Studies on the development of dourine specific diagnostic methods

2018

Batdorj Davaasuren

Doctoral program in Animal and Food Hygiene

Graduate school of Animal Husbandry

Obihiro University of Agriculture and Veterinary Medicine

媾疫特異的診断技術の確立に向けた研究

平成 **30** 年

（**2018**）

帯広畜産大学大学院畜産学研究科

畜産衛生学専攻 博士後期課程

バトドルジ　ダバスレン

Contents

List of tables

Table 1. Sample information and data summary of this study 25
Table 2. The PCR primer list for the chapter 27
Table 3. Comparison of the results of the different serodiagnostic tests 28
Table 4. Differences and similarity among the Trypanozoon 42
Table 5. Genome sequence assemblies 43
Table 6. The summary of annotated genome statistics of T. equiperdum IVM-t1 44

List of figures

Abbreviations

µL	microliter
AAT	Animal African Trypanosomosis
AFLP	amplified fragment length polymorphism
AHT	Animal Health Trust
ASFCA	Abu Dhabi Food Control Authority
BSF	Bloodstream form
CATT	Card agglutination test for trypanosomosis
CDC	Centers for Disease Control and Prevention
CDS	Coding sequences
cELISA	Competitive ELISA
CFT	Complement fixation test
CSF	cerebrospinal fluid
DNA	Deoxyribonucleic acid
ELISA	Enzyme linked immunosorbent assay
EMF	epimastigote form
FAO	Food and Agriculture Organization of the United Nations
GB	Gigabyte
GmbH	Gesellschaft mit beschränkter Haftung
ICT	immunochromatographic test
IFAT	immunofluorescent antibody test
IgG	immunoglobulin G
ITS	internal transcribed spacer
LAMP	Loop-mediated isothermal amplication
MB	Megabyte
Mbp	Mega base pair
MCF	metacyclic form
MEGA	multiplex-endonuclease genotype
ml	Milliliter
NGS	Next generation sequencing
NTTAT	Non-tsetse transmitted animal trypanosomosis
OD	Optical density
OIE	Office International des Epizooties

PAGs	Procylin-associated genes
PBS	Phosphate buffered saline
PBS-T	Phosphate buffered saline with tween 20
PCF	Procyclic form
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
RFLPs	Restriction fragment length polymorphisms
RNA	Ribonucleic acid
rTeGM6-4r	Recombinant *T. evansi* GM6 4-repeat
TBS	Tris-buffered saline
TBS-T	Tris-buffered saline with tween 20
TR	Tandem repeat
USA	United States of America
USD	United States dollar
VSG	Variable surface glycoprotein
kDNA	Kinetoplast DNA

General introduction

1.1 Taxonomy, history and biology of non-tsetse transmitted trypanosome

The taxonomic position of genus *Trypanosoma* according to J. Stevens (2004) is as follows:

Kingdom Protista

Phylum Sarcomastigophora

Class Zoomastigophorea

Order Kinetoplastida

Family Trypanosomatidae

Genus *Trypanosoma*

The genus *Trypanosoma* is divided into two sections, Stercoraria and Salivaria (Hoare, 1972). The section Stercoraria is mainly divided into three subgenera, namely *Megatrypanum, Herpetosoma* and *Schizotrypanum*. On the other hand, the Salivarian section is divided into four subgenera: *Duttonella* (*T. vivax*), *Nannomonas* (*T. congolense*), *Pycnomonas* (*T. suis*) and *Trypanozoon* (*T. brucei* sspp*., T. evansi* and *T. equiperdum)*.

Among the Stercorarian trypanosomes, *T. cruzi* is the most pathogenic for man and dogs. *T. cruzi* is cyclically transmitted via the excreta of *Triatoma*. When infected *Triatoma* feces contaminate the bite site or mucous membranes of mammalian hosts, *T. cruzi* can infect and cause symptomatic and potentially life-threatening Chagas' disease (Bern et al., 2011).

In section of the Salivarian trypanosomes, *T. vivax* (subgenus *Duttonella*) causes animal African trypanosomosis (AAT, known as Nagana) mainly in West Africa and South America together with *T. congolense*, *T. simiae*, *T. b. brucei* and *T. suis* (J. Stevens, 2004). *T. vivax* is responsible for more than half of the total animal trypanosomosis in West Africa where it is considered the major pathogen for livestock and small ruminants (Chamond et al., 2010).

T. congolense (subgenus *Nannomonas*) is probably the most prevalent and widespread virulent species for broad range of domestic mammals throughout the tsetse belt. *T. congolense* is cyclically transmitted by tsetse flies in Africa (Peacock et al., 2012).

Subgenus *Trypanozoon* includes three species: *T. brucei* sspp*., T. evansi and T. equiperdum*. *T. brucei* sspp. (*T. b. brucei*, *T. b. rhodesiense*, and *T. b. gambiense*) are distributed only in the tsetse belt in sub-Saharan Africa. They are cyclically transmitted by the bite of tsetse fly (*Glossina* spp.) (Simarro et al., 2010). Bloodstream form (BSF) of the trypanosome are propagated in blood. When tsetse fly bite the infected vertebrate host, BSF trypanosomes are ingested into the midgut of tsetse fly with blood meal. In the tsetse midgut the BSFs are differentiated into procyclic form (PCF), and then epimastigote form (EMF) which migrates to the salivary glands and transforms into metacyclic form (MCF). Finally, MCF can be injected into a new mammalian host during blood meal of tsetse fly (CDC, 2017).

On the other hand, *T. evansi* and *T. equiperdum* are not cyclically transmitted by tsetse fly but mechanically transmitted by several blood sucking insects and via coitus, respectively. Whereas these two trypanosomes and trypanosomosis are called “Non-tsetse transmitted animal trypanosomes” causing “Non-tsetse transmitted animal trypanosomosis (NTTAT)” (J. Stevens, 2004).

T. equiperdum can infect only equids and cause dourine (Brun et al., 1998). Unlike other trypanosomes, *T. equipedum* is not transmitted by insect vectors, but it is transmitted from infected horse to healthy horse via coitus. Thus, *T. equiperdum* has previously been distributed worldwide beyond tsetse belt in sub-Saharan Africa, the epidemic area of African trypanosome. *T. equiperdum* primarily parasitize the genital organ mucosa after coitus, and subsequently invade into several tissues, subcutaneous tissues, mammary glands and cerebrospinal fluid (CSF) (Mehlhorn and Aspöck, 2008). They rarely parasitize and propagate in blood stream unlike other trypanosome, therefore it is very difficult to detect *T. equiperdum* in blood.

Whilst, *T. evansi* infects a wide range of domestic and wild animals, and cause surra which presents as acute disease in camels, horses, dogs and chronic disease in cattle and buffaloes (Desquesnes et al., 2013). Wild animals such as capybaras and coatis may act as reservoirs of *T. evansi* (Silva, 1995). *T. evansi* is distributed worldwide beyond tsetse belt due to the mechanical transmission of blood sucking insects, typically tabanids. In addition, *T. evansi* can also be transmitted by vampire bats, which act as the reservoirs and vectors in South America (Jensen et al., 2008). Transmission can be vertical, horizontal, iatrogenic and *per* oral, with various epidemiological significances depending on the season, location and host species (Desquesnes et al., 2013). *T. evansi* is an obligate blood parasite that spreads through the entire body via the blood stream similar to *T. brucei* sub-spp.

The kDNA of *T. brucei* contains heterogeneous mini-circles and complete gene sets on maxi-circles. The genes in maxi-circle mainly form oxidative respiratory chain in mitochondria, and mainly function in insect stage trypanosome (Johnson et al., 1982). A mini-circle encodes a few number of guide RNA for RNA editing of the transcripts from maxi-circle genes. In the insect stage of *T. brucei*, these mitochondria related genes on maxi-circle are translated after RNA editing by heterogeneous guide RNA from heterogeneous sequence of mini-circles (Clement et al., 2004). On the other hand, *T. evansi* lacks mini-circle sequence heterogeneity and maxi-circle. *T. equiperdum* also has homogeneous mini-circle and various condition of maxi-circle (depending on the strains) (Songa et al., 1990). They lack functional kinetoplast, therefore, they cannot completely translate mitochondrial respiratory chain genes from the partial maxi-circle and/or cannot edit the transcripts from maxi-circle DNA due to the homogeneous set of guide RNA. This means that the mitochondrion of *T. evansi* and *T. equiperdum* cannot be adapted into low O_2 concentration condition inside the insect. Thus, they cannot differentiate into insect stage trypanosome and cannot be cyclically transmitted by tsetse fly (Borst et al., 1987).

T. equiperdum initially documented by Rouget in dourine horse blood at 1894 (Rouget, 1896). Buffard and Schneider inoculated it in dog and replicated infection in healthy horse and they published the result in 1900 (Buffard, 1900). Based on these evidence, etiological agent of the dourine was named *T. equiperdum* (Latin: equus = horse; perdere = to loose) by Doflein in 1901(Doflein, 1901).

A type of *Trypanosoma* initially reported as agent of surra by Griffith Evans in 1880 when he worked in endemic disease in equines and camels in the Punjab in India (modern days in Pakistan)(Steel et al., 1886). It was initially described as *T. evansi* by John Henry Steel in 1885 and the subject of second detailed description by Edouard-Gérard Balbiani in 1888 (Antoine-Moussiaux and Desmecht, 2008).

There are various ecological differences among *Trypanozoon* trypanosomes, for example host range, mechanisms of transmission and so on; however, species classification among *Trypanozoon* has been controversial topic (Claes et al., 2005). They cannot be morphologically distinguished from each other because it is thought that they are recently evolved from their common ancestor trypanosome (Desquesnes et al., 2013). Moreover, Hoare suggested that *T. equiperdum* evolved from *T. evansi* through adaptation of tissue parasitism (Hoare, 1972). Although this idea has not been accepted because once the parasite lost kDNA it is not possible to regain it (Brun et al., 1998). Opposing suggestion is that *T. equiperdum* and *T. evansi* are derived from ancient strain *of T. brucei* at different time frames (Gizaw et al., 2017). Very few

single copy genes are available for genetic differentiation such as SRA gene of *T. b. rhodesiense* (Radwanska et al., 2002), TgsGP gene of *T. b. gambiense* is absent in some foci (Berberof et al., 2001) and RoTat1.2 gene of *T. evansi* which is also absent in some foci (Claes et al., 2003b; Ngaira et al., 2005).

1.2 Disease impact, clinical symptoms and epidemiology of non-tsetse transmitted horse trypanosomosis

1.2.1 Epidemiology

Dourine

Dourine has been reported in Asia, Africa, South America, and south-eastern Europe. North America and western European countries were reported to have successfully eradicated dourine (Claes et al., 2005). In Italy, dourine was originally eradicated in the 1940s, but in the 1970s a serious epidemic occurred and in 2011, the disease reemerged in Italy (Podaliri Vulpiani et al., 2013). In Ethiopia, the first official report of the dourine was made in 1980 and since then in 1997, 1999 and 2010 the disease was reported with seroprevalence up to 28% (Hagos et al., 2010a). Dourine has occasionally occurred in Mongolia near the border where the horse movement was free in the past (unpublished data). In 2003, dourine was reported with 5.5 % seroprevalence in Mongolia (Clausen et al., 2003). Our ongoing surveillance project in Mongolia also revealed surra epidemics in other domestic animals (Musinguzi, 2017). Recently, a new *T. equiperdum* strain was isolated from the urethral mucosa of a horse in Mongolia. That horse was shown typical clinical sign of dourine (Suganuma et al., 2016). Thereafter, four other strains were isolated from Dongovi, Hentei and Tov provinces of Mongolia (unpublished data).

Surra

The *T. evansi* is a widely distributed haemoflagellate of veterinary importance that infects a variety of larger mammals including horses, mules, camels, buffalo, cattle and deer in Africa, Asia and South America (Brun et al., 1998). In Africa, *T. evansi* is present in all countries where camels are present, north of a line extending from Senegal (15° North) to Kenya (equator) (Desquesnes et al., 2013). In Asia, *T. evansi* is epidemic from the Arabian peninsula to all over Asia (Luckins, 1988). An example of enzootic status among camels in Ethiopia was reported to be around 2.3 % (Olani et al., 2016). Furthermore, seroprevalence of surra ranged from 17.4% to 39.8% in water buffaloes in Vietnam (Nguyen et al., 2014). On the other hand, new outbreak of surra was reported in Canary Islands in 1997 and it spread to Spain in 2008 and to metropolitan

France in 2006 (Gutierrez et al., 2010). An outbreak of horse trypanosomosis caused by *T. evansi* was reported in the early 1980s in Kazakhstan and Russia near Mongolian borders (Luckins, 1988). Despite these report, according to Mongolian local veterinary bureau (unpublished data), surra was first diagnosed in Bactrian camel in Khovd province in 1936 (Gungaa, 1990).

1.2.2 Clinical symptoms

Dourine

The *T. equiperdum* is a tissue parasite and is transmitted directly from equine to equine via coitus. *T. equiperdum* is highly host specific and infects only equid such as horse and donkey. Generally, horses are more susceptible than donkeys (Brun et al., 1998). Additionally, it has been very difficult to conduct experimental infection of *T .equiperdum* to laboratory animals, e.g. rabbits, rats and mice (OIE, 2016). The disease development is generally divided into three phases. The first phase is characterized by oedema and swelling of the genitalia, and it begins one to two weeks after infection. The second phase is typical cutaneous plaques ("silver dollar" plaques) appear, with thickening of the skin, this sign is considered pathognomonic for dourine. The third phase of dourine is characterized by progressive anaemia, neurological disorders and paresis of the hindquarters, often ending in death. Mortality rate is high, at around 50 % (Claes et al., 2005; Podaliri Vulpiani et al., 2013). The likelihood of infection with *T. equiperdum* depends on whether the infected host is in an infectious or noninfectious stage. The severity and duration of this disease vary with the virulence of the strain, the nutritional status of the host, and stress factors (Gizaw et al., 2017).

Surra

In contrast to *T. equiperdum*, many blood-sucking insects, especially horseflies (*Tabanus* spp.) and stableflies (*Stomoxys* spp.) can mechanically transmit *T. evansi* from one infected host to another; however, their role as vectors may vary under different conditions (Luckins, 1988). *T. evansi* has the widest host range (It is especially pathogenic in camelids and equids.) amongst salivarian trypanosomes (Desquesnes et al., 2013). Surra is basically a disease of camelids and equines, in which typical clinical expression is described, but various pathogenic effects are observed depending on the various domestic and wild hosts concerned (Desquesnes et al., 2013). Clinical symptoms of *T. evansi* infection are fever, anaemia, loss of appetite and weight, loss of condition and productivity. May develop oedema of the feet, brisket and underbelly,

nervous signs and/or abortion, cachexia, and death, with or without more peculiar signs related to the host species. In camel, the odor of the urine is characteristic, it may be due to ketone bodies, which were found to be elevated in trypanosome infected camels (ADFCA, 2017; Desquesnes et al., 2013).

1.2.3 Disease impact

Dourine

Dourine is a disease of great economic importance (B category of the OIE manual), which is a trade barrier for movement of horses (OIE, 2016). Ten countries in the world have a horse population of more than a million, such as the USA with 9,500,000, China (7,402,450); Mexico (6,260,000); Brazil (5,787,249); Argentina (3,655,000); Columbia (2,533,621); Mongolia (2,029,100); Ethiopia (1,655,383); the Russian Federation (1,319,358); and Kazakhstan (1,163,500) (FAO, 2016). Among these countries only USA is dourine free (Gizaw et al., 2017).

Dourine can deliver devastating effect on equine industry, especially in Mongolia where horse breeding always goes under natural way without any disease control and no movement restriction. In contrast, in dourine eradicated countries, artificial insemination widely used for horse breeding and in case of natural mating, health status of mare and stallion is carefully examined.

In herd level, if a mare is carrier of *T. equiperdum*, the disease transmission from the mare to stallion at coitus, from stallion to other mares will happen. If the infected individual is not stamping out, reproductive potential of the herd will collapse and no longer sustain itself (AHT, 2011).

Thus, reports and evidence clearly show the potential risk of dourine in Asia, Africa, South America, and southeastern European equine industry, especially, free roaming, grazing based production systems in the Mongolian equine industry (Gizaw et al., 2017; Purevsuren, 2004).

Surra

Surra is an important single cause of economic losses, causing morbidity of up to 30.0% and mortality of around 3.0% in camels in Ethiopia (Njiru et al., 2000). In Africa, surra affects mainly camels with acute and chronic infections that cause death. Infection is contracted during the rainy season when there is a peak level of biting insects. Camels constitute the main reservoir of *T. evansi*. In case of the Llanos region, Venezuela, Moreno et al. (2013) assume that, in absence

of treatment, around 4,716 horses would have died by surra, implying a monetary loss equivalent to US$7,486,178.40. Diagnosis and treatment aimed to reduce mortality as much as possible. However, it would have required an investment of US $804,672.08.

1.3 Diagnosis of non-tsetse transmitted animal trypanosomosis

The clinical diagnosis of dourine and surra are not always possible because their clinical signs are not specific for the disease (OIE, 2016). There are of parasitologically, serologically, and DNA based diagnostic methods. Blood smear and serological assays are authorized and gold standard methods for several trypanosomosis in OIE manuals. DNA based test are not yet fully authorized. However, each method has problems, such as low sensitivity and specificity. Neither parasitological, serological nor molecular tests are sensitive and specific enough, thus leading to various kinds of genetic and molecular methods which have been continually updated in order to enhance greater precision in diagnosis of trypanosomosis (Wei et al., 2011).

Due to the extensive similarity between *T. evansi* and *T. equiperdum,* their diagnostic methods are almost the same. *T. evansi* is usually found in the blood during high parasitemia. On the other hand, *T. equiperdum* primarily parasitizes tissues and is rarely found in blood. Therefore, sample preparation for given diagnostic techniques is different. In other words, blood is a primary source of detection for *T. evansi,* and genital organ swab and wash samples are primary source for the detection of *T. equiperdum*.

1.3.1 Parasitological diagnosis

Parasitological examinations are conducted on using blood, cerebrospinal fluid, joint fluid and lymph node fluid. The detection limit of microscopic observation (wet blood film method) is 10^5 trypanosomes/ml, and enhancement techniques such as microhematocrit centrifugation technique and mini anion exchange column centrifugation technique have increased sensitivity of 100-200 trypanosomes/ml (Buscher et al., 2009). If high sensitivity is required, inoculation of samples to laboratory rodents (xenodiagnosis) can be done. This technique can propagate trypanosome from the minimum level of parasitaemia around 20–50 parasites/mL (Desquesnes et al., 2013). In case of *T. equiperdum* infection, the parasites are difficult to find in the infected animal blood, even in oedematous areas (OIE, 2016). Laboratory rodents cannot be utilized, often fail due to the high host specificity of *T. equiperdum*.

1.3.2 Serological diagnosis

Serological diagnosis techniques are suitable for both individual diagnosis and mass-epidemiological surveillance. Almost all serological tests that are specifically developed for *T. evansi* can be applied for detection of *T. equiperdum* due to the high similarity of both trypanosomes (Gari et al., 2010). Generally, *Trypanozoon* subgenus-specific trypanosomal antigen could be applied for antibody detection assays for the diagnosis of *T. equiperdum* infection (Gizaw et al., 2017).

A CATT is which commercially available, can be used under field conditions for detecting antibodies against major surface antigen of *T. evansi*, RoTat1.2 VSG (OIE, 2016). However, it has been demonstrated that the low sensitivity of CATT against RoTat1.2 VSG negative trypanosome infection like as *T. evansi* type B and some strains of *T. equiperdum* (Njiru et al., 2006).

A CFT for dourine is the prescribed test by international animal health code for the testing of equines before their international trade (OIE, 2016). However, the CFT has low sensitivity and sometimes gives inconclusive results. Moreover, experienced technicians are required due to the numerous and cumbersome preparatory steps (Bishop et al., 1995).

An IFAT for *T. evansi* and *T. equiperdum* can also be used as a confirmatory test or to resolve inconclusive results obtained by the CFT (OIE, 2016). However, IFAT is not suitable for large-scale epidemiological studies but only applicable for limited number of samples in laboratory.

Estimates of predictive values indicate that ELISA for detecting IgG is more likely to correctly classify uninfected animals, whereas the CATT is more likely to correctly classify truly infected animals (MedCalc, 2017). ELISA would thus be suitable for verifying the disease-free status of animals prior to movement or during quarantine (OIE, 2016). ELISA and CFT require laboratory equipment, therefore, these methods are inconvenient for field use. Concordance rate between CFT and crude antigen ELISA is 96 % in Mongolian horse sample (Clausen et al., 2003). Also there are some evidence that ELISA are indeed can give better result for the epidemiological and diagnosis usage (Wassall et al., 1991). The cELISA) procedures may actually be closer to gold standards as serodiagnostic tests than the CFT (Katz et al., 2000). Therefore, ELISA diagnosis is a beneficial for daily bases and epidemiological purpose due to its relatively basic labor and equipment requirement and applicability for mass screening. Crude

antigen based ELISA (OIE, 2016), cELISA (Katz et al., 1999), and recombinant antigen based ELISA such as rTeGM6-4r based ELISA (Nguyen et al., 2015a) are already developed.

1.3.3 DNA based molecular diagnosis

Due to the high similarity among species in subgenus *Trypanozoon*, it is still difficult to identify the causative trypanosome species by PCR. In other hand, some of the primer sets can apply for detection both of *T. evansi* and *T. equiperdum*. Since *T. equiperdum* is primary tissue parasite, it is difficult to detect parasite DNA extracted from blood samples. PCR detection of *T. evansi* using TBR primers and internal transcribed spacer (ITS) primers can be useful. Sensitivity of detection depends on DNA extraction methods, sample and disease stages (Masiga et al., 1992; Njiru et al., 2005). All of these primers can be used for detection of *T. equiperdum*. DNA extracted from genital organ swab or wash are optimal sample for the detection of *T. equiperdum*.

PCR based analysis such as restriction fragment length polymorphisms (RFLPs), repetitive DNA probes, mini satellite DNA analysis, amplified fragment length polymorphism (AFLP) and multiplex-endonuclease genotyping (MEGA) has been applied for characterization for *T. equiperdum* and *T. evansi* rather than diagnosis. However, none of these methods solve clearly species-specific identification within the subgenus *Trypanozoon* (Gizaw et al., 2017).

Because perfect performance of diagnostic methods has not yet been established for animal trypanosomosis, it is important to validate and standardize serological tests whether new methods are to be used to correctly identify infected animals. Cross evaluation in different laboratories is therefore required (Desquesnes, 1997). Many of the new techniques in immunology and molecular biology, which have provided a great deal of fundamental information on the tsetse-transmitted trypanosomes, have yet to be applied to *T. evansi* (Luckins, 1988). Thus, good diagnostic methods and counter measures are urgently needed for dourine and surra in horse.

1.4 Next generation sequencing (NGS) and its application for trypanosome study

The NGS is a sequencing technology that can sequence large amount of DNA at a reasonable cost. For almost 30 years, traditional sanger sequencing technology has been used for DNA sequencing with low productivity and high cost. The NGS technologies read the DNA templates randomly along the entire genome. This is accomplished by breaking the entire genome into small pieces, then ligating those small pieces of DNA to designated adapters for random read during DNA synthesis (sequencing-by-synthesis). Therefore, NGS technology is often called massively parallel sequencing (Zhang et al., 2011). The NGS technology has many advantages over sanger sequencing technology in terms of cost-effective and high-throughput. These advantages open a gate for whole genome sequencing of many organisms that has large amount of genomic DNA. In 2015, more than 2.000 genome reports were published, and 2016 was predicted to bring even more (Smith, 2017). The steady reduction in sequencing costs associated with the advent of the new NGS technologies, is one of the major drivers of the so called “genomic revolution” (Chiara and Pavesi, 2017).

Thus far, diversity and taxonomy studies have been based on PCR amplification of molecular markers followed by DNA sequencing. This field is facing a dynamic and tremendous revolution. Over the past decade, the development of generations of sequencing technologies has resulted in an almost exponential increase in throughput and accuracy. Despite being relatively new, current sequencing techniques and associated bioinformatics analyses are highly accurate and reasonably priced. Under the Influence of this development of sequencing technologies, whole-genome sequencing of eukaryotes is becoming a standard methodology. Complete genomic data of reference organisms are the best sources of information for diversity and phylogenetic studies (d'Avila-Levy et al., 2015).

