We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Abstract

Apicomplexa is a eukaryotic phylum of intracellular parasites with more than 6000 species. Some of these single-celled parasites are important pathogens of livestock. At present, 128 genomes of phylum Apicomplexa have been reported in the GenBank database, of which 17 genomes belong to five genera that are pathogens of farm animals: Babesia, Theileria, Eimeria, Neospora and Sarcocystis. These 17 genomes are Babesia bigemina (five chromosomes), Babesia divergens (514 contigs) and Babesia bovis (four chromosomes and one apicoplast); Theileria parva (four chromosomes and one apicoplast), Theileria annulata (four chromosomes), Theileria orientalis (four chromosomes and one apicoplast) and Theileria equi (four chromosomes and one apicoplast); Eimeria brunetti (24,647 contigs), Eimeria necatrix (4667 contigs), Eimeria tenella (12,727 contigs), Eimeria acervulina (4947 contigs), Eimeria maxima (4570 contigs), Eimeria mitis (65,610 contigs) and Eimeria praecox (53,359 contigs); Neospora caninum (14 chromosomes); and Sarcocystis neurona strains SN1 (2862 contigs) and SN3 (3191 contigs). The study of these genomes allows us to understand their mechanisms of pathogenicity and identify genes that encode proteins as a possible vaccine antigen.

Keywords: Apicomplexa, genomics, parasitic protists, Babesia, Theileria, Eimeria, Sarcocystis, Neospora

1. Introduction

Apicomplexa (also called Apicomplexia) is a group of protists comprising a eukaryotic phylum of obligate intracellular parasites with more than 6000 described species [1]. Many of these cell single parasites are important pathogens of humans, domestic animals and livestock, with a health and economic relevance worldwide [2–5]. Apicomplexa microorganisms are intracellular eukaryotes thriving within another eukaryotic cell [6].
This phylum includes *Plasmodium falciparum* and four other *Plasmodium* species, the etiological agents for malaria in humans, a mosquito-transmitted and potentially deadly disease [6]. *Toxoplasma gondii* is a source of toxoplasmosis disease and congenital neurological birth defects (for example, encephalitis and ocular disease) in humans [7–9]. *Cryptosporidium* and *Cyclospora* parasites cause opportunistic human infections associated with immunosuppressive conditions (including AIDS) through contaminated food or water supplies [10, 11], while the invertebrate parasites of genus *Gregarina* are used as models for studying Apicomplexa motility [12].

Apicomplexa parasites infect a wide range of animals from mollusks to mammals [13]. Their life cycles involve only a single host, whereas others require sexual recombination in a vector species for transmission. The life cycle of these parasites has three stages: sporozoite (infective stage), merozoite (a result of asexual reproduction) and gametocyte (germ cells) [12]. These parasites are characterized by the presence of specific organelles (including rhoptries, micronemes and dense granules) involved in the establishment of an intracellular parasitophorous vacuole within the host cell [12].

A defined feature of these microorganisms is the presence of extracellular zoite forms that are usually motile and include an apical complex that gives the phylum its name [14]. With the exception of the genera *Cryptosporidium* and *Gregarina*, all species of the phylum Apicomplexa possess an apicoplast [12, 15–17].

The Apicomplexa parasites causing diseases of veterinary importance are *Babesia*, *Theileria*, *Eimeria*, *Neospora* and *Sarcocystis* [11, 18, 19]. This chapter focuses on genomics of these five genera.

2. Apicomplexa genome

2.1. Apicoplast genome

Twenty years ago, a remnant chloroplast, known as apicoplast, was discovered in *Plasmodium* [20–23]. This apicoplast lost the ability to perform photosynthesis, however, is an essential organelle, and its inhibition is lethal. The apicoplast arose from a secondary endosymbiosis event occurred where an ancestor to *Plasmodium* engulfed a photosynthetic alga [24–26]. This organelle is involved in critical metabolic pathways such as the biosynthesis of fatty acids and heme group degradation [27, 28]. Some of these metabolic pathways are considered as potential targets for antiparasitic drug designs [29, 30].

