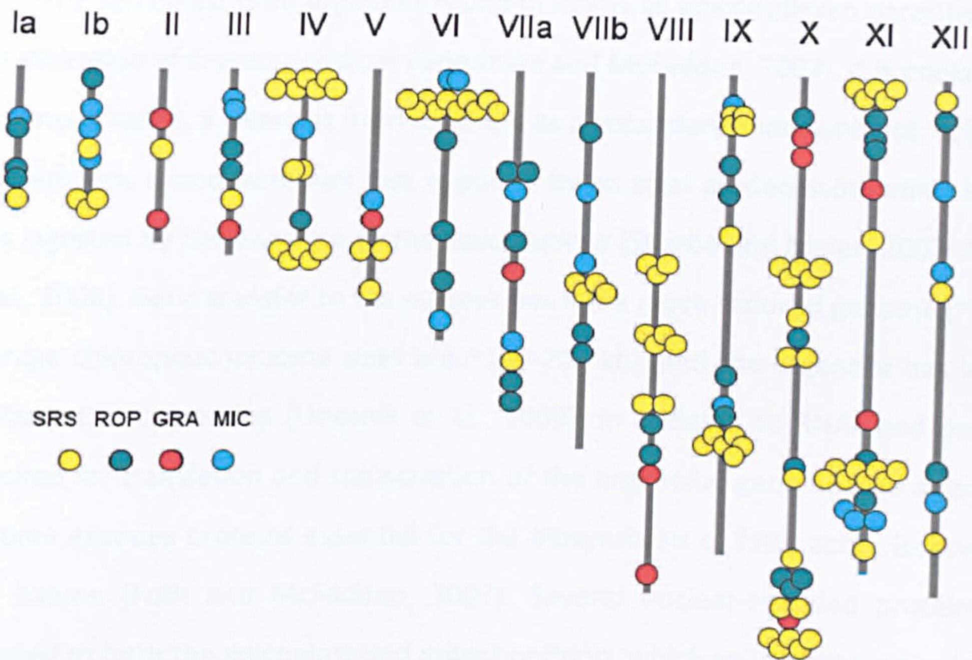


Figure 1-4. Distribution of major protein families over the 14 chromosomes (Ia-XII). No special clustering is seen except for the SRS proteins, which are often found in tandem arrays and frequently towards the subtelomeric regions. This picture summarizes localization searches for genes with keywords "SAG or SRS", "ROP or RON", "GRA" and "MIC" in the *Toxoplasma* genome database (ToxoDB v.4.3) and does not include genes with high sequence homology that have not been assigned to these groups by the administrators.

T. gondii contains two endosymbiontic organelles, a mitochondrion and a multimembrane plastid, the apicoplast (see Figure 1-5). The single mitochondrion is large and elongated, enveloped by double membranes with abundant cristae (Mather and Vaidya, 2008). As in other *Apicomplexa*, the mitochondrial genome has been reduced to ~6 kb, but *T. gondii* is unusual since it contains numerous copies of mitochondrial derived sequences in the chromosomal DNA (Ossorio et al., 1991). Due to the difficulties introduced by these gene multiplications, the *T. gondii* mitochondrial genome is still rather poorly characterized, and the sequence is not yet available in the *Toxoplasma* genome database (ToxoDB). However, similar to the situation in *Plasmodium* and *Theileria*, the reduced mitochondrial genome of *Toxoplasma* is believed to encode three proteins (cytochrome c oxidase I and III, and apocytochrome b), which participate in the electron transport chain, plus a small number of unique rRNAs (Feagin, 2000).

The apicoplast is an organelle found in nearly all apicomplexan parasites, with the exception of *Cryptosporidium* (Goodman and McFadden, 2007). It is enclosed by four membranes, a remnant from its origin as a secondary endosymbiont; originally a prokaryotic cyanobacterium was engulfed by an algal predecessor, which in turn was ingested by the ancestor of the *Apicomplexa* (Stoebe and Maier, 2002; Moore et al., 2008). Gene transfer to the nucleus has left a much reduced genome (~35 kb, average chloroplast genome sizes are ~150-200 kb), and the organelle has lost its photosynthetic function (Obornik et al., 2009). In addition to RNAs and proteins required for translation and transcription of the organellar genome, the apicoplast genome encodes proteins essential for the biosynthesis of fatty acids, isoprenoids and haeme (Foth and McFadden, 2003). Several nuclear-encoded proteins are targeted to both the apicoplast and mitochondrion, which collaborate in the haeme biosynthesis and possibly additional metabolic pathways (Obornik and Green, 2005; Pino et al., 2007) and these organelles are located in close proximity to each other (Figure 1-5).

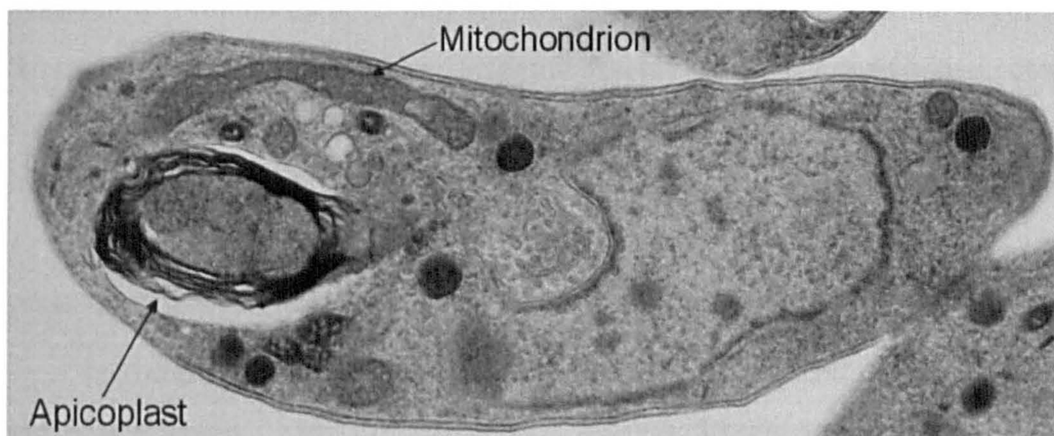


Figure 1-5. Morphology of the endysymbiotic organelles. The elongated single mitochondrion is close to the multi-membrane enclosed apicoplast, since they share at least one metabolic pathway. TEM picture of strain Me49, by Irene Bontell.

Since its recent discovery, just over a decade ago (McFadden et al., 1996; Kohler et al., 1997), the apicoplast has emerged as an important drug target since it possesses metabolic pathways of prokaryotic origin. Not only could the presence of this organelle explain the action mechanisms of known parasiticidal drugs like clindamycin and spiromycin, but it has also led to development of new drugs

currently in the clinical trial phase (Waller and McFadden, 2005; Wiesner et al., 2008).

Due to the straight-forward propagation and manipulation of *T. gondii*, a vast amount of biological data has been generated and this is increasingly being made available at the *Toxoplasma* genome database ToxoDB at www.toxodb.org (Gajria et al., 2008). Currently this database contains three annotated genomes, each spanning approximately 61 Mb divided on 14 chromosomes. In addition other data types have been added, including transcriptomic, proteomic, epigenomic data and metabolic pathway reconstructions, providing an excellent platform for hypothesis testing on the system biology level (Kim and Weiss, 2008).

1.1.4 Invasion, stage conversion and egress

Invasion of host cells is an active process that is largely conserved throughout the phylum, beginning with the parasite moving across the cell surface using a gliding motility without visible conformational changes (Heintzelman, 2006). The movement is driven by actin and myosin structures in the glideosome, a complex located between the outer plasma membrane and the inner membrane complex (Opitz and Soldati, 2002), and motility has been shown to be essential for active invasion (Soldati-Favre, 2008). Entry into the host cell is normally completed within 15-20 s and can be divided into several steps mediated by different groups of proteins, which are outlined in Figure 1-6 (Carruthers and Boothroyd, 2007).

Apicomplexan parasites have a strong polarity in their cell structure and both move and invade with the apical end leading. During invasion the parasite reorientates upon attachment so that the apex is opposed the host cell membrane (Sasono and Smith, 1998). SAG proteins are involved in the early attachment stage (Grimwood and Smith, 1996) where they, and subsequently the MIC proteins, anchor the parasite to sialylated oligosaccharides on the host cell membrane (Besteiro et al., 2009). Thereafter the moving junction (MJ) is formed by the micronemal protein Apical Membrane Antigen 1 (AMA1) and RONs 2, 4, 5 and 8 (Besteiro et al., 2009).

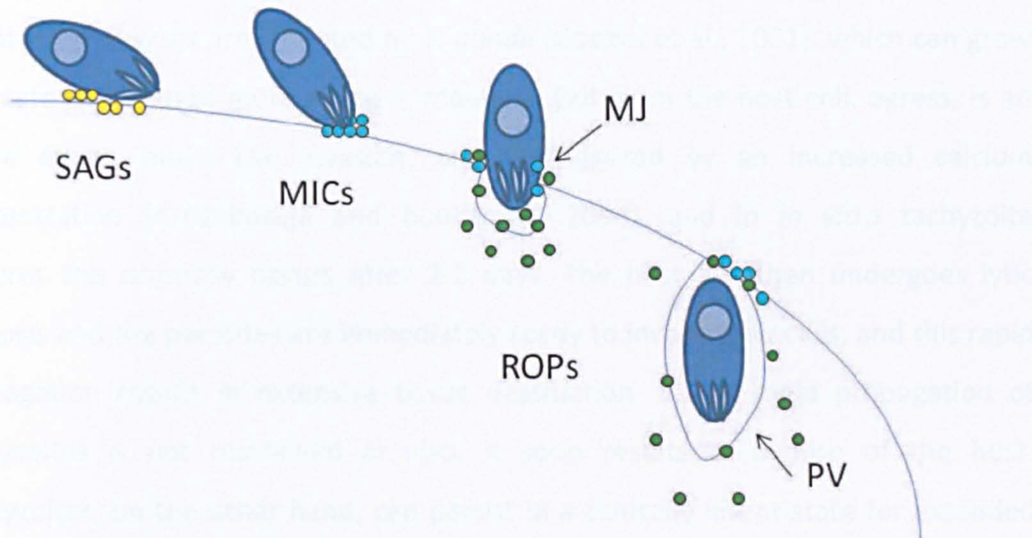


Figure 1-6. Overview of the invasion process and mediating proteins. Invasion consists of several sequential steps mediated by different groups of proteins: initial reversible attachment to surface receptors (SAGs), apical attachment (MICs), formation of the moving junction (MJ) by rhoptry neck proteins and the microneme protein AMA1, rhoptry discharge (ROPs) and invasion when the parasite apical end creates an invagination in the host cell membrane and squeezes into a confined vacuole (PV). The schematic is adapted from Carruthers and Boothroyd, 2007.

The moving junction is a conserved feature of invasion in the *Apicomplexa*, a ring formation through which the parasite squeezes into the cell. It serves as support for the parasite during invasion and efficiently excludes transmembrane host cell surface proteins from the parasitophorous vacuole (PV) (Mordue et al., 1999). Although the PV membrane mainly consists of host cell membrane, it is different from phagosomal vacuoles and does not acidify or fuse with lysosomes, (Sibley et al., 1985; Black and Boothroyd, 2000). On the contrary, it provides an optimal compartment for parasite growth and soon after formation, it associates with host mitochondria and endoplasmic reticulum to facilitate the salvage of nutrients (Magno et al., 2005).

Within the PV, the tachyzoites replicate by endodyogeny, with a replication time of approximately 6-10 hours. The division process begins with the formation of two apical membranes in the middle of the mother cell, followed by organellar division between the two daughter cells, which mature inside the original cell (Radke et al., 2001). The division within a tachyzoite vacuole is normally synchronous, resulting in 2^n parasites, but occasionally the formation of three or

more daughters within a mother cell has been observed (Hu et al., 2002). Host cell apoptotic pathways are inhibited by *T. gondii* (Goebel et al., 2001), which can grow in a safe haven until more space is required. Exit from the host cell, egress, is an active event, much like invasion, and is triggered by an increased calcium concentration (Arrizabalaga and Boothroyd, 2004), and in *in vitro* tachyzoite cultures this normally occurs after 2-3 days. The host cell then undergoes lytic necrosis and the parasites are immediately ready to invade new cells, and this rapid propagation results in extensive tissue destruction. If this rapid propagation of tachyzoites is not restrained *in vivo*, it soon results in demise of the host. Bradyzoites, on the other hand, can persist in a clinically latent state for extended periods, thereby increasing the probability of successful transmission.

In vitro, stage conversions from tachyzoites to bradyzoites can be induced by stresses like heat shock, chemical induction or high or low pH (Soete et al., 1994). *In vivo*, pro-inflammatory cytokines, most notably IFN-γ, are important inducers of stage conversion and for keeping the infection under control (Yano et al., 2002). However, spontaneous stage conversion can occur and recent evidence suggests that the cellular environment in brain and muscle may trigger bradyzoite development (Ferreira da Silva et al., 2008). In mice, bradyzoites start to appear around six days post-infection (Sahm et al., 1998) and, although tissue cysts are most commonly found in muscle and neural tissue, they can persist in a wide range of organs including heart, lung, liver, spleen, lymph nodes and the small intestine (Yano et al., 2002). Young tissue cysts can contain as few as two bradyzoites and have a diameter of only 5 µm, while a mature cyst may contain thousands of parasites and reach a diameter of 50-70 µm or more (Weiss and Kim, 2000 and Figure 1-7).

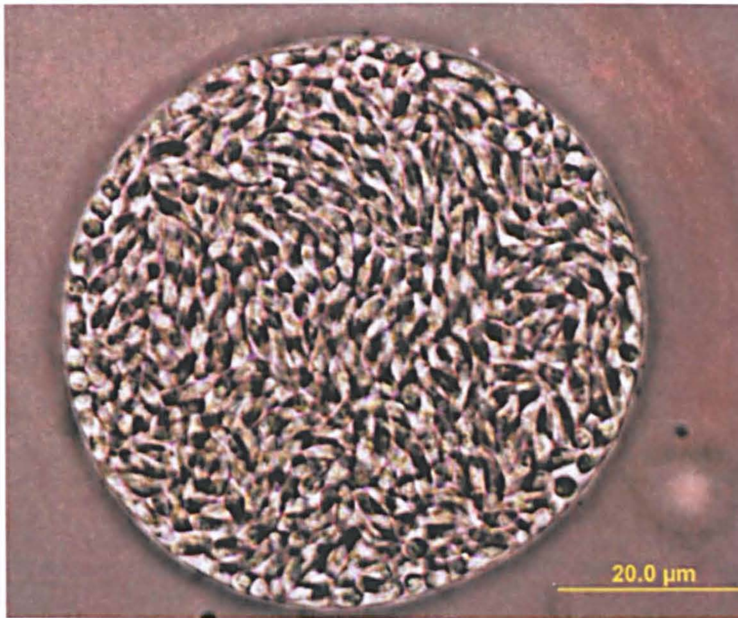


Figure 1-7. *T. gondii* bradyzoites in tissue cyst. Picture of a mature tissue cyst, containing hundreds of bradyzoites. The cyst was isolated from the brain of a laboratory mouse at JP Dubey's lab and the parasite strain originated from a chicken from Costa Rica (Picture: Irene Bontell).

Bradyzoites multiply in a slow and asynchronous fashion, using both endodyogeny and endopolygeny (formation of > 2 daughters within the mother cell), and if a tissue cyst is disrupted, bradyzoites can reinvade neighbouring cells without intermediate transformation to the tachyzoite stage (Dzierszinski et al., 2004), thus efficiently evading the specific immune response directed against tachyzoite surface antigens (Kim and Boothroyd, 2005). Reactivation of chronic toxoplasmosis, when bradyzoites revert to the tachyzoite stage, can rapidly lead to local tissue destruction as well as dispersal of parasites throughout the body. Reactivated toxoplasmosis is a problem in hosts with a weakened immune defence, including transplant patients and people with the acquired immunodeficiency syndrome (AIDS), and the mechanisms behind this stage conversion is further discussed in section 1.2.3.

1.2 The disease, toxoplasmosis

T. gondii is highly prevalent in humans, domestic animals, wild birds and mammals virtually all over the world (Tenter et al., 2000), but clinically severe disease is relatively rare. Two main types of pathogenicity are seen in *T. gondii* infections: tissue damage due to an overwhelming immune response (in the acute phase) or unimpeded parasite growth in the brain due to a deficient cellular immune response (reactivation of chronic infection). The outcome of infection is largely dependent on the host immune defence and is highly variable between different host species. Australian marsupials and lemurs from Madagascar evolved in areas without felines, only encountering *T. gondii* in recent times and these animals are highly susceptible to acute toxoplasmosis (Dubey et al., 1988; Spencer et al., 2004). In mice, the genotype of the parasite strain is a major determinant of infection severity (see 1.3.1), while rats are largely resistant and develop disease only after high doses of inoculation (Dubey and Frenkel, 1998). Among domestic animals, sheep are the most sensitive with a high level of abortions due to congenital toxoplasmosis (Morley et al., 2008), while goats, pigs and cattle are more resistant to disease, but nevertheless important for transmission (Dubey and Beattie, 1988). Humans have an intermediate sensitivity to *T. gondii* (Innes, 1997); the large majority of infections are controlled by the immune system, but clinical symptoms of various severity can occur under certain circumstances, as described in this chapter.

1.2.1 Toxoplasmosis in immunocompetent adults

Acquired *T. gondii* infections are generally asymptomatic or cause a mild, non-specific disease that does not require medical treatment, but around 10-20% report with enlarged or swollen lymph nodes, while infrequent symptoms include myocarditis, pneumonia, hepatitis and encephalitis (Montoya and Liesenfeld, 2004), but most infections go unnoticed into the persistent bradyzoite stage.

In recent years, several studies have shown that presumed innocuous chronic *T. gondii* infections may alter host behaviour, either through histopathological

changes or secretion of neuromodulators (Webster, 2007). Experiments in mice have shown that *T. gondii* infection specifically eliminated their natural fear of cat odour (Vyas et al., 2007), and correlations between *T. gondii* infection and mental disorders including schizophrenia, has been reported in human populations (Flegr, 2007; Mortensen et al., 2007). Although the precise mechanisms remain to be elucidated, the presence of two aromatic amino acid hydroxylases in the genome of *T. gondii* and the recent evidence of l-dopa production (Gaskell et al., 2009) suggests that dopamine biosynthesis could be one way through which *T. gondii* can manipulate its host (Henriquez et al., 2009).

1.2.2 Congenital toxoplasmosis

Congenital infection is the major complication of *T. gondii* infection in Europe and the United States, with incidence rates of up to around 10 cases per year in a population of one million (Jones et al., 2003; Havelaar et al., 2007; Benard et al., 2008). Women who are exposed to *T. gondii* for the first time during pregnancy may transmit the infection to the foetus, which due to its immature immune response may suffer from severe sequelae including ocular and/or mental impairment, hydrocephalus, epilepsy or intrauterine death (Jones et al., 2003). The risk of transplacental transmission is higher towards the end of the pregnancy, but the risk of severe sequelae is highest during the first trimester (Dunn et al., 1999), and there is also a risk that infected children who are asymptomatic at birth develop symptoms such as chorioretinitis later in life (Frenkel, 1974; Delair et al., 2008). Following the initial exposure, protective immunity is developed and subsequent pregnancies are normally not affected (Foulon et al., 2000). Screening of all pregnant women is performed in France, Germany, Italy and Denmark (Benard et al., 2008) and combinatory treatment consisting of pyrimethamine, sulfadiazine and folinic acid can significantly reduce the severity of symptoms in infected babies (Foulon et al., 1999).

1.2.3 Toxoplasmosis in immunocompromised hosts

Individuals with deficient cell-mediated immunity are highly susceptible to *T. gondii* and may develop acute toxoplasmosis immediately after initial infection, or as a result of reactivation of latent parasites in tissue cysts. This patient group includes transplant recipients and people taking immunosuppressive drugs, as well as those suffering from immune deficiency diseases. The largest group consists of HIV-infected individuals that are further discussed in section 2.1.3. Toxoplasmic encephalitis (TE) is the most common manifestation, but reactivation may also occur in eyes, heart, lungs or other organs (Rodgers and Harris, 1996; Eza et al., 2006).

In reactivated toxoplasmosis, conversion from bradyzoites to tachyzoites occurs several weeks before symptoms appear (Reiter-Owona et al., 2000) and therefore little is known about the direct trigger mechanism. However, a recent study (Takashima et al., 2008), detected tachyzoites inside tissue cysts in brains of mice with recent reactivation, thus showing that stage conversion followed by active egress and transmission is one starting point of reactivation, but accidental cyst rupture followed by stage conversion is another possibility (Frenkel and Escajadillo, 1987; Luft and Remington, 1992). The ability of the immune defence to prevent dissemination through killing of extracellular parasites depends on the proinflammatory cytokine IFN-γ (Suzuki et al., 1989) and depletion of IFN-γ producing CD4+ and CD8+ T-cells in chronically infected mice leads to severe encephalitis and death (Gazzinelli et al., 1992). IFN-γ induces several different parasite control mechanisms, including NO production (Gazzinelli et al., 1993), tryptophan starvation (Pfefferkorn, 1984), iron deprivation (Dimier and Bout, 1998) and production of reactive oxygen intermediates (Murray et al., 1985). Cerebral or disseminated toxoplasmosis in immunocompromised patients cause severe disease symptoms including mental confusion, blindness, aphasia, paralysis and death, but with the right diagnosis and treatment parasite replication and tissue damage can be diminished. For more on diagnosis and treatment, see 2.1.4.

1.2.4 Ocular toxoplasmosis

T. gondii has a predilection for the central nervous system (CNS) including the eyes, where it can cause acute inflammation and lasting retinal scars that may lead to impaired vision or blindness. Ocular toxoplasmosis can be a result of acquired *T. gondii* infection or a late manifestation of congenital toxoplasmosis, and the disease is highly prevalent in some regions of Brazil, where it affects up to 17% of the population (Glasner et al., 1992). The reason for the high prevalence of symptomatic disease in that region may be due to oocyst transmitted infections from contaminated drinking water (Bahia-Oliveira et al., 2003), but it has also been shown that strains from this area have unusual genotypes (Khan et al., 2006b).

1.2.5 Outbreaks of clinical toxoplasmosis

A number of waterborne toxoplasmosis outbreaks have been reported and the most well-documented took place in British Columbia, Canada, where infection was traced to a drinking water reservoir and the symptoms included fevers, headache, swollen lymph nodes and a high level of retinal lesions (Bowie et al., 1997). The reservoir was later found to be frequented by cats and cougars (Aramini et al., 1999), but no isolates were retrieved and the genotype of the strain causing the outbreak remains unknown. Other outbreaks have occurred in Atlanta, where a cat had shed oocysts in a riding stable (Teutsch et al., 1979), in American soldiers drinking river water in Panama (Sulzer et al., 1986) and more recently at a boarding school in Turkey (Doganci et al., 2006) and ocular toxoplasmosis related to a municipal water supply in India (Palanisamy et al., 2006). The overt clinical symptoms in otherwise healthy individuals during these outbreaks may be due to a high dose of ingested oocysts, since this stage is known to be more virulent in intermediate hosts compared with bradyzoites (Dubey, 2006).

In addition, there are several reports of severe and fatal toxoplasmosis in immunocompetent humans in French Guiana, who suffer from high fever, chorioretinitis, respiratory distress and multiorgan failure (Darde et al., 1998; Carme et al., 2002; Bossi and Bricaire, 2004; Demar et al., 2007). The source of infection is often unknown, but a number of these cases have been linked to the consumption

of wild game from the Amazon region (Carme et al., 2002). Genetic characterization of strains from this area has shown an abundance of highly atypical genotypes, and it is likely that the *T. gondii* strain type is an important determinant of the infection outcome in otherwise healthy people in French Guiana (Ajzenberg et al., 2004).

1.3 Diversity of *Toxoplasma gondii* strains

1.3.1 Phenotypic and genetic differences between strains

Different *T. gondii* strains display significant variation in their growth rate *in vivo* and *in vitro*, and have long been classified as 'virulent' or 'avirulent' based on their pathogenicity in mice, as determined by the LD50 (the dose resulting in 50% of the mice dying). Interestingly, nearly all isolated strains can readily be classified as virulent (LD50 ≈ 1) or avirulent (LD50 > 1000), and intermediate phenotypes are rarely seen, although they can arise through experimental crosses (Grigg et al., 2001a; Su et al., 2002). In the 1990's, genetic analysis through enzyme digestion and PCR-based methods demonstrated that all virulent strains of *T. gondii* constitute a single clonal lineage (Sibley and Boothroyd, 1992a), while the avirulent strains comprise two genetic groups (Howe and Sibley, 1995).

The three clonal lineages are called I, II and III; type I strains are uniformly lethal, while types II and III show a dose- and stage dependent disease spectrum and normally form a chronic, subclinical infection in mice (Sibley et al., 2002). The mechanisms behind the increased virulence of type I strains are not fully understood, but they have been shown to have a higher growth rate (Kaufman et al., 1959) and enhanced migratory ability (Barragan and Sibley, 2003) compared with avirulent strains. Mordue et al (2001) showed that a high inoculum (10^5 parasites) of a type II-strain resulted in similar tissue burdens (≈10^6 parasites/g) as a low inoculum (10^2 parasites) of a type I-strain. These infections resulted in comparable levels of the cytokines IFN-γ and IL-18, which caused lethal liver and lymphoid damage. After a low inoculum (10^2 parasites) of the same type II-strain, the tissue burden peaked after 8-10 days, but the immune system managed to control the infection and transform it to a latent state with moderate, non-damaging cytokine levels (Mordue et al., 2001), which implies that rapid proliferation is an important virulence factor.

Although type I strains and are less prone to form tissue cysts in mice and in cell culture (Smith, 1995), they have been shown to do so in rats (Lecomte et al., 1992), and the virulence of different strains is not an absolute quantity, but

depends on the host. No definitive correlations have been established between the genotype of the infecting strain and clinical severity in humans, since other factors such as immune status, co-infections and infectious dose are important as well. However, type I strains have been found to be overrepresented in cases of congenital toxoplasmosis in Spain (Fuentes et al., 2001) as well as among American HIV-patients (Khan et al., 2005a), and atypical genotypes which fall outside this classification appear to be more prevalent among patients with ocular toxoplasmosis (Grigg et al., 2001b).

1.3.2 Genotyping markers

Studies of the genetic diversity of *T. gondii* were first performed using a panel of enzymes that gave polymorphic digestion patterns of tachyzoite proteins (Darde et al., 1988). Twelve groups, named zymodemes Z1-Z12, were identified over the years, where Z1, Z2+Z4 and Z3 were later found to correspond to genotype I, II and III respectively, while Z5-Z12 were found in a small number of isolates, mostly from French Guiana (Darde et al., 1992; Cristina et al., 1995). Sibley and Boothroyd showed in 1992 that all virulent *T. gondii* strains had nearly identical digestion patterns using nine PCR-RFLP markers, and a few years later Howe and Sibley (1995) identified and named the archetypal lineages I, II and III. Since then a large number of genotyping studies have been based on one or multiple PCR-RFLP markers, one of the most frequently used assays is the separate amplification and digestion of the 5' and 3' end of the SAG2-gene (Howe et al., 1997), which rapidly differentiates between the three archetypal lineages. Several other PCR-RFLP markers have also been developed (Su et al., 2006) and the method has sometimes been complemented by direct sequencing of polymorphic regions in order to increase the resolution beyond the three clonotypes (Lehmann et al., 2000; Khan et al., 2005a; Zakimi et al., 2006).

Other genotyping methods include random amplification of polymorphic DNA (RAPD) and microsatellite analysis. In RAPD, short non-specific primers generate a pattern of amplicons of variable length, which can be seen as a "genetic fingerprint" of the strain (Williams et al., 1990), and this method has been shown to readily

differentiate virulent from avirulent *T. gondii* strains and also been used for mapping the genetic relatedness between strains (Guo et al., 1997). Microsatellite sequences consist of short tandem repeats (2-6 bp) which are known to mutate rapidly due to DNA polymerase slippage, and since each marker can contain numerous alleles, microsatellite analysis provide higher resolution than single SNP typing methods. Analyses of *T. gondii* strains using 5-6 different microsatellite markers confirmed previous classification into type I, II, III or atypical, but in addition revealed significant intra-lineage variation, particularly within genotype II (Blackston et al., 2001; Ajzenberg et al., 2004).

1.3.3 Genotype distribution

The classification of strains into lineage I, II and III were principally based on isolates from humans or domestic animals in Europe and North America, where type II is the predominant lineage and atypical or recombinant strains are remarkably rare. For example, four different studies of parasites from human infections in France detected a total prevalence of 10% type I strains, 83% type II, 5% type III and only 1.4% atypical strains (Ajzenberg et al., 2002; Costa et al., 1997; Honore et al., 2000; Howe et al., 1997), also see Figure 1-8 and Appendix 1. Interestingly, the proportion of strains appears to be very different among wild animals from the same continents; type II is still the most widespread, but atypical genotypes are much more prevalent. Notably, new clonal lineages have been detected in Californian sea otters and these may represent an adaptation to that specific environmental niche, where transmission appears to involve accumulation in filter-feeding invertebrates (Miller et al., 2004). The profile of strains is highly divergent in South America, where typing with conventional PCR-RFLP markers usually reveal a mixture of type I and type III alleles, while type II strains are completely absent from most areas (Dubey et al., 2003a). Figure 1-8 summarizes the distribution of different *T. gondii* genotypes, as it was known at the beginning of this work (for references and details on the studies, see Appendix 1).

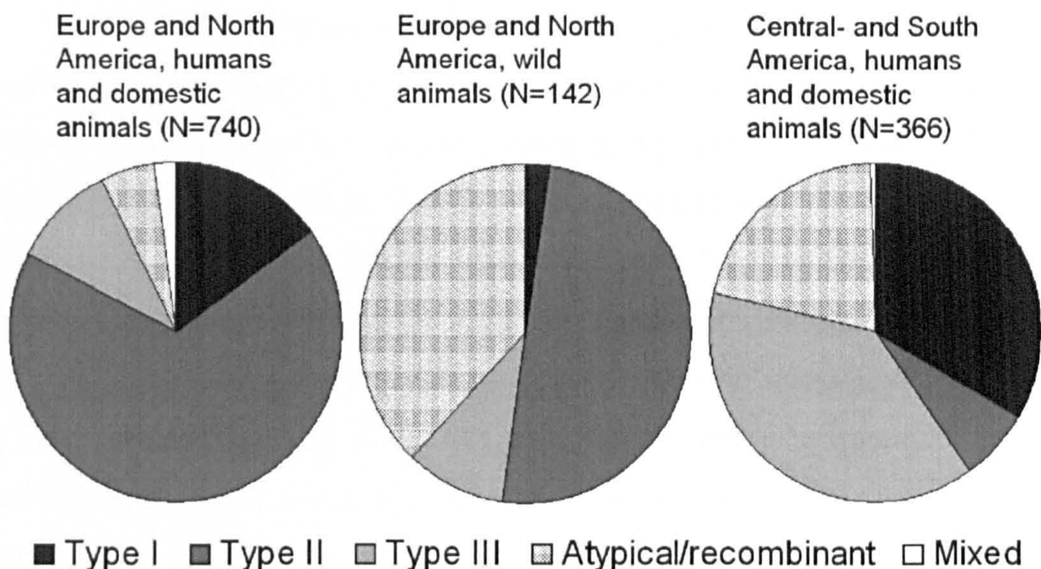


Figure 1-8. *T. gondii* genotype distribution in different hosts and geographic regions. The prevalence of different genotypes differs between humans and animals from Europe/North America compared with Central/South America. Atypical/recombinant indicates strains that do not fall in one of the three archetypal lineages, while mixed means infection with more than one genotype. These graphs summarize published studies up until early 2006, which were largely based on single-locus SAG2-genotyping. For details and references, see Appendix 1.

Extended sampling and the use of multiple genotyping markers have recently revealed a multitude of new genotypes in South America, and many strains previously classified as type I or III based on SAG2 typing were shown to contain a mixture of type I and III alleles and have been reclassified (Dubey et al., 2008). Furthermore, it has been shown that some of these "atypical" strains are actually clonal expansions, for example four clonal lineages called BrI-IV are highly prevalent in Brazil (Pena et al., 2008). Data from other regions has been practically non-existent until a few years ago and is still missing from Oceania, Central Asia and most of Africa. In Asia, multi-locus typing have recently revealed clonal type I in Korea (Quan et al., 2008) and types II and III in Iran (Zia-Ali et al., 2007), while clonal expansions of novel genotypes have been found in Sri Lanka, Vietnam and China (Dubey et al., 2007a; Dubey et al., 2007b; Dubey et al., 2007d). The large amount of new data has recently led to a better understanding of the diversity and population structure of *T. gondii* and advanced theories for the timing and events that led to the rise of the three archetypal lineages. This is discussed further in the introduction to Chapter 4.

At the beginning of the work that led to this thesis, only one report had been published on *T. gondii* genotypes in Africa. That study used SAG2-typing and showed the presence of type II (14%) and III (86%) strains in 20 chickens and one duck from Egypt (Dubey et al., 2003b). In 2005 another paper showed that all three lineages (6% I, 24% II and 70% III) were present in 17 chickens from Congo, Mali, Burkina Faso and Kenya (Dubey et al., 2005b), and these have recently been shown by multilocus typing to be highly similar to the archetypal lineages (Velmurugan et al., 2008). In addition, four strains had been genotyped from symptomatic cases in individuals of mostly unspecified African origin and in all these cases, atypical strains were detected (Ajzenberg et al., 2004; Khan et al., 2005a). No study had investigated the *T. gondii* genotype in relation to disease in African HIV-patients.

1.4 Aims

Toxoplasma gondii is an opportunistic pathogen in AIDS-patients, causing severe encephalitis with fatal outcome. Most HIV-infections occur in sub-Saharan Africa, but very little is known about the importance of AIDS-associated toxoplasmosis in Africa or the *T. gondii* strains present there. The general aim of this thesis was to characterize *T. gondii* strains from Uganda genotypically and phenotypically, which is important in relation to AIDS and also for a better understanding of the global population structure.

More specifically, the first aim was to assess the importance of *T. gondii* as an opportunistic pathogen in Ugandan HIV-patients and to determine the genotypes of disease-causing parasites. The next aim was to explore the abundance of different strain types generally present in this region, using material from asymptomatic animals. A key objective was to retrieve Ugandan *T. gondii* isolates and this was achieved from chickens, with eight different strains finally recovered. The isolation success enabled phenotypic experiments where the aim was to compare the *in vivo* and *in vitro* growth in relation to genotype. *In vitro* culturing made detailed genetic analysis feasible, and the identification of an unusual recombinant strain led to the whole genome sequencing of a type II/III strain. In this last study the aims were to elucidate the genetic composition of this strain, to identify novel mutations and genes under selective pressure, and to assess the level of divergence compared with the type II and III reference strains.

Chapter 2 Infection with *Toxoplasma gondii* is highly prevalent in Ugandan HIV-patients and all three lineages cause disease

2.1 Introduction

T. gondii has emerged as one of the most common and detrimental opportunistic pathogens in HIV/AIDS patients. Normally kept in a quiescent state by the cell-mediated immune defence, parasites may reactivate and cause extensive tissue damage and death when the immune system is compromised. Based on the seroprevalence rates, around ten million people in sub-Saharan Africa are co-infected with HIV and *T. gondii* and are thus at risk of developing toxoplasmic encephalitis, unless properly diagnosed and treated. This chapter investigates the prevalence of latent and reactivated toxoplasmosis among HIV-patients in Uganda and discusses the diagnosis of this disease in a resource-poor setting.

2.1.1 HIV and opportunistic diseases in Africa

The human immunodeficiency virus (HIV) is a relatively new pathogen for humans, but closely related simian viruses are naturally present in many species of African primates (Holmes, 2001). Based on molecular evidence, the transmission to humans is believed to have happened in central Africa about 100 years ago (Worobey et al., 2008). According to the latest comprehensive report from WHO, the HIV-viruses have killed at least 25 million people over the years, and the epidemic is now causing around 2 million annual deaths worldwide and 33 million people are currently infected (UNAIDS, 2008). In the past thirty years, HIV has spread throughout the world, but the worst affected region is sub-Saharan Africa, where AIDS is the leading cause of death in adults (Egger and Boulle, 2008). Two thirds of all infected people live in sub-Saharan Africa (Table 2-1), and this is also the

region where most children are affected: 90% of HIV-infected children and over 75% of all AIDS orphans live in Africa (UNAIDS, 2008).

Table 2-1. Global overview of HIV-infections in 2007. Data summary by continent, with specifics for Uganda shown in bold text. The data was retrieved from "Report on the global AIDS epidemic" by UNAIDS/WHO (UNAIDS, 2008).

	Total number of infections	Deaths in 2007	Adult prevalence (15-49 y)	Infected children (0-14 y)	Orphans (0-17 y)
Global	33,000,000	2,000,000	0.8%	2,000,000	15,000,000
Sub-Saharan Africa	22,000,000	1,500,000	5.0%	1,800,000	11,600,000
Uganda	**940,000**	**77,000**	**5.4%**	**130,000**	**1,200,000**
East Asia	740,000	40,000	0.1%	7,800	NA
Oceania	74,000	1,000	0.4%	1,100	NA
South and South-East Asia	4,200,000	340,000	0.3%	140,000	NA
East Europe and Central Asia	1,500,000	58,000	0.8%	12,000	NA
Western and Central Europe	730,000	8,000	0.3%	1,300	NA
North Africa and Middle East	380,000	27,000	0.3%	26,000	NA
North America	1,200,000	23,000	0.6%	4,400	NA
Caribbean	230,000	14,000	1.1%	11,000	NA
Latin America	1,700,000	63,000	0.5%	44,000	NA

The research effort on HIV/AIDS has been immense. Since the HIV virus was identified in 1983 (Barre-Sinoussi et al., 1983), over 100,000 scientific articles with major topic "HIV" or "AIDS" have been published, but no definitive cure or vaccine has yet been developed. Treatment and immunization attempts are greatly aggravated by the extraordinary genetic variety of HIV, even within a single patient (Saag et al., 1988), and over 200 known mutations are associated with drug resistance (Shafer and Schapiro, 2008). HIV is a retrovirus, consisting of a single-stranded RNA sequence of approximately 9200 nt. Two distinct viruses have been discovered: the more aggressive HIV-1, which originates from chimpanzees (Gao et al., 1999) and HIV-2, which has a prolonged pre-patent period and is mainly present in West Africa, where its original host is the sooty mangabey (Clavel, 1987). Many different genotypes of HIV-1 have been identified and while subtype B is the most common in Western Europe and USA, much more variation is seen in Africa

(Hemelaar et al., 2006). Subtype C is responsible for most cases in Southern Africa, but in Uganda, subtype D is the most common, followed by subtype A and a small proportion of subtype C (Nakasujja et al., 2005; Hemelaar et al., 2006). The importance of the HIV clades for clinical outcome is largely unknown, but recent studies comparing disease progression in African patients infected with subtypes A, C and D show that infections with recombinant subtypes and clade D have a faster progression to severe disease and death (Vasan et al., 2006; Sacktor et al., 2007; Kiwanuka et al., 2008).

The primary host cell target for the HIV-virus are CD4+ T-lymphocytes (CD4-cells), and after invasion double-stranded DNA is synthesized and integrated into the host DNA as a provirus. A clinically latent period follows, when the virus slowly but continuously replicates and adapts to the host immune system and builds up an ineradicable reservoir (Chun and Fauci, 1999). During the prepatent period the number of CD4-cells gradually declines and the patient develops AIDS, which is defined as "all HIV-infected persons who have less than 200 CD4+ T-lymphocytes/μl of blood, or a CD4+ T-lymphocyte percentage of total lymphocytes of less than 14%" (CDC, 1993). In addition, there are a number of AIDS-defining diseases. Some of these, like AIDS dementia, are caused directly by the HIV-virus and are poorly understood, but most are caused by opportunistic pathogens. The risk of opportunistic illnesses increases as the CD4-count (number of CD4-cells per μl blood) declines and the relative importance of different pathogens depend on their prevalence in each region. Tuberculosis is the biggest threat to HIV-patients in many parts of the world (Havlir et al., 2008), but a wide range of bacterial, viral, fungal and protozoal infections can cause severe disease and deaths in immunosuppressed individuals (Table 2-2). In Uganda, tuberculosis, cryptococcal meningitis, toxoplasmic encephalitis and Kaposi's sarcoma are among the main problems associated with HIV infection, but other common reasons for seeking medical care include chronic diarrhoea, bacterial meningitis, pneumonia, oral candidiasis and herpetic ulcers (Nakasujja et al., 2005).

Table 2-2. Overview of common opportunistic pathogens of HIV-patients. (Data from Jung and Paauw, 1998; Holmes et al., 2003; Davaro and Thirumalai, 2007).

Bacterial infections	**Main disease location(s)**	**CD4-count[a]**
Mycobacterium tuberculosis	Pulmonary, neurological	<500
Salmonella, Shigella, Campylobacter spp.	Gastrointestinal	<500
Streptococcus, Haemophilus, Staphylococcus, Legionella spp.	Pulmonary, sepsis	Variable
Mycobacterium avium complex	Multi-organ	<100
Treponema pallidum (syphilis)	Neurological	<350
Viral infections		
Herpes viruses	Kaposi's sarcoma, skin, retina	Variable
Cytomegalovirus	Neurological, retina	Variable
Hepatitis C virus	Liver	Variable
Epstein-Barr virus	Lymphomas	Variable
Protozoal infections		
Toxoplasma gondii	Neurological, retina	<200
Cryptosporidium spp.	Gastrointestinal	<200
Isospora belli	Gastrointestinal	<200
Microsporidia spp.	Multi-organ	<200
Fungal infections		
Pneumocystis jirovecii	Pulmonary	<250
Aspergillus spp.	Pulmonary, multi-organ	<50
Cryptococcus neoformans	Neurological	<200
Histoplasma capsulatum	Multi-organ	<50
Candida albicans	Esophagus, oral	<100

[a]Indicates at which stage of immunodeficiency HIV-patients are considered to be at risk of disease with the respective pathogens.

2.1.2 Epidemiology and transmission of *T. gondii*

T. gondii is the most common protozoan infection in humans world-wide, and since it is mostly post-natally acquired the seroprevalence increases with age (Frenkel, 1973). The prevalence may vary depending on food habits, water supply, climate and the presence of felines, and high seroprevalence rates of *T. gondii* infections have been reported from diverse population groups such as the Inuit in

northern Canada (59.8%) (Messier et al., 2008), strict vegetarians in India (47%) (Hall et al., 1999), and pregnant Parisian women (67.3%) (Jeannel et al., 1988). In Europe and USA, meat consumption is believed to be the main transmission route, and countries such as France and Serbia with a tradition of eating raw or undercooked meat have very high seroprevalence rates (Baril et al., 1999; Bobic et al., 2007). In other places, including the United Arab Emirates (Abu-Zeid, 2002), India (Yasodhara et al., 2004) and Mexico (Alvarado-Esquivel et al., 2006), seroprevalence was linked to residential area and socioeconomic status, and risk factors included poor housing and close contact with animals. Several studies from Brazil, where ocular toxoplasmosis is common, have shown that the seroprevalence rate is associated with socioeconomic status and drinking unfiltered water (Bahia-Oliveira et al., 2003; de Amorim Garcia et al., 2004; de Moura et al., 2006). In addition, many of these studies reported an increasing seroprevalence with age, and a large population based study from the Netherlands, including > 7000 samples, found an increase from 1% in the > 1-year old to 79% in the 75-79 year age group, with the steepest increase between 25 and 44 years of age (Koortbeek et al., 2004).

In Africa, the prevalence of *T. gondii*-infection is highly variable, both between and within countries (Figure 2-1 and Appendix 2), with reported rates ranging from 4% in a Tanzanian Muslim village (Gille et al., 1992) to 83.5% among pregnant women in Antananarivo, Madagascar (Lelong et al., 1995). Apart from the previously mentioned life-style factors, climatic conditions appear to be important, with a general trend of lower human seroprevalence in dry compared with humid areas, which probably reflects the reduced oocysts survival time in arid climates. For example, a study from the Ivory Coast shows a prevalence of 55-70% in humid, tropical areas and only 36% in the pre-desert zone (Dumas et al., 1989), and the corresponding figures in Senegal were 58% and 27% (Dumas et al., 1990b). The compilation of *T. gondii* seroprevalence studies from Africa shown in Figure 2-1 also supports a climatic effect, where the humid regions in West Africa and Madagascar have the highest prevalence, while drier areas in the Northern and Southern parts have a lower rate of infection.

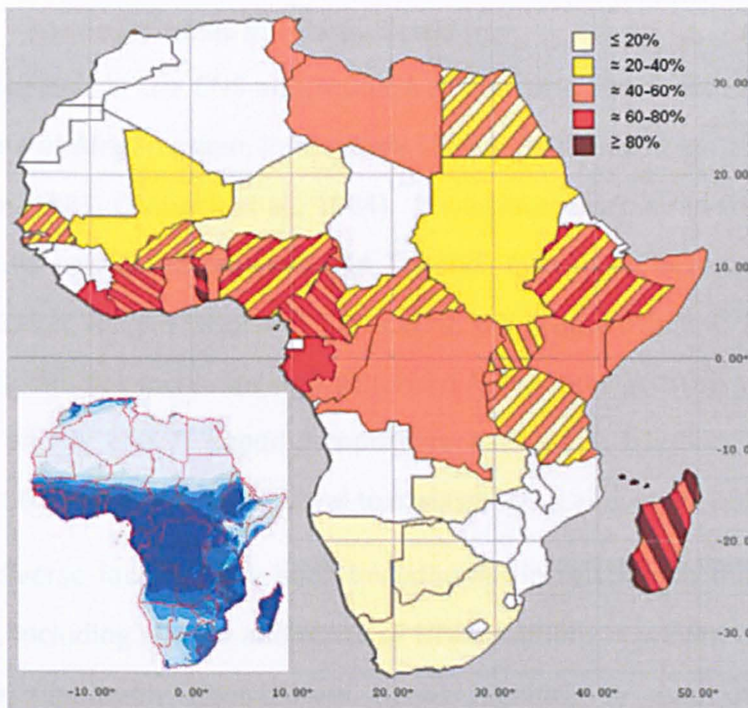


Figure 2-1. *T. gondii* seroprevalence in Africa. Compilation of seroprevalence studies on African adults between 1975 and 2006 (countries for which no data was available are uncoloured, references in Appendix 2). The insert shows the annual rainfall with darker blue indicating a higher amount of precipitation (UNEP, 2009). The highest *T. gondii* prevalence was found in West Africa and Madagascar, which have a humid climate, while a lower prevalence was found in most of Northern and Southern Africa.

The relative contribution of the different transmission routes may vary between different communities; oocysts transmission has been suggested as the major source of *T. gondii* infection in Somalia (Ahmed et al., 1988), Congo (Candolfi et al., 1993) and Gabon (Duong et al., 1992), while meat was found to be more significant in Sudan (Abdel-Hameed, 1991). Both routes were considered important in Nigeria (Uneke et al., 2005) and the Ivory Coast (Adou-Bryn et al., 2004). Since there are no direct methods to determine the route of acquisition, these studies are based on analyses of the prevalence in different age groups and/or questionnaires. One seroprevalence study had been executed in Uganda before the start of the work in this thesis, which showed a seroprevalence of 27% in the general population and 34% among HIV-patients, but no analysis of transmission or risk factor was performed (Zumla et al., 1991).

2.1.3 Toxoplasmosis in HIV-patients

Toxoplasmosis in the CNS and retina was first associated with AIDS patients, including some of African origin, in the early 1980's (Pitchenik et al., 1983; Schuman and Friedman, 1983; Clumeck et al., 1984). It was later discovered that around 25-50% of patients coinfected with HIV and *T. gondii* eventually develop TE (Luft and Remington, 1992), usually after the CD4-count has dropped below 200 (Jung and Paauw, 1998). The risk increases with advanced AIDS, and a Scottish study revealed that 48% of all HIV and *T. gondii* seropositive patients in Edinburgh with a CD4-count below 50 had developed cerebral toxoplasmosis (Laing et al., 1996).

Many diverse factors have been investigated in relation to the risk of TE in HIV-patients, including gender and route of HIV-infection. However, only a few risk factors have repeatedly been found to be significantly associated with the development of TE: low CD4-count, positive serology and absence of *T. gondii* prophylaxis and anti-retroviral treatment (Dunlop et al., 1996; Meisheri et al., 1997; Nascimento et al., 2001; Antinori et al., 2004; Nissapatorn et al., 2004). The clear association with positive *T. gondii* serology strongly suggest that reactivation of latent infection is a major cause of disease. For example, in the Norwegian study the seroprevalence in 698 HIV-patients was 17.8%, but among the 19 patients who developed TE, 18 (95%) had positive serology (Dunlop et al., 1996). However, in places where the yearly acquisition rate is high recent infections may be more important, and superinfection with a different parasite strain has been shown to cause disease in immunosuppressed patients (Ghosn et al., 2003). The risk of reactivation may also increase through additional impairment of the immune function due to concurrent infections. A case-control study from Ethiopia showed that helminth infections, which induce a very strong anti-inflammatory response, facilitate the reactivation of latent tuberculosis (Elias et al., 2006), and the same may be true for toxoplasmosis, which like tuberculosis require a pro-inflammatory response to be kept in the latent state.

Since the introduction of highly active antiretroviral therapy (HAART) in 1996, the mortality among AIDS-patients has decreased dramatically. As treatment became available, the mortality caused by opportunistic infections was reduced

more than 5-fold between 1994-97 in a US cohort consisting of patients starting with a CD4-count below 100 (Palella et al., 1998), and today carefully optimized antiretroviral treatment enables patients to lead normal lives (Rosen-Zvi et al., 2008). With the correct treatment, immune reconstitution including recovery of *T. gondii* specific responses, can be achieved in most AIDS-patients and if the CD4-count is kept > 200 for three months the patients may be considered for discontinuation of *T. gondii* prophylaxis (Furco et al., 2008). In the western world, toxoplasmosis and other opportunistic pathogens are no longer major causes of CNS related disease in HIV-patients, but the problem remains where HAART is not available (Smith et al., 2008).

2.1.4 Diagnosis and treatment

Differential diagnosis of HIV-associated opportunistic diseases can be very difficult, and the diagnosis of TE in Uganda relies mainly on clinical signs plus neuroradiological findings for patients who can afford computer tomography (CT). TE clinical symptoms include headache, convulsions, paresis and mental confusion, which is a similar spectrum as that seen for other opportunistic infections in the CNS, such as cerebral tuberculosis and neurosyphilis, rendering differential diagnosis difficult (Modi et al., 2004). CT-scans of TE abscesses show large, necrotizing lesions, which are often surrounded by several smaller foci (Offiah and Turnbull, 2006) and suspected TE can be verified by response to specific treatment, but this method requires at least two weeks follow-up (Dupon et al., 1995; Raffi et al., 1999)

A number of serological methods have been developed for specific detection of anti-*T. gondii* IgG antibodies, for example the Sabin-Feldman dye test (Sabin and Feldman, 1948), agglutination tests (Fulton and Turk, 1959; Mitchell and Green, 1960) and enzyme-linked immunosorbent assays (Walls et al., 1977). IgG-antibodies arise about 2 weeks post-infection and normally persist at detectable levels throughout life, even though the production may be reduced in immunodeficient individuals. Since IgG-detection assays cannot distinguish between latent and active infections, serial measurement of IgG and IgM-antibody levels is used for diagnosis

of active toxoplasmosis in immunocompetent people (Montoya, 2002; Lappalainen and Hedman, 2004). However, IgM is often undetectable in HIV-patients and direct detection of live parasites or of *T. gondii*-DNA outside tissue cysts is required for diagnosis in immunocompromised hosts (Bastien, 2002; Joseph et al., 2002). The gold standard for detection of TE is histological examination of brain biopsies (Dupouy-Camet et al., 1993), but this is an invasive procedure which poses a health risk for the patient. Direct detection can also be achieved through isolation of the parasite in mice or cell culture, but these methods are expensive and time-consuming. PCR detection of parasite DNA is rapid, inexpensive and has the possibility of high throughput. Cerebrospinal fluid or peripheral blood can be used for detection of active toxoplasmosis, and one of the most common targets is the 35-fold tandem repeat "B1" (Burg et al., 1989).

A range of anti-parasitic agents is available, which inhibit different stages of tachyzoite growth through interference with the folate metabolism (Fung and Kirschenbaum, 1996) or through inhibition of the purine salvage pathway (Gherardi and Sarciron, 2007; Hyde, 2007). However, no regimen is able to eradicate the intracellular bradyzoites and in order to prevent reactivation medication must continue as long as the CD4-count stays below 200 (Miro et al., 2006). In Uganda, Septrin (co-trimoxazole, trimethoprim + sulfamethoxazole) and Fansidar (pyrimethamin + sulfadoxin) are used for treatment of toxoplasmosis in HIV-patients, and these drugs are readily available as they are also used to treat malaria. Both drugs exploit the inability of the parasite to salvage folic acid from the environment by inhibition of the folate synthesis pathway, and the diet should therefore be complemented with extra folic acid. Treatment of TE usually results in a markedly improved clinical condition within 14 days (Amogne et al., 2006), but unfortunately these drugs can cause adverse effects such as rash, anaemia and leukopenia (Fung and Kirschenbaum, 1996).

2.1.5 Aims

The aims of this study were to determine *T. gondii* seroprevalence and rate of active toxoplasmosis in Ugandan HIV-patients, and to find out which parasite genotypes were present. Furthermore, we wanted to investigate the relationship between the laboratory results (serology and PCR-detection) and the clinical phenotype (level of immunosuppression and disease symptoms) in order to assess the current situation of toxoplasmosis in Ugandan HIV-patients and identify indicators of reactivation that can facilitate diagnosis.

2.2 Materials and Methods

2.2.1 Study site

Uganda is a relatively small, landlocked state in east Africa, traversed by the equator and with an elevation range between 621 and 5,110 m. The rapidly growing population of about 31 million people consist of around 20 ethnic groups, which can be divided into the Nilotic people of the North and the Bantu people of the South. Half of the population are children up to 15 years of age, around 35% live below the poverty line, infant mortality is 6.6% and the life expectancy is around 52 years (CIA, 2006). In the capital, Kampala, 67% have access to clean water but only 7% to sewerage and in many other places these figures are even lower (National Water and Sewerage Corporation, 2006). The HIV rate is estimated to be around 5.4% in the 15-49 age group, but somewhat higher in urban areas (Nakasujja et al., 2005; UNAIDS, 2008) and other common infections include tuberculosis, malaria and a range of helminth infections (Kolaczinski et al., 2007).

2.2.2 Study populations and sample collection

Patients were recruited from Mulago Hospital, which is the largest government referral hospital in Uganda and the main teaching hospital of Makerere University. Mulago Hospital is situated close to central Kampala and most patients derive from a lower socioeconomic urban population. Inclusion criteria for this study were positive HIV-serology (or previously documented infection) and willingness to participate after being informed about the purpose of the study. The samples were collected on different occasions from two different wards, and have therefore been divided into five groups, G1 – G5 (Table 2-3).

The first samples (G1) were collected in March-April 2005 from patients with known HIV-infection who presented at Mulago with neurological symptoms (N=130). For the next sampling round (G2) in Jan-Feb 2006, this category of patients was not accessible, instead blood was collected from patients who had recently

LEEDS UNIVERSITY LIBRARY

tested positive for HIV (N=130). G3 consisted of patients from that same HIV-testing clinic in February 2007 (N=45). Blood samples and clinical data from patient groups 4 and 5 (N=254 and 116 respectively) were collected by Dr Erima Bernard during 2006, who also performed the *T. gondii* IgG agglutination test for G4 patients. The G4 samples, like G1, originated from symptomatic patients ("Sympt"), whereas G5, like G2 and G3, consisted of patients from the HIV-testing clinic ("HIV-test"). In total, 675 patients were included in the study. Their average age was 33.5 years and the overall male to female ratio was 1:1.51.

Peripheral blood samples (0.5-5 ml) were obtained with sterile, disposable needles by qualified health workers and collected in EDTA-tubes. Physicians at Mulago Hospital performed medical examinations, while laboratory personnel at the hospital carried out the HIV-ELISA and CD4-count assays. The current study formed part of the medical evaluation of these patients, but was not the sole objective for taking blood specimens, why it was not always possible to obtain full clinical records and CD4-counts. However, in order to obtain a good sample size for the laboratory investigation, all HIV-positive patients were included, but naturally only those with the appropriate records were taken into account for each test. Therefore, the numbers of patients vary between tests, and the exact group size for every test is stated for all results separately.

Table 2-3. Overview of the five patient groups and the corresponding data sets.

Group	N	Collection[a]	Patients[b]	Sex (M:F)	Mean age ±SD	Clinical Data[c]	CD4-count[c]
G1	130	2005	Sympt	1:1.06	36.4±8.8	Partial	No
G2	130	2006	HIV-test	1:1.98	31.8±9.3	No	Partial
G3	45	2007	HIV-test	1:1.87	32.5±6.7	Yes	No
G4	254	EB, 2006	Sympt	1:1.33	34.2±9.3	Yes	Yes
G5	116	EB, 2006	HIV-test	1:1.76	32.9±10.1	Yes	Yes

[a]Samples from patient groups G1-G3 were collected on three occasions, from 2005 to 2007 (Irene Bontell), G4 and G5 were collected by Dr Erima Bernard in 2006.
[b]Samples were obtained from one ward where patient sought care because of neurological signs ("Sympt"), or from a HIV-testing clinic ("HIV-test").
[c]Unfortunately, clinical data and CD4-counts were only available for a subset of patients

2.2.3 *T. gondii* serology (Toxo-Screen DA)

Blood samples were incubated at 4°C overnight in order to separate the blood components, and the following day serum and buffy coat were separated and stored at -20°C until used. Toxo-Screen DA (BioMérieux) is a commercial test for detection of *T. gondii* IgG in serum through direct agglutination with formalin-treated *T. gondii* parasites of the RH-strain. Non-specific agglutination is suppressed by the addition of 2-mercapthoethanol, which denatures IgM antibodies by breaking the disulfide bonds. All reagents are supplied with the kit; antigen, diluents (coloured albumin buffer), 2-mercaptoethanol, positive and negative controls (goat serum from one animal with known titre and one uninfected animal), powdered PBS and round-bottom 96-well plates with self-adhesive sheets. Serum samples including controls were diluted in PBS to 1:20, and 25 µl serum dilution, 25 µl 0.2M 2-mercaptoethanol and 50 µl antigen solution were dispensed in each well. The plates were covered with adhesive sheets and incubated at room temperature for 6-24 hours. Plates were read by visual examination and interpreted according to the manufacturer's instructions: positive if agglutination of antibody and antigen had formed a mat covering about half of the well base and negative if antigen had sedimented at the bottom (Figure 2-2).

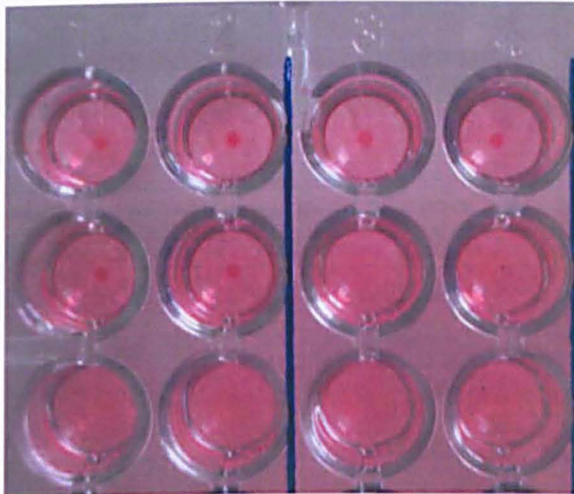


Figure 2-2. *T. gondii* direct agglutination test (Toxo-Screen DA). The uppermost row and first two wells on the second row had negative reactions, where the antigen had sedimented at the bottom; the other wells had positive reactions, where antibodies and antigen have formed a sprawling complex.

Serial dilutions were not performed since the value of this has not been established in HIV-patients, and because of cost and limited availability of reagents. However, samples tested in duplicate at the same and different time points always showed the same result and the controls were always valid. The test result is therefore a binomial yes or no, rather than a serological titre. Indecisive results were very rarely seen, and could always be resolved by a repeated testing.

2.2.4 DNA-extraction

DNA was extracted using the QIAamp DNA-kit (Qiagen) according to the instructions for the Blood Mini Spin-protocol. Briefly, 200 µl of buffy coat was mixed with 20 µl Proteinase K and 200 µl lysis buffer and incubated at 56°C for 20-40 minutes, until the cells had undergone lysis. Thereafter, 200 µl of pure ethanol was added; the mixtures were vortexed and transferred to spin columns and centrifuged for 1 minute in a benchtop microfuge at 8,000 rpm. Following two washing steps the DNA was eluted for 5 minutes using 200 µl elution buffer and the resulting DNA extracts were stored at -20°C.

2.2.5 Diagnostic and genotyping PCRs

A nested PCR-protocol was used for detection of *T. gondii*-DNA, based on the original method from 1989 targeting the 35-fold repetitive B1-gene (Burg et al., 1989). Basic genotyping, which classifies strains into one of the three archetypal lineages (I, II or III), was performed through PCR-RFLP on both ends of the SAG2-gene (Howe et al., 1997). Total reaction volume was originally 50 µl with 20 µl DNA template for the G1-samples, but this was later reduced to 20 µl with 5 µl template DNA. The mastermix was made of 4 µl GoTaq 5* buffer, 2 µl dNTPs (2mM), 1 µl forward + reverse primer (10 µM) and 1 µl GoTaq polymerase (Promega). For the nested PCR, 1 µl of the product from the first round was used as template. Primers and restriction enzymes are listed in Table 2-4.

Table 2-4. Primers and restriction enzymes used for detection and genotyping of *T. gondii*.

Gene (chrom)		Forward primer[a]	Reverse primer[a]	Length	Rest. enz.[b]
B1 (IX)	ext	GGAACTGCATCCGTTCATGAG	TCTTTAAAGCGTTCGTGGTC	194	
	int	TGCATAGGTTGCAGTCACTG	GGCGACCAATCTGCGAATACACC	97	N/A
SAG2	ext, 5'	GCTACCTCGAACAGGAACAC	GCATCAACAGTCTTCGTTGC	335	
(VIII)	int, 5'	GAAATGTTTCAGGTTGCTGC	GCAAGAGCGAACTTGAACAC	242	Sau3AI
	ext, 3'	TCTGTTCTCCGAAGTGACTCC	TCAAAGCGTGCATTATCGC	327	
	int, 3'	ATTCTCATGCCTCCGCTTC	AACGTTTCACGAAGGCACAC	222	HhaI

[a]B1 primers from Burg et al, 1989 and SAG2 primers and enzymes from Howe et al, 1997.
[b]*Sau3AI* digests amplicons with the type III allele (186+56 bp), while *HhaI* digests products with the type II allele (169+53 bp).

DNA quality and absence of PCR-inhibitors were controlled by a PCR protocol generating a 121-bp product from the human β-globin gene (Primers: ACCACCAACTTCATCCACGTTCACC (F) and CTTCTGACACAACTGTGTTCACTAGC (R)). This amplification was, however, not performed for groups G4 and G5, since the amount of DNA was limited. The PCR cycling program consisted of an initial denaturation step at 95°C for 3 min, followed by 35 cycles of denaturation at 94°C (45 s), assay specific annealing temperature (30 s), and elongation at 72°C (1 min). All programs ended with a final elongation step of 72°C for 7 min. Annealing temperatures were: B1 1st round: 55°C, B1 nested: 52°C, SAG2-3' and SAG2 5' 1st: 56°C, SAG2-3' and SAG2 5' nested: 59°C. Amplified fragments were separated on 1-2% agarose gels stained with ethidium bromide, and were visualized under UV-light.

T. gondii DNA from cultured RH parasites was used as positive control and the sensitivity and specificity of the assays were tested through serial dilutions, ranging from 0.1-100 genome equivalents of parasite DNA (Figure 2-3), with DNA from human foreskin fibroblasts (HFFs) and water as negative controls. All methods were shown to reliably detect the equivalent of 10 DNA copies, while lower concentrations were often, but not always, discovered.

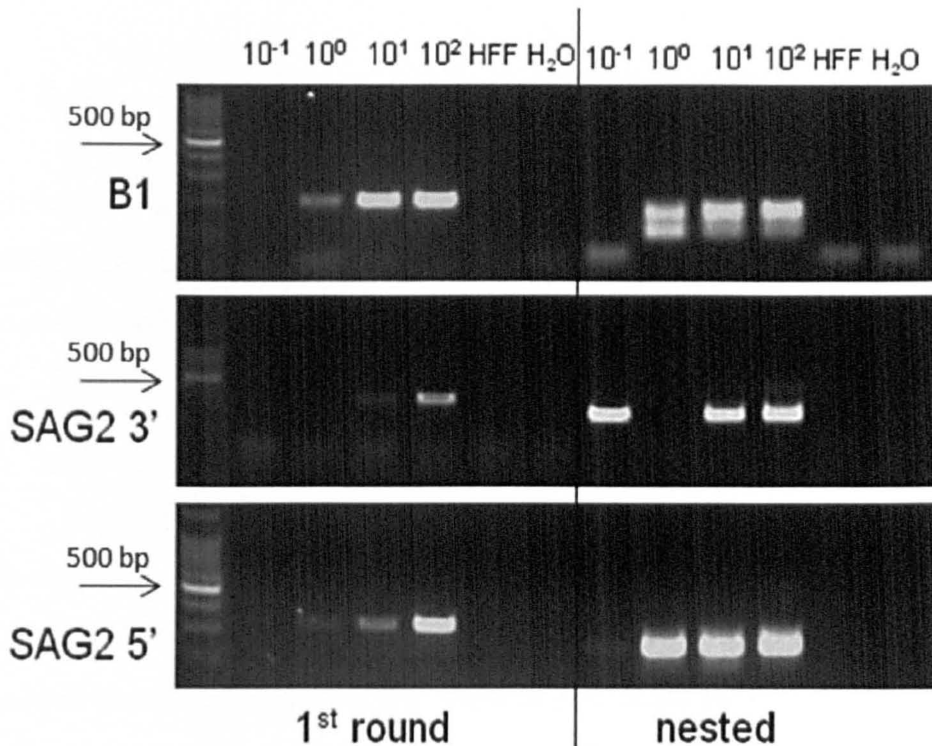


Figure 2-3. Sensitivity testing of PCR assays used for *T. gondii* detection. Parasites of the RH strain were harvested from cell culture and counted in a haematocytometer. DNA was extracted from a volume containing 10^6 parasites, and serial 10-fold dilutions were made. Extracts of uninfected HFF-cells and water were used as negative controls (short bands visible for B1 are primer-dimers). Fragment lengths are indicated by a 100 bp-marker, where the 500 bp is the brightest band.

2.2.6 Quantitative real-time PCR

Quantitative real-time PCR (Q-PCR) was used on a subset of samples in order to estimate the parasite density in blood samples from patients with reactivated disease. Primers (GCTCCTCCAGCCGTCTTG (F), TCCTCACCCTCGCCTTCAT (R)) and a TaqMan minor groove binder probe (6-FAM-AGGAGAGATATCAGGACTGTA) directed against the approximately 300-fold repetitive fragment AF146527 (Homan et al., 2000) were designed by Dr. Daniel Palm at the Swedish Institute for Infectious Disease Control (SMI) and purchased from Applied BioSystems. ROX was used as a passive reference and the program utilized consisted of 2 minutes at 50°C and 10 min 95°C, followed by 45 cycles of 95°C (15 s) and 60°C (1 min). Each sample was run in duplicate and the dilution series used to generate the standard curve was in triplicate. The standard curve consisted of a 10-fold dilution series, ranging from 0.1-1000 parasites per µl, plus three wells with no template control (H_2O), and this was included on each plate.

2.2.7 Statistical analysis

Most graphs and calculations were performed in Microsoft Excel, including the chi-square and t-tests, which were used to test for statistically significant differences between groups. The box-plot diagram was generated with the statistical software R (www.r-project.org).

2.2.8 Isolation in cell culture

Attempts were made to isolate *T. gondii* strains from seropositive HIV-patients directly into cell culture at Makerere University. For more information on standard cell culture techniques and reagents, see section 4.2.2. The same HFF cells and cell culture medium components that were used in Leeds (see 4.2.2) were brought and cultures were set up at Makerere. 200 µl buffy coat from seropositive patients was mixed with 10 ml pre-warmed (37°C) fresh DMEM medium and added to a confluent HFF monolayer. The parasites were allowed 24 hours to invade and the following day fresh medium was added. Cultures were monitored under an inverted microscope several times a week and medium was changed every 2-3 days.

All cell culture pipetting work was performed in a sterile hood at the Department of Biochemistry, the cultures were kept in a 37°C, 5% CO_2 incubator at the Department of Bacteriology and the inverted microscope was located at the Infectious Disease Institute at Mulago. The cultures were transported between these sites in a clean container with closed lid, but contamination proved difficult to avoid. The isolation attempt in cell culture probably failed because of the lack of appropriate facilities at Makerere University at the time. Space limitations in the incubator resulted in the culture flasks being shuffled around and once all ended up in a vertical position, resulting in 90% of the cells dying. While some cultures survived, yeast and bacterial contamination was a major problem, and although *T. gondii* like organisms were observed no isolates were retrieved from human blood using this method.

2.2.9 Isolation in mice

Twenty-one seropositive blood samples from group G3 were brought from Uganda to Sweden in EDTA-tubes kept in cold boxes. The buffy coat (approximately 200 µl) was mixed 50:50 with sterile PBS, and the mixture was inoculated subcutaneously into two mice per sample within 2-4 days after the drawing of blood. The mouse work was performed at the section for Experimental Biomedicine at SMI, where technical personnel performed inoculations, bleedings, daily animal care and putting the mice to death using the anaesthetic isofluorane. The animals were monitored daily for disease symptoms and bled for serology (Toxo-Screen DA, as in 2.2.3) after six weeks. Seropositive mice were sacrificed and dissected, and the brain was used for inoculation in cell culture and parasite quantification using real-time PCR (for further details on trypsin digestion, inoculation in cell culture and quantification of parasites in mouse organs, see 4.2.4 and 4.2.5).

As with direct isolation in cell-culture, mouse-isolation of *T. gondii* from human blood proved difficult. Three mice infected with blood from seropositive HIV-patients from G3, were positive by serology at a titre of 1:25. Two of these mice were inoculated with blood from the same patient (H07-29), a 26-year old female suffering from photophobia, while the last mouse was infected with blood from a 33-year old female reporting with fever and headache (H07-31). Inoculation of these mouse brain homogenates in culture did not result in any parasite growth, but after a second mouse passage (1:3) three out of six mice infected with H07-29 were positive by Q-PCR detection, but not by serology. The parasite densities were very low, however, at 110, 290 and 515 parasites per gram and the subsequent transfers to cell culture were unsuccessful. Small quantities of these brain homogenates have been cryopreserved for future reference and possibly renewed isolation attempts, but were not further used in this study.

2.3 Results

2.3.1 Patient demographics, CD4-counts and symptoms

Five different groups of samples (G1-G5) were obtained from two different wards over a period of three years (summary in Table 2-3, section 2.2.2). The patients in groups G1 and G4 ("Sympt") had a previous diagnosis of HIV and came to the hospital to seek care for neurological symptoms, while the samples in groups G2, G3 and G5 ("HIV-test") were collected from patients who came to test their HIV-status. The symptomatic patients were slightly older (on average 34.6 years) and had a more equal gender distribution (M:F, 1:1.28) than the people coming for HIV-testing, who were 32.3 years on average and predominantly female (M:F, 1:1.91). The male patients of both categories were older than the females and more than 90% were aged 45 or younger. The age distribution by sex and patient category is shown in Figure 2-4 and the full dataset used for the analysis in this chapter is provided in Appendix 3.

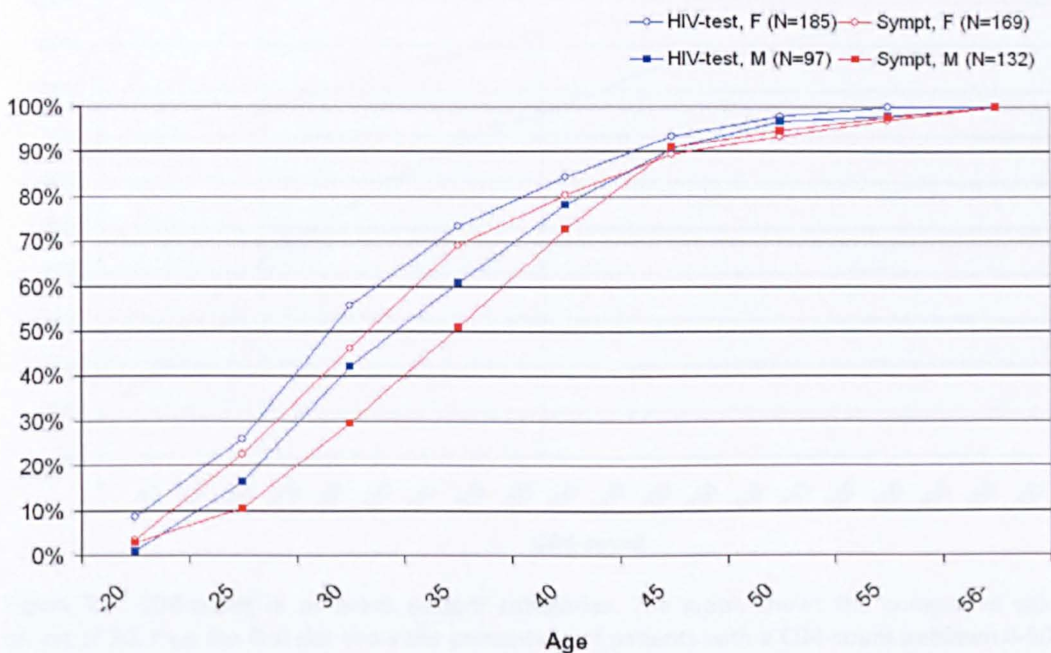


Figure 2-4. Cumulative age distribution by gender and patient category. The graph shows the proportion of patients up to a given age, in five-year intervals. The symptomatic patients were older than those who attended the clinic for a first HIV-test, and the men were on average a few years older than the female patients. The majority of all patients were between 25 and 45 years old.

The normal range of CD4-positive cells in healthy individuals is around 700-1000 CD4-cells/µl (Mair et al., 2008), while the average CD4-count from the current study was 334 for "HIV-test" patients and only 145 for the "Sympt" category. This difference between the patient groups was highly significant, as determined with the chi-square test ($P < 10^{-10}$). A cumulative plot (Figure 2-5) of the CD4-counts in these groups showed that 75% of the HIV-test group and over 90% of the symptomatic patients had CD4-counts below 500, thus being in the risk zone of several opportunistic diseases. One of the CDC definitions for AIDS is a CD4-count below 200, and this applied to 74.8% of the symptomatic patients, and 48% of the individuals in this group had severe immunodeficiency with a CD4-count below 50. The corresponding numbers for the "HIV-test" category was 42.5% below 200 and 13.4% under 50 CD4-cells /µl. No significant differences were seen between males and females, either within each group or in the cohort as a whole (P > 0.05, two-sided t-test).

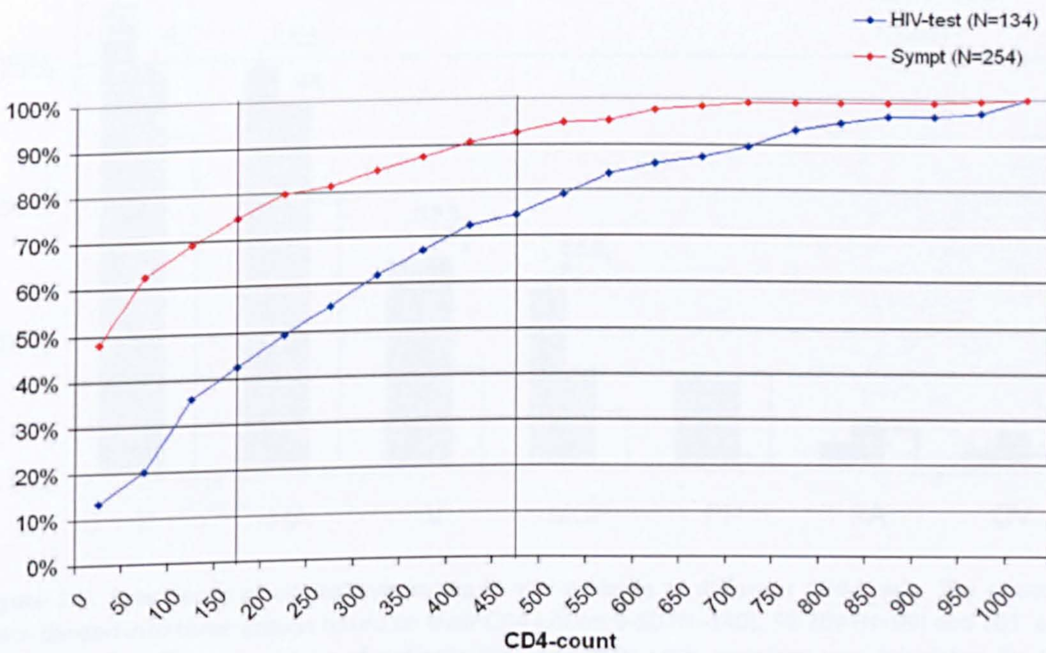


Figure 2-5. CD4-count in different patient categories. The graph shows the cumulative value in blocks of 50, thus the first dot show the percentage of patients with a CD4-count between 0-50, the next dot the percentage of patients with a CD4-count between 0-100 and so on. Nearly two thirds of all patients (247/388) had advanced immunodeficiency, with CD4-counts of 200 or less and 86% (334/388) had levels below 500 (indicated by vertical lines). The symptomatic patients had a significantly lower average level compared with the HIV-test group ($P < 10^{-10}$).

Commonly reported symptoms that were included in the analysis were fever, severe or recurrent headaches, vomiting, mental confusion, photophobia, weakness or paresis (mostly half-sided) and convulsions. Other clinical signs included rash, oral thrush, coughs, diarrhoea and neck pain. Patients with a severely compromised immune status (CD4 ≤ 50) had a considerably worse clinical picture compared to those with a moderate immunosuppression (CD4 > 200), and had significantly higher prevalence of fever, headache, vomiting and mental confusion ($P < 0.001$). The patients with intermediate CD4-levels, between 51 and 200, had an intermediate level of fever, headache and mental confusion. However, there were no significant differences in the level of photophobia, paresis or convulsions between the three groups. The percentage of patients with each of the seven symptoms and the significant differences between groups are shown in Figure 2-6.

