University of South Bohemia in České Budějovice

Faculty of Science

Development of DiCre parental lineage of *Babesia divergens*

Master thesis

Bc. Dominika Reichensdörferová

Supervisor: RNDr. Daniel Sojka, PhD.

Co-supervisor: RNDr. Marie Jalovecká, PhD.

České Budějovice 2022

Reichensdörferová D., (2022). Development of DiCre parental lineage of *Babesia divergens*. Ms. Thesis, in English – 50 p., Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic.

Annotation

Babesia divergens is a tick-transmitted parasite of the order Piroplasmida, an important causative agent of bovine and human babesiosis. Even though *Babesia* represent an important worldwide veterinary threat and an emerging risk to humans, some of the fundamental principles of their parasitic lifestyle are still poorly understood. One of the approaches, which help us to better understand the biology of *Babesia* parasites and their interaction with hosts or vectors, is functional genomics. Hence, the aim of this thesis is to establish a DiCre parental lineage of *B. divergens* that may provide the absent genome editing technique of conditional gene knock-out. Such a tool will help to functionally characterize essential genes/proteins and validate them as novel targets for the yet missing specific treatment. We have selected and identified an eligible recipient locus in the *B. divergens* genome into which a DiCre cassette will be integrated via homology recombination. We have validated and determined the optimal concentration of WR99210 and G-418 for their use as selection markers for transfected *B. divergens* cultures. Finally, we have designed and constructed the plasmid vector containing the DiCre cassette for the generation of DiCre parental *B. divergens* lineage that may be universally applicable to functionally validate any genes/proteins of interest.

I hereby declare that I am the author of this qualification thesis and that I have prepared it using only the sources and literature cited in the list of references.

Date……...

Signature……

Financial support

This work was supported by the Mobility plus project JSPS-21-08, the Czech National Grant Agency (GAČR) standard project 21-11299S and the ERDF/ESF Centre for research of pathogenicity and virulence of parasites (No.CZ.02.1.01/0.0/0.0/16_019/0000759).

Acknowledgments

My first thanks goes to Dr. Kopáček himself for accepting me as a rightful member of the team and allowing me to participate in the research in his laboratory. I would like to express my sincere gratitude to my supervisor Dr. Sojka and co-supervisor Dr. Jalovecká, who taught me a fair portion of my research skills, explained to me everything I sought to know, and guided me through the process of experiments as well as writing this thesis. My last, but not least thanks are to my parents, my family and my friends because without them it would be impossible.

Table of contents

A. INTRODUCTION

1. *Babesia* genus

Babesia spp. are tick-transmitted unicellular organisms, which multiply exclusively inside red blood cells of various vertebrates including humans. The parasite is of the order Piroplasmida of the phylum Apicomplexa which also includes widely known parasitic genera such as *Plasmodium*, *Cryptosporidium*, and *Toxoplasma* genera (White & Suvorova, 2018).

One of the characteristic attributes of Apicomplexa is the apical complex – a set of unique organelles found in all invasive stages of the parasite (Sam-Yellowe, 1996). The apical complex, located in the anterior pole of their cell, consists of several components (Votýpka et al., 2016). Skeletal components include conoid, polar rings, and microtubular protrusions, all of them characterized as non-secretory organelles (del Carmen Terrón et al., 2016). The complex also comprises of secretory components – rhoptries, dense granules, and micronemes specifically (Soldati et al., 2004). These organelles are essential for various processes, including motility of the parasite, adhesion to the host cell, subsequent invasion, development of a parasitophorous vacuole, and egress out of the host cell (Soldati et al., 2004). In contrast to other Apicomplexa, species within the genus *Babesia* have considerably reduced apical complex. The main distinctions are the absence of a conoid, dense granules, polar rings, and subpellicular microtubules, associated with the non-development of the parasitophorous vacuole (Votýpka et al., 2016). The apical complex is compiled from a single large rhoptry, a few micronemes, and a spherical body, that present an organelle equivalent to dense granules (Jalovecká, et al., 2018).

The taxonomy of piroplasms is established by morphological and developmental hallmarks and multi-gene analysis, specifically the molecular analysis of the 18S rRNA unit (Jalovecká et al., 2019). The genus *Babesia* contains more than 100 species found worldwide since their presence is reliant on their hosts, mainly ticks and vertebrates (Schnittger et al., 2012). *Babesia* infects a broad spectrum of wild and domesticated animals, especially livestock. Humans are considered as accidental hosts even though the incidence of patients confirmed with the disease babesiosis is growing around the globe (Yabsley & Shock, 2013). Even though *Babesia* parasites are not studied as intensively as their closely related plasmodium and toxoplasma (Jalovecká et al., 2019), they present great economic, veterinary, and medical impact (Gohil et al., 2013).

The lifecycles of *Babesia* and *Plasmodium* genera share some key features, such as asexual reproduction in host erythrocytes and sexual reproduction in the gut lumen of the definitive host of phylum Arthropoda (Smith et al., 2002). In contrast to *Plasmodium* spp. transmitted via mosquitos of genus *Anopheles*, *Babesia* spp. apply transmission strategy via hard ticks of genus *Ixodes*. *Babesia* spp. also differs from *Plasmodium* spp. in the lack of synchronized life cycle in host erythrocytes, which means that different phases of development occur in the bloodstream simultaneously. In-depth knowledge of *Babesia* parasitic lifestyle strategies may provide tools for the identification of suitable targets for the development of specific drugs. Given the relatively short evolutionary distance between *Plasmodium* and *Babesia* (Arisue & Hashimoto, 2015; Burki et al., 2009; Schreeg et al., 2016), approaches applied in the *Plasmodium* research may also be efficient in the development of selective *Babesia* treatment.

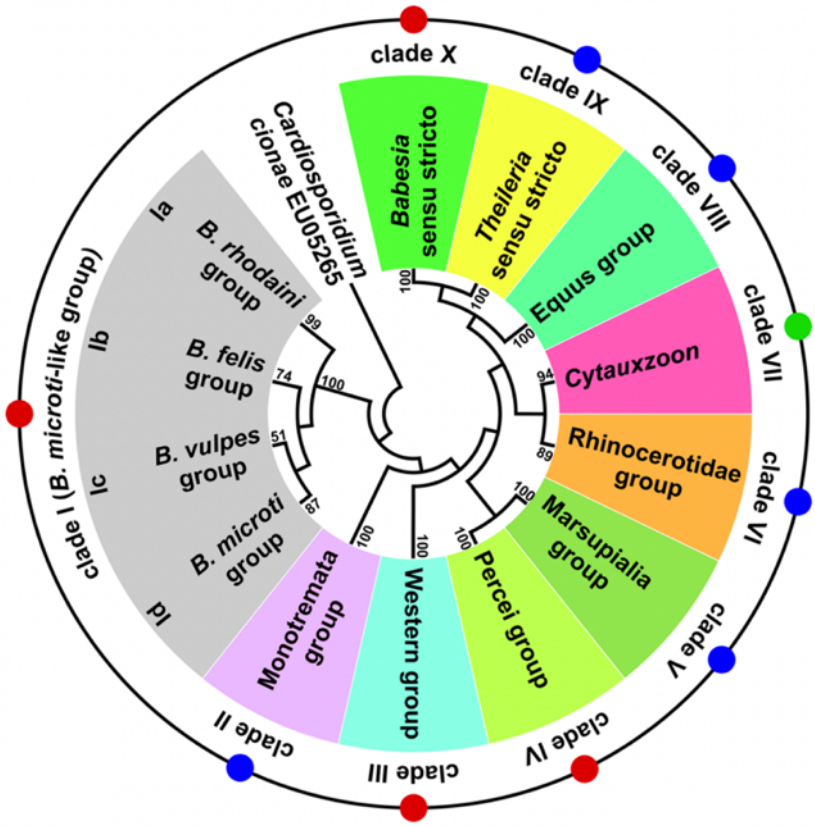

Figure 1: Molecular phylogeny of the order of Piroplasmida. The order Piroplasmida is comprised of 10 clades. *Babesia* composes 4 of those lineages – *Babesia* sensu stricto (clade X), Percei group (clade IV), Western group (clade III), and *B. microti*-like group (clade I). Red, blue and green dots depict clades associated with *Babesia*, *Theileria,* and *Cytauxzoon* species, respectively (Jalovecká et al., 2019).

1.1. *Babesia divergens*

Babesia divergens is classified as a *Babesia* sensu stricto (clade X in Figure 1) and the transmission is mediated by *Ixodes ricinus*, a member of the family of hard ticks (*Ixodidae*). *I. ricinus* has wide geographical distribution across Europe, Russia, and North Africa, which demonstrates the ability to survive under numerous conditions. The tick has a three-host life cycle, spanning three years, and humans serve as a host to all tick developmental stages (Sonenshine & Roe, 2014).

B. divergens is the most common blood parasite in Europe causing a disease named babesiosis. The reservoir host of *B. divergens* is cattle, and bovine babesiosis is considered the most important tick-transmitted disease in cattle (Zintl et al., 2003). *B. divergens* is one of the *Babesia* species able to infect humans, but they represent accidental hosts (Vannier et al., 2008). Transmission of *B. divergens* into the vertebrate host is possible during tick feeding or via blood transfusions. Even though little attention has been paid to transition via blood transfusions on a global scale, it occurs very frequently in endemic areas, particularly in people with reduced immunity due to old age or illness (Hildebrandt et al., 2007).

1.2. The life cycle of *B. divergens*

Babesia lifecycle includes two types of hosts (Figure 2). The intermediate host is a vertebrate in which asexual reproduction takes place in erythrocytes, and the definitive host is a tick, whose midgut tissue serves as the place of *Babesia* sexual reproduction (Vannier et al., 2015). The lifecycle of *B. divergens* can be divided into three consecutive phases: sporogony, merogony, and gamogony (Figure 2). During tick feeding, sporozoites are transported into the vertebrate host by saliva secretion. Sporozoites then invade red blood cells where they develop into trophozoites. Inside erythrocytes, trophozoites divide into merozoites in a process called merogony and subsequently egress to infect other red blood cells (Jalovecká et al., 2019). During asexual reproduction of merozoites, in the bloodstream of the intermediate host gametocytes occur and represent the so-called pre-sexual stages (Hildebrandt et al., 2013; Vannier et al., 2015). These transient pre-sexual stages are then imbibed into the tick gut lumen along with the mixture of asexual stages (merozoites) during tick feeding. While merozoites rapidly degrade, gametocytes further develop into gametes, called spiky-rayed bodies or Strahlenkörper (Mehlhorn & Schein, 1993). Fertilization and zygote formation occur upon the connection of two gametes of different types and the penetration of one gamete by the other

one. The motile zygote termed ookinete then penetrates the peritrophic matrix, a semipermeable membrane covering the inner surface of tick gut epithelium, and invades the tick gut cells (Schnittger et al., 2012). Inside the midgut epithelial cells, the ookinete undergoes meiotic division resulting in numerous haploid kinetes. Kinetes are then distributed all over the tick body, even to the ovaries. This leads to *Babesia* transovarial transmission, a process where one infected tick female spreads parasites to her numerous (up to thousands) eggs – a great strategy of *Babesia* widespread to multiple tick offspring (Vannier et al., 2015). Kinetes are spread to various tick organs, multiply and create secondary kinetes that invade tick salivary glands, where they transform into a single membrane syncytium called the sporont. The sporont develops into multinuclear sporoblast and stays dormant until the next instar of the tick starts feeding on a vertebrate host. After tick attachment on the host, the maturation of the sporoblast is initiated and the apical complex is formed before maturation of sporozoites is completed. During infected tick feeding on a host, mature sporozoites are released into the host bloodstream (Jalovecká, Hajdušek, et al., 2018; Mehlhorn & Schein, 1993).

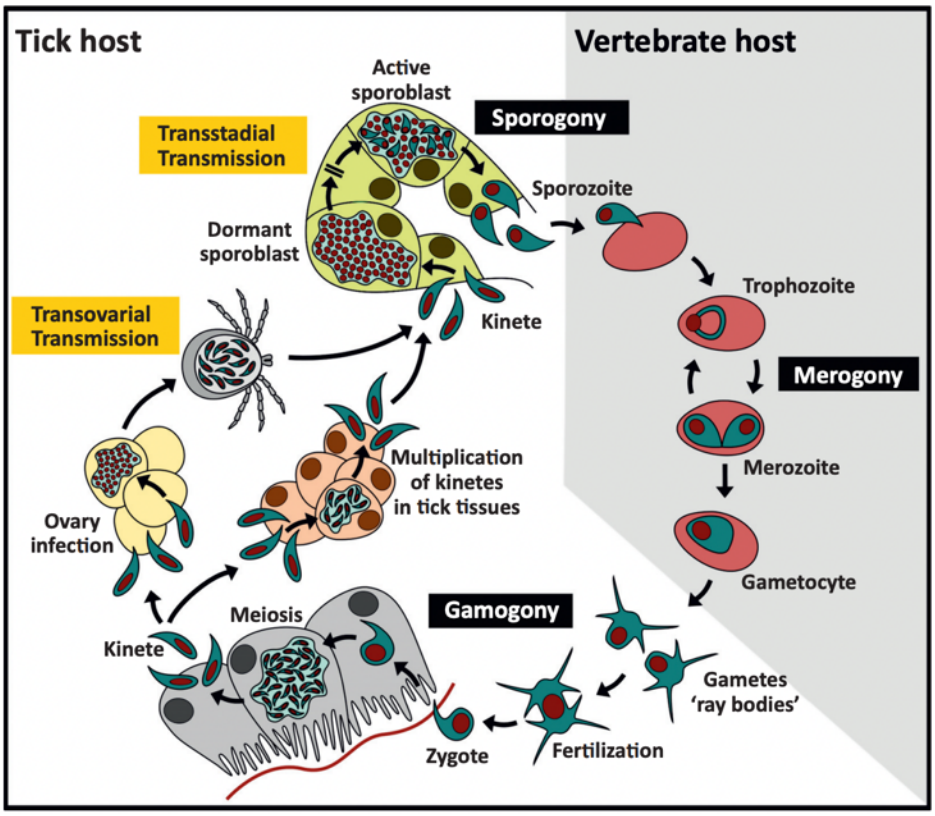

Figure 2: The life cycle of *Babesia divergens*. Sporogony takes place in the salivary glands of a tick. Sporozoites are transmitted via the saliva during tick feeding into the bloodstream of a vertebrate host. After erythrocyte invasion, sporozoites undergo merogony and develop into merozoites. Sexual commitment occurs and pre-gametocytes appear in the bloodstream. Gamogony is completed in the tick gut lumen after tick feeding on the host when newly formed

gametes fuse in the zygote that meiotically divides in the epithelium of a tick gut. Kinetes disperse over the tick body and multiply in various tick tissues, including salivary glands, where sporogony occurs. The result of this final multiplication cycle is a production of numerous sporozoites, host-invasive stages of the parasite, which enter the tick saliva and infect the host when the subsequent instar of the tick is feeding (Jalovecká, Hajdušek, et al., 2018).

2. Babesiosis

2.1. Bovine babesiosis

Babesia species responsible for bovine babesiosis have a major global economic importance in the livestock industry because babesiosis ranks among their most common blood diseases (Hunfeld et al., 2008). Specifically, *Babesia bovis* and *Babesia bigemina* found mainly in tropical and subtropical areas, and *Babesia divergens* geographically located in Europe, are the main causative agents of bovine babesiosis (Gohil et al., 2013). *Babesia* parasites can cause acute and persistent infections that appear without obvious clinical symptoms. The infections induced by *B. bovis* are linked with greater severity of the disease and increased livestock mortality and are therefore considered to be the most virulent agent of bovine babesiosis (Suarez et al., 2019).

Control of bovine babesiosis can be achieved by administration of anti-babesia drugs, such as quinuronium sulfate, amicarbalide isethionate, diminazene aceturate, or imidocarb diproprionate. However, imidocarb is the only product on today's market in most of Europe because it is the only babesicide that consistently clears the host of parasites (Ord & Lobo, 2015). A significant part of research is directed towards developing vaccines as prevention against babesiosis (Rathinasamy et al., 2019). Live vaccines containing virulence-attenuated parasites have been used to prevent the development of clinical disease nonetheless, widespread use of live vaccines is hampered by several drawbacks, such as variable degree of protection, potential tick transmissibility, reversion to virulence, viability, and logistical difficulties associated with vaccine production and deployment (de Waal & Combrink, 2006). Babesiosis vaccine research is now focused on the development of safe and efficacious non-live vaccines to overcome the limitations of live vaccines by using vaccines containing killed parasites, native proteins, soluble parasite antigens (SPAs) or recombinant parasite proteins (Brown et al., 2006; Schetters & Montenegro-James, 1995; Timms, 1989). The prevention remains poorly controlled and thus novel vaccines to prevent the development of acute disease are essential (Florin-Christensen et al., 2014).

