An investigation into the diversity of genes of the innate immune system in the European Badger (*Meles meles*) and possible associations with trypanosome infection.

Volume 1 of 1

Submitted by Andrew Mason Whiteoak

For the Qualification of Master of Science by Research

2015

The University of Salford

School of Environment and Life Sciences

Table of Contents

Acknowledgements

I would like to offer my deepest gratitude for the help given by the following people. My supervisor, Geoff Hide; always eager to help, encouraging and all together a great friend as well as advisor. Working alongside me in labs, my good friend Eze Justin Ideozu. I found in Justin a like mind, and a passionate scientist. Members of the technical team at the University of Salford have helped me with all aspects of my research; sourcing the best deals on lab supplies; providing, maintaining and advising on equipment; and giving expert scientic advice; all the time showing an obvious love for what they do - Belgees Boufana, Helen Bradshaw, Tony Bodell, Cath Hide, Sally Shepherd and Arlene Armstrong. Alex Tomlinson and Richard Delahey at Woodchester Park provided the majority of samples used in this research. Pauline Kidner, Elizabeth Mullineux and Sara Cowen at Secret World Wildlife Rescue provided blood samples taken from badger cubs rescued during the 2013 Somerset floods and from cull zones across the country. Eamonn Gormley from University College Dublin, provided samples from his on-campus badger cohort. The Leicestershire and Rutland Badger Group along with the family of the late Bill Cunnington were very kind in awarding me a substantial grant which allowed me to attend the 2014 Badger Trust Conference. Finally, I would like to thank my girlfriend, Jamie Angus, for her encouragement and great home cooking, and her children, Hazel and Mike, for giving me such a wonderful family.

This thesis is dedicated to the memory of Seamus 'No Second Troy' Black 17.01.03 - 24.01.14

Abstract

Toll-like receptor (TLR) genes encode for conserved proteins of the innate immune system which trigger pathways in response to pathogen invasion by recognising molecules essential for endoparasite survival. An endoparasite lives within the host and relies upon the host for its nutrition, here we consider it a pathogen when its presence causes detriment to the host. A pathogen is broadly anything which causes disease; causing specific symptoms in a specific location – in this case the symptoms would be the innate immune response to the parasite at its location, which would inevitably divert host energies to the creation of innate immune molecules. TLR genetic variation is rare and small mutations, such as single nucleotide polymorphisms (SNPs), can be correlated with enhanced disease resistance in some hosts. This research investigates TLR sequence diversity in European Badger (*Meles meles*) populations across the UK and possible associations with trypanosome infection. DNA from a collection of badgers from Woodchester Park, and other UK sites were available for investigation. As no badger TLR sequences were available or published, five PCR primer sets were successfully designed, using closely related species, to amplify the full badger TLR2 exon and exon 3 of TLR4. Sixty-one and fifty-nine badgers were sequenced across TLR2 and TLR4 (exon 3) respectively. Three TLR2 amino acid haplotype variants were found with two linked and one unlinked. No polymorphisms were found in the same badgers for TLR4. There was no significant relationship of polymorphism with trypanosome infection. Badgers ranging from St. Ives to Birmingham were found to be highly homogeneous with respect to TLR DNA sequence variation, and such low level variation was highly unusual in comparison to other TLR population studies.

1.0 Introduction

Variation in susceptibility to parasites has been linked to variation in Toll-like Receptor (TLR) genes involved in their recognition, as has been suggested by Morger et al (2014). TLRs are highly conserved, however some significant variations are known to occur. Cargill and Womack, 2007, found 123 novel SNPs across four TLRs differentiating nine breeds of cow. And, Tschirren et al (2013) demonstrated a correlation between TLRs and *Borrelia afzelii* susceptibility. Reismann (2009) demonstrated a link between TLR4 polymorphisms and hyporesponsiveness to inhaled LPS, as well as increased incidence of gram negative infections and bowel disease. Further links have been suggested between TLR variations and risk of cancer (El-Omar et al 2008; Gomaz et al 2012), diabetes (Liu et al 2012), asthma (Schwartz and Cook, 2005) and tubercolosis (Zhang et al 2013). TLR variation has not yet been studied in Eurasian badgers (*Meles meles*). To begin this study, we had available to us, a population of badgers characterised for trypanosome infection. This project aims to investigate TLR sequence variation in badgers and, if there is any variation, to investigate whether this is related to *Trypanosoma pestanai* infection in the British badger population.

1.1 The European Badger *Meles meles*

1.1.1 Badgers and their role as Reservoirs for Infection

The badger innate immune system has unique differences to that of other mammals (Higgins, 1985). For a wide variety of antigens, antibody response in badgers is poor in comparison with other mammals, producing profoundly low lymphocyte levels (Mahmood et al 1987). *M. bovis* was the first antigen shown to cause recognisable lymphocytic reaction in badgers. Recent studies show a robust IgG response in badgers (Lizundia et al 2011,) and a delayed hypersensitivity reaction to *Mycobacterium bovis* in which eosinophils and basophils are recruited. Eosinophil chemotactic factor is produced by T cells in badger epithelial layers in two phases about six and twenty-four hours after infection by *M. bovis* (Higgins, 1985). At this stage, little is known about the receptors of the innate immune system, such as Toll-Like Receptors, in badgers.

A current topic of interest, in light of the 2013 badger culls, is the concept of badgers as a reservoir for tuberculosis (TB) infection in British cattle. TB is progressive and ultimately fatal respiratory disease caused by *Mycobacterium bovis*. Pathology includes granuloma and

lesions (Allen et al 2010). *Mycobacterium sp.* enter the body through the respiratory route before the bacilli join the blood, transporting them to the vital organs; lungs, kidneys, liver and intestines. Pale, spherical tubercles, or lesions, form within these organs. It is very difficult to detect in the early stages. Rare clinical signs may include lethargy, emaciation, fever and pneumonia.

Following the domestication of cattle, it is speculated that *Mycobacterium tuberculosis* jumped species, becoming *M. bovis*, a cattle-based form of the disease. In the 1930s, one third of cows in Europe suffered from bovine TB (Huard et al 2006). The bacteria can linger in water, animal feed and soil (Velayati et al 2015). It is also a zoonotic disease and at its height in Britain in the twentieth century there were 50,000 human cases (Barkham, 2013). Historically, bovine TB is believed to have been a contributor to human TB cases worldwide, and it remains a concern in both developed and developing countries (Biek et al 2012).

The case of badgers and their relationship to bovine TB is widely debated. Badgers are widely known to be reservoir hosts for bTB (Donnelly and Hone, 2010; Weber et al 2013). However, arguments occur in relatively recent publications that suggest instead that they may simply be spillover hosts (Balseiro et al 2011). Their contribution to bTB persistence is still not clear (Karolemaeas et al 2012). Previously, wild animals were generally accepted as incidental, non-significant victims of the disease (Bengis and Erasmus, 1988). However, wild animals worldwide have since been implicated as TB reservoirs; for example, brushtail possums (*Trichosurus vulpecula*) in New Zealand and white-tailed deer (*Odocoileus virginianus*) in the USA (O'Brien et al 2011) and wild ungulates in Africa (Gortazar et al 2008). Studies of badgers have found that TB is spread amongst members of the sett via bite wound infections. This is a rare occurrence due to the low prevalence (6.6%) of heavily infected badgers (Balseiro et al 2011). Infectious late stages occur when culture positive badgers fail to return to the latent state. This occurs significantly more often in male badgers than female badgers, with infective male badgers having significantly greater bite wounding than non-infective male badgers. In females, an immunocompromise such as pregnancy or extreme age are the contributing factors towards infective stages (Wilkinson et al 2000). Although gene analyses of *M. bovis* isolates in badgers and their nearest cattle populations provided evidence for recent transmissions between the two hosts, directionality of transmission could not be

inferred (Biek et al 2012). TB tests on cattle have been prone to unreliability over the years. *M. bovis* shares many antigens with other mycobacteria including *M. avium*, which are commonly found in the environment, often confounding diagnostic tests (Allen et al 2010). Transmission may occur via environmental contamination, contiguous contact between farms, or the unrestricted network of livestock movements across the country (Biek et al 2012). Since mid-2013 cattle movement restrictions have been in place to ensure identification and traceability of all cattle as part of measures for disease control and to maintain consumer confidence in farm produce (Draper, 2013).

Badger communities are very stable, staying in one location over generations for centuries. Under natural conditions badgers will live in a single location without moving or expanding borders. A cull causes a perturbation effect where boundaries to break down and leads to increased transience when home females are lost (Lesellier et al 2006). At low densities badger social structure becomes fluid. Long distance movements increase with decreasing population. Territories become large and ill-defined (Byrne et al 2012). Males are territorial, and aggressively defend home setts from transients seeking mates. This accounts for heightened aggression, bite wounds and poor health of neighbouring badgers following culls (Stewart et al 1999).

A study by Weber et al (2013) found that TB infected animals tend to be transients; those animals less well connected to their own groups than uninfected badgers, but important in linking the flow of infection between groups. Culling is based on an assumption that transmission is density dependent and that a population density threshold exists below which the disease cannot persist. This assumption is supported by field evidence for exponential relationships between host density and disease transmission for bTB infection in the New Zealand brush tale possum (Carter et al 2007). Long-term data gathered by Cheeseman's team at Woodchester Park shows there is no link between the density of badgers and prevalence of TB (Karolemaeas et al 2012). The animal most likely to pass TB to a cow is another cow; the disease has destroyed entire herds on the Isle of Man, where there are no badgers (Skuce et al 2011). The Manx government now have strict regulations in place to monitor the import of cattle (Department of Environment, Food and Agriculture, 2014).

For badgers living in areas where there is high incidence of cattle with bTB, there is a risk of infection, but also a proportion of badgers will develop immunity (Gallagher and Clifton-Hadley, 2000). The badger cull creates a vacuum, into which badgers from outside setts will arrive. New arrivals are naïve to the disease, and thus a higher percentage of badgers become infected (Kidner, 2014). Culling increases bTB prevalence in neighbouring setts, and increases immigration into culled areas (Byrne et al 2012). Outside culled areas range sizes increase with proximity to the cull area, revealing a gradation of social disruption. Following culls repopulation takes less than five years (Pope et al 2006). During this time no clear latrine boundaries exists. Culls also cause sufficient stress to individual badgers to induce immunosuppression and heightened susceptibility to infection (Riordan et al 2011). TB has consistently been seen to increase in badgers post culling (Carter et al 2007).

In the UK, badgers were culled between 1973 and 1998 with the aim of reducing infection in cattle. Since the mid-1980s the incidence of infection has continued to increase, particularly in southwest England. The Randomised Badger Culling Trial was conducted in England between 1998 and 2005 and overseen by the Independent Scientific Group (ISG). Bielby et al 2014 found that culls during this trial exacerbated the perturbation effect. Herd breakdowns directly affect farmers as quarantines are imposed while cattle movement and the sale of meat and dairy products are restricted. Farmers are required to produce to completely clear tests for every cow before restrictions are lifted. The skin test is inaccurate (Al-Orainey, 2009). It fails to correctly detect an infected cow in over 25% of cases (Grabau and Novick, 1995). These undetected carriers of disease may be responsible for many cases of cow to cow transmission (Barkham, 2013). The 2013 cull involves widespread proactive slaughter. Proactive culling results in a 23% decrease in disease incidence; however this is accompanied by a 25% increase in incidence in a 2km wide buffer zone (Carter et al 2007).

Clifton-Hadley et al 1993 performed an 18 year study which showed TB to be rare in badgers. Only one sixth of over 300 badgers tested over the period were positive for TB, and two thirds of these were clear of infection prior to subsequent recapture; the badgers immune systems were able to clear the infection independently. In the Republic of Ireland trials of the Bacillus Calmette-Guerin (BCG) vaccine are underway (e.g. County Kilkenny as of 2011).

(Byrne et al 2012). In 2011, the Gloucestershire wildlife trust launched the first ever independent bovine TB vaccination programme for badgers. Field trials found that an injectable vaccine, BCG, based on the human version, reduced the risk of badgers testing positive by 54% (Carter et al 2012). Even higher success rates for vaccination trials were also reported by DEFRA (Veterinary Laboratories Agency, 2010) which found a 74% reduction in badgers testing positive for TB, over 800 badgers vaccinated. And, in 2012 the Welsh Assembly established a five year badger vaccination project, vaccinating 1424 badgers in 2012 and 1352 in 2013; analysis of project success is to come at the end of the trial in 2017 (Llywodraeth Cymru, 2014). As well as these several NGOs are participating in badger vaccination programs, including the RSPB, the Badger Trust and the National Trust. Vaccinating badgers with BCG would be a long term cost-effective alternative strategy for control of bovine TB. TB is a disease of poverty, in humans as well as animals (Grange and Zumla, 1999). As well as overcrowding, muddy yards, poorly ventilated sheds, irregular cleaning of water troughs, (Hybu Cig Cymru report, 2013) British cattle are of poor immune health due to inbreeding via selective breeding (Oltenacu and Broom, 2010). A range of afflictions; BSE, foot and mouth, dermatitis as well as TB are a reflection of genetic poverty in modern farms.

1.1.2 Badger Evolutionary History

Badgers are Mustelids, of which there are 59 extant species, which originate from Eurasia, and share a common ancestor. They vary between taxa that are fossorial, semi arboreal and semi aquatic. Their diets vary from specialization on rodents to piscivory. Mustelidae are divided into five sub families based on phylogeny: Melinae (hog nosed and Eurasian badgers), Lutrinae (otters), Mellivorinae (honey badger), Mephitinae (skunks and stink badgers) and Mustelinae (martens and weasels). Of all Mustelids, the Melinae, Arctonyx and Meles are the most variable taxa (Koepfli et al 2008). Badgers are one of the most widely distributed Mustelid found across the Palaearctic (Pope et al 2006). Badgers colonized Britain twice, first during the Pleistocene glaciations over 15,000 years ago, returning in the current interglacial, the Holocene that began 10,000 years ago (Fairnell and Barrett 2006). These badgers have homologues in Central Europe, with Norway marking the northern edge of their distribution. Britain was colonised by a brief land bridge across the Baltic Sea, subsequently isolated from additional gene flow (Pope et al 2006). Irish and Scandinavian badgers originate from Spain,

having arrived via human introduction around the seventh century. Unlike England, Ireland has one single unified badger population. Genetic analyses have shown that these two gene pools have not mixed, giving completely separate genetic populations (O'Meara et al 2012).

British badgers have historically endured a pattern of repeated colony disruption. Badgers have been killed, transported and disturbed for the past millennia, and yet continue to overcome adversity. In the early Bronze Age (6th to 7th centuries A.D.) badgers were transported from Britain to continental Europe as trophies or pelts (Fairnell and Barrett 2006). In seventh century Ireland the badger was regarded as a valuable both as a pet and as food. Law tracts from the time show regulations regarding compensation for the bites of a neighbours' pet badger, and badger meat was preserved in salt as food, and the fat used for medicinal purposes. Bone evidence has traced their presence and use as food in Ireland through the tenth and eleventh centuries (Benecke, 1999). In eighteenth and nineteenth century Industrial times, badgers were translocated to areas where there were no foxes to dig setts to encourage them (O'Meara et al 2012). At the start of the twentieth century, badger numbers were reduced across much of continental Europe through the gassing of setts for rabies control (Pope et al 2006).

Today, contemporary habitat fragmentation has had the largest impact on dispersal and population structure. At present, several distinct populations exist in Britain. Distinct family clusters are found in N. Scotland, S. Scotland, Carlisle, N. W. Wales, S. Wales, Ipswich and Woodchester, N.E. to N. Wales, and from Middlesbrough to Scarborough to Leeds (Pope et al 2006). Even small sample areas have been found to show complex genetic variations. Yung et al (2011) found four different alleles in the only seven individuals in one sett.

1.1.3 Badger Ecology

The British badger population is behavioually unique in comparison with its continental counterparts (Hunford, 2013). In Europe they are seen as largely solitary animals, but the more compact home ranges and patchy resources available in Britain have been overcome by a facultative sociability that sees multiple family groups sharing small spaces in an unexpectedly communal manner (Pope et al 2006, and, Johnson et al 2001). This may be in part due to the badger's very specific preferences for home sites. The steepest hillsides are chosen, at low altitudes, with soils suitable for good drainage and ease of digging (loam, sand or clay,) on a side away from the prevailing winds (Stewart et al 1999). Each social group has an average nine shelters per territory, which allow badgers to avoid expending energy on long distance returns to the main sett each night (Hounsome et al 2005). The largest number of shelters are used in summer, when male badgers will also visit neighbouring setts in search of mates. At this time, males are spread more widely across the various setts, while females with young stay in the main sett. Males also range within the setts from chamber to chamber far more than females (Kowalczyk et al 2004).

Badgers have a long mating season, delayed implantation and superfoetation (Yung et al 2012). In England badgers produce offspring in January-February while Irish females begin in February-March (Byrne et al 2012). They are polygynandrous, moving from chamber to chamber within the sett every few days, constantly changing their sleeping arrangements (Stewart et al 1999). Roving badger males allow for a greater genetic diversity than would be attributed to an otherwise non-dispersal oriented species. Microsatellite analysis has shown the most badger young are fathered by transients from other families (MacDonald et al 2008). The badger flea, *Paraceris melis*, identified as the agent for badger trypanosome (*Trypanosoma pestanaii*) transmission, will therefore experience a wider variety of new hosts due to roving badgers (Stewart et al 1999). In winter one main sett is used by all family members. At this time all badgers will be found in one small section of the main sett, while the remainder of that sett provides a refuge for other species including foxes (*Vulpes vulpes*), rabbits (*Oryctolagus cuniculus*), wood mice (*Apodemus sylvaticus*), bank vole (*Myodes glareolus*), pine marten (*Martes martes*) and otters (*Lutra lutra*) (Byrne et al 2012).

Despite their communal nature, foraging is an individual pursuit (Johnson et al 2001). Badgers are opportunistic and predate on lagomorphs and hedgehogs, amphibians, insects, molluscs, annelids, roots, and fruits (Byrne et al 2012). Badger movements are tied to latrine network, which is marked by clearly visible landmarks, such as large trees or road verges (Hounsome et al 2005). Latrines are a means of communication with members of the same and different social groups, communicating olfactory information such as foraging success and resource depletion. While the network marks out separate setts, it also provides a path which all badgers in adjacent groups will follow and interact upon while foraging (Rey et al 2010). British badgers are found in high density communities, and so display a strong pattern of isolation by distance. This is a regular increase in genetic differences between individuals with geographical distance due to limited dispersal. Larger rivers have been shown to produce barrier effects, creating genetically unique sets of badger families (Frantz et al 2010 b). Badgers are not much impeded by smaller natural barriers, such as rivers less than 50m wide, nor by by motorways, However, their territorial boundaries are fairly permanent. Over an eighteen year study, Clifton-Hadley et al (1993) found boundaries remained stable, provided that the badger populations remain undisturbed. MacDonald et al (2008) found that 40% visited only two social groups while 35.8% stayed permanently within their natal social group. Most activities are confined to a six km^2 range about the main sett.

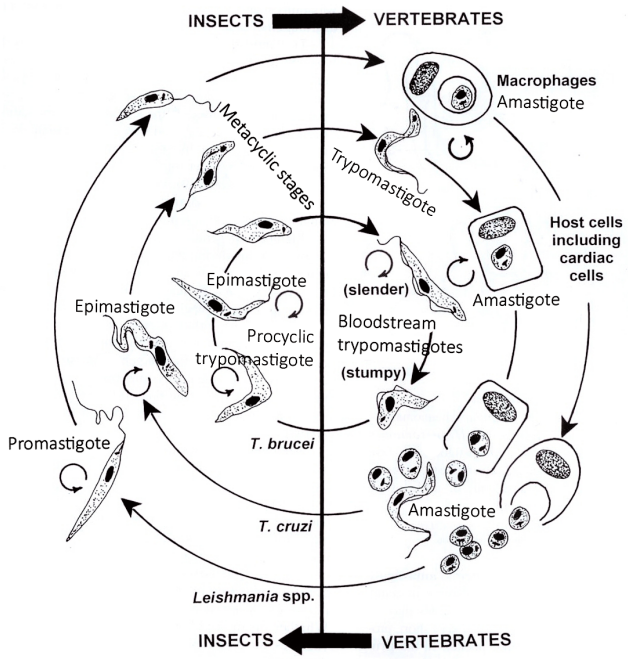
The largest social group in Europe exists at our study area, Woodchester Park, it is observed that the large size of setts at this location allows females to avoid interference while breeding (Byrne et al 2012; Balseiro et al 2011). It might be infered that this would allow for better genetic mixing than at smaller study sites, and the possibility of positive selection of genetic characteristics among this population.

1.2 Trypanosomes

1.2.1 Trypanosome Infection in Humans

Trypanosomes are small protozoan parasites of the genus '*Trypanosoma'*. They diverged early in evolution, forming one of the first groups of eukaryotes (Lima et al 2010). They have evolved to infect all land based mammals including humans. There are three distinct types of human infective trypanosomatid; the stercorarian trypanosomes, such as South American *Trypanosoma cruzi*, the salivarian human African trypanosomes and the leishmanids. All are blood borne parasites that follow a two-host lifecycle between an blood-sucking insect intermediate host and a definitive host mammal. A generalised lifecycle is shown in Figure 1. human African trypanosomiasis, chagas disease and leishmaniasis are all 'Cinderella' diseases; they affect the world's poorest people and thus are of little interest to drug companies (Simarro et al 2010). Trypanosmatid lifecycles begin with ingestion of the parasite in bloodmeal consumed by the vector; for human African trypanosomes the tsetse fly (*Glossina sp.*) and for human S. American trypanosomes the triatomine bug (*Triatoma sp.*). Both are ingested as trypomastigotes, while leismania species are ingested as amastigotes by the sandfly (*Phlebotomus sp.*) (Bates, 2007). The next stop is the respective insect's midgut where the majority are lysed (De Souza et al 2010).

Figure 1. Lifecyles of the Trypanosomatids, adapted from Brigand et al (2006).



The increase in pH and decrease in temperature in the vector midgut trigger the next stage of development (Kamhawi, 2006). Surviving parasites transform to procyclic trypomastigotes (HAT,) epimastigotes (SAT,) or nectomonad promastigotes (leishmania) (Minning et al 2009). Next, all undergo division, where replication occurs due to binary fission, after which parasites leave the midgut. The ability to remain attached rather than dispersed with faeces is key to this stage, and is mediated by surface lipophosphoglycans (LPGs) (Sacks and Kamhawi, 2001). HAT (now an epimastigote) and leishmania head towards the salivary glands (Wheeler et al 2011,) while SAT move into the hindgut (Parikh et al 2012). Adhesion of *T. cruzi* in the hind gut acts as a trigger, transforming the non-infective epimastigotes to highly infective metacyclic trypomastigotes (Nogueira et al 2007). This attachment is mediated by surface anchored glycoinositolphospholipids, which are dependent on low sugar levels. The presence of most simple sugars inhibits this attachment (Tyler and Engman, 2001).

From here the *T.cruzi* is either in faeces which is then injected into the bloodstream by the triatomids' habit of taking meals from spots on which it has recently defecated. Or in the case of HAT and leishmania invasion takes an oral route. HAT transform to a metacyclic trypomastigote once in the salivary gland and are injected during a blood meal. For leishmania, transmission is via regurgitation of a mucus plug secreted by promastigotes blocking the sand fly's anterior midgut, which forces it to regurgiate before it can take a blood meal. The sandfly saws into the skin and feeds from pooled blood (Lane, 1993) this draws up the macrophages needed for leishmania's invasion of the definitive host which the parasite will then use as a taxi-ride to deeper organs.

Once in the human the passage of each varies widely. HAT transform to bloodstream trypomastigotes and are carried to other sites, avoiding the immune system via antigenic variation and multiplying by binary fission in various body fluids before re-entering the bloodstream. *T. cruzi* promastigotes are phagocytized by macrophages (De Carvalho and De Souza, 1989) they then lyse the vacuole using porins triggered by acidity from the host lysosomes and break into the cytoplasm of the cell where transform into amastigotes and multiply to form a pseudocyst (Andrew, 1993). *T. cruzi* also multiplies in cells of various tissues to which they arrive by the macrophage as transport. Leishmania transform to amastigotes, which fool the host immune system by mimicking host molecules. They multiply

by binary fission in infected tissues before transforming to trypomastigotes and re-entering the bloodstream. Once back in the bloodstream parasites can once more be ingested by the intermediate host.

Subspecies of *Trypanosoma brucei* cause sleeping sickness (Hide et al 1991) affecting over three million people in Africa (De Souza et al 2010). In east Africa, infection with *Trypanosoma brucei rhodesiense*, results in acute disease and death within months (Hide et al 1990). A brief recurring fever, headaches and joint pain, precede the parasites' invasion of the circulatory system (Nadjm et al 2009). The haemolymphatic phase is diagnosed by the appearance of a Winterbottom's Sign; swollen lymph nodes on the back of the neck (Sinha et al 1999). Left untreated, anaemia, cardiac, and kidney dysfunctions develop. In advanced stages, parasites invade the central nervous system (MacLean et al 2010). Sleep is disrupted, and irreversible neurosis ensues, marked by confusion, tremor, general muscle weakness or paralysis. Patients present with psychotic reactions, aggression or apathy which can confuse diagnosis. This can progress to coma, organ failure (Paul et al 2014) and death. In west Africa as well as the Congo, Central African Republic and Angola, *Trypanosoma brucei gambiense* causes a chronic form of the disease, which lingers for years before death, and is asymptomatic until its advanced stages (Tiberti et al 2014). Both subspecies have distinct, non-overlapping geographical distributions (Echodu et al 2015). The main mode of transmission to humans is bite of a tsetse fly (*Glossina sp.*) but parasites can also pass from mother to infant and through blood contamination (Norman and López-Vélez 2014).

Humans have developed defensive mechanisms in the arms race against trypanosomes, namely Trypanosome Lytic Factors (TLFs) (Thomson et al 2009). TLFs are powerful toxins released when Human serum binds with high affinity to haptoglobin-haemoglobin receptors in the trypanosome flagella pocket. Thirty minutes after exposure to these toxins, non-resistant trypanosomes become swollen and immobilised. In *T. b. rhodesiense*, the single protein, Serum Resistance Antigen (SRA), confers resistance to TLFs (De Greef and Hamers 1994). SRA not occur in procyclic forms which occur in the insect, but only in bloodstream forms. Sensitivity to human serum provided the first means of distinguishing human-infective *T. b. rhodesiense* and non-human-infective *T. b. brucei* (Hide, 1999). The sequence of the SRA shows it is structurally related to Variable Surface Glycoproteins (Hide, 1999). *Trypanosoma*

brucei brucei does not have SRA, but instead has evolved avoidance mechanisms, such as down regulation of their haptoglobin-haemoglobin receptors (Stephens et al 2012). T*. b. gambiense* uses antigenic variation to constantly evade the immune system, resulting in waves of parasitaemia (Kennedy, 2005). Antigenic switching is controlled by just one major protein, which manages a repertoire of over a thousand protein coats (Hide, 2003). The parasite simply changes coat as the immune system begins to recognise and kill it, and new strains build up once more.

Post colonial independence in Africa led to political instability and economic ruin which reduced population screening, stalled progress of treatment, and caused an epidemic surge of trypanosome related disease by the 1970s (Steverding, 2008). Drug companies took little interest in a disease with no profit making enticements. A drug was not found for the west African form until 1990 (Simarro et al 2011). A key moment in the fight against Sleeping Sickness was the application of the Restriction Fragment Length Polymorphism technique to parasite identification (Hide et al 1994). This differentiated morphologically identical and genetically similar infective and non-infective sub-species for the first time (Welburn et al 2001). Since 2001, a move prompted by WHO, toward the free supply of medication, tsetse fly trapping and release of sterile males, and a 2006 initiative to develop better diagnostic tools all aim at one day eradicating Sleeping Sickness (Steverding, 2008).

The core reservoir for *T. b. rhodesiense* infection is cattle (Hide et al 1996). Wild animals in Africa, such as buffalo, elephant, hippo, rhino and warthogs maintain a reservoir for trypanosome infection in the environment (Wambwa, 2003). However the spread of the East African strain has been strongly linked to cattle movements (Hide et al 1998). Fevre et al (2001) highlighted this link by showing an outbreak which started in 1998, in the Soroti district of Uganda, was more prevalent with proximity to a cattle market which sold cattle originating from endemic areas. Molecular and epidemiological analysis has shown that people are five times more likely to acquire sleeping sickness from cattle than another human (Hide and Tait, 2009). Trypanosomes have a profound effect on domestic and farm animals, causing severe, often fatal disease, while by contrast in wild animals infection is relatively mild (Steverding, 2008). Disease in cattle is largely due to *Trypanosoma congolense* (54.3% of cases) and *Trypanosoma vivax* (45.7% of cases), with a minor contribution from *T.*

brucei (Bitew et al 2010). 3,000,000 cattle die every year from related infections, which has a significant impact on rural communities that rely on them. Trypanosomiasis removes people from caregiving or food production roles in the community, causing them to become a burden to their peers for far longer than diseases such as leishmaniasis: Despite putting relatively fewer people at risk (55,000,000 in 36 countries vs. 350,000,000 in 88 countries for leishmaniasis) the amount of Disability Adjusted Life Years lost is of significant magnitude (1,780,000 vs 2,060,000) (Hide, 1999). Many researchers suggest restriction of cattle movement (Finelle, 1972) or vaccination of cattle as a way forward in the fight against East African sleeping sickness (Maudlin et al 1990; Cox et al 2010).

Trypanosoma cruzi confers Chagas disease, affecting 16-18 million people in South America. It is the largest parasitic disease burden on that continent causing around 50,000 deaths per year (Savino et al 2007). Early stages are characterized by swelling of lymph nodes. In children, the Romana's Sign, a swelling around the eye, is characteristic in early stages (Epting et al 2010). One in three carriers develop the chronic condition, which can persist without symptoms for over thirty years (Clayton, 2010). *T. cruzi* remains hidden from the immune system inside cardiac muscle. Heart damage results from the persistence of pro-inflammatory cytokines (Rodriguez et al 2012). Intracellular replication of the parasite leads to the cell bursting (Kierszenbaum, 1999) which acts as a triggering event, whereas T and B cells are the effector cells in pathogenesis (Teixeira et al 2002). Asymptomatic patients die of heart failure due to a build-up of pseudocysts in the heart muscle and a sudden immune response (Ropert et al 2002).

Tissue resident macrophages are critical targets for early infection. Once encased in the acidic vacuole, the parasite is protected by surface trans-sialidases (Epting et al 2010). Sialic acid residues metabolised from the host are attached to the parasite cell surface by Glycosyl Phosphatidyl Inositol (GPI) anchored glycoproteins (Ropert et al 2002). These are responsible for the negative electrical charge which allows the parasite to attach to host cells (Kubarenko et al 2010). More pathogenic strains are more strongly negatively charged (Cuervo et al 2010). The flagella is positively charged, which allows the parasite to avoid host immune cell attachment (Souto-padron, 2002).

T. cruzi is transmitted by *Triatoma infestans* (Dias, 2007), which live in the wall cracks of dwellings, dropping out of roof spaces at night to feed on human hosts. Animal reservoirs include bats (multiple species), opossums (*Didelphis virginiana*), dogs (*Canis lupus familiaris*), and rats (*Rattus sp.*) (Clayton, 2010). *T. cruzi* is transmitted in a stercorarian manner. Infected vectors defecate as they feed then rotate 180^{o} biting the host a second time through freshly laid infected faeces (Lima et al 2010). Macrophages phagocytose the parasite, which fools the host cell into mistaking it for a self molecule (Davies, 1997). It then survives to reproduce within the cell, using it as transport to invade other organs while avoiding host defences, infecting heart muscle, and the reticulo-endothelial system. 70% of all cardiac deaths in young adults in South America occur due to *T. cruzi*. There are no vaccines or chemotherapies (Kierszenbaum, 1999) and just one anti-parasitic efficient in the acute-phase of disease, Benznidazole (Savino et al 2007). Pyrethroid insecticides were introduced in the 1980s, and closer monitoring of bloodbanks following the emergence of AIDS. Scientists soon realised the problem of fighting Chagas disease lay not with finding biotechnological solutions but with political and economic apathy to a non-profitable disease. The Southern Cone Initiative started in the 1990s in an effort to increase vector control and more closely monitor bloodbanks. This was incredibly successful; elimation of *Triatoma infestans* drastically reduced transmission of the disease across Uraguay, Chile, Brazil, Paraguay and Argentina by 2005 (Dias, 2007).

Various cell surface molecules aid in evasion of host immune molecules. For example *T. cruzi* mammalian stages bare a gp90 molcule which removes sugar residues necessary for parasite internalization (Covarrubias et al 2007). Tran-sialidases are and are expressed on GPI anchors, which are known to be recognised by TLR2 and (Campos et al 2001) and TLR4 (Schmitz et al 2009). They are secreted into the extra cellular mileau to aid in invasion of non-phagocytic cells (Pereira et al 1996).

Leishmaniasis is spread by the female sandfly, of which there exists over 60 species (MacMorris-Adix, 2009) in two genera *Phlebotomus* (Old World) and *Lutzomia* (New World) (Hide et al 2007). Not only is the sandfly mode of feeding crucial to parasite invasion, but its saliva's vasodilatory effects increase the parasite burden and persistance (Sacks and Kamhawi, 2001). The saliva also inhibits TNF-alpha, interferon-gamma and IL-12 and nitric

oxide production (Belkaid et al 2000). However saliva antibodies build up over time providing resistance with increased age in endemic communities (Davies et al 2000).

There are twenty one infective species of leishmania (Herwaldt, 1999) ranging through >60 countries across the Americas and Afro-Eurasia (Piscopo and Azzopardi, 2006). Leishmania are morphologically similar but genetically diverse (Croan et al 1997). Until recently, they have been characterised by epidemiological, geographical or clinical data as morphological differences were insufficient to differentiate them. Genome sequencing has disproved many of the hypothesised classifications leaving their evolutionary relationships yet to be resolved (Kerr, 2006). Leishmaniasis affects over 12 million people (Campino et al 1997) However, more than 90% of these cases are restricted to just 11 of the poorest countries: India, Bangladesh, Nepal, Sudan, Brazil, Bolivia, Peru, Afghanistan, Iran, Saudia Arabia and Syria (Hide et al 2007). Childhood infections account for more than 50% of these cases (Grech et al 2000). Leishmania is an opportunistic parasite in humans (Berman, 1997) and disease is currently on the rise in previously non-endemic areas. One reason is their adaptation to domestic animals such as dogs (*Canis lupus familiaris*) and rodents (e.g. *Rattus sp.*) (Piscopo and Azzopardi, 2006) as a result of urbanisation and deforestation (Desjeux, 2001). Another is co-infection with Human Immunodeficiency Virus (HIV) (Sharma and Singh, 2009). Relevant to this project is the finding that genetic variations in family cluster studies have an impact on susceptibility to leishmaniasis (Castellucci et al 2005).

Leismania results in cutaneous, mucocutaneous and visceral leishmaniasis. Species such as *Leishmania major* cause cutaneous leishmaniasis (McCall et al 2012,) which is self-resolving within 18 months in 90% of cases. It is localised to the bite site, which ulcers and then crusts. No man-made vaccines exist, however in endemic countries people self-vaccinate by encouraging sand-fly bites on the buttocks so as to avoid potential disfiguring scars on the face. This is known as “leishmanisation.” (Piscopo and Azzopardi, 2006). Facial scars have been known to cause ostracization of people from their communities (Yanik et al 2004). Mucocutaneous leishmaniasis may appear several years after a cutaneous lesion has self-healed, as organisms can lay dormant for years and metastise via the blood stream, causing ulceration of the nasal mucosa and performation of the septum (Marsden, 1986). This never self-heals. Deaths from secondary infection due to dessimination into the mucosa occur at a

rate of 70,000 per year (Reithinger et al 2007). The human immune respose in terms of IL-12 and TNF-alpha production and Th cell activation is stronger in response to *Leishmania major* than to visceral species which may be why it is restricted to cutaneous layers (Meddeb-Garnaoui et al 2009).

Species such as *Leishmania donovani* (McCall et al 2013) cause the most severe form of the disease; (Sharma and Singh, 2009) visceral leishmaniasis, which has a 75% to 95% mortality rate. Symptoms are fever, weightloss and enlargment of visceral organs or hepatosplenomgaly (MacMorris-Adix, 2009). Sometimes this is displayed outwardly as black pigmented spots on the skin known as Kala Azar. Incubation time can be from 3 to 34 months (Piscopo and Azzopardi, 2006). It is spread from the bite site when dendritic cells and macrophages take up infected neutrophils and migrate to deeper organs (Ribeiro-Gomes et al 2012). Visceral leishmanids are differentiated from cutaneous leishmanids by their ability to withstand higher temperatures exceeding 40^{o}c and to be more resistant to oxidative species such as nitric oxide and hydrogen peroxide (Sarkar et al 2012). The parasite inhibits oxidation using acid phophateases on its surface, and uses a proton pump to diffuse lysozymal enzymes (Alexander and Russell, 1992).

1.2.2 Trypanosomes in Other Species

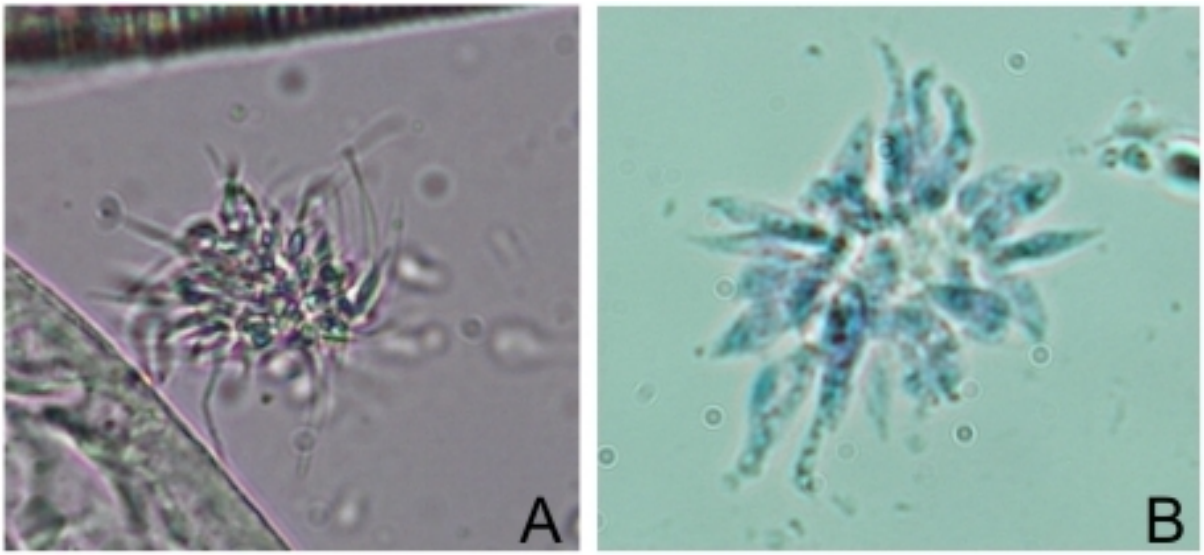
Some 472 vertebrate-parasitic species of trypanosome have been identified; 153 of fish, 60 of amphibians, 79 of reptiles, 78 of birds, and 102 of mammals (multiple species for all of the above) (Noyes et al 2002). There are also insect specific trypanosome parasites, such as *Crithidia mellificae* in honey bees (*Apis Mellifera*) and *Crithidia bombus* in bumble bees (*Bombus sp.*), both of which are implicated in current colony collapse crises (Runckel et al 2014). *Phytomonas* are plant specific trypanosomes, but as well as those known for colonizing the phloem sap of economically valuable plants such as coffee (*Coffea sp.*), coconut (*Cocos nucifera*) and oil palms (*Elaeis sp.*), there are others in the genus which cause no apparent damage and live inside fruit or seeds (Porcel et al 2014). Bodonids can be ectoparasites of fish (multiple species), or free living in thermal vents or damp soil. Moreira et al (2004) showed that trypanosomatids evolved from biflagellate bodonids by using 18S rRNA sequence comparison (Deschamps et al 2010).

It was once thought that trypanosomes co-evolved with very specific invertebrate and vertebrate hosts, but they have now been found to be far less specific than expected. Voles (*Myodes sp.*) and wood mice (*Apodemus sylvaticus*) share one species of trypanosome alongside four species of flea (Hamilton et al 2007). Trypanosome infection can have detrimental effects on heavily immuno compromised animals, that may be old, pregnant, or displaced from their habitats. In displaced wombat populations in Australia, trypanosome infection is concurrent with poor host body condition and strong inflammatory reaction leading to tissue degeneration in the heart and oesophagus (Botero et al 2013). However, Lizundia et al (2011) found no evidence of association between body condition and prevalence of *T. pestanaii* in badgers. The mouse trypanosome, *Trypanosoma musculi* produces a self-limiting infection (Jarvinen et al 1977). Trypanosomes are relient on the host for growth promoting substances. Albright and Albright (1981) found that by removing the spleen from test subjects, parasitaemia was wiped out. This placed the need for macrophages in control of the parasite as secondary to the parasite's need for host products from the spleen.

1.2.3 *Trypanosoma pestanai*, parasite of the Badger

Trypanosoma pestanai is the badger specific trypanosome. Most wild animal Trypanosomes are grouped as Herpetosoma. However, *T. pestanai* is grouped as a Megatrypanum alongside domestic and agricultural animal trypanosomes (Lizundia et al 2011). *T. pestanai* is stercorarian (Stevens et al 1998). Howevergb , unlike *Trypanosoma cruzi*, the best known Stercorarian trypanosome, which is transmitted by the rubbing of the vector faeces into abrasions such as the bite site, instead *T. pestanai* is transmitted through oral ingestion of contaminated *Paraceris melis* (badger flea) faeces during grooming (Thekisoe et al 2013). Its manner of transmission makes it unique among Megatrypanum (Hamilton et al 2007). The badgers' extensive grooming habit provides continuous opportunity for infection. *T. pestanai* infection in the host occurs in waves suggesting a variable surface glycoprotein may be in action (Lizundia et al 2011). Figure 2 shows the only image of *T. pestanai* currently available in the literature (Lizundia et al 2011).

Figure 2. Wet-smear of *Trypanosoma pestanai* from Lizundia et al 2011.



Legend: A= Presence of live motile *T. pestanai* parasites in the hindgut of an infected flea (×40). B= Detection of rosette in the hindgut of an infected flea (×100).

Badger fleas (*Paraceris melis*) are infected by ingestion of infected bloodmeals, after which, trypomastigote forms penetrate the epithelial cells of the flea stomach and replicate. Upon rupture of the infected cell, daughter trypomastigotes enter the lumen and migrate to the hindgut, where they transform into epimastigotes. They divide by binary fission, undergoing further multiplication to produce large numbers of infective metacyclic forms, which are discharged in the faeces. This parasite has been classified on the basis of morphology from blood smears; it is large, with a small kinetoplast located close to the nucleus (Lizundia et al 2011). Common features of trypanosomes include the nucleus, kinetoplast, undulating membrane, flagella pocket and the flagellum (Thekisoe et al 2013). Principal forms of *T. pestanaii* are sphaeromastigotes and epimastigotes (27 µm x 1.6 µm) with a very fine tapered posterior end (Noyes et al 1999). Lizundia et al (2011) observed various morphologies of *T. pestanai* during dissection of *P. melis* including slender, broad and intermediate forms and rosette like aggregates, the latter of which are found in the hindgut of the vector.

1.2.3.1 The Intermediate Host; *Paraceris melis*

The intermediate host for *T. pestanai* is the badger specific flea, *Paraceris melis*. Fleas evolved from freeliving ancestors. *Paraceris melis* is from the family Ceratophyllidae (Medvedev, 2004), the youngest of flea families, evolved from squirrel parasitic ancestors in the early Eocene (Beaucournu et al 2005). Diversification is thought to have occurred via host switching. Individual fleas disperse long distances via dispersal from host to host through body contact (Rust, 1994). The main determinant of the size of their geographical range is the geographical range of the host (Pulliam, 2000). Setts in temperate and colder regions are deeper, and frequented by a wider range of animals than those in warm regions; which facilitates host switching. *P. melis* is host specific, but will feed on alternate hosts.

Paraceris melis is dioecious and both sexes take blood meals. They mate up to six times an hour, producing large numbers of eggs (around 50 per day) (Vashchenok, 1997). Larval and nymph stages last from two to eight months depending on environmental conditions (longer during the cold season) and result in an average three generations a year (Vashchenok, 1993). Fleas can vary the lengths of their pre-adult stages so that the period from egg to adult varies from twenty-three days to nineteen months. Eggs are laid in deep parts of the sett where there is a stable microclimate, and hatch in two to fourteen days. Larvae are negatively phototactic and positively geotactic, keeping them within the confines of the sett. They feed on faeces from adult fleas, which has higher protein content than blood taken up by adults (Shryock and Houseman, 2005). During the third instar, the larva spins a silken cocoon from its salivary glands, which it camouflages by rolling about and in particles in the substrate (Qi, 1990). Badger flea pupae form firm cocoons, characteristic of nest fleas, rather than the soft cocoons characteristic of body fleas (Medvedev, 2004). They can reach adulthood within a week or remain cocooned for up to a year. Emergence is triggered by vibration, temperature increase or an increase in CO_2 concentration in the surrounding air, all indicating the presence of badgers (de Albuquerque Cardoso and Linardi 2006). From their emergence, fleas are capable of distinguishing between host species via odour (Taylor et al 2005). Newly emerged fleas mate immediately after emergence (Dean and Meola, 2002). However, eggs mature only after a blood meal (Ribeiro et al 1990). Feeding triggers further mating behaviour, egg maturation and oviposition. (Kirk, 1991).

Adult *P. melis* are large, measuring up to 4mm (Medvedev, 2004). Morphological features of fleas complement the behaviour and fur morphology of their hosts. Sharply pointed pro-natal spines allow them to attach to badgers' coarse fur and resist grooming. The fleas' laterally compressed body, high narrow head capsule and flexible joints allow them to run through the fur (de Albuquerque Cardoso and Lenardi, 2006) to avoid detachment via grooming. However, they detach voluntarily in winter to avoid the cold outside. Badgers skin temperature can fall to unfavourably low levels in the winter cold, as internal fat is their primary insulation. While virgin and mated female fleas, and virigin male fleas, overwinter, sexually experienced males die off immediately at the end of the reproductive period (Michelsen, 1997). Overwintering fleas have reproductive diapause, and rapidly develop fat tissue to survive prolonged starvation. This allows them to find new hosts. When separated from a host they lighten up by vomiting, then jump about for around half an hour, before seeking shelter until a stimulus is present (Cox et al 1999). Jumping distance decreases after 35 days. If separated from the host, life expectancy is 89 days (Cox et al 1999).

The full life cycle of *T. pestanai* within the flea is unknown, though it follows a stercorarian mode of transmission. The first description of a trypanosome's passage through a flea was given by Swingle (1911) describing *T. lewisi* in the rat flea. After a blood meal, *T. lewisi* migrates into the intestine to undergo a series of modifications. Division then takes place with trypanosomes attached by their flagella ends to the intestinal wall. They form cysts and are passed out with flea faeces, to be taken up by the rat during grooming.

Prevalence of trypanosomes in badgers was recorded as 43% in one study (McCarthy et al 2009,) while flea burden is comparatively low with 19.7% prevalence, and only 2.6 fleas per animal on average (San, 2007). On juvenile badgers ticks outcompete fleas, especially in summer when tick density increases with soil temperature. Thus fleas tend to largely be found only on adult badgers (Krasnov, 2008). Lice compete for the host from spring to autumn but fleas retain a constant low level burden throughout the year (San 2007). Badger trypanosome infection is directly related to the proportion of fleas infected. Non-infected badgers showed variable results where 5% (Cox et al 1999), 18% (Noyes et al 2002), and 16% of their fleas carried trypanosomes, while the flea populations found on trypanosome-infected badgers were 61% trypanosome-infected (Lizundia et al 2011).

Badgers switch their preferred sleeping sites in response to flea accumulation (Roper et al 2001). They do not usually return to the previous nights bed chamber for 6 to 7 days. Fleas left in bedding material (perhaps via grooming) abandon it within six days, moving phototactically and against gravity to gather around the entrance of the den. This effectively redistributes the fleas among the badgers. They jump in response to chemical and physical stimuli, such as carbon dioxide sources at respiratory levels (4%,) dark shapes and vibrations indicating animal movement (Cox et al 1999). Burrow dwelling fleas are negatively phototactic after a bloodmeal, but strongly positively phototactic when starving (Eiseman and Binnengton 1994). Given their life expectancy and the tendency of badgers to move between setts, infected fleas have plenty of time and opportunity to bite multiple badgers.

Lizundia et al 2011 found that young or male badgers show a higher prevalence of flea infestation than the old or female badgers (males 42%, cubs 40%, young adults 35%, females 27% and adults 16%). With age, repeated infestations evolve into a prolonged or intermittent carrier state. This may relate to the toll on male fitness due to mate seeking. Fleas tend to aggregate in badger communities (Jokela et al 2000). Most host individuals are weakly infested or not at all, while a few individuals harbour most of the population (Stanko et al 2002). Fleas tend to choose lower quality hosts over fitter hosts whose blood may be more nutritive, due to their lower ability to allocate immune and behavioural efforts towards flea removal (Pulliam, 2000). Flea association with male hosts allows them to disperse (Smith et al 2005). Male fleas are more likely to be found on found on roving juvenile male badgers, allowing them to increase their chances of meeting unrelated females. Female fleas are more abundant on female badgers, so that by staying on shorter dispersing hosts their offspring are more likely to benefit from the guaranteed resources within the host sett (Krasnov, 2008).

1.2.3.2 Other Endoparasites Found in Badgers

Other endoparasites of badgers are *Toxoplasma gondii*, *Eimeria melis* and various helminth species. *T. gondii* is more prevalent in agricultural sites (67 to 77%) than urban sites (39%) (Anwar et al 2006). *T. gondii*'s definitive host is the cat, which is far less abundant in the countryside than in urban areas (Murphy et al 2008). Similar to the findings of Hide et al (2009) in sheep; *T. gondii* was found to be transferred from badger to badger via transplacental transmission (Anwar et al 2006). Coccidial parasites, such as *Eimeria melis* and *Isospora melis*, cause swollen abdomens and diarrhoea in badgers cubs (Newman et al 2001). Male cubs were most affected, followed by female cubs, male adults and female adults (Cottrell, 2011). Infected cubs suffer pronounced physiological stress, including diarrhoea and dehydration, which can lead to morbidity. Disease impacts affected cubs for up to four years, and has been traced by the preponderance of assymetrical upper palate markings compared to healthy litter mates; parasite prevalence decreases and symmetry increases after the age of four (Nouvellet et al 2000). Coccidial parasites are monoxenous, completing their life cycle within a single host. Parasites multiply in the gut releasing oocysts which are then defecated, and transmitted via contamination of ingested foods (Anwar et al 2000). In adults, low gut levels of coccidia are normal and in most cases asymptomatic; infection occurs only in immunocompromised animals or pregnant and lactating females.

