Investigation into the Genetic Variation of Toll-Like Receptor 9 in Cattle Using Both Sanger and Next-Generation Sequencing from FTA-Cards

A thesis submitted to The University of Salford for the Degree of Doctor of Philosophy (PhD) in the School of Science, Engineering and Environment

2020

Afnan A. Al-Qurashi

Supervised by:

Professor Geoff Hide

Table of Contents page

Words counts: 51899

List of Figures

List of Tables

Acknowledgment

I would like to thank Allah for what I have achieved in my Ph.D. project and for always being there for me. Then, I would like to humbly acknowledge all those who helped me to complete this research successfully.

Firstly, I would like to acknowledge my supervisor, Professor Geoff Hide, for his supervision, advice, and guidance during my Ph.D. project. He provided me with encouragement and an ideal environment to perform scientific research. I would like to thank him for his valuable advice in scientific discussion, which has helped me to pursue my future career. I hope to maintain a relationship with him long into the future and continue to learn from him.

Besides, this thesis would not have been possible without the help and guidance of my second supervisor Dr. Ian Goodhead. He helped me a lot especially with the NGS data analysis. A special thanks to Dr.Charles Baillie, who help me with the preparation NGS libraries. I would also like to thank all members of the molecular and genomic labs at the University of Salford for their help particularly, Dr. Marian Denson. Additionally, I am thankful to Dr. Justin Ideozu and Mr. Dominic West for sharing their data.

Finally, I would like to thank my husband Dr.Faisal Minshawi, for his love, care, and support, and for keeping me sane during my writing. Thank you for being my muse, proof-reader. I would like to thank my children Toleen and Hamza, for encouraging me to finish this thesis. They have been my strength when I was weak. Last, but not least, I would like to thank my sisters and brothers, who have been there for support.

This thesis is dedicated to the memory of my mother and my father, I miss them every day...

Afnan A. Al-Qurashi

July 2020

Abbreviations

TLRs	Toll-Like Receptors
PAMPs	Pathogen Associated Molecular Patterns
DCs	Dendritic cells
NK	Natural killer cells
PRRs	Pattern Recognition Receptors
LPS	Lipopolysaccaride
DAMPs	Damage Associated Molecular Patterns
TNF	Tumour Necrosis Factor
RIG	Retinoic acid-inducible gene
RLRs	Retinoic acid-inducible gene (RIG)-I-like receptors
NOD	Nucleotide-binding oligomerization domain
NLRs	Nucleotide-binding oligomerization domain-like receptor
LRRs	Leucine-Rich Repeats
IL-1R	Interleukin-1 receptor
ER	Endoplasmic Reticulum
dsRNA	Double-Stranded RNA
ssRNA	Single-Stranded RNA
HIV	Human Immunodeficiency Virus
pDCs	Plasmacytoid Dendritic Cells
cDCs	Conventional DCs
MyD88	Myeloid Differentiation Factor 88
E. coli	*Escherichia coli*
T. gondii	*Toxoplasma gondii*
rRNA	Bacterial 23S ribosomal RNA
Mal	MyD88 adaptor-like
IRAK	Stimulation IL-1 Receptor-Associated Kinase
TRAF	Tumour necrosis factor receptor (TNFR) -associated factor
AP1	Activator protein
NF- κB	Nuclear Factor kappa-light-chain-enhancer of activated B cells
SOCS1	Suppressor of Cytokine Signalling 1
MAP3K	Mitogen-Activated protein kinase Kinase kinase
TGF-β	Transforming Growth Factor-β

TAK1	Transforming Growth Factor (TGF-β)-activated kinase 1
JNK	c-Jun *N*-terminal kinase
IKK	I kappa B kinase
RNA-Seq	RNA sequencing
ChIP	Chromatin Immunoprecipitation
TIR	Toll-interleukin receptor
INF	Type I interferons
CpG	Cytosine-phosphate-guanine
HAV	Hepatitis A Virus
Map	*Mycobacterium avium subspecies paratuberculosis*
SLE	Systemic lupus erythematosus
FTA	Flinders Technology Associates
NGS	Next Generation Sequencing
dNTPs	deoxynucleotide triphosphates
SNP	Single Nucleotide Polymorphism
WGA	Whole Genome Amplification
MHC	Major Histocompatibility Complex
NF- Kb	Nuclear Factor Kappa-light-chain-enhancer of activated B cells
IBD	Inflammatory bowel disease

Abstract

The susceptibility or resistance to infectious diseases depends on the host immune response. The genetic variation of immune-relevant genes such as Toll-Like Receptor (TLR) genes is associated with susceptibility or resistance to the pathogen. In this thesis, the genetic variation of Toll-like receptor 9 (TLR9) was investigated; this gene is known to have a function in parasite recognition. In a previous study from the Hide lab, 80 bovine blood samples were collected from bulls in Ahoada, Nigeria and the samples were stored on Whatman FTA cards. In this study bovine genomic DNA was extracted from the FTA cards using three different approaches: 5% Chelex resin, the REPLI-g Mini (Qiagen) and the DNeasy Blood & Tissue Kit (Qiagen). The most efficient method for DNA extraction was via the DNeasy Blood & Tissue (Qiagen) kit and this allowed a PCR protocol to be developed for amplification of a fragment of the bovine TLR9 gene. Subsequently, Sanger sequencing and Next Generation Sequencing (NGS) approaches validated the existence of genetic variation across a 554 bp region of exon 2 (f1TLR9) of the bovine TLR9 gene. Moreover, our data revealed that there is no significant correlation between *Trypanosoma* infection and the existence of genetic variations of f1TLR9 in *Bos. indicus* cattle from Nigeria; the interrater reliability between NGS and Sanger sequencing was intermediate. In conclusion, high-quality bovine DNA was extracted from FTA cards to allow PCR amplification and detection of genetic variation within exon 2 of the TLR9 gene by two independent sequencing approaches. These results provide a methodology for facilitating future studies into the existence of genetic variation in other TLR and immune-relevant genes and whether, or not, these correlate with disease susceptibility, or resistance, in both humans and animals.

CHAPTER ONE

Introduction

1.1. Innate immunity

The immune system is composed of a complicated network of cells, tissues and organs that work tightly together to fight pathogens. The vertebrate immune system is split into two branches: the immediate or innate immune system and the highly specific adaptive immune system. The innate and adaptive immune systems work together to initiate an immune response and eliminate pathogens. The innate immune system detects an infection using a limited number of germ-line encoded receptors that recognize molecular structures that are unique to certain classes of infectious microbes. The adaptive immune system uses randomly generated, clonally expressed, highly specific receptors with seemingly limitless specificity (Palm and Medzhitov, 2009).

Innate immunity, also called natural immunity, is considered not only the first line of defence against infectious agents but also a barrier to entry that prevents infection and recruits cellular components to identify and activate antimicrobial responses (Skevaki et al., 2015). Recent studies have shown that the innate immune system demonstrates some specificity, which simplifies the ability to identify differences between the self and foreign antigens and induces adaptive immune cells to respond to invading pathogens (Hoebe et al., 2004). The innate immune system is based on rapid, primary, and early barrier responses to an infectious agent. It exhibits a vast stereotyped response, acting immediately to recognize and eliminate invading pathogens with few microbial stimuli and reacting appropriately to control infection (Iwasaki and Medzhitov, 2010). Innate immunity works physically or chemically via the epidermis, ciliated respiratory epithelium, vascular endothelium and mucosal surfaces, which provide antimicrobial secretions, biological barriers, specialized cells and soluble molecules (Mogensen, 2009). There are different immune cells in the complement system that secrete various soluble mediators and constitute an effective defence response, such as dendritic cells (DCs), macrophages, neutrophils, mast cells and natural killer cells (NKs) (Kaufmann, 2008).

In 1989, Janeway devised the principles of molecular recognition by the innate immune system (Janeway, 1989). It has been proposed that a set of germlines encoded Pattern recognition receptors (PRRs) expressed by macrophages, neutrophils and DCs can detect conserved evolutionary Pathogen-associated molecular patterns (PAMPs) (Medvedev, 2013). These PRRs are responsible for recognizing a pathogen protein, carbohydrate, lipoprotein, DNA and RNA to enabling an immediate response against it. Moreover, it provides signals to the adaptive immunity to secrete proinflammatory cytokines (Kumar et al., 2011). There are two types of PRRs: non-signalling and signalling. Non-signalling PRRs include soluble factors and transmembrane proteins, and they play an essential role in the phagocytosis of microorganisms or recognition by the complement system followed by protease cascades. Signalling PRRs not only recognize the conserved molecular structures shared by a group of microorganisms known as PAMPs but also stimulate signalling pathways to eliminate an invading pathogen (Kaisho and Akira, 2006). PAMPs have certain features that make them efficient. For example, they must be unique to microbes and absent from eukaryotic cells to accurately signal infections of, for instance, unmethylated CpG in bacteria and viruses. Second, they are common to a broad class of microbes, enabling a limited number of germ-line encoded receptors, such as lipopolysaccharides (LPSs), to detect all infections (Palm and Medzhitov, 2009). Recently, evidence has been published showing that PRRs can sense endogenous molecules released from damaged cells that exhibit so-called damage-associated molecular patterns (DAMPs) (Amarante-Mendes et al., 2018). Recognition of PAMPs and DAMPs amplified the transcription of inflammatory genes and stimulated the expression of inflammatory cytokines, like type I interferons (IIFNs), tumour necrosis factor (TNF-α) and IL-6, and other chemokines and antimicrobial proteins (Blasius and Beutler, 2010). Signalling PRRs are classified into five families: TLRs, nucleotide-binding and oligomerization domain (NOD)-like receptors (NLRs), retinoic acid-inducible gene I (RIG-I)-like receptors (RLRs), C-type lectins (CTLs), and absent-in-melanoma (AIM)-like receptors (ALRs). (Kumar et al., 2011, Novák, 2014, Kim et al., 2016). Not only immune cells, such as macrophages, and DCs but also various non-

immune cells such as epithelial cells, endothelial cells, and fibroblasts express PRRs that recognize the foreign antigen. TLRs are transmembrane-signalling PRRs, which are responsible for realizing the specific component of the invading microbial including fungi, protozoa, and viruses not only outside of the cell but also in intracellular endosomes and lysosomes of a cell (Kharaji and Haghparast, 2010).

1.2. Toll-like receptors

1.2.1. Structure and function

Toll-like receptors (TLRs) are type one integral membrane glycoproteins with molecular weights ranging from 90–115kD (Bell et al., 2003). The first TLR was discovered in *Drosophila*, a fruit fly, at the end of the 20th century by Jules Hoffmann (Anderson, 2000). The Toll protein was demonstrated to be essential for fighting against fungal infection (Lemaitre et al., 1996). A Toll receptor was also shown in a mammalian homologue of the *Drosophila* Toll receptor, which featured similar cytoplasmic portions and induced the expression of genes implicated in inflammatory responses (Medzhitov et al., 1997). Since then, the Toll receptor has been identified as hyporesponsive to LPSs when the TLR4 gene is mutated in mice (Poltorak et al., 1998). TLRs are expressed in a variety of cells though they are most highly expressed by immune cells, including DCs, macrophages, mast cells and specific types of T cells, and less defined in non-immune cells, including fibroblasts and epithelial cells (Kaisho and Akira, 2006). The basic structure of TLRs consists of extracellular and cytoplasmic domains. The N-terminal part is either extracellular or in the endosomal membrane and contains tandem arrays of a short leucine-rich repeats (LRRs), which form a horseshoe structure (Fig. 1.1). Each LRR unit consists of a β-strand and an α-helix connected by loops that are responsible for TLR–ligand interactions. The intracellular C-terminal domain and cytoplasmic signalling domain are homologous and have high similarity to the vertebrate interleukin-1 receptor (IL-1R), also known as the Toll-interleukin

receptor (TIR; (Kumar et al., 2011). The TIR, which contains three immunoglobulin-like domains, is essential for signal transduction (Gay and Gangloff, 2007).

Activated TLRs also are vital molecules for microbial elimination. This process involves recruitment of phagocytes to the site of infection, killing of microbes, activation of APCs, and full activation of T cells (Iwasaki and Medzhitov, 2015). To date, 13 different TLRs have been characterized in mammals, 11 in humans and 13 in mice (Nie et al., 2018). TLR1 and TLR9 are conserved in mice and humans whereas mice lack TLR10 but have TLR11 and TLR 13, which humans lack. TLRs can be divided into several groups, each of which recognizes related PAMP derived from various microorganisms, including bacteria, viruses, protozoa, and fungi (Harsini et al., 2014).

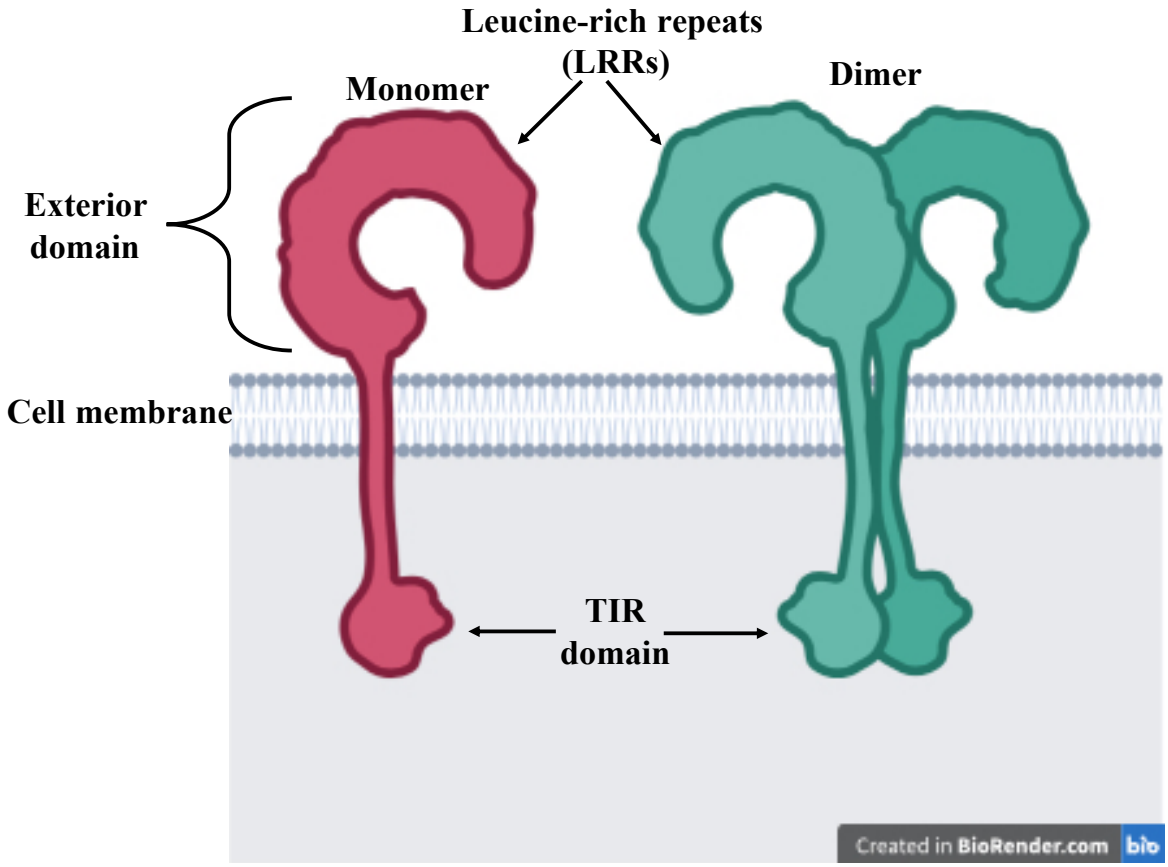


Figure 1. 1 Cartoon illustration of TLR structure. The dimer is the biologically active TLR protein. The exterior domain, containing the LRRs, senses PAMPs such as LPS, lipoproteins. The Toll-interleukin receptor (TIR) is located internally for signalling activation. The image was generated using BioRender.com.

1.2.2. TLR ligands

Different approaches are used to classify TLRs based on either their primary structure and function or their localization and respective PAMP ligands. The TLRs are subdivided into five subfamilies: TLR-2, TLR-3, TLR-4, TLR-5 and TLR-9. The TLR-2 subfamily consists of four members: TLR1, TLR2, TLR6 and TLR10. The TLR-9 subfamily includes TLR7, TLR8 and TLR9, which are highly homologous and trigger a signalling cascade, leading to cytokine production and activation of the adaptive immune response (Dasu et al., 2010). The TLR-3, TLR-4 and TLR-5 subfamilies comprise a single member, which works alone or in combination with other receptors or molecules (Skevaki et al., 2015). The second approach relies on localization and corresponding ligands, which some TLRs (such as TLR1, TLR2, TLR4, TLR5, TLR6, TLR10 and TLR11) express on the cell surface and others (TLR3, TLR7, TLR8 and TLR9) express in intracellular vesicles, like the endoplasmic reticulum (ER), endosomes, lysosomes and endolysosomes (Hennessy et al., 2010, Kawai and Akira, 2010). In cooperation with a TLR-associated adaptor molecule, TLRs can recognize fungal pathogens such as *Candida albicans*, *Aspergillus fumigatus*, *Cryptococcus neoformans* and *Pneumocystis carinii* (Roeder et al., 2004). Each TLR has specific ligands, which are mainly derived from bacteria, fungi, protozoa, and viruses. In general, TLR1, TLR2, TLR4, TLR5, TLR6, TLR10 and TLR11 mainly recognize microbial membrane components, such as lipids, lipoproteins, and proteins, whereas TLR3, TLR7, TLR8 and TLR9 recognize microbial nucleic acids (Kawai and Akira, 2010) (Table 1.1). TLR1, TLR2 and TLR6 recognize PAMP lipids. TLR1 and TLR6 are functionally associated with TLR2 and recognize triacylated lipopeptides (Alexopoulou et al., 2002, Vu et al., 2016).

TLR2 is the most studied TLR, and it recognizes different microbial components. Specifically, it recognizes lipoproteins from various gram-negative bacteria and mycoplasma (Aliprantis et al., 1999, Hirschfeld et al., 1999). Moreover, it recognizes peptidoglycan and lipoteichoic acid from gram-positive bacteria (Lehner et al., 2001, Schwandner et al., 1999). Furthermore, it recognizes

lipoarabinomannan from mycobacteria (Means et al., 1999, Underhill et al., 1999), glycosylphosphatidylinositol anchors from *Trypanosoma cruzi* (Campos et al., 2001) and phenol-soluble modulin from *Staphylococcus epidermis* (Hajjar et al., 2001). TLR2 has been identified as playing an essential role in the detection of zymosan from fungi (Underhill et al., 1999) and glycolipids from *Treponema maltophilum* (Opitz et al., 2001).

TLR3 has unique structural features in comparison to other TLRs. It lacks a highly conserved proline residue in the BB loop of the cytosolic TIR signalling domains (Verstak et al., 2013), and it is preferentially expressed in mature dendritic cells (Muzio et al., 2000). TLR3 is essential for recognition of virus-derived double-stranded RNA (dsRNA), which is produced by most viruses during the replicative cycle. The dsRNA induces synthesis of type I interferons (IFN-α and IFN-β), which result in various physiological effects, including anti-viral and immunostimulatory activities (Takeda et al., 2003).

TLR4 is a receptor of LPSs that are an essential component of the outer membrane of gram-negative bacteria. On the other hand, TLR4 recognizes peptidoglycans and lipoproteins, which are expressed on the gram-positive bacterial cell wall (Kaisho and Akira, 2006). TLR4 forms a complex with MD2 on the cell surface, and together they serve as the main LPS-binding component (Vu et al., 2016). Additionally, they can bind to CD14 or a glycosylphosphatidylinositol-anchor in monocytes and myeloid cells and detect bacterial LPS (da Silva Correia et al., 2001). TLR4 is also involved in the recognition of endogenous ligands such as heat shock proteins (HSP60 and HSP70). These proteins activate macrophages and DCs, leading to secretion of pro-inflammatory cytokines (Galucci and Matzinger, 2001). Indeed, a previous study confirmed that TLR4 mutant mice are more susceptible to pulmonary tuberculosis, and their splenocytes produced a low level of IFN-γ (Branger et al., 2004).

TLR5 recognizes the flagellin protein from both Gram-positive and Gram-negative bacteria (Hayashi et al., 2001). Flagellin protein, a highly complex structure that extends out from the outer membrane, is the primary protein component of the flagellum (Eaves-Pyles et al., 2001). Once

TLR5 is activated, it activates the nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) and the mobilization, stimulation and secretion of cytokines (Hayashi et al., 2001).

TLR7 is usually present in endosomes and it recognizes uridine-rich single-stranded RNA (ssRNA) viruses, such as the human immunodeficiency virus (HIV), and influenza A virus (Heil et al., 2004). TLR7 also recognizes synthetic poly-U RNA and specific small interfering RNAs (Hornung et al., 2005). Expression of TLR7 is highly dependent on plasmacytoid dendritic cells (pDCs), which can produce large amounts of type I interferons and cytokines after virus infection (Akira et al., 2006, Kawai and Akira, 2006). Moreover, TLR7 is expressed on conventional DCs (cDCs), senses RNA species from bacteria (e.g. group B *Streptococcus*), and induces type I interferon expression (Mancuso et al., 2009). TLR8 is similar to TLR7 in that it also recognizes viral ssRNA (Desnues et al., 2014).

TLR10 expression has been detected only in humans (Lee et al., 2014). It is expressed in B cells and pDCs (Bourke et al., 2003; Hasan et al., 2005). Previous work showed that TLR10 works with TLR2 to recognize triacylated lipopeptides and recruit myeloid differentiation factor 88 (MyD88) (Guan et al., 2010). Recently, a study showed that TLR10 could recognize viral infection by sensing the viral structural proteins such as HIV-1 gp41 (Lee et al., 2014; Henrick et al., 2019).

Expression of TLR11 has been shown only in kidney and bladder epithelial cells of mice; it does not exist as a functional protein in humans (Zhang et al., 2004). It is highly expressed in uropathogenic *Escherichia coli* (*E. coli*) infection (Anders and Patole, 2005) and can activate NF-κB (Zhang et al., 2004). Another known role of TLR11 is recognition of profilin-like protein, which is expressed in *Toxoplasma gondii* (*T. gondii*; (Lauw et al., 2005). Recently, it was shown that flagellin from *Salmonella spp.* in the intestines of mice is recognized by TLR11 (Mathur et al., 2012).

TLR12 is like TLR11, recognizing profilin protein from the protozoan parasite *T. gondii* and regulating IL-12 production by DCs in response to the parasite (Raetz et al., 2013).

TLR13 is an endosomal murine TLR that has been described to interact with the bacterial 23S ribosomal RNA (rRNA) sequence or UNC93B protein (Hidmark et al., 2012).

Table 1. 1 Summary of Toll-like receptor ligands and origins (Dos-Santos et al., 2016, Harsini et al., 2014).

TLRs	Ligand	Origin of ligand
TLR1	Triacyl lipopeptides Soluble factors Surface proteins	Bacteria and mycobacteria *Neisseria meningitides* *Borrelia burgdorferi*
TLR2	Lipoprotein/lipopeptide peptidoglycan Lipoteichoic acid Lipoarabinomannan Phenol-soluble modulin glycoinositolphospholipids Glycolipids Porins Atypical lipopolysaccharide Structural viral proteins Zymosan Heat shock protein 70	Various pathogens Gram-positive bacteria Gram-positive bacteria *Staphylococcus epidermidis* *Trypanosoma cruzi* *Treponema maltophilum* *Neisseria* *Leptospira interrogans* Herpes simplex virus Saccharomyces *Listeria monocytogenes*
TLR3	Double-stranded RNA	Viruses
TLR4	Lipopolysaccharide acid Envelope protein Heat shock protein 60	Gram-negative bacteria Respiratory syncytial virus *Chlamydia pneumonia*
TLR5	Flagellin	Flagellated bacteria
TLR6	Diacyl lipopeptides Lipoteichoic acid Zymosan Heat-labile soluble factor	Mycoplasma Gram-positive bacteria *Saccharomyces* Group B streptococci
TLR7	Single-stranded RNA synthetic poly-U RNA	Viruses
TLR8	Single-stranded RNA	Viruses
TLR9	Unmethylated CpG DNA Hemozoin	Bacteria and viruses Plasmodium
TLR10	Acylated lipopeptides	Bacteria
TLR11	Uropathogenic Profilin-like protein	*Escherichia coli* *Toxoplasma gondii*
TLR12	Profilin-like protein	*Toxoplasma gondii*
TLR13	23S ribosomal UNC93B	Bacterial

1.2.3. TLR signalling

The TLR signalling cascade activates when a TLR recognizes a PAMP site of an invading pathogen. A series of intercellular signals are activated to eliminate the pathogen and activate the adaptive immune response. During infection, microbial TLR ligands can cause DCs to mature. This response is characterized by upregulation of surface MHC–peptide complexes and co-stimulatory molecules involved in the activation of T cells and the chemokine receptor CCR7 as well as by the migration of DCs to the T cell zones in a draining lymph node (Iwasaki and Medzhitov, 2004). Once in the lymphoid tissue, mature DCs promote the activation of antigen-specific T cells and secrete cytokines and other factors that can promote effector T cell differentiation. However, microbial TLR ligands can also activate in different tissue-resident cell types to secrete inflammatory mediators, such as IFN-α, IFN-β and TNF-α. These cytokines could promote the maturation of DC and enable them to induce T cell immune responses, providing an indirect mode of DC activation (Kapsenberg, 2003). TLRs have TIR domain-containing adaptor molecules, including MyD88, TRIF, TRAM and MAL, which are responsible for activating diverse TLR signalling pathways and thus the expression of inflammatory cytokines (Fig. 1.2) (Botos et al., 2011). The first adaptor molecule to be discovered was MyD88, which is activated with all TLRs except TLR3 (Ghosh and Stumhofer, 2013). MyD88 is a cytosolic adaptor protein that plays essential roles in both innate and adaptive immune responses by mediating the signal transduction pathways that are initiated by TLRs and the IL-1 and interleukin-18 (IL-18) receptors (IL-1R and IL-18R; (Ohnishi et al., 2009). MyD88 contains an amino N-terminal death domain (DD) with approximately 90 amino acid residues, a C-terminal TIR domain with about 150 amino acid residues and a short connecting linker (Bonnert et al., 1997, Venugopal et al., 2009). The TIR domain in MyD88 interacts with TLRs in combination with another TIR-containing adaptor protein, such as Mal, and stimulates IL-1 receptor-associated kinase (Ohnishi et al., 2009). The IRAK family (IRAK1, IRAK2, IRAK4) plays a role in the response of microbial components that

stimulate TLR2, TLR3, TLR4 and TLR9 (Suzuki et al., 2002). TRAF6 is a member of the tumour necrosis factor receptor (TNFR) associated factor (TRAF) family, which mediates cytokine signalling pathways (Bradley and Pober, 2001). TRAF6 works with IRAK-1 to activate two distinct signalling pathways. The first pathway utilizes mitogen-activated proteins (MAP) kinases to activate activator protein (AP1) transcription factors. The second pathway is the TAK1/TAB activation complex, which increases the activity of the IκB (IKK) kinase complex. This complex encourages the phosphorylation and degradation of IκB, provoking nuclear translocation of the transcription factor NF-κB (Takeda and Akira, 2005).

Another adaptor protein TRIF, the TIR-domain-containing adaptor protein, is found in MyD88-deficient TLRs, such as TLR3. TRIFs have N-terminal and C-terminal domains, which recruit signalling molecules. The N-terminal region and C domain of TRAF activate NF-κB promoters, and IFN-β is activated by just the N-terminal region (Kawai and Akira, 2007).

The IRF family of transcription factors (IRF1–9) plays a significant role in TLR signalling. IRF3 is expressed by dsRNA and activated by the TRIF-dependent signalling pathway. It can produce TNF that able to activate the NF-κB pathway. IRF7 is expressed in pDC and produces a considerable amount of IFN-γ in response to viral infection (Kawai and Akira, 2007).

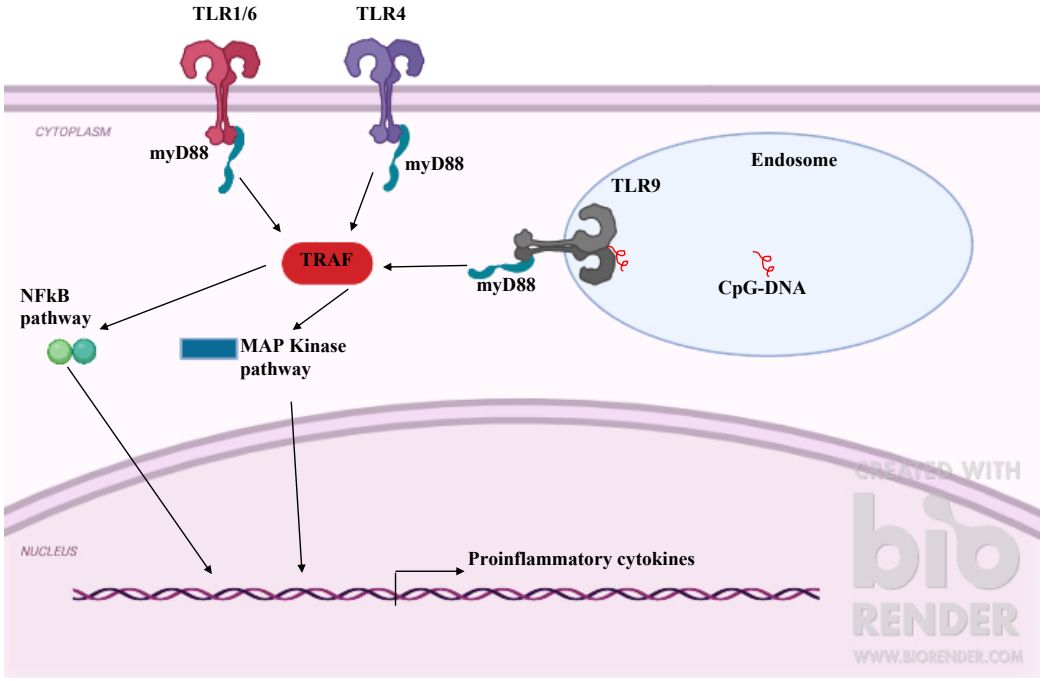


Figure 1. 2 TLR signalling pathways. MyD88-dependent signalling pathways are stimulated by all TLRs except TLR3, either directly (e.g. TLR5, TLR7, TLR8, TLR9, TLR10) or in combination with an additional adaptor molecule, TIRAP (e.g. TLR2 and TLR4. TLR3 and TLR4). This initiates a TRIF-dependent signalling pathway. The IRAK family and TRAF6 are recruited through the MyD88 pathway, which ultimately activates the TAK1/TAB1/TAB2 complex. The MAPK and IKK complex involving activated TAK1 leads to activation and nuclear translocation of the transcriptional factors AP-1 and NF-B, respectively. NF-B and AP-1 induce the gene transcription of inflammatory cytokines. The image was created using BioRender.com.

1.2.4. Role of TLR signalling in adaptive immunity

When innate immunity is ineffective for clearing an infectious agent, adaptive or acquired immunity is stimulated, producing a specific response to each infectious agent (Janeway Jr and Medzhitov, 2002, Medzhitov and Janeway, 1997). Generally, adaptive immunity recognizes a non-self (e.g. pathogen) and generates a specific immune response (e.g. antibodies). It develops memory cell types that rapidly respond to subsequent infections (e.g. plasma cells) (Bonilla and Oettgen, 2010). Adaptive immune cells are T cells that are activated by the action of APCs. B cells are another critical adaptive immune cell that produce antigen-specific antibodies. They prime naive T cells by promoting the interaction of a T cell receptor (TCR) with an MHC-II molecule coupled with secondary co-stimulatory signals. Engagement of TLRs on APC leads to up-regulation of the MHC and co-stimulatory molecules on APCs (Banchereau and Steinman, 1998).

Moreover, as described for TLR signalling, TLR activates the NF-κB and MyD88 pathways, which are critical for the initiation of the adaptive immune response. In a previous study on MyD88-deficient mice, the T helper-1 response was not activated (Schnare et al., 2001).

TLRs are also expressed by adaptive immune cells, such as B cells and T cells. Previous research demonstrated that TLR1–9 is expressed in mouse B cells and induces cytokine secretion with a pattern that is distinct from other APCs, such as DCs (Barr et al., 2007). Furthermore, TLRs are expressed by T cells which suggests that PAMPs have a direct impact on the survival of activated T cells and implies that TLRs on T cells could modulate the adaptive immune response (Gelman et al., 2004).

1.2.5. Negative regulation

Although TLRs protect against a wide variety of pathogens, inappropriate or unregulated activation of TLR signalling can lead to chronic inflammatory and autoimmune disorders. Negative regulation of TLR signalling is essential to block or limit the induction of inflammatory cytokines. The excess production of cytokines leads to systemic disorders that are associated with a high mortality rate in the host; for example, endotoxic shock can be induced by the TLR4 ligand LPS (Takeda and Akira, 2005). However, the association of many intracellular protein molecules with TLR signalling molecules plays a critical role in negative regulation, preventing TLR signalling molecules from activating inflammatory cytokines. IRAK-M, expressed in monocytes and macrophages, prevents the dissociation of IRAK4 and IRAK1 from MyD88. Therefore, it prevents the formation of the IRAK1–TRAF6 complex (Takeda and Akira, 2005). The suppressor of cytokine signalling 1 (SOCS1) is a member of the SOCS family of proteins. Eight members of this family have been reported in mammals. SOCS1 consists of two conserved motifs, a central SH2 domain and a C-terminal SOCS box (Fujimoto and Naka, 2003). SOCS1 is expressed by various cytokines, including IL-4 and IFN-γ, and some TLR ligands, such as LPS and CpG-DNA (Naka et al., 2005). Expression of SOCS1 leads to a reduction of the response of cells to TLR

ligands and directly induces negative regulation of TLR signalling. According to Kinjyo et al. (2002) and Nakagawa et al. (2002), SOCS1-deficient macrophages induce inflammatory cytokines, including TNF-α, IL-6 and IL-12. MyD88s is an alternatively spliced variant of MyD88 that lacks the intermediary domain and leads to the shutdown of IL-1/lipopolysaccharide-induced NF-κB activation and, therefore, inhibition of the MyD88-dependent pathway (Burns et al., 2003). SOCS1 molecules might suppressed in the MyD88-dependent pathway (Nakagawa et al., 2002). On the other hand, SARM (Belinda et al., 2008) and TAG (Palsson-McDermott et al., 2009) are suppressed by the TRIF-dependent pathway. TRIF interacts with TRAF3 and TRAF6 to induce the suppression of inflammatory cytokines. TRAF3 activation is negatively regulated by SOCS3, DUBA (Kayagaki et al., 2007), and TRAF6 is targeted by several inhibitory molecules such as A20 or TANK (Kondo et al., 2012, Skaug et al., 2011). Triad3A is an E3 ligase for which overexpression promotes the degradation of TLR4 and TLR9 via a proteasome-dependent pathway. As a result, Triad3A can inhibit the expression of TLR4 and TLR9 (Chuang and Ulevitch, 2004).

1.3. Toll-like receptor 9

1.3.1. Origin of Toll-like receptor 9

Bovine TLR9 has been mapped to the distal part of *Bos taurus* chromosome 22 (BTA22; (McGuire et al., 2006). The coding gene of TLR9 has two exons, and the central coding region is in the second exon, which features two CpG islands. The first CpG island is located +3095 to +3353 bp from the first base of the start codon, and the second CpG island is located +3556 to +3976 bp from the first base of the start codon (Cargill and Womack, 2007). TLR9 is localized within the ER of cells and is recruited to lysosomes only after cells are stimulated with CpG DNA, such as that from bacteria (Latz et al., 2004). Bovine TLR9 has substantial structural homology with other TLR members, and it has been grouped into a subfamily with TLR7 and TLR8 based on its amino

acid sequence. TLR7 and TLR9 have a long insertion loop (Z-loop) between LRR14 and LRR15, which is necessary for biological functional (Onji et al., 2013). The Z-loop in TLR9 is ring-shaped or has a monomer structure when it is unliganded or bound to inhibitory DNA, respectively. Otherwise, TLR9 and the CpG-DNA complex adopt an M-shaped or dimeric form when binding to ligands (Ohto et al., 2015). TLR9 is expressed intracellularly from immune cells like pDCs, macrophages and B cells, and it is not expressed from mDCs. Moreover, TLR9 is expressed from non-immune cells, including cardiomyocytes and neurones (Sarafidou et al., 2013; Shintani et al., 2013). Mammalian TLR9 sequences revealed that bovine TLR9 shared 79% homology with human TLR9 and 73% homology with murine TLR9 (Griebel et al., 2005).

1.3.2. TLR9 ligands

TLR9 has been identified as a receptor that senses CpG-DNA (Hemmi et al., 2000). TLR9 detects unmethylated cytosine-phosphate-guanine (CpG) regions that mainly present in bacterial and viral genomes (Fig. 1.3) (Vu et al., 2016). Prior investigation revealed that TLR9-deficient mice did not show any response to CpG DNA and exhibited resistance to the lethal effect of CpG DNA without elevation of serum pro-inflammatory cytokine levels (Hemmi et al., 2000). In mammals, the frequency of CpG dinucleotides is low, and the cysteine residues of CpG motifs are highly methylated.

There are two types of CpG DNA. The first is B/K-type CpG DNA, which is conventional and a potent inducer of inflammatory cytokines such as IL-12 and TNF-α. The second is A/D-type CpG DNA, which is structurally different from traditional CpG DNA and has a more remarkable ability to induce IFN-α production from pDCs but less ability to produce IL-12 (Krug et al., 2001, Verthelyi et al., 2001). Subsequently, according to Ohto et al. (2015), TLR9 develops a higher affinity for single-strand DNA than for double-strand DNA. Another study showed that TLR9 has an affinity to RNA-DNA hybrids, which are found more often in particular viruses, like cytomegalovirus and hepatitis B virus, than in ssDNA or dsDNA (Rigby et al., 2014). Moreover,

Trypanosoma brucei also promote activation of TLR9 in macrophages by the MyD88 pathway and boost the production of proinflammatory cytokines (Dos-Santos et al., 2016).

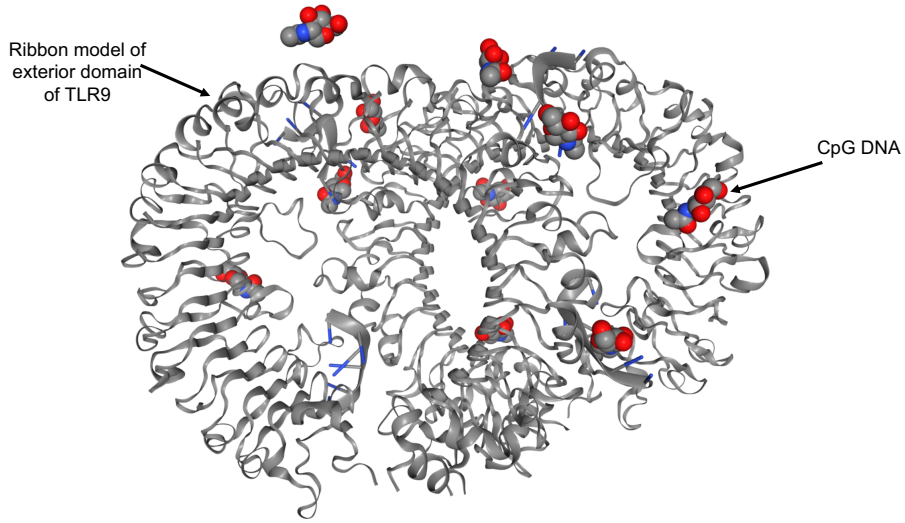


Figure 1. 3 Ribbon model of TLR9 structure. The spherical shape represents the CpG-DNA. This figure adapted from PDB ID:5Y3M (Ohto et al., 2018).

1.3.3. TLR9 signalling

The interaction between TLR9 and CpG motif-containing DNA leads to activation of TLR9 intracellular signalling pathways from the TIR cytoplasmic domain. As a result of this interaction, the conformation of TIR domains changes and the MyD88 domain is activated (Melisi et al., 2014). TLR9 binds with a TIR domain in the C-terminal portion of MyD88 that is linked to IRAK. IRAK-1 and IRAK-4 recruit and activate the RING-domain E3 ubiquitin ligase TRAF6. The serine/threonine kinase in the mitogen-activated protein kinase (MAP3K) family transforms into growth factor-β (TGF-β) activated kinase 1 (TAK1; (Melisi et al., 2011). TAK1 activate NF-κB, c-Jun N-terminal kinase (JNK) and p38 via phosphorylation of I kappa B kinase (IKK). NF-κB and other transcription factors downstream of p38 and JNK are activated, resulting in the transcription of genes that play an essential role in inflammatory and immune responses. The adaptor protein AP3 complex is necessary for TLR9-mediated IFN activation (Sasai et al., 2010).

1.4. TLR9 polymorphisms with diseases

1.4.1. Background

There are different processes for generating continuous genetic variation, such as mutation, single nucleotide polymorphism (SNP), insertion, deletion and inversion. Additionally, the persistence of genetic variation in the genome is determined by different historical and genomic factors (Wright, 2005). Polymorphism is defined as a variation or point mutation in the DNA sequence that occurs in a population with a frequency of at least 1% (Brookes, 1999). SNP is more common than other types of polymorphisms. It is a single base change occurring in a population, wherein the least frequent allele has an abundance of 1%. It is the most straightforward type of genetic difference among individuals. It may involve a nucleotide changing from one purine base to another (e.g., adenine to guanine) called a transition; on the other hand, it may change from a purine to a pyrimidine base (e.g., guanine to cytosine), which is called a transversion (Saint Pierre and Génin, 2014). SNPs may be responsible for diversity among individuals, genome evolution and the most common familial traits, such as curly hair or skin colour (Krawczak et al., 1992, Lohrer and Tangen, 2000, Pitarque et al., 2001).

Additionally, some SNPs may cause inter-individual differences in drug response (Shastry, 2003). They can also lead to common diseases, such as diabetes, obesity, hypertension and psychiatric disorders (Karki et al., 2015). In the human genome, there are an estimated 84.7 million SNP variants with a frequency higher than 1% (Genomes Project, 2015). Further, there is around one SNP per 434 bp in *Bos taurus*, per 194 bp in *B. indicus* and per 225 bp in chicken (Gurgul et al., 2018).

SNPs exist throughout the entire genome, both within and outside coding sequences (Shastry, 2002). They can be categorized into two groups: synonymous (or ‘silent’) and non-synonymous. Synonymous SNPs occur if the SNP is not located in a region of predicted or known functional interest, which can only be inferred from a gene’s location. Some synonymous SNPs are located

within coding regions but do not lead to a change in the amino acids incorporated at the site of their occurrence (Sauvage et al., 2007). Non-synonymous SNPs or a coding-region SNP may alter the sequence of the protein encoded by the modified gene or the transcribed codon and, as a result, influence promoter activity (gene expression) and messenger RNA (mRNA) conformation (stability). This, in turn, may produce diseases.

1.1.1. Genetic variation of TLR9 in human diseases

Prior studies have identified numerous variations in TLR genes such as SNPs and analysed the effects of these variations to understand their impact on gene function and human and animal health (Gavan et al., 2015, Skevaki et al., 2015, Tschirren et al., 2013). As a result, they found that some TLR polymorphism is related to susceptibility or resistance to infectious diseases. The genetic variation in TLR9 has been linked to susceptibility and resistance to bacterial, viral and parasite-based infectious diseases (Carvalho et al., 2008, Costa et al., 2017, McElrath et al., 2009, Oliveira et al., 2013, Sanders et al., 2011). Another study found that TLR9 polymorphism is a significant risk factor for hepatitis A virus (HAV) in the population of India (Kashyap et al., 2016). Moreover, recent studies showed different effects of TLR9 polymorphism in patients infected with malaria infection (Costa et al., 2017, Dhangadamajhi et al., 2017). On the other hand, a previous investigation showed no association between the analysed SNPs in TLR9 and children's susceptibility to meningococcal, pneumococcal bacterial meningitis and tuberculosis (Gowin et al., 2017, Graustein et al., 2015).

1.1.2. Genetic variation of TLR9 in bovine disease

Based on studies of other TLRs it has been discovered that bovine TLR1 and TLR2 SNPs occur that are linked with susceptibility and resistance to *Trypanosoma vivax* infection (Adenaike et al., 2018). Limited studies showed that bovine TLR9 has linked to the associations with pathogens. SNPs in TLR9 in *B. taurus indicus* were significantly associated with susceptibility to bovine

tuberculosis (bTB) (Bhaladhare et al., 2018). In contrast, another study suggests that variants in the TLR9 gene in Chinese Holstein cattle (*Bos taurus*) are not associated with susceptibility to bTB (Sun et al., 2012). However, despite all these efforts more research still needs to be done to build up much-needed evidence-based literature on bovine TLR9 and how genetic alterations may influence resistance or susceptibility to infectious diseases, especially those specific for trypanosomiasis.

1.2. Trypanosoma

1.2.1. Background

Trypanosomes belongs to the protozoan branch of the Trypanosomatida family in the order Kinetoplastida of the Protozoa kingdom. Trypanosomatids are distinguished by having only a single flagellum and are characterized by a modified mitochondrial genome known as the kinetoplast (Lukeš et al., 2005, Vickerman, 1965). Trypanosomes are ubiquitous protozoan parasites that cause infections in a wide range of human and livestock (Votypka et al., 2013). Most species of *Trypanosoma* are heteroxenous as they require more than one host to complete their life cycle. Moreover, they are haemoflagellates since they live in the blood of their definitive host and in addition to mammals, they can infect birds (Zídková et al., 2012) and fish (Overath et al., 1999). Trypanosomes are principally transmitted from one vertebrate to another by blood-sucking arthropods or leeches. They live and replicate in the bloodstream in the definitive host, and they can invade bodily fluids, such as lymph and cerebrospinal fluid (CSF) (Franco et al., 2014). The genera *Trypanosoma* are divided into two branches based on their development in a vector: Salivaria, which develops in the anterior part of the insect digestive tract, and Stercoraria, which develops in the posterior part of the vector digestive tract. Salivarian pathogenic trypanosomes have four subgenera: *Trypanozoon*, *Duttonella*, *Pycnomonas* and *Nannomonas*, which are responsible for most human and animal infections and include species such as *Trypanosoma brucei*

and *Trypanosoma evansi* (Radwanska et al., 2018). Stercoraria pathogenic trypanosomes are transmitted by contamination with the faeces of the insect, such as *Trypanosoma cruzi* (Desquesnes et al., 2013).

The main pathogenic trypanosomes in mammals include genera whose species can cause severe diseases in humans like *Trypanosoma brucei, Trypanosoma cruzi,* which are responsible for the diseases Human African trypanosomiasis (HAT) and Chagas disease respectively. Many species of trypanosome cause African animal trypanosomiasis or nagana (AAT). The three most common species that causes AAT are *Trypanosoma congolense*, *Trypanosoma vivax* and *Trypanosoma brucei brucei* (Bargul et al., 2016, Yaro et al., 2016).

Some species of trypanosomes are non-pathogenic to domesticated animals, such as *T. theileri* of bovines, *T. melophagium* of sheep and *T. caninum* of dogs and these must be differentiated for diagnosis purposes (Madeira et al., 2014, Mansfield, 1977). Only the main species of pathogenic trypanosomes affecting human and animals will be covered in this study as summarised in Table 1.2.

Table 1. 2 General characteristics of trypanosomes affecting human and animals.

Subgenus	Parasite	Hosts	Vector
Trypanozoon	*Trypanosoma brucei gambiense*	Human	*Glossina* sp.
Trypanozoon	*Trypanosoma brucei rhodesiense*	Human	*Glossina* sp.
Schizotrypanum	*Trypanosoma cruzi*	Human, Dogs, cats	Triatomine
Trypanozoon	*Trypanosoma brucei brucei*	Bovine, sheep, goat	*Glossina* sp.
Nannomonas	*Trypanosoma congolense*	Bovine, sheep, goat	*Glossina* sp.
Duttonella	*Trypanosoma. vivax*	Bovine, sheep, goat	*Glossina* sp. Tabanidae *Stomoxys*
Trypanozoon	*Trypanosoma evansi*	Bovine, equine	Tabanidae

1.2.2. Morphology

Trypanosomes vary in size from 8 to over 50 µm. Most *Trypanosoma* species the same basic morphology: a microtubule cytoskeleton, flagellar pocket, flagellum, nucleus, the kinetoplast, and mitochondrion (Fig. 1.4) (Wheeler et al., 2012). This makes it complex to distinguish causative agents, especially when performing a microscopic diagnosis (Hide and Tait, 2009).

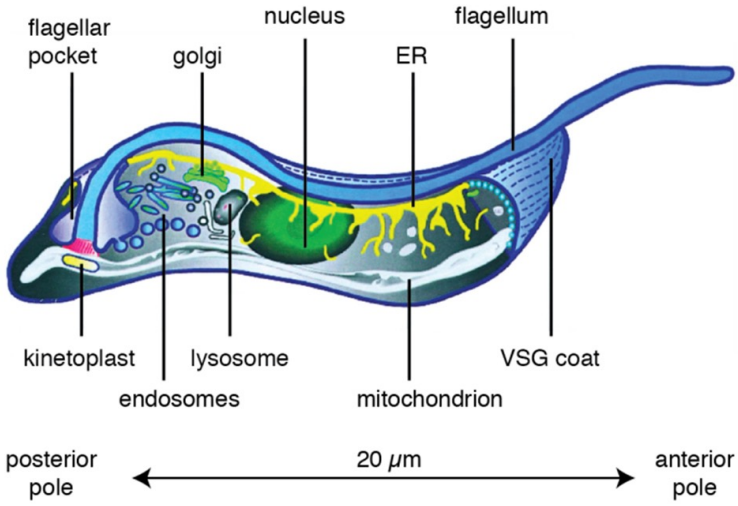


Figure 1. 4 The cell architecture of *Trypanosoma brucei*. The model presents a simplified version of the endocytic compartment of *Trypanosoma brucei.* The anterior and posterior pole are indicated as well as the length of the cell. All major organelles (microtubule cytoskeleton, flagellar pocket, flagellum, nucleus, the kinetoplast, and mitochondrion) and their location are shown (Grunfelder et al., 2003).

The trypanosome cell is elongated and has a highly polarized microtubule cytoskeleton. The microtubule cytoskeleton is responsible for defining *Trypanosoma* species' cell shape, and it remains intact throughout the cell cycle (Vaughan and Gull, 2008). The most basic posterior structure in a trypanosome cell is the flagellar pocket. This pocket is the exit point for the flagellum, on which the motility of a trypanosome cell is dependent (Vaughan and Gull, 2003). The nucleus of trypanosomes, which may be regarded as the command centre of the cell, also plays a significant part in reproduction. It contains DNA, which is arranged in the form of genes and

chromosomes and includes a repertoire of variant surface glycoprotein (VSG) genes and other proteins and enzymes required by the parasite (Wickstead et al., 2004).

The mitochondrion is a single elongated structure that runs from the posterior to the anterior of the trypanosome cell and it contains the kinetoplast DNA. The glycosomes, the cytosol, and the mitochondrion cooperate in the energy metabolism of trypanosomatids (Parsons, 2004). The kinetoplast is seen near to the posterior end of the trypanosome and differs in size and position according to the species. It is the most structurally complex DNA, being composed of two types of DNA molecule; maxicircles and minicircles, which can vary in number and catenation depending on the species. Maxicircles typically range from 20 to 40 kb, depending on the species, and they are present in a few dozen copies per network and they primarily encode critical mitochondrial enzymes. Minicircles, present in several thousand copies per network, are broadly similar in size (0.5 to 10 kb, depending on the species) but are heterogeneous in sequence (Lukeš et al., 2002) and they serve to function in the process of RNA editing. The kinetoplast has essential roles in reproduction, metabolism and segregation of the basal body and flagellum and it is probably critical for cyclical transmission by a vector (Gull, 2003).