Like mitochondria, the apicoplast possesses its own genome [29, 31–37]. The apicoplast genome is ~35 kbp smaller than chloroplasts due to the absence of genes encoding proteins involved in photosynthesis. The genome of this plastid has been reduced and contains ribosomal (rRNA) and transfer RNA (tRNA) genes that play an important role in organelle replication [24]. The characteristics of the structure of apicoplast genomes have difficult comparisons with other plastids [20].
2.2. Apicomplexa genomes in GenBank

New drug targets identification, and novel antiparasitic therapeutics are necessary due to the emergence of parasite strains resistant to treatments available today [12, 38–40]. With the recent advancements in genome sequencing technologies, the research of new drug targets can be the focus on genomics analyses.

At present (August 2016), 128 complete and draft genomes of phylum Apicomplexa have been reported in the GenBank database (http://www.ncbi.nlm.nih.gov/genbank/), of which 17 genomes belong to five genera that are pathogens of farm animals: Babesia, Theileria, Eimeria, Neospora, and Sarcocystis (3, 4, 7, 1, and 2 genomes, respectively). The study and comparison of these genomes will allow us to understand pathogenicity mechanisms and identify genes and proteins with potential drug targets in order to develop novel antiparasitic compounds of veterinary importance.

3. Classification of phylum Apicomplexa

The National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) divides the phylum Apicomplexa into two classes: Aconoidasida and Conoidasida (Figure 1). The class Aconoidasida is divided into two orders: Haemosporida and Piroplasmida.
(containing the genera Babesia and Theileria), while the class Conoidasida is divided into two subclasses: Coccidiasina (containing the genera Eimeria, Neospora and Sarcocystis, that belong to order Eucoccidiorida) and Gregarinasina (Figure 1).

It is estimated that subclass Coccidiasina separated from the class Aconoidasida ~705 million years ago [41, 42]. Moreover, in 2004, Douzery et al. calculated it as 495 million years ago [41–43].

4. Babesia

Babesia is a genus of intracellular protozoa that cause babesiosis. These parasites are transmitted by ticks and infect erythrocytes in their mammalian hosts. Babesiosis was first described in sheep and cattle in 1888 by Victor Babes, in honor of which is called the genus [44] and is characterized by hemolytic anemia and fever, with occasional hemoglobinuria and death [45].

The genus Babesia includes over 100 species that are highly specific for their hosts. Only a few Babesia species cause infections in humans, especially immunocompromised individuals. Most cases identified in humans are caused by Babesia microti and Babesia divergens, parasites of rodents and cattle, respectively [44, 46, 47].

Species affecting animals are: Babesia bigemina, Babesia major, Babesia divergens and Babesia bovis that infect cattle [44, 48–51]; Babesia ovis and Babesia motasi cause infections in sheep [44, 52, 53]; and Babesia equi and Babesia caballi cause infections in horses [44, 54].

Three genomes of Babesia species have been reported in the GenBank database. The B. bigemina strain Bond genome is 13,840,936 bp of total length divided into five chromosomes (2.5, 2.8, 3.5, 0.9 and 0.5 Mbp; GenBank accession number from NC_027216.1 to NC_027220.1, respectively). The B. divergens strain Rouen 1987 genome is 10,797,556 bp divided into 514 contigs (GenBank accession number CCSG00000000.1).

B. bovis strain T2Bo genome is 8,179,706 bp divided into four chromosomes (1.2, 1.7, 2.6 and 2.6 Mbp, respectively) and one apicoplast (35,107 bp, GenBank accession number NC_011395.1). The chromosomes I and IV of B. bovis genome are divided into seven and three contigs, respectively; chromosomes II and III GenBank accession numbers are NC_010574.1 and NC_010575.1, respectively.