2.2. Human babesiosis

Human babesiosis is an emerging zoonotic disease, caused by several species of *Babesia* genus. The most predominant agent of human babesiosis in North America is *Ixodes scapularis* transmitted *Babesia microti* (Joseph et al., 2012), also spread via blood transfusions (Ord & Lobo, 2015), while *Babesia divergens* is the most common agent causing human babesiosis in Europe (Hildebrandt et al., 2021). Unlike babesiosis derived from *B. divergens*, infections caused by *B. microti* are in most cases asymptomatic, however in immunocompromised patients or elderly people it can develop into a mild form of symptoms such as fever, weakness, fatigue, muscle pain, headache, splenomegaly, and hepatomegaly, and in the worst cases can result in death (Homer et al., 2000; Yabsley & Shock, 2013). The symptoms manifestation is due to merozoite multiplication in host cells, associated with erythrocyte lysis. Studies have also suggested an important role for cytokines in their pathogenesis (Krause et al., 2007). Treatment is based on the administration of antibiotics alongside with antimalarials, specifically clindamycin or atovaquone in combination with azithromycin or quinin, respectively (Vannier et al., 2008). Medication is currently nonspecifically aimed against *Babesia* and might often induce adverse effects (Healy & Ristic, 2018). Hence the effort to develop selective, targeted treatment or vaccine for humans and livestock is essential (AbouLaila et al., 2010).

2.2.1. Possible future treatments

There are several possible directions that drug research on babesiosis is taking. Some approaches are studied in our laboratory. For example, the proteasome inhibition of a parasite appears to be a promising strategy for the development of selective drugs. Targeted proteasome inhibition is an effective therapeutic strategy for the treatment of cancer and autoimmune diseases (Li et al., 2012; Reynolds et al., 2007) since the proteasome inhibition results in a toxic buildup of abnormal proteins (improperly synthesized or packed, potentially toxic, mutated or damaged) leading to apoptosis of the cell (Basler et al., 2015). Recently, selective inhibitors of parasitic proteasomes have been developed for the causative agents of malaria, leishmaniasis, Chagas disease, or sleeping sickness (Bibo-Verdugo et al., 2017; Khare et al., 2016; Li et al., 2016). Our laboratory tested the effect of selective inhibitors (Carfilzomib, Bortezomib) in *Babesia divergens in vitro* and a potential therapeutic effect was

demonstrated with several proofs, e.g. by the accumulation of polyubiquitinated proteins in parasites by antibodies (Jalovecká et al., 2018).

Calcium-dependent kinases might present another therapeutic target. These enzymes play a crucial role in the regulation of invasion into the host erythrocytes and egress from them. Calcium-dependent kinases occur in other Alveolates and plants, but not in animals which makes them a selective therapeutic target (Choi et al., 2020; Hortua Triana et al., 2018). Many of them are not yet characterized and their regulatory role in the process of parasite entry into the host cell and persistence within it is not well studied (Keyloun et al., 2014). This characterization of *Babesia* is currently being attempted by Dr. Jalovecká and Dr. Sojka as a part of ongoing projects of our laboratory.

3. Functional genomics tools

Thorough research of the fundamental biological principles of *Babesia* parasitic lifestyle is necessary to find suitable candidates for the development of specific drugs and cannot be done without effective tools for functional genomics. The increased number of techniques for *in vitro* culture, transfection, and availability of genome sequencing of various *Babesia* species has allowed the introduction of certain methods of genetic manipulation. The genomic sequence of *B. divergens* has been published (González et al., 2019) and allows us to understand and study various molecular processes including the life cycle, gene regulation, metabolism, and the mechanism of red blood cell invasion (Young et al., 2019). Transfection methods are commonly used to study mechanisms of gene regulation and functional gene analysis using gene knock-out and knock-in strategies (Garcia et al., 2008). There is a solid *in vitro* cultivation technique for some *Babesia* species of such as *B. bovis, B. ovata, B. bigemina, B. gibsoni*, and a stable transfection line for this species has been established (Hakimi et al., 2016; Liu et al., 2018; Silva et al., 2018). Even though for our model organism *B. divergens in vitro* culture in bovine erythrocytes has been optimized (Jalovecká et al., 2016), the stable transfection line has not yet been introduced.

3.1. *Babesia* transfection system

The transfection system is related either to the permanent expression of the gene of interest through the integration of the transfected DNA directly into the nuclear genome via homology recombination (Chapter A.3.2.), or the maintenance of a transfected plasmid as an

extrachromosomal replicating episome. The transfected DNA is introduced into the genome via linearized plasmid and the transgene is by single or double-crossover homologous recombination inserted into the *Babesia* genome (Hakimi et al., 2021). The main advantage of this approach insists on a controlled number of copies leading to better reproducibility of obtained results. On the other hand, episomal expression of the transgene is usually performed by the transfection of a circular plasmid, which remains free in the cell. Since episomal plasmids do not interfere with the genome of the cell, there is no risk that the expression of the desired gene is under the unexpected control of an internal promoter or under the control of other possible regulatory factors that have not been described yet (Mulia et al., 2021).

An episomal transfection involves designing a transfection plasmid containing a reporter gene and a gene encoding the target protein product. Reporter genes serve as markers of whether a particular gene has been incorporated and expressed in a cell population, therefore reporters, as well as a gene of interest, must be flanked by 5' UTRs and 3' UTRs (untranslated regions on 5' or 3' end) for gene expression to occur (Suarez et al., 2017). A reporter marker is usually a gene that provides cells with a selection advantage, often resistance to a selection drug. Selective inhibition of parasite growth in *in vitro* culture is essential for the development of transfection lines. BSD (blasticidin-S-deaminase) and WR99210 (hdhfr) are functional selection drugs commonly used for *Babesia* selective inhibition (Asada et al., 2015).

The plasmid is introduced into the cell by electroporation, nucleofection, microinjection, or via the liposome vector. In the case of *Babesia*, the most commonly used technique for plasmid delivery is nucleofection (Suarez & McElwain, 2010). The procedure of nucleofection is based on the transfer of nucleic acid into cells by applying a specific voltage and reagents. The transfected plasmid can be either linearized (intragenomic integration) or circular (episomal and intragenomic integration).

Transfection techniques improve a functional analysis of genes and identification of virulence factors, leading to the discovery of several genes that are required for the survival of apicomplexan parasites, the transmission of the parasites by their tick vector, and thus help the development of vaccines as well as improve characterization of mechanisms involved in babesiosis (Goswami et al., 2019). Transfection systems also allow detailed study of promoters related to the regulation of gene expression of essential and non-essential genes. They have enabled the identification of different promoters (e.g. the two-way promoter ef-1α) for several *Babesia* species, such as *B. bovis, B. bigemina* (Silva et al., 2016). It is therefore possible that regulatory mechanisms may be similar in different *Babesia* species and that

knowledge of the regulatory mechanism in one species may help to understand regulation in another (Suarez et al., 2017).

3.2. Homologous recombination

Homologous recombination is a type of genetic recombination in which genetic information is exchanged between two similar or identical molecules of DNA. It is widely used by cells to repair harmful double-strand breaks that occur on both strands of DNA, in a process called homologous recombinational repair. Homologous recombination produces new combinations of DNA sequences during meiosis and is also used in horizontal gene transfer to exchange genetic material between different species of bacteria (Hoshijima et al., 2016). Genetic manipulation methods, based on the insertion of desired DNA sequence into the genome using homologous recombination mechanisms, have recently been successfully applied to several of the apicomplexan species (Suarez et al., 2017).

Gene targeting is an approach by which a gene of interest or site-specific mutation is permanently or conditionally inserted into the targeted locus. To incorporate the DNA sequence into the genome, it first needs to be flanked from both ends by an identical sequence as the gene in the cell's genome. The homologous regions are then recognized by the cell and the sequence is replaced from the original sequence by homologous recombination initiated by the cell itself as a part of its DNA reparation mechanisms.

This process is therefore used for causing gene knock-out, which prevents the expression of a specific gene or gene knock-in, that alters the genetic locus by addition of DNA sequence (Hoshijima et al., 2016). Even though stable integration of exogenous genes in *Babesia* has proven difficult, Suarez & McElwain (2010) have integrated GFP plasmid into the genome of *B. bovis* using blasticidin as a selection marker, creating a stable fluorescent reporter line. Gene targeting is widely used for studying gene and protein function *in vivo*. In various Apicomplexa parasites, including *Babesia*, homologous recombination revealed diverse mechanisms of cell invasion by different stages of the parasite, sexual differentiation, or the mechanism of adherence to infected red blood cells (de Koning-Ward et al., 2015).

3.3. Conditional regulation of gene expression

Conditional or inducible expression systems can regulate the expression of a target gene of interest at the genome, transcriptome, or protein level. At the DNA level, inducible site-

specific recombinases allow conditional genome editing. At the RNA level, regulation can be achieved during transcription, using stage-specific or regulatable promoters, or post-transcriptionally through alteration of mRNA stability or translation. At the protein level, several systems have been developed for the inducible degradation or displacement of a protein of interest.

Given that the *Babesia* genome is haploid in erythrocytic and most of the developmental stages in the tick, conventional non-inducible knock-out systems cannot be used for functional characterization of genes that are essential for parasite invasion, growth or egress in different parasite life stages. Many of these genes represent potential therapeutic targets (Hakimi et al., 2021). Conditional genome editing can also be applied to non-essential genes, in order to avoid adaptation of the parasite that can occur during long selection procedures associated with conventional approaches (Briquet et al., 2021). The combination of both types of transfection systems gave rise to a new method of functional analysis of essential genes, where a gene is knocked-out and its function is then restored, the so-called conditional knock-out system (Asada et al., 2015). Conditional genome editing can be achieved by activating site-specific recombinases (e.g., Flp/FRT, DiCre), exchange of stage-specific promoters (tetracyclin-regulated promoters), inducible ribozymes (GlmS) or TetR-aptamer systems, or via destabilization domains causing protein degradation (FKBP-DD, DHFR, auxin-inducible degron system)(Briquet et al., 2021).

3.3.1. DiCre technique

Conditional gene knock-out methods are not established for *Babesia* yet. The technique of conditional deletion of a target gene has been established for *Plasmodium* spp. and *Toxoplasma* spp. and was shown to be efficient for several targets in the *in vitro* and *in vivo* models (de Koning-Ward et al., 2015; Knuepfer et al., 2017). One of the ways of conditional modification of essential genes is the DiCre system, modified Cre recombinase which can be regulated using ligand-induced dimerization. The first component of this system is a dimerizable Cre (DiCre) recombinase, in which its inactive two proteins are fused by rapamycin addition through cross-linking the protein fragments FKBP12 and FRB (Hakimi et al., 2021). Cre is a site-specific recombinase that catalyzes the excision of a DNA segment flanked by two identical short (34 bp) asymmetric sequences, called loxP sites (Jullien et al., 2007). Thus, two plasmid vectors are needed for a successful introduction of DiCre system to a selected organism/species. The first plasmid contains the DiCre enzymatic machinery and

the second comprises loxP sites, that define the locus in the genome where the gene manipulation will take place.

The expression of DiCre *in vivo* is complicated by the fact that this system requires the simultaneous expression of two constructs in the organism (Knuepfer et al., 2017). The DiCre system has proven to be an effective approach for conditional gene deletion of numerous Apicomplexa species, such as *Toxoplasma gondii* and *Plasmodium falciparum* (Andenmatten et al., 2013; Collins et al., 2013). The DiCre parental lineage of *Babesia* expressing Cre recombinase and the design and construct of this modified *B. divergens* lineage is the main topic of my thesis.

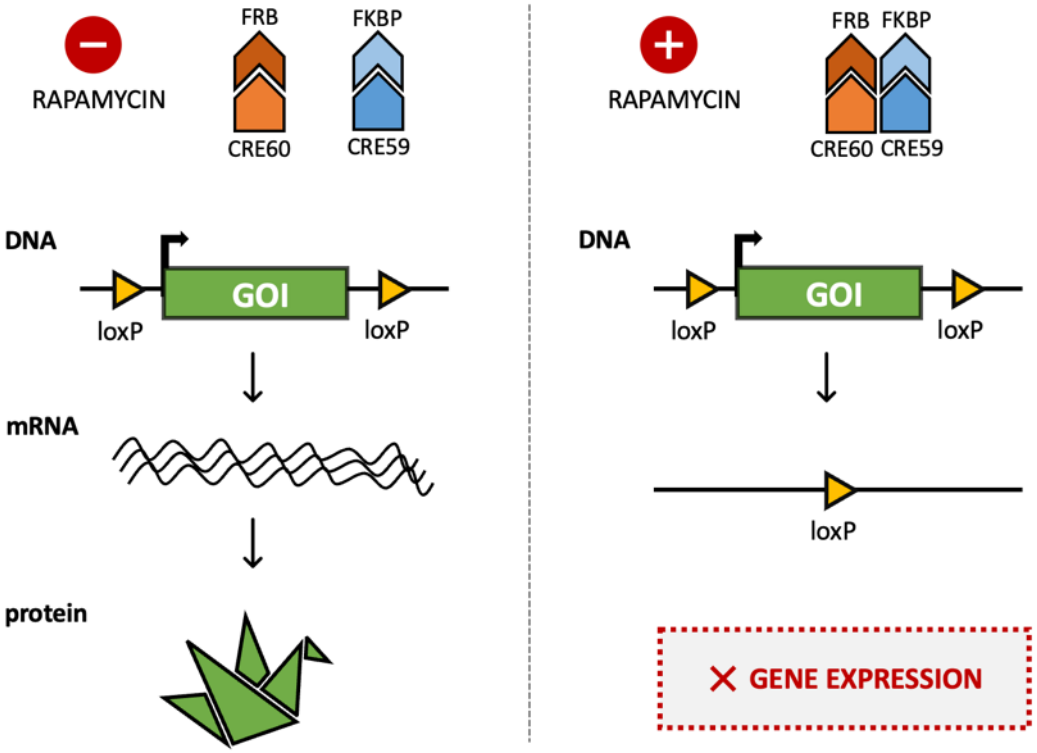

Figure 3: Conditional genome editing via DiCre technique. In the DiCre system, the Cre recombinase is expressed as two subunits fused to FKBP and FRB, respectively, that dimerize in the presence of rapamycin. After the fusion, the subunits restore recombinase activity, inducing site-specific recombination between loxP sites and excision of the floxed gene of interest (GOI).

B. OBJECTIVES

- Identification and sequencing of eligible recipient locus (homology region) in the *B. divergens* genome into which a DiCre cassette will be integrated via homology recombination
- Validation of WR99210 and G-418 as selection markers and determination of optimal concentrations for their use in *B. divergens in vitro* cultures
- Design and construct plasmid vector for the generation of parental DiCre *B. divergens* lineage

C. MATERIALS AND METHODS

1. Cultivation of *Babesia divergens*

Cultivation of *B. divergens* was carried out in the suspension of bovine erythrocytes *in vitro* in a Galaxy 170 S incubator (New Brunswick, Eppendorf) under constant conditions of 37 °C, 5% CO_2. The manipulation with the cultures was done in an Airstream Class II laminar box (BSCO) under basic sterile conditions. *B. divergens* strain 2210A G2 was used in all experiments (Jalovecká et al., 2016).

Whole bovine blood, obtained from indoor bred cows from a local slaughterhouse, was manually defibrinated and 3 times washed (Multi Centrifuge 3SR, Heraeus; 2000×*g*, 4 °C, for 10 min) in the washing medium: 500ml RPMI 1640 (Lonza), 0,5ml amphotericin B (Sigma-Aldrich, c=250µg/ml), 2,5ml gentamycin sulfate (Sigma-Aldrich, c=10mg/ml). The fibrin layer and leukocytes (called the "buffy coat") were removed from the pellet surface with a pipette after each centrifugation. Washed erythrocytes were stored in the washing medium (1:2 ratio) at 4˚C and used for *in vitro* cultivation for 3 to 6 weeks. In order to avoid possible contamination, erythrocytes were regularly tested for possible contamination: 10µl of erythrocytes with washing medium were pipetted onto blood agar (Dulab), which was placed in the incubator and observed for next 3-5 days.

