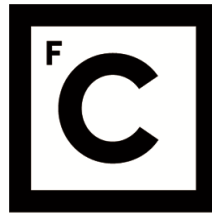UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

Genetic study of *Toxoplasma gondii* **strains isolated from humans and animals**

"Documento Definitivo"

Doutoramento em Biologia
Especialidade de Microbiologia

Anabela Barreiro Gomes Vilares Bendada Esteves

Tese orientada por:

Doutor João Paulo dos Santos Gomes
Professor Doutor Rogério Paulo de Andrade Tenreiro

Documento especialmente elaborado para a obtenção do grau de doutor

2019

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

Genetic study of *Toxoplasma gondii* strains isolated from humans and animals

Doutoramento em Biologia

Especialidade de Microbiologia

Anabela Barreiro Gomes Vilares Bendada Esteves

Tese orientada por:

Doutor João Paulo dos Santos Gomes

Professor Doutor Rogério Paulo de Andrade Tenreiro

Júri:

Presidente:

- Professor Doutor Rui Manuel dos Santos Malhó, Professor Catedrático, Faculdade de Ciências da Universidade de Lisboa.

Vogais:

- Doutor Agostinho Luís da Silva Cruz, Professor Coordenador com Agregação, Escola Superior de Saúde do Instituto Politécnico do Porto;
- Doutor João Paulo dos Santos Gomes, Investigador Auxiliar com Habilitação, Instituto Nacional de Saúde Doutor Ricardo Jorge (orientador);
- Doutora Mónica Alexandra de Sousa Oleastro, Investigadora Auxiliar, Instituto Nacional de Saúde Doutor Ricardo Jorge;
- Doutor Luís Manuel Madeira de Carvalho, Professor Associado com Agregação, Faculdade de Medicina Veterinária da Universidade de Lisboa;
- Doutor Ricardo Pedro Moreira Dias, Investigador Auxiliar Convidado, Faculdade de Ciências da Universidade de Lisboa.

Documento especialmente elaborado para a obtenção do grau de doutor

2019

E aprendes que realmente podes suportar...

que realmente és forte, e que podes ir muito

mais longe depois de pensar que não se pode mais.

William Shakespeare

AGRADECIMENTOS

À Maria João Gargaté (INSA), obrigada por seres como és, que sem ti este trabalho não teria sido possível, de todo, obrigada por teres sempre acreditado que eu era capaz, pelo apoio incondicional, pela amizade e por todo o teu esforço e empenho no meu futuro, nunca terei palavras ou ações que te possam compensar.

Às minhas companheiras de laboratório (INSA) (Assunção António, Cláudia Júlio, Idalina Ferreira, Susana Martins e Tânia Reis) que fizeram parte de todo este processo, que me ajudaram, trabalharam, encorajaram, me ouviram nos dias mais terríveis, foram indiscutivelmente umas queridas, um muito obrigado a todas para sempre. No entanto, tenho que agradecer em especial à Idalina Ferreira por tudo o que me ensinaste, à Cláudia Júlio por teres sido uma das mentoras deste processo e me encorajares, à Susana Martins por seres a boa disposição deste laboratório e por fim, pelo companheirismo, não apenas neste processo, mas também nos longos 18 anos de convivência.

Ao meu orientador Doutor João Paulo Gomes (INSA), pela oportunidade que me deu, pelo conhecimento, experiência, paciência e disponibilidade com que sempre me presenteaste, foi sem dúvida uma agradável surpresa.

Doutor Vítor Borges (INSA) por seres uma pessoa tão especial, sempre com um sorriso, amabilidade e saber único. Obrigado por todo o conhecimento que me transmitiste neste último trabalho (capítulo VI), sem ti teria sido tudo muito, mas mesmo muito mais difícil.

Ao meu orientador Professor Doutor Rogério Paulo de Andrade Tenreiro (BioISI) que permitiu que este trabalho fosse possível, que sempre se disponibilizou para me ouvir e ajudar e por todo o seu conhecimento e simpatia.

A todos os colegas e chefias do Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA),

e pela amizade com que sempre me presenteou. Lembrei-me várias vezes das suas sábias palavras.

Ao Dr Correia Neves, Dr. Jorge Canena, Dra Margarida Canas, Dr Manuel Liberato e Dr António Marques um obrigada muito especial nestes tempos difíceis, que sem o vosso saber e disponibilidade eu não me encontraria neste mundo, quanto mais a defender este trabalho.

À minha família, a ti mãe, por seres a minha referência, o meu ídolo, o meu herói, por acreditares que eu venceria desde os tempos de gestação, LUTADORAS!!!! A ti pai pelo teu humor, por me passares a tua perseverança e princípios únicos de vida que fazem de mim o que sou hoje. A vocês, mana e manos por me ajudarem a crescer desta forma tão maravilhosa. The last but not the least, a ti Tiago, Dinis e Duarte por me fazerem rir, chorar, encorajar e amar incondicionalmente em todo este processo, que se revelou um caminho árduo, com muitos obstáculos para TODOS.

x

ABSTRACT:

Toxoplasma gondii is responsible for toxoplasmosis, a zoonosis spread worldwide. This PhD thesis aimed to provide a more comprehensive overview of the genetic characteristics and virulence of *T. gondii* strains isolated from humans and animals in Portugal. We enrolled *T. gondii* strains collected from intermediary (farm, urban and sylvatic animals) and the definitive (cat) hosts, as well as from human toxoplasmosis. In general, we applied *Sag*2 and microsatellites classical genotyping, next-generation sequencing and inoculation in mice.

As major results, we highlight: *i)* the majority of isolated strains were avirulent in mice bioassay, where virulent strains belonged almost exclusively to *Sag*2 type I; *ii)* type II strains were the most identified type in strains from human toxoplasmosis as well as from all animal species, followed by type I strains; *iii)* this is the first report that shows type II *T. gondii* strains isolated from cattle, boars, foxes and hares in north of Portugal and recombinant strains isolated both from infected cats and pigeons; *iv)* the development of a rapid multi-loci-based NGS scheme, which allowed evaluating multi loci polymorphism and genomic mosaicism simultaneously. This schema is now implemented as the surveillance tool in the National Reference Laboratory of Parasitic and Fungal Infections (URSZ-INSA) to perform molecular surveillance of *T. gondii*; *v) w*ith this extended typing scheme, we identified a high rate (63 %) of recombinant strains isolated from humans, which had been previously identified as type I and II, as well as new putative genetic markers of virulence.

Overall, this PhD shed some light on the genetic and virulence diversity of *T. gondii* strains circulating in Portugal and revealed a surprising scenario of rampant genomic exchange in this parasite. These data highlight the need for further extended genomic studies to better understand the genotype/phenotype associations in this important human pathogen and thus contribute to better infection control measures.

Key Words: Genotypes, Portugal, *Toxoplasma gondii*, Loci, Microsatellites

RESUMO:

Toxoplasma gondii é o parasita responsável pela toxoplasmose e pode infetar a grande maioria dos animais vertebrados, incluindo o Homem. Este parasita tem uma distribuição mundial e apresenta um elevado interesse médico e veterinário. A infeção no Homem está maioritariamente associada ao consumo de carnes pouco cozinhadas ou cruas e à contaminação ambiental, sendo esta infeção na sua maioria assintomática. No entanto, pode ser grave ou mesmo levar à morte quando associada ao hospedeiro imunodeprimido ou à infeção primária materna por transmissão fetal, podendo ainda causar aborto ou doenças congénitas graves (ex. corioretinite, cegueira, linfoadenopatia, etc). A progressão e a gravidade da infeção têm sido associadas a diferentes fatores, incluindo características genéticas do parasita. Diferentes estudos têm sido desenvolvidos na genotipagem de *T. gondii*, os quais têm revelado que a maioria das estirpes isoladas na Europa e nos Estados Unidos da América pertencem ao tipo II e raramente se verificam isolados do tipo I, recombinantes ou atípicos. As estirpes recombinantes e atípicas apresentam normalmente uma maior gravidade na infeção, podendo levar à morte do hospedeiro imucomprometido ou mesmo do imunocompetente. No entanto, em Portugal, o conhecimento da distribuição, do tipo e da virulência dos genótipos circulantes é extremamente diminuto quer em humanos quer em animais. Para colmatar esta falta de conhecimento no nosso país, esta tese de doutoramento teve como objetivo geral a caracterização genética de estirpes de *T. gondii* de origem humana e animal, isoladas em Portugal. Neste sentido, delinearam-se diferentes estudos com objetivos específicos, correspondendo a diferentes capítulos desta tese e estando alguns deles publicados em jornais científicos.

No primeiro estudo (capítulo III) pretendeu-se caracterizar geneticamente estirpes de *T. gondii* isoladas de 209 produtos animais provenientes de dois grupos diferentes (silvestres e domésticos) da região Norte de Portugal (Grupo I – 12 bovinos, 34 ovinos, 5 caprinos e 16 porcos, Grupo II – 6 cães errantes, 59 raposas, 49 javalis e 28 lebres). Destes animais, serologicamente positivos para IgG anti *T. gondii*, foram colhidas diferentes amostras, nomeadamente coração, cérebro, diafragma e língua. Em todas as amostras foi realizada a técnica de PCR clássica do gene B1 onde 89 (43 %) animais revelaram resultados positivos. No que se refere à genotipagem de *T. gondii* realizaram-se PCRs das extremidades 5' e 3' do gene *Sag*2 onde 29 % (60/209) dos animais revelaram resultados positivos. Destas 60 estirpes, 77 % (46/60) pertenciam ao tipo II, 10 % (6/60) ao tipo I e em 13 % (8/60) das amostras não foi possível diferenciar o tipo I do tipo II. Neste estudo foi demonstrado pela primeira vez no

norte de Portugal a existência de estirpes do tipo II em vacas, javalis, raposas e lebres e foi também confirmada a predominância de estirpes do tipo II em porcos nesta mesma região. Foi também a primeira vez em que se procedeu à caracterização genética de *T. gondii* em cabras, javalis, raposas e lebres em Portugal. Demonstrou-se assim um maior leque de possíveis fontes de infeção para o Homem e para outros animais, bem como a importância da contaminação ambiental do ciclo silvático e sinantrópico de *T. gondii*.

Estes resultados vieram demonstrar a necessidade de extender os estudos de caracterização genética a outros hospedeiros, nomeadamente ao hospedeiro definitivo (gato), e a outras regiões do país com maior densidade populacional (capítulo IV). No segundo estudo (capítulo IV) pretendeu-se proceder à caracterização genética e fenotípica de estirpes de *T. gondii* isoladas de 41 cérebros de pombos e de 164 cérebros de gatos com anticorpos IgG anti-*T. gondii* provenientes da região de Lisboa. Os gatos e os pombos representam excelentes modelos de contaminação ambiental, e o gato é o hospedeiro definitivo de *T. gondii* sendo o único animal com capacidade de excretar oocistos (forma de resistência) nas zonas urbanas, logo mais populosas. Os cérebros de gatos e pombos foram colhidos e inoculados em ratinhos de laboratório de forma a estudar a virulência das estirpes, onde três isolados apresentaram fenótipo de virulência, sendo um proveniente de gato e dois de pombo. De forma a determinar o genótipo das estirpes de *T. gondii* realizaram-se duas técnicas distintas, nomeadamente a PCR clássica das extremidades 5' e 3' do gene *Sag*2 e a análise por electroforese capilar dos produtos da PCR multiplex de 5 microssatélites (B17, B18, TgM-A, W35 e *Tub*2). Com base no gene *Sag*2, a genotipagem foi alcançada em 70.7 % (29/41) dos isolados de pombos e em 50 % (82/164) dos isolados de gatos. Dos 29 isolados de pombos, 26 foram identificados como tipo II, dois como tipo III e um como tipo I, enquanto que dos 82 isolados de gatos, 72 foram identificados como tipo II e em 10 não se conseguiu fazer a distinção entre tipo I e tipo II. Com o intuito de aumentar o conhecimento na diferenciação de *T. gondii,* realizou-se a análise por electroforese capilar dos produtos da PCR multiplex de 5 microssatélites onde se identificaram duas estirpes (uma proveniente de gato e uma de pombo) recombinantes que tinham sido anteriormente identificadas como tipo II pela técnica de PCR clássica das extremidades do gene *Sag*2. Neste estudo, foi possível, pela primeira vez, identificar os genótipos de *T. gondii* circulantes no hospedeiro definitivo em Portugal e também a existência de estirpes recombinantes a circular neste país. Estes resultados vêm reforçar a importância destes animais no ciclo sinantrópico bem como demonstram a

importância da contaminação ambiental, neste caso na região de Lisboa, zona mais populosa de Portugal.

Ao estudar estes modelos de contaminação ambiental e possíveis fontes de infeção, quer animal quer humana, foram-se estudar também geneticamente e fenotipicamente (análise de virulência no ratinho) estirpes circulantes em Portugal isoladas de produtos humanos (capítulo V). O estudo da virulência e da diversidade genética de estirpes de *T. gondii* isoladas de produtos humanos nunca tinha sido realizado em Portugal anteriormente. Para alcançar este objetivo realizaram-se as mesmas técnicas utilizadas no capítulo anterior. Analisaram-se 48 estirpes (provenientes de casos de toxoplasmose congénita e adquirida) colhidas durante as duas últimas décadas, pertencentes à coleção do Laboratório Nacional de Referência de Infecções Parasitárias e Fúngicas do Departamento de Doenças Infecciosas do Instituto Nacional de Saúde Doutor Ricardo Jorge (URSZ-INSA). Sete das 48 estirpes (14.6 %) analisadas foram virulentas no ratinho de laboratório, em que três foram identificadas como tipo I, três recombinantes e uma do tipo II. A genotipagem pela PCR clássica das extremidades 5' e 3' do gene *Sag*2 não permite a identificação de estirpes recombinantes, no entanto permitiu a identificação de 35 (73 %) estirpes do tipo II e 13 (27 %) estirpes do tipo I. Contudo, com a análise por electroforese capilar dos produtos de PCR multiplex dos 5 microssatélites foi possível a identificação de 10 estirpes (21 %) recombinantes, anteriormente identificadas como type II ou como tipo I pela sequenciação das extremidades do gene *Sag*2. Este estudo constitui a primeira evidência de estirpes recombinantes a circular em Portugal provenientes da infeção congénita em humanos.

A identificação destas estirpes recombinantes provenientes da infeção humana gerou o interesse em efetuar a caracterização genética mais profunda das estirpes de *T. gondii,* a qual constituiu o objetivo do capítulo VI. Desenhou-se assim uma estratégia de caracterização genética, baseada em sequenciação de nova geração (NGS), tendo como alvos vários loci, os quais estão distribuídos pelos vários cromossomas e pensa-se estarem associados a funções de adaptação e virulência do parasita. A opção por esta abordagem *multi-loci* prende-se com o facto da metodologia de sequenciação total do genoma ter ainda um custo e dificuldade técnica muito elevados para *T. gondii,* uma vez que o genoma deste parasita é constituído por 14 cromossomas com um total de cerca de 65 Mb. A estratégia implementada permitiu, não só descriminar geneticamente a maioria das 68 estirpes estudadas [52 estirpes isoladas de produtos humanos em Portugal (48 isolados estudados no capítulo V, mais três novos isolados e uma estirpe de referência RH mantida em laboratório por vários anos no URSZ-INSA e

16 estirpes de referência mundial (GT1, ME49, VEG, ARI, CAST, COUG, CtCo5, FOU, GAB2-2007-GAL-DOM2, MAS, P89, RUB, TgCATBr5, TgCATBr9, TgCatPRC2 e VAND)], como também revelou um marcado cenário de mosaicismo genómico. De facto, cerca de dois terços (n = 43) das estirpes foram classificadas como recombinantes, sendo que esta recombinação parece estar ligada diretamente às linhagens das tipologias I, II e III ("archetypes"). As estirpes anteriormente classificadas pela sequenciação das extremidades do gene *Sag*2 como pertencendo ao tipo I revelaram um mosaicismo genómico superior (92 %) em relação aos outros tipos, nomeadamente às estirpes do tipo II, das quais 45 % revelaram recombinação. Contrariamente ao observado para o tipo I, a maioria das estirpes do tipo II, independentemente da recombinação existente, não apresentou virulência no ratinho de laboratório. De realçar ainda que, no âmbito da virulência, para além do *Sag*2, novos *loci* estudados (nomeadamente os alelos tipo I e tipo I-like dos loci *Sag*1, B17, PK1 e *Sag*3 e os alelos tipo III e III-like do TgM-A) revelaram-se candidatos muito promissores para uma rápida identificação de virulência no ratinho de laboratório. Esta abordagem baseada em NGS, a qual foi, no decurso deste estudo, já implementada no URSZ-INSA para efeitos de vigilância, permitiu também identificar variações no tamanho dos microssatélites, o que diminui drasticamente o trabalho do anterior método por deteção através de PCR clássico e electroforese capilar.

Em conclusão, o conjunto dos resultados desta tese contribui para um maior e melhor conhecimento da dinâmica da população de *T. gondii* que circula em Portugal, nomeadamente em termos da sua diversidade genética e identificação de recombinação entre estirpes. Em última análise, esta tese de doutoramento, em particular a metodologia de NGS recentemente implementada, a qual revelou um elevado poder discriminatório e de caracterização genética, poderá ser utilizada como uma ferramenta útil para estudos de epidemiologia molecular, os quais são fundamentais para o controlo desta importante infeção em saúde pública.

Palavras chave: Genótipo, Portugal, *Toxoplasma gondii*, *Loci*, Microssatélites

TABLE OF CONTENTS:

LIST OF TABLES:

LIST OF FIGURES:

LIST OF ABBREVIATIONS:

3R – Reduce, Reuse, Recycle

ABI – Applied Biosystems

AIDS – Acquired Immunodeficiency Syndrome

AF – amniotic Fluid

BB – Baby Blood

bp – Base pair

CB – Cerebral Biopsy

CI – Confidence Intervals

cM – Centimorgan

cm - centimetres

CML – Câmara Municipal de Lisboa

DMEM - Minimal Essential Medium

DNA – Desoxyribonucleic acid

ELISA – Enzyme-Linked Immunosorbent Assay

ENA – European Nucleotide Archive

Gra – Dense granule protein

Hsd: ICR (CD-1®) – Harlan Sprague Dawley, Inc., obtained a breeding stock from Charles River Breeding Laboratories.

HIV – Human Immunodeficiency Virus

HG – Haplogroup

Hp – Haplotype

ID – Index Diversity

IgG – Immunoglobulins G

IntraRec – Intra Recombinant

LD – Lethal Dosis

MAT – Modified Agglutination Test

Mb – Mega base pair

ml – millilitres

MLST – Multilocus Sequence Typing

MST – Minimum Spanning Tree

NGS – New Generation Sequencing

NIH – National Institute of Health

PBS – Phosphate Buffered Saline

PCR – Polymerase Chain Reaction

p.i. – post inoculation

PL – Placenta

PT – Portuguese

RAPD - Random Amplified Polymorphic DNA

RFLP – Restriction Fragment Length Polymorphisms

RNA – Ribonucleic Acid

Rop – Roptry

Sag – Surface Antigen

SNP – Single Nucleotide Polymorphism

TgMA – Myosine A

Tub2 – Beta Tubulina

UCB – Umbilical Cord Blood

UI – International Units

UK – United Kingdom

URSZ-INSA – Laboratório Nacional de Referência de Infecções Parasitárias e Fúngicas do Instituto Nacional de Saúde Doutor Ricardo Jorge

USA – United States of America

UTAD – Universidade de Trás os Montes e Alto Douro

WGS – Whole Genome Sequencing

µg – micrograms

NOTES OF THE AUTHOR: THESIS ORGANIZATION

This PhD thesis is organized in three parts. Its contents are essentially based on four manuscripts (listed below) that are presented as individual chapters (III to VI) in PART II, where three of them have already been published (the last one was submitted for publication at the time this thesis was submitted) in peer reviewed international journals.

The chapter arrangement does not perfectly reflect the chronological order of the manuscripts' publication, not only because some studies were developed simultaneously, but also because their publication time depended on the review process and journal requirements. In addition, despite of the main goals defined in the beginning of this PhD project, specific goals were raised on the course of the results obtained throughout the entire project, which influenced the publication priorities. As such, for the sake of clarity, the order of the chapters was decided to provide a better understanding of the approaches that were followed to achieve the proposed goals instead of reflecting the chronological order of their release to the scientific community. The three parts are the following:

PART I - Presents the literature overview **(CHAPTER I)** and the major objectives of this work **(CHAPTER II)**. It aims to provide the reader with an overview of the state of art underlying the subject of this thesis. **CHAPTER I** addresses a general introduction about *Toxoplasma gondii*, clinical features, biology, life cycle of the parasite, history, general geographical mapping, virulence, an approach of some widely used molecular methods to detect and differentiate *T. gondii*, and the epidemiology that concerns serology and genotyping of *T. gondii* in Portugal. **CHAPTER II** comprises the questions that lead to the objectives of this work.

PART II – This part describes the experimental approaches that were undertaken to achieve the proposed objectives. It contains several chapters, where most of them correspond to data that were already published in *peer-reviewed* scientific journals. **CHAPTER III –** Genotyping characterization of *Toxoplasma gondii* in animals from the North of Portugal. Contains a published study that was enriched with additional information. Lopes A.P., **Vilares A.**, Neto F., Rodrigues A., Martins T., Ferreira I., Gargaté M.J., Rodrigues M., Cardoso L. 2015. Genotyping characterization of *Toxoplasma gondii* in Cattle, Sheep, Goats and Swine from the North of Portugal. Iranian Journal of Parasitology. 10(3):465-72. **CHAPTER IV –** Contains a published study: **Vilares A.**, Gargaté M.J., Ferreira I., Martins S., Júlio C., Waap H., Ângelo H., Gomes J.P. 2014. Isolation and molecular characterization of *Toxoplasma gondii* isolated from pigeons and stray cats in Lisbon, Portugal. Veterinary

Parasitology. 205(3-4):506-11. doi: 10.1016/j.vetpar.2014.08.006. **CHAPTER V** – Contains the following published study: **Vilares A.**, Gargaté M.J., Ferreira I., Martins S., Gomes J.P. 2017. Molecular and virulence characterisation of *Toxoplasma gondii* strains isolated from humans in Portugal. Parasitology Research. 116(3):979-985. doi: 10.1007/s00436-017-5374-5. **CHAPTER VI** – Contains the following study that was recently submitted for publication: **Vilares A.**, Borges V., Sampaio D., Ferreira I., Martins S., Vieira L., Gargaté M.J., Gomes J.P. 2019. Towards a rapid sequencing-based molecular surveillance and mosaicism investigation of *Toxoplasma gondii*. In review at Parasitology Research.

PART III – This part is constituted by **CHAPTER VII** that provides a general discussion of the results obtained throughout the previous chapters and conclusions. Of note, in order to avoid redundancies, this section solely summarizes and discusses the major findings of this work because a detailed discussion of specific results was already provided in each chapter. It also presents the future perspectives enrolling specific research lines that may be addressed on the course of the obtained results. Considering the different layouts required by the different journals where the manuscripts were published, including tables, figures and references, all chapters were formatted in a unique style, with all references being cited by the name of the first author and year of the publication and listed in a single section - "References" - according alphabetical order.

PART I

CHAPTER I

INTRODUCTION

INTRODUCTION

1.1 General introduction

T. gondii belongs to the Apicomplexa phylum and it is the causal agent of toxoplasmosis. It is a wide spread protozoan parasite of humans and warm-blooded animals, including mammals and birds. In humans, this intracellular parasite is commonly acquired by the oral ingestion of tissue cysts containing bradyzoites, however it can also be transmitted by the ingestion of oocysts containing sporozoites that are the product of a sexual cycle shed by cats, the definitive host (Halonen and Weiss, 2013) and which can contaminate water and food. Classically, consumption of undercooked meat has been described to be the major risk factor for acquisition of toxoplasmosis (Tenter *et al.*, 2000). *T. gondii* is the only known species in the genus *Toxoplasma* and is considered one of the most successful eukaryotic pathogen in the world in terms of the number of host species and percentage of animals infected worldwide (Tenter *et al.*, 2000; Grigg and Suzuki, 2003; Su *et al.*, 2010). Serological studies suggest that chronic infection rates in humans can vary from less than 10 % to more than 70 % depending on geographic region and various risk factors (Joynson, 2003).

1.2 History

Toxoplasma gondii was simultaneously described in tissues of different animals by Nicolle and Manceaux, and Splendore (Table 1.1). The genus was named by Nicolle and Manceaux as *Toxoplasma* for its bow-like shape (from Greek: toxo = bow or arc; plasma = creature). Over the years, different authors described *T.* gondii-like organisms, however it is believed that these cases were likely associated with infection by *Leishmania* spp. (Weiss and Dubey, 2009).

Janku (1923) and Torres (1927) described similar cases in infants who died with convulsions at two days of age. Although Janku only referred to the observed organisms as Sporozoa and Torres as *Encephalitozoon chagasi*, it was Levaditi (1928) who suggested that both cases were due to *T. gondii*. Later, *T. gondii* still be mistakenly named as *E. brumpti* when identified in the spinal fluid of a 17 year old boy from Corsica who died of meningitis.

The first viable organism was isolated by Albert Sabin in 1937 from laboratory mice during a routine virus testing experiment (Sabin and Olitsky, 1937), and later from humans by inoculating infected tissues into mice (Wolf *et al.*, 1940). Sabin speculated that the parasite was an important zoonosis and demonstrated that this human strain was neither biologically

INTRODUCTION

nor immunologically different from isolates from other animals. It was since then that the parasite was progressively recognized as the agent of a widespread zoonosis (Halonen and Weiss, 2013).

In 1939 Sabin isolated and inoculated *T. gondii* RH (Table 1.1) in mice and on the thirteenth day after inoculation, one of the mice exhibited a distended abdomen and a large amount of peritoneal exudate in which *Toxoplasma* organisms (Sabin, 1938) . Reinoculation of the tissues from the mice in healthy mice showed that *Toxoplasma* produced a fatal toxoplasmosis. This laboratory strain, it has been passed in mice in many laboratories. After these prolonged passages its pathogenicity for mice has been stabilized (Dubey *et al.*, 1977) and it has lost the capacity to produce oocysts in cats (Frenkel *et al.*, 1976).

Table 1.1 History of the identification of *Toxoplasma gondii*.

Year	Authors	Host	Fact	Symptoms	References
1908*	Nicolle and Manceaux	*C. gundi*	First identification	-	(Weiss and Dubey, 2009)
1908*	Splendor	Rabbit	First identification	-	
1914	Castellani**	Child	Identification in smears of blood and spleen	Severe anaemia Fever Spleenomegaly	(Weiss and Dubey, 2009)
1916	Fedorovitch**	Child	Identification in blood	Severe anaemia Fever Spleenomegaly	(Weiss and Dubey, 2009)
1920	Chalmers and Kamar**	Soldier	Identification in blood	Chronic headache Fever Diarrhoea	(Weiss and Dubey, 2009)
1923	Janku	Child	Identification of humans as susceptible hosts	Died with convulsions	(Weiss and Dubey, 2009)
1937	Sabin	Mice	First viable organism isolated	-	(Sabin and Olitsky, 1937)
1939	Sabin	Child	Isolation of RH strain	Encephalitis	(Sabin, 1941)

T. gondii was identified for the first time simultaneously by these authors in different regions; **Authors could miss associated *Leishmania* spp to *T. gondii*.

INTRODUCTION

Other forms of *Toxoplasma* including tissue cysts were recognized to exist by several researchers in 1951, but it was not until the 1960s and 1970s that the parasite was identified as a coccidian and *Apicomplexa* phylum. In 1972, the cat was identified as the definitive host harbouring the sexual parasitic cycle and spreading oocysts through faeces (Jewell *et al.*, 1972).

The importance of maternal and congenital transmission has long been recognized since 1939; when a neonate from New York developed toxoplasmosis (Jones *et al.*, 2001), and the growing role of *Toxoplasma* infection in immunocompromised patients was acknowledged in the mid-1970s (Luft and Remington, 1992). A more complete appreciation of the symptoms associated with acute acquired toxoplasmosis was achieved with the report of outbreaks of acute toxoplasmosis in adults in the USA (Teutsch *et al.*, 1979) and Canada (Bowie *et al.*, 1997).

1.3 *Toxoplasma gondii* infection

1.3.1 Clinical features

In animal husbandry toxoplasmosis is the major cause of reproductive failure by leading to early embryonic and fetal death with abortions and resorption with fetal mummification (Dubey and Kirkbride, 1989; Dubey, 2009, 2010a). The severity of infection is associated with the stage of gestation at which the livestock becomes infected. The earlier infection in gestation, the more severe will be the consequences (Dubey, 2009). However, the majority of *T. gondii* infections in most host species are subclinical or asymptomatic, and chronic (Dubey, 2010a) and the parasite can remain dormant in the tissues until reactivation or until the host is eaten by a predator. Also, the associations of chronic infection with changes in behaviour, memory, and neurologic disorders (Berdoy *et al.*, 1995, 2000; Webster, 2001; Torrey and Yolken, 2007; Berenreiterová *et al.*, 2011; Dass *et al.*, 2011; House *et al.*, 2011; Pedersen *et al.*, 2011; Torrey *et al.*, 2012; Webster *et al.*, 2013) are fascinating, as illustrated for example, by the cases of the preys that lost the natural fear of predators (Haroon *et al.*, 2012). Consequently, the healthy aspect and the fear loss of preys could increase the perpetuation of *T. gondii* cycle in human's and in other hosts.

In humans, toxoplasmosis was implicated, about two decades ago, as the third most common cause of food borne infection in the USA (Mead *et al.*, 1999). It is recognized as a category B priority pathogen by the National Institutes of Health, Bethesda, USA due to its

INTRODUCTION

risk of transmission through contamination of food or water and importance as an opportunistic pathogen. Hence, studies performed in *T gondii* are often used as model to related but less tractable pathogens such as *Cryptosporidium*, another category B Biodefense Agent that causes severe diarrheal disease, and *Plasmodium*, the causative agent of malaria (Kim and Weiss, 2004).

Most human *T. gondii* infections are often asymptomatic, however, in several of its hosts it is associated with congenital infection and abortion (Gilbert *et al.*, 1999; McLeod *et al.*, 2006; Oz, 2014). Congenital toxoplasmosis can manifest with severe complications, such as miscarriage, fetal developmental retardation, encephalitis, neurological and mental illnesses, visual and auditory inflammatory disorders, cardiovascular abnormalities, and pains (Gilbert *et al.*, 1999; McLeod *et al.*, 2006; Oz, 2014). In addition, *T. gondii* can also cause encephalitis or systemic infections in the immunocompromised, particularly in individuals with HIV/AIDS. *Toxoplasma* infection was also implicated in etiologies of neurodevelopmental and neurocognitive disorders like-schizophrenia (Torrey and Yolken, 2007). Severe cases of toxoplasmosis have been reported in immunocompetent patients in association with atypical *T. gondii* genotypes (Ajzenberg *et al.*, 2004; Demar *et al.*, 2007; Elbez-Rubinstein *et al.*, 2009; Vaudaux *et al.*, 2010; Wendte *et al.*, 2011; Pomares *et al.*, 2011; Sobanski *et al.*, 2013). Although current available drugs, like pyrimethamine and sulphadiazine, can control the proliferative form of the parasite and treat *Toxoplasma* infections, they are poorly tolerated, they have severe side effects such as allergic reactions, and they are ineffective against chronic *Toxoplasma* infections. In addition, resistance to some of these drugs has been noted (Aspinall *et al.*, 2002b; Baatz *et al.*, 2006; Blader and Saeij, 2009). Pathogenesis is typically associated with the ability of these parasites to replicate and proliferate within host cells. However, pathogenicity can also vary with the host and the morphological stage ingested (Dubey, 2010a; Dubey *et al.*, 2012).

A study performed in the USA estimated that the annual cost of illnesses caused by *Toxoplasma* is about \$3 billion (Hoffmann *et al.*, 2012). A study in ewes (n=1613) from Uruguay estimated annual losses due to toxoplasmosis during gestation of 1.4 to 3.9 %, accounting to approximately \$ 1.4 to 4.7 million (Freyre *et al.*, 1997). The economic losses due to lamb mortality and missed lactation are estimated at 10 million Euros per year in Italy (Masala *et al.*, 2003). In UK, the impact of toxoplasmosis in sheep industry is between £ 12 million and £ 24 million each year (http://www.apd.rdg.ac.uk/AgEcon/livestockdisease/index.htm). However, clinical and detailed economic information is lacking in Europe and particularly in

INTRODUCTION

Portugal. Overall, toxoplasmosis results in increased production costs, diminished marketability of meat, less replacement animals, retardation of genetic progress and constitutes a major source of human infection (Leighty, 1990; Freyre *et al.*, 1997).

1.4 Biology and Life Cycle

T. gondii is a tissue cyst-forming coccidia with a heteroxenous life cycle in which an asexual reproduction in intermediate hosts is linked to a sexual reproduction in definitive hosts (Tenter and Johnson, 1997).

T. gondii has three morphological stages: tachyzoite (2 x 6 μm), bradyzoite (1-3 x 5-8.5 μm) and sporozoite (2 x 6-8 μm) (Sabin and Olitsky, 1937; Jones *et al.*, 1972; Cowper *et al.*, 2012). Bradyzoites and sporozoites are the infective forms typically acquired by ingestion. Bradyzoites are enclosed in tissue cysts in chronically infected hosts (Figure 1.1).
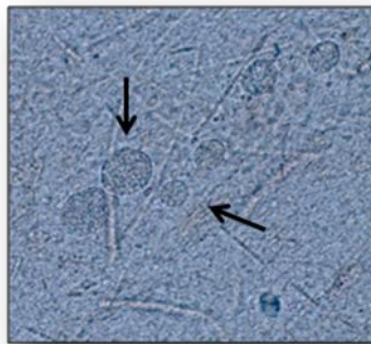
Figure 1.1. *T. gondii* tissue cysts with multiple bradyzoite enclosed in mouse brain cells.
Note: This figure belongs to the URSZ-INSA acquis.

Eight sporozoites are present in each completed sporulated oocyst (11 x 13 μm), millions of which are shed into the environment, unsporulated, by the definitive hosts (felids) of the parasite, over 4 days after the first infection. It is an obligatory intracellular parasite of the nucleated cells of its hosts. They possess a nucleus, a mitochondrion, a Golgi complex, ribosome, an endoplasmic reticulum, and an apicoplast (Roos, 1999). The oocyst is the only form of life of the parasite that can persist outside the host for a long period of time (Yilmaz and Hopkins, 1972; Frenkel and Dubey, 1973; Dubey, 1998c, 2010a).

After ingestion, tissue cysts or oocysts invade the host cells and differentiate into tachyzoites which divide rapidly within the host cells and together with the host immune response are responsible for the clinical manifestations of infection. Tachyzoites infect

INTRODUCTION

nucleated host cells and utilize monocytes, macrophages, and dendritic cells as "Trojan Horses" to escape the host immune defence (Elsheikha and Khan, 2010), to bypass the blood–brain barrier (Bierly *et al.*, 2008) and the placenta barricade, and to spread and cause systemic disease.

Tachyzoites differentiate into latent bradyzoites, which can be induced by exposure of the organism to stress conditions such as an immune response. Tissue cysts can persist indefinitely for the life of the host (Figure 1.2). If an individual becomes immunocompromised these tissue cysts serve as a reservoir from which disseminated or local infections can develop. Tissue cysts have a predilection for neural and muscle tissue as well as the eye in humans, with most cases of reactivation disease presenting as encephalitis or chorioretinitis (Kim and Weiss, 2008).
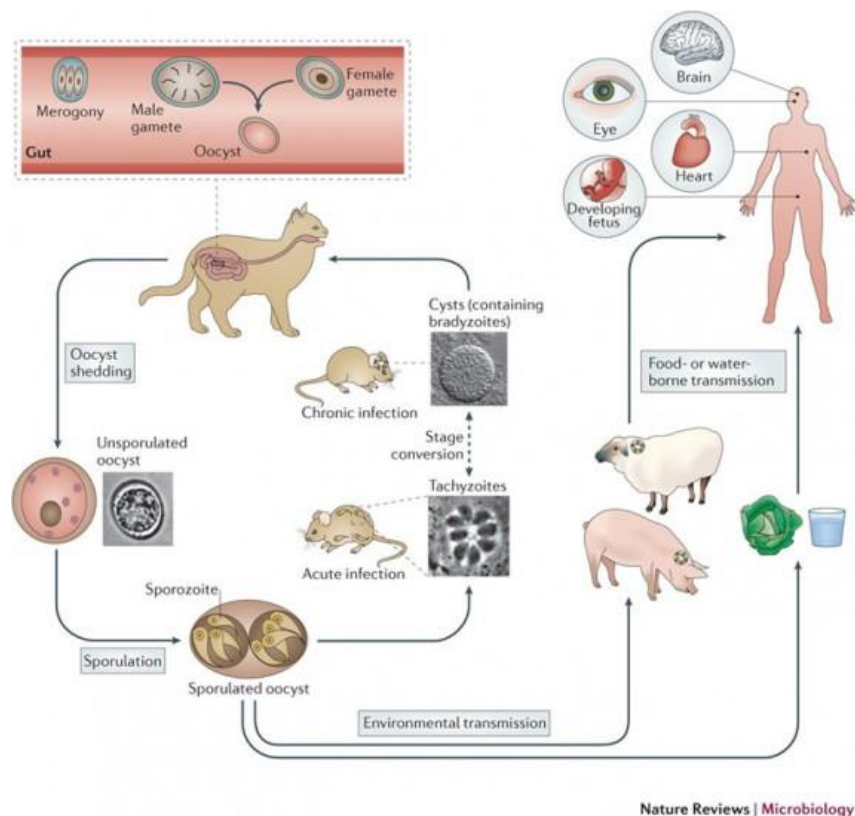
Figure 1.2. *Toxoplasma gondii* life cycle. **Note:** Adapted from Nature Reviews

1.5 Molecular epidemiology

1.5.1 Genotypes

The genome size of *T. gondii* is between 60 - 70 Mb. In 2001 only 11 chromosomes had been identified (Ajioka *et al.*, 2001) but now it is known that *T. gondii* genome consists of

INTRODUCTION

14 chromosomes (Khan *et al.*, 2005a). In the early days, isolates of *T. gondii* were grouped according to virulence in mice. Later on, the first phylogenetic studies of *T. gondii* strains indicated that their genetic complexity was much smaller than expected (Sibley and Boothroyd, 1992a; Dardé *et al.*, 1992) so, for a long time, *T. gondii* was considered to be clonal with three genetic types (Types I, II, III) (Dardé *et al.*, 1992; Howe and Sibley, 1995; Ajzenberg *et al.*, 2004) and very small differences between clonal lineages. In fact, comparative sequence analysis of individual genes indicated low allelic diversity within the clonal lineages (about 1 % divergence). In addition, limited genetic diversity between and within clonal lineages indicated that they have evolved quite recently from a common ancestor, 10,000 years ago at the most (Su, 2003). Also, recent advances in the knowledge of the virulence associated with some genotypes have been achieved (Saeij *et al.*, 2005) with the development of new genotyping tools and the multiplication of field studies (Mercier *et al.*, 2011). Also, recent studies on *T. gondii* in different animal populations worldwide started to reveal the extensive diversity of the parasite (Dubey *et al.*, 2002; Lehmann *et al.*, 2006; Pena *et al.*, 2008; Khan *et al.*, 2011a).

A fourth clonal lineage, designated haplotype 12, is largely confined to North America, where it is more common in wild animals (Khan *et al.*, 2011a). Other studies subdivided *T. gondii* in I, II, III and 4 to 14 haplotypes based on only five loci (Khan *et al.*, 2011b; a). Also, Su and colleagues classified *T. gondii* strains into 15 different haplotypes, defining six major clades (Figure 1.3) (Su *et al.*, 2012). However, Minot and colleagues showed that the observed extent of *Toxoplasma* genetic diversity does not fit the scenario of the proposed haplotypes. Instead, most strains appear to have been originated through recent recombination events and it is believed that some of them led to particularly fit genotypes that geographically spread or swept (Minot *et al.*, 2012). The majority of genotypes were identified either by polymerase chain reaction followed by restriction fragment length polymorphism (PCR-RFLP) or by microsatellite analysis using different genetic markers and variable number of loci. In America, 85.7 % of parasites were typed by using three to nine protein-coding genetic markers by PCR-RFLP, while the majority of *T. gondii* genotyping in Europe (67.4 %, n = 126) and Oceania (100 %, n = 2) was performed by microsatellite analysis. All variants from Asia and most from Africa (70.4 %) were identified with one coding-gene marker (Rico-Torres *et al.*, 2016).

INTRODUCTION

Currently, there is no gold-standard for genotype designation. As long as different methods and genes are used to type a variety of isolates, and each method has its own classification scheme, genotype designation remains a confusing issue.
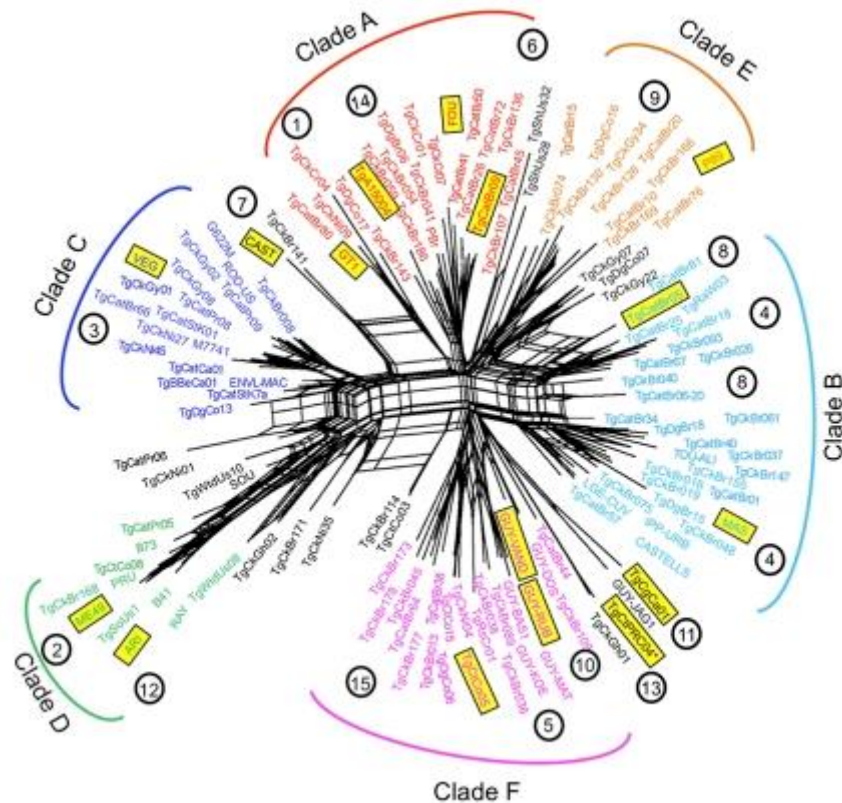
Figure 1.3. Population genetic structure of *T. gondii*. **Note:** adapted from Su et al 2012. Neighbor-net analysis was conducted using 11 multilocus RFLP markers plus 1 marker for the apicoplast and four intron sequences from the 138 representative strains representing unique haplotypes. Neighbor-net analysis showed various routes for gene flow between different populations (interconnecting lines between representative strains). Six major clades (A through F) are indicated on the basis of STRUCTURE analysis. Strains in black lettering do not correspond to major clades. Haplogroups are shown in circled numbers. Representative strains for each haplogroup are indicated by yellow boxes.

1.5.2   Genotyping Methods

Numerous studies of the *T. gondii* population structure were based on genotyping using a single marker, mostly *Sag*2 (Howe *et al.*, 1997; Fuentes *et al.*, 2001; Sabaj *et al.*, 2010) and particularly, due to its polymorphisms and sensitivity, *Gra*6 (Fazaeli *et al.*, 2000; Messaritakis *et al.*, 2008). However, genotyping with a single marker does not allow identification of nonclonal strains. To determine more precisely the presence of polymorphisms in the population, application of multilocus PCR and microsatellite analysis of multiple markers is necessary (Ajzenberg *et al.*, 2005; Su *et al.*, 2006). The paragraphs below describe the most used techniques in *T. gondii* genotyping.

