Characterisation of regulatory regions of the *PRNP* gene and their effect on susceptibility to sporadic and variant Creutzfeldt-Jakob Disease

James E. McCormack

A thesis submitted in partial fulfillment of the requirements of the University of Edinburgh for the degree of Doctor of Philosophy

The programme of research was carried out at the Institute for Animal Health , Edinburgh and the department of Biological sciences , University of Warwick.

February 2001

Declaration

I declare that this thesis has been composed entirely by myself and that the work presented herein is my own, except where otherwise stated. All experiments were designed by myself, in collaboration with my supervisors. No part of this thesis has been, or will be, submitted for any other degree, diploma or qualification.

James McCormack

February 2001

Acknowledgements

I would like to thank my supervisors, Dr. Jean Manson, NPU Edinburgh, Prof. James Ironside and Prof. Bob Will of the National CJD Surveillance Unit for their constant guidance and support during my PhD. I would also like to thank Herbert Baybutt, Nadia Tuzi, Rona Barron, Maurice Canham, Gerry O'Neill, Angie Chong and everyone else in Molecular Biology for assistance and advice throughout these three years particularly Nora Hunter, Wilfred Goldmann, Dory Cairns, Allister Smith and Arlene Ross for help with automated sequencing. A special note of thanks must go to Robert Proudfoot who is sadly missed by everyone here in NPU. I would like to thank Dr. Peter Estibeiro for the gift of a lambda library. In the CJD unit I would like to thank Dawn Everington, Kathy Estibeiro, Lisa Strain, Matthew Bishop and Marian Bruce. I would also like to thank Dave Stott, Sharon Cunningham, Dominic Murphy and everyone else in the University of Warwick for invaluable assistance when I was down south and Irene McConnell, Val Thompson, Joyce Grierson and all the rest of the barrier staff for showing me the ropes when I was upstairs. I would also like to thank Sara Knott for advice on linkage analysis. I would also like to thank the Wellcome trust for funding my work and everyone associated with the four year PhD programme particularly Dorothy Watson.

I really want to thank my family for encouraging me and putting up with my rare appearances at home, sorry Mam, Edel, Eimear and Hugh, I will be home more often in future. I would also like to hope that Dad would have approved of the course I've taken and of the final product. Finally I must thank Laura for keeping me sane these last three years and for her support, advice and love.

Abstract

The protease resistant protein (PrP) plays an essential role in transmissible spongiform encephalopathies, a group of fatal neurodegenerative diseases in mammals which include sporadic CJD and variant CJD in humans. While polymorphisms in the PrP open reading frame affect an individual's susceptibility to sCJD and vCJD, other factors must also act in determining susceptibility. One possible candidate is the level of PrP gene expression, as transgenic mice that over or under express PrP have shorter or longer TSE incubation periods. Therefore, this study aimed to characterise regulatory regions of the *PRNP* gene, which encodes the human PrP protein, to determine if polymorphisms in these regions affect an individual's susceptibility to sCJD or vCJD.

Based on analysis of PrP gene regulatory regions in other species 4.8 Kb of human genomic sequence containing 1.5Kb of upstream sequence the first exon and 3.1Kb of intron was cloned from a human genomic lambda library into a CAT reporter gene. Transient transfections in human neuroblastoma cells of this construct and of nested deletions of this construct identified two regulatory regions, one between –43 and –9 bases upstream of the transcription initiation site and one within the intron between +292 and +622 basepairs.

Having identified regulatory regions in neuroblastoma cells the effect of the two regulatory regions on expression *in vivo* was examined by making three constructs where expression of a LacZ reporter gene was driven by the 4.8Kb human fragment, the upstream and exon sequence or the exon and intron sequence. A construct which contained 3.4Kb of equivalent murine sequence was used as a control. Although all four constructs were successfully injected into fertilised eggs giving rise to transgenic embryos and adults, no LacZ expression was detected suggesting that additional sequences may be required for full expression of *PRNP*.

The upstream and intronic regions were sequenced in sCJD, vCJD and control individuals to determine if any polymorphisms or mutations exist in the regulatory

regions which cause or affect susceptibility to either form of CJD. Three polymorphisms were identified, a C to G transversion at position -101 relative to the start of exon one was found at an allele frequency of 13% in control individuals, a G to C transversion at +310 found at a frequency of 6% in controls and a T to C transition at position +385 found at a frequency of 5% in controls. Analysis of the frequency of these polymorphisms in CJD patients and controls showed that the -101G and +310C alleles each independently increase an individual's risk of developing sporadic CJD. The effect of the -101G allele was strongest in PrP ORF codon 129 methionine homozygotes even when the linkage between –101G and codon 129 methionine was accounted for. The +385C allele increased susceptibility to vCJD but this was not significant.

These results suggest that variation in the levels of expression of PrP in humans affect susceptibility to sCJD. Further work is needed to confirm these associations and to determine in which tissues alteration of PrP gene expression affects susceptibility to disease.

Abbreviations

AD-Alzheimer's Disease
alb-albumin
ANOVA-Analysis of variance
APOE-Apolipoprotein E
BSE-Bovine Spongiform Encephalopathy
CAT-Chloramphenicol Acetyl Transferase
Ci-Curie
CJD-Creutzfeldt-Jakob Disease
CNS-Central Nervous System
CWD-Chronic Wasting Disease
DMSO-Dimethyl Sulfoxide
DNA-Deoxyribonucleic acid
Dpl-Doppel protein
EDTA- Ethylenediaminetetraacetic acid.
EEG-Electroencephalogram
EGTA- Ethyleneglycoltetraacetic acid.
fCJD-familial CJD
FDC-Follicular Dendritic Cells
FFI-Fatal Familial Insomnia
GABA- γ–Amino Butyric Acid.
GH-Growth Hormone
GPI-Glycophosphoinositol
GSS-Gerstmann Straussler Scheinker disease
HSP-Heat shock protein
ic-Intra-cerebral
iCJD-iatrogenic CJD
IMS-Industrial methylated spirits
io-Intra-ocular
ip-Intra-peritoneal
IPTG- Isopropyl-β-D-thiogalactopyranoside
IRF-Interferon regulatory factor
J-Joule
Kb-Kilobase
LB-Luria Bertani Broth
LRS-Lymphoreticular system
LTP-Long term potentiation
M-Methionine
mA-miliAmp
MBM-Meat and Bone meal
mRNA-messenger RNA
OD-optical density

ORF-Open reading frame
PAGE-Poly-Acrylamide Gel Electrophoresis
PBS-Phosphate Buffered saline
PCR-Polymerase Chain Reaction
Pfu-plaque forming units
Prnd-Downstream prion like gene (Doppel protein gene)
PRNP-Human PrP gene
Prnp-Murine PrP gene
PrP-Proteinase resistant Protein
PrP^{C} –Cellular PrP
PrP^{Sc} –PrP scrapie
sc-Sub-cutaneous
SCID-Severe Combined Immunodeficiency Disease
sCJD-Sporadic CJD
SNP-Single Nucleotide Polymorphisms
SOD-Super-Oxide Dismutase
TAE- Tris Acetic acid EDTA buffer
TBE- Tris Boric acid EDTA buffer
TE-Tris EDTA buffer
TEMED-N,N,N1,N'-Tetramethylethylenediamine
TLC-Thin Layer Chromatography
TSE-Transmissible Spongiform Encephalopathy
UV-Ultra-Violet
V-Valine
vCJD-Variant Creutzfeldt-Jakob Disease
wt-Wild-type
X-gal- 5-Bromo-4 Chloro-3-indolyl-β–D-galactopyranoside

List of Figures

List of Tables

1 Chapter One: Introduction

1.1 Transmissible spongiform encephalopathies

1.1.1 Overview

Transmissible Spongiform Encephalopathies (TSEs) are a group of fatal neurodegenerative diseases of mammals (Table 1-1). Clinically TSEs are characterised by ataxia, myoclonus and dementia. Affected individuals typically show neuronal cell death, astrocyte gliosis and spongiform degeneration in the grey matter of the brain. The most characteristic feature of these diseases is the deposition of PrP (Proteinase resistant protein) in the brain. The deposition can occur either in diffuse deposits, or in large plaques of amyloid, depending on the particular TSE. The extent and anatomical location of each of the characteristic pathological features of TSEs can vary between different forms of TSE, and between individuals with the same form of TSE.

Analysing PrP protein from infected animals shows the presence of a form of PrP that is proteinase K resistant. This form of PrP is PrP^{Sc} (the "Sc" referring to scrapie, a sheep TSE where PrP^{Sc} was first identified) while normal PrP is referred to as PrP^{C} ("c" meaning cellular). TSEs also have very long incubation periods ranging from months in experimental rodent models to over 40 years in one form of human TSE. The aetiology of TSEs is varied, they can be acquired through medical procedures or eating infected animals, occur genetically through mutations in the PrP Open Reading Frame (ORF) or they can develop spontaneously with no identifiable cause. The most remarkable feature of TSEs is their transmissibility, for regardless of their aetiology, TSEs can be experimentally transmitted to other animals (Brown et al., 1994) by injecting brain homogenate of an infected animal into the brain of the experimental animal. Such transmission studies have also shown that TSEs exist as strains (Bruce, 1993) which upon passage (infection) in mice have a distinct phenotype, consisting of

a characteristic incubation period and patterns of brain lesions that are retained on subsequent passages.

PrP^{Sc} has a different secondary structure from PrP^{C} with an increase in the amount of β-sheet structure (Caughey et al., 1991) (Pan et al., 1993). The N-terminal region remains sensitive to proteinase K resulting in the three bands running at 27-30kDa on a PAGE gel after limited proteinase K digestion. The amount of glycosylation on PrP^{Sc} is used to discriminate between strains of TSE (Collinge et al., 1996) (Hill et al., 1997). One study has found four different patterns of glycosylation in different types of CJD, but this method of strain identification has been questioned as other groups have only found two patterns (Parchi et al., 1997) or found that the pattern varies in different regions of the brain (Somerville, 1999) and changes on transmission between species (Somerville et al., 1997).

Although TSEs are infectious diseases and exist as strains thus resembling a virus (TSEs were originally referred to as slow viruses due to their long incubation periods) there is a large body of evidence which suggests that a novel infectious agent is involved. This is the prion theory, which argues that the infectious agent is solely comprised of PrP^{Sc} and that the conversion of PrP^{C} to PrP^{Sc} is catalysed by PrP^{Sc} itself. Two lines of evidence supporting this theory are that firstly the mutated PrP proteins produced in genetic forms of disease can give rise to an infectious disease and secondly infectivity co-purifies with PrP^{Sc}. Treating infectious material with UV-light, DNase and RNase has no effect on infectivity implying no involvement of a nucleic acid, yet proteinase or urea treatment destroys infectivity implying that protein plays an important role in disease (Prusiner et al., 1981) (Millson et al., 1976).

Transmission studies of many different TSEs have shown it is more difficult to transmit disease between species (e.g. TSE infected mouse to hamster) than within species (e.g. mouse to mouse), with a lower proportion of animals developing disease and with longer incubation periods. However once the second species has been infected subsequent passages (infection) within that species (e.g. hamster to hamster) results in a higher proportion of animals developing disease with shorter incubation periods than the cross-

species transmission. This reduced efficiency of infection between species is termed the species barrier.

Table 1-1 Distribution and aetiology of Transmissible Spongiform Encephalopathies.

Disease	Species	Aetiology	Reference
CJD	Human	Sporadic, Acquired or Familial	(Sternbach et al., 1997)
GSS	Human	Familial	(Ghetti et al., 1994)
FFI	Human	Familial	(Rossi et al., 1998)
Kuru	Human	Acquired	(Gajdusek and Zigas, 1959)
vCJD	Human	Acquired	(Will et al., 1996)
Scrapie	Sheep	Unknown, possibly maternal	(Brown and Bradley, 1998)
BSE	Cattle	Acquired	(Wells et al., 1987)
CWD	Deer	Unknown	(Sigurdson et al., 1999)
TME	Mink	Unknown/Oral?	(McKenzie et al., 1996)
FSE	Cats	Oral	(Cadore, 1996)
Other SE	Zoo mammals	Oral	(Kirkwood and Cunningham, 1994)

1.1.2 Human TSEs

Four TSEs have been identified in humans, Creutzfeldt-Jakob Disease (CJD) Gerstmann-Straussler-Scheinker syndrome (GSS), Fatal Familial Insomnia (FFI) and kuru. All GSS and FFI cases and about 10% of CJD cases are familial (fCJD), linked to dominant acting mutations in the PrP ORF. A small proportion of CJD cases are iatrogenic (iCJD) i.e. caused by medical procedures which used equipment or products inadvertently contaminated by, or derived from CJD patients. The majority (90%) of CJD cases have no known cause, with no mutation in the PrP ORF and no association with occupation, medical procedures or area of residence (vanDuijn et al., 1998), these

cases are therefore referred to as sporadic CJD (sCJD). Kuru is a TSE spread by cannibalism in one tribe in New Guinea. A new form of CJD was identified in 1996 that has different clinical and pathological signs to other forms of CJD. Variant CJD (vCJD) as it has been termed is believed to be causally related to bovine spongiform encephalopathy (BSE) a TSE of cattle.

1.1.3 Human Familial TSEs

Familial TSEs are linked to mutations in the PrP ORF that act in an apparently autosomal dominant manner (see Figure 1-1). Each mutation causes one of three diseases, fCJD, GSS or Fatal Familial Insomnia (FFI). These mutations can be point mutations that change the identity of an amino acid or insertions of two to nine copies of the 24bp region which encodes the octapeptide repeat found in the N-terminal of the PrP protein. Although diseases caused by mutations in the PrP gene are classed as GSS, CJD or FFI there can be some overlap in clinical signs depending on the particular mutation involved (Parchi et al., 1996).

Familial CJD accounts for about 10% of total CJD cases. The average age of onset of clinical signs is 60 years of age with a median duration of disease of 4 months (Belay, 1999). The most common clinical sign is a rapidly progressing dementia accompanied by pyramidal and extra-pyramidal signs leading to spasticity, rigidity, tremors and myoclonus sometimes leading to akinetic mutism and cortical blindness. (Parchi et al., 1996). Most fCJD patients exhibit periodic tri-phase sharp waves electroencephalograph (EEG) patterns (Parchi et al., 1999). About 10% of cases have amyloid plaques in the brain containing PrP^{Sc} although the rest show diffuse PrP^{Sc} deposits. Mutations which have been associated with CJD are point mutations in codons 178, 180, 188, 200, 202, 203, 208, 210, 211, 212 or 232 of the PrP ORF or insertions of two or six copies of the octapeptide repeats at the N-terminal end of the protein (Figure 1-1) (Belay, 1999) (Salvatore and et al., 1996).