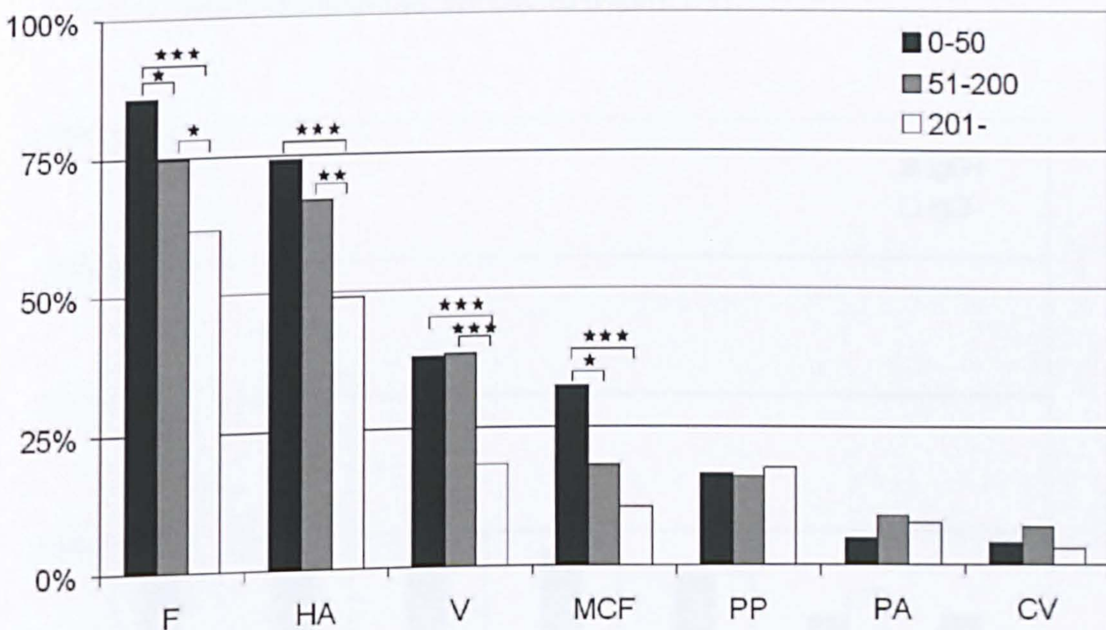


Figure 2-6. Prevalence of clinical symptoms in HIV-patients at different CD4-levels. The patients were divided into three groups based on their CD4-count: 0-50 (N=140), 51-200 (N=99) and 201 and higher (N=131). The percentage of patients suffering from each symptom was calculated for the three groups and pair wise statistical comparisons were made between all three groups using the chi-square test. Significance levels are indicated in the figure, where one star means $P < 0.05$, two stars $P < 0.01$ and three stars $P < 0.001$. (F=fever, HA=headache, V=vomiting, MCF=mental confusion, PP=photophobia, PA=paresis or weakness, CV=convulsions).

2.3.2 *T. gondii* seroprevalence

Sera from 675 HIV-positive Ugandans were tested for *T. gondii* specific IgG-antibodies using a direct agglutination test. Three hundred and eighty two patients were found to be seropositive, which gives a total infection rate of 56.6% in this population (full result list in Appendix 3). A comparison of the seroprevalence rate between patient groups with different levels of immunosuppression showed a tendency towards a higher infection rate in the patients with the lowest CD4-counts; 65.0% in the CD4 0-50 group, 58.5% in 51-200, and 55.0% in the CD4 >200 group, but these differences were not significant. There was, however, a significantly higher (P < 0.005) seroprevalence of *T. gondii* infection in the symptomatic patients in G1 and G4 (61.7%) compared with the HIV-test patients in G2, G3 and G5 (49.8%). Positive *T. gondii* serology was associated with a higher prevalence of headache (P < 0.05, chi-square), but no further significant differences were found regarding the clinical symptoms (Figure 2-7).

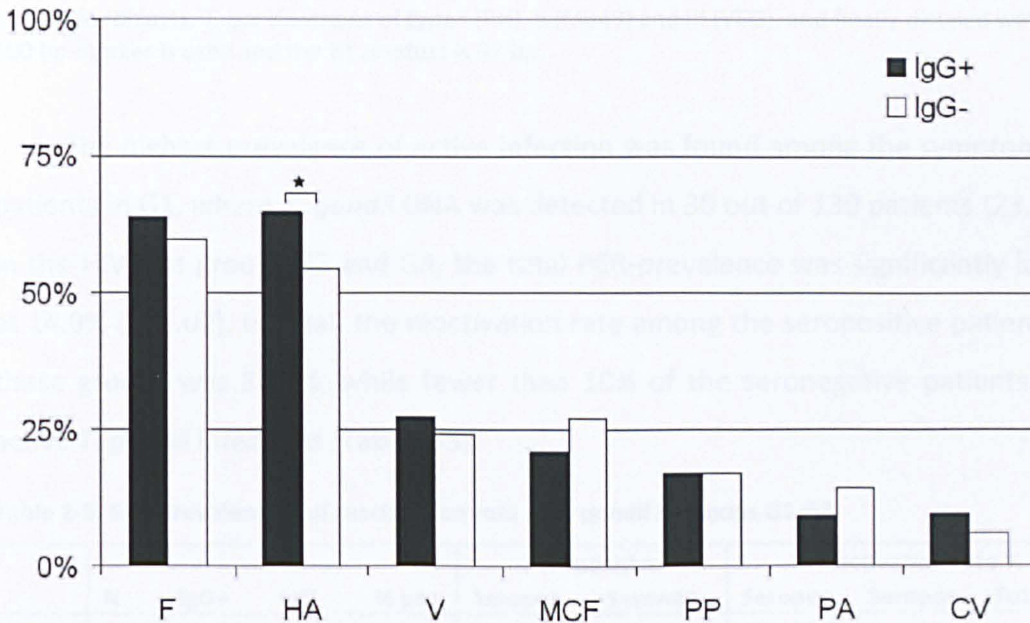


Figure 2-7. Prevalence of clinical symptoms in seropositive and seronegative patients. Patients with a positive serology ("IgG+", N=382) were more likely to suffer from severe or recurrent headaches, compared with the seronegative patients ("IgG-", N=293) at the P <0.05 level indicated by one star in the graph). Apart from that, the clinical picture was highly similar in the two groups. (F=fever, HA=headache, V=vomiting, MCF=mental confusion, PP=photophobia, PA=paresis or weakness, CV=convulsions).

2.3.3 Diagnostic PCR

Samples were first screened by nested B1-PCR, and thereafter amplified at both ends of the SAG2-gene (Figure 2-8). The B1 locus was more sensitive and many samples that had been freeze-thawed or stored for long periods were difficult to amplify with SAG2, although they were repeatedly positive for the 35-fold repeated B1 target.

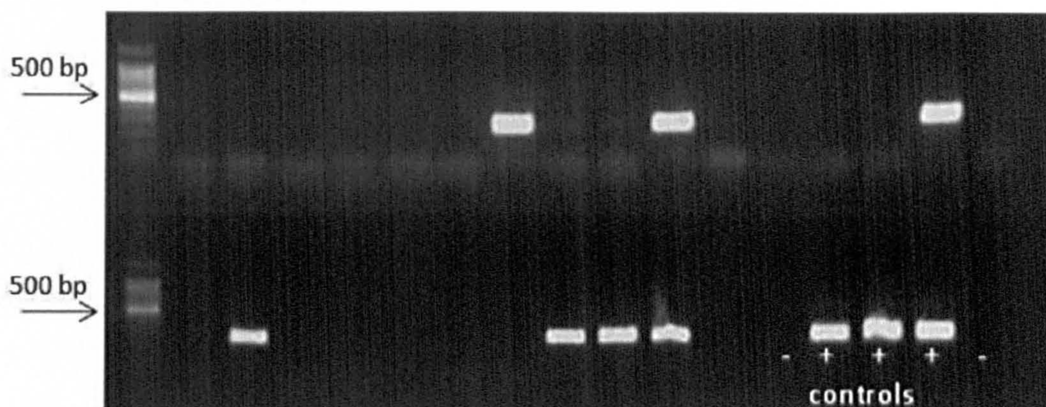


Figure 2-8. Gel picture of nested B1-PCR. This screening revealed seven positive and 20 negative samples. The final five samples on the second row constitute the controls; DNA from uninfected human fibroblasts, *T. gondii* strains of type I (RH), II (Me49) and III (VEG), and finally distilled water. A 100 bp-marker is used and the B1 product is 97 bp.

The highest prevalence of active infection was found among the symptomatic patients in G1, where *T. gondii* DNA was detected in 30 out of 130 patients (23.1%). In the HIV-test groups G2 and G3, the total PCR-prevalence was significantly lower at 14.9% ($P<0.02$). In total, the reactivation rate among the seropositive patients in these groups was 26.1%, while fewer than 10% of the seronegative patients had active *T. gondii* infections (Table 2-5).

Table 2-5. Seroprevalence and reactivation rate of *T. gondii* in groups G1-G3.

					PCR-positive		Reactivation rate %		
	N	IgG+	IgG-	% pos	Seropos	Seroneg	Seroneg	Seropos	Total
G1	130	70	60	53.8	22	8	13.3	31.4	23.1
G2	130	70	60	53.8	14	6	10.0	20.0	15.4
G3	45	21	24	46.7	6	0	0.0	28.6	13.3
Total	**305**	**161**	**144**	**52.8**	**42**	**14**	**9.7**	**26.1**	**18.4**

The number and percentage of seropositive and PCR-positive patients is shown for each group (G1-G3). The seroprevalence varied between 46.7%-53.8%. As expected most PCR-positive patients also had antibodies, but for G1 and G2 a total of 14 IgG-negative, PCR-positive patients were found.

The DNA samples from groups G4 and G5 were difficult to amplify, probably due to degradation caused by repeated freezing / thawing, which made the screening results unreliable, and only two samples were found positive by nested SAG2-PCR. Q-PCR screening of 137 samples from G4 resulted in 11 positives (8%), which had an average quantity of 0.097 (range: 0.002-0.3) copies of *T. gondii* DNA per PCR reaction tube. In total, 68 of the 675 HIV-patients investigated in this study were PCR-positive (for a complete list, see Appendix 3).

Positive serology was strongly associated with reactivation; 73.5% of the PCR-positive patients were seropositive, compared with 54.9% of the PCR-negative patients (P<0.001, chi-square test). Yet, *T. gondii* DNA was detected in the blood of 18 seronegative individuals, meaning that more than one in four Ugandan HIV-patients with active toxoplasmosis did not have detectable levels of specific IgG-antibodies. The lack of specific antibodies may be explained by the reduced capacity of antibody production in patients at an advanced stage of immunosuppression, and a closer look at the 12 PCR-positive patients for which CD4-counts were available is consistent with this assumption (Figure 2-9).

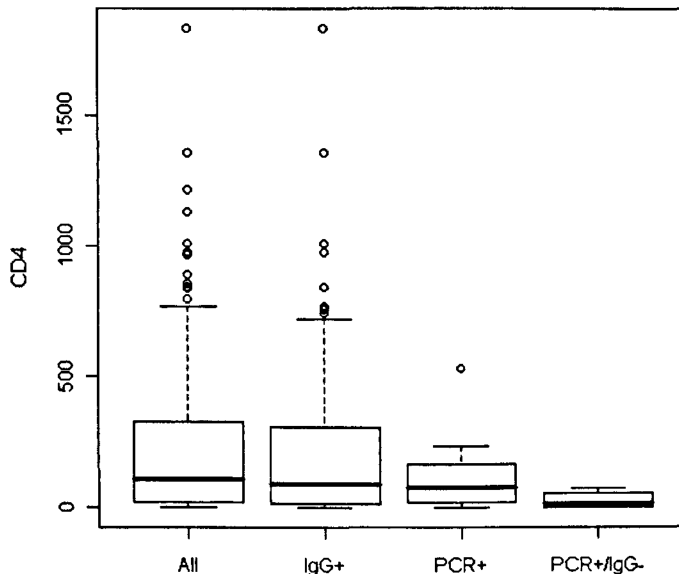


Figure 2-9 Box-plots showing the distribution of CD4-counts in different patient groups. Each box depicts the inter-quartile range with median values marked as a line. Range shows last value within 1.5 IQR, and outliers are marked by dots. 'All' includes all patients with a CD4-count (N=388, median 109.5), 'IgG+' seropositive patients (N=229 median 90), 'PCR+' all PCR-positive patients regardless of serological status (N=12 median=75), and 'PCR+/IgG-' PCR-positive patients with negative serology and known CD4-count (N=4 median 15.5).

Four PCR-positive / seronegative individuals were found in the lower end of the scale, with counts of 0, 3, 28 and 73 CD4-cells/μl, while the eight patients with positive PCR and positive serology had counts of 7, 45, 75, 94, 166, 175, 237, 336 and 532. The distribution of CD4-counts in the patients with active toxoplasmosis was compared with that of all seropositive patients and the general study population in Figure 2-9. Although there appears to be a trend towards lower CD4-count in patients with positive PCR, the differences did not reach significance.

The correlation between clinical symptoms and active toxoplasmosis was investigated through pair wise comparison of the PCR-positive and PCR-negative patients. Headache was the only symptom significantly associated with reactivation ($P<0.01$), and there was support on the $P<0.05$ level that fever was less common among patients with toxoplasmosis. However, the predictive value of these symptoms in clinical practice remains low, as headache and fever are highly prevalent in both patient groups. The proportion of patients presenting with each symptom in the respective groups are shown in Figure 2-10.

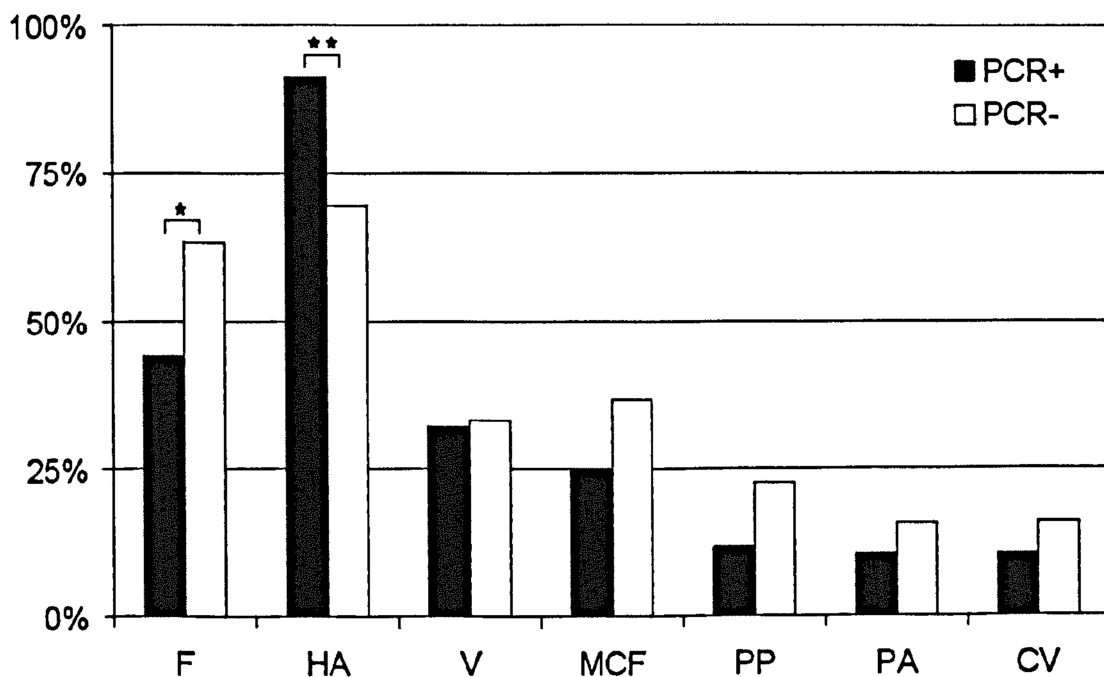


Figure 2-10. Prevalence of clinical symptoms in PCR-positive and PCR-negative patients. Headache was the most common symptom, and it was associated with reactivation, affecting 92% of the PCR-positive patients. On the contrary, fever was more common in the PCR-negative group, but was nevertheless the second most common symptom in the PCR+ group, with a prevalence of 44%.

2.3.4 Genotyping

Classification of the 30 PCR-positive samples from G1 was achieved through restriction enzyme digestion of the SAG2 products. Separate amplifications of the 5' and 3' end of the gene is necessary for accurate typing into the canonical three lineages and therefore only partial typing was possible for three samples where the 3'-end amplification was unsuccessful. SAG2-genotyping was attempted on the other patient groups but gave inconclusive results.

All three SAG2-types were found in this cohort of Ugandan HIV-patients. Type II was the most common, comprising 60% of the strains (18/30) followed by type I at 20%. Three cases of type III infections (10%) were detected and the remaining three strains were classified as being either type I or II, but not III. The sample set was too small for statistical analysis of the risk factors associated with reactivation of a particular genotype, but this data, shown in Figure 2-11, demonstrates that the three lineages are all present in Uganda and able to cause disease in HIV-patients.

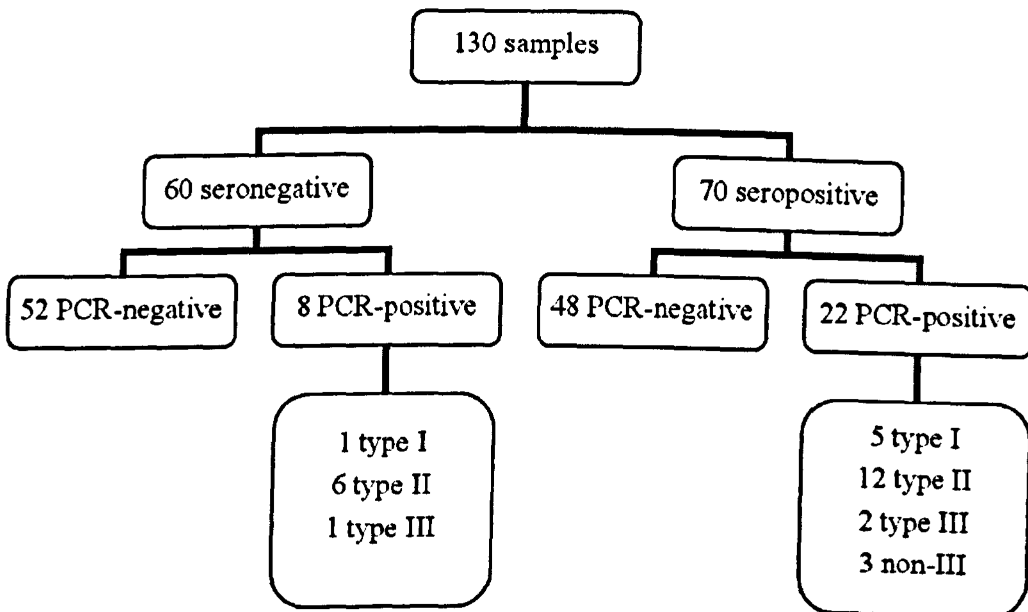


Figure 2-11. Overview of serology, PCR-results and SAG2 genotyping for G1. Genotyping showed that representatives of the three clonal lineages were present, and that type II was the most common lineage. The genotype distribution was similar among seropositive and seronegative patients.

2.4 Discussion

2.4.1 High seroprevalence and slow increase with age

The total seroprevalence rate for *T. gondii* among 675 Ugandan HIV-patients was found to be 56.6%. This is similar to the levels of exposure in neighbouring countries; 54% in Kenya (Griffin and Williams, 1983), 50% in Rwanda (Gascon et al., 1989) and 50% in Kinshasa, DRC (Dumas et al., 1990a). Two serological studies have previously been performed in Uganda: in 1966 a seroprevalence of 11.7% was reported among healthy blood donors aged 15-25 years (Ludlam and Somers, 1966), and in 1991 the prevalence in HIV-negative and symptomatic HIV-positive individuals was found to be 27% and 34% respectively, using the Sabin-Feldman dye test (Zumla et al., 1991). However, the overall prevalence in the latter HIV-positive population was 45% using a latex agglutination test, but the authors judged the dye test to be more reliable (Zumla et al., 1991). In a study from Kenya the same year, other authors found that the agglutination test was more sensitive and more specific than the dye test, and this study showed a seroprevalence of 54% in HIV-patients from Nairobi (Brindle et al., 1991). Based on the above results it appears that the *T. gondii* incidence is increasing over time in Uganda, but other factors like differences in study populations (age, place of origin, HIV-infection etc) or sensitivity of the test may also be of importance. Gender-related differences in *T. gondii* seroprevalence have been found in some populations and been traced to handling of animals and meat, either at work or in the household (Arene, 1986; Abdel-Hameed, 1991; Youssef, 1993). However, gender did not appear to be a significant factor for the prevalence of infection or reactivation in Uganda; the seroprevalence was 1.1% higher in men, but this difference was not statistically significant.

The majority of *T. gondii* infections in Uganda appear to be acquired at a relatively young age, since only a slow increase in seroprevalence with age was seen in this study population. An interesting finding was the considerably higher seroprevalence among the symptomatic patients compared with the ones who

came to the hospital for HIV-testing (Figure 2-12). Between 20 and 40 years of age this difference was in the order of 10-20 percentage units, but in the older age groups (41-45 and 46 and above) the seroprevalence among the symptomatic patients declined. Although speculative, one feasible explanation is decreased survival to this age among those who are severely immunocompromised and simultaneously infected with *T. gondii*.

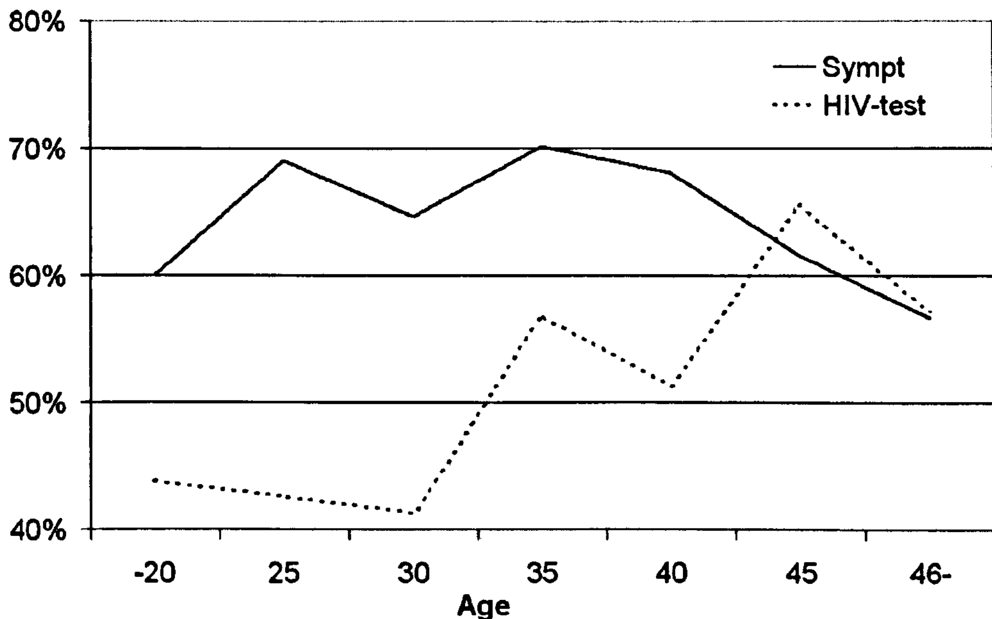


Figure 2-12. Seroprevalence in different age groups in symptomatic and HIV-test patients. An increase with age was seen in the 'HIV-test' group, but among the symptomatic patients, the prevalence declined in the older age groups. The age is shown in 5-year intervals, where 25 means 21-25, 30 means 26-30 and so on.

Taking the high rate of active toxoplasmosis in the symptomatic patients into account, this sample set is likely to be skewed towards people who are infected, which implies that their main reason for seeking care was the symptoms caused by toxoplasmosis. The same may be true also for a smaller fraction of the HIV-test patients and, hence, while we have shown that *T. gondii* is a common and important infection in HIV-patients in Uganda, the seroprevalence rate reported here cannot be assumed to faithfully mirror the infection rate in the general population.

2.4.2 Reactivation is common and occurs in seronegative patients

Sixty-eight patients, corresponding to 15.2% of the screened samples, were found to have active toxoplasmosis through PCR-detection of parasites in the peripheral blood. However, the real reactivation rate may be somewhat higher, since the DNA in G4 was scarce and degraded and only showed a reactivation rate of 8% in symptomatic patients. In groups G1-G3, the total reactivation rate was 18.4%, but among individuals with positive *T. gondii* serology it reached 26.4%.

Studies investigating HIV-associated toxoplasmosis in other parts of the world have shown similar rates of reactivation. A prospective study from France showed that 25.4% of patients with positive *T. gondii* serology and a CD4-count below 100 developed TE within 12 months (Oksenhendler et al., 1994), and a 27% reactivation rate was found among seropositive Danish AIDS-patients (Garly et al., 1997). However, during the last decade the introduction of highly active antiretroviral treatments has decreased the occurrence of toxoplasmosis and many other opportunistic diseases in AIDS-patients in Europe (Gray et al., 2003). Studies on HIV-patients from sub-Saharan Africa report a variable degree of toxoplasmosis. Several studies from West Africa consider toxoplasmosis among the most common causes of death among AIDS-patients (Lucas et al., 1993; Millogo et al., 2000), while it appears to be infrequent (<1%) in South Africa (Hari et al., 2007; Modi et al., 2007). This difference could partly be due to methodological differences; Lucas et al performed autopsies, which are likely to detect the majority of TE cases, Millogo et al looked at response to treatment, which may lead to an overestimate of the number of cases. The South African studies on the other hand used serology, clinical presentation and to some extent CT-scans, which may not be sufficient to detect all cases of active toxoplasmosis. The predominance of different subtypes of HIV in these regions might have some influence, but HIV-clade was not found to be significantly associated with disease presentation in the literature summary presented by Modi et al (2007). Consistent with the regional differences in seroprevalence shown in Figure 2-1, the *T. gondii* seroprevalence in the study population of Hari et al was only 8%, and it does appears that the prevalence of *T.*

gondii and other infectious diseases in a population is the principal factor behind the opportunistic infections seen in HIV-patients.

Two medical dissertations at Makerere University have investigated the prevalence of HIV-associated opportunistic infections of the central nervous system in Ugandan patients. D. Wamala studied 44 post-mortem samples and found that cryptococcal meningitis was the most common cause of death (33%) followed by cerebral tuberculosis (16%), bacterial infections (16%) and toxoplasmosis (13%) (Wamala, 1998). K. R. Byanyima, on the other hand found toxoplasmosis to be the most important opportunistic infection. She investigated 105 symptomatic HIV-patients, whose main clinical symptoms were headache (76.2%), fever (67.6%), focal neurological signs (63.8%), convulsions (37.1%), reduced consciousness (30.5%), vomiting (27.6%) and photophobia (12.4%). The final diagnosis was based on clinical presentation, CT-scans, laboratory tests for tuberculosis and Cryptococcus, response to therapy and in some cases post mortem examination. In the Byanyima study, toxoplasmosis was the most common disease, affecting 29 patients (27.6%), followed by cryptococcal (N=18), bacterial (N=9) and tuberculous (N=4) meningitis, and viral encephalitis (N=4) (Byanyima, 1998).

An unexpected finding in the current study was the high proportion of patients who were PCR-positive in spite of negative serology. The absence of specific antibodies has sometimes been considered as a diagnostic signal which eliminates the risk of reactivated toxoplasmosis (Raffi et al., 1999), but several other research groups have found active infections in *T. gondii*-seronegative HIV-patients (Porter and Sande, 1992; Laing et al., 1996; Joseph et al., 2002). In some cases this might be explained by recent infection, but considering the low acquisition rate in these age groups and the very advanced stage of immunosuppression among the symptomatic patients (22.4% had a CD4-count of 10 or lower), a more plausible explanation in this particular cohort is that the severe immune deficiency prohibited production of detectable antibody levels.

2.4.3 Diagnosis of latent and active toxoplasmosis in Uganda

Active *T. gondii* infection can cause a wide range of different symptoms, and the clinical outcome of reactivated infection in the brain is directly linked to the site of parasite proliferation and host cell necrosis. *In vivo* imaging results suggest that parasites use the capillary system to spread within the brain, and the preferred sites of growth are the parietal cortex and frontal lobes, where many cognitive functions are located (Dellacasa-Lindberg et al., 2007), but all types of brain cells can be infected (Suzuki, 1999) and differential diagnosis between HIV-associated neurological diseases cannot be based on clinical presentation alone. Local reactivation in the brain is followed by dissemination to other organs, including lungs and liver, and by this stage, which can last up to two weeks after the first symptoms, parasites start to become detectable in peripheral blood (Takashima et al., 2008).

Currently, there are no simple ready-to-use tests for detection of TE in AIDS-patients and expensive methods like brain scanning are not feasible in resource-limited settings. A study from Peru revealed that 82% of clinically important AIDS-related diseases remained undiagnosed until the post mortem investigation (Eza et al., 2006), which illustrates the need for reliable diagnostic tools. The IgG-agglutination test used in this study is robust and easy to perform even in less equipped laboratories, and although there is a risk of false negative results in patients with advanced immunosuppression, positive serology is useful for identifying patients at risk of reactivation. PCR, although sensitive and specific in controlled tests, is not trivial to perform on clinical samples where the amount of parasite DNA is very small in comparison with the host DNA. In this study, Q-PCR quantification showed that the parasite density in the blood of HIV-patients with reactivated toxoplasmosis was really on the limit of detection. However, rapid processing of samples without freeze-thawing, increased amount of template DNA and the use of a multi-copy target sequence increases the likelihood of detection.

In conclusion, serology is useful for identification of patients at risk of reactivation, but can be unreliable in individuals with a very low level of CD4-cells. *T. gondii* serology ought to be performed on all patients at the time of HIV-testing

and patients who are seronegative should be advised on how to avoid becoming infected, while positive patients should be advised to seek care if they get severe recurrent headaches. PCR should be used at the suspicion of TE and has a good positive predictive value, but a negative result does not rule out active toxoplasmosis since local reactivation may occur some time before parasites are detected in the blood. Even in the absence of positive serology or PCR, HIV-patients should be considered for *T. gondii* prophylaxis if their CD4-count drops below 100 and they suffer from severe headaches.

2.4.4 Three lineages are present and cause reactivation

SAG2-genotyping data was retrieved from the 30 PCR-positive samples in the first cohort and showed that all three clonal *T. gondii* genotypes were present and caused disease in Ugandan HIV-patients. The most common lineage in this study was genotype II, which was responsible for 60% of cases, followed by type I and type III, and the genotype distribution was thus reminiscent of that seen in humans and domestic animals in Europe and North America, and clearly distinct from the pattern in South America (Figure 1-8). The absence of genetic data on *T. gondii* strains generally present in Uganda, prohibited comparisons of the strain distribution in symptomatic vs. asymptomatic infections that could show whether certain genotypes are more prone to reactivate in AIDS-patients.

Single-locus typing is insufficient for detection of recombinant or atypical strains, but although attempts were made with the PCR-RFLPs described in 3.2.9, it proved difficult to obtain more genetic information from these samples. PCR-screening of blood samples is an unbiased method for surveying the strain distribution in a human population, but in order to characterize strains more thoroughly a starting material with a higher parasite density is necessary, ideally isolates propagated *in vitro*. The wish for extended genotyping possibilities and desire to obtain *T. gondii* genotype information on the general strain distribution in the same part of Uganda led on to the next study, where free-range chickens were used as a sentinel species for circulating strains.

2.5 Conclusions

The high seroprevalence rate of *T. gondii* in Uganda is alarming in light of the AIDS epidemic. Extrapolation of the data from this study to the rest of the country implies that around 500,000 people in Uganda are co-infected with HIV and *T. gondii* and therefore at risk of developing TE. The high rate of reactivation in this study emphasizes the urgent need for implementation of better diagnosis of toxoplasmosis, using a combination of serology, clinical diagnosis, PCR-detection and response to chemotherapy. Strains belonging to all the three major lineages of *T. gondii* were present in the study population, but the scarcity of parasite DNA in the clinical samples did not allow for thorough genetic analysis. Therefore, further studies exploring the genotypes and phenotypes of Ugandan *T. gondii* strains are needed in order to find out if and how they differ from the archetypal lineages. The next chapter applies multilocus genotyping to tissue samples and mouse isolates from Ugandan chickens.

Chapter 3 Multiple *Toxoplasma gondii* infections are common in Ugandan chickens, providing opportunities for sexual recombination.

3.1 Introduction

Three highly conserved genetic lineages of *T. gondii* are predominant in humans and animals in Europe and North America, and the work presented in Chapter 2 indicated that the same three genotypes are present and cause disease in Ugandan HIV-patients. The importance of AIDS-associated toxoplasmosis highlights the need to understand the population structure and genetic composition of *T. gondii* in Africa, and the following study utilizes samples from free-range Ugandan chickens to assess the diversity of *T. gondii* through multi-locus genotyping.

3.1.1 Chicken as indicator species

Different strains of *T. gondii* display highly divergent levels of virulence in laboratory mice, and the parasite genotype may be important for the disease progression in humans (see section 1.3.1). In individuals suffering from acute toxoplasmosis, it is possible to isolate parasites or perform direct genotyping from blood or cerebrospinal fluid. In asymptomatic infections, however, the parasites are hidden inside tissue cysts and DNA cannot easily be obtained, making it difficult to assess the general parasite strain distribution in a region. Since *T. gondii* is a zoonotic parasite without defined host barriers (Sibley, 2003), genetic analysis of strains present in animals provides an opportunity to estimate the prevalence of different parasite genotypes present in a region.

The domesticated chicken (*Gallus domesticus*) was chosen as sentinel species by Prof. J. P. Dubey to facilitate a direct comparison of the *T. gondii* strains present in the same host species but in different geographic locations. Free-range chickens are ideal for this purpose since they are common worldwide and often stray on or

around human settlements, where they feed from the ground. This behaviour makes them excellent indicators of oocyst contaminated soil and water, a major route of transmission to humans (Bahia-Oliveira et al., 2003; Dubey, 2004). In addition, consumption of infected chicken meat provides a route of direct transmission from chickens to humans. Chickens are popular food animals, and even though meat is generally eaten well cooked in Uganda, handling practices may allow for cross contamination of pathogens from raw to cooked chicken meat (Wanyenya et al., 2004). Isolation and genotyping of *T. gondii* in chickens thus provides important background information on strains that may infect humans, either through oocysts contaminated food/water or through bradyzoites in meat. Collectively, these traits in combination with the general availability and low cost make free-range chickens the ideal model for estimation of the *T. gondii* strain distribution in different parts of the world. Importantly, chickens are largely resistant to overt toxoplasmosis and rarely have symptoms (Dubey et al., 2007c), thus sampling of chickens can be considered an unbiased survey of the *T. gondii* population, compared with typing of parasites from human blood samples, which only reveal disease-causing strains.

3.1.2 *T. gondii* in chickens from different global regions

A large number of studies have investigated the seroprevalence and genotype distribution of *T. gondii* in chickens from different parts of the world. Nearly all of them have been performed in the same laboratory (the lab of Prof. J. P. Dubey, USDA, Beltsville, USA) using the same techniques, and are therefore suitable for direct comparisons. However, the earlier studies, up until 2005, generally relied on SAG2-typing only. The marker kit was later expanded to include 5 RFLP markers and thereafter expanded further to 10-11 different markers. The number of markers used influences the possibility of detecting atypical or recombinant strains, and therefore some of the early studies classified a large number of isolates as clonal even though they were probably not. For example, many Brazilian strains, from chickens as well as other sources, were initially classified as type I or III based on SAG2, but it has later been shown that the population structure in Brazil is very

different from the clonal lineages seen in Europe (see 1.3.3). Many Brazilian strains have a mixture of type I and III alleles and a recent retyping of 151 strains recognised as many as 58 different genotypes in Brazilian chickens, half of which were found only once, while a few genotypes were spread over wide geographic areas (Dubey et al., 2008). The first study on *T. gondii* strains in chickens from sub-Saharan Africa was performed in 2005, resulting in ten isolates from Congo, Kenya, Mali and Burkina Faso, which were typed using SAG2 (Dubey et al., 2005b). This study was later expanded with isolates from Nigeria and Ghana to include a total of 20 isolates, and all samples including the previous ones were retyped using 10 PCR-RFLP markers (Velmurugan et al., 2008). The three samples from Ghana (2) and Nigeria (1) had atypical genotypes with a mixture of alleles, while the rest of the samples appeared to be clonal (with some exceptions for the apicoplast marker), and the predominant genotype was type III (Dubey, 2005, Velmurugan, 2008).

Taken together, these chicken studies provide important background information about the diversity and distribution of *T. gondii* genotypes that are present in different geographic locations and may cause disease in humans. The overall *T. gondii* seroprevalence in chickens, as determined through a modified agglutination test (see 3.2.4), varied from around 25% in sub-Saharan Africa and Asia to nearly 50% in Central- and South America, and the variation in isolation success from seropositive samples ranged from less than 10% from Asian chickens to 80% in seropositive chickens from Central- and South America (Table 3-1). The difference in isolation success rate depends largely on the transport time and temperature conditions, and batches with severely degraded tissues rarely yield viable isolates (J. P. Dubey, personal communication). Table 3-1 provides a summary of the genotypes of *T. gondii* found in chickens from different parts of the world, and details of all included studies sorted by geographic region are listed in Appendix 4.

Table 3-1. Summary of chicken genotyping studies across the world.

	N[a]	Seropos[b]	Isolates[c]	Type I[d]	Type II[d]	Type III[d]	Atypical[d]	Multiple[d]
Europe and North America	1284	411 (32%)	114 (28%)	0 (0%)	94 (82%)	18 (16%)	2 (2%)	0 (0%)
Central- and South America	1508	714 (47%)	568 (80%)	125 (22%)	28 (5%)	141 (25%)	267 (47%)	7 (1%)
Asia	524	143 (27%)	13 (9%)	0 (0%)	5 (38%)	6 (46%)	2 (15%)	0 (0%)
North Africa and Middle East	262	117 (45%)	44 (38%)	0 (0%)	20 (45%)	24 (55%)	0 (0%)	0 (0%)
Sub-Saharan Africa	311	75 (24%)	20 (27%)	1 (5%)	3 (15%)	13 (65%)	3 (15%)	0 (0%)

[a] Number of chickens used for determination of seroprevalence.
[b] Total number of seropositive chickens (seroprevalence)
[c] Number of isolates retrieved (isolation success rate of seropositive chickens).
[d] Number (%) of isolates classified as type I, II, III, atypical or multiple infections.
The data presented summarize studies performed by Dubey et al, and details of all studies and the references are listed in Appendix 4.

3.1.3 *T. gondii* in rats and mice

Mice and rats are considered important reservoirs for *T. gondii* in nature (Webster, 1994; Murphy et al., 2008), and rodents are a direct source of infection of the definitive hosts. Although mice and rats are normally not sources of direct transmission to human hosts, the strains circulating in rodents may enter the human/domestic animal cycle through the domestic cat. Mice are more susceptible to overt toxoplasmosis compared with rats, and are able to transmit the infection vertically from generation to generation without reinfection (Beverley, 1959). Rats, on the contrary, are highly resistant to acute toxoplasmosis, even type I strains cause chronic infections, often with a low parasite burden (Lecomte et al., 1992; Dubey et al., 1999).

The seroprevalence of *T. gondii* in wild-caught mice (*Mus musculus*) and rats (*Rattus norvegicus* and *Rattus rattus*) have not been extensively studied, but the accumulated knowledge has recently been reviewed (Dabritz et al., 2008). Generally, a low seroprevalence (0.0-2.1%) was found in mice, while the results in rats were more variable with several studies reporting an infection rate of 20% or more. However, recent PCR-screening of rodents in the UK has showed a high prevalence (>40%) in both mice and rats (Marshall et al., 2004; Hughes et al., 2006). Although *T. gondii* was originally isolated in an African rodent, no recent reports

have been published regarding the importance of rodents as a reservoir for *T. gondii* in Africa.

3.1.4 Aims

The main objective of this study was to assess the general distribution of *T. gondii* genotypes in Uganda and put this in relation to the genotypes previously found to cause disease in HIV-patients in the same region (Chapter 2). This was achieved through screening and multilocus genotyping of tissues from free-range chickens and wild-caught rodents from residential areas in and around Kampala.

3.2 Materials and Methods

3.2.1 Free-range chickens

Between Sep 4th-6th 2006, 85 free-range chickens were purchased from four different areas in and around Kampala (see Figure 3-3 in the results section). The locations were Gayaza, a suburban area north of the city (N=25), Buloba, a suburb in the west (N=21), Mukono, a rural village in the east (N=19) and the Mulago area, close to the city centre (N=20). The chickens were brought to the Veterinary Faculty of Makerere University where they were bled from the wing vein and thereafter killed. Age estimates were based on the length of the tarometatarsus and the size of the comb, and the chickens were divided into four different age groups. Group 1: 6-9 months (N=27), 2: 10-13 months (N=13), 3: 14-17 months (N=32) and 4: ≥18 months (N=13). The bleeding, killing and age categorization was mainly executed by veterinary doctors William Muyombya and Maria Nakamya at the Veterinary Department of Makerere University.

The chest cavity of each chicken was opened and the heart removed using a sterile, disposable scalpel. The head and heart were placed in separate, pre-labelled zip-lock bags and kept at 4°C. The blood was allowed to clot, first for 1h in room temperature and then overnight at 4°C, and thereafter centrifuged for 5 minutes at 13,000 rpm. Approximately 1 ml of serum was transferred to screw-cap tubes and stored at -20°C.

3.2.2 Wild-caught mice and rats

Ten walk-in traps with bait were placed in different homesteads in the Mulago area in central Kampala. The traps were checked every morning for five days in September 2006. Five mice *(Mus musculus)* and nine rats *(Rattus rattus)* were trapped and brought to the Veterinary Faculty of Makerere University on the day of capture. The traps were placed in a container with tight lid, in which chloroform-drenched cotton was used to euthanize and kill the animals. The intact bodies were placed in zip-lock bags in refrigerators until transported on cold packs to USDA,

Beltsville. One rat and four mice were trapped and put to death in the same manner during two weeks in January 2007, but since the number was so small, they were not sent away, but processed at Makerere (serology only, no bioassay). Dr. William Muyombya was in charge of the rodent catching and euthanization.

3.2.3 Transport

Chicken heads, hearts and sera, and the intact bodies of the rodents were placed on cold packs in polystyrene boxes, brought as checked-in luggage from Kampala to Washington DC. The import was cleared in customs according to US regulations. All samples were in good condition upon arrival and the sample processing at USDA began within 3-6 days of the death of the animals.

3.2.4 *T. gondii* serology (MAT)

The Modified Agglutination Test (MAT) has been developed in the lab of Prof. J. P. Dubey and exploits the same principles as the Toxo-Screen DA test (section 2.2.3). Antigen solution was mixed fresh for each test (96 wells) and consisted of: 2.5 ml antigen diluting buffer, 35 μl 2-mercaptoethanol, 50 μl Evans blue dye solution and 100 μl antigen (formalin-fixed whole parasites). The control in this study was serum from a pig with a known titre of 1:200, used in 8 serial dilutions (1:25, 1:50, 1:100, 1:200, 1:400, 1:800, 1:1,600 and 1:3,200). The reaction was considered valid if the control serum gave a positive result for the first four and a negative result for the last four dilutions. Twenty-five μl of antigen solution was mixed with 25 μl serum dilution and incubation at 37°C over night was followed by visual inspection and interpretation. As for Toxo-Screen DA, a positive reaction lead to the formation of a faint mat consisting of bound antibodies and antigen, while a negative reaction resulted in accumulation of antigen in the bottom of a well visible as a distinct dot or ring.

All chicken sera were first tested for dilutions 1:5, 1:10, 1:20 and 1:40. Sera that were positive at 1:40 were further diluted and tested for 1:80, 1:160, 1:320 and 1:640. For safety reasons, the rodents were not bled while alive, but killed and transported intact. During dissection, their chest cavity was opened and a few drops

of 0.85% saline solution were added, the drawn mixture was spun at 10,000 rpm for 5 min, and the supernatant was used for serology. The serology was performed with the help of with Dr. Oliver Kwok at USDA.

3.2.5 Bioassay and isolation in mice

Bioassay in mice was used to isolate live parasites from rodent and chicken samples and to assess the mouse-virulence of the parasite strains, and a schematic overview of the process is presented in Figure 3-1. Chickens were divided into three groups based on their serological titre. Seronegative samples were pooled in three groups of 15 and bioassayed in cats (one cat per group), samples with a titre of 1:5 or 1:10 were pooled in groups of five and bioassayed in mice (four mice per pool), while samples with a titre of 1:20 or higher were inoculated into four mice each.

First, pepsin digestion was performed on chicken brains and hearts, in order to increase the concentration of parasites and release bradyzoites from tissue cysts. Chicken heads were opened with sterile scissors and bone-cutters, the brain was taken out and mixed together with the heart from the same individual in a sterile blender with saline (0.85% NaCl), until homogenized. For individual samples 50 ml saline was added, for pooled samples from 3-5 chickens 100 ml saline was used. After homogenization an equal volume of prewarmed (37°C) pepsin solution* was added, and the samples were incubated in a 37°C water bath for 45 minutes. Filtration through double layers of gauze removed undigested tissue, and cell sedimentation was achieved through centrifugation for 10 min at 2,000 g. The supernatant was discarded and the bottom phase resuspended in approximately 20 ml of 1.2% sodium bicarbonate ($NaHCO_3$) with neutral red as a pH indicator (added until orange in colour), mixed and spun for 10 min at 2,000 g. The supernatant was discarded and the pellet resuspended in 5 ml of saline with 1000 u penicillin and 100 µg streptomycin per ml. Four outbred, female Swiss-Webster mice were infected subcutaneously with one ml of this suspension, and 400 µl was used to extract DNA for direct parasite strain typing. Mice were bled for MAT-serology from

* 100 ml pepsin solution: 1 g NaCl, 2.80 ml 6N HCl, 0.52 g porcine pepsin and sdH_2O up to a final volume of 100 ml was mixed. Enzyme is activated at 37°C.

the tail vein six weeks post-inoculation, sacrificed through cervical dislocation and dissected in a sterile hood. Brains were used for light microscopy examination, subpassage in mice, cryopreservation and genotyping of infecting parasites. Prof. J. P. Dubey, Dr. Natrajan Sundar, Ms. Katherine Hopkins and Ms. Lucia Bandini performed the bioassay work at USDA, including mouse serology, microscopy and cryopreservation.

3.2.6 Cryopreservation of *T. gondii* from mouse brains

Brains of infected Swiss-Webster mice were ground in sterile mortars with 1 ml PBS until homogenized. The homogenate was transferred to a 15 ml tube with 1 ml freshly made, pre-warmed pepsin solution. This mix was incubated for 45-60 s and thereafter neutralized with $NaHCO_3$ neutral red until orange, whereby the tube was filled up with sterile filtered PBS and centrifuged at 3,000 g for 10 minutes. The supernatant was discarded and the pellet dissolved in cryomedium (DMEM with 25% DMSO and 8% BSA), and aliquoted in 1.5 ml cryovials that were placed in -20°C for 2-3 hours, then -80°C over night before final transfer to liquid nitrogen for long-term storage. The cryopreservation of these samples at USDA was performed by Dr. Natrajan Sundar.

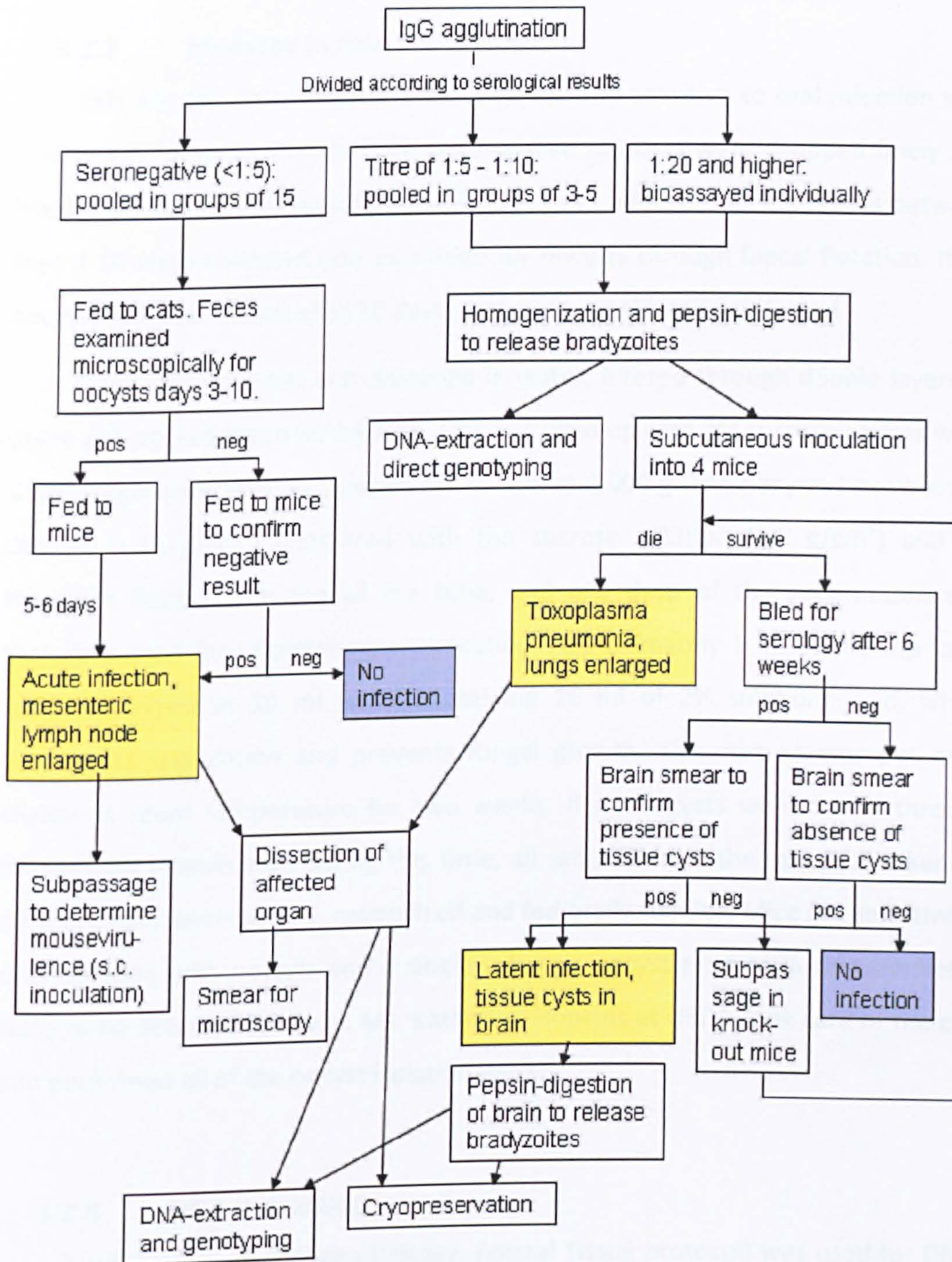


Figure 3-1. Sample processing overview for Ugandan chickens. Samples were divided into three groups based on their serology titre, and thereafter bioassayed. Genotyping was performed directly on chicken tissues as well as tissues from infected mice. Three possible clinical outcomes in infected mice are highlighted in yellow and negative endpoints in blue.

3.2.7 Bioassay in cats

Cats are the definitive hosts for *T. gondii* and sensitive to oral infection with bradyzoites in meat. Tissues from seronegative chickens were chopped finely and mixed with the cats' normal food, which they ate *ad libido*. All cat faeces between days 3-10 were collected and examined for oocysts through faecal flotation. If no oocysts had been found after 10 days, cats were considered uninfected.

Litter was collected and dissolved in water, filtered through double layers of gauze and poured into a 50 ml tube. This was filled up with 2M sucrose*; tubes were shaken vigorously and centrifuged for 10 min at 2,000 g. The oocysts have a lower density (1.11 g/cm^3) compared with the sucrose solution (1.8 g/cm^3) and do therefore float at the top of the tube, and one drop of the supernatant was therefore used for microscopic examination. Approximately 1 ml of the top layer was transferred to 50 ml tubes containing 10 ml of 2% sulphuric acid, which accelerates sporulation and prevents fungal growth. The mixture was put on a shaker at room temperature for two weeks. If no oocysts were found through microscopic examination during this time, all samples from the same cat taken on different days were pooled, neutralized and fed orally to mice. Mice are sensitive to oral infection with oocysts and a single infective oocyst present in the samples is likely to be detected this way. Ms. Katherine Hopkins at USDA took care of the cats and performed all of the oocyst isolation work.

3.2.8 DNA-extraction

A commercial kit (QIAgen DNeasy, Animal Tissue protocol) was used for DNA-extraction. For direct genotyping from chickens, the pellet from 400 µl brain-heart pepsin digest homogenate (section 3.2.5) was used as template. For genotyping of parasites isolated through mouse bioassay, DNA was extracted from 200 µg brain (by Dr. Natrajan Sundar) and sent to Leeds University, where genotyping was performed.

---

* 20 litre batch of 2M sucrose solution: 13 kg sucrose was boiled with water, then 100 ml of phenol was added to prevent fungal growth.

Briefly, 180 µl tissue lysis buffer and 20 µl Proteinase K was added to the pellet/brain tissue and mixed by vortexing. This was incubated in a water bath at 55°C, for 3 hours. Thereafter, the procedure was the same as for the human blood samples (see 2.2.4).

3.2.9 Genotyping PCRs

Five different PCR-RFLPs were used for molecular detection and genotyping of *T. gondii*, both directly from chicken and rodent tissues and from infected laboratory mice. The reactions amplified the loci SAG1, SAG2 (5′ and 3′ –end), SAG3, BTUB and GRA6, according to a published method (Dubey et al., 2006a). Primers and restriction enzymes are listed in Table 3-2.

Table 3-2. Primers and enzymes used for detection and genotyping.

Gene (Chrom)		Forward primer	Length	Rest. enz.	Fragment lengths[a] Type I	Type II	Type III
SAG1 (VIII)	ext	GTTCTAACCACGCACCCTGAG (F)	476 bp	Sau96I,	334	293	293
		AAGAGTGGGAGGCTCTGTGA (R)		HaeII	56	97	97
	int	CAATGTGCACCTGTAGGAAGC (F)	390 bp				
		GTGGTTCTCCGTCGGTGTGAG (R)					
SAG2-5′ (VIII)	ext	GCTACCTCGAACAGGAACAC (F)	335 bp	Sau3AI	242	242	186
		GCATCAACAGTCTTCGTTGC (R)					56
	int	GAAATGTTTCAGGTTGCTGC (F)	242 bp				
		GCAAGAGCGAACTTGAACAC (R)					
SAG2-3′ (VIII)	ext	TCTGTTCTCCGAAGTGACTCC (F)	327 bp	HhaI	222	169	222
		TCAAAGCGTGCATTATCGC (R)				53	
	int	ATTCTCATGCCTCCGCTTC (F)	222 bp				
		AACGTTTCACGAAGGCACAC (R)					
BTUB (IX)	ext	TCCAAAATGAGAGAAATCGT (F)	527 bp	TaqI,	220	191	127
		AAATTGAAATGACGGAAGAA (R)		BsiEI	118	127	118
	int	GAGGTCATCTCGGACGAACA (F)	411 bp		73	93	93
		TTGTAGGAACACCCGGACGC (R)					73
GRA6 (X)	ext	ATTTGTGTTTCCGAGCAGGT (F)	546 bp	MseI	258	183	161
		GCACCTTCGCTTGTGGTT (R)			86	161	97
	int	TTTCCGAGCAGGTGACCT (F)	344 bp				86
		TCGCCGAAGAGTTGACATAG (R)					
SAG3 (XII)	ext	CAACTCTCACCATTCCACCC (F)	311 bp	NciI	100	226	162
		GCGCGTTGTTAGACAAGACA (R)			64		64
	int	TCTTGTCGGGTGTTCACTCA (F)	226 bp		62		
		CACAAGGAGACCGAGAAGGA (R)					

[a]A double digest of SAG1 amplicons with Sau96I and HaeII can differentiate type I from II/III, and also detect an unusual allele "u-1" if both enzymes cut the product (237+97+56 bp), see Figure 3-8.

All the six external primer pairs were run together in a multiplex PCR-reaction. The total reaction volume was 25 µl: 20 µl H_2O, 2.5 µl 10*PCR-buffer (New England BioLabs), 0.5 µl dNTPs (10mM), 0.1 µl of each primer (50 pM/ µl), 0.1 µl Taq polymerase (New England BioLabs, 5 U/ µl) and 1 µl DNA-template. The amplification program for the external multiplex PCR was 95˚C for 4 min followed by 25 cycles of 94˚C (30 s), 55˚C (1 min) and 72˚C (2 min). One µl of the product was used as template for the nested PCRs, where each locus was amplified separately. The mastermix and PCR-program for the nested rounds was the same as above but with the following exceptions: 21 µl H_2O was added since only two instead of twelve primers were used, the annealing temperature was increased to 60˚C for all reactions, and the number of cycles was increased to 35.

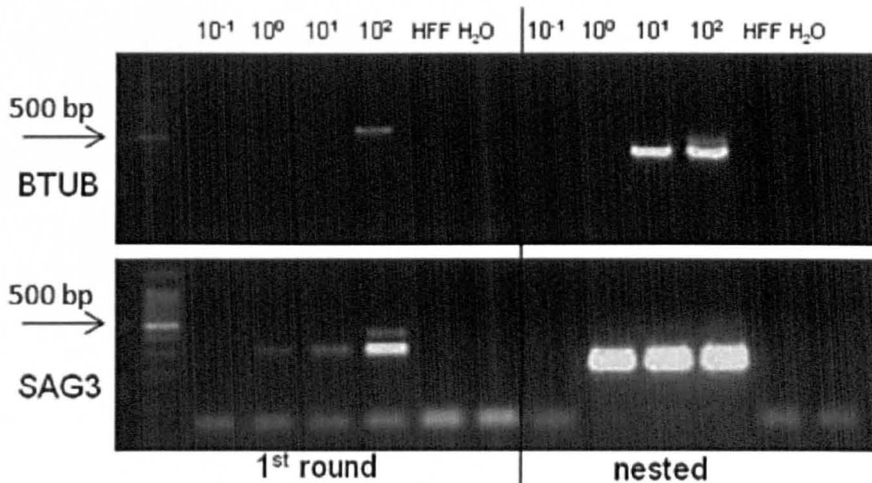


Figure 3-2. Sensitivity testing of the genotyping PCRs for BTUB and SAG3. BTUB had lower sensitivity, needing around 10 copies of parasite DNA for detection, and it was also the marker yielding the lowest number of positive results when applied to chicken and rodent samples. A 100-bp size marker was used, and the short fragments seen at the bottom are primer-dimers.

Restriction enzyme digests were performed in 20 µl reaction volumes, consisting of 12.6 µl H_2O, 2 µl buffer, 0.2 µl BSA, 0.1 µl enzyme and 5 µl PCR-product. The reactions were incubated for 1 hour at 37˚C, except for the BTUB-gene where the temperature was 60˚C for the double digest BsiEI and TaqI. Enzymes were from New England BioLabs, except HhaI that was from Promega, and the corresponding buffers were enclosed with the enzymes. Fragments were visualized on 2% agarose gels under UV-light.

3.2.10 Sequencing and sequence analysis

PCR-amplicons selected for sequencing were purified using the QIAgen PCR purification kit, which removes buffer components, primers and free nucleotides. The elution volume was 30 µl, of which 10 µl was used for each sequencing reaction, together with 4 µl of the appropriate primer. Technical personnel at the University of Leeds performed the sequencing reaction using an ABI377 DNA Sequencer machine (Applied BioSystems).

Chromatograms and automatic reads generated by the sequencer were viewed and manually controlled and edited in the BioEdit Sequence Alignment Editor software (http://bioedit.software.informer.com/). BLAST searches at NCBI Blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and ToxoDB (www.toxodb.org) were used to retrieve sequences for the same loci from different strains and alignments were made directly in BioEdit or in ClustalW (http://www.ebi.ac.uk/ Tools/clustalw2/index.html). Successfully sequenced amplicons were submitted to GenBank (Accession numbers EF585673-EF585715).

3.3 Results

3.3.1 Serological screening of *T. gondii* in chickens and rodents

Eighty-five free-range chickens were obtained from four different areas in around Kampala: Gayaza (N=25), Buloba (N=21), Mukono (N=19) and Mulago (N=20), see Figure 3-3.

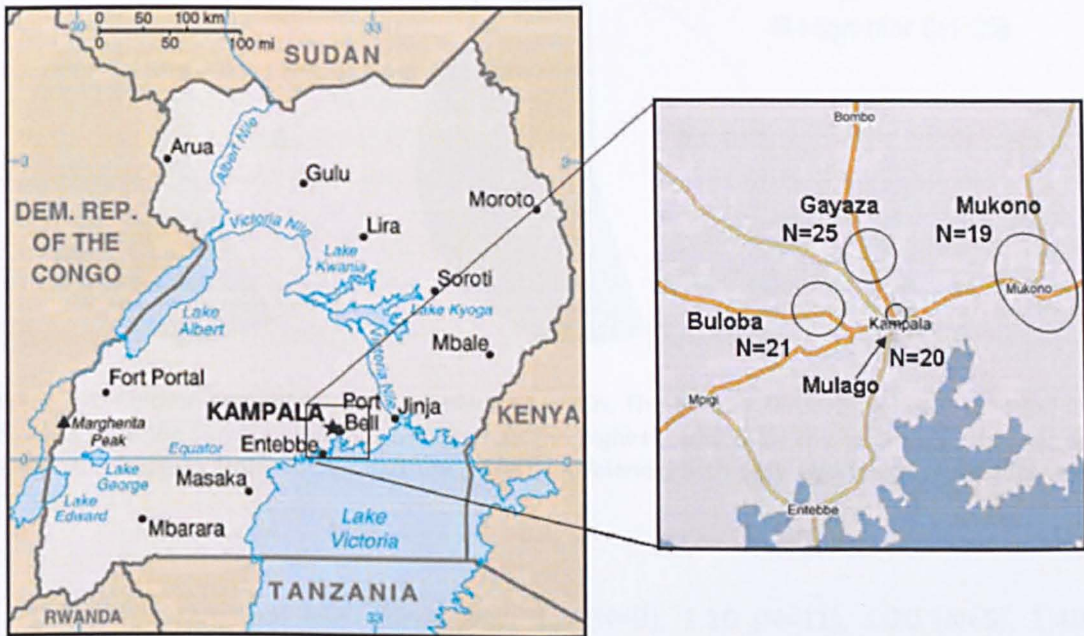


Figure 3-3. Map of sampling areas. Twenty chickens originated from the Mulago area, close to the city centre, 21 and 25 came from sub-urban settings in Buloba and Gayaza respectively, while the last 19 were obtained from rural areas around Mukono. The map is taken from the CIA factbook (https://www.cia.gov/library/publications/the-world-factbook/geos/ug.html).

Forty chickens were seropositive at a MAT-titre of 1:5 or higher, giving a total seroprevalence of 47%. There was a significant difference in total seroprevalence among the chickens from Buloba compared with those from the other locations ($P<0.001$, chi-square test). The two positive chickens from Buloba both had a low titre (1:5 or 1:10), while 20 chickens, originating from Gayaza (N=6), Mukono (N=6) and Mulago (N=8) had a titre of 1:20 or higher. The prevalence of negative, low-titre (1:5-1:10) and high titre (≥1:20) chickens from different locations are shown in Figure 3-4.

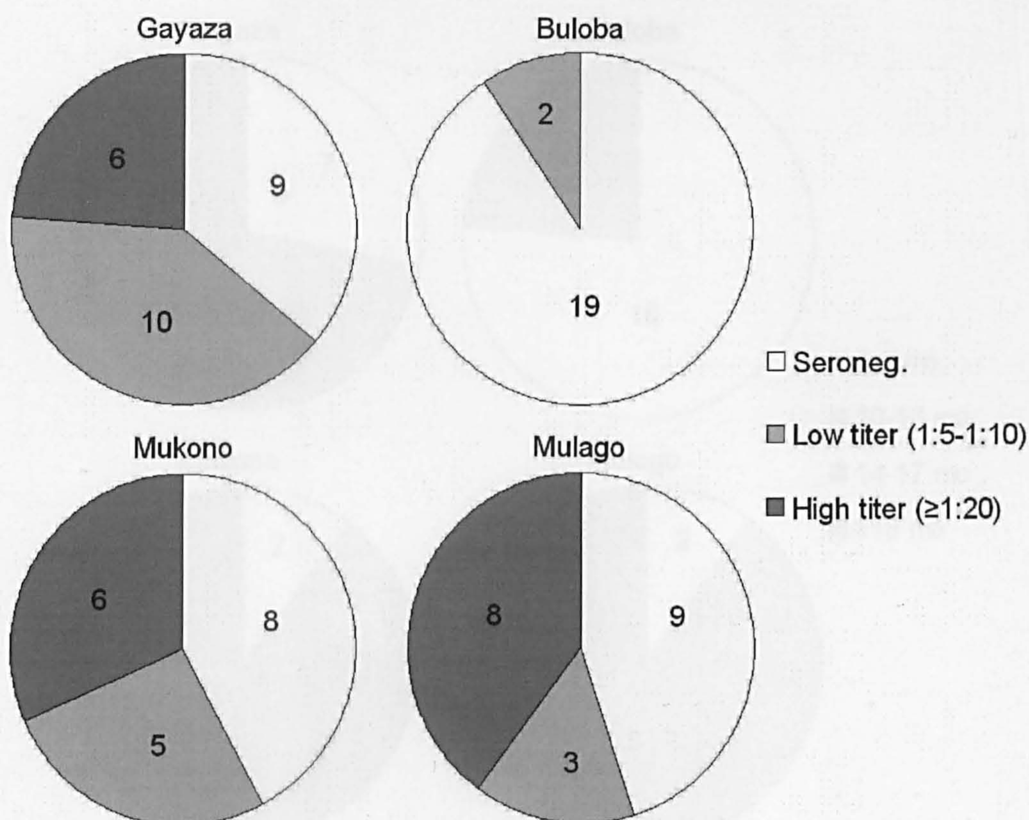


Figure 3-4. Chicken serology results, by place of origin. The highest seroprevalence was found in chickens from the Gayaza area, but Mulago had the highest number of chickens with a titre of at least 1:20. Chickens from Buloba had the lowest prevalence, with only two low-titre seropositive chickens.

The full range of MAT-titres was; 1:5 (N=9), 1:10 (N=11), 1:20 (N=5), 1:40 (N=12), 1:80 (N=1), 1:160 (N=1), 1:320 (N=1), and a table listing all the site of origin, sex, age categorization, MAT-titre and the pooling of samples for PCR and bioassay for each chicken is provided in Appendix 5. Briefly, chickens with a low MAT-titre were pooled in groups of 3-5, while the 20 chickens with a titre of at least 1:20 were bioassayed individually, and the same was true for the direct PCR-detection and genotyping. This cut-off value was based on Prof. Dubey's experience of isolation success in chickens with different MAT-titres.

The chickens from Buloba were considerably younger then the chickens from the other areas, which may be one explanation for the lower seroprevalence seen in this group. As described in section 3.2.1, chickens were divided into four age groups, and 76% of Buloba chickens belonged to age group 1 (6-9 mo), compared with 10-28% in the other groups (Figure 3-5).

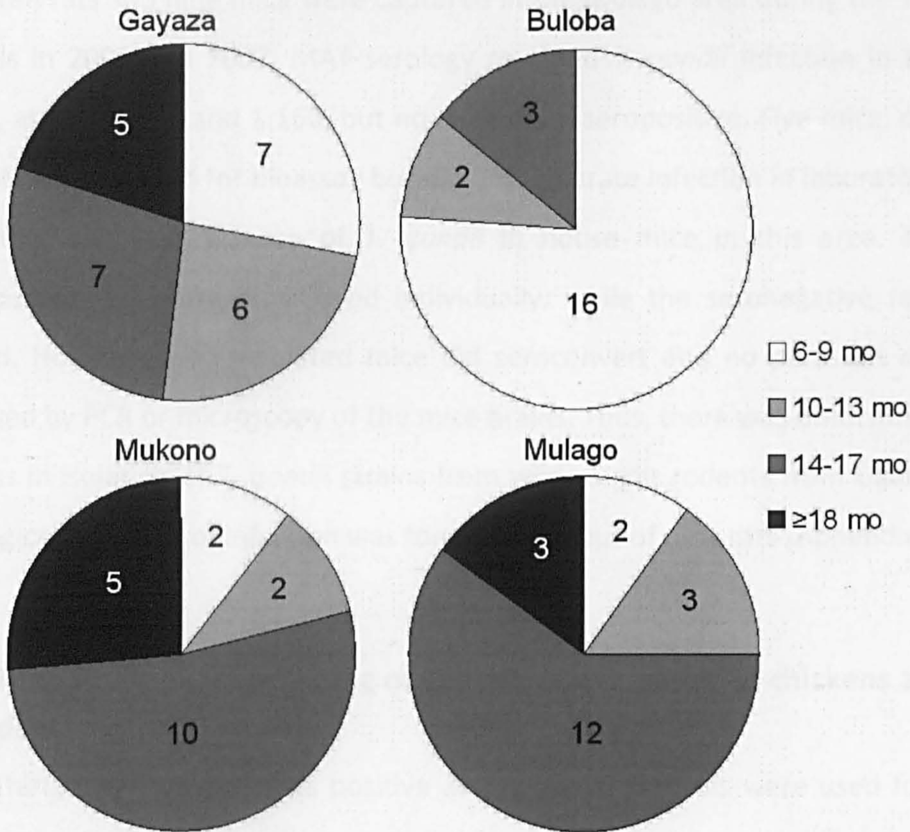


Figure 3-5. Chicken age distribution, by place of origin. The chickens from Buloba had a higher proportion of young animals (6-9 months old), which is probably correlated to the lower level of *T. gondii* exposure seen in this group.

A tendency towards an increase in seroprevalence with age was observed, and the prevalence rate reached 69% in chickens aged ≥18 months (Figure 3-6), but the difference between age groups was not statistically significant (as tested by the chi-square test) and chickens of all ages carried *T. gondii* infections.

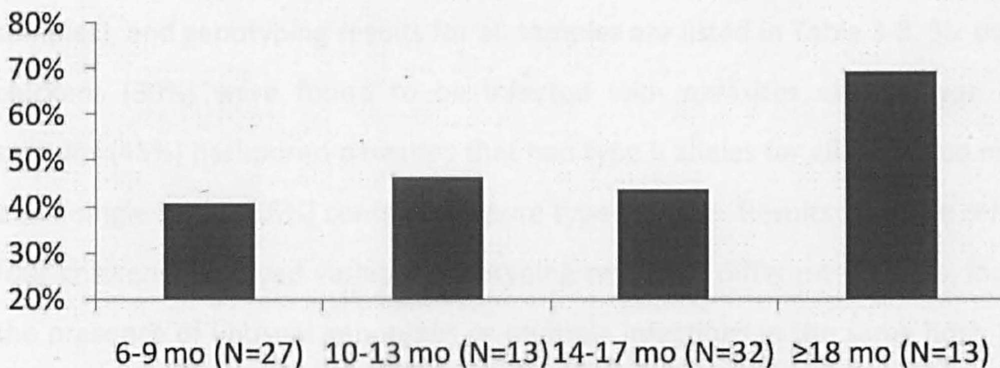


Figure 3-6. Chicken seroprevalence in different age groups. Older chickens are more likely to have been exposed to *T. gondii* infections, but the difference in seroprevalence was not significant between age groups among the Ugandan chickens.

Ten rats and nine mice were captured in the Mulago area during the sampling periods in 2006 and 2007. MAT serology revealed *T. gondii* infection in two rats (22%), at titres 1:40 and 1:160, but no mice were seropositive. Five mice, captured in 2006, were pooled for bioassay but did not generate infection in laboratory mice, indicating a low prevalence of *T. gondii* in house mice in this area. The two seropositive rats were bioassayed individually, while the seronegative rats were pooled. However, no inoculated mice did seroconvert and no parasites could be detected by PCR or microscopy of the mice brains. Thus, there was unfortunately no success in isolation of *T. gondii* strains from wild-caught rodents from Uganda, but serological evidence of infection was found in two out of nine rats (Appendix 6).

3.3.2 Direct genotyping of *T. gondii* from Ugandan chickens and rodents

Thirty DNA samples plus positive and negative controls were used for direct genotyping from chicken and rodent tissues, where a multiplex external PCR was followed by separate amplifications with internal primers for five different genes (section 3.2.9). The samples consisted of twenty individual chickens, three groups of pooled chickens with a MAT-titre of 1:10 (10A, 10B, 10C), two pools of chickens with titre 1:5 (pools 5A and 5B), two individual rats, two groups of pooled rats (RA and RB) and the pooled samples from five mice (M), as listed in Appendices 5 and 6.

T. gondii-DNA was detected directly from tissue samples in all the 20 chickens that were bioassayed individually (MAT-titre ≥ 1:20). Figure 3-7 shows the first screening at locus SAG2-3' (repeat amplification later yielded three more positive samples), and genotyping results for all samples are listed in Table 3-3. Six out of 20 chickens (30%) were found to be infected with parasites of genotype I. Nine samples (45%) harboured parasites that had type II alleles for all amplified markers, and a single sample (5%) contained a pure type III strain. Results from the remaining four chickens displayed variable genotyping results at different markers, indicating the presence of unusual genotypes or multiple infections in the same host. For Ch-62 and Ch-78 alternating type I and type II alleles were detected, while Ch-2 and Ch-70 contained parasites with alleles from types II and III.

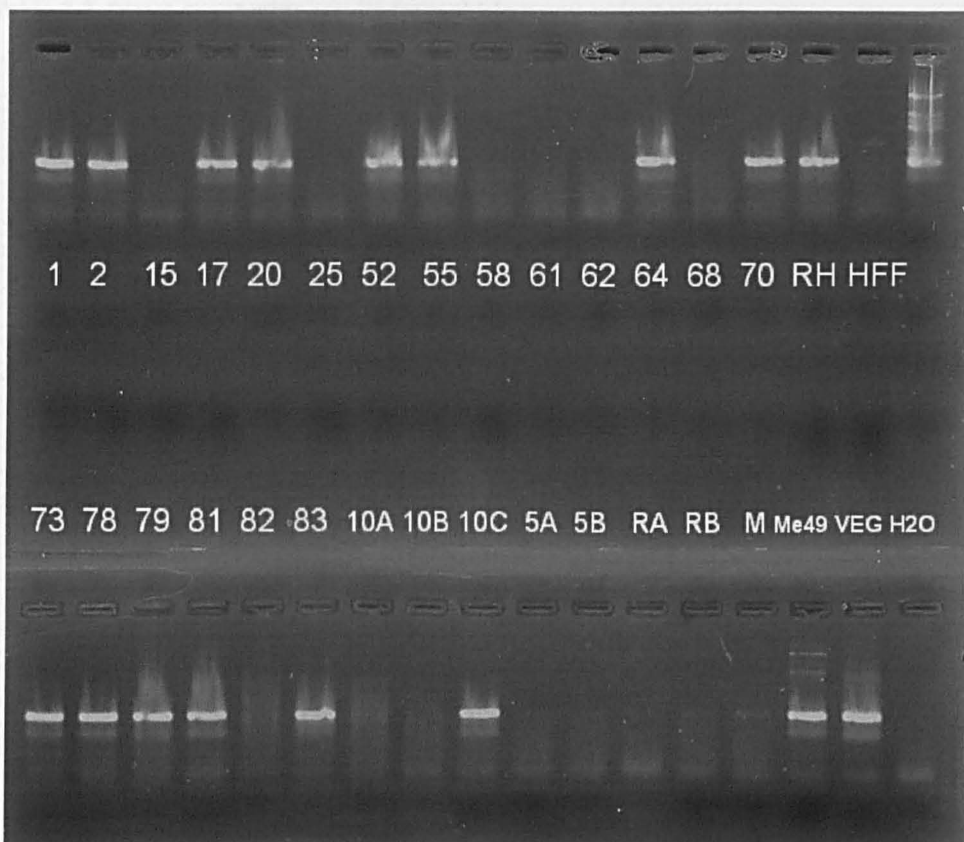


Figure 3-7. Nested SAG2-3' first screening results for Ugandan chicken and rodent samples. RH (*T. gondii* type I) and DNA from human fibroblasts were used as positive and negative controls at the first line of the gel, where a ladder is present in the last well, while Me49 (*T. gondii* type II), VEG (*T. gondii* type III) and water were used as controls on the second row. Numbers indicate the samples, 1=Ch-1, 2=Ch2 and so on. The amplified fragments are 222 bp of length.

In addition, the sample from Ch-70 contained an atypical allele "u-1" at the SAG1 locus (Figure 3-8). This allele is digested by both restriction enzymes (Sau96I and HaeII), and had previously been found in *T. gondii* strains from Colombia (Dubey et al., 2006b), and later also in Sri Lanka, China and Brazil (Dubey et al., 2007b; Dubey et al., 2007d).

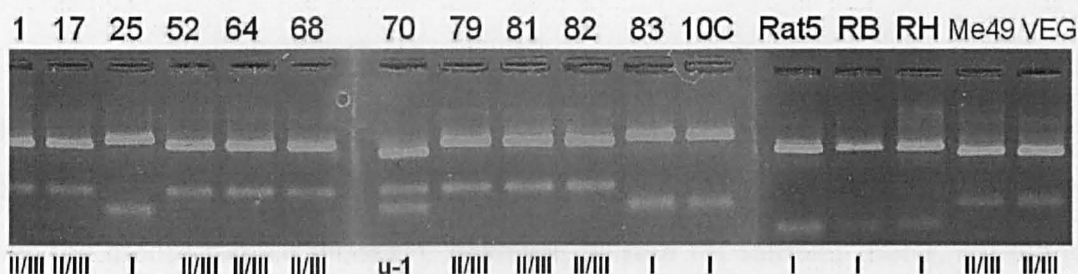


Figure 3-8. Double restriction enzyme digests of SAG1 amplicons. *HaeII* cuts the 390 bp product of type II or type III strains into two fragments of 293 and 97 bp, while *Sau96I* digests type I amplicons into 334+56 bp. For the parasites present in Ch-70 a double digest was seen, resulting in fragments of 237+97+56 bp, and this allele has also been detected in South America and Asia.