INTRODUCTION

1.5.2.1 Multilocus Enzyme Electrophoresis (MLEE)

MLEE differentiates strains by assessment of differences in migration of metabolic enzymes by electrophoresis. These migration differences are determined by the sequences of amino acids (molecular weight and enzyme charge) and could be associated to allelic variation at the corresponding gene locus. When a group of enzymes are analyzed all together, the different migration produces a unique finger print of each strain representing a multilocus genotype (Dardé *et al.*, 1988, 1992).

The first study on strain diversity in *T. gondii* based on MLEE analysis was developed by Dardé et al. (Dardé *et al.*, 1988). Six polymorphic enzymatic systems exhibited 12 zymodemes with the majority of stocks clustering into three main zymodemes Z1, Z2 and Z3 (Dardé *et al.*, 1988, 1992; Ajzenberg *et al.*, 2002a). The disadvantages of MLEE are the low resolution in *T. gondii* genetic studies, the requirement of large numbers of parasites (approximately 7×10^{6} tachyzoites) and the purity of the samples (Dardé *et al.*, 1992).

1.5.2.2 RAPD-PCR

RAPD-PCR is based on the amplification of genomic DNA using arbitrary primers to identify unknown DNA polymorphisms. *T. gondii* have been classified into virulent and avirulent strains based on the murine virulence by RAPD-PCR using arbitrary primers (Guo *et al.*, 1997), however nowadays the study of virulence strains require a much more complex interpretation. The results of this technique are difficult to reproduce and it requires a high DNA purity (Guo and Johnson, 1995; Ferreira *et al.*, 2004).

1.5.2.3 PCR- RFLP

The PCR-RFLP is based on the ability of restriction endonucleases to recognize single nucleotide polymorphisms (SNPs), digest PCR products and then display distinct DNA banding patterns on agarose gels by electrophoresis (Sibley *et al.*, 1992; Howe and Sibley, 1995). How and Sibley identified for the first time *T. gondii* genotypes with PCR-RFLP using 6 markers. Since then, several different sets of multilocus PCR-RFLP markers have been developed to genotype *T. gondii* isolates (Howe and Sibley, 1995). As the conventional multilocus PCR-RFLP requires a large amount of DNA, multiplex nested PCR-RFLP was developed to overcome this problem. However, the possibility of contamination is high and to avoid this problem multiple negative controls must be used in all PCRs (Sibley *et al.*, 1992; Cristina *et al.*, 1995; Howe and Sibley, 1995; Khan *et al.*, 2005b; de Melo Ferreira *et al.*,

2006; Dubey *et al.*, 2007a; Pena *et al.*, 2008; Soares *et al.*, 2011; Su *et al.*, 2012; Bacci *et al.*, 2015).

1.5.2.4 Microsatellite analysis

Microsatellite sequences are tandem short DNA motif repeats that are widespread in eukaryotic genomes and the sequences usually change due to insertion or deletion of repeat units. The numbers of repeat units differ in a population, thus producing multiple alleles at a microsatellite locus. The tandem repeats in *T. gondii* are composed of 2 to 6 nucleotides and occur 2–20 times (Blackston *et al.*, 2001; Ajzenberg *et al.*, 2002a, 2010; Li *et al.*, 2014). This length polymorphism of microsatellite regions can be assessed with fluorescent primers after electrophoresis on an automatic sequencer. A total of 5 markers have been used (*Tub*2, W35, TgM-A, B18, B17) to differentiate types I, II, III, recombinants and atypical genotypes (Ajzenberg *et al.*, 2002b). More recently, Ajzenberg and colleagues developed a method for *T. gondii* genotyping in a single multiplex PCR assay using 15 microsatellite markers (*Tub*2, W35, TgM-A, B18, B17, M33, IV.1, XI.1, M48, M102, N60, N82, AA, N61, and N83) (Ajzenberg *et al.*, 2010). Although microsatellite multiplex has been considered an accurate assay to differentiate *T. gondii* strains and to identify mixed infections (Ajzenberg *et al.*, 2009; Mercier *et al.*, 2010) the presence of small amounts of DNA (usually from clinical samples or wild strains with low rate of taquizoites or cysts) may impair the results (Ajzenberg *et al.*, 2010).

1.5.2.5 Sequencing

MLST was first proposed in 1998 as a typing approach enabling the unambiguous characterization of bacterial isolates in a standardized, reproducible, and portable manner using the human pathogen *Neisseria meningitides* as the model organism (Maiden, 2006). The typical MLST for *T. gondii* is based on DNA sequence polymorphisms of seven housekeeping genes, including the single nucleotide polymorphisms (SNPs), deletion and insertion of nucleotides (Su *et al.*, 2012). However different virulence loci have been used by different authors and also different sequencing approaches, such as Sanger (classic) or NGS (Khan *et al.*, 2011a, 2014; Cheng *et al.*, 2015; Lorenzi *et al.*, 2016). Once more, although this approach is the most discriminatory method it is hardly applicable to clinical samples, as a large quantity of genomic DNA is required for the successful amplification of several loci.

INTRODUCTION

1.5.2.6 Serotyping

Serotyping is a typing method based on a serological test using strain-specific peptides. For *T. gondii,* different authors have used specific peptides derived from dense granule antigens, like GRA5, GRA6 and GRA7 (Sousa *et al.*, 2008; Maksimov *et al.*, 2013). Serotyping was shown to be capable of distinguishing type II from non-type II infections (Kong *et al.*, 2003; Xiao *et al.*, 2009) and may be used to determine which strains are associated with symptomatic or asymptomatic infections (Sousa *et al.*, 2008). As serotyping is fast, inexpensive, relatively noninvasive, and there is no need to isolate parasites, this technique could have the potential to become the method of choice for typing *T. gondii* in humans and animals. However, some limitations arise, such as cross-reactivity issues, the low sensitivity of the selected peptides in detecting recombinant strains (Sousa *et al.*, 2009a) and the potential no viability of the serum (dead animals). In particular, immunosuppressed patients may not produce sufficient specific antibodies to reach the detection threshold (titer of 1:64) and infection with the rare genotypes may induce highly different humoral responses that may not be detectable using the custom polymorphic polypeptides(Kong *et al.*, 2003).

1.5.3 Genotypes and virulence

In terms of mouse virulence, Type I isolates are considered as the most virulent, and can lead to death of mice less than 10 days after the first inoculation. In contrast, strains of Type II and Type III are avirulent (although type II isolates can cause low virulence in mice) and usually cause chronic infection and shaped tissue-cysts. Cystogenic *T. gondii* strains can be grouped according to their virulence, however, all non-cystogenic isolates present high virulence in mice (de Melo Ferreira *et al.*, 2006).

In Humans, type I and recombinant strains isolated from immunocompetent individuals have been associated with some severe forms of toxoplasmosis or atypical ocular toxoplasmosis (Howe *et al.*, 1997; Fuentes *et al.*, 2001; Grigg *et al.*, 2001a). Also, in southern Brazil and French Guiana, atypical strains have been associated with severe disseminated disease (Dardé *et al.*, 1998; Blaizot *et al.*, 2019) and with debilitating ocular toxoplasmosis and death in healthy adults (Carme *et al.*, 2002; Jones *et al.*, 2006a; b; Demar *et al.*, 2012; Silveira *et al.*, 2015).

INTRODUCTION

1.5.4 Geographical distribution of *T. gondii* genotypes

Several studies have been conducted in diverse countries aiming at analyzing *T. gondii* genotypes of isolates circulating in diverse hosts and countries. Most of them performed a multilocus genotyping by PCR and RFLP markers (Sundar *et al.*, 2008; Sun *et al.*, 2013; Wang *et al.*, 2015), and less frequently by microsatellite markers (Ajzenberg *et al.*, 2010), gene sequencing or NGS (Lorenzi *et al.*, 2016). Studies are most often carried out in birds, as sentinels of the environmental contamination (Lehmann *et al.*, 2006), whereas genotypes infecting wild species or other domestic animals are less frequently studied. Although epidemiological studies presenting genotypes infecting humans may be biased due to travel or to consumption of imported food, they usually reflect circulating genotypes (Ajzenberg *et al.*, 2015).

In Europe and North America, *T. gondii* exists as four distinct clonal lineages that show marked virulence differences in laboratory mice (Howe and Sibley, 1995; Khan *et al.*, 2011a). It is accepted that type II is the predominant type in cases of congenital toxoplasmosis in Europe and North America (Howe and Sibley, 1995; Howe *et al.*, 1997; Ajzenberg *et al.*, 2002b). A fourth clonal lineage, referred to as type 12, has recently been described in North America where it is commonly found in wildlife (Khan *et al.*, 2011a). All four clonal lineages show evidence of overly abundant, highly similar multilocus genotypes and high levels of linkage disequilibrium (i.e., infrequent recombination). While it has been established that type II is predominant in Europe and North America (Dardé *et al.*, 1992; Howe and Sibley, 1995; Howe *et al.*, 1997), there are significant regional differences. Type I and some recombinant strains were isolated from immunocompetent individuals suffering from severe or atypical ocular toxoplasmosis in United States (Grigg *et al.*, 2001a).

One French study has shown that of the 86 isolates (85 %) from cases of suspected and confirmed congenital toxoplasmosis and 88 isolates from immunosuppressed patients (61.4 %) were type II (Ajzenberg *et al.*, 2002b, 2009). In other studies with immunosuppressed patients, also from France, it was shown that type II isolates were also predominant, while types I and III were rarely isolated (Howe *et al.*, 1997; Honoré *et al.*, 2000). Research in congenital toxoplasmosis from Spain and Egypt showed the presence of type I (Fuentes *et al.*, 2001; Eldeek *et al.*, 2017), while genotyping of isolates from Crete, Cyprus showed the predominance of type III (Messaritakis *et al.*, 2008), however it must be noted that these studies have been conducted using only one or four markers.

INTRODUCTION

Furthermore, strains of atypical genotypes were isolated from immunocompetent patients in French Guiana and Brazil (Carme *et al.*, 2002; Demar *et al.*, 2012; Silveira *et al.*, 2015).

1.6 *Toxoplasma gondii* in Portugal

1.6.1 *T. gondii* epidemiology in Portugal

In Portugal, there is a lack of human data regarding *T. gondii* as few studies are available, and most of them are serological estimates of *T. gondii* seroprevalence in humans (Table 1.2) and in several animal hosts (Table 1.3).

There were a limited serologic studies performed in humans, in Portugal (Ângelo, 1983, 2003; Antunes, 1984; Sevivas, 2011; Lopes *et al.*, 2012b; Gargaté *et al.*, 2016), however, only two were performed in the general population. The first serological survey reported 47 % of seropositivity, which is considered a high value when compared with the one obtained in 2013 (Ângelo, 1983; Gargaté *et al.*, 2016), where the overall seroprevalence was 22 % (Gargaté *et al.*, 2016). Of note, despite this apparent strong decreasing trend of the *T. gondii* seroprevalence in Portugal during the last decades, the risk of congenital toxoplasmosis is higher, since these data reveal that the majority of potential pregnant women are thus susceptible to primary infection (Table 1.2).

Table 1.2 Seroprevalence of *T. gondii* infection in humans from Portugal

Years	Region	Population	N tested	Prevalence (%)	Reference
1981	L	Pregnant women	686	64	(Antunes *et al.*, 1981)
1979–1980	all Portugal	General population	1675	47	(Ângelo, 1983)
2000	L	Imunosupressed VIH	63	57	(Ângelo, 2000)
2009–2010	N	Women of childbearing age	401	24	(Lopes *et al.*, 2012b)
2013	all Portugal	General population	1440	22	(Gargaté *et al.*, 2016)

L: Lisbon; N: North

T. gondii antibodies were also detected in cats (Lopes *et al.*, 2008; Duarte *et al.*, 2010; Esteves *et al.*, 2014), sheep (Sousa *et al.*, 2009b; Lopes *et al.*, 2012a), cattle (Lopes *et al.*, 2012a), dogs (Lopes *et al.*, 2011b), wild boar (Lopes *et al.*, 2011a; Coelho *et al.*, 2014), goats

INTRODUCTION

(Lopes *et al.*, 2012a), pigs (Sousa *et al.*, 2006; Lopes *et al.*, 2012a; Esteves *et al.*, 2014), horses (Lopes *et al.*, 2013), chickens (Dubey *et al.*, 2006; Rodrigues *et al.*, 2019), birds (Waap *et al.*, 2008, 2012; Lopes *et al.*, 2011a), genets, wolf, red foxes and badgers (Lopes *et al.*, 2011a) (Table 1.3). Waap and colaborators studied a large number of pegeons describing a low prevalence rate, however, when considered at flock level, seropositive animals were identified in nearly one third of feeding sites sampled, with seroprevalence rates ranging between 5 % and 62.5 % (Waap *et al.*, 2008, 2012). In cats, the serological prevalence varied from 20 to 44 %; the highest being from stray cats from Lisbon (Table 1.3). The stray cats were captured from different areas of Lisbon and revealed the high serological titres, 18000 to 162000 (Waap *et al.*, 2012). These results may be due to the *T. gondii* type strains circulating or a very high environmental contamination leading to repeated infections in cats (Lopes *et al.*, 2014, 2017). In meat animal consumption, *T. gondii* prevalence varies between 7 % in cattle and pigs and 100 % in wild boars (Table 1.3). The three surveys performed in pigs demonstrated prevalences between 7 % and 16 %, which are not different from the heterogeneous values that have been reported in Europe (Dubey, 2009; Santoro *et al.*, 2017). Of note, most authors have used in-house ELISAs and thus these results are difficult to be compared due to the need for standardization of cut-off values, antigens and specific methodology procedures. In addition to meat from domestic pigs, meat from wild boars may also be an important source of *T. gondii* infection in humans.

Lopes and collaborators (2011) found 100 % of *T. gondii* prevalence in wild boars, however another study performed also in north of Portugal achieved the prevalence of 21 % when increased the number of wild boars studied (Table 1.3) (Lopes *et al.*, 2011a; Coelho *et al.*, 2014). In Portugal there is some tradition of wild boar meat consumption. This meat is widely used to cook uncontrolled homemade sausages that prepared without any heat processing and, therefore, are consumed raw, increasing the likelihood of *T. gondii* as a foodborne pathogen for humans (Coelho *et al.*, 2014). Additionally, infected viscera and carcasses left by hunters may pose a threat to scavenging susceptible animal.

It is believed that in Europe up to 63 % of human toxoplasmosis is attributed to the consumption of undercooked or cured meat products, however all the animals, like cats, wild animals or birds, could represent the zoonotic transmission but also the environmental contamination (Cook *et al.*, 2000; Berger-Schoch *et al.*, 2011; Belluco *et al.*, 2018).

Table 1.3 Seroprevalence of *T. gondii* in animals from Portugal

Host	Years	Region	Tested animals (n)	Prev (%)	Reference
BIRDS					
Booted eagle, Black kite, Bubo bubo, Common buzzard, Eurasian Falco peregrinus, sparrowhawk, Griffin vulture, Northern goshawk, Short-toed eagle, Tyto alba, White stork	2008-2010	N and C	52	50	(Lopes *et al.*, 2011a)
Common pigeon (*Columba livia*)	NS	L	695	5	(Waap *et al.*, 2008)
	2009-2010	L	1507	3	(Waap *et al.*, 2012)
Chicken (*Gallus domesticus*)	NS	N and C	225	27	(Dubey *et al.*, 2006)
	NS	C	178	5,6	(Esteves *et al.*, 2012; Rodrigues *et al.*, 2019)
MAMMALS					
Cat (*Felis catus*)	2003-2005	L	231	24	(Duarte *et al.*, 2010)
	2004-2005	N	204	36	(Lopes *et al.*, 2008)
	2007-2008	L	215	20	(Esteves *et al.*, 2014)
	2009-2010	L	423	44	(Waap *et al.*, 2012)
Dog (*Canis familiaris*)	2008-2009	N	673	38	(Lopes *et al.*, 2011b)
Cattle (*Bos taurus*)	2008-2010	N	161	7	(Lopes *et al.*, 2012a)
Sheep (*Ovis aries*)	2005-2006	N	1467	17	(Sousa *et al.*, 2009b)
	2008-2010	N	119	34	(Lopes *et al.*, 2012a)
Goat (*Capra hircus*)	2008-2010	N	184	18	(Lopes *et al.*, 2012a)
Pig (*Sus scrofa*)	2004-2005	N	333	16	(Sousa *et al.*, 2006)
	2007-2008	C and S	381	7	(Esteves *et al.*, 2014)
	2008-2010	N	254	10	(Lopes *et al.*, 2013)
Horse (*Equus caballus*)	2008-2010	N	173	13	(Lopes *et al.*, 2013)
Deer, European badger, Genet, Red fox, Wolf	2008-2010	N and C	12	83	(Lopes *et al.*, 2011a)
Wild boar (*Sus scrofa*)	2008-2010	N	8	100	(Lopes *et al.*, 2011a)
	2011-2012	N	97	21	(Coelho *et al.*, 2014)

L: Lisbon; C: Centre; N: North; NS: Not stated; Prev: Prevalence

1.6.2.   Genotypes circulating in Portugal

Infection with *T. gondii* can present severe symptoms and the disease may be fatal. However, it is generally asymptomatic, making difficult the identification and isolation of *T. gondii* strains in immunocompetent humans and animals, which are mandatory requisites for further genotyping procedures. Frequently, complete and accurate genotyping data requires DNA extracted from large numbers of viable parasites from cell culture or mouse ascitic fluid.

To our knowledge, genetic studies integrating intermediary and definitive hosts are lacking in Portugal, but the three genotypes have already been described in animals (chickens, pigs, bovine fetus and pigeons), revealing a majority of type II strains (Sousa *et al.*, 2006, 2010; Frenkel *et al.*, 2006; Waap *et al.*, 2008; Verma *et al.*, 2015). Of note, some results suggesting recombination and other genome rearrangements were already reported as the case of four chickens infected with *T. gondii* strains previously identified as type II and type III by PCR- RFLP that were later identified as new genotypes (genotype #254) and having combinations of types II and III specific alleles (Verma *et al.*, 2015).

Globally, there is a lack of data concerning genetic diversity of different loci. The scenario in Portugal is even worse than in Europe as only two studies exist. These consisted of two limited evaluations in humans performed in France with samples collected in Portugal (Sousa *et al.*, 2008; Ajzenberg *et al.*, 2009). One was based on serotyping (antibodies anti-GRA6) and revealed a majority of type II strains but also an unexpected high prevalence of non-type II strains that could not be typed as type I or III with this method. The high number of no type II strains showed a different typing profile for these samples when compared with French sera (Sousa *et al.*, 2008). The other study enrolled two Portuguese isolates (HIV) (Ajzenberg *et al.*, 2009) where both were type II.

Thus, the understanding of the molecular epidemiology and population genetics of *T. gondii* in Portugal is a mandatory issue.

CHAPTER II

AIMS AND OUTLINES

2.1 AIMS AND OUTLINES:

The main objective of this thesis was to perform a global genetic study of *Toxoplasma gondii* strains isolated from humans and animals in Portugal.

Considering that there is a lack of knowledge on the molecular characterization of *T. gondii* strains isolated in Portugal, both from human and animal sources, namely on the distribution of the different genotypes, on the identification of circulating strains with atypical/recombinant genomic backbones, and on the putative association between genotype and virulence, we pursued the following specific objectives:

i)   To study the molecular epidemiology of *T. gondii* strains isolated from animals and humans, by the classic *Sag*2 and microsatellites typing;

ii)   To study virulence differences of isolated *T. gondii* strains by macroscopic and cytological procedures using the mouse model;

iii) To evaluate major dissimilarities in the genomic backbone of *T. gondii* circulating strains in Portugal;

iv)   To evaluate the extent of the genomic mosaic structures of the two highly virulent recombinant strains that were previously isolated at URSZ-INSA;

The results obtained from this thesis may contribute for a better knowledge on the dynamics of *T. gondii* population circulating in Portugal, including the population diversity and the identification of sexual recombination. This may ultimate contribute for the development of more effective control strategies.

PART II

CHAPTER III

Genotyping characterization of *Toxoplasma gondii* in animals from the North of Portugal

Personal contribution

Anabela Vilares participated in the design of the study, performed most of the experimental work, analysis, interpreted data and wrote part of the article.

3.1 Abstract

Epidemiological investigations on *Toxoplasma gondii* infection have found a significant association between human toxoplasmosis, consumption of raw or undercooked meat and environmental contamination. The present study aimed to characterize genotypes of *T. gondii* strains isolated from 209 animals from the north of Portugal. Different types of samples revealing antibodies to *T. gondii* were used, namely heart, brain, diaphragm and tongue from two different groups of animals (Group I – 12 cattle, 34 sheep, 5 goats and 16 pigs, Group II – 6 stray dogs, 59 foxes, 49 boars and 28 hears). Eighty-nine animals (43 %) yielded positive results when B1 classical PCR was performed and *Sag*2-based genotyping of *T. gondii* was achieved in 29 % (60/209) of the animals. Within the later, 77% (46/60) of strains were type II, 10 % (6/60) were type I, and it was not possible to differentiate between *Sag*2 type I and type II alleles in 13% (8/60). *Sag*2 type II *T. gondii* strains have been shown for the first time to infect cattle, boars, foxes and hares in north of Portugal, and *Sag*2 type II has also been confirmed as the predominant genotype in pigs in the same region. This study constitutes the first molecular characterization report of *T. gondii* in goats, wild boars, foxes and hares from Portugal, and expands the range of the potential sources of human infection. Highlighting the importance of environmental contamination in the synanthropic and silvatic cycles.

3.2 Introduction

Toxoplasmosis, caused by the protozoan *Toxoplasma gondii*, is a global parasitic zoonosis. Domestic and wild felids are the definitive hosts of the parasite and a great variety of other homoeothermic animals constitute intermediate hosts. Herbivores acquire infection mainly by the ingestion of oocysts in water or contaminated food. Carnivores and omnivores, including human beings, can additionally become infected by ingesting meat with cysts (bradyzoites) (Dubey, 2010a).

Epidemiological investigations have found a significant association between human toxoplasmosis and the consumption of raw or undercooked meat or its products (Dubey, 2010a). Infections with the parasite are also an important cause of foetal mortality in sheep and goats (Chanton-Greutmann *et al.*, 2002) and it has originated outbreaks of abortion in small ruminant and swine farms, leading to serious economic and reproductive losses (Dubey, 2009). Most of the identified strains of *T. gondii* obtained from humans and animals in Europe, using polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) and microsatellite analysis, are grouped into one of three clonal lineages of type I, II or III (Ajzenberg *et al.*, 2002b, 2004, 2015). These genotypes have been described in animals in Portugal, with types II and III being common in pigs and chicken (Sousa *et al.*, 2006; Dubey *et al.*, 2006). Studies in other European countries, including France, Spain, Switzerland, and Germany, have also revealed a predominance of genotype II in animals (Dumètre *et al.*, 2006; Montoya *et al.*, 2008; Berger-Schoch *et al.*, 2011; Herrmann *et al.*, 2013).

Due to the relatively scarce information on the presence and molecular characterization of *T. gondii* in animals for human consumption in Portugal, this study aimed at detecting parasitic DNA in tissues from different animals that were raised, hunted or slaughtered in the north of the country and to genetically characterize the isolated *T. gondii* DNA.

3.3 Methods

3.3.1 Animals and samples

Between March 2008 and March 2012, different tissue samples were obtained from different animals that were divided in two groups, where group I included slaughter animals and group II included wild and stray animals. Sampling was done as following:

Group I - Heart tissue samples were obtained from cattle (n = 20), sheep (n = 40) and goats (n = 15), slaughtered in abattoirs from the north of Portugal. Heart (n = 4), brain (n = 10), diaphragm (n = 6) and tongue (n = 1) tissues were randomly sampled from 16 slaughtered pigs. All sampled animals had been born and raised in the north of Portugal for human consumption. Data collected for each sampled animal included gender, breed, age and husbandry system (intensive, semi-intensive or extensive). Twelve cattle, 34 sheep, 5 goats and the 16 pigs were seropositive to *T. gondii*, as determined (Department of Veterinary Sciences of UTAD) by an antibody titer ≥ 20 in the modified agglutination test (Lopes *et al.*, 2012a). Only the samples from these serology-positive animals were used in the present study.

Group II – Heart (n = 77), diaphragm (n = 22), tongue (n = 6), brain (n = 64), pool of heart and diaphragm (n = 37) and pool of heart, diaphragm and tongue (n = 15) tissues were collected from stray dogs (n = 6), foxes (n = 59), hunted boars (n = 49) and hares (n = 28) with positive serology to *T. gondii*. Tissues from stray dogs and foxes were collected from animals sacrificed in the highways. All animals were hunted or collected from the north of Portugal, where wild boars and hares were to human consumption.

Individual tissue samples from the two groups were stored at -20 ℃ until DNA extraction.

3.3.2 DNA extraction and identification of *T. gondii*

Tissue samples were digested in a trypsin or acid-pepsin solution for 2 hours at 37 ℃, washed with PBS, and centrifuged at 2500 rpm for 10 min, with pellets being stored at -20 ℃ until use. The DNA extraction was performed as previously described by Miller et al. (Miller *et al.*, 1988) and the DNA was resuspended in 100 µl of DNAse/RNase free water and stored at -20 ℃ until use.

The identification of *T. gondii* was performed by PCR amplification of the B1 locus, as previously described (Jones *et al.*, 2000).

3.3.3 Genotyping

Previously described nested PCRs were used to amplify separately the 5' and 3' ends of the surface antigen 2 gene (*Sag*2) (Howe *et al.*, 1997). Positive (*T. gondii* RH DNA) and negative (DNAse/RNase free water or AmpliTaq DNA polymerase® buffer) controls were included in every amplification batches. All amplification products were analyzed in 2 %

agarose gel electrophoresis with soaking GelRed^{TM} (Biotium, Hayward, USA). Visualization of a 241 bp (*Sag2 5'* region) or a 221 bp band (*Sag2* 3' region) was considered a positive result.

Sequence analysis for comparison of *Sag*2 genetic profiles was performed in order to assess the frequency of *T. gondii Sag*2 type strains (I, II and III). Thus, PCR products were purified with ExoSAP-IT$^®$ (Affymetrix, UK) and sequenced using BigDye^{TM} (Applied Biosystems, USA), following manufacturer's instructions and using the internal primers. Products were processed in an automatic sequencer 3130 XL Genetic Analyzer (Applied Biosystems, USA). The resulting sequences were aligned with MEGA5 software and were compared with the sequence data from *T. gondii* reference strains available from GenBank, namely: AM055943.1 RH type I, XM002371960.1 ME49 type II, AF357579.1 NED type III, AF357582.1 CASTELLS and AF357580.1 MAS different/atypical strains (GenBank).

3.4 Results

Regarding Group I, from the 91 studied animals, 67 (73.6 %) were positive for anti-*T. gondii* IgG and 35.8 % (24/67) of these seropositive animals were also found to be positive by *B*1 classical PCR. *Sag*2 PCR identified 27 % (18/67) of positive animals. Of the 142 animals studied from Group II (all with positive anti-*T. gondii* IgG), 65 (46 %) yielded positive results by using the B1 classical PCR (Table 4) and 60 (42.3 %) were *Sag*2 PCR positive. *Sag*2 genotyping allowed the discrimination among the three "classical" *T. gondii* types for 29% (60/209) of the animals and the genotyping success was similar in group I (27 % - 18/67) and in group II (30 % - 42/142) (Table 3.1 and Table 3.2).

Forty six strains from animals were type II, six were type I, and it was not possible to differentiate between *Sag*2 type I and type II alleles in other eight (13 %) out of the 60 characterized isolates.

Table 3.1 *Sag*2 genotype of *T. gondii* strains from Group I

Strain	Gender	Breed	Age group (months)	Husbandry system	MAT titter	*Sag*2 Type
Cattle						
TgPtB$_1$	M	Crossbreed [H]	8-12	SI	100	II
TgPtB$_2$	F	Crossbreed [H]	8-12	SI	100	II
TgPtB$_3$	M	Crossbreed [H]	8-12	SI	100	II
TgPtB$_4$	F	Mirandesa [H]	8-12	SI	100	neg
TgPtB$_5$	M	Crossbreed [H]	5-8	SI	100	neg
TgPtB$_6$	M	Arouquesa [H]	5-8	SI	100	neg
Sheep						
TgPtS$_1$	F	Cross of Churra [H]	≥ 19	E	20	II
TgPtS$_2$	M	Churra Terra Quente [H]	7-18	SI	≥ 6400	I or II
TgPtS$_3$	M	Churra Badana [H]	≤ 6	SI	20	II
TgPtS$_4$	M	Churra Terra Quente [H]	≤ 6	SI	20	II
TgPtS$_5$	M	Churra Terra Quente [H]	≤ 6	SI	20	II
TgPtS$_6$	F	Churra Terra Quente [H]	≤ 6	SI	400	I or II
Goat						
TgPtG$_1$	M	Preta de Montesinho [H]	≤ 6	SI	20	I or II
TgPtG$_2$	M	Preta de Montesinho [H]	≤ 6	SI	20	I or II
TgPtG$_3$	M	Preta de Montesinho [H]	≤ 6	SI	20	I or II
Pig						
TgPtP$_1$	M	Cross of Bísaro [H]	≤ 3	I	400	I or II
TgPtP$_2$	F	Cross of Bísaro [H]	≤ 3	I	400	II
TgPtP$_3$	F	Cross of Bísaro [H]	4-9	I	20	II
TgPtP$_4$	F	Piétrain [Br]	4-9	I	20	neg
TgPtP$_5$	F	Piétrain [D]	4-9	I	20	II

TgPtP$_6$	F	Piétrain [D]	≥ 10	I	20	neg
TgPtP$_7$	F	Piétrain [Br]	≥ 10	I	20	I or II
TgPtP$_8$	F	Bísaro [Br]	≥ 10	SI	20	I or II
TgPtP$_9$	F	Bísaro [D]	≥ 10	SI	20	neg

B: bovine; Br: brain; D: diaphragm; F: female; G: goat; H: heart; M: male; MAT: modified agglutination test; neg: negative; P: pig; Pt: Portugal; S: sheep; Tg: *Toxoplasma gondii*; [a] same individual pig. SI: Semi Extensive; E: Extensive; I: Intensive

Table 3.2 *Sag*2 genotype of *T. gondii* strains from animals of Group II

Host	PCR B1 positive % (n)	PCR *Sag*2 positive % (n)	*Sag*2 Type		
			I	II	III
Dogs	50 (3)	50 (3)	0	3	0
Foxes	42 (25)	32 (19)	2	10	0
Boars	35 (17)	37 (18)	2	8	0
Hares	71 (20)	71 (20)	2	15	0
Total	46 (65)	42 (60)	6	36	0

Table 3.3 *T. gondii* strains detected in samples from Group II

Host Product (n)	PCR B1 positive % (n)	PCR *Sag*2 positive % (n)	Sanger Sequencing	*Sag*2 Type (n)
Dogs		%	%	
Br (6)	17 (1)	17 (1)	-	-
H (6)	33 (2)	33 (2)	50 (1/2)	II
D (6)	33 (2)	33 (2)	50 (1/2)	II
T (6)	33 (2)	33 (2)	50 (1/2)	II
Foxes	%			
Br (50)	38 (19)	22 (11)	64 (7/11)	II (5) and I (2)
H (42)	12 (5)	12 (5)	100 (5/5)	II
H+D+T (15)	20 (3)	20 (3)	100 (3/3)	II
H+D (1)	0	0	-	-
Boars				

Br (8)	0	0	-	-
H (29)	24 (7)	24 (7)	100 (7/7)	II (6) and I (1)

Host Product (n)	PCR B1 positive % (n)	PCR *Sag*2 positive % (n)	Sanger Sequencing	*Sag*2 Type (n)
D (16)	50 (8)	50 (8)	38 (3/8)	II (2) and I (1)
H+D (8)	38 (3)	38 (3)	33 (1/3)	II (1)
Hares				
H+D (28)	71 (20)	71 (20)	85 (17/20)	II (15) and I (2)

Br: brain; D: diaphragm; H: heart; T: tongue.

3.5 Discussion

The farm animals are an integrative part of human and animal food chain and share the same ecosystems than humans, representing a risk of *T. gondii* transmission. Also, wild boars are extensively hunted and consumed by humans. This meat is widely used to make different meals and uncontrolled homemade sausages that are prepared without any heat processing. Therefore, this meat is consumed raw, increasing the likelihood of *T. gondii* as a foodborne pathogen for humans (Coelho *et al.*, 2014). In contrast, in Portugal, foxes don't have a major role in human food chain and don't have predators, however they could represent the environmental contamination and could take part in scavenging other animals. Different approaches have been reported worldwide to identify *T. gondii* in wild animals. The application of these methods has generated invaluable information to enhance our understanding of the *T. gondii* prevalence, however, since most studies focused solely on the detection rather than on the genetic characterization of *T. gondii*, the information obtained is limited. In Portugal, despite the existence of four seroprevalence studies (Lopes *et al.*, 2011a; b; Coelho *et al.*, 2014, 2015) the scenario is even worst because the knowledge of genotypes circulating in animals is very scarce.

In the present study, we demonstrated not only the presence of *T. gondii* in cattle, sheep, goats, pigs, dogs, foxes, wild boars and hares but have also identified the archetypes evolved in the infection. In group I, 73.6 % (67/91) of studied animals presented positive serology for *T. gondii* and yielded 35.8 % (24/67) molecular positivity by B1 classical PCR. This is in

agreement with the wide distribution of *T. gondii* described in the rest of Europe (Gustafsson and Uggla, 1994; Hejlícek *et al.*, 1995; Aspinall *et al.*, 2003; Frölich *et al.*, 2003; Dubey *et al.*, 2004a; Hůrková and Modrý, 2006; Dumètre *et al.*, 2006; Richomme *et al.*, 2009; Aubert *et al.*, 2010; Bártová *et al.*, 2010; De Craeye *et al.*, 2011; Jokelainen *et al.*, 2011, 2012a; Fernández-Aguilar *et al.*, 2013; El Behairy *et al.*, 2013; Machacova *et al.*, 2014; Calero-Bernal *et al.*, 2015; Kornacka *et al.*, 2016). Of note, although a small number of goats was analysed, it is worth mentioning that positive animals were from three geographically distant pasture areas, which suggests the existence of a wide distribution of contamination. In pigs, for which existence of *T. gondii* is an indicator of the hygienic status and the risk of infection for a given pig farm, the molecular identification of *T. gondii* in the present study (43.8 %) was higher than the one observed in pigs from the centre and south of Portugal (35 %) (Esteves *et al.*, 2014). Possible due to hypothesis that environment is more contaminated in the north of Portugal than south and because of that, pigs could have itself more cysts which facilitates the molecular identification. We observed lower infection rates in sheep than the ones reported in several European countries (Armand *et al.*, 2016; Amdouni *et al.*, 2017). However, the molecular detection of the parasite in animals for which a positive serology was detected is sometimes a tricky task, as described in several studies (Aubert *et al.*, 2010; Fernández-Aguilar *et al.*, 2013; Turčeková *et al.*, 2014). In fact, the difficulty of achieving reliable results in detecting *T. gondii* DNA in tissues may be explained due to the juvenile animal age, the chosen organ (e.g., liver), or the limited size of the samples (since 100 g of beef can contain only one parasite) (Dubey and Thulliez, 1993; Dubey *et al.*, 1995; Gajadhar *et al.*, 1998; Esteban-Redondo *et al.*, 1999; Mason *et al.*, 2010; Fernández-Aguilar *et al.*, 2013). To overcome this problem several authors opt by performing the bioassay (i.e., mice inoculation) to increase the number of parasites and thus obtain high-quality DNA, before proceeding with the molecular identification and genotyping steps (Dubey *et al.*, 2008; Prestrud *et al.*, 2008; Yai *et al.*, 2008). However, as occurred in our study, this is not always possible in farm animals and wildlife as sampling activities may occur days after the death of the animals, hindering the recovery of viable *T. gondii*. Furthermore, the bioassay is extremely expensive and time-consuming. In these scenarios, B1 PCR may be useful in these sort of samples for the identification of the parasite, while *Sag2* is widely used for genotyping *T. gondii* strains into one of the archetypal lineages (I, II, III) (Howe *et al.*, 1997). The option for the B1-based PCR for identification purposes instead of *Sag2*-based PCR relies on the fact that, similarly to other studies (Lass *et al.*, 2009), the former shows higher sensitivity than the latter (in our study, 42.6 % and 37.3 %, respectively).

In this study, the majority of isolates from farm animals were type II, mirroring the scenario observed in the rest of Europe (Aspinall *et al.*, 2002a; Jungersen *et al.*, 2002; Dumètre *et al.*, 2006; Halos *et al.*, 2010; Berger-Schoch *et al.*, 2011), with the exception of cattle isolates where the majority of type I strains prevails in Europe (Mahami-Oskouei *et al.*, 2017; Battisti *et al.*, 2018). Although in the present study only type II strains were identified in cattle, Verma and colleagues previously described one type I *T. gondii* isolated from a deceased bovine foetus in Portugal (the only one genotyping report in cattle) (Verma *et al.*, 2015).

More unpredictable is the situation in wild life. For example, Italy have highlighted the presence of all three clonal genotypes and a number of atypical genotypes with different prevalence in wild animals (Mancianti *et al.*, 2013; Verin *et al.*, 2013; Bacci *et al.*, 2015; Formenti *et al.*, 2016). A higher prevalence of *T. gondii* genotype I has been observed in omnivorous domestic species, represented by pigs, than in omnivorous and carnivorous sylvatic species (wild boars and foxes). In contrast, recombinant genotypes seemed to be more prevalent in wildlife than in domestic species (Grigg and Sundar, 2009; Battisti *et al.*, 2018). However, this apparent increased genetic diversity (e.g. recombinant strains) from wildlife samples is not absolute, as studies of wildlife in France, Germany and Norway have reported a high prevalence of archetypal strains, where the majority of type II prevails (Dubey *et al.*, 2004a; Prestrud *et al.*, 2008; Richomme *et al.*, 2009; Aubert *et al.*, 2010; Herrmann *et al.*, 2013) or 50 % in Spain (Calero-Bernal *et al.*, 2015). In our study, in opposition to group I, wildlife animals revealed type I strains (although type II strains remains the more prevalent type in this group) (Richomme *et al.*, 2009; Calero-Bernal *et al.*, 2015). Type I was detected in wild boars, foxes and hares, which is somehow surprising because this genotype is not frequently reported in animals in Europe (Berger-Schoch *et al.*, 2011; De Craeye *et al.*, 2011). In support of the genotypes heterogeneity of *T. gondii* in the wild life, fourth clonal type strains have been described as the most prevalent in the USA (Dubey *et al.*, 2011a).

In summary, the present study constitutes the first report of *T. gondii* genotyping in wild animals and goats in Portugal and shows that different animals for human consumption and the ecosystem from the north of Portugal are infected and contaminated with *T. gondii*. Of note, studies conducted in the same region showed a low prevalence of anti-*T. gondi* antibodies in women of childbearing age (24.4 %) and 13 % in general population from north of Portugal, elevating the risk of primary infection during the pregnancy in this region (Lopes *et al.*, 2012b, 2014; Gargaté *et al.*, 2016). The finding of a higher prevalence of type II strains

in both domestic and sylvatic animals is not surprising since the moderate virulence of type II strains does not lead to a frequent death of the host, allowing the animals to survive until their slaughtering or hunting (Weiss and Kim, 2000). Subsequently, it is likely that seropositive animals have infective forms in their tissues, and the consumption of undercooked meat or meat products from these animals remains a risk factor for humans. Future studies enrolling more Portuguese territories and a higher number of animal species will be needed to better understand the scenario of *T. gondii* animal infection and the environmental contamination in Portugal, which will be important to control this important infection in both humans and animals.

CHAPTER IV

Isolation and molecular characterization of *Toxoplasma gondii* isolated from pigeons and stray cats in Lisbon, Portugal

Personal contribution

Anabela Vilares designed the study, performed almost the experimental work and analysis, interpreted data and wrote the manuscript.

Isolation and molecular characterization of *Toxoplasma gondii* in pigeons and stray cats in Lisbon, Portugal

4.1 Abstract

Cats and pigeons are important factors in the epidemiology of *Toxoplasma gondii* as felids are the only definitive hosts that can excrete environmentally resistant oocysts, and pigeons share the same places of cats and humans constituting a good model and indicator of the ground field contamination. We aimed to study the virulence and genotypes of *T. gondii* isolated from pigeons and stray cats in Lisbon - Portugal. Fresh samples of brain from 41 pigeons and 164 cats revealing antibodies to *T. gondii* were inoculated in mice. Three isolates (one isolated from a cat and two isolated from pigeons) were virulent in the mouse model. *Sag*2-based genotyping of *T. gondii* was achieved in 70.7 % (29/41) of samples isolated from pigeons (twenty-six samples were type II, two were type III, and one strain was type I). From the cat brain samples, 50 % (82/164) yielded *Sag*2 positive results, where 72 belonged to genotype II and 10 were no type III (it was not possible to discriminate between type I and II). Further genotyping was obtained by multiplex PCR of 5 microsatellites (*Tub*2, TgM-A, W35, B17, B18), allowing the identification of two recombinant strains that had been previously identified as type II by *Sag*2 amplification (one isolated from cat brain and the other from pigeon brain). This is the first evidence of recombinant strains circulating in Portugal and the first report of *T. gondii* genotyping from cats in this country. This study also highlights the importance of environmental contamination in the synanthropic cycle constituting a potential source of human infection.

Isolation and molecular characterization of *Toxoplasma gondii* in pigeons and stray cats in Lisbon, Portugal

4.2 Introduction

Toxoplasma gondii is a coccidian and is one of the most worldwide prevalent parasite infecting warm-blooded vertebrates (Frenkel *et al.*, 2006; Weiss and Kim, 2007), especially in Europe (primarily Southern Europe), South and Central America and Africa (Aubert and Villena, 2009; Schlüter *et al.*, 2014; Sepúlveda-Arias *et al.*, 2014). The *T. gondii* infectious stages involve the tachyzoites, which facilitate expansion during acute infection, the bradyzoites, which maintain chronic infection, and the sporozoites, which are disseminated in the environment within oocysts (Dubey, 1998a). The ingestion of undercooked meat containing tissue cysts has been considered the most common means for acquisition of toxoplasmosis, followed by the ingestion of vegetables or water contaminated with sporulated oocysts (Tenter *et al.*, 2000).

Felids are the only known definitive hosts that can release oocysts into the environment through their faeces (Dubey, 2010a), while tissue cysts containing bradyzoites may be found in both intermediate and definitive hosts. Sporulated infective oocysts are extremely robust and, under moderate conditions may survive for months or even years in soil (Yilmaz and Hopkins, 1972). As definitive hosts, stray cats play an important role in the life cycle of *T. gondii*, facilitating the genetic recombination between strains, as well as environmental contamination (Howe and Sibley, 1995; Grigg *et al.*, 2001a). The high *T. gondii* seroprevalence rates found among various groups of strict vegetarians (Hall *et al.*, 1999) support the importance of environmental or water contamination as a source of human infection (Dubey, 1997; Montoya and Liesenfeld, 2004). Subsequently, animal species that feed on contaminated soil are easily infected by oocysts present in the environment (soil, food or water). Pigeons share the same "habitat" of cats and humans, as flocks are observed in recreational areas such as city parks and playgrounds. The interaction between cats, pigeons and humans is quite evident, which favours the fecal-oral transmission of *T. gondii* between the definitive and intermediate hosts.

