SEROPREVALENCE AND GENETIC RESISTANCE OF DAIRY CATTLE TO BRUCELLOSIS IN SMALLHOLDER DAIRY FARMING SYSTEM IN SELECTED REGIONS OF TANZANIA

Isaac Joseph Mengele

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Life Sciences of the Nelson Mandela African Institution of Science and Technology

Arusha, Tanzania

January, 2025

ABSTRACT

Brucellosis in livestock is caused by different species such as *B. abortus* (cattle), *B. melitensis* (goats), *B. ovis* (sheep), *B. suis* (pigs) and *B. canis* (dogs). Recent studies in different regions have reported an increasing trend in smallholder dairy cattle. Furthermore, neighboring countries have reported identification of *B. melitensis* in dairy cattle, this is not known in Tanzania. In addition, the use of brucellosis resistant cattle has shown promising impact in controlling the disease in other countries, however it is not known if Tanzanian dairy cattle population contains genetic markers which are associated with the resistance/susceptibility to brucellosis. To address these questions, a cross-sectional study design was conducted in two agroecological zones, the Northern zone (Arusha, Kilimanjaro and Tanga regions) and the Southern highland zone (Iringa, Njombe and Mbeya regions) to determine brucellosis seroprevalence, risk factors, *Brucella* species circulating in smallholder dairy cattle farming systems and conduct Genome-wide association studies on cattle genome to identify SNP markers associated with brucellosis resistance/susceptibility. Seroprevalence was calculated at different administrative scales, and spatial tests were used to detect disease hotspots. A generalized mixed-effects regression model was built to explore the relationships among *Brucella* serostatus, animals and farm management factors. The overall seroprevalence was 2.39% (49/2048 cattle, 95% CI 1.7-3.1) across the study area and the Njombe Region represented the highest percentage with 15.5% (95% CI 11.0-22.0). In addition, hotspots were detected in the Njombe and Kilimanjaro Regions. Mixed-effects regression models showed that having goats around (OR 3.02, 95% CI 1.22-7.46) and abortion history (OR 4.91, 95% CI 1.43-16.9) were significant risk factors for brucellosis. The GEMMA analysis results identified a SNP marker BovineHD0900011750 located in chromosome number 9 responsible for brucellosis susceptibility (odds ratio >1) in 5.6% of the 1977 study animals. The web-based Ensembl variant effect predictor showed that the marker is found in a protein-coding region called ATPase Family Gene 1 homology (AFG1L) which produces mitochondrial integral membrane protein responsible for mitochondrial protein homeostasis and its impact in gene expression has been categorized as modifier as the direction of the impact is not yet known. Education of dairy farmers about brucellosis and its control, particularly in relation to goats in this setting is advised. In addition, a One-Health approach is needed to further study the role of small ruminants in dairy cattle brucellosis and the status of brucellosis in dairy farmers in the Njombe and Kilimanjaro Regions.

DECLARATION

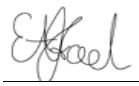
I, Isaac Joseph Mengele do hereby declare to the Senate of Nelson Mandela African Institution of Science and Technology that this thesis is my own original work and that it has neither been submitted nor being concurrently submitted for degree award in any other institution

Isaac Joseph Mengele Date

The declaration is hereby confirmed by the followings:

Prof. Gabriel Shirima Date

Prof. Mark Bronsvoort Date



Dr. Elizabeth Cook Date

COPYRIGHT

This thesis is copyright material protected under the Berne Convention, the Copyright Act of 1999, and other international and national enactments, in that behalf, on intellectual property. It must not be reproduced by any means, in full or in part, except for short extracts in fair dealing; for researcher private study, critical scholarly review or discourse with an acknowledgment, without a written permission of the Deputy Vice Chancellor for Academic, Research and Innovation, on behalf of both the author and the Nelson Mandela African Institution of Science and Technology.

CERTIFICATION

The undersigned certify that have read and hereby recommend for examination of a thesis titled: "*Seroprevalence and genetic resistance of dairy cattle to brucellosis in smallholder dairy farming system in selected regions of Tanzania"* in fulfillment of the requirements for the award of the Degree of Doctor of Philosophy in Life Sciences of The Nelson Mandela African Institution of Science and Technology (NM-AIST).

Prof. Gabriel Shirima Date

Prof. Mark Bronsvoort Date

Dr. Elizabeth Cook Date

ACKNOWLEDGEMENTS

I thank the Almighty God who had plans for me and who made everything possible to make sure that this academic journey materialized. I acknowledge the African Dairy Genetic Gain (ADGG)-Zoonoses through the International Livestock Research Institute (ILRI) for funding my PhD study and supporting various specialized and soft skills pieces of training to make sure that I became the best next-generation scientist. I would like to express my sincere gratitude to my supervisors Prof. Gabriel Shirima (NM-AIST-TZ), Prof. Mark Bronsvoort (UoE-UK) and Dr. Elizabeth Cook (ILRI-KE) for their trust, tireless supervision, and assistance toward this academic achievement.

Furthermore, I would like to extend my sincere gratitude to my employer Chief Executive Officer of Tanzania Veterinary Laboratory Agency for giving me a study leave to allow me to focus on my study. I also extend my sincere gratitude to the ADGG Tanzania team, Local Government Authorities (LGAs) in the study areas, my fieldwork team during data collection particularly Mr. Khamis Ali, ILRI and NM-AIST laboratory scientists for their wet-lab technical assistance, University of Edinburgh-The Roslin Institute-Economic Epidemiology and Risk Assessment (EERA) group, in particular Dr. Enrique Luis Hernandez-Castro for statistical and genetics analyses support and Center for Tropical Livestock Genetics and Health (CTLGH) for their technical support during my research stay in Edinburgh, UK.

My sincere gratitude goes to my Wife Mrs. Angela Mengele, my sons: Baraka Mengele and Benson Mengele, and my daughters: Glory Mengele and Gladys Mengele for their daily prayers, patience, and love throughout the time of my absence. My sincere gratitude to my mother Mrs. Nicoline Joseph and my late father Mr. Joseph Eliah Mengele who passed away when I was in Nairobi for research work, rest in peace father. Thank you, my parents, for giving me life, support and for being a cornerstone of my academic achievements.

DEDICATION

I dedicate this piece of my research work to my father, the late Mr. Joseph Eliah Mengele who passed away during my research time and I could not meet him in person during his last days when he was fighting for his life. Rest in eternal peace Father and Almighty GOD be with us all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ABBREVIATIONS AND SYMBOLS

%	Per cent
µl	Microliter
ADGG	African Dairy Genetic Gains
ASREML	Analystical Software for Residual Maximum Likelihooh
BADH	Betaine aldehyde dehydrogenase
BCV	Brucella-containing vacuole
BLAST	Basic local alignment such tool
cELISA	Competitive Enzyme Linked Immunosorbent Assay
CFSPH	The Center for Food Security and Public Health
CO2	Carbon dioxide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediamine tetraacetic acid
GDP	Gross Domestic Product
GEMMA	Genome Efficient Mixed Model Algorithm
GLM	Generalized linear model
GLMM	Generalized Linear mixed model (Mixed Effect Model)
GWAS	Genome-wide association study
HIC	High income countries
IgM	Immunoglobulin M
ILRI	International Livestock Research Institute
LD	Linkage disequilibrium
LMIC	Low and middle-income countries
LPS	Lipopolysaccharide
MAF	Minor allele frequency
MEGA	Molecular Evolutionary Genetics Analysis
MRT	Milk ring test
NM-AIST	Nelson Mandela African Institution of Science and Technology
oC	Degree centigrade
OD	Optical density
OMP	Outer membrane protein
OR	Odds ratio

O-SLPS	Ornithine smooth lipopolysaccharide
PAMPs	Pathogen associated molecular patterns
PCR	Polymerase chain reactions
qPCR	Quantitative real-time PCR
QTL	Quantitative trait locus
RER	Rough endoplasmic reticulum
RFLP	Restriction fragment length polymorphism
SAT	Slow Agglutination Tests
S-LPS	Smooth lipopolysaccharide
SNP	Single nucleotide polymorphism
SSA	Sub-Saharan Africa
TIR	Toll-interleukin receptor
TLRs	Toll like Receptors
TZS	Tanzanian shilling
UoE	University of Edinburgh
US$	United States dollar
USA	United States of America
WHO	World Health Organization
WOAH	World Organization for Animal Health

CHAPTER ONE

INTRODUCTION

1.1 Background of the problem

Brucellosis is a widespread bacterial zoonotic disease and highly prioritized for control in Africa (WHO, 2006). Brucellosis has been known to have an endemic status in most countries over the world and negatively impacts human and animal health (Corbel, 2006; Franc *et al.,* 2018; Moreno, 2014). Brucellosis is considered the most neglected non-malarial febrile illness zoonosis by the World Health Organization (WHO) because of its significant impact in developing countries. This calls for a One-Health approach to control and eliminate the disease (Pappas *et al.,* 2006; WHO, 2020).

In high income countries (HIC), the disease has been eradicated in livestock through stringent managerial and vaccination campaigns but continues to pose a significant economic threat in low and middle-income countries (LMIC) due to inadequate resources to control the disease and the husbandry complexities (Marcotty *et al.,* 2009; Tadesse, 2016). Studies in India show that, the disease in livestock is estimated to cause an economic loss of 3.4 billion US$ per annum, which significantly degrades the socioeconomic status of livestock-dependent communities and the general public (Singh *et al.,* 2015). Also, a study carried out in Zambia showed an overall economic loss of 134 131 US$ per year among livestock keepers (Mwinyi *et al.,* 2016).

The common etiological agent for brucellosis in dairy cattle is *B. abortus.* Nevertheless, *Brucella* species have the potential to spill over to other hosts especially where there are interactions between domestic species (Corbel, 2006; OIE, 2009). In East African countries and elsewhere, *B. melitensis* has been identified in cattle (Akoko *et al.,* 2021; Kaoud *et al.,* 2010; Makita *et al.,* 2011; Muendo *et al.,* 2012; Ntivuguruzwa *et al.,* 2022). This spillover transmission between domestic species may complicate any vaccination program. Also, there are some biological variations within *Brucella* species with, *B. melitensis* having three biotypes (1-3), *B. abortus* having eight biotypes (1-7,9) and *B. suis* having five biotypes (1-5) (Whatmore, 2009). In Tanzania, *B. abortus* biotype 3 has been isolated in Njombe from abortion cases in dairy cattle (Mathew *et al.,* 2015) and biotype 1 from local cattle in Katavi (Assenga *et al.,* 2015).

In animals, brucellosis is contracted by direct and indirect contact with the organism. Direct contact can be through infected vaginal discharge and birth products (placenta, fetus, fetal fluids) with the organism gaining access through broken skin or mucus membranes. Most animals; however, are thought to be infected indirectly by ingestion of contaminated feed and water (Richey, 1997; Yaeger & Holler, 2007), indirect infection by ingestion plays the major role of infection in domestic animals (Yaeger & Holler, 2007).

In cows the disease is manifested by abortion during the last trimester, lowered milk production, infertility, hygroma in knee carpal joints, and sometimes weak newborn calves, whereas in bulls, the disease causes orchitis, hygroma, epididymitis, and sterility (Currò *et al.*, 2012; Neta *et al.*, 2010; Radostitis *et al.*, 2000).

Some studies in SSA reported brucellosis prevalence in livestock at the variable range of 0-68.8% in cattle, 0.4-20% in camel, 0-88.8% in sheep and goats, and 0-12.9% in other species mostly pigs and dogs (McDermott & Arimi, 2002). In Tanzania, studies showed that all the production systems have been affected by brucellosis due to the detection of bacteria and risk factors for brucellosis transmission. Studies showed that cattle kept by pastoral communities had brucellosis seroprevalence as high as 30% (Kanuya *et al.*, 2006; Shirima, 2005) while agropastoral communities had variable seroprevalences to as high as 11.3% (Karimuribo *et al.*, 2007; Mengele *et al.*, 2018; Swai & Schoonman, 2010) and smallholder had less than 2% seroprevalence.

Following the rapid expansion of the smallholder dairy farming sector in Tanzania, the government instituted a brucellosis control program in dairy cattle in the late 1970s to late 1990s. Reports show that, during that time active surveillance, awareness creation, and calf vaccination using the *Brucella abortus* S19 vaccine were implemented which lead to reduction of brucellosis prevalence to as low as 2 % (Shirima, 2005; Shirima *et al.*, 2018). However, since then there has been no other similar control program instituted by the government. As a result, recent studies in the dairy subsector reported an increasing trend of brucellosis seroprevalence in some regions to as high as 21% (Karimuribo *et al.*, 2007; Mathew *et al.*, 2015; Mdegela *et al.*, 2004; Swai & Schoonman, 2010; Weinhäupl *et al.*, 2000). Although several studies were conducted in other farming systems, none were conducted extensively to elucidate the reasons for such trend of brucellosis incidences in smallholder dairy farming systems.

Since cross-transmission may complicate some of the conventional control interventions such as vaccination, knowledge on *Brucella* species circulating in smallholder dairy farming systems is required and alternative approaches for controlling brucellosis need to be developed. Compulsory vaccination, active surveillance, "test and slaughter" policies have managed to eradicate the disease in HIC, but those strategies have been difficult to sustainably implement in countries like Tanzania (Corbel, 2006; Refai, 2002) where farming systems are complex and resources are limited.

Successful breed selection showed that *Bos indicus* x *Bos taurus* crosses were significantly resistant to some diseases such as mastitis, tick-borne diseases, helminths, bovine tuberculosis, trypanosomiasis, and others (Banos *et al.,* 2017; Berry *et al.,* 2011; Roberts & Gray, 1973; Stone & Cundiff, 1985). A significant natural resistance against brucellosis has been demonstrated in Nellore Zebu in Brazil (Macedo *et al.,* 2013). These crossbreds of dairy cattle may have a potentially inherent resistance to *Brucella* infections. Breed selection has been considered effective, low risk, and with clear economic and social benefits to productivity and disease resistance (Martinez *et al.,* 2008). This is further supported by the view that the " ability to effectively and economically apply genetic selection to the problem of controlling infectious diseases in domestic animals will become increasingly feasible in the future" (Garry & Schutta, 2010).

Therefore, this research was carried out to establish the current status and risk factors associated with increased brucellosis seroprevalence in dairy cattle, identify circulating *Brucella* species, and establish susceptibility molecular markers for brucellosis under the smallholder dairy farming systems in six regions of Tanzania with high density of smallholder dairy cattle.

1.2 Statement of the problem

In Tanzania, brucellosis was first diagnosed in imported dairy cattle kept in parastatal farms in 1928 (Kitalyi, 1984). This was followed by the institution of control measures including active surveillance of the disease and calf vaccination between early 1980s to late 1990s. The disease prevalence was then reported to decline to below 2% (Shirima *et al.,* 2018) from 15.2% (Mahlau, 1967). Recent studies show the disease is on an increasing trend from a prevalence of <1 to 30% (Assenga *et al.,* 2015; Karimuribo *et al.,* 2007; Shirima *et al.,* 2018). Increasing the increasing prevalence calls for extensive epidemiological studies to understand risk factors for the increasing trend of brucellosis particularly in the smallholder dairy farming system.

The first isolates of *B. abortus* from cattle and *B. melitensis* from goats in Tanzania were in 1967 but these isolates were not characterized (Mahlau, 1967). In 2015, *B. abortus* was isolated and characterized and biovar 3 was identified from aborted materials of dairy cattle in Njombe region (Mathew *et al.*, 2015). Kenya, Uganda and Rwanda have identified *B. melitensis* in dairy cattle (Makita *et al.*, 2011; Muendo *et al.*, 2012; Ntivuguruzwa *et al.*, 2022). In Tanzania, it is not known yet if the increasing trend is also attributed to pathogen cross-transmission hence calling for an extensive and detailed study to investigate *Brucella* species affecting dairy cattle in Tanzania.

1.3 Rationale of the study

Control and prevention of brucellosis in dairy cattle have been a top priority in the last three decades, which included active surveillance (test and slaughter on parastatal farms), vaccination, and awareness creation but they have not been successful due to lack of sustainability, lack of resources, commitment, and poor policy enforcement (Shirima, 2005). Therefore, a new approach must be sought to control brucellosis in dairy cattle. The use of brucellosis-resistant breeds has shown promising results such as the use of the Nellore breed in Brazil (Macedo *et al.*, 2013). In Tanzania, there are no known dairy cattle resistant to brucellosis despite most of them being cross breeds of Bos indicus and Bos taurus, thus calling for the investigation to identify the dairy breeds with inherent resistance to brucellosis by identifying molecular markers associated with resistance/susceptibility to brucellosis.

1.4 Research objectives

1.4.1 General objective

To identify risk factors for increasing trend of brucellosis, *Brucella* species and assessment of genetically brucellosis resistant cross breeds of dairy cattle to improve control strategies and enhance productivity in Tanzania.

1.4.2 Specific objectives

(i) To establish the risk factors for brucellosis in dairy cattle in smallholder farming system in selected regions of Tanzania.

(ii) To identify *Brucella* species circulating in dairy cattle in smallholder dairy farming systems in selected regions of Tanzania.

(iii) To identify molecular markers associated with resistance/susceptibility to brucellosis found in dairy cattle cross breeds in smallholder dairy farming systems of Tanzania.

1.5 Research questions

(i) What are the factors driving brucellosis increasing trend in smallholder dairy cattle in Tanzania?

(ii) What are the Brucella species circulating in the dairy cattle population in Tanzania?

(iii) Which molecular marker(s) in dairy cross breeds is associated with resistance/susceptibility to brucellosis?

1.6 Significance of the study

The study is going to elucidate farming practices that are associated with the current increase in brucellosis prevalence in smallholder dairy cattle that will help to mitigate the disease control in the dairy subsector. In addition, identification of Brucella species circulating in dairy cattle is going to fill the knowledge gap as to what Brucella species are circulating in smallholder dairy cattle populations which could likely attribute to the increasing brucellosis prevalence but also could tell if we can depend on *B. abortus* S.19 vaccine to control brucellosis in smallholder dairy cattle. Furthermore, identification of genetic marker associated with brucellosis in smallholder dairy cattle will add value in disease control by doing breeding selection of animals with inherent resistance (with resistance markers) to complement other strategies for disease control.

1.7 Delineation of the study

This study is focusing on the study of brucellosis in smallholder dairy cattle in Tanzania found in two agroecological zone to determine brucellosis seroprevalence, risk factors, Brucella species circulating in smallholder dairy cattle farming systems and conduct Genome-wide association studies on cattle genome to identify SNP markers associated with brucellosis resistance/susceptibility. Seroprevalence was calculated at different administrative scales, and spatial tests were used to detect disease hotspots. A generalized mixed-effects regression model was built to explore the relationships among Brucella serostatus, animals and farm management factors.

CHAPTER TWO

LITERATURE REVIEW

2.1 Brucellosis causative agents, disease history, and risk factors

Brucellosis is one of the globally neglected bacterial zoonosis (Asmare *et al.,* 2013; Schelling *et al.,* 2003). The etiological agent is a gram-negative coccobacillus, a nonencapsulated, aerobic, and intracellular organism belonging to the genus *Brucella* (Bruce, 1887; Godfroid *et al.,* 2005). Currently, there are twelve known *Brucella* species, the three known *Brucella* species of great livestock importance include *B. abortus, B. melitensis,* and *B. suis* affecting cattle, small ruminants, and pigs respectively (Pappas, 2010). Other *Brucella* species with their preferential hosts include *B. canis* (dog) *B. ovis* (sheep) *B. ceti* (cetacea), *B. microti* (common voles), *B. inopinata* (human), *B. neotomae* (rodents), *B. pinnipedialis* (pinnipeds/seals), *B. papionis* (baboons) and B. *vulpis* (foxes) (Alton *et al.,* 1988; Bricker, 2002; O'Callaghan & Whatmore, 2011; Olsen & Palmer, 2014; Rajendhran, 2020). Among these, the ones with the zoonotic potential are *B. melitensis, B. abortus*, *B. suis,* and *B. canis*, with the first having the highest virulence and the last having the least (Corbel, 2006; Marzetti *et al.*, 2013).

Human beings get infected by direct contact which occurs on broken skin or mucus membranes with bacteria from infected fetal membranes, uterine discharges, fetal fluid, and aborted fetuses when helping animals during calving or abortion without wearing personal protective equipment (Ferrero *et al.,* 2014; Poester *et al.,* 2013; Tadesse, 2016). Livestock keepers, butcher men, those working in slaughterhouses, those working in the veterinary profession, and laboratory workers are occupationally at the highest risk of infection (Sagamiko *et al.,* 2018; Swai & Schoonman, 2012). In addition, ingestion of contaminated milk from infected animals expose humans to *Brucella* infection (Yaeger & Holler, 2007). There is a possibility of human-to-human transmission through breastfeeding or transplacentaly but transmission is rare by blood transfusion and sexual intercourse (Tuon *et al.*, 2017).

In humans the main symptoms include fever, fatigue, sweats, and general malaise, and other complications not limited to arthritis, endocarditis, and neurological signs (Dean *et al.*, 2012). These symptoms and signs are non-specific and can be confused with other febrile illnesses like typhoid fever, rheumatic fever, malaria, and arthroses (Makita *et al.,* 2011). About 0.5 million new human cases are reported globally each year (Pappas *et al.,* 2006). However, the

global true incidence is estimated to be between 0.5 million to 12.5 million cases (Godfroid *et al.,* 2013; Hull & Schumaker, 2018). This comes from the fact that there are a lot of cases missed by the available surveillance structures, resulting in a gross underestimation of the disease burden from individual countries globally (Dean *et al.,* 2012). In Tanzania, a recent study has synthesized several factors associated with the underestimation of true incidences in humans and livestock, such factors are inadequate knowledge of brucellosis among stakeholders in the livestock value chain, lack of epidemiological background among human and livestock health workers, and challenges associated with diagnostic tests (Mengele *et al.,* 2023a).

In Africa, the first case of brucellosis in cattle was reported in Zimbabwe (1906) followed by Kenya (1914) and South Africa (1915) (Chukwu, 1985), since then brucellosis has been wide-spreading and has assumed the endemic status (Ducrotoy *et al.,* 2014; Mangen *et al.,* 2002; McDermott *et al.,* 2013) with countries in sub-Saharan Africa (SSA) estimated to bear the highest burden (Delia *et al.,* 2012). Brucellosis in different wildlife populations is considered endemic and spill overs of infection between the domestic and wildlife occurs in the interface areas (Godfroid, 2002), especially when animals share water sources and grazing land (Pandey *et al.,* 1999). In SSA, the prevalence of brucellosis in livestock is unclear and poorly understood (Tumwine *et al.,* 2015) due to the lack of intensive and extensive surveillance to create valid epidemiological data. Therefore countries and livestock keepers in SSA suffer severe socioeconomic impact (Delia *et al.,* 2012; Marcotty *et al.,* 2009; McDermott *et al.,* 2013), due to limited access to international trade of their live animals and animal by-products (Akakpo *et al.,* 2010).

In Tanzania, the first brucellosis outbreak occurred in exotic imported dairy cattle in 1927 during the colonial era (Mahlau, 1967) and the first diagnosis was done in 1928 by using a serum agglutination test (Kitalyi, 1984) and the disease prevalence was found to be 15.2% (Mahlau, 1967). Up to the early 1990s the government of Tanzania imported dairy cattle as an effort to improve dairy production in the parastatal farms which were found in different zones of the country. The prevalence of the disease was reported to be less than 2% following previous government efforts to control the disease by active surveillance and vaccination in late 1970s to early1990s (Shirima *et al.,* 2018). A study conducted later in the late 1990s as a follow-up study in the lake zone of Tanzania found that, brucellosis in dairy cattle in parastatal farms and nearby private farms in the lake zone was 6.3% (Jiwa *et al.,* 1996). Since then, studies

have been carried out in different agroecological zones and production systems to establish different levels of disease prevalence and risk factors associated with the transmission of the disease. The major risk factors for the disease in the dairy sector were herd size, age, purchase of unscreened animals, farmers' knowledge of the disease, direct contact between dairy and traditional cattle, and others (Assenga *et al.,* 2015; Karimuribo *et al.,* 2007; Mengele *et al.,* 2018; Shirima *et al.,* 2018; Weinhäupl *et al.,* 2000). Recent studies show that brucellosis is re-emerging in dairy cattle (Karimuribo *et al.,* 2007; Mdegela *et al.,* 2004; Sagamiko *et al.,* 2018; Shirima *et al.,* 2018) after collapsing of the stringent control measures enforced in the last four decades in Tanzania due to resources constraints.

2.2 Dairy industry production, success, and challenges in Tanzania

Tanzania is estimated to have a population of 33 928 391 million cattle, out of which 2.6% are improved dairy breeds, crossbreds of Friesian, Jersey, and Ayrshire with Tanzania short horn zebus (Development, 2016; Statistics, 2021). This statistic makes Tanzania the second largest in livestock resource in Africa after Ethiopia, which has 60.39 million cattle (Statistics, 2021). The current statistics show that Tanzania is having a total of 865 628 improved dairy cattle. Furthermore, according to livestock master plan initiatives, the growth of the dairy industry's contribution to GDP is expected to increase by 75% by 2021/2022 (Michael *et al.,* 2018).

The livestock sub-sector plays a role in contributing to National strategic goals and National development Vision 2025. According to the National economic survey of 2019/2020, the livestock subsector grew at the rate of 5% and contributed 7.1% of the National gross domestic product (GDP) (Planning, 2021). The contribution of the dairy industry to the GDP is one-third (1/3) of the livestock sector's contribution (Njombe *et al.,* 2011). According to the national agricultural survey, for the year 2019/2020 Tanzania produced 3.1 billion liters of milk and 30% of annual production comes from improved dairy cow breeds (Lunogelo *et al.,* 2020; Statistics, 2021). The average milk production for indigenous and improved cows in Tanzania is 3 and 8 liters per day respectively. This suggests that there is more work to do to improve the production potential of dairy cattle which can as well improve the contribution to national GDP.

Livestock keeping is one of the major agricultural means of income in Tanzania. According to an economic survey of 2012, there are 4.9 million agricultural households in the country of which, 36% are pastoralists and 35% are agropastoral communities and dairy productivity

increased from 1.74 billion liters in 2011 to 1.85 billion liters in 2012 (Finance, 2013). Out of 1.85 billion liters, 0.597 billion liters equivalent to 32% come from crossbreed dairy cows and 1.256 billion liters come from traditional cows kept by pastoral and agropastoral farmers (Finance, 2013). According to the agriculture sample census of 2019-2020, milk production has reached as much as 3.1 billion liters of cow milk (Statistics, 2021). The milk production system in Tanzania is carried out in two systems, the traditional and the dairy production systems.

In the traditional production system, milk is produced by indigenous cows in rural areas, the system is characterized by low input and low output and limited pastures during the dry season therefore animals have to move longer distances in search of pastures. In addition, collection and marketing of milk is a major challenge due to farm remoteness and poor infrastructure (Njombe *et al.,* 2011). This is the main milk production system comprising 90% of the dairy cattle population and produces 70% of milk produced in Tanzania but only 10% reach the urban market and processing plants (Njombe & Msanga, 2009). Milk processing in Tanzania is undertaken by 62 small to medium plants with a capacity between 500 to 30 000 liters per day and most of them are privately owned including previously state-owned plants after privatization producing various products including pasteurized milk, fermented milk, cheese, yogurt, ghee, and butter (Njombe & Msanga, 2009; Njombe *et al.*, 2011).

Dairy production in Tanzania started in the colonial era whereby European dairy breeds (*Bos taurus*) were imported for milk production. During colonial times, there were only a few large dairy farms owned by the colonial government and few a settlers. However, after independence, most large dairy farms were taken over by the government (Kurwijila & Boki, 2003) and deliberate efforts were made to improve the dairy industry. Between 1975 and 1993, 1039 heifers and bulls were imported from the USA and New Zealand for breeding more heifers and the surplus was distributed to smallholder dairy farmers (Kurwijila & Boki, 2003). Currently, there are significant indicators of real growth in the dairy sector, including the availability of improved indigenous crossbreeds (crosses of Friesian, Jersey, and Ayrshire), increased herd numbers, increased milk output, improved herd health, diversified processed dairy products and high value of dairy products (Lokuruka, 2016).

The dairy industry is divided into three production systems, smallholder, medium, and large dairy production systems. The smallholder production system is characterized by 1-5 dairy cows per household, practicing zero grazing in the urban and peri-urban areas and feeding

grasses and some annual crop residues. The system enjoys the accessibility of the market, veterinary services, supplemental feed, and extension services, however, feed availability during the dry season is a major challenge. This system is considered a transition toward medium-scale dairy production (Njombe *et al.,* 2011; Nyange & Mdoe, 1995). The medium-scale dairy farming system is a transition toward large-scale farming, consisting of about 10-50 cows, grazing in the paddock and farms are found in peri-urban cities (Njombe *et al.,* 2011). Large-scale dairy farms are those with 50 or more dairy cows, most of these were owned by the government in parastatal farms, however currently after privatization some of them are privately owned and other newly established farms are privately owned.

Several challenges are limiting the development of the dairy industry in Tanzania, some are associated with milk production, distribution, and marketing, however, livestock diseases like transboundary, vector-borne, zoonosis, and emerging diseases further cripple the dairy industry (Lunogelo *et al.,* 2020; Njombe & Msanga, 2009; Statistics, 2021).

The challenge of milk production is associated with the availability of quality pasture and supplements to meet cow nutritional needs for milk production and access to high-yield genetics heifers (Lunogelo *et al.,* 2020). According to the livestock modernization initiative, in Tanzania, there is a high demand for heifers compared to the supply and limited technical advice due to weak extension services (Development, 2016). The dominance of geographically scattered and poorly organized smallholder dairy farmers has contributed to the high costs of milk collection by milk processors and therefore lead to low prices of milk offered by processors (Lunogelo *et al.,* 2020). For smallholder dairy farmers, to overcome the challenges of milk production, collection, distribution, and access to better markets, some have organized cooperatives to buy and process their milk and search for the markets, such examples of a cooperative are Mwakaleli dairy cooperative in Rungwe district and Nronga women dairy cooperative in Hai district (Kurwijila *et al.,* 2012).

2.3 Brucella structure, antigenicity, and susceptibility to disinfectants

Bacteria of the Genus *Brucella* are Gram-negative coccobacillus, aerobic, nonmotile, non-spore-forming, facultative intracellular bacterium of the size: 0.6-1.5 µm x 0.5-0.7 µm (Mantur *et al.,* 2007). The genome of *Brucella* has two chromosomes with the size of 2.1 and 1.2 Mbp and the G+C contents for chromosome I and chromosome II are 57.2% and 57.3% respectively (Halling *et al.,* 2005). The genome is highly conserved across all species at the nucleotide level,

this is the basis for differentiation of *Brucella* spp. by molecular methods (Michaux-Charachon *et al.,* 1997; Rajendhran, 2020). *Brucellae* are partially acid-fast with oxidase, nitrate reductase, catalase and urease activities (Khurana *et al.,* 2021).

Several antigenic structures have been characterized in *Brucella.* The lipopolysaccharide (LPS) and ornithine lipids are found on the outer membrane of the bacteria, but LPS is the main antigen that dominates the induction of antibody response but cannot induce innate immunity (Moreno *et al.*, 1981; Palacios-Chaves *et al.,* 2011).

Some bacteria have smooth-LPS (S-LPS) which contains an immunodominant O chain-polysaccharide (O-PS) chemically known as a homopolymer of 4,6-dideoxy-4-formamide-alpha-Dmannose linked via glycosidic linkages and O-polysaccharide bridged to lipid by an intermediate core oligosaccharide to form O chain smooth lipopolysaccharides (O-SLPS) (Bundle *et al.,* 1987; Gil-Ramírez *et al.,* 2014). *Brucella* species affecting livestock contain the O-SLPS (i.e. *B. abortus, B. melitensis,* and *B. suis*) hence they bear common epitopes (A and M) except for *B. ovis* and *B. canis* which have rough-LPS (R-LPS) on their outer membrane (Riezu-Boj *et al.,* 1990).

The LPS of rough strains (R-LPS) is like the LPS of smooth strains (S-LPS) except that the O-chain is either absent or reduced to a few residues. The role of O chain-polysaccharide (O-PS) confers bacterial resistance against the host's innate immunity during infection, when O-chain disruption exposes LPS to innate immunity particularly bactericidal peptidase and complement factors, this increases proinflammatory responses and subsequent weakening of the bacteria (Barquero-Calvo *et al.,* 2007; Conde-Alvarez *et al.,* 2012; De Tejada *et al.,* 1995; Eisenschenk *et al.,* 1999).

Brucella spp. are easily killed by most commonly available disinfectants such as hypochlorite solutions, 70% ethanol, isopropanol, iodophors, phenolic disinfectants, formaldehyde, glutaraldehyde, and xylene (CFSPH, 2018; Khurana *et al.,* 2021). Furthermore, most *Brucellae* are quickly inactivated at an acid pH <3.5 except for *B. microti* which seems to be resistant to a lower pH, they can as well be destroyed at moist heat of 121℃ for at least 15 mins, dry heat of 160-170℃ for at least 1 hour, gamma irradiation and pasteurization and boiling of liquids for 10 mins (CFSPH, 2018).

2.4 Pathogenesis and host immunity evasion

Brucella pathogens get into the body through ingestion of contaminated milk or contact with broken skin or mucus membranes (Brinley & Corbel, 1990; Rossetti *et al.,* 2013). Vertical transmission to the young ones from an infected dam and inhalation are all possible modes of infection (Ko & Splitter, 2003; Radostits *et al.*, 2007). Once *Brucellae* are in the body, the bacteria has predilection sites which include the gravid uterus, mammary glands, testes, lymph nodes, and joints (Greenfield *et al.,* 2002). In the gravid uterus, the bacteria use sugar-alcohol called erythritol produced by placental tissues for its growth and later attract other bacteria to localize in the uterus and cause subsequent abortion (Sperry & Robertson, 1975).

The use of erythritol by *Brucella* activates the bacteria virulence pathways which include type IV secretion system *Vir*B and flagella protein, this suggests the importance of erythritol in the pathogenicity of *Brucella* (Petersen *et al.,* 2013) The type IV secretion system *Vir*B is involved in the inhibition of the host immune response and intracellular survival during infection (Ke *et al.,* 2015).

For intracellular survival, brucellae use *Vir*B type 4 secretory system (T4SS), two-component sensory, cyclic β-glucan and regulatory system BvrS/BvrR, *Brucella* LPS (BrLPS) and pathogen-associated molecular patterns (PAMPs) (Gopalakrishnan *et al.,* 2016) and for infection use BacA, BmaC, outer membrane proteins (Omps), SagA, MucR, BtaE, and BetB (Comerci *et al.,* 2001; Del Giudice *et al.,* 2013; Lee *et al.,* 2014; Lim *et al.,* 2012; Martín-Martín *et al.,* 2012; Mirabella *et al.,* 2013; Posadas *et al.,* 2012; Ruiz-Ranwez *et al.,* 2013; Vizcaíno & Cloeckaert, 2012). When the *Brucella* is phagocytosed by neutrophils or macrophages, the T4SS is activated and there will be secretion and translocation of 15 effector proteins in the host cytosol which modulate the host cellular response and intracellular lifestyle of the bacteria (Boschiroli *et al.,* 2002; Döhmer *et al.,* 2014; Foulongne *et al.,* 2000). This causes bacteria to subvert lysosome fusion and therefore to produce *Brucella*-containing vacuole (BCV) which allows replication and binding with the endoplasmic reticulum (Celli *et al.,* 2003; Marchesini *et al.,* 2011).

Two-component sensory and regulatory system BvrS/BvrR: Co-ordinate the outer membrane proteins (OMP) architectural transformation which is necessary for pathogen metabolism during invasion and intracellular survival (Guzman-Verri *et al.,* 2002; López-Goñi *et al.,* 2002). *Brucella* strain with less or no Opolysaccharide are less pathogenic than smooth strain (Ko &

Splitter, 2003). The Brucella LPS (*Br*LPS) is a virulence factor, it is less immunogenic when compared to LPS of other bacteria (Zähringer *et al.,* 2004) induces less biological activity, a reason for maintaining the durability of the pathogens within phagocytic cells (Gopalakrishnan *et al.,* 2016).

Pathogen-associated molecular patterns (PAMPs) are virulence factors that weakly stimulate Toll-Like Receptors (TLRs) (Lapaque *et al.,* 2009). Inexpression of *Brucella* PAMPs in *Brucella* outer membrane hide the bacteria to early detection by innate immunity (Lapaque *et al.,* 2009), this causes weak innate immunity activation (Martirosyan *et al.,* 2011); therefore, there will be a failure in the development of antigen-specific acquired immunity and hence hides from the host immunity.

The *B. abortus* cyclic β (1–2) glucan gene also known as *cgs* encodes for cyclic β-1,2-glucan-synthetase enzyme which is necessary for preventing maturation and final replication of the *Brucella*-containing vacuole (BCV) preventing lysosome vacuoles fusion which is necessary for the maturation of BCV (Briones *et al.,* 2001; Dermine *et al.,* 2001).

These five virulence factors all together, contribute to multiplication and virulence mechanisms for intracellular survival of *Brucella*. The transporter-like protein BacA of *Brucella* is necessary for the intracellular survival of the bacteria (Roop *et al.,* 2009). The BetB molecule which is a short form of a Betaine aldehyde dehydrogenase (BADH) catalyses the oxidation of betaine aldehyde to glycine betaine which act as an osmotic regulator (Kempf & Bremer, 1998). *Brucella* BmaC is an adhesin responsible for attachment to the host cells or with soluble macromolecules (Posadas *et al.,* 2012). Bacterial attachment to the host target cells is necessary for some pathogens to inject effectors (Kline *et al.,* 2009). A member of zinc-finger proteins which is abbreviated as MucR found in the α-proteobacteria. Their main function is to regulate expression of all genes of the bacteria required for the successful pathogenic and symbiotic interactions with the eukaryotic hosts (Pirone *et al.,* 2018). Furthermore, MucR zinc finger proteins have dual roles in exopolysaccharides biosynthesis pathways by regulating succinoglycan biosynthesis genes and repressing the synthesis of galactoglucan, these pathways are essential for quorum sensing (Bertram-Drogatz *et al.,* 1998; Keller *et al.,* 1995; McIntosh *et al.,* 2008). A unipolar trimeric autotransporter adhesion abbreviated as BtaE, is required for full virulence/infectivity of the bacteria (Ruiz-Ranwez *et al.,* 2013).

2.5 Diagnosis of Brucellosis

Different diagnostic tests can be used for different purposes, such purposes are certification of animals free from disease, screening or prevalence studies, confirmatory diagnosis, and in countries where brucellosis is eradicated, targeted surveillance to avoid the reintroduction of brucellosis through the importation of infected animals or animal products (Godfroid *et al.,* 2010). There is no pathognomonic sign for clinical diagnosis of brucellosis as the signs are common to many other disease conditions. Therefore, laboratory testing is a key for correct detection and confirmation of the disease. Isolation and identification of the causative agent is required for definitive diagnosis (Mantur *et al.,* 2007). However, the definitive isolation and identification is time-consuming, must be performed by highly skilled personnel, and poses risk of infection to laboratory personnel. For these reasons, serological tests are normally preferred in most settings (Padilla *et al.,* 2010). The DNA technologies have been developed and used to genetically characterize *Brucella* spp. Different PCR-based assays have been used for rapid recognition of *Brucella* genus and for differential identification of species and strains (Padilla *et al.,* 2010). Direct tests diagnostic methods involve microbiological analysis and *Brucella* DNA detection by Polymerase Chain Reaction (PCR)-based methods and indirect tests which are applied either in-vitro (milk or blood or serum) or in-vivo (allergic test) (Godfroid *et al.,* 2010).

2.5.1 Direct tests

(i) Microbiological analysis

Isolation of the bacteria is required for biotyping and DNA extraction. Predominant clinical signs observed determine the choice of samples for isolation. For clinical brucellosis, preferred samples include fetal membranes, vaginal discharges, colostrum, milk, sperm, fluid from arthritis or hygroma and aborted fetuses (stomach, spleen, and lung) (Godfroid *et al.,* 2010). Different commercial culture media are available for growing *Brucella*. The commonly used basal media include Triptcase soy (BBL®), Bacto Tryptose (Difco®), Tryptic soy (Gibco®), Tryptone soya (Oxoid®) (Padilla *et al.,* 2010) and some selective media for *Brucella* spp are such as *Brucella* agar, Farrell media, and CITA (Vicente *et al.,* 2014). Usually, field samples are contaminated therefore selective medium contains different antibiotics which inhibit the growth of other bacteria (Marin *et al.,* 1996). For example, Farrell´s medium contains the following antibiotics and quantities per liter of basal medium: polymyxin B sulfate (5 mg),

bacitracin (25 mg), nystatin (100 000 units), nalidixic acid (5 mg), natamycin (50 mg), vancomycin (20 mg) (Marin *et al.*, 1996).

Brucella spp. such as *B. abortus* wild-type (biovars 1-4) need CO_2 for growth, while others like *B. abortus* wild-type (biovars 5, 6, 9), *B. abortus* S19 vaccine strain, *B. melitensis*, and *B. suis*, do not. For bacteria growth conditions in a primary culture, the recommended incubation conditions are in a 10% CO_2 atmosphere at 37°C. For culturing liquid samples such as milk or blood, the use of broth or a biphasic medium like the Castaneda medium is recommended as they increase the sensitivity because *Brucella* is present in a small number. Colonies takes time to grow but may appear after 2-3 days, but sample is considered negative after 2-3 weeks of incubation period (Alton *et al.*, 1988).

The bacteria can be seen and identified through stained colony smears prepared on a microscopic glass slide. The commonly used staining methods are Gram staining, the modified Ziehl-Nielsen, and the modified Köster (Alton *et al.*, 1988). *Brucellae* are Gram-negative coccobacilli (short rods), arranged singly but sometimes in pairs or small groups (Alton *et al.*, 1988; Godfroid *et al.*, 2010). Other features and biological properties such as colonial morphology, staining, slide agglutination with anti-*Brucella* serum (smooth or rough), urease (+), catalase (+), and oxidase (+) tests are the basis for a culture to be identified as belonging to the genus *Brucella* (Godfroid *et al.*, 2010).

Different tests are used for biotyping of *Brucella* spp., the most important being agglutination tests with antibodies against rough or smooth LPS, i.e., against the A or M epitopes of 'O' chain polysaccharides (O-LPS); dependence on CO_2 for growth, measured usually in primary cultures; lysis by phages, growth in the presence of basal fuchsine or thionine; production of H_2S; and dye sensitivity tests such as the crystal violet or acriflavine (Alton *et al.*, 1988). These tests are sensitive, time-consuming, and hazardous; therefore, it is recommended to be done in reference laboratories using standardized procedures by qualified personnel.

(ii) Molecular methods for Brucella DNA detection and typing

Polymerase Chain Reaction (PCR) has been used to identify *Brucella* DNA at genus, species and biovar levels has improved diagnostic tests and different methods have been developed (Padilla *et al.*, 2010). The PCR as well has solved the challenges associated with the isolation of bacteria, the longer time it takes to grow them, and the potential for exposure to laboratory

personnel (Ouahrani-Bettache *et al.,* 1996). The first brucellosis PCR-based test was introduced in 1990 and was targeted to a gene encoding a 43-KDa outer membrane protein from *B. abortus* strain 19 (Fekete *et al.,* 1990). Genus-specific PCR assays targeted at highly conserved *Brucella Bcsp*31gene encoding for 31kDa protein and 16S-23S rRNA operon were first designed (Bricker, 2002), later other target genes such as IS711 and *per* were used to identify *Brucella* at the genus level (Akoko *et al.,* 2021; Halling *et al.,* 1993; Lübeck *et al.,* 2003). The sensitivity of PCR increases with an increase in the number of targeted genes like IS711 due to being in multiple copies (Ouahrani-Bettache *et al.,* 1996). At first, PCR was designed to detect DNA from bacteria isolates, but now has been modified to detect DNA from clinical samples as well, however, validations have been in process to eliminate challenges associated with DNA extraction from clinical samples which affect the sensitivity of the PCR assay (Dauphin *et al.,* 2009).

For species-specific assays, the multiplex AMOS PCR, has been named after the *Brucella* species it dentify which are "abortus, melitensis, ovis, suis". The method is based on the existing polymorphism which is species-specific localization of the insertion sequence IS711 in the *Brucella* chromosome (Bricker & Halling, 1994). Another multiplex PCR based method called "Bruce ladder" (Bruce ladder PC) which is the first method designed to identify and differentiate all of the known *Brucella species,* wild strains, and the vaccine strains in the same assay (García-Yoldi *et al.,* 2006; Godfroid *et al.,* 2010; Lopez-Goni *et al.,* 2011). It was difficult to differentiate *Brucella* into species by using older molecular techniques. However, due to the advancement of techniques like multiple loci variable number tandem repeats analysis (MLVA-VNT) it has been possible to differentiate *Brucella* species (Bricker *et al.,* 2003; Mathew *et al.,* 2015). The *AlkB/IS711* and *BMEI1162/IS711* targets have been used to identify *B. abortus* and *B. melitensis* (Akoko *et al.,* 2021; Al Dahouk *et al.,* 2007; Sambu *et al.,* 2021).

Currently, different methods have been developed for genotyping *Brucella* strain, the published methods to serve the purpose include: random-amplified polymorphic DNA-PCR (RAPD-PCR), repetitive intergenic palindromic sequence-PCR (REP-PCR), and enterobacterial repetitive intergenic consensus sequence-PCR (ERIC-PCR) (Mercier *et al.,* 1996; Tcherneva *et al.,* 2000; Tcherneva *et al.,* 1996). Others include arbitrary primed-PCR (AP-PCR), and restriction fragment length polymorphism-PCR (RFLP-PCR) of the omp2 locus, and the latest Hypervariable Octameric Oligonucleotide Finger-Print (HOOF-Print) (Bricker *et al.,* 2003; Cloeckaert *et al.,* 1995; Cloeckaert *et al.,* 2001; Kılıç *et al.,* 2021). However, some of the

methods (ERIC-PCR and RAPD-PCR) have been difficult to reproduce the results in other laboratories due to assay conditions and environmental effects and therefore made them non-universal (Mercier *et al.,* 1996; Tcherneva *et al.,* 1996). The RFLP-PCR of omp2 has been reproducible and useful for the identification of *Brucella* species, however as an epidemiological tool, the technique has a challenge of a low rate of natural sequence variability of the locus (Bricker *et al.,* 2003; Cloeckaert *et al.,* 1995; Foster *et al.,* 2009). The HOOF-Print technique is the most discriminatory test for the characterized *Brucella* species and non-characterized field isolates, the test is rapid and reproducible (Bricker *et al.,* 2003).

2.5.2 Indirect Tests

Since the causative agents of brucellosis have been identified, several diagnostic tests have been developed and used. Isolation and identification of the bacterium remains the gold standard test, however, serologic tests have become useful and are presumptive evidence of infection (Padilla *et al.,* 2010). Brucellosis was first diagnosed serologically by using a simple tube agglutination test (Wright & Smith, 1897) which was followed by various modifications to the tube agglutination test and development of new serological tests to increase testing accuracy (Nielsen & Yu, 2010).

For serological tests, a smooth lipopolyscaccharide (S-LPS) *Brucella* antigen has been useful to detect brucellosis. The S-LPS contains an immunodominant O chain-polysaccharide (O-SLPS) (Bundle *et al.,* 1987). However, O-SLPS epitope is found in other Gram-negative enterobacteria like *Yersinia enterocolitica*, O:9, *Francisella tularensis, Eschelichia* coli O:157, *Salmonella urbana* O:3, *Vibrio cholera* and others (Corbel, 2006). These bacteria produce antibodies like those of *Brucella*. Therefore, the antigen used in the serological tests to detect antibodies against *Brucella* react equally with antibodies from other bacteria to produce false positive results, this is similar to vaccine antibodies like that of *B. abortus* S19; therefore, their use requires cautions (Godfroid *et al.,* 2010; Schurig *et al.,* 1991).

Following infection with *Brucella*, the first type of antibodies produced in two weeks is Immunoglobulin M (IgM) later on IgG1, IgG2, and IgA (Allan *et al.,* 1976; Beh, 1974). Therefore, detection of IgM as an indication of exposure to *Brucella* would be ideal, however, cross-reactivity with other non-*Brucella* bacteria gives a higher chance of producing false positive results and therefore reduce the specificity of the test. Detection of IgG2 and IgA

reduces the assay sensitivity, hence for brucellosis diagnosis, detection of IgG1 by serological tests gives accurate results (Allan *et al.,* 1976; Butler *et al.,* 1986; Lamb *et al.,* 1979).

Serological tests can be categorized into two main groups, conventional tests, and primary binding assays. Conventional tests rely on the antibody performing a secondary function, for instance, fixation of complement, and in primary binding assays, the only function of the antibody is attachment to its corresponding antigen (Nielsen & Yu, 2010). The conventional tests include:

(i) Agglutination tests

Agglutination tests include; Slow Agglutination Tests (SAT) and Rapid Agglutination Tests. The SAT may take up to 24 hours, whereas the rapid agglutination test gives results in a min. Examples of SAT are a Milk Ring Test (MRT) and the commonly used rapid agglutination tests are Rose Bengal, Modified Rose Bengal, Buffered Antigen Plate Agglutination, Card test, and Rivanol.

Rose Bengal Plate test or its variants are commonly and widely used in both humans and livestock since its cheap, simple, and rapid with high sensitivity thus remain the first-line diagnostic test (De Glanville *et al.,* 2017). The test is considered a screening test, with high sensitivity of up to 99% (OIE, 2018), however, it is challenged by its low specificity which increases the number of false positive results and lowers positive predictive values (Sammartino *et al.,* 2006). Rose Bengal Plate test can give positive results to cattle vaccinated with the *Brucella abortus* S.19 vaccine which has been used in Tanzania and can as well cross-react with antibodies produced from enterobacterial infections and give false positive results.

(ii) Complement fixation tests

Complement fixation tests; is complex to perform however, the complement fixation test allows to detect anti-*Brucella* antibodies that can activate the complement which are IgG and IgM (Godfroid *et al.,* 2010). The test requires sheep red blood cells (SRBCs) as an indicator for hemolysis presence or not, rabbit Anti-SRBCs serum (hemolysin), and guinea pig complement. The *Brucella* antigen must be prepared from smooth strains of *B. abortus,* and standardized according to OIE standards before use. The test procedure and results interpretation are well explained in OIE manual of 2008 where, hemolysis suggests no *Brucella* antibodies in a test sample (OIE, 2008).