By effectively combining host and pathogen genome-wide transcriptome profiling with interspecies protein–protein interaction screens, we can begin addressing a need for a global approach to dissect effectively the structural and functional genomics and proteomics of intracellular parasite infections (Choi and El-Sayed, 2012).

Recently, genome of a *T. evansi* strain from China was sequenced and compared to the *T. b. brucei* reference strain. The annotated *T. evansi* genome shows extensive similarity to the reference, with 94.9% of the predicted *T. b. brucei* coding sequences (CDS) having an ortholog in *T. evansi*, and 94.6% of the non-repetitive orthologs having a nucleotide identity of 95% or greater. Interestingly, several procyclin-associated genes (PAGs) were disrupted or not found in

this *T. evansi* strain, suggesting a selective loss of function in the absence of the insect life-cycle stage (Carnes et al., 2015).

The impact of NGS technology is indeed egalitarian in that it allows both small and large research groups the possibility to provide answers and solutions to many different problems and questions in the fields of genetics and biology, including those in medicine, agriculture, forensic science, virology, microbiology, and marine and plant biology (Kulski, 2016).

1.5 Current situation and problems of non-tsetse transmitted animal trypanosome in Mongolia

In Mongolia, animal husbandry has not only been a tradition but also a main source of nutrition and raw materials for centuries. The horse is important in the present-day lives of many Mongolians. Horses are used not only for transportation but also for the herding of other livestock, entertainment and for the myriad goods they yield (Yazdzik, 2011). The livestock sector contributes 90% of the total agricultural production, which accounts for 11% of Mongolia's total GDP, and horses comprise 5.9% of the total livestock of the country (NRSO, 2015). Annually, horse meat production is worth approximately 48 million USD, which represents 4.1% of the livestock production in 2013 (FAO, 2016). The Mongolian agricultural sector has been severely affected by various infectious diseases, and the impact of these diseases on the national economy is currently a pertinent issue (Altangerel et al., 2012). Non-tsetse transmitted animal trypanosome and trypanosomosis have been reported in Mongolia. An outbreak of non-tsetse-transmitted horse trypanosomosis caused by *T. evansi* was reported in the early 1980s in Kazakhstan and Russia near Mongolian borders (Luckins, 1988). In contrast, dourine and *T. equiperdum* were not reported in Mongolian horses until a survey carried out using a lyophilized *T. equiperdum* crude antigen-based ELISA. The study estimated that the sero-prevalence of *T. equiperdum* in Mongolian horses is 5.5% (Clausen et al., 2003). Recently, a new *T. equiperdum* strain was isolated from the urethral mucosa of a horse in Mongolia that was clinically and parasitologically defined as a dourine case (Suganuma et al., 2016). Moreover, suggested that there is high prevalence values of non-tsetse transmitted animal trypanosomosis among domestic animals in Hovd and Bayan-Ölgii provinces. During recent years, however, cases of non-tsetse-transmitted horse trypanosomosis in Mongolia have tended to increase due to uncontrolled importation and cross-breeding of horses. Thus, the impact of non-tsetse transmitted horse trypanosomosis is an issue that cannot be ignored. To overcome the endemic status of non-tsetse transmitted animal trypanosomosis among Mongolian livestocks, the

establishment of rapid, effective, and field-friendly diagnostic tools are urgently needed because of large country side and inadequacy of local veterinary facility in Mongolia.

Since, the recombinant antigen based (rTeGM6-4r based) ICT and ELISA are already developed and tested for *T. evansi*-infected water buffalo, cattle, goats and sheep (Nguyen et al., 2015a; Nguyen et al., 2014) but not evaluated for *T. equiperdum* infected horse. Field friendly diagnostic tool such as ICT will be highly effective for rural areas that are far from high level laboratory and without special equipment. ICT can diagnose the disease without any tool or skilled veterinarian. Therefore, this easy to use feature has advantage to reach every remote corner of big country like Mongolia. The ICT for non-tsetse transmitted animal trypanosomosis can help veterinarians, owners and farmer to diagnose the disease and decide the proper treatment of the disease.

Until now, there are no characteristic diagnosis markers for distinguishing between dourine and surra due to the lack of high quality genome information of *T. equiperdum*. The species definition markers might be found by the whole genome comparison between *T. evansi* and *T. equiperdum* using NGS. Despite the development of genetic and molecular techniques by different scholars to clear species-specific identification within the subgenus trypanosome remains difficult. So far, the discovery of a simple and reliable way to entirely distinguish all species and subspecies of *Trypanozoon* remains a big challenge (Gizaw et al., 2017).

1.6 Objectives of the study

The aims of the study are the establishment and evaluation of diagnostic tools for non-tsetse transmitted horse trypanosomosis and the whole genome analysis of newly isolated *T. equiperdum* from Mongolia for the discovery of novel diagnostic markers in future studies.

Chapter 1: The evaluation of GM6-based ELISA and ICT as diagnostic methods on a Mongolian farm with an outbreak of non-tsetse transmitted horse trypanosomosis

2.1 Introduction

Protozoan parasites are a common causative factor of equine diseases throughout the world. Horses can be infected by *T. b. brucei*, *T. vivax* and *T. congolense* in Africa, *T. cruzi* and *T. vivax* in South America and *T. evansi* and *T. equiperdum* worldwide (Nimpaye et al., 2011). Mongolia is listed as the epidemic countries of dourine and surra (non-tsetse transmitted trypanosomosis) by OIE. In addition, some of the previous reports also revealed the epidemic situations of dourine and surra, and the potential risk of these diseases for equine industry in Mongolia (Clausen et al., 2003; Luckins, 1988; Musinguzi, 2017; Suganuma et al., 2016). However, there are no effective countermeasures formulated for non-tsetse transmitted horse trypanosomosis in Mongolia due to the lack of effective diagnostic tools for field on-site use. To overcome the risk of non-tsetse transmitted horse trypanosomosis, the effective, simple and field-friendly diagnostic tools must be developed.

Some protozoan antigens are composed of repeated amino acid sequences that display immunological dominance. The GM6 antigen, which is located within the flagellum on the microtubular fibers of the membrane skeleton was identified by its high ability of immunoreaction with the early phase *T. evansi* infection in bovine (Müller et al., 1992). GM6 is highly conserved among the trypanosomes (Thuy et al., 2012). The recombinant GM6 antigen derived from *T. evansi*, which consists of four repeats of amino acids unit, was named rTeGM6-4r. This recombinant protein has already shown good diagnostic value in *T. evansi*-infected water buffalo, cattle, goats and sheep (Nguyen et al., 2015a; Nguyen et al., 2014). I therefore hypothesize that rTeGM6-4r may have good diagnostic value for non-tsetse transmitted horse trypanosomosis.

The objective of this chapter was to evaluate the diagnostic potential of rTeGM6-4r based diagnostic methods on a Mongolian horse farm with an outbreak of non-tsetse transmitted horse trypanosomosis.

2.2 Materials and methods

2.2.1 Study area

This study was conducted in the middle of March 2016 in Ulziit Khoroo, Ulaanbaatar city, which is located in the central part of Mongolia. A horse owner requested the diagnosis of a stallion that showed edema of unknown cause on its genital organ. The herd included approximately 50 horses.

2.2.2 Sample information

In the present study, 50 blood samples were collected from a herd of horses that included 9 males and 41 females (Table 1). Permission was obtained according to the standards of animal experimentation in Obihiro University of Agriculture and Veterinary Medicine (Approval No. 28-45). Approximately 5 mL of blood was drawn from the jugular vein into vacutainer tubes; one with EDTA-2Na for DNA extraction, and the other for serum separation. In addition to the blood samples, genital organ swabs were collected from three selected horses with clinical signs of dourine. The total DNA of each whole blood and genital organ swab sample was purified by phenol-chloroform isoamilalchohol methods, as described by Sambrook and Russell (2001). The serum was separated from the clotted blood. All samples were stored at -30ºC until use.

2.2.3 Cloning and sequence analysis of GM6 from *T. equiperdum*

Total DNA was extracted from a culture-adapted *T. equiperdum* IVM-t1 strain (Suganuma et al., 2016). The GM6 gene of *T. equiperdum* was amplified by a PCR with GM6 primer sets (Table 2). The amplicon was cloned into pCR™2.1-TOPO® vector, and the nucleotide sequence of *T. equiperdum* GM6 was determined using an ABI Prism 3100 Genetic Analyzer (Thermo Fisher Scientific, MA, USA). The amino acid sequence of *T. equiperdum* GM6 was compared to the GM6 of *T. evansi* (Thuy et al., 2012), *T. b. brucei* (accession number: XP_828202.1) and *T. b. gambiense* (accession number: Q26755.1) by ClustalW multiple alignment (Bioedit 7.2.5).

2.2.4 Indirect immunofluorescence antibody test (IFAT) for *T. equiperdum*

Smears prepared from swabs of the genital organ from horses with suspected dourine infection were fixed with methanol, followed by blocking with 5% skim milk in TBS with 0.05%

Tween 20 (TBS-T) at room temperature for 1 hour. The smears were washed 3 times with TBS-T and were then incubated with mice anti-TeGM6-4r antibodies. The smears were subsequently incubated with anti-mouse IgG-DyeLight 488 (Thermo Fisher Scientific) and Hoechest 33348 (Dojindo, Inc., Kumamoto, Japan) for kinetoplast staining and nucleus staining, respectively. The prepared IFAT slides were examined using a confocal laser scanning microscope (TCS-NT, Leica Microsystems GmbH, Wetzlar, Germany).

2.2.5 Positive and negative control sera

One positive and 20 negative serum samples were obtained from Mongolian horses. The positive serum sample was collected from a horse that was definitively diagnosed as dourine-positive based on the microscopic observation of active movement of *T. equiperdum* that were obtained from a genital organ swab (Suganuma et al., 2016). Negative sera were selected from Mongolian horses by a KIN-PCR using total DNA extracted from whole blood samples and genital organ swabs. The status of surra of each animal was also confirmed by CATT/*T. evansi* (Institute of Tropical Medicine, Antwerp, Belgium), according to the OIE manual for the diagnosis of surra (OIE, 2016). After subtracting the blank well OD value, the standard deviation and average values of the OD values of the negative samples were calculated. The cut-off value for ELISA was determined based on the summation of the mean value of negative samples plus 3 times the standard deviation of the OD values of the negative samples.

2.2.6 PCR

All of the DNA samples that were extracted from the whole blood (n=50) and swabs of the genital organs of selected horses (IDs 14, 15 and 20) showing the characteristic symptoms of dourine were analyzed by an internal transcribed spacer 1 (ITS1) PCR (Njiru et al., 2005) (Table 2). The PCR cycles were as follows: initial denaturation at 94°C for 5 min, followed by 35 cycles of 94°C for 40 s, 58°C for 40 s, 72°C for 90 s, and final extension at 72°C for 5 min.

2.2.7 Crude trypanosome antigen and rTeGM6-4r-based ELISA

The rTeGM6-4r-based ELISA and the trypanosome cell lysate crude ELISA plates were prepared in accordance with a previous study and the OIE manual, respectively (Nguyen et al., 2014; OIE, 2016) with minor modifications. In brief, Maxisorp 96-well plates (Nalgene-Nunc, NY,

USA) were coated with 1 µg/well of the crude antigen or 200 ng/well of rTeGM6-4r diluted in a coating buffer (50 mM carbonate-bicarbonate buffer [pH 9.6]) for 4 h. Then the coated plates were blocked overnight with 3% skim milk in PBS-T. After washing with phosphate-buffered saline containing 0.05% Tween 20 (PBS-T), 200 times diluted sera in dilution buffer (PBS-T containing 3% skim milk) was added to each well and incubated at room temperature for 2 h. The plates were then washed with PBS-T before adding 5,000 times diluted anti-horse IgG rabbit antibodies conjugated with horseradish peroxidase (Thermo Fisher Scientific). After 2 h of incubation, the plates were washed with PBS-T. Then, 2'-azino-bis 3-ethylbenzothiazoline-6-sulphonic acid (ABTS) solution and 0.0075% hydrogen peroxide were applied, and the plates were incubated for 30 min at 37ºC. After stopping the reaction with 0.05 M citric, phosphoric acid, the absorbance of each well at 600 nm was read with an MTP-500 microplate reader (Corona Electric, Ibaraki, Japan).

2.2.8 rTeGM6-4r-based immunochromatographic test (ICT)

The rTeGM6-4r-based ICT was prepared in accordance with the methods of a previous study (Nguyen et al., 2015b). To perform the ICT, 10 µL of a serum sample was diluted five times with PBS and loaded on the sample pad. An ICT was determined to be positive if it displayed both the test and control lines and was considered to be negative if it only displayed the control line.

2.2.8.1 Production of anti-rTeGM6-4r polyclonal antibody

BALB/c mice were immunized intraperitoneally with 100 µl emulsion of 1 mg/ml rTeGM6-4r with Titer Max Gold adjuvant (Sigma-Aldrich, St. Louis, USA) in 1:1 ratio and three times at 2-week interval. Antibody titer was determined using TeGM6-4r-based ELISA. Mice producing high antibody titer were sacrificed for blood collection and serum extraction. Sera were then subjected to polyclonal antibody purification using aMAbTrap kit (Sigma-Aldrich, St. Louis, USA) in accordance with the manufacturer's instructions. The concentration of the polyclonal antibody was detected by bicinchoninic acid assay. Anti-sera were stored at −80°C until use.

2.2.8.2 Construction of TeGM6-4r-based ICT

Three-hundred microgram per milliliter of rTeGM6-4r was bound into 10 times volume of 20 nm Gold colloid at pH 7.3, followed by blocking with 0.05 % polyethylene glycol (PEG) 20,000

and 1 % bovine serum albumin. The complex gold colloid labeled rTeGM6-4r was washed twice with blocking solution containing 0.05 % PEG 20,000 and 0.5 % bovine serum albumin, and the concentration was adjusted to OD520 1.5 with dilution buffer (10 mM Tris-HCl pH 8.2 and 5% sucrose). This diluted conjugate was applied into a glass fiber membrane and dried for 5-6 hours at room temperature.

rTeGM6-4r (0.5 mg/ml for test line) and anti-rTeGM6-4r polyclonal antibody (1 mg/ml for control line) were dispensed on 10 mm, 15 mm from edge of the nitrocellulose membrane using an XYZ Dispensing Platform (BioDot, Irvine, USA). The membrane was then blocked with 0.5% casein and dried overnight at room temperature. Subsequently, all of the ICT components were assembled manually and cut into 3-mm strips by CM 4000 (BioDot, Irvine, USA). For running the ICT, 10 µl of serum sample was diluted five times with PBS and loaded on the sample pad. The result was shown in 10 – 20 min. An ICT was determined to be positive if it displayed both test and control lines and negative if it only displayed the control line.

2.2.9 Statistical analysis

The degree of agreement between diagnostic methods was quantified by kappa statistics and 95% confidence intervals using the Microsoft Excel software program (Microsoft, Redmond, WA, USA) according to the methods of Jacob Cohen (Cohen, 1960) and Clopper and Pearson (Clopper and Pearson, 1934), respectively. The distribution of OD values for ELISA data was visualized using the GraphPad Prism 6.0 software program (GraphPad Software, Inc., CA, USA).

2.3 Results

2.3.1 Characterization *of T. equiperdum* GM6 as a candidate serodiagnostic antigen

The amino acid sequences of GM6 from different *Trypanozoon* parasites were aligned to examine its homology (Fig. 1). The amino acid sequences of the GM6 genes from *T. equiperdum, T. evansi, T. brucei* and *T. b. gambiense* showed 97–98% homology. Among the 68 amino acid repeat units within *Trypanozoon*, there were only two different amino acids. This indicates that the sequence of tandem repeat units in GM6 is highly conserved among the *Trypanozoon* subgenus. In addition, the cellular localization *of T. equiperdum* GM6 was analyzed by IFAT using anti-*T. evansi* GM6 antibodies. *T. equiperdum* GM6 was localized along the flagellum of the parasite, which was identical to the previously reported cellular localization of GM6 in the *Trypanozoon* subgenus (Hayes et al., 2014; Müller et al., 1992) (Fig. 2). Moreover, anti-*T. evansi* GM6 antibodies showed a strong cross-reaction with *T. equiperdum* parasites.

2.3.2 Field trial of rTeGM6-4r-based serodiagnostic methods for horses

Samples from 50 horses collected from one farm, on which an epidemic of trypanosomoses was clinically suspected, were screened. The sample and result summary was shown in Table 1. Of the 50 samples, 46%, 42% and 28% were found to be positive using the rTeGM6-4r-based ELISA (Fig. 3-A), crude antigen ELISA (Fig. 3-B) and rTeGM6-4r-based ICT, respectively. The diagnostic value of the rTeGM6-4r-based ELISA and the rTeGM6-4r-based ICT was measured in comparison to the crude antigen ELISA. The sensitivity and specificity of the rTeGM6-4r-based ELISA were 81% and 79%, respectively (Table 3). The Cohen’s kappa value between rTeGM6-4r-based ELISA and crude antigen-based ELISA was 0.60, and the strength of agreement between the rTeGM6-4r-based ELISA and crude antigen-based ELISA was considered 'moderate' (Table 3). The kappa value between the rTeGM6-4r-based ICT and crude antigen-based ELISA was 0.53. a mare (ID34) was seropositive by ICT although it was negative crude antigen based and rTeGM6-4r based ELISA. Additionally, a 2-year-old horse (ID 2) was seropositive by all three tests, even though this horse had not yet mated with a stallion. The horse (ID2) was born to a mare (ID 31) and stallion (ID29) that were found to be positive by all three tests (Table 1).

2.3.3 Results of the ITS1 PCR-based diagnosis of horses

Genital organ swabs from three selected horses (IDs 14, 15 and 20) that were found to be strongly positive by both the ELISA and ICT (Table 1), were clearly positive on the PCR (Fig. 4). Moreover, the active movement of trypanosomes was observed from a genital swab sample from a horse (ID14) by microscopy. In contrast, all of the whole blood samples were found to be negative by both Giemsa staining and a PCR (data not shown).

2.4 Discussion

T. equiperdum and *T. evansi* are cosmopolitan trypanosomes. The former trypanosome causes dourine via sexual transmission in Equidae, whereas the latter causes surra via mechanical transmission through bloodsucking insects such as *Tabanus* spp. Nowadays, Mongolian herdsmen and veterinarians have a strong need for diagnostic and mass screenings that can be applied on-site due to the relatively high prevalence of these diseases (Clausen et al., 2003; Desquesnes et al., 2013).

In this study, the *T. evansi* GM6 antigen was evaluated with two diagnostic tools: an ICT and an ELISA for non-tsetse-transmitted horse trypanosomosis, especially dourine. In all previous studies, crude antigen ELISA for trypanosomes exhibited high sensitivity and specificity, generally >90%. However, there is no gold standard for serological tests in trypanosomes; rather, the ELISA is a recommended serological test in trypanosomes (Kocher et al., 2015). Because the measured sensitivity and specificity are calculated using crude antigen ELISA as a reference, they are relative to crude antigen based ELISA rather than true values.

T. equiperdum GM6 cross-reacted with anti-*T. evansi* GM6 antibody, with localization detected along the flagella (Fig. 2). The GM6 amino acid sequence was observed to have high similarity among *Trypanozoon,* therefore anti-rTeGM6-4r antibody can recognize all three species (Fig. 1). These results suggested that in addition to surra in Equidae, these rTeGM6-based serodiagnostic methods would be useful for epidemiological studies involving dourine and for its diagnosis. Similarly, Pillay et al. (2013) reported that the *T. vivax* GM6 is an excellent candidate antigen for the development of a point-of-care test for the diagnosis of animal African trypanosomoses in cattle that are caused by *T. vivax* and *T. congolense*. A previous report noted that the considerable sensitivity of rTeGM6-4r would make it a useful antigen for the diagnosis of surra in future surveillance programs (Nguyen et al., 2014). A similar study of TcoCB1 and TvGm6 antigens, which are homologous proteins of TeGM6, showed high levels of sensitivity and specificity of their rapid tests (Boulange et al., 2017).

In the current study involving horses, the sensitivity and specificity of rTeGM6-4r-based ELISA were moderate at 81% and 79%, respectively (Table 3). This higher sensitivity of the ELISA in comparison to the ICT could be due to the enzyme-substrate reaction, which enhances the detection process when there is a low antibody titer. While, the rTeGM6-4r based ICT was less sensitive than the ELISA. However, ICT was relatively specific, simple and rapid (Nguyen et al., 2015a). The higher specificity of the rTeGM6-4r-based ICT in this study might be explained

by lower cross-reactivity among the tested samples. The slightly higher OD values in the negative control wells in the crude antigen-coated plates suggests the possibility of cross-reaction with a non-specific antibody in the sera of the horses (Fig. 3-B). It is not surprising that cross-reactions occur with other trypanosome species and even other parasites in most instances in which crude lysates of the antigen of a given trypanosome strain are cultured *in vivo* or *in vitro* (Magez and Radwanska, 2014). Among the samples that showed a positive reaction in the rTeGM6-4r-based ICT, only one sample (ID34) was found to be negative by the ELISAs. This sample might represent a newly infected horse in which the infection could not be detected using GM6 and the crude antigen ELISA because of the low IgG titer. However, ICT can detect both IgG and IgM, which is the dominant immunoglobulin at the early stage of dourine infection (Nguyen et al., 2015b).

The DNA extracted from the genital organs from the three selected horses with clinical symptoms were found to be positive by a PCR, whereas none of the DNA samples from whole blood were found to be positive (Fig. 4). These PCR results strongly suggested that these horses had dourine. This is also supported by the fact that *T. equiperdum* is primarily a tissue parasite; thus, although *T. equiperdum* can often be found in smears taken from the genitalia or plaques, it is not usually present in the circulating blood (Gunn and Pitt, 2012).

Based on the observation of the active movement of *T. equiperdum* by microscopy and the identification of its DNA in a PCR using genital swab samples, we concluded that the horses were truly infected with dourine. This conclusion is strongly supported by the study of Suganuma et al. (2016). The PCR results in our ongoing surveillance project in Mongolia have also revealed surra epidemics in other domestic animals (data not shown). Suganuma et al. (2016) recently isolated a new *T. equiperdum* strain from a dourine-infected horse in Mongolia. The recorded prevalence (40%) of the non-tsetse-transmitted horse trypanosomosis in the herd indicated that the disease has been spreading for a long time. However, the majority of the horses showed no clinical signs.

During the last decades, importation of horses mainly from foreign countries such as Russia and other European countries to Mongolia has tended to result in crossbreeding. Proper veterinary checkups are often not done on imported horses in animal quarantine. Moreover, stallions without veterinary checkup frequently migrate with their owners within the country and spread dourine in their harems in Mongolia (Hund, 2008). These two domestic and international horse migration events might play a key role in the spread of disease in Mongolia. The main risk

factor is the herders' inadequate knowledge in relation to the diagnosis and treatment of dourine. In this emerging market economy, inadequate incentives exist for herders to invest in disease control and animal health (Goodland et al., 2009).

The possible infection in a young horse (ID 2), which had not been involved in mating, suggests the possibility of vertical transmission of the disease because the young horse was sired by an infected stallion (ID29) and dammed by an infected mare (ID31), and the serodiagnosis of both animals was clearly positive (Table 1). This evidence agreed with a report that suggested that the transmission can be vertical, horizontal, iatrogenic and peroral, each of which has varying degrees of epidemiological significance, depending on the season, location and host species (Desquesnes et al., 2013).

The results of the present study revealed that the GM6 gene of *T. evansi* and *T. equiperdum* are almost identical in terms of the tandem repeat units. This finding allows this antigen to be used for diagnosing dourine. In the present study, rTeGM6-4r showed good diagnostic value in testing the sera of *T. equiperdum*-infected horses. Similar results were found in domestic animals infected with *T. evansi* (Nguyen et al., 2015a; Nguyen et al., 2014). In conclusion, the rTeGM6-4r-based ELISA and ICT may offer alternative diagnostic methods for large-scale epidemiological studies and the on-site diagnosis of non-tsetse-transmitted horse trypanosomoses in the field.

