Host genetic factors in susceptibility to mycobacterial disease in the African buffalo, *Syncerus caffer*.

by
Nikki le Roex

Dissertation presented for the degree of *Doctor of Philosophy (Molecular Biology) in the Faculty of Medicine and Health Sciences at Stellenbosch University*

Supervisor: Prof Eileen Hoal
Co-supervisor: Prof Paul van Helden; Assoc Prof Ad Koets

April 2014

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes 3 original papers published in peer-reviewed journals or books and 3 unpublished publications. The development and writing of the papers (published and unpublished) were the principal responsibility of myself and, for each of the cases where this is not the case, a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors.

Date: 2 October 2013

Summary

Bovine tuberculosis (BTB) is a chronic, infectious disease found in domestic livestock and wildlife, and has serious biodiversity, economic and public health implications. African buffalo act as a wildlife reservoir of BTB, maintaining and transmitting the disease within the environment. The research presented in this thesis addresses the role of host genetic variation in resistance to BTB infection in African buffalo, and reviews the possible practical application of such information. Annual BTB prevalence within the African buffalo population in Hluhluwe iMfolozi Park, South Africa, was evaluated over a seven year period in order to define the extent of *M. bovis* infection. Prevalence changes over time suggest that the test and cull operation currently in place is performing successfully with respect to the original aims of the programme. A review of genetic studies of BTB in livestock and wildlife collated previous findings in this field and provided a collection of possible candidate genes and variants. It also highlighted a lack of research in wildlife, and the limitations of working with species with insufficient genetic data. To overcome the absence of whole-genome data, next-generation sequencing was performed on nine African buffalo, in order to identify novel genetic variants in this species. Upwards of 76 000 novel SNPs within gene regions were identified, and subsequent fluorescent genotyping of 173 SNPs showed a 57% validation rate. From the validated set, 69 SNPs located in genes related to the immune system were selected for association testing with BTB status in African buffalo, and were fluorescently genotyped in 868 individuals. Three SNPs, in the Solute Carrier family 7, member A13 (SLC7A13), Deleted in Malignant Brain Tumour-1 (DMBT1) and Interleukin 1 alpha (IL1α) genes, were identified as significantly associated with BTB status. Very little sequence information of the NRAMP1 (SLC11A1) gene was obtained from the next-generation sequencing performed, and this gene has been associated with brucellosis, salmonella and paratuberculosis in other animal species, making it an excellent candidate for BTB resistance. To characterise this gene in African buffalo, Sanger sequencing was performed to generate the complete coding region, and partially sequence the 5'UTR, intronic and 3'UTR regions. Fifteen novel polymorphisms and three microsatellites were identified within the gene. Finally, a review was prepared to assess the applicability of genetic information on BTB resistance to selective breeding programmes for African buffalo. Phenotypic, marker-assisted and genomic breeding strategies were discussed, with particular emphasis on their suitability to African buffalo. Identifying genes and variants involved in BTB resistance in African buffalo provides potential targets for drug or vaccine development, as well as information that could be incorporated into selective breeding programmes. This may support new management options for controlling the BTB epidemic in the game parks of South Africa, as an alternative to, or in conjunction with, lethal control.

Opsomming

Beestuberkulose (BTB) is ‘n chroniese, aansteeklike siekte wat in vee en wild voorkom en wat ernstige gevolge vir die ekonomie, biodiversiteit en openbare gesondheid inhou. Die Kaap-buffel is ‘n wild reservoir vir BTB wat die siekte onderhou en versprei in die omgewing. Die navorsing wat in hierdie tesis aangebied word fokus op die rol van gasheer genetiese variasie in die weerstand teen BTB infeksie in Kaap-buffels en gee ‘n oorsig van die moontlike praktiese toepassing van die resultate. Die jaarlikse BTB voorkomsyfer in die Kaap-buffel bevolking in die Hluhluwe iMfolozi Park in Suid-Afrika is oor ‘n tydperk van sewe jaar geëvalueer om die omvang van *M. bovis* infeksie te bepaal. Die verandering in voorkomsyfer oor tyd dui daarop dat die toets-en-slag operasie wat tans gebruik word die oorspronklike doelwitte van die program suksesvol bereik. ‘n Oorsig en vergelyking van vorige genetiese studies van BTB in vee en wild het ‘n versameling van moontlike kandidaatgene en –variante verskaf. Dit het ook die gebrek aan navorsing in wildediere uitgewys en die navorsingsbeperkinge wanneer ‘n spesie met onvoldoende genetiese data bestudeer word benadruk. Aangesien daar nie heel genoom data beskikbaar is nie, is volgende-generasie volgordebepaling van 9 Kaap-buffels gedoen om nuwe genetiese variasies in hierdie spesie te identifiseer. Meer as 76 000 nuwe enkel-nukleotied polimorfismes (ENPs) binne geen-areas is geïdentifiseer en die daaropvolgende genotipering van 173 ENPs het ‘n bevestigingskoers van 57% gehad. Vanuit die bevestigde stel ENPs is 69 gekies vir assosiasietoetse met BTB status in die Kaap-buffel en genotipering van 868 individue is gedoen. Drie ENPs, in die *Solute Carrier family 7, member A13 (SLC7A13), Deleted in Malignant Brain Tumour-1 (DMBT1)* en *Interleukin 1 alpha (IL1α)* gene, was beduidend geassosieer met BTB status. Baie min volgorde inligting van die *NRAMP1* (*SLC11A1*) geen is verkry uit die volgende-generasie volgordebepaling. Aangesien hierdie geen voorheen met brucellose, salmonella en paratuberkulose in ander dierespesies geassosieer is, is dit ‘n uitstekende kandidaat vir BTB weerstand. Hierdie geen is in Kaap-buffels gekarakteriseer deur Sanger volgordebepaling van die volledige koderende, gedeeltelike 5’UTR, introniese en 3’UTR areas te doen. Vyftien nuwe polimorfismes en drie mikrosatelliete is geïdentifiseer. Ten slotte is ‘n oorsigstudie gedoen om die toepaslikheid van BTB genetiese weerstandsdata in selektiewe telingsprogramme van Kaap-buffels te evalueer. Fenotipiese, merkerbemiddelde en genomiese teling strategieë is bespreek, met spesifieke klem op die geskiktheid van die metodes vir Kaap-buffels. Identifisering van gene en variante wat betrokke is by BTB weerstand in die Kaap-buffel bied potensiële teikens vir medikasie of entstof ontwikkeling, sowel as inligting wat in selektiewe telingsprogramme gebruik kan word. Dit kan nuwe bestuursopsies vir die beheer van die BTB-epidemie in die parke van Suid-Afrika bied as 'n alternatief vir, of in samewerking met, dodelike beheermetodes.

Acknowledgements

There are many people I wish to thank for their support during this time. To Prof Paul van Helden and Prof Eileen Hoal, for providing me with this opportunity, and for their continued guidance and enthusiasm. Their passion for wildlife and conservation brought about this project and for that I am exceedingly grateful. Thank you to Dr Ad Koets, my co-supervisor in The Netherlands, for support and helpful discussion from afar.

To the funding bodies that made this adventure possible: WOTRO (Science for Global Development), for supporting both this work, and me, for the duration, and the Research Development and Support department of Stellenbosch University, for funding various travel opportunities around the globe.

To the wonderful team at Veterinary Wildlife Services in the Kruger National Park, for the BTB brainstorming sessions, interesting management discussions, and for allowing me to take part in the exciting capture operations. And most of all, in conjunction with eKZN Wildlife, for the collection of blood samples over the years that made this study possible

To everyone in Office 416 – for the fun chats, motivational trips to Cups ‘n Saucers, and of course, for all the advice, support and ideas, both work and non-work related. And to Nico GvP for the great biweekly coffee meetings and life discussions.

To all my family and friends, who have been so supportive and encouraging for the last 4 years. And for those friends, old and new, who have indulged and forgiven my thesis-finishing craziness and general absenteeism for the last 6 months.

And finally, I would like to acknowledge coffee. Without which, I would not be here today.

Abbreviations

ANP	Addo National Park
BCG	Bacille Calmette-Guerin
BTB	Bovine tuberculosis
CA	Brucellosis
CD	Corridor disease
CMI	Cell-mediated immunity
DNA	Deoxyribonucleic acid
FMD	Foot and mouth disease
GLTP	Greater Limpopo Transfrontier Park
HIP	Hluhluwe iMfolozi Park
HIV	Human immunodeficiency virus
IGRA	Interferon gamma release assay
IFNγ	Interferon gamma
IUCN	International Union for the Conservation of Nature
KNP	Kruger National Park
MHC	Major Histocompatibility Complex
NRAMP1	Natural resistance-associated macrophage protein 1
PPD	Purified protein derivative
QTL	Quantitative trait locus
RFLP	Restriction fragment length polymorphism
SCITT	Single comparative intradermal tuberculin test
SNP	Single nucleotide polymorphism
TB	Tuberculosis
VNTR	Variable number of tandem repeats
YBP	Years before present

Contents

Declaration		ii
Abstract		iii
Opsomming		iv
Acknowledgements		v
Abbreviations		vi
Contents		vii
Chapter 1	General Introduction	1
Chapter 2	Disease control in wildlife: evaluating a test and cull programme for bovine tuberculosis in African buffalo.	24
Chapter 3	Bovine TB in livestock and wildlife: what's in the genes?	47
Chapter 4	Novel SNP discovery in African buffalo, *Syncerus caffer*, using high-throughput sequencing.	64
Chapter 5	Gene polymorphisms in African buffalo associated with susceptibility to bovine tuberculosis infection.	87
Chapter 6	The NRAMP1 gene of the African buffalo: characterisation and variation.	108
Chapter 7	Breeding TB-resistant African buffalo: the future of BTB management?	120
Chapter 8	Summarising discussion	134
Appendix A	Contribution to publications	141

Chapter 1

General Introduction

1. African buffalo (*Syncerus caffer*)

1.1 General

Buffalo are one of the most abundant and ecologically important species of megafauna in the savannah ecosystem. They are gregarious grazers that, because of their size, can process taller grasses than most other species, thus fulfilling a potentially facilitative role for the smaller grazers [1]. They inhabit most vegetation types where permanent water is found, and are considered a riverine species, preferring areas close to rivers or lakes. They are an important prey species, and host a vast array of nematodes, pathogens and infectious diseases [1,2]. Population reductions or loss in this species would thus have significant consequences for the ecological balance of the ecosystem, as the African buffalo can be considered a keystone species in the savannah biome.

Buffalo become sexually mature at around 3 years of age, and cows may have their first calf as early as 4 years old. Due to the age dominance hierarchy of the herd, male buffalo are unlikely to mate before 7 or 8 years of age [3]. The gestation period for buffalo is around 11 months, and evidence suggests that mating and/or parturition may be timed with arrival of the wet season, taking advantage of the emergence of new grass with high protein content [4]. The main factors influencing buffalo mortality are drought and predation, primarily by lions, and their maximum life expectancy under natural conditions is suggested be around 20 years [3].

1.2 Population Biology

African buffalo are found in herds ranging from a few hundred to several thousand individuals. Herd size varies between geographical areas, and is influenced by vegetation density, food abundance and water availability [4]. During the breeding season large mixed herds are seen, while smaller mixed herds and bachelor groups are more common during the remainder of the year. Buffalo herds have home ranges that range from 5-100 km^2 in size, and display a predominantly male-dominated hierarchy, although adult female buffalo form the core of the herd. Short and long-range dispersal is undertaken by both male and female buffalo at all times of the year [1,4,5].

Numerous studies have investigated the population differentiation and genetic diversity in African buffalo populations from eastern and southern Africa, using markers such as

microsatellites, MHC, mitochondrial DNA and Y chromosomal loci [2,6–8]. Population differentiation is consistently very low and comparable with other migratory species, which suggests that there is significant dispersal and migration between herds and/or populations in buffalo. Simonsen et al. (1998) found microsatellite and mitochondrial diversity to be comparably high across buffalo populations within a number of African countries, and significant genetic differentiation could be consistently detected only at a continental level. Buffalo also showed comparable heterozygosity when compared to five other African bovid species, and the highest within population nucleotide diversity [5]. Overall, buffalo show low population differentiation and high levels of genetic variability for the MHC DRB3 gene, microsatellites and mitochondrial DNA.

1.3 History and Distribution

The high genetic diversity exhibited by African buffalo suggests that the ancestral population size was extremely large. Evidence of a massive population decline of 75-95% in the East African buffalo population was found for the mid-Holocene period, approximately 4500 YBP, which coincides with the aridification of Africa [9]. This shift in climatic conditions is suggested to be the cause of this enormous population decline, as buffalo have been shown to be extremely sensitive to drought conditions [2]. The concurrent Neolithic expansion of humans and their cattle in Africa, characterised by a shift to pastoralism and agriculture, may also have contributed to the decline in the buffalo population as a result of the farmers‟ acquisition of land and the introduction of cattle-borne diseases [10].

Prior to the arrival of European settlers, the distribution of the African buffalo extended across the majority of southern Africa, Angola, central and East Africa to the southern borders of Ethiopia and Sudan [11,12]. In 1889, rinderpest (a virus carried by cattle) was introduced into north-eastern Africa by humans and led to a continent-wide pandemic that killed 90-95% of the buffalo population. This situation was exacerbated by a foot and mouth outbreak, and resulted in the isolation of the remaining buffalo in South Africa into three remnant populations in the Kruger National Park (KNP), Hluhluwe iMfolozi Park (HiP) and Addo National Park (ANP) [11,13]. Despite this historical population crash, no evidence of a genetic bottleneck has been found in buffalo from South Africa or other African countries. This suggests that either the rinderpest epidemic was not as devastating as previously thought, or that factors such as rapid growth, gene flow and high genetic diversity masked the detrimental effects [2,5,11]

The current distribution of African buffalo extends throughout much of sub-Saharan Africa, although within this range the distribution is fragmented and largely confined to protected areas (Figure 1). The International Union for the Conservation of Nature (IUCN) SSC Antelope Specialist Group classified African buffalo as „Least Concern" in 2008, which reflects an improvement in status since the 1996 „Lower Risk/Conservation Dependant" designation. This is a result of their widespread distribution and global population size, which was estimated at approximately 900 000 individuals in 1999, with savannah buffalo accounting for around 830 000, and forest buffalo about 60 000 individuals [14]. The two largest free-ranging buffalo populations in South Africa at present are those found in the KNP/Greater Limpopo Transfrontier Park (GLTP), and HiP.

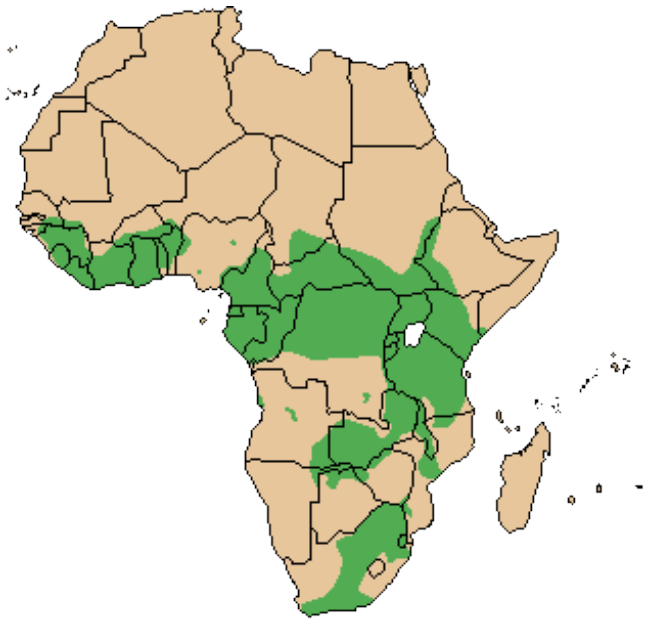

Figure 1: Extant distribution of the African buffalo (*Syncerus caffer*) [14].

1.4 Phylogeny

African buffalo belong to the family Bovidae (Mammalia, Artiodactyla, Ruminantia), which comprises more than 137 extant species classified into 45 genera [15]. The Bovidae family is a diverse group which includes domestic animals such as goats, sheep and cattle, large plains antelopes, and small forest dwellers. A conservative approach proposes that the 12 recognised tribes within the Bovidae family be grouped into six subfamilies, namely Bovinae,

Cephalophinae, Hippotraginae, Alcelaphinae, Antilopinae and Caprinae. Phylogenetic evidence suggests that these subfamilies form two distinct clades, a Bovinae clade (containing the Bovinae subfamily) and an Antilopinae clade (comprised of all the remaining non-Bovinae subfamilies). Within the Bovinae subfamily, the buffalo (*Syncerus* and *Bubulas*) and cattle (*Bos* and *Bison*) species appear to form two distinct Bovini tribes (Figure 2), along with the Boselaphini and Tragelaphini [15].

There are three recognised subspecies of African buffalo – the Cape buffalo (*Syncerus caffer caffer*), the West African buffalo (*Syncerus caffer brachyceros*) and the Forest buffalo (*Syncerus caffer nanus*). Both the Cape and West African buffalo are found in African savannah regions, while the forest buffalo inhabits the rainforests of western and central Africa [16,17]. A fourth subspecies, the Central African Savannah buffalo (*Syncerus caffer aequinoctialis*), has also been proposed. The majority of the population structure, ecology and genetic research on African buffalo thus far has been done on the Cape buffalo subspecies.

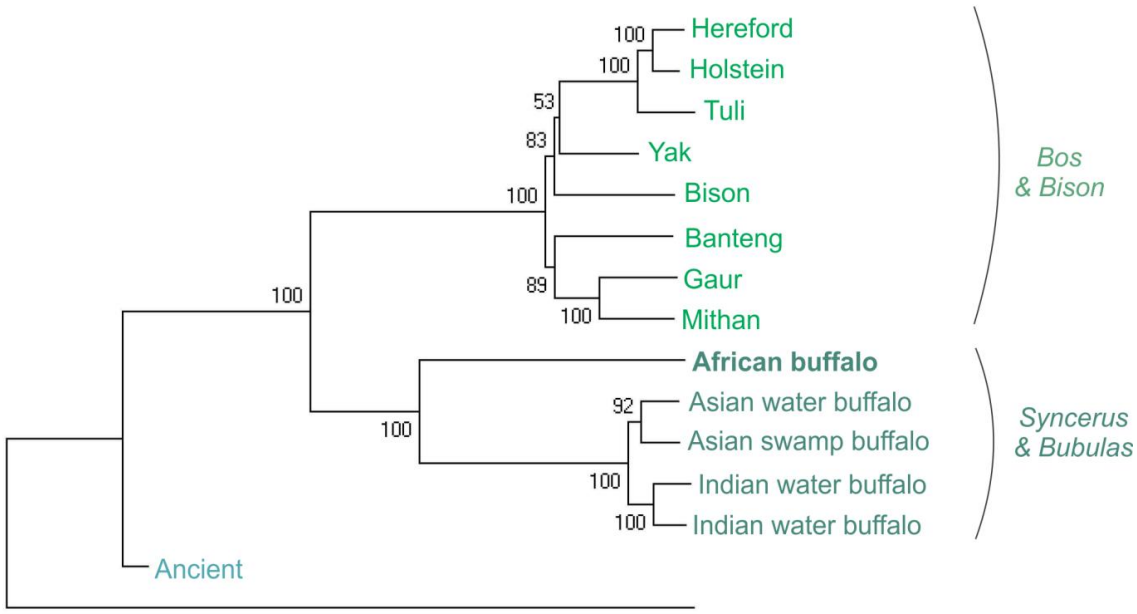

Figure 2: Neighbour joining analysis for members of the Bovini tribe (adapted from MacEachern et al., 2009).

1.5 Disease Risk Management (South Africa)

The South African Department of Agriculture, Forestry and Fisheries established the Disease Risk Management Directive (2002) to provide guidelines for disease risk management in buffalo [19]. This document outlines the regulations with regard to buffalo movement, breeding and disease-testing within the provinces and disease zones in South Africa. The role of buffalo as a maintenance host and reservoir of foot-and-mouth disease (FMD), corridor disease (CD), bovine tuberculosis (BTB) and brucellosis (CA) make these diseases the primary focus of the control policy.

Any land in South Africa used for keeping buffalo must be registered, and the disease status of the buffalo must be disclosed. All registered farms and their fences must be inspected annually by a State Veterinarian, and all buffalo must be tested for all relevant diseases before movement. For BTB, the SCIT (single comparative intradermal tuberculin) test is the required standard of testing. In BTB endemic areas, three to five consecutive negative skin tests are required 3 months apart, while in quarantine, before an animal is accepted as BTB free. Outside endemic areas, the interferon-gamma test may be used to determine disease status. The breeding of disease-free buffalo in South Africa for stocking other farms and conservancies is considered essential, given the current BTB prevalence and spread in certain areas. The movement of all buffalo and buffalo products, to and from breeding facilities, is strictly controlled with permits.

1.6 Economic significance

African buffalo are a species of great economic significance. Designated one of the „Big Five" most dangerous African animals, the buffalo plays an important role in the tourism industry. The ecotourism industry in South Africa alone generates billions of rands annually, so the importance of the Big Five cannot be underestimated. Wildlife utilization in the form of trophy hunting is also a lucrative industry, and sustainable hunting tourism, or conservation hunting, is becoming widely accepted as an important tool in conservation, job-creation and rural development [20,21]. Buffalo are the most economically important species in the hunting industry in southern Africa, as they are one of the most sought after dangerous game trophies, and the addition of buffalo to hunting packages significantly increases their value and marketability [22]. In South Africa, the average price for a buffalo trophy is approximately US $11,000, far exceeding the price in surrounding countries. Lindsey et al.

(2007) calculated that more than 160,000km^2 of land in South Africa is utilised for trophy hunting, hosting approximately 8500 foreign hunters and generating upward of US $100 million revenue per year.

The establishment of breeding programmes for hunting, ecotourism and the stocking of other parks and conservancies create a constant demand for buffalo, and prices have continued to increase accordingly. The market value of buffalo within South Africa also varies according to disease status, with „disease-free" buffalo fetching approximately 10 times that of their counterparts at auction. As demand and prices for „disease-free" buffalo (free of BTB, corridor disease, brucellosis and foot and mouth disease) increased, more breeding farms were established to benefit from this opportunity. Today there are more than 130 buffalo breeding farms in the Mpumalanga province alone [23]. The breeding of genetically and phenotypically superior buffalo bulls for hunting purposes places an even higher price on excellence - the auction price of an exceptionally large-horned buffalo bull for trophy breeding recently reached R40 million [24].

2. Bovine tuberculosis

2.1 Origin and History

Evidence suggests that the common ancestor of the *Mycobacterium tuberculosis* complex emerged approximately 40,000 years ago in East Africa. Two independent clades then evolved: one diversified in humans, resulting in the *M. tuberculosis* lineages, and the other spread from humans to animals, resulting in the other members of the complex, including *Mycobacterium bovis* [25]. The disease caused by *M. bovis* infection was termed „bovine tuberculosis" as reference to the principal host of this pathogen, cattle. However, the term is commonly used to describe *M. bovis* infection in any animal species. It has been suggested that BTB emerged in northern Italy and spread to Western Europe and the United Kingdom, whereafter it was transported across the world via infected cattle during the colonial expansions of the Netherlands and the United Kingdom [26,27]. BTB distribution across the world in the latter half of 2012 can be seen in Figure 3.

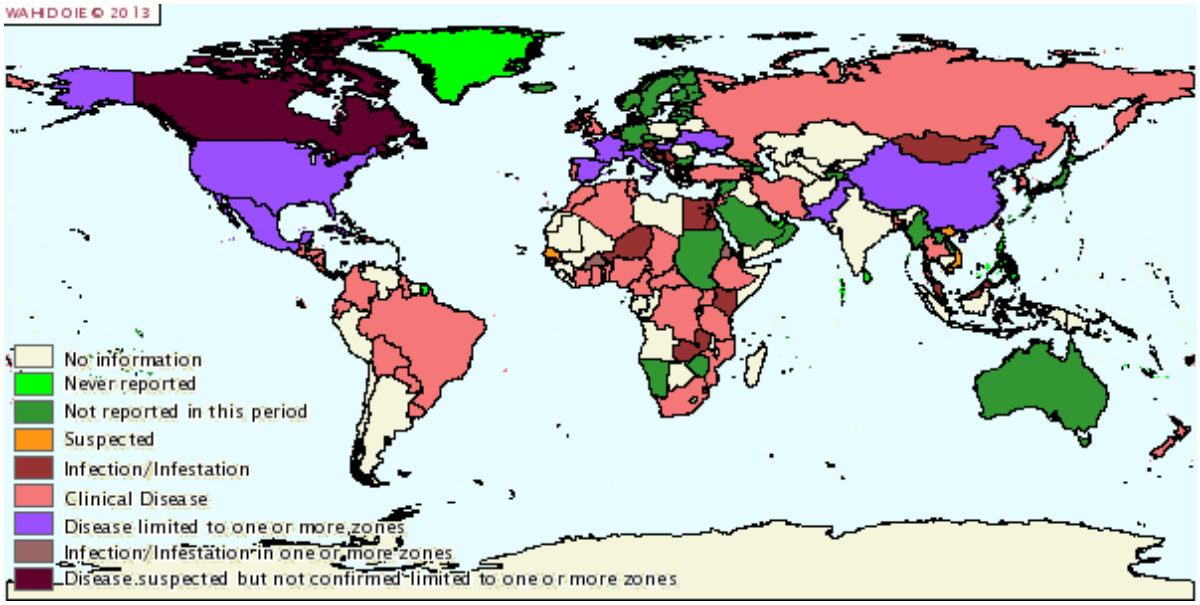

Figure 3: Reported disease information for BTB from July-December 2012, in cattle or wildlife [28].

In South Africa, bovine TB was first reported in cattle in 1880, and in wildlife (Greater kudu, *Tragelaphus strepsiceros*) in the Eastern Cape Province in 1928. Circumstantial evidence indicates that BTB entered the KNP from the south in the late 1950"s or early 1960"s, after buffalo grazed with infected cattle in the Komatipoort/Malelane region [27]. The economic impact of increasing incidence of BTB in cattle led to the implementation of a National BTB control and eradication scheme in 1969 [29]. The first case of BTB in buffalo in HiP was diagnosed in 1986, and in KNP in 1990 [30,31].

2.2 Mycobacterium bovis

Mycobacterium bovis is an aerobic, non-motile Gram-positive bacillus that forms part of the *Mycobacterium tuberculosis* complex (Figure 4). Other members of this complex include *M. tuberculosis*, *M. pinnipedii*, *M. africanum*, *M. microti*, *M. canettii* and *M. caprae* [32]. *M. tuberculosis* is specifically adapted to human hosts, and although it has occasionally been isolated in other animals, it does not appear to sustain itself in these populations, but rather results as an opportunistic infection through the close proximity of humans and animals. *M. pinnipedii*, *M. africanum* and *M. microti* and *M. caprae* are commonly found in seals, primates, voles and goats, respectively [32]. *M. bovis* shows the widest host range of any of the members of the tuberculosis complex; while predominantly isolated from cattle, *M. bovis* has been found to infect many other domestic and wildlife species such as pig, sheep, dogs,

cats, buffalo, badgers, possums, antelope, deer, bison, cheetahs, lions and many others, including humans [33]. It is a hardy pathogen that has been reported to survive in cow faeces for up to 5 months, and in soil for up to 2 years [34].

Despite the different host specificities of the strains, the *M. tuberculosis* complex members show a 99.9% similarity at the nucleotide level, and *M. tuberculosis* and *M. bovis* show a sequence divergence of only 0.05% [32]. The complete genome sequence assembly of *M. bovis* identified approximately 4000 genes within the 4.3 MB genome [35]. When identifying *M. bovis* isolates/strains, the most commonly used DNA typing methods are spoligotyping, which measures the unique number of spacer sequences located in the direct repeat region of the chromosome, and variable number of tandem repeats (VNTR) typing, which measures variation at multiple loci across the chromosome [32]. Michel et al. (2009) examined the genetic diversity of *M. bovis* strains in KNP and HiP (and its associated game reserves) using an integrated genotyping and spatial analysis study. Strain typing of 189 *M. bovis* isolates showed that unrelated strains were responsible for the BTB epidemics in KNP and HiP, which is indicative of two independent introduction events into these parks.

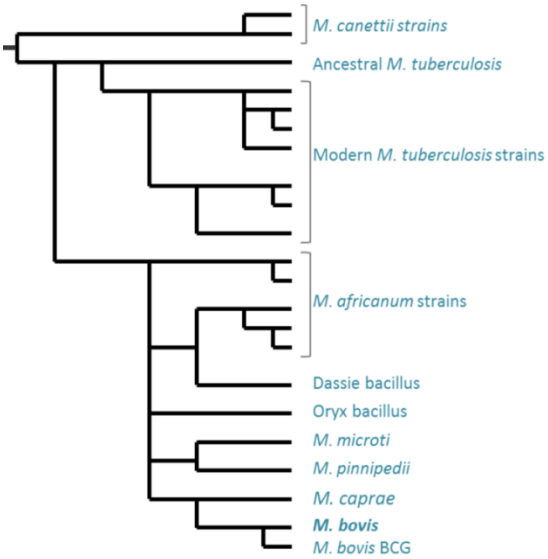

Figure 4: Evolutionary relationships of the *M. tuberculosis* complex (adapted from Huard et al., 2006).

2.3 Transmission and pathogenesis

M. bovis transmission occurs mainly via respiratory or alimentary processes. Respiratory transmission occurs when open pulmonary lesions result in mycobacteria becoming aerosolized in the respiratory tract and transmitted via breathing or coughing. This is the most common form of transmission within human and cattle populations [34,38]. Alimentary transmission occurs either via the uptake of mycobacteria-contaminated material such as sputum, draining sinuses, urine or faeces, or the consumption of primary infected material such as meat or internal organs. This is the main route of infection in carnivores. Percutaneous transmission has also been documented in kudu, where contaminated thorns scratch the skin of the animal and introduce mycobacteria in this manner, and in large predators with the introduction of *M. bovis* into fight wounds [27].

In cattle, it is generally accepted that infection occurs primarily through inhalation of aerosolised bacteria that lodge within the respiratory tract, most likely the alveolar surface of the lung. The bacteria are then phagocytosed by macrophages, the preferred host cell of mycobacteria, eliciting the innate immune response. The aggregation of macrophages and T lymphocytes at the site of infection results in the formation of granulomas, or „tubercles" [34]. The innate immune response includes, but is not limited to, the production of lysosomal hydrolases, bactericidal peptides and toxic reactive oxygen. If these measures fail to destroy the bacilli, the macrophages process the mycobacterial antigens and present them to the T-lymphocytes [38,34]. This stimulation of T-lymphocytes activates the acquired immune response and causes the release of IFNγ and stimulation of the CD8+ T cell cytolytic activity, which enhance the defence system [39]. The cell-mediated immune (CMI) response is most active in earlier stages of diseases, and wanes as disease progresses, while antibodies develop with increased bacterial load and disease progression (Figure 5). Under particular circumstances, such as disseminated TB, cattle can become anergic and have no detectable CMI response, although they still exhibit high levels of antibodies [38].

Buffalo generally display a lesion development pattern that is consistent with aerosol infection, thus inhalation is assumed to be the primary route of transmission. BTB positive buffalo typically develop lesions in the lymph nodes of the head, tonsils, lungs and thoracic lymph nodes within the first 3-6 months of infection, after which the infection spreads by local expansion or through the blood/lymph vessels to more distal sites. Lesions may progress

to caseous necrosis, cavitation and liquefaction, at which stage buffalo are considered super-infective [27].

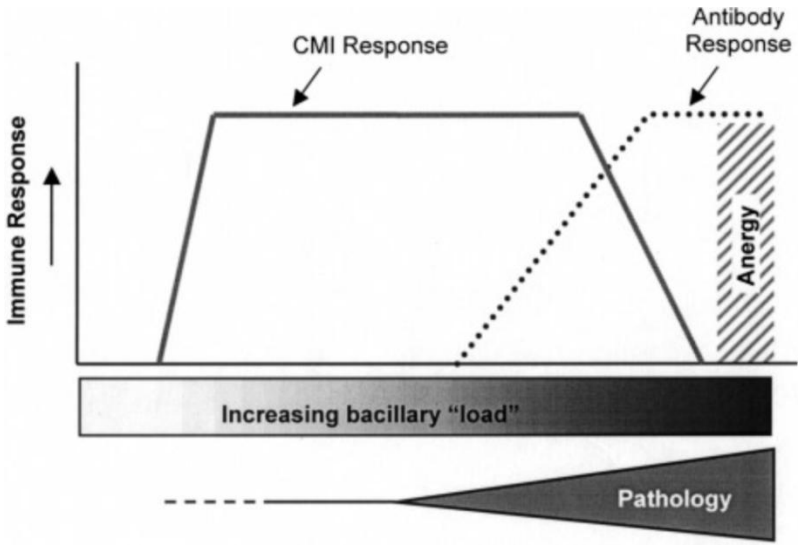

Figure 5: Comparison of the CMI and antibody immune responses with increasing disease pathology in cattle [38].

2.4 Maintenance and spill-over hosts

A species is considered a reservoir/maintenance host for a particular disease when it maintains an infection through horizontal transfer within a population in the absence of other sources of infection [27]. Many countries have successfully eradicated or significantly reduced bovine tuberculosis in cattle within their borders, but the effectiveness of BTB control programmes is drastically reduced in countries that have a wildlife maintenance host [40]. Buffalo are particularly well suited to act as wildlife maintenance hosts due to aspects of their population dynamics and ecology, such as their social herd structure, large population size and high susceptibility to BTB, all of which promote easy transmission of *M. bovis* between individuals [40,41]. BTB maintenance hosts have been identified in other countries, and include the brushtail possum (*Trichosurus vulpecula*) in New Zealand, the Eurasian badger (*Meles meles*) in the United Kingdom and the European wild boar (*Sus scrofa*) in Spain [42].

Maintenance hosts enable the persistence of BTB within wildlife populations, and the transmission of *M. bovis* between wildlife species and to livestock. Spill-over hosts are those that acquire the infection from maintenance hosts, but have a limited ability to maintain the disease within the population in the absence of repeat infection events [27]. Many spill-over hosts of BTB have been discovered, including baboons (*Papio ursinus*), leopards (*Panthera pardus*), cheetah (*Acinonyx jubatus*), impala (*Aepyceros melampus*), kudu (*Tragelaphus strepsiceros*), warthogs (*Phacochoerus africanus*) and lions (*Panthera leo*) [29]. Established transmission routes of bovine TB within South Africa can be seen in Figure 6.

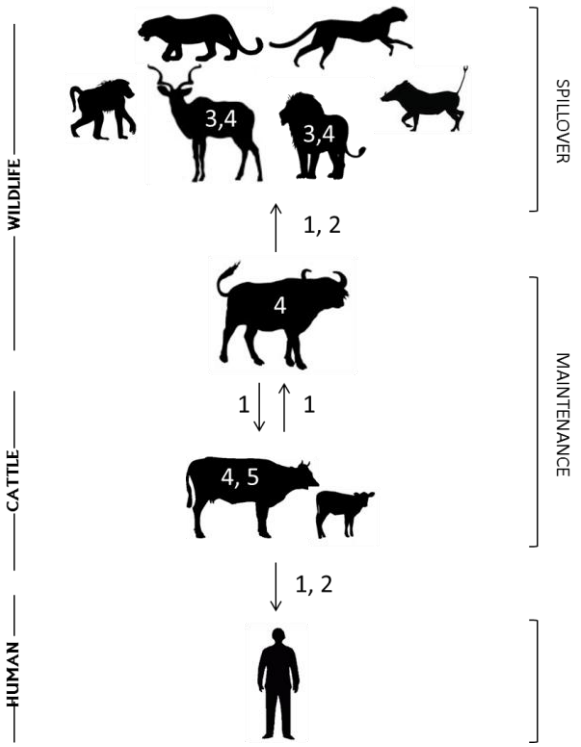

Figure 6: Established transmission routes between the maintenance and spill-over hosts in a South African context. 1–aerosol transmission, 2-ingestion, 3-percutaneous transmission, 4–horizontal transmission, 5–vertical transmission. Adapted from Biet et al. (2005).

2.5 Impact of infection

The role of buffalo as pathogen reservoirs within the game reserves of South Africa could result in major economic, ecological and public health impacts, on other animal species as well as the surrounding livestock and farming communities [34,43]. Economic losses are

suffered in the agricultural sector through the loss of productivity or slaughter of infected livestock, as well as by the negative impact of *M. bovis* infection on the international trade of animals and their products [34,40,44]. The maintenance of BTB also results in enormous economic losses through the continuing need for a BTB eradication scheme, which includes the cost of the cattle and/or wildlife testing programmes, as well as re-imbursement for slaughtered animals, where countries offer such compensation. The sale of wildlife between game reserves and to zoos around the world is compromised or prevented by the risk of BTB movement; at the very least, expensive testing and quarantine measures must be implemented. This is also problematic for species for which there is no established diagnostic test available, requiring a huge investment in research and development of tests prior to animal sale [26]. Spill-over of BTB into other wildlife species of ecological and economic importance, such as the lion, is of additional concern. Predators become infected either through the consumption of infected prey, or through aerosol transmission during prey capture. Lions are severely and fatally infected by BTB, and as a key predator species, this may have enormous consequences for the ecosystem [29].

Zoonotic diseases such as BTB can also be transmitted from animals to humans, which poses a growing public health concern. The incidence of TB in humans caused by *M. bovis* is accepted to be negligible in the developed world, but the situation is very different in developing countries. The close proximity of human, wildlife and livestock at interface zones, the regular consumption of unpasteurised milk and other dairy products, and the lack of BTB testing and/or control in many cattle herds make the risk of TB in humans due to *M. bovis* in developing countries significantly greater [26,42]. This is particularly concerning in areas with a high HIV prevalence, such as the rural communities surrounding many of South Africa‟s national parks, since HIV patients have significantly increased susceptibility to tuberculosis. This may also place an increased burden on health care systems [26,29,40,42].

3. Host Genetics

3.1 General

Investigating the genetic variation found between individuals, populations or species facilitates the understanding of the heritability of traits and phenotypes, and is a commonly utilised approach in conservation, disease susceptibility, and molecular ecology studies [45].

In terms of disease, there are many aspects involved in the outcome of infection by a pathogen. The host genome is one such factor, and plays an important role in determining whether an exposed individual will remain uninfected, become infected but remain asymptomatic, or become infected and progress to active disease. Other variables include environmental factors and the genome of the pathogen, and together these elements determine the disease risk profile of an individual (Figure 7) [44] . It is now widely accepted that the variability seen in disease response in cattle, both between breeds and between individuals, has a genetic component. Concerns about vaccination and/or drug therapy in livestock animals has generated increased interest in the development of alternative, non-chemical means of disease prevention, such as the breeding of genetically resistant animals. For many diseases, the genetic resistance component is polygenic, i.e. under the influence of many genes, and thus the resistance bred into a breed or population is unlikely to be absolute. However, breeding for genetic resistance could be a powerful tool when used in conjunction with other management strategies [47].

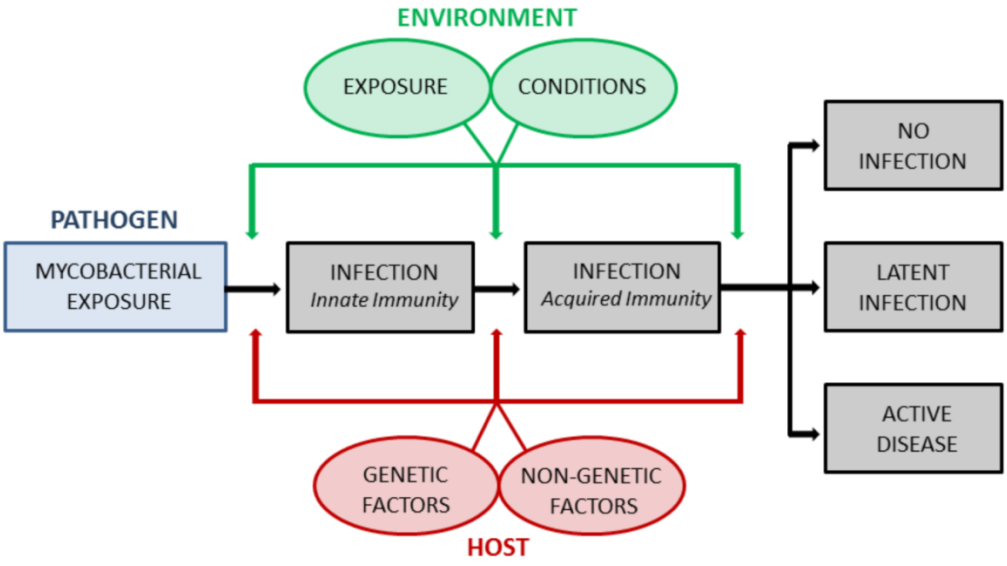

Figure 7: Factors influencing the disease profile of an individual following mycobacterial infection (adapted from Casanova and Abel, 2002).