(iii) Precipitation tests

Precipitation tests including; Agar gel immunodiffusion (AGID) and Single Radial Immunodiffusion (SRD) were the first tests to be able to distinguish vaccinal antibodies and those from natural infection of *B. abortus* (Diaz *et al.*, 1979). The AGID is a diagnostic test that uses serum from a suspect animal for the detection of antibody produced against the infection (antigen). The test serum is placed in a well in the agar and multiple antigenic peptides (MAP) antigen (known antigen) is placed in a nearby well. These two test components passively diffuse out of the well into the agar. If the test serum sample contains antibodies to antigens of MAP they bind, forming an interlaced antigen-antibody complex precipitates in the agar. The precipitate can be seen by using naked eyes as a thin white line between the wells. This same technology is used for the diagnosis of other diseases. It is important and a good practice that the test results being interpreted by using a known positive control serum sample (available in commercial kits) in the assay for comparison and validity of the test results (Michael & Elizabeth, 2018; Nielsen, 2002).

In SRD, a polysaccharide B antigen derived from *B. melitensis* containing glucose molecules and OPS the active component is normally used and mixed with agar matrix added to a well in agar matrix adjacent to another containing test serum. If the antibody is present, a ring of precipitate is formed in a few hours or after long incubation if the sera contain few antibodies (Nielsen, 2002). The precipitation tests for brucellosis have a challenge of reproducibility making their data not available to others, however, their sensitivity is not sufficient for wide usage and is neither recommended by OIE for bovine brucellosis (Jones *et al.*, 1980).

Primary Binding Assays include Fluorescence immunoassay, Radioimmunoassay, Particle counting fluorescence immunoassay, competitive enzyme immunoassay, indirect enzyme immunoassay, and fluorescence polarization assay (Nielsen & Yu, 2010; Padilla *et al.*, 2010). No serological test is 100% accurate, therefore diagnosis is based on the results of two or more tests, a screening test with high sensitivity and less specificity and a confirmatory test with high specificity and relatively high sensitivity (Nielsen & Yu, 2010; Padilla *et al.*, 2010).

In Tanzania, competitive Enzyme Linked Immunosorbent Assay (cELISA) is the most frequently used to confirm brucellosis in animals and human being (Mengele *et al.*, 2023a).This test detect IgG1 which signifies the exposure status of the animal. The sensitivity and specificity evaluation of cELISA in Tanzania livestock population has shown to be 98% and

99%, respectively (Bodenham *et al.*, 2021). In this study cELISA will be used to detect Ant-*Brucella* antibodies to assess seroprevalence status of dairy cattle in small holder dairy cattle.

2.6 Brucellosis prevention by vaccination

The control of brucellosis cattle in endemic areas is only achievable via mass vaccination (Godfroid *et al.*, 2011; Schurig *et al.*, 2002). The available vaccines (Albertian *et al.*, 2006; Schurig *et al.*, 2002) for cattle are all live (need a cold chain) and have the potential to become virulent in animals. Other drawbacks include conferring partial protection, resistance to some antibiotics, and interfere with diagnostic assays as they cannot be differentiated with naturally infected animals, which is important to achieve eradication (Ganesh *et al.*, 2014). However, the vaccines available and logistic arrangements to vaccinate all animals are costly and unaffordable for most resource-poor developing economy countries where brucellosis is prevalent and consequently amounts to no control (Ganesh *et al.*, 2014). For the past few decades in the world, researches have been carried out to develop safer *Brucella* vaccines for controlling brucellosis in animals (Avila-Calderón *et al.*, 2013). In the world, only three types of live attenuated *Brucella* vaccines have been recommended to control brucellosis in cattle. These are *B. abortus* strain 19 (S19) a smooth strain, and two rough mutant strains *B. abortus* RB51 and *B. abortus* 45/20 (Moriyón *et al.*, 2004; Schurig *et al.*, 2002; Sterne *et al.*, 1971).

The *B. abortus* S19 is a live naturally attenuated vaccine against bovine brucellosis that was developed in 1930, the used *B. abortus* was isolated from the milk of a Jersey cow. The virulent culture of *B. abortus* was left at room temperature for one year and was later found to be an avirulent and effective vaccine candidate (Buck, 1930; Thomas *et al.*, 1981). The *B. abortus* S19 has been effective in the control of brucellosis in adult cattle including the prevention of abortion. However, the O-ring which causes the smooth nature of the bacteria induces a strong humoral immune response, and the antibodies produced cannot be discriminated from those produced by natural infection (Avila-Calderón *et al.*, 2013). Discrimination between vaccinated from naturally infected animals (antibodies) can be done by using the competitive ELISA (Adams, 1990; Moriyón *et al.*, 2018; Nielsen *et al.*, 1995). Apart from that, *B. abortus* S19 vaccine is also infectious to humans and causes abortions in cows (Meyer, 1985; Nicoletti & Winter, 1990). However, the abortion rate due to the S19 vaccine is low, as reported in a study of 10 000 vaccinated pregnant cows at a gestation period between 7 and 8 months, the abortion rate was found to be less than 1% (Nicoletti & Winter, 1990).

The failure to differentiate the infected from vaccinated animals caused by the use of *B. abortus* S19 vaccine triggered further research to get a new vaccine candidate to solve the problem. In 1922, the *B. abortus* smooth strain 45/0 was isolated from a cow, after 20 passages in guinea pigs, a rough mutant was isolated and named *B. abortus* strain 45/20, which protected cattle from brucellosis (McEwen, 1940). Furthermore, this rough strain was heat-killed to avoid reversion to a virulent strain, carried in an adjuvant based on water and oil emulsion for vaccinating adult cattle, the vaccine was tested and found to be safer in pregnant animals and does not interfere with serological diagnosis (Moriyón *et al.,* 2004). However, when compared to *B. abortus* S19 vaccine in protection level, it is inferior especially when given as a single shot, therefore it requires two shots (Schurig *et al.,* 2002). Variability in protection level, severe local reaction at the injection site, unknown genetic defects, and unpredictable serological effect made the *B. abortus* strain 45/20 vaccine unpopular and eventually discontinued from use (Moriyón *et al.,* 2004; Schurig *et al.,* 2002).

The *B. abortus* RB51 is also a live vaccine and spontaneous rough mutant obtained following subculturing of virulent *B. abortus* 2308 on a medium containing antibiotics rifampicin and penicillin (Moriyón *et al.,* 2004; Schurig *et al.,* 1991). Later studies showed that the bacteria contain spontaneously inserted IS711 element which destructed the *wboA* gene encoding for a glycosyl transferase responsible for the O-side chain synthesis (Moriyón *et al.,* 2004; Vemulapalli *et al.,* 1999). Unlike the *B. abortus* strain 45/20 which is unstable and can revert to virulence, the *B. abortus* strain RB51 is stable and has been relatively widely used in some countries replacing the *B. abortus* S19 (Avila-Calderón *et al.,* 2013). The vaccine *B. abortus* RB51 is less virulent to humans than S19, however, there is little evidence to justify it as there was one case of RB51 in a veterinarian who was confirmed by bacteriological isolation and identification (Ashford *et al.,* 2004; Villarroel *et al.,* 2000). The use of monovalent vaccines (*B. abortus* S19 and *B. abortus* RB51) in controlling brucellosis in cattle is now challenged by the spill-over behavior of the *Brucella* species affecting domestic ruminants.

Breed selection is another way of controlling infectious diseases, however, it is difficult to select animals with resistance to all diseases; attention may be given to a few diseases of importance like those affecting production and public health especially in dairy farms while maintaining other desirable traits. In the dairy industry, milk yield has received more emphasis in genetic improvement programs and has produced significant increases in milk yield (Nash & Freeman, 2004).

However, the potential of genetic resistance to disease has not received direct emphasis in genetic improvement programs (Stone & Cundiff, 1985). The scientific findings suggest that genetic selection can be used to improve health in dairy cattle (Freeman & Kehrli Jr, 1994) to supplement good management practices (Nash & Freeman, 2004). Natural immune response trait has a heritability of 0.25-0.35 similar to those of milk production traits and above those of reproductive traits (Thompson-Crispi *et al.,* 2012). This potential can be used to control infectious diseases like brucellosis. The main challenge in selecting for disease resistance is to precisely identify the phenotype for disease resistance and to have consistent genetic markers with high predictive values for a disease phenotype (Wakchaure & Ganguly, 2016).

2.7 Immunity to Brucella

The host immune response is functionally divided into innate and adaptive or specific immunity. The innate immune system is the first line of defense which includes anatomical barriers (skin and internal epithelial layers), secretory molecules (various chemokines and cytokines, complement system, and opsonins), and cellular populations, such as phagocytes (neutrophils, monocytes/macrophages, dendritic cells, and natural killer cells). Adaptive immunity consists of T lymphocytes, which are involved in cytokine production and cytotoxicity (cellular immunity) as well as antibody-producing B lymphocytes (humoral immunity) (Parkin & Cohen, 2001). Macrophages and dendritic cells represent the professional antigen-presenting cells (APCs) that play a fundamental role in innate immunity, recognition, and in the induction of robust adaptive immunity against *Brucella* spp. (Skendros & Boura, 2013).

It has been known that about 2000 to 3000 genes control the host immunity system, providing mammals with a wide range of immune responses to fight against infectious organisms. Research shows that some genes have been associated with natural resistance to infection in cattle (Feng *et al.,* 1996; Hasenauer *et al.,* 2013; Qureshi *et al.,* 1996) In cattle, it has been known that the resistance to brucellosis is genetically determined (Borriello *et al.,* 2006). At first, the Nramp1 gene (natural resistance-associated macrophage protein 1) also known as Solute Carrier family 11 member A1(SLC11A1) was associated with the innate immune system and linked with natural resistance to brucellosis, tuberculosis, and salmonellosis (Feng *et al.,* 1996; Hasenauer *et al.,* 2013; Martinez *et al.,* 2008; Price *et al.,* 1990; Qureshi *et al.,* 1996). However, later studies on biological validation refuted the claim that SLC11A1 was associated with natural resistance to *Brucella* in cattle (Paixão *et al.,* 2007).

Furthermore, the gene was found to have polymorphisms because of variations in the number of GT dinucleotide repeats in the 3' untranslated region (3' UTR; positions 1781–1804) in Czech red pied and Czech black pied bovine breeds (Martinez *et al.,* 2008) in which 13 to 16 GT repeats have been identified (Adams & Templeton, 1998; Horín *et al.,* 1999). Based on single-strand conformation polymorphism analysis (SSCP) and Genome Wide Association Studies (GWAS), authors found a strong relationship between gene variation and natural resistance (Banos *et al.,* 2017; Barthel *et al.,* 2001; Hasenauer *et al.,* 2013). Genome-wide association study (GWAS) is the latest and best tool evolved to find out the association between complex diseases and genetic markers (variants) of an individual (Bush & Moore, 2012; Riancho, 2012).

2.8 Genome-wide association studies (GWAS)

Before the development and widespread use of Genome-wide association study (GWAS), complex human diseases and traits were studied using the candidate-gene association and linkage studies to genome approach, variants, genes, and pathways but failed to deliver definitive results (Consortium, 2007). Association studies are considered more powerful than linkage studies for the identification of genes attributing to the risk for common and complex diseases even before the GWAS studies came into use (Risch & Merikangas, 1996). Currently, genomic studies investigating human complex diseases and traits are done using GWAS (Naidoo *et al.,* 2011). The first GWAS study was published in 2005 which identified risk variants with a large effect size on the factor H gene associated with macular degeneration (Klein *et al.,* 2005), this was followed by large-scale Wellcome Trust case-control consortium work published in 2007 which set a stage for many more GWAS (Consortium, 2007). Currently, GWAS has been widely used in livestock for prediction or identifying qualitative and quantitative traits loci in livestock for improving productivity and disease control (Banos *et al.,* 2017; Mrode *et al.,* 2021; Petersen *et al.,* 2013; Raymond *et al.,* 2018).

Genome-wide association studies (GWAS) has been developed over the past twenty years to become a reliable and powerful tool for investigating the genetic structure of humans and their diseases (Bush & Moore, 2012). The rapid advancement of whole genome genotyping technologies such as Illumina and Affymetrix SNPs chips have contributed substantially to GWAS advancement.

In the year 2007, GWAS was recognized as a breakthrough in studying human genetic variation by the *Science* journal. During this year nearly 100 new GWAS publications relating to various complex diseases and traits were published (Pennisi, 2007). Apart from the long use and significant contribution of the GWAS and Single Nucleotide Polymorphisms (SNPs) genotyping microarrays analysis in human genetic studies, the same technologies have been used in studying livestock genetic diversity, in breeding and mapping of livestock diseases trait loci (Kijas *et al.*, 2012; Petersen *et al.*, 2013; Raymond *et al.*, 2018).

In genome-wide association studies, phenotypes are correlated with genotypes (genetic variations/SNPs) (Naidoo *et al.*, 2011), this is by identifying SNPs that exhibit high linkage disequilibrium (LD). Linkage disequilibrium is a non-random association of alleles at different loci (Montgomery, 2008). It is possible for different genomic regions with strong LD to be associated with phenotype (Visscher *et al.*, 2012). Statistically, one can calculate a measure for the probability that a single SNP is correlated with a phenotype of interest (i.e. disease) called odds ratio (OR). For example for analysis of G and T alleles in a 2*2 contingency table in a case and control groups, if OR =1 no association, OR>1 = G allele increases the risk of the disease (susceptibility), and if OR < 1 = T allele increases the risk of disease while G reduces the risk (resistance) (Zeng *et al.*, 2015). Apart from the OR, the SNP must have very high significant difference above the genome-wide association threshold, this can easily be seen visually on the Manhattan plot.

In disease studies, GWAS uses the hypothesis that common disease-common variant, which is also known as CD/CV model, and the additive genetic variance which means each SNP contributes partly to the heritability. Therefore, the observed phenotypic variation is associated with sets of SNPs in GWAS. Common diseases are attributable to SNPs which are represented in more than 1-5% of the population (Bleidorn, 2017).

The perfect method is yet to be found, GWAS has been challenged over the validity of its CD/CV model assumption especially when researchers found that 80% of the phenotypic variability remained unexplained in Crohn's disease, in a phenomenon known as "missing heritability" (Maher, 2008). Furthermore, the GWAS assumption of additive genetic variance has also been challenged as it ignores the fact that gene-gene interactions can be non-additive (Mckinney & Pajewski, 2012). However, some other studies did not consider the GWAS assumptions problematic and commented that the possible failure of GWAS may come from sampling errors (few SNPs for investigation) and model misspecification during statistical

analysis (Marjoram *et al.,* 2014). Notwithstanding, the GWAS approach has changed dramatically human disease genetics by identifying and mapping genetic variants by revealing statistically robust SNPs (Altshuler *et al.,* 2008; Donnelly, 2008) and GWAS now is used in livestock for diseases and production studies (Kijas *et al.,* 2012; Raymond *et al.,* 2018). The GWAS analysis can be done using different software to perform permutations, each with its advantage and disadvantage, such software includes PLINK, GEMMA, PRESTO, PERMONY, and others (Browning, 2008; Pahl & Schäfer, 2010; Purcell *et al.,* 2007; Zeng *et al.,* 2015).

2.8.1 Bovine Illumina 50K SNPs chip genotyping technology

In 2003, the first whole genome SNPs genotyping array could detect 10 thousand (10K) SNPs (Affymetrix GeneChip 10K, Santa Clara, CA). However, the current SNPs genotyping array (Illumina SNP chip, San Diego, CA) is capable of detecting more than 5 million SNPs per array (Illumina SNP chip 5M) (Kennedy *et al.,* 2003; Ragoussis, 2009). Genotyping technologies have significantly contributed to population genetics studies. The Bovine 50K Illumina™ SNP chip with a capacity of detecting 51 386 polymorphic SNP markers has been widely used to genotype cattle and other domestic animals. The SNP chip was designed by using a compilation of highly informative novel publicly available Bovine SNPs discovered using a reduced representation and next-generation sequencing technology platforms. The Bovine Illumina 50K SNP chip platform has been used to identify SNP markers from the DNA of the animals by the hybridization method. The SNP markers have been used for mapping qualitative trait genes/loci associated with disease phenotypes or predicting quantitative trait locus (QTL) associated with productivity in cattle and other livestock (Mrode *et al.,* 2021; Petersen *et al.,* 2013; Raymond *et al.,* 2018).

2.8.2 Linkage disequilibrium (LD)

Linkage disequilibrium is defined as a nonrandom association of alleles at two or more loci in the genome. The term is commonly used in population genetics for gene mapping as in each genomic region it reflects the gene conversion, mutation, the history of natural selection, and other forces that causes gene-frequency evolution in a population (Montgomery, 2008). For GWAS, information about LD structures of the study population is important for the interpretation of GWAS results and genomic selection for the improvement of viable traits of interest (Goddard & Hayes, 2009; Habier, 2010). The LD can help to learn the genetic

relationships among different breeds and the phylogenetic relationship between domestic animals and their primitive ancestors (Li & Kim, 2015).

Different statistical measures have been proposed to quantify LD depending on the context. The two most popular measures of LD between pairs of biallelic markers are D' and r^2. The r^2 is useful when the focus is on the prediction of one polymorphism (marker) given the other and the value range between 0 and 1, therefore is mostly used in power studies for association. On the contrary, D' is the measure of choice to assess recombination pattern and its value is between -1 and 1, thus haplotype blocks have been defined on account of D' (Chen *et al.,* 2006).

2.8.3 Minor allele frequency (MAF)

Any SNP has two alleles, minor and major alleles, and therefore two commonly occurring base-pair possibilities for a SNP location in a population (Altshuler *et al.,* 2010; Bush & Moore, 2012). The frequency of the SNP is explained in terms of the minor allele frequency (less common allele in a population). For example, the SNP with a minor allele (G) frequency of 0.40 implies that 40% of a population has the G allele versus the more common allele (the major allele), which is found in 60% of the population (Altshuler *et al.,* 2010).

It is the MAF of an individuals which causes phenotypic variations among members of the population. Many GWAS analysis has been discarding (cleaning) genotypes with low MAF of a particular threshold such as less than 10% (Cupples *et al.,* 2007; Florez *et al.,* 2007) as are considered to have a lower power to detect weak genotypic risk ratios than loci with high MAF i.e. 40%. The discarding causes a substantial loss of data by removing substantial SNPs for analysis and this affects the power to detect rare disease-causing polymorphisms (SNPs) (Gorlov *et al.,* 2008). The justifications for removing those SNPs with MAF <10% are such low genotyping rate/call rate of the genotyping platform used, perception (distrust) about statistical inference (inflate false positive results) that results from analyzing such SNPs (Tabangin *et al.,* 2009). However, recent studies have demonstrated that removing the SNPs with MAF of a certain threshold i.e <10%, is not necessary under the aspect of distrust of the results as they inflate false positive results (Moskvina *et al.,* 2006; Tabangin *et al.,* 2009). For errors associated with the genotyping platform (low call rate), imputation is normally done to replace the missed-out SNPs.

2.8.4 Heritability estimation by using analystical software for residual maximum likelihooh

Heritability is a statistic commonly used in genetics and breeding disciplines to estimate the degree of variation in phenotypic traits in a population that is due to genetic variation between individuals in a population. The estimation of heritability of dairy cattle with bovine brucellosis was calculated based on phenotypic trait and additive genetic variance components analysis. Variance components analysis is normally based on individual animal phenotypic traits of interest (disease status) as dependent outcome of interest and fixed and random model terms and general relationship matrix of individual animal. After calculating the variance components, by using prediction function for linear combination of variance components, the heritability is estimated by using heritability prediction function.

The ASREML version 4.1 statistical package in a UNIX environment is used to calculate variance components for estimation of heritability (h^2). The ASREML fits linear mixed models using Residual Maximum Likelihood (REML). The ASREML version 4 offers structural specification in the syntax, the variance structures for fixed effect model terms, random model terms, and residual error terms are specified accordingly in the linear mixed model definition by wrapping terms with the respective variance model functions. This approach is less error-prone, more concise, and more automatic for specifying multi-section residual variances. The ASREML has wide application in the analysis of repeated measures data, (un)balanced longitudinal data, multi-environment trials, (un)balanced designed experiments, and meta-analysis, univariate and multivariate animal breeding and genetics data (involving a relationship matrix for correlated effects), and regular or irregular spatial data. For specifying model formulae in ASREML, the general linear mixed model is given by: If y (n X 1) denotes the vector of observations, the general linear mixed model fitted by ASReml can be written as:

$$y = X\boldsymbol{\tau} + Z\mathbf{u} + \mathbf{e}$$

Where $\boldsymbol{\tau}$ (p X 1) is a vector of fixed effects, X (n×p) is the design matrix that associates observations with the appropriate combination of fixed effects, u (q × 1) is a vector of random effects, Z is the design matrix that associates observations with the appropriate combination of random effects and **e** (n×1) is the vector of residual errors. The ASReml assumes the vectors u and e are uncorrelated with each other. The linear mixed model is specified in ASREML as a series of model terms and qualifiers, model terms include factor and variate labels. The general

syntax linear mixed effect model is represented below, response is the label for the response variable(s) to be analyzed:

response [*qualifiers*] ~ *fixed* [**!r** *conrandom*] [**!f** *sparse_fixed*]
[**residual** *conresidual*]

~ is read as ' modeled as' and separates the response from the list of fixed and random terms in the linear mixed model, fixed represents the list of primary fixed explanatory terms/variables, that is, variates, factors, interactions, and special terms for which Wald F statistics are required, conrandom represents the list of consolidated model terms specifying both random effects and variance structures, residual statement allows specification of the residual error variance structure, conresidual is the list of residual consolidated terms specifying both random effects and variance structures. And as a general rule, all elements in the model must be space separated, elements in the model may be separated by + which is ignored except when it is at the end of a line which implies the model continues onto the next line.

The ASREML has a procedure to calculate certain functions of variance components. These functions enable calculation of heritability and correlations from simple variance components. Different modifiers for predicting (VPREDICT) the linear combination of variance components (F) for phenotypic and genetic variance components and (H) for heritability estimation which is a ratio of genetic variance over the phenotypic variance components.

2.8.5 Genome-wide association by using genome efficient mixed model algorithm

The GEMMA is the software implementing the Genome-wide Efficient Mixed Model Association algorithm for a standard linear mixed model and for GWAS (Zhou *et al.,* 2013; Zhou & Stephens, 2012). It fits a univariate linear mixed model (LMM) for markers (SNPs markers) association tests with a single phenotype to account for population stratification and sample structure, and for estimating the proportion of variance in phenotypes explained (PVE) by typed genotypes (Zhou *et al.,* 2013; Zhou & Stephens, 2012). Gemma also fits a multivariate linear mixed model for testing marker associations with multiple phenotypes simultaneously while controlling population stratification and is also used for estimating genetic correlation among complex phenotypes (Zhou & Stephens, 2012; Zhou & Stephens, 2014). The GEMMA requires four main input files containing genotypes (SNPs Markers), phenotypes (disease status of individuals), and a generalized relatedness matrix (grm) based on SNP genotypes and

covariates (optional). Genotype and phenotype files can be in two formats, either both in the PLINK binary ped format or both in the BIMBAM format, mixing up files format will result in unwanted errors. The ped format has three files with the extension: *.bed, *.bim, and *.fam, all with the same prefix. The second column of the *.fam file contains the individual id (animal Id) and the sixth column is the phenotype (disease status, labeled control 0 and cases as 1), column 5 of the *.fam file contain minor allele and column 6 major allele of which Gemma is going to read them during analysis. The *.bed file is binary and contains computer-readable information (nonhuman readable information) which is used during analysis. The binary PLINK *.bed files (*.bed/*.bim/*.fam) were made from standard ped files by using PLINK software, the file format GEMMA uses for GWAS analysis. The GEMMA, as a linear mixed model software, requires a relatedness matrix file for GWAS analysis in addition to ped files (genotype and phenotype files). For the relatedness matrix, the original general relatedness matrix used during heritability estimation was used. The GEMMA fits a linear mixed model with an intercept term if no covariate file is available, for my analysis covariate file was provided to be included in the model analysis.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study area

The study was conducted across six administrative regions and 23 local government authorities of Tanzania in 2 agroecological zones; with the Northeastern highland regions comprised of Arusha, Kilimanjaro, and Tanga and Southern highland regions comprised of Iringa, Njombe, and Mbeya (Fig. 1). In these agroecological zones, there is a range of dairy breeds crossed with either Ankole or East African Shorthorn Zebus (EASZ) (Njombe *et al.*, 2011).

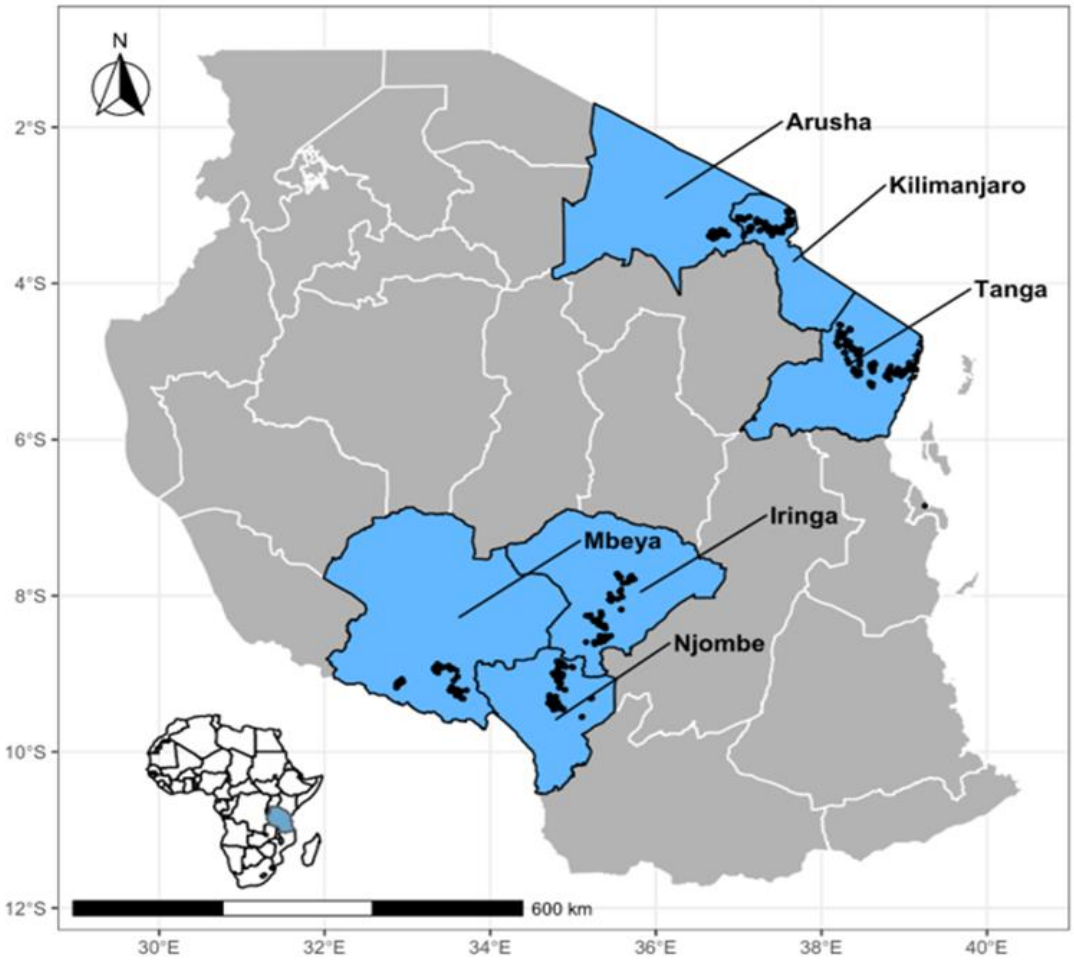


Figure 1: Map of Tanzania (with inset map of Africa) showing the six study regions with a high population of smallholder dairy cattle (blue) and the approximate locations of the sampled animals (black dots)

The 6 studied regions included animals from cross 23 local government authorities. Table 1 summarizes the local government authorities involved in the study.

Table 1: List of regions and local government authorities participated in the study

Region	Local Authority	Region	Local Authority
Arusha	Arusha CC	Iringa	Iringa MC
	Arusha DC		Iringa DC
	Meru DC		Mufindi DC
Kilimanjaro	Hai DC	Njombe	Mafinga TC
	Rombo DC		Njombe TC
	Moshi Rural DC		Njombe DC
	Siha DC		Makambako TC
Tanga	Tanga CC	Mbeya	Mbeya DC
	Korogwe TC		Mbeya CC
	Korogwe DC		Mbozi DC
	Lushoto DC		Rungwe DC
	Muheza DC		
	12		**11**

Key: DC= District Council, TC= Town Council, CC= City Council

The two agroecological zones are characterized by high altitude, humidity, and a high chance of experiencing high rainfall and temperature in the first six months of the year (TMA, 2021). The median altitude of northerneastern zone and that of Southern highland zone are 2700 m and 2520 m, respectively. Furthermore, the Northereastern zone has two mountains, Mount Kilimanjaro (5895 m) and Mount Meru (as Mount 4566 m) and Southern highland zone has Mount Rungwe (2981 m). The weather condition in these two agroecological zones supports the growth of several cash and food crops like coffee, tea, banana, potatoes, maize, and beans just to mention a few.

Figure 2 shows the temperature and rainfall distribution pattern over the year in the study areas. Heavy rainfall starting in Mach to June with a low probability of rainfall from July to December apart from short rainfall during October and November. This study was carried out throughout the year. Some places were very difficult to reach due to rainfall, mountainous terrain and slippery mud road, particularly in the Mufindi district in Iringa region.

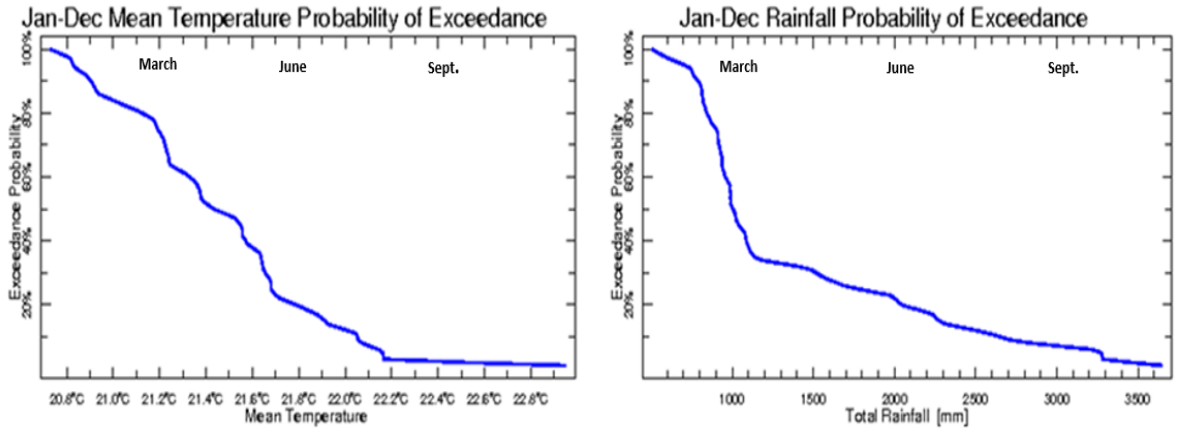


Figure 2: Mean Temperature and Rainfall pattern in study areas (Tanzania Meteorological Authority [TMA] 2021)

The two zones' human population structure is characterized by an average household size of 4.4 persons, average illiteracy of 12.08% for males above 15 years and 14.8%, for females above 15 years. The average population density is 66.5 persons per square kilometer, which is above the national population density of 51 persons per square kilometer (Statistics, 2012).

According to the 2019-2020 national agricultural census report, the three regions with the highest number of dairy cattle in decreasing order were Tanga, Arusha, and Mbeya. The other 3 regions, Kilimanjaro, Iringa, and Njombe included in this study has the 5th , 9th and 11th largest dairy cattle populations (NBS) (Statistics, 2021) (Fig. 3). According to the household budget survey of 2018, all six regions were above the food poverty line of Tanzanian Shillings (TZS) 33 748 per person per month (Planning, 2019).

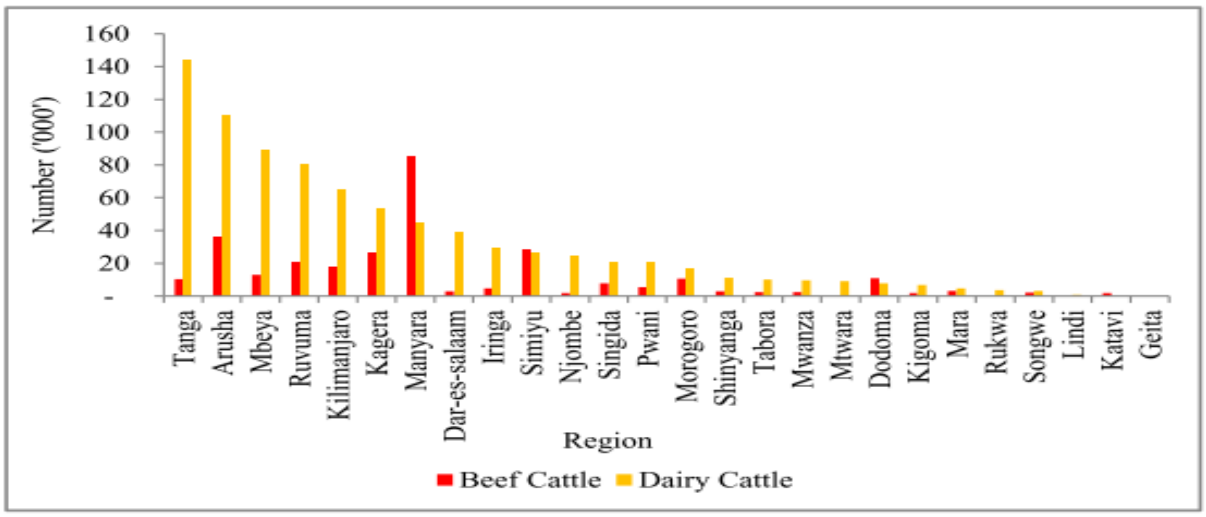


Figure 3: Number of improved cattle by type per region as of 1st August 2020 in Tanzania Mainland (NBS, 2021)

3.2 Study design

A cross-sectional study design was employed from July 2019 to October 2020. A cross-sectional study is one of the observational studies that involve the selection of the sample of n individuals from a large population regardless of the presence or absence of an outcome of interest. This is always followed by the determination of disease status in each selected individual animal (Thrusfield, 2018).

3.3 Sampling design

3.3.1 Sampling units and sample size

Dairy cattle in this study were selected from a subset of the dairy cattle registry of the Africa Dairy Genetics Gains (ADGG) (https://data.ilri.org/portal/dataset/adgg-tanzania) program. Over 4000 dairy cattle were randomly selected from the genotyped ADGG cattle registry. These cattle had known genetic characteristics and could be identified by their preliminary information such as the owner, ear tag number, age, and sex from the ADGG database.

The ADGG project used multistage sampling approach as was explained by Dahoo (Dohoo *et al.,* 2014) involving a random selection of herds and individual animals. The primary sampling unit of the study was herd and the secondary sampling unit was individual animals in the herd. The ADGG project randomly selected dairy herds from a reference point of an influential dairy farmer in the respective area and then recruiting other herds in a circle moving out from that particular herd and farmers' were asked to consent for their participation in the project.

The estimation of an appropriate sample for studying the disease prevalence in a finite population was calculated using the simplified formula deduced by Yamane in 1967. Yamane formula assumes a 95% confidence level, the target population is more homogenous of the attribute of interest (disease), and therefore the degree of variability of the target population (P) is 0.5, level of precision/sampling error is 0.05 (±5%) (Gicheru *et al.,* 2015; Israel, 1992; Kasiulevičius *et al.,* 2006).

$$n = N/(1 + Ne^2)$$ Equation…… 1

Where:

n = sample size, N is population size and e is level of precision.

According to the National Bureau of Statistics, the improved dairy cattle population is 865 628 (Statistics, 2021).

Therefore: n = 865, 628/1+ 865, 628 $(0.05)^2$ = 385 dairy cattle.

Therefore, a conservative sample size calculation for an unknown prevalence of disease (50% is assumed) in a finite population gives a sample size of 385 animals (this does not adjust for potential clustering within herds). However, to increase the statistical power and there were 4000 animals on bigger African Dairy Genetic Gain (ADGG) programme original lists and 400 animals would have been sufficient; however, 2048 dairy cattle were sampled across the study regions as I wanted to estimate risk factors and collect materials for sequencing.

3.3.2 Sample size estimation for genetic association

The formula for the Genetic Power Calculator (http://zzz.bwh.harvard.edu/gpc/cc2.html) (Purcell, 2003) requires the following variables/parameters: the disease prevalence, the minor allele frequency (MAF), the linkage disequilibrium - non-random association of alleles at two or more loci in a general population (LD), the ratio of cases to controls, an odds ratio for the effect size and the error rate (α). The parameters were estimated as below:

Disease prevalence 5 % (4.4 % was considered a brucellosis prevalence in dairy cattle when introducing dairy animals in the country) (Shirima, 2005), MAF 5 %, Case to control ratio 1:4 since there will be approximately 200 cases (based on the prevalence of 5 %) out of 4000 dairy cattle expected to be sampled, LD = 0.514. Estimate of odds ratio 2.0. No references were found for brucellosis therefore this is estimated from the study where the odds of bovine tuberculosis in cattle with different gene polymorphisms was 2.02-2.17, power = 90 % and α=0.05 (Tsairidou *et al.,* 2014). Therefore, the number of cases (phenotypic trait positive=PCR positive individuals) required for a SNP marker is 531. However, the project was expecting to sample about 4000 dairy cattle, this would increase the power to detect smaller effects. According to the number of dairy cattle registered and genotyped by ADGG 2300 cattle would be sampled in Arusha and Kilimanjaro, 891 cattle in Tanga and 809 in the Southern highland zone across the 3 regions of Mbeya, Iringa and Njombe.

3.4 Data collection and Analysis

Questionnaire data collection was done by interviewing the head of the household or someone knowledgeable with the farm. This part will answer specific objective number one, risk factors. A pre-tested structured questionnaire (Appendix 3) was administered to individual farmers from each herd by a trained researcher. The aim and objectives of the study, the benefits, and his or her role in the study were explained and if farmers consented (Appendix 1), they were interviewed and their cattle sampled. The questionnaire collected information on the farmer's livestock background, general animal husbandry practices, diseases control measures, probable risk factors for brucellosis transmission, and animal biodata such as sex, age, breed, parity, and body condition. A tablet installed with the Open Data Kit tool (ODK) was used to collect the information (Fig. 4). The tool was developed and piloted before a final version was uploaded using the ODK. In the evening of sampling day, all filled questionnaires were sent to International Livestock Research Institute servers in Kenya for storage. Sample tube barcode, animals identification number, and collection date were recorded in a Microsoft Access ® database which was later linked to the ODK questionnaire metadata for data cleaning, coding and analysis using R software.

Figure 4: Field questionnaire interview: (a) A farmer holds his copy of the signed consent form prior to the personal interview in Njombe region, (b) Actual data collection through farmer's interview in Iringa region

Blood samples were collected from all identified dairy cattle. Blood sampling from dairy cattle was done aseptically as was explained by other studies (Mugizi *et al.*, 2015) and (Shirima & Kunda, 2016). In this study, dairy cattle were manually head restrained using ropes (halter), and blood was collected through jugular venipuncture. A total of 26 ml of blood was collected from each animal using a sterile needle in 2 plain (red-top) and 1 EDTA (purple–top)) tubes (BD Vacutainer®, Becton, Dickinson, and Company, Franklin Lakes, USA). Tubes were then labeled with the date of collection, last 4 digits of the animal ear tag number and also labeled using a field barcode. Labelled tubes were placed in a rack upright and stored in a cool box containing ice packs, and then transported to the district laboratory on daily basis. Blood samples in plain tubes were allowed to pelletize overnight and then centrifuged at 3000 revolutions per minute for 5 minutes and the serum was collected and aliquoted into the 4 x 2 ml cryovial tubes which were also labeled with animal identification number, date of collection and new laboratory barcode and stored in a deep freezer at -20℃.

The field barcode, laboratory barcode, animal identification number, and date of the collection were captured in a Microsoft Access® database (to be linked with the ODK questionnaire metadata) during serum preparation. For EDTA blood, blood was gently hand mixed by inverting the tube two times and aliquoted into 2 x 2 ml cryovial tubes which were also labeled with animal identification number, date of collection and new laboratory barcode and stored in a deep freezer at -20℃. Finally, while maintaining the cold chain, the cryovials (serum and whole blood) were transported to the Nelson Mandela African Institution of Science and Technology (NM-AIST) in Arusha, Tanzania where they were stored at -20℃ in a deep freezer until serological and molecular analysis were conducted.

3.4.1 Seroprevalence and risk factors

For this objective, the data partly were collected through questionnaire survey for the re-emergence of brucellosis in the smallholder dairy farms in selected study regions, and partly were collected through blood samples that were collected from all identified dairy cattle.

(i) Serological test for detection of anti-*Brucella* antibodies

Data collection procedure

All samples (2048) were tested using a competitive enzyme-linked immunosorbent assay (cELISA) according to the manufacturer's indications (COMPELISA 160 & 400, APHA

Scientific). The test was conducted in a 96-well polystyrene plate that was precoated with purified *B. melitensis* lipopolysaccharide (LPS).

Briefly, the test serum and all ELISA reagents were allowed to come to room temperature on the bench, except for conjugate stored at -20℃ which was taken out just before use. The 20 µl of each test serum was placed in a well from columns 1 to 10 and in columns 11-12 there were 6 well for positive control (20 µl) serum, 6 wells for negative control (20 µl) serum, and 4 wells for the conjugate (20 µl) control. Immediately after reconstitution, 100 µl of the conjugate was added to all 96 wells.

The plate was then covered with a lid and then vigorously shaken (200 rpm) in a microtiter plate shaker for 2 minutes to allow mixing up serum and conjugate and later incubated at room temperature (21±6℃) for 45 minutes (instead of 30 minutes) on a rotary shaker at 160 rpm. Then, the contents of the plate were shaken out and the plate was washed 5 times using washing solution under low pressure, the plate was then dried up by taping onto a layer of absorbent towels until no more liquid was removed.

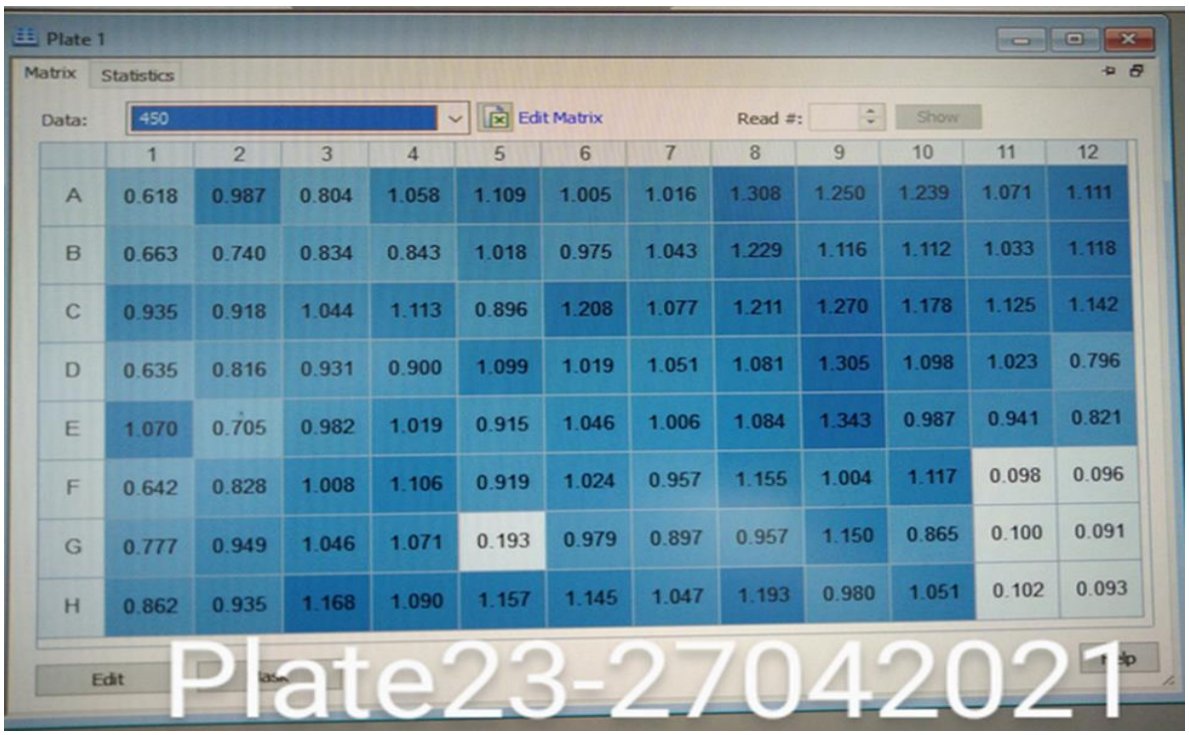
Then 100 µl of prepared o-phenylenediamine dihydrochloride (OPD)-chromogen/substrate solution was added to all wells and the plate was incubated at room temperature for a further 20 minutes, while incubating the plate (Fig. 5), the microplate reader was switched on. The reactions then stopped by adding 100 µl of provided "stopping solution" to all wells. Condensation at the bottom of a plate was removed using the paper towel before it was placed on a SYNERGY|HTX multi-mode reader (BioTek Instruments, Winooski, USA), the plate was read at 450 nm within 10 minutes.

Results interpretation

Lack of color development indicates that, the tested serum sample was positive (Fig. 5) at position G5 circled red. A positive/negative cut-off value was calculated at 60% of the mean of optical density (OD) of the four conjugate control wells (Fig. 6 and 7). Any test sample giving an OD equal to or below this value was considered as being positive (Bronsvoort *et al.,* 2009; Stack *et al.,* 1999).

Figure 5: Prepared ELISA plate after the addition of OPD and incubation ready to be set in a microplate reader spectrophotometer

Key: C- - = negative control well (3*2 wells), C+ = conjugate control wells (2*2 wells), and C++ = positive control wells (3*2 wells)



	1	2	3	4	5	6	7	8	9	10	11	12
A	0.618	0.987	0.804	1.058	1.109	1.005	1.016	1.308	1.250	1.239	1.071	1.111
B	0.663	0.740	0.834	0.843	1.018	0.975	1.043	1.229	1.116	1.112	1.033	1.118
C	0.935	0.918	1.044	1.113	0.896	1.208	1.077	1.211	1.270	1.178	1.125	1.142
D	0.635	0.816	0.931	0.900	1.099	1.019	1.051	1.081	1.305	1.098	1.023	0.796
E	1.070	0.705	0.982	1.019	0.915	1.046	1.006	1.084	1.343	0.987	0.941	0.821
F	0.642	0.828	1.008	1.106	0.919	1.024	0.957	1.155	1.004	1.117	0.098	0.096
G	0.777	0.949	1.046	1.071	0.193	0.979	0.897	0.957	1.150	0.865	0.100	0.091
H	0.862	0.935	1.168	1.090	1.157	1.145	1.047	1.193	0.980	1.051	0.102	0.093

Figure 6: The Optic density (OD) readings from the spectrophotometer for the plate above

Plate acceptance criteria

The plate was considered valid when the mean OD of the 6 negative control wells was greater than 0.700 (the optimal mean negative OD is 1.00), the mean OD of the 6 positive control wells was less than 0.100, the mean OD of the 4 conjugate control wells was greater than 0.700 (the optimal mean conjugate control wells is 1) and lastly when the binding ratio was greater than 10.

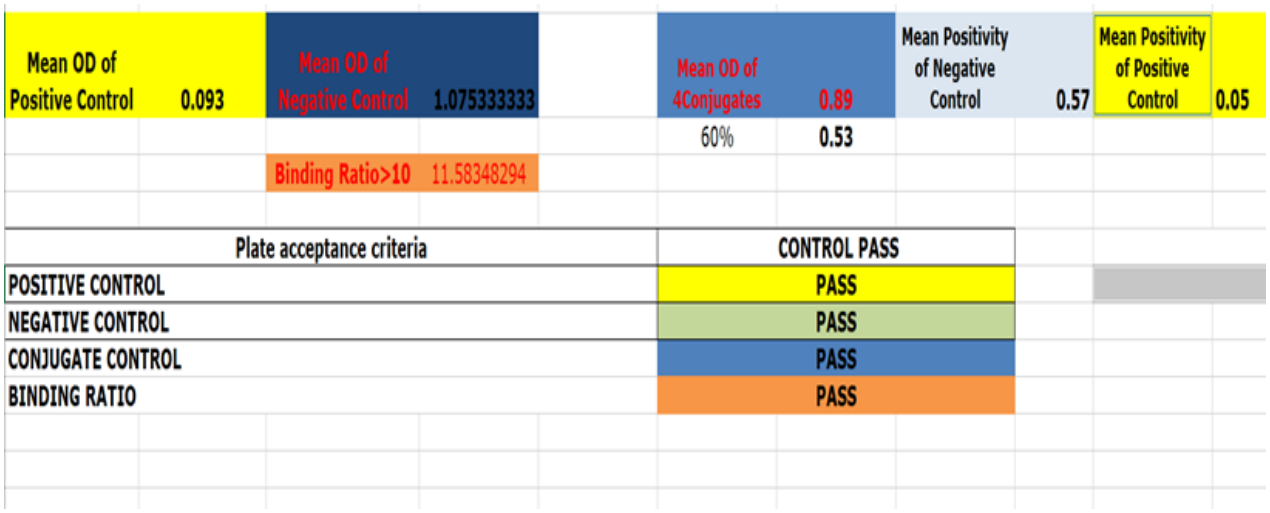


Figure 7: Conditions for analysis and plate acceptance criteria

(ii) Risk factors for brucellosis

Data collection procedure

Different aspects of risk factors were included in the questionnaire such as farmer's particulars, animal's particulars, herd management practices, and diseases control practices. Farmer's data such as age, gender, education level, experience in years of keeping dairy cattle, if attended livestock training (formal or informal), awareness and knowledge of zoonotic diseases from cattle, reasons of keeping dairy cattle was collected. Animal data such as age, dentition score, sex breed, body condition score, parity, last date served, pregnancy status, history of abortion, genital discharges, udder condition, excessive salivation, and lameness were also collected.