1.2.3. Life cycles

Trypanosome parasites need a vertebrate host and an invertebrate vector to complete their life cycle (Awuoche, 2012). The two life cycle adaptations are related to whether the trypanosome belongs to the stercorarian or salivarian branch. The life cycle of a stercorarian parasite such as *Trypanosoma cruzi* begins when the blood feeding vector (triatomine bugs) ingests trypomastigotes along with its blood meal. These parasites differentiate into epimastigotes (proliferative form) within the insect hindgut and attach to the perimicrovillar membrane to allow metacyclogenesis and differentiation into the infective metacyclic trypomastigote (Fig 1.5A). When the insect host feeds, it often defecates at the same time, releasing metacyclic trypanosomes in the faeces, which may contaminate the wound or mucosa (Onyekwelu, 2019).

The vector of the salivarian branch, such as a tsetse fly, also ingests trypomastigotes whilst taking a blood meal. The parasite undergoes divisions in the insect mid-gut over a period of 1-2 weeks and then they differentiate into replicative procyclic forms. Over the next week, procyclic form migrate anteriorly in the gut, eventually entering the salivary gland (Fig 1.5B). Within the salivary gland procyclics further differentiate into replicative non-infective epimastigotes and then infective metacyclic forms (Smith et al., 2017). *Trypanosoma brucei* metacyclics are produced in the salivary glands whereas *T. congolense* metacyclics develop in the insect labrum and hypopharynx (Lloyd and Johnson, 1924, Thevenaz and Hecker, 1980). After inoculation into the bloodstream of a new vertebrate host, the cycle starts again, and the parasite continues to divide in this form.

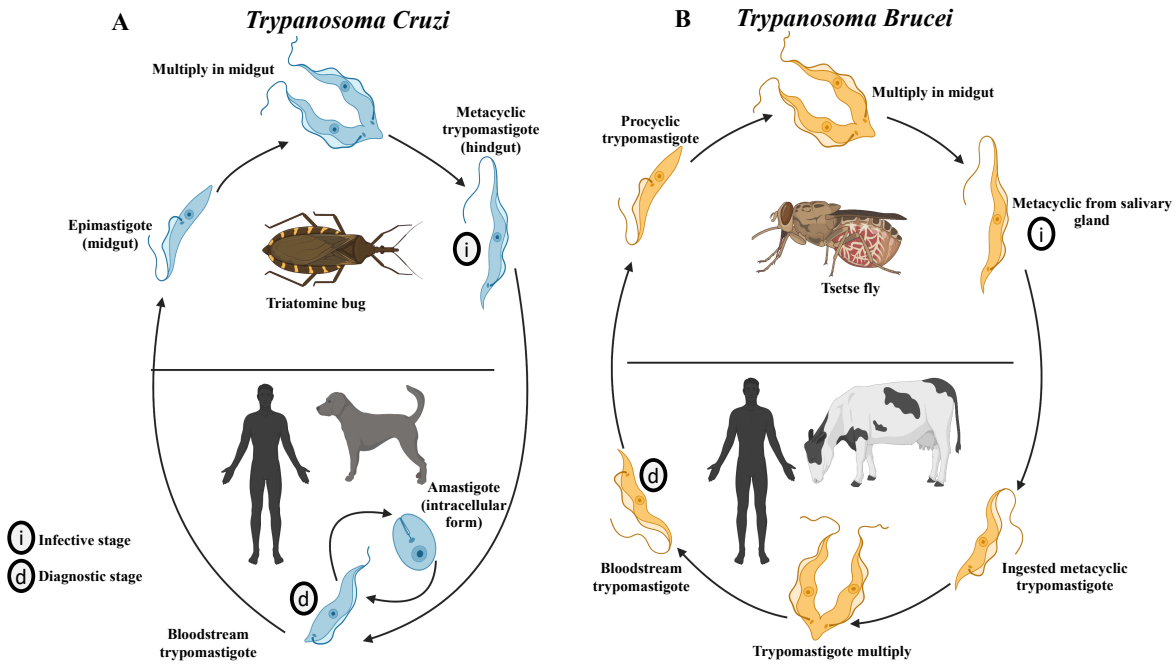


Figure 1. 5 The stercorarian and salivarian trypanosome life cycles as exemplified by *T. cruzi* (A) and *T. brucei* (B) respectively. infective stage = (i), diagnostic stage = (d). This figure was created with BioRender.com

Trypanosoma brucei trypomastigotes are maintained in the host bloodstream by rapidly proliferating (Smith et al., 2017), whereas *T. cruzi* trypomastigotes are phagocytosed by cells of the host immune system, or they enter other cells, such as cardiac muscle cells. During this intracellular phase of existence, *Trypanosoma cruzi* trypomastigotes differentiate into rounded

amastigotes (proliferative form). Amastigote multiplication inside cells or muscle fibres forms a pseudocyst. After a period of intracellular accumulation, the amastigotes are released into the blood whereupon they may invade other cells or tissues. Eventually, these amastigotes may transform again into trypomastigotes, escape the host cell by lysis, and enter into circulation. This latter stage is the one that may be ingested by a vector or infect a new host cell, again becoming an amastigote and repeating the cycle (Florin-Christensen and Schnittger, 2018).

1.2.4. Vector and transmission

Some *Trypanosoma* species transmit biologically as the parasites multiple and differentiate within the vector whereas others transmit mechanically since the vector acts like a flying syringe (Dagnachew and Bezie, 2015). Most African *Trypanosoma* species are transmitted by the bite of the tsetse fly (genus: *Glossina*), which inhabits the humid regions of Africa (Brun et al., 2010).

In sub-Saharan Africa, tsetse flies survive in temperatures of 16–38°C and 50–80% relative humidity (Franco et al., 2014). There are more than thirty species of flies and taxonomists have divided the tsetse fly family into three subgenera according to their habitat: *Nemorhina*, *Glossina sensustricto* and *Austenina* (Control et al., 1998). Both female and male tsetse flies are haematophagous and capable of transmitting infection (Holec-Gasior et al., 2013). In the rainy season, the tsetse fly population is increased, resulting in an increased risk of trypanosomiasis (Abenga, 2014). Tsetse flies can transmit and acquire parasites like other arthropods during feeding, and they serve as the intermediate host that is capable of transmitting thousands of trypanosomes via a single bite (Schmidt et al., 1977). The saliva of tsetse flies plays a significant role in parasite transmission and stimulation of the host's immunity system (Awuoche, 2012).

Some *Trypanosoma* species, such as *Trypanosoma vivax*, also are found in South America and they are transmitted by horse flies (from the Tabanidae family) and stable flies (from the *Stomoxys calcitrans* family) (Jones and Dávila, 2001, Oliveira et al., 2009).

Trypanosoma cruzi is transmitted through the faecal droppings of dozens of hematophagous triatomine bugs (kissing bugs) belonging to the order Hemiptera and the subfamily Triatominae. Triatomine bugs are widely distributed from the southern United States to Patagonia (Noireau et al., 2009).

Trypanosoma evansi is transmitted mechanically from the infected carrier animal to healthy animals by the mouthpart of a biting Insect; no parasite multiplication or biological association occurs between the parasite and vector. Transmission can occur through hematophagous Dipteran flies belonging the genera *Tabanus*, *Stomoxys* and *Hippobosca* (Desquesnes et al., 2009). In South America, *Trypanosoma evansi* is also transmitted by the biological vector-like vampire bat, *Desmodus rotundus* during its blood meal, which maintains the multiplicative cycle of *Trypanosoma evansi* in blood and tissues as a natural reservoir (Hoare, 1972, Raina et al., 1985). In addition to the above, there are other ways for *Trypanosoma* species to be transmitted between hosts. For instance, *Trypanosoma* infection can be acquired via the congenital route (Yadon and Schmunis, 2009). Laboratory accidents, often involving contaminated needles, can also lead to an infection; however, this is very rare. Furthermore, blood transfusion and organ transplantation can transmit *Trypanosoma* infection from one person to another, but again, these cases are also sporadic (Franco et al., 2014).

1.3. Trypanosomiasis

1.3.1. American trypanosomiasis: *Trypanosoma cruzi*

Trypanosoma cruzi (*T. cruzi*) is the etiologic agent of Chagas disease in human. *T. cruzi* can also cause infections in domesticated animals called American trypanosomiasis (Chagas, 1909).

1.3.1.1. Epidemiology

The parasite is transmitted primarily through contact with faeces of the insect vector (hematophagous triatomine bug) (Noireau et al., 2009). The endemic foci of *T. cruzi* infection is

restricted to South America, Central America, and parts of North America (Mexico and the southern United States) and is mostly associated with poor rural regions where houses are infested with domestic species of Triatominae (Schofield, 2000). Currently, oral infection by ingestion of contaminated food with the faeces of triatomine bugs is the most frequent route of transmission in Brazil and other Latin American countries (Santos et al., 2018, Shikanai-Yasuda and Carvalho, 2012). Animal trypanosomiasis is also widely distributed from the southern United States to Patagonia. More than 180 domestic, synanthropic and wild species of mammals, especially nest-building rodents and opossums (Marsupialia, Didelphidae), are likely to be infected by *T. cruzi* and hence probably involved in the contaminative transmission by metacyclic forms (Deane et al., 1984). Dogs and cats are primary reservoir hosts of the parasite in the domestic environment throughout the Americas and display much higher infectiousness to triatomine bugs than humans (Enriquez et al., 2014). Household dogs are estimated to be at a threefold higher risk for transmission of *T. cruzi* than cats (Gürtler and Cardinal, 2015). The primary mode of transmission in canine species seems to be the ingestion of infected vectors (Greene, 2006).

1.3.1.2. Pathology

T. cruzi infects all mammalian nucleated cells in any tissue and the diverse infectious forms contribute to the maintenance of *T. cruzi* in the host (Tomasini, 2018). Chagas disease occurs in 2 phases: acute and chronic. A critical period of Chagas disease is the acute phase, often accompanied by nonspecific clinical symptoms, such as fever, asthenia, face and limb oedema, headache, myalgia, and other ailments including minor bleeding; most commonly from the nose, skin petechiae, or bruising. Occasionally, the risk of thromboembolism and digestive bleeding is reported and may cause death (Souza et al., 2018). After the acute phase, parasites disappear from the blood, and infected people become asymptomatic for a prolonged or indefinite which is called the indeterminate stage. A chronic *T.cruzi* infection phase occurs years after the initial infection and some people develop signs of organ failure. *T.cruzi* mainly affects the heart and digestive system (Hidron et al., 2010, Jr et al., 2000).

The acute phase of *T. cruzi* infection in an animal is very similar to that observed in humans and involves fever and nonspecific signs of illness such as anorexia, lethargy, an unkempt hair coat, lymphadenopathy, enlargement of the liver and spleen, diarrhoea, ascites and palpebral oedema. After the acute phase, experimentally infected dogs become asymptomatic for a prolonged period, but some animals eventually develop congestive heart failure, visceral damage, and progressive neurological disturbances.

1.3.2. Human African Trypanosomiasis (HAT): *Trypanosoma brucei gambiense, Trypanosoma brucei rhodesiense*

The causative agents of HAT or sleeping sickness are *Trypanosoma brucei gambiense* (*T.b. gambiense*) and *Trypanosoma brucei rhodesiense* (*T.b. rhodesiense*), which are both salivarian pathogenic trypansosmes (Gibson, 2001).

1.3.2.1. Epidemiology

In 2012, over 7000 cases of trypanosomiasis were reported in 30 countries across sub-Saharan Africa (Control et al., 2013). In west and central Africa, *T.b. gambiense* was responsible for human trypanosome disease (Radwanska et al., 2018); in contrast, *T.b. rhodesiense* was responsible for human trypanosomiasis in eastern, central and southern Africa (Steverding, 2008). Humans can be an epidemiological reservoir for *T.b. gambiense* (Control et al., 2013), whereas, wild and domesticated animals, including giraffes, lions, zebras and cattle, are the reservoir hosts for *T.b. rhodesiense* (Hide et al., 1996, Von Wissmann et al., 2011).

A bite from a tsetse fly of the genus glossina can transmit *Trypanosoma brucei* between humans or from animals to humans (Schmidt et al., 1977). The palpalis group of tsetse flies, which includes *G.p. palpalis* and *G.p. gambiensis*, is the main subgenus of tsetse fly in West Africa (Ravel et al., 2006). *G. morsitans* and *G. pallidipes* are distributed mainly in East Africa, and they might be responsible for the transmission of *T.b. rhodesiense* (Holmes, 2015). *G. fuscipes* is present in South Africa, and it transmits *T. b. rhodesiense* by feeding on humans and animals (Clausen et al., 1998).

African trypanosomiasis is mainly seen in rural communities and impoverished areas (Maxfield and Bermudez, 2020). Sporadic cases are observed in short-term tourists and visitors, but these cases are reported primarily in immigrants and refugees who reside for long periods in rural areas. For example, two cases of *T.b. gambiense* have been reported in Italian nationals that resided in an area of high endemicity (Bisoffi et al., 2005).

About 98% of HAT cases between 2009 to 2013 were due to *T.b. gambiense* infection and 2% were due to *T.b. rhodesiense* (Franco et al., 2014). Most reported cases (~25,000 cases) in 1990 were *T. b.gambiense* in the Democratic Republic of Congo (Lumbala et al., 2015). In contrast, most of the reported cases of *T. b.rhodesiense* disease are in Zambia (15 cases) and Malawi (91 cases) according to the WHO data in 2019 (WHO,2020).

The WHO has worked with multiple non-governmental organizations to organize a massive campaign to eliminate gambiense HAT in epidemic areas by exterminating tsetse flies and donating the drugs required for treatment of the disease (WHO, 2020). As a result, in 2015, the number of gambiense HAT cases dropped below 3000 for the first time for 50 years (Attardo et al., 2019), and fewer than 1500 cases of gambiense HAT were reported in 2017 (Barrett, 2018). Controlling *T.b. rhodesiense* is also important due to studies that have shown that wild and domesticated animals, including cattle (Selby et al., 2013), pigs (Hamill et al., 2013), sheep and goats (Ruiz et al., 2015) can act as reservoir hosts for human infection.

1.3.2.2. Pathology

T.b. gambiense is responsible for the chronic form of sleeping sickness (Steverding, 2008). Most people infected with chronic *T.b. gambiense* can live asymptomatically for many years. During this time, they can be a carrier and transmit the infection to other humans (Checchi et al., 2008). *T.b. rhodesiense*, discovered over 100 years ago (Hide, 1999), causes the acute form of sleeping sickness, which rapidly progresses toward death within six months if untreated or inadequately treated (Kennedy, 2013, Kennedy, 2019). In the United Kingdom, two cases of HAT due to *T.b. rhodesiense* were reported in two men who travelled to Africa. One case returned from vacation

in the in the Luangwa Valley of southern Zambia. A second returning traveller had the infection during vacation to Kenya and Tanzania (Moore et al., 2002).

Early stage sleeping sickness, also called the haemolymphatic phase, is defined by restriction of trypanosomes to the blood and lymph system. In this stage, symptoms are usually nonspecific and may include a fever up to 40–41ºC, headaches, joint pains and itching (Steverding, 2008) As the disease develops, patients might develop various symptoms, including lymphadenopathy, congestive cardiac failure and enlargement of the spleen and liver. During the late stage or neurological phase of sleeping sickness, the parasites cross the blood-brain barrier (BBB) and enter the cerebrospinal fluid (CSF). Patients in this stage are characterized by confusion, a disturbed sleep pattern, sensory disturbances, tremors in the hands and tongue, extreme lethargy, and coma (Kennedy, 2008). In total, 74% of HAT patients experience common sleep disturbances, giving the disease its name (Buguet et al., 2005).

1.4. Animal Trypanosomiasis

1.4.1. Surra disease: *Trypanosoma evansi*

Trypanosoma evansi (T. evansi), which is closely related to *T.b. brucei,* is a significant veterinary problem that affects domestic, wild and livestock animals (Singh and Chhabra, 2008). *T. evansi* was the first pathogenic trypanosome identified by Griffith Evans in 1880 in India (Fallis, 1986). Its main difference from the other trypanosomatids is the lack of maxicircle kinetoplast DNA (Lun and Desser, 1995). Moreover, *T. evansi* is different from *T. congolense* and *T. vivax*, which are known as blood parasites, because the former can penetrate both blood and tissue due to its ability to invade the host nervous system. Its principal host is originally the camel, but it is present in horses and dogs but also cattle, buffaloes and pigs (Hoare, 1972).

1.4.1.1. Epidemiology

T. evansi spread outside the tsetse belt in Africa, towards the Middle East and Southern Asia, and was exported with livestock to Latin America and Europe (Diall et al., 1993, Garcia et al., 2005, Holland et al., 2005, Silva et al., 1995, Tuntasuvan et al., 1997, Gutierrez et al., 2006). *T. evansi* is the most prevalent trypanosome affecting livestock in Asia (Kumar et al., 2012, Mekata et al., 2013). *T. evansi* is mainly mechanically transmitted by hematophagous flies from the genera *Stomoxys* and *Tabanus* (Hoare, 1972). *T. evansi* transmission can occur by the common vampire bat or oral transmission (Aregawi et al., 2019).

1.4.1.2. Pathology

T. evansi was found in the blood of horses and camels suffering from an endemic disease known as trypanosomosis or surra. Surra is a significant disease in camels (Zewdu et al., 2016), horses (Rodrigues et al., 2009), dogs, pigs, cattle and buffalo (Holland et al., 2003), in which it can often be fatal in the absence of treatment. Surra leads to progressive emaciation, anaemia, pyrexia, reduced weight gain, lowered milk and meat yields, abortion and mortality in livestock (Desquesnes et al., 2013).

1.4.2. African Animal Trypanosomiasis (AAT): *Trypanosoma congolense, Trypanosoma vivax*

The most common *Trypanosoma* species in animals are *Trypanosoma congolense* (*T. congolense*), discovered by Broden in the blood of sheep and donkeys (Hoare, 1972), and *Trypanosoma vivax* (*T. vivax*), which was described for the first time in Africa by Ziemann in 1905 (Leger and Vienne, 1919). These trypanosomes infect wild and domestic animals, causing African Animal Trypanosomiasis or nagana disease which manifests as a relatively mild infection in some wild animals but in domesticated animals the disease is often severe and even fatal (Chitanga et al., 2013, Yaro et al., 2016). *Glossina morsitans spp.* and *Glossina pallidipes* are the main species of tsetse flies that transmit AAT (Van den Bossche et al., 2010).

1.4.2.1. Epidemiology

T. congolense is spread across much of sub-Saharan Africa, being reported in more than 37 countries in Africa, from the southern edge of the Sahara Desert to Zimbabwe, Angola and Mozambique (Cecchi et al., 2014). The cyclical transmission of *T. congolense* to susceptible hosts is mainly achieved through the infected tsetse fly (*Glossina* species). For example, the *G. morsitans* group can transmit the savannah, and Kilifi types and the *G. palpalis* group can transmit the Riverine-Forest type (Malele et al., 2011, Rodrigues et al., 2014, Simo et al., 2013). *T. congolense* has been classified into three different types that are morphologically identical (Godfrey, 1961). The savannah type is the most pathogenic, while the riverine-forest and Kilifi type are moderate and have low pathogenicity (Abrham et al., 2017, Knowles et al., 1988). In south, east and central Africa, *T. congolense* savannah and Kilifi types have a wide distribution and affect any livestock species, while *T. congolense* savannah and riverine forest types are distributed in west and central Africa (Rodrigues et al., 2014).

T. vivax is also spread across much of sub-Saharan Africa (Auty et al., 2015), as well as being found in 13 South American countries. The parasite probably arrived in South America in the eighteenth and nineteenth centuries in infected Zebu cattle exported from West Africa (Jones and Dávila, 2001). In Africa *T. vivax* is predominantly transmitted by tsetse flies following cyclical development; there are at least nine groups of *Glossina* vectors that can transmit *T. vivax*, including *G. morsitans*, *G. longipalpis*, *G. palpalis* and *G. tachinoides* (Ooi et al., 2016). Outside tsetse Fly areas, including South America, *T. vivax* is transmitted by hematophagous flies such as horseflies (tabanids) and stable flies (*Stomoxys*) (Ooi et al., 2016, Otte and Abuabara, 1991).

1.4.2.2. Pathology

T. congolense was the first identified causative agent of AAT, and it is responsible for more than 80% of cases and losses of domestic animals (Namangala and Odongo, 2014). In cattle, *T. congolense* is the most widespread and virulent species that cause AAT (Desta et al., 2013). Many have demonstrated that *T. congolense* is a highly pathogenic parasite in cattle (Muhanguzi et al.,

2017, Mungube et al., 2012, Ngomtcho et al., 2017). *T. congolense* mainly invades circulating blood and it is rarely found in tissues (Silva Pereira et al., 2019); the exception being during development of an infection at the site of inoculation, where the parasite is found in the skin, outside the vascular system and in the lymphatic system (Luckins and Gray, 1978, Luckins et al., 1994).

T. vivax is the most prevalent trypanosome species in cattle reared in west and east Africa (Gardiner, 1989). A few east African isolates of *T. vivax* cause an acute haemorrhagic syndrome with a mortality rate of 6–35% (Njiru et al., 2011, Magona et al., 2008). *T. vivax* causes different patterns of AAT, which is more severe and acute in Africa, while in South America, AAT occurs as rare epizootic outbreaks in cattle (Osório et al., 2008).

AAT is considered a significant factor leading to severe economic loss related to livestock production and it threatens financial and food security in sub-Saharan Africa and South America (Holt et al., 2016). The disease causes about 3 million deaths of livestock annually, leading to annual livestock production losses of about 1.2 billion USD (Ooi et al., 2016). Acute clinical signs of AAT include fever, listlessness, decreased fertility, emaciation, reduction of milk yield, hair loss, discharge from the eyes, oedema, anaemia, paralysis and, in the absence of treatment, death (Eisler et al., 2004). As the illness progresses, animals become weakened and, eventually, unfit for work. Post-mortem signs include swollen lymph nodes, enlarged spleen and liver, excessive fluid in the body cavities and petechial haemorrhages (Tesfaye et al., 2012).

1.5. The immune response against trypanosome infection: innate and adaptive immunity

The immune response to the infection of *Trypanosoma* species is determined by the interaction of immune-related molecules of the host cells with parasitic molecules (Tarleton, 2007). It is believed that *Trypanosoma* infection is controlled at three levels. First, *Trypanosoma* is detected and

directly eliminated by professional antigen-presenting cells (APCs), innate immune cells, such as macrophages, and dendritic cells (DCs). The most important molecules involved in the detection of *Trypanosoma* infection by innate immune cells are pattern recognition receptors (PRRs), such as Toll-like receptors (TLRs). PRRs are defined as receptors that detect pathogen-associated molecular patterns (PAMPs), which can stimulate and activate innate immune responses. Known PAMP molecules in *Trypanosoma cruzi* are glycosylphosphatidylinositol (tGPI), glycoinositolphospholipid (eGIPL), and unmethylated CpG motifs. These molecules interact with TLR2, TLR4 and TLR9, respectively (Dos-Santos et al., 2016). Although the innate immune system recognizes pathogens, its response does not relate to previous infections, and it does not change qualitatively or quantitatively after contact (Medzhitov and Janeway, 2000). The interaction between PRRs and *Trypanosoma* species induces potent pro-inflammatory cytokines, such as interleukin-12 (IL-12) and type-1 interferon (INF-1), in innate immune cells to activate the adaptive immune response (Fig. 1.6; (Bafica et al., 2006). A defective adaptive immune response for controlling *Trypanosoma cruzi* was reported in a TLR-deficient mice model (Bafica et al., 2006).

However, adaptive immune cells, such as T cells ($CD8^+$ T cells), can recognize trypanosome in different tissues. During a trypanosome invasion, *Trypanosoma* species release proteins that present in the form of a major histocompatibility complex type 1 (MHC-1), which interacts with $CD8^+$ T cells (Garg et al., 1997). Moreover, the detection of trypanosomes by non-immune cells, such as cardiocytes and adipocytes, is enabled by the interaction of TLR-PAMP molecules. For instance, cardiocytes stimulated with trypomastigotes or supernatants from trypomastigote cultures induce IL-1-β, which is dependent on TLR2 (Petersen et al., 2005).

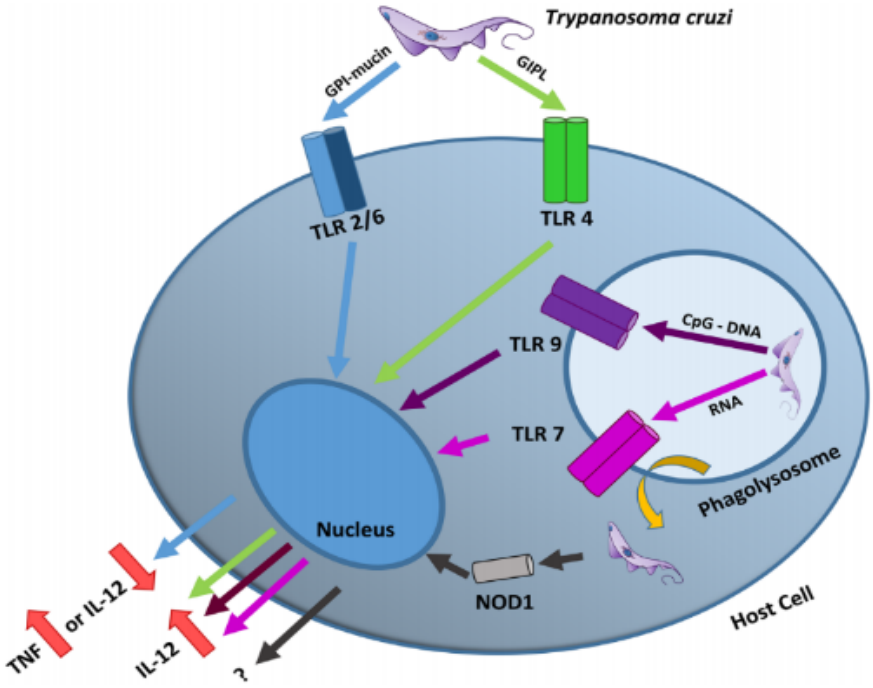


Figure 1. 6 *Trypanosoma cruzi* and TLR activation. The extracellular TLR such as TLR4 and heterodimer TLR2/6 detect GIPL and GPI-mucin. The genomic pattern CpG-DNA is detected by intracellular TLR such as TLR7 and TLR9. The activation of these TLR induces expression of inflammatory cytokines. This figure is adapted from (Cardoso et al., 2016).

1.5.1. Immune evasion strategies of trypanosome infection

Trypanosomes have evolved strategies to prevent killing by the host immune response and, thus, maintain a long-term infection in a mammal. Parasites undergo many life cycles changes in vectors and mammalian hosts with significant morphological and metabolic changes (Li, 2012, Rodrigues et al., 2014). The human plasma contains trypanosome lytic factor (TLF) or high-density lipoproteins that can kill extracellular protozoon parasites of the genus *Trypanosoma.* However, two subspecies of *Trypanosoma brucei (T. b. gambiense* and *T. b. rhodesiense*) have evolved a resistance mechanism to TLFs, resulting in the disease human sleeping sickness (Capewell et al., 2011). The TLFs contains a unique multi-protein complex consisting of the components called haemoglobin-binding protein, haptoglobin-related protein and pore-forming protein apolipoprotein- L1 (apoL-1) (Thomson et al., 2009). Impairment in any of these components in

humans may cause a zoonotic spill over of *Trypanosoma* as previously reported for human infection with *T. evansi* (Vanhollebeke et al., 2006).

Additionally, the trypanosome genome encodes variant surface glycoproteins (VSGs) which are expressed on the surface (Kennedy, 2013). The VSG repertoire can avoid the innate immune response during infection (Mansfield and Paulnock, 2005) and most importantly, through variable expression of VSG genes, enable the parasite to evade the host IgM and IgG immune responses (Brun et al., 2010). Therefore, developing a vaccine based on humoral immunity could be extremely challenging, if not impossible, to achieve.

1.6. Tools and techniques used to determine genetic diversity

1.6.1. Flinders Technology Associates cards

Flinders Technology Associates (FTA) paper was patented by Whatman to simplify the handling and processing of nucleic acids at ambient temperatures (Smith and Burgoyne, 2004). Then, Lee Burgoyne at Flinders University in Australia developed a method for storing DNA (Butler, 2011). The FTA card is a novel, easy and fast technology that is suitable for human, animal and plant biological samples (Grund et al., 2010, Smith and Burgoyne, 2004). It is used for collection, stabilization and storage of biological samples at room temperature for several years (Jignal et al., 2014). FTA cards can also be used to collect and transport the DNA of pathogens from hazardous or dangerous environments or harmful viruses, such as HIV (Pulido-Landínez et al., 2012). Some previous studies have evaluated the possibility of using FTA cards to store RNA from viruses (Biswal et al., 2016), mosquitoes, protozoa, trypanosome, plasmodia, bacteria and plant cells (Natarajan et al., 2000).

An FTA card is an absorptive cellulose paper card that consists of chemicals that lyse cells, denature proteins, inactivate pathogens and protect DNA from nucleases, oxidation and UV damage. There are two main formats of FTA cards that differ in terms of their chemical

composition. The first is the FTA classical card, on which DNA remains tightly bound to the FTA paper during analysis and washing of proteins and other inhibitors. The classical FTA card has chemical denaturants and a free radical scavenger that encourage DNA to bind to the paper (Ali et al., 2017). The second format is the FTA elute card, on which DNA is eluted, and proteins are bound in a matrix before analysis. FTA elute cards contain a chaotropic salt to elute DNA from the card (Cook 2015). FTA classic cards preserve the high molecular weight of DNA and produce higher DNA yield than the FTA elute cards (Saieg et al., 2012). Applying samples, such as bone marrow aspirates, cultured cells, plasma, blood, saliva, or tissue samples from animals, insects, microorganisms, soil, water or processed food, to an FTA card is very simple. A low quantity of the sample (a few microliters or a few milligrams) is placed onto the card and left to air dry (Cook, 2015). Then, after applying the sample to the FTA card, a sample disc is removed from the card. A paper disc puncher is used to obtain a small disc ranging from 1.2–6 mm (Cook, 2015). DNA released from FTA cards proved to be a suitable substrate for molecular methods (Mas et al., 2007).

However, there are many limitations of FTA cards that makes them impossible to use for all biological analyses. A previous study suggested using more than two discs (3 mm in size) from FTA cards to increase the sensitivity for detection of microorganisms (Cox et al., 2010). Moreover, FTA cards lack morphological assessment and cannot be used for some biological tests, such as total cell count, determination of the tumour cell percentage and other histological analyses (da Cunha Santos, 2018, da Cunha Santos et al., 2018).

1.6.2. Next-generation sequencing

1.6.2.1. Overview

DNA sequencing is revolutionizing biology, as it allows researchers to determine the order of nucleotides in each DNA molecule. Using knowledge from DNA sequencing, we can identify genetic information, transcripts and proteins downstream and clarify the biochemical processes that exist in an organism (Sanger et al., 1977). In 1975, Sanger and colleagues devised a method of DNA sequencing called first-generation sequencing, which produced detailed knowledge about many prokaryotic and several eukaryotic genomes. The main project achieved with first-generation sequencing was the mapping of the human genome (Lander et al., 2001). In 2005, next-generation sequencing (NGS) was developed, solving some of the limitations of first-generation sequencing. For example, its increased throughput by sequencing many DNA molecules in parallel decreased the cost of sequencing, improved accuracy and required only a small amount of starting martial. NGS has been applied to address important questions related to genomics, epigenomics, transcriptomics and metagenomics (Schadt et al., 2010). The principle of NGS is like Sanger sequencing, but instead of the normal deoxynucleotide triphosphates (dNTPs) used in Sanger sequencing, NGS uses fluorescently labelled dNTPs (Mardis, 2013). Also, in Sanger sequencing, cloning is required, but in NGS, this step is not required.

The NGS method has four general steps. Initially, DNA is split into small random fragments either enzymatically or via sonication, which are ligated to short adaptors or double-stranded pieces of synthetic DNA to allow production of a DNA library. The second step is the amplification of the library to increase the efficiency and accuracy with which the signal can be detected from the sequencer. Thirdly, to ensure successful cluster generation for the sequencing process, library fragments are loaded into the flow cell and then amplified into a clonal cluster through bridge amplification. The last step is sequencing by adding fluorescently labelled nucleotides. Each DNA

cluster releases fluorescent signals of different lengths and intensities that identify the bases (Mardis, 2013).

1.6.2.2. Applications

1.6.2.2.1. Genomic

The genomic applications of NGS include sequencing a new genome to address fundamental questions about an organism. Genomic applications are divided into two groups: *de novo* genome sequencing and variant calling applications. *De novo* genome sequencing is the main application of NGS technology. This process produces an enormous amount of data for a *de novo* genome that does not have an available reference sequence that is assembled of contigs. The *de novo* genome sequencing strategy can be performed routinely, and the necessary data analysis tools are readily accessible (Salzberg and Yorke, 2005). Variant calling applications are used to identify genetic variants. These applications depend on the information provided in a sequence alignment map (SAM) and binary alignment map (BAM). Meticulous inspection and confirmation of mismatches to the reference reveal variations due to, for example, SNPs or insertions and deletions (Buermans and Den Dunnen, 2014).

1.6.2.2.2. Transcriptomic

The transcriptomic application of NGS identifies the RNA sequences (RNA-Seq) in a given RNA sample that a ribosome removes since cells have different types of RNA (e.g. ribosomal RNA, mRNA or transfer RNA). Before starting an NGS library, obtained RNA samples must be converted to cDNA via reverse transcription. RNA-Seq is widely used to study gene expression, find novel spliced junctions, apply the antisense regulation mode and compare gene expression differences under varying conditions or disease states (Tarazona et al., 2011).

1.6.2.2.3. Epigenomics

Epigenomic applications are different from the other applications because they are not focussed on genetic information or quantification of gene expression, but the functions of a genome change and chemical modifications of genomic DNA and histone proteins (Zhang and Jeltsch, 2010). NGS is combined with chromatin immunoprecipitation (ChIP) to create ChIP-seq, a powerful method for identifying genome-wide DNA binding sites of proteins or associated histones (Nakato and Shirahige, 2016). The bisulfite (BS) sequencing of DNA has become the gold standard for analysis of DNA methylation, which is the conversion of unmodified cytosines to uracil whilst maintaining the fifth ring position in cytosine (5mC) unchanged. NGS has been coupled with BS to yield reduced representation (RRBS) or whole genome (WGBS) data on the global genomic distribution of 5mC (Harris et al., 2010). Moreover, immunoprecipitation sequencing (MeDIPseq) is commonly used to study 5mC or 5hmC modification by using specific antibodies (Clark et al., 2012).

1.6.2.2.4. Metagenomics

Metagenomics claims that the human body has more bacterial cells than human cells and that it is important to closely study the microbial communities in a human body because they could influence human health. According to prior studies, it is hypothesized that there is a strong relationship between the alteration of microbial compositions and obesity, inflammatory bowel disease and other immune-inflammatory disorders (Guinane and Cotter, 2013). For this reason, there has been great interest in understanding the microbiome using NGS. The Human Microbiome Project was intended to identify variations in the microbiomes of different subjects (Malla et al., 2019). For microbiome profiling, 16S ribosomal RNA is widely used to identify the composition of a microbial community and determine microbial diversity with phylogenetic clustering. The advantage of using 16S ribosomal RNA is that it is an excellent phylogenetic

marker, containing nine variable regions surrounded by a constant sequence that can be used as a target for a primer for PCR amplification. After that, amplicons can be sequenced using NGS technologies to understand the composition and diversity of the microbial community. Alternatively, metagenomic sequencing can aim to sequence the whole DNA of a microbial community, not focusing on the variety of composition of the microbiome but on identifying the gene content and the functional potential of a specific gene (Wang et al., 2013).

1.7. Hypothesis and aims of the project

The hypothesis in this study is that genetic variation of TLR9 could play an essential role in susceptibility or resistance to parasitic infection. Therefore, we aim to investigate the genetic diversity in the bovine TLR9 gene from blood spots taken from FTA cards using Sanger sequencing and NGS, and how this variation is linked to the parasitic infection. Three goals will be taken to achieve these aims:

- Development of efficient methods and techniques for DNA extraction from FTA cards for TLR9 genotyping.
- Identification of genetic variants using Sanger sequencing in the TLR9 coding region and their association with bovine trypanosome infection.
- Identification of genetic polymorphisms within TLR9 coding region using the NGS technique and their association with bovine trypanosome infection.

CHAPTER TWO

Materials and Methods

2.1. DNA Samples

All bovine blood samples used were collected from a previous research project and these blood samples were stored on Whatman FTA cards at room temperature (Ideozu., 2015). According to the Ideozu (2015) thesis, a of total of 80 bovine blood samples used in this study were of northern origin within Nigeria; however, the cattle were subsequently transported down south by herdsmen. All the samples were derived from the *Bos indicus* subspecies; a major feature for identifying these species is the possession of a humpback (Loftus et al., 1994). The blood from each sample was collected by butchers from an abattoir in Ahoada, a town in the Orashi region of Rivers State, southern Nigeria (Ideozu, 2015).

2.2. Preparation of FTA cards for DNA extraction

The FTA cards were placed on a clean mat before a 2mm punch was taken using a Harris Micro Punch (Sigma-Aldrich Company). After that, these small discs (2-5 disc) were placed in a sterile Eppendorf tube (1.5ml). To ensure there was no cross-contamination between samples, blank filter paper or blank FTA card punches were used. In this study, three extraction methods: 5% Chelex resin, REPLI-g Midi Kit (QIAGEN), and DNeasy Blood & Tissue Kit (QIAGEN) were used. DNA concentrations were measured using Spectrophotometry (Nanodrop) (ThermoFisher Scientific).

2.3. Chelex 100 Matrix Extraction

The FTA card discs were washed moderately and incubated at room temperature for 15min using the FTA purification reagent (Thermo Fisher Scientific). This step was repeated three times before incubating the discs for 15min with TE-1 Buffer (10mM Tris-HCL, 0.1mM EDTA, pH 8.0) at room temperature to avoid DNA degradation. This step was repeated twice. Then, the discs were

air dried for one hour, or at 56 °C for 10 min. Then, DNA was eluted by adding 200 µl of Chelex solution (Sodium form, 100-200 mesh; Bio-Rad Laboratories, Hemel Hempstead, Hertfordshire, UK) to the discs and then incubating them at 90 °C for 30 min. All purified DNA samples are stored in a freezer at -20 ˚C for later use.

2.4. Multiple displacement amplification

The multiple displacement amplification MDA or whole-genome amplification (WGA) procedure is a non-PCR method to amplify a small amount of DNA (Spits et al., 2006). The amplification and extraction of DNA from FTA cards were carried out according to the manufacturer instructions. In brief, firstly, buffers D1 (Denaturation buffer) and N1 (Neutralisation) were prepared. D1 buffer contains 9 µl of reconstituted buffer (DLB) and 32 µl Nuclease-free water. The N1 buffer contains 12 µl of stop solution and 68µl Nuclease-free water. Then, 2 discs from each FTA cards were mixed with 2.5µl D1 solution in centrifuge tubes. Then, all centrifuge tubes were vortexed and centrifuged 17 000× g for 30s. The mixture was incubated at room temperature for 3 min before 5µl of N1 solution was added, followed by centrifugation at 17 000× g for 30s. After that, 29 µl REPLI-g Mini reaction buffer was added to all sample tubes followed by 1µl REPLI-g Mini DNA Polymerase and 10µl of Nuclease-free water before incubation overnight at 30 °C. To inactivate the DNA Polymerase, the samples were incubated at 65 °C for 3min. after finishing the procedure, all products were cleaned using ISOLATE II PCR and Gel Kit (Bioline) to remove all reagents. The purified DNAs were stored at -20 °C.

2.5. DNeasy Blood & Tissue Extraction Kit (QIAGEN)

The manufacturer instructions were followed with minimal modification. Briefly, 280 µl of Tissue Lysis (ATL) buffer was added to 2 FTA discs, followed by 20µl of Proteinase K (QIAGEN) then

each tube was vigorously vortexed for 30s. The tubes were incubated at 56°C for 60 min and then each sample was transferred to a microcentrifuge tube and centrifuged at 11000×g for 30s. An aliquot of 300 µl of lysis buffer (AL) was added to each tube and then samples were incubated at 70 °C for 10 min before centrifugation at 11000×g for 30s. Each sample was transferred to a mini spin column and centrifuged at 6000×g for 1 min. Then, 700 µl of AW2 (wash buffer) was pipetted onto each spin column and centrifuged for 1 min at 6000× g. The spin column was transferred to a new collection tube and 700 µl of 100% ethanol was added before the sample was centrifuged for 1 minute at 6000×g. After that, the spin columns were transferred to a new collection tube and centrifuged at 17 000× g for 5 min. The column lid for each sample was opened to dry for 10 min at room temperature. After drying, the spin column was placed in a 1.5 ml centrifuge tube, and 20 µL of elution buffer (EB) was added onto each spin column, incubated at room temperature for 5 min at room temperature and centrifuged 17 000× g for 3 min. This flow-through was then stored at -20°C.

2.6. PCR Amplification

Each PCR reaction was carried out using a 25 µl reaction mixture containing 12 µl High-Fidelity DNA Polymerases & Master Mixes (Thermo Fisher Scientific), 5µM of each primer (forward and reverse) (Table 2.1) (Table 2.2), and either 20-25 ng/µl DNA extracted from FTA cards or 1µl bovine genomic DNA as a positive control (Amsbio) (http://www.amsbio.com). Nuclease-free water was used as the negative control in all PCR reactions.

2.6.1. Primer-BLAST

Primer-BLAST is one of the most used tools for designing primers. This tool is also available for free on NCBI (http://www.ncbi.nlm.nih.gov/). All primers were created by a software called snap-gene (http://www.snapgene.com) and the specificity of primers was checked on Primer Blast

(http://www.ncbi.nlm.nih.gov) before they were ordered from Eurofins MWG Operon (http://www.eurofinsgenomics.eu/). The nuclease-free water was added to primers to a stock concentration of 100 pM/µl. The primer solution was then thoroughly mixed by vortex and centrifuged at 11,000×g for 30sec. We prepared 20µM working primer concentration by making a 1:10 dilution of the stock primer. Aliquots were made of the 20µM working concentration. Both stock and working stock primers were frozen at -20 °C.

2.6.2. Amplification of Housekeeping Gene

For detection of tubulin gene has been used to evaluate the quality of the genomic DNA in humans, rats, mice, sheep, cattle, voles, badgers, foxes, deer and bats (Rebouças et al., 2013). In this report, the primer sequence designed for detection the tubulin gene was obtained from the previous study (Terry et al., 2001) in our lab (Table 2.1). PCR amplification was carried out with 5 min at 94 °C, followed by 40 cycles of the 40s at 94 °C, 40s at 60 °C, 90 min at 72 °C with a final cycle of 10 min at 72 °C.

2.6.3. TLR9 Amplification PCR

B. indicus TLR9 gene sized 4831bp (Accession no KC174788.1) was obtained from NCBI website (https://www.ncbi.nlm.nih.gov) for primer design purposes. The primers targeting TLR9 was designed by to cover most of TLR9 gene (Table 2.1) (Ideozu, 2015). The reaction was carried in the following thermocycling conditions: 5 min at 95 ^{0}C, followed by 35 cycles of the 30s at 95 ^{0}C, 45s at 55 ^{0}C, 90 min at 72 ^{0}C and a final cycle of 5 min at 72 ^{0}C.

2.6.4. TLR9 Hemi-nested PCR

We prepared hemi-nested PCR by using two amplification rounds. The first round was amplified using the TLR9 amplification conditions as previously described in (section 2.6.3). The second PCR round was targeted at CpG islands in TLR9 using another forward primer (Hemi-CpG_F) (Table 2.1) with the whTLR9_R primer. The reaction was carried out with 10 min at 94 °C,

followed by 35 cycles of the 30s at 95 °C, 30s at 60 °C, 30s at 72 °C and a final cycle of 7 min at 72 °C.

2.6.5. Amplification of the CpG island Island within the TLR9 gene

This PCR was designed to amplify the CpG island 1 within exon 2 of the bovine TLR9 gene. In this PCR, primers were designed (CpG_F and CpG_R) (Table 2.1) and the reaction was carried out with 10 min at 94 °C, followed by 40 cycles of the 30s at 95 °C, 30s at 60 °C, 30s at 72 °C and a final cycle of 10 min at 72 °C.

Table 2. 1 Forward (F) and reverse (R) primers sequence for tubulin and TLR9 amplification.

PCR	Primer	Size (bp	Sequence
Housekeeping Gene (tubulin)	MTUB_F	1196	5'-CGT GAG TGC ATC TCC ATC CAT-3'
	MTUB_R		5'-GCC CTC ACC CAC ATA CCA GTG -3'
TLR9	whTLR9_F	4156bp	5' – GGA GAA GCC GCA TTC CCT G - 3'
	whTLR9_R		5' –TGT GGG GTT AAA GGA GTG CTG- 3'
TLR9 Hemi-nested	Hemi-CpG_F	1161	5'-CTG CGT CTC CGG GAC AAT AAC-3'
CpG Island	CpG_F	1129	5-' GCGTCTCCGGGACAATAA -3'
	CpG_R		5' – CTCTGTGCTATTCGGCTG - 3'

2.6.6. Fragmentation of Exon-2 within TLR9 gene

To amplify the entire 1129bp exon 2 of the TLR9 gene different primers were designed (Table 2.2) in order to obtain a high resolution of DNA sequencing reads. In summary, exon 2 of the TLR9 gene was divided into four segments to allow four independent and overlapping PCR fragments to be generated. These four fragments were amplified by used primers in Table 2.2. The reaction was carried out with 10 min at 94 °C, followed by 40 cycles of the 30s at 95 °C, 30s at 53 °C -56 °C -60 °C -63 °C, 30s at 72 °C and a final cycle of 10 min at 72 °C.

Table 2. 2 Forward and reverse primer sequences targeting four fragments of CpG islands within exon two of bovine TLR9.

Primer	Size (bp)	Forward	Reverse
f1TLR9	554	5'-GCGTCTCCGGGACAATAA-3'	5'- GGT CCC AGC CAC AGA G-3'
f2TLR9	316	5'-AAGACAGTGGATCCCTCCTG -3'	5'- GGT CCC AGC CAC AGA G-3'
f3TLR9	616	5'-CTC TGT GGC TGG GAC C -3'	5'- GGT TAA AGG AGT GCTGGG CA-3'
f4TLR9	318	5'-ACA CGG ACC GGG TCA-3'	5'-GGT TAA AGG AGT GCT GGG CA-3'

2.7. Gel electrophoresis

A 1% (w/v) agarose gel was prepared by weighing out 1.00g of agarose powder (Sigma-Aldrich) diluted in 100 ml of 1X Tris-Borate-EDTA (TBE) buffer (ThermoFisher Scientific). The mixture was added to the flask and heated in a microwave on maximum power for 30s, mixed gently and heated for a further 30s until all the agarose had dissolved. The molten gel was placed on a shaker and allowed to cool to approximately 50°C and then 1 µl of Gel Red (Biotium) was added and mixed by swirling gently. The molten agarose was then poured into a Perspex plate which had a comb already placed inside to form wells and taped on the sides to prevent spillage. After 30min, the gel was placed in an agarose electrophoresis chamber and submerged with 1X TBE buffer.

A 100 ng/µl kb ladder (Bioline) was prepared by adding 700 µl of nuclease-free water to a 2.0 ml screw cap tube followed by 100µL of 1Kb DNA ladder (Bioline) and 200µL bromophenol blue loading dye (Bioline). The solution was mixed by vortexing and stored at -20 °C. The comb was removed from the gel, and then 10µL DNA ladder was loaded. PCR products (10 µl PCR product was mixed with 5 µl loading buffer) were loaded into the gel wells. The electrophoresis tank was connected to a power supply providing 110V and the power was switched off when the gel dye

had migrated approximately halfway down the gel. The gel was visualised under a UV transilluminator (SynGene), and bands were identified and subsequently, purified and sequenced.

2.8. DNA purification

Migrated DNA was purified by the ISOLATE II PCR and Gel Kit (Bioline) by following the manufacturer protocols without modification. PCR products and associated primers were then sent in labelled tubes according to Source Bioscience specifications (http://www.sourcebioscience.com/) for Sanger sequencing.

2.9. Next-generation sequencing (NGS) and Library preparation

2.9.1. NGS library preparation

2.9.1.1. Amplification of the CpG-1 region of TLR9

PCR amplification of the CpG-1 region of TLR9 was carried out using PCR max (Alpha Cycler 1) and primers as previously described (Table 2.2). A PCR reaction was prepared using 25 µl volume containing 12.5 µl High-Fidelity DNA Polymerases & Master Mixes (Thermo Fisher Scientific), 5 µM of each primer and 20-25 ng/µl of DNA extracted from the FTA cards. As a positive control, 1 µl of commercially acquired bovine DNA (Amsbio) was added in place of the DNA extracted from the FTA cards. Reactions were carried out with 10 min at 94 °C, followed by 30 cycles of the 30s at 95 °C, 30s at 60 °C, 30s at 72 °C and a final cycle of 10 min at 72 °C.

2.9.1.2. Amplification of fragment 1 of exon two of the bovine TLR9 gene with an adapter primer

In this PCR amplification, we linked the fragment 1 TLR9 CpG-1 region sequence to an overhanging adapter sequence provided by (Illumina) (Table 2.3). A High-Fidelity DNA Polymerases & Master Mixes (Thermo Fisher Scientific) was used with 5 µM of each primer, and

2.5µl DNA took from previous PCR amplified (2.9.1.1) after cleaning. The PCR cycling conditions were as follows: DNA denaturation at 94 °C for 10 min for one cycle, followed by 25 cycles of DNA denaturation at 94 °C for 1 min; for 1 min with an extension at 72 °C for 1 min. A final extension step at 72 °C for 10 min was carried out.

Table 2. 3 Forward and reverse primer with adapter sequence.

Primer	Sequence
Adapter + F primer	5-’ TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG GCG TCT CCG GGA CAA TA -3’
Adapter + R primer	5’ – GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGG TCC CAG CCA CAG AG - 3’

2.9.1.3. Index

The unique index for each sample was provided by using the Nextera XT Index Kit (Illumina). The PCR reaction was prepared in a total volume of 50 µl using 25 µl volume containing 25 µl High-Fidelity DNA Polymerases & Master Mixes (Thermo Fisher Scientific), 5 µl of each Index primer, 5 µl DNA from the adapter PCR products after bead cleaning. Samples were heated to 95 °C on a thermocycler for 3 mins and then the following 8 cycles were performed: 95 c for 30s, 55 °C for 30s and 72 °C for 30s with a final extension time of 5 min at 72 °C.

2.10. Size Selection

The library was purified between each PCR and prior to pooling all the libraries by using the AMPure XP beads magnetic beads PCR Clean up Kit (Beckman Coulter). The appropriate volume of beads (Bv) was added according to the PCR reaction volume (Pv) using this equation.

$$Bv = Pv*0.7$$

A low concentration of magnetic beads was added to each sample to bind the larger DNA fragments of 637 bp. The beads and PCR products were gently mixed and incubated for 10min at

room temperature before they were placed into a magnetic stand to allow the supernatant to clear. The supernatant was removed, and the beads were washed twice with 70% ethanol before incubation at room temperature for 30s. Then, the supernatant was discarded and the sample allowed to dry for 3-5min at room temperature. A 10 µl aliquot of elution buffer (reagent grade water, TRIS-HCl pH 8.0 or TE buffer) was mixed with each PCR product and incubated for 10min at room temperature. The DNA was then purified with the magnetic stand for 5 minutes before transferring the supernatant to a new collection tube.

2.11. TapeStation

The high sensitivity D1000 ScreenTape assay with the 2200 Tapestation system (Agilent Tech) was used to determine the size of the PCR generated DNA amplicons for each library post bead cleaning. The protocol was followed according to the manufacturer instructions (Agilent Tech).

2.12. Library quantification, normalisation and pooling

The DNA concentration of each library was measured by fluorometry using the Qubit 4 Fluorometer system (ThermoFisher Scientific). The quantification protocol recommended within the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific.) was followed without modification. To convert the DNA concentration from ng/µl (output from the Qubit 4 Fluormeter) to nM the following formula was used (concentration in ng/µl) / (660 g/mol × average library size) × 10^6. Each library was diluted to a final concentration of 4 nM by using Resuspension Buffer (RSB) (10 mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM $MgCl_2$). Finally, 5 µl of diluted DNA was taken from each library and pooled together in a 1.0ml centrifuge tube.