While health and reproductive traits have been selected for in breeding programs for many years, it is only more recently that regional or national schemes have begun to incorporate

disease traits. Despite relatively low heritability estimates, breeding programs have been successfully implemented in cattle for resistance to diseases such as mastitis (Scandinavia) and brucellosis (Texas). The decreasing cost of molecular genotyping will facilitate the large-scale typing of Quantitative Trait Loci (QTL) and DNA markers, making genetic selection for disease resistance using breeding programs a more feasible option in the future. This principal could be applied to the breeding of buffalo with greater BTB resistance [47].

3.2 SNP discovery

The genome assembly of many different species has provided a wealth of genetic information for use in research. However, when the species of interest does not have an assembled genome and there is only limited genetic information available, sufficient data must first be generated in order to identify the genetic variants present in a population [49,50]. Sequence information can take many forms – microsatellite loci, mitochondrial DNA, single nucleotide polymorphisms (SNPs) and restriction fragment length polymorphisms (RFLPs) are all commonly used types of genetic data, and each has specific advantages and disadvantages. African buffalo studies to date have primarily utilised markers such as microsatellites, mitochondrial DNA and Y chromosomal loci to answer population and genetic diversity questions [7,2].

For association studies in disease resistance, SNP data is likely to be the most informative, as particular genes known to be involved in the immune system can be targeted. SNPs also have numerous other benefits: they are abundant, easy to score and enable high-throughput screening [51]. Large-scale SNP discovery in the African buffalo would greatly facilitate further research in disease genetics, as well as inform other studies. The development of next-generation sequencing technologies has enabled a faster and more cost-effective approach to large-scale SNP discovery. The sequence reads obtained from next-generation sequencing platforms provide a plentiful source for the identification and characterisation of novel SNPs, which can then be used for future applications. For non-model organisms, however, there must be a closely related species with an assembled genome against which the sequence reads can be mapped [45,52].

3.3 BTB genetic studies

Experimental animal models, using gene knockout, recombinant and transgenic individuals, have provided the foundation for the elucidation of complex disease traits, such as TB. Early

studies into BTB resistance utilised inbred strains of rabbits to demonstrate the heritability of BTB resistance/susceptibility. After infection with *M. bovis*, the rabbits displayed either a resistant or susceptible phenotype, suggesting that this phenotype was genetically inherited [53]. Mouse models have also been extensively utilised in TB studies. Gros et al. (1981) discovered that susceptibility to the *M. bovis* BCG (Bacille Calmette-Guerin) strain was controlled by a single genetic factor. This gene was named *Bcg*, which later became the natural resistance-associated macrophage protein gene, *Nramp1*. As early as the 1930‟s, evidence suggested that different cattle breeds showed differential susceptibility to BTB. Zebu cattle were shown to have increased resistance to BTB in both India and Uganda, and within breeds, Simmental and black pied lowland cattle showed differential BTB susceptibility in offspring from different sires [55]. More recent studies on host genetic susceptibility to BTB in both wildlife and cattle are extensively covered in Chapter 3.

4. Disease Prevalence and Management

4.1 BTB Control strategies

Many developed countries have had great success in eradicating or drastically reducing BTB prevalence in cattle using regular tuberculin test and slaughter policies. Repeated test and cull events have resulted in the essential eradication of bovine TB in Australia, Switzerland, and Canada. Unfortunately, this is often not affordable or feasible in many developing countries, and these efforts are also severely compromised in any country with a wildlife maintenance host of BTB [56,26]. Other constraints in the eradication of this disease include the limitations and availability of diagnostic tests and the control of animal movements and trade [57]. Standard test and slaughter schemes using the tuberculin skin test involve the confinement of animals for 72 hours until the test results can be read, after which BTB positive individuals are destroyed, and BTB negative animals are released. One of the main disadvantages of this policy is the low or variable sensitivity of the tuberculin test – low sensitivity estimates mean that infected individuals can be missed and released back with the herd. In addition, there are concerns that the close confinement of a herd for 72 hours, under stressful conditions, may exacerbate the transmission of BTB during that time [27].

The use of a vaccine to control the spread and infection of BTB would provide a publicly acceptable and feasible method of control, as an alternative to, or in conjunction with, lethal sampling. The BCG vaccine has been extensively tested in humans and cattle and is safe, inexpensive, and suitable for oral delivery, all of which make it appropriate for wildlife [56]. However, the vaccine sensitises animals to tuberculin, which can compromise the effectiveness of future diagnostics. The protection it confers is not complete, but in a wildlife context, incomplete protection may be enough to substantially slow the spread of disease. The delivery of a vaccine can be problematic - while oral bait has met with limited success in some wildlife species, it is far more challenging for those in protected areas and/or where supplementary feeding is inappropriate [56]. Vaccination also does not hold equal promise for all species. Vaccination in possums showed a 95% efficacy, while vaccination studies in cattle have found highly variable protective immunity [56]. In buffalo, the BCG vaccine showed no significant protection against experimental challenge with *M. bovis*, and consequently this does not seem to be a viable control strategy for this species at this time [58].

4.2 Diagnostics

The accuracy of the diagnostic method used to detect *M. bovis* infection is crucial in managing BTB in cattle and wildlife. The most commonly used method in South Africa is the tuberculin skin test, specifically the single comparative intradermal tuberculin (SCIT) test. The SCIT test measures dermal swelling due the cell-mediated immune response 72 hours after the injection of purified protein derivative in the skin of the neck or caudal fold. Comparisons are made between the reaction to avian and bovine tuberculin, to eliminate cross reactions due to exposure to other non-tuberculous mycobacteria. Disadvantages of the SCIT test include the difficulty/logistics of administration, the variability in interpretation of results and the low degree of standardisation. However, despite these limitations, it remains the most accurate test available, and shows a sensitivity of 55.1-93.5% and a specificity of 88.8-100% in cattle [57]. Unpublished data suggests that the SCIT test in buffalo may have a sensitivity and specificity of 80.9% and 90.2%, respectively, but this has yet to be officially validated [31].

The interferon-gamma (IFNγ) test is a lab-based test that measures the *in vitro* cell-mediated immune response by quantifying the relative IFNγ levels after stimulation with avian and bovine tuberculin [57]. The advantages of this test are increased sensitivity, rapid testing and

that no follow up is required. However, the disadvantages include reduced specificity, high cost and increased likelihood of non-specific response in young animals. In cattle, sensitivity estimates are 73-100%, and specificity estimates are 85-99.6% [57]. In buffalo, IFNγ release assays (IGRA) have shown sensitivities of approximately 93% and specificities of 68-85% [59]. False positives have been observed when using the standard protocol for IGRA, and this is thought to be a result of sensitisation from environmental mycobacteria. Modification of the standard protocol with the addition of *M. fortuitum* PPD increased the specificity without compromising sensitivity, and this protocol has been used in the Kruger National Park since 2000 [59].

5. Scope of the thesis

The research reported in this thesis addresses the role of genetic factors in the susceptibility to bovine tuberculosis in the African buffalo, and the practical applications of such information.

The annual prevalence of BTB within the African buffalo population in Hluhluwe iMfolozi Park was investigated over a seven year period, at the herd and area level, to identify the extent of the epidemic and investigate changes in BTB prevalence over time (**Chapter 2**). This allowed the evaluation of the test and cull operation that is currently in place in the park, and which parameters of the operation may be the most important. A test and cull operation functions essentially like a small-scale breeding scheme, where susceptible animals are removed over time, and it is important to understand the parameters that are required for this to effect a change.

While much research has been done in characterising the genetic component of susceptibility to human tuberculosis, the same is not true for bovine tuberculosis, particularly in wildlife species. A review of studies of the genetic susceptibility to BTB in cattle and wildlife was compiled in order to efficiently gather all the information in this area, address some considerations that must be taken into account in studies such as these, and make future suggestions for work in this area (**Chapter 3**).

In order to generate new genetic data for this non-model species, next-generation sequencing was performed on African buffalo samples and the sequence reads were mapped to the

domestic cow (*Bos taurus*) genome (**Chapter 4**). This enabled the identification of novel SNPs throughout the buffalo genome. Different software packages were used to compare and optimise the mapping of non-model organism sequences to a reference genome of a different species, as well as to identify the variants found. A subset of the SNPs identified were validated using fluorescent genotyping. A total of 69 candidate polymorphisms, located in genes involved in the immune response, were then selected from the SNP subset validated in Chapter 4. These SNPs were fluorescently genotyped in the full sample set of African buffalo (n=868) and tested for genetic associations with BTB infection status using a case-control approach (**Chapter 5**).

When performing next-generation sequencing, only a percentage of the genome of interest is represented in the final data set. Because of this, specific regions or genes of interest to the researcher may not have been sequenced using this technology, and may benefit from a more thorough characterisation. This characterisation was performed on the African buffalo *NRAMP1* gene, in order to provide a more extensive picture of the variation and structure of this gene (**Chapter 6)**. Novel SNPs identified will then be available for use in future association studies.

Finally, the suitability of using genetic information, such as was identified in this body of work, in a selective breeding program for increasing BTB resistance in African buffalo, was reviewed as an alternative and sustainable disease management strategy (**Chapter 7**). Main lines of evidence for the heritability of BTB resistance were discussed, factors that must be taken into account in selective breeding programmes were considered, and the three main types of breeding programmes were evaluated, with emphasis on how these approaches might be applied to African buffalo.

References

1. Jolles AE (2007) Population biology of African buffalo (*Syncerus caffer*) at Hluhluwe-iMfolozi Park, South Africa. Afr J Ecol 45: 398–406.

2. Heller R, Lorenzen ED, Okello JB, Masembe C, Siegismund HR (2008) Mid-Holocene decline in African buffalos inferred from Bayesian coalescent-based analyses of microsatellites and mitochondrial DNA. Mol Ecol 17: 4845–4858.

3. Pienaar U de V (1969) Observation on developmental biology, growth and some aspects of the population ecology of the African Buffalo (*Syncerus caffer caffer Sparrman*) in the Kruger National Park. Koedoe 12: 29–52.

4. Ryan SJ, Knechtel CU, Getz WM (2006) Range and habitat selection of African buffalo in South Africa. J Wildlife Manage 70: 764–776.

5. Simonsen BT, Siegismund HR, Arctander P (1998) Population structure of African buffalo inferred from mtDNA sequences and microsatellite loci: high variation but low differentiation. Mol Ecol 7: 225–237.

6. Greyling BJ, Kryger P, du Plessis S, van Hooft WF, van Helden P, et al. (2008) Development of a high-throughput microsatellite typing approach for forensic and population genetic analysis of wild and domestic African Bovini. Afr J Biotechnol 7: 655–660.

7. Van Hooft WF, Groen AF, Prins HHT (2003) Genetic structure of African buffalo herds based on variation at the mitochondrial D-loop and autosomal microsatellite loci: Evidence for male-biased gene flow. Conserv Genet 4: 467–477.

8. Wenink PW, Groen AF, Roelke-Parker ME, Prins HH (1998) African buffalo maintain high genetic diversity in the major histocompatibility complex in spite of historically known population bottlenecks. Mol Ecol 7: 1315–1322.

9. deMenocal, Ortiz, Guilderson, Sarnthein (2000) Coherent high- and low-latitude climate variability during the holocene warm period. Science 288: 2198–2202.

10. Heller R, Brüniche-Olsen A, Siegismund HR (2012) Cape buffalo mitogenomics reveals a Holocene shift in the African human-megafauna dynamics. Mol Ecol 21: 3947–3959. doi:10.1111/j.1365-294X.2012.05671.x.

11. Van Hooft WF, Groen AF, Prins HH (2000) Microsatellite analysis of genetic diversity in African buffalo (*Syncerus caffer*) populations throughout Africa. Mol Ecol 9: 2017–2025.

12. Winterbach HEK (1998) The status and distribution of Cape buffalo *Syncerus caffer caffer* in southern Africa. S Afr J Wildl Res 28: 82–88.

13. O'Ryan C, Harley EH, Bruford MW, Beaumont M, Wayne RK, et al. (1998) Microsatellite analysis of genetic diversity in fragmented South African buffalo populations. Anim Conserv 1: 85–94.

14. IUCN Red List (2012). IUCN Red List. Available: http://www.iucnredlist.org/apps/redlist/details/21251/0. Accessed 13 July 2011.

15. Hassanin A, Douzery EJ (1999) The tribal radiation of the family Bovidae (Artiodactyla) and the evolution of the mitochondrial cytochrome b gene. Mol Phylogenet Evol 13: 227–243.

16. Melletti M, Penteriani V, Boitani L (2007) Habitat preferences of the secretive forest buffalo (*Syncerus caffer nanus*) in Central Africa. J Zool 271: 178–186.

17. Melletti M, Penteriani V, Mirabile M, Boitani L (2007) Some behavioral aspects of forest buffalo (*Syncerus caffer nanus)*: From herd to individual. J Mammal 88: 1312–1318.

18. MacEachern S, McEwan J, Goddard M (2009) Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (Bovidae, Bovinae). BMC Genomics 10: 177. doi:10.1186/1471-2164-10-177.

19. Disease Risk Management Directive (2002). National Department of Agriculture, Forestry and Fisheries. Available: http://www.nda.agric.za/.

20. Lindsey PA (2008) Trophy hunting in Sub Saharan Africa: Economic scale and conservation significance. Best Practices in Sustainable Hunting: 41–47.

21. Lindsey PA, Roulet PA, Romañach SS (2007) Economic and conservation significance of the trophy hunting industry in sub-Saharan Africa. Biol Cons 134: 455–469. doi:10.1016/j.biocon.2006.09.005.

22. Booth VR (2009) A comparison of the prices of hunting tourism in southern and eastern Africa. Badapest, Hungary: International Council for Game and Wildlife Conservation and Food and Agriculture Organisation of the United Nations.

23. Laubscher L, Hoffman L (2012) An Overview of Disease-Free Buffalo Breeding Projects with Reference to the Different Systems Used in South Africa. Sustainability 4: 3124–3140. doi:10.3390/su4113124.

24. Rupert Buys R40m Million Buffalo (2013). IOL News. Available: http://www.iol.co.za/business/news/rupert-buys-r40m-million-buffalo-1.1580626#.UkKpMz_hd9A.

25. Wirth T, Hildebrand F, Allix-Béguec C, Wölbeling F, Kubica T, et al. (2008) Origin, Spread and Demography of the *Mycobacterium tuberculosis* Complex. PLoS Pathog 4: e1000160. doi:10.1371/journal.ppat.1000160.

26. Michel AL, Muller B, van Helden PD (2010) *Mycobacterium bovis* at the animal-human interface: a problem, or not? Vet Micro biol 140: 371–381.

27. Renwick AR, White PCL, Bengis RG (2007) Bovine tuberculosis in southern African wildlife: a multi-species host-pathogen system. Epidemiol Infect 135: 529–540. doi:10.1017/S0950268806007205.

28. OIE - World Organisation for Animal Health (n.d.). OIE - World Organisation for Animal Health. Available: http://www.oie.int/wahis_2/public/wahid.php/Diseaseinformation/Diseasedistributionmap.

29. Michel AL, Bengis RG, Keet DF, Hofmeyr M, Klerk LM, et al. (2006) Wildlife tuberculosis in South African conservation areas: implications and challenges. Vet Microbiol 112: 91–100.

30. Rodwell TC, Whyte IJ, Boyce WM (2001) Evaluation of population effects of bovine tuberculosis in free-ranging African buffalo (*Syncerus caffer*). J Mammal 82: 231–238.

31. Jolles AE, Cooper DV, Levin SA (2005) Hidden effects of chronic tuberculosis in African buffalo. Ecology 86: 2358–2364.

32. Smith NH, Gordon SV, de la Rua-Domenech R, Clifton-Hadley RS, Hewinson RG (2006) Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. Nat Rev Microbiol 4: 670–681. doi:10.1038/nrmicro1472.

33. O'Reilly LM, Daborn CJ (1995) The epidemiology of *Mycobacterium bovis* infections in animals and man: A review. Tuber Lung Dis 76: 1–46. doi:16/0962-8479(95)90591-X.

34. Ayele WY, Neill SD, Zinsstag J, Weiss MG, Pavlik I (2004) Bovine tuberculosis: an old disease but a new threat to Africa. Int J Tuberc Lung Dis 8: 924–937.

35. Garnier T, Eiglmeier K, Camus J-C, Medina N, Mansoor H, et al. (2003) The complete genome sequence of *Mycobacterium bovis*. Proc Natl Acad Sci USA 100: 7877–7882. doi:10.1073/pnas.1130426100.

36. Michel AL, Coetzee ML, Keet DF, Maré L, Warren R, et al. (2009) Molecular epidemiology of *Mycobacterium bovis* isolates from free-ranging wildlife in South African game reserves. Vet Microbiol 133: 335–343. doi:16/j.vetmic.2008.07.023.

37. Huard RC, Fabre M, Haas P de, Lazzarini LCO, Soolingen D van, et al. (2006) Novel Genetic Polymorphisms That Further Delineate the Phylogeny of the *Mycobacterium tuberculosis* Complex. J Bacteriol 188: 4271–4287. doi:10.1128/JB.01783-05.

38. Pollock JM, Neill SD (2002) *Mycobacterium bovis* infection and tuberculosis in cattle. Vet J 163: 115–127.

39. Hope JC, Villarreal-Ramos B (2008) Bovine TB and the development of new vaccines. Comp Immunol Microbiol Infect Dis 31: 77–100.

40. Biet F, Boschiroli ML, Thorel MF, Guilloteau LA (2005) Zoonotic aspects of *Mycobacterium bovis* and *Mycobacterium avium*-intracellulare complex (MAC). Vet Res 36: 411–436.

41. Michel AL (2002) Implications of tuberculosis in African wildlife and livestock. Ann NY Acad Sci 969: 251–255.

42. Amanfu W (2006) The situation of tuberculosis and tuberculosis control in animals of economic interest. Tuberculosis (Edinb) 86: 330–335.

43. Van Rhijn I, Godfroid J, Michel A, Rutten V (2008) Bovine tuberculosis as a model for human tuberculosis: advantages over small animal models. Microbes Infect 10: 711–715.

44. Ramirez-Villaescusa AM, Medley GF, Mason S, Green LE (2009) Herd and individual animal risks associated with bovine tuberculosis skin test positivity in cattle in herds in south west England. Prev Vet Med 92: 188–198.

45. Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. Mol Ecol Resour 10: 915–934. doi:10.1111/j.1755-0998.2010.02891.x.

46. Abel L, Dessein AJ (1997) The impact of host genetics on susceptibility to human infectious diseases. Curr Opin Immunol 9: 509–516. doi:16/S0952-7915(97)80103-3.

47. Morris CA (2007) A review of genetic resistance to disease in *Bos taurus* cattle. Vet J 174: 481–491.

48. Casanova JL, Abel L (2002) Genetic Dissection of Immunity to Mycobacteria: The Human Model. Annu Rev Immunol 20: 581–620.

49. Imelfort M, Duran C, Batley J, Edwards D (2009) Discovering genetic polymorphisms in next-generation sequencing data. Plant Biotechnol J 7: 312–317.

50. Tautz D, Ellegren H, Weigel D (2010) Next Generation Molecular Ecology. Mol Ecol 19: 1–3. doi:10.1111/j.1365-294X.2009.04489.x.

51. Morin PA, Luikart G, Wayne RK, the SNP workshop group (2004) SNPs in ecology, evolution and conservation. Trends Ecol Evol 19: 208–216. doi:16/j.tree.2004.01.009.

52. Everett MV, Grau ED, Seeb JE (2011) Short reads and nonmodel species: exploring the complexities of next-generation sequence assembly and SNP discovery in the absence of a reference genome. Mol Ecol Resour 11: 93–108. doi:10.1111/j.1755-0998.2010.02969.x.

53. Lurie MB (1944) Experimental epidemiology of tuberculosis : hereditary resistance to attack by tuberculosis and to the ensuing disease and the effect of the concentration of tubercle bacilli upon these two phases of resistance. J Exp Med 79: 573–589.

54. Gros P, Skamene E, Forget A (1981) Genetic control of natural resistance to *Mycobacterium bovis* (BCG) in mice. J Immunol 127: 2417–2421.

55. Allen AR, Minozzi G, Glass EJ, Skuce RA, McDowell SWJ, et al. (2010) Bovine tuberculosis: the genetic basis of host susceptibility. Proc R Soc B 277: 2737–2745. doi:10.1098/rspb.2010.0830.

56. Buddle BM, Wedlock DN, Denis M, Vordermeier HM, Hewinson RG (2011) Update on vaccination of cattle and wildlife populations against tuberculosis. Vet Microbiol 151: 14–22. doi:16/j.vetmic.2011.02.021.

57. Schiller I, Oesch B, Vordermeier HM, Palmer MV, Harris BN, et al. (2010) Bovine Tuberculosis: A Review of Current and Emerging Diagnostic Techniques in View of their Relevance for Disease Control and Eradication. Transbound Emerg Dis 57: 205–220. doi:10.1111/j.1865-1682.2010.01148.x.

58. De Klerk L-M, Michel AL, Bengis RG, Kriek NPJ, Godfroid J (2010) BCG vaccination failed to protect yearling African buffaloes (*Syncerus caffer*) against experimental intratonsilar challenge with *Mycobacterium bovis*. Vet Immunol and Immunopathol 137: 84–92. doi:10.1016/j.vetimm.2010.04.013.

59. Michel AL, Cooper D, Jooste J, de Klerk LM, Jolles A (2011) Approaches towards optimising the gamma interferon assay for diagnosing *Mycobacterium bovis* infection in African buffalo (*Syncerus caffer*). Prev Vet Med 98: 142–151.

Chapter 2

Disease control in wildlife: Evaluating a test and cull programme for bovine tuberculosis in African buffalo.

Nikki le Roex[1*], Dave Cooper[2], Paul D van Helden[1], Eileen G Hoal[1], Anna E Jolles[3]

1 DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/ MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, Stellenbosch University, Cape Town, South Africa. **2** Ezemvelo Kwazulu-Natal Wildlife, St. Lucia, South Africa. **3** College of Veterinary Medicine and Department of Zoology, Oregon State University, Corvallis, OR 97331, USA.

*Corresponding author nikkileroex@sun.ac.za

Manuscript in prep.

Abstract

Providing an evidence base for wildlife population management is difficult, due to limited opportunities for experimentation and adequate study replication at the population level. We took advantage of a rare opportunity to assess the outcome of a test and cull programme aimed at limiting the spread of an important pathogen, *M. bovis*, in African buffalo. Buffalo act as reservoirs of *Mycobacterium bovis*, the causative agent of BTB, which can have major economic, ecological and public health impacts through the risk of infection to other wildlife species, livestock and surrounding communities. BTB prevalence data were collected in conjunction with disease control operations in Hluhluwe-iMfolozi Park (HiP), South Africa, from 1999-2006. The control programme utilised a mobile boma to capture buffalo, targeting different areas within the park each year. Throughout the study period, 4733 buffalo (250-950 per year) were tested for BTB using the single comparative intradermal tuberculin (SCIT) test, with BTB-positive animals culled, and negative animals released back into the park. BTB prevalence in HiP was extremely spatially and temporally variable, ranging from 2.3% to 54.7%. Geographic area was a strong predictor of BTB transmission in HiP, owing to relatively stable herds and home-ranges. Our data suggests that HiP‟s test and cull programme is effective at reducing BTB transmission in buffalo herds: herds experiencing more intensive and frequent captures showed reduced per capita disease transmission risk and less increase in herd prevalence over time, disease hotspots did not expand spatially over time, and BTB prevalence in all but the hotspot areas was maintained at a prevalence between 10-15% throughout the study period. We hope that future collaborative work between wildlife managers, veterinarians and scientists will build on these encouraging findings to optimise wildlife disease control programmes and mitigate conflict at the interface of conservation, agricultural and urban areas.

Keywords: *bovine tuberculosis, Mycobacterium bovis, African buffalo, prevalence, test and cull, wildlife management, disease control*

1. Introduction

Despite increasing awareness and research in animal diseases, wildlife-livestock-human interfaces continue to expand, and the role of wildlife species in the transmission of pathogens continues to grow [1]. Interface expansion increases with human population

growth and encroachment into wildlife areas, as well as intensifying and expanding livestock practices. These practices will continue to grow over time, thus highlighting the need for cross-disciplinary collaboration in the design and implementation of viable disease management strategies [1]. Commonly referred to as bovine tuberculosis (BTB), *M. bovis* infection is a highly infectious disease that is widespread amongst domestic cattle (*Bos taurus*) and endemic in a number of wildlife populations around the globe, including the brush-tailed possum (*Trichosurus vulpecula;* [2]), European badger (*Meles meles;* [3]), and African buffalo (*Syncerus caffer*; [4,5]). In southern Africa, buffalo act as the primary maintenance host for BTB. Several aspects of their population dynamics and ecology, such as their social herd structure, large population size and high susceptibility to BTB, promote easy transmission of *M. bovis* among buffalo [6,7]. BTB has been detected in several buffalo populations in southern Africa and can reach high prevalence within herds. In the Kruger National Park, zone prevalence ranged from 0-27.1% in 1991/2, and 1.5-38.2% in 1998 [8], and in Queen Elizabeth National Park (QENP) in Uganda, BTB prevalence was 21.6% in 1997 [9]. Studies have shown that buffalo herds with a high BTB prevalence exhibit increased parasite load and reduced body condition, adult survival and fecundity when compared to herds with a low BTB prevalence [4,10,11].

Wildlife pathogen reservoirs for BTB pose a risk of infection to domestic livestock and humans, particularly in developing countries where humans, livestock and wildlife live in close proximity to one another [5,12]. Economic losses are incurred through the loss of productivity and/or slaughter of infected livestock, as well as by the negative impact of *M. bovis* infection on the international trade of animals and their products [7,13,14]. The impact of transmission to other wildlife species has enormous consequences for biodiversity, ecotourism and commercial game farming, and can severely compromise species conservation efforts [5,7,15]. *M. bovis* is also zoonotic and can thus be transmitted from animals to humans, which poses a potential public health concern, particularly in HIV prevalent areas such as southern Africa, since HIV patients have a significantly increased susceptibility to tuberculosis [5,15].

Bovine tuberculosis was first detected in African buffalo in Hluhluwe iMfolozi Park (HiP), located in the KwaZulu Natal region of South Africa, in 1986 [4]. HiP is one of South Africa‟s flagship conservation areas due to its rich mammal and avian fauna, and is the source of most of southern Africa‟s restored white rhino population [16]. HiP‟s ability to export rhino and other animals to protected areas around the subcontinent is thus at the heart

of its biodiversity conservation mission, and BTB poses a clear threat to this mission. Buffalo are considered to be the primary maintenance host of BTB in HiP, with *M. bovis* infection subsequently documented in lion (*Panthera leo*), baboon (*Papio ursinus*) and greater kudu (*Tragelaphus strepsiceros*) [12]. Comprehensive surveys were conducted from 1998 to assess the prevalence of BTB in African buffalo in different regions of HiP (Figure 1). This data identified Masinda and Makhamisa as „trouble" areas with particularly high BTB prevalence, and prompted the development of a test and cull disease control programme that began in 1999.

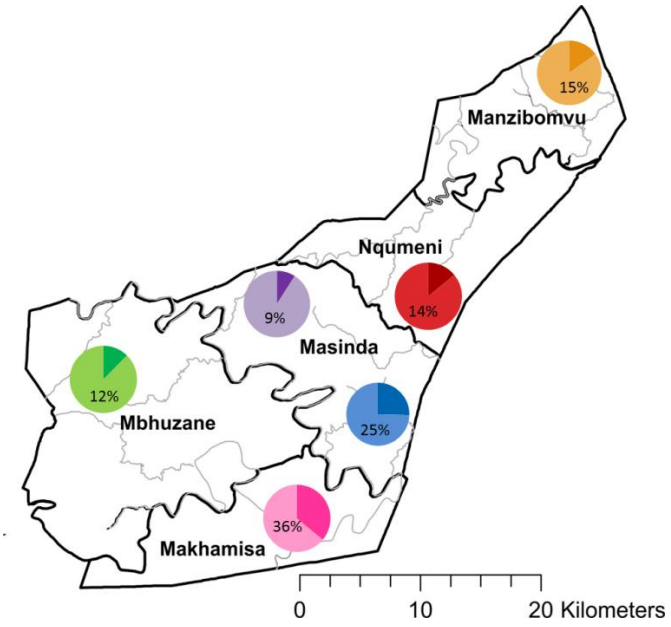

Figure 1: Map depicting regions of Hluhluwe iMfolozi Park. Pie charts depict the proportion of BTB infected and uninfected individuals in the first year of testing in that region, with numbers indicating the percentage infected. Mbhuzane, Makhamisa and Manzibomvu capture sites correspond with area names. Red pie chart = Gunjaneni capture site. Purple pie chart= Corridor capture site (spanning northern Masinda and Nqumeni). Blue pie chart= Masinda south capture site. The total number of animals in the first year of testing, for each area, were: Makhamisa=261, Mbhuzane=420, Masinda south=122, Corridor=520, Gunjaneni=207 and Manzibomvu=132.

BTB is notoriously hard to control in wildlife. Test and cull programmes have been implemented in the United Kingdom for many years in an attempt to control the prevalence and transmission of BTB in the European badger and cattle, but have met with mixed success

[17,18]. Thus far, the only example of complete eradication of BTB was achieved in Australia, where the culling and eradication of an introduced wildlife host was aggressively undertaken. A similar approach in New Zealand, with the culling of brush-tailed possums, has made significant progress, although it has yet to achieve the success of the Australian programme. However, such an aggressive approach to culling (and local eradication of a species) cannot be employed where the host species is endemic, and/or has an economic or social value [19].

HiP‟s BTB control programme aims to limit the prevalence of the disease in buffalo, in an attempt to protect the park‟s other mammal species from infection, and enable continued movement of animals from HiP to other public and privately-owned conservation areas. The programme uses a mobile capture boma to target different areas within the park on a rotational basis. Each area is processed at least every 4-5 years, with more heavily infected areas targeted more frequently. All captured buffalo are tested for BTB, positive animals are culled and animals with a negative test result are released back into the park. The objectives in this study were to determine the annual prevalence of BTB within the African buffalo population in Hluhluwe iMfolozi Park over the seven year period from 1999-2006. Prevalence changes over time were analysed to evaluate the effectiveness of the test and cull programme in limiting BTB prevalence in the buffalo population. Programme effectiveness was addressed with respect to the original aims of the operation, which were: (1) to reduce prevalence in disease hotspots, and (2) prevent increase of BTB in the more moderately affected areas.

2. Materials and Methods

2.1 Study Area, BTB Control Programme and Sampling

Data was collected from Hluhluwe-iMfolozi Park in the KwaZulu Natal province of South Africa from 1999-2006. The park, situated at 28° S and 31-32° E, was established in 1895 and occupies approximately 950km^2. During the 1999-2006 study period, the park‟s buffalo population numbered between 3000-5000 animals [20], which occurred in herds of 70-200 animals occupying well-defined home ranges [21]. Between 250-950 buffalo were tested for BTB annually. Buffalo herds were captured at different focal sites each year, representing distinct areas and buffalo subpopulations within the park - Makhamisa, Mbhuzane, Masinda

south, Corridor (northern Masinda and Nqumeni), Gunjaneni and Manzibomvu (Figure 1). The design and implementation of this disease control program required methods and equipment suitable for mass capture in the field. Moveable bomas were utilised to contain the buffalo herds that were brought in from distances between 0.5-5km away (Figure 2). Animals were spotted from the air, and a helicopter was used to guide the animals into a capture funnel that connected to the boma (Figure 2a). The funnel area was closed off with wire and hardware-cloth curtains as soon as a herd was inside, to prevent the buffalo from running back out (Figure 2b). Once in the boma structure, buffalo were partitioned into small groups and immobilised (Figure 2c) using etorphine hydrochloride (M99) and azaperone according to standard operating procedures (SOPs) developed by KwaZulu-Natal Wildlife‟s game capture veterinarians. Buffalo remained immobilised for approximately 10-30 minutes for BTB testing and data collection, followed by reversal of immobilisation using diprenorphine (M5050).

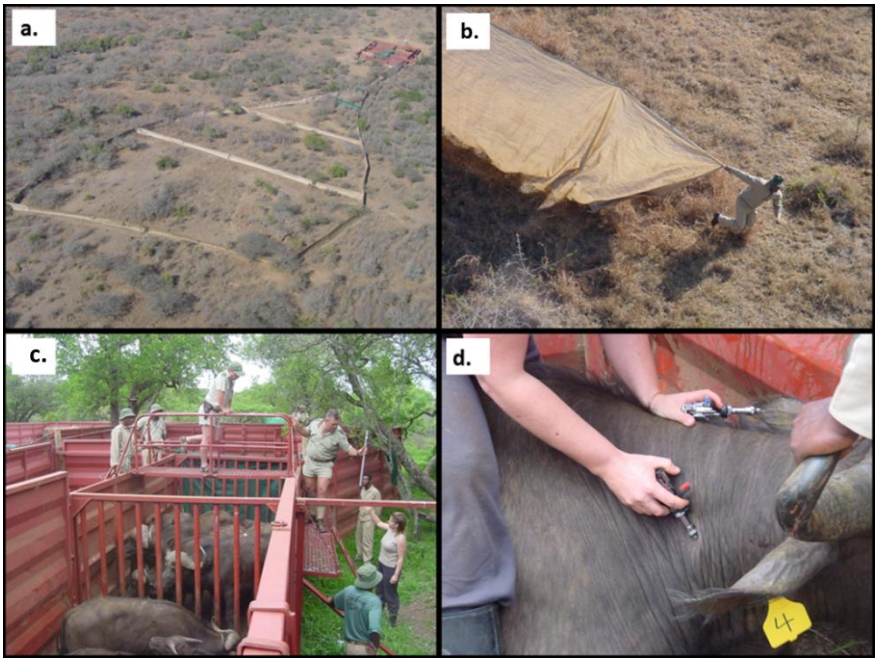

Figure 2: Buffalo capture set-up for the BTB disease control programme at HiP. a. A mobile metal boma and capture funnel are set up within a targeted buffalo population‟s home range. b. The funnel curtains are drawn closed as soon as the buffalo herd has entered the funnel. c. Gangways leading to the holding pens are equipped with sliding doors for separating animals into small groups. Buffalo are immobilized from a bridge, using a pole syringe. d. BTB intradermal skin testing: BTB antigen is injected on the left side of the neck, avian tuberculin is injected on the right side.

Not all animals from a given herd were typically captured, because herds would split into several groups while being herded towards the boma by the helicopter. While the animals were immobilised, each animal‟s age, sex, and any brand marks from previous captures were recorded. Age was estimated using tooth emergence, body size and horn development in animals less than 5 years old. Adult buffalo were aged using tooth wear as an indicator, by measuring the height of incisor 1 [20]. All captured buffalo were branded with a year-specific brand on one buttock, and a herd-specific brand on the other. The herd provenance of previously captured animals could thus be established, and the years of their previous negative BTB tests.

2.2 BTB diagnosis

Buffalo were tested for BTB using the single comparative intradermal tuberculin (SCIT) test. Small amounts of pure protein-derived (PPD) tuberculin prepared from *M. bovis* (PPD-B) and *Mycobacterium avium* (PPD-A) were injected into opposite sides of the neck of the animal (Figure 2d). Skin fold thickness was measured prior to injection and after 72 hours, with swelling indicating prior exposure to *M. bovis* or *M. avium*. A test result was considered positive when the difference in skin fold thickness at the injection site, at the time of injection vs 72 hours later, exceeded 2mm, taking into account the physical characteristics of the reaction. The avian tuberculin is used to distinguish between non-specific mycobacterial exposure and BTB infection, in order to exclude false positive results and thus improve test specificity. Thus, test results were interpreted as negative if swelling in response to avian tuberculin was as pronounced as swelling at the bovine tuberculin injection site. The comparative intradermal tuberculin test has been shown to have a sensitivity of 77-95% and a specificity of 96-99% in cattle [22]. Data from African buffalo, (J.P. Raath, unpubl.) suggest a test sensitivity and specificity of 80.9% and 90.2% in buffalo, respectively [4]. SCIT testing was performed by park and state veterinarians, and upon diagnosis BTB positive animals were slaughtered and BTB negative animals released back into the park.

2.3 Data Analysis

2.3.1 Data Compilation

The herd brands recorded during each capture were used to determine which animals were resampled in multiple years, and thus which sampling groups were associated with one another. Multiple sampling groups were considered to be components of a single herd if five

or more individuals of one group were resampled in another group at a subsequent time point. A sampling group association matrix was compiled to designate herds. Each herd was assigned to an area within HiP based on the capture site of the herd or component groups (Table 1).

Table 1: Areas and herds sampled, by year.

Area	**Herds**							
	1999	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**
Makhamisa					24,25,26,28 **[261]**			
Mbhuzane				K,L,16 **[420]**	K,L **[276]**		L **[205]**	L **[106]**
Masinda south	M,N,46 **[122]**	M, 38 **[260]**	M,N,5,9,11 **[494]**	20,21,23 **[108]**			66 **[62]**	71,72 **[109]**
Corridor	O,P,Q,R, 54,60 **[520]**		O,R **[149]**	P,Q **[274]**			P,Q **[300]**	
Gunjaneni	S,T,48 **[207]**		S,T **[175]**				S **[127]**	S **[72]**
Manzibomvu			12,13,14 **[132]**			31,32,33,34 **[302]**		

*Letters refer to herds sampled at multiple time points; numbers refer to herds only sampled once. Square brackets contain total number of animals tested for that area in that year.

Individuals were assigned to the following age cohorts: 0-1 years, 2-3 years and 4+ years, as this was the highest level of accuracy obtainable from the data when including all capture years. Analyses were performed at both the area and herd levels. A contact network between areas was compiled to depict the number of branded individuals from area X that were subsequently re-captured in area Y, in order to assess the validity of the area-based classifications.

2.3.2 Prevalence and incidence estimation

Prevalence was calculated as the proportion of skin test positive animals out of the total number of animals tested in a given herd or area, and capture year. Herd level prevalence calculations were complicated by the fact that most captures did not include the entire herds,

and captured herd fragments are often biased in age composition (Jolles, pers obs). For example, adult females with calves may tend to run together, or groups of sub-adult animals. To adjust for this possible age bias, herd level prevalence was determined by calculating the age cohort BTB prevalence values for each herd. Herd prevalence was then estimated based on observed prevalence for each age group, weighted according to the age distribution seen in the entire dataset (approximately 4500 individuals). The age distribution observed in buffalo herds is largely insensitive to BTB prevalence [4]; as such, using the whole population average age distribution will not have introduced systematic error with respect to BTB prevalence. Area level prevalence was calculated using all animals within a particular region for each time point. All data reported are apparent prevalence values, as the true prevalence cannot be calculated without published diagnostic sensitivity/specificity data; error bars for prevalence data were generated according to the sensitivity and specificity estimates by Raath (unpubl. data).