Herd management practices particulars were also collected: Distance between dairy farm, number of female cattle (cows and heifers), number of bulls, breeding method, keeping bull, keeping of goats, sheep, dogs, and pigs in the cattle house, feeding management, introduction and source of new animals in the herd for the past 12 months, and source of drinking water. For disease control management practices details such as history of routine vaccination and

against which diseases, milk preparations before drinking, disposal methods of fetal membranes, who milks the cow, etc. The farmer or someone knowledgeable about the farm was interviewed by face to face, occasionally the interview was done by phone call. These risk factor data were considered as explanatory variable to the outcome variable brucellosis serostatus in the univariable and multivariable analysis.

Statistical data analysis for brucellosis risk factors

Questionnaire data was downloaded from ODK and laboratory data in the Microsoft® Access 2013 database were imported into RStudio and were joined and cleaned before analysis. All statistical analyses were performed using RGui (64-bit) version 4.0.4 (2021-02-15) and RStudio version 4.1.2 (2021-11-01).

Animal-level seroprevalence was calculated as a proportion of seropositive animals to the total number of animals tested (animals tested positive/total number of animals tested) for both overall animal level and regional seroprevalences. In addition, an adjusted seroprevalence with 95% exact binomial confidence interval was estimated across the study area by adjusting for a stratified sampling design with regions of variable estimated cattle populations using *svydesign* and *svyciprop* functions of the survey R package (Lumley, 2011). The 95% confidence interval (CI) for binomial proportions was implemented in the *binomCI* function for univariable and multivariable analysis and two-sided *binom.test* function of *stats* package for individual level seroprevalences.

To measure associations between all explanatory variables to the binary response variable (seropositive or seronegative), estimated odds ratios and confidence intervals (0.95 confidence level) using conditional maximum likelihood and normal approximation, respectively were all implemented in the *epitools* R package (Aragon *et al.*, 2017). Variables with statistically significant (p-value < 0.05) associations and those with p-value <0.25 in univariable analyses were chosen for multivariable analyses. To model the relationship between our ELISA binomial results and a set of covariates, the binomial (logistic) generalized linear model (GLM) was built to understand the behaviors of variables without the random term effect and followed by binomial (logistic) generalized linear mixed effects models using *glmmTMB* package with a log link function implemented in the template model builder *glmmTMB* package (Bates *et al.*, 2015) and assuming a linear relationship.

$$Yij \sim \text{Bin}(1,\ pij)$$

$$E(Yij)\ =\ \sim (\ pij) \qquad \text{Equation…….2}$$

$\text{logit}(pij) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots \beta_{ij} x_{ij} + \alpha_i$

$\alpha_i \sim N(0, \sigma^2)$

Where, Yij is the *j*th ELISA result binomially distributed with a conditional probability, pij, in district *i*, and $i = 1,\ldots, 23$, and district, α_i, is the random intercept, which is assumed to be normally distributed with mean 0 and variance σ^2.

Continuous fixed effects variables were mean-centered and scaled to standard deviation using the *scale* function. To avoid multicollinearity, the Pearson correlation test was run on the variable pairs implemented in the ggpairs function from the GGally R package (rho < 0.29) based on Cohen (Jacob, 1992). A backward stepwise model selection approach was carried out to eliminate one variable at a time based on our model best-fit criteria. For instance, I kept a nested model with the lowest Akaike Information Criterion (AIC= 420.1 for GLM and AIC 331.2 for GLMM) and significant (*p*-value < 0.05) statistics from likelihood ratio tests were kept.

In parallel, marginal and conditional R^2 for GLMMs (Nakagawa *et al.,* 2017; Nakagawa & Schielzeth, 2013) were calculated using the *rsquaredGLMM* function implemented in the *MuMIn* package (Barton, 2022) to select the model explaining most of our data variance. The results were, for conditional R^2 (variance explained by random and fixed factors) was 0.811 and for marginal R^2 (variance explained by fixed factor-district name) was 0.171. The conditional R^2 results suggest that the model was good. The intra-cluster correlation coefficient (icc) of the best-fit GLMM model for Kilimanjaro,Tanga and Njombe with high number of seropositive animals was calculated using the packages (lme4, insight, and performance) for calculating the variances and icc.

Furthermore, the validation of the best model was done by plotting the model-predicted values and fixed effects against randomized scaled quantile residuals simulated using the *simulateResiduals* function from the Diagnostics HierAchical Regression Models (*DHARMa)* package (Hartig, 2022). The model was valid if simulated residuals plotted versus predicted

values and each fixed effect showed no clear clustering patterns (Fig. 8a), and deviations from the expected quantile distribution were not significant (p-value > 0.05) (Fig. 8b).

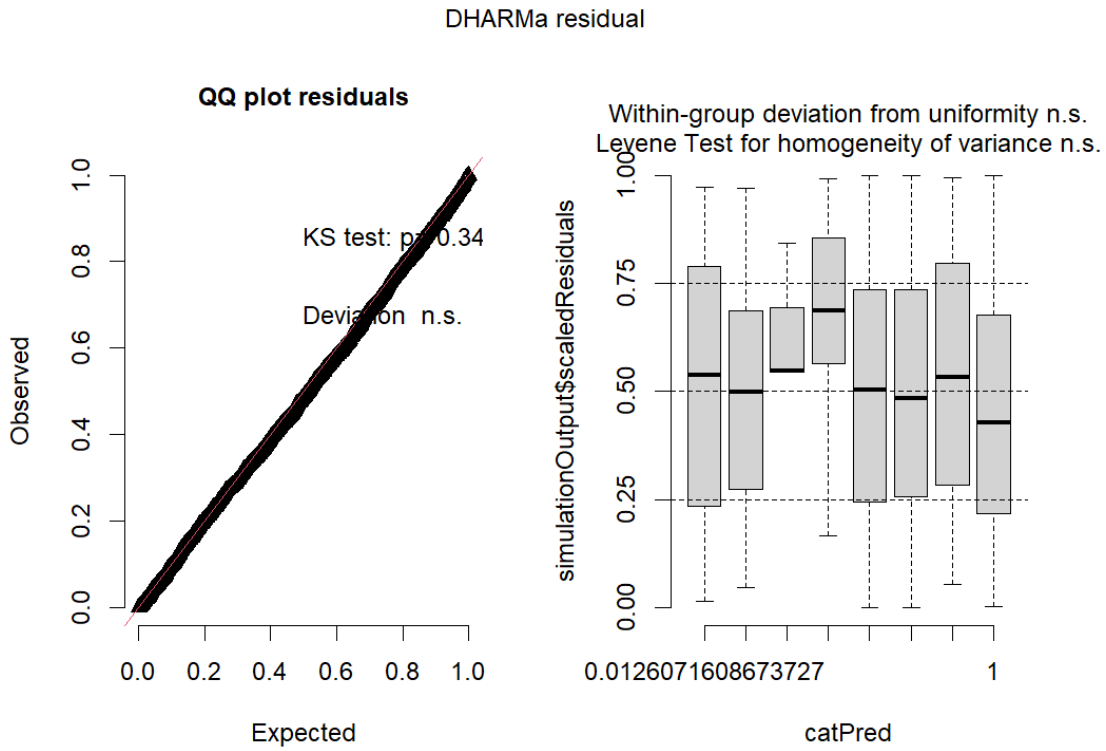


Figure 8: Modal validation plots, (a) QQ-plot residuals (Deviation not significant at *p*>0.05 and (b) Residual (spline line) vs predicted values (flat lines) plot, showing not significant different deviations at 0.5 quantile distribution (50% interquartile range)

Additionally, a Q-Q plot was visualized to detect overall deviations from the expected distribution which includes tests for correct distribution (KS test), overdispersion, and outliers. The results of the dispersion test of the simulated value via the standard deviation (sd) of the residual were drawn and found that the dispersion (within group deviation from uniformity) was not significant (n.s) (p>0.05), which mean no dispersion of the expected (simulated values) from observed values and are in correct distribution, in addition, the Levene test for homogeneity of variance was not significant (n.s) suggesting that the variances of observed and expected values were not significantly different (Fig. 8b).

The nonparametric dispersion tests via standard deviation (sd) of residuals fitted vs simulated values showed that the simulated values (red line) fitted the model (p=>0.05), suggesting that, there was no significant difference between the observed and simulated values (Fig. 9).

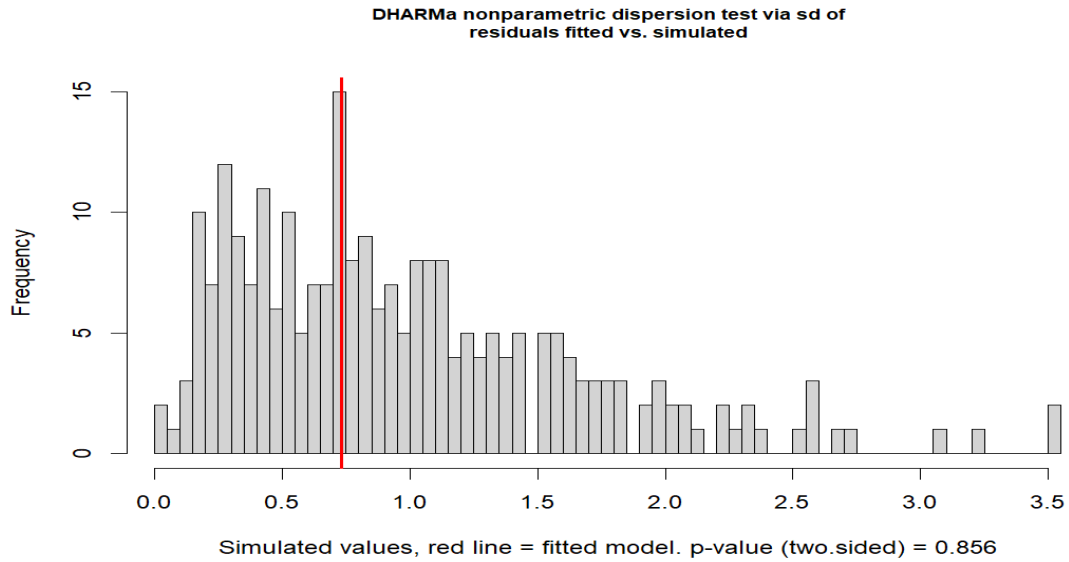


Figure 9: Model testing for dispersion of the residuals from observed values as predicted by the model showing no significant difference between the values (*p*=0.856)

Zero inflation is the situation when more zeros appear in the observation than expected under the fitted model, it compares the distribution of expected zero against the observed zeros. The results showed that the model did not count zero values (no zero-value histogram at x-axis = 0) suggesting that the modal was perfectly predicting the correct values accordingly (Fig. 10). However, the zero inflation result was not one of the best-fit model selection techniques such as AIC, BIC, WAIC, the test was only run after the selection was done to validate the decision that was taken in choosing the best-fit model.

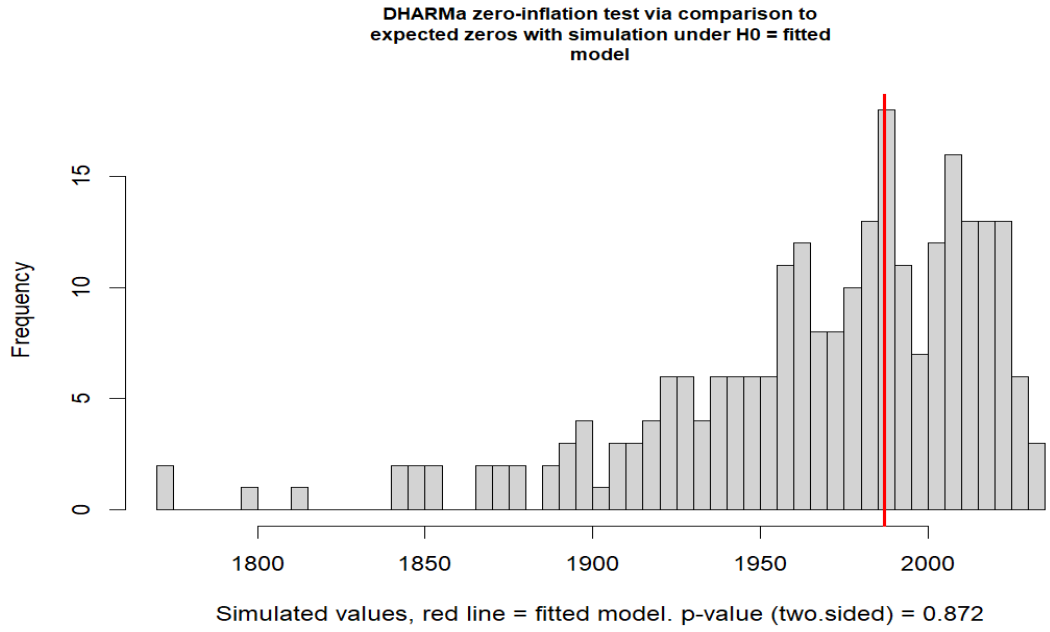


Figure 10: Zero inflation test for the best-fitted modal showing no zero inflation (no zero values histogram at x-axis = 0)

For outlier tests of the fitted model, the simulated value can be higher or smaller than the observed values in which case they got the residual values of 0 or 1 respectively. The results showed that the outlier test was not significant as there were too few outliers at 1 (Fig. 11). A lack of outliers would not be expected as would have been caused by under dispersion of the simulated values.

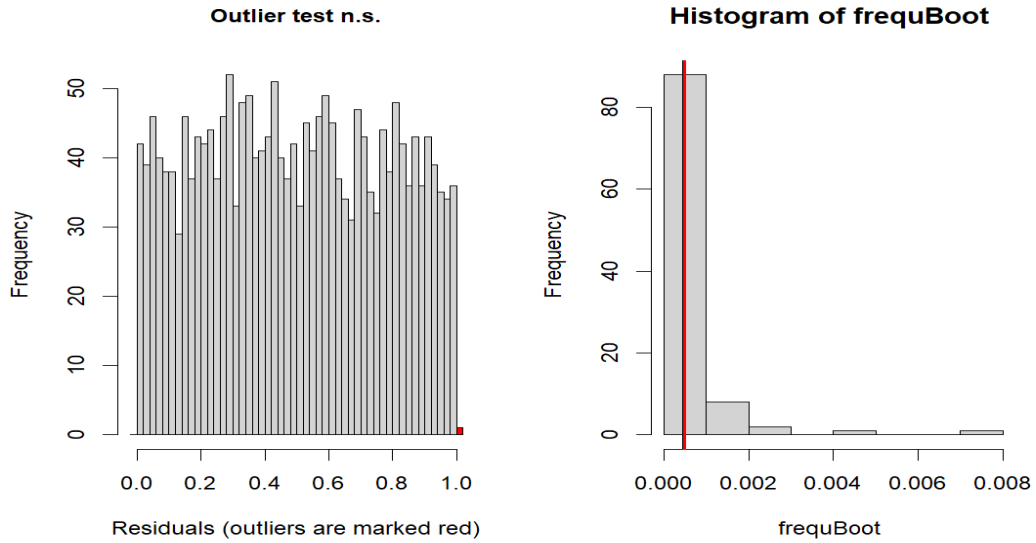


Figure 11: The outlier results of the fitted model showing too few outliers residual marked red on the x-axis = 1 on the left figure and a tiny discrete histogram on frequBoot axis on the right figure

Bernoulli spatial scan analysis

A spatial scan statistic was used to detect statistically significant spatial clusters of seropositive animals in Njombe and Kilimanjaro regions. Cluster analyses were performed using SaTScan™ v10.1 software (Kulldorff, 2009) with a Bernoulli model for binary events (i.e. seropositive/seronegative). The SaTScan uses Monte Carlo hypothesis testing to obtain the p-values and SaTScan adjusts for the underlying spatial inhomogeneity of a background population. For each location and size of the scanning window, the alternative hypothesis was that there was an elevated risk within the window as compared to outside and a likelihood ratio test was performed. Multiple different window sizes were used and the locations were the latitude/longitude for each animal with slight jittering to avoid more than one animal being at a location. The window with the maximum likelihood was the most likely cluster, that was, the cluster least likely to be due to chance. A *p*-value was assigned to this cluster. For this analysis we used 9999 Monte Carlo replications, and a cluster was considered statistically significant if the *p*-value was < 0.05.

3.4.2 Identification of Brucella species circulating in smallholder dairy cattle

(i) Genomic DNA extraction from blood, and vaginal swabs samples

Data collection procedure

- **DNA extraction**

Genomic DNA extraction from different samples was done manually using Qiagen DNeasy® Blood & Tissue kit (QIAGEN, German). Briefly, for DNA extraction from blood, cryovials containing EDTA blood were taken out from the freezer and allowed to thaw to room temperature. The 20 µl (20 microliterr) of proteinase K was pipetted out into a 2 ml microcentrifuge tube. Followed by the addition of 100 µl of anticoagulant-treated blood after vortexing for 30 seconds. The final volume was then adjusted to 220 µl by the addition of 100 µl of PBS. Then 200 µl of Buffer AL was added followed by vortex mixing and the mixture was then incubated at modified temperature of 58.1°C (instead of 56°C) for 20-30 min modified time (instead of 10 min) for optimal results.

After incubation, 200 µl of absolute ethanol was added and the mixture was thoroughly mixed by vortexing and then spanned for 30 sec at 8000 rpm. The whole mixture was pipetted into a DNeasy Mini spin column placed in a 2 ml collection tube which was then centrifuged for 1 min at 8000 rpm. The flow through in a collection tube was discarded and the spin column was placed in a new 2 ml collection tube. Then 500 µl of buffer AW1 was added and centrifuged for 1min at 8000 rpm. The flow through in a collection tube was discarded and the spin column was placed in a new 2 ml collection tube.

Then 500 µL of buffer AW2 was added and centrifuged for 3 min at 14 000 rpm. The flow through and collection tube was discarded, and the spin column was placed in a new 2 mls microcentrifuge tube. Elution of DNA was done by the addition of 100 µl of buffer AE at the center of the spin column membrane, followed by incubation for 20 min at room temperature. This was followed by centrifugation for 1min at 8000 rpm. The spin column was discarded, while DNA collected in a microcentrifuge tube and stored in a freezer at -20°C.

For vaginal swab samples, DNA extraction procedures were the same except at the initial stage where 100 µl of more swab was added instead of addition of 100 µl of PBS.

- **Assessment of DNA quality after extraction**

Two methods, agarose gel electrophoresis and nanodrop spectrophotometer for DNA quality and purity checks were used on different random samples. Nanodrop 8000 spectrophotometer (Thermo Scientific) with Nanodrop 8000 –V2.2.1 software was used to check for DNA quality and purity. For a spectrophotometer to estimate the purity and quality of DNA, nucleic acids absorb different UV light at a different wavelengths, most absorption occurs at UV light wavelengths of 260 nm due to the resonance structure of the purine and pyrimidine bases. Then, the absorbance was converted into ng/µl of double-stranded DNA (dsDNA) using the established conversion factor of 50 ng/µl for 1 optical density unit at 260 nm. For DNA purity estimation, a ratio of absorbance at 260 nm/280 nm is used, if the ratio is ~1.8, the DNA is generally considered pure. Good-quality DNA will have an A260/A280 ratio of 1.7–2.0. However, a reading of 1.6 still makes DNA suitable for use (Scientific, 2019; Wang, 2016).

The results from the nanodrop spectrophotometer DNA quality check (Table 2), showed that, most of the samples passed the quality check except for sample number 1, 6, and 8 giving an extraction efficiency of 75%. Ten microliters (10 µl) of the DNA sample were used to set on the nanodrop spectrophotometric eye for reading as indicated by the manufacturer. The eye of the nanodrop spectrophotometer was dry-cleaned using soft tissue before placing another DNA sample.

Table 2: Nanodrop quality checks for the DNA from random blood samples extracted using Qiagen DNeasy® Blood & Tissue kit

Well	Sample ID	Conc.(ng/ul)	A260	A280	A260/A280
A1	1	5.737	0.115	0.089	1.3
A2	2	14.75	0.295	0.165	1.79
A3	3	28.49	0.57	0.319	1.79
A4	4	11.93	0.239	0.127	1.88
A5	5	5.485	0.11	0.061	1.80
A6	6	21.7	2.455	1.792	1.37
A7	7	13.55	0.271	0.16	1.69
A8	8	58.54	1.171	0.756	1.55
A9	9	10.22	0.204	0.112	1.82
A10	10	20.04	0.401	0.208	1.93
A11	11	12.98	0.16	0.085	1.87
A12	12	10.35	0.207	0.108	1.91

Agarose gel electrophoresis was also done on random samples to check for DNA quality (Fig. 12). A 1% agarose gel electrophoresis was used to check for DNA degradation. If the results of agarose gel electrophoresis show only high molecular weight genomic DNA bands and no other bands or smears, then it means that the DNA sample has not been degraded during the extraction process (Wang, 2016). Results show that, of the first twelve random samples (1-12) (Fig. 12), samples 4 and 6 did not have enough DNA. For the second twelve samples (13-24) (Fig. 12), samples 20 and 22 had low quantities of DNA. In both rows, the first and second wells were used for positive and negative controls. There was no degradation in any of the 24 samples making an extraction efficiency of 100%. Twenty microliters (20 µl) of the DNA sample were placed in respective wells except for negative control wells where 20 µl of molecular grade water was used and 20 µl of the Lambda for control wells (quality DNA).

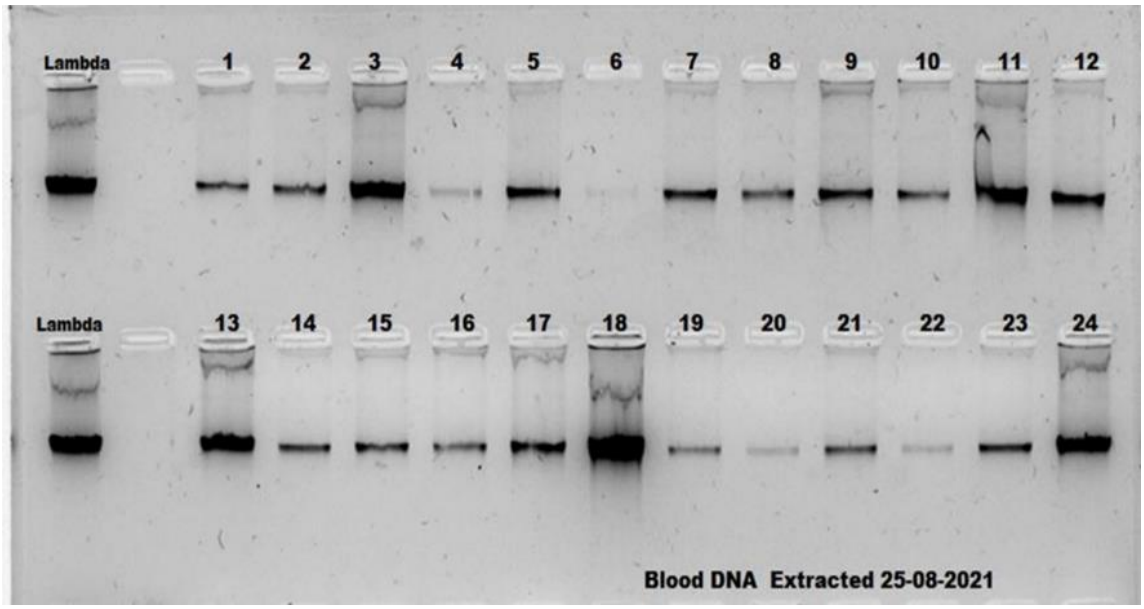


Figure 12: DNA quality check results from 24 random blood DNA samples using 1% agarose gel electrophoresis

(ii) Polymerase chain reaction procedures for *Brucella* genus and species detection

Primers and probes used for genus and species detection

- **Details for Primers and probes manufacturer and designs**

Different types of primers are available and have been used for the detection of *Brucella* species at genus and species level. For this study, *Brucella*-specific insertion sequence IS711 primer pair and probe were used. For *Brucella* species detection, IS711_downstream_of *_BMEI1162* primers and probe were used to detect *B. melitensis*, and IS711_downstream_of_*alkB* primers

and probe were used for *B. abortus* detection. All primers and their probes used were the produced by Macrogen Europe (https://www.macrogen-europe.com).

The primers and probes details for both genus (IS711) and species (AlkB and BMEI1162) detection are available in Appendix 2. For genus detection, the lyophilized forward and reversed primers were mixed with 360 µl and 320 µl of molecular grade water respectively to reconstitute to a concentration of 100 mol/µl as a stock. The probe was mixed with 440 µl of molecular grade water to reconstitute to a concentration of 100 pmol/µl as a stock. The reconstitutions were done according to a manufacturer instructions and were all stored at -20℃.

For species detection, the lyophilized forward and reversed the primer were mixed with 320 µl of molecular grade water and the lyophilized probe was mixed with 370 µl of molecular grade water. Primers and probe were reconstituted to a stock concentration of 100 pmol/µl according to the manufacturer instructions and stored at -20℃.

The *B. melitensis* primers and probe were designed to target the insertion of an IS711 element downstream of *BMEI*1162 and for *B. abortus*, the primers and probe were designed to target the specific insertion of an IS711 element downstream of the alpha-ketoglutarate-dependent dioxygenase (*alkB*) gene. The *B. melitensis* and *B. abortus* share the same IS711 reverse primer and the only difference is in their forward primers which target *BMEI1162* and *alkB* respectively. The TaqMan probes for *B. melitensis* and *B. abortus* target the *BMEI1162* and *alkB* genes, respectively (Probert *et al.*, 2004).

- **Dilution of primers and probes to a working concentration**

All primers and probes were at a stock concentration of 100 pmol/µl which needed to be diluted to a working concentration of 10 pmol/µl.

To calculate the working concertation of 10 pm/µl (primers and probe), the following popular formula for dilution was used to deduce the final concentration of the working concentration.

$$C1V1 = C2V2 \qquad \text{Equation…….3}$$

Where:

C1= The concentration of the starting solution, V1= The Volume of the starting solution, C2= Final concentration of the solution, and V2 = Volume of the diluent/ molecular grade water required to dilute the reagents to a desired concentration (10 pmol/µl).

For this case:

C1=100 pmol/µl, V1=45 µl (15 µl for each primer and probe from the stock)

C2= 10 pmol/µl, V2= not known.

V2= 100 pmol/µl x 45µl /10 pmol/µl = 450 µl final volume of the mixture of primers and probe.

To get the Volume of molecular grade water = 450 – 45 µl of primers and probe = 405 µl of water.

Therefore, 405 µl of water was mixed with 15 µl of forward primer +15 µl of reverse primer + 15 µl of the probe.

Therefore, the working concentration of primer and probe was 10 pmol/µl.

- **Primer pairs in silico specificity and sensitivity analysis**

For primer pair specificity test, primer-BLAST tool in a National Center for Biotechnology Information (NCBI) website (https:// www. ncbi. nlm. nih. gov/ tools/ primer-blast/ index. cgi? GROUP_ TARGET=on) was used, the respective forward and reverse primers were used in the primer parameters window as a query sequence and blasted into a nucleotide database to get the returned sequences/amplicons. Stringent parameters such as primer must have at least 1 total mismatch to unintended targets, including at least 1 mismatch within the last 1 base pair at the 3' end were set for primer specificity analysis. The system was set to ignore targets that have 2 or more mismatches to the primer. The maximum target amplicon size was set at 1000 bps. The returned amplicons were recorded and checked if there were any other bacteria species picked/amplified other than *Brucella*. For genus detection, the IS711 primer pairs were expected to amplify bacteria of the genus *Brucella* only as well as for the species primer pairs (BMEI1162 and alkB) to have a perfect specificity.

For primer pair sensitivity test, the aim was to check how many intended targets in GenBank are picked up and or missed by the primer. Therefore, foremost was to identify all intended targets present in the GenBank by using the first returned sequence (amplicon) during the

primer blast to conduct a nucleotide blast (BLASTn) in the GenBank. The accession number of the first returned sequence (amplicon) was used as a query sequence and the amplified region by the reverse primer was used as a query subrange in a BLASTn window. For search set database, the standard databases nucleotide collection was selected, and program selection was optimized for a somewhat similar sequence (BLASTn). For algorithm parameters section, the general paramers setting such as the maximum target sequences was set at 1000, Expect-threshold at 0.05, word size at 16. For Scoring parameters setting such as match/mismatch was set at 2,-3, Scores Gap Cost at existence: 5 and Extension: 2. Filters and masking setting such as Filter was set at low complexity region and Mask was set at lookup table only. After blasting, the results (amplicons) were again filtered to get the amplification with 95% to 100% query coverage.

The BLASTn results showed that one thousand amplicons (those with significant alignment) were returned. To assess the sensitivity of the primer, the blastn amplifications were identified and compared to primer blast amplifications by using their accession numbers. For the perfect sensitivity of the primer pair, it was expected that all primer blast amplicons would be returned during the blastn together with other somewhat similar sequence, similarly happened during the blastn. The new other amplicons were returned due to the parameter set during program selection as was optimized for a somewhat similar sequence (blastn), However, GenBank is not error-free, some sequences can be subjected to erroneous species annotation, such as different names given to the same species (sequence), or sequences assigned to an incorrect species, this can only be visible and identified by building a dendrogram.

For a dendogram building, the first 50 blastn amplicons were further downloaded into the fasta file, and imported into the MEGA7 software (Molecular Evolutionary Genetics Analysis-MEGA) for alignment and editing by deleting the excessive length of extremely long sequences for dendrogram generation to check for the taxa topological relationship of the amplicons. After alignment, the fasta file was transformed into the mega file format and closed the alignment explorer of the MEGA7 software. The mega file format was imported into the software and tree building was done using a phylogeny button. The tree was made using the neighbor-joining method for phylogenetic reconstruction. In addition, 150 number of bootstrap was set, other parameters set were, substitution type was set at nucleotide, mode/method was set at p-distance, substitution to include was set at transitions + transversions, rates among sites was set at uniform rates, pattern among lineages was set at same (homogeneous) and gaps/missing data

treatment was set at pairwise deletion and the software was allowed to compute. The dendrogram (cladogram) generated by bootstrapping with the node frequencies present for each primer pair is available in the results Chapter.

Preparation of the master mix and mixture of primers and probe before use for genus detection

The Lunar Universal Probe qPCR Master Mix (New England BioLabs, MA, USA), stored at -20°C was used in the assays for genus detection, before use the master mix was taken from the deep freezer and allowed to thaw to room temperature and followed by brief mixing by vortexing before use. The master mix formulation was supplied at 2X concentration and contains all PCR components required for amplification and quantitation of DNA except primers/probes and DNA template. Primers (forward and reverse) and probe were provided at stock concentration (100 pmol/µl) and were diluted to working concentration of 10 pmol/µl. The dilution was 15 µl of each primers and probe were taken and mixed into 405 µl of molecular grade water in eppendorf tube and thoroughly mixed together by brief vortexing.

- **Assay mix (cocktail) preparation for uniplex real-time PCR for Brucella genus detection**

For this study, the assay mix was prepared according to the protocol developed by Akoko *et al.* (2021) for reaction mix assay and real-time PCR conditions. However, because relatively smaller reaction volumes were used, components were scaled proportionally. The following volumes of each reagent were used per sample:

Master Mix	5 µl
Primer and Probe	1.25 µl
Molecular grade water	1.75 µl
DNA template	3 µl
Total volume	11 µl per sample.

The reaction components were added in a PCR plate according to the list above and thoroughly mixed in a centrifuge for 5 minutes at 2500 rpm before setting the plate into the qPCR machine.

- **Assay optimization, ten-fold serial dilution of reference DNA, and standard curve plot for Brucella genus detection**

To ensure good assay performance and identify the limit of detection (LOD) of the assay which is the lowest concentration at which 95% of the positive samples are detected. The positive control DNAs (DNA extracts from *Brucella* strains: *B. melitensis* 16 M and *B. abortus* 544) used in this study were both sourced from the Friedrich-Loeffler-Institute *Brucella* Reference Laboratory in Germany) (Akoko *et al.,* 2021). *Brucella abortus* references DNA at a concentration of 1 µg/µl was used. Ten-fold serial dilutions of the reference DNA was done manually by transferring 10 µl of *Brucella* reference DNA then mix into 90 µl of molecular grade water. The serial dilution started at the concentration of 10^{-1} to 10^{-7} µg/µl (Fig. 13).

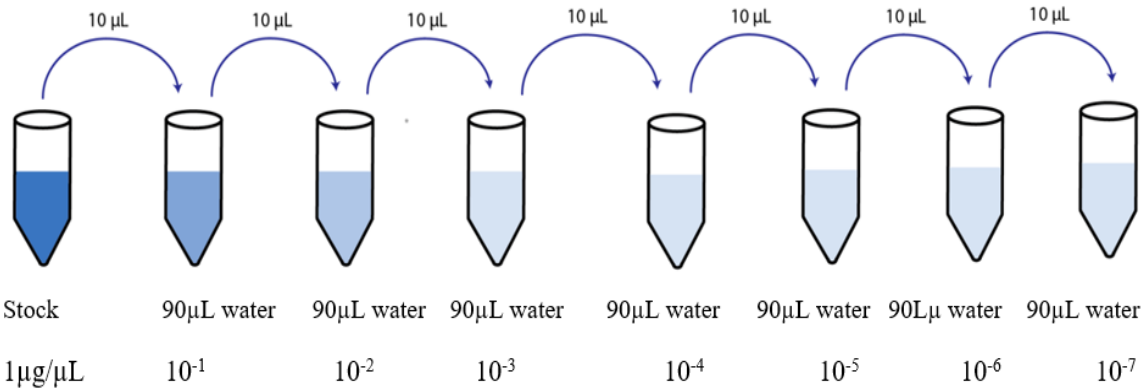


Figure 13: Illustration of serial dilution of the Brucella reference DNA at a stock concentration of 1µg/µl

- **The PCR steps and conditions setting (Temperature and Time) for Brucella genus detection using quantitative real-time PCR (qPCR)**

The fluorescence-based quantitative real-time PCR for 9 well strips with 0.2 ml capacity block was used. A QuantiStudio™ (5) (Applied Biosystems, Singapore) installed with QuantiStudio Design and Analysis software v1.5.1 was used. Fluorescence-based quantitative real-time PCRs are extremely sensitive and excellent for molecular diagnostics in life sciences (Bustin *et al.,* 2009; Kubista *et al.,* 2006).

The PCR conditions used were adopted as explained by Akoko *et al.,* 2018. All PCR stages were set, pretreatment stage (UDG step) was set at 50°C for 2 min to cleave all contaminating templates containing Uracil N glycosylase (UNG)/Uracil DNA glycosylase (UDG) bases to prevent carryover contamination, DNA denaturation at 95°C for 10 min, amplification at 95°C for 15 sec, annealing at 60°C or 1min and lastly for this reaction, 42 number of cycles was set

(Fig. 14). For genus detection, the fluorophore used was FAM and the quencher was nonfluorescent quencher NFQ-MGB (Akoko *et al.*, 2021). For more details of primer pairs and probes check on Appendix 2 on the Appendix page.

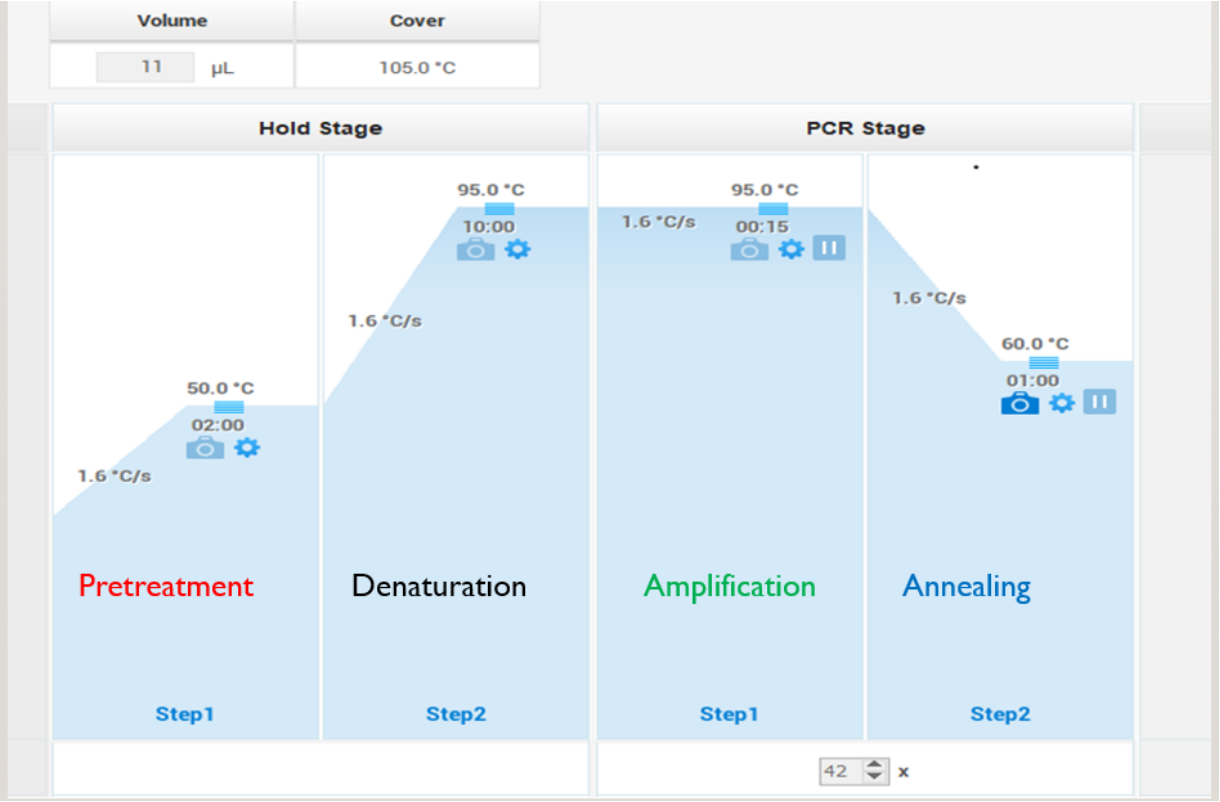


Figure 14: The PCR condition settings used for running the DNA samples in the quantitative real-time PCR machine

- **The Assay's limit of detection and standard curve for Brucella genus detection**

(a) To determine the limit of detection (LOD) to detect *Brucella* genus

The limit of detection for an assay is the lowest concentration of analyte in a sample that can be consistently detect with 95% probability. For genus-specific detection, the serial dilutions of the reference DNA (*B. abortus* 544) were run under the set conditions of the real-time PCR machine, and an analysis of the results was done. The reference DNA was run in duplicate and therefore two sigmoid curves can be seen (Fig. 15) from the same dilution except for the last sigmoid curve. The serial dilutions used were from 10^{-1}µg/µl to 10^{-6}µg/µl, therefore six duplicate sigmoid curves were expected.

The amplification plot shows that, from the left, the first sigmoid curves at cycles number 20 represent the amplification curve of the reference DNA at a dilution of 10^{-1}µg/µl, the second at a dilution of 10^{-2}µg/µl, and the last at 10^{-6}µg/µl (Fig. 15). There is only one sigmoid curve

instead of two in the last amplification curve at a dilution of 10^{-6}µg/µl, only one curve turned positive and the other turned negative(undetected) and therefore signifying the end of the detection limit for the assay, this means that, this dilution was the limit of detection of the amount of DNA in a sample for this assay and any amount below this dilution will not be detected, therefore the last but one (10^{-5}µg/µl) was used as the limit of detection for the assay for the *Brucella* genus detection.

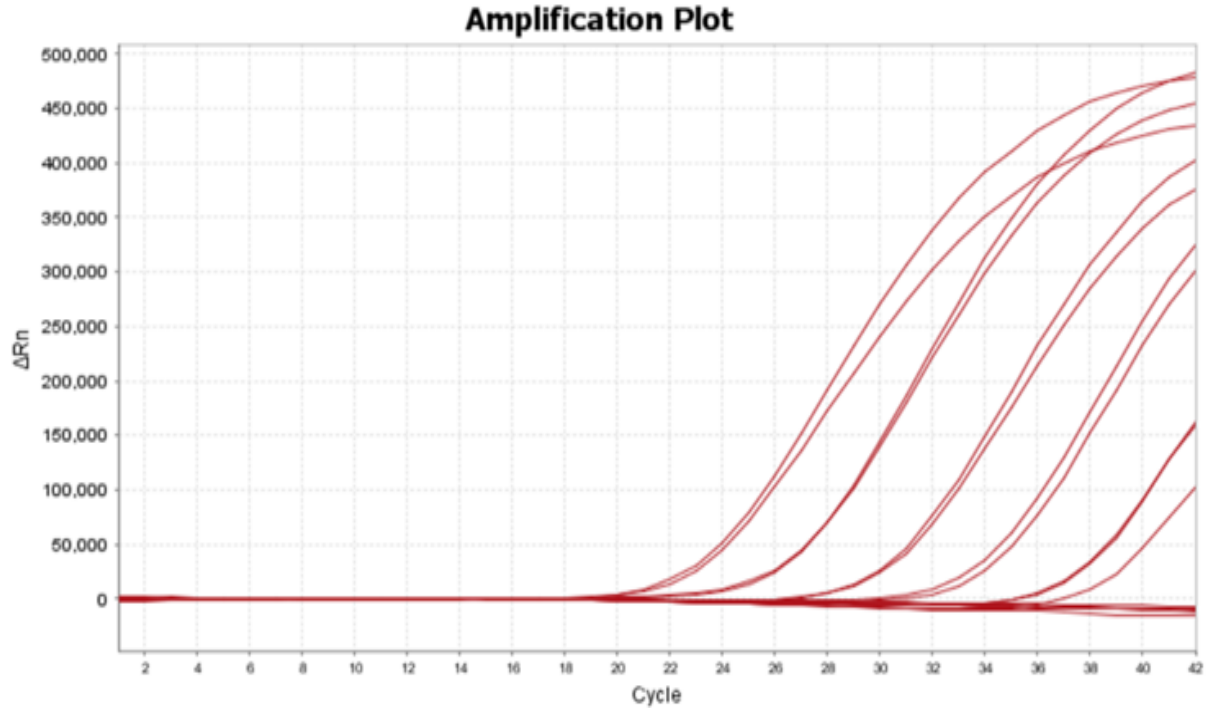


Figure 15: Amplification plot of the serial dilutions of the reference DNA for optimization of the assay for Brucella genus detection using *B. abortus* reference material and IS711 primers and probe

(b) Standard curve for Brucella genus detection

The standard curve has been drawn by using the standard/reference sample serial dilution as explained in Section 3.6.8 above to test the efficiency of assay and primers. Getting a good standard curve needs at least five data points over several orders of magnitude of dilution from the same sample. In this case, a reference sample was diluted at six orders of dilutions (10^{-1} to 10^{-6}) and therefore, the standard curve was drawn by using the threshold cycles on the Y-axis and DNA quantity on the X-axis. There was no contamination observed in the negative controls.

The values calculated to explain the efficiency of the assay include the assay efficiency of 90.5% which falls within the acceptable range of 90 -110%. To assess the good linear

relationship of data points of the standard sample, the R^2 value of 0.988 was found and the perfect correlation has the value of R^2 =1(Fig. 16). Signifying a good coefficient of correlation of data points which provides good confidence within the correlation and the slope of -3.56 was found whereas perfect qPCR assay has a slope of -3.32 with the acceptable slope range of -3.1 to -3.58 (Bilodeau, 2022; Vandesompel, 2008).

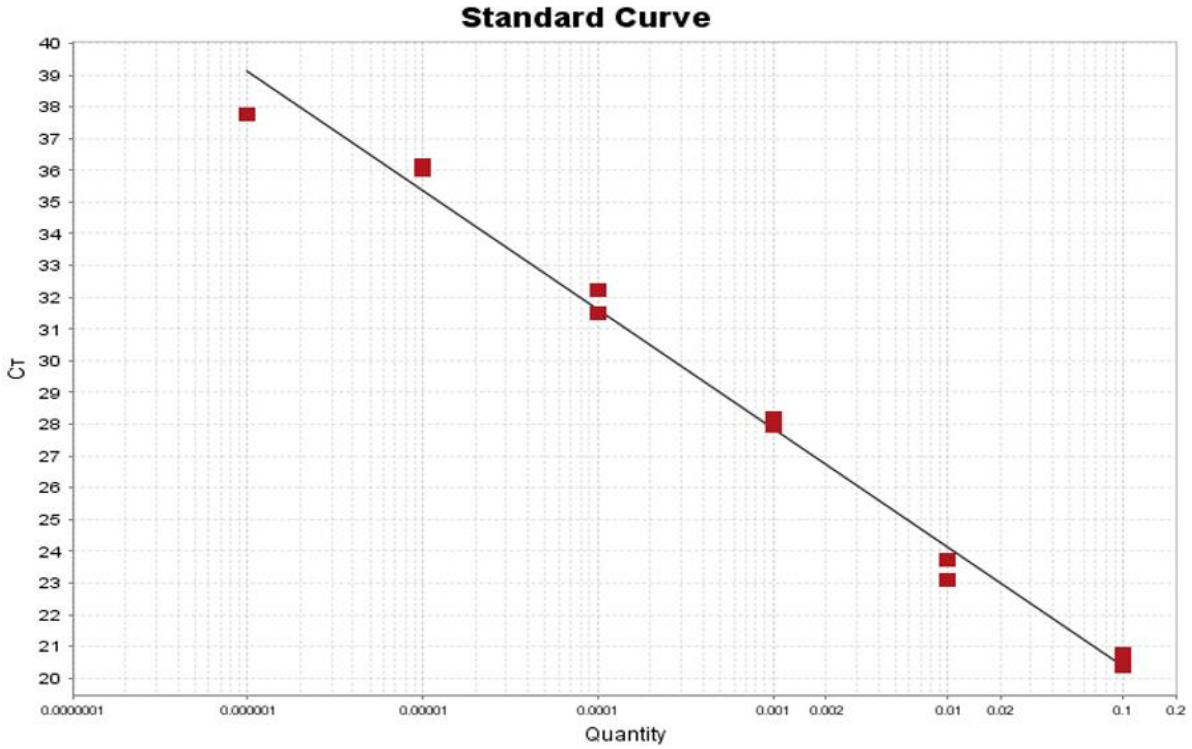


Figure 16: The IS711 standard curve for detection of Brucella genus. Efficiency (Eff: 90.55%, R2: 0.988, slope: -3.56, and y-intercept: 16.63)

The results suggest that the assay was good for *Brucella* genus detection as the parameters for assay efficiency were good and therefore the method could be used for genus detection with confidence to produce reliable results.

The assays for Brucella species detection

- **Preparation of master mix and mixture of primers and probes for *Brucella* species detection from DNA positive samples**

The master mix PerfeCta qPCR ToughMix UNG Low ROX (Quantabio ® USA) was used for the species detection for all genus-positive DNA samples. The product was a 2X concentrated ready-to-use reaction cocktail for PCR amplification of DNA templates that overcomes many known inhibitors of PCR often present in crude samples. It was a versatile and robust real-time qPCR reagent that provided maximum sensitivity and PCR efficiency.

The PerfeCta qPCR ToughMix UNG Low ROX containing an optimized concentration of $MgCL_2$, dNTPs (dATP, dCTP, dGTP, dUTP), hot-start DNA polymerase, Uracil DNA glycosylase (UDG/UNG) and stabilizers. The product was stored in a constant temperature freezer at -20℃, repeated freezing and thawing do not impair product performance. Before use, the master mix was taken out of deep freezer and allowed to thaw to room temperature and later followed by gentle mixing by vortexing. For *Brucella* species detection, primers (alkB and BMEI1162) and probe were provided at stock concentration (100 pmol/µl) and were diluted to working concentration (10 pmol/µl). The dilution was 15 µl of each primers and probe were taken and mixed into 405 µl of molecular grade water in eppendorf tube and thoroughly mixed together by brief vortexing.

- **The PCR steps and conditions (temperature and time) settings for Brucella species detection**

For *Brucella* species detection, all DNA samples detected positive at the genus level were tested for the presence of *B. abortus* or *B. melitensis* by uniplex qPCR reactions. The *alkB* and *BMEI1162* primers for detecting *B. abortus* and *B. melitensis* respectively were used.

All PCR stages were set, the pretreatment stage (UNG step) was set at 45°C for 5 min to cleave all contaminating templates containing U bases, DNA denaturation at 95°C for 5 min, amplification at 95°C for 15 seconds, annealing at 60°C for 30 seconds and 42 number of cycles was set (Fig. 17).

The final reaction volume per well was 15 µl, comprising 7.5 µl Master mix (Perfecta qPCR ToughMix UNG Low ROX), 0.75 µl Primer and probe (species-specific), 1.75 µl Molecular grade water, and 5 µl of DNA template. The final reaction mix was thoroughly mixed in a centrifuge for 5 minutes at 2500 rpm. The fluorophore used was JOE for *B. abortus* and Texas Red for *B. melitensis* and the quencher used was nonfluorescent (NFQ-MGB) (Akoko *et al.*, 2021).

Figure 17: The PCR steps and conditions (Temperature and Time) for the *B. abortus* and *B. melitensis* detection

(a) To determine the limit of detection (LOD) for an assay to detect *B. abortus*

To determine the limit of detection of an assay to detect *B. abortus*, the serial dilutions of the reference DNA (*B. abortus* 544) were run into the set conditions of real-time PCR machine, and an analysis of the results was done. The reference DNA was run in duplicate and therefore two sigmoid curves expected from each dilution. The dilutions used were from 10^{-1}µg/µl to 10^{-6}µg/µl.

The serial dilution amplification plot shows that, from the left, the first sigmoid curves at 21 cycles represent the amplification curve of the reference DNA at dilution of 10^{-1}µg/µl, the second at dilution of 10^{-2}µg/µl and the last at dilution of 10^{-6}µg/µl (Fig. 18). There was no amplification curve for the last dilution (10^{-6}µg/µl) as the assay could not detect. This signifies the end of the detection limit for the assay, this means that this dilution was the limit of detection of the amount of DNA in a sample for this assay; therefore, the last but one dilution (10^{-5}µg/µl) was used as the limit of detection for the assay for the *B. abortus* detection.