4.1. Babesia bovis genome

In 2007, Brayton et al. reported the analysis of comparative genomic between B. bovis, Theileria parva and P. falciparum genomes [33]. The B. bovis genome has 3671 protein-coding genes and 41.8% of GC content, an analysis of enzymatic pathways revealed a reduced metabolic potential. The results of comparative genomic showed that B. bovis genome (8.2 Mbp) is similar in size to that of T. parva (8.3 Mbp) [34] and Theileria annulata (8.35 Mbp) [55], the smallest Apicomplexa genomes sequenced to date.
In contrast, *B. bovis* and *P. falciparum*, which have similar clinical and pathological features, have major differences in genome size (8.2 and 22.8 Mbp, respectively) and chromosome number (4 and 14, respectively). Additionally, many stage-specific and immunologically important genes from *P. falciparum* are absent in *B. bovis* [33]. The *B. bovis* genome sequence has allowed analyses of the polymorphic variant erythrocyte surface antigen protein (*ves1* gene and discovery of the novel *smorf* gene family) that are postulated to play a role in cytoadhesion and immune evasion (similar to *var.* genes of *P. falciparum*). The ~150 *ves1* genes are distributed in clusters throughout each chromosome [33]. Finally, comparative analyses have identified several novel vaccine candidates into *B. bovis* genome, including homologs of p36 and Pf12 (*P. falciparum*); p67 and four of six proteins (*T. parva*) targeted by CD8+ cytotoxic T cells [33].

Brayton et al. also reported that the *B. bovis* apicoplast genome is 33 kbp of total length and encodes 32 putative protein coding genes, 25 tRNA genes, and small and large subunit rRNA genes. This organelle genome displays similarities in size and gene content to apicoplasts of *Eimeria tenella, P. falciparum, T. parva* and *T. gondii* [33, 35, 56]. The *B. bovis* apicoplast genome has 78.2% of AT content (21.8% of GC content) [33].

5. Theileria

The genus *Theileria* infects leukocytes [57], and they are the only eukaryotic pathogens known to transform lymphocytes [11]. These parasites infect a wide range of both domestic and wild animals and are transmitted by Ixodid ticks of the genera *Amblyomma*, *Haemaphysalis*, *Hyalomma* and *Rhipicephalus* [58, 59]. *Theileria* parasites can be grouped into schizont transforming (*T. parva, T. annulata* and *Theileria lestoquardi*) [60–62] and nontransforming (*Theileria orientalis*) species [63, 64]. The uncontrolled proliferation of schizonts results in the pathologies associated with corridor disease and East Coast fever (*T. parva*), tropical theileriosis (*T. annulata*) in cattle and malignant theileriosis (*T. lestoquardi*) in goats and sheep [59, 65]. *T. orientalis* (frequently been referred to as *T. sergenti* [66]) causes bovine piroplasmosis [67–69] and can generate anemia and icterus in cattle but rarely cause fatal disease [64]. *T. orientalis* is classified into two major genotypes: the Chitose (throughout the world) and Ikeda (eastern Asian countries) types [70]. Finally, equine piroplasmosis of horses, mules, donkeys, and zebras is caused by *Theileria equi* [71]. *T. equi* has been renamed several times [72], and molecular phylogenetic analyses indicate an intermediate position between *B. bovis* and *Theileria* spp. [73, 74].

Four genomes of *Theileria* species have been reported in the GenBank database. The *T. parva* strain Muguga genome is 8,347,606 bp divided into four chromosomes (2.5, 2.0, 1.9 and 1.9 Mbp) and one apicoplast (39,579 bp, GenBank accession number NC_007758.1). The chromosomes I and II of *T. parva* genome have the GenBank accession number NC_007344.1 and NC_007345.1, respectively, while the chromosomes III and IV are divided into four and two contigs, respectively. The *T. annulata* strain Ankara isolate clone C9 genome is 8,358,425 bp divided into four chromosomes (2.6, 2.0, 1.9 and 1.8 Mbp; GenBank accession number
NC_011129.1, NC_011099.1, NC_011100.1 and NC_011098.1, respectively). The *T. orientalis* strain Shintoku (Ikeda type) genome is 9,010,364 bp divided into four chromosomes (2.7, 2.2, 2.0 and 2.0 Mbp; GenBank accession number from NC_025260.1 to NC_025263.1, respectively) and one apicoplast (24,173 bp into one contig).