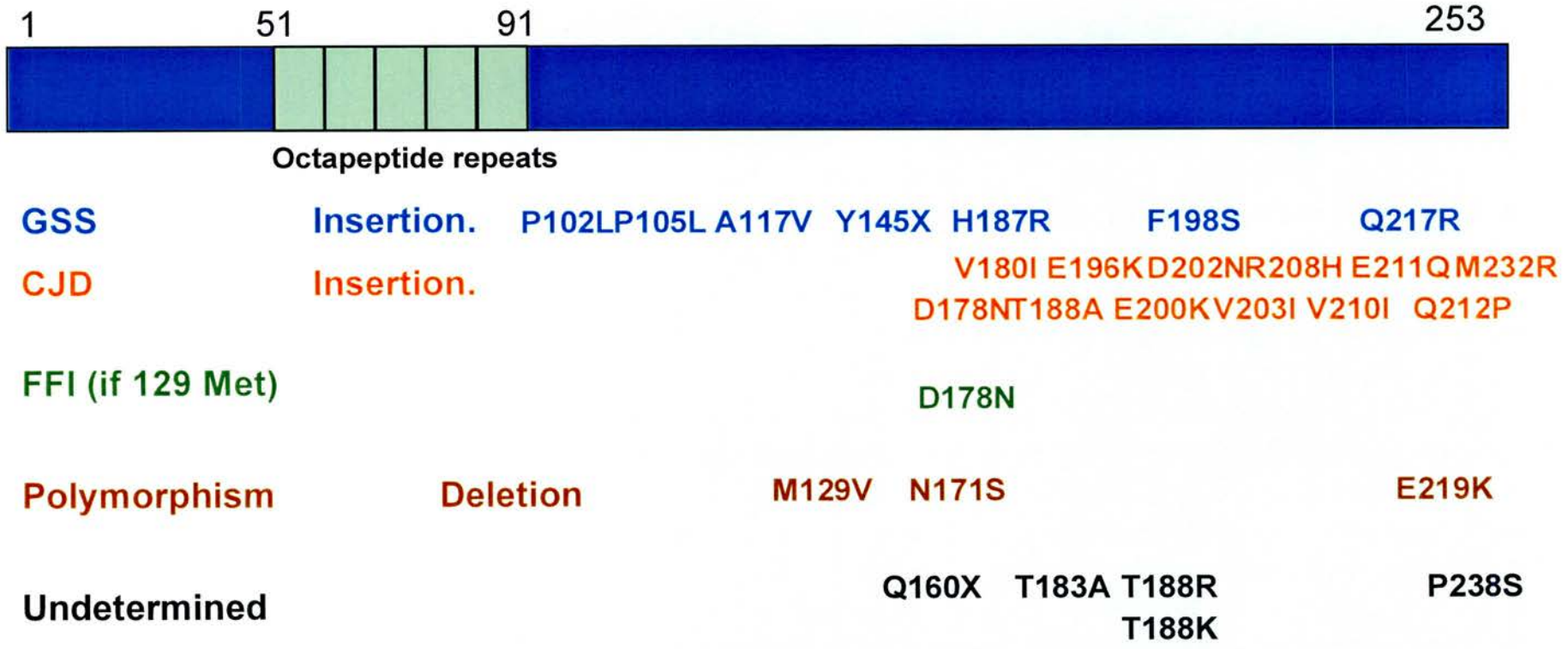


Fig.1-1 Mutations and polymorphisms of the human PrP protein .All familial TSE mutations appear to act in an autosomal manner. GSS associated mutations in Blue, CJD in red and fatal familial insomnia in Green. Polymorphisms in maroon with the most common allele on the left and mutations with undetermined phenotypes due to indeterminate clinical signs or pathology in black although some of the mutations listed here are classed as familial forms of TSE solely on the basis of their occurrence in individuals with a TSE and their absence in control population. Q160X and Y145X are nonsense (stop codons) mutations. Green rectangles between amino acids 51 and 91 are octa-peptide repeats.

GSS appears to be exclusively genetic, always being associated with dominant acting mutations at codons 102, 105, 117, 187, 198 or 287 or a stop codon mutation at codon 145 or insertions of 5,6,7,8 or 9 copies of the octapeptide repeats in the PrP open reading frame more (Belay, 1999) (Owen et al., 1990). Clinical features include dementia and extrapyramidal signs. The most consistent pathological finding is PrP amyloid plaques deposited in the brain (Bugiani et al., 2000) and the average course of GSS is five years. This contrasts with familial CJD where PrP plaques are found in only 10% of cases and the duration of disease is generally takes less than a year. There can be an overlap in pathology between GSS and CJD especially in GSS kindreds with insertional mutations (Parchi et al., 1996) there can also be variation in pathology between individuals with different GSS mutations (Piccardo et al., 1998).

Fatal Familial Insomnia (FFI) is the rarest familial TSE and only 24 affected kindreds have been found (Parchi et al., 1996) (Belay, 1999). Symptoms include reduction in sleep time, dementia and ataxia while the pathological effect is atrophy of the thalamus and olivary regions with marked neuronal loss. There is mild gliosis while spongiform degeneration or amyloid plaques are rare. However, as PrP^{Sc} can be detected in FFI it is classed as a TSE. FFI is caused by a point mutation in the PrP gene at codon 178, which changes an aspartic acid to asparagine. However, the disease is influenced by the common polymorphism at codon 129 that encodes valine or methionine. FFI only occurs if the codon 178 aspartic acid to asparagine mutation occurs on an allele which is methionine at codon 129, if codon 129 is valine familial CJD occurs. As the symptoms and pathology of FFI and CJD are quite distinct, the interaction between the asparagine at 178 and either valine or methionine at 129 must affect the structure of the protein sufficiently to alter the type of disease. However an FFI kindred from Australia (McLean et al., 1997) where the mutated allele was 178N 129M showed heterogeneity in clinical and pathological symptoms between FFI and CJD. The allele/phenotype relationship may therefore be more complicated and other factors may be involved. In both 178CJD and 178FFI kindreds the duration of disease is shorter and age of onset lower in homozygotes (MM in FFI and VV in CJD) than heterozygotes (Goldfarb and al., 1992).

Six cases have been reported of patients with a neurological disease similar to fatal familial insomnia but where there is no mutation of the PrP ORF. These may be cases of fatal sporadic insomnia (Nixon et al., 1996) (Parchi et al., 1999) analogous to sporadic CJD.

The clinical and pathological features of familial TSEs varies from mutation to mutation, from kindred to kindred and there can be considerable variation even within kindreds. For example, the average age-of-onset in kindreds with the F198S GSS varies from 34 to 71 years. While the penetrance (proportion of individuals carrying a mutation who develop disease) of D178N CJD is 100% while that of E200K CJD is estimated at 56% averaging over all affected kindreds. (Goldfarb et al., 1991) However, there is considerable variation between E200K kindreds as the disease is fully penetrant by the age of 80 in Libyan Jewish kindreds (Chapman et al., 1994) but not in a Chilean kindred. These two kindreds not only have the same mutation, but are thought to be related to each other as haplotype analysis of highly polymorphic markers around the *PRNP* gene show that both kindreds carry the same E200K mutation (Lee et al., 1999). The course of familial disease can be affected by factors other than the PrP coding sequence. A GSS kindred has been reported with the P102L mutation where affected individuals could be divided into two groups, one of whom suffered from dementia and a relatively short illness lasting just a year which is not "typical " of GSS course while the other group had ataxia over a longer period (five years), but had no cognitive impairment which is typical of GSS. There was no correlation between the two clinical courses and the codon 129 polymorphism (Barbanti et al., 1996), implying that there may be other genetic or environmental factors affecting the course of the disease.

A number of individuals with neurological disease have mutations in the ORF but have not been classified as GSS or CJD. This is due either to lack of any family history ((Finckh et al., 2000) mutations Q160X and T188K) or a lack of material to carry out pathological analysis in a suspected CJD patient ((Windl et al., 1999) T188R). The T183A mutation (Nitrini et al., 1997) has distinct clinical and pathological features which differ from most CJD cases, it rarely has the abnormal EEG pattern which is found in most sporadic and familial CJD patients and has a long course of disease of 50

months compared with the average of 8 months for sCJD. It cannot therefore be classified as GSS or CJD with any confidence. It has been argued that the CJD/GSS division is somewhat arbitrary and that each mutation should be classified as a sub-type of a single disease (Collinge et al., 1992).

1.1.4 Sporadic CJD

The commonest TSE in humans is sCJD. This neurodegenerative occurs at a rate of approximately one case per million people per year in every country and ethnic group examined. The average age of onset of clinical signs is 60 years of age with a median duration of disease of 6 months (Belay, 1999). The most common clinical sign is a rapidly progressive dementia accompanied by pyramidal and extra-pyramidal signs leading to spasticity, rigidity, tremors and myoclonus sometimes leading to akinetic mutism and cortical blindness. Most sCJD patients exhibit periodic tri-phase sharp waves electroencephalograph (EEG) patterns. About 10% of cases have amyloid plaques in the brain containing PrP^{Sc} although the rest show a variety of PrP^{Sc} deposition patterns.

Although there are no mutations of the PrP ORF in sCJD patients, two polymorphisms in the PrP ORF affect susceptibility to the disease. The most important polymorphism clinically is at codon 129 which can encode Methionine or Valine (129MV) (Owen et al., 1990). Although the 129MV polymorphism does not cause disease it exerts an influence on the susceptibility of Caucasians to sCJD (Collinge et al., 1991; Palmer et al., 1991) (Deslys et al., 1998; Zimmermann et al., 1999). In Caucasians 39% of the population are methionine homozygotes, 50% heterozygotes and 11% Valine homozygotes. Whereas analysis of over 1500 European sCJD cases shows that 71% of cases are methionine homozygotes, 13% of cases heterozygous and 16% valine homozygotes (Alperovitch et al., 1999). The reason for the increased susceptibility of codon 129 homozygotes but reduced susceptibility of heterozygotes is proposed to be that interactions between similar PrP proteins enable the development of disease; this will be discussed further in section 1.3.6

A second polymorphism at codon 219 where glutamine is replaced by lysine (E219K) (Kitamoto and Tateishi, 1994) also affects susceptibility to sCJD. This polymorphism is found at a frequency of 6% in the Japanese population but has not been found in Caucasians (Petraroli and Pocchiari, 1996). A study of 85 Japanese sCJD patients has found that all were glutamine homozygotes; this lack of lysine alleles in sCJD patients is statistically significant (Shibuya et al., 1998). Therefore, codon 219 heterozygotes (12% of the Japanese population) have a reduced susceptibility to sCJD. The codon 129 methionine/valine polymorphism is also present in Japanese populations but the frequency of valine is much lower at 4% (34% in Caucasians) analysis of Japanese sCJD patients shows no difference in susceptibility between codon 129 methionine homozygotes and heterozygotes (Doh-ura et al., 1991).

There are several other polymorphisms of the PrP gene which have no apparent effect on any form of CJD. Two are coding region polymorphisms which affect the coding region A polymorphism (N171S) has been found in a Brazilian family and been tentatively linked to atypical schizophrenia, although none of the affected individuals suffered from any symptoms of TSEs (Samaia et al., 1997) and not all carriers suffered from the disease. This mutation has also been reported in healthy individuals (Fink et al., 1994). Another polymorphism occurs where one of the five imperfect octapeptide repeats in the N-terminal region of the protein has been deleted although this mutation appears to have occurred several times independently as different repeats have been deleted (Salvatore et al., 1994) (Palmer et al., 1993). This rare polymorphism has not been associated with disease. There are also several synonymous substitutions of the open reading frame and a single nucleotide polymorphism at the 3' end of the human intron.

Analysis of sCJD patients in Europe and Japan has shown that different PrP ORF polymorphisms have different affects on sCJD susceptibility. The incidence of sCJD is the same in Japan as in Europe so perhaps different genetic factors affect susceptibility in different populations. This suggests that sCJD is a complex disease affected by several genetic loci. The examination of genetic factors to date has focused on the PrP ORF but it is possible that other genetic factors outside of this region influence

susceptibility to sCJD.

1.1.5 Iatrogenic CJD

CJD has been transmitted by medical procedures where hormones such as growth hormone and gonadotropin derived from the pituitary glands of CJD cadavers were injected into individuals with hormonal deficiencies. Tissues such as dura mater and corneas taken from CJD cadavers for grafting have also transmitted CJD to the recipients. Iatrogenic CJD (iCJD) has also been transmitted by medical equipment such as intra-cerebral electrodes, which cannot be sterilised adequately after use on the brains of CJD patients (Bernoulli et al., 1977) (Will and Matthews, 1982).

As the date of infection in iCJD cases can be determined, the incubation periods of iCJD can be calculated, this cannot be done in sporadic CJD as the cause of disease is unknown. The largest and best-characterised group of iatrogenic CJD patients are French individuals who developed CJD due to Growth hormone treatment in the 1980s (dAignaux et al., 1998). Growth hormone collected from cadavers was pooled into batches and presumably at least one cadaver used in the preparation of some batches was a pre-clinical CJD case. Analysis of 1361 individuals treated between 1982 and 1985 with growth hormone batches now known to be contaminated has shown that 55 individuals (4%) developed CJD. The mean incubation period has been estimated at 9 to 10 years. The 129MV polymorphism had an effect on the incubation period and susceptibility to iCJD that was different to that seen in sCJD. Heterozygotes were less susceptible (comparing the codon 129 genotype frequency of iCJD patients with that of the normal French population) and had longer incubation periods than homozygotes. There was no difference in incubation period between methionine and valine homozygotes. Valine homozygotes were actually slightly more susceptible to iatrogenic CJD than methionine homozygotes, which contrasts with sCJD where methionine homozygotes are more likely to develop sCJD. Analysis of iatrogenic CJD in Britain showed that 35 of 1880 (1.9%) growth hormone treated individuals developed CJD. Twenty of these patients have been sequenced at codon 129 of whom 11 (55%) were

valine homozygotes and only one was a methionine homozygote (Brown et al., 2000). The excess of valine homozyogtes among growth hormone iCJD patients may mean that valine homozygotes are more susceptible to either the route of infection or the inoculum. Perhaps a valine homozygote CJD patient was one of the infected donors and therefore valine homozygotes are more susceptible, although as several batches of growth hormone were contaminated it is unlikely that every batch just contained valine homozygote CJD patients. The small proportion of individuals treated with contaminated GH who developed CJD (4% in France 1.9% in Britain) must mean that factors other than codon 129MV status determines an individual's susceptibility to iCJD, these factors could include variation in PrP expression levels or patterns which could affect the rate of replication of the agent after infection.