Helminth parasites identified include one cestode and four nematodes specific to badgers. Badger behaviours of fossorial feeding and anal scent marking other group members, facilitate infection by helminths (Rosalino et al 2006). The badger GI parasite, *Capillaria putorii* is a nematode, commonly known as a hair worm, which causes vomiting and diarrhoea, weight loss and decreased breeding success at high levels. However, at low levels *C. putorii* presence is normal and rarely fatal, and most badgers exhibit no obvious symptoms (Torres et al 2001). The common route of transmission is via the consumption of intermediate hosts, in the case of *C. putorii* the intermediate host is a hedgehog, which are known to be consumed by badgers when other resources are low. *C. putorii* is also thought to be transmitted via consuming eggs from soil contaminated hedgehog faeces (Taylor et al 2013). Cubs are most significantly affected, and bear the heaviest parasite burdens (Cottrell, 2011). In adults, parasite burden is often due to a combinatorial effect of host condition and environmental factors (Hillegrass et al 2010; Marcogliese and Pietrock, 2011; Pross, 2012).

1.3 The Mammalian Immune System

The principal weapon evolved to fight off parasites is the immune system, which in vertebrates consists of the innate and acquired systems (Murphy et al 2008). The innate immune system is the first line of defence against invading pathogens and parasites, present in the form of physical and chemical barriers, phagocytes, dendritic and natural killer cells and plasma proteins that form the complement cascade. It forms an immediate non-specific response, recognising broad classes of organisms such as bacteria, fungi, viruses and parasites. Phagocytosis triggers degradation of pathogens and presentation of antigens (Takeda and Akira, 2005). If a pathogen cannot be immediately destroyed, it is contained in the phagocyte until the adaptive system can act. The innate system has no memory and so acts with equal potency no matter how regularly exposure occurs. Its main functions are coagulation, opsonisation, phagocytosis, apoptosis and activation of complement (Janssens and Beyaert, 2003). The first two of three pathways involved in the complement cascade are innate immune responces: the alternative pathway (activated directly by microorganisms) and the lectin pathway (activated by lectins binding to microorganisms). The third, classical pathway, is activated by immunoglobin-antigen complexes, and evolved rapidly from previously non-immune functions (Bartl et al 2003).

The adaptive response developed in an ancestor of jawed vertebrates. It involves T or B lymphocytes which detect foreign antigens using highly specific T-Cell receptors or Immunoglobins (Ig) respectively, and memory cells (Hood et al 1985). Ig binds free intact antigens while T-cell receptors process antigenic fragments presented by the Major Histocompatibility Complex (MHC) (Bartl et al 2003). T cells mature into helper or cytotoxic T cells, forming cell mediated immunity, while B cells mature into antibody secreting plasma cells, forming humoral immunity. It is a slower response taking up to five days to act, but is much more potent and specific, creating memory cells after exposure leading to a faster and more potent response when reinfection occurs. Until recently the innate and adaptive systems were seen as two completely separate entities. The discovery of Toll-Like Receptor (TLR) function linked the two for the first time. The first responders of the innate system, such as phagocytes, were found to bear both external and internal receptors, which signal molecules of the adaptive system which are of particular interest in this research.

1.3.1 Toll-like Receptors

The innate immune system has a critical role in the control of adaptive immunity (Medzhitov, 2009). Each cell is bound by a phospholipid bilayer, proteins in this layer allow the cell to sense and interact with the cellular environment. It is here that Toll-Like Receptors are located. They project through the plasma membrane of white blood cells (Rodriguez et al 2012), where they act as cellular inboxes, standing ready to accept messages in the form of Pathogen Associated Molecular Patterns (PAMPs), and ultimately initiate a signal transduction cascade which activates transcription factors leading to increased production and secretion of cytokines and T-cells to initiate a wider immune response (Christmas, 2010). Stimulation of most TLRs leads to Th1 rather than Th2 differentiation. Activation is characterised by a polarised type I response with high production of microbicidal intermediates such as nitric oxide and reactive oxygen species (Savino et al 2007).

Pathogen Associated Molecular Patterns are evolutionarily conserved motifs unique to microorganisms, essential for their metabolism and survival. Mutation can be lethal to pathogens and they are therefore quite conserved; they are invariant between micro-organisms of a given class (Janssens and Beyaert, 2003). This means a limited number of Pattern Recognition Receptors (PRRs) are needed to detect the presence of infection (Areal et al 2011). They are not produced by host cells, (Christmas, 2010,) allowing the innate immune system to distinguish between self and non-self (Muzio et al 2000). Redundancy also occurs through combination of different TLRs. Although TLRs do not always recognise PAMPs directly, response is always determined by the TLR alone, not by PAMP recognition proteins. TLRs act as the control mechanism which will either activate a cascade or not in response to signals received (Janssens and Beyaert, 2003).

The Toll family of receptors are conserved throughout the evolution of the kingdom Animalia, homologues are found in all animal species from flies to humans (Vasselon and Detmers, 2002). They belong to the type I transmembrane glycoprotein receptor family (Areal et al 2011). The Toll gene was first identified as a site of developmental mutation in *Drosophila melanogaster*, essential for determining dorsal ventral polarity during embryogenesis (Lu et al 2008). It was dubbed 'Toll' or "great" by its German discoverers (Christmas, 2010). When cloned in 1997 it was shown to switch to encoding an essential Drosophilan membrane

receptor used in defence against fungal infection in adulthood (Bowie and O'Neill, 2000). TLRs were the first physical proof of the self vs. non-self concept. By 1998 a mammalian homologue, involved in promoting inflammatory responses, TLR 4, was found (Williams, 2013). No other members of the *Drosophila* toll family (18 wheeler/toll 2 to toll 9) have immune function, suggesting an alternate ancestral function. In contrast all mammalian TLRs are specialised for immune function, indicating they have a common ancestor (Takeda and Akira 2005).

Most mammals have 10-13 different TLRs (Roach et al 2005). TLRs can be categorised as either ubiquitous, restricted (which includes TLR2 and TLR4) (Muzio et al 2000), or specific. TLRs can be divided into viral, and non-viral types (both TLR2 and TLR4 are non-viral). Viral TLRs recognise foreign nucleic acids but also target self-components (Krieg and Vollmer, 2008). They have the dual role of maintaining their function while avoiding autoimmunity, so are not expected to accumulate mutations as this would affect their functional integrity (Barreiro et al 2010).

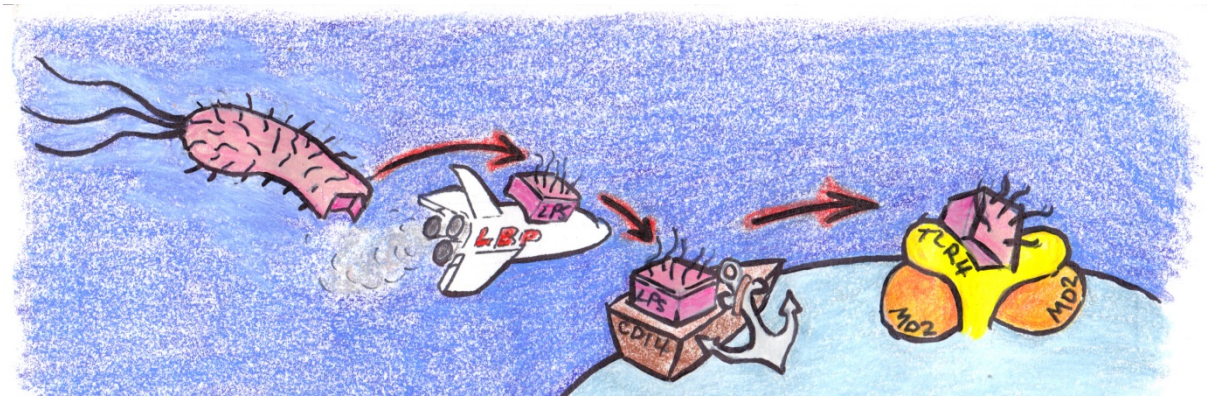
Non-viral TLRs exist on the cell surface and have a more flexible evolution, so can easily tolerate mutations (Barreiro et al 2009). These mutations can be subject to positive selection and become fixed in some populations. Positive selection may have been imposed by interacting pathogens (Grueber et al 2014). The function of non-viral TLRs is more redundant than that of viral TLRs. One microorganism may be recognised by several non-viral TLRs, so that a mutation in one does not necessarily compromise immunity (Skerrett et al 2007). Non-viral TLR4 has the highest percentage of positively selected codons (2.65%,) which are found in the Leucine-Rich Repeat (LRR) domain (Areal et al 2011). Single Nucleotide Poltmorphisms (SNPs) in TLR3, TLR7 and TLR9 which sit within endosomes and lysomes, are relatively infrequent compared with those on the cell surface such as TLR2 and TLR4. The latter are highly conserved between species because of their involvement in intracellular signalling (Langefeld et al 2009). By contrast, those expressed extracellularly exhibit significantly higher divergences, reflecting their involvement in PAMP recognition from multiple microbial sources. The most diverging part mediates the recognition of the ligand with surface charges in this region, which differ considerably between species (Werling et al 2009).

All TLRs have a common domain organisation, with an extracellular domain, a helical signal transmembrane domain, and an intracellular Toll/IL-1 receptor homology domain (TIR) (Gong et al 2010). The transmembrane domain determines the subcellular localisation of TLRs. The TIR domain is conserved across all TLRs and IL-1 receptors, and is also shared by downstream signalling adapter molecules. Upon receptor ligation, the TIR signalling complex is formed between the receptor and the adapter TIR domains (Areal et al 2011). The ectodomain is a horseshoe shaped solenoid structure and is directly involved in the recognition of a variety of pathogens (Gong et al 2010). Despite the low sequence homology, the change of one single conserved residue of the TIR in TLR 2, 4, and 6 (Pro to His) makes these proteins either inactive or active as dominant negative mutants (Vasselon and Detmers, 2002).

TLR structure is characterised by an extracellular leucine rich repeat domain (Janssens and Beyaert, 2003) housing a ligand binding site and one or two cysteine rich regions (Park et al 2009). Signalling domains of all TLRs are highly conserved, while LRR ligand recognition domains are more diverse to accommodate recognition of different PAMPs (White et al 2003). Leucine Rich Repeats (LRRs) are arrays of 20 to 30 amino acid long protein sequences enriched with hydrophobic amino acid leucine. LRRs reside in the ectodomain. All LRRs sequences can be divided into a highly conserved segment and a variable segment. The highly conserved segment consists of an 11 or 12 residue stretch with the consensus sequence; LXXLXLXXN (CX) XL (Gong et al 2010) where, L = leucine, isoleucine, valine or phenylalanine; N = asparagine, threonine, serine or cysteine; C= cysteine and X = any amino acid. Each of the 20 or more LRRs features about 10 relatively conserved residues that contribute to the concave surface of the extracellular domain. The remaining portion of each LRR is more variable and contributes to the convex surface (Kubarenko et al 2010). All LRRs in TLRs are capped by N and C terminal LRRs that are usually irregular and do not match any type of LRR consensus sequence (Gong et al 2010). TLRs have 3 particular regions of significant homology; box 1, 2, and 3. Box 1 is a signature sequence of the family. Box 2 and 3 contain amino acids important in signalling (Bowie and O'Neill, 2000). TLRs are expressed most commonly on neutrophilic granulocytes, myeloid and monocyte derived dendritic cells, monocytes and plasmacytoid dendritic cells (in that order), to a lesser degree in natural killer cells, B cells and CD4 T-cells, and rarely in CD8 T-cells (TLR3 only) and eosinophilic granulocytes (TLR7 only) (Ku et al 2007, and, Muzio et al 2000).

Each TLR signalling cascade is slightly different with different input received and outcomes (Vasselon and Detmers, 2002). Signal transduction events result in the production of cytokines and antimicrobial effector molecules (An et al 2002). The TLR4 signalling cascade is useful to look at for an overview of TLR signalling. Different TLRs use different combinations of adapter proteins to determine downstream signalling; however TLR4 is the only TLR to utilise all adapter proteins used by every other TLR combined (Lu et al 2008). With the exception of TLR3, the universal adapter protein Myeloid Differentiation Primary Response Gene 88 (MyD88) is the common denominator for all remaining TLRs (Drouin-Ouellet and Cicchetti, 2012). The MyD88 independent pathway also mediates the induction of type I interferons and interferon inducible genes (Lu et al 2008). MyD88 independent pathways are dependent on a Toll/IL-1 receptor domain containing an adaptive inducing interferon beta (TRIF). TLR4 features both MyD88 dependent and independent pathways. The TLR4 MyD88 independent pathway occurs in the phagosome, not the cell surface (Drouin-Ouellet and Cicchetti, 2012). All TLRs except TLR3 can activate downstream signalling cascades through the adapter protein MyD88. MyD88 dependent signalling via direct TLR stimulation activates dendritic cells and adaptive immune responses (Hou et al 2008). Given that this study focuses on TLRs 2 and 4 and their interaction with Trypanosome GPIs, only the MyD88 dependent pathway will be considered in detail here; Figure 3.1 to 3.6, drawn by myself, give a brief overview of TLR4 MyD88 dependent signalling.

Figure 3.1 TLR4 Signal Transduction Cascade: MyD88 Dependent Pathway.



Legend: Pink animal= invading bacteria, pink square= LPS, shuttle= LBP, anchor= CD14, globe= white blood cell, yellow and orange= TLR4/MD 2 receptor complex.

Figure 3.1 shows that upon bacterial invasion, lipopolysaccaride (LPS) is extracted from the membrane and transferred to the TLR4 MD2 complex (Park et al 2009). By a series of interactions with several proteins including LPS Binding Protein (LBP), CD 14, MD 2 and TLR4. LBP is a soluble shuttle protein which directly binds to LPS and facilitates the association between LPS and CD 14. CD 14 is a GPI anchored protein which facilitates the transfer of LPS to the TLR4/MD 2 receptor complex (Lu et al 2008). Ferguson (1999) found that the CD14 LBP receptors function equally well with a GPI anchor as does its mobile counterpart. TLR4 and myeloid differentiation factor 2 (MD2) form a heterodimer that recognises structurally diverse LPS molecules (Park et al 2009). An example of redundancy in this system is that RP105 and MD 1 activate NF-kB in response to LPS through TLR4 in the absence of MD 2. Nod receptors are also able to do this independent of TLR4, MyD88, and 'TNF receptor associated factor' (TRAF) 6 (Vasselon and Detmers, 2002).

Figure 3.2 TLR4 Signal Transduction Cascade: MyD88 Dependent Pathway.

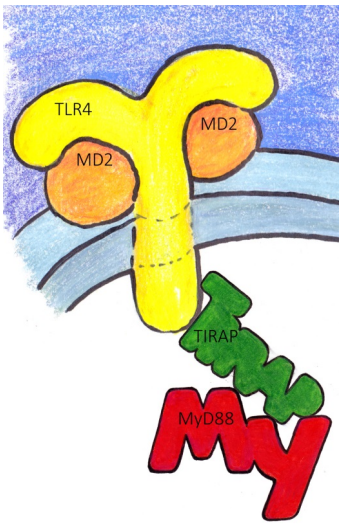


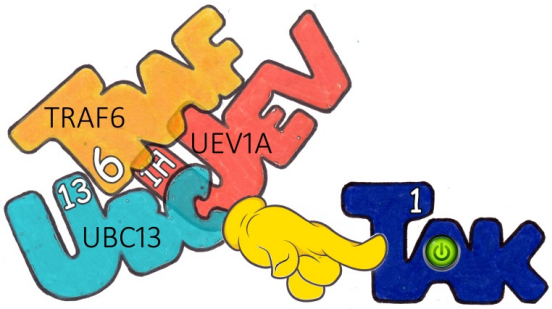
Figure 3.3 TLR4 Signal Transduction Cascade: MyD88 Dependent Pathway.

Legend: Yellow= TLR4, orange= MD2, green= TIRAP, Red= MyD88, violet= IRAK4, Navy blue= IRAK 1.

Figure 3.2 shows that TLR4 then recruits downstream signal transduction adaptor proteins through interactions with TIR domains. There are five TIR domain containing adaptor proteins: MyD88, 'TIR domain containing adaptive protein' (TIRAP) also known as the MyD88 adapter like protein (MAL), TRIF, 'TRIF related adaptor molecule' (TRAM), and a 'Sterile Alpha and HEAT-Armadillo motif containing protein' (SARM) (Lu et al 2008). Toll/IL1 receptor associated protein (TIRAP) is required for signalling cascades following LPS recognition, but

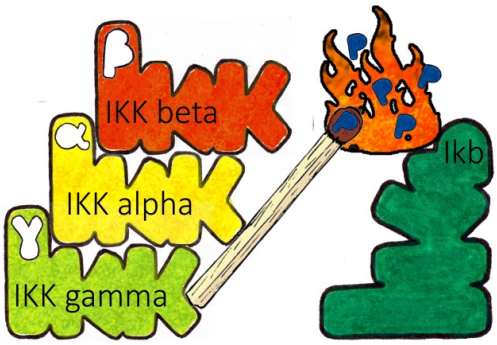
not for cascades following recognition of CpG island patterns from parasite DNA (Krieg, 2002). It is also required for signalling the presence of Trypanosome GPI anchors (Kayama and Takeda, 2010). MyD88 association is essential for inflammatory cytokine production through all TLRs and also for the inhibition of apoptosis (Takeda and Akira 2005). Elements of TLR signalling cascade are critical. TRIF and MyD88 deficiency lead to lowered ability to control parasite levels and earlier onset of death (Rodriguez et al 2012).

Figure 3.4 TLR4 Signal Transduction Cascade: MyD88 Dependent Pathway.



Legend: The TRAF6-UBC13-UEV1A complex activating TAK1.

Figure 3.5 TLR4 Signal Transduction Cascade: MyD88 Dependent Pathway.

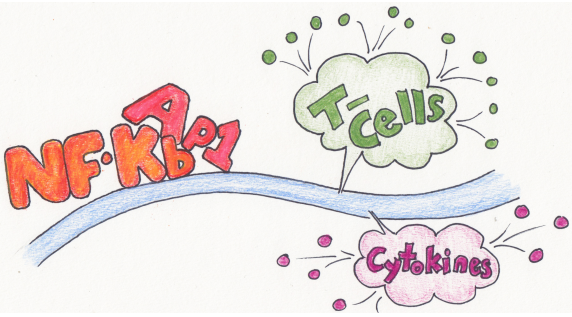


Legend: The IKKα-IKKβ-IKKγ complex phosphorylating Ikb. The little P's represent phosphates.

Figure 3.3 shows MyD88, which contains a TIR domain and Death Domain (DD), signalling another adaptor protein IL-1 Receptor Associated Kinase (IRAK.) IRAK4 is an essential DD containing kinase, responsible for subsequent recruitment, activation and degradation of IRAK 1. IRAK for is vital for childhood immunity, deficiency causes death in the young, but it shows redundancy in adulthood (Ku et al 2007). The cascade continues as IRAK1 activates TRAF6, which is known to be a critical downstream protein. TRAF 6 forms a complex with 'Ubiquitin Conjugating Enzyme' (UBC) 13 and 'UBC variant 1 isoform A' (UEV1A), shown it figure 2.4. This complex activates 'transforming growth factor beta activated kinase' (TAK) 1. TAK 1 then activates the IkB Kinase (IKK) and Mitogen Activated Protein Kinase (MAPK) pathways (Lu et al 2008). The first pathway (IKK) causes IKKα, IKKβ and IKKγ to form a complex as seen in figure 2.5, and phosphorylate IkB proteins. Their phosphorylation leads to their degradation, and the subsequent translocation of transcription factor NF-kB which controls the expression of pro-inflammatory cytokines (Lu et al 2008). MAPKs induce AP-1 dependent gene transcription, by the addition of phosphate groups to certain amino acids

(Christmas, 2010). AP1 pathways can also be activated by TLRs 2, 6 and 9 (Vasselon and Detmers, 2002).

Figure 3.6 TLR4 Signal Transduction Cascade: MyD88 Dependent Pathway.



Legend: NF-kB and AP-1 stimulating T-Cell Production

Figure 3.6 shows NF-kB and AP-1 binding to the promoter of a wide range of a wide of different target genes, stimulating proliferation and production of key pro-inflammatory mediators such as T-cells and cytokines (Janssens and Beyaert, 2003). As part of a positive feedback loop; pro-inflammatory cytokines such as IL-1β, TNFα and interferon-γ also induce TLR4 transcription (Muzio et al 2000).

1.3.2 TLRs and Their Interactions with Trypanosomes

Four of the TLRs are known to activate cytokine production in response to trypansomes. TLR2 and TLR4 sit on the outside of the cell and accept molecules that recognise GPI anchors (McCarthy et al 2009). GPI anchors are evolutionarily conserved structures found on all eukaryotic animals which provide stable membrane anchorage, and sequester proteins into lipid rafts (Ferguson, 1999). Upon GPI recognition they induce production of nitric oxide (Savino et al 2007) and synthesis of pro-inflammatory cytokines and chemokines (Teixeira et al 2002). While TLR 7 and TLR 9 project into the cell and recognise CpG sequences, revealed once the parasite has been phagocytosed (Christmas, 2010).

TLR 2 is commonly expressed on the cell surface but can also be recruited to the lysosomal compartment of macrophages just like TLR 4 (Takeda and Akira 2005). It has a wide repertoire of PAMP recognition abilities due to its ability to form heterodimers with both non-TLR-family receptors, and other TLRs, and these include recognising trypanosome GPI anchors (Vasselon and Detmers, 2002; Rodgrigues et al 2012). Pro-inflammatory response to *T. cruzi* GPI, by TLR 2, is attributed to its recognition of unsaturated fatty acids and a long

glycan core. (Kayama and Takeda, 2010). TLR4 stimulated production of pro-inflammatory cytokines as well as nitric oxide are also important in early parasiticidal events.

Studies in mice have shown that TLR4 deficiency alone does not lead to mortality, and that TLR2 also plays a role in trypanosome resistance (Rodriguez et al 2012). TLR 2 and TLR 4 also cooperate during *T. gondii* infection, optimising TNF production after recognition of GPI anchors. Whether TLR 4 or TLR 2 are more critical to trypanosome resistance is unresolved. In vitro murine studies by Yarovinsky (2008) found that deficiency of TLR 4 leads to a more severe immune pathology than TLR 2 deficiency. Conversely, in vivo murine studies by Gowda (2007) and in vitro human studies by Boutlis et al (2006) showed that GPI is recognised mainly by TLR2 and to a lesser extent by TLR4. Rodriguez et al (2012) differentiated the role of both TLRs according to trypanosome life stage. Trypomastigote GPI anchors activate TLR2 due to their Alkylacylglycerol composition, while metacyclic trypomastigote GPI anchors activate TLR4 due to their mixed ceramide and dihydroceramide structure. Rodriguez et al (2012) argues that since tGPI anchors display a hundredfold superior activity to eGPI anchors, this must attribute higher significance to TLR2 versus TLR4 in anti-trypanosome immunity.

While both TLR4 and TLR2 lead to activation of NF-kB and MAPK family members (Werling et al 2009) they are not functionally equivalent. TLR4 produces a TH 1 response, while TLR2 produces a TH 2 response. Only TLR4 is able to stimulate production of the IL-12 cytokine which leads to IFN-γ production, and IP-10 production, a chemo attractant for monocytes, NK cells and TH 1 cells (Janssens and Beyaert, 2003). TLR 2 produces greater amounts of the chemokine IL-8, a chemo attractant for neutrophils (Re and Strominger, 2001). An autoimmune issue has also been identified in the TLR response, where sustained pro-inflammatory response can potentially damage cells, contributing to pathogenesis (Moncada et al 2003) (Shio et al 2012).

Other TLRs known to be responsible for trypanosome resistance are TLR 7 and TLR 9 (Werling et al 2009). They are found in endosome and lysosome membranes (Pifer and Yarovinsky, 2011). Their ligand binding site projects into the interior of cells from their attachment site in the membrane (Christmas, 2010). TLR9 is highly specific, responding only to foreign, unmethylated, CpG sequences (Vasselon and Detmers, 2002). Rodriguez et al (2012) found

that both TLR7 and TLR9 recognise *T. cruzi* CpG motifs and are responsible for activating a host response. These sequences are revealed by degradation of the pathogen, where the TLR binds directly (Kreig, 2002). Its specificity is determined solely by the polymorphic sequence of the TLR9 receptor (Janssens and Beyaert, 2003; Krieg, 2002). Following stimulation, TLR 9 gene expression reaches peak levels within three hours, (An et al 2002,) and leads to nitric oxide production in macrophages (Takeda and Akira 2005). Both TLR 7 and TLR 9 induce a MyD88 dependent pathway in response to trypanosome CpG DNA. In vitro studies on mice have found them to be among the major TLRs involved in the recognition of *T. cruzi* (Kayama and Takeda, 2010). The relative importance of TLR 9 versus TLR 7 versus parasite GPI recognition is also debated. Khan (2006) found that TLR 9 knock out mice had less ability to clear parasites than the wild type, but suffered less pathology and reduced mortality.

1.4 Badger TLRs – The Missing Link

There is evidence that TLR sequence variation is linked to pathogen susceptibility. Single Nucleotide Polymorphisms (SNPs) in the TLR 2 gene have been associated with tuberculosis and leprosy, and in the TLR4 gene with gram negative sepsis (Everett, 2009; Areal et al 2011). Studies on humans, cattle and mice have indicated evidence for genetic variation in TB resistance. In dairy cattle, exploitation of genetic variability has already been established and used in selection programs for mastitis resistance. A microsatellite in the 3' untranslated region of the bovine NRAMP gene has been found to be associated with survival of the BCG vaccine in macrophages (Allen et al 2010). Polymorphisms within TLR genes have been associated with both disease resistance in livestock, and disease susceptibility. An Arg753Gln polymorphism in TLR2 has been associated with predisposition to staphylococcal infection, tuberculosis, rheumatic fever and urinary tract infection in children (Werling et al 2009). In a study of 870 children in Ghana, an SNP in TLR4 was found to influence malaria susceptibility (Gowda, 2007). Thus a study in the genetic variability of resistance to parasites has implications for current issues.

As yet there are no published TLR sequences for the badger. There is no information on variation in TLRs in badgers and, there is no information on the relationship between TLR sequences and trypanosome infection in the badger. Four of the TLRs are known to interact with trypanosomes. They are split into two pairs that interact either with the external

structure of the parasite or its DNA once phagocytosed. Of each pair it is still debated as to which is more important in these different roles, and a comparison of the relative importance of the two types of Trypanosome detection has not been perfomed. This study was time limited to one year, one researcher and a small budget, thus the scope was limited to two of the four possibilities. TLR2 and TLR4 are known to heterologously desensitize each other; in other words, either one will produce self-inhibitory molecules which inhibit the other, or, for example in the case of response to LPS, the activation of TLR2 creates an inhibitory LPS tolerance in TLR4 (Sabroei et al 2003). This degree of mutual inhibition or redundancy makes TLR2 and TLR4 important to study in tandem; of the TLRs known to be responsive to trypanosome invasion, these were chosen because they are non-viral. TLR4 is known to have the highest percentage of positively selected codons (Areal et al 2011). TLR7 and TLR9 are viral TLRs, so are highly conserved as they must maintain avoidance of autoimmunity (Takeda and Akira 2005; Muzio et al 2000). It was postulated that it may be more likely to find variations within TLR2 and TLR4 sequences than TLR7 and TLR9 sequences.

2.0 Research Methods

The broad objectives to achieve of this study are; to assemble a collection of badger samples, and extract DNA; to perform bioinformatic searches for TLR genes and Design Badger TLR primers; to then optimize PCR protocols for amplification of target regions; to sequence and analyse TLR2 and TLR4; to compare any sequence variation with trypanosome infection status; and also to compare badger TLR sequences with those of other species.

To perform DNA extraction, 400µl blood or $1cm^3$ tissue, were placed in a 1.5ml Eppendorf. Then, 400µl lysis buffer (a solution used for lysing cells composed of 10 mM Tris-HCl [pH 8.0], 1 mM EDTA, 500 mg pronase B, 1 % SDS) was added, and 10µl Proteinase K (20mg/ml) in the case of tissue, or 20µl Proteinase K in the case of blood. The samples were incubated at 56^{o} overnight. Then, 500µl Tris buffered phenol-chloroform pH 8.0 was added and mixed for 10 minutes. The samples were then centrifuged for 10 minutes at 13000g and the supernatant transferred to a fresh tube. This was repeated twice more, until the supernatant was clean of protein and debris. Following this, 90µl sodium acetate (3M pH 5.3) and 900µl 100% ethanol were added to the supernatant. Samples were stored at $-20^{o}C$ overnight. Cold samples were

centrifuged for 20 minutes at 13000g, and the supernatant was discarded. The pellet was washed in 500µl 70% ethanol and centrifuged for 10 minutes at 13000g, and the supernatant was discarded. The pellet was air dried at room temperature in a sealed dessicating jar, then dissolved in 100µl of Tris-EDTA buffer. This is a solution used to solubilize the DNA and to protect it from degradation, composed of 10 mM Tris-HCl and 1 mM disodium EDTA at pH 8.0. The solution was then left for at least 24 hours in the fridge to ensure all DNA was loosened from the side of the tube and dissolved. Extracted DNA samples were tested for purity and concentration on a Nanodrop Lite Spectrophotometer using the Thermo Scientific equipment protocol. The spectrophotometer gave a purity ratio score out of 2, where ratings over 1.8 were considered very pure.

To detect trypanosome infections, the ITS-Nested PCR method (Cox et al 2005) was carried out by Eze Justin Ideozu. Primer sets used are listed in Table 1.

Table 1. ITS-Nested PCR primer sequence for trypanosome detection

Outer Primer Sequences	ITS 1	5'-GAT TAC GTC CCT GCC ATT TG-3'
	ITS 2	5' TTG TTC GCT ATC GGT CTT CC-3'
Inner Primer Sequence	ITS 3	5' GGA AGC AAA AGT CGT AAC AAG G- 3'
	ITS 4	5' TGT TTT CTT TTC CTC CGC TG-3'

Nested PCR amplification was performed on a Stratagene Robocycler. The first round reaction volume totalled 25 µl consisted KCl buffer (10 mM Tris- HCl, pH 9.0, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X -100, and 0.01% (w/v) stabilizer), 2 µM of each outer primer ITS1 and ITS2, 1 mM dNTP, 1.25 U of Biotaq and 1 µl DNA. PCR was set for 5 minutes at 94 ^{o}C; 35 cycles with 40 seconds at 94 ^{o}C, 40 seconds at 58 ^{o}C and 1 minute 30 seconds at 72 ^{o}C; and finally 5 minutes at 72 ^{o}C. For the second round reaction, 1 µl of the resulting PCR products from the first round reaction were placed in fresh tubes containing 24 µl of the same reaction mixture, with the exception of the substitution of the outer primers (ITS 1 and ITS 2) with the inner primers (ITS 3 and ITS 4).

PCR for TLR variation was performed by Andrew Whiteoak. In order to prepare for the TLR PCR stage Master Mixes (MM) were made, and Statagene Robocycler PCR machines switched on and left to arrive at the required temperatures ahead of time. For all PCRs denaturing temperature was taken as 94^{o}C and extension temperature as 72^{o}C, run for 5 and 10 minutes respectively, at the beginning and end of the cycle. To make the MM, 0.2ml PCR tubes are used. All reagents except Taq polymerase were removed from the freezer and allowed to thaw fully. MM in this research were always made to x4 scale (to supply four PCR tubes) as it was found that this produced more reliable results than MM of grander scales as ingredients could be more thoroughly mixed without harming the Taq polymerase enzyme. This also allowed for zero wastage. Working on ice, 75µl sterile water were pipetted into each 4x MM tube, as well as 10µl NH4 buffer, 4µl of 50mM $MgCl_2$, and 1µl of dNTP. Then, being careful to change the pipette tip each time, 2µl of 25mM forward primer and 2µl of 25mM reverse primer were added. The lid was closed and the solution mixed gently by flicking the side of the tube. Taq polymerase (Taq) was added last and left in the freezer until all other reagents had been added.

When all MMs were prepared as above, 2µl Taq were added to each tube, which were then closed and flicked only once gently before being placed back on ice. Upon completion of MM, three lots of 24µl were pipetted into three additional PCR tubes so that four tubes containing the same amount of MM were arrived at. After this, 1ul of DNA sample were pipetted into its respective tube.

Table 2 is a list of primer sequences for all primer sets used, and shows the cycle times needed for each primer. Where a G, A, T or C base could not be found to be conserved across species during MSA a degenerate base was used; these can bind to one or more base allowing for a region to be primed without knowing every base within its 5' or 3' end. Once PCR was complete tubes were removed and immediately put onto ice. Following PCR, amplicons were run on 1% agarose gel using a Consort EV243 power supply. 7µl loading buffer and 7µl of sample waspipetted into each well. 5µl of Hyperladder was pipetted into one of the wells. When complete, the gel was brought to the gel imager.

Table 2. Primer sequences and PCR settings used in this research for TLR2 and TLR4.

Gene	Primer	Sequence 5'- 3'	Target	Td	Ta	Te	Cy
TLR2	2s F	ATGTCACGTGTTYTGTGGACA (21 bp)	1110 bp	94°c 0:40	60°c 1:10	72°c 1:30	35
	2s R	GTCRCTGAGATCCAAATATTCTA (23 bp)					
	2p2 F	CTCAGSGAMAATYTRATGGTT (21 bp)	936 bp				
	2p2 R	AAYRAAGTCCCGCTTGTG (18 bp)					
	2x F	GATGGAAGTTTTAGCGAACTTGTG (24 bp)	854 bp	94°c 0:40	56°c 0:40	72°c 1:30	40
	2x R	TCCCGAGTGAAAGACAGGAAT (21 bp)					
	2fin F	GCTCCTGTGAATTCCTGTCTTTC (23 bp)	822 bp	94°c 0:40	64°c 1:10	72°c 1:30	35
	2fin R	GCCGTGTCAGAATAAGCTACCAC (23 bp)					
TLR4	4e2 F	ATGTGATCATTGCAGGTGGTT (21 bp)	178 bp				
	4e2 R	TGACATAAGAAYGAARYGATAGAAC (25 bp)					
	4p1 F	CTGTATCTCTTTTCCCTRTAGGTGTGA(27 bp)	1110 bp	94°c 0:40	56°c 1:10	72°c 1:30	35
	4p1 R	GTTTAGGAAGCGAGCCAAATG (21 bp)					
	4p2 F	TTACCTTGATATTTCTTATAC (21 bp)	838 bp	94°c 0:40	50°c 0:40	72°c 1:30	40
	4p2 R	GAANCCTTCCTGGATGAT (18 bp)					

Legend: Td= Denaturing Temp., Ta= Annealing Temp., Te= Extension Temp., Cy= amount of cycles. Primer 2p2 was later rejected (discussed in detail below) and does not form part of the final sequencing cohort. Primer 4e2, (exon 2 TLR4) was not used in final analysis; settings adequate for amplication of all samples were not reached.

Gels were imaged using the programme Gene Snap by Syngene. Lighting was set at 200ms. Photos were saved as tiffs, and brought into photoshop which was used to count the distance in pixels of each hyperladder band from the wells, as well as the amplified bands. The known and measured lengths were recorded and entered into a Matlab function (Figure 4) which resulted in a log-lin graph. This created best fit equations which were used to estimate amplicon size. Amplicons were then pipetted into a labelled 1.5ml Eppendorf, and frozen for sequencing. Sequencing was performed via SourceBioscience. Primers were prepared by pipetting 18µl sterile distilled water and 2µl primer in one direction into a 1ml labelled tube. Both strands of each amplicon were sequenced separately in order to ensure due diligence and confirm sequences were fully complementary FinchTV 1.5.0 by Geospiza was used for sequence analysis. Sequences were collated and run through Multiple Sequence Alignment. Sequence graphics were checked by eye at each peak. Double peaks, closely matching in height, yet around half the height of the homozygote peaks surrounding them, indicated the possible presence of a heterozygote. Samples underwent a minimum of three rounds of PCR and sequencing to confirm observations. Polymorphic homozygotes also occurred, these were seen when a base was different than the consensus sequence at that data point.

Figure 4. Matlab function to create log-lin graph and line of best fit equation

```
>> x=[*];                    *Distance in pixels, separated by semi-colons
>> y=[1500;1000;800;600;400;200];        #Hyperladder bands
>> figure
>> semilogy(x,y,'+')
>> grid
>> xx=x(1):x(**);        ** number of Hyperladder bands
>> P1=polyfit(x,log(y),1);
>> yy=exp(polyval(P1,xx));
>> hold
>> semilogy(xx,yy,'r')
>> Y1 = exp(polyval(P1,***))        *** gives predicted bandsize
>>P1                        Gives b and c in closest fit equation
```

Bioinformatic databases were used to search for consensus sequences across species: PubMed-NCBI and ENSEMBLE for nucleotide sequences and UniProt for amino acid sequences. Methods of using these databases are somewhat self-explanatory, for example, for PubMed-NCBI, “Nucleotide” was chosen by the search bar. The gene name was typed in the search box which brought up lists of sequences. “Carnivores” were selected so as to narrow species choices. This gave a list and hits referring to the TLR4 genes in Carnivores, where sequences, including both exons and introns, could be obtained. For ENSEMBLE and UniProt, queries were written in the search box, and relevant species were selected, which gave further sequences. Clustal Omega was used for both Multiple Sequence Alignment and Phylogenetic comparison. This returned an alignment based on Nearest Neighbour matrices. It was then possible produce a cladogram using these programs which was returned in PHYLIP code, which was then copied into FigTree software to produce phylogenetic trees.

3.0 Results

3.1 Laboratory Analytical and Refinement Processes

The majority of samples in this study were provided by Alexandra J. Tomlinson and Richard J. Delahey, forming a core set of samples from seventy-nine badgers based at Woodchester Park. DNA from this core sample set was extracted by Hadil Alkathiry. These were tested for the presence of trypanosomes by Eze Justin Ideozu. Supplementary to this, it was decided that outgroup samples would be useful so as to compare data on a wider geographical basis. Eamonn Gormley of University College Dublin provided samples from a population of Irish badgers. Sara Cowen and Elizabeth Mullineaux of Secret World Wildlife Rescue in Gloucestershire provided samples from across the midlands and southern England. The project supervisor, Geoff Hide, provided tissue from a roadkilled badger. The Woodchester Park set had been tracked annually, and whereabouts of both their birth sett and current location was available. Of the seventy-nine, only fifteen were found at a new location. Table 3 gives badger locations. A map of the park area is shown in figure 5.

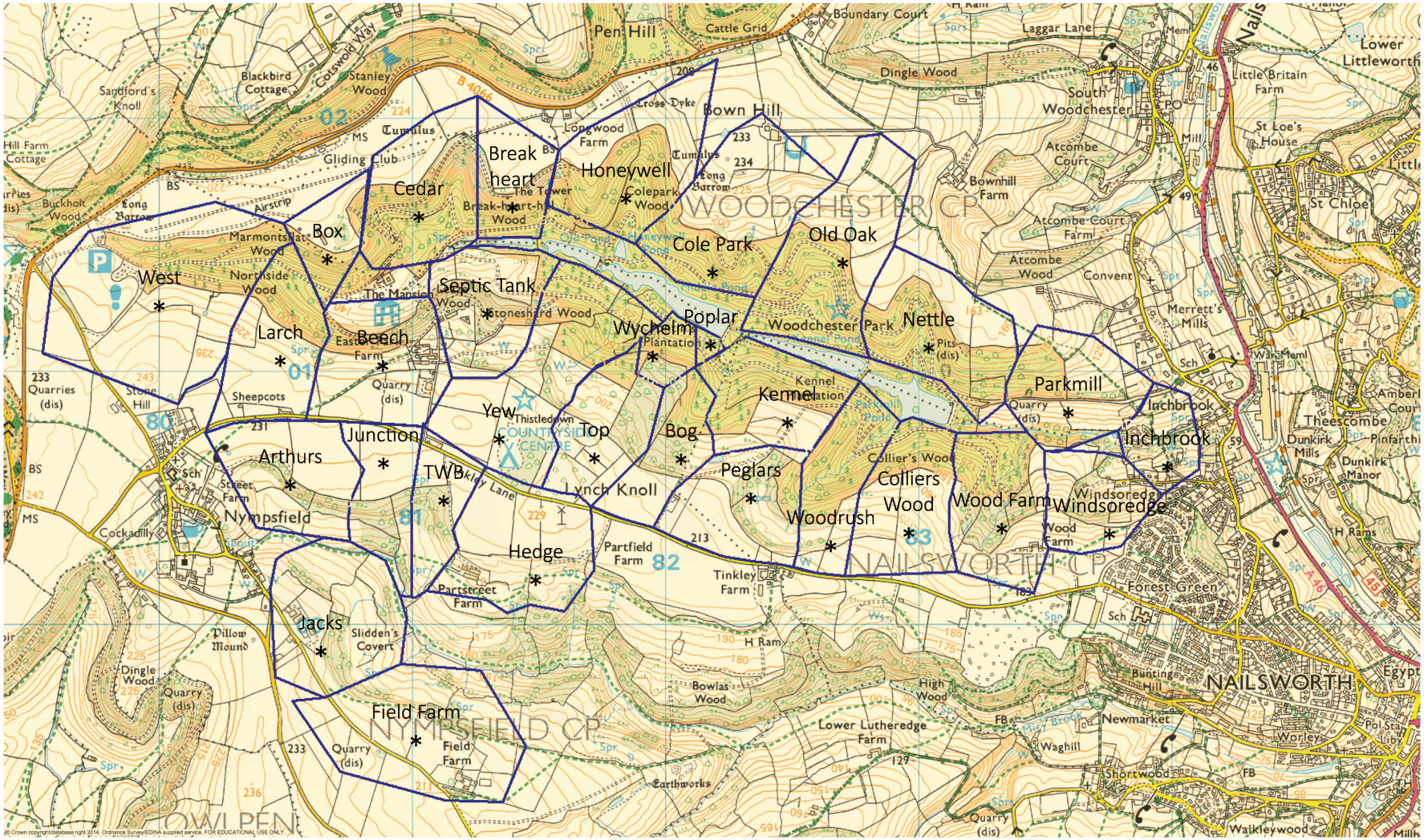
Table 3. Badger Locations within Woodchester Park (WP)

Non-movers	
Sett	Badgers
Beech	13N
Cedar	23N, 26K, 30K, 31K, 33K
Colliers Wood	45R
Honeywell	13F, 24K, 26N, 27K, 29K, 32K, 34K, 37K, 38L, 39K, 40K
Inchbrook	1L, 2F, 25K, 25N, 41K
Jacks	6L, 12L, 50N
Junction	4K, 9N, 10N, 66B
Kennel	12K, 17K, 45P
Larch	31L, 41L, 82W
Mead	14N
Nettle	58R
Old Oak	29N, 32N, 35K, 36K
Park Mill	5L, 38K, 38Y, 59B, 63Y
Scotland Bank	24P, 28K
West	5K, 8L, 9L, 34P, 67Y

Non-movers		
Sett	Badgers	
Windsoredge	21K, 22K, 23K, 33Y	
Wood Farm	13K	
Woodrush	9P	
Wychelm	10K	
Yew	11K, 24L	
Movers		
Home	Found	Badgers
Box	Larch	8I,40L
Boxwood	Breakhart	1K
Breakhart	Boxwood	2K
Chestnut	Beech	3K, 13L, 14L, 15L, 17L, 35P
Evergreen	Cole Park	28F
Mead	Arthurs	2P, 7K
Poplar	Wychelm	15K
Windsoredge	Hopper	17F

Legend: Non-movers indicates badgers who were both born, and found as adults, in the same location. Movers indicates badgers who moved from their place of birth. The shading indicates sett names provided in Excel data which do not correlate to names on the map; there are likewise eight new names on the map not on the spreadsheet. Efforts to obtain this information remain unsuccessful.

Figure 5. GIS map of Woodchester Park showing sett names and boundaries.



Legend: Scale is shown by blue grid lines, each square= 1km^2. Backdrop reproduced from the 2008 Ordnance Survey map © Crown Copyright 2008.

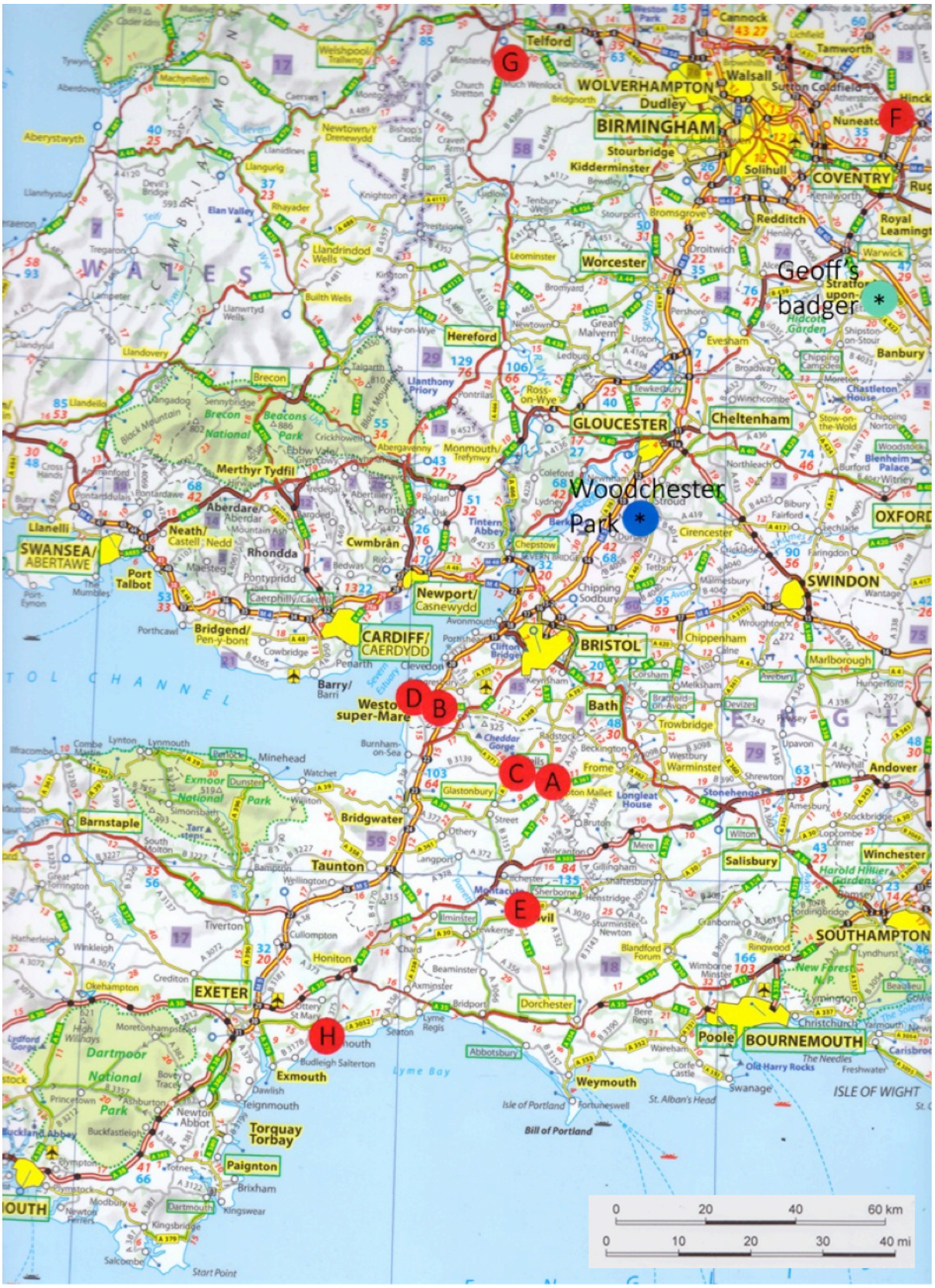
The twenty badger samples donated by Secret World Wildlife Rescue (SWWR) were all cubs orphaned following recent culls and flooding, but were in good condition except where noted in Table 4. Locations were mapped against the Woodchester Park study group, giving seven distinct new locations for this trial, seen in Figure 6. Seven badger samples were donated by the University College Dublin, from badgers raised on site as part of their academic observations, which have a shared familial history within the same location. The three sibling cohorts are listed in Table 5. The roadkill Badger 'G' was found in a mangled, decomposed, state which made sexing the animal impossible. Samples of its hide were taken for DNA extraction. Badger G were found close to Leamington Spa, with latitude and longitude 52°11'25.8"N 1°32'39.0"W. This location has been added to the map in Figure 6.

Table 4. Recording details for Secret World Badger Samples

ID	Location	Distance from WP	On map		Age	Gender	Weight	Details
13	Shepton Mallet	43.7 miles S	A	Within 5.9 miles	16 weeks	Male	4.82kg	
8	BA5 1QQ	46.4 miles S	C		14 weeks	Male	3.94kg	
20	BA5 1QS	46.7 miles S			7 weeks	Male	Not given	
4	Locking BS24 8PL	44.3 miles SW	B	Within 4.5 miles	12 weeks	Both M	Both 2.7kg	Siblings *
1								
5	Weston Super Mare	46.7 miles SW	D		16 weeks	Both M	4.28kg	Siblings
10							4.36kg	
11	BA21 4PW	62.8 miles S	E		12 weeks	Female	3.5kg	
17	CV11 6LA	76.3 miles NE	F		16 weeks	Male	Not given	Siblings
15						Male		
7						Female		
6	Shropshire	76.8 miles N	G		11 weeks	Male	Not given	*
16	Shropshire				8 weeks	Male	2.99kg	Siblings
12					10 weeks	Male	2.85kg	
19	Venn Ottery, EX10 9EX.	95.0 miles SW	H		8 weeks	Male	1.75kg	
14					12 weeks	Male	4.64kg	
9	St. Ives TR26 2JB	202.0 miles SE	I		8 weeks	Male	2.46kg	
2	Not Given	Unknown	Not possible to map		14 weeks	Male	Not given	Siblings
3					12 weeks	Male		
18					12 weeks	Male		

Legend: SWWR badgers are grouped into 8 groups according to geographical distribution, marked by alternation of the colours yellow and green. "*" indicates found emaciated next to deceased mother (culled).

Figure 6. Map showing relative geographical location of outgroup badgers.



Legend: Red labels correspond to locations in figure X by alphabetical reference, Blue=WP, Green= Badger G. The map gives only a sense of the breadth of geographical distribution and is cropped for clarity, there were also samples from St. Ives (off map.)

