University of Alberta

The Genetics of Classical BSE in European Holstein Cattle

by

Brenda Mae Murdoch

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Animal Science

Agricultural, Food and Nutritional Science

Fall 2010
Edmonton, Alberta

Examining Committee

Stephen S. Moore, Agricultural, Food and Nutritional Science

Walter T. Dixon, Agricultural, Food and Nutritional Science

Erasmus K. Okine, Agricultural, Food and Nutritional Science

Judd Aiken, Agricultural, Food and Nutritional Science

William T. Allison, Biological Sciences

Simon Mead, MRC Prion Unit, Institute of Neurology, University College of London

Abstract

Classical bovine spongiform encephalopathy (BSE) is an acquired prion disease that is invariably fatal in cattle and has been implicated as a significant human health risk. Polymorphisms that alter the prion protein of sheep or humans have been associated with transmissible spongiform encephalopathy susceptibility or resistance. In contrast, there is no strong evidence that non-synonymous mutations in the bovine prion gene (*PRNP*) are associated with classical BSE disease susceptibility. However, two bovine *PRNP* insertion/deletion polymorphisms, one within the promoter region and the other in intron 1, have been associated with susceptibility to classical BSE. These associations do not explain the full extent of BSE susceptibility, and loci outside of *PRNP* appear to be associated with disease incidence in some cattle populations**.** To test for associations with BSE susceptibility, we conducted a genome wide scan using a panel of 3,072 single nucleotide polymorphism markers on 814 animals representing cases and control Holstein cattle from the United Kingdom BSE epidemic. A higher resolution, 50K SNP, genome-wide scan was used with a sub-set (N=330) of animals to test for association with BSE incidence. Further, the bovine prion gene was tested, 19 haplotype tagging SNPs and 2 insertion/deletion alleles for association with BSE disease incidence.

The Holstein case and control cattle analyzed here consisted of two sets, one set with known family relationships and the second set with BSE cases with paired controls. The family set comprised of half-sibling progeny from six sires, revealed 27 SNPs representing 18 chromosomes associated with incidence of BSE disease, confirming previously reported chromosomal regions. Further analysis of the BSE case and paired control samples using the 50K SNP chip a polymorphic locus was identified on chromosome 1 at 29,147,078 bp with a moderate significant association (p=3.09E-5) and

over-represented in BSE affected animals. Further, analysis of the prion gene and the paired BSE case and control animals revealed a haplotype within the region of high LD to associate with BSE un-affected animals (p-value =0.000114). It is clear from our analyses that several regions of the genome are statistically associated with the incidence of classical BSE in European Holstein cattle.

Hypotheses

The hypotheses tested in this research project include the following:

1) Genetic variation in the bovine prion gene, *PRNP*, is associated with classical BSE disease status.

2) Chromosomal regions in the bovine genome are associated with BSE disease incidence.

3) High-resolution genome scanning of paired Holstein BSE case and control cattle will identify positional candidate genes that affect disease status.

Acknowledgements

This research was supported by PrioNet (Stephen Moore, John Williams, Denny Crews and Mike Coulthart) and the Alberta Prion Research Institute project #200500696 (Stephen Moore, Stephanie Czub, Catherine Graham and Mike Heaton) for their financial support.

I would like to thank the staff at the Roslin Institute, United Kingdom and in particular Emma Borthwick, Auriol Williams for their assistance in the laboratory. Additionally, I would like to thank Phil Steele at the Moredun Institute for his willingness to allow me to isolate DNA from BSE infected blood in their biosafety level three laboratory.

I would like to offer a special thanks to the Alberta Bovine Genomics program, my fellow graduate students and technicians in the Agriculture and Proteomic Center for their interactions and support throughout my graduate program.

Further, I wish to thank my supervisory and examining committee for their constructive advice. I would like to offer special gratitude to my supervisor Dr. Stephen Moore for allowing me this academic opportunity. Additionally, thank you for the unique and rewarding experience of presenting my results internationally and thus allowing me to interact with renowned international prion scientists. In addition, I would like to thank Dr. Dixon for all of his assistance and intellectually stimulating conversations throughout my program.

Finally, I wish to dedicate this thesis to my family who have always encouraged me to follow my heart. In particular my mother, Marieanne and sister, Charlene have supported me throughout my childhood and youthful adventures and continue to be there for my every need. My husband, Gordon and children Xavier and Breanna continually offer their optimistic support for all of my endeavours despite the sacrifices. Their faith in my ability and their unwavering enthusiasm gave me the confidence to pursue my ambitions.

Table of Contents

List of Tables

Chapter 4

List of Figures

Chapter 1

List of Abbreviations

a.a.	amino acid
BSE	bovine spongiform encephalopathy
BTA	*Bos taurus* autosome
C1	complement component 1
C3	complement component 3
CJD	Creutzfeldt-Jakob disease
gCJD	genetic Creutzfeldt-Jakob disease/familial Creutzfeldt-Jakob disease
fCJD	familial Creutzfeldt-Jakob disease/ genetic Creutzfeldt-Jakob disease
iCJD	iatrogenic Creutzfeldt-Jakob disease
sCJD	sporadic Creutzfeldt-Jakob disease
vCJD	new variant Creutzfeldt-Jakob disease
CNS	central nervous system
CR1	complement receptor 1
CR2	complement receptor 2
CWD	chronic wasting disease
DC	dendritic cell
DFAM	family-based association for disease traits
DNA	deoxyribonucleic acid
Dpl	doppel protein
ER	endoplasmic reticulum
FDC	follicular dendritic cell
FFI	fatal familial insomnia
GAG	glycoaminoglycans
GALT	gut associated lymphoid tissue
GI	gastrointestinal
GPI	glycosylphosphatidyl inositiol
GSS	Gertsmann-Sträussler-Scheinker syndrome
HSA	*homo sapien* autosome
HSPG	heparin sulfate proteoglycan
HGH	human growth hormone
IAP	intracisternal A particle
Indel	insertion deletion allele

kDa	kilodalton
LD	linkage disequilibrium
LTR	long tandemm repeats
LR	laminin receptor
LRP	laminin receptor protein
M-cell	microfold cell
ORF	open reading frame
PS-ON	phosphorothiolated-oligonucleotide
PrP	prion protein
PRNP	prion protein gene (human)
$Prnp^{0/0}$	prion protein gene mouse knock-out (only promoter region removed)
$Prnp^{0/+}$	prion protein gene mouse knock-out with one allele and one wild type allele
$Prnp^{-/-}$	prion protein gene mouse knock-out (exon region removed)
$Prnp^{+/+}$	wild type of mouse prion protein gene
PrP^{c}	cellular prion protein
PrP^{sc}	scrapie prion protein
PrP^{res}	protease resistant prion protein
PRND	doppel protein gene
Prnt	prion family alternatively spliced transcripts
QTL	quantitative trait loci
RNA	ribonucleic acid
SPRN	shadoo gene
Tg	transgenic
TSE	transmissible spongiform encephalopathy

Chapter 1
General Introduction

1.1 Introduction

Prion diseases, also referred to as transmissible spongiform encephalopathies (TSEs), are a group of neurodegenerative diseases that affect humans as well as other mammals and are generally believed to be mediated by a protein pathogen. Although this class of diseases manifest themselves through a variety of neuropathological symptoms, the central nervous system (CNS) is the primary target tissue and is where numerous observed pathologies have been associated with the various disease etiologies. These pathologies include; spongiform changes, neuronal loss, glial cell activation and most notably the accumulation of amyloid aggregates. The pathological manifestations are progressive, cumulative and inevitably fatal given sufficient time. In humans there are several neurological diseases which have been identified and are currently regarded as prion diseases, including; Creutzfeldt-Jakob disease (CJD), fatal familial insomnia (FFI) and Gertsmann-Sträussler-Scheinker syndrome (GSS). There are other known TSE's in many different species. For instance, scrapie in sheep and goats has been documented as far back as the 18th century. Bovine spongiform encephalopathy (BSE) in cattle is probably one of the most infamous TSE, at least partly due to the sensationalization through media reports of "mad cow" disease. The media has also served to publicize the various occurrences of the most prevalent TSE in wildlife; chronic wasting disease (CWD), which has recurrently affected white-tailed deer and elk populations, and is one of the emergent diseases in North America. Prion diseases may be genetic (heritable), infectious or sporadic in origin though, importantly, all involve the modification of the native prion protein into a misfolded infectious form. This misfolding cascade event has been effectively described by Watt *et al.* (2006) as "epigenetic templated protein misfolding", although the process by which one misfolded infectious prion protein initiates an alteration in the folding of an existing native prion protein remains somewhat unclear. Regardless, the hallmark feature of the misfolded protein is an induced change in its secondary structure, more specifically an increase in the ß-pleated sheet content of the protein. The increase in the β-pleated sheet content of the protein greatly enhances its protease resistance and is, in part, responsible for the accumulation of this abnormal

cellular protein, rather than normal cellular clearance as occurs with the native protein. As would be expected, this new class of diseases has spawned much research interest in relation to the similarity and differences between TSE diseases, their etiology, transmissibility, causation and most importantly prevention. It is hoped that this brief review will effectively summarize some of what is generally known about TSE's, with a primary focus on BSE and its associated prion gene and protein.

1.1.1 Brief history of transmissible spongiform encephalopathies TSE's

With respect to the brief history of prion disease discovery and characterization, a recent review article by Onodera *et al.* (2006), referred to the history of prion disease discovery as having occurred in "four separate phases". Phase one was the observation and description of similarities between Creutzfeldt-Jakob Disease (CJD) and scrapie in sheep while phase two was the elucidation of the similarities in prion disease etiologies as well as the mode of transmission of human CJD. It was during this second phase that etiological similarities between the human prion disease "Kuru" and scrapie-infected sheep were noted, more specifically, "spongiform degeneration and astrocytotic gliosis of the central nervous system" (Prusiner, 1998). Additionally, in phase two, cross-species transmission experiments were successfully employed, thereby playing a significant role in the general acceptance that these diffuse neurological manifestations were, in fact, the result of a classifiable and infectious disease. The specific transmission experiments were performed by Gajudsek *et al.* (1966) who transmitted Kuru to chimpanzees and by Tateishi *et al.* (1981) who successfully transmitted human prions to rodents. Upon discovery of transmissibility, many theories were advanced as to the mode of transmission, including the early front-runner pathogen considered to be a 'slow virus'.

It wasn't until the third phase in the prion disease discovery history that the "Protein only theory" was proposed by Prusiner. It was discovered through Prusiner and colleagues' work, that a protein isolate from infectious tissue was partially resistant to proteolytic degradation. In addition, it was noted that the experimental conditions which were used in transmission experiments would most certainly degrade most if not all nucleic acids (ie. DNA and RNA). Prusiner (1998) explicitly stated that, "both the UV and ionizing radiation studies as well as physical studies have eliminated the possibility

of a large nucleic acid hiding within purified preparations of proteins". Thus, Prusiner proposed that the infectivity was a result of the protein present and subsequently coined the now accepted description of this protein as a "prion", which means proteinaceous infectious agent. This work was followed by the development of the transgenic mouse model, in the fourth phase.

The human genetically-inherited prion disease Gertsmann-Straussler-Scheinker syndrome (GSS) was modeled in the transgenic (Tg) mouse (MoPrP-P101L). This not only demonstrated that the PrP gene mutation in humans resulted in disease but further elucidated that nucleic acid changes in the prion gene can also result in manifestation of disease. With the development of the prion "knock-out" mouse, it was further established that the prion protein was necessary for disease development, a significant step in meeting Koch's postulates though, in this case, not for a microbial infectious agent.

1.2 Prion Disease Etiology

1.2.1 Classical disease characteristics

With classical prion diseases, following the initial onset of noticeable symptoms, most humans present with a rapidly progressive dementia which is typically accompanied by the manifestation of ataxia. Other characteristics are generally neuropathological features within the CNS and include spongiform changes, neuronal loss, glial activation and accumulation of amyloid aggregates due to misfolded host prion protein. Typically, TSE diseases have long, variable and somewhat unpredictable incubation periods prior to the onset of clinical signs and symptoms. For instance, in humans the incubation time for vCJD and kuru, can range from 1.5 to 40 years. In addition, there have been reports in which the host does not exhibit clinical signs and further does not succumb to the pathologies associated with prion disease within their normal life span.

1.2.2 Transmission of prion disease

As was briefly mentioned earlier, the source, causation or acquisition of human prion diseases can be genetic (heritable), sporadic (random), or infectious (exposure to

foreign source of misfolded prion protein). The genetic development of prion disease is due to nucleotide mutation(s) within the gene that codes for the prion protein and is subsequently inherited through the germ line. Prion diseases which have not been inherited or transmitted through other known infectious sources are generally termed sporadic, and, in general, the scientific community does not have a good understanding as to the cause of sporadic TSE's. Patients who have acquired a prion disease through infection either consumed the infectious agent or obtained the infectious agent through contamination during medical procedures (iatrogenic).

According to Prusiner (1998) the P102L mutation in the prion protein gene, was the first mutation to be genetically linked to CNS dysfunction in GSS patients. This mutation and four others have subsequently been linked to the genetic transmission of familial human prion disease. In patients where there has not been any prion gene mutations found, these cases are typically classified and considered to be sporadic in nature. Interestingly, sporadic forms constitute the most abundant number of cases of CJD and a few cases of GSS. With respect to infectious human TSE diseases, the classical historical case is that of "Kuru". It is well founded empirically that the epidemiology of Kuru disease involved the prion protein being ingested, and as such transmitted by, ritualistic cannibalism; where the tribe specifically ingested the brain of deceased tribesmen, so as to "preserve" their spirit and knowledge. Other documented infectious sources have been associated with the medical use of material from cadavers, this includes; human growth hormone (HGH), gonadotropin, cornea transplants, and *dura mater* grafts. There have also been cases of iatrogenic transmission from improperly sterilized surgical equipment, especially neurosurgical, though it is feasible that many of these early occurrences went undocumented.

Prions originating in one species typically do not result in disease in another species; this is considered to be due to a species barrier in the infectious transmission of TSEs. There are notable exceptions to this species specificity and when there is cross species transmission it is usually characterized by prolonged incubation times during the first passage to the new host. Upon subsequent passages in a homologous host the incubation time decreases dramatically and observed pathological etiologies are noted. According to Prusiner (1998) there are three factors which contribute to the species barrier. One major factor is PrP, more specifically the sequence differences in PrP which

exist between the donor and the host. For example, if a transgenic mouse produces hamster prion protein it is susceptible to hamster prion disease but not mouse prion disease. Further, if a transgenic mouse expresses both mouse and hamster prion proteins it has a diminished susceptibility to hamster prion disease (Weissmann and Flechsig, 2003). It is worth mentioning that transgenic mice which express bovine prion proteins are susceptible to BSE, scrapie as well as vCJD. The second factor is the strain of the infectious prion particle, and the final factor includes the presence of other species-specific proteins which are thought to bind to or interact with the prion protein and facilitate conformational change. Since this accessory protein(s) is currently unknown and undefined, Prusiner designated this hypothesized protein or one such protein as; "protein X". Although there is considered to be a marked species barrier in the transmission of some prion diseases it is generally accepted that prion disease in cattle (BSE) is the source of a specific human prion disease (via ingestion of the infectious prion protein). Such human cases are termed new variant or variant CJD (vCJD) and are differentiated from heritable CJD.

The mode of transmission of prion protein can vary within species, as observed in humans, as well as between species as observed in ovine and cervids. The following table taken from Mabbott and MacPherson (2006), describes some of the suggested routes of TSE transmission in various species as well as providing reference to instances where the transmission is unclear such as the case with "horizontal and vertical transmission" of scrapie and CWD.

TSE disease	Natural host species affected	Route of transmission
Variant Creutzfeldt–Jakob disease (vCJD)	Humans	Ingestion of BSE-contaminated food. Two cases associated with transfusion of blood from vCJD-infected blood donor
Sporadic Creutzfeldt–Jakob disease	Humans	Unknown. Somatic mutation or spontaneous conversion of PrP^{c} into PrP^{Sc}?
Iatrogenic Creutzfeldt–Jakob disease	Humans	Accidental medical exposure to CJD-contaminated tissues or tissue products
Familial Creutzfeldt–Jakob disease	Humans	Associated with germline mutations in *PRNP* gene
Gerstmann–Sträussler–Scheinker syndrome	Humans	Associated with germline mutations in *PRNP* gene
Fatal familial insomnia	Humans	Associated with germline mutations in *PRNP* gene
Kuru	Humans	Ritualistic cannibalism
Scrapie	Sheep, goats and mouflon	Acquired (for example, ingestion), horizontal transmission, vertical transmission unclear
Bovine spongiform encephalopathy (BSE)	Cattle	Ingestion of BSE-contaminated meat and bonemeal
Chronic wasting disease	Mule deer, white-tailed deer, Rocky mountain elk and moose	Acquired (for example, ingestion), horizontal transmission, vertical transmission unclear
Feline spongiform encephalopathy	Domestic and zoological cats	Ingestion of BSE-contaminated food
Transmissible mink encephalopathy	Farmed mink	Acquired (ingestion) but source unknown
Exotic ungulate encephalopathy	Zoological greater kudu, nyala and oryx	Ingestion of BSE-contaminated food

Table 1-1: List of Transmissible spongiform encephalopathy (TSE) diseases, the host species affected and the route of transmission. (Mabbott and MacPherson, 2006)

Aside from human sporadic and genetic cases many species acquire prion diseases via the ingestion of some form of infectious prion agent. However, in the case of ovine and cervids it is thought that these species have the unfortunate and confounding ability to transmit disease both horizontally (from one cohort to another) and possibly vertically (from parent to offspring). CWD has been clinically recognized as far back as 1967 and is the only known prion disease to affect free-ranging mule deer, white-tailed deer, and rocky mountain elk. The process of horizontal transmission of prion disease such as in the case of CWD is currently not known. Although it is thought that this is due to oral transmission (Johnson *et al.*, 2006) via the alimentary canal and lymphoid tissues, which specific bodily fluids or secretions contain the infectivity has not yet been identified. As can be imagined, this diversification in terms of possible means of infection makes the plausibility of developing effective means of control and perhaps even eradication of CWD and scrapie much less likely in the immediate future. Furthermore, as mentioned earlier, the enhanced stability of the infectious prion protein

will certainly contribute to the horizontal transmission of this pathogen and may even indicate that there need not be direct contact between the host and newly infected animal. It is further possible that considerable time may elapse between potential host contamination of the environment, (soil or plant material), and transmission to a new host animal. (Johnson *et al.*, 2006) Long-term studies related to this are currently underway.

1.2.2.1 Oral route of disease transmission

The oral transmission of infectious prion particles from contaminated feed sources is thought to be the basis of the BSE outbreak in the United Kingdom, though definitive proof of this may be impossible to obtain retroactively. However, there is certainly substantial circumstantial evidence that this is the case, not the least of which is the dramatic reduction in BSE following the imposed ban of animal byproducts in formulations of animal feed. In addition to cattle, other mammals including humans have been known to contract TSE diseases from oral ingestion. Since it had been established that the kuru people of New Guinea contracted prion disease through the practice of ritualistic cannibalism, it was not a far stretch to subscribe to the theory that humans contracted vCJD from ingesting tissues from BSE infected beef. However, it is important to emphasize that the exact nature of this transmission is still not fully understood.

Mabbott and MacPherson (2006) describe a hypothesis for a proposed route from oral introduction to CNS disease manifestations. In addition, they hypothesized that genetic variations within this pathway contribute to the idea of a heritable aspect of susceptibility and/or resistance to disease and to the variability in the associated disease pathology as well as disease incubation timelines. Therefore a greater understanding of this interaction between host genetics and prion infection is important and may be the key to understanding the disease etiology and transmission. The elucidation of the genetic components of prion disease resistance and susceptibility may provide an opportunity to use livestock selection procedures to reduce the frequency and transmission of prion disease.

In order for an ingested prion protein to influence the structural folding of proteins within the CNS of the infected animal, the pathogenic protein must survive degradation within the gut and enter the systemic internal tissues to reach the brain. More

specifically, following ingestion of the pathogenic prion associated with a contaminated feed source, the infectious agent must retain its misfolded structure and cross the intestinal epithelium. This is not an easy task since the tight junctions between gut epithelial cells prevent the unimpeded diffusion of molecules from within the gut lumen into the serosa or vascular layers of the GI tract. However, the presence of microfold cells "M-cells" found within the Peyer's patches of the epithelium enable the hosts immune system to sample the luminal content to ensure that it can initiate the appropriate immune response when necessary. It is thought that the M-cell can actively transport the infectious agent through the gut wall to the basal lateral side on the epithelium. In addition, it is thought that intestinal epithelial cells may endocytose prion fragments which have formed protein complexes (i.e. with ferritin) which may represent an alternative and transcellular route for prions to move from the gut lumen into the host. Finally dendritic cells (DC) have the capacity to insert their dendrites between the epithelial tight junctions and acquire antigen from the lumen, which may directly or inadvertently facilitate the prion transport between these cells. These hypothesized routes of prion entry into the host from the gut lumen are outlined in the following figure taken from Mabbott and MacPherson (2006).

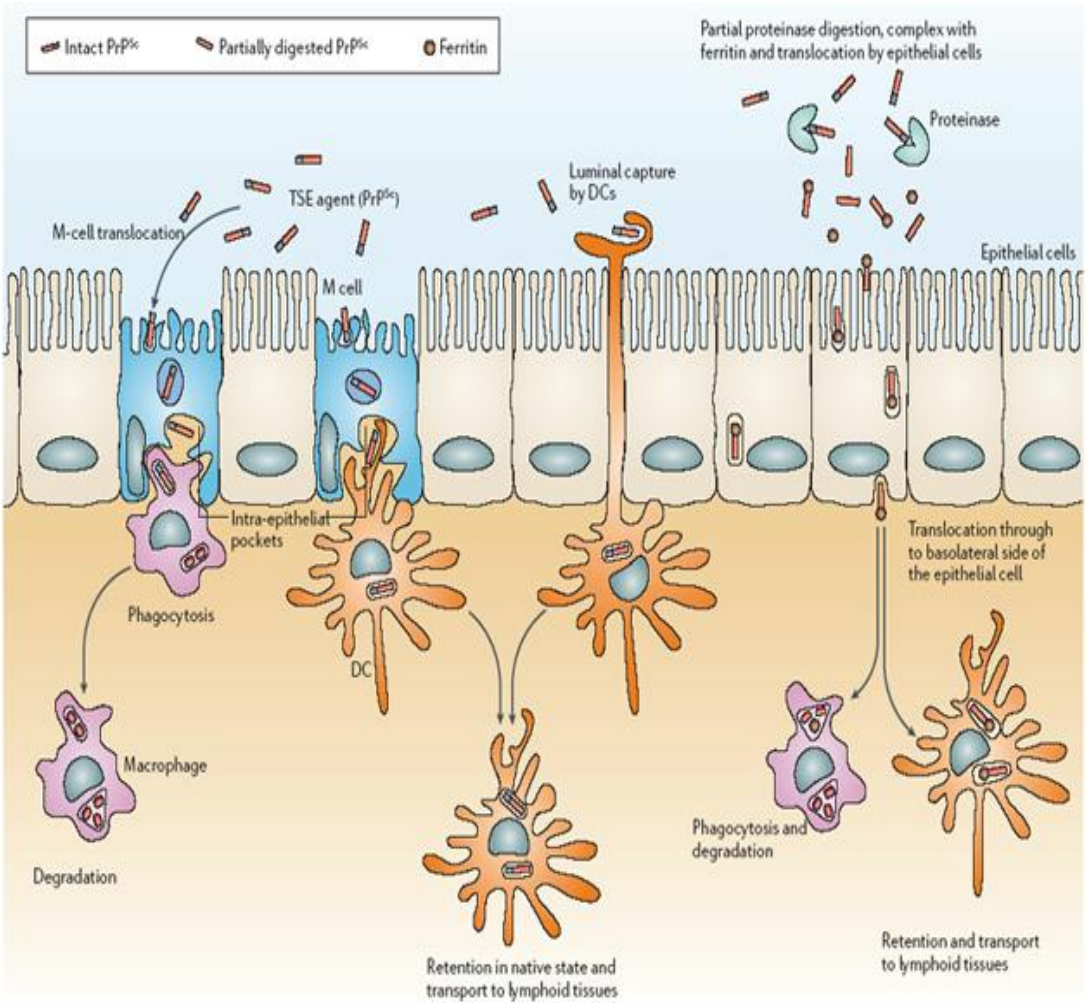


Figure 1-1: Diagrammatic representation of potential mechanism of TSE agent translocation across the intestinal epithelium. M cells are specialized cells for transepithelial transport of macromolecules and particles within the epithelium (plausible sites for transport of TSE agents across the intestinal epithelium). In addition TSE agents may occur independent of M cell via ferritin complexes and translocation. Dendritic cells (DCs) can also acquire antigens directly from the intestinal lumen by opening up tight junctions via dendrite insertion. (Mabbott and MacPherson, 2006)

Following ingestion of the infectious agent and its crossing of the gut epithelium, there is an accumulation of the prion proteins in several lymphoid tissues including the

spleen, lymph nodes, tonsils, and appendix as well as in the Peyer's patches (still within the gut lumen). Depending on the TSE agent and the host there is a large variation in the magnitude and the duration of lymphoid-tissue involvement, and this may also have an as yet undisclosed genetic association. It has been shown that there is an impairment of the typical neuroinvasion in oral challenges of animals where the Peyer's patches, follicular dendrites, or the spleen are absent. This would imply that these tissues or cells are important in the progression of disease (Mabbott and MacPherson, 2006).

Follicular dendritic cells (FDCs) reside in the B-cell follicles and germinal center of lymphoid tissue. As pointed out by Mabbott and MacPherson (2006), cells must express normal cellular prion protein (PrP^{c}) to sustain TSE infection and FDCs express high levels of PrP^{c} in comparison to lymphocytes. The dendritic processes that FDC possess can extend over several lymphocyte diameters. This results in a large surface area which can trap and retain antigens in their native state. Because FDCs are "long-lived-cells" antigens can be retained for months and even years. (Mabbott and MacPherson, 2006) Antigens retained by FDCs form immune complexes (antigen-antibody and/or complement components) through Fc antibody receptors and complement receptors CR1 and CR2. The deficiencies of complement components (C1 and C3) have been observed to impair the accumulation of TSE agent in the spleen. It is hypothesized that TSE agent-compliment-complex could be acquired by receptors CR1/CR2 on the surface of the FDC (see Figure 1-2 b and c). Specifically, the infectious prion protein (PrP^{sc}) accumulates within the germinal centers of the B cell follicles on follicular dendritic cells (FDC). It is here that the PrP^{sc} accumulates on the plasmalemma as well as in the extra cellular space surrounding the dendrite. In addition, tingible body macrophages within the B cell follicles acquire TSE agent (see Figure 1-2 f). Although the exact mechanism of macrophage involvement is unknown it is thought that the macrophage, through phagocytosis, acquires the TSE agent from the FDCs and degrades them, though may be it only partial degradation. This hypothetical process is shown in the subsequent figure from Mabbott and MacPherson (2006). Of interest is the fact that some studies have shown that the small to intermediate particles originating from the larger aggregates are more infectious in nature. Caughey and Baron (2006)

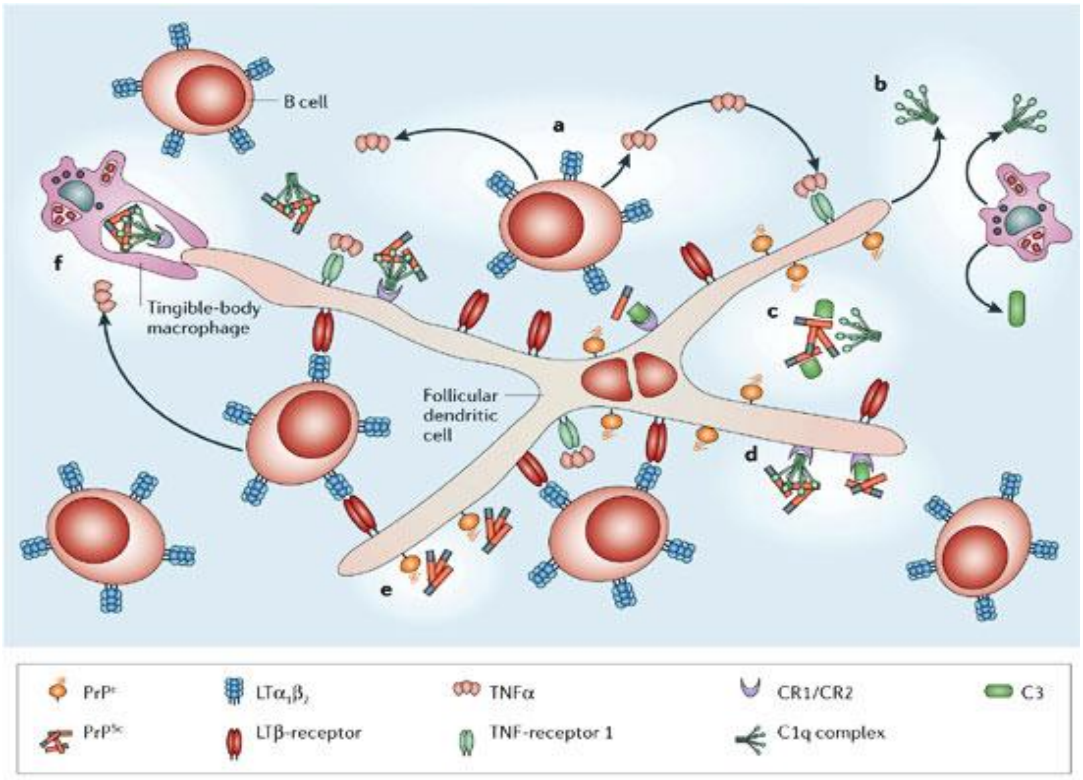


Figure 1-2: A model of the molecular and cellular components of TSE pathogenesis in lymphoid tissue. **a**) Cytokines produced by B cells induce FDC maturation and maintain FDCs in their differentiated state. **b, c)** Deficiency in compliment component impairs TSE accumulation in the spleen and C1q can bind to a conformationally modified form of PrP, current hypothesis suggests soon after exposure TSE agent become opsonized by compliment components. **d)** TSE-agents-compliment complexes could be acquired by compliment receptor (CR1/CR2) on the surface of FDCs. **e)** Cellular PrP^{c} on the surface FDCs also act as receptors for PrP^{sc}. **f)** Tingible-body macrophages within the B cell follicles also acquire TSE agents. (Mabbott and MacPherson, 2006)

A recent study has demonstrated that PrP^{sc} is released in the infectious form in association with exosomes. Exosomes contain many glycosylphosphatidyl inositiol (GPI) anchored proteins, including PrP^{c}. Exosomes may act as an intercellular transfer system for FDCs, DCs or tingible body macrophages to the peripheral nervous system.

Through studies of scrapie orally-inoculated rodents, there are indications that the infectious agent spreads from the gut associated lymphoid tissues (GALT) to the CNS

through the enteric nervous system. As can be observed in Figure 1-3 taken from Mabbott and MacPherson (2006), neuroinvasion can occur through two distinct neuroanatomical pathways. Neuroinvasion can occur via the sympathetic nervous system where it progresses through sympathetic fibres of the splanchnic nerve to the intermediolateral cell column of the mid-thoracic spinal cord. Subsequently, the agent spreads to the brain in a caudal-to-cranial direction along the spinal cord. Alternatively, neuroinvasion can occur through the parasympathetic system. In this case, the agent spreads along parasympathetic fibres of the vagus nerve to the dorsal motor nucleus of the vagus nerve within the medulla oblongata of the brain. Figure 1-3 illustrates this mode of neural transmission from peripheral to CNS domains.