Incidence, i.e. the per capita risk of becoming infected with BTB over a given time period, was calculated as the proportion of new cases since last capture. To estimate incidence – a direct measure of disease transmission risk – we needed to distinguish new BTB cases that occurred in the time period between one capture and the next, from older infections that could have occurred prior to the initial capture time point. This was done using individuals that were younger than the time since last capture, and older individuals that were branded, i.e. had been caught and tested before. The branded individuals would have had to be negative at last capture, or they would have been culled. These criteria ensure that the incidence estimates were calculated using only new cases. Incidence over the time period between captures (1-4 years) was thus estimated as the fraction of BTB-positive animals (new cases) in the candidate groups (young and branded) described above.

2.3.3. Spatial and temporal patterns of BTB prevalence

Spatial and temporal patterns of prevalence were evaluated at the area level, because no herds were re-sampled consistently through time. Linear regression was used to detect overall temporal trends in BTB prevalence in the areas sampled over the total time period. Fisher‟s exact tests were used to calculate whether there were significant short-term changes in prevalence within the areas. Geographic variation in BTB prevalence was analysed visually by plotting prevalence over time for all areas (Figure 4a); statistical comparisons between areas are included in the analyses described below.

2.3.4. Analysis of factors driving changes in disease transmission in buffalo herds

General linear mixed models were used to test whether the test and cull programme achieved reductions in disease transmission over time. We assessed two outcomes: (i) incidence (per capita risk of infection) from one capture time point to the next, and (ii) change in herd prevalence from one capture time point to the next. Variability among herd capture events in terms of exposure to the disease control programme was expressed in terms of (a) number of animals captured (capture effort), and (b) time interval between captures. Herd was included in our models as a random effect. Thus, in the incidence analysis we asked: did greater capture effort, and/or more frequent captures, result in reduced BTB incidence? The independent variables considered were:

- Capture effort: the number of animals caught in each herd during the previous time step. If the test and cull programme works, then a more thorough capture, where most animals in the herd are tested, and most BTB-positives removed, should result in reduced BTB incidence subsequently: *We hypothesized that BTB incidence would be negatively associated with capture effort.*
- Capture interval: the number of years between successive test and cull events. The longer the interval between captures, the more opportunities for disease transmission to have accrued over time. *We hypothesized that BTB incidence would be positively associated with capture interval.*
- Area: Different geographic areas within the park vary in their precipitation and temperature patterns, their vegetation and animal densities. Area might therefore affect BTB transmission, and we included area as a candidate variable in our models.

We used Akaike‟s Information Criterion (AIC) to select the models most strongly supported by our data from the set of models, including all combinations of our candidate independent variables. We considered all models within 2 points of the lowest AIC score and selected our final models to include independent variables that were statistically significant at a cut-off value of $\alpha = 0.05$. All analyses were performed in R version 2.15.2 (freely available from www.r-project.org). Reductions in disease incidence are the most direct measure of success for disease control programs; however, our estimate of incidence was based on a sub-sample of captured animals, because of the challenge of distinguishing new from established cases of BTB (see section 2.3.3. above). We ran our analyses with change in herd prevalence as the dependent variable as well, to test whether using an outcome that allowed us to include the

whole set of captured animals would return results that were congruent with the incidence analysis. Change in herd prevalence was computed simply as the difference in observed herd prevalence from one capture time point to the next. We included the same candidate independent variables and modelling approach in our analysis of change in herd prevalence as described above. Analogous to the incidence analyses, we hypothesised that change in herd BTB prevalence would be negatively associated with capture effort, and change in herd BTB prevalence would be positively associated with capture interval.

3. Results

3.1. Change over time in area BTB prevalence

The number of individuals re-captured in their previously assigned area far exceeded the number of individuals re-captured in a different area, for every area combination tested (Figure 3). This supports the area-based classifications used for the subsequent analyses.

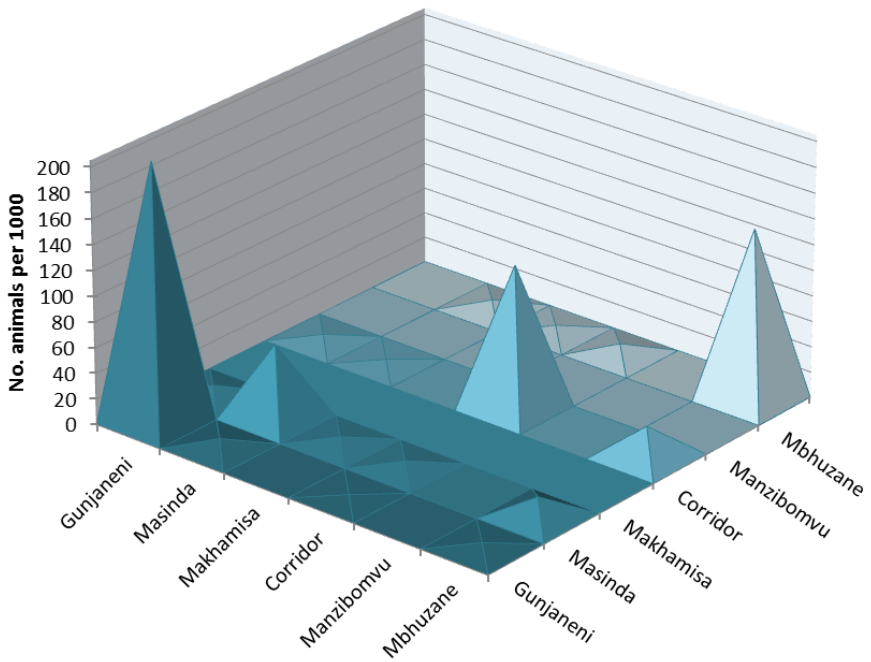

Figure 3: Contact network of animals in different areas. Amount of contact between regions is measured in the number of branded animals from area X re-captured in area Y, per 1000 animals captured in area Y.

Area level prevalence was calculated for the Makhamisa, Mbhuzane, Masinda south, Corridor, Gunjaneni and Manzibomvu areas and ranged between 2.3% and 54.7% (Table 2). The south-west area of the park clearly stood out as a BTB hotspot. In Masinda south, BTB prevalences varied between 12 and 54%. BTB at Masinda increased in prevalence from 1999 to 2000 (Fisher‟s exact test, $p<0.0001$), and decreased from 2001 to 2002 ($p<0.0001$), after which Masinda‟s area prevalence stabilized. The neighbouring Makhamisa area also showed high BTB prevalence (36%), but was sampled only once during the study period (due to difficult access in this road-less wilderness area). Areas outside the south-west corner of the park – Mbhuzane, Gunjaneni, Corridor and Manzibomvu showed more moderate BTB prevalences, mostly below 15%. BTB prevalence at Mbhuzane remained fairly stable around 10% throughout the study period, as did BTB prevalence in the Corridor area, except for an unusually low prevalence of 5.5% observed in 2002 (prevalence reduction 2001-2002: Fisher‟s exact test, $p=0.034$; prevalence increase 2002-2005, $p=0.044$). Gunjaneni showed a similar pattern (prevalence reduction 1999-2001, Fisher‟s exact test, $p<0.0001$; prevalence increase 2001-2005, $p<0.0001$). Manzibomvu was sampled only twice during the study period. BTB prevalence was moderate with a decrease from 2001 to 2004 (Fisher‟s exact test, $p=0.0136$). Linear regression indicated no significant temporal trend for any of the area datasets over the total time period (Supplementary table S1).

Table 2: Prevalence statistics for the HiP areas

Year	Makhamisa		Mbuzane		Masinda		Corridor		Gunjaneni		Manzibomvu	
	Prev	**N**	**Prev**	**N**	**Prev**	**N**	**Prev**	**N**	**Prev**	**N**	**Prev**	**N**
1999					0.254	122	0.090	520	0.140	207		
2000					0.473	260						
2001					0.547	494	0.114	149	0.023	175	0.152	132
2002			0.124	420	0.120	108	0.055	274				
2003	0.360	261	0.120	276								
2004											0.073	302
2005			0.112	205	0.145	62	0.103	300	0.205	127		
2006			0.066	106	0.138	109			0.125	72		

*N: total number of animals used for the prevalence calculations in each area at each time point

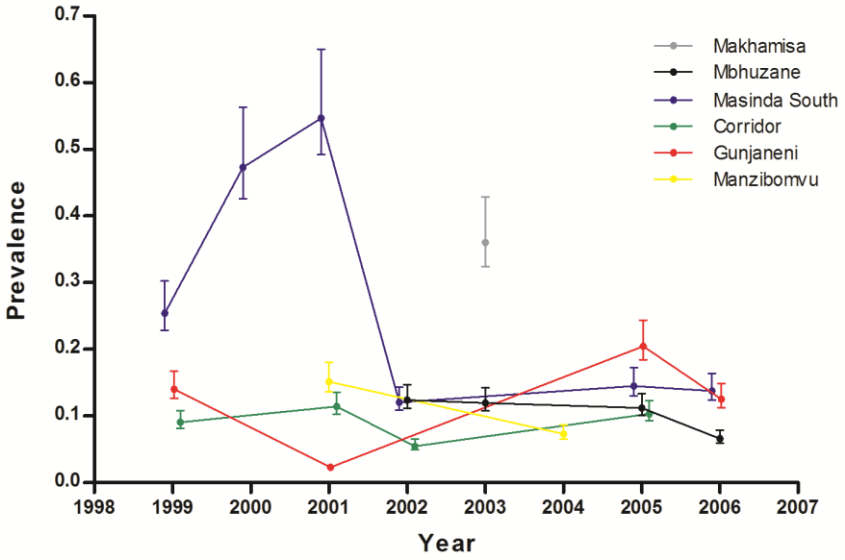

Figure 4a: Prevalence of BTB infection from 1999-2006, in six areas at HiP. Vertical lines denote standard error estimates.

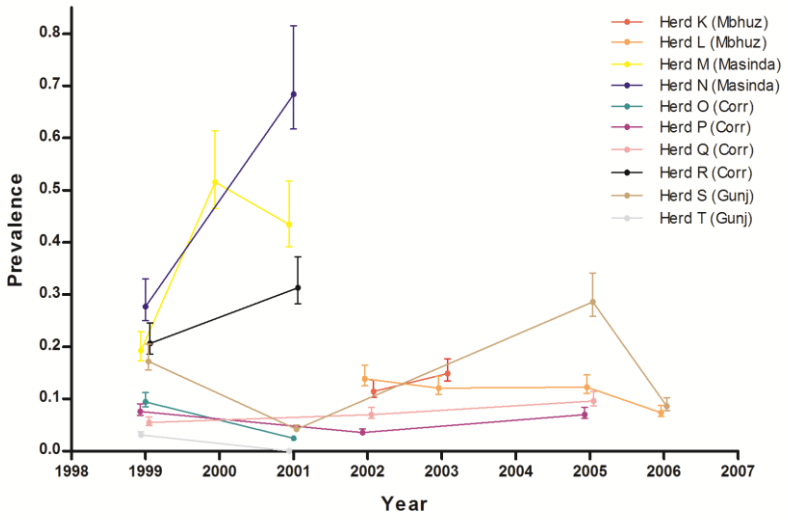

Figure 4b: Observed prevalence of BTB infection from 1999-2006 in ten buffalo herds. Vertical lines denote standard error bars.

3.2 Does the test-can-cull programme reduce BTB transmission in buffalo herds?

Our dataset of herds that were resampled over time consisted of ten herds. Five herds were sampled twice (resulting in a single line of data, each, for change in BTB prevalence, or incidence), three herds were sampled three times (resulting in 2 lines of data each), and two herds were sampled four times (three lines of data), i.e. a total of 17 data points from the ten herds. AIC values are shown in supplementary tables S2 and S3, for BTB incidence and change in herd prevalence, respectively.

For BTB incidence, our final model included area, capture effort and capture interval as independent variables (Table 3). BTB incidence was much higher in the southern Masinda area than in the Corridor, Gunjaneni or Mbhuzane. Concordant with our hypothesis, capture effort was negatively associated with BTB incidence – i.e. a larger number of buffalo being tested per herd was associated with a lower subsequent disease incidence. Capture interval was supported as a model parameter by AIC and returned a positive estimate (as hypothesised), but had no statistically significant effect on BTB incidence.

Table 3: Results of general linear mixed model evaluating BTB incidence in HiP buffalo herds.

	Estimate	**MCMC mean**	**95% CI**	**pMCMC**	**Pr(>|t|)**
Intercept	12.950	9.7967	-8.669, 31.087	0.3040	0.1499
capt. interval	3.801	4.2164	-2.383, 11.1987	0.2014	0.1182
capture effort	-0.105	-0.0967	-0.181, -0.0122	0.0282	0.0139
area: Gun	-4.495	-0.0001	-14.569, 15.0531	0.9986	0.6318
area: Mas	52.427	54.0948	37.371, 71.2783	0.0002	0.0002
area: Mbu	13.607	14.6796	-2.972, 33.8154	0.1048	0.1857

The models for change in herd BTB prevalence yielded very similar results, with the best model also including capture interval, capture effort, and area (Table 4). Herds residing in Masinda south and Mbhuzane showed larger increases in prevalence from one capture to the next than herds in Gunjaneni or Corridor. Capture interval was positively associated with change in herd prevalence (Figure 5a), i.e. the greater the time interval between successive captures, the more herd BTB prevalence increased from the initial time point to the subsequent capture. Finally, capture effort was negatively associated with change in

prevalence, meaning that a more intensive test and cull effort was associated with a reduced increase (or possibly decrease) in herd BTB prevalence (Figure 5b).

Table 4: Results of GLMM evaluating change in BTB prevalence in HiP buffalo herds.

	Estimate	MCMC mean	95% CI	pMCMC	Pr(>|t|)
intercept	-13.694	-13.5203	-32.7768, 4.9969	0.1416	0.1365
capt. interval	11.713	11.6819	5.1939, 18.0682	0.0020	0.0024
capture effort	-0.108	-0.1087	-0.1912, -0.0306	0.0120	0.0125
area: Gun	-4.434	-4.6271	-18.0372, 9.2556	0.4672	0.4688
area: Mas	30.628	30.5115	14.7904, 46.7031	0.0018	0.0012
area: Mbu	20.556	20.3045	3.5970, 37.9821	0.0290	0.0212

Overall, these results highlight massive variability in BTB transmission dynamics among geographic areas in HiP, and support the hypothesis that more intensive disease control efforts (in terms of test and cull programme frequency and efficacy) were associated with reduced BTB transmission in HiP buffalo herds.

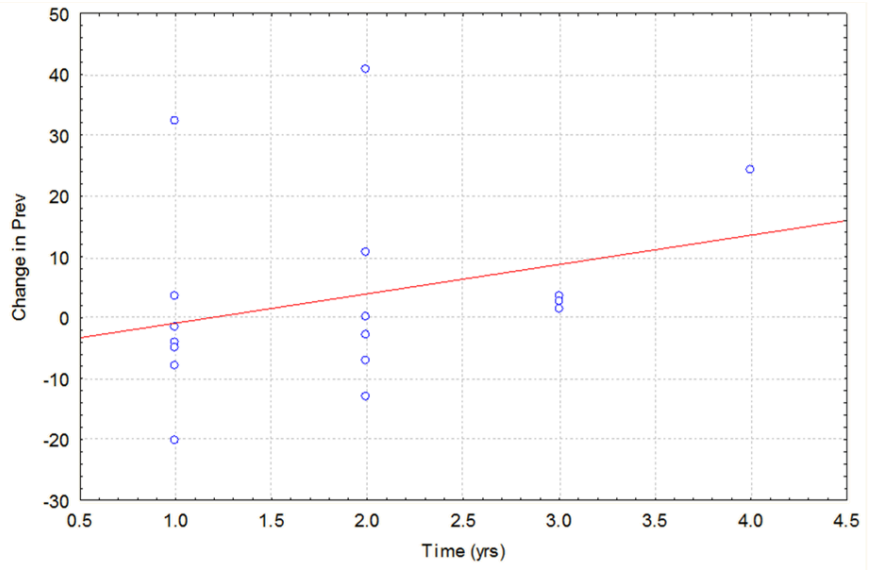

Figure 5a: Positive association between capture interval and change in prevalence. Change in Prevalence = -5.684 + 4.8123*x.

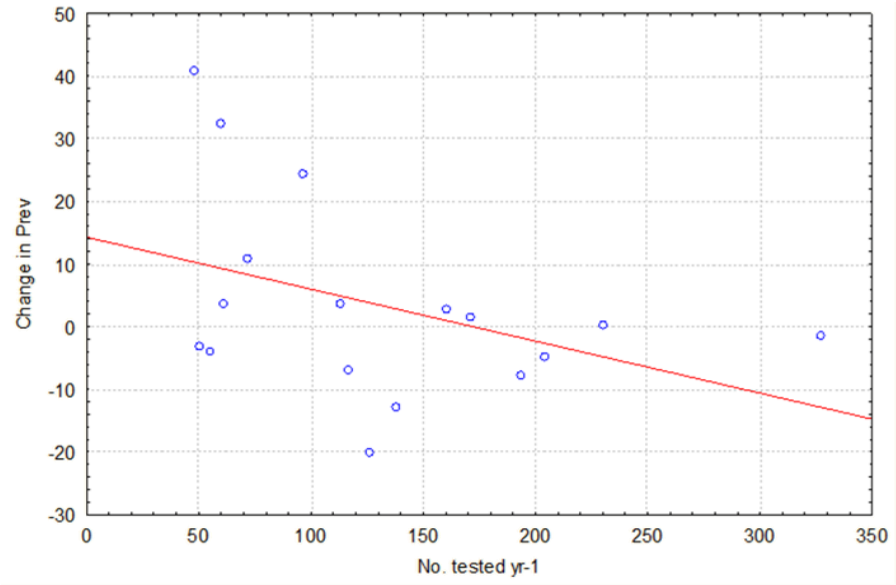

Figure 5b: Negative association between capture effort and change in prevalence. Change in Prevalence = 14.2685 - 0.083*x.

4. Discussion

BTB prevalence in HIP was extremely spatially and temporally variable, ranging between 2.3% and 54.7%. Geographic area was a strong predictor of BTB transmission in HiP, as assessed by our estimates of BTB incidence and change in BTB prevalence in buffalo herds. This may appear surprising, given the relatively small size of the reserve, and ability of buffalo to travel long distances. However, previous studies of the spatial ecology of buffalo at HiP, based on extensive radio-tracking, also concluded that buffalo herds in HiP appear to be remarkably stable and occupy well-delineated home ranges (Jolles, unpubl., [21]). Our analysis of recapture locations further supports that buffalo tend to remain within their original subpopulations at HiP. Variation in BTB transmission among geographic areas has many possible causes. Variability in microclimate, substrate and habitat can affect host population contact patterns and susceptibility to infection, pathogen persistence in the environment, and exposure to parasites that might affect BTB transmission [23,24]. BTB prevalences observed at HiP bracket the range of prevalences seen in other buffalo populations in southern and East Africa (e.g. Kruger National Park, South Africa, 0-27.1% in 1991/2, 1.5-38.2% in 1998 [8]; Queen Elizabeth National Park, Uganda, 21.6% in 1997 [9]). Although comparisons across studies are complicated by differences in time frame and

diagnostic methods used, it is clear that some areas within HiP had unusually high BTB prevalences, highlighting the fact that *M. bovis* is capable of infecting a very substantial proportion of animals in some buffalo populations.

The data suggests that HiP‟s test and cull programme is effective at reducing BTB transmission in buffalo herds: herds experiencing more intensive and more frequent captures showed reduced disease transmission risk and lower increase in herd prevalence over time, disease hotspots did not expand spatially over time, and BTB prevalence (in all but the hotspot areas) was maintained at a low prevalence between 10-15% throughout the study period. HiP‟s test and cull programme for BTB control, with its innovative, field-based mass capture technique for buffalo, thus represents a rare success story in the management of wildlife disease. However, these results should be interpreted with caution, particularly in contexts beyond HiP‟s buffalo population, for a number of reasons. First, despite our massive sampling effort, our dataset of re-sampled herds varying in disease control intensity was small, with only 17 data points from 10 buffalo herds: only moderately strong conclusions can be drawn from this data. Second, even though intensive, frequent captures may be most effective at reducing BTB transmission, they also reduce buffalo numbers [20]. Disease management must therefore operate within tight bounds if buffalo are to be maintained as an important ecological component of a reserve‟s mammal fauna. Optimising efficiency of the disease control programme to achieve the greatest reduction in BTB transmission without compromising buffalo population viability is a challenging goal. Future work explicitly modelling BTB dynamics under different disease control scenarios – varying the frequency, extent, spatial targeting of captures, and management action triggers – might assist with progress towards it. Third, test and cull disease management can have unintended consequences, which need to be weighed against the benefits when evaluating the merits of this management strategy. Captures are extremely stressful to the animals and can disrupt their social structure – both of which may add to their infection risk after being released back into the park [3,25]. This may be especially problematic in areas with high disease prevalence, where released susceptible animals experience a high encounter rate with infected buffalo. Comprehensive coverage of high prevalence hotspots may be essential to circumvent this pitfall. Moreover, any large-scale capture operation of wild animals will incur some rate of injury or death to the animals. HiP‟s buffalo capture operation has evolved over more than a decade, and every effort is made to keep the animals unharmed. As a result, animal loss rates during capture are very low (Jolles, unpubl.). Nonetheless, losses, including deaths

post-release, are difficult to assess with certainty. Perhaps of more concern, recent genetic evidence suggests that buffalo removals associated with BTB control may reduce immunogenetic variation in the population (de Lane-Graaf et al., *in review*). If so, management of disease may carry costs in terms of adaptability to future parasite and pathogen challenges – an undesirable outcome for modest-sized, isolated populations. Finally, even though the test and cull program may help limit BTB transmission, it is unlikely to be able to eradicate the disease, not least because of unknown diagnostic test sensitivity. Thus, park management are potentially faced with an indefinite, costly commitment to disease control. As such, economic, conservation and veterinary considerations must be carefully weighed in deciding what disease control strategy is optimal – a formidable challenge in the face of financial and ecological uncertainty.

Acknowledgements

Field work was conducted by Jen Britt, Shawn Cassatt, Cathy Dora, Vanessa Ezenwa, Tiffany Garcia, Flo Gardipee, Cleo Gosling, Alice Henshaw, Kim Kanapeckas, Liz Kleynhans, Mvuyi Matokazi, Zodwa Mnyandu, Thando Nomatshila, Mike O'Brien, Dave Paoletti, Maxime Roulet, Nikki Stevens, Wendy Turner, Sithembile Vilikazi, Matt Waldram, Michelle Walters, and many shorter-term volunteers. Many thanks to Marina Nel, Debbie Cook, Alicia and Warren McCall of the State Veterinary Office in Hluhluwe, who conducted the buffalo BTB testing. Buffalo were captured by KZN Wildlife game capture staff, and thanks are due in particular to Paul Jennings and Quentin Rochat. Management staff of Hluhluwe-iMfolozi Park, Pete Hartley, Jabulani Ngubane, Paul Havelman, Dave Robertson and Jeff Cook ran an efficient buffalo capture operation, and Dave Balfour, Sue van Rensburg, Ruth and Owen Howison, facilitated AEJ's work at HUP. Erin Gorsich helped with data analysis in R. Between 1999-2004 AEJ received financial support for this project from the National Geographic Society, the Theresa Heinz Foundation, the Wildlife Conservation Society, NSF grant no. DEB-0090323 to Wayne Getz, the Mellon Foundation, the Explorer's Club and SigmaXi. From 2004-2006, this research was funded by NSF DEB 0541981 to AEJ and approved by the Oregon State University IACUC (Protocol number OSU: 3267).

References

1. Miller RS, Farnsworth ML, Malmberg JL (2013) Diseases at the livestock–wildlife interface: Status, challenges, and opportunities in the United States. Preventive Veterinary Medicine 110: 119–132. doi:10.1016/j.prevetmed.2012.11.021.

2. Fulford GR, Roberts MG, Heesterbeek JAP (2002) The metapopulation dynamics of an infectious disease: tuberculosis in possums. Theor Popul Biol 61: 15–29. doi:10.1006/tpbi.2001.1553.

3. Woodroffe R, Donnelly CA, Jenkins HE, Johnston WT, Cox DR, et al. (2006) Culling and cattle controls influence tuberculosis risk for badgers. ProcNatlAcadSciUSA 103: 14713–14717.

4. Jolles AE, Cooper DV, Levin SA (2005) Hidden effects of chronic tuberculosis in African buffalo. Ecology 86: 2358–2364.

5. Amanfu W (2006) The situation of tuberculosis and tuberculosis control in animals of economic interest. Tuberculosis(Edinb) 86: 330–335.

6. Michel AL (2002) Implications of tuberculosis in African wildlife and livestock. AnnNYAcadSci 969: 251–255.

7. Biet F, Boschiroli ML, Thorel MF, Guilloteau LA (2005) Zoonotic aspects of Mycobacterium bovis and Mycobacterium avium-intracellulare complex (MAC). VetRes 36: 411–436.

8. Rodwell TC, Kriek NP, Bengis RG, Whyte IJ, Viljoen PC, et al. (2001) Prevalence of bovine tuberculosis in African buffalo at Kruger National Park. JWildlDis 37: 258–264.

9. Kalema-Zikusoka G, Bengis RG, Michel AL, Woodford MH (2005) A preliminary investigation of tuberculosis and other diseases in African buffalo (Syncerus caffer) in Queen Elizabeth National Park, Uganda. Onderstepoort JVetRes 72: 145–151.

10. Caron A, Cross PC, Du Toit JT (2003) Ecological implications of bovine tuberculosis in African buffalo herds. Ecological Applications 13: 1338–1345.

11. Cross PC, Heisey DM, Bowers JA, Hay CT, Wolhuter J, et al. (2009) Disease, predation and demography: assessing the impacts of bovine tuberculosis on African buffalo by monitoring at individual and population levels. Journal of Applied Ecology 46: 467–475.

12. Michel AL, Bengis RG, Keet DF, Hofmeyr M, Klerk LM, et al. (2006) Wildlife tuberculosis in South African conservation areas: implications and challenges. VetMicrobiol 112: 91–100.

13. Ayele WY, Neill SD, Zinsstag J, Weiss MG, Pavlik I (2004) Bovine tuberculosis: an old disease but a new threat to Africa. IntJTubercLung Dis 8: 924–937.

14. Ramirez-Villaescusa AM, Medley GF, Mason S, Green LE (2009) Herd and individual animal risks associated with bovine tuberculosis skin test positivity in cattle in herds in south west England. PrevVetMed 92: 188–198.

15. Michel AL, Muller B, van Helden PD (2010) Mycobacterium bovis at the animal-human interface: a problem, or not? VetMicrobiol 140: 371–381.

16. Waldram MS, Bond WJ, Stock WD (2008) Ecological Engineering by a Mega-Grazer: White Rhino Impacts on a South African Savanna. Ecosystems 11: 101–112. doi:10.1007/s10021-007-9109-9.

17. MacDonald DW, Riordan P, Mathews F (2006) Biological hurdles to the control of TB in cattle: A test of two hypotheses concerning wildlife to explain the failure of control. Biological Conservation 131: 268–286.

18. Jenkins HE, Woodroffe R, Donnelly CA (2010) The duration of the effects of repeated widespread badger culling on cattle tuberculosis following the cessation of culling. PLoSOne 5: e9090.

19. O'Brien DJ, Schmitt SM, Rudolph BA, Nugent G (2011) Recent advances in the management of bovine tuberculosis in free-ranging wildlife. Veterinary Microbiology 151: 23–33. doi:16/j.vetmic.2011.02.022.

20. Jolles AE (2007) Population biology of African buffalo (Syncerus caffer) at Hluhluwe-iMfolozi Park, South Africa. African Journal of Ecology 45: 398–406.

21. Dora CA (2004) The influences of habitat structure and landscape heterogeneity on African buffalo (Syncerus Caffer) group size in Hluhluwe-iMfolozi Park, South Africa [MSc]. University of Oregon.

22. Monaghan ML, Doherty ML, Collins JD, Kazda JF, Quinn PJ (1994) The tuberculin test. VetMicrobiol 40: 111–124.

23. Jolles AE, Ezenwa VO, Etienne RS, Turner WC, Olff H (2008) Interactions between macroparasites and microparasites drive infection patterns in free-ranging African buffalo. Ecology 89: 2239–2250.

24. Ezenwa VO, Etienne RS, Luikart G, Beja-Pereira A, Jolles AE (2010) Hidden consequences of living in a wormy world: nematode-induced immune suppression facilitates tuberculosis invasion in African buffalo. Am Nat 176: 613–624. doi:10.1086/656496.

25. McDonald RA, Delahay RJ, Carter SP, Smith GC, Cheeseman CL (2008) Perturbing implications of wildlife ecology for disease control. Trends EcolEvol 23: 53–56.

Supplementary information

Table S1: Linear regression statistics for area datasets

	Masinda S	Mbhuzane	Gunjaneni	Corridor	Manzibomvu
Best-fit values					
Slope	-0.03988 ± 0.02675	-0.01229 ± 0.005587	0.009207 ± 0.01474	0.0008209 ± 0.007279	-0.02622 ±
Y-intercept when X=0.0	80.12 ± 53.57	24.74 ± 11.20	-18.32 ± 29.52	-1.553 ± 14.57	52.62 ±
X-intercept when Y=0.0	2009	2013	1989	1891	2007
1/slope	-25.08	-81.36	108.6	1218	-38.14
95% CI					
Slope	-0.1141 to 0.03439	-0.03633 to 0.01175	-0.05421 to 0.07263	-0.03050 to 0.03214	Perfect line
Y-intercept when X=0.0	-68.58 to 228.8	-23.44 to 72.92	-145.3 to 108.7	-64.25 to 61.15	Perfect line
X-intercept when Y=0.0	2004 to +infinity	2007 to +infinity	-infinity to	-infinity to 2000	
Goodness of Fit					
r^2	0.3571	0.7076	0.1633	0.006318	1
Sy.x	0.1667	0.01767	0.08434	0.03152	
Is slope significantly non-zero?					
F	2.222	4.84	0.3902	0.01272	
DFn, DFd	1.000, 4.000	1.000, 2.000	1.000, 2.000	1.000, 2.000	1.000, 0.0000
P value	0.2104	0.1588	0.5959	0.9205	
Deviation from zero?	Not Significant	Not Significant	Not Significant	Not Significant	Perfect line
Data					
Number of X values	6	4	4	4	2
Maximum number of Y replicates	1	1	1	1	1
Total number of values	6	4	4	4	2
Number of missing values	2	4	4	4	6

Table S2: Model selection for BTB incidence in buffalo herds at HIP. Statistically significant terms are shown in red for best-fitting models (i.e. within 2 AIC points of the model with the lowest AIC). The model selected as final model is marked with asterisks. "n.c." stands for not competitive.

Model: incidence ~	model rank	AIC
area	3	114.2
capture interval	n.c.	136.1
capture effort	n.c.	147.1
area + capture interval	1	113.8
area + capture effort	n.c.	118.2
capture interval + capture effort	n.c.	143.9
area + capture interval + capture effort	2	114.0

Table S3: Model selection for change in BTB prevalence in buffalo herds at HIP. Statistically significant terms are shown in red for best-fitting models (i.e. within 2 AIC points of the model with the lowest AIC). The model selected as final model is marked with asterisks. "n.c." stands for not competitive.

Model: incidence ~	**model rank**	**AIC**
area	n.c.	124.1
capture interval	n.c.	138.5
capture effort	n.c.	145.9
area + capture interval	2	115.1
area + capture effort	n.c.	127.3
capture interval + capture effort	n.c.	141.6
area + capture interval + capture effort	1	114.8

Chapter 3

Bovine TB in livestock and wildlife: what's in the genes?

Nikki le Roex[1], Paul D van Helden[1], Ad P Koets[2,3], Eileen G Hoal[1]

1 DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Stellenbosch University, Cape Town, South Africa. **2** Department of Farm Animal Health, **3** Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands.

Physiological Genomics (2013); 45(15): 631-637.

Abstract

Bovine tuberculosis (BTB) is a chronic, infectious disease found in domestic livestock and wildlife. It is caused predominantly by *Mycobacterium bovis,* which forms part of the *Mycobacterium tuberculosis* complex. BTB has serious implications for the movement of animals and animal products, biodiversity and public health, and is of significant economic concern. The existence of wildlife maintenance hosts makes it extremely difficult to eradicate BTB, even when established control strategies are in place, creating the need for alternative methods for controlling this disease. There are multiple factors that influence the outcome of infection by a pathogen, one of which is the host‟s genome. The identification of genetic variants involved in the susceptibility to bovine TB would supply a new selection of potential drug targets as well as the possibility for the breeding of animals with greater disease resistance. In this review, we collate the results of the BTB heritability and association studies performed in cattle and wildlife, discuss considerations and other methodologies (such as gene expression work) to be taken into account when performing genetic studies, and make some recommendations for future work in this area.

Keywords: *bovine, tuberculosis, genetics, susceptibility*

1. Introduction

Bovine tuberculosis (BTB) in domestic livestock and wildlife is caused by *Mycobacterium bovis,* which forms part of the *Mycobacterium tuberculosis* complex [1]. It is a chronic, infectious disease that is transmitted predominantly by inhalation and is characterised by the formation of granulomas in the lungs, lymph nodes and various other organs [2]. The exact origin of this disease is unclear, but it has been suggested that bovine TB emerged in Europe and spread to the United Kingdom, and was subsequently introduced to many parts of the world through the movement of cattle from the UK and the Netherlands to former colonies [3]. BTB is widespread in cattle around the world, and is present in many wildlife species and domestic animals. This has serious implications for international trade and export of animals and animal products, including beef. Additional economic losses are suffered through the slaughter of livestock, and regular BTB testing incurs a vast expense [4]. Bovine TB is also a public health concern, as it is a zoonotic disease and can be transmitted to humans. This is of

particular concern in developing countries, where people and livestock live in close proximity to each other [4,5].

Much interest has recently been focused on the role of wildlife reservoirs in the maintenance and spread of BTB, both within wildlife populations and to domestic livestock and other spill-over hosts. The brushtail possum (*Trichosurus vulpecula*) in New Zealand, the African buffalo (*Syncerus caffer*) in South Africa, the Eurasian badger (*Meles meles*) in the United Kingdom, the white-tailed deer (*Odocoileus virginianus*) in the USA and the European wild boar (*Sus scrofa*) in Spain are all considered maintenance hosts of BTB in their respective environments [3]. The list of spill-over hosts is far more extensive, and includes goats (*Capra hircus*), lynx (*Lynx pardinus*), lions (*Panthera leo*), dogs (*Canis familiaris*), cats (*Felis catus*), sheep (*Ovis aries*), mink (*Lutreola vison*) and many more [6,5].

Multiple factors influence the outcome of infection by a pathogen. Host genetic factors interact with both environmental factors and the genome of the pathogen to determine the disease profile of an individual, with a particular interplay between them resulting in active disease. Many lines of evidence convincingly demonstrate that the host‟s genetic make-up plays a role in the outcome of infection with TB in both humans and animals. Twin studies have shown that when monozygotic and dizygotic twins are compared with respect to development of TB, the concordance is higher in monozygotic twins than in dizygotic twins [7]. In 1926, 251 babies unintentionally injected with a live dose of virulent *M. tuberculosis* showed a range of responses, demonstrating that there were innate factors involved in the outcome of infection [8]. With respect to *Mycobacterium bovis* susceptibility, in the 1940‟s it was found that inbred strains of rabbits displayed one of two phenotypes (susceptible or resistant) following infection with *M. bovis*. This inheritance pattern suggested that the disease profile displayed by the rabbits was largely genetically inherited [9]. The differential susceptibility of mice to the *M. bovis* Bacille Calmette-Guerin (BCG) vaccine led to the identification of the first TB susceptibility locus, namely the Nramp1 gene [10].

There are many different prevention and treatment strategies, used by different countries on different BTB hosts – test and slaughter policies, the BCG vaccine, wide-scale culling and restricted movement and legislation. All of these strategies have met with limited success in countries where a wildlife BTB reservoir exists, and there is a pressing need for alternative methods of controlling this disease [4,5]. The identification of genetic variants involved in susceptibility to bovine TB would supply a new selection of potential drug targets as well as

the possibility of breeding animals with greater disease resistance. Research into genetic resistance to BTB has begun in a number of countries (Figure 1) and the application of genetic resistance information to breeding programmes has already been implemented in cattle for other diseases such as mastitis [11]. Genetic and immunological information form an essential part of this approach, and breeding strategies for BTB resistance, either by marker-assisted selection or the incorporation of marker information into estimated breeding values (EBVs) for a whole-genomic selection approach [12], could be invaluable for both the commercial and private sectors. No single study design or approach can be relied on to identify all the genetic components of TB susceptibility, as each different approach has its own advantages and limitations. Heritability and association studies are the most commonly used approaches in investigating the genetic basis of disease susceptibility, and will be the focus of this review.

Figure 1: Countries highlighted green represent those where genetic studies investigating resistance to BTB have taken place.

Below, we summarise the results of the BTB heritability and association studies performed in cattle and wildlife, suggest considerations and other methodologies to be taken into account, and make some recommendations for future work in this area.

2. Heritability Studies

Heritability studies are used to assess the extent of the genetic contribution to a particular phenotype, i.e. the extent to which the disease phenotype is controlled by the host‟s genetic make-up. In humans, this is most commonly done using twin studies, or more recently, using families and household contacts to determine the degree of concordance of genotype and phenotype [13]. In animals, this can be more thoroughly investigated using breeding programs and experimental infection. Heritability is an important aspect of any disease, and is an essential factor in the successful implementation of breeding programmes. Imperfect diagnostics, such as the limited sensitivity of the skin test in BTB detection, raises the possibility that the true heritability of this disease may be even higher than observed [14]. Historical data on differential BTB susceptibility in cattle breeds has been outlined by Vordermeier et al. [15] and will not be addressed here.

2.1 Livestock

Numerous studies have shown cattle breeds to display different levels of susceptibility to bovine TB. A study done by Ameni et al. [16] with Zebu, Zebu-Holstein crosses and Holstein cattle under identical field husbandry conditions in Ethiopia, showed both a higher prevalence and severity of BTB in the Holstein cattle compared to the Zebu or Zebu-Holstein crosses. Another study of Ethiopian cattle [17] identified breed as a significant factor affecting TB manifestation. Breed types were classified as local, exotic or cross-bred, and showed exotic cattle to have a higher risk of severe BTB pathology.

Bermingham et al. [18] investigated the heritability of single comparative intradermal tuberculin (SCIT) test responsiveness and confirmed *M. bovis* infection in Irish Holstein-Friesian dairy cattle. Approximately 600-800 herd tests encompassing 14,000 cows showed heritability estimates of 0.14 and 0.18 for SCIT test responsiveness and confirmed *M. bovis* infection, respectively. A similar study done on British Holstein-Friesian dairy cattle calculated heritability estimates of SCIT test responsiveness and confirmed *M. bovis* infection to be 0.16 and 0.18, respectively [19].

2.2 Wildlife

Red deer are found as both livestock and wildlife in various parts of New Zealand, Australia, the USA and Canada. They are highly susceptible to bovine TB and act as a reservoir for transmission to cattle [20]. A study by Mackintosh et al. [21] showed differential

susceptibility and transmission in red deer, which was attributed to host genetic variation. Experimental infection of stags showed a range of disease phenotypes, which was closely mirrored in the response of sired offspring, suggesting a strong genetic basis for resistance. The heritability of BTB resistance was calculated to be as high as 0.48. An investigation of BTB in white-tailed deer in Michigan, USA, by Blanchong et al. [22], using microsatellites, found BTB-infected deer to be significantly more closely-related than non-infected deer. The heritability studies performed in livestock and wildlife can be seen in Table 1.