Table 5. UCD Badger sample details

ID	Age (yrs)	Gender	Weight	Details
21	2	F	7.7kg	Siblings
22	2	F	8.3kg	B. 2012
21	1	F	8.5kg	Siblings
24	1	M	8.8kg	B. 2013

ID	Age (yrs)	Gender	Weight	Details
25	2	M	9.2kg	Siblings
26	2	M	9.7kg	B. 2012
27	2	F	8.6kg	

DNA extraction from was performed after practice on a non-project related primate blood source, Sample J, which yielded 290.8 ng/µl DNA with a purity of 1.81/2 and was deemed satisfactory. Samples 1 through 20 comprised 400µl to 800µl red blood cells per tube. Additional extraction attempts were made on 2, 3, 6, 19 and 20 after first attempts yielded low DNA levels. No sample remained for re-extraction after the first attempt on 10. Badger 20 was excluded from the trial as DNA remained impure following further extraction and clean up. Dilutions were made to reach a concentration of 100ng/µl. Extraction protocols for badger samples 21 to 27 were optimized to extract from the 40µl white blood cells per sample, using 100µl lysis buffer, with succeeding stages performed with 90µl phenol-chloform per step. While this gave good purity ratings, DNA yields were low (Table 6, left). Prior to extraction from Badger G two trials were performed using store-bought beef. This resulted in strong streaking when run on agarose gel and spectrophotometry testing gave purity readings of 1.93/2 and 1.94/2 and concentrations of 1469.1 µg/ml and 1687.0 µg/ml.

Tubulin PCR testing of extracted DNA was performed. Tubulin is a highly repeated gene found in multiple copies throughout the genome. A test for tubulin indicates that DNA is capable of amplification, lacking inhibitors that may prevent this (Terry et al 2001). The primers and PCR settings used were as follows: Primer MtubF (5' -CGTGAGTGCATCTCCATCCAT-3') and MtubR (5' -GCCCTCACCCACATACCAGTG-3') underwent PCR at 94^{O}C for 5 minutes, 40 cycles of 94 ^{O}C for 50 seconds, 55 ^{O}C for 1 minute and 72 ^{O}C for 1 minute 30 seconds, then 72 ^{O}C for 10 minutes. This was performed in order to ensure that the DNA is viable for amplification with untested primers. However, all extractions from badger G failed initial PCR testing using tubulin primers. Spectrophotometery (Table 6, right) showed badger G DNA was highly contaminated, but had good concentrations of DNA. Clean up of all five extractions combined achieved higher purity (1.55) and a concentration of 513.1 µg/ml, and were retained for PCR.

Table 6. Results of Spectrophotometer purity and quantity test for outgroup badgers

Badger G only		
Ex.	Purity n/2	Conc. µg/ml
1	1.39	718.3
2	1.40	323.4
3	1.42	574.2
4	1.42	476.7
5	1.31	768.6
All	1.55	513.1

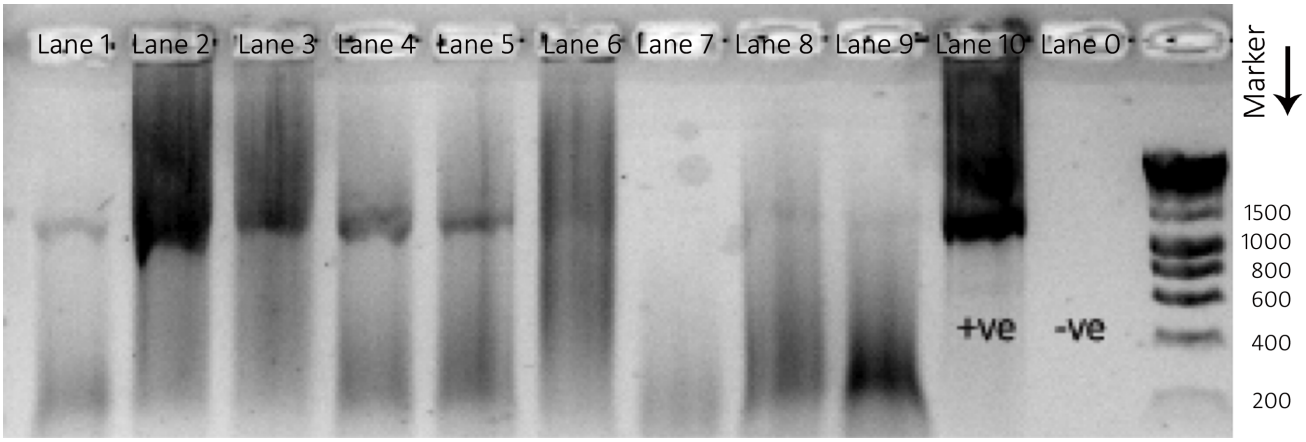
ID	Purity n/2	Conc. µg/ml
1	1.89	361.3
2*	1.65	133.6
3*	1.77	327.3
4	1.81	258.1
5	1.81	224.6
6*	1.85	334.3
7	1.82	229.8
8	1.79	167.9
9	1.85	298.9
10	1.37	128.8
11	1.82	203.4
12	1.75	100.4
13	1.79	127.6
14	1.78	170.4

ID	Purity n/2	Conc. µg/ml
15	1.80	155.0
16	1.78	134.0
17	1.78	149.1
18	1.88	476.8
19*	1.87	124.4
20**	1.55	10.3
21*	1.74	40.7
22*	2.0	2.3
23*	1.7	0.6
24*	1.71	18.4
25*	1.74	5.6
26*	1.79	52.1
27*	1.38	0.4

Legend: Samples highlighted in red were excluded from the PCR stage due to extreme low purity or concentration. * indicates additional attempts at extraction. Ex= Extraction number for Badger G. All= Result of clean up of all 5 badger G extractions combined.

To find the optimal DNA concentration in preparation for the PCR stage, two fold serial dilutions of the second beef steak extraction were made and tested using a Tubulin PCR. The best result, seen in Figure 7, was produced at around 100 µg/ml, because this dilution produced the strongest band with the greatest similarity to the strength of the positive. This was taken as the optimal concentration and all extracted DNA was diluted to this level.

Figure 7. Beef DNA dilution experiment for tubulin PCR.



Legend: Lanes 1-9 are tubulin primer PCR of the beef steak extraction in decreasing concentration; Lane 1=210.8 µg/ml, 2=105.4 µg/ml, 3= 52.7 µg/ml, 4= 26.36 µg/ml, 5= 13.18 µg/ml, 6= 6.59 µg/ml, 7= 3.29 µg/ml, 8= 1.65 µg/ml, 9= 0.82 µg/ml, 10= positive tubulin control, 0= negative control.

In order to design primers to select targets for amplification in the badger, known TLR sequences from other species were collated. The following databases were investigated for published badger TLR sequences, or a whole badger genome, and all TLRs in all species: ARNIE, ARSA, BLAST, Bioinformatic Harvester, Cosmic, DECIPHER, DGVa, DNA Databank of Japan, DRASearch, EBI, EBI-metagenomics, ENCODE, Ensembl, euGenes, European Genome Archive, EXOMISER, Gene Cards, GeneDB, GLIDERS, Google, GWAS, MEROPS, PubMed-NCBI, PATRIC, Pfam, Phenodigm, Rfam, Sanger Centre, Source (Stanford,) Quantomics, Tiffin, Treefam, TXSearch, Vega, Wormbase, and ZF-models. Of these, PubMed-NCBI and ENSEMBL held the greatest abundance of TLR sequences for different species. Cosmic, EBI, and VEGA held some additional, limited, TLR data. None of the databases bore a TLR sequence for the badger, or the full badger genome at the time of searching (September, 2013). Relevant TLR sequences for all available mammals were searched, and over forty were found with sequences available for TLR2 and TLR4; *Mustela putorius furo, Odobenus rosmaus divergens, Pteropus alecto, Myotis brandtii, Equus asinus, Equus caballus, Orcinus orca, Lageno rhynchus obliquidens, Tursiops truncatus, Bos indicus, Bos taurus, Bison bison, Bubalus bubalis, Cerco cebusatys, Macaca fascicularis, Otolemur garnettii, Loxodonta africana, Pantroglodytes verus, Hylobateslar, Callithrix jacchus, Sorex araneus, Homo sapiens, Pan paniscus, Papio anubus, Cricetulus griseus, Felis catus, Gorilla gorilla, Pongo pygmaeus, Ailuropoda melanoleuca, Sus scrofa, Boselaphustrago camelus, Myodes, Eothenomys, Arvicola, Microtus, Chionomys, Micromys, Apodemus, Canis familaris, Mus musculus, Rattus norvegicus, Condylura cristata* and *Capra hircus*. The most relevant were the ferret (*Mustela putorius furo*), cat (*Felis catus*), dog (*Canis familaris*), panda (*Ailuropoda melanoleuca*) and walrus (*Odobenus rosmaus divergens*) which are all found in the order Carnivora along with the badger. Having gathered sequences of interest, there was scope to produce Multiple Sequence Alignments.

First, it was important to identify the targetted proteins so as to differentiate exons and introns. Exons form the DNA code which results in the final protein configuration, while introns do not code for proteins. Since the targets were TLR proteins, the exons were of greatest interest and would form the target of any primers designed. A Matlab code was written in order to convert strings of base pairs of any length into their final protein sequence

(Figure 8). This was then used to compare nucleotide sequences to protein sequences found in online databanks and thus identify exons and introns.

Figure 8. 'DNAtoProtein' Matlab function (written by Andrew Whiteoak.)

```
function AminoAcidSequence=DNAtoProtein(bpSequence)
N=floor(length(bpSequence)/3); M=N*3;
Threesomes=reshape(bpSequence(1:M),3,N)';
Result=zeros(1,N);
Result(strmatch('AAA',Threesomes))='K';
Result(strmatch('AAG',Threesomes))='K';
Result(strmatch('AAC',Threesomes))='N';
Result(strmatch('AAT',Threesomes))='N';
Result(strmatch('ACA',Threesomes))='T';
Result(strmatch('ACC',Threesomes))='T';
Result(strmatch('ACG',Threesomes))='T';
Result(strmatch('ACT',Threesomes))='T';
Result(strmatch('ACN',Threesomes))='T';
Result(strmatch('AGA',Threesomes))='R';
Result(strmatch('AGG',Threesomes))='R';
Result(strmatch('CGA',Threesomes))='R';
Result(strmatch('CGC',Threesomes))='R';
Result(strmatch('CGG',Threesomes))='R';
Result(strmatch('CGT',Threesomes))='R';
Result(strmatch('CGN',Threesomes))='R';
Result(strmatch('AGC',Threesomes))='S';
Result(strmatch('AGT',Threesomes))='S';
Result(strmatch('TCA',Threesomes))='S';
Result(strmatch('TCC',Threesomes))='S';
Result(strmatch('TCG',Threesomes))='S';
Result(strmatch('TCT',Threesomes))='S';
Result(strmatch('TCN',Threesomes))='S';
Result(strmatch('ATA',Threesomes))='I';
Result(strmatch('ATC',Threesomes))='I';
Result(strmatch('ATT',Threesomes))='I';
Result(strmatch('ATG',Threesomes))='M';
Result(strmatch('CAA',Threesomes))='Q';
Result(strmatch('CAG',Threesomes))='Q';
Result(strmatch('CAC',Threesomes))='H';
Result(strmatch('CAT',Threesomes))='H';
Result(strmatch('CCA',Threesomes))='P';
Result(strmatch('CCC',Threesomes))='P';
Result(strmatch('CCG',Threesomes))='P';
Result(strmatch('CCT',Threesomes))='P';
Result(strmatch('CCN',Threesomes))='P';
Result(strmatch('CTA',Threesomes))='L';
Result(strmatch('CTC',Threesomes))='L';
Result(strmatch('CTG',Threesomes))='L';
Result(strmatch('CTT',Threesomes))='L';
Result(strmatch('CTN',Threesomes))='L';
Result(strmatch('TTA',Threesomes))='L';
Result(strmatch('TTG',Threesomes))='L';
Result(strmatch('GAA',Threesomes))='E';
Result(strmatch('GAG',Threesomes))='E';
Result(strmatch('GAC',Threesomes))='D';
Result(strmatch('GAT',Threesomes))='D';
Result(strmatch('GCA',Threesomes))='A';
Result(strmatch('GCC',Threesomes))='A';
Result(strmatch('GCG',Threesomes))='A';
Result(strmatch('GCT',Threesomes))='A';
Result(strmatch('GCN',Threesomes))='A';
Result(strmatch('GGA',Threesomes))='G';
Result(strmatch('GGC',Threesomes))='G';
Result(strmatch('GGG',Threesomes))='G';
Result(strmatch('GGT',Threesomes))='G';
Result(strmatch('GGN',Threesomes))='G';
Result(strmatch('GTA',Threesomes))='V';
Result(strmatch('GTC',Threesomes))='V';
Result(strmatch('GTG',Threesomes))='V';
Result(strmatch('GTT',Threesomes))='V';
Result(strmatch('GTN',Threesomes))='V';
Result(strmatch('TAA',Threesomes))='X';
Result(strmatch('TAG',Threesomes))='X';
Result(strmatch('TGA',Threesomes))='X';
Result(strmatch('TAC',Threesomes))='Y';
Result(strmatch('TAT',Threesomes))='Y';
Result(strmatch('TGG',Threesomes))='W';
Result(strmatch('TGC',Threesomes))='C';
Result(strmatch('TGT',Threesomes))='C';
Result(strmatch('TTC',Threesomes))='F';
Result(strmatch('TTT',Threesomes))='F';
Result(Result==0)='*';
AminoAcidSequence=char(Result);
End
```

In order to infer the position of exons and introns, the TLR2 and TLR4 protein sequences for the closest relatives of the badger; the ferret, cat, dog, panda and walrus (most closely related due to being of the order Carnivora) were compared to their nucleotide sequences. To this end it was found that TLR2 has a single, unbroken, 2355 bp exon (Figure 9), and TLR4 has three exons and two introns (Figure 10). This was an incredibly useful venture because it allows the researcher to make useful decisions as to where to design primers. Clearly, for TLR2, it is a simply a matter of crawling along the single exon searching for areas of high

sequence conservation between the collection of sequences. While for TLR 4, avoidance of intron spaces between the three exons would be crucial.

Figure 9. The overall nucleotide structure of TLR2

Figure 10. The overall nucleotide structure of TLR4

Multiple Sequence Alignments (MSA) allow researchers to see where nucleotides or proteins are conserved across species. These regions indicate possibilities for primer design for unknown species sequences due to their higher probability of their also being conserved in the same manner. Current bioinformatics researchers recommend Clustal software as the fastest (Sievers et al 2011) and most sensitive method (Soding, 2005) of performing MSAs, outperforming other packages (Stevens et al 1998). It is recommended due to its capacity to align any number of protein sequences using modest amounts of memory (Blackshields et al 2010). To sequence large regions of DNA, it is necessary to subdivide a large fragment into smaller ones to be individually sequenced (Slatko et al 1999). A full protein alignment for TLR4 for 46 species was made, which showed that the minke whale *Balenoptera acutostrata*, Omura's whale *Balaenoptera omuraim*, the Chinese white dolphin *Sousa chinensis*, the killer whale *Orcinus orca*, the common dolphin *Delphinus capensis* and the bottle nosed dolphin *Tursiops truncates* do not carry exon 1 or exon 2, but begin their protein sequence at exon 3. Alignments amongst the order Carnivora were used to identify conserved sequences, so that primers could be drawn from them. Conserved areas were identified from which primers were designed to cover all of TLR2 and exons 2 and 3 of TLR4. A Computer Code was created to give reverse primer sequences which was highly useful during the primer design stage. Forward primers are read straight from a given sequence, in a 5' to 3' direction. However, reverse primers are the reverse complement of a given sequence. A Matlab function was created to perform this task (McMahon, 2007). The code is shown in Figure 11. Letters other

than GATC indicate base pair redundancy (degenerate bases) and their reverse complement. Using this function researchers can copy-paste in the sequence where a primer possibility lies, and have Matlab output the sequence as a reverse primer.

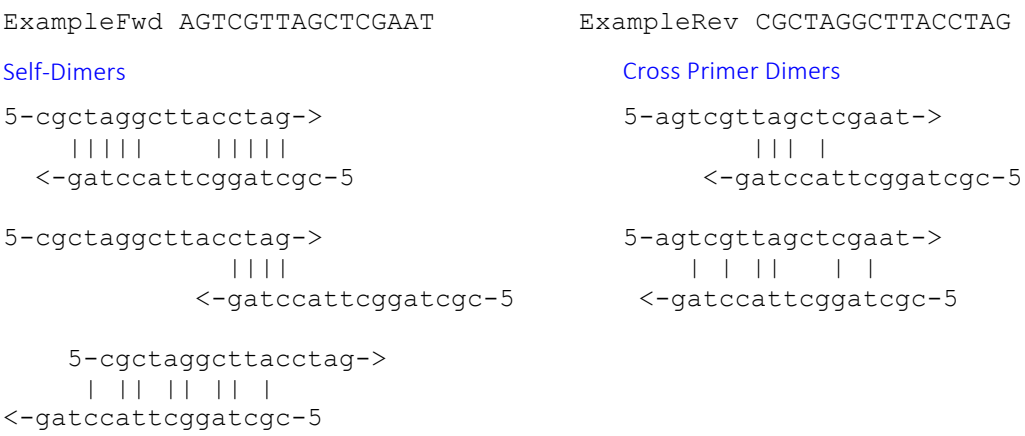
Figure 11. Function for creating Reverse Primers in Matlab (written by Whiteoak, 2013).

```
function Comp = ReversePrimer(Primer)
%This function gives the reverse primer including degenerate bases
b=fliplr(Primer);
%fliplr(Primer) reverses the direction of the forward primer but does not complement it.
Comp(b=='A')='T';
Comp(b=='T')='A';
Comp(b=='C')='G';
Comp(b=='G')='C';
Comp(b=='B')='V';
Comp(b=='V')='B';
Comp(b=='D')='H';
Comp(b=='H')='D';
Comp(b=='K')='M';
Comp(b=='M')='K';
Comp(b=='R')='Y';
Comp(b=='Y')='R';
Comp(b=='S')='W';
Comp(b=='W')='S';
Comp(b=='N')='N';
End
```

Lvovsky et al (1998) discussed the effect of secondary structure on very small primer lengths with only eight bases (8-mers) and two degenerate positions. Degenerate bases are manufactured replacements for single nucleotides which can code for more than one nucleotide and therefore attach to more than one possible complementary. These short 8-mer primers were significantly affected by local structure, and hence the folding energy of the primers, leaving them more highly prone to dimerisation and requiring strong PCR optimization with respect to temperature and cycling to improve their efficiency. Dimerisation causes the amplification of non-target lengths of amplified DNA which form when primers complement one another and join together, amplifying themselves rather than the target band. Even when the target band is amplified, the dimer creates contamination which will hamper sequencing results, and thus must be avoided. This can occur when forward and reverse primers complement, or when either the forward or the reverse complements itself. There is method by which primers can be checked manually where the reverse complement of each primer is written by first writing the primer letters out backwards, then writing out their complement. If this is a close match to either the other primer or itself then a primer dimer is likely. The programme created in MatLab 'ReversePrimer' can also perform the former function while the latter is manual. An example of this is given in Figure 12, showing the various dimers given with two hypothetical primers. During primer design, conserved sequences around target areas were investigated until

primer pairs with no dimerization were found. Lvovsky et al (1998) pointed out that this became less of an issue for conventional longer primers that have more selectivity.

Figure 12. Two hypothetical primers and their dimers



Jaric et al (2013) looked at the design of universal primers, in which they combined the standard design rules with an automated database search and optimization to design primers that could be used over a range of species more effectively to identify their presence as well as their prevalence. Here, the bases were optimized to achieve a limited specificity that was able to work efficiently across a large range species of species. By contrast, our research requires a very high degree of specificity. We were concerned about the **Specificity** of primers with a large number of full or partial degeneracies and could find no literature regarding this effect. Hence we developed the following simple combinatorial model to evaluate the effect of degeneracies on the primer's **Specificity.**

Table 7. Standard code for degenerate base specification, paired by complement.

Degenerate base	B	V	D	H	K	M	R	Y	S	W	N
Bases covered		A	A	A		A	A			A	A
	C	C		C		C		C	C		C
	G	G	G		G		G		G		G
	T		T	T	T			T		T	T

All degenerate bases available are show in Table 7. Six of them allow selection to waver between two possible target bases ("x"). Four allow selection to waver between three possible target bases ("y"). One allows selection to waver between all four possible target bases ("z").

An equation was derived to measure relative specificity if the primer length ("L") the the amount and type of degenerate bases are known using the following deduction. Where no degenerate bases are involved, specificity must be a function of length. Let us take, for simplicity, a primer with a length of three nucleotides. There are 64 possible combinations using four base pairs over those three nucleotides; this is a function of 4^3 or 4^L where 4 equals the amount of bases in DNA (G, A, T and C,) and L=primer length. With no degenerate bases, the primer would have a specificity of 1 in 64, as it would target a specific 3 bp sequence from 64 possibilities; and would rewrite as $\frac{1}{4^L}$. If one of the three bases in this primer is a two-way degenerate base, then while there are only 64 permutations of GATC in three positions, there are now two chances of matching one of them. Therefore specificity drops to 2 in 64. Likewise, if two bases are two-way degenerate bases, specificity will drop to 4 in 64. From the deductions above we can see that specificity becomes a function of the power of the amount of variation in the degenerate bases. The final equation was written as:

$$Specificity = \frac{2^x \times 3^y \times 4^z}{4^L}$$

This equation was extended to show that length is more critical for specificity than the ratio of specific nucleotides to degenerate bases. The presence of degenerate bases reduces specificity but not as much as reducing primer length does. To express the reduction in effective sequence length in base pairs as a product of any combination of degenerate bases on an (x,y,z) plot it was useful to know that a four way degenerate base has an equivalent effect on specificity as two two-way degenerate bases, both reduce specificity by ¼. The equation was shown as follows:

$$Specificity = \frac{2^x \times 3^y \times 2^{2z}}{4^L} = \frac{2^{(x+2z)} \times 3^y}{4^L}$$

This gave a function in two variables. Given that $2 = \sqrt{4} = 4^{1/2}$, then,

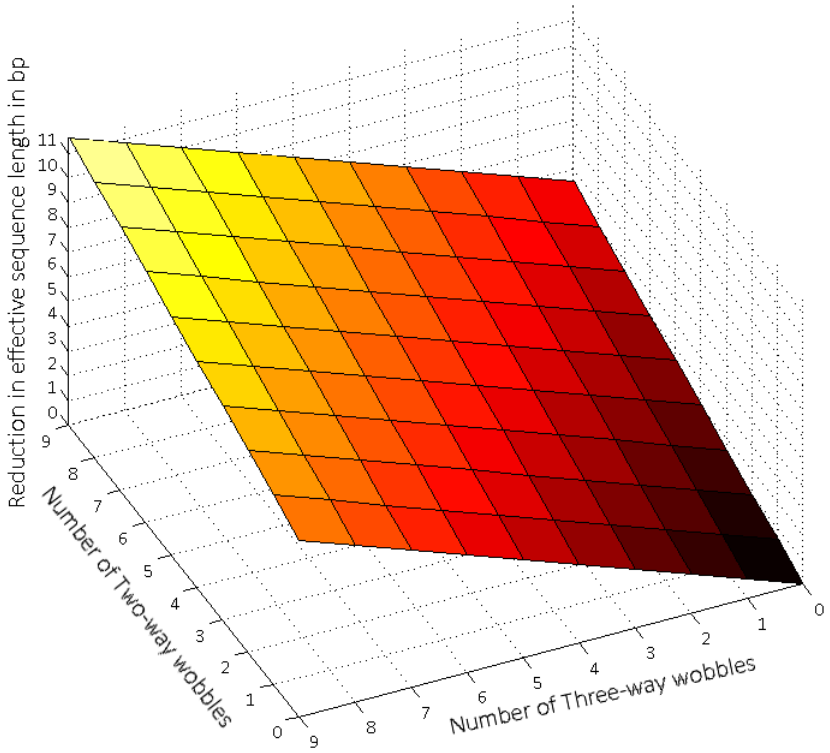
$$Specificity = \frac{4^{\left(\frac{x}{2}+z\right)} \times 3^y}{4^L} = 4^{-[L-\left(\frac{x}{2}+z\right)-log4(3^y)]}$$

Reduction of specificity in effective length in bp = $\frac{x}{2} + log4(3^y) + z$

Figure 13 was plotted using two and three way degenerate bases to show their combined effect on specificity, as expressed by how many base pairs their presence would effectively remove. The reductive slope caused by three-way degenerate bases is much steeper than that caused by two-way degenerate bases, as would be expected. Length has the largest

impact on a primer's specificity. Any increase in sensitivity caused by a degenerate base is far outweighed by simply increasing the nucleotide length of the designed primer. For each additional nucleotide, primers become four times more specific. However, researchers have found that the longer the primer, the smaller the fraction of primed templates will occur in each annealing step, leading to a significant decrease in the volume of final product. Dieffenbach et al (1993) recommend between 18-24 nucleotide length primers.

Figure 13. Graph showing reduction in effective primer length due to degenerate bases.



Legend: The X and Y axes show possible numbers of two-way and three-way degenerate bases, with each cross-hair above either axis being a possible combination of marked amounts of both types. The Z axis takes into account the relative specificity levels of each combination with respect to reduced length; if a primer of 'N' length in base pairs possessed the combination of amounts of degenerate bases shown by the X and Y axis, its specificity would be reduced by 'N-r' where r is the number on the Z axiz. Specificity refers to how specific a primer is, it is the opposite of sensitivity. Highly specific primers will only capture the intended target and thus may miss the target if there are any minor variables; highly sensitive primers will capture the intended target but may also capture a number of unwanted non-targets. The colours go from dark red (most specific) to pale yellow (least specific).

The position of the degenerate base effects efficiency not specificity (Jaric et al 2013) and this equation looks to specificity alone. It is considered useful to have at least three G or C bases at the 3' end, which helps primers bind to their targets and therefore increases yield; in the same respect, degenerate bases grouped at the 3' end should be avoided as this will lower

primer efficiency (http://www.premierbiosoft.com/tech_notes/PCR_Primer_Design.html). Clearly this is only part of the design. Primers also must be designed with regard to primer-dimers, GC content, and melting and annealing temperatures as decribed in the following analysis. An investigation was made into the best methods for Tm calculation. The most commonly given Tm equation was: Tm = 4(G+C)+2(A+T). The primer manufacturer, Eurofins MWG, provided another equation:

$$TM\ [°C]\ = 69.3 + \frac{41(nG+nC)}{L} - \frac{650}{L}$$

Where n = number of nucleosides of type G or C, and, L= number of all nucleotides per sequence (personal communication Wamack, 2013). To compare equations, both were reformulated as functions of the percentage GC content.

The crude equation became: $TM\ [°C]\ = 4(pL) + 2(L - pL)$

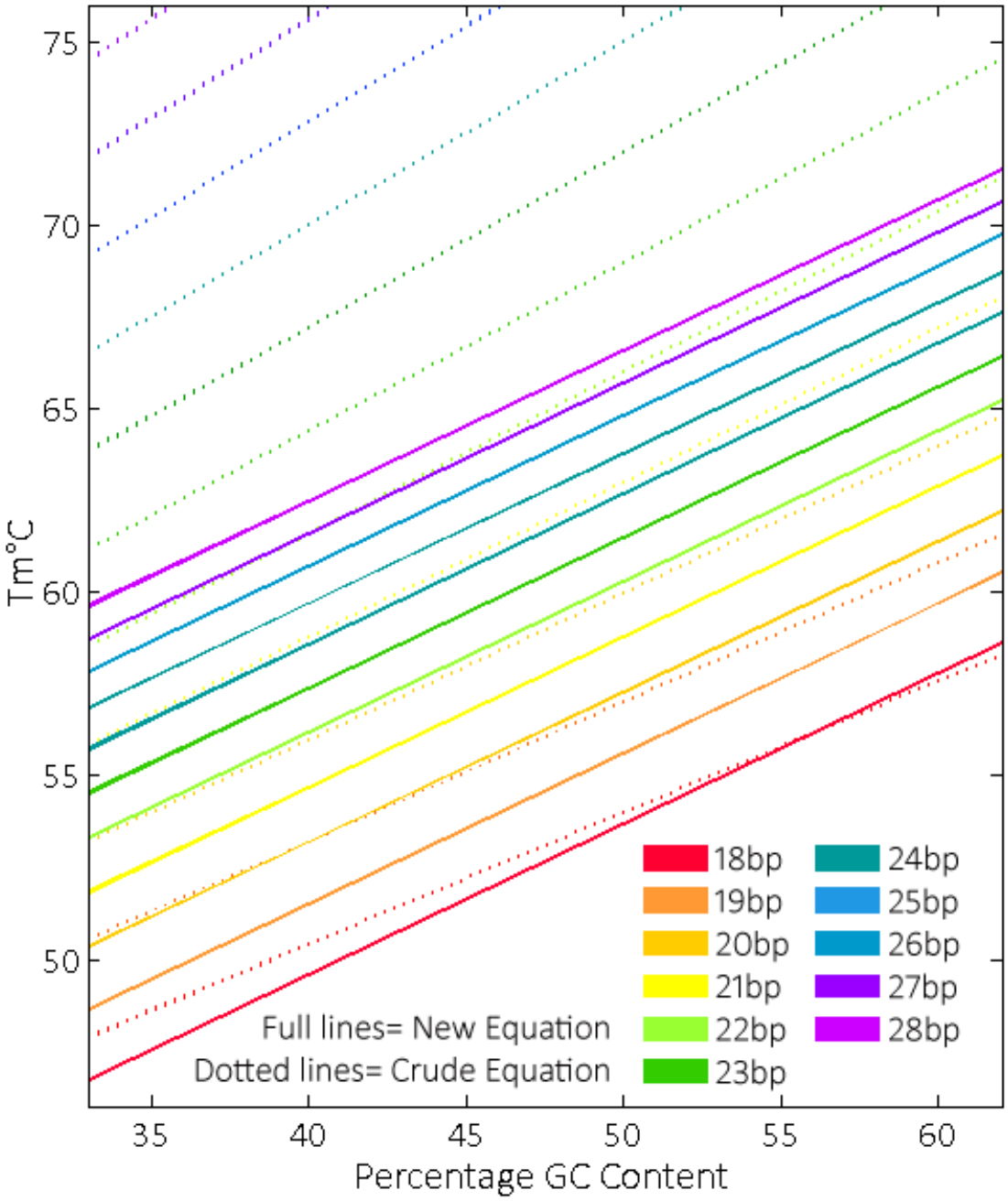
The MWG equation became: $TM\ [°C]\ = 69.3 + \frac{41(pL)}{L} - \frac{650}{L}$

Where, p= the percentage GC content expressed as a decimal e.g. 0.10 = 10% GC content, and, L= Nucleotide length of primer. A comparative graph was plotted (Figure 14). It shows that for the crude equation, the longer the primer becomes, the hotter Tm is predicted to be. This prediction is done in a linear fashion (given it is a linear equation) so that with a length off 23bp and 50% GC content, the predicted Tm is already equal to the normal Extension Temperature, which would make PCR impossible if it were correct. The graph also shows, by the nature of spaces between lines representing the MWG equation in its application to different primer length, that it allows for a logarythmic reduction in Tm with length.

The MWG equation was then plotted alone (Figure 15) to examine the relationship between GC content and length with respect to their effect on Tm according to this equation. We can see the equation has been formulated so that Tm never becomes as high as the extenstion temperature. In this graph it is not length that makes the biggest difference to Tm but GC content. For example at 50% GC content, the 16bp primer (red line) with 8 total G+C, and the 24bp primer (violet line) with 12 total GC, both have a Tm of 60^{o}c even though they are 33% to 50% different in length. Meanwhile a 10% difference in GC content, for example for a 20bp primer with 10 total GC vs a 20bp primer with 9 total GC the Tm changes from 57^{o}c to 55^{o}c. Neither of the above equations take into account the temporal aspect of primer design for PCR, in which one base pair must always first 'unzip' from its complement before the next in

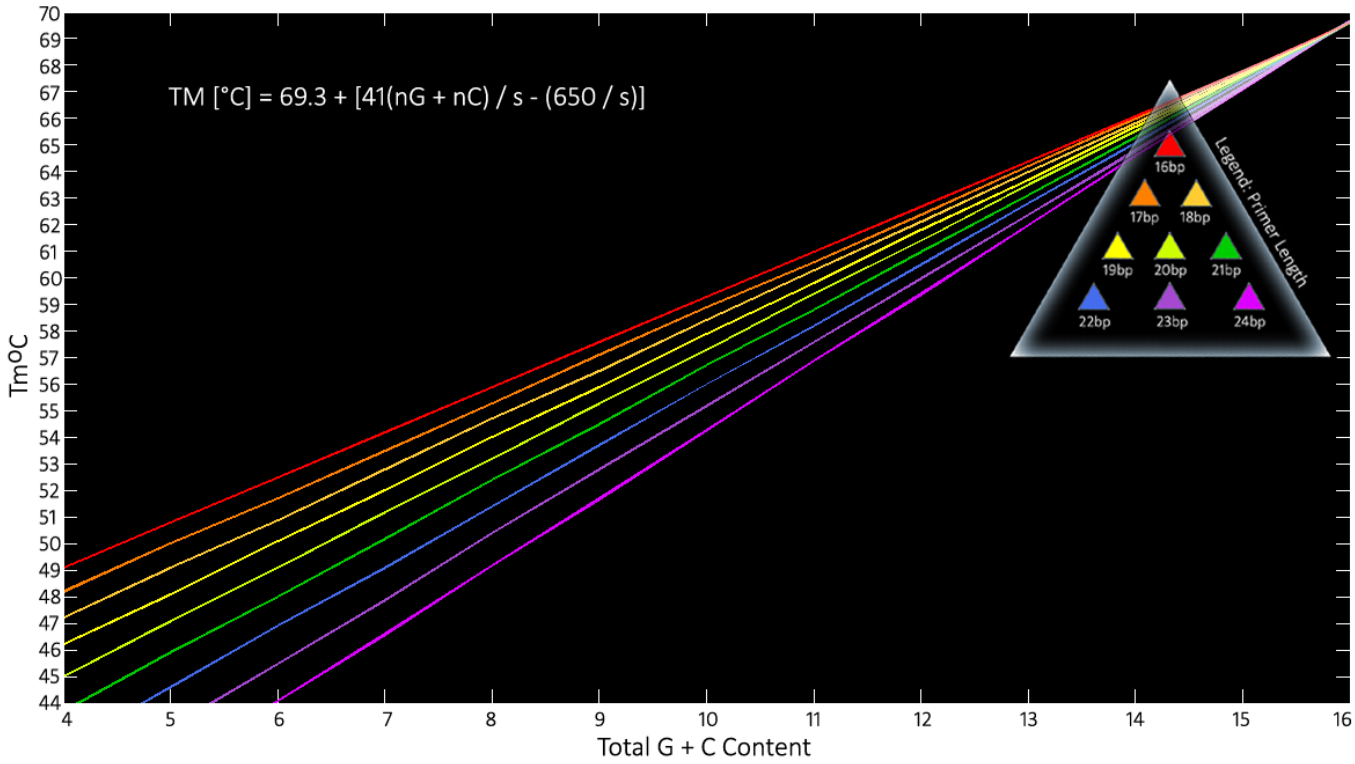
a step wise fashion. Dieffenbach et al (1993) indicated that Tm prediction based on nearest neighbour thermodynamic parameters are slightly more accurate. 'Nearest Neighbour' is an expression of adjacent nucleotide interaction. This uses a boundary element method to calculate melting temperature taking into account that a G next to a T will be different from a G next to an A and so forth. A website was found that performed these calculations (http://www.thermoscientificbio.com/webtools/multipleprimer/).

Figure 14. Graph showing Tm values given by the crude vs. MWG equation.



Legend: Full lines= MWG equation, dotted lines= Crude equation.

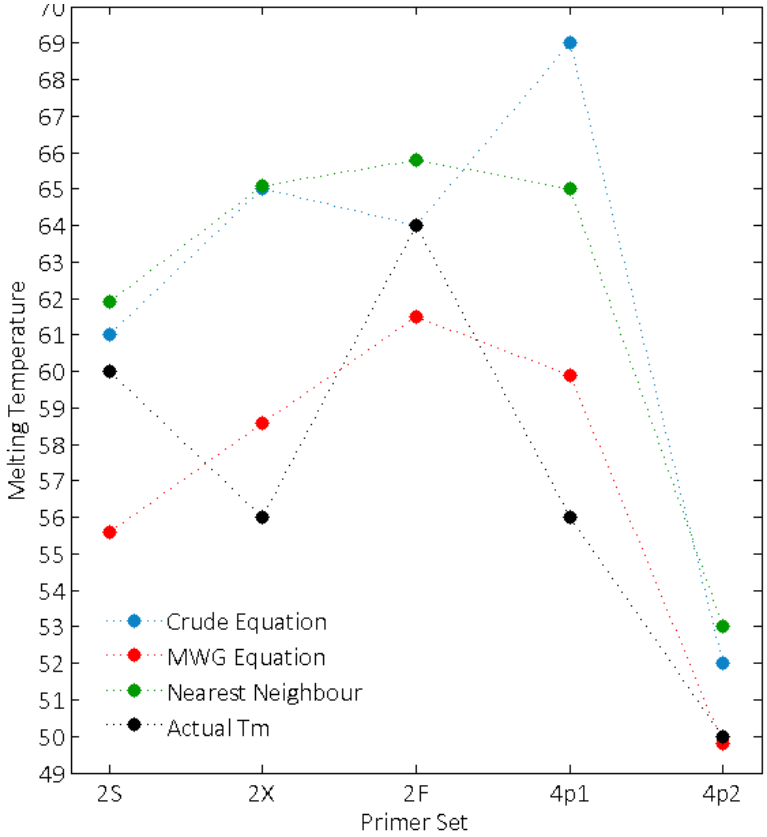
Figure 15. Graph showing Tm vs. GC content and primer length using the MWG equation.



Legend: Lines represent different primer lengths as shown by the key.

Gradient testing in lab found optimal Tm settings for successful PCR. A table of Tm results for all successful primers across all three equations compared to actual Tm results as found by gradient PCR tests is shown in Table 8. The mean of each primer pair was taken as the predicted balanced Tm for each equation and was shown as a graph (Figure 16.) The graph shows that the Nearest Neighbour and Crude equations come out with surprisingly similar estimates for each primer. Points were joined by lines on the graph, in spite of primers being unrelated, in order to capture the overall pattern of predictions. The MWG equations' pattern was consistently in line with the Nearest Neighbour predicts but on average 5°c lower with little variance. Meanwhile, the actual successful Tm of each primer found by gradient PCR did not correspond to the pattern of any equation, veering up and down in Tm result; matching the MWG equation prediction only with primer set 4p2, and the crude equation with primer set 2F, and otherwise adhering to know particular set of predictions by any equation. The actual Tm must therefore be reliant on far wider parameters than any equation can predict, which starkly shows the importance of the Gradient PCR step.

Figure 16. Graph showing comparison of all three equations to actual lab-tested Tm.



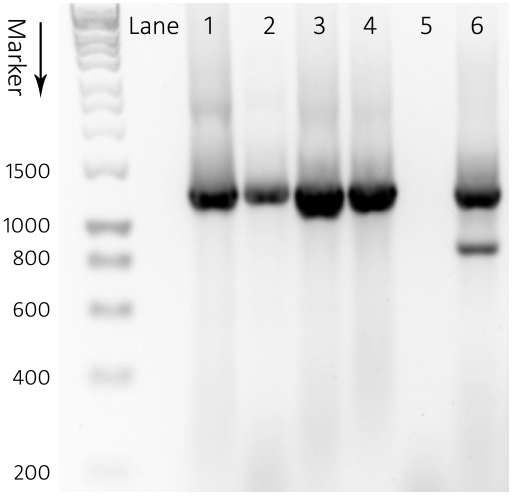
Legend: Each reaction is independent of one another, thus dotted lines do not indicate a connection but simply allow the reader to more easily visualise the dispersal of each separate equation or actual Tm. Some data points overlay one another, the lines show what is otherwise obscured by data sharing the same point.

Table 8. Comparison of equations versus final Tm used in PCR

	Equations for $TM\ [°C]$ and balances between two primer results						Actual Tm
	4(G+C)+2(A+T)	Mean	$69.3 + \frac{41(nG+nC)}{L} - \frac{650}{L}$	Mean	N. Neighbour	Mean	
2s Fwd	60	61	55.9	55.6	64.5	61.9	60
2s Rev	62		55.3		59.2		
2x Fwd	68	65	59.3	58.6	64.9	65.1	56
2x Rev	62		57.9		65.2		
2fin Fwd	68	64	60.6	61.5	65.5	65.8	64
2fin Rev	60		62.4		66.0		
43p1 Fwd	76	69	61.9	59.9	65.0	65.0	56
43p1 Rev	62		57.9		64.9		
43p2 Fwd	52	52	48.1	49.8	47.8	53.0	50
43p2 Rev	52		51.4		58.2		

Once primer design was complete, primers were manufactured by MWG Eurofins. All badger blood samples were tested for capacity to amplify by using a mammalian tubulin PCR to confirm good quality DNA and to check for the presence of PCR inhibitors (Terry et al 2001). Figure 17 shows an example of a successful tubulin run on Woodchester samples 34K, 8i, 15K and 8L, as well as a negative (H2O) and positive control DNA. The positive control in column 6 is taken from a sample of cattle blood. Expected band size varies from mammalian species to species and its size in badgers has not been published. This band size is not easily predictable by bioinformatics because the tubulin genes are a family of genes. Band sizes vary across different mammalian species, however analysis of the gel in Figure 17 shows it is 1301bp. Table 9 shows band sizes vs distance travelled along the gel, this data was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 18. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.00873 and c= 8.44589. The mean error for hyperladder bands was 4.28%.

Figure 17. Example of a successful tubulin PCR across four badgers



Legend: Lane 1= 1K, 2= 2P, 3= 4K, 4= 5K, 5= Negative Control, 6= Positive Control

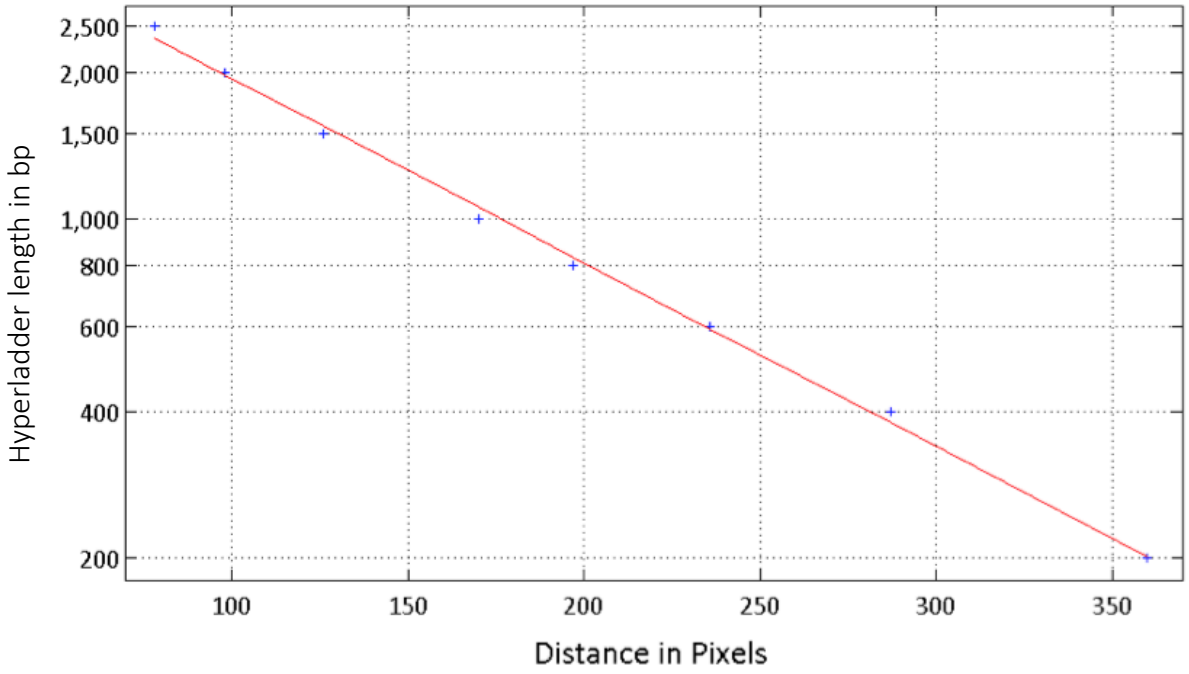
Table 9. Band sizes vs distance travelled along the gel for tubulin PCR.

Hyperladder (bp)	200	400	600	800	1000	BT	1500	2000	2500
Distance (px)	360	287	236	197	170	146	126	98	78
Predicted length by closest fit (bp)	201	380	593	834	1055	1301	1550	1979	2356
% Error	0.50	5.00	1.17	4.25	5.50		3.33	1.05	5.76

Legend: BT= Badger tubulin band

Column 6 has more than one band, which occurs in some but not all species; within the tandemly repeated genes of some species there are some repeats that are different lengths. All Woodchester samples tested positive for tubulin. Of the samples donated by Secret World 1-8, 11- 14, and 16, 18 and 19 tested positive for tubulin. As was expected, samples 10 and 20 failed. Samples 9, 15 and 17 were trialled additional times and later run through all TLR PCRs but failed to produce any results. All samples donated by UCD tested positive.

Figure 18. Standard curve of band size vs distance travelled for tubulin PCR.

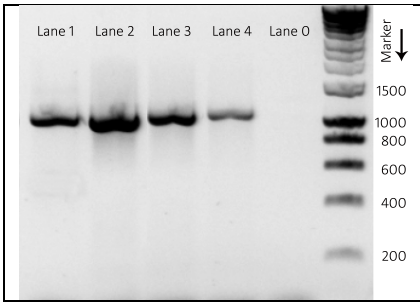


PCR optimisation was performed for all primer sets. TLR 2 was described by three of four redundant primer sets, 2s, 2x and 2 fin. 2p2 was later abandoned in favour of the other primers for reasons described in the following sections. Primer set 2s covered the first 1110bp of the single exon gene from the 5' end. The reverse primer of 2s overlaps the forward primer of 2x; this gene set covers the middle 854bp of TLR2. Primer set 2fin overlaps the region covered by primer set 2x, and covers the final 822bp as well as 79bp of the intron. Primer set 2p2 overlapped the regions covered by the 5' end of primer sets 2s and the 3' end of primer set 2x. Final optimization for all TLR2 primers is discussed here.

According to nearest neighbour calculations primer 2s Fwd had a Tm of 64.5^{o}C while 2s Rev had a Tm of 59.2^{o}C. In order to find which temperature was best for amplification, gradient

PCRs were run between 52°C and 64°C. The strongest bands appeared at the 60°C to 62°C level. A problematic, faded non-target band (slightly larger than the desired band size of 1110bp) was removed by increasing the annealing time to 1:10 (m:s). Presence of this non-target band was further reduced by lowering the cycle time to 35 cycles.

Table 10. PCR amplification using primer 2s and predicted length using the Hyperladder

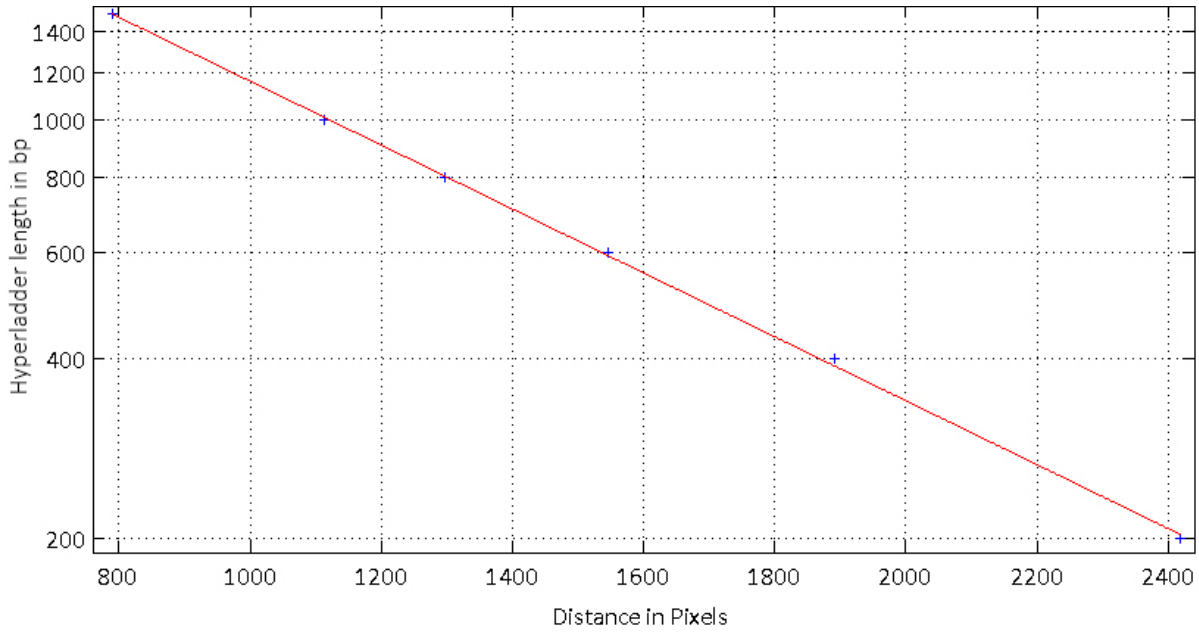


Hyperladder	Distance	Pred.	% Error
1500 bp	791px	1501bp	0.67%
Band 2s Exp. 1110bp	1070px	1066bp	3.60%
1000 bp	1113px	1011bp	1.10%
800 bp	1297px	807bp	0.88%
600 bp	1547px	594bp	1.00%
400 bp	1891px	389bp	2.75%

Legend: Lane 1= 35P, 2= 12L, 3= 27K, 4= 1L, 0= Negative Control, Pred= Predicted Length by Closest Fit

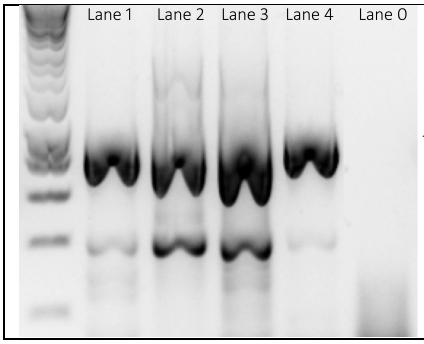
The sample gel photo was measured by counting pixels. This data (Table 10) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 19. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0013 and c= 8.3093. The amplified bands were estimated to be 1066 base pairs, 40 bp more than the desired product. The mean error for hyperladder bands was 1.28%.

Figure 19. Graph of distanced travelled vs. hyperladder size for 2s



According to nearest neighbour calculations primer 2p2 Fwd had a Tm of 61.4°C while 2p2 Rev had a Tm of 61.6°C. In order to find which Tm was best for amplification, gradient PCRs were run between 54°C and 64°C. The strongest bands appeared at 60°C. Using this primer, a non target band appeared no matter what was done to optimise the reaction. To aid in analysis of this contamination, discussed later, the non-target band was isolated by increasing the $MgCl_2$ concentration during PCR. The sample gel photo shown in Table 11 was measured by counting pixels on a high resolution screen. It is clear, given the nature of bands in this gel that accuracy would be questionable, and errors were calculated. This data was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 20. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0012 and c= 8.3856. The amplified bands were estimated to be 915 base pairs, 21 bp less than the desired product. The mean error for hyperladder bands was 0.91%.