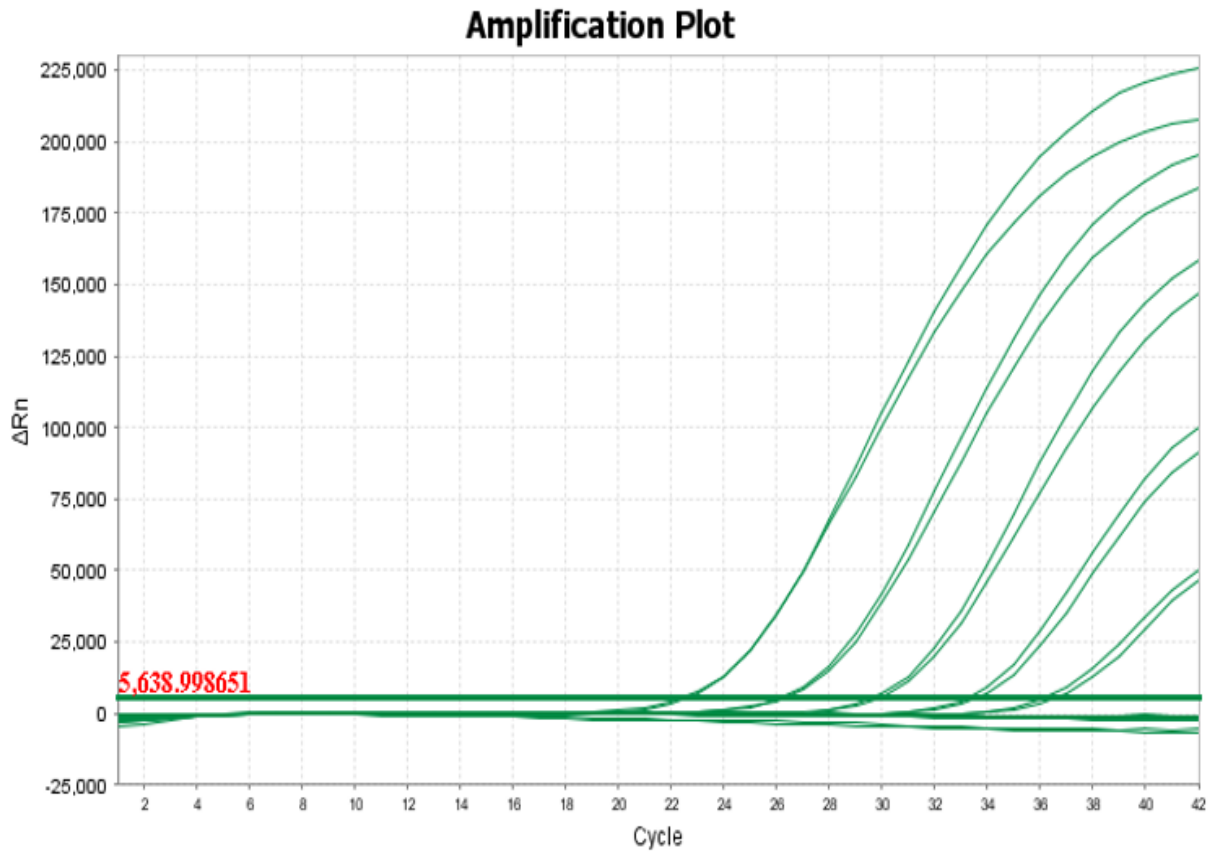


Figure 18: Amplification plot of the serial dilution of the *B. abortus* reference DNA for optimization of the assay for *B. abortus* detection using alkB primer pair

(b) The standard curve for *B. abortus* detection

The standard curve has been drawn by using the *B. abortus* reference DNA after serial dilution during testing the efficiency of primers. The measures for assay efficiency can be seen in Fig. 19.

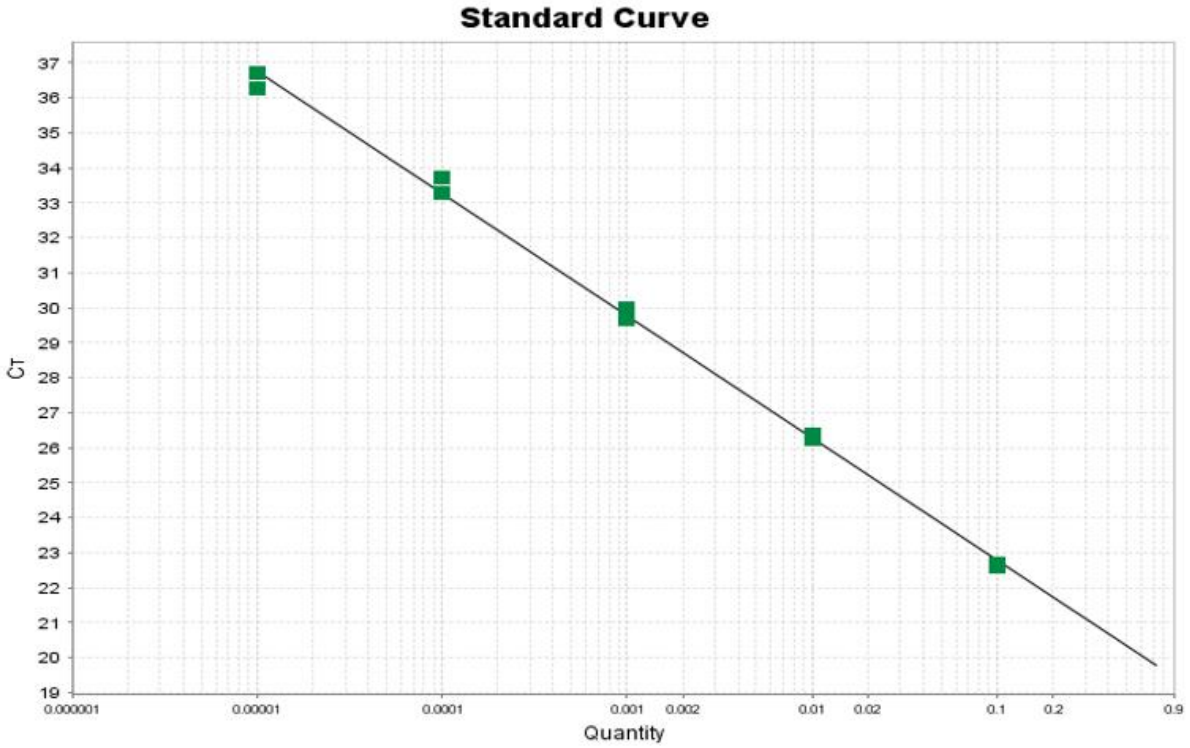


Figure 19: The alkB target standard curve for detection of *B. abortus*. Efficiency (Eff: 93.496%, R2: 0.998, slope: -3.488, and y-intercept: 19.288)

(c) To determine the limit of detection (LOD) for an assay to detect *B. melitensis*

To determine the limit of detection of an assay for *B. melitensis*, the serial dilutions of the *B. melitensis* 16 M reference DNA at 1µg/µl were run into the set conditions of real-time PCR machine, and analysis of the results was done. The reference DNA was run in duplicate and therefore two sigmoid curves expected from each dilution. The serial dilutions used were from 10^{-1}µg/µl to 10^{-6}µg/µl. The reference DNA dilutions were run in duplicate and therefore two sigmoid curves were expected for each pair of dilutions.

The serial dilution amplification plot below shows that, from the left, the first sigmoid curves at 21 cycles represent the amplification curve of *B. melitensis* reference DNA at dilution 10^{-1}µg/µl, the second at dilution of 10^{-2}µg/µl and the last at dilution of 10^{-6}µg/µl (Fig. 20). There was no amplification curve for the last dilution (10^{-6}µg/µl) as the assay could not detect, this signifies the end of the detection limit of the amount of DNA in a sample for the assay, this means that this dilution was the limit of detection for this assay, therefore the last but one (10^{-5}µg/µl) was used as the limit of detection for the assay for the *B. melitensis* detection.

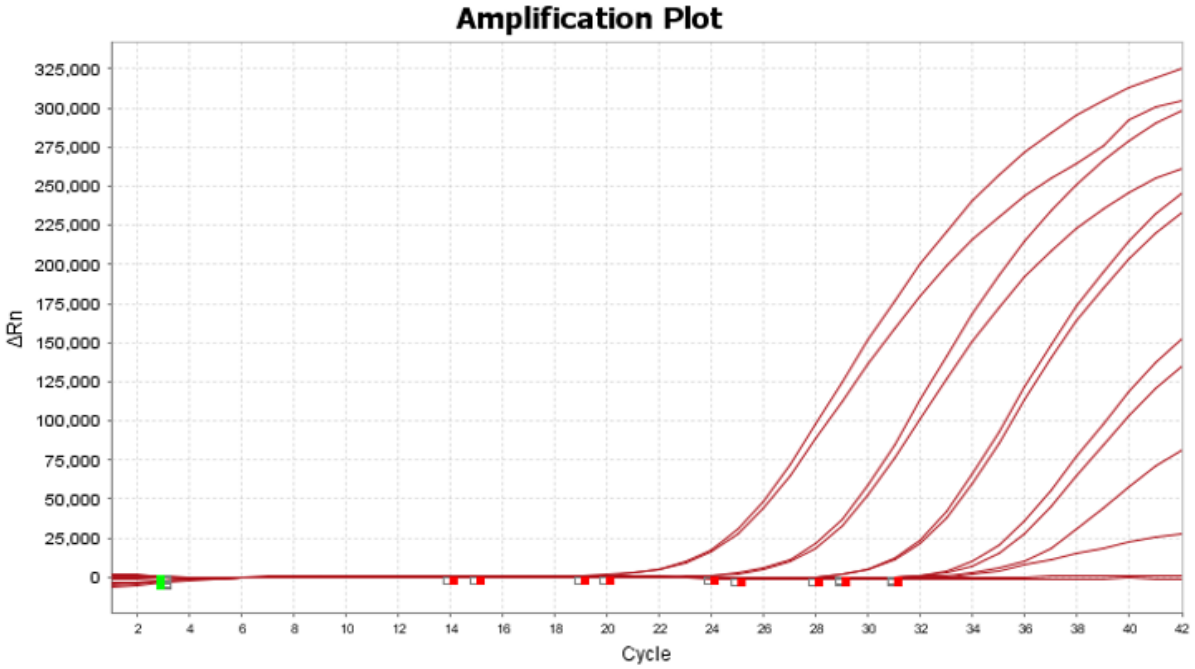


Figure 20: Amplification plot of serial dilutions of *B. melitensis* reference DNA to detect the limit of detection of the assay using BMEI1162 primers

(d) The standard curve for *B. melitensis* detection

The standard curve has been drawn by using the *B. melitensis* reference DNA after serial dilution during testing the efficiency of primers and the efficiency of the assay is seen in Fig. 21.

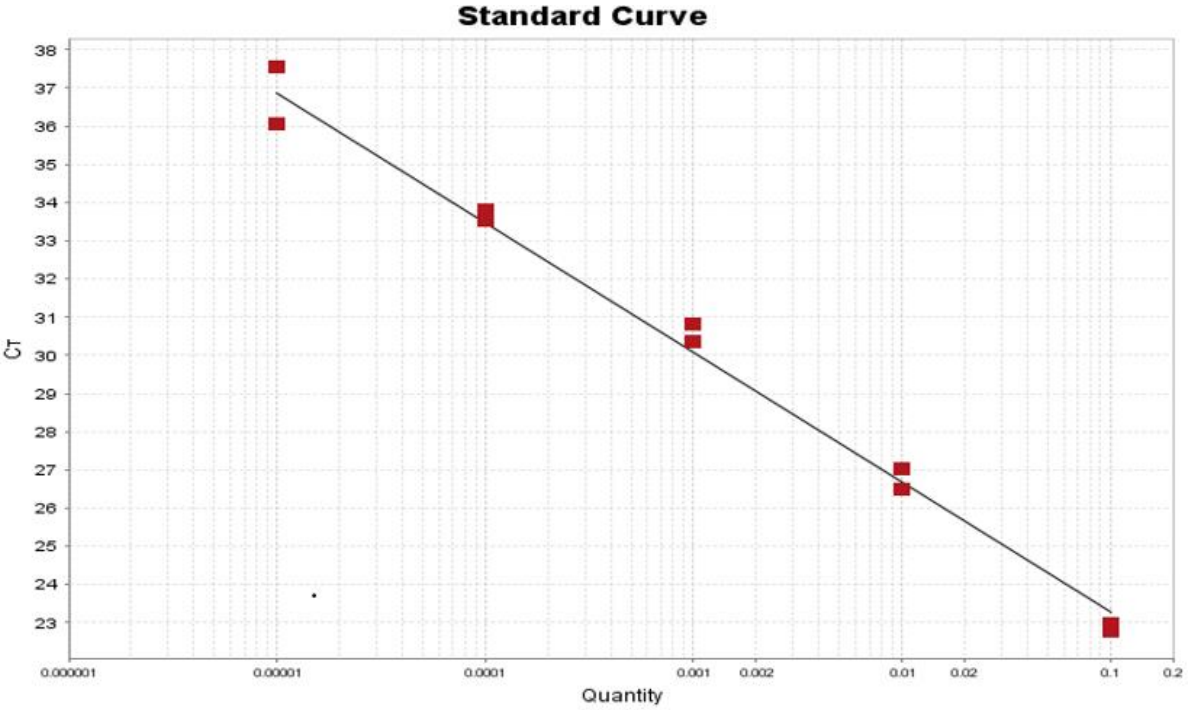


Figure 21: The BMEI1162 target standard curve for detection of *B. melitensis*. Efficiency (Eff: 97.149%, R2: 0.984, slope: -3.392, and y-intercept: 19.885)

3.4.3 Identification of SNP markers in dairy cattle associated with genetic resistance/susceptibility against brucellosis

(i) Bovine Illumina 50K SNP chip Genotyping Technology for identification of Single Nucleotide Polymorphism markers from cattle blood DNA

For SNP genotyping of the dairy cattle in the smallholder farming system, the Bovine 50K Illumina™ SNP chip (51 386 polymorphic SNP markers) was used. Manufacturer designed the SNP chips by using a combination of publicly available Bovine SNPs along with highly informative novel Bovine SNPs discovered using a reduced representation and next-generation sequencing technology strategies. The Bovine Illumina 50K SNP chip was used to identify SNP markers in the animal genome which were then associated with the phenotypes of interest (Mrode *et al.*, 2021; Petersen *et al.*, 2013; Raymond *et al.*, 2018), for this study phenotype of interest was a PCR positivity. The microtubes (flow-cells) are embedded with oligonucleotide probes containing predetermined SNPs, by the method of differential hybridization detect the respective SNPs from the DNA sample and send corresponding signals to the system for recording.

In this study, 1997/2048 dairy cattle had an overall base call rate of 95% on genome wide-genotyping using Bovine 50K Illumina™ SNP chip and therefore passed the quality control check for inclusion in the analysis (Aliloo *et al.*, 2018; Mrode *et al.*, 2021). To increase the statistical power for GWAS analysis, imputation of genotypes was done to increase the number of SNPs from 50K to 700K. The genotyping work was subcontracted to Neogenomics laboratories in the United States of America. The genotyping file results in American Standard Codes for Information Interchange (ASCII) format PED/MAP were provided from the genotyping laboratory for downstream analysis.

(ii) Heritability estimation for brucellosis PCR positivity by using Residual Maximum likelihood software (ASREML)

Heritability estimation was done by using command line statistical software called Residual Maximum Likelihood (ASREML) version 4 software (Gilmour *et al.*, 2015). Heritability for bovine brucellosis was estimated based on individual animal PCR test results (PCR status) as the trait of interest, SNPs data, and other genetically oriented factors such as breed, sex, the age of individual animals and environmental factors such as feeding management, herd size,

contact with other livestock and others were considered. A total of 1997 individual dairy cattle data were available for heritability estimation. Out of the 1997 individual cattle with genotypes, only 36 were PCR-positive (1.8%). Dairy cattle with PCR-positive results were defined as cases (infected/diseased) and those with negative results as controls (healthy). For analysis purposes (ASREML requirements), PCR results were considered as the response variable, negative individuals (controls/healthy) and positive individuals (cases/infected) were coded with values of 0 and 1, respectively. The coding was also done to all other explanatory variables involved in the analysis by using R studio software downloaded form such website https://www.r-studio.com.

The ASREML uses a inverse general relationship matrix (GIV) as one of the input files for the estimation of the variance components. The PLINK software version 1.9 (http://pngu.mgh.harvard.edu/purcell/plink/) was used to convert the pedigree file (*.ped) to generate a general relationship matrix (GRM) and ASREML converted the GRM using *nrm()* variance model function to a general inverse relationship matrix (GIV) input file. The ASREML needs 3 input files, one is an inverse relationship matrix file (*.giv), a data file (*.csv), and a job file (*.as) containing all the commands for the variance components analysis and parameter predictions. Figure 22 is an example of a job file that has 6 main parts: Title, data field definitions from the data file (*.csv), the inverse relationship matrix file (*.giv), the data file name (*.csv), linear mixed model definition, and prediction functions.

*cattle_pheno - Notepad
File Edit Format View Help

```
!NS !WORKSPACE 20 !RENAME !ARGS 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Title: cattle_pheno. !DOPART $1
 id  !A !LL 19      # TZN000192809878
 givid  1977      # 1
 Bruce   1  # 0
 sex  *      # 2
 age  *       # 1
 breed  *       # 1
 herdsize  *       #
 management  *       # 1
 bull  *       # 2
 cont_dogs *   # 1
 cont_pigs  *       # 1
 cont_goats  *       # 1
 cont_sheep  *    # 2
 distance  *       # 1
 placenta *       # 2
 cattleyears *       # 2
 water  *       # 1
 PC1    1    # -0.0342149
 PC2    1    # -0.0287136

cattle_ibsinv_03.giv

# Check/Correct these field definitions.
adgg_bruce060223.csv  !SKIP 1 !MAXIT 10 !EXTRA 10 !MVINCLUDE  !FCON

!Part 1
!GDENSE

Bruce ~ mu sex age breed herdsize management bull cont_dogs cont:pigs cont_goats cont_sheep distance !r grm1(id)

!VPREDICT !DEFINE
F addit 1  #pheno var 2
H Heri 1 2 #additive heritability
```

Figure 22: The ASREML job file for variance components analysis and heritability prediction

The ASREML assumes a linear mixed model effect to estimate different parameters by first calculating variance components of fixed and random model terms and then by using the variance components results to calculate other functions such as a linear combination of variance components (F) and heritability (H).

The variance structures for fixed effect model terms such as sex, age, breed, environmental factors and principal components, and random model effects term (animal identification number) were estimated. All of them were specified in the linear mixed effect model by wrapping terms with the required variance model functions and qualifiers. At first univariable analysis was conducted where each variable was tested to explain the response variable with brucellosis PCR positivity first by fixed effect alone and later by fixed effect and random effect term. Multivariable analysis was also done by starting with the complex model first without random effect term and later with the random effect term to the simplest model in the same order. The ASREML calculates logL for each model with or without random effect, to compare the nested model logL and calculate the residual maximum likelihood ratio test (REMLRT). The REMLRT is twice the difference in the likelihoods for two nested models and if the result is greater than 3.8 then the model is considered significant at 5%. Non-nested models (multivariable models) were compared by using the Akaike information criterion (AIC), a model with the smallest AIC was considered the best fit to explain the brucellosis PCR positivity and for calculating variance components for heritability estimation.

The complex model was specified as follows in the job file:

Bruce ~ mu sex age breed herd_size, management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta cattleyears water!r grm1(id) 0.1 !GP.

! VPREDICT !DEFINE

F addit 1 + 2 #pheno var 3

H Heri 1 3 #additive heritability

The model terms in the mixed effect model above can be explained as follows:

The outcome variable Bruce (PCR positive/negative) is explained by the fixed effects variables animal sex (sex), animal age (age), animal breed (breed), herd size (herd_size), feeding management (management), keeping bull (bull), farm has dog around (cont_dogs), farm has

pigs around (cont_pigs), farm has goat around (cont_goats), farm has sheep around (cont_sheep), distance between dairy farm (distance), placenta management (placenta), how long a farmer has been keeping cattle (cattleyears) and source of drinking water for cattle (water) as the fixed effect terms under the constant/intercept mu and the random effect term which is individual animal identification number wrapped in a variance model function generalized relationship matrix ***(grm1())***. The ASREML plug-in values are derived from the observed phenotypic variance matrix to estimate the relationship of variance components, the qualifier ***!GP*** requests that the resulting estimated generalized relationship matrix be kept within the parameter space (positive definite), if not just use the smallest positive defined value 0.1 to avoid the model from stopping running. This analysis gives variance components for a fixed effect (var 1) and a random effect (var 2) which are the input results for heritability estimation.

Instruction to calculate the heritability function is headed by the line ***!VPREDICT !DEFINE***. Functions of the variance components were specified by lines in the form of letter label coefficients. *Letter* is basically an ASREML function for which different estimates of interests such as F and H can be calculated, *label* names the results of the function of variance components such as addit and Heri for this case and *coefficient* is the list of arguments/coefficients for the linear function such as 1, 2 and 3.

Letter **F** forms linear combinations of variance components 1 (var 1) and 2 (var 2) to form a third (var 3) estimate of the function of the variance component called additive phenotypic variance (addit). Letter **H** is for forming heritability, which is a ratio of two variance components (var 1 and var 2) for forming an estimate called Heritability (Heri).

(iii) Population structure of the study animals for genome wide association study (GWAS) analysis

Genome wide association study analysis requires a homogenous population with a small genetic distance between individuals which suggest that the individuals in the population came from a common ancestry. To understand the genetic divergence of the study population PLINK software (Purcell *et al.*, 2007) was used in a supercomputer Eddie belonging to the University of Edinburgh.

To estimate population structure of study animal, principal component analysis (PCA) a dimensionality reduction method was used to reduce the dimensions of large data set of PED/FAM files to estimate eigenvectors which gives measures of the genetic distance of study animals. The PLINK function *–pca* was used to calculate the PCA from the binary PLINK files which had the following components (variables/dimensions) in PED/FAM file: Family ID (FID), Individual ID (IID), Paternal ID (PID), Maternal ID (MID), Sex and Phenotype and in MAP/BIM file components are Chromosome number, SNP identifier, Genetic distance (in centi-Morgans) and physical base pair position and base pair. The line of code used was *./plink -bfile bruc.pcr.gemma.gwas -cow -pca -out bruc.pcr.eigenvev.*

The -bfile command the PLINK to use the binary plink file for analysis, *bruc.pcr.gemma.gwas*, a common prefix for binary plink files, -cow command plink to analyze data considering this is cattle data, -pca function command the plink to analyze for principal components -out command the plink to send the analysis output results into a file with prefix *bruc.pcr.eigenvev*.

The -pca produces three files with the extensions (suffix), *.eigenval, *.eigenvec, and *.nosex files. The *.eigenvec file contains twenty principal components analysis for each individual. It is this *.eigenvec file that was imported in R and used to draw a plot by using the default first two principal components PC1 and PC2 by using the *ggplot2* package.

(iv) Genome wide association study (GWAS) analysis to identify significant SNPs to brucellosis PCR positivity

Genome wide association study analysis was done by using GEMMA version 0.98.1 software, a command line implemented in a high-performing computer based at the University of Edinburgh called Eddie. Binary Plink files format containing genotypes and phenotypes was used (*.bed, *.fam, and *.bim) for GWAS analysis. The SNPs data were transformed into a plink data format by using the PLINK software. From the pedigree file (*.ped) the general relationship matrix (GRM) was generated and was first used during ASREML analysis for heritability estimation and was also used here for GWAS analysis. The command line for GWAS analysis was as follows:

gemma -bfile bruc.pcr.gemma.gwas -lm 4 -o bruc.pcr.gemma.gwas.p1 -k cattle_ibsinv_03.giv -km 2 -c bruc.pcr.gemma.covarp1.txt

The function -bfile commands the gemma to read the bed files during analysis (all had the same prefix bruc.pcr.gemma.gwas), -lm 4 function command gemma to which statistical test to use, 4 specifies the 3 statistical test (Wald test (1), Likelihood ration test (2) and Score test (3)). The *p*-values of these tests were used to determine which SNP marker had the smallest *p*-value and therefore most significant. Other functions in the model, -o tell the gemma to name the output file prefix (bruc.pcr.gemma.gwas.p1), -k function specifies the relatedness matrix file to be used in the analysis which is cattle_ibsinv_03.giv, -km 2 specify input kinship/relatedness file type (default 1; valid value 1 or 2) and -c specify input covariates file name (optional); an intercept term is needed in the covariates file.

Unless specified in the model (not specified in this analysis), GEMMA uses its default functionalities that, non-polymorphic SNPs, SNPs with missingness below 5%, SNPs with minor allele frequency below 1%, SNPs with r^2 correlation with any of the covariates above 0.9999 and SNPs with Hardy-Weinberg *p*-values below 0.001 will not be included in the analysis.

The GEMMA analysis produces two results files, one is an association (assoc) text file (*.txt), and the other is *.log file. The log file contains some computation parameters and time, the proportion of variance in phenotype explained (PVE) results, and standard error for the null linear mixed model. The log file results are important for computation only and will not be presented here. The association file contains the GWAS association analysis results, the results which were used in R software using the *qqman* package to plot Manhattan and QQ-plots for visualization of the markers in each chromosome significantly associated with PCR positivity (disease phenotype).

Visualization of SNP marker with significant association above the genome-wide threshold line in a Manhattan plot (smallest p-value/ highest height in Manhattan plot) was done by using the *qqman* package in R software and association data file. The GWAS significant SNP was further checked for its effect on cattle genome by using the web-based Ensembl Variant Effect Predictor (VEP) (Martin *et al.,* 2023; McLaren *et al.,* 2016) which is an open-source tool for the genomic annotation and filtering of genomic variants by identifying the gene affected by the variant (SNP), location of the variant on the gene (noncoding or coding region) and consequence of the variant on protein sequence as was explained by others (Hunt *et al.,* 2022; McLaren *et al.,* 2010).

3.5 Conceptual framework of the methodological events

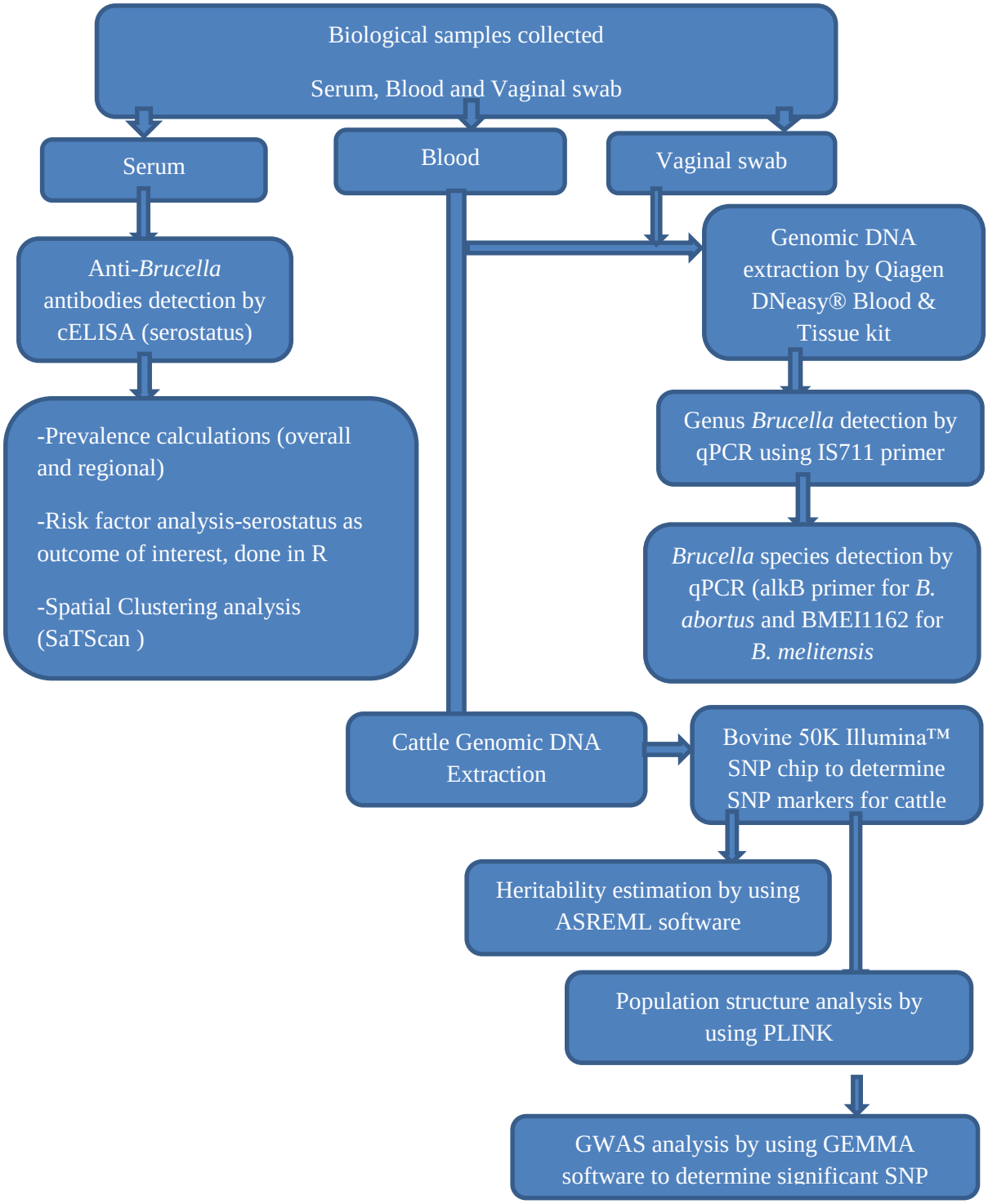


Figure 23: Conceptual framework of the methodological events used in this study

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Results

4.1.1 Visualization of study regions and number of study farms

The number of farms sampled was 1371 with median herd size of two cattle in two zones. The Northeastern zone comprised Arusha, Kilimanjaro and Tanga regions and Southern highland zone comprised Iringa, Njombe and Mbeya regions. The highest number of farms sampled were in the Kilimanjaro region (374), followed by the Tanga region (319), Arusha region (222), Mbeya region (165), Iringa (147), and least Njombe (144). The map shows the distribution of the number of farms by to local authority levels (Districts/Town council) as shown in Fig. 24. The legend color coding signifies the number of study farms at the district level/local authority in each region. The darker the color the higher the number of herds sampled.

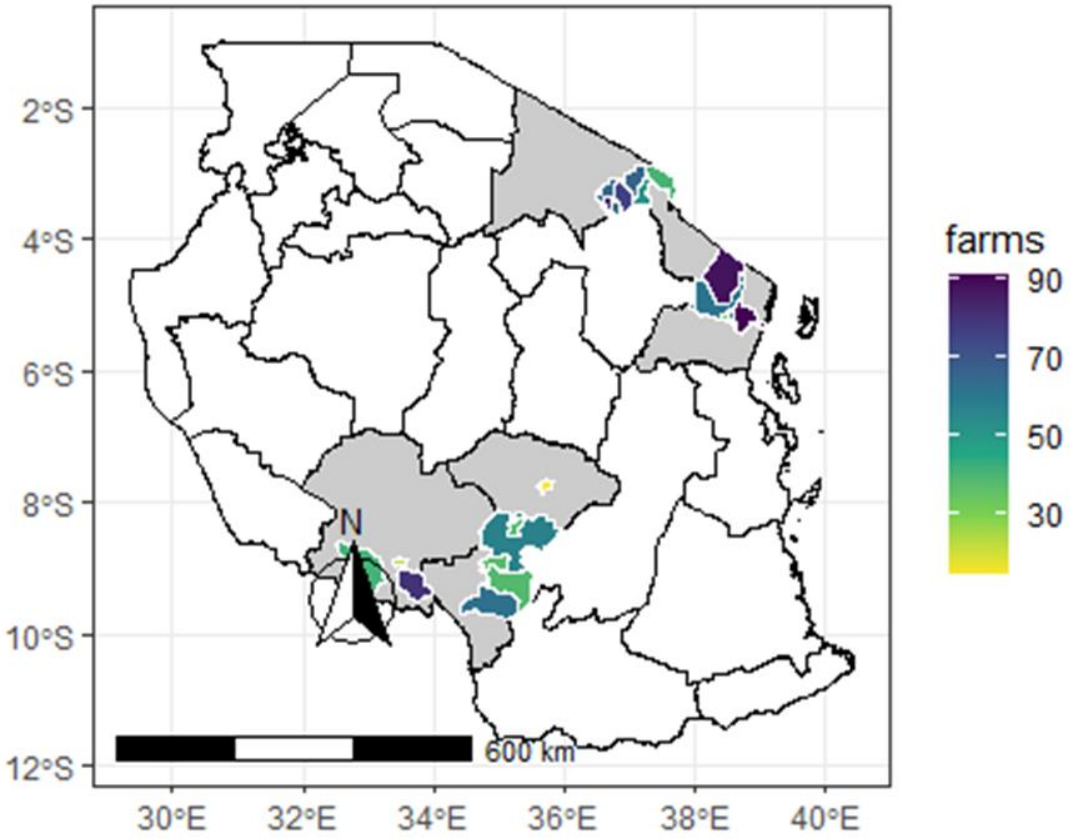


Figure 24: Tanzania map showing the number of study farms visited in the study regions

4.1.2 Seroprevalence and risk factors for Brucella spp on smallholder dairy farms in the six study regions in Tanzania

The cELISA was used to detect the circulating antibodies against Brucella species for all the samples. A total of 49 dairy cattle tested positive on the cELISA out of 2048. The individual unadjusted seroprevalence was 2.39% (95% CI: 1.7-3.1) and then adjusted (for unequal sample weights per region) individual seroprevalence was estimated to be 1.82% (95%CI 1.71–1.94). Out of the six regions, Njombe had the highest individual seroprevalence of 15.5% (95% CI: 10.64-21.51) whereas none of the dairy cattle was identified with anti-Brucella antibodies in Mbeya region (Table 3).

Table 3: Brucella seroprevalence at animal level in the study regions of Tanzania

	Animal level seroprevalence_cELISA						
Region	**Negat.**	**Posit.**	**Total**	**Prev %**	**95% CI**	**Pop**	**Weight**
Arusha	317	1	318	0.31	0.017-1.74	78 637	247
Tanga	519	5	524	0.96	0.31-2.22	41 639	395
Kilimanjaro	508	13	521	2.5	1.34-4.24	161 984	4043
Mbeya	217	0	217	0	0.00-1.60	72 724	0
Njombe	158	29	187	15.51	10.64-21.51	7177	1102
Iringa	280	1	281	0.36	0.019-1.90	7081	25
TOTAL	**1999**	**49**	**2048**	**2.39**	**1.7-3.1**	**369242**	

Key: Negat. = Negative, Posit. = Positive, Pop=population, Prev = Prevalence, %=percent, CI= Confidence Interval, Weight= Expected number positives scaled to regional population

4.1.3 Distribution of seropositive animals within study regions

The spatial choropleth map shows that the seropositive animals were clustered within a small number of local authorities in the Njombe and Kilimanjaro regions. There were a total of 29/187 seropositive animals in Njombe region from 8/136 farms sampled. Only three seropositive cattle came from farms that reported an abortion in the previous 12 months. In Kilimanjaro, there were a further 13/521 seropositive animals in the region representing 5/379 farms sampled (Fig. 25).

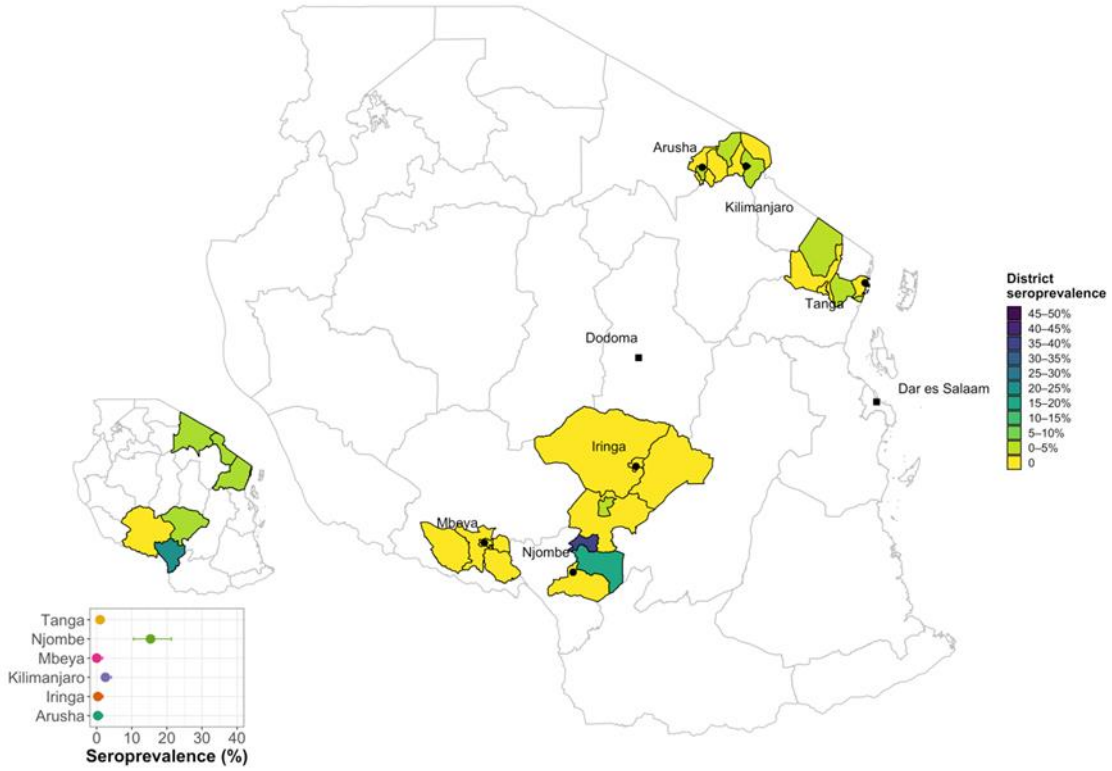


Figure 25: Choropleth map showing the regional seroprevalences (insets) and the detailed seroprevalence by local authorities sampled in each study region

4.1.4 Spatial clustering of the areas with seropositive (cases) and seronegative (controls) dairy cattle

To explore the spatial clustering of seropositive animals, Njombe and Kilimanjaro regions were mapped (Fig. 22) and a SatScan spatial analysis conducted (Kulldorff, 2009). Visually there appeared to be a cluster of positive animals (circled red), one (1) cluster in the northern part of the Njombe region and possibly two (2) clusters in the Kilimanjaro region. This was more formally tested using the spatial clustering test ($p < 0.05$) which identified a cluster of 84 animals within which were all 29 seropositive with a relative risk of 26.4 (95%CI: 3.7 – 190.6) and a radius of 21.14 km in Njombe (Fig. 26a) and a further statistically significant cluster comprising 49 animals with 7 positives (with a relative risk of 11.2 (95%CI: 3.9 – 32.1) and radius of 3.93 km in Kilimanjaro (Fig. 26b). In the Kilimanjaro region, the top left cluster was not statistically significant.

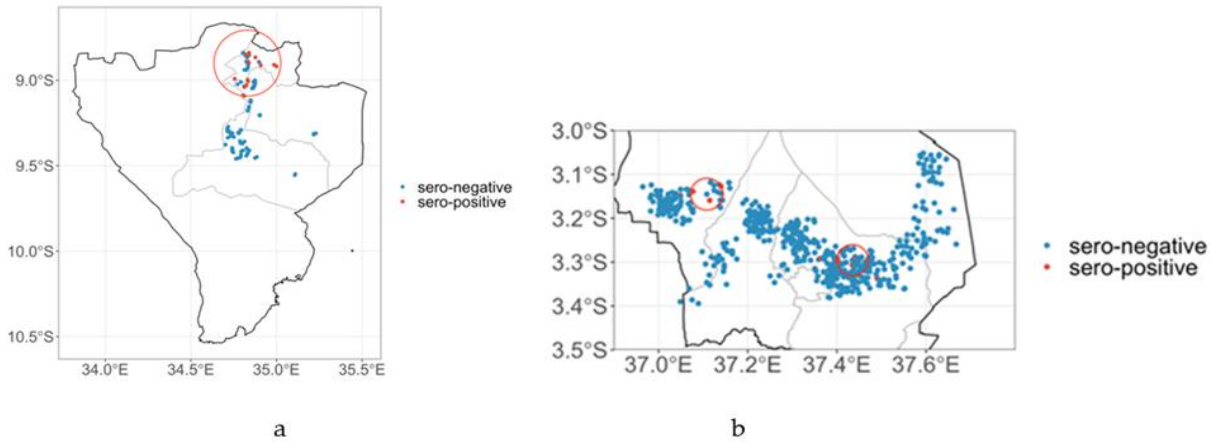


Figure 26: Map of Njombe (a) and Kilimanjaro (b) regions showing district boundaries, the location of seropositive and seronegative animals (jittered), and the radius (red circle) of the significant clusters identified by the SaTScan analysis

To try and assess if this was a single outbreak or more of an endemic expansion, the individual age to seroprevalences analysis was done and plotted (Fig. 27). The overall results shows that there was relatively high prevalence of nearly 0.1 in young individuals with 1 year of age and then dropped to 0.05 at age of 2 years and then dropped further to less than 0.05 throughout to the age of 8 years and rise again to 0.05 at age of 9 years, fall again to less than 0.05 at age of ten years and rise to nearly 0.05 at age of greater that eleven years. The overall individual age to seroprevalences do not suggest any strong increase with age. Overall, seroprevalence was very low across all ages with the possible exception of the over 8 years old animals. Given the apparently clustered pattern of seropositivity, Kilimanjaro and Njombe were separated and plotted on their own as shown in Fig. 27 (Inset: Age to brucellosis seroprevalence for the Kilimanjaro and Njombe regions only with a red dashed line for mean regional seroprevalences). In Kilimanjaro, seroprevalence was relative high just above 0.02 at the age of two and three years and in Njombe region the prevalence was just above 0.2 at the age of two years and 0.3 at age of nine years and droped down to 0.2 at age of ten years and rise again to nearly 0.4 at age of greater than 10 years. Again, the seroprevalence appears nearly uniform across ages in Njombe at below 0.2 and Kilimanjaro at just above 0.0 (2.5%) (red dashed line). Overall, there was no apparent steady increase in prevalence with age.

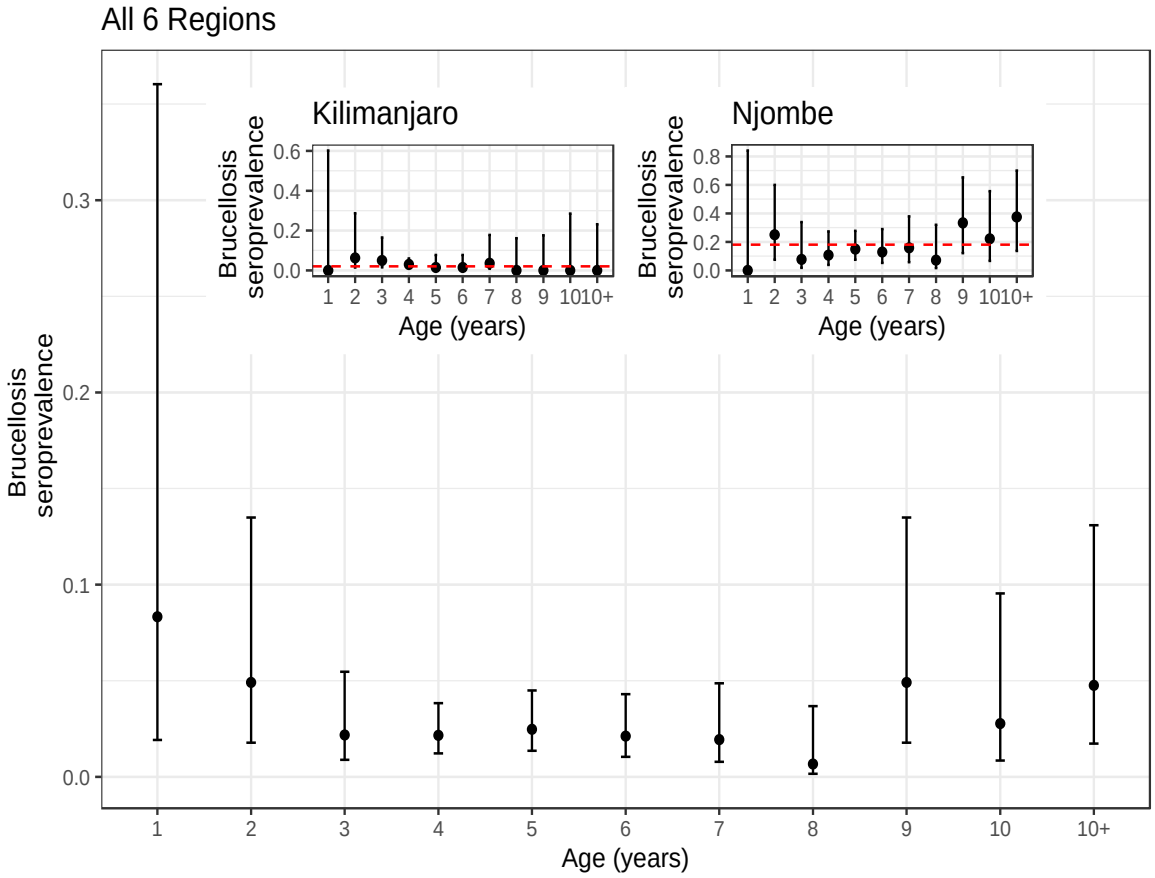


Figure 27: Age to brucellosis seroprevalence analysis across all six regions

4.1.5 Farmers' socio-economic predictors for brucellosis seropositivity

Of the respondent interviewed, 59.2% were males and 40.8% were females, and in terms of education, 74.1% had primary education, 11.7% had secondary education and 14.17% had tertiary education. Regarding livestock training, the study found that 70.6% of farmers had never had formal livestock training and 29.4 % had informal training. In addition, 8.0% of the farmers had experience of keeping dairy cattle for less than 5 years, and 91.98% had an experience of more than 5 years. It was further found that majority of farmers (60.5%) keep dairy cattle as a primary source of income. Regarding knowledge of milk-borne zoonosis, this study found that 65.4% of the farmers had no knowledge and 33.6% had some knowledge. In addition, 93.2% of the respondent did not know about dairy cattle abortifacient infectious diseases human can be infected with by handling infected aborted materials and fetuses.

The univariable analysis results showed that there was no significant difference between brucellosis seropositivity and knowledge of milk-borne zoonoses, knowledge of zoonosis

causing abortion, experience of keeping dairy cattle, gender of the farmer, livestock training status, reason for keeping dairy cattle (p>0.05) (Table 4).

Table 4: Univariable analysis of farmer-level variables and their association with animal-level Brucella seropositivity

		Number		Propor		95% Confid. Interv		P-
Variable	**Level**	**Negat.**	**Posit.**	**tion %**	**OR**	**Lower**	**Upper**	**value**
Farmers' gender	female	818	17	40.8	1			
	male	1179	32	59.2	1.3	0.72	2.32	0.46
Livestock training	no	1415	29	70.6	1			
	yes	582	20	29.4	1.67	0.94	2.99	0.08
Education	primary	1473	44	74.1	1			
	secondary	238	2	11.7	0.28	0.067	1.17	0.07
	tertiary	287	3	14.1	0.35	0.11	1.13	0.07
Knowledge of milk-borne zoonosis	no	1301	36	65.4	1			
	yes	675	13	33.6	0.71	0.38	1.34	0.45
Knowledge of abortion zoonosis	no	1839	48	93.2	1			
	yes	137	1	6.8	0.41	0.08	2.12	0.58
Experience	≤5	155	9	8.0	1			
	>5	1842	40	92.0	0.49	0.22	1.15	0.09
	home consumption	87	3	3.5	1			
Reason for keeping dairy cattle	primary income	1211	26	60.5	0.62	0.18	2.09	0.44
	Secondary income	699	20	35.1	0.82	0.24	2.83	0.73

OR= Odds ratio, Negat. =Negative, Posit.= Positive, +Ve= Positive, Confid. Interv= Confidence Interval, Proportion%= proportion of each level in a variable

4.1.6 Animal factors as predictors for brucellosis seropositivity

Few animal-level factors and management variables were studied. 97% of dairy cattle sampled were female and only 2.87% were males. The majority of the studied farms (73.66%) had no bulls while 26.34% had bulls for breeding purposes. Of the 585 bulls kept by dairy farms, 47.1% were used for hire to breed cows from other farms. Different crossbreds of dairy cattle were studied, of the 2048 sampled cattle, 3.86% was other breed comprised of local breed Tanzania short horn zebu (TSHZ), and their crosses with beef cattle breeds like boran and beef master. The dairy farms studied were predominated (81.43%) by a herd size of less than 5 dairy

cattle. Although 1.12% of the farmers reported vaccinating their dairy cattle against brucellosis, there was no evidence of vaccination and when vaccinators were consulted, they reported vaccinating against other diseases and not brucellosis. Furthermore, the majority (54.75%) of sampled cattle had good body condition, fed largely under intensive management (74.8%) and tap water was their main source of drinking water (63.8%).

Univariable analysis of the animal-level factors and management practices showed that there was no significant association between brucellosis seropositivity and cattle sex, herd size, bull hire, brucellosis vaccination, breed, and age of cattle ($p>0.05$). Results also showed that keeping your own bull was significantly associated with protection against brucellosis ($p<0.05$, OR<1) (Table 5).

Univariable analysis suggested that extensive feeding management was statistically associated with a reduced likelihood of being brucellosis seropositivity (OR=0.4, $p<0.05$). There was a statistical association between water source risk of being seropositive with animals getting well water at higher odds (OR=5.51) of being brucellosis seropositive compared to those drinking from rivers. The body condition score of greater than 3 was found to be statistically significantly associated with brucellosis seropositivity ($p=0.005$) and cattle with a body condition score greater than 3 were 0.43 times less likely to be Brucella seropositive (OR=0.43). Furthermore, the introduction of new animals in the herd for the past 12 months was found to be not significantly associated with brucellosis seropositivity ($p>0.05$) (Table 5).