Table 3-3. Direct genotyping results from chicken and rodent tissues.

	IgG-titre	SAG1[a]	SAG2 5'[b]	SAG2 3'[c]	SAG3	BTUB	GRA6	Genotype
Ch-1	1:80	II / III	I / II	II	II	II	II	II
Ch-2	1:40	-	I / II	II	-	-	III	II + III
Ch-15	1:40	-	I / II	-	II	-	II	II
Ch-17	1:160	II / III	I / II	II	II	II	II	II
Ch-20	1:20	-	-	II	II	-	-	II
Ch-25	1:20	I	I / II	I / III	-	-	-	I
Ch-52	1:40	II / III	I / II	II	II	II	II	II
Ch-55	1:20	-	-	I / III	I	-	I	I
Ch-58	1:40	-	I / II	-	I	-	I	I
Ch-61	1:40	-	I / II	I / III	I	-	I	I
Ch-62	1:20	-	-	II	I	II	II	I + II
Ch-64	1:320	II / III	III	I / III	III	III	III	III
Ch-68	1:40	II / III	-	-	-	-	II	II
Ch-70	1:40	u-1	I / II	II	III	III	II	II + III
Ch-73	1:40	-	-	I / III	I	-	-	I
Ch-78	1:20	II / III	I / II	I / III	I	-	II	I + II
Ch-79	1:40	II / III	I / II	II	II	II	II	II
Ch-81	1:40	II / III	I / II	II	II	II	II	II
Ch-82	1:40	II / III	I / II	-	II	II	II	II
Ch-83	1:40	I	I / II	I / III	I	-	I	I
Pool 5A	1:5	-	-	-	-	-	-	-
Pool 5B	1:5	-	-	-	-	-	-	-
Pool 10A	1:10	-	-	-	-	-	-	-
Pool 10B	1:10	-	-	-	-	-	-	-
Pool 10C	1:10	I	-	I / III	-	-	I	I
Rat-2	1:160	-	I / II	-	-	-	-	I or II
Rat-5	1:40	I	-	-	-	-	-	I
Pool RA	0	-	I / II	-	-	-	-	I or II
Pool RB	0	I	I / II	-	-	-	-	I
Pool M	0	-	-	-	-	-	-	-
RH	NA	I	I / II	I / III	I	I	I	I
Me49	NA	II / III	I / II	II	II	II	II	II
VEG	NA	II / III	III	I / III	III	III	III	III
water	NA	-	-	-	-	-	-	

[a] SAG1 can only differentiate between type I and non-type I.
[b] SAG2-5' can only differentiate between type III and non-type III.
[c] SAG2-3' can only differentiate between type II and non-type II.

One of the pooled chicken samples was PCR-positive (Pool 10C, consisting of three Mulago chickens with MAT-titre 1:10), and was found to contain a type I strain. The samples from chickens with titre 1:5 or 1:10 were negative by PCR, as was the mouse sample (Pool M). However, all four rat samples (Rat-2, Rat-5 and pools RA and RB) were positive at one or more genotyping loci. Two of these, Rat-5 and Pool RB, were shown to harbour parasites of genotype I, while the other two (Rat-2 and Pool RA) could only be classified as being either type I or II (Table 3-3).

3.3.3 Bioassay results and genotyping of isolates

Tissues from 20 chickens with MAT-titre of 1:20 or higher were homogenized, digested in acidic pepsin and inoculated into four mice per sample. Mice from nine of these twenty groups were seropositive, while none of the pooled samples caused the mice to seroconvert, indicating an insufficient quantity of viable parasites for infection. Furthermore, none of the cats fed seronegative chicken tissues shed oocysts, which confirms that chickens with undetectable antibodies were uninfected, and no mice inoculated with rat or mouse tissues did seroconvert. The isolation rate was thus 45% among chickens with MAT ≥1:20.

The nine *T. gondii* strains that were infective to mice originated from Ch-1, 2, 17, 52, 68, 70, 79, 81 and 82. Out of 35 mice infected with these samples (nine groups of four, minus one mouse who died), 33 were seropositive at a MAT-titre of 1:25. The parasite density was low for all of these strains; only a small number of tissue cysts were visible in the brains, and for nearly a third of the positive mice no parasites were detected by careful microscopic examination performed by an experienced parasitologist (Prof. J. P. Dubey). All the nine isolates thus appeared to be slow growing, and since none of the Ugandan strains caused symptoms in laboratory mice, they were classified as avirulent. Table 3-4 lists all mice that became infected, and all mice where no tissue cysts could be seen in the brain are shaded in grey.

Six of the retrieved isolates originated from samples that had been found to be infected with type II strains through direct genotyping (Ch-1, 17, 52, 79, 81 and 82), and genotyping of DNA extracted from the brains of infected mice using four markers (SAG2-5', SAG2-3', SAG3 and GRA6) confirmed this typing (Table 3-4). These samples had all been readily genotyped on most markers straight from tissue samples and all mice in these groups seroconverted (4/4 mice, but 3/3 in the Ch-1 group, where one mouse died before the end of the experiment), which indicates a comparably high parasite density. Tissue cysts were visible in the brain for most of these mice, except one mouse from the Ch-81 and two from the Ch-82 group (Table 3-4).

Table 3-4. Genotyping results from mice inoculated with tissue from infected chickens

	Mouse-ID[a]	SAG2 5'	SAG2 3'	SAG3	GRA6	Total[b]
	A1	I / II	II	II	II	
	A2	I / II	II	II	II	
Ch-1	A3	I / II	II	II	II	II
	A4	I / II	I / III	I + II	II	
	A5	III	I / III	III	II	
Ch-2	A6	III	I / III	III	II	Mix
	A8	I / II	II	II	II	
	A9	I / II	II	II	II	
	A10	I / II	II	II	II	
Ch-17	A11	I / II	II	II	II	II
	A12	I / II	II	II	II	
	A13	I / II	II	II	II	
	A14	I / II	II	II	II	
Ch-52	A15	I / II	II	II	II	II
	A19	I / II	I / III	I	I	
	A20	I / II	II	I	I + II	
Ch-68	A21	I / II	II	II	II	Mix
	A22	I / II	I / III	I	I	
	A23	I / II	II	I + III	I + II	
	A24	I / II	I / III	II + III	II	
Ch-70	A25	I / II	I / III	-	-	Mix
	A26	I / II	II	II	II	
	A27	I / II	II	II	II	
	A28	I / II	II	II	II	
Ch-79	A29	I / II	II	II	II	II
	A30	I / II	II	II	II	
	A31	I / II	II	II	II	
	A32	I / II	II	II	II	
Ch-81	A33	I / II	II	II	II	II
	A34	I / II	II	II	II	
	A35	I / II	II	II	II	
	A36	I / II	II	II	II	
Ch-82	A37	I / II	II	II	II	II

[a]Seropositive mice without visible tissue cysts are shaded in grey.

[b]Six strains, from chickens Ch-1, 17, 52, 79, 81 and 82, were type II strains as determined by direct genotyping from chicken tissues as well as the typing from brains of infected mice shown above. Strains originating from chickens Ch-2, Ch-68 and Ch-70 all had a mixture of alleles and divergent patterns seen in different mice. The presence of more than one allele at the same locus showed that these were multiple infections of two (Ch-68) or three (Ch-2 and Ch-70) different strains.

The last three isolates came from samples that had been found by direct genotyping to contain parasites with type II alleles (Ch-68) or a mixture of type II and III alleles (Ch-2 and 70). Three out of four mice (A4-A6) from the Ch-2 group had seroconverted, and brain cysts were visible in two of them (A5 and A6). Genotyping of A4 showed a mixture of type I and II alleles: type II for GRA6, type I for SAG2 and

a sequence pattern with double peaks at SAG3 (Figure 3-9). A5 and A6 both had the same allelic pattern, but it was different from A4: type III at loci SAG2 and SAG3, and type II at GRA6. This suggests multiple infections in the same chicken with three different strains, since three different alleles were detected at the same locus (SAG-3). From these data, it is not possible to deduce whether the strains in question were true types I, II and III or had a combination of alleles.

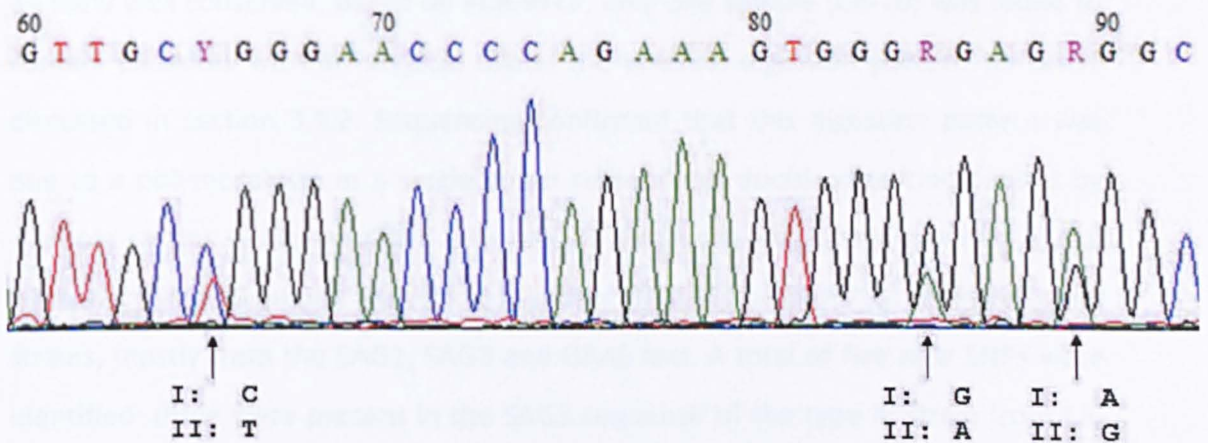


Figure 3-9. Chromatogram with double peaks indicative of co-infection. The figure shows a stretch of sequence in the SAG3 gene from Ch-2, mouse A4, and arrows point to the differences between genotype I and II (all other bases within this region are identical in these lineages). The proportional height of the type I and II is conserved in the three SNPs, and type I appeared to be the predominant strain in this infection.

Out of four mice in the Ch-68 group, three were seropositive (A19-A21) and cysts were seen in one (A20). Here two different genotypes were detected, type I in mouse A19, type II in mouse A21, and both I and II in mouse A20. From this result it was concluded that Ch-68 was infected with two different strains belonging to lineages I and II.

The most complex pattern was seen in the group of mice infected with tissues from Ch-70 (Table 3-4). All four mice (A22-A25) were seropositive, but none had visible brain cysts, and the genotype pattern found in the different mice contained alleles from all three lineages: the strains in A22 and A25 had type I alleles at all amplified loci, while A23 and A24 appeared to contain strains of genotypes I, II and III. Although it was not possible to deduce the genetic make-up of the individual infecting strains when a complex pattern was seen, it was shown that three distinct

alleles were present at the SAG3 locus, which indicates a multiple infection with three different strains.

When mixed genotypes were detected via RFLP, sequencing was performed to verify the results. Analysis of these sequences frequently revealed double peaks at known polymorphic sites indicating the presence of multiple alleles (Figure 3-9), which strengthens the evidence for co-infection with several *T. gondii* strains in the same host. The short gene fragments utilized for genotyping in this study appear to be fairly well conserved. Based on PCR-RFLP, only one sample (Ch-70) was found to contain parasites with an atypical allele; the unusual digestion pattern for SAG1 discussed in section 3.3.2. Sequencing confirmed that this digestion pattern was due to a polymorphism in a single strain rather than double digestion caused by multiple strains being amplified. Overall, 47 sequence fragments covering >15kb were retrieved (GenBank Accession numbers EF585673-EF585715) from different strains, mostly from the SAG1, SAG3 and GRA6 loci. A total of five new SNPs were identified: three were present in the SAG3 sequence of the type III strain from Ch-70, one in the SAG1 sequence from Ch-25, and one in the SAG2-3' sequence from Ch-83 (see Appendix 7). These results indicate a high level of conservation between the Ugandan strains and the three archetypal lineages.

3.4 Discussion

3.4.1 Serotitre was related to amplification and isolation success

In the current study, there was a clear correlation between the serological titre of the chicken and the success of direct PCR-genotyping and isolate recovery (Figure 3-10). Parasites from all chickens with a titre of 1:20 or higher could be genotyped directly from the chicken tissue homogenate by at least two markers, and all *T. gondii* strains from chickens with titres of 1:80 or higher were readily amplified at all six loci. It is therefore likely that the amplification success and serological titres were correlated to the parasite density in the organs used for bioassay and DNA-extraction.

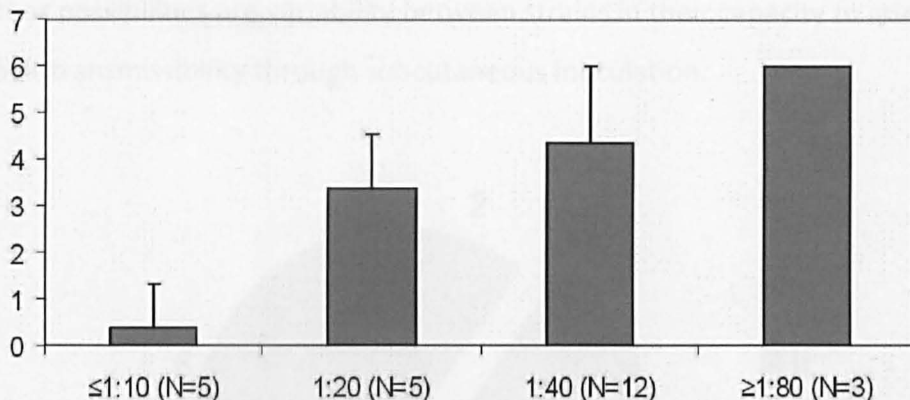


Figure 3-10. Amplification success in relation to serological titre in chicken. The higher the serological titre in the host, the easier was amplification of parasite DNA, suggesting a correlation between serological titre and parasite density in chickens. Bars show the average number of amplified markers plus standard deviation.

Parasite isolation from Ugandan chickens through mouse inoculation was successful for nine *T. gondii* strains, all of which originated from birds with a MAT-titre of at least 1:40, and the total isolation rate was 22.5% (9/40 seropositive chickens). MAT-titres as low as 1:5 have been shown to be meaningful in chickens and occasionally result in isolates (Dubey et al., 2003a; Dubey et al., 2005a). However, isolate retrieval is accomplished more often from animals with a higher

titre (Dubey et al., 2003b), and the isolation success rate of 22.5% seen in this study is similar to that found in most parts of the world (Table 3-1).

3.4.2 Genotype related bias in isolation success

An unexpected finding in this study was the preferential selection of genotype II in the mouse bioassay. Mice from nine out of 20 groups showed serological evidence of infection 42 days post-inoculation. Six of the causing strains were pure type II, while three samples originated from chickens infected with type II and at least one other strain. Among the eleven strains that failed to grow in mice, six belonged to genotype I, two were type II, two were mixed I/II infections and one was a type III (Figure 3-11). The type I strains in this study were difficult to amplify by several PCR–markers and limited parasite numbers may be the explanation for isolation failure. However, the same logic cannot be applied to the type III strain (Ch-64) which came from a high titre chicken (1:320) and was easily amplified at all loci. Other possibilities are variability between strains in their capacity to infect mice or in their transmissibility through subcutaneous inoculation.

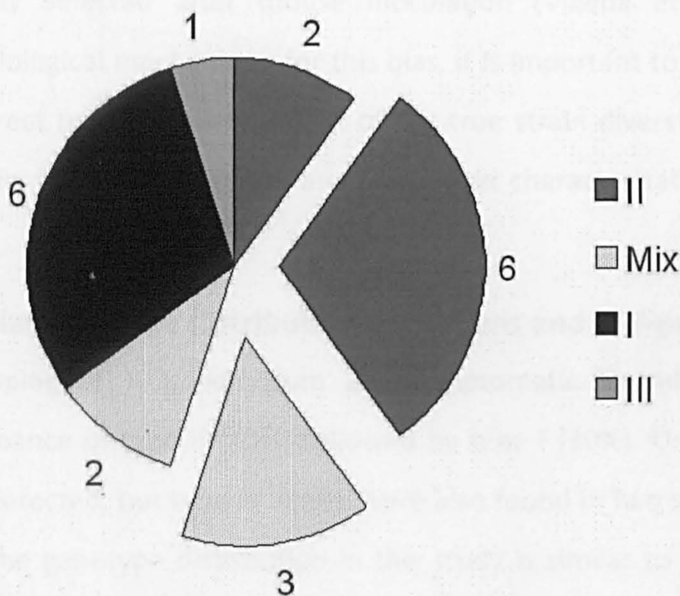


Figure 3-11 Genotype related bias in mouse bioassay. Six of eight type II strains did successfully infect mice, while none of the type I strains did, and this difference was highly significant ($P<0.01$ between type I and II, chi-square test). In addition, three of five isolation attempts on samples from chickens with multiple infections were successful.

The ratio of tachyzoites to encysted bradyzoites can vary between strains and influence their survival during transport and their susceptibility to treatment with acidic pepsin. Tachyzoites are normally killed by pepsin, while bradyzoites are resistant and remain infectious upon this treatment (Dubey, 1998c), and inefficient tissue cyst formation has been associated with decreased oral infectivity in mice (Fux et al., 2007). The pepsin treatment applied to the chicken tissues in this study mimics the environment in the digestive tract and aims to release bradyzoites from cysts. Strains with immature cysts or a low proportion of bradyzoites are therefore less likely to survive and remain infectious after pepsin digestion, and while the cystogenic ability of different strains has not been specifically investigated in chickens, several highly virulent type I strains are non-cystogenic in mice (Ferreira Ade et al., 2006). Although type I strains have been isolated in the same laboratory according to the same protocol, these are almost exclusively from South America (see Appendix 4) and there might be inherent differences between type I strains from different parts of the world that render them more or less susceptible to this isolation bias. It has been noted previously that studies which employ direct typing often detect a comparatively high number of type I and mixed infection, while type II are preferentially selected after mouse inoculation (Villena et al., 2004). Regardless of the biological mechanisms for this bias, it is important to be aware of this and employ direct typing for assessment of the true strain diversity, although isolates are necessary for in depth genetic and phenotypic characterizations.

3.4.3 Similar genotype distribution in chickens and HIV-patients

Direct genotyping of *T. gondii* from 20 asymptomatic Ugandan chickens showed a predominance of type II (40%) followed by type I (30%). Only one pure type III strain was detected, but type III alleles were also found in two samples with mixed infections. The genotype distribution in this study is similar to that seen in HIV-patients from the same part of Uganda (Figure 3-12).

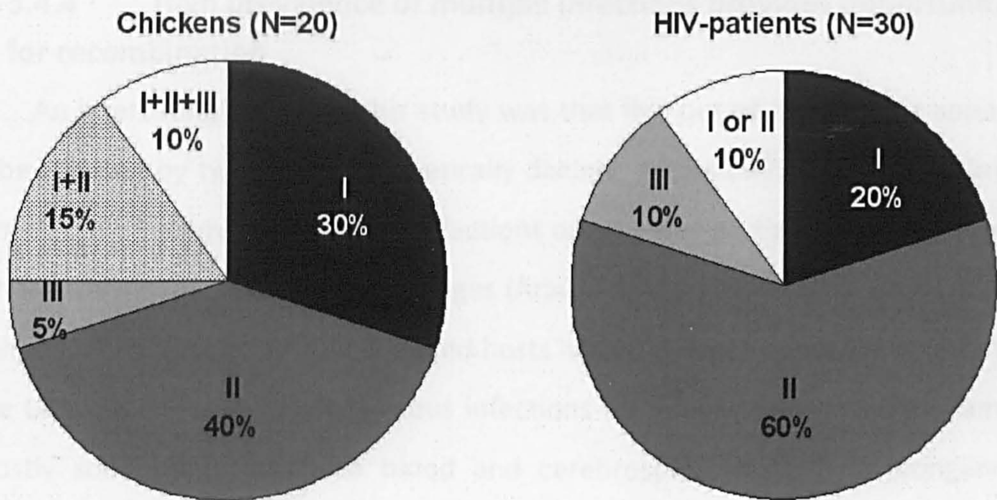


Figure 3-12. *T. gondii* strain distribution in chickens and HIV-patients from Uganda. The strain distribution was similar in HIV-patients and chickens from Uganda, with a predominance of type II strains. The main difference was the discovery of multiple infections in 25% of the chickens.

The main difference was a somewhat higher proportion of type II strains in the human patients, but since all the multiple infections in chickens contained type II alleles a significant difference cannot be established, and from the current data there is no evidence that any particular *T. gondii* genotype is more likely to cause disease in HIV-patients. The pattern seen in Uganda, with a predominance of type II followed by type I and a small number of type III strains, is similar to what has been found through direct genotyping in European patients (Fuentes et al., 2001; Ajzenberg et al., 2002; Aspinall et al., 2003). However, the genotype distribution in Uganda is different from previous studies on chickens from Africa, Europe, USA, Middle East and Asia (results and references in Appendix 4), where types II and III were common and type I very rarely found. It is also very different from the distribution in chickens from Central and South America (Appendix 4), where there is a predominance of types I, III and atypical/recombinant strains, while genotype II is exceedingly rare. However, it is important to note that all the other chicken studies were performed on isolates only. Considering the isolation bias in this study, where all the type I strains were lost, it is possible that some of the previous genotyping studies from outside South America underestimate the true prevalence of type I strains, which may be more widely distributed than previously thought.

3.4.4 High prevalence of multiple infections provides opportunities for recombination

An interesting finding in this study was that five out of 20 chickens appeared to be infected by two or three genetically distinct strains of *T. gondii*. Experiments in mice have shown that previous infections only confer partial protection towards reinfection with strains of other lineages (Araujo et al., 1997; Dao et al., 2001), but multiple infections in naturally infected hosts have only been reported occasionally. One UK study detected simultaneous infections with type I and II in 6/19 samples (mostly solid tissue, but also blood and cerebrospinal fluid) from congenitally infected infants and immunosuppressed patients through SAG2-amplification and sequencing, as indicated by double peaks at polymorphic sites (Aspinall et al., 2003). In a study of chickens from Nicaragua, four out of 48 successful bioassays resulted in mixed genotype patterns in strains obtained in different mice infected with material from the same chicken (Dubey et al., 2006c); similar to what was found in the current study. These results suggest that multiple infections are relatively common, but the low detection frequency may be due to the fact that most studies employ methods which are designed to detect and genotype a single strain per host, i.e. PCR-RFLP without sequencing or isolation where only one mouse per group is used for typing.

Considering the high prevalence of infection in chickens and the presence of multiple parasite genotypes within the relatively small geographic areas of this study, it is not surprising that some individuals have been exposed to more than one parasite genotype during their lifetime. Consistent with the increased risk of exposure with time (Figure 3-6), it was revealed that all chickens with multiple infections belonged to one of the older age categories, (14-17 months or ≥18 months), and both the individuals with triple infections were ≥ 18 months old.

The presence of multiple *T. gondii* strains in a single chicken conveys an excellent opportunity for recombination if the host is consumed by a cat. Inter-strain recombination can only occur if the definitive host is co-infected within a short time frame, and immunity to oocyst reshedding is developed within 6 weeks (Freyre et al., 2007), therefore ingestion of multiple strains from an intermediate

host provides the best possibility for the successive generation of recombinant progeny in the feline host. Recombinant progeny can exhibit new phenotypic properties and in some cases be more virulent than either parent (Grigg et al., 2001a), and the high level of multiply infected chickens discovered in this study could potentially prompt the creation of a multitude of novel recombinant strains.

3.5 Conclusions

Nearly half of the Ugandan chickens had serological evidence of *T. gondii* infection and, like in HIV-patients from the same region, strains of genotype II were the most prevalent followed by type I. Typing of five loci on four different chromosomes and generation of over 15 kb of sequence revealed only five novel SNPs, indicating that African strains, unlike those found in South America, are highly similar to the archetypal lineages predominant in Europe. However, a high number of multiple infections were found which provides ample opportunities for simultaneous infections in cats and successive generation of recombinant strains that may be transmitted to humans. The genetic composition of the individual strains in these multiple infections could not be clarified from mouse tissues, which lead on to the work presented in Chapter 4, where the isolates are adapted to in vitro culture for thorough genotypic and phenotypic characterization.

Chapter 4 Characterization of a natural recombinant *Toxoplasma gondii* isolate from Uganda reveals an intermediate phenotype *in vivo.*

4.1 Introduction

T. gondii has an unusual population structure, where a small number of highly successful strains have expanded clonally through self-fertilization and asexual propagation to occupy large environmental niches. The most well-documented lineages are genotypes I, II and III, which are highly predominant in Europe and North America and, as shown in Chapters 2 and 3, these are also present in humans and chickens in Uganda. Sexual recombination between strains appears to be relatively infrequent in *T. gondii*, or at least very few progeny survive to tell the tale, but when the isolates from Ugandan chickens were secured *in vitro,* one natural recombinant between type II and III was found. The following chapter explores genotypic and phenotypic characteristics of this recombinant strain and closely related but non-recombinant type II and type III strains from the same source.

4.1.1 Sexual recombination in *T. gondii*

Sex, the combination of genes from more than one source in a single cell (Xu, 2004), is together with mutations and natural selection one of the major forces that drive evolution (Heitman, 2006). The first sex may have occurred among the first cellular life forms around 3.85 billion years ago, and been driven by selfish genetic elements like bacteriophages and transposons (Hickey, 1982). Eukaryotic organisms, which have probably existed for less than 1.5 billion years (Rivera, 2007), have developed a different mode of sex, which involves meiosis and gamete formation, and sexual reproduction is obligate in most multicellular organisms (Xu, 2004). Sexual recombination can benefit a population through the accumulation of beneficial mutations and removal of deleterious ones, and the greatly enhanced

genetic variation can be a major advantage for rapid adaptation to a changing environment and evasion of infectious diseases (Xu, 2004). However, there is also a risk of breaking up favourable gene combinations and there is often a cost involved in the process of finding a mate, and for unicellular organisms the sexual process is often much slower than asexual replication (Agrawal, 2006). Several parasitic and fungal pathogens exhibit largely clonal population structures even though they remain capable of sexual reproduction (Heitman, 2006), and predominantly asexual reproduction with infrequent sexual contributions may provide the optimal combination for rapid spread of successful clones while retaining the ability of efficient adaptation to environmental changes. In the case of *T. gondii*, experimental crosses (Pfefferkorn and Pfefferkorn, 1980; Sibley et al., 1992) have shown that a single mating opportunity between two strains in the definitive host can result in a large number of new strains with altered phenotypic properties, yet a very limited number of clonal types are highly overrepresented in nature (see Chapter 1.3), indicating a mainly clonal dispersal with limited sexual recombination (Grigg and Sundar, 2009).

The details of the genotyping of progeny from three experimental crosses (one between a type I and III and two between types II and III) have been made publicly available by LD Sibley at Washington University of St Louis, through the *Toxoplasma* Genome Mapping Database (http://toxomap.wustl.edu/interface.htm). In total, 40 type II/III progenies have been typed at 132-134 loci, but the markers were unevenly distributed and chromosomes VI and XI were only typed at a single locus. Thirty-five progeny from the type I/III cross have been genotyped at 177 markers (6-25 per chromosome), and genetic recombination was detected at a high level in all fourteen chromosomes with this close typing (see Table 4-1).

Table 4-1. Summary of recombination patterns in progeny from experimental crosses.

Chrom	Length (Mb)	Cross between II/III (N=40) No of loci[a]	II[b]	III[b]	rec[b]	Cross between I/III (N=35) No of loci[a]	I[b]	III[b]	rec[b]
Ia	1.89	3	40%	38%	**23%**	6	9%	29%	**63%**
Ib	1.95	3	45%	55%	**0%**	6	6%	46%	**49%**
II	2.30	7	50%	45%	**5%**	5	46%	14%	**40%**
III	2.47	6-7	73%	25%	**3%**	6	9%	34%	**57%**
IV	2.57	12	30%	30%	**40%**	15	37%	14%	**49%**
V	3.14	13-14	50%	28%	**23%**	6	17%	31%	**51%**
VI	3.60	1	28%	73%	**0%**	9	31%	20%	**49%**
VIIa	4.50	21-22	15%	68%	**18%**	21	26%	11%	**63%**
VIIb	5.02	6	45%	48%	**8%**	9	23%	20%	**57%**
VIII	6.92	11	38%	23%	**40%**	17	26%	20%	**54%**
IX	6.38	14-15	10%	75%	**15%**	16	3%	34%	**63%**
X	7.41	23	13%	60%	**28%**	12	20%	17%	**63%**
XI	6.57	1	83%	18%	**0%**	25	43%	0%	**57%**
XII	6.87	10	30%	18%	**53%**	24	11%	29%	**60%**

[a]No of genotyping markers tested at each chromosome (PCR-RFLP).
[b]Percentage of progeny containing chromosomes of type I/II/III or recombinant chromosomes.
The table is a compiled summary of the raw data available at http://toxomap.wustl.edu/interface.htm.

Recombination can occur through assortment of chromosomes, cross-over within a chromosome or a combination of both these processes. Chromosome sorting is present in all the progeny listed in the *Toxoplasma* Genome Mapping Database, where the number of chromosomes inherited in their entirety from either parent ranged from one to ten. Furthermore, all but five strains had at least one case of chromosomal cross-over. In all the 75 recombinant progeny, the average number (± standard deviation) of recombinant chromosomes was 3.5 (±2.2), with a higher degree found in the more closely typed I/III cross (Figure 4-1). A maximum of 10 recombinant chromosomes was found in one strain, and there were also a small number of cases with up to four cross-over events within the same chromosome, suggesting gene conversions (Khan et al., 2005b). This analysis of the data from experimental crosses shows that *T. gondii* strains undergo sexual recombination with a high degree of inter- and intra-chromosomal recombination and emphasizes the need for high density genotyping for accurate genetic characterization of recombinant strains.

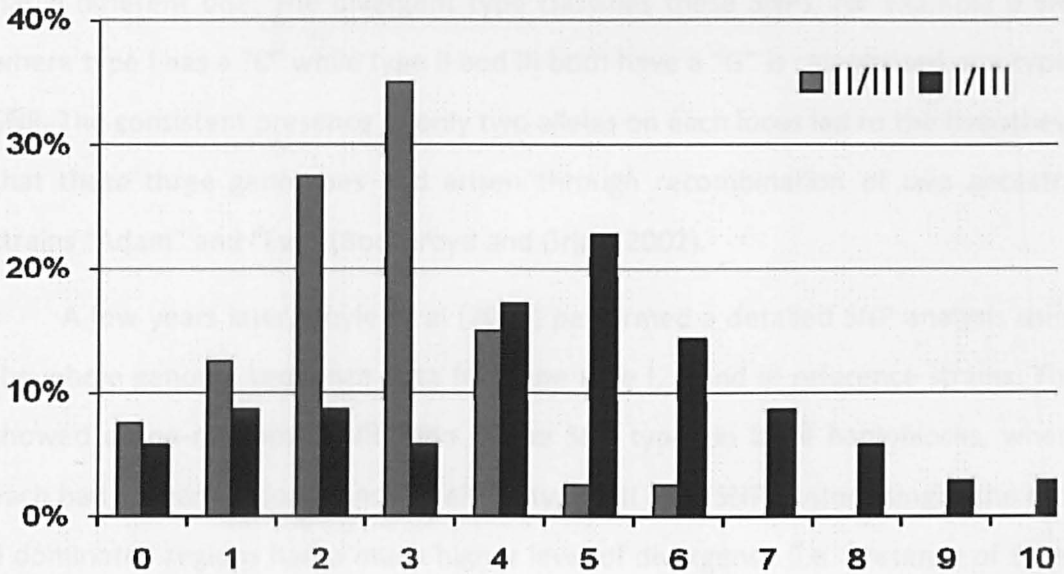


Figure 4-1. Prevalence of recombinant chromosomes in experimental *T. gondii* crosses. In the crosses between genotypes II and III most progeny had two or three chromosomes containing alleles from both parental strains (grey). The maximum number of recombinant chromosomes among these strains was six (one strain), while three strains had no evidence of chromosomal recombination. In the cross between I and III (black) the number of recombinant chromosomes was higher, which may be due to the higher density of genotyped loci. Most I/III recombinant strains had between four and six recombinant chromosomes, but one had as many as ten and two had none. The raw data were retrieved from the *Toxoplasma* Genome Mapping Database (http://toxomap.wustl.edu/interface.htm) and analyzed in Excel.

4.1.2 Origin of three clonal lineages I, II and III

The global population structure of *T. gondii* is more complex than three clonal lineages, and the classification into type I, II and III is of little use in South America (Sibley and Ajioka, 2008). The designation is, however, of use for classification of the majority of strains from Europe, North America and probably also for parts of Africa, and these lineages have been found in many hosts and locations across the globe. Where appropriate, this categorization is very useful since these three lineages have been extensively studied at both the genotypic and phenotypic level. Representative strains from all three lineages (GT1 (I), Me49 (II) and VEG (III)) have had their full genome sequenced and annotated (see Chapter 5.1), and additional transcriptome and proteome data is also available in ToxoDB and detailed sequence analysis have provided a good understanding of the origin of the three archetypal lineages, I, II and III.

One interesting feature of the three lineages is the bi-allelic SNP pattern: for practically all SNPs, two of the three strains share one allele while the third strain

has a different one. The divergent type classifies these SNPs, for example a SNP where type I has a "C" while type II and III both have a "G" is categorized as a type I SNP. The consistent presence of only two alleles on each locus led to the hypothesis that these three genotypes had arisen through recombination of two ancestral strains "Adam" and "Eve" (Boothroyd and Grigg, 2002).

A few years later, Boyle et al (2006) performed a detailed SNP analysis using the whole genome sequence data from the type I, II and III reference strains. This showed a non-random distribution of the SNP types in large haploblocks, where each had a clear predominance of either type I, II or III SNPs. Interestingly, the type II dominated regions had a much higher level of divergence (i.e. presence of type I and III SNPs) than the regions where type I or type III SNPs were predominant. This pattern was used to make predictions for the origin of these lineages, and a model with recombination and back-crosses between five ancestral strains; II_1, II_2, II_3, α and β, was proposed (Figure 4-2). According to this model, the three type II strains were very closely related, while α and β were distinct, but closer to each other than to the type II strains). Briefly, the model proposes that modern day type I derive from a cross between II_1 and α plus a backcross of II_1 with the F1 progeny, while type II descends directly from II_2 and type III is a recombinant of II_3 and β (Boyle et al., 2006). This model was found to be highly relevant for the SNP distribution found in the recombinant strain from Uganda, and is discussed in more detail in Chapter 5.

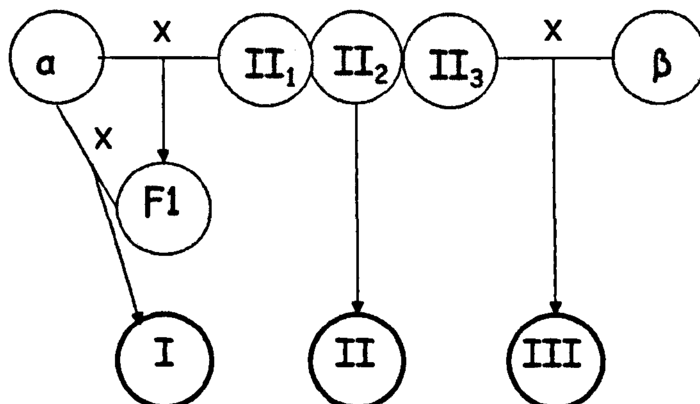


Figure 4-2. Schematic of the cross that gave rise to the canonical three lineages I, II and III. Adapted from Boyle et al., 2006. The three type II strains were close relatives, while α and β were closer to each other than to the type II strains.

Phylogenetic analysis of >270 kb of published sequence from atypical strains as well as strains within the I, II, III lineages led to the proposition that the most recent common ancestor (MRCA) of all modern *T. gondii* strains (including the parental strains that gave rise to types I, II and III) existed around 150,000 years ago (Morrison, 2005). The actual crosses that yielded the clonal lineages occurred later, but the timing of this is under debate. Grigg and Suzuki suggested two hypotheses "intermixing early" and "intermixing late", where the first theory suggests that the crosses occurred relatively soon after the time of the MRCA of modern strains but the rapid clonal expansion was delayed, the second hypothesis states a recent recombination event around 10,000 years ago with immediate expansion (Grigg and Suzuki, 2003). "Intermixing late" was supported by intron sequence comparisons between clonal and "atypical" strains (Su et al., 2003), while Morrison's reanalysis of this and additional data supports the early recombination theory, with the cross taking place around 100,000 years ago (Figure 4-3).

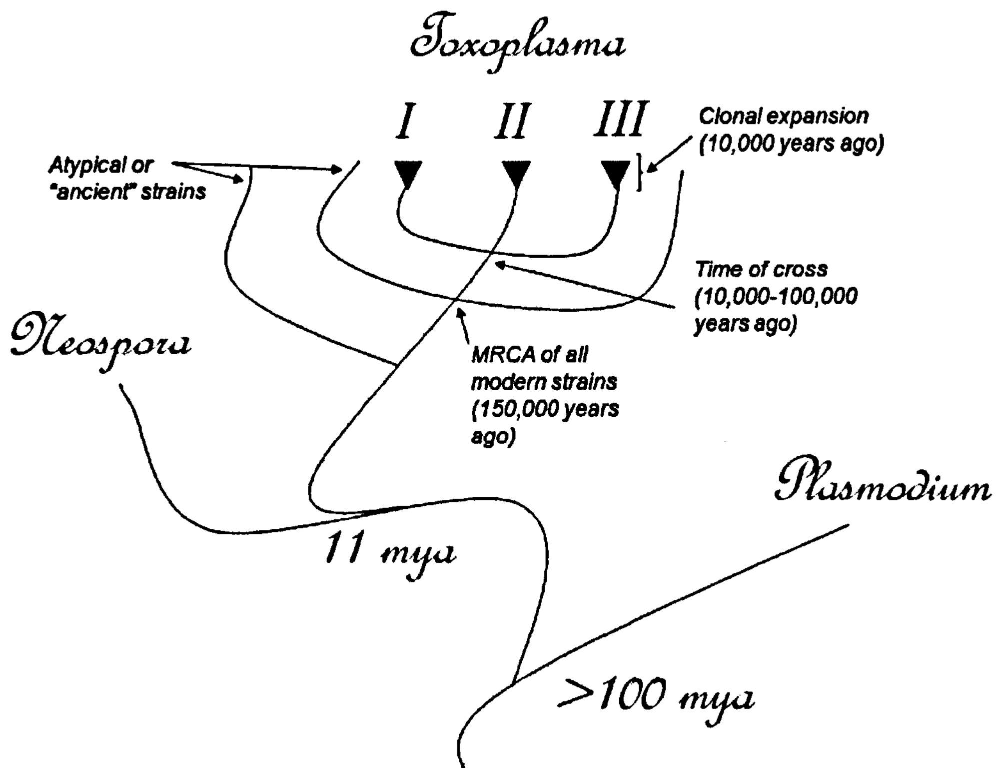


Figure 4-3. Overview of *T. gondii* divergence over time. *Toxoplasma* has existed as a genus for around 11 million years, and probably originates from Eurasia, where modern felines evolved (O'Brien et al., 2008). Most modern strains exhibit relatively little divergence and have gone through a genetic bottle neck some 150,000 years ago (Morrison, 2005), but ancient strains may remain in areas where they have not been outcompeted. A few crosses within the past 100,000 years created the highly successful lineages I, II and III, and these populations are believed to have undergone rapid expansion which is correlated to the domestication of animals around 10,000 years ago.

The three lineages appear to have undergone a rapid expansion correlated in time with the domestications of animals around 10,000 years ago, and it is plausible that the new situation with many species including cats living close together provided optimal conditions for a subset of the circulating strains, leading to the predominance of the clonal types seen today (Su et al., 2003).

4.1.3 Global population structure

In a worldwide population survey initiated by Prof. J. P. Dubey and Dr. Chunlei Su, the aim is to collect strains from a range of hosts and geographic location, and genotype all parasites with a set of 11 PCR-RFLP markers. More than 950 strains have been investigated so far, representing 131 different allelic combinations. Interestingly, most strains do belong to one of the archetypal lineages, I, II or III, and a smaller but significant number belong to a handful of other clonal types (for example the Brazilian lineages Br I-IV (Pena et al., 2008), or differ from these lineages on the apicoplast locus only, while only one single representative strain has been identified for most of the new multi-locus genotypes (Su, Dubey et al,, unpublished). A comparative study using a subset of these isolates was performed by Lehmann et al in 2006, when microsatellite analysis of seven loci were used for genotyping of 275 isolates originating from chickens from five continents. Four major haplogroups were found: one with a world-wide distribution (type I), one genotype corresponding to types II and III that was found everywhere except South America, and two genotypes that were confined to South America. The authors then hypothesized that *T. gondii* had existed for a long time as a rare parasite in the wild felines in South America, which increased in prevalence after the introduction of the domestic cat and spread to other continents through human activities (Lehmann et al., 2006).

However, another study published the following year reported no less than 11 haplotype groups among 46 strains from Europe, North America and South America through sequence analysis of eight intronic regions from five genes spanning >3.7 kb (Khan et al., 2007), and all these haplogroups appeared to derive from four ancestral strains. While two of these haplogroups were found on all three

continents, the archetypal I, II and III were confined to Europe and North America, while the remaining six types were exclusively found in South America. Their analyses reached the conclusion that the parasite was introduced in South America when the big cats first crossed the Panamian land bridge about 1-2 million years ago and that the separation between North and South American strains began by that time. Furthermore, they linked the ability of clonal expansion with direct oral transmission, which was correlated with the presence of a monomorphic version of chromosome Ia (Khan et al., 2007). It thus appears that *T. gondii* has a much more diverse population structure than previously thought, but it also appears that strains which possess advantageous genetic combinations for the niche(s) they live in can undergo rapid clonal expansion and outcompete less successful siblings.

4.1.4 Aims

The extent of *T. gondii* sexual recombination in nature is unknown, but it is clear that a single recombination event can exert a dramatic effect on the population structure and transmission patterns of this parasite. Among the strains isolated from Ugandan chickens (Chapter 3), there were representatives from both the type II and type III lineages, but also one unusual strain that was found to contain alleles typical of both type II and III. The objectives of the work described in this chapter was to assess the level of divergence on the genetic level among the Ugandan isolates with particular reference to the recombinant strain, in order to see if this isolate was likely to be a recent recombinant or belong to an ancient lineage, and also to investigate whether the recombinant strain had any unusual phenotypic properties in relation to the clonal representatives.

4.2 Materials and methods

4.2.1 *Toxoplasma gondii* strains

Nine *T. gondii* isolates obtained in mice at USDA and cryopreserved (described in Chapter 3) were thawed and injected into outbred, female Swiss-Webster mice at USDA by Dr. Natrajan Sundar. The animals were then euthanized six weeks post-inoculation and half the brain was cryopreserved (Dr. Sundar), while the other half was sent to SMI, where the rest of the mouse work and initial transfer to cell culture was performed. Isolate names were applied according to the Dubey lab system, which denotes the parasite (*Toxoplasma gondii* = Tg), the host species (Chicken = Ck), country of origin (Uganda = Ug) and isolate number. Thus, the nine isolates from chickens Ch1, 2, 17, 52, 68, 70, 79, 81 and 82 were named TgCkUg1-9. Originally, each chicken sample was inoculated into four mice (section 3.2.5), but due to the large number of samples handled at the Dubey lab, only one brain from each group was cryopreserved. If possible the tissue selected was from a source in which tissue cysts had been detected by microscopy and the cryopreserved brains that were used for the retrieval of isolates in this study originated from mice A1, A5, A8, A12, A20, A23, A26, A30 and A35 (Table 3.4)

Two laboratory strains of *T. gondii* were kept in culture at Leeds University and used as reference strains in the growth experiments; RH (type I) and Me49 (type II). DNA from the type III strains VEG (from USDA) and NED (from SMI) were used as reference strains for molecular work. The genomic data from GT1 (type I), Me49 (type II) and VEG (type III) that is available at ToxoDB, was utilized for sequence references.

4.2.2 Cell culture methods

Human foreskin fibroblasts (HFFs) were cultured in DMEM (Dulbecco's Modified Eagle's Medium with L-Glutamine, Invitrogen), supplemented with 10% FBS (heat-inactivated foetal bovine serum, GIBCO) and 1% PEST (10,000 U penicillin and 10 mg streptomycin/ml in 0.9% NaCl, SIGMA). During the first passages after

inoculation of parasites from mouse brains 1% Fungizone (250 µg/ml Amphotericin B, GIBCO) was added to prevent fungal contamination. The cells were grown at 37°C in a humidified incubator with 5% CO_2 in flasks with a vented lid or with loosened cap to allow gas exchange.

HFFs were maintained through subpassage every 4-7 days in T25 culture flasks and all the work was performed in a sterile flow hood. When the cells had formed a nearly confluent monolayer, the old medium was discarded and the monolayer was rinsed with sterile PBS. The cells were incubated with 2 ml Trypsin-EDTA (0.5%, Invitrogen) for 2-3 min at 37°C, until they detached after tapping the flask on the hood floor to dislodge cells, full detachment of cells was checked in an inverted microscope. Cells were then pelleted by centrifugation for 5 min and resuspended in 24 ml fresh medium, which was distributed into four new T25 flasks (6 ml/flask), or a 24 well plate (1 ml/well).

Freeze-down of HFFs and parasites was completed when the cells were in the exponential phase of growth. HFFs were trypsinized and centrifuged, as for sub-passage (described above), while the parasites were released from their host cells by rapping the flask on the hood floor and passing them through a 26-gauge needle. Cells were pelleted through centrifugation (4000 g, 10 min) and the pellet from each flask was resuspended in 1 ml cryomedium (80% DMEM, 10% FBS and 10% DMSO) and transferred to cryovials and stepwise frozen at -20ºC for 3-4 hours, -80ºC over night, liquid nitrogen for long-term storage.

4.2.3 Sub-passage in mice

Infected mouse brains were received from USDA on cold packs, and each of them was homogenised in 1 ml sterile PBS by repeated passage through a blunt 19 gauge needle[1]. 250 µl of the homogenate was inoculated directly in one NMRI mouse through the subcutaneous route, while the remaining 750 µl was digested in trypsin according to a procedure adapted from a published protocol (Dubey, 1998c).

[1] This was a larger needle than the one used for separation of tachyzoites from cell culture since the brain is a tougher material and this procedure aimed for homogenization, not release of parasites,

Briefly, brain homogenate was transferred mixed with an equal volume of 0.5% Trypsin-EDTA, incubated at 37ºC for 20 min, washed in 5 ml PBS, thoroughly mixed before pelleted through centrifugation at 2,000 g for 10 min. The supernatant was removed and the pellet resuspended in 250 µl PBS with 1% PEST giving a total volume of approximately 400 µl, which was inoculated into two NMRI mice. All mice were bled (by technical personnel at the animal house at SMI) for serology six weeks post-inoculation. The serology was performed using the Toxo-Screen DA as described in 2.2.3, at dilutions 1:25 and 1:250.

4.2.4 Isolation of *T. gondii* in cell culture from infected mouse brains

Infected mice were euthanized (by personnel at the animal house) through exposure to the anaesthetic isofluoran until dead[1], and thereafter dissected in a sterile flow hood. A small piece of brain was removed for DNA-extraction and the rest was homogenized in 1 ml of PBS through needle passage. 500 µl of brain homogenate was mixed with 10 ml cell culture medium and added directly to a confluent monolayer of HFFs in T25 flasks. The remaining 500 µl was subject to the trypsin digestion procedure described in 4.2.3 before the pellet was resuspended in 10 ml medium and added to a confluent HFF-monolayer[2]. The parasites were allowed 24 hours to invade before the monolayer was washed with sterile PBS to remove brain cells and 6 ml fresh medium was added. Once transferred to cell culture the isolates were propagated according to standard protocols, but as the time to consume a monolayer was several weeks compared with a few days for laboratory strains, interim medium changes were necessary approximately twice a week, and the removed media was used for quantification of free parasites.

---

[1] Procedures were different at the animal facilities at USDA and SMI.

[2] Trypsin digestion releases the bradyzoites from tissue cysts and cultures with trypsinized samples were found to be cleaner and the brain cells were easier to remove at the first media change.

4.2.5 *In vivo* growth assay

Small samples of brain, heart, lung and quadriceps muscle were dissected from each mouse in order to determine the parasite burden in these organs. Each tissue sample was weighed and the DNA was extracted using the QIAamp DNA kit (as in 2.2.4) and eluted in exactly 100µl. Five µl (1/20 of the total extraction) were used as template in each Q-PCR reaction (performed in duplicate according to the protocol described in 2.2.6). The parasite burden per gram of tissue was calculated using the following formula: ((Average number of parasites detected per Q-PCR-reaction)*20)/(Weight used for DNA-extraction (grams)).

4.2.6 Separation of multiple strains

Chickens 2, 68 and 70 had been found to be infected with several strains of *T. gondii* by sequencing and mouse bioassay (see Chapter 3.3.4). In order to secure the propagation of single clones in culture, dilutions were made on 24-well plates. Parasites were counted in a haematocytometer and a volume containing 10 parasites were added to the HFF monolayer in each well. The plate was then incubated for four days without being moved and thereafter the wells were inspected under an inverted microscope, and all wells with a single plaque were marked. The medium was then changed and the parasites were allowed growing for another week before the DNA-extraction was done for ten single-plaque wells from each sample. The DNA was then genotyped at the SAG3 and SAG2-5' loci as described previously (3.2.9).

However, no differences were seen between any samples originating from the same chicken. All samples from Ch-2 (TgCkUg2) and Ch-70 (TgCkUg6) were type III for SAG2-5' as well as SAG3, while Ch-68 (TgCkUg5) showed type II alleles for these loci. It thus appears that the multiple infections were lost in earlier passages and a single isolate was retrieved from each chicken.

4.2.7 *In vitro* growth assay

The growth during the very first cell culture passage was monitored for each strain in order to compare the efficiency of direct inoculation versus trypsin

digestion. In order not to impair the growth of the parasites, the measurements were performed on the medium as it was changed. DNA was extracted from the pellet from 5 ml of medium (out of a total volume of 6 ml), which was eluted in 100 µl and quantified by Q-PCR in the same manner as the *in vivo* growth assay (section 4.2.5).

The growth *in vitro* was also quantified at passage eight, when the strains had had some time to adjust to growth in cell culture, and the same method with quantification of parasites in 5 ml medium was used. This time, 100,000 parasites were added to each T25 flask, and medium was changed and used for extraction on days 2, 4, 7, 10 and 13 post-inoculation. Reference strains RH and Me49 were also included in this experiment.

4.2.8 Transmission Electron Microscopy

Parasites from nearly lysed monolayers were collected and pelleted in 1.5 ml tubes. Samples were fixed over night in 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer ($Na_2(CH_3)_2ASO_2.3H_2O$, pH 7.0). Following two subsequent washes in the same buffer they were stained for two hours in 1% osmium tetroxide. The samples were dehydrated in ethanol-water dilutions of increasing alcohol concentration (20%, 40%, 60%, 80% and 100%), washed with propylene oxide and embedded in Spurr's resin by overnight incubations in ascending concentration and finally allowed to solidify in 100% resin at 70°C.

The blocks were trimmed in a microtome by glass knife and thereafter cut to ultra-thin sections (80 nm) using a diamond knife. Sections were collected on plastic coated copper grids, stained with electron-dense metals (uranyl acetate for 2 hours and lead citrate for 20 minutes), washed and thereafter examined on a JEOL 1200EX transmission electron microscope. This work was done with the assistance of Mr. Adrian Hick at Leeds University.

4.2.9 Sequencing and sequence analysis

Five genomic loci with a high SNP density were identified using the 'Identify genes based on SNP-characteristics'-tool on ToxoDB (Figure 4-4).

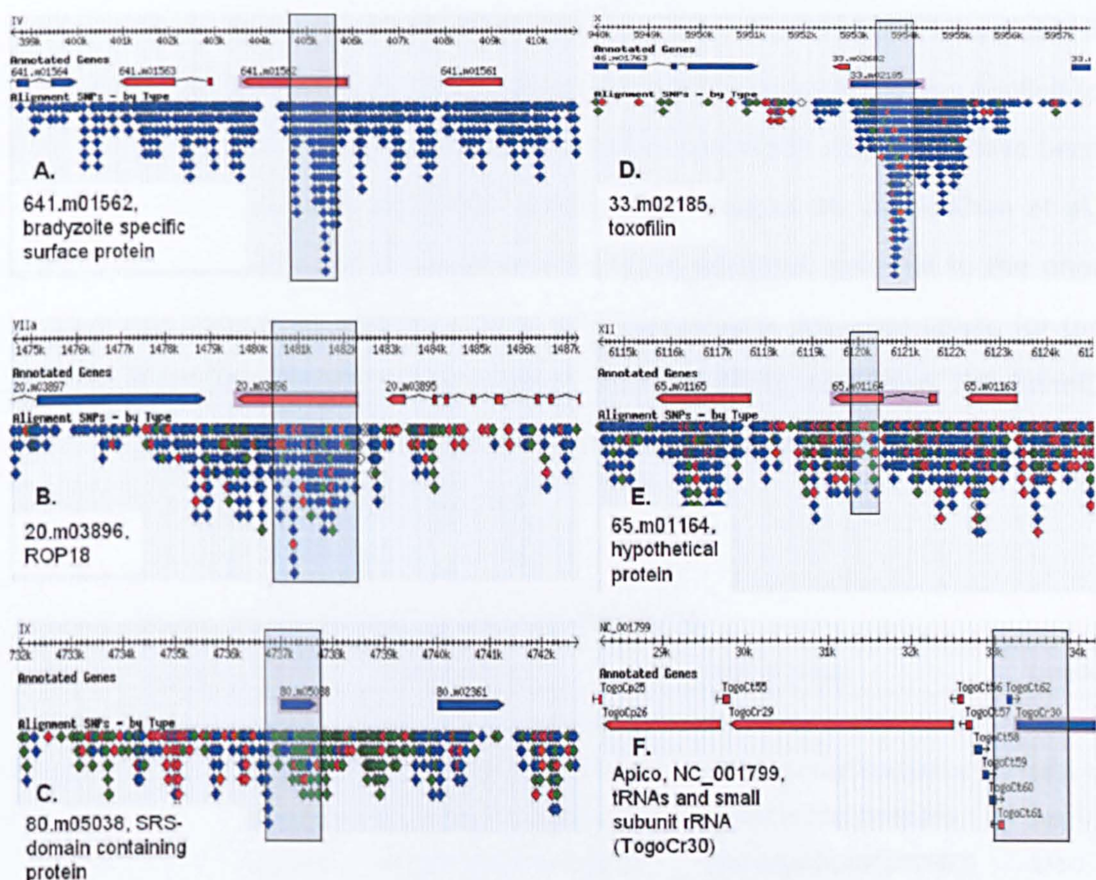


Figure 4-4. Polymorphic loci selected for sequencing. The sequenced loci are marked with boxes and the coloured diamond shapes indicate SNPs (red = type I, green = type II, blue = type III). The chosen regions are located on five chromosomes (IV, VIIa, IX, X, XII) and the apicoplast. No SNP data is available for the apicoplast and the locus was randomly chosen by the primer design tool.

Alignments for the selected loci of the three sequenced strains were compiled in BioEdit and used to find conserved regions, where the primers could be placed. The primer design tool at the website of University of Massachusetts Medical School (http://biotools.umassmed.edu/ bioapps/primer3_www.cgi), were used for selection of forward and reverse primers. The best option was chosen by manual controls of each pick: absence of hairpin loops or palindrome sequences was checked at http://www.cybergene.se/primer.html, melting temperature for all primers was adjusted to be approximately 55˚C, and a general comparison with all available *T. gondii* sequence was performed to avoid unspecific products

(http://www.toxodb.org/toxo/showQuestion.do?questionFullName=GenomicSequenceQuestions.SequencesBySimilarity).

Primers for the apicoplast were designed in the same manner, but from an arbitrary region as only one apicoplast sequence (from the RH strain) is available it is not possible to perform SNP comparisons. In addition, sequencing was applied to the UPRT-1 intron and two microsatellite loci (B17 and W35), since they have been utilized in other *T. gondii* population studies (Ajzenberg et al., 2005; Khan et al., 2005a). For the microsatellites new primers were designed, external to the ones used by Ajzenberg et al (2005), in order to achieve longer sequence reads; for the UPRT-1 intron the original primers by Khan, Su et al (2005) were used. The GeneID (from ToxoDB version 4.3), chromosomal location, primers and amplicon lengths for all sequenced loci are listed in Table 4-2.

Table 4-2. Genomic locations and primers used for sequencing.

GeneID[a]	Chrom	Start	Forward primer	Reverse primer	Length[b]
641.m01562	IV	404650	GCCTTGACAGTCACCAACAC	GTGAATGACAGTTTCCGTTGA	1037
20.m03896	VIIa	1480563	CCTGCTGTTCGAAGTTTTGC	GATACAGCCGTTGACAAAGC	1601-7
80.m05038	IX	4736806	GTGAGTAATTGCGGCAACG	TGTCACTCTCACTTTGCATG	949
33.m02185	X	5953214	CGCAATACAAGTCACGCCCT	TGAGGGAATGCGTCTGTGCC	874-7
65.m01164	XII	6119813	TGGAACAGGGCTGTGCTG	GACTCATTCGTGTCCCTTCC	631-4
Apico	Apico	33056	CAGCCGAAAGCTCTAACCAC	GCCATGCACCTCCTGTAAAT	904
W35	II	633228	GCTCGCAGAGGAGGTTCACT	AGTCATCCGTGCGTTAGCGT	196
UPRT-1	XI	2709204	TCCACAGGGCTTCTAAAAT	GAGTTGAGAACAGGCTTCAG	688
B17	XII	6452115	CCAGTTTCGCGAACAAGTGA	CATAGTTCAGTCTCGGGTGA	342

[a]Gene products are shown in Figure 4-4.
[b]Length indicates the actual good quality sequence length obtained for all isolates, which was used for comparative analysis. This does not always extend to the full length encompassed by the primer pairs.

Sequencing reactions were performed as described in 3.2.10, and all sequences were manually controlled and edited in BioEdit. All sequences from the same locus were truncated to the same length as the strain with the shortest good quality sequence, and the corresponding sequences for GT1, Me49 and VEG were retrieved from ToxoDB. Alignments and creation of Nexus-files were performed using ClustalX2 (http://clustalx2.software.informer.com/2.0/), the phylogenetic

trees were generated in MrBayes (Huelsenbeck and Ronquist, 2001) until P<0.01 and thereafter visualized in TreeView (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html). For the microsatellite sequences B17 and W35 all reference strains available in GenBank were retrieved and analyzed in the same manner together with the sequences from four Ugandan strains: TgCkUg2, 5, 6 and 8. The sequence data generated for these isolates has been submitted to GenBank (Accession numbers: FI274744-FI274973).

4.2.10 Random Amplification of Polymorphic DNA

In addition to the targeted genotyping efforts, random amplification of polymorphic DNA (RAPD) was used as a more general approach to detect genetic variation. In addition to the Ugandan strains, DNA from five reference strains were included, Me49 (type II), NED (type III), and three atypical strains: VAND, TONT and RUB (DNA from these strains were kindly provided by the BRC ToxoBS group through Dr. Benjamin Edvinsson at SMI). Two 10 nucleotide oligos (B12 and C20) were selected because they have been found to readily differentiate between virulent and avirulent strains (Guo et al 1997).

Eight additional random primers, four 10 nt and four 12 nt oligos, were randomly chosen with the help of a dice where 1=A, 2=C, 3=G and 4=T. Five of these were found to be suitable for differentiation between strains either alone or in pairs: 10.1: CTGGGGTCAC, 10.2: AAGCCTTCGA, 12.2: ACCCCCAATGGA, 12.3: ACTACCCCAGGA, 12.4: GTGAAGGTCGCT. The DNA template concentration was standardized and 10 ng was used in each reaction which was performed under the same conditions as for the PCR-RFLP reactions (3.2.9) with the exception of decreased annealing temperatures (34-40°C) and prolonged extension time (1 min 45s). The fragments were separated on 2% agarose gels and the absence or presence of bands was recorded and analysed using the Treecon software (Van de Peer and De Wachter, 1994) with the Nei & Li method, Neighbour joining tree topology and Bootstrap analysis of 2000 samples.

4.3 Results

4.3.1 Isolation in mice

Previously cryopreserved mouse brains infected with *T. gondii* from Ugandan chickens (Chapter 3) were thawed and passaged in mice at USDA, sent to SMI where they were inoculated in mice and transferred to cell cultures, which were later brought to Leeds where the molecular genotyping took place. An overview of the whole isolation procedure is shown in Figure 4-5.

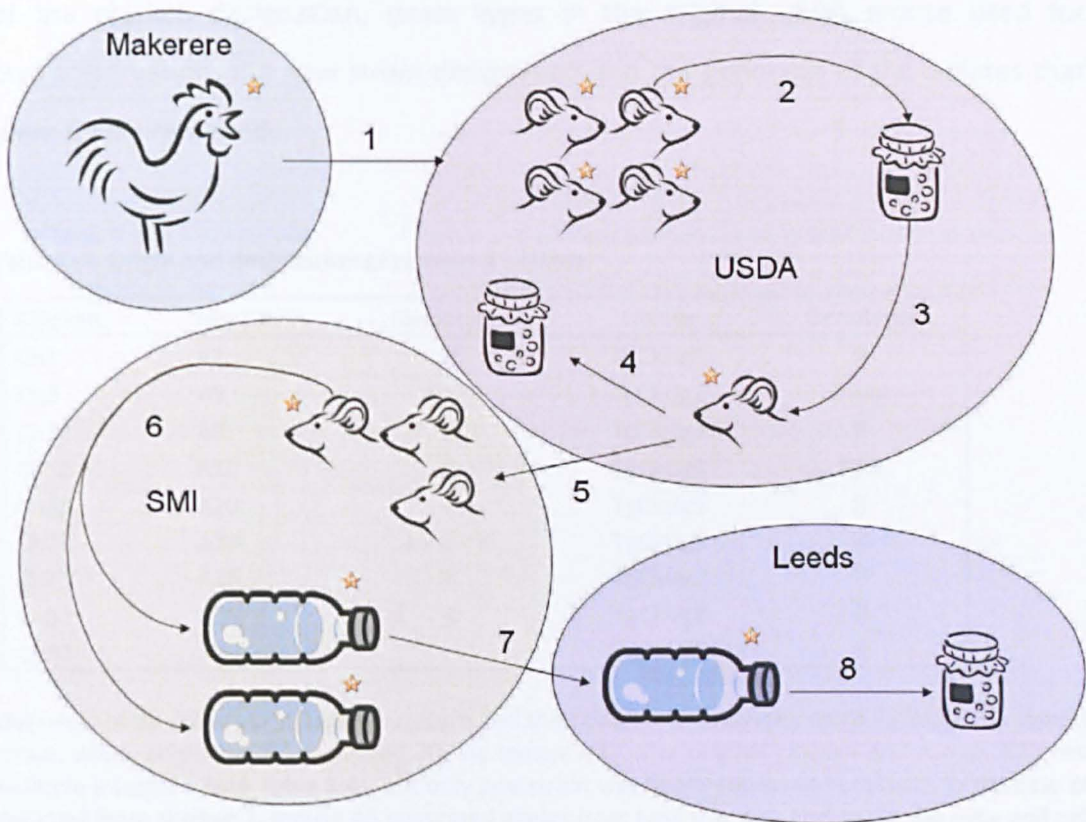


Figure 4-5. Overview of the isolate retrieval procedure. Coloured circles indicate where each step took place and the stars specify samples that were genotyped. **1.** Brain and heart tissues from 85 chickens were brought from Makerere to USDA and inoculated into four mice per chicken, as described in Chapter 3. **2.** Brain material from one infected mouse from each group was cryopreserved (preferentially those with visible tissue cysts). **3.** Later, these isolates were thawed and inoculated into one mouse per sample at USDA, and the infection was allowed to progress to brain cyst formation (6 weeks). **4.** Half the brain was cryopreserved at USDA and the other half shipped to SMI. **5.** At SMI, each sample was used to infect three mice, and the tissues from these mice were used for quantification of parasites in different tissues (the *in vivo* growth assay) **6.** Infections were allowed to progress to tissue cyst formation and thereafter brains were used for inoculation into cell culture, with and without trypsin digestion. Here the *in vitro* growth assay for passage 1 was performed. **7.** Isolates were brought to Leeds where subsequent cell culture work, including the *in vitro* growth assay for passage 8, and molecular typing was performed. **8.** All isolates are preserved in liquid nitrogen at the University of Leeds.

Eight of the nine isolates sent from USDA were successfully passaged in mice at SMI, but *T. gondii* DNA was not detected by nested PCR or by Q-PCR of the brain homogenate from the last strain (TgCkUg4 from Ch-52). Basic PCR-RFLP genotyping (SAG2, SAG3 and GRA6 as described in 3.2.9) was performed on DNA extracted from the mouse brains from USDA and the mice at SMI, to check which isolates had survived. All strains originally typed as II continued to type as plain type II strains, while the multiple infections appeared to have been decimated to one type II strain from Ch-68 (mouse A20), one type III strain from Ch-70 (mouse A23) and a recombinant type II/III strain from Ch-2 (mouse A5). See Table 4-3 for an overview of the chicken designation, strain types in the original USDA mouse used for cryopreservation, the new strain designation and the genotype of the isolates that were finally retrieved.

Table 4-3. Origin and designation of retrieved isolates.

Chicken	Mouse	Genotype	Isolate	Genotype
Ch1	A1	II	TgCkUg1	II
Ch2	A5	II + III	TgCkUg2	II + III
Ch17	A8	II	TgCkUg3	II
Ch52	A12	II	TgCkUg4	N/A
Ch68	A20	I + II	TgCkUg5	II
Ch70	A23	I + II + III	TgCkUg6	III
Ch79	A26	II	TgCkUg7	II
Ch81	A30	II	TgCkUg8	II
Ch82	A35	II	TgCkUg9	II

Overview of the isolates retrieved in culture and their origin. For example, strain TgCkUg6 is a type III strain, which originates from chicken 70, via mouse A23. The original chicken and mouse A23 had multiple infections (see Table 3-4), but only one strain was finally retrieved in culture. In the case of TgCkUg2 from chicken 2, mouse A5 possessed alleles from type II and III, and so did the mice and cell cultures at SMI, and this turned out to be a recombinant strain.

All infections with Ugandan *T. gondii* isolates were asymptomatic in mice; this was true for the first two passages in Swiss-Webster mice (USDA) as well as for the following passage in NMRI mice (SMI). Two conventional nested PCRs (SAG2-5' and SAG2-3') and two Q-PCRs (B1 and AF) were used to test for the presence of parasites in mouse brains from USDA. Six samples (TgCkUg1, 2, 5, 6, 7, 8) were positive by all methods, but the two samples with the lowest parasite numbers (TgCkUg3 and 9) were only detected by Q-PCR of multiple copy genes. The most

sensitive method was the Q-PCR of the highly repetitive AF-sequence (described in section 2.2.6), therefore this was the method of choice for detection and quantification in the *in vivo* and *in vitro* growth assays. TgCkUg4 was negative for all PCRs. An overview of these results is shown in Table 4-4.

Table 4-4. Detection and quantification of parasites and isolation success in mice.

Isolate	SAG2-5'	SAG2-3'	Q-PCR B1	Q-PCR AF	Brain homog (μg)	*T. gondii* /reaction	No of *T. gondii* / g brain	Seropos mice (direct)	Seropos mice (tryps)	Sub-passage (direct)
TgCkUg1	+	+	+	+	203	3	2,956	1/1	2/2	ND
TgCkUg2	+	+	+	+	298	105	70,470	1/1	0/2	2/2
TgCkUg3	-	-	-	+	5	1	40,000	1/1	2/2	ND
TgCkUg4	-	-	-	-	258	-	-	0/1	0/2	ND
TgCkUg5	+	+	+	+	263	4	3,042	1/1	2/2	ND
TgCkUg6	+	+	+	+	95	140	294,737	1/1	2/2	ND
TgCkUg7	+	+	+	+	213	27	25,352	1/1	2/2	ND
TgCkUg8	+	+	+	+	303	37	24,422	1/1	2/2	ND
TgCkUg9	-	-	+	+	319	2	1,254	1/1	2/2	ND

DNA was extracted from the amount of brain homogenate (brain in PBS, 1:10) remaining after inoculation of the brain material from USDA into mice at SMI. This was carefully weighed and from this amount, the DNA elution volume (100μl) and volume used as template in the PCR-reaction (5μl), the number of parasites per gram of each brain could be calculated based on the Q-PCR result per reaction. DNA from at least three parasites per reaction was necessary for detection with all markers.

At SMI, most of the mouse infections were successful, whether trypsin digestion was used or not. Seven strains caused infection in all three mice, but only one mouse got infected with the TgCkUg2 strain, and this was with the sample that had not been subject to trypsin digestion (Table 4-4). The brain from this infected mouse was homogenized and subpassaged into two new mice, which both got infected. As seen in Table 4.4 the parasite density in the brain of TgCkUg2 was high, so the lack of infectivity is more likely to be due to trypsin treatment sensitivity than low parasite numbers. TgCkUg4 did not infect any mice, but since the original brain sample from USDA was negative for all PCR-reactions it appears to have been uninfected or had a very low number of viable parasites to begin with.

4.3.2 Quantification of parasites in mouse organs

The density of *T. gondii* parasites in brain, heart, lungs and leg muscle from infected mice was determined through Q-PCR of the highly repetitive AF-fragment. This method revealed significant genotype-related differences in the parasite density and also detected a considerable tissue tropism (Table 4-5). The type III strain (TgCkUg6) produced the highest tissue burden in all organs; lungs, brain, muscle and heart contained 4, 30, 40 and 200 times higher parasite densities respectively compared with the type II average (significant for brain ($P<0.001$), heart ($P<0.002$) and muscle ($P<0.02$) using the t-test), while the recombinant type II/III strain (TgCkUg2) showed an intermediate phenotype for all organs investigated (Figure 4-6). In brain tissue, the average density as determined by Q-PCR was $4.5*10^6$ parasites per gram for type III, $1.2*10^6$ for the recombinant, and $1.5*10^5$ for the six type II strains. In heart tissue, the mean values for parasite density were $1.2*10^5$ (III), $8.9*10^3$ (II/III) and $6.2*10^2$ (II) parasites per gram. The parasite burden caused by the recombinant strain was significantly higher compared with type II strains ($P<0.03$ for brain, heart and muscle), while the difference between TgCkUg2 and TgCkUg6 did not reach significance.

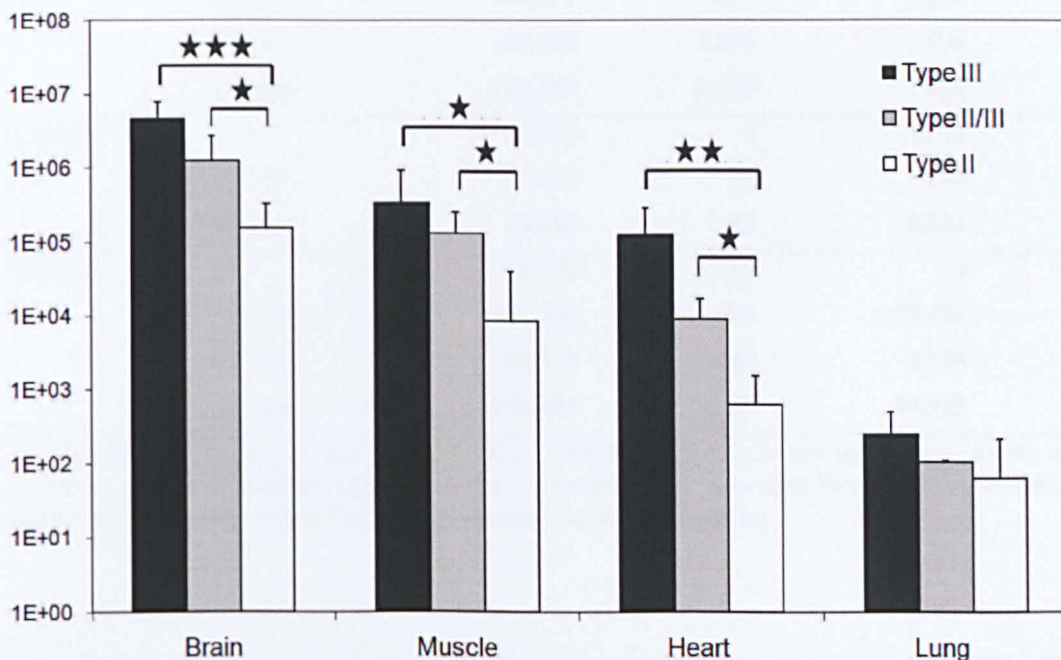


Figure 4-6. Parasite density per gram of infected mouse organs. The type III strain produced a high tissue density for all organs compared with all of the six type II strains, and the recombinant type II/III had intermediate values. All mice had the highest number of parasites in the brain. Significance levels are indicated in the figure, where one star means $P<0.05$, two stars $P<0.01$ and three stars $P<0.001$.

Table 4-5. Parasite density per gram of infected mouse organs

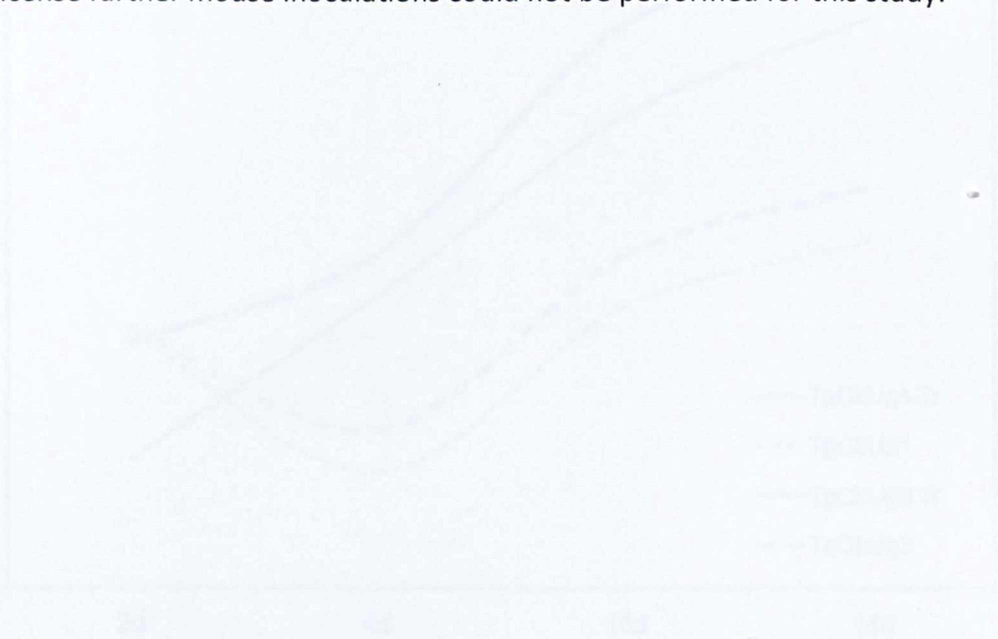
	Mouse[a]	Brain	Heart	Muscle	Lung
TgCkUg1	m34	27,383	78	1,911	5
(II)	m35 (tr)	354,404	160	0	38
	m36 (tr)	17,855	52	597	14
	average	**133,214**	**97**	**836**	**19**
TgCkUg2	m6	168,779	14,594	43,914	108
(II/III)	m8	2,285,529	3,327	217,195	103
	average	**1,227,154**	**8,960**	**130,554**	**106**
TgCkUg3	m38	12,690	16	2	0
(II)	m47 (tr)	234,607	140	67	65
	m48 (tr)	136,235	566	1	10
	average	**127,844**	**240**	**23**	**25**
TgCkUg5	m41	92,615	39	0	0
(II)	m53 (tr)	16,242	1,150	0	27
	m54 (tr)	541,394	276	1	0
	average	**216,750**	**488**	**0**	**9**
TgCkUg6	m40	7,697,924	306,000	0	244
(III)	m51 (tr)	4,614,821	56,553	1,006,725	487
	m52 (tr)	1,252,527	7,181	6,224	2
	average	**4,521,757**	**123,245**	**337,650**	**244**
TgCkUg7	m42	8,007	0	146	0
(II)	m55 (tr)	444,171	853	2,122	0
	m56 (tr)	82,689	3,280	206	597
	average	**178,289**	**1,378**	**825**	**199**
TgCkUg8	m43	15,503	0	5,703	0
(II)	m57 (tr)	51,014	86	559	2
	average	**33,258**	**43**	**3,131**	**1**
TgCkUg9	m44	61,775	1,597	0	41
(II)	m59 (tr)	433,236	1,761	129,121	0
	m60 (tr)	60,024	483	3,548	262
	average	**185,012**	**1,280**	**44,223**	**101**

[a] Three mice per sample, except TgCkUg2, where only the two mice in the second inoculation were used (see Table 4-4), and TgCkUg8, where the third mouse had been used for cell culture inoculation before this test was initiated. Trypsinized samples are indicated by (tr).

The brain was the only organ where parasites were detected in 100% of the seropositive mice. Regardless of genotype, all isolates showed a similar tissue tropism, with the highest affinity for the brain, although the actual numbers varied

from nearly eight million parasites per gram in mouse m40 (TgCkUg6) to just eight thousand in mouse m42 (TgCkUg7). The second highest abundance was found in the heart of some mice and the muscle of others, but the variation was large and more than one in three mice (8/22, five different isolates) had undetectable or very low parasite levels (0-2 parasites/gram). This can probably be explained by an uneven distribution of tissue cysts within muscle. Hearts of all but two mice were found to harbour *T. gondii* parasites, but the relative level compared with the brain was only about 1:50. The lowest density was detected in lung tissue, where the maximum density was around 600 parasites per gram, less than one per cent of the level found in the brain of the same mouse (m51).