Table 11. PCR amplification using primer 2p2 and predicted length using the Hyperladder

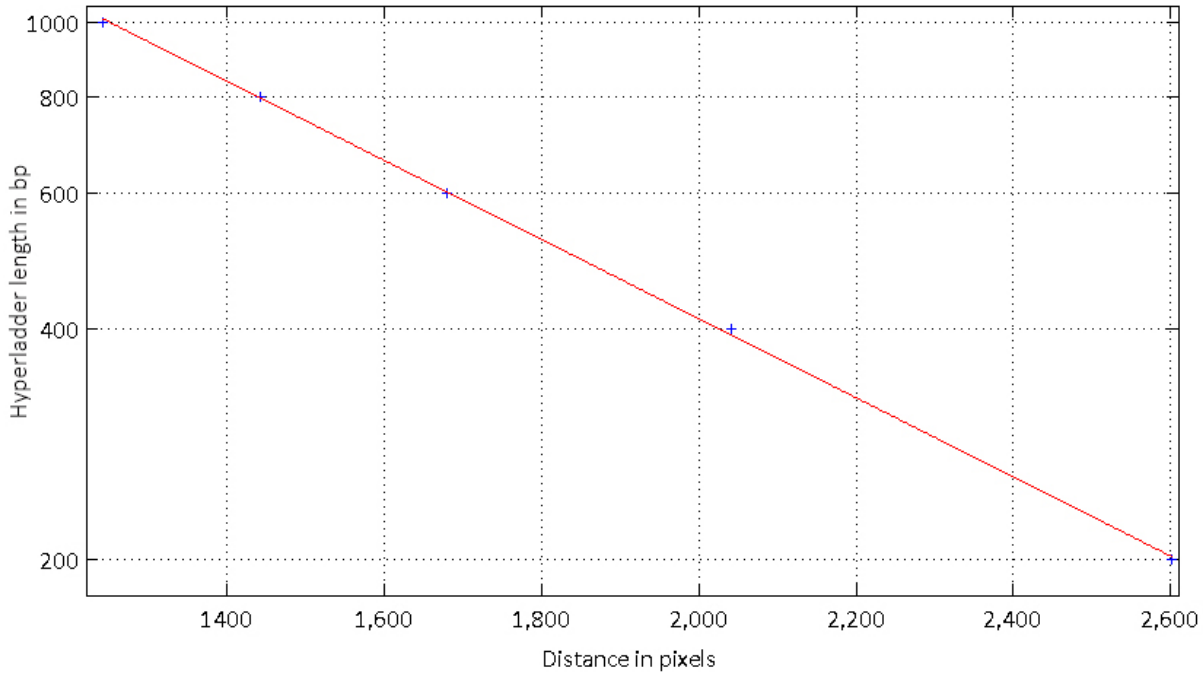


Hyperladder	Distance	Pred.	% Error
1000 bp	1241px	1010bp	1.00%
2p2 (Exp 936bp)	1325px	915bp	2.24%
800 bp	1442px	797bp	0.38%
600 bp	1680px	601bp	0.17%
400 bp	2041px	392bp	2.00%
Non-target	2071px	379bp	
200 bp	2601px	202bp	1.00%

Legend: Lane 1= 3K, 2= 13F, 3= 13N, 4= 15R, 0= Negative control, Pred= Predicted Length by Closest Fit.

The digital method for measuring distanced travelled along the gel allowed for greater accuracy, however a clearer gel image would have been more desirable. At the time of experimentation equipment failure in the lab meant that electrophoresis equipment was running above set voltage, which heated gels and caused DNA to smear. Primer set 2p2 was later abandoned for reasons discussed in depth in the following pages, and returning to this primer set simply to produce a better image was not considered critical to completion of the necessary research given more successful redundant primers had been found.

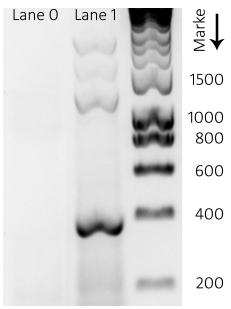
Figure 20. Graph of distanced travelled vs. hyperladder size for 2p2



The following describes the same working, done for the isolated non-target band in primer set 2p2. The sample gel photo was measured by counting pixels. This data (Table 12) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 21. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0011 and c= 8.0944. The amplified bands were estimated to be 354 base pairs. The mean error for hyperladder bands was 1.68%. The isolated band, when sequenced, came to 374bp; the area contaminated in target sequence 2p2 is 406 bp from the 255th nucleotide to the end. Thus, the estimate given by the first hyperladder is more accurate both because of its lower mean error and its closer match to the actual strand length measured.

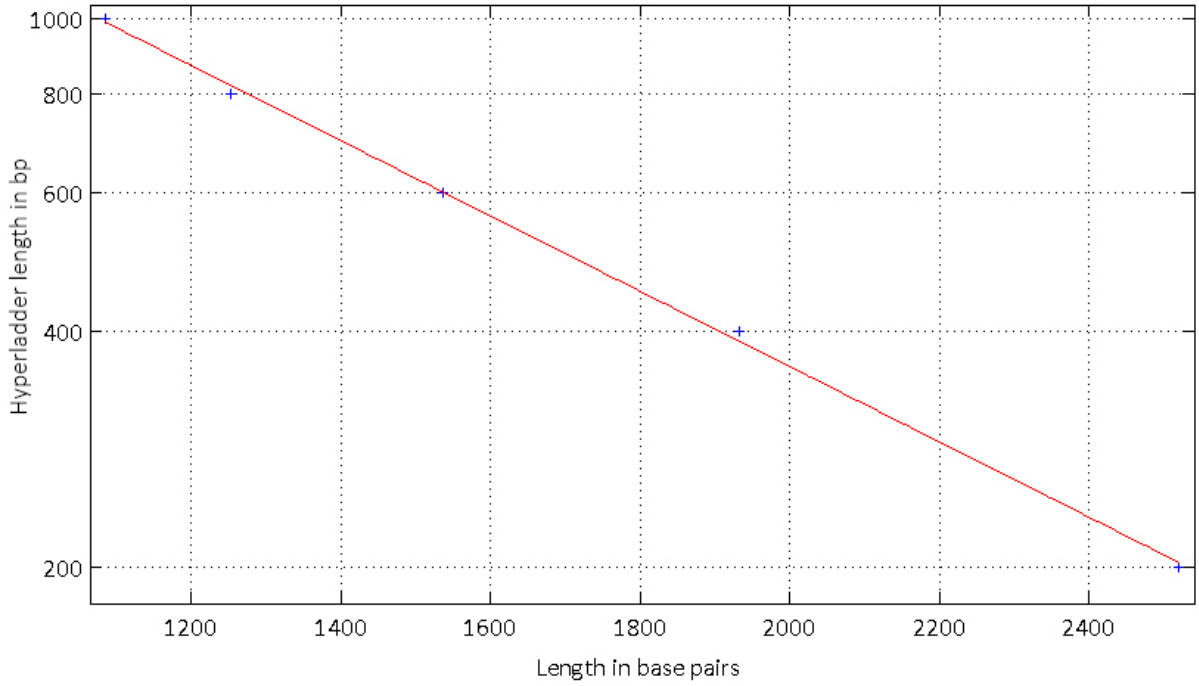
Table 12. Example gel 2p2^MgCl and predicted length using Hyperladder

Hyperladder	Distance	Predicted Length	% Error
1000bp	1085px	989bp	1.10%
800 bp	1254px	821bp	2.63%
600 bp	1537px	601bp	0.16%
400 bp	1933px	388bp	3.00%
Non-target	2015px	354bp	N/A
200 bp	2520px	203bp	1.50%



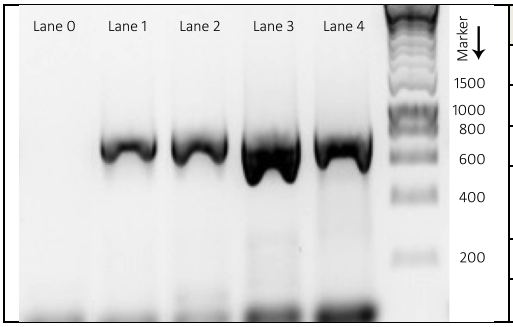
Legend: 1= 46P, 0= -ve Control

Figure 21. Graph of distanced travelled vs. hyperladder size for 2p2 non target



According to nearest neighbour calculations primer 2X Fwd had a Tm of 64.9°C while 2x Rev had a Tm of 65.2°C. In order to find which temperature was best for amplification, gradient PCRs were run between 54°C and 64°C. The strongest bands appeared at 56 °C level. The sample gel photo was measured by counting pixels. This data (Table 13) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 22.

Table 13. PCR amplification using primer 2X and predicted length using Hyperladder



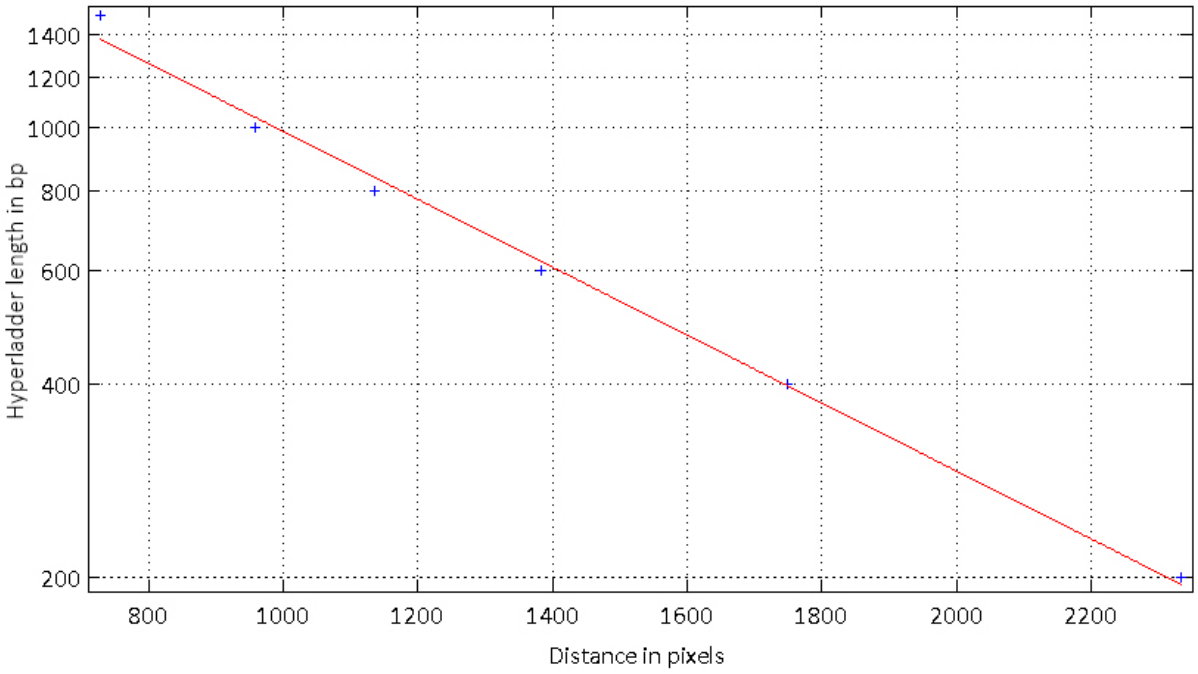
Hyperladder	Distance	Pred	Error
1500 bp	728px	1377	8.13%
1000 bp	958px	1041	4.10%
800 bp	1136px	838	4.75%
Band 2x Exp. 845bp	1216px	777	8.05%
600 bp	1384px	619	3.17%
400 bp	1750px	397	0.75%

Legend: Lane 0= Negative Control, 1= 82W, 2= 26K, 3= 31L, 4= 2F, Pred= Predicted length by closest fit.

This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0012 and c= 8.1143. The amplified bands were estimated to be 777 base pairs, 68 less than the predicted 845bp. The mean

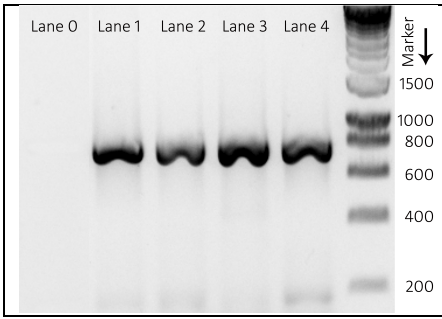
error for hyperladder bands was 4.18%; this is possibly due to the smeared nature of the hyperladder.

Figure 22. Graph of distanced travelled vs. hyperladder size for 2x



According to nearest neighbour calculations primer 2F Fwd had a Tm of 65.4°C while 2 fin Rev had a Tm of 66.0°C. In order to find which temperature was best for amplification, gradient PCRs were run from 56°C to 66°C. The strongest bands appeared at 64 °C. A larger non-target was removed by increasing Ta to 1:10 (m:s) and reducing cycles to x35.

Table 14. PCR amplification using primer 2F and predicted lengths using the hyperladder

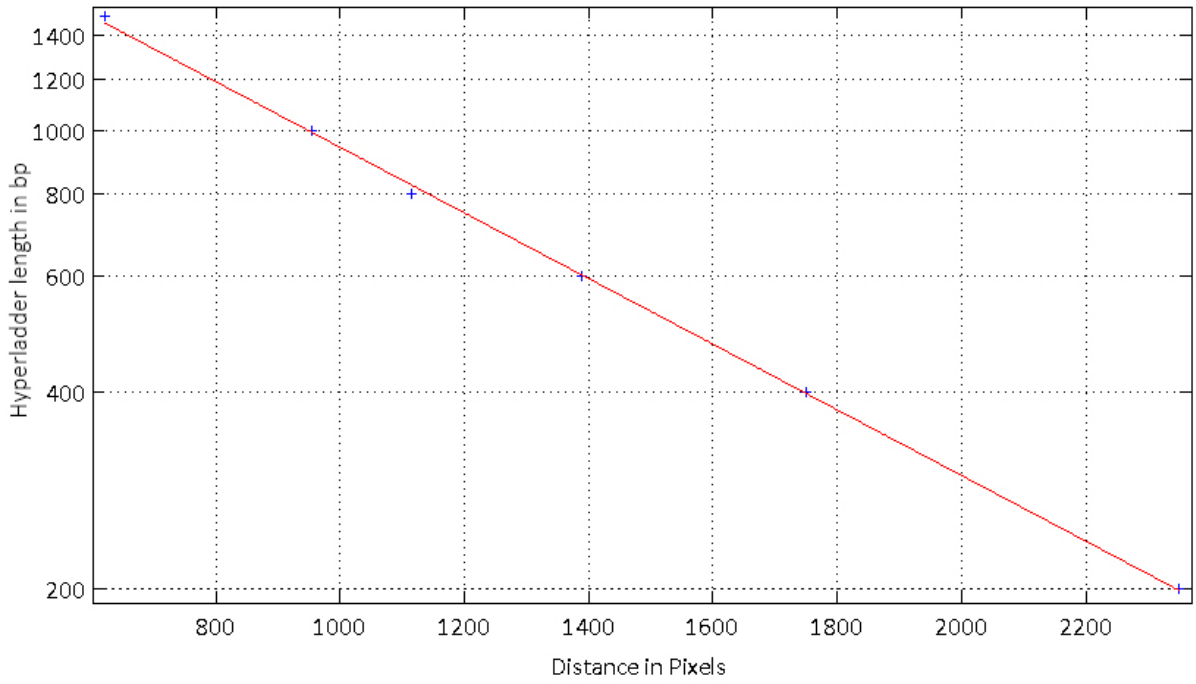


Hyperladder	Distance in px	Pred.	% Error
1500 bp	620	1466	2.27%
1000 bp	953	998	0.20%
800 bp	1115	827	3.38%
Band 2f Exp. 743bp	1205	746	0.40%
600 bp	1389	603	0.50%
400 bp	1749	399	0.25%

Legend: Lane 0= Negative Control, 1= 58R, 2= 41L, 3 =40L, 4= 9P, Pred= Predicted length by closest fit.

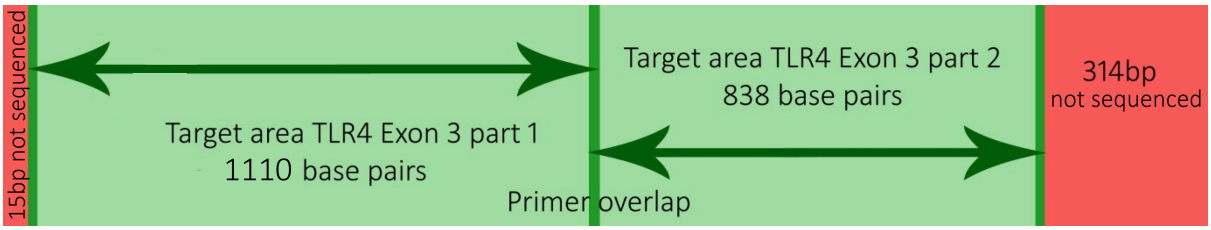
The sample gel photo was measured by counting pixels. This data (Table 14) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 23. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0012 and c= 8.0064. The amplified bands were estimated to be 746 base pairs, 3 more than the predicted 743bp. The mean error for hyperladder bands was 1.32%.

Figure 23. Graph of distanced travelled vs. hyperladder size for 2F



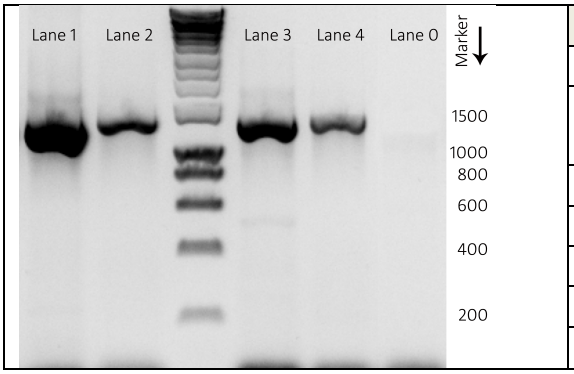
Most of Exon 3 of TLR4 was described by two primer sets (Figure 24). This section will go on to show what was done to successfully work with these two primers sets. Exon 2 of TLR4 was also sequenced in its entirety. According to nearest neighbour calculations primer 4p1 Fwd had a Tm of 65°C while 4p1 Rev had a Tm of 64.9°C. In order to find which temperature was best for amplification, gradient PCRs were run for all temperatures between 54°C and 66°C. The strongest bands appeared at 56°C . A larger non-target band was removed by increasing annealing time to 1:10 and reducing cycles to x35.

Figure 24. TLR4 Exon 3, as described by two successful primer sets.



The sample gel photo was measured by counting pixels. This data (Table 15) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 25. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0012 and c= 8.3516. The amplified bands were estimated to be 1146 base pairs, 36 more than the predicted 1110bp. The mean error for hyperladder bands was 1.79%.

Table 15. PCR amplification using primer 4p1 and predicted length using Hyperladder



Hyperladder	Distance	Pred.	% Error
1500 bp	845	1508	5.33%
Band 4p1 Exp. 1110bp	**1090**	**1146bp**	**3.24%**
1000 bp	1189	990	1.00%
800 bp	1352	812	1.50%
600 bp	1604	596	0.67%
400 bp	1942	395	1.25%
200 bp	2490	202	1.00%

Legend: Lane 1= 26K, Lane 2= 26N, Lane 3= 27K, Lane 4= 30L, Lane 0= Negative, Pred= Predicted Length

According to nearest neighbour calculations primer 4p2 Fwd had a Tm of 48.1°C while Primer 4p2 Rev had a Tm of 51.4°C. In order to find which temperature was best for amplification, gradient PCRs were run between 44°C and 54°C. The strongest bands appeared at 50 °C. The sample gel photo was measured by counting pixels. This data (Table 16) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 26. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0127 and c= 4.9067. The amplified bands were estimated to be 847 base pairs, just 9bp more than the desired product. The mean error for hyperladder bands was 1.37%.

Figure 25. Graph of distanced travelled vs. hyperladder size for 4p1

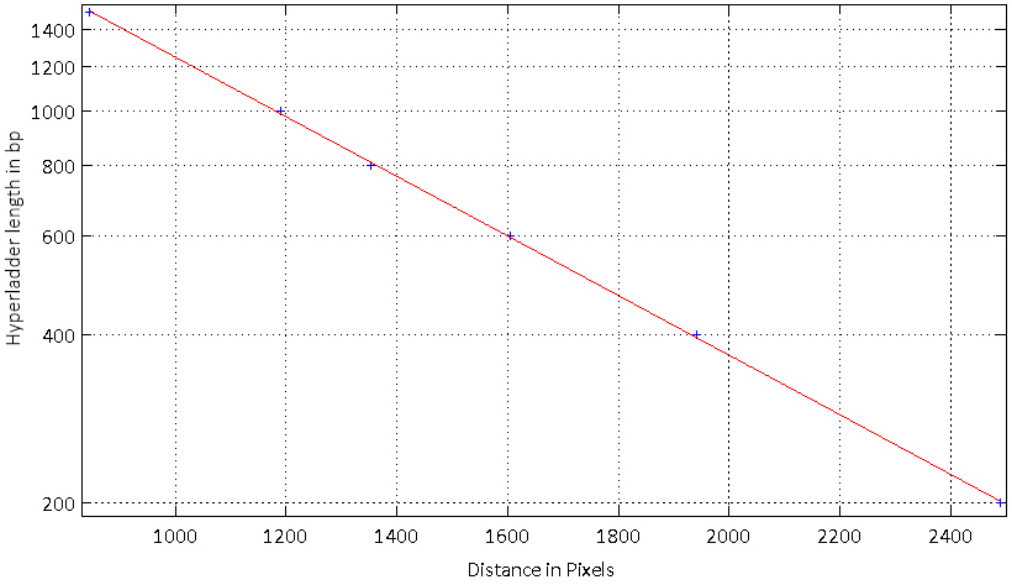
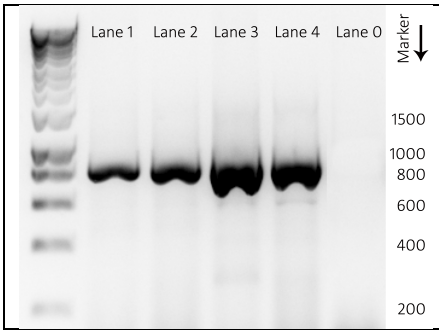


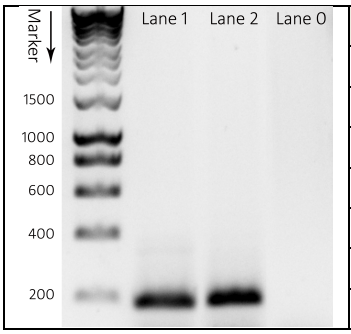
Table 16. PCR amplification using primer 4p2 and predicted length using Hyperladder



Hyperladder	Distance	Pred.	Error
1000bp	158 px	999bp	0.10%
Band 4p2 Exp 838bp	**145px**	**847bp**	**1.07%**
800bp	142px	816bp	2.00%
600bp	117px	594bp	1.00%
400bp	84px	391bp	2.25%
200bp	32px	203bp	1.50%

Legend: Lane 1= 12L, 2= 13L, 3= 14L, 4= 14N, 0= Negative Control, Pred= Predicted length by closest fit.

Table 17. PCR amplification using primer 4e2 and predicted length using Hyperladder



Hyperladder	Distance	Pred.	% Error
1500 bp	819px	1484bp	1.07%
1000 bp	1103px	1016bp	1.60%
800 bp	1267px	816bp	2.00%
600 bp	1504px	594bp	1.00%
400 bp	1825px	387bp	3.25%
200 bp	2306px	204bp	2.00%
4e2 (Exp 178bp)	**2349px**	**192bp**	**7.87%**

Legend: Lane 1= 9P, 2= 50N, 0= Negative Control, Pred= Predicted length by closest fit.

According to nearest neighbour calculations primer 4e2 Fwd had a Tm of 64.6°C while 4e2 Rev had a Tm of 60.9°C. In order to find which temperature held the best compromise for amplification gradient PCRs were run for all temperatures between 46°C and 64°C, before

some success was had at 48°C. Best results occurred when cycle timings were extended as much as possible, here Td:Ta:Te was 1:00;1:00;2:15 for 40 cycles. Optimal settings were not found and redesign of primers in this area proved even less successful. however three samples yielded good bands, giving 178bp amplicons, and were sequenced. The sample gel photo was measured by counting pixels. This data (Table 17) was used to compare distance travelled along the gel with length in base pairs, and plotted in Figure 27. This gave a closest fit equation of $y = e^{(bx+c)}$, where b= -0.0013 and c= 8.3969. The amplified bands were estimated to be 192 base pairs, 14 more than the predicted 178bp. The mean error for hyperladder bands was 1.82%.

Figure 26. Graph of distanced travelled vs. hyperladder size for 4p2

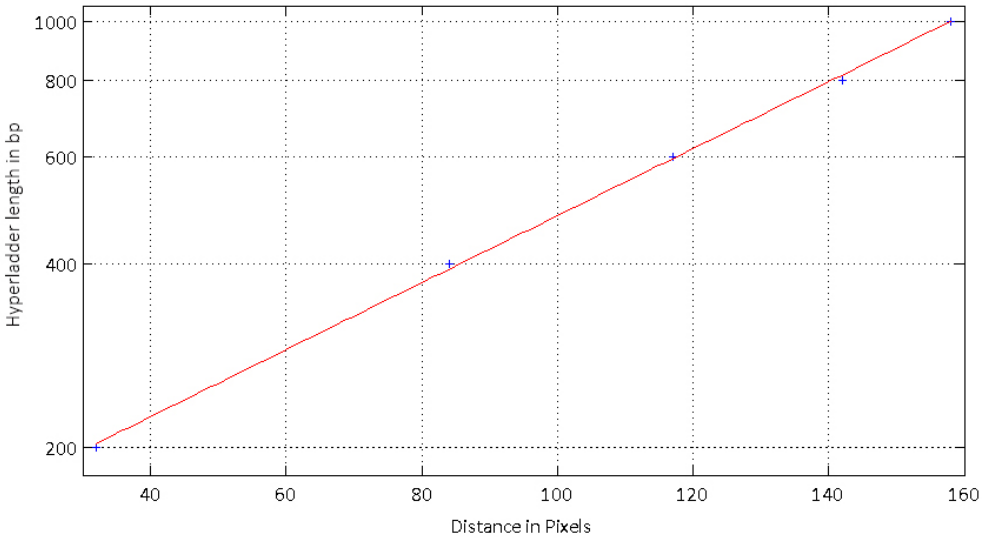
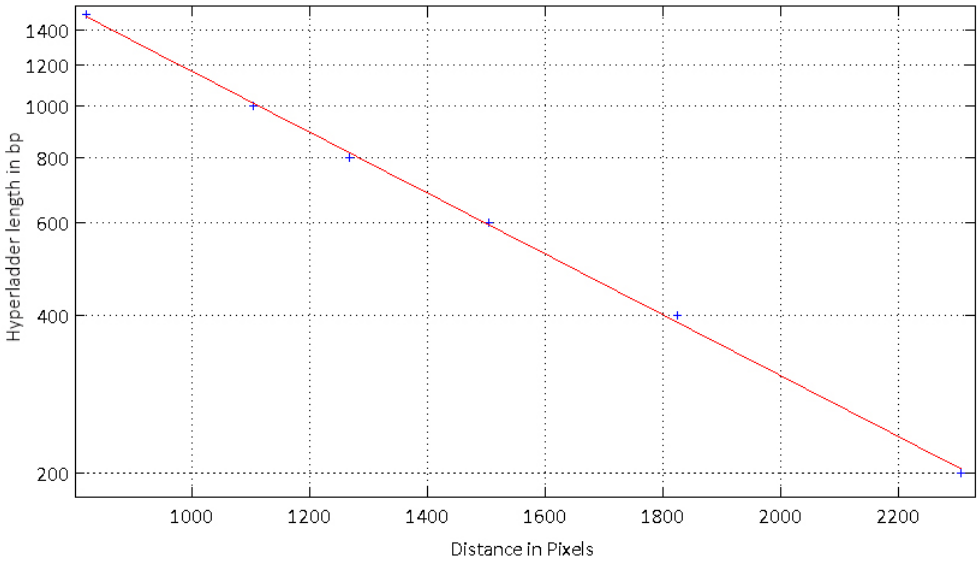


Figure 27. Graph of distanced travelled vs. hyperladder size for 4e2



3.2 Sequencing and Analysis of TLRs

PCR products were frozen at -20°c before being sent for forward and reverse sequencing. Read outs were returned in the form of .abi graphics which were checked by eye for variants. Fifty-two samples from Woodchester Park, as well as nine outgroup samples provided by Secret World Wildlife Rescue, were sequenced across TLR2 in its entirety (Table 18).

Table 18. Sixty-one Badgers sequenced across ALL primers in TLR2.

Woodchester Sample set	1K, 1L, 2F, 2P, 3K, 4K, 5K, 7K, 8I, 8L, 9P, 12K, 12L, 13L, 13N, 14L, 14N, 15K, 17K, 17L, 21K, 22K, 22N, 24L, 24P, 25K, 26K, 26N, 27K, 28F, 30L, 31L, 32N, 33Y, 34K, 35P, 36K, 38K, 38L, 39K, 40K, 40L, 41L, 45P, 45R, 50N, 58R, 59B, 63Y, 66B, 67I, 82W.
Outgroups	1, 2, 3, 4, 6, 13, 14, 16, 19.

A further three Woodchester badger samples 9N, 37K, and 38Y were sequenced across primers 2S and 2X only. Four additional SWWR badgers were partially sequenced: badger 12 across 2S and 2X, and badgers 5 and 18 were across 2X only. Badger 24 provided by University College Dublin was sequenced across 2S and 2X. Forty-nine samples from Woodchester Park, as well as ten outgroup samples provided by Secret World Wildlife Rescue were sequenced across all primers for TLR4 exon 3 (Table 19). A further eleven Woodchester samples were sequenced across primer 4p2 only: 3K, 13F, 13N, 15R, 17F, 22K, 28F, 28K, 31K, 32K, and 46P, as well as a further three SWWR samples: 1, 4 and 18.

Table 19. Fifty-nine Badgers sequenced for across ALL primers in TLR4 exon 3.

Woodchester Sample set	1K, 1L, 2F, 2P, 4K, 5K, 7K, 8I, 8L, 9N, 9P, 12K, 12L, 13L, 14L, 14N, 15K, 17K, 17L, 21K, 22N, 24L, 24P, 25K, 26K, 26N, 27K, 30L, 32N, 33Y, 34K, 35P, 36K, 38K, 38L, 38Y, 39K, 40K, 40L, 41L, 45P, 45R, 50N, 58R, 59B, 63Y, 66B, 67I, 82W.
Outgroups	2, 3, 5, 6, 11, 12, 13, 14, 16, 19

Sequences were run through multiple sequence alignment to gain an overall consensus sequence for each primer set, and the .abi graphic of every sample was scanned by eye, for every nucleotide, in a search for variant or heterozygous bases. Lyons (2014) advised that heterozygotes are marked by the presence of both peaks with matching curves sitting at

roughly half the height of homozygous peaks, or where one peak is directly under the other but lower in height. There is a precedent for sequencing both strands, as heterozygous sequences can sometimes be missed on one strand and seen on the other. Taylor et al (1999) found a T>A heterozygosity which was called correctly in one direction, but missed on reverse strand, and an A>G heterozygosity missed on the forward stand, but called on the reverse strand. They found that due to signal weakening as the sequence progresses, heterozygotes were lost unless read from the complimentary strand. Error rates increase rapidly after 690 bases, due to weakened peak resolution. No variations were found using primer 2S, though noise in the data sometimes produced false positives that were proven not to be a variation through further trials. One of these was a C/G Heterozygous nucleotide in sample 3K which would have equated to a stop codon had its veracity been proven. This was the 146th base along TLR2. Several trials were performed to check its veracity (Figure 28.) A repeated trial always refers to where both repeated PCR and repeated sequencing in both directions has occurred. In this case a large amount of repeated trials occurred in order to ensure due dilegence. Sequences were often returned with 'noise' so that any possible polymorphism was visible but questionable. A single retrial returning a polymorphism with low noise levels was not considered definitive, so trials were done until at least three produced sequences with noise low enough to determine that the polymorphism could not be dismissed as noise within noisy data.

Further sample testing revealed no variations across all samples for primer set 2s. Noise in other sample data was resolved by further PCR. For example, the first run of 4K was noisy and was resampled to show peaks identical in sequence to all other badgers tested (Figure 29.)

Figure 29. Two PCR samplings and sequencings of sample 4K

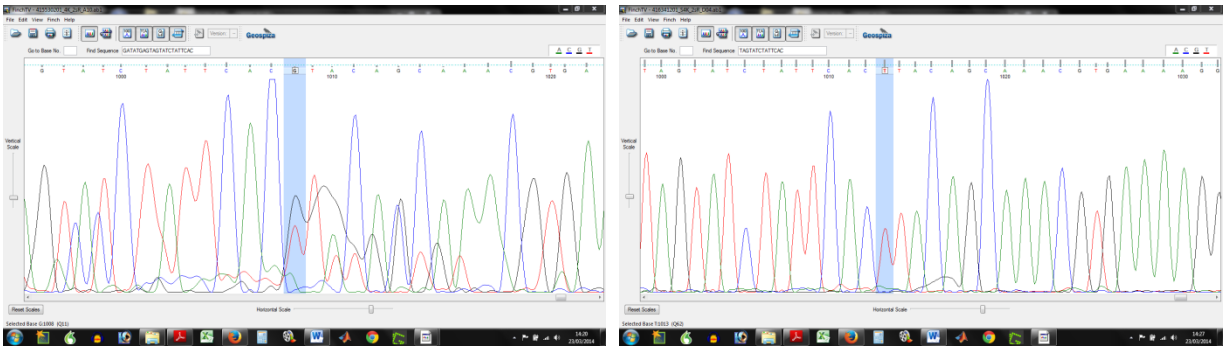
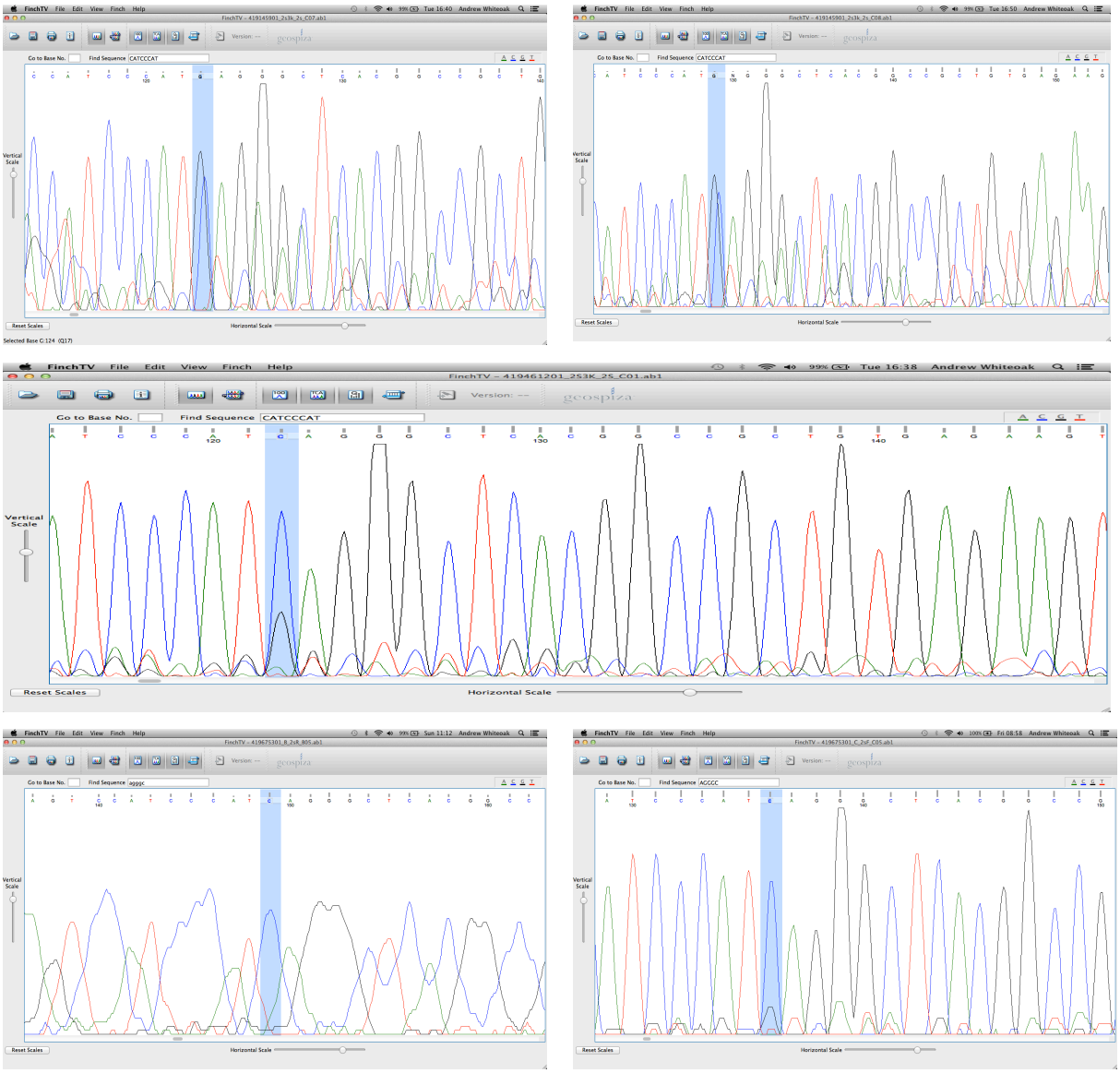


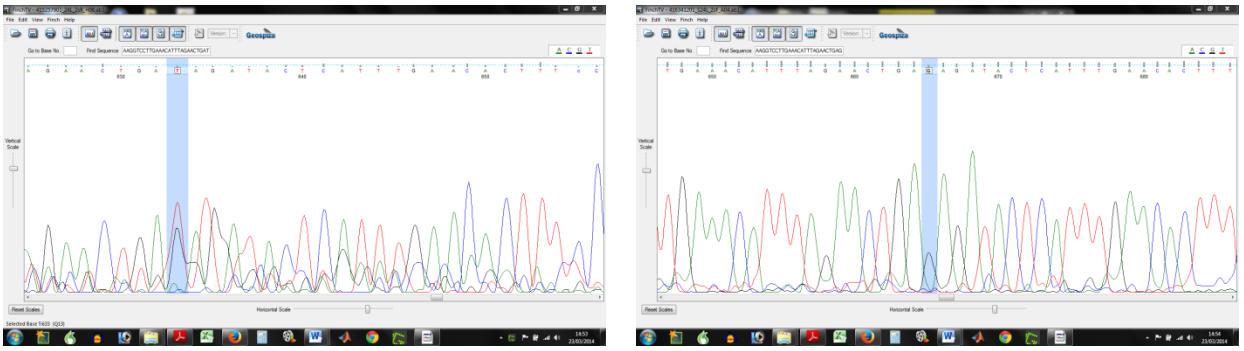
Figure 28. Tests for variation in sample 3K.



Legend: Left-right top-to-bottom shows early, noisy, trials, and a final, seventh trial with a clear single C peak.

The first sampling of 24L produced a messy forward sequence. T appeared where a G was expected. The resampled data made it clear that the nucleotide was in fact a G (Figure 30). Early sampling of 66B gave a G peak underneath the expected A peak. Resampling found no G peak, showing that even in cleaner samples, noise can be an issue (figure 31). Other samples required multiple trials. The first and second trials of 82W were contaminated from the 909th nucleotide. The sequence identified as consensus across all other badgers can be seen clearly underneath the taller waves which graph the contamination (figure 32). A third and fourth sampling revealed the data points to be identical to the majority consensus data.

Figure 30. Two PCR samplings and sequencings of sample 24L



TLR2 primer set 2p2 was sequenced across only 13 Woodchester badgers; 1K, 7K, 8L, 9N, 12K, 14L, 24L, 25K, 26N ,32N, 38K, 63Y and 66B before sequencing issues led to a redesign of primers for this area. The forward strand in all 13 samples showed a perfectly 'clean' sequence with a single, evenly spaced peak at every nucleotide. However the reverse strands in all 13 began to be noisy at the exact same point, with increasing noise until the 3' end was reached. This suggested that there was a contaminating sequence, and this underlying sequence looked to be the same each time. It was hypothesised that the reverse primer may have captured a smaller region, perhaps a pseudogene. In order to solve this mystery, first, an attempt was made to decipher the contaminating sequence by eye. Noting underlying peaks the sequence in Figure 33 was reached. In an attempt to isolate this sequence, further PCR was performed using these primers. $MgCl_2$ loading was doubled and a single short band appeared. This band was the same size as the length of strand contaminating the target, and was sequenced (Figure 34).

Figure 31. Two PCR samplings and sequencings of sample 66B

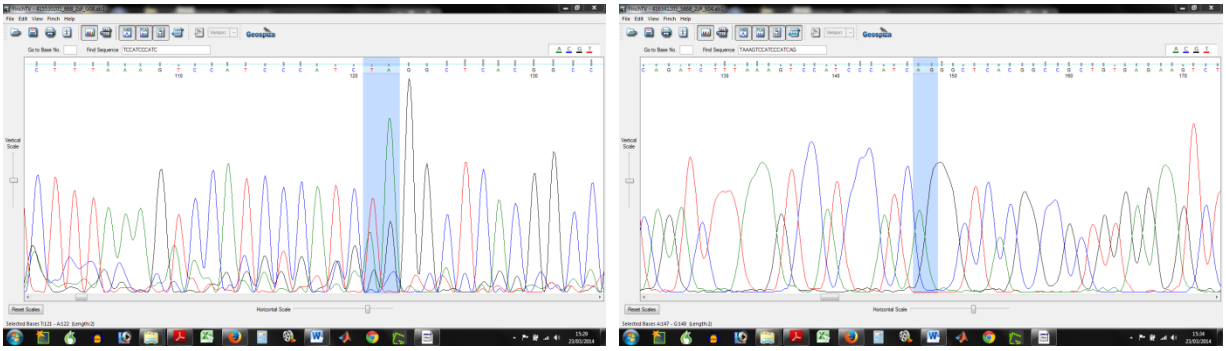


Figure 32. The first two PCR samplings and sequencings of sample 82W

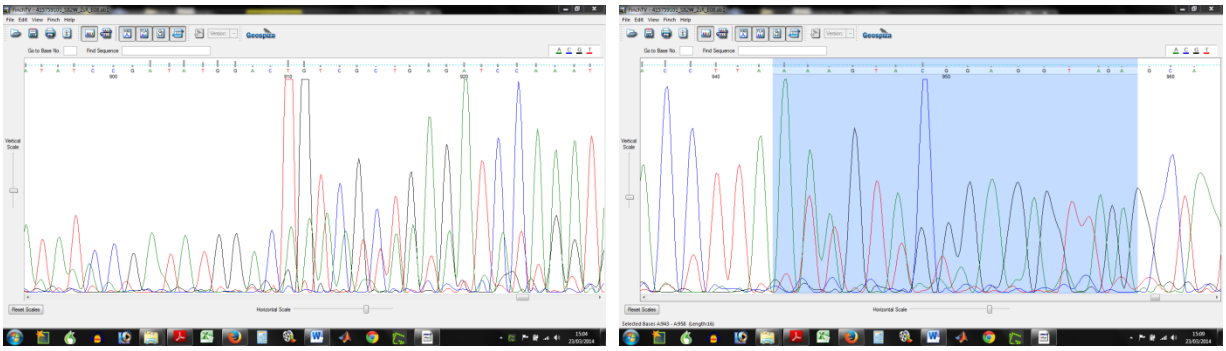


Figure 33. Primer 2p2 Underlying sequence called by eye

AAGTNCCGNNNGTGATNGAGCWCWNTNYNGRNTAAWTNGKNGTTTTANRRAWAATTTRMARGAGAGGNGKTANGG
ATANAGGYATNTCNNTCTTTTNGCANANNGTNTTCACNTTCNCGTKTMTATNTAGNAANCARMCKRANNNNNNAT
NAGANAAGGGGAGRNGNNTGCTGTGNAAANNGAGNGTGTYANGACAGGCCTTGGAGAARTGNCNCNGTAAGRCAC
NNNMTATGCGTNNNNNCTTGANGNNNNNGCAGGGAGCANNNNNCGATGAAAANTGCNTWGANAAGRGCCNGTG

Legend: To obtain this sequence the abi graphic was viewed by eye. All peaks had at least two bases and the smaller, or 'underlying', sequence peaks were noted. Where peaks could be one of a number of base calls a degenerate base value was used to indicate this.

Figure 34. Isolated band sequence (double $MgCl_2$ with 2p2 primers)

AAGTCCCGCTTGTGCCCGNGCNCTCCNNGGGATAAATNNNCGGTACAAGGAAAGAATACAGGCGTCACGGNTGAA
AGGGCTTGNNCATTGACTNGCACNNNCACAAACNNTTGTAGNNATGGGTTTGCAAGCAAGCACATNGGGGCGGAG
AAAAGGGGACGAATTCGCTGGTCAAAAATTAGTGTTTTTTAGCGGCATCCGGNGAGAAACACGAAGTTAAGCCCA
GCAGATCAGCCCCAGAANTCCAGACGCAGGAGNTATGTTNGCAAGNGATCATTCGGGAGGAAGGGCGNNCAGGGA
GGCCAANGANGAAGCCTTGAGCAGAGNCTNNGTCGNNTNNTNCNNNNNNNNNNNNNNTNNNNNNNNNNG

The $MgCl_2$ band and underlying sequence were matched closely via multiple sequence alignment (Figure 35). Efforts were made to understand where the contamination had appeared. It was found that the reverse complement of sequence 2p2 also matched the underlying sequence closely. It was postulated that the contaminating sequence was in fact a section of reverse strand read forwards. The sequence at the start of contamination was also closely complementary to the forward primer. This created a unpredicted pseudo dimer. This differs from a primer dimer which is simply the interation of either two primers attaching to one another, or a primer self-attaching. Instead the reverse strand of the amplified sequence from primer set 2p2, contained a central sequence which complemented the forward primer of 2p2, so that during PCR the forward primer was able to palindromically attach to both strands of the target sequence, creating an additional, shorter amplicon of one section of the target sequence.

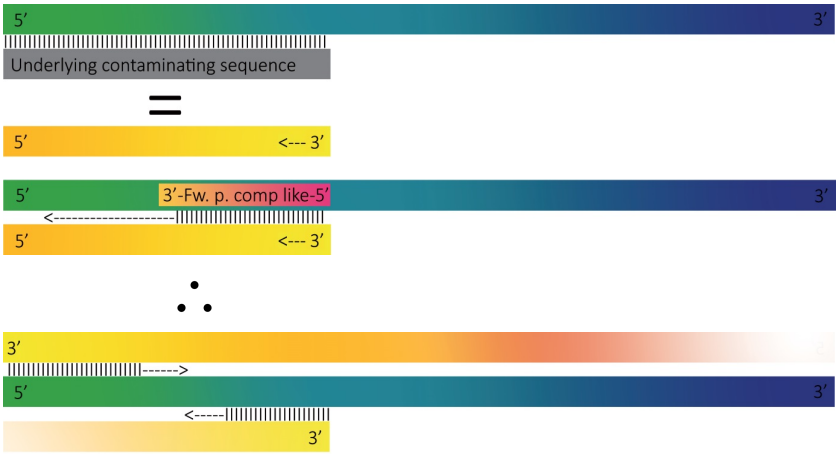
Figure 35. First 120 bases of Multiple sequence alignment

```
Fwd Primer 5'CTCAGSGAMAATYTRATGGTT
            !!!!!!  !!      !!
2p2RevStrand AAGTCCCGCTTGTGAAGGCACAGCTTGAAGGGAGGGTCGAAGTGCTCCAGCTCCTGGACC
MgClband     AAGTCCCGCTTGTGCTCGAGCWCTCCNNGGGATAAATNGGCGGTTTAAGGAAAGAATACA
underlying   AAGTCCCGNNNGTGATNGAGCWCTNTNYNGRNTAAWTNGKNGTTTTANRRAWAATTTRMA
U-match-M    ********   *** * *******     ** ***** ** * **** *****   ****
U-match-Rev  ********   ****  *  **   *   **     * *  **  *  ** *  *  **

2p2RevStrand ATCATGTTCTCCACCCAGTAGGAATCGTGCTCACTGTAGGACACGAAGGCATCATAGCAC
MgClband     GGCGTCACGGNTGAAAGGGCTTGNNCATTGACTNGCACNNNCACAAACNNTTGTAGNNAT
Underlying   RGAGAGGNGKTANGGATANAGGYATNTCNNTCTTTTNGCANANNGTNTTCACNTTCNCGT
U-match-M    ** *    **     *                **                    *     *
U-match-Rev  *    *   *         *** **     **  *         *    **   *  *
```

Legend: U-match-M indicates, with asterixes, where the underlying contaminating sequence matches the isolated $MgCl_2$ band. U-match-Rev indicates, with asterixes, where the underlying contaminating sequence matches the reverse strand of amplicon 2p2. "!" indicate where the Fwd primer complements these sequences.

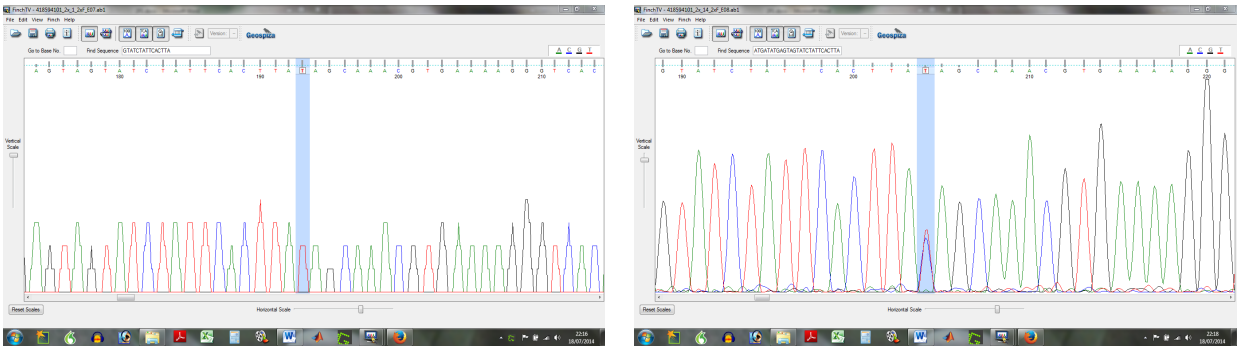
Figure 36. Resolution of primer set 2p2 self-contamination.



Alignment was strong for the first few bases; 42/60 bp of the underlying sequence match the $MgCl_2$ isolated band. Twenty-nine of these match the reverse complment of the 2p2 amplicon. The source of the problem was the palindromic nature of the forward target strand, which in collaboration with the nearly complimentary sequence mid way through the reverse strand, caused self contamination to occur (figure 36). The solution was to create primer sets 2X and 2F.

Two confirmed haplotype variations at one nucleotide position were found by primer set 2X. Outgroup badger 1 showed a homozygous T, and outgroup badger 14 showed a heterozygous C/T (Figure 37), for a base homozygous for C in all other badgers.