The vast majority of *T. gondii* strains in North America and in some European countries are classified into the three known lineages (types I, II and III). However, recent studies revealed that strains from South America and Africa are genetically more diverse, including recombinant and atypical strains. These strains have been associated with severe acute disseminated toxoplasmosis in both immunocompromised and immunocompetent patients (Grigg *et al.*, 2001a; Bossi and

Bricaire, 2004; Lehmann *et al.*, 2006; Cavalcante *et al.*, 2007; Vaudaux *et al.*, 2010; Demar *et al.*, 2012; Micic *et al.*, 2013). Different studies of *T. gondii* infection in cats and pigeons have been performed since 1972 (Pendergraph, 1972; Kapperud, 1978; Kirkpatrick *et al.*, 1990; Mushi *et al.*, 2000; Tsai *et al.*, 2006; Waap *et al.*, 2008, 2012; Salant *et al.*, 2009; Kulasena *et al.*, 2011; Yan *et al.*, 2011), however, the methodologies used in those studies are quite variable, as well as the positivity criteria and the number of animals that were manipulated. Concerning *T. gondii* genotyping in pigeons, the only study performed so far was conducted in Portugal. It enrolled 695 pigeons and showed 4.6 % prevalence and a majority of type II strains with no identification of atypical or recombinant strains (Waap *et al.*, 2008). Regarding studies in cats, most published data show a majority of type II strains (Montoya *et al.*, 2008; Jokelainen *et al.*, 2012b) and one study held in China reported a majority of atypical strains (Qian *et al.*, 2012).

Little is known about the circulating strains in Portugal, either from the point of view of its virulence or their genotypic characterization. We have previously analyzed the serological state of 1507 pigeons and 467 cats and reported a 2.6 % and 44.2 % prevalence rate, respectively (Waap *et al.*, 2012). In the present study, we intended to evaluate the *T. gondii* genotyping distribution in cats and pigeons in Lisbon, and to study the virulence of the isolated strains in mice.

4.3 Methods

4.3.1 Samples

We studied 41 pigeons and 167 cats seropositive to *T. gondii* by the modified agglutination test (MAT) from 1507 free-living pigeons that were captured from 64 different feeding sites of Lisbon, and 467 cats from periodic catches made by Lisbon City Hall services, between 2009 and 2011 (Waap *et al.*, 2012).

4.3.2 Sample preparation

Brain tissues were collected under aseptic conditions and within 24 hours after euthanasia from 41 pigeons (including two pigeons with doubtful serology) and from 164 cats that had revealed antibodies to *T. gondii*. Three cats' brains were discarded as they have not reached the

laboratory in proper conditions. Brain homogenates were prepared from approximately 1 cm^3 of brain tissue and 3 ml of phosphate buffer solution (PBS) at pH 7.4 with 21 G needles, prior to their inoculation into mice. We also collected fifteen muscle samples from cats, in aseptic conditions, and digested them with acid-pepsin before inoculation and DNA extraction, according to a previously described procedure (Dubey, 1998a).

4.3.3 *In vivo* assay - Inoculation in mice

The inoculation, maintenance and euthanasia of mice were performed under the standards of the URSZ-INSA, established for the pre- and post-natal diagnosis of toxoplasmosis. These guidelines are in accordance with the Protocol of International Guiding Principles for Biomedical Research Involving Animals as issued by the Council for the International Organizations of Medical Sciences. Briefly, each homogenate was mixed with antibiotics (250 UI / ml penicillin and 500 μg / ml streptomycin) and was used to inoculate (1 ml / mouse) two female mice Hsd: ICR (CD-1$^®$) with negative serology to endoparasites (Harlan Ibérica, Barcelona), by intraperitoneal injection. The animals were monitored serologically by MAT, in order to check for the presence of *T. gondii* antibodies at 15, 30, 60 and 90 days post-inoculation (p.i.). At the end of this period, mice were euthanized by inhalation of 5 % halothane and confirmed death by physical examination. The general morphological state of organs was observed; sample tissues were collected from the heart, liver, spleen and brain, and were visualized by optical microscopy in order to search for *T. gondii* cysts.

4.3.4 DNA extraction and identification of *T. gondii*

The DNA was extracted from three different products: the original brain samples from pigeons and cats, the muscles from cats treated with acid-pepsin and also the brain samples from MAT positive mice. The DNA extraction was performed from 400 μl of each product, with 200 μl of extraction buffer, 40 μl of proteinase K by QIAamp blood kit (Qiagen, Chatsworth, USA) according to the manufacturer's protocol. The samples were eluted with 50 μl of elution buffer.

The identification of *T. gondii* was performed by PCR amplification of the B1 locus, as previously described (Jones *et al.*, 2000).

4.3.5 Genotype analysis of *T. gondii*

For genotyping purposes, all samples (cat and pigeon's brains and the brains from inoculated mice) were subjected to amplification and sequencing of the two ends of *Sag*2 gene, and to a multiplex PCR of five microsatellites: the beta-tubulin (*Tub*2), the myosin A (TgM-A) genes and three EST (W35, B17, B18) markers, as described elsewhere (Howe *et al.*, 1997; Ajzenberg *et al.*, 2002a, 2004). For *Sag2*, the PCR products were visualized in GelRED (Biotarget, Lisbon, Portugal) stained 2 % agarose gel electrophoresis, purified and sequenced in ABI 3130 xl Genetic Analyzer (Applied Biosystems). For microsatellites, electrophoresis of PCR products was performed on the ABI 3130 xl Genetic Analyzer (Applied Biosystems), and the data were stored and analyzed with Gene Mapper (version 3.7; Applied Biosystems). MEGA5 software (Tamura et al 2011) was used to identify the genotypes by comparing the obtained sequences with the ones from reference strains available in GenBank. Nucleotide sequence data reported in this paper are available in the GenBank™ databases under the accession numbers: KJ754389 - KJ754489 and KJ754490 - KJ754499. A conservative approach was used when analyzing the data from microsatellites: as none of the generated fragments was sequenced and as only atypical strains reveal insertion/deletion or mutation events in microsatellites analysis (Ajzenberg *et al.*, 2005), all non-atypical strains were considered to share the same sequence in these markers as the reference strain belonging to the same type (I, II or III). This strategy has been previously used by other authors (Mercier *et al.*, 2010).

4.4 Results

4.4.1 Identification of *T. gondii*

PCR targeting the B1 gene was performed on *T. gondii* DNA extracted from brain homogenates of 41 pigeons and 164 cats, and from 15 cats muscles, yielding a positivity rate of 68 % (28/41), 51 % (84/164), and 20 % (3/15), respectively. No inhibitors were found in the negative samples.

Isolation and molecular characterization of *Toxoplasma gondii* in pigeons and stray cats in Lisbon, Portugal

4.4.2 *In vivo* assay

For pigeons, the inoculation of brain samples in mice revealed a *T. gondii* isolation rate of 58.5 % (24/41), whereas for cats the isolation rate was 11 % (18/164). The inoculation of muscle tissues was not successful. All but five inoculated mice, regardless the origin of the sample, showed antibodies to *T. gondii* from the 15th day p.i. One of the exceptions occurred for one mouse inoculated with a brain sample from a cat, as it died 48 hours p.i. The "post-mortem" microscopy analysis of this mouse's ascite revealed a large number of bacteria and the absence of *T. gondii* tachyzoites. Three other mice inoculated with samples from cats' brains yielded positive serology after two months p.i. (two mice) or one month p.i. The last scenario was also observed for one mouse inoculated with a brain sample from a pigeon.

In the "post-mortem" analysis of mice with positive serology, the organs were macroscopically normal and cysts were found in all brain tissues, with a sole exception previously inoculated with a cat's brain. Three samples (one isolated from a cat and two isolated from pigeons) were virulent in the mouse model: one mouse revealed a detour in the spine, another one paralyzed on the left side and the other died at 1 month p.i.

4.4.3 Genotype analysis

Considering the low sensitivity of the B1 PCR, we opted by performing the *Sag*2 PCR for all 41 homogenates of pigeons brains that were inoculated in mice, as all of them had yielded positive serology.

*Sag*2 genotyping allowed the discrimination among the three "classical" *T. gondii* types. Twenty-six samples were type II, two were type III, and one strain was type I. Microsatellites' analysis was applied to all samples that had revealed positive results in the *Sag*2 PCR amplification, but showed low sensitivity likely due to the low *T. gondii* DNA concentration, or to the initial organic matrix product. In fact, only 55.2 % (16/29) of *Sag*2 positive samples were also positive by microsatellites' PCR. This methodology allowed the identification of 12 type II, two type III, one type I and one recombinant strain (B17 - II/III, *Tub*2 - II/III, W35 - II/III, TgM-A - II, B18 - I/III). Only the latter did not confirm the type previously identified by *Sag*2 typing (Type II) (Table 4.1).

Isolation and molecular characterization of *Toxoplasma gondii* in pigeons and stray cats in Lisbon, Portugal

Table 4.1. Typeable *Toxoplasma gondii* isolates from pigeons.

		Microsatellites				
		I	II	III	Rec	NT
	I	1	–	–	–	–
Sag2	**II**	–	12	–	1[a]	13
	III	–	–	2	–	–

12 strains were non-typeable for both methods; Rec – recombinant strain, NT – non-typeable.
[a] B17 – 334 bp (II/III), *Tub*2 – 289 bp (II/III), W35 – 242 bp (II/III), TgM-A – 202 bp (II), B18 – 158 bp (I/III)

As for pigeons, *Sag*2 PCR was performed for all 179 homogenates from cats (164 from brains and 15 from muscles) that were inoculated in mice, as all of them yielded positive serology in infected mice. From the brain samples, 50 % (82/164) yielded *Sag*2 positive results, where 72 belonged to genotype II and for 10 samples it was not possible to discriminate between type I and II due to the impossibility to amplify the 2nd fragment of *Sag*2 (Table 4.2).

Table 4.2. Typeable *Toxoplasma gondii* isolates from cats.

		Microsatellites				
		I	II	III	Rec	NT
	I	–	–	–	–	–
Sag2	**II**	1[a]	16	–	–	55
	III	–	–	–	–	–
	Non-type III	–	–	–	–	10

82 strains were non typeable for both methods; Rec – recombinant strain, NT – non-typeable. [a] Recombinant strain.

Only three muscle samples yielded type results, which matched the genotypes obtained with the "corresponding" brains. All 84 B1 PCR positive samples were also positive for *Sag2*, which also recovered two B1 negative samples. The microsatellites typing strategy (applied to all 86 *Sag2* positive samples) was highly inefficient also in samples originated from cats as only 17 samples were successfully typed. From these 17 samples, 16 were identified as type II strains (in agreement with *Sag2* typing) and one as type I (previously classified as type II by *Sag2*), indicating that the later was a recombinant strain.

The differentiation of strains was also performed in mice's brains inoculated with homogenates from cats and pigeons and the results were consistent with those previously obtained directly from biological products (brain from cats or pigeons).

It is worth noting that genotyping was achieved for *T. gondii* strains from all mice with positive serology (24 out of 24 pigeons and 18 out of 18 cats). Concerning all original products (brains from pigeons and cats), the genotyping rates were 61 % (25/41) for pigeons, and 42 % (69/164) for cats.

4.5 Discussion

Cats are the major source of transmission of *T. gondii*, since they are the only urban host capable of excreting oocysts to the environment. Subsequently, animal species that feed on contaminated soil are easily infected by oocysts presents in the environment (soil, food or water).

In Lisbon, as in most cities of the world, stray cats proliferate due to the ease of finding food, water, and shelter in these urban centres, and thus they may be responsible for serious health problems in humans. The distribution of cats depends heavily on socio-demographic characteristics of certain Lisbon areas, as the oldest areas of the city register a special abundance of stray cats and the subsequent close contact with humans. Pigeons share the same "habitat" of cats and humans, as flocks are observed in recreational areas such as city parks and playgrounds. The interaction between cats, pigeons and humans is quite evident, which favours the fecal-oral transmission of *T. gondii* between the definitive and intermediate host in the urban cycle.

Different studies involving ground level feeding animals (such as cats, pigs, and chickens) have been performed worldwide revealing the presence of *T. gondii* (various genotypes were

Isolation and molecular characterization of *Toxoplasma gondii* in pigeons and stray cats in Lisbon, Portugal

identified in some of those studies) (Dubey *et al.*, 2006, 2007b; Su *et al.*, 2006; Cavalcante *et al.*, 2007; Salant *et al.*, 2009; Dubey, 2010b; Yan *et al.*, 2011; Qian *et al.*, 2012; Herrmann *et al.*, 2013). Some *T. gondii* studies focusing on diverse animals have been conducted also in Portugal, but none of them have combined genetic and virulence characterization of *T. gondii* strains in cats. In pigeons, this dual evaluation was previously performed by our group over a very small number of samples (Waap *et al.*, 2008). In the present study, we have obtained an isolation rate after inoculation in mice of 58.5 % (24/41), which was considerably higher than our previous study (39 %; 9/23). Also, it was much more successful than the one performed in Brazil (Godoi *et al.*, 2010) enrolling 126 pigeons, where no isolation was achieved. In the latter, 22 (6 %) mice died, none of them presented any form of *T. gondii,* and all the remaining living mice were seronegative at the end of the *in vivo* assay. A possible explanation for the difference in the isolation rates could be the amount of brain used and parasite concentration in brain samples, which are difficult to control when a small amount of tissue is available for infection (Dubey *et al.*, 2005a).

Concerning the inoculation of biological products derived from the definitive host, the isolation rate *in vivo* was quite low (11 %; 18/164) when compared with previous studies that reported isolation rates in mice of 25 % (2/8) (Dubey *et al.*, 2007c) and 32.1 % (36/112) (Al-Kappany *et al.*, 2010). Mice inoculated with cats' muscles showed an isolation rate of 0 %, in contrast with a previous study revealing an isolation rate of 52.6 % (10/19) (Dubey *et al.*, 2007a). Considering that the brain samples of this group of cats also yielded negative results in this *in vivo* assay, we hypothesize that their infectious load was lower when compared with the remainder animals. Although the costs of a laboratory animal care and animal sacrifice remain a reality, our data indicate that the genotyping results after mice inoculation may be considerably enhanced.

Regarding the *T. gondii* identification through PCR directed to the B1 gene, the identification rate was again higher (68 %, 28/41) in initial brain homogenates from pigeons than from cats (51 %, 84/164). The differences in effectiveness of identification may be due to the low concentration of parasites or even to the absence of cysts in the brain area analyzed in cats. In fact, the cat's brain is larger, and the *T. gondii* appetence to the brain tissue of the definitive host seems to be lower than to the intermediate host (Dubey, 1997). Although B1 gene has been

traditionally used for identification purposes (Jones *et al.*, 2000), we opted by performing the *Sag2* based genotyping in all biological samples (homogenates) as we had obtained positive serology in mice that were B1 negative. This strategy seemed adequate, as we were also able to use *Sag2*-based genotyping in one B1 negative *T. gondii* strain. The *T. gondii* differentiation by *Sag2* gene and microsatellites through the use of mice's brains was easier to perform than in the original biological products (brains from pigeons or cats), which was likely due to the fact that the mice inoculation enhances the amount of parasites. Another possible explanation is that the brain from both pigeons and cats are more fibrous, and thus more difficult to digest.

The majority of the isolates from both hosts were classified as type II strains, which is consistent with what has been reported worldwide (Howe *et al.*, 1997; Fazaeli *et al.*, 2000; Honoré *et al.*, 2000; Ajzenberg *et al.*, 2005; Waap *et al.*, 2008). We have also isolated types I, III and one recombinant strain in pigeons, whereas in cats only one strain was non-type II (one recombinant strain that was type II by *Sag2* but type I by microsatellites). Mouse death with the type III strain was observed after 48 hours p.i. which was somehow surprising since type III are classified as nonvirulent strains ($LD_{50} > 10^5$ parasites) (Science Polymorphic Secreted Kinases are key virulence factors in Toxoplasmosis). The identification of type III strains in animals have been reported by other authors. In fact, Sousa (2006) revealed that in 37 seropositive pigs four out of 15 isolates were type III (11 were type II) (Sousa *et al.*, 2006). In chickens, Dubey (2006b) identified four type III out of 12 isolates (eight were type II), where none of the isolates were lethal for mice (Dubey *et al.*, 2006). Type I strains have been rarely found (Dubey *et al.*, 2004b; Montoya *et al.*, 2008) and none have been previously described in Portugal, except in the preliminary study carried out by our team in 2006 (Waap *et al.*, 2008). These strains are usually associated with high virulence in humans and generally cause mouse death within days (Sibley and Boothroyd, 1992b; Sibley *et al.*, 1999; Grigg *et al.*, 2001b). In the present study, the type I strain identified from one pigeon was not isolated *in vivo* but it was genotyped from the original product (the pigeon's brain). The difficulty of isolation of these strains may be related to the small number of cysts present in the different organs, and also because of the virulence of type I strains that can cause the death of animals before parasite encystment (Dubey *et al.*, 2005b; c; d). Atypical and recombinant strains have been described diverging from the classical phenotype of genotypes I, II and III (Ajzenberg *et al.*, 2004; Bossi and Bricaire, 2004; Lehmann *et al.*, 2006;

Cavalcante *et al.*, 2007). Type I or type I-like atypical isolates are more likely involved in severe retinochoroiditis in humans (Grigg *et al.*, 2001b) and the atypical isolates often cause severe acute disseminated toxoplasmosis in immunocompetent patients (Bossi and Bricaire, 2004). Different studies in humans and animals (namely chickens and dogs) have revealed that strains from South America and Africa seem to be genetically more diverse, frequently including recombinant (Dubey *et al.*, 2007d; Lindström *et al.*, 2008) and atypical strains (Ajzenberg *et al.*, 2004; Lehmann *et al.*, 2006; Cavalcante *et al.*, 2007). However, a recent study in Italy, have also reported the identification of two recombinant strains in goats (Mancianti *et al.*, 2013). In those studies as well as in the present work, the microsatellites typing method, which uses one multiplex PCR to perform multilocus typing with five markers (Ajzenberg *et al.*, 2005), was an essential tool for the identification of recombinant and atypical strains.

Recombinant strains had never been described in Portugal before and we could expect a *T. gondii* genotype distribution in Portugal similar to the one observed in other European countries (Howe *et al.*, 1997; Fazaeli *et al.*, 2000; Honoré *et al.*, 2000; Ajzenberg *et al.*, 2005; Waap *et al.*, 2008). However, Portugal has a long history of trade and social interaction with South American countries (such as Brazil) and China (where atypical and recombinant strains are not uncommon), which involve human migrations, the import of food products (like meat or vegetables with the potential of transmitting the parasite), and also the presence of rats and cats (*T. gondii* intermediate and definitive hosts, respectively) in commercial ships. These factors may explain the identification of recombinant strains in the present study held in Portugal.

CHAPTER V

Molecular and virulence characterization of *Toxoplasma gondii* strains isolated from humans in Portugal

Personal contribution

Anabela Vilares designed the study, performed all most the experimental work and bioinformatics analyses, interpreted data and wrote the manuscript.

5.1 Abstract

Toxoplasma gondii is an apicomplexan parasite responsible for toxoplasmosis which infects all warm-blooded vertebrates, including mammals and birds. The majority of studies conducted in Europe have revealed that more than 80 % of strains isolated from human infections belong to genotype II, whereas genotype I and III are responsible for a small number of cases. Atypical and recombinant strains are generally associated with more severe infections. In Portugal, there is a lack of data concerning genetic diversity, as the classical typing studies in humans have never been performed.

We aimed to determine the *Sag*2 and microsatellite-based (*Tub*2, TgM-A, W35, B17, B18) genotypes of *T. gondii* isolated from humans in Portugal, as well as to study their virulence in mice. We analyzed 48 strains from congenital and acquired toxoplasmosis collected during the last two decades. *Sag*2-based genotyping of *T. gondii* was achieved in all 48 strains where 35 (73 %) were classified as type II and 13 (27 %) were type I. The multilocus PCR of 5 microsatellites allowed the classification of 10 strains (21 %) as recombinant strains that had been previously identified as type II or I by *Sag*2 typing. Seven out of the 48 strains, including three type I, three recombinant and one type II, were virulent in mice. This study constitutes the first evidence of recombinant strains circulating in Portugal in humans from congenital infection, highlighting the need for a better evaluation of these strains as their phenotype is still barely understood.

5.2 Introduction

Toxoplasma gondii is an obligate intracellular protozoan parasite which is responsible for toxoplasmosis in different species of birds and mammals, including humans (Dubey, 2010a). Infections in healthy adults are generally benign, although toxoplasmic retinitis is a frequent cause of serious eye disease in otherwise healthy adults (Gilbert *et al.*, 1999). More profound disease occurs in immunocompromised hosts (Luft and Remington, 1992) or as the result of congenital infections (Hohlfeld *et al.*, 1989). The primary infection during pregnancy is asymptomatic to mother and sometimes to the fetus, however a new born deprived of treatment may develop later symptomatology with recurrent ocular toxoplasmosis, which may lead to blindness and neurological problems during childhood and adolescence (Montoya and Liesenfeld, 2004; Dubey, 2010a; Groër *et al.*, 2011; Tedla *et al.*, 2011; Zhou *et al.*, 2011; Horacek *et al.*, 2012). The determinants of disease severity are not well understood, although a variety of factors including host genotype, nutrition, immune status, infection load and parasite genotype have been suggested to influence the outcome of infection (Sibley *et al.*, 2002).

The distribution of *T. gondii* genotypes varies by geographic region (Lehmann *et al.*, 2006). A highly clonal population structure has been observed in North America and Europe (Fazaeli *et al.*, 2000; Honoré *et al.*, 2000; Ajzenberg *et al.*, 2002b, 2005; Sibley and Ajioka, 2008; Waap *et al.*, 2008), where the three main clonal lineages that are designated types I, II, and III, comprise the vast majority of isolates (Howe and Sibley, 1995; Sibley and Ajioka, 2008; Khan *et al.*, 2011a). Moreover, a fourth clonal lineage ("haplotype 4") has been described as the most common type in wildlife in North America (Dubey *et al.*, 2011a; Rajendran *et al.*, 2012). These findings are based on microsatellite typing, PCR-RFLP, and/or on intron sequencing, and isolates have been usually classified into types II and III clonal genetic lineages (Waap *et al.*, 2008; Jokelainen *et al.*, 2011; Vilares *et al.*, 2014; Verma *et al.*, 2015). In both animal and human samples, atypical and recombinant strains have been rarely reported (Ajzenberg *et al.*, 2004; Bossi and Bricaire, 2004; Lehmann *et al.*, 2006; Cavalcante *et al.*, 2007; Vilares *et al.*, 2014). In the south of Europe (Mediterranean countries), genetic diversity seems to be higher, revealing genotype I, III and recombinant strains (Messaritakis *et al.*, 2008; Mancianti *et al.*, 2013; Vilares *et al.*, 2014; Bacci *et al.*, 2015). Curiously, different studies in humans and animals (namely chickens and dogs) have revealed that strains from South America and Africa seem to be genetically more diverse than in European countries, frequently including recombinant (Dubey *et al.*, 2007e; f) and atypical strains

(Ajzenberg *et al.*, 2004; Lehmann *et al.*, 2006; Cavalcante *et al.*, 2007; Boughattas *et al.*, 2011). Whereas it has been demonstrated that these genotypes are associated with more severe symptoms in immunocompromised hosts, the atypical genotypes are frequently associated with severe toxoplasmosis also in the immunocompetent human patients (Grigg *et al.*, 2001a; b; Carme *et al.*, 2002; Bossi and Bricaire, 2004; Dubey and Su, 2009; Delhaes *et al.*, 2010; Demar *et al.*, 2012).

In Portugal there is a lack of data regarding the *T. gondii* genotype distribution in humans, and this issue gains special relevance since Portugal has a long history of trade and social interaction with Brazil and African countries, involving human and animal migrations and also import/export of food products. In fact, only one study has been published involving human samples (Sousa *et al.*, 2008), but it was based on serotyping (antibodies anti-GRA6) rather than on molecular approaches. That study revealed a majority of type II strains but also an unexpected high prevalence of non-type II strains that could not be typed as type I or III. The existence of recombinant strains was later reported in animals (Vilares et al. 2014). These results suggested the existence of a typing profile different from the one found in the other European countries (Sousa *et al.*, 2008). In the present study, we intended to evaluate the *T. gondii* genotyping distribution and study the virulence in mice of strains from the collection (1994 until 2014) of the Portuguese National Institute of Health.

5.3 Methods

5.3.1 Samples

We studied 48 *T. gondii* strains collected from different sources: from confirmed congenital toxoplasmosis, 16 strains were isolated from 11 placentas (PL), three from umbilical cord blood (UCB), and two from baby blood (BB), whereas from suspected congenital toxoplasmosis, 31 strains were isolated from 20 PL, five from amniotic fluids (AF) and six from UCB. A single strain was isolated from a cerebral biopsy (CB) of a positive HIV with cerebral toxoplasmosis (Figure 5). The 48 strains revealed the following geographic distribution: 66.7 % were collected from the central region of Portugal, 25 % from the north, 6.3 % from the south and 2 % from the Azores islands. All strains (collected since 1994 until 2014) belong to the strain collection of the URSZ-INSA which is the responsible lab in Portugal for the pre- and post-natal diagnosis for the toxoplasmosis.

The inoculation, maintenance and euthanasia of mice were performed according to the

Portuguese standards of the NIH, established for the pre- and post-natal diagnosis of toxoplasmosis, which are in accordance with the Protocol of International Guiding Principles for Biomedical Research Involving Animals as issued by the Council for the International Organizations of Medical Sciences. Briefly, all clinical samples were inoculated by intraperitoneal injection (1 ml / mouse) in female mice [Hsd: ICR (CD-1®); Harlan Ibérica, Barcelona], two each for the AF, BB, UCB and CB and four for PL samples with negative serology to endoparasites. AF were centrifuged and resuspended in 1 ml, and PL were pre treated with trypsine (1:250) and antibiotics (250 UI/ml penicillin and 500 μg / ml streptomycin) before the inoculation. The remainder biological samples were not subjected to any pre-treatment protocol. The inoculated animals were monitored serologically by MAT at 10, 21 and 42 days post-inoculation (p.i.), in order to check the presence of *T. gondii* specific antibodies. At the end of this period, mice were euthanized by inhalation of 5 % halothane and confirmed death by physical examination. The general morphological state of organs was observed, and sample tissues were collected from the brain, and were visualized by optical microscopy in order to search for *T. gondii* cysts. Although tissues cysts may develop in visceral organs, including the lungs, liver, and kidneys, they are more prevalent in the neural tissues, including the brain (Suzuki *et al.*, 2010). Intact tissue cysts are believed not to cause any harm and can persist for the life of the immunocompetent host without causing a host inflammatory response (Dubey, 1998b).

We defined virulence at isolation (without knowledge of infecting dose) based on the mortality of mice within four months of infection and categorized isolates into two groups: virulent (death of 100 % of mice) and non-virulent (< 30 % death). In this work we did not observe moderate virulence (30 % to less than 100 % mortality).

5.3.2 *T. gondii* DNA extraction

T. gondii DNA was extracted from two different mice samples, namely, the ascitic fluid or the brain, depending on the concentration of tachyzoites in microscopy of ascitic fluids. The *T. gondii* DNA extraction from ascitic fluids was performed by boiling one ml of the sample for 30 minutes. From tissue cysts (+/- 1 cm^3), DNA was extracted by using the QIAamp DNA mini kit (Qiagen, Chatsworth, USA) according to the manufacturer's tissue protocol, with slight modifications. These consisted of the homogenization of samples with 200 μl of extraction buffer and 40 μl of proteinase K, and the DNA elution with 50 μl of elution buffer.

5.3.3 Genotype analysis of *T. gondii*

For genotyping purposes, all DNA samples from the inoculated mice were subjected to amplification and sequencing of the two ends of *Sag*2 gene, and to a multiplex PCR of five microsatellites. The later included the beta-tubulin (*Tub*2) and the myosin A (TgM-A) genes as well as three expressed sequence tag (W35, B17, B18) markers (Howe *et al.*, 1997; Ajzenberg *et al.*, 2002b, 2004). For *Sag2*, the PCR products were visualized in GelRED (Biotium Inc., Fremont, USA) stained 2 % agarose gel electrophoresis, purified and sequenced with an ABI 3130xl Genetic Analyzer (Applied Biosystems). For microsatellite analysis, capillary electrophoresis of PCR products was performed on the ABI 3130xl Genetic Analyzer (Applied Biosystems), and data were stored and analyzed with Gene Mapper (version 3.7; Applied Biosystems). MEGA5 software (Tamura *et al.*, 2007) was used to identify the genotypes by comparing the obtained sequences with the ones from reference strains available in GenBank (ME49 - XM_018781602.1, VEG - LN714498.1, RH - AY941252.1 and AY895019.1, RUB - AF357581.1, MAS - AF357580.1).
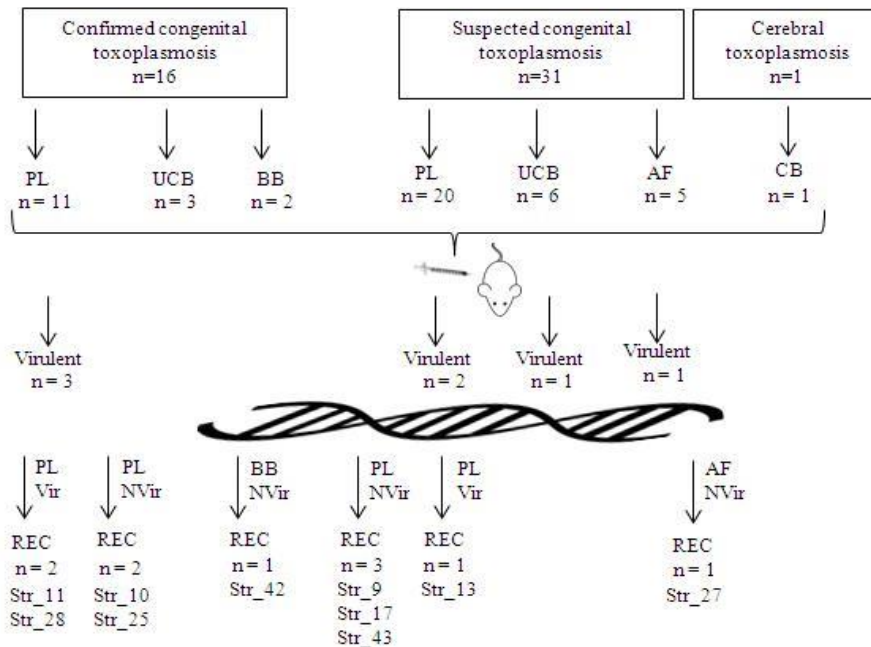
Fig 5.1 – Flowchart of the study. The figure summarizes the mice inoculation results as well as the genotyping data concerning exclusively the recombinant strains. **Note:** These strains are aligned vertically regarding the biological product they were isolated from. PL-placenta; UCB-umbilical cord blood, BB-baby blood, AF-amniotic fluid, CB-cerebral biopsy; Vir- virulent strain; Nvir- non-virulent strain; REC-Recombinant strains.

5.4 Results

5.4.1 Inoculation in mice

Seven (Str_1, Str_3, Str_6, Str_11, Str_13, Str_18, Str_28) out of the 48 strains showed to be virulent in mice (Table 5.1 and Fig 5.1), as they killed the mice after the first inoculation.

Considering the location of tissue cysts is we searched for the presence of cysts in mice brain and found that, except for ten cases, all were positive. Two of the strains (Str_11, Str_3) that did not yield cysts in mice brain were virulent to mice.

5.4.2 Genotypes analysis

*Sag*2 genotyping was achieved in 100 % of samples and allowed the differentiation among the three "classical" *T. gondii* strain types. Thirty-five (73 %) strains were classified as type II, where 23 were isolated from mice inoculated with PL, four with AF, two with BB and six with UCB. Thirteen (27 %) belonged to type I, where eight strains were isolated from mice inoculated with PL, one with AF, one with CB and three with UCB (Table 10). PCR multiplex microsatellites analysis was applied to all samples and was highly efficient. It allowed the identification of thirty five (73 %) strains as type II (25 PL, 3 AF, 1 BB and 6 UCB), six (12.5 %) as type I (1 PL, 1 AF, 1 CB and 3 UCB), one (2 %) as type III (1 PL), and six (12.5 %) as recombinant strains (4 PL, 1 AF and 1 BB). Ten strains (21 %) identified by microsatellites analysis did not confirm the type previously identified by *Sag*2 typing. In fact, from these 10 samples, three were identified as type II (previously classified as type I by *Sag*2), one as type III (previously classified as type I by *Sag*2), three as I/III recombinant strains (previously classified as type I by *Sag*2) and three more as I/II recombinant strains (previously classified as type II by *Sag*2) (Table 5.1 and Fig 5.1).

Molecular and virulence characterization of *Toxoplasma gondii* strains isolated from humans in Portugal

Table 5.1 Virulence and genotyping analysis of *Toxoplasma gondii* strains

Strain number*	Sample"	Mice virulence (death day after the first inoculation)"	Genotyping[+]
Str_1	AF	V (9)	I
Str_2	UCB	NV	I
Str_3	UCB	V (9)	I
Str_4	AF	NV	II
Str_5	AF	NV	II
Str_6	PL	V (11)	I
Str_7	PL	NV	II
Str_8	PL	NV	II
Str_9	PL	NV	REC I/II
Str_10	PL	NV	REC I/III
Str_11	PL	V (7)	REC I/III
Str_12	PL	NV	II
Str_13	PL	V (42)	REC I/III
Str_14	PL	NV	II
Str_15	UCB	NV	II
Str_16	PL	NV	II
Str_17	PL	NV	REC I/II
Str_18	PL	V (120)	II
Str_19	UCB	NV	II
Str_20	PL	NV	II
Str_21	UCB	NV	II
Str_22	PL	NV	II
Str_23	PL	NV	II
Str_24	UCB	NV	II
Str_25	PL	NV	REC I/II
Str_26	PL	NV	II
Str_27	AF	NV	REC I/II
Str_28	PL	V (10)	REC I/II
Str_29	PL	NV	II
Str_30	AF	NV	II
Str_31	CB	NV	I
Str_32	PL	NV	II
Str_33	UCB	NV	II
Str_34	UCB	NV	I
Str_35	PL	NV	II
Str_36	UCB	NV	II
Str_37	PL	NV	II

Molecular and virulence characterization of *Toxoplasma gondii* strains isolated from humans in Portugal

Strain number*	Sample"	Mice virulence (death day after the first inoculation)"	Genotyping[+]
Str_38	PL	NV	II
Str_39	PL	NV	II
Str_40	PL	NV	II
Str_41	PL	NV	II
Str_42	BB	NV	REC I/II
Str_43	PL	NV	REC I/II or II/III
Str_44	PL	NV	II
Str_45	PL	NV	II
Str_46	PL	NV	II
Str_47	PL	NV	II
Str_48	BB	NV	II

* Strains were numbered solely for simplification purposes considering their complex laboratory designation. "AF - amniotic fluid, UCB - Umbilical cordon blood, PL – Placenta, CB – Cerebral biopsy, BB – Baby blood. " NV-Non virulent; V-Virulent. [+] REC – Recombinant.

5.5 Discussion

In Portugal, as in most European countries, *T. gondii* isolates collected from animals (which enrol the majority of the studies) with chronic infections have shown remarkably little diversity, and type I or recombinant strains have been rarely found (Herrmann *et al.*, 2010, 2012a, 2013; Vilares *et al.*, 2014). Concerning strains collected from human beings, type II predominates in Europe (Howe *et al.*, 1997; Ajzenberg *et al.*, 2005), mirroring the scenario found in animals. However, there is a lack of knowledge of the genetic diversity of *T. gondii* strains from human samples in Portugal. The only previously published data concerns the report of type II strains in two HIV seropositive individuals (Ajzenberg *et al.*, 2009). In the present study, we enrolled 48 *T. gondii* strains collected throughout about two decades, and observed that 67 % (32/48) of strains were classified as type II simultaneously by *Sag*2 and microsatellites procedures, which was quite similar to the genotypes distribution found in other European and USA studies (Sibley *et al.*, 2009; Herrmann *et al.*, 2014). The results of multilocus genotyping were in agreement with the *Sag*2 genotyping for 38/48 isolates. The ten discrepant strains reinforce the interest of using a multilocus approach. Particularly, seven *Sag*2 type I isolates exhibited a mixture of type I/III or I/II alleles and three type II isolates exhibited a mixture of type I/II, being considered as recombinant strains. In this regard, the

Molecular and virulence characterization of *Toxoplasma gondii* strains isolated from humans in Portugal

microsatellites typing method, which uses one multiplex PCR to perform multilocus typing with five markers (Ajzenberg *et al.*, 2005) was an efficient and essential tool for the identification of a molecular epidemiology scenario where 21 % (10/48) of the investigated strains were found to be recombinant strains. These were described in Portugal for the first time in 2014 (Vilares *et al.*, 2014). We speculate that the higher proportion of type I and recombinant isolates in the human Portuguese sampling compared to other European studies (Howe *et al.*, 1997; Nowakowska *et al.*, 2006; Ajzenberg *et al.*, 2010) may be explained by the fact that Portugal has a long history of trade and social interaction with South American countries (such as Brazil) and China (where atypical and recombinant strains are not uncommon), which involve human migrations, the import of food products (like meat or vegetables with the potential of transmitting the parasite), and also the presence of rats and cats (*T. gondii* intermediate and definitive hosts, respectively) in trading ships.

Concerning the inoculation in mice, the virulence rate *in vivo* was quite low, 15 % (7/48), which is similar to the one observed in other countries (Ajzenberg *et al.*, 2002b; Gebremedhin *et al.*, 2014). Two recombinant type I/III (Str_11, Str_13) and one recombinant type I/II (Str_28) were lethal to mice at the first inoculation (Fig 5.1) and two of these (Str_11, Str_28) had the ability of transplacental transmission to the fetus. Unexpectedly, two type I strains were not lethal to mice, and mice death inoculated with one type II strain was observed. One could speculate that these strains may not be typical type I and II strains but instead recombinant strains that could not be identified with the limited number of molecular markers used. Another explanation for the observed virulence of the type II strain could be an elevated parasite load in the inoculum. In fact, the biological samples are inoculated in mice according to the standard clinic diagnostic procedures, preventing the *a priori* determination of the number of parasites.

As concluding remarks, this study presents the first data in Portugal concerning *T. gondii* genotyping of human samples, and reveals the existence of a considerable proportion of recombinant strains. The finding of genetic variations in the predominant clonal lineages may have important implications for public health as, in the future, drug resistance or immune evasion in these lineages must be acquired and they could be responsible for emergent toxoplasmosis in different countries. It is thus imperative to better characterize the genetic background of these unusual strains, in order to gain more insight into genotype - phenotype associations.

CHAPTER VI

Towards a rapid sequencing-based molecular surveillance and mosaicism investigation of *Toxoplasma gondii*

Personal contribution

Anabela Vilares contributed to the design of the study, performed all most the experimental work, performed part of bioinformatics analyses, interpreted data and wrote "part" of the manuscript.

6.1 Abstract

Advances in molecular epidemiology of *Toxoplasma gondii* are hampered by technical and cost-associated hurdles underlying the acquisition of genomic data from parasites. In order to implement an enhanced genotyping approach for molecular surveillance of *T. gondii*, we applied a multi-locus amplicon-based sequencing strategy (targeting genome-dispersed polymorphic loci potentially involved in adaptation and virulence) to samples associated with human infection. This approach enabled us not only to genetically discriminate almost all 68 studied strains, but also to reveal a scenario of marked genomic mosaicism. Two-thirds (n=43) of all strains could be classified as recombinant, although recombination seemed to be linked to the classical archetypal lineage. While 92 % of the *Sag*2 archetype I strains revealed genetic mosaicism, only 45 % of *Sag*2 archetype II strains could be identified as recombinant. Contrarily to the virulence-associated archetype I, most type II strains (regardless their recombination background) were non-virulent in mouse. Besides *Sag*2, some of the newly studied loci (namely the Type I/I-like alleles of *Sag*1, B17, PK1 and *Sag*3 and Type III/III-like alleles of TgM-A) constitute promising candidates to rapidly infer *T. gondii* mouse virulence. Our successful attempt to capture microsatellite length variation launches good perspectives that the laborious intensive historical method can be straightforwardly transfer to more informative next-generation sequencing (NGS)/bioinformatics-based methodologies. Overall, while *T. gondii* whole-genome sequencing will be hardly feasible in most laboratories, this study shows that a discrete loci panel has the potential to improve the molecular epidemiology of *T. gondii* towards a better monitoring of circulating genotypes with clinical importance.

6.2 Introduction

Toxoplasma gondii is a widespread parasite of animals that causes opportunistic infection in humans (Carruthers, 2002). It has two distinct reproductive mechanisms: (a) asexual reproduction in intermediate hosts, in which tissue cysts containing haploid parasites infect other hosts through carnivorism and (b) sexual reproduction in definitive hosts (felines), shedding millions of thick-walled cysts known as oocysts (Dubey, 1995). Oocysts undergo meiosis in the environment to form eight haploid sporozoites that are highly infectious to a variety of hosts (Elmore *et al.*, 2010). Because a single parasite can produce both micro- and macrogametes and self-mate progeny, novel recombined genotypes can only be generated in the rare events in which a feline is infected simultaneously with two or more strains (Sibley and Ajioka, 2008).

Studies on the population genetic structure of *T. gondii* have revealed complex patterns that differ among geographic regions (Sibley and Ajioka, 2008). Early analyses showed that the majority of strains isolated in Europe and North America belong to three clonal haplotypes. These are types I, II, and III (clonal or archetypal lineages), as defined by the traditional analysis of RFLP recognition sites of the 5' and 3' ends of *Sag*2 locus (Dubey *et al.*, 2012), being the type II the predominant lineage causing human toxoplasmosis (Hosseini *et al.*, 2019). A fourth clonal lineage, prevalent in North America wild animals, from haplogroup (HG) 12, was recently described based on the diversity of mutational profiles found in five introns and *Sag*1, *Gra*6 and *Gra*7 (Khan *et al.*, 2011a). Strains with atypical or novel allele combinations, based in different markers, such as the chromosome Ia and several discrete loci (M6, M33, M48, M95, M163, M102) have also been isolated (Lehmann *et al.*, 2006; Khan *et al.*, 2007; Pena *et al.*, 2008). Some studies revealed a marked distinctive trend between the Northern Hemisphere, where a few, highly clonal and abundant lineages predominate (most from type II) and South America countries, where the diversity seems to show greater evidence of recombination. These strains isolated in South America date to a much older time period and it is believed that they have largely remained genetically isolated from those reported in the North America (Su *et al.*, 2012). Noteworthy, type I or type I-like atypical isolates seem to be involved in severe retinochoroiditis in humans (Grigg *et al.*, 2001a) and they are often associated with severe acute disseminated toxoplasmosis in immunocompetent patients (Bossi and Bricaire, 2004). Recently, although the emergence of new haplotypes has been described, Minot and colleagues pointed out that the

observed extent of *Toxoplasma* genetic diversity does not seem to fit the scenario of the proposed haplotypes (Minot *et al.*, 2012).The authors argued that observed extent of *Toxoplasma* genetic diversity cannot be explained by the mating of four to six ancestral strains, instead, most strains appear to have formed through recent recombination events (Minot *et al.*, 2012).

In Portugal, there is still an important lack of knowledge about the genotypes diversity of circulating strains in humans. In fact, besides a first report of microsatellites-based genotyping of two isolates (from HIV-positive patients) in 2009 (Ajzenberg *et al.*, 2009), only in 2017 it has been reported the application of classic *Sag*2-based genotyping and microsatellites size analyses to 48 strains from congenital and acquired toxoplasmosis (Vilares *et al.*, 2017). This recent study reported the first recombinant strains isolated from congenital toxoplasmosis circulating in Portugal, and suggest that large-scale genotyping studies focused on chromosome-dispersed analyses may be important for a better understanding of genotype-phenotype associations and degree of genetic mosaicism.