Since the variation between mice in the same group was substantial and the number of mice was small it is difficult to make a ranking of the type II isolates. The highest densities were found in the brains of three mice infected with TgCkUg5, 7 and 9, all of which had over 400,000 parasites per gram. A larger experiment with known inoculation loads is required to draw definitive conclusions about differences in growth rate between the type II strains *in vivo*, but due to the nature of the animal license further mouse inoculations could not be performed for this study.



4.3.3 Isolation and growth in cell culture

When the isolates were transferred to cell culture, homogenized brain tissue was either transferred directly to HFF monolayers or treated with trypsin for 30 minutes and washed in PBS before inoculation. Growth data was collected by Q-PCR quantification of free parasites in the medium (except TgCkUg8, which was processed earlier) during their very first passage in cell culture. Parallel comparisons with and without trypsin digestion for each sample showed that both methods were successful for all isolates, but that trypsinization greatly enhanced the growth rate. A growth characteristic observed in the samples which were not treated with trypsin (and for all Ugandan samples in subsequent cell culture passages) was an extended lag-phase. The number of free parasites decreased to very low levels around day four and never recovered to reach the same levels as the samples that had been subject to trypsin digestion (Figure 4-7).

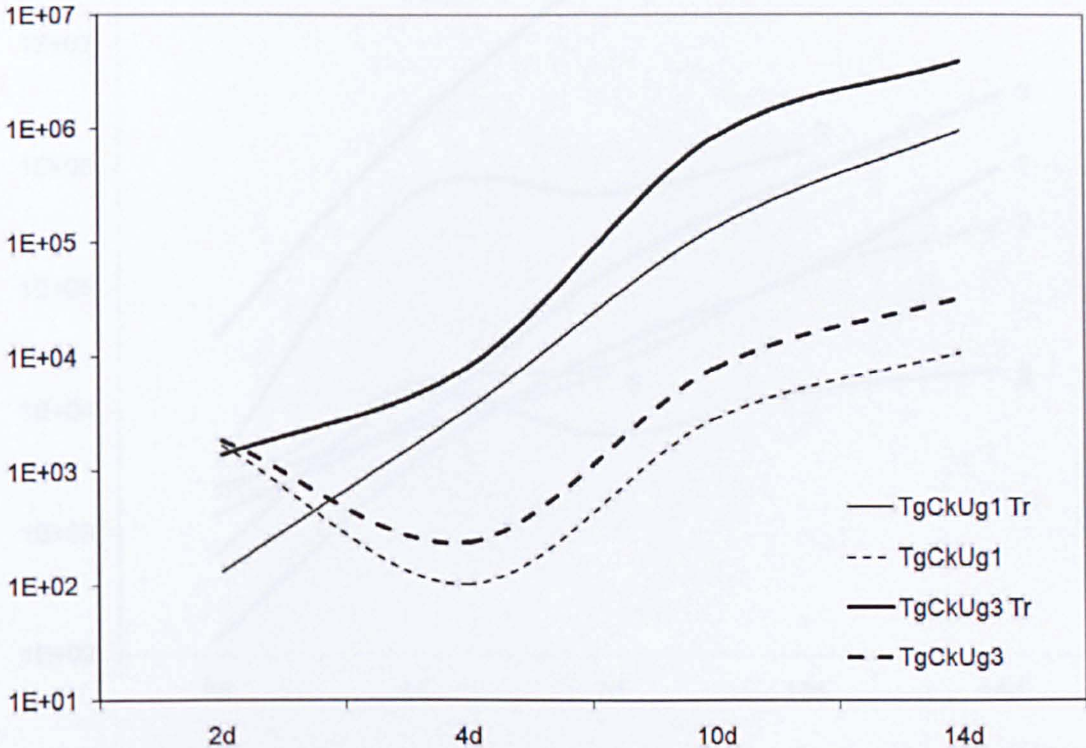


Figure 4-7. Growth rate comparison *in vitro*, between directly inoculated versus trypsinized parasites. The graph shows the average number of free parasites per ml of medium (average from two Q-PCR replicates) for strains TgCkUg1 and 3 with direct inoculation and after trypsin digestion ('Tr'). Trypsin digestion resulted in enhanced growth rate for all parasite strains.

The difference in strain growth rate was remarkable in the first passage, where the type III strain (TgCkUg6) had reached 2.1 million parasites per ml after four days, while the type II strain range by that time was between 3,500-16,500 parasites per ml for the trypsinized samples. TgCkUg2 had an intermediate value of 580,000 *T. gondii*/ml. In this initial culture the quantity of parasites in the inocula was directly dependent on the parasite density in mouse brain, thus the growth advantage seen with TgCkUg2 and 6 (Figure 4-8) may be correlated to this. Among the type II strains, the distinction between the faster growing TgCkUg1, 3 and 7 and the slower TgCkUg5 and 9 was observable from the start and remained throughout the months of culturing. Interestingly, this was not correlated with the growth in vivo, where mice infected with TgCkUg5 and 9 had high parasite densities in the brain.

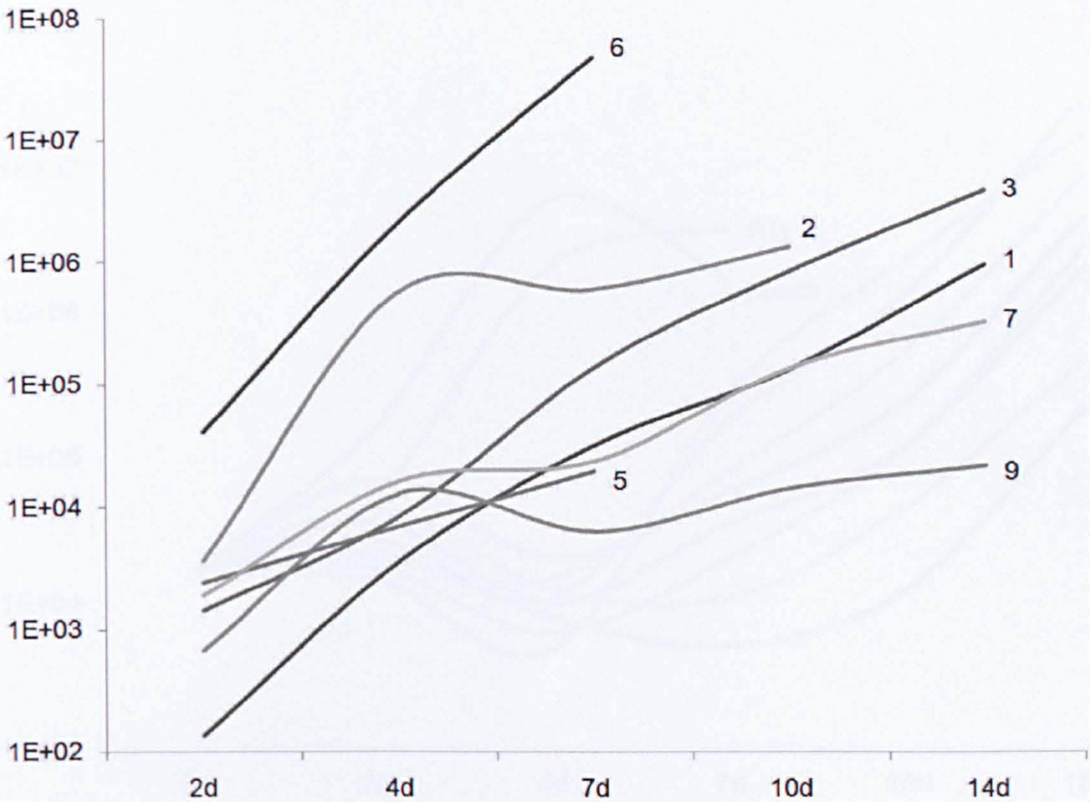


Figure 4-8. Growth during the first passage in culture (trypsinized samples). The strains that had the highest parasite density on day 2 (the type III TgCkUg6 and the recombinant TgCkUg2) also had the highest densities on days 4 and 7. Among the type II strains, which all started with a lower density, strains TgCkUg1, 3 and 7 had a good growth rate, while TgCkUg5 (data for days 2, 4 and 7 only) and TgCkUg9 were more difficult to culture.

During the eighth passage in cell culture, a controlled growth experiment was initiated where 100,000 parasites were added to each flask (approximately 16,600 per ml medium). The genotype related growth bias seen in mice was not consistent in passage eight, suggesting different selective pressures. There was considerable variability among the Ugandan strains, but the fastest and slowest growing strains *in vitro* all belonged to genotype II, while the recombinant and type III strains were found in the middle of the spectrum. As shown in Figure 4-9, even the fastest growing Ugandan strains experienced a lag phase around day 4 and they were all far behind the established lab strains RH and Me49. Based on their growth rate in culture, the Ugandan strains could be divided into three groups: the fast growing strains TgCkUg3 and 8, the intermediate ones: TgCkUg1, 2, 6 and 7, and the slow growing TgCkUg5 and 9.

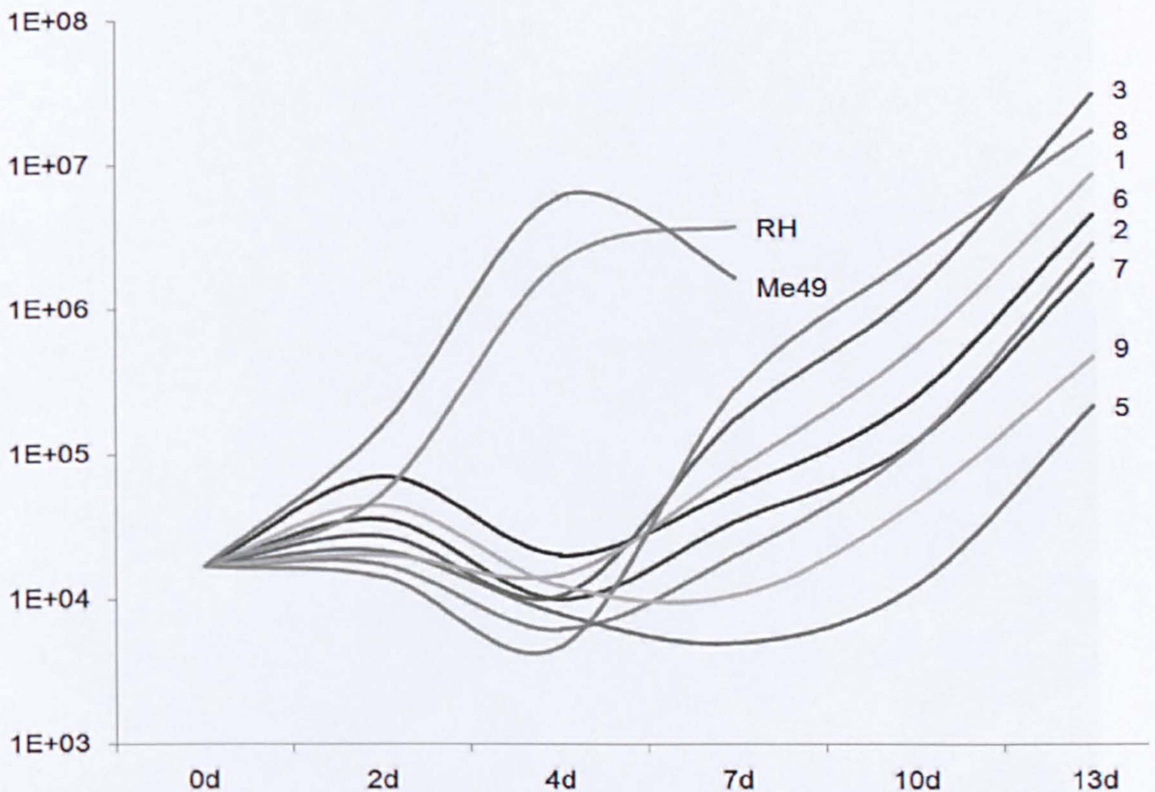


Figure 4-9. Growth during the eighth round of cell culture. RH and Me49 quickly destroyed the monolayer and few HFF cells remained around day 4, while all Ugandan strains experienced a lag phase and slow growth at this time. The type III (TgCkUg6, '6') and the recombinant strain (TgCkUg2, '2') had intermediate growth rates compared with the other strains from Ugandan chickens. TgCkUg5 ('5') and TgCkUg9 ('9') were consistently more slow-growing *in vitro* compared with the other type II strains.

4.3.4 Morphology

Transmission electron microscopy (TEM) was performed to examine the morphology of Ugandan *T. gondii* isolates from tissue culture (Figure 4-10). Samples were taken from TgCkUg6 (III), TgCkUg8 (II), TgCkUg2 (II/III) and Me49 (II, ref) after seven days in culture (four days in the case of Me49), and processed as described in 4.2.8. For the Ugandan parasites, a high number of polysaccharide storage granules and micronemes were observed in many parasites and, in addition, the wall of the parasitophorous vacuole was thickened and the rhoptries were dense. These features are typical of bradyzoite development within tissue cysts. It thus appears that the Ugandan parasites spontaneously convert to bradyzoites to some extent under standard cell culture conditions.

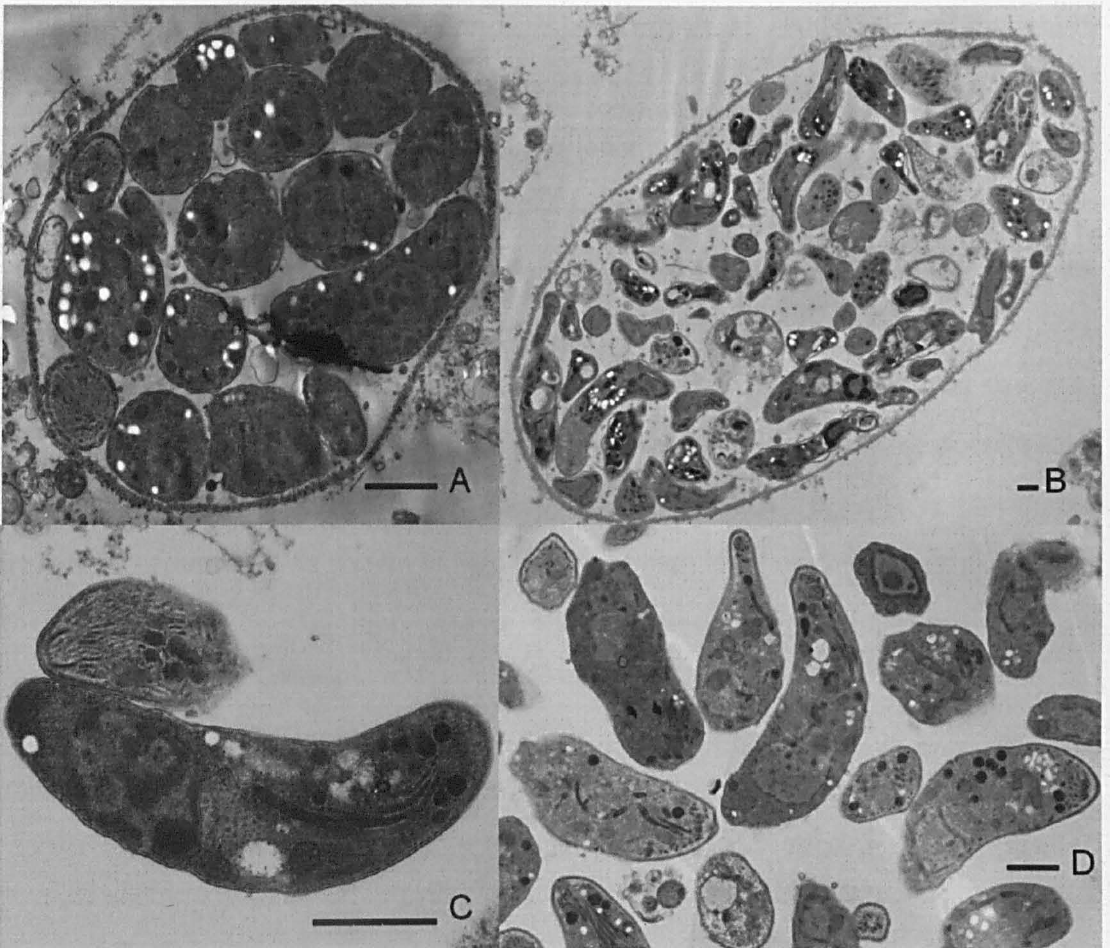


Figure 4-10. TEM pictures of Ugandan isolates and the Me49 reference strain. A & B show parasites (TgCkUg6 (A) and TgCkUg8 (B)) inside vacuoles/cysts. The thickness and relatively loose fit of the membranes indicate that these are bradyzoite tissue cysts. The large vacuole in B contains parasites with odd shapes, which may be dying. **C** shows one and a half parasite of the TgCkUg2 strain. Note the posterior position of the nucleus, another sign of bradyzoites. **D** shows free parasites of the Me49 strain. Scale bars represent 1 µm.

In some instances parasites inside vacuoles appeared to be dead or dying (note the odd shapes among parasites in Figure 4-10 B), which may be another reason why their growth was delayed. In contrast, the Me49 strain, which is well adapted to cell culture, consistently consumed the HFFs within a few days and no cysts or vacuoles were found at day 4, but instead there were many free parasites (Figure 4-10 D).

4.3.5 Microsatellites and sequence analysis

Microsatellite analysis is based on the presence of length polymorphisms due to a variable number of repeats. The microsatellite markers B17 and W35 have been applied to a large number of strains and been shown to be highly variable between different strains of *T. gondii* (Ajzenberg et al., 2004). Fragment length analysis at these loci showed that amplicons from all eight Ugandan strains were of the same length for these markers, and sequencing of four amplicons (TgCkUg2, 5, 6 and 8) did not reveal any substitution polymorphisms. This analysis showed that the Ugandan type II strains (including TgCkUg2) were identical to Me49 and the Ugandan type III to VEG, which confirmed that the Ugandan strains are indeed highly similar to each other and the reference strains. In addition to the four Ugandan isolates, sequences for B17 and W35 from 46 strains were downloaded from GenBank, and a Bayesian phylogeny was generated from the concatenated B17+W35 sequences. As shown in Figure 4-11, a high level of divergence exists over these loci, but the Ugandan strains were conserved and clustered with the archetypal II and III strains.

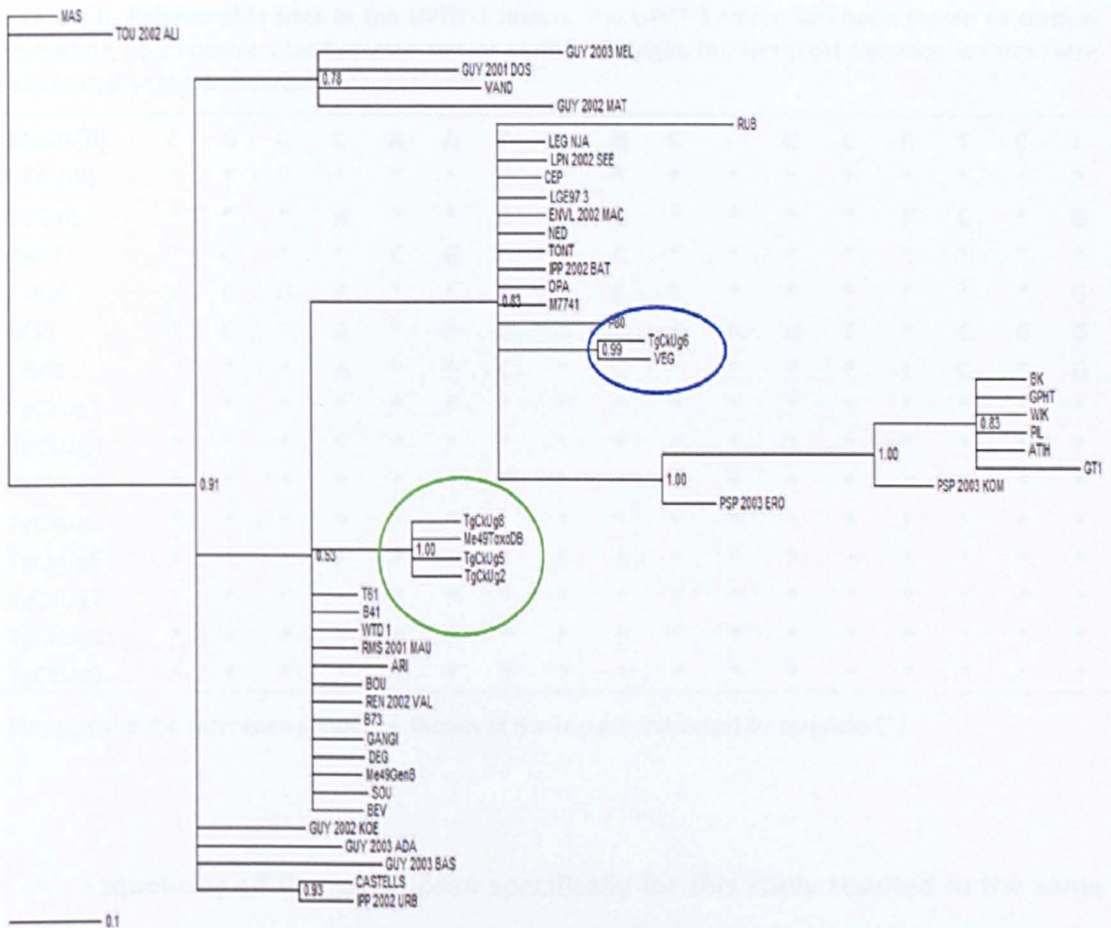


Figure 4-11. Phylogram for microsatellite markers B17 and W35. Sequence from 46 strains (mostly of South American origin) have been generated by Ajzenberg et al (2004) and these data were retrieved from GenBank and analyzed together with the sequence from four Ugandan strains. This tree shows the high degree of conservation between TgCkUg6 and VEG (blue circle) and also between TgCkUg2, 5 and 8 and Me49 (green circle), while there is a high level of divergence among the other strains for these two polymorphic loci.

Targeted sequencing of selected polymorphic loci (see 4.2.9) was performed in order to further characterize these strains and potentially find new SNPs. The UPRT-1 intron is located on chromosome XI, which is dominated by type I SNPs, meaning that type I has polymorphisms compared with types II and III, which are mostly identical. Several new polymorphisms have been detected at this locus for various strains (Table 4-6), but it was not useful for differentiating between the Ugandan strains, which were all identical to Me49 and VEG.

Table 4-6. Polymorphic sites in the UPRT-1 intron. The UPRT-1 intron has been shown to contain numerous polymorphic sites between strains of diverse origin, but the eight Ugandan isolates were allidentical to the archetypal lineages II and III.

Me49 (II)	C	G	C	C	A	A	T	G	G	C	-	C	C	C	T	C	T
VEG (III)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
GT1 (I)	*	*	*	A	*	*	C	*	C	*	*	*	*	T	C	*	G
CAST	*	C	*	*	C	G	*	*	C	*	*	*	*	*	*	*	*
COUG	T	C	G	A	*	*	C	C	C	*	*	*	*	*	*	*	G
MAS	*	C	*	A	*	*	C	*	C	A	A	G	T	*	C	G	G
RH88	*	*	*	A	*	*	C	*	C	*	*	*	*	T	C	*	G
TgCkUg1	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg2	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg3	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg5	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg6	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg7	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg8	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg9	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*

Identities to the reference sequence shown at the top are indicated by asterisks (*).

Sequencing of the loci chosen specifically for this study resulted in the same conclusion; the *T. gondii* isolates from Uganda are highly conserved and fit well into the clonal lineages observed in Europe and USA (Table 4-7). Based on analysis of 5.1 kb high-quality sequence from five polymorphic chromosomal regions it was possible to detect a low level of variation between the type II strains, and discover SNPs that were conserved in several of the Ugandan type II strains, but not present in Me49. The Ugandan type III strain(s) (TgCkUg6, plus TgCkUg2 for one locus), was however identical with VEG at all positions. The very limited divergence between Me49 and Ugandan type II and between VEG and Ugandan type III is in sharp contrast to the comparably high level of difference between VEG and Me49.

The amplified sequence of the apicoplast genome encompassed 904 bp and contained two SNPs compared with the type I strain RH (the only sequence presently available in ToxoDB). Both SNPs (one deletion and one C-A substitution) were present in the coding sequence of the rRNA coding gene TogoCr30, and were conserved in all eight Ugandan isolates.

Table 4-7. Summary of divergences between strains over >5 kb of sequence

Locus	Chrom.	Length (bp)	SNPs between Me49/VEG	Divergence Me49/VEG	SNPs between Me49/UgII[a]	SNPs between VEG/UgIII[b]
641.m01562	IV	1037	98	9.5%	2	0
20.m03896 (ROP18)	VIIa	1601-7	97	6.0%	0	0
80.m05038 (SRS36A)	IX	949	43	4.5%	0	0
33.m02185 (Toxofilin)	X	874-7	138	15.7%	1 - 3	0
65.m01164	XII	631-4	14	2.2%	0	0
Total		**5107**	**390**	**7.6%**	**3 - 5**	**0**

[a]UgII=TgCkUg1, 3, 5, 7, 8, 9 plus TgCkUg2 for loci 641.m01562, 20.m03896, 80.m05038, 33.m02185.
[b]UgIII=TgCkUg6 for all loci, plus TgCkUg2 for locus 65.m01164.

Toxofilin (33.m02185), is a highly expressed rhoptry protein that binds to host actin, and may affect the virulence of the parasite (Bradley et al., 2005). This protein is highly polymorphic with a predominance of type III SNPs (Figure 4-4 D). Sequence analysis revealed between one and three SNPs in the Ugandan type II strains compared with Me49 for this gene. Two of these SNPs were also found in the type I strain GT1, but one was exclusively found in TgCkUg7 and TgCkUg9. Interestingly, all of these SNPs were non-synonymous and resulted in amino acid substitutions, as shown in Figure 4-12.

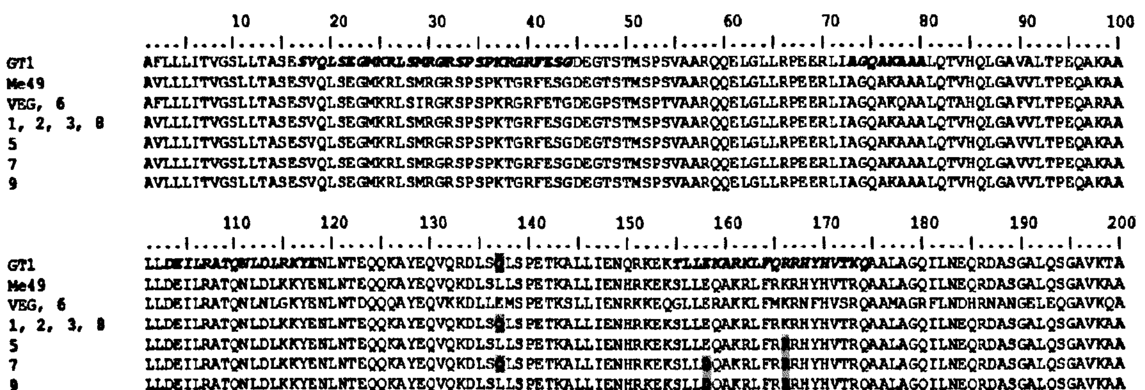


Figure 4-12. Amino acid substitutions in Toxofilin. The Toxofilin amino acid sequences for GT1 (I), Me49 (II) and VEG (III), plus the eight strains from Uganda (abbreviated to 1, 2, 3 ...). This protein has a relatively high level of SNPs between the reference strains, especially type III SNPs. TgCkUg6 was identical to VEG, but all type II were different from Me49 for at least one position. TgCkUg1, 2, 3 and 8 had one amino acid substitution, TgCkUg5 and 9 had two and TgCkUg7 had three (grey shading). The bold italic text mark the predicted actin-binding motifs (Jan et al., 2007).

The bradyzoite specific surface protein (641.m01562) contained two SNPs, which were conserved in five of the Ugandan type II strains (TgCkUg1, 2, 3, 7 and 8). One of these SNPs was non-synonymous, leading to a change from alanine to glutamic acid at position 75. The two remaining type II strains, TgCkUg5 and 9, were identical to Me49, and TgCkUg6 was identical to VEG. No new SNPs were found in the other polymorphic genes ROP18 (20.m03896), SRS36A (80.m05038) and the hypothetical protein 65.m01164; at all positions the Ugandan type II strains were identical to Me49 and the Ugandan type III strain was identical to VEG. TgCkUg2 was shown to possess type II sequence for all these loci except 65.m01164, where it was identical to the type III strains (see Figure 4.14). Additional sequencing was performed later and is reported in Chapter 5, and the final phylogenetic trees are shown in Figure 5-12.

4.3.6 RAPD analysis

PCR amplification with random primers was used to compare the "genetic fingerprints" of the strains without the bias of targeted amplification. Results from seven RAPD reactions (using primers C20, B12, 10.1, 10.2, 12.2+12.3, 12.2+12.4 and 12.3+12.4) immediately showed that the Ugandan strains were similar to each other and the clonal types Me49 (II) and NED (III), while the atypical strains VAND, TONT and RUB were different from the clonal types and more similar to one another (Figure 4-13).

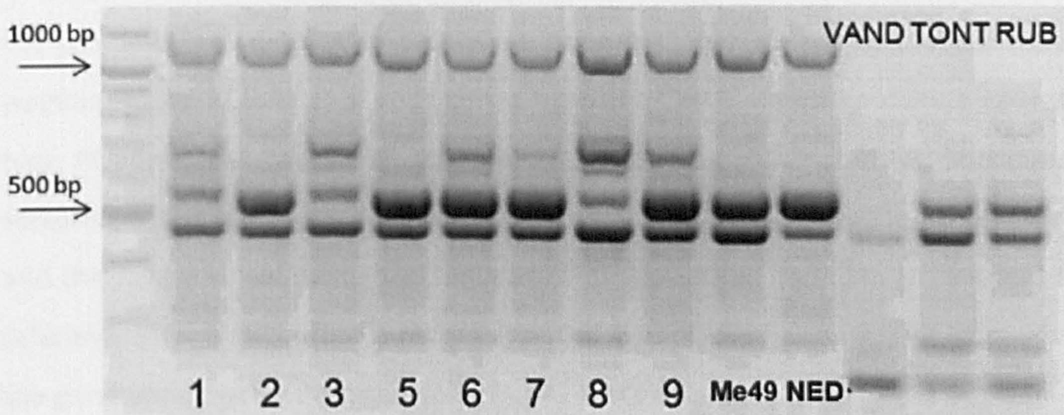


Figure 4-13. Amplification with random primers 12.2+12.4. Ugandan isolates (1-9) were highly similar to each other and to the clonal strains Me49 (II) and NED (III). DNA from atypical strains VAND, TONT and RUB were easily differentiated from the Ugandan and clonal strains by all primers. The size marker is a 100 bp marker with the 1000 bp and 500 bp bands indicated by arrows.

The final analysis included 60 bands, which were noted as present or absent for each strain, and the overall difference between strains ranged from 1-46 bands, but excluding the atypical strains, this was narrowed down to 1-14 fragment differences (Table 4-8). The highest level of similarity was found between TgCkUg1 and 3 (98%), TgCkUg5 and 9 (98%) and between TgCkUg7 and 9 (97%).

Table 4-8. Summary of RAPD results.

Strain	Ug1	Ug2	Ug3	Ug5	Ug6	Ug7	Ug8	Ug9	Me49	NED	VAND	TONT	RUB
TgCkUg1	-	5	1	5	8	6	9	4	13	6	*45*	*31*	*32*
TgCkUg2	92%	-	6	4	3	3	12	3	8	3	*42*	*26*	*27*
TgCkUg3	**98%**	90%	-	6	9	5	10	5	14	7	*44*	*32*	*31*
TgCkUg5	92%	93%	90%	-	7	3	12	1	10	7	*44*	*30*	*29*
TgCkUg6	87%	95%	85%	88%	-	6	9	6	7	4	*43*	*27*	*28*
TgCkUg7	90%	95%	92%	95%	90%	-	13	2	9	6	*43*	*29*	*29*
TgCkUg8	85%	80%	83%	80%	85%	78%	-	11	12	11	*46*	*30*	*29*
TgCkUg9	93%	95%	92%	**98%**	90%	**97%**	82%	-	9	6	*45*	*29*	*28*
Me49	78%	87%	77%	83%	88%	85%	80%	85%	-	9	*40*	*26*	*25*
NED	90%	95%	88%	88%	93%	90%	82%	90%	85%	-	*43*	*29*	*30*
VAND	25%	30%	27%	27%	28%	28%	23%	25%	33%	28%	-	*22*	*21*
TONT	48%	57%	47%	50%	55%	52%	50%	52%	57%	52%	63%	-	7
RUB	47%	55%	48%	52%	53%	52%	52%	53%	58%	50%	65%	88%	-

For the horizontal row the strain names of the Ugandan strains are abbreviated for clarity. Ug1=TgCkUg1 *etc.*

The upper right part of the table shows the number of bands which were present in one but not the other of two strains (>20 in italics), while the lower part shows the degree of identity between each strain pair (>95% in bold).

4.3.7 Composite genotype of the recombinant strain

The TgCkUg2 strain was isolated from a chicken with multiple infections and the genotype was difficult to determine prior to clonal propagation in cell culture (section 4.2.6). PCR-RFLP of four genes from the first isolation in culture revealed type III alleles for the SAG2 and SAG3 loci, but type II for GRA6. Subsequent sequencing disclosed type II alleles for five additional loci and type III for one locus, and the high level of sequence similarity with the other Ugandan strains and the reference strains suggested a recombinant of two modern strains. The locations of the genotyped markers are shown in Figure 4-14.

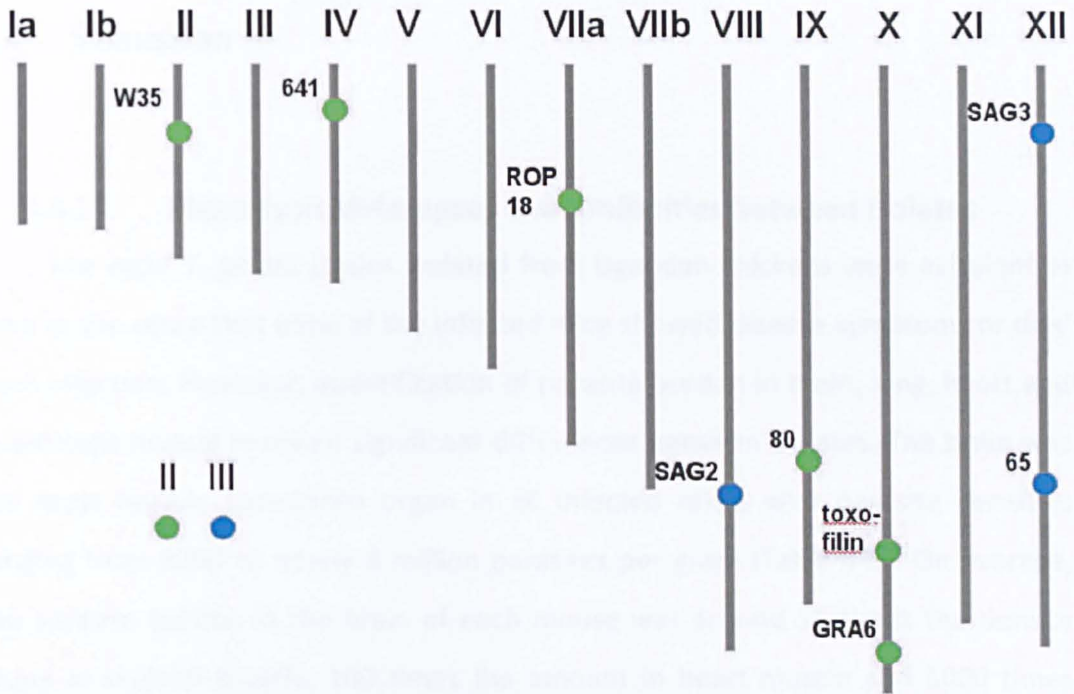


Figure 4-14. Mapping of genotyped loci in TgCkUg2. The approximate location of informative genotyped loci in the recombinant strain TgCkUg2 is mapped to the fourteen chromosomes (Ia-XII). The results of nine loci, which distinguish between genotype II and III, showed that TgCkUg2 held type II alleles at six loci (green) and type III alleles at three (blue), providing evidence of its recombinant nature. 641 = 641.m01562, 80=80.m05038, 65=65.01164.

4.4 Discussion

4.4.1 Phenotypic differences and similarities between isolates

The eight *T. gondii* strains isolated from Ugandan chickens were avirulent in mice in the sense that none of the infected mice showed disease symptoms or died from infection. However, quantification of parasite burden in brain, lung, heart and quadriceps muscle revealed significant differences between isolates. The brain was the most heavily parasitized organ in all infected mice, with parasite densities ranging from 8000 to nearly 8 million parasites per gram (Table 4-5). On average, the parasite burden in the brain of each mouse was around 15 times the density found in skeletal muscle, 100 times the amount in heart muscle and 1000 times more than in lung tissue. The preferential location of *T. gondii* in the central nervous system of chronically infected mice is in agreement with previous studies, which have shown an early, transient infection in lungs and lymphoid organs, followed by long-term persistence in the brain (Sumyuen et al., 1995; Dubey, 1997b). The current study utilized Q-PCR for quantification of *T. gondii* in mouse tissues, and the accuracy of this method revealed substantial variation between individual mice infected with the same *T. gondii* strain (occasionally more than 100-fold within the same organ), and this variation indicates that host genotype and immune response plays an important role in the outcome of *T. gondii* infection. Nevertheless, parasite genotype appeared to be the major determinant for tissue burden in mice in this study, where the type III strain caused significantly higher parasite densities compared with the type II strains, and the recombinant strain had an intermediate phenotype (Figure 4-6).

Virulence of *T. gondii* strains is traditionally determined by mouse mortality, where type I strains generally kill all mice, type II less than ten per cent and type III strains display more variability and intermediate levels of virulence (Araujo and Slifer, 2003). However, quantification of tissue burden through Q-PCR provides insight into the more subtle differences between strains and also differences between individual mice. A high number of parasites in the mouse brain may

augment the behavioural influence, which enhances the possibility of *T. gondii* transmission to the definitive host (Vyas et al., 2007). Furthermore, an overall high parasite burden contributes to a higher infectious dose in the cat (or other carnivore), increasing the likelihood of successful transmission. A parasite that causes a high tissue burden without critical impairment of the host fitness is most likely to be successfully transmitted (de Roode et al., 2008). In the current study, the Ugandan type III strain caused a higher parasite density in mice without causing disease symptoms and thus appears to be better adapted than the type II strains to murine hosts. The genotype related differences detected in mice was not correlated to the growth rate in culture, where three of the type II strains grew more rapidly than the recombinant and type III strain (Figure 4-9). Thus, the variation in mouse tissue burden could not be explained by intrinsic differences in replication rate between the lineages, instead it might be caused by strain-specific differences in the activation of the mouse immune response. Consistent with this, type II strains have been shown to be more potent inducers of the immune activators IFN-γ and IL-12 compared with type III strains, due to differential immunogenic pathway interactions of polymorphic rhoptry proteins (Saeij et al., 2005; Saeij et al., 2007), and the contribution of several genes can explain the intermediate virulence of the recombinant strain. The nature of the recombinant strain TgckUg2 is further investigated and discussed in Chapter 5.

An unexpected trait common to all the new isolates, regardless of their genotype or subsequent growth pace, was the slow initial growth at every new passage in cell culture, with the exception of trypsin treated parasites in passage 1 (Figures 4.7, 4.8 and 4.9). Trypsin is used to release bradyzoites from tissue cysts and promotes transition to the rapidly dividing tachyzoite form (Dubey, 1998a), therefore application of trypsin digestion to half of the samples explains the variable growth rates in the first passage, where parasites were retrieved from chronically infected mice with bradyzoites in the brain. However, the lag-phase seen in subsequent passages, where parasites were transferred from one culture flask to another, was unexpected. This growth retardation may be caused by spontaneous conversion to the bradyzoite stage, as indicated by transmission electron

microscopy (Figure 4-10), and may be due to their very recent isolation. Few papers describe the *in vitro* growth characteristics of primary *T. gondii* isolates, but a study from Zimbabwe also reported that new isolates required several weeks to consume a monolayer and that the numbers of free and viable parasites were low compared with well-adapted strains (Hove et al., 2005). It is possible that some strains are inherently less agile in tissue culture, but more data on primary cell culture isolation of *T. gondii* strains from different parts of the world would be needed to find out whether this is a specific attribute of some African strains or if a higher level of 'bradyzoiteness' is commonly seen during the first months in culture.

4.4.2 Limited genetic divergence among *T. gondii* from Uganda

Five unlinked genetic loci, situated on chromosomes IV, VIIa, IX, X and XII, were sequenced in the Ugandan isolates and compared with the reference strains. More than 5 kb high quality sequence was retrieved from each of the isolates, but a very limited level of diversity was detected. These loci were selected due to their polymorphic nature; a total of 390 SNPs are present between types II and III, corresponding to a nucleotide difference of 7.6% (the genome wide difference between these lineages is just over 0.5%, see Table 5-1 in section 5.1.2). Over the total length of 5107 bp, TgCkUg6 was identical to VEG, TgCkUg2 was identical to VEG and TgCkUg6 for one of the sequenced loci (65.m01164), but otherwise identical to three of the Ugandan type II strains (TgCkUg1, 3 and 8). The total number of SNPs between the Ugandan type II strains compared with the reference strain Me49 was 3-5 (<0.1%), and there were no genetic differences at all between the Ugandan type III strain and the VEG strain isolated from USA two decades ago. This minimal genetic drift implies that the strains from the same lineage share a very recent common ancestor and it is possible that there has been a recent human-mediated movement of these strains between continents (Lehmann et al., 2006). In that case, the strains isolated from Uganda may represent strains present in the human and domestic animal cycle, which is connected through inter-continental travel and trade, and may have less resemblance to ancient African strains which could have been introduced when the big cats arrived from Asia

millions of years ago. The presence of these supposed ancient African strains has not been shown, but the relative tolerance to overt toxoplasmosis in 'Old World' animals suggests that *T. gondii* has evolved together with these species for a long time (Innes, 1997), and it is possible that non-clonal strains exist in other habitats or regions in Africa, perhaps relying on other definitive hosts than the domestic cat.

The sequence divergence that was found among the type II strains indicated a division into two groups with TgCkUg1, 3, 8 and the type II part of TgCkUg2 on one side and TgCkUg5, 9 on the other, while TgCkUg7 had SNPs representative of both groups (Figure 4-15 A). Interestingly, this correlated well with the growth in the eighth cell culture passage, where TgCkUg1, 3 and 8 had a more rapid growth compared with TgCkUg5 and 9, while TgCkUg2 and 7 were in between (Figure 4-9). Furthermore, the close relationship between strains TgCkUg1 and 3, as well as that between TgCkUg5 and 9 was confirmed by RAPD analysis (Figure 4-15 B), and together these results point to at least two populations among the Ugandan type II strains which have a genotype related growth rate *in vitro*.

A.

Me49	T	A	A	C	G
TgCkUg1	A	.	.	T	T
TgCkUg2	A	.	.	T	T
TgCkUg3	A	.	.	T	T
TgCkUg5	.	.	G	.	.
TgCkUg7	A	T	G	T	T
TgCkUg8	A	.	.	T	T
TgCkUg9	.	T	G	.	.

B.

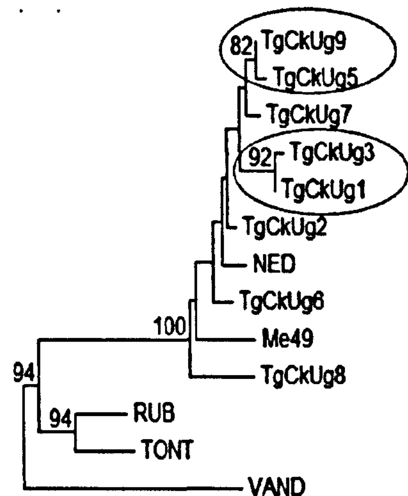


Figure 4-15. Genetic subgroups among the Ugandan type II strains. A. Sequencing revealed a maximum of five SNPs over 5109 bp between the Ugandan type II strains and Me49, which showed that TgCkUg1, 3, 2 and 8 were identical to each other, while TgCkUg5, 7 and 9 showed different SNP patterns. **B.** RAPD analysis confirmed the close relationship between the isolates from Uganda and the clonal lineages and the relationships suggested by the SNPs were partly confirmed by RAPD analysis (encircled). The tree is a neighbour-joining tree generated in Treecon (see 4.2.10).

4.4.3 Isolation of a natural recombinant strain

The discovery of multiple infections in chickens and the isolation of a recombinant strain suggest that, under the right circumstances, the sexual cycle may be more common than previously thought. The recombinant TgCkUg2 did not exhibit more divergence from the archetypal lineages at any single locus, but was identical to type III on loci tested from chromosomes VIII and XII, and to type II on chromosomes II, IV, VIIa, IX and X, and is thus highly likely to be a recent recombinant rather than an ancient, atypical strain. Recombination between different strains has been assumed to be a rare event in nature (Boothroyd, 1993). First, more than one genotype must be present in the environment and then they must be ingested by the definitive host within a short time frame. The most likely way for this to happen is for a cat to eat a host animal which is infected by two different strains, but the strong immune response elicited by the primary infection may prohibit the establishment of a second infection in some host species. However, the possibility of multiple infections has been proven in laboratory mice. Araujo and co-workers demonstrated that 100% of mice chronically infected with a type II strain were successfully infected with a type III strain via the oral route, and the mice brains were later found to harbour tissue cysts from both strains (Araujo et al., 1997) . These results have later been confirmed for types II and III (Dao et al., 2001) as well as for type II infections followed by type I (Dzitko et al., 2006). However, the relative abundance of the different strains in the host tissues may vary considerably and if one strain is largely predominant, self-fertilization is likely to be more frequent than recombination between strains. For the Ugandan strains we found that the difference in parasite burden in mice was up to 1000-fold and if this relationship would be true in a multiple infection it is likely that only the most abundant strain would prevail. Indeed, starting with three multiply infected chickens, three passages in mice and possibly unequal tissue burdens in the original host resulted in the propagation of only one strain from each chicken. Concern about the loss of strains from multiple infections in the process of isolation has also been raised previously (Villena et al., 2004).

Although there appear to be many hurdles on the way to a successful recombination event, this may be more frequent in some environments. The study presented in Chapter 3 detected a high rate of multiple infections in chickens from Kampala, where two and even three strains were present at detectable levels. As the chickens used in this study were free-range and there are plenty of cats around, which may prey on the chickens, scavenge from rubbish piles or be fed chicken viscera, it is possible that recombination events are more common in that setting than in sparsely populated areas or industrialized cities.

The benefit of recombination and genetic variation may be limited in the domestic cycle, since the three clonal lineages appear to be very well-adapted to these hosts. During the cross that gave rise to TgCkUg2, thousands and maybe millions of genetically distinct progeny must have arisen, but many of these might have been inferior and quickly outcompeted. While recombinant strains are sometimes discovered by direct genotyping in tissues or isolated in mice, this is the first time a natural recombinant strain and its putative parental strains have been propagated in culture and compared *in vivo* and *in vitro*. Although it is not possible to tell how long TgCkUg2 has existed, it has been successfully transmitted in nature. Therefore this strain is interesting for further investigation, in order to find clues to its success in the environment where other recombinant strains have failed.

4.5 Conclusions

Toxoplasma gondii strains from domestic chickens in Uganda are closely related to the globally distributed clonal lineages, with a very low level of polymorphisms at the nucleotide level, but with dramatic phenotypic differences both *in vivo* and *in vitro*. It is possible that the clonal lineages of *T. gondii* have been introduced recently in Uganda, perhaps through trade in food or animals, and it remains to be elucidated if these strains, which are readily propagated between intermediate hosts, have outcompeted ancient African strains or if such strains can be found in habitats less affected by human activities. The very low level of genetic variation in *T. gondii* emphasizes the need for in depth genotyping to detect genetic differences between strains, which may have dramatically different phenotypic characteristics. Identification of a recombinant strain with intermediate growth rate *in vivo* shows that new genotypes with altered phenotypic properties may suddenly appear in places where different strains are circulating and cats, humans and other animals live close together. The unique isolation of a natural recombinant *T. gondii* isolate from Africa led on to the next phase of study where the whole genome of this strain was sequenced in order to fully probe the nature of the recombination event and divergence of this African isolate relative to the archetypal strains.

Chapter 5 Whole genome sequencing of a recombinant *Toxoplasma gondii* strain from Uganda reveals chromosome sorting and local allelic variants

5.1 Introduction

Sequence comparisons of entire genomes provide unparalleled opportunities for detection of genetic relatedness, new polymorphisms and genes under selection in an efficient large-scale and unbiased manner. The aim of the research described in this chapter was to generate the first whole genome sequencing of a recombinant *T. gondii* strain and make a thorough comparison with the reference type II and III strains to pinpoint new mutations that might be of biological significance.

5.1.1 The *T. gondii* genome and ToxoDB

Early mapping studies of the *T. gondii* genome using pulse-field electrophoresis and hybridization led to the identification of eleven chromosomes and a genome size estimate of 80 Mb (Sibley and Boothroyd, 1992b; Sibley et al., 1992). In 2005 a new composite map was assembled where the number of genetic markers was expanded from 57 to 250, which led to the identification of additional linkage groups resulting in the naming of 14 chromosomes (Khan et al., 2005b), and this was used as a scaffold for the assembly of the first, 10x shotgun whole genome sequence of *T. gondii*.

ToxoDB (www.toxodb.org) is a functional genomics database for *T. gondii*, which was first released in 2001 (Gajria et al., 2008) and the original sequencing of B7 (a clone of Me49) was performed by The Institute for Genomic Research (now part of the J. Craig Venter Institute) in collaboration with David Roos at the University of Pennsylvania, and funded by the NIH/NIAID. This work has not been

published, but the methodology is described on http://www.tigr.org/tdb/e2k1/tga1/. Briefly, random genomic shotgun libraries were inserted in plasmids, which were sequenced from both ends using conventional sequencing (Sanger, 1977). In addition, random selections of larger sequences inserted in BACs were sequenced to provide a scaffold for linking the contigs. The first 0.7x assembly was released in March 2002; thereafter the coverage increased through more sequencing and enhanced contig assembly (http://v3-0.toxodb.org/about_data.shtml). In September 2003, ToxoDB 3.0 was released with 10x coverage of the Me49 strain. Version 4.3 was released in December of 2007 and is the version that was used for the sequence comparisons in this thesis. The main difference between v.4.3 and the new release (v.5.0, Nov 2008) is the annotation of GT1 and VEG and the new nomenclature for chromosomal regions and genes (Figure 5-1). However, the old version can still be accessed and the previous gene names remain searchable in ToxoDB 5.0.

ToxoDB can be accessed from the umbrella database EuPathDB (www.eupathdb.org), which facilitates comparisons between several eukaryotic pathogens including *Plasmodium*, *Cryptosporidium* and *Giardia*, with *Leishmania* and *Trypanosoma* soon to be incorporated. The latest version contains the full annotated genomes of GT1 (I), Me49 (II) and VEG (III), which were all isolated in the USA: GT1 from a goat (Dubey, 1980), Me49 from a sheep (Kaufman et al., 1958) and VEG from a human AIDS-patient (Parmley et al., 1992). In addition, chromosome Ia and Ib from another American type I strain, RH (sequenced by the Sanger Institute, Hinxton, UK) is available, this strain was isolated from a 6-year old child in 1939 (Sabin, 1941). The apicoplast genome is available for RH only, while the mitochondrion sequence has not yet been made available. In addition, the draft *Neospora* genome (Sanger) is now available through ToxoDB. Through the "Queries and Tools"-option it is possible to browse the genome, retrieve sequences, perform BLAST queries and identify genes based on their SNP characteristics, expression, cellular location, function, motifs and many other options, and it is also possible to combine queries to perform more complex tasks.

Figure 5-1. Comparison of the genome browser in ToxoDB versions 4.3 (A) and 5.0 (B). The two pictures show the same region on chromosome II in the previous and new version of ToxoDB. The main difference is the possibility of direct comparison between the three strains, which have all been annotated in the new release. Also note the re-naming of the chromosomal regions and the genes. The SNP coding system is the same, where red indicate type I SNPs (type I different from type II and III, which have the same allele), green for type II SNPs and blue for type III SNPs. Black diamonds indicate the rare instances where all three strains have different nucleotides in a certain position.

The full *T. gondii* genome sequence extends to 61.6 Mb, distributed over 14 chromosomes ranging in size from <2 Mb (Ia and Ib) to >7 Mb (X). The number of chromosomes is the same as for *P. falciparum* (Gardner et al, 2005), but the genome size is nearly three times larger. The AT-richness characterizing the genome of *P. falciparum* is not present in *T. gondii,* which typically has a GC-content of around 50% and higher in coding regions (Khan et al, 2006a). Approximately 9,000 genes have been predicted in the three sequenced strains of *T. gondii*, but around

half of these still do not have any function assigned. Many *T. gondii* genes are heavily spliced, around 75% have at least one intron, and over 1000 genes have got ten or more exons, with dynein heavy chain proteins being the most extreme with >60 intronic regions (ToxoDB). Over 2,000 *T. gondii* proteins have recently been analyzed through proteomic analysis using 2D-gel separation followed by mass spectrometry, and intron spanning peptide fragments mostly confirmed the predicted splice sites (Xia et al,. 2008). Centromeric regions have not been defined in the *T. gondii* chromosomes (Gubbels et al., 2008) and in spite of its large genome size *T. gondii does* not have the huge amount of repetitive DNA elements found in some other protozoan parasites like *Trypanosoma* and *Plasmodium* (Wickstead et al., 2003). However, clusters of tandem repeats (Matrajt et al., 1999) and mobile genetic elements (Terry et al., 2001) have been identified, and in addition gene duplications are common. A recent bioinformaticanalysis reported the detection of 787 duplicate genes (Dybas et al., 2008), and this is thought to be particularly important for the variation of surface antigens (He et al., 2002).

5.1.2 Genetic differences between strains

The clonal lineages I, II and III are highly similar on the nucleotide level, with an overall divergence of less than one per cent. The largest difference (as determined through ToxoDB searches) is >0.9%, for chromosomes Ib and XII between genotypes I and II, while the highest level of similarity is seen for chromosome XI between lineage II and III, where the dissimilarity is limited to 128 SNPs over more than 6.5 Mb, while there are more than 50,000 SNPs to type I (Table 5-1). This uneven distribution shows that chromosome XI has an overwhelming dominance of type I SNPs (where type I has a different allele compared with type II and III). Chromosome Ia is highly conserved in all three lineages and was probably jointly inherited in all three lineages through the original crosses (see section 4.1.2). One chromosome (IV) has a predominance of type III SNPs, while most of the others have different regions dominated by two or three different SNP types (Boyle et al., 2006).

Table 5-1. Nucleotide differences between types I, II and III.

		No of SNPs			Nucleotide difference (%)		
Chrom.	Length (bp)	I-II	I-III	II-III	I-II	I-III	II-III
Ia	1,896,408	324	329	56	0.017%	0.017%	0.003%
Ib	1,956,324	17,845	16,264	6,795	0.912%	0.831%	0.347%
II	2,302,931	20,059	15,604	20,394	0.871%	0.678%	0.886%
III	2,470,845	21,713	15,177	18,312	0.879%	0.614%	0.741%
IV	2,576,468	485	21,601	21,626	0.019%	0.838%	0.839%
V	3,147,601	25,579	22,199	26,177	0.813%	0.705%	0.832%
VI	3,600,655	31,400	29,308	8,852	0.872%	0.814%	0.246%
VIIa	4,502,211	32,096	28,967	39,409	0.713%	0.643%	0.875%
VIIb	5,023,822	43,363	36,392	24,102	0.863%	0.724%	0.480%
VIII	6,923,375	61,101	51,333	28,961	0.883%	0.741%	0.418%
IX	6,384,456	41,661	32,625	37,391	0.653%	0.511%	0.586%
X	7,418,475	59,497	41,950	59,496	0.802%	0.565%	0.802%
XI	6,570,290	54,354	54,366	128	0.827%	0.827%	0.002%
XII	6,871,637	62,373	54,638	21,075	0.908%	0.795%	0.307%
Total	61,645,498	471,850	420,753	312,774	0.765%	0.683%	0.507%

The data was retrieved from ToxoDB by submitting queries for each chromosome between the different strains. (Tool: "Find SNPs by chromosomal location"). Grey shading indicate extreme levels of similarity, where entire chromosomes are dominated by one SNP type, or in the case of Ia, are virtually void of SNPs.

The differences between strains are more pronounced on the gene level: out of the over 9,000 predicted or verified genes in the *T. gondii* genome, only around 7,000 are found in all three strains, and 6905 are present at exactly one instance in each strain. The number of unique and shared genes between the three strains is shown in Figure 5-2.

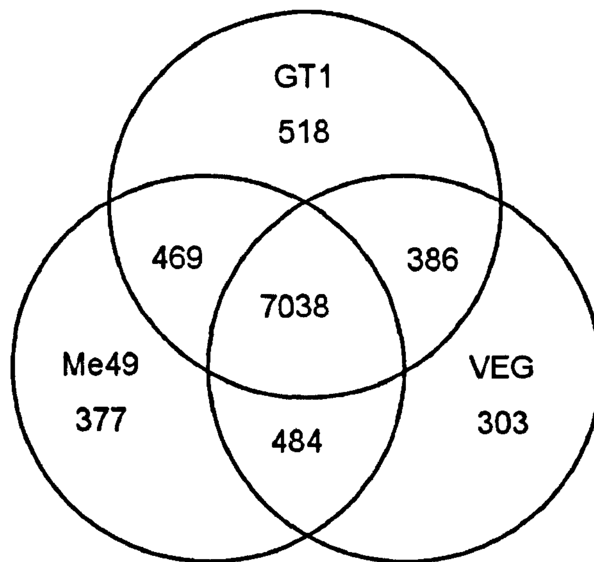


Figure 5-2. Shared and unique predicted genes among the three annotated strains. GT1 (I) has the highest number of unique genes, while Me49 (II) and VEG (III) have the greatest number of shared genes. (Data compiled from numerous queries in ToxoDB).

Among the 7038 genes present in all three strains, 5282 (75%) have at least one coding, non-synonymous SNP between at least two strains. Of these, 644 are under positive selection with a dN/dS ratio of >2, which mean that the number of non-synonymous SNPs is twice as high as the number of synonymous SNPs. Most of these genes are currently annotated as hypothetical proteins, but among the 184 with some kind of function assigned, proteins from a few families are overrepresented, including five dense granule proteins, eight rhoptry proteins and more than twenty surface antigens or SAG1 related sequences (SRS). The highest selective pressure is seen on the dense granule proteins GRA6 (dN/dS=16) and GRA3 (dN/dS=14.5), which are both located on chromosome X, while the highest absolute number of non-synonymous SNPs (245) is found in a hypothetical protein on chromosome XII (all information retrieved through queries in ToxoDB). In summary, the level of divergence on the nucleotide level is below 1% on all chromosomes between the three archetypal lineages, but the variation among predicted genes is in the order of 10%. Moreover, three quarters of all shared genes contain amino acid changing differences and many genes involved in invasion and host interactions appear to be under diversifying selective pressure.

5.1.3 Massively parallel genome sequencing technologies

In recent years, several new sequencing technologies have been developed that have in common that rather than the long, high-confidence reads obtained by conventional Sanger sequencing, they produce many millions of short reads (≈30-250 bp, and increasing). These new developments provide new opportunities since no cloning of source DNA is needed, thus the bias introduced in this step is eliminated. Furthermore the read density is directly proportional to the number of repeats, providing a direct quantification of minor allelic variants, tandem repeats etc (Swaminathan et al., 2007). The biggest advantage is the drastically reduced cost and increased throughput, which for the first time have made whole genome sequencing of eukaryotes possible outside the big genome research institutes. The Roche 454 platform was the first available new generation sequencing technique and to date the most widely used, with over 300 peer-reviewed articles (www.454.com, January 2009) and the principle is shown in Figures 5-3 and 5-4.

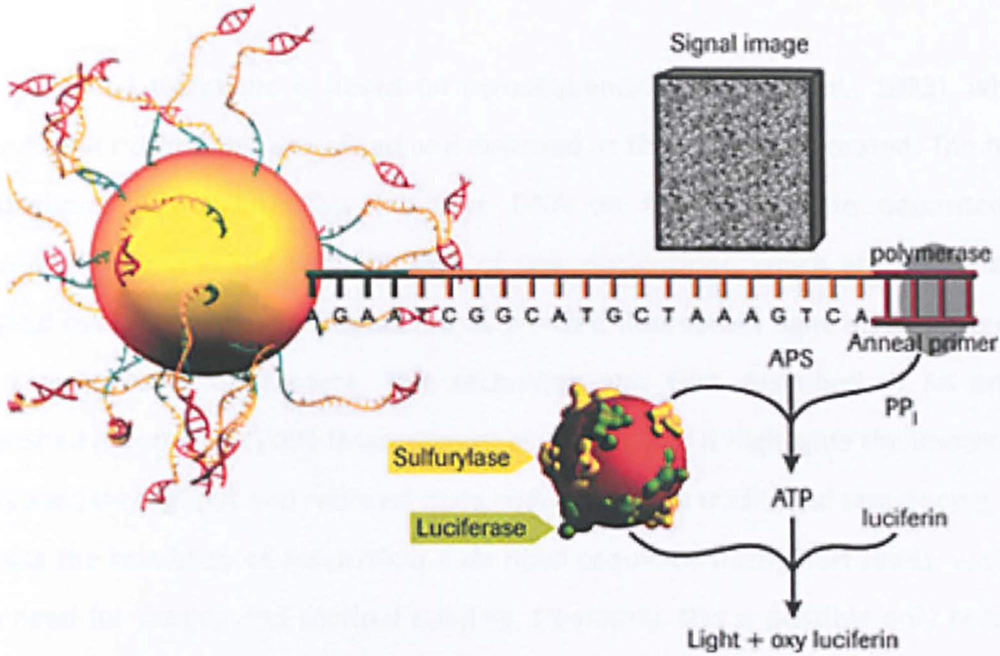


Figure 5-3. Chemistry of 454 sequencing. A single-stranded DNA library is prepared and short adaptors are attached to the 3' and 5' ends of each fragment, these are used for attachment to beads, primer-binding and purification. One unique fragment is bound to each bead and amplification takes place in separate microreactors (PCR-mixture droplets in an oil emulsion) resulting in several millions of copies of identical DNA fragments on each bead. The beads are deposited in microwells, containing one bead each and luciferase-labelled nucleotides are flushed sequentially over the wells. Each bead gives a separate read detected by a camera, which converts optical brightness to electrical amplitude signals. The picture and information was retrieved from the 454 website (www.454.com).

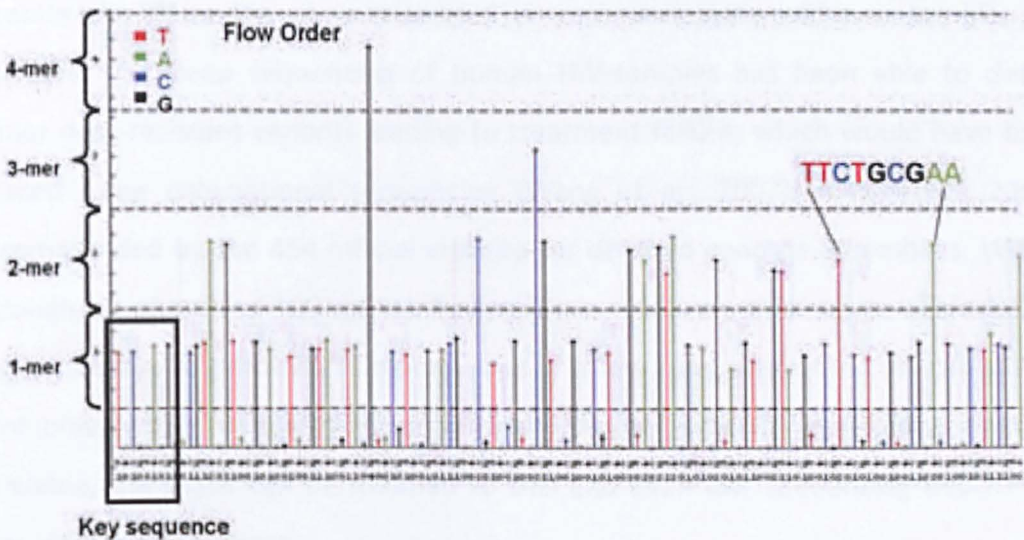


Figure 5-4. Reads from a single bead in 454 sequencing. As a control measure, all sequences start with a key sequence incorporated in the adaptors (the use of different adaptors enables the differentiation between multiple samples run in the same reaction), thereafter the actual sequence follows. The nucleotides follows a set flow order, and when incorporated the light emitted results in a signal proportional to the number of bases in a row. The picture and information was retrieved from the 454 website (www.454.com).

The 454 technique is based on pyrosequencing (Nyren et al., 1993), where fluorescent nucleotides are added and detected as they are incorporated. The high-throughput format amplifies template DNA on beads that are deposited in individual wells and the incorporation of new nucleotides, which are sequentially flushed over the wells, are registered as emitted fluorescent light that is detected by a high resolution camera. This technique was first described in an article published in Nature in 2005 (Margulies et al., 2005), and it highlights the immensely increased throughput and reduced costs compared with traditional sequencing, and proves the feasibility of assembling a *de novo* sequence from short reads, without the need for cloning and scaffold building. Obviously, this is possible only because the computer capacity and bioinformatics software have developed rapidly as well, and the capacity of these methods is continuously increasing. The possibility of ultra-deep coverage of small genomes and the capacity to use bioinformatics to assemble separate genomes even from closely related species present in the same sample have opened up completely new research fields like metagenomics – the investigation of all microbial species present in a certain niche such as soil, sea water or the human gut (Hall, 2007). Furthermore, the quantitative ability also enables the detection of rare species or rare mutations within a sample. For example, 454 deep sequencing of human HIV-samples has been able to detect minor drug-resistant variants leading to treatment failure, which would have been missed using conventional sequencing (Wang et al., 2007). A coverage 20x is recommended by the 454 official website for *de novo* genome assemblies. With a throughput of around 100 million base pairs in each run, this can be achieved for most prokaryotic genomes using only one or a few runs, while the cost can still be prohibitively high for eukaryotic organisms. However, when a reference sequence is available, the reads can be mapped to this and even low sequencing depths can provide a good coverage.

5.1.4 Aims

The aim of the final study of this thesis was to sequence and analyze the genome of the TgCkUg2 strain using the 454-platform, and this was the first *T. gondii* sequencing project to target a recombinant strain and one of the first whole genome sequencing projects of eukaryote pathogens employing one of the new generation sequencing technologies. Through alignment and comparison with the previously sequenced *T. gondii* reference strains, the goal was to unravel the genetic constitution of TgCkUg2, and to look for novel allelic variants and genes under selection in this recombinant *T. gondii* strain from Uganda.

5.2 Materials and Methods

5.2.1 Parasite strain

The parasite strain chosen for whole genome sequencing, TgCkUg2, was isolated from chicken number 2 (see Chapters 3 and 4 for more information). This chicken was a male, around 1.5 years of age and originated from Wakiso in the Gayaza area north of Kampala. The chicken was killed on Sep 4th 2006 and the *Toxoplasma* strain had been through three passages in mice, including cryopreservation for four months, and ten passages in cell culture when DNA was extracted in January 2008.

Parasites were grown in six T75 flasks and allowed to eat through the fibroblast monolayer. The total quantity of medium was 150 ml and the density of parasites in the medium was counted in a haematocytometer and found to be $3.45*10^6$ per ml, which gave a total harvest from all flasks of at least $5*10^8$ parasites. Only a small proportion of human cells were visible under the microscope but due to their large DNA-content, additional efforts were taken to reduce their number. Three PBS-washes (3,500 rpm, 5 min) was used to get rid of cell debris and was followed by the release of intracellular *T. gondii* by passage through 26g needles. Thereafter two more washes were performed at a lower speed (2000 rpm, 5 min) in order to collect only the dense parasites, but excluding the fibroblasts. Pellets were suspended in two aliquots of 1 ml PBS and new counts were performed. The total harvest was around $4*10^8$ parasites.

5.2.2 DNA extraction and quality control

Serial DNA extractions were performed using the QIAgen DNA extraction kit (see 2.2.4), using approximately $2*10^7$ purified parasites per reaction. The concentration and purity of the extractions were measured using a Nanodrop 1000 machine (Thermo Scientific). The DNA samples contained between 25 and 80 ng/µl and only extracts with a 260/280 > 1.8 and 260/230 > 1 were used in order to avoid protein and solvent contaminations. DNA from several extractions, which met the

required standard, were mixed and concentrated in a DNA concentrator to a volume of 50 µl and a total DNA content of 5 µg.

Two *T. gondii* genes (BTUB and SAG3) and one human gene (β-globin), all single-copy, were used to estimate the DNA contributions from the respective organisms to the final extracts, through amplification of serial dilutions (primers and conditions for these PCR-reactions are described in chapter 3.2.9), and there were around 100 parasite genomes per human copy. In addition, the BTUB and SAG3 amplicons were sequenced, and the type II result for BTUB and type III for SAG3 confirmed that this was indeed the recombinant strain.

5.2.3 454-sequencing, creation of output files and bioinformatic analysis

The whole genome sequencing including creation of the genomic library, annotation with other genomes and assembly of FASTA and Artemis files described in this section was performed by Prof. Neil Hall and Dr. Kevin Ashelford at the University of Liverpool. First, a DNA library with fragments of around 250 bp was created through nebulization and enzymatic attachment of the adaptors needed for amplification and sequencing. Thereafter, three runs were performed on the 454 platform described in section 5.1.3. Reads were assembled into contigs using the Newbler software (Roche) and the sequence from TgCkUg2 was aligned with the published sequences for the type II (Me49) and III (VEG) reference strains retrieved from ToxoDB. Sequences of human origin were filtered away. Since the strain was a type II/III recombinant, the 454 sequence was not aligned to the ToxoDB type I strain (GT1) as this would have increased the complexity of the analysis, while adding little useful information.

The output files were FASTA text files containing all contigs (their length, number of reads, location (mapped to Me49) and sequence, (see Appendix 9 for a sample) and FASTA text files with details of all the SNP calls (location, number of reads, nucleotide present in each strain, per cent of concordant reads and the actual reads shown. SNPs were considered valid if three or more reads, of which at least one in the forward and one in the reverse direction, were in agreement. If

there were additional reads not in agreement, or if a SNP could be mapped to several locations in the genome they were not taken into account. These strict criteria prohibited the detection of all existing SNPs, but those that were called can be assumed to be correct with a high level of confidence (see Appendix 10 for examples of valid and invalid SNPs).

Furthermore, files were prepared for analysis in Artemis (http://www.sanger.ac.uk/Software/Artemis /v10/) which made it easy to distinguish between the features that could be properly mapped and those that were in repeat regions or had insufficient reliability. Three different types of SNPs were called for TgCkUg2: novel SNPs, where the TgCkUg2 allele was different from both Me49 and VEG (orange), type II SNPs, where TgCkUg2 is identical to Me49, but different from VEG (green), and type III SNPs, where TgCkUg2 was identical to VEG, but different from Me49 (blue). In some cases, SNPs were called against one of the reference strains, while the relationship to the other could not be determined according to the strict criteria employed. In these instances a conservative interpretation was chosen, where TgCkUg2 was assumed to be identical to the other strain, rather than having a novel allele. In addition, SNPs between Me49 and VEG where the affiliation of TgCkUg2 could not be ascertained (yellow), and SNPs which had ambiguous mapping positions in the genome (white), were available in the Artemis file, but these were not included in the subsequent analysis. The SNP types and the colour codes are listed in Table 5-2.

Table 5-2. SNP colour scheme

TgCkUg2 vs. Me49[a]	TgCkUg2 vs. VEG[a]	Total number of SNPs[b]	Colour[c]
different	different	1,252	Orange
different	same	29,056	Blue
different	NA	735	
same	different	41,153	Green
NA	different	550	
SNPs between Me49 and VEG		165,225	Yellow
Repeat regions		11,483	White

[a]Result from comparisons between TgCkUg2 and either Me49 or VEG, where "same" means that the strains had the same allele, "different" means that they had different alleles and "NA" indicates that the comparison did not yield a definite answer. In these cases a conservative approach were chosen and these SNPs were assumed to be similar to one of the parental strains rather than being a novel SNP.

[b]SNPs of each type detected across the genome.

[c]This colour scheme was used to differentiate between SNPs in all the subsequent analyses, but the yellow and white SNPs were not included.

Additional features implemented in the Artemis files were genes, exons, introns, mRNA, tRNA, coding sequence (CDS), and information for each SNP on whether it was non-coding, synonymous or non-synonymous. A screenshot from Artemis with some of these features is shown in Figure 5-5.

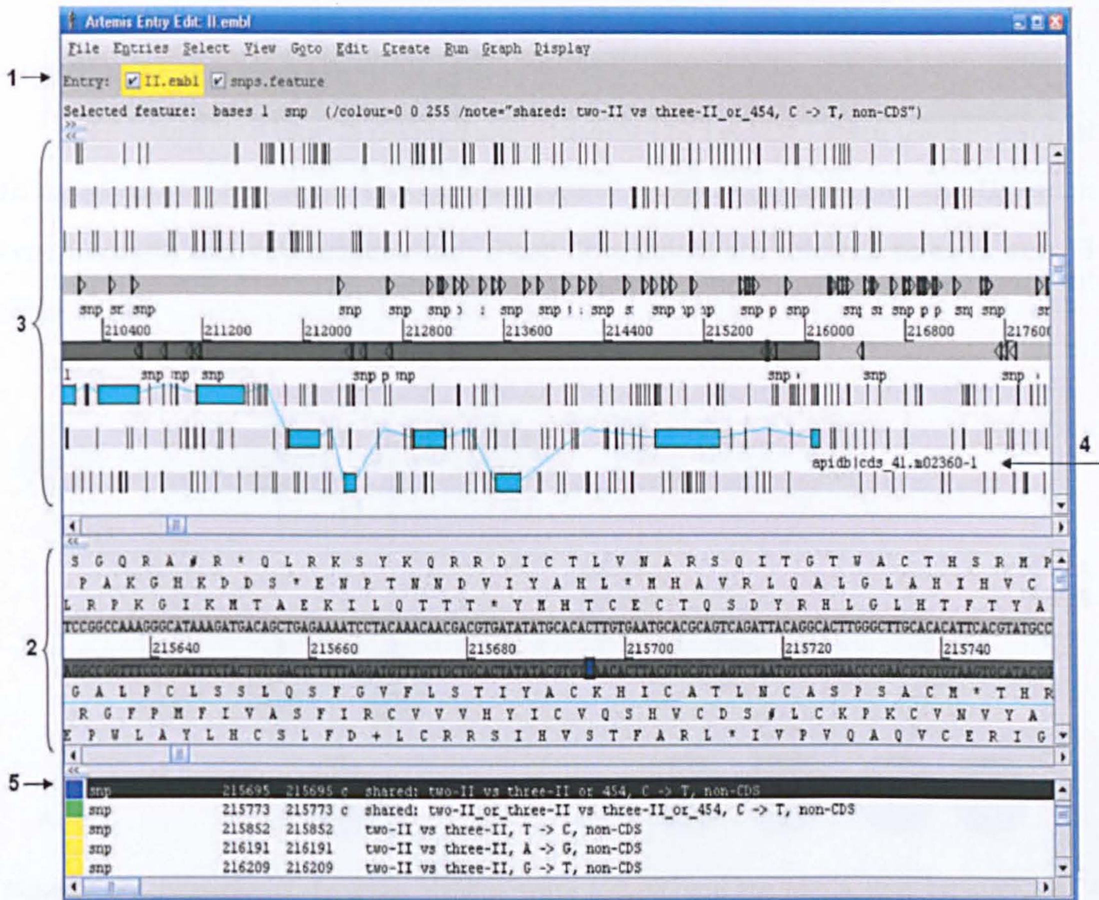


Figure 5-5. Artemis screenshot showing SNP features. The Me49 sequence from chromosome II is imported in embl format, which includes the gene annotations from ToxoDB, and then a feature file is overlaid, which includes the SNPs between Me49, VEG and TgCkUg2 **(1)**. The sequence and six-frame translation, position and features are shown in detail **(2)** and in larger scale **(3)**, both panels are zoomable. It is possible to show a range of features that Artemis reads from the sequence, here stop codons are displayed as vertical black lines in all reading frames in **(3)**, and one can see how the gene 41.m02360 **(4)** is encoded within several open reading frames (exons are shown as boxes, introns as lines and the whole mRNA is shown as a dark grey box). The selected feature here **(5)** is a non-coding SNP situated in an intron, at position 215695. All blue-coloured SNPs are shared between VEG (three-II) and TgCkUg2 (454), but different from Me49 (two-II). The colour codes (0 0 255 for blue) were used to retrieve all SNPs of the same type and export these to Excel for creation of graphs and other analyses.

Selected features were retrieved from Artemis and imported into Microsoft Excel, where calculations were performed and graphs made. To enable quantitative analysis of the FASTA files all contig information, except the actual sequences, were

imported to Excel where calculations regarding average contig length, distance between contigs, percent mapped etc could be easily performed. All contig, SNP and mapping data generated for TgCkUg2 has been submitted to ToxoDB and the inclusion of this data into the main database is currently in progress.

5.2.4 Divergence time estimate

The type II and type III derived chromosomes of TgCkUg2 were used in parallel to calculate the MRCA of the Ugandan type II (UgII) and Me49 and the Ugandan type III (UgIII) and VEG respectively. These time points are referred to as T2 and T3 (Figure 5-6).