B. divergens in vitro cultures were maintained in 24-well culture plates (Techno Plastic Products). Cultivation medium (2ml per well) consisted of: 500ml RPMI 1640 (Lonza), 0.5ml amphotericin B (Sigma-Aldrich, c=250µg/ml), 2.5ml gentamycin sulfate (Sigma-Aldrich, c=10mg/ml), 10ml fetal bovine serum (Capricorn). Apart from medium well contained 50µl bovine erythrocytes and 10µl of infected passaged erythrocyte sediment. The resulting 2.5% hematocrit (the ratio between the volume of erythrocytes and the volume of cultivation medium) was maintained throughout the cultivation of *B. divergens*. When 5-10% parasitemia was reached (Chapter C.1.1.), 10µl of infected erythrocytes were collected from the well and pipetted in a new well with a suspension of medium and erythrocytes of the determined hematocrit. Cultures were maintained by regular passaging ideally 3 times a week and observed microscopically to estimate parasitemia and avoid the possible contamination (Chapter C.1.1.). When necessary (Chapter C.2.) the culture was cultivated into larger volumes: 25cm^2 or 75cm^2 cultivation bottle (VWR), while maintaining the determined hematocrit.

1.1. Determination of parasitemia by light microscopy

Determination of parasitemia (number of infected red blood cells in the total amount of blood) was defined by the Giemsa-like stained blood smears. The slide was air-dried, stained with the commercial Diff Quick Staining Set (Siemens) according to the provided protocol, rinsed with 70% ethanol and distilled water. The stained slide was analyzed under the BX53F a light microscope (Olympus) at a magnification of 1000×, using immersion oil. Photos were obtained by the Olympus DP72 camera and the CellSens imaging software. Parasitemia was counted as the number of *B. divergens* infected erythrocytes per 1000 erythrocytes.

1.2. Determination of parasitemia by flow cytometry

For parasitemia determination by flow cytometry a volume of 50 µl of pelleted infected red blood cells taken out from *B. divergens* cultures was centrifuged (Multi Centrifuge 3SR, Heraeus; 600×*g*, 3 min) and washed with 1×PBS (8g NaCl, 0.2g KH_2PO_4, 2.9g $Na_2HPO_4 \cdot 12H_2O$, 0.2g KCl, in 1 liter of dH_2O). Then the fixative solution (4% paraformaldehyde and 0.025% glutaraldehyde in 1×PBS) was added and the sample was incubated at room temperature for 30 min. After incubation, the sample was centrifuged (Multi Centrifuge 3SR, Heraeus; 600×*g*, 3 min) and washed twice with 1×PBS. Finally, 50 µl of 1×PBS were added to the pellet. Fixed cells were stored in a refrigerator at 4 °C for a maximum of three weeks. To visualize the parasitemia, cells were stained using Ethidium Homodimer 1 - EthD-1 (Biotium): after centrifugation (Multi Centrifuge 3SR, Heraeus; 600×*g*, 3 min), the pellet was incubated at 37 °C for 30 min with 50 µl of 0.02mM EthD-1 (diluted in 1×PBS). The sample was then centrifuged (Multi Centrifuge 3SR, Heraeus; 600×*g*, 3 min) and washed twice with 1×PBS and transferred into a flow cytometry tube. Stained samples were subsequently analyzed on a flow cytometer FACS CantoII using the provided Diva software (BD Biosciences).

2. Nucleic acid isolation

B. divergens genomic DNA was isolated from the *in vitro* culture of 5-10% parasitemia. A volume of 200 µl of 1× PBS was added to the 50 µl of infected erythrocytes. DNA was then isolated from this suspension using the NucleoSpin® Blood DNA isolation kit (Macherey-Nagel) according to the manufacturer's instructions and the final concentration of

isolated DNA was measured using the NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific), and DNA was stored at -20°C. The success of DNA isolation was tested via PCR and subsequent gel electrophoresis (Chapter C.4.) using primers previously validated in the laboratory (Table 1).

Table 1: Primers for verification of DNA isolation.

Number	Name	Sequence	Annealing temperature
1	**BdSub1 pet100 F**	CACCACACTGGACCAG GAGACACCTTC	60°C
2	**BdSub1 pet100 R**	CTATGGAAAGAAGTAGA CTTGCAACACCCG	60°C

3. Determination of WR99210/G-418 effect on *B. divergens in vitro* culture

Experiments were conducted in 96-well plates (Techno Plastic Products) in biological triplicates according to the scheme provided in Figures 4 and 5. Each well contained 100 µl of the culture medium (Chapter C.1.) with defined concentration (Figures 4, 5) of selection drugs WR99210 (kindly provided by Jacobus Pharmaceutical) or G-418 (Roche) and 5 µl of bovine erythrocytes. The culture medium containing selection drugs at defined concentrations was exchanged in an interval of 48 hours. Samples were collected in biological triplicates at two-day intervals for 8 consecutive days in the case of WR99210 and 10 consecutive days in the case of G-418 (2, 4, 6, 8, 10 DPI – days post infection) and immediately fixed (Chapter C.1.2.). Blood smears from each time point were prepared to control parasitemia parallelly to flow cytometry parasitemia determination (Chapter C.1.1.). Fixed and subsequently stained samples were evaluated using flow cytometry (Chapter C.1.2.). Statistical analyses of obtained data were performed in GraphPad Prism software (version 7.0a) by AUC (area under the curve) calculations followed by one-way analysis of variance (ANOVA). IC_{50} values were determined, and graphs were also created in GraphPad Prism software (version 7.0a).

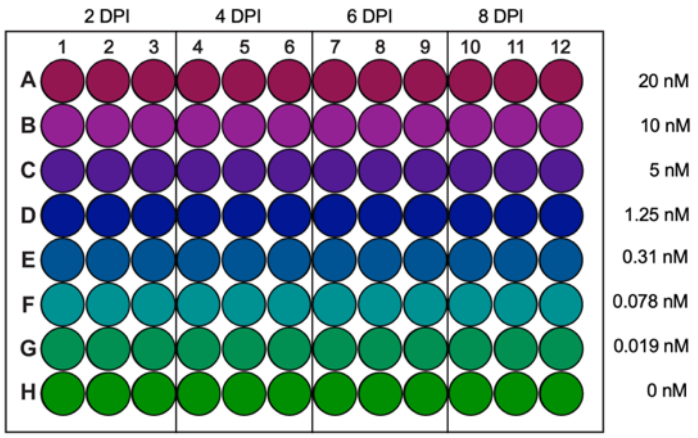

Figure 4: Design of experiment with WR99210 drug. Final concentrations of WR99210 drug in cultivation media: 20 nM, 10 nM, 5 nM, 1.25 nM, 0.31 nM, 0.078 nM, 0.019 nM, 0 nM. Each timepoint is represented by biological triplicates. DPI – days post infection.

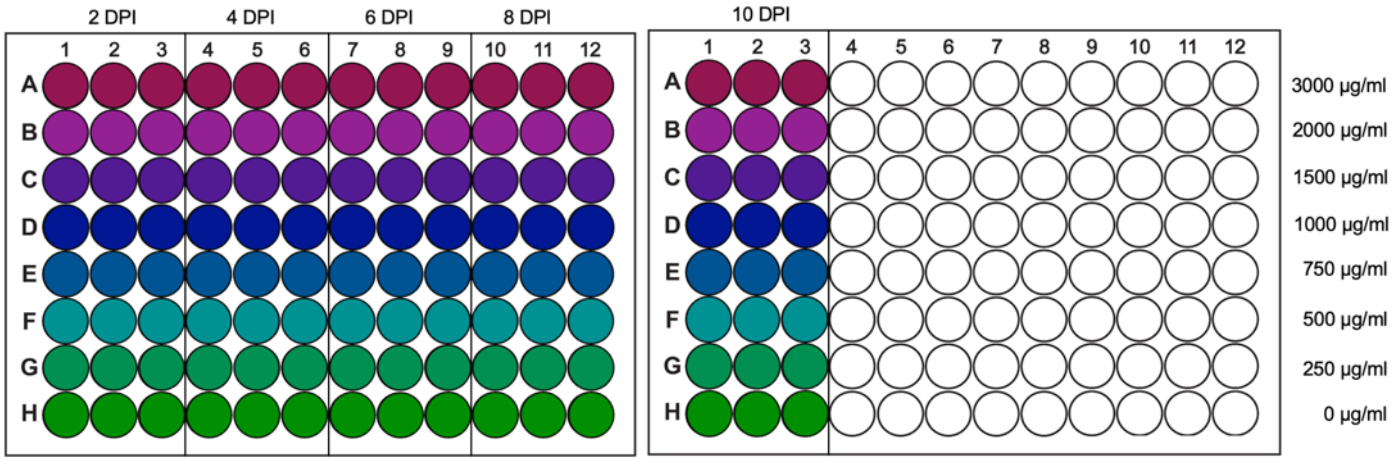

Figure 5: Design of experiment with G-418 drug. Final concentrations of a G-418 drug in cultivation media: 3000 µg/ml, 2000 µg/ml, 1500 µg/ml, 1000 µg/ml, 750 µg/ml, 500 µg/ml, 250 µg/ml, 0 µg/ml. Each timepoint is represented by biological triplicates. DPI – days post infection.

4. PCR and agarose gel electrophoresis

Polymerase chain reactions (PCR) were executed in the T100 Thermal Cycler (BioRad). The volume of PCR reaction for colony testing was 10 or 25 µl (Table 2). PCR primers were designed using the Geneious Prime software ® (version 2022.0.1). The generic PCR program is described in Table 3. The annealing temperature of PCR primers was established by gradient PCR (annealing temperature: 45-65°C). PCR products were separated on Ethidium Bromide (Sigma-Aldrich) stained 1% agarose gel consisting of Agarose (Sigma-Aldrich) and 1 × TAE

buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0). PCR samples were mixed with Gel Loading Buffer II (Invitrogen), loaded on the agarose gel, electrophoretically separated and visualized using Gel logic 112 transilluminator (Sigma-Aldrich). The size of PCR products was determined using GeneRuler™ 100bp or 1kb DNA Ladders (Thermo Fisher Scientific). For cloning and sequencing purposes, resulting PCR products were isolated from residual PCR reactions or from agarose gel slices using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel). DNA concentration was measured using the NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific).

Table 2: Standard PCR reaction set-up.

Reagent	Total volume 10 µl	Total volume 25 µl
FastStart™ PCR Master (Roche)	5 µl	12.5 µl
Forward primer (10 µM)	1 µl	1 µl
Reverse primer (10 µM)	1 µl	1 µl
DNA	1.6 µl	4 µl
Nuclease-free water (Top-Bio)	1.4 µl	6.5 µl

Table 3: Standard PCR amplification program.

	Temperature	Time	Repetition
Initial denaturation	94°C	5 min	
Denaturation	94°C	1 min	30×
Annealing	55 – 60 °C	30 sec	
Elongation	72°C	1 – 4 min	
Final elongation	72°C	4 min	

5. DNA cloning and sequencing

The pGEM®-T Easy Vector cloning reaction (Promega) and One Shot TOP10 *Escherichia coli* cells (Thermo Fisher Scientific) were used for conventional cloning and sequencing purposes. TOP10 *E. coli* cells were then spread onto preheated selection Luria-Bertani (LB) broth plates supplemented with ampicillin (AMP) at 37°C, for 12 hours. Plates were prepared using 10 g of broth (Amresco), 6 g of agar (Amresco), 400 ml of distilled water, and AMP with final c = 50 µg/ml.

Colonies were tested for the presence of cloned PCR amplicons using colony PCR (Chapter C.4.) with insert specific primers (Table 4). The volume of the reaction was 10 µl. Each colony was dissolved in 20 µl PCR H_2O and used as a DNA template. Positive *E. coli* colonies were inoculated into 4 ml of LB-AMP medium and incubated (37°C; 200 rpm; 15 hours). The plasmid DNA was isolated using the NucleoSpin® Plasmid kit (Macherey-Nagel) following the provided protocol. Resulting plasmid DNA concentration of isolated plasmid DNA was determined using the NanoDrop™ 1000 spectrophotometer (Thermo Fisher Scientific). Plasmid DNA was sequenced by the external service company (SEQme s.r.o.) and DNA sequences were analyzed using the Geneious Prime software.

Table 4: Insert specific primers for colony testing.

Number	Name	Sequence	Annealing temperature
1	**Bdiv_004560c F1**	GCCATGATGACACTTAA ACAAG	56°C
2	**Bdiv_004560c R1**	CAGCCCTAGGACGACT GAAA	59°C

6. Bioinformatic analyses and DNA construct design

DNA sequences were analyzed using the Clustal Omega Multiple Alignment in the Geneious Prime software® (version 2022.0.1) and compared to reference sequences obtained from PiroplasmaDB (https://piroplasmadb.org/piro/app/).

Orthologues were identified using a The Basic Local Alignment Search Tool (BLAST) search in PiroplasmaDB database (https://piroplasmadb.org/piro/app/) and PlasmoDB database (https://plasmodb.org/plasmo/app). Multiple amino acid sequence alignments and calculation of sequence identities of orthologues were carried out using the Clustal Omega Multiple Alignment in the Geneious Prime software® (version 2022.0.1).

The plasmid construct with site-specific unique restriction sites was designed in SnapGene Viewer (version 6.0.2). Sequences for the vector map were kindly provided through the collaboration of Dr. Jalovecká and Dr. Sojka with Dr. Ganter (Heidelberg University, Germany), and Dr. Asada (Obihiro University of Agriculture and Veterinary Medicine, Japan).

7. Gibson assembly technique

Plasmid construct linearization was performed via DNA restriction by the *Not*I enzyme (New England Biolabs) for 6 hours at 37°C. Subsequently, the Quick CIP phosphatase was added to avoid recircularization and re-ligation by the removal of phosphate groups from both 5' ends of linearized DNA strands. The restriction reaction was heat-inactivated at 80°C for 2min. Gibson reaction was set up (20 µl reaction, incubation for 60 min at 50°) according to the NEBuilder HiFi DNA Assembly protocol (New England Biolabs) using 7:1 (insert:vector) molar ratio calculated by the NEBioCalculator (https://nebiocalculator.neb.com/#!/ligation). The desired insert was PCR amplified (Chapter C.4., sample volumes: 25 µl) using specific primers for Gibson cloning (Table 5). The chemical transformation of the plasmid was conducted with NEB 5-alpha Competent *E. coli* cells (New England Biolabs). 2 µl of the reaction was added into the vial with competent cells, incubated on ice for 30 min, heat-shocked, and then incubated with SOC medium (New England Biolabs) for 1 hour. Bacterial cells were then spread onto tempered selection LB broth/agar plates supplemented with AMP and cultivated at 37°C for 12 hours (Chapter C.5.). Selection of positive plasmid colonies, isolation of plasmid DNA and sequencing was conducted according to Chapter 6. Primers used for plasmid DNA sequencing are listed in Table 6.

Table 5: Primers for DiCre cassette part A/B PCR amplification and Gibson assembly.

Number	Name	Sequence	Annealing temperature
1	**PF 1 gibson**	TAACCCTTCAAAGCGATCGCATGGCC CCTAAGAAGAAGAG	57°C
2	**PR 2 gibson**	TGGATGGTCTTCACAGCCAGGCCTCT GGCTTGCAGG	57°C
3	**PF 3 gibson**	TGTACCTGCAAGCCAGAGGCCTGGCTG TGAAGACCATC	57°C
4	**PR 4 gibson**	CTAGTCTACGGTGCGGCCGCATCGAT ACGGTGATTAATTAATCAGT	57°C
5	**PF 1.1. gibson**	CTGATTAATTAATCACCGTATCGATTT CGATATCAGATTTAATGCTAGTTAATTC	59°C
6	**PR 1.2 gibson promoter CAM**	CTGATTTCATAAATGCTTCACATAGC	58°C
7	**PF 2.1 gibson promoter CAM**	CTTATAGGAGTATCAACCTAGAATCT	55°C
8	**PR 2.2 gibson**	GGATCCGATACTAGTCTACGGTGCGGC CGCATCGGCCGGCCCCCT	59°C

Table 6: Primers used for sequencing.