In the last decade, with advances in the high-throughput sequencing technologies, as well as the implementation of a genome and functional genomic database (Gajria *et al.*, 2008), it has been possible not only to characterize the genome of the three archetypes of *T. gondii* strains (type I/GT1, II/ ME49, and III/VEG), but also to compare the genome of dozens of globally distributed *T. gondii* isolates (Lorenzi *et al.*, 2016). In general, these genome-scale studies highlighted both the complexity of the evolutionary and population dynamics of *T. gondii* and the remarkable mosaicism of its genome, which is shaped by mutational events as diverse as SNP-based diversification, tandem amplification, or large haploblock inheritance. Still, despite these increasing efforts to decode *T. gondii* genetic diversity, a few important methodological hindrances still hamper our ability to investigate which specific genotypic traits are linked to differential *T. gondii* transmission, host range and pathogenicity patterns, such as: *i)* current genotyping methods offer insufficient resolution power to get insight on the *T. gondii* evolutionary dynamics, namely to capture the recombination events; *ii)* whole-genome sequencing (WGS) technologies are not available for less resourced laboratories, and even for labs with such resources WGS of *T. gondii* constitutes a cost-prohibitive and bioinformatically defiant task due to its up to 60 Mb genome size spread by 14 chromosomes; and *iii)* there is a lack of harmonized, standardized, reproducible methods (and loci panels) to allow reliable inter-

laboratory comparisons. In the present study, we combined classical *Sag*2- and microsatellites-based typing with a novel amplicon-based sequencing approach targeting several genome dispersed loci potentially involved in adaptation and virulence to more profoundly genetically characterize strains isolated from humans. While this study constitutes an important first attempt to transit to a more discriminatory NGS-based *T. gondii* genotyping, it contributes to a better knowledge about the diversity and genetic mosaicism of *T. gondii*.

6.3 Methods

6.3.1 Strains

The Portuguese National Institute of Health (NIH) holds the largest collection of *T. gondii* strains isolated from humans in Portugal (64 strains collected from 1994 onwards). In this study, we were able to study 52 strains, including 48 isolates that had been genotyped by *Sag*2 (archetypal types) and microsatellites in 2017 (Vilares *et al.*, 2017), as well as three isolates (Str_49, Str_50 and Str_51) and one reference RH strain (maintained for several years in the laboratory) for which traditional genotyping (*Sag*2 and 5 microsatellites) was performed in the present study, as previously described (Vilares *et al.*, 2017).

From the total of 51 selected *T. gondii* clinical strains, 17 strains are associated with confirmed congenital toxoplasmosis (CT) [12 placentas (PL), 3 umbilical cord blood (UCB), and 2 baby blood (BB)], 33 with suspected CT [22 PL, 5 amniotic fluids (AF), 6 UCB], and one strain was collected from cerebral biopsy (CB) from a HIV-positive patient. Most isolates, 70 % (36/51), are from the central region of Portugal, 24 % (12/51) from the north, 4 % (2/51) from the south and 2 % (1/51) from Azores islands (Supplementary Table S1). Strains were also characterized regarding their virulence trait in laboratory mice, and minimal data was also collected regarding the associated clinical manifestations in humans (Supplementary Table S1) (Vilares *et al.*, 2017). The virulence in mice was defined (without knowledge of infecting dose) based on the mortality of mice within four months of infection and the isolates were categorized into three groups: "virulent" (death of 100 % of mice); "moderately virulent" (30 % to less than 100 % mortality) and "not virulent" (< 30 % death) (Supplementary Table S1). DNA samples

were extracted from the 51 selected *T. gondii* clinical strains after a single passage in mice (mice inoculated with human product) using commercial QIAamp DNA Mini Kit (Qiagen, UK).

6.3.2 PCR, Sanger sequencing and Next-generation Sequencing

PCR amplification of seven loci (CB21-4, PK1, L363, *Sag*1, *Gra*6, *Sag*3, M102) located in different chromosomes (Chr III, Chr VI, Chr VIIb, Chr VIII, Chr X, Chr XII, Chr VIIa) (http://toxomap.wustl.edu/verticalmap08_01-2005high.jpg) were performed to generate high quality amplicons for subsequent Sanger sequencing or high-throughput amplicon-based NGS (PCR primers are described in Supplementary Table S2). These loci were selected based on the following criteria: *i)* being encoded in different chromosomes (and inserted in phenotypically relevant haploblocks); *ii)* being polymorphic among representative strains of the classical types I, II, III; *iii)* presenting literature support for their potential involvement in adaptation/virulence (Khan *et al.*, 2005a; de Melo Ferreira *et al.*, 2006; Ajzenberg *et al.*, 2010; Su *et al.*, 2010; Lauron *et al.*, 2015).

PCR products were visualized in GelRED (Biotarget, Lisbon, Portugal) stained 2 % agarose gel electrophoresis, purified and sequenced in ABI 3130xl Genetic Analyzer (Applied Biosystems). To perform NGS, purified PCR products from each strain were pooled and purified using Agencourt AMPure XP PCR Purification kit, before proceeding to Nextera XT DNA Library Preparation (Illumina Inc, San Diego, CA, USA), according to manufacture instructions. Libraries were subsequently sequenced (2 x 150 bp paired-end reads) using a MiSeq (Illumina) equipment available at the Portuguese NIH.

6.3.3 Bioinformatics analyses

Analysis of NGS data was conducted using the mapping-based bioinformatics pipeline implemented in INSaFLU (https://insaflu.insa.pt/), which is a web-based platform for amplicon-based NGS data analysis (Borges *et al.*, 2018). Briefly, the core bioinformatics steps (Borges *et al.*, 2018) involved: i) raw NGS reads quality analysis and improvement using FastQC v. 0.11.5; (https://www.bioinformatics.babraham.ac.uk/projects/fastqc) and Trimmomatic v. 0.27 (http://www.usadellab.org/cms/index.php?page=trimmomatic), respectively; and, ii) reference-based mapping, consensus generation and variant detection using the multisoftware tool Snippy v. 3.2-dev (https://github.com/tseemann/snippy), using a multi-FASTA file with representative

sequences of a type II strain of each one of the target amplicons as reference sequence. Single Nucleotide Polymorphisms (SNPs) and indels were inspected and confirmed through visual inspection using the Integrative Genomics Viewer (http://www.broadinstitute.org/igv). Consensus sequences were exclusively validated for locus displaying 100 % of their size covered by at least 10-fold of depth of coverage (the mean depth of coverage per sample for validated amplicons ranged from ~973 to 7341-fold) (Supplementary Table S3). All raw sequence reads used in the present study were deposited in the European Nucleotide Archive (ENA) (BioProject PRJEB32939). Detailed ENA accession numbers are described in Supplementary Table S3.

Alignment of consensus sequences *per* locus was subsequently performed using MAFFT v. 7.313 (https://mafft.cbrc.jp/alignment/software/). In order to better integrate the novel sequence data on the available *T. gondii* genome diversity, we analyze 14 additional reference strains for which the genome is publicly available at ToxoDB platform (Supplementary Table S4) and RH strain maintained in the laboratory (Supplementary Table S1). MEGA 7.0 (https://www.megasoftware.net/) was further applied to alignment visual inspection/refinement. Refined consensus nucleotide alignments used for SNP/indel mapping are available at https://doi.org/10.5281/zenodo.3244024. In order to better visualise the genetic mosaicism and better link it to both the *Sag*2 archetypal types and the profiles of the newly sequenced loci, a Neighbor-joining phylogenetic tree was generated for *Sag*2 refined alignment (upon concatenation of the two sequenced regions at both 5' and 3' ends) using the Kimura 2-parameter model (1000 bootstraps) using MEGA. Of note, *Sag*2 was selected since: i) it is the traditional gene used for the archetypal classification in three main types; ii) sequence data was obtained for this locus for all isolates; and iii) it revealed the highest discriminatory power among the studied genes. Also, to ensure that the SNP/indel genetic diversity found for *Sag*2 was truly represented in the tree, 1bp-indels were artificially replaced by single bases, thus counting as SNPs.

6.3.4 Allele nomenclature

Each locus-specific refined nucleotide alignment was manually converted in a MLST-like allelic matrix (Supplementary Table S5A), i.e., numerical allele identifiers were attributed for each unique sequence profile found for each locus (numerical classification). "1", "2" and "3" allele identifiers were attributed for profiles 100 % matching the profiles observed for representative strains of the three classical types GT1 (I), Me49 (II) and VEG (III), respectively,

while the remaining identifiers were attributed sequentially (differential microsatellites lengths found for M102 and CB21-4 were not account for this numerical classification - see the last section of Material and Methods section). The MLST-like matrix was further used to infer the resolution power provided by each locus or combination of loci, by estimating Simpson's Index of Diversity (ID) through the web-based framework "Comparing partitions" (http://www.comparingpartitions.info/) (Carriço *et al.*, 2006).

Each locus sequence was additionally subjected to nominal classification. Allele profiles other than I, II and III (corresponding to allele numbers 1, 2 and 3, respectively) were inspected and nominally classified as Type I-like, Type II-like, Type III-like if they harbor SNPs/indels markers of a given classical type and not other type-specific markers from other types (even though they may not enroll all markers from that type and/or may present additional SNPs/indels that are not observed in any type reference strains) (see Supplementary Table S5B). A few cases where these criteria could not be fulfilled were indicated as "not classified" (including intriguing rare cases where the locus harbors SNP and/or indels "specific" of different classical types – these were labeled as "IntraRec"). SNP/indel profiles detected by Sanger/NGS that were used for numerical and nominal classification can be found in detail in Supplementary Table S6. Microsatellite-containing loci were additionally classified when the number of tandem repeats within the microsatellites does not match the number described for the three archetypes reference strains, as historically performed. For the loci enrolled in traditional microsatellites-based analyses (B18, B17, TgM-A, *Tub*2 and W35), the designation "atypical" was kept. For the newly sequencing loci also found to harbor microsatellites (M102 and CB21-4; see the last section of Material and Methods section), we opted by using the "atypical-like" label to avoid confusion with the former.

6.3.5 Strain nomenclature

Strains were classified as "I", "II" and "III" when they presented a "clonal" signature, i.e., when all loci were from the same type or type-like. In an opposite scenario, strains were classified as "interlocus recombinants" when they simultaneously harbored profiles matching more than one classical type or type-like (locus-specific "atypical" or "atypical-like" profiles were not considered for this classification, although they were identified) (see details in Supplementary Table S5B). To assess the gain introduced by newly sequenced loci in the

identification of interlocus recombinant strains, strains were first classified solely based on the traditional *Sag*2- and microsatellites-based analysis, and then, based on the combination of all loci under analysis. In order to better evaluate interlocus recombination (see next section), the nominal allele and strain classification were converted into an intuitive color-coded representation using Microreact (https://microreact.org/showcase). Finally, the web-based framework "Comparing partitions" (http://www.comparingpartitions.info/) was also applied to measure the concordance between locus-specific classification and mouse virulence through the Adjusted Rand coefficient (Carriço *et al.*, 2006).

6.3.6 Intra-host heterogeneity targeting tandem repeats of microsatellites

In order to capture the most abundant repeat number profile within the microsatellites-containing loci subjected to NGS (M102 and CB21-4), allowing the assessment of potential intra-patient sub-population diversity, we applied a previously described Python script (Pinto *et al.*, 2016) directly to NGS reads (after quality improvement with Trimmomatic). This strategy consists on the extraction and counting (directly from reads, both forward and reverse) DNA sequences (i.e., the microsatellites sequences) that are flanked by two conserved, small DNA strings. The defined strings for CB21-4 correspond to the conserved 10bp (alignment position "214-CATACAATAT-223" and "257-GTACATTCTT-266") contiguously flanking each side of the tandem repeat region (the repeat unit is "AT") (alignment available at https://doi.org/10.5281/zenodo.3244024). For M102, since the microsatellite (tandem repeat of the "AT" unit) is not contiguously flanked by conserved regions for all three "types" (Supplementary Table S6), the strings (alignment position "38-CACATATATAC-48" and "98-CATATATATTTT-108") (alignment available at https://doi.org/10.5281/zenodo.3244024) were defined in order to capture both the tandem repeat and the variable flanking regions. After sequence capture, sequence profiles that do not match any of expected tandem repeats profiles where excluded, even though the proportion of every single erroneous profile did not exceed 5 % of the total counts obtained for each sample. Subsequently, considering the expectedly low population size of *T. gondii* in context of human infection, only validated profiles that account for at least 5 % of the filtered counts for each sample are presented. For these two loci, sequences representing the most abundant profiles were inserted in the respective alignment of consensus sequences (available at https://doi.org/10.5281/zenodo.3244024; Supplementary Table S6).

6.4 Results

6.4.1 Analysis of the discriminatory power by locus and genotyping strategy

Genotyping data was collected for 13 loci (dispersed in 9 chromosomes) (Table 6.1) for 68 *T. gondii* human strains in order to get insight into the genetic diversity among *T. gondii* human isolates (the strains for which genotypic data could be obtained for each loci is detailed in Supplementary Table S5A). This study enrolls 52 strains, which represent about 80 % of the *T. gondii* collection of the Portuguese NIH (including 51 clinical strains and the reference strain "RH") (Supplementary Table S1), as well as 16 additional strains (Supplementary Table S4) showing diverse classification (I, II, III and atypical) and for which sequence data could be retrieved from public databases for the studied loci (Lorenzi *et al.*, 2016). In addition, it includes data generated by three distinct genotyping approaches: detection of classical *Sag*2 ends "RFLP" recognition sites (Classical genotyping) (Howe *et al.*, 1997), microsatellites size analysis and nucleotide sequencing methods (Sanger sequencing and NGS) (Table 6.1).

Table 6.1 *T. gondii* loci analyzed in the present study.

Chr	Locus	Number of strains	Number of types [a]	Methodology	Size of the target sequence
VIII	*Sag* 2 5' and 3' ends	68	3	Sanger Sequencing (recognition sites) [b]	359 (211 plus 148)[c]
VIIa	**B18**	68	3	Microsatellites[d]	158-160 [e]
II	**W35**	68	3	Microsatellites[d]	242-248 [e]
IX	*Tub* 2	68	3	Microsatellites[d]	289-291 [e]
XII	**B17**	68	3	Microsatellites[d]	336-342 [e]
X	**TgM-A**	68	4	Microsatellites[d]	204-209 [e]
VIII	*Sag* 1	32	3	Nucleotide sequencing (Sanger and/or NGS)	199
X	*Gra* 6	65	10	Nucleotide sequencing (Sanger and/or NGS)	234
VI	PK1	50	10	Nucleotide sequencing (Sanger and/or NGS)	734
XII	*Sag* 3	65	11	Nucleotide sequencing (Sanger and/or NGS)	148
VIIb	L363	39	13	Nucleotide sequencing (Sanger and/or NGS)	746
III	CB21-4	61	9	Nucleotide sequencing (Sanger and/or NGS)	310-324[e]
VIIa	M102	59	9	Nucleotide sequencing (Sanger and/or NGS)	78-104[e]
VIII	*Sag* 2 5' and 3' ends	68	23	Nucleotide sequencing (Sanger and or NGS)[c]	359 (211 plus 148)[c]

[a] Number of types obtained upon MLST-like numerical allele classification (see details in Supplementary Table S5). [b] Detection of the traditional RFLP recognition sites (archetypal) after Sanger sequencing. [c] 211bp from 5' *Sag*2 end and 148bp from 3' *Sag*2 end (analysis of nucleotide diversity refers to the concatenation of these two regions). [d] Detection of the microsatellites length by Sanger Sequencing (diversity within microsatellites was not taken into account for allele numerical classification). [e] Size range is due to small indelks and/or the differential size of the microsatellites.

Classical genotyping of *Sag*2 5' and 3' ends allowed the differentiation of the three "classical" *T. gondii* strain types among the 68 studied strains (I, 36.8 %; II, 58.8 %; III, 4.4 %) (Supplementary Table S5 and S7). For the 51 clinical isolates from Portugal (thereafter designated "PT" isolates), Type II was predominant (71 %; n = 36) followed by type I (29 %; n = 15). No type III strains were identified by this *Sag*2-based classical approach. A similarly low discriminatory level was reached by traditional microsatellites-based analyses, with no more than 4 genotypes being differentiated *per* locus (Table 6.1). In contrast, the genetic analysis of additional 8 loci by nucleotide sequencing-based methods (Sanger or NGS) considerably increased the resolution power, reaching a maximum of 23 distinct allelic types out of 68 strains (20 types among the 51 PT isolates) for *Sag*2 loci (Figure 6.1) (Table 6.1; Supplementary Table S5). As previously described (Vilares *et al.*, 2017), virulent strains in mice mainly belong to the "classical" type I (Figure 6.1). Of note, classification of genetic profiles (alleles) observed for each locus as I (allele identifier "1"), II (allele identifier "2"), and III (allele identifier "3"), based on matching profiles observed for representative strains (GT1, Me49 and VEG, respectively) revealed that type II was the most frequently observed in 7 out of the 13 loci (relative frequencies of allele "2" per locus ranged from 32 to 65 %) (Supplementary Table S7). The dominant profile for the other 6 loci (W35, *Tub*2, B17, *Sag*1, CB21-4, *Sag*3) was a non-matching profile (i.e., a profile other than the 3 "archetypes"). In particular, for W35, Tub2, B17 and *Sag*1, the dominant profile was shared by all type II and Type III strains, which could not be discriminated by these four loci. Noteworthy, type I was "top 2" in ten loci, including five loci where type II was dominant. In an opposite scenario, type III strains revealed a low frequency (< 6 %) regardless of the loci under analysis, with exception of TgM-A and *Sag*3 where they represented about 16 % of the strains analysed.

Towards a rapid sequencing-based molecular surveillance and mosaicism investigation of *Toxoplasma gondii*
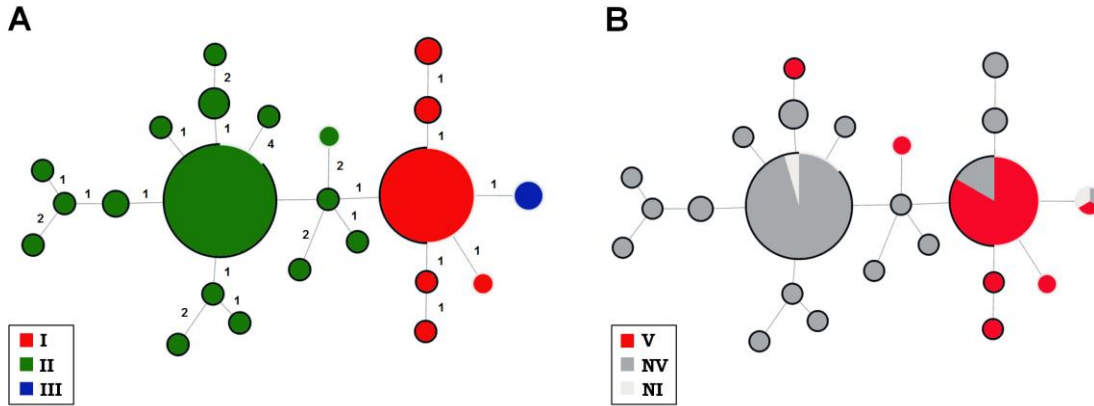
Figure 6.1. Minimum spanning tree (MST) of the 68 *T. gondii* isolates based on *Sag*2 5' and 3' ends nucleotide diversity colored according to classical genotyping classification (A) and mouse virulence (B). **Note:** The MST was constructed using the goeBURST algorithm implemented in the PHYLOViZ Online platform based on a concatenated nucleotide alignment enrolling the two *Sag* 5' and 3' ends for *Sag*2 locus. Each circle represents a unique SNP/indel profile and the numbers on the connecting lines in panel A represent the number of SNP/indel differences between strains. The black outer rings reflect the proportion of PT isolates with each profile. V, virulent; NV, non-virulent, NI, no information available.

We also performed a deeper evaluation of the discriminatory power (evaluated by Simpson's Index of Diversity) conferred by each locus individually and by the combination of loci analysed by the same genotyping strategy (Figure 6.2).
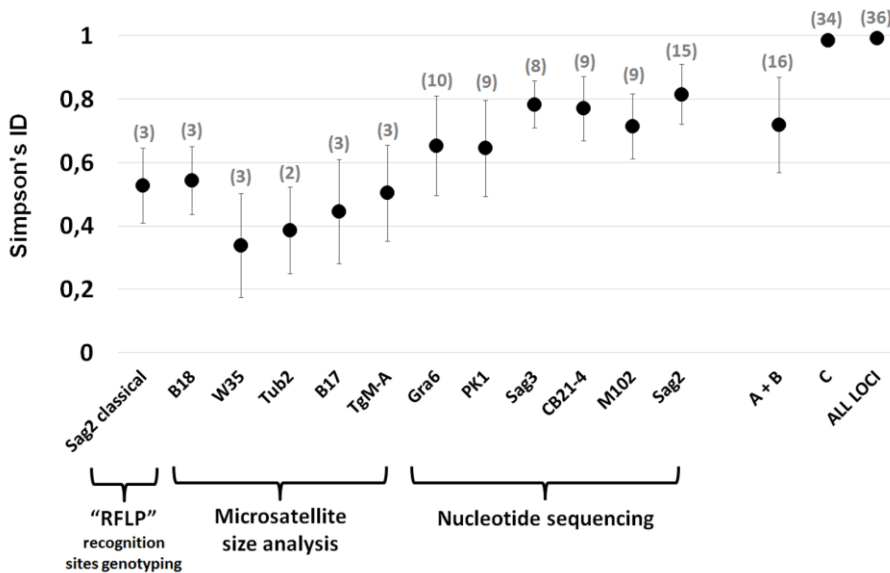
Figure 6.2. Discriminatory power provided by each locus or combination of loci. **Note:** The graph displays the discriminatory power (based on the Simpson's Index of Diversity, with 95 % confidence intervals) and number of different genotypes (numbers above the dots) discriminated using allelic profiles of each single locus or combination of loci (organized by genotyping strategy). This comparison exclusively enrolls strains (n = 40) for which data could be collected for all loci with exception of L363 and *Sag*1 (these loci were excluded since experimental success was considerably low).

Considering that no data could be obtained for all loci and for all strains (Table 6.1; Supplementary Tables S3 and S5), this analysis enrols 40 strains (including 25 "PT" clinical strains) (Supplementary Table S5) for which data could be collected for all loci with exception of L363 and *Sag*1 (these loci were excluded since experimental success was considerably low). As expected, both the classical genotyping based on *Sag*2 ends recognition sites and the microsatellites-based strategies conferred low discriminatory resolution, with up to three types being discriminated per locus. Still, it is noteworthy that the combination of these profiles allowed the discrimination of 16 different types among the subset of 40 isolates. Similarly, despite the six newly sequenced loci proved to be polymorphic (8-15 alleles found among the 40 strains), a higher gain in resolution was obtained when combining their profiles. In fact, their concatenation yielded 34 different types (Simpson's ID of 0.985; 95 % confidence interval of 0.964-1.000), which is very close to the maximum resolution reached when combining all profiles (36 types; Simpson's ID of 0.991; 95 % confidence interval of 0.977-1.000) (Figure 6.3). Of note, the minimum spanning tree based on concatenated profiles showed a good level of concordance not only with both *Sag*2 classical classification and sequencing-derived alleles, but also with mouse virulence profiles.

Towards a rapid sequencing-based molecular surveillance and mosaicism investigation of *Toxoplasma gondii*
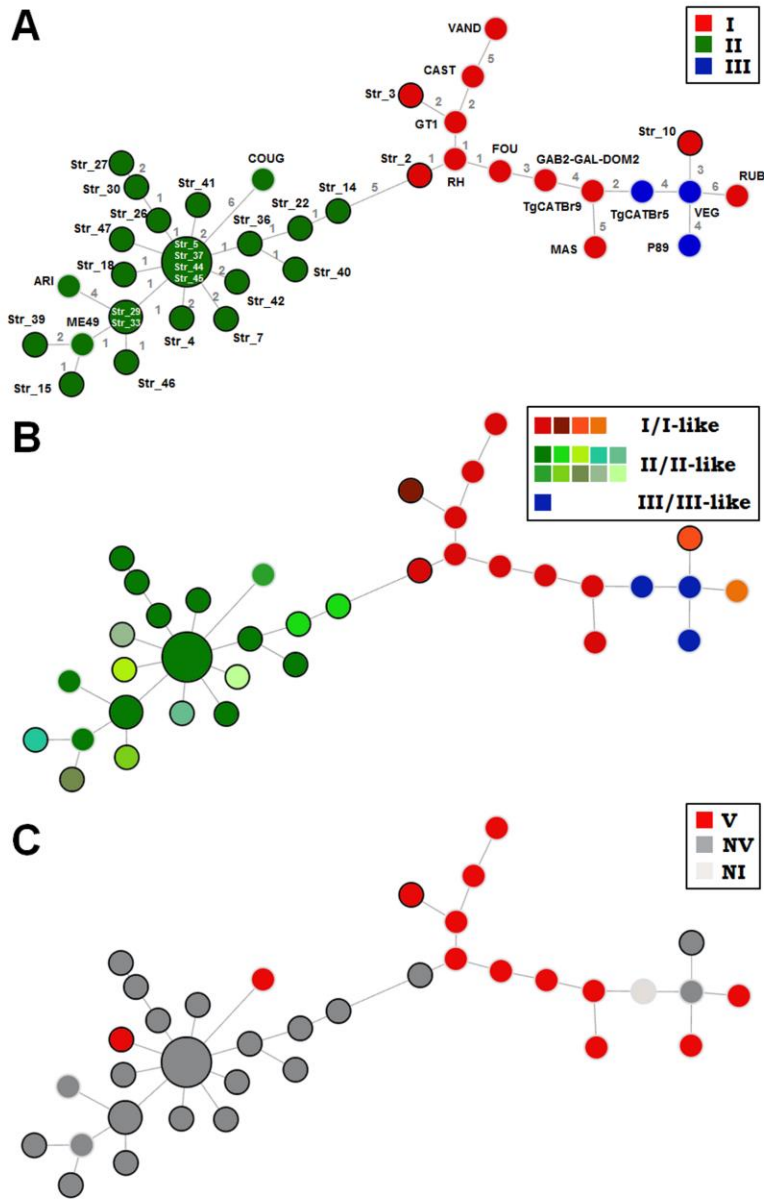
Figure 6.3. Minimum spanning tree (MST) based on the multi-loci-based combined allelic profiles colored according to classical *Sag*2 archetypes (A), *Sag*2 sequence alelles (B) and mouse virulence (C). **Note:** MST was constructed using the goeBURST algorithm implemented in the PHYLOViZ Online platform, and is based on combination of allelic profiles obtained for all studied loci (Supplementary Table S5A), with exception of L363 and *Sag*1 (these loci were excluded since experimental success was considerably low). It enrolls 40 strains for which data could be collected for all loci. Each circle represents a unique combined allelic profile and the numbers on the connecting lines in panel A represent the allele differences between strains. The black outer rings reflect the proportion of PT isolates with each profile. Strain's designation is detailed in panel A. V, virulent; NV, non-virulent, NI, no information available.

6.4.2 Interlocus recombination

While the strategies focused on the diversity at nucleotide level clearly improved the resolution for *T. gondii* genotyping, their integration with the traditional genotyping data consolidated the existence of a scenario of remarkable interlocus recombination (i.e., the same strain simultaneously harbored loci with profiles matching more than one classical type or type-like) (Figure 6.4). In fact, while the combination of the classical *Sag*2 genotyping and the microsatellites approach alone would identify 24 out of 65 strains (36.9 %) as interlocus recombinant ("parental" strains assumed as "clonal" are not taken into account for counting), the evaluation of all loci under study allowed the identification of a total of 43 (66.2 %) recombinant strains, including 29 out of the 51 (56.9 %) strains isolated in Portugal (Table 6.2) (detailed classification can be found in Supplementary Table S5B). The most marked gain was observed for type I, as seven out of the eight type I strains classified by the traditional approach (combination of *Sag*2 and microsatellites) were indeed found to be recombinant when using the extended approach. Of note, although data could not be detected for all loci for all strains, the loci involved in this shift were essentially *Gra*6 and *Sag*3 – each one exchanged in at least four out of the seven recombinant strains. Regarding type II, 11 out of the 32 strains displaying type II/II-like profiles (combining *Sag*2 and microsatellites) could be afterwards identified as recombinant strains I/II (Table 6.2). The mostly affected loci here were M102, CB21-4 and *Gra*6 (each being responsible for shifts in at least three strains) (Supplementary Table S5B).

Towards a rapid sequencing-based molecular surveillance and mosaicism investigation of *Toxoplasma gondii*
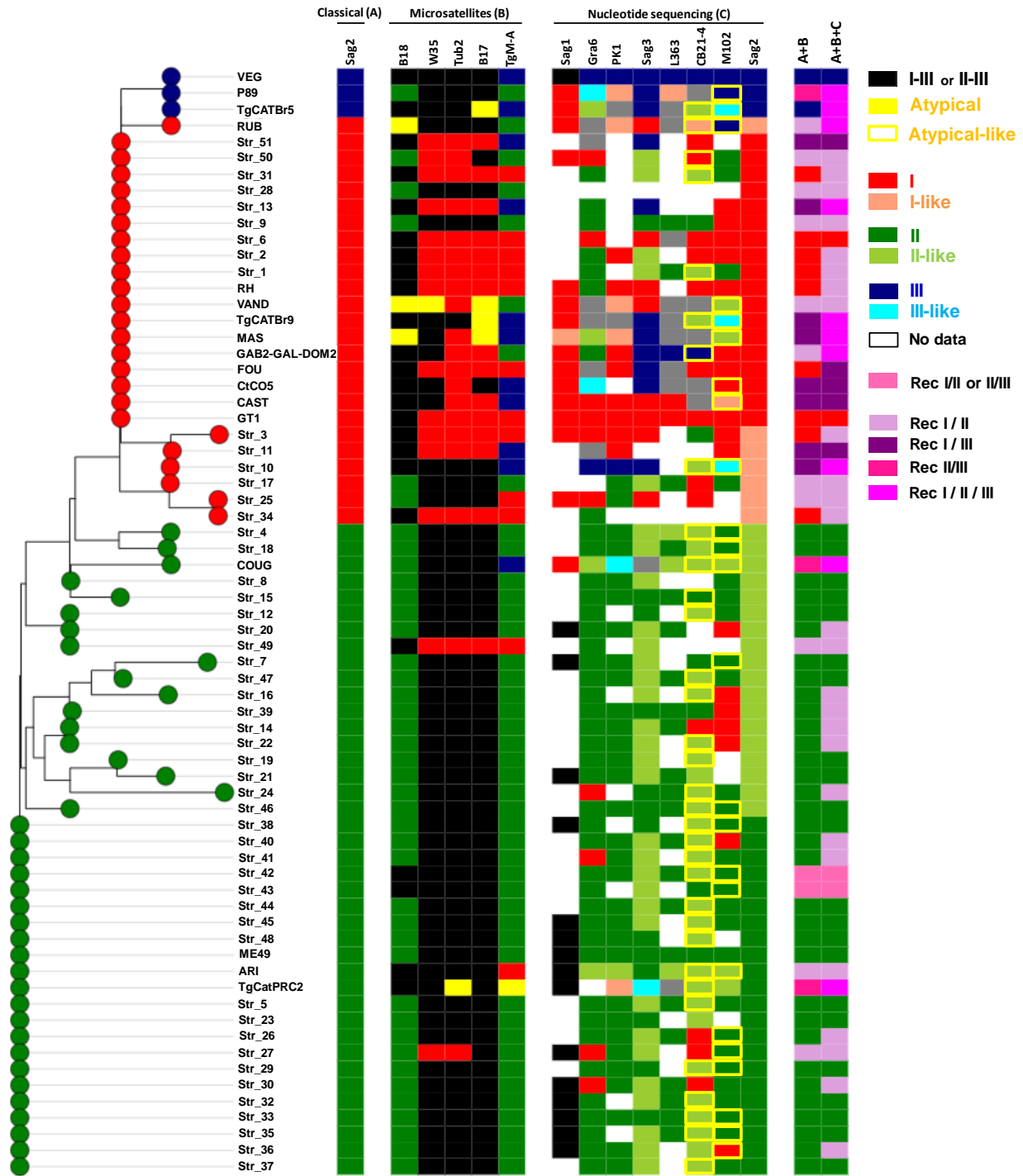
Figure 6.4. *T. gondii* genetic mosaicism. **Note:** Schematic representation of the allelic diversity and mosaicism of the 68 *T. gondii* strains (including 51 PT clinical strains) under study. The panel illustrates the nominal allele classification (color-coded) obtained for each locus (organized by genotyping strategy, as indicated above the panel), as well as the final strain classification (two right columns) obtained after the combination of loci used for traditional *Sag*2 and microsatellites genotyping (A + B) or the combination of all loci (A + B + C). Strains are grouped according with the topology of a Neighbor-joining phylogenetic tree constructed upon concatenation of the two sequenced regions at both *Sag*2 5' and 3' ends, which shows congruence with the *Sag*2 archetypes classification (represented by the colored nodes: I, red; II, green; III; blue). All detailed data can be found in Supplementary Table S5B.

Of note, both inter and intra-chromosomal recombination could be detected. For instance, 13 out of the 50 strains classified as I, II or III in *Gra*6 revealed discordant classification in the same-chromosome locus TgM-A. It is also noteworthy that the extended strategy allowed the identification of ten recombinant I/II/III strains (including two PT strains), which could not be detected when exclusively taking into account the classical *Sag*2- and microsatellites-derived profiles (Table 6.2; Supplementary Table S5B).

Table 6.2 Identification of interlocus recombination: use of two traditional loci panels (*Sag*2 classical + Microsatellites) *versus* combination of all loci[a]

| | | *Sag*2 classical + Microsatellites | | | | | | | | Total |
		I	II	III	Rec I/II	Rec I/III	Rec II/III	Rec I/II or II/III	Rec I/II/III	
Combination of all loci	I	1 (1)								1 (1)
	II		21 (21)							21 (21)
	III									0 (0)
	Rec I/II	6 (5)	11 (11)		9 (7)					26 (23)
	Rec I/III	1 (0)				4 (2)				5 (2)
	Rec II/III									0 (0)
	Rec I/II or II/III							2 (2)		2 (2)
	Rec I/II/III			1 (0)	2 (0)	4 (2)	3 (0)			10 (2)
	Total	8 (6)	32 (32)	1 (0)	11 (7)	8 (4)	3 (0)	2 (2)	0 (0)	*65 (51)*

[a] Strains assumed as parental for type I, II, and III classification are not taken into account for counting. Counts in parenthesis refer to PT strains. Light blue cells reflect the recombination "zone", where dark blue highlights recombination strains that were identified solely when combining all loci.

In general, these data are consistent with a scenario of remarkable genetic mosaicism in *T. gondii* (Figure 6.4).

6.4.3 Atypical and atypical-like profiles and insights on microsatellites-driven intra-host diversification

T. gondii strains have been traditionally classified at genetic level as "atypical" when the number of tandem repeats within microsatellites-containing loci does not match the number described for the three archetypes reference strains (Ajzenberg *et al.*, 2005). In order to keep the backwards compatibility with this historical classification, we have also applied it based on the diversity found within the loci enrolled in the traditional microsatellites-based analyses (B18, B17, TgM-A, *Tub*2 and W35). Although none of the PT isolates exhibited an "atypical" profile for these five loci, these profiles could be observed for six reference strains (Supplementary Table S5B; Figure 6.4). It is however worth noting that other references have also been classified as atypical when additional loci (not studied in the present study) were taken into account (Ajzenberg *et al.*, 2010). While we cannot directly transpose this classification for the two new loci sequenced by Sanger/NGS that also reveal tandem repeats (CB21-4 and M102), as the methodology is different, the small size of the tandem repeat still allowed verifying whether the consensus number of tandem repeats does not match the number described for the three archetypes reference strains (these cases were marked as "atypical-like") (Supplementary Table S5 and S6). For CB21-4, 33 (including 26 PT strains) out of the 45 strains (73.3 %) (for which the microsatellite length could be evaluated) revealed "atypical-like" microsatellites lengths (Supplementary Table S6). For M102, 24/56 (42.9 %) were classified as "atypical-like" strains (including 14 PT strains) (Supplementary Table S6). In summary, when combining the classification of all microsatellites-containing loci, a total of 43 (including 32 PT strains) out of 65 (66.2 %) studied strains (excluding three parental strains) revealed at least one locus with atypical or atypical-like length (Supplementary Table S5).

The application of in-depth NGS to the two newly sequenced loci displaying tandem repeats (CB21-4 and M102) further allowed us to investigate whether different microsatellites sizes were observed within a single sample, which could reflect *T. gondii* intra-host population diversity. While a most abundant profile (assumed as the consensus) was identified for both loci for most strains analysed (42 strains for M102 and 32 for CB21-4) (Supplementary Figure 1), we observed a scenario of considerable intra-population diversity marked by the co-existence of sub-populations carrying different microsatellite repeat number.

6.4.4 Insights on the association between loci/strain classification and mouse virulence

Previous studies (Sibley and Boothroyd, 1992a; Khan *et al.*, 2005a; Vilares *et al.*, 2017) have pointed out that most *T. gondii* strains exhibiting an enhanced virulence in mouse are classified as Type I or I-like by the classical *Sag*2 ends "RFLP" recognition sites genotyping. This trend is also observed in the present study (Figure 6.1), with 18 out of 21 (85.7 %) mouse virulent strains being Type I or I-like by *Sag*2 (Supplementary Table S5). From another point-of-view, 18 out of 25 (72 %) Type I or I-like by *Sag*2 were found to be mouse virulent. As this could find a parallel for other loci, we sought to evaluate the potential correlation between mouse virulence and the nominal classification obtained for each individual locus. Interestingly, we observed that > 80 % of strains classified as Type I or I-like by using B17, *Sag*1, PK1 and *Sag*3 individually were mouse virulent (although L363 showed the same trend, it was not included because there were only three type I strains). The highest Adjusted Rand coefficient (0.742) was obtained for *Sag*1 locus which exclusively discriminates I from II-III profiles. Thirteen out of the 15 strains with *Sag*1 type I (for which virulence was determined) were mouse virulent, whereas none of the 15 non-type I displayed a virulence profile. Noteworthy, although Type III or III-like profiles were unfrequently found, > 80 % strains displaying this classification for TgM-A and *Sag*3 were mouse virulent. In an opposite scenario, Type II or II-like strains are particularly overrepresented among non-virulent strains regardless of the locus (more than 75 % Type II or II-like are non-virulent). Finally, we inspected the concordance between the mouse virulence phenotype and the final strains' classification. Taking into account that 23 out of the 24 Type I/I-like strains (*Sag*2 classification only) presented a recombinant profile when combining the classification of all analysed loci, most virulent strains (18/21; 85.7 %) were expectedly found to be recombinant (5 Rec I/III, 6 Rec I/II and 7 Rec I/II/III strains) (Supplementary Table S5). Only 51 % of non-virulent strains present a recombinant profile. Also not unexpected, 21 out of 22 strains (95.5 %) that present a type II or II-like profile throughout all loci displayed a non-virulent phenotype in mouse.

6.5 Discussion

The acquisition of genomic data from *T. gondii* causing human infection is very challenging due to the limited access to human tissues for parasite isolation and characterisation. Indeed, more than 90 % of existing *T. gondii* genotyping data is derived from animal sources (Robertson *et al.*, 2019). Moreover, our current knowledge about the global *T. gondii* population structure and genetic diversity essentially relies on limited genomic data obtained through the analysis of discrete genetic markers (e.g., RFLP markers of *Sag*2 loci), microsatellites or the concatenation of sequence data from a few chromosome-dispersed loci (Ajzenberg *et al.*, 2005, 2009; Su *et al.*, 2010; Dubey *et al.*, 2011b; Ferreira *et al.*, 2011; Vijaykumar *et al.*, 2018). Although these *data* have historically sustained that *T. gondii* population can be structured into three main archetypal lineages, marked by both low inter- (1 to 3 %) and intra-lineage (< 0.01 %) genetic variation and distinct global geographic distribution (Su, 2003; Boyle *et al.*, 2006; Sibley and Ajioka, 2008; Dardé, 2008; Minot *et al.*, 2012), our knowledge about the real *T. gondii* genetic landscape is still certainly skewed since *T. gondii* is known to undergo marked sexual recombination. In fact, this cannot be captured by single-loci-based strategies and may be masked by the simple concatenation of different loci (even if they have completely different evolutionary histories) (Khan *et al.*, 2005b; Grigg and Sundar, 2009; Sibley *et al.*, 2009; Minot *et al.*, 2012). Most surveillance laboratories worldwide are still not capacitated to perform sequencing of a large loci panels   (Herrmann *et al.*, 2010, 2012b, 2013; Vilares *et al.*, 2014), and WGS of such complex parasite is rather cost-prohibitive for routine surveillance purposes in a near future (Talundzic et al., 2017). WGS-based studies have however been key for understanding that *T. gondii* presents an evolutionary history marked not only by clonal expansion and selective sweeps (Sibley and Ajioka, 2008; Minot *et al.*, 2012), but also inheritance of large haploblocks expectedly impacting transmission, host range and pathogenicity (Khan *et al.*, 2005a; Dubey and Jones, 2008; Su *et al.*, 2012; Lorenzi *et al.*, 2016).

In the present study, we gave a little step towards the better characterization of *T. gondii* genetic diversity by applying a novel genotyping NGS-based strategy, and comparing it with traditional approaches. We aimed to implement such strategy as the molecular typing tool for routine surveillance of *T. gondii* carried out by the URSZ-DDI. Although a restricted number of loci was analyzed, they are genome dispersed and some of them (e.g., *Gra*6, *Sag*1 and *Sag*3) are

inserted into genetic inherited haploblocks likely shaping transmission, host range and pathogenicity (Lorenzi *et al.*, 2016). This panel enabled not only to substantial increase the genetic discriminatory power (Figure 6.3), but also revealed a scenario of marked *T. gondii* genomic mosaicism (Figure 6.4). For example, the combined MLST-like allelic profiling increased the resolution power to up to 36 allelic types among 40 strains (Simpson's ID of 0.991), whereas no more than three types would have been identified if classical *Sag*2-based "RFLP" classical genotyping or microsatellites locus were analyzed individually. Also, the nucleotide sequencing approach alone (34 types; Simpson's ID of 0.985) was considerably more discriminatory than the combination of the two traditional methodologies (16 types; Simpson's ID of 0.718) (Figure 6.3). Interestingly, individual *Sag*2 nucleotide sequencing analysis not only ensured the backwards compatibility with the *Sag*2 classical "RFLP recognition sites genotyping" (i.e., full congruence with the classification into archetypes I, II and III), but also largely increased the discrimination (from three archetypes to 23 distinct allelic types out of 68 strains) (Fazaeli and Ebrahimzadeh, 2007) (Figure 6.2).

Regarding the detection of recombination, although the combination of *Sag*2 classical genotyping and microsatellites have already allowed inferring some degree of mosaicism, our approach almost doubled the detection of recombinant strains (a total of 56.9 % of PT strains and 66.2 % of all strains). This scenario is underestimated, since only a tiny part of the genome was studied, and allows us to speculate that virtually all circulating *T. gondii* strains resulted from genetic admixture. The scenario of recombination is particularly marked among strains traditionally classified as Type I. In fact, from the 25 type I strains classified solely by *Sag*2 classical genotyping, 92 % (23/25) were further classified as recombinant when combining all loci (Figure 6.4). This contrasts with *Sag*2 Type II strains, where only 45 % (18/40) could be identified as recombinant. This apparent less recombining character observed for the highly prevalent type II strains supports that this lineage might have been involved in selective sweep towards clonal expansion driven by a likely more fitted genome make-up (Sibley and Ajioka, 2008; Minot *et al.*, 2012). Again, this assumption needs validation through large-scale population studies enrolling large loci panels (or even WGS). *Sag*2 Type II strains were also considerably less linked to virulence, supporting that their potentially more fitted genome make-up favours the parasite maintenance and transmission (instead of promoting enhanced immunopathogenicity and

"dead end" infections), and minimize recombination due to its potential fitness cost. Of note, whereas *Sag*2 Type I strains are expectedly overrepresented among mouse virulent strains (Dubey *et al.*, 2007a; Eldeek *et al.*, 2017), other loci seemed to be good predictor of mouse virulence. That was the case for B17, PK1, *Sag*3 and *Sag*1, since 80 % of strains classified as Type I or I-like by these loci were mouse virulent. Hypothetically, these genes may constitute good candidates to integrate markers databases aiming to infer *T. gondii* virulence through rapid molecular screening.