2.13. Quantification PCR (qPCR)

qPCR was used to quantify the DNA concentration of the pooled library accurately. The pooled library was diluted 1:1000, 1:10,000 and 1:20,000 by using buffer EB (10 mM Tris-Cl, pH 8.5) (QIAGEN). Then, the reaction solution was prepared as follow: 10µl from A KAPA SYBR FAST qPCR Master Mix (Sigma-Aldrich Company), 1 µl (5µM) of each primer (Table 2.4), and 4 µl nuclease-free water. Then, 4 µl of diluted library or library quantification DNA standards (Roche) was added to the reaction solution to have a total volume in each qPCR reaction of 20 µl. To increase the accuracy of the assay, duplicates were prepared for each standard reaction and triplicates for each pooled library reaction. The reaction was carried out in the Mic machine (BioMolecular Systems) using the following reaction conditions: hold at 95 °C for 5 min; 40 cycles 30s at 95 °C, 45s at 60 °C.

Table 2. 4 Forward and reverse qPCR primers for quantifying the DNA concentration of the libraries. The forward and reverse primer was annealed to the adapter primer of the library DNA and can therefore be used to quantify the library.

Primer	Sequence
Forward primer	5-’ ATGATACGGCGACCACCGAG -3’
Reverse primer	5’ – CAAGCAGAAGACGGCATACGAG - 3’

2.14. DNA Denaturation

We denatured dsDNA from the pooled libraries was denatured by combining 5 µl from the pooled library with 5 µl of 0.2% freshly diluted sodium hydroxide (NaOH) (SIGMA) in a 1.5ml microcentrifuge tube. As a control, 5 µl of Phix control v3 (Illumina) was combined with 5 µl of 0.2% freshly diluted NaOH in a 1.5ml microcentrifuge tube. Samples were vortexed thoroughly,

spun down and incubated for 5 min at room temperature and then immediately, 550 µl of HT1 hybridisation buffer (Illumina) was added to the microcentrifuge tubes.

2.15. MiSeq Sample loading

To a 1.5 ml microcentrifuge tube was added 188 µl of denatured pooled libraries (stock concentration 20 pM), 20 µl of denatured Phix internal control DNA (stock concentration 20 pM) and 392 µl of HT1 hybridisation buffer (Illumina). Aliquots of 600 µl of the 8-12 pM libraries were loaded to the MiSeq system (Illumina) using the Miseq Reagent Nano Kit v2 (500-cycles) (Illumina).

2.16. SnapGene

SnapGene was used to view and analyse DNA sequences (http://www.snapgene.com/products/snapgene_viewer/). SnapGene was also used to align sequences together, including the reference gene (Accession number KC174788.1) deposited within the NCBI database (https://www.ncbi.nlm.nih.gov/nucleotide/KC174788.1?report=genbank&log$=nucltop&blast_rank=4&RID=4UCCPX3J016.

2.17. EMBOSS CpGplot

The EMBOSS CpGplot is a free tool (https://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/) provided by the EBI and it was used to identify CpG islands in bovine TLR9.

2.18. Statistical analysis

Statistical analysis was carried out using IBM SPSS Statistics for Mac (version 20) and Microsoft Excel. The Fisher's exact test on GraphPad software (http://graphpad.com/) was used to determine and compare the prevalence of parasitic infection in the 70 Nigerian cattle. SPSS was also used to measure the frequency distribution of parasitic infection in badgers. In addition, SPSS was used to investigate the frequency distribution of single nucleotide polymorphisms (SNPs), and also to measure the correlation between infected samples and genetic variation using the Chi-Square test. Finally, SPSS was also used to measure the degree of agreement (Kappa test) between Sanger sequencing and NGS.

CHAPTER THREE

Establishment of a DNA-Extraction Method from FTA-Cards for PCR Genotyping and Sequencing the Bovine TLR9 Gene

3.1. Introduction

Toll-like receptors (TLR)s are pattern recognition receptors (PRRs), which have been described as the first line of defence in an immune response. TLRs play an essential role in activating the innate immune response to a wide variety of pathogen-associated molecular patterns (PAMPs), such as those of lipopolysaccharides (LPSs) and CpG-DNA (Kashyap et al., 2016). TLR9 was identified as a receptor for unmethylated CpG as well as bacterial DNA (Cornélie et al., 2004). Unmethylated CpG motifs are present in the DNA of microorganisms such as viruses, fungi and other invading pathogens (Sarafidou et al., 2013). The TLR9 signalling pathway can culminate in two outcomes: (1) activation of interferon regulatory factors (IRFs) in IRF-signalling endosomes, which leads to the production of type I interferons, and (2) activation of nuclear factor kappa B (NF-κB) in NF-κB signalling endosomes, which induces expression of pro-inflammatory cytokines (Sasai et al., 2010). Flinders Technology Associates (FTA) cards are a simple, economical and sensitive method of collecting biological samples from different kinds of organisms to fix and store DNA directly from the tissue while preventing bacteria growth. In this study, we used the FTA card to collect the blood samples from bovine sources in Nigeria as described previously (Ideozu, 2015). FTA cards have proven to be very successful for DNA preservation and extraction and with different methods (McClure et al., 2009) (Serra et al. 2018). This chapter aims to develop an efficient method for recovery and amplification of DNA from FTA cards that can be used for techniques such as genotyping and Next Generation Sequencing (NGS). Specifically, this could be used for the detection of pathogens and genotyping host immune genes.

The primary purpose of this chapter is to develop efficient methods for extracting DNA from FTA cards whilst simultaneously evaluating the quantity and quality of DNA extracted from FTA cards.

The second purpose was to genotype TLR9 using PCR. To achieve these goals, the following objectives were set out:

- Optimisation of the best methods for extracting DNA from FTA cards, including whole genome amplification (WGA) to enrich DNA.
- Use of a NanoDrop system to evaluate the quality and quantity of extracted DNA.
- Perform PCR amplification of a mammalian tubulin gene to evaluate the extraction of host DNA.
- Design of primers to amplify either the whole bovine TLR9 gene or specific exons.
- To develop PCR amplification protocols to amplify the bovine TLR9 gene from FTA cards.
- To confirm the purity DNA and identity of the bovine TLR9 gene using Sanger sequencing.

3.2. Methods

The bovine blood samples were collected on FTA card as previously described (Ideozu, 2015). To optimise the methods for DNA extraction from FTA card, we extracted DNA from two to five punches using three methods: 5% Chelex resin, the REPLI-g Mini or the WGA Kit (Qiagen) and the DNeasy Blood & Tissue Kit (Qiagen). The DNA concentration was measured using a NanoDrop Spectrophotometer (ThermoFisher Scientific) and stored at -20 °C for downstream application. Then, we used the internal control or bovine tubulin gene to test the quality of extracted DNA. After confirming DNA quality, the PCR technique was optimised to amplify the bovine TLR9 gene using different methods (details in chapter 2). Ultimately, we established the amplified bovine TLR9 gene with Sanger DNA sequencing and alignment of data to the reference bovine TLR9 gene (accession no. KC174788.1) (NCBI database (http://www.ncbi.nlm.nih.gov/).

3.3. Results

3.3.1. DNA extraction from FTA cards and qualitative and quantitative assessment of DNA yield

Whatman FTA Elute cards are impregnated with protein denaturants, which cause lysis of cells, to fix and store DNA directly from the tissue while preventing the growth of pathogens. The goal of this process was to develop a successful method for extracting DNA from FTA cards. We used two to five discs with diameters of approximately 2 mm punched from FTA cards. Different DNA extraction protocols were evaluated using a few different samples applied to FTA cards. We extracted DNA from FTA cards by using 5% Chelex resin, the WGA Kit and the DNeasy Blood & Tissue Kit. However, the DNA yields estimated by the NanoDrop Spectrophotometer varied considerably between samples and extraction methods. The results from the samples were averaged (n=5) based on the number of discs and the method of extraction (Table 3.1). There were significant differences between the extraction methods and DNA yields. For example, 5% Chelex resin extracted a very low concentration of DNA, which indicated that not all DNA bound to the FTA cards was released onto all FTA discs.

In contrast, the WGA Kit extracted a very high concentration of nucleic acid because the kit was designed both to extract and amplify (enrich) the DNA found on FTA discs. The WGA Kit mixture contained buffer, DNA polymerase and unincorporated dNTPs, which might be sources of contamination. Therefore, we purified the mixture using spin column chromatography to ensure that only DNA was eluted. A significant difference in the DNA concentration extracted using the WGA Kit was observed when we increased the number of discs. However, little difference was observed when we added more discs using the DNeasy Blood & Tissue Kit as an extraction method.

Table 3. 1 Average DNA concentration by the number of discs and extraction method (n=5).

Extraction Methods	ng/µl			
	2 Discs (260/280)	3 Discs (260/280)	4 Discs (260/280)	5 Discs (260/280)
5% Chelex resin	7.4 ±2.6 (1.4)	6.7 ±0.9 (1.4)	7.3 ±0.8 (1.6)	9.00 ±1.9 (1.5)
WGA Kit	38.2 ±5.9 (0.7)	63.9 ±5.2 (0.9)	43.8 ±4.1 (0.7)	44.6 ±4.9 (0.9)
DNeasy Blood & Tissue Kit	18.5 ±2.7 (1.6)	17.68 ±3.4 (1.8)	23.9 ±1.7 (1.7)	23.4 ±2.3 (1.8)

The results demonstrated no association between DNA concentration and the number of FTA discs for the Chelex and DNeasy Blood & Tissue Kit extraction methods, and the average DNA concentrations obtained from the WGA Kit demonstrated little association with the number of FTA discs. The DNeasy Blood & Tissue Kit was associated with lower DNA yield than the WGA Kit. However, the average 260/280 absorbance ratio of DNeasy Blood & Tissue Kit was closer to 1.7, which indicates a higher quality DNA than the WGA Kit (>1.00). Overall, this experiment indicates that the DNeasy Blood & Tissue Kit method produces higher quality and purity of DNA extracted from FTA cards compared to the WGA kit and Chelex extraction approaches.

3.3.2. Use of PCR amplification of mammalian tubulin genes

Housekeeping genes are expressed constitutively by different cell types and can be used to check the quality of extracted DNA. For this purpose, we used the mammalian tubulin gene, which encodes part of the cytoskeleton of eukaryotic cells that participate in various cellular processes. We conducted a PCR using mammalian tubulin as the internal control gene to determine the

quality and purity of DNA after extraction from FTA cards. The generic tubulin primers were designed by previous study (Terry et al., 2001). The forward primer sequence was 5'-CGT GAG TGC ATC TCC ATC CAT-3', and the reverse primer sequence was 5'-GCC CTC ACC CAC ATA CCA GTG -3'. using these primers to assess extracted DNA from FTA cards using Chelex resin, WGA Kit and DNeasy Blood & Tissue Kit, they were used with commercial bovine genomic DNA (AMSBIO) as a positive control and water as a negative control. The PCR reaction was carried out with 0.5 pmole/µl of each primer and 20–25 ng/µl of DNA. The equipment and detailed protocol used for PCR are presented in Chapter 2.2.1.

After the PCR was completed, the PCR products were mixed with dye and then loaded in a 1.0 % agarose gel and gel was performed for approximately 45 min to identify the expected band size (i.e. 1196 bp). The result was visualised under UV light, as shown in Fig. 3.1A. A clear band corresponding to the tubulin gene appeared in lane 2 on the gel, indicating that that the PCR amplification was successful. Moreover, there was no band in lane three because the water was used as a negative control.

We tested the quality of DNA extracted from FTA cards after confirming that the tubulin PCR amplification worked with bovine control DNA. We started with DNA eluted from FTA cards using Chelex resin because it was the first method used to extract the DNA, as described in Chapter 2.1.1. As shown in Fig. 3.1B, lanes 2–8 on the agarose gel did not produce bands, indicating that the DNA yield resulting from the Chelex resin extraction method was insufficient for amplification by PCR. We used positive and negative controls in lanes 9 and 10, respectively, to ensure that the results can be attributed to the independent variable.

Because amplification of the DNA extracted using the Chelex resin method was unsuccessful, we tried DNA extracted from FTA cards using the WGA Kit following the procedure detailed in Chapter 2.4. This PCR was carried out using 25–30 ng/µl DNA extracted using the WGA Kit with

0.5 pmole/µl tubulin primers. This resulted in obvious bands in lanes 2 and 3 of the gel as shown in Fig. 3.1C, indicating that the WGA Kit extraction method could extract DNA with low purity and a fair quality to detect the tubulin gene.

The DNA quality test was also performed on DNA eluted using the DNeasy Blood & Tissue Kit with FTA cards; again, the PCR was carried out using tubulin primers. Amplification was successful when we used DNA extracted with the DNeasy Blood & Tissue Kit. As shown in Fig. 3.1D, clear bands were produced after gel electrophoresis in lanes 3–7, and there were no bands in lanes 2 and 8 because of the low DNA yield from these samples. To sum up, we showed that the DNeasy Blood & Tissue Kit is an efficient method to extract the DNA from the FTA card.

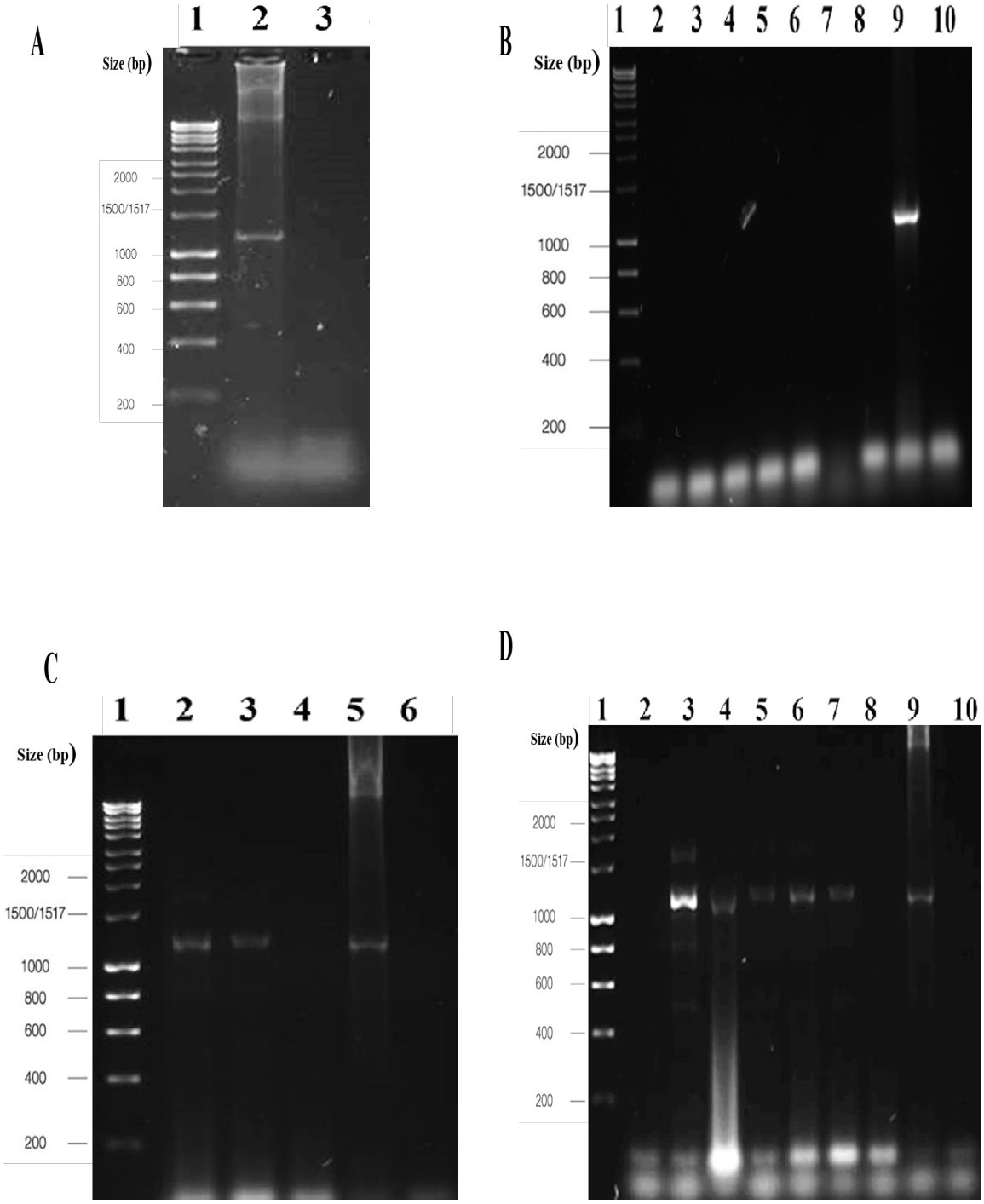


Figure 3.1 Assessment of DNA quality using PCR amplification of template DNA using generic mammalian tubulin primers. (**A**) Gel picture of successful tubulin gene amplification. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: clear band obtained using positive control (DNA). Lane 3: negative control (water). (**B**) Gel picture of unsuccessful tubulin gene amplification. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2–8: no bands seen for DNA extracted using Chelex resin, which indicates unsuccessful amplification. Lane 9: positive control

(commercial DNA). Lane 10: negative control (water). (**C**) Gel picture of successful tubulin gene amplification. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2 and 3: clear bands obtained using DNA extracted by the WGA Kit. Lane 4: no bands, which indicates inefficient DNA extraction. Lane 5: positive control (commercial DNA). Lane 6: negative control (water). (**D**) Gel picture of successful tubulin gene amplification. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 3–7: clear bands obtained using DNA extracted by the DNeasy Blood & Tissue Kit. Lanes 2 and 8: no bands, which indicates inefficient DNA extraction. Lane 9: positive control (commercial DNA). Lane 10: negative control (water).

3.3.3. Bovine TLR9 PCR amplification

3.3.3.1. Amplification of the bovine TLR9 gene

The primary purpose of this section was to amplify the whole bovine TLR9 gene (4156bp). The whTLR9 primers were designed to cover most of the bovine TLR9 gene. Fig. 7.1A (appendices) shows a schematic diagram of the selected primers after alignment with the program clustal omega. Initially, we tried to test TLR9 primers in bovine commercial DNA purchased from AMSBIO (http://www.amsbio.com) as a positive control and then tested the DNA extracted from FTA cards. Unfortunately, although gel electrophoresis was carried out many times after each PCR reaction to identify the expected band size, no bands appeared under UV light.

This failure led us to optimise certain aspects. For example, we used different concentrations of primers and changed the PCR conditions to achieve more specific bands. Specifically, we carried out PCR using various concentrations of whTLR9 primers (0.5 pmole/µl, 0.2 pmole/µl, 0.4 pmole/µl and 0.6 pmole/µl), we tried to increase and decrease the primer concentrations from the first PCR reaction with bovine commercial DNA. However, the different primer concentrations could not amplify regions of the bovine TLR9 gene, even after several attempts. As a result, there are no bands shown in Fig. 7.1B (appendices).

As a part of optimization, the experiment was repeated with various temperature gradients (55 ºC, 58 ºC, 60 ºC and 61 ºC) to determine the specific primer annealing temperature. As shown in Fig.

7.1C, there were still no visible bands after gel electrophoresis, indicating that the various temperature gradients carried out could not solve the problem. All reagents used for amplification of TLR9 (see Chapter 2.2.2) were replaced to troubleshoot the experiment. Also, we used the MyTaq™ Hs PCR reaction mixture (Bioline) instead of the High-Fidelity DNA Polymerases & Master Mixes (Thermo Fisher Scientific) to improve the efficiency of amplification for large templates and enable high annealing temperatures. However, these optimisation efforts did not affect the result; amplification of TLR9 was still unsuccessful. To sum up, we showed that genotyping of the whole TLR9 gene (4156bp) was not achieved in our study.

3.3.3.2. Development of aTLR9 hemi-nested PCR

Following unsuccessful amplification of the whole TLR9 gene, we employed the hemi-nested PCR technique to increase the specificity of TLR9 gene amplification. Hemi-nested PCR includes two PCR reactions. In the first, we used the same primers as we used for TLR9 amplification (outside primers). In the second, we used the products of the first PCR reaction as a template and designed a new forward primer (inside primer) with a reverse outside primer. These primers amplified the targeted region of the bovine TLR9 gene that covered CpG islands on exon 2. According to previous studies (Cargill and Womack, 2007, Seabury et al., 2007), bovine TLR9 has two CpG islands in the TLR9 genomic sequence, as determined by CpGPlot (http://www.ebi.ac.uk/emboss/cpgplot/), but none have been detected in TLR1, TLR 3, TLR5, TLR7, TLR8 or TLR10. The first CpG island detected in bovine TLR9 is located from 3095–3353 bp, while the second CpG island is located from 3556–3976 bp. We designed a new forward primer (CpG_F) to cover the CpG island regions Fig. 7.2A (appendices). The forward primer, (CpG_F) 5'–CTGCGTCTCCGGGACAATAAC-3', started at 3601 bp. We used the same reverse primer as in previous experiments (whTLR9_R) 5'-TGT GGG GTT AAA GGA GTG CTG -3', which started at 4691 bp.

The expected band size for the bovine TLR9 in the hemi-nested PCR experiment was 1161 bp. Bovine commercial DNA was used as a positive control. Our study showed no specific bands were identified with control bovine DNA. The experiment was repeated several times with different optimisation procedures, as shown in Fig. 7.2B (appendices). No visible bands appeared, even when using different mixes (MyTaq™ Hs mix). In addition, there were no bands identified with various temperature gradients in the second hemi-nested PCR reaction Fig. 7.2C (appendices). Overall, the hemi-nested PCR was not working in our study.

3.3.3.3. PCR of CpG island 1 and 2 of bovine TLR9

Since there was not even a single band when we performed amplification of the whole TLR9 gene and hemi-nested PCR amplification, we suspected that the primers were nonspecific. Thus, we conducted amplification again to detect the TLR9 coding region by targeting the CpG islands located on exon 2 of the bovine TLR9. For this amplification, we designed new short primers to amplify the target sequences (i.e. CpG islands), as shown in Fig. 3.2A. The expected band size was 1129 bp. Before testing whether DNA eluted from FTA cards could be amplified using these new primers, we tested the new primers on bovine control DNA (commercial DNA). As shown in Fig. 3.2B, a visible band of bovine commercial DNA appeared in the second lane, but the band size was smaller than expected (1129 bp). Therefore, we diluted 1 µl of the control DNA in nuclease-free water at different concentrations (1:100, 1:50 and 1:10). Also, we used various concentrations (0.2 pmole/µl, 0.4 pmole/µl and 0.6 pmole/µl) of the CpG primers. The result still was visible bands but smaller than the expected band size.

To confirm the identity of the bovine control DNA PCR product as the bovine TLR9 gene, the isolated band was sent to a sequencing company (Source Bioscience) for Sanger sequencing. The DNA sequencing data were received electronically and analysed using SnapGene. Fig. 3.3A shows clear sequence data and peaks for the bovine control DNA sample.

Then, we searched for highly identical sequence matches using the National Centre for Biotechnology (NCBI) database (http://www.ncbi.nlm.nih.gov/). This produced BLAST hits that confirmed alignment with the bovine (*B. indicus*) TLR9 gene, as shown in Fig. 3.3B. Thus, amplification of the target sequence using CpG primers seemed to be successful.

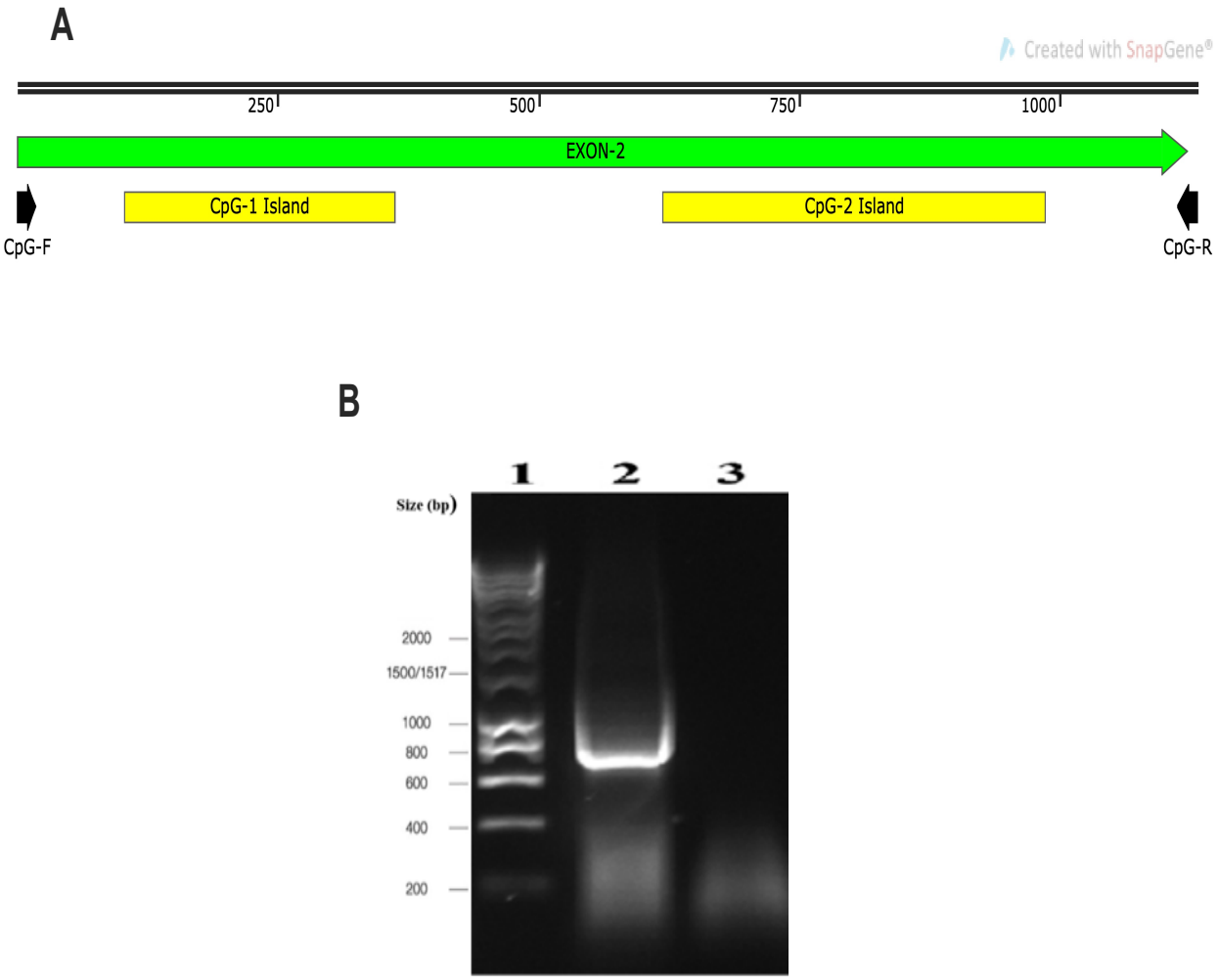
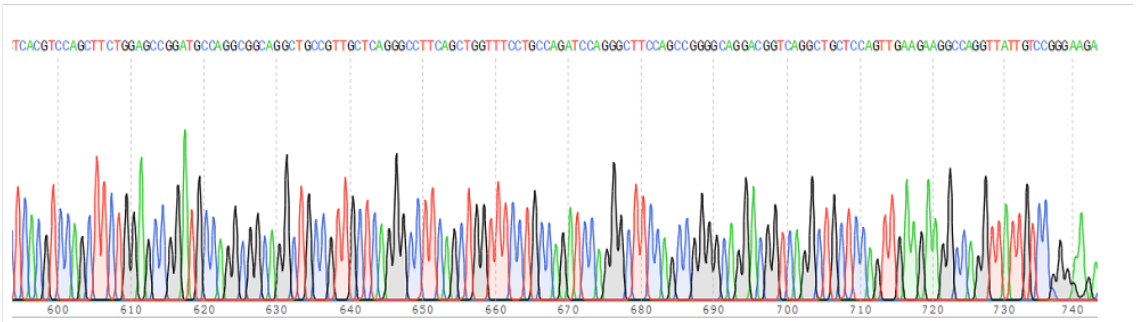


Figure 3. 2 Amplification of CpG islands in TLR9 using control DNA. (**A**) Designed primers used to target CpG islands. The black line represents the amplified bovine TLR9 gene. The green arrow represents exon 2 of bovine TLR9. The yellow labels indicate the locations of CpG islands 1 and 2 in the bovine TLR9 gene. The black arrows indicate the locations of the designed forward and reverse CpG island primers. (**B**) Gel picture of amplified CpG islands. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: clear band, which indicates successful amplification of positive control bovine DNA with CpG island two primers. Lane 3: negative control (water).

A



B

Bos indicus isolate 9 toll-like receptor 9 (TLR9) gene, complete cds

Sequence ID: KC174788.1 **Length:** 4831 **Number of Matches:** 1

Range 1: 4174 to 4741 GenBank Graphics ▼ Next Match ▲

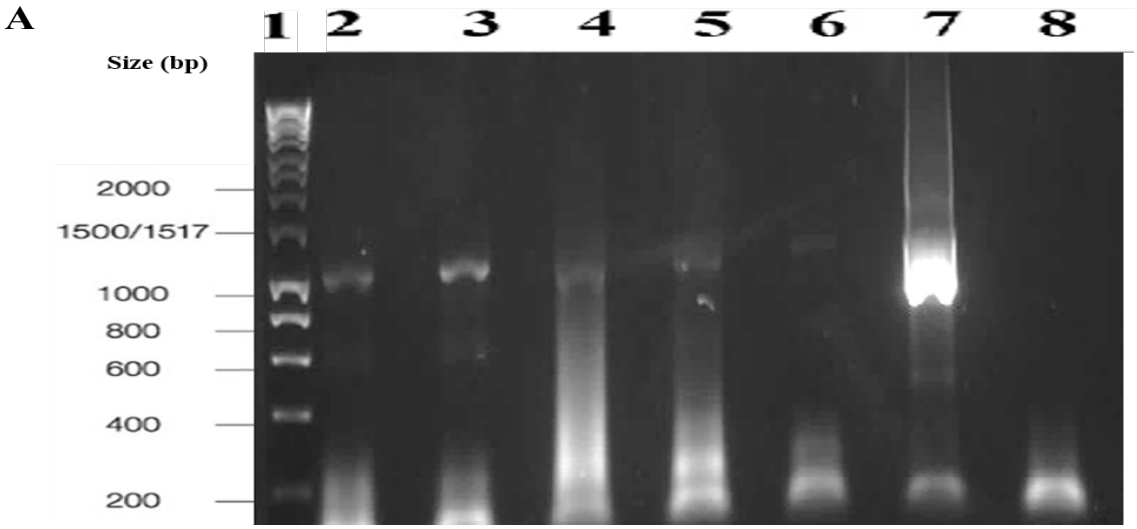
Score	Expect	Identities	Gaps	Strand
1027 bits(556)	0.0	563/568(99%)	1/568(0%)	Plus/Plus

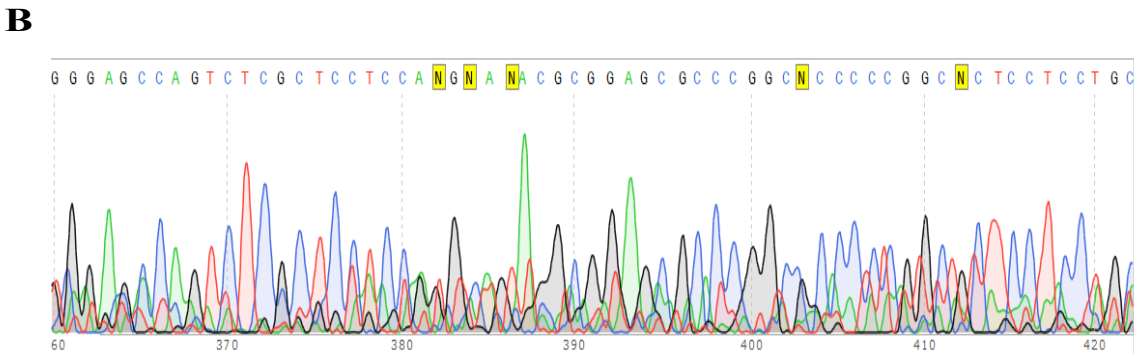
```
Query  1     TTG-CCCGACGGCGGCGGCAGCGGGGCGAGGACACCCTGCTCTATGATGCCTTCGTGGTC  59
             ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4174  TTGCCCCGACGGCGGCGGCAGCGGGGCGAGGACACCCTGCTCTATGATGCCTTCGTGGTC  4233

Query  60    TTCGACAAGGTGCANAGTGCAGTGNNTGATTGGGTGTACAACGAGCTCCGCGTGCAGCTG  119
             ||||||||||||||| |||||||||  ||||||||||||||||||||||||||||||||
Sbjct  4234  TTCGACAAGGTGCAGAGTGCAGTGGCTGATTGGGTGTACAACGAGCTCCGCGTGCAGCTG  4293

Query  120   GAGGAGCGCCGGGGGCGCCGGGCGCTCCGCCTCTGCCTGGAGGAGCGAGACTGGCTCCCT  179
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4294  GAGGAGCGCCGGGGGCGCCGGGCGCTCCGCCTCTGCCTGGAGGAGCGAGACTGGCTCCCT  4353

Query  180   GGTAAGACGCTCTTCGAGAACCTGTGGGCCTCGGTCTACAGCAGCCGCAAGACCATGTTC  239
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4354  GGTAAGACGCTCTTCGAGAACCTGTGGGCCTCGGTCTACAGCAGCCGCAAGACCATGTTC  4413

Query  240   GTGCTGGACCACACGGACCGGGTCAGCGGNCTCCTGCGCGCCAGCTTCCTGCTGGCCCAG  299
             |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  4414  GTGCTGGACCACACGGACCGGGTCAGCGGCCTCCTGCGCGCCAGCTTCCTGCTGGCCCAG  4473

Query  300   CAGCGCCTGTTGGAGGACCGCAAGGACGTCGTAGTGCTGGTGATCCTGCGCCCCGCCGCC  359
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4474  CAGCGCCTGTTGGAGGACCGCAAGGACGTCGTAGTGCTGGTGATCCTGCGCCCCGCCGCC  4533

Query  360   TATCGGTCCCGCTACGTGCGGCTGCGCCAGCGCCTCTGCCGCCAGAGCGTCCTCCTCTGG  419
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4534  TATCGGTCCCGCTACGTGCGGCTGCGCCAGCGCCTCTGCCGCCAGAGCGTCCTCCTCTGG  4593

Query  420   CCCCACCAGCCCAGTGGCCAGGGTAGTTTCTGGGCCAACCTGGGCATAGCCCTGACCAGG  479
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4594  CCCCACCAGCCCAGTGGCCAGGGTAGTTTCTGGGCCAACCTGGGCATAGCCCTGACCAGG  4653

Query  480   GACAACCGTCACTTCTATAACCGGAACTTCTGCCGGGGCCCCACGACAGCCGAATAGCAC  539
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4654  GACAACCGTCACTTCTATAACCGGAACTTCTGCCGGGGCCCCACGACAGCCGAATAGCAC  4713

Query  540   AGAGTGACTGCCCAGCACTCCTTTAACC  567
             ||||||||||||||||||||||||||||
Sbjct  4714  AGAGTGACTGCCCAGCACTCCTTTAACC  4741
```

Figure 3. 3 Sanger sequencing data for the PCR amplification of bovine control DNA with the primers design to target the CpG island. (**A**) Screenshot of the chromatogram showing the nucleotide peaks (green for adenine, blue for cytosine, black for guanine and red for thymine), which indicate good DNA quality. (**B**) Screenshot of the control DNA sequence aligned to exon 2 in bovine TLR9. The highest similarity (99%) was observed with *B. indicus* TLR9 using the program BLAST.

Following successful amplification of the control bovine DNA with CpG primers, we used these primers for PCR amplification of some of the bovine samples collected on Whatman FTA cards following DNA elution using the WGA kit. The amplification of DNA extracted from FTA cards produced faint bands in all sample wells, and a clear band appeared in the bovine positive control, implying that PCR of the CpG island coding region was successful. The experiment was repeated several times with different optimisation procedures, such as the increased concentration of the DNA eluted from FTA cards (25 to 40 ng/µl) and increased concentration of primers (0.4 to 0.6 pmole/µl). However, the bands did not change; they were still faint at 1129 bp. Therefore, we changed some of the PCR conditions, increasing the annealing temperature from 58 ºC to 62 ºC and extending the number of cycles from 25 to 35. The gel image in Fig. 3.4A shows that very clear bands were produced after gel electrophoresis due to amplification of CpG islands in the TLR9 coding region using DNA extracted from FTA cards with the WGA Kit.

Following the development of the PCR technique, we sent the PCR products after purification (Isolate II PCR and Gel kit), with associated primers, for Sanger DNA sequencing (Source Bioscience). As shown in Fig. 3.4B and Fig. 3.4C, the DNA sequencing data yielded some overlap peaks, and similarity to bovine TLR9 was confirmed using the NCBI database. This indicated that even though the DNA eluted from FTA cards using the WGA Kit did not yield very clear sequence data, amplification and production of a sequence that was identical to that of bovine TLR9 were successful.





C

Bos indicus TLR9 (TLR9) mRNA, complete cds

Sequence ID: KX138608.1 **Length:** 3158 **Number of Matches:** 1

Range 1: 2723 to 3086 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
403 bits(218)	3e-108	310/365(85%)	4/365(1%)	Plus/Minus

```
Query  1     CCTGGATCAGGGCTNNNNCNAGGTTGGCCC-TTANCTACCCTGGCCTCTGGGCTGGTGGG  59
             ||||| ||||||||     | ||||||||||    | ||||||||||| ||||||||||||||
Sbjct  3086  CCTGG-TCAGGGCTATGCCCAGGTTGGCCCAGAAACTACCCTGGCCACTGGGCTGGTGGG  3028

Query  60    GCCAGAGGAAGACGTTCTGGNGGGA-AGGAATTGGTTTTGCCGCTTTTTACGAGACCGAT  118
             ||||||||| |||| |||||| || | |||    |||     |||||    |  || |||||||
Sbjct  3027  GCCAGAGGAGGACGCTCTGGCGGCAGAGGCGCTGGCGCAGCCGCACGTAGCGGGACCGAT  2968

Query  119   ATGCGGCGGGTCGCAGGATCNCCANCACTACNACGTCCTTGCAGTCTTCCAACAGGCGCT  178
             | |||||||| |||||||||| ||| |||||| |||||||||| ||| ||||||||||||||
Sbjct  2967  AGGCGGCGGGGCGCAGGATCACCAGCACTACGACGTCCTTGCGGTCCTCCAACAGGCGCT  2908

Query  179   GCTGGGCCAGCAAGAAGCTGGNGCGCAGGANGCCGCTGACCCGGTCCGTGTGGACCA-TT  237
             |||||||||||| |||||||| |||||||| |||||||||||||||||||||||| |||
Sbjct  2907  GCTGGGCCAGCAGGAAGCTGGCGCGCAGGAGGCCGCTGACCCGGTCCGTGTGGTCCAGCA  2848

Query  238   TNAACATGGTCTTGCNGCTGCTGTAGACCTANGCCCACAGGTTCTCGAAGAGCGTCTTAC  297
               |||||||||||||| |||||||||||||| | ||||||||||||||||||||||||||
Sbjct  2847  CGAACATGGTCTTGCGGCTGCTGTAGACCGAGGCCCACAGGTTCTCGAAGAGCGTCTTAC  2788

Query  298   CAGGGAGCCAGTCTCGCTCCTCCANGNANACGCGGAGCGCCCGGCNCCCCCGGCNCTCCT  357
             |||||||||||||||||||||||| | | | |||||||||||||||| |||||||| |||||
Sbjct  2787  CAGGGAGCCAGTCTCGCTCCTCCAGGCAGAGGCGGAGCGCCCGGCGCCCCCGGCGCTCCT  2728

Query  358   CCTGC  362
             || ||
Sbjct  2727  CCAGC  2723
```

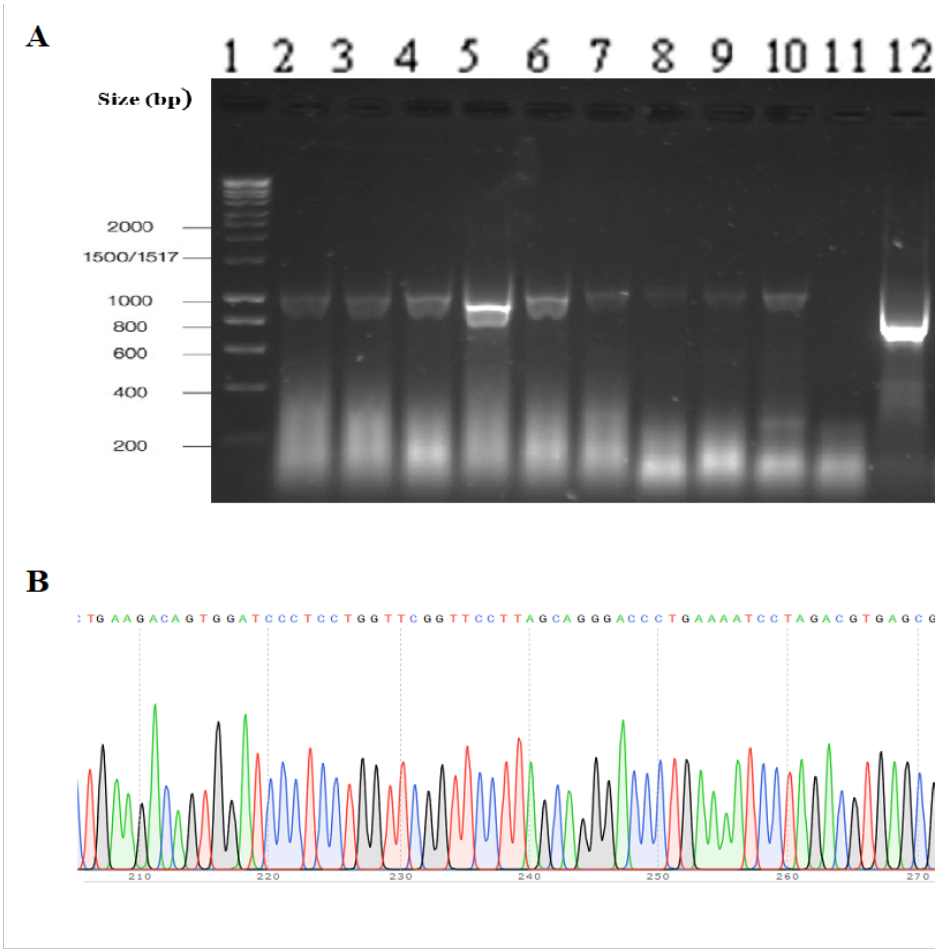
Figure 3. 4 Amplification and Sanger sequencing of CpG islands in the bovine TLR9 gene following extraction of DNA from FTA cards using the WGA kit. (**A**) Gel picture of successfully amplified CpG islands. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2–6: bands that became clear after optimisation. Lane 7: clear band obtained using positive control (bovine DNA). Lane 8: negative control (water). (**B**) Screenshot of the sequencing chromatogram showing

the respective nucleotide peaks (green for adenine, blue for cytosine, black for guanine and red for thymine). Overlapping peaks indicate the low quality of the DNA. (**C**) Screenshot of the alignment of DNA eluted by the WGA Kit to exon 2 in bovine TLR9. The highest similarity (99%) is observed for *B. indicus* TLR9 using the program BLAST.

To produce high-yield DNA data sequence with clean peaks, we tried to amplify the CpG island 2 TLR9 coding regions using DNA extracted from FTA cards with the DNeasy Blood & Tissue Kit. The fact that all nine lanes featured the estimated band size (1129 bp) indicated that the DNeasy Blood & Tissue Kit produced good-quality DNA for amplification of the target sequence using PCR as shown in Fig. 3.5A. Then, the NCBI database (http://www.ncbi.nlm.nih.gov/) was searched to confirm the identity of the PCR product as the bovine TLR9 gene.

To determine whether the DNeasy Blood & Tissue Kit yielded better sequence data than the WGA Kit, we purified PCR products using the ISOLATE II PCR and Gel Kit. We then sent the products and their associated primers to Source Bioscience for Sanger sequencing.

Fig. 3.5B shows clear sequence data was obtained for the bovine DNA samples eluted from FTA cards using the DNeasy Blood & Tissue Kit. The alignment in Fig 3.5C shows the sequence is 99.5% similar to the reference *B. indicus* TLR9 data in GenBank. In conclusion, our data showed that the DNeasy Blood & Tissue Kit is a more efficient method than the WGA Kit to extract DNA from FTA card and produce high quality and purity DNA suitable for PCR amplification.



C

Bos indicus isolate 9 toll-like receptor 9 (TLR9) gene, complete cds

Sequence ID: KC174788.1 Length: 4831 Number of Matches: 1

Range 1: 3877 to 4141 GenBank Graphics ▼ Next Match ▲

Score	Expect	Identities	Gaps	Strand
483 bits(261)	2e-132	263/265(99%)	0/265(0%)	Plus/Plus

```
Query  1     CTAGACGTGAGCGCCAACCCGCTCCACTGCGCCTGCGGGGCGGCCTTTGTGGACTTCCTG  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3877  CTAGACGTGAGCGCCAACCCGCTCCACTGCGCCTGCGGGGCGGCCTTTGTGGACTTCCTG  3936

Query  61    CTGGAGAGACAGGAGGCCGTGCCCGGGCTGTCCAGGCGCGTCACATGTGGCAGTCCGGGC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  3937  CTGGAGAGACAGGAGGCCGTGCCCGGGCTGTCCAGGCGCGTCACATGTGGCAGTCCGGGC  3996

Query  121   CAGCTCCAGGGCCGCAGCATCTTCACACAGGACCTGCGNCTCTGCCTGGATGAGACCCTC  180
             |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  3997  CAGCTCCAGGGCCGCAGCATCTTCACACAGGACCTGCGTCTCTGCCTGGATGAGACCCTC  4056

Query  181   TCCTTGGACTGCTTTGGCCTCTCNCTGCTAATGGTGGCGCTGGGCCTGGCAGTGCCCATG  240
             |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Sbjct  4057  TCCTTGGACTGCTTTGGCCTCTCGCTGCTAATGGTGGCGCTGGGCCTGGCAGTGCCCATG  4116

Query  241   CTGCACCACCTCTGTGGCTGGGACC  265
             |||||||||||||||||||||||||
Sbjct  4117  CTGCACCACCTCTGTGGCTGGGACC  4141
```

Figure 3. 5 Amplification and Sanger sequencing of CpG islands in the bovine TLR9 gene following extraction of DNA from FTA cards using the DNeasy Blood & Tissue Kit (**A**) Gel picture of amplification of CpG islands. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2–10: clear bands indicate CpG islands were successfully amplified in DNA extracted using the DNeasy Blood & Tissue Kit. Lane 11: negative control (water). Lane 12: successful amplification of

positive control (bovine DNA). (**B**) Screenshot of the chromatogram of DNA extracted using the DNeasy Blood & Tissue Kit. The nucleotide peaks (green for adenine, blue for cytosine, black for guanine and red for thymine) indicate good DNA quality. (**C**) Screenshot of the alignment of exon 2 in bovine TLR9 from DNA extracted using DNeasy Blood & Tissue Kit. The highest similarity (99%) is observed for *B. indicus* TLR9 using the program BLAST.

3.3.3.4. Development of PCR amplification of sections of the TLR9 gene within exon 2

To confirm that the PCR technique is working consistently and can provide useful DNA sequence data for the whole exon 2 of TLR9, we designed six primers (see Fig. 3.6A) to amplify small fragments of CpG island 2 in exon 2 of the bovine TLR9 gene. The first fragment (f1TLR9) was located from +3603 to +4156 bp in the bovine TLR9 gene, and its expected band size was 554 bp. The second fragment (f2TLR9) was located from +3842 to +4156 bp, and the expected band size was 316 bp. The third fragment (f3TLR9) was located from +4141 to +4757 bp, and its expected band size was 616 bp. The fourth fragment (f4TLR9) was located from +4439 to +4756 bp, and its expected band size was 318 bp.

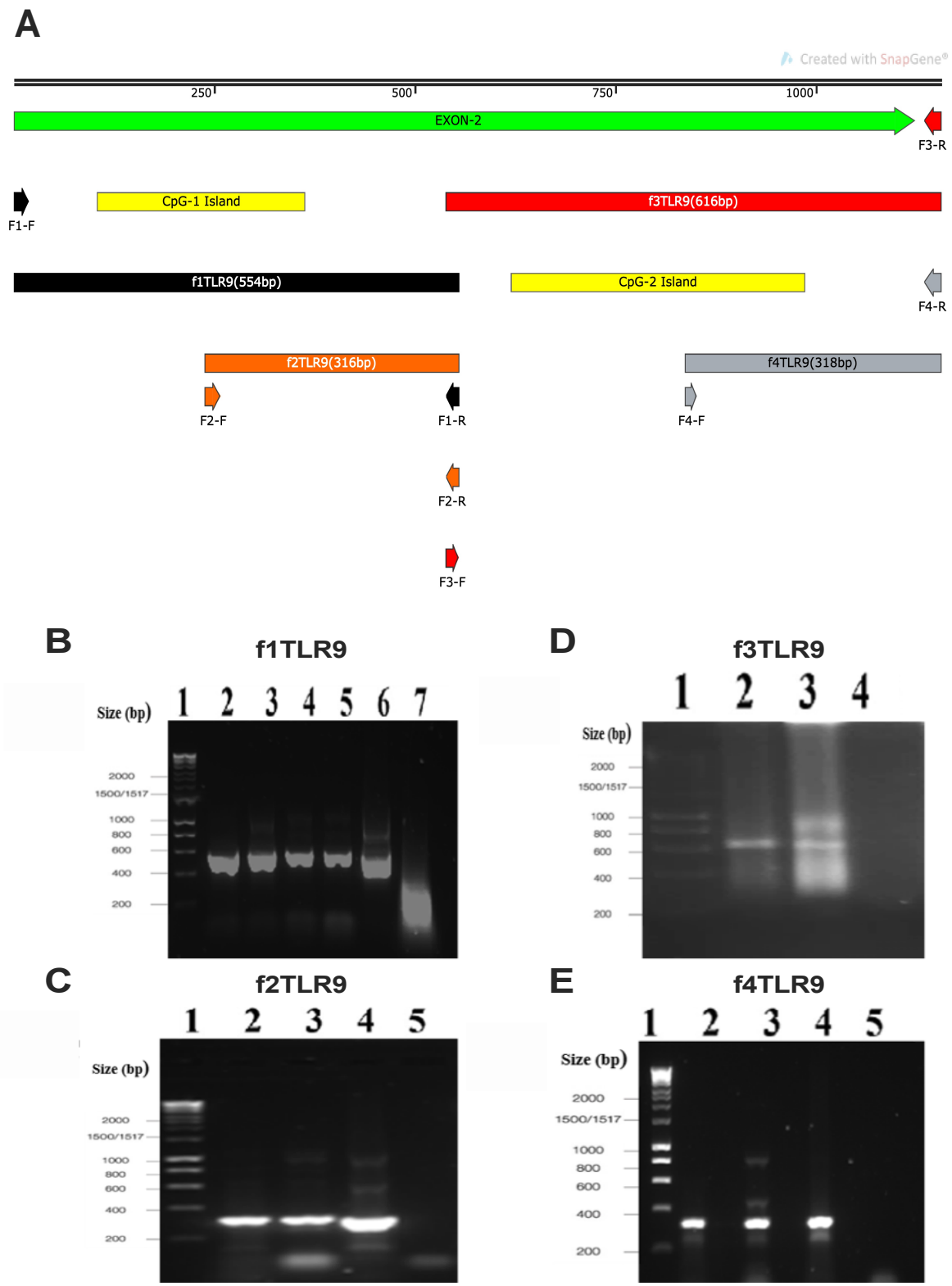


Figure 3. 6 Fragment amplification at exon 2 of the bovine TLR9 gene. (A) The relative locations of the 6 designed PCR primers targeting exon 2 of the bovine TLR9 gene. The green label indicates the part of exon 2 within TLR9. The yellow labels indicate the locations of CpG

islands 1 and 2 in the bovine TLR9 gene. The black label indicates the location of fragment 1, and black arrows indicate the forward and reverse primers. The orange label indicates the location of fragment 2, and orange arrows indicate the forward and reverse primers. The red label indicates the location of fragment 3, and red arrows indicate the forward and reverse primers. The grey label indicates the location of fragment 4, and grey arrows indicate the forward and reverse primers. **(B)** Gel picture of successful amplification of fragment 1 within exon 2 of the bovine TLR9 gene. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2–5: clear bands for tested samples of DNA. Lane 6: positive control (bovine DNA). Lane 7: negative control (water). **(C)** Gel picture of successful amplification of fragment 2 within exon 2 of the bovine TLR9 gene. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2 and 3: clear bands for bovine DNA samples. Lane 4: positive control (bovine DNA). Lane 5: negative control (water). **(D)** Gel picture of successful amplification of fragment 3 within exon 2 of the bovine TLR9 gene. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: clear bands for bovine DNA samples. Lane 3: positive control (bovine DNA). Lane 4: negative control (water). (**E)** Gel picture of successful amplification of fragment 4 within exon 2 of the bovine TLR9 gene. Lane 1: hyper ladder 1 kb marker (Bioline). Lanes 2 and 3: clear bands for bovine DNA samples. Lane 4: positive control (bovine DNA). Lane 5: negative control (water).