Table 5: Univariable analysis of animal and management-level variables associated with Brucella seropositivity

Variable	Level	Number		Proportion %	OR	95%Confid. Interval		P-value
		Negative	Positive			Lower	Upper	
Cattle sex	female	1939	49	97.17	1			
	male	58	0	2.83	0.33	0.02	5.49	0.4
Bull	no	1459	48	73.66	1			
	yes	538	1	26.34	0.056	0.008	0.41	8.65E-06
Bull hire	no	285	1	52.9	1			
	yes	300	0	47.1	0.26	0.011	6.52	0.19
Vaccination	no	1974	49	98.88	1			
	yes	23	0	1.12	0.85	0.05	14.17	1
Breed	Other	77	2	3.86	1			
	SHZxAyrshire	416	6	20.63	0.48	0.11	2.1	0.61
	SHZxFriesian	1363	40	68.57	0.92	0.25	3.38	1
	SHZxJersey	141	1	6.94	0.33	0.04	2.55	0.62
Cattle age	<5	1138	29	57.03	1			
	≥5	859	20	42.97	0.91	0.51	1.61	0.77
Feeding management	intensive	1487	43	74.78	1			
	extensive	510	6	25.22	0.4	0.17	0.95	0.03
Herd size	<5	1621	45	81.43	1			
	≥5	376	4	18.57	0.38	0.14	1.07	0.06
Water source	river	334	3	16.47	1			
	tap	1279	27	63.83	2.35	0.71	7.79	0.18
	well	384	19	19.70	5.51	1.61	18.78	0.002
Body condition	≤3 score	894	32	45.25	1			
	>3 score	1103	17	54.75	0.43	0.23	0.77	0.005
New animals	no	1857	48	93.11	1			
	yes	140	1	6.89	0.28	0.04	2.02	0.25

OR= Odds ratio, Proportion%= proportion of each level in a variable

4.1.7 Management variables predictor for brucellosis seropositivity

Several management variables were studied revealing that 73.8% of the dairy farms were situated within a distance of less than 100 meters, while 26.2% of dairy farms were located at a distance exceeding 100 meters from one another. The majority of these dairy farms engaged in mixed livestock rearing, with over 50 % of them keeping goats, dogs, and pigs in addition to dairy cattle.

This study also enquired about the individuals responsible for milking dairy cows (milkers). The findings revealed that, 54.2% of the dairy farms utilized family members for milking, 6.7% employed outside help who regularly came to milk the cows, and 39.2% relied on a hired person living with the family (respondent). When it comes to the history of dairy farm abortions, 91.0% reported no history of abortions, while only 9.0% had a history of abortion on their farms. In terms of placenta management practices after a cow gave birth, 17.0% of the farmers exhibited poor practices, such as feeding the placenta to other animals, leaving it in a place, or disposing of it in the vicinity. In contrast, 83.0% followed good practices such as burying or burning the placenta.

The univariable analysis of farm factors and other management variables revealed an association between the distance between farms and brucellosis seropositivity ($p<0.05$). Dairy cattle kept on farms located less than 100 meters apart were found to be less likely to be *Brucella* seropositive (OR=0.33). Furthermore, the presence of dogs, goats or pigs on dairy farms was significantly associated with brucellosis seropositivity ($p<0.05$), and dairy cattle kept on these farms were more likely to be *Brucella* seropositive (OR>1) (Table 6).

Further univariable analysis revealed that dairy animals reared in Njombe region were 39 times more likely to contract brucellosis compared to those reared in other regions included in the current study, and this association was found to be significantly significant ($p<0.05$). Additionally, the history of abortion and placenta management were not significantly associated with brucellosis seropositivity ($p>0.05$) (Table 6).

Table 6: Univariable analysis results of farms and management level variables associated with Brucella seropositivity

Variable	Level	Number		Proportion %	OR	95% Confid. Interval		p-value
		Negative	Positive			Lower	Upper	
Distance	≤100	512	25	26.24	1			
	<100	1485	24	73.76	0.33	0.18	0.58	<0.001
Dogs	no dog	236	1	11.58	1			
	have dog	1763	48	88.42	6.42	0.88	46.76	0.037
Goats	no goats	671	8	33.15	1			
	have goats	1328	41	66.85	2.58	1.21	5.55	0.013
Sheep	no sheep	1514	31	75.47	1			
	have sheep	484	18	24.53	1.82	1	3.27	0.062
Pigs	no pigs	947	11	46.82	1			
	have pigs	1050	38	53.18	3.11	1.58	6.13	<0.001
Milker	family	1079	29	54.15	1			
	outside	133	3	6.66	0.84	0.25	2.81	1
	respondent	785	17	39.19	0.81	0.44	1.48	0.55
Region	Arusha	317	1	15.54	1			
	Iringa	280	1	13.73	1.13	0.117	10.91	1
	Kilimanjaro	507	13	25.42	5.63	1.036	30.62	0.037
	Mbeya	217	0	10.61	0.48	0.0197	11.96	0.517
	Njombe	158	29	9.14	39.27	7.55	204.79	<0.001
	Tanga	518	5	25.56	2.24	0.37	13.68	0.717
Abortion history	no	1820	42	91.0	1			
	yes	177	7	9.0	1.72	0.76	3.88	0.19
Placenta Mngnt	bad practices	337	11	17.01	1			
	good practices	1660	38	82.99	0.69	0.35	1.37	0.33

OR= Odds ratio, Mngnt= Management, Proportion%= proportion of each level in a variable

4.1.8 Multivariable analysis: logistic regression (backward elimination)

The summary result of the logistic regression analysis for the best-fit model indicated that a distance of less than 100 m between farms, extensive management, keeping goats, keeping pigs, and a history of abortion were all significantly associated with brucellosis seropositivity (p<0.05).

The findings suggest that, dairy cattle on farms located within a distance of less or equal to 100 m between dairy farms were 0.23 times less likely to contracting brucellosis (OR= 0.23, 95%CI: 0.13-0.43) compared to those on farms situated more than 100 m apart from each other. Additionally, the results demonstrated that, management practices significantly influenced brucellosis seropositivity. Dairy cattle in farms practicing extensive management were 0.16 times less likely to be *Brucella* seropositive (OR=0.16, 95%CI: 0.06-0.37) compared to those in farms employing intensive management.

Furthermore, keeping goats or pigs on the farm was significantly associated with brucellosis seropositivity (p<0.05), and dairy cattle on farms with these animals were 2 times more likely to be *Brucella* seropositive (goat: OR=2.7, 95%CI: 1.25-6.45, pigs: OR=2.3, (OR=2.3, 95%CI: 1.15-4.78) compared to those on farms that do not keep goats or pigs. History of abortion was found to be significantly associated with brucellosis seropositivity (p<0.05) and dairy cattle on a farm with a history of abortion were more likely to be *Brucella* seropositive than those on a farm with no history of abortion (OR=2.55, 95%CI: 0.99-5.84). Analysis also showed that keeping dogs and keeping sheep was associated with an increased risk of *Brucella* infection however the association was not significant (p>0.05) (Table 7).

Table 7: Multivariable analysis results based on logistic regression

Variable	Level	OR	95%CI	P-value
Distance	<=100 m	-	-	
	>100 m	0.23	0.13-0.43	<0.05
Management	intensive	-	-	
	extensive	0.16	0.06-0.37	<0.05
Goats	no goats	-	-	
	have goats	2.7	1.25-6.45	<0.05
Dogs	no dogs	-	-	
	have dogs	5.98	3.14-12.58	>0.05
Pigs	no pigs	-		
	have pigs	2.3	1.15-4.78	<0.05
Sheep	no sheep	-	-	
	have sheep	1.6	0.68-3.12	>0.05
History of Abortion	no	-		
	yes	2.55	0.99-5.84	<0.05

OR = Odds Ratio, CI = Confidence interval

4.1.9 Multivariable mixed effect model (Backward elimination)

The multivariable mixed effect model with district as a random effect term revealed that, dairy farms keeping goats, the introduction of new animals in the past 12 months, and a history of abortion were significantly associated with brucellosis seropositivity (p<0.05), as shown in Table 8. The presence of goats around dairy cattle was found to be significantly associated with brucellosis seropositivity. Dairy cattle on a farm with goats around them were more likely to test positive for brucellosis (p<0.05, OR=2.89, 95%CI: 1.19-7.03). A history of abortions on dairy farms was also significantly associated with brucellosis seropositivity (OR=5.7, 95% CI: 1.92-17). Dairy cattle on farms with a history of abortion were more likely to test positive for *Brucella* antibodies than cattle on farms with no history of abortion. Additionally, the introduction of new animals into a farm over the past 12 months was found to be significantly associated with brucellosis seropositivity (p<0.05). Dairy cattle in farms that introduced new cattle were less likely to test positive for brucellosis than those on farms without such introductions (p<0.05, OR=0.06, 95% CI: 0.01-0.51), as shown in Table 8.

Table 8: Results of multivariable analysis based on mixed effect model

Variable	Level	OR	95%CI	*P*-value
Goats	no goats	-	-	
	have goats	2.89	1.19-7.03	0.01
New animals	no	-	-	
	yes	0.06	0.01-0.51	0.01
Abortion history	no	-	-	
	yes	5.7	1.92-17	0.001

OR = Odds Ratio, CI = Confidence interval

4.1.10 Multivariable mixed effect model for Brucella seropositivity in regions with highest number of seropositive animals

A more focused analysis was conducted by subsetting the dataset to include only the three regions with relatively higher numbers of seropositive animals (Tanga, Kilimanjaro and Njombe). The final mixed-effects model revealed that having goats around and history abortion in the past 12 months were significantly associated with *Brucella* seropositivity (p<0.05). Dairy cattle kept on farm with goats around were more likely to contract the infection than those on a farm without goats (OR=3.02, 95% CI: 1.22-7.46). Furthermore, dairy cattle kept on a farm with a history of abortion in the past 12 months had higher odds of being (OR=4.91, (95% CI: 1.46 – 16.9) *Brucella* seropositive than those on a farm with no history of abortion during the same period (Table 9).

Table 9: Multivariable mixed effect model for Brucella seropositivity in dairy cattle in Tanga, Kilimanjaro, and Njombe regions of Tanzania

Risk Factor	Level	OR	95% CI	*p*-value
Goats	no goat	-	-	
	have goats	3.02	1.22-7.46	0.017
Abortion history	no	-	-	
	yes	4.91	1.43-16.9	0.012

OR = Odds Ratio, CI = Confidence interval

4.1.11 Molecular detection of Brucella species from genus-positive blood and vaginal DNA samples

(i) *Brucella* genus IS711 primer pair *in silico* specificity and sensitivity testing in GenBank

For the IS711 in silico specificity test, more than 299 amplicons were returned from the primer blast and all of them were bacteria of the genus *Brucella* suggesting that the specificity of the IS711 primer was good (100%). However, different *Brucella* species were identified, such as *B. abortus, B. melitensis, B. canis, B. suis, B. pinnipedialis,* were returned from the *in-silico* amplification. For the primer pair matching, different positions in the chromosome 1 and chromosome 2 of the amplicons were matched and amplified (Fig. 28).

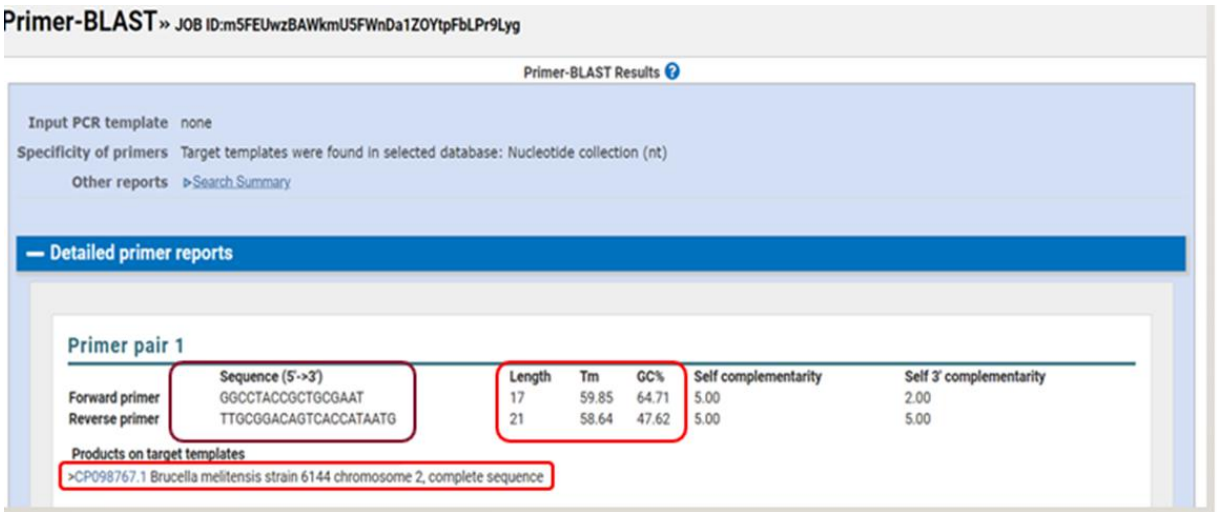


Figure 28: Primer-blast results for genus detection showing the query sequences used (IS711 primer pair sequences) and the first amplicon (product) with accession number CP098767.1 which was used for the nucleotide blasting (BLASTn)

For the primer pair sensitivity test, the first returned sequence in the primer blast with an accession number CP098767.1 (*B. melitensis strain* 6144 chromosome 2, complete sequence)

was used at the query sequence and the amplified region of the reverse primer from 472369bp to 472388bp was used as a query subrange in a BLASTn window.

The blastn results showed that one thousand amplifications (those with significant alignment) were returned and each amplicon was identified. The blastn amplicons were manually compared to primer blast amplicons by using their accession numbers. Because all the returned sequences during the primer blast were again returned during the BLASTn (did not miss out any amplicons returned during primer blast) (Fig. 29). The results suggested that the primer pair had good sensitivity (100% sensitivity).

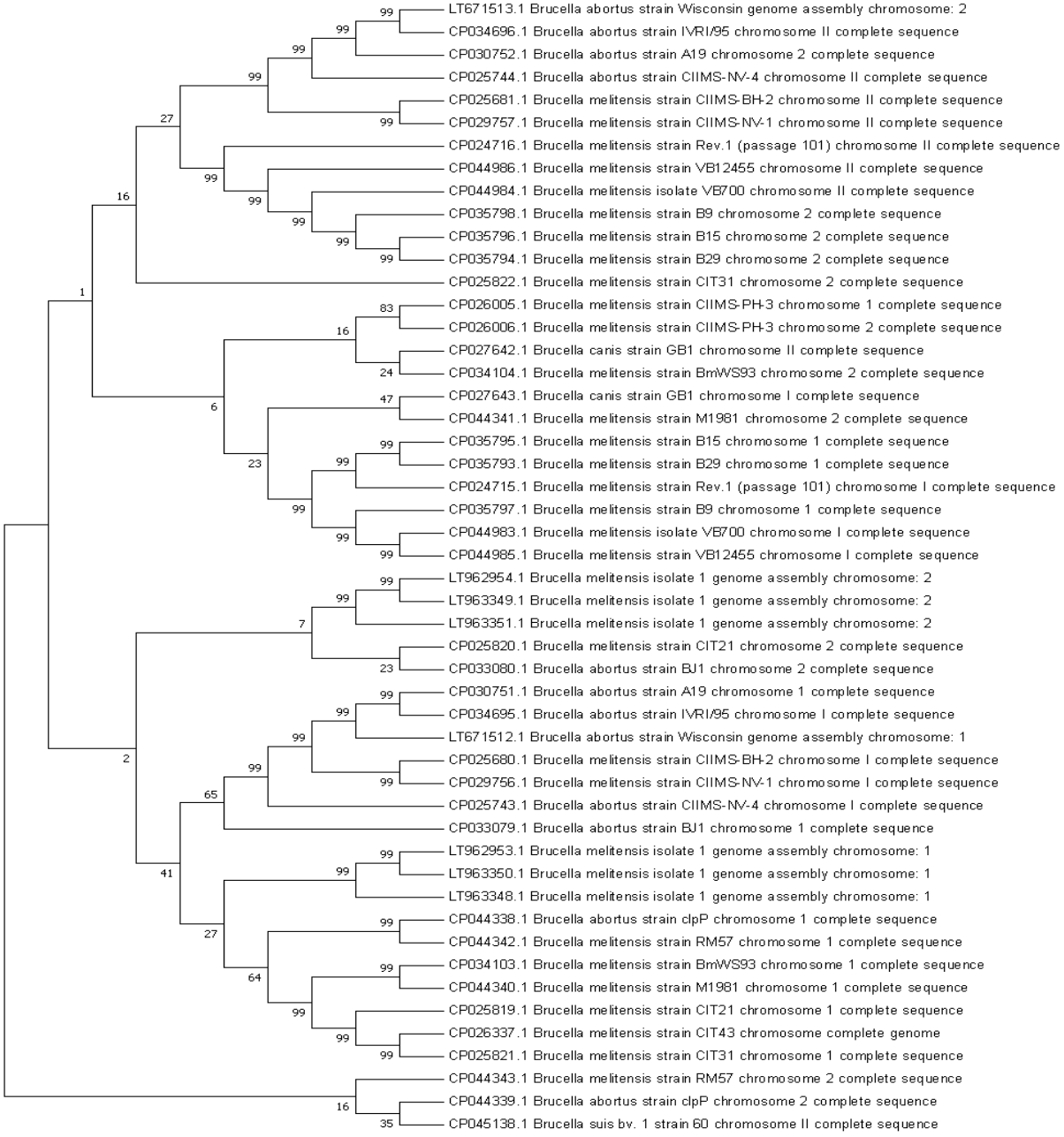


Figure 29: The cladogram of the first 50 BLASTn results showing taxa topological relationship of *Brucella spp.* amplified for genus detection

The confidence levels of tree topology are shown by bootstrap value in percentage. Branches with >80% bootstrap value are closely related. The cladogram does not show the root because the outgroup was not included in the cladogram reconstruction. According to these findings, the primer pair IS711 had good specificity and sensitivity, and therefore the laboratory results from this primer pair were reliable and were used for upstream analysis.

(ii) *Brucella abortus* alkB primer pair specificity and sensitivity testing results

For specificity, a total of 118 amplicons with significant alignment to the pair primer were returned. However, six (6) out of the 118 returned sequences (cross-reactive products) were not B. abortus but rather were B. melitensis biovar Abortus (3), B. ovis, and B. suis (2). This suggests that 95% of the amplicons were B. abortus and therefore the primer pair had good specificity (Fig. 30).

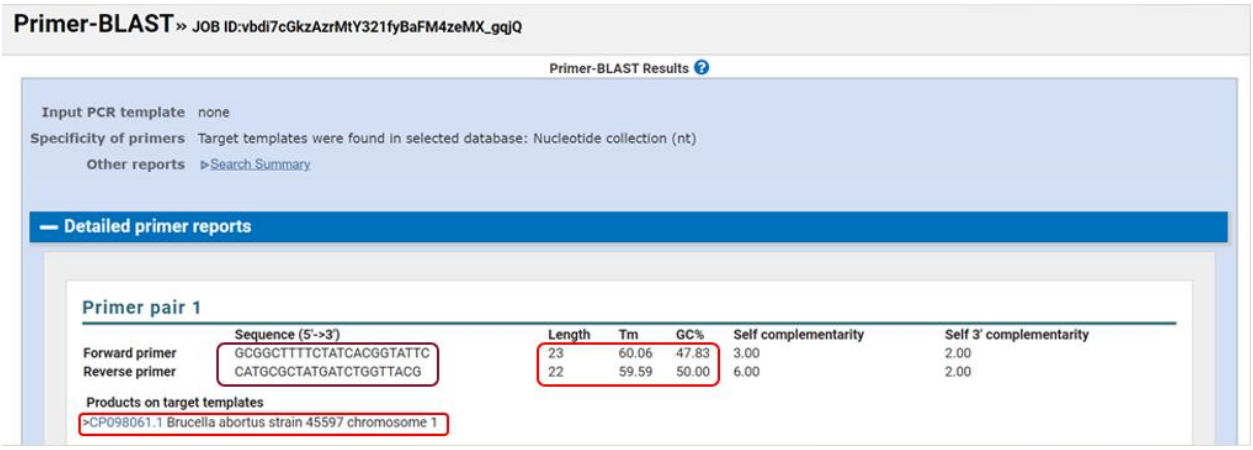


Figure 30: The alkB pair primer blast results, showing primers sequences used (alkB primer pair sequences) and the first returned sequence with an accession number CP098061.1 which was used for the nucleotide blasting (BLASTn)

For the in-silico sensitivity testing of the alkB, the first returned sequence from the primer-blast with an accession number CP098061.1 (*B. abortus strain* 45 597 chromosome 1) was used as a query sequence in the BLASTn. The amplified region of the reverse primer from 138 936bp to 138 957bp was used as a query subrange in a BLASTn window.

The nucleotide blasting results showed that, one thousand amplifications (those with significant alignment) were returned including those returned during the primer blast, and were identified and compared to primer blast results by using their accession numbers. In addition, *B. melitensis*, *B. suis*, and *B. ovis* were amplified. Because all the returned sequences (those that produced significant alignment to the primer pair) during the primer blast were all again returned during the BLASTn, this suggested that the primer pair had good sensitivity (100%).

As per Fig. 31 the confidence levels of tree topology are shown by bootstrap value in percentage. Branches with >80% bootstrap value are closely related.

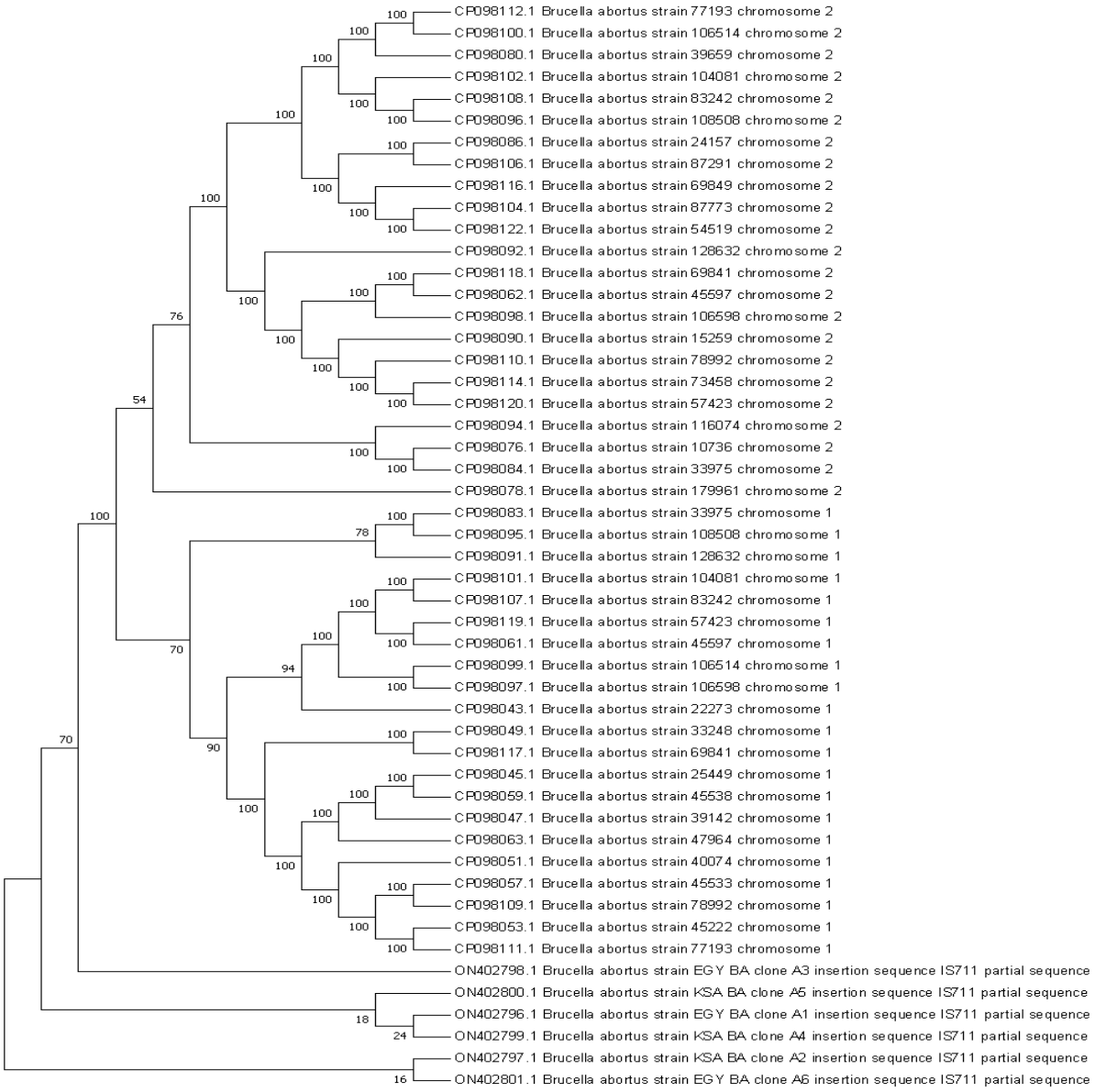


Figure 31: The cladogram of the first 50 BLASTn results showing the taxa topological relationship of the amplified *B. abortus*

(iii) *Brucella melitensis* BMEI1162 primer pair specificity and sensitivity testing in GenBank

For specificity testing, a total of 166 amplicons with significant alignment to the primer pair were returned. However, three (3) out of the 166 returned sequences (cross-reactive products) were not *B. melitensis* but rather *B. suis* (1) and *B. ceti* (2). This suggests that 98.2% of the amplicons were *B. melitensis* and therefore the primer pair had good specificity (Fig. 32).

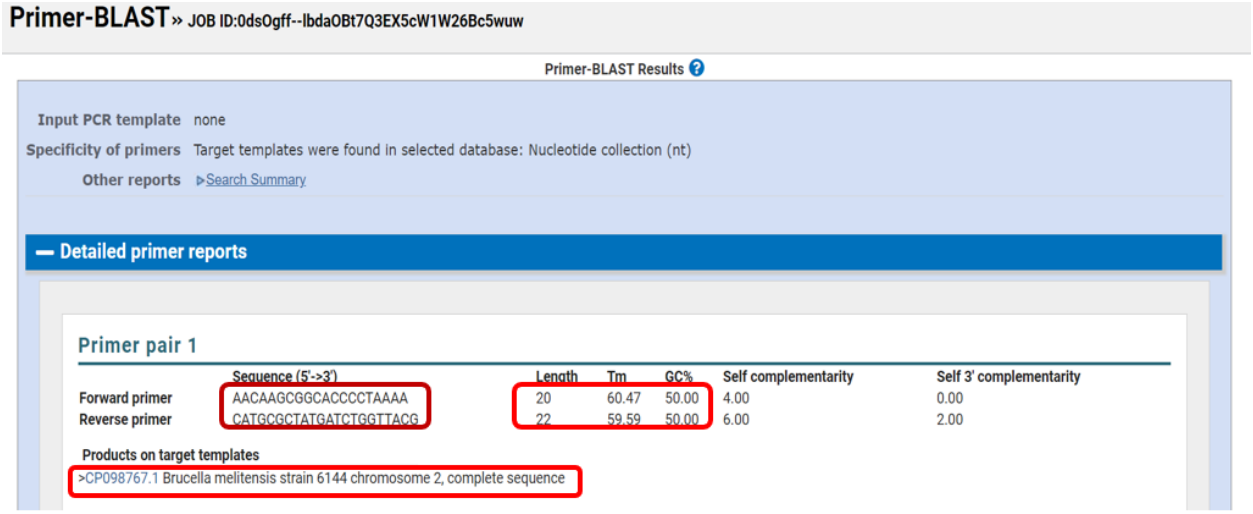


Figure 32: Results of BMEI1162 primer pair blast, showing the first returned sequence (product) with accession number CP098767.1 which was used for the nucleotide blasting (BLASTn)

For sensitivity testing of the BMEI1162, the first returned sequence from the primer-blast with an accession number CP098767.1 (*Brucella melitensis* strain 6144 chromosome 2 complete sequence) was used as a query sequence in the BLASTn. The amplified region of the reverse primer from 79 495bp to 795 16bp was used as a query subrange in a BLASTn window. The blasting results showed that, one thousand amplifications (those with significant alignment) were returned including those returned during the primer-blast, and were identified and compared to primer-blast results by using their accession numbers. Because all the returned sequences (those that produced significant alignment to the pair primer) during the primer blast were all again returned during the BLASTn, this suggested that the primer pair had good sensitivity (100%). The cladogram (Fig. 30) shows the topological relationship of the first 50 BLASTn amplicons results including some few *B. abortus* which had a similar sequence to the query sequence (Fig. 33). The confidence levels of tree topology are shown by bootstrap value in %. Branches with >70% bootstrap value are closely related.

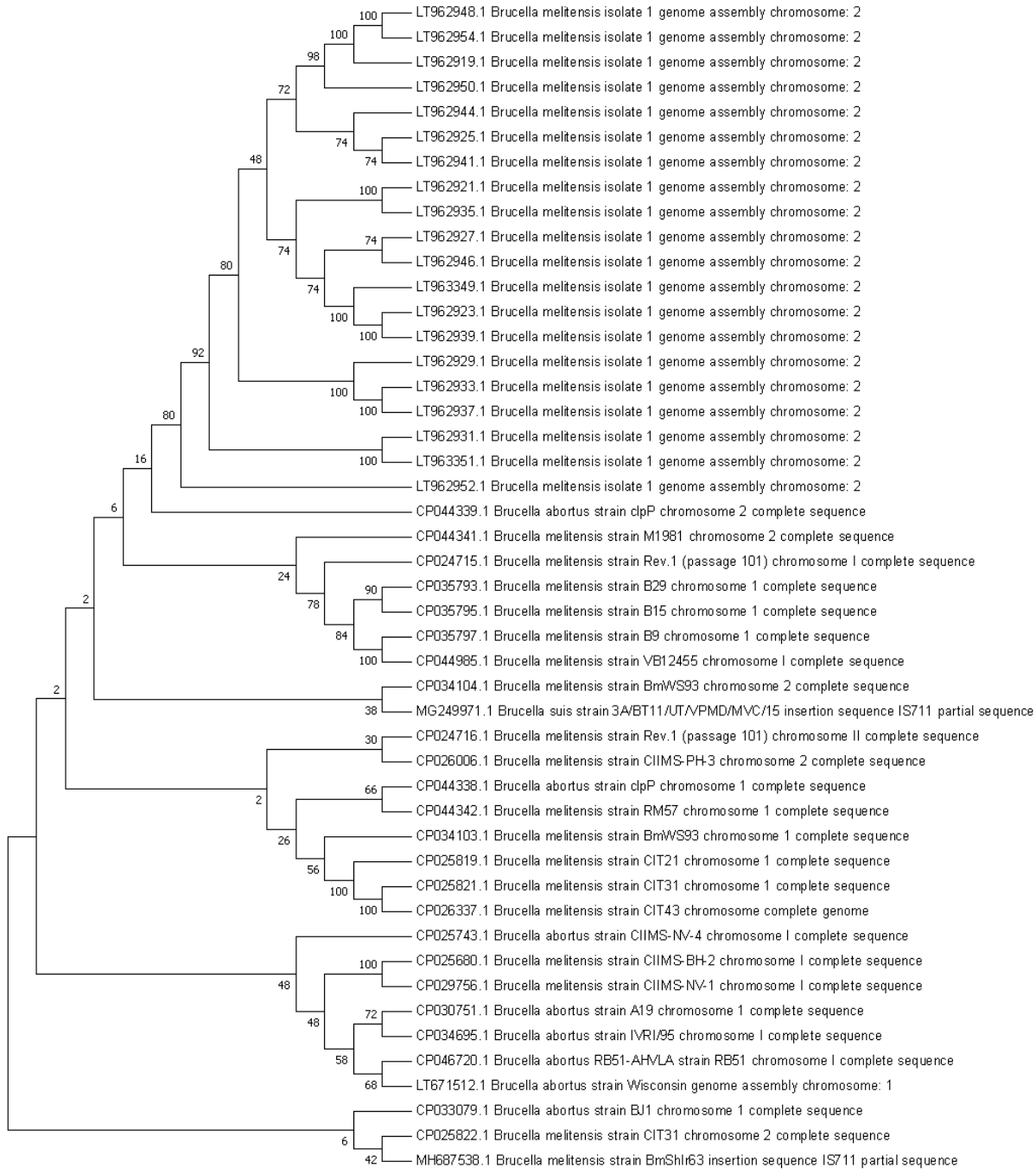


Figure 33: Topological relationship of *Brucella spp*. amplified for *B. melitensis* detection

4.1.12 The Brucella genus detection from blood DNA and swab samples

For Brucella genus detection, the methodology used was as described in Chapter three using IS711 primer pair. The amplification plot in Fig. 34 shows the genus positive blood DNA samples (b).

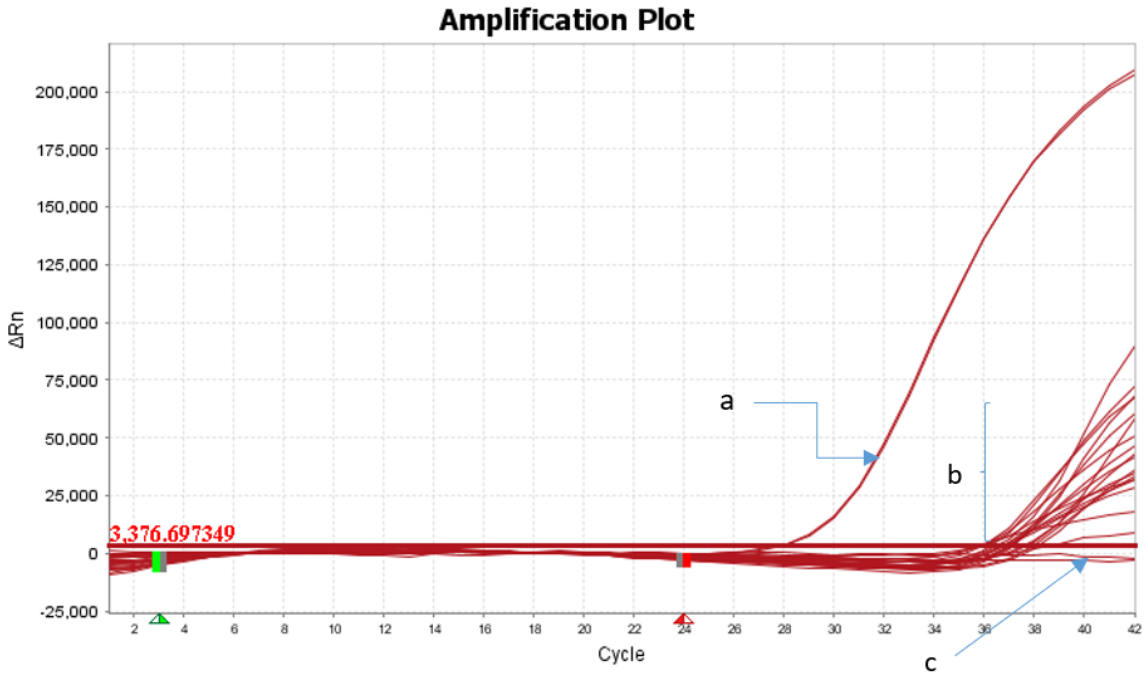


Figure 34: Amplification curves from blood DNA samples: (a) amplification curves for positive control, (b) amplification curves for positive samples, and (c) control negatives and negative samples

Similar methodology was used for swabs DNA samples to detect *Brucella* at genus level, the results are represented by amplification curve in Fig. 35.

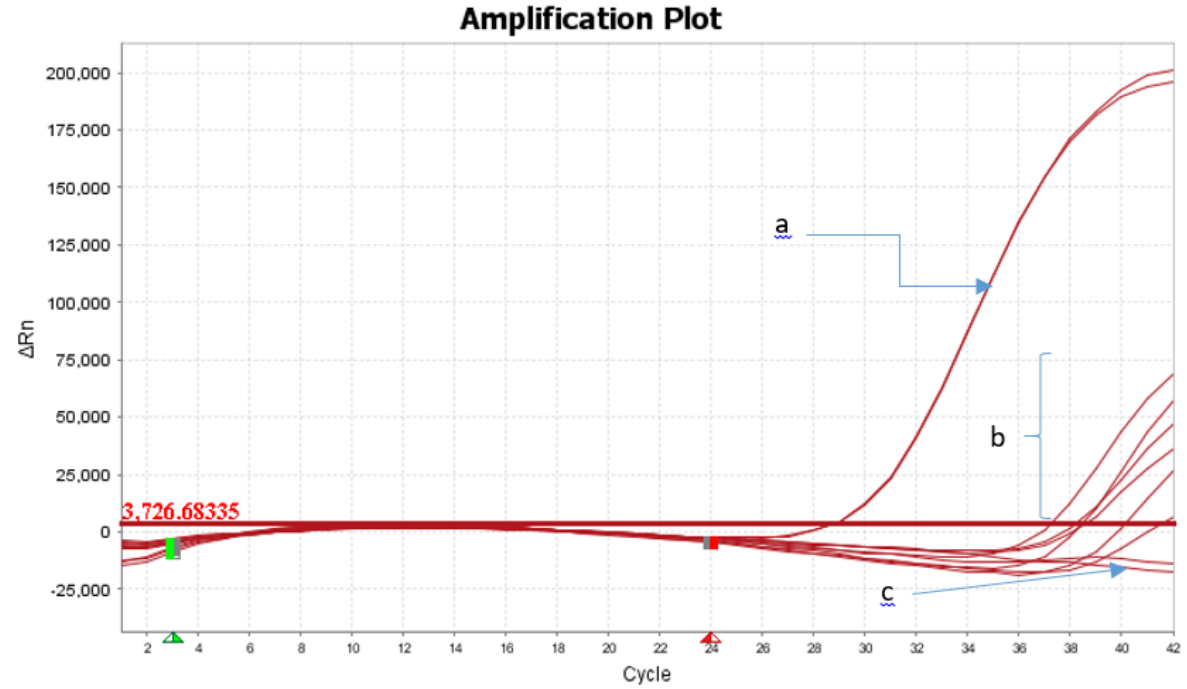


Figure 35: Amplification curves from swab DNA samples: (a) amplification curves for positive control, (b) amplification curves for positive samples, and (c) control negatives and negative samples

A total of 36 genus-positive blood DNA samples and 37 genus positive swab DNA samples were identified. The distribution of the positive blood DNA samples and positive swab DNA samples for each region was as shown in Table 10.

Table 10: Regional distribution of Brucella genus positives from blood DNA and swab DNA samples

Region	Total Animals Sampled	Number of Positive Blood Samples (%)	Number of Positive Swab Samples (%)
Arusha	318	5/318 (1.6%)	10/294 (3.4%)
Tanga	524	5/524 (1.0%)	1/412 (0.2%)
Kilimanjaro	521	12/521 (2.1%)	8/513 (1.6%)
Iringa	281	7/281 (2.5%)	1/273 (0.4%)
Njombe	187	2/187 (1.1%)	14/186 (7.5%)
Mbeya	217	5/217 (2.3%)	3/215 (1.4%)
Total	**2048**	**36/2048 (1.76%)**	**37/1893 (2.0%)**

4.1.13 Brucellosis PCR-prevalence in selected study regions of Tanzania

The results showed that the Kilimanjaro region had the highest number of positive blood DNA samples for the *Brucella* genus (12/521), followed by the Iringa region (5/281), Tanga region (5/524), Arusha (5/318), Mbeya region had 5/217, and lastly, Njombe region 2/187 genus positive blood DNA samples (Table 11).

The overall unadjusted PCR prevalence of *Brucella* spp. in smallholder dairy cattle was 1.76 % (95% CI: 1.23-2.43) with Iringa region having the highest unadjusted brucellosis PCR-prevalence of 2.49% (95%CI:1.01-5.06), followed by Kilimanjaro 2.31% (95%CI: 1.2-3.99), Mbeya region 2.3% (95%CI: 0.75-5.29), Arusha region 1.6% (95%CI: 0.51-3.63), Njombe region 1.07% (95%CI: 0.13-3.81) and Tanga region 0.95% (95%CI: 0.03-2.22) (Table 11).

Table 11: Genus positive results from blood DNA samples

Region	Negative	Positive	Total	PCR_Prevalence%	95% CI
Arusha	313	5	318	1.6	0.51-3.63
Tanga	519	5	524	1.00	0.03-2.22
Kilimanjaro	509	12	521	2.3	1.2-3.99
Iringa	274	7	281	2.5	1.01-5.06
Njombe	185	2	187	1.1	0.13-3.81
Mbeya	212	5	217	2.3	0.75-5.29
Total	**2012**	**36**	**2048**	**1.76**	**1.23-2.43**

4.1.14 The Brucella species detection from Brucella genus-positive DNA samples

Brucella species identification was done on the IS711 genus-positive blood DNA samples as described in methods sections. The amplification plot below shows the sigmoid curves for the *B. abortus* positive blood DNA samples (b), the control positive (a), and negative blood DNA and negative control (c) (Fig. 36). Among the genus *Brucella* blood DNA positive samples, 13 were positive to *B. abortus*.

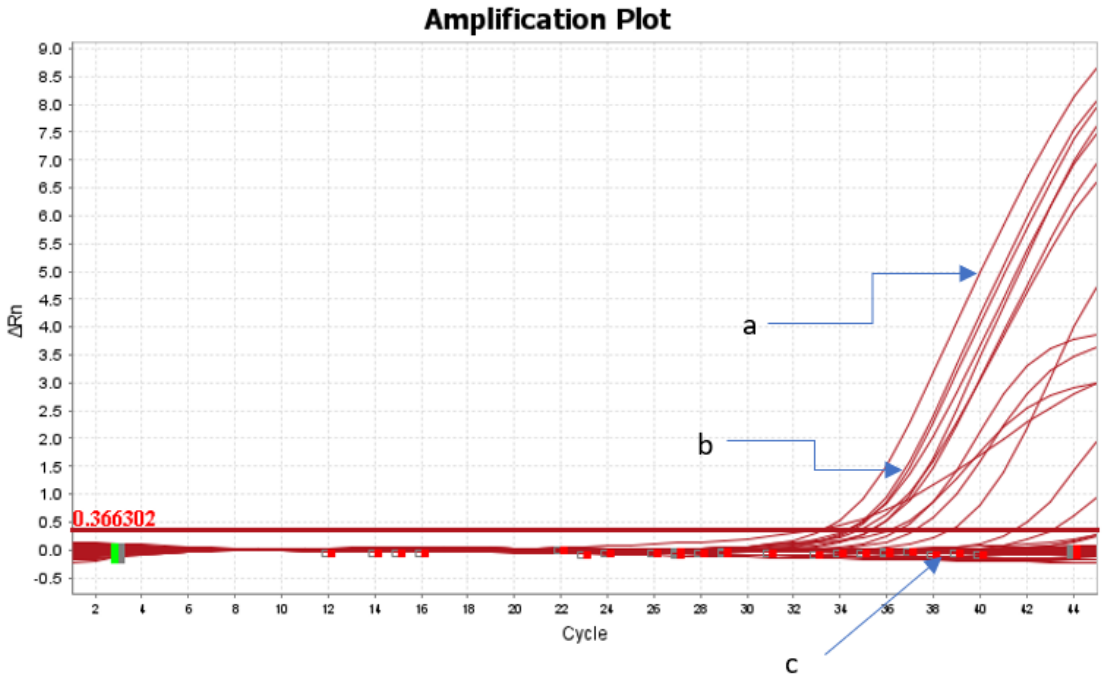


Figure 36: Amplification plot for the *B. abortus* showing amplification curves for control positive (a), positive samples (b), and negative control and negative samples (c)

The amplification plot in Fig. 37 shows the sigmoid curves for the *B. melitensis* positive blood DNA samples (b), the control positives (a), and negative blood DNA and negative control (c). Among the genus *Brucella* blood DNA positive samples, 30 were positive to *B. melitensis*.

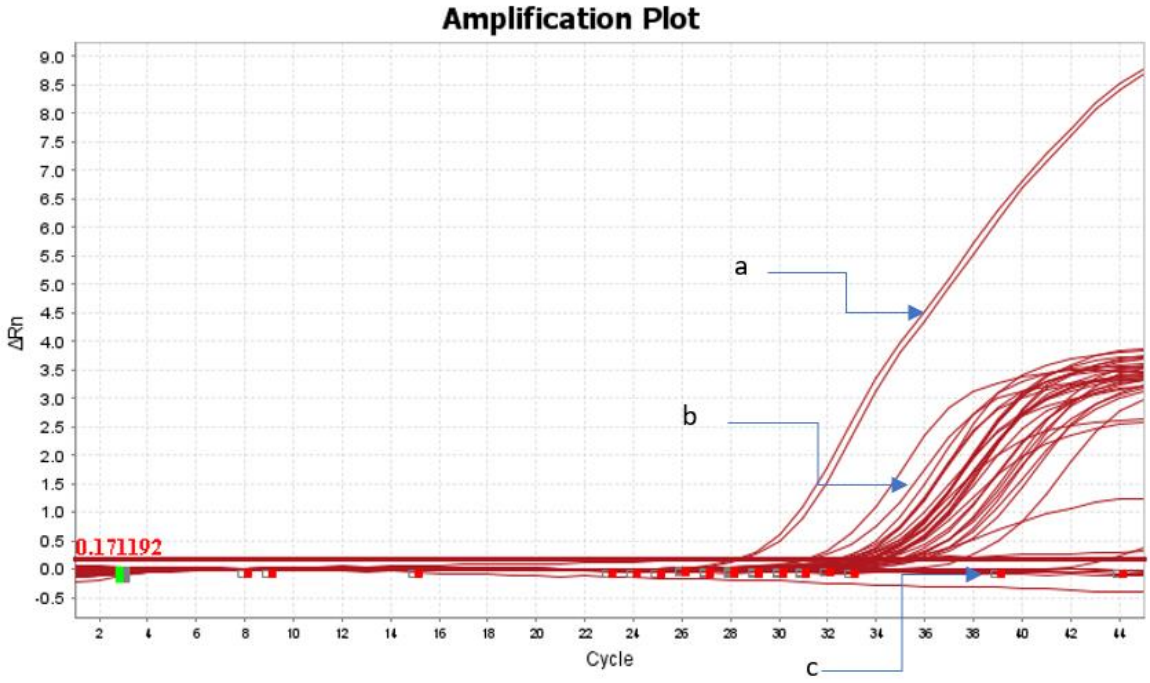


Figure 37: The amplification plot for BMEI1162 primer for *B. melitensis* detection showing the amplification curves for control positive (a), positive samples (b), and control negative and negative samples (c)

4.1.15 Brucella species identified from Brucella genus positive from blood DNA and vaginal swab DNA samples

The majority of blood samples (19/36 (52.8%)) and vaginal swabs (29/37 (78.4%)) were PCR-positive for *B. melitensis* only, with a further 11/36 (30.5%) blood samples and 7/37 (18.9%) swabs found to be PCR-positive for both (*B. melitensis* and *B. abortus*), *B. abortus* occurred on its own in 2/36 (5.6%) blood samples. There were four blood samples (11.4%) and one swab sample (2.7%) in which no species was determined. The majority of blood samples (19/35 (54.3%)) and vaginal swabs (29/37 (78.4%)) were PCR-positive for *B. melitensis* only, with a further 10/35 (28.6%) blood samples and 7/37 (18.9%) swabs found to be PCR-positive for both (*B. melitensis* and *B. abortus*), meaning that the vast majority of infections involved *B. melitensis* (Table 12). The *B. abortus* occurred on its own in 2/35 (5.7%) blood samples. There were four blood samples (11.4%) and one swab sample (2.7%) in which no species was determined (Table 12).

Table 12: Real-time polymerase chain reaction (qPCR) results for Brucella species identified from genus-positive blood and swab samples

Region	Sample	*B. abortus*	*B. melitensis*	Mixed	Undetermined
Arusha	Blood n = 5	0	3	2	0
	Swabs n = 10	0	9	1	0
Kilimanjaro	Blood n = 12	1	9	1	1
	Swabs n = 8	0	6	2	0
Tanga	Blood n = 5	0	2	3	0
	Swabs n = 1	0	1	0	0
Njombe	Blood n = 2	0	2	0	0
	Swabs n = 14	0	11	2	1
Iringa	Blood n = 7	0	3	2	2
	Swabs n = 1	0	1	0	0
Mbeya	Blood n = 5	1	1	3	0
	Swabs n = 3	0	1	2	0
Total	Blood n = 36	2(5.6%)	19(52.8%)	11(30.5%)	4(11.4%)
	Swabs n = 37	0(0%)	29(78.4%)	7(18.9%)	1(2.7%)

Mixed = PCR-positive for both *B. abortus* and *B. melitensis*.

4.1.16 Genome-wide association studies (GWAS) to identify SNP markers in dairy cattle associated with brucellosis inherent resistance/susceptibility

(i) Heritability Prediction results

Univariable and multivariable analysis was conducted using ASREML software to identify the explanatory variable that was significantly associated with the brucellosis PCR positivity and later be used for GWAS analysis to avoid over-parametrization.

(ii) Univariable analysis results for prediction of PCR positivity and heritability by using analystical software for residual maximum likelihooh

The results in Table 13 showed that the REMLRT was the same throughout the univariable analysis and none of the explanatory variables were significantly associated with *Brucella* PCR positivity (p>0.05). All models had the same REMLRT of 263.3. The general relationship matrix for the random effect (Grm1(id)) was found to be significantly associated with PCR positivity in all models with a p-value of less than 0.15E-07 (p<0.05). The univariable analysis results also showed that none of the explanatory variables (fixed effects) was a good predictor (had a p-value less than 0.05) for heritability and therefore heritability estimation is zero for all explanatory variables.