Number	Name	Sequence	Annealing temperature
1	**PF 1 seq**	CGGCGCCGTTCACATATTTTG	61°C
2	**PF 2 seq**	GATGGAACGGGGCCCC	59°C
3	**PF 3 seq**	GAACTCTGACAGATGCCAGG	58°C
4	**PF 4 seq**	GGGACATGGCCAGGGC	59°C
5	**PR 5 seq**	CGGCTCGTATGTTGTGTGG	55°C
6	**hDHFR F**	CAGGTCAGGATCCGTGC	57°C
7	**hDHFR R**	GCGTTTGTGGAGGCCTTG	59°C
8	**CRE 59 F**	GGTCTGCCTTTGGCGC	59°C
9	**CRE 59 R**	GCCGACGACCTTCGTTTG	59°C
10	**HR 1 F**	CGACAAGGATAGTATTATCG	51°C
11	**HSP 70 F**	GCCTAGCACGCGTTG	55°C
12	**HSP 70 F 2**	CTGTCGCCGATACACAC	55°C
13	**ACT F**	GCTGGTATTATTAATCCACG	52°C
14	**ACT F 2**	GTTACTCGACAATAGTTAAATTGC	55°C
15	**TER CRT F**	GCATCGTGCTACGGC	55°C
16	**TER CRT F 2**	CAGGTCAGCGAACGAG	54°C
17	**CAM F**	GACACATGTACCAACATAC	51°C
18	**CAM F 2**	CAGATCGTTATATATTTAGGTATC	51°C
19	**HR 2 F**	GAAGGAGGATTCGAAAAACG	56°C
20	**M13 F**	GTAAAACGACGGCCAGT	55°C
21	**CRE 60 R**	GGATCAGCATTCTCCCAC	55°C
22	**Bdiv_004560c F2**	GTGTCCTTAGCATACGGTTGTG	59°C
23	**Bdiv_004560c F3**	CTGGTTCTAAGAACAGAGGTGG	59°C

D. RESULTS

1. Inhibitory effect of WR99210 and determination of IC_{50} value

The inhibitory effect of WR99210 selection drug on *in vitro B. divergens* culture was analyzed by parasitemia determination using flow cytometry (Chapter C.1.2.). Parasitemia levels varying due to different concentrations of the selection drug from all four time points (2, 4, 6, 8 DPI – days post infection) are depicted in Figure 6. In comparison to higher concentration levels, the scale of concentrations of 1.25, 0.31, 0.07, and 0.02 nM didn't have an influence on parasite's growth *in vitro* and the AUC (area under the curve) values were not statistically significantly different from the negative control (NC) AUC. The inhibitory effect of WR99210 was demonstrated at a concentration of 5 nM, when parasitemia on 8 DPI was 0.1 %. The AUC value (4.3 ± 0.3) analyzed from the overall parasite growth curve was significantly different ($p < 0.001$) compared to the AUC of NC (95.7 ± 28.5). The same effect was demonstrated at a concentration of 10 nM, where the parasitemia on 8 DPI was 0.1 %, and the AUC (5.0 ± 0.7) was significantly different ($p < 0.001$) compared to the AUC of NC. The concentration of 20 nM was shown to be inhibitory as well, with 0.1% parasitemia on 8 DPI, and the AUC (4.3 ± 0.5) was again confirmed as significantly different ($p < 0.001$) when compared to the AUC of NC. Figure 7 compares changes in parasitemia for the NC and for cultures treated with WR99210 medium at 20, 10, and 5 nM including the display of 8 DPI blood culture smears. A difference in parasitemia was observed at effective WR99210 concentrations of 20, 10, and 5 nM compared to the initial parasitemia (1%, 0 DPI). Statistically significant parasitemia reductions were observed for the highest concentration of WR99210 tested (20 nM) for intervals analyzed: 4 DPI (0.6 % ± 0.1, $p < 0.05$), 6 DPI (0.3 % ± 0.1, $p < 0.01$), and 8 DPI (0.1 %, $p < 0.001$), except for 2 DPI (0.7 % ± 0.2). Cultures treated with 10 nM concentration of drug showed statistically significant differences in parasitemia for all intervals analyzed: 2 DPI (0.6 % ± 0.08, $p < 0.05$), 6 DPI (0.3 %, $p < 0.001$), and 8 DPI (0.1 %, $p < 0.001$), except for 4 DPI (1.1 % ± 0.2). Last statistically significant differences were observed for 5 nM concentration WR99210 for all intervals analyzed: 2 DPI (0.6 % ± 0.5, $p < 0.001$), 4 DPI (0.7 % ± 0.1, $p < 0.01$), 6 DPI (0.3 %, $p < 0.001$), and 8 DPI (0.1 %, $p < 0.001$). Figure 8 shows the statistical determination of the IC_{50} value on 8 DPI after triplicates were averaged, their standard deviation was determined, and the data was normalized (logarithmized). Final IC_{50} was 2.49 nM and R^2 was 0.8849. WR99210 concentration of 5 nM appears optimal for selective cultivation of parasites transfected with

resistance holding plasmid. Even at the highest tested concentrations, no toxic effect of WR99210 on red blood cells was observed (Figure 7).

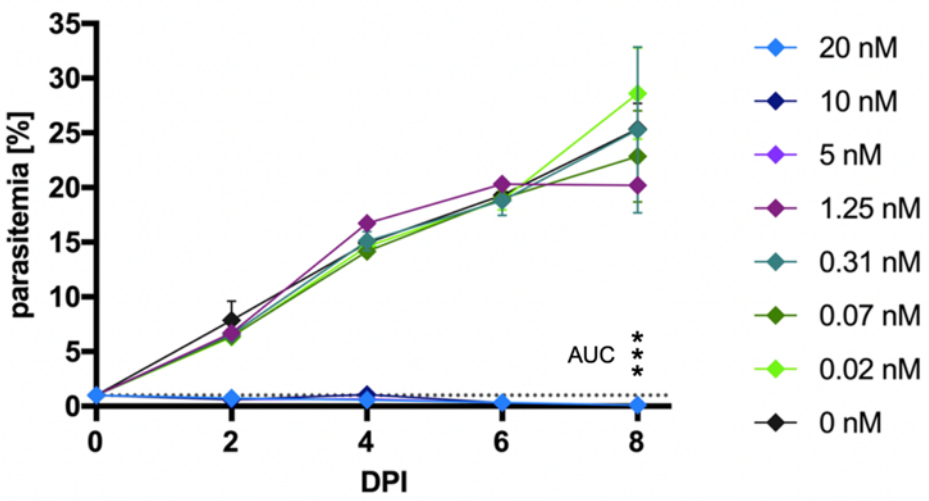

Figure 6: Treatment of *B. divergens in vitro* culture with different concentrations of WR99210 drug. Parasites were cultivated *in vitro* in standard conditions (37°, 5% CO_2) for 8 consecutive days (Chapter C.1.). The result represents the mean of three independent replicates and error bars indicate standard deviations. Parasitemia was determined by flow cytometry, for staining cultures *B. divergens* was used for EthD-1 as explained in Chapter C.1.2. The grey dotted line represents initial parasitemia (1%); DPI - days post infection; AUC - area under the curve. *** = $p < 0.001$. Statistical test performed: one-way ANOVA.

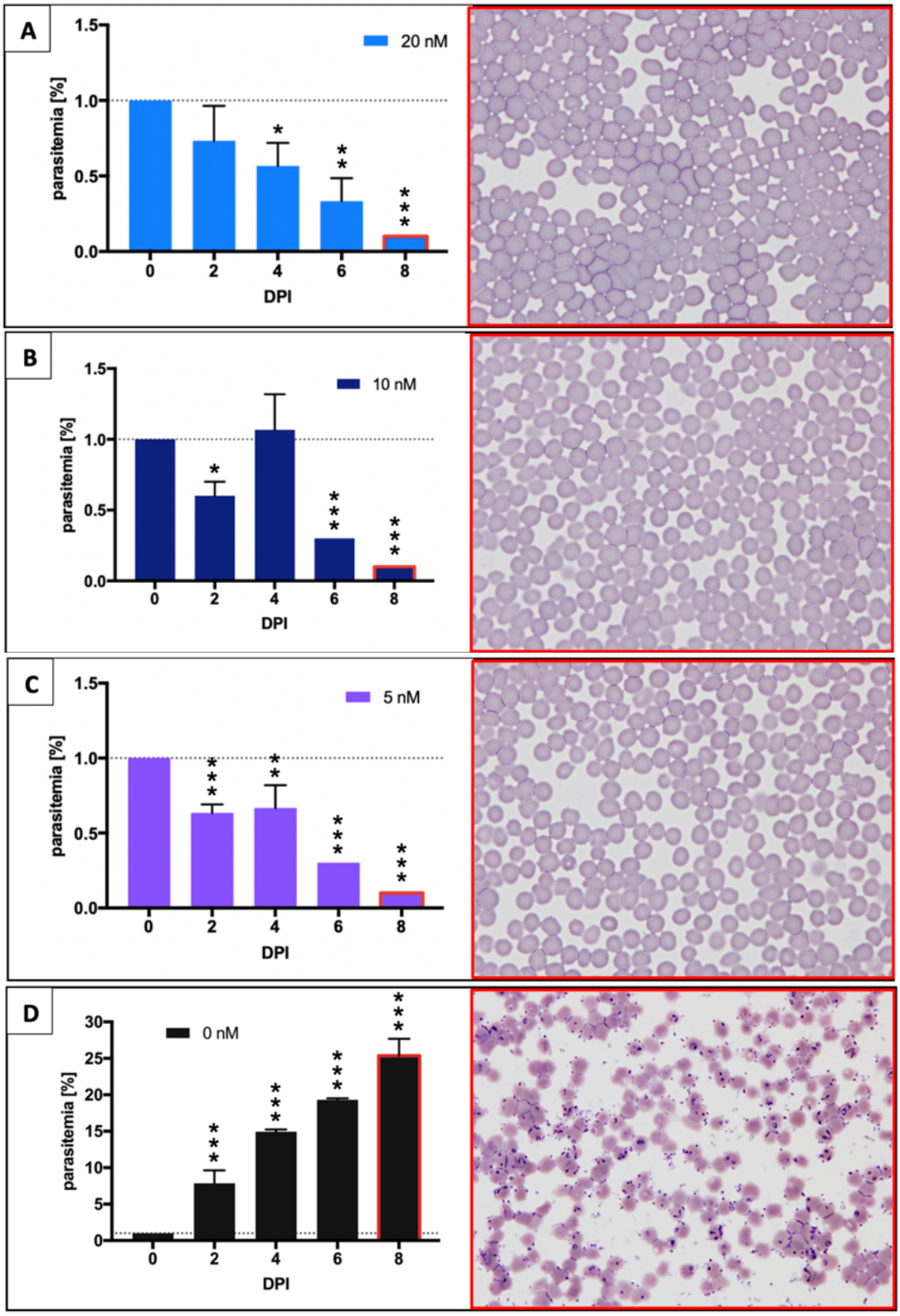

Figure 7: Treatment of *B. divergens in vitro* culture with effective concentrations of WR99210 drug. (A-C) Drug concentrations of 20nM, 10nM, 5nM, respectively, and untreated/negative control (D). Parasites were cultivated *in vitro* in standard conditions (37°, 5% CO_2) for 8 consecutive days (Chapter C.1.). Blood smears were taken on 8th day of the experiment and were stained and visualized as described in Chapter C.1.1. The result represents the mean of three independent replicates and error bars indicate standard

deviations. Parasitemia was determined by flow cytometry, for staining cultures *B. divergens* was used for EthD-1 as explained in Chapter C.1.2. The grey dotted line represents initial parasitemia (1%); DPI - days post infection; * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Statistical test performed: one-way ANOVA.

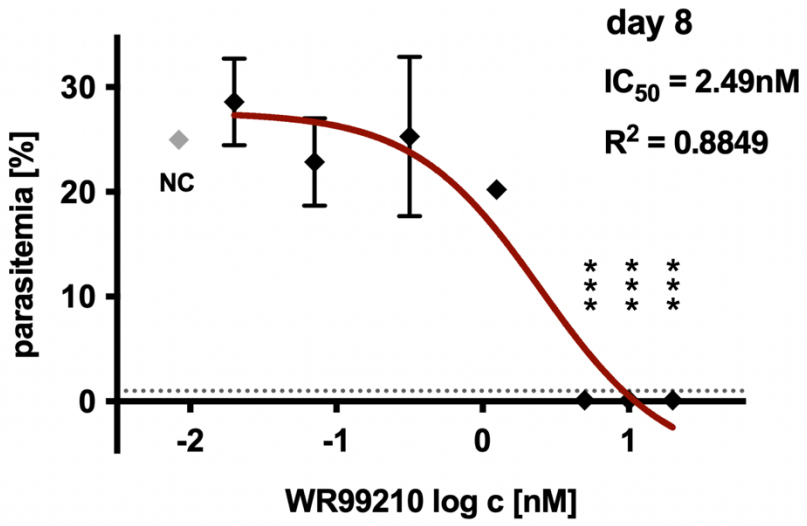

Figure 8: Determination of WR99210 IC_{50} value in *B. divergens in vitro* culture. The IC_{50} was calculated from respective *B. divergens* parasitemia levels on the 8th day by non-linear regression using a dose-response curve as well as the factor of regression (R^2). Prior the analysis, individual concentrations were transformed (log c). The result represents the mean of three independent replicates and error bars indicate standard deviations. For some points, error bars are invisible due to the minimal variance of the multiplicate. The grey dotted line represents initial parasitemia (1%); NC – negative control; IC_{50} – half-maximal inhibitory concentration; *** = $p < 0.001$. Statistical test performed: one-way ANOVA.

2. Inhibitory effect of G-418 and determination of IC_{50} value

The inhibitory effect of G-418 selection drug on *in vitro B. divergens* culture was analyzed by parasitemia determination using flow cytometry (Chapter C.1.2.). Parasitemia levels varying due to different concentrations of the selection drug from all five time points (2, 4, 6, 8, 10 DPI) are depicted in Figure 9. In comparison to higher concentration levels, the scale of concentrations 1000, 750, 500, and 250 µg/ml didn't have an influence on parasite's growth *in vitro,* and the AUC values were not statistically significantly different from the negative control (NC) AUC. The inhibitory effect of G-418 was demonstrated at a concentration of 1500 µg/ml, when parasitemia on 10 DPI was 6.5 % ± 1.4. The AUC value (86.8 ± 9.7) analyzed from the overall parasite growth curve was significantly different ($p < 0.001$) compared to the AUC of NC (129 ± 13.1). The same effect was demonstrated at a concentration of 2000 µg/ml, where the parasitemia on 10 DPI was 3.7 % ± 0.3, and the AUC

(64.2 ± 5) was significantly different ($p < 0.001$) compared to the AUC of NC. The concentration of 3000 μg/ml was shown to be inhibitory as well, with parasitemia 2.2 % ± 0.6 on 10 DPI, and the AUC (33.5 ± 2) was again confirmed as significantly different ($p < 0.001$) when compared to the AUC of NC. Figure 10 compares changes in parasitemia for the NC and for cultures treated with G-418 medium at 3000, 2000, and 1500 μg/ml including the display of 10 DPI blood culture smears. Differences in parasitemia were observed at effective G-418 concentrations of 3000, 2000, and 1500 μg/ml compared to the initial parasitemia (1%). Statistically significant differences were observed for the highest concentration of G-418 tested (3000 μg/ml) for all intervals analyzed: 2 DPI (4 % ± 0.2, $p < 0.001$), 4 DPI (5 % ± 0.8, $p < 0.001$), 6 DPI (3.2 % ± 0.2, $p < 0.01$), and 8 DPI (2.8 % ± 0.4, $p < 0.05$), except for 10 DPI (2.2 % ± 0.6), where parasitemia decreasing trend occurred. Cultures treated with 2000 μg/ml concentration of drug showed statistically significant differences in parasitemia for all intervals analyzed: 2 DPI (7 % ± 0.5, $p < 0.001$), 4 DPI (12 % ± 1, $p < 0.001$), 6 DPI (5.9 % ± 1.8, $p < 0.01$), and 8 DPI (4.8 % ± 0.7, $p < 0.05$), except for 10 DPI (3.7 % ± 0.3), where also parasitemia decreasing trend occurred. Last statistically significant differences were observed for 1500 μg/ml concentration of G-418 for all intervals analyzed: 2 DPI (7.3 % ± 0.6, $p < 0.001$), 4 DPI (14.9 % ± 1.3, $p < 0.05$), 6 DPI (10.3 % ± 1.6, $p < 0.001$), 8 DPI (7.2 % ± 0.3, $p < 0.001$) and 10 DPI (6.5 % ± 1.4, $p < 0.01$). Figure 11 shows the statistical determination of the IC_{50} value on 10 DPI after triplicates were averaged, their standard deviation was determined, and the data was normalized (logarithmized). Final IC_{50} was 1293 μg/ml and R^2 was 0.9279. G-418 concentration of 3000 μg/ml appears optimal for selective cultivation of parasites transfected with resistance holding plasmid. No toxic effect of G-418 on red blood cells was observed (Figure 10).