Finally, the *T. equi* strain WA genome is 11,674,479 bp divided into four chromosomes (3.7, 2.3, 2.1 and 3.5 Mbp) and one apicoplast (47,880 bp into one contig). The chromosomes I and III of *T. equi* genome have the GenBank accession number NC_021366.1 and NC_021367.1, respectively, while the chromosomes II and IV are divided into two and six contigs, respectively.

### 5.1. *Theileria parva* genome

The complete genome sequence of *T. parva* was reported in 2005 [34]. *T. parva* genome has 4035 protein encoding genes (20% fewer than *P. falciparum*) and 34.1% of GC content. Putative functions were assigned to 38% of the predicted proteins. Like *P. falciparum*, the chromosomes of *T. parva* contain one extremely A + T-rich region (>97%) about 3 kbp in length that may be the centromere [34]. Unlike *P. falciparum*, *T. parva* genome contains two identical, unlinked 5.8S-18S-28S rRNA units, which suggest that it does not possess functionally distinct ribosomes [75]. The infection of T and B lymphocytes by *T. parvum* results in a reversible transformed phenotype with uncontrolled proliferation of host cells that remain persistently infected. Parasite proteins that may modulate host cell phenotype are described by [55]. Telomeres of *T. parvum* have a conserved (~140 bp) sequence adjacent to the telomeric repeat and several subtelomeric regions exhibit 70–100% sequence similarity [34, 76]. The apicoplast genome of *T. parva* differs from *P. falciparum* in that all of its genes are transcribed in the same direction, and 26 of the 44 protein-coding genes share 27–61% sequence similarity with proteins encoded by the *P. falciparum* apicoplast genome [34].

### 5.2. *Theileria annulata* genome

The *T. annulata* genome sequence was also reported in 2005 [55]. The nuclear genome of *T. annulata* is similar in size (8.35 Mbp) to that of *T. parva* (8.3 Mbp). *T. annulata* genome has 3792 protein encoding genes (243 genes fewer than *T. parva*), 49 tRNA and 5 rRNA genes, and 32.54% of GC content. In addition, 3265 orthologous genes were predicted between *T. annulata* and *T. parva* genomes. Pain et al. predicted 3265 orthologous genes between the *T. annulata* and *T. parva* genomes. Additionally, 34 (*T. annulata*) and 60 (*T. parva*) genes are single-copy genes and their functions have been not described [55].

The parasite genes involved in host-cell transformation require a signal peptide or a specific host-targeting signal sequence. Some candidates include TashAT and SuAT protein families in *T. annulata* [77, 78] and related host nuclear proteins (TPhNs protein family) in *T. parva*. A cluster of 17 SuAT1 and TashAT-like genes was identified in the *T. annulata* genome [55].

### 5.3. *Theileria orientalis* genome

In 2012, Hayashida et al. reported the comparative genomic analyses between *T. orientalis*, *T. parva*, *T. annulata* and *B. bovis*. The genome size of *T. orientalis* (9 Mbp) is approximately
8% larger than the reported genome sizes of *T. parva* (8.3 Mbp), *T. annulata* (8.35 Mbp) and *B. bovis* (8.2 Mbp). The number of predicted protein-coding (3995) genes identified in *T. orientalis* is similar to that found in *T. parva* (4035). The number of predicted protein-coding (3995) genes identified in *T. orientalis* is similar to that found in *T. parva* (4035). The GC content of the *T. orientalis* genome (41.6%) is higher than *T. parva* and *T. annulata* (34.1% and 32.5%, respectively) but similar to *B. bovis* (41.8%). Unlike *T. parva* and *T. annulata*, *T. orientalis* does not induce uncontrolled proliferation of infected leukocytes and multiplies predominantly within infected erythrocytes [79]. *T. orientalis* is the first genome sequence of a nontransforming *Theileria* species that occupies a phylogenetic position close to that of the transforming species [79].