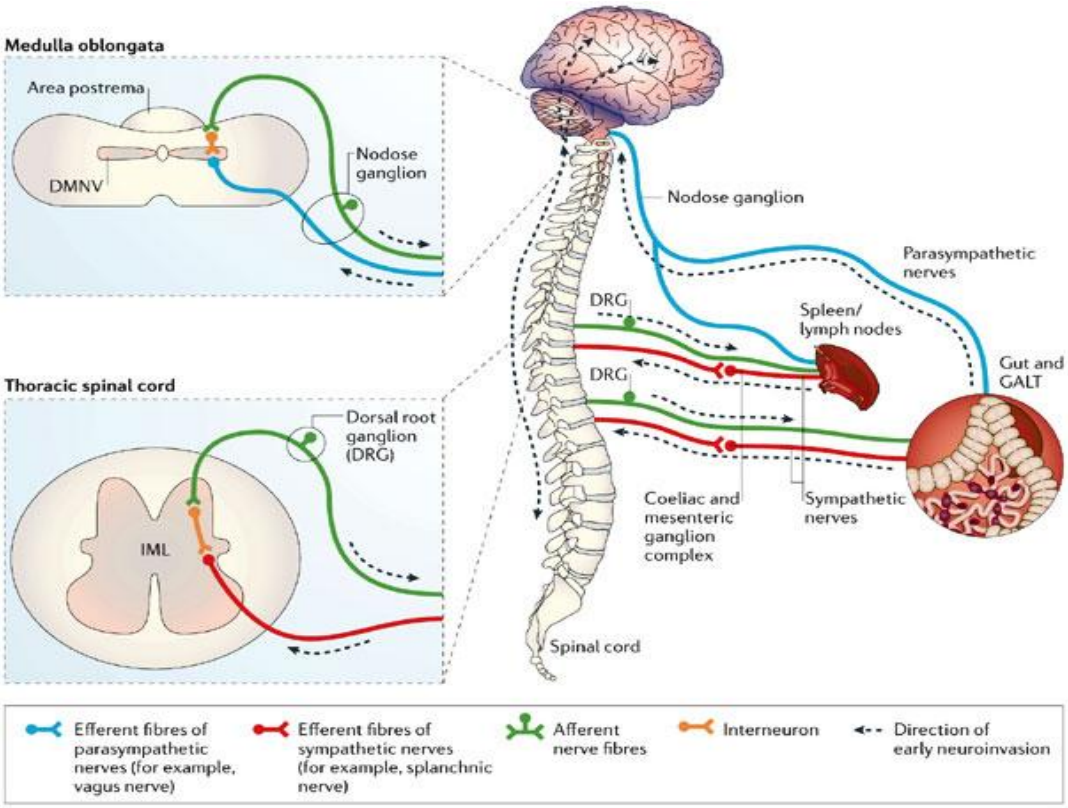


Figure 1-3: Initial pathways of TSE agent neuroinvasion originating from the intestinal lumen. Analysis of the progression of disease-specific prion protein (PrP) accumulation within the nervous system of experimentally inoculated rodents indicates that the scrapies agent spreads from the intestine to the CNS through two distinct neuroanotomical pathways. In each case the initial spread is retrograde direction along efferent (motor) pathways. One route to the CNS occurs along sympathetic fibres of the splanchnic nerve to the intermediolateral cell column of the mid thoracic spinal cord. Subsequently, the agent spreads to the brain in a caudal-to-cranial direction along the spinal cord. Neuroinvasion from the intestine can also occur independently of the spinal cord. In this case, the agent spreads along parasympathetic fibres of the vagus nerve to the dorsal motor nucleus of the vagus nerve within the medulla oblongata of the brain. (Mabbott and MacPherson, 2006)

1.2.3 Classification of human prion disease

Currently the manner by which prion diseases in humans are classified is based on two factors; one criterion is the fragment or banding profile produced in a Western blot and the other is the amino acid residues which occupies position 129 in the protein sequence. In general, these two features have been combined, resulting in a system which is referred to as strain typing. Strain typing has been described by Watts *et al.* (2006) as the classification of conformational sub-varieties. Further, the various strains have been intracerebrally inoculated in mice to establish the inoculation period as well as pathological characteristics of each. Armed with strain type information, research has now discovered some associations between the specific neuropathological characteristics and the specific strain type present (Wadsworth *et al.*, 2003).

1.2.3.1 Strain-typing of Prion protein

Strain typing has been further categorized based on the following criteria: degree of N-glycosylation, and degree of resistance to proteinase K digestion. Western blots are used to analyze the molecular variants by identifying the molecular weight and glycosylation types. Parchi *et al.* (1999) and Hill *et al.* (2003) both established independent classification systems. In an attempt to avoid confusion in what Onodera *et al.* (2006) states is a complex relationship between the molecular type and the prion strain, we will focus on the Parchi *et al.* (1999) classification scheme. Within this classification system, there are two categories: type-1 and type-2 with type-2 having subtypes A and B. In the following example the CJD phenotype is linked to the amino acid Methionine at codon position 129 and classified as type 1. It should be noted that with this banding pattern the smallest fragment observed is 21 kDa. In contrast, in type-2 the smallest fragment is 19 kDa and either Valine or Methionine (in the case of variant CJD) is present at position 129. The type-2 strain can be further segregated based on the different relative amounts of each glycoform as observed in the case of the variant CJD. The following figure of Type-1 and -2 indicate different molecular weights, while type -2A and -2B differ in the relative amount of each glycoform. (Parchi *et al.*, 1999)

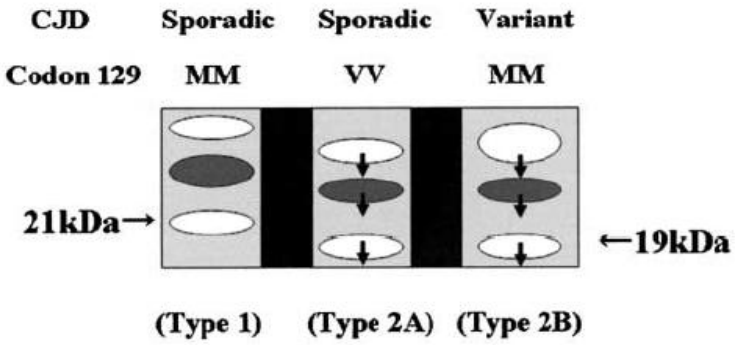


Figure 1-4: Illustration of the molecular weight and glycoform patterns of sporadic and variant CJD. (Onodera *et al.*, 2006) Type-1 and -2 indicate different molecular weights, type -2A and -2B differ in the relative amount of each glycoform. (Parchi *et al.*, 1999)

Truchot (2004) observed a correlation of type-1 with synaptic deposits and a correlation of type-2 in diffuse deposits. In addition, Pouti *et al.* (1999) reported the co-existence of both types in the same patient. It is for this reason that Onodera *et al.* (2006) suggests the strain typing is a quantitative rather than a qualitative process.

In the case of BSE there are two categories in the classification of BSE. The first and most common is classical. This category includes cattle which acquired BSE through oral transmission. This strain profile is similar to that of vCJD (type 2B). The other category is termed atypical and in general considered to be acquired spontaneously. The atypical form has two distinct profiles; H-type (referring to a higher molecular weight band) and L-type (referring to a low molecular weight band).

1.3 Prion biology

1.3.1. PrP protein characteristics

1.3.1.1 Native PrP^c protein

The prion protein is a plasma membrane glycoprotein which is generally well conserved among species with both a similar amino acid sequence and secondary protein

structure. Essential proteins generally show higher levels of conservation across species than less essential proteins or ones that are components of more redundant biological processes or pathways. Therefore, it was initially speculated that the native prion protein must play some essential biological role. However, when *Prnp* knock-out mice were shown to lack any dramatic phenotypic effect as compared to wild-type mice, the high degree of conservation of *Prnp* is a bit more 'evolutionarily perplexing'. The PrP polypeptide is approximately 253 amino acids (a.a.) long, but this is slightly variable and species-dependent, prior to any post-translational modifications. The N-terminal of the mature protein is cleaved in the endoplasmic reticulum (ER) and amino acids (a.a.) 232-253 are modified by a glycosyl-phospotidyl-inositol (GPI) anchor. The GPI anchor allows the prion protein to be targeted and attached to the exterior of cholesterol and sphingolipid rich membrane subdomains termed "lipid rafts". These rafts are detergent resistant subdomains of the plasma membrane phospholipid bilayer. The mature PrP^{C} undergoes a cycle where it is endocytosed, and either shuttles back to the cell surface in the lipid rafts, or is eventually targeted for degradation in lysosomes (Mastrangelo and Westaway, 2001). In the ER, two N-linked oligosaccharide chains are added to a.a. 181 and 197 in humans, and 180 and 197 in the mouse. At this point it is worth noting that although there are two glycosylation sites on the prion protein it has been routinely observed to be present in a mixed state. That is to say mixtures of mono-, di-, and unglycosylated prion protein banding patterns have all been observed in Western blots of proteins extracted from mice and humans. The mature human prion protein consist of 209 a.a. (210 a.a. in mouse), however the final length of the prion protein is species specific.

The prion protein has two domains; the N-terminal domain which is flexible and considered disordered, and the C terminal domain which is stable and ordered. The N-terminal region, a.a. 23 - 125, is generally unstructured although it contains an octa-repeat region. This octa-repeat region has the ability to bind copper, as depicted in Figure 1-5 taken from Caughey and Baron (2006), the significance of which has yet to be fully determined.

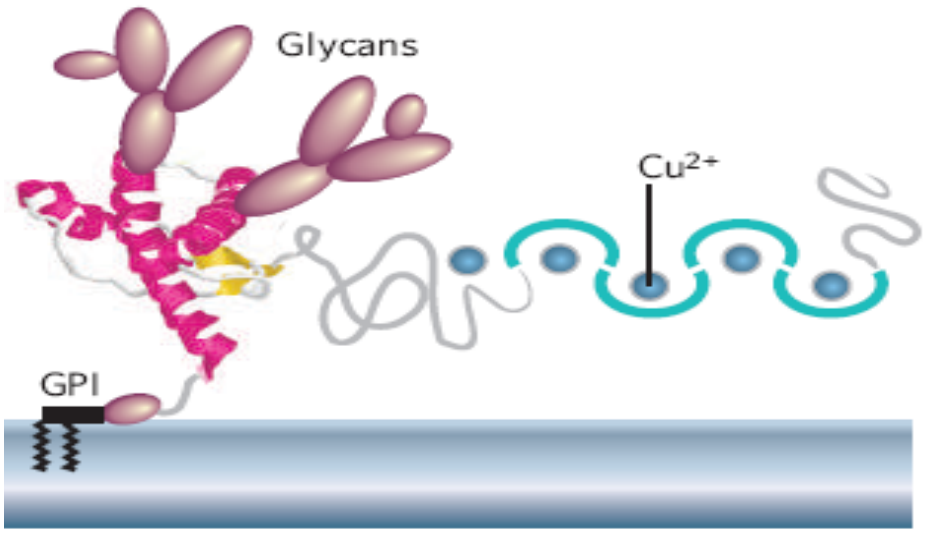


Figure 1-5: Three dimensional view of the structural features of the prion protein. The C terminal structured region, a.a. 126 – 231, has three alpha (α) helixes shown pink coils, and two short beta (ß) sheet domains, shown as yellow ribbons. The α-helixes are stabilized via the formation of disulfide bridges. The octa repeat region is shown in blue. (Caughey and Baron, 2006)

1.3.1.2 Characterization of the misfolded PrP^{res} protein

The prion protein can exist in two isoforms; the native cellular wild type isoform (PrP^{c}) and the pathological isoform (PrP^{sc}). It is important to emphasize that although both of these isoforms have the same amino acid sequence they have a different secondary structure. The secondary structure of the native isoform consists of 42% α-helices and 3% ß-pleated sheets. This isoform is typically soluble in non-denaturing detergents and is sensitive to proteinase K digestion. In contrast; the isoform associated with disease consists of 30% α helices and 43% ß sheet. As a result this "misfolded" structure is generally insoluble and resistant to proteinase K digestion. PrP^{c} is monomeric however PrP^{sc} is multimeric and can assemble into amyloid fibrils (see Figure 1-6). (Caughey and Baron, 2006) One key aspect of the disease etiology is that the presence of the misfolded isoform can convert the native isoform to the misfolded structure. However, this conversion of the wild type prion protein to the disease-

associated isoform has a significant energy barrier. At a neutral pH the α-helixes are favoured, however upon initiation of conversion the kinetics are believed to be rapid (Watt *et al.*, 2006). Given that the pH of the lipid raft, where the native protein is believed to reside, is acidic, it has been hypothesized that this reduces the energy barrier and favours aggregate formation.

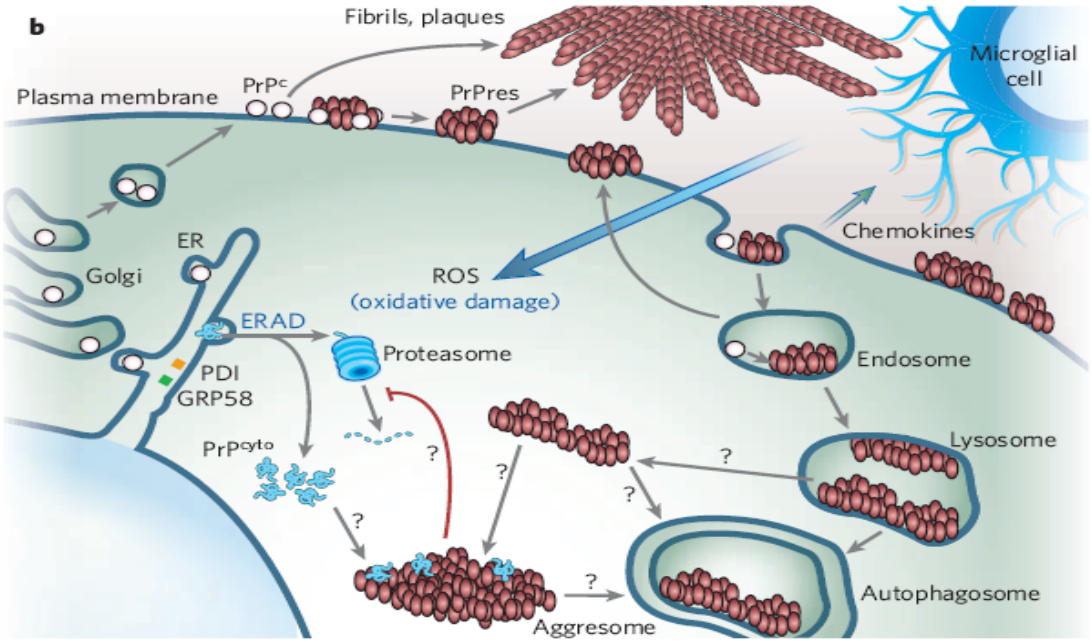


Figure 1-6: The cell biology of PrP^{c}, PrP^{res} and membrane subdomains 'rafts' (hypothesized). Under conditions of proteasome inhibition, cytoplasmic forms of PrP aggregates associated with neurotoxicity (such as PrPcyto and aggresomes). However, the mechanisms by which such aggregates are generated (for instance, the translocation of PrP^{res} from the lumen of endocytic and lysosomal vesicles into cytosolic aggresomes, and the cause of proteasome inhibition) are unknown. The release of reactive oxygen species (ROS) from chemokine-activated microglial cells could contribute to ER stress and/or the ERAD process by inactivation of ER-chaperones. Excessive levels of misfolded proteins in the cytosol might impair proteasome function. PrP^{res} can accumulate on the cell surface, in intracellular vesicles such as lysosomes or autophagosomes, or in extracellular deposits. PrP^{res} and TSE infectivity can be released from cells such as exosomes and infected neuronal cells by incorporating PrP^{res} from the extracellular mechanisms that involves HSPG and LRP/LR. (Caughey and Baron, 2006)

1.3.2. Function of PrP^c

Although the function of the native prion protein is not yet fully understood, it is believed that the fact that it is well-conserved across many species is evidence that it has an important cellular function. Furthermore, its abundance in the brain implies that its primary function is neuronal, although it is also present in leukocytes and haematopoietic stem cells. It has been postulated by Caughey and Baron that, within the neuronal system, prion protein interactions may modulate cell adhesion, homing of neurites as well as serve in synapse formation and survival. They proposed a sort of restructuring or chaperoning function since prion protein is associated with raft membranes, adhesion molecules, signaling pathways and is consistent with assembly of new cellular structures. Examples of this would be the aforementioned neurites, synapses in neurons and follicular dendrites. Figure 1-7 represents neurite outgrowth and synapse formation modulated by PrP^C interactions with NCAM and STI-1. This can ultimately lead to the activation of intracellular signaling pathways. In the case of the NCAM's, which are known to mediate contacts between neurons and other cells, the signaling pathway is mediated by activation of Fyn kinase. It is presumed to function through the receptor-type protein phosphatase-α (RTPTα). PrP^C binds various adhesion and extracellular-matrix molecules, HSPG, laminin and the laminin receptor (LR) and its precursor (LRP), and help to guide growing neurites to their appropriate destinations. Caughey and Baron postulate that these molecules are delivered to the growing tips (growth cones) of neurites through transport vesicles, perhaps as preformed complexes.

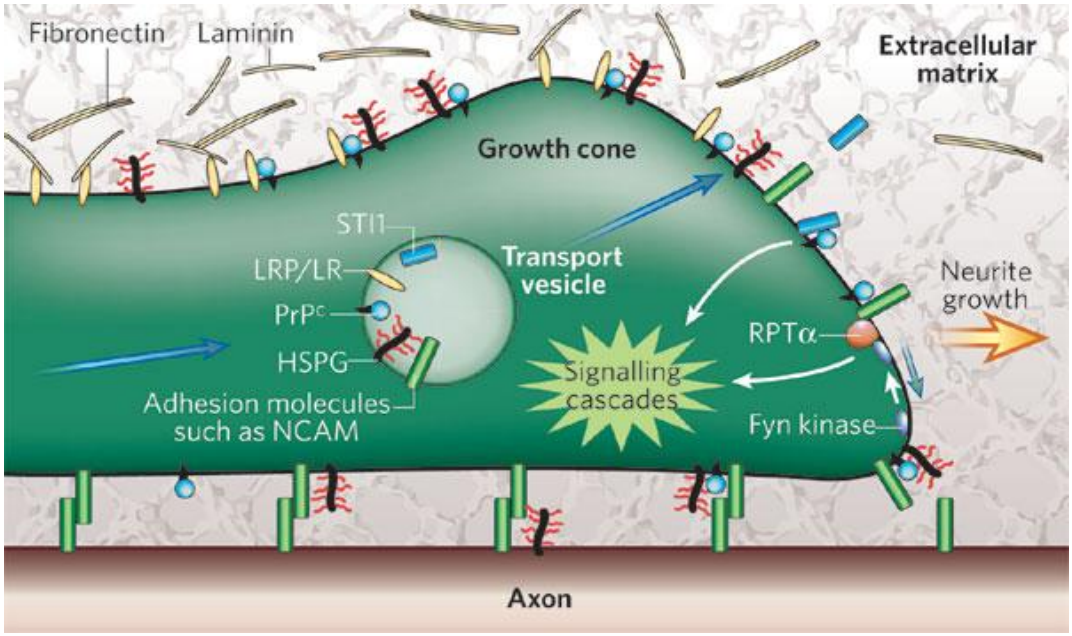


Figure 1-7: Hypothesized model of PrP^c interactions associated with axonal growth. PrP^c seems to be important for neurite (nascent axon and dendrite) growth and synapse formation in neurons. PrP^c interactions with NCAM and STI-1 modulate neurite outgrowth, and activate intracellular signalling pathways. NCAM signalling pathway is mediated by activation of Fyn kinase, presumably via receptor-type protein phosphatase-α (RTPTα). In addition to NCAM, PrP^c binds HSPG, laminin, laminin receptor (LR) and laminin receptor precursor (LRP), which are known to mediate contacts between neurons, other cells and the extracellular matrix, thus helping to guide growing neurites to their appropriate destinations. (Caughey and Baron, 2006)

Although the prion protein can also bind copper, the affinity of this protein for metal ions is low under physiological conditions (although it can be enhanced via glycosaminoglycans). However it may serve to either clear, transport, or as a sensor system for metal ions (Caughey and Baron, 2006) in attempts to control oxidative stress. Both Brown *et al.* (1997) and Shaked *et al.* (2001) agree that the prion protein plays an important neuroprotective role with respect to copper toxicity and oxidative stress. White *et al.* (1999) was able to demonstrate that the prion protein aids in the ability of cells to deal with oxidative stress through the modulation of glutathione reductase activity. In agreement with these hypothesized functions, prion proteins can negatively modulate

phagocytic activity and may aid in the control of inflammatory response or apoptosis of cells in the brain.

Onodera *et al.* (2006) also describes many of the molecules which have been reported to interact with PrP. Table 1-2 contains a list of such molecules, and their implied physiological functions, that have been reported in papers by the respective authors.

Molecules	Physiological function	References
Bcl-2	Anti-apoptosis	Kurschner, C. (1995)
Caveolin 1	Trigger FYN activation in signaling	Mouillen-Richaard, S. (2000)
CK2	Phosphotransferase activity	Meggio, F. (2000)
GAG	Biomolecular transport	Gonzalez-Inlesias (2002)
GFAP	Cell repair	Eng, L.F. (1994)
Grb2	Adaptor proetin	Spielhaupter, C. (2001)
Copper	SOD like activity and anti-oxidant	Brown, D.R. (1999), Kretzschmar, M.A. (2000)
HSP60	Chaperone	Edenhofer, F. (1996)
Laminin	Adhesion, neurogenesis	Graner, E. (2000)
Laminin receptor	Laminin binding	Rieger, R (1997)
N-CAM	Adhesion	Schmitt-Ulms, G. (2001)
Nrf2	Apoptosis inhibitor	Yehiely, F. (1997)
Pint 1	unknown	Spielhaupter, C. (2001)
STI1	Neurogenesis and neuroprotection	Lopes, M.H. (2005)
Synapsin 1b	Regulation of nuerotransmitter	Spielhaupter, C. (2001)
Tubulin	Intracellular transportation	Nieznanski, K. (2005)

Table 1-2: List of molecules with reported ability to interact with cellular prion protein. (Onodera *et al.*, 2006)

Given the delicate nature of the brain, and its cells, where increased inflammation and its downstream effects are often damaging, irreparable and irreversible, the increased expression of the prion protein appears to be consistent with this associated protective function.

1.3.2.1. PrP as an interactive molecule

As a consequence of mouse transgenic work, Prusiner (1997), postulated an interaction between prion proteins and another species-specific protein, termed "protein X". During these experiments it was observed that if a transgenic mouse expresses both human and mouse prion proteins then it was resistant to infectious human prion protein. However if the transgenic mouse expressed only human prion or a chimeric human/mouse protein variant they were susceptible to infection. Prusiner (1997) interpreted these results as evidence that the mouse prion protein is bound by another protein, rendering it unavailable for structure conversion by the human infectious protein. Conversely, the human prion expressed in the transgenic mouse doesn't bind or weakly binds the mouse protein X therefore the prion protein is still available for conversion. This work was followed by failed attempts to isolate and or identify "protein X".

Caughey and Baron (2006) discuss a number of interesting ligands which reportedly have the ability to bind to, and or interact with PrP^c. The characterization of these ligands and their respective function may provide insight as to the function of PrP^c itself in health and disease. However as described by, Caughey and Baron (2006) the overall picture is chaotic. Adding to the confusion is the fact that $Prnp^{0/0}$ mice (see knock-out section to follow) appear to have no overt clinical problems and have a normal life span. Nonetheless, these authors describe reports of alteration in circadian rhythms, hippocampal neuronal function, spatial learning, brain copper and cuproenzyme levels, oxidative tissue damage, phagocytosis and inflammatory response, haematopoietic-stem-cell renewal, neural-stem-cell differentiation and stress response. (Caughey and Baron, 2006) They further suggest that there may be a backup system in place, which at least partially compensates when there is a loss of PrP function, hence masking obvious phenotypic changes in the knock-out mice.

Given the broad array of reported functions and the wide variety of tissues in which prion proteins are expressed, the localized functions may be cell type and physiologically process dependent. PrP ligands such as: sulfate glycosaminoglycans (GAGs) like those found in heparin sulfate proteoglycans (HSPG), extracellular-matrix proteins such as laminin receptor (LR), laminin receptor protein (LRP) and raft membrane lipids, as well as heat shock proteins like STI1 are thought to serve a variety of specific functions. It is interesting to note that most of the known cellular adhesion

ligands, NCAMs, laminins, laminin receptors and HSPGs, and their binding partners are notorious for being polymorphic. The precise developmental and physiological control of expression of the different polymorphs allows their interactions with PrP^{c} to be highly context-dependent and to have diverse functional consequences. (Caughey and Baron, 2006) This hypothesis could be extended to include functional consequences such as variation in disease susceptibility, disease incubation time, and possibly disease resistance.

In addition, some of these ligands have been reported to play important roles in modulating the conversion of PrP^{c} to PrP^{res}. In fact, Caughey and Baron refer to some of these ligands as anti-TSE compounds (see Figure 1-8). Such ligands include the molecules mentioned above as well as phosphorothiolated-oligonucleotides (PS-ON). They speculate that these ligands are able to bind directly to PrP^{c} (and in some cases PrP^{res}) and are capable of assembling a common structural feature, a linear polyanionic chain. It is due to this feature that these authors in Figure 1-8 suggest that the inhibition is as a result of competitive binding of the same or overlapping sites. There are three regions of the prion protein that are implicated as potential binding domains. The first is the amino-terminal which is highly cationic. The second region is the octa-peptide repeat region, capable of binding copper and the third is the carboxy terminus which contains the α-helix (review Figure 1-5). It is hypothesized that these inhibitor sites are also sites of normal binding by physiological ligands. (Caughey and Baron, 2006)

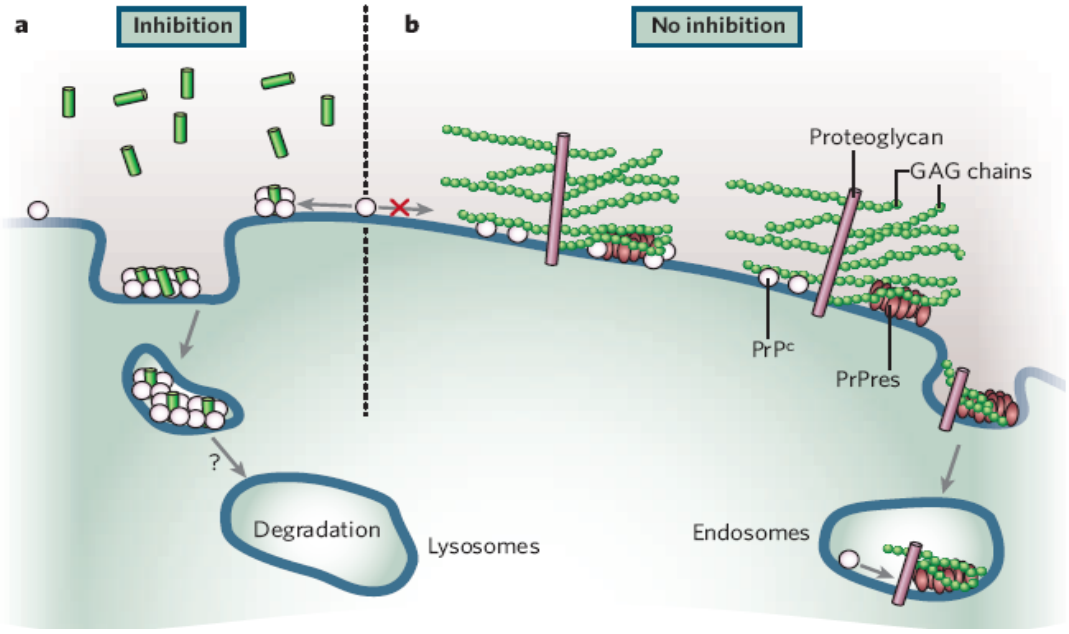


Figure 1-8: Anti TSE compounds sequester PrP^{c} inhibiting PrP^{res} conversion. **a)** This results in the sequestration of PrP^{c} either in a state and/or subcellular location that is incompatible with conversion to PrP^{res}. Some inhibitors may also bind to PrP^{res} in a way that interferes with its ability to convert PrP^{c}. Because several structurally different classes of PrP^{res} inhibitors share properties (like PrP-binding sites and cause PrP aggregation and endocytosis) it raises the question of how these properties might relate to normal function of PrP^{c} and the mechanism of PrP^{res} conversion. Specifically, it may be that these inhibitors bind to site(s) reserved for normal physiological ligands and influences the conversion to PrP^{res}. Prime candidates for this type of ligand are GAGs such as heparin sulphate. **b)** Many of the PrP^{res} inhibitors can be viewed as GAG analogues or mimics including nucleic acids, which, are often linear polyanionic polymers with repeating anionic and hydrophobic surfaces. Interactions with other cellular components such as laminin receptors and membrane rafts might also be affected (directly or indirectly) by inhibitor binding. Finally, various cyclic tetrapyrroles (such as haemin) exhibit strong inhibition. (Caughey and Baron, 2006)

A 3D structural model of the prion protein Figure 1-9 illustrates some of the functional domains of the native protein as well as sites of structural variation and

modification. It is often useful in the quest to discern the function(s) of a novel or uncharacterized protein, to compare the conservation of site-specific domains across several animal species, where possible. This allows the identification of regions that may be critical for biochemical post-translational modification, signal sequences and the establishment of binding domains that may be related to physiological functions. The two distinct sites for glycosylation that were briefly mentioned earlier in this review are sites where the native protein is found with varying degrees of glycosylation. Interestingly, the paper by Caughey and Baron (2006) reported that LRP/LP binding to PrP^{c} maps to residue 144-179 which is also the domain proposed to be involved in binding to PrP^{res}. They hypothesize that dissociation of LRP/LR from PrP^{c} may be necessary to allow the availability of PrP^{c} for conversion to PrP^{res}.

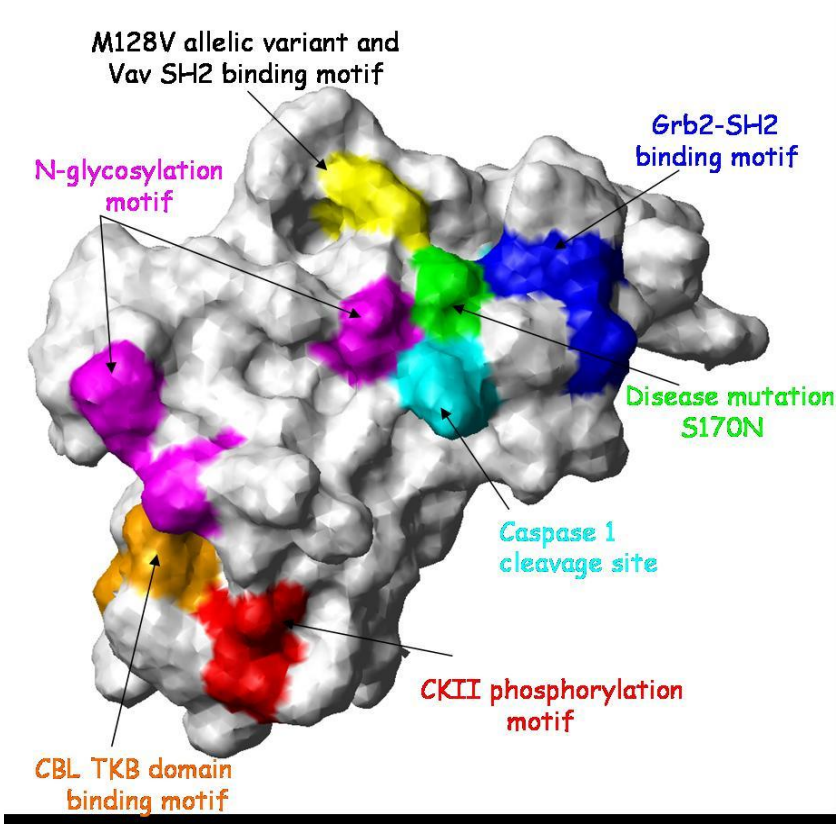


Figure 1-9: Mini-motif prediction of how mutations cause disease in the human prion protein (NP_000302). (Http://sms.engr.uconn.edu/priondetails.htlm)

1.3.3. The prion protein family

In humans, the prion gene is located on human autosome 20 (HSA20). The synteny block of the *PRNP* locus has been cytogenetically identified as HSA20 P12-13 and BTA 13 q17. Choi *et al.* (2006) determined that the synteny boundary was 7.1kb upstream of the *PRNP* gene in humans and 7.8 kb upstream in bovine. Within this locus or domain there are three genes including the prion gene (*PRNP*), the Doppel gene (*PRND*), and testis-specific alternatively spliced transcriptional product (*PRNT*). To date, all of these genes have been mapped within a 55 kb region in humans.