1.1.6 Variant CJD

A new variant of CJD (vCJD) was identified in 1996 (Will et al., 1996). Up to the start of January 2001 there have been 88 confirmed or suspected cases in the UK, two in France and one in the Republic of Ireland (Table 1-2). The clinical and pathological symptoms are slightly different from classical CJD (the collective term used to describe sCJD, fCJD and iCJD). Most strikingly the age of onset is much lower (median age 28) with some patients being in their teens whereas very few (Belay, 1999) classical CJD cases under the age of 42 have been reported. The duration of the disease is also longer (median 13 months) than classical CJD (median 6 months) while the initial clinical symptoms are behavioural changes and ataxia compared with the dementia and myoclonus seen in classic CJD. Periodic sharp waves in electroencephalographs (EEG) found in most classical CJD cases are absent in vCJD. Pathologically vCJD is characterised by distinctive "florid" plaques in the brain containing PrP only rarely seen in classical CJD. Much evidence has accumulated that vCJD is caused by exposure to BSE (Bovine Spongiform Encephalopathy) infected material. The incubation periods and lesion profiles (the amount of vacuolar pathology in particular regions of the brain) of a panel of inbred mouse lines infected with BSE cattle brains or vCJD brains are indistinguishable from each other but different from those of scrapie or sCJD (Bruce et

al., 1997)(Figure 1-2). Epidemiological evidence also points to a link as almost all vCJD cases have occurred in Britain after the peak of the BSE epidemic.

All vCJD patients to date have been homozygous for the methionine polymorphism at codon 129 (Ironside, 1998) (Zeidler et al., 1997) a genotype shared with 40% of the Caucasian population. This genotype also increases the risk of developing sCJD. The absence of any heterozygous or valine homozygote vCJD patients may indicate that human methionine homozygotes are much more susceptible to vCJD. This may be due to the fact that, along with almost all species examined so far (Schatzl et al., 1995) (Schatzl et al., 1997), cattle are also methionine homozygotes at codon 129, which could increase the efficiency of interactions between bovine PrP^{Sc} and human PrP. Alternatively, all genotypes might be equally susceptible to BSE, but methionine homozygotes may have the shortest incubation period and valine homozygous and heterozygote individuals are still incubating the disease. Although the study of French Growth Hormone iCJD patients suggests that codon 129 genotype can affect both susceptibility and incubation period which could mean that VV and MV individuals will be less susceptible and have increased incubation periods. At present there is no way of determining which of these scenarios will occur.

Table 1-2 Number of vCJD cases in the UK 1995-1st January 2001.

Source CJD surveillance unit. There have also been two cases in France and one case in the Republic of Ireland.

Year	Number
1995	3
1996	10
1997	10
1998	18
1999	15
2000	25+7 probable cases

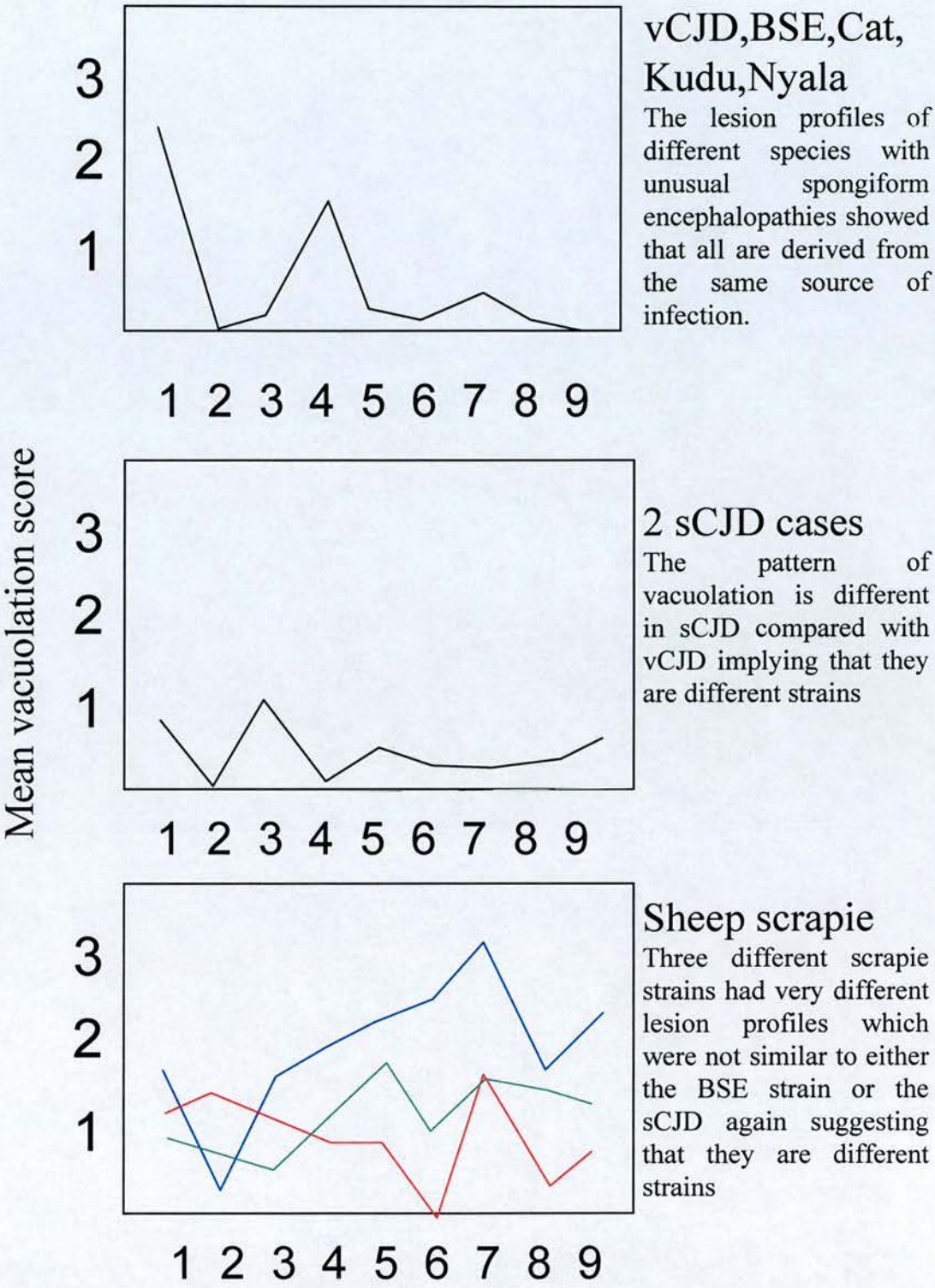


Figure 1-2 Lesion profiles of BSE,vCJD scrapie and sCJD. Brain homogenate from infected individuals were injected into RIII mice and and upon development of disease the vacuolation in various brain regions was examined and the amount of vacuolation scored from 1 to 5. Brain regions examined were 1=dorsal medulla, 2=cerebellar cortex, 3=superior colliculus, 4=hypothalmus, 5=thalmus, 6=hippocampus, 7=septum, 8=retrosplenial and adjacent motor cortex, 9=cingulate and adjacent motor cortex.Adapted from Bruce et al. 1997.

It is difficult to predict the number of individuals who will eventually develop vCJD from the numbers who have so far developed the disease to date (Table 1-2), as many of the parameters which will determine the size of the epidemic (e.g. amount of infectivity that entered the food chain, average incubation period, genetic variations in susceptibility and/or incubation period) are unknown. The individuals who have developed vCJD may represent the people with the shortest incubation periods who were infected at the beginning of the BSE epidemic, (in which case the epidemic will be very large) or they may be uniquely susceptible individuals who were infected at the height of the BSE epidemic, in which case the epidemic will be small. Computer modelling of the eventual scale of the epidemic has focused on two parameters, the average incubation period and the number of people who could be infected by one infected animal. Using data on vCJD deaths to the end of 1999 (Ghani et al., 2000) one model predicts that the total number of cases will be less than 6000, assuming that the mean incubation period is less than 60 years and that each infected bovine carcass infects two people on average. However, this model assumed that only methionine homozygotes would develop disease therefore assumed that 60% of the population is resistant to vCJD. If susceptibility or incubation period is affected by additional loci outside of the PrP open reading frame then the eventual scale of the epidemic could be larger than these predictions.

1.1.7 Kuru

Kuru is a TSE which has been described in the Fore tribe of New Guinea. It is thought to have first occurred at the beginning of the 20th, century but only came to the attention of Western doctors when it reached epidemic proportions in the 1950s (Gajdusek, 1996). Clinically it is characterised by extra-pyramidal signs, strabismus, dysphagia, mutism and inappropriate laughter. The pathology is characterised by many amyloid plaques composed of PrP in addition to diffuse deposits of PrP gliosis and vacuolation (Lantos et al., 1997) (Hainfellner et al., 1997). It is thought to have been transmitted by the

practice of cannibalistic funeral feasts of members of the community who had died of Kuru. The origin of the initial case which started the epidemic is unknown but may have been a fCJD or sCJD case. The disease was the largest cause of death of Fore women and was common in prepubescent children of both sexes. The route of infection is uncertain, the oral route is a distinct possibility with the distorted sex ratio due to the practice of women eating the brains (presumably with a high titre of infectivity) while men ate the flesh of the deceased. However kuru transmits poorly to primates via the oral route, this fact and the unusual sex and age distribution has led to the proposition of an alternative route of infectivity through cuts in the hands or mucous membranes. This may have occurred more often in women and children, as women prepared the bodies for funeral feasts and children were cared for exclusively by women possibly becoming infected at the same time. The ending of cannibalism from the 1960s onward led to a reduction in the number of cases of Kuru, but the disease has an exceptionally long incubation period as a few individuals have developed Kuru 40 years after the cessation of cannibalism.

Analysis of the M129V genotype of 91 kuru victims (Cervenakova et al., 1998) shows a strong correlation between genotype and age of onset. Younger patients (<15 years) tend to be pre-dominantly homozygous with more Methionine homozygotes then Valine homozygotes yet in older patients (>30 years) the situation is reversed with more heterozygotes than homozygotes. This difference may reflect longer incubation periods in heterozyogtes resulting in heterozygotes infected in childhood not developing disease until adulthood. This elongation of heterozygote incubation periods is similar to that seen in iatrogenic CJD. The frequency of methionine and valine alleles at codon 129 in the pre-Kuru Fore population is unknown so the effect of the codon 129 polymorphism on susceptibility to Kuru cannot be determined.

1.1.8 Scrapie

The most common TSE is scrapie, a fatal disease of sheep and goats. Scrapie was first mentioned in 1755 (Brown and Bradley, 1998) and the disease is endemic in all

countries with sheep with the exception of New Zealand and Australia. The pathology is characterised by vacuolation of the neuronal perikarya and grey matter neuropil, neuronal loss and astrocytosis (Wood et al., 1997). The clinical signs include pruritis, leading to continuous rubbing against posts resulting in wool loss and laceration of the skin, abnormal gait and posture. Ataxia occurs later and results in the animals being unable to walk (Brugere-Picoux, 1996).

Scrapie was first shown to be transmissible experimentally in 1936 when sheep were infected by inoculation from diseased animals (Cuille and Chelle, 1936). The cause of scrapie is unknown. It has been shown to be transmitted maternally (Foster et al., 1992) but maternal transmission alone cannot maintain an epidemic. Although not a genetic disease, polymorphisms in the PrP ORF play a large role in affecting scrapie susceptibility. Three polymorphisms at codons 136 (Valine or Alanine), 154 (Arginine or Histidine) and 171 (Arginine, Histidine or Glutamine) in the ovine PrP ORF affect susceptibility (Figure 1-3). The haplotypes which have been observed are (listed in order of increasing susceptibility to scrapie) ARR AHR ARQ ARH VRQ. No ARR homozygote has ever been observed to have scrapie, with the exception of one sheep from Japan (Ikeda et al., 1995), and sheep with this genotype are resistant to experimental infection with scrapie or BSE. In contrast, VRQ homozygotes are highly susceptible to natural scrapie, with most animals of this genotype developing the disease if there is an outbreak in a flock. However, the breed of sheep (and therefore genetic background) also plays a role in scrapie susceptibility as sheep with identical genotypes but from different breeds can have different susceptibilities. For example, Suffolk ARQ homozygotes are relatively susceptible to scrapie while Cheviot ARQ homozygotes are relatively resistant (Hunter, 1997).