In these amplifications, we used a DNA template eluted from FTA cards using the REPLI-g Mini Kit because it demonstrated an ability to amplify and produce DNA sequence data suitable for observing SNPs. Each fragment had different annealing temperatures based on its primers. Small fragments could be amplified at low annealing temperatures (around 53–56 ºC) while large fragments could be amplified at high annealing temperatures (approximately 60–62 ºC). As shown in Fig. 3.6, clear bands were observed for all fragments, indicating that all PCR amplifications using the newly designed primers were successful. Following successful amplification of exon 2 of bovine TLR9 in DNA extracted from bovine blood collected on FTA cards, all PCR products were cleaned using the ISOLATE II PCR and Gel Kit and were sent with their associated primers to Source Bioscience for Sanger sequencing to confirm their identity. All four fragments of exon 2 bovine TLR9 sequence featured clear, evenly spaced nucleotide peaks. Otherwise, the yield of DNA sequencing data covered almost all nucleotides on the exon 2 fragment. In brief, these results indicate that we could investigate the variation in exon 2 of TLR9 between the bovine blood

samples collected on FTA cards. When retrieved, the four fragments of DNA sequencing data were used to search for highly similar sequences in the NCBI database (Table 3.2). We confirmed alignment of all fragments with the reference bovine TLR9 gene.

Table 3. 2 Summary of the sequencing of the TLR9 fragments following NCBI BLAST analysis. The four fragments demonstrated high similarity to the gene from *B. indicus* accession number (KC1747788.1).

Sequencing sample	Max score	Total score	Query cover	E-Value	Indent
f1TLR9	924	924	100%	0.0	99.60%
f2TLR9	483	483	99%	2e-132	99.25%
f3TLR9	970	970	97%	0.0	97.83%
f4TLR9	457	457	93%	1e-124	99.21%

3.4. Discussion

In this study, we extracted DNA and optimised the conditions to obtain high-resolution TLR9 gene sequencing from bovine blood samples applied on Whatman FTA cards. The bovine blood samples were obtained from *B. indicus* bulls in Nigeria by a previous thesis study at Salford University (Ideozu, 2015). The samples were collected and stored on FTA cards at room temperature.

Two to five punches with approximate diameters of 2 mm were obtained from the same sample from each FTA card. These punches were used to determine the quantity and quality of the eluted DNA (based on the number of punches). Use of FTA cards in PCR has been proven to detect and diagnose protozoa infection in many recent studies (Ahmed et al., 2011, Ahmed et al., 2013, Cox

et al., 2005, Cox et al., 2010, Hashimoto et al., 2019, Hide et al., 2003). We found no significant difference in DNA concentration when the number of FTA punches was increased. A previous study showed minimal variation between DNA yields from one versus three punches from the same individual, similar to our findings. Still, unlike us, they used a manual FTA elution method (Parsons and Bright, 2012).

Three different methods were used to extract DNA from the FTA cards: 5% Chelex resin, a WGA Kit and a DNeasy Blood & Tissue Kit. The aim of these methods was to determine the best extraction method for eluting DNA with high yield and quality to perform PCRs in order to generate sequence data for the bovine TLR9 gene. The yield of extracted nucleic acids differed significantly differed between the three extraction methods. Although the Chelex-100 resin is the method of choice for protocols requiring rapid extraction of DNA from trace amounts of biological samples (Ardura et al., 2010, Casquet et al., 2012, Walsh et al., 1991), this method yielded the lowest DNA concentrations in our study.

The REPLI-g Mini Kit, which is considered a multiple displacement amplification (MDA) application, eluted the highest DNA concentration compared to the other methods. Several studies have demonstrated the efficacy of MDA for analysing genomic DNA preparations from a limited number of bacterial cells (Kvist et al., 2007, Marcy et al., 2007). Although we could amplify the DNA extracted from the cards using the REPLI-g Mini Kit, the 260/280 absorbance ratio, which assesses DNA purity, was low and poor, with overlapping reads obtained from Sanger sequencing. Other studies showed the same results as our study: WGA can amplify DNA, but it produces poor sequencing reads and overlapping with a non-specific template (Ahsanuddin et al., 2017, Guo et al., 2015). Moreover, a recent study applied WGA as a way to extract DNA incorporated with tetramethylammonium chloride in order to improved subsequent amplification and sequencing coverage (Oyola et al., 2014).

The DNeasy Blood & Tissue Kit allowed high quality DNA to be extracted from FTA cards, as demonstrated in some previous studies (Moura et al., 2016, Ward et al., 2019). Indeed, the DNeasy Blood & Tissue Kit produced higher-purity DNA than the REPLI-g Mini Kit because the DNeasy Blood & Tissue Kit contains lysis buffer, which breaks biological cells to obtain molecular material. Moreover, it has a washing step to remove remaining contaminants and enzyme inhibitors. Also, a spin column is used to adsorb DNA onto a silica membrane to permit multiple washes prior to elution. This is consistent with other reports that have shown that the quantity of DNA extracted with Chelex might be high but the DNeasy Blood and tissue kit produced a higher purity of DNA (Silva et al., 2014).

To estimate the quality of DNA extracted from FTA cards, we tested the eluted DNA from all three methods by PCR amplification of the bovine tubulin gene as a representative housekeeping gene. The DNA eluted using 5% Chelex resin was not amplified with tubulin primers due to insufficient DNA concentration and low efficiency of removing PCR inhibitors (Hu et al., 2015, Miller et al., 1999). The REPLI-g Mini Kit and DNeasy Blood & Tissue Kit methods successfully extracted and amplified DNA with tubulin primers.

This study aimed to use the DNA extracted from blood spotted on FTA cards for identification and sequencing of the bovine TLR9 gene and detection of genomic variation. For this purpose, we amplified the bovine control DNA sample by targeting the whole gene (4761bp). The primers could not amplify the whole bovine TLR9 gene after attempts; most likely due to the large DNA template, which, as shown in a previous study, makes amplification difficult (Cha and Thilly, 1993). After that, we tried to target exon 2, which contains a CpG island, within the TLR9 gene using hemi-nested PCR. The CpG island within exon 2 of TLR9 was a short DNA sequence in which the frequency of the CG dinucleotide was higher than in another region of the bovine TLR9

gene. This region was chosen because bovine TLR9 has been reported to have two CpG islands, while other TLRs have no CpG islands (Cargill and Womack, 2007, Seabury et al., 2007). Moreover, CpG islands are hotspots for pathological mutations that cause genetic diseases. The plan was to amplify the whole TLR9 gene in the first round and then use other primers to target exon 2 within the TLR9 gene in the second round. We repeated the hemi-nested PCR several times with different optimisation measures, such as different annealing temperatures, new mixes and different stocks of the same primers were used. However, all these measures failed, even though previous studies have described and showed that hemi-nested PCR is a highly sensitive and specific method to amplify genes (Bachman et al., 2004, Pujol-Riqué et al., 1999, Singh et al., 2008).

We designed a new set of primers to improve specificity of amplification of exon 2 of TLR9 gene. The data demonstrate that we successfully amplified exon 2 of the TLR9 gene and Sanger sequencing showed that it had 99% similarity to bovine TLR9 (*B. indicus*, accession number (KC1747788.1).

Following successful genotyping of exon 2, the PCR was tested on bovine samples applied to FTA cards and extracted from blood using the REPLI-g Mini Kit; an approach proven to be effective by many researchers (Jasmine et al., 2008, Rykalina et al., 2014). Our data showed clear amplification of test samples and the sequence data aligned with bovine TLR9; however, there were overlapping peaks indicative of poor sequencing data. DNA extracted using the DNeasy Blood & Tissue kit was purer compared to the DNA generated by the REPLI-g Mini Kit. Highly efficient extraction of DNA using the DNeasy Blood & Tissue Kit has also been demonstrated previously (Abekhti et al., 2017, Simonelli et al., 2009).

To prepare the DNA for further study, such as investigation of genetic variation in the bovine TLR9 gene, we fragmented the exon 2 of the bovine TLR9 gene into four regions. All fragments

were amplified with PCR and produced clear bands corresponding to the expected fragment sizes. Additionally, the DNA sequence data for each fragment showed good quality with fewer overlapping peaks.

In conclusion, this study demonstrates an efficient method for extracting DNA from bovine blood spotted on FTA cards. Moreover, it establishes a set of conditions to PCR amplify exon 2, which contains the CpG islands within the TLR9 gene from the FTA cards. These methods could be used to study the genetic variation within CpG islands using different sequencing platforms.

CHAPTER FOUR

An Investigation into the Genetic Variation of the Bovine TLR9 Gene, Using Sanger Sequencing, in African Cattle with Parasitic Infections

4.1. Introduction

Genetic variation within mammalian TLRs may be associated with susceptibility or resistance to a parasitic infection such as *Trypanosoma* infection (Weitzel et al., 2012). In a previous study, a total of 70 bovine blood samples were obtained from bulls in Nigeria and stored on FTA cards (Ideozu, 2015; West, 2018). Multiple punches were obtained from each FTA card to extract parasitic DNA and identify the parasite using a next-generation sequencing (NGS) strategy (West, 2018). Also, the previous lab workers detected *Trypanosoma* species infection using an ITS-nested PCR method (Ideozu, 2015) which depended upon the generation of visible species-specific PCR products detected by agarose gel electrophoresis (Cox et al., 2005). In this study, we extracted the DNA from 70 blood samples using the DNeasy Blood & Tissue kit (Bioline). The quality and concentration of DNA were measured by NanoDrop Spectrophotometry (Table 7.1, appendices). Then, we amplified the bovine f1TLR9 and sent all the purified DNA samples for Sanger sequencing to determine the TLR9 polymorphism. The rationale for selecting the f1TLR9 region was that it contains one of the CpG islands and leucine-rich repeats (Cargill and Womack, 2007). TLR9 polymorphism has been demonstrated to lead to predisposition to inflammatory and autoimmune disorders, such as asthma or psoriasis (Lazarus et al., 2003, Zabłotna et al., 2017). Moreover, TLR9 polymorphism may play an essential role in the development of different kinds of cancer, as shown in prior research (Mandal et al., 2012, Roszak et al., 2012). Some studies have also observed that TLR9 is significantly associated with infectious diseases such as HIV (Bochud et al., 2007), pulmonary aspergillosis (Carvalho et al., 2008), meningitis (Gowin et al., 2017) and trypanosomiasis (Drennan et al., 2005).

The main aim of this chapter is to identify genetic variations in bovine TLR9 and the influence of these variations on the risk of *Trypanosoma* infection. African bovine blood samples with a known

Trypanosoma infection status, which were collected on FTA cards, were produced from previous research in our lab (Ideozu, 2015). To achieve the aims, the following objectives were pursued:

- To identify and analyse the genetic variation by aligning all TLR9 DNA sequences and comparing them to the reference sequence for the bovine gene in the NCBI database.
- To investigate the relationship between the genetic variation in bovine f1TLR9 gene sequences and *Trypanosoma* infection.
- To study the correlation between homozygosity and heterozygosity patterns in f1TLR9 alleles regions with respect to *Trypanosoma* infection.
- To detect modifications in the CpG islands within the bovine f1TLR9 and assess the link between these SNPs and *Trypanosoma* infection.

4.2. Methods

The DNeasy Blood & Tissue kit (Bioline) was used to extract DNA from 70 bovine blood samples initially collected from bulls in Nigeria on FTA cards, as previously described (Ideozu, 2015). We amplified fragment 1 from bovine TLR9 (named f1TLR9) as described in Chapter 3 and the products were sequenced and aligned with the reference gene (accession no. KC174788.1) in order to determine the extent of genetic variation. This approach has allowed the extent of genetic variation in the f1TLR9 to be assessed with respect to parasitic infection using the Chi-Squared test.

4.3. Results

4.3.1. Prevalence of parasitic infection in southern Nigerian cattle

After combining the data from both previous studies (Ideozu, 2015, West, 2018), we found that the prevalence of parasitic infection in southern Nigerian cattle was 30% (95% CI: 24.8–34.5%),

as shown in Fig. 4.1A. The majority of parasitic infections were *Trypanosoma* infections (28.6%; 95% CI: 22–32.7%), followed by *Theileria*, *Sarcocystis* and *Eimeria* infections (<5%), as shown in Fig. 4.1B (West, 2018). To further investigate *Trypanosoma* infection, we closely studied the prevalence of *Trypanosoma* species in the Nigerian cattle using multiple punches from Whatman FTA cards (Ideozu, 2015). In our previous research, the majority of samples were infected with *Trypanosoma vivax* (*T. vivax*) (n = 10, 25%), *Trypanosoma theileri* (*T. theileri*) (n = 7, 20%) and *Trypanosoma simiae* (*T. simiae*) (n = 5, 10%) (Ideozu, 2015).

In addition, a number of the cattle (n = 7, 20%) also showed infection with more than one *Trypanosoma* species (i.e., mixed infections). The same number of cattle (n = 7, 20%) were also categorised as having an unknown T*rypanosoma* infection since the PCR amplicon generated by the ITS-nested PCR approach was inconclusive with respect to species identity (Ideozu, 2015). Overall, our data indicate that trypanosomes are significant parasites that affect Nigerian cattle and *T. vivax* has the highest identified prevalence among the *Trypanosoma* species in the study population.

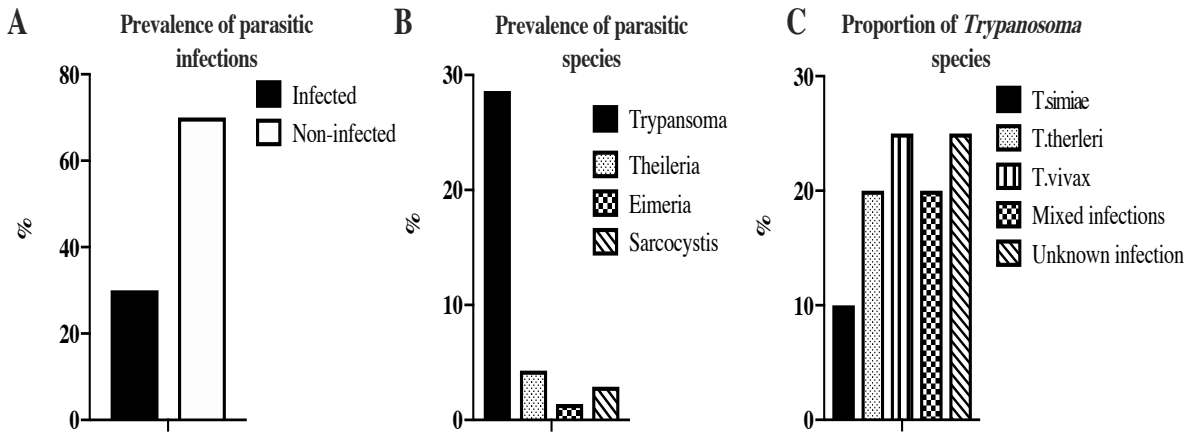


Figure 4. 1 The parasite infection profiles within the Nigerian cattle population. (A) The majority of all parasitic infections detected from FTA cards using NGS and the *Trypanosoma* spp. using ITS-nested PCR and NGS, combined data compiled from Ideozu, (2015) and West, (2018). **(B)** The prevalence of different parasitic infections detected in blood spots from southern Nigerian cattle stored on FTA cards (West, 2018). **(C)** Using ITS-nested PCR, the relative proportions of different *Trypanosoma* species infecting the Nigerian cattle were identified (Ideozu, 2015).

4.3.2. The relationship between parasitic infections and the presence of LRR-24 haplotypes in southern Nigerian cattle

To investigate the genetic variation in f1TLR9 with parasitic diseases, we aimed to identify haplotypes within the f1TLR9 of southern Nigerian cattle. We defined a haplotype as the inheritance of a cluster of SNPs at single positions in the leucine-rich repeat 24 (LRR-24) of bovine TLR9, as illustrated in Fig. 4.2A. The positions of three variants G/A, C/A, and C/A are located at nucleotide numbers 3760, 3768 and 3780 of the *B. indicus* TLR9 gene sequence (Table 4.1).

Table 4. 1 SNPs identified in the LRR-24 of bovine f1TLR9.

SNP allele	SNP position in *B. indicus* TLR9 gene KC174788.1	SNP position in *B. indicus* TLR9 mRNA KX138608.1	SNP position in *B. indicus* TLR9 protein APQ40216.1	Amino acid change
G/A	3760	2193	731	Valine/methionine
C/A	3768	2201	733	No
C/A	3780	2213	737	No

Our results indicated eight haplotypes in the f1TLR9 of southern Nigerian cattle. Of the 70 samples that were investigated, 61 sequences aligned to the same sequencing region. We classified haplotypes based on protein translation (transcript haplotypes)—H1 or H2—as shown in Fig. 4.2B. H1 haplotypes (H1a, b, c, d) did not change the protein sequence; all had the amino acid valine at position 731 of the protein. In contrast, H2 haplotypes (H2a, b, c, and d) all substituted valine at position 731 to methionine and they were identical at the protein level whilst being distinct from the H1 haplotype protein sequence. In total, 68 % of samples had H2 haplotypes (n=79) and 32% of samples had H1 haplotypes (n=38) (Table 4.2). The most well-distributed haplotype among the 70 samples was H2a (A-C-C position 3760,3768,3780), followed by H1a (G-C-C), H2c (A-A-A),

H2b (A-A-C), H2d (A-C-A), H1b (G-A-C), H1c (G-C-A) and H1d (G-A-A). Thus, H2 haplotypes were the predominant haplotype in this region.

Our analysis indicated that 29 (47.5%) of 61 cows had H2/H2 haplotypes, 18 (29.5%) had H1/H2 haplotypes, and only 14 (23%) had H1/H1 haplotypes (Table 4.3). Calculation of the frequency of each protein haplotype shows that H1 (p = 0.38) is less frequent than H2 (q = 0.62). Expected frequencies for H1/H1, H1/H2 and H2/H2 can be calculated by using Hardy-Weinberg ($p^2 + 2pq + q^2$) and for the population of 61 would be, respectively, 8.8, 28.7 and 23.5 animals of each genotype. In comparison to the observed numbers 14, 18, 29, there is a significantly reduced level of heterozygosity (p = 0.0154, $\chi^2 = 8.35$, df = 2) suggesting that these animals are probably derived from an inbred population.

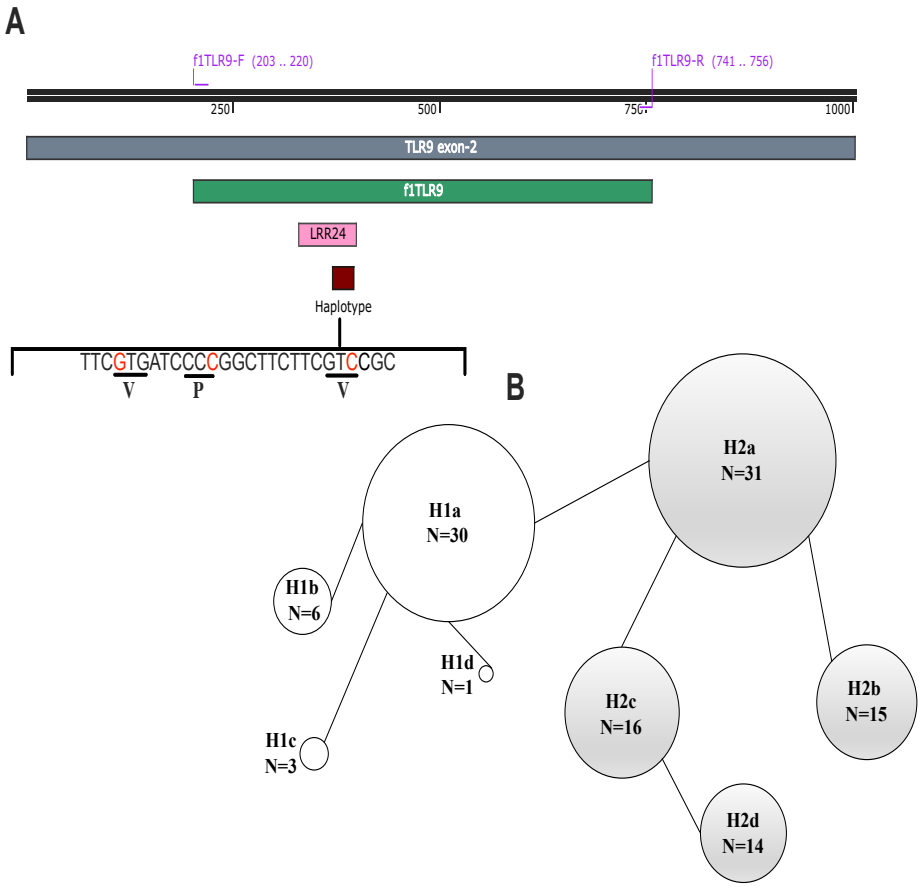


Figure 4. 2 Identified haplotypes within the LLR-24 region of southern Nigerian cattle. (A) Map of the f1TLR9 (554 bp), LLR-24 (69 bp) and haplotypes (28 bp) within the exon 2 region of

the bovine TLR9 gene. **(B)** Schematic illustration of the distribution of haplotypes within the tested samples. Circle sizes reflect the number of haplotype copies observed in the tested samples and the same haplotypes numbers (e.g., H1a, H1b) identify haplotypes that are identical at the amino acid level.

Table 4. 2 Identified transcript haplotypes within the LRR-24 of bovine f1TLR9. In the genotype column, the three letters refer to each polymorphic position (i.e. positions 3760, 3768, 3780).

Genotype	Haplotype	Transcript haplotype	Amino acid	Frequency within the total haplotypes	Percentage of haplotypes
G-C-C	H1a	H1	V-P-V	38	32%
G-A-C	H1b				
G-C-A	H1c				
G-A-A	H1d				
A-C-C	H2a	H2	M-P-V	79	68%
A-A-C	H2b				
A-A-A	H2c				
A-C-A	H2d				

Table 4. 3 Frequency of identified transcript haplotypes of bovine f1TLR9.

	Transcript haplotype	Frequency	% occurrence within total population	Valid %
Valid	H1/H1	14	20%	23%
	H2/H2	29	41.4%	47.5%
	H1/H2	18	25.7%	29.5%
	Total	61	87.1%	100%
Missing		9	12.9%	
Total		70	100%	

Table 4. 4 Evaluation of the association between parasitic infection and the presence of identified haplotypes in bovine f1TLR9.

Transcript Haplotype	Count/percent within transcript haplotype	Infected	Non-infected	Total	Significance
H1/H1	Count	4	10	14	No
	%	28.6%	71.4%	100%	
H2/H2	Count	8	21	29	No
	%	27.6%	72.4%	100%	
H1/H2	Count	7	11	18	No
	%	38.9%	61.1%	100%	
Total	Count	19	42	61	0.70
	%	31.1%	68.9%	100%	

To investigate the association between protein substitution from valine to methionine in bovine f1TLR9 and susceptibility or resistance to parasitic infection (*Trypanosoma*, *Theileria*, *Sarcocystis* and *Eimeria*), we studied the relationship between the transcript haplotypes and parasite-infected and non-parasite-infected cows. In total, 19 (31.1%) of the 61 cows had a parasitic infection, and 42 (68.9%) were not infected. Eight (27.6%) of the 19 parasite-infected cows were found to have H2/H2 haplotypes, seven (38.9%) had H1/H2 haplotypes, and four (28.6%) had H1/H1 haplotypes. Also, 21 (72.4%) of the 42 non-parasite-infected cows were found to have H2/H2 haplotypes, 11 (61.1%) had H1/H2 haplotypes, and 10 (71.4%) had H1/H1 haplotypes. Our data showed no significant differences among different haplotypes with infected and non-infected cows (p-value:

0.70; Pearson's chi-square). Overall, the results demonstrated that there was no correlation between the transcript haplotypes and susceptibility or resistance to parasitic infection (Table 4.4). Furthermore, our previous data showed that the primary source of parasitic infection in southern Nigerian cattle was *Trypanosoma* (Ideozu, 2015). Therefore, we closely studied the impact of different transcript haplotypes on susceptibility and resistance to *Trypanosoma* species infection. Many *Trypanosoma* species have been detected in southern Nigerian cattle, including *T. vivax*, *T. theirleri* and *T. simiae*. We found that, in total, 18 of 61 cows were infected with *Trypanosoma* species: four (28.6%) cows had H1/H1 haplotype, eight (27.6%) cows had H2/H2 haplotypes, and six (33.3%) cows had H1/H2 haplotypes. The H1/H1 haplotype was found on 10 (71.4%) non-*Trypanosoma* -infected cows, one (7.1%) cow infected with *T. theileri,* one (7.1%) cow infected with *T. vivax*, and two (14.3%) cows infected with unknown *Trypanosoma* species. The H2/H2 haplotype was found in 21 (72.4%) non- *Trypanosoma* -infected cows, one (3.4%) cow infected with *T. simiae,* three (10.3%) cows infected with *T. theirleri,* two (6.9%) cows infected with *T. vivax,* and two (6.9%) cows infected with mixed *Trypanosoma* infections. The H1/H2 haplotype was found on 12 (66.7%) non- *Trypanosoma* -infected cows, one (5.6%) cow infected with *T. simiae,* two (11.1%) cows infected with *T. vivax,* two (11.1%) cows infected with mixed *Trypanosoma* infections, and one (5.6%) cow infected with mixed *Trypanosoma* infections. Moreover, we considered multiple *Trypanosoma* species infection. In total, 6.6% of cows had such infection: 6.9% of these cows had H2/H2 haplotypes, and 11.1% had H1/H2 haplotypes. In addition, 4.9% of cows were infected with an unidentified *Trypanosoma* species (i.e. unknown infection). Two (14.3%) of these cows were discovered to have H1/H1 haplotypes, and one (5.6%) had an H1/H2 haplotype (Table 4.5). There was no association between genetic variation in Nigerian cows and infection with different *Trypanosoma* species (p-value: 0.91; Pearson's chi-square). Also, there was no relationship between transcript haplotypes and non-infected samples.

To sum up, there was no significant correlation between the transcript haplotypes and susceptibility or resistance to infection of different *Trypanosoma* species in southern Nigerian cattle.

Then, we analysed genetic variants of the eight haplotypes identified earlier to define the patterns of homozygosity (HO) of heterozygosity (HT) in the context of parasitic infection. The most frequent haplotypes were heterozygous (65.6%), although some homozygous haplotypes were found (34.4%; Table 4.6). Among the heterozygous sample, 35% of cows had a parasitic infection, and 65% were not infected. Among the homozygous samples, 23.8% of cows had a parasitic infection and 76.2% were not infected. Our study showed that none of the heterozygous or homozygous haplotypes was significantly related to parasitic infection (p-value: 0.275; Pearson's chi-square) and there was no correlation between the heterozygous or homozygous haplotypes and susceptibility or resistance to parasitic infection (Table 4.7).

Table 4. 5 Evaluation of the association between *Trypanosoma* infection and the presence of identified haplotypes in bovine f1TLR9.

Transcript haplotype	Count/percent within transcript haplotype	Non-infected	*T. simiae*	*T. theileri*	*T. vivax*	Mixed infections	Unknown infection	Total	Significance
H1/H1	Count	10	0	1	1	0	2	14	No
	%	71.4%	0%	7.1%	7.1%	0%	14.3%	100%	
H2/H2	Count	21	1	3	2	2	0	29	No
	%	72.4%	3.4%	10.3%	6.9%	6.9%	0%	100%	
H1/H2	Count	12	1	0	2	2	1	18	No
	%	66.7%	5.6%	0%	11.1%	11.1%	5.6%	100%	
Total	Count	43	2	4	5	4	3	61	0.91
	%	70.5%	3.3%	6.6%	8.2%	6.6%	4.9%	100%	

Since the main parasites infecting the tested samples were *Trypanosoma* species, we examined the association between heterozygous or homozygous haplotypes and *Trypanosoma* species infection. The data showed that heterozygous haplotypes were found in 27 (67.5%) of the cows with no *Trypanosoma* infection and 13 (32.5%) of the cows infected with *Trypanosoma*. Of the cows with heterozygous alleles, four (10%) cows were infected with *T. vivax* and multiple *Trypanosoma* species, two (5%) cows were infected with *T. theileri*, two (5%) cows were infected with *T. simiae,* and one (2.5%) cow was infected with unknown *Trypanosoma* species. Homozygous haplotypes were found in 16 (76.2%) non-infected cows and 5 (23.8%) cows infected with *Trypanosoma*. Of these, two (9.5%) cows were infected with *T. theileri* two (9.5%) cows unknown *Trypanosoma* species, and one (4.8%) was infected with *T. vivax*.

Overall, our results showed no significant correlation between heterozygous or homozygous haplotypes and susceptibility or resistance to infection of different *Trypanosoma* species (p-value: 0.35; Pearson's chi-square) (Table 4.8).

Table 4. 6 Distribution of homozygous and heterozygous alleles in bovine f1TLR9.

Alleles		Frequency	% occurrence within total population	Valid %
Valid	HO	21	30%	34.4%
	HT	40	57.1%	65.6%
	Total	61	87.1%	100%
Missing		9	12.9%	
Total		70	100%	

Table 4. 7 Evaluation of the association between parasitic infection and homozygous and heterozygous alleles in bovine f1TLR9.

Alleles	Count/percent of alleles	Infected	Non-infected	Total	Significance
HO	Count	5	16	21	No
	%	23.8%	76.2%	100%	
HT	Count	14	26	40	No
	%	35%	65 %	100%	
Total	Count	19	42	61	0.275
	%	31.1%	68.9%	100%	

Table 4. 8 Evaluation of the association between infection of different *Trypanosoma* species and homozygous and heterozygous alleles in bovine f1TLR9.

Alleles	Count/percent within alleles	Non-infected	*T. simiae*	*T. theileri*	*T. vivax*	Mixed infections	Unknown infection	Total	Significance
HO	Count	16	0	2	1	0	2	21	No
	%	76.2%	0%	9.5%	4.8%	0%	9.5%	100%	
HT	Count	27	2	2	4	4	1	40	No
	%	67.5%	5%	5%	10%	10%	2.5%	100%	
Total	Count	43	2	4	5	4	3	61	0.35
	%	70.5%	3.3%	6.6%	8.2%	6.6%	4.9%	100%	

4.3.3. Investigation of the relationship between genetic variants at CpG sites in bovine f1TLR9 and parasitic infection

Previous study found two CpG islands in the bovine TLR9 genomic sequence but no such islands in TLR3, 7 or 8 (Cargill and Womack, 2007). Similarly, we detected two CpG islands in the bovine TLR9 genomic sequence (accession no: KC174788.1) using CpGPlot from EMBOSS (https://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/). The first CpG island detected in bovine TLR9 has a length of 259 bp and is located at 3706–3964 bp, while the second CpG island has a length of 421 bp and is located at 4167–4587 bp. Based on further investigations, we found that f1TLR9 contains CpG island 1, and we detected genetic variation at two CG positions identified earlier in the LRR-24 within the bovine f1TLR9 (see Fig. 4.3). The first position (P1) is at 3760 bp and the second position (P2) is at 3768 bp. Three haplotypes were discovered at each position in 61 of the 70 tested samples, and then we examined the correlation between these haplotypes and susceptibility or resistance to different parasitic infections (*Trypanosoma*, *Theileria*, *Sarcocystis* and *Eimeria*). In P1, 49.2% of samples were found to have a CA haplotype, 27.9% were found to have a CA/CG haplotype, and 23% were found to have a CG haplotype (Table 4.9).

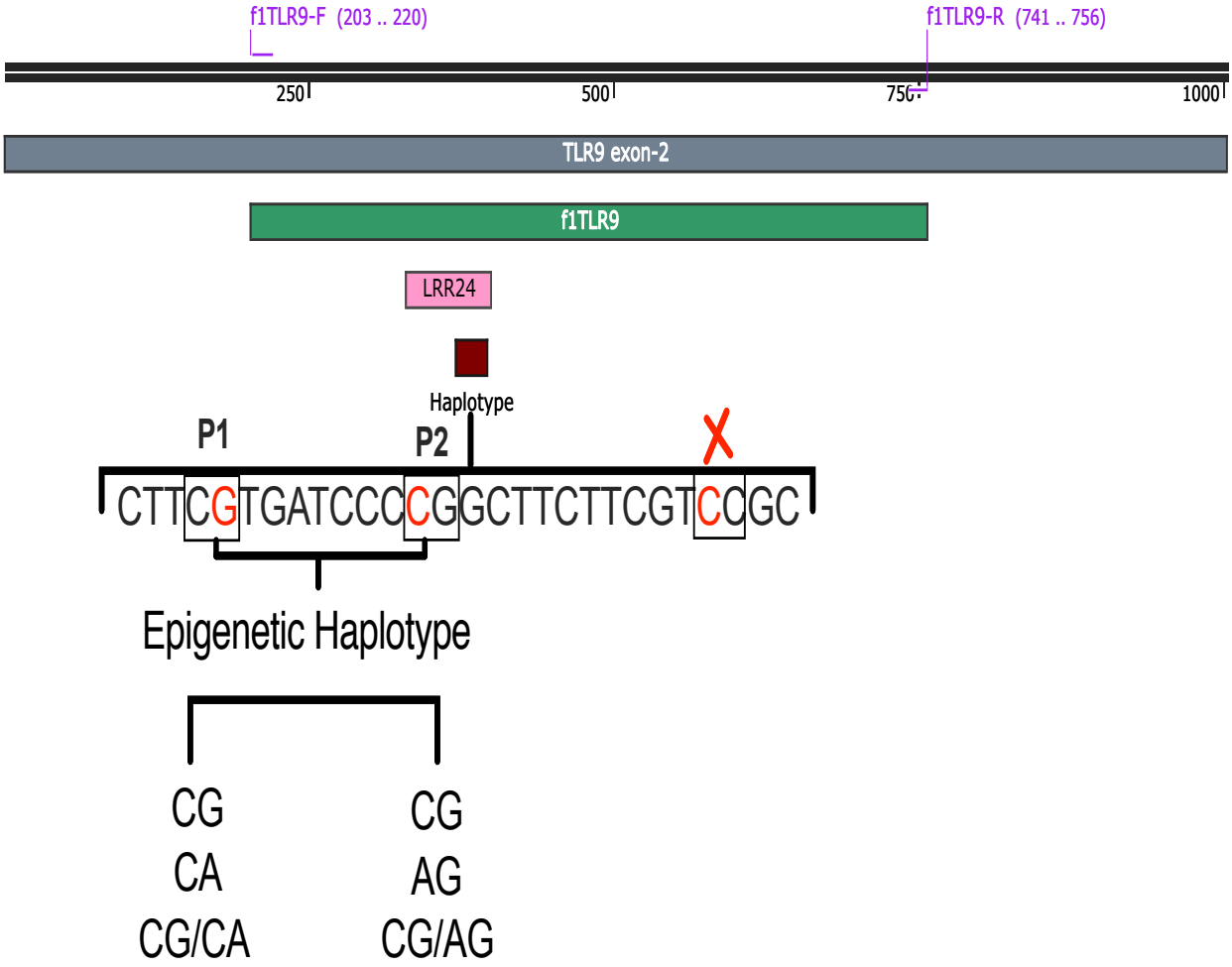


Figure 4. 3 Mapping of haplotype patterns located in the CpG site within bovine f1TLR9. Two genetic variations were detected at position 1 (P1) and position 2 (P2) in f1TLR9. The red cross highlights that SNPs are excluded since it is not a valid CpG site.

Table 4. 9 Frequency of haplotype patterns located at P1 of the CpG site within bovine f1TLR9.

Haplotype		Frequency	% occurrence within total population	Valid %
Valid	CA	30	42.9%	49.2%
	CG	14	20%	23%
	CA/CG	17	24.3%	27.9%
	Total	61	85.7%	100%
Missing		9	12.9%	
Total		70	100%	

Table 4. 10 Evaluation of the association between parasitic infection and haplotype patterns located at P1 of the CpG site within bovine f1TLR9.

Genotype	Count/percent at P1	Infected	Non-infected	Total	Significance
CA	Count	10	20	30	No
	%	33.3%	66.7%	100%	
CG	Count	5	9	13	No
	%	35.7%	64.3%	100%	
CA/CG	Count	4	13	17	No
	%	23.5%	76.5%	100%	
Total	Count	19	42	61	0.66
	%	31.1%	68.9%	100%	

To determine whether the haplotypes at P1 are related to parasitic infection, we analysed our data, and the results showed that, of the 61 cattle samples with different haplotypes, 68.9% were not infected with parasites and 31.1% were infected. Of the cows with CA haplotypes, 20 (66.7%)

cows were non-infected, and ten (33.3%) cows were infected. Of the cows with CG haplotypes, nine (64.3%) samples were not infected, while five (35.7%) samples were infected. Lastly, the CA/CG haplotypes were found in 13 (76.5%) cows non-infected samples and 4 (23.50%) cows infected samples.

Investigation of the relationship between haplotypes and susceptibility or resistance to parasitic infection (Table 4.10) revealed no significant correlation (p-value: 0.658; Pearson's chi-square) between P1 haplotypes discovered in the P1 of f1TLR9 and parasitic infection. Further investigation of P1 haplotypes in cattle samples infected with different *Trypanosoma* species (*T. vivax*, *T. theileri* and *T. simiae*) revealed that CA haplotypes were found in 30 cattle samples. Of the cows with CA haplotypes, 20 (66.7%) cows were not infected with *Trypanosoma*, four (13.3%) cows were infected with *T. vivax*, three (10%) cows were infected with *T. theirleri*, one (3.3%) was infected with *T. simiae*, and two (6.7%) cows were infected with mixed *Trypanosoma* infection. CG haplotypes were found in 14 cattle samples, 9 (64.3%) of which had no *Trypanosoma* infection and five of which were infected with *Trypanosoma*. Of the infected samples, one (7.1%) cow had *T. vivax*, one (7.1%) cow had *T. theileri*, one (7.1%) cow had mixed *Trypanosoma* infection, and two (14.3%) cows were infected with unknown *Trypanosoma* species. Lastly, CA/CG-haplotypes were found in 17 cattle samples. Of the samples, 14 (82.4%) samples were not infected with *Trypanosoma*, and 3 were infected. Of the infected samples, one (5.9%) cow had *T. simiae*, one (5.9%) cow had mixed *Trypanosoma* infection, and one (5.9%) cow had unknown *Trypanosoma* infection. This analysis indicated no significant correlation (p-value: 0.471; Pearson's chi-square) between P1 haplotypes and infection with different *Trypanosoma* species (Table 4.11).

Table 4. 11 Evaluation of the association between infection with different *Trypanosoma* species and haplotype patterns located at P1 of the CpG site within bovine f1TLR9.

Genotype	Count/percent within P1	Non-infected	*T. simiae*	*T. theileri*	*T. vivax*	Mixed infections	Unknown infection	Total	Significance
CA	Count	20	1	3	4	2	0	30	No
	%	66.7%	3.3%	10%	13.3%	6.7%	0%	100%	
CG	Count	9	0	1	1	1	2	14	No
	%	64.3%	0%	7.1%	7.1%	7.1%	14.3%	100%	
CA/CG	Count	14	1	0	0	1	1	17	No
	%	82.4%	5.90%	0%	0%	5.9%	5.9%	100%	
Total	Count	43	2	4	5	4	3	61	0.47
	%	70.5%	3.3%	6.6%	8.2%	6.6%	4.9%	100%	

The P1 haplotypes at the CpG site within f1TLR9 demonstrated no significant correlation with parasitic infection.

We also explored whether the second position (P2) haplotypes were related to parasitic infection. In total, 36 of the 61 cattle samples (60%) were found to have CG haplotypes, 17 (27.9%) had AG/CG haplotypes, and 7 (11.5%) had AG haplotypes (Table 4.12). Forty-two (68.9%) samples did not have a parasitic infection, and 19 (31.1%) were infected. Of those with the CG haplotype, 25 (67.6%) cows were not infected with parasites, and 12 (32.4%) had a parasitic infection. Of those with AG/CG haplotypes, 11 (64.7%) had no parasitic infection, and 6 (35.3%) had a parasitic infection. Last, of those with AG haplotypes, six (85.7%) were not infected with parasites, and one (14.3%) had a parasitic infection. The results of our investigation of the relationship between haplotypes and susceptibility or resistance to parasitic infection illustrated that there was no significant correlation (p-value: 0.569; Pearson's chi-square) between P2 haplotypes discovered within f1TLR9 and parasitic infection (Table 4.13).

Table 4. 12 Frequency of haplotype patterns located at P2 of the CpG site within bovine f1TLR9.

Haplotype		Frequency	% occurrence within total population	Valid %
Valid	AG	7	10%	11.5%
	CG	37	52.9%	60%
	AG/CG	17	24.3%	27.9%
	Total	61	85.7%	100%
Missing		9	12.9%	
Total		70	100%	

Table 4. 13 Evaluation of the association between parasitic infection and haplotype patterns located at P2 of the CpG site within bovine f1TLR9.

Genotype	Count/percent within P1	Infected	Non-infected	Total	Significance
AG	Count	1	6	7	No
	%	14.3%	85.7%	100%	
CG	Count	12	25	36	No
	%	32.4%	67.6%	100%	
AG/CG	Count	6	11	17	No
	%	35.3%	64.7%	100%	
Total	Count	19	42	61	0.569
	%	31.1%	68.9%	100%	

Since the primary purpose of this chapter is to identify the association between genetic variation and parasitic infection, we analysed the relation between P2 haplotypes and infection with different *Trypanosoma* species (*T. vivax*, *T. theirleri* and *T. simiae*). The data showed that CG haplotypes were found in 36 samples. Of these, 25 (67.6%) were not infected with *Trypanosoma*, while 11 (32.4%) were infected. Of those with GC haplotypes, one (2.7%), the sample was found to be infected with *T. vivax,* three (8.1%) with *T. theirleri*, two (5.4%) with *T. simiae*, three (8.1%) with mixed *Trypanosoma* species and three (8.3%) with unknown *Trypanosoma* species. Among those with AG/CG haplotypes, 12 of 17 (70.6%) had no *Trypanosoma* infection, and five (29.4%) were infected. Of these, four (23.5%) were infected with *T. vivax,* and one (5.9%) was infected with *T. theirleri*. Of the seven samples with AG haplotypes, six (85.7%) were not infected with *Trypanosoma*, while one (14.3%) had mixed *Trypanosoma* infection. Analysis of the results revealed no significant correlation (p-value: 0.24; Pearson's chi-square) between P2 haplotypes and infection with different *Trypanosoma* species (Table 4.14).

Table 4. 14 Evaluation of the association between infection with different *Trypanosoma* species and haplotype patterns located at P2 of the CpG site within bovine f1TLR9.

Genotype	Count/percent within P1	Non-infected	*T. simiae*	*T. theileri*	*T. vivax*	Mixed infections	Unknown infection	Total	Significance
AG	Count	6	0	0	0	1	0	7	No
	%	85.7%	0%	0%	0%	14.3%	0%	100%	
CG	Count	25	2	3	1	3	3	36	No
	%	67.6%	5.4%	8.1%	2.7%	8.1%	8.1%	100%	
AG/CG	Count	12	0	1	4	0	0	17	No
	%	70.6%	0%	5.9%	23.5%	0%	0%	100%	
Total	Count	42	2	4	5	4	3	61	0.24
	%	70.5%	3.3%	6.6%	8.2%	6.6%	4.9%	100%	

4.4. Discussion

TLR9 is known to recognise unmethylated CpG dinucleotides from microbial DNA. Genetic variation in TLR9 has been postulated to play a significant role in a host's susceptibility or resistance to an infectious disease (Mukherjee et al., 2019). Our report identified genetic variation in TLR9 genes infected with *Trypanosoma* from Nigerian cattle. All the blood samples obtained in this study were previously collected and stored on FTA cards to determine the prevalence of *Trypanosoma* infection in Nigerian cattle (Ideozu, 2015). The data showed that the overall prevalence of parasitic infection is 30% (95% CI: 24.8–34.5%), including infection with *Trypanosoma*, *Eimeria*, *Sarcocystis* and *Theileria* (Fig. 4.1B). Of these parasitic infections, trypanosomiasis accounts for the majority of infections (28.6%; 95% CI: 22–32.7%) (Fig. 4.1C). It is primarily caused by *T. vivax* among *B. indicus* bulls in Ahoada, a town in southern Nigeria. The previous study revealed that the overall prevalence of trypanosomiasis in Nigeria from 1960 to 2017 is 16.1% (95% CI: 12.3–20.3%; (Odeniran and Ademola, 2018). Another study showed that the prevalence of trypanosomiasis in Nigeria detected by using the buffy coat technique and Giemsa thin blood films is 8.4% (Enwezor et al., 2009).

An essential factor that may play a role in evaluations of the prevalence of trypanosomiasis is the type of diagnostic technique that is employed. Previously, our lab detected trypanosomiasis using a PCR technique called internal transcribed spacer-nested PCR (ITS-nested) (Cox et al., 2005; Cox et al., 2010). We previously reported that the ITS-nested technique is an effective tool for diagnosing various species of trypanosomes (Ideozu, 2015).

Since Sanger sequencing is designed for sequencing relatively short pieces of DNA (about 300 to 1000 bp), we separated exon 2 of bovine TLR9 into four fragments to obtain high resolution of data after DNA sequencing, as described in Chapter 3. Fragment 1 (f1TLR9) was selected as it contains a CpG island (Cargill and Womack, 2007) and also, a region of 24 amino acids

corresponding to a LRR motif (NCBI Reference Sequence: NP_898904.1) that may bind foreign DNA.

To date, the genetic diversity of this region has not been studied with parasitic infection in *B. indicus*. However, a previous study shows that there are associations between TLR9 signalling and parasitic infection (Drennan et al., 2005). Therefore, we postulated that genetic variation within this region might be correlated with susceptibility or resistance of Nigerian *B. indicus* cattle to parasitic infection.

We identified and categorised the genetic variation within f1TLR9 into two categories: protein haplotype and CG haplotype. For the protein haplotype category, we identified eight haplotypes and sorted them based on their translation profile. The transcript or protein haplotype is defined as a non-synonymous missense SNP that is detected due to a nucleotide change of G to A, which leads to amino acid substitution in the functional domain at position 3760 (V3760M) of the encoded protein. We analysed the coding sequence of the functional domain (LRR-24 region) in *B. indicus* to determine its biological significance with susceptibility or resistance to parasitic infection. A previous study by Matsumoto et al. (2012) reported that 13 SNPs within the coding sequence of the fatty acid synthase (FASN) gene, of which five were non-synonymous mutations, were used to investigate the association between genetic diversity and economic traits in 198 Holstein cattle. Two of the non-synonymous SNPs affected milk fat content and mirystisic acid (C14) (Matsumoto et al., 2012). A different study identified one non-synonymous missense SNP in bovine TLR9 at position 369 (R396H), which may have some biological significance (Cargill and Womack, 2007). We could not make a direct comparison and identify whether the SNP and allele identified in that study are similar to those identified in this study because they are not located within the f1TLR9.

In our study, the H2/H2 haplotype was the more frequent haplotype found in cattle (29 of 70 samples, 47.5%). Nine samples (12.8%) were missing due to poor recovery of the PCR product.

The protein haplotypes found in f1TLR9 showed no significant correlation to parasitic infection in Nigerian *B. indicus* cattle. However, a recent study showed a significant association between TLR9 polymorphism and susceptibility or resistance to bovine tuberculosis (Sun et al., 2012). In contrast, Oliveira et al. (2013) found no consistently significant associations between TLR9 polymorphism and human papillomavirus. Our explanation for the lack of correlation between amino acid substitution within f1TLR9 and parasitic infection is that valine and methionine amino acids share most of the same chemical properties, except methionine, has a sulphur group, which may contribute to the structural stability of proteins (Lim et al., 2019).

Furthermore, we analysed the patterns of homozygosity and heterozygosity within the f1TLR9. Our data revealed that there is no correlation in both the heterozygosity and homozygosity of Nigerian *B. indicus* cattle genes and parasitic infection. We noticed that the prevalence of heterozygosity is lower than that of homozygosity, which may be derived from an inbred population. A previous study of East African cattle reported a strong correlation between low heterozygosity pattern and both death and illness caused by infectious diseases (Murray et al., 2013).

We considered the SNPs in CpG island 1 as potential sites for methylation patterns, which, once established and maintained at CpG sites, may be altered and affect the regulation of the TLR9 gene. Previous studies reported that the presence of an SNP at a CpG site was positively correlated with methylation of CpG islands and was associated with gene expression (Hellman and Chess, 2010; Dayeh et al., 2013). Moreover, Maldonado et al. (2019) identified methylation SNPs (meSNPs) at a CpG site from DNA extracted from the tissue of cattle with different feed efficiencies, and they expected that these meSNPs were mechanistically responsible for creating or deleting the methylation targets that are responsible for differential expression of genes. Our data showed that two SNPs were present at P1 and P2 in the CpG island 1 (Fig. 4.3). Therefore, we hypothesised that variations caused by SNPs at CpG island 1 may alter the epigenetic profile

of the gene, leading to an association with susceptibility to parasitic infection. However, our data revealed that there is no correlation between CpG-SNPs and the incidence of parasitic infection. In conclusion, in this chapter, we investigated three SNPs at CpG island 1 within exon 2 of bovine TLR9. Of these, one is a non-synonymous SNP resulting in a change in protein sequence, while the other two are synonymous and do not result in a change in the protein sequence. We found no significant association between parasitic infection and these SNPs in the CpG island 1 within f1TLR9.

CHAPTER FIVE

Comparison of Sanger Sequencing with Next Generation Sequencing (NGS) Tools for Investigating Variation in the Bovine TLR9 Gene

5.1.Introduction

The first human genome sequencing project (HGP) was completed in 2004 at an estimated cost of 2.7 billion dollars (Hood and Rowen, 2013). The HGP was instrumental in the development of high-throughput technologies for preparing, mapping and sequencing DNA. Moreover, sequencing eukaryotic genomes have had a significant scientific, economic and cultural impact that has changed the nature of biomedical research and medicine (Xuan et al., 2013).

The development of next-generation sequencing (NGS) technologies has improved the ability to sequence multiple genes and allow either the whole genome or parts of it to be sequenced in a short time and at great depth with increased sensitivity. In addition, NGS has provided a new opportunity for a fast, low-cost, accurate way to direct a shift from traditional Sanger sequencing methods towards a new high-throughput genomic methodology that generates millions of sequences simultaneously from one sample. In 2005, the first massively parallel DNA sequencing platforms emerged, ushering in a new era of NGS (Margulies et al., 2005, Shendure et al., 2005). Since 2008, NGS technologies have been in use primarily in research on different strategies for DNA sequencing. The cyclic array sequencing strategy can sequence a dense array of DNA molecules by performing repeated cycles of enzymatic manipulation and imaging-based data collection and dispenses with the amplification step (Shendure et al., 2004).