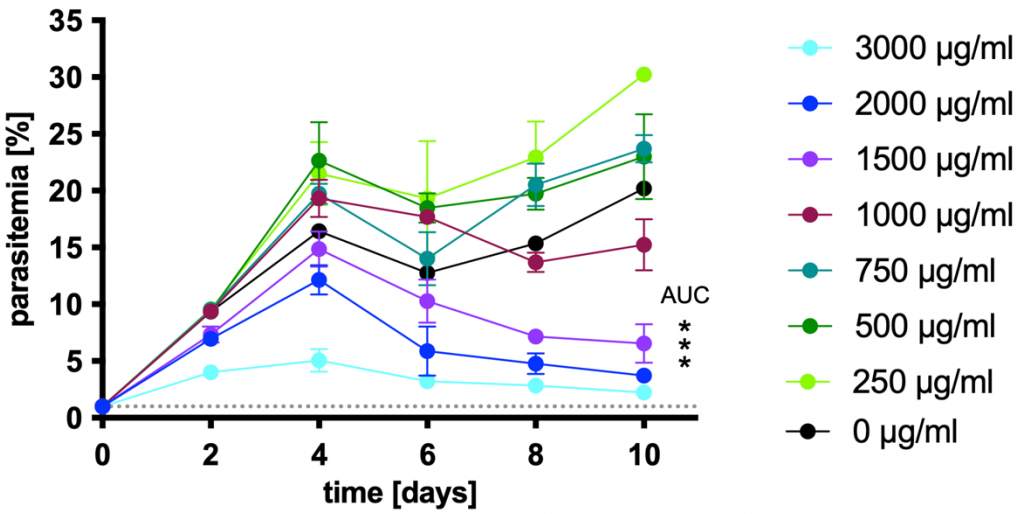

Figure 9: Treatment of *B. divergens in vitro* culture with different concentrations of G-418 drug. Parasites were cultivated *in vitro* in standard conditions (37°, 5% CO_2) for 10 consecutive days. The result represents mean of three independent replicates and error bars indicate standard deviations. The grey dotted line represents initial parasitemia (1%); DPI - days post infection; AUC - area under the curve; *** = $p < 0.001$. Statistical test performed: one-way ANOVA.

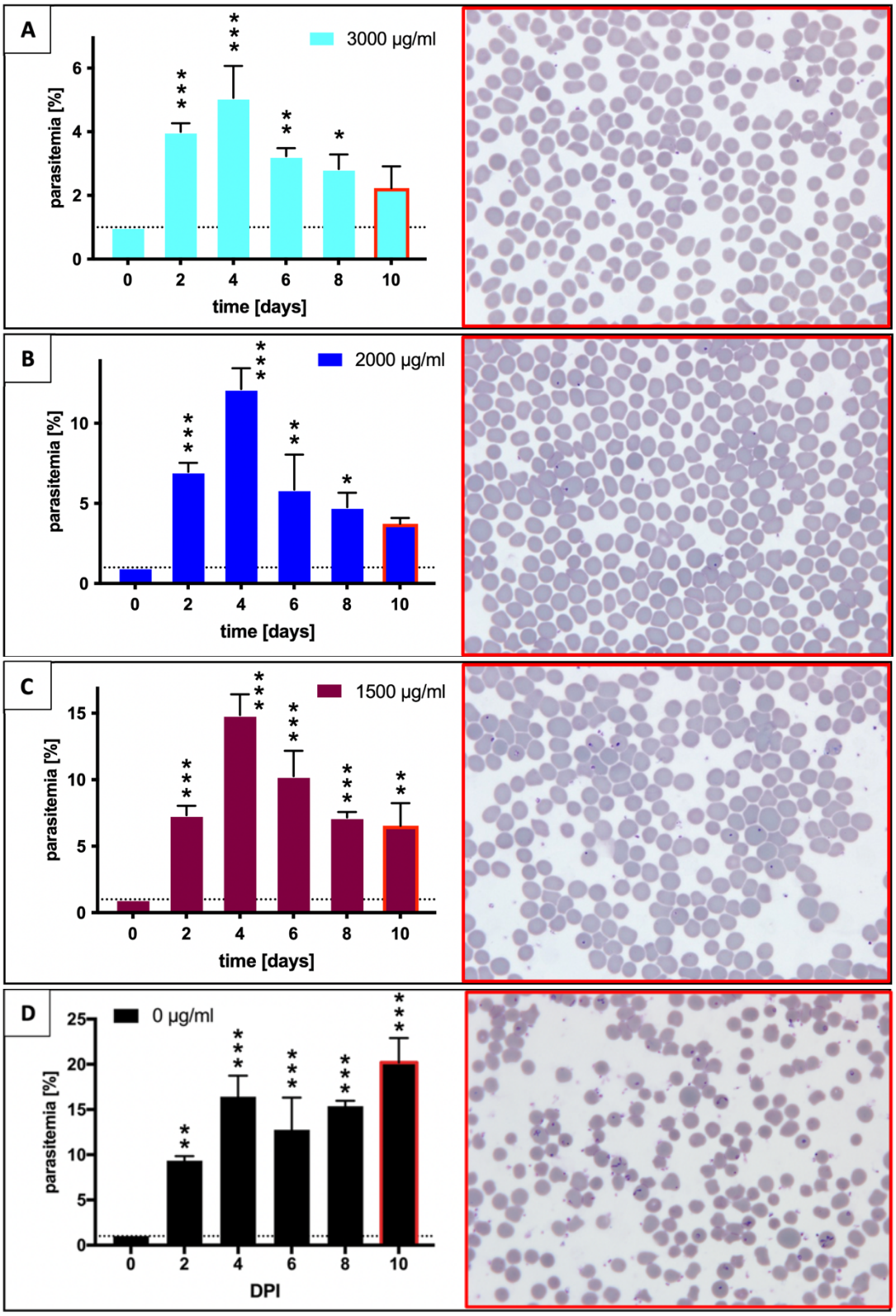

Figure 10: Treatment of *B. divergens in vitro* culture with effective concentrations of G-418 drug. (A-C) Drug concentrations of 3000, 2000 and 1500 µg/ml and untreated (negative) control (D). Parasites were cultivated *in vitro* in standard conditions (37°, 5% CO_2) for 10 consecutive days (Chapter C.1.). Blood smears were taken on 10th day of the experiment and were stained and visualized as described in Chapter C.1.1. The result

represents the mean of three independent replicates and error bars indicate standard deviations. Parasitemia was determined by flow cytometry, for staining cultures *B. divergens* was used for EthD-1 as explained in Chapter C.1.2. The grey dotted line represents initial parasitemia (1%); DPI - days post infection; * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Statistical test performed: one-way ANOVA.

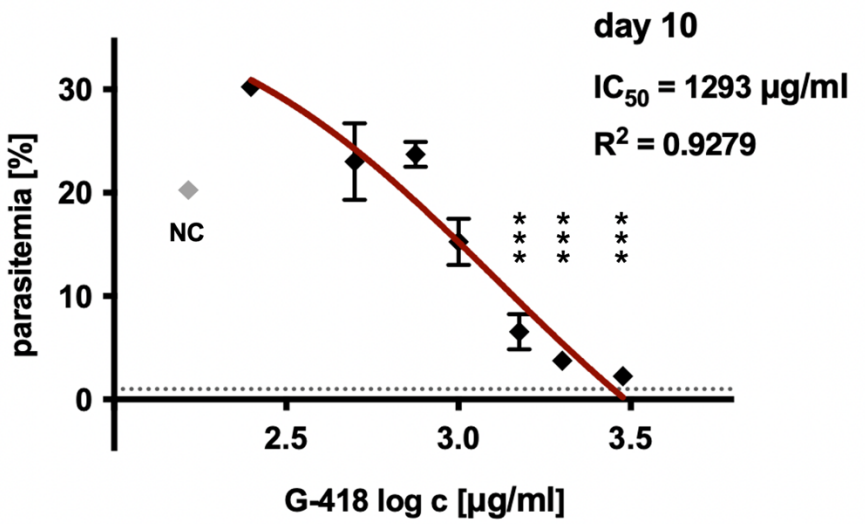

Figure 11: Determination of G-418 IC_{50} value in *B. divergens in vitro* culture. The IC_{50} was calculated from respective *B. divergens* parasitemia levels on the 10th day by non-linear regression using a dose-response curve as well as the factor of regression (R^2). Prior the analysis, individual concentrations were transformed (log c). The result represents the mean of three independent replicates and error bars indicate standard deviations. For some points, error bars are invisible due to the minimal variance of the multiplicate. The grey dotted line represents initial parasitemia (1%); NC – negative control, IC_{50} - half-maximal inhibitory concentration; *** = $p < 0.001$. Statistical test performed: one-way ANOVA.

3. Identification and sequencing of eligible recipient locus (HR)

B. divergens DNA was isolated and the Bdiv_04560c segment was PCR amplified using designed insert specific primers (Table 4). The PCR product was cloned and transformed into competent cells. Resulting bacterial clones were PCR tested for holding the DNA insert and further cultivated. Plasmid DNA of four positive bacterial clones was sequenced and isolated. The alignment of the resulting consensus clone sequence and reference sequence (PiroplasmaDB), presented in the appendix at the end of the thesis (Chapter H.1.), displays only one single nucleotide polymorphism (SNP) (Chapter H.1.).

4. DiCre cassette vector design

The design of the plasmid vector for the transformation of *B. divergens* for stable expression of the two DiCre subunits is depicted in Figure 12. The construct was designed to hold the CRE59 subunit with *B. divergens*-specific 5'UTR region containing promoter HSP 70 (heat-shock protein 70) and 3'UTR region containing terminator CAM (calmodulin), CRE60 subunit with *B. divergens*-specific 5'UTR region containing promoter ACT (actin), and 3'UTR region containing terminator CRT (chloroquine transporter resistance), hDHFR resistance (human dihydrofolate reductase) with *B. divergens*-specific 5'UTR region containing promoter CAM (calmodulin) and 5'HR 1 and 3'HR 2 (homology regions) with recipient intragenomic site. Red triangles represent unique one-cutter restriction sites, which were located around 5'HR 1, 5'UTR region containing promoters HSP 70, ACT, CAM and 3'HR 2. Restriction sites were designed so that the sequence between them could be eventually replaced by another sequence in the future. The pUC57-BsaI-Free Vector was picked as the plasmid backbone. The total size of the whole plasmid construct is 14 061bp.

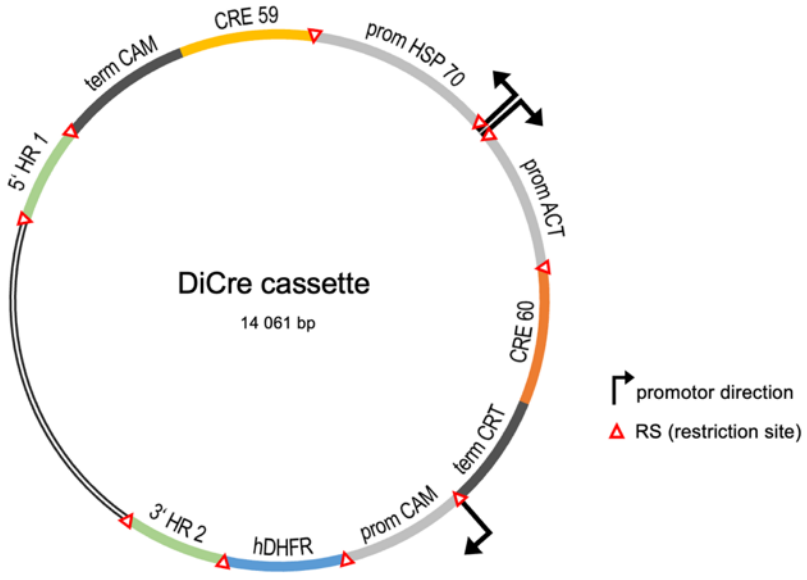

Figure 12: The *B. divergens* DiCre plasmid vector containing restriction sites. The plasmid contains CRE59 subunit (yellow box) with a 5'UTR region containing promoter HSP 70 (heat-shock protein 70) and a 3'UTR region containing terminator CAM (calmodulin), CRE60 subunit (orange box) with a 5'UTR region containing promoter ACT (actin) and 3'UTR region containing terminator CRT (chloroquine transporter resistance), hDHFR resistance (human dihydrofolate reductase; blue box) with a 5'UTR region containing promoter CAM (calmodulin) and 5'HR 1 and 3'HR 2 (homology regions; green boxes) specifically. The backbone is shown as a black double continuous line. Arrows represent the orientation of promoters and red triangles restriction sites. The total size of the original DNA construct is 14 061bp.

5. Gibson assembly of the DiCre plasmid

Although the whole plasmid depicted in Figure 12 was originally ordered as one circular synthetic DNA construct from the external company (Gene Universal), due to unspecified problems during the DiCre cassette synthesis, the construct was delivered as two different plasmids holding two separated parts of the DiCre cassette tagged as A and B, respectively (Figure 13 A, B). Thus, we decided to reconstruct the whole plasmid by PCR reamplification of DNA fragments from template plasmids A and B by Gibson assembly. The scheme for the sequential assembly is outlined in Figure 14. Primarily, the whole CRE60 subunit was assembled (Chapter C.8.) as shown in Figure 14, A. Specific Gibson primers with appropriate overlaps (Table 6) were designed to incorporate the rest of the CRE60 subunit. After successful assembly of the DiCre cassette plasmid with the whole CRE60, the rest of the cassette was incorporated into the plasmid (Figure 14, B). Due to the size of the whole DNA fragment, two strategies were selected – incorporation of one large segment (Figure 14, PCR product 4) and incorporation of two smaller segments (Figure 14, PCR product 2 and PCR product 3). The DiCre plasmid containing the whole DiCre cassette is now ready for the transfection of *B. divergens* parasite cultures.

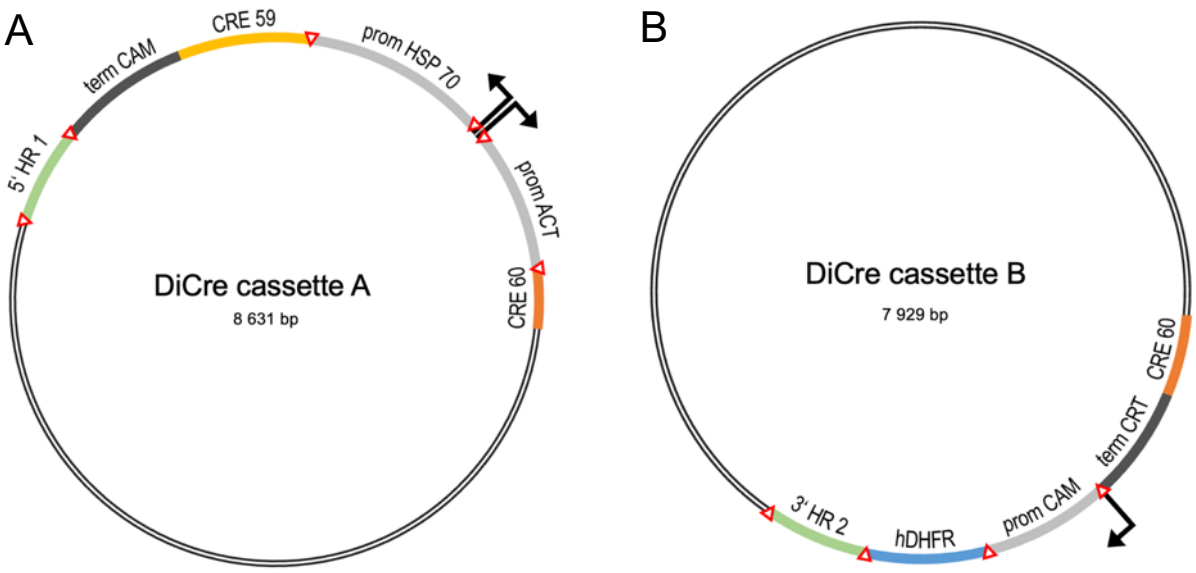

Figure 13: Delivered DiCre cassette vectors A and B containing synthetized DiCre cassette fragments. (A) The DiCre cassette A plasmid contains CRE59 subunit (yellow box) with a 5'UTR region containing promoter HSP 70 (heat-shock protein 70) and a 3'UTR region containing terminator CAM (calmodulin), 5'HR 1 (homology region; green box) and part of CRE60 subunit (orange box) with promoter ACT (actin). **(B)** The DiCre cassette B plasmid contains hDHFR resistance (human dihydrofolate reductase; blue box) with a 5'UTR region

containing promoter CAM (calmodulin), 3'HR 2 (homology region; green box) and part of CRE60 subunit (orange box) with a 3'UTR region containing terminator CRT (chloroquine transporter resistance). Backbones are shown as a black double continuous line. The arrows represent the orientation of promoters and triangles restriction sites. The size of the DiCre cassette plasmid A and B is 8 631bp and 7 929bp, respectively.