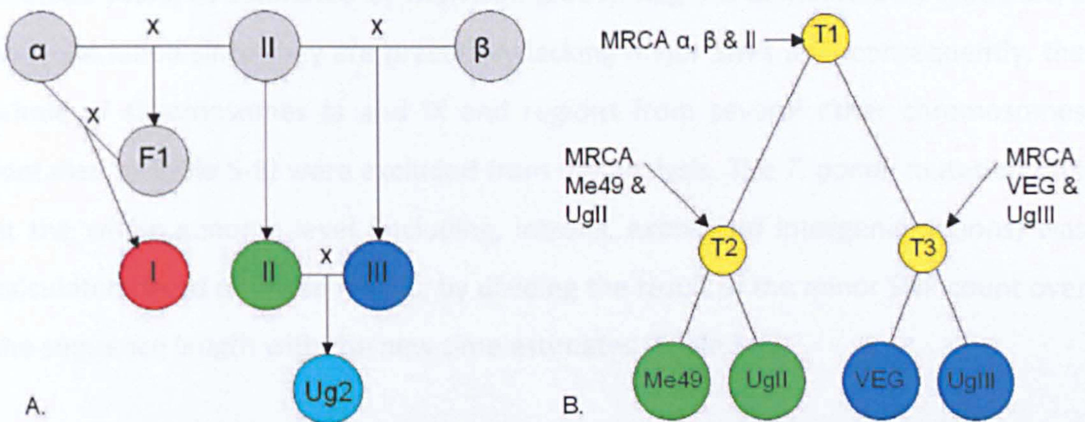


Figure 5-6. Schematics of the cross yielding types I, II, III, and the MRCA time estimates. A. A simplified overview of the crossing events between three ancestral strains (α, β and II) that resulted in the three modern lineages I, II and III (adapted from Boyle, 2006. Also see section 4.1.2). Ug2 stands for TgCkUg2, which is a recent recombinant between modern type II and III from Uganda. The type II chromosomes from this strain are used to represent Ugandan type II strains (UgII) and the type III chromosomes represent Ugandan type III strains (UgIII). **B.** The time point of the MRCA to strains α β and II is called T1 and is estimated at 150,000 years ago (Morrison, 2005). T2 is the MRCA of Me49 and UgII and T3 is the MRCA of VEG and UgIII.

Two different methods were used to calculate T2 and T3. The first was based on the *T. gondii* intron mutation rate, calculated by Su et al (2003). Intronic SNPs were defined as SNPs present in the mRNA, but not in the CDS, and the lengths of introns were calculated analogously. The intron minor SNP count divided by the intron length (the SNP density per nucleotide) was divided by the *T. gondii* intron mutation rate of $1.94*10^{-8}$ SNPs per nucleotide per year (Su et al., 2003). In addition, an estimate of the divergence time between type II and III (T1) was

calculated using the calculated estimates for T2 and T3 and relating these to the major SNP type found in introns (see Table 5-7 in the results section).

For the second set of calculations, the relation between major and minor SNPs was assumed to be directly proportional to the relation in time between the MRCA of all type II and III strains (T1) and the more recent MRCAs of Me49/UgII (T2) and VEG/UgIII (T3) (Figure 5-6 B), meaning that the mutation rate was assumed to be constant over time, in accordance with the 'molecular clock' principle (Zuckerkandl and Pauling, 1965). Minor SNPs were defined as SNPs against the background (*i.e.* against Me49 on type II chromosomes and against VEG on type III chromosomes) and included the unique (orange) SNPs. Major SNPs were defined as SNPs against the non-background type (but excluding unique SNPs). Here, T1 was assumed to be 150,000 years, as estimated by Morrison (2005). Regions dominated by type I SNPs were excluded since they are practically lacking major SNPs and, consequently, the whole of chromosomes Ia and IX and regions from several other chromosomes (detailed in Table 5-8) were excluded from the analysis. The *T. gondii* mutation rate at the whole genome level (including, introns, exons and intergenic regions) was calculated based on these results, by dividing the result of the minor SNP count over the sequence length with the new time estimates (Table 5-8).

5.2.5 Targeted sequencing and phylogeny.

Sanger sequencing was used to verify SNPs discovered by the 454 sequencing of TgCkUg2 and to look at the genetic relationship between the eight Ugandan isolates (see Chapter 4). Thirty-two new primer pairs were designed across the genome (Table 5-3), around loci with a high prevalence of unique SNPs as well as more conserved loci (see 4.2.9 for more details on primer design, sequence analysis and tree building methodology). All primers were designed with a similar melting temperature (60-65°C) and all but the X-27 locus (which was not further used) were readily amplified using the same 25 µl mix (17.5 µl water, 1.25 µl of each primer (10 µM), 1.25 µl dNTPs (4 mM), 5 µl buffer and 0.125 GoTaq (Promega) plus 1 µl DNA (10 ng)) and touchdown PCR program (94°C 5 minutes; 10 cycles of 94°C 30s, 60°C 30s, 72°C 90s; 25 cycles of 94°C 30s, 56°C 30s, 72°C 90s; and finally 72°C for 7

minutes). PCR-products were purified and sequenced as described in 3.2.10. All loci were first sequenced in four strains (TgCkUg2, 5, 6, 8), and when informative polymorphisms were found, TgCkUg1, 3, 7 and 9 were also sequenced. A total of 230 sequences, including those generated previously (see section 4.3.5) were submitted to GenBank (Accession numbers: FI274744-FI274973).

Table 5-3. Primers used for complementary sequencing of additional Ugandan isolates

Locus	Chrom	Start pos.	Forward primer	Reverse primer
Ia-1	Ia	6169	CTTGAACATGCTGCCCATC	AACGAGACAATGCTCCAACA
Ib-2	Ib	188452	GAGGAATGGCAAGGTCTTCA	GTTGTACATCTCCGCCATCA
Ib-3	Ib	1638147	TCCAGGTCCAGCACTTCTTC	GTTCTTCGGCAGGGAAGTCT
II-4	II	74885	AGAAGCTCCTGTGAACGATGA	TGACGCAAACTCTACAAAAACG
II-5	II	1460420	TCCAGCAGTCCTTGACAATG	TCACAGTTCACTACGACAAAG
III-6	III	3650	CTGTCCCGACTGTTTCAGTG	CGTCTGATCCTAAGTACGCAGT
III-7	III	2457553	ATCAGAATTGTGCGAACACG	TGCATCGTGATAGCGAAGAC
IV-8	IV	8518	AGGGGATGCTGTCCAGTTTA	GAATTCGCCTGAACAAGCAG
IV-9	IV	1096127	GAGAAAGCATGGCGACAAAC	AGGAAAGCGAAATGCGTGT
V-10	V	2692201	TTGCTGCCTTATCACACAGG	ATAACTGGCCAGGCATCTCA
V-11	V	1466	CGAGAATGCGAACATGGAC	CACCACTACATCACTGGACGA
V-12	V	2500	TCATCGATATGCTGGTCTGC	TAATTACCGCGCTTCCAAGA
VI-13	VI	266750	GCCTGCTGCACGAATTAAG	TCTCGCTCCTGACCCTTG
VI-14	VI	2222524	CAAGCGATGTTGAGGGAACT	ATGTTTGCATGCCTTTTTGT
VIIa-15	VIIa	4388405	GAAACCTACCAGGCCACACT	CTCTCTCAGTTCCCGTGGTC
VIIa-16	VIIa	1061357	GGGAGCTTTTTCATTGTTGC	CCAGACAATTGTATCCGTCCT
VIIb-17	VIIb	4210302	CGCATTGTCCAGCCTTTC	AGAAGCAGCCACTGATCCAT
VIIb-18	VIIb	1506358	AAGCCTCCTGTGGGCATAC	GCGTCGAAAACAAGCAGAG
VIII-19	VIII	904801	CGCGAAACTTTCTTTTGGAT	CTCCCACTGAGTGCTTTCGT
VIII-20	VIII	2074625	GCAACTTTCAGACGGGAATC	CGAAGTGACACGTGGAGTTG
VIII-21	VIII	5857660	GACGACCGCAGAATTTTTGT	GCATCATGTGTGCGTCATTA
IX-22	IX	96023	AACAAGGCGAGGTGAAAGTG	GGCGTCGGTAGAAGCACATA
IX-23	IX	105770	GCCCACACCTTCTCAGTGAT	AGCGAGGTGCTGTTTTGAAC
IX-24	IX	5634308	CCGGCAAGAAAAGAAGTGAC	TCTGTCTCGTGAGAGCTTCG
X-25	X	30863	CGGTTTCCGTGTCACACTC	GAGCAGGCAAGGACAGAACT
X-26	X	4139809	CGCAGTCGCGTTGTAAAAAT	CTGGGATCCTTCTTGGCATA
X-27	X	7275823	CACATCGTAGGTGCTGTGCT	CGAAACGTGTGAGTACACAT
XI-28	XI	727	GCCAGACTCGAACTGGACTT	CGACGCAGCTGTCACTACAT
XI-29	XI	235815	GTACGGAGCCAGGGGATT	CATCGTGCTTGCGTCTCTAC
XI-30	XI	4244028	CTGTCAACCAAGCTGCTGAA	TACTTATGTTGGGCGCTGCT
XII-31	XII	72111	AAAGCAAGACAGAGGGTCGT	CTCGGTTCCCCTGTCTCTC
XII-32	XII	5851643	GACCGGACTGGTTCTATGCT	GCAACTTTAAGCGCGGTATC

Primers were designed for complementary sequencing of all the Ugandan isolates. The selected regions were distributed over all chromosomes and included loci with a high prevalence of unique SNPs as well as more conserved sections.

5.3 Results

5.3.1 Coverage and contig metrics

Three runs at the 454 GSFLX (Roche) generated approximately 4x coverage of the 61.6 Mb *T. gondii* genome. Human DNA contamination was unavoidable, but those reads could be removed bioinformatically. In total, 673,878 reads from TgCkUg2 were assembled into 67,013 contigs, ranging from 95 to 12,769 bp. The total contig length was 51.8 Mb, spanning over 84% of the genomic sequence. The average contig length was 774 bp, while the mean distance between contigs was 148 bp, and these figures were similar over all chromosomes (Table 5-4), thus there does not appear to be any coverage bias. Many of the gaps are due to the fact that regions that could not be unambiguously mapped, for example repeat regions, and deeper sequencing would be needed for accurate mapping of the remaining sequence.

Table 5-4. Contig and coverage summary for all chromosomes and the apicoplast.

	Chrom. length	Number of reads	Number of contigs	Total contig length	Percent covered by contigs	Read density[a]	Average distance between contigs	Average contig length	Median contig length
Ia	1,896,408	20,649	1,988	1,585,140	83.59%	13.03	155	797	595
Ib	1,956,324	20,583	2,127	1,639,876	83.82%	12.55	147	771	563
II	2,302,931	24,968	2,478	1,939,495	84.22%	12.87	147	783	561
III	2,470,845	26,771	2,730	2,060,909	83.41%	12.99	149	755	568
IV	2,576,468	27,510	2,885	2,150,897	83.48%	12.79	148	746	546
V	3,147,601	33,619	3,433	2,582,080	82.03%	13.02	163	752	558
VI	3,600,655	39,723	3,850	3,042,491	84.50%	13.06	146	790	575
VIIa	4,502,211	48,365	4,884	3,797,608	84.35%	12.74	145	778	572
VIIb	5,023,822	53,768	5,469	4,231,651	84.23%	12.71	144	774	562
VIII	6,923,375	75,308	7,501	5,851,305	84.52%	12.87	143	780	575
IX	6,384,456	72,298	7,039	5,337,365	83.60%	13.55	150	758	566
X	7,418,475	85,205	8,056	6,298,377	84.90%	13.53	140	782	575
XI	6,570,290	70,653	7,222	5,549,592	84.46%	12.73	142	768	568
XII	6,871,637	74,458	7,351	5,770,139	83.97%	12.90	150	785	564
Total	**61,645,498**	**673,878**	**67,013**	**51,836,925**	**84.09%**	**13.00**	**148**	**774**	**567**
Apico	34,996	3,055	1	25,117	71.77%	121.47	-	-	-

[a] Reads per kb over the contigs.

The read density, contig lengths, distance between contigs and proportion covered by contigs showed little variance between the 14 chromosomes. However, most of the apicoplast was represented by a single long contig with a read density over nine times higher than that for any of the chromosomes, indicating multiple copies of the apicoplast genome in each parasite.

5.3.2 SNP distribution

To determine the relative contribution of type II and type III regions to the recombinant isolate, the sequence generated for TgCkUg2 was aligned with published data for the reference strains Me49 (type II) and VEG (type III), and mapped to the 14 chromosomes. The total number of unambiguous SNPs identified, excluding repeat regions mapping to several locations, were 72,746. This number corresponds to about a quarter of the >300,000 known polymorphisms between Me49 and VEG (ToxoDB). The distribution of SNPs called against the two reference strains was highly disproportionate and indicated that each one of the chromosomes was inherited in its entirety from a single parental strain, without evidence of recombination events. Chromosomes II, IV, VI, VIIa, IX and X originated from a type II strain while chromosomes Ia, Ib, III, V, VIIb, VIII and XII originated from a type III parent, see Table 5-5.

Table 5-5. SNP summary for all chromosomes.

						Average distance		Density / kb	
	Length (Mb)	Type	SNPs: TgCkUg2 vs. Me49	SNPs: TgCkUg2 vs. VEG	SNPs: Unique to TgCkUg2	All SNPs	Unique SNPs	Major SNPs[a]	Minor SNPs[b]
Ia	1.9	III	128	2	7	13,842	270,915	0.067	0.005
Ib	2.0	III	1,483	1	10	1,309	195,632	0.758	0.006
II	2.3	II	73	4,370	52	512	44,287	1.898	0.057
III	2.5	III	4,224	27	73	571	33,847	1.710	0.040
IV	2.6	II	176	4,731	60	519	42,941	1.836	0.038
V	3.1	III	5,725	54	26	542	121,062	1.819	0.025
VI	3.6	II	168	1,985	99	1,599	36,370	0.551	0.043
VIIa	4.5	II	79	8,663	56	512	80,397	1.924	0.017
VIIb	5.0	III	5,530	62	41	892	122,532	1.101	0.021
VIII	6.9	III	6,775	123	158	981	43,819	0.979	0.041
IX	6.4	II	325	8,119	269	733	23,734	1.272	0.047
X	7.4	II	236	13,459	209	534	35,495	1.814	0.033
XI	6.6	ND	60	49	117	29,072	56,156	-	-
XII	6.9	III	4,809	58	75	1,390	91,622	0.7	0.005
Total	**61.6**	**II / III**	**29,791**	**41,703**	**1,252**	**847**	**49,238**	**1.264**	**0.023**

[a] Major SNPs = The dominant SNP type: SNPs against Me49 on type III chromosomes and against VEG on type II chromosomes.

[b] Minor SNPs = All SNPs where TgCkUg2 differs from the background type, for example on chromosome V this is the sum of SNPs against VEG (54) and unique SNPs (26), since all these represent nucleotides where TgCkUg2 is different from VEG.

The density of major SNPs was in the order of 0.5-2 SNPs per kb, while the minor SNP density was around 0.005-0.05 per kb. Also note the extremely low total SNP density on chromosomes Ia and XI (see Figures 5-7 and 5-8 for more on this).

Most SNPs were mapped against either Me49 or VEG, *i.e.* TgCkUg2 was identical to one of the reference strains but different from the other. Dissimilarities to Me49 are marked in blue and to VEG in green (see Table 5-2 and Figures 5-7 and 5-8), which means that a type II background is shown in green and regions inherited from type III in blue, and this corresponds to the SNP type colours in ToxoDB. In addition to these SNP types, 1,252 unique polymorphisms were found where TgCkUg2 was different from both Me49 and VEG, these SNPs are coloured in orange and are discussed in more detail in section 5.3.5. Due to the extreme scarcity of SNPs between types II and III on chromosome XI, which is dominated by type I SNPs (Figure 5-9), it has not been possible to ascertain the source of origin of this chromosome. In total, 226 SNPs were called over its length of >6.5 Mb, averaging one SNP per 29 kb. Over half of these (117 SNPs) were unique to TgCkUg2, while 49 positions were identical to Me49 and 60 to VEG. There was no obvious separation between the SNP types on this chromosome and therefore no evidence of a cross-over (Figure 5-7 (XI)).

Some chromosomes were completely dominated by one SNP type over the full length (Figure 5-7 (II)), but several chromosomes showed dramatic changes in major SNP density (Figure 5-7 (VIII)). The major SNPs correspond to the divergence between the type II and III lineages, and were therefore scarce in regions where these two lineages are highly similar, that is regions inherited from closely related parental strains in the original cross (see section 4.1.2). These drastic changes in SNP density were observed in chromosomes Ib, III, VI, VIIb, VIII, IX and XII, but in all cases this coincided with type I dominance and is thus consistent with a uniparental origin in TgCkUg2. Detailed graphs of the SNP densities in all chromosomes are provided in Appendix 11. Altogether, there were six type II and seven type III chromosomes encompassing 26.8 Mb and 28.3 Mb respectively, and in addition there was one chromosome (XI) which might be either II or III. This nearly equal contribution from both parental strains to the genome of TgCkUg2 is consistent with a single event of sexual recombination. An overview of the SNPs found in TgCkUg2 when aligned with Me49 and VEG is shown in Figure 5-8.

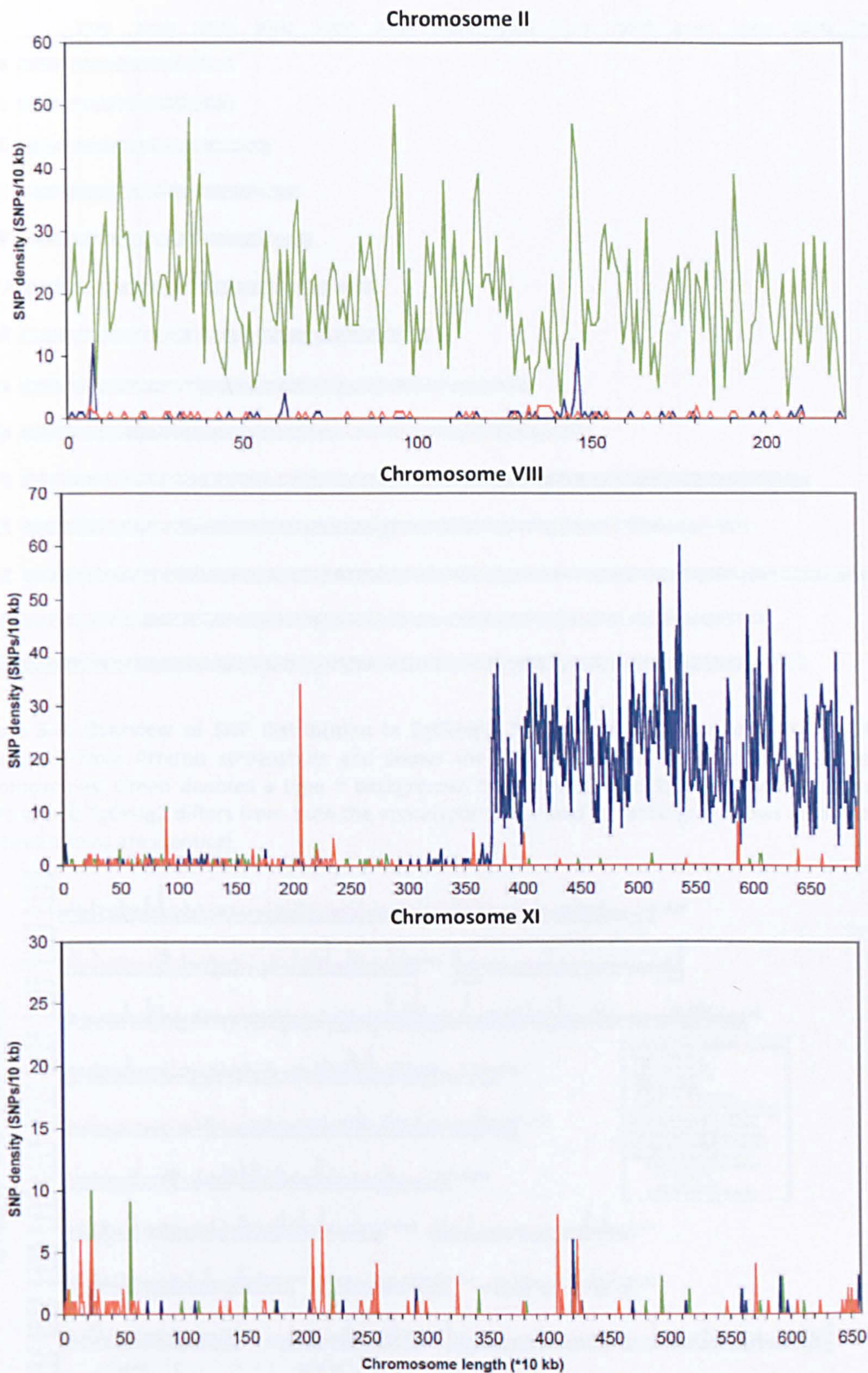


Figure 5-7. SNP density graphs of chromosomes II, VIII and XI. The X-axis represents the chromosome length *10 kb and the Y-axis shows the density of the different SNP types per 10 kb. Chromosome II has a type II background, (the major SNP type is green, *i.e.* TgCkUg2 vs. VEG) and VIII is a type III chromosome (dominated by blue SNPs (TgCkUg2 vs. Me49)). Note the radical change in SNP density on VIII and the constant low SNP density on XI, which can be explained by the low level of genetic difference between type II and III over these regions (as indicated by a predominance of type I SNPs in Figure 5-9). The SNP density graphs for all chromosomes were created in Excel and are shown in Appendix 11.

Figure 5-8. Overview of SNP distribution in TgCkUg2. The picture was created by K. Ashelford (Liverpool) from Artemis screenshots and shows the SNP distribution of TgCkUg2 over the 14 chromosomes. Green denotes a type II background, blue a type III background, orange indicates SNPs where TgCkUg2 differs from both the archetypal type II and III, while grey shows areas where all three strains are identical.

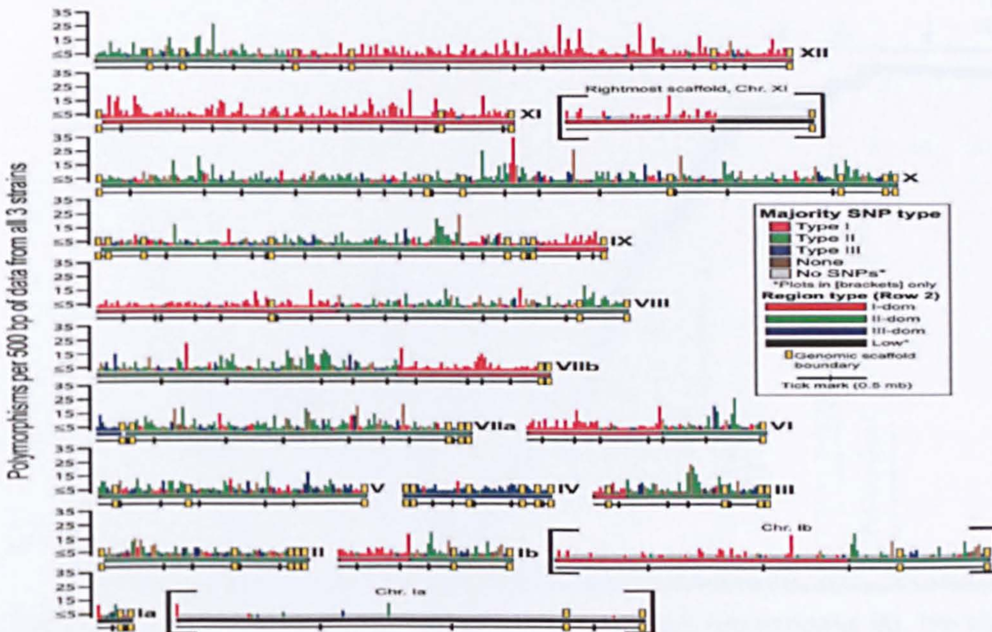


Figure 5-9. Overview of dominant SNP types in different chromosomal regions (reproduced with permission from Boyle et al, PNAS 103:10514-9). This picture was generated by Boyle et al (2006) through a comparison of the whole genome sequences of GT1 (I), Me49 (II) and VEG (III). Red indicates a dominance of type I SNPs, green type II and blue type III. When compared with the TgCkUg2 SNP map, the type I SNP dominant areas in red correspond very well with the areas void of SNPs (grey) in Figure 5-8.

Comparisons of the TgCkUg2 454 sequence data with the data reproduced in Figure 5-9 showed that regions without major SNPs in TgCkUg2 coincide with the regions dominated by type I SNPs (except in the case of chromosome Ia, where all the three lineages are highly similar). The chromosomes in Figure 5-9 are not depicted at a linear scale, however, and to enable a more detailed comparison Dr. Jon Boyle at the University of Pittsburgh reanalyzed the original sequence comparison data according to a different algorithm. For each position in each chromosome, the program looked for SNPs, if there was a type I SNP it was marked at the same level, type II SNPs were "plus one" and type III SNPs "minus one". Thus, regions dominated by type I SNPs were shown as straight lines, type II SNPs by ascending and type III by descending lines. An overlay of these images on the SNP density graphs at the same scale, revealed a perfect match between SNP scarce regions in TgCkUg2 and regions dominated by type I SNPs in the comparison between the three reference strains (Figure 5-10 and Appendix 11). This provides strong evidence that TgCkUg2 is the progeny of a cross between modern type II and III strains, where no chromosomal recombination took place.

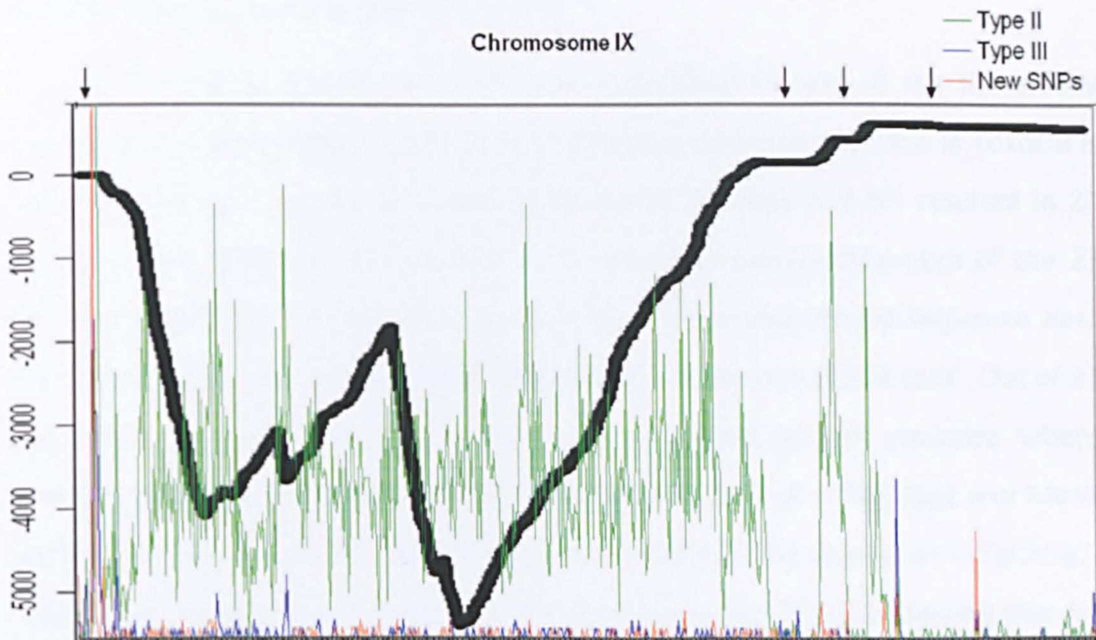


Figure 5-10. SNP density graph overlaid with SNP type graph (chromosome IX). The black line is straight at type I SNPs, rises +1 for every type II SNPs and falls -1 for every type III SNP (scale on the Y-axis). However, the horizontal parts of the line (indicated by arrows)match closely to the regions with low SNP density, which provides robust evidence that IX was inherited in its entirety from the UgII parent in the last cross. Note that the SNP density for TgCkUg2 is not affected by changes in major SNP type between type II and III – as long as type II and III are different TgCkUg2 differs from the non-parental type of each chromosome. Graphs for all chromosomes are provided in Appendix 11.

5.3.3 The apicoplast genome

The largest continuous sequence obtained in this sequencing effort was mapped to the apicoplast: this contig spanned over 25,000 bp which mapped to positions 4,871 to 29,940 of the sequenced 34,996 bp apicoplast genome from RH (ToxoDB). Apart from that, there was a high proportion of the apicoplast genome that was not mapped, which may be due to the AT-richness and presence of low complexity regions found in the genome of this organelle (for example poly-A stretches, see GenBank entry NC_001799), which complicates unambiguous alignments. As shown in Table 5-4, less than 72% of the apicoplast genome was covered by contigs, compared with 82-85% for the chromosomes. Interestingly, the read density of the apicoplast was considerably higher than the average read density for the chromosomes: 121 reads / kb for the apicoplast compared with an average of 13 reads / kb for the chromosomal sequences. The unbiased mechanism of 454 sequencing results in automatic quantification of amplified regions (Swaminathan et al., 2007), and the higher read density thus implies an average apicoplast genome copy number of 9 or 10.

The apicoplast genome has not been completed for any of the three fully sequenced reference strains, and the only sequence currently available in ToxoDB is from RH. Alignment of the apicoplast genomes of TgCkUg2 and RH resulted in 23 SNP calls over 25,069 bp of sequence. The sequence surrounding each of the 23 SNPs from TgCkUg2 was therefore compared with the fragmented sequence data from Me49, VEG and GT1 in the NCBI Trace Archive, using the BLAST tool[1]. Out of 23 high-confidence SNPs detected between TgCkUg2 and RH, all positions where matching data was available were confirmed to be identical in TgCkUg2 and Me49 and VEG, thus it was not possible to deduce the origin of the apicoplast in TgCkUg2. Analogous comparisons using the trace archive sequences for GT1 showed that for 20 of these 23 positions GT1 and TgCkUg2 were identical, while GT1 had the same allele as RH at the three remaining bases (Table 5-6).

---

1 http://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Nucleotides&PROGRAM=blastn&BLAST_SPEC=TraceArchive&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch

Table 5-6. Details of SNPs in the Apicoplast genome.

Position	RH	TgCkUg2	Me49	VEG	GT1
5442	G	T	T	T	T
7848	C	T	T	T	T
7969	G	A	A	A	A
8864	G	T	T	T	T
11253	-	A	not found	not found	A
12858	C	T	T	T	T
14396	C	T	T	T	T
14918	G	T	T	T	T
18101	C	T	T	T	*C*
18506	T	C	C	C	*T*
18669	C	A	A	A	A
18675	C	A	A	A	A
18701	C	T	T	T	T
19606	G	A	A	A	A
19959	G	A	not found	A	A
22554	G	A	A	A	A
24528	C	A	A	A	A
24794	A	G	G	G	*A*
25468	A	T	not found	T	T
25508	C	T	T	T	T
25962	C	T	T	T	T
26380	A	T	T	T	T
29819	G	TT	TT	TT	TT

Details of the 23 SNPs detected between TgCkUg2 and RH, and their counterparts in the Me49, VEG and GT1 strains.

5.3.4 Divergence time estimates

SNP data from TgCkUg2 was used to determine the time of the most recent common ancestor of the Ugandan types II and III (UgII and UgIII) and their respective reference strains. The divergence time for Me49/UgII is called T2 and that for VEG/UgIII is called T3 (see Figure 5-6). The calculations were made for all chromosomes separately and also for all chromosomes of the same genotype together and the two starting points were the *T. gondii* intron mutation rate of $1.94*10^{-8}$ mutations per nucleotide per year (Su et al., 2003), and an estimate of T1 of 150,000 years (Morrison, 2005).

For the first method all minor SNPs found in introns were retrieved for each chromosome, since the mutation rate used was calculated for *T. gondii* intronic regions. From the introns of the type II chromosomes (II, IV, VI, VIIa, IX and X) all SNPs where TgCkUg2 was different from Me49 were retrieved and for the type III chromosomes (Ia, Ib, III, V, VIIb, VIII and XII) the SNPs where TgCkUg2 differed from

VEG were used. In total, the type II intronic regions contained 381 SNPs over 1.13 Mb, compared with 229 SNPs over 1.28 Mb for type III, corresponding to SNP densities of $3.38*10^{-4}$ and $1.79*10^{-4}$ SNPs per base pair. Division of the SNP density with the intron mutation rate resulted in overall estimates of T2 and T3 of around 17,400 and 9,200 years, which implies that the Ugandan type II strains are older than their type III counterparts. The obtained results were used for calculation of the MRCA of the parental strains to all lineages (T1), and the estimates in both cases were much higher than the 150,000 years calculated by Morrison (2005). Calculations based on the data from the type II chromosomes of TgCkUg2 landed at around 1 million years, while calculations based on the type III parts gave an estimate of around half a million years. The results of these calculations are shown in Table 5-7.

Table 5-7. Divergence time estimates based on the *T. gondii* intron mutation rate.

	All intronic regions			Regions with a major SNP type[a]			
Type II chrom	**Intron length (bp)**	**Minor SNPs in introns**	**MRCA[b] years (T2)**	**Minor SNPs in introns**	**Major SNPs in introns**	**Intron length (bp)**	**MRCA[c] years (T1)**
II	89,003	34	19,691	34	1,800	89,003	973,654
IV	85,471	58	34,979	58	1,758	85,471	1,700,620
VI	152,385	65	21,987	21	656	46,869	1,570,910
VIIa	198,902	28	7,256	28	3,264	198,902	442,187
IX	266,876	117	22,598	65	2,667	175,224	1,484,722
X	334,181	79	12,186	79	5,044	334,181	807,328
Total II	**1,126,818**	**381**	**17,429**	**285**	**15,189**	**929,650**	**1,066,743**
Type III chrom	**Intron length (bp)**	**Minor SNPs in introns**	**MRCA[b] years (T3)**	**Minor SNPs in introns**	**Major SNPs in introns**	**Intron length (bp)**	**MRCA[c] years (T1)**
Ia[d]	80,522	2	1,280	...	NA	...	NA
Ib	85,805	8	4,806	6	466	38,069	392,610
III	82,660	9	5,612	6	1,381	66,453	270,064
V	119,536	15	6,468	15	2,015	119,536	383,720
VIIb	220,517	39	9,116	16	2,052	115,228	511,919
VIII	337,557	78	11,911	25	2,296	138,216	717,022
XII	352,605	78	11,403	25	1,394	78,096	638,808
Total III	**1,279,202**	**229**	**9,228**	**93**	**9,604**	**555,598**	**533,830**

[a] Whole chromosomes for II, IV, VIIa, X, V. No part of Ia. Ib >1.2 Mb, III < 1.9 Mb, VI >2.6 Mb, VIIb < 2.5 Mb, VIII > 4 Mb, IX 0.5-4 Mb, XII, 1.5 Mb. Only regions with a major SNP type can be used for this calculation since T1 is based on the number of major SNPs.
[b] (Minor SNPs in introns / Intron length) / $1.94*10^{-8}$
[c] (Region with major SNP type intron length * MRCA T2 or T3) / Major SNPs in introns.
[d] Since there are no major SNPs in chromosome Ia, T1 could not be calculated

The second approach related the number of 'major SNPs' to the number of 'minor SNPs', where the former type was assumed to represent the divergence between the type II and III lineages (beginning at T1), and the latter type the divergence between Ugandan strains and the reference strains (starting at time points T2 and T3), see Figure 5-6. This method is not applicable to regions without major SNPs, since the recombination events between the ancestral strains that gave rise to the present lineages occurred much more recently then did the MRCA of all the present strains, and types II and III have had much less time to acquire mutations in the regions they inherited jointly in this cross. Therefore, chromosomes Ia, XI and other chromosomal regions where type I SNPs are predominant (see Boyle 2006 and Figure 5-9) were excluded. All chromosomal regions included are listed in Table 5-8. These calculations resulted in considerably more recent divergence time estimates of 4,600 years for type II and 1,600 years for type III, which would correspond to an overall genomic mutation rate of approximately $1.29*10^{-8}$ per year.

Table 5-8. Divergence time estimate based on the relation between major and minor SNPs.

Type II chrom	Length[a]	Minor SNPs	Major SNPs	MRCA[b] years (T2)	Genomic mutation rate[c]
II	2,302,931	125	4,370	4,291	1.27E-08
IV	2,576,468	236	4,731	7,483	1.22E-08
VI >2.6 Mb	1,000,655	70	1,913	5,489	1.27E-08
VIIa	4,502,211	135	8,663	2,338	1.28E-08
IX 0.5-4 Mb	3,500,000	205	6,349	4,843	1.21E-08
X	7,418,475	445	13,459	4,960	1.21E-08
Total II	**21,300,740**	**1,216**	**39,845**	**4,578**	**1.25E-08**
Type III chrom	**Length[a] (bp)**	**Minor SNPs**	**Major SNPs**	**MRCA[b] years (T3)**	**Genomic mutation rate[c]**
Ib >1.2 Mb	756,324	6	1,375	655	1.21E-08
III <1.9 Mb	1,900,000	33	3,953	1,252	1.39E-08
V	3,147,601	80	5,725	2,096	1.21E-08
VIIb <2.5 Mb	2,500,000	43	5,069	1,272	1.35E-08
VIII >4 Mb	2,923,375	49	6,268	1,173	1.43E-08
XII <1.5 Mb	1,500,000	69	3,259	3,176	1.45E-08
Total III	**12,727,300**	**280**	**25,648**	**1,638**	**1.34E-08**

[a]Only chromosomal regions that were not jointly inherited by type II and type III in the lineage cross, *i.e.* regions that have a major SNP type, were included in this analysis. The length is in base pairs.
[b](Minor SNPs * 150,000) / Major SNPs. [c] (Minor SNPs / Length) / MRCA

As with the intron mutation rate method, substantial differences among chromosomes of the same genotype were observed. However, with the exception of chromosome VIIa, all estimates of T2 were higher than those of T3.

5.3.5 Novel SNPs and genes under selection

In the alignments of TgCkUg2 with Me49 and VEG, 1,252 positions were found where the two reference strains were identical but the Ugandan strain TgCkUg2 had a different base (marked in orange in Figures 5-7, 5-8, 5-10 and Appendix 11). These novel SNPs were dispersed all over the genome with an average distance between them of about 50 kb. Based on the SNP discovery rate with the coverage and cut-off criteria used, the real density of novel SNPs is however likely to be around four times higher (see 5.3.2). Novel SNPs were often found towards the chromosomal ends, in the subtelomeric regions. The proportion of all SNPs located in the outermost 20% of each chromosome (10% on each end) was 21.4%, close to the expected value, but a significantly higher proportion (38%) of the unique SNPs were found in these regions (P<<0.001). Several chromosomes had one or a few "hot-spots" with a high level of new mutations, see Figure 5-11.

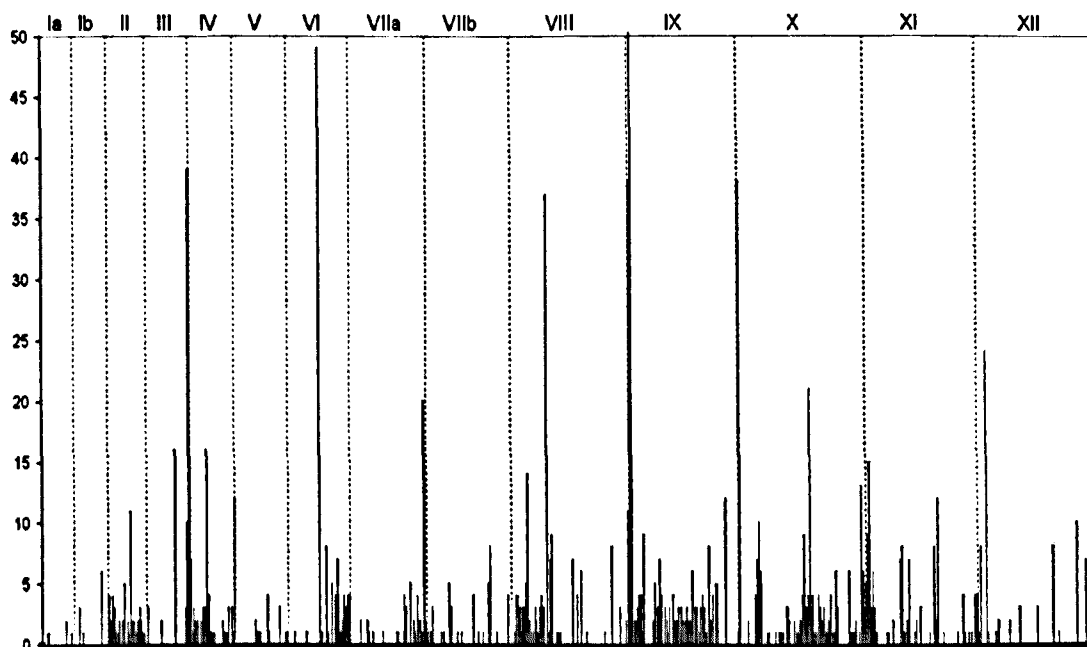


Figure 5-11. Distribution of novel SNPs in TgCkUg2. The graph shows the number of unique SNPs per 100 kb of sequence over the 14 chromosomes (separated by dashed lines). Several hot-spots with a high SNP density are seen and these were often located near the chromosomal ends, most prominently for chromosomes IX, X and III. Large peaks were also seen in central regions of chromosomes VI and VIII.

The highest concentrations of new SNPs were found in the subtelomeric regions of III, IX and X, and in more central regions of VI and VIII, but smaller clusters were found on all chromosomes except Ia. The majority of the novel SNPs (N=1,086, 86.7% of the total 1,252) were non-coding mutations in intergenic or intronic regions, but among the coding SNPs there were twice as many non-synonymous (N=111) as synonymous (N=55) mutations.

Fifteen genes, which had at least two new SNPs in the coding sequence, were identified and these are listed in Table 5-9.

Table 5-9. Genes with ≥2 novel SNPs in the coding region of TgCkUg2.

		Novel SNPs in TgCkUg2[b]		SNPs between the three lineages[b]		
Gene ID[a]	Chrom.	Syn	Non-syn	Syn	Non-syn	Protein description[a]
641.m01562	IV	1	1	12	86	SRS16B
641.m02553	IV	1	1	6	1	WD-40 repeat protein, putative
49.m03276	VI	1	1	0	4	ROP29
49.m03279	VI	2	8	1	13	Hypothetical
49.m03372	VI	1	1	13	3	Long chain fatty acid CoA ligase
55.m04829	VIIb	1	3	1	2	SRS26A
44.m02583	VIII	0	2	9	11	Hypothetical
44.m05903	VIII	0	2	7	8	Hypothetical
57.m01765	IX	0	2	143	231	Protein kinase domain containing
2.m00067	IX	3	7	0	0	Hypothetical
57.m01732	IX	0	2	7	8	Hypothetical
80.m02252	IX	1	1	4	2	Phosphoenolpyruvate carboxykinase, putative
42.m07434	X	0	2	0	0	Hypothetical
551.m00238	XII	3	6	8	44	ROP5
65.m00001	XII	4	0	9	6	NTPase I

[a]The GeneID is from ToxoDB 4.3, but the product description was updated from the more recent annotation in v.5.0.
[b]All listed genes have at two or more novel SNPs in TgCkUg2. The number of SNPs between the three lineages for each gene (data from ToxoDB) is listed for comparison. For example 57.m01765 is highly polymorphic, while 2.m00067 has a lot of novel SNPs but no divergence at all between the three archetypal strains.

Ten of the genes with novel SNPs in the CDS, including two SRS genes, had more non-synonymous than synonymous SNPs. Many of these genes, which appear to be under positive selection, have a high number of coding SNPs between all the lineages, but two of the hypothetical genes listed in Table 5-9 (57.m01765 and 42.m07434) did not have any previously identified SNPs (ToxoDB search). Three genes contained six or more amino acid changing mutations: hypothetical proteins 2.m00067 (IX), 49.m03279 (VI) and the rhoptry protein ROP5 (551.m00238, XII).

5.3.6 Local allelic variants and relationship between strains

In order to assess the level of divergence between all the eight Ugandan isolates and to investigate whether the novel SNPs in TgCkUg2 were conserved in the other strains as well, 32 new primer pairs were designed (see 5.2.6). These were distributed over all 14 chromosomes and while some of them were targeting novel SNP containing loci, others were simply distributed across the genome to provide an estimate of the overall divergence. Amplifications were primarily done in four strains: the recombinant (TgCkUg2), the type III (TgCkUg6) and two type II strains (TgCkUg5 and 8), and if informative polymorphisms were detected, the remaining type II strains were also amplified.

A high level of sequence homology was seen between the novel isolates from Uganda and the reference strains. The type III strain, TgCkUg6, was very closely related to the type III reference strain VEG as well as to the type III regions of TgCkUg2. In comparison to VEG, TgCkUg6 had 39 SNPs over 20.9 kb and most of these SNPs were concentrated to two loci: II-4 (10 SNPs over 598 bp) and VI-13 (18 SNPs over 368 bp). Apart from these regions the sequence identity between the type III strains was >99.9%. Locus II-4 consisted of non-coding sequence in the subtelomeric region of chromosome II, where TgCkUg6 shared some alleles with strains of genotype II (including Me49). The second locus, VI-13, included 220 bp of coding sequence from a SRS-gene (SRS22H, 49.m03110), where several new, non-synonymous SNPs were found for TgCkUg6, TgCkUg2 and three of the Ugandan type II strains. Interestingly, most new SNPs found within coding regions, including

SRS22H, Toxofilin (33.m02185) and SRS16B (641.m01562), resulted in amino acid changes, suggesting positive selection (Table 5-10).

Table 5-10. Amino acids at three polymorphic loci for Ugandan and reference strains.

Locus	SRS22H (VI-13)										Toxofilin			SRS16B
aa position	111	113	138	139	140	141	143	144	146	150	147	168	176	77
Me49 (II)	E	E	K	P	S	A	H	R	T	D	L	E	K	A
TgCkUg5	*	G	*	G	*	*	*	*	*	*	*	*	R	*
TgCkUg9	*	*	*	*	*	*	*	*	*	V	*	D	R	*
TgCkUg7	*	*	*	*	*	*	*	*	*	*	Q	D	R	E
TgCkUg1	*	*	*	*	*	*	*	*	*	*	Q	*	*	E
TgCkUg8	*	*	*	*	*	*	*	*	*	*	Q	*	*	E
TgCkUg3	*	*	*	*	*	*	*	*	*	V	Q	*	*	E
TgCkUg2	D	G	T	G	T	G	R	*	P	V	Q	*	*	E
TgCkUg6	D	G	N	G	*	G	R	S	P	V	E	*	*	T
VEG (III)	*	*	*	*	*	*	*	*	*	*	E	*	*	T
GT1 (I)	*	G	S	A	T	E	R	S	D	G	Q	*	R	A

Identities to the reference sequence shown at the top are indicated by asterisks (*).

The Ugandan type II isolates, including the type II regions of TgCkUg2, were closely related to Me49 (>99.5% sequence identity), but with some allelic variation. The new SNPs were largely concentrated at a few loci and many were shared among several Ugandan isolates, suggesting that these are local allelic variants. At one locus on chromosome IV (IV-8), several of the Ugandan type II strains shared sequence homology with GT1 (Table 5-11). This was only seen for a short stretch near the chromosomal end (positions 8805-8907), and this was the only instance where a high level of similarity with the type I reference strain was detected.

Table 5-11. Sequence similarity between Ugandan type II strains and the reference type I.

Locus	IV-8															
Me49 (II)	T	C	T	A	T	T	G	G	A	G	G	G	A	A	G	T
VEG (III)	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg6	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg5	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg9	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgCkUg7	*	*	A	T	C	C	C	*	*	T	*	*	G	C	C	*
TgCkUg1	C	A	A	T	C	C	C	C	C	T	A	A	G	C	C	A
TgCkUg2	C	A	A	T	C	C	C	C	C	T	A	A	G	C	C	A
TgCkUg3	C	A	A	T	C	C	C	C	C	T	A	A	G	C	C	A
TgCkUg8	C	A	A	T	C	C	C	C	C	T	A	A	G	C	C	A
GT1 (I)	C	A	A	T	C	C	C	C	C	T	A	A	G	C	C	A

Identities to the reference sequence shown at the top are indicated by asterisks (*).

Complementary Sanger sequencing confirmed the assignment of TgCkUg2 chromosomes according to the 454 SNP analysis, but possible cross-over events were detected in two regions without major SNPs. Chromosome VIII was identified as derived from type III, based on the major SNP density in the second half of the chromosome (Figure 5-7). However, for loci VIII-19 and VIII-20 (located at approximately 0.9 and 2.1 Mb) TgCkUg2 was more similar to Me49 than VEG and even contained four allelic variants, which were present in the Ugandan type II isolate TgCkUg8 as well. However, locus VIII-21 located around 5.8 Mb on the same chromosome identified TgCkUg2 as a type III strain. Similarly, comparison of TgCkUg2 and TgCkUg6 sequence for the VI-13 locus, located around 0.3 Mb, indicated the presence of a type III region in the otherwise type II derived chromosome VI. These results provide indications of chromosomal recombination in TgCkUg2 (Table 5-12).

Table 5-12. Different SNP types on TgCkUg2 chromosome VIII suggests recombination.

Locus Position	VIII-19 0.9 Mb							VIII-20 2.1 Mb				VIII-21 5.8 Mb		
Me49 (II)	A	T	G	C	T	T	T	C	T	G	A	G	C	G
TgUgCk5	*	*	*	*	*	*	*	*	*	*	*	*	*	*
TgUgCk8	*	*	*	*	A	C	*	T	*	A	*	*	*	*
TgUgCk2	*	*	*	*	A	C	*	T	*	A	*	A	A	T
TgUgCk6	T	*	C	G	*	*	C	*	C	*	T	A	A	T
VEG (III)	T	C	C	G	*	*	C	*	C	*	T	A	A	T

Identities to the reference sequence shown at the top are indicated by asterisks (*).

Few informative SNPs were detected on chromosome XI, but one unique allele within locus XI-28 was present in the recombinant as well as all of the type II Ugandan strains, but not in Me49, VEG or TgCkUg6. The fact that all of the Ugandan type II strains had this mutation implies that chromosome XI in TgCkUg2 was inherited from the type II parent.

Based on the sequence divergence over all amplified loci (including those described in Chapter 4), it was possible to resolve that TgCkUg5 and TgCkUg9 were the isolates most similar to Me49 and that TgCkUg3 was the strain most similar to the type II component of the recombinant TgCkUg2 (Figure 5-12).

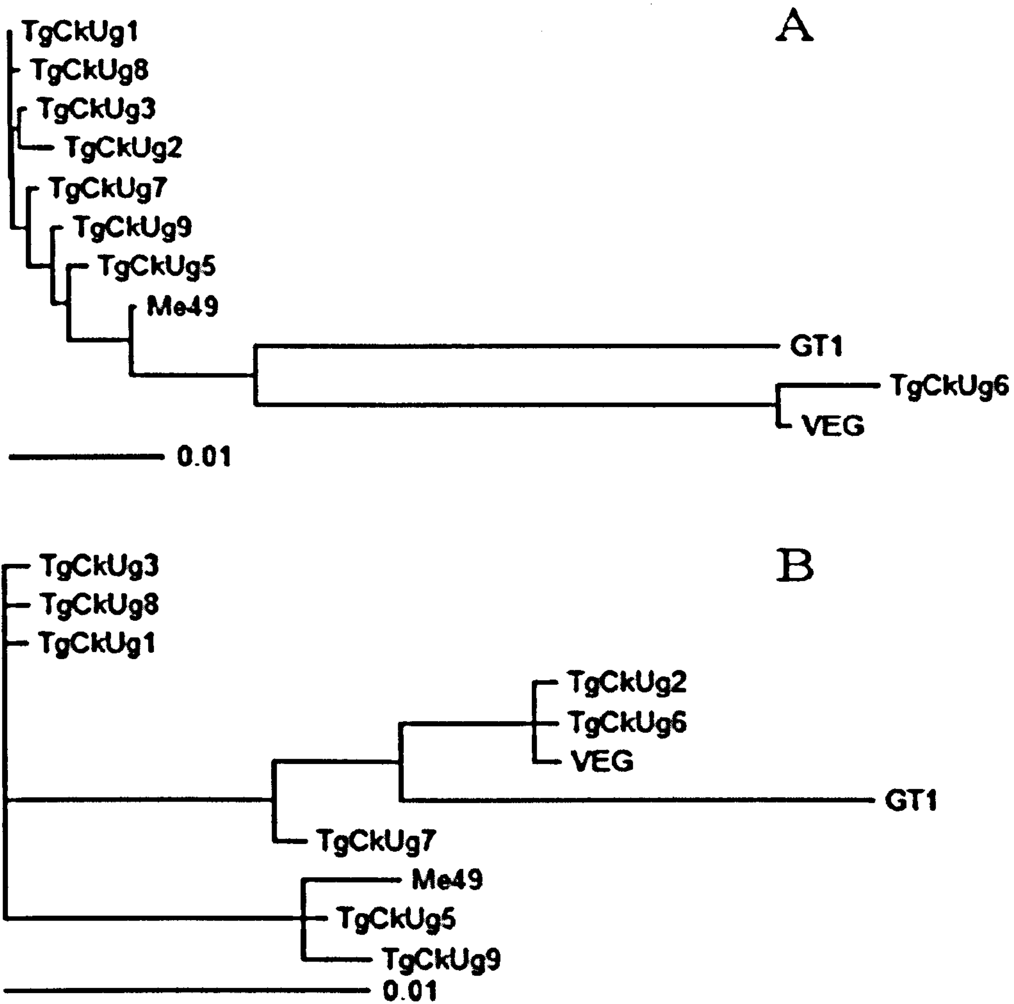


Figure 5-12. Phylogenetic trees of the relationship between Ugandan and reference strains. Trees based on sequence from chromosomes where TgCkUg2 were type II (A) and III (B). The type II part of the recombinant TgCkUg2 is most similar to TgCkUg3 and this strain can therefore be considered the putative type II parent in the cross. The slow-growing TgCkUg5 and 9 associate with Me49, while TgCkUg7 showed a higher level of similarity with non-type II strains. The methodology is described in section 4.2.9.

5.4 Discussion

5.4.1 Recombination through chromosome sorting

The first whole genome sequencing of a recombinant *T. gondii* strain revealed a nearly equal contribution of type II and III alleles, and the TgCkUg2 strain is likely to have arisen through a single recombination event between modern type II and III strains. The origin of most chromosomes could easily be deduced from the SNP density, since TgCkUg2 was either similar to VEG (type III origin) or Me49 (type II origin). However, extended sections with an apparent paucity of SNPs were observed, and these were found to be consistent with the type I SNP dominated regions in the comparison between the three archetypal lineages (Boyle et al., 2006). Even though novel SNPs were observed in the Ugandan strains, the whole genome sequence of TgCkUg2 showed almost perfect synteny with either Me49 or VEG for any given region. Thus, the sequencing of this recombinant isolate does not provide evidence of exotic or ancient strains in Uganda, but demonstrates recombination between two strains that are closely related to the lineages in Europe and North America.

Analysis of recombinant progeny from experimental crosses have shown that recombination can occur on all *T. gondii* chromosomes and that the frequency is related to the physical size (Khan et al., 2005b). The genetic map unit was estimated at around 104 kb/cM (Khan et al., 2005b), which is high compared with the human genome 1000 kb/cM (Sved et al., 2008) but represent a considerably lower recombination rate compared with the related apicomplexan *P. falciparum*, which has an estimated cross-over frequency of 17 kb/cM (Su et al., 1999). The study by Khan et al also detected an unexpectedly high prevalence of closely spaced crossovers, which may be due to gene conversions as seen in other protozoa (Le Blancq et al., 1988; Su et al., 1999; Palmer and Brayton, 2007). For TgCkUg2, no chromosomal crossovers were detected by whole genome SNP analysis. However, complementary sequencing of regions that had an unusually high number of minor SNPs showed that two loci on the otherwise type III derived chromosome VIII had a

high similarity to type II strains. This could indicate a regular crossover event, but the limited extension of the peaks (Figure 5-13) suggests that these may represent gene conversion events. Similarly, the identity of several Ugandan type II strains to the type I reference strain in the subtelomeric sequence of chromosome IV (Table 5-11), was limited to around 0.1 kb, and is likely to have arisen through one double cross-over rather than two separate events.

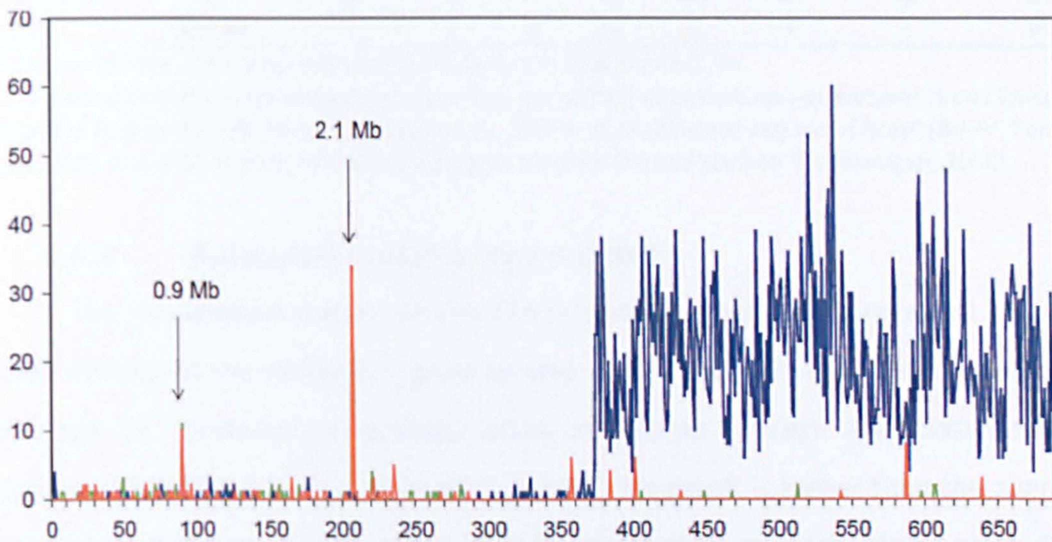


Figure 5-13. SNP density graph of chromosome VIII, with polymorphic loci suggestive of recombination. The area around the two polymorphic loci at 0.9 Mb and 2.1 Mb were sequenced for four Ugandan strains and showed that TgCkUg2 was identical to TgCkUg8 (a type II strain) for these loci, while it was clearly similar to type III for the region after 3.7 Mb, as shown by the predominance of blue SNPs.

Clonal propagation of non-archetypal genotypes has been observed in South America (Pena et al., 2008) and Asia (Dubey et al., 2007a; Dubey et al., 2007d), but comparison of the ten loci used in these studies showed that TgCkUg2 had a unique genotype, which has not been observed in any of the >900 strains analyzed to date (Dr. Chunlei Su, personal communication). However, apart from locus c22-8 on chromosome Ib, TgCkUg2 was highly similar to one strain from Ghana (Table 5-13).

Table 5-13. TgCkUg2 compared with circulating non-archetypal lineages and African strains.

Locus[a]	TgCkUg2	Vietnam, China + S. Amer.	BrI	BrII	BrIII	BrIV	Nigeria	Ghana1	Ghana 2
c22-8 (Ib)	III	II	u-1	I	II	u-1	III	II	II
c29-2 (III)	III	III	I	III	III	I	I	III	III
L358 (V)	III	II	I	I	III	I	I	II	III
PK1 (VI)	II	II	I	II	III	III	III	III	II
SAG1 (VIII)	III	u-1	I	I	I	u-1	u-1	u-1	II or III
SAG2 (VIII)	III	II	I	I	III	I	I	II	III
BTUB (IX)	II	III	I	III	III	III	III	III	II
GRA6 (X)	II	II	II	III	III	III	III	II	II
SAG3 (XII)	III	III	III	III	III	III	III	III	III
Apico	II or III	I	I	III	III	I	I	I	III

[a]Loci used in the worldwide population study by J.P. Dubey and C. Su.
Comparative strains were chosen because they are widely dispersed non-archetypal clonal lineages, found in Asia and South America (Dubey et al., 2007a, d) or different regions of Brazil (BrI-IV, Pena et al., 2008). In addition three African strains were used for the comparison (Velmurugan, 2008).

5.4.2 Apicoplast genome copy number

The quantitative nature of the 454 sequencing technology revealed that the read density of the apicoplast genome was more than nine times higher than the average for chromosomal contigs, which implies an average apicoplast genome copy number of 9.3 in this strain of *T. gondii*. This result is higher than the average apicoplast genome copy number of 5-6 reported in an early study using DNA hybridization (Fichera and Roos, 1997), but significantly lower than the average of 25 copies (occasionally up to 80) detected by microscope photon counting (Matsuzaki et al., 2001). These differences could be due to inherent differences between strains or methodological discrepancies, but the genome copy number may also vary over the different life cycle stages (Matsuzaki et al., 2001).

5.4.3 Level of divergence and relatedness to other strains

The estimated time of the most recent ancestors relies on the theory of the molecular clock: that the mutation rate in a species remains relatively constant over time (Zuckerkandl and Pauling, 1965; Bromham and Penny, 2003). Since there are no fossil records for *T. gondii*, calibration of the clock is difficult, and furthermore it is reasonable to assume that the accumulation of mutations in an organism with a life cycle that includes rapid division as well as prolonged periods of latency depends heavily on factors such as the transmission frequency and life span of

infected hosts. For the first calculations of the MRCA of the TgCkUg2 parental strains (UgII and UgIII) and their respective reference strains (Me49 and VEG), the density of minor SNPs was divided by the *T. gondii* intron mutation rate and the estimates for the two splits landed at around 17,400 and 9,200 years ago, and using these estimates of T2 and T3 resulted in an estimate of T1 of 500,000-1,000,000 years. Interestingly, these figures correspond relatively well with the estimates calculated by the Sibley group who suggest a MRCA for the clonal lineages around 10^4 years ago and of all *T. gondii* strains around 10^6 years ago (Khan et al., 2007). In the second approach, we used a previous estimate of the MRCA of all present *T. gondii* strains, which represents the time when the divergence between the different strains through mutations began (T1). It would not be possible to substitute this for the time estimates for the crossings which have given rise to the clonal types, since much of the nucleotide divergence had occurred previously in the ancestral strains, before participating in the cross. The resulting approximations of 4,600 for type II and 1,600 years for type III correspond to an overall genomic mutation rate of $1.29*10^{-8}$ SNPs per bp per year, which is about 68% of the rate calculated for introns. The two methods, which partly use different data sets, end up with noticeably different results, and until we have better tools for calibration of the clock it is difficult to tell when these strains began their divergence from their common ancestor. However, the relative timing is fairly consistent and it is clear that the type II split (T2) preceded that of type III (T3). Comparisons of the density of major and minor SNPs between the chromosomes derived from type II and III (Table 5-7) showed significant differences between the groups, where the level of major SNPs was significantly higher in the type III derived chromosomes, while the minor SNP density was higher in the type II chromosomes (Figure 5-14). Furthermore, the difference between the minor SNP density in introns was significantly higher ($P<0.01$) compared with the overall minor SNP density, for type II as well as for type III derived chromosomes (not shown in graph). This is consistent with the calculations which showed that the genomic mutation rate was lower compared with the intron mutation rate.

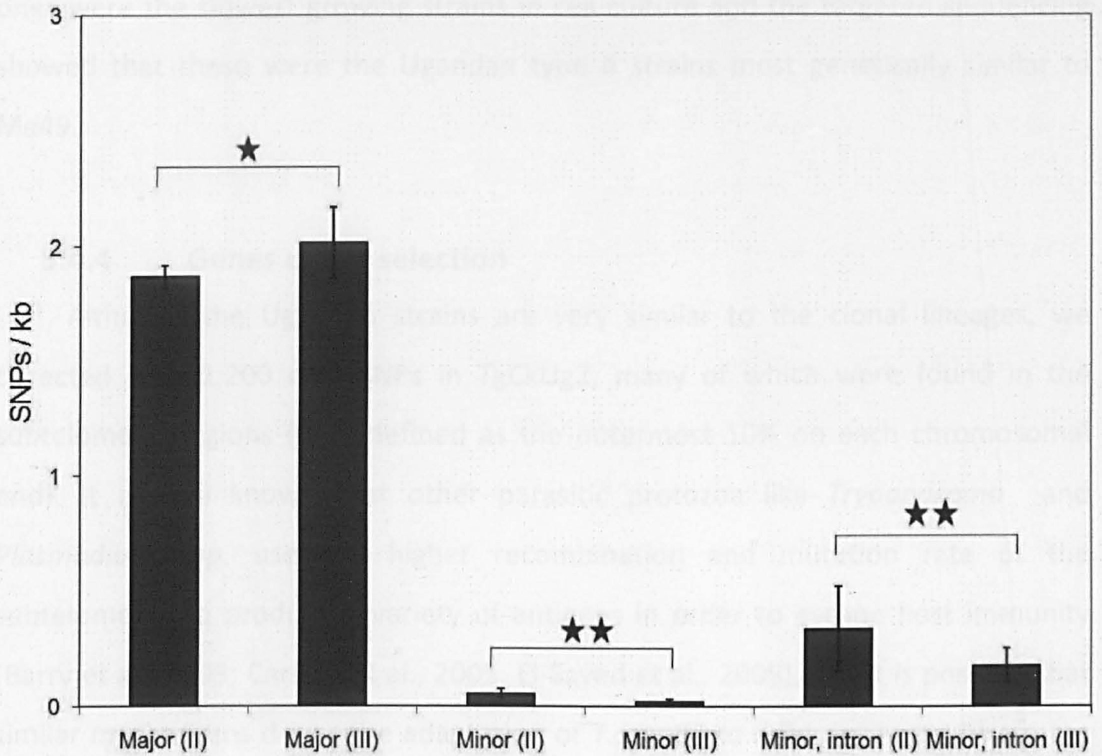


Figure 5-14. Average SNP density in type II and type III derived chromosomes. Average values and standard deviations for major and minor SNPs and minor SNPs in introns are shown for the type II and type III derived parts of TgCkUg2. These SNP densities were the basis for calculations of the divergence time estimates, and the real numbers are listed in Table 5-7 in section 5.3.4.

Sequencing of selected polymorphic regions in all the Ugandan strains made it possible to look at the relationship between the type II component of TgCkUg2 and the other type II strains. From this it was inferred that the closest relative to the type II parent of the recombinant strain was the fast-growing TgCkUg3. The original chicken host of TgCkUg3, Ch17, had the highest antibody titre of all chickens from which isolates were retrieved (1:160 compared with 1:80 for Ch1/TgCkUg1 and 1:40 for all the others) and the parasite DNA was readily amplified directly from the chicken tissues (Chapter 3). All this points towards a relatively high parasite burden in the chicken, and this may have increased the likelihood of recombination. Chickens 2 and 17, the hosts of TgCkUg2 and 3, both originated from the same district (Gayaza) as did the host of TgCkUg1. All the other isolates were obtained from chickens from Mulago and, as might be expected, there was a closer genetic relationship between *T. gondii* strains from chickens living in the vicinity of each other. The relatedness between TgCkUg1 and TgCkUg3 was also confirmed by the RAPD analysis, as was the closeness betweenTgCkUg5 and TgCkUg9. These latter

ones were the slowest growing strains in cell culture and the targeted sequencing showed that these were the Ugandan type II strains most genetically similar to Me49.

5.4.4 Genes under selection

Although the Ugandan strains are very similar to the clonal lineages, we detected over 1,200 new SNPs in TgCkUg2, many of which were found in the subtelomeric regions (here defined as the outermost 10% on each chromosomal end). It is well known that other parasitic protozoa like *Trypanosoma* and *Plasmodium spp.* use the higher recombination and mutation rate of the subtelomeres to produce a variety of antigens in order to escape host immunity (Barry et al., 2003; Carlton et al., 2005; El-Sayed et al., 2005), and it is possible that similar mechanisms drives the adaptation of *T. gondii* to different hosts. There are clusters of genes encoding surface antigens and other proteins important in the host-pathogen interaction towards the chromosomal ends of *T. gondii* (Jung et al., 2004 and ToxoDB), and future gene annotations and functional studies of hypothetical proteins and regulatory factors may appoint a biological meaning to the clusters of novel SNPs detected in TgCkUg2.

Pathogenicity in *T. gondii* is a multigenic trait and even though all the contributing genes have not yet been identified, a genome scale sequence analysis of recombinant and atypical isolates is a powerful tool for mapping genes important for survival and transmission in different natural environments. Early genetic linkage analysis of 6 RFLP markers in five strains showed that *T. gondii* virulence in mice was associated with the SAG1 gene on chromosome VIII (Howe et al., 1996). This analysis has later been followed up by experimental gene knockouts as well as large-scale bioinformatic analyses, and it is now clear that several surface antigens and secretory proteins affect the virulence of *T. gondii*. Knockout of the GRA2 gene led to attenuation of the virulent RH strain, while still allowing for chronic infection (Mercier et al., 1998), and a similar reduction in virulence was seen in SAG3 deficient mutants (Dzierszinski et al., 2000). Linkage analysis of recombinant progeny from an experimental cross led to the first identification of a major

virulence locus of the then called chromosome VII (Su et al., 2002), which has later been the focus of much attention under its new designation VIIa. In 2006, there was one report showing that knockout of MIC2 (located on VIIa) resulted in severely restrained attachment and invasion, completely eliminating the virulence in mice (Huynh and Carruthers, 2006), and later the same year two studies published in the same issue of Science revealed a major virulence locus on VIIa and proved through knockout experiments that the introduction of a virulent ROP18 allele in an avirulent strain increased the mouse mortality rate by 4-5 logs (Saeij et al., 2006; Taylor et al., 2006). Although novel SNPs were not present in the ROP18 gene (20.m03896) in our Ugandan isolates, positive selection was detected for several other genes active in the interface with the host, including two rhoptry antigens and two SRS genes (Table 5-9). In the study by Saeij and others, a total of four distinct virulence loci were identified in an experimental cross between types II and III, and these were located on chromosomes VIIa, VIIb, X and XII (Saeij et al., 2006). The TgCkUg2 strain possessed the avirulent allele for two of these loci: type III on VIIb and type II on X, and the virulent allele for two: type II on VIIa and type III on XII. The candidate virulence gene at the VIIa locus was ROP18, where genotype II is associated with virulence (Saeij et al., 2006), but no polymorphisms were seen for ROP18 in TgCkUg2, it was identical to Me49 and the Ugandan type II strains. SAG3 and ROP5 were the candidate genes on XII, where genotype III is the virulent type, and while no new SNPs were found for TgCkUg2 in the SAG3 gene, ROP5 was one of the genes with the highest number of novel SNPs (Table 5-9). The presence of six non-synonymous and three synonymous new SNPs suggests that ROP5 may be under diversifying selection in TgCkUg2.

While it is currently not possible to determine which loci are responsible for the intermediate phenotype of this recombinant strain the high density of SNP data from the current study could provide the basis for closer mapping of genotype-phenotype associations and genes under selection for niche adaptation within African isolates.

5.5 Conclusions

Whole genome sequencing of the recombinant TgCkUg2 revealed a nearly equal contribution from the type II and III parental strains. Recombination had mainly occurred through chromosome sorting and although indication of gene conversion events were recognized these were in regions where type II and III are highly similar and is probably of little biological significance. In experimental crosses intrachromosomal recombination is frequent and it is possible that the successful propagation in nature for this strain is due to the conservation of entire chromosomes, which may contain advantageous gene combinations.

The amount of data generated in this study proved that the *T. gondii* strains present in Uganda are indeed highly similar to the archetypal lineages, and they were calculated to have diverged from the North American reference strains around 10^4 years ago, with the type III divergence being more recent than type II. Novel SNPs were detected throughout the genome, but with a non-random distribution. A higher mutation frequency was detected in subtelomeric regions, but also in the coding sequence of a handful of genes. Interestingly, one of the genes indicated to be under positive selection (ROP5) have previously been indicated to be important for *T. gondii* virulence, and the alternate possession of virulent and avirulent alleles at different loci may explain the intermediate *in vivo* phenotype detected for this recombinant strain.

Chapter 6 General discussion and final conclusions

Infectious and parasitic diseases kill around ten million people globally every year, and approximately half of these cases occur in Africa making communicable diseases the most important cause of death on the continent (WHO, 2004). In addition, many infections confer a substantial level of disability, with enormous social and economic consequences. The economic impact of malaria alone has been estimated to decrease the gross national income in many African countries by around 10% (Sachs and Malaney, 2002) and HIV/AIDS adds another 2-4% growth reduction (Dixon et al., 2001). Research on infectious diseases is thus important in order to reduce morbidity and mortality as well as for fighting poverty, and a thorough understanding of the disease-causing microorganisms is crucial for the development of effective diagnostic, prevention and treatment strategies.

A microbe is considered a pathogen if it is capable of causing damage in its host (Casadevall and Pirofski, 1999), but the level of virulence depends on a combination of factors including the host immune response, concurrent infections and the genetic make-up of the microorganisms. Host-pathogen interactions represent very complex systems, where the same infectious agents can give rise to completely different outcomes, highlighting the importance of factors encountered by the microbe in the host upon infection. For instance, while most Epstein-Barr virus infections are asymptomatic, some otherwise healthy people develop infectious mononucleosis, while oral hairy leukoplakia is a common complication in severely immunocompromised patients (Greenspan et al., 1985) and lymphoproliferative cancers may develop upon co-infections with Epstein-Barr virus and malaria parasites (Chene et al., 2007). The recent spread of HIV has lead to a massive increase in diseases rarely seen before, caused by a number of viral, bacterial, protozoal and fungal pathogens which normally do not cause disease in immunocompetent hosts. In the pre-AIDS era *T. gondii* was recognised as an important veterinary pathogen and causative agent of severe congenital infections,

but cases of clinically important post-natally acquired toxoplasmosis were unusual before HIV-induced immunosuppressions became widespread (Moskowitz et al., 1983). High levels of *T. gondii* reactivation were seen in AIDS-patients in high-income countries before the introduction of HAART, but nowadays toxoplasmic encephalitis is uncommon in countries where appropriate antiretroviral treatment is provided (Miro et al., 2006). It is clear that the immune status of the host is highly important for the development of *T. gondii*, and this was confirmed through the investigation of latent and active toxoplasmosis in Ugandan HIV-patients described in Chapter 2 of this thesis, where nearly one out of five HIV-patients were shown to have active *T. gondii* infections. The affected patients had reached a more advanced stage of immunosuppression compared with the patients who had managed to keep their infections in the latent state, and these seropositive, PCR-negative individuals are at risk of reactivation at a later stage of AIDS. The seroprevalence rates of *T. gondii* exceed 50% in many countries in East-, West-, and Central Africa (Figure 2-1), and previous autopsy studies have shown that TE is an important cause of death in AIDS-patients in this part of the world (Byanyima, 1998; Lucas et al., 1993). Until efficient HIV-treatment is widely available in Africa opportunistic diseases including TE will continue to cause significant morbidity and mortality, unless they are properly diagnosed and treated. To achieve disease control, we must know the enemy. The major aim of the work presented in this thesis was therefore to investigate the genotypic and phenotypic properties of *T. gondii* strains from Africa, where the problem with HIV-associated toxoplasmosis is worst.

The success of an invasive microorganism is dependent of its capacity to evade the host immune response and pharmaceutic treatments, and different types of organisms use different approaches to meet the challenge of a changing environment. HIV-viruses are adaptation specialists that can escape from immune defence mechanisms as well as antiviral drugs through very high reproduction and mutation rates; every possible mutation in the genome can be generated on a daily basis in a single viremic patient (Richman et al., 2004). In bacteria, large blocks of genes, 'pathogenicity islands', can be transmitted horizontally through conjugation or with bacteriophages, and these may confer dramatic phenotype alterations such

as adaptations to an intracellular lifestyle (Gal-Mor and Finlay, 2006). Eukaryotic organisms typically have a slower mutation and replication rate compared with viruses and bacteria, and only rarely acquire genes through horizontal uptake (Opperdoes and Michels, 2007). Sexual recombination provides an opportunity of extensive genetic exchange and is the driving force of evolution in metazoa, however many protozoan parasites appear to use this option sparingly. A clonal population structure was first discovered in *Trypanosoma cruzi*, where isozyme analysis of 121 samples revealed clonal strains without evidence of mating, although some clones derived from the same individual or even the same insect vector (Tibayrenc et al., 1986). Different levels of clonality were later detected in several other organisms including *Leishmania*, *Giardia*, *Trypanosoma brucei*, *Entamoeba*, *Plasmodium* and *Toxoplasma* (Tibayrenc et al., 1990, 1991). The observed clonality could be due to inherent properties of the parasites favouring asexual transmission or self-fertilization over recombination, or it might be the result of downstream purifying selection of non-recombinant progeny (Tibayrenc, 2002). These explanations may not be mutually exclusive and could vary between species, for example *Plasmodium* goes through an obligatory sexual stage in the mosquito, while *Toxoplasma* can be transmitted directly between intermediate hosts, which enable prolonged transmission of true clones. It is noteworthy that a predominantly clonal population structure does not exclude sexual recombination. In the case of *T. gondii*, a limited number of clonal lineages are found over vast geographic areas, but a high level of non-clonal strains have been found in South America (Pena et al., 2008), which may be due to increased mating opportunities or selection of a diversity of strains. As more areas and host species are being investigated, more diversity has been unravelled, and the work described in this thesis provides the first characterization of an African *T. gondii* population. The focus of this work was on strains present in the domestic cycle, since these are most relevant for HIV-associated toxoplasmosis.

The discovery of multiple infections in five Ugandan chickens gave the first indication that mating opportunities between genetically distinct genotypes may be more frequent in the human / domestic animal cycle in Africa compared with

Europe or North America. Small scale rearing of animals like goats and chickens around the houses is commonplace and the transmission between these animals and domestic cats can easily take place. The subsequent isolation of a recombinant strain confirmed that sexually derived strains were present in Uganda. Recombination is difficult to detect if it occurs between strains with a similar genetic background, such as two type II strains. However, the deep sequencing performed on TgCkUg2 revealed different levels of novel SNPs among the chromosomes derived from the same lineage. For example, the MRCA of the Ugandan type II and Me49 was estimated to be 3-5 times higher based on chromosome IV compared with VIIa, using the two different methods (Tables 5-7 and 5-8) and this may be an indication of a previous chromosome sorting event between two type II strains. Furthermore, the gene conversion event where several (but not all) Ugandan type II strains shared SNPs with GT1 indicate that this recombination occurred at a time point prior to the type II/III cross that gave rise to TgCkUg2. Thus, lack of opportunity for sexual recombination, which has been suggested to be an important explanation for the dominance of three clonal types in Europe and North America (Sibley and Boothroyd, 1992, Dardé, 1996, Ajzenberg, et al., 2002), does not appear to hold true for Africa.