Table 1: Bovine TB Heritability studies

Authors	**Animal**	**Breed**	**Factor**	**Result/Heritability estimate**
Mackintosh et al. (2000)	Red deer		BTB resistance	0.48
Ameni et al. (2007)	Cattle	Zebu/ Holstein/ Zebu-Holstein	BTB prevalence and severity	Higher in Holstein
Blanchong et al. (2007)	White-tailed deer		BTB infection	Infected animals more related
Bermingham et al. (2009)	Cattle	Irish Holstein-Friesian	Skin test response; M. bovis culture	0.14; 0.18
Brotherstone et al. (2010)	Cattle	British Holstein-Friesian	Skin test response; M. bovis culture	0.16; 0.18
Biffa et al. (2012)	Cattle	Exotic/ local/ cross bred	BTB severity	Higher in exotic breeds

3. Association studies

Association studies compare the allele frequencies of specific markers between infected/diseased and uninfected/not diseased individuals, and are currently the most widely-utilised in studying susceptibility to infection or disease. These case-control association studies are better than linkage studies at detecting genes that have a small effect [23], and as TB is known to be under polygenic control, there are likely to be many genes that each have a small effect on susceptibility. Genetic association studies can be roughly divided into two categories – candidate gene or whole-genome studies. The candidate gene approach investigates SNPs or other variation in genes deemed to be of biological relevance to the

question at hand. For example, in the case of TB, one might select genes involved in the immune response, or that have been shown to be associated with other infectious diseases, even in other species. This approach utilises a relatively small number of markers, but can be expanded depending on budget and time. One of the main disadvantages of this approach, however, is that the marker associated with disease resistance may not be causative, but may be in linkage disequilibrium (LD) with the causative variant. This is particularly applicable to those species with small effective population sizes and limited ranges, where LD blocks are likely to be longer [23]. The whole genome approach involves genotyping polymorphic markers across the whole genome, and therefore makes no prior assumptions as to which genes may or may not be involved. This can be a distinct advantage, as it is likely that there will be many more genes involved in disease resistance than those of the immune system. This approach thus utilises a large number of markers and is more likely to identify novel genes or pathways involved in susceptibility [24]. The candidate gene approach is less commonly employed in animals with assembled genomes, where sufficient data exists to interrogate the whole genome. However, in animals without sufficient genomic data, such as most wildlife species, candidate genes are a frequently utilised starting point, even though the implicit assumptions are that relatively few loci underpin the phenotype, and their effect is detectable.

3.1 Livestock

Kadarmideen et al. [25] investigated polymorphisms of the 3‟UTR microsatellite of the SLC11A1 (NRAMP1) gene and their association with BTB in African Zebu cattle. They compared four phenotypes – namely SCIT test responsiveness, presence of visible lung lesions, culture test outcome and the predicted true infection using a Bayesian model, and found two alleles to be significantly associated with reduced incidence of BTB traits. Driscoll et al. [26] analysed genetic factors in British cattle using microsatellite loci and found two markers, BMS2753 and INRA111, to be strongly associated with BTB reactor status. The BMS2753 marker lies in close proximity to IFNGR1 and several other immune response genes, and INRA111 was previously associated with mastitis [27]. A genome-wide association scan of bovine TB susceptibility in Holstein-Friesian dairy cattle by Finlay et al. [28] found three SNPs in a region on BTA 22 to be associated with bovine TB susceptibility. This region contains the taurine transporter gene *SLC6A6*. Sun et al. [29] investigated the role of polymorphisms in the toll-like receptor 1 and 9 genes with susceptibility to BTB in Chinese Holstein cattle. Significant associations were found with TLR1 genotype variants at

TLR1-G1596A and BTB susceptibility, and a trend towards significance with TLR1-A1475C.

3.2 Wildlife

The European wild boar is abundant on the Iberian Peninsula, and its range and population size are increasing steadily. It has become a wildlife reservoir for BTB in regions of Spain and concern has been raised about the spread and management of this disease in both wild and contained populations [30]. Acevedo-Whitehouse et al. [30] investigated the genetic component of resistance in European wild boars from southern Spain, and found that increased genetic heterozygosity had a significant positive effect on BTB resistance, both in terms of infection status and disease severity/dissemination. Several of the loci tested also revealed high homology to regions of the genome with immune-related functions. Naranjo et al. [31] investigated polymorphisms of a methylmalonyl-CoA mutase (MUT) microsatellite and their relationship to BTB susceptibility in European wild boar. The microsatellite is located next to exon 2 of the gene, and exhibited three polymorphisms within the population. The authors found evidence of a balancing polymorphism, with one allele conferring significant protection, another conferring significant increased risk of infection, but the presence of both alleles conferring the highest level of protection against infection. African buffalo are a maintenance host of BTB in southern Africa and pose a risk of infection to cattle and surrounding communities, and a threat to biodiversity and ecotourism [32]. Le Roex et al. [32] explored polymorphisms in immune-related genes and their association with BTB susceptibility in the African buffalo using fluorescent genotyping. SNPs located in three genes – SLC7A13, IL1α and DMBT1 – were found to be significantly associated with BTB status in the buffalo. The association studies performed in livestock and wildlife can be seen in Table 2.

Table 2: BTB Association studies

Author	Animal	Breed	Associated factor	Gene/marker
Acevedo-Whitehouse et al. (2005)	European wild boar		Infection status; BTB severity	Heterozygosity associated with resistance
Naranjo et al. (2008)	European wild boar		Lesions & culture	Methylmalonyl CoA mutase (MUT)
Driscoll et al. (2011)	Cattle	Multiple	Skin test response	BMS2753; INRA111
Kadarmideen et al. (2011)	Cattle	African Zebu	Skin test response; visible lesions	SLC11A1
Finlay et al. (2012)	Cattle	Irish Holstein-Friesian	EBV (based on skin test response)	Chr 22, SLC6A6
Sun et al. (2012)	Cattle	Chinese Holstein	Skin test response	TLR1
le Roex et al. (2013)	African buffalo		Skin test/ELISA response	SLC7A13; IL1a; DMBT1

4. Considerations in genetic studies

Despite the popularity of association studies and their ability to identify susceptibility loci, they may also identify spurious associations [23]. These should be minimised in the first instance in order to avoid reporting inaccurate results, and any associations need to be rigorously replicated and/or validated before being accepted as probable. In order to reduce the likelihood that spurious associations are identified, there are a number of considerations that must be taken into account.

4.1 Sample collection and phenotype classification

The first consideration is the accuracy of the disease diagnosis. The best diagnostic tool available - i.e. the test with the highest specificity and sensitivity - should be used for designating individuals as cases or controls, as too many misclassifications could severely bias the association tests [33,34]. The phenotypes must also be clearly defined according to the question/objective at hand – for example, if one is interested only in active disease, the SCIT test alone may not be an appropriate diagnostic. The cases and controls should have comparable exposure to the pathogen in order to be correctly classified as having a susceptible or resistant phenotype. Cases and controls should also ideally be matched as closely as possible in all other ways (age, environment etc.) so as not to introduce additional variables that may confound the association testing [33]. The total number of samples in a

study is an important consideration, often overlooked. This must be sufficiently large to generate enough statistical power to detect variants that have a relatively small effect, as is the usual situation in a complex disease. Statistical power is a function of the size of the study population, the allele frequency of the SNP, the size of the effect of the SNP and the level of linkage disequilibrium. Power calculations of a real as opposed to theoretical sample set can only be done retrospectively, so in order to increase power at the outset, sample numbers should be in the order of hundreds, or even thousands. Lack of study power through inadequate samples numbers is one of the most frequent causes of non-reproducible association results [23].

4.2 Experimental design and analysis

The experimental design of a study is crucial in order to produce results that are not false positives or false negatives. Batch effects occur when different sets of samples are run on different days or in batches under slightly different conditions, caused by human error, different reagents or hardware, and can result in spurious associations. If samples must be run in different batches, controlling for batch effects can be done by the randomisation of samples and variables across batches. Population stratification is the other most common cause of spurious associations. If subgroups within the study population have different allele frequencies, associations may be found that are a function of the population substructure rather than disease susceptibility [13,23]. Any possible subgroups should be randomised across batches and controlled for during statistical analyses. In studies where large numbers of markers are compared across many genes, multiple comparisons are made. In order to minimise the occurrence of type 1 errors, correction for multiple testing, e.g. Bonferroni correction, should be applied where appropriate [13].

4.3 Wildlife

When working with wildlife species, there are additional factors which may hamper a genetic study. Many wildlife species have not had their full genomes sequenced, and thus the availability of genetic information is limited. This means that much work must be done prior to beginning a genetic association study. This includes the identification and development of suitable markers and, ideally, the elucidation of the complex genetic interactions and pathways that may operate within the species. Even in well-established models, such as cattle, many proteins and gene products have not been characterised and much annotation is merely hypothesised rather than experimentally established. Associated markers may also not

be the causal variant in a disease, but rather may be in linkage disequilibrium (LD) with the causal variant [13,23]. In the absence of established LD information for a particular species, this can be difficult to ascertain. Wildlife species are not managed in the same fashion that domestic livestock may be, and thus relatedness can be a concern. A lack of population ecology data on a particular species leaves researchers unable to establish the level of relatedness in a population and this can introduce confounding effects. Phenotype determination in wildlife may also be more difficult due to absent/reduced diagnostic capabilities for many species, and the logistical complications of capture and/or recapture in a natural environment.

4.4 Additional challenges

The type of association study should reflect the aim of the study, with candidate gene studies having greater power to detect small effects, and genome-wide studies being better suited to identifying novel genes/pathways [24]. In the same way that the host‟s genetic make-up influences the outcome of infection, so too does the genetic make-up of the pathogen. Information regarding the genetic variation and lineage of the pathogen should be included, and reported, where known. BTB is influenced by many genes, and the role of gene-gene interactions and gene-environment interactions cannot be excluded. A genetic variant not found to be associated on its own may still play a role in conjunction with another gene or variant, and the same may also be true for studies undertaken under different environmental conditions [33,13]. Current methodologies seldom include this type of analysis and thus these genes are missed. The type of markers used in a study is also important – while microsatellite markers are often well-characterised, and work very well in linkage studies, the high allelic diversity and reduced LD at these loci may reduce the power to detect associations with a disease phenotype in a population-level study [35].

5. Gene expression studies

Studying the gene expression pattern of a particular phenotype provides another approach to identifying candidate genes. Gene expression studies look to identify gene transcripts that are differentially expressed in groups of animals, in order to identify markers of infection. The primary goal is generally for use in diagnostics, but these studies can also inform genetic researchers of candidate genes worthy of investigation, as well as contributing information

that can be used in conjunction with genetic and other data in a systems biology approach to build a more complete understanding of the pathways/systems involved in disease [36,37]. Gene expression studies will not be covered extensively in this review, as they are not strictly genetic studies, but nevertheless can be of great assistance. Studies using natural models of infection, i.e. that investigate differential gene expression in cells/tissues of animals that were naturally infected with BTB, are particularly useful, as this is the closest approximation of what might happen in nature.

Studies using natural BTB infection in cattle have identified many genes that are differentially expressed between *M. bovis*-infected and uninfected animals – these include, but are not limited to: ADAM17, CXCR3, IER5, PHB2, CD84, TBK1, TLR2, TLR3, BCL2, NFATC4, IFNG, IFNGR1, TNFSF13B, KIAA1971, SLAMF1, CASP1, DEFB10, IFNAR1, KIR3DS1, MYD88, PTPN2, STAT1, STAT2, TREM1, TYK2, TYROBP, CD83, CTLA4, IL1A, IL8 and IL15 [38–41]. In wildlife, studies of naturally infected wild boar have identified the differential expression of many genes, including C3, MUT, ANX, CUL, VCAM-1, C7, ARG, OPN, CXCR4, VDR, BAP29, GAL-1, SLA class I and II, Ig, CC1qB [42–44]. A gene expression study in Iberian red deer showed the differential expression of SMAP-29 and CK2 between M. *bovis*-infected and uninfected individuals [45]. It is evident that a number of the genes listed above are involved in the immune system, many of which, such as VDR, IFNG, and TLR2, are also implicated in human TB association studies [46–48], lending further credence to the identification of the gene and its product as an important factor in mycobacterial disease.

6. Conclusions and future directions

The identification of genetic markers linked to a resistant BTB phenotype can inform studies in other species, and may enable the breeding of animals with greater disease resistance. This would, at the very least, be a useful additional measure in conjunction with the other control strategies currently available. The ability to breed animals with a greater BTB resistance may be of considerable commercial benefit to the cattle farming industry, and would certainly be invaluable in other species, such as the African buffalo, which forms part of a multi-million dollar hunting and breeding industry in South Africa, with “disease-free” buffalo particularly prized. This also pertains to other wildlife species kept in a farm or reserve environment, where selective breeding can produce very tangible benefits. Allen et al. [34] suggested that

several other bovine pathogens such as *Brucella abortus* and *Mycobacterium avium paratuberculosis* may operate within their hosts in a similar manner to *M. bovis*, and thus the breeding of BTB resistance may also inadvertently select for resistance to those pathogens as well. However, it is also possible that such animals may show increased susceptibility to other diseases.

To date, most of the genetic research has focused on heritability and candidate gene association studies investigating single markers. While association studies remain the most popular choice in disease susceptibility studies, there are a number of considerations to be taken into account, in the collection and classification of samples, the experimental design and analyses as well as the species under investigation. Current human studies in TB susceptibility have included a number of genome-wide association studies, and this may be a direction future animal work will take, provided an assembled genome of the animal of interest is available. The large number of subjects required, and the relatively low power of genome-wide association studies to detect associations of small effect does hamper this approach, but the identification of novel loci may become increasingly important [49]. Gene expression studies, particularly of natural infections, provide an excellent pool of differentially expressed candidate genes for selection, and should be utilised to inform genetic studies wherever possible.

Future genetic work may also need to progress from the strategy of looking for individual loci. There are three main approaches to investigating genetic markers and their associations: (i) to investigate individual markers in isolation, and their association with the biological phenotype, (ii) to investigate gene-gene interactions, by testing marker/SNP combinations for association with the phenotype, and (iii) to investigate pathways, and determine whether a group of related genes/markers are associated with the phenotype. As we know, genes do not work in isolation, and there is increasing interest in the analysis of gene-gene interactions and pathways. Studies in humans have begun investigating these processes [50,51] but this has yet to be embarked upon in animal disease studies.

In summary, investigations into the role of the host‟s genetic make-up in bovine TB susceptibility are in their infancy, and there is an enormous amount of work still to be done before any true understanding of the interplay between host genetics and BTB susceptibility can be definitively achieved. The genome assembly of a growing number of species will provide researchers with the tools and information needed to tackle this imposing question

from a number of angles, and the co-operation of multiple disciplines will ensure the best success at identifying the host genetic component of bovine TB resistance.

Acknowledgements

We thank the anonymous reviewers for their thoughtful comments and suggestions.

Grants

NLR was supported by a student stipend from WOTRO Science for Global Development (grant W01.65.321.00) awarded to PVH, EGH and AK.

Disclosure

The authors declare no conflict of interest.

References

1. Smith NH, Gordon SV, de la Rua-Domenech R, Clifton-Hadley RS, Hewinson RG (2006) Bottlenecks and broomsticks: the molecular evolution of Mycobacterium bovis. Nat Rev Microbiol 4: 670–681. doi:10.1038/nrmicro1472.

2. Pollock JM, Neill SD (2002) Mycobacterium bovis infection and tuberculosis in cattle. VetJ 163: 115–127.

3. Renwick AR, White PCL, Bengis RG (2007) Bovine tuberculosis in southern African wildlife: a multi-species host-pathogen system. Epidemiol Infect 135: 529–540. doi:10.1017/S0950268806007205.

4. Michel AL, Muller B, van Helden PD (2010) Mycobacterium bovis at the animal-human interface: a problem, or not? VetMicrobiol 140: 371–381.

5. Ayele WY, Neill SD, Zinsstag J, Weiss MG, Pavlik I (2004) Bovine tuberculosis: an old disease but a new threat to Africa. IntJTubercLung Dis 8: 924–937.

6. Maas M, Michel AL, Rutten VPMG (2012) Facts and dilemmas in diagnosis of tuberculosis in wildlife. Comp Immunol Microbiol Infect Dis. doi:10.1016/j.cimid.2012.10.010.

7. Comstock GW (1978) Tuberculosis in twins: a re-analysis of the Prophit survey. Am Rev Respir Dis 117: 621–624.

8. Rieder HL (2003) [Clarification of the Luebeck infant tuberculosis]. Pneumologie 57: 402–405. doi:10.1055/s-2003-40555.

9. Lurie MB (1944) Experimental epidemiology of tuberculosis : hereditary resistance to attack by tuberculosis and to the ensuing disease and the effect of the concentration of tubercle bacilli upon these two phases of resistance. J Exp Med 79: 573–589.

10. Gros P, Skamene E, Forget A (1981) Genetic control of natural resistance to Mycobacterium bovis (BCG) in mice. J Immunol 127: 2417–2421.

11. Rupp R, Boichard D (2003) Genetics of resistance to mastitis in dairy cattle. Vet Res 34: 671–688. doi:10.1051/vetres:2003020.

12. Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.

13. Romero R, Kuivaniemi H, Tromp G, Olson J (2002) The design, execution, and interpretation of genetic association studies to decipher complex diseases. Am J Obstet Gynecol 187: 1299–1312.

14. Bishop SC, Woolliams JA (2010) On the genetic interpretation of disease data. PLoS ONE 5: e8940. doi:10.1371/journal.pone.0008940.

15. Vordermeier M, Ameni G, Berg S, Bishop R, Robertson BD, et al. (2012) The influence of cattle breed on susceptibility to bovine tuberculosis in Ethiopia. Comp Immunol Microbiol Infect Dis 35: 227–232. doi:10.1016/j.cimid.2012.01.003.

16. Ameni G, Aseffa A, Engers H, Young D, Gordon S, et al. (2007) High prevalence and increased severity of pathology of bovine tuberculosis in Holsteins compared to zebu breeds under field cattle husbandry in central Ethiopia. Clin Vaccine Immunol 14: 1356–1361. doi:10.1128/CVI.00205-07.

17. Biffa D, Bogale A, Godfroid J, Skjerve E (2012) Factors associated with severity of bovine tuberculosis in Ethiopian cattle. Trop Anim Health Prod 44: 991–998. doi:10.1007/s11250-011-0031-y.

18. Bermingham ML, More SJ, Good M, Cromie AR, Higgins IM, et al. (2009) Genetics of tuberculosis in Irish Holstein-Friesian dairy herds. JDairy Sci 92: 3447–3456.

19. Brotherstone S, White IMS, Coffey M, Downs SH, Mitchell AP, et al. (2010) Evidence of genetic resistance of cattle to infection with Mycobacterium bovis. J Dairy Sci 93: 1234–1242. doi:10.3168/jds.2009-2609.

20. Griffin JF, Mackintosh CG (2000) Tuberculosis in deer: perceptions, problems and progress. Vet J 160: 202–219. doi:10.1053/tvjl.2000.0514.

21. Mackintosh CG, Qureshi T, Waldrup K, Labes RE, Dodds KG, et al. (2000) Genetic resistance to experimental infection with Mycobacterium bovis in red deer (Cervus elaphus). InfectImmun 68: 1620–1625.

22. Blanchong JA, Scribner KT, Kravchenko AN, Winterstein SR (2007) TB-infected deer are more closely related than non-infected deer. Biol Lett 3: 103–105.

23. Cardon LR, Bell JI (2001) Association study designs for complex diseases. Nat Rev Genet 2: 91–99. doi:10.1038/35052543.

24. Amos W, Driscoll E, Hoffman JI (2011) Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? Proc Biol Sci 278: 1183–1188. doi:10.1098/rspb.2010.1920.

25. Kadarmideen HN, Ali AA, Thomson PC, Müller B, Zinsstag J (2011) Polymorphisms of the SLC11A1 gene and resistance to bovine tuberculosis in African Zebu cattle. Anim Genet 42: 656–658. doi:10.1111/j.1365-2052.2011.02203.x.

26. Driscoll EE, Hoffman JI, Green LE, Medley GF, Amos W (2011) A preliminary study of genetic factors that influence susceptibility to bovine tuberculosis in the British cattle herd. PLoS ONE 6: e18806. doi:10.1371/journal.pone.0018806.

27. Schulman NF, Viitala SM, de Koning DJ, Virta J, Mäki-Tanila A, et al. (2004) Quantitative trait Loci for health traits in Finnish Ayrshire cattle. J Dairy Sci 87: 443–449. doi:10.3168/jds.S0022-0302(04)73183-5.

28. Finlay EK, Berry DP, Wickham B, Gormley EP, Bradley DG (2012) A genome wide association scan of bovine tuberculosis susceptibility in Holstein-Friesian dairy cattle. PLoS ONE 7: e30545. doi:10.1371/journal.pone.0030545.

29. Sun L, Song Y, Riaz H, Yang H, Hua G, et al. (2012) Polymorphisms in toll-like receptor 1 and 9 genes and their association with tuberculosis susceptibility in Chinese Holstein cattle. Vet Immunol Immunopathol 147: 195–201. doi:10.1016/j.vetimm.2012.04.016.

30. Acevedo-Whitehouse K, Vicente J, Gortazar C, Hofle U, Fernandez-de-Mera IG, et al. (2005) Genetic resistance to bovine tuberculosis in the Iberian wild boar. MolEcol 14: 3209–3217.

31. Naranjo V, Acevedo-Whitehouse K, Vicente J, Gortazar C, de la FJ (2008) Influence of methylmalonyl-CoA mutase alleles on resistance to bovine tuberculosis in the European wild boar (Sus scrofa). Anim Genet 39: 316–320.

32. le Roex N, Koets AP, van Helden P, Hoal EG (2013) Gene Polymorphisms in African buffalo associated with susceptibility to bovine tuberculosis infection. PLoS ONE 8(5): e64494. doi:10.1371/journal.pone.0064494.

33. Stein CM (2011) Genetic Epidemiology of Tuberculosis Susceptibility: Impact of Study Design. PLoS Pathog 7. doi:10.1371/journal.ppat.1001189.

34. Allen AR, Minozzi G, Glass EJ, Skuce RA, McDowell SWJ, et al. (2010) Bovine tuberculosis: the genetic basis of host susceptibility. Proc R Soc B 277: 2737–2745. doi:10.1098/rspb.2010.0830.

35. Curtis D, Vine AE, Knight J (2008) Investigation into the Ability of SNP Chipsets and Microsatellites to Detect Association with a Disease Locus. Ann Hum Genet 72: 547–556. doi:10.1111/j.1469-1809.2008.00434.x.

36. Williams RBH, Chan EKF, Cowley MJ, Little PFR (2007) The influence of genetic variation on gene expression. Genome Res 17: 1707–1716. doi:10.1101/gr.6981507.

37. Cheung VG, Spielman RS (2002) The genetics of variation in gene expression. Nature Genetics 32: 522–525. doi:10.1038/ng1036.

38. Killick KE, Browne JA, Park SDE, Magee DA, Martin I, et al. (2011) Genome-wide transcriptional profiling of peripheral blood leukocytes from cattle infected with Mycobacterium bovis reveals suppression of host immune genes. BMC Genomics 12: 611. doi:10.1186/1471-2164-12-611.

39. Blanco FC, Schierloh P, Bianco MV, Caimi K, Meikle V, et al. (2009) Study of the immunological profile towards Mycobacterium bovis antigens in naturally infected cattle. Microbiol Immunol 53: 460–467. doi:10.1111/j.1348-0421.2009.00141.x.

40. Meade KG, Gormley E, Doyle MB, Fitzsimons T, O'Farrelly C, et al. (2007) Innate gene repression associated with Mycobacterium bovis infection in cattle: toward a gene signature of disease. BMCGenomics 8: 400.

41. Meade KG, Gormley E, Park SDE, Fitzsimons T, Rosa GJM, et al. (2006) Gene expression profiling of peripheral blood mononuclear cells (PBMC) from Mycobacterium bovis infected cattle after in vitro antigenic stimulation with purified protein derivative of tuberculin (PPD). Vet Immunol Immunopathol 113: 73–89. doi:10.1016/j.vetimm.2006.04.012.

42. Galindo RC, Ayoubi P, Naranjo V, Gortazar C, Kocan KM, et al. (2009) Gene expression profiles of European wild boar naturally infected with Mycobacterium bovis. Vet Immunol Immunopathol 129: 119–125. doi:10.1016/j.vetimm.2008.12.012.

43. Naranjo V, Ayoubi P, Vicente J, Ruiz-Fons F, Gortazar C, et al. (2006) Characterization of selected genes upregulated in non-tuberculous European wild boar as possible correlates of resistance to Mycobacterium bovis infection. VetMicrobiol 116: 224–231.

44. Naranjo V, Höfle U, Vicente J, Martín MP, Ruiz-Fons F, et al. (2006) Genes differentially expressed in oropharyngeal tonsils and mandibular lymph nodes of tuberculous and nontuberculous European wild boars naturally exposed to Mycobacterium bovis. FEMS Immunol Med Microbiol 46: 298–312. doi:10.1111/j.1574-695X.2005.00035.x.

45. Fernández de Mera IG, Pérez de la Lastra JM, Ayoubi P, Naranjo V, Kocan KM, et al. (2008) Differential expression of inflammatory and immune response genes in mesenteric lymph nodes of Iberian red deer (Cervus elaphus hispanicus) naturally infected with Mycobacterium bovis. Dev Comp Immunol 32: 85–91. doi:10.1016/j.dci.2007.05.001.

46. Möller M, Hoal EG (2010) Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. Tuberculosis (Edinb) 90: 71–83. doi:10.1016/j.tube.2010.02.002.

47. Babb C, van der Merwe L, Beyers N, Pheiffer C, Walzl G, et al. (2007) Vitamin D receptor gene polymorphisms and sputum conversion time in pulmonary tuberculosis patients. Tuberculosis 87: 295–302.

48. Rossouw M, Nel HJ, Cooke GS, van Helden PD, Hoal EG (2003) Association between tuberculosis and a polymorphic NFkappaB binding site in the interferon gamma gene. Lancet 361: 1871–1872.

49. Hill AVS (2012) Evolution, revolution and heresy in the genetics of infectious disease susceptibility. Philos Trans R Soc Lond, B, Biol Sci 367: 840–849. doi:10.1098/rstb.2011.0275.

50. De Wit E, van der Merwe L, van Helden PD, Hoal EG (2011) Gene-gene interaction between tuberculosis candidate genes in a South African population. Mamm Genome 22: 100–110. doi:10.1007/s00335-010-9280-8.

51. Velez DR, Hulme WF, Myers JL, Stryjewski ME, Abbate E, et al. (2009) Association of SLC11A1 with tuberculosis and interactions with NOS2A and TLR2 in African-Americans and Caucasians. Int J Tuberc Lung Dis 13: 1068–1076.

Chapter 4

Novel SNP discovery in African buffalo, *Syncerus caffer*, using high-throughput sequencing

Nikki le Roex[1], Harry Noyes[2], Andrew Brass[3], Daniel G Bradley[4], Steven J Kemp[2,5], Suzanne Kay[2], Paul D van Helden[1], Eileen G Hoal[1]

1 DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/ MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, Stellenbosch University, South Africa. **2** Institute of Integrative Biology, University of Liverpool, L69 7ZB, Liverpool, UK. **3** School of Computer Science, University of Manchester, M13 9PL, Manchester, UK. **4** Molecular Population Genetics, Smurfit Institute of Genetics, Trinity College Dublin, Dublin 2, Ireland. **5** International Livestock Research Institute P.O. Box 30709, Nairobi, 00100 Kenya.

PLoS ONE (2012); 7(11): e48792.

Abstract

The African buffalo, *Syncerus caffer*, is one of the most abundant and ecologically important species of megafauna in the savannah ecosystem. It is an important prey species, as well as a host for a vast array of nematodes, pathogens and infectious diseases, such as bovine tuberculosis and corridor disease. Large-scale SNP discovery in this species would greatly facilitate further research into the area of host genetics and disease susceptibility, as well as provide a wealth of sequence information for other conservation and genomics studies. We sequenced pools of Cape buffalo DNA from a total of 9 animals, on an ABI SOLiD4 sequencer. The resulting short reads were mapped to the UMD3.1 *Bos taurus* genome assembly using both BWA and Bowtie software packages. A mean depth of 2.7X coverage over the mapped regions was obtained. Btau4 gene annotation was added to all SNPs identified within gene regions. Bowtie and BWA identified a maximum of 2,222,665 and 276,847 SNPs within the buffalo respectively, depending on analysis method. A panel of 173 SNPs was validated by fluorescent genotyping in 87 individuals. 27 SNPs failed to amplify, and of the remaining 146 SNPs, 43-54% of the Bowtie SNPs and 57-58% of the BWA SNPs were confirmed as polymorphic. dN/dS ratios found no evidence of positive selection, and although there were genes that appeared to be under negative selection, these were more likely to be slowly evolving house-keeping genes.

1. Introduction

Investigating genetic variation between individuals, populations or species provides the basis for understanding the heritability of traits and phenotypes, and provides researchers with the opportunity to study complex issues in conservation, disease susceptibility, molecular ecology and many other disciplines [1]. Animal models have long provided the foundation for the genetic analysis of complex traits, using gene knockouts, recombinant and transgenic animals. The advent of whole-genome sequencing and the subsequent assembly of the genomes of the domestic cow, mouse, rat and others provides a resource pool of genetic markers with which to work in subsequent studies [2,3]. However, when studying a species for which limited sequence information is available, it is necessary to generate sequence data in order to identify the genetic variants present in the population. The discovery of genetic variants, in particular large-scale SNP discovery, identifies genetic markers that may have the

power to answer multiple research questions, and can be performed using a wide variety of technologies [4,5].

The development of next-generation sequencing technologies, such as the ABI SOLiD, Illumina GA and HiSeq and Roche 454 platforms, has enabled a faster and more cost-effective approach to generating sequence data and SNP discovery. Each of these platforms has its own particular chemistry and combination of template preparation, sequencing and data analysis, and thus each has its own advantages and disadvantages [6,7]. Benefits of working with SNP data include the fact that they are abundant, distributed throughout the genome, are easy to score and can be used in high-throughput screening [1,8]. The sequence data obtained using next-generation technologies are typically short sequence reads of approximately 50-750bp. Whole genome shotgun or re-sequencing using short reads requires the alignment of millions of sequence reads to a high quality reference genome sequence. Once the reads have been mapped to the reference genome, nucleotide variation between the sample and the reference can be identified [4]. For most species that do not have a fully sequenced genome, this can be problematic. However, if the genome of a closely related species is available, this can be used as a reference, although the proportion of data that can be mapped to the genome may be significantly reduced, and the data obtained will indicate variation seen between species as well as between individuals of the same species [9–11]. Other considerations such as depth of coverage and quality must be taken into account, particularly when using this approach.

The African buffalo, *Syncerus caffer*, is one of the most abundant and ecologically important species of megafauna in the savannah ecosystem. There are three recognised subspecies of African buffalo – the Cape buffalo (*Syncerus caffer caffer*), the West African buffalo (*Syncerus caffer brachyceros*) and the Forest buffalo (*Syncerus caffer nanus*). A fourth subspecies, the Central African Savannah buffalo (*Syncerus caffer aequinoctialis*), has also been proposed. All individuals used in this study were *Syncerus caffer caffer*. African buffalo are an important prey species, as well as a host for a vast array of nematodes, pathogens and infectious diseases [12]. Some of the infectious diseases for which the African buffalo is a wildlife host are bovine tuberculosis (BTB), corridor disease and foot-and-mouth disease [13]. Their current distribution extends throughout much of sub-Saharan Africa, although within this range the distribution is fragmented and largely confined to protected areas. The most recent IUCN census data estimates the global African buffalo population to be approximately 900,000 [14]. Previous genetic studies on buffalo have focused primarily on

population differentiation and genetic diversity, using markers such as microsatellites, mitochondrial DNA and Y chromosomal loci [15–17]. Large-scale SNP discovery in this species would greatly facilitate further research into the area of host genetics and disease susceptibility, as well as provide a wealth of sequence information for other conservation and genomics studies.

Our goals in this study were to perform next-generation sequencing and mapping of the African buffalo genome using two different software packages, implement large-scale identification of novel SNPs within the African buffalo genome, and determine and compare the SNP validation rates using fluorescent genotyping. This data can then inform future studies.

2. Materials and Methods

2.1 Ethics

The Stellenbosch University Animal Care and Use Committee (SU ACU) deemed it unnecessary to obtain ethical clearance for this study as the blood samples used for DNA extraction were collected under the directive of SANParks and KwaZulu Natal Wildlife for other purposes, and their use in the present study is incidental.

2.2 Samples

EDTA blood samples from nine Cape buffalo from Hluhluwe iMfolozi Park, in the KwaZulu Natal province of South Africa, were obtained during an annual BTB test and cull operation. Genomic DNA was extracted from the blood samples using the Qiagen DNeasy Blood and Tissue Kit, according to the manufacturer‟s instructions. Equimolar amounts of DNA from four and five animals were made into two barcoded paired end libraries. Fragment libraries were constructed according to manufacturer‟s instructions using the SOLiD™ Fragment Library Construction Kit, and sequenced on a single plate of an ABI SOLiD4 sequencer (Lifetech) to generate 50bp reads.

2.3 Mapping and SNP detection

Due to the absence of an assembled African buffalo reference genome, the *Bos taurus* UMD3.1 genome sequence, produced by the Centre for Bioinformatics and Computational Biology (CBCB) at the University of Maryland, was used as the reference genome in this

study [18] . The short reads were mapped to the *Bos taurus* genome in colour space using two separate packages, BWA [19] and Bowtie [20]. SNPs were called from Bowtie reads with pileup and mpileup, and SNPs were called from BWA reads with pileup and Genome Analysis Tool Kit 1.1-3 (GATK) [21]. Reads were initially mapped with Bowtie with the following parameters -C (colorspace); --snpfrac 0.01 (1% of positions expected to be polymorphic); -3 1 (trim one base from the 3‟ end of the read). SNPs were extracted from the reads mapped with Bowtie using pileup in SAMtools with the –c and –r 0.01 (1% of positions expected to be polymorphic) options [22]. 173 SNPs identified by this pipeline were submitted for validation. Subsequently the data was reanalysed using Bowtie with --snpfrac 0.001, followed by Picard Tools v1.48 to remove duplicates, and SAMtools mpileup with options -B (Disable probabilistic realignment); -d 29 (maximum depth 29) to identify SNPs. BWA was run with option -q 20 (trim low quality reads from 3‟ end). SNPs were extracted from reads mapped with BWA with SAMtools using the same options as for Bowtie, and also with Picard Tools and GATK [21] for base quality score recalibration, indel realignment, duplicate removal, and SNP and INDEL discovery and genotyping [23]. Shellscripts for calling SNPs with GATK are available from the authors on request. Nucleotide differences called by SAMtools were classified as SNPs if (i) the alternate allele was supported by a minimum coverage of 2, and (ii) the alternate allele had a minimum Phred quality score of 20. Nucleotide differences called by GATK were classified as SNPs if they were flagged as “PASS”. The use of the *Bos taurus* reference genome in this study enabled the identification of two classes of SNPs - those that occur between the cow and buffalo, and those that occur within the buffalo population.

2.4 Annotation

Gene annotation was added to the SNPs identified within whole gene regions, which were defined as from 1000bp upstream of the 5‟ end to 1000bp downstream of the 3‟ end of any gene. A local Perl script was then used to submit the SNPs within gene regions to the Ensembl v66 Application Programme Interface (API) to determine functional consequence [24,25]. SNPs and annotations were written to a local MySQL database for storage and interrogation.

2.5 Signatures of selection

The ratio of non-synonymous to synonymous substitutions (dN/dS) was calculated for all buffalo genes using PAML [26]. The Ensembl v66 API was used to identify the positions of

SNPs within the coding sequence of bovine genes (*Bos taurus* genome build UMD3.1), and Cape buffalo SNPs were substituted into the gene coding sequences at these positions before submission to a local copy of PAML to obtain the dN/dS ratios. P values were obtained by taking twice the absolute difference between the log likelihood of the observed dN/dS ratio and the ratio obtained with dN/dS set to 1 and comparing it to χ^2 with one degree of freedom. The Database for Annotation, Visualization, and Integrated Discovery (DAVID) was used to identify whether there were pathways that were overrepresented in the gene lists [27].

2.6 Validation

A selection of 173 SNPs was made from the collective BWA pileup and Bowtie pileup SNP pool based on the predicted consequence and/or gene location of the SNP, in order to use the validation data in a subsequent case-control association study. SNPs were fluorescently genotyped in 87 Cape buffalo by KBioscience (www.kbioscience.co.uk), using their competitive allele-specific KASP SNP genotyping platform. Assays were designed for each of the SNPs by KBioscience. The KASP assay system is a competitive allele-specific PCR incorporating a FRET quencher cassette (www.kbioscience.co.uk). All but six of the SNPs validated as polymorphic within Cape buffalo were also validated in the original 9 individuals from the sequencing pool; the six SNPs not genotyped in the original samples were omitted due to technical reasons. KBioscience assay IDs are shown in Table S1 and are available to other users on application to KBiosciences.

3. Results

3.1 Mapping

The number of reads generated by the SOLiD sequencing was 442,984,249 in total. Sequence reads have been submitted to the sequence read archive at NCBI (accession number SRA051668.1). Given the relatively large genetic distance between *Bos taurus* and Cape buffalo, we initially used a high value of snpfrac (0.01) in the Bowtie parameters; this is the proportion of positions that are expected to be polymorphic. With snpfrac at 0.01, 23% of the reads mapped to the reference genome but the validation rate of the SNPs was relatively low (43%; see below). We therefore repeated the mapping with snfrac = 0.001, which reduced the percentage mapping to 21% and increased the validation rate to 54%. The mapped Bowtie reads were all 47 bases in length and covered 1.58GB, which is equivalent to 56% of the

bovine genome being covered by at least one read, and represents 2.7X coverage of aligned data. Even fewer reads mapped with BWA (19%), of which 9.7% were PCR duplicates. The BWA reads were trimmed based on the base quality; 61% of reads were 48 bases and the mean length was 44.5 bases. The percentage of 35bp reverse reads that mapped was very low with both alignment programmes (<2%) and these data were not used. The percentage of exon sequence covered by at least one read was slightly higher (64%) than that of the genome as a whole, which is consistent with exons occurring in more highly conserved regions of the genome.

3.2 SNP discovery and annotation

In order to be identified as SNPs, nucleotide variants called by SAMtools had to be supported by at least two reads, and have a base quality of at least 20, indicating 99% accuracy in the call. SNPs identified by GATK had to be classified as „PASS". A flow diagram of the analysis is shown in Figure 1. The number of SNPs called by different pipelines varied substantially, with BWA followed by pileup having about a tenth of the other methods (Figure 1 and Table 2).

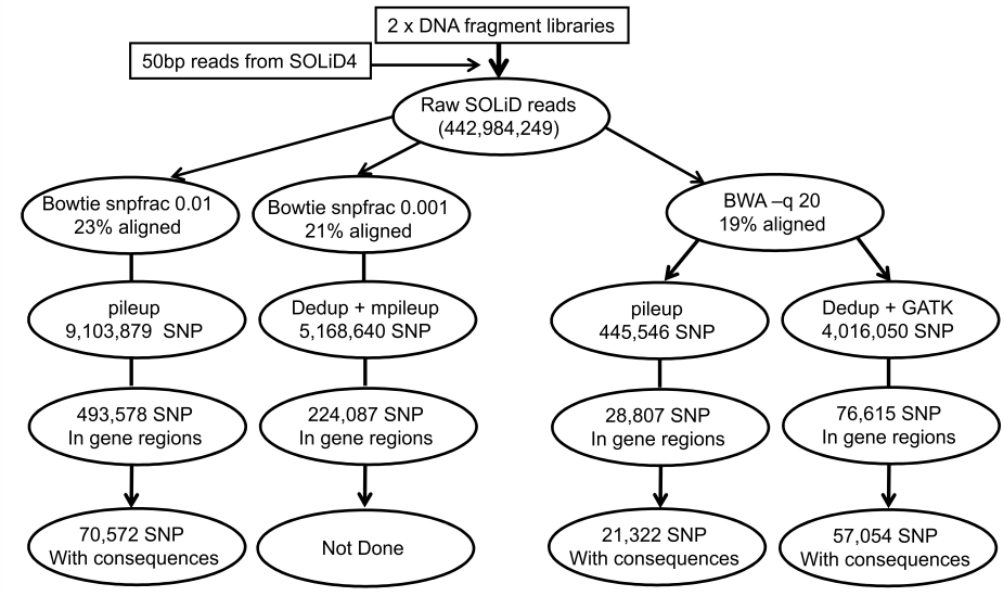

Figure 1: SNP discovery pipeline using Bowtie and BWA

GATK was run on the same BWA generated bam files as pileup and found 7 times more SNPs, so the difference in this case was due to the SNP calling and not the mapping. Bowtie and mpileup generated about half the number of SNPs as Bowtie and pileup; however, in this case Bowtie was run with snpfrac at 0.001 and 0.01 respectively and Picard tools was used to remove duplicates prior to mpileup but not pileup, so it is likely that the number of SNPs observed was reduced at several points in the pipeline leading to mpileup compared to that leading to pileup.