Table 1. Sample information and data summary of this study

ID	Sex[a]	Age	Breed	Crude antigen ELISA	rGM6-4r based ELISA	rGM6-4r based ICT	Blood DNA PCR	Smear DNA PCR[b]	Microscopic examination[b]	Clinical symptoms[b]
1	F	2	Mixed	-	-	-	-	n/a	n/a	ND
2	F	2	Mixed	+	+	+	-	n/a	n/a	ND
3	F	2	Mixed	-	-	-	-	n/a	n/a	ND
4	F	2	Mixed	+	-	-	-	n/a	n/a	ND
5	F	2	Mixed	-	-	-	-	n/a	n/a	ND
6	F	2	Mixed	-	+	-	-	n/a	n/a	ND
7	F	2	Mixed	+	+	-	-	n/a	n/a	ND
8	F	2	Mixed	-	-	-	-	n/a	n/a	ND
9	F	2	Mixed	-	-	-	-	n/a	n/a	ND
10	F	2	Mixed	-	-	-	-	n/a	n/a	ND
11	F	2	Mixed	+	+	-	-	n/a	n/a	ND
12	F	2	Mixed	-	-	-	-	n/a	n/a	ND
13	F	2	Mixed	+	-	-	-	n/a	n/a	ND
14	M	4	Mixed	+	+	+	-	+	+	genital organ Smegma
15	M	3	Arab	+	+	+	-	+	-	genital organ Smegma
16	M	7	Mixed	+	+	-	-	n/a	n/a	ND
17	M	4	Mixed	-	+	+	-	n/a	n/a	ND
18	M	6	Mixed	-	-	-	-	n/a	n/a	ND
19	F	22	Mixed	+	+	-	-	n/a	n/a	ND
20	F	22	Mixed	+	+	+	-	+	-	weight loss
21	F	4	Mixed	+	-	-	-	n/a	n/a	ND
22	F	3	Mixed	+	+	+	-	n/a	n/a	ND
23	F	6	Mixed	+	+	+	-	n/a	n/a	ND
24	M	17	Arab	-	+	-	-	n/a	n/a	ND
25	M	6	Mixed	-	+	-	-	n/a	n/a	ND
26	F	6	Mixed	+	+	+	-	n/a	n/a	ND
27	M	15	Arab	-	-	-	-	n/a	n/a	ND
28	F	6	Mixed	-	-	-	-	n/a	n/a	ND
29	M	7	Arab	+	+	+	-	n/a	n/a	ND
30	F	2	Mixed	-	-	-	-	n/a	n/a	ND
31	F	8	Mixed	+	+	+	-	n/a	n/a	ND
32	F	8	Mixed	-	-	-	-	n/a	n/a	ND
33	F	10	Mixed	-	-	-	-	n/a	n/a	ND
34	F	6	Mixed	-	-	+	-	n/a	n/a	ND

35	F	6	Mixed	-	-	-	-	n/a	n/a	ND
36	F	3	Mixed	-	-	-	-	n/a	n/a	ND
37	F	6	Mixed	-	-	-	-	n/a	n/a	ND
38	F	9	Mixed	-	-	-	-	n/a	n/a	ND
39	F	7	Mixed	-	-	-	-	n/a	n/a	ND
40	F	8	Mixed	-	-	-	-	n/a	n/a	ND
41	F	7	Mixed	-	-	-	-	n/a	n/a	ND
42	F	9	Mixed	+	+	+	-	n/a	n/a	ND
43	F	9	Mixed	-	-	-	-	n/a	n/a	ND
44	F	7	Mixed	-	+	-	-	n/a	n/a	ND
45	F	6	Mixed	+	+	-	-	n/a	n/a	ND
46	F	4	Mixed	+	-	-	-	n/a	n/a	ND
47	F	3	Mixed	-	+	-	-	n/a	n/a	ND
48	F	6	Mixed	+	+	+	-	n/a	n/a	ND
49	F	6	Mixed	+	+	+	-	n/a	n/a	ND
50	F	4	Mixed	-	-	-	-	n/a	n/a	ND
	F41, M9		Arab 5, Mixed breed 45	21 (42%)[c]	23 (46%)[c]	14 (28%)[c]	0 (0%)	3	1	3

[a] F: Female and M: Male; [b] PCR and microscopic observation of genital organ swab were performed in only three horses (ID 14, 15 and 20) which showed some clinical symptoms; [c] Positive sample No and percentage.

Table 2. The PCR primer list for the chapter

Method	Sequence 5'-3'	Target locus	Size (bp)	Reference
ITS1 PCR	CCGGAAGTTCACCGATATTG TTGCTGCGTTCTTCAACGAA	ITS1 region	540	Njiru *et al.*
GM6 PCR	GGATCCATGGAGCTTGCTAAA GAATTCCTAATGTGAATGCTC	GM6 TR units	Variable	Thuy *et al.*

Table 3. Comparison of the results of the different serodiagnostic tests

		Crude antigen ELISA[a]			Sensitivity	95% CI[c]	Specificity	95% CI[c]	Agreement[d]	Kappa[b]
		Positive	Negative	Total						
rTeGM 6-4r ELISA	Positive	17	6	50	81%	58%-95%	79%	60%-92%	80%	0.60
	Negative	4	23							
rTeGM 6-4r ICT	Positive	12	2	50	57%	32%-78%	93%	77%-99%	78%	0.53
	Negative	9	27							

[a] Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2016

[b] Cohen's Kappa value

[c] Confidence intervals at 95% for sensitivity and specificity

[d] Number of observed agreement

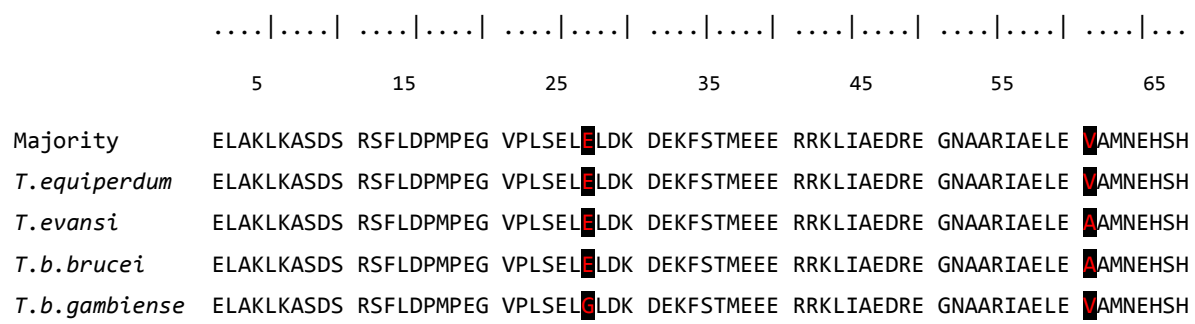
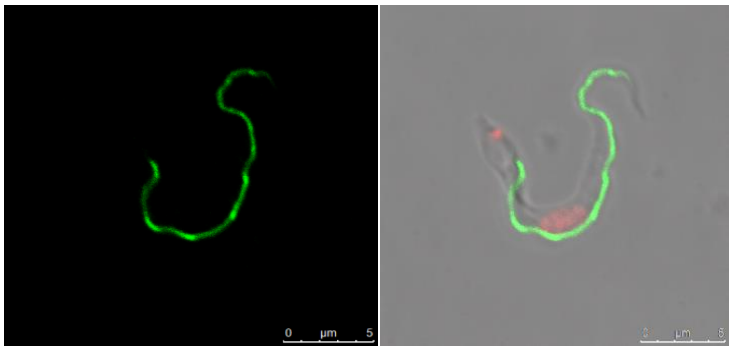
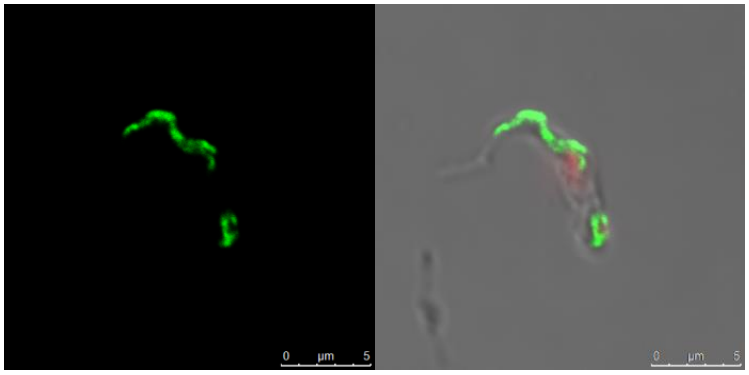


Figure 1. The comparison of the amino acids in the tandem repeat unit of GM6 among the Trypanozoon subgenus.

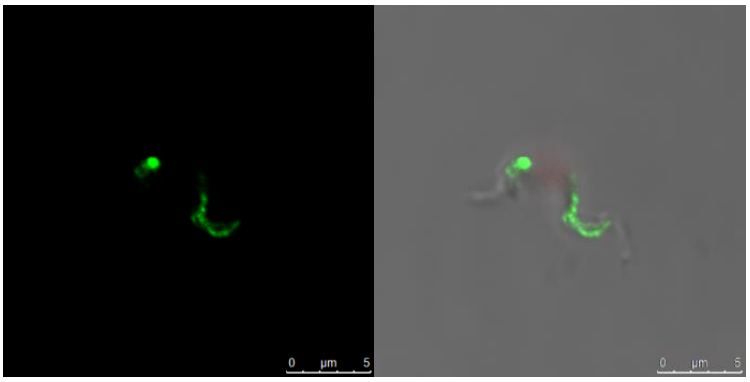
Ninety-seven percent amino acid sequence homology was observed between the GM6 genes from *T. equiperdium* IVM-t1 strain, *T. evansi* Tansui strain, *T. b. brucei* (accession no. XP_828202.1) and *T. b. gambiense* (accession no. Q26755.1). A single unit of tandem repeat consists of 68 amino acids. The red colored amino acids indicate differences.



T. equiperdum



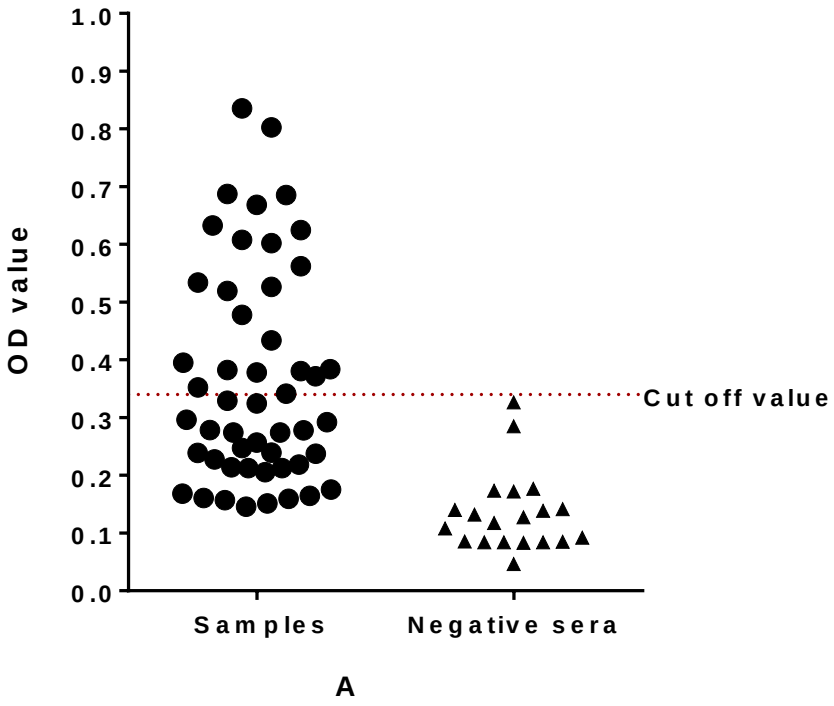
T. brucei



T. evansi

Figure 2. The cellular localization of GM6 antigen.

T. equiperdum, *T. evansi* and *T. brucei* GM6 was cross-reacted with anti-*T. evansi* GM6 antibody. *T. equiperdum* GM6 was localized along the flagella. Left panel: GM6 antigen localization (Green signal). Right panel: A merged picture, Red: A kinetoplast and the nucleus.



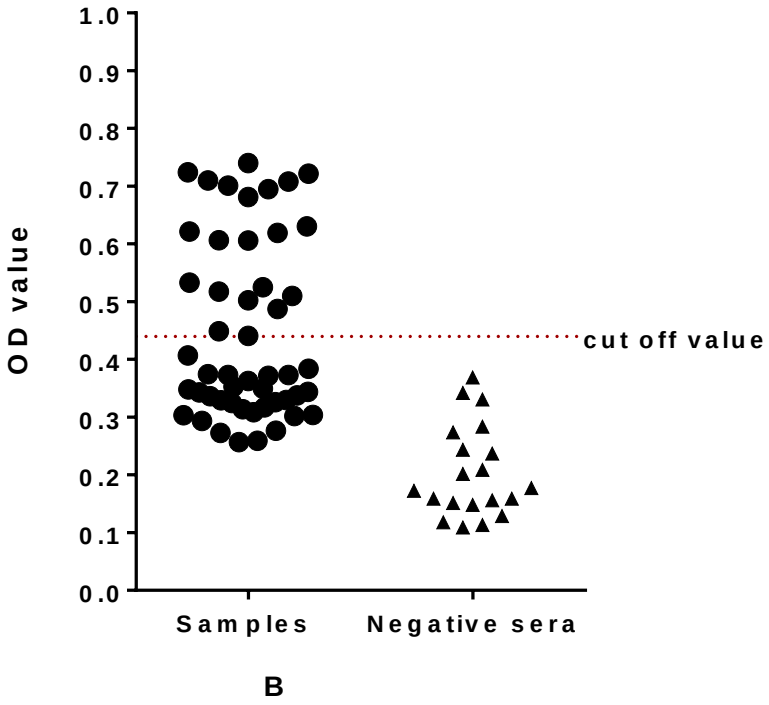


Figure 3. Optical density distribution of two ELISA methods

Optical density distribution of farm samples (N=50) and negative controls (N=20) by rTeGM6-4r based ELISA (A) and crude antigen ELISA (B). Cut off values of rTeGM6-4r-based ELISA and crude antigen ELISA were 0.34 and 0.44, respectively.

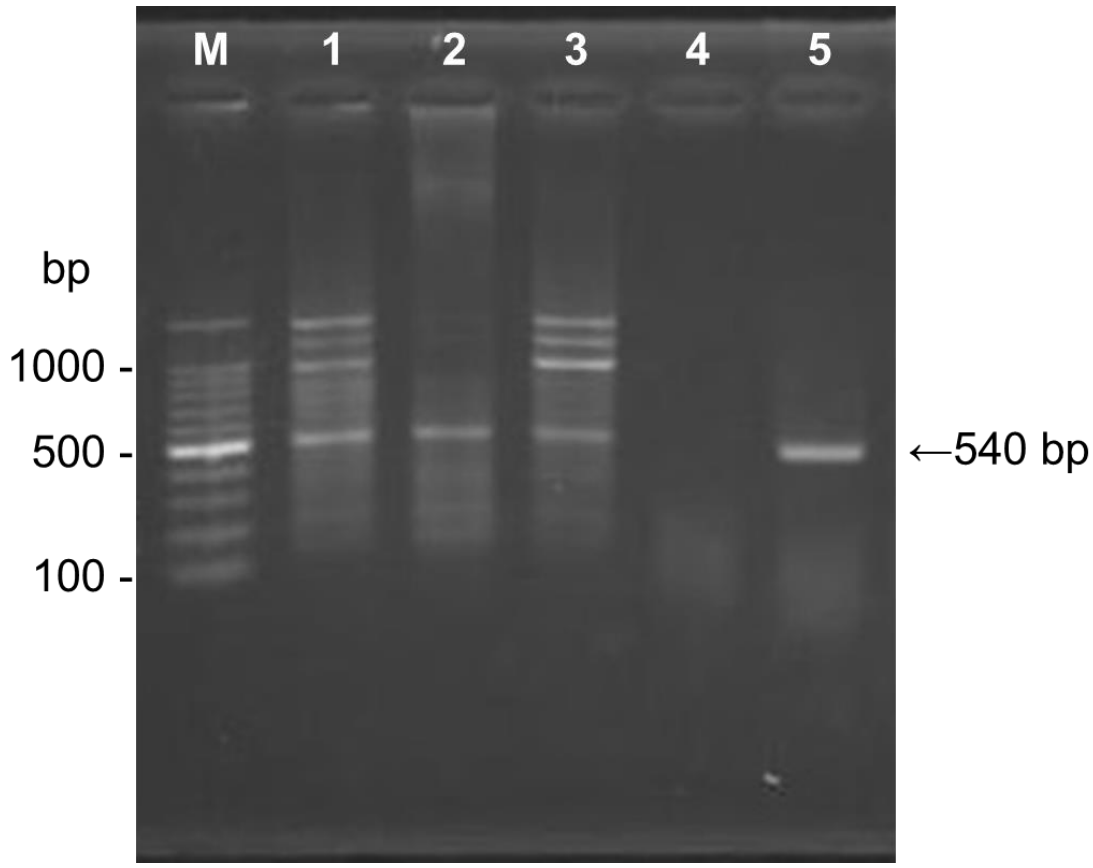


Figure 4. The ITS1-PCR results from a genital organ swab.

The PCR results for selected swab samples. M, A 100 bp DNA marker; 1, Sample ID 14; 2, Sample ID 15; 3, Sample ID 20; 4, A negative control (a swab sample from a healthy horse); 5, A positive control (Total DNA of *T. equiperdum* IVM-t1). Dourine infection was strongly indicated in all three samples based on positive results in three of the serodiagnostic tests and the clinical signs.

Chapter 2: Whole genome and transcriptome assembly of *T. equiperdum* IVM-t1 strain

3.1 Introduction

Trypanosoma equiperdum causes dourine in equines in some parts of Asia, Africa, Europe and South America (Brun et al., 1998). *T. equiperdum* belongs to subgenus *Trypanozoon* with *T. brucei* sspp. and *T. evansi*. *T. brucei* sspp. and *T. evansi* which initially parasitize in the bloodstream of host animal during blood meals of tsetse fly or other blood sucking insects. Whilst the *T. equiperdum* initially parasitizes in genital organs of infected equids, and is transmitted through coitus. Despite these unique features, *T. equiperdum* shares high degree of morphological and genetic similarity with other *Trypanozoon* and cannot be clearly distinguished from each of them morphologically and molecularly (Table 4.) (Desquesnes et al., 2013).

Many *T. equiperdum* strains were isolated in the past 50 years (Claes et al., 2005). Recently, some *T. equiperdum* strains were newly isolated from Italia, Ethiopia (Dodola strain) and Venezuela (TeAp-N/D1, TeGu-ND1 strains) (Hagos et al., 2010b; Pascucci et al., 2013; Sanchez et al., 2015). Among the *Trypanozoon* group, whole genome sequences of *T. brucei* TREU927 strain, *T. evansi* STIB805 strain and *T. equiperdum* OVI strain were published in 2005 (Berriman et al., 2005), 2015 (Carnes et al., 2015), 2017 (Cuypers et al., 2017; Hebert et al., 2017), respectively. It has been hoped that species-specific PCR for diagnosis of *T. equiperdum* can be developed by comparative genomic analyses using high quality of whole genome data. However, the whole genome sequence of *T. equiperdum* OVI strain was not adequate for the further analyses due to the low quality of sequence data; small number of the N50, and too much number of contigs and many numbers of gaps *etcetera*. Moreover, *T. equiperdum* OVI strain and other previously isolated *T. equiperdum* strains might have been misidentified *T. evansi* or *T. brucei* because all of them are not isolated from genital organ but isolated from blood. Therefore, high quality whole genome sequence using true *T. equiperdum* strains is strongly required for the further *T. equiperdum* characterization and dourine research.

Recently, a new *T. equiperdum* IVM-t1 strain was isolated from the urethral tract of a dourine-diagnosed stallion in Mongolia and the strain was successfully cultured *in vitro* (Suganuma et al., 2016). The molecular characterization of this *T. equiperdum* strain was conducted based on the 18S rRNA, ITS and maxicircle gene sequences. Thus, new *T. equiperdum* strain, which is directly isolated from the infectious lesions of the horse, is more suitable for whole genome analysis.

Bioinformatics is a new science which is the combination of biology, mathematics, computer science and engineering (Spengler, 2000). Since Next Generation Sequencing (NGS) was introduced in the year 2000 era, the massive data produced by NGS also presents a significant challenge for data storage, analyses, and management solutions. Advanced bioinformatics tools are essential for the successful application of NGS technology (Zhang et al., 2011). There is a long way from massive raw nucleotide sequence data to well processed whole genome data. At first, raw sequence data are obtained from genomic DNA using several types of NGS equipments. For example, Illumina sequencing technology uses individual hydrofill chambers to extend individual molecules. This method tends to give a high degree of accuracy of nucleotide sequence although it generates shorter sequence reads which are 100-200 bases long. On the other hand, Pacific bio (PacBio) sequencing based on immobilized polymerase based machine can generate longer reads, around 10 kilo bases with a lower degree of accuracy of nucleotide sequence. These two methods have their own advantages and disadvantages, therefore, their advantages could be enhanced by combining them using several bioinformatics tools. Secondly, raw sequence data (relatively short read sequences) can be applied for the compensation and as combination to produce long sequences (contigs). In recent years, whole genome sequence data has been increased as a result of NGS technology improvement (Quail et al., 2012). Sequence processing software tools have also accordingly increased giving a diverse range of choice. When we enter the bioinformatics field, software application choices are critical. There are several different softwares that can align short read sequences with long read sequences. The goal of whole-genome shotgun assembly is to represent each genomic sequence in one scaffold; however, this is not always possible (Nordberg et al., 2014). The identification of species-specific genes and the development of species-specific diagnosis molecular diagnostics can be achieved by comparative genomics among *Trypanozoon* using their high quality of whole genome sequence data. Therefore, objective of this chapter is assembling and processing whole genome sequence of *T. equiperdum* IMV-t1 strain.

3.2 Materials and methods

3.2.1 DNA extraction and purification

The total DNA of culture-adapted *T. equiperdum* IVM-t1 propagated on soft agarose medium was extracted and purified using TE-saturated phenol (Sigma-Aldrich Japan) and phenol-chloroform isoamyl-alcohol solution (Sigma-Aldrich Japan) (Sambrook and Russell, 2006). Purified DNAs were kept at -30ºC until use.

3.2.2 Whole genome sequencing using NGS

The total DNA of *T. equiperdum* IVM-t1 was sequenced using Miseq (Illimina, Inc., CA, USA) and PacBio RSII (Pacific Bioscience, Inc., CA, USA). Genomic sequence data which generated from MiSeq had a high degree of accuracy of nucleotide sequencing although it had shorter reads (100 – 200 bp/read). On the other hand, PacBio generated a longer sequence reads (~ 10 kbp/read) but less accuracy of nucleotide sequencing. The NGS analyses of *T. equiperdum* IVM-t1 were performed by Associate Professor Junya Yamagishi, Research Center for Zoonosis Control, Hokkaido University.

3.2.3 Genome assembly

3.2.3.1 Outline of genome assembly

Whole genome assembly were carried out with Bowtie-1.2.1.1 (Langmead et al., 2009), Pilon (Walker et al., 2014) and ABySS- 2.0.2 (Simpson et al., 2009). Two generated data were integrated into the draft genome of *T. equiperdum* by Metassembler-1.5 (Wences and Schatz, 2015). (Fig. 5).

3.2.3.2 Genome assembly based on PacBio long reads

Bowtie software aligned Illumina short reads to PacBio long reads as reference. In other words, Illumina short reads were mapped on PacBio long reads. During this processing by Bowtie software, several analyses such as database building (indexing) and actual aligning were also processed for next step. The output of bowtie's alignment, which is the SAM file format, was converted to the BAM file format using SAMtools (Li et al., 2009). Prior to the final base calling by Pilon software, bowtie's alignment BAM file was sorted and indexed using SAMtools once

more. The variant calling was processed using data which was produced by previous Bowtie and SAMtool steps using Pilon software. In other words, the sequence errors by PacBio long reads (low accuracy) were corrected using mapped Illumina short reads (high accuracy) by Pilon software. A FASTA file containing an improved nucleotide sequence data and an optional VCF file containing nucleotide variation between PacBio reads and Illumina reads were produced (Walker et al., 2014).