Table 13: Univariable analysis results of the explanatory variables for PCR results and Heritability prediction

Variable (fixed)	Random effect	LogL	Remlrt	P-value	Grm1(id)	Residual	Heritability
sex	grm1(id)	3030.98 2899.33	263.3	0.955	0.1366E-07	0.1694E-01	0.00
age	grm1(id)	3028.03 2896.39	263.28	0.474	0.1642E-08	0.1693E-01	0.00
breed	grm1(id)	3031.97 2900.33	263.28	0.128	0.1300E-07	0.1692E-01	0.00
herdsize	grm1(id)	3026.27 2894.63	263.28	0.496	0.1301E-07	0.1693E-01	0.00
management	grm1(id)	3031.47 2899.82	263.3	0.086	0.1299E-07	0.1691E-01	0.00
bull	grm1(id)	3030.07 2898.43	263.28	0.681	0.1493E-07	0.1693E-01	0.00
cont_dog	grm1(id)	3027.92 2896.27	263.3	0.824	0.1473E-07	0.1694E-01	0.00
cont_pigs	grm1(id)	3027.17 2895.52	263.3	0.964	0.1445E-07	0.1694E-01	0.00
cont_goats	grm1(id)	3027.39 2895.75	263.28	0.943	0.1507E-07	0.1694E-01	0.00
cont_sheep	grm1(id)	3027.52 2895.87	263.3	0.790	0.1322E-07	0.1694E-01	0.00
distance	grm1(id)	3030.00 2898.35	263.3	0.942	0.1503E-07	0.1694E-01	0.00
placenta	grm1(id)	3030.16 2898.52	263.28	0.895	0.1445E-07	0.1694E-01	0.00
cattleyears	grm1(id)	3027.44 2895.79	263.3	0.721	0.1328E-07	0.1694E-01	0.00
water	grm1(id)	3029.91 2898.26	263.3	0.926	0.1420E-07	0.1694E-01	0.00

(iii) Multivariable analysis results for prediction of PCR positivity and heritability by using analystical software for residual maximum likelihooh

Multivariable analysis was conducted in all explanatory variables to find out which variable would be significantly associated with PCR positivity as an outcome variable as none of the variable was significant at univariable analysis. The analysis was started with the complex model without random effect and then with the random effect so that the nested models can be compared using REMLRT and between non-nested model comparison was done by using AIC. Table 14 summarizes 11 models and their statistics for comparison and heritability estimation results. The nested models were compared using REMLRT, and the findings show that all models had the same REMLRT value of 263.3. Non-nested models were compared by using AIC, model 11 was found to have the smallest AIC value of -6022.02 and -5756.73 for the one without random effect and with random effect respectively. The generalized relationship matrix model function for individual animals (grm(id)) has been significantly associated with PCR positivity with p-values of less than 0.1308E-07 ($p<0.05$) in all models. Results also show that none of the variables was significantly associated with PCR positivity (none had a p-value <0.05) that could explain the heritability. Heritability was found to be zero in all models.

Table 14: Multivariable analysis results of explanatory variables for brucellosis PCR positivity and Heritability prediction

	Model(s)	LogL	Remlrt	Ss	Aic	Grm(id)	Heritability
1	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta cattleyears water	2951.16	263.3	0.17034E-01	-5900.32		
	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta cattleyears water !r grm1(id)	2819.51		0.17034E-01	-5635.02	0.1309E-07	0.00
2	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta cattleyears	2956.07	263.3	0.17027E-01	-5910.14		
	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta cattleyears !r grm1(id)	2824.42		0.17027E-01	-5644.85	0.130844E-07	0.00
3	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta	2963.42	263.3		-5924.84		
	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance placenta !r grm1(id)	2831.77		0.17019E-01	-5659.55	0.1308E-07	0.00
4	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance	2968.23	263.3	0.17010E-01	-5934.46		
	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance !r grm1(id)	2836.58		0.17010E-01	-5669.17	0.1307E-07	0.00
5	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep	2972.95	263.3	0.17005E-01	-5943.89		
	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats cont_sheep distance !r grm1(id)	2841.30		0.17005E-01	-5678.60	0.130673E-07	0.00
6	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats	2980.53	263.3	0.16989E-01	-5959.06		
	Age sex breed herdsize management bull cont_dogs cont_pigs cont_goats !r grm1(id)	2848.88					0.00

	Model(s)	LogL	Remlrt	Ss	Aic	Grm(id)	Heritability
				0.16989E-01	-5693.77	0.130551E-07	
7	Age sex breed herdsize management bull cont_dogs cont_pigs	2988.11	263.3	0.16973E-01	-5974.22		
	Age sex breed herdsize management bull cont_dogs cont_pigs !r grm1(id)	2856.47		0.16973E-01	-5708.93	0.1304E-07	0.00
8	Age sex breed herdsize management bull cont_dogs	2995.89	263.3	0.16956E-01	-5989.79		
	Age sex breed herdsize management bull cont_dogs !r grm1(id)	2864.25		0.16956E-01	-5724.50	0.1303E-07	0.00
9	Age sex breed herdsize management bull	3002.79	263.3	0.16945E-01	-6003.58		
	Age sex breed herdsize management bull !r grm1(id)	2871.14		0.16945E-01	-5738.29	0.1302E-07	0.00
10	Age sex breed herdsize management	3007.69	263.3	0.16937E-01	-6013.38		
	Age sex breed herdsize management !r grm1(id)	2876.04		0.16937E-01	-5748.09	0.1301E-07	0.00
11	Age sex breed herdsize	3012.01	263.3	0.16939E-01	-6022.02		
	Age sex breed herdsize !r grm1(id)	2880.36		0.16939E-01	-5756.73	0.1302E-07	0.00

Remlrt: Residual Maximum likelihood Ratio test = 2 [log(`R2) − log(`R1)] – compares models with the same fixed effects:
Aic: Akaike Information Criterion: compares non-nested models (models with different fixed effects)
Ss: Sum of Squares of the model
Grm(): generalized relationship matrix function
id: animal identification number as a random effect.
LogL: log likelihood test statistic

(iv) Genome-wide association studies analysis results by using Genome Efficient Mixed Model Algorithm

The Gemma analysis association file results are presented in Table 15. The table shows that, there are 13 columns, column one shows the chromosome number (Chr), column two the SNP id (Rs), column three base pair position on the chromosome (Ps), the number of missing individuals for a given SNP (N_mis), number of non-missing individuals for a given SNP (N_obs), minor allele (Allele1), major allele (Allele0), allele frequency (Af), beta estimates (Beta), standard errors for beta (Se), p-values for Wald test (P_wald), p-values for likelihood ratio test (P_lrt) and p-values for score test (P_score).

Table 15 shows the 23 SNPs markers at different positions on chromosome 1, cattle have 28 chromosomes and there were 1977 cattle and 668 911 SNP markers for analysis, not easy to show all the results here in this table. The results also show that the p-values for the 3 tests were almost the same for each SNP on the same animal (very small differences). For visualization of each SNP based on its genomic coordinate and p-value, the Manhattan plot was plotted and shown in the next section.

Table 15: Gemma analysis results for the association of SNPs markers with the disease phenotype

Chr	Rs	Ps	N_mis	N_obs	Allele1	Allele0	Af	Beta	Se	P_wald	P_lrt	P_score
1	BovineHD2700011070	530998	0	1977	T	C	0.355	-1.49E-03	4.45E-03	7.38E-01	7.37E-01	7.37E-01
1	BovineHD0100000015	716721	0	1977	G	A	0.416	3.60E-03	4.39E-03	4.13E-01	4.10E-01	4.11E-01
1	BovineHD0100000026	719789	0	1977	G	A	0.405	3.59E-04	4.40E-03	9.35E-01	9.35E-01	9.35E-01
1	BovineHD0100000027	724546	0	1977	A	G	0.214	2.82E-03	5.17E-03	5.86E-01	5.84E-01	5.84E-01
1	BovineHD0100046367	730858	0	1977	C	T	0.24	1.00E-03	5.02E-03	8.42E-01	8.41E-01	8.41E-01
1	BovineHD0100000035	761316	0	1977	T	G	0.483	1.96E-04	4.23E-03	9.63E-01	9.63E-01	9.63E-01
1	BovineHD0100000037	769018	0	1977	G	T	0.105	-4.39E-03	6.91E-03	5.26E-01	5.24E-01	5.24E-01
1	BovineHD0100000038	778681	0	1977	A	G	0.094	-2.77E-03	7.29E-03	7.05E-01	7.03E-01	7.03E-01
1	BovineHD0100000042	790905	0	1977	A	C	0.489	-2.03E-03	4.29E-03	6.37E-01	6.35E-01	6.35E-01
1	BovineHD0100000043	792193	0	1977	C	T	0.489	-2.03E-03	4.29E-03	6.37E-01	6.35E-01	6.35E-01
1	BovineHD0100000044	793507	0	1977	T	C	0.489	-2.03E-03	4.29E-03	6.37E-01	6.35E-01	6.35E-01
1	BovineHD0100000046	797069	0	1977	G	A	0.419	4.63E-04	4.38E-03	9.16E-01	9.15E-01	9.15E-01
1	BovineHD0100000048	799951	0	1977	T	C	0.486	1.99E-03	4.25E-03	6.40E-01	6.38E-01	6.39E-01
1	BovineHD0100000049	801138	0	1977	G	A	0.404	1.54E-03	4.31E-03	7.21E-01	7.20E-01	7.20E-01
1	BovineHD0100000050	802335	0	1977	A	G	0.212	-3.06E-03	5.13E-03	5.51E-01	5.49E-01	5.49E-01
1	BovineHD0100000051	805126	0	1977	G	A	0.482	-1.87E-03	4.28E-03	6.61E-01	6.60E-01	6.60E-01
1	BovineHD0100000052	805814	0	1977	T	G	0.493	1.87E-03	4.24E-03	6.59E-01	6.58E-01	6.58E-01
1	BovineHD0100000053	809505	0	1977	G	A	0.127	-2.58E-03	6.33E-03	6.84E-01	6.83E-01	6.83E-01
1	BovineHD0100000054	810128	0	1977	T	C	0.228	-1.52E-05	5.06E-03	9.98E-01	9.98E-01	9.98E-01
1	BovineHD0100000055	812618	0	1977	G	A	0.297	2.05E-03	4.60E-03	6.57E-01	6.55E-01	6.55E-01
1	BovineHD0100046368	815169	0	1977	T	G	0.167	7.01E-04	5.87E-03	9.05E-01	9.05E-01	9.05E-01
1	BovineHD0100000056	817590	0	1977	C	T	0.167	7.01E-04	5.87E-03	9.05E-01	9.05E-01	9.05E-01
1	BovineHD0100000057	824172	0	1977	C	T	0.315	1.38E-03	4.53E-03	7.61E-01	7.60E-01	7.60E-01

(v) Study population stratification for GWAS analysis

Genome-wide association analysis requires a homogenous population that suggests nearly common ancestry. Figure 38 is a plot drawn using the principal component 2 on the x-axis and principal component 1 on the y-axis representing the genetic distance variation between individuals under the study. The results showed that most of the animals were clustered at the center showing an area of high genetic homogeneity with few animals on the periphery showing a slightly increased genetic distance from the cluster at the center. Generally, this study population is homogenous with only one cluster.

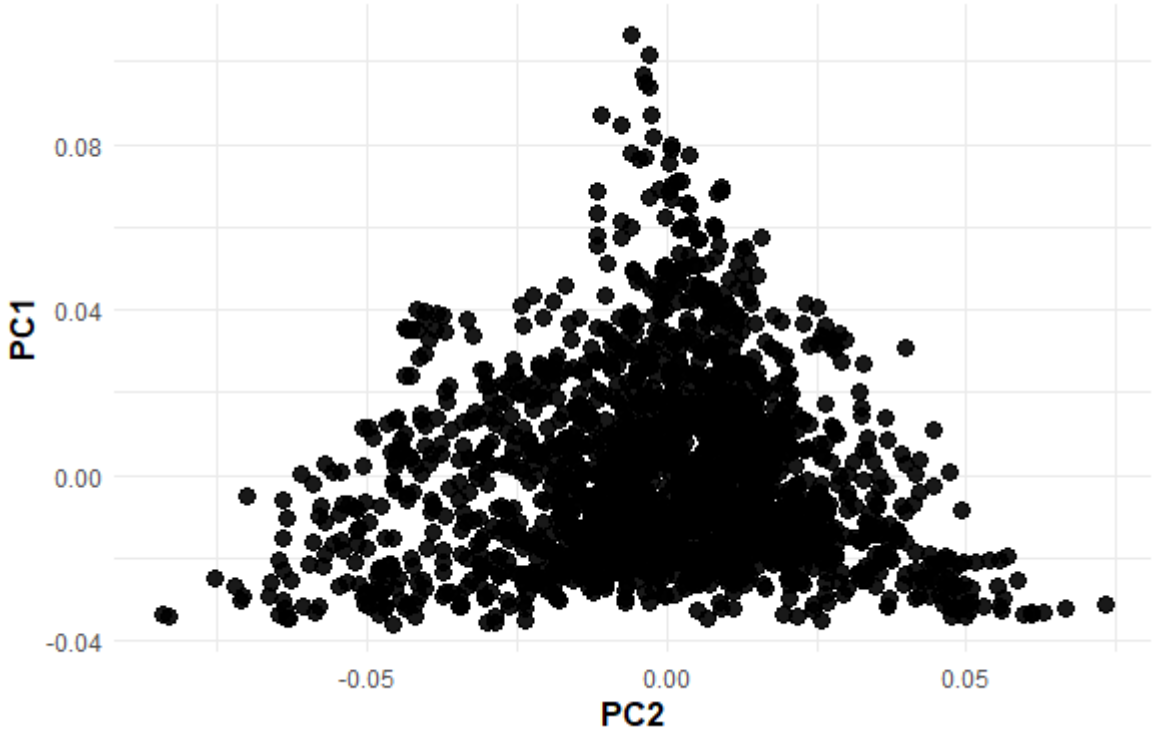


Figure 38: A plot showing the study population genetic distance variation between individuals in a population (population stratification)

(vi) Quantile-Quantile (Q-Q) plot for GWAS analysis results on disease phenotype

A Q-Q plot is a scatterplot created by plotting two sets of quantiles (Observed (Y-axis) vs Expected (X-axis) p-values) against one another. If both sets of quantiles came from the same distribution, one should see the points forming a roughly straight line. The qq-plot checks the overall distribution of –log10 (*p*-values) with the expectation under the null hypothesis of no association (the diagonal line shows where the points should fall under the null hypothesis). Figure 39 shows that the observed and expected SNP p-values had the same distribution and therefore lie on the diagonal. At the point of expected -log10 (*p*-values) between 3 and 4, some

SNPs started to assume different distributions suggesting a violation of the null hypothesis. One SNP circled red has shown the highest observed p-value different from others and the SNP mostly violated the null hypothesis therefore the alternative hypothesis is chosen as that was the most significant association.

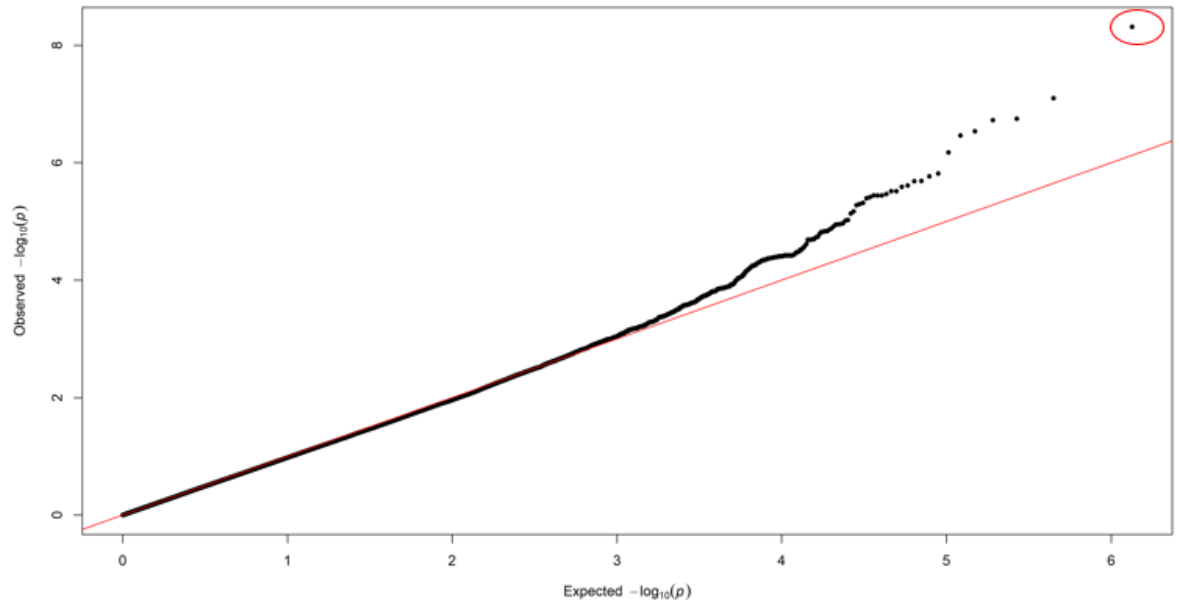


Figure 39: The Q-Q plot for GWAS analysis results shows a distribution of -log10 (p-values) of 668 911 SNPs tested for PCR positivity with some violating the null hypothesis

(vii) Manhattan plot for GWAS analysis results on disease phenotype

Manhattan plot was drawn by using the -log10 of the association p-values for each SNP against their genomic coordinates. The strongest associations must have the smallest p-values and the -log of these p-values will have the highest height in the plot. In Fig. 40, a red horizontal line is a genome-wide association threshold line and a blue horizontal line is a suggestive line. All the SNPs from chromosomes 1 to 29 were represented. Figure 40 shows that, at chromosome number 9, there is a point circled green with the highest height above the genome-wide association threshold line. This point is a SNP with the smallest p-value. There are some other neighboring SNPs (circled red), one on chromosome 8 and two on chromosome 9, and two distant SNPs (circled purple) one on chromosome 26 and chromosome 28, all above the suggestive line.

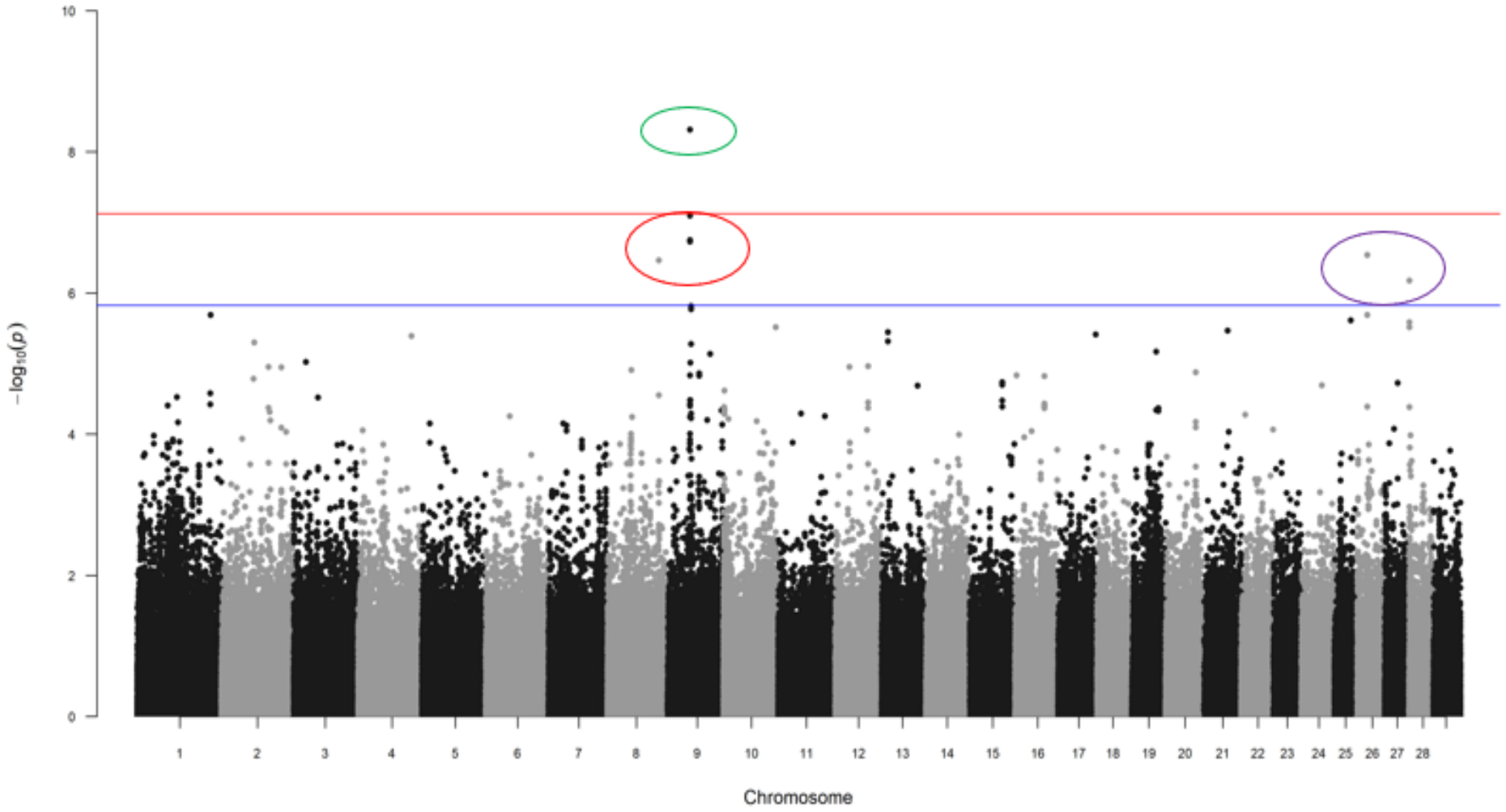


Figure 40: **Manhattan plot of association p-values of SNPs in 29 chromosomes (The x-axis shows the location and the y-axis displays the significance of the association (-log10 (p-values))**

(viii) The trait locus of core SNP marker with highest association to brucellosis PCR positivity

Genome-wide association study has identified one SNP significantly associated with brucellosis PCR positivity (core SNP) as it had the highest height on the Manhattan plot (Fig. 40). The core SNP is located on chromosome number 9 at the 41 852 753 base pair position. The core SNP name is BovineHD0900011750 and has A/G base pair as all SNPs are considered biallelic. The core SNP has a beta estimate value of 5.46E-02 when exponentiated giving a value of a measure of its effect size (Odds Ratio) of 1.05. This is true also for all the suggestive SNPs that each had an odds ratio of greater than 1. Other suggestive SNPs which have shown high LD with the core SNP are shown in Table 16. Each SNP has been identified by its chromosome location, SNP name, base-pair position, alleles, allele frequency, beta estimate (odds ratio), standard error (se), and p-values for different association tests done.

Table 16: The SNP markers showed the highest association with brucellosis PCR positivity and high LD

Chr	Rs	Ps	Allele1	Allele0	Af	Beta	Se	P_wald	P_lrt	P_score
9	BovineHD0900011750	41 852 753	A	G	0.056	5.46E-02	9.28E-03	4.83E-09	4.09E-09	5.59E-09
9	BovineHD0900011751	41 860 505	G	T	0.056	5.04E-02	9.36E-03	7.98E-08	6.93E-08	8.66E-08
9	BovineHD0900011748	41 836 969	G	A	0.064	4.59E-02	8.75E-03	1.77E-07	1.55E-07	1.90E-07
8	BovineHD0800029026	96 637 272	A	G	0.034	6.12E-02	1.19E-02	3.42E-07	3.011E-07	3.61E-07
26	BovineHD2600004903	19 059 082	C	T	0.06	4.44E-02	8.62E-03	2.89E-07	2.54E-07	3.07E-07
28	BovineHD2800000058	1 126 396	C	T	0.25	2.36E-02	4.72E-03	6.66E-07	5.90E-07	6.96E-07

(ix) Variant Effect Prediction (VEP) of the core single nucleotide polymorphism

The core SNP marker is associated with other 7 protein-coding regions (genes) located in chromosome 9. These genes include FOXO3, AFG1L, SNX3, NR2E1, OSTMI, SEC63 and SCML4. By location, the core SNP lies within the protein-coding gene called ATPase Family Gene 1 homology (AFG1L) (Fig. 41), a gene responsible for a mitochondrial integral membrane protein that plays a role in mitochondrial protein homeostasis.

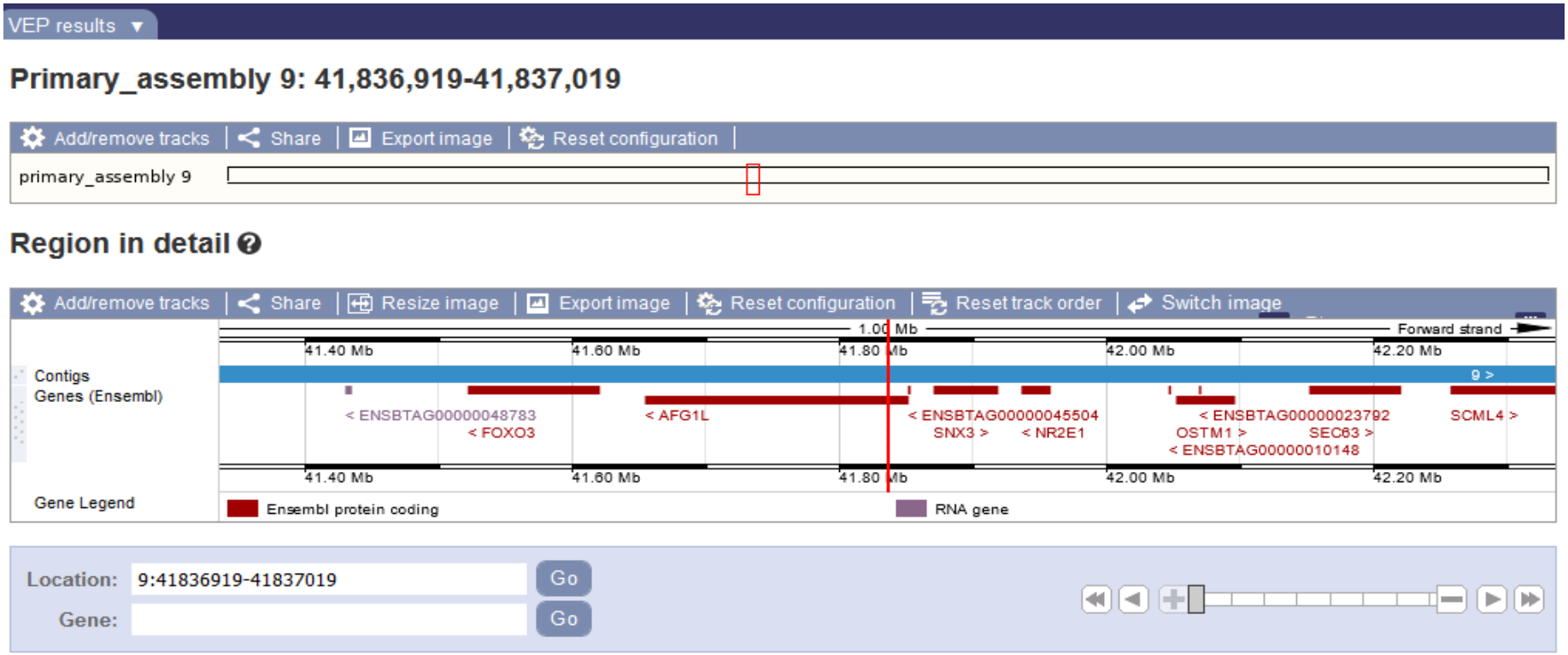


Figure 41: Web-based Ensembl VEP analysis results showing the location (vertical red line) and gene affected (AFG1L) by GWAS significant SNP (BovineHD0900011750) on cattle genome

4.2 Discussion

4.2.1 Brucellosis seroprevalence in dairy cattle in selected regions of Tanzania

Brucellosis is one of the globally neglected bacterial zoonosis (Asmare *et al.,* 2013; Schelling *et al.,* 2003). Even though brucellosis in domestic animals has been eradicated in high-income countries (HICs), the disease keeps causing significant economic losses in lower and middle-income countries (LMICs) including Tanzania because of uneffective control measures and lack of a well-established surveillance system (McDermott *et al.,* 2013; Terefe *et al.,* 2017). Several studies on brucellosis in Tanzania were conducted in different cattle production systems and established varying prevalences ranging from 0.6-30% and risk factors such as older age, contact with other animals, contact between infected and clean herds, management system, watering points, poor disposal methods of aborted materials (Karimuribo *et al.,* 2007; Mathew *et al.,* 2017; Mengele *et al.,* 2023b; Shirima *et al.,* 2018; Swai & Schoonman, 2010). Furthermore, recent studies showed that, brucellosis in dairy cattle is on increasing trend (Shirima *et al.,* 2018). Therefore, this study aimed to determine the current burden and related risk factors for brucellosis in dairy cattle kept under smallholder farming system in selected high milk-producing regions of Tanzania.

The findings from this study showed the presence of anti-*Brucella* antibodies circulating in dairy cattle populations in some of the 6 regions suggesting the presence of *Brucella* bacteria and brucellosis in some of the study regions particularly Njombe. The findings of this study are similar to other brucellosis studies carried out in dairy cattle in the country and some in similar regions (Karimuribo *et al.,* 2007; Mathew *et al.,* 2015; Sagamiko *et al.,* 2018; Swai *et al.,* 2005; Swai & Schoonman, 2010) which found varying levels of brucellosis.

The overall animal level seroprevalence recorded was 2.4%. It has to be acknowledged that, cELISA test apart from being widely used in Tanzania to report brucellosis in animals (Mengele *et al.,* 2023a), it has been tested to have a high sensitivity and high specificity of 98% and 99% respectively (Bodenham *et al.,* 2021) which gives enough confidence on the results. However, the assay is not 100% perfect; therefore, the number of true positive and true negative animals presented are likely being affected.

Brucellosis is a trans-boundary disease, and smallholder dairy cattle in neighboring countries are also affected. Cross-sectional and random sampling studies carried out in Zambia, Rwanda,

Burundi, Malawi, and Uganda have recorded higher brucellosis seroprevalences at 23.1%, 6% 7.7%, 14.7%, and 6% respectively (Bernard *et al.,* 2005; Chagunda *et al.*, 2014; Muma *et al.,* 2012; Musallam *et al.,* 2019; Segwagwe *et al.,* 2018) than the overall findings from this study, suggesting that dairy cattle in those countries have higher levels of brucellosis and possibly experience a higher impact of brucellosis compared to Tanzania. This also suggests that, dairy cattle coming from those countries to Tanzania should be given the highest level of alert as potential sources of brucellosis infection to Tanzania dairy cattle. Countries like Ethiopia and Kenya have recorded seroprevalences between 1% to 1.9%, which is lower than the current result (Asmare *et al.,* 2013; Gicheru *et al.,* 2015; Terefe *et al.,* 2017). Studies carried out in other countries like Nigeria, Ecuador, and India have recorded seroprevalences between 5.5% and 29.6% (Lindahl *et al.,* 2018; Mai *et al.,* 2012; Poulsen *et al.,* 2014), which was higher than the findings from this study. These findings suggest that brucellosis is present and affecting cattle in SSA and more so in smallholder dairy cattle, hence regional and global integration efforts are required for the unified strategic control of brucellosis.

At the regional level, the highest animal level seroprevalence was recorded in the Njombe region (15.5%) while the Mbeya region was the least with no seropositive animal (0%). A similar research study carried out by Mathew *et al.* (2015) in Njombe provides further evidence that Njombe region may be the current brucellosis hot spot under the smallholder dairy farming system in Tanzania. Njombe being the region with the highest seroprevalence may be attributed by the fact that the dairy sector is growing very fast, dairy cattle sales within the region is high, and mixed farming practice where goats are reared together with dairy cattle. Furthermore, smallholder dairy cattle in Njombe region depend on private or communal open well water as a source of drinking water, it was witnessed in the field as a common practice that water from the well was poured on a common water trough for animals from different herds to drink, these attributes are likely to cause high transmission rate of *Brucella* and hence high prevalence.

The current study found a seroprevalence of 15.5% in Njombe region, a previous study in Njombe also found a seroprevalence of 0.3% in dairy cattle (Mathew *et al.,* 2017) suggesting that brucellosis is on increasing trend. In Tanzania brucellosis in dairy cattle has been reported in other regions, in Morogoro region a seroprevalence of 21.3% has been reported (Shirima *et al.,* 2018) which is higher than that of Njombe and in contrast to lower seroprevalences ranging from 0.6 to 4.2% found in Iringa, Mbeya and Coast regions (Asakura *et al.,* 2018; Karimuribo *et al.,* 2007; Mdegela *et al.,* 2004; Sagamiko *et al.,* 2018). Such regional variations may be

attributed to the genuine outbreak, inadequate brucellosis surveillance systems, and lack of brucellosis testing schemes and farm management systems due to increased intensification of dairy production in urban and peri-urban areas. Therefore, these findings are calling for immediate attention and strategic mitigation measures against brucellosis in high-risk areas such as Njombe region.

The region-specific seroprevalences in smallholder dairy cattle for Kilimanjaro and Arusha were 2.5% and 0.31%, respectively. A similar study conducted in dairy cattle between 2013 and 2015 around Manyara, Arusha and Kilimanjaro regions found animal level seroprevalence of 0.01% (Bodenham *et al.*, 2021). The current study in Kilimanjaro and Arusha suggest that, brucellosis in smallholder dairy cattle is on increasing trend and therefore, further improvement and strengthening of control measures against brucellosis by both local authorities and farmers is required to avoid further spread of the disease.

The seroprevalence encountered in Mbeya region was contrary to earlier findings in dairy cattle where the seroprevalence ranged from 2.8% to 17.8% (Mathew *et al.*, 2017; Mfune, 2015; Sagamiko *et al.*, 2018). The low seroprevalence encountered might have been attributed to efficient implementation of control strategies following earlier studies or discrepancies in sample size, target population, sampling strategies, and testing schemes used.

Previous brucellosis seroprevalence studies in dairy cattle carried out in Iringa and Tanga reported relatively higher seroprevalence levels than the findings from this study (Karimuribo *et al.*, 2007; Swai & Schoonman, 2012). The previous studies potentially alarmed the public health concern and One-Health stakeholders to control the disease. This study found relatively lower seroprevalence levels in those regions suggesting that some control measures might be going on to control the disease in those regions. The presence of dairy industries that buy milk from smallholder dairy farmers in Tanga, Iringa, and Mbeya might be the contributing factor to the low seroprevalence of brucellosis in these regions as they request from farmers a negative brucellosis screening results of the animals before accepting the milk. In Njombe region, the dairy industry is led by smallholder farmers through a farmer's cooperative initiative with no restriction on the disease testing unlike in Tanga, Iringa and Mbeya.

4.2.2 Spatial clustering of brucellosis

The spatial clustering analysis results identified two significant clusters one in Kilimanjaro and the other in Njombe region. Furthermore, the age related seroprevalence did not suggests steady increase with age of animals in both Njombe and Kilimanjaro which suggest more of a local outbreak rather than endemic expansion. This outbreak situation requires further investigation and rolling out control measure to stop further spread of the disease.

4.2.3 Risk factors for increasing trend of brucellosis in smallholder dairy farming systems in selected regions of Tanzania

Different possible risk factors which could likely been associated with the increasing trend of the disease in dairy cattle kept under the smallholder farming systems have been studied and some were significantly associated with brucellosis at univariable and multivariable levels of analysis.

Farms that kept their own bull for breeding were significantly less likely to have seropositive cattle. This study revealed that keeping one's own bull was protective (OR<1) to *Brucella* infection. This finding was also reported by other research studies, indicating that using one's own bulls for breeding was more protective against brucellosis (Cárdenas *et al.,* 2019). This study revealed that, 47.1% of dairy farms were hiring bulls for breeding and 52.9% of dairy farms were not. This study found that hiring bull was not significantly associated with brucellosis seropositivity. The use of bulls from certified brucellosis free herds or brucellosis free bulls for breeding is encouraged as a measure to control brucellosis in smallholder dairy farming in areas where artificial insemination is not practiced.

The source of drinking water for cattle has been significantly associated with bovine brucellosis seropositivity. In the current study, 63.8% of farmers reported using tap water, 16.5% used river water, and 19.7% used well water as a source of drinking water for their cattle. Of the 49 seropositive animals, 27 come from tap water and 19 come from well water. The study found that well water was significantly associated to brucellosis seropositivity compared to other sources, with dairy cattle drinking well water being more likely to contract brucellosis. This finding is similar to the results of other studies where cattle were provided with unsafe water (Coelho *et al.,* 2007) or surface drinking water (Swai & Schoonman, 2010). Surface and open well water are prone to various sources of contaminations such as dung, urine, milk, placenta,

and uterine discharges, making them potential mediums for brucellosis transmission to cattle when used as the source of drinking water. Therefore, it is crucial to provide dairy cattle with clean and safe water.

Distance between herds was found to be significantly associated with brucellosis seropositivity. In this study, seventy-three percent (73.76%) of sampled cattle came from dairy farms located at a distance of less than 100 m between herds. Significant association was found between distance and brucellosis seropositivity ($p<0.05$) and dairy cattle in farms at a distance of less than 100 m between dairy farms were less likely of being brucellosis seropositive (OR<1). This finding is contrary to other research studies which found that cattle in closer herds were more likely to be brucellosis seropositive than cattle in distant herds (Omer *et al.,* 2000; Richey & Harrell, 1997; Soomro *et al.,* 2014). The current finding is probably due to the nature of zero-grazing, milkers are largely family members (54%) and few seropositive animals. An increase in the proximity of dairy farms in urban and peri-urban areas is due to a rapid increase in milk demand in urban settings (Njombe *et al.,* 2011).

Dairy cattle kept with or in contact with other domestic animals have been significantly associated with brucellosis seropositivity. Domestic animals like dogs, goats, and pigs were investigated and their statistical association with brucellosis seropositivity revealed. The current study revealed that, 88.4% of dairy farms kept dogs and dairy cattle kept on a farm that keep dog was more likely (OR>1) of being brucellosis seropositive than those on a farm which do not keep dog. This finding is consistent with the finding of other studies which found that the odds of disease in dairy cattle kept on a farm which keeps dog was 2.55 (Shome *et al.,* 2023). Although this study found that 83% of dairy farms practice good management of aborted material including placenta, dogs in a dairy farms eat aborted fetus and placenta and may get infected if contaminated. Dogs infected with *Brucella* may later on infect cattle (The Center for Food Security and Public Health [CFSPH], 2018).

The current study also revealed that, 66.8% of dairy farms keep goats and goats were found to be significantly associated with brucellosis seropositivity ($p<0.05$). Furthermore, dairy cattle on a farm which keeps goats was found to be more likely of being *Brucella* seropositive compared to those in a farm that does not keep goats (OR>1). This finding is consistent with other studies findings which found that cattle reared together with goats were 8.9 times more likely of contracting brucellosis (Anka *et al.,* 2014), this was further supported by other studies (Behera *et al.,* 2020; Calistri *et al.,* 2013; Ducrotoy *et al.,* 2017).

The current study found that, 53.2% of dairy farms had pigs around and only 46.8% did not have pigs around and pigs were found to be significantly associated with brucellosis seropositivity ($p<0.01$). Dairy cattle kept on a farm which had pigs around were more likely of being *Brucella* seropositive (OR>1) compared to those kept in a farm that does not keep pigs. Even though there was no evidence in the data, pigs are the natural host of *B. suis* which can infect and cause brucellosis in cattle (Tulu, 2022).

The good body condition score (>3 score) of dairy cattle was found to be significantly associated with brucellosis seropositivity ($p<0.05$). Furthermore, good body condition of dairy cattle was found to be protective to being seropositive (OR<1). The finding was consistent with the results of the previous studies which found higher *Brucella* infection in animals with poor body condition scores (Tsegay *et al.,* 2015). However, the finding was contrary to the findings of other research studies which found no association between body condition scores and brucellosis seropositivity (Awah-Ndukum *et al.,* 2018; Islam *et al.,* 2021; Makita *et al.,* 2011). However, unless animals were not well managed that compromise body immunity brucellosis may infect all animals irrespective of their body condition.

The multivariable mixed effect model results found having goats around and history of abortion, were significantly associated with brucellosis seropositivity ($p<0.05$). Keeping goats or having goats around dairy cattle was significantly associated with brucellosis seropositivity ($p<0.05$). The majority of cattle herds (67%) were found to keep goats where 84% of the seropositive cattle came from herds mixed with goats. Dairy cattle on farm with goats were more likely of being *Brucella* seropositive (OR>1) than those on farms with no goats. This is further proven by the fact that, the largest number of *B. melitensis* positive farms (33.3%) were found in Kilimanjaro where dairy cattle are kept together with goats and pastures are cut from communal land where cattle and goats graze. This finding is consistent with the findings of other studies which found that cattle kept with goats around are at higher risk of contracting *Brucella* infection (Anka *et al.,* 2014; Asmare *et al.,* 2013; Cárdenas *et al.,* 2019). Furthermore, it has been reported that interaction of cattle with small ruminants was associated with brucellosis reemergence in cattle in SSA (Ducrotoy *et al.,* 2017). Therefore, education for dairy farmers is mandatory to encourage them not to keep dairy cattle together with small ruminants or to ensure that small ruminants are free from *Brucella* infections.

Herds with history of abortion were significantly associated with brucellosis seropositivity ($p<0.05$) compared to those without history of abortion, this suggests that dairy cattle in a farm

with a history of abortion were more likely to be *Brucella* seropositive (OR>1) than those in a farm without a history of abortion. This finding is consistent with other studies' findings in dairy cattle (Mfune *et al.,* 2021; Segwagwe *et al.,* 2018; Yanti *et al.,* 2021) where history of abortion of the herd was significantly associated with brucellosis seropositivity. Although abortion is not a pathognomonic sign for brucellosis it should be considered in regions where the disease is endemic.

4.2.4 Brucella species circulating in smallholder dairy cattle farming system in the study regions of Tanzania

The current study identified *Brucella* at genus and species level by using blood and vaginal swab samples. The study identified both *B. abortus* and *B. melitensis* circulating in this dairy cattle population and the unidentified *Brucella* species. The current study report for the first time that dairy cattle in Tanzania are infected predominantly by *B. melitensis*, a *Brucella* species generally responsible for goat brucellosis. Dairy cattle being predominantly infected with *B. melitensis* can be explained by the fact that most dairy farms (67%) were keeping goats and it has been revealed that dairy cattle kept on a farm that keep goats or has goats around were more likely of being *Brucella* seropositive when compared to dairy cattle on a farm that do not keep goats (Mengele *et al.,* 2023b). However, the finding of this study was contrary to the findings of other studies where *B. abortus* was found to be more predominant than *B. melitensis* (Aliyev *et al.,* 2022; Njeru *et al.,* 2022).

This study also reports for the first time that dairy cattle in smallholder farming systems in Tanzania are naturally co-infected by *B. abortus* and *B. melitensis* (mixed infection). This finding was in agreement with the findings from other studies in neighboring country Rwanda and other African countries (Abnaroodheleh *et al.,* 2023; Aliyev *et al.,* 2022; Kolo *et al.,* 2019; Mitterran *et al.,* 2020; Ntivuguruzwa *et al.,* 2022) which also identified co-infected cattle with *B. abortus* and *B. melitensis*. Other studies revealed that, *Brucella* species co-infection has been attributed to the interaction of cattle with other domestic animals (Kolo *et al.,* 2019; Morales-Estrada *et al.,* 2016; Morales-Estrada *et al.,* 2012). Furthermore, the interaction of cattle with small ruminants has been associated with increased trend of brucellosis in dairy cattle (Ducrotoy *et al.,* 2017; Mengele *et al.,* 2023b).

It was not within the scope of this study to identify other *Brucella* species than *B. abortus* and *B. melitensis*. The presence of other domestic animals such as sheep, dogs and pigs together

with cattle highlights the potential of infection with other *Brucella* species such as *B. ovis, B. canis* and *B. suis* but can also be *B. neotomae* (Motto *et al.,* 2018). The presence of other *Brucella* species in dairy cattle has also been reported by other studies, *Brucella* species such as *B. suis* and *B. canis* have already been identified in dairy cattle following natural infection by other studies in other countries (Baek *et al.,* 2012; Ewalt *et al.,* 1997; Fretin *et al.,* 2013). The possibility of other *Brucella* species circulating in dairy cattle may be a result of their interaction with other domestic animals and cross-species infection potential of *Brucella* species.

It has to be acknowledged that, the PCR results presented by this study were likely be affected by the fact PCR and DNA extraction assays were not perfect such that might have affected the DNA quality, and the total number of positive animals. Furthermore, the LOD of the assays might have left out the undetermined species. Both DNA extraction and PCR assay were done in a quality assurance and controlled environment to avoid possible contamination and misinterpretation of the results, however contamination of the DNA samples do occurs in the laboratory (Aslanzadeh, 2004) and gives unexpected result.

This study also report that, there was poor agreement between cELISA and PCR results. This discrepancy could have been attributed to the long-term storage of samples, which were kept in a deep freezer at a very low temperature (−20°C) for over a year because of delayed laboratory reagents and consumables due to the lockdown resulted after COVID 19 pandemic; these conditions are likely to have degraded the samples. Poor agreement could have as well been attributed to individual animals' varying bacteremic and immunologic phases and abortion status (Maksimović *et al.,* 2022). This finding is similar to other studies in cattle which found poor agreement between serological and PCR results in different samples from the same animal (El-Diasty *et al.,* 2018). Therefore, to provide the highest disease detection rate, it is important to run the serological and molecular detection methods using different samples and therefore reduce the number of false-negative results.

Brucella species are considered host-specific, however, studies over the recent years have shown that cross-species infections are not uncommon in domestic animals. The detection of *B. melitensis* in dairy cattle represents a paradigm shift in disease control, necessitating different strategies for both human and livestock in One-Health approach. Furthermore, human cases diagnosed with the most virulent *B. melitensis* in Tanzania (Bodenham *et al.,* 2020), may be hypothesized to originate from cattle, in addition to small ruminants. Controlling the disease

by vaccination using the monovalent *B. abortus* S19 vaccine in dairy cattle population might be challenging, as the vaccine *B. abortus* S19 vaccine does not provide cross-immunity. This challenge may be further complicated in individuals with mixed infections.

The *B. abortus* S19 vaccine has been developed and validated for cattle. The current finding on involvement of small ruminants particularly goats in the epidemiology of brucellosis in dairy cattle require validation of the vaccine for the use in goats, otherwise *B. melitensis* Rev.1 vaccine should be used to control the disease in small ruminants. Controlling brucellosis in cattle should now go in hand with controlling the disease in goats. Surveillance studies should also focus on knowing the disease burden not only in cattle but also in other domesticated animals such as goats. Therefore, education to dairy farmers and livestock extension officers is mandatory, for farmers to encourage them not to keep dairy cattle together with small ruminants might not be practical but encouraging them to regularly check on brucellosis status so that to have brucellosis-free small ruminants and dairy cattle is practical and might be the way forward to control the disease in dairy herds.

To control brucellosis by vaccination in dairy cattle, the development of a bivalent vaccine containing both *B. abortus* and *B. melitensis* strains should be a way forward in SSA. Studies show that *B. abortus* S19 is protective against *B. melitensis* in naturally infected cattle (van Straten *et al.,* 2016) but no information about its effectiveness in cases of mixed infection or its effectiveness in controlling the disease in goats. The *B. melitensis* Rev. 1 vaccine for goats is available but has not been used in Tanzania, and therefore strategies must be in place for the government or private sector to start producing the vaccine locally so that it can be available at affordable prices to farmers. There is no evidence on effectiveness of *B. melitensis* Rev 1 in cattle and therefore should not be used to vaccinate cattle (WOAH, 2022).

4.2.5 Genome-wide association studies analysis to identify SNP markers associated with brucellosis resistance/susceptibility in dairy cattle

The potential of genetic resistance to infectious diseases has not been given attention in genetic improvement programs to improve cattle health in addition to good management practices in controlling infectious diseases (Nash & Freeman, 2004). It has been found that, about 3000 genes in mammals control natural host immunity and hence resistance against infectious organisms (Breuer *et al.,* 2013; Hasenauer *et al.,* 2013). In cattle, it has been known that resistance to brucellosis is genetically determined and heritable (Borriello *et al.,* 2006;

Thompson-Crispi *et al.,* 2012). Furthermore, later studies found that the NRAMP1 gene was associated with resistance to brucellosis, however, it was later refuted after it failed on biological validation (Feng *et al.,* 1996; Paixão *et al.,* 2007). Based on single-strand conformation polymorphism analysis (SSCP) and Genome Wide Association Studies (GWAS), science found a strong relationship between gene variation and natural resistance to brucellosis (Banos *et al.,* 2017; Barthel *et al.,* 2001; Hasenauer *et al.,* 2013). Genome-wide association study (GWAS) is the latest and best tool evolved to find out the association of complex diseases and genetic markers (variants) of an individual in a population (Bush & Moore, 2012; Riancho, 2012).

In this study, the heritability component in terms of additive genetic variance of dairy cattle to brucellosis was found to be zero, suggesting that there was no enough evidence in the data that the observed variations of brucellosis PCR status between individual cattle was due to genetic variation and rather due to environmental difference or stochastic variation. However, the result was also influenced by the low statistical power due to the fewer number of cases (36 disease positive) compared to the controls (1941 disease negative) and 14 explanatory variable. The few number of PCR positives were likely to cause model over-fitting and hence inaccurate prediction. The zero heritability finding was contrary to other studies which found the heritability due to additive genetic variance of dairy cattle to brucellosis between 0.25-0.5, similar to other productive traits (Martínez *et al.,* 2010; Thompson-Crispi *et al.,* 2012).

Association analyses are generally conducted in a homogenous population and not genetically distant. To satisfy the condition before the GWAS analysis was done, the principal component analysis was conducted and principal components 1 and 2 were used to draw a plot that justified that the population was rather homogenous (Fig. 37). A homogenous population reflects the small genetic distance between individuals under the study. Most of the dairy cattle in Tanzania are crossbred of a local breed (TSHZ) with the exotic dairy breeds mainly of Friesians, Ayrshire, and Jersey lines. The central cluster in Fig. 37 suggests a homogenous population and few in the periphery suggest little genetic distance due to the exotic nature of their genetic makeup. Generally, dairy cattle in the study population were homogenous with little genetic distance between them and therefore, fit for genome-wide association analyses.

The GWAS analysis revealed that the SNP marker BovineHD0900011750 with minor and major alleles A/G, respectively located in chromosome number 9 was most significantly associated with brucellosis and was found in 0.056 (5.6%) of the study individuals.