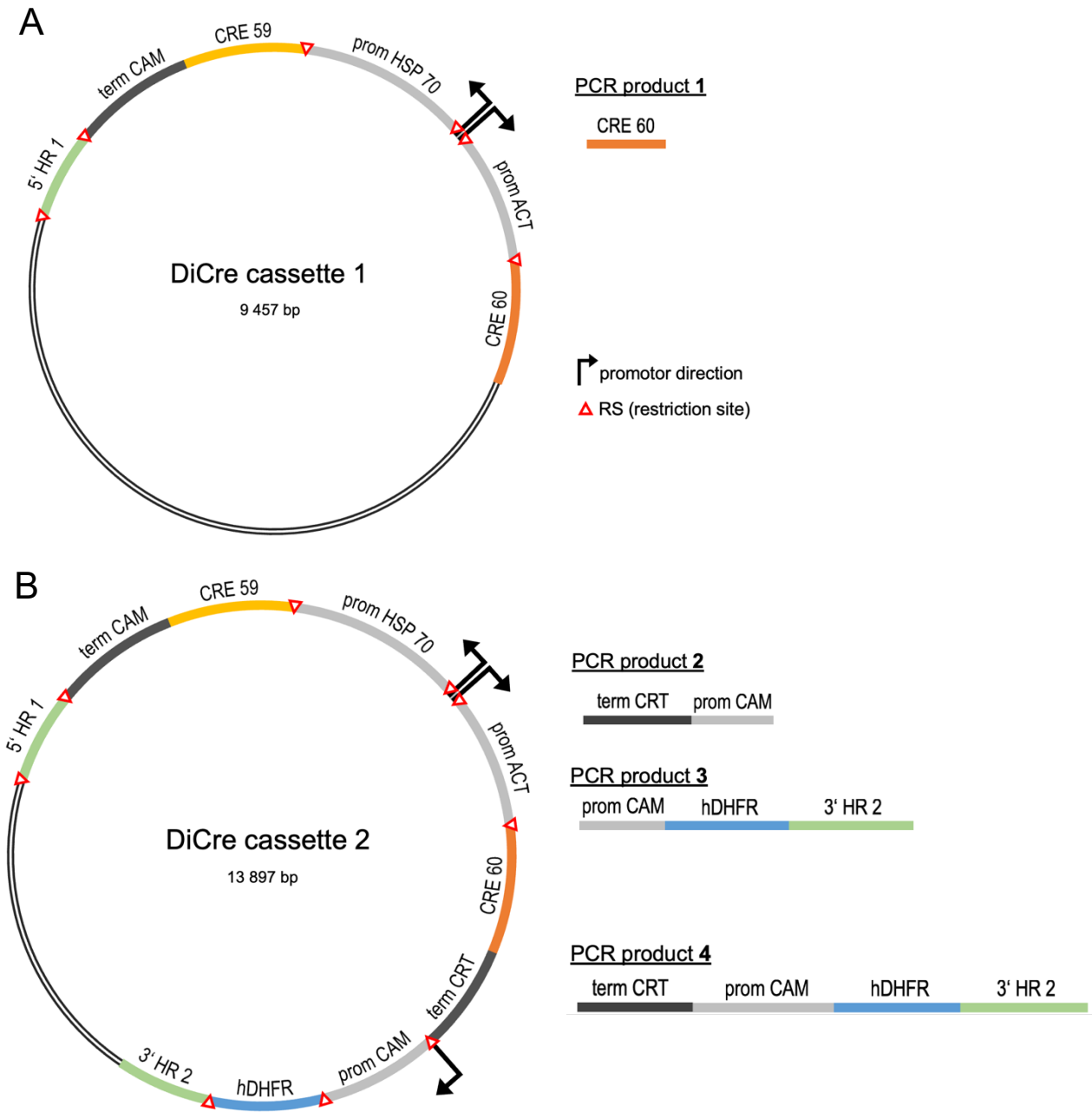

Figure 14: Outline of the PCR and Gibson assembly reconstruction of the original plasmid construct from DiCre cassette vectors A and B. (A) Primarily, the DiCre cassette plasmid containing DiCre cassette with whole CRE 60 (**DiCre cassette 1**) sequence was assembled by incorporating one insert (PCR product 1). The size of the assembled construct was 9 457 bp. **(B)** In the next step, the rest of the DiCre cassette (**DiCre cassette 2**) was incorporated into the plasmid formed in the previous step by incorporating either two short

inserts (PCR product 2 and PCR product 3) or one long insert (PCR product 4). The total size of the final assembled plasmid is 13 897 bp. Backbones are shown as a black double continuous line. Arrows represent the orientation of promoters and red triangles restriction sites. *5'HR 1* - 5' homology region 1, *term CAM* - 3'UTR region containing terminator CAM (calmodulin), *CRE59* - one of the DiCre subunits, *prom HSP 70* - 5'UTR region containing promoter HSP 70 (heat-shock protein 70), *prom ACT* - 5'UTR region containing promoter ACT (actin), *CRE60* - one of the DiCre subunits, *term CRT* - 3'UTR region containing terminator CRT (chloroquine transporter resistance), *prom CAM* - 5'UTR region containing promoter CAM (calmodulin), *hDHFR* - hDHFR resistance (human dihydrofolate reductase), *3'HR 2* - 3' homology region 2.

E. DISCUSSION

Babesiosis caused by *B. divergens* is considered an emerging life-threatening zoonotic disease due to the growing incidence of human infections with *B. divergens* in Europe (Hildebrandt et al., 2021). *B. divergens* is one of the less studied parasites within the *Babesia* group, and many fundamental processes in the biology of this parasite remain poorly understood (Lobo et al., 2019). One of the approaches, which help us to better understand the biology of *Babesia* parasites and their interaction with hosts or vectors, is functional genomics. The introduction of some of the rapidly developing functional genomic tools to *Babesia* would also help to identify and validate novel potential therapeutic targets for the yet missing specific treatment for babesiosis (Hildebrandt et al., 2021). Currently, an increasing number of techniques for parasite *in vitro* cultivation, transfection, and availability of genome sequences of various *Babesia* species has enabled the introduction of important functional genomics tools. Even though a wide number of genetic manipulation methods has already been established for major Apicomplexan model organisms of *Plasmodium* and *Toxoplasma* genera, only a limited number of them have been made available for *Babesia* (Hakimi et al., 2021). Stable transfection lines have already been created for some *Babesia* spp., including *B. bovis*, *B. bigemina, B. gibsoni, B. ovata* (Hakimi et al., 2016; Liu et al., 2018; Suarez & McElwain, 2010). Although methods for *B. divergens in vitro* cultivation and genome sequence are known, a stable transfection system has not yet been established (Jalovecká et al., 2016; Young et al., 2019).

Since non-inducible knock-out systems cannot be used for functional characterization of essential *Babesia* genes, the conditional (inducible) expression systems represent a unique strategy to study *Babesia* enzymes/proteins playing crucial roles in parasite invasion, growth and egress from host cells (Hakimi et al., 2021). DiCre system with modified dimerizable Cre recombinase is one of the methods of conditional genome editing and has been successfully introduced to *T. gondii* and *P. falciparum* (Andenmatten et al., 2013; Collins et al., 2013). For the successful introduction of the DiCre system to *B. divergens*, the plasmid vector holding the DiCre cassette and regions for homologous recombination into a suitable recipient region in the *B. divergens* genome (*B. divergens*-specific homology regions 5'HR 1 and 3'HR 2) needs to be designed and produced. In this thesis, I have thus designed and produced the DiCre plasmid containing sequences for hDHFR (gene for WR99210 resistance), CRE59 and CRE60 subunits with *B. divergens*-specific 5'UTR regions containing promoters and 3'UTR region containing terminators was constructed. Unique restriction sites were added to the DiCre

plasmid so that the 5'UTR and 3'UTR regions may be replaced for future DiCre cassette editing.

Based on a literature search, we have selected the 6-Cys-E gene (Bdiv_04560c) as the homologous recipient locus of the *B. divergens* genome for the DiCre cassette incorporation. This gene is highly homologous to the p230p gene (PF3D7_0208900) that has been used as a stable recipient locus for the DiCre cassette transfections of the *Babesia* related malaria parasite *P. falciparum* (Knuepfer et al., 2017). Additionally, targeted disruption of *B. bovis* Bbo-6cys-E gene (BBOV_II006600), an orthologue of Bdiv_04560c and PF3D7_0208900 genes, indicated its dispensability for parasite blood stages (Silva et al., 2011). The Bdiv_04560c sequence was PCR amplified and then verified via sequencing. The resulting sequence was compared to the reference sequence obtained from PiroplasmaDB and one SNP was detected. Even though this mutation changes one amino acid in primary protein structure residue (arginine to serine), no effect on gene expression is expected.

Besides the plasmid backbone, including an antibiotic resistance gene for ampicillin and other sequences essential for multiplication in bacteria, the plasmid for developing transfection lines must also contain a selection marker, a resistance gene against a selection drug that inhibits parasite growth, the locus for homology recombination and DiCre cassette. The selection drug (e.g., BSD, WR99210, G-418) is used to apply selection pressure on cells, so the parasite cells are forced to retain the plasmid (Asada et al., 2015; Hakimi et al., 2021). For our purpose, we have decided to use WR99210 and G-418 selection markers that have been demonstrated to work efficiently in *in vitro* cultivation of other *Babesia* species (Asada et al., 2015) and *P. falciparum*, respectively (Birnbaum et al., 2017). The usage of two selection markers is required to introduce two separate plasmids necessary for DiCre functioning. One of the plasmids must carry a DiCre cassette containing the sequence for dimerizable Cre recombinase, which catalyzes the excision of a selected DNA segment that is flanked by introduced loxP sites. The second plasmid contains the sequence for loxP sites themselves. Therefore, we designed two independent experiments to test the effect of different concentrations of two selection drugs WR99210 and G-418 on *B. divergens in vitro* culture. WR99210 effectively inhibited the growth of *B. divergens in vitro* at 20, 10, and 5 nM concentrations when the AUC (area under the curve) was significantly ($p < 0.001$) reduced compared to the AUC of negative control (Figure 7). 10 or 5 nM concentrations of WR99210 drug are commonly used for *P. falciparum* (Knuepfer et al., 2017) and also for other *Babesia* species (Asada et al., 2012, 2015; Hakimi et al., 2016; Liu et al., 2018) selection systems.

G-418 effectively inhibited the growth of *B. divergens in vitro* at concentrations of 3000, 2000, and 1500 µg/ml when the AUC was significantly ($p < 0.001$) reduced compared to the AUC of negative control (Figure 10). 400 µg/ml concentration of G-418 drug is commonly used for *P. falciparum* selection system (Birnbaum et al., 2017; Wichers et al., 2019). Additionally, the G-418 selection marker is known for the delay in the growth inhibition effect (Birnbaum et al., 2017), thus significant differences in growth inhibition were observed during last tested days of the experiment. The standard method used to determine parasitemia in *B. divergens in vitro* culture is the evaluation of Giemsa-stained blood smears under light microscopy (Jalovecká et al., 2017). Due to the time-consuming nature of processing a larger number of samples, the method of parasitemia determination by flow cytometry was selected. The parasitemia determination by flow cytometry is established for *P. falciparum* (Jang et al., 2014) and was optimized by the previous bachelor's student Šárka Borsodi (Borsodi, 2021). Parasitemia was determined for each of the selection drug (WR99210 and G-418) concentrations, and the IC_{50} value, which is essential to establish in order to use a selection marker, was statistically determined from the data obtained (Figures 8 and 11). No toxic effect of WR99210 or G-418 on red blood cells was monitored (Figures 7 and 10). Therefore, WR99210 concentration of 5 nM and G-418 concentration of 3000 µg/ml will be used as optimal for transfection systems of *B. divergens*.

Even though the plasmid construct was originally ordered as one product, the delivered synthetic DNA vector was due to unspecified problems of the company (Gene Universal) delivered in two plasmids. Therefore, a plasmid design procedure (Figure 14) was outlined to assemble the whole construct using several subsequent Gibson assembly reactions containing the appropriate PCR products with the appropriate overlaps (Table 6). The Gibson assembly system is also a commonly used approach for plasmid construction of Apicomplexa transfection systems (Wagner et al., 2013). Gibson assembly is a robust universal DNA assembly method that removes some of the disadvantages of conventional cloning, such as the incompatibility of restriction enzymes, the need for ligation, and the limited size of segments (McAllaster et al., 2016) and moreover, it is timesaving, since the whole construct can be assembled in a single step and directly transformed into competent cells. The DiCre plasmid containing the whole DiCre cassette was successfully assembled and will be introduced into the *B. divergens* cell culture via subsequent nucleofection. To verify that DiCre recombinase-expressing *B. divergens* lineage can mediate rapamycin-induced excision, wild type (WT) and

parental DiCre lineages will be transfected with the green fluorescent protein (GFP) reporter episomal plasmid where the GFP gene will be flanged by loxP sites.

The establishment of a DiCre *B. divergens* line will be universally applicable to functionally validate any genes/proteins of interest. This represents an excellent experimental platform that will well serve the whole community of *Babesia* researchers and will also be further used to achieve goals within currently running projects in the laboratory, e.g. to functionally analyze calcium-dependent protein kinases (CDPKs) that regulate key biological processes of the host cell attachment and invasion, gliding motility and parasite egress (Bisio & Soldati-Favre, 2019; Lim et al., 2012). CDPKs are considered as one of the most promising new therapeutic targets for diseases caused by Apicomplexa (Choi et al., 2020; Keyloun et al., 2014) since they are absent in animal hosts which excludes on-target toxicity of CDPKs inhibitors to host cells (Billker et al., 2009). Plasmepsin-like proteases of *Babesia* with essential involvement in egress and invasion during blood-stage babesiosis are other great candidates to be functionally characterized by the Cre/loxP technology (Šnebergerová et al., 2021).

F. CONCLUSION

- Eligible recipient locus (homology region) in the *B. divergens* genome into which a DiCre cassette will be integrated via homology recombination was identified and sequenced.
- WR99210 and G-418 were validated as selection markers and optimal concentrations for their use in *B. divergens in vitro* cultures were determined.
- Plasmid vector for the generation of parental DiCre *B. divergens* lineage was designed and constructed.

G. REFERENCES

AbouLaila, M., Nakamura, K., Govind, Y., Yokoyama, N., & Igarashi, I. (2010). Evaluation of the in vitro growth-inhibitory effect of epoxomicin on *Babesia* parasites. *Veterinary Parasitology*, *167*, 19–27. https://doi.org/10.1016/j.vetpar.2009.09.049

Andenmatten, N., Egarter, S., Jackson, A. J., Jullien, N., Herman, J.-P., & Meissner, M. (2013). Conditional genome engineering in *Toxoplasma gondii* uncovers alternative invasion mechanisms. *Nature Methods*, *10*, 125–127. https://doi.org/10.1038/nmeth.2301

Arisue, N., & Hashimoto, T. (2015). Phylogeny and evolution of apicoplasts and apicomplexan parasites. *Parasitology International*, *64*, 254–259. https://doi.org/10.1016/j.parint.2014.10.005

Asada, M., Tanaka, M., Goto, Y., Yokoyama, N., Inoue, N., & Kawazu, S. I. (2012). Stable expression of green fluorescent protein and targeted disruption of thioredoxin peroxidase-1 gene in *Babesia bovis* with the WR99210/dhfr selection system. *Molecular and Biochemical Parasitology*, *181*, 162–170. https://doi.org/10.1016/j.molbiopara.2011.11.001

Asada, M., Yahata, K., Hakimi, H., Yokoyama, N., Igarashi, I., Kaneko, O., Suarez, C. E., & Kawazu, S. (2015). Transfection of *Babesia bovis* by Double Selection with WR99210 and Blasticidin-S and Its Application for Functional Analysis of Thioredoxin Peroxidase-1. *Public Library of Science One*, *10*, e0125993. https://doi.org/10.1371/journal.pone.0125993

Basler, M., Mundt, S., Bitzer, A., Schmidt, C., & Groettrup, M. (2015). The immunoproteasome: A novel drug target for autoimmune diseases. *Clinical and Experimental Rheumatology*, *33*, 74–79.

Bibo-Verdugo, B., Jiang, Z., Caffrey, C. R., & O'Donoghue, A. J. (2017). Targeting proteasomes in infectious organisms to combat disease. In *Federation of European Biochemical Societies Journal*, *284*, 1503–1517). https://doi.org/10.1111/febs.14029

Billker, O., Lourido, S., & Sibley, L. D. (2009). Calcium-Dependent Signaling and Kinases in Apicomplexan Parasites. *Cell Host and Microbe*, *5*, 612–622. https://doi.org/10.1016/j.chom.2009.05.017

Birnbaum, J., Flemming, S., Reichard, N., Soares, A. B., Mesén-Ramírez, P., Jonscher, E., Bergmann, B., & Spielmann, T. (2017). A genetic system to study *Plasmodium*

falciparum protein function. *Nature Methods*, *14*, 450–456. https://doi.org/10.1038/nmeth.4223

Bisio, H., & Soldati-Favre, D. (2019). Signaling Cascades Governing Entry into and Exit from Host Cells by *Toxoplasma gondii*. *Annual Review of Microbiology*, *73*, 579–599. https://doi.org/10.1146/annurev-micro-020518-120235

Borsodi, Š. (2021). Vývoj fluorescenční reportérové linie *Babesia divergens*. Jihočeská univerzita v Českých Budějovicích, Přírodovědecká fakulta.