Nevertheless, the whole genome analysis of this strain and sequencing of >20 kb from each of the eight additional Ugandan strains showed a remarkable level of similarity with the reference strains isolated decades ago in North America. Unlike the findings from South America, the Ugandan strains were highly conserved and although the recombinant strain provided evidence of the sexual cycle, all the 14 chromosomes were nearly identical to either type II or III, and virtually no mixing of genes from different origin was seen on the same chromosome. Indications of recombinatory events were detected for small regions located on chromosomes VI and VIII, but these occurred in type I dominated regions where type II and III share a high level of similarity and did therefore not alter the overall gene composition of the chromosomes. Although diverse genotypes are likely to be generated through sexual recombination in this setting, the environment does not seem to favour the persistence of multiple diverse genotypes. A prerequisite for recombinant strains to

be found is their ability to survive and propagate, and in order to persist in nature they must be able to compete with the established lineages. *T. gondii* is known to undergo sexual recombination given the chance, and the average map unit size has been calculated to 104 kb/cM (Khan et al., 2005b). The data from TgCkUg2 shows evidence of at least three past recombination events, but the gene content on each chromosome has remained intact. A single cross yields countless genetically distinct progeny, but only a handful may survive in nature, where the predominant strain type in a defined environment can be assumed to be the one that is best adapted to the prevailing conditions. Thus, in contrast to the upstream inhibition of sexual recombination proposed previously (Ferguson, 2002) the data obtained in this study favour the model of downstream selection through environmental pressure rather than absence of opportunity for sexual recombination. This conclusion is in agreement with a recent review on the subject, where the population structure is described as "punctuated by clonal sweeps" when a new recombinant happens to be superiorly adapted to the prevailing conditions (Grigg and Sundar, 2009).

The research presented in this thesis revealed that the *T. gondii* strains present in humans and domestic animals in Uganda are closely related to the clonal lineages, and began diverging from these approximately 10,000 years ago. Although recombination occurs the selective pressures appear to favour conservation of the genetic make-up of the clonal strains, but some genetic drift inevitably occurs over time. But are these good representatives for African strains? The strains were isolated from the area in and around Kampala, and they are likely to be representative of the strain population important for the HIV-associated toxoplasmosis in this region. However, it is possible that other strains are present elsewhere in Africa. A recent serotyping study reported a predominance of non-type II strains in West Africa (Ivory Coast, Congo, Angola, Cameroon and Gabon) and several samples reacted with both peptides, which may indicate multiple infection in human patients (Sousa et al., 2008). Furthermore, a recent case report found a type I/III recombinant strain causing severe toxoplasmic encephalitis and chorioretinitis in a HIV-patient from Ghana (Genot et al., 2007). None of these studies provide much detail on the genetic composition of the strains involved, but

the latter emphasize the clinical importance of the *T. gondii* genotype, and the disease severity often associated with atypical strains. No study has yet looked further afield into the African sylvatic cycle, where different selective pressures may shape the population structure and transmission may be driven by other feline species.

The studies presented in this thesis extend the understanding of the *T. gondii* population structure and provides the first in depth characterization of African strains. The whole genome sequencing project was one of the first to utilize a new generation sequencing platform for a protozoan genome, and it was the first large scale sequencing of any *T. gondii* strain apart from the three clonal references. Evidence of recombination through chromosome sorting, in an organism where meiotic recombination is frequent in experimental crosses, was a significant result, and furthermore the novel African alleles and genes under selection provide clues to the selective pressures acting in this environment. The unique situation with a fully sequenced genome from a natural recombinant strain and the isolation of its putative parents provide excellent opportunities for further studies of the genetic basis for virulence and transmission of *T. gondii*.

References

Abdel-Hameed, A.A., 1991. Sero-epidemiology of toxoplasmosis in Gezira, Sudan. J Trop Med Hyg 94, 329-332.

Abu-Zeid, Y.A., 2002. Serological evidence for remarkably variable prevalence rates of *Toxoplasma gondii* in children of major residential areas in United Arab Emirates. Acta Trop 83, 63-69.

Adou-Bryn, K.D., Ouhon, J., Nemer, J., Yapo, C.G., Assoumou, A., 2004. [Serological survey of acquired toxoplasmosis in women of child-bearing age in Yopougon (Abidjan, Cote d'Ivoire)]. Bull Soc Pathol Exot 97, 345-348.

Agrawal, A.F., 2006. Evolution of sex: why do organisms shuffle their genotypes? Curr Biol 16, 696-704.

Ahmed, H.J., Mohammed, H.H., Yusuf, M.W., Ahmed, S.F., Huldt, G., 1988. Human toxoplasmosis in Somalia. Prevalence of *Toxoplasma* antibodies in a village in the lower Scebelli region and in Mogadishu. Trans R Soc Trop Med Hyg 82, 330-332.

Ahn, H.J., Kim, S., Kim, H.E., Nam, H.W., 2006. Interactions between secreted GRA proteins and host cell proteins across the paratitophorous vacuolar membrane in the parasitism of *Toxoplasma gondii*. Korean J Parasitol 44, 303-312.

Ajzenberg, D., Cogne, N., Paris, L., Bessieres, M.H., Thulliez, P., Filisetti, D., Pelloux, H., Marty, P., Darde, M.L., 2002. Genotype of 86 *Toxoplasma gondii* isolates associated with human congenital toxoplasmosis, and correlation with clinical findings. J Infect Dis 186, 684-689.

Ajzenberg, D., Banuls, A.L., Su, C., Dumetre, A., Demar, M., Carme, B., Darde, M.L., 2004. Genetic diversity, clonality and sexuality in *Toxoplasma gondii*. Int J Parasitol 34, 1185-1196.

Ajzenberg, D., Dumetre, A., Darde, M.L., 2005. Multiplex PCR for typing strains of *Toxoplasma gondii*. J Clin Microbiol 43, 1940-1943.

Alvarado-Esquivel, C., Sifuentes-Alvarez, A., Narro-Duarte, S.G., Estrada-Martinez, S., Diaz-Garcia, J.H., Liesenfeld, O., Martinez-Garcia, S.A., Canales-Molina, A., 2006. Seroepidemiology of *Toxoplasma gondii* infection in pregnant women in a public hospital in northern Mexico. BMC Infect Dis 6, 113.

Amogne, W., Teshager, G., Zenebe, G., 2006. Central nervous system toxoplasmosis in adult Ethiopians. Ethiop Med J 44, 113-120.

Antinori, A., Larussa, D., Cingolani, A., Lorenzini, P., Bossolasco, S., Finazzi, M.G., Bongiovanni, M., Guaraldi, G., Grisetti, S., Vigo, B., Gigli, B., Mariano, A., Dalle Nogare, E.R., De Marco, M., Moretti, F., Corsi, P., Abrescia, N., Rellecati, P., Castagna, A., Mussini, C., Ammassari, A., Cinque, P., d'Arminio Monforte, A., 2004. Prevalence, associated factors, and prognostic determinants of AIDS-related toxoplasmic encephalitis in the era of advanced highly active antiretroviral therapy. Clin Infect Dis 39, 1681-1691.

Aramini, J.J., Stephen, C., Dubey, J.P., Engelstoft, C., Schwantje, H., Ribble, C.S., 1999. Potential contamination of drinking water with *Toxoplasma gondii* oocysts. Epidemiol Infect 122, 305-315.

Araujo, F., Slifer, T., Kim, S., 1997. Chronic infection with *Toxoplasma gondii* does not prevent acute disease or colonization of the brain with tissue cysts following reinfection with different strains of the parasite. J Parasitol 83, 521-522.

Araujo, F.G., Slifer, T., 2003. Different strains of *Toxoplasma gondii* induce different cytokine responses in CBA/Ca mice. Infect Immun 71, 4171-4174.

Arene, F.O., 1986. The prevalence of toxoplasmosis among inhabitants of the Niger Delta. Folia Parasitol (Praha) 33, 311-314.

Arrizabalaga, G., Boothroyd, J.C., 2004. Role of calcium during *Toxoplasma gondii* invasion and egress. Int J Parasitol 34, 361-368.

Aspinall, T.V., Guy, E.C., Roberts, K.E., Joynson, D.H., Hyde, J.E., Sims, P.F., 2003. Molecular evidence for multiple *Toxoplasma gondii* infections in individual patients in England and Wales: public health implications. Int J Parasitol 33, 97-103.

Bachmeyer, C., Mouchnino, G., Thulliez, P., Blum, L., 2006. Congenital toxoplasmosis from an HIV-infected woman as a result of reactivation. J Infect 52, e55-57.

Bahia-Oliveira, L.M., Jones, J.L., Azevedo-Silva, J., Alves, C.C., Orefice, F., Addiss, D.G., 2003. Highly endemic, waterborne toxoplasmosis in north Rio de Janeiro state, Brazil. Emerg Infect Dis 9, 55-62.

Baril, L., Ancelle, T., Goulet, V., Thulliez, P., Tirard-Fleury, V., Carme, B., 1999. Risk factors for *Toxoplasma* infection in pregnancy: a case-control study in France. Scand J Infect Dis 31, 305-309.

Barragan, A., Sibley, L.D., 2003. Migration of *Toxoplasma gondii* across biological barriers. Trends Microbiol 11, 426-430.

Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W., Montagnier, L., 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). Science 220, 868-871.

Barry, J.D., Ginger, M.L., Burton, P., McCulloch, R., 2003. Why are parasite contingency genes often associated with telomeres? Int J Parasitol 33, 29-45.

Bastien, P., 2002. Molecular diagnosis of toxoplasmosis. Trans R Soc Trop Med Hyg 96 Suppl 1, S205-215.

Benard, A., Petersen, E., Salamon, R., Chene, G., Gilbert, R., Salmi, L.R., 2008. Survey of European programmes for the epidemiological surveillance of congenital toxoplasmosis. Euro Surveill 13, 1-7.

Besteiro, S., Michelin, A., Poncet, J., Dubremetz, J.F., Lebrun, M., 2009. Export of a *Toxoplasma gondii* Rhoptry Neck Protein Complex at the Host Cell Membrane to Form the Moving Junction during Invasion. PLoS Pathog 5, e1000309.

Beverley, J.K., 1959. Congenital transmission of toxoplasmosis through successive generations of mice. Nature 183, 1348-1349.

Black, M.W., Boothroyd, J.C., 2000. Lytic cycle of *Toxoplasma gondii*. Microbiol Mol Biol Rev 64, 607-623.

Blackston, C.R., Dubey, J.P., Dotson, E., Su, C., Thulliez, P., Sibley, D., Lehmann, T., 2001. High-resolution typing of *Toxoplasma gondii* using microsatellite loci. J Parasitol 87, 1472-1475.

Blumenschein, T.M., Friedrich, N., Childs, R.A., Saouros, S., Carpenter, E.P., Campanero-Rhodes, M.A., Simpson, P., Chai, W., Koutroukides, T., Blackman, M.J., Feizi, T., Soldati-Favre, D., Matthews, S., 2007. Atomic resolution insight into host cell recognition by *Toxoplasma gondii*. Embo J 26, 2808-2820.

Bobic, B., Nikolic, A., Klun, I., Vujanic, M., Djurkovic-Djakovic, O., 2007. Undercooked meat consumption remains the major risk factor for *Toxoplasma* infection in Serbia. Parassitologia 49, 227-230.

Boothroyd, J.C., 1993. Population biology of *Toxoplasma*: clonality, virulence, and speciation (or not). Infect Agents Dis 2, 100-102.

Boothroyd, J.C., Grigg, M.E., 2002. Population biology of *Toxoplasma gondii* and its relevance to human infection: do different strains cause different disease? Curr Opin Microbiol 5, 438-442.

Boothroyd, J.C., Dubremetz, J.F., 2007. Kiss and spit: the dual roles of *Toxoplasma* rhoptries. Nat Rev Microbiol 6, 79-88.

Bossi, P., Bricaire, F., 2004. Severe acute disseminated toxoplasmosis. Lancet 364, 579.

Bowie, W.R., King, A.S., Werker, D.H., Isaac-Renton, J.L., Bell, A., Eng, S.B., Marion, S.A., 1997. Outbreak of toxoplasmosis associated with municipal drinking water. The BC *Toxoplasma* Investigation Team. Lancet 350, 173-177.

Boyle, J.P., Rajasekar, B., Saeij, J.P., Ajioka, J.W., Berriman, M., Paulsen, I., Roos, D.S., Sibley, L.D., White, M.W., Boothroyd, J.C., 2006. Just one cross appears capable of dramatically altering the population biology of a eukaryotic pathogen like *Toxoplasma gondii*. Proc Natl Acad Sci U S A 103, 10514-10519.

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

Bradley, P.J., Sibley, L.D., 2007. Rhoptries: an arsenal of secreted virulence factors. Curr Opin Microbiol 10, 582-587.

Brindle, R., Holliman, R., Gilks, C., Waiyaki, P., 1991. *Toxoplasma* antibodies in HIV-positive patients from Nairobi. Trans R Soc Trop Med Hyg 85, 750-751.

Bromham, L., Penny, D., 2003. The modern molecular clock. Nat Rev Genet 4, 216-224.

Burg, J.L., Grover, C.M., Pouletty, P., Boothroyd, J.C., 1989. Direct and sensitive detection of a pathogenic protozoan, *Toxoplasma gondii*, by polymerase chain reaction. J Clin Microbiol 27, 1787-1792.

Byanyima, K.R., 1998. Cranial CT findings in the HIV-1 infected patients in Mulago Hospital. M Med dissertation, B99C Albert Cook Library, Makerere University.

Candolfi, E., Berg, M., Kien, T., 1993. [Prevalence of toxoplasmosis in Pointe-Noire in Congo. Study of the sampling of 310 subjects]. Bull Soc Pathol Exot 86, 358-362.

Carlton, J., Silva, J., Hall, N., 2005. The genome of model malaria parasites, and comparative genomics. Curr Issues Mol Biol 7, 23-37.

Carme, B., Bissuel, F., Ajzenberg, D., Bouyne, R., Aznar, C., Demar, M., Bichat, S., Louvel, D., Bourbigot, A.M., Peneau, C., Neron, P., Darde, M.L., 2002. Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. J Clin Microbiol 40, 4037-4044.

Carruthers, V., Boothroyd, J.C., 2007. Pulling together: an integrated model of *Toxoplasma* cell invasion. Curr Opin Microbiol 10, 83-89.

Carruthers, V.B., 1999. Armed and dangerous: *Toxoplasma gondii* uses an arsenal of secretory proteins to infect host cells. Parasitol Int 48, 1-10.

Casadevall, A., Pirofski, L.A., 1999. Host-pathogen interactions: redefining the basic concepts of virulence and pathogenicity. Infect Immun 67, 3703-3713.

CDC, 1993. Revised Classification System for HIV Infection and Expanded Surveillance Case Definition for AIDS Among Adolescents and Adults http://www.cdc.gov/mmwr/preview/mmwrhtml/00018871.htm.

Chene, A., Donati, D., Guerreiro-Cacais, A.O., Levitsky, V., Chen, Q., Falk, K.I., Orem, J., Kironde, F., Wahlgren, M., Bejarano, M.T., 2007. A molecular link between malaria and Epstein-Barr virus reactivation. PLoS Pathog 3, e80.

Chun, T.W., Fauci, A.S., 1999. Latent reservoirs of HIV: obstacles to the eradication of virus. Proc Natl Acad Sci U S A 96, 10958-10961.

CIA, 2006. Country Information - Uganda. https://www.cia.gov/library/publications/the-world-factbook/print/ug.html

Clavel, F., 1987. HIV-2, the West African AIDS virus. AIDS 1, 135-140.

Clumeck, N., Sonnet, J., Taelman, H., Mascart-Lemone, F., De Bruyere, M., Vandeperre, P., Dasnoy, J., Marcelis, L., Lamy, M., Jonas, C., Eyckmans, L., Noel, M., Vanhaeverbeek, M., Butzler, J.P., 1984. Acquired immunodeficiency syndrome in African patients. N Engl J Med 310, 492-497.

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

Cornelissen, A.W., Overdulve, J.P., 1981. Sex determination and sex differentiation in Isospora (*Toxoplasma*) *gondii*. Trans R Soc Trop Med Hyg 75, 594-596.

Cornelissen, A.W., Overdulve, J.P., van der Ploeg, M., 1984. Determination of nuclear DNA of five eucoccidian parasites, Isospora (*Toxoplasma*) *gondii*, Sarcocystis cruzi, Eimeria tenella, E. acervulina and Plasmodium berghei, with special reference to gamontogenesis and meiosis in I. (T.) *gondii*. Parasitology 88, 531-553.

Costa, J.M., Darde, M.L., Assouline, B., Vidaud, M., Bretagne, S., 1997. Microsatellite in the beta-tubulin gene of *Toxoplasma gondii* as a new genetic marker for use in direct screening of amniotic fluids. J Clin Microbiol 35, 2542-2545.

Cristina, N., Darde, M.L., Boudin, C., Tavernier, G., Pestre-Alexandre, M., Ambroise-Thomas, P., 1995. A DNA fingerprinting method for individual characterization of *Toxoplasma gondii* strains: combination with

isoenzymatic characters for determination of linkage groups. Parasitol Res 81, 32-37.

Cruz, M.L., Cardoso, C.A., Saavedra, M.C., Santos, E.D., Melino, T., 2007. Congenital toxoplasmosis infection in an infant born to an HIV-1-infected mother. Braz J Infect Dis 11, 610-611.

Dabritz, H.A., Miller, M.A., Gardner, I.A., Packham, A.E., Atwill, E.R., Conrad, P.A., 2008. Risk factors for *Toxoplasma gondii* infection in wild rodents from central coastal California and a review of T. *gondii* prevalence in rodents. J Parasitol 94, 675-683.

Dao, A., Fortier, B., Soete, M., Plenat, F., Dubremetz, J.F., 2001. Successful reinfection of chronically infected mice by a different *Toxoplasma gondii* genotype. Int J Parasitol 31, 63-65.

Darde, M.L., Bouteille, B., Pestre-Alexandre, M., 1988. Isoenzymic characterization of seven strains of *Toxoplasma gondii* by isoelectrofocusing in polyacrylamide gels. Am J Trop Med Hyg 39, 551-558.

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

Darde, M.L., 1996. Biodiversity in *Toxoplasma gondii*. Curr Top Microbiol Immunol 219, 27-41.

Darde, M.L., Villena, I., Pinon, J.M., Beguinot, I., 1998. Severe toxoplasmosis caused by a *Toxoplasma gondii* strain with a new isoenzyme type acquired in French Guyana. J Clin Microbiol 36, 324.

Davaro, R.E., Thirumalai, A., 2007. Life-threatening complications of HIV infection. J Intensive Care Med 22, 73-81.

de Amorim Garcia, C.A., Orefice, F., de Oliveira Lyra, C., Gomes, A.B., Franca, M., de Amorim Garcia Filho, C.A., 2004. Socioeconomic conditions as determining factors in the prevalence of systemic and ocular toxoplasmosis in Northeastern Brazil. Ophthalmic Epidemiol 11, 301-317.

de Moura, L., Bahia-Oliveira, L.M., Wada, M.Y., Jones, J.L., Tuboi, S.H., Carmo, E.H., Ramalho, W.M., Camargo, N.J., Trevisan, R., Graca, R.M., da Silva, A.J., Moura, I., Dubey, J.P., Garrett, D.O., 2006. Waterborne toxoplasmosis, Brazil, from field to gene. Emerg Infect Dis 12, 326-329.

de Roode, J.C., Yates, A:J., Alitzer, S., 2008. Virulence-transmission trade-offs and population divergence in virulence in a naturally occurring butterfly parasite. Proc Natl Acad Sci U S A 105, 7489-7494.

Delair, E., Monnet, D., Grabar, S., Dupouy-Camet, J., Yera, H.E., A, P.B., 2008. Respective Roles of Acquired and Congenital Infections in Presumed Ocular Toxoplasmosis. Am J Ophthalmol 146, 851-855.

Dellacasa-Lindberg, I., Hitziger, N., Barragan, A., 2007. Localized recrudescence of *Toxoplasma* infections in the central nervous system of immunocompromised mice assessed by in vivo bioluminescence imaging. Microbes Infect 9, 1291-1298.

Demar, M., Ajzenberg, D., Maubon, D., Djossou, F., Panchoe, D., Punwasi, W., Valery, N., Peneau, C., Daigre, J.L., Aznar, C., Cottrelle, B., Terzan, L., Darde, M.L., Carme, B., 2007. Fatal outbreak of human toxoplasmosis along the

Maroni River: epidemiological, clinical, and parasitological aspects. Clin Infect Dis 45, e88-95.

Desmonts, G., Couvreur, J., Alison, F., Baudelot, J., Gerbeaux, J., Lelong, M., 1965. [Epidemiological study on toxoplasmosis: the influence of cooking slaughter-animal meat on the incidence of human infection]. Rev Fr Etud Clin Biol 10, 952-958.

Di Cristina, M., Marocco, D., Galizi, R., Proietti, C., Spaccapelo, R., Crisanti, A., 2008. Temporal and spatial distribution of *Toxoplasma gondii* differentiation into Bradyzoites and tissue cyst formation in vivo. Infect Immun 76, 3491-3501.

Dimier, I.H., Bout, D.T., 1998. Interferon-gamma-activated primary enterocytes inhibit *Toxoplasma gondii* replication: a role for intracellular iron. Immunology 94, 488-495.

Dixon, S., McDonald, S., Roberts, J., 2004. The impact of HIV and AIDS on Africa's economic development. Trans Am Clin Climatol Assoc 115, 289-303.

Doganci, L., Tanyuksel, M., Araz, E.R., Besirbellioglu, B.A., Erdem, U., Ozoguz, C.A., Yucel, N., Ciftcioglu, A., 2006. A probable outbreak of toxoplasmosis among boarding school students in Turkey. Clin Microbiol Infect 12, 672-674.

Dubey, J.P., Miller, N.L., Frenkel, J.K., 1970. The *Toxoplasma gondii* oocyst from cat feces. J Exp Med 132, 636-662.

Dubey, J.P., Frenkel, J.K., 1972. Cyst-induced toxoplasmosis in cats. J Protozool 19, 155-177.

Dubey, J.P., Frenkel, J.K., 1976. Feline toxoplasmosis from acutely infected mice and the development of *Toxoplasma* cysts. J Protozool 23, 537-546.

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

Dubey, J.P., 1987. *Toxoplasma gondii* cysts in placentas of experimentally infected sheep. Am J Vet Res 48, 352-353.

Dubey, J.P., Beattie, C.P., 1988. Toxoplasmosis of Animals and Man. CRC Press, Boca Raton, Florida.

Dubey, J.P., Ott-Joslin, J., Torgerson, R.W., Topper, M.J., Sundberg, J.P., 1988. Toxoplasmosis in black-faced kangaroos (Macropus fuliginosus melanops). Vet Parasitol 30, 97-105.

Dubey, J.P., 1997a. Distribution of tissue cysts in organs of rats fed *Toxoplasma gondii* oocysts. J Parasitol 83, 755-757.

Dubey, J.P., 1997b. Tissue cyst tropism in *Toxoplasma gondii*: a comparison of tissue cyst formation in organs of cats, and rodents fed oocysts. Parasitology 115 (Pt 1), 15-20.

Dubey, J.P., 1997c. Bradyzoite-induced murine toxoplasmosis: stage conversion, pathogenesis, and tissue cyst formation in mice fed bradyzoites of different strains of *Toxoplasma gondii*. J Eukaryot Microbiol 44, 592-602.

Dubey, J.P., 1998a. Advances in the life cycle of *Toxoplasma gondii*. Int J Parasitol 28, 1019-1024.

Dubey, J.P., 1998b. *Toxoplasma gondii* oocyst survival under defined temperatures. J Parasitol 84, 862-865.

Dubey, J.P., 1998c. Refinement of pepsin digestion method for isolation of *Toxoplasma gondii* from infected tissues. Vet Parasitol 74, 75-77.

Dubey, J.P., Frenkel, J.K., 1998. Toxoplasmosis of rats: a review, with considerations of their value as an animal model and their possible role in epidemiology. Vet Parasitol 77, 1-32.

Dubey, J.P., Lindsay, D.S., Speer, C.A., 1998. Structures of *Toxoplasma gondii* tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. Clin Microbiol Rev 11, 267-299.

Dubey, J.P., Shen, S.K., Kwok, O.C., Frenkel, J.K., 1999. Infection and immunity with the RH strain of *Toxoplasma gondii* in rats and mice. J Parasitol 85, 657-662.

Dubey, J.P., Graham, D.H., da Silva, D.S., Lehmann, T., Bahia-Oliveira, L.M., 2003a. *Toxoplasma gondii* isolates of free-ranging chickens from Rio de Janeiro, Brazil: mouse mortality, genotype, and oocyst shedding by cats. J Parasitol 89, 851-853.

Dubey, J.P., Graham, D.H., Dahl, E., Hilali, M., El-Ghaysh, A., Sreekumar, C., Kwok, O.C., Shen, S.K., Lehmann, T., 2003b. Isolation and molecular characterization of *Toxoplasma gondii* from chickens and ducks from Egypt. Vet Parasitol 114, 89-95.

Dubey, J.P., 2004. Toxoplasmosis - a waterborne zoonosis. Vet Parasitol 126, 57-72.

Dubey, J.P., Gomez-Marin, J.E., Bedoya, A., Lora, F., Vianna, M.C.B., Hill, D., Kwok, O.C., Shen, S.K., Marcet, P.L., Lehmann, T., 2005a. Genetic and biologic characteristics of *Toxoplasma gondii* isolates in free-range chickens from Colombia, South America. Veterinary Parasitology 134, 67-72.

Dubey, J.P., Karhemere, S., Dahl, E., Sreekumar, C., Diabate, A., Dabire, K.R., Vianna, M.C., Kwok, O.C., Lehmann, T., 2005b. First biologic and genetic characterization of *Toxoplasma gondii* isolates from chickens from Africa (Democratic Republic of Congo, Mali, Burkina Faso, and Kenya). J Parasitol 91, 69-72.

Dubey, J.P., 2006. Comparative infectivity of oocysts and bradyzoites of *Toxoplasma gondii* for intermediate (mice) and definitive (cats) hosts. Vet Parasitol.

Dubey, J.P., Patitucci, A.N., Su, C., Sundar, N., Kwok, O.C., Shen, S.K., 2006a. Characterization of *Toxoplasma gondii* isolates in free-range chickens from Chile, South America. Vet Parasitol 140, 76-82.

Dubey, J.P., Su, C., Cortes, J.A., Sundar, N., Gomez-Marin, J.E., Polo, L.J., Zambrano, L., Mora, L.E., Lora, F., Jimenez, J., Kwok, O.C., Shen, S.K., Zhang, X., Nieto, A., Thulliez, P., 2006b. Prevalence of *Toxoplasma gondii* in cats from Colombia, South America and genetic characterization of T. *gondii* isolates. Vet Parasitol 140, 69-75.

Dubey, J.P., Sundar, N., Pineda, N., Kyvsgaard, N.C., Luna, L.A., Rimbaud, E., Oliveira, J.B., Kwok, O.C., Qi, Y., Su, C., 2006c. Biologic and genetic characteristics of *Toxoplasma gondii* isolates in free-range chickens from Nicaragua, Central America. Vet Parasitol 142, 47-53.

Dubey, J.P., Huong, L.T., Sundar, N., Su, C., 2007a. Genetic characterization of *Toxoplasma gondii* isolates in dogs from Vietnam suggests their South American origin. Vet Parasitol 146, 347-351.

Dubey, J.P., Rajapakse, R.P., Wijesundera, R.R., Sundar, N., Velmurugan, G.V., Kwok, O.C., Su, C., 2007b. Prevalence of *Toxoplasma gondii* in dogs from Sri Lanka and genetic characterization of the parasite isolates. Vet Parasitol 146, 341-346.

Dubey, J.P., Webb, D.M., Sundar, N., Velmurugan, G.V., Bandini, L.A., Kwok, O.C., Su, C., 2007c. Endemic avian toxoplasmosis on a farm in Illinois: clinical disease, diagnosis, biologic and genetic characteristics of *Toxoplasma gondii* isolates from chickens (Gallus domesticus), and a goose (Anser anser). Vet Parasitol 148, 207-212.

Dubey, J.P., Zhu, X.Q., Sundar, N., Zhang, H., Kwok, O.C., Su, C., 2007d. Genetic and biologic characterization of *Toxoplasma gondii* isolates of cats from China. Vet Parasitol 145, 352-356.

Dubey, J.P., Velmurugan, G.V., Chockalingam, A., Pena, H.F., de Oliveira, L.N., Leifer, C.A., Gennari, S.M., Bahia Oliveira, L.M., Su, C., 2008. Genetic diversity of *Toxoplasma gondii* isolates from chickens from Brazil. Vet Parasitol 157, 299-305.

Dubremetz, J.F., Garcia-Reguet, N., Conseil, V., Fourmaux, M.N., 1998. Apical organelles and host-cell invasion by Apicomplexa. Int J Parasitol 28, 1007-1013.

Dubremetz, J.F., 2007. Rhoptries are major players in *Toxoplasma gondii* invasion and host cell interaction. Cell Microbiol 9, 841-848.

Dumas, N., Cazaux, M., Ferly-Therizol, M., Seguela, J.P., Gueunier, M., Charlet, J.P., 1989. [Epidemiology of toxoplasmosis in Ivory Coast]. Bull Soc Pathol Exot Filiales 82, 513-519.

Dumas, N., Cazaux, M., Tshikom, K., Mwinba, K., Nsinba, Salaun, J.J., Brochier, B., Seguela, J.P., 1990a. [Epidemiological study of toxoplasmosis at Kinshasa and in lower Zaire]. Ann Soc Belg Med Trop 70, 289-296.

Dumas, P.N., Le Guenno, B., Digoutte, J.P., Seguela, J.P., 1990b. [Toxoplasmosis in the Republic of Senegal. Sero-epidemiological survey]. Bull Soc Pathol Exot 83, 283-285.

Duncanson, P., Terry, R.S., Smith, J.E., Hide, G., 2001. High levels of congenital transmission of *Toxoplasma gondii* in a commercial sheep flock. Int J Parasitol 31, 1699-1703.

Dunlop, O., Rootwelt, V., Sannes, M., Goplen, A.K., Abdelnoor, M., Skaug, K., Baklien, K., Skar, A., Melby, K., Myrvang, B., Bruun, J.N., 1996. Risk of toxoplasmic encephalitis in AIDS patients: indications for prophylaxis. Scand J Infect Dis 28, 71-73.

Dunn, D., Wallon, M., Peyron, F., Petersen, E., Peckham, C., Gilbert, R., 1999. Mother-to-child transmission of toxoplasmosis: risk estimates for clinical counselling. Lancet 353, 1829-1833.

Duong, T.H., Dufillot, D., Martz, M., Richard-Lenoble, D., Kombila, M., 1992. [Seroepidemiological study of toxoplasmosis in Libreville, Gabon]. Ann Soc Belg Med Trop 72, 289-293.

Dupon, M., Cazenave, J., Pellegrin, J.L., Ragnaud, J.M., Cheyrou, A., Fischer, I., Leng, B., Lacut, J.Y., 1995. Detection of *Toxoplasma gondii* by PCR and tissue culture in cerebrospinal fluid and blood of human immunodeficiency virus-seropositive patients. J Clin Microbiol 33, 2421-2426.

Dupouy-Camet, J., de Souza, S.L., Maslo, C., Paugam, A., Saimot, A.G., Benarous, R., Tourte-Schaefer, C., Derouin, F., 1993. Detection of *Toxoplasma gondii* in venous blood from AIDS patients by polymerase chain reaction. J Clin Microbiol 31, 1866-1869.

Dybas, J.M., Madrid-Aliste, C.J., Che, F.Y., Nieves, E., Rykunov, D., Angeletti, R.H., Weiss, L.M., Kim, K., Fiser, A., 2008. Computational analysis and experimental validation of gene predictions in *Toxoplasma gondii*. PLoS ONE 3, e3899.

Dzierszinski, F., Mortuaire, M., Cesbron-Delauw, M.F., Tomavo, S., 2000. Targeted disruption of the glycosylphosphatidylinositol-anchored surface antigen SAG3 gene in *Toxoplasma gondii* decreases host cell adhesion and drastically reduces virulence in mice. Mol Microbiol 37, 574-582.

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

Dzitko, K., Staczek, P., Gatkowska, J., Dlugonska, H., 2006. *Toxoplasma gondii*: Serological recognition of reinfection. Exp Parasitol 112, 134-137.

Egger, M., Boulle, A., 2008. Population effect of scaling up ART in resource-poor settings. Lancet 371, 1558-1559.

El-Sayed, N.M., Myler, P.J., Blandin, G., Berriman, M., Crabtree, J., Aggarwal, G., Caler, E., Renauld, H., Worthey, E.A., Hertz-Fowler, C., Ghedin, E., Peacock, C., Bartholomeu, D.C., Haas, B.J., Tran, A.N., Wortman, J.R., Alsmark, U.C., Angiuoli, S., Anupama, A., Badger, J., Bringaud, F., Cadag, E., Carlton, J.M., Cerqueira, G.C., Creasy, T., Delcher, A.L., Djikeng, A., Embley, T.M., Hauser, C., Ivens, A.C., Kummerfeld, S.K., Pereira-Leal, J.B., Nilsson, D., Peterson, J., Salzberg, S.L., Shallom, J., Silva, J.C., Sundaram, J., Westenberger, S., White, O., Melville, S.E., Donelson, J.E., Andersson, B., Stuart, K.D., Hall, N., 2005. Comparative genomics of trypanosomatid parasitic protozoa. Science 309, 404-409.

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

Elbez-Rubinstein, A., Ajzenberg, D., Darde, M.L., Cohen, R., Dumetre, A., Yera, H., Gondon, E., Janaud, J.C., Thulliez, P., 2009. Congenital toxoplasmosis and reinfection during pregnancy: case report, strain characterization, experimental model of reinfection, and review. J Infect Dis 199, 280-285.

Elias, D., Mengistu, G., Akuffo, H., Britton, S., 2006. Are intestinal helminths risk factors for developing active tuberculosis? Trop Med Int Health 11, 551-558.

Eza, D., Cerrillo, G., Moore, D.A., Castro, C., Ticona, E., Morales, D., Cabanillas, J., Barrantes, F., Alfaro, A., Benavides, A., Rafael, A., Valladares, G., Arevalo, F., Evans, C.A., Gilman, R.H., 2006. Postmortem findings and opportunistic infections in HIV-positive patients from a public hospital in Peru. Pathol Res Pract 202, 767-775.

Fazaeli, A., Carter, P.E., Darde, M.L., Pennington, T.H., 2000. Molecular typing of *Toxoplasma gondii* strains by GRA6 gene sequence analysis. Int J Parasitol 30, 637-642.

Feagin, J.E., 2000. Mitochondrial genome diversity in parasites. Int J Parasitol 30, 371-390.

Ferguson, D.J., 2002. *Toxoplasma gondii* and sex: essential or optional extra? Trends Parasitol 18, 355-359.

Ferguson, D.J., Henriquez, F.L., Kirisits, M.J., Muench, S.P., Prigge, S.T., Rice, D.W., Roberts, C.W., McLeod, R.L., 2005. Maternal inheritance and stage-specific variation of the apicoplast in *Toxoplasma gondii* during development in the intermediate and definitive host. Eukaryot Cell 4, 814-826.

Ferguson, D.J.P., Dubremetz, J.F., 2007. The Ultrastructure of *Toxoplasma gondii*. In: Toxoplasma gondii the model Apicomplexan, edited by L. Weiss and K. Kim. P 19-48.

Ferreira Ade, M., Vitor, R.W., Gazzinelli, R.T., Melo, M.N., 2006. Genetic analysis of natural recombinant Brazilian *Toxoplasma gondii* strains by multilocus PCR-RFLP. Infect Genet Evol 6, 22-31.

Ferreira da Silva, F., Barbosa, H.S., Gross, U., Luder, C.G., 2008. Stress-related and spontaneous stage differentiation of *Toxoplasma gondii*. Mol Biosyst 4, 824-834.

Fichera, M.E., Roos, D.S., 1997. A plastid organelle as a drug target in apicomplexan parasites. Nature 390, 407-409.

Flegr, J., 2007. Effects of *Toxoplasma* on human behavior. Schizophr Bull 33, 757-760.

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

Foulon, W., Villena, I., Stray-Pedersen, B., Decoster, A., Lappalainen, M., Pinon, J.M., Jenum, P.A., Hedman, K., Naessens, A., 1999. Treatment of toxoplasmosis during pregnancy: a multicenter study of impact on fetal transmission and children's sequelae at age 1 year. Am J Obstet Gynecol 180, 410-415.

Foulon, W., Naessens, A., Ho-Yen, D., 2000. Prevention of congenital toxoplasmosis. J Perinat Med 28, 337-345.

Frenkel, J.K., 1973. *Toxoplasma* in and around us. Biol Sci 23, 343-352.

Frenkel, J.K., 1974. Pathology and pathogenesis of congenital toxoplasmosis. Bull N Y Acad Med 50, 182-191.

Frenkel, J.K., Escajadillo, A., 1987. Cyst rupture as a pathogenic mechanism of toxoplasmic encephalitis. Am J Trop Med Hyg 36, 517-522.

Freyre, A., Falcon, J., Mendez, J., Gastell, T., Venzal, J.M., 2007. *Toxoplasma gondii*: cross-immunity against the enteric cycle. Exp Parasitol 115, 48-52.

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

Fulton, J.D., Turk, J.L., 1959. Direct agglutination test for *Toxoplasma gondii*. Lancet 2, 1068-1069.

Fung, H.B., Kirschenbaum, H.L., 1996. Treatment regimens for patients with toxoplasmic encephalitis. Clin Ther 18, 1037-1056.

Furco, A., Carmagnat, M., Chevret, S., Garin, Y.J., Pavie, J., De Castro, N., Charron, D., Derouin, F., Rabian, C., Molina, J.M., 2008. Restoration of *Toxoplasma gondii*-specific immune responses in patients with AIDS starting HAART. Aids 22, 2087-2096.

Fux, B., Nawas, J., Khan, A., Gill, D., Su, C., Sibley, L.D., 2007. *Toxoplasma gondii* strains lacking oral transmission are defective in stage differentiation. Infect Immun 75, 2580-2590.

Gajria, B., Bahl, A., Brestelli, J., Dommer, J., Fischer, S., Gao, X., Heiges, M., Iodice, J., Kissinger, J.C., Mackey, A.J., Pinney, D.F., Roos, D.S., Stoeckert, C.J., Jr., Wang, H., Brunk, B.P., 2008. ToxoDB: an integrated *Toxoplasma gondii* database resource. Nucleic Acids Res 36, D553-556.

Gal-Mor, O., Finlay, B.B., 2006. Pathogenicity islands: a molecular toolbox for bacterial virulence. Cell Microbiol 8, 1707-1719.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., Sharp, P.M., Hahn, B.H., 1999. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature 397, 436-441.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B., 2002. Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419, 498-511.

Garly, M.L., Petersen, E., Pedersen, C., Lundgren, J.D., Gerstoft, J., 1997. Toxoplasmosis in Danish AIDS patients. Scand J Infect Dis 29, 597-600.

Gascon, J., Torres-Rodriguez, J.M., Soldevila, M., Merlos, A.M., 1989. [Seroepidemiology of toxoplasmosis in 2 communities of Rwanda (Central Africa)]. Rev Inst Med Trop Sao Paulo 31, 399-402.

Gaskell, E.A., Smith, J.E., Pinney, J.W., Westhead, D.R., McConkey, G.A., 2009. A unique dual activity amino acid hydroxylase in *Toxoplasma gondii*. PLoS ONE 4, e4801.

Gazzinelli, R., Xu, Y., Hieny, S., Cheever, A., Sher, A., 1992. Simultaneous depletion of CD4+ and CD8+ T lymphocytes is required to reactivate chronic infection with *Toxoplasma gondii*. J Immunol 149, 175-180.

Gazzinelli, R.T., Eltoum, I., Wynn, T.A., Sher, A., 1993. Acute cerebral toxoplasmosis is induced by in vivo neutralization of TNF-alpha and correlates with the down-regulated expression of inducible nitric oxide synthase and other markers of macrophage activation. J Immunol 151, 3672-3681.

Genot, S., Franck, J., Forel, J.M., Rebaudet, S., Ajzenberg, D., de Paula, A.M., Darde, M.L., Stein, A., Ranque, S., 2007. Severe *Toxoplasma gondii* I/III Recombinant-Genotype Encephalitis in a Human Immunodeficiency Virus Patient. J Clin Microbiol 45, 3138-3140.

Gherardi, A., Sarciron, M.E., 2007. Molecules targeting the purine salvage pathway in Apicomplexan parasites. Trends Parasitol 23, 384-389.

Ghosn, J., Paris, L., Ajzenberg, D., Carcelain, G., Darde, M.L., Tubiana, R., Bossi, P., Bricaire, F., Katlama, C., 2003. Atypical toxoplasmic manifestation after discontinuation of maintenance therapy in a human immunodeficiency virus type 1-infected patient with immune recovery. Clin Infect Dis 37, e112-114.

Gille, E., Bjorkman, A., Rooth, I., Ljungstrom, I., Linder, E., 1992. Low seroprevalence of *Toxoplasma gondii* antibodies in a Tanzanian village. Trans R Soc Trop Med Hyg 86, 263-265.

Glasner, P.D., Silveira, C., Kruszon-Moran, D., Martins, M.C., Burnier Junior, M., Silveira, S., Camargo, M.E., Nussenblatt, R.B., Kaslow, R.A., Belfort Junior, R., 1992. An unusually high prevalence of ocular toxoplasmosis in southern Brazil. Am J Ophthalmol 114, 136-144.

Goebel, S., Gross, U., Luder, C.G., 2001. Inhibition of host cell apoptosis by *Toxoplasma gondii* is accompanied by reduced activation of the caspase cascade and alterations of poly(ADP-ribose) polymerase expression. J Cell Sci 114, 3495-3505.

Goodman, C.D., McFadden, G.I., 2007. Fatty acid biosynthesis as a drug target in apicomplexan parasites. Curr Drug Targets 8, 15-30.

Gray, F., Chretien, F., Vallat-Decouvelaere, A.V., Scaravilli, F., 2003. The changing pattern of HIV neuropathology in the HAART era. J Neuropathol Exp Neurol 62, 429-440.

Greenspan, J.S., Greenspan, D., Lennette, E.T., Abrams, D.I., Conant, M.A., Petersen, V., Freese, U.K., 1985. Replication of Epstein-Barr virus within the epithelial cells of oral "hairy" leukoplakia, an AIDS-associated lesion. N Engl J Med 313, 1564-1571.

Griffin, L., Williams, K.A., 1983. Serological and parasitological survey of blood donors in Kenya for toxoplasmosis. Trans R Soc Trop Med Hyg 77, 763-766.

Grigg, M.E., Bonnefoy, S., Hehl, A.B., Suzuki, Y., Boothroyd, J.C., 2001a. Success and virulence in *Toxoplasma* as the result of sexual recombination between two distinct ancestries. Science 294, 161-165.

Grigg, M.E., Ganatra, J., Boothroyd, J.C., Margolis, T.P., 2001b. Unusual abundance of atypical strains associated with human ocular toxoplasmosis. J Infect Dis 184, 633-639.

Grigg, M.E., Suzuki, Y., 2003. Sexual recombination and clonal evolution of virulence in *Toxoplasma*. Microbes Infect 5, 685-690.

Grigg, M.E., Sundar, N., 2009. Sexual recombination punctuated by outbreaks and clonal expansions predicts *Toxoplasma gondii* population genetics. Int J Parasitol, *In Press*.

Grimwood, J., Smith, J.E., 1992. *Toxoplasma gondii*: the role of a 30-kDa surface protein in host cell invasion. Exp Parasitol 74, 106-111.

Grimwood, J., Smith, J.E., 1996. *Toxoplasma gondii*: the role of parasite surface and secreted proteins in host cell invasion. Int J Parasitol 26, 169-173.

Gubbels, M.J., Lehmann, M., Muthalagi, M., Jerome, M.E., Brooks, C.F., Szatanek, T., Flynn, J., Parrot, B., Radke, J., Striepen, B., White, M., 2008. Forward genetic analysis of the Apicomplexan cell division cycle in *Toxoplasma gondii*. Plos Path 4, e36.

Guo, Z.G., Gross, U., Johnson, A.M., 1997. *Toxoplasma gondii* virulence markers identified by random amplified polymorphic DNA polymerase chain reaction. Parasitol Res 83, 458-463.

Hager, K.M., Carruthers, V.B., 2008. MARveling at parasite invasion. Trends Parasitol 24, 51-54.

Hall, N., 2007. Advanced sequencing technologies and their wider impact in microbiology. J Exp Biol 210, 1518-1525.

Hall, S.M., Pandit, A., Golwilkar, A., Williams, T.S., 1999. How do Jains get *Toxoplasma* infection? Lancet 354, 486-487.

Hari, K.R., Modi, M.R., Mochan, A.H., Modi, G., 2007. Reduced risk of *Toxoplasma* encephalitis in HIV-infected patients--a prospective study from Gauteng, South Africa. Int J STD AIDS 18, 555-558.

Havelaar, A.H., Kemmeren, J.M., Kortbeek, L.M., 2007. Disease burden of congenital toxoplasmosis. Clin Infect Dis 44, 1467-1474.

Havlir, D.V., Getahun, H., Sanne, I., Nunn, P., 2008. Opportunities and challenges for HIV care in overlapping HIV and TB epidemics. Jama 300, 423-430.

He, X.L., Grigg, M.E., Boothroyd, J.C., Garcia, K.C., 2002. Structure of the immunodominant surface antigen from the *Toxoplasma gondii* SRS superfamily. Nat Struct Biol 9, 606-611.

Heintzelman, M.B., 2006. Cellular and molecular mechanics of gliding locomotion in eukaryotes. Int Rev Cytol 251, 79-129.

Heitman, J., 2006. Sexual reproduction and the evolution of microbial pathogens. Curr Biol 16, R711-725.

Hemelaar, J., Gouws, E., Ghys, P.D., Osmanov, S., 2006. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. Aids 20, W13-23.

Henriquez, S.A., Brett, R., Alexander, J., Pratt, J., Roberts, C.W., 2009. Neuropsychiatric disease and *Toxoplasma gondii* infection. Neuroimmunomodulation 16, 122-133.

Hickey, D.A., 1982. Selfish DNA: a sexually-transmitted nuclear parasite. Genetics 101, 519-531.

Hill, D.E., Chirukandoth, S., Dubey, J.P., 2005. Biology and epidemiology of *Toxoplasma gondii* in man and animals. Anim Health Res Rev 6, 41-61.

Holmes, C.B., Losina, E., Walensky, R.P., Yazdanpanah, Y., Freedberg, K.A., 2003. Review of human immunodeficiency virus type 1-related opportunistic infections in sub-Saharan Africa. Clin Infect Dis 36, 652-662.

Holmes, E.C., 2001. On the origin and evolution of the human immunodeficiency virus (HIV). Biol Rev Camb Philos Soc 76, 239-254.

Homan, W.L., Vercammen, M., De Braekeleer, J., Verschueren, H., 2000. Identification of a 200- to 300-fold repetitive 529 bp DNA fragment in *Toxoplasma gondii*, and its use for diagnostic and quantitative PCR. Int J Parasitol 30, 69-75.

Honoré, S., Couvelard, A., Garin, Y.J., Bedel, C., Hénin, D., Dardé, M.L., Derouin, F., 2000. Genotyping of *Toxoplasma gondii* strains from immunocompromised patients. Pathol Biol (Paris) 48, 541-547.

Howe, D.K., Sibley, L.D., 1995. *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. J Infect Dis 172, 1561-1566.

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

Howe, D.K., Honore, S., Derouin, F., Sibley, L.D., 1997. Determination of genotypes of *Toxoplasma gondii* strains isolated from patients with toxoplasmosis. J Clin Microbiol 35, 1411-1414.

Hove, T., Lind, P., Mukaratirwa, S., 2005. Preliminary characterisation of *Toxoplasma gondii* isolates from Zimbabwe, with stage-specific monoclonal antibodies. Ann Trop Med Parasitol 99, 377-382.

Hu, K., Mann, T., Striepen, B., Beckers, C.J., Roos, D.S., Murray, J.M., 2002. Daughter cell assembly in the protozoan parasite *Toxoplasma gondii*. Mol Biol Cell 13, 593-606.

Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17, 754-755.

Hughes, J.M., Williams, R.H., Morley, E.K., Cook, D.A., Terry, R.S., Murphy, R.G., Smith, J.E., Hide, G., 2006. The prevalence of Neospora caninum and co-infection with *Toxoplasma gondii* by PCR analysis in naturally occurring mammal populations. Parasitology 132, 29-36.

Hutchison, W.M., 1965. Experimental transmission of *Toxoplasma gondii*. Nature 206, 961-962.

Hutchison, W.M., 1967. The nematode transmission of *Toxoplasma gondii*. Trans R Soc Trop Med Hyg 61, 80-89.

Huynh, M.H., Carruthers, V.B., 2006. *Toxoplasma* MIC2 Is a Major Determinant of Invasion and Virulence. PLoS Pathog 2, 753-762.

Hyde, J.E., 2007. Targeting purine and pyrimidine metabolism in human apicomplexan parasites. Curr Drug Targets 8, 31-47.

Innes, E.A., 1997. Toxoplasmosis: comparative species susceptibility and host immune response. Comp Immunol Microbiol Infect Dis 20, 131-138.

Ismael, A.B., Dimier-Poisson, I., Lebrun, M., Dubremetz, J.F., Bout, D., Mevelec, M.N., 2006. Mic1-3 Knockout of *Toxoplasma gondii* Is a Successful Vaccine against Chronic and Congenital Toxoplasmosis in Mice. J Infect Dis 194, 1176-1183.

Jan, G., Delorme, V., David, V., Revenu, C., Rebollo, A., Cayla, X., Tardieux, I., 2007. The toxofilin-actin-PP2C complex of *Toxoplasma*: identification of interacting domains. Biochem J 401, 711-719.

Jeannel, D., Niel, G., Costagliola, D., Danis, M., Traore, B.M., Gentilini, M., 1988. Epidemiology of toxoplasmosis among pregnant women in the Paris area. Int J Epidemiol 17, 595-602.

Jones, J., Lopez, A., Wilson, M., 2003. Congenital toxoplasmosis. Am Fam Physician 67, 2131-2138.

Joseph, P., Calderon, M.M., Gilman, R.H., Quispe, M.L., Cok, J., Ticona, E., Chavez, V., Jimenez, J.A., Chang, M.C., Lopez, M.J., Evans, C.A., 2002. Optimization and evaluation of a PCR assay for detecting toxoplasmic encephalitis in patients with AIDS. J Clin Microbiol 40, 4499-4503.

Jung, A.C., Paauw, D.S., 1998. Diagnosing HIV-related disease: using the CD4 count as a guide. J Gen Intern Med 13, 131-136.

Jung, C., Lee, C.Y., Grigg, M.E., 2004. The SRS superfamily of *Toxoplasma* surface proteins. Int J Parasitol 34, 285-296.

Kaufman, H.E., Remington, J.S., Jacobs, L., 1958. Toxoplasmosis: the nature of virulence. Am J Ophthalmol 46, 255-260; discussion 260-251.

Kaufman, H.E., Melton, M.L., Remington, J.S., Jacobs, L., 1959. Strain differences of *Toxoplasma* gondil. J Parasitol 45, 189-190.

Khan, A., Su, C., German, M., Storch, G.A., Clifford, D.B., Sibley, L.D., 2005a. Genotyping of *Toxoplasma gondii* strains from immunocompromised patients reveals high prevalence of type I strains. J Clin Microbiol 43, 5881-5887.

Khan, A., Taylor, S., Su, C., Mackey, A.J., Boyle, J., Cole, R., Glover, D., Tang, K., Paulsen, I.T., Berriman, M., Boothroyd, J.C., Pfefferkorn, E.R., Dubey, J.P., Ajioka, J.W., Roos, D.S., Wootton, J.C., Sibley, L.D., 2005b. Composite genome map and recombination parameters derived from three archetypal lineages of *Toxoplasma gondii*. Nucleic Acids Res 33, 2980-2992.

Khan, A., Böhme, U., Kelly, K.A., Adlem, E., Brooks, K., Simmonds, M., Mungall, K., Quail, M.A., Arrowsmith, C., Chillingworth, T., Churcher, C., Harris, D., Collins, M., Fosker, N., Fraser, A., Hance, Z., Jagels, K., Moule, S., Murphy, L., O'Neil, S., Rajandream, M.A., Saunders, D., Seeger, K., Whitehead, S., Mayr, T., Xuan, X., Watanabe, J., Suzuki, Y., Wakaguri, H., Sugano, S., Sugimoto, C., Paulsen, I., Mackey, A.J., Roos, D.S., Hall, N., Berriman, M., Barrell, B., Sibley, L.D., Ajioka, J.W., 2006a. Common inheritance of chromosome Ia associated with clonal expansion of *Toxoplasma gondii*. Genome Res 16, 1119-1125.

Khan, A., Jordan, C., Muccioli, C., Vallochi, A.L., Rizzo, L.V., Belfort, R., Jr., Vitor, R.W., Silveira, C., Sibley, L.D., 2006b. Genetic divergence of *Toxoplasma gondii* strains associated with ocular toxoplasmosis, Brazil. Emerg Infect Dis 12, 942-949.

Khan, A., Fux, B., Su, C., Dubey, J.P., Darde, M.L., Ajioka, J.W., Rosenthal, B.M., Sibley, L.D., 2007. Recent transcontinental sweep of *Toxoplasma gondii* driven by a single monomorphic chromosome. Proc Natl Acad Sci U S A 104, 14872-14877.

Kim, K., Weiss, L.M., 2004. *Toxoplasma gondii*: the model apicomplexan. Int J Parasitol 34, 423-432.

Kim, K., Weiss, L.M., 2008. *Toxoplasma*: the next 100years. Microbes Infect 10, 978-984.

Kim, S.K., Boothroyd, J.C., 2005. Stage-specific expression of surface antigens by *Toxoplasma gondii* as a mechanism to facilitate parasite persistence. J Immunol 174, 8038-8048.

Kim, S.K., Karasov, A., Boothroyd, J.C., 2007. Bradyzoite-specific surface antigen SRS9 plays a role in maintaining *Toxoplasma* persistence in the brain and in host control of parasite replication in the intestine. Infect Immun 75, 1626-1634.

Kiwanuka, N., Laeyendecker, O., Robb, M., Kigozi, G., Arroyo, M., McCutchan, F., Eller, L.A., Eller, M., Makumbi, F., Birx, D., Wabwire-Mangen, F., Serwadda, D., Sewankambo, N.K., Quinn, T.C., Wawer, M., Gray, R., 2008. Effect of human immunodeficiency virus Type 1 (HIV-1) subtype on disease progression in persons from Rakai, Uganda, with incident HIV-1 infection. J Infect Dis 197, 707-713.

Kohler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J., Palmer, J.D., Roos, D.S., 1997. A plastid of probable green algal origin in Apicomplexan parasites. Science 275, 1485-1489.

Kolaczinski, J.H., Kabatereine, N.B., Onapa, A.W., Ndyomugyenyi, R., Kakembo, A.S., Brooker, S., 2007. Neglected tropical diseases in Uganda: the prospect and challenge of integrated control. Trends Parasitol 23, 485-493.

Kong, J.T., Grigg, M.E., Uyetake, L., Parmley, S., Boothroyd, J.C., 2003. Serotyping of *Toxoplasma gondii* infections in humans using synthetic peptides. J Infect Dis 187, 1484-1495.

Koortbeek, L.M., de Melker, H.E., Veldhuijzen, I.K., Conyn van Spaendonck, M.A.E, 2004. Population-based *Toxoplasma* seroprevalence study in the Netherlands. Epidemiol Infect 132, 839-845.

Laing, R.B., Flegg, P.J., Brettle, R.P., Leen, C.L., Burns, S.M., 1996. Clinical features, outcome and survival from cerebral toxoplasmosis in Edinburgh AIDS patients. Int J STD AIDS 7, 258-264.

Lappalainen, M., Hedman, K., 2004. Serodiagnosis of toxoplasmosis. The impact of measurement of IgG avidity. Ann Ist Super Sanita 40, 81-88.

Le Blancq, S.M., Swinkels, B.W., Gibson, W.C., Borst, P., 1988. Evidence for gene conversion between the phosphoglycerate kinase genes of Trypanosoma brucei. J Mol Biol 200, 439-447.

Leander, B.S., 2008. Marine gregarines: evolutionary prelude to the apicomplexan radiation? Trends Parasitol 24, 60-67.

Lecomte, V., Chumpitazi, B.F., Pasquier, B., Ambroise-Thomas, P., Santoro, F., 1992. Brain-tissue cysts in rats infected with the RH strain of *Toxoplasma gondii*. Parasitol Res 78, 267-269.

Lehmann, T., Blackston, C.R., Parmley, S.F., Remington, J.S., Dubey, J.P., 2000. Strain typing of *Toxoplasma gondii*: comparison of antigen-coding and housekeeping genes. J Parasitol 86, 960-971.

Lehmann, T., Marcet, P.L., Graham, D.H., Dahl, E.R., Dubey, J.P., 2006. Globalization and the population structure of *Toxoplasma gondii*. Proc Natl Acad Sci U S A 103, 11423-11428.

Lekutis, C., Ferguson, D.J., Grigg, M.E., Camps, M., Boothroyd, J.C., 2001. Surface antigens of *Toxoplasma gondii*: variations on a theme. Int J Parasitol 31, 1285-1292.

Lelong, B., Rahelimino, B., Candolfi, E., Ravelojaona, B.J., Villard, O., Rasamindrakotroka, A.J., Kien, T., 1995. [Prevalence of toxoplasmosis in a population of pregnant women in Antananarivo (Madagascar)]. Bull Soc Pathol Exot 88, 46-49.

Levine, N.D., 1977. Taxonomy of *Toxoplasma*. J Protozool 24, 36-41.

Levine, N.D., 1988. Progress in taxonomy of the Apicomplexan protozoa. J Protozool 35, 518-520.

Lucas, S.B., Hounnou, A., Peacock, C., Beaumel, A., Djomand, G., N'Gbichi, J.M., Yeboue, K., Honde, M., Diomande, M., Giordano, C., 1993. The mortality and pathology of HIV infection in a west African city. Aids 7, 1569-1579.

Ludlam, G.B., Somers, K., 1966. Incidence of *Toxoplasma* antibodies in Ugandans with special reference to cardiomyopathy. Trans R Soc Trop Med Hyg 60, 621-625.

Luft, B.J., Remington, J.S., 1992. Toxoplasmic encephalitis in AIDS. Clin Infect Dis 15, 211-222.

Magno, R.C., Straker, L.C., de Souza, W., Attias, M., 2005. Interrelations between the parasitophorous vacuole of *Toxoplasma gondii* and host cell organelles. Microsc Microanal 11, 166-174.
Mair, C., Hawes, S.E., Agne, H.D., Sow, P.S., N'Doye, I., Manhart, L.E., Fu, P.L., Gottlieb, G.S., Kiviat, N.B., 2008. Factors associated with CD4 lymphocyte counts in HIV-negative Senegalese individuals. Clin Exp Immunol 151, 432-440.
Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He, W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., Rothberg, J.M., 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376-380.
Marshall, P.A., Hughes, J.M., Williams, R.H., Smith, J.E., Murphy, R.G., Hide, G., 2004. Detection of high levels of congenital transmission of *Toxoplasma gondii* in natural urban populations of Mus domesticus. Parasitology 128, 39-42.
Mather, M.W., Vaidya, A.B., 2008. Mitochondria in malaria and related parasites: ancient, diverse and streamlined. J Bioenerg Biomembr 40, 425-433.
Matrajt, M., Angel, S.O., Pszenny, V., Guarnera, E., Roos, D.S., Garberi, J.C., 1999. Arrays of repetitive DNA elements in the largest chromosomes of *Toxoplasma gondii*. Genome 42, 265-269.
Matsuzaki, M., Kikuchi, T., Kita, K., Kojima, S., Kuroiwa, T., 2001. Large amounts of apicoplast nucleoid DNA and its segregation in *Toxoplasma gondii*. Protoplasma 218, 180-191.
McFadden, G.I., Reith, M.E., Munholland, J., Lang-Unnasch, N., 1996. Plastid in human parasites. Nature 381, 482.
Meisheri, Y.V., Mehta, S., Patel, U., 1997. A prospective study of seroprevalence of Toxoplasmosis in general population, and in HIV/AIDS patients in Bombay, India. J Postgrad Med 43, 93-97.
Mercier, C., Howe, D.K., Mordue, D., Lingnau, M., Sibley, L.D., 1998. Targeted disruption of the GRA2 locus in *Toxoplasma gondii* decreases acute virulence in mice. Infect Immun 66, 4176-4182.
Messier, V., Levesque, B., Proulx, J.F., Rochette, L., Libman, M.D., Ward, B.J., Serhir, B., Couillard, M., Ogden, N.H., Dewailly, E., Hubert, B., Dery, S., Barthe, C., Murphy, D., Dixon, B., 2008. Seroprevalence of *Toxoplasma gondii* Among Nunavik Inuit (Canada). Zoonoses Public Health 56, 188-197.
Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R., Jessup, D.A., Boothroyd, J.C., Brownstein, D., Conrad, P.A., 2004. An unusual genotype of *Toxoplasma gondii* is common in California sea otters (Enhydra lutris nereis) and is a cause of mortality. Int J Parasitol 34, 275-284.

Millogo, A., Ki-Zerbo, G.A., Traore, W., Sawadogo, A.B., Ouedraogo, I., Peghini, M., 2000. [*Toxoplasma* serology in HIV infected patients and suspected cerebral toxoplasmosis at the Central Hospital of Bobo-Dioulasso (Burkina Faso)]. Bull Soc Pathol Exot 93, 17-19.

Miro, J.M., Lopez, J.C., Podzamczer, D., Pena, J.M., Alberdi, J.C., Martinez, E., Domingo, P., Cosin, J., Claramonte, X., Arribas, J.R., Santin, M., Ribera, E., 2006. Discontinuation of primary and secondary *Toxoplasma gondii* prophylaxis is safe in HIV-infected patients after immunological restoration with highly active antiretroviral therapy: results of an open, randomized, multicenter clinical trial. Clin Infect Dis 43, 79-89.

Mitchell, R.G., Green, C.A., 1960. The haemagglutination test for *Toxoplasma* antibodies. J Clin Pathol 13, 331-335.

Modi, G., Hari, K., Modi, M., Mochan, A., 2007. The frequency and profile of neurology in black South African HIV infected (clade C) patients - a hospital-based prospective audit. J Neurol Sci 254, 60-64.

Modi, M., Mochan, A., Modi, G., 2004. Management of HIV-associated focal brain lesions in developing countries. Qjm 97, 413-421.

Montoya, J.G., 2002. Laboratory diagnosis of *Toxoplasma gondii* infection and toxoplasmosis. J Infect Dis 185 Suppl 1, S73-82.

Montoya, J.G., Liesenfeld, O., 2004. Toxoplasmosis. Lancet 363, 1965-1976.

Moore, R.B., Obornik, M., Janouskovec, J., Chrudimsky, T., Vancova, M., Green, D.H., Wright, S.W., Davies, N.W., Bolch, C.J., Heimann, K., Slapeta, J., Hoegh-Guldberg, O., Logsdon, J.M., Carter, D.A., 2008. A photosynthetic alveolate closely related to apicomplexan parasites. Nature 451, 959-963.

Mordue, D.G., Desai, N., Dustin, M., Sibley, L.D., 1999. Invasion by *Toxoplasma gondii* establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring. J Exp Med 190, 1783-1792.

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

Morley, E.K., Williams, R.H., Hughes, J.M., Thomasson, D., Terry, R.S., Duncanson, P., Smith, J.E., Hide, G., 2008. Evidence that primary infection of Charollais sheep with *Toxoplasma gondii* may not prevent foetal infection and abortion in subsequent lambings. Parasitology 135, 169-173.

Morris, G.M., Gasser, R.B., 2006. Biotechnological advances in the diagnosis of avian coccidiosis and the analysis of genetic variation in Eimeria. Biotechnol Adv 24, 590-603.

Morrison, D.A., 2005. How old are the extant lineages of *Toxoplasma gondii*? Parassitologia 47, 205-214.

Mortensen, P.B., Norgaard-Pedersen, B., Waltoft, B.L., Sorensen, T.L., Hougaard, D., Yolken, R.H., 2007. Early infections of *Toxoplasma gondii* and the later development of schizophrenia. Schizophr Bull 33, 741-744.

Moskowitz, L.B., Kory, P., Chan, J.C., Haverkos, H.W., Conley, F.K., Hensley, G.T., 1983. Unusual causes of death in Haitians residing in Miami. High prevalence of opportunistic infections. Jama 250, 1187-1191.

Murphy, R.G., Williams, R.H., Hughes, J.M., Hide, G., Ford, N.J., Oldbury, D.J., 2008. The urban house mouse (Mus domesticus) as a reservoir of infection for the human parasite *Toxoplasma gondii*: an unrecognised public health issue? Int J Environ Health Res 18, 177-185.

Murray, H.W., Gellene, R.A., Libby, D.M., Rothermel, C.D., Rubin, B.Y., 1985. Activation of tissue macrophages from AIDS patients: in vitro response of AIDS alveolar macrophages to lymphokines and interferon-gamma. J Immunol 135, 2374-2377.

Nakasujja, N., Musisi, S., Robertson, K., Wong, M., Sacktor, N., Ronald, A., 2005. Human immunodeficiency virus neurological complications: An overview of the Ugandan experience. J Neurovirol 11 Suppl 3, 26-29.

Nascimento, L.V., Stollar, F., Tavares, L.B., Cavasini, C.E., Maia, I.L., Cordeiro, J.A., Ferreira, M.U., 2001. Risk factors for toxoplasmic encephalitis in HIV-infected patients: a case-control study in Brazil. Ann Trop Med Parasitol 95, 587-593.

National Water and Sewerage Corporation, 2006. Kampala. www.nwsc.co.ug.

Nicolle, C., Manceaux, L., 1908. On a Leishman body (or close organisms) gondi infection. Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences 147, 763-766.

Nicolle, C., Manceaux, L., 1909. The new protozoan of Gondi. Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences 148, 369-372.

Nissapatorn, V., Lee, C., Quek, K.F., Leong, C.L., Mahmud, R., Abdullah, K.A., 2004. Toxoplasmosis in HIV/AIDS patients: a current situation. Jpn J Infect Dis 57, 160-165.

Nyren, P., Pettersson, B., Uhlen, M., 1993. Solid phase DNA minisequencing by an enzymatic luminometric inorganic pyrophosphate detection assay. Anal Biochem 208, 171-175.

O'Brien, S.J., Johnson, W., Driscoll, C., Pontius, J., Pecon-Slattery, J., Menotti-Raymond, M., 2008. State of cat genomics. Trends Genet 24, 268-279.

Obornik, M., Green, B.R., 2005. Mosaic origin of the heme biosynthesis pathway in photosynthetic eukaryotes. Mol Biol Evol 22, 2343-2353.

Obornik, M., Janouskovec, J., Chrudimsky, T., Lukes, J., 2009. Evolution of the apicoplast and its hosts: from heterotrophy to autotrophy and back again. Int J Parasitol 39, 1-12.

Offiah, C.E., Turnbull, I.W., 2006. The imaging appearances of intracranial CNS infections in adult HIV and AIDS patients. Clin Radiol 61, 393-401.

Oksenhendler, E., Charreau, I., Tournerie, C., Azihary, M., Carbon, C., Aboulker, J.P., 1994. *Toxoplasma gondii* infection in advanced HIV infection. Aids 8, 483-487.

Omata, Y., Umeshita, Y., Murao, T., Kano, R., Kamiya, H., Kudo, A., Masukata, Y., Kobayashi, Y., Maeda, R., Saito, A., Murata, K., 2005. *Toxoplasma gondii* does not persist in goldfish (Carassius auratus). J Parasitol 91, 1496-1499.

Opitz, C., Soldati, D., 2002. 'The glideosome': a dynamic complex powering gliding motion and host cell invasion by *Toxoplasma gondii*. Mol Microbiol 45, 597-604.

Opperdoes, F.R., Michels, P.A., 2007. Horizontal gene transfer in trypanosomatids. Trends Parasitol 23, 470-476.

Ossorio, P.N., Sibley, L.D., Boothroyd, J.C., 1991. Mitochondrial-like DNA sequences flanked by direct and inverted repeats in the nuclear genome of *Toxoplasma gondii*. J Mol Biol 222, 525-536.

Palanisamy, M., Madhavan, B., Balasundaram, M.B., Andavar, R., Venkatapathy, N., 2006. Outbreak of ocular toxoplasmosis in Coimbatore, India. Indian J Ophthalmol 54, 129-131.

Palella, F.J., Jr., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., Aschman, D.J., Holmberg, S.D., 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 338, 853-860.

Palmer, G.H., Brayton, K.A., 2007. Gene conversion is a convergent strategy for pathogen antigenic variation. Trends Parasitol 23, 408-413.

Parmley, S.F., Goebel, F.D., Remington, J.S., 1992. Detection of *Toxoplasma gondii* in cerebrospinal fluid from AIDS patients by polymerase chain reaction. J Clin Microbiol 30, 3000-3002.

Pena, H.F., Gennari, S.M., Dubey, J.P., Su, C., 2008. Population structure and mouse-virulence of *Toxoplasma gondii* in Brazil. Int J Parasitol 38, 561-569.

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

Pfefferkorn, E.R., 1984. Interferon gamma blocks the growth of *Toxoplasma gondii* in human fibroblasts by inducing the host cells to degrade tryptophan. Proc Natl Acad Sci U S A 81, 908-912.

Pfefferkorn, L.C., Pfefferkorn, E.R., 1980. *Toxoplasma gondii*: genetic recombination between drug resistant mutants. Exp Parasitol 50, 305-316.

Pino, P., Foth, B.J., Kwok, L.Y., Sheiner, L., Schepers, R., Soldati, T., Soldati-Favre, D., 2007. Dual targeting of antioxidant and metabolic enzymes to the mitochondrion and the apicoplast of *Toxoplasma gondii*. PLoS Pathog 3, e115.

Pitchenik, A.E., Fischl, M.A., Dickinson, G.M., Becker, D.M., Fournier, A.M., O'Connell, M.T., Colton, R.M., Spira, T.J., 1983. Opportunistic infections and Kaposi's sarcoma among Haitians: evidence of a new acquired immunodeficiency state. Ann Intern Med 98, 277-284.

Plattner, F., Soldati-Favre, D., 2008. Hijacking of host cellular functions by the Apicomplexa. Annu Rev Microbiol 62, 471-487.

Porter, S.B., Sande, M.A., 1992. Toxoplasmosis of the central nervous system in the acquired immunodeficiency syndrome. N Engl J Med 327, 1643-1648.

Qiu, W., Wernimont, A., Tang, K., Taylor, S., Lunin, V., Schapira, M., Fentress, S., Hui, R., Sibley, L.D., 2009. Novel structural and regulatory features of rhoptry secretory kinases in *Toxoplasma gondii*. Embo J 8, 969-979

Quan, J.H., Kim, T.Y., Choi, I.U., Lee, Y.H., 2008. Genotyping of a Korean isolate of *Toxoplasma gondii* by multilocus PCR-RFLP and microsatellite analysis. Korean J Parasitol 46, 105-108.

Radke, J.R., Striepen, B., Guerini, M.N., Jerome, M.E., Roos, D.S., White, M.W., 2001. Defining the cell cycle for the tachyzoite stage of *Toxoplasma gondii*. Mol Biochem Parasitol 115, 165-175.

Raffi, F., Franck, J., Pelloux, H., Derouin, F., Reliquet, V., Ambroise-Thomas, P., Aboulker, J.P., Leport, C., Dumon, H., 1999. Specific anti-toxoplasmic IgG antibody immunoblot profiles in patients with AIDS-associated *Toxoplasma* encephalitis. Diagn Microbiol Infect Dis 34, 51-56.

Reiter-Owona, I., Seitz, H., Gross, U., Sahm, M., Rockstroh, J.K., Seitz, H.M., 2000. Is stage conversion the initiating event for reactivation of *Toxoplasma gondii* in brain tissue of AIDS patients? J Parasitol 86, 531-536.

Rich, S.M., Ayala, F.J., 2003. Progress in malaria research: the case for phylogenetics. Adv Parasitol 54, 255-280.

Richman, D.D., Little, S.J., Smith, D.M., Wrin, T., Petropoulos, C., Wong, J.K., 2004. HIV evolution and escape. Trans Am Clin Climatol Assoc 115, 289-303.

Rivera, M.C., 2007. Genomic analyses and the origin of the eukaryotes. Chem Biodivers 4, 2631-2638.

Robinson, S.A., Smith, J.E., Millner, P.A., 2004. *Toxoplasma gondii* major surface antigen (SAG1): in vitro analysis of host cell binding. Parasitology 128, 391-396.

Rodgers, C.A., Harris, J.R., 1996. Ocular toxoplasmosis in HIV infection. Int J STD AIDS 7, 307-309.

Rosen-Zvi, M., Altmann, A., Prosperi, M., Aharoni, E., Neuvirth, H., Sonnerborg, A., Schulter, E., Struck, D., Peres, Y., Incardona, F., Kaiser, R., Zazzi, M., Lengauer, T., 2008. Selecting anti-HIV therapies based on a variety of genomic and clinical factors. Bioinformatics 24, i399-406.

Saag, M.S., Hahn, B.H., Gibbons, J., Li, Y., Parks, E.S., Parks, W.P., Shaw, G.M., 1988. Extensive variation of human immunodeficiency virus type-1 in vivo. Nature 334, 440-444.

Sabin, A.B., 1941. Toxoplasmic encephalitis in children. JAMA 116, 801-807.

Sabin, A.B., Feldman, H.A., 1948. Dyes as Microchemical Indicators of a New Immunity Phenomenon Affecting a Protozoon Parasite (*Toxoplasma*). Science 108, 660-663.

Sachs, J., Malaney, P., 2002. The economic and social burden of malaria. Nature 415, 680-685.

Sacktor, N., Nakasujja, N., Robertson, K., Clifford, D.B., 2007. HIV-associated cognitive impairment in sub-Saharan Africa--the potential effect of clade diversity. Nat Clin Pract Neurol 3, 436-443.

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

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

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

Sahm, M., Gross, U., Reiter-Owona, I., Seitz, H.M., 1998. Infection and stage conversion during murine pulmonary toxoplasmosis: a study with three different strains of *Toxoplasma gondii*. J Parasitol 84, 723-729.

Sasono, P.M., Smith, J.E., 1998. *Toxoplasma gondii*: an ultrastructural study of host-cell invasion by the bradyzoite stage. Parasitol Res 84, 640-645.
Schuman, J.S., Friedman, A.H., 1983. Retinal manifestations of the acquired immune deficiency syndrome (AIDS): cytomegalovirus, candida albicans, cryptococcus, toxoplasmosis and Pneumocystis carinii. Trans Ophthalmol Soc U K 103 (Pt 2), 177-190.
Shafer, R.W., Schapiro, J.M., 2008. HIV-1 drug resistance mutations: an updated framework for the second decade of HAART. AIDS Rev 10, 67-84.
Sibley, L.D., Weidner, E., Krahenbuhl, J.L., 1985. Phagosome acidification blocked by intracellular *Toxoplasma gondii*. Nature 315, 416-419.
Sibley, L.D., Pfefferkorn, E.R., Boothroyd, J.C., 1991. Proposal for a uniform genetic nomenclature in *Toxoplasma gondii*. Parasitol Today 7, 327-328.
Sibley, L.D., Boothroyd, J.C., 1992a. Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. Nature 359, 82-85.
Sibley, L.D., Boothroyd, J.C., 1992b. Construction of a molecular karyotype for *Toxoplasma gondii*. Mol Biochem Parasitol 51, 291-300.
Sibley, L.D., LeBlanc, A.J., Pfefferkorn, E.R., Boothroyd, J.C., 1992. Generation of a restriction fragment length polymorphism linkage map for *Toxoplasma gondii*. Genetics 132, 1003-1015.
Sibley, L.D., Niesman, I.R., Asai, T., Takeuchi, T., 1994. *Toxoplasma gondii*: secretion of a potent nucleoside triphosphate hydrolase into the parasitophorous vacuole. Exp Parasitol 79, 301-311.
Sibley, L.D., Mordue, D.G., Su, C., Robben, P.M., Howe, D.K., 2002. Genetic approaches to studying virulence and pathogenesis in *Toxoplasma gondii*. Philos Trans R Soc Lond B Biol Sci 357, 81-88.
Sibley, L.D., 2003. Recent origins among ancient parasites. Veterinary Parasitology 115, 185-198.
Sibley, L.D., Ajioka, J.W., 2008. Population Structure of *Toxoplasma gondii*: Clonal Expansion Driven by Infrequent Recombination and Selective Sweeps. Annu Rev Microbiol 62, 329-351.
Sinai, A.P., 2007. The *Toxoplasma* kinase ROP18: an active member of a degenerate family. PLoS Pathog 3, e16.
Smith, A.B., Smirniotopoulos, J.G., Rushing, E.J., 2008. From the archives of the AFIP: central nervous system infections associated with human immunodeficiency virus infection: radiologic-pathologic correlation. Radiographics 28, 2033-2058.
Smith, J.E., 1995. A ubiquitous intracellular parasite: the cellular biology of *Toxoplasma gondii*. Int J Parasitol 25, 1301-1309.
Soete, M., Camus, D., Dubremetz, J.F., 1994. Experimental induction of bradyzoite-specific antigen expression and cyst formation by the RH strain of *Toxoplasma gondii* in vitro. Exp Parasitol 78, 361-370.
Sogin, M.L., Silberman, J.D., 1998. Evolution of the protists and protistan parasites from the perspective of molecular systematics. Int J Parasitol 28, 11-20.
Soldati-Favre, D., 2008. Molecular dissection of host cell invasion by the apicomplexans: the glideosome. Parasite 15, 197-205.
Sousa, S., Ajzenberg, D., Vilanova, M., Costa, J., Darde, M.L., 2008. Use of GRA6-derived synthetic polymorphic peptides in an immunoenzymatic assay to

serotype *Toxoplasma gondii* in human serum samples collected from three continents. Clin Vaccine Immunol 15, 1380-1386.
Speer, C.A., Tilley, M., Temple, M.E., Blixt, J.A., Dubey, J.P., White, M.W., 1995. Sporozoites of *Toxoplasma gondii* lack dense-granule protein GRA3 and form a unique parasitophorous vacuole. Mol Biochem Parasitol 75, 75-86.
Speer, C.A., Dubey, J.P., 2005. Ultrastructural differentiation of *Toxoplasma gondii* schizonts (types B to E) and gamonts in the intestines of cats fed bradyzoites. Int J Parasitol 35, 193-206.
Spencer, J.A., Joiner, K.S., Hilton, C.D., Dubey, J.P., Toivio-Kinnucan, M., Minc, J.K., Blagburn, B.L., 2004. Disseminated toxoplasmosis in a captive ring-tailed lemur (Lemur catta). J Parasitol 90, 904-906.
Splendore, A., 1908. Un nuovo parassita de'conigli incontrato nelle lesioni anatomiche d'une malattia che ricorda in molti punti il Kala-azar dell'uomo. Rev. Soc. Scient. Sao Paulo 3, 109-112.
Stoebe, B., Maier, U.G., 2002. One, two, three: nature's tool box for building plastids. Protoplasma 219, 123-130.
Su, C., Howe, D.K., Dubey, J.P., Ajioka, J.W., Sibley, L.D., 2002. Identification of quantitative trait loci controlling acute virulence in *Toxoplasma gondii*. Proc Natl Acad Sci U S A 99, 10753-10758.
Su, C., Evans, D., Cole, R.H., Kissinger, J.C., Ajioka, J.W., Sibley, L.D., 2003. Recent expansion of *Toxoplasma* through enhanced oral transmission. Science 299, 414-416.
Su, C., Zhang, X., Dubey, J.P., 2006. Genotyping of *Toxoplasma gondii* by multilocus PCR-RFLP markers: A high resolution and simple method for identification of parasites. Int J Parasitol 36, 841-848.
Su, X., Ferdig, M.T., Huang, Y., Huynh, C.Q., Liu, A., You, J., Wootton, J.C., Wellems, T.E., 1999. A genetic map and recombination parameters of the human malaria parasite Plasmodium falciparum. Science 286, 1351-1353.
Sulzer, A.J., Franco, E.L., Takafuji, E., Benenson, M., Walls, K.W., Greenup, R.L., 1986. An oocyst-transmitted outbreak of toxoplasmosis: patterns of immunoglobulin G and M over one year. Am J Trop Med Hyg 35, 290-296.
Sumyuen, M.H., Garin, Y.J., Derouin, F., 1995. Early kinetics of *Toxoplasma gondii* infection in mice infected orally with cysts of an avirulent strain. J Parasitol 81, 327-329.
Suzuki, Y., Conley, F.K., Remington, J.S., 1989. Importance of endogenous IFN-gamma for prevention of toxoplasmic encephalitis in mice. J Immunol 143, 2045-2050.
Suzuki, Y., 1999. Genes, cells and cytokines in resistance against development of toxoplasmic encephalitis. Immunobiology 201, 255-271.
Swaminathan, K., Varala, K., Hudson, M.E., 2007. Global repeat discovery and estimation of genomic copy number in a large, complex genome using a high-throughput 454 sequence survey. BMC Genomics 8, 132.
Sved, J.A., McRae, A.F., Visscher, P.M., 2008. Divergence between human populations estimated from linkage disequilibrium. Am J Hum Genet 83, 737-743.