Figure 37. Variant badger 1 (left) and variant badger 14 (right)



Just as with primer 2S, other samples suffered noise and were checked again by resampling. An example is sample 58R. On first sampling an A peak was seen in place of an expected G. Further sampling still showed the A close underneath the normal G peak on an unotherwise clear graphic. However, this was still questionable; unlike the heterozygous nucleotide given by matching half-size peaks in badger 14, the 58R peaks were of similar height to peaks surrounding them. A third sampling showed that there was actually only a G at this point. These three trials are shown in Figure 38, clockwise from left to right.

In sample 3K there there seemed a very real possibility that a heterozygous variation had been found. Though the data was noisy, the C peak noted superimposed over the expected G peak, and was half the height of surrounding peaks. However a second trial found only the expected G peak (Figure 39). Negotiable amounts of noise in other samples meant that even where it seemed unlikely that a variation would be found, samples were repeated. In 37K there is clearly a large amount of noise. However, there appears to be a T and G curve simultaneously overlapping at a lower height than surrounding peaks where normally the sequence is G homozygous. A clean resampling shows that it is indeed homozygous for G as expected (Figure 40).

Primer set 2F bore one confirmed variation. Multiple samplings of badger 30L showed C/G heterozygous peak. Early trials were noisy but further PCR came up with clean data where the double peak remained (Figure 41).

Figure 38. Resolution of noise in sample 58R for primer 2X

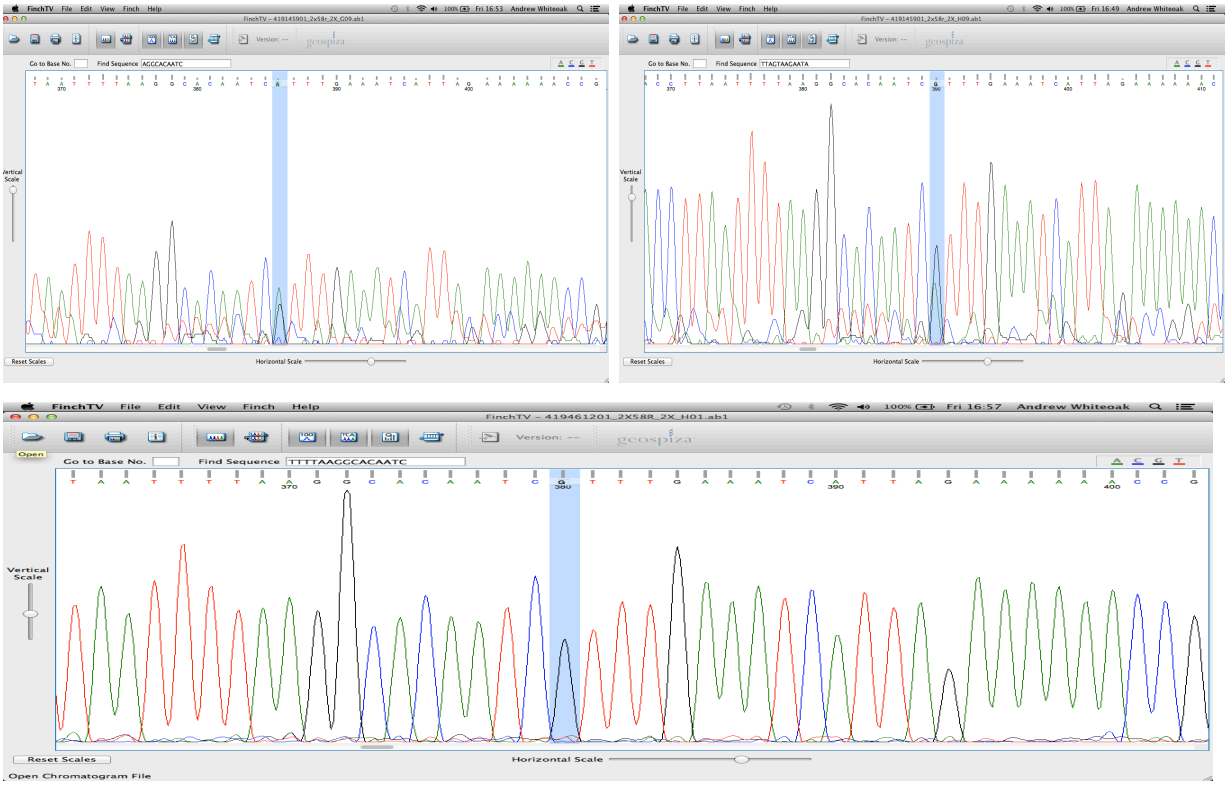


Figure 39. Two PCR samplings and sequencings of sample 3K

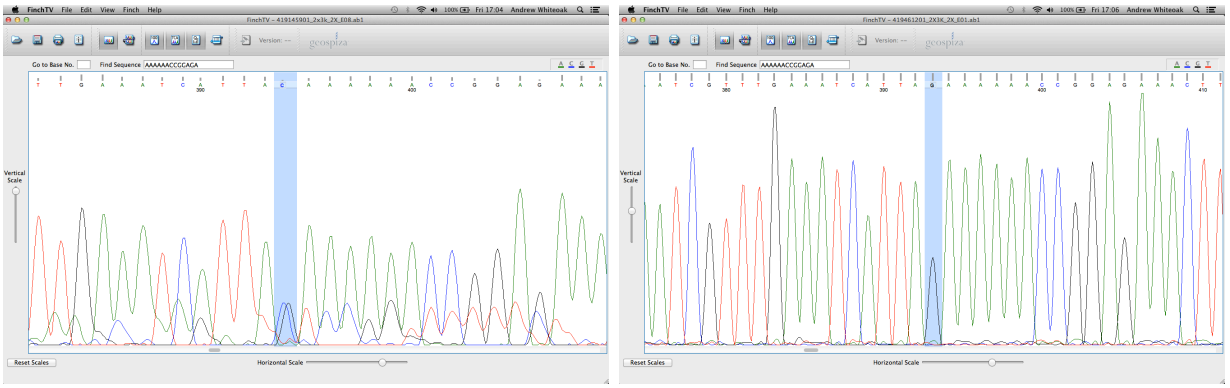


Figure 40. Two PCR samplings and sequencings of sample 37K

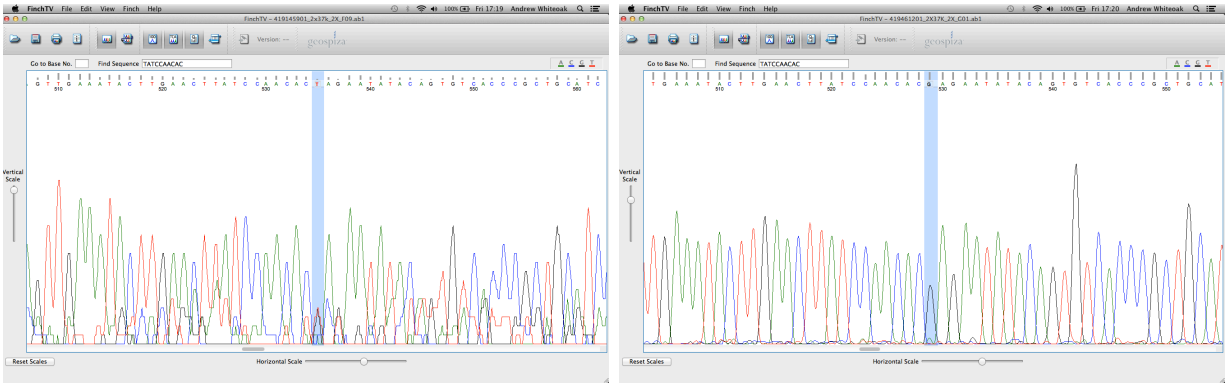


Figure 41. Three sequencings of PCR samples from 30L showing C/G heterozygous base.

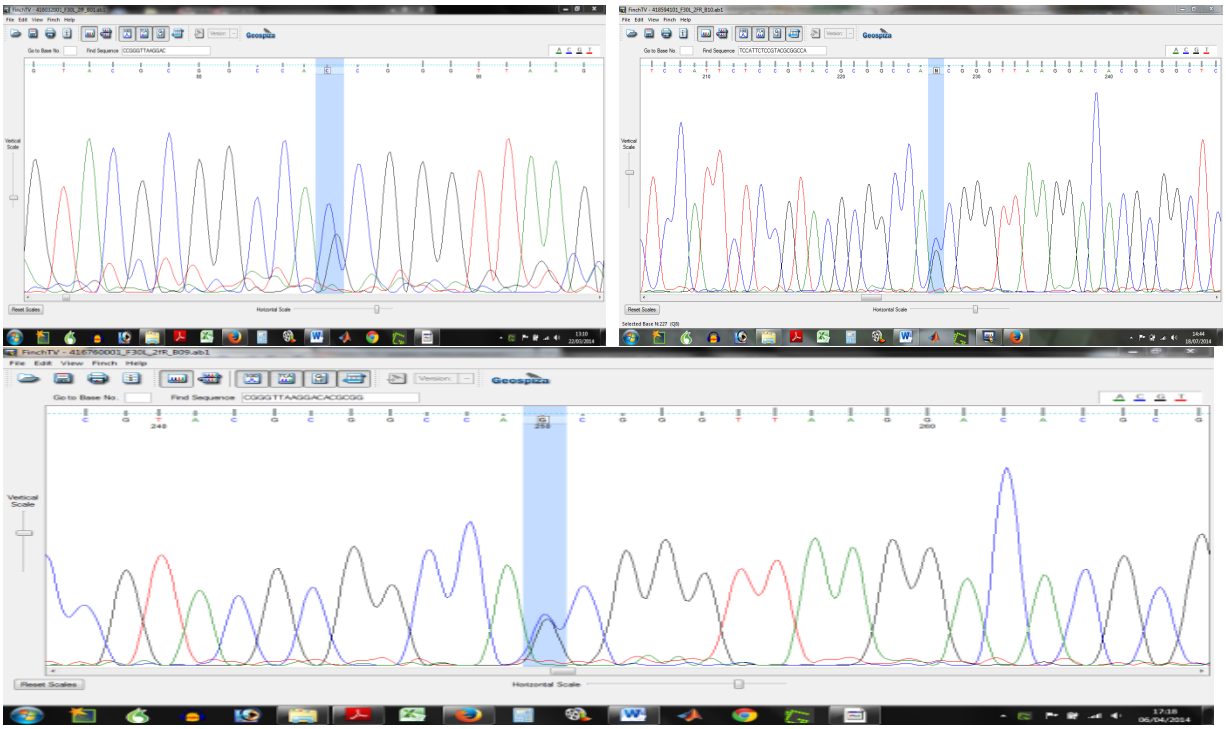


Figure 42. Two PCR samplings and sequencings of sample 22N

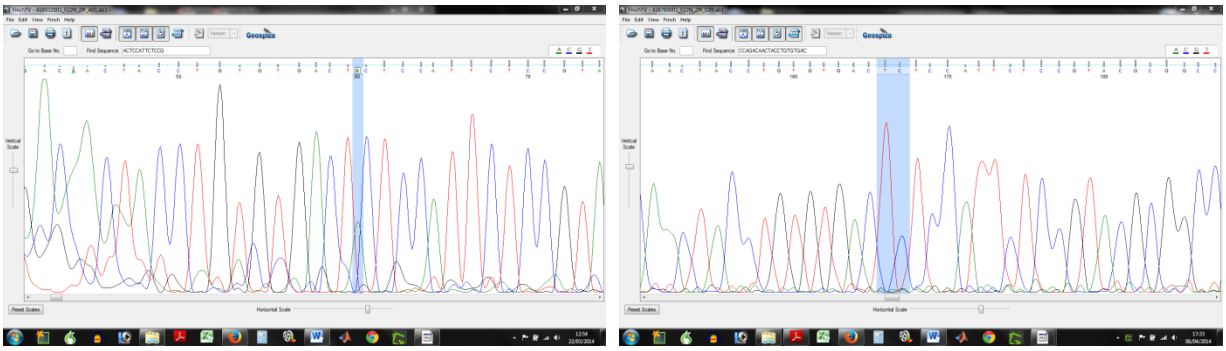
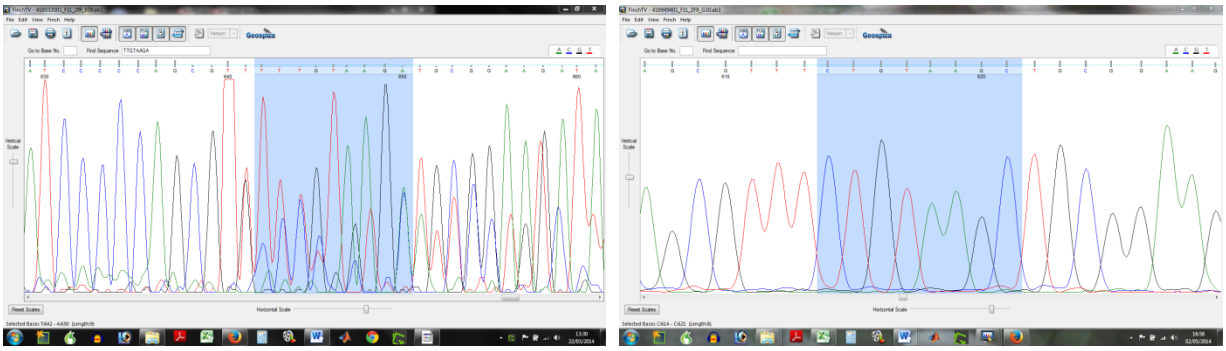


Figure 43. Two PCR samplings and sequencings of sample 1L



Noise in other samples was investigated and bases were proven to be none variant. Noise created the question of an A nucleotide insertion in the first sampling of 22N, and was shown to be noise which created a software based sequence misreading after resampling (Figure 42). Noise in sample 1L was resolved with further sampling (Figure 43).

By aligning sequenced sections end to end, a full sequence for TLR2 in the badger was obtained (Figure 45) which was then converted to an amino acid sequence using the function designed in Matlab (Figure 44). Variations found were examined to find their affect on the final protein. All variations were SNPs which are either transitions or transversions (Seabury et al 2007). Transitions are like replaced by like; a purine replaced by a purine (A or G) or a pyrimidine by a pyrimidine (C or T). Transversions are like replaced by non-like; e.g. a purine replaced by a pyrimidine (Friedberg and Lawrence, 2001).

Figure 44. Amino Acid Sequence of Badger TLR2.

MSRVLWTVWVLGAVTKLSKEEGPDQASSLSCDLTGVCDGRFRSLKSIPSGLTAAVRSLDLSNNEITYIRNRDLRG
CVNLKALKLASNRINAIEEDSFISLRSLEHLDLSYNLLSNLSSSWFRPLSSLKFLNLLGNRYKSLGEMPLFSPLT
NLQILKVGSIDSFTELQEKDFAGLSFLEELEIDASNLQRYEPESLKSIQNISYLALRMKQPIFLLEIFGDLSRSL
KHLELRDTHLNTFQFSKASIRETNTLIKKWTFRNVKITDGSFSELVKLLNCVSGVLEVEFEGCTLDGLGNFDISD
MDKIKNIGGIETLIVRRLAIPHFYSFYDMSSIYSL**T**ANVKRVTVESSKVFLVPCLLSQHLKSLEYLDLSDNLMVE
ESLRNSACDQAWPLLQTLILRHNRLKSLEKTGETLLSLKNLTKLDISKNNYVSMPETCQWPDKLKYLNLSNTRIY
SVTRCIPWMLEILDISNNNLDSFSLILPRLKELYISGNKLKTLPDASFLPTLRILRISRNIISTFTKEQLDSFST
LEALEAGGNNFFCSCEFLSFTREQQSLAQILTDWPDNYLCDSPFSVRG**Q**RVKDTRLPASECHRVALVSAVCSVLF
LLILLTGVLCHHFHGLWYLKMMWAWLQAKRKPRKAPPRDVCYDAFVSYSEHDSYWVENMMVQELEHFDPPFKLCL
HKRDFIPGKWIIDNIIDSIEKSHKTIFVLSENFVKSEWCKYELDFSHFRLFDENNDAAILVLLEPIEKKAIPQRF
CKLRKIMNTKTYLEWPTDETQQEGFWLNLRMAIKS

Legend: Pink highlights amino acids in consensus sequence which are changed by non-synonymous variation.

Sixty-one badgers were sequenced using all TLR2 primers covering all 2355bp of this single exon gene. Four haplotypes were found, all of which led to nonsynonymous mutations. Two percent of badgers from the Woodchester Park set showed variation from the majority consensus sequence, while in outgroup sets 22% of badgers showed TLR2 sequence variation. This refers to 1/52 badgers from the core sample set and 2/9 badgers from outgroups. A Fisher's Exact Test was performed to resolve whether there was a significant difference between core and out group variation (Table 20).

Table 20. Fisher's Exact Test on TLR2 variation levels in core and out group badgers.

	Woodchester	Outgroup	column sum	
Variant	1	2	3	
Non-Variant	51	7	58	
row sum	52	9	61	
		P value		0.054

The test found a marginally significant difference between the two groups; significant at the 90% level and just outside of the 95% significance level with a P value of 0.054. This may suggest variation due to Geographical Isolation by Distance (GID) which implies that a broader geographical study across badger TLR2 is in order.

Figure 45. Complete single exon TLR2 for the badger

ATGTCACGTGTTTTGTGGACAGTATGGGTGTTGGGGGCTGTAACCAAACTCTCCAAGGAAGAGGGCCCTGATCAG
GCTTCTTCTCTGTCCTGTGACCTCACGGGTGTCTGTGATGGCCGCTTCAGATCTTTAAAGTCCATCCCATCAGGG
CTCACGGCCGCTGTGAGAAGTCTTGACCTCTCCAACAATGAGATCACCTACATCCGCAACAGGGACCTGCGGGGC
TGCGTGAACCTCAAGGCTCTGAAGCTGGCATCTAATAGAATTAACGCCATAGAGGAAGATTCTTTTATTTCCCTG
CGGAGTCTCGAACATTTGGACTTATCCTATAATCTCTTATCGAACTTATCATCCTCCTGGTTCAGGCCCCTTTCT
TCCCTAAAGTTCTTAAACTTACTGGGAAATCGTTACAAATCCCTCGGGGAAATGCCTCTTTTTTCTCCGCTTACA
AATTTGCAAATTCTGAAAGTAGGGAGTATTGACAGCTTCACCGAACTTCAGGAAAAGGATTTTGCTGGGCTCAGT
TTTCTTGAGGAACTCGAGATTGATGCTTCAAATCTCCAGAGGTACGAGCCAGAGAGTTTGAAATCAATTCAGAAC
ATCAGCTACCTGGCTCTTCGTATGAAGCAGCCTATTTTCTTGCTGGAGATTTTTGGAGATCTTTCAAGGTCCTTG
AAACATTTAGAACTGAGAGATACTCATTTGAACACTTTCCAGTTTTCGAAAGCATCCATCAGGGAAACAAATACA
TTGATTAAAAAGTGGACGTTTAGAAACGTGAAAATCACTGATGGAAGTTTTAGCGAACTTGTGAAACTGCTGAAT
TGTGTCTCTGGAGTGTTAGAAGTCGAGTTTGAGGGCTGTACGCTCGATGGGCTTGGTAATTTTGACATATCCGAT
ATGGACAAAATTAAAAATATAGGTGGGATAGAGACCTTAATAGTACGGAGGTTGGCTATTCCACATTTTTACTCA
TTTTATGATATGAGTAGTATCTATTCACTTACAGCAAACGTGAAAAGGGTCACAGTAGAAAGCAGTAAGGTTTTT
CTGGTTCCTTGCTTACTCTCACAACATTTAAAGTCGCTAGAATATTTGGATCTCAGTGACAATTTAATGGTTGAG
GAATCCTTGAGAAACTCAGCCTGTGACCAGGCCTGGCCCCTCCTGCAAACCTTAATTTTAAGGCACAATCGTTTG
AAATCATTAGAAAAAACCGGAGAAACTTTGCTTAGTCTGAAAAACCTGACTAAGCTTGACATTAGTAAGAATAAT
TATGTTTCTATGCCTGAAACTTGTCAGTGGCCAGACAAGTTGAAATACTTGAACTTATCCAACACGAGAATATAC
AGTGTCACCCGCTGCATCCCCTGGATGCTGGAAATTTTAGATATTAGCAATAACAACCTCGATTCCTTTTCCCTG
ATTTTGCCACGACTCAAAGAACTTTATATTTCCGGAAATAAGTTGAAGACCCTACCAGATGCCTCCTTCTTACCC
ACGTTACGCATCTTGAGAATCAGCAGAAATATAATAAGTACTTTCACTAAGGAGCAACTTGATTCTTTTAGCACA
TTAGAGGCTTTGGAAGCTGGCGGTAACAATTTCTTTTGCTCCTGTGAATTCCTGTCTTTCACTCGGGAGCAGCAG
TCGCTGGCCCAGATCCTGACCGACTGGCCAGACAACTACCTGTGTGACTCTCCATTCTCCGTACGCGGCCAGCGG
GTTAAGGACACGCGGCTCCCGGCCTCTGAATGCCACCGGGTGGCTCTGGTGTCTGCTGTGTGCTCGGTCCTCTTC
CTGCTGATACTGCTCACGGGGGTCCTGTGCCACCACTTCCATGGGCTGTGGTACCTGAAAATGATGTGGGCCTGG
CTCCAGGCCAAGAGGAAGCCCAGGAAAGCACCCCCCAGGGACGTGTGCTATGATGCCTTCGTGTCCTACAGTGAG
CACGATTCCTACTGGGTGGAGAACATGATGGTCCAGGAGCTGGAGCACTTCGACCCTCCCTTCAAGCTGTGCCTT
CACAAGCGGGACTTTATCCCCGGCAAATGGATCATTGACAATATCATTGACTCCATCGAGAAGAGCCACAAGACC
ATCTTTGTGCTCTCAGAGAACTTTGTGAAGAGCGAGTGGTGCAAATATGAGCTGGACTTCTCCCATTTTCGCCTT
TTTGATGAGAACAACGATGCTGCCATTCTCGTTCTTCTGGAGCCCATTGAGAAGAAGGCCATCCCCCAGCGTTTC
TGTAAGCTGCGGAAGATAATGAACACCAAGACCTACTTGGAGTGGCCCACAGATGAAACTCAGCAGGAAGGGTTT
TGGTTAAATTTGAGAATGGCAATAAAGTCCTAGATTCCTTTATTAAAGGCTAGTCTGGGTCTGCTGGTGGTCTTT
GTGTCACTAGTGGTAGCTTATTCTGACACGGC

Legend: Red letters: Intron, Pink highlights: loci at which variation was found, Underlined: Base-pairs coding for amino acid codons affected by nucleotide variation.

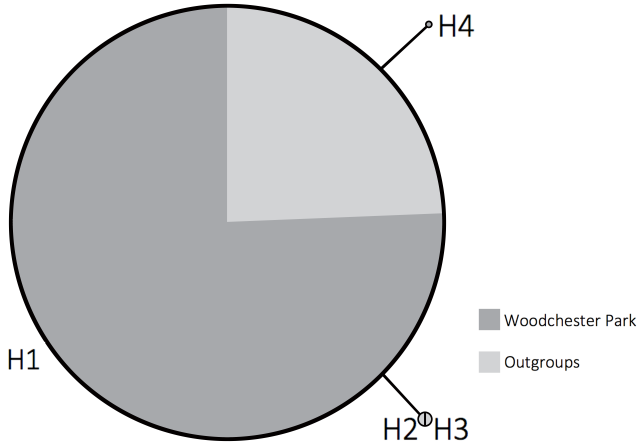
A summary of badgers sampled is shown in Table 21. A haplotype network was created by aligning the 57 H1 sequences against the single H4 and linked H2 and H3 sequences, and is displayed in Figure 46.

Table 21. Summary of sixty-one badgers sequenced and their haplotypes for TLR2

WP	H	WP	H	WP	H	WP	H	WP	H	WP	H	SWWR	H
1K	H1	8L	H1	17K	H1	26N	H1	36K	H1	50N	H1	**1**	**H2**
1L	H1	9P	H1	17L	H1	27K	H1	38K	H1	58R	H1	2	H1
2F	H1	12K	H1	21K	H1	28F	H1	38L	H1	59B	H1	3	H1
2P	H1	12L	H1	22K	H1	**30L**	**H4**	39K	H1	63Y	H1	4	H1
3K	H1	13L	H1	22N	H1	31L	H1	40K	H1	66B	H1	6	H1
4K	H1	13N	H1	24L	H1	32N	H1	40L	H1	67I	H1	13	H1
5K	H1	14L	H1	24P	H1	33Y	H1	41L	H1	82W	H1	**14**	**H3**
7K	H1	14N	H1	25K	H1	34K	H1	45P	H1			16	H1
8I	H1	15K	H1	26K	H1	35P	H1	45R	H1			19	H1

Legend: H= Haplotype, WP= Woodchester Park, SWWR= Secret World Wildlife Rescue

Figure 46. Haplotype network showing low variation from majority concensus sequence



Legend: Circle size proportional to number of badgers per haplotype (H1=57, H4=1, linked haplotypes H2 and H3=2). Dark grey= Woodchester Park sample set, Pale grey= Outgroup sample set.

Haplotypes 2 and 3 (H2 and H3) are both transitions occuring at nucleotide 1007 which form a homozygote, and heterozygote, respectively, changing cytosine to thymine. This causes a missense mutation translating a threonine to an isoleucine at amino acid 336. Haplotype 4 (H4) is a transversion of guanine to cytosine at nucleotide 1722, leading to a missense mutation of glutamine to histidine at amino acid 574 (Table 22).

Table 22. Three variations found in TLR2

Haplotype	Nt. Variation	SNP type	A.A. Variation	Mutation	Badger	% Pop.
1	WILD TYPE					95%
2	T HoZ 1007 A**C**A > A**T**A	Transition	Threonine (T) >Isoleucine (I) 336	Missense	1	1.6%
3	C/T HeZ 1007 A**C**A > A**T**A	Transition	Threonine (T) >Isoleucine (I) 336	Missense	14	1.6%
4	C/G HeZ 1722 CA**G** > CA**C**	Transversion	Glutamine (Q) >Histidine (H) 574	Missense	30L	1.6%

Legend: HoZ= Homozygote, HeZ= Heterozygote, number denotes position along nt or AA sequence.

H2 and H3 were seen in outgroup badgers located 73 miles apart. It is interesting to note that H2, the homozygous mutation of nt1007 was found in a badger with a sibling of identical age and gender (both 2.7kg males when found) and yet this sibling was homozygous for the wild type nucleotide at this point. These were both found in Locking, 44.3 miles southwest of Woodchester Park. Linked haplotype H3 was found in a badger found in Venn Ottery, 95 miles from Woodchester Park. This badger was sampled along with a cub from the same sett, but which bore the wildtype nucleotide at this point. H4, the unlinked haplotype, was found in a Woodchester Park badger, but was not shared by any member of its sett, nor of the setts in the area. Figure 47 shows locations where badgers with variable genes were found, but does not include the entire range of badgers sampled and found to bear the wild type from locations near Birmingham to St. Ives.

No variation was found across all samples for TLR4. In primer set 4p1, a sampling of 36K showed what looked to be an A/C heterozygote polymorphism. Peaks matched and were smaller than surrounding peaks. However, further PCR and sequencing showed that 36K was homozygous for C as with all other badgers (Figure 48). Sequences for almost all 4p1 primed PCR samples were very clear and showed no variation. It was only in some of the more degraded out group samples that noise occurred enough to create need for a second round of PCR. A low resolution sample of Badger 12 from Secret World Wildlife Rescue, had a questionable A peak in an area normally homozygous for C. A second trial had clear, sharp peaks, and proved there to be no variation (Figure 49).

Figure 47. Map showing locations where variable haplotypes were found.

Legend: WP=Woodchester Park (location of badger with H4), Lo= Locking (location of badger with H2), SM= Shepton Mallet (sample area, no variable haplotypes found), VO=Venn Ottery (location of badger with H3),

Figure 48. Two PCR samplings and sequencings of sample 36K

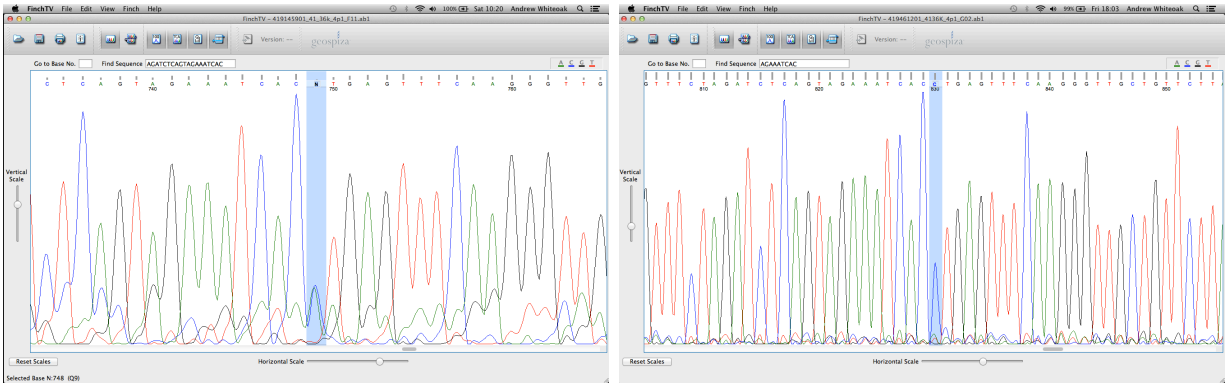
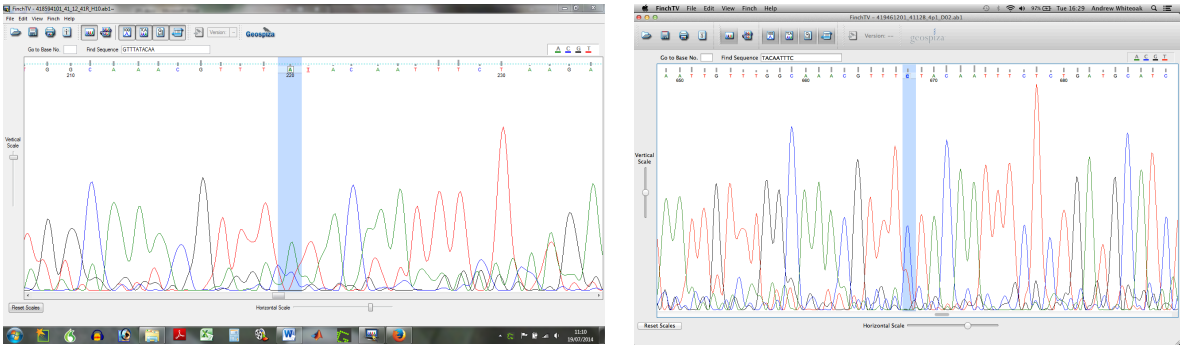
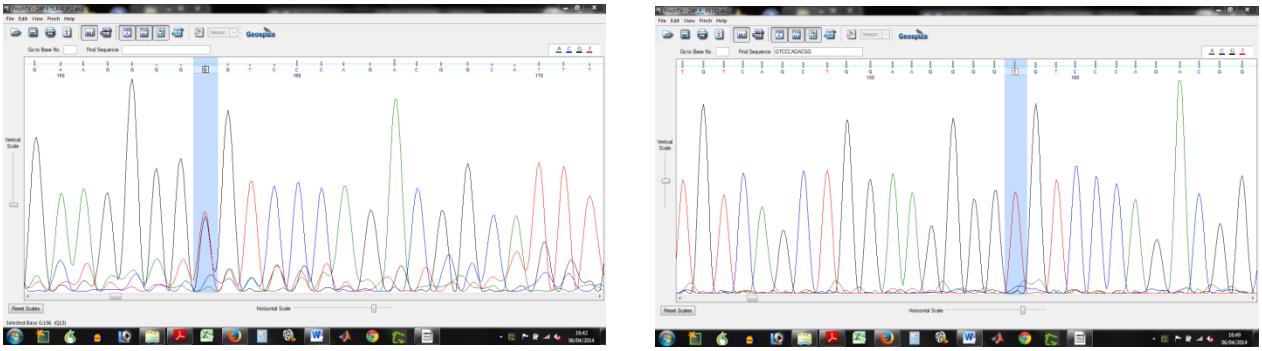


Figure 49. Anomolous peak in low resolution data resolved after a second PCR trial



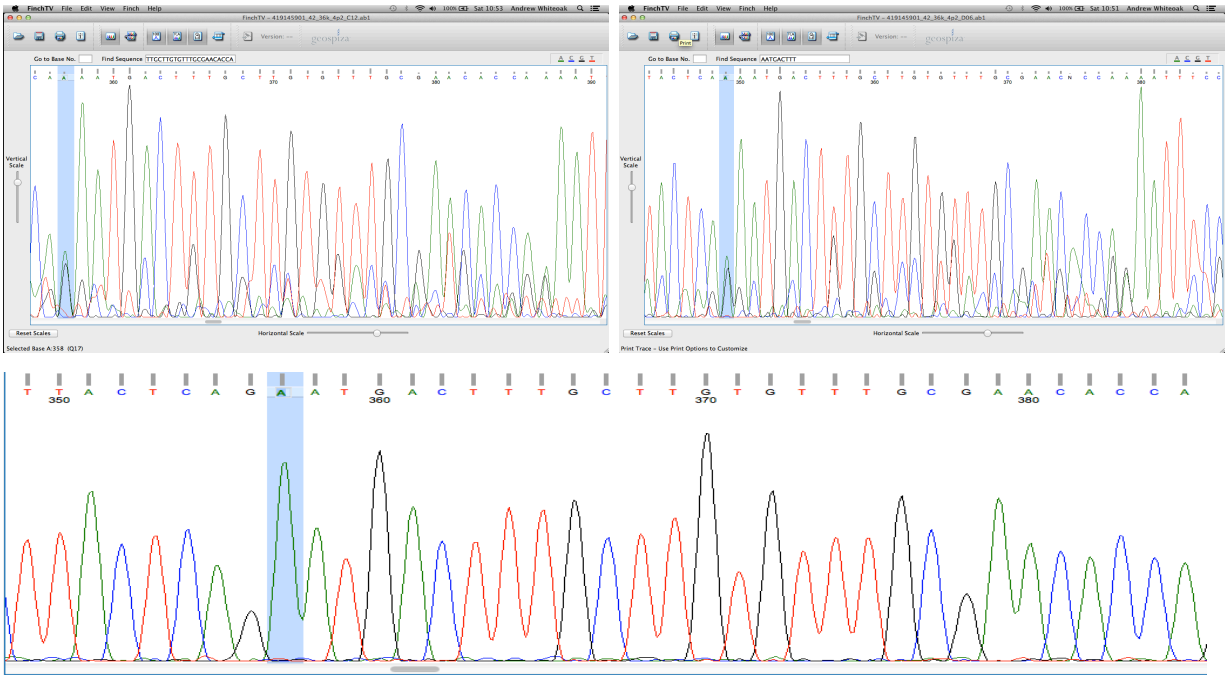
Neither was there any variation found in primer 4p2. A possible heterozygous G/T nucleotide was seen in the first sampling of 24P. This sample underwent fresh PCR and was shown to in fact bare no such variation (Figure 50).

Figure 50. Questionable peak for 24P primer 4p2 resolved within two PCR trials.



Two possible variations were disproved in sample 36K; A/G peaks were seen where wild types are normally G homozygous, are seen at points 358 and 387 in the first trial, and at points 349 and 379 in second trial. Finally, a third trial found these points to be G homozygous once all noise was cleared up (points 357 and 386) (Figure 51).

Figure 51. Three trials to find no variation for 36K with primer 4p2



There were no variations found across all nucleotides sequenced, for all fifty-nine badgers tested over 1909 base pairs of TLR4 exon 3 which covers all of the LRR region. The primer set for TLR4 exon 2 was sequenced across only three Woodchester badgers; 9P, 13L and 50N. All three produced an identical sequence but no more were pursued due to extreme difficulty in

amplifying this area. By aligning sequence sections so that the end of one primer set was aligned with the beginning of the next, the known TLR4 sequence was arrived at (Table 23). This was translated to a partial amino acid sequence using the function in MatLab (Table 24).

Table 23. Badger sequence for TLR4

Exon 1 (96bp) Unknown

Exon 2 (165bp) Complete

GTTCCCAACATTACTTACAAATGCATGGAGCTGAATCTCAACAAAATTCCCAACAACATTCCTACATCAACCAAGAAACTGGA
CCTGAGCTTTAATCCCCTGAGGCATTTAGGCAACCATAGCTTCTCCAACTTCCCAGAACTACAGGTGTTGGATTTATCCACA

Exon 3 (1909/2238 bp) First 15 unknown, middle 1909 known, final 314 unknown.

ACTCGAATTATAGCATATCAGGGCCTAAACCACCTCTCCATCTTGATATTGACAGGAAACCCTATCCAGAAGTTTTCCACAGA
GGCCTTTTCTGGACTTTCAAGTTTACAGACTCTGGTGGCCGTGGAGACAAACCTACGGTCTCTAAAGGACCTACCCATTGGAC
ATCTCAAAAACTTGAAGGAGCTTAACGTGGCTCACAATCTTATCTGCTCCTTCAAGTTACCTGAATATTTTTGTAACCTGACC
AACCTGGAGTACTTGGATCTTTCCAATAACAACATCAAAGATATTTATCATAGTGACTTGCAGGTTCTACATCAAATGCCTCT
ACTCAACCTTTCTTTAGACTTGTCCTTGAACCCTTTATCCTTTATCCAGCCAGGTGCCTTTAAAGAAATTAAACTCCATGAAC
TGACTTTGAGAAGTAATTTTAATAGTACAGATGTAATGAAAACTTGTATTCAAGGTCTGGCTGGTTTAAAGATCCATCACTTG
GTTCTGGGAGAATTTAAAAATGAAAGGAACTTGGAAAGCTTTGACAAATATCTCCTGGAGGGACTGTGCAATTTGACCATTGA
AAAATTCCGCATAGCATTCTTTAATGAGTTCTCAGAGGATATCACTGACTTATTTAATTGTTTGGCAAACGTTTCTACAATTT
CTCTGATGCATCTGTTTTTAAACAGACCACGACACCTTCCTAAAAATCTCAGATGGCAACGGCTGGAAATGGTTAACTGTGAA
TTTGAAGAATTTCCCAAATGGGAGCTGGACTCTCTCAAGGAGTTTGTTTTCACTGCCAACAAAGGCGTGAGCACTTTTACTGA
GATGAAGCTGGAAAGCCTTGAGTTTCTAGATCTCAGTAGAAATCACCTGAGTTTCAAGGGTTGCTGTTCTTACTCTGACTTGG
GGGCAACCAGACTGAAGCATTTAGATCTGAGCTTCAATGATATTATTACCATGAGTTCAAACTTCTTGGGCTTAGAACAGCTA
GAATATCTAGATTTCCAGCATTCCAATTTGAAACAGGGCAGTGACTTTTCAGTATTCCTATCCCTCAGAAACCTCCGTTACCT
TGATATTTCTTATACTCATACCCAAGTTGTCTTCCGGGGCATTTTTGATGGCTTGGTCAGCCTCCAAGTCTTGAAAATGGCTG
GCAATTCTTTTCAGGACAACTTTCTCCCAAATATTTTCAAAGACCTGACTAACTTGACCATTCTAGACCTCTCTAAGTGTCAG
CTGGAAGGGGTGTCCCAGACGGCATTTGGCTCACTTCCTAAACTTCAGTTGATAAATATGAGTCACAACAACCTCTTGTCATT
GGATATACTCCCTTATGAGCCTCTCCTCTCTCTCCAAATTCTGGATTGCAGTTTTAATCGAATAGTGGCCTCCACGGAGCAAG
TACGACAGCATTTCCCAAGTAATCTAGTTTCCTTAAATCTTACTCAGAATGACTTTGCTTGTGTTTGCGAACACCAGAATTTC
CTGCAGTGGGTCAAGGACCACAGGCAGCTCTTGGTGGAAGTGGAAAAAATGGTGTGTACCAAACCTTTAGACATGCAGGACAT
GCCCCTGCTGAGTTTTAGGAATGCCACCTGTCAGAGGAGCAAGACTATCATTACTGTGTCAGTGTTCACTGTACTCATGGTTT
CTTTGGTAGCAGTTTTGGTGTATAAGTTCTATTTCCACCTGATGCTTCTTGCTGGTTGCAAAAAGTACAGCAGAGGCAAGAGC
ACCTACGATGCCTTCGTTATCTACTCAAGCCAGGATGAAGACTGGGTGAGGAATGAACTGGTAAAGAACTTGGAGGAGGGGGT
GCCTCCCTTTCAGCTCTGCCTTCACTACAGAGACTTTATCCCTGGTGTGGCCATCGCCGCCAATATCATCCAGGAAGGCTTCA

Table 24. Partial TLR4 Amino acid sequence for the badger

VPNITYKCMELNLNKIPNNIPTSTKKLDLSFNPLRHLGNHSFSNFPELQVLDLST*****TRIIAYQGLNHLSILILTGNPIQK
FSTEAFSGLSSLQTLVAVETNLRSLKDLPIGHLKNLKELNVAHNLICSFKLPEYFCNLTNLEYLDLSNNNIKDIYHSDLQVLH
QMPLLNLSLDLSLNPLSFIQPGAFKEIKLHELTLRSNFNSTDVMKTCIQGLAGLKIHHLVLGEFKNERNLESFDKYLLEGLCN
LTIEKFRIAFFNEFSEDITDLFNCLANVSTISLMHLFLNRPRHLPKNLRWQRLEMVNCEFEEFPKWELDSLKEFVFTANKGVS
TFTEMKLESLEFLDLSRNHLSFKGCCSYSDLGATRLKHLDLSFNDIITMSSNFLGLEQLEYLDFQHSNLKQGSDFSVFLSLRN
LRYLDISYTHTQVVFRGIFDGLVSLQVLKMAGNSFQDNFLPNIFKDLTNLTILDLSKCQLEGVSQTAFGSLPKLQLINMSHNN
LLSLDILPYEPLLSLQILDCSFNRIVASTEQVRQHFPSNLVSLNLTQNDFACVCEHQNFLQWVKDHRQLLVEVEKMVCTKPLD
MQDMPLLSFRNATCQRSKTIITVSVFTVLMVSLVAVLVYKFYFHLMLLAGCKKYSRGKSTYDAFVIYSSQDEDWVRNELVKNL
EEGVPPFQLCLHYRDFIPGVAIAANIIQEGF

Legend: The first 32 amino acids of exon 1 are unknown, the 55 amino acids of exon 2 are shown. The first 5 amino acids of exon 3 are unknown as are the final 105; the central 636 amino acids are shown.

3.3 Variation in relation to trypanosome Infection

Badger infection status was determined by Eze Justin Ideozu, using DNA extracted from the Woodchester Park sample group by Hadil Alkathiry by the phenol-chloroform method described earlier. PCR methods used are described in section 2.0. Twenty-nine out of eighty-two badger DNA samples amplified using ITS-Nested PCR producing band sizes of 1271 bp which indicated the badgers were positive for trypanosomes. This gave a prevalence of 35.4% (25.9% - 46.2%; 95% CI) (Ideozu et al unpublished). Raw data giving infected and non-infected individuals is shown in Table 25.

Table 25. Raw data showing badger infection status.

Badger ID	Status	Badger ID	Status	Badger ID	Status	Badger ID	Status
24P	1	28K	1	12K	1	24L	0
59B	0	2P	0	58R	1	67Y	0
10N	0	15K	0	7K	0	21K	1
13L	1	38K	0	38L	1	40K	0
26K	1	50N	1	13N	0	5L	1
3K	0	31K	1	66B	0	33K	1
1K	0	25N	1	9N	1	23N	0
17K	0	1L	0	9L	0	2K	0
39K	0	40L	0	33Y	0	24K	0
22K	0	31L	0	30L	0	29K	1
17F	1	4K	1	37K	0	29N	1
2F	0	14N	1	32K	0	30K	0
36K	0	34K	0	23K	0	6L	0
38Y	1	27K	0	41K	0	10K	0
5K	0	35K	1	8L	0	11K	0
12L	1	17L	1	41L	1	13K	0
35P	0	28F	0	15L	1	9K	0
82W	0	8I	0	63Y	1	3N	0
45P	1	25K	0	32N	0	22N	0
26N	0	13F	1	45R	0		
9P	1	34P	0	14L	0		
Total	8		9		7	5	29

Legend: 1=infected (shaded green for visual reference) , 0=Not infected.

The Infection status and its relation to the badger's movements were compared. Sixty-four badgers never left their home sett, while fifteen were found in a new location. A null hypothesis was formed, stating that *infection status is the same for both static and mobile badgers*. For this data it was not possible to use Chi-square test because the number of

mobile infected badgers is only four, which well is below the statistical threshold for Chi-square validity for one degree of freedom, which suggests that all elements in the table should be above ten. This is because Chi Square uses the normal distribution as an approximation for the binomial distribution which is valid when 'n' is large.

Because badgers are either infected or not, and mobile or not, we are dealing with dichotomous categorical variables that can have a value of '1' or '0' and thus Fisher's Exact Test may be used. Unlike Chi-square it directly uses a combinatorial calculation that relates the probability to the number of successes in a set of 'n' events, and so is valid when dealing with very small numbers. Fisher's Exact Test also takes account of the fact that for every 'success' the probability of further success is altered as proportions are changed. For example, when testing for infected badgers in a fixed population, the discovery of an infected badger is a 'success'. This alters the probability of finding another as proportions of 'successful' badgers in the remaining, untested, population are lowered. This behaviour means that the hyper geometric distribution should be used as opposed to the binomial distribution. By using the hyper geometric distribution Fisher's Exact Test allows us to examine whether our small data set has significant differences between mobile and static badgers regarding trypanosome infection. The number of infected vs. noninfected and mobile vs. static badgers gives a two by two contingency table (Table 26).

Table 26. Fisher's "Exact" Test

	Infected	Not infected	Column sum
Mobile Badgers	4	11	15
Static Badgers	25	39	64
Row sum	29	50	79
		P value	0.16

This can then be used to work out the exact probability that the observed data could happen by chance. The contingency table shows what was observed, and these observations indicate that infection rates are not the same; so the test asks the data whether the observed difference is significant.

Fisher's Exact Test calculates the P value using the following equation:

$$P = \frac{(a+b)!\ (c+d)!\ (a+c)!\ (b+d)!}{a!\ b!\ c!\ d!\ n!}$$

Where a = mobile infected badgers, b = mobile non-infected badgers, c = static infected badgers, d = static non-infected badgers and n = all badgers. When this is calculated in Excel the P value is 0.16. Thus there is 16% chance the data could have occurred randomly. Therefore, we must fail to reject the null hypothesis. Either, there is no significant relationship between infection status and badger movement, or, we do not have enough diversity of observations to find significance.

Of the fifty-two badgers which were fully sequenced across TLR2, only one (30L) held a variant haplotype (Table 27) Fifteen were infected by trypanasomes, while thirty-seven were not. 30L was among those not infected. Due to the low diversity of variant gene data it is not possible to perform a correlation which would show any significant relation between variation and infection.

Table 27. Fifty-two Woodchester Badgers sequenced across ALL primers in TLR2.

1K, 1L, 2F, 2P, 3K, **4K**, 5K, 7K, 8I, 8L, **9P**, **12K**, **12L**, **13L**, 13N, 14L, **14N**, 15K, 17K, **17L**, **21K**, 22K, 22N, 24L, **24P**, 25K, **26K**, 26N, 27K, 28F, **30L**, 31L, 32N, 33Y, 34K, 35P, 36K, 38K, **38L**, 39K, 40K, 40L, **41L**, **45P**, 45R, 50N, **58R**, 59B, **63Y**, 66B, 67I, 82W.

Red highlight= Infected with Trypanosomes 15/52

Black underlined (30L) = Single Woodchester badger with variant haplotype G1722C/Q574H

3.4 Comparing Badger TLR Sequences with Those of Other Species

Variant haplotypes exist at two locations on TLR2. The first part of analysis involved plotting where in the protein these haplotypes are located, and what changes they make to the protein. Basic Local Alignment Search Tool (BLAST) software was used to compare the amino-acid sequences found to homologous proteins across species. A Position-Specific Iterated BLAST (PSI-BLAST) was chosen because it outputs all known proteins within the query sequence according to their position along the protein, so that their precise location along the TLR can be identified. Table 28 was created by zooming in on a BLAST output of TLR2 by eye to examine each amino acid one by one and identify where each known protein begins, ends or overlaps another. Haplotypes two and three move the position of a Leucine Rich Repeat named COG4886 (unknown function) from amino acids 330-494 to 359-535, effectively enlarging this domain by 12 amino acids. These same haplotypes also alter the position of an E3 Ubiquitin Protein Ligase named PRK15370 SirP from amino acids 334-559 to 392-538, shortening this domain by 79 proteins. Haplotype 4 adds a three additional amino acids to a Leucine Rich Repeat called SMART0082 in the C-terminal domain, which normally sits between amino acids 534-584, but in haplotype 4 sits between 534-587.

Table 28. Badger TLR2 with locations of related proteins highlighted

MSRVLWTVWVLGAVTKLSKEEGPDQASSLSCDLTGVCDGRFRSLKSIPSGLTAAVRSLDLSNNEITYI
RNRDLRGCVNLKALKLASNRINAIEEDSFISLRSLEHLDLSYNLLSNLSSSWFRPLSSLKFLNLLGNR
YKSLGEMPLFSPLTNLQILKVGSIDSFTELQEKDFAGLSFLEELEIDASNLQRYEPESLKSIQNISYL
ALRMKQPIFLLEIFGDLSRSLKHLELRDTHLNTFQFSKASIRETNTLIKKWTFRNVKITDGSFSELVK
LLNCVSGVLEVEFEGCTLDGLGNFDISDMDKIKNIGGIETLIVRRLAIPHFYSFYDMSSIYSLTANVK
RVTVESSKVFLVPCLLSQHLKSLEYLDLSDNLMVEESLRNSACDQAWPLLQTLILRHNRLKSLEKTGE
TLLSLKNLTKLDISKNNYVSMPETCQWPDKLKYLNLSNTRIYSVTRCIPWMLEILDISNNNLDSFSLI
LPRLKELYISGNKLKTLPDASFLPTLRILRISRNIISTFTKEQLDSFSTLEALEAGGNNFFCSCEFLS
FTREQQSLAQILTDWPDNYLCDSPFSVRGQRVKDTRLPASECHRVALVSAVCSVLFLLILLTGVLCHH
FHGLWYLKMMWAWLQAKRKPRKAPPRDVCYDAFVSYSEHDSYWVENMMVQELEHFDPPFKLCLHKRDF
IPGKWIIDNIIDSIEKSHKTIFVLSENFVKSEWCKYELDFSHFRLFDENNDAAILVLLEPIEKKAIPQ
RFCKLRKIMNTKTYLEWPTDETQQEGFWLNLRMAIKS

Legend: Four Leucine Rich Repeats (LRR) were identified; three LRR8 (shown by text coloured magenta, purple and red) and one LRRCT (text colour lime). Two of the LRR8 proteins were found to overlap and this area is highlighted by green text. There is also an LRR multi domain, COG4886, highlighted in grey. The function of these Leucine Rich Repeats are as yet unknown. There is also one Toll-interleukin 1 receptor domain (TIR) 2 superfamily (blue text) known to function in signal transduction and is found at the cell surface interface.