Briquet, S., Gissot, M., & Silvie, O. (2021). A toolbox for conditional control of gene expression in apicomplexan parasites. *Molecular Microbiology*, *117*, 618–631. https://doi.org/10.1111/mmi.14821

Brown, W. C., Norimine, J., Goff, W. L., Suarez, C. E., & Mcelwain, T. F. (2006). Prospects for recombinant vaccines against *Babesia bovis* and related parasites. *Parasite Immunology*, *28*, 315–327. https://doi.org/10.1111/j.1365-3024.2006.00849.x

Burki, F., Inagaki, Y., Bråte, J., Archibald, J. M., Keeling, P. J., Cavalier-Smith, T., Sakaguchi, M., Hashimoto, T., Horak, A., Kumar, S., Klaveness, D., Jakobsen, K. S., Pawlowski, J., & Shalchian-Tabrizi, K. (2009). Large-Scale Phylogenomic Analyses Reveal That Two Enigmatic Protist Lineages, Telonemia and Centroheliozoa, Are Related to Photosynthetic Chromalveolates. *Genome Biology and Evolution*, *1*, 231–238. https://doi.org/10.1093/gbe/evp022

Choi, R., Hulverson, M. A., Huang, W., Vidadala, R. S. R., Whitman, G. R., Barrett, L. K., Schaefer, D. A., Betzer, D. P., Riggs, M. W., Doggett, J. S., Hemphill, A., Ortega-Mora, L. M., McCloskey, M. C., Arnold, S. L. M., Hackman, R. C., Marsh, K. C., Lynch, J. J., Freiberg, G. M., Leroy, B. E., … van Voorhis, W. C. (2020). Bumped Kinase Inhibitors as therapy for apicomplexan parasitic diseases: lessons learned. *International Journal for Parasitology*, *50*, 413–422. https://doi.org/10.1016/j.ijpara.2020.01.006

Collins, C. R., Das, S., Wong, E. H., Andenmatten, N., Stallmach, R., Hackett, F., Herman, J. P., Müller, S., Meissner, M., & Blackman, M. J. (2013). Robust inducible Cre recombinase activity in the human malaria parasite *Plasmodium falciparum* enables efficient gene deletion within a single asexual erythrocytic growth cycle. *Molecular Microbiology*, *88*, 687–701. https://doi.org/10.1111/mmi.12206

de Koning-Ward, T. F., Gilson, P. R., & Crabb, B. S. (2015). Advances in molecular genetic systems in malaria. *Nature Reviews Microbiology*, *13*, 373–387. https://doi.org/10.1038/nrmicro3450

de Waal, D. T., & Combrink, M. P. (2006). Live vaccines against bovine babesiosis. *Veterinary Parasitology*, *138*, 88–96. https://doi.org/10.1016/j.vetpar.2006.01.042

del Carmen Terrón, M., González-Camacho, F., González, L. M., Luque, D., & Montero, E. (2016). Ultrastructure of the *Babesia divergens* free merozoite. *Ticks and Tick-Borne Diseases*, *7*, 1274–1279. https://doi.org/10.1016/j.ttbdis.2016.07.001

Florin-Christensen, M., Suarez, C. E., Rodriguez, A. E., Flores, D. A., & Schnittger, L. (2014). Vaccines against bovine babesiosis: where we are now and possible roads ahead. *Parasitology*, *141*, 1563–1592. https://doi.org/10.1017/S0031182014000961

Garcia, C. R. S., de Azevedo, M. F., Wunderlich, G., Budu, A., Young, J. A., & Bannister, L. (2008). *Plasmodium* in the Postgenomic Era: New Insights into the Molecular Cell Biology of Malaria Parasites. *International Review of Cell and Molecular Biology*, *266*, 85–156. https://doi.org/10.1016/S1937-6448(07)66003-1

Gohil, S., Herrmann, S., Günther, S., & Cooke, B. M. (2013). Bovine babesiosis in the 21st century: Advances in biology and functional genomics. *International Journal for Parasitology*, *43*, 125–132. https://doi.org/10.1016/j.ijpara.2012.09.008

González, L. M., Estrada, K., Grande, R., Jiménez-Jacinto, V., Vega-Alvarado, L., Sevilla, E., Barrera, J. de la, Cuesta, I., Zaballos, Á., Bautista, J. M., Lobo, C. A., Sánchez-Flores, A., & Montero, E. (2019). Comparative and functional genomics of the protozoan parasite *Babesia divergens* highlighting the invasion and egress processes. *Public Library of Science Neglected Tropical Diseases*, *13*, e0007680. https://doi.org/10.1371/journal.pntd.0007680

Goswami, D., Minkah, N. K., & Kappe, S. H. I. (2019). Designer Parasites: Genetically Engineered *Plasmodium* as Vaccines To Prevent Malaria Infection. *The Journal of Immunology*, *202*, 20–28. https://doi.org/10.4049/jimmunol.1800727

Hakimi, H., Asada, M., & Kawazu, S. (2021). Recent Advances in Molecular Genetic Tools for *Babesia*. *Veterinary Sciences*, *8*, 222. https://doi.org/10.3390/vetsci8100222

Hakimi, H., Yamagishi, J., Kegawa, Y., Kaneko, O., Kawazu, S., & Asada, M. (2016). Establishment of transient and stable transfection systems for *Babesia ovata*. *Parasites and Vectors*, *9*, 171. https://doi.org/10.1186/s13071-016-1439-z

Healy, G., & Ristic, M. (2018). Human babesiosis. In *Babesiosis of Domestic Animals and Man*. https://doi.org/10.1201/9781351070027

Hildebrandt, A., Gray, J. S., & Hunfeld, K. P. (2013). Human Babesiosis in Europe: What clinicians need to know. In *Infection*, *41*, 1057–1072. https://doi.org/10.1007/s15010-013-0526-8

Hildebrandt, A., Hunfeld, K. P., Baier, M., Krumbholz, A., Sachse, S., Lorenzen, T., Kiehntopf, M., Fricke, H. J., & Straube, E. (2007). First confirmed autochthonous case of human *Babesia microti* infection in Europe. *European Journal of Clinical Microbiology and Infectious Diseases*, *26*, 595–601. https://doi.org/10.1007/s10096-007-0333-1

Hildebrandt, A., Zintl, A., Montero, E., Hunfeld, K.-P., & Gray, J. (2021). Human Babesiosis in Europe. *Pathogens*, *10*, 1165. https://doi.org/10.3390/pathogens10091165

Homer, M. J., Bruinsma, E. S., Lodes, M. J., Moro, M. H., Telford, S., Krause, P. J., Reynolds, L. D., Mohamath, R., Benson, D. R., Houghton, R. L., Reed, S. G., & Persing, D. H. (2000). A polymorphic multigene family encoding an immunodominant protein from *Babesia microti*. *Journal of Clinical Microbiology*, *38*, 362–368.

Hortua Triana, M. A., Márquez-Nogueras, K. M., Vella, S. A., & Moreno, S. N. J. (2018). Calcium signaling and the lytic cycle of the Apicomplexan parasite *Toxoplasma gondii*. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1865*, 1846–1856. https://doi.org/10.1016/j.bbamcr.2018.08.004

Hoshijima, K., Jurynec, M. J., & Grunwald, D. J. (2016). Precise genome editing by homologous recombination. *Methods in Cell Biology*, *135*, 121–147. https://doi.org/10.1016/bs.mcb.2016.04.008

Hunfeld, K. P., Hildebrandt, A., & Gray, J. S. (2008). Babesiosis: Recent insights into an ancient disease. *International Journal for Parasitology*, *38*, 1219–1237. https://doi.org/10.1016/j.ijpara.2008.03.001

Jalovecká, M., Bonsergent, C., Hajdušek, O., Kopáček, P., & Malandrin, L. (2016). Stimulation and quantification of *Babesia divergens* gametocytogenesis. *Parasites and Vectors*, *9*, 439. https://doi.org/10.1186/s13071-016-1731-y

Jalovecká, M., Hajdušek, O., Sojka, D., Kopáček, P., & Malandrin, L. (2018). The Complexity of Piroplasms Life Cycles. *Frontiers in Cellular and Infection Microbiology*, *8*, 248. https://doi.org/10.3389/fcimb.2018.00248

Jalovecká, M., Hartmann, D., Miyamoto, Y., Eckmann, L., Hajdušek, O., O'Donoghue, A. J., & Sojka, D. (2018). Validation of *Babesia* proteasome as a drug target. *International*

Journal for Parasitology: Drugs and Drug Resistance, *8*, 394–402. https://doi.org/10.1016/J.IJPDDR.2018.08.001

Jalovecká, M., Sojka, D., Ascencio, M., & Schnittger, L. (2019). *Babesia* Life Cycle – When Phylogeny Meets Biology. In *Trends in Parasitology*, *35*, 356–368. https://doi.org/10.1016/j.pt.2019.01.007

Jalovecká, M., Urbanová, V., Sojka, D., Malandrin, L., Šíma, R., Kopáček, P., & Hajdušek, O. (2017). Establishment of *Babesia microti* laboratory model and its experimental application. *Tick and Tick-Borne Pathogen Conference and 1st Asia Pacific Rickettsia Conference*, 140.

Jang, J. W., Kim, J. Y., Yoon, J., Yoon, S. Y., Cho, C. H., Han, E. T., An, S. S. A., & Lim, C. S. (2014). Flow Cytometric Enumeration of Parasitemia in Cultures of *Plasmodium falciparum* Stained with SYBR Green I and CD235A. *The Scientific World Journal*, *2014*, 1–6. https://doi.org/10.1155/2014/536723

Joseph, J. T., Purtill, K., Wong, S. J., Munoz, J., Teal, A., Madison-Antenucci, S., Horowitz, H. W., Aguero-Rosenfeld, M. E., Moore, J. M., Abramowsky, C., & Wormser, G. P. (2012). Vertical transmission of *Babesia microti*, United States. *Emerging Infectious Diseases*, *18*, 1318–1321. https://doi.org/10.3201/eid1808.110988

Jullien, N., Goddard, I., Selmi-Ruby, S., Fina, J.-L., Cremer, H., & Herman, J.-P. (2007). Conditional Transgenesis Using Dimerizable Cre (DiCre). *Public Library of Science One*, *2*, e1355. https://doi.org/10.1371/journal.pone.0001355

Keyloun, K. R., Reid, M. C., Choi, R., Song, Y., Fox, A. M. W., Hillesland, H. K., Zhang, Z., Vidadala, R., Merritt, E. A., Lau, A. O. T., Maly, D. J., Fan, E., Barrett, L. K., van Voorhis, W. C., & Ojo, K. K. (2014). The gatekeeper residue and beyond: homologous calcium-dependent protein kinases as drug development targets for veterinarian Apicomplexa parasites. *Parasitology*, *141*, 1499–1509. https://doi.org/10.1017/S0031182014000857

Khare, S., Nagle, A. S., Biggart, A., Lai, Y. H., Liang, F., Davis, L. C., Barnes, S. W., Mathison, C. J. N., Myburgh, E., Gao, M. Y., Gillespie, J. R., Liu, X., Tan, J. L., Stinson, M., Rivera, I. C., Ballard, J., Yeh, V., Groessl, T., Federe, G., Supek, F. (2016). Proteasome inhibition for treatment of leishmaniasis, Chagas disease and sleeping sickness. *Nature*, *537*, 229–233. https://doi.org/10.1038/nature19339

Knuepfer, E., Napiorkowska, M., van Ooij, C., & Holder, A. A. (2017). Generating conditional gene knockouts in *Plasmodium* - A toolkit to produce stable DiCre recombinase-

expressing parasite lines using CRISPR/Cas9. *Scientific Reports*, *7*, 3881. https://doi.org/10.1038/s41598-017-03984-3

Krause, P. J., Daily, J., Telford, S. R., Vannier, E., Lantos, P., & Spielman, A. (2007). Shared features in the pathobiology of babesiosis and malaria. *Trends in Parasitology*, *23*, 605–610. https://doi.org/10.1016/j.pt.2007.09.005

Li, H., O'Donoghue, A. J., van der Linden, W. A., Xie, S. C., Yoo, E., Foe, I. T., Tilley, L., Craik, C. S., da Fonseca, P. C. A., & Bogyo, M. (2016). Structure-and function-based design of *Plasmodium*-selective proteasome inhibitors. *Nature*, *530*, 233–236. https://doi.org/10.1038/nature16936

Li, H., Ponder, E. L., Verdoes, M., Asbjornsdottir, K. H., Deu, E., Edgington, L. E., Lee, J. T., Kirk, C. J., Demo, S. D., Williamson, K. C., & Bogyo, M. (2012). Validation of the proteasome as a therapeutic target in *Plasmodium* using an epoxyketone inhibitor with parasite-specific toxicity. *Chemistry and Biology*, *19*, 1535–1545. https://doi.org/10.1016/j.chembiol.2012.09.019

Lim, D. C., Cooke, B. M., Doerig, C., & Saeij, J. P. J. (2012). *Toxoplasma* and *Plasmodium* protein kinases: Roles in invasion and host cell remodelling. *International Journal for Parasitology*, *42*, 21–32. https://doi.org/10.1016/j.ijpara.2011.11.007

Liu, M., Adjou Moumouni, P. F., Asada, M., Hakimi, H., Masatani, T., Vudriko, P., Lee, S.-H., Kawazu, S., Yamagishi, J., & Xuan, X. (2018). Establishment of a stable transfection system for genetic manipulation of *Babesia gibsoni*. *Parasites and Vectors*, *11*, 260. https://doi.org/10.1186/s13071-018-2853-1

Lobo, C. A., Cursino-Santos, J. R., Singh, M., & Rodriguez, M. (2019). *Babesia divergens*: A Drive to Survive. *Pathogens*, *8*, 95. https://doi.org/10.3390/pathogens8030095

McAllaster, M. R., Sinclair-Davis, A. N., Hilton, N. A., & de Graffenried, C. L. (2016). A unified approach towards *Trypanosoma brucei* functional genomics using Gibson assembly. *Molecular and Biochemical Parasitology*, *210*, 13–21. https://doi.org/10.1016/j.molbiopara.2016.08.001

Mehlhorn, H., & Schein, E. (1993). The piroplasms: "A long story in short" or "Robert Koch has seen it." *European Journal of Protistology*, *29*, 279–293. https://doi.org/10.1016/S0932-4739(11)80371-8

Mulia, G. E., Picanço-Castro, V., Stavrou, E. F., Athanassiadou, A., & Figueiredo, M. L. (2021). Advances in the Development and the Applications of Nonviral, Episomal

Vectors for Gene Therapy. *Human Gene Therapy*, *32*, 1076–1095. https://doi.org/10.1089/hum.2020.310

Ord, R. L., & Lobo, C. A. (2015). Human Babesiosis: Pathogens, Prevalence, Diagnosis, and Treatment. *Current Clinical Microbiology Reports*, *2*, 173–181. https://doi.org/10.1007/s40588-015-0025-z

Rathinasamy, V., Poole, W. A., Bastos, R. G., Suarez, C. E., & Cooke, B. M. (2019). Babesiosis Vaccines: Lessons Learned, Challenges Ahead, and Future Glimpses. *Trends in Parasitology*, *35*, 622–635. https://doi.org/10.1016/j.pt.2019.06.002

Reynolds, J. M., el Bissati, K., Brandenburg, J., Günzl, A., & Mamoun, C. ben. (2007). Antimalarial activity of the anticancer and proteasome inhibitor bortezomib and its analog ZL3B. *BMC Clinical Pharmacology*. https://doi.org/10.1186/1472-6904-7-13

Sam-Yellowe, T. Y. (1996). Rhoptry organelles of the apicomplexa: Their role in host cell invasion and intracellular survival. *Parasitology Today*, *12*, 308–316. https://doi.org/10.1016/0169-4758(96)10030-2

Schetters, T. P. M., & Montenegro-James, S. (1995). Vaccines against babesiosis using soluble parasite antigens. *Parasitology Today*, *11*, 456–462. https://doi.org/10.1016/0169-4758(95)80059-X

Schnittger, L., Rodriguez, A. E., Florin-Christensen, M., & Morrison, D. A. (2012). *Babesia*: A world emerging. *Infection, Genetics and Evolution*, *12*, 1788–1809. https://doi.org/10.1016/J.MEEGID.2012.07.004

Schreeg, M. E., Marr, H. S., Tarigo, J. L., Cohn, L. A., Bird, D. M., Scholl, E. H., Levy, M. G., Wiegmann, B. M., & Birkenheuer, A. J. (2016). Mitochondrial Genome Sequences and Structures Aid in the Resolution of Piroplasmida phylogeny. *Public Library of Science One*, *11*, e0165702. https://doi.org/10.1371/journal.pone.0165702

Silva, M. G., Knowles, D. P., Mazuz, M. L., Cooke, B. M., & Suarez, C. E. (2018). Stable transformation of *Babesia bigemina* and *Babesia bovis* using a single transfection plasmid. *Scientific Reports*, *8*, 6096. https://doi.org/10.1038/s41598-018-23010-4

Silva, M. G., Knowles, D. P., & Suarez, C. E. (2016). Identification of interchangeable cross-species function of elongation factor-1 alpha promoters in *Babesia bigemina* and *Babesia bovis*. *Parasites and Vectors*, *9*, 576. https://doi.org/10.1186/s13071-016-1859-9

Silva, M. G., Ueti, M. W., Norimine, J., Florin-Christensen, M., Bastos, R. G., Goff, W. L., Brown, W. C., Oliva, A., & Suarez, C. E. (2011). *Babesia bovis* expresses Bbo-6cys-E,

a member of a novel gene family that is homologous to the 6-cys family of *Plasmodium*. *Parasitology International*, *60*, 13–18. https://doi.org/10.1016/j.parint.2010.09.004

Smith, T. G., Walliker, D., & Ranford-Cartwright, L. C. (2002). Sexual differentiation and sex determination in the Apicomplexa. In *Trends in Parasitology*, *18*, 315–323. https://doi.org/10.1016/S1471-4922(02)02292-4

Šnebergerová, P., Bartošová-Sojková, P., Jalovecká, M., & Sojka, D. (2021). Plasmepsin-like Aspartyl Proteases in *Babesia*. *Pathogens*, *10*, 1241. https://doi.org/10.3390/pathogens10101241

Soldati, D., Foth, B. J., & Cowman, A. F. (2004). Molecular and functional aspects of parasite invasion. *Trends in Parasitology*, *20*, 567–574. https://doi.org/10.1016/j.pt.2004.09.009

Sonenshine, D. E., & Roe, R. M. (2014). *Biology of ticks*, *2*. Oxford University Press.