Takashima, Y., Suzuki, K., Xuan, X., Nishikawa, Y., Unno, A., Kitoh, K., 2008. Detection of the initial site of *Toxoplasma gondii* reactivation in brain tissue. Int J Parasitol 38, 601-607.

Taylor, S., Barragan, A., Su, C., Fux, B., Fentress, S.J., Tang, K., Beatty, W.L., Hajj, H.E., Jerome, M., Behnke, M.S., White, M., Wootton, J.C., Sibley, L.D., 2006. A Secreted Serine-Threonine Kinase Determines Virulence in the Eukaryotic Pathogen *Toxoplasma gondii*. Science 314, 1776-1780.

Tenter, A.M., Heckeroth, A.R., Weiss, L.M., 2000. *Toxoplasma gondii*: from animals to humans. Int J Parasitol 30, 1217-1258.

Terry, R.S., Smith, J.E., Duncanson, P., Hide, G., 2001. MGE-PCR: a novel approach to the analysis of *Toxoplasma gondii* strain differentiation using mobile genetic elements. Int J Parasitol 3, 155-161

Teutsch, S.M., Juranek, D.D., Sulzer, A., Dubey, J.P., Sikes, R.K., 1979. Epidemic toxoplasmosis associated with infected cats. N Engl J Med 300, 695-699.

Tibayrenc, M., Ward, P., Moya, A., Ayala, F.J., 1986. Natural populations of Trypanosoma cruzi, the agent of Chagas disease, have a complex multiclonal structure. Proc Natl Acad Sci U S A 83, 115-119.

Tibayrenc, M., Kjellberg, F., Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structures of Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas, and Trypanosoma and their medical and taxonomical consequences. Proc Natl Acad Sci U S A 87, 2414-2418.

Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Breniere, S.F., Darde, M.L., Ayala, F.J., 1991. Are eukaryotic microorganisms clonal or sexual? A population genetics vantage. Proc Natl Acad Sci U S A 88, 5129-5133.

Tibayrenc, M., Ayala, F.J., 2002. The clonal theory of parasitic protozoa: 12 years on. Trends Parasitol 18, 405-410.

UNAIDS, 2008. Report on the global AIDS epidemic. http://www.unaids.org/en/KnowledgeCentre/HIVData/GlobalReport/2008/2008_Global_report.asp.

Uneke, C.J., Duhlinska, D.D., Njoku, M.O., Ngwu, B.A., 2005. Seroprevalence of acquired toxoplasmosis in HIV-infected and apparently healthy individuals in Jos, Nigeria. Parassitologia 47, 233-236.

UNEP, 2009. United Nations Environment Programme, Africa Environment Outlook, Climate Variability. http://www.unep.org/dewa/Africa/ publications/aeo-1/032.htm.

Waller, R.F., McFadden, G.I., 2005. The apicoplast: a review of the derived plastid of apicomplexan parasites. Curr Issues Mol Biol 7, 57-79.

Walls, K.W., Bullock, S.L., English, D.K., 1977. Use of the enzyme-linked immunosorbent assay (ELISA) and its microadaptation for the serodiagnosis of toxoplasmosis. J Clin Microbiol 5, 273-277.

Wamala, D., 1998. Brain infections in acquired immune deficiency syndrome due to human immune deficiency virus in Mulago hospital., Makerere University.

Van de Peer, Y., De Wachter, R., 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. Comput Appl Biosci 10, 569-570.

Wang, C., Mitsuya, Y., Gharizadeh, B., Ronaghi, M., Shafer, R.W., 2007. Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance. Genome Res 17, 1195-1201.

Wanyenya, I., Muyanja, C., Nasinyama, G.W., 2004. Kitchen practices used in handling broiler chickens and survival of Campylobacter spp. on cutting surfaces in Kampala, Uganda. J Food Prot 67, 1957-1960.

Vasan, A., Renjifo, B., Hertzmark, E., Chaplin, B., Msamanga, G., Essex, M., Fawzi, W., Hunter, D., 2006. Different rates of disease progression of HIV type 1 infection in Tanzania based on infecting subtype. Clin Infect Dis 42, 843-852.

Webster, J.P., 1994. Prevalence and transmission of *Toxoplasma gondii* in wild brown rats, Rattus norvegicus. Parasitology 108 (Pt 4), 407-411.

Webster, J.P., 2007. The effect of *Toxoplasma gondii* on animal behavior: playing cat and mouse. Schizophr Bull 33, 752-756.

Weiss, L.M., Kim, K., 2000. The development and biology of bradyzoites of *Toxoplasma gondii*. Front Biosci 5, D391-405.

Velmurugan, G.V., Dubey, J.P., Su, C., 2008. Genotyping studies of *Toxoplasma gondii* isolates from Africa revealed that the archetypal clonal lineages predominate as in North America and Europe. Vet Parasitol 155, 314-318.

WHO, 2004. http://www.who.int/healthinfo/global_burden_disease/GBD_report _2004update_AnnexA.pdf

Wickstead, B., Ersfeld, K., Gull, K., 2003. Repetitive elements in genomes of parasitic protozoa. Microbiol Mol Biol Rev 67, 360-375

Wiesner, J., Reichenberg, A., Heinrich, S., Schlitzer, M., Jomaa, H., 2008. The plastid-like organelle of apicomplexan parasites as drug target. Curr Pharm Des 14, 855-871.

Villena, I., Marle, M., Darde, M.L., Pinon, J.M., Aubert, D., 2004. *Toxoplasma* strain type and human disease: risk of bias during parasite isolation? Trends Parasitol 20, 160-162; discussion 162-163.

Williams, J.G., Kubelik, A.R., Livak, K.J., Rafalski, J.A., Tingey, S.V., 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18, 6531-6535.

Wolf, A., Cowen, D., Paige, B., 1939. Human Toxoplasmosis: Occurrence in Infants as an Encephalomyelitis Verification by Transmission to Animals. Science 89, 226-227.

Worobey, M., Gemmel, M., Teuwen, D.E., Haselkorn, T., Kunstman, K., Bunce, M., Muyembe, J.J., Kabongo, J.M., Kalengayi, R.M., Van Marck, E., Gilbert, M.T., Wolinsky, S.M., 2008. Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960. Nature 455, 661-664.

Vyas, A., Kim, S.K., Giacomini, N., Boothroyd, J.C., Sapolsky, R.M., 2007. Behavioral changes induced by *Toxoplasma* infection of rodents are highly specific to aversion of cat odors. Proc Natl Acad Sci U S A 104, 6442-6447.

Xia, D., Sanderson, S.J., Jones, A.R., Prieto, J.H., Yates, J.R., Bromley, E., Tomley, F.M., Lal, K., Sinden, R.E., Brunk, B.P., Roos, D.S., Wastling, J.M., 2008. The proteome of *Toxoplasma gondii*: integration with the genome provides novel insights into gene expression and annotation. Genome Biol 9, R116

Xu, J., 2004. The prevalence and evolution of sex in microorganisms. Genome 47, 775-780.

Yano, A., Mun, H.S., Chin, M., Norose, K., Hata, K., Kobayashi, M., Aosai, F., Iwakura, Y., 2002. Roles of IFN-gamma on stage conversion of an obligate intracellular protozoan parasite, *Toxoplasma gondii*. Int Rev Immunol 21, 405-421.

Yasodhara, P., Ramalakshmi, B.A., Lakshmi, V., Krishna, T.P., 2004. Socioeconomic status and prevalence of toxoplasmosis during pregnancy. Indian J Med Microbiol 22, 241-243.

Youssef, M.E., 1993. Profile of toxoplasmosis in two different localities in Dakahlia Governorate. J Egypt Soc Parasitol 23, 423-430.

Zakimi, S., Kyan, H., Oshiro, M., Sugimoto, C., Xuenan, X., Fujisaki, K., 2006. Genetic Characterization of GRA6 Genes from *Toxoplasma gondii* from Pigs in Okinawa, Japan. J Vet Med Sci 68, 1105-1107.

Zia-Ali, N., Fazaeli, A., Khoramizadeh, M., Ajzenberg, D., Darde, M., Keshavarz-Valian, H., 2007. Isolation and molecular characterization of *Toxoplasma gondii* strains from different hosts in Iran. Parasitol Res 101, 111-115.

Zuckerkandl, E., Pauling, L., 1965. Molecules as documents of evolutionary history. J Theor Biol 8, 357-366.

Zumla, A., Savva, D., Wheeler, R.B., Hira, S.K., Luo, N.P., Kaleebu, P., Sempala, S.K., Johnson, J.D., Holliman, R., 1991. *Toxoplasma* serology in Zambian and Ugandan patients infected with the human immunodeficiency virus. Trans R Soc Trop Med Hyg 85, 227-229.

Appendix 1: T. gondii genotypes in different regions and host species

Abbreviations for genotyping methods are: RFLP = restriction fragment length polymorphisms, MS = microsatellites, Z = zymodemes, SNP = single nucleotide polymorphisms, seq = sequencing.

Recombinant strains are strains which show alleles typical for different genotypes on different loci, and atypical are strains with non-archetypal alleles on at least one locus. Mixed means co-infection with two or more strains in the same host.

Host	Origin	N	Method(s)	No of markers	Type I	Type II	Type III	Recomb.	Atypical	Mixed	Reference
Cats	Brazil	38	RFLP	1	15		23				Dubey et al., 2004d
Chickens	Argentina	9	RFLP	1	1	1	7				Dubey et al., 2003d
Chickens	Argentina	17	RFLP	1	4	3	10				Dubey et al., 2005g
Chickens	Austria	67	RFLP	1		67					Dubey et al., 2005a
Chickens	Brazil	48	RFLP	1	34		13			1	Dubey et al., 2003a
Chickens	Chile	22	RFLP	≥5		17	4	1			Dubey et al., 2006a
Chickens	Colombia	24	RFLP	1	7		17				Dubey et al., 2005b
Chickens	Costa Rica	32	RFLP	≥5	5		1	26			Dubey et al., 2006b
Chickens	DRC, Mali, Burkina Faso, Kenya	17	RFLP	1	1	4	12				Dubey et al., 2005d
Chickens	Grenada	36	RFLP	1	5	1	29		1		Dubey et al., 2005i
Chickens	Guatemala	8	RFLP	1	3		5				Dubey et al., 2005f
Chickens	India	7	RFLP, MS	≥5		2	5				Sreekumar et al., 2003
Chickens	Israel	19	RFLP	1		17	2				Dubey et al., 2004f
Chickens	Mexico	6	RFLP	1	1		5				Dubey et al., 2004c
Chickens	Peru	10	RFLP	1	7		3				Dubey et al., 2004b
Chickens	Sri Lanka	12	RFLP	1		6	6				Dubey et al., 2005h
Chickens	USA	19	RFLP	1		5	14				Dubey et al., 2003c
Chickens	Venezuela	13	RFLP	1		3	10				Dubey et al., 2005e
Chickens, duck	Egypt	20	RFLP	1		3	17				Dubey et al., 2003b
Dogs	Brazil	9	RFLP	1	4		5				da Silva et al., 2005
Domestic animals	Brazil	16	RFLP	≥5				16			Ferreira Ade et al., 2006

Host	Origin	N	Method(s)	No of markers	Type I	Type II	Type III	Recomb.	Atypical	Mixed	Reference
Domestic animals	Colombia	19	RFLP	1	19						Gallego et al., 2006
Domestic animals	Europe, USA	5	SNP + MS	≥5		2	1		2		Ajzenberg et al., 2004
Domestic animals	Europe, USA	9	RFLP	3	2	4		1	2		Sibley and Boothroyd, 1992
Domestic animals	Europe, USA	10	Z, MS	1	1	5	4				Costa et al., 1997
Domestic animals	Europe, USA	23	MS, Z, RFLP	≥5	1	14	4		4		Ajzenberg et al., 2002a
Domestic animals	Uruguay, Barbados	3	SNP + MS	≥5			2		1		Ajzenberg et al., 2004
Domestic animals	USA, Canada, U.K	4	RFLP, seq	≥5	1	2	1				Lehmann et al., 2000
Humans	Africa	3	SNP + MS	≥5		1		1	1		Ajzenberg et al., 2004
Humans	Australia	1	RFLP	3					1		Sibley and Boothroyd, 1992
Humans	Brazil	1	RFLP	3	1						Sibley and Boothroyd, 1992
Humans	Brazil	4	RFLP	≥5				4			Ferreira Ade et al., 2006
Humans	Colombia	14	RFLP	1	12		1		1		Gallego et al., 2006
Humans	Europe, USA	14	SNP + MS	≥5	2	1	3	1	7		Ajzenberg et al., 2004
Humans	Europe, USA	14	RFLP	3	6	1		2	5		Sibley and Boothroyd, 1992
Humans	Europe, USA	24	Z, MS	1	6	17	1				Costa et al., 1997
Humans	Europe, USA	60	MS, Z, RFLP	≥5	15	37	4		4		Ajzenberg et al., 2002a
Humans	France	37	MS	1		37					Costa et al., 1997
Humans	France	68	RFLP	1	7	55	6				Howe et al., 1997
Humans	France	86	MS	≥5	7	73	2		4		Ajzenberg et al., 2002b
Humans	France	90	RFLP	1	14	69	7				Honore et al., 2000
Humans	French Guyana	16	MS	≥5					16		Carme et al., 2002

Host	Origin	N	Method(s)	No of markers	Type I	Type II	Type III	Recomb.	Atypical	Mixed	Reference
Humans	Japan	1	RFLP	3					1		Sibley and Boothroyd, 1992
Humans	Spain	25	RFLP	1	10	10	5				Fuentes et al., 2001
Humans	U.K.	32	RFLP	1	10	11	1			10	Aspinall et al., 2003
Humans	Uruguay, Guadeloupe, Reunion, French Guyana	13	SNP + MS	≥5	1		1		11		Ajzenberg et al., 2004
Humans	USA	12	RFLP	≥5	3	3	1	5			Grigg et al., 2001
Humans	USA, Canada, U.K	9	RFLP, seq	≥5	3	2	3		1		Lehmann et al., 2000
Pigs	Brazil	1	RFLP	3	1						Sibley and Boothroyd, 1992
Pigs	Brazil	7	RFLP	1	2		5				dos Santos et al., 2005
Pigs	Japan	1	RFLP	3		1					Sibley and Boothroyd, 1992
Pigs	Portugal	15	RFLP, MS	≥5		10	4		1		de Sousa et al., 2006
Pigs	USA	7	RFLP, MS	1	2	3	2				Dubey et al., 2005c
Pigs	USA	25	RFLP, MS	≥5		20	5				Lehmann et al., 2003
Pigs	USA	43	RFLP	1		36	7				Mondragon et al., 1998
Pork, beef, lamb	UK	27	RFLP	1	21					6	Aspinall et al., 2002
Sea otters	USA	13	RFLP	2		6			7		Cole et al., 2000
Sea otters	USA	35	RFLP, RAPD	≥5		14			21		Conrad et al., 2005
Sea otters	USA	35	RFLP, seq	≥5		14			21		Miller et al., 2004
Sheep	U.K.	15	RFLP	1		15					Owen and Trees, 1999
Wildlife	Europe, USA	3	SNP + MS	≥5				2	1		Ajzenberg et al., 2004
Wildlife	USA	10	RFLP	1		2	8				Dubey et al., 2004e
Wildlife	USA	43	RFLP	1	3	35	5				Dubey et al., 2004a
Wildlife	USA, Canada, U.K	3	RFLP, seq	≥5			1	2			Lehmann et al., 2000

References, Appendix 1

Ajzenberg, D., Banuls, A.L., Tibayrenc, M., Darde, M.L., 2002a. Microsatellite analysis of Toxoplasma gondii shows considerable polymorphism structured into two main clonal groups. Int J Parasitol 32, 27-38.

Ajzenberg, D., Cogne, N., Paris, L., Bessieres, M.H., Thulliez, P., Filisetti, D., Pelloux, H., Marty, P., Darde, M.L., 2002b. Genotype of 86 Toxoplasma gondii isolates associated with human congenital toxoplasmosis, and correlation with clinical findings. J Infect Dis 186, 684-689.

Ajzenberg, D., Banuls, A.L., Su, C., Dumetre, A., Demar, M., Carme, B., Darde, M.L., 2004. Genetic diversity, clonality and sexuality in Toxoplasma gondii. Int J Parasitol 34, 1185-1196.

Aspinall, T.V., Marlee, D., Hyde, J.E., Sims, P.F., 2002. Prevalence of Toxoplasma gondii in commercial meat products as monitored by polymerase chain reaction--food for thought? Int J Parasitol 32, 1193-1199.

Aspinall, T.V., Guy, E.C., Roberts, K.E., Joynson, D.H., Hyde, J.E., Sims, P.F., 2003. Molecular evidence for multiple Toxoplasma gondii infections in individual patients in England and Wales: public health implications. Int J Parasitol 33, 97-103.

Carme, B., Bissuel, F., Ajzenberg, D., Bouyne, R., Aznar, C., Demar, M., Bichat, S., Louvel, D., Bourbigot, A.M., Peneau, C., Neron, P., Darde, M.L., 2002. Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. J Clin Microbiol 40, 4037-4044.

Cole, R.A., Lindsay, D.S., Howe, D.K., Roderick, C.L., Dubey, J.P., Thomas, N.J., Baeten, L.A., 2000. Biological and molecular characterizations of Toxoplasma gondii strains obtained from southern sea otters (Enhydra lutris nereis). J Parasitol 86, 526-530.

Conrad, P.A., Miller, M.A., Kreuder, C., James, E.R., Mazet, J., Dabritz, H., Jessup, D.A., Gulland, F., Grigg, M.E., 2005. Transmission of Toxoplasma: Clues from the study of sea otters as sentinels of Toxoplasma gondii flow into the marine environment. Int J Parasitol 35, 1155-1168.

Costa, J.M., Darde, M.L., Assouline, B., Vidaud, M., Bretagne, S., 1997. Microsatellite in the beta-tubulin gene of Toxoplasma gondii as a new genetic marker for use in direct screening of amniotic fluids. J Clin Microbiol 35, 2542-2545.

da Silva, A.V., Pezerico, S.B., de Lima, V.Y., d'Arc Moretti, L., Pinheiro, J.P., Tanaka, E.M., Ribeiro, M.G., Langoni, H., 2005. Genotyping of Toxoplasma gondii strains isolated from dogs with neurological signs. Vet Parasitol 127, 23-27.

de Sousa, S., Ajzenberg, D., Canada, N., Freire, L., da Costa, J.M., Darde, M.L., Thulliez, P., Dubey, J.P., 2006. Biologic and molecular characterization of Toxoplasma gondii isolates from pigs from Portugal. Vet Parasitol 135, 133-136.

dos Santos, C.B.D., de Carvalho, A.C.F.B., Ragozo, A.M.A., Soares, R.M., Amaku, M., Yai, L.E.O., Dubey, J.P., Gennari, S.M., 2005. First isolation and molecular characterization of Toxoplasma gondii from finishing pigs from Sao Paulo State, Brazil. Veterinary Parasitology 131, 207-211.

Dubey, J.P., Graham, D.H., da Silva, D.S., Lehmann, T., Bahia-Oliveira, L.M., 2003a. Toxoplasma gondii isolates of free-ranging chickens from Rio de Janeiro, Brazil: mouse mortality, genotype, and oocyst shedding by cats. J Parasitol 89, 851-853.

Dubey, J.P., Graham, D.H., Dahl, E., Hilali, M., El-Ghaysh, A., Sreekumar, C., Kwok, O.C., Shen, S.K., Lehmann, T., 2003b. Isolation and molecular characterization of Toxoplasma gondii from chickens and ducks from Egypt. Vet Parasitol 114, 89-95.

Dubey, J.P., Graham, D.H., Dahl, E., Sreekumar, C., Lehmann, T., Davis, M.F., Morishita, T.Y., 2003c. Toxoplasma gondii isolates from free-ranging chickens from the United States. J Parasitol 89, 1060-1062.

Dubey, J.P., Venturini, M.C., Venturini, L., Piscopo, M., Graham, D.H., Dahl, E., Sreekumar, C., Vianna, M.C., Lehmann, T., 2003d. Isolation and genotyping of Toxoplasma gondii from free-ranging chickens from Argentina. J Parasitol 89, 1063-1064.

Dubey, J.P., Graham, D.H., De Young, R.W., Dahl, E., Eberhard, M.L., Nace, E.K., Won, K., Bishop, H., Punkosdy, G., Sreekumar, C., Vianna, M.C., Shen, S.K., Kwok, O.C., Sumners, J.A., Demarais, S., Humphreys, J.G., Lehmann, T., 2004a. Molecular and biologic characteristics of Toxoplasma gondii isolates from wildlife in the United States. J Parasitol 90, 67-71.

Dubey, J.P., Levy, M.Z., Sreekumar, C., Kwok, O.C., Shen, S.K., Dahl, E., Thulliez, P., Lehmann, T., 2004b. Tissue distribution and molecular characterization of chicken isolates of Toxoplasma gondii from Peru. J Parasitol 90, 1015-1018.

Dubey, J.P., Morales, E.S., Lehmann, T., 2004c. Isolation and genotyping of Toxoplasma gondii from free-ranging chickens from Mexico. J Parasitol 90, 411-413.

Dubey, J.P., Navarro, I.T., Sreekumar, C., Dahl, E., Freire, R.L., Kawabata, H.H., Vianna, M.C., Kwok, O.C., Shen, S.K., Thulliez, P., Lehmann, T., 2004d. Toxoplasma gondii infections in cats from Parana, Brazil: seroprevalence, tissue distribution, and biologic and genetic characterization of isolates. J Parasitol 90, 721-726.

Dubey, J.P., Parnell, P.G., Sreekumar, C., Vianna, M.C., De Young, R.W., Dahl, E., Lehmann, T., 2004e. Biologic and molecular characteristics of Toxoplasma gondii isolates from striped skunk (Mephitis mephitis), Canada goose (Branta canadensis), black-winged lory (Eos cyanogenia), and cats (Felis catus). J Parasitol 90, 1171-1174.

Dubey, J.P., Salant, H., Sreekumar, C., Dahl, E., Vianna, M.C., Shen, S.K., Kwok, O.C., Spira, D., Hamburger, J., Lehmann, T.V., 2004f. High prevalence of Toxoplasma gondii in a commercial flock of chickens in Israel, and public health implications of free-range farming. Vet Parasitol 121, 317-322.

Dubey, J.P., Edelhofer, R., Marcet, P., Vianna, M.C.B., Kwok, O.C.H., Lehmann, T., 2005a. Genetic and biologic characteristics of Toxoplasma gondii infections in free-range chickens from Austria. Veterinary Parasitology 133, 299-306.

Dubey, J.P., Gomez-Marin, J.E., Bedoya, A., Lora, F., Vianna, M.C.B., Hill, D., Kwok, O.C., Shen, S.K., Marcet, P.L., Lehmann, T., 2005b. Genetic and biologic characteristics of Toxoplasma gondii isolates in free-range chickens from Colombia, South America. Veterinary Parasitology 134, 67-72.

Dubey, J.P., Hill, D.E., Jones, J.L., Hightower, A.W., Kirkland, E., Roberts, J.M., Marcet, P.L., Lehmann, T., Vianna, M.C.B., Miska, K., Sreekumar, C., Kwok, O.C.H., Shen, S.K., Gamble, H.R., 2005c. Prevalence of viable Toxoplasma gondii in beef, chicken, and pork from retail meat stores in the United States: Risk assessment to consumers. Journal of Parasitology 91, 1082-1093.

Dubey, J.P., Karhemere, S., Dahl, E., Sreekumar, C., Diabate, A., Dabire, K.R., Vianna, M.C., Kwok, O.C., Lehmann, T., 2005d. First biologic and genetic characterization of Toxoplasma gondii isolates from chickens from Africa (Democratic Republic of Congo, Mali, Burkina Faso, and Kenya). J Parasitol 91, 69-72.

Dubey, J.P., Lenhart, A., Castillo, C.E., Alvarez, L., Marcet, P., Sreekumar, C., Lehmann, T., 2005e. Toxoplasma gondii infections in chickens from Venezuela: Isolation, tissue distribution, and molecular characterization. Journal of Parasitology 91, 1332-1334.

Dubey, J.P., Lopez, B., Alvarez, M., Mendoza, C., Lehmann, T., 2005f. Isolation, tissue distribution, and molecular characterization of Toxoplasma gondii from free-range chickens from Guatemala. Journal of Parasitology 91, 955-957.

Dubey, J.P., Marcet, P.L., Lehmann, T., 2005g. Characterization of Toxoplasma gondii isolates in free-range chickens from Argentina. Journal of Parasitology 91, 1335-1339.

Dubey, J.P., Rajapakse, R.P.V.J., Ekanayake, D.K., Sreekumar, C., Lehmann, T., 2005h. Isolation and molecular characterization of Toxoplasma gondii from chickens from Sri Lanka. Journal of Parasitology 91, 1480-1482.

Dubey, J.P., Patitucci, A.N., Su, C., Sundar, N., Kwok, O.C., Shen, S.K., 2006a. Characterization of Toxoplasma gondii isolates in free-range chickens from Chile, South America. Vet Parasitol 140, 76-82.

Dubey, J.P., Su, C., Oliveira, J., Morales, J.A., Bolanos, R.V., Sundar, N., Kwok, O.C., Shen, S.K., 2006b. Biologic and genetic characteristics of Toxoplasma gondii isolates in free-range chickens from Costa Rica, Central America. Vet Parasitol.

Dubey, J.R., Bhaiyat, M.I., de Allie, C., Macpherson, C.N., Sharma, R.N., Sreekumar, C., Vianna, M.C., Shen, S.K., Kwok, O.C., Miska, K.B., Hill, D.E., Lehmann, T., 2005i. Isolation, tissue distribution, and molecular characterization of Toxoplasma gondii from chickens in Grenada, West Indies. J Parasitol 91, 557-560.

Ferreira Ade, M., Vitor, R.W., Gazzinelli, R.T., Melo, M.N., 2006. Genetic analysis of natural recombinant Brazilian Toxoplasma gondii strains by multilocus PCR-RFLP. Infect Genet Evol 6, 22-31.

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

Gallego, C., Saavedra-Matiz, C., Gomez-Marin, J.E., 2006. Direct genotyping of animal and human isolates of Toxoplasma gondii from Colombia (South America). Acta Trop 97, 161-167.

Grigg, M.E., Ganatra, J., Boothroyd, J.C., Margolis, T.P., 2001. Unusual abundance of atypical strains associated with human ocular toxoplasmosis. J Infect Dis 184, 633-639.

Honore, S., Couvelard, A., Garin, Y.J., Bedel, C., Henin, D., Darde, M.L., Derouin, F., 2000. [Genotyping of Toxoplasma gondii strains from immunocompromised patients]. Pathol Biol (Paris) 48, 541-547.

Howe, D.K., Honore, S., Derouin, F., Sibley, L.D., 1997. Determination of genotypes of Toxoplasma gondii strains isolated from patients with toxoplasmosis. J Clin Microbiol 35, 1411-1414.

Lehmann, T., Blackston, C.R., Parmley, S.F., Remington, J.S., Dubey, J.P., 2000. Strain typing of Toxoplasma gondii: comparison of antigen-coding and housekeeping genes. J Parasitol 86, 960-971.

Lehmann, T., Graham, D.H., Dahl, E., Sreekumar, C., Launer, F., Corn, J.L., Gamble, H.R., Dubey, J.P., 2003. Transmission dynamics of Toxoplasma gondii on a pig farm. Infect Genet Evol 3, 135-141.

Miller, M.A., Grigg, M.E., Kreuder, C., James, E.R., Melli, A.C., Crosbie, P.R., Jessup, D.A., Boothroyd, J.C., Brownstein, D., Conrad, P.A., 2004. An unusual genotype of Toxoplasma gondii is common in California sea otters (Enhydra lutris nereis) and is a cause of mortality. Int J Parasitol 34, 275-284.

Mondragon, R., Howe, D.K., Dubey, J.P., Sibley, L.D., 1998. Genotypic analysis of Toxoplasma gondii isolates from pigs. J Parasitol 84, 639-641.

Owen, M.R., Trees, A.J., 1999. Genotyping of Toxoplasma gondii associated with abortion in sheep. J Parasitol 85, 382-384.

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

Sreekumar, C., Graham, D.H., Dahl, E., Lehmann, T., Raman, M., Bhalerao, D.P., Vianna, M.C., Dubey, J.P., 2003. Genotyping of Toxoplasma gondii isolates from chickens from India. Vet Parasitol 118, 187-194.

Appendix 2: T. gondii-seroprevalence in adult Africans

Country	N	Seroprev. (%)	Reference
Benin	211	54	Rodier et al., 1995
Burkina Faso	1828	25	Millogo et al., 2000
Burkina Faso	336	25	Simpore et al., 2006
Burkina-Faso	148	50	Ledru et al., 1995
Burundi	622	44	Excler et al., 1988
Cameroon	421	48.5	Marty et al., 1985
Cameroon	192	77	Ndumbe et al., 1992
Central African Republic	814	40 (overall)	
		25 (pre-desert)	Dumas et al., 1990b
Central African Republic	1953	51	Morvan et al., 1999
Ceuta (Morocco)	100	51	Perez-Rendon Gonzalez and Lopez Caminero, 1992
Comoros	871	84	Julvez et al., 1994
Congo-Brazzaville	2500	40	Dumas et al., 1990a
Congo-Brazzaville	310	42	Candolfi et al., 1993
DRC		50-65%	Dumas et al., 1990c
Egypt	150	43	el-Nawawy et al., 1996
Egypt	700	15	Hamadto et al., 1997
Egypt	152	58	Hussein et al., 2001
Ethiopia	144	34	Flatau et al., 1993
Ethiopia	1016	75	Guebre-Xabier et al., 1993
Ethiopia	170	80	Woldemichael et al., 1998
Gabon	268	61	Billiault et al., 1987
Gabon	1178	68	Duong et al., 1992
Gabon	767	71	Nabias et al., 1998
Ghana	364	57	Anteson et al., 1978
Ivory Coast	2000	37 (pre-desert)	
		56-70 (humid)	Dumas et al., 1989
Ivory Coast	1025	60	Adou-Bryn et al., 2004
Kenya	322	54	Griffin and Williams, 1983
Kenya	180	54	Brindle et al., 1991
Liberia	133	70	Negro Ponzi et al., 1976
Liberia	390	58	Omland et al., 1977
Libya	2300	50.5	Khadre and el Nageh, 1987
Libya	369	47	Kassem and Morsy, 1991
Madagascar	599	84	Lelong et al., 1995
Madagascar	2354	52	Dromigny et al., 1996
Mali	100	34	Maiga et al., 1984
Niger	400	18.2	Develoux et al., 1988
Niger	218	15	Develoux et al., 1989
Niger	371	18	Julvez et al., 1996
Nigeria	352	78	Onadeko et al., 1992
Nigeria	162	68	Ekweozor et al., 1994
Nigeria	363	32	Uneke et al., 2005
Nigeria	1650	59	Arene, 1986

Country	N	Seroprev. (%)	Reference
Rwanda		50	Gascon et al., 1989
Senegal	415	25 (dry)	
		58 (humid)	
		33 (Dakar)	Dumas et al., 1990d
Senegal	720	40	Diallo et al., 1996
Senegal	353	40	Faye et al., 1998
Somalia	356	53	Zardi et al., 1980
Somalia		40-56%	Ahmed et al., 1988
South Africa	600	10	Brink et al., 1975
South Africa, Namibia, Botswana	3379	20	Jacobs and Mason, 1978
Sudan	386	42	Abdel-Hameed, 1991
Sudan	487	34	Elnahas et al., 2003
Tanzania	357	48	Gill and Mtimavalye, 1982
Tanzania	450	4	Gille et al., 1992
	32	47	Doehring et al., 1995
Tanzania	849	35	Agbo et al., 1991
Togo	819	59	Deniau et al., 1991
Togo	618	62-88%	Bouratbine et al., 2001
Tunisia	1421	58	Zumla et al., 1991
Uganda	279	33	This study
Uganda	677	56	Zumla et al., 1991
Zambia	376	7	Dumas et al., 1990d

References, Appendix 2

Abdel-Hameed, A.A., 1991. Sero-epidemiology of toxoplasmosis in Gezira, Sudan. J Trop Med Hyg 94, 329-332.

Adou-Bryn, K.D., Ouhon, J., Nemer, J., Yapo, C.G., Assoumou, A., 2004. [Serological survey of acquired toxoplasmosis in women of child-bearing age in Yopougon (Abidjan, Cote d'Ivoire)]. Bull Soc Pathol Exot 97, 345-348.

Agbo, K., David, M., Amavi-Tete, T., Deniau, M., 1991. [The diagnosis of toxoplasmosis in the Lome University Hospital Center]. Bull Soc Pathol Exot 84, 659-663.

Ahmed, H.J., Mohammed, H.H., Yusuf, M.W., Ahmed, S.F., Huldt, G., 1988. Human toxoplasmosis in Somalia. Prevalence of Toxoplasma antibodies in a village in the lower Scebelli region and in Mogadishu. Trans R Soc Trop Med Hyg 82, 330-332.

Anteson, R.K., Sekimoto, S., Furukawa, S., 1978. Studies on toxoplasmosis in Ghana. I. Prevalence of toxoplasma antibodies as measured by the haemagglutination (Eiken) test Ghana Med J 17, 147-149.

Arene, F.O., 1986. The prevalence of toxoplasmosis among inhabitants of the Niger Delta. Folia Parasitol (Praha) 33, 311-314.

Billiault, X., Collet, M., Dupont, A., Lefevre, S., 1987. [Toxoplasmosis in pregnant women in Haut Ogooue Province (Gabon)]. Bull Soc Pathol Exot Filiales 80, 74-83.

Bouratbine, A., Siala, E., Chahed, M.K., Aoun, K., Ben Ismail, R., 2001. [Sero-epidemiologic profile of toxoplasmosis in northern Tunisia]. Parasite 8, 61-66.

Brindle, R., Holliman, R., Gilks, C., Waiyaki, P., 1991. Toxoplasma antibodies in HIV-positive patients from Nairobi. Trans R Soc Trop Med Hyg 85, 750-751.

Brink, J.D., De Wet, J.S., Van Rensburg, A.J., 1975. A serological survey of toxoplasmosis in the Bloemfonetin area. S Afr Med J 49, 1441-1443.

Candolfi, E., Berg, M., Kien, T., 1993. [Prevalence of toxoplasmosis in Pointe-Noire in Congo. Study of the sampling of 310 subjects]. Bull Soc Pathol Exot 86, 358-362.

Deniau, M., Tourte-Schaefer, C., Agbo, K., Dupouy-Camet, J., Heyer, C., Lapierre, J., 1991. [Evaluation of the risk of congenital toxoplasmosis in Togo]. Bull Soc Pathol Exot 84, 664-672.

Develoux, M., Candolfi, E., Hanga-Doumbo, S., Kien, T., 1988. [Toxoplasmosis in Niger. A serological analysis of 400 subjects]. Bull Soc Pathol Exot Filiales 81, 253-259.

Develoux, M., Chandenier, J., Tinni, A., 1989. [Toxoplasmosis in pregnant women in Niamey (Niger)]. Bull Soc Pathol Exot Filiales 82, 406-409.

Diallo, S., Ndir, O., Dieng, Y., Leye, A., Dieng, T., Bah, I.B., Diop, B.M., Gaye, O., Faye, O., 1996. [Seroprevalence of toxoplasmosis in Dakar (Senegal) in 1993: study of women in their reproductive years]. Sante 6, 102-106.

Doehring, E., Reiter-Owona, I., Bauer, O., Kaisi, M., Hlobil, H., Quade, G., Hamudu, N.A., Seitz, H.M., 1995. Toxoplasma gondii antibodies in pregnant women and their newborns in Dar es Salaam, Tanzania. Am J Trop Med Hyg 52, 546-548.

Dromigny, J.A., Pecarrere, J.L., Leroy, F., Ollivier, G., Boisier, P., 1996. [Prevalence of toxoplasmosis in Tananarive. Study conducted at the Pasteur Institute of Madagascar (PIM) on a sample of 2354 subjects]. Bull Soc Pathol Exot 89, 212-216.

Dumas, N., Cazaux, M., Ferly-Therizol, M., Seguela, J.P., Gueunier, M., Charlet, J.P., 1989. [Epidemiology of toxoplasmosis in Ivory Coast]. Bull Soc Pathol Exot Filiales 82, 513-519.

Dumas, N., Cazaux, M., Carme, B., Seguela, J.P., Charlet, J.P., 1990a. [Toxoplasmosis in the Congo Republic. Seroepidemiological study]. Bull Soc Pathol Exot 83, 349-359.

Dumas, N., Cazaux, M., Meunier, D.M., Seguela, J.P., Charlet, J.P., 1990b. [Toxoplasmosis in the Central African Republic. Complementary study in a rural area]. Bull Soc Pathol Exot 83, 342-348.

Dumas, N., Cazaux, M., Tshikom, K., Mwinba, K., Nsinba, Salaun, J.J., Brochier, B., Seguela, J.P., 1990c. [Epidemiological study of toxoplasmosis at Kinshasa and in lower Zaire]. Ann Soc Belg Med Trop 70, 289-296.

Dumas, P.N., Le Guenno, B., Digoutte, J.P., Seguela, J.P., 1990d. [Toxoplasmosis in the Republic of Senegal. Sero-epidemiological survey]. Bull Soc Pathol Exot 83, 283-285.

Duong, T.H., Dufillot, D., Martz, M., Richard-Lenoble, D., Kombila, M., 1992. [Seroepidemiological study of toxoplasmosis in Libreville, Gabon]. Ann Soc Belg Med Trop 72, 289-293.

Ekweozor, C.C., Okpala, L.E., Bamgboye, A.E., Jegede, O., 1994. Toxoplasmosis: antitoxoplasma IgG antibody levels in patients with lymphoid neoplasia in Ibadan, Nigeria. Afr J Med Med Sci 23, 347-353.

el-Nawawy, A., Soliman, A.T., el Azzouni, O., Amer el, S., Karim, M.A., Demian, S., el Sayed, M., 1996. Maternal and neonatal prevalence of toxoplasma and cytomegalovirus (CMV) antibodies and hepatitis-B antigens in an Egyptian rural area. J Trop Pediatr 42, 154-157.

Elnahas, A., Gerais, A.S., Elbashir, M.I., Eldien, E.S., Adam, I., 2003. Toxoplasmosis in pregnant Sudanese women. Saudi Med J 24, 868-870.

Excler, J.L., Pretat, E., Pozzetto, B., Charpin, B., Garin, J.P., 1988. Sero-epidemiological survey for toxoplasmosis in Burundi. Trop Med Parasitol 39, 139-141.

Faye, O., Leye, A., Dieng, Y., Richard-Lenoble, D., Diallo, S., 1998. [Toxoplasmosis in Dakar. Seroepidemiologic sampling of 353 women of reproductive age]. Bull Soc Pathol Exot 91, 249-250.

Flatau, E., Nishri, Z., Mates, A., Qupty, G., Reichman, N., Raz, R., 1993. Seroprevalence of antibodies against Toxoplasma gondii among recently immigrating Ethiopian Jews. Isr J Med Sci 29, 395-397.

Gascon, J., Torres-Rodriguez, J.M., Soldevila, M., Merlos, A.M., 1989. [Seroepidemiology of toxoplasmosis in 2 communities of Rwanda (Central Africa)]. Rev Inst Med Trop Sao Paulo 31, 399-402.

Gill, H.S., Mtimavalye, L.A., 1982. Prevalence of Toxoplasma antibodies in pregnant African women in Tanzania. Afr J Med Med Sci 11, 167-170.

Gille, E., Bjorkman, A., Rooth, I., Ljungstrom, I., Linder, E., 1992. Low seroprevalence of Toxoplasma gondii antibodies in a Tanzanian village. Trans R Soc Trop Med Hyg 86, 263-265.

Griffin, L., Williams, K.A., 1983. Serological and parasitological survey of blood donors in Kenya for toxoplasmosis. Trans R Soc Trop Med Hyg 77, 763-766.

Guebre-Xabier, M., Nurilign, A., Gebre-Hiwot, A., Hailu, A., Sissay, Y., Getachew, E., Frommel, D., 1993. Sero-epidemiological survey of Toxoplasma gondii infection in Ethiopia. Ethiop Med J 31, 201-208.

Hamadto, H.A., Rashid, S.M., el-Fakahany, A.F., Lashin, A.H., 1997. Seroepidemiological studies for toxoplasmosis among out- and inpatients in Benha University Hospitals, Qualyobia Governorate. J Egypt Soc Parasitol 27, 223-231.

Hussein, A.H., Ali, A.E., Saleh, M.H., Nagaty, I.M., Rezk, A.Y., 2001. Prevalence of toxoplasma infection in Qualyobia governorate, Egypt. J Egypt Soc Parasitol 31, 355-363.

Jacobs, M.R., Mason, P.R., 1978. Prevalence of Toxoplasma antibodies in Southern Africa. S Afr Med J 53, 619-621.

Julvez, J., Magnaval, J.F., Baixench, M.T., Maron, I., 1994. [Seroepidemiology of toxoplasmosis on Mayotte (Comores Archipelago)]. Med Trop (Mars) 54, 343-344.

Julvez, J., Magnaval, J.F., Meynard, D., Perie, C., Baixench, M.T., 1996. [Seroepidemiology of toxoplasmosis in Niamey, Niger]. Med Trop (Mars) 56, 48-50.

Kassem, H.H., Morsy, T.A., 1991. The prevalence of anti-Toxoplasma antibodies among pregnant women in Benghazi, (S.P.L.A.J.) Libya. J Egypt Soc Parasitol 21, 69-74.

Khadre, M.A., el Nageh, M.M., 1987. Serological survey for toxoplasmosis in Tripoli, S.P.L.A.J. Libya. Trans R Soc Trop Med Hyg 81, 761-763.

Ledru, E., Diagbouga, S., Ledru, S., Cauchoix, B., Yameogo, M., Chami, D., Sanou, O., Chiron, J.P., 1995. A study of Toxoplasma and Cytomegalovirus serology in tuberculosis and in HIV-infected patients in Burkina Faso. Acta Trop 59, 149-154.

Lelong, B., Rahelimino, B., Candolfi, E., Ravelojaona, B.J., Villard, O., Rasamindrakotroka, A.J., Kien, T., 1995. [Prevalence of toxoplasmosis in a population of pregnant women in Antananarivo (Madagascar)]. Bull Soc Pathol Exot 88, 46-49.

Maiga, Y., Samake, M., Marjolet, M., 1984. [Toxoplasmosis in Bamako (Republic of Mali). Incidence of the disease in women of child-bearing age]. Med Trop (Mars) 44, 319-322.

Marty, P., Reynes, J., Le Fichoux, Y., 1985. [Study of toxoplasmosis in pregnant women in Cameroon]. Bull Soc Pathol Exot Filiales 78, 623-628.

Millogo, A., Ki-Zerbo, G.A., Traore, W., Sawadogo, A.B., Ouedraogo, I., Peghini, M., 2000. [Toxoplasma serology in HIV infected patients and suspected cerebral toxoplasmosis at the Central Hospital of Bobo-Dioulasso (Burkina Faso)]. Bull Soc Pathol Exot 93, 17-19.

Morvan, J.M., Mambely, R., Selekon, B., Coumanzi-Malo, M.F., 1999. [Toxoplasmosis at the Pasteur Institute of Bangui, Central African Republic (1996-1998): serological data]. Bull Soc Pathol Exot 92, 157-160.

Nabias, R., Ngouamizokou, A., Migot-Nabias, F., Mbou-Moutsimbi, R.A., Lansoud-Soukate, J., 1998. [Serological investigation of toxoplasmosis in patients of the M.I.P. center of Franceville (Gabon)]. Bull Soc Pathol Exot 91, 318-320.

Ndumbe, P.M., Andela, A., Nkemnkeng-Asong, J., Watonsi, E., Nyambi, P., 1992. Prevalence of infections affecting the child among pregnant women in Yaounde, Cameroon. Med Microbiol Immunol (Berl) 181, 127-130.

Negro Ponzi, A., Petruzzelli, E., Cappuccinelli, P., 1976. [Antibodies against the cytomegalic virus and Toxoplasma gondii in a sample of the Monrovia population]. G Batteriol Virol Immunol 69, 220-228.

Omland, T., Tonjum, A., Frentzel-Beyme, R.R., 1977. Prevalence of Toxoplasma gondii antibodies in different populations of native Liberians. Tropenmed Parasitol 28, 372-376.

Onadeko, M.O., Joynson, D.H., Payne, R.A., 1992. The prevalence of Toxoplasma infection among pregnant women in Ibadan, Nigeria. J Trop Med Hyg 95, 143-145.

Perez-Rendon Gonzalez, J., Lopez Caminero, A., 1992. [Seroprevalence of human toxoplasmosis in Ceuta]. Aten Primaria 9, 109-110.

Rodier, M.H., Berthonneau, J., Bourgoin, A., Giraudeau, G., Agius, G., Burucoa, C., Hekpazo, A., Jacquemin, J.L., 1995. Seroprevalences of Toxoplasma, malaria, rubella, cytomegalovirus, HIV and treponemal infections among pregnant women in Cotonou, Republic of Benin. Acta Trop 59, 271-277.

Simpore, J., Savadogo, A., Ilboudo, D., Nadambega, M.C., Esposito, M., Yara, J., Pignatelli, S., Pietra, V., Musumeci, S., 2006. Toxoplasma gondii, HCV, and HBV seroprevalence and co-infection among HIV-positive and -negative pregnant women in Burkina Faso. J Med Virol 78, 730-733.

Uneke, C.J., Duhlinska, D.D., Njoku, M.O., Ngwu, B.A., 2005. Seroprevalence of acquired toxoplasmosis in HIV-infected and apparently healthy individuals in Jos, Nigeria. Parassitologia 47, 233-236.
Woldemichael, T., Fontanet, A.L., Sahlu, T., Gilis, H., Messele, T., Rinke de Wit, T.F., Yeneneh, H., Coutinho, R.A., Van Gool, T., 1998. Evaluation of the Eiken latex agglutination test for anti-Toxoplasma antibodies and seroprevalence of Toxoplasma infection among factory workers in Addis Ababa, Ethiopia. Trans R Soc Trop Med Hyg 92, 401-403.
Zardi, O., Adorisio, E., Harare, O., Nuti, M., 1980. Serological survey of toxoplasmosis in Somalia. Trans R Soc Trop Med Hyg 74, 577-581.

Appendix 3. Result table, Ugandan HIV-patients

All HIV-samples used in the study, listed by group and sample ID. HA=headache, PP=photophobia, MCF=mental confusion, CV=convulsions, F=fever, W=weakness or paresis, V=vomiting. '+' indicates a positive test result or the presence of a symptom, '-' indicates a negative test result or the absence of a symptom. Blanks show that the information was not obtained.

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G1	H05-1	40	F		+	+	+	-	-	-	-	-	-	
G1	H05-2	56	F		+	-	-	+	-	+	-	+	-	
G1	H05-3	35	M		+	-	+	-	-	-	-	+	-	
G1	H05-4	34	F		+	-	-	+	-	+	-	+	-	
G1	H05-5	31	F		+	+	-	-	-	-	-	+	-	
G1	H05-6	45	M		+	-	-	+	-	-	-	+	-	
G1	H05-7	36	F		+	-	-	-	-	-	-	+	-	
G1	H05-8											+	-	
G1	H05-9	29	M		+	-	-	-	-	+	-	+	-	
G1	H05-10	28	F		+	+	-	+	-	-	-	+	-	
G1	H05-11	23	F		+	-	-	+	-	-	-	+	-	
G1	H05-12											+	-	
G1	H05-13											-	-	
G1	H05-14											+	-	
G1	H05-15	32	M		+	-	+	+	-	-	-	+	-	
G1	H05-16											-	-	
G1	H05-17	44	M		+	-	-	+	-	+	-	-	-	
G1	H05-18											-	-	
G1	H05-19	37	F		+	-	-	+	-	-	+	+	+	II
G1	H05-20	32	F		+	-	-	+	-	-	+	+	+	I/II
G1	H05-21	36	F		+	-	+	+	-	-	-	-	+	III
G1	H05-22		M		-	-	-	+	-	+	-	-	-	
G1	H05-23	27	F		-	-	-	+	-	+	-	+	-	
G1	H05-24		F		-	-	+	+	-	-	-	+	-	
G1	H05-25		M		-	-	+	-	-	+	-	-	-	
G1	H05-26		M		+	-	-	-	-	-	-	+	+	II
G1	H05-27	39	M		-	-	-	-	-	+	-	-	-	
G1	H05-28	40	F		+	-	-	+	-	-	-	+	-	
G1	H05-29											-	-	
G1	H05-30	27	F		-	-	-	+	-	-	-	-	-	
G1	H05-31											-	+	II
G1	H05-32	43	M		-	-	-	-	-	+	-	-	-	
G1	H05-33											-	-	
G1	H05-34	40	M		-	-	-	-	-	+	-	-	-	
G1	H05-35											+	+	II
G1	H05-36		M		-	-	-	-	-	+	-	+	-	
G1	H05-37											-	+	II
G1	H05-38											+	-	
G1	H05-39											+	-	
G1	H05-40											-	-	
G1	H05-41											+	-	
G1	H05-42		M		-	-	+	-	-	-	-	+	-	
G1	H05-43											+	+	II
G1	H05-44											-	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G1	H05-45											-	+	II
G1	H05-46											-	-	
G1	H05-47											+	-	
G1	H05-48											-	-	
G1	H05-49											-	-	
G1	H05-50											-	-	
G1	H05-51											+	+	I
G1	H05-52											+	+	I
G1	H05-53											+	-	
G1	H05-54											-	-	
G1	H05-55											-	+	I
G1	H05-56												-	
G1	H05-57											+	-	
G1	H05-58											+	-	
G1	H05-59											+	-	
G1	H05-60											+	+	II
G1	H05-61											-	-	
G1	H05-62											-	-	
G1	H05-63											-	-	
G1	H05-64											+	+	I
G1	H05-65											+	+	I/II
G1	H05-66											-	-	
G1	H05-67											+	+	II
G1	H05-68											-	+	II
G1	H05-69		F									-	-	
G1	H05-70	42	M									-	-	
G1	H05-71	54	M									-	+	II
G1	H05-72											-	-	
G1	H05-73	60	M									-	-	
G1	H05-74											-	-	
G1	H05-75		F									-	-	
G1	H05-76		M		+	-	+	-	-	-	-	-	-	
G1	H05-77	22	M		+	+	-	+	-	-	-	-	-	
G1	H05-78		M		+	-	+	+	-	+	-	-	-	
G1	H05-79				-	-	-	-	-	+	-	+	-	
G1	H05-80	30	F		-	-	-	-	+	-	-	-	-	
G1	H05-81				+	-	-	-	-	+	-	+	+	II
G1	H05-82	35	M		+	+	-	-	-	-	-	+	+	II
G1	H05-83	37	M		-	-	-	-	-	-	-	+	-	
G1	H05-84		M		-	-	+	+	-	+	-	+	-	
G1	H05-85	25	M		+	-	-	-	+	-	-	+	-	
G1	H05-86				-	-	+	-	-	+	-	-	-	
G1	H05-87				-	-	-	-	+	-	-	-	-	
G1	H05-88				+	-	-	-	+	-	-	+	-	
G1	H05-89	25	F		-	-	-	-	-	+	-	-	-	
G1	H05-90	27	M		+	-	-	-	-	+	-	+	+	III
G1	H05-91				+	-	+	-	-	-	-	+	-	
G1	H05-92	39	M		-	-	-	-	-	-	-	-	-	
G1	H05-93				-	-	-	-	-	+	-	-	-	
G1	H05-94				-	-	+	-	-	+	-	-	-	
G1	H05-95				-	-	-	-	-	-	-	+	-	
G1	H05-96		M		-	-	+	-	-	-	-	-	-	
G1	H05-97				-	-	+	-	+	-	+	+	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G1	H05-98		M		-	-	-	-	-	-	-	+	-	
G1	H05-99		F		-	-	+	-	-	+	-	-	-	
G1	H05-100		F		-	-	-	-	-	-	-	+	-	
G1	H05-101		F		-	-	-	-	-	-	-	-	-	
G1	H05-102	46	M		-	-	-	-	-	-	-	+	-	
G1	H05-123		F		+	-	-	-	-	+	-	+	+	II
G1	H05-124		F		+	+	-	-	-	-	-	-	-	
G1	H05-125				+	-	-	-	+	-	-	+	+	II
G1	H05-126				-	-	+	-	-	-	-	+	+	II
G1	H05-127	23	F		+	-	-	-	-	-	-	+	+	I
G1	H05-128		F		+	+	+	-	-	-	-	-	-	
G1	H05-129				-	-	-	-	-	+	-	+	-	
G1	H05-130		F		-	-	-	+	+	-	-	+	-	
G1	H05-131	50	F		-	-	+	-	-	-	-	+	-	
G1	H05-132				+	-	-	-	-	-	-	-	+	II
G1	H05-133				+	-	+	-	-	-	-	+	-	
G1	H05-134				-	-	+	-	-	-	-	-	-	
G1	H05-135	42	F		+	-	-	-	-	-	-	+	-	
G1	H05-136	31	F		-	-	-	-	-	+	-	+	-	
G1	H05-137	35	F		+	-	+	-	-	-	-	+	+	I/II
G1	H05-138	48	M		+	-	-	-	-	-	-	+	+	II
G1	H05-139	32	M		+	+	+	-	-	-	-	+	-	
G1	H05-140	38	M		-	-	+	-	-	-	-	-	-	
G1	H05-141	38	F		+	-	-	+	-	-	-	+	-	
G1	H05-142	25	M		+	-	-	-	+	-	-	+	-	
G1	H05-143		M		-	-	-	-	-	-	-	-	-	
G1	H05-144	41	M		-	-	+	-	-	-	-	+	-	
G1	H05-145	35	F		-	-	-	-	-	-	-	+	+	I
G1	H05-146		F		-	-	+	-	-	-	-	-	-	
G1	H05-147		F		+	-	+	-	-	-	-	-	-	
G1	H05-148	35	F		+	+	-	-	-	-	-	+	+	III
G1	H05-149		F		-	-	+	-	-	-	-	+	-	
G1	H05-150		F		+	-	-	-	+	+	-	-	-	
G1	H05-151	43	F		-	-	+	-	-	-	-	-	-	
G2	H06-1	34	M	263								-	-	
G2	H06-2	17	F									-	+	
G2	H06-3	40	F									+	-	
G2	H06-4	1	F									-	+	
G2	H06-5	37	M	234								-	-	
G2	H06-6	45	M									+	-	
G2	H06-7	25	M									-	-	
G2	H06-8	28	F									+	-	
G2	H06-9	20	F									-	+	
G2	H06-10	19	F	600								+	-	
G2	H06-11	36	F	175								+	+	
G2	H06-12	22	M	22								+	-	
G2	H06-13	27	F									+	-	
G2	H06-14	45	M									+	-	
G2	H06-15	5	F									-	+	
G2	H06-16	32	F									+	-	
G2	H06-17	40	F	344								-	-	
G2	H06-18	25	F	150								-	-	
G2	H06-19	18	F									+	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G2	H06-20	30	F									-	-	
G2	H06-21	30	F									+	-	
G2	H06-22	17	M									-	-	
G2	H06-23	42	F									-	-	
G2	H06-24	17	F									+	-	
G2	H06-25	25	F									+	-	
G2	H06-26	43	M									+	-	
G2	H06-27	50	F									-	-	
G2	H06-28	26	F									+	-	
G2	H06-29	55	F									+	-	
G2	H06-30	35	F									+	-	
G2	H06-31	26	F									-	-	
G2	H06-32	30	F									-	-	
G2	H06-33	30	M									-	-	
G2	H06-34	30	F	68								+	-	
G2	H06-35	40	F	142								-	-	
G2	H06-36	31	F									+	-	
G2	H06-37	24	F									-	-	
G2	H06-38	34	M									+	-	
G2	H06-39	30	F									+	+	
G2	H06-40	29	F									-	-	
G2	H06-41	28	M									-	-	
G2	H06-42	38	F									+	-	
G2	H06-43	42	F	71								-	-	
G2	H06-44	25	F	648								-	-	
G2	H06-45	45										-	-	
G2	H06-46	46	M	323								-	-	
G2	H06-47	26	F									-	-	
G2	H06-48	47	F	143								+	-	
G2	H06-49	45	F	244								+	-	
G2	H06-50	30	M									-	-	
G2	H06-51		M									-	-	
G2	H06-52	49	F									+	+	
G2	H06-53		F									-	-	
G2	H06-54		M									+	-	
G2	H06-55		M									+	+	
G2	H06-56	44	F									+	-	
G2	H06-57		F									+	-	
G2	H06-58		F									+	+	
G2	H06-60	24	M									+	+	
G2	H06-61	37	M									-	-	
G2	H06-62	22	F									+	+	
G2	H06-63	19	F									-	-	
G2	H06-64	30	F									+	-	
G2	H06-65	33	F									+	+	
G2	H06-66	41	M									+	-	
G2	H06-67	37	F	502								+	-	
G2	H06-68	38	M									+	+	
G2	H06-69	24	M									-	+	
G2	H06-70	18	F									+	+	
G2	H06-71	42	F									+	+	
G2	H06-72	25	M									-	-	
G2	H06-73	36	M	70								-	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G2	H06-74	46	F	103								+	-	
G2	H06-75	30	F									+	+	
G2	H06-76	36	M									-	-	
G2	H06-77	31	F									+	-	
G2	H06-78	23	F									-	-	
G2	H06-79	27	F									+	-	
G2	H06-80	20	F									-	-	
G2	H06-81	35	F									+	+	
G2	H06-82	44	F									+	-	
G2	H06-83	30	M									-	-	
G2	H06-84	38	M									+	+	
G2	H06-85	23	F	1837								+	-	
G2	H06-86	47	M									-	-	
G2	H06-87	35	M									+	-	
G2	H06-88	23	F									-	-	
G2	H06-89	27	M									-	-	
G2	H06-90	30	F	329								-	-	
G2	H06-91	25	M									+	-	
G2	H06-92	32	M									+	-	
G2	H06-93	32	F									+	-	
G2	H06-94	26	F									+	-	
G2	H06-95	23	F									-	-	
G2	H06-96	33	M									-	-	
G2	H06-97	36	F									+	-	
G2	H06-98	32	F									-	-	
G2	H06-99	34	F									-	-	
G2	H06-100	26	F									+	-	
G2	H06-101	55	M									+	-	
G2	H06-102	37	F									-	-	
G2	H06-103	29	F									-	-	
G2	H06-104	25	M									-	+	
G2	H06-105	40	M									-	-	
G2	H06-106	25	F									-	-	
G2	H06-107	24	M									+	-	
G2	H06-108	40	M									+	-	
G2	H06-109	25	F									+	-	
G2	H06-110	40	F									-	-	
G2	H06-111	45	M									+	-	
G2	H06-112	26	M									-	-	
G2	H06-113	30	M									-	-	
G2	H06-114	33	F									+	-	
G2	H06-115	30	F									-	-	
G2	H06-116	24	F									-	-	
G2	H06-117	34	M									-	-	
G2	H06-118	40	M									+	-	
G2	H06-119	21	F									-	-	
G2	H06-120	33	F									+	-	
G2	H06-121	27	F									+	-	
G2	H06-122	30	F									-	-	
G2	H06-123	41	F									+	-	
G2	H06-124	33	F									-	-	
G2	H06-125	35	F									+	-	
G2	H06-126	21	F									+	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G2	H06-127	39	F									+	-	
G2	H06-128	30	F									-	-	
G2	H06-129	45	M									+	-	
G2	H06-130	34	F									+	-	
G2	H06-131	26	F									+	-	
G3	H07-1	35	M		-	-			-			-	-	
G3	H07-2	45	M		-	+			-			+	-	
G3	H07-3	35	M		+	-			+			-	-	
G3	H07-4	26	F		+	-			+			+	+	
G3	H07-5	33	F		-	-			-			+	-	
G3	H07-6	43	M		+	-			-			+	+	
G3	H07-7	34	M		-	-			-			-	-	
G3	H07-8	35	F		+	+			-			-	-	
G3	H07-9	42	F		-	-			-			-	-	
G3	H07-10	34	M		-	-			-			+	-	
G3	H07-11	45	F		-	-			+			-	-	
G3	H07-12	29	F		+	-			+			+	-	
G3	H07-13	32	F		+	-			+			-	-	
G3	H07-14	30	M		+	-			+			-	-	
G3	H07-15	26	M		-	+			+			-	-	
G3	H07-16	29	F		-	-			+			-	-	
G3	H07-17	21	F		+	-			-			+	+	
G3	H07-18	22	F		-	+			-			+	-	
G3	H07-19	23	M		-	-			+			+	-	
G3	H07-20	40	M		+	-			+			-	-	
G3	H07-21	45	F		+	+			+			+	-	
G3	H07-22	33	M		+	+			+			+	-	
G3	H07-23	32	F		-	+			-			+	-	
G3	H07-24	36	F		+	+			+			+	-	
G3	H07-25	34	F		-	-			-			+	-	
G3	H07-26	42	F		+	+			-			-	-	
G3	H07-27	23	M		-	-			-			-	-	
G3	H07-28	30	F		-	-			-			-	-	
G3	H07-29	26	F		-	+			-			+	+	
G3	H07-30	38	F		+	-			+			-	-	
G3	H07-31	33	F		+	-			+			+	+	
G3	H07-32	38	F		+	-			+			-	-	
G3	H07-33	39	F		+	+			-			+	-	
G3	H07-34	30	M		+	+			+			+	-	
G3	H07-35	36	M		-	-			+			-	-	
G3	H07-36	28	F		+	+			+			-	-	
G3	H07-37	23	F		+	+			+			+	-	
G3	H07-38	25	F		-	-			+			-	-	
G3	H07-39	24	F		+	-			+			-	-	
G3	H07-40	30	F		+	-			+			-	-	
G3	H07-41	37	F		+	-			+			-	-	
G3	H07-42	27	M		+	-			-			+	+	
G3	H07-43	30	F		+	-			+			-	-	
G3	H07-44											+	-	
G3	H07-45											-	-	
G4	E06-1	26	F	19	+	-	-	-	+	-	-	+	-	
G4	E06-2	31	F	152	+	-	-	-	+	+	-	+	-	
G4	E06-3	45	F	37	+	-	-	+	+	-	-	+	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G4	E06-4	30	M	5	+	+	+	-	-	-	-	+	-	
G4	E06-5	34	M	42	+	-	+	+	-	-	-	+	-	
G4	E06-6	31	M	33	+	+	+	-	-	-	-	-	-	
G4	E06-7	28	F	0	+	-	+	-	+	-	-	-	+	
G4	E06-8	33	F	397	+	-	-	-	+	-	-	+	-	
G4	E06-9	43	M	73	+	-	-	-	+	-	-	-	+	
G4	E06-10	39	M	182	+	-	+	+	+	+	-	+	-	
G4	E06-11	30	F	336	+	-	+	-	+	-	-	+	+	
G4	E06-12	30	M	17	+	+	-	-	+	-	-	+	-	
G4	E06-13	36	M	15	+	-	-	-	+	-	+	-	-	
G4	E06-14	45	M	16	-	-	-	-	-	-	+	+	-	
G4	E06-15	47	M	49	+	+	-	-	-	-	+	-	-	
G4	E06-16	38	M	49	+	-	+	-	+	-	-	-	-	
G4	E06-17	40	M	13	+	-	-	-	+	-	-	+	-	
G4	E06-18	20	F	235	+	-	+	-	+	-	-	-	-	
G4	E06-19	54	F	31	+	-	-	-	+	-	-	+	-	
G4	E06-20	38	M	9	+	-	-	-	+	-	-	+	-	
G4	E06-21	29	F	611	+	-	-	-	+	-	+	+	-	
G4	E06-22	41	F	97	+	+	+	+	-	-	+	-	-	
G4	E06-23	21	F	20	+	-	-	-	+	-	+	+	-	
G4	E06-24	42	F	128	+	+	+	+	+	-	+	+	-	
G4	E06-25	43	F	6	+	+	+	-	+	+	-	-	-	
G4	E06-26	53	M	26	+	+	+	-	+	-	-	-	-	
G4	E06-27	50	F	22	+	-	+	-	+	-	-	-	-	
G4	E06-28	30	F	89	-	+	-	-	+	-	-	-	-	
G4	E06-29	35	F	166	+	+	+	-	+	-	+	+	+	
G4	E06-30	35	F	4	+	+	-	-	+	-	-	+	-	
G4	E06-31	30	M	226	+	-	+	-	+	-	-	+	-	
G4	E06-32	35	F	524	-	+	-	-	+	-	-	-	-	
G4	E06-33	20	M	292	+	-	-	-	+	-	-	+	-	
G4	E06-34	43	F	68	-	+	+	+	+	-	+	-	-	
G4	E06-35	38	M	6	-	-	-	-	+	-	-	+	-	
G4	E06-36	35	F	1	-	+	-	-	+	+	-	+	-	
G4	E06-37	20	F	71	+	+	-	-	+	-	+	+	-	
G4	E06-38	35	M	354	+	-	-	-	+	-	+	+	-	
G4	E06-39	30	M	164	+	-	-	-	+	-	-	-	-	
G4	E06-40	23	M	44	+	-	+	-	+	-	-	+	-	
G4	E06-41	33	M	46	+	+	-	-	+	-	-	+	-	
G4	E06-42	25	M	248	+	-	-	-	+	-	+	-	-	
G4	E06-43	45	M	9	+	+	+	-	+	-	+	-	-	
G4	E06-44	33	M	689	+	-	-	-	+	-	-	+	-	
G4	E06-45	27	F	47	+	+	-	-	+	-	+	+	-	
G4	E06-46	35	F	150	+	-	-	-	+	-	+	+	-	
G4	E06-47	21	F	61	+	-	-	-	-	-	+	+	-	
G4	E06-48	34	M	11	-	+	-	-	+	-	-	+	-	
G4	E06-49	45	M	4	+	-	-	-	+	-	-	+	-	
G4	E06-50	43	M	45	+	-	-	-	+	-	-	+	+	
G4	E06-51	28	F	8	+	-	-	-	-	-	-	+	-	
G4	E06-52	25	F	93	+	-	-	-	+	-	+	+	-	
G4	E06-53	30	F	240	+	-	-	-	+	-	-	-	-	
G4	E06-54	32	M	210	+	-	-	-	+	-	-	+	-	
G4	E06-55	47	F	353	-	-	-	-	-	-	-	+	-	
G4	E06-56	36	M	75	+	-	-	-	+	-	+	+	+	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G4	E06-57	29	M	48	+	-	+	-	+	-	+	-	-	
G4	E06-58	19	F	304	+	-	-	-	+	-	-	+	-	
G4	E06-59	30	F	1	+	-	-	-	+	-	+	-	-	
G4	E06-60	25	F	83	+	-	-	-	+	-	+	-	-	
G4	E06-61	25	F	23	+	-	-	-	+	-	-	+	-	
G4	E06-62	45	F	78	+	-	-	-	+	-	-	+	-	
G4	E06-63	55	F	3	+	-	+	-	+	-	+	-	-	
G4	E06-64	28	F	94	+	-	-	-	+	-	+	+	+	
G4	E06-65	25	F	19	+	-	-	+	+	-	+	+	-	
G4	E06-66	22	F	419	-	-	-	-	+	-	-	+	-	
G4	E06-67	60	F	442	-	-	-	-	+	-	+	+	-	
G4	E06-68	28	F	1361	+	-	-	-	+	-	+	+	-	
G4	E06-69	35	F	18	+	-	-	-	+	-	+	+	-	
G4	E06-70	36	M	532	+	-	-	-	+	-	+	+	+	
G4	E06-71	35	M	303	+	-	-	-	+	-	-	-	-	
G4	E06-72	33	M	26	+	+	+	-	+	-	+	+	-	
G4	E06-73	15	M	4	+	-	-	-	+	-	+	-	-	
G4	E06-74	27	M	429	+	-	-	-	+	-	-	-	-	
G4	E06-75	28	F	33	+	-	-	-	+	-	+	-	-	
G4	E06-76	27	F	16	+	+	+	-	+	-	+	+	-	
G4	E06-77	30	F	20	+	-	-	-	+	-	+	-	-	
G4	E06-78	22	F	100	+	-	+	-	+	-	-	+	-	
G4	E06-79	40	M	15	+	-	+	-	+	-	-	+	-	
G4	E06-80	38	F	67	+	-	+	-	+	+	-	+	-	
G4	E06-81	60	F	463	+	-	-	-	+	-	+	+	-	
G4	E06-82	23	F	181	+	-	-	-	+	-	+	+	-	
G4	E06-83	37	M	356	+	+	-	-	+	-	+	+	-	
G4	E06-84	22	F	111	+	-	-	-	+	-	+	+	-	
G4	E06-85	34	M	9	+	-	-	-	+	+	-	+	-	
G4	E06-86	24	M	3	+	-	-	-	+	-	+	+	-	
G4	E06-87	40	M	81	+	-	+	-	+	-	-	+	-	
G4	E06-88	44	M	143	+	+	+	-	+	-	+	+	-	
G4	E06-89	42	M	456	+	-	+	+	+	-	+	+	-	
G4	E06-90	32	M	240	+	+	+	-	+	-	-	+	-	
G4	E06-91	29	M	1	+	-	+	-	+	-	-	+	-	
G4	E06-92	41	M	181	+	-	-	-	+	-	-	+	-	
G4	E06-93	38	M	48	+	-	+	-	+	-	+	-	-	
G4	E06-94	35	F	16	+	+	+	-	+	-	+	+	-	
G4	E06-95	28	F	493	+	+	+	+	+	+	+	-	-	
G4	E06-96	26	F	208	+	+	+	+	+	-	+	-	-	
G4	E06-97	28	F	7	+	-	-	-	+	-	+	+	+	
G4	E06-98	32	F	97	+	-	-	-	+	-	+	-	-	
G4	E06-99	20	M	129	+	-	-	-	+	-	+	-	-	
G4	E06-100	28	F	70	+	-	-	-	+	-	+	+	-	
G4	E06-101	22	F	403	+	-	-	-	+	-	-	-	-	
G4	E06-102	44	F	5	+	-	-	-	+	-	+	+	-	
G4	E06-103	45	M	162	-	-	-	-	+	-	+	+	-	
G4	E06-104	25	F	312	+	-	-	-	+	-	+	+	-	
G4	E06-105	36	M	10	+	-	-	-	+	-	+	+	-	
G4	E06-106	50	F	46	+	-	-	-	+	-	-	+	-	
G4	E06-107	32	F	141	+	-	-	-	+	+	+	-	-	
G4	E06-108	26	F	82	+	-	+	+	+	-	-	+	-	
G4	E06-109	55	F	9	+	-	-	-	+	-	+	+	-	

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G4	E06-110	49	M	39	+	+	-	+	+	-	-	+	-	
G4	E06-111	40	F	240	-	-	-	-	+	-	+	+	-	
G4	E06-112	38	M	47	-	-	+	-	+	-	-	+	-	
G4	E06-113	32	F	559	+	-	-	-	+	-	-	-	-	
G4	E06-114	25	F	0	+	-	+	-	+	-	-	-	-	
G4	E06-115	35	F	419	+	+	+	+	+	-	+	-	-	
G4	E06-116	45	F	130	+	+	+	-	+	-	+	-	-	
G4	E06-117	37	M	0	+	-	+	-	+	-	-	+	-	
G4	E06-118	34	F	8	+	-	+	-	+	-	+	+	-	
G4	E06-119	26	F	2	+	+	+	+	+	-	-	+	-	
G4	E06-120	29	F	38	+	-	+	-	+	-	-	+	-	
G4	E06-121	25	F	2	+	-	+	-	+	-	-	+	-	
G4	E06-122	30	M	21	+	-	-	-	+	-	-	+	-	
G4	E06-123	47	M	21	+	-	+	-	+	-	+	+	-	
G4	E06-124	37	M	37	+	-	-	-	+	-	-	+	-	
G4	E06-125	30	F	324	+	+	-	-	+	-	+	+	-	
G4	E06-126	28	M	4	+	-	-	-	+	-	+	+	-	
G4	E06-127	38	F	330	-	-	-	-	-	-	-	+	-	
G4	E06-128	44	M	175	+	+	-	-	-	-	-	-	-	
G4	E06-129	27	M	2	+	-	+	-	+	-	+	+	-	
G4	E06-130	60	M	16	+	-	-	-	+	-	+	+	-	
G4	E06-131	40	M	3	+	-	+	-	+	-	-	-	+	
G4	E06-132	28	F	28	+	-	-	-	+	-	+	-	+	
G4	E06-133	60	F	59	+	-	-	-	+	-	+	-	-	
G4	E06-134	32	F	89	+	-	-	-	+	-	+	+	-	
G4	E06-135	28	F	41	+	-	+	-	+	-	+	-	-	
G4	E06-136	34	M	237	+	-	-	-	+	-	+	+	+	
G4	E06-137	40	M	18	+	-	+	-	+	-	+	+	-	
G4	E06-138	32	M	16	+	-	-	-	+	-	+	+	-	
G4	E06-139	28	F	22	+	-	-	-	+	-	+	+		
G4	E06-140	35	F	21	+	-	-	-	+	-	+	+		
G4	E06-141	40	M	23	+	-	-	-	+	-	-	+		
G4	E06-142	28	M	195	+	-	-	-	+	-	-	+		
G4	E06-143	40	F	61	+	-	-	-	+	-	-	+		
G4	E06-144	30	M	2	+	+	-	-	+	-	+	+		
G4	E06-145	33	F	67	+	-	-	-	+	-	-	-		
G4	E06-146	35	F	128	+	-	-	-	+	-	+	-		
G4	E06-147	52	F	41	+	-	-	-	+	-	+	-		
G4	E06-148	35	F	14	+	-	+	-	+	-	-	-		
G4	E06-149	22	F	13	-	-	-	-	+	-	+	-		
G4	E06-150	58	M	1	+	-	+	-	+	-	-	-		
G4	E06-151	33	F	172	+	-	-	-	+	-	-	-		
G4	E06-152	31	M	296	+	-	-	-	+	-	-	-		
G4	E06-153	24	F	15	+	-	+	-	+	-	-	+		
G4	E06-154	43	M	715	-	-	-	-	+	-	-	+		
G4	E06-155	40	M	78	-	-	-	-	+	+	+	+		
G4	E06-156	29	M	122	-	-	-	-	+	+	+	+		
G4	E06-157	41	M	674	-	-	-	-	+	+	-	+		
G4	E06-158	30	M	515	+	-	-	-	+	+	+	-		
G4	E06-159	43	M	42	-	-	-	-	+	+	+	-		
G4	E06-160	30	F	976	+	-	+	-	+	-	-	+		
G4	E06-161	23	F	20	+	-	-	-	+	-	+	+		
G4	E06-162	27	F	77	+	-	-	-	+	-	+	+		