3.2.3.3 Genome assembly based on Illumina short reads

Paired end Illumina short reads were assembled using ABySS-2.0.2 software. Before the assembly, the contamination of DNA from other organisms (horse, human, bovine, virus, bacteria) were filtered out from Illumina reads by DeconSeq software (Schmieder and Edwards, 2011). In order to improve assembly quality, the quality filtering of raw Illumina short reads was performed using NGS QC Toolkit-2.3.3 software (Patel and Jain, 2012) in order to improve assembly quality. To produce most ideal contigs, the optimized k-mer value was analyzed among 25 to 80 at intervals of 5. Better k-mer values were obtained among 65 to 70, therefore, the optimized the k-mer was analyzed among 65 to 70 at intervals of 1. Through the k-mer optimization step, the k-mer value was optimized at 68. Using this k-mer parameter, *de novo* genome based on Illumina short read was constructed.

3.2.3.4 Construction of the draft genome of T. equiperdum

Two generated data, which were PacBio based sequence data (3.1.3.2) and Illumina based sequence data (3.1.3.3), were integrated into more long and short numbers of contigs by Metassembler-1.5 (Wences and Schatz, 2015) and finally produced as the draft genome of *T. equiperdum* IVM-t1.

3.2.4 Reference-based gene prediction (genome annotation) and phylogenetic analysis

Chromosome construction, gene prediction and gene annotation of each predicted gene of *T. equiperdum* IVM-t1 were performed using the published genome of *T. brucei* TERU927 as reference by companion pipeline; a web server for annotation and analysis of parasite genomes (https://companion.sanger.ac.uk/) (Steinbiss et al., 2016).

In the first step by companion pipeline, the generated contigs of the draft genome of *T. equiperdum* were re-constructed to be each eleven chromosomes and one bin-contig using the chromosome structure of the genome of *T. brucei* as reference. In the next step, genes of *T. equiperdum* were predicted based on homology using the gene information of *T. brucei* as reference by transfer of highly conserved gene models. In addition, the *de novo* gene prediction of *T. equiperdum* IVM-t1 was performed using self-learning system of the companion pipeline based on nucleotide sequence, genes, intron and exon structures of *T. brucei* as references. Through these two gene prediction steps, the homologous genes and *T. equiperdum* specific genes, which were not sharing any homology in *T. brucei* reference genome, were identified and annotated. Also, the Companion pipeline visualized the gene homology of each chromosome between *T. equiperdum* IVM-t1 and *T. brucei* TERU927. The summary of gene prediction is shown in Table 6.

The phylogenic tree was constructed based on the amino acid sequence of 50 core proteins, which were automatically selected by the companion pipeline, of *T. equiperdum IVM-t1*, *T. equiperdum* OVI, *T. evansi* STIB805, *T. brucei* TERU927, *Leishmania donovani* BPK282A1*, L. infantum* JPCM5*, L. braziliensis* MHOM/BR/75/M2903 and *L. major* Friedlin. The alignment was performed using ClustalW multiple alignment in Bioedit (Hall, 1999; Thompson et al., 1994). The phylogenetic tree was built based on the alignment using Fitch-Margoliash method in Bioedit (Fitch and Margoliash, 1967).

3.3 Results

3.3.1 NGS sequencing

The whole genome sequencing of culture adapted *T. equiperdum* IVM-t1 was conducted with 300ng/µl concentration of the genomic DNA using Illumina and PacBio platforms. The 23.6 MB FASTA and 46.4 MB FASTAQ files were generated by PacBio NGS. The 18.5 GB pair-end (forward and reverse) sequence data file was produced by Illumina NGS. The coverage of read depth of *T. equiperdum* IVM-t1 genome was fifty-fold and five fold using paired-end sequencing of Illumina and long read sequencing of PacBio, respectively.

3.3.2 Genome assembly

Raw sequence reads from PacBio and Illumina were successfully assembled by Bowtie2 and ABySS softwares, respectively. ABySS assembled genome based on Illumina reads generated 102,473 contigs with N50 value of 17,434 and the cumulative length of 35,594,521 bp. On the other hand, Bowtie and Pilon assembled genome based on PacBio reads produced 418 contigs with N50 value of 460,897 and the cumulative length of 35,271,224 bp. The draft genome which was integrated Illumina and PacBio based analyzed data was assembled into 45 contigs with N50 value of 859,849 and the cumulative length of 26,988,997 bp (≈ 27 Mbp) (Table 5 and Figure 6 and Figure 7). Comparison of the draft genome of *T. equiperdum* IVM-t1 with the preciously published genomes of *T. brucei* (Berriman et al., 2005) and *T. evansi* (Carnes et al., 2015) as reference sequences, showed each evaluative parameter of the draft genome of *T. equiperdum* IVM-t1 were at the same level. Moreover, all of the parameters were significantly improved in comparison with previously published genome of *T. equiperdum* OVI (Table 5). The total genome size of *T. equiperdum* IVM-t1 which is ≈ 27 Mbp is almost the same as *T. equiperdum* OVI, and slightly bigger than *T. brucei* TREU927 which ≈ 26 Mbp and *T. evansi* STIB805 which is ≈ 25.2 Mbp)

3.3.1 Reference-based Gene prediction and phylogenetic analysis

The 27 Mbp *T. equiperdum* IVM-t1 genome contains 7,995 genes, 7,893 coding genes, 102 non-coding genes, and 2,298 pseudogenes (Table 6.). Comparison of the predicted genes of *T. equiperdum* IVM-t1 with the reference gene sets of *T. brucei* (TREU927), revealed that 6,265 protein-coding gene clusters were shared in both trypanosome species, 218 protein-coding

gene clusters were *T. equiperdum* IVM-t1 specific, and 420 singletons (i.e. genes without orthologs and paralogs in either species) (Figure 8.).

The contigs of *T. equiperdum* IVMt1 were re-arranged into 11 chromosomes and one bin-contigs based on genomes of *T. brucei* (TREU927) as reference. These circular plots showed alignments between each genes of reference chromosome and genes of newly annotated *T. equiperdum* IVM-t1 chromosome (Figure 9. A-L). Several inversions, translocation, singletons, gaps and missing core genes were observed in *T. equiperdum* IVM-t1 genome.

A phlylogenetic tree was built with *T. equiperdum* IVM-t1, *T. equiperdum* OVI, *T. evansi* STIB 805, *T. congolense* IL3000 and *Leishmania* spp. (Figure 10.). Amongst the subgenus *Trypanozoon* cluster, the *T. equiperdum* IVM-t1 and *T. evansi* STIB805 were grouped together, whilst *T. brucei* TREU 927 and *T. equiperdum* OVI were grouped together.

3.4 Discussion

The draft genome of *T. equiperdum* IVM-t1 was successfully assembled in this study. This is the second draft genome of *T. equiperdum*. The first draft genome of *T. equiperdum* has already been published (Hebert et al., 2017). However, assembly quality and contig numbers of the draft genome of *T. equiperdum* IVM-t1 were significantly improved comparison with the genome of *T. equiperdum* OVI (Hebert et al., 2017). This improved *T. equiperdum* genome data has long been needed for further studies on *T. equiperdum* and dourine.

Overall genome size of the assembled genome of *T. equiperdum* IVM-t1 (≈ 27 Mbp) is same level with existing genome of *T. equiperdum* OVI (≈ 27 Mbp) (Hebert et al., 2017), *T. evansi* STIB805 (≈ 25.2 Mbp) (Carnes et al., 2015), and *T. b. gambiese* DAL 972 (≈ 26 Mbp) (Jackson et al., 2010). Large insertions, deletions, and rearrangements of the genome of *T. equiperdum* IVM-t1 were found in chromosome level assembly comparison with *T. brucei* (TREU927). Assembly errors, which contains a significantly large insertion, deletion, inversion, or rearrangement in assembled regions, were sometimes produced as artifacts of assembly program (Muggli et al., 2015). Similarly, these artifacts (insertions, deletions, and rearrangements) were mainly caused by assembly error.

The total genome size of *T. equiperdum* IVM-t1 is slightly bigger (≈ 1Mb) than reference *T. brucei* (TREU927). The size of chromosome 1 (0.9 Mb) and 10 (4.0 Mb) of *T. equiperdum* IVM-t1 has same size with counterpart reference chromosome. The size of chromosome 2 (0.8 Mb), 5 (1.4 Mb), 7 (2.1 Mb), 8 (1.5 Mb), 9 (2.4 Mb) and 11 (3.6 Mb) is smaller than the size of counterpart reference chromosome. The size of chromosome 3 (2.3 Mb), 4 (2.4 Mb), 6 (2.2 Mb) is bigger than the size of corresponding reference chromosome. Moreover, most of the unallocated contigs were arranged into ‘bin-contigs’ (Figure 9.L). Many of the homologous genes of *T. brucei* were located in same chromosomes of *T. equiperdum*. While, some of the homologous genes in *T. brucei* were located in bin-contig of *T. equiperdum* IVM-t1. Therefore, I need to improve the quality of gene prediction by re-analysis of genome using other software and RNA-sequence to identify the actual gene. Most of the protein-coding genome clusters (6,265 protein-coding gene clusters) were shared in both *T. equiperdum* IVM-t1 and *T. brucei* TERU927, however, 218 protein-coding gene clusters were *T. equiperdum* IVM-t1 specific. In addition, the 1,353 protein-coding gene clusters and 735 singletons were specifically found in *T. brucei*. Similarly, several species-specific genes were also found during comparative genome analysis between *T. evansi* and *T. brucei* (Carnes et al., 2015). Species identification and species-specific diagnostic methods can will be developed using these species-specific DNA markers. Moreover,

these results suggested that the singletons and species-specific protein coding clusters were related to species-specific parasitism. Comparative genomics among trypanosomes is strongly expected for discovery of species-specific genes and identification of their function in further studies.

Phylogenetic analysis has revealed the close relationship among *Trypanozoon* group like as previously reported. Among *Trypanozoon*, *T. equiperdum* IVM-t1 was more closely related to *T. evansi* STIB 805 than *T. equiperdum* OVI. This result suggests that *T. equiperdum* OVI might have been misidentified *T. brucei* variant. Claes et al. (2005) reported that *T. equiperdum* strains, which lacks the RoTat 1.2 VSG gene (BoTat and OVI) are more closely related to *T. b. brucei*, whereas all other *T. equiperdum* strains, which has the RoTat 1.2 VSG gene, clustered in a homogenous group of *T. evansi* (Claes et al., 2005; Claes et al., 2003a). However, Suganuma et al (2016) revealed that *T. equiperdum* IVM-t1 strain lacks the RoTat1.2 VSG gene. Alternatively, this phylogenetic tree indicates polyphyletic origin of *T. equiperdum*. This is in good agreement with observations of Carnes et al. (2015) who also concluded that extant strains of *T. evansi* or *T. equiperdum* are polyphyletic and evolved on at least four independent occasions. I have to reveal the taxonomical relationship of *T. equiperdum* by whole genomic analyses using other *T. equiperdum* strains.

In conclusion, whole genome draft assembly produced by the current study provides a resource for future trypanosome genetic studies and has revealed some *T. equiperdum* specific genes.

Table 4. Differences and similarity among the Trypanozoon

Character	***T. evansi***	***T. equiperdum***	***T. brucei***
Natural hosts	Variety of domestic and wild animals.	Equines only	Variety of domestic and wild animals.
Vectors and transmission mode	Biting flies (e.g.Tabanus spp, Stomoxys spp), vampire bats, mechanical transmission	Direct transmission through coitus	Tsetse flies, cyclical
kDNA	Minicircles only	Minicircles; some case Maxicircles also present	Complete set of Mini and Maxicircles
RFLPs in rDNAs and VSG genes	No difference		

Table 5. Genome sequence assemblies

Assembly statistics	*T. equiperdum* IVM-t1 (Analyzed after ABySS)	*T. equiperdum* IVM-t1 (Analyzed after Bowtie)	*T. equiperdum* IVM-t1 (Draft genome)	*T. equiperdum* OVI V2 (REF)
# contigs (>1000 bp)	1,536	418	45	2,026
# contigs (>5000 bp)	886	411	45	923
# contigs (>10000 bp)	582	383	45	607
# contigs (>25000 bp)	263	215	44	281
# contigs (>50000 bp)	120	93	44	122
Total length (>0 bp)	35,594,521	35,271,224	26,988,976	26,228,029
Total length (>1000 bp)	24,398,065	35,271,224	26,988,976	26,228,029
Total length (> 5000 bp)	22,690,418	35,242,381	26,988,976	23,549,827
Total length (>10000 bp)	20,458,552	35,025,764	26,988,976	21,281,451
Total length (>25000 bp)	15,419,530	32,025,134	26,965,930	16,219,410
Total length (>50000 bp)	10,279,886	28,063,110	26,965,930	10,619,978
# contigs	102,473	418	45	2026
Largest contig	196,715	1,700,184	1,700,184	367571
Total length	35,594,521	35,271,224	26,988,976	26,228,029
GC (%)	43.69	45.41	46.23	45.95
N50[a]	17,434	460,897	859,849	38,149
N75[b]	205	78,181	484,939	13,817
L50[v]	379	20	12	180
L75[f]	7,621	67	22	470
# N's per 100 kbp	113.13	0	0	437.4

[a] the N50 (length) is defined as the shortest sequence length at 50% of the genome (large numbers suggests good quality)

[b] the N75 (length) is defined as the shortest sequence length at 75% of the genome

[c] the L50 count is defined as the smallest number of contigs whose length sum produces N50 (small numbers suggests good quality)

[d] the L75 count is defined as the smallest number of contigs whose length sum produces N75 (small numbers suggests good quality)

Table 6. The summary of annotated genome statistics of *T. equiperdum* IVM-t1

Genome statistics	Value
Number of annotated regions/sequences	12
Number of genes	7,995
Gene density (genes/megabase)	292.45
Number of coding genes	7,893
Number of pseudogenes	2,298
Number of genes with function	4,736
Number of pseudogenes with function	1,314
Number of non-coding genes	102
Number of genes with multiple CDSs	173
Overall GC%	46.23
Coding GC%	51.03

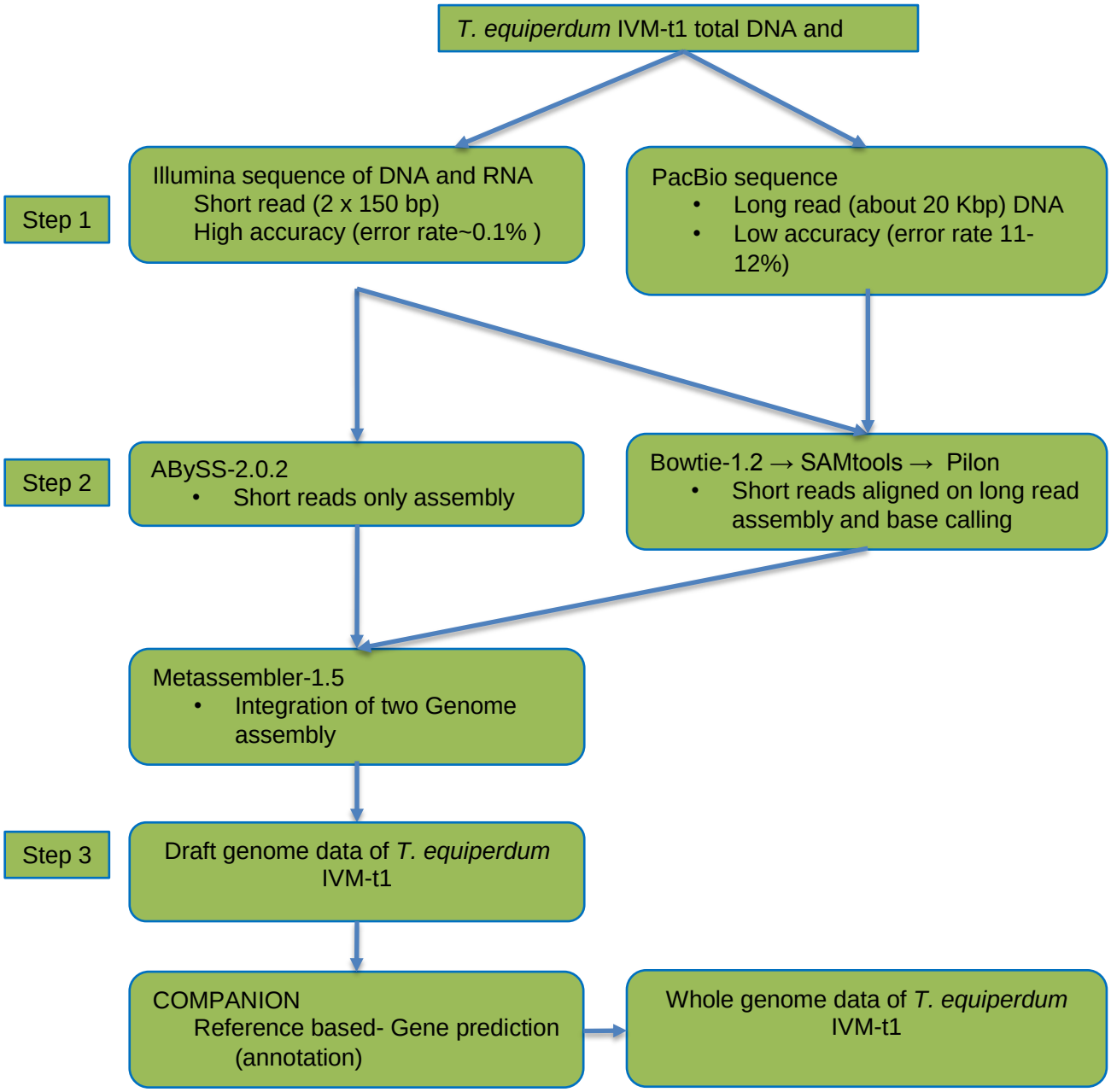


Figure 5. The general pipeline of NGS analysis

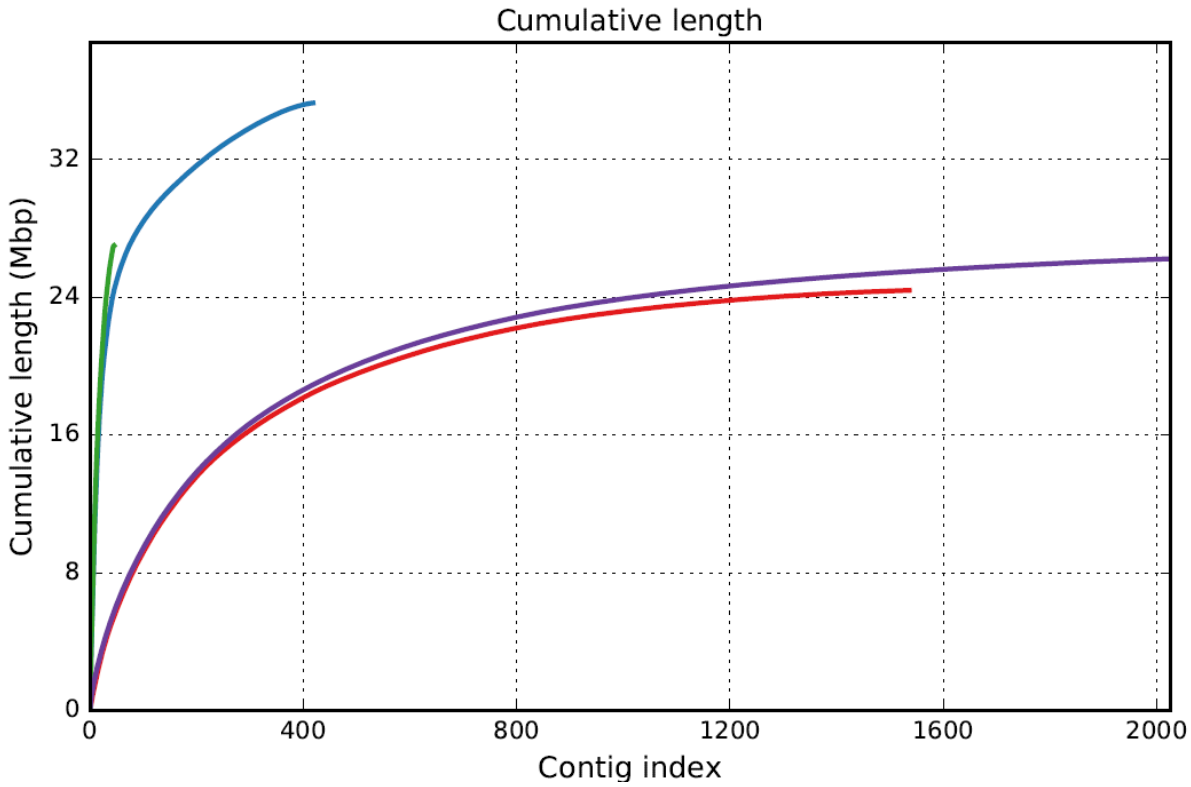


Figure 6. Quality of constructed genome data. Vertical axis shows the cumulative length of the genome assembly by Megabasepair scale. Horizontal axis shows the contigs are ordered from largest (contig #1) to smallest. *T. equiperdum* IVM-t1 draft genome (green), *T. equiperdum* IVM-t1 Bowtie assembly (blue), *T. equiperdum* ABySS assembly (red) and *T. equiperdum* OVI (purple)

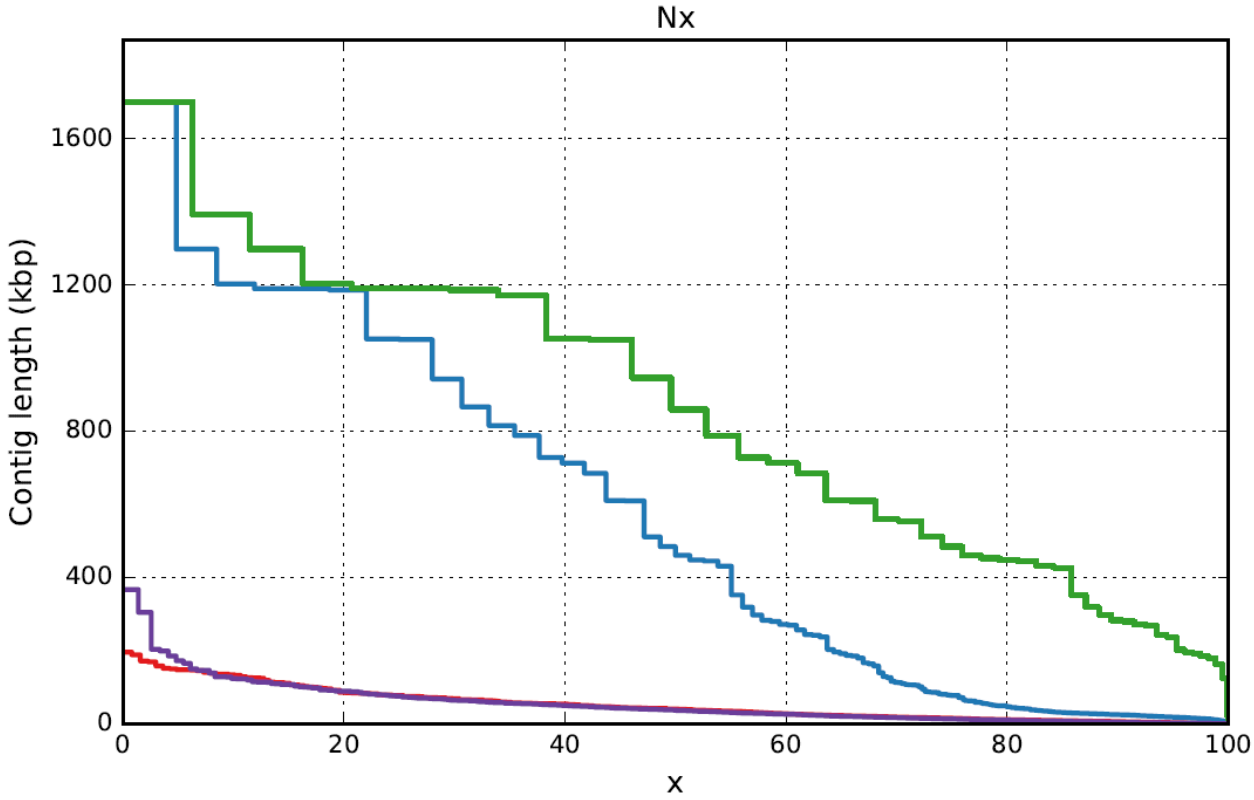


Figure 7. Nx plot of constructed genome data.