Statistically, the SNP marker measure of effect in the cattle genome was found to be associated with the risk (brucellosis susceptibility) as it had a beta coefficient of 5.46E-02 (Odds ratio of 1.05). The variant effect predictor analysis showed that, the SNP marker was found within a protein-coding region called AFG1L (Fig. 40) and it's impact is considered as a modifier as the direction of impact is not yet known. The AFG1L gene is responsible for the production of mitochondrial integral membrane protein which is responsible for mitochondrial protein homeostasis. The AFG1L gene has been associated with susceptibility to brucellosis, however previous studies found that the NRAMP1 gene was associated with the natural resistance of brucellosis in cattle (Feng *et al.,* 1996).

Five SNP markers were found above the suggestive line, as they have high LD to genome-wide significant maker BovineHD0900011750. Two markers in chromosome 9 (BovineHD0900011751 and BovineHD0900011748), one marker in chromosome 8 (BovineHD0800029026), one marker in chromosome 26 (BovineHD2600004903), and one marker in chromosome 28 (BovineHD2800000058). The GWAS also identify SNPs (Alleles) which exhibit non-random association (LD) which are normally co-inherited and work together with the most significant SNP marker (core SNP) to affect the additive genetic variance to the trait of interest (Montgomery, 2008).

The AFG1L gene was demonstrated to be associated with brucellosis susceptibility. Further studies need to be conducted to determine if individuals with the GWAS significant SNP marker are more susceptible to brucellosis than those with no significant marker. Initially population wide studies might be conducted to detect the marker in the population and increase the sample size for GWAS or linkage studies. Secondly experimental studies in vivo might be conducted to show the relationship between the presence of this SNP in cattle and increased susceptibility to brucellosis like how it was done in other disease (Wragg *et al.,* 2022). Finally, functional assays could investigate the biological mechanism of these genotypes-phenotype association to approve or refute the finding just like how was done in other studies (Barthel *et al.,* 2001; Horˇ ín *et al.,* 1999; Paixão *et al.,* 2007).

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

The findings from the current study on brucellosis in smallholder dairy cattle in Njombe, Kilimanjaro, and Tanga regions of Tanzania confirmed that, brucellosis in smallholder dairy cattle was present during the study period, with an overall animal level prevalence of 2.4%. Among the studied regions, Njombe and Kilimanjaro regions were the hotspots of brucellosis in dairy cattle and, Njombe region was a region with emerging smallholder dairy farming with the highest prevalence of 15.5%. The current study found out that, 66.8% of dairy farms kept goats and mixed effect model analysis showed interaction between dairy cattle and goats whereby keeping goats or having goats around with dairy cattle was a significant risk factor attributable for increased trend of brucellosis in dairy cattle in smallholder farming system. Moreover, the current study revealed for the first time that *B. melitensis spp* are circulating in smallholder dairy population in Tanzania with also existence of co-infection of *B. abortus* and *B. melitensis* in dairy cattle and *B. melitensis* infection was predominant infection. Presence of *B. melitensis* and *B. abortus* in dairy cattle is probably another cause of increasing trend of brucellosis in dairy cattle. In addition, there was likely potential for other *Brucella* species such as *B. suis*, *B. ovis* and *B. canis* circulating in dairy cattle population. The GWAS analysis of the current study confirmed that, the SNP marker BovineHD0900011750 found in a AFG1L gene was associated with brucellosis susceptibility in dairy cattle and there was no enough evidence in the data to show that the observed phenotypic trait (brucellosis PCR results) variation in the study population was due to genetics but was rather due to environmental or stochastic variations. This study had limitations especially in genomics part. The main challenge was to precisely identify animals with phenotype (diseased animals) for disease resistance gene with high predictive values for the genetic markers. The challenge was associated with the inherent nature of a cross-sectional study design. The expert opinion was to use the PCR results as a criteria for case (PCR positive) and control (PCR negative) rather than using serology (cELISA) results. Nevertheless, the study was expecting to get a prevalence of at least 5% in the study population, however the prevalence was less than 5% in both serology and PCR results. This significantly affected the statistical power of analysis. To have a good statistical power for heritability estimation and GWAS, more than 5% of cases should have sufficed with the available sample size. For future studies, GWAS based on cross-

sectional study design, PCR results should be used as basis for case and control definitions and PCR prevalence should be more than 5% to have best statistical power of analysis and study population can change to different production system, domestic species and study areas. Another limitation was that, the study only targeted the burden of *Brucella* infection among cattle in smallholder farming while omitting studying the brucellosis burden among the surround smallholder farmers and milk consumers in the studied regions. This is a significant study design omission, especially when addressing *Brucella* infection under the One Health approach.

5.2 Recommendations

Presence of brucellosis in dairy cattle is posing a threat to public health and dairy cattle production, the situation is worse in most prevalent regions such as Njombe and Kilimanjaro and rapid strategic responses to control the disease in dairy cattle and small ruminants is appealing in Njombe and Kilimanjaro regions. Furthermore, the disease status among dairy farmers and the general public is equally important to be known and addressed in One-Health approach.

This study also recommends further studies on *Brucella* species circulating in dairy cattle and epidemiological studies on the roles of small ruminants to brucellosis in dairy cattle. Finally, the study is recommending immediate development and studies on cost benefit analysis on the use of bivalent *Brucella* vaccine (for *B. abortus* and *B. melitensis*) and or local production of *B. melitensis* Rev. 1 vaccine to control brucellosis in cattle and small ruminants respectively. Further studies should also focus on biological validation of the BovineHD0900011750 SNP marker on AFG1L gene which is associated with brucellosis susceptibility in dairy cattle. This study was conducted in dairy cattle, similar study can also be done in cattle under extensive farming in agro and pastoral communities where brucellosis may be more prevalent.

REFERENCES

Abnaroodheleh, F., Emadi, A., Dashtipour, S., Jamil, T., Khaneghah, A. M., & Dadar, M. (2023). Shedding rate of *Brucella* spp. in the milk of seropositive and seronegative dairy cattle. *Heliyon*, *9*(4), 1-8.

Adams, L., & Templeton, J. (1998). Genetic resistance to bacterial diseases of animals. *Revue Scientifique et Technique-Office International des Epizooties*, *17*(1), 200-219.

Adams, L. G. (1990). *Development of Live Brucella Vaccines*. Advances in Brucellosis Research, Texas A & M University Press, College Station. https://scholar.google.com

Akakpo, A. J., Têko-Agbo, A., & Koné, P. (2010). *The Impact of Brucellosis on the Economy and Public Health in Africa*. https://scholar.google.com

Akoko, J. M., Pelle, R., Lukambagire, A. S., Machuka, E. M., Nthiwa, D., Mathew, C., Fèvre, E. M., Bett, B., Cook, E. A., Othero, D., & Bonfoh, B. (2021). Molecular epidemiology of Brucella species in mixed livestock-human ecosystems in Kenya. *Scientific Reports*, 11(1), 8881.

Al Dahouk, S., Nöckler, K., Scholz, H. C., Pfeffer, M., Neubauer, H., & Tomaso, H. (2007). Evaluation of genus-specific and species-specific real-time PCR assays for the identification of *Brucella* spp. *Clinical Chemistry and Laboratory Medicine*, *45*(11), 1464–1470. https://doi.org/10.1515/CCLM.2007.305

Albertian, M., Fedorov, A., & Iskandarov, M. (2006). *Brucellosis Vaccines: Past, Present and Future.* https://scholar.google.com

Aliloo, H., Mrode, R., Okeyo, A., Ni, G., Goddard, M., & Gibson, J. P. (2018). The feasibility of using low-density marker panels for genotype imputation and genomic prediction of crossbred dairy cattle of East Africa. *Journal of Dairy Science*, *101*(10), 9108-9127.

Aliyev, J., Alakbarova, M., Garayusifova, A., Omarov, A., Aliyeva, S., Fretin, D., & Godfroid, J. (2022). Identification and molecular characterization of *Brucella abortus* and *Brucella melitensis* isolated from milk in cattle in Azerbaijan. *BMC Veterinary Research*, *18*(1), 71.

Allan, G., Chappel, R., Williamson, P., & McNaught, D. J. E. (1976). A quantitative comparison of the sensitivity of serological tests for bovine brucellosis to different antibody classes. *Epidemiology and Infection*, *76*(2), 287-298.

Alton, G. G., Jones, L. M., Angus, R. D., & Verger, J. M. (1988). *Techniques for the Brucellosis Laboratory, Institut National de la Recherche Agonomique*. Paris, France.

Altshuler, D., Daly, M. J., & Lander, E. S. (2008). Genetic mapping in human disease. *Science*, *322*(5903), 881-888.

Altshuler, D., Gibbs, R., Peltonen, L., Altshuler, D., & , G. R. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature*, *467*, 52–58.

Anka, M. S., Hassan, L., Khairani-Bejo, S., Zainal, M. A., Mohamad, R. b., Salleh, A., & Adzhar, A. (2014). A case-control study of risk factors for bovine brucellosis seropositivity in Peninsular Malaysia. *PloS One*, *9*(9), e108673.

Aragon, T. J., Fay, M. P., Wollschlaeger, D., Omidpanah, A., & Omidpanah, M. A. (2017). *Package 'epitools'*. https://scholar.google.com

Asakura, S., Makingi, G., Kazwala, R., & Makita, K. (2018). Brucellosis risk in urban and agro-pastoral areas in Tanzania. *EcoHealth*, *15*, 41-51.

Ashford, D. A., Di Pietra, J., Lingappa, J., Woods, C., Noll, H., Neville, B., Weyant, R., Bragg, S. L., Spiegel, R. A., Tappero, J., & Perkins, B. A. (2004). Adverse events in humans associated with accidental exposure to the livestock brucellosis vaccine RB51. *Vaccine*, 22(25-26), 3435-3439.

Aslanzadeh, J. (2004). Preventing PCR amplification carryover contamination in a clinical laboratory. *Annals of Clinical Laboratory Science*, *34*(4), 389-396.

Asmare, K., Sibhat, B., Molla, W., Ayelet, G., Shiferaw, J., Martin, A. D., Skjerve, E., & Godfroid, J. (2013). The status of bovine brucellosis in Ethiopia with special emphasis on exotic and cross bred cattle in dairy and breeding farms. *Acta Tropica*, 126(3), 186-192.

Assenga, J. A., Matemba, L. E., Muller, S. K., Malakalinga, J. J., & Kazwala, R. R. (2015). Epidemiology of Brucella infection in the human, livestock and wildlife interface in the Katavi-Rukwa ecosystem, Tanzania. *BMC Veterinary Research*, 11, 1-11.

Avila-Calderón, E. D., Lopez-Merino, A., Sriranganathan, N., Boyle, S. M., & Contreras-Rodríguez, A. (2013). A history of the development of *Brucella* vaccines. *BioMed Research International*, *2013*, 1-8.

Awah-Ndukum, J., Mouiche, M. M. M., Bayang, H. N., Ngwa, V. N., Assana, E., Feussom, K. J. M., Manchang, T. K., & Zoli, P. A. (2018). Seroprevalence and associated risk factors of brucellosis among indigenous cattle in the Adamawa and north regions of Cameroon. *Veterinary Medicine International*, 2018(1), 3468596.

Baek, B. K., Park, M. Y., Islam, M. A., Khatun, M. M., Lee, S. I., & Boyle, S. M. (2012). The first detection of Brucella canis in cattle in the Republic of Korea. *Zoonoses and Public Health*, 59(2), 77-82.

Banos, G., Winters, M., Mrode, R., Mitchell, A., Bishop, S., Woolliams, J., & Coffey, M. (2017). Genetic evaluation for bovine tuberculosis resistance in dairy cattle. *Journal of Dairy Science*, *100*(2), 1272-1281. https://doi.org/10.3168/jds.2016-11897

Barquero-Calvo, E., Chaves-Olarte, E., Weiss, D. S., Guzmán-Verri, C., Chacón-Díaz, C., Rucavado, A., & Moriyón, I., Moreno, Edgardo (2007). *Brucella abortus* uses a stealthy strategy to avoid activation of the innate immune system during the onset of infection. *PLoS One*, *2(7)*, e631.

Barthel, R., Feng, J., Piedrahita, J. A., McMurray, D. N., Templeton, J. W., & Adams, L. G. (2001). Stable transfection of the bovine NRAMP1 gene into murine RAW264. 7 cells: Effect on *Brucella abortus* survival. *Infection and Immunity*, *69*(5), 3110-3119.

Barton, K. (2022). *MuMIn: Multi-Model Inference. R package version 1.46.0*. https://CRAN.R-project.org/package=MuMIn.

Bates, D. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-51.

Beh, K. (1974). Quantitative distribution of Brucella antibody amongst immunoglobulin classes in vaccinated and infected cattle. *Research in Veterinary Science*, *17*(1), 1-4.

Behera, S. K., Das, D., Balasubramani, K., Chellappan, S., Rajaram, K., Mohanta, H. K., & Nina, P. B. (2020). Seroprevalence and risk factors of brucellosis in livestock in the wildlife and livestock interface area of Similipal Biosphere Reserve, India. *Veterinary World*, *13*(3), 465.

Bernard, F., Vincent, C., Matthieu, L., David, R., & James, D. (2005). Tuberculosis and brucellosis prevalence survey on dairy cattle in Mbarara milk basin (Uganda). *Preventive Veterinary Medicine*, *67*(4), 267-281.

Berry, D. P., Bermingham, M. L., Good, M., & More, S. J. (2011). Genetics of animal health and disease in cattle. *Irish Veterinary Journal*, *64*(1), 1-10.

Bertram-Drogatz, P., Quester, I., Becker, A., & Pühler, A. (1998). The Sinorhizobium meliloti MucR protein, which is essential for the production of high-molecular-weight succinoglycan exopolysaccharide, binds to short DNA regions upstream of exoH and exoY. *Molecular and General Genetics*, *257*(4), 433-441.

Bilodeau, S. (2022, 20/06/2022). *The qPCR Standard Curve: The Key to Good qPCR Data*. https://bitesizebio.com/31470/obligate-qpcr-standard-curve/

Bleidorn, C. (2017). *Genomes. In Phylogenomics (pp. 1-20).* Springer International Publishing AG, Gewerbestrasse 11, 6330 Cham, Switzerland.

Mbwambo, G. A., Van Zwetselaar, M., Sonda, T., Lukambagire, A. S., Njau, J. S., Wadugu, B., Ignass, I. P., Amani, N. B., Hugho, E. A., Rubach, M. P., & Sakasaka, P. (2024). Complete genome sequence of Brucella abortus isolated from a human blood culture sample in Tanzania. *Microbiology Resource Announcements*, 13(2), e00930-23.

Bodenham, R. F., Mazeri, S., Cleaveland, S., Crump, J. A., Fasina, F. O ., De Glanville, W. A., Haydon, D. T., Kazwala, R. R., Kibona, T. J., Maro, V. P., & Maze, M. J. (2021). Latent class evaluation of the performance of serological tests for exposure to *Brucella* spp. in cattle, sheep, and goats in Tanzania. *PLoS Neglected Tropical Diseases*, *15*(8), e0009630.

Borriello, G., Capparelli, R., Bianco, M., Fenizia, D., Alfano, F., Capuano, F., Ercolini, D., Parisi, A., Roperto, S., & Iannelli, D. (2006). Genetic resistance to *Brucella abortus* in the water buffalo (*Bubalus bubalis*). *Infection and Immunity, 74*(4), 2115-2120.

Boschiroli, M. L., Ouahrani-Bettache, S., Foulongne, V., Michaux-Charachon, S., Bourg, G., Allardet-Servent, A., O'Callaghan, D. (2002). The *Brucella suis* virB operon is induced intracellularly in macrophages. *Proceedings of the National Academy of Sciences, 99*(3), 1544-1549.

Breuer, K., Foroushani, A. K., Laird, M. R., Chen, C., Sribnaia, A., Lo, R., Lynn, D. J. (2013). InnateDB: Systems biology of innate immunity and beyond recent updates and continuing curation. *Nucleic Acids Research, 41*(D1), D1228-D1233.

Bricker, B. J. (2002). PCR as a diagnostic tool for brucellosis. *Veterinary Microbiology, 90*(1-4), 435-446.

Bricker, B. J., Ewalt, D. R., & Halling, S. M. (2003). Brucella'HOOF-Prints': Strain typing by multi-locus analysis of variable number tandem repeats (VNTRs). *BMC Microbiology, 3*(1), 1-13.

Bricker, B. J., & Halling, S. M. (1994). Differentiation of *Brucella abortus* bv. 1, 2, and 4, *Brucella melitensis, Brucella ovis,* and *Brucella suis* bv. 1 by PCR. *Journal of Clinical Microbiology, 32*(11), 2660-2666.

Brinley, W., & Corbel, M. (1990). *Brucella infections in man and animals: Contagious equine metritis. In Topley Wilson's Principles of Bacteriology, Virology and Immunology(8th Ed.).* Edward Arnold, London, England.

Briones, G., Iñón de Iannino, N., Roset, M., Vigliocco, A., Paulo, P. S., & Ugalde, R. A. (2001). Brucella abortus cyclic β-1, 2-glucan mutants have reduced virulence in mice and are defective in intracellular replication in HeLa cells. *Infection and Immunity, 69*(7), 4528-4535.

Bronsvoort, B. M. D., Koterwas, B., Land, F., Handel, I. G., Tucker, J., Morgan, K. L., Tanya, V. N., Abdoel, T. H., & Smits, H. L. (2009). Comparison of a flow assay for brucellosis antibodies with the reference cELISA test in West African Bos indicus. *PLoS One, 4*(4), e5221.

Browning, B. L. (2008). Rapid calculation of order statistic distributions and multiple-testing adjusted P-values via permutation for one and two-stage genetic association studies. *Bioinformatics, 9*, 1-5.

Bruce, D. (1887). *Note on the Discovery of a Microorganism in Malta Fever*. John Brigg. https://scholar.google.com

Buck, J. (1930). Studies of vaccination during calfhood to prevent bovine infectious abortion. *Journal of Agricultural Research, 41*(9), 667-689.

Bundle, D. R., Cherwonogrodzky, J. W., Caroff, M., & Perry, M. B. (1987, January). *The Lipopolysaccharides of Brucella abortus and B. melitensis. In Annales de l'Institut Pasteur/Microbiologie (Vol. 138, No. 1, pp. 92-98).* Elsevier Masson.

Bush, W. S., & Moore, J. H. (2012). Genome-wide association studies. *PLoS Computational Biology, 8*(12), e1002822.

Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., & Vandesompele, J. (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clinical Chemistry, 55*(4), 611-622.

Butler, J., Seawright, G., McGivern, P., & Gilsdorf, M. (1986). Preliminary evidence for a diagnostic immunoglobulin G1 antibody response among culture-positive cows vaccinated with *Brucella abortus* strain 19 and challenge exposed with strain 2308. *American Journal of Veterinary Research*, *47*(6), 1258-1264.

Calistri, P., Iannetti, S., Atzeni, M., Di Bella, C., Schembri, P., & Giovannini, A. (2013). Risk factors for the persistence of bovine brucellosis in Sicily from 2008 to 2010. *Preventive Veterinary Medicine*, *110*(3-4), 329-334.

Cárdenas, L., Peña, M., Melo, O., & Casal, J. (2019). Risk factors for new bovine brucellosis infections in Colombian herds. *BMC Veterinary Research, 15*, 1-8.

Celli, J., de Chastellier, C., Franchini, D. M., Pizarro-Cerda, J., Moreno, E., & Gorvel, J. P. (2003). Brucella evades macrophage killing via VirB-dependent sustained interactions

with the endoplasmic reticulum. *The Journal of Experimental Medicine*, *198*(4), 545-556.

CFSPH. (2018). *Technical Factsheet Brucellosis. The Center for Food Security and Public Health. https://www.cfsph.iastate.edu/Factsheets*

Chagunda, M. G., Mapemba, J. P., Awah-Ndukum, J., Tebug, S. F., & Wiedemann, S. (2014). Risk, knowledge and preventive measures of smallholder dairy farmers in northern Malawi with regard to zoonotic brucellosis and bovine tuberculosis. *Onderstepoort Journal of Veterinary Research, 81*(1), 1-6. http://dx.doi.org/10.4102/ojvr.v81i1.594.

Chen, Y., Lin, C. H., & Sabatti, C. (2006). Volume measures for linkage disequilibrium. *BMC Genetics, 7*(1), 1-8.

Chukwu, C. (1985). Brucellosis in Africa. I. The prevalence. *Bulletin of Animal Health and Production in Africa, 33*(3), 193-198.

Cloeckaert, A., Verger, J. M., Grayon, M., & Grepinet, O. (1995). Restriction site polymorphism of the genes encoding the major 25 kDa and 36 kDa outer-membrane proteins of *Brucella Microbiology*, *141*(9), 2111-2121.

Cloeckaert, A., Verger, J. M., Grayon, M., Paquet, J. Y., Garin-Bastuji, B., Foster, G., & Godfroid, J. (2001). Classification of *Brucella* spp. isolated from marine mammals by DNA polymorphism at the omp2 locus. *Microbes and Infection, 3*(9), 729-738.

Coelho, A., Coelho, A., Roboredo, M., & Rodrigues, J. (2007). A case–control study of risk factors for brucellosis seropositivity in Portuguese small ruminants herds. *Preventive Veterinary Medicine, 82*(3-4), 291-301.

Comerci, D. J., Martínez-Lorenzo, M. J., Sieira, R., Gorvel, J. P., & Ugalde, R. A. (2001). Essential role of the VirB machinery in the maturation of the *Brucella abortus*-containing vacuole. *Cellular Microbiology, 3*(3), 159-168.

Conde-Alvarez, R., Arce-Gorvel, V., Iriarte, M., Manček-Keber, M., Barquero-Calvo, E., Palacios-Chaves, L., Chacon-Diaz, C., Chaves-Olarte, E., Martirosyan, A., von Bargen, K., & Grilló, M. J. (2012). The lipopolysaccharide core of Brucella abortus acts as a shield against innate immunity recognition. *PLoS Pathogens, 8*(5), e1002675.

Consortium, C. C. (2007). Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. Wellcome Trust. *Nature*, *447*(7145), 661.

Corbel, M. J. (2006). *Brucellosis in Humans and Animals. World Health Organization. Food and Agriculture Organization and World Organization for Animal Health*. https://scholar.google.com

Cupples, L. A., Arruda, H. T., Benjamin, E. J., D'Agostino, R. B., Demissie, S., DeStefano, A. L., Dupuis, J., Falls, K. M., Fox, C. S., Gottlieb, D. J., & Govindaraju, D. R. (2007). The Framingham Heart Study 100K SNP genome-wide association study resource: Overview of 17 phenotype working group reports. *BMC Medical Genetics*, *8*(1), 1-19.

Currò, V., Marineo, S., Vicari, D., Galuppo, L., Galluzzo, P., Nifosì, D., Pugliese, M., Migliazzo, A., Torina, A., & Caracappa, S. (2012). The isolation of *Brucella* spp. from sheep and goat farms in Sicily. *Small Ruminant Research, 106,* S2-S5.

Dauphin, L. A., Hutchins, R. J., Bost, L. A., & Bowen, M. D. (2009). Evaluation of automated and manual commercial DNA extraction methods for recovery of *Brucella* DNA from suspensions and spiked swabs. *Journal of Clinical Microbiology*, *47*(12), 3920-3926.

De Glanville, W. A., Conde-Alvarez, R., Moriyon, I., Njeru, J., Diaz, R., Cook, E. A., Morin, M., Bronsvoort, B. M. D. C., Thomas, L. F., Kariuki, S., & Fevre, E. M. (2017). Poor performance of the rapid test for human brucellosis in health facilities in Kenya. *PLoS Neglected Tropical Diseases*, *11*(4), e0005508.

De Tejada, G. M., Pizarro-Cerda, J., Moreno, E., & Moriyon, I. (1995). The outer membranes of *Brucella* spp. are resistant to bactericidal cationic peptides. *Infection and Immunity*, *63*(8), 3054-3061.

Dean, A. S., Crump, L., Greter, H., Schelling, E., & Zinsstag, J. (2012). Global burden of human brucellosis: A systematic review of disease frequency. *PLoS Neglected Tropical Diseases*, *6*(10), e1865.

Del Giudice, M. G., Ugalde, J. E., & Czibener, C. (2013). A lysozyme-like protein in Brucella abortus is involved in the early stages of intracellular replication. *Infection and Immunity*, *81*(3), 956-964.

Grace, D., Mutua, F. K., Ochungo, P., Kruska, R. L., Jones, K., Brierley, L., Lapar, M. L., Said, M. Y., Herrero, M. T., Phuc, P. M., & Thao, N. B. (2012). *Mapping of Poverty and Likely Zoonoses Hotspots.* https://scholar.google.com

Dermine, J. F., Duclos, S., Garin, J., St-Louis, F., Rea, S., Parton, R. G., & Desjardins, M. (2001). Flotillin-1-enriched lipid raft domains accumulate on maturing phagosomes. *Journal of Biological Chemistry*, *276*(21), 18507-18512.

Development, Ministry of Livestock and Fisheries (2016). *Tanzania Livestock Modernization Initiative.* The Government of United Republic of Tanzania. https:// scholar. google. com

Diaz, R., Garatea, P., Jones, L. M., & Moriyon, I. (1979). Radial immunodiffusion test with a Brucella polysaccharide antigen for differentiating infected from vaccinated cattle. *Journal of Clinical Microbiology*, *10*(1), 37-41.

Döhmer, P. H., Valguarnera, E., Czibener, C., & Ugalde, J. E. (2014). Identification of a type IV secretion substrate of *Brucella abortus* that participates in the early stages of intracellular survival. *Cellular Microbiology*, *16*(3), 396-410.

Dohoo, I., Martin, W., & Stryhn, H. (2014). *Sampling. In S. M. McPike (Ed.), Veterianary Epidemiologic Research (2nd Ed.).* AVC Inc., University of Prince Edward Island, 550 University Avenue, Charlottetown, Prince Edward Island, Canada.

Donnelly, P. (2008). Progress and challenges in genome-wide association studies in humans. *Nature*, *456*(7223), 728-731.

Ducrotoy, M., Bertu, W. J., Matope, G., Cadmus, S., Conde-Álvarez, R., Gusi, A. M., Welburn, S., Ocholi, R., Blasco, J. M., & Moriyón, I. (2017). Brucellosis in Sub-Saharan Africa: Current challenges for management, diagnosis and control. *Acta Tropica*, *165*, 179-193.

Ducrotoy, M. J., Bertu, W. J., Ocholi, R. A., Gusi, A. M., Bryssinckx, W., Welburn, S., & Moriyon, I. (2014). Brucellosis as an emerging threat in developing economies: Lessons from Nigeria. *PLoS Neglected Tropical Diseases*, *8*(7), e3008.

Eisenschenk, F. C., Houle, J. J., & Hoffmann, E. M. (1999). Mechanism of serum resistance among *Brucella abortus* isolates. *Veterinary Microbiology*, *68*(3-4), 235-244. https://doi.org/10.1016/S0378-1135(99)00075-9

El-Diasty, M., Wareth, G., Melzer, F., Mustafa, S., Sprague, L. D., & Neubauer, H. (2018). Isolation of *Brucella abortus* and *Brucella melitensis* from seronegative cows is a serious impediment in brucellosis control. *Veterinary Sciences*, *5*(1), 28.

Ewalt, D. R., Payeur, J. B., Rhyan, J. C., & Geer, P. L. (1997). Brucella suis biovar 1 in naturally infected cattle: A bacteriological, serological, and histological study. *Journal of Veterinary Diagnostic Investigation*, *9*(4), 417-420.

Fekete, A., Bantle, J., Halling, S. M., & Sanborn, M. (1990). Preliminary development of a diagnostic test for *Brucella* using polymerase chain reaction. *Journal of Applied Bacteriology*, *69*(2), 216-227.

Feng, J., Li, Y., Hashad, M., Schurr, E., Gros, P., Adams, L. G., & Templeton, J. W. (1996). Bovine natural resistance associated macrophage protein 1 (Nramp1) gene. *Genome Research*, *6*(10), 956-964.

Ferrero, M. C., Hielpos, M. S., Carvalho, N., Barrionuevo, P., Corsetti, P., Giambartolomei, G. H., Costa O. S., & Baldi, P. C. (2014). Key role of Toll-like receptor 2 in the inflammatory response and major histocompatibility complex class II downregulation in *Brucella abortus*-infected alveolar macrophages. *Infection and Immunity*, *82*(2), 626-639.

Florez, J. C., Manning, A. K., Dupuis, J., McAteer, J., Irenze, K., Gianniny, L., Mirel, D. B., Fox, C. S., Cupples, L. A., & Meigs, J. B. (2007). A 100K genome-wide association scan for diabetes and related traits in the Framingham Heart Study: Replication and integration with other genome-wide datasets. *Diabetes*, *56*(12), 3063-3074. https://doi.org/10.2337/db07-0451

Foster, J. T., Beckstrom-Sternberg, S. M., Pearson, T., Beckstrom-Sternberg, J. S., Chain, P. S., Roberto, F. F., Hnath, J., Brettin, T., & Keim, P. (2009). Whole-genome-based phylogeny and divergence of the genus Brucella. *Journal of Bacteriology*, *191*(8), 2864-2870.

Foulongne, V., Bourg, G., Cazevieille, C., Michaux-Charachon, S., & O'Callaghan, D. (2000). Identification of *Brucella suis* genes affecting intracellular survival in an in vitro human macrophage infection model by signature-tagged transposon mutagenesis. *Infection and Immunity*, *68*(3), 1297-1303. https://doi.org/10.1128/IAI.68.3.1297-1303.2000

Franc, K., Krecek, R., Häsler, B., & Arenas-Gamboa, A. (2018). Brucellosis remains a neglected disease in the developing world: A call for interdisciplinary action. *BMC Public Health*, *18*(1), 125.

Freeman, A., & Kehrli, M. (1994). *Improved Disease Resistance in Dairy Cattle: Correlation of Health Disorders with Measures of Immune Status; and Genetic Improvement for Disease Resistance by Identifying Sires whose Daughters have Fewer Disease Problems.* https://scholar.google.com

Fretin, D., Mori, M., Czaplicki, G., Quinet, C., Maquet, B., Godfroid, J., & Saegerman, C. (2013). Unexpected *Brucella suis* biovar 2 infection in a dairy cow, Belgium. *Emerging Infectious Diseases*, *19*(12), 2053.

Ganesh, N. V., Sadowska, J. M., Sarkar, S., Howells, L., McGiven, J., & Bundle, D. R. (2014). Molecular recognition of Brucella A and M antigens dissected by synthetic oligosaccharide glycoconjugates leads to a disaccharide diagnostic for brucellosis. *Journal of the American Chemical Society*, 136(46), 16260-16269.

García-Yoldi, D., Marín, C. M., De Miguel, M. J., Muñoz, P. M., Vizmanos, J. L., & López-Goñi, I. J. C. C. (2006). Multiplex PCR assay for the identification and differentiation of all *Brucella* species and the vaccine strains *Brucella abortus* S19 and RB51 and *Brucella melitensis* Rev1. *Clinical Chemistry*, *52*(4), 779-781.

Garry, A. L., & Schutta, C. J. (2010). Natural resistance against brucellosis: A review. *The Open Veterinary Science Journal*, *4*(1), 61-71. 10.2174/1874318801004010061.

Gicheru, M. N., Mwangi, E., & Mwangi, R. (2015). Prevalence and Knowledge of Brucellosis in Dairy Cattle in Makuyu Division, Murang'a County, Kenya. *International Journal of Scientific Engineering and Technology*, *4*(12), 549-555.

Gil-Ramírez, Y., Conde-Álvarez, R., Palacios-Chaves, L., Zúñiga-Ripa, A., Grilló, M. J., Arce-Gorvel, V., Hanniffy, S., Moriyón, I., & Iriarte, M. (2014). The identification of wadB,

a new glycosyltransferase gene, confirms the branched structure and the role in virulence of the lipopolysaccharide core of *Brucella abortus*. *Microbial Pathogenesis, 73*, 53-59.

Gilmour, A. R. (2015). *ASReml User Guide Release 4.1 Structural Specification*. VSN International Ltd. www.vsni.co.uk.

Goddard, M. E., & Hayes, B. J. (2009). Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews Genetics, 10*(6), 381-391.

Godfroid, J. (2002). Brucellosis in wildlife. *Revue Scientifique et Technique-Office International Des Épizooties*, 21(1), 277-286.

Godfroid, J., Al-Dahouk, S., Pappas, G., Roth, F., Matope, G., Muma, J., Marcotty, T., Pfeiffer, D., & Skjerve, E. (2013). A "One Health" surveillance and control of brucellosis in developing countries: Moving away from improvisation. *Comparative Immunology, Microbiology and Infectious Diseases, 36*(3), 241-248.

Godfroid, J., Cloeckaert, A., Liautard, J. P., Kohler, S., Fretin, D., Walravens, K., Garin-Bastuji, B., & Letesson, J. J. (2005). From the discovery of the Malta fever's agent to the discovery of a marine mammal reservoir, brucellosis has continuously been a re-emerging zoonosis. *Veterinary Research, 36*(3), 313-326.

Godfroid, J., Nielsen, K., & Saegerman, C. (2010). Diagnosis of brucellosis in livestock and wildlife. *Croatian Medical Journal*, *51*(4), 296-305.

Godfroid, J., Scholz, H. C., Barbier, T., Nicolas, C., Wattiau, P., Fretin, D., Whatmore, A. M., Cloeckaert, A., Blasco, J. M., Moriyon, I., & Saegerman, C. (2011). Brucellosis at the animal/ecosystem/human interface at the beginning of the 21st century. *Preventive Veterinary Medicine*, *102*(2), 118-131.

Gopalakrishnan, A., Dimri, U., Saminathan, M., Yatoo, M. I., Priya, G. B., Gopinath, D., Sujatha, V., Ajith, Y., Suthar, A., Lawrence, C., & Dhama, K. (2016). Virulence factors, intracellular survivability and mechanism of evasion from host immune response by *Brucella*: an overview. *Journal of Animal and Plant Science*, *26*(6), 1542-1555.

Gorlov, I. P., Gorlova, O. Y., Sunyaev, S. R., Spitz, M. R., & Amos, C. I. (2008). Shifting paradigm of association studies: Value of rare single-nucleotide polymorphisms. *The American Journal of Human Genetics*, *82*(1), 100-112. #

Greenfield, R. A., Drevets, D. A., Machado, L. J., Voskuhl, G. W., Cornea, P., & Bronze, M. S. (2002). Bacterial pathogens as biological weapons and agents of bioterrorism. *American Journal of the Medical Sciences*, *323*(6), 299-315.

Guzman-Verri, C., Manterola, L., Sola-Landa, A., Parra, A., Cloeckaert, A., Garin, J., Gorvel, J. P., Moriyon, I., Moreno, E., & López-Goñi, I. (2002). The two-component system BvrR/BvrS essential for *Brucella abortus* virulence regulates the expression of outer membrane proteins with counterparts in members of the Rhizobiaceae. *Proceedings of the National Academy of Sciences*, *99*(19), 12375-12380.

Habier, D. (2010). More than a third of the WCGALP presentations on genomic selection. *Jorunal of Animal Breeding and Genetics, (127)*, 336-337.

Halling, S. M., Peterson-Burch, B. D., Bricker, B. J., Zuerner, R. L., Qing, Z., Li, L. L., Kapur, V., Alt, D. P., & Olsen, S. C. (2005). Completion of the genome sequence of *Brucella abortus* and comparison to the highly similar genomes of *Brucella melitensis* and *Brucella suis*. *Journal of Bacteriology*, *187*(8), 2715-2726.

Halling, S. M., Tatum, F. M., & Bricker, B. J. (1993). Sequence and characterization of an insertion sequence, IS711, from *Brucella ovis*. *Gene*, *133*(1), 123-127.

Hartig, F. (2022). *DHARMa Computer Software: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. R package version 0.4.5*. https://CRAN.R-project.org/package=DHARMa.

Hasenauer, F., Caffaro, M., Czibener, C., Comerci, D., Poli, M., & Rossetti, C. (2013). Genetic analysis of the 3′ untranslated region of the bovine SLC11A1 gene reveals novel polymorphisms. *Molecular Biology Reports*, *40*, 545-552.

Horˇ ín, P., Rychlik, I., Templeton, J., & Adams, L. (1999). A complex pattern of microsatellite polymorphism within the bovine NRAMP1 gene. *European Journal of Immunogenetics*, *26*(4), 311-313.

Hull, N. C., & Schumaker, B. A. (2018). Comparisons of brucellosis between human and veterinary medicine. *Journal of Infection Ecology and Epidemiology*, *8*(1), 1500846.

Hunt, S. E., Moore, B., Amode, R. M., Armean, I. M., Lemos, D., Mushtaq, A., Parton, A., Schuilenburg, H., Szpak, M., Thormann, A., & Perry, E. (2022). Annotating and prioritizing genomic variants using the Ensembl Variant Effect Predictor: A tutorial. *Human Mutation*, *43*(8), 986-997.

International Organisation for Animal Health. (2008). *Manual of Diagnostic Tests and Vaccine for Terestrial Animals.* https://scholar.google.com

International Organisation for Animal Health. (2009). OIE terestrial Manual. *Bovine Brucellosis*, 2009, 1-35.

International Organisation for Animal Health. (2018). *Chapter 3.1.4- Infection with B. abortus, B. mellitensis and B. suis. Terestrial manual*. https://scholar.google.com

Islam, S., Barua, S. R., Moni, S. P., Islam, A., Rahman, A. A., & Chowdhury, S. (2021). Seroprevalence and risk factors for bovine brucellosis in the Chittagong Metropolitan Area of Bangladesh. *Veterinary Medicine and Science*, *7*(1), 86-98.

Israel, G. D. (1992). *Determining Sample Size: University of Florida Fact Sheet PEOD-6*. https://scholar.google.com

Jacob, C. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Jones, L. M., Berman, D., Moreno, E., Deyoe, B., Gilsdorf, M., Huber, J., & Nicoletti, P. (1980). Evaluation of a radial immunodiffusion test with polysaccharide B antigen for diagnosis of bovine brucellosis. *Journal of Clinical Microbiology*, *12*(6), 753-760.

Kanuya, N., Matiko, M., Kessy, B., Mgongo, F., Ropstad, E., & Reksen, O. (2006). A study on reproductive performance and related factors of zebu cows in pastoral herds in a semi-arid area of Tanzania. *Theriogenology*, *65*(9), 1859-1874.

Kaoud, H., Zaki, M. M., El-Dahshan, A., & Nasr, S. A. (2010). Epidemiology of brucellosis among farm animals. *Nature and Science, 8*(5), 190-197.

Karimuribo, E., Ngowi, H., Swai, E., & Kambarage, D. (2007). Prevalence of brucellosis in crossbred and indigenous cattle in Tanzania. *Livestock Research for Rural Development, 19*(10), 148-152.

Kasiulevičius, V., Šapoka, V., & Filipavičiūtė, R. (2006). Sample size calculation in epidemiological studies. *Gerontologija, 7*(4), 225-231.

Ke, Y., Wang, Y., Li, W., & Chen, Z. (2015). Type IV secretion system of Brucella spp. and its effectors. *Frontiers in Cellular and Infection Microbiology, 5*, 72.

Keller, M., Roxlau, A., Weng, W. M., Schmidt, M., Quandt, J., Niehaus, K., Jording, D., Arnold, W., & Puhler, A. (1995). Molecular analysis of the Rhizobium meliloti mucR gene regulating the biosynthesis of the exopolysaccharides succinoglycan and galactoglucan. *MPMI-Molecular Plant Microbe Interactions, 8*(2), 267-277.

Kempf, B., & Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Archives of Microbiology, 170*(5), 319-330.

Kennedy, G. C., Matsuzaki, H., Dong, S., Liu, W. M., Huang, J., Liu, G., Su, X., Cao, M., Chen, W., Zhang, J., & Liu, W. (2003). Large-scale genotyping of complex DNA. *Nature Biotechnology, 21*(10), 1233-1237. https://doi.org/10.1038/nbt869

Khurana, S. K., Sehrawat, A., Tiwari, R., Prasad, M., Gulati, B., Shabbir, M. Z., Chhabra, R., Karthik, K., Patel, S. K., Pathak, M., & Iqbal-Yatoo, M. (2021). Bovine brucellosis: A comprehensive review. *Veterinary Quarterly, 41*(1), 61-88.

Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R., San Cristobal, M., Servin, B., McCulloch, R., Whan, V., Gietzen, K., & Paiva, S. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biology, 10*(2), e1001258.

Kılıç, S., Çelebi, B., & Turan, M. (2021). *Brucella melitensis* and *Brucella abortus* genotyping via real-time PCR targeting 21 variable genome loci. *Journal of Microbiological Methods, 180*, 106125. https://doi.org/10.1016/j.mimet.2020.106125

Kitalyi, J. (1984). Bovine brucellosis in Government parastatal and ujamaa village dairy farms in the central zone of Tanzania: Assessment of control measures in some of farms in Proceedings of *2nd Tanzania Veterinary Association Scientific Conference,* AICC, 4-6 December, Arusha, Tanzania.

Klein, R. J., Zeiss, C., Chew, E. Y., Tsai, J. Y., Sackler, R. S., Haynes, C., Henning, A. K., SanGiovanni, J. P., Mane, S. M., Mayne, S. T., & Bracken, M. B. (2005). Complement factor H polymorphism in age-related macular degeneration. *Science, 308*(5720), 385-389.

Kline, K. A., Fälker, S., Dahlberg, S., Normark, S., & Henriques-Normark, B. (2009). Bacterial adhesins in host-microbe interactions. *Cell Host and Microbes*, *5*(6), 580-592. https://doi.org/10.1016/j.chom.2009.05.011

Ko, J., & Splitter, G. A. (2003). Molecular host-pathogen interaction in brucellosis: Current understanding and future approaches to vaccine development for mice and humans. *Clinical Microbiology Reviews*, *16*(1), 65-78.

Kolo, F. B., Adesiyun, A. A., Fasina, F. O., Katsande, C. T., Dogonyaro, B. B., Potts, A., Matle, I., Gelaw, A. K., & Van Heerden, H. (2019). Seroprevalence and characterization of *Brucella* species in cattle slaughtered at Gauteng abattoirs, South Africa. *Veterinary Medicine and Science*, *5*(4), 545-555.

Kubista, M., Andrade, J. M., Bengtsson, M., Forootan, A., Jonák, J., Lind, K., Sindelka, R., Sjöback, R., Sjögreen, B., Strömbom, L., & Ståhlberg, A. (2006). The real-time polymerase chain reaction. *Molecular Aspects of Medicine*, *27*(2-3), 95-125.

Kulldorff, M. (2009). *Information Management Services Isv. Software for the Spatial and Space-Time Scan Statistics*. https://www.satscan.org/

Kurwijila, L., & Boki, K. (2003). *Milk and Dairy Products, Post-Harvest Losses and Food Safety in Sub-Saharan Africa and the Near East: A Review of the Small Scale Dairy Sector-Tanzania.* https://scholar.google.com

Kurwijila, L. R., Omore, A. O., & Grace, D. (2012). *Tanzania Dairy Industry Overview*. https://scholar.google.com

Lamb, V. L., Jones, L., Schurig, G., & Berman, D. (1979). Enzyme-linked immunosorbent assay for bovine immunoglobulin subclass-specific response to *Brucella abortus* lipopolysaccharides. *Infection and Immunity*, *26*(1), 240-247.

Lapaque, N., Muller, A., Alexopoulou, L., Howard, J. C., & Gorvel, J. P. (2009). *Brucella abortus* induces Irgm3 and Irga6 expression via type-I IFN by a MyD88-dependent pathway, without the requirement of TLR2, TLR4, TLR5 and TLR9. *Microbial Pathogenesis*, *47*(6), 299-304. https://doi.org/10.1016/j.micpath.2009.09.005

Lee, J. J., Kim, J. H., Kim, D. G., Kim, D. H., Simborio, H. L., Min, W. G., Rhee, M. H., Lim, J. H., Chang, H. H., & Kim, S. (2014). Characterization of betaine aldehyde dehydrogenase (BetB) as an essential virulence factor of *Brucella abortus*. *Veterinary Microbiology*, *168*(1), 131-140. https://doi.org/10.1016/j.vetmic.2013.10.007

Li, Y., & Kim, J. J. (2015). Effective population size and signatures of selection using bovine 50K SNP chips in Korean native cattle (Hanwoo). *Evolutionary Bioinformatics*, *11*, S24359.

Lim, J. J., Kim, D. H., Lee, J. J., Kim, D. G., Min, W., Lee, H. J., Rhee, M. H., & Kim, S. (2012). Protective effects of recombinant *Brucella abortus* Omp28 against infection with a virulent strain of *Brucella abortus* 544 in mice. *Journal of Veterinary Science*, *13*(3), 287-292.

Lindahl, J. F., Goyal Kumar, N., Deka, R. P., Shome, R., & Grace, D. (2018). Serological evidence of *Brucella* infections in dairy cattle in Haryana, India. *Infection Ecology and Epidemiology*, *8*(1), 1555445.

Lokuruka, M. N. (2016). Overview of dairy processing and marketing in East African dairy value chains: Opportunities and challenges. *African Journal of Food Science*, *10*(11), 254-262.

López-Goñi, I., García-Yoldi, D., Marín, C. M., De Miguel, M. J., Barquero-Calvo, E., Guzman-Verri, C., Albert, D., & Garin-Bastuji, B. (2011). New Bruce-ladder multiplex PCR assay for the biovar typing of *Brucella suis* and the discrimination of *Brucella suis* and *Brucella canis*. *Veterinary Microbiology*, *154*(1-2), 152-155.

López-Goñi, I., Guzman-Verri, C., Manterola, L., Sola-Landa, A., Moriyon, I., & Moreno, E. (2002). Regulation of Brucella virulence by the two-component system BvrR/BvrS. *Veterinary Microbiology*, *90*(1-4), 329-339.

Lübeck, P. S., Skurnik, M., Ahrens, P., & Hoorfar, J. (2003). A Multiplex PCR-Detection Assay for Yersinia enterocolitica Serotype O: 9 and Brucella spp. Based on the Perosamine Synthetase Gene: Application to Brucella Diagnostics. *The Genus Yersinia: Entering the Functional Genomic Era,* 2003, 451-454.

Lumley, T. (2011). *Complex surveys: A Guide to Analysis Using R*. John Wiley & Sons.

Lunogelo, H., Makene, F., & Gray, H. (2020). *Dairy Processing in Tanzania Prospects for SME Inclusion: Innovation and Social Inclusion in Agro-processing.* https://scholar.google.com

Macedo, A., Costa, E., Silva, A., Paixão, T., & Santos, R. (2013). Monocyte-derived macrophages from Zebu (*Bos taurus indicus)* are more efficient to control *Brucella abortus* intracellular survival than macrophages from European cattle (*Bos taurus taurus*). *Veterinary Immunology and Immunopathology*, *151*(3-4), 294-302.

Maher, B. (2008). The case of the missing heritability. *Nature*, *456*(7218), 18-21.

Mahlau, E. (1967). Further brucellosis surveys in Tanzania. *Bulletin of Epizootic Diseases of Africa*, *15*, 373-378.

Mai, H. M., Irons, P. C., Kabir, J., & Thompson, P. N. (2012). A large seroprevalence survey of brucellosis in cattle herds under diverse production systems in northern Nigeria. *BMC Veterinary Research*, *8*(1), 144.

Makita, K., Fèvre, E., Waiswa, C., Eisler, M., Thrusfield, M., & Welburn, S. (2011). Herd prevalence of bovine brucellosis and analysis of risk factors in cattle in urban and peri-urban areas of the Kampala economic zone, Uganda. *BMC Veterinary Research*, *7*(1), 1-8.

Maksimović, Z., Jamaković, A., Semren, O., & Rifatbegović, M. (2022). Molecular detection of Brucella spp. in clinical samples of seropositive ruminants in Bosnia and

Herzegovina. *Comparative Immunology, Microbiology and Infectious Diseases*, *86*, 101821.

Mangen, M., Otte, J., Pfeiffer, D., & Chilonda, P. (2002). Bovine brucellosis in sub-Saharan Africa: Estimation of sero-prevalence and impact on meat and milk offtake potential. *FAO Livestock Policy Discussion Paper*, *8*, 1-58.

Mantur, B., Amarnath, S., & Shinde, R. (2007). Review of clinical and laboratory features of human brucellosis. *Indian Journal of Medical Microbiology*, *25*(3), 188-202.

Marchesini, M. I., Herrmann, C. K., Salcedo, S. P., Gorvel, J. P., & Comerci, D. J. (2011). In search of *Brucella abortus* type IV secretion substrates: Screening and identification of four proteins translocated into host cells through VirB system. *Cellular Microbiology*, *13*(8), 1261-1274.

Marcotty, T., Matthys, F., Godfroid, J., Rigouts, L., Ameni, G., Gey-van-Pittius, N., Kazwala, R., Muma, J., Van Helden, P., Walravens, K., & De Klerk, L. M. (2009). Zoonotic tuberculosis and brucellosis in Africa: Neglected zoonoses or minor public-health issues? The outcomes of a multi-disciplinary workshop. *Annals of Tropical Medicine and Parasitology*, *103*(5), 401-411.

Marin, C., Alabart, J., & Blasco, J. (1996). Effect of antibiotics contained in two Brucella selective media on growth of *Brucella abortus*, *B. melitensis*, and *B. ovis*. *Journal of Clinical Microbiology*, *34*(2), 426-428.

Marjoram, P., Zubair, A., & Nuzhdin, S. (2014). Post-GWAS: Where next? More samples, more SNPs or more biology? *Heredity*, *112*(1), 79-88.

Martín-Martín, A. I., Sancho, P., De Miguel, M. J., Fernández-Lago, L., & Vizcaíno, N. (2012). Quorum-sensing and BvrR/BvrS regulation, the type IV secretion system, cyclic glucans, and BacA in the virulence of *Brucella ovis*: Similarities to and differences from smooth brucellae. *Infection and Immunity*, *80*(5), 1783-1793.

Martin, F. J., Amode, M. R., Aneja, A., Austine-Orimoloye, O., Azov, A. G., Barnes, I., Becker, A., Bennett, R., Berry, A., Bhai, J., & Bhurji, S. K. (2023). Ensembl 2023. *Nucleic Acids Research*, *51*(D1), D933-D941.

Martínez, R., Dunner, S., Toro, R., Tobón, J., Gallego, J., & Cañón, J. (2010). Effect ofpolymorphisms in the Slc11a1 coding region on resistance to brucellosis by macrophages in vitro and after challenge in two Bos breeds (Blanco Orejinegro and Zebu). *Genetics and Molecular Biology*, *33*, 463-470.