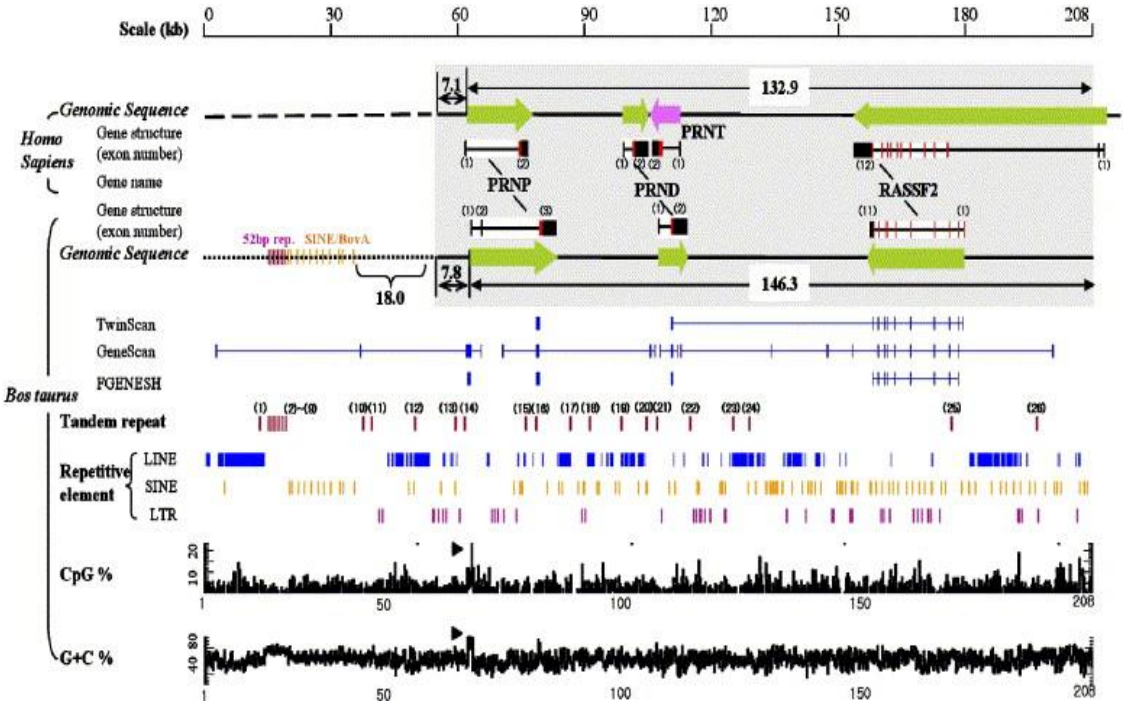


Figure 1-10: Comparative genomic organization of the *PRNP* locus between cattle and human. (Choi *et al.*, 2006)

In humans, the Doppel gene (*PRND)* contains two exons; however, there are three known exons in the same gene of sheep, mice, and cattle. Although the amino acid sequence has only 25% identity with *PRNP* (because it lacks the repeat region and the hydrophobic domain (Rivera-Milla *et al.*, 2006) it is still structurally and topologically similar (Choi *et al.*, 2006). Like PrP, Doppel is a GPI-anchored glycoprotein with three α-helices and two ß-sheets (Rivera-Milla *et al.*, 2006). The Doppel protein (Dpl) is

expressed in many tissues though interestingly it is not expressed at appreciable levels in brain. It is of interest, and perhaps a hint relating to the function of the doppel protein, that its absence in the testis causes sterility (Weissmann and Flechsig, 2003). Of further functional interest is that ectopic expression of Dpl in the brain of *Prnp* $^{0/0}$ mice results in ataxia. This suggests that tissue specific regulation of Dpl expression is important.

Figure 1-11, taken from Mastrangelo, and Westaway (2001) shows the structural similarity of mouse PrP and Dpl proteins. The schematic shows the positions of the predicted features and its relationship to posttranslational modifications and secondary structure (modified from Moore *et al.*, 1999).

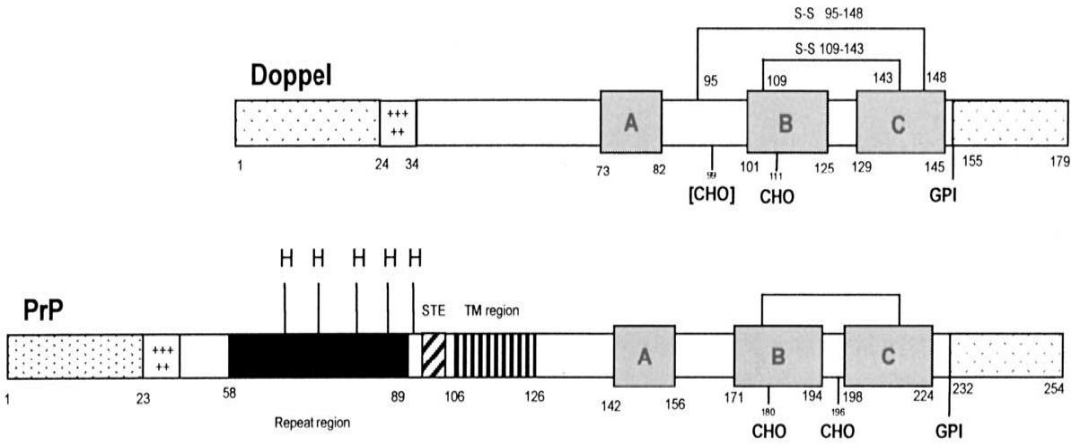


Figure 1-11: Structure of mouse prion and doppel protein. Three α-helices found in PrP and Dpl are shown as gray boxes A, B, and C (Mo *et al.*, 2001). The residue numbers for the predicted disulfide bridges and Asn-linked glycosylation sites (CHO) are also shown. The box with vertical stripes shows the hydrophobic transmembrane region containing the highly-conserved sequence motif AlaGlyAlaAlaAlaAlaGlyAla that is found in all PrPs, and the stop transfer effector region (STE) is represented by a box filled with diagonal stripes. The white box with the '+' symbols indicates a cluster of basic residues. 'H' represents the histidine residues in the repeat region and at position 95 of PrP, these residues are related to the octa-repeat domain that was reported earlier in the context of ability for copper binding. (Mastrangelo and Westaway, 2001)

In addition to *PRNP* and *PRND*, the other gene found within the gene cluster is *PRNT*. The *PRNT* gene is quite complex as it encodes three alternatively spliced

transcripts and is expressed exclusively in the adult testis. The *PRNT* gene is thought to be absent in cattle but it is believed that the *PRND* and *PRNT* genes may be evolutionary and functionally related and, as such, may be somewhat redundant in function such that "loss" or "absence" of this gene in cattle is not deleterious. (Choi *et al.*, 2006) In a more recent publication by Premzl and Gamulin (2007) the *PRNT* is shown to be absent in mouse, rat and cow and these authors then propose that it is closer to *PRND* in evolutionary origin then *PRNP*.

In addition to the three previously mentioned genes, a bovine prion-like Shadoo (*SPRN*) gene has been published. (Uboldi *et al.*, 2006) Due to the fact that *SPRN* is more highly conserved then *PRNP*, it is hypothesized that it might be the ancestor of the prion gene itself and that after a duplication event, there was independent evolution of these genes. Specifically, the model (based on bioinformatics data) proposes that the duplication of *SPRN* gave rise to *SPRNA* and *SPRNB*. A second duplication event of *SPRNB* yielded *SPRNB1* (which eventually resulted in the *PRNP* cluster of *PRNP*, *PRND*, and *PRNT*). *SPRNB2* is observed only in fish and is likely a result of the separate gene duplication event that occurred in teleosts. Through the use of radiation hybrid mapping, the prion-like gene *SPRN* was assigned to BTA26 q23, which corresponds to HSA10q24.3-26.3. Bovine *SPRN* has 74.8% identity and 84.7% similarity to human *PRNP*. Like *PRNP*, *SPRN* has two exons and has the appearance of a housekeeping gene with CpG islands spanning the entire exon one and no TATA-box.

1.4. Prion Molecular Modifications

1.4.1. Prion knock-out in mouse

Weissmann and Flechsig (2003) authored; "PrP knock-out and PrP transgenic mice in prion research". In this paper, Weissmann and Flechsig (2003) describe several lines of transgenic mice which were generated with the use of embryonic stem cells. They incorporated two strategies termed 'conservative' and 'radical' in which the conservative strategy involved disrupting the open reading frame of *Prnp* while the radical strategy involved complete deletion of the open reading frame plus the flanking regions. Mice from which the 'conservative' knock-out strategy was used (i.e. the Zürich

I *Prnp*$^{0/0}$ mouse) appeared to develop and reproduce normally without any significant and notable phenotypic differences from the wild type mice. However, it was noted that, upon ageing, these mice exhibited demyelination of the peripheral nervous system, without exhibiting clinical symptoms. In addition, the authors noted that there were some alterations in circadian activity and sleep rhythm. In contrast the mice which received the 'radical' knock-out strategy, (i.e. the Zürich II *Prnp*$^{-/-}$ mouse), also developed normally, however, later in life these animals exhibited severe ataxia and Purkinje cell loss. It was later discovered that just downstream (16kb) of the *Prnp* gene was *Prnd* gene (encoding the doppel (Dpl) protein). In the wild type mouse, *Prnd* transcription is under the control of its own promoter, however in the 'radical' experimental strategy this promoter was removed. With the *Prnp* promoter still intact, the doppel protein (Dpl) was ectopically expressed in the brain of the *Prnp*$^{-/-}$ mice. *Prnp*$^{0/0}$ mice, which were intracerebrally injected with scrapie prion remained symptom free for at least two years and these mice failed to propagate prions in either their brain or spleen. This was in contrast to the wild type mice (*Prnp*$^{+/+}$) where, following intracerebral injection, these mice exhibited pathological symptoms in 160 days and died at 170 days. Of further interest and importance was the fact that mice, hemizygous for the *Prnp* allele (*Prnp*$^{0/+}$), had a prolonged disease incubation time of 290 days. According to Caughey and Baron (2006), the knock-out of the PrP gene eliminates the endogenous source or substrate necessary for PrP^{res} formation and thus prevents both disease and prion replication. Alternatively, the over expression of PrP^{c} increases the substrate availability and thus the rate of PrP^{res} formation, which shortens the incubation time observed for disease onset (Prusiner, 1997). These experiments strongly support the hypothesis that the disease manifestation is as a result of induced transformation of the native prion protein into the misfolded prion, and that the progression of the disease is proportional to the quantity of the native prion protein present in the host neuronal tissue.

1.5. Prion Genetics

1.5.1. Prion genomic organization

Prusiner (1998) has suggested that the observed species variation in PrP sequences are potentially responsible, at least in part, for the species barrier that has been observed with attempts to induce between-species infections, when prions are introduced from one host species into another. A related argument can be made that the homology between the cattle and human prion proteins is essential for the apparent cross species infectivity of BSE to human vCJD. In fact, in an experiment by Wadsworth *et al.* (2004), it was reported that Tg mice expressing the human PrP 129V form exhibit a significant barrier to infection from BSE or vCJD whereas mice expressing human PrP 129M displayed the typical vCJD infectivity pattern. Further, this is supported with work by Vorberg *et al.* (2003), where they report that rabbits are the only mammalian species to be completely resistant to TSE agents from other species. They attribute this resistance to critical amino acid residues located throughout the rabbit sequence which results in a structure that is unable to refold to the abnormal isoform.

As such, it is important to understand the structure of these genes, other members of this gene family, as well as structures of other important players in its potential "induced folding" pathway. It is through this understanding, as well as understanding of observed variations in the structure, that the role and function of the PrP protein can be fully appreciated. For example the N-terminal region which contains a variable number of repeats, also referred to as the octa-peptide repeat region, is able to bind metal ions and is similar to some other proteins/genes implicated in neurodegenerative disorders (i.e. Alzheimers, Parkinsons and Huntingtons disease). Ideally, transgenic models can highlight potential roles and functions of protein pathways in health and disease through "loss-of-use" mutants. However, when this approach failed, it should not be interpreted as an indication of a lack of an important physiological function or role of the prion protein.

According to Uboldi *et al.* (2005) the *PRNP* genes of human, hamster, sheep, mouse and cattle have a GC rich exon 1 and TATA-less promoters. However, have a CpG-enriched region of about 2.6 kb that contains numerous regulatory DNA elements (Geldermann *et al.*, 2006). The human *PRNP* gene consists of two exons (90 and 2384

bp) and one intron (12,698 bp). The open reading frame (ORF) length is 762 bp in humans and 795 bp in cattle. There is 83 % sequence homology at the nucleotide level and 86% at the amino acid level, which represents a very high degree of homology between these divergent species. The genomic organization of the bovine prion gene (*PRNP*) reveals no specific promoter region. The bovine *PRNP* gene contains three exons (53, 98, and 4091 bp long) and these are separated by two introns (2436 and 13,551 bp long respectively) with exon three being the only translated exon. Figure 1-12 shows a schematic of the bovine *PRNP* gene.

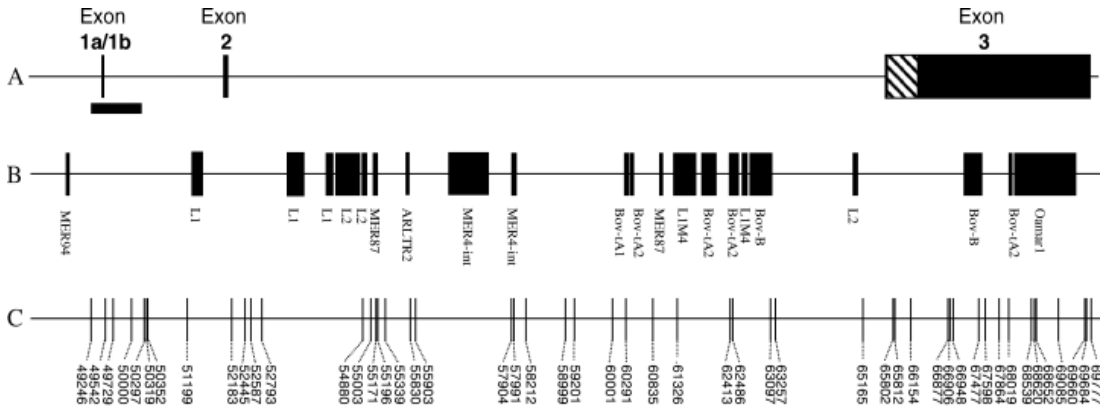


Figure 1-12: Schematic of the bovine prion gene. (Hills *et al.*, 2003) **A)** Shows the three exons and the hatched box illustrates the ORF at the beginning of exon three. **B)** Shows the location of the repeat elements. **C)** Shows the location of the sequence variation identified in Genbank Accession number AJ298878.

The *PRNP* gene in sheep has three exons (52, 98, and 4028 bp) with two introns (2421 and 14031 bp). The ovine PrP, like the other species, has three α-helices and two β-sheets. Helix one spans residues 147-155 followed by the two short β strands, the second helices spans residues 176-197 and the third helices spans residues 203-229 (Bujdoso *et al.*, 2005). The 3'-UTR region of the prion gene in the sheep is 3246 bp longer than that in other species.

The *Prnp* in the mouse also has three exons (47, 98, 2008 bp) with two introns (2190 and 17,733). Two different alleles of the PRNP gene are present in mouse where

the A allele is larger than the B allele by 6.6 kb (Lee *et al.*, 2006). These authors identified 44 inbred strains with A alleles and concluded that the increase in the size of allele A is due to an intracisternal A particle (IAP) element insertion. The IAP element is inserted in the opposite transcriptional orientation to the *Prnp* gene. These IAP- long terminal repeats or IAP-LTR are known to contain regulatory signals for promotion, initiation, and polyadenylation of the transcript. Further, Baybutt and Manson (1997) have shown that the mouse intron 1 contains two distinct sections that promote (10,114-10307) or alternatively suppress (9744-9932) transcription.

Lee *et al.* (2006) identified three regions which are conserved outside the coding region, among human, sheep, and mice and are therefore hypothesized to be functionally important. The first such region is 500bp upstream of exon 2, the second region is at the 3' end of the transcript and the third region described is 3-4Kb downstream of the poly-adenylation site. It is speculated that transcription factor LYF-1 may be a candidate for binding to the promoter region.

Another member of the prion family which is present in cattle and humans is the Doppel gene (*PRND).* The *PRND* gene contains two exons in cattle (89 bp and 3219 bp) (Choi *et al.*, 2006) and in humans is 60bp and 3922bp. The coding sequence starts at the second exon in both cattle and humans and is 537bp and 531 bp, respectively. There is 81.5 % sequence identity at the nucleotide level and 75% identity at the amino acid level between the two species. The promoter region responsible for expression regulation of Dpl in bovine, was identified by Del Vecchi *et al.*, (2005) as a short segment of 355bp. In addition, they found that the expression of Dpl was tightly down regulated in tissue other than testis and suggest a complex tissue-specific regulation of Dpl gene expression. Other than spermatogenesis, the physiological role of the doppel protein, like prion, remains elusive.

Shadoo *SPRN,* the other prion protein family member present in cattle, contains two exons. These exons are separated by an intron of 779 bases in humans and 726 bases in cattle, with the translated sequence fully contained in exon 2. In humans, the entire *SPRN* gene is 3398bp in length with exon 1 starting at 96bp and exon 2 at 3123bp. In comparison, bovine *SPRN* which is only 1434bp in length with exon 1 starting at 111bp and exon 2 starts at 599bp. In bovine *SPRN* encodes a 143 amino acid protein whereas it is 151 and 147 a.a., in human and mouse respectively. *SPRN* is expressed in high levels

within the brain, and lower levels in the testis and lung. As the identification of this protein is relatively new, to date, its' function and physiological role remains unreported (Uboldi *et al.*, 2006).

1.5.2. *PRNP* genetic variation

1.5.2.1. Human *PRNP*

Prusiner (1998) stated that, "more than 20 mutations of the *PRNP* gene are now known to be present in the inherited human prion diseases, and significant genetic linkage has been established for 5 of these mutations." Susceptibility and disease resistance has been associated with a variety of genotypic variation in humans, mouse and sheep. As stated previously, prion diseases in humans can be familial, sporadic or acquired by infection. Each of these forms has a number of mutations associated with different disease phenotypes. At amino acid codon 129 there is variation in the nucleotide sequence which results in the amino acid coding for either a Methionine (M) or a Valine (V). In the Caucasian population, 52% are homozygous MM, 36% are heterozygous MV and 12% are homozygous VV. In addition, there have been associations with the genotype at amino acid codon 129 and Sporadic (sCJD). Further, 95% of sporadic CJD patient who are MM are type 1 (see section on strain typing) and 86% of patient who are VV or MV are type 2 (Gambetti *et al.*, 2003). This suggests that the presence of one or two V allele favours strain type 2. According to Gambetti *et al.*, (2003) there is a significant link that infers that codon 129 acts as a modifier of the disease phenotype in human prion disease. It is worth noting that about 20% of the sCJD cases exhibited both type 1 and type 2 phenotypes and this is thought not to be random but rather genotypically related.

There have been other mutations associated with codon 129 and disease. Familial CJD (fCJD), or genetic (gCJD) is inherited through the germline and not acquire through an infectious source. The most common mutation is a form of familial CJD (fCJD), found in over 90% of the cases, where there is an a.a. substitution at position 200 in addition to there being a M at codon 129. This is referred to as a haplotype and is written as CJD E200K-129M. Other less common haplotypes which have been documented in the Gambetti *et al.*, (2003) paper: CJD E200K-129V, CJD D178N-129V,

CJD V210I-129M, CJD V180I-129M, CJD T183A-129M, CJD M232R-129M. As well as point mutations, there has been fCJD disease associated with insert mutations in the *PRNP* gene. For example there have been 32 families that have disease associated with 1-9 (never 3) extra 24 bp repeats. Genetic fatal familial insomnia (FFI) is a result of *PRNP* D178N mutation which is cis to a 129M polymorphism. This haplotype, D178N-129M, is the most common (25 pedigree and five additional subjects) in FFI cases.

The table below is comprised of a list of sequence variations in the human prion gene and the translated amino acid. The sequences with the superscript 1 denote spontaneous de-amination of methylated C in CpG and may account for enhanced frequency since replacement of methylated cytosine residues would be expected to enhance expression. The codon with the superscript 2 denotes an acidic to basic amino acid substitution. The codon with the superscript 3 is the polymorphism that has been reported to modulate disease susceptibility and outcome. The codon with the superscript 4 is cleaved off in mature prion in forming lipid PGI anchor. The codon with the superscript 5 results in a termination signal. The codon with the superscript 6 is a silent mutation in third codon position.

Initial	Final	Initial	Final AA	Codon	Reference
aagCCGagt[1]	CTG	Pro	Leu	P102L	Hsiao (1989) Nature 338:342.
aagCCAaaa	CTA	Pro	Leu	P105L	Kitamoto (1993) BBRC 191:709.
gcaGCAagc	GTA	Ala	Val	A117V	Doh-Ura (1989) BBRC 163:974.
gcaGCAagc[6]	GCG	Ala	Ala	A117A	Hsiao (1989) Nature 338:342
gggGGCctt	GGT?	Gly	Gly	G124G	Prusiner (1997) Science 278:245
tacATGctg	GTG	Met	Val	M129V[3]	Doh-Ura (1989) BBRC 163:974.
gacTATgag[5]	TAG	Tyr	stop	Y145S	Ghetti (1996) PNAS 93:744
caaGTGtac[6]	GTA?	Val	Val	V161V	Prusiner (1997) Science 278:245
agcAACcag	AGC	Asn	Ser	N171S[5]	Samaia (1997)Nature 390:241
cacGACtgc[1]	AAC	Asp	Asn	D178N	Goldfarb (1991) Lancet 337:425
tgcGTCaat[1]	ATC	Val	Ile	V180I	Kitamoto (1993) BBRTC 191:709.
atcACAatc	GCA	Thr	Ala	T183A	Ann.Neurol (1997)42:138
aacTTCacc	TCC	Phe	Ser	F198S	Hsiao (1992) Nature Genet.1:68
accGAGacc[1]	AAG	Glu	Lys	E200K	Goldgaber (1989)Exp.Neurol.106:204.
gagCGCgtg[1]	CAG	Arg	His	R208H	Mastrianni (1996) Neurol.47:1305
gtgGTTgag	ATT	Val	Ile	V210I	Pocchiari (1993) Ann. Neurol.34:802.
accCAGtac	CGG	Gln	Arg	Q217R	Hsiao (1992) Nature Genet. 1:68.
tacGAGagg[1]	AAG	Glu	Lys	E219K[2]	Barbanti (1996) Neurobiology 47:734
agcATGgtc	AGG	Met	Arg	M232R[4]	Kitamoto (1993) BBRC 191:709.

Table 1-3: Summary of currently identified sequence variations in the human prion gene. Superscript 1 denotes spontaneous de-amination of methylated C in CpG and may account for enhanced frequency since replacement of methylated cytosine residues. Superscript 2 denotes an acidic to basic amino acid substitution. Superscript 3 is the polymorphism that has been reported to modulate disease susceptibility and outcome. Superscript 4 is cleaved off in mature prion in forming lipid PGI anchor. Superscript 5 results in a termination signal. Superscript 6 is a silent mutation in third codon position. (http://www.mad-cow.org/prion_point_mutations.html)

Variant CJD (vCJD) is found in patients who have been orally exposed to the BSE prion agent. In addition, all clinical cases of vCJD have at least one methionine allele at codon 129 and in most cases are homozygous for methionine. The etiologies of

vCJD patient have the same disease characteristics as cattle with BSE. Although these reports of disease associations with genotypes are just that, association reports, it is not unreasonable to conclude that a Valine genotype may confer resistance.

In Figure 1-13 (modified from Prusiner 1997) the human PrP sequence is shown schematically. Here the five octa-repeats (in the grey boxes), the H1 through H4 as well as the three α helices A, B, and C, and the two ß strands S1 and S2 can be observed. In addition the mutations causing inherited human prion disease and polymorphisms in human, mouse, and sheep can be observed. The mutations shown above the line of the human sequence are mutations that cause prion disease. The mutation shown below the lines are polymorphisms, some but not all of which are known to influence the phenotype of disease. The X axis for mouse and sheep also have polymorphism and the parentheses indicate the corresponding human codon.

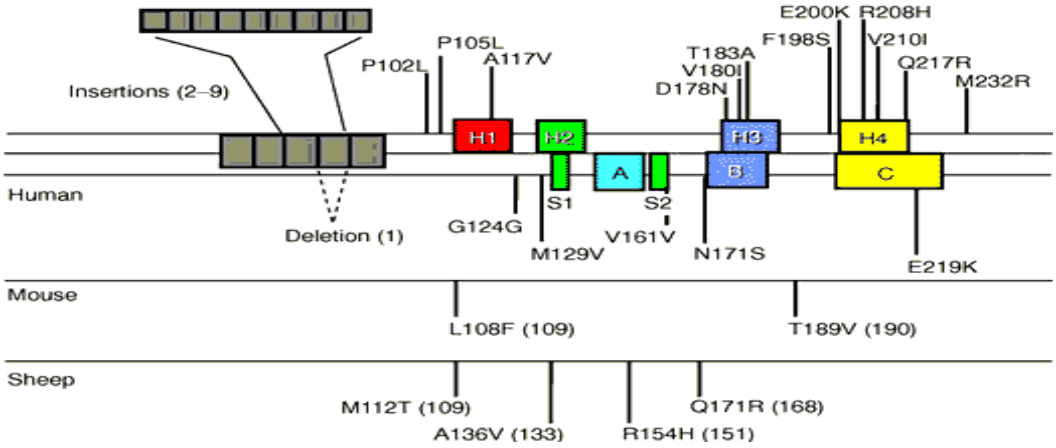


Figure 1-13: Schematic of the human PrP sequence. Here the five octa-repeats (in the grey boxes), the H1 through H4 as well as the three α helices A, B, and C, and the two ß strands S1 and S2 can be observed. In addition the mutations causing inherited human prion disease and polymorphisms in human, mouse, and sheep can be observed. (Prusiner, 1997)

1.5.2.2. Sheep *PRNP*

Differences in susceptibility to natural scrapie have been associated with substitutions at amino acid residue 136, 154, and 171. Specifically, animals with the codon haplotype V at 136, R at 154 and Q at 171 (VRQ) or A136 R154 Q 171 (ARQ) have been associated with susceptibility. Conversely A136, R154, R171 (ARR) has been associated with resistance (Alvarez *et al.*, 2005). Animals with susceptible genotypes which contract scrapie have PrP^{sc} deposits in the medulla oblongata and are referred to as typical cases. Conversely, atypical cases are seen in the occurrences of scrapie in supposedly resistant genotypes. The Nor98, named after a Norwegian isolate, is an example of atypical scrapie. In addition to the resistant genotype, the protease resistant PrP results in a shorter fragment, 12kDa, that is deposited in the cerebral cortex (Watts *et al.*, 2006). This isolate has been successfully transmitted to "ovinized" mice expressing PrP^{VRQ} and still produces the differentiating 12kDa fragment.

1.5.2.3. Cattle *PRNP*

Although there are known polymorphisms in PrP coding regions in cattle, unlike humans and sheep none of these known polymorphisms has been shown to influence disease susceptibility or resistance to date. In addition, there has not been any reported association with the octapeptide repeat regions of the prion gene and BSE. Hills *et al.* (2001) re-sequenced the *PRNP* region in order to identify sequence variation that might affect BSE susceptibility and/or resistance and they identified 42 single nucleotide polymorphisms and nine insertion deletion (including a 12bp insertion deletion in intron one) but did not identify any associations with disease. In a paper by Sander *et al.* (2004), there was a report of a tentative association between prion promoter polymorphisms and BSE susceptibility in German Cattle. They reported an association with two insertion/deletion (indel) sites, the first of which was a 23 bp indel polymorphism containing a repression factor RP-58 binding site. The second was a 12 bp indel in intron one which contained a SP-1 binding site. They, and others (Haase *et al.*, 2007), hypothesized that the polymorphisms in the promoter region cause changes in expression which result in variations in susceptibility and/or resistance. In a follow up paper by Sander *et al.* (2005), they used a reporter gene assay to examine PrP expression levels. They observed higher expression levels with the del/del allele compared to the ins/ins allele. In addition they noted that the frequency of the del/del haplotype was

higher in BSE affected groups. They proposed that the del/del genotype is associated with susceptibility to BSE in German cattle. They also proposed that single nucleotide polymorphisms within the putative promoter region of *PRNP* are in strong linkage disequilibrium with the 23 bp indel and that they presume that one of the polymorphisms might have an influence on the promoter activity leading to modified susceptibility.

Haase *et al.* (2007) examined the indel polymorphisms in several different cattle breeds and made a similar conclusion to that of Sander *et al.* (2004); that the deletion alleles confer a higher risk of developing BSE. The latter paper did, however cite a few differences and exceptions. They identified that there was a breed exception since they did not observe any association with either indels in Brown breeds (Swiss Brown and German Brown). In addition, they did not always find significance with both indels. They concluded that the effect of the 23bp indel was stronger than that of the 12bp indel. Conversely, Juling *et al.* (2006) report a stronger association with the 12bp deletion indel in UK Holstein and German breeds (Holstein and Brown). It is apparent that there is a need for subsequent study of these domains to reconcile these differences in observations. In addition to the expression work completed by Sander *et al.* (2005), Kaskevich *et al.*, (2007) reported their own results of an independent expression study. They suggest that the main effect of the statistical variation that they observed was as a consequence of the 12bp indel. However, they did not believe that the 12bp indel itself is always associated with an increased expression level of *PRNP*. Rather "this study demonstrates that the 12-bp indel and its flanking SNP's are important for the *PRNP* expression." Finally this paper describes a specific haplotype which includes the 12bp insertion allele as having low expression in comparison to the other haplotypes they examined.

Geldermann *et al.*, (2006) also reported an association with the 12bp del indel and the incidence of BSE. This study found that although there was statistical significance across all of the breeds (Simmental, Black Holstein, Red Holstein and Brown Swiss) the largest differences with incidence of BSE were observed in Brown Swiss. They also examined two microsatellite markers, one in intron one (R16) and then other in intron two (R18). Across breeds they observed that the 12bp del-R18 173bp haplotype (highest frequency) had an association with BSE incidence and the 12bp ins-R18 175bp haplotype (second highest frequency) had an association with control animals. They suggested that these two haplotypes might harbour causative DNA variant(s).

The fact that this region of the prion gene is in extensive linkage disequilibrium, as reported by Hasse *et al.*, (2007) as well as Clawson *et al.*, (2006) makes it difficult to separate the effects of one indel from the other, in addition to the effects of single nucleotide polymorphisms. Inoue *et al.*, (1997) identified that the bovine prion protein gene requires interaction between the promoter and the intron but, given the diversity of genotypes as well as the variation of the expression results, it appears that there is still a lot of work remaining to be done.

If one accepts the argument that an increase in PrP expression results in an increase in disease susceptibility, then it is reasonable to expect that any factors that are known to increase expression should be present in the most susceptible genotypes. Based on this assumption, the absence of a repression factor (RP-58) should lead to an increase in expression as observed in the del/del genotype. However, this genotype is also missing the proposed activator, SP-1. Therefore, if this one is taken to the next logical step, then the highest expression level should be the Del/Ins allele. Indeed, in some experiments by Sander *et al.*, (2004) this has been observed to be the case such as in the KOP cell line expression results. However, it was not substantiated in their PT cell expression results, which failed to identify a similar expression pattern. Given this genotype (del/ins) is very rare and affects the interactive repressor site (23bp del) it is difficult to disentangle the single genotypic expression effects. Of course, there is always the possibility that although the 12bp insertion contains an SP-1 binding site, it is not SP-1 that is always binding there. It is possible that this site may be able to bind the activator SP-1 as well as other repressor(s) like SP-3 or 4. This may aid in explaining the expression variability observed. It is also possible, and even probable, that the site of the 12bp insertion, in and of itself is not the only driving force of the expression. Instead the site may work in concert with sequences on either side of it to exert its effects.