Vertical axis shows contig length of the genome assembly by Megabasepair. Horizontal axis shows percentage of total assembly. *T. equiperdum* IVM-t1 draft genome (green), *T. equperdum* IVM-t1 Bowtie assembly (blue), *T. equiperdum* ABySS assembly (red) and *T. equiperdum* OVI (purple)

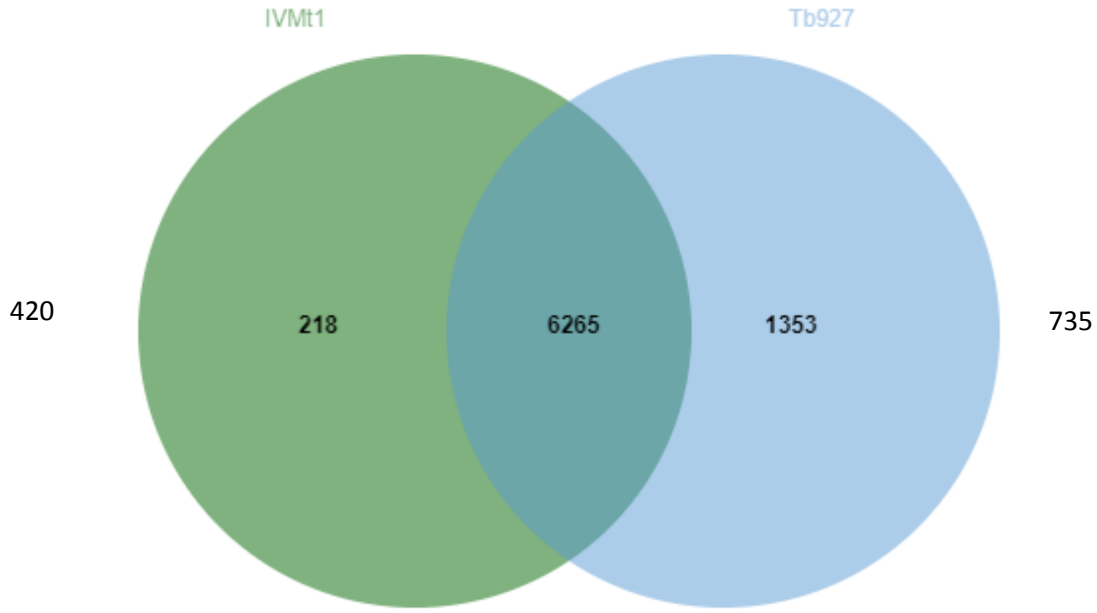


Figure 8. Venn diagram of orthologous CDSs between *T. equiperdum* IVM-t1 and *T. brucei* TREU927.

This diagram shows shared and species-specific protein-coding gene clusters in the target genome of *T. equiperdum* (IVM-t1) (left, green) and the *T. brucei* (TREU927) reference (right, blue). Singletons, i.e. genes without orthologs and paralogs in either species, are placed outside the Venn diagram to the left and right.

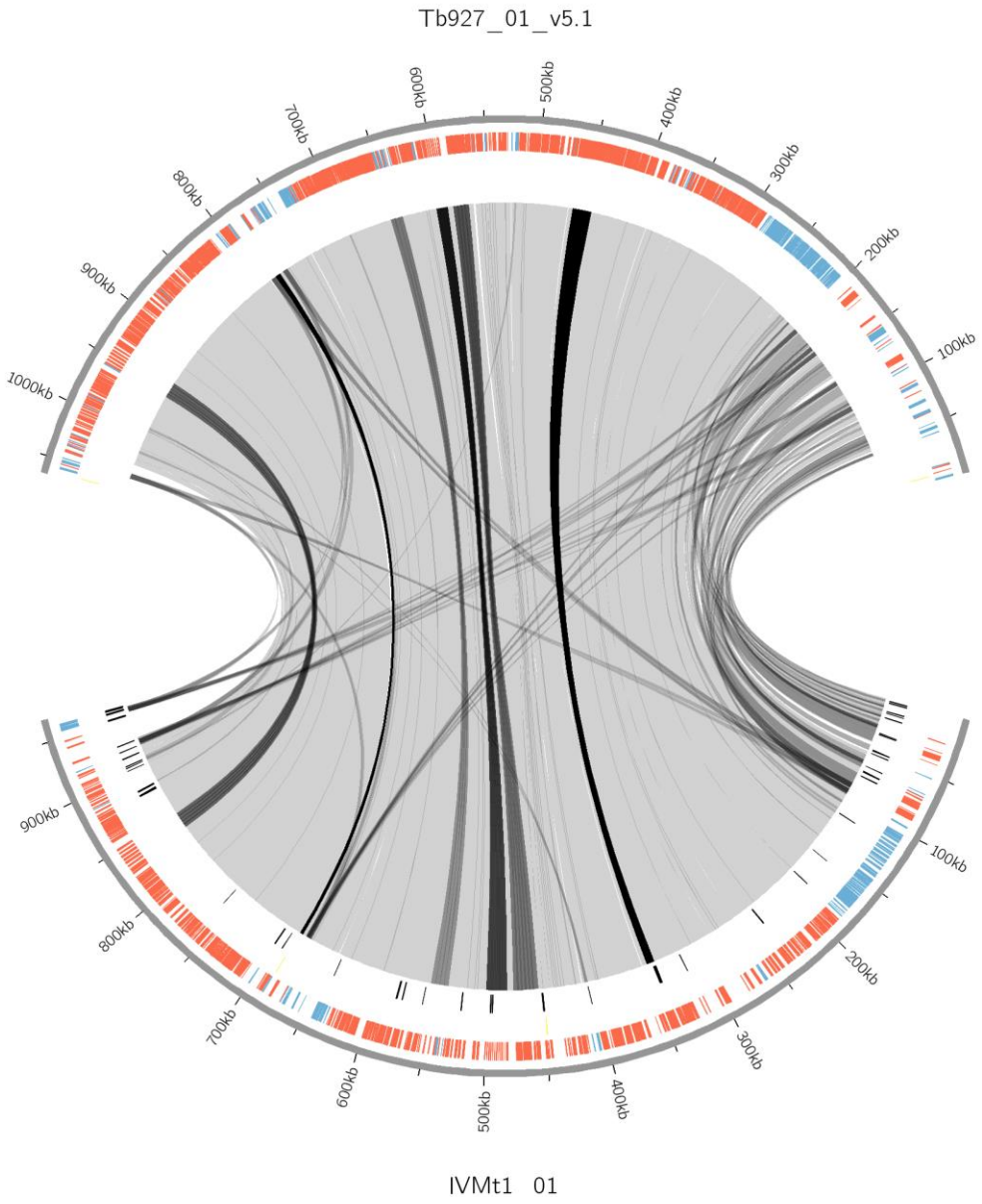


Figure 9. Circular plots of alignments between genes of reference *T. brucei* TREU927 and genes of newly annotated *T. equiperdum* IVM-t1.

Figure 9A. Chromosome 1. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 1 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 1 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are

annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

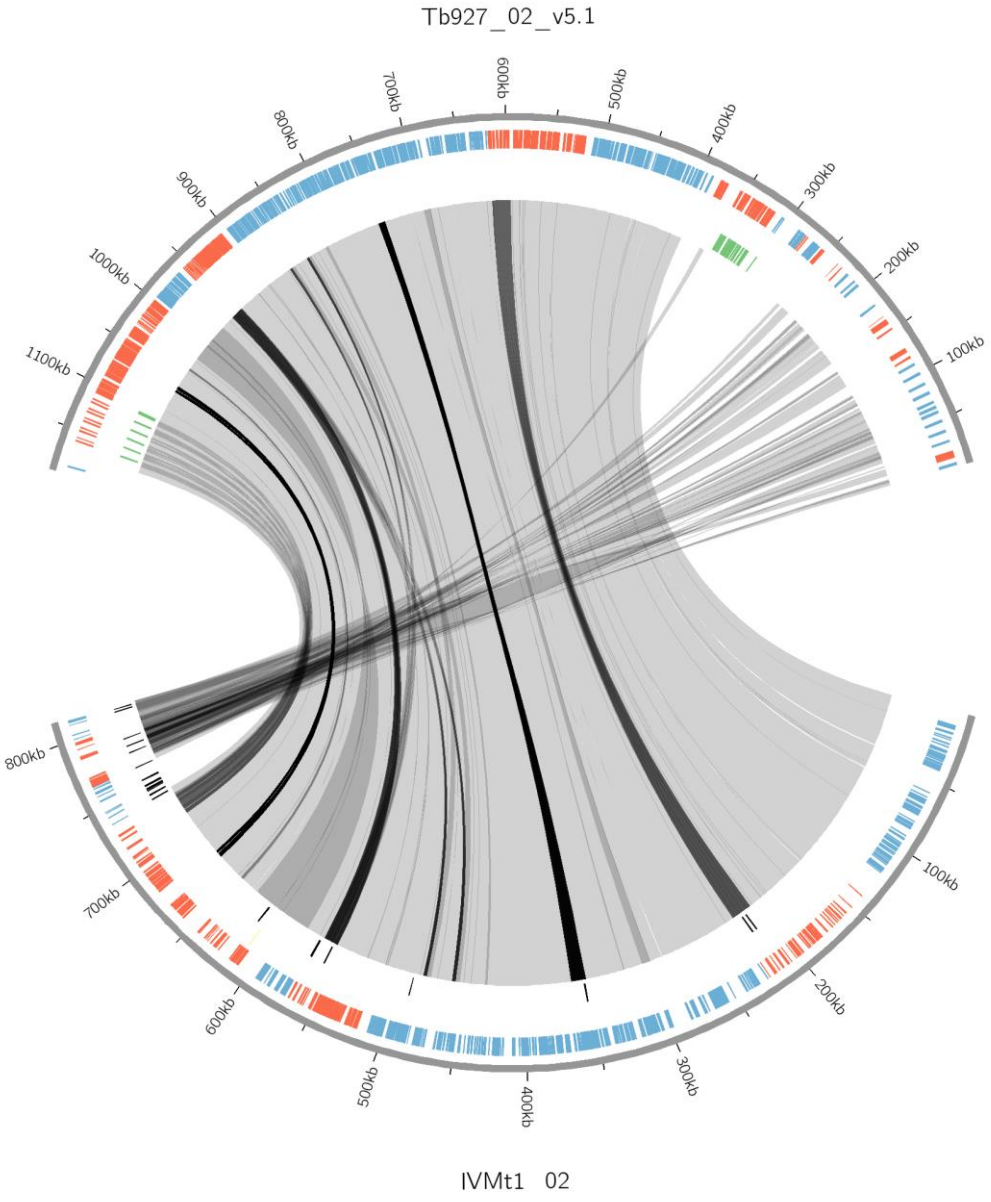


Figure 9B. Chromosome 2. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 2 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 2 (bottom). Grey ribbons between both represent similar regions as identified by

BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

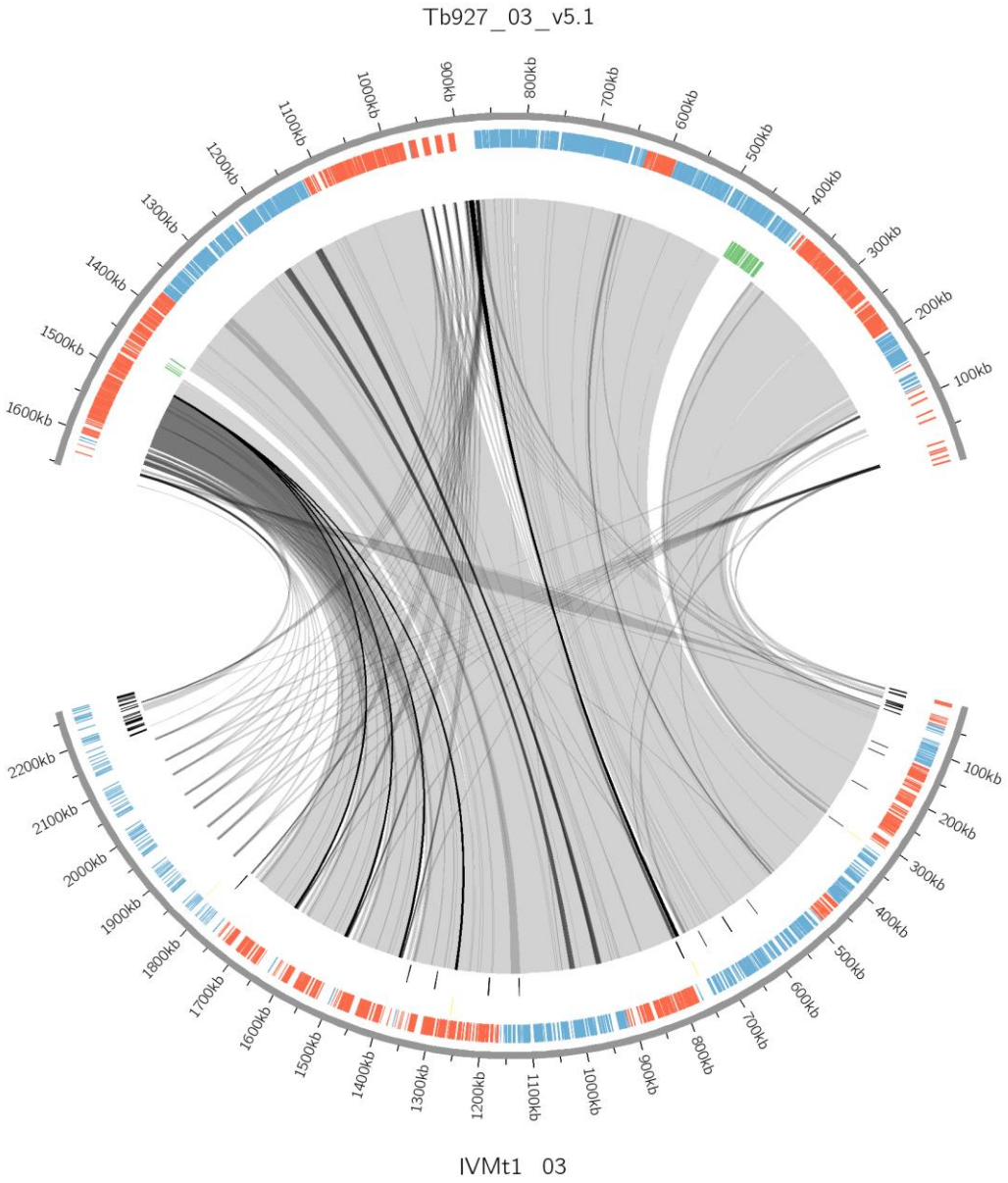


Figure 9C. Chromosome 3. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 3 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 3 (bottom). Grey ribbons between both represent similar regions as identified by

BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

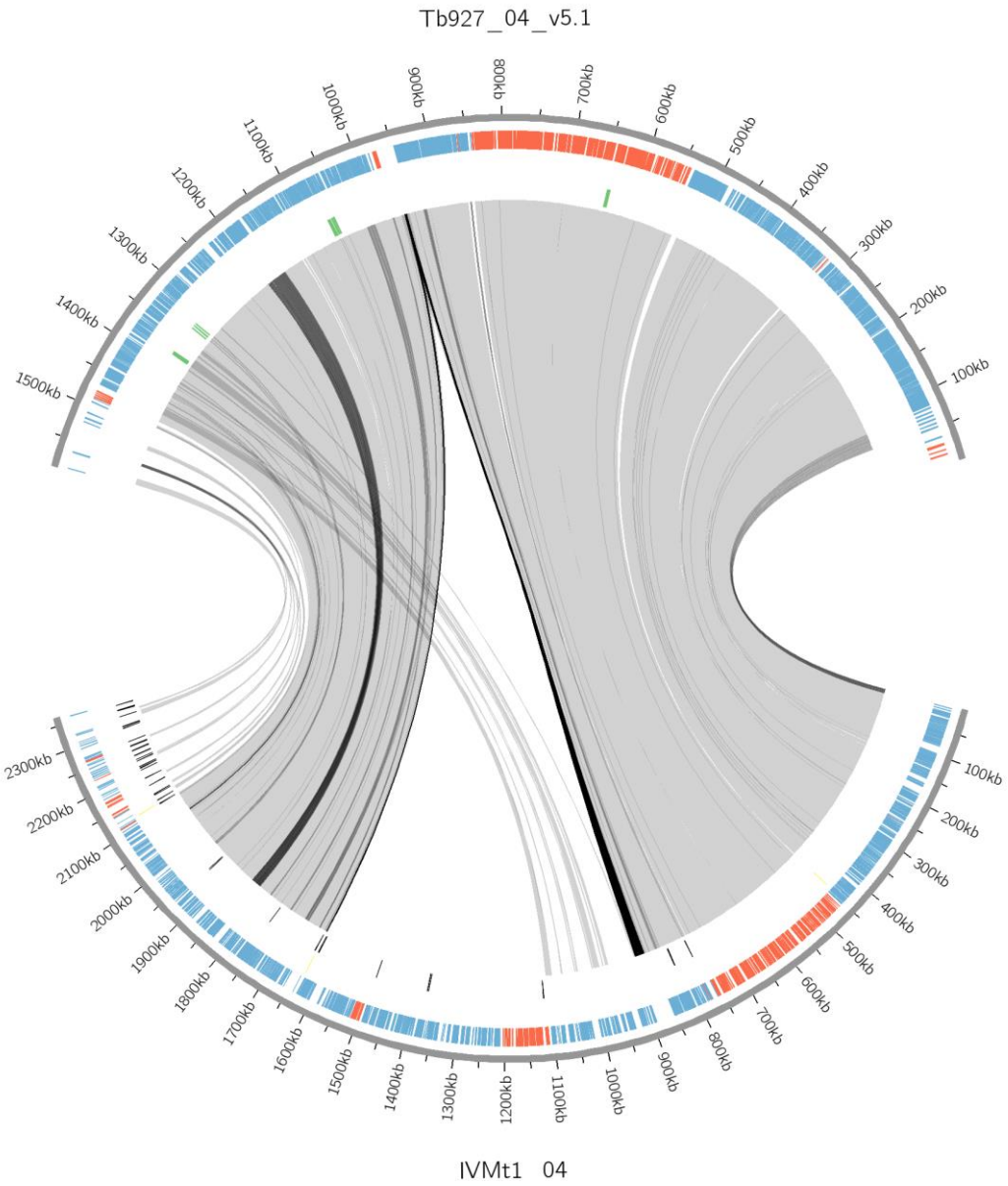


Figure 9D. Chromosome 4. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 4 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 4 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

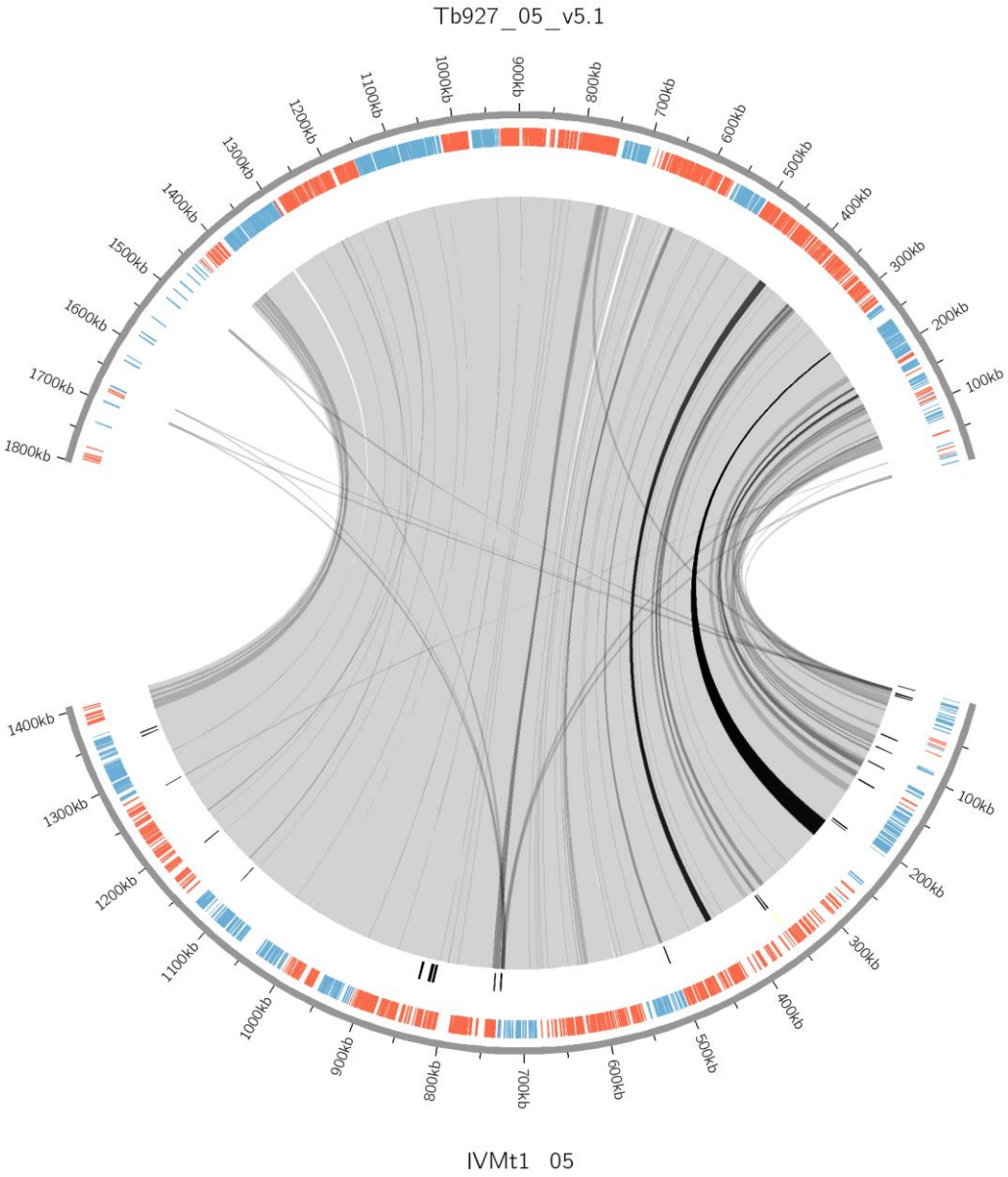


Figure 9E. Chromosome 5. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 5 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 5 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