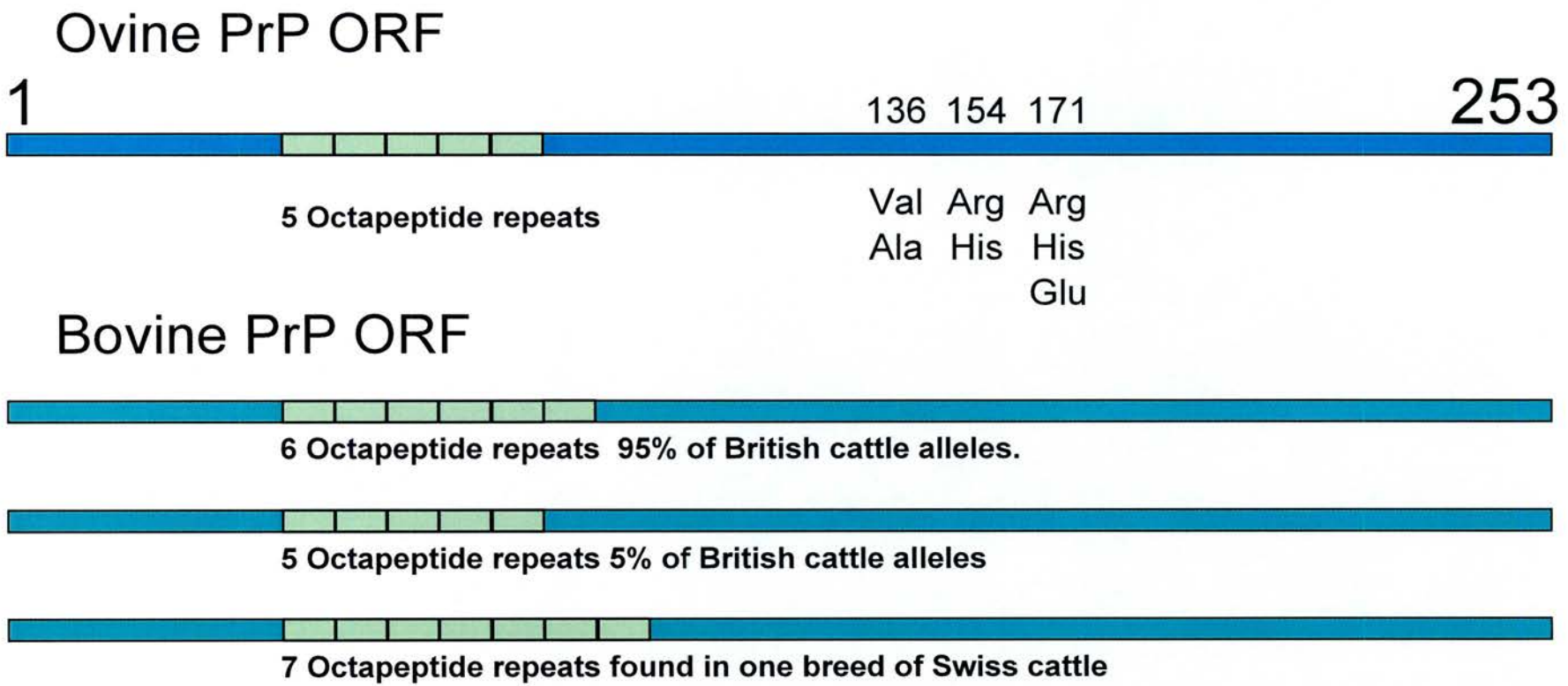


Fig.1-3 Polymorphisms of the ovine and bovine PrP ORF. The ovine PrP ORF contains three polymorphic sites which affect scrapie susceptibility. Codon 136 can encode Valine or Alanine, codon 154 can encode Arginine or Histidine and codon 171can encode Arginine, Histidine or Glutamine. The bovine ORF has only one polymorphism, a variation in the number of octapeptide repeats in the N-terminal of the protein. While most species including sheep have five repeats most cattle have six copies although 5% of alleles contain five repeats.

1.1.9 Bovine Spongiform Encephalopathy

The first reported case of BSE occurred in the UK in 1985 (Wells et al., 1987). Since then the majority of cases have occurred in the UK although several other European countries have also reported cases (Table 1-3). Clinical signs include change in behaviour which develops into ataxia. The pathology is characterised by deposition of PrP mostly in the basal nuclei and medulla with spongiform degeneration found predominantly in the brain-stem (Wells et al., 1992).

Table 1-3 Number of BSE cases by country. Data correct as of January 2001 except UK and Swiss data (December 2000) and Irish data (August 2000) Source:Office International des Epizooties. www.oie.int

Country	Cases
United Kingdom	180501
Republic of Ireland	587
Portugal	503
Switzerland	366
France	218
Germany	25
Belgium	21
Netherlands	9
Spain	7
Denmark	3
Italy	2
Liechtenstein	2
Luxembourg	1

It is thought that BSE spread through the British herd by the use of spinal tissue and offal from infected animals in meat and bone meal (MBM) which was fed to cattle as a food supplement. The epidemic reached a peak in 1992 but has since declined probably due to the banning of bovine materials from MBM in 1988.

The original source of the BSE outbreak is unknown; one possibility is a scrapie-infected sheep that may have been used in MBM preparation and that BSE is therefore

a form of scrapie. An alternative origin of BSE is that it arose spontaneously in a cow as "sporadic" BSE (in an analogous manner to sCJD in humans) and this case was included in MBM and started the epidemic. The scrapie hypothesis is contradicted by the lesion profile of BSE infected mice which is different from that of any known scrapie strain. Although as only a few strains of scrapie have had their lesion profiles typed, BSE could have originated from a scrapie strain which has not been characterised or from a strain which changed significantly when crossing the species barrier from sheep to cattle.

It is unlikely that the question of the origin of BSE can ever be resolved completely. If a scrapie strain identical to BSE was isolated in sheep, it would be difficult to determine if this was a scrapie strain or a case of BSE being transmitted to sheep. The identification of a sporadic case of BSE in cattle (if such cases occur) would be extremely difficult, as it would probably resemble BSE contracted orally.

Unlike sheep and humans, the bovine PrP protein has very little variation; the only polymorphism is in the number of octapeptide repeats. Analysis of common breeds of cattle in Britain has shown that approximately 95% of bovine *Prnp* alleles have six repeats while 5% have five. This contrasts with most other mammalian species where almost all individuals have 5 octapeptide repeats, neither allele has any discernible effect on BSE susceptibility (Hunter et al., 1994). Recently a rare breed of Swiss cattle has been identified with seven octapeptide repeats (Schlapfer et al., 1999) (figure 1-3).

In addition to vCJD in humans, several species of captive or domestic felines (domestic cats, lions, ocelots, pumas) and ungulates (nyala, kudu, gemsbok, oryk, eland and ankole cow) also contracted spongiform encephalopathies at the time of the BSE epidemic (Kirkwood and Cunningham, 1994) (Ministry of Agriculture, 1996). Incubation periods and lesion profiles in mice inoculated with brain homogenates from these animals and vCJD patients show that these TSEs are of the same strain as BSE (Bruce et al., 1994; Bruce et al., 1997), which strongly indicates that these infections were caused by eating BSE infected material. Interestingly no TSE has been reported in dogs despite their similar diet to domestic cats; perhaps the species barrier between dogs and cattle is

greater than that between cattle and felines. Alternatively, perhaps the rate of replication of the agent is so slow in dogs that their life span is not long enough to develop clinical signs of disease.

1.1.10 Transmissible Mink encephalopathy

Transmissible mink encephalopathy (TME) was first observed in farmed mink in the US in 1947 but has since been observed in Canada and Eastern Europe (McKenzie et al., 1996). Affected animals changed behaviour, became hyper-excitable and gradually lost co-ordination. Analysis of TME brains showed the existence of proteinase K resistant PrP and spongiform degeneration in the cerebral cortex and midbrain regions (Marsh, 1992). Inoculation of Syrian Hamsters with TME brain homogenates led to the development of a spongiform encephalopathy showing that TME is transmissible. The source of TME is though to be animal feed, possibly the practice of US farmed mink being fed "downer" cows i.e. cattle unable to stand due to some form of neurological disease. TME has been experimentally transmitted to cattle orally but not to sheep. After three passages of a TME case in hamsters two distinct clinical phenotypes emerged, Hyper (HY) and Drowsy (DY). These two phenotypes are considered strains of TME as they have different average incubation periods 65 days and 168 days for HY and DY respectively. End-stage HY animals have a 100-fold higher titre of infectious units than end-stage DY and the proteinase K resistant PrP proteins produced by the two strains also have distinct characteristics with HY PrP being more proteinase K resistant than DY. PK resistant DY PrP protein has a smaller mass than HY due to differing degrees of PK sensitivity at the N-terminal of the protein.

1.1.11 Chronic Wasting Disease

Chronic wasting disease (CWD) was first identified in 1967 and occurs in farmed and wild mule deer, white tailed deer and elk in the US. The pathology is characterised by spongiform degeneration in the grey matter, astrocytosis and the deposition of amyloid plaques (Williams and Young, 1993). The mode of transmission is unknown. Oral

dosing of CWD brain homogenate to mule deer fawns did result in PK resistant PrP being present in gut associated lymphoid tissue after 80 days, but the experiment did not continue long enough to determine if CWD would develop or if PK resistant PrP developed in the central nervous system (CNS) (Sigurdson et al., 1999). CWD inoculated intra-cerebrally into ferrets caused disease but hamsters are resistant to infection. If a CWD infected ferret is used to infect a hamster then it does develop disease. This is an example of a species barrier being by-passed by infecting a third species.

1.1.12 Experimental TSEs

Although there is no naturally occurring rodent TSE, mice and hamsters have been infected successfully with most strains of TSE and have acted as the main experimental animals in TSE research. Experiments using different inbred lines of mice showed that a particular mouse line could be classed as a long or short incubating line when infected with a TSE strain. Analysing the PrP gene in mice shows that there are two alleles, PrP A and PrP B which are found in short and long incubating lines of mice respectively. However some TSE strains, such as 301V (mouse passaged BSE), show the opposite pattern of incubation periods with allele b mice having shorter incubation periods. Allele A has a leucine residue at codon 108 and threonine residue at codon 189, allele B has phenylalanine at codon 108 and valine at codon 189 (Westaway et al., 1987) it also has a 6kb deletion in the second intron and differences in the upstream putative promoter region (Westaway et al., 1994).

The difference in incubation periods of the two murine PrP alleles A and B may reflect different abilities of the two alleles to adopt PrP^{Sc} conformation due to the two amino acids differences. (Moore et al., 1995). Experiments using a transgenic mouse model where 129/Ola mice (which normally have allele A) had their PrP gene targeted with a construct which altered codons 108 and 189 to allele B resulted in an incubation period with mouse adapted BSE strain 301V of 133 days. This is much shorter than the 244 days seen in wild-type 129/Ola mice, but similar to the incubation period of an allele B

mouse such as the VM/Dk line which has an incubation period of 119 days. There were also differences in pathology and lesion profile between wild type and transgenic B alleles. This difference between the TSE incubation periods of transgenic allele B mice and wild-type allele B mice could be due to the differing genetic backgrounds of the two lines of mice or the intron deletion and upstream region polymorphisms of the *Prnp* gene which were not altered in the transgenic line. Therefore while the amino acid differences of murine allele A and B account for most of the difference in incubation periods between the two lines, variation in sequence outside of the PrP ORF also affects TSE incubation periods.

1.2 The *PRNP* gene and the PrP protein.

1.2.1 Gene structure

The genomic structure of the PrP gene has been determined in rats, mice, cows, sheep, hamsters and humans. All species except humans have three exons. The first two exons are short (50-160bp) and form part of the 5' UTR of the processed mRNA. They are separated by an intron of approximately 2.5Kb. The PrP gene first exon is not well conserved across species but the second exon is highly conserved (Figure 1-4). The presence of two non-coding exons may suggest that they play a regulatory role in either controlling transcription or mRNA stability. In all species, the open reading frame (ORF) is found in the last exon, which is located 10Kb downstream of the second exon. The third exon is approximately 2Kb in length and consists of a short (10bp) 5' untranslated region, the open reading frame and the 3' untranslated region.

The human *PRNP* gene is found on the short arm of chromosome 20. The genomic structure of the human PrP gene is similar to that of other species (Lee et al., 1998), it contains two exons which correspond to the first and third exons in other species with the open reading frame being found in the last exon. It also contains an exon-like sequence within the intron which is highly conserved to the second exons of other

species. The sequence also has consensus splice sites and is found in the same position within the intron relative to the other two exons as the second exon in other species. However this exon has not been found in any PrP transcripts or expressed sequence tagged sequences (EST) and therefore appears to be constituitively spliced out of human mRNA. The PrP mRNA undergoes alternative splicing in hamsters and cattle (Li and Bolton, 1997) (Horiuchi et al., 1997). Both these alternative splicing events involve the first two exons and thus do not affect the open reading frame. In cattle, the first exon can be 53 or 168 nucleotides long depending on which of the two donor splice sites is used. The donor site that produces the 53bp exon 1 is homologous to that used in other species. Both transcripts are found in all tissues except spleen, where only the 53bp exon 1 form is found, and there is no difference in the efficiency of translation of the two transcripts *in vitro*.

In hamsters the second exon is only found in a small proportion of mRNAs as the second exon is usually spliced out of the final mRNA (Li and Bolton, 1997). The level of alternative splicing is tissue dependent with relatively more exon 2 containing mRNA in the frontal cortex, than in the hippocampus or superior and inferior colliculi. In scrapie infected hamsters the levels of exon 2 containing mRNA relative to non-exon 2 mRNA remain constant during the course of an infection in the hippocampus and frontal cortex but rise in the superior and inferior colliculi. The rise in exon 2 mRNA expression coincides with a rise in expression of glial fibrillary acidic protein (GFAP) and a drop in synaptophysin. GFAP is expressed in hamster astrocytes while synaptophysin is expressed in neuronal cells during the formation of synaptic vesicles. The changes in expression levels may represent the proliferation of astrocytes and the death of neuronal cells. Levels of PrP exon 2 containing mRNA are linked to GFAP levels and may therefore be expressed primarily in astrocytes.

In sheep two mRNAs of 4.6Kb and 2.1Kb in length are produced not by alternative splicing but by the termination of transcription at different points through the recognition of different polyadenylation sites (Hunter et al., 1994). The relative ratio of the two transcripts varies depending on the tissue with the greatest amount of 4.6Kb in brain while the 2.1 Kb mRNA is found at its highest levels in spleen.

Figure 1-4 CLUSTALW alignment of the first and second exon sequences of mammalian PrP genes.

There is relatively little conservation of the first exon sequence across mammalian genera (A) although it is highly conserved within ruminants and rodents. The bovine exon 1b has been omitted as it contains 100bp downstream of the end of the conventional bovine exon.
In contrast the second exon (B) is very highly conserved in mammals. Despite the human exon apparently not being used it does not appear to have accumulated mutations in the exon or the consensus splice sites in the flanking regions of intron.