Our first insights on intra-host diversity were not conclusive at this scale, but showed that microsatellites may be drivers of *T. gondii* population diversification on behalf of host-pathogen interaction, considering the observed non-static character of these genomic regions (Supplementary Figure S1). In fact, the NGS-based analysis of two loci (CB21-4 and M102) harbouring tandem repeats (i.e., microsatellites) showed not only a no conservation of microsatellites size among strains from the same type, but also traces of intra-host sub-population diversity. This kind of hypermutable regions are common mediators of pathogen intra-host genetic heterogeneity, a phenomenon highly described in bacteria field (van der Woude and Baumler, 2004), but also in some parasites (Gelanew *et al.*, 2014; Pajuelo *et al.*, 2017). In another perspective, our novel approach for capturing diversity within a specific microsatellite launches good perspectives that the traditional microsatellites analyses, which rely on laborious and error-prone electrophoresis-based DNA sizing, might be straightforwardly transfer to more straightforward, robust and informative NGS/bioinformatics-based methodologies.

In summary, while the transition to the application of novel sequencing technologies for *T. gondii* genotyping will not likely overcome some important hurdles in a near future, as discussed above, this study may represent a straightforward model for acquiring more consistent and robust genomic data. In fact, the approach applied here for *T. gondii* genetic characterization (taking advantage of the nucleotide sequencing of *Sag*2 and seven additional genome-dispersed loci, coupled with classical microsatellites-derived data) constitutes a relevant proof-of-concept that the establishment of wide multi-locus typing strategies have the potential to improve molecular epidemiology of *T. gondii* towards the identification and monitoring of circulating genotypes with clinical and epidemiological importance both locally and globally. In fact, the present study corroborates not only the expectation that more prevalent strains (frequently

associated with *Sag*2 type II profile) revealed a more homogeneous genomic structure (Su *et al.*, 2012; Minot *et al.*, 2012; Lorenzi *et al.*, 2016), but also that global *T. gondii* diversity is rampant and strongly linked to genetic admixture events (Khan *et al.*, 2005b; Boyle *et al.*, 2006). This consolidates that tracking the genetic diversity (and mosaicism) of *T. gondii* may be of special relevance not only for the study of genotype-phenotype associations, but also to strengthen molecular epidemiology of *T. gondii*, which may be useful for both routine surveillance and source attribution in eventual outbreak context.

PART III

CHAPTER VII

Final discussion, conclusions and future directions

DISCUSSION, CONCLUSIONS AND FUTURE DIRECTIONS

7.1 FINAL DISCUSSION AND CONCLUSIONS:

Since the studies published in the 1990s by Howe, Sibley and Dardé (Dardé *et al.*, 1992; Howe *et al.*, 1997), the goal to understand *Toxoplasma gondii* genotyping persists and much has been written about this parasite regarding genetic characteristics such as clonality and recombination. *T. gondii* is widespread in both human and animal hosts, however, most studies on *T. gondii* in Portugal have focussed on serological investigations and some discrete studies consisted of genotyping the parasite isolated from animals. In order to overcome this lack of information, this PhD thesis aimed to provide a more comprehensive overview of the genetic characteristics and virulence of *T. gondii* strains circulating in Portugal.

Thus, in chapter III, we performed *Sag* 2 genotyping in order to identify the *T. gondii Sag*2 archetype strains (I, II and III) in farm (cattle, sheep, goats, pigs) and wild animals (stray dogs, foxes, boars, hares). While farm animals were slaughtered for human consumption, boars and hares play an important role in human food chain as intermediate hosts and they are hunted in large scale in Portugal (Coelho *et al.*, 2014, 2015; Lopes *et al.*, 2014). Although stray dogs and foxes do not have a role in human food chain and do not have predators, they may reflect the environmental contamination and can be a reservoir for the scavenging by other animals. We took advantage of the collaboration with a research group from the Department of Veterinary Sciences of UTAD, allowing us to have access to 317 samples (from 209 animals representing eight species) as part of a seroprevalence study that was on course in the north region of Portugal (Lopes *et al.*, 2012a).

Our study constituted the first molecular characterization report of *T. gondii* in goats, wild boars, foxes and hares in Portugal. The lack of previous evaluations is likely due to the difficulties associated with the collection of biological material in perfect conditions to carry on this kind of studies. In fact, the collection of enough amount of high quality DNA to characterize *T. gondii* is not always possible, especially in wildlife in which sampling activities and death animals may occur in different spaces and time-windows, hampering the recovery of viable *T. gondii* to inoculate in mice. Furthermore, the bioassay in mice is extremely costly and time-consuming, hampering its use for epidemiological studies enrolling an extensive number of samples. Eighty-nine animals (43 %) yielded positive results when B1 classical PCR was

performed, demonstrating that environmental contamination in the synanthropic and silvatic cycles constitute a potential source of human infection. Although no information could be obtained, the high infection rate observed for farm animals (35,8 %) could eventually be explained if no proper control measures were undertaken, such as: *i)* lack of confinement of farm animals; *ii)* no rodent control; and *iii)* lack of control for contamination of feed and water for farm animals.

We also observed that the majority of *T. gondii* strains for which *Sag*2 classical genotyping was successfully applied (29 %; 60/209) were type II (77 %; 46/60), followed by the type I (10 %; 6/60), where the latter was observed solely in strains isolated from wild species. *Sag*2 type II *T. gondii* strains were shown here, for the first time, to infect cattle, boars, foxes and hares in north of Portugal. As in our study, the predominance of type II *T. gondii* strains in animals has been widely reported by other authors (Aspinall *et al.*, 2002a; Jungersen *et al.*, 2002; Dubey *et al.*, 2004a; Dumètre *et al.*, 2006; Prestrud *et al.*, 2008; Richomme *et al.*, 2009; Halos *et al.*, 2010; Aubert *et al.*, 2010; Berger-Schoch *et al.*, 2011; Herrmann *et al.*, 2013; VERGARA *et al.*, 2018). Noteworthy, in our study type I was detected in wild boars, foxes and hares, which is a remarkable finding because this genotype is not frequently reported in animals in Europe (Berger-Schoch *et al.*, 2011; De Craeye *et al.*, 2011) with exception of one study conducted in Italy. In that study, in opposition to our study, the higher prevalence of genotype I was observed at higher rates in omnivorous domestic species, represented by pigs, than in omnivorous and carnivorous sylvatic species (wild boars and foxes) (Battisti *et al.*, 2018). However, the apparently increased genetic diversity from wildlife samples is not absolute, as studies of wildlife in France, Germany and Norway have reported a high prevalence of archetypal strains, where the majority of type II prevails (Dubey *et al.*, 2004a; Prestrud *et al.*, 2008; Richomme *et al.*, 2009; Aubert *et al.*, 2010; Herrmann *et al.*, 2013; Calero-Bernal *et al.*, 2015). Overall, the results of this study highlighted the need to extend such genetic evaluations to other animals and locations, such as the focus on the definitive host (cat), particularly in high-density population areas.

As such, in Chapter IV, we aimed to perform a molecular characterization of *T. gondii* strains collected from stray cats and pigeons from the Lisbon region, since these animals share the same "habitat" as humans and feed themselves from the ground, favoring the fecal-oral transmission of *T. gondii* (Waap *et al.*, 2012). Moreover, cats are the definitive host and embody a

large responsibility in the spread of this parasite since they are the only felid in urban cycle that can shed oocysts (Elmore *et al.*, 2010). We took advantage of the collaboration with the Sanitary Control Department of Lisbon City Hall and with INIAV, allowing us to have access to samples from 1507 free-living pigeons captured from 64 neighbourhoods of Lisbon, and tissue samples from 467 cats (Waap *et al.*, 2012). These animals were captured and killed under the pre-existing animal control programs by the city hall services. In Portugal, previous studies focusing genetic and virulence characterization of *T. gondii* strains in cats had never been performed and a first evaluation in pigeons was previously achieved by our group over a very small number of samples (Waap *et al.*, 2008).

In the present study, we observed a *T. gondii* positivity rate of 68 % (28/41) in pigeons by PCR of B1 gene. This value was considerably higher than the one obtained in our previous study (39 %; 9/23) (Waap *et al.*, 2008), which may be due either to the dissimilar sample size, or to the difficulties associated with the control of the parasite concentration in brain samples (Dubey *et al.*, 2005a). In cats, the *T. gondii* molecular identification by B1 gene was 51 % (84/164), a higher value than the ones previously described by other authors (Esteves *et al.*, 2014; Tian *et al.*, 2014). The inoculation of brain homogenates from pigeons in mice yielded a *T. gondii* isolation rate of 58.5 % (24/41), whereas for cats the value was much lower (11 %; 18/164). This difference was expected since the pigeons' brain favour the isolation of the parasite due to its smaller size (yielding a higher concentration of the parasite) and less fibrous character, when compared with the cats' brain. We found three virulent strains (isolated from one cat and two pigeons) out of 42 (from 24 pigeons and 18 cats) after mice inoculation. The *in vivo* assays, besides allowing the study of strains' virulence, also facilitated their genetic characterization by improving the parasite concentration and sample purity. According to our experience, the *T. gondii* DNA is more easily extracted from the mice brains than from the pigeons and even from the cats' brains and muscles. After recovering the *T. gondii* strains using essentially the mice bioassays, genotyping of *T. gondii* was achieved in 70.7 % (29/41) in pigeons and 50 % (82/164) in cats by *Sag*2 classical PCR. Once more, the majority of *T. gondii* strains were type II from both pigeons and cats, followed by type I strains. However, additional PCR of 5 microsatellites (*Tub*2, TgM-A, W35, B17, B18), allowed the identification of two recombinant strains (one from a pigeon and one from a cat) previously identified as type II by *Sag*2 amplification. Of note, as expected, the two

95

recombinant and one type I strain were virulent in mice. In this study, we presented the first data in Portugal concerning the *T. gondii* genotyping from the definitive host, and revealed the first report of recombinant strains both in cats and in pigeons.

Although these previous studies were merely humble contributions to a better understanding of the toxoplasmosis in animals isolated in Portugal, they constituted the first steps towards deciphering the scenario of the genetic diversity of *T. gondii* circulating in multiple animal species that constitute important reservoirs for human infection. In chapter V, we precisely focused our efforts in characterizing the *T. gondii* strains isolated from humans. We analyzed 48 strains from congenital and acquired toxoplasmosis collected during the last two decades. These samples were sent to the URSZ-DDI on behalf of the toxoplasmosis diagnosis (Pre, Post-Natal and surveillance of toxoplasmosis). Similarly to the earlier studies described in this PhD thesis, a majority of type II strains (73 %; 35/48) were identified by *Sag*2 classical PCR, mirroring the results reported by other authors in Europe that also focused in congenital toxoplasmosis (Nowakowska *et al.*, 2006; Ajzenberg *et al.*, 2015). A major outcome of this work was the finding of 21 % (10/48) of recombinant strains by PCR targeting five microsatellites (previously used in chapter IV, in samples from pigeons and cats) that had been previously identified as type II or I by *Sag*2 classical genotyping. Although type I strains are typically described as more virulent in mice than the other archetypes, unexpectedly, this study identified only three type I strains (out of 6) demonstrating virulence in mice. In contrast, one type II strain killed the mice in 120 days after the first inoculation, and kept killing mice in less time in the subsequent passages. We speculate that these strains classified as type I and II by classical *Sag*2 and 5 microsatellites, may eventually be recombinant strains if other loci were analyzed. The observed virulence rate of 15 % (7/48) in the mice bioassay was quite low, but was similar to the one observed in other countries enrolling human samples (Ajzenberg *et al.*, 2002b; Gebremedhin *et al.*, 2014).

The identification of *T. gondii* strains displaying a recombinant character in both human (this chapter) and animal samples (previous chapters) may be of particular relevance, as it is believed that they are associated with more severe symptoms in immunocompetent and immunocompromised hosts and may pass through the placenta more easily than the archetypal strains (Grigg *et al.*, 2001b). Nevertheless, revealing a recombinant character by simply using the traditional typing loci is certainly a too modest approach as the probability of finding

recombinant strains increases if broader genomic regions are evaluated. As such, in chapter VI, we aimed to evaluate the genomic recombination of *T. gondii* isolated from humans in Portugal. We combined classical *Sag*2- and microsatellites-based typing with additional sequencing (Sanger and NGS) of seven loci dispersed by the different *T. gondii* chromosomes, including genes believed to have a role in virulence and adaptation, like CB21-4, PK1, L363, *Sag*1, *Gra*6, *Sag*3 and M102 (Khan *et al.*, 2005a; de Melo Ferreira *et al.*, 2006; Ajzenberg *et al.*, 2010). The putative identification that virulent genes are transferred between strains may be a strong indicator that genome plasticity is a major mechanism used by *T. gondii* for infection and survival purposes. CB21-4 and M102 are microsatellites markers largely used in the differentiation of *T. gondii*, namely in the identification of recombinant and atypical strains (de Melo Ferreira *et al.*, 2006; Ajzenberg *et al.*, 2010). PK1 encodes a pyruvate kinase that converts phosphoenolpyruvate to pyruvate in glycolysis and it is a central metabolic regulator in most organisms (Khan *et al.*, 2005a; Bakszt *et al.*, 2010). L363 is a possible marker of highly virulent strains with unknown function (de Melo Ferreira *et al.*, 2006). *Sag*1 is the major surface protein of *T. gondii* and, together with *Sag*3, are involved in parasite-host cell attachment allowing the entry of the parasite into host cell (Mineo and Kasper, 1994; Jacquet *et al.*, 2001). Dense granule antigens of *T. gondii* ("GRA") are engaged not only in virulence, but also in parasite survival (Mercier *et al.*, 2005). These GRA constitute an important fraction of antigens that circulate in the bloodstream during the first hours following infection (Golkar *et al.*, 2008). IgG antibody response to GRA6 has been reported stronger in acute infection than chronic infection (Gatkowska *et al.*, 2006).

We implemented a rapid NGS-based schema to perform molecular surveillance of *T. gondii*, which allowed evaluating multi loci polymorphism and genomic mosaicism simultaneously in a single tube. We highlight the major finding of this study, the identification of a high rate (63 %; 43/68) of recombinant strains, where the majority of those had been previously identified as type I or II by classical *Sag*2. Remarkably, the genetic mosaicism was observed more frequently in "previous" *Sag*2 type I strains (92 %) than in type II (45 %) strains. Based on the results of mice inoculation, we also concluded that other loci [namely *Sag*1, B17 (microsatellite), PK1, *Sag*3 and TgM-A (microsatellite)] besides *Sag*2, may also constitute good predictors of *T. gondii* mouse virulence, as specific alleles of those loci seemed to be associated with virulence in mice. We also observed a scenario of considerable intra-population diversity

when analyzing by NGS the microsatellite profile of M102 and CB21-4, suggesting the existence of intra-host adaptation during the infection process. Nevertheless, studies at higher scale, both in the number of strains and target loci, would be needed to support such speculation. The results from this study raise good perspectives that the laborious intensive historical method can be directly transferred to this new methodology (NGS) as the proposed NGS-based approach may represent a straightforward model for acquiring more consistent and robust genomic data in a laboratory "user-friendly" fashion. In this regard, this new *T. gondii* genetic surveillance system was firmly implemented in the URSZ-DDI in Portugal and is already being used in different research studies.

Overall, this PhD thesis shed some light on the genetic and virulence diversity of *T. gondii* strains infecting animals and humans and revealed a surprising scenario of rampant genomic exchange in this parasite, where virtually, all strains possess mosaic chromosomes. Although this major finding brings some complexity to the interpretation of the genotype-phenotype ("virulence") association, it is notable that "type-II-like" loci are predominant in the *T. gondii* strains that were analyzed. As such, the high prevalence of these type-II-like strains should not be underestimated, as human infections with strains belonging to these genotypes can be fatal or even severe. The majority of this potentially dangerous genotype in the environment and in different hosts (definitive and intermediate) is both a risk factor and a welfare issue for animals and humans. Although the impact of such findings cannot be immediately addressed at this stage, future genotype-phenotype association studies enrolling multiple loci may reveal genomic structures potentially associated with an increased virulence phenotype, launching basis for the development of proper prophylactic or therapeutic measures.

7.2 FUTURE DIRECTIONS:

The ability to acquire knowledge about microorganisms' genetics has entered into a new level due to the rampant progress of sequencing and bioinformatics technologies, which are expected to become easily available (at both handling and cost level) to the scientific community in the next years. In addition, the investments on the 3R's allows the increase of knowledge in the phenotypic behaviour of microorganisms using, for instance, the cell culture as an alternative methodology to mice inoculation. The application of these approaches to *T. gondii,* even if they are still considered a *long shot* due to the difficulties associated with both the genome size and laboratory handling of this parasite, may help the clarification of some of the *T. gondii* recognized enigmas, such as: *i)* Why type II strains that are avirulent in mice are virulent in humans (mainly in ocular disorders); *ii)* How and which genotypes (i.e., multi-loci based genotypes) are responsible for human disease? *iii)* What justifies the observed genomic differences of *T. gondii* strains in different world regions? *iv)* Why, in some countries, the genotypes isolated from wild animals are so different from the ones isolated in other animals and humans that shares the same habitat? Finally, some new questions show up on the course of this Ph.D. thesis, namely those arising from the investigation of the genomic structure of *T. gondii* strains. In particular, it may be speculated that a specific genome make-up or specific combination of alleles from multiple loci may underlie the observed phenotypic properties of the parasite (e.g., virulence in different hosts). Considering the above exposed aspects and the laboratory capacity of the URSZ-INSA, we have two major goals for the near future:

- **On behalf of the 3R's goal (Reduce, Reuse, Recycle), to reduce the mice used in future projects, mainly the ones enrolling genotyping procedures**

Experimental approach:

- Propagation of multiple *T. gondii* strains in different cell lines with continuous assessment of progeny and virulence. This task is already on course in our lab;

Major question to be answered:

- Is it possible to replace the mouse model for *T. gondii* propagation in order to perform typing assays without losing the virulence characteristics?

- **To better understand the genomic characteristics of** *T. gondii* **strains isolated in Portugal.**

Experimental approach:

- By using the multi-loci NGS-based approach, developed in chapter VI and already implemented in the URSZ-INSA, we will assess the genomic profile of *T. gondii* strains isolated from animals;

Major questions to be answered:

- Are the genomic profiles of *T. gondii* strains isolated from animals "similar" to the ones observed in strains isolated from humans (identified in chapter VI)?
- Can we better understand the infection sources responsible for the human infection?

REFERENCES

Ajioka, J. W., Fitzpatrick, J. M. and Reitter, C. P. (2001). *Toxoplasma gondii* genomics: shedding light on pathogenesis and chemotherapy. *Expert Reviews in Molecular Medicine* **3**, 1–19. doi: 10.1017/S1462399401002204.

Ajzenberg, D., Bañuls, A. L., Tibayrenc, M. and Dardé, M. L. (2002a). Microsatellite analysis of *Toxoplasma gondii* shows considerable polymorphism structured into two main clonal groups. *International journal for parasitology* **32**, 27–38.

Ajzenberg, D., Cogné, N., Paris, L., Bessières, M., Thulliez, P., Filisetti, D., Pelloux, H., Marty, P. and Dardé, M. (2002b). Genotype of 86 *Toxoplasma gondii* Isolates Associated with Human Congenital Toxoplasmosis, and Correlation with Clinical Findings. *The Journal of Infectious Diseases* **186**, 684–689. doi: 10.1086/342663.

Ajzenberg, D., Bañuls, A. L., Su, C., Dumètre, A., Demar, M., Carme, B. and Dardé, M. L. (2004). Genetic diversity, clonality and sexuality in *Toxoplasma gondii*. *International Journal for Parasitology* **34**, 1185–1196. doi: 10.1016/j.ijpara.2004.06.007.

Ajzenberg, D., Dumètre, A. and Dardé, M. L. (2005). Multiplex PCR for typing strains of *Toxoplasma gondii*. *Journal of Clinical Microbiology* **43**, 1940–1943. doi: 10.1128/JCM.43.4.1940-1943.2005.

Ajzenberg, D., Pelloux, H., Quinio, D., Thulliez, P., Cuisenier, B., Desbois, N., Aubert, D., Robert-Gangneux, F., Pratlong, F., Kauffmann-Lacroix, C., Nevez, G., Delhaes, L., Rabodonirina, M., Dardé, M., Duhamel, C., Totet, A., Gay-Andrieu, F., Dalle, F., Bessières, M., Yera, H., Villena, I., Menotti, J., Duong, T. H., Marty, P., Bougnoux, M., Franck, J., Paris, L., Pujol, S., Bonnabau, H., Carme, B., Flori, P. and Filisetti, D. (2009). Genotype of 88 *Toxoplasma gondii* Isolates Associated with Toxoplasmosis in Immunocompromised Patients and Correlation with Clinical Findings. *The Journal of Infectious Diseases* **199**, 1155–1167. doi: 10.1086/597477.

Ajzenberg, D., Collinet, F., Mercier, A., Vignoles, P. and Dardé, M. L. (2010). Genotyping of *Toxoplasma gondii* isolates with 15 microsatellite markers in a single multiplex PCR assay. *Journal of Clinical Microbiology* **48**, 4641–4645. doi: 10.1128/JCM.01152-10.

Ajzenberg, D., Collinet, F., Aubert, D., Villena, I., Dardé, M.-L. and Devillard, S. (2015). The rural–urban effect on spatial genetic structure of type II *Toxoplasma gondii* strains involved in human congenital toxoplasmosis, France, 2002–2009. *Infection, Genetics and Evolution* **36**, 511–516. doi: 10.1016/J.MEEGID.2015.08.025.

Al-Kappany, Y. M., Rajendran, C., Abu-Elwafa, S. a, Hilali, M., Su, C. and Dubey, J. P. (2010). Genetic diversity of Toxoplasma gondii isolates in Egyptian feral cats reveals new genotypes. *The Journal of parasitology* **96**, 1112–1114. doi: 10.1645/GE-2608.1.

Amdouni, Y., Rjeibi, M. R., Rouatbi, M., Amairia, S., Awadi, S. and Gharbi, M. (2017). Molecular detection of *Toxoplasma gondii* infection in slaughtered ruminants (sheep, goats and cattle) in Northwest Tunisia. *Meat Science* **133**, 180–184. doi: 10.1016/J.MEATSCI.2017.07.004.

Ângelo, H. (1983). Prevalência dos anticorpos antitoxoplasmose - Inquérito Serológico Nacional, Portugal Continental, 1979/1980. *Arquivos INSA* **8**, 105–111.

Ângelo, H. (2000). Metodologia de Identificação da Infecção por *Toxoplasma gondii* no Hospedeiro Imunocomprometido.

Ângelo, M. H. (2003). Legal dispositions and preventive strategies in congenital toxoplasmosis in Portugal. *Archives de Pediatrie* **10**, 25–26.

Antunes, F. (1984). Toxoplasmose - Estudo da epidemiologia e da infecção congénita na região de Lisboa.

Antunes, F., Bacelar, F., Janz, J. G., Saraiva, J. and Araújo, F. C. (1981). Prevalência e incidência da toxoplasmose adquirida em 686 grávidas da região de Lisboa. *Médico* **101**, 904–912.

Armand, B., Solhjoo, K., Shabani-Kordshooli, M., Davami, M. H. and Sadeghi, M. (2016). Toxoplasma infection in sheep from south of Iran monitored by serological and molecular methods; risk assessment to meat consumers. *Veterinary World* **9**, 850–855. doi: 10.14202/vetworld.2016.850-855.

Aspinall, T. V., Marlee, D., Hyde, J. E. and Sims, P. F. G. (2002a). Prevalence of *Toxoplasma gondii* in commercial meat products as monitored by polymerase chain reaction – food for thought? *International Journal for Parasitology* **32**, 1193–1199. doi: 10.1016/S0020-7519(02)00070-X.

Aspinall, T. V., Joynson, D. H. M., Guy, E., Hyde, J. E. and Sims, P. F. G. (2002b). The Molecular Basis of Sulfonamide Resistance in *Toxoplasma gondii* and Implications for the Clinical Management of Toxoplasmosis.

The Journal of Infectious Diseases **185**, 1637–1643. doi: 10.1086/340577.

Aspinall, T. V, Guy, E. C., Roberts, K. E., Joynson, D. H. M., Hyde, J. E. and Sims, P. F. G. (2003). Molecular evidence for multiple *Toxoplasma gondii* infections in individual patients in England and Wales: public health implications. *International journal for parasitology* **33**, 97–103.

Aubert, D. and Villena, I. (2009). Detection of *Toxoplasma gondii* oocysts in water: proposition of a strategy and evaluation in Champagne-Ardenne Region, France. *Memorias do Instituto Oswaldo Cruz* **104**, 290–5. doi: 10.1590/s0074-02762009000200023.

Aubert, D., Ajzenberg, D., Richomme, C., Gilot-Fromont, E., Terrier, M. E., de Gevigney, C., Game, Y., Maillard, D., Gibert, P., Dardé, M. L. and Villena, I. (2010). Molecular and biological characteristics of *Toxoplasma gondii* isolates from wildlife in France. *Veterinary Parasitology* **171**, 346–349. doi: 10.1016/J.VETPAR.2010.03.033.

Baatz, H., Mirshahi, A., Puchta, J., Gümbel, H. and Hattenbach, L.-O. (2006). Reactivation of toxoplasma retinochoroiditis under atovaquone therapy in an immunocompetent patient. *Ocular immunology and inflammation* **14**, 185–7. doi: 10.1080/09273940600659740.

Bacci, C., Vismarra, A., Mangia, C., Bonardi, S., Bruini, I., Genchi, M., Kramer, L. and Brindani, F. (2015). Detection of Toxoplasma gondii in free-range, organic pigs in Italy using serological and molecular methods. *International Journal of Food Microbiology* **202**, 54–56. doi: 10.1016/j.ijfoodmicro.2015.03.002.

Bakszt, R., Wernimont, A., Allali-Hassani, A., Mok, M. W., Hills, T., Hui, R. and Pizarro, J. C. (2010). The Crystal Structure of *Toxoplasma gondii* Pyruvate Kinase 1. *PLoS ONE* **5**, e12736. doi: 10.1371/journal.pone.0012736.

Bártová, E., Sedlák, K., Syrová, M. and Literák, I. (2010). *Neospora* spp. and *Toxoplasma gondii* antibodies in horses in the Czech Republic. *Parasitology Research* **107**, 783–785. doi: 10.1007/s00436-010-1929-4.

Battisti, E., Zanet, S., Trisciuoglio, A., Bruno, S. and Ferroglio, E. (2018). Circulating genotypes of *Toxoplasma gondii* in Northwestern Italy. *Veterinary Parasitology* **253**, 43–47. doi: 10.1016/J.VETPAR.2018.02.023.

Belluco, S., Simonato, G., Mancin, M., Pietrobelli, M. and Ricci, A. (2018). *Toxoplasma gondii* infection and food consumption: A systematic review and meta-analysis of case-controlled studies. *Critical Reviews in Food Science and Nutrition* **58**, 3085–3096. doi: 10.1080/10408398.2017.1352563.

Berdoy, M., Webster, J. P. and Macdonald, D. W. (1995). Parasite-altered behaviour: is the effect of *Toxoplasma gondii* on *Rattus norvegicus* specific? *Parasitology* **111 (Pt 4**, 403–9.

Berdoy, M., Webster, J. P. and Macdonald, D. W. (2000). Fatal attraction in rats infected with *Toxoplasma gondii*. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **267**, 1591–1594. doi: 10.1098/rspb.2000.1182.

Berenreiterová, M., Flegr, J., Kuběna, A. A. and Němec, P. (2011). The distribution of *Toxoplasma gondii* cysts in the brain of a mouse with latent toxoplasmosis: implications for the behavioral manipulation hypothesis. *PloS one* **6**, e28925. doi: 10.1371/journal.pone.0028925.

Berger-Schoch, A. E., Herrmann, D. C., Schares, G., Müller, N., Bernet, D., Gottstein, B. and Frey, C. F. (2011). Prevalence and genotypes of Toxoplasma gondii in feline faeces (oocysts) and meat from sheep, cattle and pigs in Switzerland. *Veterinary Parasitology* **177**, 290–297. doi: 10.1016/j.vetpar.2010.11.046.

Bierly, A. L., Shufesky, W. J., Sukhumavasi, W., Morelli, A. E. and Denkers, E. Y. (2008). Dendritic cells expressing plasmacytoid marker PDCA-1 are Trojan horses during *Toxoplasma gondii* infection. *Journal of immunology (Baltimore, Md. : 1950)* **181**, 8485–91. doi: 10.4049/jimmunol.181.12.8485.

Blackston, C. R., Dubey, J. P., Dotson, E., Su, C., Thulliez, P., Sibley, D. and Lehmann, T. (2001). High-resolution typing of *Toxoplasma gondii* using microsatellite loci. *Journal of Parasitology* **87**, 1472–1475. doi: 10.1645/0022-3395(2001)087[1472:HRTOTG]2.0.CO;2.

Blader, I. J. and Saeij, J. P. (2009). Communication between *Toxoplasma gondii* and its host: impact on parasite growth, development, immune evasion, and virulence. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica* **117**, 458–76. doi: 10.1111/j.1600-0463.2009.02453.x.

Blaizot, R., Nabet, C., Blanchet, D., Martin, E., Mercier, A., Dardé, M.-L., Elenga, N. and Demar, M. (2019). Pediatric Amazonian Toxoplasmosis Caused by Atypical Strains in French Guiana, 2002-2017. *The Pediatric infectious disease journal* **38**, e39–e42. doi: 10.1097/INF.0000000000002130.

Borges, V., Pinheiro, M., Pechirra, P., Guiomar, R. and Gomes, J. P. (2018). INSaFLU: an automated open web-

based bioinformatics suite "from-reads" for influenza whole-genome-sequencing-based surveillance. *Genome Medicine* **10**, 46. doi: 10.1186/s13073-018-0555-0.

Bossi, P. and Bricaire, F. (2004). Severe acute disseminated toxoplasmosis. *Lancet (London, England)* **364**, 579. doi: 10.1016/S0140-6736(04)16841-4.

Boughattas, S., Abdallah, R. Ben, Siala, E., Aoun, K. and Bouratbine, A. (2011). An atypical strain associated with congenital toxoplasmosis in Tunisia. *New Microbiologica* **34**, 413–416.

Bowie, W. R., King, A. S., Werker, D. H., Isaac-Renton, J. L., Bell, A., Eng, S. B. and Marion, S. A. (1997). Outbreak of toxoplasmosis associated with municipal drinking water. *The Lancet* **350**, 173–177. doi: 10.1016/S0140-6736(96)11105-3.

Boyle, J. P., Rajasekar, B., Saeij, J. P. J., Ajioka, J. W., Berriman, M., Paulsen, I., Roos, D. S., Sibley, L. D., White, M. W. and Boothroyd, J. C. (2006). Just one cross appears capable of dramatically altering the population biology of a eukaryotic pathogen like *Toxoplasma gondii*. *Proceedings of the National Academy of Sciences* **103**, 10514–10519. doi: 10.1073/pnas.0510319103.

Calero-Bernal, R., Saugar, J. M., Frontera, E., Pérez-Martín, J. E., Habela, M. A., Serrano, F. J., Reina, D. and Fuentes, I. (2015). Prevalence and genotype identification of *Toxoplasma gondii* in wild animals from southwestern Spain. *Journal of Wildlife Diseases* **51**, 233–238. doi: 10.7589/2013-09-233.

Carme, B., Bissuel, F., Ajzenberg, D., Bouyne, R., Aznar, C., Demar, M., Bichat, S., Louvel, D., Bourbigot, A. M., Peneau, C., Neron, P. and Dardé, M. L. (2002). Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana. *Journal of clinical microbiology* **40**, 4037–44. doi: 10.1128/jcm.40.11.4037-4044.2002.

Carriço, J. A., Silva-Costa, C., Melo-Cristino, J., Pinto, F. R., de Lencastre, H., Almeida, J. S. and Ramirez, M. (2006). Illustration of a common framework for relating multiple typing methods by application to macrolide-resistant *Streptococcus pyogenes*. *Journal of clinical microbiology* **44**, 2524–32. doi: 10.1128/JCM.02536-05.

Carruthers, V. B. (2002). Host cell invasion by the opportunistic pathogen *Toxoplasma gondii*. *Acta Tropica* **81**, 111–122. doi: 10.1016/S0001-706X(01)00201-7.

Cavalcante, A. C. R., Ferreira, A. M., Melo, M. N., Fux, B., Brandão, G. P. and Vitor, R. W. A. (2007). Virulence and molecular characterization of *Toxoplasma gondii* isolated from goats in Ceará, Brazil. *Small Ruminant Research* **69**, 79–82. doi: 10.1016/J.SMALLRUMRES.2005.12.023.

Chanton-Greutmann, H., Thoma, R., Corboz, L., Borel, N. and Pospischil, A. (2002). Abortion in small ruminants in Switzerland: investigations during two lambing seasons (1996-1998) with special regard to chlamydial abortions. *Schweizer Archiv für Tierheilkunde* **144**, 483–492. doi: 10.1024/0036-7281.144.9.483.

Cheng, W., Liu, F., Li, M., Hu, X., Chen, H., Pappoe, F., Luo, Q., Wen, H., Xing, T., Xu, Y. and Shen, J. (2015). Variation detection based on next-generation sequencing of type Chinese 1 strains of *Toxoplasma gondii* with different virulence from China. *BMC Genomics* **16**, 1–9. doi: 10.1186/s12864-015-2106-z.

Coelho, C., Vieira-Pinto, M., Faria, A. S., Vale-Gonçalves, H., Veloso, O., Paiva-Cardoso, M. das N., Mesquita, J. R. and Lopes, A. P. (2014). Serological evidence of Toxoplasma gondii in hunted wild boar from Portugal. *Veterinary Parasitology* **202**, 310–312. doi: 10.1016/j.vetpar.2014.03.013.

Coelho, C., Lopes, A. P., Mesquita, J. R., Cardoso, L. and Vieira-Pinto, M. (2015). *Toxoplasma gondii* Infection in Hunted Wild Boars (*Sus scrofa*): Heart Meat Juice as an Alternative Sample to Serum for the Detection of Antibodies. *EcoHealth* **12**, 685–8. doi: 10.1007/s10393-015-1073-9.

Cook, A., Gilbert, R. E., Buffolano, W., Zufferey, J., Petersen, E., Jenum, P. A., Foulon, W., Semprini, A. and Dunn, D. (2000). Sources of toxoplasma infection in pregnant women: European multicentre case-control study. European Research Network on Congenital Toxoplasmosis. *BMJ (Clinical research ed.)* **321**, 142–7. doi: 10.1136/bmj.321.7254.142.

Cowper, B., Matthews, S. and Tomley, F. (2012). The molecular basis for the distinct host and tissue tropisms of coccidian parasites. *Molecular and Biochemical Parasitology* **186**, 1–10. doi: 10.1016/J.MOLBIOPARA.2012.08.007.

Cristina, N., Dardé, M. L., Boudin, C., Tavernier, G., Pestre-Alexandre, M. and Ambroise-Thomas, P. (1995). A DNA fingerprinting method for individual characterization of *Toxoplasma gondii* strains: combination with isoenzymatic characters for determination of linkage groups. *Parasitology Researche* **81**, 32–7.

Dardé, M. L. (2008). *Toxoplasma gondii* , "new" genotypes and virulence. *Parasite* **15**, 366–371. doi: 10.1051/parasite/2008153366.

Dardé, M. L., Pestre-Alexandre, M. and Bouteille, B. (1988). Isoenzymic characterization of seven strains of *Toxoplasma gondii* by isoelectrofocusing in polyacrylamide gels. *The American Journal of Tropical Medicine and Hygiene* **39**, 551–558. doi: 10.4269/ajtmh.1988.39.551.

Dardé, M. L., Bouteille, B. and Pestre-Alexandre, M. (1992). Isoenzyme analysis of 35 *Toxoplasma gondii* isolates and the biological and epidemiological implications. *The Journal of parasitology* **78**, 786–94.

Dardé, M. L., Villena, I., Pinon, J. M. and Beguinot, I. (1998). Severe toxoplasmosis caused by a *Toxoplasma gondii* strain with a new isoenzyme type acquired in French Guyana. *Journal of clinical microbiology* **36**, 324.

Dass, S. A. H., Vasudevan, A., Dutta, D., Soh, L. J. T., Sapolsky, R. M. and Vyas, A. (2011). Protozoan parasite *Toxoplasma gondii* manipulates mate choice in rats by enhancing attractiveness of males. *PloS one* **6**, e27229. doi: 10.1371/journal.pone.0027229.

De Craeye, S., Speybroeck, N., Ajzenberg, D., Dardé, M. L., Collinet, F., Tavernier, P., Van Gucht, S., Dorny, P. and Dierick, K. (2011). *Toxoplasma gondii* and Neospora caninum in wildlife: Common parasites in Belgian foxes and Cervidae? *Veterinary Parasitology* **178**, 64–69. doi: 10.1016/j.vetpar.2010.12.016.

de Melo Ferreira, A., Vitor, R. W. A., Gazzinelli, R. T. and Melo, M. N. (2006). Genetic analysis of natural recombinant Brazilian *Toxoplasma gondii* strains by multilocus PCR–RFLP. *Infection, Genetics and Evolution* **6**, 22–31. doi: 10.1016/J.MEEGID.2004.12.004.

Delhaes, L., Ajzenberg, D., Sicot, B., Bourgeot, P., Dardé, M.-L., Dei-Cas, E. and Houfflin-Debarge, V. (2010). Severe congenital toxoplasmosis due to a *Toxoplasma gondii* strain with an atypical genotype: case report and review. *Prenatal Diagnosis* **30**, 902–905. doi: 10.1002/pd.2563.

Demar, M., Daigre, J.-L., Peneau, C., Aznar, C., Panchoe, D., Cottrelle, B., Darde, M.-L., Valery, N., Maubon, D., Ajzenberg, D., Carme, B., Punwasi, W., Djossou, F. and Terzan, L. (2007). Fatal Outbreak of Human Toxoplasmosis along the Maroni River: Epidemiological, Clinical, and Parasitological Aspects. *Clinical Infectious Diseases* **45**, e88–e95. doi: 10.1086/521246.

Demar, M., Hommel, D., Djossou, F., Peneau, C., Boukhari, R., Louvel, D., Bourbigot, A.-M., Nasser, V., Ajzenberg, D., Darde, M.-L. and Carme, B. (2012). Acute toxoplasmoses in immunocompetent patients hospitalized in an intensive care unit in French Guiana. *Clinical Microbiology and Infection* **18**, E221–E231. doi: 10.1111/J.1469-0691.2011.03648.X.

Duarte, A., Castro, I., da Fonseca, I. M. P., Almeida, V., de Carvalho, L. M. M., Meireles, J., Fazendeiro, M. I., Tavares, L. and Vaz, Y. (2010). Survey of infectious and parasitic diseases in stray cats at the Lisbon Metropolitan Area, Portugal. *Journal of Feline Medicine and Surgery* **12**, 441–446. doi: 10.1016/j.jfms.2009.11.003.

Dubey, J. P. (1995). Duration of immunity to shedding of Toxoplasma gondii oocysts by cats. *The Journal of parasitology* **81**, 410–5.

Dubey, J. P. (1997). 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. (1998a). *Refinement of pepsin digestion method for isolation of Toxoplasma gondii from infected tissues*. doi: 10.1016/S0304-4017(97)00135-0.

Dubey, J. P. (1998b). Advances in the life cycle of *Toxoplasma gondii*. *International journal for parasitology* **28**, 1019–24.

Dubey, J. P. (1998c). *Toxoplasma gondii* oocyst survival under defined temperatures. *The Journal of parasitology* **84**, 862–5.

Dubey, J. P. (2009). Toxoplasmosis in pigs—The last 20 years. *Veterinary Parasitology* **164**, 89–103. doi: 10.1016/j.vetpar.2009.05.018.

Dubey, J. P. (2010a). *Toxoplasmosis of animals and humans*, 2nd Edn. ed. Vectors, P. CRC Press.

Dubey, J. P. (2010b). *Toxoplasma gondii* Infections in Chickens (*Gallus domesticus*): Prevalence, Clinical Disease, Diagnosis and Public Health Significance. *Zoonoses and Public Health* **57**, 60–73. doi: 10.1111/j.1863-2378.2009.01274.x.

Dubey, J. P. and Jones, J. L. (2008). *Toxoplasma gondii* infection in humans and animals in the United States.

International Journal for Parasitology **38**, 1257–1278. doi: 10.1016/J.IJPARA.2008.03.007.

Dubey, J. and Kirkbride, C. (1989). Economic and public health considerations of congenital toxoplasmosis in lambs. *Journal of the American Veterinary Medical Association* **195**, 1715–6.

Dubey, J. P. and Su, C. (2009). Population biology of *Toxoplasma gondii*: what's out and where did they come from. *Memorias do Instituto Oswaldo Cruz* **104**, 190–5. doi: 10.1590/s0074-02762009000200011.

Dubey, J. P. and Thulliez, P. (1993). Persistence of tissue cysts in edible tissues of cattle fed *Toxoplasma gondii* oocysts. *American journal of veterinary research* **54**, 270–3.

Dubey, J. P., Christie, E. and Pappas, P. W. (1977). Characterization of *Toxoplasma gondii* from the feces of naturally infected cats. *The Journal of infectious diseases* **136**, 432–5.

Dubey, J. P., Weigel, R. M., Siegel, A. M., Thulliez, P., Kitron, U. D., Mitchell, M. A., Mannelli, A., Mateus-Pinilla, N. E., Shen, S. K., Kwok, O. C. H. and Todd, K. S. (1995). Sources and reservoirs of *Toxoplasma gondii* infection on 47 swine farms in Illinois. *The Journal of Parasitology* **81**, 723. doi: 10.2307/3283961.

Dubey, J . ., Graham, D. ., Blackston, C. ., Lehmann, T., Gennari, S. ., Ragozo, A. M. ., Nishi, S. ., Shen, S. ., Kwok, O. C. ., Hill, D. . and Thulliez, P. (2002). Biological and genetic characterisation of *Toxoplasma gondii* isolates from chickens (Gallus domesticus) from São Paulo, Brazil: unexpected findings. *International Journal for Parasitology* **32**, 99–105. doi: 10.1016/S0020-7519(01)00364-2.

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. B., Shen, S. K., Kwok, O. C. H., Sumners, J. A., Demarais, S., Humphreys, J. G. and Lehmann, T. (2004a). Molecular and biologic characteristics of *Toxoplasma gondii* isolates from wildlife in the United States. *Journal of Parasitology* **90**, 67–71. doi: 10.1645/GE-110R.

Dubey, J. P., Navarro, I. T., Sreekumar, C., Dahl, E., Freire, R. L., Kawabata, H. H., Vianna, M. C. B., Kwok, O. C. H., Shen, S. K., Thulliez, P. and Lehmann, T. (2004b). *Toxoplasma gondii* infections in cats from Paraná, Brazil: seroprevalence, tissue distribution, and biologic and genetic characterization of isolates. *The Journal of parasitology* **90**, 721–6. doi: 10.1645/GE-382R.

Dubey, J. P., Gomez-Marin, J. E., Bedoya, A., Lora, F., Vianna, M. C. B., Hill, D., Kwok, O. C. H., Shen, S. K., Marcet, P. L. and Lehmann, T. (2005a). Genetic and biologic characteristics of *Toxoplasma gondii* isolates in free-range chickens from Colombia, South America. *Veterinary Parasitology* **134**, 67–72. doi: 10.1016/J.VETPAR.2005.07.013.