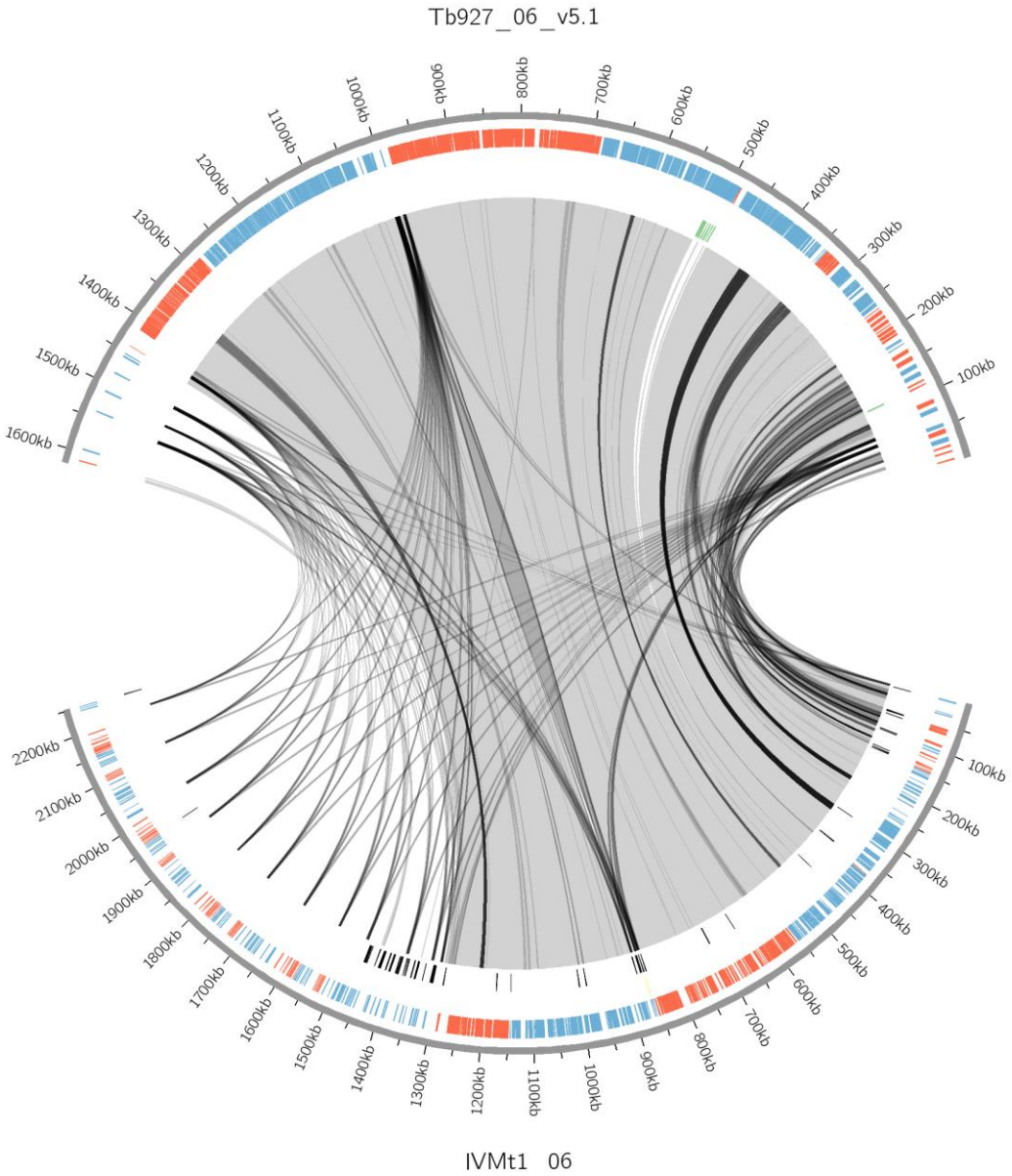


Figure 9F. Chromosome 6. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 6 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 6 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

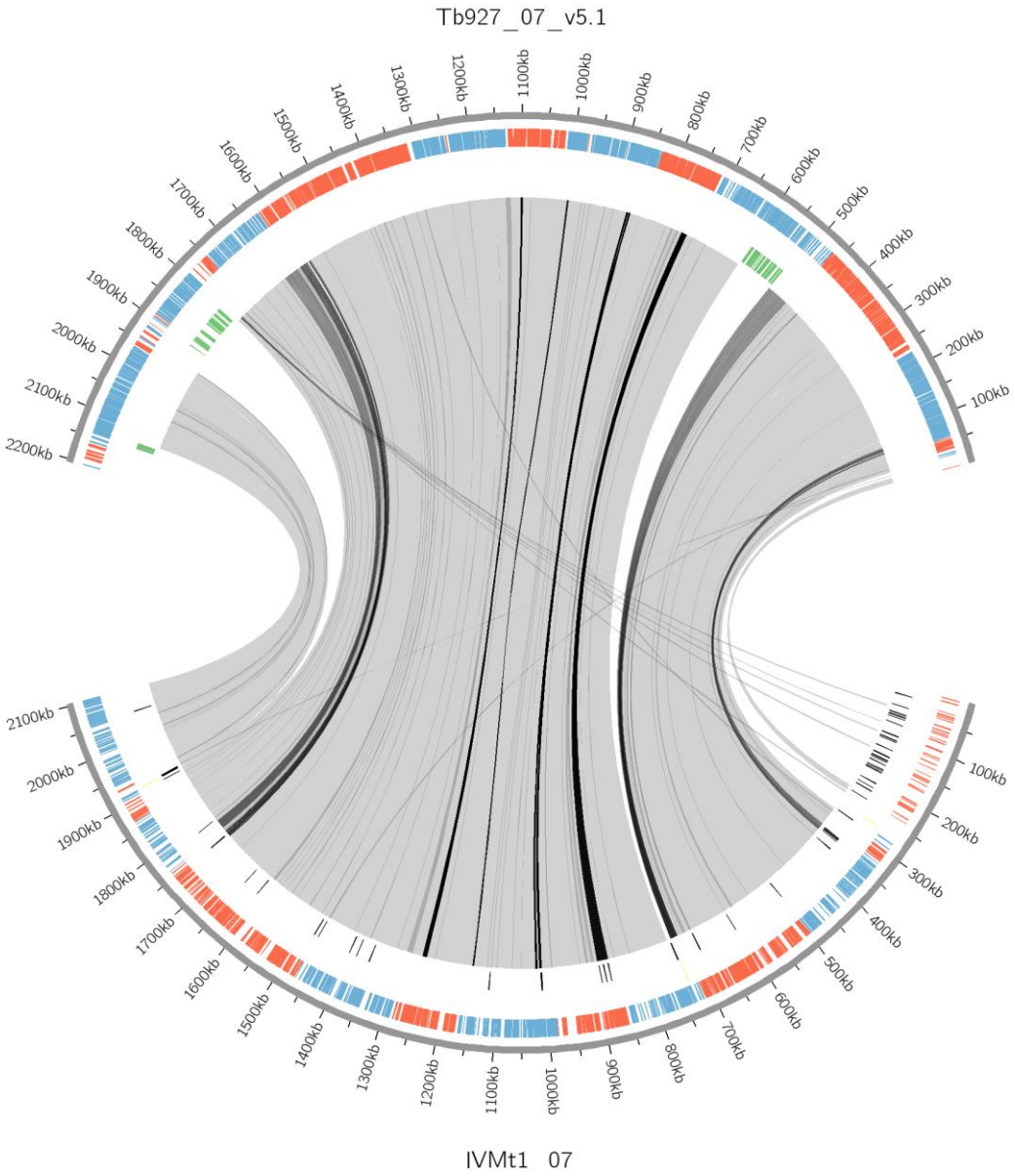


Figure 9G. Chromosome 7. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 7 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 7 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

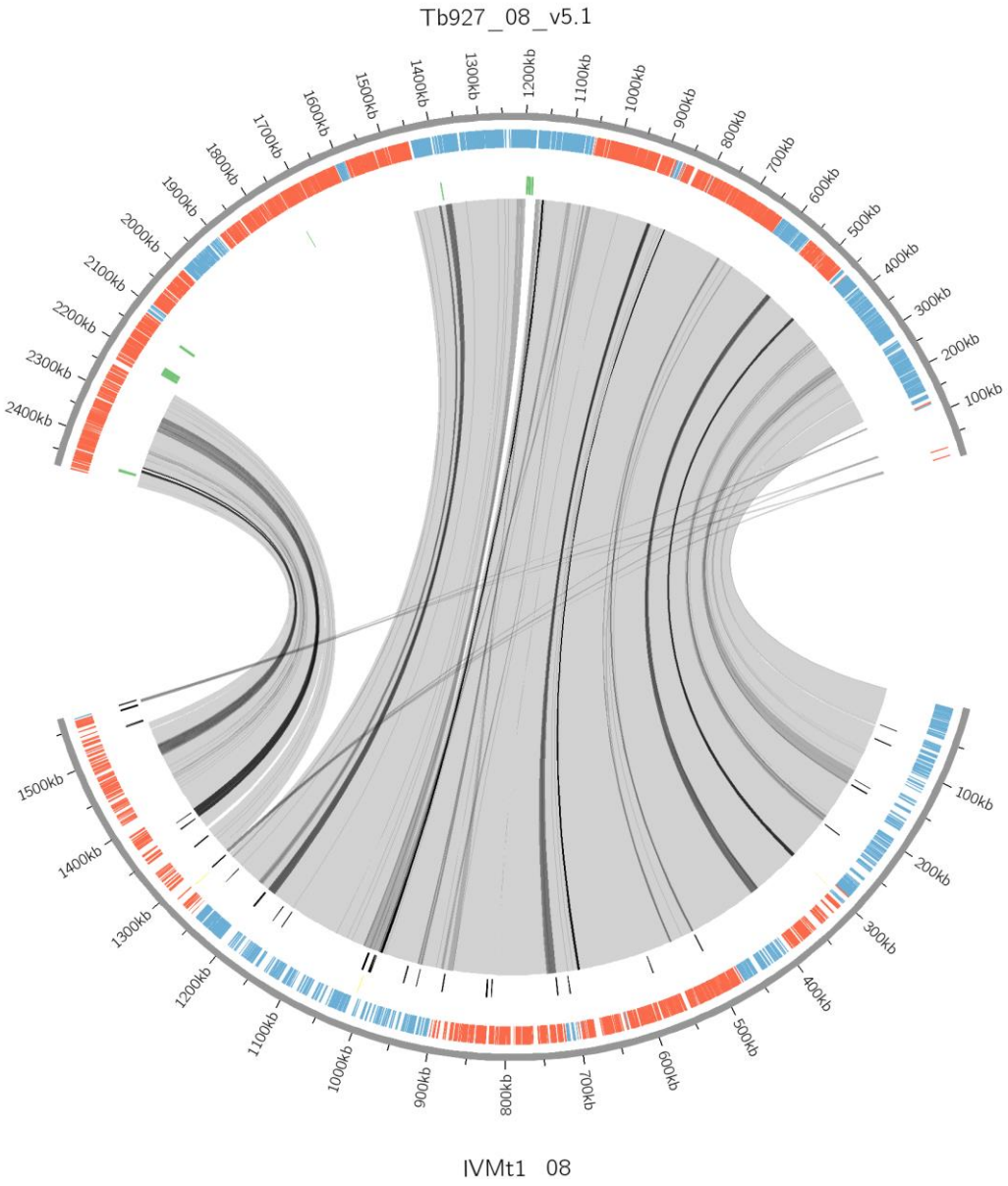


Figure 9H. Chromosome 8. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 8 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 8 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

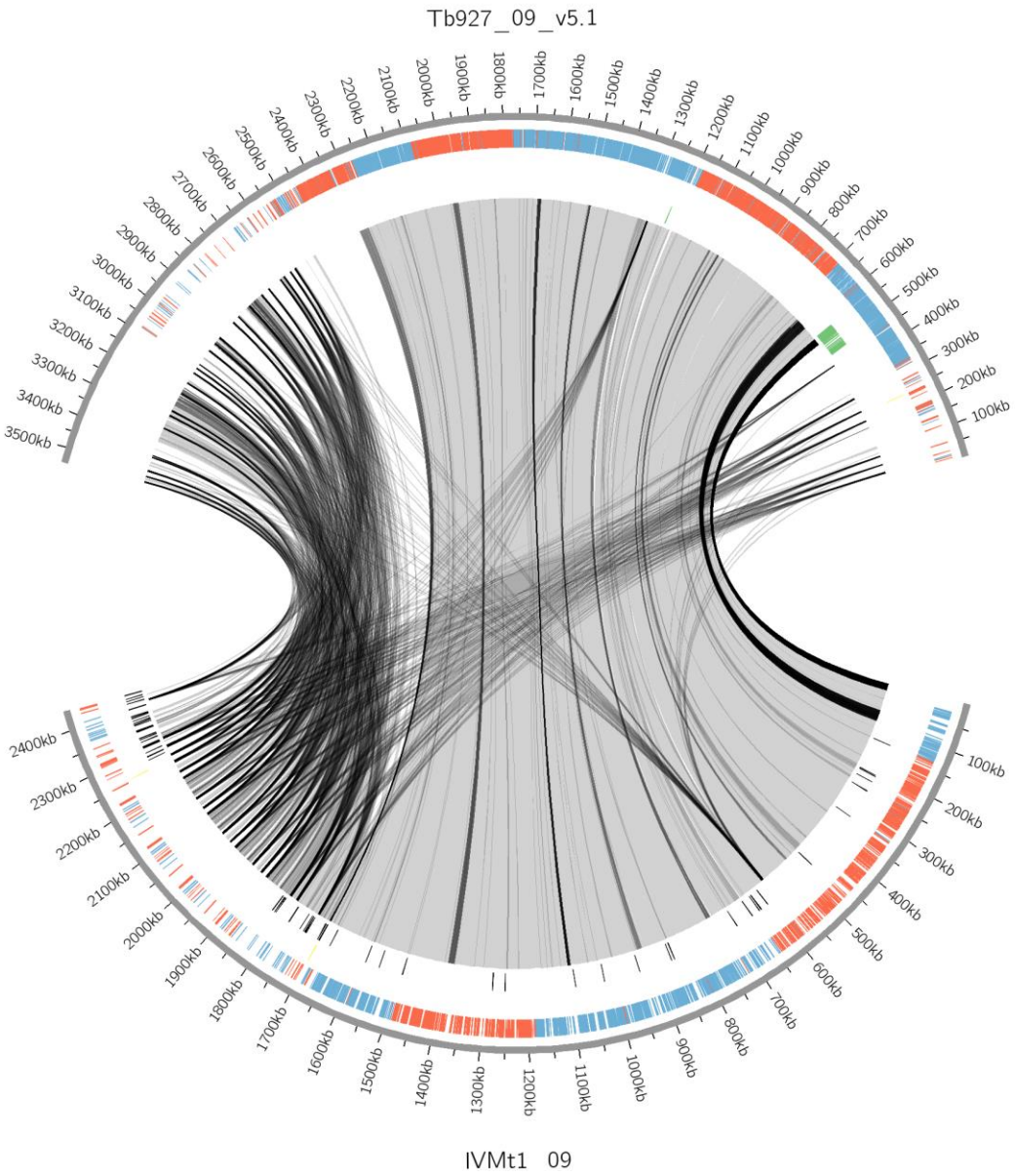


Figure 9I. Chromosome 9. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 9 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 9 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

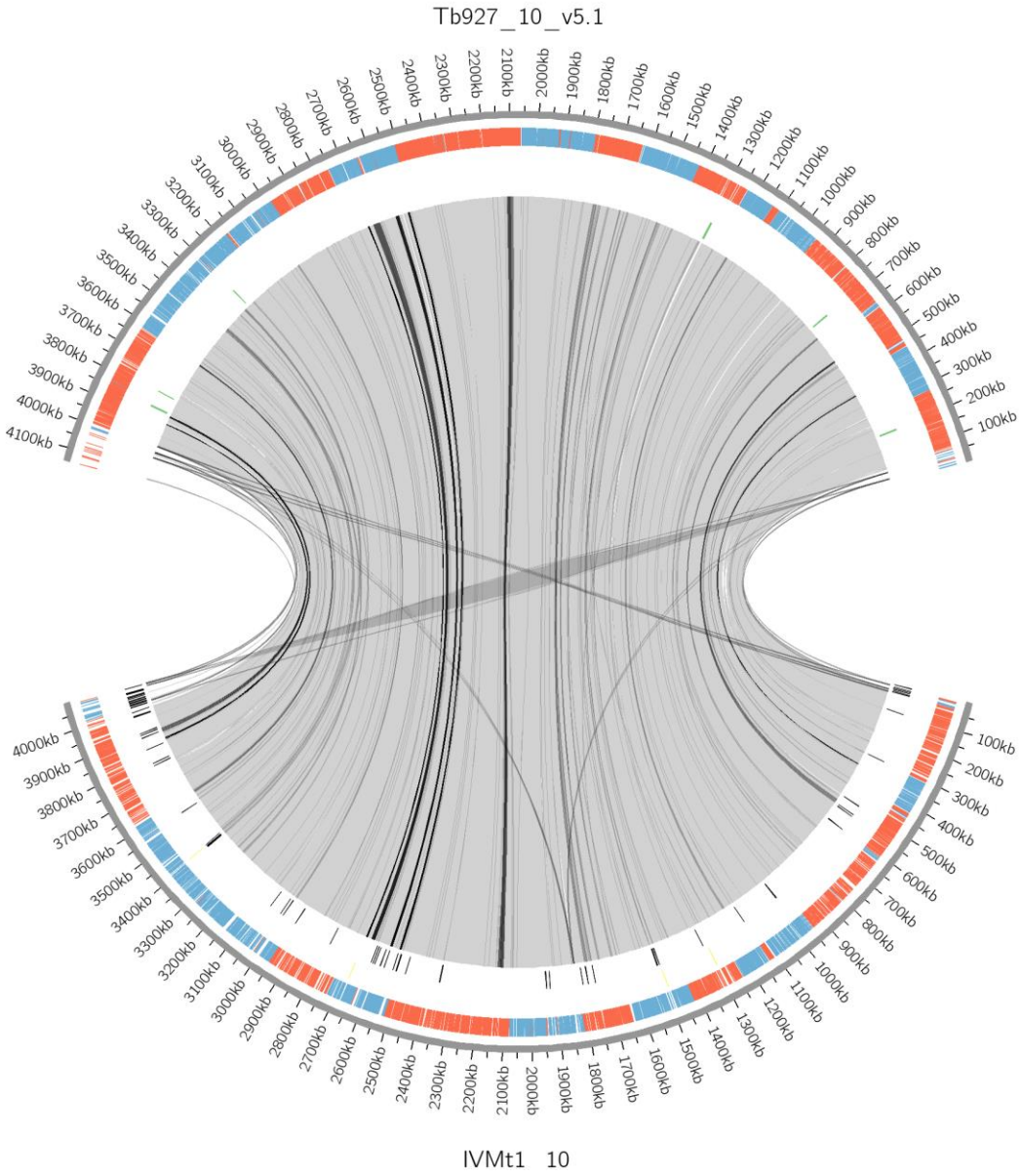


Figure 9J. Chromosome 10. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 10 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 10 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

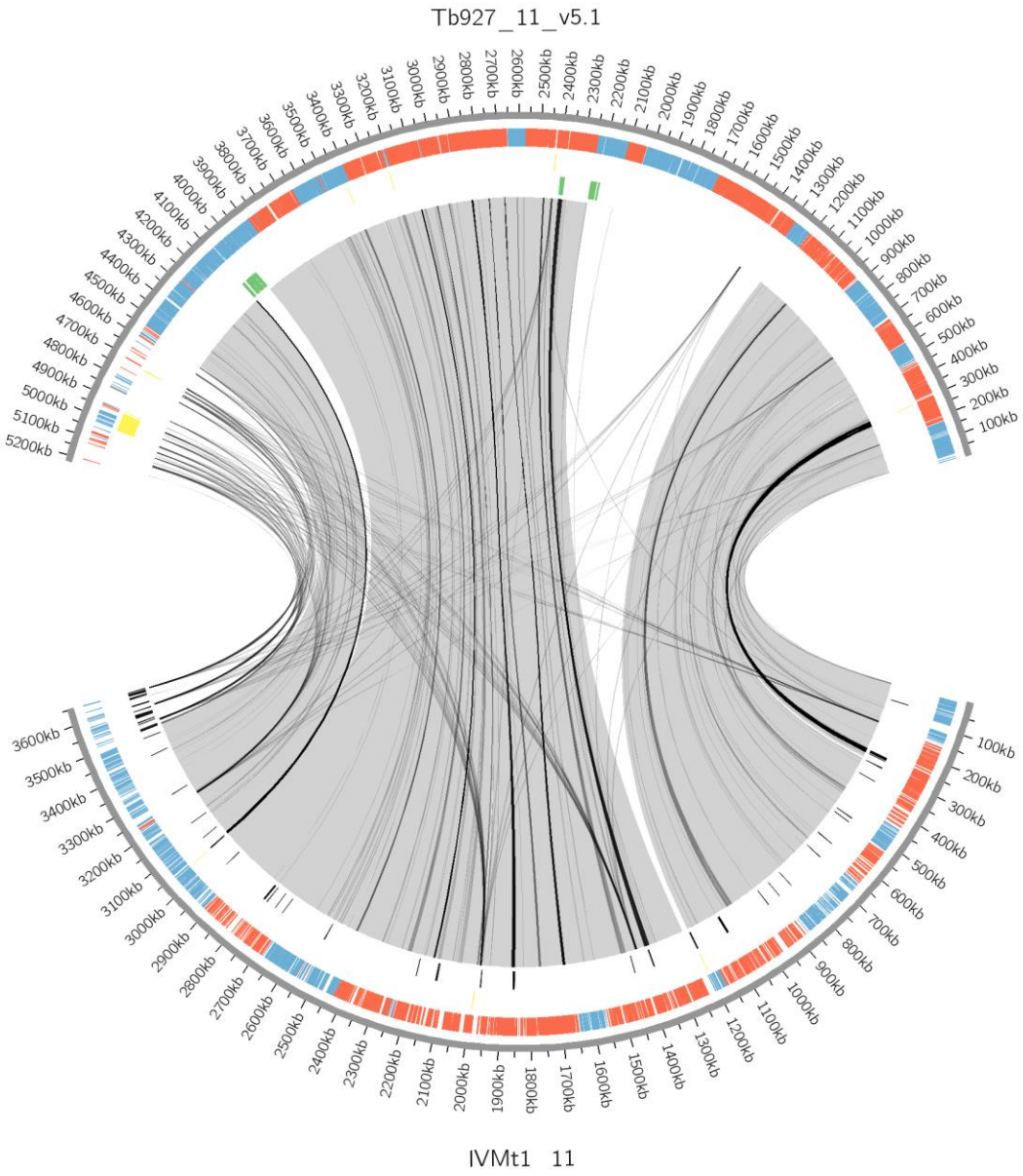


Figure 9K. Chromosome 11. Circular plots of alignments between each genes of reference *T. brucei* (TREU927) chromosome 11 (top) and genes of newly annotated *T. equiperdum* IVM-t1 chromosome 11 (bottom). Grey ribbons between both represent similar regions as identified by BLASTN matches. Genes on the forward strand (blue) and the reverse strand (red) are annotated on the chromosomes, as well as gaps (yellow), singletons (black) and missing core genes (green), each in separate tracks.

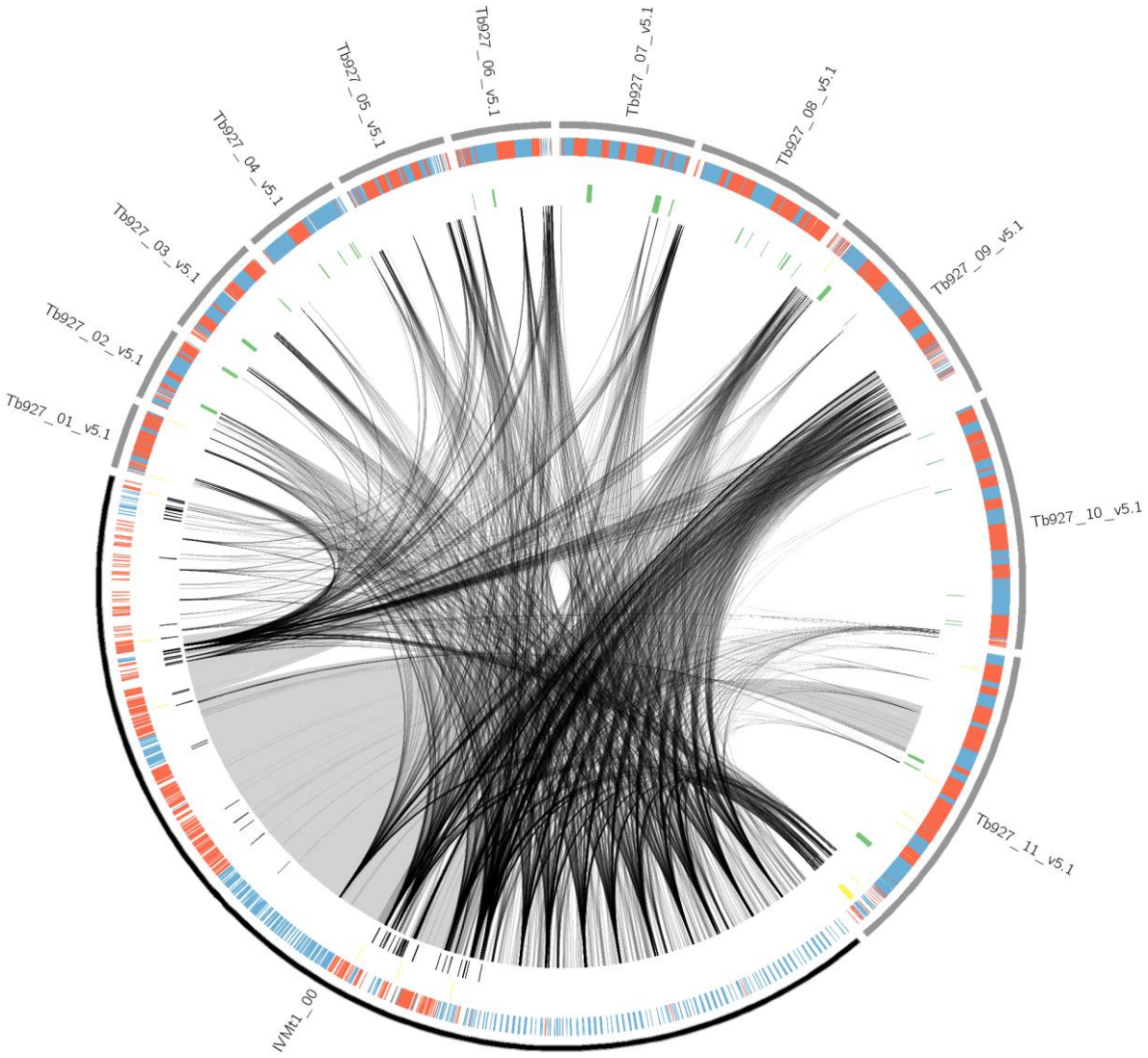


Figure 9L.” bin” sequences. Circular plots of uncontiguated input sequences concatenated into a 'bin' sequence (not drawn to scale) compared to all reference *T. brucei* (TREU927) chromosomes. The reference chromosomes are shown at the top and the uncontiguated part (could not be located in any chromosome) of *T. equiperdum* IVM-t1 genome is shown at the bottom.