A

```
Bovine  ------------------------------------------------------------
Sheep   ------------------------------------------------------------
Human   CCGCCCGCGAGCGCCGCCGCTTCCCTTCCCCGCCCCGCGTCCCTCCCCCTCGGCCCCGCG
Rat     ------------------------------------------------------------
Murine  ------------------------------------------------------------GG
Hamster ------------------------------------------------------------CACT


Bovine  -----------------GCCAGTCGCTGACAGCCGCAGAGCTGAGAGCGTCTTCTCTCTC
Sheep   -----------------GCCAGTCGCTGACAGCCGCAGAGCTGAGAGCGTCTTCTCTCC-
Human   CGTCGCCTGTCCTCCGAGCCAGTCGCTGACAGCCGCGGCGCCGCGAGCTTCTCCTCTCCT
Rat     -----------------GCGTTGTCAGAGCAGCAGACGGAGTCTGAGCGTCGCGTC----
Murine  ATACTCCCGGCTCCCCCGCGTTGTCGGATCAGCAGACCGATTCTGGGCGCTGCGTCCG--
Hamster GCCCCGCCCGCTCCCCCGCGGCGTCCGAGCAGCAGACCGAGAAGGCACATCGAGTCCACT
                         **          **** *          *  *      **

Bovine            ------G-CAGAAGCAG
Sheep             --------CAGAGGCAG
Human             CACG--A-CCGAGGCAG
Rat               ---------GGTGGCAG
Murine            ------ATCGGTGGCAG
Hamster           CGTCGCGTCGGTGGCAG
                             ****
```

B

```
Rat     GACTCCTGAATATATTTCAAAACTGAACCATTTCAACCCAACTGAAGTATTCTGCCTTCT
Murine  GACTCCTGAGTATATTTCAGAACTGAACCATTTCAACCGAGCTGAAGCATTCTGCCTTCC
Hamster GACTCCTGAATATATTCCAAAACTGAACAATTTCAACTGAGCTGAAGTACTCTGTTTTTC
Ovine   GACTTCTGAATATATTTGAAAACTGAACAGTTTCAACCAAGCTGAAGCAT-CTGTCTTCC
Bovine  GACTTCTGAATATATTTGAAAACTGAACAGTTTCAACCAAGCCGAAGCAT-CTGTCTTCC
Human   GACTCCTGAATATTTTTCAAAACTGAACAATTTCAGCCATGTCTGAGCTTTCCGTCTTCC
        **** **** *** **  * ********  ***** *        **    * *  **

Rat     TAGCGGTACCAGTCCGGTTT-AGGAGAGCCA-AGCCGACT
Murine  TAGTGGTACCAGTCCAATTT-AGGAGAGCCA-AGCAGACT
Hamster TAGAGGTACCAGTTCAGTTT-AGGAGAGTCACAGCAGATC
Ovine   CAGAGACACAGATCCAACTTGAGCTGAATCACAGCAGAT-
Bovine  CAGAGACACAAATCCAACTTGAGCTGAATCACAGCAGAT-
Human   TGGAGGCACAAATCTAGTTT-AGCTGAACCACAACAGATT
          * *  **   *     ** **  **  ** * * **
```

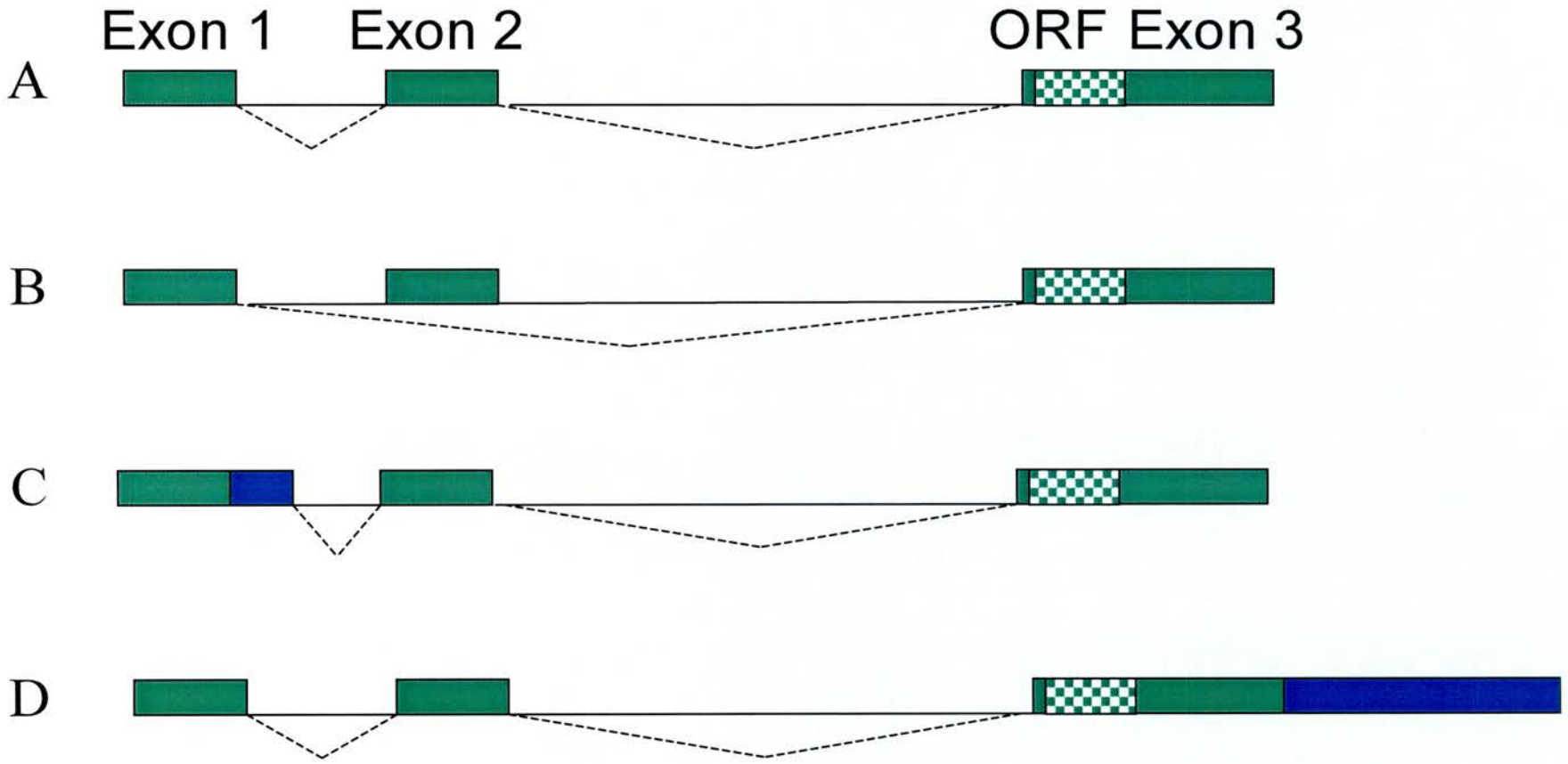


Fig 1-5 Alternative splicing of mammalian PrP gene transcripts: The majority of transcripts in most species are produced by method A where all three exons are included in the final mRNA.All human transcripts and the majority of hamster transcripts are produced by method B where the second exon is spliced out of the final mRNA. Method C is used in cattle where two donor splice sites generate two differently sized exon 1. Method D is found in sheep where two mRNAs are produced not by alternative splicing but by the use of two different poly-adenylation sites in the 3' UTR. Diagram not to scale.

Thus three different species use three different means of producing different PrP mRNAs, cows use different donor splice sites, hamsters use different acceptor splice sites, while sheep use different polyadenylation sites (Figure 1-5). Untranslated regions are thought to play a role in mRNA stability, and it is here that all the mRNA variation occurs. Perhaps differing levels of PrP protein expression in different tissues is achieved by alternative splicing between two mRNA variants of differing stability. Therefore, the putative human second exon may be a regulatory region if it is used in some human transcripts.

1.2.2 The PrP protein

PrP is a di-glycosylated GPI-anchored protein that is highly expressed in astrocytes and neurones but also on a wide variety of other cell types. The human PrP protein has glycosylation sites at amino acids 181 and 197, although both sites are not always glycosylated, allowing PrP to be separated by PAGE into three bands with molecular masses between 33 and 35 kDa which are unglycosyated, mono- and di-glycosylated protein. The amino acid sequence of the PrP protein is highly conserved in mammals. In humans the PrP protein contains 253 amino acids, with a signal peptide of 22 amino acids at the N-terminal which is cleaved off in the endoplasmic reticulum and a 23 amino acid signal sequence at the C-terminal which is cleaved to allow attachment of the glycophosphoinositol anchor (GPI). The N-termini of all PrP proteins have five imperfect glycine-rich octapeptide repeats between amino acids 51 and 91 (Figure1-1). PrP has also been found in chickens (Gabriel et al., 1992) and turtles (Simonic et al., 2000) although in these species the protein has 10 hexa-repeats in its N-terminal rather than the 5 octa-peptide repeats seen in mammals.

NMR analysis of the structure of human, and bovine recombinant mature PrP protein expressed in bacteria in H_2O (Wuthrich et al., 1999) (Zahn et al., 2000) has shown that the structure of PrP is highly conserved between in both species.

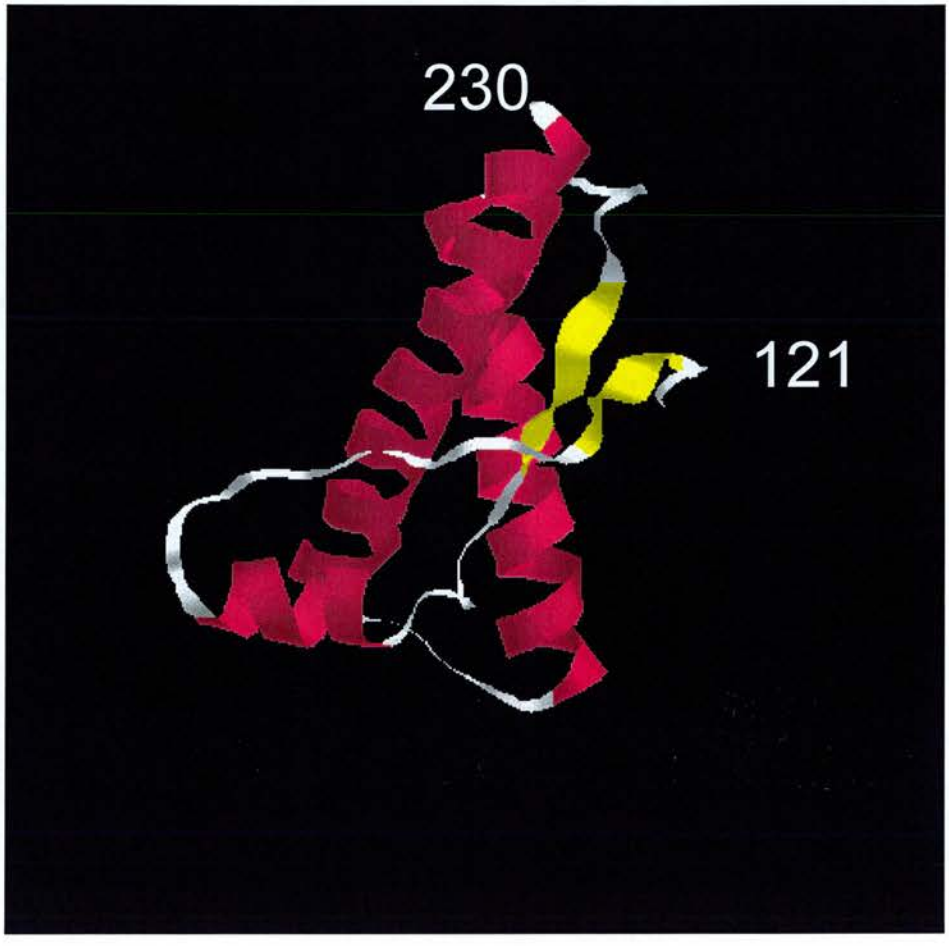


Figure 1-6 NMR structure of human PrP protein. Adapted from Zahn R, et al. 2000. Recombinant PrP consists of an unstructured domain from the first amino acid to approximately amino acid 120 (not shown) The rest of the protein forms globular domain of as defined by NMR by consists of two beta sheets (yellow) and three alpha helices (red). As this protein is produced in *E.coli* it is not glycosylated which could affect its structure, the octapeptides in the N-terminal unstructured region are though to bind copper which could also affect expression.

The region between amino acids 23 and 124 is unstructured in solution while the remainder of the protein (AA125-253) forms a globular domain which contains two β-sheets and three α-helices (Figure 1-6). Codon 129 which affects susceptibility to TSEs in humans is found in the first beta sheet while the five octapeptides are found in the unstructured region.

Examining the distribution of familial TSE mutations shows that, apart from the insertion of octapeptide repeats, most mutations occur in the C-terminal globular domain rather than the unstructured N-terminal (Figure 1-1). This may reflect a change in protein sequence which disrupts the structure of the protein enabling conversion to the PrP^{Sc} conformation which leads to disease.

1.2.3 PrP Knockout transgenic models.

In order to understand the normal role of PrP, the murine *Prnp* gene has been inactivated by gene targeting resulting in five lines of null mice, Zurich I (Bueler et al., 1993), Edinburgh (Manson et al., 1994), Nagasaki (Sakaguchi et al., 1995), RCM (Moore et al., 1999) and Zurich II (Weissmann and Aguzzi, 1999). Abolishing expression of the PrP gene does not affect the embryonic development or reproductive ability of the five lines of mice (Manson et al., 1994). The lack of any major phenotype in null mice does not mean that PrP plays no role in development or adult life but that in its absence other genes may compensate for its absence by up-regulation. Closer examination of the Zurich I and Edinburgh null lines have uncovered minor abnormalities in sleep patterns and circadian activity rhythms (Tobler et al., 1996) which may be related to the severe disruptions in sleep patterns seen in FFI. Weakened GABA receptor mediated fast inhibition and reduction of long term potentiation (LTP) has also been reported in these two lines of null mice (Collinge et al., 1994) (Manson et al., 1995) although this effect has not been replicated in another study (Lledo et al., 1996). Whether these minor phenotypes offer any insight into the function of PrP is unclear.

In order to determine which region of the PrP protein was required to restore

susceptibility to TSEs, transgenic mice were made which had deletions of the N-terminal region PrP (Fischer et al., 1996) (Shmerling et al., 1998). Infecting such mice with scrapie showed that deletion of the region between amino acids (AA) 69 and 84 or between AA 32 and 80 still resulted in proteins which when expressed on a null mouse background made the mice susceptible to scrapie infection. However, a further set of deletions which removed AA 32 to 121 or 32 to 134 resulted in mice which developed severe ataxia at 3-8 weeks of age depending on the size of the deletion and the level of expression. A deletion between AA32 and 106 did not result in this phenotype implying that the region between AA106 and 121 is responsible for this phenotype. However, this phenotype can be rescued by introducing a wild-type gene or even an AA32-93 deleted transgene.