Table 2: Comparison of SNPs identified after mapping with BWA and Bowtie

Number of SNPs	BWA pileup	BWA GATK	Bowtie pileup	Bowtie mpileup
Relative to reference sequence	445,546	3,739,203	6,881,214	4,234,692
Within buffalo population	109,361	276,847	2,222,665	933,948
Within gene regions*	28,807	76,615	493,578	224,087

*Gene regions are from the gene start to the gene end and include introns

3.3 SNP validation

173 SNPs were sent for validation on 87 Cape buffalo DNA samples, and 146 were amplified successfully. Of these 146 SNPs, 71 appeared to be monomorphic and 75 were polymorphic. The different mapping and SNP calling methods produced different numbers of validated SNPs (Table 1).

Table 1: Counts of SNP that converted to successful assays predicted by different pipelines

Validation result	BWA pileup	BWA GATK	Bowtie pileup	Bowtie mpileup
Monomorphic	20 (12)*	31 (23)	71 (11)	46 (21)
Polymorphic	45 (6)	68 (12)	75 (4)	73 (13)
Not amplified. Number, (%)	12, 16% (1)	20, 17% (1)	27, 16% (0)	23, 13% (0)
Validated %	58%	57%	43%	51%
Percent of assays correctly identified as polymorphic	58% (67%)	57% (67%)	43% (45%)	54% (74%)

*Figures in brackets are the number of loci that were predicted monomorphic (fixed for alternate allele) by the corresponding pipeline. The percentage correctly identified as polymorphic shows the percentage of all assays that were polymorphic and in brackets the percentage of assays predicted polymorphic that were confirmed.

The difference in validation rates between the three software packages was statistically significant (Chi sq test, χ 28, 4df, $p < 0.001$). BWA followed by SAMtools Pileup or GATK did not detect 20 or 7 SNPs respectively that were detected by Bowtie and Pileup and validated as polymorphic. SNPs called by GATK were scored as "PASS" or "LowQual"; there was no difference in validation rate between these two classes of SNPs (ChiSq =0.6, p = 0.45). Validated SNPs and KBioscience assay IDs can be seen in Table S1. Given the higher validation rate from the BWA data, this data was used for the subsequent analysis. An unbiased pattern of substitution would be expected to show a transition/transversion (Ti/Tv) ratio of 0.5, as transversions are twice as likely as transitions. However, a ratio of around 2.1 is usually observed in mammals, and a ratio that is significantly lower than 2.1 can be an indicator of poor quality sequencing data [23]. The observed Ti/Tv ratio for homozygous SNPs in this study was 2.18. Therefore, despite the relatively low validation rate, the Ti/Tv ratio suggested that SNP calls were in the expected proportions and were therefore not random.

3.4 SNP Consequences

The predicted consequences of the SNPs were obtained from Ensembl and a summary of the numbers of the different types of consequences obtained using the BWA and GATK pipeline is shown in Table 3.

Table 3: Counts of BWA/GATK SNPs of each consequence within gene regions

SNP CONSEQUENCE	COUNT
DOWNSTREAM	3,400
STOP_LOST	1
UPSTREAM	3,265
INTRONIC	67,542
WITHIN_MATURE_miRNA	1
STOP_GAINED	26
SYNON	1,408
SPLICE_SITE	205
NON-SYNON	982
3'UTR	690
5'UTR	54
WITHIN_NON_CODING_GENE	106
ESSENTIAL_SPLICE_SITE	8

11,054 SNP had multiple annotations either because they were in multiple transcripts or because a SNP had more than one consequence e.g. SPLICE_SITE and INTRONIC and 3,105 SNP within gene regions had no annotation. A complete list of SNPs in exons with consequences is shown in Table S2. DAVID was used to discover whether any specific pathways were particularly affected by premature stop codons but none were significantly overrepresented after correction for multiple testing (minimum p value =0.706) suggesting that these genes were affected at random. A web page was created where SNPs could be filtered by genomic region, gene, SNP consequence or restriction site (http://www.genomics.liv.ac.uk/tryps/resources.html) [28]. The inclusion of the restriction sites affected makes it possible to obtain lists of SNPs that can be used in RFLP assays. These are very useful where a set of markers is required across a genome region but the particular SNPs assayed are not important. The website implements Primer3 [29] to design primers around SNPs according to user-specified criteria.

3.5 Signatures of selection

Genes were screened for particularly high or low non-synonymous/synonymous substitution ratios (dN/dS) that might be indicative of positive or negative selection, respectively (Table S3). No genes had dN/dS ratios that were significantly higher than 1 (χ^2 $p < 0.05$). DAVID analysis of the 205 genes with no synonymous SNPs suggested that C-type lectins might be overrepresented in this set (Table S4). In contrast 1,231 genes had dN/dS ratios significantly less than 1 ($p<0.05$) suggesting that they were more conserved than expected. Genes involved in the regulatory processes “nucleoside binding” and “phosphorylation” were the most overrepresented (Table S5).

4. Discussion

The African buffalo has become a species of interest in recent years due to its role as a wildlife maintenance host for a variety of infectious and zoonotic diseases, such as corridor disease, foot-and-mouth disease and bovine tuberculosis [13]. There is no African buffalo reference genome available for use in disease association studies, and much benefit would be gained by the generation of sequence data and large-scale SNP discovery in this species. The cost efficiency, large data output and fast turnaround time of the next-generation sequencing technologies have greatly facilitated the generation of novel sequence data as well as large-scale SNP discovery in non-model organisms. ABI SOLiD technology was used in this study

to generate over 400 million 50bp reads of African buffalo genome sequence, and for the preliminary identification of approximately a quarter of a million novel SNPs within the buffalo genome. When investigating a species for which a complete genome sequence is not available, the reference genome of a related species can be used for mapping and SNP discovery. The success of this approach was recently shown by mapping sequence reads of the great tit, *Parus major*, to the reference genome of the zebra finch, *Taeniopygia guttata* [10]. Using this method, the authors were able to identify 20,000 novel SNPs. A similar approach was used in mapping sequence reads of the turkey, *Meleagris gallopavo*, to the sequenced genome of the chicken, and approximately 8000 SNPs were identified in the turkey genome [30].

In this study, the Cape buffalo sequence reads were mapped to the reference genome of the domestic cow, *Bos taurus.* The most recent common ancestor of the domestic cow and the African buffalo is estimated to have existed approximately 5-10 million years ago (MYA), at the time of the divergence of the subtribes Bubalina, which consists of the *Syncerus* and *Bubalus* genera, and Bovina, which is comprised of the *Bos* and *Bison* genera [31]. Despite the relatively recent split of these two genera, only 19% to 23% of the buffalo short reads mapped to the reference cow genome using BWA and Bowtie. These percentages are comparable to those achieved by van Bers *et al*. [10], where 26% and 32% of the great tit sequence reads generated in two pools were mapped to the zebra finch genome using data generated by the Illumina Genome Analyser. Similarly, Kerstens *et al*. [30] mapped approximately 30% of the raw sequence reads generated in the turkey to the chicken genome. The low percentage mapping resulted in a depth of coverage that was substantially below a level that would be ideal for SNP discovery. It would appear that 2-4 times more raw data is required when using a distant relative as a reference sequence, although the longer reads that are now available from the Illumina sequencers (100-150bp) might improve the percentage mapping. Nevertheless the extra data required when using a distant relative is still substantially less than would be required for a *de novo* assembly (>80X coverage), and mapping to a related species dramatically simplifies the SNP discovery and annotation compared to that required for a *de novo* assembly.

Significantly different results were found in this study when using different mapping and SNP calling methods (Table 2). From the combined BWA pileup and Bowtie pileup SNP pool, 173 SNPs from within the Cape buffalo population were selected for validation. The selection was based on gene function and SNP consequence, in order to use the validation

data in a subsequent case-control association study. Validation required SNPs to pass two tests 1) that the loci amplified and 2) that the loci were polymorphic in Cape buffalo. The number of SNPs detected by each method that were validated ranged from 45–75, and the percentage ranged from 43–58% (Table 1). BWA and GATK had highest percent validated (57%), although it is probably not statistically significantly better than the Bowtie/mpileup combination (54% validated), which yielded more SNPs. The BWA/GATK SNPs were used for the functional analyses discussed below. It is not possible to accurately estimate false positive rates for any method except Bowtie and pileup since all SNPs assayed were originally identified with this method. The identification of false positives in SNP discovery may be a result of sequencing errors, alignment errors or the occurrence of paralogous sequence variants [1]. The error rate of the ABI SOLiD is estimated at 0.0006% [32] , which would lead to about 50,000 erroneous base calls or 10% of the BWA and 0.5% of the Bowtie SNPs; in either case this is insufficient to explain the observed error rate, particularly since the Ti/Tv ratio was 2.18 as would be expected for a mammalian genome, rather than a rate nearer 0.5 which would be expected from random SNP calls. Therefore it seems more likely that the errors resulted from the low coverage (2.7X) and problems with aligning to the genome of a different species [33]. Since selected SNPs from this study will be used in future candidate gene association studies, it was desirable to identify as many SNPs as possible in order to obtain the largest number of candidates, and therefore false positives were considered less problematic than false negatives. SNPs within genes were annotated using the Ensembl SNP annotation API. Although this provided consequences for 57,054 SNPs, no general conclusions could be drawn from the data. There were no pathways overrepresented amongst the genes associated with the 27 SNPs that modified stop codons. It should be noted that the 1000 genomes project found 250-300 loss of function variants within genes in each individual [34], so the discovery of 27 in a pool of DNA from nine individuals probably does not have significant biological implications.

The non-synonymous/synonymous substitution ratios provided very little evidence for genes being under positive selection, although a large number appeared to be under purifying selection. This may be because this approach was developed for comparing distantly related taxa, and the relatively small number of SNPs between the Cape buffalo and the cow (mean 4.2 per gene) means that the power to detect positive selection is very limited [35]. The large number of genes that are apparently under purifying selection in the buffalo may contain many that are too slowly evolving to exhibit a detectable signal between these two species.

Housekeeping genes appeared to be overrepresented in this list but this may be a consequence of their tendency to be relatively slowly evolving rather than evidence of purifying selection.

This investigation of the Cape buffalo, a species without an assembled genome, has yielded a wealth of data that will provide useful tools for further study of this species and related aspects.

Acknowledgements

We gratefully acknowledge Dave Cooper, the Hluhluwe iMfolozi Park staff and SANParks for providing the samples used in this study. We thank Ad Koets for helpful discussion.

Supporting Information

All data based on BWA mapping.

Table S1 List of the 69 validated SNPs, their Btau4 and UMD3 positions and KBioscience assay IDs
ValidatedSNPs.txt

Table S2 Complete list of BWA/GATK SNPs with consequences and annotated gene names
BwaSNPwithConsequences.txt

Table S3 List of genes that had at least one synonymous or non-synonymous SNP, with dN/dS ratios and p values calculated with PAML
Annotated_dNdSratios.txt

Table S4 David annotation of genes with dN/dS ratios of 99, which are mainly those with no synonymous SNPs
DAVIDAnalysisDnDS=99

Table S5 David annotation of genes that had dN/dS ratios significantly less than 1
DAVIDAnalysisDnDS_P0.05

References

1. Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. Molecular Ecology Resources 10: 915–934. doi:10.1111/j.1755-0998.2010.02891.x.

2. Shendure J, Ji H (2008) Next-generation DNA sequencing. NatBiotechnol 26: 1135–1145.

3. Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, et al. (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources 11: 1–8. doi:10.1111/j.1755-0998.2010.02979.x.

4. Imelfort M, Duran C, Batley J, Edwards D (2009) Discovering genetic polymorphisms in next-generation sequencing data. Plant BiotechnolJ 7: 312–317.

5. Tautz D, Ellegren H, Weigel D (2010) Next Generation Molecular Ecology. Molecular Ecology 19: 1–3. doi:10.1111/j.1365-294X.2009.04489.x.

6. Voelkerding KV, Dames SA, Durtschi JD (2009) Next-Generation Sequencing: From Basic Research to Diagnostics. Clin Chem 55: 641–658. doi: 10.1373/clinchem.2008.112789 .

7. Metzker ML (2010) Sequencing technologies - the next generation. NatRevGenet 11: 31–46.

8. Morin PA, Luikart G, Wayne RK, the SNP workshop group (2004) SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution 19: 208–216. doi:16/j.tree.2004.01.009.

9. Collins LJ, Biggs PJ, Voelckel C, Joly S (2008) An approach to transcriptome analysis of non-model organisms using short-read sequences. Genome Inform 21: 3–14.

10. Van Bers NEM, van Oers K, Kerstens HHD, Dibbits BW, Crooijmans RPMA, et al. (2010) Genome-wide SNP detection in the great tit Parus major using high throughput sequencing. Molecular Ecology 19: 89–99. doi:10.1111/j.1365-294X.2009.04486.x.

11. Everett MV, Grau ED, Seeb JE (2011) Short reads and nonmodel species: exploring the complexities of next-generation sequence assembly and SNP discovery in the absence of a reference genome. Molecular Ecology Resources 11: 93–108. doi:10.1111/j.1755-0998.2010.02969.x.

12. Jolles AE (2007) Population biology of African buffalo (Syncerus caffer) at Hluhluwe-iMfolozi Park, South Africa. African Journal of Ecology 45: 398–406.

13. Grobler JP, Van der Bank FH (1996) Genetic diversity and isolation in African buffalo (Syncerus caffer). Biochemical Systematics and Ecology 24: 757–761.

14. IUCN Red List (2012). IUCN Red List. Available: http://www.iucnredlist.org/apps/redlist/details/21251/0. Accessed 13 July 2011.

15. Van Hooft WF, Groen AF, Prins HH (2002) Phylogeography of the African buffalo based on mitochondrial and Y-chromosomal loci: Pleistocene origin and population expansion of the Cape buffalo subspecies. MolEcol 11: 267–279.

16. Van Hooft WF, Groen AF, Prins HHT (2003) Genetic structure of African buffalo herds based on variation at the mitochondrial D-loop and autosomal microsatellite loci: Evidence for male-biased gene flow. Conservation Genetics 4: 467–477.

17. Heller R, Lorenzen ED, Okello JB, Masembe C, Siegismund HR (2008) Mid-Holocene decline in African buffalos inferred from Bayesian coalescent-based analyses of microsatellites and mitochondrial DNA. MolEcol 17: 4845–4858.

18. Zimin AV, Delcher AL, Florea L, Kelley DR, Schatz MC, et al. (2009) A whole-genome assembly of the domestic cow, Bos taurus. Genome Biol 10: R42. doi:10.1186/gb-2009-10-4-r42.

19. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25: 1754–1760. doi:10.1093/bioinformatics/btp324.

20. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25. doi:10.1186/gb-2009-10-3-r25.

21. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. (2010) The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297–1303. doi:10.1101/gr.107524.110.

22. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079. doi:10.1093/bioinformatics/btp352.

23. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature Genetics 43: 491–498. doi:10.1038/ng.806.

24. Rios D, McLaren WM, Chen Y, Birney E, Stabenau A, et al. (2010) A database and API for variation, dense genotyping and resequencing data. BMC Bioinformatics 11: 238. doi:10.1186/1471-2105-11-238.

25. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, et al. (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26: 2069–2070. doi:10.1093/bioinformatics/btq330.

26. Yang Z (1997) PAML: A Program Package for Phylogenetic Analysis by Maximum Likelihood. Comput Appl Biosci 13: 555–556. doi:10.1093/bioinformatics/13.5.555.

27. Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, et al. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 4: R60–R60.

28. Noyes H (2011) Cape Buffalo SNP. Available: http://www.genomics.liv.ac.uk/tryps/capeSNP.php.

29. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365–386.

30. Kerstens H, Crooijmans R, Veenendaal A, Dibbits B, Chin-A-Woeng T, et al. (2009) Large scale single nucleotide polymorphism discovery in unsequenced genomes using second generation high throughput sequencing technology: applied to turkey. BMC Genomics 10: 479. doi:10.1186/1471-2164-10-479.

31. MacEachern S, McEwan J, Goddard M (2009) Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (Bovidae, Bovinae). BMC Genomics 10: 177. doi:10.1186/1471-2164-10-177.

32. Applied Biosystems (2010) Applied Biosystems SOLiD4 System. Available: http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_080173.pdf.

33. Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92: 255–264. doi:16/j.ygeno.2008.07.001.

34. 1000 Genomes Project Consortium (2010) A map of human genome variation from population-scale sequencing. Nature 467: 1061–1073. doi:10.1038/nature09534.

35. Kryazhimskiy S, Plotkin JB (2008) The Population Genetics of dN/dS. PLoS Genet 4: e1000304. doi:10.1371/journal.pgen.1000304.

Supplementary information

Table S1: List of the 69 validated SNPs, their Btau4 and UMD3 positions and KBioscience assay IDs

Name	Assay ID	UMD3.1 Position	Btau4 Position	Consequence of SNP	Gene Name	Ensembl Gene ID	Description of gene
SNP_1	1538-0001	Chr11:7179447	Chr11:7403981	NON-SYNON	Q75R58_BOVIN	ENSBTAG00000001034	IL-18 receptor alpha Fragment [Source:UniProtKB/TrEMBL;Acc:Q75R58]
SNP_4	1538-0004	Chr17:63612902	Chr17:64549854	NON-SYNON	NP_001019728.1	ENSBTAG00000014628	2-prime-5-prime-oligoadenylate synthetase 2 [Source:RefSeq peptide;Acc:NP_001019728]
SNP_6	1538-0006	Chr1:111031389	Chr1:112144359	NON-SYNON	PTX3_BOVIN	ENSBTAG00000009012	Pentraxin-related protein PTX3 Precursor (Pentaxin-related protein PTX3) [Source:UniProtKB/Swiss-Prot;Acc:Q0VCG9]
SNP_7	1538-0007	Chr1:143152388	Chr1:144420892	NON-SYNON	MX2_BOVIN	ENSBTAG00000008471	Interferon-induced GTP-binding protein Mx2 (Myxovirus resistance protein 2) [Source:UniProtKB/Swiss-Prot;Acc:Q9BDI7]
SNP_8	1538-0008	Chr22:53599791	Chr22:54390942	NON-SYNON	CCR5_BOVIN	ENSBTAG00000008099	C-C chemokine receptor type 5 (C-C CKR-5)(CC-CKR-5)(CCR-5)(CCR5)(CD195 antigen) [Source:UniProtKB/Swiss-Prot;Acc:Q2HJ17]
SNP_13	1538-0013	Chr3:78427756	Chr3:83795862	NON-SYNON	B2MU63_BOVIN	ENSBTAG00000014983	Interleukin 23 receptor [Source:UniProtKB/TrEMBL;Acc:B2MU63]
SNP_16	1538-0016	Chr5:36884961	Chr5:39910725	NON-SYNON	IRAK4_BOVIN	ENSBTAG00000021105	Interleukin-1 receptor-associated kinase 4 (IRAK-4)(EC 2.7.11.1) [Source:UniProtKB/Swiss-Prot;Acc:Q1RMT8]
SNP_17	1538-0017	Chr5:36890957	Chr5:39916721	NON-SYNON	IRAK4_BOVIN	ENSBTAG00000021105	Interleukin-1 receptor-associated kinase 4 (IRAK-4)(EC 2.7.11.1) [Source:UniProtKB/Swiss-Prot;Acc:Q1RMT8]
SNP_18	1538-0018	Chr5:47797051	Chr5:51485425	NON-SYNON	B3FXL5_BOVIN	ENSBTAG00000007636	Interleukin-1 receptor-associated kinase M transcript variant 3 [Source:UniProtKB/TrEMBL;Acc:B3FXL5]
SNP_19	1538-0019	Chr5:47817597	Chr5:51505676	NON-SYNON	B3FXL5_BOVIN	ENSBTAG00000007636	Interleukin-1 receptor-associated kinase M transcript variant 3 [Source:UniProtKB/TrEMBL;Acc:B3FXL5]
SNP_25	1538-0025	Chr7:98739431	Chr7:97743988	NON-SYNON	ERAP2_BOVIN	ENSBTAG00000039275	Endoplasmic reticulum aminopeptidase 2 (EC 3.4.11.-) [Source:UniProtKB/Swiss-Prot;Acc:A6QPT7]
SNP_30	1538-0030	Chr20:33550138	Chr20:35669518	3‟UTR	CO7_BOVIN	ENSBTAG00000011766	Complement component C7 Precursor [Source:UniProtKB/Swiss-Prot;Acc:Q29RQ1]
SNP_41	1538-0041	Chr14:79862771	Chr14:76297528	NON-SYNON	A7E3E3_BOVIN	ENSBTAG00000040461	SLC7A13 protein Fragment [Source:UniProtKB/TrEMBL;Acc:A7E3E3]
SNP_56	1538-0056	Chr22:55714039	Chr22:56497049	3‟UTR	NP_001071304.1	ENSBTAG00000005857	"solute carrier family 6 (neurotransmitter transporter, GABA), member 1 [Source:RefSeq peptide;Acc:NP_001071304]"
SNP_57	1538-0057	Chr3:103930723	Chr3:110248862	3‟UTR	GTR1_BOVIN	ENSBTAG00000009617	"Solute carrier family 2, facilitated glucose

Name	Assay ID	UMD3.1 Position	Btau4 Position	Consequence of SNP	Gene Name	Ensembl Gene ID	Description of gene
							transporter member 1 (Glucose transporter type 1, erythrocyte/brain)(GLUT-1) [Source:UniProtKB/Swiss-Prot;Acc:P27674]"
SNP_59	1538-0059	Chr5:45246904	Chr5:48631315	3"UTR	S35E3_BOVIN	ENSBTAG00000031919	Solute carrier family 35 member E3 [Source:UniProtKB/Swiss-Prot;Acc:A4IFK2]
SNP_62	1538-0062	Chr7:23330398	Chr7:20718957	3"UTR	NP_001039967.1	ENSBTAG00000000668	solute carrier family 22 member 5 [Source:RefSeq peptide;Acc:NP_001039967]
SNP_63	1538-0063	Chr9:63303181	Chr9:64264202	3"UTR	NP_001029809.1	ENSBTAG00000014057	"solute carrier family 35 (CMP-sialic acid transporter), member A1 [Source:RefSeq peptide;Acc:NP_001029809]"
SNP_64	1538-0064	Chr11:96772566	Chr11:100196170	3"UTR	NP_001039918.1	ENSBTAG00000013314	pre-B-cell leukemia transcription factor 3 [Source:RefSeq peptide;Acc:NP_001039918]
SNP_65	1538-0065	Chr11:96772626	Chr11:100196230	3"UTR	NP_001039918.1	ENSBTAG00000013314	pre-B-cell leukemia transcription factor 3 [Source:RefSeq peptide;Acc:NP_001039918]
SNP_67	1538-0067	Chr13:59903453	Chr13:60149955	3"UTR	NP_001068977.1	ENSBTAG00000002065	transcription factor AP-2 gamma [Source:RefSeq peptide;Acc:NP_001068977]
SNP_75	1538-0075	Chr20:8038747	Chr20:8480635	3"UTR	NP_001020485.1	ENSBTAG00000010012	basic transcription factor 3 [Source:RefSeq peptide;Acc:NP_001020485]
SNP_76	1538-0076	Chr23:27968593	Chr23:28088619	5"UTR	NP_001094527.1	ENSBTAG00000010698	"general transcription factor IIH, polypeptide 4, 52kDa [Source:RefSeq peptide;Acc:NP_001094527]"
SNP_80	1538-0080	Chr25:33291910	Chr25:34830312	3"UTR	GTF2I_BOVIN	ENSBTAG00000009780	General transcription factor II-I (GTFII-I)(TFII-I) [Source:UniProtKB/Swiss-Prot;Acc:A7MB80]
SNP_82	1538-0082	Chr25:33297474	Chr25:34835878	SYNON	GTF2I_BOVIN	ENSBTAG00000009780	General transcription factor II-I (GTFII-I)(TFII-I) [Source:UniProtKB/Swiss-Prot;Acc:A7MB80]
SNP_95	1538-0095	Chr5:28826280	Chr5:31609528	3"UTR	NP_001076866.1	ENSBTAG00000019312	transcription factor CP2 [Source:RefSeq peptide;Acc:NP_001076866]
SNP_99	1538-0099	Chr5:29279068	Chr5:32123465	NON-SYNON	ATF1_BOVIN	ENSBTAG00000018131	Cyclic AMP-dependent transcription factor ATF-1 (Activating transcription factor 1) [Source:UniProtKB/Swiss-Prot;Acc:Q08DA8]
SNP_100	1538-0100	Chr5:95801232	Chr5:102318724	SYNON	NP_001095795.1	ENSBTAG00000003221	activating transcription factor 7 interacting protein [Source:RefSeq peptide;Acc:NP_001095795]
SNP_103	1538-0103	Chr7:47384295	Chr7:44983209	3"UTR	NP_001092656.1	ENSBTAG00000001002	"transcription factor 7 (T-cell specific, HMG-box) [Source:RefSeq peptide;Acc:NP_001092656]"
SNP_104	1538-0104	Chr9:29122178	Chr9:29953026	3"UTR	NP_001076874.1	ENSBTAG00000012946	heat shock transcription factor 2 [Source:RefSeq peptide;Acc:NP_001076874]
SNP_105	1538-0105	Chr3:83571563	Chr3:89312654	NON-SYNON	DOCK7	ENSBTAG00000019877	Dedicator of cytokinesis protein 7 [Source:UniProtKB/Swiss-Prot;Acc:Q96N67]
SNP_109	1538-0109	Chr6:92798767	Chr6:94259546	3"UTR	NP_001095623.1	ENSBTAG00000010956	"scavenger receptor class B, member 2

Name	Assay ID	UMD3.1 Position	Btau4 Position	Consequence of SNP	Gene Name	Ensembl Gene ID	Description of gene
SNP_113	1538-0113	Chr6:38121032	Chr6:37512128	SYNON	OSTP_BOVIN	ENSBTAG00000005260	[Source:RefSeq peptide;Acc:NP_001095623]" Osteopontin Precursor (Bone sialoprotein 1)(Secreted phosphoprotein 1)(SPP-1) [Source:UniProtKB/Swiss-Prot;Acc:P31096]
SNP_114	1538-0114	Chr5:102960154	Chr5:110568539	NON-SYNON	B6ULZ4_BOVIN	ENSBTAG00000034678	WC1 [Source:UniProtKB/TrEMBL;Acc:B6ULZ4]
SNP_118	1538-0118	Chr9:99649203	Chr9:102168606	5‟UTR	NP_001070495.1	ENSBTAG00000011106	PARK2 co-regulated [Source:RefSeq peptide;Acc:NP_001070495]
SNP_119	1538-0119	Chr3:19695912	Chr3:21166345	3‟UTR	TP8L2_BOVIN	ENSBTAG00000020394	"Tumor necrosis factor, alpha-induced protein 8-like protein 2 (TIPE2) [Source:UniProtKB/Swiss-Prot;Acc:Q3ZBK5]"
SNP_120	1538-0120	Chr13:60253499	Chr13:60552967	SPLICE_SITE	NSF1C_BOVIN	ENSBTAG00000006533	NSFL1 cofactor p47 (p97 cofactor p47) [Source:UniProtKB/Swiss-Prot;Acc:Q3SZC4]
SNP_121	1538-0121	Chr3:27020461	Chr3:29341010	SPLICE_SITE	AT1A1_BOVIN	ENSBTAG00000001246	Sodium/potassium-transporting ATPase subunit alpha-1 Precursor (Sodium pump subunit alpha-1)(EC 3.6.3.9)(Na(+)/K(+) ATPase alpha-1 subunit) [Source:UniProtKB/Swiss-Prot;Acc:Q08DA1]
SNP_122	1538-0122	Chr29:22710735	Chr29:23625098	SPLICE_SITE	VGLU2_BOVIN	ENSBTAG00000038347	Vesicular glutamate transporter 2 (VGluT2)(Solute carrier family 17 member 6) [Source:UniProtKB/Swiss-Prot;Acc:A6QLI1]
SNP_123	1538-0123	Chr17:14434651	Chr17:15210331	STOP_GAINED	FREM3	ENSBTAG00000044051	FRAS1-related extracellular matrix protein 3 Precursor [Source:UniProtKB/Swiss-Prot;Acc:P0C091]
SNP_124	1538-0124	Chr5:99860724	Chr5:106747054	5‟UTR	NP_001002884.1	ENSBTAG00000022861	killer cell lectin-like receptor family J member 1 [Source:RefSeq peptide;Acc:NP_001002884]
SNP_125	1538-0125	Chr1:6491257	Chr1:6409151	NON-SYNON	NP_001070335.1	ENSBTAG00000020122	ubiquitin specific protease 16 [Source:RefSeq peptide;Acc:NP_001070335]
SNP_126	1538-0126	Chr10:28787955	Chr10:28257660	NON-SYNON	AVEN	ENSBTAG00000044029	
SNP_128	1538-0128	Chr13:46685530	Chr13:46498304	NON-SYNON	IDI1_BOVIN	ENSBTAG00000004075	Isopentenyl-diphosphate Delta-isomerase 1 (EC 5.3.3.2)(Isopentenyl pyrophosphate isomerase 1)(IPP isomerase 1)(IPPI1) [Source:UniProtKB/Swiss-Prot;Acc:Q1LZ95]
SNP_130	1538-0130	Chr16:44300796	Chr16:40372961	NON-SYNON	NP_001095974.1	ENSBTAG00000010401	ubiquitination factor E4B [Source:RefSeq peptide;Acc:NP_001095974]
SNP_133	1538-0133	Chr2:89583887	Chr2:93460963	NON-SYNON	ADO_BOVIN	ENSBTAG00000009725	Aldehyde oxidase (EC 1.2.3.1) [Source:UniProtKB/Swiss-Prot;Acc:P48034]
SNP_136	1538-0136	Chr26:11124520	Chr26:11485676	NON-SYNON	NP_001069166.1	ENSBTAG00000017367	interferon-induced protein with tetratricopeptide repeats 5 [Source:RefSeq peptide;Acc:NP_001069166]
SNP_137	1538-0137	Chr26:42797281	Chr26:43046766	NON-SYNON	Q28908_BOVIN	ENSBTAG00000022715	Mucin Fragment

Name	Assay ID	UMD3.1 Position	Btau4 Position	Consequence of SNP	Gene Name	Ensembl Gene ID	Description of gene
							[Source:UniProtKB/TrEMBL;Acc:Q28908]
SNP_139	1538-0139	Chr4:70052100	Chr4:72322760	NON-SYNON	SNX10_BOVIN	ENSBTAG00000001822	Sorting nexin-10 [Source:UniProtKB/Swiss-Prot;Acc:Q0IIL5]
SNP_141	1538-0141	Chr1:58938059	Chr1:59327116	3″UTR	VATA_BOVIN	ENSBTAG00000002703	V-type proton ATPase catalytic subunit A (V-ATPase subunit A)(EC 3.6.3.14)(Vacuolar proton pump subunit alpha)(V-ATPase 69 kDa subunit) [Source:UniProtKB/Swiss-Prot;Acc:P31404]
SNP_142	1538-0142	Chr10:12412917	Chr10:12288215	3″UTR	PTAD1_BOVIN	ENSBTAG00000015155	Protein tyrosine phosphatase-like protein PTPLAD1 (Protein-tyrosine phosphatase-like A domain-containing protein 1) [Source:UniProtKB/Swiss-Prot;Acc:A7YY55]
SNP_144	1538-0143	Chr11:46349799	Chr11:48188084	3″UTR	IL1A_BOVIN	ENSBTAG00000010349	Interleukin-1 alpha Precursor (IL-1 alpha) [Source:UniProtKB/Swiss-Prot;Acc:P08831]
SNP_145	1538-0144	Chr13:43330508	Chr13:43253392	3″UTR	GDIB_BOVIN	ENSBTAG00000005316	Rab GDP dissociation inhibitor beta (Rab GDI beta)(Guanosine diphosphate dissociation inhibitor 2)(GDI-2) [Source:UniProtKB/Swiss-Prot;Acc:P50397]
SNP_146	1538-0145	Chr15:17460234	Chr15:15622104	3″UTR	ELMD1_BOVIN	ENSBTAG00000002691	ELMO domain-containing protein 1 [Source:UniProtKB/Swiss-Prot;Acc:Q0IIE6]
SNP_147	1538-0146	Chr15:3493351	Chr15:2115975	3″UTR	CASPD_BOVIN	ENSBTAG00000020884	Caspase-13 Precursor (CASP-13)(EC 3.4.22.-)(Evolutionary related interleukin-1-beta-converting enzyme)(ERICE) [Contains Caspase-13 subunit 1;Caspase-13 subunit 2] [Source:UniProtKB/Swiss-Prot;Acc:O75601]
SNP_149	1538-0148	Chr15:47331292	Chr15:45808076	3″UTR	PRDBP_BOVIN	ENSBTAG00000019754	Protein kinase C delta-binding protein [Source:UniProtKB/Swiss-Prot;Acc:A4FV37]
SNP_150	1538-0149	Chr19:5818851	Chr19:4770650	3″UTR	NP_001069069.1	ENSBTAG00000000205	monocyte to macrophage differentiation-associated [Source:RefSeq peptide;Acc:NP_001069069]
SNP_151	1538-0150	Chr2:125811079	Chr2:129403297	3″UTR	PTAFR_BOVIN	ENSBTAG00000027051	Platelet-activating factor receptor (PAF-R)(PAFr) [Source:UniProtKB/Swiss-Prot;Acc:Q9TTY5]
SNP_152	1538-0151	Chr29:5056001	Chr29:5264699	3″UTR	CHRD1_BOVIN	ENSBTAG00000013615	Cysteine and histidine-rich domain-containing protein 1 (CHORD domain-containing protein 1) [Source:UniProtKB/Swiss-Prot;Acc:Q29RL2]
SNP_153	1538-0152	Chr3:78644503	Chr3:84013404	3″UTR	S35D2_BOVIN	ENSBTAG00000005445	UDP-N-acetylglucosamine/UDP-glucose/GDP-mannose transporter (Solute carrier family 35 member D2) [Source:UniProtKB/Swiss-Prot;Acc:A2VE55]
SNP_156	1538-0155	Chr13:7542738	Chr13:7503555	SYNON	Q0IIB7_BOVIN	ENSBTAG00000004302	C13H20ORF7 protein Fragment

Name	Assay ID	UMD3.1 Position	Btau4 Position	Consequence of SNP	Gene Name	Ensembl Gene ID	Description of gene
SNP_157	1538-0156	Chr15:6203634	Chr15:4895096	SYNON	NP_001092477.1	ENSBTAG00000040567	[Source:UniProtKB/TrEMBL;Acc:Q0IIB7] matrix metalloproteinase 27 [Source:RefSeq peptide;Acc:NP_001092477]
SNP_160	1538-0159	Chr22:1807200	Chr22:1605654	SYNON	NP_001092396.1	ENSBTAG00000018227	"solute carrier family 4, sodium bicarbonate cotransporter, member 7 [Source:RefSeq peptide;Acc:NP_001092396]"
SNP_161	1538-0160	Chr22:1807251	Chr22:1605705	SYNON	NP_001092396.1	ENSBTAG00000018227	"solute carrier family 4, sodium bicarbonate cotransporter, member 7 [Source:RefSeq peptide;Acc:NP_001092396]"
SNP_162	1538-0161	Chr22:29391700	Chr22:29991401	SYNON	RYBP	ENSBTAG00000022689	RING1 and YY1-binding protein (Death effector domain-associated factor)(DED-associated factor)(YY1 and E4TF1-associated factor 1)(Apoptin-associating protein 1)(APAP-1) [Source:UniProtKB/Swiss-Prot;Acc:Q8N488]
SNP_163	1538-0162	Chr23:25183813	Chr23:25961459	SYNON	ELOV5_BOVIN	ENSBTAG00000003359	Elongation of very long chain fatty acids protein 5 [Source:UniProtKB/Swiss-Prot;Acc:Q2KJD9]
SNP_165	1538-0164	Chr26:32981542	Chr26:33115701	SYNON	PLSB_BOVIN	ENSBTAG00000011917	"Glycerol-3-phosphate acyltransferase, mitochondrial Precursor (GPAT)(EC 2.3.1.15) [Source:UniProtKB/Swiss-Prot;Acc:Q5GJ77]"
SNP_166	1538-0165	Chr26:42711970	Chr26:42951722	SYNON	O97658_BOVIN	ENSBTAG00000008389	Serine protease Fragment [Source:UniProtKB/TrEMBL;Acc:O97658]
SNP_168	1538-0167	Chr28:8444446	Chr28:6881215	SYNON	LYST_BOVIN	ENSBTAG00000016804	Lysosomal-trafficking regulator [Source:UniProtKB/Swiss-Prot;Acc:Q9TTK4]

Table S2: Complete list of BWA/GATK SNPs with consequences and annotated gene names

This table is too large to be reproduced here and can be accessed at

http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info:doi/10.1371/journal.pone.0048792.s002

Table S3: List of genes that had at least one synonymous or non-synonymous SNP, with dN/dS ratios and p values calculated with PAML

This table is too large to be reproduced here and can be accessed at

http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info:doi/10.1371/journal.pone.0048792.s003

Table S4: David annotation of genes with dN/dS ratios of 99, which are mainly those with no synonymous SNPs

Annotation Cluster 1	**Enrichment Score: 2.404780122112932**											
Category	**Term**	**Count**	**%**	**P**	**Genes**	**List Total**	**Pop Hits**	**Pop Total**	**Fold Enrichment**	**Bonf**	**Benj**	**FDR**
INTERPRO	IPR001304:C-type lectin	4	2.380952	0.003094	ENSBTAG00000002135, ENSBTAG00000020755, ENSBTAG00000039980, ENSBTAG00000009252	56	49	9171	13.3688	0.385	0.385	3.661
SMART	SM00034:CLECT	4	2.380952	0.004365	ENSBTAG00000002135, ENSBTAG00000020755, ENSBTAG00000039980, ENSBTAG00000009252	31	49	4364	11.49177	0.138	0.138	3.774
INTERPRO	IPR016186:C-type lectin-like	4	2.380952	0.00452	ENSBTAG00000002135, ENSBTAG00000020755, ENSBTAG00000039980, ENSBTAG00000009252	56	56	9171	11.6977	0.509	0.299	5.306

Table S5: David annotation of genes that had dN/dS ratios significantly less than 1

This table is too large to be reproduced here and can be accessed at

http://www.plosone.org/article/fetchSingleRepresentation.action?uri=info:doi/10.1371/journal.pone.0048792.s005

Chapter 5

Gene polymorphisms in African buffalo associated with susceptibility to bovine tuberculosis infection

Nikki le Roex[1*], Ad P Koets[2,3], Paul D van Helden[1], Eileen G Hoal[1]

1 DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/ MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, Stellenbosch University, Cape Town, South Africa. **2** Department of Farm Animal Health, **3** Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands.

PLoS ONE (2013); 8(5): e64494.