### 5.4. *Theileria equi* genome

The *T. equi* genome sequence was reported in 2012 [80]. *T. equi* genome size (11.6 Mbp) is larger than *T. parva* (8.3 Mbp), *T. annulata* (8.35 Mbp), *T. orientalis* (9 Mbp) and *B. bovis* (8.2 Mbp). *T. equi* genome has two rRNA operons, 46 tRNA genes and 5330 nuclear protein coding genes, ~25% greater than found for *T. parva*, *T. annulata* and *B. bovis*. Furthermore, *T. equi* genome contains 1985 unique genes, and 366 and 137 homologs of genes found only in the two *Theileria* spp. or *B. bovis*, respectively. The apicoplast genome of *T. Equi* has 43 unidirectionally coding sequences, which includes each of the 20 rRNA, and two rRNA genes are present [80].

### 6. *Eimeria*

*Eimeria* is a genus that includes species capable of causing the disease coccidiosis in cattle and poultry. *Eimeria* parasites exhibit immense diversity in host range including mammals, birds, reptiles, fish and amphibians [81–86]. It is estimated that there are many thousands of *Eimeria* species [87]. Coccidiosis is primarily associated with enteric disease with few exceptions [88–90]. The avian coccidiosis can be subdivided into hemorrhagic and malabsorptive pathologies related to *Eimeria brunetti*, *Eimeria necatrix* and *Eimeria tenella*; or *Eimeria acervulina*, *Eimeria maxima*, *Eimeria mitis* and *Eimeria praecox*, respectively [91]. *E. tenella* is among the most pathogenic avian parasites causing weight loss, reduced feed efficiency, reduced egg production and death [92]. The total loss is estimated at around USD 2.4 billion annually [93], including the costs of control and prevention worldwide.

Seven genomes of *Eimeria* species have been reported in the GenBank database. The *E. brunetti* strain Houghton genome is 66,890,165 bp divided into 24,647 contigs (GenBank accession number CBUX00000000.1). The *E. necatrix* strain Houghton genome is 55,007,932 bp divided into 4667 contigs (GenBank accession number CBUZ00000000.1). The *E. tenella* strain Houghton genome is 51,859,607 bp divided into 12,727 contigs (GenBank accession number CBUW00000000.1). The *E. acervulina* strain Houghton genome is 45,830,609 bp divided into 4947 contigs (GenBank accession number CBUS00000000.1). The *E. maxima* strain Weybridge genome is 45,975,062 bp divided into 4570 contigs (GenBank accession number CBUY00000000.1). The *E. mitis* strain Houghton genome is 60,415,144 bp divided into 65,610 contigs (GenBank accession number CBUT00000000.1). *E. praecox* strain Houghton genome is 60,083,328 bp divided into 53,359 contigs (CBUU00000000.1).
E. tenella strain Houghton was isolated in the United Kingdom in 1949. The E. tenella genome size is ~60 Mbp with a GC content of ~53%. Its molecular karyotype comprises 14 chromosomes of between 1 and >6 Mbp, and the genome is available in http://www.sanger.ac.uk/Projects/E_tenella/. Moreover, parallel projects have been undertaken to generate the complete sequences of chromosomes I (~1 Mbp) and II (~1.2 Mbp), which are associated with resistance to the anticoccidial drug arprinocid and precocious development, respectively [94]. In 2007, Ling et al. reported the sequencing and analysis of the first chromosome of E. tenella [95]. The chromosome I of E. tenella is 1,347,714 pb of total length and has the GenBank accession number AM269894.1.

7. Neospora

The genus Neospora is constituted by only two species: Neospora caninum and Neospora hughesi. N. caninum is the etiologic agent of the disease neosporosis and is a close relative of T. gondii [96]. They share many common morphological and biological features [97]. Neospora parasite appears not to be zoonotic, having a more restricted host range [98, 99], and shows a striking capacity for highly efficient vertical transmission in bovines [100]. N. caninum is one of the leading causes of infectious bovine abortion [101, 102].