Recently, the cyclic array sequencing strategy has been implemented in commercial products developed by biotechnology companies such as Roche and Illumina. While these products differ in the specifics of their biochemistry as well as in how the array of DNA is generated, they use essentially similar sample preparation (Magi et al., 2010). The sample preparation involves the construction of a sequencing library, utilises the universal adaptors to allow amplification, and sequencing of all the DNA fragments in parallel using a single primer pair and sequencing primer (Metzker, 2010). Currently, there are two major amplification methods used in libraries for the

NGS reaction that are carried out: single-molecule templates and clonally amplified templates. The single-molecule template method requires a small amount of genomic DNA material and does not require PCR amplification, which increases the chance of creating mutations in clonally amplified templates that masquerade as sequence variants. The single-molecule template technology depends upon using specific probes to interrogate sequences such as AT-rich and GC-rich sequences (Slatko et al., 2018).

The clonally amplified template preparation method requires amplification of target fragments to initiate the NGS reaction. The two most common amplifying methods are emulsion PCR (emPCR) and solid-phase amplification or bridge PCR (bPCR) (Ma et al., 2013). Both PCR methods amplify by ligating adaptors containing universal priming sites to the fragment ends. The emulsion PCR uses beads that are immobilised in a polyacrylamide gel on a standard microscope slide or wells in which the NGS chemistry can be performed (Kanagal-Shamanna, 2016). In contrast, the solid-phase amplification produces template clusters on a glass slide (Cao et al., 2017).

The production of high-throughput data with low-cost budget by NGS technology makes it useful in a variety of areas in biology. NGS technologies are currently used to generate the primary genomic sequence of an organism and define the genetic variations. Several projects have successfully generated and assembled a *de novo* eukaryotic genome based on NGS technology. Li et al. (2010) generated the first draft of the *Ailuropoda melanoleuca* (panda) genome, covering approximately 94% of the whole genome. Recent studies have generated a high-quality genome assembly for bacteria (Charette et al., 2012, Jeong et al., 2013, Kumar et al., 2012), fungi (DiGuistini et al., 2009), yeast (Liti et al., 2013) and viruses (Marston et al., 2013, Montmayeur et al., 2017). NGS technology has also been applied to the entire body of genome resequencing data to identify genetic variations, including single nucleotide polymorphisms (SNP), insertions/deletions and structural variations. For example, a previous study used the NGS method to define the diversity of human platelet antigen polymorphisms (Davey et al., 2017). Other studies

used NGS to identify single nucleotide variants, such as identifying intraspecies genetic variations in cattle (Barris et al., 2012) and structural variants in the human genome (Bentley et al., 2008, Pelak et al., 2010). Furthermore, NGS technology is among the most promising approaches to identify whether SNPs are associated with diseases. Indeed, recent studies have highlighted the relationship between SNPs and some genetic diseases, such as retinal dystrophies (Glöckle et al., 2014), amyotrophic lateral sclerosis (ALS) (Wu et al., 2012) and cancer (Agrawal et al., 2011, Liu et al., 2012).

Currently the Illumina NGS platform is the most commonly used as it is cheaper and more accurate that other platforms (Shendure and Ji, 2008). The Illumina sequencing approach or bridge amplification process involves clonal amplification of adaptor-ligated DNA fragments immobilised at one end on the surface, where each template creates a 'bridge' structure by hybridising with its free end to the complementary adapter on the surface of a glass slide called a flow cell (Kumar et al., 2012). Each single-stranded fragment is created as a copy close to the original fragment by a mixture containing the PCR amplification reagents. The adapters on the surface act as primers for the subsequent PCR amplification (Ansorge, 2009). After several PCR cycles, about 1,000 copies of single-stranded DNA fragments are created on the surface, forming a surface-bound colony or cluster. Bases are read by using a cyclic reversible termination strategy, which floods the flow cell with cleavable fluorescent nucleotides. The four fluorescent nucleotides are labelled with 3'-O-azidomethyl-dNTPs to pause the polymerisation reaction and enable removal of unincorporated bases. The four fluorescent nucleotides are added for all clusters at once through progressive rounds of base incorporation, washing, imaging and cleaving (Guo et al., 2008, Turcatti et al., 2008).

The first Illumina sequencing platform was the Genome Analyzer (GA), which produced ~35 bp reads and generated more than 1 Gb of high-quality sequence data per run in 2–3 days. Across all

Illumina models, the overall error rates were below 1%, and the most common type of error was a substitution (Dohm et al., 2008).

However, the Illumina is limited by sequence read lengths, which are short due to dephasing effects or signal decay, and error rate, which increases with reads length. The dephasing effects happen in any cycle of sequencing, such as decreased or increased efficacy of the nucleotide, or some terminating moiety's failure to remove or add that leads to incomplete extension or overextension of the growing strand along with the template. Furthermore, the signal dephasing, such as decay in the fluorescent signal or dark nucleotides, which are nucleotides without a fluorescent label, are incorporated into clusters.

This chapter aims to develop a protocol to extract genomic DNA from previously collected bovine blood on FTA card and using NGS as a tool for detecting polymorphisms in fragment 1 of bovine TLR 9 (f1TLR9) (Fig. 3.8: chapter 3). Additionally, we compare the quality of the data generated from the NGS method with data from the Sanger method and assess the degree of agreement between these two methods. The study was specifically aimed at:

- To link the overhang adapter to the target fragment of bovine TLR9.
- To purify the amplicon away from free primers and primer-dimer species.
- To link Illumina index sequencing to the f1TLR9 amplicon.
- To determine the size of the amplicon for each library, then quantifying, normalising and pooling all libraries.
- Assess the quality of NGS data.
- Detect the genetic variants in NGS data.

5.2. Methods

Bovine blood samples were taken from our previous study (Ideozu, 2015). DNA was extracted from two punches of FTA cards using the DNeasy Blood & Tissue extraction kit (QIAGEN). PCR amplification was carried out in four rounds to prepare our libraries and the workflow of NGS

sequencing is illustrated in Fig. 5.1. The first round amplified bovine TLR9, the second and third rounds connected the adapter with the amplicon DNA, and the fourth round attached the Illumina sequencing index to the amplicon DNA. Between each PCR, the PCR products were cleaned by using the AMPure XP beads magnetic PCR Clean up Kit (Beckman Coulter). The high sensitivity measured DNA amplicons for each library D1000 ScreenTape Assay with the 2200 TapeStation System (Agilent Technologies). The final library was then adjusted to 4nM with 10 mM Tris pH 8.5 using the Qubit 4 Fluorometer (ThermoFisher Scientific). The library was pooled with 5 µl of diluted DNA extracted from each library with the unique indexes. After that, the concentration of pooled libraries was quantified by utilising a quantification PCR. Before the beginning of the MiSeq sequencing run, we denatured the pooled libraries and diluted them with a hybridisation buffer. We then loaded an 8–12 pM concentration from the denatured pooled libraries and a 20 pM concentration from the denatured Phix (Illumina), diluting in an HT1 hybridisation buffer (Illumina). Sequencing was performed on the MiSeq system (Illumina) using the MiSeq Reagent Nano Kit v2 (500 cycles) (Illumina). The sequencing data for each sample was traced back with a unique barcode sequence. The raw sequenced data generated by the MiSeq was stored in a FASTQ file to analyse the quality control of the nucleotide sequence reads using the FastQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). After the quality control, output reads of the NGS, output reads were examined, the reads were mapped to sequences to the reference genome accession no KC174788.1 from NCBI (https://www.ncbi.nlm.nih.gov/nucleotide/KC174788.1?report=genbank&log$=nucltop&blast_rank=4&RID=4UCCPX3J016). The BBTool (https://jgi.doe.gov/data-and-tools/bbtools/) was used to align the sequencing data to the reference genome sequence. The Variant Calling Format (VCF) tools program (https://vcftools.github.io/index.html) was used to store and analyse DNA polymorphism data such as SNPs, insertions, deletions and structural changes, as well as variants, and store the data as VCF files. In addition, the Microsoft Excel programme was used to analyse

read quality control. We used the Tablet software (https://ics.hutton.ac.uk/tablet/download-tablet/) to show all reads for each sample. The association between genetic variation in Nigerian cattle and susceptibility or resistance to parasitic infection was then compared using Pearson's Chi-square. The Kappa coefficient analysis was then applied to measure the degree of agreement between the Sanger sequencing and NGS methods.

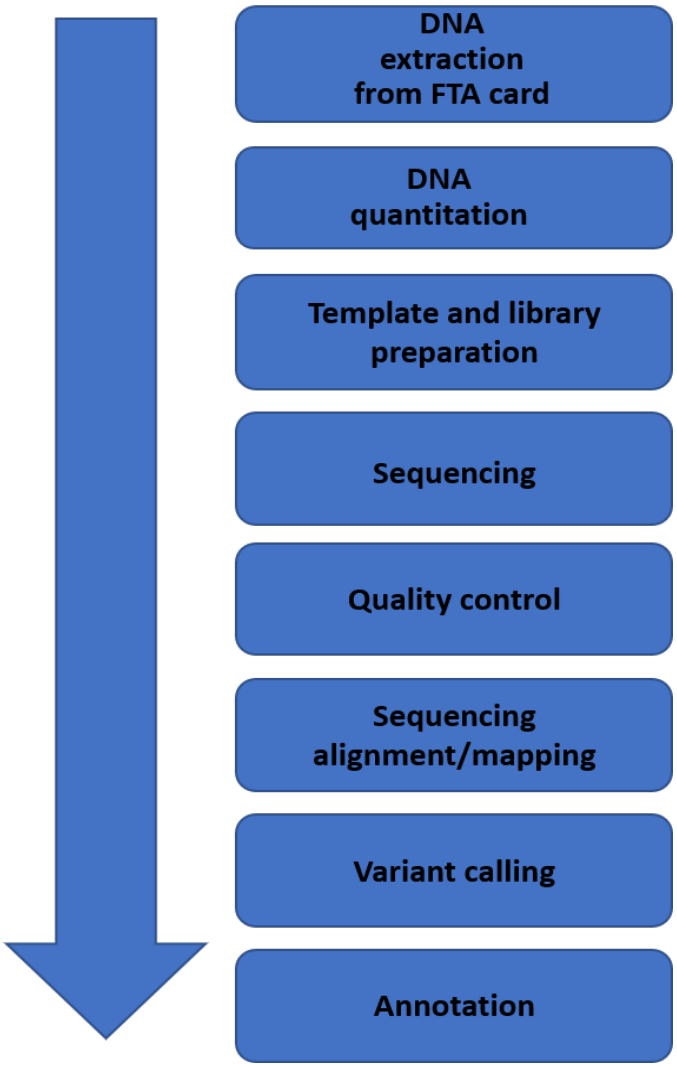


Figure 5. 1 Next-generation sequencing workflow. Library preparation, sequencing and data analysis.

5.3.Results

5.3.1. Optimising conditions for library construction

Preparation of the NGS libraries started with the amplification of the starting material, which was fragment 1 on exon 2 bovine TLR 9 (f1TLR9) taken from 80 blood samples spotted onto FTA cards. We took two punches (2 mm) from each sample and the genomic DNA was extracted using a DNeasy Blood & Tissue extraction kit. The quality and concentration of DNA were measured using a NanoDrop Spectrophotometer. The DNA extraction procedure and results were described in Chapter 2.4 and Chapter 3.1 but, briefly, involved considerable optimisation of DNA extraction from the FTA cards to obtain sufficient quantity and quality of DNA. Then, sequence adaptors were connected to f1TLR9 to allow the enrichment of those fragments. In our first attempt to prepare the NGS libraries, we used a long primer that contains the sequencing adapter and gene-specific primer to amplify the f1TLR9 from FTA cards as well as on bovine commercially obtained control DNA. However, our data showed that this amplification was not successful. The experiment was repeated several times, using different optimisation procedures such as increasing and/or reducing cycles and extension times, changing reagents (Myfi ™ mix and MyTaq ™ HS mix) and temperature gradients.

To address this problem, we amplified the f1TLR9 before we integrated the adapter sequence overhang to the f1TLR9 amplicon. This result showed that the successful first-round PCR amplification reaction was successful in amplifying fragment 1 on exon 2 of bovine TLR9 of both sample and control (expected band size is 554bp, Table 2.2) (Fig. 5.2A). In the second-round PCR, 2 µl from the first-round PCR product were amplified with the adapter using “adapter+ f1TLR9” primer. Our analysis showed faint bands of the control and no band of the sample on the gel, which was visualised under UV light as shown in Fig. 5.2B. Then, in the third-round PCR, 2 µl from second-round PCR product were amplified using “adapter primer”. The successful PCR amplification appeared on the gel when we visualised under UV light; the result was shown as a

clear band that appeared to be the expected size of 621bp as shown in Fig. 5.2C. This was necessary to confirm that the Illumina overhang adapter had linked to the f1TLR9 amplicon. The PCR products were sent for Sanger sequencing after purification. Analysis of the sequence data was performed using the program SnapGene, which showed the Illumina adapter linked to our amplicon of interest (Fig. 5.3) and also, it indicates a successful exploration of this PCR technique. Ultimately, we integrated the Illumina index and confirmed the integration using Illumina index primers as shown in Fig. 5.2D.

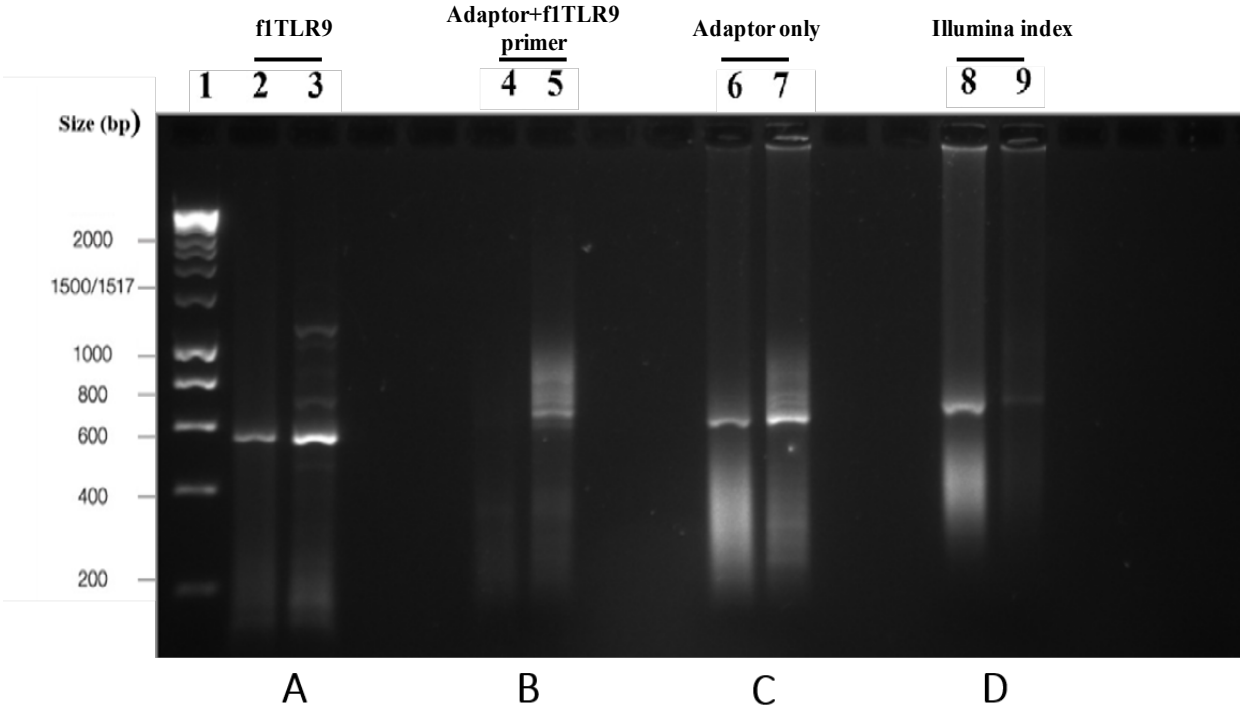


Figure 5. 2 The four PCR amplification rounds for f1TLR9. The first lane was the HyperLadder 1 kb marker (Bioline). **(A)** The second (sample) and third (control) lanes had clear bands, indicating the successfully amplified f1TLR9 using bovine DNA sample and bovine positive control DNA (554bp). **(B)** The fourth (sample) and fifth (control) lanes amplified the first-round PCR product with adapter sequences. **(C)** The sixth (sample) and seventh (control) lanes showed obvious bands, indicating a successful link to the Illumina overhang adapter sequences to the f1TLR9 amplicon using bovine DNA sample and bovine positive control DNA (621bp). **(D)** The eighth (sample) and ninth (control) lanes showed clear bands indicative of the Illumina Index being attached to the f1TLR9 amplicon using bovine DNA sample and bovine positive control DNA.

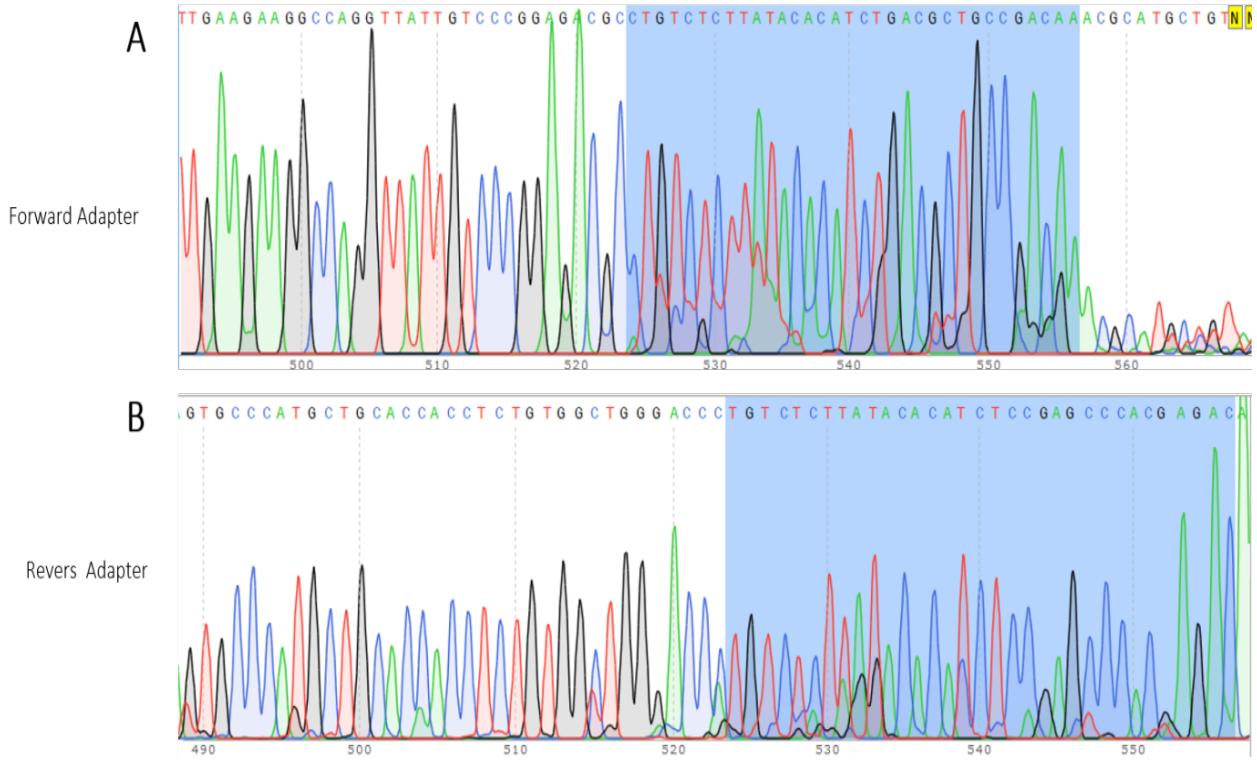


Figure 5. 3 Screenshot of Sanger sequencing data. Chromatograms of the nucleotide peaks of the f1TLR9 amplicon using forward (**A**) and reverse (**B**) primers (green for adenine, blue for cytosine, black for guanine and red for thymine). The blue shadow highlights the region of the Illumina overhang primer linked to the f1TLR9 amplicon. The SnapGene software was used to analyse the DNA sequencing data.

To determine the size of DNA amplicons and confirm purifying the adapter f1TLR9 amplicon away from free primers and primer-dimer species, we used the TapeStation technique for the fourth-round PCR after cleaning up the PCR product. Our analysis showed the poor quality of some DNA amplicons since there were unspecified products appearing that were under 500 bp as shown in Fig 5.4A. To address this problem, we applied serial cleaning between each reaction step. Our data showed that the serial cleaning of the PCR product efficiently removed all the non-specific bands, as illustrated in Fig. 5.4B and Fig. 5.4C.

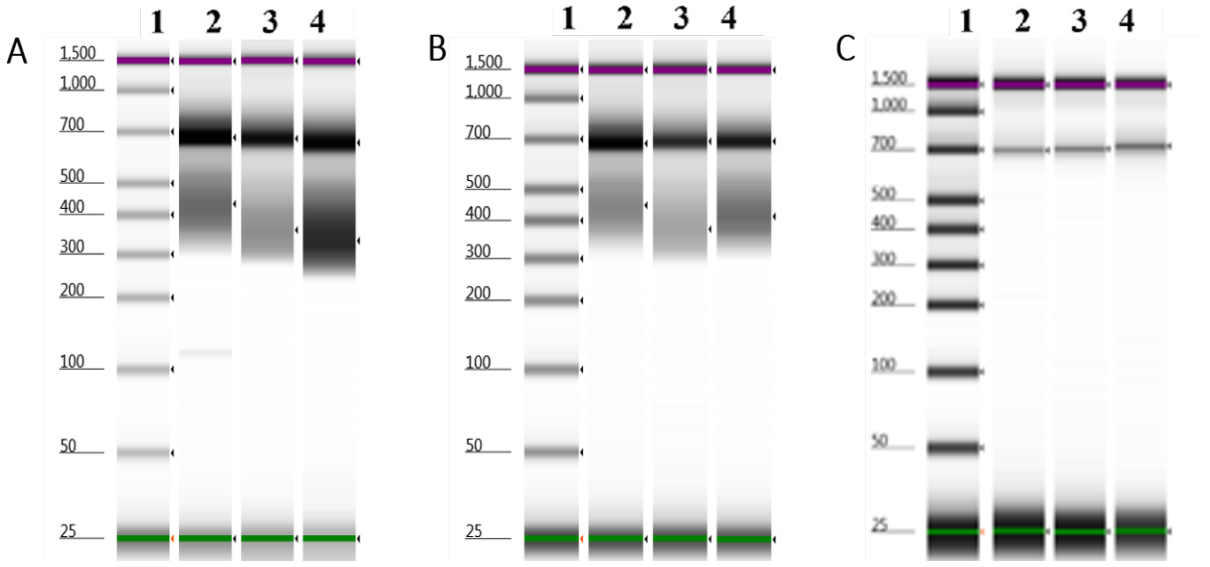


Figure 5. 4 TapeStation images for purification of the first library preparation. (A) The first lane was the HyperLadder 25 bp marker (Agilent Technologies). The unclean libraries in the second to the fourth lanes after the first cleaning using a 1.8:1 ratio of magnetic beads. **(B)** The first lane, HyperLadder 25 bp marker (Agilent Technologies). The second to the fourth lanes, after the libraries were cleaned twice using a 1.8:1 ratio of magnetic beads. As shown, the libraries were still not sufficiently pure. **(C)** The first lane, HyperLadder 25 bp marker (Agilent Technologies). The second to the fourth lanes refer to high-quality libraries produced after the third cleaning, using a 1.8:1 ratio of magnetic beads.

Although the libraries were very clean, the concentration of the libraries, determined by fluorimetry (Qubit), was very low < 4 nM for each library. To sum up, this quantification failed to prepare NGS libraries of sufficient quantity to run the NGS.

The failure of the first attempt to prepare NGS libraries led us to optimise the cleaning methods to select only the large DNA fragments. After repeating the four PCR rounds with a newly optimised cleaning technique, in which we used a 0.7:1 ratio of magnetic beads to PCR reaction instead of a 1.8:1 ratio, we found some unspecific DNA fragments that appeared on the TapeStation image Fig. 5.5A. Therefore, we further optimised the cleaning process by reducing the magnetic beads ratio to 0.5:1. Our data showed that libraries became cleaner, and the concentration of the DNA was good; however, a faint band of approximately 300 bp appeared as shown in Fig. 5.5B. The NGS reaction did not generate reads, and the run was terminated because of over-clustering. The flow cell chart that is useful for visualizing data per tile across the flow cell showed the over-

cluster densities across all tiles of the flow cell (Fig .7.1, appendices). To investigate this issue, we performed PCR for the pooling libraries using the index primers to see whether the small fragments would amplify. The result demonstrated that the second NGS pooling libraries were not pure, as shown in Fig. 5.6.

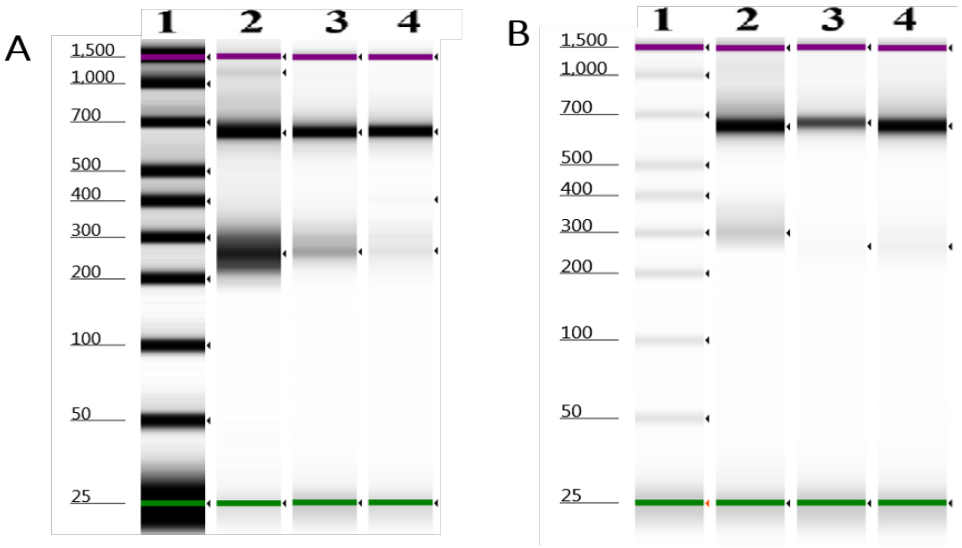


Figure 5. 5 TapeStation images for purification of the second NGS libraries. (A) The first lane was the HyperLadder 25 bp marker (Agilent Technologies). The second, third and fourth lanes show libraries with unspecific bands appear after cleaning using a 0.7:1 ratio of magnetic beads. **(B)** The first lane, HyperLadder 25 bp marker (Agilent Technologies). The second and fourth lanes showed reasonably pure NGS libraries after cleaning using a 0.5:1 ratio of magnetic beads with very faint bands around 300 bp.

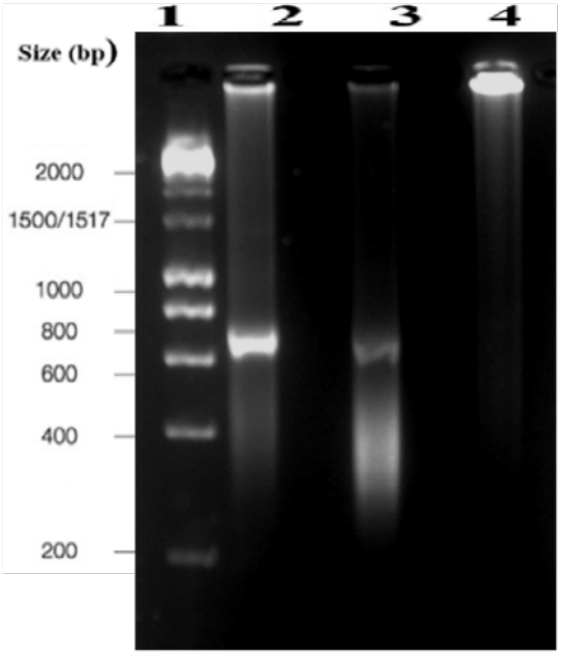


Figure 5. 6 Agarose gel image of the pooled libraries using the index primers. The first lane was the HyperLadder1 kb marker (Bioline). The second and third lanes showed the smearing of small DNA fragments after amplification using the Illumina Index primer, indicating that the second NGS pooling libraries were not pure. The fourth lane was the negative control (water).

5.3.2. NGS library preparation

After adjusting our understanding of an efficient library preparation protocol, we amplified the f1TLR9 before integrating the adapter and index sequence, as previously described. Our data showed a successful PCR amplification for the four PCR rounds, as illustrated in Fig. 5.7A. We then loaded all libraries to a 4% agarose gel and excised the f1TLR9 fragments to ensure that short fragments were separated from the long fragments. After that, the library was purified using a 0.5:1 ratio of magnetic beads and the resulting libraries appeared to be of good quality (Fig. 5.7B) and quantity (Table 7.2, appendices). For additional confirmation, the pooling was amplified using the Illumina Index primer to check whether small fragments were being amplified. The data showed no unspecific band as shown in Fig. 5.7C. After that, the pooled libraries were quantified using TapeStation to detect those that showed a clear band at 700 bp as shown in Fig. 5.7D. Before loading the libraries into the NGS instrument, we used quantitative PCR (qPCR) to determine the

quantity of targeted DNA fragments within the pooled libraries. The concentration of the pooled libraries was determined, based on the fragment size of 700 bp, to be 3.5 nM and 8 pM was added to the machine for analysis.

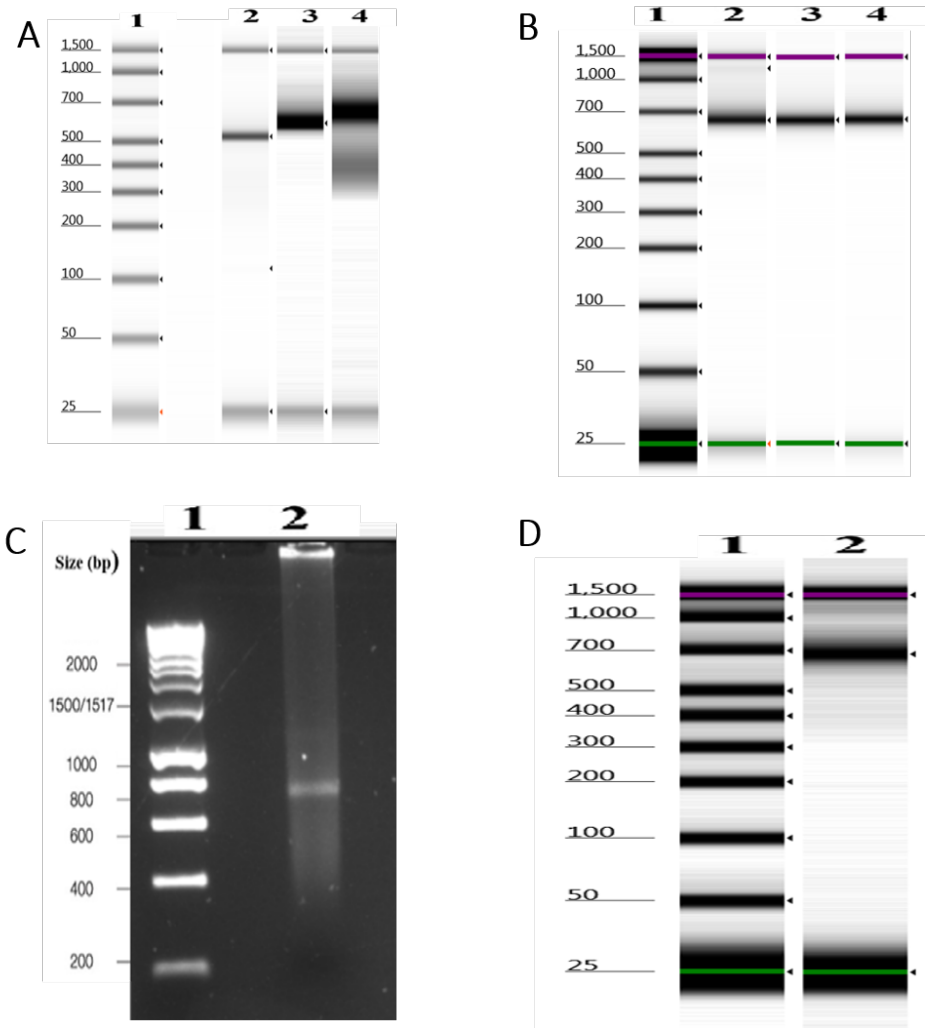


Figure 5. 7 Agarose gel and TapeStation images for the purification of NGS libraries. (A) The first lane was the HyperLadder 25 bp marker (Agilent Technologies). TapeStation images for successful PCR were shown in the second lane which contains the first-round PCR product; the third lane which contains the adapter primer PCR products and the fourth lane which contains the Illumina Index PCR products. **(B)** The first lane was the HyperLadder 25 bp marker (Agilent Technologies). The second, third and fourth lanes showed clean libraries after cleaning using a 0.5:1 ratio of magnetic beads. **(C)** An agarose gel image with the HyperLadder 1 kb marker (Bioloine) in the first lane and the second lane shows the amplification of pooled libraries using the Illumuna Index primer; note the absence of small fragments in this lane. **(D)** The first lane was the HyperLadder 25 bp marker (Agilent Technologies). The second lane was the TapeStation images for the pooled libraries.

5.3.3. Quality assessment of NGS throughput for f1TLR9

The successful sequencing of pooled libraries (80 cattle samples) generated an enormous volume of sequencing data (>7 million reads), which was stored in the BeasSpace Hub; the data was accessed and analysed by navigating to the BeasSpace Hub site. The data was obtained and saved in the FastQC program in order to evaluate the quality of the NGS data. Each sample produced two reads from both ends of a linear fragment (paired-end sequencing), in which the average read length per sequence was between 35 bp and 251 bp. As errors occur in Illumina sequencing, these were expressed in quality scores of the base call using the Phred score; therefore, we assessed the quality scores for each base calling from each sample. The Phred score was determined by using the quality scores equation, which is defined as $P = 10^{-Q/10}$, where Q is the quality score obtained from NGS data, represented as ASCII. As shown in Fig. 5.8A, the nucleotide calling bases on the X-axis corresponded to the Phred quality scores on the Y-axis, which were very high (~30); this indicated that the error probability of each base call was very low. Moreover, we evaluated the Phred quality scores per sequence, which was generated from each sample, as shown in Fig. 5.8B. Most of the sequences generated from tested samples had a high Phred quality score; indeed, the average score of 36.5 indicated that it was a high-quality read (Supplementary data Table 7.3).

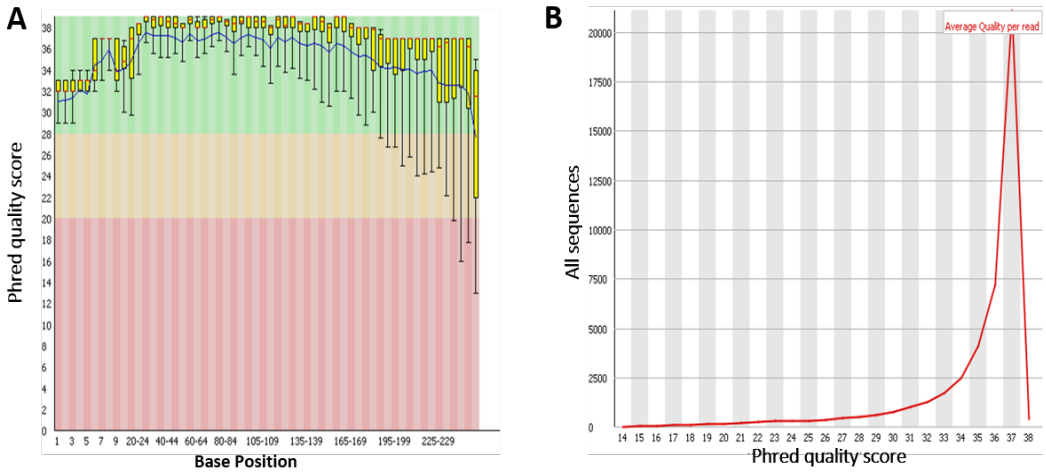


Figure 5. 8 Screenshot of Phred quality score graphs. (A) Phred quality score per base showing all bases located in the green area, indicating the Phred quality score was very high. **(B)** Phred quality score per sequence showing no sequence had a low Phred quality score, indicating that the sample quality was good.

On the other hand, we loaded 80 samples from 80 cattle with a unique index. Only 75 samples produced sequence data; five samples (23, 34, 38, 47 and 68) did not produce data (Fig. 5.9A) (Table7.3, appendices). The total number of reads from 75 samples were >7 million. The MiSeq 500-cycle format reads 250bp form each end of the DNA sequence, and as the TLR9 fragment is 554 bp this results in a gap of 54 bp in the central region of the target. Despite this gap, the percentage of high-quality reads was high at 97.0% (Fig. 5.9B). We also estimated the percentages of each read that did or did not align to the reference; as shown in Fig. 5.9C, we found that 76% of the yielded sequence reads from all cattle samples were aligned to the reference, while 24% of the sequence reads were not aligned (Supplementary data Table 7.3).

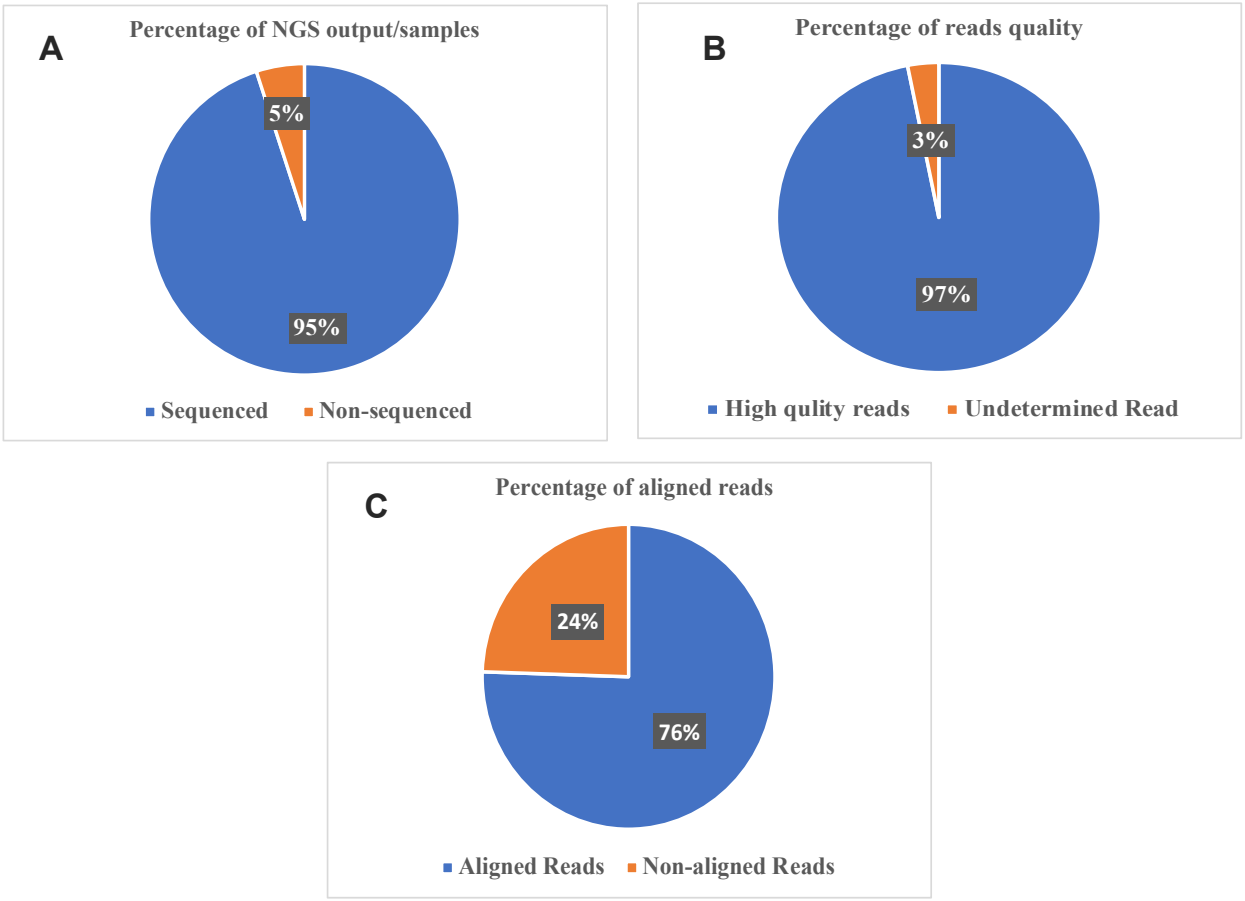


Figure 5. 9 Summary of the different NGS quality control statistics. (A) The percentage of sequencing output per sample showed that most of the samples generated a sequence, and only a few samples did not work. **(B)** The percentage of high-quality reads and low-quality reads from all NGS output. **(C)** The percentage of aligned reads and non-aligned reads to the reference, as generated from all samples.

5.3.4. Investigation of the variant calling in bovine f1TLR9 using NGS and Sanger data

To detect the variant base calling, we used The Variant Calling Format to align the tested sample sequences to the reference sequence, which was obtained from the NCBI database (KC174788.1). The Tablet software was used to display all reads for each sample that aligned to the reference and then we investigated whether the variants that we found by VCF were illustrated among the reads. As a result of analysing NGS data for 75 cattle samples, we identified two variants in region 2 of *B. indicus* f1TLR9. The first variant base was T/C, located on nucleotide number 4035 of the *B. indicus* TLR9 gene sequence. The second variant base was G/A, located on nucleotide number

4080 of the *B. indicus* TLR9 gene sequence (Table 5.1). According to the two variants, four haplotypes were discovered. Three of these haplotypes were frequently present, and one was very rare; the haplotypes were TG (H1), CA (H2), CG (H3) and TA (H4) (Fig. 5.10).

Out of the 75 cattle samples, 22 cows had H1/H1, which were the most common haplotypes (29.3%). H1/H4 were found in 15 cows (20.0%), H3/H3 were found in 12 cows (16.0%), and H1/H3 were found in 5 cows (6.7%). Some haplotypes, which were found very rarely among the cattle samples, were collected from only one group and named "Rare" haplotypes (Table 5.2).

Moreover, we examined the fidelity of the NGS data using the Sanger sequencing technique as a confirmatory method. We aligned all Sanger sequences for the tested samples with the reference sequence (KC174788.1). Out of 75 (93.8%) cattle samples that were investigated, 36 cows (48.0%) had H1/H1 haplotypes, 17 cows (22.7%) had H1/H4 haplotypes, 12 cows (16.0%) had H1/H3 haplotypes, and 10 cows (13.3%) had rare haplotypes (Table 5.3). Overall, the distribution of haplotypes between NGS and Sanger had a similar pattern.

To address the question of how the data from NGS and Sanger correspond, we assessed the degrees of reliability, correlation and validity of the NGS data with the Sanger data. The percentages of haplotype similarity between NGS and Sanger were >60.1% (H1/H1), 58.8% (H1/H4), 33.3% (H1/H3) and 70% (rare haplotype). The statistical analysis showed that the interobserver agreement for identifying the genetic variation between NGS and Sanger sequencing was moderate (kappa = 0.421). Overall, our data showed that there was a significant correlation of gene variation scattering between the Sanger and NGS methods (Kendall's tau-b=0.48; p-value, <0.001) (Table 5.4).

Table 5. 1 The identified SNPs with bovine f1TLR9.

SNP Allele	SNP position in *B. indicus* TLR9 gene KC174788.1	SNP position in *B. indicus* TLR9 mRNA KX138608.1	SNP position in *B. indicus* TLR9 Protein APQ40216.1	Amino Acid Change
T/C	4035	2468	822	NO
G/A	4080	2513	837	NO

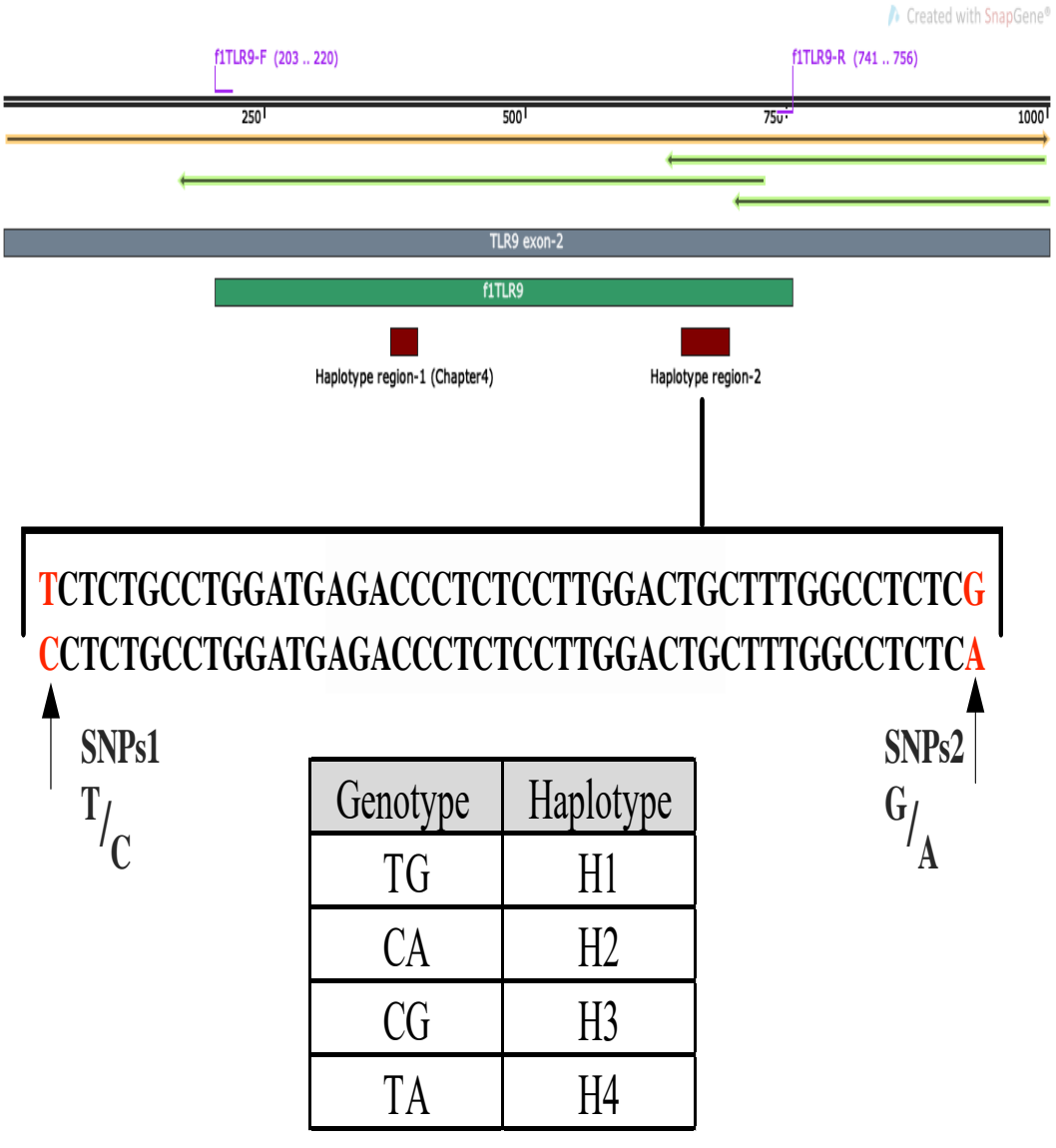


Figure 5.10 Identified haplotype mapping within f1TLR9 and list of the discovered haplotypes. The grey label indicates the exon 2 of the bovine TLR9 gene. The green label indicates the region of interest within exon 2 of the bovine TLR9 gene. The brown label 1 shows the region of a haplotype that was found by Sanger sequencing (Chapter 4) and the brown label 2 indicates

the region of a haplotype that was found by NGS (data within this Chapter). A magnified view of the DNA sequence to indicate the site of the SNPs within region 2 of the f1TLR9 fragment. The table summarises the haplotypes identified in region 2.

Table 5. 2 The frequency of identified haplotypes identified from NGS data in bovine f1TLR. The valid percent corrects for the data that is absent from five samples.

NGS	Frequency	% occurrence within total population	Valid %
H1/H1	22	27.5%	29.3%
H1/H4	15	18.8%	20.0%
H1/H3	5	6.3%	6.7%
H3/H3	12	15.0%	16.0%
Rare	21	26.3%	28.0%
Total	75	93.8%	100.0%
Missing	5	6.3%	
Total	80	100.0%	

Table 5. 3 The frequency of identified haplotypes identified from Sanger sequencing data in bovine f1TLR. The valid percent corrects for the data that is absent from five samples.

Sanger	Frequency	% occurrence within total population	Valid %
H1/H1	36	45.0%	48.0%
H1/H4	17	21.3%	22.7%
H1/H3	12	15.0%	16.0%
Rare	10	12.5%	13.3%
Total	75	93.8%	100.0%
Missing	5	6.3%	
Total	80	100.0%	

Table 5. 4 Evaluation of the correlation between NGS and Sanger sequencing methods.

NGS vs Sanger crosstab			Sanger				Total
			H1/H1	H1/H4	H1/H3	Rare	
NGS	H1/H1	Count	22	0	0	0	22
		% within Sanger	61.10%	0.00%	0.00%	0.00%	29.30%
	H1/H4	Count	3	10	0	2	15
		% within Sanger	8.30%	58.80%	0.00%	20.00%	20.00%
	H1/H3	Count	1	0	4	0	5
		% within Sanger	2.80%	0.00%	33.30%	0.00%	6.70%
	H3/H3	Count	4	1	6	1	12
		% within Sanger	11.10%	5.90%	50.00%	10.00%	16.00%
	Rare	Count	6	6	2	7	21
		% within Sanger	16.70%	35.30%	16.70%	70.00%	28.00%
Total		Count	36	17	12	10	75
		% within Sanger	100.00%	100.00%	100.00%	100.00%	100.00%

Interrater reliability (kappa)	0.412
Correlation coefficient (Kendall's tau-b)	0.482
P-value	<0.001

5.3.5. Association of genetic variants with *Trypanosoma* infection

The data on parasitic infection status was available for *B. indicus* cattle samples from previous studies (Ideozu, 2015; West, 2018). These studies used ITS-nested PCR and NGS for the detection of parasitic infections; these were found to be *Trypanosoma*, *Theileria*, *Sarcocystis* and *Eimeria*. In Chapter 4 of this study, we investigated the relationship between the genetic variation within CpG Island-1 of TLR9 using Sanger sequencing and susceptibility, or resistance, to parasitic infection; however, no significant correlation was found between genetic variants and parasitic infection (Chapter 4). Therefore, we applied the same analysis with the haplotypes identified by NGS to ascertain whether, or not, these might associate with susceptibility, or resistance to parasite infection.

In total, 80 tested samples were sequenced using NGS; however, only the 65 cattle, used previously for the detection of parasitic infections (Ideozu, 2015, West, 2018), were included in this analysis. Out of 65 cattle samples, 20 (30.8%) cows were found to have parasitic infections, and 45 (69.2%) cows were not infected with parasites. H1/H1 was found in 6 (35.3%) infected cows and 11 (46.7%) non-infected cows. Moreover, 25% of the infected cows and 75% of the non-infected cows were found to have H1/H4. In addition, H1/H3 was found in 20% of the infected cows and 80% of the non-infected cows. Furthermore, 41.7% of the infected cows and 58.3% of the non-infected cows had the H3/H3 haplotype. The rare haplotypes were found in 5 (26.3%) infected cows and 14 (73.7%) non-infected cows. We measured the correlation between haplotypes and susceptibility or resistance to parasitic infection. Our data did not show a significant correlation between haplotypes and parasite infection (p-value: 0.832; Pearson Chi-Square) (Table 5.4).

We tried to examine the association between susceptibility or resistance to parasitic infection and each haplotype that we observed for the Sanger sequencing results. This test used 65 cattle samples for which the parasite infections had been previously identified (Ideozu, 2015). In total, 20 (30.8%) out of 65 cows were infected with parasites (*Trypanosoma*, *Theileria*, *Sarcocystis* or *Eimeria*),

while 45 (69.2%) cows were not infected with parasites. Of the 30 cows that had H1/H1 haplotypes, 12 (40.0%) were positive for a parasitic infection, and 18 (60.0%) were negative. Of the 14 cows that had H1/H4 haplotypes, 4 (28.6%) were infected with parasites, and 10 (71.4%) were not infected. Of the 12 cows that had H1/H3 haplotypes, 3 (25.0%) had parasites, and 9 (75.0%) were not infected. Only 1 (11.1%) out of 9 cows with rare haplotypes had a parasitic infection and 8 (88.9%) of these cows were not infected with parasites. H3/H3 haplotypes were not found in the Sanger sequencing results. Therefore, the relationship between haplotype and susceptibility or resistance to parasitic infection was investigated. Our data showed that neither haplotype had a significant association with parasitic infection (p-value: 0.384; Pearson Chi-Square) (Table 5.5).