Abstract

Bovine tuberculosis (BTB) is a chronic, highly infectious disease that affects humans, cattle and numerous species of wildlife. In developing countries such as South Africa, the existence of extensive wildlife-human-livestock interfaces poses a significant risk of *Mycobacterium bovis* transmission between these groups, and has far-reaching ecological, economic and public health impacts. The African buffalo (*Syncerus caffer*), acts as a maintenance host for *Mycobacterium bovis*, and maintains and transmits the disease within the buffalo and to other species. In this study we aimed to investigate genetic susceptibility of buffalo for *Mycobacterium bovis* infection. Samples from 868 African buffalo of the Cape buffalo subspecies were used in this study. SNPs (n=69), with predicted functional consequences in genes related to the immune system, were genotyped in this buffalo population by competitive allele-specific SNP genotyping. Case-control association testing and statistical analyses identified three SNPs associated with BTB status in buffalo. These SNPs, SNP41, SNP137 and SNP144, are located in the SLC7A13, DMBT1 and IL1α genes, respectively. SNP137 remained significantly associated after permutation testing. The three genetic polymorphisms identified are located in promising candidate genes for further exploration into genetic susceptibility to BTB in buffalo and other bovids, such as the domestic cow. These polymorphisms/genes may also hold potential for marker-assisted breeding programmes, with the aim of breeding more BTB-resistant animals and herds within both the national parks and the private sector.

Keywords: *bovine, TB, buffalo, genetic, association, genotyping, mycobacteria, susceptibility*

1. Introduction

Bovine tuberculosis (BTB) is a chronic, highly infectious disease which affects humans, cattle and wildlife. BTB is caused by infection with *Mycobacterium bovis*, which is a member of the *Mycobacterium tuberculosis* complex. In developing countries such as South Africa, the existence of extensive wildlife-human-livestock interfaces poses a significant risk of transmission between these groups, and has far-reaching ecological, economic and public health impacts [1–3]. In addition to the impact on biodiversity and ecotourism, rural communities at these wildlife interfaces can also be severely affected, with bovine TB compromising their health, food supply and livelihoods. Many people are reliant on

unpasteurised milk and meat from their livestock for food supply and high levels of HIV prevalence within these communities dramatically increases susceptibility to BTB infection [1,3].

In the absence of a wildlife reservoir, most developed countries have effectively used test and cull schemes to reduce BTB prevalence in cattle to extremely low levels. This system is seldom utilised, or effective, in developing countries with a wildlife reservoir that maintains BTB within the environment [4]. In South Africa, the African buffalo (*Syncerus caffer*), acts as a maintenance host/reservoir for *M. bovis*, and thus maintains and transmits the disease within the buffalo populations and to other species [1,5]. It is a species of enormous economic and ecological importance, due to its occurrence in large numbers in the savannah ecosystem, commercial game farming, tourism, and for hunting purposes [1]. This poses a significant obstacle to the management and control of BTB. Different BTB management strategies with regards to buffalo are implemented in the different game parks and private reserves within South Africa, and new methods of detection and control are constantly under investigation.

Studies in humans have reported multiple genetic factors and polymorphisms in a number of different genes to be associated with tuberculosis susceptibility [6]. In animal studies, the heritability of BTB resistance has been shown to be 0.48 ± 0.09 in red deer [7], 0.18 ± 0.04 in British dairy cattle [8] and 0.18 ± 0.04 in Irish cattle [9], suggesting a genetic component in BTB susceptibility in these animals, which may be more pronounced in wildlife. A recent study by Sun et al. [10] investigated the role of host genetic factors in susceptibility to BTB in Chinese Holstein cattle, and found the G1596A polymorphism in the toll-like receptor 1 (TLR1) gene to be associated with BTB infection status. Finlay et al. (2012) performed a genome-wide association scan in Irish cattle and identified a genomic region on BTA 22 associated with BTB susceptibility [11]. The identification of host genetic factors involved in BTB susceptibility in African buffalo may allow for marker-assisted selection for breeding programs, as a possible long-term disease management alternative, particularly in areas currently free of BTB or with low prevalence levels.

Mycobacterium bovis is primarily transmitted between animals by inhalation, and to a lesser degree, ingestion [3,12]. After inhalation, bacilli adhere to the alveolar surface of the lung and are phagocytosed by macrophages. Macrophages then process the mycobacterial antigens

and present them to T-lymphocytes, which are considered essential recognition components of the immune response. Typical tuberculous lesions in buffalo occur most often in the lungs and lymph nodes, but can also be found in other, more distal sites [5]. Resistance to infection and disease could be mediated by either the innate or adaptive immune systems and therefore the genes related to either of these processes may be candidate genes for BTB susceptibility loci.

In this study, candidate polymorphisms in genes involved in the immune response were selected from the SNPs that we recently identified and validated in the African buffalo [13]. These SNPs were tested for genetic associations with BTB infection status in African buffalo, using a case-control approach.

2. Materials & Methods

2.1 Ethics

The Stellenbosch University Animal Care and Use Committee (SU ACU) deemed it unnecessary to obtain ethical clearance for this study as the blood samples used for DNA extraction were collected under the directive of SANParks and Ezemvelo KZN Wildlife for other purposes, and their use in the present study is incidental.

2.2 Study Population

The DNA used in the association study was derived from 868 African buffalo (198 cases, 670 controls) of the Cape buffalo subspecies, consisting of two independent populations. The population subgroups consisted of 434 buffalo from the Kruger National Park (KNP), and 434 buffalo from Hluhluwe iMfolozi Park (HiP) (Figure 1). Both KNP and HiP have known histories of BTB in both buffalo and other wildlife species, and given the social herd structure of this species, both cases and controls were considered equally exposed. Blood samples were taken from the buffalo subpopulations during routine BTB testing operations, by qualified SANParks or Ezemvelo KZN Wildlife veterinary staff. In the KNP samples, cases and controls were defined by the positive or negative outcome of the Bovigam® ELISA assay. HiP samples were defined as cases or controls based on the results of the standard single comparative intradermal tuberculin (SCIT) test, administered following capture, with a test result considered positive when the difference in skin fold swelling between the avian and bovine inflammatory responses was greater than or equal to 2mm. Recent publications have

suggested that that a cut-off of >2mm for positive test interpretation in cattle is most appropriate in sub-Saharan Africa [14,15]. Overall BTB prevalence in our samples in the KNP and HiP subpopulations was 22% and 23.5% respectively (own unpublished data).

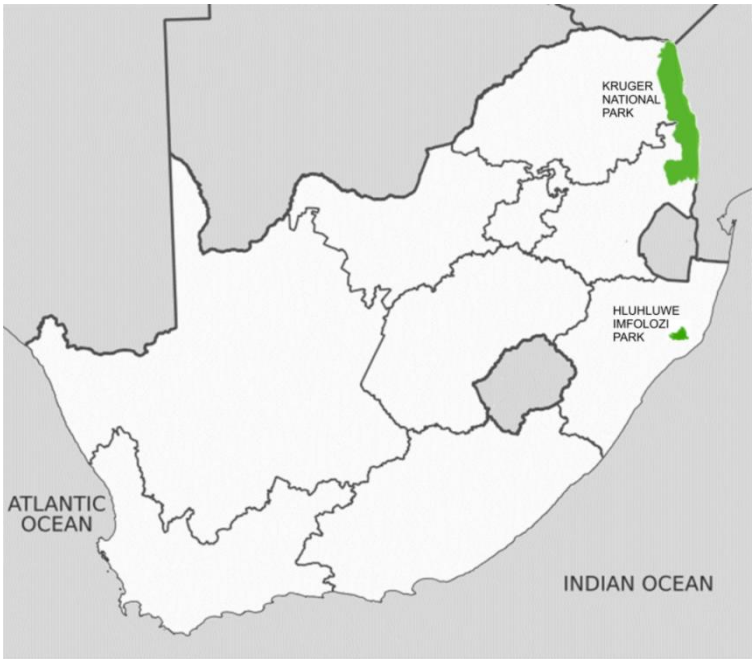

Figure 1: Location of the African buffalo populations in South Africa used in this study.

2.3 Preparation of genomic DNA

The Kruger National Park DNA samples were extracted from blood, using a salt-chloroform extraction method [16]. KNP Samples which contained insufficient concentrations of DNA for the purposes of this study were subjected to whole genome amplification (WGA) using the Illustra Genomiphi DNA Amplification Kit (GE Healthcare Life Sciences, Piscataway, USA), which employs a strand displacement amplification technique, according to the manufacturer‟s instructions. DNA was extracted from the HiP EDTA blood samples using the Qiagen DNeasy Blood and Tissue Kit (Qiagen Group®), according to the manufacturer‟s instructions.

2.4 SNP selection and genotyping

SNP selection was made from previously validated SNPs [13], with priority given to those with higher frequencies in the validation panel. These SNPs all had potentially functional

consequences, due to their positions in the 3‟UTR, 5‟UTR and coding regions, in genes related to the immune system. In total, 69 SNPs were fluorescently genotyped in the full buffalo population by LGC Genomics KBioscience [17] (London, United Kingdom), using their competitive allele-specific KASP® SNP genotyping platform. The KASP® assay system relies on the discrimination power of a FRET-based homogenous form of competitive allele-specific PCR to determine the alleles at a specific locus within genomic DNA [17]. A block randomisation technique was employed to evenly distribute the samples across the 10 plates used for the genotyping in order to minimise batch effects. Variables considered included BTB infection status, location and sex. Positive controls were included on all plates to ensure consistent results. The selected SNPs and their assay IDs can be seen in Table S1.

2.5 Statistical analysis

Genetic differentiation between the KNP and HiP subpopulations was tested using FST values in Genepop v4.0.10 [18,19]. Further statistical analyses were performed in Plink v1.07 (http://pngu.mgh.harvard.edu/purcell/plink/) [20]. SNPs in this study were not removed if they were out of Hardy-Weinberg Equilibrium (HWE), as due to the social structure of a buffalo herd, non-random mating is observed and thus the expectations of HWE do not hold for this data. African buffalo display a male dominance hierarchy, with a select number of dominant bulls mating with the sexually mature females, who are often related. This suggests a degree of relatedness within a herd [21]. The populations were analysed together in order to minimise the effect of any possible relatedness present in the individual herds. Quality control consisted of removing individuals with more than 10% missing data (genotyping failure) and SNPs that had a call rate of less than 90%. Pairwise linkage disequilibrium measures were calculated for all SNP pairs using correlations based on genotype allele counts, including those on different chromosomes. Allele frequencies, odds ratios (OR) and 95% confidence intervals were determined and the associated p-values were calculated using Fisher‟s Exact test. A p-value <0.05 was regarded as significant. The Cochran-Mantel-Haenszel (CMH) test was used to obtain an average odds ratio and associated p-value for each SNP, adjusting for the two subpopulations. Genotypic associations with BTB infection were tested using the Cochran-Armitage trend test for the additive allele effect. The Cochran-Armitage trend test was considered the most applicable genotypic test for this study, as it does not assume HWE. Logistic regression analysis was also used to calculate odds ratios and 95% confidence intervals, assuming an additive effect of allele dosage. We controlled for subpopulation and sex by including those covariates in the logistic regression model.

Permutation testing was performed using the max(T) permutation method with 5000 permutations, to obtain an empirical p-value for each SNP. SNPs found to be significantly associated with BTB status in the buffalo after correcting for the population structure were investigated further. Non-synonymous SNPs were submitted to the Ensembl Variant Effect Predictor [22] in order to establish the type and position of the amino acid changes caused.

3. Results

The SNPs in this study comprised 69 loci across 64 genes, previously selected for possible functional consequence and gene location. The F_{ST} value for the genetic differentiation between the two subpopulations was 0.1345. Quality control removed 24 individuals due to missing data. No SNPs were excluded due to call rate. Because of the non-random mating structure within the buffalo herds, HWE was not tested. Two loci, SNP105 and SNP160, were identified as being in linkage disequilibrium with other SNPs ($r^2 > 0.5$) and were excluded. Fisher‟s Exact test for allelic association identified four SNPs with significantly different frequencies between cases and controls ($p < 0.05$): SNP30, SNP41, SNP137 and SNP144. After adjusting for differences due to the subpopulations using the CMH test, three of the four previously identified SNPs remained significantly associated. Following correction for false positives using the max(T) permutation method, SNP137 was identified as significantly associated with BTB status. Associated SNPs in the allelic association and CMH tests, and the accompanying odds ratios and p values, can be seen in Table 1. SNP30 and SNP144 are located in the 3‟UTRs of the complement component C7 gene and the interleukin 1-alpha (IL1α) gene, respectively. SNP41 and SNP137 are located in the coding regions of the solute carrier family 7 (anionic amino acid transporter), member 13 (SLC7A13) gene and Deleted in Malignant Brain Tumour (DMBT1) genes, respectively. All SNPs, allele frequencies, odds ratios and p values for the CMH test can be seen in Table S2.

The Cochran-Armitage trend test for genotypic association identified the same four SNPs (SNP30, SNP41, SNP137 and SNP144) as significant. No significant associations remained after correction using max(T) permutation.

Table 1: Allele frequencies, odds ratios and p-values of four significant SNPs

SNP	Allele	Fcase	Fcont	OR (95% CI)	P	CMH OR (95% CI)	CMH P	P_{emp}
30	G	0.225	0.276	0.761 (0.581-0.995)	0.048	0.765 (0.584-1.002)	0.0514	0.973
41	C	0.347	0.285	1.335 (1.047-1.702)	0.022	1.377 (1.075-1.764)	0.0116	0.559
137	G	0.423	0.332	1.472 (1.163-1.863)	0.001	1.608 (1.243-2.079)	0.0003	0.024
144	T	0.353	0.294	1.312 (1.03-1.67)	0.032	1.421 (1.095-1.843)	0.008	0.431

P_{emp}: Empirical p-value obtained by max(T) permutation

Logistic regression analysis incorporating both sex and subpopulation as covariates identified SNP41, SNP137 and SNP144 as significantly associated with BTB status. Logistic regression analysis incorporating only subpopulation as a covariate produced the same result, and sex was not identified as a significant factor in this study (data not shown). After correction using max(T) permutation, SNP137 remained significantly associated. We note that the SNP137 was also significantly associated in the allelic, genotypic and logistic regression tests in both the HiP and KNP populations when analysed separately (data not shown). The significant logistic regression model results, odds ratios and p-values can be seen in Table 2, and the logistic regression results for all SNPs in Table S3.

Table 2: Logistic regression model with significant SNPs, odds ratios and p-values

SNP	Allele	Test	OR (95% CI)	P	P_{perm}
41	C	ADD	1.349 (1.06-1.715)	0.01474*	0.6265
41	C	COV1 [a]	1.222 (0.8746-1.708)	0.2397	
137	G	ADD	1.574 (1.221-2.029)	0.000461*	0.02599*
137	G	COV1	1.47 (1-2.16)	0.04982*	
144	T	ADD	1.435 (1.101-1.87)	0.007569*	0.3893
144	T	COV1	1.396 (0.9605-2.029)	0.08032	

[a]COV1 refers to the population subgroup
* denotes statistical significance, $p<0.05$

The amino acid changes caused by SNP41 and SNP137, as well as their positions in their respective proteins, were identified by the Ensemble Variant Effect Predictor (Table 3). Although both the DMBT1 and SLC7A13 proteins are uncharacterised in the cow, SNP137 was predicted to occur in a CUB domain within the DMBT1 protein. CUB domains are

evolutionarily conserved protein domains found commonly in extracellular and plasma-membrane associated proteins [23].

Table 3: Ensembl Variant Effect Predictor predictions of amino acid changes and positions

SNP	Location	Allele	Ensembl Gene ID	Consequence	cDNA position	Position in CDS[a]	Protein position	AA change
41	14:79862771	C	ENSBTAG00000040461	missense	1444	1330	444	I/V
137	26:42797281	G	ENSBTAG00000022715	missense	3350	3350	1117	H/R

[a]CDS refers to the coding sequence

4. Discussion

To our knowledge, this is the first report of genetic susceptibility to bovine TB in the African buffalo. In other bovids, specifically the domestic cow, very few loci have been identified, apart from microsatellite loci [24], TLR1 [10] and a genomic region on BTA 22 [11]. In this study, we investigated 69 SNPs located in genes relating to the immune system for association with BTB infection status in African buffalo, using a large sample set of 868 buffalo samples from two independent populations, the Kruger National Park and Hluhluwe iMfolozi Park. Prior to correction for the population subgroups, four SNPs were identified as associated with BTB status – SNP30, SNP41, SNP137 and SNP144. The F_{ST} value of 0.13 suggests that there is a degree of differentiation between the buffalo populations, which necessitates controlling for. After correction for population subgroups using the Cochran-Mantel-Haenszel test and logistic regression analysis, three SNPs remained significantly associated: SNP41, SNP137 and SNP144.

SNP41 occurs in the SLC7A13 gene, and is predicted to cause a non-synonymous amino acid substitution at position 444 in this protein, from isoleucine to valine. The SLC7A13 gene forms part of the family of heteromeric amino acid transporters (HATs). HATs are comprised of a heavy and a light chain/subunit, and the SLC7A13 gene encodes one of the light subunits [25,26]. The light subunits are highly hydrophobic and are suggested to have a 12 transmembrane-domain topology [25]. While SLC7A13 is relatively uncharacterised, other transporters, such as SLC11A1 (NRAMP1), have been associated with TB infection in humans [27,28].

SNP137 is located in the coding region of the Deleted in Malignant Brain Tumour-1 (DMBT1) gene. This SNP is predicted to cause a substitution at position 1117 of the protein, from histidine to arginine. The DMBT1 gene produces a number of different protein isoforms, all of which belong to the scavenger receptor cysteine-rich (SRCR) superfamily of proteins. SRCR proteins are highly conserved throughout the metazoa, and contain two CUB domains and one zona pellucida domain [29]. The glycoprotein isoforms have been shown to have a range of functions that include the binding of both gram-positive and gram-negative bacteria, binding to viral proteins, and binding to host surfactant proteins A and D, among others, which are collectins that play an important role in the innate immune response [30,31]. Both DMBT1 and the mouse homolog of this protein (CRP-ductin) are regarded as pattern recognition receptors (PRRs), which are a well-established part of the human host defence system against invading organisms such as *M. tuberculosis*. PRRs recognize conserved structures in bacterial cell walls, yeast cell walls, and bacterial DNA that are vital to the survival of the invading organism [29,32]. Many members of the SRCR superfamily have been shown to be located on immune cells such as B-lymphocytes, T-lymphocytes and macrophages, and the role of the DMBT1 proteins in the innate immune response has been well established [30]. While evidence is lacking for the bovine protein, the conserved nature of these proteins suggests a similar function in the domestic cow.

SNP144 is found in the 3‟ untranslated region (UTR) of the Interleukin 1 alpha (IL1α) gene. The IL1α protein belongs to the IL1 family of pro-inflammatory cytokines, and plays a role in the activation of T-lymphocytes, the recruitment of leucocytes to sites of inflammation, and mediating local and systemic inflammation, and is expressed predominantly by macrophages and neutrophils [33,34]. A study by Bellamy et al. [35] found a significant association between the IL1 gene cluster and *M. tuberculosis* in humans, although this was lost for IL1α after Bonferroni correction.

While the results of this study are promising, there were some limitations, most notably the difficulties of working with a non-model organism without an assembled genome or well-established population data. Because our buffalo herd data falls somewhere between that of population data and family data, and the degree of relatedness among the cases and controls prevented the investigation of SNPxSNP interactions. The second limitation faced was the different diagnostic methods used for the two population groups. The Bovigam® ELISA assay was used to diagnose BTB infection in the Kruger National Park buffalo, and the SCIT test was used in Hluhluwe iMfolozi Park, but since both of these tests are based on a cell-

mediated immune response and therefore expected to detect animals at a similar stage of infection [36,37], this was not considered problematic in this study.

In conclusion, our study has identified three genetic polymorphisms, in the SLC7A13, IL1α and DMBT1 genes, that are associated with bovine TB infection status in the African buffalo. The association of the three polymorphisms listed above were robust to correction for buffalo population, and one remained significant after permutation testing. The large number of individuals used in this study, the randomisation procedure and statistical controls employed, and the plausibility of the genes identified suggest that these polymorphisms are located in excellent candidate genes for further exploration into genetic susceptibility to BTB in both the buffalo and other bovids, such as the domestic cow. These genes may also have potential for marker-assisted breeding programmes, with the aim of breeding more BTB resistant animals and herds within both the national parks and the private sector.

Acknowledgements

We would like to gratefully acknowledge Dave Cooper and the staff of Ezemvelo KZN Wildlife, as well as Roy Bengis, state veterinarians and SANParks Veterinary Wildlife Services (VWS), for the collection of buffalo blood samples. Without those samples, the study would not have been possible.

Supporting Information

Table S1: SNP identifiers, Assay IDs, UMD3 positions, Ensembl gene ID and names, consequences **(**TableS1.txt)

Table S2: All allele frequencies, odds ratios and p-values for Cochran-Mantel-Haenszel test

(TableS2.txt)

Table S3: Logistic regression odds ratios and p-values for all SNPs

(TableS3.txt)

References

1. Michel AL, Coetzee ML, Keet DF, Maré L, Warren R, et al. (2009) Molecular epidemiology of Mycobacterium bovis isolates from free-ranging wildlife in South African game reserves. Veterinary Microbiology 133: 335–343. doi:16/j.vetmic.2008.07.023.

2. Michel AL, Bengis RG, Keet DF, Hofmeyr M, Klerk LM, et al. (2006) Wildlife tuberculosis in South African conservation areas: implications and challenges. VetMicrobiol 112: 91–100.

3. Ayele WY, Neill SD, Zinsstag J, Weiss MG, Pavlik I (2004) Bovine tuberculosis: an old disease but a new threat to Africa. IntJTubercLung Dis 8: 924–937.

4. Caron A, Cross PC, Du Toit JT (2003) Ecological implications of bovine tuberculosis in African buffalo herds. Ecological Applications 13: 1338–1345.

5. Renwick AR, White PCL, Bengis RG (2007) Bovine tuberculosis in southern African wildlife: a multi-species host-pathogen system. Epidemiol Infect 135: 529–540. doi:10.1017/S0950268806007205.

6. Möller M, Hoal EG (2010) Current findings, challenges and novel approaches in human genetic susceptibility to tuberculosis. Tuberculosis (Edinb) 90: 71–83. doi:10.1016/j.tube.2010.02.002.

7. Mackintosh CG, Qureshi T, Waldrup K, Labes RE, Dodds KG, et al. (2000) Genetic resistance to experimental infection with Mycobacterium bovis in red deer (Cervus elaphus). InfectImmun 68: 1620–1625.

8. Brotherstone S, White IMS, Coffey M, Downs SH, Mitchell AP, et al. (2010) Evidence of genetic resistance of cattle to infection with Mycobacterium bovis. J Dairy Sci 93: 1234–1242. doi:10.3168/jds.2009-2609.

9. Bermingham ML, More SJ, Good M, Cromie AR, Higgins IM, et al. (2009) Genetics of tuberculosis in Irish Holstein-Friesian dairy herds. JDairy Sci 92: 3447–3456.

10. Sun L, Song Y, Riaz H, Yang H, Hua G, et al. (2012) Polymorphisms in toll-like receptor 1 and 9 genes and their association with tuberculosis susceptibility in Chinese Holstein cattle. Vet Immunol Immunopathol 147: 195–201. doi:10.1016/j.vetimm.2012.04.016.

11. Finlay EK, Berry DP, Wickham B, Gormley EP, Bradley DG (2012) A genome wide association scan of bovine tuberculosis susceptibility in Holstein-Friesian dairy cattle. PLoS ONE 7: e30545. doi:10.1371/journal.pone.0030545.

12. Hope JC, Villarreal-Ramos B (2008) Bovine TB and the development of new vaccines. Comp ImmunolMicrobiolInfectDis 31: 77–100.

13. Le Roex N, Noyes H, Brass A, Bradley DG, Kemp SJ, et al. (2012) Novel SNP discovery in African buffalo, Syncerus caffer, using high-throughput sequencing. PLoS ONE 7: e48792. doi:10.1371/journal.pone.0048792.

14. Ameni G, Hewinson G, Aseffa A, Young D, Vordermeier M (2008) Appraisal of Interpretation Criteria for the Comparative Intradermal Tuberculin Test for Diagnosis of Tuberculosis in Cattle in Central Ethiopia. Clinical and Vaccine Immunology 15: 1272–1276. doi:10.1128/CVI.00114-08.

15. Ngandolo BN, Muller B, guimbaye-Djaibe C, Schiller I, Marg-Haufe B, et al. (2009) Comparative assessment of fluorescence polarization and tuberculin skin testing for the diagnosis of bovine tuberculosis in Chadian cattle. PrevVetMed 89: 81–89.

16. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16: 1215.

17. LGC Genomics KBioscience (n.d.). Available: http://www.lgcgenomics.com/.

18. Raymond M, Rousset F (1995) GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. J Hered 86: 248–249.

19. Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8: 103–106. doi:10.1111/j.1471-8286.2007.01931.x.

20. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. American Journal of Human Genetics 81: 559.

21. Simonsen BT, Siegismund HR, Arctander P (1998) Population structure of African buffalo inferred from mtDNA sequences and microsatellite loci: high variation but low differentiation. MolEcol 7: 225–237.

22. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, et al. (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. Bioinformatics 26: 2069–2070. doi:10.1093/bioinformatics/btq330.

23. Blanc G, Font B, Eichenberger D, Moreau C, Ricard-Blum S, et al. (2007) Insights into how CUB domains can exert specific functions while sharing a common fold: conserved and specific features of the CUB1 domain contribute to the molecular basis of procollagen C-proteinase enhancer-1 activity. J Biol Chem 282: 16924–16933. doi:10.1074/jbc.M701610200.

24. Driscoll EE, Hoffman JI, Green LE, Medley GF, Amos W (2011) A preliminary study of genetic factors that influence susceptibility to bovine tuberculosis in the British cattle herd. PLoS ONE 6: e18806. doi:10.1371/journal.pone.0018806.

25. Palacín M, Nunes V, Font-Llitjós M, Jiménez-Vidal M, Fort J, et al. (2005) The genetics of heteromeric amino acid transporters. Physiology (Bethesda) 20: 112–124. doi:10.1152/physiol.00051.2004.

26. Verrey F, Closs EI, Wagner CA, Palacin M, Endou H, et al. (2004) CATs and HATs: the SLC7 family of amino acid transporters. Pflugers Arch 447: 532–542. doi:10.1007/s00424-003-1086-z.

27. Velez DR, Hulme WF, Myers JL, Stryjewski ME, Abbate E, et al. (2009) Association of SLC11A1 with tuberculosis and interactions with NOS2A and TLR2 in African-Americans and Caucasians. Int J Tuberc Lung Dis 13: 1068–1076.

28. Hoal EG, Lewis LA, Jamieson SE, Tanzer F, Rossouw M, et al. (2004) SLC11A1 (NRAMP1) but not SLC11A2 (NRAMP2) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. Int J Tuberc Lung Dis 8: 1464–1471.

29. Ligtenberg AJM, Karlsson NG, Veerman ECI (2010) Deleted in Malignant Brain Tumors-1 Protein (DMBT1): A Pattern Recognition Receptor with Multiple Binding Sites. Int J Mol Sci 11: 5212–5233. doi:10.3390/ijms1112521.

30. Madsen J, Mollenhauer J, Holmskov U (2010) Review: Gp-340/DMBT1 in mucosal innate immunity. Innate Immun 16: 160–167. doi:10.1177/1753425910368447.

31. Palaniyar N (2010) Antibody equivalent molecules of the innate immune system: parallels between innate and adaptive immune proteins. Innate Immun 16: 131–137. doi:10.1177/1753425910370498.

32. Kleinnijenhuis J, Oosting M, Joosten LAB, Netea MG, Van Crevel R (2011) Innate immune recognition of Mycobacterium tuberculosis. Clin Dev Immunol 2011: 405310. doi:10.1155/2011/405310.

33. Schmidt DR, Kao WJ (2007) The interrelated role of fibronectin and interleukin-1 in biomaterial-modulated macrophage function. Biomaterials 28: 371–382. doi:10.1016/j.biomaterials.2006.08.041.

34. Beltan E, Horgen L, Rastogi N (2000) Secretion of cytokines by human macrophages upon infection by pathogenic and non-pathogenic mycobacteria. Microb Pathog 28: 313–318. doi:10.1006/mpat.1999.0345.

35. Bellamy R, Ruwende C, Corrah T, McAdam KPWJ, Whittle HC, et al. (1998) Assessment of the interleukin 1 gene cluster and other candidate gene polymorphisms in host susceptibility to tuberculosis. Tubercle and Lung Disease 79: 83–89. doi:10.1054/tuld.1998.0009.

36. Vordermeier M, Goodchild A, Clifton-Hadley R, de la Rua R (2004) The interferon-gamma field trial: background, principles and progress. Vet Rec 155: 37–38.

37. de la Rua-Domenech, Goodchild AT, Vordermeier HM, Hewinson RG, Christiansen KH, et al. (2006) Ante mortem diagnosis of tuberculosis in cattle: a review of the tuberculin tests, gamma-interferon assay and other ancillary diagnostic techniques. ResVetSci 81: 190–210.

Supplementary information

Table S1: SNP identifiers, Assay IDs, UMD3 positions, Ensembl gene ID and names, consequences

Name	KBioscience Assay	UMD3 Chr	UMD3 position	EnsGeneId	Gene	Consequence
SNP_1	1538-0001	Chr11	7179447	ENSBTAG00000001034	IL-18Ra	Non-Synon
SNP_4	1538-0004	Chr17	63612902	ENSBTAG00000014628	Oas2	Non-Synon
SNP_6	1538-0006	Chr1	111031389	ENSBTAG00000009012	PTX3	Non-Synon
SNP_7	1538-0007	Chr1	143152388	ENSBTAG00000008471	MX2	Non-Synon
SNP_8	1538-0008	Chr22	53599791	ENSBTAG00000008099	CCR5	Non-Synon
SNP_13	1538-0013	Chr3	78427756	ENSBTAG00000014983	IL-23R	Non-Synon
SNP_16	1538-0016	Chr5	36884961	ENSBTAG00000021105	IRAK4	Non-Synon
SNP_17	1538-0017	Chr5	36890957	ENSBTAG00000021105	IRAK4	Non-Synon
SNP_18	1538-0018	Chr5	47797051	ENSBTAG00000007636	IrakM	Non-Synon
SNP_19	1538-0019	Chr5	47817597	ENSBTAG00000007636	IrakM	Non-Synon
SNP_25	1538-0025	Chr7	98739431	ENSBTAG00000039275	ERAP2	Non-Synon
SNP_30	1538-0030	Chr20	33550138	ENSBTAG00000011766	C7	3' UTR
SNP_41	1538-0041	Chr14	79862771	ENSBTAG00000040461	SLC7A13	Non-Synon
SNP_56	1538-0056	Chr22	55714039	ENSBTAG00000005857	SLC6A1	3' UTR
SNP_57	1538-0057	Chr3	103930723	ENSBTAG00000009617	SLC2A1	3' UTR
SNP_59	1538-0059	Chr5	45246904	ENSBTAG00000031919	SLC35E3	3' UTR
SNP_62	1538-0062	Chr7	23330398	ENSBTAG00000000668	SLC22A5	3' UTR
SNP_63	1538-0063	Chr9	63303181	ENSBTAG00000014057	SLC35A1	3' UTR
SNP_64	1538-0064	Chr11	96772566	ENSBTAG00000013314	PBX3	3'UTR
SNP_65	1538-0065	Chr11	96772626	ENSBTAG00000013314	PBX3	3'UTR
SNP_67	1538-0067	Chr13	59903453	ENSBTAG00000002065	TFAP2C	3'UTR
SNP_75	1538-0075	Chr20	8038747	ENSBTAG00000010012	Btf3	3'UTR
SNP_76	1538-0076	Chr23	27968593	ENSBTAG00000010698	GTF2H4	5'UTR
SNP_80	1538-0080	Chr25	33291910	ENSBTAG00000009780	GTF2I	3'UTR
SNP_82	1538-0082	Chr25	33297474	ENSBTAG00000009780	GTF2I	Synon
SNP_95	1538-0095	Chr5	28826280	ENSBTAG00000019312	TFCP2	3'UTR
SNP_99	1538-0099	Chr5	29279068	ENSBTAG00000018131	ATF1	Non-Synon
SNP_100	1538-0100	Chr5	95801232	ENSBTAG00000003221	ATF7IP	Synon
SNP_103	1538-0103	Chr7	47384295	ENSBTAG00000001002	TCF7	3'UTR
SNP_104	1538-0104	Chr9	29122178	ENSBTAG00000012946	HSF2	3'UTR
SNP_105	1538-0105	Chr3	83571563	ENSBTAG00000019877	DOCK7	Non-Synon
SNP_109	1538-0109	Chr6	92798767	ENSBTAG00000010956	SCARB2	3'UTR
SNP_113	1538-0113	Chr6	38121032	ENSBTAG00000005260	SPP1	Synon
SNP_114	1538-0114	Chr5	102960154	ENSBTAG00000034678	Unchar	Non-Synon
SNP_118	1538-0118	Chr9	99649203	ENSBTAG00000011106	PACRG	5'UTR

Name	KBioscience Assay	UMD3 Chr	UMD3 position	EnsGeneId	Gene	Consequence
SNP_119	1538-0119	Chr3	19695912	ENSBTAG00000020394	TNFAIP8L2	3'UTR
SNP_120	1538-0120	Chr13	60253499	ENSBTAG00000006533	NSFL1C	Splice_Site
SNP_121	1538-0121	Chr3	27020461	ENSBTAG00000001246	ATP1A1	Splice_Site
SNP_122	1538-0122	Chr29	22710735	ENSBTAG00000038347	SLC17A6	Splice_Site
SNP_123	1538-0123	Chr17	14434651	ENSBTAG00000044051	FREM3	Stop_Gained
SNP_124	1538-0124	Chr5	99860724	ENSBTAG00000022861	KLRJ1	5'UTR
SNP_126	1538-0126	Chr10	28787955	ENSBTAG00000044029	AVEN	Non-Synon
SNP_128	1538-0128	Chr13	46685530	ENSBTAG00000004075	IDI1	Non-Synon
SNP_130	1538-0130	Chr16	44300796	ENSBTAG00000010401	UBE4B	Non-Synon
SNP_136	1538-0136	Chr26	11124520	ENSBTAG00000017367	IFIT5	Non-Synon
SNP_137	1538-0137	Chr26	42797281	ENSBTAG00000022715	DMBT1	Non-Synon
SNP_139	1538-0139	Chr4	70052100	ENSBTAG00000001822	SNX10	Non-Synon
SNP_141	1538-0141	Chr1	58938059	ENSBTAG00000002703	ATP6V1A	3'UTR
SNP_142	1538-0142	Chr10	12412917	ENSBTAG00000015155	PTPLAD1	3'UTR
SNP_144	1538-0143	Chr11	46349799	ENSBTAG00000010349	IL1A	3'UTR
SNP_145	1538-0144	Chr13	43330508	ENSBTAG00000005316	GDI2	3'UTR
SNP_146	1538-0145	Chr15	17460234	ENSBTAG00000002691	ELMOD1	3'UTR
SNP_147	1538-0146	Chr15	3493351	ENSBTAG00000020884	CASP13	3'UTR
SNP_149	1538-0148	Chr15	47331292	ENSBTAG00000019754	PRKCDBP	3'UTR
SNP_150	1538-0149	Chr19	5818851	ENSBTAG00000000205	MMD	3'UTR
SNP_151	1538-0150	Chr2	125811079	ENSBTAG00000027051	PTAFR	3'UTR
SNP_152	1538-0151	Chr29	5056001	ENSBTAG00000013615	CHORDC1	3'UTR
SNP_153	1538-0152	Chr3	78644503	ENSBTAG00000005445	SLC35D1	3'UTR
SNP_156	1538-0155	Chr13	7542738	ENSBTAG00000004302	C13H20orf7	Synon
SNP_157	1538-0156	Chr15	6203634	ENSBTAG00000040567	MMP27	Synon
SNP_160	1538-0159	Chr22	1807200	ENSBTAG00000018227	SLC4A7	Synon
SNP_161	1538-0160	Chr22	1807251	ENSBTAG00000018227	SLC4A7	Synon
SNP_162	1538-0161	Chr22	29391700	ENSBTAG00000022689	RYBP	Synon
SNP_163	1538-0162	Chr23	25183813	ENSBTAG00000003359	ELOVL5	Synon
SNP_165	1538-0164	Chr26	32981542	ENSBTAG00000011917	GPAM	Synon
SNP_166	1538-0165	Chr26	42711970	ENSBTAG00000008389	HTRA1	Synon
SNP_168	1538-0167	Chr28	8444446	ENSBTAG00000016804	LYST	Synon
SNP_169	1538-0168	Chr3	48853655	ENSBTAG00000020344	SLC44A3	Synon
SNP_171	1538-0170	Chr5	80508595	ENSBTAG00000005365	OVCH1	Synon

Table S2: All allele frequencies, odds ratios and p-values for Cochran-Mantel-Haenszel test

SNP	A1	MAF	P	OR	L95	U95
SNP1	T	0.2257	0.994	0.999	0.7589	1.315
SNP4	A	0.1924	0.8609	1.029	0.7501	1.411
SNP6	A	0.3726	0.8837	1.021	0.7782	1.338
SNP7	C	0.4656	0.7015	1.046	0.8309	1.317
SNP8	A	0.4618	0.923	1.011	0.8027	1.275
SNP13	C	0.3865	0.5633	1.074	0.8434	1.367
SNP16	T	0.2125	0.8461	1.028	0.7773	1.359
SNP17	G	0.3683	0.5312	0.9238	0.7188	1.187
SNP18	A	0.3582	0.723	1.046	0.8105	1.351
SNP19	C	0.2088	0.5743	1.088	0.8118	1.458
SNP25	G	0.3714	0.6563	1.058	0.825	1.357
SNP30	G	0.2648	0.05136	0.7651	0.5843	1.002
SNP41	C	0.2992	0.01157	1.377	1.075	1.764
SNP56	A	0.4222	0.8318	0.9745	0.7678	1.237
SNP57	T	0.4334	0.09275	1.22	0.9675	1.538
SNP59	A	0.4839	0.2167	0.8609	0.6791	1.091
SNP62	A	0.2749	0.8018	1.034	0.7976	1.34
SNP63	A	0.117	0.9125	0.9799	0.6827	1.407
SNP64	G	0.07895	0.3054	0.7869	0.4976	1.244
SNP65	G	0.3943	0.4663	0.9036	0.6832	1.195
SNP67	A	0.2822	0.6623	0.9439	0.7277	1.224
SNP75	A	0.2313	0.7808	0.9618	0.7315	1.265
SNP76	A	0.2287	0.06302	1.281	0.9855	1.664
SNP80	G	0.2643	0.9927	0.9987	0.7623	1.308
SNP82	G	0.4393	0.7198	1.043	0.8284	1.314
SNP95	C	0.09476	0.9157	1.022	0.6826	1.53
SNP99	C	0.1752	0.6263	0.9258	0.6783	1.264
SNP100	A	0.07355	0.7649	0.9352	0.6015	1.454
SNP103	A	0.07873	0.7784	0.9378	0.5999	1.466
SNP104	T	0.2173	0.9697	0.9946	0.7533	1.313
SNP105	A	0.2038	0.6905	0.9365	0.6779	1.294
SNP109	A	0.3966	0.9377	1.011	0.7758	1.317
SNP113	G	0.3521	0.8984	0.9819	0.7434	1.297
SNP114	T	0.3255	0.8603	1.022	0.7985	1.309
SNP118	A	0.3237	0.8256	0.9722	0.7562	1.25
SNP119	A	0.1358	0.9725	0.9941	0.709	1.394
SNP120	T	0.2151	0.07114	0.7481	0.5473	1.023
SNP121	A	0.1105	0.8343	0.9615	0.6656	1.389
SNP122	A	0.0427	0.3563	0.7502	0.4067	1.384
SNP123	C	0.1615	0.4722	1.119	0.823	1.523
SNP124	T	0.4119	0.9227	1.012	0.7988	1.282
SNP126	A	0.08523	0.5599	0.8807	0.5753	1.348
SNP128	A	0.1866	0.07736	1.299	0.9717	1.736
SNP130	G	0.3132	0.788	0.9655	0.7477	1.247
SNP136	G	0.04111	0.2946	1.337	0.7763	2.303
SNP137	G	0.3529	0.0002763	1.608	1.243	2.079
SNP139	T	0.2506	0.8931	1.019	0.7709	1.348
SNP141	A	0.2506	0.697	1.053	0.8108	1.369
SNP142	A	0.2557	0.9318	1.012	0.7713	1.327
SNP144	T	0.3076	0.007927	1.421	1.095	1.843
SNP145	G	0.3854	0.6788	0.9505	0.7475	1.209
SNP146	G	0.2385	0.7778	0.9615	0.7321	1.263
SNP147	A	0.3824	0.8902	1.018	0.7941	1.304
SNP149	T	0.1955	0.3136	0.8561	0.633	1.158
SNP150	A	0.3508	0.8912	0.9819	0.7562	1.275
SNP151	C	0.391	0.2565	1.145	0.9064	1.446
SNP152	C	0.4187	0.5018	0.9191	0.7161	1.18