Grey ribbons between both represent similar regions as identified by BLASTn matches. Genes on the forward strand (blue), the reverse strand (red), gaps (yellow), singletons (black) and missing core genes (green) are shown on of each chromosomes.

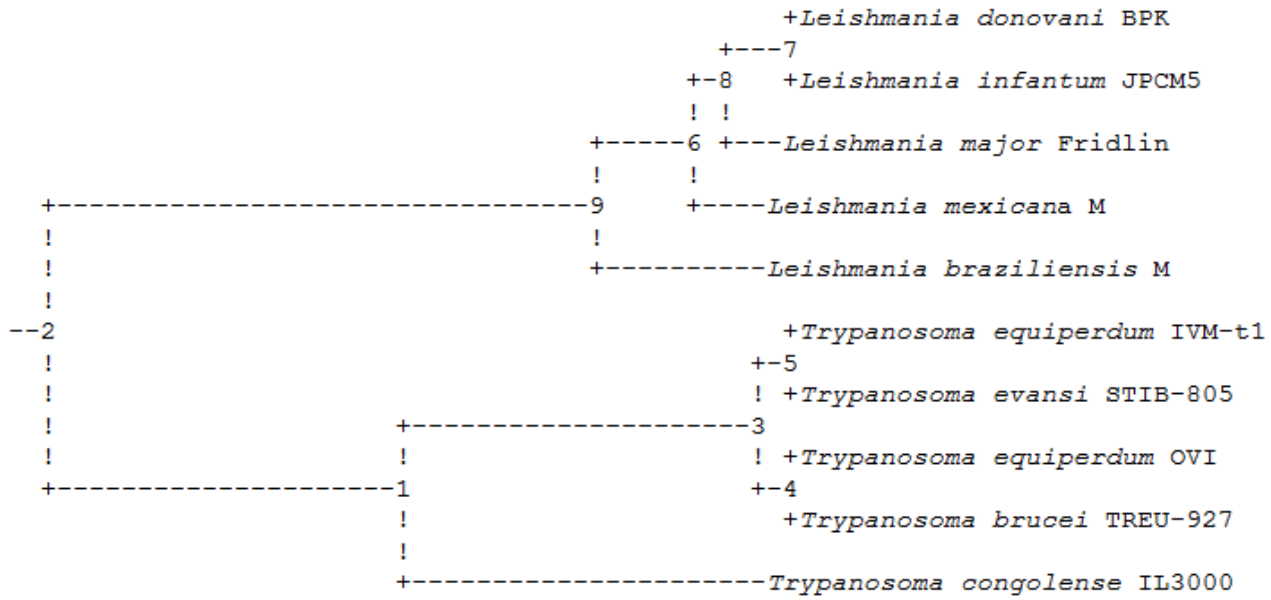


Figure 10. Phylogenetic tree based on 50 core genes by Fitch-Margoliash method.

General discussion

Non-tsetse transmitted animal trypanosomoses are distributed in most of the continents in the world. The diseases are mainly caused by *T. evansi* infections mechanically transmitted via blood sucking insects causing surra and *T. equiperdum* infections transmitted via coitus from animal to animal causing dourine. The epidemics of horse trypanosomosis were highly suspected and caused huge economic damage in Mongolia. These etiological trypanosomes belong to subgenus *Trypanozoon* together with *T. brucei* spp. In case of non-tsetse transmitted trypanosomosis in horse, it has been very difficult to identify etiological species of trypanosome in Mongolia due to the lack of species-specific diagnosis marker among subgenus *Trypanozoon*.

Recombinant GM6 antigen based ELISA and ICT were developed for diagnosis of surra in cattle and water buffalo (Nguyen et al., 2015a; Nguyen et al., 2015b; Nguyen et al., 2014). The results using rGM6 based methods were well correlated with the results using OIE recommended crude antigen ELISA. Thus, rTeGM6-4r based ELISA and ICT can be applied in sero-epidemiological studies of animal trypanosomosis in cattle and water buffaloes and can potentially be applied as alternative or supplementary methods of crude antigen ELISA. However, the diagnostic potential of rGM6 based ELISA and ICT were not evaluated for horse trypanosomosis, especially dourine. Therefore, the diagnostic potential of rGM6 based ELISA and ICT was evaluated using horse serum samples which were collected in a horse farm that was a suspected dourine epidemic, in chapter 1. Moreover, whole genome analyses of newly isolated *T. equiperdum* IVM-t1 strain was performed to in order to discover specific gene targets that can be used for development of DNA-based species-specific diagnostic assay for *T. equiperdum* infections and dourine in the future as reported in chapter 2.

The highly conserved amino acid sequence of GM6 and cross-reaction of anti-TeGM6 antigens were suggested that they can result in good diagnosis for *T. equiperdum* and dourine. The GM6 based serological assays showed high sensitivity (ELISA: 81%, ICT: 51%), specificity (79%, 93%) and good correlation (0.6, 0.53) with the results of crude antigen ELISA as reference test using horse samples from Mongolia. The horses were highly suspected dourine cases from PCR and microscopic observations using blood and genital organ swab. Therefore, these results suggested that rTeGM6-4r based methods can be applied for non-tsetse transmitted horse trypanosomosis, especially dourine and *T. equiperdum*. Moreover, rTeGM6-4r-based ELISA was more suitable for epidemiological surveillance in epidemic countries due to the higher sensitivity than ICT and easy handling for many numbers of samples. On the other hand, rGM6 based ICT

showed potential to be suitable for on-site diagnosis by local veterinarians due to the easy and time saving confirmation of the result.

Whole genome analyses were performed using *T. equiperdum* IVM-t1 by integration of two different types of NGS data. The draft genome of *T. equiperdum* IVM-t1 show the smaller numbers of total contig (IVM-t1: 45, OVI: 2,026), longer of largest contig (1.7 Mb, 0.37 Mb), larger N50 (859,849, 38,149), smaller L50 (12, 180) and smaller numbers of gaps per 100 kbp (0, 437.4) than previously published the draft genome of *T. equiperdum* OVI. These results clearly showed that constructed draft genome of *T. equiperdum* IVM-t1 in this study had higher quality than previously published genome of *T. equiperdum* OVI strain. Moreover, the results of gene prediction revealed that there are some *T. equiperdum* specific genes which have not yet been identified. It has been thought that *T. equiperdum* and *T. evansi* have evolved from common *T. brucei* type trypanosome. The smaller numbers of *T. equiperdum* IVM-t1 specific genes than *T. brucei* TREU927 specific genes suggested that some of the genes lost their functionality during evolution. The same phenomenon was also observed in the comparative analysis between *T. evansi* and *T. brucei* (Carnes et al., 2015). Moreover, these species-specific genes might be related to the species-specific parasitism. These predicted genes were only estimated only by genome sequence from statistical properties of the protein-coding sequence itself, therefore, I have to improve the gene prediction based on transcriptome data through RNA sequencing. Transcriptome based gene prediction will reveal the functional genes of *T. equiperdum in vivo* and *in vitro*. In future, I have to compare the genome of *T. equiperdum* IVM-t1 with the genome of other *Trypanozoon* (*T. brucei, T. evansi*, and other *T. equiperdum* strains) for identification of not only *T. equiperdum* but also other trypanosome species-specific DNA marker for development of diagnostic assays. The comparative genomics will hopefully give new insight of species classification and evolution among *Trypanozoon*.

General summary in English

Livestock industry is a major economic division in many developing countries. Mongolian livestock sector contributes 90% of the total agricultural production, which accounts for 11% of total GDP in Mongolia. Horse population is over 2 million heads, and horse are considered as the most important domestic animal with high economic value due to their central role in nomadic pastoralism as well as source of meat and dairy product in Mongolia. Non-tsetse transmitted horse trypanosomosis (dourine and surra), which are important infectious protozoan diseases, are distributed in many countries including Mongolia. Dourine a lethal animal trypanosomosis in equidae which is caused by Trypanosoma equiperdum infection. Many trypanosome are transmitted by blood sucking insects, while, T. equiperdum is transmitted from infected horse to healthy horse via coitus. Because of this unique infectious strategy of T. equiperdum, the disease can easily transmit by a stallion to all of mares in herd. Therefore, dourine is a disease of great economic importance for horse production sector. Effective diagnostic methods for dourine to control the disease is highly expected, however, none of the effective and field- friendly diagnostic methods are developed until now. Additionally, there are no practical diagnosis markers for distinguishing between dourine and surra due to the lack of genome information of T. equiperdum. The diagnosis marker based on species specific nucleotide sequences and/or genes may be found by the whole genome comparison between T. evansi and T. equiperdum.

For these reasons, in this study, I aimed that evaluation of T. evansi GM6 (rTeGM6-4r)-based diagnositic methods (ELISA and immunochromatographic test (ICT)) for the sero-diagnosis and sero-epidemiological study against dourine in chapter 1. Furthermore, in order to develop novel diagnosis tools based on species-specific nucleotide sequences of T. equiperdum in future, the construction of draft genome of T. equiperdum IVM-t1 strain and comparative genomics were performed in chapter 2.

In chapter 1, I evaluated the potential diagnostic abilities of rTeGM6-4r based diagnostic methods for dourine using Mongolian horse samples. This recombinant protein has already shown good diagnostic value in surra in other livestock species. First, genetic diversities of GM6 antigen were determined by sequencing among subgenus Trypanozoon. Ninety-seven percent amino acid sequence homology of GM6 was found among subgenus Trypanozoon. Moreover, anti-rTeGM6-4r antibody universally recognized GM6 antigens in Trypanozoon trypanosomes in indirect fluorescence antibody test (IFAT). These findings suggested that the utilities of rTeGM6-4r based sero-diagnostic methods for both surra and dourine. Next, I examined 50 blood

samples collected from a trypanosomosis outbreak-suspected horse farm in Mongolia. In addition to the blood samples, genital organ swabs were collected from three selected horses with clinical signs of dourine. The diagnostic values of an rTeGM6-4r based ELISA or ICT were measured in comparison to the result of a T. evansi crude antigen based ELISA as a reference test, which is recommended by OIE terrestrial manual. The positive serum samples were detected in 46%, 42%, and 28% of the tested horses, using the rTeGM6-4r based ELISA, crude antigen-based ELISA, and rTeGM6-4r based ICT, respectively. Moreover, the sensitivity and specificity of rTeGM6-based ELISA were 81% and 79%, respectively. The kappa value between reference test and rTeGM6-based ELISA was 0.6, which was considered as moderate. While, the sensitivity and specificity of rTeGM6-4r based ICT were 57% and 93%, respectively. The kappa value between reference test and rTeGM6-4r based ICT was 0.53. Additionally, the motile trypanosomes were found by microscopic observation and the trypanosome specific band was shown by PCR using the samples collected from genital organs with clinical symptoms. On the other hand, none of the PCR positive sample were found using DNA extracted from whole blood. These results highly indicated that the horses were infected in dourine, but not infected in surra. In general, the rTeGM6-4r based ELISA and ICT represented as useful diagnostic options for non-tsetse transmitted horse trypanosomosis, especially for dourine.

In chapter 2, I focused on sequencing and processing of whole genome DNA of newly isolated T. equiperdum IVM-t1 strain. First, total genomic DNA was extracted from culture adapted T. equiperdum IVM-t1 strain and sequenced by two different types of next generation sequencers, MiSeq and PacBio. Genomic sequence data, which generated from MiSeq, had a high degree of accuracy of nucleotide sequencing although it had shorter reads. On the other hand, PacBio generated a longer sequence reads but less accuracy of nucleotide sequencing. By assembling the raw data from two next generation sequencers by multiple softwares, I finally constructed the draft genome of T. equiperdum IVM-t1. Assembly quality and contig numbers of the draft genome of T. equiperdum IVM-t1 were significantly improved in comparison with previously published the genome of T. equiperdum OVI. This improved T. equiperdum genome data has long been needed for further studies on T. equiperdum and dourine. Additionally, chromosome construction, gene prediction of T. equiperdum IVM-t1 were performed using the published genome of T. brucei TERU927 as reference by Companion pipeline. As a result of comparative genomics analyses between T. equiperdum IVM-t1 and T. brucei TERU927, 11 chromosomes and around 8,000 gene candidates including around 600 T. equiperdum IVM-t1-specific genes were predicted. Moreover, these results suggest that species-specific genes were related to

species-specific parasitism. Comparative genomics among trypanosomes is strongly expected for discovery of species-specific genes and identification of their function in further studies. At this stage of the study, whole genome draft assembly and intrinsic gene prediction produced by the current study provides a resource for future trypanosome genetic studies and has revealed some T. equiperdum specific genes. The phylogenic analysis based on the amino acid sequence of 50 core proteins revealed the close relationship among subgenus Trypanozoon. Furthermore, these results can be extended by transcriptome based more accurate gene annotation and more sophisticated genetic analysis using this available resource. Species-specific diagnostic methods will be developed based on these data resource of whole genome sequencing.

The two chapters indicated that rTeGM6-4r based ELISA and ICT can be applied in sero-epidemiological study and in diagnosis of dourine in horse. And the whole genome data of T. equiperdum IVM-t1 must be an important research resource for the development of new diagnostic method for dourine in the future.

General summary in Japanese

学位論文要旨

媾疫特異的診断技術の確立に向けた研究

畜産業は、多くの発展途上国において重要な経済の一部門である。モンゴル国における畜産業は農業生産部門の 90%を占め、全 GDP の 11%に相当する。モンゴル国においてウマは 200 万頭を大きく超えて飼養され、乳製品等の生産及び使役動物としての役割に加え、その経済価値の高さから非常に重要な家畜である。非ツェツェ媒介性ウマトリパノソーマ病（媾疫及びスーラ病）は、モンゴル国をはじめとする多くの国で感染が認められる重要家畜疾病である。中でも媾疫トリパノソーマ（*Trypanosoma equiperdum*）感染によって発症する媾疫はウマ属に特異的であり、感染ウマは致死的な経過をたどる。さらに媾疫トリパノソーマは他種トリパノソーマと異なり交尾感染により感染環が成立するため、種ウマの感染が飼養する馬群全てに感染を拡大し大きな経済被害を与える。しかしながら、媾疫に対する正確で簡便な診断法は未だに確立されていない。さらに媾疫トリパノソーマのゲノム情報が整備されていないため、媾疫トリパノソーマ特異的塩基配列をターゲットとした媾疫特異的診断法の開発も未開発である。よって本研究では、これまでに他種家畜のスーラ病診断技術としてすでに確立されている遺伝子組換え *T. evansi* GM6 タンパク質を用いた Enzyme-linked immunosorbent assay（ELISA）法及び Immunochromatographic test (ICT) 法を、媾疫の診断に適応することを目的とし

て研究を実施した（第 1 章)。さらに将来的により高感度・高精度な媾疫特異的診断法を開発するための情報基盤を整備するため、モンゴル国で新たに分離した媾疫トリパノソーマの全ゲノム配列情報を解読し、ゲノム情報を整備することを目的として研究を実施した（第 2 章)。

第 1 章ではモンゴル国のウマを用いて、媾疫に対する既存の診断法の媾疫診断に対する有用性を検証した。まず他種家畜のスーラ病の診断用抗原としてこれまでに確立されている GM6 タンパク質の遺伝的多様性を検証した。その結果、媾疫トリパノソーマを含む *Trypanozoon* 亜属のトリパノソーマ間で高度に保存されていることが明らかとなった。さらに媾疫トリパノソーマの GM6 抗原に対する抗体の反応性を間接蛍光抗体法で検証した結果、他種トリパノソーマの GM6 と免疫交差性が認められた。以上の結果より、GM6 抗原を用いた診断法は媾疫の診断においても適用可能であることが示唆された。そこで、GM6 抗原を用いた診断法を用いて媾疫の診断が可能であるか、媾疫の集団発生が疑われた農場で飼養されているウマ 50 頭を用いて検討した。rTeGM6-4r ELISA 法及び rTeGM6-4r ICT 法を用いて診断した結果を、OIE 推奨診断法である *T. evansi* Crude Antigen ELISA 法による診断結果とそれぞれ比較解析したところ、rTeGM6-4r ELISA 法と ICT 法の媾疫診断の感度・特異性はそれぞれ、81%・79%及び 57%・93%であることが明らかとなった。さらに診断結果の一致度は、rTeGM6 ELISA で 0.60、rTeGM6 ICT で 0.53 であることが示された。一方で、本農場のウマ血液検体を用いた PCR では 50 頭全てが陰性であったにもかかわらず、一部のウマ生殖器検体を用いた顕微鏡検査及び PCR ではトリパノソーマ陽性検体が見出された。この結果から、本農場では媾疫が蔓延している可能性が強く示唆された。以上のことから rTeGM6-4r 抗原を用いた

診断法はウマトリパノソーマ病、特に媾疫に対して有用であることが明らかとなった。

第 2 章では、モンゴル国において新たに分離した媾疫トリパノソーマ（*T. equiperdum* IVM-t1 株）の全ゲノム配列情報を解読し、同原虫のゲノム情報を整備することを目的として研究を実施した。まず培養馴化した *T. equiperdum* IVM-t1 株から精製された総 DNA を、2 種類の次世代シーケンサー（Next Generation Sequencer: NGS）Miseq 及び PacBio で解読した。続いて NGS で解読された多数の短いフラグメントを複数のゲノム解析ソフトウェアを用いて解析することで、少数の長い塩基配列であるコンティグに集約した。その結果、既存の *T. equiperdum* OVI 株の全ゲノム解析結果に比べて、塩基配列読み取り精度及び各コンティグ長が大幅に改善されたドラフトゲノムの構築に成功した。さらに近縁種である *T. brucei* のゲノム情報を参照とし、得られたドラフトゲノムを各染色体に再構築するとともに、*T. equiperdum* IVM-t1 の遺伝子を推定した。その結果、11 本の染色体と約 8,000 の遺伝子配列が推計された。また *T. brucei* と *T. equiperdum* IVM-t1 の比較ゲノム解析の結果、大部分の遺伝子は両者に共通であったが、*T. equiperdum* IVM-t1 特異的な約 600 の遺伝子群が特定された。以上の結果から、本研究で得られた *T. equiperdum* IVM-t1 のゲノム情報は、既存の媾疫トリパノソーマゲノム解析を大きく上回る精度とコンティグ長を有することから、塩基配列多様性を基盤とした媾疫の新規診断法開発のための基礎的知見として十分な成果であると言える。また *T. equiperdum* 特異的遺伝子群を対象とした媾疫特異的な遺伝子診断マーカーの開発が期待される。また今回の解析で想定された T. equiperdum 特異的遺伝子群は、交尾感染や生殖器粘膜への寄生等の *T. equiperdum* に特徴的な寄生戦略に関与してい

る可能性がある。さらに保存性の高い 50 遺伝子のアミノ酸配列を用いてキネトプラスト綱（リーシュマニア属及びトリパノソーマ属）の系統関係を解析した結果、*T. equiperdum* IVM-t1 株は他種 *Trypanozoon* 亜属トリパノソーマと同一のクレードを形成した。今後、遺伝子トランスクリプトーム解析を実施することで正確な遺伝子発現配列情報並びに遺伝子発現情報を得ることができ、より高精度な媾疫トリパノソーマ遺伝子情報の構築とそれを利用した媾疫特異的診断法の開発が期待される。

以上より、本研究において媾疫に対する rGM6-4r を診断抗原として用いた診断法の適応が確認された。さらに将来的な新規媾疫診断法開発の礎となる媾疫トリパノソーマ全ゲノム情報が得られた。今後、本研究成果をもとにした効果的なウマトリパノソーマ病対策の実施が期待される。

Acknowledgements

I am thankful to the Graduate School of Animal Husbandry, Obihiro University of Agriculture and Veterinary Medicine (OUAVM) where many international students have had opportunity to learn and gain knowledge and experience at high level of scientific approach in the doctoral program of Animal and Food Hygiene. I am special thanks for the Research Unit for Vector Biology, National Research Center for the Protozoan Disease (NRCPD), for hosting my PhD studies. Special thanks to the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and Science and Technology Research Partnership for Sustainable Development (SATREPS) which financially supported my study.

This thesis would have not been completed without great help and support from my supervisors, colleagues, friends and family.

First and foremost, I would like to express my deepest gratitude to Prof. Noboru Inoue for allowing me as his student, for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge and to Asst. Prof. Keisuke Suganuma who has kindly given me countless support and has patiently taught me important lessons from laboratory techniques to scientific writing.

I also wish to express my gratitude to my supervisor Prof. Xuenan Xuan for his supervision and guidance. I am thankful to Prof. Naoaki Yokoyama for his valuable and constructive supervision.

I also owe my great debt of gratitude Dr. Thu Thuy Nguyen, for introducing me to all the detail of her work and helping me to start my study.

My sincere thanks also goes to Assoc. Prof. Junya Yamagishi, who provided me an opportunity to join his team as internship fellow, and for giving me access to the research facilities and for teaching me bioinformatics. Without their precious support it would not be possible to conduct this research.

I am extremely thankful to Simon Peter Musinguzi, Shino Yamasaki, Ehab Mossad and Oriel Thekisoe for their help in editing my writings and for their kind guidance and encouragement.

I would like to thank my laboratory colleagues Trinh Minh Anh, Adrian Miki C. Macalanda, Molefe Nthatisi Innocentia, Jose Ma. M. Angeles, Karleezelle J. Moendeg and Daiki Mizushima.

I would like to thank Prof. Shin-Ichiro Kawazu and all members of the Research Unit for Advanced Preventive Medicine.

I would like to thank Assoc. Prof. Shinya Fukumoto and members of the Research Unit for Vector Biology.

I would like to thank Prof. Ikou Igarashi, all members of Research Unit for Molecular Diagnosis.

I greatly appreciate my dearest friends: Aemiro Kehaliew, Ragab Makhlouf Mahmoud Fereig, Takahito Hattori, Danushi Udayanga, Ms. Maki Ishiwata, Ros Burbidge, Akira Soga and many other friends in Obihiro for their friendship and beneficial discussions.

Special thanks to Tsogtbaatar Tugsbaatar for kindly and generously helping me in my study and ordinary life in Obihiro.

Special thanks to Ms. Chikako Hashimoto, Ms. Yayoko Onishi and Ms. Kobayashi Shiho for their technical assistance related to Japanese language documents and life in Obihiro.

I would like to express my sincere gratitude to the Institute of Veterinary Medicine of Mongolia (IVM) and SATREPS project members for allowing me to study in Japan.

I also thank all my colleagues it IVM, especially Dr. Badgar Battsetseg, Dr. Banzragch Battor Dr. Punsantsogvoo Myagmarsuren, Mr. Otgonsuren Davaajav, Mr. Sandagdorj Narantsatsral and Mr. Tovuu Amgalanbaatar.

I am indebted to my parents for their immense support and persistent encouragement. I am grateful to my wife Natsagdorj Oyungerel for always understanding my work and giving me necessary support.

And finally, to everyone who influenced me positively in one way or the other ...