Overall, the correlation coefficient analysis was performed between haplotypes and parasitic infection for the NGS and Sanger sequencing results. The data analysis showed that the correlation coefficient to parasitic infection was approximately similar between the NGS haplotypes (R=0.035) and the Sanger sequencing haplotypes (R=0.195).

Table 5. 5 Evaluation of the association between parasitic infection and haplotype patterns in bovine f1TLR9 using NGS and Sanger sequencing.

	Haplotype	Count/ % within NGS	Infected	Non-infected	Total	Sig	Haplotype	Count/ % within Sanger	Infected	Non-infected	Total	Sig	
NGS	H1/H1	Count	6	11	17	No	H1/H1	Count	12	18	30	No	Sanger
		%	35.3%	64.7%	100.0%			%	40.0%	60.0%	100.0%		
	H1/H4	Count	3	9	12	No	H1/H4	Count	4	10	14	No	
		%	25.0%	75.0%	100.0%			%	28.6%	71.4%	100.0%		
	H1/H3	Count	1	4	5	No	H1/H3	Count	3	9	12	No	
		%	20.0%	80.0%	100.0%			%	25.0%	75.0%	100.0%		
	H3/H3	Count	5	7	12	No							
		%	41.7%	58.3%	100.0%								
	Rare	Count	5	14	19	No	Rare	Count	1	8	9	No	
		%	26.3%	73.7%	100.0%			%	11.1%	88.9%	100.0%		
	Total	Count	20	45	65	No	Total	Count	20	45	65	No	
		%	30.8%	69.2%	100.0%			%	30.8%	69.2%	100.0%		

Test	NGS	Sanger
Correlation coefficient (Kendall's tau-b)	0.035	0.195
P-value	0.832	0.384

5.4.Discussion

In this study, we established a method to sequence genomic material from bovine blood stored on FTA cards for *B. indicus* bulls from Ahoada, a town in southern Nigeria, using the NGS technique. We prepared the NGS library and assessed the quality of sequenced data before the downstream analysis. We also assessed the impact of the genetic variation obtained from the NGS data to the susceptibility or resistance of parasitic infection. Finally, we evaluated the reliability of NGS to the Sanger sequencing method in determining variant calling specificity and sensitivity.

Initially, a Harris micro punch was used to obtain two 2 mm punches from the FTA cards and DNA was extracted using the DNeasy Blood & Tissue kit. The extracted DNA was amplified using primer pairs targeting bovine f1TLR9, a region in exon 2 of the bovine TLR9 gene, in the first round of PCR. As part of the library preparation, three rounds of PCR were carried out in order to integrate the Illumina adapter to the bovine f1TLR9 template. We performed the final PCR round (the fourth) to attach Illumina indices to each cattle sample. The PCR amplicon (700 bp) was then purified to remove short fragments, such as primer dimers and off-target amplicons, which may affect the target sequencing yield (Head et al., 2014, Quail et al., 2012). The primer-dimers, which can be formed by chance base pairing and extension of each primer using the other primer as a template during PCR (Chou et al., 1992, Poritz and Ririe, 2014), can form clusters and consume valuable space on the flow cell without generating any useful data. Thus, we typically used either magnetic beads or agarose gels to purify DNA of the expected size.

The solid-phase reversible immobilisation (SPRI) paramagnetic beads method, initially developed by DeAngelis et al. (1995), is a good method for purifying pure DNA based on size selection for preparing NGS libraries (Lundin et al., 2010, Stranneheim et al., 2011). Therefore, we used SPRI magnetic beads (AMPure XP beads) to purify our target amplicon (f1TLR9) and eliminate primer-dimers and any non-specific amplicons in each step of library preparation. To evaluate the

purification quality, we used the TapeStation system, which is a very sensitive system that displays and analyses the size distribution of DNA fragments (McEvoy et al., 2020, Sato et al., 2019). According to the manufacturer's instructions, we purified our target amplicon using a 1.8:1 ratio of AMPure XP-to-PCR volume. However, this dilution factor did not efficiently purify our target amplicon, which was ~700 bp, as we found another non-specific amplicon of ~300 bp. The high-ratio magnetic beads increased the chance that short fragments, such as primer dimers, would attach to the beads and decrease the sensitivity. Moreover, as the result of repeated purification at each step of library preparation, we lost most of the target amplicon.

Therefore, we established different dilution factors to enhance the purification quality of the target amplicon from other library preparations. This was done according to previous studies, which developed convenient and reliable protocols to accurately determine the critical concentration of SPRI magnetic beads for targeting and precipitating the interesting DNA fragments (Borgström et al., 2011, He et al., 2013). We used a dilution factor of 0.7:1 of AMPure XP-to-PCR reaction after each PCR round to improve the purification of f1TLR9. The TapeStation system showed that the purification was improved in some libraries; however, a few libraries had faint bands of non-target amplicons. Subsequently, our first trial of NGS sequencing failed because of the over-clustering of DNA fragments on the flow cell due to insufficient purification.

The BluePippin Prep method is beneficial for narrowing the size selection (Quail et al., 2012). We tried to use this method in our NGS libraries, but this technique requires extra instrumentation that was not available at our facility. Therefore, we established a technique to increase the efficiency of library purification by combining SPRI magnetic beads flowing by agarose gel electrophoresis separation. Briefly, we purified the f1TLR9 with AMPure XP (at a ratio of 0.5:1), followed by loading each library in 4% agarose gel (Lee et al., 2012). Then, we used qPCR to quantify the libraries concentration instead of the TapeStation system because it is highly sensitive and accurate

for quantifying specific DNA molecules (Hussing et al., 2018). Consequently, this method was successfully implemented to generate bovine f1TLR9 sequences for most of the samples.

We assessed the quality of sequencing data to provide information about the accuracy of our library preparation through the yield of sequence coverage, read length, and base quality. The accuracy of sequencing data is essential for various downstream analyses, such as sequence assembly, read alignment and single nucleotide polymorphism identification. In contrast, low-quality scores can affect the accuracy of identifying simple nucleotide variation calls, including single nucleotide polymorphisms and small insertions and deletions (Nakamura et al., 2011). We used the FastQC program, which is a popular tool for performing quality checks on raw NGS data (Andrews, 2010). The Phred quality score, a logarithmic error probability, expresses the base call-in quality scores. Overall, our data showed that the quality of the DNA sequence was high. The Phred quality scores per nucleotide were high (~30), which refers to a base with a 1-in-1,000 probability of being incorrect or 99.9% accuracy (Ewing et al., 1998).

The most common source of concern for quality is the DNA library, which can also be directly reflected by the number of reads sequenced. Although the amount of DNA in each library was strictly measured before the sequencing step, it was difficult to equalise the DNA amount and molecular size of the libraries. This is reflected in the variability in the number of reads in each sample. On the other hand, we observed some low-quality reads, such as undetermined reads, which were produced from incorrect indexes specified in the sample sheet, missing index sequences (i.e. PhiX reads, which have no index) and sequencing errors in the index read.

We postulated that genetic variations in TLR9 in our region of interest (region 2) are associated with susceptibility or resistance to parasitic infection. The previous results showed that SNPs in TLR9 and TLR4 are associated with a viral infection, which increases the risk of cervical cancer (Pandey et al., 2019). However, our data showed that the genetic variations detected by both methods were not significantly correlated with parasitic infection. We believe our region of

interest might have little or no impact on the susceptibility or resistance to parasitic infection because it has little influence on biological functions. Moreover, the polymorphisms that we found were synonymous, did not change the amino acid sequence; as such, they would not result in a functional change to the protein.

Moreover, the polymorphisms did not locate in a CpG site, which could not cause an epigenetic change if the C residue was capable of being methylated. This result gave us an indication of the similar outcomes of both methods. Therefore, we studied the reliability of NGS and Sanger sequencing to detect the genetic variation of our region of interest in exon2 of the bovine TLR9 gene.

Sanger sequencing has been used as a gold-standard method for detecting DNA mutation in a single gene; however, Sanger sequencing has low sensitivity and is unable to perform a parallel investigation of multiple targets, such as somatic cancer mutations. On the other hand, NGS allows determination of a broad spectrum of mutations such as SNPs, insertions, deletions and rare variations. A previous study showed that using NGS is more sensitive than Sanger sequencing to detect *PIK3CA* gene mutations, which is associated with hepatocellular carcinoma (Arsenic et al., 2015). However, because we sequenced a single gene, we used Sanger sequencing as a standard or control to validate NGS data sequencing. Our data showed intermediate reliability between the NGS and Sanger sequencing data. We believe that due to the multiple PCR rounds during library preparations, potential exists for some errors in NGS base calling due to polymerase misincorporation, structure-induced template-switching, PCR-mediated recombination and DNA damage (Potapov and Ong, 2017). Indeed, Oyola et al. (2012) found the PCR-free NGS libraries preparation to be significantly better than that dependent upon PCR (Oyola et al., 2012). Moreover, our region of interest has CpG island and that may be the factor behind the errors or discrepancies of sequencing. A previous study also showed that CpG regions were significantly related to the higher false-negative rate of mutation calls, which highlights that CpG regions might be a

significant factor behind inconsistency and/or false-negative errors in NGS analysis (Kim et al., 2019).

Previous studies have also examined the validity of the NGS data by comparing it to the data from a gold-standard method such as Sanger. For example, NGS has been reported to successfully detect pathogenic SNPs and insertion/deletion mutations ranging in size from 1 to 18 bp that were previously characterised by Sanger sequencing (Yohe et al., 2015). Therefore, NGS has made much progress in its use in clinical settings to diagnose the genetic makeup that causes genetic disease (Di Resta and Ferrari, 2018). Moreover, NGS has recently been used in genetic testing, such as cell-free DNA in prenatal genetic testing and human leukocyte antigen (HLA) typing (Ordulu et al., 2016, Weimer et al., 2016). In this study, we used both techniques (Sanger sequencing and NGS) to sequence exon two of bovine TLR9. Using Sanger sequencing is time consuming and highly costly, depending on the number of samples. In contrast, for NGS, all samples are simultaneously processed and this improves efficiency and potentially saves time and reduces costs.

In conclusion, we established a protocol to investigate the genetic variation of the TLR9 gene from FTA cards using the NGS technique. The quality of NGS data was high and included sequence coverage, read length and base quality. We identified a novel genetic variation located between 4035 bp and 4080 bp of the *B. indicus* TLR9 gene sequence. Moreover, we used Sanger sequencing as a confirmatory method to validate the NGS data. This validation was measured by calculating the extent of the agreement, which showed intermediate reliability between the Sanger and the NGS data. These SNPs are not associated with parasitic infection.

CHAPTER SIX

General Discussion

DNA sequencing is considered a gold-standard technology for biotechnological approaches, generating data concerning the entire genome and facilitating genetic variation analysis for a wide variety of organisms from mammals to pathogens. In our study, we established a methodological approach to study the genetic variation of TLR9 in a mammalian host using both Sanger and NGS (Fig. 6.1). TLR9 is a part of the innate immune system that plays an important role in recognising CpG motifs in microbes such as bacteria, fungi and parasites. It has been hypothesised that TLR9 polymorphisms play a role in susceptibility or resistance of the host to an infectious disease (Mukherjee et al., 2019). Consequently, we analysed the correlation of identified SNPs in TLR9 with parasitic infection previously identified in Nigerian cattle (Ideozu, 2015). In Chapter Three, we established and compared the different methods for extracting genomic material from Whatman FTA cards. In our previous study (Ideozu, 2015), we collected blood samples from Nigerian cattle on Whatman FTA cards to investigate the prevalence of parasitic infection among tested samples.

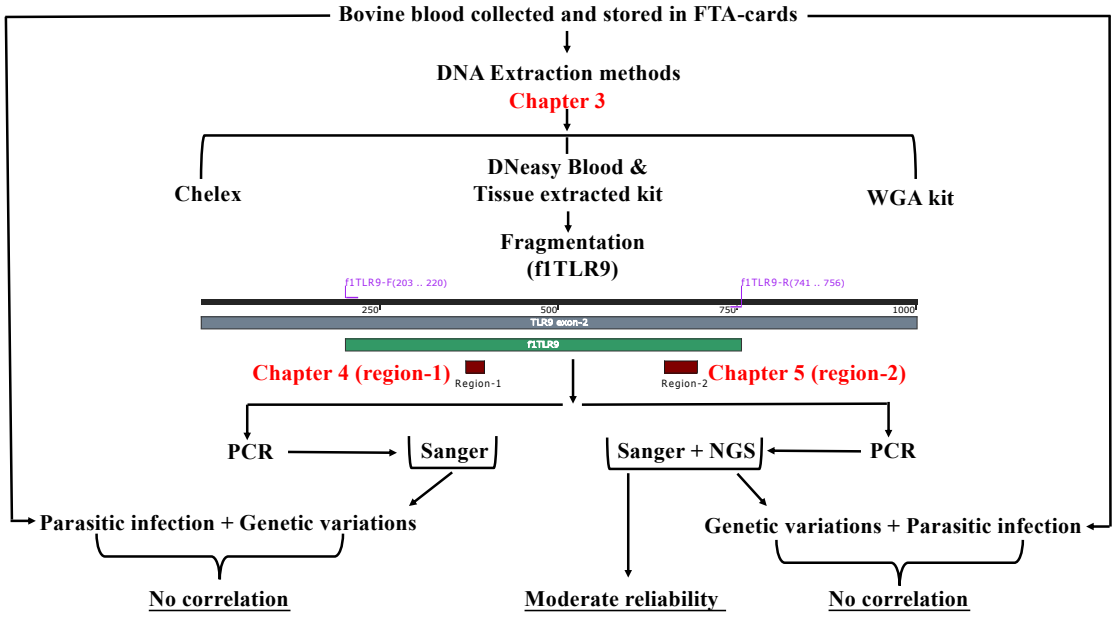


Figure 6. 1 Schematic illustration of the main objectives of this project. Blood spots were collected from a previous project (Ideozu, 2015) and three methods were then used to extract DNA from FTA cards. The DNeasy Blood & Tissue extraction kit was used for downstream applications. Initially we studied region 1 of f1TLR9 using Sanger sequencing to investigate SNPs and any correlation that may exist with parasite infection profiles in the cattle. Similarly, we studied the presence of SNPs in region-2 using both Sanger and NGS. In this part of the project,

we studied the correlation of the genetic variation with parasitic infection; moreover, we investigated the reliability relationship between the Sanger and NGS methods.

FTA cards are the most reliable method for DNA storage and extraction. FTA cards can store samples for up to a year and be transported at room temperature in areas where cold storage is not available or when the transport of biohazardous agents is hindered by international regulations. FTA cards are becoming increasingly popular for the rapid collection and archiving of a large number of samples such as blood, saliva, sperm and stool (Jignal et al., 2014, Lalani et al., 2018, Stangegaard et al., 2013). FTA cards are chemically treated so that they lyse cell membranes on contact, denature proteins and inhibit detritivores during drying to ensure the safe handling of the card without risk of biohazards and pathogen, while nucleic acids are physically entrapped, immobilised and stabilised for storage at room temperature (Ahmed et al., 2011, Stangegaard et al., 2013). FTA cards also protect nucleic acids from nucleases, oxidation, UV damage and microbial and fungal attacks. FTA card-based technologies have proven a great success for genetic analysis (McClure et al., 2009, Saieg et al., 2012) and have been widely used in forensic laboratories that need to process a great number and wide variety of crime scene samples for DNA analysis (Rahikainen et al., 2016). Moreover, Whatman FTA cards have proven a great success for detecting bacteria (Almeida et al., 2018, Lalani et al., 2018), viruses (Jóźwiak et al., 2016, Montmayeur et al., 2017) and protozoa (Adams et al., 2006, Ahmed et al., 2011, Hide et al., 2003, Ahmed et al., 2013).

Our analysis showed that using the DNeasy Blood & Tissue extraction kit is the most efficient method of the approaches evaluated in this project (Figure 6.1). Previous data showed that the DNeasy Blood & Tissue extraction kit was used to extract pure DNA from tissue and Whatman FTA cards (Moura et al., 2016, Ward et al., 2019). A pure genomic DNA extract is best obtained using the DNeasy Blood & Tissue extraction kit because its lysis buffer may provide optimal DNA purification. Consequently, we tried to amplify and sequence the whole gene of TLR9; however,

this goal was not achieved due to the large DNA template >4kb, as shown previously (Cha and Thilly, 1993). Moreover, it is hard to trace the genetic variation of a large DNA fragment using Sanger sequencing; therefore, we focused on the CpG islands, which are present within ~1 kb (3588–4717) of the TLR9 gene. Furthermore, CpG islands have a high mutational rate that is linked to genetic diseases (Cooper et al., 2010). We then PCR amplified multiple fragments in order to have a high resolution of the DNA sequence, as shown in Fig. 3.8. Finally, we used the first fragment, labelled f1TLR9, which contains the CpG island-1 and a functional domain; leucine-rich repeats-24 (LRR24), as a region of interest for further investigation.

In Chapter Four, we investigated the correlation of the genetic variation in TLR9 with parasitic infection. According to our study, we found that the most common parasitic infection in Nigerian cattle was *Trypanosoma*. *Trypanosoma* belongs to the protozoan class and has a single flagellum. This parasitic infection is transmitted in bovines by the bite of the tsetse fly called *Glossina* spp. (Brun et al., 2010). The impact of this parasite on livestock is relatively mild for some types of wild animals, while it can be severe and often fatal in others (Chitanga et al., 2013, Yaro et al., 2016).

There are different methods for detecting parasitic infection, such as the molecular, microscopical and serological approaches. In our lab, we used a molecular assay called internal transcribed spacer (ITS)-nested PCR and next-generation sequencing (NGS) to identify the parasitic infection. ITS-nested PCR is an effective and sensitive method for detecting species and subspecies of parasitic infection, such as *Trypanosoma* in mammalian hosts (Cox et al., 2005; Cox et al., 2010; Wastling and Welburn, 2011; Ideozu et al., 2015). Our study also re-evaluated the infection rate from FTA cards using either ITS-nested PCR or NGS methods. The prevalence of trypanosomes in this study was shown to be 28.6% (95% CI: 22%–32.7%) among 70 samples. Moreover, we also identified other common zoonotic and enzootic parasites such as *Theileria*, *Eimeria* and *Sarcocystis* in cattle (West, 2018). On the other hand, for many years, the microscopical method has been used to detect

parasitic infection from blood, tissue, stool, lymph node and bone marrow (Ndao, 2009). However, this method is time-consuming and requires high technical expertise. Moreover, it requires a second independent reader of the slide to increase the accuracy of the diagnosis. The other method for detecting parasites is serological assay. The serological assay is used to detect the immune response to a parasitic infection when the biological sample or tissue specimen is unavailable (Ndao, 2009).

The immune response to a parasitic infection such as *Trypanosoma* is based on molecular recognition of cytosine-phosphate-guanosine (CpG) and cellular glycoproteins. For instance, immune cells, such as macrophages and dendritic cells (DC), can recognise the cellular or molecular motifs by pattern recognition receptors (PRRs) such TLR2 and TLR9, respectively (Tarleton, 2007). In response to parasite recognition, immune cells induce proinflammatory cytokines such as Interleukin-12 (IL-12) and Type-1 interferon (INF-1) to amplify the immune response and activate the adaptive immune response. TLR has been reported to play a crucial role in activation of adaptive immune response (Bafica et al., 2006). Therefore, we were interested in studying the TLR9 polymorphisms and their relationship to parasitic infection.

In our report, we examined the genetic variants of TLR9 in association with the intracellular parasitic infection using the Sanger sequencing method. By studying the genetic variants, it is possible to understand the molecular bases of functional impairment in biological components, such as immunological molecules, which make us more susceptible to infectious diseases (Chapman and Hill, 2012). For instance, a collection of evidence has been reviewed for studying the correlation of single-nucleotide polymorphisms of TLRs and susceptibility to infectious diseases in humans (Skevaki et al., 2015). Some single-nucleotide polymorphisms (SNPs), called nonsynonymous polymorphisms, result in a change of the amino acid sequence and subsequently affect the protein structure. For example, nonsynonymous polymorphisms in TLR2 have an impaired response, which is shown as a significantly low serum level of IL-12 to *Mycobacterium*

leprae (Kang et al., 2002). On the other hand, the synonymous type of SNP does not result in a change in the amino acid sequence. This type of SNP may not affect the protein structure or function; however, it may be linked to the progress of certain diseases. For instance, synonymous polymorphisms in TLR9 have been shown to play a role in the rapid progression of human immunodeficiency virus-1 (HIV-1) infection (Bochud et al., 2007). In our study, we identified one nonsynonymous SNP and seven synonymous SNPs within TLR9; none of which showed a correlation to parasitic infection.

Consequently, we evaluated the linking of parasitic infection in Nigerian cattle with the allelic homozygosity or heterozygosity of the TLR9 gene. It is believed that the heterozygosity of an immune-related gene such as the HLA loci might be more resistant to infectious diseases than homozygotes due to a wider repertoire to present antigens. Indeed, the heterozygous HLA class I has been reported to be associated with lower viral load than homozygous HLA (Jeffery et al., 2000). However, according to our data and the literature, it appears that heterozygotic status in the TLR gene does not play a role in resistance to infectious diseases. It has been reported that heterozygotic or homozygotic SNPs in TLR2 at position 2258 (G>A) were associated with cytomegalovirus disease and high viral load in liver transplant recipients (Kang et al., 2012, Kijpittayarit et al., 2007). Furthermore, heterozygote SNPs of TLR9 have been reported to increase the risk of cytomegalovirus disease in children (Paradowska et al., 2016).

We then investigated possible epigenetic mechanisms of DNA methylation in relation to parasitic infection. DNA methylation is a process of adding a methyl group by DNA methyltransferase (DNMT) to the cytosine residues in CpG islands (Kulis and Esteller, 2010). In a mammalian host, DNA methylation is a common mechanism to 'switch off' genes (Moore et al., 2013). High methylation in the TLR4 promoter was reported to be associated with diabetic foot ulcers (Singh et al., 2013).

Conversely, another study showed that low methylation in TLR9 is associated with diopathic pulmonary fibrosis progression (Hogaboam et al., 2012). However, the role of methylation in TLR function is not well understood. In addition, the impact of polymorphism in CpG islands with regard to resistance or susceptibility of parasitic infection is still limited. Therefore, we studied the impact of DNA polymorphism on the cytosine-guanine (CG) dinucleotide, a potential site for DNA methylation, within the CpG island of TLR9. Our data showed that there are two SNPs in CG dinucleotides within our region of interest (f1TLR9) and these were found to have no association with resistance or susceptibility to parasitic infection.

Future studies should consider the potential effects of TLR9 expression by quantifying mRNA from FTA cards. RNA could be efficiently isolated from FTA cards for application in clinical and forensic genetics (Skonieczna et al., 2016). This experiment, if possible, could be used to investigate TLR9 gene expression in mutated and unmutated CG locations in CpG islands. Generally, the expression profile of TLRs is tissue-specific, and the role of DNA methylation is not well understood. A previous study showed that DNA methylation significantly reduced the level of TLR2 and TLR9 in the decidual tissue of women with spontaneous preterm labour (Walsh et al., 2017).

In chapter five, we used the NGS sequencing method to sequence our region of interest within the TLR9 gene. In research, NGS has been used as a tool for *de novo* genome sequencing, DNA sequencing, RNA or transcriptomic sequencing and epigenomics. For example, NGS technology was used to study and characterise the genomic material of SARS-CoV-2, associated with severe human respiratory diseases in Wuhan, China (Wu et al., 2020). In a clinical setting, NGS has been used to diagnose genetic diseases and tissue typing, such as HLA typing (Ordulu et al., 2016, Weimer et al., 2016).

Sanger sequencing, which is based on the same principle as NGS, is considered an ideal method to sequence short DNA sequence. However, it has limited sensitivity or power to characterise

mosaic mutations. We believed that sequencing of a single gene, TLR9 in our study, by the Sanger method is the gold standard; therefore, we used it as a control to confirm the data obtained from the NGS method. Our data showed that genetic variations found in the NGS and Sanger methods have intermediate reliability. The validation of data from NGS was evaluated previously using Sanger sequencing to detect SNPs and insertion/deletion mutations (Yohe et al., 2015).

Moreover, we used both data from NGS and Sanger sequencing to correlate genetic variation in the bovine TLR9 gene with parasitic infection. However, our data did not find a significant correlation between SNPs found using either method with parasitic infection. As we discussed previously, we postulate that our region of interest is likely not to have a significant impact on the biological function of TLR9. In addition, another limitation of this project was that the sample size of 70 cattle may not be sufficient to determine the correlation of TLR9 polymorphisms with parasitic infection. However, all the SNPs identified from our samples are novel according to our knowledge of the literature on bovine TLR9 polymorphisms.

In conclusion, using FTA cards is a valuable tool to study the genetic material of both blood-borne diseases such as African trypanosomes and the genetic makeup of the mammalian host. In our study, the DNeasy Blood & Tissue extraction kit was an efficient method to extract mammalian genomic material from FTA cards. Both the Sanger and the NGS methods could be used to determine the genetic variation of host genomic DNA from FTA cards.

Future work

We demonstrated the methodological concepts of DNA extraction from FTA cards and sequencing part of the mammalian host gene, TLR9, using Sanger sequencing and NGS. We are interested to improve these methods, for instance, to minimise the number of PCRs which therefore could enhance the resolution and purity of sequencing especially with NGS.

On the other hand, our project focused on a small fragment of TLR9 with the aim being to have a high resolution of both Sanger sequencing and NGS. Our future implication is to investigate other immune relevant genes (e.g. other TLRs and NLRs). However, with respect to TLR9, we are interested to investigate SNPs on other regions of the gene and especially those that encode regions predicted to influence biological function (e.g. LRR domains). Another consideration is to evaluate the biological functions of mutated TLR9 (e.g. LRRs regions) using a recombinant protein technology. This notion could be applied to other immune-relevant such as other TLRs and NLRs. We believed that genetic differences in immunological status probably involve SNPs in many genes; therefore, we need to scale up to large numbers of cattle samples to investigate the relationship between susceptibility or resistance to parasitic infection. Consequently, we also wish to be able to sequence multiple immunological genes from lots of animals with known disease infections since this will enable us to link infection susceptibility with individual SNPs. In our project, we draw a baseline method to use the FTA cards to collect and store lots of samples, and NGS to sequence lots of genes (i.e., immunological genes) to study the association with susceptibility or resistance to infectious diseases.

CHAPTER SEVEN

Appendices

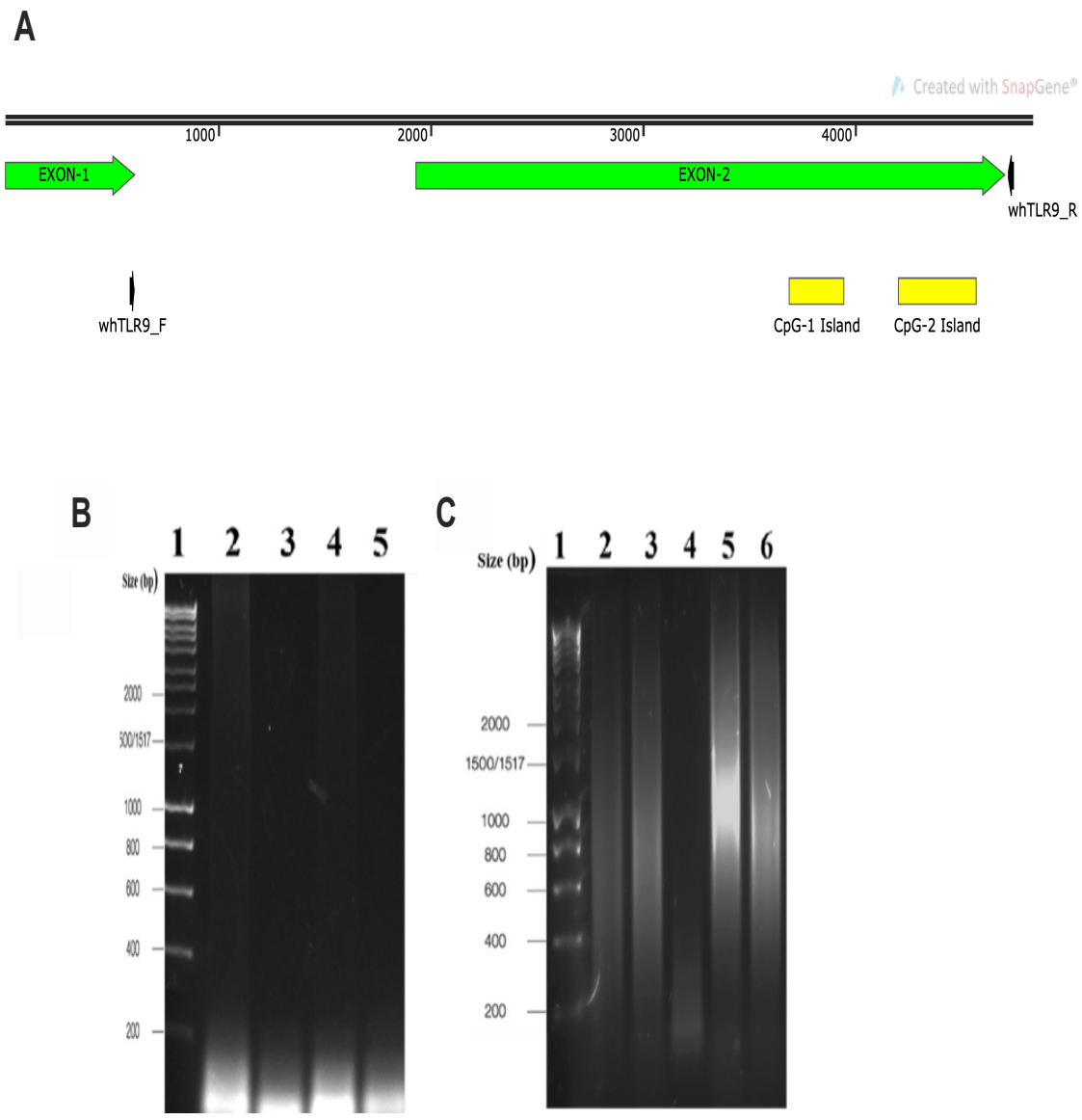


Figure 7. 1 Figure PCR amplification of the bovine TLR9 gene. (**A**) Designed primers targeting the bovine TLR9 gene. The black line represents the bovine TLR9 gene (4156 bp). The green arrows indicate exons 1 and 2, and the yellow labels indicate CpG islands 1 and 2 in the bovine TLR9 gene. The black arrows indicate the locations of the designed forward and reverse primers. (**B**) Gel picture of unsuccessful bovine TLR9 gene amplification using different primer concentrations with positive control bovine DNA. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: 0.2 pmole/µl. Lane 3: 0.4 pmole/µl. Lane 4: 0.6 pmole/µl. Lane 5: negative control (water). (**C**) Gel picture of unsuccessful bovine TLR9 gene amplification using different annealing temperatures with positive control bovine DNA. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: 55 ºC. Lane 3: 58 ºC. Lane 4: 60 ºC. Lane 5: 61 ºC. Lane 6: negative control (water).

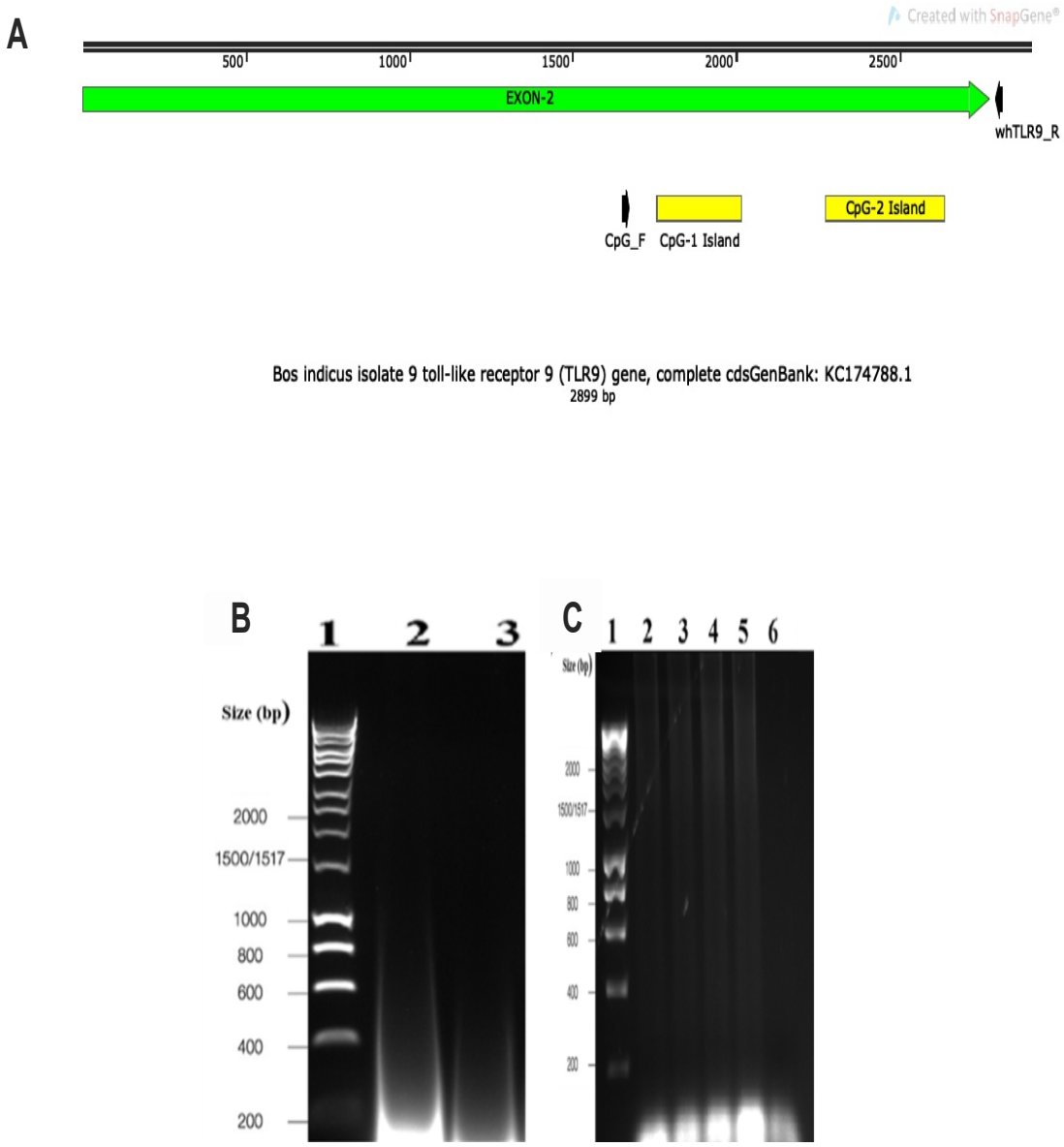


Figure 7. 2 Amplification of the bovine TLR9 gene using hemi-nested PCR. (**A**) Designed primers targeted CpG islands. The black line represents the amplified bovine TLR9 gene. The green label indicates the exon 2 of the bovine TLR9 gene. The black arrows show the locations of the designed forward and reverse hemi-nested PCR primers. The yellow labels indicate the locations of CpG islands 1 and 2 in the bovine TLR9 gene. (**B**) Gel picture of hemi-nested PCR. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: positive control (commercial DNA) using MyTaq™ Hs mix. Lane 3: negative control (water). (**C**) Gel picture of hemi-nested PCR. Lane 1: hyper ladder 1 kb marker (Bioline). Lane 2: annealing temperature of 55 ºC. Lane 3: 58 ºC. Lane 4: 61 ºC. Lane 5: 63 ºC. Lane 6: negative control (water).

Table 7. 1 The concentration of extracted DNA using the DNeasy Blood & Tissue extraction kit (QIAGEN) Columns 1 and 3 present the number of cattle samples. Columns 2 and 4 present the concentration of DNA that extracting from FTA cards using DNeasy Blood & Tissue extraction kit (QIAGEN).

Sample No.	Conc ng/ µl	Sample No.	Conc ng/ µl	Sample No.	Conc ng/ µll	Sample No.	Conc ng/ µl
1	29	21	13.2	41	14.8	61	19.2
2	10.5	22	12.4	42	14	62	16.4
3	22.5	23	21.5	43	12.3	63	24.8
4	14.8	24	28	44	20.6	64	18.1
5	17.2	25	19.1	45	19.2	65	22.5
6	19.3	26	23.2	46	10.2	66	11.3
7	13.6	27	22.3	47	23.4	67	30
8	14.9	28	20.4	48	10.3	68	15.9
9	22.8	29	15.5	49	12.2	69	21,8
10	13.2	30	16.6	50	15.9	70	14.2
11	11.3	31	24.1	51	15.6	71	19.7
12	20.3	32	12.8	52	16.2	72	16.6
13	12	33	24	53	14.6	73	22.6
14	13.7	34	12.8	54	13.5	74	24.3
15	16.5	35	15.5	55	20.6	75	10.3
16	15.1	36	17.9	56	10.6	76	22.7
17	14.1	37	11.3	57	12.8	77	22.8
18	18.7	38	20.,3	58	17.5	78	11.3
19	13.3	39	17.5	59	21.4	79	18.3
20	20.4	40	11.3	60	15.9	80	15.8

Table 7.2 The concentration of NGS libraries determined using the Qubit 4 Fluorometer system (ThermoFisher Scientific). Columns 1 and 3 present the number of cattle samples. Columns 2 and 4 present the concentration of DNA that extracting from FTA cards using DNeasy Blood & Tissue extraction kit (QIAGEN).

Sample No.	Concentration (ng/ µl)	Sample No.	Concentration(ng/µl)
1	7.13	41	6.30
2	4.34	42	3.39
3	5.72	43	3.69
4	5.62	44	3.14
5	5.50	45	4.92
6	4.46	46	3.42
7	6.18	47	7.40
8	4.17	48	2.88
9	6.40	49	3.25
10	4.44	50	5.02
11	3.28	51	3.67
12	3.10	52	3.29
13	3.73	53	4.03
14	4.01	54	3.59
15	4.10	55	3.10
16	3.22	56	3.09
17	3.68	57	3.50
18	3.78	58	4.53
19	3.46	59	4.91
20	4.18	60	3.21
21	5.09	61	2.44
22	3.46	62	2.80
23	6.13	63	3.32
24	3.02	64	4.48
25	3.66	65	3.41
26	3.39	66	5.02
27	4.39	67	5.58
28	5.02	68	5.07
29	6.84	69	6.33
30	3.98	70	5.42
31	3.66	71	5.95
32	6.04	72	2.91
33	4.85	73	3.55
34	5.79	74	4.47
35	3.71	75	3.55
36	3.03	76	3.84
37	4.09	77	3.66
38	6.39	78	5.41
39	3.89	79	3.58
40	3.56	80	4.50

Table 7. 3 The NGS data (the number of reads per samples, Phred score, error probability) Columns 1 present the number of cattle samples. Columns 2 present the sequence reads pre samples. Columns 3 present the number of the sequences per sample aligned to the reference sequencing. Columns 4 present the quality score per sample. Columns 5 present the percentage of the quality of each base per sample.

Sample number	Number of reads	Total reads X2	Total number of reads alignment to ref	Phred Score (QC score)	Base Call Accuracy
1	9893	19786	17192	37	99.9800474
2	4822	9644	6380	36	99.9748811
3	8600	17200	11572	37	99.9800474
4	10896	21792	19228	37	99.9800474
5	13069	26138	22094	36	99.9748811
6	15091	30182	26984	36	99.9748811
7	11762	23524	19688	37	99.9800474
8	16016	32032	29464	37	99.9800474
9	4285	8570	6018	36	99.9748811
10	9257	18514	15282	37	99.9800474
11	8213	16426	13830	37	99.9800474
12	5905	11810	8029	36	99.9748811
13	9771	19542	13822	37	99.9800474
14	8602	17204	15636	36	99.9748811
15	12521	25042	22030	37	99.9800474
16	9035	18070	15872	37	99.9800474
17	7322	14644	6498	37	99.9800474
18	8036	16072	10514	37	99.9800474
19	6812	13624	6566	36	99.9748811
20	9710	19420	14310	36	99.9748811
21	8310	16620	13496	37	99.9800474
22	8714	17428	9710	37	99.9800474
23	No Read	No Read	No Read	No Read	No Read
24	6410	12820	7348	36	99.9748811

25	7029	14058	11416	36	99.9748811
26	8140	16280	9638	37	99.9800474
27	1432	2864	1562	36	99.9748811
28	3036	6072	1904	37	99.9800474
29	6957	13914	5032	37	99.9800474
30	3780	7560	1958	36	99.9748811
31	891	1782	898	35	99.9683772
32	5111	10222	3124	36	99.9748811
33	11650	23300	8224	37	99.9800474
34	No Read	No Read	No Read	No Read	No Read
35	1381	2762	1094	36	99.9748811
36	4311	8622	5069	37	99.9800474
37	20924	41848	36304	36	99.9748811
38	52	104	80	36	99.9748811
39	8135	16270	12508	36	99.9748811
40	6285	12570	7716	37	99.9800474
41	9769	19538	17180	36	99.9748811
42	6830	13660	10867	36	99.9748811
43	8115	16230	12370	37	99.9800474
44	6103	12206	5988	37	99.9800474
45	9442	18884	15544	36	99.9748811
46	3983	7966	4410	37	99.9800474
47	No Read	No Read	No Read	No Read	No Read
48	15541	31082	24070	36	99.9748811
49	6902	13804	6000	37	99.9800474
50	3409	6818	6048	37	99.9800474
51	15799	31598	19338	37	99.9800474
52	13894	27788	22400	37	99.9800474
53	8338	16676	14146	36	99.9748811

54	4997	9994	8152	36	99.9748811
55	8694	17388	14194	36	99.9748811
56	8129	16258	9858	36	99.9748811
57	5419	10838	8496	37	99.9800474
58	10253	20506	18344	37	99.9800474
59	12607	25214	23368	36	99.9748811
60	15240	30480	27086	36	99.9748811
61	13328	26656	24512	37	99.9800474
62	10132	20264	17866	37	99.9800474
63	16263	32526	24682	36	99.9748811
64	1922	3844	3448	37	99.9800474
65	12091	24182	18902	37	99.9800474
66	7625	15250	12646	36	99.9748811
67	9982	19964	14870	37	99.9800474
68	No Read	No Read	No Read	No Read	No Read
69	5671	11342	8762		
70	28740	57480	51140	37	99.9800474
71	5213	10426	9094	37	99.9800474
72	36860	73720	44602	36	99.9748811
73	2635	5270	3234	37	99.9800474
74	12806	25612	18194	36	99.9748811
75	9727	19454	13682	37	99.9800474
76	15023	30046	18722	36	99.9748811
77	5952	11904	9658	37	99.9800474
78	11056	22112	16000	36	99.9748811
79	21548	43096	34138	36	99.9748811
80	18442	36884	21804	37	99.9800474
Undetermined Read	25454	50908		36.5131579	99.9775146

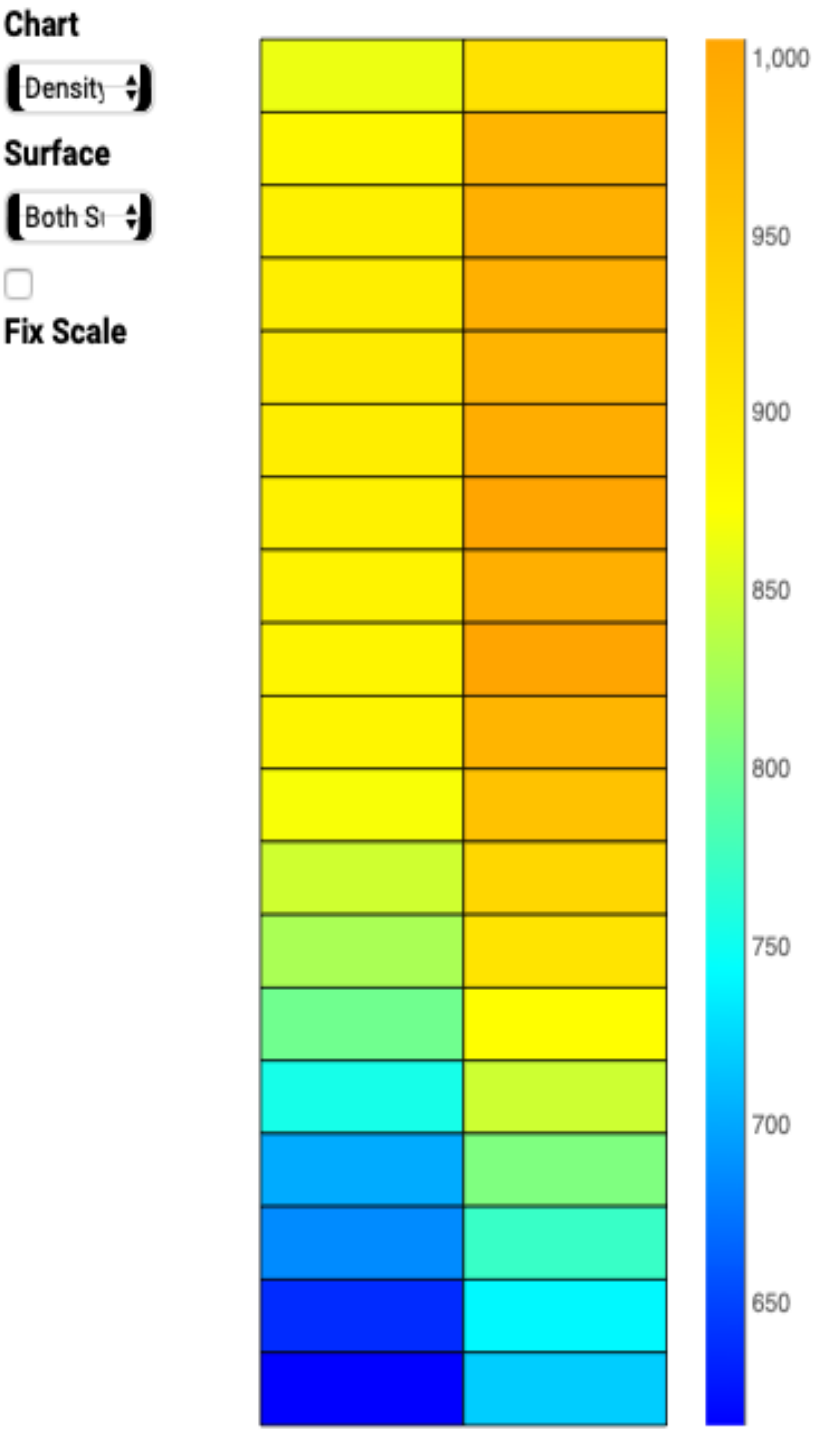


Figure 7. 3 Screenshot of flow cell charts used for visualizing data per tile across the flow cell. The colour scale provides colour coding for easy identification of density levels per tile. Our flow cell shows that most of tile has yellow and orange colour which are the same colour at the higher end of the colour range indicting that the template densities in flow cell increase beyond the optimal range (over-clustered).

References

ABEKHTI, A., DGEGHATI, S., TAMINIAU, B., KIHAL, M. & DAUBE, G. 2017. Optimization of DNA extraction from the Algerian traditional date's product "Btana". *Applied Biology in Saharan Areas,* 1**,** 23-32.

ABENGA, J. N. 2014. A comparative pathology of Trypanosoma brucei infections. *Glob Adv Res J Med Med Sci,* 3**,** 390-399.

ABRHAM, A., SHEMELES, D., TAKELE, A. & BERHANU, D. 2017. Review on Characterization of Trypanosoma congolense; A Major Parasite of Cattle in Africa. *Acta Parasitol. Glob,* 8**,** 39-49.

ADAMS, E., MALELE, I., MSANGI, A. & GIBSON, W. 2006. Trypanosome identification in wild tsetse populations in Tanzania using generic primers to amplify the ribosomal RNA ITS-1 region. *Acta Tropica,* 100**,** 103-109.

ADENAIKE, A., WHETO, M., OLURODE, S., SMITH, O., TALABI, A. & IKEOBI, C. 2018. Single nucleotide polymorphisms in two TLR genes and their effects on trypanosomosis traits in muturu and White Fulani cattle challenged with Trypanosoma vivax in Nigeria. *Bulletin of Animal Health and Production in Africa,* 66**,** 115-122.

AGRAWAL, N., FREDERICK, M. J., PICKERING, C. R., BETTEGOWDA, C., CHANG, K., LI, R. J., FAKHRY, C., XIE, T.-X., ZHANG, J. & WANG, J. 2011. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science,* 333**,** 1154-1157.

AHMED, H. A., MACLEOD, E. T., HIDE, G., WELBURN, S. C. & PICOZZI, K. 2011. The best practice for preparation of samples from FTA® cards for diagnosis of blood borne infections using African trypanosomes as a model system. *Parasites & vectors,* 4**,** 68.

AHMED, H. A., PICOZZI, K., WELBURN, S. C. & MACLEOD, E. T. 2013. A comparative evaluation of PCR-based methods for species-specific determination of African animal trypanosomes in Ugandan cattle. *Parasites & Vectors,* 6**,** 316.

AHSANUDDIN, S., AFSHINNEKOO, E., GANDARA, J., HAKYEMEZOĞLU, M., BEZDAN, D., MINOT, S., GREENFIELD, N. & MASON, C. E. 2017. Assessment of REPLI-g Multiple Displacement Whole Genome Amplification (WGA) Techniques for Metagenomic Applications. *Journal of Biomolecular Techniques: JBT,* 28**,** 46.

AKIRA, S., UEMATSU, S. & TAKEUCHI, O. 2006. Pathogen recognition and innate immunity. *Cell,* 124**,** 783-801.

ALEXOPOULOU, L., THOMAS, V., SCHNARE, M., LOBET, Y., ANGUITA, J., SCHOEN, R. T., MEDZHITOV, R., FIKRIG, E. & FLAVELL, R. A. 2002. Hyporesponsiveness to vaccination with Borrelia burgdorferi OspA in humans and in TLR1-and TLR2-deficient mice. *Nature Medicine,* 8**,** 878-884.

ALI, N., RAMPAZZO, R. D. C. P., COSTA, A. D. T. & KRIEGER, M. A. 2017. Current nucleic acid extraction methods and their implications to point-of-care diagnostics. *BioMed research international,* 2017.

ALIPRANTIS, A. O., YANG, R.-B., MARK, M. R., SUGGETT, S., DEVAUX, B., RADOLF, J. D., KLIMPEL, G. R., GODOWSKI, P. & ZYCHLINSKY, A. 1999. Cell activation and apoptosis by bacterial lipoproteins through toll-like receptor-2. *Science,* 285**,** 736-739.

ALMEIDA, C. N. D., FURIAN, T. Q., BORGES, K. A., PERDONCINI, G., MAUEL, M. J., ROCHA, S. L. D. S., NASCIMENTO, V. P. D., SALLE, C. T. P. & MORAES, H. L. D. S. 2018. Assessment of FTA card employment for Pasteurella multocida DNA transport and detection of virulence-associated genes in strains isolated from fowl cholera in the United States. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia,* 70**,** 1855-1861.

AMARANTE-MENDES, G. P., ADJEMIAN, S., BRANCO, L. M., ZANETTI, L. C., WEINLICH, R. & BORTOLUCI, K. R. 2018. Pattern recognition receptors and the host cell death molecular machinery. *Frontiers in Immunology,* 9**,** 2379.

ANDERS, H.-J. & PATOLE, P. S. 2005. Toll-like receptors recognize uropathogenic Escherichia coli and trigger inflammation in the urinary tract. *Nephrology Dialysis Transplantation,* 20**,** 1529-1532.

ANDERSON, K. V. 2000. Toll signaling pathways in the innate immune response. *Current Opinion in Immunology,* 12**,** 13-19.