There have been studies of TSE transmission in mice which have shown that genetic loci other than the *PRNP* gene are likely to be involved in the genetic control of prion disease. (Lloyd *et al.*, 2001, Moreno *et al.*, 2003) A broader perspective was undertaken in work completed by Hernandez-Sanchez *et al.* (2002). In this paper, they looked at the entire genome to examine the genetic basis for incidence of BSE in offspring of four Holstein sires. In their examination of 166 microsatellite markers genome wide, they identified three chromosomal regions of interest located on BTA 5,

10, 20 respectively. The marker on chromosome 10, INRA107, was in a similar location comparatively to an area previously identified in the mouse. There are two genes in that area which the authors identified as candidate genes. One of the genes identified is HEX A encodes an enzyme hexosaminidase A and has been implicated in Tay-Sachs disease. The second gene identified is P450 family XIX (CYP19) and this gene has also been implicated in several human disorders. The authors were unable to identify any candidate gene on the other two chromosomes, BTA 5, and 20. Following this work Zhang *et al.* (2004) used the Holstein pedigree to perform a quantitative trail loci (QTL) analysis. Here they found two regions, one on BTA 17 and one on the X/Y pseudo autosomal region at the genome-wide significance level. In addition, they report four genome-wide suggestive regions (BTA 1, 6, 13, and 19). Although the prion gene is on chromosome 13 they do not report a QTL in the same region as the prion gene, although, there is a discrepancy in the maps used. Zhang *et al.* (2004) suggested a candidate gene (NF 1) for the chromosome 19 region. Geldermann *et al.* (2006), in addition to the indel work, report an association with BSE incidence with a microsatellite marker (RM222) allele on chromosome 19 as well as with a genotype of marker DIK4009. They specifically examined three polymorphisms in the NF1 region; however, the RM222 which showed significance is about 10cM away from NF1. These authors were able to identify a cluster of genes in the area around RM222 and DIK4009 which are all plausible candidate genes for susceptibility/resistance. They include CLAPB1/AP2B1 (clathrin-associated/ assembly adaptor protein)/(adapter-related protein complex 2, beta 1 subunit), MCP1/CCL2 (monocyte chemotactic protein 1)/(chemokine (C-C motif) ligand2), NF1 (neurofibromin 1), CRYBA1 (crystalline, beta polypeptide 1), TP53 (tumor protein 53), CHRNB1 (cholinergic receptor, nicotinic, beta 1, muscle), OGCP/SLC25A11 (mitochondrial 2-oxoglutarate/malate carrier protein)/(solute carrier family 25, member 11), ASP/ASPA (aspartoacylase), MDCR (Miller-Dieker syndrome chromosome region), 12-LOX/ALOX12 (12-lipoxygenase)/(arachidonate 12 lipoxygenase), RCV1 (recoverin).

1.6 Future Direction of Prion Work

Although great strides have been made over the last 20 years in prion research there is still much to do. For one thing, it is important to gain a better understanding of

the normal physiological role(s) and function(s) of the prion protein itself. As we understand more about the innate function of the endogenous prion protein and its biological role, better decisions relating to proactive selection for or against the cognate expression as a tool for reducing prion disease susceptibility, severity or even incidence may become apparent. It is of further importance to understand how genetic variations i.e. nucleotide polymorphisms in the *PRNP* gene as well as its interactive partners can lead to altered gene(s) function and expression. Again this will provide a mechanism for integration and utilization of identified associations of various genes, including specific sequence information, with variation in susceptibility to various prion diseases. Therefore a greater understanding of the regulation of the prion gene itself, as well as how genetic variations within identified disease-associated regions are related, will prove to be invaluable in the effort to better control the prevalence and dissemination of this disease. More specifically, with association studies of genome wide single nucleotide polymorphisms both as a single marker as well as the haplotype level it may be possible to determine a relationship between the manifestation and non-manifestation of prion disease. This in turn may lead to a greater capacity to elucidate both the normal role of the prion protein as well as the initiation process and progression of the prion disease.

It is equally important to gain a better understanding of the nature of TSE transmissibility. The exact mechanism of how TSE infectivity is conferred is something that still eludes the research community. Although there are many theories, the specific details have yet to be rigorously evaluated and determined with any high degree of confidence. It is noted that the accumulation of TSE is inhibited by the absence of the spleen and complement components (Mabbott and MacPherson, 2006). This implies that oral transmission of TSE disease requires a biological pathway from GI to CNS and therefore disruption of this pathway affects the onset of disease. This coupled with the fact that there is a large variation in disease incubation times further implies that there may be variations in the process of disease onset. Some researchers have pointed out that one major difference between PrP^c and other amyloidogenic polypeptides is its GPI anchor and hypothesize that the GPI anchor is the key to the unique transmissibility of TSE diseases (Caughey and Baron, 2006). This area of research may provide insight into methods by which to reduce or interrupt routes of oral transmission of prion diseases that

may include interventions as simple as modified handling of feedstuff to pharmaceutical development.

Another important focus of future prion research will need to include better biochemical and physiological characterization of the prion protein folding. It is evident that much will be gained by the elucidation of the exact mechanism involved in the transformation of the normal prion protein folding to the protease resistant folded protein. It has been pointed out that the refolding mechanism requires surmounting of a substantial energy barrier at neutral pH, however, the more acidic pH of lipid rafts may aid in this conversion process (Caughey and Baron, 2006). It may be possible to influence the conversion rate through manipulation and imposition of energy barriers that favour the cognate protein structure. Furthermore this research focus may yield potential mechanisms for reversing or reducing this transformation and as such may yield treatments or even cures for the diseases associated with prions in both animal and human species.

Overall a better understanding of the nature of prion proteins with their assortment of genetic variations, how they function normally and in a disease state, would not only aid in the discovery of therapeutic treatments but also in the prevention and dissemination of the disease. It is essential to evaluate the implications that may be associated with selection of perceived resistance to prion disease. For instance, the selection of animals for perceived resistance may effectively lead to the delay in prion disease development is simply an extension of the incubation time and symptom development then it is possible that this may impede the effectiveness of our identification of new or existing cases.

1.7. References

Alvarez, L., Arranz, J.J., and San Primitivo, F. (2005) Identification of a new leucine haplotype (ALQ) in codon 154 in the ovine prion protein gene in Spanish sheep. *J. Anim. Sci.* **84,** 259-265.

Baybutt, H., and Manson, J. (1997) Characterisation of two promoters for prion protein (PrP) gene expression in neuronal cells. *Gene* **184,** 125-131.

Brown, D.R., Qin, K., Herms, J.W., Madlung, A., Madson, J., Strome, R., Fraser, P.E., Kruck, T., von Bohlen, A., Schulz-Schaeffer, W., Giese A., Westaway, D., and Kretzschmar H. (1997) The cellular prion protein binds copper in vivo. *Nature* **390,** 684-687.

Bujdoso, R., Burke, D.F., and Thackray, A.M. (2005) Structural differences between allelic variants of the ovine prion protein revealed by molecular dynamics simulations. *Proteins* **61,** 840-849.

Caughey, B., and Baron, G.S. (2006) Prions and their partners in crime. *Nature* Oct **443(7113)**, 803-810.

Choi, S.H., Kim, I.C., Kim, D.S., Kim, D.W. Chae, S.H., Choi, H.H., Choi, I., Yeo, Y.S., Song, M.N., and Park, H.S. (2006) Comparative genomic organization of human and bovine PRNP locus. *Genomics* **87,** 598-607.

Clawson, M.L., Heaton, M.P., Keele, J.W., Smith, T. P., Harhay, G., and Laegreid, W.W. (2006) Prion gene haplotypes of U.S. cattle. *BMC Genetics* **7,** 51

Del Vecchi, I., Azzalin, A., Guidi, E., Amati, G., Caramori,T., Uboldi, C., Comincini, S., and Ferretti, L. (2005) Functional mapping of the bovine doppel gene promoter region. *Gene* **356,** 101-108.

Gajudsek, D.C., Gibbs, C.J., and Alpers, M. (1966) Experimental transmission of kuru-like syndrome to chimpanzees. *Nature* **209**, 794-796.

Gambetti, P., Kong, Q., Zou, W., Parchi, P., and Chen, S.G. (2003) Sporadic and familial CJD: classification and characterization. *Brit. Med. Bulletin* **66,** 213-239.

Geldermann, H., He, H., Bopal, P., Bartenschlager, H., and Prue, B. (2006) Comparison of DNA variant in the PRNP and NF1 regions between bovine spongiform encephalopathy and control cattle. *Anim Genet* **37,** 469-474.

Haase, B., Doherr, M., Seuberlich, T., Drogemuller, C., Dolf, G., Nicken, P., Schiebel, K., Ziegler, U., Groschup, M., Zurbriggen, A., and Leeb, T. (2007) PRNP promoter polymorphisms are associated with BSE susceptibility in Swiss and German cattle. *BMC Genet.* **8,** 15.

Hernandez-Sanchez, J., Waddington, D., Wiener, P., Haley, C.S., and Williams, J.L. (2002) Genome-wide search for markers associated with bovine spongiform encephalopathy. *Mamm Genome* **13,** 164-168.

Hill, A.F., Joiner, S., Wadsworth, J.D., Sidle, K.C., Bell, J.E., Budka, H., Ironside,J.W., and Collinge, J. (2003) Molecular classification of sporadic Creutzfeldt-Jacob disease. *Brain* **126**, 1333-1346.

Hills, D., Cominicini, S., Schlaepfer, J., Dolf, G., Ferretti, L., and Williams, J.L. (2001) Complete genomic sequence of the bovine prion gene (PRNP) and polymorphism in its promoter. *Animal Genetics* **32**, 231-232.

Hills, D., Schlaepfer, J., Comincini, S., MacLean, I., Dolf, G., Ferretti, L., Olsaker, I., and Williams, J.L. (2003) Sequence variation in the bovine and ovine PRNP genes. *Anim. Genetics* **34,**183-190.

Inoue, S., Tankaka, M., Horiuchi, M., Ishiguro, N., Shinagawa, M. (1997) Characterization of the bovine prion protein gene: the expression requires interaction between the promoter and intron. *J. Vet. Med. Sci.* **59,** 175-183.

Johnson, C.J., Phillips, K.E., Schramm, P.T., McKenzie, D., Aiken, J.M. and Pederson, J.A. (2006) Prions adhere to soil minerals and remain infectious. *PLoS Pathog.* **2(4)**,e32

Juling, K., Schwarzenbacher, H., Williams, J.L., and Fries, R. (2006) A major genetic component of BSE susceptibility. *BMC Biol.* **4,** 33.

Kashkevich, K., Humeny, A., Zielger, U., Groschup, M.H., Nicken, P., Leeb, T., Fischer, C., Becker, C., and Schiebel, K. (2007) Functional relevance of DNA polymorphisms within the the promoter region of the prion gene and their association to BSE infection. *FASEB* **7,** 1547-1555.

Lee, I.Y., Westaway, D., Smit, A., Wang, K., Seto, J., Chen, L., Acharya, C., Ankener, M., Baskin, D., Cooper, C., Yso, H., Prusiner, S.B., and Hood, L.E. (2006) Complete genomic sequence and analysis of the prion protein gene from three mammalian species. *Genome Res*. **8,** 1022-1037.

Lloyd , SE., Onwuazor ON, Beck JA, Mallinson G, Farrall M, Targonski P, Collinge J, Fisher EM. (2001) Identification of multiple quantitative trait loci linked to prion disease incubation period in mice. *Proc Natl Acad Sci* **98(11),** 6279-6283.

Mabbott, N. and MacPherson, G. (2006) Prions and their lethal journey to the brain. *Nat Rev. Microbiol.* Mar **4(3),** 201-211.

Mastrangelo,P. and Westaway, D. (2001) Biology of the prion gene complex. *Biochem Cell Biol.* **79,** 613-628.

Mo, H., Moore, R.C., Cohen, F.E., Westaway, D., Prusiner, S.B., Wright, P.E., and Dyson, H.J. (2001) Two different neurodegenerative disease caused by proteins with similar structures. *Proc Natl Acad Sci* **98(5),** 2352-2357.

Moore, R.C, Lee, I.Y, Silverman, G.L., Harrison, P.M., Strome, R., Heinrich, C., Karunaratne, A., Pastemak, S.H., Chishti, M.A., Liang, Y., Mastrangelo, P., Wang, K., Smit, A.F., Katamin, S., Carlson, G.A., Cohen, F.E., Prusiner, S.B., Melton, D.W., Tremblay, P., Hood, L.E., and Westaway, D. (1999) Ataxia in prion protein (PrP)-deficient mice is associated with upregulation of the novel PrP-like protein doppel. *J. Mol. Biol.* **292(4)**, 797-817.

Moreno CR, Lantier F, Lantier I, Sarradin P, Elsen JM: Detection of new quantitative trait Loci for susceptibility to transmissible spongiform encephalopathies in mice. *Genetics* 2003, 165(**4**):2085-91.

Onodera, T., Sakudo, A., Wu, G., and Saeki, K. (2006) Bovine spongiform encephalopathy in Japan: history and recent studies on oxidative stress in prion disease. *Microbiol. Immunol.* **50,** 565-578.

Parchi, P., Castellani, R., Capllari, S., Ghetti, B., Young,K., Chen, S.G. Farlow, M., Dickson. D.W., Sima, A.A., Trojanowski, J.Q., Peterson, R.B., and Gambetti, P. (1996) Molecular basis of phenotypic variability in sporadic Creutzfeldt-Jakob disease. *Ann. Neurol.* **39**, 767-778.

Pouti, G., Giaccone, G., Rossi, G., Canciani, B., Bugiani, O., and Tagliavini, F. (1999) Sporadic Creutzfeldt-Jakob disease: co-occurrent of different types of PrP(Sc) in the same brain. *Neurology* **53,** 2173-2176.

Premzl, M. and Gamulin, V. (2007) Comparative genomic analysis of prion genes. *BMC Genom.* **8**, 1.

Prusiner, S.B. (1997) Prion diseases and the BSE crisis. *Science* **278,** 245-251.

Prusiner, S.B. (1998) Prions *Proc. Natl. Acad. Sci.* **95,** 13363-13383.

River-Milla, E., Stuermer, C.A., and Malaga-Trillo, E. (2003) An evolutionary basis for scrapies disease: identification of a fish prion mRNA. *Trend Genet.* **19,** 72-75.

Sander, P., Hamann, H., Drogmuller, C., Kaschkevich, K., Schiebel, K., and Leeb, T. (2005) Bovine prion protein gene (PRNP) promoter polymorphisms modulate PRNP expression and may be responsible for differences in bovine sponigform encephalopathy susceptibility. *J. Biol Chem* **280,** 37408-37414.

Sander, P., Hamann, H., Pfeiffer, I., Wemheuer, W., Brenig, B., Groschup, M.H., Ziegler, U., Distl, O., and Leeb, T. (2004) Analysis of sequence variability of the bovine prion protein gene (PRNP) in German cattle breeds. *Neurogenetics* **5,** 19-25.

Shaked, G.M., Meiner,Z., Avaham. I., Taraboulos, A., and Gabizon, R. (2001) Reconstitution of prion infectivity from solubilized protease-resistant PrP and nonprotein components of prion rods. *J. Biol. Chem.* **276,** 14324-14328**.**

Tateishi, J., Koga, M., and Mori, R. (1981) Experimental transmission of Creutzfeldt-Jakob disease. *Act Pathol. JPN* **31:** 943-951.

Truchot, L., Bencsik, A., Perret-Liaudet, A., Biacabe, A.G. Richard, M., Ironside, J., Kopp, N., and Streichenberger, N. (2004) Quantitative study of spongiform changes in putaen of 24 case of Creutzfeldt-Jakob disease. *J. Neuropahol. Exp. Neurol.* **63**, 193-198.

Uboldi, C., Paulis, M., Guidi, E., Bertoni, A., Di Meo, G., Perucatti, A. Iannuzzi, L., Raimondi, E., Brunner, R.M., Eggen, A., and Ferretti, L. (2006) Cloning of the bovine prin-like shadoo *(SPRN)* gene by comparative analysis of the predicted genomic locus. *Mamm Genome* **17,** 1130-1139.

Vorberg, I., Groschup, M.H., Pfaff, E., and Priola, S.A. (2003) Multiple amino acid residues within the rabbit prion protein inhibit formation of its abnormal isoform. *J. Virol.* **77,** 2003-2009.

Wadsworth, J.D., Asante, E.A., Desbruslais, M., Linehan, J.M., Joiner, S, Gowland, I., Welch, J., Stone, L., Lloyd, S.E., Hill, A.F., Brandner, s., and Collinge, J. (2004) Human prion protein with Valine 129 prevents expression of variant CJD phenotype. *Science* **306,** 1793-1796.

Wadsworth, J.D., Hill, A.F., Beck, J.A., and Collinge, J. (2003) Molecular and clinical classification of human prion disease. *Br. Med. Bull.* **66,** 241-254.

Watts, J.C., Balachandran, A., Westaway, D. (2006) The expanding universe of prion diseases. *PLoS Pathog.* Mar **2(3),** e26.

Weissmann, C., and Flechsig, E (2003) PrP knock-out and PrP transgenic mice in prion research. *Br. Med. Bull. Rev*. **66**, 43-60.

White, A.R., Collins, S.J., Maher, F., Jobling, M.F., Stewart, L.R., Thyer, T.M., Beyreuther, K., Masters, C.L., and Cappai, R. (1999) Prion protein-deficient neurons reveal lower glutathione reductase activity and increased susceptibility to hydrogen peroxide toxicity. *Am. J. Pathol.* **5,** 1723-1730.

Zhang, C., de Koning, D., Hernandez-Sanchez, J., Haley, C.S., Williams, J.L., and Wiener, P. (2004) Mapping of multiple quantitative trait loci affecting bovine spongiform encephalopathy. *Genetics* **167,** 1863-1872.

Chapter 2
A 2cM genome-wide scan of European Holstein cattle affected by classical BSE

2.1 Background

Transmissible spongiform encephalopathies (TSEs) are fatal neurodegenerative diseases that have been identified in a number of mammalian species (humans, cattle, sheep, mice, etc.) [1]. One of the common characteristics of these diseases is the accumulation of abnormally folded prion protein within the central nervous system. The prion protein is a glycosyl-phosphatidylinositol (GPI) anchored protein that has a native form (PrP^{c}) for which the secondary structure consists mainly of alpha-helices. The disease-associated or misfolded form, PrP^{Res}, has a substantial increase in the beta-pleated sheets content and reduction of the alpha-helices in comparison to the native form [2]. This altered confirmation is associated with an increased resistance to digestion with proteinase K [3]. Furthermore, the presence of PrP^{Res} behaves like a seed that promotes the conversion of further native PrP^{c} to PrP^{Res} via a mechanism that is to date not completely understood [4].

Specific *PRNP* alleles of non-synonymous polymorphisms in humans and sheep are associated with acquired TSE susceptibility [5-8]. This is not the case in cattle, although the deletion alleles of a 23 base pair insertion/deletion (InDel) polymorphism in the bovine *PRNP* promoter region and of a 12 base pair InDel within intron 1 have been associated with incidence of classical bovine spongiform encephalopathy (BSE) [9-11]. Both of these polymorphisms are contained in a block of high linkage disequilibrium (LD) within *PRNP* that appears conserved in many cattle populations, and is entirely outside of the coding region [12]. Thus, the full extent of *PRNP* association with classical BSE is currently not known. Previous genetic studies of BSE cattle have identified putative loci other than *PRNP* (located on chromosome 13 at 47.2Mb) that are associated with incidence of disease [13, 14]. These studies carried out low density whole-genome scans with microsatellite markers approximately every 20cM in female European Holstein cattle which contracted BSE and unaffected half-sib controls.

A version of this chapter has been published. Murdoch BM *et al.* A 2 cM genome-wide scan of European Holstein cattle affected by classical BSE. *BMC Genet* 2010, 11:20.

The objective of this study was to test loci throughout the bovine genome for an association with classical BSE using markers at 2cM resolution. This resolution is approximately a ten-fold improvement over previous genome scans of BSE samples. Two animal sample sets were used, allowing for two analysis approaches: a case-control association study and family based sib-transmission disequilibrium test (sib-TDT) study [13, 14]. In both sample sets the cattle used were female European Holstein cattle which contracted BSE in the late 1980's and early 1990's (family analysis) and in the mid-1990's (case-control analysis) and were identified on commercial farms where the most likely source of disease was through consumption of contaminated feed during the United Kingdom's BSE epidemic. Although the United Kingdom imposed a ruminant feed ban of meat and bone meal in June 1988, it wasn't until August 1996 that a total ban on bovine meat and bone meal was implemented [15].

2.2 Results

The panel of SNP markers used in this study consisted of 3,072 SNPs dispersed across the genome at an approximate average interval of 2cM [16]. Of these, an average 2,853 passed quality control measures and had an average genotyping success rate of 0.99. The samples used in this study came from two sets of animals with different relatedness that allowed for the use of several statistical analyses to test the association between SNP genotypes with classical BSE incidence. By using both offspring from six sires (sib-TDT) and (case-control) Holstein animals, this study examined within-breed and within-family SNP association with disease incidence.

2.2.1 Family based association testing for BSE incidence

The related sample set (N=481) was comprised of six paternal half-sib sire families of which four were scanned previously with microsatellites [13, 14]. Samples were not available from either the sires or the dams; therefore, the sib-TDT analysis method [17] was used. Additionally, PLINK software [18] was used to establish an empirical p value and determine significance. The results for this sib-TDT analysis (412 animals across 2,827 SNPs after quality control) identified 46 SNPs that passed the

Bonferroni correction with a p<0.05 and 27 SNPs that passed the Bonferroni correction with a p<0.01 (Table 2-1). In addition, a 10,000,000 permutation test was performed on this data using the PLINK "max T" to establish genome wide empirical p values. The genome wide 10,000,000 permutation identified 31 SNPs with an empirical permutation of p<0.01. This group of SNPs included the 27 SNPs which passed the Bonferroni at p<0.01 plus four additional SNPs. Many of these SNPs are located in very close proximity [16] to one another and are potentially in LD in the population.

		Location		unadjusted	permuted	
SNP ID	**CHR**	**bp**	**Gene**	**p value**	**p value**	**Bonferroni**
AAFC02065030	2	37,055,124		1.93E-08	3.61E-05	5.45E-05
rs29020694	4	90,418,076		1.87E-06	4.05E-03	5.29E-03
AAFC02132123_1	4	117,126,683	LOC100138299	9.38E-07	1.98E-03	2.65E-03
AAFC02132123_2	4	117,126,810	LOC100138299	3.38E-06	7.33E-03	9.55E-03
AAFC02012009	5	30,554,866	LOC507184	3.53E-06	7.69E-03	9.98E-03
rs29003193	5	66,501,993		3.12E-08	5.79E-05	8.82E-05
rs29012226	5	67,701,052	ANKS1B	6.63E-07	1.38E-03	1.87E-03
SCAFFOLD106936	6	98,767,642		6.35E-07	1.32E-03	1.80E-03
rs29016161	7	24,604,791		2.45E-06	5.31E-03	6.93E-03
rs29017305	9	62,946,258	BACH2	2.30E-06	4.97E-03	6.49E-03
rs29013631	10	38,725,566		2.25E-06	4.87E-03	6.36E-03
rs29022366	12	80,076,729	LOC786668	2.74E-06	5.95E-03	7.75E-03
rs29010388	14	4,145,186		4.20E-07	8.56E-04	1.19E-03
AAFC02138417	15	70,632,844		3.31E-06	7.19E-03	9.36E-03
rs29010371	16	63,558,950	FAM129A	1.06E-06	2.26E-03	3.00E-03
rs29009572	17[#]	37,575,810		1.29E-06	2.76E-03	3.65E-03
rs29021871	17	44,213,230		1.06E-06	2.25E-03	2.99E-03
CART	20[#]	4,922,252	CART	3.57E-08	6.75E-05	1.01E-04
rs29018531	20	38,814,738		1.32E-06	2.82E-03	3.72E-03
AAFC02028192	21	5,702,403	MEF2A	1.24E-07	2.49E-04	3.49E-04
rs29022862	21	12,595,937		1.85E-07	3.73E-04	5.22E-04
rs29009825	21	24,827,395		1.90E-06	4.11E-03	5.38E-03
rs29012664	22	42,415,179	FHIT	5.30E-07	1.10E-03	1.50E-03
NW_930303_1	24[#]	29,286,364		1.02E-07	2.05E-04	2.90E-04
NW_930303_2	24[#]	29,286,314		3.06E-07	6.17E-04	8.64E-04
SCAFFOLD176855	28	26,871,628	SLC29A3	1.03E-06	2.18E-03	2.90E-03
SCAFFOLD68962	X	7,641,750		3.15E-06	6.86 E-03	8.91E-03

Table 2-1 - The results of sib-TDT model analysis using the large family sample set. Locations were determined by blast to bovine sequence version 4.0. All other locations denoted by [#] were either determined from bovine sequence version 2.0, 4.1 or the Maryland sequence assembly. Permuted p values reported here are from a 10,000,000 genome-wide SNP permutation.

2.2.2 Case-control association testing for BSE incidence

The case-control samples were comprised of 149 BSE case and 184 control animals. The control samples include a least one animal collected from the same farm as each of the BSE cases as well as the controls for, and 15 BSE negative animals. The genotyping data on these animals was analyzed for an association with disease status using the case-control allelic test within the PLINK software [18]. This analysis (320 animal across 2,872 SNPs after quality control) revealed 20 SNPs with a $p<0.01$, 14 SNPs with a $p<0.005$ and 6 SNPs with a $p<0.001$. In order to determine the number of these SNPs that may have occurred by chance, an empirical p-value for each single SNP and across all SNPs (genome wide) was calculated using the max (T) permutation procedure with 10,000 permutations. Following correction for the false discovery rate no significant associations at $p\leq0.05$ genome-wide significance were identified in this data set. This was consistent with the use of Bonferroni multi-test correction on this data set, where again none of the SNPs achieved significance of $p\leq0.05$. To assess significance a threshold was set at $p\leq1.7x10^{-5}$ (Bonferroni calculation). Twenty SNPs were identified with $p<0.01$ (Table 2-2), where four SNPs had a $p\leq5x10^{-4}$ and one SNP had a $p\leq1x10^{-4}$. Thus a single SNP had a suggestive association with BSE incidence on chromosome 14 ($p=7.25x10^{-5}$). SNPs which did not quite reach this threshold but had a $p\leq5x10^{-4}$ were found on chromosomes 4, 10, 14, and 15. The number of loci on each chromosome, the identity of these specific SNPs and their corresponding p-values are reported in Table 2-2.

SNP ID	CHR	Location bp	Gene	unadjusted p value
rs29013431	2	15,312,400		5.99E-03
rs29012194	2	26,095,592		6.28E-03
rs29016537	2	29,654,789		1.30E-03
rs29020907	4	59,252,723	IMP2	4.97E-04
rs29024570	10	4,359,586		4.30E-04
AAFC02107025	10	21,153,203	DHRS1	4.73E-03
rs29024728	13	28,618,368		3.34E-03
PRNP08	13	47,214,453	PRNP	4.38E-03
rs29021171	14	10,106,755		3.84E-03
SCAFFOLD51887	14	43,984,153		7.25E-05*
rs29014819	15	35,713,980		7.17E-04
rs29014820	15	35,714,095		6.05E-04
rs29014821	15	35,714,113		4.66E-04
AAFC02014662	16	65,943,749		7.95E-03
AAFC02012500_1	21	11,820,092	MCTP2	1.82E-03
AAFC02012500_	21	11,820,201	MCTP2	3.93E-03
rs29026011	21	13,254,954	LOC618464	3.00E-03
rs29019629	21	33,209,686	CSPG4	2.01E-03
rs29011202	24	18,493,192		9.08E-03
AJ496776	28	31,495,721		5.74E-03

Table 2-2 - PLINK case-control association results of the case-control sample set. All locations determined by blast to bovine sequence version 4.0. Threshold of significance is $p \leq 1.7 \times 10^{-5}$ and suggestive significance $p \leq 10^{-4}$ is denoted as *.

This data set was also subjected to a best fit model test where the standard allelic, trend, dominance, recessive and genotypic association tests were performed, and the test with the lowest p value was reported. All of the SNPs identified above were also identified in the best fit model as either allelic or trend, however additional SNPs with recessive, dominant and genotype associations were also identified. In the best fit model there were a total of fourteen SNPs with a recessive mode of action, seven dominant SNPs and eight genotypic SNPs with a $p<0.01$ (see Additional Table 2-1). Using the same thresholds as described above, (i.e. significant with a $p \leq 10^{-5}$ and suggestive with a $p \leq 10^{-4}$), one SNP on chromosome 14 had a suggestive association with BSE incidence.

2.3 Discussion

2.3.1 Classical BSE as a phenotype

Clinical presentation of BSE disease is a difficult phenotype to test for genetic associations. Animals that have developed BSE are clearly susceptible, however, those which are clinically healthy are difficult to assess. In the present study, clinically healthy animals were used as controls, however, these animals may have been incubating disease or may not have ingested enough infectious agent to become symptomatic, or alternatively, they may in fact have been resistant to disease. Therefore, the analysis was performed with the realization that phenotypic noise in the controls will have reduced the power to detect associations. Another consideration is that classical BSE is a complex trait which may be more consistent with interactive and possibly subtle effects of multiple contributing loci. Therefore, multiple testing corrections applied to results such as these may be overly prone to type II errors (i.e. discarding real associations). Consequently, it is important to examine the results for supporting evidence of associations between disease and genetic loci, as discussed below.

2.3.2 *PRNP* gene

PRNP variation was not exhaustively tested for an association with classical BSE as the focus of this study was primarily genome wide. Over 380 polymorphisms are known to reside throughout the coding and non-coding regions of *PRNP* [12]. Of these, 13 *PRNP* haplotype tagging SNPs (htSNPs) were used in the scans of which 8 were informative for both data sets. The htSNPs used in this study capture a large portion of *PRNP* haplotype variation observed in a diverse assemblage of U.S cattle, spanning the promoter region into the last exon, however, they do not capture all of it. Additionally, the two InDels previously identified [9-11] as having allele associations with classical BSE were not genotyped in this study and thus no information is available for this sample set and the InDels *PRNP* haplotypes. One SNP (PRNP08) had a p value of 4.38×10^{-3} for an association with classical BSE but did not pass multiple test corrections. Consequently, no significant associations of *PRNP* variation (the 8 informative htSNPs) with classical BSE were identified in this single marker analysis. Given that the single

marker analysis of the informative htSNPs did not capture all of the htSNPs in the *PRNP* region or the two InDels it should not be considered an exhaustive analysis of the *PRNP* gene region.

2.3.3 Family-based analysis

Family-based analyses, although in general being less powerful than case-control studies [19], offer robustness to non-random mating. The transmission/disequilibrium test developed in 1993 by Spielman [17] is intended to test for linkage between complex diseases and genetic markers. The sib-TDT approach used here does not reconstruct parental genotypes in their absence, but uses marker data from unaffected half siblings instead [17]. The DFAM analysis model fits the structure of the related half-sibling sample set and has been utilized in other species as well [20]. The study presented here used a much larger number of markers than the previous studies and used an analysis approach that is robust to population stratification [21]. The advantage this analysis has over the case-control approach is that with 302 affected animals it has twice the number of affected individuals, and thus, a higher study power and likelihood of detecting markers associated with disease loci.

Many of the samples included in the family sample set used in this study were also used by Hernández-Sánchez *et al.* [13] and Zhang *et al*. [14], while the case-control sample set was analyzed here for the first time. Hernández-Sánchez *et al.* [13] also used a TDT approach, however, their analysis method requires heterozygous parents to allow the parent of origin of alleles to be unequivocally determined and as a result many animals had to be disregarded in their analysis. In addition, progeny with the same genotype as the predicted genotype of the sires or progeny that were themselves homozygous were excluded. As a result, although the TDT method used by Hernández-Sánchez *et al.* [13] has the potential to be powerful, the use of this approach with microsatellite marker based data was limited by the number of genotypes which could be used and ranged from 92 (in the case of marker BMS1658) to 210 (in the case of INRA36). The analysis method used to localize QTLs by Zhang *et al*. [14] was a regression approach which does not require the parents to be heterozygous and hence all individuals could be included in the analysis. However, the QTL approach is not robust

to population stratification. Moreover, the total number of samples (360) used in the Zhang *et al.* [14] study was smaller than that of Hernández-Sánchez *et al.* [13] (530) as well as this study (412). The two previous analyses of overlapping family samples yielded different results: the TDT analysis of Hernández-Sánchez *et al.* [13] found evidence for associations with BSE incidence on chromosomes 5, 10 and 20, whereas the QTL analysis by Zhang *et al*. [14] identified BSE associated QTL chromosomes 1, 6, 13, 17, 19 and X/Y_{ps}. SNPs in the regions matching regions found in previous studies are detailed in Table 2-3. With regards to the family data this study offers a similar power to that of the TDT analysis by Hernández-Sánchez *et al.* [13], however uses half-sib controls as opposed to inferring the sires genotype. In addition, due to necessary genotypic restrictions (only heterozygous genotypes can be used) of the TDT method, the approach used here allowed for a greater number of animals to be included in the analysis. Although the total number of animals used in this analysis includes two smaller families (6 families versus 4 families in the previous studies [13, 14]) this analysis method does not infer sire genotypes and therefore the inclusion of the two smaller families do not reduce the overall power of this analysis.