The fact that this phenotype is not seen in PrP null mice and can be rescued by a wild-type gene has led to the proposal of a model where PrP binds to an unknown ligand. In null mice a protein functionally similar to PrP binds to the ligand and effects signal transduction, preventing any phenotype from occurring (Figure 1-7) However removal of the flexible tail (ΔAA32-121) results in a PrP protein which is proposed to bind to the ligand but to be unable to effect signal transduction resulting in cell death leading to the severe ataxic phenotype. It is proposed that the globular domain (as characterised *in vitro* by Wuthrich (Wuthrich et al., 1999)) of PrP between amino acids 125-250 contains a ligand binding site while the flexible domain between AA 32 and 124 effects signal transduction, although presumably the region between 106 and 121 is sufficient to induce signal transduction as the truncated proteins have effectively removed the flexible tail.

1.2.4 Doppel

Three of the null lines (Sakaguchi, RCM and Zurich II lines) develop a syndrome characterised by ataxia, Purkinje cell degeneration and cerebellar atrophy at 70 weeks of age (Sakaguchi et al., 1996). This phenotype can be abolished by introducing a PrP transgene into the line (Nishida et al., 1999). Why these symptoms did not occur in the

other lines may be due to differences in the construct used to knockout the gene. Comparing the transgenic constructs used to inactivate the *Prnp* gene (Figure 1-8) shows that the two non-ataxic lines (Zurich I, Edinburgh) replaced the central part of the third exon (Zurich I) or inserted a reporter gene within the third exon (Edinburgh). In contrast the ataxic lines replaced most of the third exon and some of the 3' end of intron 2 (Nagasaki and RCM lines) or removed the entire third exon, some intronic sequence and sequence downstream of the third exon (Zurich II line).

Analysing expression patterns in the brains of RCM and Nagasaki (Moore et al., 1999) (Li et al., 2000) ataxic mice shows that they have high levels of a chimeric transcript that contains PrP non-coding exons and the coding exon of another gene called *Prnd* (prion-like downstream), which produces a protein called Doppel. *Prnd* is found 16kb downstream of *Prnp* in mice and has a similar genomic structure with two untranslated exons upstream of a single ORF-containing exon although it also has an untranslated exon 3' of the coding exon. The gene is normally expressed in testis and heart but not in the brain or lungs (Moore et al., 1999). Analysing the chimeric transcripts in the brains of ataxic mice showed that they contain *Prnp* exons 1 or 1 and 2 fused to *prnd* exon 3. The chimeric doppel transcripts are transcribed from the PrP promoter and therefore have a PrP-like expression pattern. Transcription runs through the deleted third exon to the doppel gene. The PrP untranslated exons are spliced to the third exon of doppel which contains its ORF although some transcripts also have one or two intergenic non-coding exons. This aberrant splicing is thought to be due to the removal of the PrP third exon acceptor splice site in the construct used to make the three ataxic lines (Figure 1-9).

It is proposed that in a wild-type mouse any transcripts which continues through the PrP third exon into doppel would normally not result in a mRNA containing Dpl exons, as splicing to the third PrP exon would lead to a mRNA containing the poly-A signal in the PrP third exon and the doppel portion of the transcript would be removed. Therefore the removal of the splice donor site in the ataxic knockouts leads to increased doppel mRNA in the brain, presumably leading to ectopic expression of doppel in the brain, which causes the ataxic phenotype.

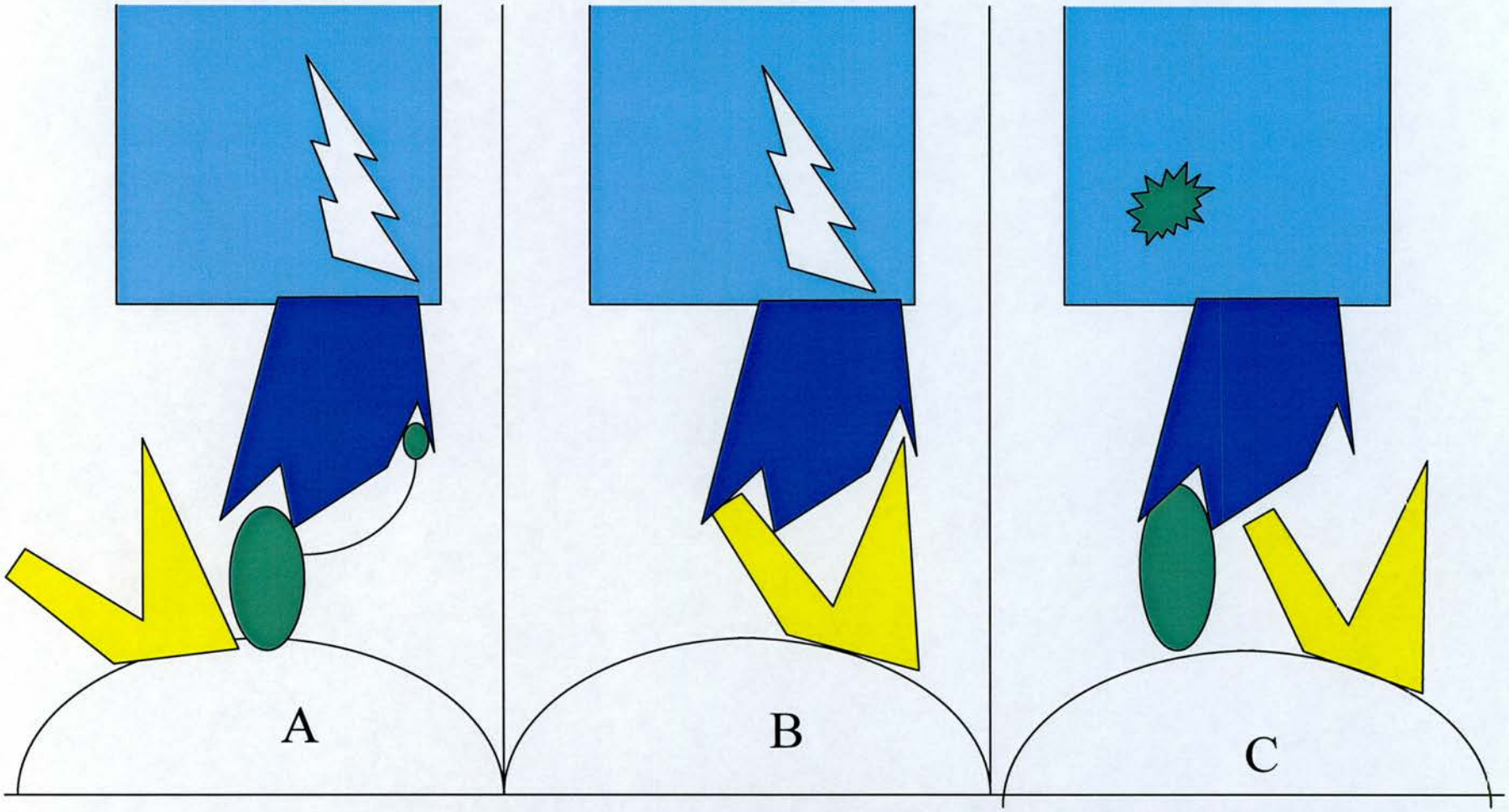


Figure 1-7 The PrP-ligand Model of Shmerling, et al. 1998. A) PrP (green) binds to a ligand (Blue) effecting signal transduction (lightning) in the target cell (light blue) through the flexible N-terminal region (curved line) In null animals (B) another molecule (yellow) binds to the ligand resulting in signal transduction. In Δ32:121 transgenic animals (C) PrP binds but cannot effect signal transductipon resulting in severe ataxic phenotype.

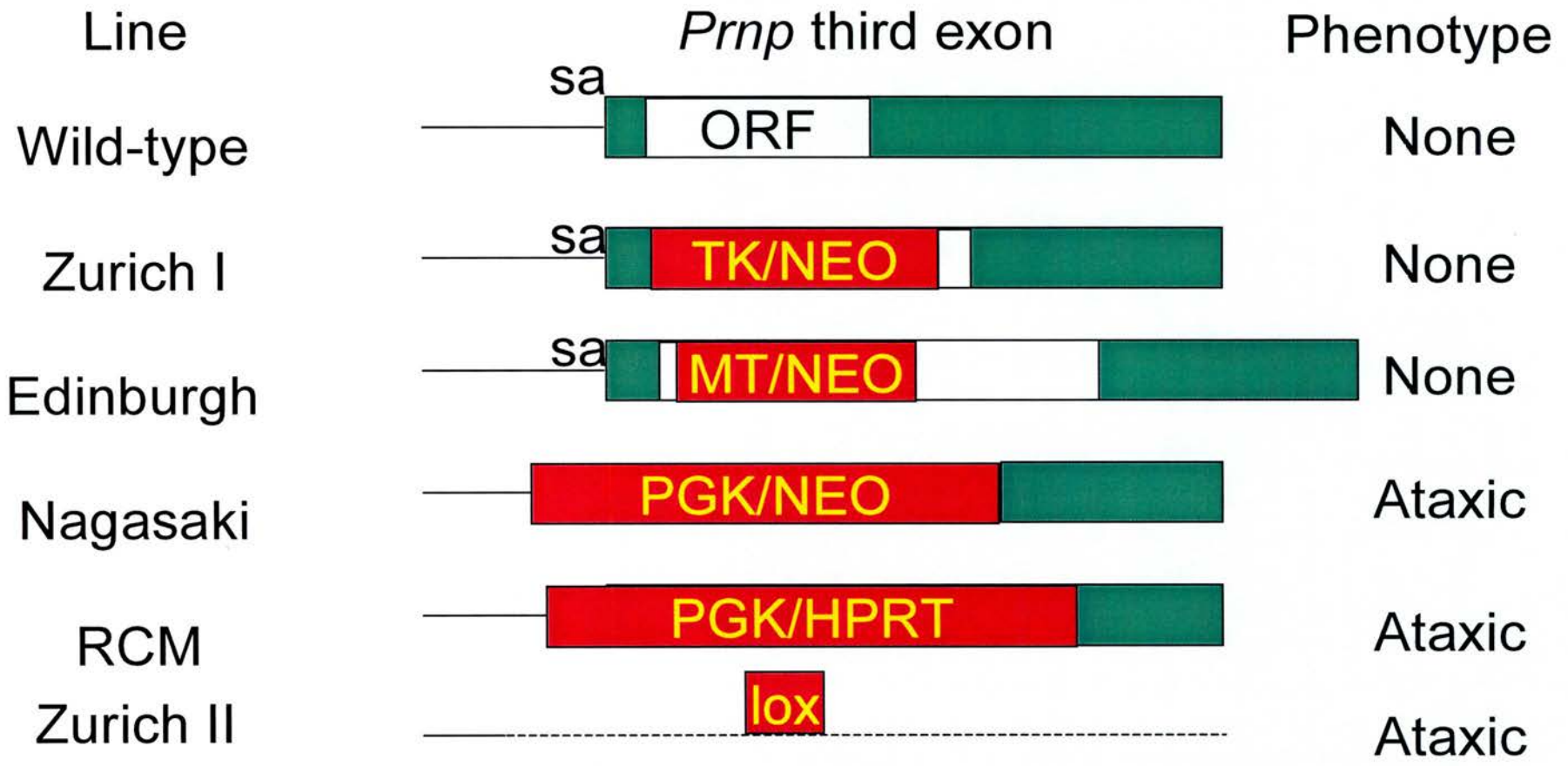


Figure 1-8 Constructs used in the creation of *Prnp* null lines. Five null lines have been made. The Zurich I line replaced most of the PrP ORF (white box) with a selection cassette (red box), the Edinburgh line inserted a selection cassette within the ORF. Both these constructs did not remove the splice acceptor (sa) site at the 5' end of the third exon The Nagasaki and RCM lines replaced the ORF and the 3' end of the intron with the selection cassette while the Zurich II line removed the entire third exon and flanking regions. The ataxic phenotype only occurred in lines which removed the splice acceptor site and the 3' end of the second intron. Diagram not to scale. Adapted from Weissman and Aguzzi 1999.

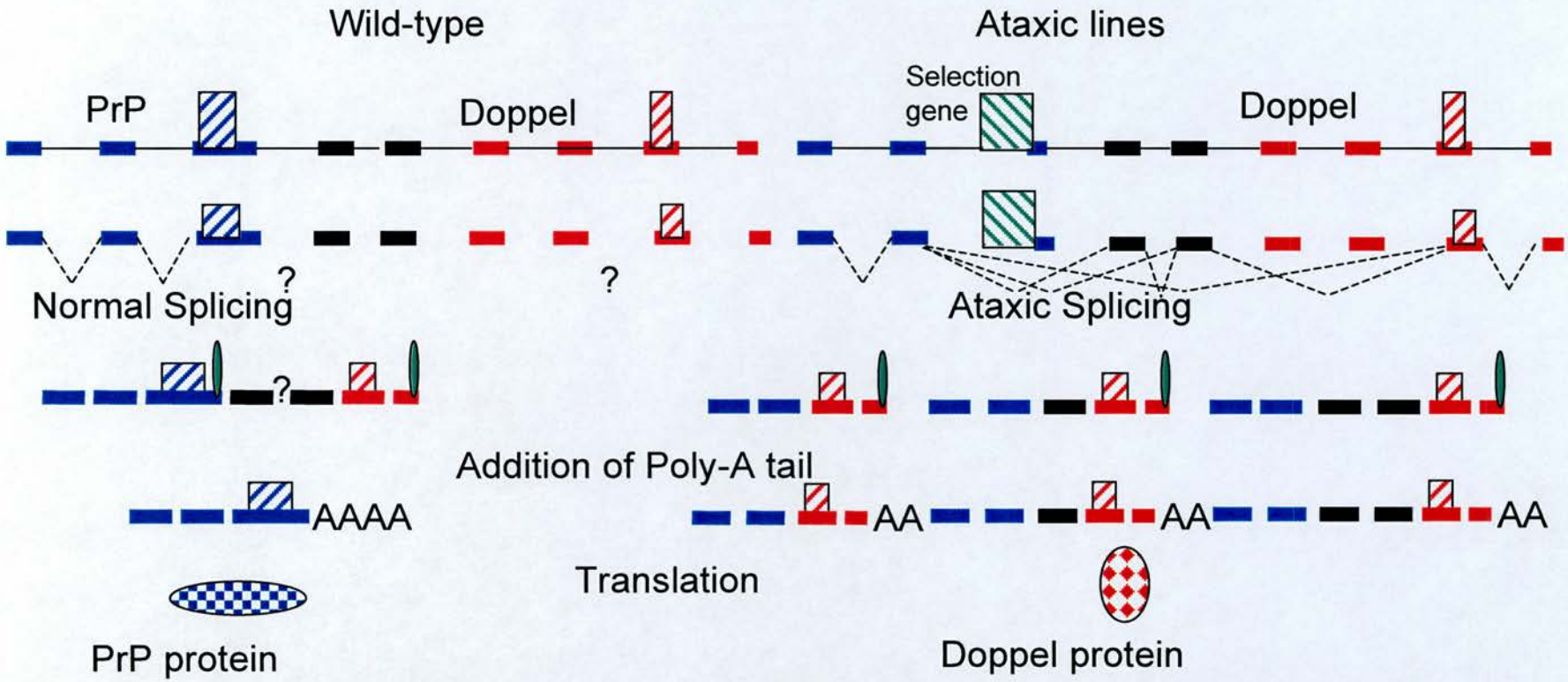


Figure 1-9 Model of abnormal splicing resulting in Doppel expression in brain. PrP (blue) is transcribed normally in wild-type mice but some transcripts continue into Doppel 16kb downstream. Normally splicing results in a transcript containing the PrP coding exon (blue striped box) which contains a Poly-A signal site (green oblong) poly-adenylation at this position leads to the removal of any doppel exons. Translation results in PrP protein (blue oblong). In the ataxic lines the presence of a selection gene(green striped box) removes the splice acceptor site from PrP exon 3 therefore splicing occurs from the PrP second exon to the doppel coding exon(red striped box) and some transcripts also splice to one or two inter-genic non-coding exons (black line). the Doppel poly-a signal is used resulting in the ectopic production of doppel protein (red oval) driven by the PrP regulatory regions. Diagram not to scale. Adapted from Tremblay et al, 1999.