A new search of TLR2 in the order Carnivora was made (September, 2014) and two new TLR sequences were found, *Ursus maritimus* and *Panthera tigris*. These, in addition to the TLR2 sequences collected at the beginning of the project, as well as badger haplotype 1 were collated for phylogenetic analysis. A clustal alignment of these sequences was performed to check whether the changes were in regions of mutability for other species. Haplotypes 2 and 3 refer to T>I variation at aa position 336, which at the nucleotide level are homozygous and heterozygous respectively. This was found to be a T across all Carnivores with no variation, so the local variation here is unusual (Figure 52.)

Figure 52. Clustal Alignment showing amino acid position 336 in TLR2.

```
Tiger          VGKIGSVGGIETLTVRRLVIPYFYSFYDLSSVYSLTERVKRITVESSKVFLVPCLLSQHL
Cat            MGKIGSVGGIETLTVRRLVIPYFYSFRDLSSVYSLTERVKRVTVESSKVFLVPCLLSQHL
Dog            VDKIKNIGQIETLTVRRLHIPHFYSFYDMSSIYSLTEDVKRITVESSKVFLVPCSLSQHL
Badger         MDKIKNIGGIETLIVRRLAIPHFYSFYDMSSIYSLTANVKRVTVESSKVFLVPCLLSQHL
Ferret         MDKIKNIGGIETLIVRRLAIPYFYSFYDMSSIYSLTANVKRVTVESSKVFLVPCLLSQHL
Walrus         TDKIKSLGGIEILTVRRLYIPYFYSFYDMSSIYSLTVEVKRVTVESSKVFLVPCSLSQHL
Polar_Bear     MDKIKNIGGIETLTVRRLYIPYFYSFYDMSSIYSLTEDVKRVTVESSKVFLVPCSLSQHL
Panda          MDTIKNIGGIETLTVRRLYIPYFYSFYDMSSIYSLTEDVKRVTVESSKVFLVPCSLSKHL
                 .* .:* ** * **** **:**** *:**:****  ***:************ **:**
```

Legend: Green highlight T amino acid for all Carnivores at posn 336.

Haplotype 4 is a Q>H heterozygous varaiation at position 574. It too shows no variation from Q across all Carnivores, and thus the local H variation is unusual (Figure 53.)

Figure 53. Clustal Alignment showing amino acid position 574 in TLR2.

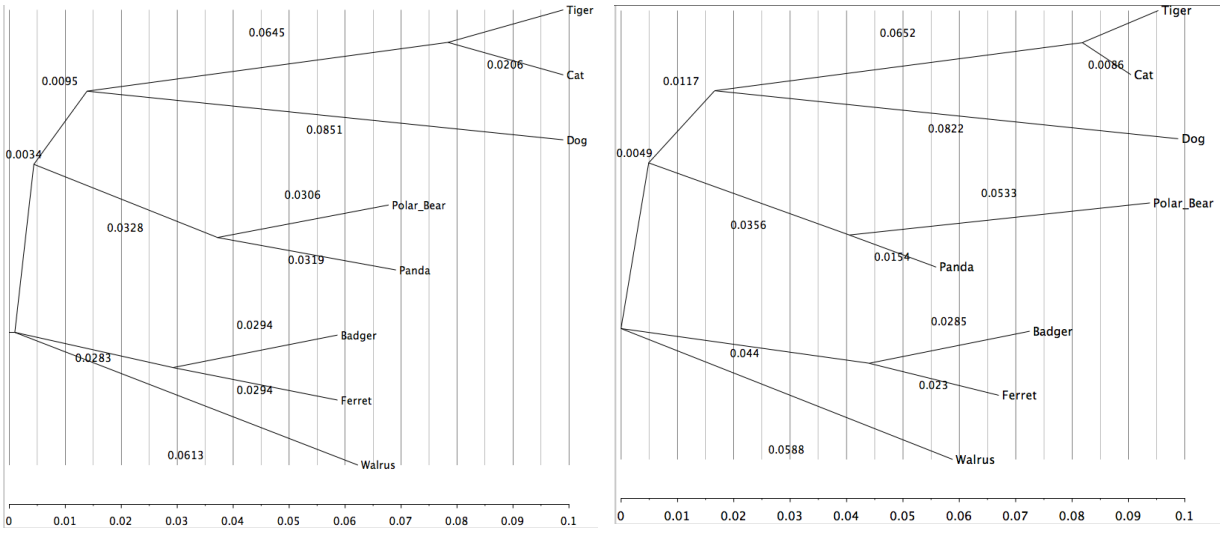
```
Tiger          EFLSFTREQQALAQTLTDWPHNYLCDSPFYVRGQRVQDTHLPASECHRAALVSAVCCVLF
Cat            EFLSFTREQQALAQILTDWPDNYLCDSPFYVRGQRVQDTHLPASECHRAALVSVVCCVLF
Dog            EFLSFTQEQQALAGLLVGWPEDYLCHSPSYVRGQRVGTARLPASECHRTALVAAVCCVLL
Badger         EFLSFTREQQSLAQILTDWPDNYLCDSPFSVRGQRVKDTRLPASECHRVALVSAVCSVLF
Ferret         EFLSFTREQQSLAQILTDWPDNYLCDSPFSVRGQRVKDTRLPASECHQVALVSAVCSVLF
Walrus         EFLSFTQEQRALVQILIDWPENYLCDSPFSVRGQQVQDTRLPASECHRAALVSAVVSVLL
Polar_Bear     EFLSFTQGQQALAQILTDWPENYLCDSPFYVRGQRVRDTRLPVSECHRAALVSAVCSVLF
Panda          EFLSFTQGQQALAQILTDWPENYLCDSPFYVRGQRVRDTRLPVSECHRAALVSAVCSVLF
               ******: *::*.  *  **.:***.**  ****:*  ::**.****:.***:.* .**:
```

Legend: Green highlight T amino acid for all Carnivores at posn 336.

In order to understand how TLR2 compares to a published phylogenetic tree of the order Carnivora, phylogenetic tree maker software was used. Prefered software changes quickly, and while PHYLIP was the product of choice until recently, today available programmes are no longer maintained. A current programme of choice is FigTree (Dudas and Rambaut, 2014; Bouckaert et al 2014), both as a graphical viewer of phylogenetic trees and a program for

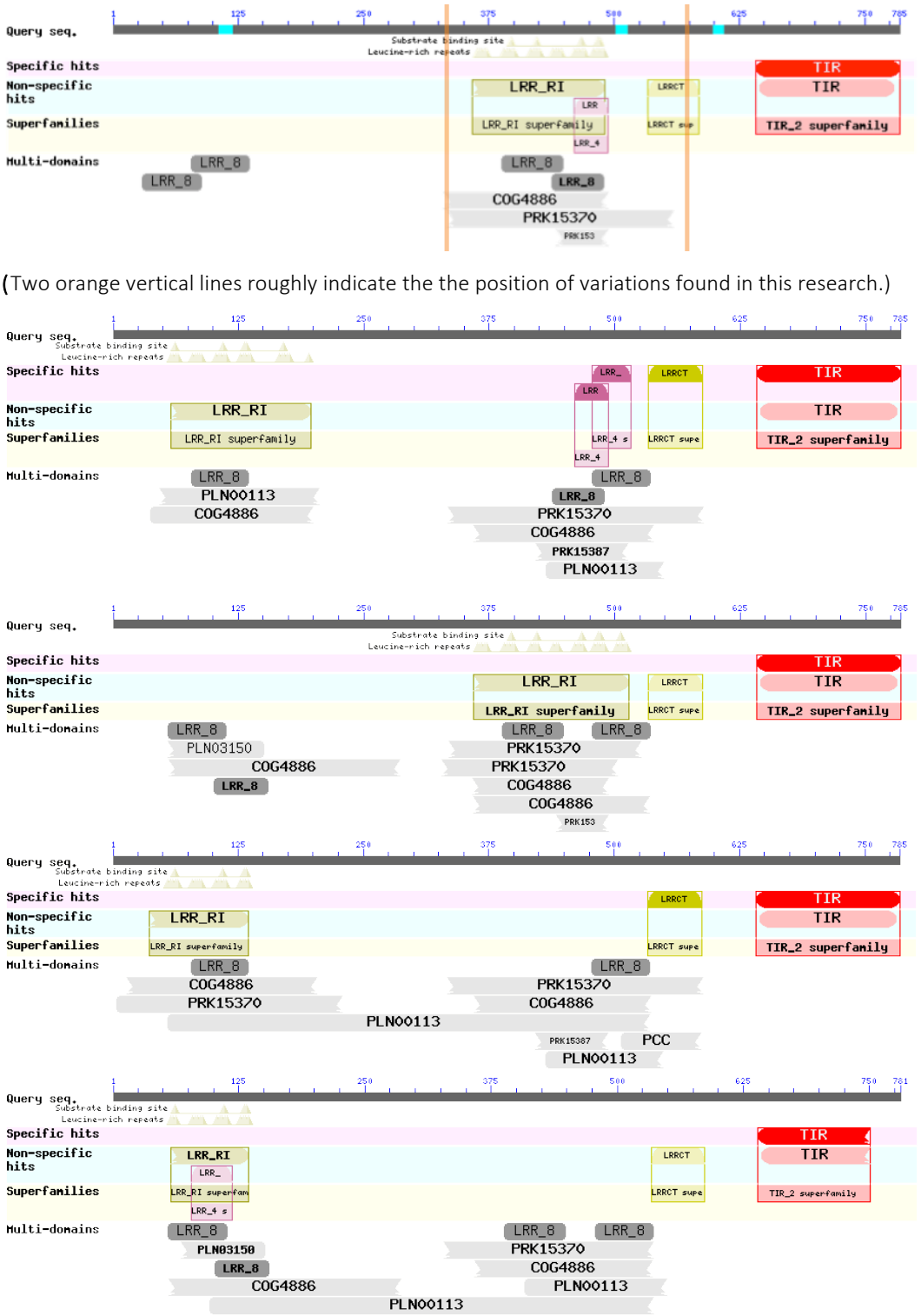
producing publication-ready figures. This method estimates phylogenies from distance matrix data under the "additive tree model" where distances are equal to the sums of branch lengths between species. Clustal alignments for TLR2 were made by ClustalW Phylogeny, formatted with FigTree, and placed next to a published Carnivora phylogenetic tree for comparison (Figure 54). While our presentation of the tree used for comparison is a simplification of Nyakatura and Bininda-Emonds' (2012) tree containing sixty animals (distances were divided down as distance increases with the amount of species included) it is noted that the tree using TLR2 sequences involves an originally limited array of carnivores as only this limited set of TLR2 sequences are currently available. Due to this the relative lengths of branches are artificial, and the following interpretations, though interesting, are merely putative and will benefit from further work into TLR sequence studies across other carnivore populations. Our comparison postulates that, for TLR2, the polar bear may have evolved far more rapidly than the panda; and, to a more subtle degree the tiger TLR2 shows a more rapid arms race than the cat; while the badger seems to have far more evolved TLR2 than the ferret; and the panda appears to have comparatively slow development in TLR2. PSI-BLAST was performed for all available carnivores in order to compare TLR2 configurations with *Meles meles*. Percentage similarities between TLRs across species were calculated (Cargill and Womack, 2007) (Figures 55 and 56).

Figure 54. Phylogenetic tree of Carnivora adapted from Nyakatura and Bininda-Emonds, (2012) vs. TLR2 Cladogram using distance matrix method.



Legend: Published tree left, TLR2 based tree right. Scale shows relative evolutionary distances from a common ancestor

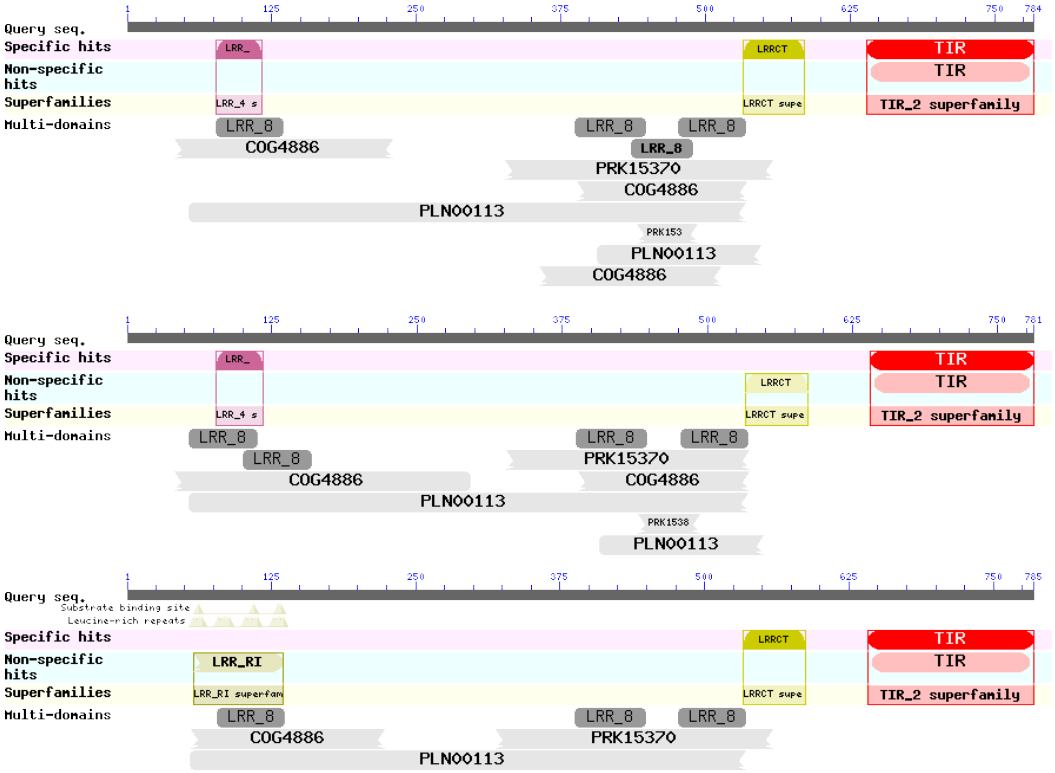
Figure 55. Badger TLR2, then ferret, panda, walrus and polar bear TLR2 which have 95%, 89%, 88% and 87% similarity to the badger.



(Two orange vertical lines roughly indicate the the position of variations found in this research.)

Legend: COG4886, CD00116, SMART00082, pfam12799, pfam13855 and PLN00113 are all Leucine Rich Repeat domains. TIGR00864 is a Polycystin Cation Channel protein. PRK15370 and PRK15387 are E3 Ubiquitin protein ligases. smart00255 and pfam01582 are in the TIR homology domain. Functions unknown.

Figure 56. Cat, tiger and dog TLR2 which have 86%, 85% and 83% similarity to the badger.



Legend: COG4886, CD00116, SMART00082, pfam12799, pfam13855 and PLN00113 are all Leucine Rich Repeat domains. TIGR00864 is a Polycystin Cation Channel protein. PRK15370 and PRK15387 are E3 Ubiquitin protein ligases. smart00255 and pfam01582 are in the TIR homology domain. Functions unknown.

The protein domains of each animal were compared in detail and it was found that while proten types were homologous across species, the numbers of homologous proteins varied between species (Table 29). These comprise Leucine Rich Repeat domains; COG4886, CD00116, SMART00082, pfam12799, pfam13855 and PLN00113: The Polycystin Cation Channel protein TIGR00864: The E3 Ubiquitin protein ligases PRK15370 and PRK15387 and smart00255 and pfam01582 of unknown function, which are found in the TIR homology domain. The biggest difference between TLR2 genes of different species is the total number of Leucine Rich repeat domains COG4886 and pfam13855, while the number of proteins located in the TIR homology domain were most similar between species (Figure 57). The dicovery of polymorphisms within the LRR region is in line with other TLR studies. Cuscó et al 2014 listed locations of 31 non-synonymous SNPs found across all ten canine TLRs across 435 canids. Six were in or close to Leucine Rich Repeat regions, three were in or close to TIR domains, and the remaining twenty two were located in the sensor domains of the TLRs.

Table 29. Domain Hits for all Carnivores listed for TLR2

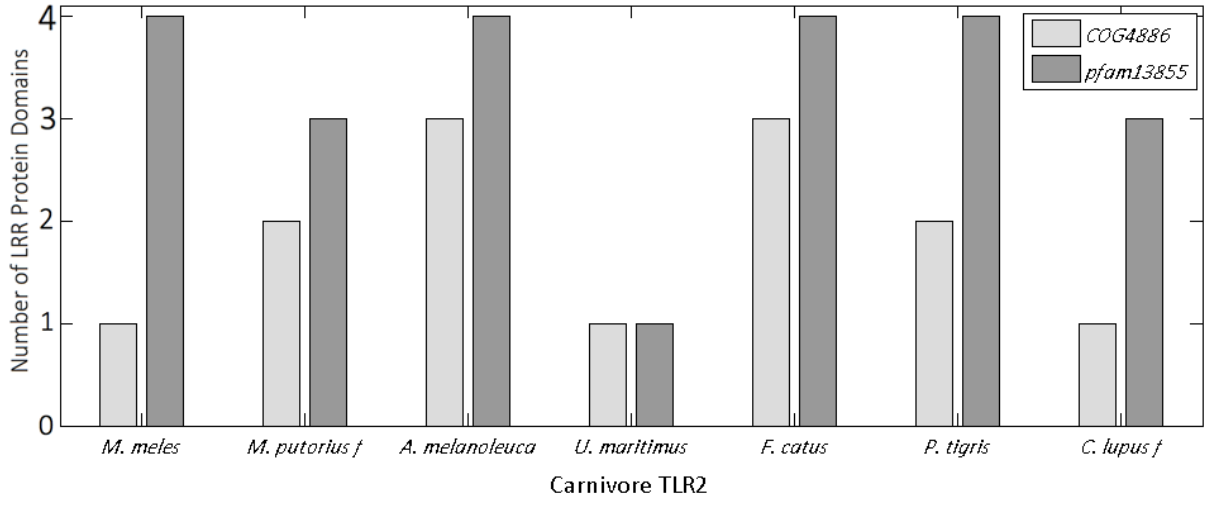
Domain	Badger	Ferret	Panda	Walrus	Polar Bear	Cat	Tiger	Dog
COG4886	330-494	359-539; 38-200	358-494; 359-535; 56-286	14-204; 360-537	359-535	389-536; 42-230; 357-513	389-536; 42-296	56-223
CD00116	358-491	58-197	359-514	37-135	58-135			58-135
SMART00082	534-584	534-587	534-587	534-587	534-587	533-586	533-586	534-587
pfam12799	460-494	478-516; 460-494			78-118; 55-113; 388-449; 101-155; 478-536	77-117	77-117	
pfam13855	388-449; 78-136; 438-490; 30-89	78-135; 478-536; 438-490	55-113; 388-449; 478-536; 101-155	78-135; 478-536	101-155	77-135; 477-535; 387-448; 436-489	54-112; 477-535; 387-448; 100-159	78-136; 478-536; 388-449
PLN00113		47-206; 432-549		55-536; 432-549	96-536; 408-549	54-535; 407-548	54-535; 407-548	55-536
TIGR00864				507-586				
PRK15370	334-559	334-588	328-503; 335-526	4-229; 334-588	328-536	327-558	327-536;	320-559
PRK15387	442-494	421-513	442-494	421-494		441-493	441-493	
smart00255	641-785	641-785	641-785	641-785	639-751	640-784	640-781	641-785
pfam01582	644-780	644-782	644-782	644-782	642-751	643-781	643-788	644-782

Legend: Internal structure count:

One Two Three Four Five

In the domain column, proteins highlighted in red are Leucine Rich Repeats, all are of unknown function, however it is known that CD00116 belongs to a Ribonuclease Inhibitor -like subfamily, SMART00082 is found in the C-terminal domain, pfam12799 and pfam13855 are structurally different and these structures are identified as LRR4 and LRR8 respectively, and PLN00113 is an LRR-like protein kinase. The protein highlighted in green is a Polycystin Cation Channel protein, proteins highlighted in purple are E3 Ubiquitin protein ligases and proteins highlighted in orange belong to the TIR homology domain.

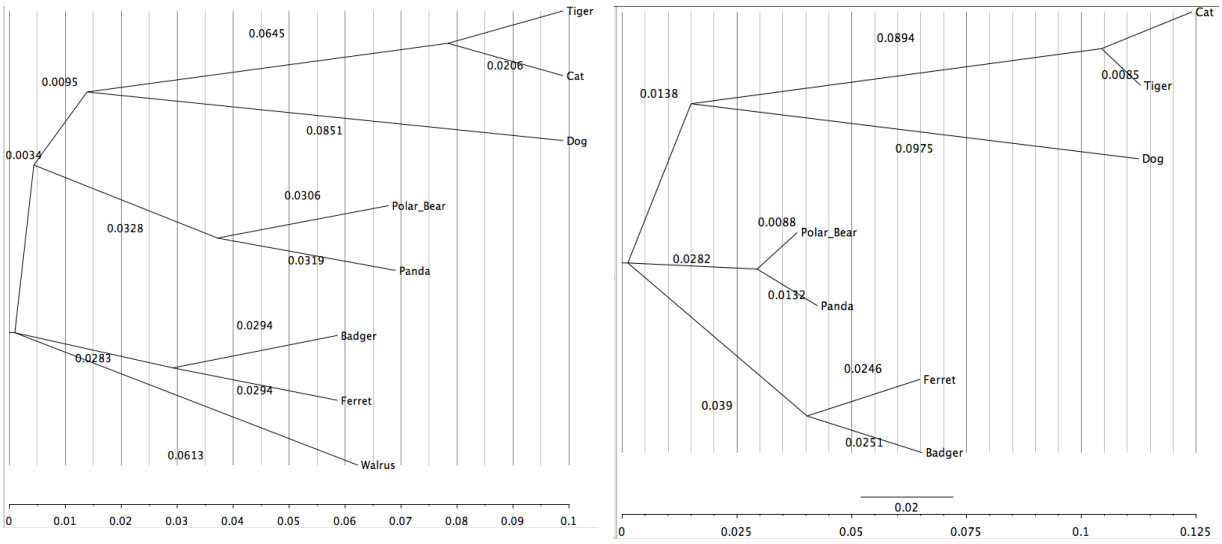
Figure 57. Leucine Rich Repeat density across species in TLR2



Legend: COG4886 is a Leucine Rich Repeat superfamily and pfam13855 is a Leucine Rich Repeat, which are most variable in their numbers found in TLR2 are most variable across species.

As with TLR2, phylogenetic analysis was also performed for TLR4 and compared to the phylogenetic tree for Carnivora (Nyakatura and Bininda-Emonds, 2012) (Figure 58). The walrus was excluded as its relative position in the TLR4 tree was extremely advanced in comparison to all other carnivores. Again, due to the low number of carnivore TLR4 sequences available, the relative lengths of branches in the TLR4 tree are artificial and all interpretations are putative and would benefit from further investigation into TLR4 sequences in other carnivores. The putative tree suggests that TLR4 phylogeny the cat is possibly more positively selected than the tiger, which contrasts with published phylogenetic relationships. It also postulates that phylogeny whereas the polar bear appeared to show positive in the TLR2 phylogeny, in TLR4 its postion does not vary from the published tree. And, that the badger and ferret show no putative increase in positive selection on TLR4, which may explain why no variation was found in the badger population for TLR4 while variation was found for TLR2.

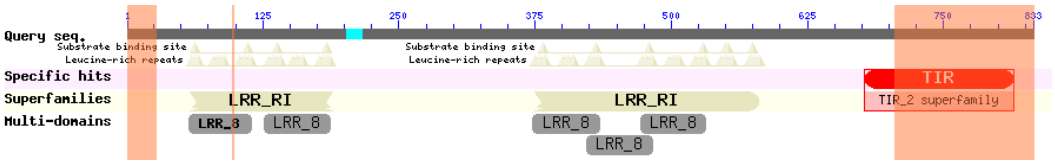
Figure 58. Phylogenetic tree of Carnivora adapted from Nyakatura and Bininda-Emonds, (2012) vs. TLR4 Cladogram using distance matrix method.



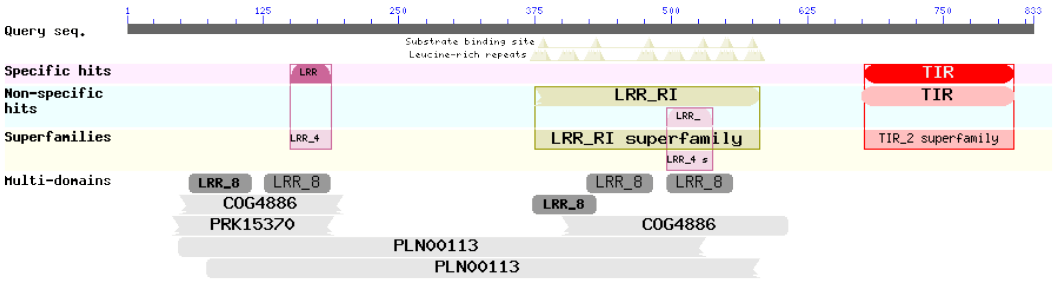
Legend: Published tree left, TLR4 based tree right. Scale shows relative evolutionary distances from a common ancestor

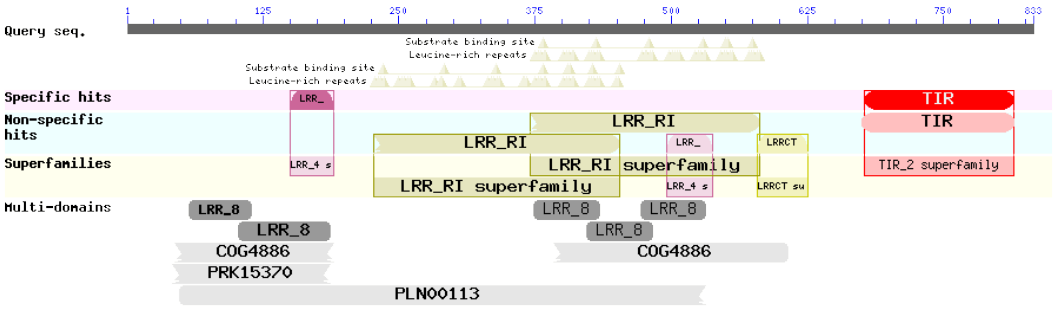
PSI-BLAST was performed for all available carnivores in order to compare TLR4 configurations with *Meles meles*. Percentage similarities between TLRs across species were calculated for the portion sequenced in the badger (Cargill and Womack, 2007) (Figures 59 and 60).

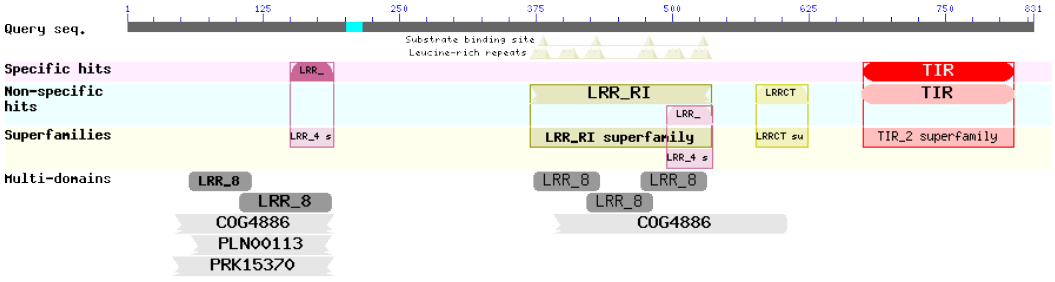
Figure 59. Badger TLR4, then ferret, polar bear, panda and dog TLR4 which have 95%, 91%, 90% and 84% similarity to the badger.

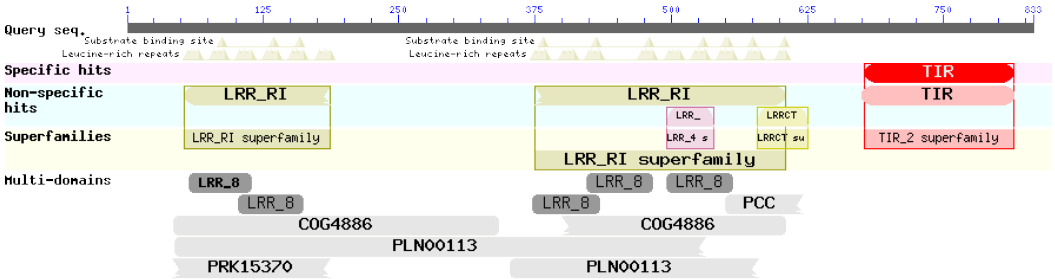


(Areas NOT yet sequenced in the badger are highlighted in orange)



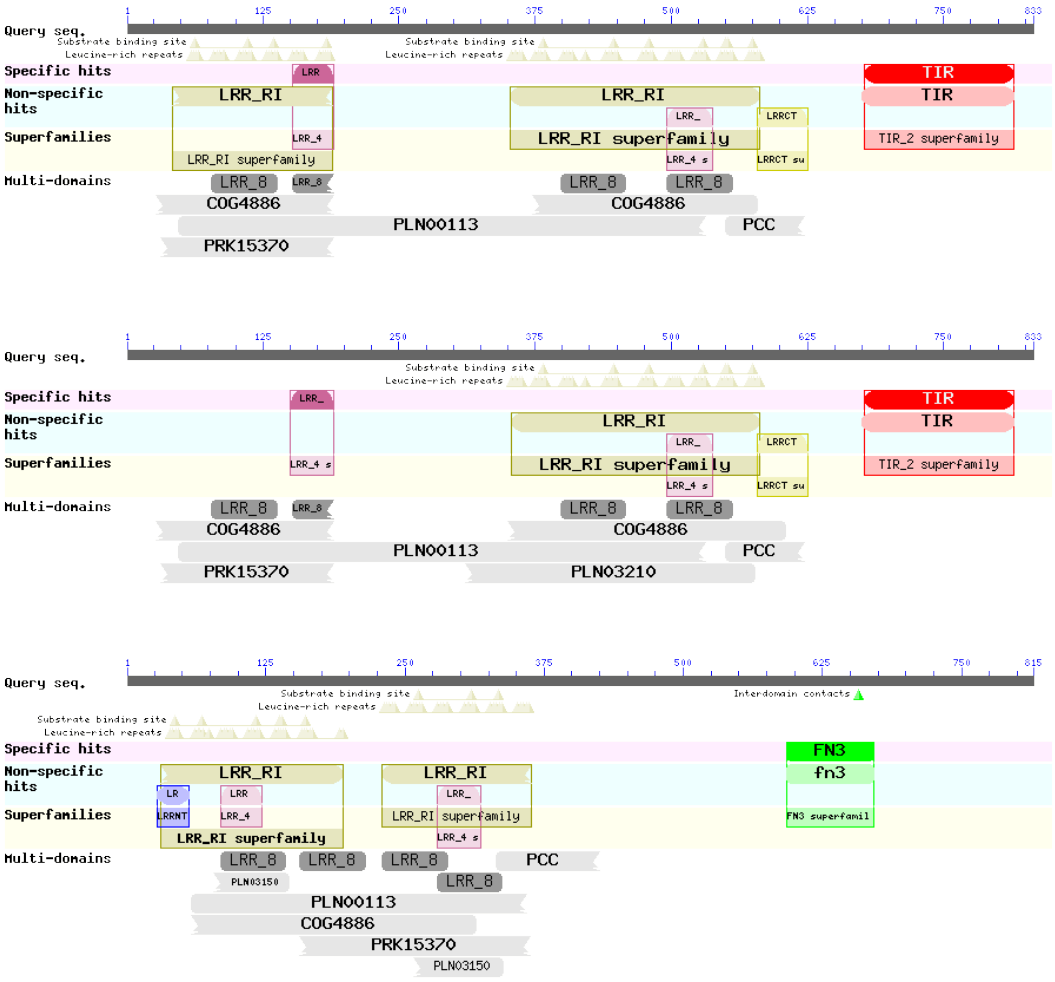






Legend: COG4886, CD00116, SMART00082, pfam12799, pfam13855 and PLN00113 are all Leucine Rich Repeat domains. TIGR00864 is a Polycystin Cation Channel protein. PRK15370 and PRK15387 are E3 Ubiquitin protein ligases. smart00255 and pfam01582 are in the TIR homology domain. Functions unknown.

Figure 60. Tiger, cat and walrus TLR4 which have 84%, 83% and no similarity to the badger.



Legend for X through Y: COG4886, CD00116, SMART00082, pfam12799, pfam13855 and PLN00113 are all Leucine Rich Repeat domains. TIGR00864 is a Polycystin Cation Channel protein. PRK15370 and PRK15387 are E3 Ubiquitin protein ligases. smart00255 and pfam01582 are in the TIR homology domain. Functions unknown.

The protein domains of each animal were compared in detail and it was found that while proten types were homologous across species, the numbers of homologous proteins varied somewhat less between species than they had done in TLR2 (Table 30). The proteins found were the same as those found in TLR2. The TIR domain is the largest area not sequenced in this project and also appears to be the area where variation is least likely as it shows no significant variation across species. TLR4 shows greater homology than TLR2 across species. There is no variation in the amounts of protein COG4886 which was a highly variant factor across TLR2 in different species. In protein Pfam13855 only the felids differ from other carnivores in levels of this protein.

Table 30. Domain Hits for all Carnivores listed for TLR4

Domain	Badger	Ferret	Polar Bear	Panda	Dog	Tiger	Cat
COG4886	44-190; 400-607	49-199; 400-607	44-190; 392-607	44-190; 391-605	43-341; 400-605	27-190; 373-579	27-190; 350-605
CD00116	58-189; 375-581	375-581	371-581; 227-453	370-536	53-187; 375-605	352-581; 42-189	353-581
SMART00082			579-626	577-624	579-626	579-626	579-626
pfam12799	150-190; 496-538	150-188; 496-538	150-190; 496-538	150-190; 495-537	496-539	96-538; 152-190	150-190; 496-538
pfam13855	57-114; 126-187; 373-434; 422-483; 472-532;	57-144; 126-187; 373-431; 422-483; 496-556;	57-114; 472-532; 103-187; 422-483; 374-434	57-114; 103-197; 373-433; 471-531; 421-482;	57-114; 103-162; 373-434; 422-483; 496-556;	78-138; 496-556; 399-458; 152-190	79-138; 152-190; 399-458; 496-556;
PLN00113	47-532	47-532 73-581	48-532	59-187	44-532; 352-580	47-532	47-532
TIGR00864					550-621	550-622	550-622
PRK15370		42-190	42-187	42-190	42-187	31-190	31-190
smart00255	674-816	674-816	674-816	672-814	674-816	674-816	674-816
pfam01582	677-815	677-815	677-815	675-813	677-815	677-815	677-815

Legend: Internal structure count

One Two Four Five

In the domain column, proteins highlighted in red are Leucine Rich Repeats, all are of unknown function, however it is known that CD00116 belongs to a Ribonuclease Inhibitor -like subfamily, SMART00082 is found in the C-terminal domain, pfam12799 and pfam13855 are structurally different and these structures are identified as LRR4 and LRR8 respectively, and PLN00113 is an LRR-like protein kinase. The protein highlighted in green is a Polycystin Cation Channel protein, proteins highlighted in purple are E3 Ubiquitin protein ligases and proteins highlighted in orange belong to the TIR homology domain.

4.0 Discussion

DNA was extracted from the tissues of one hundred and six badgers, which were represented by eighty-two from Woodchester Park and twenty-eight from outgroup sources. Bioinformatic searches at the beginning and end of the study found no TLRs have yet been sequenced for the badger. Primers designed in this study successfully amplifed TLR2 and TLR4 (exon 3) after PCR optimization. The single exon gene of TLR2 was sequenced in its entirety across sixty-one badgers, fifty-two from Woodchester Park and nine from outgroup sources. TLR4 exon 3 was sequenced across fifty-nine badgers; forty-nine from Woodchester Park and ten from outgroup sources. TLR2 sequencing showed three variant haplotypes of non-synonymous SNPs which produced missense mutations at the amino acid level. No DNA sequence variations were found in TLR4. No correlation was found between TLR variation and trypanosome infection status. Badger TLR2 was found to bear positive selection due to its higher ratio of nonsynonymous to synonymous mutations. When Badger TLR consensus sequences were compared to those of other species, phylogenetic comparisons showed that badger TLR2 was strongly selected for, while badger TLR4 showed little positive selection.

Early stages of the investigation revealed an absence of TLR4 (exon 3) DNA sequence variation across the first fifty-nine Woodchester Park badgers sampled, and no variation was found in early stages (first 32 badgers) sequenced across TLR2. Due to this it became clear that the results would preclude the analysis of any form of correlation with the 40% trypanosome infection found in this collection of badgers. In order to mitigate against the possibility of not being able to achieve the objective of investigating TLR variation with respect to trypanosome infection, outgroup badgers were sought to further explore TLR polymorphism. A variety of sources (veterinarians, rescue workers and academics) were contacted to investigate possible sources of specimens to act as geographical outgroups. Two groups responded, Secret World Wildlife Rescue (SWWR) and University College Dublin were able to donate samples. This could only occur from April onwards, in relation to the badger season and when cubs were two months old and able to given health checks. Furthermore, the beginning of the delayed 2013 badger cull determined when badgers were being rescued in southern England and the midlands. Thus parameters beyond the control of the study restricted the diversity of samples. Samples from SWWR ranged in all directions from

Woodchester Park, for up to 200 miles, which made it possible to rule out any inbreeding should low variation be found due to the geographical isolation by distance. Samples from Woodchester Park were given with a map detailing their sett locations and a spreadsheet details place of birth and the sett each badger was inhabiting at the time of sampling. It was interesting to note that most badgers never left their home sett. Extraction was highly successful when applied to the 400ul filled tubes of red blood cells donated by SWWR. However, the standard phenol-chloroform extraction method, used in this study may not be the best method for production of DNA from small quantities of white blood cells. Unfortunately, the small quantities of material and time constraints precluded further investigation of extraction methods on this occasion.

The bioinformatic search for TLR genes proved very fruitful. Due to the novel nature of TLR investigation, the number of available TLR gene sequences on web based databases nearly doubled during the course of the project, providing a substantial bank for comparison to the sequenced badger TLRs. However, even on the basis of the sparser amount of TLR information at the beginning of the study, primer design was very successful. Once exon regions were identified, Multiple Sequence Alignments (MSAs) allowed the conserved regions across species to be noted. From these regions the first primers were selected. It was only after some sequencing had been undertaken that it was realised how very close badger and ferret sequences were. After this it was possible to use the ferret TLR sequence as a reference when making MSAs to specifically select conserved regions were the least variance was found between the ferret and all other creatures. Investigations into primer design proved very useful in selecting the optimal parameters to target specific sequences. Primer length was found to improve specificity more than any other variable, and detection of possible primer dimers was also a key parameter. Balanced temperature settings were found to be of relatively low import, it was also found that no known equation, including nearest neighbour evaluations, can accuretely predict optimal temperature settings better than performing a gradient PCR in lab. Adding degenerate bases to code for any of two or three nucleotides at one base position was found to have little relative impact on specificity, and so could be embarked upon with the confidence of knowing that countering such choices with increased primer length would prevent the possibility of amplifying non-target regions.

Defining the most important parameters of primer design in this way lead to cost- and time-efficient creation of successful primers.

During PCR optimisation $MgCl_2$ gradients were performed as suggested in the Stratagene Robocycler manual. Following gradient PCR a handful of samples still did not amplify at optimum temperatures which successfully amplified TLRs in all other samples, or would produce additional bands not seen in the majority of samples. Where non-target bands were smaller than the target band size $MgCl_2$ gradients were run where $MgCl_2$ levels were incrementally decreased in proportion to the master mix until the larger band appeared alone. Where additional bands appeared that were larger than the target amplicon, $MgCl_2$ levels were incrementally increased in proportion to the master mix until the smaller target band appeared alone. This was especially useful when attempting to isolate and identify the contaminating band from primer set 2p2. In this case the band was isolated when $MgCl_2$ loading was doubled. It was also found that increasing cycles allowed more amplicon to be produced, this helped maintain strong bands were annealing time was reduced.

As well as being automatically analysed by FinchTV software, the data were also examined manually. All variations and noise were confirmed or retested via additional rounds of PCR and sequencing. TLR4 showed no sequence variation across both Woodchester and outgroup badgers. However, non-synonymous variation which caused missense mutations at the amino acid level was found in TLR2. Two percent of badgers from the Woodchester Park set showed variation from the majority consensus sequence, while in outgroup sets twenty-two percent of badgers showed TLR2 sequence variation, this may point to Geographical Isolation by Distance (GID). Analyses of parasite burden in Woodchester park found no significant correlation between place of birth and infection, nor any significant variation across setts; disease was fairly evenly distributed across the whole of Woodchester Park with a mean 37.2% of badgers infected. There were too few roaming badgers to make any significant correlation between roaming and disease susceptibility. As well as finding non-synonymous mutations that indicated positive selection in TLR2 but not TLR4 in the population sampled in this study; phylogenetic tree comparisons of these genes to those of other species showed the consensus badger TLR2 gene positioned much further along the distance matrix in comparison to other species than the badger TLR4 gene, inferring stronger evolutionary

selection on badger TLR2. This suggests that badger TLR2 is being driven to such selection by an evolutionary arms race against an as yet unidentified pathogen or parasite, and that further investigation of badger TLR2 in relation to key badger diseases would be useful.
TLR population studies are a recent phenomena and data is sparse. In human studies the number of possible haplotypes and total gene variance is often discarded in favour of investigating less rare, known, polymorphisms. In a study of 640 volunteers from Romania and the Netherlands Ioana et al (2012) focused on three TLR2 polymorphisms and found consistent levels of their appearance. For all three a variant homozygous nucleotide was rare, comprising from none to 0.16% of the population. The heterozygous variations found were at levels of 0.78%, 4.53% and 6.09% in populations. These results are very similar to our own findings for TLR2, with respect to low amounts of variation. Across all badgers in this study, heterozygous variation occurred in 1.6% of the total population and homozygous variation in 3.2% of the population. Similarly, in volunteer populations of 580 and 905, respectively, Mockenhaupt et al (2006) and Reismann (2009) found 1.21%, 3.45% and 0.11% had homozygous variation, while between 4.83% 22.24% and 11.49% had heterozygous variation, across three pre-selected nucleotides on TLR4. In contrast to this, badgers in this study were found to have absolutely no variation across TLR4 gene sequences. Given the populations studied above are over ten times larger than our study population, and yet yield some, low, variation in TLR4, this may point to a need to continue the TLR4 across a wider sample range.

Hashemi-Shahri et al (2014) examined known TLR8 variations in humans and found that 27% of a population of 175 had a homozygous variation, while 29% had a heterozyougous variation. Liu et al (2012) examined a known TLR9 variation and found that 19% of a population of 432 had a homozygous variation, while 33% had a heterozygous variation. Zhang et al (2014) examined four known TLR9 variations and found that on average across all four variations examined 13% of a population of 854 had a homozygous variation, while 26% had a heterozyougous variation. Carvalho et al (2007) assayed known polymorphisms across 388 Portuguese blood donors. 11.1% and 10.8% were heterozygous for TLR4 Asp299Gly and Thr399Ile (TLR4) while 0% were homozygous for either. A TLR9 variant allele (T-1237C) was present in 17.3% as heterozygous mutation and 2.1% as homozygous. Martínez-Ríos et al (2013) studied two known polymorphisms within 740 Mexican volunteers. They found that for TLR-4 Asp299Gly, the AA homozygote occurred in 93% of the population, the AG

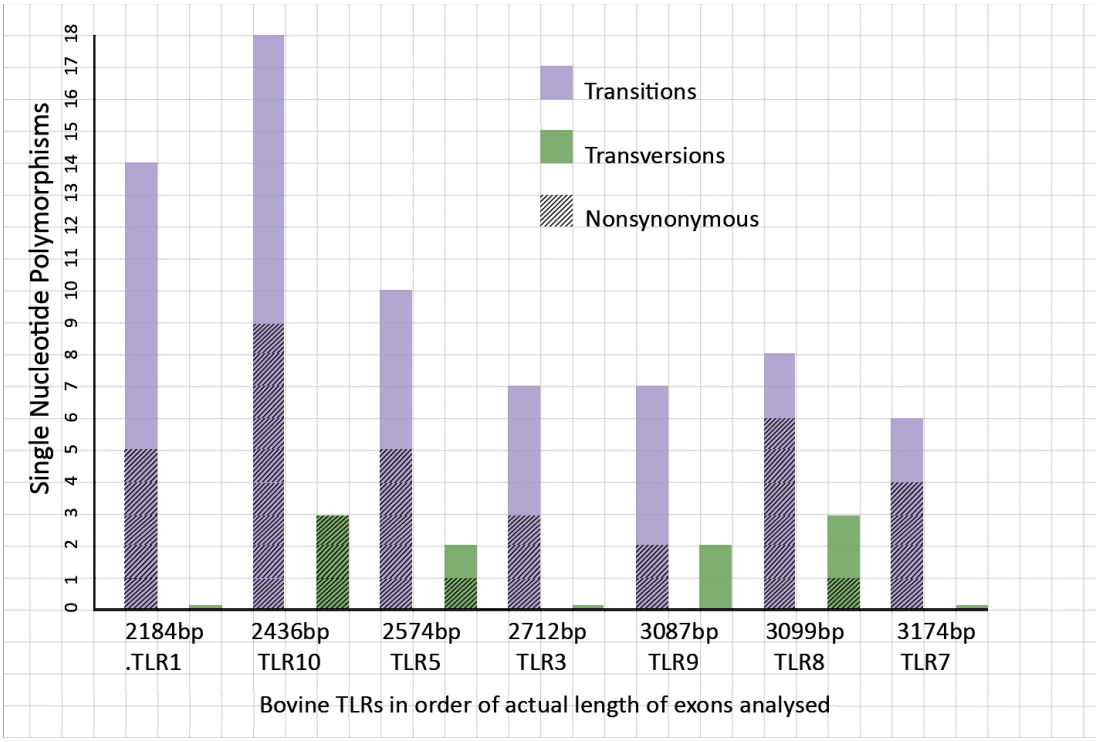
heterozygote in 7% and a GG homozygote never occurred; and that for TLR-4 Thr399Ile, the CC homozygote occurred in 97% of the population, the CT heterozygote in 3% and a TT homozygote never occurred.

These studies, taken alone, possibly indicates how unusual our findings were in terms of low levels of TLR sequence variation across a population. However, other studies of known polymorphisms found equal, or even less, variant polymorphisms. Kim et al (2012) investigated six known TLR polymorphisms; TLR1 (Arg80Thr), TLR2 (Arg753Gln and Arg677Trp), TLR4 (Asp299Gly and Thr399Ile), TLR6 (Ser249Pro) across 322 Korean volunteers and found no polymorphisms between them. Rodriguez-Osorio et al (2013) found unusually low variation in two known SNPs, which was ascribed to the ethnicity of the population tested. Of 170 Mexicans, only 5 were heterozygous for both TLR4 D299G and TLR4 T399I polymorphisms. Emingil et al (2007) found an unusually low amount of known polymorphisms in 245 Turkish subjects. For SNPs TLR2 Arg753Gln, TLR4 Asp299Gly and TLR4 Thr399Ile they found 10.1%:0%, 4.5%:0.3% and 3.4%:0% homozygous: heterozygous variation respectively. Weng et al (2014) looked at population variation in ten known TLR4 polymorphisms and found that a maximum of >2% of the 11,319 people studied bore variant for any SNP studied.

Some studies went further, looking for all polymorphisms in a single gene fragment, but often failed to report the distribution of these genes within the population studied. For TLR2 and TLR4 Mukherjee et al (2014) found eight and nine non-synonmous (NS) mutuations in a population of 266 people, while Grueber et al (2012) found one and four NS mutuations in a smaller study of 23 *Petroica australis takiura*. Charbonnel et al (2014) found ten NS mutations for TLR4 across 300 *Myodes glareolus*. Abrantes et al (2013) reported on the genetic diversity of TLR3 across 80 European rabbits from Portugal, France and Spain, where they found 14 non-synonymous variations mostly in the LRR region except for only one in the TIR domain. Morger et al (2014) performed a population study comparable to our own study of setts across Woodchester Park, examining woodmice collected from sites 500m apart around a quadrant roughly 6750m^2 near Malham Tarn. This looked at variation in TLR11 and TLR12, and found six and thirteen haplotypes, bearing respectively four and nine nonsynonymous mutations, respectively, across 120 *Apodemus Sylvaticus*. Half that number of badgers, from

Woodchester Park, were examined in this study, in an area roughly equal to the Malham study, but only two TLR2 haplotypes were found in this area; the wild type consensus sequence borne by fifty-eight badgers sequence, and haplotype 2, an SNP leading to one nonsynonymous mutation. No variation was found for TLR4 across these badgers. Additional variation found due to outgroup sampling. Thus our findings show unusually low variation within a core sample population of animals in a geographical range where contact between them is not heeded by landscape barriers.

Figure 61. SNPs in bovine TLRs from Seabury et al (2007) and Cargill and Womack (2007).



Seabury et al (2007) found 92 SNPs and 6 indels, across three TLRs in 10 cow breeds, 45 of which were located in intron areas. Cargill and Womack (2007) found 130 SNPs (97 of which were in introns) and 9 indels across four TLRs in the same range of cows. This study focused only on the exons of TLRS, however looking at the exon data alone there was still a greater range of variations found in the cow studies; a mean of 15.6 per TLR in the Seabury study and 8.25 per TLR in the Cargill and Womack study. The data is still not entirely comparable as in each of the above studies, two cow species were sampled: *Bos taurus*, *Bos indicus* as well as eight breeds derived from them; Angus, Charolais, Holstein, Limousin, Brahman, Nelore, Braford, Piedmontese and Romagnola, whereas a single species with no hybrids were

considered in this study. Geographical distribution analysis was not possible as these breeds cover an international range. Given this, the higher. level of variation found may well be a factor of the amount of breeds and the massive geographical distribution. Further analysis of the data (Figure 61) presented in both papers showed that the level of variations found did not correlate to the amount of length of fragments analysed, thus variation is not evenly spread across the TLRs and some (TLR10 in cattle especially) are more positively selected than others. Unfortunately, neither paper looked at TLR2 or TLR4, as comparative levels of variation here would have been fascinating in relation to this study.

To date, only Tschirren et al (2014) have produced a study which investigates both the range of haplotypes and their distribution within a population: Ten out of fifteen haplotypes were found to be nonsynonymous across a fragment of 49.87% of the TLR2 exon (1173/2352bp) were shared across 726 *Myodes glareolus,* in a 0.25km^2 area of Kalvs Moss, distributed across 292, 203, 50, 25, 24, 21, 20, 14, 2 and 1 *M. glareolus* respectively. Compared to our own, over ten times as many animals were sampled in this study, but just over three times as many nonsynonymous haplotypes were found. Thus, just as with our study variation was relatively low, and the comparison suggests that further investigation of badgers could well expand our current repertoire of TLR sequence variation.