Dubey, J. P., Edelhofer, R., Marcet, P., Vianna, M. C. B., Kwok, O. C. H. and Lehmann, T. (2005b). Genetic and biologic characteristics of *Toxoplasma gondii* infections in free-range chickens from Austria. *Veterinary Parasitology* **133**, 299–306. doi: 10.1016/j.vetpar.2005.06.006.

Dubey, J. P., Lenhart, A., Castillo, C. E., Alvarez, L., Marcet, P., Sreekumar, C. and Lehmann, T. (2005c). *Toxoplasma gondii* infections in chickens from Venezuela: isolation, tissue distribution, and molecular characterization. *Journal of Parasitology* **91**, 1332–1334. doi: 10.1645/GE-500R.1.

Dubey, J. P., Marcet, P. L. and Lehmann, T. (2005d). Characterization of *Toxoplasma gondii* isolates in free-range chickens from Argentina. *Journal of Parasitology* **91**, 1335–1339. doi: 10.1645/GE-575R.1.

Dubey, J. P., Vianna, M. C. B., Sousa, S., Canada, N., Meireles, S., Correia da Costa, J. M., Marcet, P. L., Lehmann, T., Dardé, M. L. and Thulliez, P. (2006). Characterization of *Toxoplasma gondii* isolates in free-range chickens from Portugal. *Journal of Parasitology* **92**, 184–186. doi: 10.1645/GE-652R.1.

Dubey, J. P., López-Torres, H. Y., Sundar, N., Velmurugan, G. V, Ajzenberg, D., Kwok, O. C. H., Hill, R., Dardé, M. L. and Su, C. (2007a). Mouse-virulent Toxoplasma gondii isolated from feral cats on Mona Island, Puerto Rico. *The Journal of parasitology* **93**, 1365–9. doi: 10.1645/GE-1409.1.

Dubey, J. P., Sundar, N., Gennari, S. M., Minervino, A. H. H., Farias, N. A. da R., Ruas, J. L., dos Santos, T. R. B., Cavalcante, G. T., Kwok, O. C. H. and Su, C. (2007b). Biologic and genetic comparison of Toxoplasma gondii isolates in free-range chickens from the northern Pará state and the southern state Rio Grande do Sul, Brazil revealed highly diverse and distinct parasite populations. *Veterinary Parasitology* **143**, 182–188. doi: 10.1016/j.vetpar.2006.08.024.

Dubey, J. P., Zhu, X. Q., Sundar, N., Zhang, H., Kwok, O. C. H. and Su, C. (2007c). Genetic and biologic characterization of *Toxoplasma gondii* isolates of cats from China. *Veterinary Parasitology* **145**, 352–356. doi: 10.1016/j.vetpar.2006.12.016.

Dubey, J. P., Cortés-Vecino, J. A., Vargas-Duarte, J. J., Sundar, N., Velmurugan, G. V., Bandini, L. M., Polo, L. J., Zambrano, L., Mora, L. E., Kwok, O. C. H., Smith, T. and Su, C. (2007d). Prevalence of Toxoplasma gondii in dogs from Colombia, South America and genetic characterization of *T. gondii* isolates. *Veterinary Parasitology* **145**, 45–50. doi: 10.1016/j.vetpar.2006.12.001.

Dubey, J. P., Huong, L. T. T., Sundar, N. and Su, C. (2007e). Genetic characterization of Toxoplasma gondii isolates in dogs from Vietnam suggests their South American origin. *Veterinary Parasitology* **146**, 347–351. doi: 10.1016/j.vetpar.2007.03.008.

Dubey, J. P., Applewhaite, L., Sundar, N., Velmurugan, G. V, Bandini, L. a, Kwok, O. C. H., Hill, R. and Su, C. (2007f). 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**, 1559–65. doi: 10.1017/S0031182007003083.

Dubey, J. P., Sundar, N., Hill, D., Velmurugan, G. V., Bandini, L. A., Kwok, O. C. H., Majumdar, D. and Su, C. (2008). High prevalence and abundant atypical genotypes of Toxoplasma gondii isolated from lambs destined for human consumption in the USA. *International Journal for Parasitology* **38**, 999–1006. doi: 10.1016/j.ijpara.2007.11.012.

Dubey, J. P., Velmurugan, G. V., Rajendran, C., Yabsley, M. J., Thomas, N. J., Beckmen, K. B., Sinnett, D., Ruid, D., Hart, J., Fair, P. A., McFee, W. E., Shearn-Bochsler, V., Kwok, O. C. H., Ferreira, L. R., Choudhary, S., Faria, E. B., Zhou, H., Felix, T. A. and Su, C. (2011a). Genetic characterisation of Toxoplasma gondii in wildlife from North America revealed widespread and high prevalence of the fourth clonal type. *International Journal for Parasitology* **41**, 1139–1147. doi: 10.1016/j.ijpara.2011.06.005.

Dubey, J. P., Rajendran, C., Ferreira, L. R., Martins, J., Kwok, O. C. H., Hill, D. E., Villena, I., Zhou, H., Su, C. and Jones, J. L. (2011b). High prevalence and genotypes of *Toxoplasma gondii* isolated from goats, from a retail meat store, destined for human consumption in the USA. *International Journal for Parasitology* **41**, 827–833. doi: 10.1016/j.ijpara.2011.03.006.

Dubey, J. P., Hill, D. E., Rozeboom, D. W., Rajendran, C., Choudhary, S., Ferreira, L. R., Kwok, O. C. H. and Su, C. (2012). High prevalence and genotypes of *Toxoplasma gondii* isolated from organic pigs in northern USA. *Veterinary Parasitology* **188**, 14–18. doi: 10.1016/j.vetpar.2012.03.008.

Dumètre, A., Ajzenberg, D., Rozette, L., Mercier, A. and Dardé, M.-L. (2006). *Toxoplasma gondii* infection in sheep from Haute-Vienne, France: Seroprevalence and isolate genotyping by microsatellite analysis. *Veterinary Parasitology* **142**, 376–379. doi: 10.1016/j.vetpar.2006.07.005.

El Behairy, A. M., Choudhary, S., Ferreira, L. R., Kwok, O. C. H., Hilali, M., Su, C. and Dubey, J. P. (2013). Genetic characterization of viable *Toxoplasma gondii* isolates from stray dogs from Giza, Egypt. *Veterinary Parasitology* **193**, 25–29. doi: 10.1016/j.vetpar.2012.12.007.

Elbez-Rubinstein, A., Ajzenberg, D., Dardé, M., Cohen, R., Dumètre, A., Yera, H., Gondon, E., Janaud, J. and Thulliez, P. (2009). Congenital Toxoplasmosis and Reinfection during Pregnancy: Case Report, Strain Characterization, Experimental Model of Reinfection, and Review. *The Journal of Infectious Diseases* **199**, 280–285. doi: 10.1086/595793.

Eldeek, H. E. M., Ahmad, A. A. R., El-Mokhtar, M. A., Abdel Kader, A. R. M. M., Mandour, A. M. and Mounib, M. E. M. (2017). Toxoplasma genotyping in congenital toxoplasmosis in Upper Egypt: evidence of type I strain. *Parasitology Research* **116**, 2393–2406. doi: 10.1007/s00436-017-5541-8.

Elmore, S. A., Jones, J. L., Conrad, P. A., Patton, S., Lindsay, D. S. and Dubey, J. P. (2010). *Toxoplasma gondii*: epidemiology, feline clinical aspects, and prevention. *Trends in Parasitology* **26**, 190–196. doi: 10.1016/j.pt.2010.01.009.

Elsheikha, H. M. and Khan, N. A. (2010). Protozoa traversal of the blood–brain barrier to invade the central nervous system. *FEMS Microbiology Reviews* **34**, 532–553. doi: 10.1111/j.1574-6976.2010.00215.x.

Esteban-Redondo, I., Maley, S. ., Thomson, K., Nicoll, S., Wright, S., Buxton, D. and Innes, E. . (1999). Detection of *T. gondii* in tissues of sheep and cattle following oral infection. *Veterinary Parasitology* **86**, 155–171. doi: 10.1016/S0304-4017(99)00138-7.

Esteves, A., Araújo, M., Cardoso, L., Clemêncio, C. and Lopes, A. P. (2012). Seropositivity to *Toxoplasma gondii* in field chickens and rabbits slaughtered in slaughterhouses in the Centre of Portugal. *Acta Parasitológica*

Portuguesa **19**, 182–183.

Esteves, F., Aguiar, D., Rosado, J., Costa, M. L., de Sousa, B., Antunes, F. and Matos, O. (2014). *Toxoplasma gondii* prevalence in cats from Lisbon and in pigs from centre and south of Portugal. *Veterinary Parasitology* **200**, 8–12. doi: 10.1016/j.vetpar.2013.12.017.

Fazaeli, A. and Ebrahimzadeh, A. (2007). A new perspective on and re-assessment of SAG2 locus as the tool for genetic analysis of *Toxoplasma gondii* isolates. *Parasitology Research* **101**, 99–104. doi: 10.1007/s00436-006-0449-8.

Fazaeli, A., Carter, P. E., Darde, M. L. and Pennington, T. H. (2000). Molecular typing of *Toxoplasma gondii* strains by GRA6 gene sequence analysis. *International Journal for Parasitology* **30**, 637–642. doi: 10.1016/S0020-7519(00)00036-9.

Fernández-Aguilar, X., Alzaga, V., Villanúa, D., Cabezón, O., García-Bocanegra, I., Dubey, J. P. and Almería, S. (2013). Epidemiology and prevalence of *Toxoplasma gondii* infection in the Iberian hare (*Lepus granatensis*). *Veterinary Parasitology* **196**, 194–198. doi: 10.1016/j.vetpar.2013.01.061.

Ferreira, A., Vitor, R., Carneiro, A., Brandão, G. and Melo, M. (2004). Genetic variability of Brazilian *Toxoplasma gondii* strains detected by random amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR) and simple sequence repeat anchored-PCR (SSR-PCR). *Infection, Genetics and Evolution* **4**, 131–142. doi: 10.1016/j.meegid.2004.03.002.

Ferreira, I. M. R., Vidal, J. E., de Mattos, C. de C. B., de Mattos, L. C., Qu, D., Su, C. and Pereira-Chioccola, V. L. (2011). *Toxoplasma gondii* isolates: Multilocus RFLP–PCR genotyping from human patients in Sao Paulo State, Brazil identified distinct genotypes. *Experimental Parasitology* **129**, 190–195. doi: 10.1016/J.EXPPARA.2011.06.002.

Formenti, N., Gaffuri, A., Trogu, T., Viganò, R., Ferrari, N. and Lanfranchi, P. (2016). Spread and genotype of *Toxoplasma gondii* in naturally infected alpine chamois (*Rupicapra r. rupicapra*). *Parasitology research* **115**, 2115–20. doi: 10.1007/s00436-016-4981-x.

Frenkel, J. K. and Dubey, J. P. (1973). Effects of freezing on the viability of toxoplasma oocysts. *The Journal of parasitology* **59**, 587–8.

Frenkel, J. K., Dubey, J. P. and Hoff, R. L. (1976). Loss of stages after continuous passage of *Toxoplasma gondii* and *Besnoitia jellisoni*. *The Journal of protozoology* **23**, 421–4.

Frenkel, J. K., Dubey, J. P. and Beattie, C. P. (2006). Toxoplasmosis of Animals and Man. *The Journal of Parasitology* **75**, 816. doi: 10.2307/3283074.

Freyre, A., Bonino, J., Falcón, J., Castells, D., Correa, O. and Casaretto, A. (1997). The incidence and economic significance of ovine toxoplasmosis in Uruguay. *Veterinary Parasitology* **73**, 13–15. doi: 10.1016/S0304-4017(97)00069-1.

Frölich, K., Wisser, J., Schmüser, H., Fehlberg, U., Neubauer, H., Grunow, R., Nikolaou, K., Priemer, J., Thiede, S., Streich, W. J. and Speck, S. (2003). Epizootiologic and ecologic investigations of European brown hares (*Lepus europaeus*) in selected populations from Schleswig-Holstein, Germany. *Journal of wildlife diseases* **39**, 751–61. doi: 10.7589/0090-3558-39.4.751.

Fuentes, I., Rubio, J. M., Ramírez, C. and Alvar, J. (2001). Genotypic characterization of *Toxoplasma gondii* strains associated with human toxoplasmosis in Spain: direct analysis from clinical samples. *Journal of clinical microbiology* **39**, 1566–70. doi: 10.1128/JCM.39.4.1566-1570.2001.

Gajadhar, A. A., Aramini, J. J., Tiffin, G. and Bisaillon, J. R. (1998). Prevalence of *Toxoplasma gondii* in Canadian market-age pigs. *The Journal of parasitology* **84**, 759–63.

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., Wang, H. and Brunk, B. P. (2008). ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic acids research* **36**, D553-6. doi: 10.1093/nar/gkm981.

Gargaté, M. J., Ferreira, I., Vilares, A., Martins, S., Cardoso, C., Silva, S., Nunes, B. and Gomes, J. P. (2016). Toxoplasma gondii seroprevalence in the Portuguese population: comparison of three cross-sectional studies spanning three decades. *BMJ open* **6**,. doi: 10.1136/bmjopen-2016-011648.

Gatkowska, J., Hiszczynska-Sawicka, E., Kur, J., Holec, L. and Dlugonska, H. (2006). *Toxoplasma gondii*: An evaluation of diagnostic value of recombinant antigens in a murine model. *Experimental Parasitology* **114**, 220–

227. doi: 10.1016/j.exppara.2006.03.011.

Gebremedhin, E. Z., Abdurahaman, M., Tessema, T. S., Tilahun, G., Cox, E., Goddeeris, B., Dorny, P., De Craeye, S., Dardé, M.-L. and Ajzenberg, D. (2014). Isolation and genotyping of viable Toxoplasma gondii from sheep and goats in Ethiopia destined for human consumption. *Parasites & vectors* **7**, 425. doi: 10.1186/1756-3305-7-425.

Gelanew, T., Hailu, A., Scho'nian, G., Lewis, M. D., Miles, M. A. and Yeo, M. (2014). Multilocus sequence and microsatellite identification of intra-specific hybrids and ancestor-like donors among natural ethiopian isolates of leishmania donovani. *International Journal for Parasitology*. doi: 10.1016/j.ijpara.2014.05.008.

Gilbert, R. E., Dunn, D. T., Lightman, S., Murray, P. I., Pavesio, C. E., Gormley, P. D., Masters, J., Parker, S. P. and Stanford, M. R. (1999). Incidence of symptomatic toxoplasma eye disease: aetiology and public health implications. *Epidemiology and infection* **123**, 283–9. doi: 10.1017/S0950268899002800.

Godoi, F. S. L. de, Nishi, S. M., Pena, H. F. de J. and Gennari, S. M. (2010). *Toxoplasma gondii*: diagnosis of experimental and natural infection in pigeons (*Columba livia*) by serological, biological and molecular techniques. *Revista Brasileira de Parasitologia Veterinaria* **19**, 238–43. doi: 10.1590/S1984-29612010000400009.

Golkar, M., Azadmanesh, K., Khalili, G., Khoshkholgh-Sima, B., Babaie, J., Mercier, C., Brenier-Pinchart, M. P., Fricker-Hidalgo, H., Pelloux, H. and Cesbron-Delauw, M. F. (2008). Serodiagnosis of recently acquired Toxoplasma gondii infection in pregnant women using enzyme-linked immunosorbent assays with a recombinant dense granule GRA6 protein. *Diagnostic Microbiology and Infectious Disease* **61**, 31–39. doi: 10.1016/j.diagmicrobio.2007.09.003.

Grigg, M. E. and Sundar, N. (2009). Sexual recombination punctuated by outbreaks and clonal expansions predicts *Toxoplasma gondii* population genetics. *International journal for parasitology* **39**, 925. doi: 10.1016/J.IJPARA.2009.02.005.

Grigg, M. E. and Suzuki, Y. (2003). Sexual recombination and clonal evolution of virulence in Toxoplasma. *Microbes and Infection* **5**, 685–690. doi: 10.1016/S1286-4579(03)00088-1.

Grigg, M. E., Ganatra, J., Boothroyd, J. C. and Margolis, T. P. (2001a). Unusual abundance of atypical strains associated with human ocular Toxoplasmosis. *The Journal of Infectious Diseases* **184**, 633–639. doi: 10.1086/322800.

Grigg, M. E., Bonnefoy, S., Hehl, A. B., Suzuki, Y. and Boothroyd, J. C. (2001b). Success and virulence in Toxoplasma as the result of sexual recombination between two distinct ancestries. *Science* **294**, 161–165. doi: 10.1126/science.1061888.

Groër, M. W., Yolken, R. H., Xiao, J.-C., Beckstead, J. W., Fuchs, D., Mohapatra, S. S., Seyfang, A. and Postolache, T. T. (2011). Prenatal depression and anxiety in *Toxoplasma gondii*-positive women. *American journal of obstetrics and gynecology* **204**, 433.e1–7. doi: 10.1016/j.ajog.2011.01.004.

Guo, Z. G. and Johnson, A. M. (1995). Genetic characterization of *Toxoplasma gondii* strains by random amplified polymorphic DNA polymerase chain reaction. *Parasitology* **111 (Pt 2**, 127–32.

Guo, Z. G., Gross, U. and Johnson, A. M. (1997). *Toxoplasma gondii* virulence markers identified by random amplified polymorphic DNA polymerase chain reaction. *Parasitology research* **83**, 458–63.

Gustafsson, K. and Uggla, A. (1994). Serologic survey for *Toxoplasma gondii* infection in the brown hare (*Lepus europaeus*) in Sweden. *Journal of wildlife diseases* **30**, 201–4. doi: 10.7589/0090-3558-30.2.201.

Hall, S., Pandit, A., Golwilkar, A. and Williams, T. (1999). How do Jains get toxoplasma infection? *The Lancet* **354**, 486–487. doi: 10.1016/S0140-6736(99)02587-8.

Halonen, S. K. and Weiss, L. M. (2013). Toxoplasmosis. *Handbook of Clinical Neurology* **114**, 125–145. doi: 10.1016/B978-0-444-53490-3.00008-X.

Halos, L., Thébault, A., Aubert, D., Thomas, M., Perret, C., Geers, R., Alliot, A., Escotte-Binet, S., Ajzenberg, D., Dardé, M.-L., Durand, B., Boireau, P. and Villena, I. (2010). An innovative survey underlining the significant level of contamination by *Toxoplasma gondii* of ovine meat consumed in France. *International Journal for Parasitology* **40**, 193–200. doi: 10.1016/j.ijpara.2009.06.009.

Haroon, F., Händel, U., Angenstein, F., Goldschmidt, J., Kreutzmann, P., Lison, H., Fischer, K.-D., Scheich, H., Wetzel, W., Schlüter, D. and Budinger, E. (2012). *Toxoplasma gondii* actively inhibits neuronal function in

chronically infected mice. *PloS one* **7**, e35516. doi: 10.1371/journal.pone.0035516.

Hejlícek, K., Literák, I. and Lhoták, M. (1995). Prevalence of antibodies to *Toxoplasma gondii* in army dogs in the Czech Republic and Slovak Republic. *Veterinarni medicina* **40**, 137–40.

Herrmann, D. C., Pantchev, N., Vrhovec, G. G., Barutzki, D., Wilking, H., Fröhlich, A., Lüder, C. G. K., Conraths, F. J. and Schares, G. (2010). Atypical *Toxoplasma gondii* genotypes identified in oocysts shed by cats in Germany. *International Journal for Parasitology* **40**, 285–292. doi: 10.1016/j.ijpara.2009.08.001.

Herrmann, D. C., Maksimov, P., Maksimov, A., Sutor, A., Schwarz, S., Jaschke, W., Schliephake, A., Denzin, N., Conraths, F. J. and Schares, G. (2012a). *Toxoplasma gondii* in foxes and rodents from the German Federal States of Brandenburg and Saxony-Anhalt: Seroprevalence and genotypes. *Veterinary Parasitology* **185**, 78–85. doi: 10.1016/j.vetpar.2011.10.030.

Herrmann, D. C., Conraths, F. J., Pantchev, N., Vrhovec, M. G., Schares, G., Maksimov, A. and Bärwald, A. (2012b). *Toxoplasma gondii* sexual cross in a single naturally infected feline host: Generation of highly mouse-virulent and avirulent clones, genotypically different from clonal types I, II and III. *Veterinary Research* **43**, 39. doi: 10.1186/1297-9716-43-39.

Herrmann, D. C., Wibbelt, G., Götz, M., Conraths, F. J. and Schares, G. (2013). Genetic characterisation of *Toxoplasma gondii* isolates from European beavers (*Castor fiber*) and European wildcats (*Felis silvestris silvestris*). *Veterinary Parasitology* **191**, 108–111. doi: 10.1016/J.VETPAR.2012.08.026.

Herrmann, D. C., Maksimov, P., Hotop, A., Groß, U., Däubener, W., Liesenfeld, O., Pleyer, U., Conraths, F. J. and Schares, G. (2014). Genotyping of samples from German patients with ocular, cerebral and systemic toxoplasmosis reveals a predominance of Toxoplasma gondii type II. *International Journal of Medical Microbiology* **304**, 911–916. doi: 10.1016/j.ijmm.2014.06.008.

Hoffmann, S., Batz, M. B. and Morris, J. J. G. (2012). Annual Cost of Illness and Quality-Adjusted Life Year Losses in the United States Due to 14 Foodborne Pathogens. *Journal of food protection* **75**, 1292–1302. doi: 10.4315/0362-028X.JFP-11-417.

Hohlfeld, P., Daffos, F., Thulliez, P., Aufrant, C., Couvreur, J., MacAleese, J., Descombey, D. and Forestier, F. (1989). Fetal toxoplasmosis: outcome of pregnancy and infant follow-up after in utero treatment. *The Journal of pediatrics* **115**, 765–9. doi: 10.1016/s0022-3476(89)80660-2.

Honoré, S., Couvelard, A., Garin, Y. J., Bedel, C., Hénin, D., Dardé, M. L. and Derouin, F. (2000). Genotyping of *Toxoplasma gondii strains from immunocompromised patients. Pathologie-biologie* **48**, 541–7.

Horacek, J., Flegr, J., Tintera, J., Verebova, K., Spaniel, F., Novak, T., Brunovsky, M., Bubenikova-Valesova, V., Holub, D., Palenicek, T. and Höschl, C. (2012). Latent toxoplasmosis reduces gray matter density in schizophrenia but not in controls: Voxel-based-morphometry (VBM) study. *The World Journal of Biological Psychiatry* **13**, 501–509. doi: 10.3109/15622975.2011.573809.

Hosseini, S. A., Amouei, A., Sharif, M., Sarvi, S., Galal, L., Javidnia, J., Pagheh, A. S., Gholami, S., Mizani, A. and Daryani, A. (2019). Human toxoplasmosis: a systematic review for genetic diversity of *Toxoplasma gondii* in clinical samples. *Epidemiology and Infection* **147**, e36. doi: 10.1017/S0950268818002947.

House, P. K., Vyas, A. and Sapolsky, R. (2011). Predator cat odors activate sexual arousal pathways in brains of Toxoplasma gondii infected rats. *PLoS ONE* **6**,. doi: 10.1371/JOURNAL.PONE.0023277.

Howe, D. K. and Sibley, L. D. (1995). *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *Journal of Infectious Diseases* **172**, 1561–1566. doi: 10.1093/infdis/172.6.1561.

Howe, D. K., Honoré, S., Derouin, F. and Sibley, L. D. (1997). Determination of genotypes of *Toxoplasma gondii* strains isolated from patients with toxoplasmosis. *Journal of clinical microbiology* **35**, 1411–4.

Hůrková, L. and Modrý, D. (2006). *PCR detection of Neospora caninum, Toxoplasma gondii and Encephalitozoon cuniculi in brains of wild carnivores*. doi: 10.1016/j.vetpar.2006.01.005.

Jacquet, A., Coulon, L., De Nève, J., Daminet, V., Haumont, M., Garcia, L., Bollen, A., Jurado, M. and Biemans, R. (2001). The surface antigen SAG3 mediates the attachment of *Toxoplasma gondii* to cell-surface proteoglycans. *Molecular and Biochemical Parasitology* **116**, 35–44. doi: 10.1016/S0166-6851(01)00297-3.

Jewell, M. L., Frenkel, J. K., Johnson, K. M., Reed, V. and Ruiz, A. (1972). Development of Toxoplasma oocysts in neotropical felidae. *The American journal of tropical medicine and hygiene* **21**, 512–7.

Jokelainen, P., Isomursu, M., Näreaho, A. and Oksanen, A. (2011). Natural *Toxoplasma gondii* infections in

European Brown Hares and Mountain Hares in Filand: Proportional mortality rate, antibody prevalence, and genetic characterization. *Journal of Wildlife Diseases* **47**, 154–163. doi: 10.7589/0090-3558-47.1.154.

Jokelainen, P., Näreaho, A., Hälli, O., Heinonen, M. and Sukura, A. (2012a). Farmed wild boars exposed to *Toxoplasma gondii* and *Trichinella* spp. *Veterinary Parasitology* **187**, 323–327. doi: 10.1016/J.VETPAR.2011.12.026.

Jokelainen, P., Simola, O., Rantanen, E., Näreaho, A., Lohi, H. and Sukura, A. (2012b). Feline toxoplasmosis in Finland. *Journal of Veterinary Diagnostic Investigation* **24**, 1115–1124. doi: 10.1177/1040638712461787.

Jones, T. C., Yeh, S. and Hirsch, J. G. (1972). The interaction between Toxoplasma gondii and mammalian cells. I. Mechanism of entry and intracellular fate of the parasite. *The Journal of experimental medicine* **136**, 1157–72. doi: 10.1084/jem.136.5.1157.

Jones, C. D., Okhravi, N., Adamson, P., Tasker, S. and Lightman, S. (2000). Comparison of PCR detection methods for B1, P30, and 18S rDNA genes of *T. gondii* in aqueous humor. *Investigative ophthalmology & visual science* **41**, 634–44.

Jones, J. L., Dietz, V. J., Power, M., Lopez, A., Wilson, M., Navin, T. R., Gibbs, R. and Schulkin, J. (2001). Survey of Obstetrician-Gynecologists in the United States About Toxoplasmosis. *Infectious Diseases in Obstetrics and Gynecology* **9**, 23–31. doi: 10.1155/S1064744901000059.

Jones, J. L., Muccioli, C., Belfort, R., Holland, G. N., Roberts, J. M., Silveira, C. and Silveira, C. (2006a). Recently acquired *Toxoplasma gondii* infection, Brazil. *Emerging infectious diseases* **12**, 582–7. doi: 10.3201/eid1204.051081.

Jones, L. A., Alexander, J. and Roberts, C. W. (2006b). Ocular toxoplasmosis: in the storm of the eye. *Parasite Immunology* **28**, 635–642. doi: 10.1111/j.1365-3024.2006.00874.x.

Joynson, D. H. M. (2003). Congenital toxoplasma infection in the UK. *Archives de pediatrie : organe officiel de la Societe francaise de pediatrie* **10 Suppl 1**, 27–8.

Jungersen, G., Jensen, L., Rask, M. R. and Lind, P. (2002). Non-lethal infection parameters in mice separate sheep Type II *Toxoplasma gondii* isolates by virulence. *Comparative Immunology, Microbiology and Infectious Diseases* **25**, 187–195. doi: 10.1016/S0147-9571(01)00039-X.

Kapperud, G. (1978). Survey for toxoplasmosis in wild and domestic animals from Norway and Sweden. *Journal of wildlife diseases* **14**, 157–62.

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. and Sibley, L. D. (2005a). Composite genome map and recombination parameters derived from three archetypal lineages of *Toxoplasma gondii*. *Nucleic Acids Research* **33**, 2980–2992. doi: 10.1093/nar/gki604.

Khan, A., Su, C., German, M., Storch, G. A., Clifford, D. B. and Sibley, L. D. (2005b). Genotyping of *Toxoplasma gondii* Strains from Immunocompromised Patients Reveals High Prevalence of Type I Strains. *Journal of Clinical Microbiology* **43**, 5881–5887. doi: 10.1128/JCM.43.12.5881-5887.2005.

Khan, A., Fux, B., Su, C., Dubey, J. P., Darde, M. L., Ajioka, J. W., Rosenthal, B. M. and Sibley, L. D. (2007). Recent transcontinental sweep of *Toxoplasma gondii* driven by a single monomorphic chromosome. *Proceedings of the National Academy of Sciences* **104**, 14872–14877. doi: 10.1073/pnas.0702356104.

Khan, A., Dubey, J. P., Su, C., Ajioka, J. W., Rosenthal, B. M. and Sibley, L. D. (2011a). Genetic analyses of atypical *Toxoplasma gondii* strains reveal a fourth clonal lineage in North America. *International Journal for Parasitology* **41**, 645–655. doi: 10.1016/j.ijpara.2011.01.005.

Khan, A., Miller, N., Roos, D. S., Dubey, J. P., Ajzenberg, D., Dardé, M. L., Ajioka, J. W., Rosenthal, B. and Sibleya, L. D. (2011b). A monomorphic haplotype of chromosome Ia is associated with widespread success in clonal and nonclonal populations of *Toxoplasma gondii*. *mBio* **2**, 1–10.

Khan, A., Ajzenberg, D., Mercier, A., Demar, M., Simon, S., Dardé, M. L., Wang, Q., Verma, S. K., Rosenthal, B. M., Dubey, J. P. and Sibley, L. D. (2014). Geographic separation of domestic and wild strains of *Toxoplasma gondii* in French Guianacorrelates with a monomorphic version of chromosome1a. *PLoS Neglected Tropical Diseases* **8**, e3182. doi: 10.1371/journal.pntd.0003182.

Kim, K. and Weiss, L. M. (2004). *Toxoplasma gondii*: the model apicomplexan. *International journal for parasitology* **34**, 423. doi: 10.1016/J.IJPARA.2003.12.009.

Kim, K. and Weiss, L. M. (2008). Toxoplasma: the next 100years. *Microbes and infection* **10**, 978–84. doi: 10.1016/j.micinf.2008.07.015.

Kirkpatrick, C. E., Colvin, B. A. and Dubey, J. P. (1990). *Toxoplasma gondii* antibodies in common barn-owls (*Tyto alba*) and pigeons (*Columba livia*) in New Jersey. *Veterinary Parasitology* **36**, 177–180. doi: 10.1016/0304-4017(90)90108-N.

Kong, J., Grigg, M. E., Uyetake, L., Parmley, S. and Boothroyd, J. C. (2003). Serotyping of *Toxoplasma gondii* infections in humans using synthetic peptides. *The Journal of Infectious Diseases* **187**, 1484–1495. doi: 10.1086/374647.

Kornacka, A., Cybulska, A., Bień, J., Goździk, K. and Moskwa, B. (2016). The usefulness of direct agglutination test, enzyme-linked immunosorbent assay and polymerase chain reaction for the detection of *Toxoplasma gondii* in wild animals. *Veterinary Parasitology* **228**, 85–89. doi: 10.1016/J.VETPAR.2016.08.010.

Kulasena, V. A., Rajapakse, R. P. V. J., Dubey, J. P., Dayawansa, P. N. and Premawansa, S. (2011). Seroprevalence of *Toxoplasma gondii* in cats from Colombo, Sri Lanka. *The Journal of parasitology* **97**, 152. doi: 10.1645/GE-2640.1.

Lass, A., Pietkiewicz, H., Modzelewska, E., Dumètre, A., Szostakowska, B. and Myjak, P. (2009). Detection of *Toxoplasma gondii* oocysts in environmental soil samples using molecular methods. *European Journal of Clinical Microbiology & Infectious Diseases* **28**, 599–605. doi: 10.1007/s10096-008-0681-5.

Lauron, E. J., Tolia, N. H., Jimah, J. R., Khan, A., Wang, Q., Behnke, M. S. and Sibley, L. D. (2015). Rhoptry proteins ROP5 and ROP18 are major murine virulence factors in genetically divergent South American strains of *Toxoplasma gondii*. *PLOS Genetics* **11**, e1005434. doi: 10.1371/journal.pgen.1005434.

Lehmann, T., Marcet, P. L., Graham, D. H., Dahl, E. R. and Dubey, J. P. (2006). Globalization and the population structure of *Toxoplasma gondii*. *Proceedings of the National Academy of Sciences* **103**, 11423–11428. doi: 10.1073/pnas.0601438103.

Leighty, J. C. (1990). Strategies for control of toxoplasmosis. *Journal of the American Veterinary Medical Association* **196**, 281–6.

Li, M., Mo, X. W., Wang, L., Chen, H., Luo, Q. L., Wen, H. Q., Wei, W., Zhang, A. M., Du, J., Lu, F. L., Lun, Z. R. and Shen, J. L. (2014). Phylogeny and virulence divergency analyses of *Toxoplasma gondii* isolates from China. *Parasites and Vectors* **7**, 1–12. doi: 10.1186/1756-3305-7-133.

Lindström, I., Sundar, N., Lindh, J., Kironde, F., Kabasa, J. D., Kwok, O. C. H., Dubey, J. P. and Smith, J. E. (2008). Isolation and genotyping of Toxoplasma gondii from Ugandan chickens reveals frequent multiple infections. *Parasitology* **135**, 39–45. doi: 10.1017/S0031182007003654.

Lopes, A. P., Cardoso, L. and Rodrigues, M. (2008). Serological survey of *Toxoplasma gondii* infection in domestic cats from northeastern Portugal. *Veterinary Parasitology* **155**, 184–189. doi: 10.1016/j.vetpar.2008.05.007.

Lopes, A. P., Sargo, R., Rodrigues, M. and Cardoso, L. (2011a). High seroprevalence of antibodies to *Toxoplasma gondii* in wild animals from Portugal. *Parasitology Research* **108**, 1163–1169. doi: 10.1007/s00436-010-2158-6.

Lopes, A. P., Santos, H., Neto, F., Rodrigues, M., Kwok, O. C. H., Dubey, J. P. and Cardoso, L. (2011b). Prevalence of Antibodies to *Toxoplasma gondii* in Dogs From Northeastern Portugal. *Journal of Parasitology* **97**, 418–420. doi: 10.1645/GE-2691.1.

Lopes, A. P., Dubey, J. P., Neto, F., Rodrigues, A., Martins, T., Rodrigues, M. and Cardoso, L. (2012a). Seroprevalence of *Toxoplasma gondii* infection in cattle, sheep, goats and pigs from the North of Portugal for human consumption. *Veterinary Parasitology* 2010–2013. doi: 10.1016/j.vetpar.2012.12.001.

Lopes, A. P., Dubey, J. P., Moutinho, O., Gargaté, M. J., Vilares, A., Rodrigues, M. and Cardoso, L. (2012b). Seroepidemiology of *Toxoplasma gondii* infection in women from the North of Portugal in their childbearing years. *Epidemiology and Infection* **140**, 872–877. doi: 10.1017/S0950268811001658.

Lopes, A. P., Sousa, S., Dubey, J. P., Ribeiro, A. J., Silvestre, R., Cotovio, M., Schallig, H. D. F. H., Cardoso, L. and Cordeiro-da-Silva, A. (2013). Prevalence of antibodies to *Leishmania infantum* and *Toxoplasma gondii* in horses from the north of Portugal. *Parasites & vectors* **6**, 178. doi: 10.1186/1756-3305-6-178.

Lopes, A. P., Dubey, J. P., Dardé, M.-L. and Cardoso, L. (2014). Epidemiological review of *Toxoplasma gondii* infection in humans and animals in Portugal. *Parasitology* **141**, 1699–1708. doi: 10.1017/S0031182014001413.

Lopes, A. P., Oliveira, A. C., Granada, S., Rodrigues, F. T., Papadopoulos, E., Schallig, H., Dubey, J. P. and

Cardoso, L. (2017). Antibodies to *Toxoplasma gondii* and *Leishmania* spp. in domestic cats from Luanda, Angola. *Veterinary Parasitology* **239**, 15–18. doi: 10.1016/j.vetpar.2017.04.009.

Lorenzi, H., Khan, A., Behnke, M. S., Namasivayam, S., Swapna, L. S., Hadjithomas, M., Karamycheva, S., Pinney, D., Brunk, B. P., Ajioka, J. W., Ajzenberg, D., Boothroyd, J. C., Boyle, J. P., Dardé, M. L., Diaz-Miranda, M. A., Dubey, J. P., Fritz, H. M., Gennari, S. M., Gregory, B. D., Kim, K., Saeij, J. P. J., Su, C., White, M. W., Zhu, X. Q., Howe, D. K., Rosenthal, B. M., Grigg, M. E., Parkinson, J., Liu, L., Kissinger, J. C., Roos, D. S. and Sibley, L. D. (2016). Local admixture of amplified and diversified secreted pathogenesis determinants shapes mosaic *Toxoplasma gondii* genomes. *Nature Communications* **7**,. doi: 10.1038/ncomms10147.

Luft, B. J. and Remington, J. S. (1992). Toxoplasmic encephalitis in AIDS. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **15**, 211–22.

Machacova, T., Bartova, E., Di Loria, A., Sedlak, K., Mariani, U., Fusco, G., Fulgione, D., Veneziano, V. and Dubey, J. P. (2014). Seroprevalence of *Toxoplasma gondii* in donkeys (*Equus asinus*) in Italy. *The Journal of veterinary medical science* **76**, 265–7. doi: 10.1292/jvms.13-0352.

Mahami-Oskouei, M., Moradi, M., Fallah, E., Hamidi, F. and Asl Rahnamaye Akbari, N. (2017). Molecular detection and genotyping of *Toxoplasma gondii* in chicken, beef, and lamb meat consumed in Northwestern Iran. *Iranian journal of parasitology* **12**, 38–45.

Maiden, M. C. J. (2006). Multilocus sequence typing of bacteria. *Annual Review of Microbiology* **60**, 561–588. doi: 10.1146/annurev.micro.59.030804.121325.

Maksimov, P., Zerweck, J., Dubey, J. P., Pantchev, N., Frey, C. F., Maksimov, A., Reimer, U., Schutkowski, M., Hosseininejad, M., Ziller, M., Conraths, F. J. and Schares, G. (2013). Serotyping of *Toxoplasma gondii* in cats (*Felis domesticus*) reveals predominance of type II infections in Germany. *PLoS ONE* **8**, 1–16. doi: 10.1371/journal.pone.0080213.

Mancianti, F., Nardoni, S., Ascenzi, C. D., Pedonese, F., Mugnaini, L., Franco, F. and Papini, R. (2013). Seroprevalence , detection of DNA in blood and milk , and genotyping of *Toxoplasma gondii* in a goat population in Italy. *Hindawi* **2013**, 1–6.

Masala, G., Porcu, R., Madau, L., Tanda, A., Ibba, B., Satta, G. and Tola, S. (2003). Survey of ovine and caprine toxoplasmosis by IFAT and PCR assays in Sardinia, Italy. *Veterinary Parasitology* **117**, 15–21. doi: 10.1016/J.VETPAR.2003.07.012.

Mason, S., Quinnell, R. J. and Smith, J. E. (2010). Detection of *Toxoplasma gondii* in lambs via PCR screening and serological follow-up. *Veterinary Parasitology* **169**, 258–263. doi: 10.1016/J.VETPAR.2010.01.021.

McLeod, R., Khan, A. R., Noble, G. A., Latkany, P., Jalbrzikowski, J., Boyer, K. and Toxoplasmosis Study Group (2006). Severe sulfadiazine hypersensitivity in a child with reactivated congenital toxoplasmic chorioretinitis. *The Pediatric Infectious Disease Journal* **25**, 270–272. doi: 10.1097/01.inf.0000202070.59190.9a.

Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. and Tauxe, R. V (1999). Food-related illness and death in the United States. *Emerging infectious diseases* **5**, 607–25. doi: 10.3201/eid0505.990502.

Mercier, C., Adjogble, K. D. Z., Däubener, W. and Delauw, M.-F.-C. (2005). Dense granules: Are they key organelles to help understand the parasitophorous vacuole of all apicomplexa parasites? *International Journal for Parasitology* **35**, 829–849. doi: 10.1016/J.IJPARA.2005.03.011.

Mercier, A., Devillard, S., Ngoubangoye, B., Bonnabau, H., Bañuls, A. L., Durand, P., Salle, B., Ajzenberg, D. and Dardé, M. L. (2010). Additional haplogroups of *Toxoplasma gondii* out of Africa: Population structure and mouse-virulence of strains from Gabon. *PLoS Neglected Tropical Diseases* **4**,. doi: 10.1371/journal.pntd.0000876.

Mercier, A., Ajzenberg, D., Devillard, S., Demar, M. P., de Thoisy, B., Bonnabau, H., Collinet, F., Boukhari, R., Blanchet, D., Simon, S., Carme, B. and Dardé, M. L. (2011). Human impact on genetic diversity of Toxoplasma gondii: Example of the anthropized environment from French Guiana. *Infection, Genetics and Evolution* **11**, 1378–1387. doi: 10.1016/j.meegid.2011.05.003.

Messaritakis, I., Detsika, M., Koliou, M., Sifakis, S. and Antoniou, M. (2008). Prevalent genotypes of *Toxoplasma*

gondii in pregnant women and patients from Crete and Cyprus. *American Journal of Tropical Medicine and Hygiene* **79**, 205–209.

Micic, D., Ristic, G., Markovic, M., Djurkovic-Djakovic, O., Stajner, T., Ivovic, V., Vasiljevic, Z., Vujic, D., Pasic, S. and Ajzenberg, D. (2013). Atypical strain of Toxoplasma gondii causing fatal reactivation after hematopoietic stem cell transplantion in a patient with an underlying immunological deficiency. *Journal of Clinical Microbiology* **51**, 2686–2690. doi: 10.1128/jcm.01077-13.

Miller, S., Dykes, D. and Polesky, H. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research* **16**, 1215. doi: 10.1093/nar/16.3.1215.

Mineo, J. R. and Kasper, L. H. (1994). Attachment of *Toxoplasma gondii* to host cells involves major surface protein, SAG-1 (P-30). *Experimental Parasitology* **79**, 11–20. doi: 10.1006/EXPR.1994.1054.

Minot, S., Melo, M. B., Li, F., Lu, D., Niedelman, W., Levine, S. S. and Saeij, J. P. J. (2012). Admixture and recombination among *Toxoplasma gondii* lineages explain global genome diversity. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13458–63. doi: 10.1073/pnas.1117047109.

Montoya, J. and Liesenfeld, O. (2004). Toxoplasmosis. *The Lancet* **363**, 1965–1976. doi: 10.1016/S0140-6736(04)16412-X.

Montoya, A., Miró, G., Mateo, M., Ramírez, C. and Fuentes, I. (2008). Molecular Characterization of *Toxoplasma gondii* Isolates from Cats in Spain. *Journal of Parasitology* **94**, 1044–1046. doi: 10.1645/GE-1403.1.

Mushi, E. Z., Binta, M. G., Chabo, R. G., Ndebele, R. and Panzirah, R. (2000). Parasites of domestic pigeons (*Columba livia* domestica) in Sebele, Gaborone, Botswana. *Journal of the South African Veterinary Association* **71**, 249–50.

Nowakowska, D., Colón, I., Remington, J. S., Grigg, M., Golab, E., Wilczynski, J. and Sibley, L. D. (2006). Genotyping of *Toxoplasma gondii* by multiplex PCR and peptide-based serological testing of samples from infants in Poland diagnosed with congenital toxoplasmosis. *Journal of clinical microbiology* **44**, 1382–9. doi: 10.1128/JCM.44.4.1382-1389.2006.

Oz, H. S. (2014). Maternal and congenital toxoplasmosis, currently available and novel therapies in horizon. *Frontiers in Microbiology* **5**, 1–6. doi: 10.3389/fmicb.2014.00385.

Pajuelo, M. J., Eguiluz, M., Roncal, E., Quiñones-García, S., Clipman, S. J., Calcina, J., Gavidia, C. M., Sheen, P., Garcia, H. H., Gilman, R. H., Gonzalez, A. E., Zimic, M. and Cysticercosis Working Group in Peru, for the C. W. G. in (2017). Genetic variability of *Taenia solium* cysticerci recovered from experimentally infected pigs and from naturally infected pigs using microsatellite markers. *PLoS neglected tropical diseases* **11**, e0006087. doi: 10.1371/journal.pntd.0006087.