The sib-TDT analysis in this study identified two significant SNPs on BTA 20 associated with BSE incidence, rs29018531 at 38.8Mb and a SNP within the cocaine and amphetamine responsive transcript peptide gene, (*CART*). *CART* is not currently on the Btau4.0 bovine sequence assembly but was previously mapped to chromosome 20 at 38.5cR [15]. In addition, the location of *CART* was reported on Btau2.0 as 4.92Mb as well as the Maryland map as 9.78Mb. Therefore it is unclear from this study if the observed associations identified here are attributable to one locus or two separate loci. The study of Hernández-Sánchez *et al.* [13] observed an association with marker INRA36 (at 37.9Mb) on BTA 20 with BSE incidence. This study also identified significant markers on chromosomes 5, 6, 10, 17 and X associated with BSE. The study by Hernández-Sánchez *et al.* [13] also reports associations on BTA 5 and 10 but did not report confidence intervals. The marker identified on BTA 6 in this study (Table 2-3) is in the same chromosomal region as the marker described by Zhang *et al.* [14] and is within the confidence interval. Moreover, the significant marker on BTA 6, Scaffold106936 at 98.7Mb corresponds with the QTL region, 51-72cM on mouse chromosome 5, previously associated with susceptibility to TSE in mice [22].

Additionally, the homologous region to that identified on BTA 6 was also identified as a QTL modulating scrapie incubation period in sheep [23]. Interestingly, the QTL region on mouse chromosome 5 described by Moreno *et al.* [22] also corresponds to the location of a significant SNP identified on BTA 17, 44.2Mb. Thus these chromosomal regions identified in the present study are also supported by studies in cattle and other species. Comparative locations were determined by using the National Center for Biotechnology information map viewer of the mouse QTL regions (build 37.1), then the human and Btrna were selected and bovine locations were determined.

CHR	SNP ID	Analysis Method	p-value	Location Mb	Previously Observed	Peak Location
1	rs29009859	sib-TDT	2.0×10^{-5}	89.7	[14]	106.5
5	AAFC02012009	sib-TDT	$^{**}9.9 \times 10^{-3}$	30.6	No	
5	E25B16-36408-3	sib-TDT	$^{**}8.8 \times 10^{-5}$	66.5	No	
5	rs29012226	sib-TDT	$^{**}1.9 \times 10^{-3}$	67.7	No	
5	rs29024670	sib-TDT	3.0×10^{-5}	112.8	[13]	107.0
6	SCAFFOLD106936_12205	sib-TDT	$^{**}1.8 \times 10^{-3}$	98.8	No	
10	rs29024570	Case control	$^{*}4.3 \times 10^{-4}$	4.4	No	
10	rs29013631	sib-TDT	$^{**}6.5 \times 10^{-3}$	38.7	No	
10	rs29015623	sib-TDT	9.0×10^{-5}	48.0	[13]	40.0
17	rs29021871	sib-TDT	$^{**}3.0 \times 10^{-3}$	44.2	No	
19	rs29027102	sib-TDT	2.0×10^{-5}	62.5	[14]	53.5
20	rs29018531	sib-TDT	$^{**}3.7 \times 10^{-3}$	38.8	[13]	46.0
20	CART-SNP	sib-TDT	$^{**}1.0 \times 10^{-4}$	$4.9^{\#}$	[13]	
X	SCAFFOLD68962_9331	sib-TDT	$^{**}1.0 \times 10^{-5}$	7.6	No	
X	SCAFFOLD285727_12117	sib-TDT	1.0×10^{-5}	85.5	[13]	112.5

Table 2-3 - Comparison of SNPs observed in this study with locations indentified in other studies. Sib-TDT p values are reported as unadjusted unless denoted by ** which indicates Bonferroni corrected values that pass 10,000,000 permutations. All case control p-values are reported are as unadjusted and suggested significance is denoted as *. Only SNPs with appreciable or significant p-values located within the confidence intervals of ref# [14] or within 10cM of ref# [13] are reported here. The location determined by older bovine sequence version 2.0 is denoted by #. All other locations were determined by blast to bovine sequence version 4.0 and are reported in mega bases. The previous studies reported location in cM however for the sake of uniformity the location is reported here, based on marker positions, in Mb.

2.3.4 Case-control analysis

In selecting the analysis approach, it is important to match the appropriate model to the data structure to maximize the power. This study used a case-control approach to analyze the unrelated BSE animals. This approach is powerful in its ability to detect loci linked with disease: however, it has been criticized by geneticists for its lack of robustness for population stratification arising from non-random mating or unknown relationships between individuals [18]. The data presented here was examined for stratification and none was observed (see additional Figure 2-1). Despite the power of the case-control approach, this study was limited by the relatively small number of animals used (149 cases and 184 controls) and no significant results were observed. An increase in the number of cases and controls included in this sample set would have a dramatic effect on the power to detect loci associated with disease [18], as would a higher density scan conducted with an increased number of genome-wide markers.

2.3.5 Shared regions identified in the sib-TDT family and implicated in the case-control analysis

Many of the SNPs included in the panel are in close proximity and are in LD in Holstein [15]. Therefore, it may be more appropriate to consider the results in terms of chromosomal regions instead of individual markers. Linkage disequilibrium will result in the alleles of several closely spaced SNPs being associated with disease status because they all fall on the same haplotype. Thus, it would be expected that several SNPs in LD with a locus involved in disease would show significant associations. Examples of this can be observed with the loci on BTA 15 in Table 2-2 as well as BTA 4 in Table 2-1. In addition, if the same regions give significant or a nearly significant association across the different sample sets, this would also increase confidence that the association is real. Chromosome 2 is a good example: the most significant marker in this study, AAFC02065030, with a Bonferroni corrected $p=5.5x10^{-5}$ in the family based analysis is at 37.1Mb on this chromosome. In the unrelated samples set, three markers were identified on Chromosome 2, one at 15.3Mb, another at 26.1Mb, and the last ($p=1.0x10^{-3}$) at 29.7Mb. From the different analyses methods and across the two sample sets four markers were identified in the chromosomal region from 15.3Mb-37.1Mb on BTA2, with

the region of 29-37Mb being significant. This data supports an association between chromosome 2 and BSE disease, however from the data it cannot be determined if the effects of one locus or several loci spread across the region.

Another chromosomal region which harbours significant markers that were identified in the family-based analyses and was observed but failed to reach significance in the unrelated analysis is on chromosome 21. Four SNPs were observed on BTA 21 from the analysis of unrelated samples, two SNPs which are located at 11.8Mb ($p=4.4x10^{-3}$ and $p=8.3x10^{-3}$) in the *MCTP2* gene, another at 13.3Mb ($p=2.5x10^{-3}$) in LOC618464, and an addition SNP at 33.2Mb ($p=2.0x10^{-3}$) in the gene *CSPG4*. The family based sib-TDT analysis identified three significant SNPs, at 5.7Mb in the *MEF2* gene ($p=3.5x10^{-4}$), 12.6Mb ($p=5.2x10^{-4}$), and 24.83Mb ($p=5.4x10^{-3}$). When considering both sample sets there appears to be an interval from 11.8Mb-13.3Mb with the outer range of approximately 5.7-33.2Mb for a locus with an effect on BSE. BTA 21 shares conservation of synteny with sheep OAR 16 where QTLs for scrapie susceptibility [22] and scrapie incubation period have been reported [21].

Regions containing loci significantly associated with disease status identified on chromosomes 14, 16 and 28 in the sib-TDT analysis of the family-based data were also observed to be of interest, but failed to reach significance, in the case-control study of unrelated-animals. On chromosome 14 the marker, rs29021171, at 10.1Mb had a $p=2.3x10^{-3}$ in the analysis of unrelated-animals and is in relative close proximity to, rs29010388, at 4.2Mb which was identified as significant in the sib-TDT analysis ($p=1.2x10^{-3}$). On chromosome 16 marker AAFC02014662 at 65.9Mb identified in the case-control ($p=7.95x10^{-3}$) is in close proximity to rs29010371 at 63.5Mb, which is in the gene *FAM129A*, that was identified as significant ($p=3.0x10^{-3}$) in the sib-TDT analysis. Finally on chromosome 28 marker AJ496776, at 31.5Mb identified in the unrelated sample analysis ($p=9.9x10^{-3}$) is in close proximity to the marker SCAFFOLD176855 (within *SLC29A3* gene) at 26.9Mb identified as significant ($p=2.9x10^{-3}$) in the sib-TDT analysis. None of these chromosomal regions have been previously reported as being associated with BSE.

2.3.6 Candidate genes identified in the case-control and /or sib-TDT family analysis

The SNPs identified from the family samples on chromosomes 4, 5, 9, 12, 16, 20, 21, 22 and 28 are all found within genes; however, the polymorphisms on chromosomes 4 and 12 are in hypothetical genes. The most notable is the polymorphism on chromosome 5, rs29012226, in the ankyrin repeat and sterile alpha motif domain containing 1B gene (*ANKS1B*). This gene is also known as amyloid beta protein precursor (*APP*) intracellular domain associated protein 1 (AIDA-1), and is associated with *APP* binding [26]. It is well known that APP generates beta amyloid and plays a key role in Alzheimers disease [27-30]. Further cellular prion protein and AIDA-1 has been implicated as a receptor for amyloid–β oligomers [31, 26], making *ANKS1B* a good candidate gene for further study.

Another candidate gene, which is in close proximity to three SNPs on BTA 2 at ~29.3Mb, with alleles that associated with BSE incidence is *B3GALT1*. Beta-1,3-galactosyltransferase (*B3GALT1*), is a transferase polypeptide gene involved in the biosynthesis of GPI anchors. The involvement of the GPI anchor, with lipid raft and TSE disease has been investigated [32] and it is thought that the GPI anchor may affect the conformation, or the association of the prion protein with specific membrane domains [33]. An additional potential candidate gene is *CART*, cocaine and amphetamine regulate transcript. This neuropeptide plays a role in a variety of physiological processes, some of which include: promotion of hippocampal neurons by upregulating brain-derived neurotrophic factors [34] and synaptogenesis [35]. In addition, the expression of *CART* has shown to be down regulated in mouse prion disease [36].

The SNP located on BTA 14 (43.9Mb), associated with BSE incidence in the case-control analysis, is in close proximity to the gene exotoses (multiple) 1, *EXT1*. McCormick et al., [33] showed that *EXT1* is an endoplasmic reticulum (ER)-resident type II transmembrane glycoprotein whose expression in cells results in the alteration of the synthesis and display of cell surface heparan sulfate glycosaminoglycans (GAGs). The N terminus of PrP contains a GAG-binding motif and it is thought that PrP binding of GAG is important in prion disease [33-35]. Additionally, this region contains another candidate gene *STMN2*, which has been identified in a whole genome association study for genetic risk factors for variant Creutzfeldt-Jakob in humans [36]. Specifically, Mead

and others [36] found an association with acquired prion diseases, including vCJD (p=5.6 x 10(-5)), kuru incubation time (p=0.017), and resistance to kuru (p=2.5 x 10(-4)), in a region upstream of *STMN2* (the gene that encodes *SCG10*). Superior cervical ganglion 10, *SCG10*, is a neuronal growth associated protein and may play a role in neuronal differentiation in modulating membrane interaction with the cytoskeleton during neurite outgrowth. *STMN2* is at 39.9Mb on chromosome 14 in cattle, which is in close proximity to *EXT1*, making both *STMN2* and *EXT1* are functional and positional gene candidates.

2.4 Conclusions

The large number of SNP markers and the two sets of animals used in this study make it the most comprehensive study to date to test genetic loci for an association with classical BSE in European Holstein cattle. The genome-wide scan of half sib families identified an association between the genetic loci on 18 chromosomes with BSE incidence in European Holstein cattle, including a region on BTA 20 associated with BSE incidence that has been reported in previous studies. The identification of markers at or near statistical significance within the same chromosomal regions in both sets of samples provides independent evidence for the association of those regions and the presence of one or more genes within the regions influencing the incidence of BSE in cattle. While this study has identified some interesting findings it is important that these results are confirmed in additional cattle populations as well as other species. It is worth noting that this study identified a large number of associations with classical BSE disease incidence throughout the bovine genome verse one single major locus with a large effect in the bovine genome. This would make it difficult to select cattle that are genetically resistant to classical BSE, however the results give some insight into gene pathways important during disease progression.

2.5. Methods

2.5.1. Animal information

This study used two sets of samples from cases and controls, but with different structures. The first sample set consisted of female European Holstein collected in the mid 90's and included 149 BSE case and 184 control animals. The control animals were collected from the same farms as the BSE cases but otherwise were not known to be related to each other. The second sample set was family based and consisted of 302 BSE affected and 179 unaffected half-sib Holsteins from six sire families. All the BSE affected and unaffected cattle within one family were paternal half sibs from the designated sire but with different dams. No DNA samples were available from any of the sires.

In both sample sets cattle designated as BSE positive were first examined by qualified veterinarians. BSE status was subsequently confirmed post-mortem by histology (by the Veterinary Laboratories Agency, New Haw, Surrey, UK). None of the control animals exhibited any clinical symptoms of disease and were presumed to be free of disease. All of the control animals were age and sex matched from the same calving season and from the same farm as the BSE cases. As such, the control animals are assumed to have been exposed to the same environment.

2.5.2. DNA isolation and genotyping

Genomic DNA from the case-control was isolated from blood using a high salt phenol/chloroform extraction method as described by Sherman *et al.* [41]. Genomic DNA from the family animal set was isolated from blood samples by phenol and chloroform extraction, as described by Hernández-Sánchez *et al.* [13].

The genotyping panel was comprised of two oligonucleotide pool assays (OPAs) as described by McKay and others in 2007 [16]. Briefly, 5,500 SNP were mapped on the Roslin-Cambridge 3,000 rad bovine–hamster whole genome radiation hybrid panel (WGRH3000) [42] and the minor allele frequency (MAF) was determined on a variety of breeds, including Holstein. Of the original SNPs, 3,072 were selected to give the greatest genome coverage and MAF >0.05 for the genome scan. An Illumina GoldenGate assay

[43] was performed using the two custom OPAs and genotypes determined using an Illumina BeadScan (Illumina Inc., San Diego, CA) and the Illumina BeadStudio software. Sequences containing SNPs were blasted on the bovine assembly (4.0) to determine the SNP locations [16] and, the location for all except 111 SNPs were determined.

2.5.3. Data quality control

Quality control analyses were carried out by removing confounding effects prior to the data analysis. For the case-control sample set from the original 3,072 SNPs in the assay, 122 SNPs were excluded based on assay failure, as observed by poor clustering of alleles using the BeadStudio software. The remaining 2,950 SNPs were submitted to the PLINK program [18] where seven duplicated SNPs were removed and 12 SNPs were removed because more than 10% of the samples failed to genotype at that locus (GENO >0.1). Additionally, 13 samples were removed due to low genotyping rate (MIND>0.1). A further 61 SNPs were removed due to a MAF <0.01. The graphical representation of the MAF for each SNP for each chromosome is shown in additional Figure 2-2. The remaining individuals that were included in the analyses had a mean genotyping rate of 0.995. Population stratification was tested for in this data set and none was observed (data in additional Figure 2-1).

For the family sample set, transmission disequilibrium analysis was performed using the sib-TDT application in the PLINK program. Following the removal of SNPs that did not cluster well, 2,904 SNPs were used in the analysis. Of these loci, a further 22 SNPs were removed due to missing data (GENO>0.1), seven duplicated SNPs, and a further 48 SNPs were removed due to low allele frequency (MAF<0.01). The MAFs for each of the SNPs in this data set is provided in additional file 3 and the graphical representation of the MAF for each chromosome is shown as in additional file 4. Of the individuals examined, 5 samples failed and 64 were removed because of low genotyping frequency (MIND>0.1), for the 412 remaining individuals (see Table 2-4) a genotyping success rate of 0.98 was obtained.

Sire	Number of BSE			Number of control		
	This study	**[13]**	**[14]**	**This study**	**[13]**	**[14]**
Barold Rock Seal	33	53	43	11	28	22
Brynhyfryd Cascade	68	93	72	28	56	18
Leighton Workboy	71	88	70	40	44	22
Maybar Juniper	90	124	83	51	44	30
Bowerchalk Polacca	4	0	0	6	0	0
Deri cascade	3	0	0	7	0	0
Total	269	358	268	143	172	92

Table 2-4- The sire identities and the number of half sib offspring analyzed in this study verses the number of animal used two previous studies.

2.5.4. Statistical analysis

The PLINK software v1.04 was used to perform the majority of the statistical analysis [18]. The data from the case-control sample set were analyzed using the basic case-control association (χ^2) test. Whereas the family sample set was analyzed with the DFAM program [18], which is an adjusted family TDT analysis, as described by Spielman [17]. To correct for multiple tests Bonferroni single-step adjusted p-values (BONF) procedures were applied. A permutation test was also used in this study, which was max(T)" permutation with 10,000 and 10,000,000 permutations, a procedure that permutes both a point-wise SNP significance and a genome-wide significance.

2.6 References

1. Prusiner SB, Hsiao KK: Human prion disease. *Ann Neuro.* 1994, 35:385-395.
2. Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, Groth D, Mehlhorn I, Huang Z, Fletterick RJ, Cohen FE, Prusiner SB: Conversion of α-helices into β-sheets features in the formation of the scrapies prion proteins. *Proc Natl Acad Sci USA* 1993, 90:10962-10966.
3. Cronier S, Gros N, Tattum MH, Jackson GS, Clarke AR, Collinge J, Wadsworth JDF: Detection and characterization of proteinase K-sensitive disease-related prion protein with thermolysin. *Biochem J* 2008, 416(2):297-305.
4. Atarashi R, Moore RA, Sim VL, Hughson AG, Dorward DW, Onwubiko HA, Priola SA, Caughey B: Ultrasensitive detection of scrapie prion protein using seeded conversion of recombinant prion protein. *Nat Methods* 2007, 8:645-650.
5. Laegreid WW, Clawson ML, Heaton MP, Green BT, O'Rourke KI, Knowles DP: Scrapie resistance in ARQ sheep. *J Virol* 2008, 82(20):10318-10320.
6. Alvarez L, Arranz JJ, San Primitivo F: Identification of a new leucine haplotype (ALQ) at codon 154 in the ovine prion protein gene in Spanish sheep. *J Anim Sci* 2006, 84(2):259-265.
7. Collinge J, Sidle KCL, Meads J, Ironside J, Hill AF: Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* 1996, 383:685-690.
8. Gambetti P, Kong Q, Zou W, Parchi P, Chen SG: Sporadic and familial CJD: classification and characterisation. *Br Med Bull* 2003, 66:213-239.
9. Sander P, Hamann H, Drögemüller C, Kashkevich K, Schiebel K, Leeb T: Bovine prion protein gene (*PRNP*) promoter polymorphisms modulate *PRNP* expression and

may be responsible for differences in bovine spongiform encephalopathy susceptibility. *J Biol Chem* 2005, 280(45):37408-37414.

10. Juling K, Schwarzenbacher H, Williams JL, Fries R: A major genetic component of BSE susceptibility. *BMC Biol* 2006, 4:33.
11. Haase B, Doherr MG, Seuberlich T, Drögemüller C, Dolf G, Nicken P, Schiebel K, Ziegler U, Groschup MH, Zurbriggen A, Leeb T: *PRNP* promoter polymorphisms are associated with BSE susceptibility in Swiss and German cattle. *BMC Genetics* 2007, 8:15.
12. Clawson ML, Heaton MP, Keele JW, Smith TPL, Harhay GP, Laegreid WW: Prion gene haplotypes of U.S. cattle. *BMC Genetics* 2006, 7:51
13. Hernández-Sánchez J, Waddington D, Wiener P, Haley CS, Williams JL: Genome-wide search for markers associated with bovine spongiform encephalopathy. *Mamm Genome* 2002, 13(3):164-168.
14. Zhang C, de Koning DJ, Hernández-Sánchez J, Haley CS, Williams JL, Wiener P: Mapping of multiple quantitative trait loci affecting bovine spongiform encephalopathy. *Genetics* 2004, 167(4):1863-1872.
15. Anderson RM, Donnelly CA, Ferguson NM, Woolhouse MEJ, Watt CJ, Udy HJ, MaWhinney S, Dunstan SP, Southwood TRE, Wilesmith JW, Ryan JB, Hoinville LJ, Hillerton JE, Austin AR, Wells GA. Transmission dynamics and epidemiology of BSE in British cattle. *Nature* 1996, 382:779–788.
16. McKay SD, Schnabel RD, Murdoch BM, Aerts J, Gill CA, Gao C, Li C, Matukumalli LK, Stothard P, Wang Z, Van Tassell CP, Williams JL, Taylor JF, Moore SS: Construction of bovine whole-genome radiation hybrid and linkage maps using high-throughput genotyping. *Anim Genet* 2007, 38(2):120-125.

17. Spielman RS, Ewens WJ: A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet* 1998, 62:450-458.
18. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC: PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007, 81:559-575.
19. Risch N, Teng J: The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases I. DNA pooling. *Genome Res.* 1998, 8:1273-1288.
20. Wiiki A, Wade C, Biagi T, Ropstad E, Bjerkas E, Lindblad-Toh K, Lingaas F: A deletion in Nephronophthisis 4 (*NPHP4*) is associated with recessive cone-rod dystrophy in standard wire-haired dachshund. *Genome research* 2008, Sep 18(9): 1415-1421.
21. Devlin B, Roeder K, Bacanu SA: Unbiased methods for population-based association studies *Genet Epidem* 2001, 21:273-284.
22. Moreno CR, Lantier F, Lantier I, Sarradin P, Elsen JM: Detection of new quantitative trait Loci for susceptibility to transmissible spongiform encephalopathies in mice. *Genetics* 2003, 165(4):2085-91.
23. Moreno CR, Cosseddu GM, Schibler L, Roig A, Moazami-Goudarzi K, Andreoletti O, Eychenne F, Lajous D, Schelcher F, Cribiu EP, Laurent P, Vaiman D, Elsen JM: Identification of new quantitative trait loci (other than the *PRNP* gene) modulating the scrapie incubation period in sheep. *Genetics* 2008 179(1):723-726.

24. Lloyd SE, Uphill JB, Targonski PV, Fisher EM, Collinge J. Identification of genetic loci affecting mouse-adapted bovine spongiform encephalopathy incubation time in mice. *Nuerogenetics* 2002, Oct 4(2): 77-81.

25. Laurent P, Schibler L, Vaiman A, Laubier J, Delcros C, Cosseddu G, Vaiman D, Cribiu EP, Yerle M. A 12 000-rad whole-genome radiation hybrid panel in sheep: application to the study of the ovine chromosome 18 region containing a QTL for scrapie susceptibility. *Anim Genet* 2007, 38(4):358-363.

26. Ghersi E, Noviello C, D'Adamio L: Amyloid –β protein precursor (AβPP) intracellular domain-associated protein-1 proteins bind to AβPP and modulate its processing in an isoform-specific manner. *JBC* Nov 19 279(47):49105-49112.

27. Ring S, Weyer SW, Kilian SB, Waldron E, Pietrzik CU, Filippov MA, Herms J, Buchholz C, Eckman CB, Korte M, Wolfer DP, Müller UC: The secreted β-amyloid precursor protein ectodomain APPsα is sufficient to rescue the anatomical, behavioral, and electrophysiological abnormalities of APP-deficient mice. *J Neurosci* 2007, 27(29):7817-7826.

28. Li ZW, Stark G, Götz J, Rülicke T, Gschwind M, Huber G, Müller U, Weissmann C: Generation of mice with a 200-kb amyloid precursor protein gene deletion by Cre recombinase-mediated site-specific recombination in embryonic stem cells. *Proc Natl Acad Sci USA.* 1996, 93(12):6158-6162.

29. Müller U, Cristina N, Li ZW, Wolfer DP, Lipp HP, Rülicke T, Brandner S, Aguzzi A, Weissmann C: Behavioral and anatomical deficits in mice homozygous for a modified β-amyloid precursor protein gene. *Cell* 1994, 79(5):755-765.

30. Zheng H, Jiang M, Trumbauer ME, Sirinathsinghji DJS, Hopkins R, Smith DW, Heavens RP, Dawson GR, Boyce S, Conner MW, Stevens KA, Slunt HH, Sisoda SS,

Chen HY, Van der Ploeg LHT: β-amyloid precursor protein-deficient mice show reactive gliosis and decreased locomotor activity. *Cell* 1995, 81(4):525-531.

31. Laurén J, Gimbel DA, Nygaard HB, Gilbert JW, Strittmatter SM: Cellular prion protein mediates impairment of synaptic plasticity by amyloid-β oligomers. *Nature* 2009, 457 (7233):1128-1132.
32. Paulick MG, Bertozzi CR: The glycosylphosphatidylinositol anchor: a complex membrane-anchoring structure for proteins. *Biochemistry* 2008, 47(27):6991-7000.
33. Campana V, Caputo A, Sarnataro D, Paladino S, Tivodar S, Zurzolo C: Characterization of the properties and trafficking of an anchorless form of the prion protein. *J Biol Chem* 2007, 282(31):22747-22756.
34. Wu B, Hu S, Yang M, Pan H, Zhu S: CART peptide promotes the survival of hippocampal neurons by upregulating brain-derived neurotrophic factor. *Biochem Biophys Res Commun* 2006, 347(3):656-661.
35. Abrahám H, Orsi G, Seress L: Ontogeny of cocaine- and amphetamine-regulated transcript (CART) peptide and calbindin immunoreactivity in granule cells of the dentate gyrus in the rat. *Int J Dev Neurosci* 2007, 25(5):265-274.
36. Sorensen G, Medina S, Parchaliuk D, Phillipson C, Robertson C, Booth SA: Comprehensive transcriptional profiling of prion infection in mouse models reveals networks of responsive genes. *BMC Genomics* 2008, Mar 3 9:114.
37. McCormick C, Leduc Y, Martindale D, Mattison K, Esford LE, Dyer AP, Tufaro F: The putative tumour suppressor EXT1 alters the expression of cell-surface heparan sulfate. *Nat Genet* 1998, 19(2):158-161.
38. Yin S, Pham N, Yu S, Li C, Wong P, Chang B, Kang SC, Biasini E, Tien P, Harris DA, Sy MS: Human prion proteins with pathogenic mutations share common

conformational changes resulting in enhanced binding to glycosaminoglycans. *Proc Natl Acad Sci USA* 2007, 104(18):7546-7551.

39. Yu S, Yin S, Pham N, Wong P, Kang SC, Petersen RB, Li C, Sy MS: Ligand binding promotes prion protein aggregation--role of the octapeptide repeats. *FEBS J* 2008, 275(22):5564-5575.
40. Mead S, Poulter M, Uphill J, Beck J, Whitfield J, Webb TEF, Campbell T, Adamson G, Deriziotis P, Tabrizi SJ, Hummerich H, Verzilli C, Alpers MP, Whittaker JC, Collinge J: Genetic risk factors for variant Creutzfeldt-Jakob disease: a genome-wide association study. *Lancet Neurol* 2009, 8(1):57-66.
41. Sherman EL, Nkrumah DJ, Murdoch BM, Moore SS: Identification of polymorphisms influencing feed intake and efficiency in beef cattle. *Anim Genet* 2008, 39(3):225-231.
42. Williams JL, Eggen A, Ferretti L, Farr CJ, Gautier M, Amati G, Ball G, Caramorr T, Critcher R, Costa S, Hextall P, Hills D, Jeulin A, Kiguwa S, Ross O, Smith AL, Saunier K, Urquhart B, Waddington D. A bovine whole-genome radiation hybrid panel and outline map. *Mamm Genome* 2002, 13(8):469-474.
43. Oliphant A, Barker DL, Stuelpnagel JR, Chee MS: BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. *Biotechniques* 2002, Suppl 56-8:60-61.

Chapter 3
PRNP Haplotype Associated with Classical BSE Incidence in European Holstein Cattle

3.1 Background

Transmissible spongiform encephalopathies (TSEs), also known as prion diseases, are a group of mammalian neurodegenerative diseases that are invariably fatal and affect humans, ruminants, cats, and mink [1] (reviewed by [2]). TSEs are characterized by abnormal deposits of a protease-resistant isoform of the host genome encoded prion protein and are unique in that they can manifest through acquired, inherited, or sporadic origins [3]. At least three distinct bovine spongiform encephalopathies (BSEs) are known to afflict cattle [4]. The most common TSE of cattle is classical BSE with 190,542 cases reported in 26 countries (http://www.oie.int/eng/info/en_esbmonde.htm). Classical bovine spongiform encephalopathy is an acquired TSE that is most likely spread through the consumption of meat and bone meal contaminated with the infectious prion agent. Dietary exposure to products from BSE infected cattle is the suspected cause of the human TSE, variant Creutzfeldt-Jakob Disease (vCJD) [5-7]. Atypical BSE or BASE occurs in two forms, "H" and "L" [8,9]. Approximately 51 atypical BSE cases have been reported worldwide (quote from Dr. Reg Butler, http://www.ibtimes.com/contents/20100318/cattle-disease-classical-bse-atypical-bse.htm).

The prion protein is essential for the development of TSE disease [10], and genetic variations in the prion gene (*PRNP*) have been associated with TSE susceptibility in humans [3,11], sheep [12,13], deer [14,15], and cattle [16-19]. Two bovine *PRNP* alleles have been associated with susceptibility to classical BSE: a 23 base pair (bp) deletion within the promoter region and a 12 bp deletion within intron 1 [16-19]. However, the deletion alleles are not entirely independent of one another as there is high linkage disequilibrium (LD) between the two polymorphic sites in *Bos taurus* cattle populations [20]. This suggests that the possible effects of variations in the *PRNP* gene on incidence of classical BSE may be better understood if *PRNP* haplotypes were considered in testing

A version of this chapter has been accepted for publication. Murdoch BM *et al.* Prion haplotype associated with classical BSE incidence European Holstein Cattle. PLoS ONE 2010.

for association with disease incidence. Moreover, *PRNP* haplotypes, containing one or both of the two insertion/deletion alleles, may have a stronger association with either susceptibility or resistance to classical BSE than if the indels are considering independently.

More than 390 polymorphisms have been described in a 25-kb region of chromosome 13 containing the *PRNP* gene [20,21]. This chromosomal segment contains distinct regions of high and low LD that is conserved across many *Bos taurus* cattle populations [20]. The region of high LD includes the promoter region, exons 1 and 2, and part of intron 2 (6.7-kb) of the *PRNP* gene. Importantly, both the 23- and 12-bp indels that have been associated with classical BSE susceptibility are contained in this region of high LD. The remainder of *PRNP*, including the entire coding region has relatively low LD. To account for the genetic architecture of the *PRNP* gene, a set of haplotype tagging single nucleotide polymorphisms (htSNPs) has been described that efficiently define haplotypes within and across each of the LD regions. These htSNPs can be used to test for association between *PRNP* haplotypes and susceptibility to either classical or atypical BSE susceptibility [20,22]. In this study, 18 htSNPs, including the 12-bp, and 23-bp indels were used to test *PRNP* haplotypes for an association with classical BSE in European Holstein cattle. Haplotypes associated with healthy control animals and classical BSE were identified.

3.2 Results

The objective of this study was to test for association between *PRNP* haplotypes and BSE disease using a set of htSNPs that effectively tag haplotypes within and across the *PRNP* locus. All htSNPs previously described by Clawson *et al.* [20] and the 23 bp and 12-bp indels described by Hill *et al.* [23] within the *PRNP* gene were genotyped in 330 European Holstein cows, of which 146 were BSE cases and 184 were controls. Single marker and haplotype analyses were performed using PLINK software [24]. Although many of these htSNPs had been previously genotyped [25] and tested for single marker associations with BSE disease status, the complete set of htSNPs and the 23 and 12-bp indels were not included in the earlier study, and therefore the haplotypes were not tested for association with BSE disease.