ANDREWS, S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

ANSORGE, W. J. 2009. Next-generation DNA sequencing techniques. *New Biotechnology,* 25**,** 195-203.

ARDURA, A., LINDE, A. R., MOREIRA, J. C. & GARCIA-VAZQUEZ, E. 2010. DNA barcoding for conservation and management of Amazonian commercial fish. *Biological Conservation,* 143**,** 1438-1443.

AREGAWI, W. G., AGGA, G. E., ABDI, R. D. & BÜSCHER, P. 2019. Systematic review and meta-analysis on the global distribution, host range, and prevalence of Trypanosoma evansi. *Parasites & Vectors,* 12**,** 67.

ARSENIC, R., TREUE, D., LEHMANN, A., HUMMEL, M., DIETEL, M., DENKERT, C. & BUDCZIES, J. 2015. Comparison of targeted next-generation sequencing and Sanger sequencing for the detection of PIK3CA mutations in breast cancer. *BMC Clin Pathol,* 15**,** 20.

ATTARDO, G. M., ABD-ALLA, A. M., ACOSTA-SERRANO, A., ALLEN, J. E., BATETA, R., BENOIT, J. B., BOURTZIS, K., CAERS, J., CALJON, G. & CHRISTENSEN, M. B. 2019. Comparative genomic analysis of six Glossina genomes, vectors of African trypanosomes. *Genome Biology,* 20**,** 1-31.

AUTY, H., TORR, S. J., MICHOEL, T., JAYARAMAN, S. & MORRISON, L. J. 2015. Cattle trypanosomosis: the diversity of trypanosomes and implications for disease epidemiology and control. *Rev Sci Tech,* 34**,** 587-98.

AWUOCHE, E. O. 2012. Tsetse fly saliva: Could it be useful in fly infection when feeding in chronically aparasitemic mammalian hosts. *Open Vet J,* 2**,** 95-105.

BACHMAN, K. E., ARGANI, P., SAMUELS, Y., SILLIMAN, N., PTAK, J., SZABO, S., KONISHI, H., KARAKAS, B., BLAIR, B. G. & LIN, C. 2004. The PIK3CA gene is mutated with high frequency in human breast cancers. *Cancer Biology & Therapy,* 3**,** 772-775.

BAFICA, A., SANTIAGO, H. C., GOLDSZMID, R., ROPERT, C., GAZZINELLI, R. T. & SHER, A. 2006. Cutting edge: TLR9 and TLR2 signaling together account for MyD88-dependent control of parasitemia in Trypanosoma cruzi infection. *J Immunol,* 177**,** 3515-9.

BANCHEREAU, J. & STEINMAN, R. M. 1998. Dendritic cells and the control of immunity. *Nature,* 392**,** 245-52.

BARGUL, J. L., JUNG, J., MCODIMBA, F. A., OMOGO, C. O., ADUNG'A, V. O., KRÜGER, T., MASIGA, D. K. & ENGSTLER, M. 2016. Species-specific adaptations of trypanosome morphology and motility to the mammalian host. *PLoS Pathog,* 12**,** e1005448.

BARR, T. A., BROWN, S., RYAN, G., ZHAO, J. & GRAY, D. 2007. TLR-mediated stimulation of APC: Distinct cytokine responses of B cells and dendritic cells. *Eur J Immunol,* 37**,** 3040-53.

BARRETT, M. P. 2018. The elimination of human African trypanosomiasis is in sight: Report from the third WHO stakeholders meeting on elimination of gambiense human African trypanosomiasis. *PLoS neglected tropical diseases,* 12**,** e0006925.

BARRIS, W., HARRISON, B., MCWILLIAM, S., BUNCH, R., GODDARD, M. & BARENDSE, W. 2012. Next generation sequencing of African and Indicine cattle to identify single nucleotide polymorphisms. *Animal Production Science,* 52**,** 133-142.

BELINDA, L. W.-C., WEI, W. X., HANH, B. T. H., LEI, L. X., BOW, H. & LING, D. J. 2008. SARM: a novel Toll-like receptor adaptor, is functionally conserved from arthropod to human. *Molecular Immunology,* 45**,** 1732-1742.

BELL, J. K., MULLEN, G. E. D., LEIFER, C. A., MAZZONI, A., DAVIES, D. R. & SEGAL, D. M. 2003. Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends in Immunology,* 24**,** 528-533.

BENTLEY, D. R., BALASUBRAMANIAN, S., SWERDLOW, H. P., SMITH, G. P., MILTON, J., BROWN, C. G., HALL, K. P., EVERS, D. J., BARNES, C. L. & BIGNELL, H. R. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature,* 456**,** 53-59.

BHALADHARE, A., SHARMA, D., CHAUHAN, A., KUMAR, A., SONWANE, A., SINGH, R. V., KUMAR, P., KUMAR, S. & BHUSHAN, B. 2018. Association study of Single Nucleotide

Polymorphisms (SNP) in Toll-like Receptor 9 gene with bovine tuberculosis. *Indian Journal of Animal Research,* 52**,** 533-537.

BISOFFI, Z., BELTRAME, A., MONTEIRO, G., ARZESE, A., MAROCCO, S., RORATO, G., ANSELMI, M. & VIALE, P. 2005. African trypanosomiasis gambiense, Italy. *Emerging Infectious Diseases,* 11**,** 1745.

BISWAL, J. K., SUBRAMANIAM, S., RANJAN, R. & PATTNAIK, B. 2016. Evaluation of FTA® card for the rescue of infectious foot-and-mouth disease virus by chemical transfection of extracted RNA in cultured cells. *Molecular and Cellular Probes,* 30**,** 225-230.

BLASIUS, A. L. & BEUTLER, B. 2010. Intracellular toll-like receptors. *Immunity,* 32**,** 305-315.

BOCHUD, P. Y., HERSBERGER, M., TAFFE, P., BOCHUD, M., STEIN, C. M., RODRIGUES, S. D., CALANDRA, T., FRANCIOLI, P., TELENTI, A., SPECK, R. F., ADEREM, A. & SWISS, H. I. V. C. S. 2007. Polymorphisms in Toll-like receptor 9 influence the clinical course of HIV-1 infection. *AIDS,* 21**,** 441-6.

BONILLA, F. A. & OETTGEN, H. C. 2010. Adaptive immunity. *J Allergy Clin Immunol,* 125**,** S33-40.

BONNERT, T. P., GARKA, K. E., PARNET, P., SONODA, G., TESTA, J. R. & SIMS, J. E. 1997. The cloning and characterization of human MyD88: a member of an IL-1 receptor related family. *FEBS letters,* 402**,** 81-84.

BORGSTRÖM, E., LUNDIN, S. & LUNDEBERG, J. 2011. Large scale library generation for high throughput sequencing. *PloS one,* 6.

BOTOS, I., SEGAL, D. M. & DAVIES, D. R. 2011. The structural biology of Toll-like receptors. *Structure,* 19**,** 447-459.

BOURKE, E., BOSISIO, D., GOLAY, J., POLENTARUTTI, N. & MANTOVANI, A. 2003. The toll-like receptor repertoire of human B lymphocytes: inducible and selective expression of TLR9 and TLR10 in normal and transformed cells. *Blood,* 102**,** 956-963.

BRADLEY, J. R. & POBER, J. S. 2001. Tumor necrosis factor receptor-associated factors (TRAFs). *Oncogene,* 20**,** 6482.

BRANGER, J., LEEMANS, J. C., FLORQUIN, S., WEIJER, S., SPEELMAN, P. & VAN DER POLL, T. 2004. Toll-like receptor 4 plays a protective role in pulmonary tuberculosis in mice. *International Immunology,* 16**,** 509-516.

BROOKES, A. J. 1999. The essence of SNPs. *Gene,* 234**,** 177-186.

BRUN, R., BLUM, J., CHAPPUIS, F. & BURRI, C. 2010. Human african trypanosomiasis. *The Lancet,* 375**,** 148-159.

BUERMANS, H. P. J. & DEN DUNNEN, J. T. 2014. Next generation sequencing technology: advances and applications. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1842**,** 1932-1941.

BUGUET, A., BISSER, S., JOSENANDO, T., CHAPOTOT, F. & CESPUGLIO, R. 2005. Sleep structure: a new diagnostic tool for stage determination in sleeping sickness. *Acta Tropica,* 93**,** 107-117.

BURNS, K., JANSSENS, S., BRISSONI, B., OLIVOS, N., BEYAERT, R. & TSCHOPP, J. 2003. Inhibition of interleukin 1 receptor/Toll-like receptor signaling through the alternatively spliced, short form of MyD88 is due to its failure to recruit IRAK-4. *Journal of Experimental Medicine,* 197**,** 263-268.

BUTLER, J. 2011. *Advanced Topics in Forensic DNA Typing: Methodology*.

CAMPOS, M. A. S., ALMEIDA, I. C., TAKEUCHI, O., AKIRA, S., VALENTE, E. P., PROCÓPIO, D. O., TRAVASSOS, L. R., SMITH, J. A., GOLENBOCK, D. T. & GAZZINELLI, R. T. 2001. Activation of Toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. *The Journal of Immunology,* 167**,** 416-423.

CAO, Y., FANNING, S., PROOS, S., JORDAN, K. & SRIKUMAR, S. 2017. A review on the applications of next generation sequencing technologies as applied to food-related microbiome studies. *Frontiers in Microbiology,* 8**,** 1829.

CAPEWELL, P., VEITCH, N. J., TURNER, C. M., RAPER, J., BERRIMAN, M., HAJDUK, S. L. & MACLEOD, A. 2011. Differences between Trypanosoma brucei gambiense groups 1 and 2 in their resistance to killing by trypanolytic factor 1. *PLoS Negl Trop Dis,* 5**,** e1287.

CARDOSO, M. S., REIS-CUNHA, J. L. & BARTHOLOMEU, D. C. 2016. Evasion of the immune response by Trypanosoma cruzi during acute infection. *Frontiers in Immunology,* 6**,** 659.

CARGILL, E. J. & WOMACK, J. E. 2007. Detection of polymorphisms in bovine toll-like receptors 3, 7, 8, and 9. *Genomics,* 89**,** 745-55.
CARVALHO, A., PASQUALOTTO, A., PITZURRA, L., ROMANI, L., DENNING, D. & RODRIGUES, F. 2008. Polymorphisms in toll-like receptor genes and susceptibility to pulmonary aspergillosis. *The Journal of Infectious Diseases,* 197**,** 618-621.
CASQUET, J., THEBAUD, C. & GILLESPIE, R. G. 2012. Chelex without boiling, a rapid and easy technique to obtain stable amplifiable DNA from small amounts of ethanol-stored spiders. *Molecular Ecology Resources,* 12**,** 136-141.
CECCHI, G., PAONE, M., FELDMANN, U., VREYSEN, M. J., DIALL, O. & MATTIOLI, R. C. 2014. Assembling a geospatial database of tsetse-transmitted animal trypanosomosis for Africa. *Parasites & Vectors,* 7**,** 39.
CHA, R. S. & THILLY, W. G. 1993. Specificity, efficiency, and fidelity of PCR. *PCR Methods Appl,* 3**,** 18-29.
CHAGAS, C. 1909. Nova tripanozomiaze humana: estudos sobre a morfolojia e o ciclo evolutivo do Schizotrypanum cruzi n. gen., n. sp., ajente etiolojico de nova entidade morbida do homem. *Memórias do Instituto Oswaldo Cruz,* 1**,** 159-218.
CHAPMAN, S. J. & HILL, A. V. 2012. Human genetic susceptibility to infectious disease. *Nat Rev Genet,* 13**,** 175-88.
CHARETTE, S. J., BROCHU, F., BOYLE, B., FILION, G., TANAKA, K. H. & DEROME, N. 2012. Draft genome sequence of the virulent strain 01-B526 of the fish pathogen Aeromonas salmonicida. Am Soc Microbiol.
CHECCHI, F., FILIPE, J. A., BARRETT, M. P. & CHANDRAMOHAN, D. 2008. The natural progression of Gambiense sleeping sickness: what is the evidence? *PLoS Neglected Tropical Diseases,* 2**,** e303.
CHITANGA, S., NAMANGALA, B., DE DEKEN, R. & MARCOTTY, T. 2013. Shifting from wild to domestic hosts: the effect on the transmission of Trypanosoma congolense to tsetse flies. *Acta Tropica,* 125**,** 32-36.
CHOU, Q., RUSSELL, M., BIRCH, D. E., RAYMOND, J. & BLOCH, W. 1992. Prevention of pre-PCR mis-priming and primer dimerization improves low-copy-number amplifications. *Nucleic Acids Research,* 20**,** 1717-1723.
CHUANG, T.-H. & ULEVITCH, R. J. 2004. Triad3A, an E3 ubiquitin-protein ligase regulating Toll-like receptors. *Nature Immunology,* 5**,** 495.
CLARK, C., PALTA, P., JOYCE, C. J., SCOTT, C., GRUNDBERG, E., DELOUKAS, P., PALOTIE, A. & COFFEY, A. J. 2012. A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip® for methylome profiling. *PloS One,* 7.
CLAUSEN, P., ADEYEMI, I., BAUER, B., BRELOEER, M., SALCHOW, F. & STAAK, C. 1998. Host preferences of tsetse (Diptera: Glossinidae) based on bloodmeal identifications. *Medical and Veterinary Entomology,* 12**,** 169-180.
CONSORTIUM, I. H. G. S. 2004. Finishing the euchromatic sequence of the human genome. *Nature,* 431**,** 931.
CONTROL, W. E. C. O. T., TRYPANOSOMIASIS, S. O. A. & ORGANIZATION, W. H. 1998. *Control and Surveillance of African Trypanosomiasis: Report of a WHO Expert Committee*, World Health Organization.
CONTROL, W. E. C. O. T., TRYPANOSOMIASIS, S. O. H. A. & ORGANIZATION, W. H. 2013. *Control and Surveillance of Human African Trypanosomiasis: Report of a WHO Expert Committee*, World Health Organization.
COOK , D. A. T. 2015. Molecular Microbial Diagnostic Methods: Pathways to Implementation for the Food and Water Industries.
COOK, N., D'AGOSTINO, M. & THOMPSON, K. C. 2015. *Molecular Microbial Diagnostic Methods: Pathways to Implementation for the Food and Water Industries*, Academic Press.
COOPER, D. N., MORT, M., STENSON, P. D., BALL, E. V. & CHUZHANOVA, N. A. 2010. Methylation-mediated deamination of 5-methylcytosine appears to give rise to mutations causing human inherited disease in CpNpG trinucleotides, as well as in CpG dinucleotides. *Hum Genomics,* 4**,** 406-10.

CORNÉLIE, S., HOEBEKE, J., SCHACHT, A.-M., BERTIN, B., VICOGNE, J., CAPRON, M. & RIVEAU, G. 2004. Direct evidence that toll-like receptor 9 (TLR9) functionally binds plasmid DNA by specific cytosine-phosphate-guanine motif recognition. *Journal of Biological Chemistry,* 279**,** 15124-15129.

COSTA, A. G., RAMASAWMY, R., IBIAPINA, H. N. S., SAMPAIO, V. S., XÁBREGAS, L. A., BRASIL, L. W., TARRAGÔ, A. M., ALMEIDA, A. C. G., KUEHN, A. & VITOR-SILVA, S. 2017. Association of TLR variants with susceptibility to Plasmodium vivax malaria and parasitemia in the Amazon region of Brazil. *PloS One,* 12.

COX, A., TILLEY, A., MCODIMBA, F., FYFE, J., EISLER, M., HIDE, G. & WELBURN, S. 2005. A PCR based assay for detection and differentiation of African trypanosome species in blood. *Experimental Parasitology,* 111**,** 24-29.

COX, A. P., TOSAS, O., TILLEY, A., PICOZZI, K., COLEMAN, P., HIDE, G. & WELBURN, S. C. 2010. Constraints to estimating the prevalence of trypanosome infections in East African zebu cattle. *Parasites & Vectors,* 3**,** 82.

DA CUNHA SANTOS, G. 2018. FTA cards for preservation of nucleic acids for molecular assays: a review on the use of cytologic/tissue samples. *Archives of Pathology & Laboratory Medicine,* 142**,** 308-312.

DA CUNHA SANTOS, G., SAIEG, M., TRONCONE, G. & ZEPPA, P. 2018. Cytological preparations for molecular analysis: A review of technical procedures, advantages and limitations for referring samples for testing. *Cytopathology,* 29**,** 125-132.

DA SILVA CORREIA, J., SOLDAU, K., CHRISTEN, U., TOBIAS, P. S. & ULEVITCH, R. J. 2001. Lipopolysaccharide is in close proximity to each of the proteins in its membrane receptor complex transfer from CD14 to TLR4 and MD-2. *Journal of Biological Chemistry,* 276**,** 21129-21135.

DAGNACHEW, S. & BEZIE, M. 2015. Review on Trypanosoma vivax. *Afr. J. Basic Appl. Sci.,* 7**,** 41-64.

DAGNACHEW, S., TEREFE, G., ABEBE, G., SIRAK, A., BOLLO, E., BARRY, D. & GODDEERIS, B. 2015. Comparative clinico-pathological observations in young Zebu (Bos indicus) cattle experimentally infected with Trypanosoma vivax isolates from tsetse infested and non-tsetse areas of Northwest Ethiopia. *BMC Veterinary Research,* 11**,** 307.

DASU, M. R., DEVARAJ, S., PARK, S. & JIALAL, I. 2010. Increased toll-like receptor (TLR) activation and TLR ligands in recently diagnosed type 2 diabetic subjects. *Diabetes Care,* 33**,** 861-868.

DAVEY, S., NAVARRETE, C. & BROWN, C. 2017. Simultaneous human platelet antigen genotyping and detection of novel single nucleotide polymorphisms by targeted next-generation sequencing. *Transfusion,* 57**,** 1497-1504.

DAYEH, T. A., OLSSON, A. H., VOLKOV, P., ALMGREN, P., RÖNN, T. & LING, C. 2013. Identification of CpG-SNPs associated with type 2 diabetes and differential DNA methylation in human pancreatic islets. *Diabetologia,* 56**,** 1036-1046.

DEANE, M. P., LENZI, H. L. & JANSEN, A. 1984. Trypanosoma cruzi: vertebrate and invertebrate cycles in the same mammal host, the opossum Didelphis marsupialis. *Memórias do Instituto Oswaldo Cruz,* 79**,** 513-515.

DESNUES, B., MACEDO, A. B., ROUSSEL-QUEVAL, A., BONNARDEL, J., HENRI, S., DEMARIA, O. & ALEXOPOULOU, L. 2014. TLR8 on dendritic cells and TLR9 on B cells restrain TLR7-mediated spontaneous autoimmunity in C57BL/6 mice. *Proceedings of the National Academy of Sciences,* 111**,** 1497-1502.

DESQUESNES, M., BITEAU-COROLLER, F., BOUYER, J., DIA, M. L. & FOIL, L. 2009. Development of a mathematical model for mechanical transmission of trypanosomes and other pathogens of cattle transmitted by tabanids. *International Journal for Parasitology,* 39**,** 333-346.

DESQUESNES, M., HOLZMULLER, P., LAI, D. H., DARGANTES, A., LUN, Z. R. & JITTAPLAPONG, S. 2013. Trypanosoma evansi and surra: a review and perspectives on origin, history, distribution, taxonomy, morphology, hosts, and pathogenic effects. *Biomed Res Int,* 2013**,** 194176.

DESTA, M., BEYENE, D. & HAILE, S. 2013. Trypanosome infection rate of Glossina pallidipes and trypanosomosis prevalence in cattle in Amaro Special District of Southern Ethiopia. *Journal of Veterinary Medicine and Animal Health,* 5**,** 164-170.

DHANGADAMAJHI, G., KAR, A., ROUT, R. & DHANGADAMAJHI, P. 2017. A meta-analysis of TLR4 and TLR9 SNPs implicated in severe malaria. *Rev Soc Bras Med Trop,* 50**,** 153-160.

DI RESTA, C. & FERRARI, M. 2018. Next Generation Sequencing: From Research Area to Clinical Practice. *EJIFCC,* 29**,** 215-220.

DIALL, O., BOCOUM, Z., DIARRA, B., SANOGO, Y., COULIBALY, Z. & WAIGALO, Y. 1993. Epidemiology of trypanosomiasis caused by T. evansi in camels in mali: results of parasitological and clinical survey. *Revue d'elevage et de medecine veterinaire des pays tropicaux,* 46**,** 455-461.

DIGUISTINI, S., LIAO, N. Y., PLATT, D., ROBERTSON, G., SEIDEL, M., CHAN, S. K., DOCKING, T. R., BIROL, I., HOLT, R. A. & HIRST, M. 2009. De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biology,* 10**,** R94.

DOHM, J. C., LOTTAZ, C., BORODINA, T. & HIMMELBAUER, H. 2008. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Research,* 36**,** e105.

DOS-SANTOS, A. L., CARVALHO-KELLY, L. F., DICK, C. F. & MEYER-FERNANDES, J. R. 2016. Innate immunomodulation to trypanosomatid parasite infections. *Exp Parasitol,* 167**,** 67-75.

DRENNAN, M. B., STIJLEMANS, B., VAN DEN ABBEELE, J., QUESNIAUX, V. J., BARKHUIZEN, M., BROMBACHER, F., DE BAETSELIER, P., RYFFEL, B. & MAGEZ, S. 2005. The induction of a type 1 immune response following a Trypanosoma brucei infection is MyD88 dependent. *The Journal of Immunology,* 175**,** 2501-2509.

DYER, N. A., ROSE, C., EJEH, N. O. & ACOSTA-SERRANO, A. 2013. Flying tryps: survival and maturation of trypanosomes in tsetse flies. *Trends in Parasitology,* 29**,** 188-196.

EAVES-PYLES, T. D., WONG, H. R., ODOMS, K. & PYLES, R. B. 2001. Salmonella flagellin-dependent proinflammatory responses are localized to the conserved amino and carboxyl regions of the protein. *The Journal of Immunology,* 167**,** 7009-7016.

EISLER, M., DWINGER, R., MAJIWA, P., PICOZZI, K., MAUDLIN, I., HOLMES, P. & MILES, M. 2004. Diagnosis and epidemiology of African animal trypanosomiasis. *The Trypanosomiases,* 253**,** 267.

ENRIQUEZ, G. F., BUA, J., OROZCO, M. M., WIRTH, S., SCHIJMAN, A. G., GÜRTLER, R. & CARDINAL, M. V. 2014. High levels of Trypanosoma cruzi DNA determined by qPCR and infectiousness to Triatoma infestans support dogs and cats are major sources of parasites for domestic transmission. *Infection, Genetics and Evolution,* 25**,** 36-43.

ENWEZOR, F. N., UMOH, J. U., ESIEVO, K. A., HALID, I., ZARIA, L. T. & ANERE, J. I. 2009. Survey of bovine trypanosomosis in the Kachia Grazing Reserve, Kaduna State, Nigeria. *Vet Parasitol,* 159**,** 121-5.

EWING, B., HILLIER, L., WENDL, M. C. & GREEN, P. 1998. Base-calling of automated sequencer traces usingPhred. I. Accuracy assessment. *Genome Research,* 8**,** 175-185.

FALLIS, A. M. 1986. Historical Column//Page d'histoire: GRIFFITH EVANS 1835-1935 DISCOVERER OF THE FIRST PATHOGENIC TRYPANOSOME. *The Canadian Veterinary Journal,* 27**,** 336.

FLORIN-CHRISTENSEN, M. & SCHNITTGER, L. 2018. *Parasitic protozoa of farm animals and pets*, Springer.

FRANCO, J. R., SIMARRO, P. P., DIARRA, A. & JANNIN, J. G. 2014. Epidemiology of human African trypanosomiasis. *Clinical Epidemiology,* 6**,** 257.

FUJIMOTO, M. & NAKA, T. 2003. Regulation of cytokine signaling by SOCS family molecules. *Trends in Immunology,* 24**,** 659-666.

GALUCCI, S. & MATZINGER, P. 2001. Danger signals: SOS to the immune system. Curr. Op. *Immunology,* 13**,** 114-9.

GARCIA, H., GARCIA, M.-E., PEREZ, H. & MENDOZA-LEON, A. 2005. The detection and PCR-based characterization of the parasites causing trypanosomiasis in water-buffalo herds in Venezuela. *Annals of Tropical Medicine & Parasitology,* 99**,** 359-370.

GARDINER, P. 1989. Recent studies of the biology of Trypanosoma vivax. *Advances in Parasitology.* Elsevier.

GARG, N., NUNES, M. P. & TARLETON, R. L. 1997. Delivery by Trypanosoma cruzi of proteins into the MHC class I antigen processing and presentation pathway. *The Journal of Immunology,* 158**,** 3293-3302.

GAVAN, M. K., OLIVER, M. K., DOUGLAS, A. & PIERTNEY, S. B. 2015. Gene dynamics of toll-like receptor 4 through a population bottleneck in an insular population of water voles (Arvicola amphibius). *Conservation Genetics,* 16**,** 1181-1193.

GAY, N. J. & GANGLOFF, M. 2007. Structure and function of Toll receptors and their ligands. *Annu Rev Biochem,* 76**,** 141-65.

GELMAN, A. E., ZHANG, J., CHOI, Y. & TURKA, L. A. 2004. Toll-like receptor ligands directly promote activated CD4+ T cell survival. *J Immunol,* 172**,** 6065-73.

GENOMES PROJECT, C. 2015. A global reference for human genetic variation. *Nature,* 526**,** 68.

GHOSH, D. & S STUMHOFER, J. 2013. Do you see what I see: recognition of protozoan parasites by toll-like receptors. *Current Immunology Reviews,* 9**,** 129-140.

GIBSON, W. 2001. Molecular characterization of field isolates of human pathogenic trypanosomes. *Tropical Medicine & International Health,* 6**,** 401-406.

GLÖCKLE, N., KOHL, S., MOHR, J., SCHEURENBRAND, T., SPRECHER, A., WEISSCHUH, N., BERND, A., RUDOLPH, G., SCHUBACH, M. & POLOSCHEK, C. 2014. Panel-based next generation sequencing as a reliable and efficient technique to detect mutations in unselected patients with retinal dystrophies. *European Journal of Human Genetics,* 22**,** 99-104.

GODFREY, D. 1961. Types of Trypanosoma Congolense: II.—Differences in the Courses of Infection. *Annals of Tropical Medicine & Parasitology,* 55**,** 154-166.

GOWIN, E., ŚWIĄTEK-KOŚCIELNA, B., KAŁUŻNA, E., NOWAK, J., MICHALAK, M., WYSOCKI, J. & JANUSZKIEWICZ-LEWANDOWSKA, D. 2017. Analysis of TLR2, TLR4, and TLR9 single nucleotide polymorphisms in children with bacterial meningitis and their healthy family members. *International Journal of Infectious Diseases,* 60**,** 23-28.

GRAUSTEIN, A., HORNE, D., ARENTZ, M., BANG, N., CHAU, T., THWAITES, G., CAWS, M., THUONG, N., DUNSTAN, S. & HAWN, T. 2015. TLR9 gene region polymorphisms and susceptibility to tuberculosis in Vietnam. *Tuberculosis,* 95**,** 190-196.

GREENE, C. E. 2006. *Infectious diseases of the dog and cat*, WB Saunders\Elsevier Science.

GRIEBEL, P. J., BROWNLIE, R., MANUJA, A., NICHANI, A., MOOKHERJEE, N., POPOWYCH, Y., MUTWIRI, G., HECKER, R. & BABIUK, L. A. 2005. Bovine toll-like receptor 9: a comparative analysis of molecular structure, function and expression. *Vet Immunol Immunopathol,* 108**,** 11-6.

GRUND, E., DARISSA, O. & ADAM, G. 2010. Application of FTA® Cards to Sample Microbial Plant Pathogens for PCR and RT-PCR. *Journal of Phytopathology,* 158**,** 750-757.

GRUNFELDER, C. G., ENGSTLER, M., WEISE, F., SCHWARZ, H., STIERHOF, Y.-D., MORGAN, G. W., FIELD, M. C. & OVERATH, P. 2003. Endocytosis of a glycosylphosphatidylinositol-anchored protein via clathrin-coated vesicles, sorting by default in endosomes, and exocytosis via RAB11-positive carriers. *Molecular Biology of the Cell,* 14**,** 2029-2040.

GUAN, Y., RANOA, D. R. E., JIANG, S., MUTHA, S. K., LI, X., BAUDRY, J. & TAPPING, R. I. 2010. Human TLRs 10 and 1 share common mechanisms of innate immune sensing but not signaling. *The Journal of Immunology,* 184**,** 5094-5103.

GUINANE, C. M. & COTTER, P. D. 2013. Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. *Therap Adv Gastroenterol,* 6**,** 295-308.

GULL, K. 2003. Host–parasite interactions and trypanosome morphogenesis: a flagellar pocketful of goodies. *Current Opinion in Microbiology,* 6**,** 365-370.

GUO, J., XU, N., LI, Z., ZHANG, S., WU, J., KIM, D. H., MARMA, M. S., MENG, Q., CAO, H. & LI, X. 2008. Four-color DNA sequencing with 3′-O-modified nucleotide reversible terminators and chemically cleavable fluorescent dideoxynucleotides. *Proceedings of the National Academy of Sciences,* 105**,** 9145-9150.

GUO, Y., LI, N., LYSÉN, C., FRACE, M., TANG, K., SAMMONS, S., ROELLIG, D. M., FENG, Y. & XIAO, L. 2015. Isolation and enrichment of Cryptosporidium DNA and verification of DNA purity for whole-genome sequencing. *Journal of Clinical Microbiology,* 53**,** 641-647.

GURGUL, A., MIKSZA-CYBULSKA, A., SZMATOŁA, T., JASIELCZUK, I., PIESTRZYŃSKA-KAJTOCH, A., FORNAL, A., SEMIK-GURGUL, E. & BUGNO-PONIEWIERSKA, M. 2018. Genotyping-by-sequencing performance in selected livestock species. *Genomics*.

GÜRTLER, R. E. & CARDINAL, M. V. 2015. Reservoir host competence and the role of domestic and commensal hosts in the transmission of Trypanosoma cruzi. *Acta Tropica,* 151**,** 32-50.

GUTIERREZ, C., CORBERA, J. A., JUSTE, M. C., DORESTE, F. & MORALES, I. 2006. Clinical, hematological, and biochemical findings in an outbreak of abortion and neonatal mortality associated with Trypanosoma evansi infection in dromedary camels. *Annals of the New York Academy of Sciences,* 1081**,** 325-327.

HAJJAR, A. M., O'MAHONY, D. S., OZINSKY, A., UNDERHILL, D. M., ADEREM, A., KLEBANOFF, S. J. & WILSON, C. B. 2001. Cutting edge: functional interactions between toll-like receptor (TLR) 2 and TLR1 or TLR6 in response to phenol-soluble modulin. *The Journal of Immunology,* 166**,** 15-19.

HAMILL, L. C., KAARE, M. T., WELBURN, S. C. & PICOZZI, K. 2013. Domestic pigs as potential reservoirs of human and animal trypanosomiasis in Northern Tanzania. *Parasites & Vectors,* 6**,** 322.

HARRIS, R. A., WANG, T., COARFA, C., NAGARAJAN, R. P., HONG, C., DOWNEY, S. L., JOHNSON, B. E., FOUSE, S. D., DELANEY, A. & ZHAO, Y. 2010. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nature Biotechnology,* 28**,** 1097.

HARSINI, S., BEIGY, M., AKHAVAN-SABBAGH, M. & REZAEI, N. 2014. Toll-like receptors in lymphoid malignancies: double-edged sword. *Critical Reviews in Oncology/Hematology,* 89**,** 262-283.

HASAN, U., CHAFFOIS, C., GAILLARD, C., SAULNIER, V., MERCK, E., TANCREDI, S., GUIET, C., BRIÈRE, F., VLACH, J. & LEBECQUE, S. 2005. Human TLR10 is a functional receptor, expressed by B cells and plasmacytoid dendritic cells, which activates gene transcription through MyD88. *The Journal of Immunology,* 174**,** 2942-2950.

HASHIMOTO, M., BANDO, M., KIDO, J.-I., YOKOTA, K., MITA, T., KAJIMOTO, K. & KATAOKA, M. 2019. Nucleic acid purification from dried blood spot on FTA Elute Card provides template for polymerase chain reaction for highly sensitive Plasmodium detection. *Parasitology International,* 73**,** 101941.

HAYASHI, F., SMITH, K. D., OZINSKY, A., HAWN, T. R., EUGENE, C. Y., GOODLETT, D. R., ENG, J. K., AKIRA, S., UNDERHILL, D. M. & ADEREM, A. 2001. The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature,* 410**,** 1099-1103.

HE, Z., ZHU, Y. & GU, H. 2013. A new method for the determination of critical polyethylene glycol concentration for selective precipitation of DNA fragments. *Applied Microbiology and Biotechnology,* 97**,** 9175-9183.

HEAD, S. R., KOMORI, H. K., LAMERE, S. A., WHISENANT, T., VAN NIEUWERBURGH, F., SALOMON, D. R. & ORDOUKHANIAN, P. 2014. Library construction for next-generation sequencing: overviews and challenges. *Biotechniques,* 56**,** 61-77.

HEIL, F., HEMMI, H., HOCHREIN, H., AMPENBERGER, F., KIRSCHNING, C., AKIRA, S., LIPFORD, G., WAGNER, H. & BAUER, S. 2004. Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science,* 303**,** 1526-1529.

HELLMAN, A. & CHESS, A. 2010. Extensive sequence-influenced DNA methylation polymorphism in the human genome. *Epigenetics & Chromatin,* 3**,** 11.

HEMMI, H., TAKEUCHI, O., KAWAI, T., KAISHO, T., SATO, S., SANJO, H., MATSUMOTO, M., HOSHINO, K., WAGNER, H. & TAKEDA, K. 2000. A Toll-like receptor recognizes bacterial DNA. *Nature,* 408**,** 740-745.

HENNESSY, E. J., PARKER, A. E. & O'NEILL, L. A. J. 2010. Targeting Toll-like receptors: emerging therapeutics? *Nature Reviews Drug Discovery,* 9**,** 293-307.

HENRICK, B. M., YAO, X. D., ZAHOOR, M. A., ABIMIKU, A., OSAWE, S. & ROSENTHAL, K. L. 2019. TLR10 Senses HIV-1 Proteins and Significantly Enhances HIV-1 Infection. *Front Immunol,* 10**,** 482.

HIDE, G. 1999. History of sleeping sickness in East Africa. *Clin Microbiol Rev,* 12**,** 112-25.

HIDE, G., HUGHES, J. M. & MCNUFF, R. 2003. A rapid and simple method of detection of Blepharisma japonicum using PCR and immobilisation on FTA paper. *BMC ecology,* 3**,** 7.

HIDE, G. & TAIT, A. 2009. Molecular epidemiology of African sleeping sickness. *Parasitology,* 136**,** 1491-1500.

HIDE, G., TAIT, A., MAUDLIN, I. & WELBURN, S. 1996. The origins, dynamics and generation of Trypanosoma brucei rhodesiense epidemics in East Africa. *Parasitology Today,* 12**,** 50-55.
HIDMARK, A., VON SAINT PAUL, A. & DALPKE, A. H. 2012. Cutting edge: TLR13 is a receptor for bacterial RNA. *The Journal of Immunology,* 189**,** 2717-2721.
HIDRON, A., VOGENTHALER, N., SANTOS-PRECIADO, J. I., RODRIGUEZ-MORALES, A. J., FRANCO-PAREDES, C. & RASSI, A. 2010. Cardiac involvement with parasitic infections. *Clinical Microbiology Reviews,* 23**,** 324-349.
HIRSCHFELD, M., KIRSCHNING, C. J., SCHWANDNER, R., WESCHE, H., WEIS, J. H., WOOTEN, R. M. & WEIS, J. J. 1999. Cutting edge: inflammatory signaling by Borrelia burgdorferi lipoproteins is mediated by toll-like receptor 2. *The Journal of Immunology,* 163**,** 2382-2386.
HOARE, C. A. 1972. The trypanosomes of mammals. A zoological monograph. *The trypanosomes of mammals. A Zoological Monograph.*
HOEBE, K., JANSSEN, E. & BEUTLER, B. 2004. The interface between innate and adaptive immunity. *Nature Immunology,* 5**,** 971-974.
HOGABOAM, C. M., MURRAY, L. & MARTINEZ, F. J. 2012. Epigenetic mechanisms through which Toll-like receptor-9 drives idiopathic pulmonary fibrosis progression. *Proc Am Thorac Soc,* 9**,** 172-6.
HOLEC-GASIOR, L., DRAPALA, D., DOMINIAK-GORSKI, B. & KUR, J. 2013. Epidemiological study of Toxoplasma gondii infection among cattle in Northern Poland. *Ann Agric Environ Med,* 20**,** 653-6.
HOLLAND, W., DO, T., HUONG, N., DUNG, N., THANH, N., VERCRUYSSE, J. & GODDEERIS, B. 2003. The effect of Trypanosoma evansi infection on pig performance and vaccination against classical swine fever. *Veterinary Parasitology,* 111**,** 115-123.
HOLLAND, W., THANH, N., DO, T., SANGMANEEDET, S., GODDEERIS, B. & VERCRUYSSE, J. 2005. Evaluation of diagnostic tests for Trypanosoma evansi in experimentally infected pigs and subsequent use in field surveys in north Vietnam and Thailand. *Tropical Animal Health and Production,* 37**,** 457-467.
HOLMES, P. 2015. On the road to elimination of rhodesiense human African trypanosomiasis: first WHO meeting of stakeholders. *PLoS Neglected Tropical Diseases,* 9**,** e0003571.
HOLT, H. R., SELBY, R., MUMBA, C., NAPIER, G. B. & GUITIAN, J. 2016. Assessment of animal African trypanosomiasis (AAT) vulnerability in cattle-owning communities of sub-Saharan Africa. *Parasit Vectors,* 9**,** 53.
HOOD, L. & ROWEN, L. 2013. The human genome project: big science transforms biology and medicine. *Genome Medicine,* 5**,** 79.
HORNUNG, V., GUENTHNER-BILLER, M., BOURQUIN, C., ABLASSER, A., SCHLEE, M., UEMATSU, S., NORONHA, A., MANOHARAN, M., AKIRA, S. & DE FOUGEROLLES, A. 2005. Sequence-specific potent induction of IFN-α by short interfering RNA in plasmacytoid dendritic cells through TLR7. *Nature Medicine,* 11**,** 263-270.
HU, Q., LIU, Y., YI, S. & HUANG, D. 2015. A comparison of four methods for PCR inhibitor removal. *Forensic Science International: Genetics,* 16**,** 94-97.
IDEOZU, E. 2015. *The use of molecular tools for Pan-Trypanosoma analysis and epigenetics of the host.* PhD, University of Salford Manchester.
IDEOZU, E. J., WHITEOAK, A. M., TOMLINSON, A. J., ROBERTSON, A., DELAHAY, R. J. & HIDE, G. 2015. High prevalence of trypanosomes in European badgers detected using ITS-PCR. *Parasit Vectors,* 8**,** 480.
IWASAKI, A. & MEDZHITOV, R. 2004. Toll-like receptor control of the adaptive immune responses. *Nature Immunology,* 5**,** 987.
IWASAKI, A. & MEDZHITOV, R. 2010. Regulation of adaptive immunity by the innate immune system. *Science,* 327**,** 291-295.
IWASAKI, A. & MEDZHITOV, R. 2015. Control of adaptive immunity by the innate immune system. *Nature Immunology,* 16**,** 343-353.
JANEWAY, C. A. 1989. A primitive immune system. *Nature*, *341*(6238), 108-108.
JANEWAY JR, C. A. & MEDZHITOV, R. 2002. Innate immune recognition. *Annual Review of Immunology,* 20**,** 197-216.

JASMINE, F., AHSAN, H., ANDRULIS, I. L., JOHN, E. M., CHANG-CLAUDE, J. & KIBRIYA, M. G. 2008. Whole-genome amplification enables accurate genotyping for microarray-based high-density single nucleotide polymorphism array. *Cancer Epidemiology and Prevention Biomarkers,* 17**,** 3499-3508.

JEFFERY, K. J., SIDDIQUI, A. A., BUNCE, M., LLOYD, A. L., VINE, A. M., WITKOVER, A. D., IZUMO, S., USUKU, K., WELSH, K. I., OSAME, M. & BANGHAM, C. R. 2000. The influence of HLA class I alleles and heterozygosity on the outcome of human T cell lymphotropic virus type I infection. *J Immunol,* 165**,** 7278-84.

JEONG, H., SIM, Y. M., KIM, H. J., LEE, D.-W., LIM, S.-K. & LEE, S. J. 2013. Genome sequence of the vancomycin-producing Amycolatopsis orientalis subsp. orientalis strain KCTC 9412T. *Genome Announc.,* 1**,** e00408-13.

JIGNAL, P., SHAIKH, M. & DARSHAN, M. 2014. Forensic conception: DNA typing of FTA spotted samples. *J. Appl. Biol. Biotechnol,* 2**,** 21-29.

JONES, T. W. & DÁVILA, A. M. 2001. Trypanosoma vivax–out of Africa. *TRENDS in Parasitology,* 17**,** 99-101.

JÓŹWIAK, M., WYROSTEK, K., DOMAŃSKA-BLICHARZ, K., OLSZEWSKA-TOMCZYK, M., ŚMIETANKA, K. & MINTA, Z. 2016. Application of FTA® Cards for detection and storage of avian influenza virus. *Journal of Veterinary Research,* 60**,** 1-6.

JR, A. R., RASSI, A. & LITTLE, W. C. 2000. Chagas' heart disease. *Clinical cardiology,* 23**,** 883-889.

KAISHO, T. & AKIRA, S. 2006. Toll-like receptor function and signaling. *Journal of Allergy and Clinical Immunology,* 117**,** 979-987.

KANAGAL-SHAMANNA, R. 2016. Emulsion PCR: techniques and applications. *Clinical Applications of PCR.* Springer.

KANG, S. H., ABDEL-MASSIH, R. C., BROWN, R. A., DIERKHISING, R. A., KREMERS, W. K. & RAZONABLE, R. R. 2012. Homozygosity for the toll-like receptor 2 R753Q single-nucleotide polymorphism is a risk factor for cytomegalovirus disease after liver transplantation. *J Infect Dis,* 205**,** 639-46.

KANG, T. J., LEE, S. B. & CHAE, G. T. 2002. A polymorphism in the toll-like receptor 2 is associated with IL-12 production from monocyte in lepromatous leprosy. *Cytokine,* 20**,** 56-62.

KAPSENBERG, M. L. 2003. Dendritic-cell control of pathogen-driven T-cell polarization. *Nature Reviews Immunology,* 3**,** 984.

KARKI, R., PANDYA, D., ELSTON, R. C. & FERLINI, C. 2015. Defining "mutation" and "polymorphism" in the era of personal genomics. *BMC Medical Genomics,* 8**,** 37.

KASHYAP, P., DEKA, M., DUTTA, S. & KUMARI, N. 2016. Association of toll-like receptor (TLR) 2, 7 and 9 gene polymorphism with hepatitis A virus infection risk in population of Assam, India. *Bioscience Biotechnology Research Communication,* 9**,** 293-299.

KAUFMANN, S. H. E. 2008. Immunology's foundation: the 100-year anniversary of the Nobel Prize to Paul Ehrlich and Elie Metchnikoff. *Nature Immunology,* 9**,** 705-712.

KAWAI, T. & AKIRA, S. 2006. Innate immune recognition of viral infection. *Nature Immunology,* 7**,** 131-137.

KAWAI, T. & AKIRA, S. 2007. TLR signaling. *Semin Immunol,* 19**,** 24-32.

KAWAI, T. & AKIRA, S. 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol,* 11**,** 373-84.

KAYAGAKI, N., PHUNG, Q., CHAN, S., CHAUDHARI, R., QUAN, C., O'ROURKE, K. M., EBY, M., PIETRAS, E., CHENG, G. & BAZAN, J. F. 2007. A deubiquitinase that regulates type I interferon production. *Science,* 318**,** 1628-1632.

KENNEDY, P. G. 2008. The continuing problem of human African trypanosomiasis (sleeping sickness). *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society,* 64**,** 116-126.

KENNEDY, P. G. 2013. Clinical features, diagnosis, and treatment of human African trypanosomiasis (sleeping sickness). *The Lancet Neurology,* 12**,** 186-194.

KENNEDY, P. G. 2019. Update on human African trypanosomiasis (sleeping sickness). *Journal of Neurology***,** 1-4.

KENNEDY, P. G. & RODGERS, J. 2019. Clinical and neuropathogenetic aspects of human African trypanosomiasis. *Frontiers in Immunology,* 10**,** 39.
KHARAJI, M. H. & HAGHPARAST, A. 2010. Simultaneous Detection of Pattern Recognition Receptors (PRRs) in Human Peripheral Blood Mononuclear Cells (PBMC) by Touchdown PCR. *World Applied Sciences Journal,* 9**,** 479-483.
KIJPITTAYARIT, S., EID, A. J., BROWN, R. A., PAYA, C. V. & RAZONABLE, R. R. 2007. Relationship between Toll-like receptor 2 polymorphism and cytomegalovirus disease after liver transplantation. *Clin Infect Dis,* 44**,** 1315-20.
KIM, Y. K., SHIN, J.-S. & NAHM, M. H. 2016. NOD-like receptors in infection, immunity, and diseases. *Yonsei Medical Journal,* 57**,** 5-14.
KINJYO, I., HANADA, T., INAGAKI-OHARA, K., MORI, H., AKI, D., OHISHI, M., YOSHIDA, H., KUBO, M. & YOSHIMURA, A. 2002. SOCS1/JAB is a negative regulator of LPS-induced macrophage activation. *Immunity,* 17**,** 583-591.
KNOWLES, G., BETSCHART, B., KUKLA, B., SCOTT, J. & MAJIWA, P. 1988. Genetically discrete populations of Trypanosoma congolense from livestock on the Kenyan coast. *Parasitology,* 96**,** 461-474.
KONDO, T., KAWAI, T. & AKIRA, S. 2012. Dissecting negative regulation of Toll-like receptor signaling. *Trends in Immunology,* 33**,** 449-458.
KRAWCZAK, M., REISS, J. & COOPER, D. N. 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human genetics,* 90**,** 41-54.
KRUG, A., ROTHENFUSSER, S., HORNUNG, V., JAHRSDÖRFER, B., BLACKWELL, S., BALLAS, Z. K., ENDRES, S., KRIEG, A. M. & HARTMANN, G. 2001. Identification of CpG oligonucleotide sequences with high induction of IFN-α/β in plasmacytoid dendritic cells. *European Journal of Immunology,* 31**,** 2154-2163.
KULIS, M. & ESTELLER, M. 2010. DNA methylation and cancer. *Adv Genet,* 70**,** 27-56.
KUMAR, H., GUPTA, M. P., SIDHU, P. K., MAHAJAN, V., BAL, M. S., KAUR, K., ASHUMA, VERMA, S. & SINGLA, L. D. 2012. An outbreak of acute Trypanosoma evansi infection in crossbred cattle in Punjab, India. *Journal of Applied Animal Research,* 40**,** 256-259.
KUMAR, H., KAWAI, T. & AKIRA, S. 2011. Pathogen recognition by the innate immune system. *Int Rev Immunol,* 30**,** 16-34.
KUMAR, N., MUKHOPADHYAY, A. K., PATRA, R., DE, R., BADDAM, R., SHAIK, S., ALAM, J., TIRUVAYIPATI, S. & AHMED, N. 2012. Next-generation sequencing and de novo assembly, genome organization, and comparative genomic analyses of the genomes of two Helicobacter pylori isolates from duodenal ulcer patients in India. Am Soc Microbiol.
KUMAR, S., BANKS, T. W. & CLOUTIER, S. 2012. SNP discovery through next-generation sequencing and its applications. *International Journal of Plant Genomics,* 2012.
KVIST, T., AHRING, B. K., LASKEN, R. S. & WESTERMANN, P. 2007. Specific single-cell isolation and genomic amplification of uncultured microorganisms. *Applied Microbiology and Biotechnology,* 74**,** 926-935.
LALANI, T., TISDALE, M. D., LIU, J., MITRA, I., PHILIP, C., ODUNDO, E., REYES, F., SIMONS, M. P., FRASER, J. A. & HUTLEY, E. 2018. Comparison of stool collection and storage on Whatman FTA Elute cards versus frozen stool for enteropathogen detection using the TaqMan Array Card PCR assay. *PloS One,* 13.
LANDER, E. S., LINTON, L. M., BIRREN, B., NUSBAUM, C., ZODY, M. C., BALDWIN, J., DEVON, K., DEWAR, K., DOYLE, M. & FITZHUGH, W. 2001. Initial sequencing and analysis of the human genome. *Nature,* 409**,** 860-921.
LATZ, E., SCHOENEMEYER, A., VISINTIN, A., FITZGERALD, K. A., MONKS, B. G., KNETTER, C. F., LIEN, E., NILSEN, N. J., ESPEVIK, T. & GOLENBOCK, D. T. 2004. TLR9 signals after translocating from the ER to CpG DNA in the lysosome. *Nature Immunology,* 5.
LAU, H. & HURT, A. C. 2016. Assessment of the RNASound RNA Sampling Card for the Preservation of Influenza Virus RNA. *Frontiers in Microbiology,* 7**,** 1736.
LAUW, F. N., CAFFREY, D. R. & GOLENBOCK, D. T. 2005. Of mice and man: TLR11 (finally) finds profilin. *Trends in Immunology,* 26**,** 509-511.

LAZARUS, R., KLIMECKI, W. T., RABY, B. A., VERCELLI, D., PALMER, L. J., KWIATKOWSKI, D. J., SILVERMAN, E. K., MARTINEZ, F. & WEISS, S. T. 2003. Single-nucleotide polymorphisms in the Toll-like receptor 9 gene (TLR9): frequencies, pairwise linkage disequilibrium, and haplotypes in three US ethnic groups and exploratory case–control disease association studies. *Genomics,* 81**,** 85-91.

LEE, S. M. Y., KOK, K.-H., JAUME, M., CHEUNG, T. K. W., YIP, T.-F., LAI, J. C. C., GUAN, Y., WEBSTER, R. G., JIN, D.-Y. & PEIRIS, J. S. M. 2014. Toll-like receptor 10 is involved in induction of innate immune responses to influenza virus infection. *Proceedings of the National Academy of Sciences,* 111**,** 3793-3798.

LEGER, M. & VIENNE, M. 1919. Epizootie à trypanosomes chez les bovidés de la Guyane Française. *Bull Soc Pathol Exot,* 12**,** 258-266.

LEHNER, M. D., MORATH, S., MICHELSEN, K. S., SCHUMANN, R. R. & HARTUNG, T. 2001. Induction of cross-tolerance by lipopolysaccharide and highly purified lipoteichoic acid via different Toll-like receptors independent of paracrine mediators. *The Journal of Immunology,* 166**,** 5161-5167.

LEMAITRE, B., NICOLAS, E., MICHAUT, L., REICHHART, J.-M. & HOFFMANN, J. A. 1996. The dorsoventral regulatory gene cassette spätzle/Toll/cactus controls the potent antifungal response in Drosophila adults. *Cell,* 86**,** 973-983.

LI, Z. 2012. Regulation of the cell division cycle in Trypanosoma brucei. *Eukaryotic cell,* 11**,** 1180-1190.

LIM, J. M., KIM, G. & LEVINE, R. L. 2019. Methionine in Proteins: It's Not Just for Protein Initiation Anymore. *Neurochem Res,* 44**,** 247-257.

LITI, G., BA, A. N. N., BLYTHE, M., MÜLLER, C. A., BERGSTRÖM, A., CUBILLOS, F. A., DAFHNIS-CALAS, F., KHOSHRAFTAR, S., MALLA, S. & MEHTA, N. 2013. High quality de novo sequencing and assembly of the Saccharomyces arboricolus genome. *BMC genomics,* 14**,** 69.

LIU, P., MORRISON, C., WANG, L., XIONG, D., VEDELL, P., CUI, P., HUA, X., DING, F., LU, Y. & JAMES, M. 2012. Identification of somatic mutations in non-small cell lung carcinomas using whole-exome sequencing. *Carcinogenesis,* 33**,** 1270-1276.

LLOYD, L. & JOHNSON, W. 1924. The trypanosome infections of tsetse-flies in Northern Nigeria and a new method of estimation. *Bulletin of Entomological Research,* 14**,** 265-288.