Martinez, R., Toro, R., Montoya, F., Burbano, M., Tobón, J., Gallego, J., Dunner, S., & Cañón, J. (2008). Bovine SLC11A1 3′ UTR SSCP genotype evaluated by a macrophage in vitro killing assay employing a *Brucella abortus* strain. *Journal of Animal Breeding and Genetics*, *125*(4), 271-279.

Martirosyan, A., Moreno, E., & Gorvel, J. P. (2011). An evolutionary strategy for a stealthy intracellular *Brucella* pathogen. *Immunological Reviews*, *240*(1), 211-234. https://doi.org/10.1111/j.1600-065X.2010.00982.x

Marzetti, S., Carranza, C., Roncallo, M., Escobar, G. I., & Lucero, N. E. (2013). Recent trends in human *Brucella canis* infection. *Comparative Immunology, Microbiology and Infectious Diseases*, *36*(1), 55-61.

Mathew, C., Klevar, S., Løken, T., Mwamengele, G., Skjerve, E., Godfroid, J., Stokstad, S., & Mdegela, R. H. (2017). Reproductive infections in cattle in Tanzania: Lessons for control priorities. *Microbiology and Infectious Diseases*, *5*(1), 1-9.

Mathew, C., Stokstad, M., Johansen, T. B., Klevar, S., Mdegela, R. H., Mwamengele, G., Michel, P., Escobar, L., Fretin, D., & Godfroid, J. (2015). First isolation, identification, phenotypic and genotypic characterization of *Brucella abortus* biovar 3 from dairy cattle in Tanzania. *BMC Veterinary Research*, *11*, 1-9. https://doi.org/10.1186/s12917-015-0476-8

McDermott, J., & Arimi, S. (2002). Brucellosis in sub-Saharan Africa: Epidemiology, control and impact. *Veterinary Microbiology*, *90*(1-4), 111-134.

McDermott, J., Grace, D., & Zinsstag, J. (2013). Economics of brucellosis impact and control in low-income countries. *Revue Scientifique et Technique*, *32*(1), 243-261.

McEwen, A. (1940). Experiments in Contagious Abortion. The Immunity of Cattle Inoculated with Vaccines of Graded Virulence. *Veterinary Record*, *52*, 815-819.

McIntosh, M., Krol, E., & Becker, A. (2008). Competitive and cooperative effects in quorum-sensing-regulated galactoglucan biosynthesis in Sinorhizobium meliloti. *Journal of Bacteriology*, 190(15), 5308-5317.

Mckinney, B., & Pajewski, N. (2012). Six degrees of epistasis: Statistical network models for GWAS. *Frontiers in Genetics*, *2*, 109.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., Flicek, P., & Cunningham, F. (2016). The Ensembl Variant Effect Predictor. *Genome Biology*, *17*(1), 122.

McLaren, W., Pritchard, B., Rios, D., Chen, Y., Flicek, P., & Cunningham, F. (2010). Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*, *26*(16), 2069-2070.

Mdegela, R. H., Kusiluka, L. J. M., Kapaga, A. M., Karimuribo, E. D., Turuka, F. M., Bundala, A., Kivaria, F., Kabula, B., Manjurano, A., Loken, T., & Kambarage, D. M. (2004). Prevalence and determinants of mastitis and milk-borne zoonoses in smallholder dairy farming sector in Kibaha and Morogoro districts in Eastern Tanzania. *Journal of Veterinary Medicine, Series B*, *51*(3), 123-128.

Mengele, I., Garcia-Gonzalez, D., Mekonnen, G., & Arenas-Gamboa, A. (2018). Bovine brucellosis seroprevalence, farmers' awareness, practices and animal health extension services inputs in Mpwapwa district, Tanzania. *Tanzania Veterinary Journal*, *36*, 135-144.

Mengele, I. J., Shirima, G. M., Bwatota, S. F., Motto, S. K., Bronsvoort, B. M. D. C., Komwihangilo, D. M., Lyatuu, E., Cook, E. A. J., & Hernandez-Castro, L. E. (2023). The Status and Risk Factors of Brucellosis in Smallholder Dairy Cattle in Selected Regions of Tanzania. *Veterinary Sciences*, *10*(2), 1-15.

Mengele, I. J., Shirima, G. M., Bronsvoort, B. M., Hernandez-Castro, L. E., & Cook, E. A. J. (2023a). Diagnostic challenges of brucellosis in humans and livestock in Tanzania: A thematic review. *CABI One Health*, *2023*(2023)

Mercier, E., Jumas-Bilak, E., Allardet-Servent, A., O'Callaghan, D., & Ramuz, M. (1996). Polymorphism in Brucella strains detected by studying distribution of two short repetitive DNA elements. *Journal of Clinical Microbiology*, *34*(5), 1299-1302.

Meyer, M. (1985). Characterization of Brucella abortus strain 19 isolated from human and bovine tissues and fluids. *American Journal of Veterinary Research*, *46*(4), 902-904.

Mfune, R. L. (2015). *Epidemiological Study of Bovine Brucellosis in Smallholder Dairy Cattle in Lushoto and Rungwe Districts.* https://scholar.google.com

Mfune, R. L., Mubanga, M., Silwamba, I., Sagamiko, F., Mudenda, S., Daka, V., Godfroid, J., Hangombe, B. M., & Muma, J. B. (2021). Seroprevalence of bovine brucellosis in selected districts of Zambia. *International Journal of Environmental Research and Public Health*, *18*(4), 1436.

Michael, C., & Elizabeth, M. (2018). *Agar Gel Immundifussion.* https://scholar.google.com

Michael, S., Mbwambo, N., Mruttu, H., Dotto, M., Ndomba, C., Silva, M. D., Makusaro, F., Nandonde, S., Crispin, J., Shapiro, B. I., & Desta, S. (2018). *Tanzania Livestock Master Plan, Ministry of Livestock and Fisheries, United Republic of Tanzania.* https://scholar.google.com

Michaux-Charachon, S., Bourg, G., Jumas-Bilak, E., Guigue-Talet, P., Allardet-Servent, A., O'Callaghan, D., & Ramuz, M. (1997). Genome structure and phylogeny in the genus Brucella. *Journal of Bacteriology*, *179*(10), 3244-3249.

Mirabella, A., Terwagne, M., Zygmunt, M., Cloeckaert, A., De Bolle, X., & Letesson, J. (2013). *Brucella melitensis* MucR, an orthologue of Sinorhizobium meliloti MucR, is involved in resistance to oxidative, detergent, and saline stresses and cell envelope modifications. *Journal of Bacteriology*, *195*(3), 453-465.

Mitterran, K. N. R., Barberine, S. A., Oumarou, F., & Simo, G. (2020). Detection of *Brucella abortus* and *Brucellla melitensis* in cattle and sheep from southern Cameroon. *Research Square*, 2020, 1-13. https://doi.org/10.21203/rs.3.rs-21575/v2

Morales-Estrada, A., Hernández-Castro, R., López-Merino, A., Singh-Bedi, J., & Contreras-Rodríguez, A. (2016). Isolation, identification, and antimicrobial susceptibility of

Brucella spp. cultured from cows and goats manure in Mexico. *Austral Journal of Veterinary Sciences*, *48*(2), 231-235.

Morales-Estrada, A. I., Joel, C. S., Ahide, L. M., Maria, R. M. G., Juan, G. V. V., & Araceli, C. R. (2012). Characterization of *Brucella* species in Mexico by Bruce-Ladder polymerase chain reaction (PCR). *African Journal of Microbiology Research*, *6*(11), 2793-2796.

Moreno, E. (2014). Retrospective and prospective perspectives on zoonotic brucellosis. *Journal of Frontiers in Microbiology*, *5*, 213.

Moreno, E., Berman, D., & Boettcher, L. (1981). Biological activities of *Brucella abortus* lipopolysaccharides. *Infection and Immunity*, *31*(1), 362-370.

Moriyón, I., Alcalá, C. M. M., Garin-Bastuj, R., Vaz, Y., Melzer, F., McGiven, J., Kadre, B. (2018). *Brucellosis: Gap Analysis Scoring for Cattle Brucellosis.* https://www. discontools. eu/ database/ 42-brucellosis.html

Moriyón, I., Grilló, M., Monreal, D., González, D., Marín, C., López-Goñi, I., Mainar-Jaime, R., Moreno, E., & Blasco, J. (2004). Rough vaccines in animal brucellosis: Structural and genetic basis and present status. *Veterinary Research*, *35*(1), 1-38.

Moskvina, V., Craddock, N., Holmans, P., Owen, M. J., & O'Donovan, M. C. (2006). Effects of differential genotyping error rate on the type I error probability of case-control studies. *Human Heredity*, *61*(1), 55-64.

Motto, S., Makondo, Z., Sailen, A., Magwisha, H., & Mayenga, E. (2018). Molecular Detection and Genotyping of *Brucella* Species in Cattle and Goat by High Resolution Melt Analysis in Tanzania. *Tanzania Veterinary Journal*, *36*, 54-59.

Mrode, R., Ojango, J., Ekine-Dzivenu, C., Aliloo, H., Gibson, J., & Okeyo, M. (2021). Genomic prediction of crossbred dairy cattle in Tanzania: A route to productivity gains in smallholder dairy systems. *Journal of Dairy Science*, *104*(11), 11779-11789.

Muendo, E. N., Mbatha, P. M., Macharia, J., Abdoel, T. H., Janszen, P. V., Pastoor, R., & Smits, H. L. (2012). Infection of cattle in Kenya with *Brucella abortus* biovar 3 and

Brucella melitensis biovar 1 genotypes. *Tropical Animal Health and Production*, *44*, 17-20.

Mugizi, D. R., Boqvist, S., Nasinyama, G. W., Waiswa, C., Ikwap, K., Rock, K., Lindahl, E., Magnusson, U., & Erume, J. (2015). Prevalence of and factors associated with *Brucella* sero-positivity in cattle in urban and peri-urban Gulu and Soroti towns of Uganda. *Journal of Veterinary Medical Science*, *77*(5), 557-564.

Muma, J. B., Pandey, G. S., Munyeme, M., Mumba, C., Mkandawire, E., & Chimana, H. M. (2012). Brucellosis among smallholder cattle farmers in Zambia: Public health significance. *Tropical Animal Health and Production,* 44, 915-920.

Musallam, I., Ndour, A. P., Yempabou, D., Ngong, C. A. C., Dzousse, M. F., Mouiche-Mouliom, M. M., Feussom, J. M. K., Ntirandekura, J. B., Ntakirutimana, D., Fane, A., & Dembele, E. (2019). Brucellosis in dairy herds: A public health concern in the milk supply chains of West and Central Africa. *Acta Tropica*, *197*, 105042-105042.

Mwinyi, M. O., Kayunze, K. A., Sitali, D. C., Simuunza, M. C., & Muma, J. B. (2016). Socio-economic Impact of Brucellosis on Livestock Farmers in Southern and Western Provinces, Zambia. *International Journal of Technical Research and Application*, *4*(6), 204-209.

Naidoo, N., Pawitan, Y., Soong, R., Cooper, D. N., & Ku, C. S. (2011). Human genetics and genomics a decade after the release of the draft sequence of the human genome. *Human Genomics*, *5*(6), 1-46.

Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, *14*(134), 1-11.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133-142.

Nash, D., & Freeman, A. (2004). Selection for Disease Resistance: Why It Is Important and How It Can Be Accomplished. *Lowa State University Animal Industry Report*, *650*(1), 58.

Neta, A. V. C., Mol, J. P., Xavier, M. N., Paixão, T. A., Lage, A. P., & Santos, R. L. (2010). Pathogenesis of bovine brucellosis. *The Veterinary Journal*, 184(2), 146-155.

Nicoletti, P., & Winter, A. (1990). *The Immune Response to B.abortus: The Cell-Mediated Response to Infection.* CRC Press: Boca Raton, FL.

Nielsen, K. (2002). Diagnosis of brucellosis by serology. *Veterinary Microbiology*, *90*(1-4), 447-459.

Nielsen, K., Kelly, L., Gall, D., Nicoletti, P., & Kelly, W. (1995). Improved competitive enzyme immunoassay for the diagnosis of bovine brucellosis. *Veterinary Immunology and Immunopathology*, *46*(3-4), 285-291.

Nielsen, K., & Yu, W. J. P. (2010). Serological diagnosis of brucellosis. *Prilozi*, *31*(1), 65-89.

Njeru, J., Nthiwa, D., Akoko, J., Oyas, H., & Bett, B. (2022). Correction: Incidence of Brucella infection in various livestock species raised under the pastoral production system in Isiolo County, Kenya. *BMC Veterinary Research*, *18*(1), 194.

Njombe, A., & Msanga, Y. (2009). *Livestock and Dairy Industry Development in Tanzania.* Department of Livestock production, Marketing Infrastructure Development, Ministry of Livestock Development. https://scholar.google.com

Njombe, A., Msanga, Y., Mbwambo, N., & Makembe, N. (2011). *The Tanzania Dairy Industry: Status, Opportunities and Prospects. 7th African Dairy Conference and Exhibition, 25-27 May, Dar Es Salaam, Tanzania.* https://scholar.google.com

Ntivuguruzwa, J. B., Babaman, K. F., Mwikarago, E. I., & van Heerden, H. (2022). Seroprevalence of brucellosis and molecular characterization of *Brucella* spp. from slaughtered cattle in Rwanda. *PloS One*, *17*(11), e0261595.

Nyange, D., & Mdoe, N. (1995). *Dairy Industry in Tanzania and the Prospect for Small Scale Milk Producers. Workshop on Strategies for Market Orientation of Small Scale Milk Producers and their Organisations, 20-24 March, Morogoro, Tanzania.* https://scholar.google.com

O'Callaghan, D., & Whatmore, A. M. (2011). Brucella genomics as we enter the multi-genome era. *Briefings in Functional Genomics, 10*(6), 334-341.

Olsen, S., & Palmer, M. (2014). Advancement of knowledge of Brucella over the past 50 years. *Veterinary Pathology, 51*(6), 1076-1089.

Omer, M., Skjerve, E., Woldehiwet, Z., & Holstad, G. (2000). Risk factors for Brucella spp. infection in dairy cattle farms in Asmara, State of Eritrea. *Preventive Veterinary Medicine, 46*(4), 257-265.

Ouahrani-Bettache, S., Soubrier, M. P., & Liautard, J. P. (1996). IS6501-anchored PCR for the detection and identification of Brucella species and strains. *Journal of Applied Bacteriology*, 81(2), 154-160.

Padilla, P. F., Nielsen, K., Ernesto, S. L., & Ling, Y. W. (2010). Diagnosis of brucellosis. *The Open Veterinary Science Journal*, 4(1), 46-60.

Pahl, R., & Schäfer, H. (2010). Permory: An LD-exploiting permutation test algorithm for powerful genome-wide association testing. *Bioinformatics, 26*(17), 2093-2100.

Paixão, T. A., Poester, F. P., Carvalho Neta, A. V., Borges, A. M., Lage, A. P., & Santos, R. L. (2007). NRAMP1 3' untranslated region polymorphisms are not associated with natural resistance to Brucella abortus in cattle. *Infection and Immunity, 75*(5), 2493–2499. https://doi.org/10.1128/IAI.01855-06

Palacios-Chaves, L., Conde-Álvarez, R., Gil-Ramírez, Y., Zúñiga-Ripa, A., Barquero-Calvo, E., Chacón-Díaz, C., Chaves-Olarte, E., Arce-Gorvel, V., Gorvel, J. P., Moreno, E., & De Miguel, M. J. (2011). *Brucella abortus* ornithine lipids are dispensable outer membrane components devoid of a marked pathogen-associated molecular pattern. *PLoS One*, *6*(1), e16030.

Pandey, G. S., Kobayashi, K., Nomura, Y., Nambota, A., Mwima, H. K., & Suzuki, A. K. (1999). Studies on sero-prevalence of brucellosis in Kafue lechwe (*Kobus leche kafuensis*) in Zambia. *Indian Veterinary Journal*, *76*(4), 275-278.

Pappas, G. (2010). The changing Brucella ecology: Novel reservoirs, new threats. *International Journal of Antimicrobial Agents, 36,* S8-S11.

Pappas, G., Papadimitriou, P., Akritidis, N., Christou, L., & Tsianos, E. V. (2006). The new global map of human brucellosis. *The Lancet Infectious Diseases*, *6*(2), 91-99.

Parkin, J., & Cohen, B. (2001). An overview of the immune system. *The Lancet*, *357*(9270), 1777-1789.

Pennisi, E. (2007). Breakthrough of the year. *Science*, *318*(5858), 1842-1843.

Petersen, E., Rajashekara, G., Sanakkayala, N., Eskra, L., Harms, J., & Splitter, G. (2013). Erythritol triggers expression of virulence traits in *Brucella melitensis*. *Microbes and Infection*, *15*(6-7), 440-449. https://doi.org/10.1016/j.micinf.2013.02.002.

Petersen, J. L., Mickelson, J. R., Cothran, E. G., Andersson, L. S., Axelsson, J., Bailey, E., Bannasch, D., Binns, M. M., Borges, A. S., Brama, P., & Da Câmara-Machado, A. (2013). Genetic diversity in the modern horse illustrated from genome-wide SNP data. *PloS One*, *8*(1), e54997.

Pirone, L., Pitzer, J. E., D'Abrosca, G., Fattorusso, R., Malgieri, G., Pedone, E. M., Pedone, P. V., Roop, R. M., & Baglivo, I. (2018). Identifying the region responsible for *Brucella abortus* MucR higher-order oligomer formation and examining its role in gene regulation. *Scientific Reports*, *8*(1), 1-8.

Planning, Ministry of Finance and Planning (2019). *Tanzania Mainland Household Budget Survey 2017-2018, Key indicators Report: Poverty Eradication Division.* National Bureau of Statistics, United Republic of Tanzania. https://scholar.google.com

Planning, Ministry of Finance and Planning (2021). *The Economic Survey 2020*. United Republic of Tanzania. https://scholar.google.com

Poester, F., Samartino, L., & Santos, R. (2013). Pathogenesis and pathobiology of brucellosis in livestock. *Revue of Scientifique Technique Office International des Epizootics*, *32*(1), 105-115.

Posadas, D. M., Ruiz-Ranwez, V., Bonomi, H. R., Martín, F. A., & Zorreguieta, A. (2012). BmaC, a novel autotransporter of *Brucella suis*, is involved in bacterial adhesion to host cells. *Cellular Microbiology*, *14*(6), 965-982.

Poulsen, K. P., Hutchins, F. T., McNulty, C. M., Tremblay, M., Zabala, C., Barragan, V., Lopez, L., Trueba, G., & Bethel, J.W., 2014. Brucellosis in dairy cattle and goats in northern Ecuador. *The American Journal of Tropical Medicine and Hygiene*, *90*(4), 712-715.

Price, R., Templeton, J., Smith, R., & Adams, L. (1990). Ability of mononuclear phagocytes from cattle naturally resistant or susceptible to brucellosis to control in vitro intracellular survival of Brucella abortus. *Infection and Immunity*, *58*(4), 879-886.

Probert, W. S., Schrader, K. N., Khuong, N. Y., Bystrom, S. L., & Graves, M. H. (2004). Real-Time Multiplex PCR Assay for Detection of *Brucella* spp., *B. abortus*, and *B. melitensis*. *Journal of Clinical Microbiology*, *42*(3), 1290-1293.

Purcell, S., Cherny, S. S., & Sham, P. C. (2003). Genetic Power Calculator: Design of linkage and association genetic mapping studies of complex traits. *Bioinformatics,* 19(1), 149-150.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559-575.

Qureshi, T., Templeton, J., & Adams, L. (1996). Intracellular survival of *Brucella abortus, Mycobacterium bovis* BCG, *Salmonella dublin*, and *Salmonella typhimurium* in macrophages from cattle genetically resistant to *Brucella abortus*. *Veterinary Immunology and Immunopathology*, *50*(1-2), 55-65.

Radostitis, O. M., Gay, C. C., Blod, C. D. A., & Hinchcliff, K. W. (2000). Brucellosis. *In Veterinary Medicine, a Text book of the diseases of cattle, sheep, pigs, goats and horses* (9th Ed.). W.B. Sounders, London, UK.

Radostits, O. M., Gay, C., Hinchcliff, K. W., & Constable, P. D. J. V. M. (2007). *Brucellosis.* Veterinary Medicine, a textbook of the diseases of Cattle, Horses, Sheep, Pigs and Goats. W.B. Sounders, Philadelphia, US, 2045-2050.

Ragoussis, J. (2009). Genotyping technologies for genetic research. *Annual Review of Genomics and Human Genetics*, *10*, 117-133.

Rajendhran, J. (2021). Genomic insights into Brucella. *Infection, Genetics and Evolution*, 87, 104635.

Raymond, B., Johansson, A. M., McCormack, H. A., Fleming, R. H., Schmutz, M., Dunn, I. C., & De Koning, D. J. (2018). Genome-wide association study for bone strength in laying hens. *Journal of Animal Science*, *96*(7), 2525-2535.

Refai, M. (2002). Incidence and control of brucellosis in the Near East region. *Veterinary Microbiology*, *90*(1-4), 81-110.

Riancho, J. A. (2012). Genome-wide association studies (GWAS) in complex diseases: Advantages and limitations. *Reumatologia Clinica*, *8*(2), 56-57.

Richey, E. J. H. (1997). *Brucella abortus Disease (Brucellosis) in Beef Cattle.* University of Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences: Gainesville, FL, USA.

Riezu-Boj, J. I., Moriyón, I., Blasco, J. M., Gamazo, C., & Díaz, R. (1990). Antibody response to *Brucella* ovis outer membrane proteins in ovine brucellosis. *Infection and Immunity*, *58*(2), 489-494.

Risch, N., & Merikangas, K. (1996). The future of genetic studies of complex human diseases. *Science*, *273*(5281), 1516-1517.

Roberts, C., & Gray, A. (1973). Studies on trypanosome-resistant cattle. II. The effect of trypanosomiasis on N'dama, Muturu and Zebu cattle. *Tropical Animal Health Production*, *5*(4), 220-233.

Roop, R. M., Gaines, J. M., Anderson, E. S., Caswell, C. C., & Martin, D. W. (2009). Survival of the fittest: how *Brucella* strains adapt to their intracellular niche in the host. *Medical Microiology and Immunology*, *198*(4), 221-238.

Rossetti, C., Drake, K., Siddavatam, P., Lawhon, S., & Nunes, J. (2013). Systems Biology Analysis of *Brucella* Infected Peyer's Patch Reveals Rapid Invasion with Modest Transient Perturbations of the Host Transcriptome. *PLoS One*, *8*(12), e81719.

Ruiz-Ranwez, V., Posadas, D.M., Van der Henst, C., Estein, S. M., Arocena, G. M., Abdian, P. L., Martín, F. A., Sieira, R., De Bolle, X., & Zorreguieta, A. (2013). BtaE, an adhesin that belongs to the trimeric autotransporter family, is required for full virulence and defines a specific adhesive pole of *Brucella suis*. *Infection and Immunity*, *81*(3), 996-1007.

Sagamiko, F., Muma, J., Karimuribo, E., Mwanza, A., Sindato, C., & Hang'ombe, B. (2018). Sero-prevalence of Bovine Brucellosis and associated risk factors in mbeya region, Southern highlands of Tanzania. *Acta Tropica*, *178*, 169-175.

Sambu, R. M., Mathew, C., Nonga, H. E., Lukambagire, A. S., Yapi, R. B., Akoko, J., Fokou, G., Keyyu, J. D., Bonfoh, B., & Kazwala, R. R. (2021). Circulating *Brucella* species in wild animals of the Serengeti ecosystem, Tanzania. *One Health Outlook*, *3*(1), 15.

Sammartino, L., Gil, A., & Elzer, P. (2006). *Capacity Building for Surveillance and Control of Bovine and Caprine Brucellosis.* Food and Agriculture Organization, Rome, Italy. https://www.fao.org/3/a0083e/a0083e0a.htm.

Schelling, E., Diguimbaye, C., Daoud, S., Nicolet, J., Boerlin, P., Tanner, M., & Zinsstag, J. (2003). Brucellosis and Q-fever seroprevalences of nomadic pastoralists and their livestock in Chad. *Preventive Veterinary Medicine*, *61*(4), 279-293.

Schurig, G. G., Roop, R. M., Bagchi, T., Boyle, S., Buhrman, D., & Sriranganathan, N. (1991). Biological properties of RB51: A stable rough strain of *Brucella abortus*. *Veterinary Microbiology*, *28*(2), 171-188.

Schurig, G. G., Sriranganathan, N., & Corbel, M. J. (2002). *Brucellosis Vaccines: Past, Present and Future. Veterinary Microbiology*, *90*(1-4), 479-496.

Scientific, T. (2019). 260/280 and 260/230 Ratios. *T042-Technical Bulletin NanoDrop Spectrophotometers*. https://scholar.google.com

Segwagwe, B. E., Samkange, A., Mushonga, B., Kandiwa, E., & Ndazigaruye, G. (2018). Prevalence and risk factors for brucellosis seropositivity in cattle in Nyagatare District, Eastern Province, Rwanda. *Journal of the South African Veterinary Association*, *89*(1), 1-8.

Shirima, G. (2005). *The Epidemiology of Brucellosis in Animals and Humans in Arusha and Manyara Regions in Tanzania*. https://scholar.google.com

Shirima, G., & Kunda, J. (2016). Prevalence of brucellosis in the human, livestock and wildlife interface areas of Serengeti National Park, Tanzania. *Onderstepoort Journal of Veterinary Research*, *83*(1), 1-4.

Shirima, G., Lyimo, B., & Kanuya, N. (2018). Re-emergence of Bovine Brucellosis in Smallholder Dairy Farms in Urban Settings of Tanzania. *Journal of Applied Life Sciences International*, *17*(2), 1-7.

Shome, R., Natesan, K., Kalleshamurthy, T., Yadav, C., Sahay, S., Skariah, S., Mohandoss, N., Kumar, O. R. V., Shome, B. R., & Rahman, H. (2023). Management of bovine brucellosis in organized dairy herds through the identification of risk factors: A cross-sectional study from Karnataka, India. *Veterinary World*, *16*(5), 1122-1130.

Singh, B., Dhand, N. K., & Gill, J. (2015). Economic losses occurring due to brucellosis in Indian livestock populations. *Preventive Veterinary Medicine*, *119*(3-4), 211-215.

Skendros, P., & Boura, P. (2013). Immunity to brucellosis. *Revue Scientifique et Technique*, *32*(1), 137-147.

Slatkin, M. (2008). Linkage disequilibrium: Understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics*, *9*(6), 477-485.

Soomro, A. H., Kamboh, A. A., Rind, R., Dawani, P., Sarwar, M., Abro, S. H., & Awais, M. (2014). A study on prevalence and risk factors of brucellosis in cattle and buffaloes in district Hyderabad, Pakistan. *Journal of Animal Health and Production*, *2*(3), 33-37.

Sperry, J., & Robertson, D. C. (1975). Erythritol catabolism by Brucella abortus. *Journal of Bacteriology*, *121*(2), 619-630.

Stack, J., Perrett, L., Brew, S., & MacMillan, A. (1999). Competitive ELISA for bovine brucellosis suitable for testing poor quality samples. *The Veterinary Record*, *145*(25), 735-736.

National Bureau of Statistics (2012). *Socio-Economic Atlas of Tanzania: Average Household Size and Literarcy.* United Republic of Tanzania. https://scholar.google.com

National Bureau of Statistics. (2021). National Sample Census of Agriculture 2019-2020. National Report. *Ministry of Finance and Planning, United Republic of Tanzania*. https://scholar.google.com

Sterne, M., Trim, G., & Broughton, E. (1971). Immunisation of laboratory animals and cattle with non-agglutinogenic extracts of *Brucella abortus* strain 45/20. *Journal of Medical Microbiology*, *4*(2), 185-194.

Stone, R. T., & Cundiff, L. V. (1985). Breeding cattle for genetic resistance to disease. *Beef Research Program Progress Report*, 2, 25-26.

Swai, E., Mshanga, D., Sanka, N., & Marandu, N. (2005). Prevalence of bovine brucellosis in smallholder dairying farming area, Moshi, Tanzania. *Bulletin of Animal Health and Production in Africa*, *53*(2), 97-105.

Swai, E., & Schoonman, L. (2012). A survey of zoonotic diseases in trade cattle slaughtered at Tanga city abattoir: A cause of public health concern. *Asian Pacific Journal of Tropical Biomedicine*, *2*(1), 55-60.

Swai, E. S., & Schoonman, L. (2010). The use of rose bengal plate test to asses cattle exposure to *Brucella* infection in traditional and smallholder dairy production systems of Tanga region of Tanzania. *Veterinary Medicine International*, *2010*, 1-8.

Tabangin, M. E., Woo, J. G., & Martin, L. J. (2009). The effect of minor allele frequency on the likelihood of obtaining false positives. *BMC Proceedings*, 3(7), S7-S41.

Tadesse, G. (2016). Brucellosis seropositivity in animals and humans in Ethiopia: A meta-analysis. *PLoS Neglected Tropical Diseases*, *10*(10), e0005006.

Tcherneva, E., Rijpens, N., Jersek, B., & Herman, L. (2000). Differentiation of *Brucella* species by random amplified polymorphic DNA analysis. *Journal of Applied Microbiology*, *88*(1), 69-80.

Tcherneva, E., Rijpens, N., Naydensky, C., & Herman, L. (1996). Repetitive element sequence based polymerase chain reaction for typing of *Brucella* strains. *Veterinary Microbiology*, *51*(1-2), 169-178.

Terefe, Y., Girma, S., Mekonnen, N., & Asrade, B. (2017). Brucellosis and associated risk factors in dairy cattle of eastern Ethiopia. *Tropical Animal Health and Production*, *49*(3), 599-606.

Thomas, E., Bracewell, C., & Corbel, M. (1981). Characterisation of *Brucella abortus* strain 19 cultures isolated from vaccinated cattle. *The Veterinary Record*, *108*(5), 90-93.

Thompson-Crispi, K., Sewalem, A., Miglior, F., & Mallard, B. (2012). Genetic parameters of adaptive immune response traits in Canadian Holsteins. *Journal of Dairy Science*, *95*(1), 401-409.

Thrusfield, M. (2018). *Obervational Studies. In Veterinary Epidemiology.* Wiley Blackwel, River Street, Hoboken, NJ 07030, USA.

Tanzania Meteorological Agency. (2021). *Climatology-Climate Analysis. https:// scholar. google. com*

Tsairidou, S., Woolliams, J. A., Allen, A. R., Skuce, R. A., McBride, S. H., Wright, D. M., Bermingham, M. L., Pong-Wong, R., Matika, O., McDowell, S.W., & Glass, E. J. (2014). Genomic prediction for tuberculosis resistance in dairy cattle. *PLoS One*, *9*(5), e96728.

Tsegay, A., Tuli, G., Kassa, T., & Kebede, N. (2015). Seroprevalence and risk factors of Brucellosis in small ruminants slaughtered at Debre Ziet and Modjo export abattoirs, Ethiopia. *The Journal of Infection in Developing Countries*, *9*(04), 373-380.

Tulu, D. (2022). Bovine brucellosis: Epidemiology, public health implications, and status of brucellosis in Ethiopia. *Veterinary Medicine: Research and Reports, 2022(13),* 21-30.

Tumwine, G., Matovu, E., Kabasa, J. D., Owiny, D. O., & Majalija, S. (2015). Human brucellosis: Sero-prevalence and associated risk factors in agro-pastoral communities of Kiboga District, Central Uganda. *BMC Public Health*, *15*(1), 1-8.

Tuon, F. F., Gondolfo, R. B., Cerchiari, N., & Health, I. (2017). Human-to-human transmission of *Brucella*: A systematic review. *Tropical Medicine*, *22*(5), 539-546.

Van Straten, M., Bardenstein, S., Keningswald, G., & Banai, M. (2016). *Brucella abortus* S19 vaccine protects dairy cattle against natural infection with *Brucella melitensis*. *Vaccine, 34*(48), 5837–5839.

Vandesompel, J. (2008). *The qPCR Guide Eurogentec*. https://scholar.google.com

Vemulapalli, R., McQuiston, J. R., Schurig, G. G., Sriranganathan, N., Halling, S. M., & Boyle, S. M. (1999). Identification of an IS711 Element Interrupting the wboA Gene of *Brucella abortus*Vaccine Strain RB51 and a PCR Assay To Distinguish Strain RB51 from Other *Brucella* Species and Strains. *Clinical and Diagnostic Laboratory Immunology*, *6*(5), 760-764.

Vicente, A. F., Antunes, J. M., Lara, G. H., Mioni, M. S., Allendorf, S. D., Peres, M. G., Appolinário, C. M., Listoni, F. J., Ribeiro, M. G., & Megid, J. (2014). Evaluation of three formulations of culture media for isolation of *Brucella* spp. regarding their ability to inhibit the growth of contaminating organisms. *BioMed Research International*, *2014*, 1-3.

Villarroel, M., Grell, M., & Saenz, R. (2000). Isolation and identification of *B.abortus* RB51 in human. First report in Chile. *Archivos de Medicina Veterinaria*, *32*(1), 89-91.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *The American Journal of Human Genetics*, *90*(1), 7-24.

Vizcaíno, N., & Cloeckaert, A. (2012). *Biology and genetics of the Brucella outer membrane. In I. Lopez- Goni & D. O'Callaghan (Eds.), Brucella Molecular Microbiology and Genomics (pp. 133–161)*. Caister Academic Press, Norfork, UK.

Wakchaure, R., & Ganguly, S. (2016). Disease Resistance in Livestock: A Review. 10.5958/0973-9149.2016.00015.0. *Journal of Immunology and Immunopathology*, *18*(2), 94-99.

Wang, M. L. (2016). *MJ Biobank DNA Sample Quality Test and Analysis.* https:// scholar. google .com

Weinhäupl, I., Schöpf, K. C., Khaschabi, D., Kapaga, A. M., & Msami, H. M. (2000). Investigations on the prevalence of bovine tuberculosis and brucellosis in dairy cattle in Dar es Salaam region and in zebu cattle in Lugoba area, Tanzania. *Tropical Animal Health and Production*, *32*(3), 147-154.

Whatmore, A. M. (2009). Current understanding of the genetic diversity of Brucella, an expanding genus of zoonotic pathogens. *Infection, Genetics and Evolution*, *9*(6), 1168-1184.

World Health Organization. (2006). *The Control of Neglected Zoonotic Diseases: A Route to Poverty Alleviation.* Report of a joint WHO/DFID-AHP meeting, 20-21 September 2005 with the participation of FAO and OIE, WHO Headquarters, Geneva, Switzerland.

World Health Organization. (2020). *Neglected Zoonotic Diseases.* https://www.who.int

World Organization for Animal Health. (2022). *Brucellosis (Infection w*ith *Brucella abortus, B. melitensis* and *B. suis*). *OIE Terrestrial Manual*. https://scholar.google.com

Wragg, D., Cook, E. A., Latre de Late, P., Sitt, T., Hemmink, J. D., Chepkwony, M. C., Njeru, R., Poole, E. J., Powell, J., Paxton, E. A., & Callaby, R. (2022). A locus conferring tolerance to Theileria infection in African cattle. *PLoS Genetics*, *18*(4), e1010099.

Wright, A., & Smith, F. (1897). On the application of the serum test to the differential diagnosis of typhoid and Malta fever: And on the further application of the method of serum diagnosis to the elucidation of certain problems in connexion with the duration of immunity and the geographical distribution of disease. *The Lancet*, *149(3836)*, 656-659.

Yaeger, M. J., & Holler, L. D. (2007). *Bacterial Causes of Bovine Infertility and Abortion. In Current Therapy in Large Animal Theriogenology (pp. 389-399)*. Elsevier, Amsterdam, Netherlands

Yanti, Y., Sumiarto, B., Kusumastuti, T. A., Panus, A., & Sodirun, S. (2021). Seroprevalence and risk factors of brucellosis and the brucellosis model at the individual level of dairy cattle in the West Bandung District, Indonesia. *Veterinary World*, *14*(1), 1-10.

Zähringer, U., Lindner, B., Knirel, Y. A., Van den Akker, W. M., Hiestand, R., Heine, H., & Dehio, C. (2004). Structure and biological activity of the short-chain lipopolysaccharide from *Bartonella henselae* ATCC 49882T. *Journal of Biological Chemistry*, *279*(20), 21046-21054.

Zeng, P., Zhao, Y., Qian, C., Zhang, L., Zhang, R., Gou, J., Liu, J., Liu, L., & Chen, F. (2014). Statistical analysis for genome-wide association study. *Journal of Biomedical Research*, *29*(4), 285.

Zhou, X., Carbonetto, P., & Stephens, M. (2013). Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genetics*, *9*(2), e1003264.

Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, *44*(7), 821-824.

Zhou, X., & Stephens, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nature Methods*, *11*(4), 407-409.

APPENDICES

Appendix 1: The Consent form for a farmer to sign or finger printing

Animal owner consent form

Epidemiology and genetic associations of key zoonoses in small-holder dairy cattle in Tanzania

Collaborating Institutions: Nelson Mandela African Institute for Science and Technology
International Livestock Research Institute (ILRI)
The Roslin Institute, University of Edinburgh
Centre for Tropical Livestock Genetics and Health (CTLGH)
African Dairy Genetic Gains Project (ADGG)
Tanzania Livestock Research Institute (TALIRI)

Invitation to participate and description of project

You are invited to participate in a research study that is designed to test dairy cattle for a range of zoonoses. Zoonoses are those diseases that can be transmitted between animals and people. We will be sampling dairy cattle in Arusha, Kilimanjaro, Tanga, Njombe, Songwe, Iringa and Mbeya Regions. Through this work we will better understand the risks of these diseases to dairy cattle in Tanzania. This information will then help us make better recommendations on disease control and prevention strategies.

Your participation is entirely voluntary. This work is being carried out by research workers at NM-AIST in collaboration with the African Dairy Genetic Gains project, and our partner institutions.

Description of study procedures

1. **Examination and sampling of selected livestock**

   The selected dairy cattle will be examined and sampled. We will collect a blood sample and a swab from the vagina/prepuce. We will take a photograph of the eartag.

2. **Short questionnaire for the animal's owner**

   We will ask you some questions about yourself, including your name, and your relation to the animal. We will then ask some questions about the selected animal, including: age, recent health events including abortion and whether the animal has ever been treated or vaccinated

   This questionnaire should last 20 minutes.

Confidentiality

We will record your answers to the questionnaire electronically using a phone or tablet. This will be transmitted securely over the internet to computers at ILRI.

Any personal data disclosed during this interview will remain confidential to the research team and will only be used for the purposes of this project. Your identity will remain anonymous throughout. Personal identifying information, such as your name and telephone number, will be securely kept.

All samples taken from the animal will be tested at the lab at NM-AIST or ILRI. We will store the samples for further research after the initial tests are completed. Some of these samples may be sent to international laboratories for further testing, and some may be used to screen for other infections.

After the project is completed, anonymous samples may be kept in a long-term storage facility at ILRI or elsewhere and used for further research. As the storage will be anonymous, you will not be personally identifiable, and it will not be possible to report back to you the results of any future studies.

Risks

Examination and sampling of your animals will be carried out by trained animal health or veterinary professionals. The risk to your animals from sampling is very low. Occasionally blood sampling may cause some bleeding, swelling or bruising where the blood is collected. If you have any concerns about your animals after sampling, please contact the PRA, and inform the study team.

Benefits

The animal clinical examination and sampling will be provided at no cost to you. If any problems are identified during the examination, advice will be provided. Research results from your animal's sample will not be returned to you.

You will not directly benefit from the research, but we will organize meetings at the end of the study to communicate the results and discuss the research. Government departments will be informed of the results, and we hope that this will help them to make better decisions regarding the control of zoonoses in Tanzania.

Participation and withdrawal

We would like to emphasise that your participation is voluntary. If you consent and later decide that you do not want that the samples taken from your animal, or your information to be used in the study, you can contact the study team and ask that your contribution be withdrawn. Contact details are provided below.

If there is anything that you have not understood, please feel free to ask questions. You are welcome to ask us to go over any aspect of this form again before you decide whether or not to participate.

Authorisation

I confirm that I have read (or someone has read to me) this form, and I have understood the purposes of the research, what my participation will involve and any risks of the research. I agree to participate in the project described.

Name of animal owner.....Signature/thumb print.....

Date.....

Investigator name .....Signature of Investigator.....

Investigator contact number .....

Please record any comments from the respondent regarding diseases in their animals here:

Appendix 2: Oligonucleotide primers and probes details used to perform PCR assays

Target	Targeted gene	Sequences of primers and probes (5' -3')	Fluorophore/ quencher	Reference
Genus *Brucella*	IS711	Probe: AAG CCA ACA CCC GGC Forward: GGC CTA CCG CTG CGA AT Reverse: TTG CGG ACA GTC ACC ATA ATG	FAM/-MGBNFQ	Matero, 2011
B. melitensis	IS711 downstream of BMEI1162	Probe CAGGAGTGTTTCGGCTCAGAATAATCCACA Forward AACAAGCGGCACCCCTAAAA Reverse CATGCGCTATGATCTGGTTACG	Texas Red/BHQ2	Probert. 2004
B. abortus	IS711 downstream of alkB	Probe: CGCTCATGCTCGCCAGACTTCAATG Forward: GCGGCTTTTCTATCACGGTATTC Reverse: CATGCGCTATGATCTGGTTACG	JOE/BHQ1	

Appendix 3: Questionnaire used in the study

Section	Question	Choices	Label	hint
Registration	District	Select/filter	District	
	Ward	Select/filter	Ward	
	Village	Select/filter	Village	
	Farmer Name	Select/filter	Farmer_Name	
	Animal ID	Select/filter	Animal_ID	
	Signed consent to allow sampling	Yes; no	Consent	If no terminate interview
	Interviewer	Mengele; Shabani; PhD1; MSc1	Interviewer	
	Date	Dd/mm/yyyy	Date	Today's date
Interviewee	Name	Free txt	Interviewee_name	
	Role in cattle management	Principle person looking after cattle; owner; occasionally look after cattle; do not look after the cattle	Role	Multiple options
	Gender	Male; female	Gender	
	Level of education	None; primary; secondary, tertiary	Education	Mark highest
	How many years experience keeping cattle?	Number/integer	Experience (years)	
	Have you ever been on a livestock training course for dairy cattle	Yes; No	Training	
If yes to above	What year did you have your training	Integer (4 digits) drop down?	Training year	Enter year
	Are you aware of any diseases you could catch from your cows milk?	Bovine TB, brucellosis, Q fever; RVF, other, none, Don't know	Milk_zoonoses	Tick all listed
	Are you aware of any diseases you could catch from an aborted calf?	Brucellosis, Q fever, leptospirosis, rift valley fever, other, none, Don't know	Abortion_zoonoses	Tick all listed
	Which of the following statements best describes this	A primary income source to owner; secondary income source to the	Reason_own_cattle	

Section	Question	Choices	Label	hint
	herds role for the owner.	owner; just for home consumption and sale to neighbours; Only home consumption		
Herd management	How many heifers and cows do you currently have in the herd?	integer	Herd_size	
	Do you keep your own bull for breeding	Yes; No	Bull	
If yes above	Do you hire out the bull to neighbours?	Yes; No	Bull_hire	
	In the last 12 months have you brought new animals into this herd?	Market; neighbour; none	New_animals	
If yes?	Did you do any pretesting?	Yes; No	Pretest	
	Do you keep sheep at the same household as these cattle?	Yes; No	Sheep	
	Do you keep goats at the same household as these cattle?	Yes; No	Goats	
	Do you keep pigs at the same household as these cattle?	Yes; No	Pigs	
	Do you keep dogs at the household?	Yes; No	Dogs	
	Which option best describes the feeding management	Only zero grazed; generally zero grazed but occasionally graze at pasture; generally grazed at pasture	Management	
	Which option best describes water provision for the herd	Well/bore hole; tap water; river or stream	Water	
	Do you vaccinate the herd routinely against any diseases	Yes; No	Vaccinations	
If yes above	FMDV	Yes; No; Don’t know	FMDV	

Section	Question	Choices	Label	hint
	Brucellosis	Yes; No; Don't know	Brucella	
	Leptospirosis	Yes; No; Don't know	Lepto	
	Pasteurella	Yes; No; Don't know	Pasteurella	
	Black leg	Yes; No; Don't know	Blackleg	
	Anthrax	Yes; No; Don't know	Anthrax	
	Other	Free text	Other_vacc	
	Which option best describes who milks the cows?	Respondent; Owner (if not respondent) Family member; Outside milker/contract milker	Milker	
If outside milker	Does the milker go to multiple farms?	Yes; No	Milker_farms	
	Which best describes preparation of milk from this herd before drinking?	Warm up on fire or stove; bring to the boil on fire or stove; consume without any heating	Milk_prep	
	Who normally assists with calving for the herd?	Respondent; Owner (if not Respondent); Family member; Outside help?	Calving_assist	
	How do you normally dispose of the after birth/placenta after a calving	Leave for cow to eat; Burn; Bury; Throw on rubbish heap; Feed to other animals (dogs/pigs)	Placenta	
	Has any cow aborted in the last 12 months as far as you are aware?	Yes; No; Don't know	Abortion	
	In your view do you have trouble getting cows in calf?	Yes; No; Don't know	Calving_trouble	
Is yes above	Do you know why you are having this problem?	Free text?	Calving_trouble_reason	
	Do you observe rodents in or around the cattle house?	Yes; No	Rodents	
If yes	Do you use any rodent control?	Yes; No	Rodent_control	

Section	Question	Choices	Label	hint
Genotyped animal	Picture ear tag number	Photo	Ear_tag	Clear photo of ear tag to cross reference with animal_ID
	Last 4 digits on animal ID		Short_ID	Needed to help sample labelling
	Animal Age	Integer	Age (years)	Age in years if known. Leave blank if not known
	Animal breed	Shorthorn; Zebu; Shorthorn-Zebu cross; Shorthorn-Zebu cross Friesian; Shorthorn-Zebu cross Ayrshire; Shorthorn-Zebu cross Jersey; Local-Grade cross; Grade; Ankole; Unknown; Other	Breed	Select one stated by owner
	Dentition score	0;1;2;3;4;5	Dentition	See sheet for dentition scoring
	Body condition score	1; 1.5; 2; 2.5; 3; 3.5; 4; 4.5; 5	BCS	See chart for BCS scoring
	Animal Sex	Male; Female; not evaluated	Sex	Select one
If female	Which option best describes this cow?	Heifer; Cow with 1 or more calves;	Cow_age	
If female	When was she last served	Never; Month/year	Service	
If had 1 or more calves				
	How many calves has this cow given birth to alive	Integer	Calf_number	
	When did she last calve?	Month/year	Calf_date	

Section	Question	Choices	Label	hint
	Which option best describes the last calf?	Normal healthy; borne weak but survived; born weak and died within first month; don't know	Calf_status	
	Which option best describes getting the cow back in calf after the last calving?	Not yet put to the bull; put to the bull but not pregnant; put to the bull and pregnant	Calving_status	
	Which option best describes her current pregnancy status?	Don't know; Inseminated but not sure if pregnant; Pregnancy tested positive; Pregnancy tested negative	Pregnancy_status	
	Has this cow ever aborted/premature dead calf?	Yes; No; Don't know?	Abortion_status	
If yes above	When did she abort/have premature calf	Month/year	Abortion_date	
	Genital discharge	No genital discharge; Serous; Mucoid; Puruent; Bloody; Other; Not Evaluated	Genital_discharge	
	Udder condition	Normal; Mastitic; Flabby; Other	Udder_status	
	Milk consistency	Normal; Bloody; Mucoid; Purulent; Other	Milk_status	
	Does the animal appear to be drooling	Yes; No; Not evaluated	Salivation	
	Does the animal appear lame or unwilling to move	Yes; No; Not evaluated	Lameness	
	FMD-like lesions	Mouth; Feet; Mouth and feet; None; Not evaluated	FMD_lesions	Useful to know for risk of spread to next far and to take risk mitigating action

Section	Question	Choices	Label	hint
Sample collection	Serum Sample 1	Yes; No	Serum1	
	Serum Sample 1 Barcode		Serum_code1	Scan bar code and hand write 4 digit animal ID and date on tube
	Reason for not collecting blood sample		Serum1_reason	
	Serum Sample 2	Yes; No	Serum2	
	Serum Sample 2 Barcode		Serum_code2	Scan bar code and hand write 4 digit animal ID and date on tube
	Reason for not collecting blood sample		Serum2_reason	
	EDTA Sample 1	Yes; No	EDTA	
	EDTA Sample 1 Barcode		EDTA_code	Scan bar code and hand write 4 digit animal ID and date on tube
	Reason for not collecting blood sample		EDTA_reason	
	Vagina swab Sample 1	Yes; No	Swab	
	Vaginal Swab Sample 1 Barcode		Swab_code	Scan bar code and hand write 4 digit animal ID and date on tube
	Reason for not collecting vaginal swab		Swab_reason	
Please estimate distance to next dairy farm	Please estimate distance to next dairy farm	Less than 100 m; 100-500 m, more than 500 m	Distance	
GPS northing				

Section	Question	Choices	Label	hint
GPS easting				

RESEARCH OUTPUTS

(i) Publications

Mengele, I. J., Shirima, G. M., Bronsvoort, B. M., Hernandez-Castro, L. E., & Cook, E. A. (2023). Diagnostic challenges of brucellosis in humans and livestock in Tanzania: A thematic review. *CABI One Health*, (2023), ohcs20230001.

Mengele, I. J., Akoko, J. M., Shirima, G. M., Bwatota, S. F., Motto, S. K., Hernandez-Castro, L. E., Komwihangilo, D. M., Lyatuu, E., Bronsvoort, B. M. D. C., & Cook, E. A. J. (2024). Brucella species circulating in smallholder dairy cattle in Tanzania. *Pathogens*, *13*(9), 815.

Mengele, I. J., Shirima, G. M., Bwatota, S. F., Motto, S. K., Bronsvoort, B. M. D. C., Komwihangilo, D. M., Lyatuu, E., Cook, E. A. J., & Hernandez-Castro, L. E. (2023).The status and risk factors of brucellosis in smallholder dairy cattle in selected regions of Tanzania. *Veterinary Sciences*, *10*(2), 155.

(ii) Poster presentation