Suarez, C. E., Alzan, H. F., Silva, M. G., Rathinasamy, V., Poole, W. A., & Cooke, B. M. (2019). Unravelling the cellular and molecular pathogenesis of bovine babesiosis: is the sky the limit? *International Journal for Parasitology*, *49*, 183–197. https://doi.org/10.1016/j.ijpara.2018.11.002

Suarez, C. E., Bishop, R. P., Alzan, H. F., Poole, W. A., & Cooke, B. M. (2017). Advances in the application of genetic manipulation methods to apicomplexan parasites. *International Journal for Parasitology*, *47*, 701–710. https://doi.org/10.1016/j.ijpara.2017.08.002

Suarez, C. E., & McElwain, T. F. (2010). Transfection systems for *Babesia bovis*: A review of methods for the transient and stable expression of exogenous genes. *Veterinary Parasitology*, *167*, 205–215. https://doi.org/10.1016/j.vetpar.2009.09.022

Timms, P. (1989). Development of babesial vaccines. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *83*, 73–79. https://doi.org/10.1016/0035-9203(89)90608-1

Vannier, E. G., Diuk-Wasser, M. A., ben Mamoun, C., & Krause, P. J. (2015). Babesiosis. *Infectious Disease Clinics of North America*, *29*, 357–370. https://doi.org/10.1016/j.idc.2015.02.008

Vannier, E., Gewurz, B. E., & Krause, P. J. (2008). Human Babesiosis. *Infectious Disease Clinics of North America*, *22*, 469–488. https://doi.org/10.1016/j.idc.2008.03.010

Votýpka, J., Modrý, D., Oborník, M., Šlapeta, J., & Lukeš, J. (2016). Apicomplexa. In *Handbook of the Protists*, 1–58. Springer International Publishing. https://doi.org/10.1007/978-3-319-32669-6_20-1

Wagner, J. C., Goldfless, S. J., Ganesan, S. M., Lee, M. C., Fidock, D. A., & Niles, J. C. (2013). An integrated strategy for efficient vector construction and multi-gene expression in *Plasmodium falciparum*. *Malaria Journal*, *12*, 373. https://doi.org/10.1186/1475-2875-12-373

White, M. W., & Suvorova, E. S. (2018). Apicomplexa Cell Cycles: Something Old, Borrowed, Lost, and New. *Trends in Parasitology*, *34*, 759–771. https://doi.org/10.1016/j.pt.2018.07.006

Wichers, J. S., Scholz, J. A. M., Strauss, J., Witt, S., Lill, A., Ehnold, L.-I., Neupert, N., Liffner, B., Lühken, R., Petter, M., Lorenzen, S., Wilson, D. W., Löw, C., Lavazec, C., Bruchhaus, I., Tannich, E., Gilberger, T. W., & Bachmann, A. (2019). Dissecting the Gene Expression, Localization, Membrane Topology, and Function of the *Plasmodium falciparum* STEVOR Protein Family. *MBio*, *10*, e01500-19. https://doi.org/10.1128/mBio.01500-19

Yabsley, M. J., & Shock, B. C. (2013). Natural history of Zoonotic *Babesia*: Role of wildlife reservoirs. In *International Journal for Parasitology: Parasites and Wildlife*, 2, 18–31. https://doi.org/10.1016/j.ijppaw.2012.11.003

Young, K. M., Corrin, T., Wilhelm, B., Uhland, C., Greig, J., Mascarenhas, M., & Waddell, L. A. (2019). Zoonotic *Babesia*: A scoping review of the global evidence. *Public Library of Science One*, *14*, e0226781. https://doi.org/10.1371/journal.pone.0226781

Zintl, A., Mulcahy, G., Skerrett, H. E., Taylor, S. M., & Gray, J. S. (2003). *Babesia divergens*, a Bovine Blood Parasite of Veterinary and Zoonotic Importance. In *Clinical Microbiology Reviews*, *16*, 622–636. https://doi.org/10.1128/CMR.16.4.622-636.2003

H. APPENDIX

Supplement 1: Alignment of reference Bdiv_04560c sequence with isolated and cloned sequence from *B. divergens* culture.

Bdiv_004560c

```
ref             CGTTGAGTGTGCCATGATGACACTTAAACAAGTGACACAACTAATTACTTTATCCAGGGT
cons            ------------------------------------------------------------


ref             ACGGATAAAAGAAGGGTTTTACGAAGCTGCTGGACTTAAACAATCAGCATCATGGCGTTC
cons            ---------AGAAGGGTTTTACGAAGCTGCTGGACTTAAACAATCAGCATCATGGCGTTC
                         **************************************************


ref             CAAAATACACGATGGAGTGTCTTTATCACCATGATGCTCGCCCTAGGCGTGCTCAATCCG
cons            CAAAATACACGATGGAGTGTCTTTATCACCATGATGCTCGCCCTAGGCGTGCTCAATCCG
                ************************************************************


ref             GGTGTACAGGCGCTTGGTGACAGTTGGTCCCTCAAAGATATATTCAAGAAGTCCGTAAAC
cons            GGTGTACAGGCGCTTGGTGACAGTTGGTCCCTCAAAGATATATTCAAGAAGTCCGTAAAC
                ************************************************************


ref             CACACCACCAGCTATGCCTTGTTGGCAAACACCATCTCATACACTAGCGAAGCTGACTTT
cons            CACACCACCAGCTATGCCTTGTTGGCAAACACCATCTCATACACTAGAGAAGCTGACTTT
                ******************************************** ************


ref             GACCCAAATGGTGAGATGATGAGAAGTGTAGAGCTGGCTCCGGGAGAGTCACTTAAATAC
cons            GACCCAAATGGTGAGATGATGAGAAGTGTAGAGCTGGCTCCGGGAGAGTCACTTAAATAC
                ************************************************************


ref             GCATGCGGAATTCCCGGCAATGGTACTAACGGTACTTTTACGGCCATACCACAGGATCCC
cons            GCATGCGGAATTCCCGGCAATGGTACTAACGGTACTTTTACGGCCATACCACAGGATCCC
                ************************************************************


ref             GTCAGATACGTGCTAGACTACAAGGGGGATGCTGATATGGATGTTGCCCTTACCAGGGTG
cons            GTCAGATACGTGCTAGACTACAAGGGGGATGCTGATATGGATGTTGCCCTTACCAGGGTG
                ************************************************************


ref             AGGCCTAGTTACTCTGTATACAGGAGTGACGCAACTATCGTTGTAGGTGCGTTCCAAAAA
cons            AGGCCTAGTTACTCTGTATACAGGAGTGACGCAACTATCGTTGTAGGTGCGTTCCAAAAA
                ************************************************************


ref             ATCGGATTTGACCATCTTTTGATCACTTACGACAAGGATAGTATTATCGCGGCGAATAAC
cons            ATCGGATTTGACCATCTTTTGATCACTTACGACAAGGATAGTATTATCGCGGCGAATAAC
                ************************************************************
```

```
ref             CCGGATAGGTTTAGCCTAAACCTCGTGTGCATGTTTCAATCTAACGGTGAGGCAACTGAG
cons            CCGGATAGGTTTAGCCTAAACCTCGTGTGCATGTTTCAATCTAACGGTGAGGCAACTGAG
                ************************************************************

ref             CCCCGTTATCGTTGGCTCCAGATAAAGTTCAAGAATGTGGTGTCCTTAGCATACGGTTGT
cons            CCCCGTTATCGTTGGCTCCAGATAAAGTTCAAGAATGTGGTGTCCTTAGCATACGGTTGT
                ************************************************************

ref             GGAAGTGTCATGTTTCACATGTTCAAGAACACGATTCCAATGAGCGCCAGTGCCTCTCCT
cons            GGAAGTGTCATGTTTCACATGTTCAAGAACACGATTCCAATGAGCGCCAGTGCCTCTCCT
                ************************************************************

ref             GAGGAGACAAAGTTGTTGAGATTTTGCTCCATTAATGCAGAACCCGGTATGGTTGTGGGG
cons            GAGGAGACAAAGTTGTTGAGATTTTGCTCCATTAATGCAGAACCCGGTATGGTTGTGGGG
                ************************************************************

ref             ATTTACTGCGACAAGGACGAGCACGTATACCCGGATAACTGCTTCGTACAGACCCAGTCC
cons            ATTTACTGCGACAAGGACGAGCACGTATACCCGGATAACTGCTTCGTACAGACCCAGTCC
                ************************************************************

ref             GTAAGAAGTGAAAAGGTTTCGAGATCAGGTCTACTCGAAACTGATTTCCCATACAACAGC
cons            GTAAGAAGTGAAAAGGTTTCGAGATCAGGTCTACTCGAAACTGATTTCCCATACAACAGC
                ************************************************************

ref             CTTGAAGGCAGACTGAGGCTCGCACGAGTCAGGGATAGGTTGGATGGGTACAACAGCGGG
cons            CTTGAAGGCAGACTGAGGCTCGCACGAGTCAGGGATAGGTTGGATGGGTACAACAGCGGG
                ************************************************************

ref             TCCGCATGTTTCTGCAAAGACAAATCCAACGTTATCACAGCTTCAGTTGAGGTCCTATTT
cons            TCCGCATGTTTCTGCAAAGACAAATCCAACGTTATCACAGCTTCAGTTGAGGTCCTATTT
                ************************************************************

ref             AACATGGAACAAGTGTGTGACTACGAGAAGTTTATGCATGTACTACTGCGCCGCCGCGCT
cons            AACATGGAACAAGTGTGTGACTACGAGAAGTTTATGCATGTACTACTGCGCCGCCGCGCT
                ************************************************************

ref             TTACCTCATTATAAATGCCACAAGGAGTTACGCGCTGGAGGAGTGGTTAAACTTGTCATA
cons            TTACCTCATTATAAATGCCACAAGGAGTTACGCGCTGGAGGAGTGGTTAAACTTGTCATA
                ************************************************************

ref             CCGAAGTTAGATCAATCCTCCGAGCCGTCCAAGGTGCTTCGCTCAAGGGGGATATACCCT
cons            CCGAAGTTAGATCAATCCTCCGAGCCGTCCAAGGTGCTTCGCTCAAGGGGGATATACCCT
                ************************************************************
```

```
ref             CGTAAGATAACGGAGTATACATACATCAGCGGCGACGAAGCGACAGGGTTGAAGCCCATC
cons            CGTAAGATAACGGAGTATACATACATCAGCGGCGACGAAGCGACAGGGTTGAAGCCCATC
                ************************************************************

ref             AAGATCGACGAGATTCTCGGTAGAGCGGGGCTTATTATCACCAAGAAGGAGGATTCGAAA
cons            AAGATCGACGAGATTCTCGGTAGAGCGGGGCTTATTATCACCAAGAAGGAGGATTCGAAA
                ************************************************************

ref             AACGAAATTTATGAGTTCAAGGCCTCCACAAACGCAATACTGGTTCTAAGAACAGAGGTG
cons            AACGAAATTTATGAGTTCAAGGCCTCCACAAACGCAATACTGGTTCTAAGAACAGAGGTG
                ************************************************************

ref             GCCAACCTCACGTATCTCTATGAGTTCTATGACACCTTCAAGCACTCTTTAATGCCAGTT
cons            GCCAACCTCACGTATCTCTATGAGTTCTATGACACCTTCAAGCACTCTTTAATGCCAGTT
                ************************************************************

ref             AGAACTATAATCTCCATGGGGATTGTGCCGACAGATCCCTTTACCTATGGATGCGGTGTC
cons            AGAACTATAATCTCCATGGGGATTGTGCCGACAGATCCCTTTACCTATGGATGCGGTGTC
                ************************************************************

ref             TCTTCATCCACCATCTTTAACCATGATGGCGTCCATTTCGAGAACAAAGCTGTGACTACA
cons            TCTTCATCCACCATCTTTAACCATGATGGCGTCCATTTCGAGAACAAAGCTGTGACTACA
                ************************************************************

ref             AACTTAGGAACGCATAACGAGACCCACTGCACGGTCAATGGGTACGCTAATTCACCCGTG
cons            AACTTAGGAACGCATAACGAGACCCACTGCACGGTCAATGGGTACGCTAATTCACCCGTG
                ************************************************************

ref             GGATTCTACTGTCCACCAAATTTCACCCTTTACCCAGAGGACTGTTTCAGTAGTGCTATT
cons            GGATTCTACTGTCCACCAAATTTCACCCTTTACCCAGAGGACTGTTTCAGTAGTGCTATT
                ************************************************************

ref             TTGGTTTCTACAGGTAGAAAGGTTGCTGTGAGTGACTATGTCCCTCTAGCCAGGGTTGTC
cons            TTGGTTTCTACAGGTAGAAAGGTTGCTGTGAGTGACTATGTCCCTCTAGCCAGGGTTGTC
                ************************************************************

ref             AAGTCCAAGAACATAAAGGTCCTTGACTTCTCTATCCCATCACGTCTGAAGACTGGTGTC
cons            AAGTCCAAGAACATAAAGGTCCTTGACTTCTCTATCCCATCACGTCTGAAGACTGGTGTC
                ************************************************************

ref             ACATATTCCAATGAAAAGCTACAGTGCAAGTGCCAGCGCAATGATGGAGCAGTTATGGCG
cons            ACATATTCCAATGAAAAGCTACAGTGCAAGTGCCAGCGCAATGATGGAGCAGTTATGGCG
                ************************************************************
```

```
ref             ACGATCACACTGGACTTGAGCAATCCCTATCCGCGAGCATGAGGCCGATCCAATGGTG
cons            ACGATCACACTGGACTTGAGCAATCCCTATCCGCGAGCATGAGAGGCCGATCCAATGGTG
                ************************************************************

ref             GCAAAATCGTCCTCCGCAATAGGTAGCAATTACCTGGTTCACCTTAAATTGGCGATTAAC
cons            GCAAAATCGTCCTCCGCAATAGGTAGCAATTACCTGGTTCACCTTAAATTGGCGATTAAC
                ************************************************************

ref             ATCAAGCATTATGTGCATTGTAACATTATGTATAAAATTGTGAGTTAGGGTTTCAGTCGT
cons            ATCAAGCATTATGTGCATTGTAACATTATGTATAAAATTGTGAGTTAGGGTTTCAGTCGT
                ************************************************************

ref             CCTAGGGCTGCC
cons            C-----------
                *
```

GCCATGATGACACTTAAACAAG PCR primer in forward direction

CAGCCCTAGGACGACTGAAA PCR primer in reverse direction

GTGTCCTTAGCATACGGTTGTG sequencing primer in forward direction 1

CTGGTTCTAAGAACAGAGGTGG sequencing primer in forward direction 2

ATG start codon

TGA stop codon

C

A mismatch

ref reference sequence

cons consensus sequence