SNP	A1	MAF	P	OR	L95	U95
SNP153	A	0.2416	0.7053	0.9394	0.6797	1.298
SNP156	G	0.3794	0.5895	0.9342	0.7301	1.196
SNP157	T	0.4117	0.595	0.9381	0.741	1.188
SNP160	C	0.446	0.5935	0.939	0.7455	1.183
SNP161	T	0.4461	0.578	0.9365	0.7432	1.18
SNP162	T	0.3966	0.2459	1.15	0.908	1.457
SNP163	T	0.278	0.773	1.038	0.8052	1.338
SNP165	G	0.1617	0.7681	0.9526	0.6896	1.316
SNP166	A	0.4898	0.2536	1.145	0.9036	1.45
SNP168	G	0.2497	0.6306	0.9352	0.7124	1.228
SNP169	T	0.09325	0.8509	1.039	0.6985	1.545
SNP171	T	0.3474	0.2393	1.16	0.906	1.485

Table 3: Logistic regression odds ratios and p-values for all SNPs

SNP	A1	TEST	OR	L95	U95	P
SNP1	T	ADD	0.9989	0.7526	1.326	0.9939
SNP1	T	COV1	1.099	0.7895	1.529	0.5769
SNP4	A	ADD	1.029	0.7505	1.41	0.8609
SNP4	A	COV1	1.058	0.7196	1.555	0.7753
SNP6	A	ADD	1.021	0.7768	1.341	0.8824
SNP6	A	COV1	1.079	0.712	1.635	0.7199
SNP7	C	ADD	1.045	0.8321	1.312	0.705
SNP7	C	COV1	1.078	0.7759	1.499	0.6535
SNP8	A	ADD	1.012	0.8021	1.276	0.9227
SNP8	A	COV1	1.074	0.7741	1.491	0.6683
SNP13	C	ADD	1.055	0.8558	1.299	0.618
SNP13	C	COV1	1.143	0.8141	1.604	0.4406
SNP16	T	ADD	1.028	0.778	1.357	0.8474
SNP16	T	COV1	1.093	0.7843	1.523	0.6003
SNP17	G	ADD	0.9217	0.7146	1.189	0.5303
SNP17	G	COV1	1.151	0.7974	1.662	0.4523
SNP18	A	ADD	1.046	0.8113	1.348	0.7289
SNP18	A	COV1	1.05	0.7168	1.538	0.8022
SNP19	C	ADD	1.087	0.8121	1.455	0.5747
SNP19	C	COV1	1.133	0.796	1.613	0.488
SNP25	G	ADD	1.057	0.8263	1.352	0.6598
SNP25	G	COV1	1.069	0.7462	1.532	0.7155
SNP30	G	ADD	0.7621	0.5805	1.001	0.05043
SNP30	G	COV1	1.055	0.7618	1.46	0.7483
SNP41	C	ADD	1.349	1.06	1.715	0.01474
SNP41	C	COV1	1.222	0.8746	1.708	0.2397
SNP56	A	ADD	0.9756	0.7728	1.232	0.8354
SNP56	A	COV1	1.079	0.7669	1.518	0.6628
SNP57	T	ADD	1.22	0.9671	1.538	0.09339
SNP57	T	COV1	1.123	0.8094	1.559	0.4867
SNP59	A	ADD	0.8678	0.6887	1.093	0.2289
SNP59	A	COV1	1.012	0.7169	1.428	0.9476
SNP62	A	ADD	1.034	0.7975	1.341	0.8008
SNP62	A	COV1	1.08	0.7726	1.508	0.654
SNP63	A	ADD	0.9806	0.6873	1.399	0.9139
SNP63	A	COV1	1.106	0.7966	1.535	0.5479
SNP64	G	ADD	0.7936	0.5056	1.246	0.3148
SNP64	G	COV1	1.148	0.8158	1.616	0.4278
SNP65	G	ADD	0.9018	0.6799	1.196	0.4733
SNP65	G	COV1	0.9724	0.6197	1.526	0.9032
SNP67	A	ADD	0.9408	0.7201	1.229	0.6549
SNP67	A	COV1	1.057	0.756	1.479	0.7448
SNP75	A	ADD	0.9447	0.6792	1.314	0.7356
SNP75	A	COV1	1.107	0.7966	1.539	0.5447
SNP76	A	ADD	1.272	0.982	1.647	0.06843
SNP76	A	COV1	1.074	0.777	1.485	0.6648
SNP80	G	ADD	0.9976	0.6841	1.455	0.9899
SNP80	G	COV1	1.125	0.7715	1.639	0.5413
SNP82	G	ADD	1.044	0.8269	1.319	0.7164
SNP82	G	COV1	1.101	0.796	1.523	0.5609
SNP95	C	ADD	1.022	0.684	1.526	0.916
SNP95	C	COV1	1.093	0.7665	1.559	0.6233
SNP99	C	ADD	0.9228	0.6718	1.268	0.62
SNP99	C	COV1	1.099	0.7879	1.534	0.5774
SNP100	A	ADD	0.9358	0.6031	1.452	0.7671
SNP100	A	COV1	1.1	0.7954	1.522	0.5636
SNP103	A	ADD	0.9391	0.6038	1.461	0.7806

SNP	A1	TEST	OR	L95	U95	P
SNP103	A	COV1	1.098	0.7733	1.558	0.6022
SNP104	T	ADD	0.9946	0.7533	1.313	0.9697
SNP104	T	COV1	1.114	0.8051	1.54	0.5156
SNP105	A	ADD	0.9317	0.6662	1.303	0.6792
SNP105	A	COV1	1.154	0.7619	1.749	0.4985
SNP109	A	ADD	1.01	0.7824	1.304	0.9396
SNP109	A	COV1	1.123	0.7557	1.668	0.5662
SNP113	G	ADD	0.9825	0.7473	1.292	0.8994
SNP113	G	COV1	1.117	0.7348	1.7	0.6036
SNP114	T	ADD	1.023	0.7972	1.312	0.8593
SNP114	T	COV1	1.09	0.7816	1.521	0.6106
SNP118	A	ADD	0.9822	0.804	1.2	0.8606
SNP118	A	COV1	1.072	0.7682	1.496	0.6831
SNP119	A	ADD	0.9937	0.7023	1.406	0.9717
SNP119	A	COV1	1.094	0.7858	1.522	0.5957
SNP120	T	ADD	0.744	0.5423	1.021	0.06678
SNP120	T	COV1	1.31	0.9031	1.901	0.1547

Chapter 6

The NRAMP1 gene of the African buffalo:

characterisation and variation

Nikki le Roex[1*], Paul D van Helden[1], Ad P Koets[2,3], Eileen G Hoal[1]

1 DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/ MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, Stellenbosch University, Cape Town, South Africa. **2** Department of Farm Animal Health, **3** Department of Infectious Diseases and Immunology, Faculty of Veterinary Medicine, Utrecht University, The Netherlands

Veterinary Immunology and Immunopathology (2013); submitted.

Abstract

NRAMP1 polymorphisms have been associated with susceptibility to salmonella, brucellosis, and paratuberculosis in numerous animal species. While the African buffalo is of high economic significance in the ecotourism and hunting industries in southern Africa, its role as a maintenance host of foot-and-mouth disease, corridor disease, bovine tuberculosis and brucellosis is of much concern. Ten African buffalo samples were used to sequence the full coding region and partial 5‟UTR, intronic and 3‟UTR regions of the NRAMP1 gene. Computer software was used to analyse the gene composition, identify variation, analyse the predicted structure of the protein and compare the translated amino acid sequence to other species. The NRAMP1 coding region spanned 1647 base pairs and generated an amino acid sequence of 548 residues and a stop codon. 15 novel SNPs and three microsatellites were identified within the gene. The predicted NRAMP1 protein contained 11 hydrophobic transmembrane domains, two N-linked glycosylation sites, four protein kinase C phosphorylation sites, and a cytoplasmic consensus transport signature. Phylogenetic analysis grouped the African buffalo NRAMP1 sequence closest to the water buffalo. This characterisation of the African buffalo NRAMP1 gene could facilitate further research in the burgeoning fields of disease genetics, animal breeding and vaccine development.

Keywords: *NRAMP1; SLC11A1; buffalo; membrane transport protein; gene characterisation; SNP*

1. Introduction

The natural resistance-associated macrophage protein 1 (NRAMP1, also known as SLC11A1) gene is a member of a small family of closely related genes, homologs of which have been identified in plants, animals, yeast and bacteria [1]. NRAMP1 is an integral membrane protein defined by a hydrophobic core, that typically consists of 10 transmembrane domains, two N-glycosylation sites and a conserved transport motif, known as the „binding protein-dependent transport systems inner membrane component signature‟ [2,3]. These domains are highly conserved, and mutations in these regions are likely to have significant consequences. The NRAMP1 protein is expressed on macrophage endosomes and lysosomes, plays a role in oxygen and nitrogen-dependent antimicrobial activity and the

maintenance of divalent metal homeostasis, and is believed to control the growth of intracellular pathogens during the early stages of infection [4].

Nramp1 was identified as the candidate gene for the Bcg/Lsh/Ity locus in mice, which confers innate resistance to intracellular pathogens such as *Mycobacteria*, *Leishmania* and *Salmonella* [2,5]. Mice with a Gly to Asp substitution at amino acid 169, located in the fourth transmembrane domain of the Nramp1 protein, are susceptible to infection by the aforementioned pathogens [6]. Human studies have linked NRAMP1 polymorphisms with susceptibility to leprosy [7], tuberculosis [8,9], Crohn‟s disease [10] and rheumatoid arthritis [11]. In animal studies, NRAMP1 polymorphisms have been associated with salmonella in chickens [12], brucellosis in water buffalo [13], and with paratuberculosis in cattle [14] and sheep [15].

African buffalo are one of the most abundant and ecologically important species of megafauna in the savannah ecosystem. They are an important prey species, as well as a species of high economic significance in the ecotourism and conservation hunting industries [16]. However, African buffalo host a vast array of nematodes, pathogens and infectious diseases, and their role as maintenance hosts of foot-and-mouth disease, corridor disease, bovine tuberculosis and brucellosis is of much concern [17]. This limits conservation options, since wildlife reservoirs of diseases such as these can lead to major economic, ecological and public health impacts on other wildlife species, livestock and surrounding farming communities [18,19].

It has become widely accepted that the variability seen in disease response in cattle, both between breeds and between individuals, has a genetic component. While genetic studies in disease resistance are relatively common in species for which there is a wealth of genetic information, this is not the case for other species, particularly wildlife, that are less well-characterised. The discovery of genetic variants involved in disease susceptibility can provide an assortment of potential drug and/or vaccine targets, as well as allow for the possibility of breeding animals with enhanced disease resistance. The application of genetic resistance information in breeding programs has been implemented in cattle for diseases such as mastitis [20], and would be invaluable in both the commercial and private sectors. For species such as the buffalo that have not had their genomes assembled, it is important that the genes potentially involved in disease resistance first be properly characterised, in order to enable future investigation.

In this paper, we sequence and report the full coding region of the African buffalo NRAMP1 gene, as well as partial 5‟UTR, intronic and 3‟UTR sequences. We identify variation present within the gene, which can be used to inform future genetic disease resistance studies, analyse the predicted structure of the protein, and compare the translated amino acid sequence to other species.

2. Materials and methods

2.1 Samples

DNA samples from ten African buffalo of the Cape buffalo subspecies (*Syncerus caffer caffer*), previously extracted from blood using the salt-chloroform extraction method [21], were used for the generation of the coding region of the NRAMP1 gene in this study. The buffalo samples originated from Kruger National Park, South Africa.

2.2 Primer design

Cattle and water buffalo *NRAMP1* sequences (AY438096.1, DQ493965.1, DQ493966.1, AY398759.1, AY398761.1, AY398765.1, AY398767.1, EF682067.2, DQ441408.1) from the NCBI database (http://www.ncbi.nlm.nih.gov/genbank/) [22] were aligned in BioEdit v7.1.11 [23] and used to generate a full NRAMP1 consensus sequence. The consensus sequence and Primer3web v4.0 [24] were used to design primers to span the exons of the NRAMP1 gene, in addition to the use of previously described primers used for water buffalo [25]. Primers were synthesized by the Synthetic DNA Laboratory, University of Cape Town, South Africa. The names and sequences of all primers used, as well as the regions covered, are listed in Table 1.

Table 1: Primer sequences, amplicon sizes and exons spanned.

Forward Primer	Reverse primer	Size* (bp)	T_A (°C)	Exons
PromF: GGGAGAGGTGCAGAACTCA	PromR: TGCTCACCTGACATGAGGA	315	60	5′UTR
1F: CAGCCACTCGCACAGAGAG	1R: CTGGTGAGATCCCGCATGAT	786	62	1 & 2
3F: CCAAGACCCCTCATTCCACT	3R: GGACTTGCAATGCTCCACTT	698	62	3 & 4
5F: TGCAGGACTGAGACAAGAGA[1]	5R: GATGAGAGGCACACACCCA[1]	675	62	5 & 6
7bF: AGTGCTCAGGATATCAGCCA[1]	7cR: GCTAAGGTCCTCTCCCCTGT	258	60	7
8F: GACTTTCACTCACCCCAC	7bR: ATTCAGAGGCAGGCGTCGAG	278	60	8
9F: ACTTCGATGCTTCCGTCCTT	9R: AGTCCCCTGCAGAACTTGAG	239	62	9
10bF: ACATCTCCTTCCTACTGCCC	10bR: CCACATCCGCGTCCTGAGTG	442	60	10 & 11
12F: CATCCCACCACCGAGTTCC	12R: GGATTGCTGTGGTTCCTGTT	247	62	12
13F: AAGGTTGCACAGTGAAGCAG	14bR: CTGTGGTCTCCTCTCCTTGC	785	60	13 & 14
14F: CACTTCCTCCATCATGGTG	UTRR: ATGGAACTCACGTTGGCTG[2]	658	62	15
UTRF: TGTGTGTGTGAAGGCAGCA[2]	UTRR: ATGGAACTCACGTTGGCTG[2]	190	60	3′UTR

* size is the inferred size from the *Bos taurus NRAMP1* complete gene sequence (DQ848779.1) and *Bos taurus NRAMP1* promoter region sequence (AY438096.1), including primer sequences.
[1]from [25]
[2]from [30]

2.3 Amplification and alignment

PCR amplification was done in 20µl reactions, containing final concentrations of 1X KAPA Biosystems ReadyMix Taq, 0.2µM forward and reverse primers, and 1-2µl DNA. PCR conditions were as follows: 94°C denaturing for 5 min, followed by 35 cycles of 94°C for 40 sec, 57-62°C for 35 sec and 72°C for 30 sec, followed by a final extension of 5 min at 72°C. PCR fragments were visualised by agarose gel electrophoresis and purified using the Promega Wizard Gel Purification Kit, according to the manufacturer's protocol. Sequencing of PCR products was performed by the Central Analytical Facility (CAF) at Stellenbosch University, South Africa. The sequences obtained for each region were subjected to BLAST analysis (www.ncbi.nlm.nih.gov/BLAST) to confirm their identity, and were checked visually and aligned in BioEdit. The nucleotide sequences were compared in BioEdit in order to identify novel polymorphisms within the gene, and used to generate a consensus sequence for each region. The African buffalo consensus sequences were aligned with the *Bos taurus* NRAMP1 cds (DQ493965.1) sequence and concatenated in order to generate a final coding region NRAMP1 sequence. The buffalo coding region nucleotide sequence was translated in BioEdit to obtain the amino acid sequence.

2.4 Analysis

The percentage homologies between the nucleotide and amino acid sequences of the African buffalo, cow (*Bos taurus*; DQ493965.1) and water buffalo (*Bubulas bubalis*; FJ827149.1)

NRAMP1 coding regions were calculated. Positions and consequences of SNPs were determined. TopPred [26] was used to predict the location of transmembrane domains within the buffalo NRAMP1 protein, and ScanProsite [27] was used to identify putative glycosylation and phosphorylation sites. Neighbour-joining phylogenetic analysis was performed in MEGA v5.2.1 [28] under default settings using the NRAMP1 amino acid sequences of *Syncerus caffer*, *Bos taurus* (NP777077.1), *Bos indicus* (ABF61464.1), *Bubulas bubalis* (ACO53852.1), *Bubulas carabanensis* (ACO53851.1), *Bison bison* (AAB17552.1), *Cervus elaphus* (AAC28240.1), *Ovis aries* (AAC28241.1), *Homo sapiens* (AAG15405.1), *Mus musculus* (NP_038640.2) *and Sus scrofa* (AAF36527.1).

2.5 3'UTR characterisation

Initial amplification and sequencing of the 3'UTR microsatellite region suggested the presence of more than two alleles per individual for this fragment. In order to clarify the situation, cloning and sequencing of the PCR amplicons was done. PCR amplification was done in 25ul reactions, containing final concentrations of 1X Supertherm Gold PCR buffer, 1.5mM MgCl2, 2mM dNTPs, 0.2µM of each primer, 2µl DNA and 0.05µl Supertherm Gold Taq. PCR conditions were as follows: 94°C denaturing for 5min, followed by 35 cycles of 94°C for 40 sec, 60°C for 35 sec and 72°C for 30 sec, followed by a final extension of 7 min at 72°C. PCR fragments were visualised by agarose gel electrophoresis under long wavelength UV, excised and gel-purified using Promega Wizard Gel Purification Kit, according to the manufacturer's protocol. Purified PCR fragments were ligated into the pGEM-T Easy vector (Promega, USA), and cloned into JMN109 competent *E. coli* cells. Plasmid DNA was isolated from each clone using the Promega Wizard SV Miniprep Purification Kit, as per the manufacturer's protocol, and sequenced by the Central Analytical Facility (CAF) at Stellenbosch University, South Africa. Sequences were checked visually, edited and aligned in BioEdit.

3. Results and discussion

The African buffalo NRAMP1 sequences were aligned with the *Bos taurus* NRAMP1 gene (DQ848779.1) and combined to produce a final sequence for the African buffalo NRAMP1 gene (accession number KF264050). The final NRAMP1 coding region spanned 1647 base

pairs (Figure 1A), had a base composition of 58.9% G+C content, and generated an amino acid sequence of 548 residues (Figure 1B).

Comparative analysis identified 15 novel single nucleotide polymorphisms and three microsatellites within the buffalo NRAMP1 gene. Of these 15 SNPs, 5 were found in the 5‟ and 3‟ untranslated regions, 9 were intronic, and one was exonic. The distribution of the SNPs, with positions in brackets relative to the African buffalo NRAMP1 sequence (KF264050), were as follows: 5‟UTR – three SNPs (104, 214, 228), intron 1 – one SNP (580), intron 3 – four SNPs (2206, 2226, 2236, 2279), intron 5 – two SNPs (3902, 3974), intron 6 – one SNP (4554), exon 7 – one SNP (4608), intron 12 – one SNP (8176), and 3‟UTR – two SNPs (10631, 10655). The exonic SNP was a synonymous substitution. The greater number of intronic versus exonic SNPs is to be expected, due to the evolutionary constraints placed on the coding regions of expressed genes. Similarly, the presence of only one exonic SNP that does not alter the amino acid sequence of this highly conserved protein is unsurprising. However, variants within the 5‟ UTR can have a significant impact on gene expression, and thus variants from this region are often the targets of genetic association studies. Three $(GT)_n$ microsatellites were present within the 3‟UTR region (positions 10538, 10631, 10667), showing a range of allele sizes. Cloning confirmed the presence of four alleles per individual for part of the 3‟UTR. The presence of multiple alleles per individual suggests that at least one duplication of this region may have occurred in the African buffalo; however, this requires further investigation for confirmation. This region has been associated with brucellosis and paratuberculosis [13,14], among other diseases, and would benefit from further research.

A

```
1    GTTGTGGGGC TCAGGCCAGG AGGGGAACAA ATGTCCTTAC TGTGGCCACC CAGACCCTTC CTCCAGAACT GGCTACTTCT GCCTTTGGAA AGTGTTTCAC
101  AACRTGCCTG CGTGTGAGGA CAGCTGAGGA TGGGTAAGAG GGCGTGGACC CGCACAATCG CCACTGCCTA GAGAGGAAGC ACAGGCTTGC CATGCCCGTG
201  AGGGGCTGCC CGGMACGCCA GCCACTCRCA CAGAGAGTGC CCGAGCCTGC AGTCCTCATG TCAGGTGACA CAGGCCCCCC AAAGCAGGGA GGGACCAGAT
301  ATGGCTCCAT CTCCAGCCCA CCCAGTCCAG GGCCACAGCA AGCGCCTCCC GGAGGGACCT ACCTAAGTGA GAAGATCCCC ATTCCGGATA CAGAATCGGG
401  TGCATTCAGC CTGCGGAAGC TGTGGGCCTT CACAGGGCCT GGATTCCTCA TGAGCATCGC ATTCCTGGAC CCAGGAAACA TTGAGTCGGA TCTTCAGACT
501  GGGGCTGTGG CTGGATTCAA ACTGCTCTGG GTGCTGCTGT GGGCCACAGT GTTGGGCTTG CTCTGCCAGC GACTGGCTGC CCGGCTGGGT GTGGTGACAG
601  GCAAGGACTT GGGCGAGGTC TGCCATCTCT ACTACCCTAA GGTGCCCCGC ATTCTCCTCT GGCTGACCAT CGAGCTAGCC ATCGTGGGCT CAGACATGCA
701  GGAAGTCATT GGCACAGCTA TTGCATTCAG TCTGCTCTCT GCCGGACGAA TCCCACTCTG GGGTGGTGTC CTCATCACTG TCGTGGACAC TTTCTTCTTC
801  CTCTTCCTCG ATAACTACGG GTTGCGGAAG CTGGAAGCCT TTTTTGGATT TCTWATTACC ATAATGGCCT TGACCTTCGG CTATGAGTAC GTGGTGGCTC
901  AGCCTGCTCA GGGAGCATTG CTTCAGGGCC TGTTCCTGCC CTCGTGCCCA GGCTGTGGCC AGCCCGAGCT GCTGCAGGCC GTGGGCATCA TTGGCGCCAT
1001 CATCATGCCC CACAACATCT ACCTGCATTC CTCCCTCGTC AAGTCTCGAG AGGTAGACCG GTCCCGGCGG GCGGACATCC GAGAAGCCAA CATGTACTTC
1101 CTGATTGAAG CCACCATCGC CCTGTCTGTC TCCTTCCTCA TCAACCTCTT TGTCATGGCT GTCTTTGGGC AAGCCTTCTA CAAGCAAACC AACCAGGCTG
1201 CGTTCAACAT CTGTGCCAAC AGCAGCCTCC ACGACTACGC GCCGATCTTT CCCAGGAACA ACCTGACCGT GGCAGTGGAC ATTTACCAAG GAGGCGTGAT
1301 CCTGGGCTGC CTCTTTGGTC CTGCAGCCCT GTACATCTGG GCCGTGGGTC TCCTGGCTGC TGGGCAGAGC TCCACCATGA CCGGCACCTA CGCGGGACAG
1401 TTTGTGATGG AGGGCTTCCT GAAGCTGCGG TGGTCACGCT TCGCCCGAGT CCTGCTCACT CGCTCCTGCG CCATCCTGCC CACTGTGCTC CTGGCTGTCT
1501 TCAGGGACCT GCGGGACCTG TCAGGCCTCA ACGACCTGCT CAATGTGCTG CAGAGCCTGC TGCTTCCCTT CGCTGTGCTG CCCATCCTCA CCTTCACCAG
1601 CATGCCCGCC CTGATGCAGG AGTTTGCCAA TGGCCTGGTG AGCAAAGTTA TCACTTCCTC CATCATGGTG CTGGTCTGCG CTGTCAACCT TTACTTCGTG
1701 ATCAGCTACG TGCCCAGCCT CCCCCACCCT GCCTACTTCA GCCTTGTAGC ACTGCTGGCC GCAGCCTACC TGGGCCTCAC CACTTACCTG GTCTGGACCT
1801 GTCTCATCAC CCAGGGAGCC ACTCTTCTGG CCCACAGTTC CCACCAACGC TTCCTGTATG GGCTTCCTGA AGAGGATCAG GAGAAGGGGA GGACCTCGGG
1901 ATGAGCTCCC ACCAGGGCCT GGCCACGGGT GTGATGAGTG GGCATAGTGG CCTGTCAGAC AAGGGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGAAG
2001 GCAGCAAGAC GGAGAGGGAG TTCTGGAAGT TGGCCAACGT GAGTTCCAGA GGGACCTRTG TGTGTGTGTG ACACACTGGC CYGCCAGACA AAGGTGTGTG
2101 TGTGTGTGTG TGTGTGTACG CACAGCAAGA CAGGGAGTTC TGGAAGCCAG CCAACGTGAG TTCCATAAT
```

B

```
1   MSGDTGPPKQ GGTRYGSISS PPSPGPQQAP PGGTYLSEKI PIPDTESGAF SLRKLWAFTG PGFLMSIAFL DPGNIESDLQ TGAVAGFKLL WVLLWATVLG
101 LLCQRLAARL GVVTGKDLGE VCHLYYPKVP RILLWLTIEL AIVGSDMQEV IGTAIAFSLL SAGRIPLWGG VLITVVDTFF FLFLDNYGLR KLEAFFGFLI
201 TIMALTFGYE YVVAQPAQGA LLQGLFLPSC PGCGQPELLQ AVGIIGAIIM PHNIYLHSSL VKSREVDRSR RADIREANMY FLIEATIALS VSFLINLFVM
301 AVFGQAFYKQ TNQAAFNICA NSSLHDYAPI FPRNNLTVAV DIYQGGVILG CLFGPAALYI WAVGLLAAGQ SSTMTGTYAG QFVMEGFLKL RWSRFARVLL
401 TRSCAILPTV LLAVFRDLRD LSGLNDLLNV LQSLLLPFAV LPILTFTSMP ALMQEFANGL VSKVITSSIM VLVCAVNLYF VISYVPSLPH PAYFSLVALL
501 AAAYLGLTTY LVWTCLITQG ATLLAHSSHQ RFLYGLPEED QEKGRTSG*
```

Figure 1: A - African buffalo NRAMP1 partial 5ʻUTR, full coding region and partial 3ʻUTR. The 5ʻUTR and 3ʻUTR segments are underlined. B - Translated NRAMP1 amino acid sequence of the African buffalo. The binding-protein-dependent transport system inner membrane component signature (BPDTSIMCS) is underlined.

At the nucleotide level, the African buffalo NRAMP1 coding region showed a 98.79% and 98.72% homology to that of the cow and water buffalo, respectively. The translated buffalo amino acid sequence was identical in length, and showed a 98.91% and 98.72% homology, to that of the cow and water buffalo, respectively. Analysis of the protein structure revealed the presence of 11 transmembrane domains, including the core Pfam NRAMP1 domain, two N-glycosylation sites, and four potential protein kinase C phosphorylation sites (Figure 2). The binding-protein-dependent transport system inner membrane component signature (BPTSIMCS) motif was found from amino acid residues 370-389, located on the intracellular

loop between transmembrane domains 7 and 8. These structural features are in agreement with the features typical of this family of transport proteins, which are: 10-12 transmembrane segments, a glycosylated extracellular loop, and a cytoplasmic consensus transport signature [3].

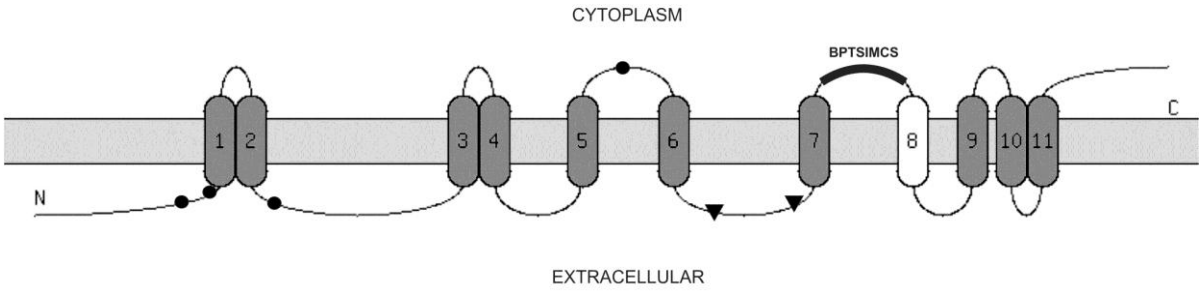

Figure 2: Schematic diagram of the 11 predicted hydrophobic transmembrane domains of the African buffalo NRAMP1 protein. Shaded segments indicate unambiguous transmembrane domains; clear segments indicate putative transmembrane domains. The BPTSIMCS is shown as a thick, bold portion of the intracellular loop. N-glycosylation sites are indicated with solid triangles; protein kinase C phosphorylation sites are indicated with solid circles.

The phylogenetic relationship between the African buffalo (*Syncerus caffer*), cattle (*Bos taurus*), cattle (*Bos indicus),* water buffalo (*Bubulas bubalis*), water buffalo (*Bubulas carabanensis*), bison (*Bison bison*), red deer (*Cervus elaphus*), sheep (*Ovis aries*), human (*Homo sapiens*), mouse (*Mus musculus*) and pig (*Sus scrofa*) NRAMP1 amino acid sequences can be seen in Figure 3. Phylogenetic analysis grouped the African buffalo NRAMP1 amino acid sequence closest to that of the water buffalo, followed by the cow and bison group. This reflects the proposed phylogenetic placement of the African buffalo in the Bovini tribe [29].

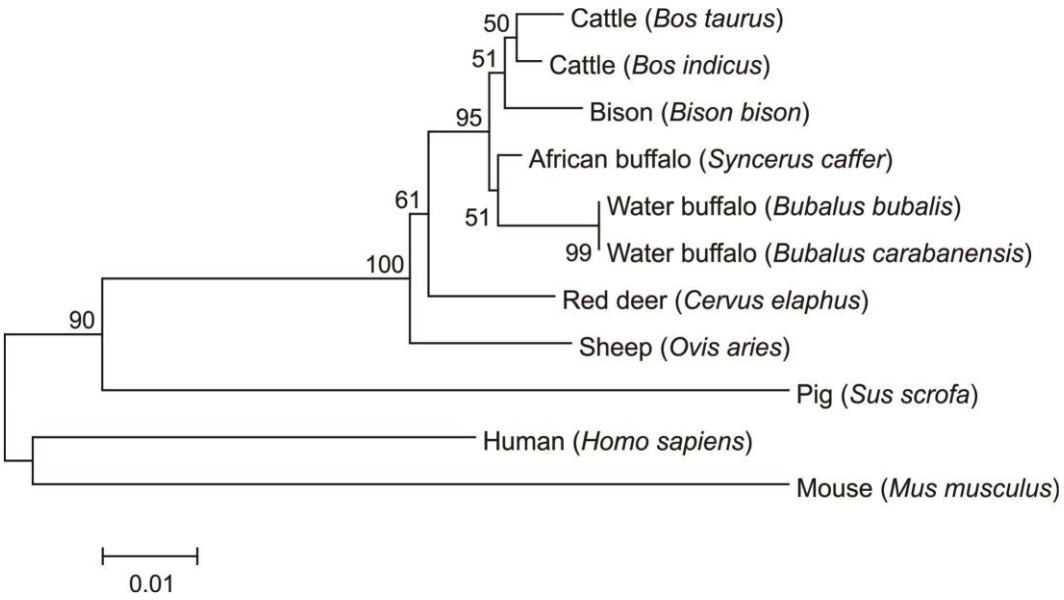

Figure 3: Unrooted neighbour-joining phylogenetic tree of NRAMP1 amino acid sequences. Bootstrap values are based on 1000 replications.

To the best of our knowledge, this is the first characterisation of the NRAMP1 gene in African buffalo. While disease association studies in animals that have fully assembled genomes, such as cattle, commonly utilise genome-wide association studies, this is not possible in many other species for which genome-wide information is not available. As such, the characterisation of the NRAMP1 gene in the African buffalo and identification of novel polymorphisms can facilitate further research in the burgeoning fields of disease genetics, animal breeding and vaccine development.

Conflict of interest

The authors declare no conflict of interest

Acknowledgements

We would like to acknowledge Roy Bengis, state veterinarians and SANParks Veterinary Wildlife Services (VWS), for the collection of buffalo blood samples. Without those samples, the study would not have been possible. This study was funded by a WOTRO Science for Global Development grant (W01.65.321.00) awarded to PVH, EGH and AK. NLR was supported by the same grant.

References

1. Saeij JP, Wiegertjes GF, Stet RJ (1999) Identification and characterization of a fish natural resistance-associated macrophage protein (NRAMP) cDNA. Immunogenetics 50: 60–66.

2. Vidal SM, Malo D, Vogan K, Skamene E, Gros P (1993) Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. Cell 73: 469–485.

3. Cellier M, Privé G, Belouchi A, Kwan T, Rodrigues V, et al. (1995) Nramp defines a family of membrane proteins. Proc Natl Acad Sci USA 92: 10089–10093.

4. Wigley P (2004) Genetic resistance to Salmonella infection in domestic animals. Res Vet Sci 76: 165–169. doi:10.1016/S0034-5288(03)00117-6.

5. Skamene E, Pietrangeli CE (1991) Genetics of the immune response to infectious pathogens. Curr Opin Immunol 3: 511–517.

6. Malo D, Vogan K, Vidal S, Hu J, Cellier M, et al. (1994) Haplotype mapping and sequence analysis of the mouse Nramp gene predict susceptibility to infection with intracellular parasites. Genomics 23: 51–61. doi:10.1006/geno.1994.1458.

7. Abel L, Sánchez FO, Oberti J, Thuc NV, Hoa LV, et al. (1998) Susceptibility to leprosy is linked to the human NRAMP1 gene. J Infect Dis 177: 133–145.

8. Bellamy R, Ruwende C, Corrah T, McAdam KP, Whittle HC, et al. (1998) Variations in the NRAMP1 gene and susceptibility to tuberculosis in West Africans. N Engl J Med 338: 640–644. doi:10.1056/NEJM199803053381002.

9. Hoal EG, Lewis LA, Jamieson SE, Tanzer F, Rossouw M, et al. (2004) SLC11A1 (NRAMP1) but not SLC11A2 (NRAMP2) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. Int J Tuberc Lung Dis 8: 1464–1471.

10. Gazouli M, Atsaves V, Mantzaris G, Economou M, Nasioulas G, et al. (2008) Role of functional polymorphisms of NRAMP1 gene for the development of Crohn‟s disease. Inflamm Bowel Dis 14: 1323–1330. doi:10.1002/ibd.20488.

11. Ates O, Dalyan L, Müsellim B, Hatemi G, Türker H, et al. (2009) NRAMP1 (SLC11A1) gene polymorphisms that correlate with autoimmune versus infectious disease susceptibility in tuberculosis and rheumatoid arthritis. Int J Immunogenet 36: 15–19. doi:10.1111/j.1744-313X.2008.00814.x.

12. Hu J, Bumstead N, Barrow P, Sebastiani G, Olien L, et al. (1997) Resistance to salmonellosis in the chicken is linked to NRAMP1 and TNC. Genome Res 7: 693–704.

13. Capparelli R, Alfano F, Amoroso MG, Borriello G, Fenizia D, et al. (2007) Protective effect of the Nramp1 BB genotype against Brucella abortus in the water buffalo (*Bubalus bubalis*). Infect Immun 75: 988–996.

14. Pinedo PJ, Buergelt CD, Donovan GA, Melendez P, Morel L, et al. (2009) Candidate gene polymorphisms (BoIFNG, TLR4, SLC11A1) as risk factors for paratuberculosis infection in cattle. Prev Vet Med 91: 189–196.

15. Reddacliff LA, Beh K, McGregor H, Whittington RJ (2005) A preliminary study of possible genetic influences on the susceptibility of sheep to Johne‟s disease. Aust Vet J 83: 435–441.

16. Booth VR (2009) A comparison of the prices of hunting tourism in southern and eastern Africa. Badapest, Hungary: International Council for Game and Wildlife Conservation and Food and Agriculture Organisation of the United Nations.

17. Jolles AE (2007) Population biology of African buffalo (*Syncerus caffer*) at Hluhluwe-iMfolozi Park, South Africa. Afr J Ecol 45: 398–406.

18. Van Rhijn I, Godfroid J, Michel A, Rutten V (2008) Bovine tuberculosis as a model for human tuberculosis: advantages over small animal models. Microbes Infect 10: 711–715.

19. Ayele WY, Neill SD, Zinsstag J, Weiss MG, Pavlik I (2004) Bovine tuberculosis: an old disease but a new threat to Africa. Int J Tuberc Lung Dis 8: 924–937.

20. Rupp R, Boichard D (2003) Genetics of resistance to mastitis in dairy cattle. Vet Res 34: 671–688. doi:10.1051/vetres:2003020.

21. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16: 1215.

22. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, et al. (2013) GenBank. Nucleic Acids Res 41: D36–42. doi:10.1093/nar/gks1195.

23. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41: 95–98.

24. Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol 132: 365–386.

25. De Matteis G, Napolitano F, Scata MC, Crisa A, Catillo G, et al. (2009) Molecular characterization of the NRAMP1 gene in buffalo. Ital J Anim Sci 8: 72–74.

26. Von Heijne G (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol 225: 487–494.

27. De Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, et al. (2006) ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 34: W362–365. doi:10.1093/nar/gkl124.

28. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28: 2731–2739. doi:10.1093/molbev/msr121.

29. MacEachern S, McEwan J, Goddard M (2009) Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (Bovidae, Bovinae). BMC Genomics 10: 177. doi:10.1186/1471-2164-10-177.

30. Ganguly I, Sharma A, Singh R, Deb SM, Singh DK, et al. (2008) Association of microsatellite (GT)n polymorphism at 3′UTR of NRAMP1 with the macrophage function following challenge with Brucella LPS in buffalo (*Bubalus bubalis*). Vet Microbiol 129: 188–196.

Chapter 7

Breeding TB-resistant African buffalo: the future of BTB management?

Courtney Berrington[1+], Nikki le Roex[1+*], Eileen G Hoal[1], Paul D van Helden[1]

1 DST/NRF Centre of Excellence for Biomedical Tuberculosis Research/ Medical Research Council (MRC) Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Faculty of Health Sciences, Stellenbosch University,Cape Town, South Africa.

[+] both authors contributed equally

Journal of Animal Breeding and Genetics (2013); submitted.

Abstract

The high prevalence of bovine tuberculosis (BTB) in buffalo herds in southern African game reserves represents not only an ecological threat to biodiversity but a health risk through the wildlife-cattle-human interface. With standard test and slaughter methods not being logistically feasible in this free range system, the co-existence of alternative BTB hosts and the limited effectiveness of the BCG vaccine in buffalo, there is a need for another method of BTB management. This review assesses breeding for increased resistance to BTB in buffalo as a possible method of BTB management and future eradication of the disease. Main lines of evidence for the heritability of BTB resistance in mammals are briefly discussed, and we consider how the information obtained from these studies can be applied to breeding programmes. The principles and challenges of selective breeding are addressed, and the different methods that can be employed in such breeding programmes are examined in more detail. Finally, conclusions and suggestions for the future are discussed.