Only one genome of N. caninum strain Liverpool has been reported in the GenBank database. This genome has 57,547,420 bp of total length divided into 14 chromosomes: Ia (2,288,409 bp), Ib (1,908,326 bp), II (2,170,133 bp), III (2,139,717 bp), IV (2,317,323 bp), V (2,735,753 bp), VI (3,360,651 bp), VIIa (3,947,736 bp), VIIb (4,923,984 bp), VIII (6,723,156 bp), IX (5,490,906 bp), X (6,985,512 bp), XI (6,081,843 bp) and XII (6,473,971 bp); GenBank accession number from NC_018385.1 to NC_018398.1.

8. Sarcocystis

More than 150 species of Sarcocystis have an indirect life cycle. They require both an intermediate and a final host, usually a herbivorous and a carnivorous vertebrate animal, respectively [103]. For this transition, Sarcocystis species produce infectious tissue cysts surrounded by glycosylated cyst walls that are largely restricted to muscle. Ingestion of tissue cysts through predation by the final hosts propagates the life cycle [104]. All vertebrates, including mammals, some birds, reptiles and possibly fish, are intermediate hosts to at least one Sarcocystis species [105, 106]. Final hosts include carnivores or omnivores, such as humans, some reptiles and raptorial birds [107].

Sarcocystis species are the causal agents of Sarcocystosis, a disease typically asymptomatic but can be associated with myositis, diarrhea or infection of the central nervous system [104]. Some species of Sarcocystis that infect farm animals (such as cattle, sheep and horses) cause fever, lethargy, poor growth, reduced milk production, abortion and death [107].
Sarcocystis cruzi, Sarcocystis hirsuta and Sarcocystis hominis used cattle as intermediate hosts, and canids, felids and humans as final hosts, respectively [108]. Additionally, Sarcocystis sinensis also used cattle as intermediate host [109], but its final host remains to be elucidated [110]. *S. hominis* causes gastrointestinal malaise [108] and *S. sinensis* may also elicit symptoms in humans [111].

*Sarcocystis neurona* is the causal agent of equine protozoal myeloencephalitis [106]. This disease destroys neural tissue and can be fatal to horses, marine mammals and several other mammals. *S. neurona* also infects many mammals asymptomatically [104]. Furthermore, three *Sarcocystis* species have been identified from pigs: *Sarcocystis miescheriana, Sarcocystis porcifelis* and *Sarcocystis suihominis* [112]. In 2015, Blazejewski et al. reported the first genome sequence of *S. neurona* strain SN1 [104].

Two genomes of *S. neurona* strains have been reported in the GenBank database. The *S. neurona* strain SN3 clone E1 genome is 124,404,968 bp divided into 3191 contigs (GenBank accession number JAQE0000000.1). *S. neurona* strain SN1 genome is 130,023,008 bp divided into 2862 contigs (GenBank accession number JXWP0000000.1). *S. neurona* strain SN1 was isolated from an otter that died of protozoal encephalitis [113]. The apicoplast genome architectures of *S. neurona* strains SN1 and SN3 are highly similar to those of *Toxoplasma gondii* and *Plasmodium falciparum* [104]. *S. neurona* strains SN1 and SN3 are the first genomes reported in the genus *Sarcocystis*. These genomes are more than twice the size of other sequenced coccidian genomes.

Acknowledgements

This work was supported by CONACyT scholarship 293,552.

Author details

Fernando Martinez-Ocampo

Address all correspondence to: fernando.martinez@uaem.mx

Laboratory of Ecogenomic Studies, Biotechnology Research Center (CEIB), Autonomous University of the State of Morelos (UAEM), Cuernavaca, Morelos, Mexico

References


[64] Onuma M, Kakuda T, Sugimoto C. Theileria parasite infection in East Asia and control of the disease. Comparative Immunology, Microbiology and Infectious Diseases (England). 1998 Jul;21(3):165-177


[110] Dubey JP, Fayer R, Rosenthal BM, Calero-Bernal R, Uggla A. Identity of Sarcocystis species of the water buffalo (Bubalus bubalis) and cattle (Bos taurus) and the suppression of Sarcocystis sinensis as a nomen nudum. Veterinary Parasitology (Netherlands). 2014 Sep;205(1-2):1-6