The link between aberrant splicing to Doppel and the ataxic phenotype in null mice could be tested by making a very small deletion of PrP which removes just the consensus splice site at the end of the second intron. This would also differentiate between aberrant splicing and the loss of regulatory elements at the 3' end of the second intron. Doppel is a GPI anchored protein, which contains two consensus glycosylation sites, however its function and role in the ataxic phenotype are unknown. Analysis of doppel sequence in humans has shown that there are several polymorphisms in the protein coding region (Mead et al., 2000), but no doppel allele shows an association with sCJD or vCJD. The ataxic phenotype seen in three of the five null is probably an artefact of the particular transgenic strategies used in those lines and as the protein is not normally expressed in brain it probably plays no role in TSEs.

1.2.5 PrP and copper

The strongest candidate for the function of PrP is that of a copper binding protein which may have superoxide dismutase activity. Analysis of the brains of PrP null mice has shown that they contain less copper in membrane-rich extracts than wild-type (Brown et al., 1997). The N-terminal region of mature human PrP protein (amino acids 23-98) produced in *E.coli* has been shown to bind copper (Hornshaw et al., 1995). While the addition of copper to murine neuroblastoma cells stimulates PrP to undergo endocytosis, this effect is absent in cells expressing mutant proteins lacking the N-terminal (Pauly and Harris, 1998).

One of the physiological functions of copper is to act as a co-enzyme for Cu/Zn super-oxide dismutase (SOD), in PrP null mice there are decreased levels of Cu/Zn SOD activity and increased levels in transgenic mice over-expressing PrP (Brown and Besinger, 1998). This effect on activity is post-translational as Northern and Western analysis shows that Cu/Zn SOD expression and protein levels are not affected, implying that PrP aids in the functioning of the enzyme, possibly by transporting copper into the cell. (Brown and Besinger, 1998). Mn SOD activity is increased in null mice probably to compensate for the decreased levels of Cu/Zn SOD activity. Despite this up-

regulation of Mn SOD, PrP null cells are more susceptible to oxidative stress *in vitro.*

The PrP protein itself has shown SOD activity when bound to copper. (Brown et al., 1999) Thus it may play a dual role in transporting copper to the cytoplasm and also acting as a SOD in its own right. Analysing PrP binding to other cations has shown if recombinant PrP protein is denatured in urea and allowed to renature in the presence of various cations it can bind to Cu, Mn, Ni or Zn but it cannot bind Mg, Fe or Ca. Mn bound PrP also exhibits SOD activity although at only 50% of the level of activity of Cu bound PrP. Intriguingly when stored at 4°C for two weeks Mn bound PrP loses its SOD activity while the copper bound PrP was unaffected. This "aged" Mn bound PrP is slightly proteinase K resistant and has a higher percentage of beta-sheet secondary structure (Brown et al., 2000). The ability to create proteinase resistant PrP with increased beta-sheet is possibly a pre-cursor of PrP^{Sc} and that the conformation change may be mediated by metal binding.

Recent studies on recombinant murine PrP (Brown et al., 2000) have shown differences in SOD activity between the two murine PrP alleles that affect incubation periods. Four recombinant proteins PrP A (codon 108=L, codon 189=T) PrP B (108=F, 189=V) PrP A108 (codon 108=F, codon 189=T) and PrP A189 (codon 108=L, codon 189=V) were compared for SOD activity and copper binding. PrP A had lower levels of SOD activity than PrP B. PrP A189 had equivalent levels to PrP B while PrP A109 had intermediate levels of activity. There was no difference in the amount of copper bound (4 atoms) between any of the proteins. Aged proteins (1 month at 4°C) saw a gradual decrease in SOD activity which was more pronounced in the two intermediate proteins (PrP A108 and PrP A189) than in the naturally occuring alleles PrP A and PrP B. When the proteins were aged the two intermediate forms spontaneously formed aggregates and fibrils and developed proteinase K resistance in solution. This may imply that PrP can spontaneously aggregate, although these experiments used recombinant proteins at high concentrations in solution, which may mean they are not applicable to the physiological behaviour of PrP *in vivo*.

1.3 PrP and TSEs

1.3.1 Models: Prion Vs nucleic acid models

There are three competing theories to explain the mechanism of TSEs, the most widely accepted of which is the Prion theory which states that PrP^{Sc} is the sole infectious agent and that the conversion of PrP^{C} to PrP^{Sc} is catalysed by PrP^{Sc}. The viral theory proposes that TSEs are viral diseases and the formation of PrP^{Sc} is a form of pathological damage rather than an integral part of the disease process. The virino theory states that a nucleic acid is the infectious agent which is associated with PrP^{Sc} and that this protects the nucleic acid from mutagens and nucleases.

1.3.2 The prion theory

The prion theory, also known as the protein-only theory states that protein-protein interactions convert PrP^{C} to PrP^{Sc} .The mechanism of conversion from PrP^{C} to PrP^{Sc} is unknown although two models have been proposed. The heterodimer model (Figure1-10) (Pruisner, 1991) proposes that a PrP^{Sc} molecule interacts directly with a PrP^{C} molecule changing its conformation to that of a PrP^{Sc} protein which in turn can act in converting more PrP^{C} molecules. In this model, a single PrP^{Sc} protein is enough to start the chain reaction of conversion. The initial PrP^{Sc} molecule is proposed to come from infection (Kuru, Iatrogenic CJD and vCJD) or the very rare misfolding of a normal protein (sCJD). The mutations in PrP that cause familial prion diseases are thought to predispose the protein to misfolding and somatic mutations in PrP are also proposed as another mechanism for the initiation of sCJD.

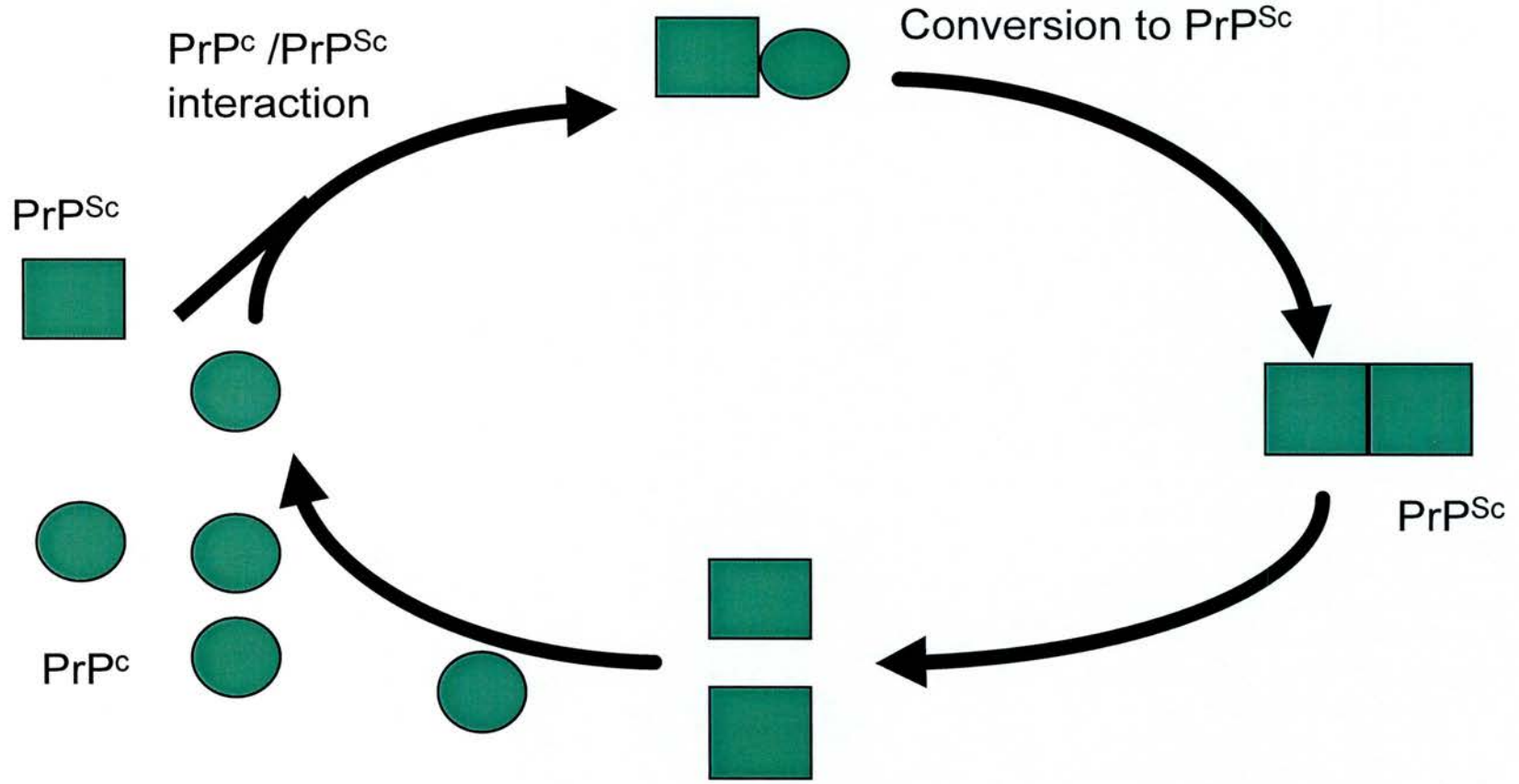


Figure 1-10 The Prion Heterodimer Model: PrP^{C} interacts with a PrP^{Sc} protein this interaction results in the conversion of PrP^{C} to PrP^{Sc} a the newly created PrP^{Sc} protein can then interact with and convert other PrP^{C} proteins resulting in a chain reaction which leads to an accumulation of PrP^{Sc}. The origin of the initial PrP^{Sc} seed could be due to spontaneous conversion of a PrP^{C} with familial TSE mutations proposed to have a very high rate of spontaneous conversion.

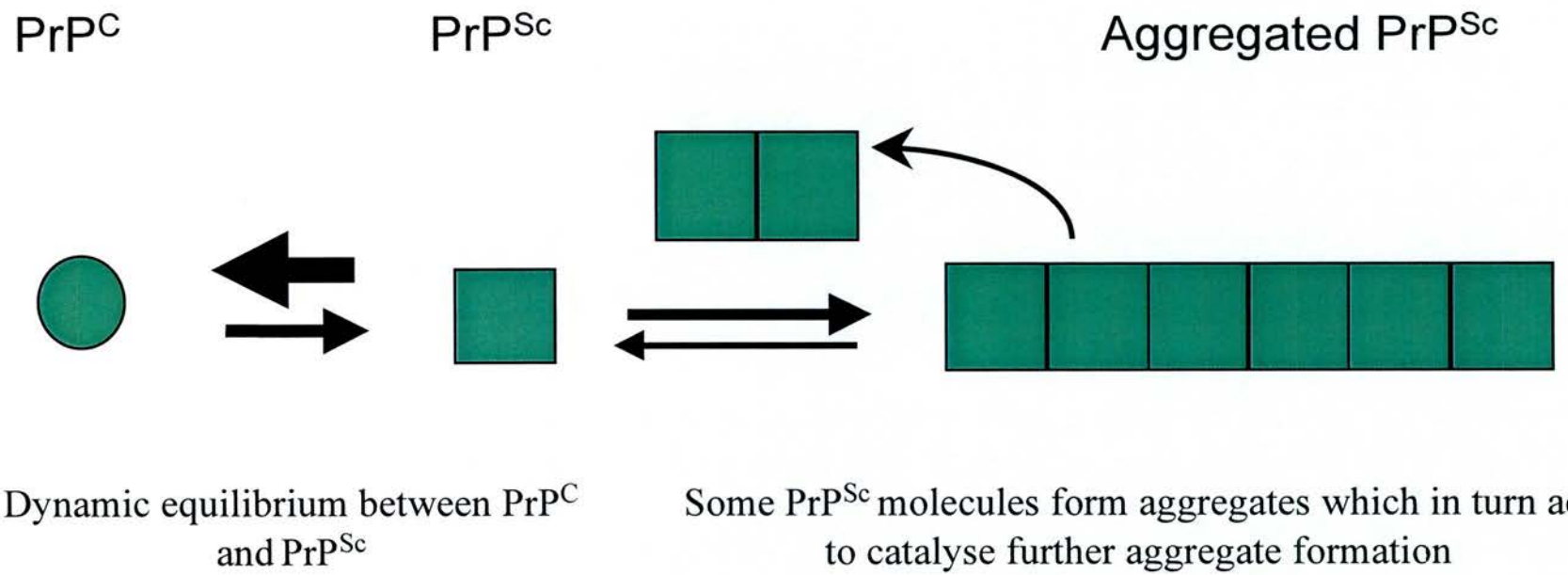


Figure 1.11 The Prion Seeding Model: The two conformations which PrP can adopt (PrP^C and PrP^{Sc}) are proposed to be in dynamic equilibrium but with the PrP^C form predominating, if some PrP^{Sc} molecules aggregate this fixes the molecule in the PrP^{Sc} structure and catalyses the aggregation of other PrP^{Sc} molecules. The origin of the initial PrP^{Sc} seed could be a spontaneous aggregate or come from infection. Familial mutations are proposed to greatly increase the likelihood of aggregation. Diagram adapted from Come *et al.* 1993