TLR variation is important in disease. Studies reveal that TLR polymorphisms lead to a change in the degree of innate responsiveness, either increasing or decreasing the inflammatory response can be both destructive or protective depending on the context of specific diseases tested. TLR polymorphisms have been found to be significantly related to both disease onset risk and prevention. In patients with pre-existing disease, where polymorphism prevalences were not significantly different to healthy controls, instead significant links were often found between either the diseases' development or protection against its development.

The following studies all found polymorphisms which bestowed a significantly increased risk of disease onset. A great deal of work focused on TLR2 and TLR4, just as in our study, but was limited to testing previously known polymorphisms. ALL disease onset associations listed here are 'significant'. TLR4 Asp299Gly, TLR4 Thr399Ile, and TLR2 Arg753Gln were associated with early-onset and severe preeclampsia in a study of 270 pregnant women (Xie et al 2010).

TLR4 Asp299Gly was also significantly higher among 137 scrub typhus patients versus 134 controls from South India (Janardhanan et al 2013). TLR2 indel (196 to 174) was associated with risk of gall bladder cancer, and, non-synonymous TLR4 Exon 4 1936C>T was associated with cancer risk in females with gallstones in a study of 490 people from India (Srivastava et al 2010). TLR2 rs893629 was associated with arterial thrombosis in a study of 3587 North American systemic lupus erythrmus patients (Kaiser et al 2014).

The prevalence of this, like a great number of disease-significant SNPs, proved to be dependent on ethnicity. TLR2 rs893629 was more common in African American patients (9% prevalence) than in white or hispanic patients (less than 2%). TLR2 (P631H) and TLR6 (P249S) were associated with increased susceptibility to skin infection across a study of 646 caucasians. TLR1 R80T resulted in lower IL-6 cytokine responses to *Staph aureus* infection (Stappers et al 2014). TLR9 1237-CC was significantly linked with cerebral malaria in 117 children studied in Uganda (Sam-Agudu et al 2010). TLR5 C1174T encodes a variant that destroys flagellin-induced signaling, and is significantly associated with increased risk recurrent cystitis (Hawn et al 2009). In a study of 358 Holstein cows four TLR9 polymorphisms (A945G, G1187A, G1401A, and C2788T), increased risks of metritis occurrence, while TLR2 polymorphism C9564T increased risk of cytologic endometritis (Pinedo et al 2013).

Further studies showed that relationships between TLRs and disease onset are not always 1:1; multiple polymorphisms can combine to promote disease where one polymorphism may have no significant affect. Sampath et al (2013) found that when two TLR4 SNPS coexisted (rs4986790 and rs4986791) there was significantly increased risk of Gram-negative infections in the 408 low birth-weight babies tested. Liao et al (2010) found that if two TLR9 gene polymorphisms (rs287084 and rs352140) occurred in tandem, they significantly increased risk of Graves' ophthalmopathy in males only, in a study of 471 Taiwanese volunteers. This study points to another issue when analysing TLR polymorphisms and their contribution to disease; many studies have found gender dimorphic affect. The reasons for this are not elucidated for TLR9, but are more clear cut for TLR7 and TLR8, which, in humans, are located on the X chromosome. Alagarasu et al (2015) found a significantly higher frequency of TLR8 rs3764879–rs3764880 haplotype C-A was in males with dengue hemorrhagic fever compared to healthy controls, but no such relationship occurred in females. Thus TLR analysis is a

multifaceted field which must take into account race, gender, and interacting effects of multiple polymorphisms. For example, eleven TLR polymorphisms, across 276 volunteers, were both significantly associated with HIV status and were also race specific: Nine SNPs in TLR1, TLR4, TLR6 and TLR8 in Caucasians, and two other SNPs, in TLR4 and TLR8, in African Americans (Willie et al 2014). Seventeen SNPs across TLRs 1 and 2, TLR4, and TLRs 6 through 10, were linked with allergic rhinitis and asthma. Prevalence of each SNP varied significantly between ethnicities (Gao et al 2010).

Conversely, the following studies all reveal polymorphisms which bestow a significantly increased protection against disease onset. Again, ALL associations listed here are significant. TLR2 Arg677Trp was associated with a lower risk of cytomegalovirus infection in adults but not infants in a study of 229 volunteers (Jabłonska et al 2014). TLR4 (Asp299Gly) significantly protected those with chronic gastritis from reaching the active ulcerative stage among 195 patients infected with *Helicobacter pylori* (Bagheri et al 2014). TLR4 T610C and T610C and TLR6 G14578A, produce lower risk of clinical endometritis (Pinedo et al 2013). TLR4 A896G was associated with protection from recurrent cystitis while TLR1 G1805T was associated with protection from pyelonephritis in a study of 987 menstruating caucasian women (Hawn et al 2009). TLR1 rs5743551-CC was associated with reduced risk of disease in 300 healthy controls compared to 702 Russian patients with coronary artery disease (Golovkin et al 2014). TLR3 rs3775296 formed significant resistance to photosensitivity and anemia compared to wild type carriers (Wang et al 2014). TLR6 Pro249Ser showed a significantly reduced risk for atherosclerosis in a study of 503 heart disease patients and 605 healthy controls (Hamann et al 2013). TLR6 rs3775073-CC was significantly associated with decreased risk of infective endocarditis in a study of 410 Russian caucasians (Golovkin et al 2015).

Where TLR polymorphisms were not found to be significantly related to either disease onset or its prevention, they were often found to play a significant role in its development in patients with disease. The following TLR polymorphisms are significantly related to disease development from early to later stages once a disease has already taken hold. ALL associations listed here are significant. Spelatas et al (2009) associated TLR4 T399I with a 2.4-fold increased risk of COPD development, in a study of 240 smokers. Nachtigall et al (2014) found that TLR2 Arg753Gln and TLR4 Asp299Gly were related to progression from sepsis to

septic shock in 145 critically ill patients. Oliveira et al (2013) found that TLR4 rs4986790 and TLR4 rs4986791 were associated with presence of autoimmune thyroiditis in 572 bipolar disorder patients vs. 202 healthy controls. Pine et al (2009) studied 201 acute-stage caucasian HIV patients from Seattle and found that TLR4 D299G and TLR4 T399I were significantly more frequent, but TLR9 1635G was significantly less frequent, among individuals with high viral loading. Zidi et al (2014) studied 130 cervical cancer patients and 200 healthy controls in Tunisia, and found that TLR3 rs3775290-CC and TLR4 rs4986790-CC are significantly associated with higher risk of developing cervical cancer following herpes infection. Matas-Cobos et al (2014) found that TLR3 rs3775291 and TLR6 rs5743795 play a significant role in disease development to severe and acute stages (respectively) across a study of 260 pancreatitis patients. Apinjoh et al (2013) found that TLR9 rs187084 was significantly associated with susceptibility to malaria, while a TLR1 rs4833095 significantly weakened bodily defences against hyperparasitaemia; both contributed to the onset of severe malaria, in a study of 971 children with malaria and 891 healthy volunteers from the Cameroon. Yang et al (2012) found a significant association between TLR9 1237T/C and increased risk of chronic kidney disease, and higher mean plasma IL-6 levels in 630 end stage renal disease patients from the Han Chinese population.

Combinations of TLR polymorphisms also acted to speed disease development where single polymorphisms had not weakened the immune system. The following combinations made a significant difference in hastening disease development. Peric et al (2014) found that while single polymorphisms taken alone showed no significant correlation to chronic Hepatitis C, if taken in combination, then, a polymorphic heterozygous TLR4 Asp299Gly together with a homozygous TLR4 Thr399Ile correlated with significantly higher viral loading. Wang et al (2014) found that intron based TLR7 SNP rs3853839-G>C played a significant role in systemic lupus erythematosus development in a study of 1957 females. Another non-synonymous polymorphism TLR8 rs3764880-G>C, based in an exon, was associated with oral ulcers. These two SNPs acting in combination were associated with pericardial effusion.

Just as some polymorphisms hasten disease progression, others protect against it. ALL of the following TLR polymorphisms confer significant protection against disease. Alagarasu et al (2015) sampled 120 dengue cases and 109 healthy controls from Maharashtra, Western

India, and found significantly lower frequencies of the TLR3 rs3775291 T allele in those with dengue hemorrhagic fever. This was associated with damage to the TLR resulting in a decreased inflammatory response which had a protective effect against disease progression. Devaraju et al (2014) studied a TLR9 (1237C/T) polymorphism and found that a wild type C allele conferred significant risk of disease development across 300 systemic lupus erythematosus patients, while a polymorphic T allele was found in a significant proportion of the 460 healthy controls. Also, just as while some polymorphisms had limited affect alone, polymorphisms also work in tandem to prevent disease progression. ALL of the following TLR polymorphic combinations had significant afffect in the prevention of disease development. A combined heterozygous TLR4 D299G/T399I was protective against heart disease among 125 Chagas Disease patients in northern Chile (Weitzel et al 2012). Holla et al (2010) found that 222 patients with chronic periodontitis bore different TLR9 haplotypes with a complex range of SNPs to the 259 unrelated controls studied in Czechoslovakia. No significant relationship was found between individual SNPs and this disease but analysis showed that the SNPs worked in combination to alter disease outcome. Castaño-Rodríguez et al (2014) found that a combination of TLR4 polymorhisms (rs10759931, rs1927911 and rs10116253) were protected against progression to gastric cancer in 310 Chinese patients infected with *Helicobacter pylori*.

As we have seen above some of the known TLR polymorphisms tested were found to be either detrimental or protective against disease and its progression. A study by Liadaki et al (2011) makes it clear that even for specific polymorphisms, results are not clear cut. In a study of two nonsynonymous TLR4 SNPs across 327 tonsillectomy patients with a history of tonsillitis, Liadaki et al found that TLR4 polymorphisms D299G and T399I predispose individuals to streptococcus infection on a three-fold scale compared to 245 healthy bone marrow donors, while the former polymorphism was protective against Haemophilus influenzae by two-fold.

As has already been discussed in detail in section "*1.1.1 Badgers and their role as Reservoirs for Infection*," badgers are implicated in bovine TB. They, alongside the brushtail possum in New Zealand, the African buffalo in South Africa, the white-tailed deer in the USA, and the European wild boar in Spain are all considered maintenance hosts (Le Roex et al 2013).

Current cull-based strategies to intervene with this are complex, controversial and generate polarised public feeling. Badger vaccination trials have shown success, but receive little financial support and no government backing. Current tuberculin skin tests on cattle have proven costly and unreliable, and depend on antibody detection which cannot differentiate between active infection and a host's immune defenses after overcoming disease.

Currently TB infects one-third of the world's population and causes 1.6 million deaths per year (Kulchavenya et al 2012). Only 5% to 15% of cases develop into active TB suggesting a key mediating role for genes of the innate immune system (Khan et al 2013). Studies suggesting that TLR polymorphism correlates with TB susceptibility are particularly pertinent to this study, and are detailed below. The absence of TLR2 in mice has been associated with TB susceptibility (Stenger et al 2002). However, more specifically, in a study of 474 African-Americans (Velez et al 2009) and a linked study of 381 Caucasians and 667 Africans (Velez et al 2010) an indel polymorphism (−196 to −174) in TLR2 was significantly associated with increased risk of TB. Khan et al (2013) added that this same polymorphism was significantly associated with TB susceptibility across 187 volunteers from Pakistan.

Not only TLR2, but others such as TLR1 have been implicated in bTB susceptibility (Le Roex et al 2013). Sun et al (2012) found a TLR1-G1596A gene variant was significantly associated with bTB susceptibility due to a reduction in PAMP recognition in a study of 586 Chinese Holstein cows. Wujcicka et al (2014) pointed out that polymorphisms, especially of TLR2 and TLR4, have been found to enhance congential transmission of TB. It is thus very possible that TLRs and the innate immune system might be involved in badger susceptibility to this disease. However, the research required to resolve this may require looking at multiplicit factors, such as polymorphisms acting in combination rather than alone, as well as variation between genders and between distinct population heritages, such as that between Irish and British badgers as discussed in section "*1.1.2 Badger Evolutionary History*". Simply looking at a small range of known polymorphisms may not be enough, and instead looking at full genes we did in this study, may be called for. For example, Selvaraj et al (2010) found no association between five known polymorphisms and TB prevalance across 206 TB patients and 212 healthy controls, having tested for TLR-1 1805T/G (Ile602Ser), TLR-2 2258G/A (Arg753Gln),

TLR-4 896A/G (Asp299Gly), TLR-4 1196C/T (Thr399Ile), and TLR-6 745C/T (Ser249Pro). However a study of all TLR variations may well have had more positive results.

As far as we are aware there are currently no publications on badger TLRs. No badger TLR has yet been published. There are no badger TLRs nor any badger full genome sequences on any DNA database. This research has begun to fill that gap.

Further development to the groundwork laid out here might include sequencing TLR2 and TLR4 from a wider sample set of badgers including outside UK. Or, sequencing all other TLRs from badgers. Cuscó et al (2014) studied the exon regions of all ten canine TLRs across 335 dogs from seven reeds and 100 wolves from two populations, and found TLR5, then TLR4, to be the most polymorphic TLR among canines. Indeed, while polymorphisms in both TLR4 and TLR5 are associated with increased resistance to inflammatory bowel disease in German Shepards, only the TLR5 polymorphism confers IBD resistance in 38 other breeds (Kathrani et al 2011). It would be interesting, therefore, to sequence all badger TLRs and deduce from their phylogenies which are most positively selected. We might also develop studies that investigate links between diseases and TLR polymorphisms, as decribed above. Or, investigate intron and flanking sequences for possible epigenetic marks that link with disease (e.g. CpG methylation); investigate other molecules of the innate immune system (e.g, MyD88, inflamasome components); or develop whole genome sequencing of badgers to reveal all possible SNPs involved with disease.

The TLR data generated by this study suggests a badger population bottleneck. However, the possibility exists that this could be an artefact of considering only this study, thus a literature search was performed which revealed a further eleven papers which considered variation in genetic markers across populations. The following studies also suggested low genetic variablity in the European badger. Domingo-Roura et al (2003) examined twelve microsatellites across a badger population in Wytham Woods, Oxfordshire, and found that only five showed variability. However targets regions were short (94bp to 359bp) and the sample group small; one primer was tested across only 13 badgers, three across 16 badgers, and the remaining eight across 40 badgers. Badgers were potentially related individuals from a 6 km^2 area. Annavi et al (2014) performed a temporal study of badgers sampled from 1987

to 2002 across the same 6 km^2 area in Wytham Woods, genotyping 1170 individuals across 35 microsatellite loci, and found an average of only 4.46 alleles per locus. Huck et al (2008) studied twenty microsatellite loci across seventy-four badgers found at a sample site of less than 2km^2 in Brighton, UK, and found a mean allelic richness value of 3·4 indicating low genetic variability. Dawnay et al (2008) sequenced ten Short Tandem Repeats from 1083 badgers from twenty populations across England and Wales and found very low allelic diversity even in this large sample range.

It is possible that postulations on genetic variability in badgers may in fact be dependent on the choice of sampled genes. The same seven microsatellites were used across three studies, all of which distinguished siblings with >99% certainty, both in British and continental populations. Scheppers et al (2007) used the seven loci to compare genetic methods of estimating population size via hair-trap sampling to observation. The sample population lived across 13km^2 of adjoining sites between Eppeldorf and Medernach in Luxembourg. While sample population was small and geographically limited, the particular loci tested proved variant enough to identify more individuals (n=55) than was estimated by eye (n=49). In a similar study, Frantz et al (2004) expanded on an observed sample of 13 badgers to show that 15 individuals lived in an area of 5·4 km^2 between Ermsdorf and Eppeldorf in Luxembourg. The same was true across a British badger population: Frantz et al 2003 investigated badger faecal samples from Woodchester Park, the core sample area used in this study, using the same seven microsatellite loci and was able to distinguish between 33 individual badgers using this method.

Higher genetic variability appears to exist within badgers on the continent vs. Britain (Frantz et al 2004; Scheppers et al 2007), and also in Asia vs. Europe (Tashima et al 2011). Indeed, Pope et al (2002) sampled allelic variation across twenty-two microsatellites in 354 badgers ranging across England, Wales and Scotland were compared to 108 badgers across nine countries in continental Europe. Mean number of alleles per locus per individual was found to be much lower in the UK (0.14) compared to the continent (0.30). O'Meara et al (2012) studied six microsatellite loci across badgers in Holland (N = 29), Ireland (N = 28), and Spain (N = 16), and found that genetic variation on a loci by loci basis was not correlated with geographical but dependent on the loci being tested. Using the microsatellite Mel116 a high

diversity of 7.6 alleles per locus and per sample was found in Northern England, whereas other loci suggested low diversity across the entire range. However, overall loci the Irish samples showed lower diversity while continental samples showed greater differntiation. Tashima et al (2011) studied mitochondrial DNA across 112 badgers from a wide geographical range spanning the European continent to the Japanese islands. They traced a common ancestor of Eastern and western lineages to the Volga River in Russia, and showed that Japanese Meles meles have greater diversity, which was attributed to Japan's island structure. Badgers based on the European continent were found to have twelve mtDNA haplotypes, while Japanese badgers had twenty-seven. A possible answer to low variability in British badgers may be to import their more genetically diverse counterparts on the continent to promote extra-group breeding which may in turn have positive effect on genetic diversity and thus immuno-adaptive capability in British badgers.

On the other hand a more recent study by Frantz et al (2010) compared sixteen microsatellite loci across 180 badgers over 500km^2 in Glousestershire and 149 badgers over 600km^2 in the Broye region of Switzerland, and found the genetic diversity across both populations not significantly different: Mean number of alleles per locus were 5.6 and 6.2 respectively. These results taken alone are contradictory and inconclusive. Genetic variability on the continent may also have regional variation. Further work should look at wider sample ranges across Britain and continental Europe, a wider variety genes, and larger sample groups.

5.0 References

Abrantes J., Areal H. and Esteves P. J. (2013) Insights into the European rabbit (*Oryctolagus cuniculus*) innate immune system: genetic diversity of the toll-like receptor 3 (TLR3) in wild populations and domestic breeds. BMC Genetics 14:73.

Akira S. and Takeda K. (2004) Toll-Like Receptor Signalling. Immunology 4 499-513.

Al-Orainey I. O. (2009). Diagnosis of latent tuberculosis: Can we do better? Annals of Thoracic Medicine. 4:1 5-9.

Alagarasu K., Bachal R. V., Memane R. S., Shah P. S., Cecilia D. (2015) Polymorphisms in RNA sensing toll like receptor genes and its association with clinical outcomes of dengue virus infection. Immunobiology 220 164-168.

Albright J. W. and Albright J. F. (1981) Differences in Resistance to *Trypanosoma musculi* Infection Among Strains of Inbred Mice. Infection and Immunity 33:2 364-371.

Alexander J. and Russell D. G. (1992) The interaction of Leishmania species with macrophages. Advanced Parasitology 31: 175.

Allen A. R., Minozzi G., Glass E. J., Skuce R. A., McDowell S. W. J., Woolliams J. A. and Bishop S. C. (2010) Bovine tuberculosis: the genetic basis of host susceptibility. Proceedings of the Royal Society B 10.1098.

An H., Xu H., Yu Y., Zhang M., Qi R., Yan X., Liu S., Wang W., Guo Z., Qin Z., Cao X. (2002) Up-regulation of TLR9 gene expression by LPS in mouse macrophages via activation of NF-kB, ERK and p38 MAPK signal pathways. Immunology Letters 81 165-169.

Annavi G., Newman C., Dugdale H. L., Buesching C. D., Sin Y. W., Burke T. and Macdonald D. W. (2014) Neighbouring-group composition and within-group relatedness drive extra-group paternity rate in the European badger (*Meles meles*). Journal of Evolutionary Biology.

Andrews N. W. (1993) Living dangerously: how *Trypanosoma cruzi* uses lysosomes to get inside host cells, and then escapes into the cytoplasm. Biology Research 26 65-67.

Anwar A., Knaggs J., Service K. M., McLaren G. W., Riordan P., Newman C., Delahay R. J., Cheeseman C., Macdonald D. W. (2006) Antibodies to *Toxoplasma gondii* in Eurasian Badgers. Journal of Wildlife Diseases 42:1 179-181.

Apinjoh T. O., Anchang-Kimbi J. K., Njua-Yafi C., Mugri R. N., Ngwai A. N., Rockett K. A., Mbunwe E., Besingi R. N., Clark T. G., Kwiatkowski D. P. and Achidi E. A. (2013) Association of Cytokine and Toll-Like Receptor Gene Polymorphisms with Severe Malaria in Three Regions of Cameroon. PLoS ONE 8(11): e81071.

Areal H., Abrantes J. and Esteves P. J. (2011) Signatures of positive selection in Toll-like receptor (TLR) genes in mammals. BMC Evolutionary Biology 11:368 1471-2148.

Arnason U., Adegoke J. A., Bodin K., Borne E. W., Esa Y. B., Gullberg A., Nilsson M., Short R. V., Xu X., and Janke A. (2002) Mammalian mitogenomic relationships and the root of the eutherian tree. Proceedings of the National Academy of Sciences of the United States of America 99:12 8151-8156.

Austen J. M., Jefferies R., Friend J.A., Ryan U., Adams P. and Reid S. A. (2009) Morphological and molecular characterization of *Trypanosoma copemani* isolated from Gilbert's potoroo (*Potorous gilbertii*) and quokka (*Setonix brachyurus*). Parasitology 136 783-792.

Bagheri N., Azadegan-Dehkordi F., Sanei H., Taghikhani A., Rahimiand G., Salimzadeh L., Hashemzadeh-Chaleshtori M., Rafieian-kopaei M., Shirzad M. and Shirzad H. (2014) Associations of a TLR4 single-nucleotidepolymorphism with *H. pylori* associatedgastric diseases in iranian patients. Research in Hepatology and Gastroenterology 10:1016.

Balseiro A., Rodríguez O., González-Quirós P., Merediz I., Sevilla I. A., Davé D.,Dalley D. J., Lesellier S., Chambers M. A., Bezos J., Muñoz M., Delahay R. J., Gortázar C. and Prieto J. M. (2011) Infection of Eurasian badgers (*Meles meles*) with *Mycobacterium bovis* and *Mycobacterium avium* complex in Spain. The Veterinary Journal 6 pages.

Barkham P. (2013) Badgerlands: the twilight world of Britain's most enigmatic animal. Granta publications. 389 pages.

Barreiro L. B. Marioni J. C., Blekhman R., Stephens M. and Gilad Y. (2010): Universal TLR response is more conserved than the immune responses to specific bacterial or viral infections. PLOS Genetics. 10.1371/journal.pgen.1001249.g002.

Barreiro L. B., Ben-Ali M., Quach H., Laval G., Patin E., Pickrell J. K., Bouchier C., Tichit M., Neyrolles O., Gicquel B., Kidd J.R., Kidd K. K., Alcaïs A., Ragimbeau J., Pellegrini S., Abel S., Casanova J-L. and Quintana-Murci L. (2009) Evolutionary Dynamics of Human Toll-Like Receptors and Their Different Contributions to Host Defense. PLoS Genet 5(7): e1000562.

Bartl S., Baish M., Weissman I. L. and Diaz M. (2003) Did the Molecules of Adaptive Immunity Evolve from the Innate Immune System? Intergrated Compedium of Biology. 43 338-346.

Bartolomei M. S. and Ferguson-Smith A. C. (2011) Mammalian Genomic Imprinting. Cold Spring Harbor Perspectives in Biology 3: A 002592 last accessed online 10/08/2013.

Bartolomei M. S. (2009) Genomic imprinting: employing and avoiding epigenetic processes. Genes and Development 23 2124–2133.

Bates P. A. (2007) Transmission of Leishmania metacyclic promastigotes by phlebotomine sandflies. Internationa Journal for Parasitology 37 10-3 1097-1106.

Beaucournu J.C., Degeilh B.and Guiguen C. (2005) The flea parasites of birds: taxonomic diversity and dispersion. Parasite 12 111-121.

Belkaid Y., Valenzuela J. G. , Kamhawi S., Rowton E., Sacks D. L., Ribeiro J. M. (2000) Delayed-type hypersensitivity to Phlebotomus papatasi sandfly bite: an adaptive response induced by the fly? Procedures of the National Academy of Science USA 97 6704–6709.

Benecke N. (1999) The Holocene History of the European Vertebrate Fauna. Modern Aspects of Research: Verlag Marie Leidorf GmbH. Pages 355-371.

Bengis R. G. and Erasmus J. M. (1988) Wildlife diseases in South Africa: a review. Rev. sci. tech. Off. International Epiz. 7:4 807-821.

Berman J. D. (1997) Human Leishmaniasis: Clinical, Diagnostic, and Chemotherapeutic Developments in the Last 10 Years. Clinical Infectious Diseases. 24 684-703.

Biek R, O'Hare A, Wright D, Mallon T, McCormick C, Orton R. J., McDowell S., Trewby H., Skuce R. A. and Kao R. R. (2012) Whole Genome Sequencing Reveals Local Transmission Patterns of Mycobacterium bovis in Sympatric Cattle and Badger Populations. PLoS Pathog 8(11): e1003008.

Bielby J., Donnelly C. A., Pope L. C., Burke T, and Woodroffe R. (2014) Badger responses to small-scale culling may compromise targeted control of bovine tuberculosis. PNAS 111: 25 9193-9198.

Bitew M., Amedie Y. and Abebe A. (2010) Prevalence of Bovine Trypanosomosis in Selected Areas of Jabi Tehenan District, West Gojjam Of Amhara Regional State, Northwestern Ethiopia. Global Veterinaria 5:5 243-247.

Blackshields G., Sievers F., Shi W., Wilm A., Higgins D. G. (2010) Sequence embedding for fast construction of guide trees for multiple sequence alignment. Algorithms for Molecular Biology 5:21

Bossdorf O.,Richards C. L. and Pigliucc M. (2008) Epigenetics for ecologists. Ecology Letters 11 106-115.

Botero A., Thompson C. K., Peacock C. S., Clode P. L., Nicholls P. K., Wayne A. F., Lymbery A. J. and Thompson R. C. A. (2013) Trypanosome genetic diversity, polyparasitism and the

population decline of the critically endangered Australian marsupial, the brush tailed bettong or woylie (*Bettongia penicillata*). International Journal for Parasitology 2 77-89.

Bouckaert R., Heled J., Kühnert D., Vaughan P., Wu C-H., Xie D., Suchard M. A., Rambaut A. and Drummond A. J. (2014) BEAST 2: A Software Platform for Bayesian Evolutionary Analysis. PLoS Comput Biol 10, e1003537.

Boutlis C. S., Yeo T. W. and Anstey N. M. (2006) Malaria tolerance – for whom the cell tolls? Trends in Parasitology 22:8 371-377.

Bowie A. and O'Neill L. A. J. (2000) The interleukin-1 receptor/Toll-like receptor superfamily: signal generators for pro-inflammatory interleukins and microbial products. Journal of Leukocyte Biology 67 508-515.

Brigand F. et al (2006) Energy metabolism of trypanosomatids: adaptation to available carbon sources. Molecular and Biochemical Parasitology. 149 1-9

Byrne A. W., Sleeman D. P., O'Keeffe J. And Davenport J. (2012) The ecology of the European badger (*Meles meles*) in Ireland: a review. Biology and environment: proceedings of the Royal Irish Academy. 112b:1 105-132.

Campino L., Santos-Gomes G. M., Pratlong F., Antunes F., Mauricio I., Dedet J. P. and Abranches P. (1997) HIV/Leishmania co-infections in Portugal: diagnosis and isoenzyme characterization of Leishmania. Annals of Tropical Medicine and Parasitology 91 433-436.

Campos M. A. S., Almeida I. C., Takeuchi O, Akira S., Valente E. P., Procópio D. O., Travassos L. R., Smith J. A., Golenbock D. T. and Gazzinelli R. T. (2001) Activation of toll-like receptor-2 by glycosylphosphatidylinositol anchors from a protozoan parasite. Journal of Immunology 167:1 416-423.

Cargill E. J. and Womack J. E. (2007) Detection of polymorphisms in bovine toll-like receptors 3, 7, 8, and 9. Genomics. 89:745-755.

Carrera-Silva E. A., Cano R. C., Guinazu N., Aoki M. P., Pellegrini A. and Gea S. (2008) TLR2, TLR4 and TLR9 are differentially modulated in liver lethally injured from BALB/c and C57BL/6 mice during *Trypanosoma cruzi* acute infection. Molecular Immunology 45 3580-3588.

Carter S. P., Delahay R. J., Smith G. C., Macdonald D. W., Riordan P., Etherington, T. R., Pimley E. R., Walker N. J. and Cheeseman C. L. (2007). Culling-induced social perturbation in Eurasian badgers *Meles meles* and the management of TB in cattle: an analysis of a critical problem in applied ecology. Proceedings of the Royal Society B: Biological Sciences, 274:1626 2769–2777.

Carter S. P., Chambers M. A., Rushton S. P., Shirley M. D. F., Schuchert P., Pietravalle S., Murray A., Rogers F., Gettinby G., Smith G. C., Delahay R. J., Hewinson R. G. and McDonald R. A. (2012) BCG Vaccination Reduces Risk of Tuberculosis Infection in Vaccinated Badgers and Unvaccinated Badger Cubs. PLoS ONE 7(12): e49833.

Carvalhoa A., Marques A., Maciela P. and Rodrigues F. (2007) Study of disease-relevant polymorphisms in the TLR4 and TLR9 genes: A novel method applied to the analysis of the Portuguese population. Molecular and Cellular Probes 21 316-320.

Castaño-Rodríguez N., Kaakoush N. O., Pardo A. L., Goh K., Fock K. M., Mitchell H. M. (2014) Genetic polymorphisms in the Toll-like receptor signalling pathway in *Helicobacter pylori* infection and related gastric cancer. Human Immunology 75 808-815.

Castellucci L., Cheng L. H., Araujo C. (2005) Familial aggregation of mucosal leishmaniasis in northeast Brazil. American Journal of Tropical Medicine and Hygeine 73 69-73.

Cestari I., Evans-Osses I., Schlapbach L. J., Messias-Reason L. and Ramirez M. I. (2013) Mechanisms of complement lectin pathway activation and resistance by trypanosomatid parasites. Molecular Immunology 53 328-334.

Charbonnel N., Pagès M., Sironen T., Henttonen H., Vapalahti O., Mustonen J. and Vaheri A. (2014) Immunogenetic Factors Affecting Susceptibility of Humans and Rodents to Hantaviruses and the Clinical Course of Hantaviral Disease in Humans. Viruses 6 2214-2241

Chen G., Severo M. S., Sohail M., Sakhon O. S., Wikel S. K., Kotsyfakis M. and Pedra J. H. F. (2012) *Ixodes scapularis* saliva mitigates inflammatory cytokine secretion during *Anaplasma phagocytophilum* stimulation of immune cells. Parasites and Vectors 5:229 10 pages.

Chen G. and Pedra J. H. F. (2010) The Inflammasome in Host Defense. Sensors 10 97-112.

Chou C., Tegenfeldt J. O., Bakajin O., Chan S. S., Cox E. C., Darnton N., Duke T. and Austin H. (2002) Electrodeless Dielectrophoresis of Single- and Double-Stranded DNA. Biophysical Journal 83 2170-2179.

Christmas P. (2010) Toll-Like Receptors: Sensors that Detect Infection. Nature Education 3:9 85.

Clayton J. (2010) Chagas disease 101. *Nature* 465 4-5.

Clifton-Hadley R. S., Wilesmith J. W. and Stuart F. A. (1993) *Mycobacterium bovis* in the European badger (*Meles meles*) Epidemiological findings in tuberculous badgers from a naturally infected population. Epidemiology of Infectious Diseases 111 9-19.

Covarrubias C., Cortez M., D. Ferreira and Yoshida N. (2007) Interaction with host factors exacerbates *Trypanosoma cruzi* cell invasion capacity upon oral infection. International Journal for Parasitology. 37: 14 1609-1616.

Cox A. P., Tosas O., Tilley A., Picozzi K., Coleman P., Hide G. and Welburn S. C. (2010) Constraints to estimating the prevalence of trypanosome infections in East African zebu cattle. Parasites and Vectors 3:82 1-8.

Cox A. P., Tilley A, McOdimba F., Fyfe J., Hide G. and Welburn S. C. (2005) A PCR based assay for detection and differentiation of African trypanosome species in blood. Experimental Parasitology 111:2 4-9.

Cox R., Stewart P. D., Macdonald D. W. (1999) The Ectoparasites of the European Badger, *Meles meles*, and the Behaviour of the Host-Specific Flea, *Paraceras melis*. Journal of Insect Behaviour 12:2.

Croan D. G., Morrison D. A., Ellis J. T. (1997) Evolution of the genus Leishmania revealed by comparison of DNA and RNA polymerase gene sequences. Molecular and Biochemical Parasitology 89 149-159.

Cuervo P., Domont G. B. and De Jesus J. B. (2010) Proteomics of trypanosomatids of human medical importance. Journal of Proteomics 73 845-867.

Cuscó A., Sánchez A., Altet L., Ferrer L. and Francino O. (2014) Non-synonymous genetic variation in exonic regions of canine Toll-like receptors. Canine Genetics and Epidemiology 1:11.

Davies C. R., Reithinger R., Campbell-Lendrum D., Feliciangeli D., Borges R. and Rodriguez N. (2000) The epidemiology and control of leishmaniasis in Andean countries. Cad Saúde Publica 16 925–950.

Davies J. M. (1997) Molecular Mimicry: can epitope mimicry induce auto-immune disease? Immunology and Cell Biology 75 113-126

Dawnay N., Ogden R., Thorpe R. S., Pope L. C., Dawson D. A. and McEwing R. (2008) A forensic STR profiling system for the Eurasian badger: A framework for developing profiling systems for wildlife species. Forensic Science International. Genetics 2 47-53.

De Albuquerque C. V. and Linardi P. M. (2006) Scanning electron microscopy studies of sensilla and other structures of the head of *Polygenis tripus*. Micron 37 557- 565.

De Carvalho T. M. U. and De Souza W. (1989) Early events related with the behaviour of *Trypanosoma cruzi* within an endocytic vacuole in mouse peritoneal macrophages. Cell Structure and Function. 14:4 383-392.

De Souza W., Ulisses de Carvalho T. M. and Barrias E. S. (2010) Review on *Trypanosoma cruzi* : Host Cell Interaction. International Journal of Cell Biology

Dean S. R. and Meola R.W. (2002) Effective diet composition on weight gain, sperm transfer, and instrumentation in the cat flea. Journal of medical entomology 39 370-375.

De Greef C. and Hamers R. (1994) The serum resistance-associated (SRA) gene of *Trypanosoma brucei rhodesiense* encodes a variant surface glycoprotein-like protein. Mol Biochem Parasitol. 68:2 277-284.

Department of Environment, Food and Agriculture (2014) Notes for guidance: Importation of Cattle into the Isle of Man from Great Britain (version 17) 3 pages https://www.gov.im/media/193659/guidance-notes-on-importation-of-cattle-v17.pdf

Deschamps P., Lara E., Marande W., Lopez-Garcıa P., Ekelund F. and Moreira D. (2011) Phylogenomic Analysis of Kinetoplastids Supports That Trypanosomatids Arose from within Bodonids. Molecular Biology and Evolution 28:1 53-58.

Desjeux P. (2001) The increase in risk factors for leishmaniasis worldwide. Transripts of the Royal Socety of Tropical Medicine and Hygiene 95 239-243.

Devarajua P., Gulatib R., Antonya P. T., Mithuna C. B. and Negia V. S. (2014) Susceptibility to SLE in South Indian Tamils may be influenced bygenetic selection pressure on TLR2 and TLR9 genes. Molecular Immunology 4524 Author's copy.

Dias J. C. P. (2007) Southern Cone Iniative for the elimination of domestic populations of *Triatoma infestans* and the interruption of transfusional Chagas disease. Historical apects, present situation and perspectives. Mem Inst Oswaldo Cruz 102:1 11-18.

Dieffenbach C. W., Lowe T. M. J. and Dveksler G. W. (1993) General Concepts for PCR Primer Design. PCR Methods and Applications 3:9 530-537.

Domingo-Roura X. ,Macdonald D . W., Roy M. S., Marmi J., Terradas J., Woodroffe R., Burke T and Wayne R . K. (2003) Confirmation of low genetic diversity and multiple breeding females in a social group of Eurasian badgers from microsatellite and field data. Molecular Ecology 12 533-539.

Donnelly C. A and Hone J. (2010) Is there an association between levels of bovine tuberculosis in cattle herds and badgers? Statistical Communications in Infectious Diseases 2:1 Accession Number 20133149921.

Draper S. (2013) An Introduction to the Farming Advice Service. FAS Presentation 47 slides.

Drouin-Ouellet J. and Cicchetti F. (2012) Inflammation and neurodegeneration: the story 'retolled.' Trends in Pharmacological Sciences 33: 10 542-552.

Dudas G. and Rambaut A. (2014) Phylogenetic Analysis of Guinea 2014 Ebolavirus Outbreak. PLoS Currents: Outbreaks 6.

Echodu R., Sistrom M., Bateta R., Murilla G., Okedi L., Aksoy S., Enyioha C., Enyaru J., Opiyo E., Gibson W. and Caccone A. (2015) Genetic Diversity and Population Structure of *Trypanosoma brucei* in Uganda: Implications for the Epidemiology of Sleeping Sickness and Nagana. PLOS 10.1371.

Eiseman C. H. and Binnengton K. C. (1994) The Peritrophic membrane: its formation, structure, chemical composition and permeability in relation to vaccination against ectoparasitic arthropods. International Journal for Parasitology 24 15-26.

Eizirik E., Murphy W.J., Koepfli K., Johnson W. E., Dragoo J. W., Wayne R. K. and O'BrienS. J. (2010) Pattern and timing of diversification of the mammalian order Carnivora inferred from multiple nuclear gene sequences. *Molecular Phylogenetics and Evolution* 56 49–63.

Ekanayake D. K., Minning T., Weatherly B., Gunasekera K., Nilsson D.,Tarleton R., Ochsenreiter T. and Sabatini R. (2011) Epigenetic Regulation of Transcription and Virulence in *Trypanosoma cruzi* by O-Linked Thymine Glucosylation of DNA. Molecular and Cellular Biology 31:8 1690-1700.

El-Omar E. M., Ng M. T. and Hold G. L. (2008) Polymorphisms in Toll-like receptor genes and risk of cancer. Oncogene 27 244-252.

Emingil G., Berdeli A., Baylas H., Saygan B. H., Gurkan A., Kose T. and Atilla G. (2007) Toll-Like Receptor 2 and 4 Gene Polymorphisms in Generalized Aggressive Periodontitis. Journal of Periodontology 78:10 1968-1977.

Epting C. L., Coates B. M. and Engman D. M. (2010) Molecular mechanisms of host cell invasion by *Trypanosoma cruzi*. Experimental Parasitology 126 283-291.

Etherington R., Pimley E. R., Walker N. J., Cheeseman C. L., Carter S. P., Delahay R. J., Smith G. C., Macdonald D. W., Riordan P. (2007) Culling-induced social perturbation in Eurasian

badgers *Meles meles* and the management of TB in cattle: an analysis of a critical problem in applied ecology. Proceedings of the Royal Society B 274

Everett B. (2009) Immunological correlates of illness severity and course in acute Q fever. Thesis submitted for the degree of Doctor of Philosophy, University of NSW. 352 pages.

Fairnell E. H. and Barrett J. H. (2007) Fur-bearing species of the Scottish islands. Journal of Archaeological Science 34 463-484.

Falstad I. and Karter A. J. (1992) Parasites, bright males, and the immunocompetence handicap. American Naturalist 139 603-622.

Ferguson M. A. J. (1999) The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. Journal of Cell Science 112 2799-2809.

Fevre E. M., Coleman P. G., Odiit M., Magona J. W., Welburn S. C. and Woolhouse (2001) the origins of a new *Trypanosoma brucei rhodesiense* sleeping sickness outbreak in eastern Uganda. The Lancet 358 625-628.

Finelle P. (1972) African animal trypanosomiasis. Disease and Chemotherapy 1-26

Frantz A. C., San E. D. L., Pope L. C. and Burke T. (2010) (a) Using genetic methods to investigate dispersal in two badger (*Meles meles*) populations with different ecological characteristics. Heredity 104 49-501.

Frantz A. C., Pope L. C. Etherington T. R., Wilson G. J.and Burke T. (2010) (b) Using isolation-by-distance based approaches to assess the barrier effect of linear landscape elements on badger (*Meles meles*) dispersal. Molecular Ecology 19 1663-1674.

Frantz A. C., Do Linh San E., Pope L. and Burke T. (2010) Using genetic methods to investigate dispersal in two badger (*Meles meles*) populations with different ecological characteristics. Heredity 104 493-501.

Frantz A. C., Schaul M., Pope L. C., Fack F., Schley L., Muller C. P. and Roper T. J. (2004) Estimating population size by genotyping remotely plucked hair: the Eurasian badger. Journal of Applied Ecology 41 985-995.

Frantz A. C., Pope L. C., Carpenter P . J ., Roper T . J ., Wilson G . J ., Delahay R . J . and Burke T. (2003) Reliable microsatellite genotyping of the Eurasian badger (*Meles meles*) using faecal DNA. Molecular Ecology. 12 1649-1661.

Friedberg E. C. and Lawrence E. (2001) Sydney Brenner: My Life in Science, as told to Lewis Wolpert. BioMed Central Ltd, London. 191 pages.

Gallagher J. and Clifton-Hadley R. S. (2000) Tuberculosis in badgers; a review of the disease and its significance for other animals. Research in Vetrinary Science 69:3 203-217.

Gao F. and Zhang C. T. (2006) GC-Profile: a web-based tool for visualizing and analyzing the variation of GC content in genomic sequences. Nucleic Acids Research 34 10.1093.

Gao Z., Rennie D. C. and Senthilselvan A. (2010) Allergic rhinitis and genetic components: focus on Toll-like receptors (TLRs) gene polymorphism. The Application of Clinical Genetics 3 109-120.

Garcia E. S., Castro D. P., Figueiredo M. B. and Azambuja P. (2012) Parasite-mediated interactions within the insect vector: *Trypanosoma rangeli* strategies. Parasites and Vectors 5:105

Golovkin A. S., Ponasenko A. V., Yuzhalin A. E., Salakhov R. R., Khutornaya M. V., Kutikhin A. G., Rutkovskaya N. V., Savostyanova Y. Y. and Barbarash L. S. (2015) An association between single nucleotide polymorphisms within TLR and TREM-1 genes and infective endocarditis. Cytokine 71: 16-21.

Golovkin A. S., Ponasenko A. V., Khutornaya M. V., Kutikhin A. G., Salakhov R. R., Yuzhalin A. E., Zhidkova I. I., Barbarash O. L. and Barbarasha L. S. (2014) Association of TLR and TREM-1 gene polymorphismswith risk of coronary artery disease in a Russian population. Gene 550: 101-109.

Gomaz A., Pavelic J. and Glavan T. M. (2012)The polymorphisms in Toll-like receptor genes and cancer risk. Periodicum Biologorum 114:4 461-469.

Gomez-Diaz E., Jorda M., Peinado M. A. and Rivero A. (2012) Epigenetics of Host–Pathogen Interactions: The Road Ahead and the Road Behind. PLoS Pathogens 8(11): e1003007.

Gong J., Wei T., Zhang N., Jamitzky F., Heckl W. M., Rössle S. C. and Stark R. W. (2010) TollML: a database of toll-like receptor structural motifs. Journal of Molecular Modelling 16:7 1283-1289

Gormley E. and Costello E. (2003) Tuberculosis and badgers: new approaches to diagnosis and control. Journal of Applied Microbiology 94 80-86.

Gortazar C., Torres M. J., Vicente J., Acevedo P., Reglero M., de la Fuente J., Negro J. J. and Aznar-Martın J. (2008) Bovine Tuberculosis in Donana Biosphere Reserve: The Role of Wild Ungulates as Disease Reservoirs in the Last Iberian Lynx Strongholds. PLoS ONE 3(7): e2776.

Gowda D. C. (2007) TLR-mediated cell signaling by malaria GPIs. Trends in Parasitology 23(12) 596- 604.

Grabau J. C. and Novick L. F.(1995) False Positive Tuberculosis Skin Test Results. Public Health Reports. Vol 110 703-706.

Grange J. and Zumla A. (1999) Tuberculosis and the poverty-disease cycle. Journal of the Royal Society of Medicine. 92:3 105-107.

Grech V., Mizzi J. and Mangion M. (2000) Visceral leishmaniasis in Malta; an 18 year paediatric, population based study. Arch Dis Child 82 381-385.

Grueber C. E., Wallis G. P. and Jamieson I. G. (2014) Episodic Positive Selection in the Evolution of Avian Toll-Like Receptor Innate Immunity Genes. PLoS ONE 9(3): e89632.

Grueber C. E., Wallis G. P., King T. M. and Jamieson I. G. (2012) Variation at Innate Immunity Toll-Like Receptor Genes in a Bottlenecked Population of a New Zealand Robin. PLoS ONE 7(9): e45011.

Guhl F. and Vallejo G. A. (2003) *Trypanosoma* (Herpetosoma) *rangeli* Tejera, 1920 - An Updated Review. Mem Inst Oswaldo Cruz 98:4 435-442.

Hamann L., Koch A., Sur S., Hoefer N., Glaeser C., Schulz S., Gross M., Franke A., Nöthlings A., Zacharowski K. and Schumann R. R. (2013) Association of a common TLR-6 polymorphism with coronary artery disease-implications for healthy ageing? Immunity and Ageing 10:43.

Hamilton P. B., Gibson W. C., Stevens J. R. (2007) Patterns of co-evolution between trypanosomes and their hosts deduced from ribosomal RNA and protein-coding gene phylogenies. Molecular Phylogenetics and Evolution 44: 15–25.

Hashemi-Shahri S. M., Taheri M., Gadari A., Naderi M., Bahari G. and Hashemi M. (2014) Association Between TLR8 and TLR9 Gene Polymorphisms and Pulmonary Tuberculosis. Gene Cell Tissue 1(1): e18316.

Hawn T. R., Scholes D., Li S. S., Wang H., Yang Y., Roberts P. L., Stapleton A. E., Janer M., Aderem A., Stamm W. E., Zhao L. P. and Hooton T. M. (2009) Toll-Like Receptor Polymorphisms and Susceptibility to Urinary Tract Infections in Adult Women. PLoS ONE 4(6).

Herwaldt B. L. (1999) Leishmaniasis. Lancet 354 1191-1199.

Hide G. and Tait A. (2009) Molecular epidemiology of African sleeping sickness. Parasitology 136 1491-1500.

Hide G., Morley E. K., Hughes J. M., Gerwash O., Elmahaishi M. S., Elmahaishi K. H., Thomasson D., Wright E. A., Williams R. H., Murphy R. G. and Smith J. E. (2009) Evidence for high levels of vertical transmission in *Toxoplasma gondii*. Parasitology 136: 14 1877-1885

Hide M., Bucheton B., Kamhawi S., Bras-Goncalves R., Sundar S., Lemesre J. L. and Banuls A. L. (2007) Understanding Human Leishmaniasis: The need for an Integrated Approach. Encyclopedia of Infectious Diseases: Modern Methodologies (ed. by M. Tibayrenc) pp. 87. Wiley-Liss, Hoboken, N.J.

Hide G. (2003) The origins of sleeping sickness epidemics. Biologist 50:6 268-272.

Hide G. (1999) History of Sleeping Sickness in East Africa. Clinical Microbiology reviews. 12:1 112-125.

Hide G., Angus S. D., Holmes P. H., Maudlin I. and Welburn S. C. (1998) *Trypanosoma brucei*: Comparison of Circulating Strains in and Endemic and Epidemic Area of a Sleeping Sickness Focus. Experimental Parasitology 89: 21-29.

Hide G. (1997) The molecular epidemiology of the trypanosomatids, in Trypanosomiasis and Leishmaniasis: Biology and Control (Eds: Hide G., Mottram J. C., Coombs G. H. and Holmes P. H). 289-304.

Hide G., Tait A., Maudlin I. and Welburn S. C. (1996) The Origins, dynamics and generation of *Trypanosoma brucei rhodesiense* epidemics in East Africa. Parasitology Today 12: 50-55.

Hide G., Tait A., Maudlin I. and Welburn S. C. (1994) Epidemiological relationships of *Trypanosoma brucei* stocks from South East Uganda: evidence for different population structures in human infective and non-human infective isolates. Parasitology 109 95-111.

Hide G., Buchanen N., Welburn S. C., Maudin I., Barry J. D. and Tait A. (1991) *Trypanosoma b. rhodesiense*: Characterization of stocks from Zambia, Kenya and Uganda using repetitive DNA probes. Experimental parasitology 72 430-439.

Hide G., Cattand P., LeRay D., Barry J. D. and Tait A. (1990) The identification of *T. brucei* subspecies using repetitive DNA sequences. Molecular and Biochemical Parasitology 39 213-226.

Higgins D. A. (1985) The skin inflammatory response of the badger (*Meles meles*) British Journal of Experimental Pathology 66: 643-653.

Ho D. H. and Burggren W. W. (2010) Epigenetics and transgenerational transfer: a physiological perspective. The Journal of Experimental Biology 213 3-16.

Holla L. I., Vokurka J., Hrdlickova B., Augustin P. and Fassmann A. (2010) Association of Toll-like receptor 9 haplotypes with chronic periodontitis in Czech population. Journal of Clinical Periodontology 37: 152-159.

Hou B., Reizis B. and DeFranco A. L. (2008) Toll-like Receptors Activate Innate and Adaptive Immunity by using Dendritic Cell-Intrinsic and -Extrinsic Mechanisms. Immunity 29 272-282.

Hounsome T. D., Young R. P., Davison J., Yarnell R. W., Trewby I. D., Garnett B. T., Delahay R. J. and Wilson G. J. (2005) An evaluation of distance sampling to estimate badger (*Meles meles*) abundance. Journal of Zoology 266 1-7

Huard R. C., Fabre M., de Haas P., Lazzarini L. C. O., van Soolingen D., Cousins D. and Ho J. L. (2006) Novel Genetic Polymorphisms That Further Delineate the Phylogeny of the *Mycobacterium tuberculosis* Complex. Journal of Bacteriology 188:12 4271–4287.

Huck M., Frantz A. C., Dawson D. A., Burke T. and Roper T. J. (2008) Low genetic variability, female-biased dispersal and high movement rates in an urban population of Eurasian badgers *Meles meles*. Journal of Animal Ecology e:01415.

Hybu Cig Cymru (2013) A review of UK and Welsh red meat production during the first quarter of 2013. HCC Market Survey. 8 pages.

Ideozu E. J., Whiteoak A. M., Tomlinson A. J., Robertson A., Delahay R. J. and Hide G. (unpublished, 2015) Detection of trypanosomes in European badgers using ITS-PCR.

Ioana M., Ferwerda B., Plantinga T. S., Stappers M. Oosting M., McCall M., Cimpoeru A., Burada F., Panduru N., Sauerwein R., Doumbo O., van der Meer J. W. M., van Crevel R., Joosten L. A. B. and Netea M. G. (2012) Different Patterns of Toll-Like Receptor 2 Polymorphisms in Populations of Various Ethnic and Geographic Origins. Infection and Immunity 80:5 1917.