Keywords: *M. bovis, BTB, buffalo, resistance, breeding*

1. Introduction

African buffalo are a significant component of the savannah ecosystem, due to their physical size, large numbers and their role as coarse grazers, which has implications for other fauna and flora [1]. They also form a considerable proportion of the prey biomass of lions, which are the apex predator in this ecosystem. Varying environmental conditions, such as drought, can cause a dramatic decrease in the buffalo population, as a result of reduced body condition and susceptibility to starvation, predation and infectious disease. This was seen during the drought of 1982/3 in the Kruger National Park (KNP), South Africa, when the buffalo population declined by more than 50% [2]. Under severe drought conditions, both the spread and severity of bovine tuberculosis may be altered, due to changes in herd density and the toll that environmental stress takes on the immune system, respectively [3,4]. Thus, although stable at present, the presence of this new, introduced disease may have profound implications for the long-term stability of the buffalo population.

The high prevalence of BTB in buffalo herds in southern African game reserves represents not only an ecological threat to biodiversity, but also a health risk through the wildlife-cattle-human interface. In a study of cattle from 27 villages in Tanzania, Cleaveland et al. [5]

reported that cattle in contact with wildlife showed a significantly higher prevalence of BTB infection. Sufficient interface contact also makes *M. bovis*, the causative agent of BTB, a zoonotic risk to human health, since infection in humans has been known to occur through the consumption of unpasteurized milk or meat of infected livestock. High HIV prevalence within the rural communities that surround South African game reserves puts people at an increased risk of acquiring TB due to compromised immune systems [6,7].

Cattle are believed to have introduced *M. bovis* into KNP through the (then) unfenced southeast corner of the park during the late 1950‟s/early 1960‟s [8]. The first confirmed BTB diagnosis in buffalo within the park occurred in 1990, and the extent of disease spread and the large herd sizes and ranges meant that standard test and slaughter methods were not logistically feasible in this free-range system [9]. Furthermore, any large-scale slaughtering would have resulted in public outcry, causing negative publicity for the park. The establishment of a BTB control strategy was also complicated by the presence of co-existing/spill-over BTB hosts, such as lion (*Panthera leo)*, chacma baboon (*Papio ursinus)*, warthog (*Phacochoerus africanus)* and greater kudu (*Tragelaphus strepsiceros*) [7].

While prevention may be better than cure, the BCG vaccine has limited effectiveness. In buffalo, the BCG vaccine showed no significant protection against experimental challenge with *M. bovis*, and consequently this does not represent a viable control strategy for this species [3,10]. Many developed countries have had great success in eradicating or drastically reducing BTB prevalence in cattle using regular tuberculin test and slaughter policies, but the logistic demands of operating such a programme in free-ranging wildlife populations of this size render this option impractical as well [11]. For the above reasons, selective breeding for increased BTB resistance or tolerance should be seriously considered as a potential control strategy in the management of bovine tuberculosis in African buffalo.

Resistance and tolerance are the two main aspects of self defence against pathogens; these aspects determine the disease severity and have been shown to exhibit substantial genetic variation in animal models [12]. Whilst resistance involves limiting the actual bacterial burden, tolerance involves the restriction of the harmful consequences caused by the bacteria. Resistance and tolerance are often shown to be negatively correlated in infectious disease, with a trade-off occurring between them as they employ opposing strategies: the cost of an increased immune control of infection is typically an increase in „collateral damage‟ of infected tissue [12]. A possible negative outcome to breeding for increased resistance is the

antagonistic co-evolution of the pathogen. Any alteration in the resistance of the host to the pathogen places selective pressure on the pathogen and corresponding counter-adaptations could occur. No such selective pressure is placed on the pathogen if selection of the host is tolerance-based, thus diminishing the antagonistic co-evolution of host and pathogen [12]. However, the presence of co-existing BTB hosts makes increasing the resistance of buffalo to BTB a more desirable goal, as it would be more effective at reducing interspecies spread by reducing BTB prevalence in a major host.

In this review, we briefly discuss lines of evidence of the heritability of BTB resistance in mammals, and consider the benefits and challenges of selective breeding. We compare the different methods of selective breeding, as well as the suitability of each of these methods for breeding BTB resistance in African buffalo. Finally, we make recommendations for future work.

2. Evidence of resistance to BTB in mammals

The heritability of resistance to infection after exposure to *M. bovis* has been noted in a number of mammals. As early as the 1940‟s, studies showed that resistance to *M. bovis* infection was heritable. Lurie [13] showed that inbred strains of rabbits displayed one of two phenotypes (susceptible or resistant) following infection with *M. bovis*. This inheritance pattern suggested that the disease profile displayed by the rabbits was largely genetically inherited.

A 3-year heritability study by Mackintosh et al. [14] on *M. bovis* resistance in farmed red deer, was the first of its kind in livestock. Red deer stags were challenged intratonsillarly with virulent *M.bovis*, and following euthanasia, lesion severity was graded and BTB infection confirmed. Lesion severity scores (LSS) calculated for each individual were used to select six stags to represent the range of resistance. Their semen was used to artificially inseminate BTB-free red deer hinds, resulting in the birth of 70 BTB-free calves. These calves were then challenged with *M. bovis* intratonsillarly, and regular skin testing took place until the calves were euthanised and necropsied. Each calf‟s parentage was determined via genetic marker analysis and its lesion severity score calculated [14]. The lesion scores in offspring corresponded with the lesion scores of stags: offspring from the non-infected stags had the lowest lesion scores and offspring from the severely infected stags had the highest lesion

scores. A heritability estimate of 0.48 was calculated for the observed resistance. None of the resistant stags produced antibody in response to the tuberculin, whilst susceptible stags produced levels of antibody to tuberculin that correlated with their severity of infection [14]. Thus the resistant and susceptible stags showed two distinctive immune response patterns which could be tested for and exploited through phenotypic selection.

In Irish Holstein-Friesian dairy herds, 15,182 cows and 8,104 heifers with a high likelihood of *M. bovis* exposure were repeatedly tested over three years for responsiveness to *M. bovis*-purified protein derivative (PPD) in single comparative intradermal tuberculin tests (SCITT), and produced heritability estimates of 0.14 and 0.12 respectively for responsiveness [15]. Data on the presence of TB lesions in the carcasses, confirming *M. bovis* infection, produced a heritability estimate of 0.18 in cows for susceptibility to infection. A strong correlation between responsiveness to *M. bovis*-PPD and resistance to confirmed *M. bovis* infection was evident, making indirect selection an option: selection for responsiveness would be effective at shifting the population mean towards decreased susceptibility [15]. The correlation also suggests that both traits may be influenced by the same or closely positioned genes.

These studies serve as examples of the heritability of *M. bovis* infection, and how this information might be utilised in selective breeding programmes. A more extensive review of the genetic studies showing evidence of heritability or association of BTB resistance in cattle and wildlife can be found in le Roex et al. [16].

3. Selective breeding

The genetic improvement of livestock through selective breeding programmes has been successfully applied to many species and traits. Selective breeding programmes seek to identify those individuals with a particular trait of interest, and utilise these individuals preferentially for breeding. Thus over time, the trait will become more prevalent in a particular population.

Heritability is a crucial parameter in breeding programs, since it provides an indication of the potential success of this strategy for a particular trait [17]. Ideally, heritability studies should first be undertaken in African buffalo in order to determine the extent to which the observed phenotypic variation of BTB resistance is caused by genetic variance and is not a function of environmental variance. The higher the heritability score, the greater the role that genetics

plays in determining the phenotype and the more potentially enhanceable the trait is through selective breeding [17]. In the case of free-ranging mixed herd wildlife populations, where the family structure or relatedness of the individuals is unknown, a breeding plan must first be developed, and the offspring from the second generation allowed to mature before heritability values can be calculated.

Selective breeding programmes could also benefit from risk management practices, such as frequent testing and status monitoring, quarantine and treatment of infected animals with antibiotics, but these practices are not possible in a free-ranging environment. In order to benefit from these management practices, wildlife populations such as buffalo would need to be maintained in a fenced environment under controlled conditions, similar to that of cattle. This practice has other advantages in African buffalo, such as increased feed quality and availability, which results in a younger age of first calving (pers. comm Bernard Wooding). Another factor that needs to be considered is the suitability of the phenotypic test. The more accurate the diagnostic test to determine the phenotype, the more reliable the association between genotype and phenotype, and the better the results from a selective breeding programme. The single comparative intradermal tuberculin test (SCITT) is the most commonly used diagnostic test for BTB infection, and sensitivity and specificity estimates for this test in buffalo have not been officially established. The direct selection for the ability to pass the SCITT test, rather than direct selection for resistance phenotypes, may also result in the diminished ability of buffalo to develop a measureable response to the SCITT. This would result in an inability to detect exposure to *M. bovis*/BTB infection, thus undermining the very usefulness of the test [18].

The selection of individuals for these breeding programmes can be based on either phenotypic or genotypic merit. A schematic diagram of the basic steps of a selective breeding programme for BTB resistance in African buffalo can be seen in Figure 1. The methods used for these approaches, and their suitability to African buffalo, are discussed below.

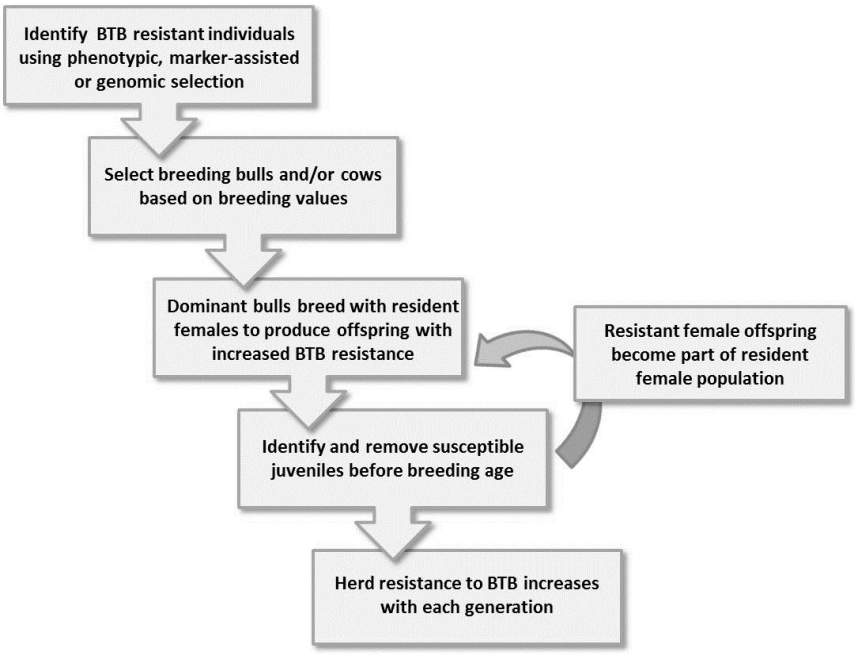

Figure 1: Schematic overview of steps for breeding BTB resistance in African buffalo

3.1 Phenotypic selection

The selection of animals for breeding programs has traditionally been based on the estimated breeding values (EBVs) of individuals, calculated using phenotypic records, pedigrees and the heritability of a particular trait [19,20]. An EBV is an estimate of the genetic merit of an individual with regards to a particular trait, without identifying or having knowledge of the actual genes involved in the trait. Many of the traits that are important in agriculture are complex or quantitative traits; these traits are influenced by many genes and produce a continuous distribution of phenotypes among individuals [19]. Beef and dairy cattle stud farms routinely measure quantitative traits across herds in order to enhance traits such as total milk yield, butter fat percentage and body composition. Extensive data is also kept on the family structure within herds, which is essential for analysing the relationships between individuals. Such analyses contribute to the calculation of each animal‟s breeding value, as incorporating the phenotypic information of the offspring gives a more reliable indication of the genetic merit of the bull than the bull‟s phenotypic measure alone [21,22]. In situations where pedigrees are unknown, appropriate breeding and record keeping must first be carried out.

Phenotypic studies have investigated the resistance of cattle to *M. bovis* through the measurement of the individual‟s inflammatory response and by trial exposures *in vivo* and *in vitro*. The continuous distribution of resistance as a trait (measured by lesion scores or SCCIT inflammatory responses), and the likelihood of this trait being polygenic and multi-factorial, suggests that resistance to *M. bovis* can be regarded as a quantitative trait [21]. Thus the same process should be applicable to breeding for BTB resistance as it is for a quantitative trait. The breeding value of a bull as a function of its resistance to BTB can be calculated using the relative phenotypic resistance of closely related animals, and is a more robust measure than the bull‟s own phenotypic measure of resistance [21,22]. This breeding value is a measure of how breeding with that male could influence the overall resistance of the herd. In order to draw the most accurate comparisons between parent and offspring resistance phenotypes, the offspring should be of immunologically comparable age when tested. For example, a study to establish the disease status in African buffalo calves performed gamma-interferon testing at 9 and 9.5 months, and skin testing from 11 months, which ensured that the immune responses of the offspring were not influenced by milk-derived maternal immunity [10].

While selection of complex traits (such as milk yield in cows) based on phenotypic records has been successful in the past, it also can be difficult, and costly. Often the phenotype of interest can only be measured at a particular stage of life, after death, in one sex, or is very expensive to test [19].

3.2 Marker-assisted selection

Marker-assisted selection incorporates the information obtained from a small number of markers into the traditional EBV calculations – that is, breeding values are calculated by combining pedigree, phenotype and marker information [19]. Markers are typically identified by candidate gene association studies in the trait of interest. The success of marker-assisted selection depends on a number of factors, such as the number of genes contributing to a trait, the relative contribution of each gene, and what has been previously attained in terms of genetic improvement using phenotypic selection [23]. For example, if phenotypic selection has been implemented for many years and substantial improvement has already been achieved, it is unlikely that the addition of a single significant genetic marker to the EBV calculation will have a considerable effect. However, in relatively new breeding programs, the incorporation of any genetic information into the calculation of breeding values could

substantially improve the EBV calculations and thus the success of the program [23]. Marker-assisted selection has been used in a number of livestock species already, to improve traits such as reproductive rate, body composition and growth rate [24].

In candidate gene association studies, genes are selected based on prior knowledge of the physiology of a particular trait, and polymorphisms/variants within the gene are tested for statistical association with that trait. While BTB resistance is a polygenic trait, if the effect of a particular gene is large enough to cause a discontinuity of the phenotype (despite the effects of other loci and non-genetic factors), it can be detected in an association study and thus has possible utility in breeding programs [21]. In cattle, strong and consistent correlations have been shown between the microsatellite markers INRA111 and BMS2753 and the reactor status of the SCITT test within nine diverse breeds [25]. Marker INRA111 has also shown association with ovine foot rot and is located in a region of the cow genome associated with mastitis. The 2-2 genotype of INRA111, common in all of these nine breeds, confers a twofold protection against BTB, on average [25]. Sun et al. [26] investigated seven SNPs in the Toll-Like Receptor (TLR) 1 and 9 genes in Chinese Holstein cattle, and found that polymorphism TLR-G1596A exhibited a significant association with *M. bovis* infection. In a Chadian breed of cattle, microsatellite polymorphisms in the Natural Resistance Associated Macrophage Protein 1 (NRAMP1/SLC11A1) gene, specifically alleles 211, 215 and 217, were significantly associated with lower incidence of infection of BTB [27].

In African buffalo, an association study done by le Roex et al. [28] tested 69 SNPs in immune-related genes for association with BTB infection status, and found three positive associations. These SNPs were located in the Solute Carrier Family 7 Member A13 (SLC7A13) gene, Deleted in Malignant Brain Tumour 1 (DMBT1) gene and Interleukin 1 alpha (IL1α) gene. SLC7A13 is a heteromeric amino acid transporter (HAT), and although relatively uncharacterised, other transporters, such as SLC11A1 (NRAMP1), have been associated with TB infection in humans [29]. DMBT1 is regarded as a pattern recognition receptor (PRR), which forms part of the human host defence system against invading organisms such as *M. tuberculosis* [30]. The IL1α protein is a pro-inflammatory cytokine, and plays a role in the activation of T-lymphocytes, and the recruitment of leucocytes to sites of inflammation [31]. These SNPs are promising candidates for the selective breeding of buffalo with increased BTB resistance. In addition to these SNPs associated with BTB in African buffalo, the high sequence homology between African buffalo and bovine genes, such as SLC11A1 (N. le Roex, unpubl.), suggests that SNPs identified in the cattle BTB

studies may be potentially useful additions to the set of candidate genes that could be utilized for screening for resistant African buffalo for breeding programs.

3.3 Genomic Selection

Genomic selection refers to the selection of individuals based on marker or haplotype information obtained from a dense array of markers that cover the entire genome, thus theoretically incorporating all the loci that contribute to a particular trait [22]. This information is compiled to calculate a genomic breeding value (GEBV), which can then be used to predict the breeding value of individuals in the absence of phenotype data [19]. Genomic selection is highly applicable in livestock species where uniformity of phenotype is desired and lack of trait diversity is considered beneficial (excluding stud farms). In order to generate the marker information required for genomic selection, a large reference population with phenotype data is required. This is one of the biggest challenges facing genomic selection, but if such a population can be established, all measurable traits of interest can be phenotyped in this same population, and thus GEBVs for all phenotypes can be calculated [19]. This is an enormous amount of invaluable data, and has the potential to dramatically alter the future of breeding programs. Success has already been achieved - studies in Holstein-Friesan cattle have shown up to a 71% concordance between GEBVs and true breeding values, averaged across a number of traits [32], and studies in broiler chickens showed a fourfold increase in the accuracy of prediction of a food conversion rate phenotype [33].

The African buffalo genome has not been assembled as yet, and no commercial SNP chip for the buffalo is available. Without the existence of a genome-wide marker map, a genomic selection approach is impossible, and the cost of generating a customised array based on the sequencing that has been done on the African buffalo genome thus far, as well as the time and cost of validating the SNPs identified, is likely to be prohibitive. In terms of BTB, genomic selection for resistance to the disease may increase resistance to other intracellular bacterial pathogens, such as *Brucella abortus* and *Mycobacterium avium paratuberculosis*, which would be an additional benefit [34]. However, variation within the immune response provides population resilience. A population with diverse immune responses to a novel pathogen has a greater chance of survival and increased adaptability. Thus care must be taken when selecting for BTB resistance that susceptibility to other types of pathogens is not increased [35].

4. Conclusions and future directions

African buffalo are a species of great importance in southern Africa, particularly in the tourism, hunting and commercial game farming industries. This creates a constant demand for buffalo, and prices have continued to increase accordingly. The market value of buffalo within South Africa varies according to disease status, with „disease-free" buffalo fetching approximately 10 times that of their counterparts at auction. However, bovine TB is prevalent in a large proportion of the country, and the maintenance of the disease by buffalo has implications for biodiversity, ecotourism and public health. The BCG vaccine has no significant effect on BTB infection in buffalo, and as such, it would be enormously beneficial to breed buffalo with increased resistance to BTB. These buffalo could then be used to re-stock reserves where BTB is endemic, and establish new herds in private game farms.

Selective breeding programmes have been successfully applied to improve numerous traits in livestock species. Advances in DNA technologies have allowed the incorporation of genetic information into the estimated breeding values of individuals, thus greatly increasing the scope for improvement and the success of these strategies. To the best of our knowledge, no selective breeding programmes for BTB resistance in buffalo have been implemented, and as such, the incorporation of any genetic information into a marker-assisted breeding strategy would be enormously advantageous. While genomic selection may ultimately be more beneficial, or provide a larger increase in genetic improvement, until such time as there is a sufficiently large reference population and/or commercial SNP chip available for the African buffalo, marker-assisted selection may be the best strategy. As the cost of genotyping and custom SNP arrays continue to decline, genomic selection may become more feasible. Future plans should look to the accurate phenotyping of a large reference population of buffalo for multiple traits, including BTB resistance.

While the breeding for BTB resistance in African buffalo offers many positive outcomes, there are some considerations which must be taken into account. As *M. bovis* is capable of infecting numerous species other than cattle and buffalo, which act as reservoir hosts, an increase in tolerance to BTB in cattle or buffalo populations is not an optimal solution to control of the disease in an ecosystem. However, it may make a considerable difference to spill-over hosts, such as greater kudu, particularly if the increased tolerance is associated with a reduction in *M. bovis* shedding and thus transmission. If a dual selection strategy that selects for both resistant and tolerant genotypes is not possible, it would be important to

consider the resulting effect on tolerant genotypes within both the breeding program population, as well as the neighbouring population of the same and/or other susceptible species.

References

1. Jolles AE (2007) Population biology of African buffalo (*Syncerus caffer*) at Hluhluwe-iMfolozi Park, South Africa. Afr J Ecol 45: 398–406.

2. Reardon, Mike (2012) Shaping Kruger. Struik Nature.

3. Cross PC, Heisey DM, Bowers JA, Hay CT, Wolhuter J, et al. (2009) Disease, predation and demography: assessing the impacts of bovine tuberculosis on African buffalo by monitoring at individual and population levels. J Appl Ecol 46: 467–475.

4. O'Brien DJ, Schmitt SM, Rudolph BA, Nugent G (2011) Recent advances in the management of bovine tuberculosis in free-ranging wildlife. Vet Microbiol 151: 23–33. doi:16/j.vetmic.2011.02.022.

5. Cleaveland S, Shaw DJ, Mfinanga SG, Shirima G, Kazwala RR, et al. (2007) *Mycobacterium bovis* in rural Tanzania: risk factors for infection in human and cattle populations. Tuberculosis (Edinb) 87: 30–43.

6. Amanfu W (2006) The situation of tuberculosis and tuberculosis control in animals of economic interest. Tuberculosis (Edinb) 86: 330–335.

7. Michel AL, Bengis RG, Keet DF, Hofmeyr M, Klerk LM, et al. (2006) Wildlife tuberculosis in South African conservation areas: implications and challenges. Vet Microbiol 112: 91–100.

8. Renwick AR, White PCL, Bengis RG (2007) Bovine tuberculosis in southern African wildlife: a multi-species host-pathogen system. Epidemiol Infect 135: 529–540. doi:10.1017/S0950268806007205.

9. Rodwell TC, Whyte IJ, Boyce WM (2001) Evaluation of population effects of bovine tuberculosis in free-ranging African buffalo (*Syncerus caffer*). J Mammal 82: 231–238.

10. De Klerk L-M, Michel AL, Bengis RG, Kriek NPJ, Godfroid J (2010) BCG vaccination failed to protect yearling African buffaloes (*Syncerus caffer*) against experimental intratonsilar challenge with *Mycobacterium bovis*. Vet Immunol and Immunopathol 137: 84–92. doi:10.1016/j.vetimm.2010.04.013.

11. Michel AL, Muller B, van Helden PD (2010) *Mycobacterium bovis* at the animal-human interface: a problem, or not? Vet Microbiol 140: 371–381.

12. Råberg L, Sim D, Read AF (2007) Disentangling Genetic Variation for Resistance and Tolerance to Infectious Diseases in Animals. Science 318: 812–814. doi:10.2307/20051510.

13. Lurie MB (1944) Experimental epidemiology of tuberculosis : hereditary resistance to attack by tuberculosis and to the ensuing disease and the effect of the concentration of tubercle bacilli upon these two phases of resistance. J Exp Med 79: 573–589.

14. Mackintosh CG, Qureshi T, Waldrup K, Labes RE, Dodds KG, et al. (2000) Genetic resistance to experimental infection with *Mycobacterium bovis* in red deer (*Cervus elaphus*). Infect Immun 68: 1620–1625.

15. Bermingham ML, More SJ, Good M, Cromie AR, Higgins IM, et al. (2009) Genetics of tuberculosis in Irish Holstein-Friesian dairy herds. JDairy Sci 92: 3447–3456.

16. Le Roex N, van Helden PD, Koets AP, Hoal EG (2013) Bovine TB in livestock and wildlife: what‟s in the genes? Physiol Genomics. doi:10.1152/physiolgenomics.00061.2013.

17. Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics. Essex, England: Longman Group.

18. Amos W, Brooks-Pollock E, Blackwell R, Driscoll E, Nelson-Flower M, et al. (2013) Genetic Predisposition to Pass the Standard SICCT Test for Bovine Tuberculosis in British Cattle. PLoS ONE 8: e58245. doi:10.1371/journal.pone.0058245.

19. Goddard ME, Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet 10: 381–391. doi:10.1038/nrg2575.

20. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009) Invited review: Genomic selection in dairy cattle: Progress and challenges. J Dairy Sci 92: 433–443. doi:10.3168/jds.2008-1646.

21. Hill WG, Mackay TFC (2004) D. S. Falconer and Introduction to quantitative genetics. Genetics 167: 1529–1536.

22. Meuwissen TH, Hayes BJ, Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819–1829.

23. Goddard M e., Hayes B j. (2007) Genomic selection. J Anim Breed Genet 124: 323–330. doi:10.1111/j.1439-0388.2007.00702.x.

24. Dekkers JCM (2004) Commercial application of marker- and gene-assisted selection in livestock: strategies and lessons. J Anim Sci 82 E-Suppl: E313–328.

25. Driscoll EE, Hoffman JI, Green LE, Medley GF, Amos W (2011) A preliminary study of genetic factors that influence susceptibility to bovine tuberculosis in the British cattle herd. PLoS ONE 6: e18806. doi:10.1371/journal.pone.0018806.

26. Sun L, Song Y, Riaz H, Yang H, Hua G, et al. (2012) Polymorphisms in toll-like receptor 1 and 9 genes and their association with tuberculosis susceptibility in Chinese Holstein cattle. Vet Immunol Immunopathol 147: 195–201. doi:10.1016/j.vetimm.2012.04.016.

27. Kadarmideen HN, Ali AA, Thomson PC, Müller B, Zinsstag J (2011) Polymorphisms of the SLC11A1 gene and resistance to bovine tuberculosis in African Zebu cattle. Anim Genet 42: 656–658. doi:10.1111/j.1365-2052.2011.02203.x.

28. Le Roex N, Koets AP, van Helden P, Hoal EG (2013) Gene Polymorphisms in African buffalo associated with susceptibility to bovine tuberculosis infection. PLoS ONE 8: e64494. doi:10.1371/journal.pone.0064494.

29. Hoal EG, Lewis LA, Jamieson SE, Tanzer F, Rossouw M, et al. (2004) SLC11A1 (NRAMP1) but not SLC11A2 (NRAMP2) polymorphisms are associated with susceptibility to tuberculosis in a high-incidence community in South Africa. Int J Tuberc Lung Dis 8: 1464–1471.

30. Ligtenberg AJM, Karlsson NG, Veerman ECI (2010) Deleted in Malignant Brain Tumors-1 Protein (DMBT1): A Pattern Recognition Receptor with Multiple Binding Sites. Int J Mol Sci 11: 5212–5233. doi:10.3390/ijms1112521.

31. Schmidt DR, Kao WJ (2007) The interrelated role of fibronectin and interleukin-1 in biomaterial-modulated macrophage function. Biomaterials 28: 371–382. doi:10.1016/j.biomaterials.2006.08.041.

32. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, et al. (2009) Invited review: reliability of genomic predictions for North American Holstein bulls. J Dairy Sci 92: 16–24. doi:10.3168/jds.2008-1514.

33. Gonzalez-Recio O, Gianola D, Rosa GJ, Weigel KA, Kranis A (2009) Genome-assisted prediction of a quantitative trait measured in parents and progeny: application to food conversion rate in chickens. Genet Sel Evol 41: 3. doi:10.1186/1297-9686-41-3.

34. Allen AR, Minozzi G, Glass EJ, Skuce RA, McDowell SWJ, et al. (2010) Bovine tuberculosis: the genetic basis of host susceptibility. Proc R Soc B 277: 2737–2745. doi:10.1098/rspb.2010.0830.

35. Van Helden P (2011) Adaptation, redundancy or resilience. EMBO Rep 12: 872. doi:10.1038/embor.2011.162.

Chapter 8

Summarising discussion

The work presented in this thesis addresses the issue of *Mycobacterium bovis* infection in African buffalo, the genetic underpinnings of host susceptibility to this disease, and the applicability of this information to the commercial and farming sectors. African buffalo are of particular interest as hosts of BTB because they functions as wildlife reservoirs of *M. bovis* within free-range game reserve settings, and it is believed that without this reservoir, spill-over hosts would not be sufficient to sustain this alien disease in the environment.

An initial investigation into the extent of the BTB epidemic in African buffalo was undertaken using data collected from Hluhluwe iMfolozi Park, South Africa (**Chapter 2**). BTB was first diagnosed in buffalo in HiP in 1986, and surveys in 1998 identified regions of particularly high BTB prevalence within the park. Since 1999, there has been a limited test and cull policy in operation in the park, involving the mass capture and testing of between 250-900 buffalo annually, and the destruction of BTB positive individuals. Test and cull policies have traditionally met with mixed success, and as such it is important to quantify changes in BTB prevalence that may be attributed to this management strategy. Data collected over a seven year period from 1999-2006 was analysed to estimate BTB prevalence at both the area and herd level within the park. Brand and capture information was used to support the validity of the area-based classifications, and prevalence changes over time were analysed with respect to the original aims of the operation, which were to reduce BTB prevalence in high prevalence areas, and to prevent the increase of disease in the more moderately affected areas. Area level prevalence for the different regions ranged from 2.3% to 54.7%, but was most commonly in the region of 10-15%. A decrease and subsequent stabilisation in the high prevalence areas suggested that a high initial prevalence may take longer to be affected by the culling of BTB positive individuals, but once a sufficient proportion had been removed, prevalence may decrease. The number of individuals captured (capture effort) and the time interval between capture events (capture interval) were significantly correlated with change in BTB prevalence – i.e. the more individuals are tested and culled, the less increase in BTB prevalence, and the greater the time period between capture events, the larger the increase in BTB prevalence. This data suggests that the test and cull programme is working in its current capacity, and that the way this programme is managed in the future might be crucial to its continued success: it may be of little benefit to perform test and cull events in a particular area if the time between events is longer than three years; likewise, a certain number of animals may need to be tested at each event in order to

effect a change. These factors should be included in the management plans for this programme in the future, in order to maximise the benefits of this strategy.

The role that host genetic variation may play in susceptibility to disease has been investigated in many different diseases and species. Tuberculosis in humans, caused by *Mycobacterium tuberculosis*, has been extensively studied, and there are many reviews regarding the role of host genetics in susceptibility to TB. Bovine tuberculosis, by comparison, has received far less attention in the genetics field, particularly in wildlife species. Before undertaking a disease association study, it is important to review what has previously been attempted in other studies, in related species or environments, and identify candidate genes that may be involved in the development of resistance/susceptibility. For this purpose, a review of the genetic susceptibility to BTB in cattle and wildlife was prepared, incorporating both heritability and association studies (**Chapter 3**). Collating both heritability and association studies of BTB in livestock and wildlife highlighted the need for further research in these areas in wildlife. Considerations that must be taken into account when undertaking genetic studies - such as phenotype classification, statistical power and population stratification – ere addressed, and suggestions made as to how these obstacles may be overcome. Finally, recommendations for future work in this area were proposed, including genome-wide association studies and investigations into more complex mechanisms such as gene-gene interactions. This review assembled the body of genetic work in BTB, and provided a foundation on which to base this research.

One of the biggest hurdles to overcome in animal genetic work is the lack of sufficient genetic data available. Model organisms, such as the mouse or cow, have fully sequenced and assembled genomes, but this is not the case for many animal species. The African buffalo is one of these species; while researchers have investigated population genetics and phylogeography in African buffalo, these studies predominantly utilised microsatellite markers and mitochondrial loci, and to date, large-scale or whole-genome information has not been available. To this end, next generation sequencing was performed on nine buffalo, in order to generate a wealth of genetic sequence information, and identify single nucleotide polymorphisms throughout the genome (**Chapter 4**). The sequencing was performed on the ABI SOLiD4 platform, and Bowtie and BWA software packages were used to map the buffalo sequence reads to the UMD3.1 domestic cow (*Bos taurus*) reference genome. The Bowtie and BWA packages identified a maximum of 2,222,665 and 276,847 SNPs within the

buffalo genome respectively, depending on the analysis method used. The BWA package combined with the Genome Analysis Tool Kit 1.1-3 (GATK) identified 76,615 SNPs within gene regions, 56,054 of which could be associated with possible consequence data using information from Ensembl. A panel of 173 SNPs were selected for validation, and BWA + GATK SNPs showed a validation rate of approximately 57%. The SNPs making up the validation panel were selected specifically with the view to utilise them in the subsequent BTB association study. This study allowed the generation of sufficient sequence information to make informed choices about candidate genes for BTB association work in African buffalo, and will serve to inform any future genetic studies across a wide range of fields.

There are many factors involved in the outcome of infection by a pathogen, one of which it the genetic make-up of the host. The host genome plays an important role in determining whether an exposed individual will remain uninfected, become infected but remain asymptomatic, or become infected and progress to active disease. In order to investigate genetic susceptibility to BTB, 69 SNPs with predicted functional consequences in genes related to the immune system were selected from those previously identified and validated in the next-generation sequencing. These SNPs were fluorescently genotyped in 868 African buffalo samples, collected from the Kruger National Park and Hluhluwe iMfolozi Park, and tested for association with BTB status (**Chapter 5**). Case-control association testing and statistical analyses identified three SNPs associated with BTB status in buffalo. These SNPs - SNP41, SNP137 and SNP144 - were found in the SLC7A13, DMBT1 and IL1α genes, respectively. The SLC7A13 gene forms part of the family of heteromeric amino acid transporters (HATs), and although this particular transporter is relatively uncharacterised, other related transporters have been associated with TB infection in humans. SNP137 is located in the coding region of the Deleted in Malignant Brain Tumour-1 (DMBT1) gene. DMBT1 is one of the pattern recognition receptors (PRRs), which recognize the conserved structures in bacterial and yeast cell walls that are vital to the survival of the invading organism. This SNP remained significantly associated after permutation testing, and was the most robust of the associations identified. SNP144 is located in the Interleukin 1 alpha (IL1α) gene, which belongs to the IL1 family of pro-inflammatory cytokines, and plays a role in mediating local and systemic inflammation. These three polymorphisms and genes are promising candidates for further exploration into genetic susceptibility to BTB, and may also hold potential for marker-assisted breeding programmes in African buffalo. There will, of

course, be many other genes and variants that contribute to the host genetic component of BTB susceptibility, and should be investigated as time and finances allow.

Despite the numerous benefits of next-generation sequencing, there are a few notable limitations, particularly when mapping to the reference genome of a different species. One of these limitations is that only a percentage of the sequenced genome is represented in the final data. There are a number of reasons why a particular region might not map correctly or have sufficient coverage – regions containing many repeats are difficult to map unequivocally, and there may be too many differences in the sequence identity of the reference and sequenced genomes at a particular locus for correct alignment to take place. As such, there may be little or no sequence information obtained on specific regions or genes of interest. One of the most commonly investigated genes in bacterial disease susceptibility is NRAMP1 (SLC11A1), and it has been associated with susceptibility to salmonella, brucellosis, and paratuberculosis in numerous animal species. Very little informative sequence data was obtained from the SOLiD4 next-generation sequencing of the buffalo for this gene. To compensate for this lack of data, Sanger sequencing was performed to more thoroughly characterise the sequence, structure and variation of the NRAMP1 gene in African buffalo (**Chapter 6**), with the intention of utilising the identified polymorphisms in future BTB susceptibility work. Fifteen novel single nucleotide polymorphisms and three microsatellites were identified within the NRAMP1 gene. The predicted protein structure was similar to that of the cow, and displayed the conserved elements characteristic of this family of proteins. Phylogenetic analysis of the amino acid sequence of the protein grouped the African buffalo with water buffalo. Of the variants identified, 5 SNPs were found in the 5‟ and 3‟ untranslated regions, 9 SNPs were intronic, and one was exonic. While the exonic SNP was not predicted to cause an amino acid change in the NRAMP1 protein, variants within the 5‟ UTR can have a significant impact on gene expression, and thus the variants identified in this region may be excellent candidates for future genetic association work.

Investigations into host genetic susceptibility seek to identify genetic variants that play a role in host resistance, in order to utilise these genetic markers either as potential drug targets or vaccine candidates, or for selective breeding purposes. While the BCG vaccine is highly effective in some animals, no significant protection has been shown in African buffalo after experimental challenge with *M. bovis*. The ineffectiveness of the BCG vaccine, combined with the logistical concerns of administering a new drug or vaccine to a free-ranging wildlife

species such as the African buffalo, suggests that the primary utility of identifying genetic variants involved in susceptibility to BTB in African buffalo may lie in the potential for selective breeding. Selective breeding has been employed for numerous production traits in cattle, such as milk yield and age of first calving, as well as for diseases such as mastitis. However, there are many aspects of this type of control program that must be understood and taken into account, particularly when performed on a different, wild and free-ranging species. To more thoroughly clarify some of these aspects, a review was compiled to address the applicability of selective breeding to African buffalo (**Chapter 7**). Phenotypic and genotypic approaches to selective breeding were evaluated, and the role of heritability studies in breeding programs and the type of information required were addressed. Consideration of these aspects in relation to African buffalo, and what has been achieved in cattle, suggests that selective breeding for BTB resistance in African buffalo is a viable strategy. Presently, the inclusion of genetic loci in a marker-assisted approach would improve on a traditional phenotypic approach, and as the cost of genotyping continues to decrease, the incorporation of additional genetic loci or genome-wide information would further enhance the potential of breeding BTB resistance in African buffalo.

In the course of this study, a vast amount of sequence data on the African buffalo genome was generated, through next-generation and Sanger sequencing, and has identified upward of 75 000 SNPs within gene regions. This is the first time that next-generation sequencing technology has been applied to the African buffalo, and the data generated can be utilised in many future studies across multiple fields. Our candidate SNP set was by no means exhaustive – many other genes and SNPs are likely to influence BTB susceptibility in African buffalo, and should be investigated further as time and finances allow. Furthermore, it may be beneficial to determine the level of success that could be achieved using a commercial cattle SNP chip for genotyping African buffalo. As technology progresses and costs decrease, this could be the most informative and cost-effective way to investigate the full complement of genetic variance involved in BTB susceptibility in African buffalo, provided that sufficient data can be obtained using cattle SNPs. And finally, the utility of the information gained from genetic association studies must be considered. With the lack of protection conferred to buffalo by the BCG vaccine, it seems likely that the primary utility of genetic association information will be in the incorporation of this information into breeding values, for selective breeding of animals with increased BTB resistance. This strategy has particular application to newly established game farms looking to introduce buffalo, or in the re-stocking of farms in

BTB endemic areas, in order to limit the maintenance and spread of BTB within these reserves, and minimise the ecological and economic impacts associated with this disease.

Bovine TB is a chronic, infectious disease that threatens the welfare of people, livestock and wildlife. In countries like South Africa, with vast human-livestock-wildlife interfaces, it is imperative that sufficient resources are committed to understanding the different aspects of BTB infection, and that new methods of control and management are constantly under development. Host genetic studies form a vital part of the research needed to inform a new set of wildlife BTB management options, so that the impact and consequences of *M. bovis* exposure may be limited in the future.

Appendix A

Contribution to Publications

Chapter 2 Bovine tuberculosis in African buffalo: estimating BTB prevalence and evaluating the test and cull method in Hluhluwe iMfolozi Park, South Africa.

- First author
- Conceptualisation and planning of study
- Data compilation
- Analysis and interpretation of data
- Writing of manuscript

Chapter 3 Bovine TB in livestock and wildlife: what's in the genes?

- First and primary author
- Writing of manuscript

Chapter 4 Novel SNP discovery in African buffalo, *Syncerus caffer*, using high-throughput sequencing.

- First author
- Analysis and interpretation of data
- Writing of manuscript

Chapter 5 Gene polymorphisms in African buffalo associated with susceptibility to bovine tuberculosis infection.

- First author
- Conceived and designed the experiments
- Performed the experiments
- Analysis and interpretation of data
- Writing of manuscript

Chapter 6 The NRAMP1 gene of the African buffalo: characterisation and variation.

- First author
- Conceived and designed the experiments
- Performed the experiments
- Analysis and interpretation of data
- Writing of manuscript

Chapter 7 Breeding TB-resistant African buffalo: the future of BTB management?

- Joint first author
- Writing of manuscript