Initially 18 of the haplotype tagging SNPs, as well as the 23 and 12-bp indels, were tested independently, for an association with disease status. Three of the htSNPs were monomorphic in the sample set and therefore were not included in the haplotype analysis. A single SNP located at 13861 bp (based on NCBI accession number DQ457195) showed a significant association with disease incidence with a Bonferroni corrected *p* value of 0.027 and an odds ratio (OR) of 1.885 (Table 3-1). A significant association with htSNP 13861 and BSE incidence was also observed in an additional family based case control sample set described in Murdoch *et al.* [25] (Additional Table 3-1). SNP 13861 which is location in intron 2 has not been identified as a regulatory region of *PRNP* [26]. Additionally, the T allele at the htSNP 4136 was over represented in control animals in the single marker association analysis (Table 3-1 and [25]), however, the association was not significant after correcting for the multiple testing. Further, in an additional family-based sample set described by Murdoch *et al.* [25], the frequency of htSNP 4136 was low (MAF < 0.05) and was not found to have a significant association with disease status (Additional Table 3-1).

SNP ID	Allele 1	Frequency in BSE affected	Frequency in unaffected	Allele 2	*p-value* (uncorrected) for BSE association	*p-value* (corrected) for BSE association	Odds Ratio
*snp 248	C	0	0	T	NA	1	NA
*snp 449	G	0.290	0.327	T	0.393	1	0.838
*snp 1392	T	0.005	0.004	C	0.807	1	1.413
*snp 1567	T	0	0	C	NA	1	NA
*snp 1701	A	0.405	0.448	G	0.365	1	0.840
*snp 1783	A	0	0	G	Na	1	NA
indel 23-bp	I	0.290	0.361	D	0.110	1	0.722
*snp 3641	C	0.290	0.351	T	0.168	1	0.754
*snp 4136	T	0.047	0.106	C	0.024	0.48	0.419
indel 1-2bp	I	0.342	0.417	D	0.104	1	0.726
snp 4732	A	0.310	0.250	G	0.182	1	1.350
*snp 4776	T	0.147	0.136	C	0.740	1	1.095
*snp 6811	T	0.021	0.011	A	0.397	1	1.900
*snp 8631	G	0.431	0.462	A	0.510	1	0.881
*snp 9162	C	0.006	0	T	0.233	1	NA
*snp 9786	C	0.447	0.485	T	0.426	1	0.859
snp 13793	G	0.483	0.402	A	0.097	1	1.387
snp 13861	C	0.574	0.417	G	0.001	0.027	1.885
snp 13925	G	0.067	0.062	C	0.820	1	1.096
snp 17284	A	0.111	0.061	G	0.130	1	1.920
*snp 20720	T	0.011	0.004	C	0.374	1	2.840
*snp 20957	T	0.058	0.063	C	0.807	1	0.907
*snp 21680	T	0.367	0.407	C	0.392	1	0.846

Table 3-1. *PRNP* htSNPs and indel frequencies within BSE and case animals. The SNP denoted by * was previously tested for an association with classical BSE and no significant associations were found (Murdoch *et al*. [25]). The presence of NA denotes that SNP was not analyzable due to the absence in the unaffected sample set.

The genetic architecture of *PRNP* that has regions of high and low LD, and is conserved across many *Bos taurus* populations, was taken into consideration in the haplotype-case association analysis. HtSNPs previously described by Clawson *et al.* [20] were used to define the haplotypes present: nine htSNPs are in the low LD region, and ten htSNPs including both the 23 and 12-bp indels are in the high LD region (Figure 3-1). A segment of *PRNP* within low LD region, containing htSNPs 13793, 13861 and 13925, is rich in GC content and contains a multi- G indel. Although 330 samples were

sequenced, 101 of the samples did not yield reliable sequence quality and thus genotypes for this region; htSNPs 13793, 13861 and 13925. These 101 samples were not analyzed further and a reduced sample set of 95 cases and 134 controls was used for testing the association between haplotype and BSE status. No haplotype associations with BSE disease status were found within the region of low LD. However, within the region of high LD, one haplotype (Table 3-2, haplotype 8) was significantly over represented in unaffected animals (p = 0.005). This haplotype has been previously identified [20] and contains insertion alleles for both indel 23 and 12, and is tagged by the 4136 htSNP (Figure 3-1, Table 3-2).

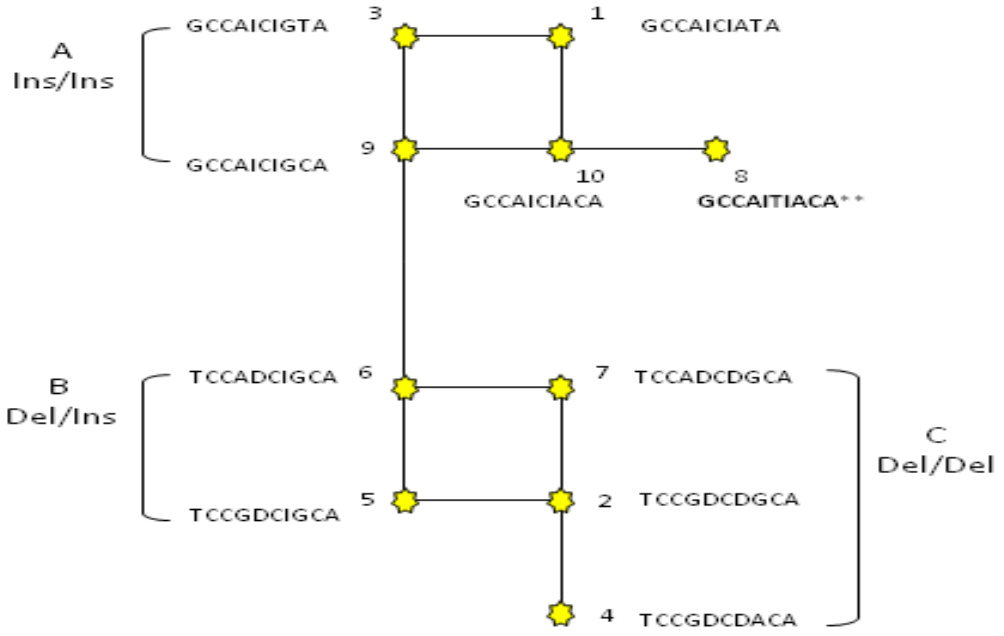


Figure 3-1. Median-joining network of haplotypes identified in the *PRNP* region of high LD (Haplotypes SNP449 | SNP1392 | SNP1576 | SNP1701 | indel23 | SNP4136 | indel12 | SNP4732| SNP4776| SNP6811 A) Haplotypes with insertion alleles for both the 23 and 12-bp indels). ** denotes haplotype 8 within this haplotype block that is significantly association (p = 0.005) with unaffected BSE animals. B) Haplotypes with the deletion allele for the 23-bp indel and the insertion allele for the 12bp-indel. C) Haplotypes with deletion alleles for both indels. Briefly, the haplotypes within the brackets only differ from the joining haplotype by one SNP. For example within bracket A haplotype 3 differs from haplotype 9 at SNP 4776, where haplotype 3 is a T and 9 is a C.

Haplotype		Frequency of BSE affected	Frequency of unaffected	*p* value
1	G C C A I C I A T A	0.132	0.130	0.968
2	T C C G D C D G C A	0.542	0.481	0.359
3	G C C A I C I G T A	0.0185	0.0156	0.871
4	T C C G D C D A C A	0.0629	0.0167	0.106
5	T C C G D C I G C A	0.0135	0.0223	0.603
6	T C C A D C I G C A	0.0355	0.0472	0.652
7	T C C A D C D G C A	0.0561	0.0491	0.817
8	G C C A I T I A C A	0.0332	0.123	0.00504**
9	G C C A I C I G C A	0.0554	0.0934	0.256
10	G C C A I C I A C A	0.0506	0.0214	0.272

Table 3-2. Analysis of network one haplotype block consisting of 8 htSNPs and 2 indels. Haplotype block for SNP 449 | SNP 1392 | SNP 1576 | SNP 1701 | indel 23 | SNP 4136 | indel 12 | SNP 4732| SNP 4776| SNP 6811 ** denote significance of $p < 0.01$.

The association between haplotype 8 and BSE (shown in Figure 3-1) was tested in comparison with haplotypes formed from the 23- and 12-bp indels alone. The alleles at these two indels were tightly linked and the haplotype defined by the insertion allele of the 23-bp indel, and the deletion allele of 12-bp indel (ID) was not observed, while the deletion/insertion (DI) haplotype was at low frequency in both the cases and controls. The deletion/deletion (DD) haplotype was the most frequent haplotype in both case and control populations, with a significantly higher frequency in the cases than controls ($p = 0.032$). Conversely, the insertion/insertion (II) haplotype was at significantly higher frequency in the controls than in the cases ($p = 0.038$) (see Table 3-3).

Haplotype	Frequency in BSE affected	Frequency in BSE unaffected	*p* value
II	0.283	0.409	0.038*
DI	0.0588	0.0682	0.764
DD	0.658	0.523	0.032*

Table 3-3. Frequency of 23-bp and 12-bp indel haplotype block. * Denotes significance of $p<0.05$

Although, the 23 and 12-bp indel haplotype was shown to be associated with BSE disease the level of significance was lower than for haplotype 8, which also contained htSNP 4136, suggesting an effect of the haplotype and that the indels alone are not

responsible for the association observed (see Figure 3-1). The haplotype which was significantly associated with the absence of disease contained the insertion alleles of the 23 and 12-bp indels in addition to the T allele of htSNP 4136. The T allele further separates the insertion/insertion haplotype into two haplotypes (Table 3-4), one of which is more tightly associated with classical BSE resistance than the haplotype defined by the insertion/insertion alleles alone (Tables 3-2, 3-3, and 3-4).

Locus	Haplotype				Frequency of BSE affected	Frequency of unaffected	p value
1	A	I	T	I	0.0373	0.173	0.000114***
2	A	I	C	I	0.246	0.236	0.854
3	A	D	C	I	0.0417	0.0448	0.905
4	G	D	C	I	0.0173	0.0234	0.731
5	A	D	C	D	0.076	0.0688	0.833
6	G	D	C	D	0.582	0.454	0.0472*

Table 3-4. Haplotype analysis results with reduced htSNP set. Haplotype defined by the alleles of SNP 1701 | indel 23 |SNP 4136 | indel 12 *** denotes significance of $p<0.001$

The alleles of htSNP 449 and those of the 23-bp indel were found to be tightly linked in this study (Tables 3-1 and 3-3), whereas the three htSNPs (1392, 1567, and 6811) had very low minor allele frequencies in both BSE cases and controls (MAF < 0.05) making the power of testing the association of these loci with disease status low. Therefore, the htSNPs with a low minor allele frequency were excluded from further haplotype analysis. Haplotypes with and without the alleles of htSNP 449 and the 23-bp indel were independently tested for association with BSE (Additional Table 3-2). A haplotype defined by a minimal set of four SNP alleles (1701/23-bp indel/4136/ 12-bp indel) was found to be significantly associated with BSE status (Additional Table 3-2 and Table 3-4). The haplotype characterized by allele A at SNP 1701, insertion allele at the 23-bp indel, T at SNP 4136 and insertion in the 12-bp indel was significantly over represented in unaffected animals ($p = 0.000114$, see Table 3-4). Whereas the haplotype characterized by the G allele at SNP 1701, deletion at the 23-bp indel, C allele at SNP 4136 and deletion in the 12-bp indel was over represented in BSE affected animals although the level of significance was lower ($p = 0.0472$).

3.3 Discussion

A haplotype defined by the alleles of four polymorphisms in the region of high LD within the *PRNP* gene was found to be associated with reduced incidence classical BSE. The haplotype includes insertion alleles at the 23 and 12-bp indels which have previously been shown to be associated with classical BSE [14-17]. The haplotypes also contains htSNP 1701 and 4136, the latter showing the strongest individual association with BSE, with the "T" allele associated with control in this study. The haplotype is in a region of high LD and the extent of this haplotype is unknown. Nonetheless, this finding indicates that a genetic determinant in or near the region of high LD in the *PRNP* has an effect on resistance or susceptibility to classical BSE.

The 23 and 12-bp indel alleles of *PRNP* were tested independently of the htSNPs for an association with classical BSE, as these loci have been associated with BSE status in previous studies [16-18]. In the present study, a haplotype defined by the insertion alleles of both 23- and 12-bp indels (the insertion/insertion haplotype) was significantly associated with unaffected animals ($p = 0.038$), and the deletion/deletion haplotype was associated with BSE affected animals ($p = 0.032$), which is in agreement with the previous reports. However this association had a lower level of significance than haplotype 8 which included the 4136 htSNP. Both the indels contain transcription factor binding sites: the 23-bp indel insertion allele contains a repression factor RP-58 binding site and the insertion allele of the 12-bp indel contains a SP-1 binding site [26]. The presence of the 12-bp insertion allele may disrupt a coordinated regulation of the prion gene [27]. In addition, different indel haplotypes within the *PRNP* gene have been shown to have different levels of expression in cell culture, with higher expression associated with the deletion/deletion allele compared to the insertion/insertion allele [17]. However, it is not known if the level of expression *per se* has an effect on BSE susceptibility.

PRNP htSNP 4136 is located in the promoter region exactly 14 bases upstream from exon 1 (GenBank file DQ457195). It has been shown that expression of the bovine prion protein gene requires interaction between the promoter and the first intron [25]. It was therefore possible that this htSNP had an effect on promoter activity. Thus bioinformatic analysis of sequence containing the htSNP 4136 [C/T] was performed using ConSite (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/). This analysis identified potential changes in transcription factor binding sites: the sequence with the

more common C allele, which is associated with disease, has a putative c-Fos binding site, whereas, the T allele has a putative NF-kappaB binding site (see Figure 3-2 and Table 3-5). c-Fos can suppress the expression of c-Jun/ATF2 which promotes neuronal apoptosis [28], while NF-KappaB may have a role in prion diseases through the inflammatory response [29-32]. It is not know if these variations are biologically relevant or have an effect on the course of BSE infection. Investigation of the effect of transcription factor binding activities in the promoter region of the bovine *PRNP* from a variety of different haplotype sequences found in the *Bos taurus* population would be interesting.

Figure 3-2. Sequence for haplotype tagged SNP 4136 [C/T] binding sites. The * denotes the position of htSNP 4136 and where the alleles differ.

SNP	Transcription factor	Sequence	Score	Strand
4136_C	c-FOS	GCGACTCA	7.051	+
	c-FOS	CGACTCAC	7.633	-
	Snail	CACCCG	6.063	-
4136_T	NF-kappaB	GCGACTTACC	7.333	+
	c-Rel	GCGACTTACC	6.623	-

Table 3-5. Putative transcription binding sites for htSNP 4136 C and T allele.

The htSNP 13861 was significantly associated with BSE disease incidence (Bonferroni corrected *p* value 0.027). A significant association (Bonferroni corrected *p* value 0.016) with BSE status was also found in an additional samples set of related BSE cases and controls (Additional Table 3-1). The htSNP 13861 is within the second intron and immediately 5' of a indel that consists of alleles with multiple G's and although that was difficult to resolve by sequencing may serve as a marker for the multi-G indel.

Although this region of the *Bos taurus PRNP* gene has not been previously identified as containing regulatory elements [25] many purine-rich regions have been shown to have regulatory function. Specifically the GGG triplet is a common splicing control element (review by [33]). Further, heterogeneous nuclear ribonucleoprotein (hnRNP) A2/B1, a gene that is associated with pre-mRNA processing, has been previously identified as an interactive partner of the prion protein [34]. It is not known if the [C/G] htSNP 13861 which is located proximally to a multiple G indel is associated with *PRNP* splicing control or if there may be a relationship that with BSE disease.

This study further defines the association of the 23 and 12-bp indel with classical BSE resistance previously described [17-19]. Specifically when htSNPs 1701 and 4136 are included in the haplotype the significance of the association increases from $p = 0.038$ to $p = 0.00011$. The association of this haplotype with classical BSE has not been previously reported. However, in a previous study [22], a haplotype in the low LD region of *PRNP* was associated with atypical BSE (BASE). The haplotype associated with atypical BSE was not found to be significantly associated with classical BSE in this study. This is consistent with a previous report [35] which suggested that variation in the *PRNP* gene that may influence classical BSE susceptibility are not associated with other TSE in cattle.

While the origins of classical and atypical BSEs are not known they, in common with all TSEs, require the prion protein for the manifestation of disease. The present study and others have shown that variations in *PRNP* are associated with either incidence of classical or atypical BSE [16-19, 5, 22], although the extensive LD within some regions of bovine *PRNP* make it difficult to determine which, if any of the variations have a functional role. Variations in PRNP are important factors for understanding susceptibility of BSE and potentially managing bovine TSE in cattle. The results reported here better define the variations within the bovine *PRNP* gene that have an effect on susceptibility or resistance to BSE and contribute to an improved definition of the genetic factors involved in the disease.

In conclusion a *PRNP* haplotype was identified that associated with classical BSE incidence. The haplotype is contained within known region of high linkage disequilibrium in *Bos taurus* cattle. The haplotype may be linked with a genetic determinant in or near *PRNP* that influences classical BSE susceptibility.

3.4 Materials and Methods

This study used DNA, extracted from blood samples of 333 Holstein cows from the UK, of which 149 were BSE cases and 184 were unaffected controls. BSE positive cattle were examined by qualified veterinarians and their BSE status was subsequently confirmed post-mortem by histology (by the Veterinary Laboratories Agency, New Haw, Surrey, UK). Control animals did not exhibit any BSE symptoms at the time of collection and were age matched with BSE cases from the same farm. Included in the control group were 15 BSE negative animals, confirmed by post-mortem histology. The case and control blood samples were all collected by the UK Veterinary Investigation Service and local veterinary practitioners in the mid 1990s in Southern England, at the peak of the BSE epidemic. The family-based samples set consisted of 302 BSE affected and 179 unaffected half-sib Holsteins from six sire families. All the BSE affected and unaffected cattle within one family were paternal half sibs from the designated sire but with different dams. DNA samples were not available from any of the sires or dams.

Genomic DNA was isolated from blood using a high salt phenol/chloroform extraction method as described by Sherman *et al.* [36]. Genotyping was performed using an oligonucleotide pool assay (OPA) as described by McKay and others [37]. This OPA comprises SNPs covering all chromosomes and included 15 of the 19 *PRNP* htSNPs reported by Clawson *et al.* [20]. An Illumina GoldenGate assay [38] was performed using the OPA and genotypes determined using an Illumina BeadScan (Illumina Inc., San Diego, CA) and the Illumina BeadStudio software. The four *PRNP* htSNPs, in the *PRNP* promoter region that were not present in the OPA were genotyped by PCR amplification of the region and Sanger sequencing as described by Clawson *et al.* [20,39]. Genotypes for the 12 and 23-bp indel were identified by PCR and the alleles were resolved on a 3% agarose gel stained with ethidium bromide. The multiplex PCR reaction which included 10 µl of genomic DNA, forward primer CTCGGTTTTACCCTCCTGG and the reverse primer GGCTAGATTCCTACACACCACC for the 12-bp indel, the forward primer CCTGATTTTCAAGTCCTCCCAG and the reverse primer TTATGCCCATGAATTGTGTAGGC for the 23-bp indel was carried out following the manufacturer's protocols (Accuprime *Taq* DNA polymerase system, Invitrogen). The

thermocycling conditions were: an initial denaturation of 2 minute at 94˚C, followed by thirty cycles at 94˚C for 2 min, 54˚C for 30 sec. and 68˚C for 1 min.

PLINK v1.04 [23] was used to phase haplotypes and perform all of the statistical analysis. The data from the case-control sample set were analyzed using the basic case-control association (χ^2) test and the haplotype-based association test was used for all of the haplotype analysis.

3.6 References

1. Prusiner SB (1998) Prions. Proc Natl Acad Sci U S A. 23: 13363-13383.
2. Mabbott NA, MacPherson GG (2006) Prions and their lethal journey to the brain. Nat Rev Microbiol 4: 201-211.
3. Prusiner S, Hsiao K (1994) Human prion diseases. Ann Neurol 35: 385-395.
4. Tester S, Juillerat V, Doherr MG, Haase B, Polak M et al. (2009) Biochemical typing of pathological prion protein in aging cattle with BSE. Virol J 6: 64.
5. Bruce ME, Will RG, Ironside JW, McConnel I, Drummond D et al. (1997) Transmission to mice indicste that ‘new variant’ CJD is caused by the BSE agent. Nature 389:498-501.
6. Hill A, Desbruslais M, Joiner S, Sidle K, Gowland I et al. (1997) The same prion strain causes vCJD and BSE. Nature 389: 448-450.
7. Collinge J, Sidle KCL, Meads J, Ironside J, Hill AF (1996) Molecular analysis of prion strain variation and the aetiology of ‘new variant’ CJD. Nature 383: 685–690.
8. Richt JA, Hall SM (2008) BSE case associated with prion protein gene mutation. PLoS Pathogens 4:9 e1000156.
9. Stack MJ, Focosi-Snyman R, Cawthraw S, Davis L, Chaplin MJ et al. (2009) Third atypical BSE case in Great Britain with an H-type molecular profile. Vet Rec 20: 605-606.
10. Weissmann C, Flechsig E (2003) PrP knock-out and PrP transgenic mice in prion research. Br Med Bull 66: 43-60.
11. Gambetti P, Kong Q, Zou W, Parchi P, Chen SG (2003) Sporadic and familial CJD: classification and characterisation. Br Med Bull 66: 213-39.
12. Laegreid WW, Clawson ML, Heaton MP, Green BT, O'Rourke KI et al. (2008) Scrapie resistance in ARQ sheep. J Virol 20: 10318-10320.

13. Alvarez L, Arranz JJ, San Primitivo F (2006) Identification of a new leucine haplotype (ALQ) at codon 154 in the ovine prion protein gene in Spanish sheep. J Anim Sci 2: 259-265.
14. Blanchong JA, Heisey DM, Scribner KT, Libants SV, Johnson C et al. (2009) Genetic susceptibility to chronic wasting disease in free-ranging white-tailed deer: complement component C1q and Prnp polymorphisms. Infect Genet Evol 6: 1329-35.
15. Wilson GA, Nakada SM, Bollinger TK, Pybus MJ, Merrill EH et al. (2009) Polymorphisms at the PRNP gene influence susceptibility to chronic wasting disease in two species of deer (Odocoileus Spp.) in western Canada. J Toxicol Environ Health A 17-18: 1025-1029.
16. Sander P, Hamann H, Pfeiffer I, Wemheuer W, Brenig B et al. (2004) Analysis of sequence variability of the bovine prion protein gene (PRNP) in German cattle breeds. Neurogenetics 1: 19-25.
17. Sander P, Hamann H, Drögemüller C, Kashkevich K, Schiebel K et al. (2005) Bovine prion protein gene (PRNP) promoter polymorphisms modulate PRNP expression and may be responsible for differences in bovine spongiform encephalopathy susceptibility. J Biol Chem 45: 37408-37414.
18. Juling K, Schwarzenbacher H, Williams JL, Fries R (2006) A major genetic component of BSE susceptibility. BMC Biol 4: 33.
19. Haase B, Doherr MG, Seuberlich T, Drögemüller C, Dolf G et al. (2007) PRNP promoter polymorphisms are associated with BSE susceptibility in Swiss and German cattle. BMC Genet Apr 8:15.
20. Clawson ML, Heaton MP, Keele JW, Smith TPL, Harhay GP et al. (2006) Prion gene haplotypes of U.S. cattle. BMC Genet 7: 51.
21. Heaton MP, Keele JW, Harhay GP, Richt JA, Koohmaraie M et al. (2008) Prevalence of the prion gene E211K variant in U.S. cattle. BMC Vet Res 4: 25.

22. Clawson ML, Richt JA, Baron T, Biacabe AG, Czub S et al. (2008) Association of a bovine prion gene haplotype with atypical BSE. PLoS ONE 3: e1830.
23. Hills D, Comincini S, Schlaepfer J, Dolf G, Ferretti L et al. (2001) Complete genomic sequence of the bovine prion gene (PRNP) and polymorphisms in it promoter region. Anim Genet 32: 231-232.
24. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA et al. (2007) PLINK: A tool set for whole-genome Association and Population-Based Linkage Analyses. Am. J. Hum. Genet 81:559–575.
25. Murdoch BM, Clawson ML, Laegreid WW, Stothard P, Settles M (2010) A 2 cM genome-wide scan of European Holstein cattle affected by classical BSE. BMC Genet 11: 20.
26. Inoue S, Tanaka M, Horiuchi M, Ishiguro N, Shinagawa M (1997) Characterization of the bovine prion protein gene: The expression requires interaction between the promoter and intron. J Vet Med Sci 3: 175-183.
27. Xue G, Sakudo A, Kim CK, Onodera T (2008) Coordinate regulation of bovine prion protein gene promoter activity by two Sp1 binding site polymorphisms. Biochem Biophys Res Commun 4: 530-535.
28. Yuan Z, Gong S, Luo J, Zheng Z, Song B et al. (2009) Opposing roles for ATF2 and c-Fos in c-Jun-mediated neuronal apoptosis. Mol Cell Biol 9: 2431-2442.
29. Kim JI, Ju WK, Choi JH, Choi E, Carp RI et al. (1999) Expression of cytokine genes and increased nuclear factor-κB activity in the brains of scrapie-infected mice. Brain Res Mol Brain Res 73: 17-27.
30. Bourteele S, Oesterle K, Weinzierl AO, Paxian S, Riemann M et al. (2007) Alteration of NF-kB activity leads to mitochondrial apoptosis after infection with pathological prion protein. Cell Microbiol 9: 2202-2217.

31. Julius C, Heikenwalder M, Schwarz P, Marcel A, Karin M et al. (2008) Prion propagation in mice lacking central nervous system NF-kB signalling. J Gen Virol 89:1545-1550.
32. Mattson MP (2005) NF-kB in the survival and plasticity of neurons. Neurochem Res 30: 883-893.
33. Majewski J, Ott J (2002) Distribution and characteristics of regulatory elements in the human genome. Genome Res Lett 12: 1827-1836.
34. Strom A, Diecke S, Hunsmann G, Stuke AW (2006) Identification of prion protein binding proteins by combined use of far-western immunoblotting, two dimensional gel electrophoresis and mass spectrometry. Proteomics 1: 26-34.
35. Brunelle BW, Hamir AN, Baron T, Biacabe AG, Richt JA et al. (2007) Polymorphisms of the prion gene promoter region that influence classical bovine spongiform encephalopathy susceptibility are not applicable to other transmissible spongiform encephalopathies in cattle. J Anim Sci 85: 3142-3147.
36. Sherman L, Nkrumah D, Murdoch B, Moore S (2008) Identification of polymorphisms influencing feed intake and efficiency in beef cattle. Anim Genet 3: 225-231.
37. McKay SD, Schnabel RD, Murdoch BM, Matukumalli LK, Aerts J et al. (2007) Whole genome linkage disequilibrium maps in cattle. BMC Genet 8: 74.
38. Oliphant A, Barker D, Stuelpnagel J, Chee M. (2002) BeadArray technology: enabling an accurate, cost-effective approach to high-throughput genotyping. Biotechniques 56-8: 60-61.
39. Clawson ML, Heaton MP, Keele JW, Smith TP, Harhay GP et al. (2008) A sequencing strategy for identifying variation throughout the prion gene of BSE-affected cattle. BMC Res Notes 1: 32.

Chapter 4
High Density Genome-wide Scan Identifies Loci Associated with Classical BSE Incidence

4.1 Background

Transmissible spongiform encephalopathies (TSEs), also known as prion diseases, are a group of mammalian neurodegenerative diseases that are invariably fatal [1]. TSEs are unique in that they can manifest through acquired, inherited, or sporadic origins [2]. One of the common characteristics of these diseases is the accumulation of abnormally folded prion protein within the central nervous system were the proteinateous infectious agent mediates conformational change in the native host protein to a protease-resistant form of the prion protein [1,2]. The central nervous system (CNS) is the primary tissue where the disease form of the protein accumulates and the diseases manifest. Classical bovine spongiform encephalopathy (cBSE) is an acquired TSE that spreads through the consumption of meat and bone meal contaminated with the infectious prion agent. Dietary exposure to classical BSE is the suspected cause of the human TSE, variant Creutzfeldt-Jakob Disease (vCJD) [3-5].

The prion protein is a glycosyl-phosphatidylinositol (GPI) anchored protein that has a native form (PrP^{c}) for which the secondary structure consists mainly of alpha-helices. The disease-associated or misfolded form, PrP^{Res}, has a substantial increase in the beta-pleated sheets content and reduction of the alpha-helices in comparison to the native form [6]. This altered confirmation is associated with an increased resistance to digestion with proteinase K [7]. Furthermore, the presence of PrP^{Res} promotes the conversion of native PrP^{c} to PrP^{Res} via a mechanism that is to date not completely understood [8]. The prion protein is essential for the development of TSEs [9], and variation in the prion gene (*PRNP*) has been associated with TSE susceptibility in humans, mice, sheep, deer and cattle [review 10].

Specific *PRNP* alleles of non-synonymous polymorphisms in humans and sheep are associated with acquired TSE susceptibility [11-14]. In general this is not the case in cattle, although alleles and haplotypes containing insertion/deletion polymorphism in the bovine *PRNP* promoter region have been associated with incidence of classical bovine

spongiform encephalopathy (BSE) [15-19]. However, previous QTL studies in sheep [20-22] and mice [23, 24] have identified regions, other than the prion gene, that contribute to TSE disease progression. Further, in cattle genetic studies of classical BSE cattle identified loci other than *PRNP* (located on chromosome 13 at 47.2Mb) that associate with disease incidence [19, 25, 26].

The objective of this study was to genotype a high-density set of loci throughout the bovine genome and test for an association with classical BSE. The animal samples used in this study were BSE case and control female European Holstein cattle which contracted BSE in the mid-1990's. The BSE cattle were identified on commercial farms where the most likely source of disease was through consumption of contaminated feed prior to, and during the United Kingdom's BSE epidemic. The BSE cases were paired with age matched controls from same calving season and from the same farm for a case-control analysis association.

4.2 Results

4.2.1 Case-control association analysis for BSE Incidence

The case-control samples were comprised of 143 BSE case and 173 control animals. The control samples include at least one animal collected from the same farm as each of the BSE cases as well as the controls for, and 15 BSE negative animals. By using (case-control) animals, this study examined within-breed (Holstein) SNP association with disease incidence. The bovine SNP50 beadchip used in this study consisted of greater than 58K SNPs dispersed across the genome at an approximate average gap interval of 37 kb [27]. Of these, an average 48,053 SNPs passed quality control measures with an average genotyping success rate of 0.966. The genotyping data on these animals was analyzed for an association with disease status using the case-control allelic test within the PLINK software [28]. An allelic association test analysis (316 animal across 48,053 SNPs after quality control) revealed 1 SNP with a p = 3.09E-5. This study identified a polymorphic locus on chromosome 1 at 29,147,078 bp with a moderate significant, as defined by Wellcome trust [29], association (p=3.09E-5) and over represented in BSE affected animals. Further, a locus on chromosome 14 at 43,984,235 bp is suggestively associated (p=5.24E-05) and over represented in BSE affected animals (see Table 4-1).

Loci on chromosomes 5, 10, 14, 15, 17, 21 and 24 have P values ($p \leq 5E\text{-}4$) that trend towards significance.

CHR	Position (bp)	SNP	A1	A2	F_A	F_U	*p-value*	OR
1	29,147,078	BTB-01333940	A	G	0.150	0.052	3.09E-05*	3.225
5	8,363,297	BFGL-NGS-117339	G	A	0.277	0.421	3.66E-04	0.528
10	29,247,747	BTB-01261410	A	C	0.521	0.376	2.55E-04	1.808
14	43,984,235	Hapmap57707-rs29024913	A	G	0.476	0.318	5.24E-05	1.945
14	45,169,821	ARS-BFGL-BAC-23887	G	A	0.549	0.402	2.23E-04	1.812
15	35,715,314	ARS-BFGL-NGS-65959	A	G	0.380	0.243	1.91E-04	1.914
17	40,958,294	Hapmap43651-BTA-67415	A	G	0.311	0.451	3.34E-04	0.550
21	13,242,468	Hapmap42965-BTA-53492	G	A	0.525	0.379	2.40E-04	1.810
24	1,490,858	ARS-BFGL-NGS-11247	C	A	0.199	0.335	1.36E-04	0.494

Table 4-1: Allelic association analysis of BSE using 50K SNP chip. The Frequency of affected animals is abbreviated F_A, frequency of unaffected animals is abbreviated F_U and OR is the abbreviation for Odds Ratio. * denotes moderate significance as defined by Wellcome Trust [29].