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G4	E06-163	45	F	507	+	-	-	-	+	-	+	-		
G4	E06-164	50	F	107	+	-	+	-	+	-	-	+		
G4	E06-165	37	F	16	+	-	-	-	+	-	-	+		
G4	E06-166	28	F	3	+	-	-	-	+	-	-	-		
G4	E06-167	34	F	5	+	-	-	-	+	-	+	+		
G4	E06-168	36	M	46	+	-	-	-	+	-	-	+		
G4	E06-169	19	F	3	-	-	-	-	+	-	-	+		
G4	E06-170	27	F	11	-	+	-	-	+	-	-	+		
G4	E06-171	22	F	232	+	-	-	-	+	-	-	+		
G4	E06-172	27	F	3	+	-	-	-	+	-	+	-		
G4	E06-173	34	F	56	+	-	-	-	+	-	+	+		
G4	E06-174	28	F	238	+	-	-	-	+	+	-	+		
G4	E06-175	35	F	518	+	-	-	-	-	-	-	+		
G4	E06-176	31	M	1	+	-	-	-	+	-	+	-		
G4	E06-177	42	M	52	+	-	-	-	+	-	-	+		
G4	E06-178	32	F	638	-	-	-	-	+	-	-	+		
G4	E06-179	35	F	635	-	-	-	-	+	-	-	+		
G4	E06-180	26	M	151	+	-	-	-	+	-	+	+		
G4	E06-181	35	M	3	-	-	-	-	+	-	+	+		
G4	E06-182	34	M	5	+	-	+	-	+	-	-	+		
G4	E06-183	54	M	152	+	-	+	-	-	+	-	-		
G4	E06-184	24	F	3	+	-	-	-	+	-	-	+		
G4	E06-185	45	M	115	+	-	-	-	+	-	-	+		
G4	E06-186	38	F	13	-	+	+	-	+	-	-	+		
G4	E06-187	36	F	143	-	-	+	+	+	-	-	+		
G4	E06-188	25	F	3	-	-	+	-	+	-	+	+		
G4	E06-189	28	F	0	+	-	-	-	+	+	+	+		
G4	E06-190	34	F	2	+	-	-	-	+	-	+	+		
G4	E06-191	26	M	619	-	-	-	-	+	-	-	+		
G4	E06-192	42	M	9	+	-	+	-	+	-	-	+		
G4	E06-193	35	M	29	+	-	+	-	+	-	-	-		
G4	E06-194	34	M	0	+	-	-	-	+	-	-	+		
G4	E06-195	28	M	54	+	-	-	-	+	-	+	+		
G4	E06-196	28	M	103	+	-	-	-	+	-	-	+		
G4	E06-197	40	M	332	+	-	-	-	+	-	-	+		
G4	E06-198	38	F	6	+	-	-	-	+	-	-	+		
G4	E06-199	25	F	448	-	-	-	-	+	-	+	-		
G4	E06-200	32	F	34	-	-	-	-	+	-	-	+		
G4	E06-201	21	F	1	+	-	-	-	+	-	+	-		
G4	E06-202	30	M	77	+	-	-	-	+	-	-	+		
G4	E06-203	34	M	6	-	-	-	-	+	-	+	-		
G4	E06-204	20	M	90	+	-	-	-	+	-	+	+		
G4	E06-205	43	M	219	-	-	+	-	+	+	-	-		
G4	E06-206	29	F	41	-	-	-	-	-	-	-	-		
G4	E06-207	29	F	128	+	-	-	-	+	-	-	-		
G4	E06-208	36	M	155	+	+	-	-	-	-	-	-		
G4	E06-209	33	F	10	+	-	-	-	+	-	-	-		
G4	E06-210	35	F	2	+	-	-	-	+	-	+	+		
G4	E06-211	25	F	305	+	+	+	-	+	-	-	-		
G4	E06-212	28	M	12	+	-	-	-	+	-	-	-		
G4	E06-213	21	M	391	+	-	-	-	+	-	-	+		
G4	E06-214	40	F	70	-	-	+	-	+	-	-	-		
G4	E06-215	25	F	12	+	+	+	-	+	-	-	-		

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G4	E06-216	45	F	2	+	-	-	-	+	-	-	+		
G4	E06-217	40	F	2	-	-	-	-	+	-	+	-		
G4	E06-218	50	F	194	+	-	-	-	+	-	+	-		
G4	E06-219	25	F	708	+	-	-	-	+	-	-	+		
G4	E06-220	27	M	363	-	-	+	-	+	-	-	+		
G4	E06-221	20	F	272	-	-	-	-	+	-	+	-		
G4	E06-222	44	F	52	+	-	-	-	+	-	-	+		
G4	E06-223	24	F	3	+	-	+	-	+	-	-	+		
G4	E06-224	32	F	95	+	-	-	-	-	+	+	+		
G4	E06-225	32	F	111	-	+	+	-	-	-	+	+		
G4	E06-226	40	F	101	+	-	-	-	+	-	+	+		
G4	E06-227	55	M	290	+	+	-	-	+	-	-	+		
G4	E06-228	35	M	62	+	-	+	-	+	-	-	-		
G4	E06-229	30	M	416	+	+	-	-	+	-	-	+		
G4	E06-230	45	M	91	+	-	-	-	+	-	-	+		
G4	E06-231	40	M	0	+	-	+	-	+	-	+	-		
G4	E06-232	43	F	442	-	-	-	-	+	-	-	-		
G4	E06-233	20	F	109	+	-	-	-	+	-	+	+		
G4	E06-234	49	F	9	-	-	+	-	+	-	-	-		
G4	E06-235	38	F	1	-	-	-	-	+	-	-	+		
G4	E06-236	25	M	514	+	+	-	-	+	-	+	-		
G4	E06-237	34	M	612	+	-	+	-	+	-	-	+		
G4	E06-238	25	M	7	+	-	-	-	+	-	+	+		
G4	E06-239	68	F	96	+	-	-	-	-	-	+	-		
G4	E06-240	34	M	20	+	-	-	-	+	-	-	-		
G4	E06-241	37	F	53	+	+	+	-	-	+	-	-		
G4	E06-242	32	F	223	+	+	+	-	+	+	-	-		
G4	E06-243	27	F	60	+	-	-	+	+	-	-	+		
G4	E06-244	23	M	11	+	-	+	-	-	-	-	+		
G4	E06-245	54	F	21	+	+	-	-	+	-	-	+		
G4	E06-246	30	F	0	-	-	-	-	+	+	-	-		
G4	E06-247	24	F	16	-	-	+	+	+	-	-	+		
G4	E06-248	52	F	0	+	-	-	-	+	-	-	+		
G4	E06-249	45	F	608	-	-	-	-	+	-	+	+		
G4	E06-250	23	F	465	+	-	-	-	+	-	-	+		
G4	E06-251	44	M	90	-	+	-	-	-	-	+	+		
G4	E06-252	33	M	13	+	-	+	-	+	-	-	-		
G4	E06-253	24	F	352	-	-	-	-	+	-	+	-		
G4	E06-254	30	M	500	+	-	-	-	-	-	-	-		
G5	E06-255	27	F	320	-	-	-	-	-	-	-	+		
G5	E06-256	25	F	66	-	-	-	-	+	-	-	+		
G5	E06-257	19	F	214	-	-	-	-	+	-	-	-		
G5	E06-258	33	F	465	-	-	-	-	-	-	-	+		
G5	E06-259	24	F	1218	-	-	-	-	-	-	-	-		
G5	E06-260	30	F	542	-	-	-	-	-	-	-	-		
G5	E06-261	24	M	159	-	-	-	-	-	-	-	-		
G5	E06-262	42	F	552	-	-	-	-	+	-	-	+		
G5	E06-263	28	F	417	-	-	-	-	-	-	-	-		
G5	E06-264	35	M	33	-	-	-	-	-	-	-	-		
G5	E06-265	26	F	250	-	-	-	-	+	-	-	-		
G5	E06-266	32	F	753	-	-	-	-	+	-	-	-		
G5	E06-267	31	F	110	-	-	-	-	+	-	-	+		
G5	E06-268	37	M	588	-	-	-	-	-	-	-	-		

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G5	E06-269	40	F	231	-	-	-	-	-	-	-	+		
G5	E06-270	20	F	142	-	-	-	-	+	-	-	-		
G5	E06-271	49	F	302	-	-	-	-	+	-	-	-		
G5	E06-272	32	M	358	-	-	-	-	+	-	-	+		
G5	E06-273	27	M	169	-	-	-	-	-	-	-	-		
G5	E06-274	27	F	160	-	-	-	-	-	-	-	-		
G5	E06-275	29	M	771	-	-	-	-	+	-	-	-		
G5	E06-276	30	M	42	-	-	-	-	-	-	-	+		
G5	E06-277	27	F	354	-	-	-	-	-	-	-	-		
G5	E06-278	31	M	134	-	-	-	-	+	-	-	-		
G5	E06-279	40	F	159	-	-	-	-	-	-	-	+		
G5	E06-280	20	F	721	-	-	-	-	-	-	-	+		
G5	E06-281	49	M	20	-	-	-	-	-	-	-	-		
G5	E06-282	32	M	102	-	-	-	-	-	-	-	+		
G5	E06-283	34	F	127	+	-	-	-	+	-	-	-		
G5	E06-284	25	M	1135	+	-	-	-	+	-	-	-		
G5	E06-285	27	M	355	+	-	-	-	+	-	-	+		
G5	E06-286	37	M	137	+	-	-	-	+	-	+	-		
G5	E06-287	25	M	265	+	-	-	-	+	-	+	+		
G5	E06-288	37	M	4	-	-	-	-	-	-	-	+		
G5	E06-289	24	M	4	-	-	-	-	+	-	-	-		
G5	E06-290	45	F	2	-	-	-	-	+	-	+	-		
G5	E06-291	35	F	57	-	-	-	-	-	-	-	-		
G5	E06-292	27	M	842	-	-	-	-	-	-	-	-		
G5	E06-293	27	F	681	+	-	-	-	-	-	-	-		
G5	E06-294	25	F	844	+	-	-	-	-	-	-	+		
G5	E06-295	30	M	42	-	-	-	-	-	-	-	-		
G5	E06-296	51	F	160	+	+	-	-	+	-	-	-		
G5	E06-297	30	M	1	+	-	-	-	+	-	-	+		
G5	E06-298	42	F	328	+	-	-	-	-	-	-	-		
G5	E06-299	35	M	356	+	-	-	-	+	-	-	+		
G5	E06-300	28	F	423	-	-	-	-	-	-	-	-		
G5	E06-301	28	M	860	-	-	-	-	-	-	-	-		
G5	E06-302	70	M	100	-	+	-	-	+	-	-	+		
G5	E06-303	57	M	530	+	-	-	-	+	-	-	-		
G5	E06-304	28	M	90	+	-	-	-	-	-	-	+		
G5	E06-305	27	F	464	+	-	-	-	-	-	-	+		
G5	E06-306	37	M	80	-	-	-	-	-	-	-	+		
G5	E06-307	29	F	252	-	-	-	-	-	-	-	+		
G5	E06-308	47	M	518	-	-	-	-	-	-	-	+		
G5	E06-309	19	F	253	+	+	-	-	-	-	-	+		
G5	E06-310	31	F	10	-	-	-	-	-	-	-	-		
G5	E06-311	34	F	110	-	-	-	-	-	-	-	+		
G5	E06-312	20	F	26	-	-	-	-	-	-	-	+		
G5	E06-313	36	F	129	+	-	-	-	+	-	-	-		
G5	E06-314	26	M	7	-	-	-	-	+	-	-	+		
G5	E06-315	32	F	85	-	-	-	-	+	-	-	-		
G5	E06-316	27	M	968	-	-	-	-	-	-	-	-		
G5	E06-317	27	M	520	-	-	-	-	+	-	-	-		
G5	E06-318	30	F	375	-	-	-	-	+	-	-	+		
G5	E06-319	31	F	149	-	-	-	-	-	-	-	-		
G5	E06-320	37	M	105	-	-	-	-	-	-	-	+		
G5	E06-321	43	M	136	-	-	-	-	+	-	-	+		

Group	Sample	Age	Sex	CD4	HA	PP	MCF	CV	F	W	V	IgG	PCR	Type
G5	E06-322	24	F	298	+	+	-	-	+	-	-	-		
G5	E06-323	45	F	404	+	+	-	-	+	-	-	-		
G5	E06-324	33	F	307	-	+	-	-	-	-	-	+		
G5	E06-325	30	F	799	-	+	-	-	-	-	-	-		
G5	E06-326	24	F	397	-	+	-	-	-	-	-	+		
G5	E06-327	37	M	117	-	-	-	-	-	-	-	-		
G5	E06-328	24	F	623	+	-	-	-	-	-	-	-		
G5	E06-329	29	F	450	+	-	-	-	-	-	-	+		
G5	E06-330	50	M	585	-	-	-	-	-	-	-	+		
G5	E06-331	44	M	9	-	-	-	-	-	-	-	+		
G5	E06-332	30	F	26	-	-	-	-	-	-	-	-		
G5	E06-333	52	F	218	-	-	-	-	-	-	-	-		
G5	E06-334	25	F	655	-	-	-	-	-	-	-	-		
G5	E06-335	23	F	600	+	+	-	-	-	+	-	-		
G5	E06-336	31	M	343	-	-	-	-	-	+	-	-		
G5	E06-337	77	M	276	+	-	-	-	-	+	-	-		
G5	E06-338	24	F	26	-	-	-	-	-	-	-	-		
G5	E06-339	46	F	263	-	-	-	-	-	+	-	+		
G5	E06-340	36	F	5	-	+	-	-	-	+	-	+		
G5	E06-341	48	F	244	+	+	-	-	-	-	-	+		
G5	E06-342	25	F		-	-	-	-	-	-	-	+		
G5	E06-343	22	F	894	-	-	-	-	-	-	-	-		
G5	E06-344	30	F	8	-	-	-	-	+	-	-	+		
G5	E06-345	29	F	165	-	-	-	-	-	-	-	+		
G5	E06-346	27	F	481	-	-	-	-	-	-	-	-		
G5	E06-347	54	F	225	-	-	-	-	-	-	-	+		
G5	E06-348	27	F	769	-	-	-	-	-	-	-	+		
G5	E06-349	26	F	326	-	-	-	-	-	-	-	-		
G5	E06-350	35	F	438	-	-	-	-	-	-	-	+		
G5	E06-351	18	F	1008	-	-	-	-	-	-	-	+		
G5	E06-352	42	F	747	-	-	-	-	-	-	-	+		
G5	E06-353	23	M	602	-	-	-	-	+	-	-	+		
G5	E06-354	33	F	710	-	-	-	-	-	-	-	-		
G5	E06-355	26	F	450	-	-	-	-	+	-	-	+		
G5	E06-356	30	M	583	-	+	-	-	-	-	-	+		
G5	E06-357	42	M	152	-	-	-	-	+	-	-	-		
G5	E06-358	50	F	759	-	-	-	-	-	-	-	+		
G5	E06-359	30	F	142	-	+	-	-	+	-	-	-		
G5	E06-360	31	M	278	-	+	-	-	+	-	-	+		
G5	E06-361	45	F	378	+	+	-	-	-	-	-	-		
G5	E06-362	28	F	190	+	-	-	-	+	-	-	-		
G5	E06-363	30	F	522	-	-	-	-	+	-	-	-		
G5	E06-364	22	F	6	-	-	-	-	-	-	-	+		
G5	E06-365	27	M	150	-	-	-	-	-	-	-	-		
G5	E06-366	28	F	126	+	-	-	-	+	-	-	-		
G5	E06-367	31	F	215	-	-	-	-	+	-	-	-		
G5	E06-368	40	F	423	-	-	-	-	-	-	-	+		
G5	E06-369	22	F	105	-	-	-	-	+	-	-	-		
G5	E06-370	42	M	136	-	-	-	-	-	-	-	+		

Appendix 4. Chicken genotyping studies

	N	Seropos (%)	Isolates (%	Genotyping markers	Type I	Type II	Type III	Atypical	Multiple	Reference
Europe and USA	**1284**	**411 (32)**	**114**		**0%**	**82%**	**16%**	**2%**	**0%**	
Austria	830	302 (36)	67	SAG2		67				Dubey et al. 2005a
Italy	80	11 (14)	3	10 RFLP		3				Dubey et al. 2008a
Poland	20	6 (30)	2	10 RFLP				2		Dubey et al. 2008ª
Portugal	225	61 (27)	12	SAG2		8	4			Dubey et al. 2006e
USA (Illinois)	11	11 (100)	11	11 RFLP		11				Dubey et al. 2007c
USA (Ohio and Mass.)	118	20 (17)	19	SAG2		5	14			Dubey et al. 2003c
Central and South America	**1508**	**714 (47)**	**568**		**13%**	**7%**	**24%**	**54%**	**1%**	
Argentina (La Plata)	29	19 (65)	9	SAG2	1	1	7			Dubey et al. 2003e
Argentina (Santiago)	61	25 (41)	17	SAG2	4	3	10			Dubey et al. 2005f
Brazil (Minas Gerais)	28	15 (54)	18	SAG2	17		1			Brandão et al. 2006
Brazil (Amazon)	50	33 (66)	24	SAG2	14		10			Dubey et al. 2006ª
Brazil (Rio Grande do Sul)	50	19 (38)	19	11 RFLP				19		Dubey et al. 2007b
Brazil (Pará)	34	20 59)	15	11 RFLP				15		Dubey et al. 2007b
Brazil (Paraná)	40	16 (40)	13	SAG2	7		6			Dubey et al. 2003d
Brazil (Rio de Janeiro)	198	129 (65)	48	SAG2	34		13		1	Dubey et al. 2003a
Brazil (São Paulo)	82	33 (40)	21	SAG2	13		8			Dubey et al. 2002
Brazil 13 states, retyping			151	10 RFLP	1		5	143	2	Dubey et al. 2008b
Chile	85	47 (55)	21	5 RFLP		17	3	1		Dubey et al. 2006b
Colombia	77	32 (42)	24	SAG2	7		17			Dubey et al. 2005b
Costa Rica	144	60 (42)	32	5 RFLP	5		1	26		Dubey et al. 2006c
Grenada	102	53 (52)	36	SAG2	5	1	29	1		Dubey et al. 2005h
Guatemala	50	37 (74)	8	SAG2	3		5			Dubey et al. 2005e
Guyana	76	50 (66)	35	11 RFLP			2	33		Dubey et al. 2007a
Mexico	208	13 (6)	6	SAG2	1		5			Dubey et al. 2004b
Nicaragua	98	84 (86)	48	5 RFLP	6	3	6	29	4	Dubey et al. 2006d

	N	Seropos (%)	Isolates (%	Genotyping markers	Type I	Type II	Type III	Atypical	Multiple	Reference
Peru	50	13 (26)	10	SAG2	7		3			Dubey et al. 2004a
Venezuela	46	16 (35)	13	SAG2		3	10			Dubey et al. 2005d
Asia	**524**	**143 (27)**	**13**		**0%**	**38%**	**46%**	**15%**	**0%**	
Indonesia	94	24 (26)	1	10 RFLP				1		Dubey et al. 2008ª
Sri Lanka	100	39 (39)	11	SAG2		5	6			Dubey et al. 2005g
Vietnam	330	80 (24)	1	10 RFLP				1		Dubey et al. 2008ª
North Africa/Middle East	**262**	**117 (45)**	**44**		**0%**	**45%**	**55%**	**0%**	**0%**	
Egypt	121	49 (41)	19	SAG2,10 RFLP for 7 samples		3	16			Dubey et al. 2003b, Velmurugan, 2008
Iran	45	23 (51)	6	GRA6 seq			6			Zia-Ali et al. 2007
Israel	96	45 (47)	19	SAG 2		17	2			Dubey et al. 2004c
Sub-Saharan Africa	**311**	**75 (24)**	**20**		**5%**	**15%**	**65%**	**15%**	**0%**	
Burkina Faso	40	0 (0)	1	10 RFLP			1			Dubey et al. 2005c, Velmurugan, 2008
Congo	50	25 (50)	10	SAG2,10 RFLP for 5 sample	1	1	8			Dubey et al. 2005, Velmurugan, 2008
Ghana	64	41 (64)	2	10RFLP				2		Dubey et al. 2008a
Kenya	30	4 (13)	1	10 RFLP		1				Dubey et al. 2005c, Velmurugan, 2008
Mali	48	0 (0)	5	SAG2,10 RFLP for 4 samples		1	4			Dubey et al. 2005c, Velmurugan, 2008
Nigeria	79	5 (6)	1	10 RFLP				1		Velmurugan et al. 2008

References, Appendix 4

Brandao, G.P., Ferreira, A.M., Melo, M.N., Vitor, R.W., 2006. Characterization of Toxoplasma gondii from domestic animals from Minas Gerais, Brazil. Parasite 13, 143-149.

Dubey, J.P., Graham, D.H., Blackston, C.R., Lehmann, T., Gennari, S.M., Ragozo, A.M., Nishi, S.M., Shen, S.K., Kwok, O.C., Hill, D.E., Thulliez, P., 2002. Biological and genetic characterisation of Toxoplasma gondii isolates from chickens (Gallus domesticus) from Sao Paulo, Brazil: unexpected findings. Int J Parasitol 32, 99-105.

Dubey, J.P., Graham, D.H., da Silva, D.S., Lehmann, T., Bahia-Oliveira, L.M., 2003a. Toxoplasma gondii isolates of free-ranging chickens from Rio de Janeiro, Brazil: mouse mortality, genotype, and oocyst shedding by cats. J Parasitol 89, 851-853.

Dubey, J.P., Graham, D.H., Dahl, E., Hilali, M., El-Ghaysh, A., Sreekumar, C., Kwok, O.C., Shen, S.K., Lehmann, T., 2003b. Isolation and molecular characterization of Toxoplasma gondii from chickens and ducks from Egypt. Vet Parasitol 114, 89-95.

Dubey, J.P., Graham, D.H., Dahl, E., Sreekumar, C., Lehmann, T., Davis, M.F., Morishita, T.Y., 2003c. Toxoplasma gondii isolates from free-ranging chickens from the United States. J Parasitol 89, 1060-1062.

Dubey, J.P., Navarro, I.T., Graham, D.H., Dahl, E., Freire, R.L., Prudencio, L.B., Sreekumar, C., Vianna, M.C., Lehmann, T., 2003d. Characterization of Toxoplasma gondii isolates from free range chickens from Parana, Brazil. Vet Parasitol 117, 229-234.

Dubey, J.P., Venturini, M.C., Venturini, L., Piscopo, M., Graham, D.H., Dahl, E., Sreekumar, C., Vianna, M.C., Lehmann, T., 2003e. Isolation and genotyping of Toxoplasma gondii from free-ranging chickens from Argentina. J Parasitol 89, 1063-1064.

Dubey, J.P., Levy, M.Z., Sreekumar, C., Kwok, O.C., Shen, S.K., Dahl, E., Thulliez, P., Lehmann, T., 2004a. Tissue distribution and molecular characterization of chicken isolates of Toxoplasma gondii from Peru. J Parasitol 90, 1015-1018.

Dubey, J.P., Morales, E.S., Lehmann, T., 2004b. Isolation and genotyping of Toxoplasma gondii from free-ranging chickens from Mexico. J Parasitol 90, 411-413.

Dubey, J.P., Salant, H., Sreekumar, C., Dahl, E., Vianna, M.C., Shen, S.K., Kwok, O.C., Spira, D., Hamburger, J., Lehmann, T.V., 2004c. High prevalence of Toxoplasma gondii in a commercial flock of chickens in Israel, and public health implications of free-range farming. Vet Parasitol 121, 317-322.

Dubey, J.P., Edelhofer, R., Marcet, P., Vianna, M.C.B., Kwok, O.C.H., Lehmann, T., 2005a. Genetic and biologic characteristics of Toxoplasma gondii infections in free-range chickens from Austria. Veterinary Parasitology 133, 299-306.

Dubey, J.P., Gomez-Marin, J.E., Bedoya, A., Lora, F., Vianna, M.C.B., Hill, D., Kwok, O.C., Shen, S.K., Marcet, P.L., Lehmann, T., 2005b. Genetic and biologic characteristics of Toxoplasma gondii isolates in free-range chickens from Colombia, South America. Veterinary Parasitology 134, 67-72.

Dubey, J.P., Karhemere, S., Dahl, E., Sreekumar, C., Diabate, A., Dabire, K.R., Vianna, M.C., Kwok, O.C., Lehmann, T., 2005c. First biologic and genetic characterization of Toxoplasma gondii isolates from chickens from Africa (Democratic Republic of Congo, Mali, Burkina Faso, and Kenya). J Parasitol 91, 69-72.

Dubey, J.P., Lenhart, A., Castillo, C.E., Alvarez, L., Marcet, P., Sreekumar, C., Lehmann, T., 2005d. Toxoplasma gondii infections in chickens from Venezuela: Isolation, tissue distribution, and molecular characterization. Journal of Parasitology 91, 1332-1334.

Dubey, J.P., Lopez, B., Alvarez, M., Mendoza, C., Lehmann, T., 2005e. Isolation, tissue distribution, and molecular characterization of Toxoplasma gondii from free-range chickens from Guatemala. Journal of Parasitology 91, 955-957.

Dubey, J.P., Marcet, P.L., Lehmann, T., 2005f. Characterization of Toxoplasma gondii isolates in free-range chickens from Argentina. Journal of Parasitology 91, 1335-1339.

Dubey, J.P., Rajapakse, R.P.V.J., Ekanayake, D.K., Sreekumar, C., Lehmann, T., 2005g. Isolation and molecular characterization of Toxoplasma gondii from chickens from Sri Lanka. Journal of Parasitology 91, 1480-1482.

Dubey, J.P., Gennari, S.M., Labruna, M.B., Camargo, L.M., Vianna, M.C., Marcet, P.L., Lehmann, T., 2006a. Characterization of Toxoplasma gondii isolates in free-range chickens from Amazon, Brazil. J Parasitol 92, 36-40.

Dubey, J.P., Patitucci, A.N., Su, C., Sundar, N., Kwok, O.C., Shen, S.K., 2006b. Characterization of Toxoplasma gondii isolates in free-range chickens from Chile, South America. Vet Parasitol 140, 76-82.

Dubey, J.P., Su, C., Oliveira, J., Morales, J.A., Bolanos, R.V., Sundar, N., Kwok, O.C., Shen, S.K., 2006c. Biologic and genetic characteristics of Toxoplasma gondii isolates in free-range chickens from Costa Rica, Central America. Vet Parasitol.

Dubey, J.P., Sundar, N., Pineda, N., Kyvsgaard, N.C., Luna, L.A., Rimbaud, E., Oliveira, J.B., Kwok, O.C., Qi, Y., Su, C., 2006d. Biologic and genetic characteristics of Toxoplasma gondii isolates in free-range chickens from Nicaragua, Central America. Vet Parasitol 142, 47-53.

Dubey, J.P., Vianna, M.C., Sousa, S., Canada, N., Meireles, S., Correia da Costa, J.M., Marcet, P.L., Lehmann, T., Darde, M.L., Thulliez, P., 2006e. Characterization of Toxoplasma gondii isolates in free-range chickens from Portugal. J Parasitol 92, 184-186.

Dubey, J.P., Applewhaite, L., Sundar, N., Velmurugan, G.V., Bandini, L.A., Kwok, O.C., Hill, R., Su, C., 2007a. Molecular and biological characterization of Toxoplasma gondii isolates from free-range chickens from Guyana, South America, identified several unique and common parasite genotypes. Parasitology 134, 1559-1565.

Dubey, J.P., Sundar, N., Gennari, S.M., Minervino, A.H., Farias, N.A., Ruas, J.L., dos Santos, T.R., Cavalcante, G.T., Kwok, O.C., Su, C., 2007b. Biologic and genetic comparison of Toxoplasma gondii isolates in free-range chickens from the northern Para state and the southern state Rio Grande do Sul, Brazil revealed highly diverse and distinct parasite populations. Vet Parasitol 143, 182-188.

Dubey, J.P., Webb, D.M., Sundar, N., Velmurugan, G.V., Bandini, L.A., Kwok, O.C., Su, C., 2007c. Endemic avian toxoplasmosis on a farm in Illinois: clinical disease, diagnosis, biologic and genetic characteristics of Toxoplasma gondii isolates from chickens (Gallus domesticus), and a goose (Anser anser). Vet Parasitol 148, 207-212.

Dubey, J.P., Huong, L.T., Lawson, B.W., Subekti, D.T., Tassi, P., Cabaj, W., Sundar, N., Velmurugan, G.V., Kwok, O.C., Su, C., 2008a. Seroprevalence and isolation of Toxoplasma gondii from free-range chickens in Ghana, Indonesia, Italy, Poland, and Vietnam. J Parasitol 94, 68-71.

Dubey, J.P., Velmurugan, G.V., Chockalingam, A., Pena, H.F., de Oliveira, L.N., Leifer, C.A., Gennari, S.M., Bahia Oliveira, L.M., Su, C., 2008b. Genetic diversity of Toxoplasma gondii isolates from chickens from Brazil. Vet Parasitol 157, 299-305.

Dubey, J.R., Bhaiyat, M.I., de Allie, C., Macpherson, C.N., Sharma, R.N., Sreekumar, C., Vianna, M.C., Shen, S.K., Kwok, O.C., Miska, K.B., Hill, D.E., Lehmann, T., 2005h. Isolation, tissue distribution, and molecular characterization of Toxoplasma gondii from chickens in Grenada, West Indies. J Parasitol 91, 557-560.

Velmurugan, G.V., Dubey, J.P., Su, C., 2008. Genotyping studies of Toxoplasma gondii isolates from Africa revealed that the archetypal clonal lineages predominate as in North America and Europe. Vet Parasitol 155, 314-318.

Zia-Ali, N., Fazaeli, A., Khoramizadeh, M., Ajzenberg, D., Darde, M., Keshavarz-Valian, H., 2007. Isolation and molecular characterization of Toxoplasma gondii strains from different hosts in Iran. Parasitol Res.

Appendix 5. Chicken samples and serology results

Chickens were pooled according to their serological titre. Seronegative chickens were pooled in groups of 15 and fed to three cats, none of which shed oocysts or seroconverted. For simplicity this has been omitted from the table.

Sample No	Site	Sex	Age group	MAT titre	PCR+bioassay
Ch-1	Gayaza	F	3	1:80	Ch-1
Ch-2	Gayaza	M	4	1:40	Ch-2
Ch-3	Gayaza	M	3	0	
Ch-4	Gayaza	M	3	0	
Ch-5	Gayaza	F	4	1:5	Pool 5A
Ch-6	Gayaza	F	4	1:10	Pool 10A
Ch-7	Gayaza	M	2	1:10	Pool 10A
Ch-8	Gayaza	M	2	1:5	Pool 5A
Ch-9	Gayaza	M	2	0	
Ch-10	Gayaza	F	1	0	
Ch-11	Gayaza	M	1	1:5	Pool 5A
Ch-12	Gayaza	F	2	1:5	Pool 5A
Ch-13	Gayaza	F	2	0	
Ch-14	Gayaza	F	1	1:10	Pool 10A
Ch-15	Gayaza	F	1	1:40	Ch-15
Ch-16	Gayaza	M	4	0	
Ch-17	Gayaza	F	2	1:160	Ch-17
Ch-18	Gayaza	M	1	0	
Ch-19	Gayaza	M	3	1:5	Pool 5A
Ch-20	Gayaza	F	1	1:20	Ch-20
Ch-21	Gayaza	F	3	1:10	Pool 10A
Ch-22	Gayaza	F	3	1:10	Pool 10B
Ch-23	Gayaza	M	3	0	
Ch-24	Gayaza	M	4	0	
Ch-25	Gayaza	F	1	1:20	Ch-25
Ch-26	Buloba	M	1	0	
Ch-27	Buloba	M	1	0	
Ch-28	Buloba	F	2	0	
Ch-29	Buloba	M	1	0	
Ch-30	Buloba	F	1	1:10	Pool 10B
Ch-31	Buloba	F	1	0	
Ch-32	Buloba	M	3	0	
Ch-33	Buloba	M	1	0	
Ch-34	Buloba	M	1	0	
Ch-35	Buloba	M	1	0	
Ch-36	Buloba	M	1	0	
Ch-37	Buloba	M	1	0	
Ch-38	Buloba	M	1	0	
Ch-39	Buloba	F	3	0	
Ch-40	Buloba	M	1	0	
Ch-41	Buloba	F	1	0	

Sample No	Site	Sex	Age group	MAT titre	PCR+bioassay
Ch-42	Buloba	M	1	1:5	Pool 5B
Ch-43	Buloba	M	1	0	
Ch-44	Buloba	F	1	0	
Ch-45	Buloba	M	2	0	
Ch-46	Buloba	F	3	0	
Ch-47	Mukono	M	3	1:5	Pool 5B
Ch-48	Mukono	M	3	0	
Ch-49	Mukono	M	3	0	
Ch-50	Mukono	F	3	0	
Ch-51	Mukono	F	2	0	
Ch-52	Mukono	M	3	1:40	Ch-52
Ch-53	Mukono	M	4	1:5	Pool 5B
Ch-54	Mukono	M	3	0	
Ch-55	Mukono	F	1	1:20	Ch-55
Ch-56	Mukono	F	4	1:10	Pool 10B
Ch-57	Mukono	F	3	0	
Ch-58	Mukono	M	2	1:40	Ch-58
Ch-59	Mukono	F	4	1:10	Pool 10B
Ch-60	Mukono	F	3	0	
Ch-61	Mukono	F	4	1:40	Ch-61
Ch-62	Mukono	F	4	1:20	Ch-62
Ch-63	Mukono	F	3	0	
Ch-64	Mukono	F	1	1:320	Ch-64
Ch-65	Mukono	F	3	1:5	Pool 5B
Ch-66	Mulago	F	4	0	
Ch-67	Mulago	F	4	0	
Ch-68	Mulago	M	3	1:40	Ch-68
Ch-69	Mulago	F	2	0	
Ch-70	Mulago	F	4	1:40	Ch-70
Ch-71	Mulago	M	3	0	
Ch-72	Mulago	M	1	1:10	Pool 10C
Ch-73	Mulago	F	1	1:40	Ch-73
Ch-74	Mulago	F	2	1:10	Pool 10C
Ch-75	Mulago	F	3	1:10	Pool 10C
Ch-76	Mulago	M	3	0	
Ch-77	Mulago	M	2	0	
Ch-78	Mulago	M	3	1:20	Ch-78
Ch-79	Mulago	M	3	1:40	Ch-79
Ch-80	Mulago	M	3	0	
Ch-81	Mulago	M	3	1:40	Ch-81
Ch-82	Mulago	M	3	1:40	Ch-82
Ch-83	Mulago	M	3	1:40	Ch-83
Ch-84	Mulago	M	3	0	
Ch-85	Mulago	M	3	0	

Appendix 6. Rodent samples and serology results

Rodents were bioassayed and pooled according to their serological titre, but in 2007 no bioassay was performed since all samples were negative.

Sample no	Year	MAT-titre	PCR/bioassay
Rat-1	2006	0	Pool RA
Rat-2	2006	160	Rat-2
Rat-3	2006	0	Pool RA
Rat-4	2006	0	Pool RA
Rat-5	2006	40	Rat-5
Rat-6	2006	0	Pool RA
Rat-7	2006	0	Pool RB
Rat-8	2006	0	Pool RB
Rat-9	2006	0	Pool RB
Rat-10	2007	0	Serology only
Mouse-1	2006	0	Pool M
Mouse-2	2006	0	Pool M
Mouse-3	2006	0	Pool M
Mouse-4	2006	0	Pool M
Mouse-5	2006	0	Pool M
Mouse-6	2007	0	Serology only
Mouse-7	2007	0	Serology only
Mouse-8	2007	0	Serology only
Mouse-9	2007	0	Serology only

Appendix 7. Location of new SNPs detected by direct sequencing

Novel or unusual allelic variants detected through direct sequencing of chicken tissue samples are marked with grey boxes. In SAG1 the parasites in Ch-70 had the u-1 allele, which is similar to type I at one position and II and III at another. The Ch-25 parasites had one new allele for SAG1, but were otherwise like type I. Ch-83 had one new allele at SAG2-3' and Ch-70 had three novel SNPs for the SAG3 sequence.

SAG1

```
                 10        20        30        40        50        60        70        80        90       100
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Ch-25    CAATGTGCACCTGTAGGAAGCTGTAGTCACTGCTGATTCTCACTGTTCTCGGCAAGGGCCGACGACCGGAGTACAGTTTTTGTGGGCAGAGCCGTTGTGC
Ch-70    CAATGTGCACCTGTAGGAAGCTGTAGTCACTGCTGATTCTCGCTGTTCTCGGCAAGGGCCGACGACCGGAGTACAGTTTTTGTGGGCAGAGCCGCTGTGC
GT-1     CAATGTGCACCTGTAGGAAGCTGTAGTCACTGCTGATTCTCACTGTTCTCGGCAAGGGCCGACGACCGGAGTACAGTTTTTGTGGGCAGAGCCGTTGTGC
Me49     CAATGTGCACCTGTAGGAAGCTGTAGTCACTGCTGATTCTCGCTGTTCTCGGCAAGGGCTGACGACCGGAGTACAGTTTTTGTGGGCAGAGCCGCTGTGC
VEG      CAATGTGCACCTGTAGGAAGCTGTAGTCACTGCTGATTCTCGCTGTTCTCGGCAAGGGCTGACGACCGGAGTACAGTTTTTGTGGGCAGAGCCGCTGTGC

                110       120       130       140       150       160       170       180       190       200
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Ch-25    AGCTTTCCGTTCTTCTCGGTTGTGTCACATGTGTCATTGTCGTGTAAACACACGGTTGTATGTCGGTTTCGCTGCACCACTTCATTATTTCTTCTGGTTT
Ch-70    AGCTTTCCGTTGTTCTCGGTTGTGTCACATGTGTCATTGTCGTGTAAACACACGGTTGTATGTCGGTTTCGCTGCACCACTTCATTATTTCTTCTGGTTT
GT-1     AGCTTTCCGTTCTTCTCGGTTGTGTCACATGTGTCATTGTCGTGTAAACACACGGTTGTATGTCGGTTTCGCTGCACCACTTCATTATTTCTTCTGGTTT
Me49     AGCTTTCCGTTGTTCTCGGTTGTGTCACATGTGTCATTGTCGTGTAAACACACGGTTGTATGTCGGTTTCGCTGCACCACTTCATTATTTCTTCTGGTTT
VEG      AGCTTTCCGTTGTTCTCGGTTGTGTCACATGTGTCATTGTCGTGTAAACACACGGTTGTATGTCGGTTTCGCTGCACCACTTCATTATTTCTTCTGGTTT

                210       220       230       240       250       260       270       280       290       300
         ....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Ch-25    TTTGACGAGTATGTTTCCGAAGGCAGTGAGACGCGCCGTCACGGCAGGGGTGTTTGCCGCGCCCACACTGATGTCGTTCTTGTGATGTGGCGTTATGGCA
Ch-70    TTTGGCGAGTATGTTTCCGAAGGCAGTGAGACGCGCCGTCACGGCAGGGGTGTTTGCCGCGCCCACACTGATGTCGTTCTTGCGATGTGGCGCTATGGCA
GT-1     TTTGACGAGTATGTTTCCGAAGGCAGTGAGACGCGCCGTCACGGCAGGGGTGTTTGCCGCGCCCACACTGATGTCGTTCTTGCGATGTGGCGTTATGGCA
Me49     TTTGGCGAGTATGTTTCCGAAGGCAGTGAGACGCGCCGTCACGGCAGGGGTGTTTGCCGCGCCCACACTGATGTCGTTCTTGCGATGTGGCGCTATGGCA
VEG      TTTGGCGAGTATGTTTCCGAAGGCAGTGAGACGCGCCGTCACGGCAGGGGTGTTTGCCGCGCCCACACTGATGTCGTTCTTGCGATGTGGCGCTATGGCA
```

Appendix 8. Parasite density in mouse organs

	Mouse	Brain	Heart	Muscle	Lung
TgCkUg1	m34	27,383	78	1,911	5
(II)	m35 (tr)	354,404	160	0	38
	m36 (tr)	17,855	52	597	14
	average	**133,214**	**97**	**836**	**19**
TgCkUg2	m6	168,779	14,594	43,914	108
(II/III)	m8	2,285,529	3,327	217,195	103
	average	**1,227,154**	**8,960**	**130,554**	**106**
TgCkUg3	m38	12,690	16	2	0
(II)	m47 (tr)	234,607	140	67	65
	m48 (tr)	136,235	566	1	10
	average	**127,844**	**240**	**23**	**25**
TgCkUg5	m41	92,615	39	0	0
(II)	m53 (tr)	16,242	1,150	0	27
	m54 (tr)	541,394	276	1	0
	average	**216,750**	**488**	**0**	**9**
TgCkUg6	m40	7,697,924	306,000	0	244
(III)	m51 (tr)	4,614,821	56,553	1,006,725	487
	m52 (tr)	1,252,527	7,181	6,224	2
	average	**4,521,757**	**123,245**	**337,650**	**244**
TgCkUg7	m42	8,007	0	146	0
(II)	m55 (tr)	444,171	853	2,122	0
	m56 (tr)	82,689	3,280	206	597
	average	**178,289**	**1,378**	**825**	**199**
TgCkUg8	m43	15,503	0	5,703	0
(II)	m57 (tr)	51,014	86	559	2
	average	**33,258**	**43**	**3,131**	**1**
TgCkUg9	m44	61,775	1,597	0	41
(II)	m59 (tr)	433,236	1,761	129,121	0
	m60 (tr)	60,024	483	3,548	262
	average	**185,012**	**1,280**	**44,223**	**101**

Appendix 9. FASTA contig output example

Sample of output file, showing data for 3 of the 67,013 contigs.

```
>contig25406  VI, 3599926..3600620  length=694   numreads=16
GTTTAGGAAGAAAAAGTAGGAaCATTACTCCGTGTGTTCTATAGAGGTTGTGCGTGCATG
GTTGCACTGGCTGAGCCATGACACGGCAGGGgACGCCTCAGTGAACGGTTCCTCCCAACG
TCCCCGTGATCGGTGTGAGGATCTCGCGTCTTACAAAGGCCCACGAATGAGGGAACCCTG
TGGgCGCCCATCACCAATGTAGTACTGGGGCACACTGACAAAACGTCTACGCCCGCGCAT
GTGACGGgACGTTTTGTCCGGAAACCATCGGgCAAGTCATCAACTGTTTCCTCtCACGAA
AAAGAAAAaGAAACGTAaTGgCTTGTTCACATGAACCAATCACCAGCGAACGCTGgTCTG
TGTGCAGCTGCAATGTTGAAAGCGGAGGAGAATGCAGCGGCGTCTGCGCGTTGCTTGAGA
TGGCGGCTCGGGTTCaGaGCTGGACTCCACTAATAGTATTTTTCACTCTGGTTTTTGTAT
GTCATACGTTTCGAGAGTGCGGGTGTGTGGCTATCGCCAAACCCGTGTGTTCGCCTTGGC
TGGCAGGATATCTCCGCTCAGCCTCGACCAGGGGCATCGATGGGCCTCTTTACCCCTTGA
ACGAGGACCCTAATCTGTTCGGGcgcagaaggggtgttttccgtagtgccaatgcgtctg
cgactatagctagcttggatgtcgaggagtttcg
>contig25407  VIIa, 28..2058  length=2027   numreads=113
ccgccgcctcaacagagtcgtcgtgTcTTCGCTTCCCACTCGgAGACTCAAGTCCGCACC
CGCGCCCCGTGGATGGATGGACGACTAGCGGATCGACCGCCCGCGAACGCAGACAAACAG
TACGCAGAACGTGCACCCGCCGTCCGACGATACACGCGAcACCTGAGCCAAGCCGCCAAA
ACGCGACAACCAGTCGCACCACGATCCCCGACGTCGACTCCAGCGACACCCTACTCAGGA
CGATGCCCCTCGTGTAACTCCCGATTCACCCTCCGAACCCGTCGACCAATCGTCGTGGCC
ACTCAGGCCACGTCCCCCGACATCCTCAGACACGAGTAgCACATCGCACCGCTGCCTCCC
GTCACGGTCTCGCGTACCGGCTCCACCTGTCGCAGACGCCACGACGTCGCCTCAAACGTA
GGAGGGCCAGCCAGCTCGTGCCGCCGCCAATGCGTACCTGCGACCTCCGACGTCGTTCCC
TGGCGCTCCTTCGCAACTCATGTCTGCACAGTGCATTCGCCGCCCGACCTACAAGTGGCC
CCCCGACCTCCACATACCTCACGACCACATCTGCcTTCACTCCCATGTGCCGCGTCCATT
CTCTGCACGTCTCCTCCTCCTCGCCGCCAACTCCCCACAGACCACCCCTCATGCAAACCA
AACCTGACCCACTAGAAACACAACGCAACAGCCGCCCACCAGATCcATCCGTCGTCCGCA
GGGTCACTCACCGACtGTTCaCAAaGaaaaaaCACcAAACACAGTGACGAGTCAAATCGT
CCTCAACCCTGTCACCACCCACGCCTTCCCACGTACTTCCGACGCATTCCGCAAGCCGCG
ACCCCCTCACCGCCATCACGCGACACCGGTCTGTGGCCGGGACACCGTCGCCCCGGCATT
AGGCTCCCGCGCCGCTCCTGGATCACCTCCGCCGCCTCAACAGAGTCGTCGTGTCTTCGC
TTCCCACTCGGAGACTCAAGTCCGCACCCGCGCCCCGTGGATGGATGGACGACTAGCGGA
TCGACCGCCCGCGAACGCAGACAAACaGTACGCAGAACGTGCACCCGCCGTCCGACGATA
CACGCGACACCTGAGCCAAGCCGCCAAAACGCGAAAACcAGTCgCACCACGATCCCCGaC
GTCGACTCCAGCGACACCCTACTCAGGACGATGCCCCTCGTGTAACTCCCGATTCACCCT
CCGAACCCGTCGACCAATCGTCGTGGCCACTCAGGCCACGTCCCCCGACATCCTCAGACA
CGAGTAACACATCGCACCGCTGCCTCCCGTCACggtctcgcgtaccggctccacctgtcg
cagaCGCCACGACGTCGCCTCAAACGTAGGAAAGCCAGCCAGCTCGTGCCGCCGCCAATG
CGTACCTGcgacctccgacgtcgttccctggcactccgtcgcacctcatgtctgcacagt
gcattcgccgcccgacctacaagtggcccccgacctccacatacctcacgaccacatct
gccttcactcccatgtgccgcgtccattctctgcacgtctcctccacctcgccgccaacT
CCCCACAGaccacccctcatgcaaaccaaacctgacccactagaaacacaacgcaacagc
cgcccaccagatccatccgtcgtccgcagggcactcaccgactgttcacaaagcaccaaa
cacagtgacgagtcaaatcGTCCTCAACCCTGTCACCACCCACGCCTTCCCACGTACTTC
CGACGCATTCCGCAAGCCGCGACCCCCCTCACCGCCATCACGCGACACCGGTCTGTGGCC
GGGACACCGTCGCCCCGGCAGTAGGCTcccgcgccgctcctggatcacctccgccGCCTC
AACAGAGTCGTCGTGTCTTCGCTTCCCACTCGGAGATACAAATCCCTACCTGCACCCCGT
GGAGGGATAGAAGACTAGCGAGTGGACTACACGCCAACGTAGAATCGCACTatgccgtca
cagttaccgtcgctgcgcgcgtgatctgtttgtctaagtttagtaag
>contig25408  VIIa, 2143..2841  length=698   numreads=5
cgcagctgactgaggtGGAGTCGCACAATAAACCCTCGGTGTTGCGACTGCAACGTTAAG
ACATTCTACATCCCGCACTTACGGGAGTGTCCTGGATGATTGTCGCGACGAGACGTCAAC
CAAACCaCGAGTCAcTGgTAgCTTGAGTGATTCTTTGCCCACTTACgTTGGGGGTTGTT
tAtCCCATCaTCGGTGtGAGcTgCCAACGAcGTGGCgACCCAAAGATGAGtCTTCTTcGG
CATGAGGTTCCTGAGCTTCAAAAATTCAACACGACACTTGCAGAAGTGTTGCTCTGGTGC
TGGATCACTCACGACGCTCTCAAACGATATGATCTGGGTCGGCCAGGCGAACCAATCAAT
TTTTGAGTatgactgtttgtaatgtactgaatacattcagtatttcgtgatacagcagca
agcacttccgtacgcaatcacgctattAAAAACATTGacccaatcatcATTTgcttgtgg
ctggttttgagatgttgccgagtgcggagatccatcttgtacagttggctgtctgccagt
ctcgacatattgggcctccgtccatagacacaccgtgtttacgccgtgtacattgttctt
ggagaagcgcgttttggcgttatgagtcagaaaaacagcaaaatgcagcgctaccttccc
gtttggtaccaccagaaactgcaccagtccgcctgccg
```

Appendix 10. Examples of valid and invalid SNP calls

Example of a valid SNP call, where 7+7 reads are in agreement and no discordant reads were found (100% concordant reads).

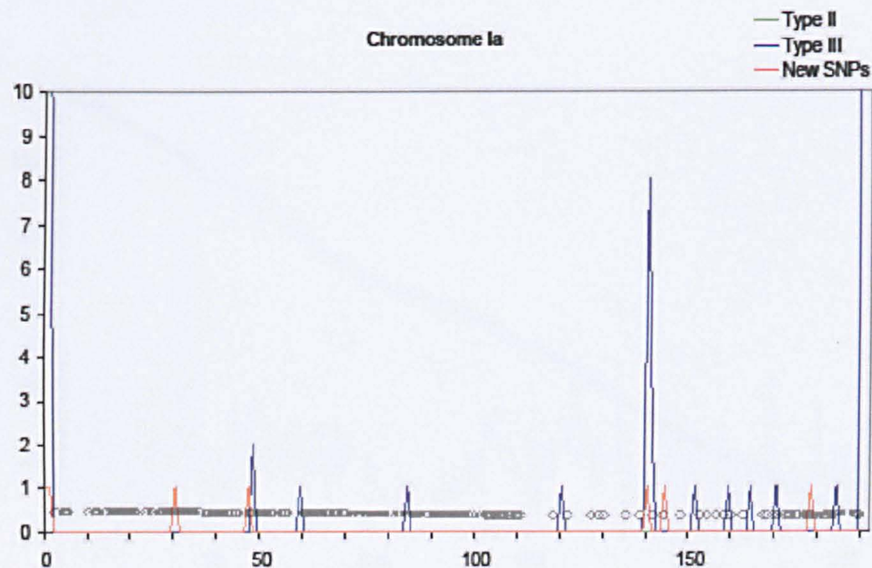
```
Reference                                  Variation           Frequency
Accno                                Start End    Seq    Seq          #fwd   #rev   #Var   #Tot    (%)
                                      Pos  Pos
____________________________________________  ________________________  ____________________
>Ia                                  34    34     G      A            7      7      14     14      100

Reads with Difference:
Ia                                 7+ GCA-CTGTATTGA-GTACGAGTG-AATCTT-GCA-TGT-GCGTGCGACGATGACGCG-AGTG 61
                                                                    **
E5TZWM301DILQP                   120- GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 66
E5G1LZ402IY10B                   139- GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 85
E5TZWM302JFG7B                   116+ GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 170
E5G1LZ401BIFWI                   148- GCA-CTGTATT-AAGTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 94
E47S0TK01CRYNG                    13+ GCAACTGTATTGA-GTACGAGTGCA-TCTTA-CA-TGTA-CGTGCGACGATGACGCGG-GTG 68
E5G1LZ401A4CEV   (2)              66+ GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 120
E5G1LZ402IKT4C                    44+ GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 98
E47S0TK01C6ZKR                   219- GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 165
E5G1LZ402I1A04                   219- GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 165
E5G1LZ401BEFOJ                    21+ GCA-CTGTATTGA-GTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 75
E5TZWM302F0M05                   227- GCA-CTGTATTGA-GTACGAGTGCA-TCTTA-CA-TGTA-CGTGCGACGATGACGCGG-GTG 173
E5G1LZ402GWVA3                     5+ GCA-CTGTATT-AAGTACGAGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 59
E47S0TK02GTOK7                     1+                    AGTG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 38
E5TZWM301DMQBA                   254-                      TG-AATCTTA-CAC-GTA-CGTGCGACGATGACGCGG-GTG 219
                                                                    **

Other Reads:


-----------------------------
```

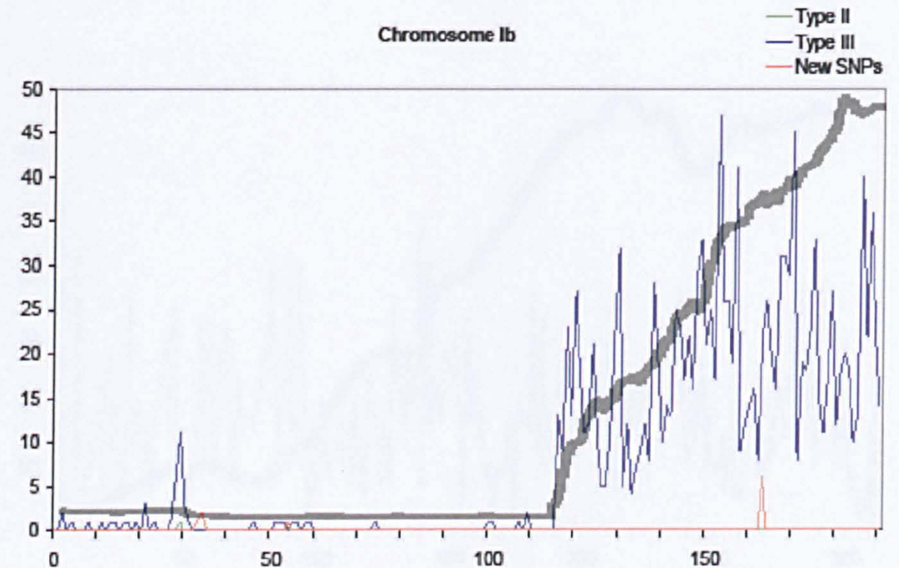
Example of an invalid SNP call, where 3+2 reads are in agreement, but 15 discordant reads were found (25% concordance).

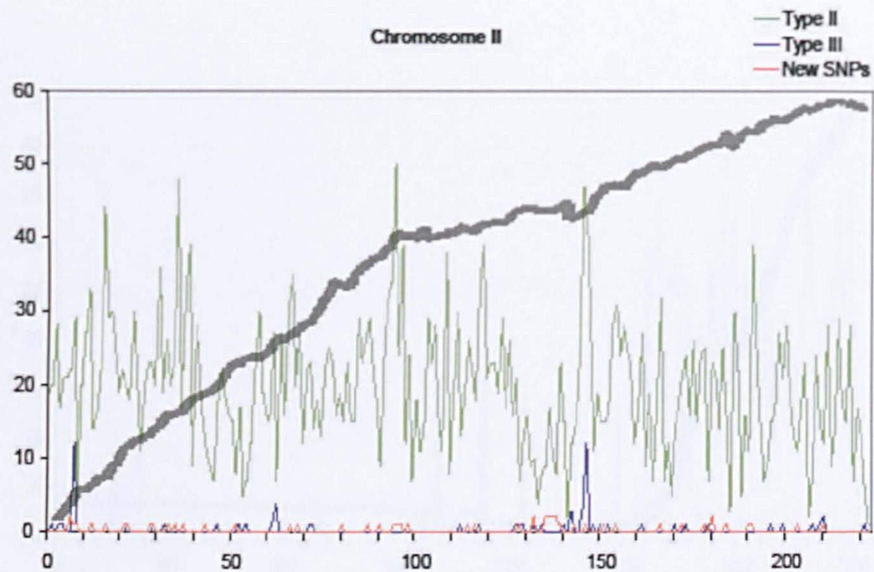
```
Reference                                     Variation       Frequency
Accno                                   Start End    Seq    Seq        #fwd   #rev   #Var   #Tot   (%)
                                         Pos  Pos
___________________________________     ________________________   _______________________
>Ia                                     85    85     G      C          3      2      5      20     25

Reads with Difference:
Ia                                    55+ GCG-AGTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATCGA-CTC-GTCTATT 115
                                                                         *******
E5G1LZ401BIFWI                       100- GCGG-GTGCGGGAGAGCAAAGCTCTACCATGCGGG--AC-AAAACAAAGCATC-AACTC-GTCTATT 40
E5G1LZ402GWVA3                        53+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATGCGGG--AC-AAAACAAAGCATC-AACTC-GTCTATT 113
E47S0TK02F8I7D                        23+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATGCGGG--AC-AAAACAAAGCATCGA-CTCC-TCTATT 83
E5TZWM302FFM0Q                       254-                       CTCTACCATGCGGG--AC-AAAACAAAGCATC-AACTC-GTCTATT 214
E5TZWM302GJ291                         1+                          TACCATGCGGG--AC-AAAACAAAGCATC-AACTC-GTCTATT 38
                                                                         *******
Other Reads:
                                                                         *******
E5TZWM301DILQP                        72- GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 12
E5G1LZ402IY10B                        91- GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 31
E5TZWM302JFG7B                       164+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 224
E47S0TK01CRYNG                        62+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 122
E5G1LZ401A4CEV   (2)                 114+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 174
E5G1LZ402IKT4C                        92+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 152
E47S0TK01C6ZKR                       171- GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 111
E5G1LZ402I1A04                       171- GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GACGAAAACAAAGCATC-AACTC-GTCTATT 110
E5G1LZ401BEFOJ                        69+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 129
E5TZWM302F0M05                       179- GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 119
E47S0TK02GTOK7                        32+ GCGG-GTGCGGGAGAGCAA-GCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 91
E5TZWM301DMQBA                       225- GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 165
E47S0TK02G69A6                         4+ GCGG-GTGCGGGAGAGCAAAGCTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 64
E5TZWM301BHM5I                       252-                       CTCTACCATG-GGG-GAC-AAAACAAAGCATC-AACTC-GTCTATT 212
E47S0TK01DOE4H                       248-                          TACCATG-GG-NGACGAAAACAAAGCATC-AACTC-GTCTATT 210
                                                                         *******

----------------------------
```

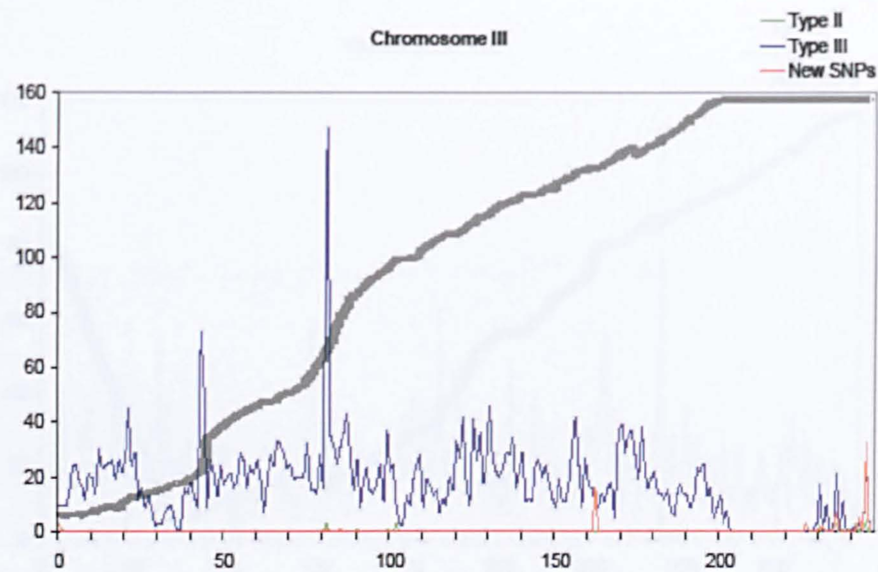
Appendix 11. SNP distribution and correlation with SNPs in the three reference strains

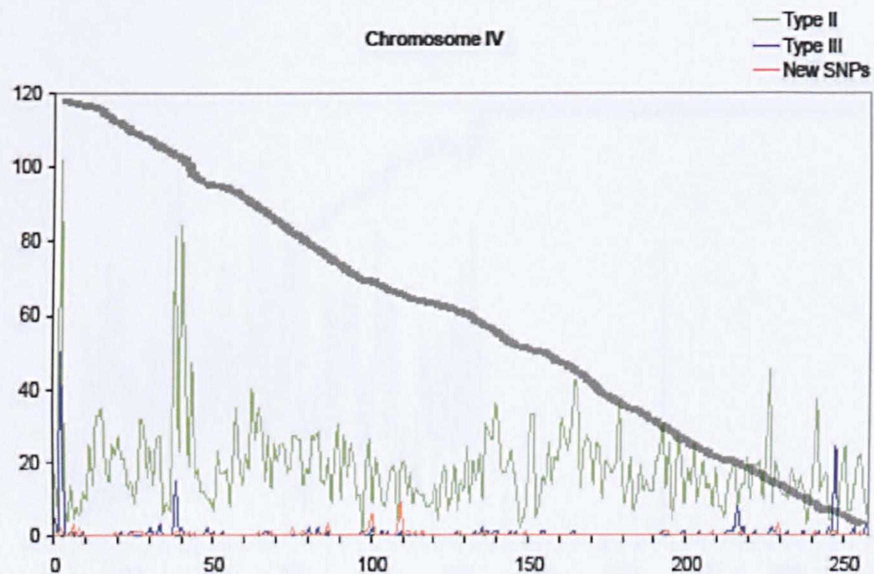
Overlay pictures for all chromosomes showing the match between SNP data forTgCkUg2 and the dominant SNP type from the comparison between GT1 (I), Me49 (II) and VEG (III). The underlying figures show SNPs for TgCkUg2 against Me49 (a type II strain) and/or VEG (a type III strain). Green depicts a type II background genotype in TgCkUg2 (TgCkUg2 = Me49, but different from VEG), a type III background is shown in blue (TgCkUg2 = VEG, but different from Me49) and novel SNPs are orange (TgCkUg2 is different from both Me49 and VEG). The overlying grey lines were created by looking for SNPs between GT1, Me49 and VEG at each position along the chromosome length. There is no change on the Y-axis for type I SNPs (where GT1 has the divergent base), but the line rises (+1) for every type II SNP and falls (-1) for type III SNPs (this scale is not shown in these graphs, but is depicted in Figure 5-10). These graph show a clear correlation between type I dominated areas (flat lines) and regions scarce in SNPs in TgCkUg2.

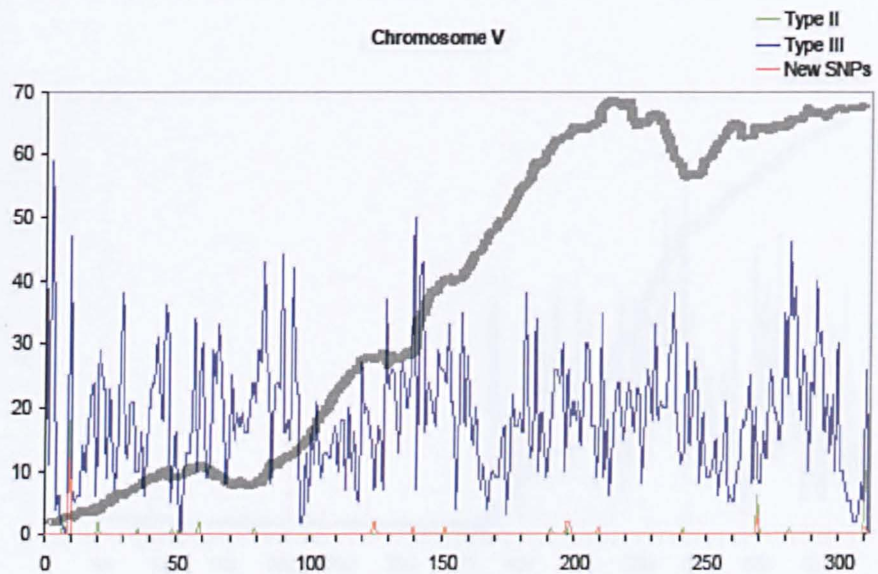


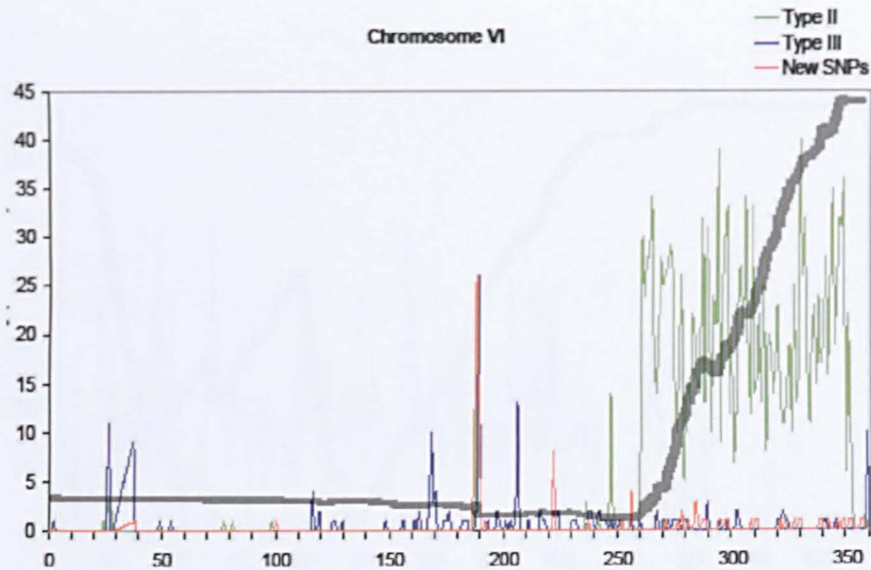


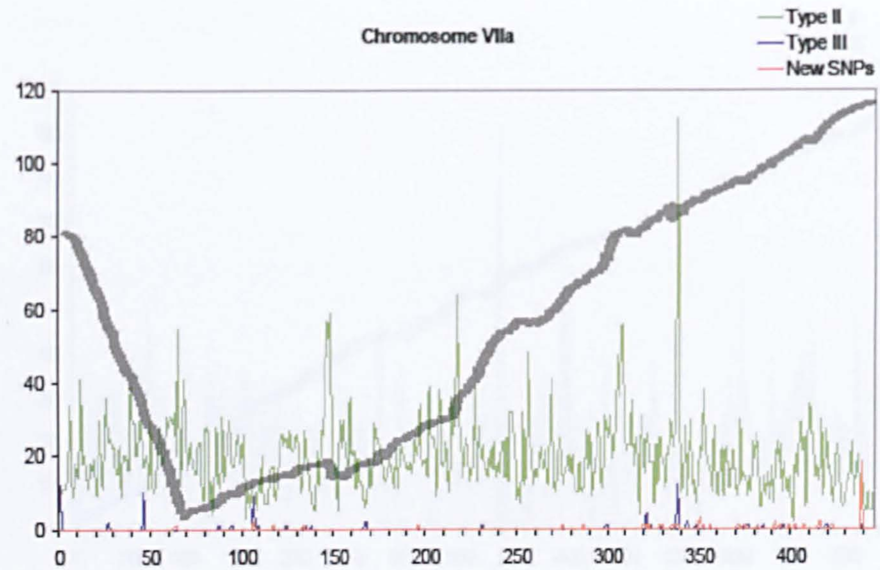

Chromosome II
Type II
Type III
New SNPs
60
50
40
30
20
10
0
0
50
100
150
200

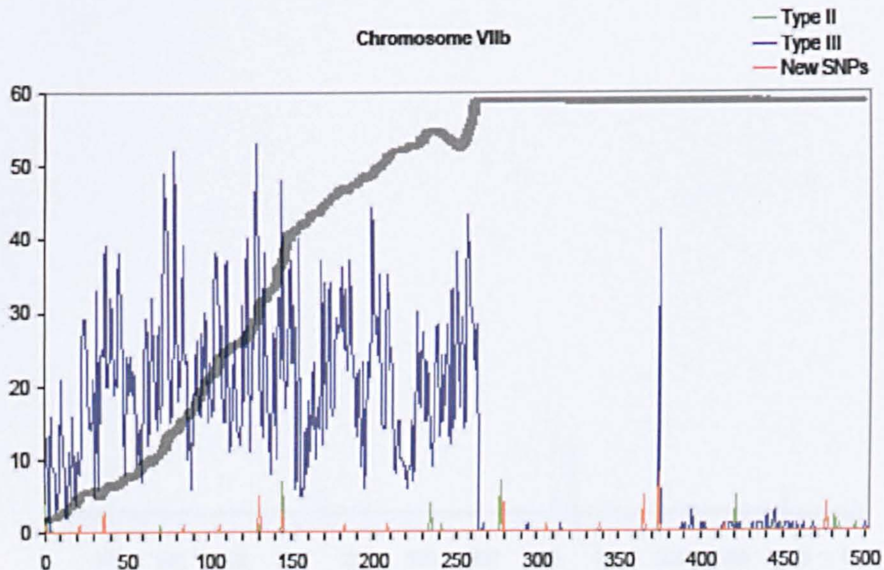

Chromosome III
Type II
Type III
New SNPs
160
140
120
100
80
60
40
20
0
0
50
100
150
200

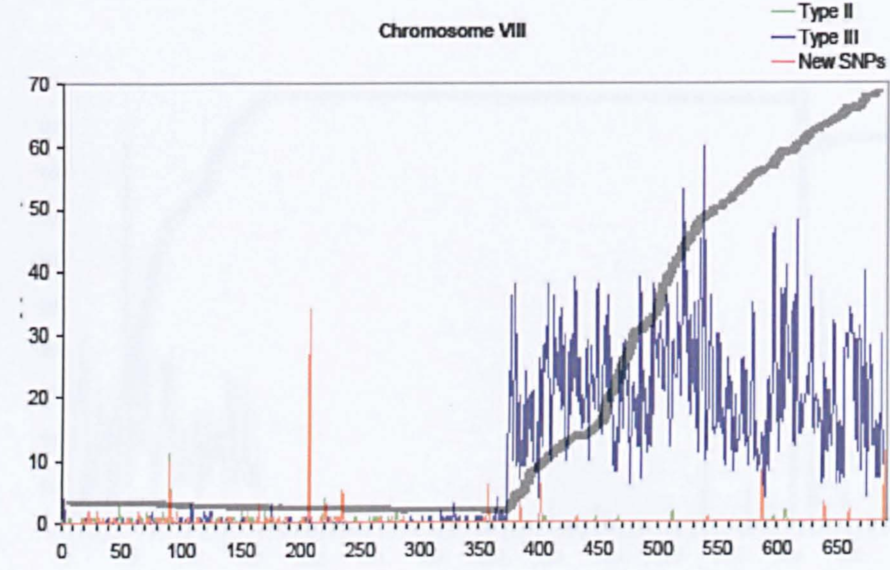

Chromosome IV
Type II
Type III
New SNPs
120
100
80
60
40
20
0
0
50
100
150
200
250

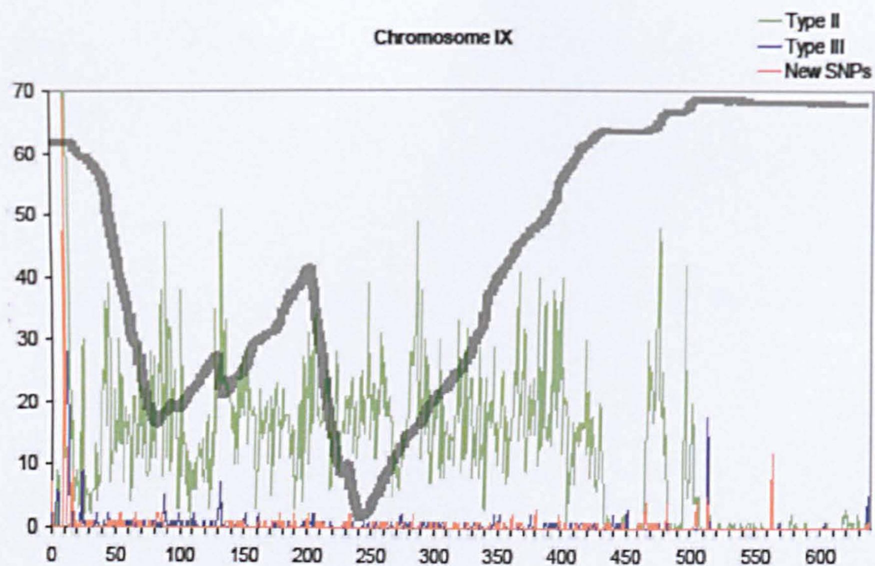

Chromosome V
Type II
Type III
New SNPs
70
60
50
40
30
20
10
0
0
50
100
150
200
250
300

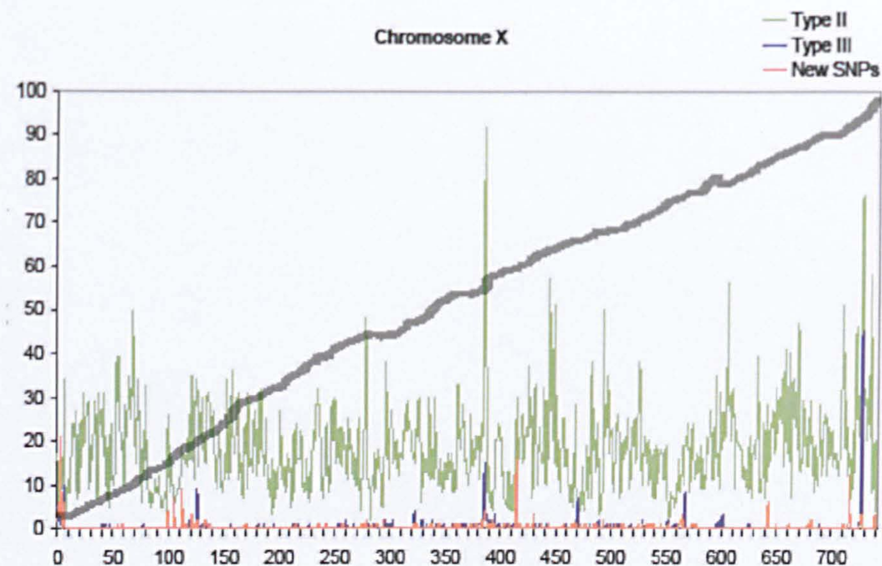

Chromosome VI
Type II
Type III
New SNPs
45
40
35
30
25
20
15
10
5
0
0
50
100
150
200
250
300
350

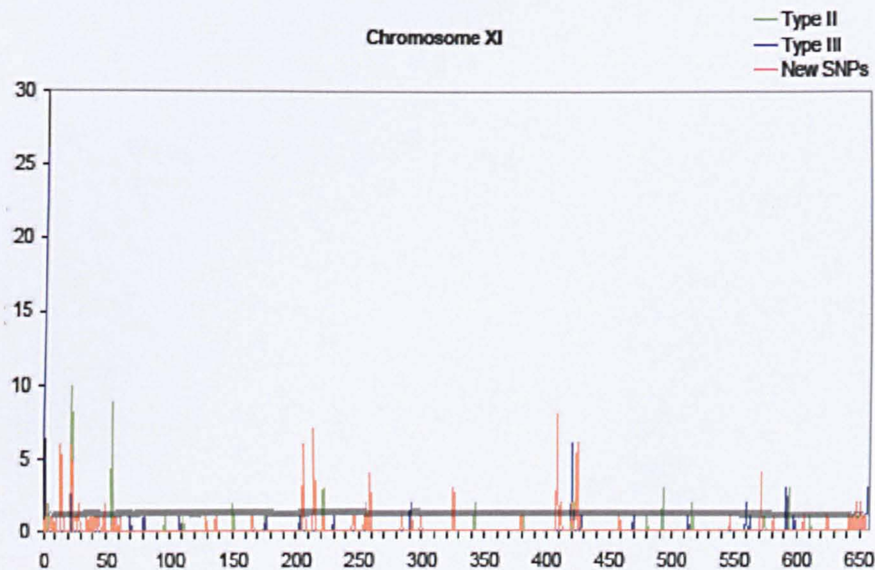

Chromosome VIIa
Type II
Type III
New SNPs
120
100
80
60
40
20
0
0
50
100
150
200
250
300
350
400

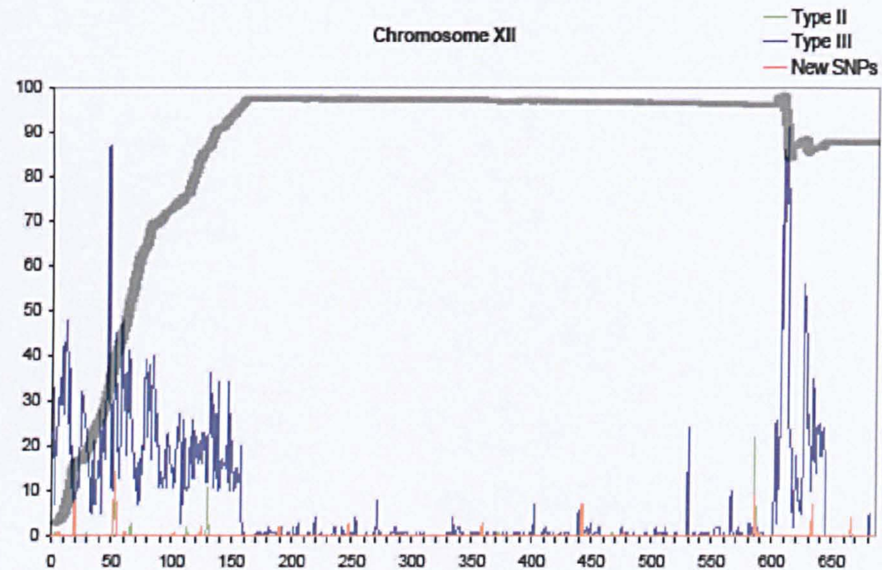

Chromosome VIIb
Type II
Type III
New SNPs
60
50
40
30
20
10
0
0
50
100
150
200
250
300
350
400
450
500


Chromosome VIII
Type II
Type III
New SNPs
70
60
50
40
30
20
10
0
0
50
100
150
200
250
300
350
400
450
500
550
600
650


Chromosome IX
Type II
Type III
New SNPs
70
60
50
40
30
20
10
0
0 50 100 150 200 250 300 350 400 450 500 550 600


Chromosome X
Type II
Type III
New SNPs
100
90
80
70
60
50
40
30
20
10
0
0 50 100 150 200 250 300 350 400 450 500 550 600 650 700


Chromosome XI
Type II
Type III
New SNPs
30
25
20
15
10
5
0
0 50 100 150 200 250 300 350 400 450 500 550 600 650


Chromosome XII
Type II
Type III
New SNPs
100
90
80
70
60
50
40
30
20
10
0
0 50 100 150 200 250 300 350 400 450 500 550 600 650