LOFTUS, R., HUGH, D. M., NGERE, L., BALAIN, D., BADI, A., BRADLEY, D. & CUNNINGHAM, E. 1994. Mitochondrial genetic variation in European, African and Indian cattle populations. *Animal Genetics,* 25**,** 265-271.

LOHRER, H. D. & TANGEN, U. 2000. Investigations into the molecular effects of single nucleotide polymorphism. *Pathobiology,* 68**,** 283-290.

LUCKINS, A. & GRAY, A. 1978. An extravascular site of development of Trypanosoma congolense. *Nature,* 272**,** 613-614.

LUCKINS, A., SUTHERLAND, D., MWANGI, D. & HOPKINS, J. 1994. Early stages of infection with Trypanosoma congolense: parasite kinetics and expression of metacyclic variable antigen types. *Acta Tropica,* 58**,** 199-206.

LUKEŠ, J., GUILBRIDE, D. L., VOTÝPKA, J., ZÍKOVÁ, A., BENNE, R. & ENGLUND, P. T. 2002. Kinetoplast DNA network: evolution of an improbable structure. *Eukaryotic Cell,* 1**,** 495-502.

LUKEŠ, J., HASHIMI, H. & ZÍKOVÁ, A. 2005. Unexplained complexity of the mitochondrial genome and transcriptome in kinetoplastid flagellates. *Current Genetics,* 48**,** 277-299.

LUN, Z. & DESSER, S. 1995. Is the broad range of hosts and geographical distribution of Trypanosoma evansi attributable to the loss of maxicircle kinetoplast DNA? *Parasitology Today,* 11**,** 131-133.

LUNDIN, S., STRANNEHEIM, H., PETTERSSON, E., KLEVEBRING, D. & LUNDEBERG, J. 2010. Increased throughput by parallelization of library preparation for massive sequencing. *PloS One,* 5.

MA, Z., LEE, R. W., LI, B., KENNEY, P., WANG, Y., ERIKSON, J., GOYAL, S. & LAO, K. 2013. Isothermal amplification method for next-generation sequencing. *Proceedings of the National Academy of Sciences,* 110**,** 14320-14323.

MADEIRA, M. F., ALMEIDA, A. B., BARROS, J. H., OLIVEIRA, T. S., SOUSA, V. R., ALVES, A. S., MIRANDA, L. F., SCHUBACH, A. O. & MARZOCHI, M. C. 2014. Trypanosoma caninum, a

new parasite described in dogs in Brazil: aspects of natural infection. *The Journal of Parasitology,* 100**,** 231-234.
MAGI, A., BENELLI, M., GOZZINI, A., GIROLAMI, F., TORRICELLI, F. & BRANDI, M. L. 2010. Bioinformatics for next generation sequencing data. *Genes,* 1**,** 294-307.
MAGONA, J., WALUBENGO, J. & ODIMIN, J. 2008. Acute haemorrhagic syndrome of bovine trypanosomosis in Uganda. *Acta tropica,* 107**,** 186-191.
MALELE, I. I., MAGWISHA, H. B., NYINGILILI, H. S., MAMIRO, K. A., RUKAMBILE, E. J., DAFFA, J. W., LYARUU, E. A., KAPANGE, L. A., KASILAGILA, G. K. & LWITIKO, N. K. 2011. Multiple Trypanosoma infections are common amongst Glossina species in the new farming areas of Rufiji district, Tanzania. *Parasites & Vectors,* 4**,** 217.
MALLA, M. A., DUBEY, A., KUMAR, A., YADAV, S., HASHEM, A. & ABD_ALLAH, E. F. 2019. Exploring the human microbiome: The potential future role of next-generation sequencing in disease diagnosis and treatment. *Frontiers in Immunology,* 9**,** 2868.
MANCUSO, G., GAMBUZZA, M., MIDIRI, A., BIONDO, C., PAPASERGI, S., AKIRA, S., TETI, G. & BENINATI, C. 2009. Bacterial recognition by TLR7 in the lysosomes of conventional dendritic cells. *Nature Immunology,* 10**,** 587-594.
MANDAL, R. K., GEORGE, G. P. & MITTAL, R. D. 2012. Association of Toll-like receptor (TLR) 2, 3 and 9 genes polymorphism with prostate cancer risk in North Indian population. *Molecular Biology Reports,* 39**,** 7263-7269.
MANSFIELD, J. M. 1977. Nonpathogenic trypanosomes of mammals. *Parasitic Protozoa,* 1**,** 297-327.
MANSFIELD, J. M. & PAULNOCK, D. M. 2005. Regulation of innate and acquired immunity in African trypanosomiasis. *Parasite Immunol,* 27**,** 361-71.
MARCY, Y., ISHOEY, T., LASKEN, R. S., STOCKWELL, T. B., WALENZ, B. P., HALPERN, A. L., BEESON, K. Y., GOLDBERG, S. M. & QUAKE, S. R. 2007. Nanoliter reactors improve multiple displacement amplification of genomes from single cells. *PLoS genetics,* 3**,** e155.
MARDIS, E. R. 2013. Next-generation sequencing platforms. *Annual review of analytical chemistry,* 6**,** 287-303.
MARGULIES, M., EGHOLM, M., ALTMAN, W. E., ATTIYA, S., BADER, J. S., BEMBEN, L. A., BERKA, J., BRAVERMAN, M. S., CHEN, Y.-J. & CHEN, Z. 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature,* 437**,** 376.
MARSTON, D. A., MCELHINNEY, L. M., ELLIS, R. J., HORTON, D. L., WISE, E. L., LEECH, S. L., DAVID, D., DE LAMBALLERIE, X. & FOOKS, A. R. 2013. Next generation sequencing of viral RNA genomes. *BMC genomics,* 14**,** 444.
MATHUR, R., OH, H., ZHANG, D., PARK, S.-G., SEO, J., KOBLANSKY, A., HAYDEN, M. S. & GHOSH, S. 2012. A mouse model of Salmonella typhi infection. *Cell,* 151**,** 590-602.
MATSUMOTO, H., INADA, S., KOBAYASHI, E., ABE, T., HASEBE, H., SASAZAKI, S., OYAMA, K. & MANNEN, H. 2012. Identification of SNPs in the FASN gene and their effect on fatty acid milk composition in Holstein cattle. *Livestock Science,* 144**,** 281-284.
MATTHEWS, K. R. 2005. The developmental cell biology of Trypanosoma brucei. *Journal of Cell Science,* 118**,** 283-290.
MAXFIELD, L. & BERMUDEZ, R. 2020. Trypanosomiasis. *StatPearls.* Treasure Island (FL).
MCCLURE, M. C., MCKAY, S. D., SCHNABEL, R. D. & TAYLOR, J. F. 2009. Assessment of DNA extracted from FTA® cards for use on the Illumina iSelect BeadChip. *BMC research notes,* 2**,** 107.
MCELRATH, M., BOCHUD, P. & PINE, S. 2009. Polymorphisms in toll-like receptor 4 and toll-like receptor 9 influence viral load in a seroincident cohort of HIV-1-infected individuals.
MCEVOY, C. R., SEMPLE, T., YELLAPU, B., CHOONG, D. Y., XU, H., MIR ARNAU, G., FELLOWES, A. P. & FOX, S. B. 2020. Improved next-generation sequencing pre-capture library yields and sequencing parameters using on-bead PCR. *BioTechniques*.
MCGUIRE, K., JONES, M., WERLING, D., WILLIAMS, J. L., GLASS, E. J. & JANN, O. 2006. Radiation hybrid mapping of all 10 characterized bovine Toll-like receptors. *Animal Genetics,* 37**,** 47-50.
MEANS, T. K., LIEN, E., YOSHIMURA, A., WANG, S., GOLENBOCK, D. T. & FENTON, M. J. 1999. The CD14 ligands lipoarabinomannan and lipopolysaccharide differ in their requirement for Toll-like receptors. *The Journal of Immunology,* 163**,** 6748-6755.

MEDVEDEV, A. E. 2013. Toll-like receptor polymorphisms, inflammatory and infectious diseases, allergies, and cancer. *Journal of Interferon & Cytokine Research,* 33**,** 467-484.
MEDZHITOV, R. & JANEWAY, C. 2000. Innate immune recognition: mechanisms and pathways. *Immunological Reviews,* 173**,** 89-97.
MEDZHITOV, R. & JANEWAY, C. A. 1997. Innate immunity: the virtues of a nonclonal system of recognition. *Cell,* 91**,** 295-298.
MEDZHITOV, R., PRESTON-HURLBURT, P. & JANEWAY, C. A. 1997. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature,* 388**,** 394-397.
MEKATA, H., KONNAI, S., MINGALA, C. N., ABES, N. S., GUTIERREZ, C. A., DARGANTES, A. P., WITOLA, W. H., INOUE, N., ONUMA, M. & MURATA, S. 2013. Isolation, cloning, and pathologic analysis of Trypanosoma evansi field isolates. *Parasitology Research,* 112**,** 1513-1521.
MELISI, D., FRIZZIERO, M., TAMBURRINO, A., ZANOTTO, M., CARBONE, C., PIRO, G. & TORTORA, G. 2014. Toll-like receptor 9 agonists for cancer therapy. *Biomedicines,* 2**,** 211-228.
MELISI, D., XIA, Q., PARADISO, G., LING, J., MOCCIA, T., CARBONE, C., BUDILLON, A., ABBRUZZESE, J. L. & CHIAO, P. J. 2011. Modulation of pancreatic cancer chemoresistance by inhibition of TAK1. *Journal of the National Cancer Institute,* 103**,** 1190-1204.
METZKER, M. L. 2010. Sequencing technologies—the next generation. *Nature reviews genetics,* 11**,** 31-46.
MILLER, D., BRYANT, J., MADSEN, E. & GHIORSE, W. 1999. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Applied and Environmental Microbiology,* 65**,** 4715-4724.
MOGENSEN, T. H. 2009. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews,* 22**,** 240-273.
MOLOO, S. & GRAY, M. 1989. New observations on cyclical development of Trypanosoma vivax in Glossina. *Acta Tropica,* 46**,** 167-172.
MONTMAYEUR, A. M., NG, T. F. F., SCHMIDT, A., ZHAO, K., MAGAÑA, L., IBER, J., CASTRO, C. J., CHEN, Q., HENDERSON, E. & RAMOS, E. 2017. High-throughput next-generation sequencing of polioviruses. *Journal of Clinical Microbiology,* 55**,** 606-615.
MOORE, D. A., EDWARDS, M., ESCOMBE, R., AGRANOFF, D., BAILEY, J. W., SQUIRE, S. B. & CHIODINI, P. L. 2002. African trypanosomiasis in travelers returning to the United Kingdom. *Emerging Infectious Diseases,* 8**,** 74.
MOORE, L. D., LE, T. & FAN, G. 2013. DNA methylation and its basic function. *Neuropsychopharmacology,* 38**,** 23-38.
MOURA, A., HODON, M. A., SOARES FILHO, P. M., ISSA, M. D. A., OLIVEIRA, A. P. F. D. & FONSECA JÚNIOR, A. A. 2016. Comparison of nine DNA extraction methods for the diagnosis of bovine tuberculosis by real time PCR. *Ciência Rural,* 46**,** 1223-1228.
MUHANGUZI, D., MUGENYI, A., BIGIRWA, G., KAMUSIIME, M., KITIBWA, A., AKURUT, G. G., OCHWO, S., AMANYIRE, W., OKECH, S. G. & HATTENDORF, J. 2017. African animal trypanosomiasis as a constraint to livestock health and production in Karamoja region: a detailed qualitative and quantitative assessment. *BMC veterinary research,* 13**,** 355.
MUKHERJEE, S., HUDA, S. & SINHA BABU, S. P. 2019. Toll-like receptor polymorphism in host immune response to infectious diseases: A review. *Scand J Immunol,* 90**,** e12771.
MUNGUBE, E. O., VITOULEY, H. S., ALLEGYE-CUDJOE, E., DIALL, O., BOUCOUM, Z., DIARRA, B., SANOGO, Y., RANDOLPH, T., BAUER, B. & ZESSIN, K.-H. 2012. Detection of multiple drug-resistant Trypanosoma congolense populations in village cattle of south-east Mali. *Parasites & Vectors,* 5**,** 155.
MURRAY, G. G., WOOLHOUSE, M. E., TAPIO, M., MBOLE-KARIUKI, M. N., SONSTEGARD, T. S., THUMBI, S. M., JENNINGS, A. E., VAN WYK, I. C., CHASE-TOPPING, M., KIARA, H., TOYE, P., COETZER, K., DE, C. B. B. M. & HANOTTE, O. 2013. Genetic susceptibility to infectious disease in East African Shorthorn Zebu: a genome-wide analysis of the effect of heterozygosity and exotic introgression. *BMC Evol Biol,* 13**,** 246.
MUZIO, M., BOSISIO, D., POLENTARUTTI, N., D'AMICO, G., STOPPACCIARO, A., MANCINELLI, R., VAN'T VEER, C., PENTON-ROL, G., RUCO, L. P. & ALLAVENA, P. 2000. Differential

expression and regulation of toll-like receptors (TLR) in human leukocytes: selective expression of TLR3 in dendritic cells. *The Journal of Immunology,* 164**,** 5998-6004.
NAKA, T., FUJIMOTO, M., TSUTSUI, H. & YOSHIMURA, A. 2005. Negative regulation of cytokine and TLR signalings by SOCS and others. *Advances in immunology,* 87**,** 61-122.
NAKAGAWA, R., NAKA, T., TSUTSUI, H., FUJIMOTO, M., KIMURA, A., ABE, T., SEKI, E., SATO, S., TAKEUCHI, O. & TAKEDA, K. 2002. SOCS-1 participates in negative regulation of LPS responses. *Immunity,* 17**,** 677-687.
NAKAMURA, K., OSHIMA, T., MORIMOTO, T., IKEDA, S., YOSHIKAWA, H., SHIWA, Y., ISHIKAWA, S., LINAK, M. C., HIRAI, A. & TAKAHASHI, H. 2011. Sequence-specific error profile of Illumina sequencers. *Nucleic Acids Research,* 39**,** e90-e90.
NAKATO, R. & SHIRAHIGE, K. 2016. Recent advances in ChIP-seq analysis: from quality management to whole-genome annotation. *Brief Bioinform*.
NAMANGALA, B. & ODONGO, S. 2014. Animal African trypanosomosis in sub-Saharan Africa and beyond African borders. *Trypanosomes and trypanosomiasis.* Springer.
NATARAJAN, P., TRINH, T., MERTZ, L., GOLDSBOROUGH, M. & FOX, D. 2000. based archiving of mammalian and plant samples for RNA analysis. *Biotechniques,* 29**,** 1328-1333.
NDAO, M. 2009. Diagnosis of parasitic diseases: old and new approaches. *Interdiscip Perspect Infect Dis,* 2009**,** 278246.
NGOMTCHO, S. C. H., WEBER, J. S., BUM, E. N., GBEM, T. T., KELM, S. & ACHUKWI, M. D. 2017. Molecular screening of tsetse flies and cattle reveal different Trypanosoma species including T. grayi and T. theileri in northern Cameroon. *Parasites & Vectors,* 10**,** 631.
NIE, L., CAI, S.-Y., SHAO, J.-Z. & CHEN, J. 2018. Toll-like receptors, associated biological roles, and signaling networks in non-mammals. *Frontiers in immunology,* 9**,** 1523.
NJIRU, Z., OUMA, J., BATETA, R., NJERU, S., NDUNGU, K., GITONGA, P., GUYA, S. & TRAUB, R. 2011. Loop-mediated isothermal amplification test for Trypanosoma vivax based on satellite repeat DNA. *Veterinary Parasitology,* 180**,** 358-362.
NOIREAU, F., DIOSQUE, P. & JANSEN, A. M. 2009. Trypanosoma cruzi: adaptation to its vectors and its hosts. *Veterinary Research,* 40**,** 1-23.
NOVÁK, K. 2014. Functional polymorphisms in Toll-like receptor genes for innate immunity in farm animals. *Veterinary Immunology and Immunopathology,* 157**,** 1-11.
ODENIRAN, P. O. & ADEMOLA, I. O. 2018. A meta-analysis of the prevalence of African animal trypanosomiasis in Nigeria from 1960 to 2017. *Parasites & Vectors,* 11**,** 280.
OHNISHI, H., TOCHIO, H., KATO, Z., ORII, K. E., LI, A., KIMURA, T., HIROAKI, H., KONDO, N. & SHIRAKAWA, M. 2009. Structural basis for the multiple interactions of the MyD88 TIR domain in TLR4 signaling. *Proceedings of the National Academy of Sciences,* 106**,** 10260-10265.
OHTO, U., ISHIDA, H., SHIBATA, T., SATO, R., MIYAKE, K. & SHIMIZU, T. 2018. Toll-like receptor 9 contains two DNA binding sites that function cooperatively to promote receptor dimerization and activation. *Immunity,* 48**,** 649-658. e4.
OHTO, U., SHIBATA, T., TANJI, H., ISHIDA, H., KRAYUKHINA, E., UCHIYAMA, S., MIYAKE, K. & SHIMIZU, T. 2015. Structural basis of CpG and inhibitory DNA recognition by Toll-like receptor 9. *Nature,* 520**,** 702-705.
OLIVEIRA, J., HERNÁNDEZ-GAMBOA, J., JIMÉNEZ-ALFARO, C., ZELEDÓN, R., BLANDÓN, M. & URBINA, A. 2009. First report of Trypanosoma vivax infection in dairy cattle from Costa Rica. *Veterinary Parasitology,* 163**,** 136-139.
OLIVEIRA, L. B., LOUVANTO, K., RAMANAKUMAR, A. V., FRANCO, E. L. & VILLA, L. L. 2013. Polymorphism in the promoter region of the Toll-like receptor 9 gene and cervical human papillomavirus infection. *The Journal of General Virology,* 94**,** 1858.
ONJI, M., KANNO, A., SAITOH, S.-I., FUKUI, R., MOTOI, Y., SHIBATA, T., MATSUMOTO, F., LAMICHHANE, A., SATO, S. & KIYONO, H. 2013. An essential role for the N-terminal fragment of Toll-like receptor 9 in DNA sensing. *Nature Communications,* 4**,** 1949.
ONYEKWELU, K. C. 2019. Life Cycle of Trypanosoma cruzi in the Invertebrate and the Vertebrate Hosts. *Biology of Trypanosoma cruzi.* IntechOpen.
OOI, C.-P., SCHUSTER, S., CREN-TRAVAILLÉ, C., BERTIAUX, E., COSSON, A., GOYARD, S., PERROT, S. & ROTUREAU, B. 2016. The cyclical development of Trypanosoma vivax in the

tsetse fly involves an asymmetric division. *Frontiers in Cellular and Infection Microbiology,* 6**,** 115.

OPITZ, B., SCHRÖDER, N. W. J., SPREITZER, I., MICHELSEN, K. S., KIRSCHNING, C. J., HALLATSCHEK, W., ZÄHRINGER, U., HARTUNG, T., GÖBEL, U. B. & SCHUMANN, R. R. 2001. Toll-like receptor-2 mediates Treponema glycolipid and lipoteichoic acid-induced NF-κB translocation. *Journal of Biological Chemistry,* 276**,** 22041-22047.

ORDULU, Z., KAMMIN, T., BRAND, H., PILLALAMARRI, V., REDIN, C. E., COLLINS, R. L., BLUMENTHAL, I., HANSCOM, C., PEREIRA, S., BRADLEY, I., CRANDALL, B. F., GERROL, P., HAYDEN, M. A., HUSSAIN, N., KANENGISSER-PINES, B., KANTARCI, S., LEVY, B., MACERA, M. J., QUINTERO-RIVERA, F., SPIEGEL, E., STEVENS, B., ULM, J. E., WARBURTON, D., WILKINS-HAUG, L. E., YACHELEVICH, N., GUSELLA, J. F., TALKOWSKI, M. E. & MORTON, C. C. 2016. Structural Chromosomal Rearrangements Require Nucleotide-Level Resolution: Lessons from Next-Generation Sequencing in Prenatal Diagnosis. *Am J Hum Genet,* 99**,** 1015-1033.

OSÓRIO, A. L. A. R., MADRUGA, C. R., DESQUESNES, M., SOARES, C. O., RIBEIRO, L. R. R. & COSTA, S. C. G. D. 2008. Trypanosoma (Duttonella) vivax: its biology, epidemiology, pathogenesis, and introduction in the New World-a review. *Memórias do Instituto Oswaldo Cruz,* 103**,** 1-13.

OTTE, M. & ABUABARA, J. 1991. Transmission of South American Trypanosoma vivax by the neotropical horsefly Tabanus nebulosus. *Acta Tropica,* 49**,** 73-76.

OVERATH, P., HAAG, J., MAMEZA, M. & LISCHKE, A. 1999. Freshwater fish trypanosomes: definition of two types, host control by antibodies and lack of antigenic variation. *Parasitology,* 119**,** 591-601.

PALM, N. W. & MEDZHITOV, R. 2009. Pattern recognition receptors and control of adaptive immunity. *Immunological Reviews,* 227**,** 221-233.

PALSSON-MCDERMOTT, E. M., DOYLE, S. L., MCGETTRICK, A. F., HARDY, M., HUSEBYE, H., BANAHAN, K., GONG, M., GOLENBOCK, D., ESPEVIK, T. & O'NEILL, L. A. 2009. TAG, a splice variant of the adaptor TRAM, negatively regulates the adaptor MyD88–independent TLR4 pathway. *Nature Immunology,* 10**,** 579.

PANDEY, N. O., CHAUHAN, A. V., RAITHATHA, N. S., PATEL, P. K., KHANDELWAL, R., DESAI, A. N., CHOXI, Y., KAPADIA, R. S. & JAIN, N. D. 2019. Association of TLR4 and TLR9 polymorphisms and haplotypes with cervical cancer susceptibility. *Sci Rep,* 9**,** 9729.

PARADOWSKA, E., JABLONSKA, A., STUDZINSKA, M., SKOWRONSKA, K., SUSKI, P., WISNIEWSKA-LIGIER, M., WOZNIAKOWSKA-GESICKA, T., NOWAKOWSKA, D., GAJ, Z., WILCZYNSKI, J. & LESNIKOWSKI, Z. J. 2016. TLR9 -1486T/C and 2848C/T SNPs Are Associated with Human Cytomegalovirus Infection in Infants. *PLoS One,* 11**,** e0154100.

PARSONS, L. & BRIGHT, J.-A. 2012. A manual and automated method for the forensic analysis of DNA from buccal samples on Whatman Indicating FTA Elute Cards. *Australian Journal of Forensic Sciences,* 44**,** 393-402.

PARSONS, M. 2004. Glycosomes: parasites and the divergence of peroxisomal purpose. *Molecular microbiology,* 53**,** 717-724.

PELAK, K., SHIANNA, K. V., GE, D., MAIA, J. M., ZHU, M., SMITH, J. P., CIRULLI, E. T., FELLAY, J., DICKSON, S. P. & GUMBS, C. E. 2010. The characterization of twenty sequenced human genomes. *PLoS Genetics,* 6.

PETERSEN, C. A., KRUMHOLZ, K. A. & BURLEIGH, B. A. 2005. Toll-like receptor 2 regulates interleukin-1beta-dependent cardiomyocyte hypertrophy triggered by Trypanosoma cruzi. *Infect Immun,* 73**,** 6974-80.

PITARQUE, M., VON RICHTER, O., OKE, B., BERKKAN, H., OSCARSON, M. & INGELMAN-SUNDBERG, M. 2001. Identification of a single nucleotide polymorphism in the TATA box of the CYP2A6 gene: impairment of its promoter activity. *Biochemical and biophysical research communications,* 284**,** 455-460.

POLTORAK, A., HE, X., SMIRNOVA, I., LIU, M. Y., VAN HUFFEL, C., DU, X., BIRDWELL, D., ALEJOS, E., SILVA, M., GALANOS, C., FREUDENBERG, M., RICCIARDI-CASTAGNOLI,

P., LAYTON, B. & BEUTLER, B. 1998. Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science,* 282**,** 2085-8.
PORITZ, M. A. & RIRIE, K. M. 2014. Getting things backwards to prevent primer dimers. *The Journal of molecular diagnostics: JMD,* 16**,** 159.
PUJOL-RIQUÉ, M., DEROUIN, F., GARCIA-QUINTANILLA, A., VALLS, M., MIRO, J. & DE ANTA, M. J. 1999. Design of a one-tube hemi-nested PCR for detection of Toxoplasma gondii and comparison of three DNA purification methods. *Journal of medical microbiology,* 48**,** 857-862.
PULIDO-LANDÍNEZ, M., LAVINIKI, V., SÁNCHEZ-INGUNZA, R., GUARD, J. & DO NASCIMENTO, V. P. 2012. Use of FTA cards for the transport of DNA samples of Salmonella spp. from poultry products from Southern Brazil. *Acta Scientiae Veterinariae,* 40.
QUAIL, M. A., GU, Y., SWERDLOW, H. & MAYHO, M. 2012. Evaluation and optimisation of preparative semi-automated electrophoresis systems for Illumina library preparation. *Electrophoresis,* 33**,** 3521-3528.
RADWANSKA, M., VEREECKE, N., DELEEUW, V., PINTO, J. & MAGEZ, S. 2018. Salivarian Trypanosomosis: A Review of Parasites Involved, Their Global Distribution and Their Interaction With the Innate and Adaptive Mammalian Host Immune System. *Frontiers in immunology,* 9.
RAETZ, M., KIBARDIN, A., STURGE, C. R., PIFER, R., LI, H., BURSTEIN, E., OZATO, K., LARIN, S. & YAROVINSKY, F. 2013. Cooperation of TLR12 and TLR11 in the IRF8-dependent IL-12 response to Toxoplasma gondii profilin. *The Journal of Immunology,* 191**,** 4818-4827.
RAHIKAINEN, A.-L., PALO, J. U., DE LEEUW, W., BUDOWLE, B. & SAJANTILA, A. 2016. DNA quality and quantity from up to 16 years old post-mortem blood stored on FTA cards. *Forensic science international,* 261**,** 148-153.
RAINA, A., KUMAR, R., SRIDHAR, V. R. & SINGH, R. 1985. Oral transmission of Trypanosoma evansi infection in dogs and mice. *Veterinary Parasitology,* 18**,** 67-69.
RAVEL, S., PATREL, D., KOFFI, M., JAMONNEAU, V. & CUNY, G. 2006. Cyclical transmission of Trypanosoma brucei gambiense in Glossina palpalis gambiensis displays great differences among field isolates. *Acta tropica,* 100**,** 151-155.
REBOUÇAS, E. D. L., COSTA, J. J. D. N., PASSOS, M. J., PASSOS, J. R. D. S., HURK, R. V. D. & SILVA, J. R. V. 2013. Real time PCR and importance of housekeepings genes for normalization and quantification of mRNA expression in different tissues. *Brazilian Archives of Biology and Technology,* 56**,** 143-154.
RIGBY, R. E., WEBB, L. M., MACKENZIE, K. J., LI, Y., LEITCH, A., REIJNS, M. A. M., LUNDIE, R. J., REVUELTA, A., DAVIDSON, D. J. & DIEBOLD, S. 2014. RNA: DNA hybrids are a novel molecular pattern sensed by TLR9. *The EMBO journal***,** e201386117.
RODRIGUES, A., FIGHERA, R., SOUZA, T., SCHILD, A. & BARROS, C. 2009. Neuropathology of naturally occurring Trypanosoma evansi infection of horses. *Veterinary pathology,* 46**,** 251-258.
RODRIGUES, A. C., ORTIZ, P. A., COSTA-MARTINS, A. G., NEVES, L., GARCIA, H. A., ALVES, J. M., CAMARGO, E. P., ALFIERI, S. C., GIBSON, W. & TEIXEIRA, M. M. 2014. Congopain genes diverged to become specific to Savannah, Forest and Kilifi subgroups of Trypanosoma congolense, and are valuable for diagnosis, genotyping and phylogenetic inferences. *Infection, Genetics and Evolution,* 23**,** 20-31.
RODRIGUES, J. C. F., GODINHO, J. L. P. & DE SOUZA, W. 2014. Biology of human pathogenic trypanosomatids: epidemiology, lifecycle and ultrastructure. *Proteins and Proteomics of Leishmania and Trypanosoma.* Springer.
ROEDER, A., KIRSCHNING, C. J., RUPEC, R. A., SCHALLER, M. & KORTING, H. C. 2004. Toll-like receptors and innate antifungal responses. *Trends in microbiology,* 12**,** 44-49.
ROSZAK, A., LIANERI, M., SOWIŃSKA, A. & JAGODZIŃSKI, P. P. 2012. Involvement of Toll-like Receptor 9 polymorphism in cervical cancer development. *Molecular biology reports,* 39**,** 8425-8430.
RUIZ, J. P., NYINGILILI, H. S., MBATA, G. H. & MALELE, I. I. 2015. The role of domestic animals in the epidemiology of human african trypanosomiasis in Ngorongoro conservation area, Tanzania. *Parasites & vectors,* 8**,** 510.

RYKALINA, V. N., SHADRIN, A. A., AMSTISLAVSKIY, V. S., ROGAEV, E. I., LEHRACH, H. & BORODINA, T. A. 2014. Exome sequencing from nanogram amounts of starting DNA: comparing three approaches. *PloS One,* 9**,** e101154.

SAIEG, M. A., GEDDIE, W. R., BOERNER, S. L., LIU, N., TSAO, M., ZHANG, T., KAMEL-REID, S. & DA CUNHA SANTOS, G. 2012. The use of FTA cards for preserving unfixed cytological material for high-throughput molecular analysis. *Cancer cytopathology,* 120**,** 206-214.

SAINT PIERRE, A. & GÉNIN, E. 2014. How important are rare variants in common disease? *Briefings in functional genomics,* 13**,** 353-361.

SALZBERG, S. L. & YORKE, J. A. 2005. Beware of mis-assembled genomes. *Bioinformatics,* 21**,** 4320-4321.

SANDERS, M. S., VAN WELL, G. T. J., OUBURG, S., LUNDBERG, P. S., VAN FURTH, A. M. & MORRÉ, S. A. 2011. Single nucleotide polymorphisms in TLR9 are highly associated with susceptibility to bacterial meningitis in children. *Clinical infectious diseases,* 52**,** 475-480.

SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the national academy of sciences,* 74**,** 5463-5467.

SANTOS, V. R. C. D., MEIS, J. D., SAVINO, W., ANDRADE, J. A. A., VIEIRA, J. R. D. S., COURA, J. R. & JUNQUEIRA, A. C. V. 2018. Acute Chagas disease in the state of Pará, Amazon Region: is it increasing? *Memórias do Instituto Oswaldo Cruz,* 113.

SARAFIDOU, T., STAMATIS, C., KALOZOUMI, G., SPYROU, V., FTHENAKIS, G. C., BILLINIS, C. & MAMURIS, Z. 2013. Toll like receptor 9 (TLR9) polymorphism G520R in sheep is associated with seropositivity for small ruminant lentivirus. *PloS One,* 8**,** e63901.

SASAI, M., LINEHAN, M. M. & IWASAKI, A. 2010. Bifurcation of Toll-like receptor 9 signaling by adaptor protein 3. *Science,* 329**,** 1530-1534.

SATO, M. P., OGURA, Y., NAKAMURA, K., NISHIDA, R., GOTOH, Y., HAYASHI, M., HISATSUNE, J., SUGAI, M., TAKEHIKO, I. & HAYASHI, T. 2019. Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. *DNA Research,* 26**,** 391-398.

SAUVAGE, C., BIERNE, N., LAPEGUE, S. & BOUDRY, P. 2007. Single nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster Crassostrea gigas. *Gene,* 406**,** 13-22.

SCHADT, E. E., TURNER, S. & KASARSKIS, A. 2010. A window into third-generation sequencing. *Human molecular genetics,* 19**,** R227-R240.

SCHMIDT, G. D., ROBERTS, L. S. & JANOVY, J. 1977. *Foundations of parasitology*, Mosby Saint Louis.

SCHNARE, M., BARTON, G. M., HOLT, A. C., TAKEDA, K., AKIRA, S. & MEDZHITOV, R. 2001. Toll-like receptors control activation of adaptive immune responses. *Nat Immunol,* 2**,** 947-50.

SCHOFIELD, C. 2000. Trypanosoma cruzi--the vector-parasite paradox. *Memórias do Instituto Oswaldo Cruz,* 95**,** 535-544.

SCHWANDNER, R., DZIARSKI, R., WESCHE, H., ROTHE, M. & KIRSCHNING, C. J. 1999. Peptidoglycan-and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2. *Journal of Biological Chemistry,* 274**,** 17406-17409.

SEABURY, C. M., CARGILL, E. J. & WOMACK, J. E. 2007. Sequence variability and protein domain architectures for bovine Toll-like receptors 1, 5, and 10. *Genomics,* 90**,** 502-515.

SELBY, R., BARDOSH, K., PICOZZI, K., WAISWA, C. & WELBURN, S. C. 2013. Cattle movements and trypanosomes: restocking efforts and the spread of Trypanosoma brucei rhodesiense sleeping sickness in post-conflict Uganda. *Parasites & vectors,* 6**,** 281.

SHASTRY, B. S. 2002. SNP alleles in human disease and evolution. *Journal of human genetics,* 47**,** 561.

SHASTRY, B. S. 2003. SNPs and haplotypes: genetic markers for disease and drug response. *International journal of molecular medicine,* 11**,** 379-382.

SHENDURE, J. & JI, H. 2008. Next-generation DNA sequencing. *Nature biotechnology,* 26**,** 1135.

SHENDURE, J., MITRA, R. D., VARMA, C. & CHURCH, G. M. 2004. Advanced sequencing technologies: methods and goals. *Nature Reviews Genetics,* 5**,** 335-344.

SHENDURE, J., PORRECA, G. J., REPPAS, N. B., LIN, X., MCCUTCHEON, J. P., ROSENBAUM, A. M., WANG, M. D., ZHANG, K., MITRA, R. D. & CHURCH, G. M. 2005. Accurate multiplex polony sequencing of an evolved bacterial genome. *Science,* 309**,** 1728-1732.

SHIKANAI-YASUDA, M. A. & CARVALHO, N. B. 2012. Oral transmission of Chagas disease. *Clinical Infectious Diseases,* 54**,** 845-852.
SHINTANI, Y., KAPOOR, A., KANEKO, M., SMOLENSKI, R. T., D'ACQUISTO, F., COPPEN, S. R., HARADA-SHOJI, N., LEE, H. J., THIEMERMANN, C. & TAKASHIMA, S. 2013. TLR9 mediates cellular protection by modulating energy metabolism in cardiomyocytes and neurons. *Proceedings of the National Academy of Sciences,* 110**,** 5109-5114.
SIEGEL, C. S., STEVENSON, F. O. & ZIMMER, E. A. 2017. Evaluation and comparison of FTA card and CTAB DNA extraction methods for non-agricultural taxa. *Applications in plant sciences,* 5**,** 1600109.
SILVA, R., AROSEMENA, N., HERRERA, H., SAHIB, C. & FERREIRA, M. 1995. Outbreak of trypanosomosis due to Trypanosoma evansi in horses of Pantanal Mato-grossense, Brazil. *Veterinary Parasitology,* 60**,** 167-171.
SILVA PEREIRA, S., TRINDADE, S., DE NIZ, M. & FIGUEIREDO, L. M. 2019. Tissue tropism in parasitic diseases. *Open biology,* 9**,** 190036.
SIMO, G., SOBGWI, P. F., NJITCHOUANG, G. R., NJIOKOU, F., KUIATE, J. R., CUNY, G. & ASONGANYI, T. 2013. Identification and genetic characterization of Trypanosoma congolense in domestic animals of Fontem in the South-West region of Cameroon. *Infection, Genetics and Evolution,* 18**,** 66-73.
SIMONELLI, P., TROEDSSON, C., NEJSTGAARD, J. C., ZECH, K., LARSEN, J. B. & FRISCHER, M. E. 2009. Evaluation of DNA extraction and handling procedures for PCR-based copepod feeding studies. *Journal of plankton research,* 31**,** 1465-1474.
SINGH, K., SINGH, V. K., AGRAWAL, N. K., GUPTA, S. K. & SINGH, K. 2013. Association of Toll-like receptor 4 polymorphisms with diabetic foot ulcers and application of artificial neural network in DFU risk assessment in type 2 diabetes patients. *Biomed Res Int,* 2013**,** 318686.
SINGH, V. & CHHABRA, M. 2008. Trypanosomoses (surra) in India: an Update. *J Parasit Dis,* 32**,** 104-110.
SINGH, V., MISHRA, S., RAO, G., JAIN, A. K., DIXIT, V., GULATI, A. K., MAHAJAN, D., MCCLELLAND, M. & NATH, G. 2008. Evaluation of nested PCR in detection of Helicobacter pylori targeting a highly conserved gene: HSP60. *Helicobacter,* 13**,** 30-34.
SKAUG, B., CHEN, J., DU, F., HE, J., MA, A. & CHEN, Z. J. 2011. Direct, noncatalytic mechanism of IKK inhibition by A20. *Molecular cell,* 44**,** 559-571.
SKEVAKI, C., PARARAS, M., KOSTELIDOU, K., TSAKRIS, A. & ROUTSIAS, J. 2015. Single nucleotide polymorphisms of T oll-like receptors and susceptibility to infectious diseases. *Clinical & Experimental Immunology,* 180**,** 165-177.
SKONIECZNA, K., STYCZYNSKI, J., KRENSKA, A., WYSOCKI, M., JAKUBOWSKA, A. & GRZYBOWSKI, T. 2016. RNA isolation from bloodstains collected on FTA cards - application in clinical and forensic genetics. *Arch Med Sadowej Kryminol,* 66**,** 244-254.
SLATKO, B. E., GARDNER, A. F. & AUSUBEL, F. M. 2018. Overview of next-generation sequencing technologies. *Current protocols in molecular biology,* 122**,** e59.
SMITH, L. M. & BURGOYNE, L. A. 2004. Collecting, archiving and processing DNA from wildlife samples using FTA® databasing paper. *BMC ecology,* 4**,** 4.
SMITH, T. K., BRINGAUD, F., NOLAN, D. P. & FIGUEIREDO, L. M. 2017. Metabolic reprogramming during the Trypanosoma brucei life cycle. *F1000Research,* 6.
SOUZA, P. F. P. D., PINE, R. C. D., RAMOS, F. L. D. P. & PINTO, A. Y. D. N. 2018. Atrial fibrillation in acute Chagas disease acquired via oral transmission: a case report. *Revista da Sociedade Brasileira de Medicina Tropical,* 51**,** 397-400.
SPITS, C., LE CAIGNEC, C., DE RYCKE, M., VAN HAUTE, L., VAN STEIRTEGHEM, A., LIEBAERS, I. & SERMON, K. 2006. Whole-genome multiple displacement amplification from single cells. *Nature protocols,* 1**,** 1965.
SUN, L., SONG, Y., RIAZ, H., YANG, H., HUA, G., GUO, A. & YANG, L. 2012. Polymorphisms in toll-like receptor 1 and 9 genes and their association with tuberculosis susceptibility in Chinese Holstein cattle. *Veterinary immunology and immunopathology,* 147**,** 195-201.

STANGEGAARD, M., BØRSTING, C., FERRERO-MILIANI, L., FRANK-HANSEN, R., POULSEN, L., HANSEN, A. J. & MORLING, N. 2013. Evaluation of four automated protocols for extraction of DNA from FTA cards. *Journal of laboratory automation,* 18**,** 404-410.

STEVERDING, D. 2008. The history of African trypanosomiasis. *Parasites & vectors,* 1**,** 3.

STRANNEHEIM, H., WERNE, B., SHERWOOD, E. & LUNDEBERG, J. 2011. Scalable transcriptome preparation for massive parallel sequencing. *PloS One,* 6.

SUN, L., SONG, Y., RIAZ, H., YANG, H., HUA, G., GUO, A. & YANG, L. 2012. Polymorphisms in toll-like receptor 1 and 9 genes and their association with tuberculosis susceptibility in Chinese Holstein cattle. *Veterinary immunology and immunopathology,* 147**,** 195-201.

SUZUKI, N., SUZUKI, S., DUNCAN, G. S., MILLAR, D. G., WADA, T., MIRTSOS, C., TAKADA, H., WAKEHAM, A., ITIE, A. & LI, S. 2002. Severe impairment of interleukin-1 and Toll-like receptor signalling in mice lacking IRAK-4. *Nature,* 416**,** 750-756.

TAKEDA, K. & AKIRA, S. 2005. Toll-like receptors in innate immunity. *Int Immunol,* 17**,** 1-14.

TAKEDA, K., KAISHO, T. & AKIRA, S. 2003. Toll-like receptors. *Annual review of immunology,* 21**,** 335-376.

TARAZONA, S., GARCÍA-ALCALDE, F., DOPAZO, J., FERRER, A. & CONESA, A. 2011. Differential expression in RNA-seq: a matter of depth. *Genome research,* 21**,** 2213-2223.

TARLETON, R. L. 2007. Immune system recognition of Trypanosoma cruzi. *Curr Opin Immunol,* 19**,** 430-4.

TERRY, R. S., SMITH, J. E., DUNCANSON, P. & HIDE, G. 2001. MGE-PCR: a novel approach to the analysis of Toxoplasma gondii strain differentiation using mobile genetic elements. *International journal for parasitology,* 31**,** 155-161.

TESFAYE, D., SPEYBROECK, N., DE DEKEN, R. & THYS, E. 2012. Economic burden of bovine trypanosomosis in three villages of Metekel zone, northwest Ethiopia. *Trop Anim Health Prod,* 44**,** 873-9.

THEVENAZ, P. & HECKER, H. 1980. Distribution and attachment of Trypanosoma (Nannomonas) congolense in the proximal part of the proboscis of Glossina morsitans morsitans. *Acta tropica,* 37**,** 163-175.

THOMSON, R., SAMANOVIC, M. & RAPER, J. 2009. Activity of trypanosome lytic factor: a novel component of innate immunity. *Future Microbiol,* 4**,** 789-96.

TOMASINI, N. 2018. Introgression of the kinetoplast DNA: an unusual evolutionary journey in Trypanosoma cruzi. *Current genomics,* 19**,** 133-139.

TSCHIRREN, B., ANDERSSON, M., SCHERMAN, K., WESTERDAHL, H., MITTL, P. R. & RÅBERG, L. 2013. Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population. *Proceedings of the Royal Society B: Biological Sciences,* 280**,** 20130364.

TUNTASUVAN, D., SARATAPHAN, N. & NISHIKAWA, H. 1997. Cerebral trypanosomiasis in native cattle. *Veterinary Parasitology,* 73**,** 357-363.

TURCATTI, G., ROMIEU, A., FEDURCO, M. & TAIRI, A.-P. 2008. A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic acids research,* 36**,** e25-e25.

UNDERHILL, D. M., OZINSKY, A., HAJJAR, A. M., STEVENS, A., WILSON, C. B., BASSETTI, M. & ADEREM, A. 1999. The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature,* 401**,** 811-815.

UNDERHILL, D. M., OZINSKY, A., SMITH, K. D. & ADEREM, A. 1999. Toll-like receptor-2 mediates mycobacteria-induced proinflammatory signaling in macrophages. *Proceedings of the National Academy of Sciences,* 96**,** 14459-14463.

VAN DEN BOSSCHE, P., DE LA ROCQUE, S., HENDRICKX, G. & BOUYER, J. 2010. A changing environment and the epidemiology of tsetse-transmitted livestock trypanosomiasis. *Trends in parasitology,* 26**,** 236-243.

VANHOLLEBEKE, B., TRUC, P., POELVOORDE, P., PAYS, A., JOSHI, P. P., KATTI, R., JANNIN, J. G. & PAYS, E. 2006. Human Trypanosoma evansi infection linked to a lack of apolipoprotein L-I. *N Engl J Med,* 355**,** 2752-6.

VAUGHAN, S. & GULL, K. 2003. The trypanosome flagellum. *Journal of Cell Science,* 116**,** 757-759.

VAUGHAN, S. & GULL, K. 2008. The structural mechanics of cell division in Trypanosoma brucei. Portland Press Limited.
VENUGOPAL, P. G., NUTMAN, T. B. & SEMNANI, R. T. 2009. Activation and regulation of toll-like receptors (TLRs) by helminth parasites. *Immunologic research,* 43**,** 252-263.
VERSTAK, B., ARNOT, C. J. & GAY, N. J. 2013. An alanine-to-proline mutation in the BB-loop of TLR3 Toll/IL-1R domain switches signalling adaptor specificity from TRIF to MyD88. *The Journal of Immunology,* 191**,** 6101-6109.
VERTHELYI, D., ISHII, K. J., GURSEL, M., TAKESHITA, F. & KLINMAN, D. M. 2001. Human peripheral blood cells differentially recognize and respond to two distinct CPG motifs. *The Journal of Immunology,* 166**,** 2372-2377.
VICKERMAN, K. 1965. Polymorphism and mitochondrial activity in sleeping sickness trypanosomes. *Nature,* 208**,** 762-766.
VON WISSMANN, B., MACHILA, N., PICOZZI, K., FÈVRE, E. M., BAREND, M., HANDEL, I. G. & WELBURN, S. C. 2011. Factors associated with acquisition of human infective and animal infective trypanosome infections in domestic livestock in western Kenya. *PLoS Neglected Tropical Diseases,* 5**,** e941.
VOTYPKA, J., SUKOVA, E., KRAEVA, N., ISHEMGULOVA, A., DUZI, I., LUKES, J. & YURCHENKO, V. 2013. Diversity of trypanosomatids (Kinetoplastea: Trypanosomatidae) parasitizing fleas (Insecta: Siphonaptera) and description of a new genus Blechomonas gen. n. *Protist,* 164**,** 763-81.
VU, A., CALZADILLA, A., GIDFAR, S., CALDERON-CANDELARIO, R. & MIRSAEIDI, M. 2016. Toll-like receptors in mycobacterial infection. *European Journal of Pharmacology*.
WALSH, P. S., METZGER, D. A. & HIGUCHI, R. 1991. Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques,* 10**,** 506-513.
WALSH, S. W., CHUMBLE, A. A., WASHINGTON, S. L., ARCHER, K. J., SAHINGUR, S. E. & STRAUSS, J. F., 3RD 2017. Increased expression of toll-like receptors 2 and 9 is associated with reduced DNA methylation in spontaneous preterm labor. *J Reprod Immunol,* 121**,** 35-41.
WANG, J., QI, J., ZHAO, H., HE, S., ZHANG, Y., WEI, S. & ZHAO, F. 2013. Metagenomic sequencing reveals microbiota and its functional potential associated with periodontal disease. *Scientific Reports,* 3**,** 1843.
WANG, X., XU, S., GAO, X., REN, H. & CHEN, J. 2007. Genetic polymorphism of TLR4 gene and correlation with mastitis in cattle. *Journal of Genetics and Genomics,* 34**,** 406-412.
WARD, A., HIDE, G. & JEHLE, R. 2019. Skin swabs with FTA® cards as a dry storage source for amphibian DNA. *Conservation Genetics Resources,* 11**,** 309-311.
WASTLING, S. L. & WELBURN, S. C. 2011. Diagnosis of human sleeping sickness: sense and sensitivity. *Trends Parasitol,* 27**,** 394-402.
WEIMER, E. T., MONTGOMERY, M., PETRAROIA, R., CRAWFORD, J. & SCHMITZ, J. L. 2016. Performance Characteristics and Validation of Next-Generation Sequencing for Human Leucocyte Antigen Typing. *J Mol Diagn,* 18**,** 668-675.
WEITZEL, T., ZULANTAY, I., DANQUAH, I., HAMANN, L., SCHUMANN, R. R., APT, W. & MOCKENHAUPT, F. P. 2012. Mannose-binding lectin and Toll-like receptor polymorphisms and Chagas disease in Chile. *The American journal of tropical medicine and hygiene,* 86**,** 229-232.
WEST, D. 2018. Use of an 18s rRNA metagenomics approach as a method of detection of multiple infections in field blood samples collected on FTA cards from cattle. Master Thesis (MSc by research), University of Salford.
WHEELER, R. J., GULL, K. & GLUENZ, E. 2012. Detailed interrogation of trypanosome cell biology via differential organelle staining and automated image analysis. *BMC Biol,* 10**,** 1.
WICKSTEAD, B., ERSFELD, K. & GULL, K. 2004. The small chromosomes of Trypanosoma brucei involved in antigenic variation are constructed around repetitive palindromes. *Genome Research,* 14**,** 1014-1024.
WRIGHT, A. F. 2005. Genetic variation: polymorphisms and mutations. *eLS*.
WORLD HEALTH ORGANIZATION, 2020. Human African trypanosomiasis*:* WHO HAT elimination partners: Implementing organizations [Online]. Available:

https://www.who.int/trypanosomiasis_african/partners/implementing_organizations/en/ [Accessed 2020].
WORLD HEALTH ORGANIZATION, W. H. 2020. Number of new reported cases (T.b. rhodesiense) Data by country [Online]. Available: https://apps.who.int/gho/data/view.main.95240 [Accessed 2020].
WU, C.-H., FALLINI, C., TICOZZI, N., KEAGLE, P. J., SAPP, P. C., PIOTROWSKA, K., LOWE, P., KOPPERS, M., MCKENNA-YASEK, D. & BARON, D. M. 2012. Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature,* 488**,** 499-503.
WU, F., ZHAO, S., YU, B., CHEN, Y.-M., WANG, W., HU, Y., SONG, Z.-G., TAO, Z.-W., TIAN, J.-H., PEI, Y.-Y., YUAN, M.-L., ZHANG, Y.-L., DAI, F.-H., LIU, Y., WANG, Q.-M., ZHENG, J.-J., XU, L., HOLMES, E. C. & ZHANG, Y.-Z. 2020. Complete genome characterisation of a novel coronavirus associated with severe human respiratory disease in Wuhan, China. *BioRxiv***,** 2020.01.24.919183.
XUAN, J., YU, Y., QING, T., GUO, L. & SHI, L. 2013. Next-generation sequencing in the clinic: promises and challenges. *Cancer Letters,* 340**,** 284-295.
YADON, Z. E. & SCHMUNIS, G. A. 2009. Congenital Chagas disease: estimating the potential risk in the United States. *The American Journal of Tropical Medicine and Hygiene,* 81**,** 927-933.
YARO, M., MUNYARD, K., STEAR, M. & GROTH, D. 2016. Combatting African animal trypanosomiasis (AAT) in livestock: the potential role of trypanotolerance. *Veterinary Parasitology,* 225**,** 43-52.
YOHE, S., HAUGE, A., BUNJER, K., KEMMER, T., BOWER, M., SCHOMAKER, M., ONSONGO, G., WILSON, J., ERDMANN, J., ZHOU, Y., DESHPANDE, A., SPEARS, M. D., BECKMAN, K., SILVERSTEIN, K. A. & THYAGARAJAN, B. 2015. Clinical validation of targeted next-generation sequencing for inherited disorders. *Arch Pathol Lab Med,* 139**,** 204-10.
ZABŁOTNA, M., SOBJANEK, M., PURZYCKA-BOHDAN, D., SZCZERKOWSKA-DOBOSZ, A., NEDOSZYTKO, B. & NOWICKI, R. J. 2017. The significance of Toll-like receptor (TLR) 2 and 9 gene polymorphisms in psoriasis. *Advances in Dermatology and Allergology/Postępy Dermatologii i Alergologii,* 34**,** 85.
ZEWDU, A., NEGASH, A., ASSEN, A. & YAREGAL, B. 2016. Camel Trypanosomosis: A Review on Diagnostic Approaches and Immunological Consequences.
ZHANG, D., ZHANG, G., HAYDEN, M. S., GREENBLATT, M. B., BUSSEY, C., FLAVELL, R. A. & GHOSH, S. 2004. A toll-like receptor that prevents infection by uropathogenic bacteria. *Science,* 303**,** 1522-1526.
ZHANG, Y. & JELTSCH, A. 2010. The application of next generation sequencing in DNA methylation analysis. *Genes (Basel),* 1**,** 85-101.
ZÍDKOVÁ, L., CEPICKA, I., SZABOVÁ, J. & SVOBODOVÁ, M. 2012. Biodiversity of avian trypanosomes. *Infection, Genetics and Evolution,* 12**,** 102-112.