Iwamoto T., Sonobe T. and Hayashi K. (2003) Loop-Mediated Isothermal Amplification for Direct Detection of *Mycobacterium tuberculosis* Complex*, M. avium, and M. intracellulare* in Sputum Samples. Journal of Clinical Microbiology 41:6 2616-2622.

Jabłonska A., Paradowska E., Studzinska M., Suski P., Nowakowska D., Wisniewska-Ligier M., Wozniakowska-Gesicka T., Wilczynski J. and Lesnikowski Z. J. (2014) Relationship between toll-like receptor 2 Arg677Trp and Arg753Gln and toll-like receptor 4 Asp299Gly polymorphisms and cytomegalovirus infection. International Journal of Infectious Diseases e11-e15.

Janardhanan J., Martin S. J., Astrup E., Veeramanikandan R., Aukrust P., Ooriapadickal C Abraham1 and George M Varghese1 (2013) Single-nucleotide polymorphisms in Toll-like receptor (TLR)-2, TLR4 and heat shock protein 70 genes and susceptibility to scrub typhus. Journal of Human Genetics 58: 707-710.

Janssens S. and Beyaert R. (2003) Role of Toll-Like Receptors in Pathogen Recognition. Clinical Microbiology Reviews 16:4 637-646.

Jaric M., Segal J., Silva-Herzog E., Schneper L., Mathee K. and Narasimhan G. (2013) Better primer design for metagenomics applications by increasing taxonomic distinguishability. BMC Proceedings 7:S4 10 pages.

Jarvinen J. A. and Dalmasso A. P. (1977) *Trypanosoma musculi* Infections in Normocomplementemic, C5-Deficient, and C3-Depleted Mice. Infection and Immunity 16:2 557-563.

Johnson D. D. P, Macdonald D. W. and Morecroft M. D. (2001) An analysis of social spacing in the European badger (*Meles meles*) in the UK. Wissenschaftliche Mitteilungen aus dem Niederösterreichischen Landesmuseum 14: 161-170.

Kaiser R., Tang L. F., Taylor K. E., Sterba K., Nititham J., Brown E. E., Edberg J. C., McGwin G., Alarcon G. C., Ramsey-Goldman R., Reveille J. D., Vila L. M., Petri M., Rauch J., Miller E., Mesznik K., Kwok P., Kimberly R. P., Salmon J. E. and Criswell L. A. (2014) A Polymorphism in TLR2 Is Associated With Arterial Thrombosis in a Multiethnic Population of Patients With Systemic Lupus Erythematosus. Arthritis and Rheumatology 66:7 1882–1887.

Kamhawi S. (2006) Phlebotomine sand flies and Leishmania parasites: friends or foes? Trends in Parasitology 22 439-445.

Karle R., Holder F. and Smith M. (2003) Application of FTA-based Technology for Sample Collection, Transport, Purification, and Storage of PCR-ready Plant DNA. Last accessed online 09/08/2013.

Karolemeas K., Donnelly C. A., Conlan A. J. K., Mitchell A. P., Clifton-Hadley R. S., Upton P., Wood J. L. N. and McKinley T. J. (2012) The Effect of Badger Culling on Breakdown Prolongation and Recurrence of Bovine Tuberculosis in Cattle Herds in Great Britain. PLoS ONE 7(12) e51342.

Kathrani A., House A., Catchpole B., Murphy A., Werling D. and Allenspach K. (2011) Breed-independent toll-like receptor 5 polymorphisms show association with canine inflammatory bowel disease. Tissue Antigens 78:94-101.

Kayama H. and Takeda K. (2010) The innate immune response to *T. cruzi* infection. Microbes and Infection 12 511-517.

Kennedy P. G. E. (2005) Sleeping Sickness: Human African Trypanosomiasis. Practical Neurology Review 5 260-267

Kerr S. F. (2006) Molecular trees of trypanosomes incongruent with fossil records of hosts. Memorias de Instituto Oswaldo Cruz 101 25-30.

Kim E. J., Chung W. C., Lee K., Paik C. N., Jung S. H., Lee B. I., Chae H. S. and Choi K. Y. (2011) Association Between Toll-Like Receptors/CD14 Gene Polymorphisms and Inflammatory Bowel Disease in Korean Population. Gastroenterology and Hepatology 3346.

Khan A., Aslam M. A., Hussain I., Naz A. G., Rana I. A., Ahmad M. M., Ali M. and Ahmad S. (2013) Role of Toll-like receptor 2 polymorphism in susceptibility to pulmonary tuberculosis in Pakistani population. International Journal of Immunogenetics 0: 1-7.

Khan I. A. (2007) Toll road for *Toxoplasma gondii*: the mystery continues. Trends in Parasitology 23(1) 1-3.

Khandelwal G. and Bhyravabhotla J. (2010) A Phenomenological Model for Predicting Melting Temperatures of DNA Sequences. PLoS ONE 5:8 e12433.

Kidner P. (2014) A Place of Safety: the highs and lows of working with wildlife. PK Publishing, Secret World Wildlife Rescue. 207 pages

Kierszenbaum F. (1999) Chagas' Disease and the Autoimmunity Hypothesis. Clinical Microbiology Reviews 12:2 210-223.

Kim J. K., Samaranayake M. and Pradhan S. (2009) Epigenetic mechanisms in mammals. Cellular and Molecular Life Sciences 66 596-612.

Kirk W. D. J. (1991) The size relationship between insects and their hosts. Ecological entomology 16 351-359.

Koepfli K. P., Deere K. A., Slater G. J., Begg C., Begg K., Grassman L., Lucherini M. and Veron G. (2008) Multigene phylogeny of the Mustelidae: Resolving relationships, tempo and biogeographic history of a mammalian adaptive radiation. BMC Biology 6:10 1741-1747.

Kong Q., Lu S., Tong Q., Lou D., Chen R., Zheng B., Kumagai T., Wen L., Ohta N. and Zhou X. (2012) Loop-mediated isothermal amplification (LAMP): Early detection of *Toxoplasma gondii* infection in Mice. Parasites & Vectors 5:2.

Kowalczyk R., Zalewski A. and Jêdrzejewska B. A. (2004) Seasonal and spatial pattern of shelter use by badgers *Meles meles* in Biaowieja Primeval Forest (Poland). Acta Theriologica 49:1 75-92.

Krasnov B. R. (2008) Functional and evolutionary ecology of fleas: a model for ecological parasitology. Cambridge University press, 593 pages.

Krieg A. M. and Vollmer J. (2008) Toll-like receptors 7, 8, and 9: linking innate immunity to autoimmunity. Immunological Reviews. 220:1 251-269.

Krieg A. M. (2002) From A to Z on CpG. Trends in Immunology 23:2 64-65.

Ku C., von Bernuth H., Picard C., Zhang S., Chang H., Yang K.,Chrabieh M., Issekutz A. C., Cunningham C. K., Gallin J., Holland S. M., Roifman C., Ehl S., Smart J., Tang M., Barrat F. J., Levy O., McDonald D., Day-Good N. K., Miller R., Takada H., Hara T., Al-Hajjar S., Al-Ghonaium A., Speert D., Sanlaville D., Li X., Geissmann F., Vivier E., Maródi L., Garty B., Chape H.l, Rodriguez-Gallego C., Bossuyt X., Abel L., Puel A. and Casanova J. (2007) Selective predisposition to bacterial infections in IRAK-4–deficient children: IRAK-4–dependent TLRs are otherwise redundant in protective immunity. The Rockefeller University Press 204: 10 2407-2422.

Kubarenko A. V., Ranjan S., Colak E., George J., Frank M.,2 and Weber A. N. R. (2010) Comprehensive modeling and functional analysis of Toll-like receptor ligand-recognition domains. Protein Science 19 558-569.

Kuboki N., Inoue N., Sakurai T., Di Cello F., Grab D. J., Suzuki C. S. and Igarashi I. (2003) Loop-Mediated Isothermal Amplification for Detection of African Trypanosomes. Journal of Clinical Microbiology 41:12 5517-5524.

Kulchavenya E., Kim C. S., Bulanova O. and Zhukova I. (2012) Male genital tuberculosis: epidemiology and diagnostic. World Journal of Urology 30: 15.

Lane R. P. (1993) Sandflies (Phlebotominae) In: Lane R. P., Crosskey R. W., editors. Medical Insects and Arachnids. Chapman and Hall, London. 78–119.

Langefeld T., Mohamed W., Ghai R. and Chakraborty T. (2009) Toll-Like Receptors and NOD-Like Receptors: Domain Architecture and Cellular Signalling. Chapter 9: Target Pattern Recognition in Innate Immunity, edited by Uday Kishore. Landes Bioscience and Springer.

Lesellier S., Palmer S., Dalley D.J., Dave D., Johnson L., Hewinson R. G. and Chambers M. A. (2006) The safety and immunogenicity of Bacillus Calmette-Guerin (BCG) vaccine in European badgers (*Meles meles*). Veterinary Immunology and Immunopathology 112 24-37.

Le Roex N., van Helden P. D., Koets A. P. and Hoal E. G. (2013) Bovine TB in livestock and wildlife: what's in the genes? Physiological Genomics 45: 631-637.

Liadaki K., Petinaki E., Skoulakis C., Tsirevelou P., Klapsa D., Germenis A. E. and Speletas M.

(2011) Toll-Like Receptor 4 Gene (TLR4), but Not TLR2, Polymorphisms Modify the Risk of Tonsillar Disease Due to Streptococcus pyogenes and Haemophilus influenzae. Clinical And Vaccine Immunology, Feb. 2011 217–222

Liao W., Chen R., Lin H., Liu Y., Chen W., Tsai Y., Wan L. and Tsai F. (2010) Toll-like receptor gene polymorphisms are associated with susceptibility to Graves' ophthalmopathy in Taiwan males. BMC Medical Genetics 11:154

Liu F., Lu W., Qian Q., Qi W., Hu J. and Feng B. (2012) Frequency of TLR 2, 4, and 9 Gene Polymorphisms in Chinese Population and Their Susceptibility to Type 2 Diabetes and Coronary Artery Disease. Journal of Biomedicine and Biotechnology. Article 373945, 7 pages.

Lizundia R., Newman C., Buesching C. D., Ngugi D., Blake D., Sin Y. W., Macdonald D. W., Wilson A. and McKeever D. (2011) Evidence for a Role of the Host-Specific Flea (*Paraceras melis*) in the Transmission of *Trypanosoma pestanai* to the European Badger. PLoS ONE 6(2): e16977.

Llywodraeth Cymru (2014) Bovine TB Eradication Programme IAA Badger Vaccination Project. Year 2 Report. Digital ISBN 978 1 4734 1222 4. 14 pages.

Lu Y., Yeh W. and Ohashi P. S. (2008) LPS/TLR4 signal transduction pathway. Cytokine 42 145-151.

Lvovsky L., Ioshikhes I., Raja M. C., Zevin-Sonkin D., Sobolev I. A., Liberzon A., Shwartzburd J. and Ulanovsky L. E. (1998) Interdependence between DNA template secondary structure and priming efficiencies of short primers. Nucleic Acids Research 26:23 5525-5532.

Lynch J. M. (1996) Postglacial colonization of Ireland by mustelids, with particular reference to the badger (*Meles meles* L). Journal of Biogeography 23 179-185.

Lyons R. (2014) http://seqcore.brcf.med.umich.edu/doc/dnaseq/interpret.html accessed 11/01/2014

MacLean L. M., Odiit M., Chisi J. E., Kennedy P. G. E., Sternberg J. M. (2010) Focus–Specific Clinical Profiles in Human African Trypanosomiasis Caused by*Trypanosoma brucei rhodesiense*. PLoS Negl Trop Dis 4(12): e906.

Macdonald D. W., Newman C., Buesching C. D. and Johnson P. J. (2008) Male-biased movement in a high-density population of the Eurasian badger (*Meles meles*). Journal of Mammalogy 89:5 1077-1086.

MacMorris-Adix M. (2008) Leishmaniasis: A review of the disease and the debate over the origin and dispersal of the causative parasite Leishmania. Macalester Reviews in Biogeography 1:2 19 pages.

Mahmood K. H., Rook G. A. W., Stanford J. L., Stuart F. A. and Pritchard D. G. (1987) The immunological consequences of challenge with bovine tubercle bacilli in badgers (*Meles meles*). Epidemiology and Infection 98 155-163.

Marsden P. D. (1986) Mucosal leishmaniasis. Transripts of the Royal Socety of Tropical Medicine and Hygiene 80 859-876.

Marshall O. J. (2004) PerlPrimer: cross-platform, graphical primer design for standard, bisulphite and real-time PCR. Bioinformatics Applications Note 20: 15 2471-2472.

Martínez-Ríos M. A., Vargas-Alarcón G., Vallejo M., Cruz-Martinez E., Pérez-Méndez O., Medina-Andrade Á., De la Torre-García M., Pena-Duquea M. A. and Fragoso J. M. (2013) Toll-like receptor 4 gene polymorphisms and acute coronary syndrome: No association in a Mexican population. Archivos de Cardiologia de Mexico. 83:4 257-262.

Martinon F., Mayor A. and Tschopp J. (2009) The Inflammasomes: Guardians of the Body. The Annual Review of Immunology 27 229-265.

Matas-Cobos A. M., Redondo-Cerezo E., Alegría-Motte C., Martínez-Chamorro A., Saenz-López P., Jiménez P., Jiménez M. R., González-Calvín J. L., de Teresa J. and Osuna F. R. (2014) The Role of Toll-Like Receptor Polymorphisms in Acute Pancreatitis Occurrence and Severity. Pancreas Journal; author's pre-release copy.

Maudlin I., Welburn S. C., Gashumba J. K., Okuns N. and Kalunds M. (1990) The role of cattle in the epidemiology of sleeping sickness in Uganda. Bulletin de la Societe Francaise de Parasitologie 8:2 788.

McCall L., Zhang W. and Matlashewski G. (2013) Determinants for the Development of Visceral Leishmaniasis Disease. PLoS Pathog 9(1): e1003053.

McCarthy G., Shiel R., O'Rourke L., Murphy D., Corner L., Costello E. and Gormley E. (2009) Bronchoalveolar lavage cytology from captive badgers. American Society for Veterinary Clinical Pathology 38:3 381-387.

Meddeb-Garnaoui A., Zrelli H. and Dellagi K. (2009) Effects of tropism and virulence of Leishmania parasites on cytokine production by infected human monocytes. Clinical and Experimental Immunology 155 199-206.

Medvedev S. G. (2004) Morphological adaptations of fleas (Siphonapetra) to parasitism. Entomological review 83 313-333.

Medzhitov R. (2009) Approaching the Asymptote: 20 Years Later. Immunity 30 766-775

Michelsen V. (1997) A revised interpretation of the mouthparts in adult fleas (Insecta, Siphonaptera). Zoologischer Anzeiger 35 217-223.

Minning T. A., Weatherly D. B., Atwood J., Orlando R., and Tarleton R. L. (2009) The steady-state transcriptome of the four major life-cycle stages of *Trypanosoma cruzi*. BMC Genomics 10:370

Misch E. A. and Hawn T. R. (2008) Toll-like receptor polymorphisms and susceptibility to human disease. Clinical Science 114 347-360.

Mockenhaupt F. P., Cramer J. P., Hamann L., Stegemann M. S., Eckert J., Oh N., Otchwemah R. N., Dietz E., Ehrhardt S., Schroder N. W. J., Bienzle U. and Schumann R. (2006) Toll-like receptor (TLR) polymorphisms in African children: Common TLR-4 variants predispose to severe malaria. PNAS 103:1 177-182.

Mohammadi M., Zahedi M. J., Nikpoor A. R., Baneshi M. R. and Hayatbakhsh M. M. (2013) Interleukin-17 Serum Levels and TLR4 Polymorphisms in Ulcerative Colitis. Iranian Journal of Immunology 10:2 83-92.

Moncada D. M., Kammanadiminti S. J. and Chadee K. (2003) Mucin and Toll-like receptors in host defence against intestinal parasites. Trends in Parasitology 19(7) 305- 311.

Moreira D., Lopez-Garcıa P. and Vickerman K. (2014) An updated view of kinetoplastid phylogeny using environmental sequences and a closer outgroup: proposal for a new classification of the class Kinetoplastea. International Journal of Systematic and Evolutionary Microbiology 54, 1861-1875.

Morger J., Bajnok J., Boyce K., Craig P. S., Rogan M. T., Lun Z., Hide G., Tschirren B. (2014) Naturally occurring Toll-like receptor 11 (TLR11) and Toll-like receptor 12 (TLR12) polymorphisms are not associated with *Toxoplasma gondii* infection in wild wood mice. Infection, Genetics and Evolution 26: 180–184

Mukherjee S., Ganguli D. and Majumder P. P. (2014) Global Footprints of Purifying Selection on Toll-Like Receptor Genes Primarily Associated with Response to Bacterial Infections in Humans. Genome Biology and Evolution 6:3 551-558.

Murphy G., Oldbury D., Hughes J., Hide G., Thomasson D., and Williams H. (2008) Role of urban mice in transmission of *Toxoplasma gondii*. Proceedings of the Sixth International Conference on Urban Pests, Robinson W. H. and Bajomi D. (editors) pp. 319-323.

Murphy K., Travers P. and Walport M. (2008) Janeway's immunobiology, 7th edition. New York, NT: Garland Science.

Muzio M., Bosisio D., Polentarutti N., D'amico G., Stoppacciaro A., Mancinelli R., van't Veer C., Penton-Rol G., Ruco L. P., Allavena P. and Mantovani A. (2000) Differential Expression and Regulation of Toll-Like Receptors in Human Leukocytes: Selective Expression of TLR3 in Dendritic Cells. Journal of Immunology 164 5998-6004.

Nachtigall I., Tamarkin A., Tafelski S., Weimann A., Rothbart A., Heim S., Wernecke K. D. and Spies C. (2014) Polymorphisms of the toll-like receptor 2 and 4 genes are associated with faster progression and a more severe course of sepsis in critically ill patients. Journal of International Medical Research 42:1 93-110.

Nadjm B, Van Tulleken C., Macdonald D. and Chiodini P. L. (2009) East African Trypanosomiasis in a Pregnant Traveler. Emerging Infectious Diseases. 15: 11 1866-1867.

Njiru Z. K., Oumab J. O., Enyaru J. C. and Dargantes A. P. (2010) Loop-mediated Isothermal Amplification (LAMP) test for detection of *Trypanosoma evansi* strain B. Experimental Parasitology 125 196-201.

Nogueira N. F. S., Gonzalez M. S., Gomes J. E. De Souza W., Garcia E. S. Azambuja P., Nohara L. L., Almeida I. C., Zingales B. and Colli W. (2007) *Trypanosoma cruzi*: involvement of glycoinositolphospholipids in the attachment to the luminal midgut surface of *Rhodnius prolixus*. Experimental Parasitology. 116:2 120-128.

Norman F. F. and López-Vélez R. (2014) Mother-to-child transmission of *Trypanosoma cruzi* infection (Chagas disease): a neglected problem. Trans R Soc Trop Med Hyg. 108:7 388-390.

Notomi T., Okayama H., Masubuchi H., Yonekawa T., Watanbe K., Amino N. and Hase T. (2000) Loop-mediated isothermal amplification of DNA. Nucleic Acids Research 28:12 62-69.

Nouvellet P., Buesching C. D., Dugdale H. L., Newman C. and Macdonald D. W. (2011) Mouthing off about developmental stress: Individuality of palate marking in the European badger and its relationship with juvenile parasitoses. Journal of Zoology 283:52 52-62.

Noyes H.A., Stevens J. R., Teixeirac M., Pheland J. and Holz P. (1999) A nested PCR for the ssrRNA gene detects *Trypanosoma binneyi* in the platypus and *Trypanosoma sp*. in wombats and kangaroos in Australia. International Journal for Parasitology 19 331-339

Noyes H.A., Ambrose P., Barker F., Begon M., Bennet M., Bown K. J. and Kemp S. J. (2002) Host specificity of Trypanosoma (Herpetosoma) species: evidence that bank voles (*Clethrionomys glareolus*) carry only one *T. (H). evotomys* 18S rRNA genotype but wood mice (*Apodemus sylvaticus*) carry at least two polyphyletic parasites. Parasitology 124 185-190.

Nyakatura K. and Bininda-Emonds O. R. P. (2012)Updating the evolutionary history of Carnivora (Mammalia): a new species-level supertree complete with divergence time estimates. BMC Biology 10:12.

O'Brien D. J., Schmitt S. M., Rudolph B. A. and Nugent G. (2011) Recent advances in the management of bovine tuberculosis in free-ranging wildlife. Veterinary Microbiology 151 23-33.

O'Meara D. B., Edwards C. J., Sleeman D. P., Cross T. F., Statham M. J., Mcdowell J. R., Dillane E. J., Coughlan J. P., O'Leary D. P., O'Reilly C., Bradley D. G. and Carlsson J. (2012) Genetic structure of Eurasian badgers *Meles meles* (Carnivora: Mustelidae) and the colonization history of Ireland. Biological Journal of the Linnean Society 106 893-909.

Oliveira J., Busson M., Etain B., Jamain S., Hamdani N., Boukouaci W., Amokrane K., Bennabi M., LeGuen E., Dargél A. A., Houenou J., Ivanova R., Bellivier F., Henry C., Kahn J., Charron D., Krishnamoorthy R., Vervoitte L., Leboyer M. and Tamouza R. (2013) Polymorphism of Toll-like receptor 4 gene in bipolar disorder. Journal of Affective Disorders, pre-press copy.

Oltenacu P. A. and Broom D. M. (2010) The impact of genetic selection for increased milk yield on the welfare of dairy cows. Animal Welfare 2010 19:S 39-49.

Oosting M., ter Hofstede H., Sturm P., Adema G. J., Kullberg B., van der Meer J. W. M., Netea M. G. and Joosten L. A. B. (2011) TLR1/TLR2 Heterodimers Play an Important Role in the Recognition of Borrelia Spirochetes. PLoS ONE 6(10): e25998.

Palm N. W. and Medzhitov R. (2009) Pattern recognition receptors and control of adaptive immunity. Immunological Reviews 227 221-233.

Parida M., Sannarangaiah S., Dash P. K., Rao P. V. L. and Morita K. (2008) Loop mediated isothermal amplification (LAMP): a new generation of innovative gene amplification technique; perspectives in clinical diagnosis of infectious diseases. Reviews in Medical Virology 18 407-421.

Park B. S., Song D. H., Kim H. M., Choi B., Lee H. and Lee J. (2009) The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. Nature 458 1191-1197.

Parikh P. P., Zheng J., Logan-Klumpler F., Stoeckert Jr. C. J., Louis C., Topalis P., Protasio A. V., Sheth A. P., Carrington M., Berriman M. and Sahoo S. S. (2012) The Ontology for Parasite Lifecycle (OPL): towards a consistent vocabulary of lifecycle stages in parasitic organisms. Journal of Biomedical Semantics 3:5

Paul M., Stefaniak J., Smuszkiewicz P., Van Esbroeck M., Geysen D. and Clerin J. (2014) Outcome of acute East African trypanosomiasis in a Polish traveller treated with pentamidine. BMC Infectious Diseases 14:111 http://www.biomedcentral.com/1471-2334/14/111

Pereira M. E. A., Zhang K., Gong Y., Herrera E. M. and Ming M. (1996) Invasive phenotype of *Trypanosoma cruzi* restricted to a population expressing trans-sialidase. Infection and Immunity. 64: 9 3884-3892.

Perić M., Bošnjak Z., Šarkanj B., Barbić J., Antolović-Požgain J., Ružman N., Roksandić-Križan I. and Vukovic D. (2014) Polymorphisms of Toll-like receptors 2 and 4 in chronically infected hepatitis C patients from north-east Croatia. Archives of Virology 159:12 Author's copy.

Pifer R. and Yarovinsky F. (2011) Innate responses to *T. gondii* in mice and humans. Trends in Parasitology 27: 9 388-394.

Pine S. O., McElrath M. J. and Bochud P. (2009) Polymorphisms in toll-like receptor 4 and toll-like receptor 9 influence viral load in a seroincident cohort of HIV-1-infected individuals. AIDS 23:2387-2395.

Pinedo P. J., Galvão K. N., Seabury C. M. (2013) Innate immune gene variation and differential susceptibility to uterine diseases in Holstein cows. Theriogenology 80: 384–390

Piscopo T. V. and Azzopardi C. M. (2006) Leishmaniasis. PostGrad Medical Journal 82 649-657.

Reithinger R., Dujardin J., Louzir H., Pirmez C., Alexander B. and Brooker S. (2007) Cutaneous leishmaniasis. Lancet Infectious Diseases 7 581–596.

Pond C. M. (2003) The Fats of Life. Cambridge University Press, Cambridge, UK; pages 34-201.

Pope L. C., Domingo-Roura X., Erven K. and Burke T. (2006) Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale. Molecular Ecology 15 371-386.

Porcel B. M., Denoeud F., Opperdoes F., Noel B., Madoui M., Hammarton T. C., Field M.C., Da Silva C., Couloux A., Poulain J., Katinka M., Jabbari K., Aury J., Campbell D. A., Cintron R., Dickens N. J., Docampo R., Sturm N. R., Koumandou V. L., Fabre S., Flegontov P., Lukes J.,

Michaeli S., Mottram J. C., Szoor B., Dan Zilberstein D., Bringaud F., Wincker P. and Dollet M. (2013) The Streamlined Genome of *Phytomonas* spp. Relative to Human Pathogenic Kinetoplastids Reveals a Parasite Tailored for Plants. PLoS Genet 10(2): e1004007.

Poulin R. and Oikos F. T. (2008) Epigenetic effects of infection on the phenotype of host offspring: parasites reaching across host generations. Oikos 117 331-335.

Prokhortchouk E. and Defossez P. (2008) The cell biology of DNA methylation in mammals. Biochimica et Biophysica Acta 1783 2167-2173.

Pulliam H.R. (2000) On the relationship between niche and distribution. Ecology letters 3 349-361.

Qi Y.M. (1990) Fine structures of the reproductive system of three flea species: development of female genitalia, and, development of male genitalia. Acta Entomologica Sinica 33 182-188 and 403-411.

Re F. and Strominger J. L. (2001) Mechanism of Signal Transduction: Toll-like Receptor 2 (TLR2) and TLR4 Differentially Activate Human Dendritic Cells. Journal of Biological Chemistry 276 37692-37699.

Reed P. and Yarovinsky F. (2011) Innate responses to *Toxoplasma gondii* in mice and humans. Trends in Parasitology 27(9) 388-393.

Reismann P. (2009)Study of the Toll-like receptor 4 gene polymorphisms in diseases presenting with subclinical and chronic inflammation (diabetes mellitus, ischemic stroke, periodontitis) Ph.D. thesis, Semmelweis University. Budapest. 12 pages.

Rey A. W., Newman C., Buesching C. D., Nouvellet P. M. and Macdonald D. W. (2009) Border or Byway: The relationship between European badger (*Meles meles*) path networks and the use and function of associated latrines. Available online at www.ptes.org/files/1600_badgers_paths_and_latrines-_final_report.pdf last accessed 11/06/2013

Ribeiro-Gomes F. L., Peters N. C., Debrabant A. and Sacks D. L. (2012) Efficient capture of infected neutrophils by dendritic cells in the skin inhibits the early anti-Leishmania response. PLoS Pathog 8: e1002536.

Ribeiro J.M.C., Vaughan J.A.and Farhang-Azad A.(1990) Characterisation of salivary apyrase activity of three rodent flea species. Comparative Biochemistry and Physiology B 95 215-218.

Riordan P., Delahay R. J., Cheeseman C., Johnson P. J. and Macdonald D.W. (2011) Culling-induced changes in badger (*Meles meles*) behaviour, social organisation and the epidemiology of bovine tuberculosis. PLoS ONE 6(12): e28904.

Roach J. C., Glusman G., Rowen L., Kaur A. Purcell M. K., Smith K. D. Hood L. E. and Aderem A. (2005) the evolution of vertebrate Toll-like receptors. Procedures of the National Academy of Science USA 102 9577-9582.

Rodrigues M. M., Oliveira A. C. and Bellio M. (2012) The Immune Response to *Trypanosoma cruzi* : Role of Toll-Like Receptors and Perspectives for Vaccine Development. Journal of Parasitology Research: Article ID 507874, 12 pages.

Rodriguez-Osorio C. A., Lima G., Herrera-Caceres J. O., Villegas-Torres B. E., Zuniga J., Ponce-de-Leon S., Llorente L. and Sifuentes-Osorniog J. (2013) Genetic variations in toll-like receptor 4 in Mexican-Mestizo patients with intra-abdominal infection and/or pneumonia. Immunology Letters 153 41-46.

Roper T. J., Ostler J. R., Schmid T. K.and Christian S. F. (2001) Sett use in European badgers. Behaviour 138 173-187.

Ropert C., Ferreira L. R. P., Campos M. A. S., Procópio D. O., Travassos L. R., Ferguson M. A. J., Reis L. F. L., Teixeira M. M., Almeida I. C. and Gazzinelli R. T. (2002) Macrophage signaling by glycosylphosphatidylinositol-anchored mucin-like glycoproteins derived from *Trypanosoma cruzi* trypomastigotes. Microbes and Infection 4 1015-1025.

Rosalino L. M., Torres J. and Santos-Reis M. (2006) A survey of helminth infection in Eurasian badgers (*Meles meles*) in relation to their foraging behaviour in a Mediterranean environment in southwest Portugal. European Journal of Wildlife Research 52 202-206.

Runckel C., DeRisi J. and Flenniken M. L. (2014) A Draft Genome of the Honey Bee Trypanosomatid Parasite *Crithidia mellificae*. PLoS ONE 9(4): e95057.

Rust M. K. (1994) Interhost movement of adult cat fleas. Journal of Medical Entomology 31 486-489.

Rychlik W., Spencer W. J. and Rhoads R. E. (1990) Optimization of the annealing temperature for DNA amplification in vitro. Nucleic Acids Research 18: 21.

Sacks D. L. and Kamhawi S. (2001) Molecular aspects of parasite-vector and vector–host interactions in leishmaniasis. Annual Reviews in Microbiology 55 453-483.

Sabroei I., Read R. C., Whyte M. K. B., Dockrell D. H., Vogel S. N. and Dower S. K. (2003) Toll-Like Receptors in Health and Disease: Complex Questions Remain. The Journal of Immunology 171: 1630-1635.

Sam-Agudu N. A., Greene J. A., Opoka R. O., Kazura J. W., Boivin M. J., Zimmerman P. A., Riedesel M. A., Bergemann T. L., Schimmenti L. A. and John C. C. (2010) TLR9 Polymorphisms are Associated with Altered IFN-γ Levels in Children with Cerebral Malaria. American Journal of Tropical Medicine and Hygeine 82(4) 548–555.

Sampath V., Mulrooney N. P., Garland J. S., He J., Patel A. L., Cohen J. D., Simpson P. M. and Hines R. N. (2013) Toll-like receptor genetic variants are associated with gram-negative infections in VLBW infants. Journal of Perinatology 10.1038.

San E. D. L. (2007) Ectoparasite Categories of The Eurasian Badger (*Meles Meles*) In Western Switzerland. Wildlife Biological Practice 3:2 52-59.

Sarkar A., Ghosh S., Pakrashi S., Roy D., Sen S. and Chetterjee M. (2012) Leishmania strains causing self-healing cutaneous leishmaniasis have greater susceptibility towards oxidative stress. Free Radicals Research 46 665–673.

Savino W., Villa-Verde D. M. S., Mendes-da-Cruz D. A., Silva-Monteiro E., Perez A. R., Aoki M. D. P., Bottasso O., Guinazu N., Silva-Barbosa S. D. and Gea S. (2007) Cytokines and cell adhesion receptors in the regulation of immunity to *Trypanosoma cruzi*. Cytokine & Growth Factor Reviews 18 107-124.

Scheppers T. L. J., Frantz A. C., Schaul M., Engel E., Breyne P., Schley L. and Roper T. J. (2007) Estimating social group size of Eurasian badgers *Meles meles* by genotyping remotely plucked single hairs. Journal of Wildlife Biology 13 195-207.

Schmitz V., Svensjo E., Serra R. R., Teixeira M. M., and Scharfstein J. (2009) Proteolytic generation of kinins in tissues infected by *Trypanosoma cruzi* depends on CXC chemokine secretion by macrophages activated via Toll-like 2 receptors. Journal of Leukocyte Biology 85:6 1005-1014.

Schnare M., Barton G. M., Holt A. C., Takeda K., Akira S. and Medzhitov R. (2001)Toll-like receptors control activation of adaptive immune responses. Nature Immun. 2:10 947-952.

Schwartza D. A. and Cooka D. N. (2005) Polymorphisms of the Toll-Like Receptors and Human Disease. Clinical Infectious Diseases 41 403-407.

Selvaraj P., Harishankar M., Singh B., Jawahar M. S. and Banurekha V. V. (2010) Toll-like receptor and TIRAP gene polymorphisms in pulmonary tuberculosis patients of South India. Tuberculosis 90: 306-310.

Service M. W. (1986) Blood-sucking insects: vectors of disease. Studies in Biology 167 54-59.

Sharma U. and Singh S. (2009) Immunobiology of leishmaniasis. Indian Journal of Experimental Biology 47 412-423.

Shio M. T., Hassani K., Isnard A., Ralph B., Contreras I., Gomez M. A., Abu-Dayyeh I. and Olivier M. (2012) Host Cell Signalling and Leishmania Mechanisms of Evasion. Journal of Tropical Medicine. 14 pages e819512.

Shryock J. A. and Houseman R.M. (2005) A comparison of faecal protein content in male and female cat fleas, *Ctenocephalides felis*. Florida entomologist 88 335-337.

Sievers F., Wilm A., Dineen D., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Soding J., Thompson J. D. and Higgins D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology 7 539.

Simarro P. P., Diarra A., Postigo J. A. R., Franco J. R. and Jannin J. G. (2011) The Human African Trypanosomiasis Control and Surveillance Programme of the World Health Organization 2000-2009. PloS Neglected Tropical Disease 5:2.

Simarro P. P., Cecchi G., Paone M., Franco J. R., Diarra A., Ruiz J. A., Fevre E. M., Courtin F., Mattioli R. and Jannin J. G. (2010) The Atlas of human African trypanosomiasis: a contribution to global mapping of neglected tropical diseases. International Journal of Health Geographics. 9:57 1-18.

Sin Y. W., Dugdale H. L., Newman C., Macdonald D. W. and Burke T. (2012) Evolution of MHC class I genes in the European badger (*Meles meles*). Ecology and Evolution 2:7 1644–1662.

Sin Y. W., Dugdale H. L., Newman C., Macdonald D. W. and Burke T. (2011) MHC class II genes in the European badger: characterization, patterns of variation, and transcription analysis. Immunogenetics s00251-011-0578-9.

Sinha A., Grace C., Alston W. K., Westenfeld F. and Maguire J. H. (1999) African Trypanosomiasis in Two Travelers from the United States. Clinical Infectious Diseases 29 840-844.

Skerrett S. J., Wilson C. B., Liggitt H. D. and Hajjar A. M. (2007) Redundant Toll-like receptor signaling in the pulmonary host response to *Pseudomonas aeruginosa*. American Journal of Physiology - Lung Cellular and Molecular Physiology 292:1 312-322.

Skuce R. A., Allen A. R. and McDowell S. W. J. (2011). Bovine tuberculosis (TB): a review of cattle-to-cattle transmission, risk factors and susceptibility. 167 pages www.dardni.gov.uk

Slatko B. E., Eckert R. L., Albright L. M. and Ausubel F. M. (1999) DNA Sequencing Strategies. Current Protocols in Molecular Biology 7.1 1-7

Smith A., Telfer S., Burthe S., Bennett M.and Begon M. (2005) Trypanosomes, fleas and field voles: ecological dynamics's of host of ectoparasite interaction. Parasitology 131 355-365.

Söding J. (2005) Protein homology detection by HMM–HMM comparison. Bioinformatics 21:7 951-960

Souto-Padrón T. (2002) The surface charge of trypanosomatids. Anais da Academia Brasileira de Ciências 74:4 649-675.

Speletas M., Merentiti V., Kostikas K., Liadaki K., Minas M., Gourgoulianis K. and Germenis A. E. (2009) Association of TLR4-T399I Polymorphism with Chronic Obstructive Pulmonary Disease in Smokers. Clinical and Developmental Immunology 260286 6 pages.

Srivastava K., Srivastava A., Kumar A. and Mittal B. (2010) Significant association between toll-like receptor gene polymorphisms and gallbladder cancer. Liver International 1478-3223.

Stanko M., Miklisova D., Guy de Belloq J. and Morand S. (2002) Mammal density and patterns of ectoparasite species richness and abundance. Oecologica 131 289-295.

Stappers M. H. T., Thys Y., Oosting M., Planting T. S., Ioana M., Reimnitz P., Mouton J. W., Netea M. G., Joosten L. A. B. and Gyssens I. C. (2014) TLR1, TLR2, and TLR6 Gene Polymorphisms Are Associated With Increased Susceptibility to Complicated Skin and Skin Structure Infections. Journal of Infectious Diseases. Advance Access; Author's Copy.

Statham M., Turner P. D. and O'Reilly C. (2005) Use of PCR amplification and restriction enzyme digestion of mitochondrial D-loop for identification of mustelids in Ireland. The Irish Naturalists' Journal 28:1 1-6.

Stenger S., Engele M., Bolcskei P. L., Rollinghoff M. and Wagner M. (2002) Induction of direct antimicrobial activity through mammalian toll-like receptors. Pneumologie 56:103-107.

Stephens N. A., Kieft R., MacLeod A., and Hajduk S. L. (2012) Trypanosome resistance to human innate immunity: targeting Achilles' heel. Trends in Parasitology 28: 12 539-546.

Stevens J., Noyes H. and Gibson W. (1998) The Evolution of Trypanosomes Infecting Humans and Primates. Mem Inst Oswaldo Cruz 93:5 669-676.

Steverding D. (2008) The History of African trypanosomiasis. Parasites & Vectors 1:3 8 p.

Stewart P. D., Bonesi L. and Macdonald D. W. (1999) Individual differences in den maintenance effort in a communally dwelling mammal: the Eurasian badger. Animal behaviour 57 153–161.

Strain S. A. J., McNair J. and McDowell S. W. J. (2011) Bovine tuberculosis: a review of diagnostic tests for *M. bovis* infection in badgers. Bacteriology Branch Veterinary Sciences Division Agri-Food and Biosciences Institute 41 pages.

Sun L., Song Y., Riaza H., Yang H., Hua G., Guoc A. and Yanga L. (2012) Polymorphisms in toll-like receptor 1 and 9 genes and their association with tuberculosis susceptibility in Chinese Holstein cattle. Veterinary Immunology and Immunopathology 147 195-201.

Swingle L. D. (1911) The Transmission of *Trypanosoma lewisi* by Rat Fleas (Ceratophyllus Sp. and Pulex Sp)., with Short Descriptions of Three New Herpetomonads. The Journal of Infectious Diseases 8:2 125-146.

Takeda K. and Akira S. (2005) Toll-like receptors in innate immunity. International Immunology 17: 1 1-14.

Tarleton R. L. (2007) Immune system recognition of *Trypanosoma cruzi*. Current Opinion in Immunology 19 430-434.

Tashima S., Kaneko Y., Anezaki T., Baba M., Yachimori S., Abramov A. V., Saveljev A. P.and Masuda R. (2011) Phylogeographic Sympatry and Isolation of the Eurasian Badgers (*Meles*, Mustelidae, Carnivora): Implications for an Alternative Analysis Using Maternally as well as Paternally Inherited Genes. Zoological Science 28 293-303.

Taylor M. A., Coop R. L and Wall R. L. (2013) Veterinary Parasitology: 3rd Edition. Wiley-Blackwell. 900 pages.

Taylor S. D., Dittmar de la Cruz K., Porter M. L. and Whiting M.F.(2005) Characterisation of the long wavelength opsin from Mecoptera and Siphonaptera: Does a flea see? Molecular biology and evolution 22 1165-1174.

Taylor G. R., Ellis L. A., Robinson M. D., Charlton R. S., Mueller R. F., Knowles M. A. and Bishop D. T. (1999) How Accurate is Heterozygote Base Calling in Dye-Terminator Sequencing? Accessed online at www.dartmouth.edu/ ~staphy/protocol/ download/ Big_Dye_Sequence.doc 11/01/2014

Teixeira M. M., Gazzinelli R. T. and Silva J. S. (2002) Chemokines, inflammation and *Trypanosoma cruzi* infection. Trends in Parasitology 18:6.

Terry R. S., Smith J. E., Duncanson P. and Hide G (2001) MGE-PCR: a novel approach to the analysis of *Toxoplasma gondii* strain differentiation using mobile genetic elements. International Journal of Parasitology 2001 31:2 155-161.

The Game Conservancy Trust (2006) Shooting as a potential tool in badger population control. Report to Defra 39p.

Thekisoe O. M. M., Honda T., Fujita H., Battsetseg B., Hatta T., Fujisaki K., Sugimoto C. and Inoue N. (2007) A trypanosome species isolated from naturally infected *Haemaphysalis hystrici*s ticks in Kagoshima Prefecture, Japan. National Research Center for Protozoan Diseases 32 pages.

Thekisoe O. M. M., Inoue N., Kuboki N., Tuntasuvan D.,Bunnoy W., Borisutsuwan S., Igarashi I. and Sugimoto C. (2005) Evaluation of loop-mediated isothermal amplification (LAMP), PCR and parasitological tests for detection of *Trypanosoma evansi* in experimentally infected pigs. Veterinary Parasitology 130 327-330.

Thermo Scientific (2012) NanoDrop Lite User Guide 269-275200 Thermo Scientific Inc.

Thomson R., Samanovic M. and Raper J. (2009). Activity of trypanosome lytic factor: a novel component of innate immunity. Future Microbiology. 4 789-796.

Thornton P. D., Newman C., Johnson P. J., Buesching C. D., Baker S. E., Slater D., Johnson D. D. P. and Macdonald D. W. (2005) Preliminary comparison of four anaesthetic techniques in badgers (*Meles meles*). Veterinary Anaesthesia and Analgesia 32 40-47.

Tiberti N., Lejon V., Ngoyi D. M., Matovu E., Enyaru J., Walter N., Fouda C., Lutumba P., Kristensson K., Bisser S., Ndungu J. M., Büscher P. and Sanchez J-C (2014) Increased acute immune response during the meningo-encephalitic stage of *Trypanosoma brucei rhodesiense* sleeping sickness compared to *Trypanosoma brucei gambiense*. Translational Proteomics 375.

Torres C., Vitalis E. A., Baker B. R., Gardner S. N., Torres M. W. and Dzenitis J. N. (2011) LAVA: An Open-Source Approach to Designing LAMP (Loop-Mediated Isothermal Amplification) DNA Signatures. BMC Bioinformatics 12:240.

Tschirren B., Andersson A., Scherman K., Westerdahl H., Mittl P. and Råberg L. (2013) Polymorphisms at the innate immune receptor TLR2 are associated with *Borrelia* infection in a wild rodent population. Proceeds of the Royal Socociety of Biology 280.

Tschirren B., Raberg I and Westerdahl H. (2011) Signatures of selection acting on the innate immunity gene Toll-like receptor 2 (TLR2) during the evolutionary history of rodents. Journal of Evolutionary Biology 24 1232-1240.

Tyler K. M. and Engman D. M. (2001) The life cycle of *T. cruzi* Revisited. International Journal for Parasitology 31 472-481.

Vashchenok V. S. (1993) Factors regulating egg production in fleas *Leptopyslla segnis*. Parasitology 27 382 – 388.

Vashchenok V. S. (1997) Ecology of flea groups of the species conformis (*Xenopsylla*) of the fauna of Russia and adjacent countries. Parazitologiia 31:6 492-513.

Vasselon T. and Detmers P. A. (2002) Toll Receptors: a Central Element in Innate Immune Responses. Infection and Immunity 70:3 1033-1043.

Velayati A. A., Farnia P and Mirsaeidi M. (2015) Persistence of *Mycobacterium tuberculosis* in environmental samples. International Journal of Mycobacteriology 4:1.

Velez D. R., Wejse C., Stryjewski M. E., Abbate E., Hulme W. F., Myers J. L., Estevan R., Patillo S. G., Olesen R., Tacconelli R., Sirugo G., Gilbert J. R., Hamilton C. D., and Scott W. K. (2010) Variants in toll-like receptors 2 and 9 influence susceptibility to pulmonary tuberculosis in Caucasians, African-Americans, and West Africans. Human Genetics 127:1 65-73.

Velez D. R., Hulme W. F., Myers J. L., Stryjewski M. E., Abbate E., Estevan R., Patillo S. G., Gilbert J. R., Hamilton C. D. and Scott W. K. (2009) Association of SLC11A1 with tuberculosis interactions with NOS2A and TLR2 in African-Americans and Caucasians. International Journal of Tuberculosis Lung Disorders 13:9 1068-1076.

Veterinary Laboratories Agency (2010) Efficacy testing of BCG vaccine in badgers. SID 5 Research Project Final Report: Department for Environment Food and Rural Affairs. 19 pages.

Wamack M. (2013) Private communication with Technical Support, Eurofins MWG GmbH.

Wambwa E. (2003) Diseases of importance at the Wildlife/ Livestock interface in Kenya. The IUCN Species Survival Commission: Conservation and Development Interventions at the Wildlife/Livestock Interface Implications for Wildlife, Livestock and Human Health. Ch3 21-25.

Wang C., Chang S., Wu Y. J, Lin J., Ho H., Chou T., Yang B., Wu J. and Chen J. (2014) Genetic variations in Toll-like receptors (TLRs 3/7/8) are associated with systemic lupus erythematosus in a Taiwanese population. Scientific Reports 4:3792.

Weber M. and Schubeler D. (2007) Genomic patterns of DNA methylation: targets and function of an epigenetic mark. Current Opinion in Cell Biology 19 273-280.

Weber N., Carter S. P., Dall S, R. X., Delahay R. J., McDonald J. L., Bearhop S. and McDonald R. A. (2013) Badger social networks correlate with tuberculosis infection. Current Biology 23:20 915-916.

Weitzel T., Zulantay I., Danquah I., Hamann L., Schumann R., Apt W. and Mockenhaupt F. P. (2012) Short Report: Mannose-Binding Lectin and Toll-Like Receptor Polymorphisms and Chagas Disease in Chile. American Journal Tropical Medicine and Hygeine 86:2 229-232.

Welburn S. C., Fevre E. M., Coleman P. G., Odiit M. and Maudlin I. (2001) Sleeping Sickness: a tale of two diseases. Trends in Parasitology 17:1 19-24.

Whatman (2008) DNA Extraction from FTA Cards Using the GenSolve DNA Recovery Kit. GenVault, Inc. Illumina GoldenGate Assay; Pub. No.370-2007-032.

Weber N., Bearhop S., Dall S. R. X., Delahay R. J., McDonald R. A., Carter, S. P. (2013) Denning behaviour of the European badger (*Meles meles*) correlates with bovine tuberculosis infection status. Behavioral Ecology and Sociobiology 67:3 471-479.

Weng P., Huang P., Page J. H., Chen J., Xu J., Koutros S., Berndt S., Chanock S., Yeager M., Witte J. S., Eeles R. A., Easton D. F., Neal D. E., Donovan J., Hamdy F. C., Muir K. R., Giles G., Severi G., Smith J. R., Balistreri C. R., Shui I. M. and Chen Y. (2014) Polymorphisms of an Innate Immune Gene, Toll-Like Receptor 4, and Aggressive Prostate Cancer Risk: A Systematic Review and Meta-Analysis. PLoS one 9:10 e110569.

Wheeler R. J., Gluenz E. and Gull K. (2011) The cell cycle of Leishmania: morphogenetic events and their implications for parasite biology. Molecular Microbiology 79:3 647-662.

Whiteoak A. M., Ideozu E. J,., Alkathiry H., Tomlinson A. J., Delahey R. J., Cowen S., Mullineaux E., Gormley E., Birtles R. J., Lun Z-R and Hide G. (unpublished, 2015) Low gene variation in Toll-like Receptors (2 and 4) in the European Badger *Meles meles*.

Wilkinson D., Smith G. C., Delahay R. J., Rogers L. M., Cheeseman C. L. and Clifton-Hadley R. S. (2000) The effects of bovine tuberculosis (*Mycobacterium bovis*) on mortality in a badger (*Meles meles*) population in England. Journal of Zoology 250 389-395.

Williams R. (2013) Bruce Beutler: The Persistent Prospector. Profiles in Cardiovascular Science. Circulation Research (March 2013) American Heart Association 751-755.

Willie B., Hall N. B., Stein C. M., Jurevic R. J., Weinberg A., Mehlotra R. K. and Zimmerman P. A. (2014) Association of Toll-like receptor polymorphisms with HIV status in North Americans. Genes and Immunity 1-9.

Wujcicka W., Wilczyński J. and Nowakowska D. (2014) Do the placental barrier, parasite genotype and Toll-like receptor polymorphisms contribute to the course of primary infection with various *Toxoplasma gondii* genotypes in pregnant women? European Journal of Clinical and Microbiological Infectious Diseases (2014) 33 703-709.

Xiao R. and Boehnke M. (2009) Quantifying and correcting for the winner's curse in genetic association studies. Genetics and Epidemiology 33(5) 453-462.

Xie F., Hu Y., Speert D. P., Turvey S. E., Peng G., Money D. M., Magee L. A. and von Dadelszen P. (2010) Toll-like Receptor Gene Polymorphisms and Pre-eclampsia Risk: A Case-Control Study and Data Synthesis. Hypertension in Pregnancy, Early Online Publication: 1-9.

Yang H., Lu K., Lee H., Huang S., Lin Y. and Wu C. (2013) Role of the Functional Toll-Like Receptor-9 Promoter Polymorphism (-1237T/C) in Increased Risk of End-Stage Renal Disease: A Case-Control Study. PLoS ONE 8(3): e58444.

Yanik M., Gurel M. S., Simsek Z. and Kati M. (2004) The psychological impact of cutaneous leishmaniasis. Clinical and Experimental Dermatology 29 464-467.

Yarovinsky F. (2008) Toll-like receptors and their role in host resistance to *T. gondii*. Immunology Letters 119 17-21.

Zhang J., Zhu Q., Meng F., Lei H. and Zhao Y. (2014) Association study of TLR-9 polymorphisms and systemic lupus erythematosus in Northern Chinese Han population. Gene 533 385-388

Zhang Y., Jiang T., Yang X., Xue Y., Wang C., Liu J., Zhang X., Chen Z., Zhao M. and Li J. (2013) Toll-Like Receptor -1, -2, and -6 Polymorphisms and Pulmonary Tuberculosis Susceptibility: A Systematic Review and Meta-Analysis. PLoS ONE 8(5): e63357.

Zidi S., Verdi H., Yilmaz-Yalcin Y., Yazici A. C. , Gazouani E., Mezlini A., Atac F. and Yacoubi-Loueslati B. (2014) Impact of Toll-Like Receptors 2/3/4/9, IL-1-α/β and TNF-α Polymorphisms in Cervical Cancer Susceptibility in Tunisia. Pathology and Oncology Research; Author's personal copy.