Thanks for all your encouragement!

References

ADFCA 2017. Trypanosoma evansi infections (surra) (www.adfca.ae/english Abu Dhabi Food Control Authority).

AHT, B., DEFRA 2011. Dourine-an emerging veneral threat to European horses (www.aht.org.uk AHT or Animal Health Trust), pp. 15-18.

Altangerel, K., Sivakumar, T., Battsetseg, B., Battur, B., Ueno, A., Igarashi, I., Yokoyama, N., 2012. Phylogenetic relationships of Mongolian Babesia bovis isolates based on the merozoite surface antigen (MSA)-1, MSA-2b, and MSA-2c genes. Vet Parasitol 184, 309-316.

Antoine-Moussiaux, N., Desmecht, D., 2008. Epidémiologie de l'infection par Trypanosoma evansi. Ann. Med. Vet. Annales de Medecine Veterinaire 152, 191-201.

Berberof, M., Perez-Morga, D., Pays, E., 2001. A receptor-like flagellar pocket glycoprotein specific to Trypanosoma brucei gambiense. Molecular and biochemical parasitology 113, 127-138.

Bern, C., Kjos, S., Yabsley, M.J., Montgomery, S.P., 2011. Trypanosoma cruzi and Chagas' Disease in the United States. Clinical microbiology reviews 24, 655-681.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., Bohme, U., Hannick, L., Aslett, M.A., Shallom, J., Marcello, L., Hou, L., Wickstead, B., Alsmark, U.C., Arrowsmith, C., Atkin, R.J., Barron, A.J., Bringaud, F., Brooks, K., Carrington, M., Cherevach, I., Chillingworth, T.J., Churcher, C., Clark, L.N., Corton, C.H., Cronin, A., Davies, R.M., Doggett, J., Djikeng, A., Feldblyum, T., Field, M.C., Fraser, A., Goodhead, I., Hance, Z., Harper, D., Harris, B.R., Hauser, H., Hostetler, J., Ivens, A., Jagels, K., Johnson, D., Johnson, J., Jones, K., Kerhornou, A.X., Koo, H., Larke, N., Landfear, S., Larkin, C., Leech, V., Line, A., Lord, A., Macleod, A., Mooney, P.J., Moule, S., Martin, D.M., Morgan, G.W., Mungall, K., Norbertczak, H., Ormond, D., Pai, G., Peacock, C.S., Peterson, J., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Reitter, C., Salzberg, S.L., Sanders, M., Schobel, S., Sharp, S., Simmonds, M., Simpson, A.J., Tallon, L., Turner, C.M., Tait, A., Tivey, A.R., Van Aken, S., Walker, D., Wanless, D., Wang, S., White, B., White, O., Whitehead, S., Woodward, J., Wortman, J., Adams, M.D., Embley, T.M., Gull, K., Ullu, E., Barry, J.D., Fairlamb, A.H., Opperdoes, F., Barrell, B.G., Donelson, J.E., Hall, N., Fraser, C.M., Melville, S.E., El-Sayed, N.M., 2005. The genome of the African trypanosome Trypanosoma brucei. Science 309, 416-422.

Bishop, S., Rae, P.F., Phipps, L.P., Boid, R., Luckins, A.G., 1995. Trypanosoma equiperdum: detection of trypanosomal antibodies and antigen by enzyme-linked immunosorbent assay. The British veterinary journal 151, 715-720.

Borst, P., Fase-Fowler, F., Gibson, W.C., 1987. Kinetoplast DNA of Trypanosoma evansi. Molecular and biochemical parasitology 23, 31-38.

Boulange, A., Pillay, D., Chevtzoff, C., Biteau, N., Come de Graca, V., Rempeters, L., Theodoridis, D., Baltz, T., 2017. Development of a rapid antibody test for point-of-care diagnosis of animal African trypanosomosis. Vet Parasitol 233, 32-38.

Brun, R., Hecker, H., Lun, Z.R., 1998. Trypanosoma evansi and T. equiperdum: distribution, biology, treatment and phylogenetic relationship (a review). Vet Parasitol 79, 95-107.

Buffard, D.G.S.M., 1900. Le trypanosome de la dourine Archives de Parasitologie 3, 124-133.

Buscher, P., Mumba Ngoyi, D., Kabore, J., Lejon, V., Robays, J., Jamonneau, V., Bebronne, N., Van der Veken, W., Bieler, S., 2009. Improved Models of Mini Anion Exchange Centrifugation Technique (mAECT) and Modified Single Centrifugation (MSC) for sleeping sickness diagnosis and staging. PLoS neglected tropical diseases 3, e471.

Carnes, J., Anupama, A., Balmer, O., Jackson, A., Lewis, M., Brown, R., Cestari, I., Desquesnes, M., Gendrin, C., Hertz-Fowler, C., Imamura, H., Ivens, A., Koreny, L., Lai, D.H., MacLeod, A., McDermott, S.M., Merritt, C., Monnerat, S., Moon, W., Myler, P., Phan, I., Ramasamy, G.,

Sivam, D., Lun, Z.R., Lukes, J., Stuart, K., Schnaufer, A., 2015. Genome and phylogenetic analyses of Trypanosoma evansi reveal extensive similarity to T. brucei and multiple independent origins for dyskinetoplasty. PLoS neglected tropical diseases 9, e3404.

CDC, C.f.D.C.a.P. 2017. Parasites -African Trypanosomiasis (also known as Sleeping Sickness).

Chamond, N., Cosson, A., Blom-Potar, M.C., Jouvion, G., D'Archivio, S., Medina, M., Droin-Bergere, S., Huerre, M., Goyard, S., Minoprio, P., 2010. Trypanosoma vivax infections: pushing ahead with mouse models for the study of Nagana. I. Parasitological, hematological and pathological parameters. PLoS neglected tropical diseases 4, e792.

Chiara, M., Pavesi, G., 2017. Evaluation of Quality Assessment Protocols for High Throughput Genome Resequencing Data. Frontiers in genetics 8, 94.

Choi, J., El-Sayed, N.M., 2012. Functional genomics of trypanosomatids. Parasite immunology 34, 72-79.

Claes, F., Buscher, P., Touratier, L., Goddeeris, B.M., 2005. Trypanosoma equiperdum: master of disguise or historical mistake? Trends Parasitol 21, 316-321.

Claes, F., C Agbo, E., Radwanska, M., Pas, M., Baltz, T., de Waal, T., Goddeeris, B., Claassen, E., Büscher, P., 2003a. How does Trypanosoma equiperdum fit into the Trypanozoon group? A cluster analysis by RAPD and Multiplex-endonuclease genotyping approach, Vol 126, 425-431 pp.

Claes, F., Verloo, D., De Waal, D.T., Majiwa, P.A., Baltz, T., Goddeeris, B.M., Buscher, P., 2003b. The expression of RoTat 1.2 variable surface glycoprotein (VSG) in Trypanosoma evansi and T. equiperdum. Veterinary parasitology 116, 209-216.

Clausen, P.H., Chuluun, S., Sodnomdarjaa, R., Greiner, M., Noeckler, K., Staak, C., Zessin, K.H., Schein, E., 2003. A field study to estimate the prevalence of Trypanosoma equiperdum in Mongolian horses. Vet Parasitol 115, 9-18.

Clement, S.L., Mingler, M.K., Koslowsky, D.J., 2004. An intragenic guide RNA location suggests a complex mechanism for mitochondrial gene expression in Trypanosoma brucei. Eukaryotic cell 3, 862-869.

Clopper, C.J., Pearson, E.S., 1934. The use of confidence or fiducial limits illustrated in the case of the binomial Biometrika 26, 404-413.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement 20, 37-46.

Cuypers, B., Van den Broeck, F., Van Reet, N., Meehan, C.J., Cauchard, J., Wilkes, J.M., Claes, F., Goddeeris, B., Birhanu, H., Dujardin, J.C., Laukens, K., Buscher, P., Deborggraeve, S., 2017. Genome-Wide SNP Analysis Reveals Distinct Origins of Trypanosoma evansi and Trypanosoma equiperdum. Genome biology and evolution 9, 1990-1997.

d'Avila-Levy, C.M., Boucinha, C., Kostygov, A., Santos, H.L., Morelli, K.A., Grybchuk-Ieremenko, A., Duval, L., Votypka, J., Yurchenko, V., Grellier, P., Lukes, J., 2015. Exploring the environmental diversity of kinetoplastid flagellates in the high-throughput DNA sequencing era. Memorias do Instituto Oswaldo Cruz 110, 956-965.

Desquesnes, M., 1997. [International and regional standardization of immunoenzyme tests: methods, concerns and limitations]. Rev Sci Tech 16, 809-823.

Desquesnes, M., Dargantes, A., Lai, D.H., Lun, Z.R., Holzmuller, P., Jittapalapong, S., 2013. Trypanosoma evansi and surra: a review and perspectives on transmission, epidemiology and control, impact, and zoonotic aspects. Biomed Res Int 2013, 321237.

Doflein, F., 1901. Die Protozoen als Parasiten und Krankheitserreger nach biologischen Gesichtspunkten dargestellt. Fischer, Jena.

FAO 2016. FAOSTAT (FAO).

Fitch, W.M., Margoliash, E., 1967. Construction of phylogenetic trees. Science 155, 279-284.

Gari, F.R., Ashenafi, H., Tola, A., Goddeeris, B.M., Claes, F., 2010. Comparative diagnosis of parasitological, serological, and molecular tests in dourine-suspected horses. Tropical animal health and production 42, 1649-1654.
Gizaw, Y., Megersa, M., Fayera, T., 2017. Dourine: a neglected disease of equids. Tropical animal health and production 49, 887-897.
Goodland, A., Sheehy, D., Shine, T., 2009. Mongolia Livestock Sector Study, Volume I–Synthesis Report. Sustainable Development Department, East Asia and Pacific Region, World Bank, Washington, DC.
Gungaa, T. 1990. Research report of countermeasurement against surra based on study result among Mongolia bacterian camel (State Central Veterinary Laboratory), pp. 57-59.
Gunn, A., Pitt, S.J., 2012. Parasitology : an integrated approach. John Wiley & Sons, Chichester, West Sussex.
Gutierrez, C., Desquesnes, M., Touratier, L., Buscher, P., 2010. Trypanosoma evansi: recent outbreaks in Europe. Vet Parasitol 174, 26-29.
Hagos, A., Abebe, G., Buscher, P., Goddeeris, B.M., Claes, F., 2010a. Serological and parasitological survey of dourine in the Arsi-Bale highlands of Ethiopia. Tropical animal health and production 42, 769-776.
Hagos, A., Goddeeris, B.M., Yilkal, K., Alemu, T., Fikru, R., Yacob, H.T., Feseha, G., Claes, F., 2010b. Efficacy of Cymelarsan and Diminasan against Trypanosoma equiperdum infections in mice and horses. Vet Parasitol 171, 200-206.
Hall, T., 1999. BIOEDIT: A USER-FRIENDLY BIOLOGICAL SEQUENCE ALIGNMENT EDITOR AND ANALYSIS PROGRAM FOR WINDOWS 95/98/ NT. Nucleic Acids Symposium Series 41, 95-98.
Hayes, P., Varga, V., Olego-Fernandez, S., Sunter, J., Ginger, M.L., Gull, K., 2014. Modulation of a cytoskeletal calpain-like protein induces major transitions in trypanosome morphology. J Cell Biol 206, 377-384.
Hebert, L., Moumen, B., Madeline, A., Steinbiss, S., Lakhdar, L., Van Reet, N., Buscher, P., Laugier, C., Cauchard, J., Petry, S., 2017. First Draft Genome Sequence of the Dourine Causative Agent: Trypanosoma Equiperdum Strain OVI. Journal of genomics 5, 1-3.
Hoare, C.A., 1972. The trypanosomes of mammals : a zoological monograph. Blackwell Scientific Publications, Oxford; Edinburgh.
Hund, A. 2008. The Stallion's Mane: The Next Generation of Horses in Mongolia (SIT Digital Collections Donald B. Watt library, Independent Study Project (ISP) Collection).
J. Stevens, S.B., 2004. The systematics of trypanosomes of medical and veterinary importance, In: Maudlin, I., Holmes, P.H., Miles, M.A. (Eds.) The trypanosomiases. CABI, Wallingford.
Jackson, A.P., Sanders, M., Berry, A., McQuillan, J., Aslett, M.A., Quail, M.A., Chukualim, B., Capewell, P., MacLeod, A., Melville, S.E., Gibson, W., Barry, J.D., Berriman, M., Hertz-Fowler, C., 2010. The genome sequence of Trypanosoma brucei gambiense, causative agent of chronic human african trypanosomiasis. PLoS neglected tropical diseases 4, e658.
Jensen, R.E., Simpson, L., Englund, P.T., 2008. What happens when Trypanosoma brucei leaves Africa. Trends Parasitol 24, 428-431.
Johnson, B.J., Hill, G.C., Fox, T.D., Stuart, K., 1982. The maxicircle of Trypanosoma brucei kinetoplast DNA hybridizes with a mitochondrial gene encoding cytochrome oxidase subunit II. Molecular and biochemical parasitology 5, 381-390.
Katz, J.B., Chieves, L.P., Hennager, S.G., Nicholson, J.M., Fisher, T.A., Byers, P.E., 1999. Serodiagnosis of equine piroplasmosis, dourine, and glanders using an arrayed immunoblotting method. Journal of veterinary diagnostic investigation : official publication of the American Association of Veterinary Laboratory Diagnosticians, Inc 11, 292-294.
Kocher, A., Desquesnes, M., Kamyingkird, K., Yangtara, S., Leboucher, E., Rodtian, P., Dargantes, A., Jittapalapong, S., 2015. Evaluation of an Indirect-ELISA Test for Trypanosoma evansi

Infection (Surra) in Buffaloes and Its Application to a Serological Survey in Thailand. Biomed Res Int 2015, 361037.
Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biology 10, R25.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.
Luckins, A.G., 1988. Trypanosoma evansi in Asia. Parasitol Today 4, 137-142.
Magez, S., Radwanska, M., 2014. Trypanosomes and trypanosomiasis. Springer, Wien, x, 294 pages pp.
Masiga, D.K., Smyth, A.J., Hayes, P., Bromidge, T.J., Gibson, W.C., 1992. Sensitive detection of trypanosomes in tsetse flies by DNA amplification. Int J Parasitol 22, 909-918.
MedCalc 2017. Free statistical calculators. In Diagnostic test evaluation calculator

(Acacialaan 22, 8400 Ostend, Belgium, MedCalc Software bvba(BE)).
Mehlhorn, H., Aspöck, H., 2008. Encyclopedia of parasitology.
Moreno, S.A., Concepcion, J.L., Nava, M., Molinari, J., 2013. Importance of the horse and financial impact of equine trypanosomiasis on cattle raising in Venezuela. Tropical animal health and production 45, 1669-1676.
Muggli, M.D., Puglisi, S.J., Ronen, R., Boucher, C., 2015. Misassembly detection using paired-end sequence reads and optical mapping data. Bioinformatics 31, i80-88.
Müller, N., Hemphill, A., Imboden, M., Duvallet, G., Dwinger, R.H., Seebeck, T., 1992. Identification and characterization of two repetitive non-variable antigens from African trypanosomes which are recognized early during infection. Parasitology 104 Pt 1, 111-120.
Musinguzi, S.P., 2017. Molecular epidemiological studies on trypanosomosis and piroplasmosis among livestock in southern Africa and central Asia Doctoral Dissertation Obihiro University of Agriculture and Veterinary Medicine
Ngaira, J.M., Olembo, N.K., Njagi, E.N., Ngeranwa, J.J., 2005. The detection of non-RoTat 1.2 Trypanosoma evansi. Experimental parasitology 110, 30-38.
Nguyen, T.T., Motsiri, M.S., Taioe, M.O., Mtshali, M.S., Goto, Y., Kawazu, S., Thekisoe, O.M., Inoue, N., 2015a. Application of crude and recombinant ELISAs and immunochromatographic test for serodiagnosis of animal trypanosomosis in the Umkhanyakude district of KwaZulu-Natal province, South Africa. J Vet Med Sci 77, 217-220.
Nguyen, T.T., Ruttayaporn, N., Goto, Y., Kawazu, S., Sakurai, T., Inoue, N., 2015b. A TeGM6-4r antigen-based immunochromatographic test (ICT) for animal trypanosomosis. Parasitol Res 114, 4319-4325.
Nguyen, T.T., Zhou, M., Ruttayaporn, N., Nguyen, Q.D., Nguyen, V.K., Goto, Y., Suzuki, Y., Kawazu, S., Inoue, N., 2014. Diagnostic value of the recombinant tandem repeat antigen TeGM6-4r for surra in water buffaloes. Vet Parasitol 201, 18-23.
Nimpaye, H., Njiokou, F., Njine, T., Njitchouang, G.R., Cuny, G., Herder, S., Asonganyi, T., Simo, G., 2011. Trypanosoma vivax, T. congolense "forest type" and T. simiae: prevalence in domestic animals of sleeping sickness foci of Cameroon. Parasite (Paris, France) 18, 171-179.
Njiru, Z.K., Constantine, C.C., Guya, S., Crowther, J., Kiragu, J.M., Thompson, R.C., Davila, A.M., 2005. The use of ITS1 rDNA PCR in detecting pathogenic African trypanosomes. Parasitol Res 95, 186-192.
Njiru, Z.K., Constantine, C.C., Masiga, D.K., Reid, S.A., Thompson, R.C., Gibson, W.C., 2006. Characterization of Trypanosoma evansi type B. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases 6, 292-300.
Nordberg, H., Cantor, M., Dusheyko, S., Hua, S., Poliakov, A., Shabalov, I., Smirnova, T., Grigoriev, I.V., Dubchak, I., 2014. The genome portal of the Department of Energy Joint Genome Institute: 2014 updates. Nucleic acids research 42, D26-31.

NRSO, M. 2015. Mongolian statistical year book Kh.Bajiikhuu, S.M.R.S.A.D., ed. (http://www.1212.mn/en, National Registration and Statistics Office of Mongolia).
OIE 2016. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals 2016. In World Organization for Animal Healt (http://www.oie.int/en/international-standard-setting/terrestrial-manual/access-online/ OIE).
Olani, A., Habtamu, Y., Wegayehu, T., Anberber, M., 2016. Prevalence of camel trypanosomosis (surra) and associated risk factors in Borena zone, southern Ethiopia. Parasitol Res 115, 1141-1147.
Pascucci, I., Di Provvido, A., Camma, C., Di Francesco, G., Calistri, P., Tittarelli, M., Ferri, N., Scacchia, M., Caporale, V., 2013. Diagnosis of dourine in outbreaks in Italy. Vet Parasitol 193, 30-38.
Patel, R.K., Jain, M., 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS One 7, e30619.
Peacock, L., Cook, S., Ferris, V., Bailey, M., Gibson, W., 2012. The life cycle of Trypanosoma (Nannomonas) congolense in the tsetse fly. Parasit Vectors 5, 109.
Pillay, D., Izotte, J., Fikru, R., Buscher, P., Mucache, H., Neves, L., Boulange, A., Seck, M.T., Bouyer, J., Napier, G.B., Chevtzoff, C., Coustou, V., Baltz, T., 2013. Trypanosoma vivax GM6 antigen: a candidate antigen for diagnosis of African animal trypanosomosis in cattle. PLoS One 8, e78565.
Podaliri Vulpiani, M., Carvelli, A., Giansante, D., Iannino, F., Paganico, D., Ferri, N., 2013. Reemergence of Dourine in Italy: Clinical Cases in Some Positive Horses. Journal of Equine Veterinary Science 33, 468-474.
Purevsuren, B., 2004. Prevalence study of Equine Trypanosomosis in Mongolia and development of diagnostic preparate Biology, Veterinary, Parasitology Mongolian State University of Agriculture publishing house of Mongolian State University of Agriculture
Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC genomics 13, 341.
Radwanska, M., Chamekh, M., Vanhamme, L., Claes, F., Magez, S., Magnus, E., de Baetselier, P., Buscher, P., Pays, E., 2002. The serum resistance-associated gene as a diagnostic tool for the detection of Trypanosoma brucei rhodesiense. The American journal of tropical medicine and hygiene 67, 684-690.
Rouget, J., 1896. Contribution a l'étude du trypanosome des mammifere. Ann. Inst. Pasteur 10, 716-728.
Sambrook, J., Russell, D.W., 2001. Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
Sambrook, J., Russell, D.W., 2006. The condensed protocols : from molecular cloning : a laboratory manual. N.Y. : Cold Spring Harbor laboratory Press, Cold Spring Harbor.
Sanchez, E., Perrone, T., Recchimuzzi, G., Cardozo, I., Biteau, N., Aso, P.M., Mijares, A., Baltz, T., Berthier, D., Balzano-Nogueira, L., Gonzatti, M.I., 2015. Molecular characterization and classification of Trypanosoma spp. Venezuelan isolates based on microsatellite markers and kinetoplast maxicircle genes. Parasit Vectors 8, 536.
Schmieder, R., Edwards, R., 2011. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS One 6, e17288.
Silva, R.A.M.S., 1995. Trypanosomosis outbreak due to Trypanosoma evansi in the Pantanal, Brazil a preliminary approach on risk factors. Revue D'Elevage Medicine Veterinaire Pays Tropicaux 48, 315-319.
Simarro, P.P., Cecchi, G., Paone, M., Franco, J.R., Diarra, A., Ruiz, J.A., Fevre, E.M., Courtin, F., Mattioli, R.C., Jannin, J.G., 2010. The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. International journal of health geographics 9, 57.

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J., Birol, I., 2009. ABySS: a parallel assembler for short read sequence data. Genome Res 19, 1117-1123.

Smith, D.R., 2017. Goodbye genome paper, hello genome report: the increasing popularity of 'genome announcements' and their impact on science. Briefings in Functional Genomics 16, 156-162.

Songa, E.B., Paindavoine, P., Wittouck, E., Viseshakul, N., Muldermans, S., Steinert, M., Hamers, R., 1990. Evidence for kinetoplast and nuclear DNA homogeneity in Trypanosoma evansi isolates. Molecular and biochemical parasitology 43, 167-179.

Spengler, S.J., 2000. Bioinformatics in the Information Age. Science 287, 1221-1223.

Steel, J.H., Evans, G., Royal College of Physicians of, L., 1886. Report of veterinary surgeon J. H. Steel, A. V. D., on his investigation into an obscure and fatal disease among transport mules in British Burma, which he found to be a fever of relapsing type, and probably identical with the disorder first described by Dr. Griffith Evans under the name "Surra", in a report (herewith reprinted) published by the Punjab Government, Military Department, No. 439-4467, of 3rd. December 1880-vide the Veterinary Journal (London), 1881-1882. Publisher not identified, Place of publication not identified :.

Suganuma, K., Narantsatsral, S., Battur, B., Yamasaki, S., Otgonsuren, D., Musinguzi, S.P., Davaasuren, B., Battsetseg, B., Inoue, N., 2016. Isolation, cultivation and molecular characterization of a new Trypanosoma equiperdum strain in Mongolia. Parasit Vectors 9, 481.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic acids research 22, 4673-4680.

Thuy, N.T., Goto, Y., Lun, Z.R., Kawazu, S., Inoue, N., 2012. Tandem repeat protein as potential diagnostic antigen for Trypanosoma evansi infection. Parasitol Res 110, 733-739.

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., Earl, A.M., 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One 9, e112963.

Wassall, D.A., Gregory, R.J., Phipps, L.P., 1991. Comparative evaluation of enzyme-linked immunosorbent assay (ELISA) for the serodiagnosis of dourine. Vet Parasitol 39, 233-239.

Wei, Y., Wen, Y.-Z., Desquesnes, M., Lun, Z.-R., 2011. Molecular epidemiology of# trypanosoma evansi# and# T. Equiperdum# and atypical human infection by animal trypanosomes.

Wences, A.H., Schatz, M.C., 2015. Metassembler: merging and optimizing de novo genome assemblies. Genome Biol 16, 207.

Yazdzik, E. 2011. The Mongolian Horse and Horseman (SIT Digital Collections Donald B. Wat Library, Independent Study Project (ISP) Collection).

Zhang, J., Chiodini, R., Badr, A., Zhang, G., 2011. The impact of next-generation sequencing on genomics. Journal of genetics and genomics = Yi chuan xue bao 38, 95-109.