Pedersen, M. G., Stevens, H., Pedersen, C. B., Nørgaard-Pedersen, B. and Mortensen, P. B. (2011). Toxoplasma infection and later development of schizophrenia in mothers. *American Journal of Psychiatry* **168**, 814–821. doi: 10.1176/appi.ajp.2011.10091351.

Pena, H. F. J., Gennari, S. M., Dubey, J. P. and Su, C. (2008). Population structure and mouse-virulence of *Toxoplasma gondii* in Brazil. *International Journal for Parasitology* **38**, 561–569. doi: 10.1016/j.ijpara.2007.09.004.

Pendergraph, G. E. (1972). A serological study of toxoplasmosis in wild pigeons and their possible role in its dissemination. *Journal of the Elisha Mitchell Scientific Society* **88**, 208–209.

Pinto, M., Borges, V., Antelo, M., Pinheiro, M., Nunes, A., Azevedo, J., Borrego, M. J., Mendonça, J., Carpinteiro, D., Vieira, L. and Gomes, J. P. (2016). Genome-scale analysis of the non-cultivable *Treponema pallidum* reveals extensive within-patient genetic variation. *Nature microbiology* **2**, 16190. doi: 10.1038/nmicrobiol.2016.190.

Pomares, C., Ajzenberg, D., Bornard, L., Bernardin, G., Hasseine, L., Dardé, M.-L. and Marty, P. (2011). Toxoplasmosis and horse meat, France. *Emerging Infectious Diseases* **17**, 1327. doi: 10.3201/EID1707.101642.

Prestrud, K. W., Dubey, J. P., Åsbakk, K., Fuglei, E. and Su, C. (2008). First isolate of *Toxoplasma gondii* from arctic fox (*Vulpes lagopus*) from Svalbard. *Veterinary Parasitology* **151**, 110–114. doi: 10.1016/J.VETPAR.2007.11.011.

Qian, W., Wang, H., Su, C., Shan, D., Cui, X., Yang, N., Lv, C. and Liu, Q. (2012). Isolation and characterization of

Toxoplasma gondii strains from stray cats revealed a single genotype in Beijing, China. *Veterinary Parasitology* **187**, 408–413. doi: 10.1016/J.VETPAR.2012.01.026.

Rajendran, C., Su, C. and Dubey, J. P. (2012). Molecular genotyping of *Toxoplasma gondii* from Central and South America revealed high diversity within and between populations. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases* **12**, 359–68. doi: 10.1016/j.meegid.2011.12.010.

Richomme, C., Aubert, D., Gilot-Fromont, E., Ajzenberg, D., Mercier, A., Ducrot, C., Ferté, H., Delorme, D. and Villena, I. (2009). Genetic characterization of *Toxoplasma gondii* from wild boar (*Sus scrofa*) in France. *Veterinary Parasitology* **164**, 296–300. doi: 10.1016/J.VETPAR.2009.06.014.

Rico-Torres, C. P., Vargas-Villavicencio, J. A. and Correa, D. (2016). Is *Toxoplasma gondii* type related to clinical outcome in human congenital infection? Systematic and critical review. *European Journal of Clinical Microbiology and Infectious Diseases* **35**, 1079–1088. doi: 10.1007/s10096-016-2656-2.

Robertson, L. J., Clark, C. G., Debenham, J. J., Dubey, J. P., Kváč, M., Li, J., Ponce-Gordo, F., Ryan, U., Schares, G., Su, C. and Tsaousis, A. D. (2019). Are molecular tools clarifying or confusing our understanding of the public health threat from zoonotic enteric protozoa in wildlife? *International Journal for Parasitology: Parasites and Wildlife*. doi: 10.1016/J.IJPPAW.2019.01.010.

Rodrigues, F., Moreira, F., Coutinho, T., Dubey, J. P., Cardoso, L. and Lopes, A. P. (2019). Antibodies to *Toxoplasma gondii* in slaughtered free-range and broiler chickens. *Veterinary Parasitology* **271**, 51–53. doi: 10.1016/J.VETPAR.2019.06.007.

Roos, D. S. (1999). The apicoplast as a potential therapeutic target in Toxoplasma and other apicomplexan parasites: some additional thoughts. *Parasitology Today* **15**, 41. doi: 10.1016/S0169-4758(98)01367-2.

Sabaj, V., Galindo, M., Silva, D., Sandoval, L. and Rodríguez, J. C. (2010). Analysis of *Toxoplasma gondii* surface antigen 2 gene (SAG2). Relevance of genotype I in clinical toxoplasmosis. *Molecular Biology Reports* **37**, 2927–2933. doi: 10.1007/s11033-009-9854-2.

Sabin, A. B. (1938). Isolation of a filtrable, transmissible agent with "Neurolity" properties from Toxoplasma-infected tissues. *Science (New York, N.Y.)* **88**, 189–91. doi: 10.1126/science.88.2278.189.

Sabin, A. B. (1941). Toxoplasmic encephalitis in children. *JAMA: The Journal of the American Medical Association* **116**, 801. doi: 10.1001/jama.1941.02820090001001.

Sabin, A. B. and Olitsky, P. K. (1937). Toxoplasma and obligate intracellular parasitism. *Science (New York, N.Y.)* **85**, 336–8. doi: 10.1126/science.85.2205.336.

Saeij, J. P. J., Boyle, J. P. and Boothroyd, J. C. (2005). Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host. *Trends in Parasitology* **21**, 476–481. doi: 10.1016/j.pt.2005.08.001.

Salant, H., Landau, D. Y. and Baneth, G. (2009). A cross-sectional survey of *Toxoplasma gondii* antibodies in Israeli pigeons. *Veterinary Parasitology* **165**, 145–149. doi: 10.1016/j.vetpar.2009.06.031.

Santoro, A., Veronesi, F., Milardi, G. L., Ranucci, D., Branciari, R., Diaferia, M. and Gabrielli, S. (2017). Sequence variation in the B1 gene among *Toxoplasma gondii* isolates from swine and cats in Italy. *Research in Veterinary Science* **115**, 353–355. doi: 10.1016/j.rvsc.2017.06.025.

Schlüter, D., Däubener, W., Schares, G., Groß, U., Pleyer, U. and Lüder, C. (2014). Animals are key to human toxoplasmosis. *International Journal of Medical Microbiology* **304**, 917–929. doi: 10.1016/j.ijmm.2014.09.002.

Science Polymorphic Secreted Kinases are key virulence factors in Toxoplasmosis.

Sepúlveda-Arias, J. C., Gómez-Marin, J. E., Bobić, B., Naranjo-Galvis, C. A. and Djurković-Djaković, O. (2014). Toxoplasmosis as a travel risk. *Travel Medicine and Infectious Disease* **12**, 592–601. doi: 10.1016/j.tmaid.2014.05.007.

Sevivas, T. (2011). Prevalência de Anticorpos anti-*Toxoplasma gondii* em grávidas da região de Lisboa e Vale do Tejo e estudo dos factores de risco.

Sibley, L. D. and Ajioka, J. W. (2008). Population Structure of Toxoplasma gondii: Clonal Expansion Driven by Infrequent Recombination and Selective Sweeps. *Annu. Rev. Microbiol* **62**, 329–51. doi: 10.1146/annurev.micro.62.081307.162925.

Sibley, L. D. and Boothroyd, J. C. (1992a). Construction of a molecular karyotype for *Toxoplasma gondii*. *Molecular*

and Biochemical Parasitology **51**, 291–300. doi: 10.1016/0166-6851(92)90079-Y.

Sibley, L. D. and Boothroyd, J. C. (1992b). Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. *Nature* **359**, 82–85. doi: 10.1038/359082a0.

Sibley, L. D., LeBlanc, A. J., Pfefferkorn, E. R. and Boothroyd, J. C. (1992). Generation of a restriction fragment length polymorphism linkage map for Toxoplasma gondii. *Genetics* **132**, 1003–15.

Sibley, L. D., Mordue, D. and Howe, D. K. (1999). Experimental approaches to understanding virulence in toxoplasmosis. *Immunobiology* **201**, 210–224. doi: 10.1016/S0171-2985(99)80061-8.

Sibley, L. D., Mordue, D. G., Su, C., Robben, P. M. and Howe, D. K. (2002). Genetic approaches to studying virulence and pathogenesis in *Toxoplasma gondii. Philosophical Transactions of the Royal Society B: Biological Sciences* **357**, 81–88. doi: 10.1098/rstb.2001.1017.

Sibley, L. D., Khan, A., Ajioka, J. W. and Rosenthal, B. M. (2009). Genetic diversity of *Toxoplasma gondii* in animals and humans. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **364**, 2749–61. doi: 10.1098/rstb.2009.0087.

Silveira, C., Muccioli, C., Hollland, G. N., Jones, J. L., Yu, F., de Paulo, A. and Belfort Jr, R. (2015). Ocular involvement following an epidemic of *Toxoplasma gondii* infection in Santa Isabel do Ivaí, Brazil. *American Journal of Ophthalmology* **159**, 1013-1021.e3. doi: 10.1016/J.AJO.2015.02.017.

Soares, R. M., Silveira, L. H., da Silva, A. V., Ragozo, A., Galli, S., Lopes, E. G., Gennari, S. M. and de Jesus Pena, H. F. (2011). Genotyping of *Toxoplasma gondii* isolates from free range chickens in the Pantanal area of Brazil. *Veterinary Parasitology* **178**, 29–34. doi: 10.1016/j.vetpar.2010.12.037.

Sobanski, V., Ajzenberg, D., Delhaes, L., Bautin, N. and Just, N. (2013). Severe toxoplasmosis in immunocompetent hosts: be aware of atypical strains. *American journal of respiratory and critical care medicine* **187**, 1143–5. doi: 10.1164/rccm.201209-1635LE.

Sousa, S. de, Ajzenberg, D., Canada, N., Freire, L., Costa, J. M. C. da, Dardé, M. L., Thulliez, P. and Dubey, J. P. (2006). Biologic and molecular characterization of *Toxoplasma gondii* isolates from pigs from Portugal. *Veterinary Parasitology* **135**, 133–136. doi: 10.1016/J.VETPAR.2005.08.012.

Sousa, S., Ajzenberg, D., Vilanova, M., Costa, J. and Dardé, 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. *Clinical and Vaccine Immunology* **15**, 1380–1386. doi: 10.1128/CVI.00186-08.

Sousa, S., Ajzenberg, D., Marle, M., Aubert, D., Villena, I., Da Costa, J. C. and Dardé, M. L. (2009a). Selection of polymorphic peptides from GRA6 and GRA7 sequences of Toxoplasma gondii strains to be used in serotyping. *Clinical and Vaccine Immunology* **16**, 1158–1169. doi: 10.1128/CVI.00092-09.

Sousa, S., Thompson, G., Silva, E., Freire, L., Lopes, D., Correia da Costa, J. M., Castro, A., Carvalheira, J. and Canada, N. (2009b). Determination of the More Adequate Modified Agglutination Test Cut-off for Serodiagnosis of *Toxoplasma gondii* Infection in Sheep. *Zoonoses and Public Health* **56**, 252–256. doi: 10.1111/j.1863-2378.2008.01187.x.

Sousa, S., Canada, N., Correia da Costa, J. M. and Dardé, M.-L. (2010). Serotyping of naturally *Toxoplasma gondii* infected meat-producing animals. *Veterinary Parasitology* **169**, 24–28. doi: 10.1016/J.VETPAR.2009.12.025.

Su, C. (2003). Recent expansion of Toxoplasma through enhanced oral transmission. *Science* **299**, 414–416. doi: 10.1126/science.1078035.

Su, C., Zhang, X. and Dubey, J. P. (2006). Genotyping of *Toxoplasma gondii* by multilocus PCR-RFLP markers: A high resolution and simple method for identification of parasites. *International Journal for Parasitology* **36**, 841–848. doi: 10.1016/J.IJPARA.2006.03.003.

Su, C., Shwab, E. K., Zhou, P., Zhu, X. Q. and Dubey, J. P. (2010). Moving towards an integrated approach to molecular detection and identification of *Toxoplasma gondii. Parasitology* **137**, 1–11. doi: 10.1017/S0031182009991065.

Su, C., Dubey, J. P., Ajzenberg, D., Khan, A., Ajioka, J. W., Rosenthal, B. M., Majumdar, D., Darde, M.-L., Zhu, X.-Q., Sibley, L. D. and Zhou, P. (2012). Globally diverse *Toxoplasma gondii* isolates comprise six major clades originating from a small number of distinct ancestral lineages. *Proceedings of the National Academy of Sciences* **109**, 5844–5849. doi: 10.1073/pnas.1203190109.

Sun, H., Wang, Y., Zhang, Y., Ge, W., Zhang, F., He, B., Li, Z., Fan, Q., Wang, W., Tu, C., Li, J. and Liu, Q.

(2013). Prevalence and genetic characterization of *Toxoplasma gondii* in bats in Myanmar. *Applied and environmental microbiology* **79**, 3526–8. doi: 10.1128/AEM.00410-13.

Sundar, N., Cole, R. A., Thomas, N. J., Majumdar, D., Dubey, J. P. and Su, C. (2008). Genetic diversity among sea otter isolates of *Toxoplasma gondii*. *Veterinary Parasitology* **151**, 125–132. doi: 10.1016/J.VETPAR.2007.11.012.

Suzuki, Y., Wang, X., Jortner, B. S., Payne, L., Ni, Y., Michie, S. A., Xu, B., Kudo, T. and Perkins, S. (2010). Removal of *Toxoplasma gondii* cysts from the brain by perforin-mediated activity of CD8+ T cells. *The American journal of pathology* **176**, 1607–13. doi: 10.2353/ajpath.2010.090825.

Talundzic, E., Ndiaye, Y. D., Deme, A. B., Olsen, C., Patel, D. S., Biliya, S., Daniels, R., Vannberg, F. O., Volkman, S. K., Udhayakumar, V. and Ndiaye, D. (2017). Molecular epidemiology of Plasmodium falciparum kelch13 mutations in senegal determined by using targeted amplicon deep sequencing. *Antimicrobial agents and chemotherapy* **61**, e02116-16. doi: 10.1128/AAC.02116-16.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599. doi: 10.1093/molbev/msm092.

Tedla, Y., Shibre, T., Ali, O., Tadele, G., Woldeamanuel, Y., Asrat, D., Aseffa, A., Mihret, W., Abebe, M., Alem, A., Medhin, G. and Habte, A. (2011). Serum antibodies to *Toxoplasma gondii* and Herpesvidae family viruses in individuals with schizophrenia and bipolar disorder: a case-control study. *Ethiopian medical journal* **49**, 211–20.

Tenter, A. M. and Johnson, A. M. (1997). Phylogeny of the tissue cyst-forming coccidia. *Advances in parasitology* **39**, 69–139.

Tenter, A. M., Heckeroth, A. R. and Weiss, L. M. (2000a). *Toxoplasma gondii*: from animals to humans. *International Journal for Parasitology* **30**, 1217–1258. doi: 10.1016/S0020-7519(00)00124-7.

Tenter, A. M., Heckeroth, A. R. and Weiss, L. M. (2000b). *Toxoplasma gondii*: from animals to humans. *International Journal for Parasitology* **30**, 1217–1258. doi: 10.1016/S0020-7519(00)00124-7.

Teutsch, S. M., Juranek, D. D., Sulzer, A., Dubey, J. P. and Sikes, R. K. (1979). Epidemic Toxoplasmosis Associated with Infected Cats. *New England Journal of Medicine* **300**, 695–699. doi: 10.1056/NEJM197903293001302.

Tian, Y.-M., Huang, S.-Y., Miao, Q., Jiang, H.-H., Yang, J.-F., Su, C., Zhu, X.-Q. and Zou, F.-C. (2014). Genetic characterization of *Toxoplasma gondii* from cats in Yunnan Province, Southwestern China. *Parasites & vectors* **7**, 178. doi: 10.1186/1756-3305-7-178.

Torrey, E. F. and Yolken, R. H. (2007). Schizophrenia and toxoplasmosis. *Schizophrenia bulletin* **33**, 727–8. doi: 10.1093/schbul/sbm026.

Torrey, E. F., Bartko, J. J. and Yolken, R. H. (2012). Toxoplasma gondii and other risk factors for schizophrenia: An update. *Schizophrenia Bulletin* **38**, 642–647. doi: 10.1093/schbul/sbs043.

Tsai, Y.-J., Chung, W.-C., Lei, H. H. and Wu, Y. (2006). Prevalence of Antibodies to *Toxoplasma gondii* in Pigeons (*Columba livia*) in Taiwan. *Journal of Parasitology* **92**, 871–871. doi: 10.1645/GE-716R2.1.

Turčeková, L., Hurníková, Z., Spišák, F., Miterpáková, M. and Chovancová, B. (2014). Toxoplasma gondii in protected wildlife in the Tatra National Park (TANAP), Slovakia. *Annals of Agricultural and Environmental Medicine* **21**, 235–238. doi: 10.5604/1232-1966.1108582.

van der Woude, M. W. and Baumler, A. J. (2004). Phase and antigenic variation in bacteria. *Clinical Microbiology Reviews* **17**, 581–611. doi: 10.1128/CMR.17.3.581-611.2004.

Vaudaux, J. D., Muccioli, C., James, E. R., Silveira, C., Magargal, S. L., Jung, C., Dubey, J. P., Jones, J. L., Doymaz, M. Z., Bruckner, D. A., Belfort, R., Holland, G. N., Grigg, M. E. and Grigg, M. E. (2010). Identification of an atypical strain of *Toxoplasma gondii* as the cause of a waterborne outbreak of toxoplasmosis in Santa Isabel do Ivai, Brazil. *The Journal of infectious diseases* **202**, 1226–33. doi: 10.1086/656397.

VERGARA, A., MARANGI, M., CARADONNA, T., PENNISI, L., PALUDI, D., PAPINI, R., IANIERI, A., GIANGASPERO, A. and NORMANNO, G. (2018). *Toxoplasma gondii* Lineages Circulating in Slaughtered Industrial Pigs and Potential Risk for Consumers. *Journal of Food Protection* **81**, 1373–1378. doi: 10.4315/0362-028X.JFP-17-496.

Verin, R., Mugnaini, L., Nardoni, S., Papini, R. A., Ariti, G., Poli, A. and Mancianti, F. (2013). Serologic,

molecular, and pathologic survey of *Toxoplasma gondii* infection in free-ranging red foxes (*VULPES VULPES*) in central Italy. *Journal of Wildlife Diseases* **49**, 545–551. doi: 10.7589/2011-07-204.

Verma, S. K., Ajzenberg, D., Rivera-Sanchez, A., Su, C. and Dubey, J. P. (2015). Genetic characterization of Toxoplasma gondii isolates from Portugal, Austria and Israel reveals higher genetic variability within the type II lineage. *Parasitology* **142**, 948–957. doi: 10.1017/S0031182015000050.

Vijaykumar, B. R., Kant, R. S., Rajendran, C., Lekshmi, S. U., Keerthana, S., Mahadevan, A., Shankar, S. K. and Jayshree, R. S. (2018). Restriction fragment length polymorphism-based genotyping of *Toxoplasma gondii* from autopsy-proven cases of acquired immunodeficiency syndrome-associated cerebral toxoplasmosis. *Annals of Indian Academy of Neurology* **21**, 250–255. doi: 10.4103/aian.AIAN_358_17.

Vilares, A., Gargaté, M. J., Ferreira, I., Martins, S., Júlio, C., Waap, H., Ângelo, H. and Gomes, J. P. (2014). Isolation and molecular characterization of *Toxoplasma gondii* isolated from pigeons and stray cats in Lisbon, Portugal. *Veterinary Parasitology* **205**, 506–511. doi: 10.1016/j.vetpar.2014.08.006.

Vilares, A., Gargaté, M. J., Ferreira, I., Martins, S. and Gomes, J. P. (2017). Molecular and virulence characterization of *Toxoplasma gondii* isolated from humans in Portugal. *Parasitology Research.* doi: 10.1007/s00436-017-5374-5.

Waap, H., Vilares, A., Rebelo, E., Gomes, S. and Ângelo, H. (2008). Epidemiological and genetic characterization of *Toxoplasma gondii* in urban pigeons from the area of Lisbon (Portugal). *Veterinary Parasitology* **157**,. doi: 10.1016/j.vetpar.2008.07.017.

Waap, H., Cardoso, R., Leitão, A., Nunes, T., Vilares, A., Gargaté, M. J., Meireles, J., Cortes, H. and Ângelo, H. (2012). *In vitro* isolation and seroprevalence of *Toxoplasma gondii* in stray cats and pigeons in Lisbon, Portugal. *Veterinary Parasitology* **187**, 542–547. doi: 10.1016/j.vetpar.2012.01.022.

Wang, L., He, L., Meng, D., Chen, Z., Wen, H., Fang, G., Luo, Q., Huang, K. and Shen, J. (2015). Seroprevalence and genetic characterization of *Toxoplasma gondii* in cancer patients in Anhui Province, Eastern China. *Parasites & Vectors* **8**, 162. doi: 10.1186/S13071-015-0778-5.

Webster, J. P. (2001). Rats, cats, people and parasites: the impact of latent toxoplasmosis on behaviour. *Microbes and Infection* **3**, 1037–1045. doi: 10.1016/S1286-4579(01)01459-9.

Webster, J. P., Kaushik, M., Bristow, G. C. and McConkey, G. a (2013). Toxoplasma gondii infection, from predation to schizophrenia: can animal behaviour help us understand human behaviour? *The Journal of experimental biology* **216**, 99–112. doi: 10.1242/jeb.074716.

Weiss, L. and Dubey, J. (2009). Toxoplasmosis: A history of clinical observations. *International Journal for Parasitology* **39**, 895–901. doi: 10.1016/j.ijpara.2009.02.004.Toxoplasmosis.

Weiss, L. M. and Kim, K. (2000). The development and biology of bradyzoites of *Toxoplasma gondii*. *Frontiers in bioscience : a journal and virtual library* **5**, D391-405.

Weiss, L. M. and Kim, K. (2007). Toxoplasma gondii*: the model apicomplexan : perspectives and methods*. Elsevier/Academic Press.

Wendte, J. M., Gibson, A. K. and Grigg, M. E. (2011). Population genetics of Toxoplasma gondii: New perspectives from parasite genotypes in wildlife. *Veterinary Parasitology* **182**, 96–111. doi: 10.1016/j.vetpar.2011.07.018.

Wolf, A., Cowen, D. and Paige, B. H. (1940). Toxoplasmic Encephalomyelitis: IV. Experimental transmission of the infection to animals from a human infant. *The Journal of experimental medicine* **71**, 187–214.

Xiao, J., Buka, S. L., Cannon, T. D., Suzuki, Y., Viscidi, R. P., Torrey, E. F. and Yolken, R. H. (2009). Serological pattern consistent with infection with type I *Toxoplasma gondii* in mothers and risk of psychosis among adult offspring. *Microbes and Infection* **11**, 1011–1018. doi: 10.1016/J.MICINF.2009.07.007.

Yai, L. E. O., Ragozo, A. M. A., Aguiar, D. M., Damaceno, J. T., Oliveira, L. N., Dubey, J. P. and Gennari, S. M. (2008). Isolation of *Toxoplasma gondii* from Capybaras (*Hydrochaeris hydrochaeris*) from São Paulo State, Brazil. *Journal of Parasitology* **94**, 1060–1063. doi: 10.1645/GE-1548.1.

Yan, C., Yue, C. L., Qiu, S. B., Li, H. L., Zhang, H., Song, H. Q., Huang, S. Y., Zou, F. C., Liao, M. and Zhu, X. Q. (2011). Seroprevalence of *Toxoplasma gondii* infection in domestic pigeons (*Columba livia*) in Guangdong Province of southern China. *Veterinary Parasitology* **177**, 371–373. doi: 10.1016/j.vetpar.2010.12.004.

Yilmaz, S. M. and Hopkins, S. H. (1972). Effects of different conditions on duration of infectivity of *Toxoplasma gondii* oocysts. *The Journal of parasitology* **58**, 938–9.

Zhou, P., Chen, Z., Li, H.-L., Zheng, H., He, S., Lin, R.-Q. and Zhu, X.-Q. (2011). *Toxoplasma gondii* infection in humans in China. *Parasites & Vectors* **4**, 165. doi: 10.1186/1756-3305-4-165.

ANNEX

Supplementary Table S1. *T. gondii* isolates from the collection of the URSZ-INSA used in the study.

Id sample	Host	Human product	Geographic origin	Mouse virulence	Confirmed CT	Clinical manifestations	Year
RH_Lab	Human	T	---	V	P	Encephalitis	1939
Str_1	Human	AF	PT- Lisbon	V	NI	NI	1994
Str_2[a]	Human	UCB	PT- Lisbon	NV	NI	NI	1994
Str_3	Human	UCB	PT- Setubal	V	NI	NI	1999
Str_4[a]	Human	AF	PT- Lisbon	NV	NI	NI	1999
Str_5[a]	Human	AF	PT- Coimbra	NV	NI	NI	1999
Str_6	Human	PL	PT- Lisbon	V	NI	NI	2000
Str_7[a]	Human	PL	PT- Leiria	NV	NI	NI	2000
Str_8	Human	PL	PT- Leiria	NV	NI	NI	2000
Str_9	Human	PL	PT- Porto	NV	NV	NI	2000
Str_10[a]	Human	PL	PT- Santarem	NV	P	NI	2000
Str_11	Human	PL	PT- Coimbra	V	P	NI	2000
Str_12	Human	PL	PT- Lisbon	NV	NI	NI	2000
Str_13	Human	PL	PT- Porto	V	NI	NI	2000
Str_14[a]	Human	PL	PT- Lisbon	NV	NI	NI	2000
Str_15[a]	Human	UCB	PT- Coimbra	NV	NI	NI	2001
Str_16	Human	PL	PT- Guarda	NV	NI	NI	2001
Str_17	Human	PL	PT- Setubal	NV	NI	NI	2001
Str_18[a]	Human	PL	PT- Porto	V	P	NI	2002
Str_19[a]	Human	UCB	PT- Lisbon	NV	NI	NI	2003
Str_20	Human	PL	PT- Coimbra	NV	P	NI	2003
Str_21	Human	UCB	PT- Lisbon	NV	N	NA	2003
Str_22[a]	Human	PL	PT- Porto	NV	NI	NI	2003
Str_23[a]	Human	PL	PT- Santarem	NV	NI	NI	2003
Str_24	Human	UCB	PT- Porto	NV	P	NI	2005
Str_25	Human	PL	PT- Porto	NV	P	Intracranial calcifications	2005
Str_26[a]	Human	PL	PT- Coimbra	NV	P	NV	2005
Str_27[a]	Human	AF	PT- Lisbon	NV	NI	NI	2005
Str_28	Human	PL	PT- Coimbra	V	P	NI	2005
Str_29[a]	Human	PL	PT- Porto	NV	NI	NI	2006
Str_30[a]	Human	AF	PT- Coimbra	NV	NI	NI	2006
Str_31	Human	T	PT- Lisbon	V	NA	NA	2000
Str_32	Human	PL	PT- Bragança	NV	P	NI	2006
Str_33[a]	Human	UCB	PT- Porto	NV	NI	NV	2007
Str_34	Human	UCB	PT- Coimbra	NV	P	NI	2007

Str_35	Human	PL	PT- Lisbon	NV	P	Intracranial calcifications; Chorioretinitis	2007
Str_36[a]	Human	UCB	PT- Faro	NV	P	Splenomegaly; IGR	2007
Str_37[a]	Human	PL	PT- Porto	NV	N	NA	2007
Str_38	Human	PL	PT- Santarem	NV	P	NI	2007
Str_39[a]	Human	PL	PT- Lisbon	NV	NI	NI	2008
Str_40[a]	Human	PL	PT- Azores	NV	NI	NI	2008
Str_41[a]	Human	PL	PT- Evora	NV	NI	NI	2010
Str_42[a]	Human	BB	PT- Santarem	NV	P	NV	2010
Str_43	Human	PL	PT- Leiria	NV	NI	NI	2011
Str_44[a]	Human	PL	PT- Aveiro	NV	N	NA	2011
Str_45[a]	Human	PL	PT- Lisbon	NV	NI	NI	2010
Str_46[a]	Human	PL	PT- Porto	NV	N	NA	2012
Str_47[a]	Human	PL	PT- Santarem	NV	P	Hidrocephalitis	2012
Str_48[a]	Human	BB	PT- Lisbon	NV	P	NV	2012
Str_49	Human	PL	PT- Porto	NV	P	NI	2013
Str_50	Human	PL	PT- Lisbon	NV	NI	NI	2008
Str_51	Human	PL	PT- Lisbon	V	NI	NI	2007

[a] Strains with data for all loci studied (with exception of Sag1 and L363)

T: tissues; AF: amniotic fluid; UCB: umbilical cordon blood; PL: placenta; BB: baby blood; PT - Portugal;

V: virulent (when clinical symptons were reported); NV: not virulent; CT: congenital toxoplasmosis;

P: positive; NI: no information available; N: negative; NA: not applicable (the parasite did not cross the placentar barrier);

IGR: intrauterine growth retardation.

Supplementary Table S2. Primers used for PCR, Sanger and/or NGS.

Locus	Chr	Primer Foward	Primer reverse	Amplicon length (bp)
Pk 1[a]	VI	FE 5'TTC CCC GCA GTA ATT GGA GGA TAC 3' FI 5'TCA TCG CTG AAT CTC ATT GC 3'	RE 5'AGG CGA TTA AAG ACA TGG CAG AAA 3' RI 5'CGC AAA GGG AGA CAA TCA GT 3'	943-980
GRA 6[b]	X	FE 5'ATT TGT GTT TCC GAG CAG GT 3' FI 5'TTT CCG AGC AGG TGA CCT 3'	RE 5'TCG CCG AAG AGT TGA CAT AG 3'	343
M 102[a]	VIIa	F 5'GAG CGA CGC CCG TAT GAT AAG G 3'	R 5'CGC GCT GAG AAG CTG ACA TAC AG 3'	497-529
L363[a]	VIIb	FE 5'CAG CGG CAG CAA GTG AAT AAA CAC3' FI 5'CAG GCG CCA CTG TGA TGC 3'	RE 5'GTC ATC GCC GCC AGG GAA GC 3' RI 5'AGA AGG CTC GCC CGT GTA GAG AA 3'	901-5
SAG 1[a]	VIII	FE 5'TGC ACC TGT AGG AAG CTG TAG TCA 3' FI 5'CAC ACG GTT GTA TGT CGG TTT C 3'	RE 5'CAA GCT GCG ATA GAG CCA ATA AGT 3' RI 5'TCC CCC GTC CAC CAG CTA TCT TCT 3'	404-6
SAG 3[b]	XII	FE 5'CAA CTC TCA CCA TTC CAC CC 3' FI 5'TCT TGT CGG GTG TTC ACT CA 3'	RE 5'GCG CGT TGT TAG ACA AGA CA 3' RI 5'CAC AAG GAG ACC GAG AAG GA 3'	225
CB21-4[b]	III	F 5'CCA GGT GTT TCG ATA TTG AT 3'	R 5'GCC TGT GTG GTG TTC GAA TC 3'	485-509

[a] Genotyping by NGS; [b] Genotyping by Sanger and NGS

Supplementary Table 3. Mean depth of coverage per locus

Sample	CB21	PK1	M102	L363	Sag1	GRA6	SAG3	Mean coverage per strain	ENA BioProject	ENA SAMPLE ID	ENA SECONDARY SAMPLE ID
RH_Lab	*	5979,15	5245,83		1419,47	6328,71	*	4743	PRJEB32939	ERS3503939	SAMEA5700278
Str_1	6489,87		5716,06	925,98		7429,27	3477,24	4808	PRJEB32939	ERS3503940	SAMEA5700279
Str_2	6935,39	7439,21	5048,51			2772,35	734,78	4586	PRJEB32939	ERS3503941	SAMEA5700280
Str_3	*	7580,49	4516,81		1595,66	*	1039,54	3683	PRJEB32939	ERS3503942	SAMEA5700281
Str_4	2506,62	6335,51	4618,32	7512,92		*	656,88	4326	PRJEB32939	ERS3503943	SAMEA5700282
Str_5	936,67	445,65	5084,94	6365,45		4654,18	769,37	3043	PRJEB32939	ERS3503944	SAMEA5700283
Str_6	6756,37		6132,99	2193,52		3107,85	5618,74	4762	PRJEB32939	ERS3503945	SAMEA5700284
Str_7	7242,95	6888,63	357,14		758,41	6826,65	*	4415	PRJEB32939	ERS3503946	SAMEA5700285
Str_8		7582,88	5509,14			*	*	6546	PRJEB32939	ERS3503947	SAMEA5700286
Str_9	*		4469,94	6933,44		*	1399,41	4268	PRJEB32939	ERS3503948	SAMEA5700287
Str_10	7078,15	7220,08	6460,91			5426,81	3459,52	5929	PRJEB32939	ERS3503949	SAMEA5700288
Str_11		7352,65	5743,53			1935,76		5011	PRJEB32939	ERS3503950	SAMEA5700289
Str_12	5143,26		3655,95			*	*	4400	PRJEB32939	ERS3503951	SAMEA5700290
Str_13			5238,99			*	7047,65	6143	PRJEB32939	ERS3503952	SAMEA5700291
Str_14	*	7403,00	4625,90	6041,34		*	*	6023	PRJEB32939	ERS3503953	SAMEA5700292
Str_15	3910,35	7119,66	5207,27	5190,73		5348,21	*	5355	PRJEB32939	ERS3503954	SAMEA5700293
Str_16	6690,55		2371,06			*	*	4531	PRJEB32939	ERS3503955	SAMEA5700294
Str_17	*	7547,99	4894,44	3670,12		*	2153,06	4566	PRJEB32939	ERS3503956	SAMEA5700295
Str_18	*	7409,81	5026,05	2246,81		*	*	4894	PRJEB32939	ERS3503957	SAMEA5700296
Str_19	4588,80	7356,79				7482,86	*	6476	PRJEB32939	ERS3503958	SAMEA5700297
Str_20		7258,23	4447,04	7000,61	990,52	*	*	4924	PRJEB32939	ERS3503959	SAMEA5700298
Str_21	*	7313,20		5406,78	1211,46	*	*	4644	PRJEB32939	ERS3503960	SAMEA5700299
Str_22	1425,39	7261,24	4980,53			1149,70	*	3704	PRJEB32939	ERS3503961	SAMEA5700300
Str_23	*	1922,51				*	*	1923	PRJEB32939	ERS3503962	SAMEA5700301
Str_24	4392,70		2214,00	50,72		*	*	2219	PRJEB32939	ERS3503963	SAMEA5700302
Str_25	*	612,61			1442,45	*	864,11	973	PRJEB32939	ERS3503964	SAMEA5700303
Str_26	*	7158,68	5374,18	2843,16		*	*	5125	PRJEB32939	ERS3503965	SAMEA5700304
Str_27	*	7258,17	5725,56		243,30	*	1781,66	3752	PRJEB32939	ERS3503966	SAMEA5700305
Str_28								---	PRJEB32939	ERS3503967	SAMEA5700306
Str_29	2737,28	223,87	6732,51			*	2392,45	3022	PRJEB32939	ERS3503968	SAMEA5700307
Str_30	*	3815,91	5713,88	4248,82	1447,42	*	*	3807	PRJEB32939	ERS3503969	SAMEA5700308
Str_31	4396,65		312,43			*	*	2355	PRJEB32939	ERS3503970	SAMEA5700309
Str_32	6105,49		5978,83	101,20	1344,96	*	*	3383	PRJEB32939	ERS3503971	SAMEA5700310
Str_33	1324,34	3265,15	4439,27	129,66	1272,40	1297,09	359,79	1727	PRJEB32939	ERS3503972	SAMEA5700311
Str_35	2814,27		4868,51	6812,19	7134,45	*	*	5407	PRJEB32939	ERS3503973	SAMEA5700312
Str_36	*	7264,58	4684,70		1453,26	*	*	4468	PRJEB32939	ERS3503974	SAMEA5700313
Str_37	7319,42	7112,60	3514,25			*	*	5982	PRJEB32939	ERS3503975	SAMEA5700314
Str_38	6607,80		1554,65			*	832,72	2998	PRJEB32939	ERS3503976	SAMEA5700315
Str_39	6827,46	96,65	4733,60	1834,67		*	*	3373	PRJEB32939	ERS3503977	SAMEA5700316
Str_40	2012,77	7507,52	5102,62	1058,50		*	*	3920	PRJEB32939	ERS3503978	SAMEA5700317
Str_41	7188,27	6900,42	6532,95			*	796,87	5355	PRJEB32939	ERS3503979	SAMEA5700318
Str_42	6617,10	7496,67	4852,56	2540,06		6463,41	*	5594	PRJEB32939	ERS3503980	SAMEA5700319
Str_43	7288,15		6788,81			*	3061,06	5713	PRJEB32939	ERS3503981	SAMEA5700320
Str_44	6604,15	7264,34	5961,95	3903,00		4970,82	*	5741	PRJEB32939	ERS3503982	SAMEA5700321
Str_45	6015,54	7346,75	5485,74		1594,81	*	*	5111	PRJEB32939	ERS3503983	SAMEA5700322
Str_46	2331,55	7315,18	5431,99	111,33		*	*	3798	PRJEB32939	ERS3503984	SAMEA5700323
Str_47	2394,01	4006,63	1800,22	4942,83		*	*	3286	PRJEB32939	ERS3503985	SAMEA5700324
Str_48	1763,69	1092,01			310,38	5671,22	663,97	1900	PRJEB32939	ERS3503986	SAMEA5700325
Str_49						7647,82	7035,15	7341	PRJEB32939	ERS3503987	SAMEA5700326
Str_50	5229,25		5755,18		1474,21	*	1769,14	3557	PRJEB32939	ERS3503988	SAMEA5700327
Str_51	6868,60				2215,10	7638,73	1563,79	4572	PRJEB32939	ERS3503989	SAMEA5700328
Mean coverage per locus	4891,96	5804,41	4718,83	3567,99	1619,27	5067,73	2260,80	4196			

* Sequences obtained by Sanger sequencing

No data.

Supplementary Table 4. Reference strains with genome sequences available at ToxoDB.

| Id sample | Host | Geographic origin | Type strain | Mouse Virulence | Human virulence | Isolation Year | ToxoDB loci | | | | | | | |
							Sag2	CB21-4	PK1	L363	SAG1	GRA6	SAG3	M102
GT1	Goat	USA	I	V	NA	1980	TGGT1_chrVIII	TGGT1_chrIII	TGGT1_chrVI	TGGT1_chrVIIb	TGGT1_chrVIII	TGGT1_chrX	TGGT1_chrXII	TGGT1_chrVIIa
ME49	Sheep	USA	II	NV	NA	1965	TGME49_chrVIII	TGME49_chrIII	TGME49_chrVI	TGME49_chrVIIb	TGME49_chrVIII	TGME49_chrX	TGME49_chrXII	TGME49_chrVIIa
VEG	Human (AIDS)	USA	III	NV	V	1988	TGVEG_chrVIII	TGVEG_chrIII	TGVEG_chrVI	TGVEG_chrVIIb	TGVEG_chrVIII	TGVEG_chrX	TGVEG_chrXII	TGVEG_chrVIIa
ARI	Human (Transplant)	USA	Atypical	NV	V	1992	AQGS01000519	AQGS01002246	AQGS01002525	AQGS01002623	AQGS01000636	AQGS01001935	AQGS01002996	AQGS01001447
CAST	Human (AIDS)	USA	Atypical	V	V	1988	AHIV01000154	AHIV01002291	AHIV01002139	AHIV01001382	AHIV01000059	AHIV01002848	AHIV01001358	AHIV01000841
COUG	Cougar	Canada	Atypical	V	NA	1996	AGQR01001675	AGQR01002991	AGQR01001084	AGQR01000083	AGQR01000539	AGQR01002501	AGQR01003479	AGQR01002247
CtCo5	Cat	Colombia	Atypical	V	NA	2005	AKIR01001216	AKIR01001614	NI	AKIR01001479	AKIR01003375	AKIR01003418	AKIR01002844	AKIR01000181
FOU	Human (Transplant)	France	Atypical	V	V	1992	AETH01000695	AETH01001179	AETH01001493	AETH01001295	AETH01000411	AETH01003480	AETH01003400	AETH01000949
GAB2-2007-GAL-DOM2	Chicken	Gabon	Atypical	V	NA	2007	AHZU01000615	AHZU01001406	AHZU01001657	AHZU01000002	AHZU01000083	AHZU01001344	AHZU01000189	AHZU01000904
MAS	Human (Congenital)	France	Atypical	V	NI	1991	AEXC01000621	AEXC01000198	AEXC01001478	AEXC01002488	AEXC01000407	AEXC01003166	AEXC01003331	AEXC01001844
P89	Pig	USA	Atypical	V	NA	1991	AETJ01000080	AETJ01000748	AETJ01001951	AETJ01001875	AETJ01000187	AETJ01003029	AETJ01002709	AETJ01000855
RUB	Human (Immunocompetent)	French Guiana	Atypical	V	V	1992	AFYV01001796	AFYV01002365	AFYV01002294	AFYV01000309	AFYV01000949	AFYV01003446	AFYV01003576	AFYV01001691
TgCATBr5	Cat	Brazil	Atypical	NI	NA	2006	AFPV01001459	AFPV01001874	AFPV01001691	AFPV01000769	AFPV01000478	AFPV01004140	AFPV01004084	AFPV01001242
TgCATBr9	Cat	Brazil	Atypical	V	NA	2006	AFHV01000310	AFHV01001434	AFHV01002112	AFHV01001131	AFHV01000155	AFHV01003616	AFHV01003563	AFHV01001286
TgCatPRC2	Cat	China	Atypical	NI	NA	2007	AHZP01000397	AHZP01001656	AHZP01000063	AHZP01000545	AHZP01000250	NI	AHZP01003659	AHZP01002005
VAND	Human (Immunocompetent)	French Guiana	Atypical	V	V	1997	AEYJ01001104	AEYJ01001003	AEYJ01001266	AEYJ01000459	AEYJ01000315	AEYJ01000797	AEYJ01001230	AEYJ01000942

V: virulent, NV: non virulent, NI: no information available; NA: not applicable

Khan et al 2011b

Supplementary Table S5A. MLST-like numerical allele classification (loci are grouped by genotyping approach).