While SNPs in close proximity may be link others may exhibit linkage disequilibrium (LD). Linkage disequilibrium is the non random association of alleles within a population. SNPs linked with significant loci provide supporting evidence that the observed associations are valid. This relationship is easily observed when the p values of the case-control association test are plotted by their base pair position (Btau4.0) (see Figure 4-1). Here it is apparent that the significant SNP on chromosome 14 has other SNPs which are linked and provide support of the association.

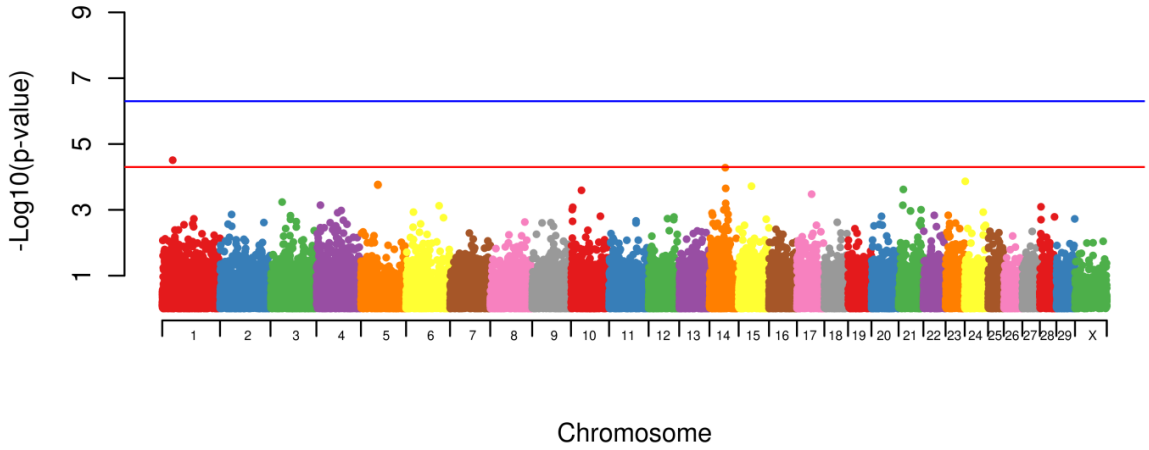


Figure 4-1: Genome-wide plot of -log10 p-values of loci from BSE case-control allelic association analysis. Bovine autosomes 1–29 and Chromosome X are separated by colour. The horizontal red line is drawn at -log10 ($5x10^5$) and the horizontal blue line is drawn at -log10 ($5x10^7$) to show significant loci at the moderate and strong levels respectively.

This data set was analyzed with a best fit model test where the standard allelic, trend, dominance, recessive and genotypic association tests were performed, and the test with the lowest p value was reported. The best fit model identified, in addition to the allelic association, a total of three recessive SNPs and two dominant SNPs. Specifically, a SNP on chromosomes 2, 3 and 21 exhibited a recessive suggestive association with BSE incidence and loci on chromosomes 14 and 24 exhibited a dominant suggestive association the BSE incidence. The dominant locus indentified on chromosome 14 at 45,169,821 bp is in close proximity to the significant allelic loci observed at 43,984,235 bp (see Table 4-2).

CHR	Position (bp)	SNP	A1	A2	TEST	AFF	UNAFF	*p-value*
1	29,147,078	BTB-01333940	A	G	ALLELIC	43/243	18/328	3.09E-05*
2	108,556,922	BTA-48618-no-rs	G	A	REC	39/103	18/155	9.10E-05
3	32,408,065	ARS-BFGL-NGS-104661	A	C	REC	10/133	41/132	5.87E-05
14	43,984,235	Hapmap57707-rs29024913	A	G	ALLELIC	136/150	110/236	5.24E-05
14	45,169,821	ARS-BFGL-BAC-23887	G	A	DOM	121/22	112/61	6.44E-05
21	13,242,468	Hapmap42965-BTA-53492	G	A	REC	38/105	17/156	9.30E-05
24	1,490,858	ARS-BFGL-NGS-11247	C	A	DOM	48/95	97/76	6.45E-05

Table 4-2: Model association analysis of BSE using 50K SNP chip. Duplicate loci results were removed. * denotes moderate significance as defined by Wellcome Trust [29].

4.3 Discussion

4.3.1 Classical BSE

Classical BSE disease is a difficult phenotype to test for genetic associations. Animals that have developed BSE are clearly susceptible, however, those which are clinically healthy are difficult to assess. In the present study, clinically healthy animals were used as controls, however, these animals may have been incubating disease or may not have ingested enough infectious agent to become symptomatic, or alternatively, they may in fact have been resistant to disease. Further, because classical BSE is an acquired disease who infectivity enters through the intestine and is primarily manifested in the brain many, multiple genes may be involved in the disease process. Another consideration is that classical BSE is a complex trait which may be more consistent with interactive and possibly subtle effects of multiple contributing loci. Therefore, multiple testing corrections applied to results such as these may be overly prone to type II errors (i.e. discarding real associations). Consequently, it is important to examine the results for evidence of associations between disease and genetic loci, as discussed below.

4.3.2 Associations with BSE affected animals

The results from this study indicate several loci on four different chromosomes are associated with BSE disease. Most notable is a single nucleotide polymorphism on

chromosome 1 at 29.15Mb is significantly (p=3.09E-05) associated with BSE disease. Another locus, on chromosome 14 at 32.4Mb, that is suggestive of significance (p=5.87E-05) was identified to be associated with BSE disease. The results on chromosome 14 are further supported by a SNP 1.1Mb away that is also associated both in a dominate model (6.44E-05) and in an allelic model (2.23E-04) with disease. Additionally, there is evidence of suggestive recessive associations on chromosome 2 at 108.6Mb (9.10E-05), and chromosome 21 at 13.2Mb (9.30E-05). Further, loci on chromosome 10 and 15 are trending towards significance (p=2.55E-04 and 1.91E-04 respectively).

4.3.3 Associations with unaffected BSE animals

There is evidence of associations with absence of disease in control animals on two chromosomes. A novel locus on chromosome 3 at 32.4Mb is recessively associated (p=5.87E-05) in control animals. Whereas the locus on chromosome 24 is dominantly associated (p=6.45E-05) to control animals. Loci trending toward significant associations with disease absence were identified on chromosome 5 at 8.3Mb and 17 at 40.9Mb (p=3.66E-04 and p=3.34E-04) respectively.

4.3.4 Shared regions identified in previous analysis and studies in Holstein cattle

It is perhaps not surprising that many of the significant findings identified in this study overlap with previous studies as the same samples were used [19]. However the most convincing evidence of associations are those identified with a different sample set in previous studies [25, 26, 30]. The locus on chromosome 1 is associated with most significance in this study. An association on this chromosome was first identified by Zhang et al. [26] and although the position of the peak LOD is at the other end of the chromosome it is within the confidence interval. The region identified in this study on chromosome 14 was also observed in a previous study [19], although the same samples were analyzed a different set of SNPs were utilized in the study and therefore represent a technical verification. With the exception of the locus on chromosome 3 all of the other suggestive and trending associations identified in this study provide confirmation of

associations identified with other Holstein animals in previous studies [25, 26] as well as in a human study [30] (see Table 4-3). The association identified on chromosome 3 is novel and has not been previously identified in cattle.

CHR	Position Mb	SNP	*p-value*	Previously observed	Study	Location Mb
1	29.1	BTB-01333940	3.09E-05	Yes	[19,26]	89.7
2	108.6	BTA-48618-no-rs	9.10E-05	Yes	[19]	37.1
3	32.4	ARS-BFGL-NGS-104661	5.87E-05	no		
5	8.4	BFGL-NGS-117339	3.66E-04	Yes	[19,25]	30.6
10	29.2	BTB-01261410	2.55E-04	Yes	[19,25]	21.1
14	44.0	Hapmap57707-rs29024913	5.24E-05	Yes	[19,30]	44.0
15	35.7	ARS-BFGL-NGS-65959	1.91E-04	Yes	[19]	35.7
17	41.0	Hapmap43651-BTA-67415	3.34E-04	Yes	[19,26]	44.2
21	13.2	Hapmap42965-BTA-53492	9.30E-05	Yes	[19]	33.2
24	1.5	ARS-BFGL-NGS-11247	6.45E-05	Yes	[19]	18.5

Table 4-3: Comparison of this study and those reported previously.

4.3.5 Candidate genes

One of the most promising of the candidate genes is hypothetical gene LOC521010, similar to FK506 binding protein 2 located on chromosome 1 at 29,316,874bp. This gene encodes a protein that is a member of the immunophilin protein family and is involved in basic cellular processes involving protein folding. This encoded protein is a cis-trans prolyl isomerase that binds the immunosuppressants FK506 and rapamycin. The SNP located on BTA 14 (43.9Mb), associated with BSE disease case-control analysis, is in close proximity to the gene exostoses (multiple) 1, *EXT1*. McCormick et al., [31] showed that *EXT1* is an endoplasmic reticulum (ER)-resident type II transmembrane glycoprotein whose expression in cells results in the alteration of the synthesis and display of cell surface heparin sulfate glycosaminoglycans (GAGs). The N terminus of PrP contains a GAG-binding motif and it is thought that PrP binding of GAG is important in prion disease [31, 32]. Additional positional candidate genes of interest on the chromosomes with trending significant results include: leucine-rich repeats and immunoglobulin-like domains 2, LRIG2, a protein which participate in protein-protein interactions, Repulsive guidance molecule family member A, a glycosyl

phosphatidylinositol-anchored glycoprotein that functions as an axon guidance protein in the central nervous system.

4.4 Methods and Materials

4.4.1 Animal Information

This study used DNA, extracted from blood samples of female European Holstein cows collected in the mid 1990s in Southern England. The samples included 144 BSE affected (case) and 177 unaffected (control) animals. The control animals were contemporaries of the BSE cases and collected from the same farms. In addition 15 BSE negative, determined by post-mortem histology, and paired control animals were included in the control set. BSE positive cattle were examined by qualified veterinarians and their BSE status was subsequently confirmed post-mortem by histology (by the Veterinary Laboratories Agency, New Haw, Surrey, UK). Control animals were age matched with BSE cases and from the same calving season and farm. As such, the control animals are assumed to have been exposed to the same environment. The realtionship of the samples ca be observed in Figure 4-2.

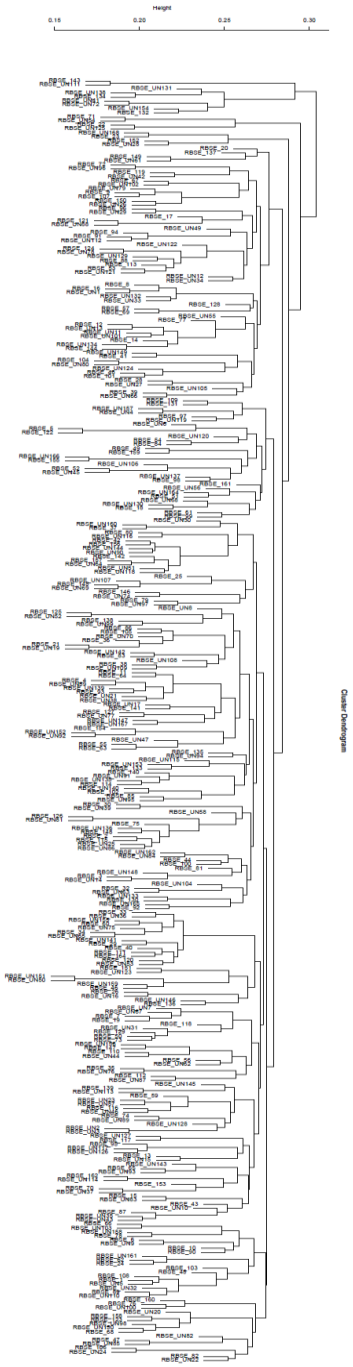


Figure 4-2: Graphical representation of the BSE Case and control sample relationship. Derived from genotypic data use R 2.9.0. cluster diagram statisical software.

4.4.2 Genotyping

In this study genotyping was performed using DNA from 312 BSE case and control Holstein animal. The bovine SNP50 beadchip was used and the assay was performed following the manufactures (Illumina Inc., San Diego, CA) recommended protocol. Briefly genomic DNA for each sample was whole-genome amplified and then fragmented, precipitated and re-suspended. The DNA fragments were hybridized to sequence specific 50mer probes bound to beads. The following base was added using an enzymatic single base extension with labeled nucleotide. The 12 sample chips were then imaged using a beadarray reader. With the use of BeadStudio software (Illumina Inc.) genotypes for all the samples were combined and processed through the automated genotype calling. The bovine SNP50 beadchip used in this study consisted of greater than 58K SNPs dispersed across the genome at an approximate average gap interval of 37 kb [27]. Of these, an average 48,053 passed quality control measures and had an average genotyping success rate of 0.966.

4.4.3 Data quality control

Quality control analyses were carried out by removing confounding effects prior to the data analysis. For the case-control sample set from the original 58,336 SNPs in the assay, 1389 SNPs were excluded based on assay failure, as observed by poor clustering of alleles using the BeadStudio software. The remaining 56,947 SNPs were submitted to the PLINK program [28] where 2,065 SNPs were removed because more than 10% of the samples failed to genotype at that locus (GENO >0.1). Additionally, 5 samples were removed due to low genotyping rate (MIND>0.1). A further 8,739 SNPs were removed due to a MAF <0.01. The remaining individuals that were included in the analyses had a mean genotyping rate of 0.966. Population stratification was tested for in this data set and none was observed (Figure 4-3).

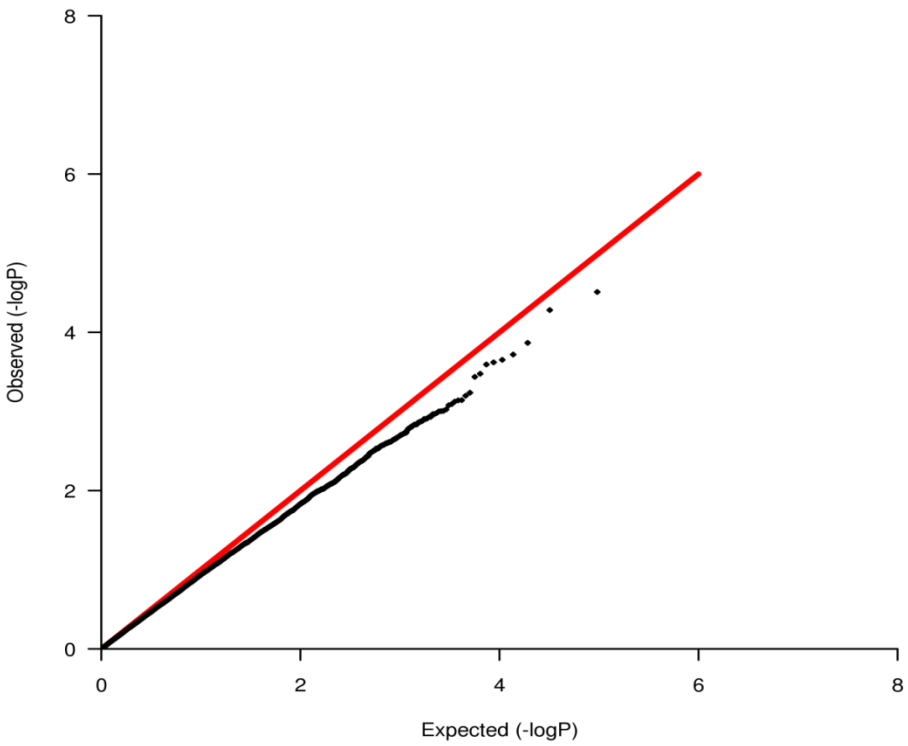


Figure 4-3: Quantile-quantile plot of p-values in BSE case-control samples. P-values from the corresponding allelic test for association plotted as $-\log_{10}$ values. The Q–Q plot shows no evidence of a deviation from the expected null distribution of P-values and thus no evidence of population substructure.

4.4.4 Statistical analysis

The PLINK software v1.04 was used to perform the statistical analysis [28]. The genotypic data from the BSE case and control sample set were analyzed using the basic case-control association (χ^2) test. The allelic association test, the dominance, and the recessive used one degree of freedom (d.f.). Uncorrected p-values of $<5X10^5$ were considered moderate evidence of statistical significance as established by the Wellcome trust case-control Consortium 2007. All of the plots and figure were produced using R statistical. Physical positions and alleles are expressed in terms of the forward strand of the reference genome (BTAU4.0, ftp://ftp.hgsc. bcm.tmc.edu/pub/data/Btaurus/).

4.5 Conclusion

This high density genome-wide scan identified a locus on bovine chromosome 1 that is significantly associated with BSE affected animals. Additionally loci on chromosomes 2, 14, and 21 also exhibited associations with BSE disease. Further, loci on chromosomes 3 and 24 exhibited associations with the absence of BSE disease. Moreover, these results suggest that loci, other than the prion gene can contribute to both the absence and presence to BSE disease.

4.6 References

1. Prusiner SB (1998) Prions. Proc Natl Acad Sci U S A. 23: 13363-13383.
2. Prusiner S, Hsiao K (1994) Human prion diseases. Ann. Neurol 35: 385-395.
3. Bruce ME, Will RG, Ironside JW, McConnell I, Drummond D, Suttie A, McCardle L, Chree A, Hope J, Birkett C, Cousens S, Fraser H, Bostock CJ (1997) Transmission to mice indicate that 'new variant' CJD is caused by the BSE agent. Nature 389: 498-501.
4. Hill A, Desbruslais M, Joiner S, Sidle K, Gowland I, Collinge J, Doey LJ, Lantos P (1997) The same prion strain causes vCJD and BSE. Nature 389: 448-450.
5. Collinge J, Sidle KCL, Meads J, Ironside J, Hill AF (1996) Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. Nature 383: 685–690.
6. Pan KM, Baldwin M, Nguyen J, Gasset M, Serban A, Groth D, Mehlhorn I, Huang Z, Fletterick RJ, Cohen FE, Prusiner SB (1993) Conversion of α-helices into β-sheets features in the formation of the scrapie prion proteins. Proc Natl Acad Sci USA 90: 10962-10966.
7. Cronier S, Gros N, Tattum MH, Jackson GS, Clarke AR, Collinge J, Wadsworth JDF (2008) Detection and characterization of proteinase K-sensitive disease-related prion protein with thermolysin. Biochem J, 416(2):297-305.
8. Atarashi R, Moore RA, Sim VL, Hughson AG, Dorward DW, Onwubiko HA, Priola SA, Caughey B (2007) Ultrasensitive detection of scrapie prion protein using seeded conversion of recombinant prion protein. Nat Methods 8:645-650.
9. Weissmann C, and Flechsig E (2003) PrP knock-out and PrP transgenic mice in prion research. Br Med Bull 66: 43-60.
10. Mabbott NA, MacPherson GG (2006) Prions and their lethal journey to the brain. Nat Rev Microbiol 3: 201-211.
11. Laegreid WW, Clawson ML, Heaton MP, Green BT, O'Rourke KI, Knowles DP (2008) Scrapie resistance in ARQ sheep. J Virol 82(20):10318-10320.

12. Alvarez L, Arranz JJ, San Primitivo F (2006) Identification of a new leucine haplotype (ALQ) at codon 154 in the ovine prion protein gene in Spanish sheep. J Anim Sci 84(2):259-265.
13. Collinge J, Sidle KCL, Meads J, Ironside J, Hill AF (1996) Molecular analysis of prion strain variation and the aetiology of ‘new variant’ CJD. Nature 383:685-690.
14. Gambetti P, Kong Q, Zou W, Parchi P, Chen SG (2003) Sporadic and familial CJD: classification and characterisation. Br Med Bull 2003, 66:213-239.
15. Sander P, Hamann H, Pfeiffer I, Wemheuer W, Brenig B, Groschup MH, Ziegler U, Distl O, Leeb T (2004) Analysis of sequence variability of the bovine prion protein gene (PRNP) in German cattle breeds. Neurogenetics 1: 19-25.
16. Sander P, Hamann H, Drögemüller C, Kashkevich K, Schiebel K, Leeb T (2005) Bovine prion protein gene (PRNP) promoter polymorphisms modulate PRNP expression and may be responsible for differences in bovine spongiform encephalopathy susceptibility. J Biol Chem 45: 37408-37414.
17. Juling K, Schwarzenbacher H, Williams JL, Fries R (2006) A major genetic component of BSE susceptibility. BMC Biol 4:33.
18. Haase B, Doherr MG, Seuberlich T, Drögemüller C, Dolf G, Nicken P, Schiebel K, Ziegler U, Groschup MH, Zurbriggin A, Leeb T (2007) PRNP promoter polymorphisms are associated with BSE susceptibility in Swiss and German cattle. BMC Genet Apr 8:15.
19. Murdoch BM, Clawson ML, Laegreid WW, Stothard P, Settles M (2010) A 2 cM genome-wide scan of European Holstein cattle affected by classical BSE. BMC Genet 11:20.
20. Laegreid WW, Clawson ML, Heaton MP, Green BT, O'Rourke KI, Knowles DP (2008) Scrapie Resistance in ARQ Sheep. J Virol 20: 10318-10320.
21. Alvarez L, Arranz JJ, San Primitivo F (2006) Identification of a new leucine haplotype (ALQ) at codon 154 in the ovine prion protein gene in Spanish sheep. J Anim Sci 2: 259-265.

22. Moreno CR, Cosseddu GM, Schibler L, Roig A, Moazami-Goudarzi K, Andreoletti O, Eychenne F, Lajous D, Schelcher F, Cribiu EP, Laurent P, Vaiman D, Elsen JM (2008) Identification of new quantitative trait loci (other than the PRNP gene) modulating the scrapie incubation period in sheep. Genetics 179(1):723-726.
23. Moreno CR, Lantier F, Lantier I, Sarradin P, Elsen JM (2003) Detection of new quantitative trait Loci for susceptibility to transmissible spongiform encephalopathies in mice. Genetics 165(4):2085-91.
24. Lloyd SE, Uphill JB, Targonski PV, Fisher EM, Collinge J.(2002) Identification of genetic loci affecting mouse-adapted bovine spongiform encephalopathy incubation time in mice. Nuerogenetics Oct 4(2): 77-81.
25. Hernández-Sánchez J, Waddington D, Wiener P, Haley CS, Williams JL (2002) Genome-wide search for markers associated with bovine spongiform encephalopathy. Mamm Genome 13(3):164-168.
26. Zhang C, de Koning DJ, Hernández-Sánchez J, Haley CS, Williams JL, Wiener P (2004) Mapping of multiple quantitative trait loci affecting bovine spongiform encephalopathy. Genetics 167(4):1863-1872.
27. Matukumalli LK, Lawley CT, Schnabel RD, Taylor JF, Allan MF, Heaton MP, O'Connell J, Moore SS, Smith TP, Sonstegard TS, Van Tassell CP (2009) Development and characterization of a high density SNP genotyping assay for cattle. PLoS One. 4(4):e5350. Epub 2009 Apr 24.
28. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC (2007) PLINK A tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559-575.
29. Wellcome Trust Case Control Consortium (2007) Genome-wide association study of 14,000 cases of seven common disease and 3,000 shared controls. Nature, 447, 661-678.
30. Mead S, Poulter M, Uphill J, Beck J, Whitfield J, Webb TE, Campbell T, Adamson G, Deriziotis P, Tabrizi SJ, Hummerich H, Verzilli C, Alpers MP, Whittaker JC, Collinge J (2009) Genetic risk factors for Creutzfeldt-Jakob disease: a genome-wide association study. Lancet Neurol. 8(1):25-26.

31. McCormick C, Leduc Y, Martindale D, Mattison K, Esford LE, Dyer AP, Tufaro F (1998) The putative tumour suppressor EXT1 alters the expression of cell-surface heparan sulfate. Nat Genet 19(2):158-161.
32. Yin S, Pham N, Yu S, Li C, Wong P, Chang B, Kang SC, Biasini E, Tien P, Harris DA, Sy MS (2007) Human prion proteins with pathogenic mutations share common conformational changes resulting in enhanced binding to glycosaminoglycans. Proc Natl Acad Sci USA 104(18):7546-7551.

Chapter 5
Summary and Future Directions

5.1 Summary

In the interest of understanding more about BSE my thesis research was aimed at examining genetic polymorphisms and their association with classical BSE. The first study used a panel of SNP markers that consisted of 3,072 SNPs dispersed across the genome at an approximate average interval of 2cM. The cattle used were female European Holstein cattle which contracted BSE in the late 1980's and early 1990's (family analysis) and in the mid-1990's (case-control analysis) and were identified on commercial farms where the most likely source of disease was through consumption of contaminated feed during the United Kingdom's BSE epidemic. The samples came from two sets of animals with different relatedness that allowed for the use of several statistical analyses to test the associations between SNP genotypes with classical BSE incidence. Through the use of both offspring from six sires (sib-TDT) and (case-control) Holstein animals, this study allowed us to examine associations' within-breed and within-family with disease incidence.

The results for this sib-TDT analysis (412 animals across 2,827 SNPs after quality control) identified 46 SNPs that passed the Bonferroni correction with a $p<0.05$ and 27 SNPs that passed the Bonferroni correction with a $p<0.01$. Although significant loci were identified on 18 different chromosomes the most significant of these loci ($P \leq E04$) were found on chromosomes 2, 5, 20, 21, and 24. The case control allelic analysis consisted of 320 animal across 2,872 SNPs after quality control) revealed 20 SNPs with a $p<0.01$, 14 SNPs with a $p<0.005$ and 6 SNPs with a $p<0.001$. Following correction for the false discovery rate, no significant associations at $p \leq 0.05$ genome-wide significance were identified in this data set. However, a single SNP had a suggestive association with BSE incidence on chromosome 14 ($p=7.25 \times 10^{-5}$). SNPs which did not quite reach this threshold but had a $p \leq 5 \times 10^{-4}$ were found on chromosomes 4, 10, 14, and 15.

Many of the SNPs included in the panel are in close proximity and are in LD in Holstein. Therefore, it was more appropriate to consider the results in terms of chromosomal regions instead of individual markers. Linkage disequilibrium will result in the alleles of several closely spaced SNPs being associated with disease status because they all fall on the same haplotype. Thus, it would be expected that several SNPs in LD with a locus involved in disease would show significant associations. Examples of this can be observed with the loci on BTA 2, 4, 15 and 21.

The large number of SNP markers and the two sets of animals used in this study make it the most comprehensive study to date to test genetic loci for an association with classical BSE in European Holstein cattle. The genome-wide scan of half sib families identified an association between the genetic loci on 18 chromosomes with BSE incidence in European Holstein cattle, including a region on BTA 20 associated with BSE incidence that has been reported in previous studies. The identification of markers within the same chromosomal regions, in both sets of samples, provides independent evidence for this regions association with the incidence of BSE in cattle.

Although this first study was completed with a well-defined panel of SNP markers new technology became available that allowed for a much higher density scan. Therefore we used the Bovine 50K SNP chip to test for associations with BSE disease at a much improved genomic resolution. An allelic association test analysis (316 animal across 48,053 SNPs after quality control) revealed 1 SNP with a $p \leq 5E\text{-}5$, 1 SNP with a p=5E-5. This study identified a polymorphic locus on chromosome 1 at 29,147,078 bp with a moderately significant, as defined by Wellcome trust [29], association (p=3.09E-5) and was over represented in BSE affected animals. Further, a locus on chromosome 14 at 43,984,235 bp is suggestively associated (p=5.24E-05) and over represented in BSE affected animals. Interestingly, this locus is just 82 bp away from the SNP identified in my previous analysis. Further loci on chromosomes 5, 10, 14, 15, 17, 21 and 24 have P values ($p \leq 5E\text{-}4$) that trend towards significance.

This data set was analysed with a best fit model test where the standard allelic, trend, dominance, recessive and genotypic association tests were performed, and the test with the lowest p value was reported. The best fit model identified, in addition to the allelic association, a total of three recessive SNPs and two dominant SNPs. Specifically,

SNPs on chromosomes 2, 3 and 21 exhibited a recessive suggestive association with BSE incidence and loci on chromosomes 14 and 24 exhibited dominant suggestive association the BSE incidence. The dominant locus indentified on chromosome 14 at 45,169,821 bp is in close proximity to the significant allelic loci observed at 43,984,235 bp. Together these results from the allelic association test and the recessive and dominance analysis confirm loci on chromosomes 1, 2, 5, 10, 14, 15, 17, 21, 24 and identified a novel locus on chromosome 3.

As a final component of my research project I examined the *PRNP* gene since more than 390 polymorphic alleles have been described in a 25-kb region of chromosome 13 containing this gene. This region contains distinct regions of high and low LD that appear conserved across many *Bos taurus* cattle populations. A region of high LD including the promoter region and importantly, both the 23- and 12-bp deletion alleles that associate with classical BSE susceptibility are contained in this region of high LD. The remainder of *PRNP*, including the entire coding region has relatively low LD. To account for the genetic architecture of the *PRNP* gene, a set of 19 haplotype tagging SNPs (htSNPs) described by Clawson et al. efficiently define haplotypes within and across each of the LD regions. These htSNPs were used for testing the association of *PRNP* haplotypes with classical BSE susceptibility. In this study, the 19 htSNPs, including the 12-bp, and 23-bp indel were used to test *PRNP* haplotypes for an association with classical BSE in European Holstein cattle. A haplotype associated with classical BSE resistance was identified.

A *PRNP* haplotype defined by the alleles of four polymorphisms was found to associate with cBSE resistance. The haplotype includes insertion alleles from the two InDels with known associations with classical BSE, and the alleles of htSNP 1701 and 4136. However, the haplotype is uniquely tagged by the "T" allele of htSNP 4136 and efficiently captured the association of classical BSE resistance. A bioinformatic analysis of sequence containing the htSNP 4136 [C/T] was performed using ConSite (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite/), which identified potential changes in transcription factor binding sites associated with the polymorphism. The sequence with the more common C allele, which is associated with disease, has an apparent c-Fos binding site, whereas, the T allele has an apparent NF-kappaB binding site.

5.2 Future directions

The large number of SNP markers and the two sets of animals used in this study make it the most comprehensive study to date to test genetic loci for an association with classical BSE in European Holstein cattle. The genome-wide scan of half sib families identified an association between the genetic loci on 18 chromosomes with BSE incidence in European Holstein cattle. The identification of markers at or near statistical significance within the same chromosomal regions in both sets of samples provides independent evidence for the association of those regions and the presence of one or more genes within the regions influencing the incidence of BSE in cattle. Furthermore the high density scan provides confirmation of BSE association on many of the same chromosomes. While this study has identified some interesting findings it is important that these results are confirmed in additional populations of cattle and of other animal species susceptible to prion diseases.

PRNP htSNP 4136, a critical SNP in the haplotype that is associated with disease absence is, located in the promoter region exactly 14 bases upstream from exon 1. Expression of the bovine prion protein gene requires interaction between the promoter and the first intron. We are not able to ascertain from this study alone whether these variations are biologically relevant to the etiology of BSE infection. Therefore, more specific investigation of the effect of altered transcription factor binding activities in the promoter region of the bovine *PRNP* is a critical step that is necessary in an improved understanding of its biological significance to BSE disease.

In the pursuit of positional and functional candidate genes one of the most promising of the candidate genes is hypothetical gene LOC521010, similar to FK506 binding protein 2 located on chromosome 1 at 29,316,874bp. This gene encodes a protein that is a member of the immunophilin protein family and is involved in basic cellular processes involving protein folding, the relevance of BSE itself being a disease manifested through altered protein folding should sufficiently emphasize the relevance of this candidate for subsequent analyses. This encoded protein is a cis-trans prolyl isomerase that binds the immunosuppressants FK506 and rapamycin. The SNP located on BTA 14 associated with BSE disease case-control analysis, is in close proximity to the gene exostoses (multiple) 1, *EXT1*. McCormick et al., showed that *EXT1* is a

transmembrane glycoprotein whose expression in cells results in the alteration of the synthesis and display of cell surface heparin sulfate glycosaminoglycans (GAGs). The N terminus of PrP contains a GAG-binding motif and it is thought that PrP binding of GAG is important in prion disease.

Although substantial progress has been made over the last 20 years in prion research there is still much that remains unclear, unknown and even contradictory. This emphasizes that prion research remains an important field of study. For one thing, it is important to gain a better understanding of the normal physiological role(s) and function(s) of the prion protein itself. It is of further importance to understand how genetic variations i.e. nucleotide polymorphisms in the *PRNP* gene, as well as its interactive partners, can lead to altered gene function and expression. Again, this will provide a mechanism for integration and utilization of identified associations of various genes, including specific sequence information, with variation in susceptibility to various prion diseases. Therefore a greater understanding of the regulation of the prion gene itself, as well as how genetic variations within identified disease-associated regions are related, will facilitate the development of strategies to better control the incidence, prevalence and dissemination of this disease.

It is equally important to gain a better understanding of the nature of TSE transmissibility. The exact mechanism of how TSE infectivity is conferred is something that still eludes the research community. It is noted that the accumulation of TSE is inhibited by the absence of the spleen and complement components (Mabbott and MacPherson, 2006). This implies that oral transmission of TSE disease requires a biological pathway from GI to CNS and therefore disruption of this pathway affects the onset of disease. This coupled with the fact that there is a large variation in disease incubation times further implies that there may be variations in the process of disease onset. Some researchers have pointed out that one major difference between PrP^c and other amyloidogenic polypeptides is its GPI anchor and hypothesize that the GPI anchor is the key to the unique transmissibility to TSE diseases (Caughey and Baron, 2006). This area of research may provide insight into methods by which to reduce or interrupt routes of oral transmission of prion diseases that may include interventions as simple as modified handling of feedstuff to pharmaceutical development.

Overall a better understanding of the nature of all prion proteins with their assortment of genetic variations, how they function normally and in a disease state, would not only aid in the discovery of therapeutic treatments but also in the prevention and dissemination of the disease. It is essential to evaluate the implications that may be associated with selection of perceived resistance to prion disease. For instance, the selection of animals for perceived resistance may effectively lead to the delay in prion disease development. If selection for resistance simply results in an extension of the incubation time and onset of symptom then it is possible that this may actually impede the effectiveness of our identification of new or existing cases. As with all research, the testing of a hypothesis generally results in the development of several additional hypotheses to be tested, and while my research has yielded a foundation for future investigation this field of study is full of potential for discovery.

Appendix A

Chapter 2 Appendix

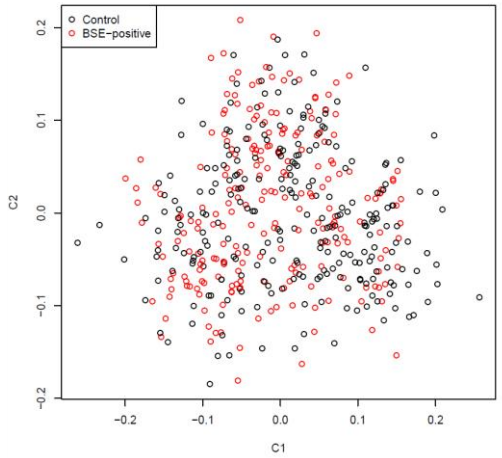
Additional Table 2-1 The best fit model analysis results (Genotypic, Recessive, and Dominant) of the case-control sample set. Locations were determined by blast to bovine sequence version 4.0. The location denoted by # was determined by blast to the Maryland sequence assembly. Threshold of suggestive significance $p \leq 10^{-4}$ is denoted as *.

CH	SNP ID	Location	TEST	AFF	UNAFF	CHIS	D	p-value
1	rs29012747	35,152,734	REC	16/125	40/137	6.85	1	8.90E-03
1	rs29026512	59,101,518	DOM	113/29	117/60	7.11	1	7.70E-03
3	rs29016848	90,945,586	DOM	80/63	71/106	7.96	1	4.80E-03
4	rs29010270	104,945,42	REC	14/129	5/172	6.87	1	8.80E-03
4	rs29009791	107,076,63	REC	6/137	26/151	9.68	1	1.90E-03
5	rs29002468	12,350,330	REC	20/122	9/167	7.63	1	5.70E-03
5	AF017143	71,118,012	DOM	111/31	114/62	6.82	1	9.10E-03
5	rs29024670	112,776,75	REC	17/125	5/172	10.27	1	1.49E-03
8	rs29012436	70,299,523	DOM	112/31	113/64	7.95	1	4.80E-03
8	rs29027876	70,303,317	DOM	111/32	113/64	7.15	1	7.50E-03
10	rs29018034	61,645,811	GEN	28/56/5	18/96/63	9.32	2	9.40E-03
11	rs29009663	16,684,267	DOM	84/59	131/45	8.84	1	2.90E-03
13	rs29019327	3,111,153	REC	18/125	45/132	8.24	1	4.10E-03
14	rs29012827	10,347,750	GEN	46/32/6	41/78/52	17.12	2	1.90E-
14	rs29021189	13,465,428	REC	42/100	29/148	7.93	1	4.90E-03
15	rs29012082	23,647,376	GEN	29/47/6	15/87/75	13.29	2	1.30E-03
15	rs29012083	23,647,575	GEN	35/51/5	23/92/62	10.96	2	4.20E-03
15	rs29012086	23,647,658	GEN	35/51/5	23/92/62	10.96	2	4.20E-03
17	rs29026692	66,954,281	REC	5/136	22/154	8.06	1	4.50E-03
18	rs29016029	41,705,003	GEN	17/74/5	32/62/83	9.26	2	9.70E-03
19	rs29015011	55,229,010	REC	31/112	18/156	7.72	1	5.50E-03
20	rs29021984	56,612,844	REC	14/129	5/172	6.87	1	8.80E-03
20	rs29009836	22,715,561	DOM	89/54	135/42	7.42	1	6.50E-03
21	rs29017681	14,702,969	REC	31/111	18/159	8.24	1	4.10E-03
22	rs29016333	40,957,965	GEN	35/58/5	36/104/3	11.54	2	3.10E-03
23	rs29013434	7,339,117	REC	23/120	51/126	7.21	1	7.30E-03
23	AAFC0207653	40,983,286	GEN	18/83/4	36/70/69	11.29	2	3.50E-03
26#	rs29017074	23,326,504	REC	42/94	31/143	7.24	1	7.10E-03

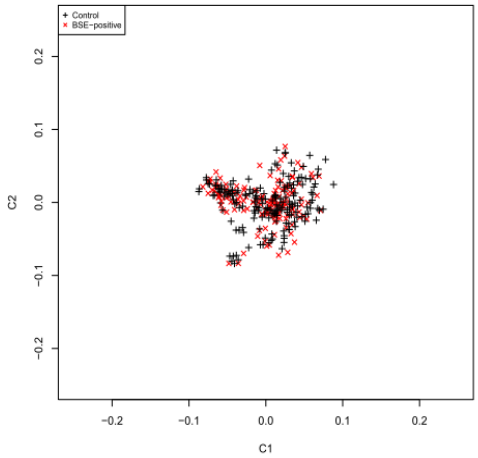
A version of this chapter has been accepted for publication. Murdoch BM *et al.* Prion haplotype associated with classical BSE incidence European Holstein Cattle. PLoS ONE 2010.

Additional Figure 2-1 Multi-dimensional Scaling (MDS) of the multiple family sample set and unrelated sample set. Multi-dimensional scaling (MDS) of the matrix of genome-wide identity-by-state (IBS) distances was used to provide a two-dimensional projection of the data onto axes representing components of genetic variation. Animals whose genetic ancestry differs significantly appear as outliers on the MDS plot. To avoid confounding the multi-dimensional scaling by extended linkage disequilibrium, the genotype data were reduced to a set in which no pair of SNPs was correlated with r^2 > 0.2. For this set of SNPs, the genome-wide IBS pair-wise identities between each pair of animals were calculated using PLINK (Purcell et al. 2007; Version 1.04). These IBS-relationships were converted into genetic distances by subtracting them from one, and the matrix of pair-wise IBS distances was used as input for multi-dimensional scaling.

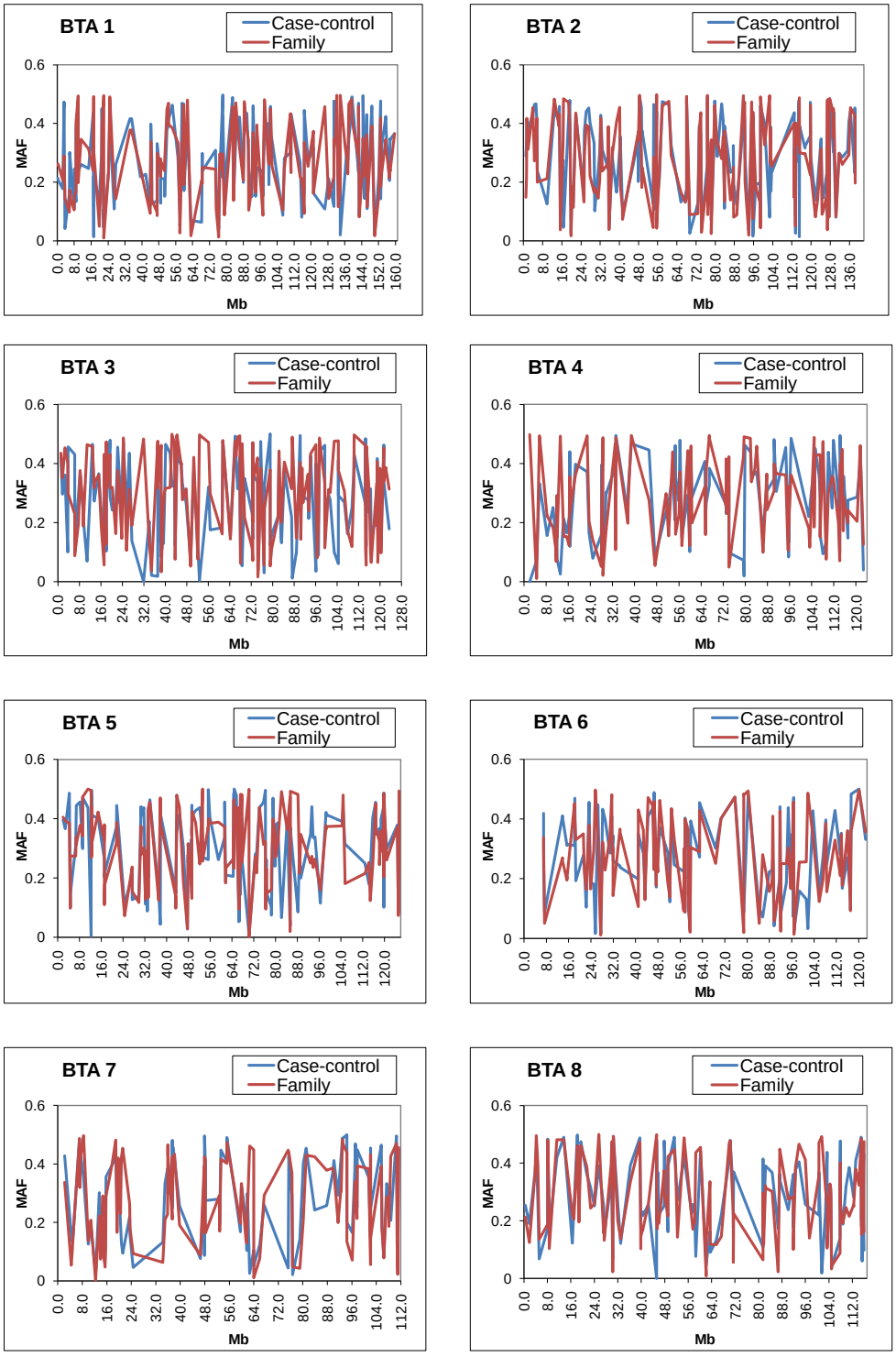
MDS plot of family sample set

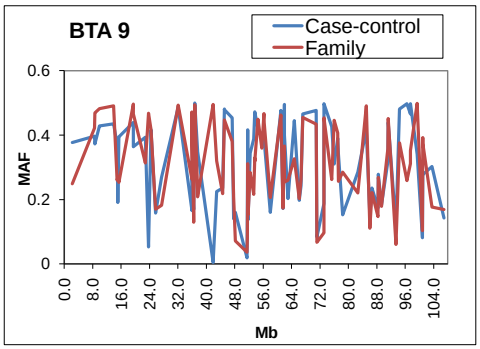


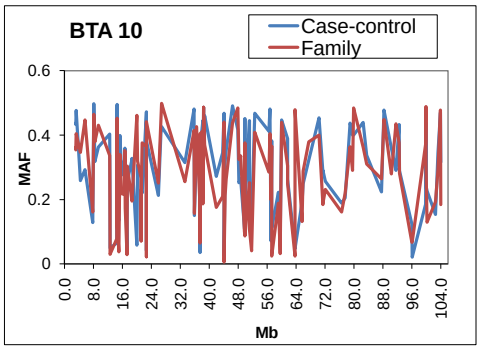
MSD plot of unrelated samples

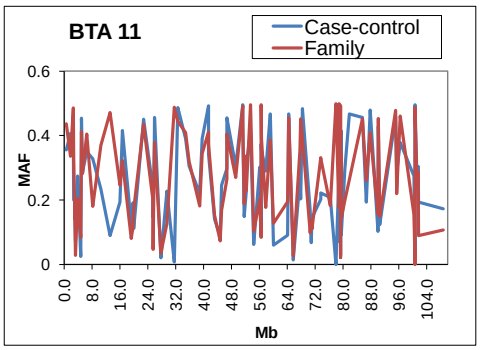


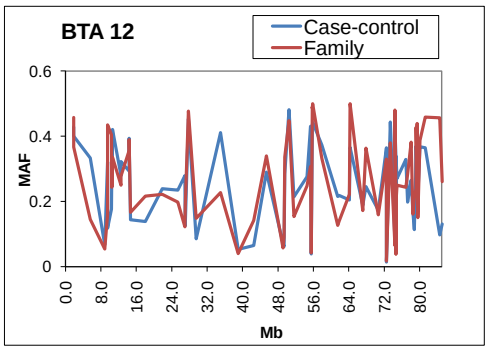
Additional Figure 2-2 -Graphical plots of the MAF versus position for each chromosome for the multiple family sample set and case-control sample set.

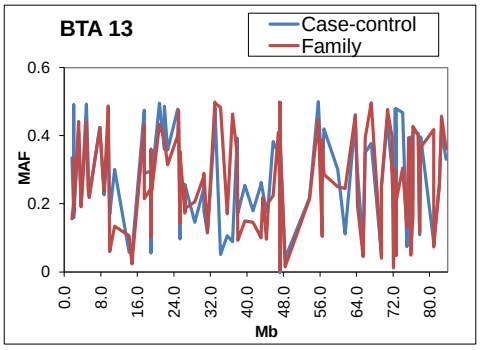


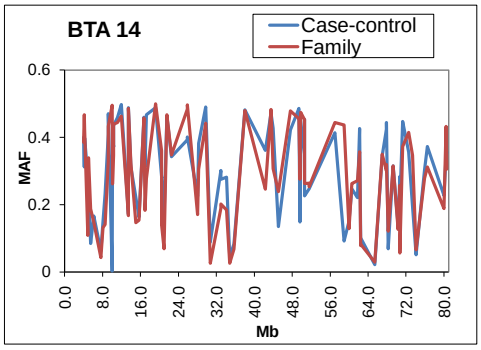

BTA 9
Case-control
Family
MAF
Mb

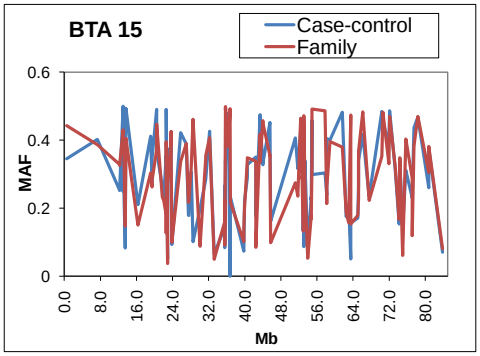

BTA 10
Case-control
Family
MAF
Mb

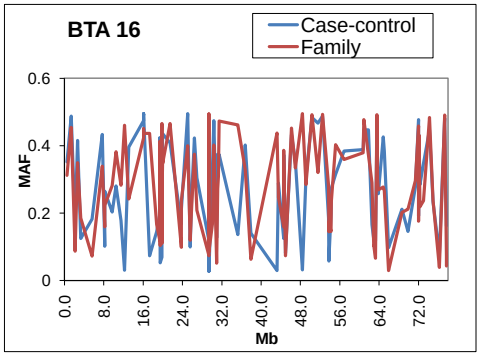

BTA 11
Case-control
Family
MAF
Mb

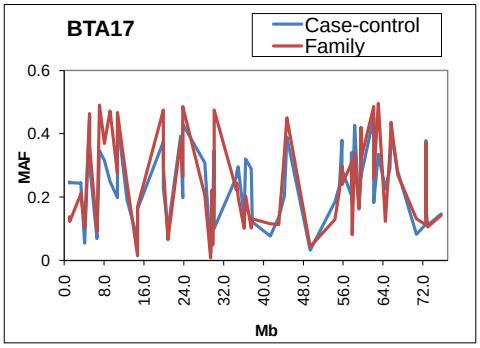

BTA 12
Case-control
Family
MAF
Mb

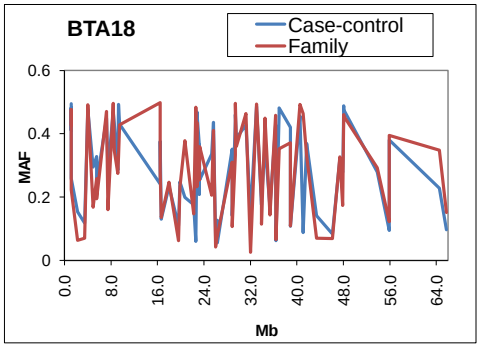

BTA 13
Case-control
Family
MAF
Mb

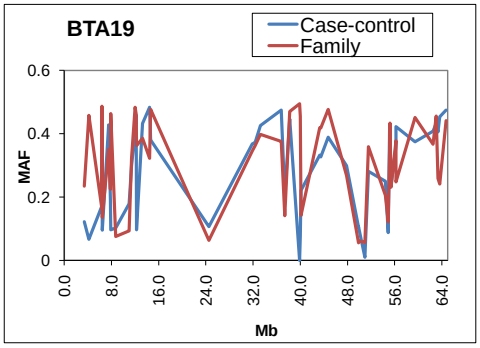

BTA 14
Case-control
Family
MAF
Mb

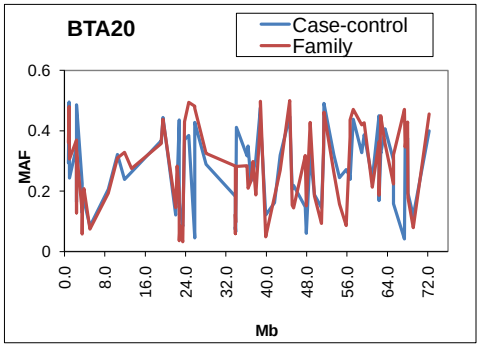

BTA 15
Case-control
Family
MAF
Mb

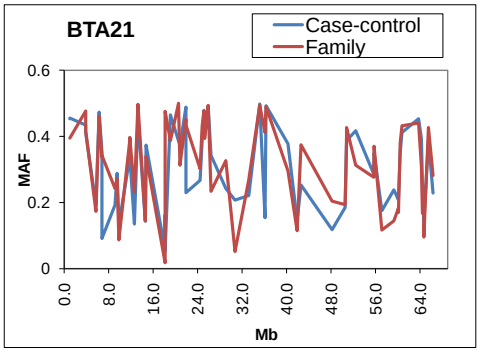

BTA 16
Case-control
Family
MAF
Mb

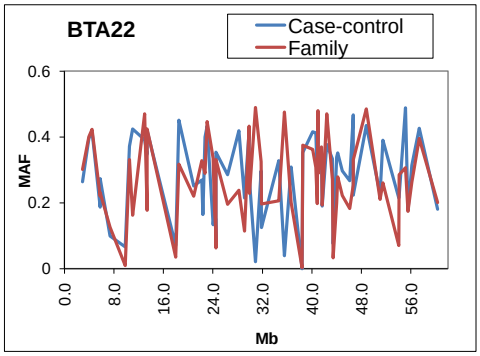

BTA17
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
72.0
Mb

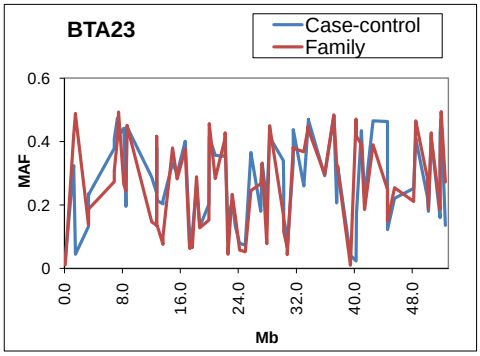

BTA18
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
Mb

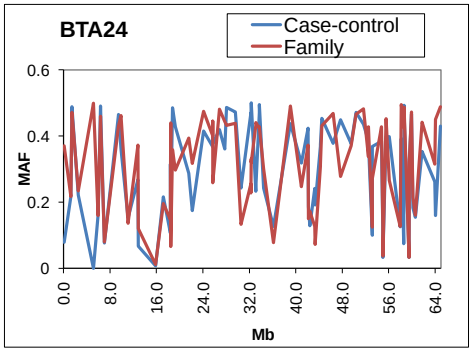

BTA19
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
Mb

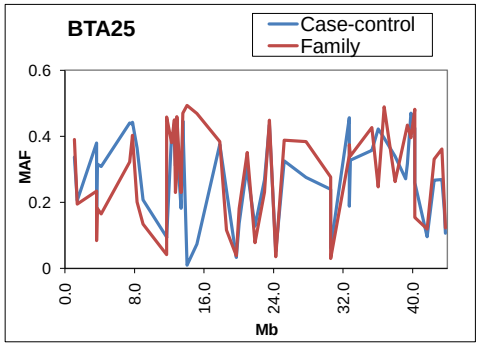

BTA20
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
72.0
Mb

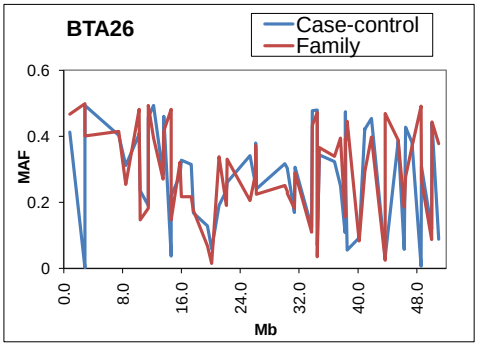

BTA21
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
Mb

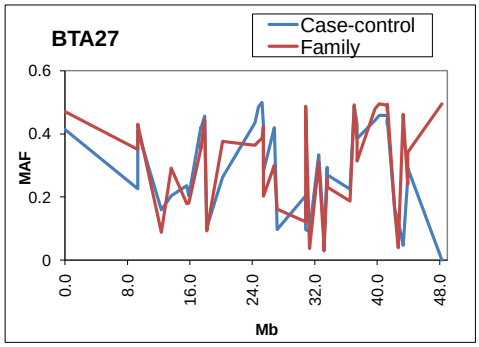

BTA22
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
Mb

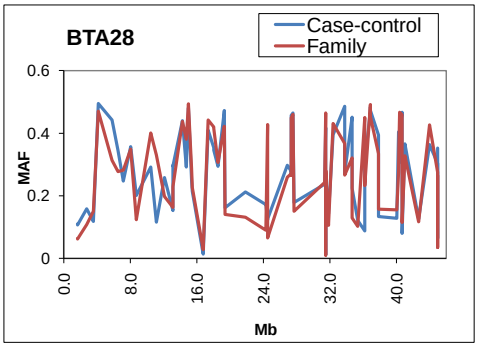

BTA23
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
Mb

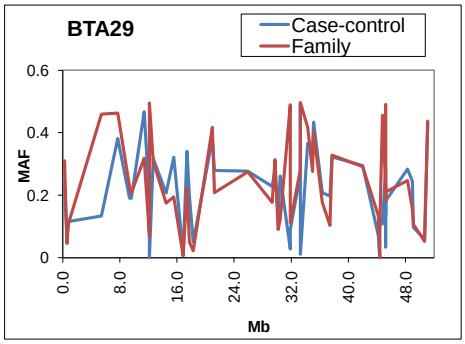

BTA24
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
Mb

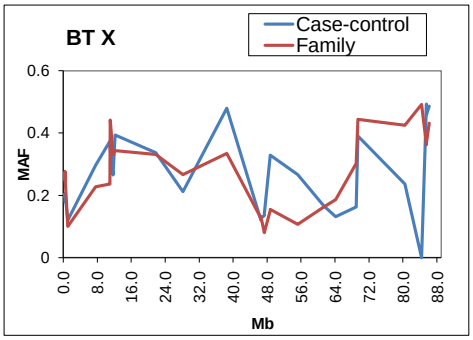

BTA25
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
Mb


BTA26
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
Mb


BTA27
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
Mb


BTA28
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
Mb


BTA29
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
Mb


BT X
Case-control
Family
MAF
0.6
0.4
0.2
0
0.0
8.0
16.0
24.0
32.0
40.0
48.0
56.0
64.0
72.0
80.0
88.0
Mb

Chapter 3 Appendix

Additional Table 3-1 *PRNP* htSNPs and indel frequencies within family BSE and case animals

SNP ID	Allele 1	Frequency in BSE affected	Frequency in unaffected	Allele 2	*p-value* (uncorrected) for BSE association	*p-value* (corrected) for BSE association	Odds Ratio
*snp 248	C	0	0	T	0.067	1	NA
*snp 449	G	0.193	0.266	T	0.010	1	0.660
*snp 1392	T	0	0	C	0.844	1	NA
*snp 1567	T	0	0	C	0.699	1	NA
*snp 1701	A	0.289	0.310	G	0.525	1	0.907
*snp 1783	A	0	0	G	0.065	1	NA
indel 23-bp	I	0.210	0.279	D	0.016	1	0.687
*snp 3641	C	0.219	0.223	T	0.891	1	0.978
*snp 4136	T	0.045	0.054	C	0.600	1	0.838
indel 12-bp	I	0.281	0.335	D	0.080	1	0.774
snp 4732	A	0.207	0.265	G	0.055	1	0.722
*snp 4776	T	0.117	0.197	C	0.002	0.039	0.543
*snp 6811	T	0	0	A	0.608	1	NA
*snp 8631	G	0.380	0.418	A	0.253	1	0.852
*snp 9162	C	0.	0	T	0.287	1	NA
*snp 9786	C	0.739	0.715	T	0.437	1	1.128
snp 13793	G	0.503	0.381	A	0.014	1	1.644
snp 13861	G	0.460	0.622	C	0.001	0.016	0.518
snp 13925	G	0.179	0.117	C	0.055	1	1.656
snp 17284	A	0.051	0.081	G	0.178	1	0.613
*snp 20720	T	0	0	C	0.065	1	NA
*snp 20957	T	0.188	0.171	C	0.521	1	1.120
*snp 21680	C	0.380	0.395	T	0.663	1	0.939

The SNP denoted by * was previously tested for an association with classical BSE and no significant associations were found (Murdoch *et al*. [25]). The presence of NA denotes that SNP was not analyzable due to the absence in the unaffected sample set.

Additional Table 3-2 Results of comparative haplotype analysis with different combinations. The first table includes both htSNP 449 and indel 23; the following tables include either htSNP 449 or indel 23.

Haplotype **snp449**|snp1701|**indel23**|snp4136|indel12|snp4732|snp4776

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
GAICIAT	0.1294	0.1284	0.000544	1	0.9814
TGDCDGC	0.5337	0.47	0.9362	1	0.3332
GAICIGT	0.01742	0.01415	0.03849	1	0.8445
TGDCDAC	0.05835	0.01609	2.396	1	0.1217
TGDCIGC	0.01128	0.02157	0.4278	1	0.5131
TADCDGC	0.06056	0.05959	0.000953	1	0.9754
TADCIGC	0.03735	0.04581	0.1074	1	0.7431
GAITIAC	0.02689	0.1306	11.1	1	0.000864
GAICIGC	0.05409	0.0921	1.353	1	0.2448
GAICIAC	0.05467	0.02144	1.509	1	0.2194
TADCDAC	0.01629	0.000276	1.33	1	0.2488

Haplotype **snp449**|snp1701|snp4136|indel12|snp4732|snp4776

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
GACIAT	0.1283	0.1253	0.004505	1	0.9465
GACIGT	0.01779	0.01387	0.05576	1	0.8133
GATIAC	0.02675	0.1251	10.4	1	0.001261
TATIAC	0.005409	0.0257	2.064	1	0.1508
GACIAC	0.05485	0.02168	1.527	1	0.2165
TACDAC	0.01621	0.000277	1.355	1	0.2445
TGCDAC	0.058	0.01556	2.491	1	0.1145
GACIGC	0.05334	0.09157	1.404	1	0.2361
TACIGC	0.03684	0.04243	0.04957	1	0.8238
TGCIGC	0.01144	0.02107	0.3818	1	0.5366
TACDGC	0.06055	0.05822	0.005648	1	0.9401
TGCDGC	0.5306	0.4592	1.197	1	0.2739

Haplotype **snp449**|snp1701|snp4136|indel12|snp4732

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
GATIA	0.02673	0.1199	9.591	1	0.001956
TATIA	0.005409	0.02579	2.077	1	0.1496
GACIA	0.1832	0.1508	0.4311	1	0.5114
TACDA	0.01627	0.000968	1.218	1	0.2697
TGCDA	0.05833	0.01714	2.306	1	0.1289
GACIG	0.07106	0.1071	1.009	1	0.3151
TACIG	0.03683	0.0409	0.02653	1	0.8706
TGCIG	0.01142	0.02107	0.3826	1	0.5362
TACDG	0.0605	0.05752	0.009328	1	0.9231
TGCDG	0.5303	0.4588	1.2	1	0.2734

Haplotype **snp449**|snp1701|snp4136|indel12

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
GATI	0.03397	0.1304	9.21	1	0.002406
TATI	0.003312	0.04053	5.538	1	0.01861
TACI	0.2493	0.2516	0.00166	1	0.9675
TACI	0.03839	0.03744	0.001484	1	0.9693
TGCI	0.01737	0.02355	0.1216	1	0.7273
TACD	0.07597	0.05688	0.3383	1	0.5608
TGCD	0.5817	0.4597	3.617	1	0.05719

Haplotype snp1701|**indel23**|snp4136|indel12|snp4732|snp4776

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
AICIAT	0.1283	0.1204	0.03455	1	0.8525
AICIGT	0.01788	0.01342	0.07414	1	0.7854
AITIAC	0.03211	0.1449	11.96	1	0.000545
AICIAC	0.04909	0.01919	1.429	1	0.2319
ADCDAC	0.01683	0.008413	0.308	1	0.5789
GDCDAC	0.05818	0.01816	2.245	1	0.134
AITIGC	0.005788	0.02915	2.489	1	0.1147
AICIGC	0.05376	0.08551	1.014	1	0.3139
ADCIGC	0.0358	0.0436	0.09951	1	0.7524
GDCIGC	0.01146	0.02028	0.3323	1	0.5643
ADCDGC	0.05994	0.05896	0.001042	1	0.9743
GDCDGC	0.5309	0.4381	2.079	1	0.1493

Haplotype snp1701|**indel23**|snp4136|indel12|snp4732

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
AITIA	0.02662	0.1384	12.76	1	0.000355
AICIA	0.183	0.1411	0.7519	1	0.3859
ADCDA	0.01693	0.01018	0.1894	1	0.6635
GDCDA	0.05843	0.02007	2.025	1	0.1548
AITIG	0.005823	0.02911	2.47	1	0.116
AICIG	0.07117	0.1036	0.8446	1	0.3581
ADCIG	0.03682	0.04137	0.03396	1	0.8538
GDCIG	0.01149	0.02027	0.3289	1	0.5663
ADCDG	0.05981	0.05847	0.001944	1	0.9648
GDCDG	0.5299	0.4374	2.065	1	0.1507

Haplotype snp1701|**indel23**|snp4136|indel12

HAPLOTYPE	F_A	F_U	CHISQ	DF	P
AITI	0.0373	0.1733	14.9	1	0.000114
AICI	0.246	0.2358	0.03403	1	0.8536
ADCI	0.04169	0.04479	0.01415	1	0.9053
GDCI	0.01729	0.02339	0.118	1	0.7312
ADCD	0.07598	0.06884	0.04463	1	0.8327
GDCD	0.5817	0.4539	3.938	1	0.0472