Id	"RFLP" recognition sites (by Sanger sequencing)	Microsatellite size analysis					Nucleotide sequencing (Sanger/NGS)							
Chr	VIII	VIIa	II	IX	XII	X	VIII	X	VI	XII	VIIb	III	VIIa	VIII
	Sag2	B18[b]	W35[b]	Tub2[b]	B17[b]	TgM-A	Sag1[b]	Gra6	PK1	Sag3	L363	CB21-4	M102	Sag2
GT1[a]	1	31	1	1	1	1	1	1	1	1	1	1	1	1
ME49[a]	2	2	23	23	23	2	23	2	2	2	2	2	2	2
VEG[a]	3	31	23	23	23	3	23	3	3	3	3	3	3	3
ARI[a]	2	31	23	23	23	1	23	4	4	2	4	4	4	2
CAST[a]	1	31	23	1	1	3	1	1	1	1	1	5	1	1
COUG[a]	2	2	23	23	23	3	1	4	5	4	5	6	5	4
CtCo5[a]	1	31	23	1	23	3	1	5	NI	3	6	5	1	1
FOU[a]	1	31	1	1	1	1	1	2	1	3	6	1	1	1
GAB2-GAL-DOM2[a]	1	31	23	1	1	2	1	6	1	3	3	3	1	1
MAS[a]	1	4	23	1	4	3	4	7	6	3	7	7	6	1
P89[a]	3	2	23	23	23	1	1	5	7	3	8	8	3	3
RUB[a]	1	4	23	23	23	2	1	8	8	1	9	9	3	5
TgCATBr5[a]	3	31	23	23	4	3	1	9	9	3	7	7	7	3
TgCATBr9[a]	1	31	23	23	4	3	1	6	9	3	6	7	7	1
TgCatPRC2[a]	2	31	23	4	23	4	23	ND	10	5	10	4	4	2
VAND[a]	1	4	4	1	4	2	1	10	8	1	11	5	8	1
RH_Lab	1	31	1	1	1	1	1	2	1	1	ND	1	1	1
Str_1	1	31	1	1	1	1	ND	2	ND	6	2	4	2	1
Str_2	1	31	1	1	1	1	ND	2	1	6	1	1	1	1
Str_3	1	31	1	1	1	1	1	1	1	1	ND	2	1	6
Str_4	2	2	23	23	23	2	ND	2	2	7	12	4	2	7
Str_5	2	2	23	23	23	2	ND	2	2	6	2	4	2	2
Str_6	1	31	1	1	1	1	ND	1	ND	1	13	1	1	1
Str_7	2	2	23	23	23	2	23	2	2	6	ND	2	2	8
Str_8	2	2	23	23	23	2	ND	2	2	6	ND	ND	2	9
Str_9	1	2	23	23	23	2	ND	2	ND	2	2	2	1	1
Str_10	1	31	23	23	23	3	ND	3	3	3	ND	4	9	10
Str_11	1	31	1	1	1	3	ND	6	1	ND	ND	ND	1	5
Str_12	2	2	23	23	23	2	ND	2	ND	2	2	4	2	11
Str_13	1	31	1	1	1	3	ND	2	2	3	ND	ND	1	1
Str_14	2	2	23	23	23	2	ND	2	2	6	2	1	1	12
Str_15	2	2	23	23	23	2	ND	2	2	2	2	2	2	13
Str_16	2	2	23	23	23	2	ND	2	ND	6	ND	4	1	14
Str_17	1	2	23	23	23	2	ND	ND	2	6	2	1	2	10
Str_18	2	2	23	23	23	2	ND	2	2	6	2	4	2	15
Str_19	2	2	23	23	23	2	ND	2	2	6	ND	4	ND	16
Str_20	2	2	23	23	23	2	23	2	2	6	2	ND	1	11
Str_21	2	2	23	23	23	2	23	2	2	6	2	4	ND	17
Str_22	2	2	23	23	23	2	ND	2	2	6	ND	4	1	12
Str_23	2	2	23	23	23	2	ND	2	2	2	ND	4	ND	2
Str_24	2	2	23	23	23	2	ND	1	ND	2	2	4	2	18
Str_25	1	2	23	23	23	1	1	1	2	1	ND	1	ND	19
Str_26	2	2	23	23	23	2	ND	2	2	6	2	1	2	2
Str_27	2	2	1	1	23	2	23	1	2	6	ND	1	2	2
Str_28	1	2	23	23	23	2	ND	ND	ND	ND	ND	ND	ND	1
Str_29	2	2	23	23	23	2	ND	2	2	2	ND	4	2	2
Str_30	2	2	23	23	23	2	23	1	2	6	2	1	2	2
Str_31	1	31	1	1	1	1	ND	2	ND	8	ND	4	2	1
Str_32	2	2	23	23	23	2	23	2	ND	6	2	4	2	2
Str_33	2	2	23	23	23	2	23	2	2	2	2	4	2	2
Str_34	1	31	1	1	1	1	ND	2	ND	ND	ND	ND	ND	20
Str_35	2	2	23	23	23	2	23	2	ND	6	2	4	2	2
Str_36	2	2	23	23	23	2	23	2	2	6	ND	4	1	2
Str_37	2	2	23	23	23	2	ND	2	2	6	ND	4	2	2
Str_38	2	2	23	23	23	2	ND	2	ND	2	ND	4	2	2
Str_39	2	2	23	23	23	2	ND	2	2	2	2	2	1	21
Str_40	2	2	23	23	23	2	ND	2	2	9	2	4	1	2
Str_41	2	2	23	23	23	2	ND	1	2	10	ND	4	2	2
Str_42	2	31	1	23	23	2	ND	2	2	6	2	2	2	2
Str_43	2	31	23	23	23	2	ND	2	ND	6	ND	2	2	2
Str_44	2	2	23	23	23	2	ND	2	2	6	2	4	2	2
Str_45	2	2	23	23	23	2	23	2	2	6	ND	4	2	2
Str_46	2	2	23	23	23	2	ND	2	2	2	2	4	2	22
Str_47	2	2	23	23	23	2	ND	2	2	6	2	4	2	23
Str_48	2	2	23	23	23	2	23	2	2	2	ND	4	ND	2
Str_49	2	31	1	1	1	1	ND	2	ND	6	ND	ND	ND	11
Str_50	1	2	1	1	23	2	23	1	ND	11	ND	1	2	1
Str_51	1	2	1	1	1	3	1	6	ND	3	ND	1	ND	1

ND - Not detected; NI - No information available; [a]Reference sequences from TOXODB; Allele classification for loci that do not differentiate the three archetypes (labeled with b) was performed as follows: "31" (no differentiation between Type I and Type III), "23" (no differentiation between Type II and Type III)

127

Supplementary Table S5B. Nominal allele and strain classification. In order to assess the gain introduced by newly sequenced loci in the identification of interlocus recombinant strains, strains were classified based on two loci combinations: i) loci used for traditional *Sag* 2 and microsatellites genotyping (Clasification 1); ii) all loci (Classification 2).

	Nominal Allele Classification														Strain Classification					
	"RFLP" recognition sites (by Sanger sequencing)	Microsatellite size analysis					Nucleotide sequencing (Sanger/NGS)								Classification 1. Sag2 classical genotyping + Microsatellites	Classification 2. Combination of all loci (and methods)				
chromosome	VIII	VIIa	II	IX	XII	X	VIII	X	VI	XII	VIIb	III	VIIa	VIII			At least one locus with microsatellites is Atypical or Atypical-like	Mouse Virulence	Confirmed CT	Human virulence
Id	Sag2	B18	W35	Tub2	B17	TgM-A	Sag1	Gra6	PK1	Sag3	L363	CB21-4	M102	Sag2	Type	Type				
GT1[a]	I	I-III	I	I	I	I	I	I	I	I	I	I	I	I	I	I	N	V	NA	NA
ME49[a]	II	II	II-III	II-III	II-III	II	II-III	II	II	II	II	II	II	II	II	II	N	NV	NA	NA
VEG[a]	III	I-III	II-III	II-III	II-III	III	II-III	III	III	III	III	III	III	III	III	III	N	NV	NA	V
ARI[a]	II	I-III	II-III	II-III	II-III	I	II-III	II-like	II-like	II	II-like	II-like (atypical-like)	II-like (atypical-like)	II	Rec I/II	Rec I/II	Y	NV	NA	V
CAST[a]	I	I-III	II-III	I	I	III	I	I	I	I	I	NC	I-like (atypical-like)	I	Rec I/III	Rec I/III	Y	V	NA	V
COUG[a]	II	II	II-III	II-III	II-III	III	I	II-like	III-like	NC (IntraRec I/III)	II-like	II-like (atypical-like)	II-like (atypical-like)	II-like	Rec I/III	Rec II/III	Y	V	NA	NA
CtCo5[a]	I	I-III	II-III	I	II-III	III	I	II-like	I	III	NC (IntraRec I/II/III)	NC	I (atypical-like)	I	Rec I/III	Rec I/III	Y	V	NA	NA
FOU[a]	I	I-III	I	I	I	I	I	NC	I	III	NC (IntraRec I/II/III)	I	I	I	I	Rec I/III	N	V	NA	V
GAB2-GAL-DOM2[a]	I	I-III	II-III	I	I	II	I	II	I	III	III	III (atypical-like)	I	I	Rec I/III	Rec I/II/III	Y	V	NA	NA
MAS[a]	I	atypical	II-III	I	atypical	III	I-like	II-like	I-like	III	NC (IntraRec I/II)	NC (IntraRec II/III)	II-like (atypical-like)	I	Rec I/II (atypical)	Rec I/II/III	Y	V	P	NI
P89[a]	III	II	II-III	II-III	II-III	II	I	II-like	I-like	III	I-like	NC	III (atypical-like)	III	Rec II/III	Rec II/III	Y	V	NA	V
RUB[a]	I	atypical	II-III	II-III	II-III	II	I	NC	I-like	I	NC (IntraRec I/II/III)	I-like (atypical-like)	III (atypical-like)	I-like	Rec I/II (atypical)	Rec I/II/III	Y	V	NA	V
TgCATBr5[a]	II	I-III	II-III	II-III	atypical	III	I	II-like	NC	III	NC (IntraRec I/II)	II-like (atypical-like)	III-like (atypical-like)	III	III (atypical)	Rec I/II/III	Y	NI	NA	NA
TgCATBr9[a]	I	I-III	II-III	II-III	atypical	III	I	NC	NC	III	II-like	II-like (atypical-like)	III-like (atypical-like)	I	Rec I/III (atypical)	Rec I/II/III	Y	V	NA	NA
TgCatPRC2[a]	I	I-III	II-III	atypical	II-III	atypical	II-III	ND	I-like	III-like	NC (IntraRec I/II)	II-like (atypical-like)	II-like	Rec I/III (atypical)	Rec I/II/III	Y	NI	NA	NA
VAND[a]	I	atypical	atypical	I	atypical	II	I	NC	I-like	I	NC (IntraRec I/II/III)	NC	II-like (atypical-like)	I	Rec I/II (atypical)	Rec I/II	Y	V	NA	V
RH_Lab	I	I-III	I	I	I	I	I	I	I	I	ND	I	I	I	I	Rec I/II	N	V	P	V
Str_1	I	I-III	I	I	I	I	ND	II	ND	II-like	II	II-like (atypical-like)	II	I	I	Rec I/II	N	V	NI	NI
Str_2	I	I-III	I	I	I	I	ND	II	I	II-like	ND	I	I	I	I	Rec I/II	N	NV	NI	NI
Str_3	I	I-III	I	I	I	I	I	I	I	ND	II	I	I-like	I	Rec I/II	N	V	NI	NI
Str_4	II	II	II-III	II-III	II-III	II	ND	II	II	II-like	II-like	II-like (atypical-like)	II (atypical-like)	II-like	II	II	Y	NV	NI	NI
Str_5	II	II	II-III	II-III	II-III	II	ND	II	II	II-like	II-like	II-like (atypical-like)	II	II	II	II	Y	NV	NI	NI
Str_6	I	I-III	I	I	I	I	ND	I	ND	I	NC (IntraRec I/II)	I	I	I	I	I	N	V	NI	NI
Str_7	II	II-III	II-III	II-III	II	II-III	II	II	II-like	ND	II	II (atypical-like)	II-like	II	II	Y	NV	NI	NI
Str_8	II	II-III	II-III	II-III	II	ND	II	II	II-like	ND	ND	II	II-like	II	II	N	NV	NI	NI
Str_9	II	II-III	II-III	II-III	II	ND	II	ND	II	II	II	Rec I/II	Rec I/II	N	NV	NV	NI
Str_10	II	I-III	II-III	II-III	II-III	III	ND	III	III	III	ND	II-like (atypical-like)	III-like (atypical-like)	II-like	Rec I/III	Rec I/II/III	Y	NV	P	NI
Str_11	I	I-III	I	I	I	III	ND	NC	I	ND	ND	II-like (atypical-like)	I	I-like	Rec I/III	Rec I/III	N	V	P	NI
Str_12	II	II	II-III	II-III	II-III	II	ND	II	ND	I	ND	II-like (atypical-like)	II	II-like	II	II	Y	NV	NI	NI
Str_13	I	I-III	I	I	I	III	ND	II	ND	III	ND	ND	I	Rec I/III	Rec I/II/III	N	V	NI	NI
Str_14	II	II-III	II-III	II-III	II	ND	II	II	II-like	II	I	I	II-like	II	Rec I/II	N	NV	NI	NI
Str_15	II	II-III	II-III	II-III	II	ND	II	II	II	II	II (atypical-like)	II	II-like	II	N	NV	NI	NI
Str_16	II	II-III	II-III	II-III	II	ND	II	ND	II-like	ND	II-like (atypical-like)	II	II-like	II	Rec I/II	Y	NV	NI	NI
Str_17	I	II-III	II-III	II-III	II	ND	ND	II-like	II	I	I-like	I-like	Rec I/II	Rec I/II	N	NV	NI	NI
Str_18	II	II-III	II-III	II-III	II	ND	II	II	II-like	II	II-like	II (atypical-like)	II-like	II	II	N	V	P	NI
Str_19	II	II-III	II-III	II-III	II	ND	II	II	II-like	ND	II-like (atypical-like)	ND	II-like	II	II	Y	NV	NI	NI
Str_20	II	II-III	II-III	II-III	II	II-III	II	II	II-like	III	ND	II-like	II	II	Y	NV	P	NI
Str_21	II	II-III	II-III	II-III	II	II-III	II	II	II-like	II	II-like	ND	II-like	II	II	N	NV	N	NA
Str_22	II	II-III	II-III	II-III	II	ND	II	II	II-like	ND	II-like (atypical-like)	ND	II-like	II	II	N	NV	NI	NI
Str_23	II	II-III	II-III	II-III	II	ND	II	II	II	ND	II-like	ND	II	II	N	NV	NI	NI
Str_24	II	II-III	II-III	II-III	II	ND	I	ND	II	II	II-like (atypical-like)	II	II-like	II	Rec I/II	Y	NV	P	NI
Str_25	I	II	II-III	II-III	II-III	I	I	II	I	ND	I	ND	I-like	Rec I/II	Rec I/II	Y	NV	P	V
Str_26	II	II	II-III	II-III	II-III	II	ND	II	II	II-like	II	I	II (atypical-like)	II	Rec I/II	Rec I/II	Y	NV	P	NV
Str_27	II	II	I	II-III	II-III	II	II-III	I	II	II-like	ND	I	II (atypical-like)	II	Rec I/II	Rec I/II	N	NV	NI	NI
Str_28	I	II-III	II-III	II-III	II	ND	ND	ND	ND	ND	II-like (atypical-like)	ND	I	Rec I/II	Rec I/II	N	V	P	NI
Str_29	II	II-III	II-III	II-III	II	ND	II	I	II	ND	II-like (atypical-like)	II (atypical-like)	II	II	Y	NV	NI	NI
Str_30	II	II-III	II-III	II-III	II	II-III	II	I	II-like	II	I	II	II	Rec I/II	N	NV	NI	NI
Str_31	I	I-III	I	I	I	I	ND	II	ND	II-like	ND	II-like (atypical-like)	II	I	I	Rec I/II	Y	V	NA	NA
Str_32	II	II-III	II-III	II-III	II	II-III	II	ND	II-like	II	II-like (atypical-like)	II	II	II	Y	NV	P	NI
Str_33	II	II-III	II-III	II-III	II	II-III	II	II	II	II-like (atypical-like)	II (atypical-like)	II	II	Y	NV	NI	NV
Str_34	I	I-III	I	I	I	ND	II	II	ND	ND	ND	ND	I-like	I	Rec I/II	Y	NV	P	V
Str_35	II	II-III	II-III	II-III	II	II-III	II	ND	II	II-like (atypical-like)	II (atypical-like)	II	II	N	NV	P	V
Str_36	II	II-III	II-III	II-III	II	II-III	II	II	II-like	ND	II-like	I (atypical-like)	II	II	Rec I/II	Y	NV	P	V
Str_37	II	II-III	II-III	II-III	II	ND	II	II	II-like	ND	II-like (atypical-like)	II	II	II	Y	NV	P	NA
Str_38	II	II-III	II-III	II-III	II	II-III	II	ND	II	II-like (atypical-like)	II (atypical-like)	II	II	Y	NV	P	NI
Str_39	II	II-III	II-III	II-III	II	ND	II	II	II	II	II	I	II-like	II	Rec I/II	N	NV	NI	NI
Str_40	II	II-III	II-III	II-III	II	ND	II	I	II	ND	II-like (atypical-like)	II	II	II	Y	NV	NI	NI
Str_41	II	II-III	II-III	II-III	II	ND	II	I	II	ND	II-like (atypical-like)	II	II	II	N	NV	NI	NI
Str_42	II	I-III	II-III	II-III	II	ND	II	II	II-like	II	II-like (atypical-like)	II (atypical-like)	II	Rec I/II or II/III	Rec I/II or II/III	Y	NV	P	NV
Str_43	II	II-III	II-III	II-III	II	ND	II	ND	II-like	ND	II	II (atypical-like)	II	Rec I/II or II/III	Rec I/II or II/III	Y	NV	NI	NI
Str_44	II	II-III	II-III	II-III	II	ND	II	II	II-like	II	II-like (atypical-like)	II	II	II	Y	NV	N	NA
Str_45	II	II-III	II-III	II-III	II	II-III	II	II	II-like	II	II-like (atypical-like)	II	II	II	Y	NV	NI	NI
Str_46	II	II-III	II-III	II-III	II	ND	II	II	II	II	II-like (atypical-like)	II	II-like	II	II	Y	NV	N	NA
Str_47	II	II-III	II-III	II-III	II	ND	II	II	II	II	II-like (atypical-like)	II	II-like	II	Y	NV	P	V
Str_48	II	II-III	II-III	II-III	II	II-III	II	II	II	ND	II-like (atypical-like)	ND	II	II	Y	NV	P	NV
Str_49	II	I-III	I	I	I	ND	II	ND	II-like	ND	ND	ND	II-like	Rec I/II	Rec I/II	N	NV	P	NI
Str_50	II	I	I	I	II-III	II	I	ND	II-like	ND	I (atypical-like)	II	I	Rec I/II	Rec I/II	Y	NV	NI	NI
Str_51	I	I-III	I	I	I	III	ND	NC	ND	III	ND	I	ND	I	Rec I/III	Rec I/III	N	V	NI	NI

[a]Reference sequences from TOXODB

ND - Not detected; NC - Not classified; NI - No information available;

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for *Sag1*.

Strain	Alignment position		Alelle classification	
	51	139	Numerical *	Nominal
GT1 (reference type I)	A	T	1	I
ME49 (reference type II)	G	C	23	II-III
VEG (reference type III)	G	C	23	II-III
ARI	G	C	23	II-III
CAST	A	T	1	I
COUG	A	T	1	I
CtCo5	A	T	1	I
FOU	A	T	1	I
GAB2-2007-GAL-DOM2	A	T	1	I
MAS	A	C	4	I-like
P89	A	T	1	I
RUB	A	T	1	I
TgCATBr5	A	T	1	I
TgCATBr9	A	T	1	I
TgCatPRC2	G	C	23	II-III
VAND	A	T	1	I
RH Lab	A	T	1	I
Str_3	A	T	1	I
Str_7	G	C	23	II-III
Str_20	G	C	23	II-III
Str_21	G	C	23	II-III
Str_25	A	T	1	I
Str_27	G	C	23	II-III
Str_30	G	C	23	II-III
Str_32	G	C	23	II-III
Str_33	G	C	23	II-III
Str_35	G	C	23	II-III
Str_36	G	C	23	II-III
Str_45	G	C	23	II-III
Str_48	G	C	23	II-III
Str_50	G	C	23	II-III
Str_51	A	T	1	I

* Allele number "23" means that there is no differentiation between Type II and Type III.

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for *Gra6*.

Gra6

Strain	29	72	102	126	137	177	193	201	202	226	Numerical	Nominal
GT1 (reference type I)	T	C	G	G	C	G	G	A	A	C	1	I
ME49 (reference type II)	.	T	T	.	.	.	.	.	G	.	2	II
VEG (reference type III)	.	.	T	.	T	.	A	.	.	.	3	III
ARI	.	T	T	A	.	T	.	.	G	.	4	II-like
CAST	.	.	G	.	.	.	.	.	.	.	1	I
COUG	.	T	T	A	.	T	.	.	G	.	4	II-like
CtCo5	.	.	T	.	.	.	A	.	.	.	5	III-like
FOU	.	.	T	.	.	.	.	C	.	.	6	NC
GAB2-2007-GAL-DOM2	.	T	T	.	.	.	.	.	G	.	2	II
MAS	.	T	T	.	.	.	.	.	.	.	7	II-like
P89	.	.	T	.	.	.	A	.	.	.	5	III-like
RUB	.	.	T	.	.	.	.	.	.	T	8	NC
TgCATBr5	C	T	T	.	.	.	.	.	.	.	9	II-like
TgCATBr9	.	.	T	.	.	.	.	C	.	.	6	NC
VAND	.	.	T	.	.	.	.	.	.	.	10	NC
RH_Lab	.	T	T	.	.	.	.	.	G	.	2	II
Str_1	.	T	T	.	.	.	.	.	G	.	2	II
Str_2	.	T	T	.	.	.	.	.	G	.	2	II
Str_3	.	.	G	.	.	.	.	.	.	.	1	I
Str_4	.	T	T	.	.	.	.	.	G	.	2	II
Str_5	.	T	T	.	.	.	.	.	G	.	2	II
Str_6	.	.	G	.	.	.	.	.	.	.	1	I
Str_7	.	T	T	.	.	.	.	.	G	.	2	II
Str_8	.	T	T	.	.	.	.	.	G	.	2	II
Str_9	.	T	T	.	.	.	.	.	G	.	2	II
Str_10	.	.	T	.	T	.	A	.	.	.	3	III
Str_11	.	.	T	.	.	.	.	C	.	.	6	NC
Str_12	.	T	T	.	.	.	.	.	G	.	2	II
Str_13	.	T	T	.	.	.	.	.	G	.	2	II
Str_14	.	T	T	.	.	.	.	.	G	.	2	II
Str_15	.	T	T	.	.	.	.	.	G	.	2	II
Str_16	.	T	T	.	.	.	.	.	G	.	2	II
Str_18	.	T	T	.	.	.	.	.	G	.	2	II
Str_19	.	T	T	.	.	.	.	.	G	.	2	II
Str_20	.	T	T	.	.	.	.	.	G	.	2	II
Str_21	.	T	T	.	.	.	.	.	G	.	2	II
Str_22	.	T	T	.	.	.	.	.	G	.	2	II
Str_23	.	T	T	.	.	.	.	.	G	.	2	II
Str_24	.	.	G	.	.	.	.	.	.	.	1	I
Str_25	.	.	G	.	.	.	.	.	.	.	1	I
Str_26	.	T	T	.	.	.	.	.	G	.	2	II
Str_27	.	.	G	.	.	.	.	.	.	.	1	I
Str_29	.	T	T	.	.	.	.	.	G	.	2	II
Str_30	.	.	G	.	.	.	.	.	.	.	1	I
Str_31	.	T	T	.	.	.	.	.	G	.	2	II
Str_32	.	T	T	.	.	.	.	.	G	.	2	II
Str_33	.	T	T	.	.	.	.	.	G	.	2	II
Str_34	.	T	T	.	.	.	.	.	G	.	2	II
Str_35	.	T	T	.	.	.	.	.	G	.	2	II
Str_36	.	T	T	.	.	.	.	.	G	.	2	II
Str_37	.	T	T	.	.	.	.	.	G	.	2	II
Str_38	.	T	T	.	.	.	.	.	G	.	2	II
Str_39	.	T	T	.	.	.	.	.	G	.	2	II
Str_40	.	T	T	.	.	.	.	.	G	.	2	II
Str_41	.	.	G	.	.	.	.	.	.	.	1	I
Str_42	.	T	T	.	.	.	.	.	G	.	2	II
Str_43	.	T	T	.	.	.	.	.	G	.	2	II
Str_44	.	T	T	.	.	.	.	.	G	.	2	II
Str_45	.	T	T	.	.	.	.	.	G	.	2	II
Str_46	.	T	T	.	.	.	.	.	G	.	2	II
Str_47	.	T	T	.	.	.	.	.	G	.	2	II
Str_48	.	T	T	.	.	.	.	.	G	.	2	II
Str_49	.	T	T	.	.	.	.	.	G	.	2	II
Str_50	.	.	G	.	.	.	.	.	.	.	1	I
Str_51	.	T	T	.	.	.	.	.	G	.	2	II

130

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for *PK1*.

PK1

Strain	8	25	95	97	118	136	154	DEL 200	266	352	403	439	464	533	536	540	606	625	645	716	Numerical	Nominal
GT1 (reference type I)	C	A	G	A	C	G	C	G	C	T	A	C	G	C	G	G	C	G	T	T	1	I
ME49 (reference type II)	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
VEG (reference type III)	.	G	.	.	.	.	.	--	.	A	.	T	A	.	.	T	.	.	C	.	3	III
ARI	.	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	4	II-like
CAST	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
COUG	.	G	C	.	.	A	.	.	T	.	.	T	A	.	.	T	T	.	C	.	5	III-like
FOU	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
GAB2-2007-GAL-DOM2	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
MAS	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	6	I-like
P89	.	A	.	T	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	7	I-like
RUB	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	8	I-like
TgCATBr5	.	G	C	.	.	A	.	.	T	.	G	.	.	T	A	T	.	.	C	.	9	NC
TgCATBr9	.	G	C	.	.	A	.	.	T	.	G	.	.	T	A	T	.	.	C	.	9	NC
TgCatPRC2	.	G	.	.	T	.	A	.	.	.	.	.	.	.	.	T	.	A	C	.	10	II-like
VAND	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	C	.	8	I-like
RH Lab	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
Str_2	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
Str_3	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
Str_4	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_5	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_7	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_8	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_10	.	G	.	.	.	.	.	--	.	A	.	T	A	.	.	T	.	.	C	.	3	III
Str_11	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	T	.	1	I
Str_14	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_15	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_17	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_18	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_19	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_20	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_21	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_22	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_23	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_25	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_26	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_27	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_29	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_30	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_33	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_36	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_37	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_39	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_40	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_41	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_42	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_44	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_45	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_46	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_47	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II
Str_48	G	G	.	.	.	.	A	.	.	.	.	.	.	.	.	T	.	.	C	C	2	II

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for *Sag3*.

Sag3

Strain	| Alignment position 7	8	10	11	14	16	36	40	41	42	62	72	79	81	88	96	107	111	139	| Numerical	Nominal
GT1 (reference type I)	G	A	A	G	G	C	G	A	G	G	A	T	G	G	A	G	A	A	G	1	I
ME49 (reference type II)	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
VEG (reference type III)	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
ARI	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
CAST	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
COUG	.	.	.	.	.	.	.	.	.	.	.	.	.	A	C	G	G	.	.	4	NC (IntraRec I/III)
CtCo5	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
FOU	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
GAB2-2007-GAL-DOM2	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
MAS	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
P89	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
RUB	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
TgCATBr5	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
TgCATBr9	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
TgCatPRC2	.	.	.	.	.	T	A	G	.	.	.	.	C	.	A	C	C	.	.	5	III-like
VAND	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
RH Lab	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
Str_1	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_2	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_3	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
Str_4	.	.	A	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	7	II-like
Str_5	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
Str_7	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_8	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_9	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_10	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
Str_12	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_13	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III
Str_14	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_15	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_16	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_17	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_18	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_19	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_20	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_21	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_22	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_23	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_24	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_25	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	G	.	.	.	1	I
Str_26	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_27	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_29	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_30	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_31	.	.	.	A	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	8	II-like
Str_32	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_33	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_35	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_36	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_37	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_38	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_39	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_40	.	G	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	9	II-like
Str_41	.	.	T	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	10	II-like
Str_42	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_43	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_44	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_45	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_46	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_47	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_48	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	.	2	II
Str_49	.	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	6	II-like
Str_50	A	.	.	.	.	T	A	G	.	.	.	.	.	A	C	C	.	G	T	11	II-like
Str_51	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	C	C	.	.	3	III

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for L363.

L363

Strain	2	19	167	174	289	308	312	322	326	345	365	394	401	424	450	454	461	535	541	554	566	617	618	650	660	667	692	735	743	Numerical	Nominal
GT1 (reference type I)	G	G	T	C	T	A	C	T	A	G	T	A	G	C	G	C	T	G	A	T	T	G	G	G	G	C	G	A	C	1	I
ME49 (reference type II)	C	A	G				A				C				G	C	A	T	T	G	G		A	A				T	C	2	II
VEG (reference type III)		A	G	A			A		C		C		T		G	C	A	T	T	G	G		A	A				T	T	3	III
ARI	A	G					A				C	A			G	C							A	A					C	4	II-Like
CAST	G	A					C				C			C	G	C	C	T	T	G	G							T	C	1	I
COUG	A	G		G	A		A				C			C	G	C	C	T	T	G	G							T	C	5	I1-Like
CtCo5	G	G					C				C																	T	C	6	NC (IntraRec I/II/III)
FOU	G	G					C				C																	T	C	6	NC (IntraRec I/II/III)
GAB2-2007-GAL-DOM2	A	G	A				A						T															T	C	3	III
MAS	G	G				A	A											T								G			C	7	NC (IntraRec I/II)
P89	G	G				A	G	C		A																			C	8	I-Like
RUB	G	G				A	C		A										T				A	A	T			T	C	9	NC (IntraRec I/II/III)
TgCATBr5	G	G					A				C															G			C	7	NC (IntraRec I/II)
TgCATBr9	G	G					C												T									T		6	NC (IntraRec I/II/III)
TgCatPRC2	G	C	G				A						T		G	C		T	T	G	T								C	10	NC (IntraRec I/II)
VAND	G	G					A						T		G	C	A	T	T	G	G		A	A		G			C	11	NC (IntraRec I/II/III)
Str_1	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_4	C	A					A				C				G	C	A	T	T	G	G		A	A			A		C	12	II-like
Str_5	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_6	C	G	C				C																						C	13	NC (IntraRec I/II)
Str_9	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_14	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_15	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_17	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_18	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_20	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_21	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_24	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_26	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_30	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_32	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_33	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_35	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_39	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_40	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_42	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_44	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_46	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II
Str_47	C	A					A				C				G	C	A	T	T	G	G		A	A					C	2	II

133

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for CB21-4

CB21-4

Strain	11	14	15	67	85	96	145	Microsatellite (cont.) [220-223] [a]	INS [224-228]	Microsatellite [229-256] [b]	289	323	Numerical	Nominal
GT1 (reference type I)	G	T	A	G	A	T	T	ATAT	-----	ATATATATATATATAT	A	T	1	I
ME49 (reference type II)	.	G	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	2	II
VEG (reference type III)	.	.	.	.	.	.	G	ATAT	GT	ATATATATATATATATATATATAT	.	C	3	III
ARI	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATATATATAT	.	C	4	II-like (Atypical-like)
CAST	.	.	.	.	.	.	.	ATAT	-----	ATATATATATAT	.	C	5	NC
COUG	.	.	A	.	C	.	.	ATAT	-----	ATATATATATATATATATATAT	.	C	6	II-like (Atypical-like)
CtCo5	.	.	.	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	5	NC
FOU	.	.	.	.	.	.	.	ATAT	-----	ATATATATATATAT	.	T	1	I
GAB2-2007-GAL-DOM2	.	.	.	.	.	.	G	ATAT	GT	ATATATATATATATATAT	.	C	3	III (Atypical-like)
MAS	.	.	G	A	C	.	.	ATAT	GT	ATATATATATATATATAT	.	C	7	NC (IntraRec II/III)
P89	.	T	.	.	.	.	.	ATAT	GTATG	ATATATATATATATATAT	G	C	8	NC
RUB	T	.	.	.	.	.	.	ATAT	-----	ATATATATATATATATAT	.	T	9	I-like (Atypical-like)
TgCATBr5	.	.	G	A	C	.	.	ATAT	GT	ATATATATATATATAT	.	C	7	II-like (Atypical-like)
TgCATBr9	.	.	G	A	C	.	.	ATAT	GT	ATATATATATATAT	.	C	7	II-like (Atypical-like)
TgCatPRC2	.	.	.	A	.	.	.	ATAT	-----	ATATATATATATATATATATATATAT	.	.	4	II-like (Atypical-like)
VAND	.	.	.	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	5	NC
RH Lab	.	.	.	.	.	.	.	ATAT	-----		.	T	1	I
Str 1	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 2	.	.	.	.	.	.	.	ATAT	-----	ATATATATATATAT	.	T	1	I
Str 3	.	G	A	.	.	.	.	ATAT	-----		.	C	2	II
Str 4	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 5	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 6	.	.	.	.	.	.	.	ATAT	-----	ATATATATATATAT	.	T	1	I
Str 7	.	G	A	.	.	.	.	ATAT	-----	ATATATATATAT	.	C	2	II
Str 9	.	G	A	.	.	.	.	ATAT	-----		.	C	2	II
Str 10	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 12	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 14	.	.	.	.	.	.	.	ATAT	-----		.	T	1	I
Str 15	.	G	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	2	II (Atypical-like)
Str 16	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 17	.	.	.	.	.	.	.	ATAT	-----		.	T	1	I
Str 18	.	.	A	.	.	.	.	ATAT	-----		.	C	4	II-like
Str 19	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 21	.	.	A	.	.	.	.	ATAT	-----		.	C	4	II-like
Str 22	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 23	.	.	A	.	.	.	.	ATAT	-----		.	C	4	II-like
Str 24	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 25	.	.	.	.	.	.	.	ATAT	-----		.	T	1	I
Str 26	.	.	.	.	.	.	.	ATAT	-----		.	T	1	I
Str 27	.	.	.	.	.	.	.	ATAT	-----		.	T	1	I
Str 29	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 30	.	.	.	.	.	.	.	ATAT			.	T	1	I
Str 31	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 32	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 33	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 35	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 36	.	.	A	.	.	.	.	ATAT	-----		.	C	4	II-like
Str 37	.	.	A	.	.	.	.	ATAT	-----	ATATATATATAT	.	C	4	II-like (Atypical-like)
Str 38	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 39	.	G	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	2	II
Str 40	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 41	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 42	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 43	.	G	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	2	II
Str 44	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 45	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 46	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 47	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 48	.	.	A	.	.	.	.	ATAT	-----	ATATATATATATAT	.	C	4	II-like (Atypical-like)
Str 50	.	.	.	.	.	.	.	ATAT	-----	ATATATATATAT	.	T	1	I (Atypical-like)
Str 51	.	.	.	.	.	.	.	ATAT	-----	ATATATATATATAT	.	T	1	I

[a] This conserved region corresponds to the beginning of the microsatellite, which is interrupted for some reference strains (including the type III reference strain VEG). As such, the full size of the microsatellite presented in Supplementary Figure 1 for PT isolates corresponds to sum of the two regions presented here.

[b] With exception of reference sequences, the size and profile of the microsatellite presented refers to the most abundant profile detected by NGS (this region is in gray for locus sequenced by Sanger sequencing). Microsatellites diversity was not used for Allele numerical classification.

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for M102.

Strain	4 7	INS [8-11]	27 33 34	INS [49-62]	Microsatellite * [63-84]	INS	92	95	Numerical	Nominal
GT1 (reference type I)	T A	----	A A C	---- ATATATATATATAT	--	G	G	1	I
ME49 (reference type II)		----	G C T	ATAA ATATATATATATATATAT	--	T	A	2	II
VEG (reference type III)	A G	TTTG	. . . TAAATATACATAA ATATATATATATAT		-- .	.	3	III
ARI	. .	----	. . T	---- ATATATATATATAT	GT	. .		4	II-like (Atypical-like)
CAST	. .	----		---- ATATATATATATAT				1	I (Atypical-like)
COUG	. .	----	. . T	ATAA ATATATATATATATAT	--	T A		5	II-like (Atypical-like)
CtCo5	. .	----	. . .	---- ATATATATAT	--	. .		1	I (Atypical-like)
FOU	.	----	. . .	---- ATATATATATAT	--	. .		1	I
GAB2-2007-GAL-DOM2	.	----	. . .	---- ATATATATATAT	--	. .		1	I
MAS	. .	----	G . T	ATAA ATATATATATATATATATAT	G-	. .		6	II-like (Atypical-like)
P89	A G	TTTG	. . . TAAATATACATAA ATATATATATATATATAT	--	. .		3	III (Atypical-like)
RUB	A G	TTTG	. . . TAAATATACATAA ATATATATATATAT	--	. .		3	III (Atypical-like)
TgCATBr5	. .	----	. . . TAAATATACATAA ATATATATATAT	--	- -		7	III-like (Atypical-like)
TgCATBr9	. .	----	. . . TAAATATACATAA ATATATATATA	--	- -		7	III-like (Atypical-like)
TgCatPRC2	. .	----	. . T	---- ATATATATATATATAT	GT	. .		4	II-like
VAND	. .	----	. . T	ATAA ATATATATAT	--	. .		8	II-like (Atypical-like)
RH_Lab	. .	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_1	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_2	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_3	.	----	. . .	---- ATATATATATAT	--			1	I
Str_4	. .	----	G C T	ATAA ATATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_5	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_6	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_7	. .	----	G C T	ATAA ATATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_8	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_9	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_10	. G	----	. . . TAAATATACATAA ATATATATATATATATATAT	--	. .		9	III-like (Atypical-like)
Str_11	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_12	.	----	G C T	ATAA ATATATATATATAT	--	T A		2	II
Str_13	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_14	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_15	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_16	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_17	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_18	. .	----	G C T	ATAA ATATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_20	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_22	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_24	.	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II
Str_26	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_27	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_29	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_30	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_31	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_32	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_33	. .	----	G C T	ATAA ATATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_35	. .	----	G C T	ATAA ATATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_36	. .	----	. . .	---- ATATATATATAT	--	. .		1	I (Atypical-like)
Str_37	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_38	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_39	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_40	.	----	. . .	---- ATATATATATAT	--	. .		1	I
Str_41	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_42	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_43	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_44	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_45	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_46	. .	----	G C T	ATAA ATATATATATATATATAT	--	T A		2	II (Atypical-like)
Str_47	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II
Str_50	.	----	G C T	ATAA ATATATATATATATAT	--	T A		2	II

* With exception of reference sequences, the size and profile of the microsatellite presented refers to the most abundant profile detected by NGS. Microsatellites diversity was not used for Allele numerical classification.

Supplementary Table S6. Mutational profiles used for numerical and nominal allele classification for *Sag2*.

Sag2

Strain	1	7	12	26	29	DEL 43	DEL 44	50	69	90	166	198	225	239	248	267	317	337	Numerical	Nominal
GT1 (reference type I)	T	G	C	G	T	C	A	T	T	G	C	A	G	C	G	G	C	T	1	I
ME49 (reference type II)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
VEG (reference type III)	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	III
ARI	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
CAST	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
COUG	.	.	.	.	.	.	.	G	.	C	.	.	.	.	.	.	.	C	4	II-like
CtCo5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
FOU	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
GAB2-2007-GAL-DOM2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
MAS	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
P89	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	III
RUB	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	5	I-like
TgCATBr5	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	III
TgCATBr9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
TgCatPRC2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
VAND	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
RH Lab	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 3	.	T	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	6	I-like
Str 4	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	A	.	C	7	II-like
Str 5	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 6	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 7	C	T	T	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	8	II-like
Str 8	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	9	II-like
Str 9	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 10	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.	.	10	I-like
Str 11	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	5	I-like
Str 12	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	11	II-like
Str 13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 14	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	12	II-like
Str 15	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	C	13	II-like
Str 16	.	T	T	C	.	.	.	.	.	.	.	.	.	.	.	.	G	C	14	II-like
Str 17	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	.	.	10	I-like
Str 18	.	.	.	.	.	-	-	.	.	.	.	.	.	G	.	.	.	C	15	II-like
Str 19	.	T	.	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	16	II-like
Str 20	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	11	II-like
Str 21	.	T	.	.	.	-	-	.	.	.	.	.	.	.	.	A	G	C	17	II-like
Str 22	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	12	II-like
Str 23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 24	.	.	T	.	.	.	.	G	.	C	G	.	.	.	.	.	G	C	18	II-like
Str 25	.	.	.	.	.	-	-	.	.	.	.	.	.	.	A	.	.	.	19	I-like
Str 26	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 29	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 30	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 31	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 32	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 33	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 34	.	.	.	.	.	-	-	.	.	.	.	.	.	.	A	.	.	.	20	I-like
Str 35	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 36	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 37	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 38	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 39	.	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	21	II-like
Str 40	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 41	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 42	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 43	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 44	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 45	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 46	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	22	II-like
Str 47	C	T	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	23	II-like
Str 48	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	C	2	II
Str 49	.	.	.	.	.	-	-	.	.	.	.	.	.	.	.	.	G	C	11	II-like
Str 50	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I
Str 51	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	I

Supplementary Table S7. Number of types detected *per* locus upon MLST-like numerical allele classification (loci are grouped by genotyping approach).

"RFLP" recognition sites (by Sanger sequening)

VIII *Sag* 2

Types	no. of strains	%
2	40	58,8
1	25	36,8
3	3	4,4
3	**68**	**100,00**

Microsatellite size analysis

VIIa B18

Type	no. of strains	%
2	41	60,3
31	24	35,3
4	3	4,4
3	**68**	**100,00**

II W35

Type	no. of strains	%
23	51	75,0
1	16	23,5
4	1	1,5
3	**68**	**100,00**

IX *Tub* 2

Type	no. of strains	%
23	47	69,1
1	20	29,4
4	1	1,5
3	**68**	**100,00**

XII B17

Type	no. of strains	%
23	50	73,5
1	14	20,6
4	4	5,9
3	**68**	**100,00**

X_TgM_A

Type	no. of strains	%
2	44	64,7
1	12	17,7
3	11	16,2
4	1	1,5
4	**68**	**100,00**

Nucleotide sequencing (Sanger/NGS)

VIII *Sag* 1

Type	no. of strains	%
23	16	50,0
1	15	46,9
4	1	3,1
3	**32**	**100,00**

X *Gra* 6

Type	no. of strains	%
2	41	63,1
1	10	15,4
6	4	6,2
3	2	3,1
5	2	3,1
4	2	3,1
7	1	1,5
8	1	1,5
9	1	1,5
10	1	1,5
10	**65**	**100,00**

VI PK1

Type	no. of strains	%
2	31	62,0
1	8	16,0
9	2	4,0
3	2	4,0
8	2	4,0
10	1	2,0
5	1	2,0
4	1	2,0
6	1	2,0
7	1	2,0
10	**50**	**100,00**

XII *Sag* 3

Type	no. of strains	%
6	26	40,0
2	13	20,0
3	11	16,9
1	8	12,3
4	1	1,5
5	1	1,5
7	1	1,5
8	1	1,5
9	1	1,5
10	1	1,5
11	1	1,5
11	**65**	**100,00**

VIIb L363

Type	no. of strains	%
2	22	56,4
6	3	7,7
1	2	5,1
3	2	5,1
7	2	5,1
4	1	2,6
5	1	2,6
8	1	2,6
9	1	2,6
10	1	2,6
11	1	2,6
12	1	2,6
13	1	2,6
13	**39**	**100,00**

III CB21_4

Type	no. of strains	%
4	30	49,2
1	13	21,3
2	7	11,5
5	3	4,9
7	3	4,9
3	2	3,3
6	1	1,6
8	1	1,6
9	1	1,6
9	**61**	**100,00**

VIIa M102

Type	no. of strains	%
2	29	49,2
1	19	32,2
3	3	5,1
7	2	3,4
4	2	3,4
5	1	1,7
6	1	1,7
8	1	1,7
9	1	1,7
9	**59**	**100,00**

VIII *Sag* 2

Type	no. of strains	%
2	22	32,4
1	18	26,5
3	3	4,4
11	3	4,4
5	2	2,9
10	2	2,9
12	2	2,9
4	1	1,5
6	1	1,5
7	1	1,5
8	1	1,5
9	1	1,5
13	1	1,5
14	1	1,5
15	1	1,5
16	1	1,5
17	1	1,5
18	1	1,5
19	1	1,5
20	1	1,5
21	1	1,5
22	1	1,5
23	1	1,5
23	**68**	**100,00**

Supplementary Figure 1. Genetic heterogeneity of *T. gondii* CB21-4 e M102 microsatellites.
Each graph displays the relative percentage of sequence reads with a particular microsatellite length (in different colors) for the two newly sequenced loci displaying tandem repeats, CB21-4 (A) and M102 (B). Data could be obtained for 32 (CB21-4) and 42 (M102) strains.