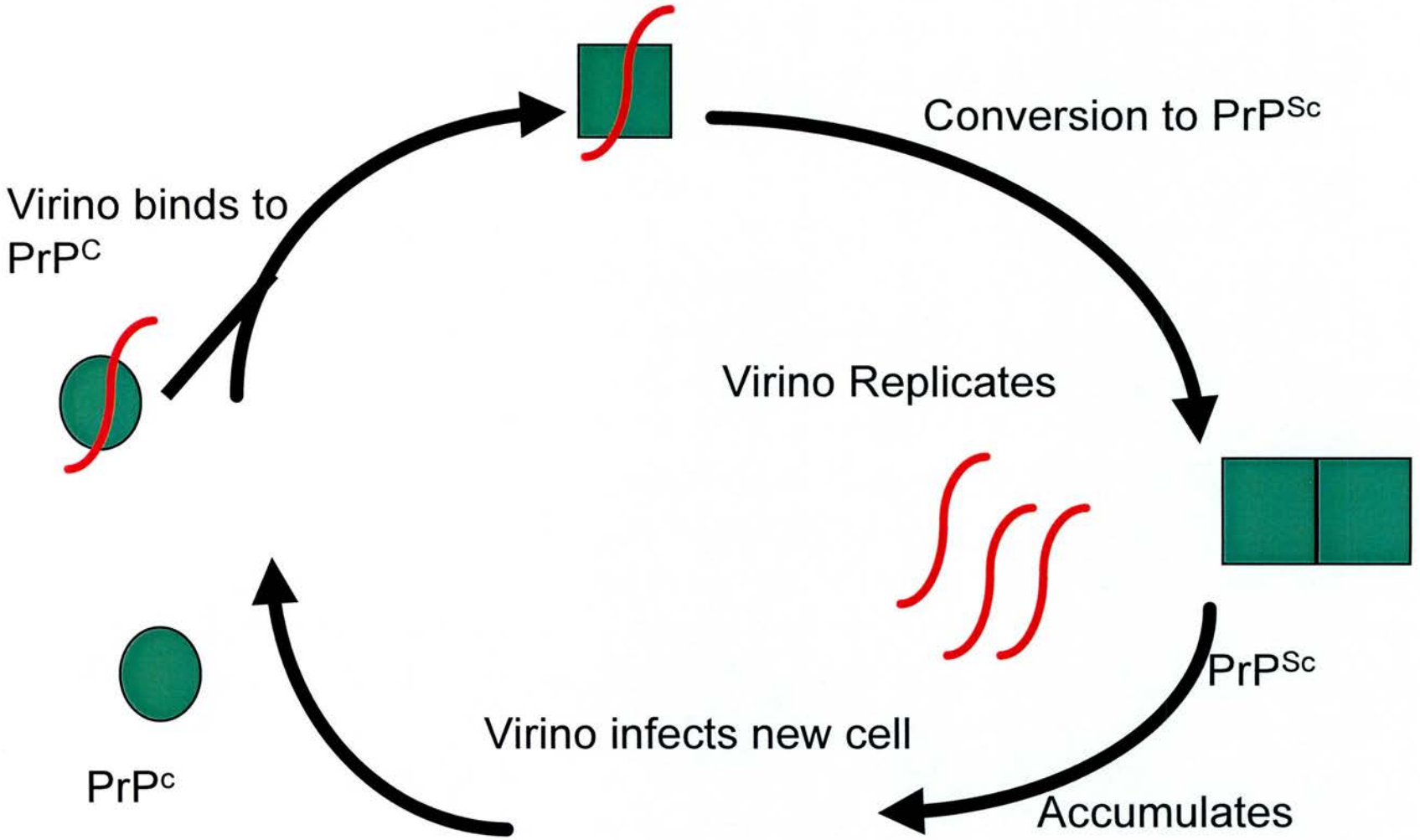


Figure 1.12 The virino model: Virinos are proposed to consist of nucleic acid (red line) which bind to PrP^{C} (green circle) converting it to PrP^{Sc} (green square). The virus replicates within the cell allowing it to infect other cells while bound to PrP.

The seeded polymerisation model of PrP^{Sc} formation (Come et al., 1993) (Figure 1-11) proposes that the PrP protein is in dynamic equilibrium between the PrP^{Sc} and PrP^{C} conformations with the PrP^{C} conformation pre-dominating. If however a large amount of PrP^{Sc} accumulates, the molecules aggregate together and the PrP^{Sc} conformation becomes favoured, this aggregate also serves as a scaffold that fixes other single molecules transiently in the PrP^{Sc} conformation into the PrP^{Sc} conformation permanently. Portions of the aggregates occasionally break off and initiate seeding in another part of the brain. This seeding mechanism requires a large aggregate of PrP^{Sc} to start the process. In TSEs this may come from infection or the chance accumulation of large numbers of PrP^{Sc} molecules, PrP proteins carrying familial TSE mutations are more likely to adopt the PrP^{Sc} structure for longer periods of time and thus have more chance to form aggregates eventually leading to familial prion diseases.

1.3.3 The viral theory

One aspect of TSEs which the prion hypothesis has not satisfactorily addressed is the existence of multiple strains of agent, whose diversity resembles that of a nucleic acid rather than a single protein. This has led many investigators to retain the original hypothesis of a novel form of virus which is able to survive processes, such as UV-irradiation or nuclease treatment (Alper, 1985), which degrade nucleic acids, but which has not yet been detected. Also not all TSEs exactly fit the prion theory, some diseases can be highly infectious with very little PrP^{Sc} produced e.g. FFI, and not all of the familial diseases have been successfully transmitted to experimental animals. Under the viral theory, the conversion of PrP^{C} to PrP^{Sc} is part of the disease process rather then the cause. Evidence supporting the viral theory is a study where BSE inoculated mice developed TSE symptoms but 50% did not have proteinase K resistant PrP (Lasmezas et al., 1997). In addition, symmetrical structures have been identified in scrapie-infected hamsters brains (Diringer, 1996) which resemble virus particles although these structures are smaller than the smallest known virus.

1.3.4 The Virino theory

A similar theory to the viral theory is the virino theory, which postulates that the agent responsible for disease is a nucleic acid, termed a virino, which binds to the PrP protein (Dickinson and Outram, 1988). The virino is postulated to catalyse the conversion of PrP^{C} to PrP^{Sc} and bind to PrP^{Sc} which may protect it from the processes (nuclease treatment, UV-radiation)(Figure 1-12) which have been used to eliminate any nucleic acids from a PrP^{Sc} preparation. Thus, brain inocula from an infected animal will contain virino as well as PrP^{Sc}. According to the viral and virino theories the mutations in PrP seen in familial prion diseases do not in themselves cause disease but make carriers extremely susceptible to particular viral/virino strains which are either ubiquitous in the environment (Manuelidis and Manuelidis, 1993) or present as a retrovirus in mammalian genomes (Murdoch et al., 1990). PrP may possibly act as a receptor for either the virus or the virino which would explain why null mice are resistant to disease and over-expression of PrP decreases incubation periods.

1.3.5 Transgenic models of familial TSEs.

According to the prion theory familial TSEs are caused by mutations in the PrP ORF that pre-dispose the protein to spontaneously convert to PrP^{Sc}. The fact that the mutations are not always fully penetrant but are transmissible has lead to the proposition under the viral and virino theories that these mutations do not directly cause disease but greatly increase an individual's susceptibility to some form of virus or virino to which most individuals are highly resistant. In order to discriminate between the two alternatives three familial TSEs have been modelled in transgenic mice.

P102L is the most common form of GSS. In order to model the disease, two different types of transgenic mice have been made one by microinjection and one by homologous recombination. In microinjection, the transgenic construct is injected into fertilised eggs where it integrates randomly into the genome. Varying numbers of copies of the transgene can be introduced by this method leading to varying levels of PrP expression.

Homologous recombination in contrast, replaces the endogenous gene with a mutated transgene. Both the microinjection and the homologous recombination constructs contained leucine at codon 101 rather than proline (codon 102 in humans is equivalent to codon 101 in mice). Using the microinjection technique several lines were made which produced varying levels of P101L (Hsiao et al., 1990). Mice with the highest levels of P101L expression spontaneously develop a neurodegenerative disease characterised by spongiform degeneration and deposition of amyloid plaques in the brain. However the PrP of these animals is not resistant to proteinase K digestion. These spontaneously sick animals were able to transmit disease to other P101L transgenic mice and hamsters (Hsiao et al., 1994). Another transgenic line which produces lower levels of 101L (Kaneko et al., 2000) only develop spontaneous disease about 20% of the time and then only at 400-600 days of age compared with 150-190 days in the high-expressing line. The possibility exists that the phenotype is due to over-expression caused by the extra transgene rather than the P101L mutation contained within it. This is similar to what occurs in transgenic mice which over-express the wild-type mouse, hamster or sheep proteins who spontaneously develop neuromuscular disease and suffer vacuolation of the grey matter in the absence of any infection. These over-expressing mice also do not produce proteinase resistant PrP (Westaway et al., 1994). Alternatively, it may be argued that the over-expressing lines are a more accurate representation of GSS in humans, which takes 50 years on average for disease to develop. Therefore in the limited life-span of a mouse over-expression is needed to compensate for the shorter life-span.

Another line of transgenic mice replaced the endogenous gene with a P101L mutant by homologous recombination (Moore et al., 1995). These mice did not develop spontaneous disease but 101LL (Homozygous mutant) mice were 100% susceptible to primary GSS P102L inoculum (i.e. brain homogenate from a GSS P102L patient) in contrast to wild-type mice where only 14% of wild-type mice developed disease (Manson et al., 1999). They also had a shorter incubation period (288 days) than wild type mice (456). Mice heterozygous for the mutation (101PL mice) were less susceptible to GSS P102L primary inoculum than 101LL mice and had longer incubation periods.

The 101LL mice infected with GSS had no detectable PrP deposition in the brain and produced extremely low levels of PrP^{Sc} when analysed by Western blot. Heterozygous mice varied considerably in the amount of deposited PrP and PrP^{Sc} produced. However, on subsequent passage from GSS infected 101LL mice into other 101LL 101PL and 101PP mice all three genotypes were 100% susceptible and had much shorter incubation periods than primary GSS passage. The pattern of incubation periods was reversed however when the mouse adapted scrapie strain ME7 was inoculated into the three genotypes, with wt mice having shorter incubation periods than 101PL or 101LL mice (although there was no significant difference between 101PL and 101LL mice). This result implies that a PrP protein containing the 101L mutation is not inherently more unstable than wild-type (101P) mouse protein but may suggest that it interacts with different efficiencies with the infectious agents that cause GSS or ME7 scrapie.

Mice containing nine extra copies of the octapeptide repeat have been made in order to model the forms of familial CJD and GSS associated with insertion of octapeptide repeats in the human PrP protein (Chiesa et al., 1998). Of five lines made on a wild-type background the two lines which expressed low levels of protein did not develop any abnormal phenotype, while the three that expressed the mutant protein at levels equal to or above that of the normal protein developed spontaneous neurodegenerative disease at 100-200 days. When bred to null mice the onset of disease was reduced to 100 days for mice hemizygous for the transgene array and 60 days for mice homozygous for the transgene array. The pathology of these mice was characterised by atrophy of the cerebellum and the production of PrP that is slightly resistant to proteinase K digestion but this disease has not been shown to be transmissible so it is difficult to determine whether or not this syndrome is a model of a TSE or an artefact of over-expression of a mutant PrP.

The GSS variant A117V is unusual in that very little proteinase resistant PrP has been found in brains and although it is termed a transmissible spongiform encephalopathies it has not yet been transmitted to experimental animals (Brown et al., 1994). Analysis of A117V Syrian Hamster proteins expressed in mice has shown that the A117V protein and proteins with mutations near codon 117 adopt an unusual topology with respect to

the membrane. *In vitro* studies that digest with cytosolically deposited proteinase K results in the creation of two partially digested forms of PrP in addition to the extra-cellular secreted PrP. These digested forms consist of a C terminal transmembrane (Ctm) form, which becomes lodged in the membrane and had its N-terminal removed by the cytosolic PK, and an N-terminal transmembrane (Ntm) form which has had its C-terminal region digested. All forms of PrP appear to have a percentage of proteins in all three conformations but the A117V mutant and other similar mutations have a high propensity to adopt the Ctm form. In transgenic mice over-expressing two of these proteins, spontaneous disease occurs after 60 days. The pathology was characterised by vacuolation and gliosis in one mutant, and neurodegeneration and gliosis in the other mutant. Neither deposited any PrP in either plaques or diffuse deposits. This contrasts with A117V GSS patients who have distinctive PrP plaques. Analysis of brain tissue from a GSS patient allowed the determination of the topology of its PrP by the fact that Ctm PrP is very slightly proteinase resistant although much less so than conventional PrP^{Sc} and also has a smaller mass (19kDa) than PrP^{Sc} (27-30kDa). Treatment of brain extracts with mild and harsh proteinase K showed that there was no PrP^{Sc} but there was PrP Ctm implying that GSS A117V may be a different type of disease from other familial and non-familial PrP diseases. According to the virus/virino hypothesis GSS A117V may be a genetic disease directly caused by the mutation and thus not infectious which would explain why GSS A117V has not been successfully transmitted to experimental animals. (Mallucci et al., 1999) and others although it is then not clear why A117V resembles other forms of GSS in clinical signs and brain pathology.

Of the three familial TSEs modelled in mice only P101L has produced spontaneous disease and then only when the protein is over-expressed. However, this mouse line did not produce proteinase K resistant PrP. The other P101L GSS model did not develop a spontaneous TSE but its susceptibility to GSS was increased while susceptibility to one strain of mouse-adapted scrapie was decreased. This could reflect either differing incubation periods to differing TSE strains under the virino hypothesis or that the single amino acid change at codon 101 has a large effect on the interactions between PrP^{C} and PrP^{Sc} molecules in different conformations according to the prion theory.

1.3.6 The 129 Polymorphism and PrP theory

The increased susceptibility of codon 129 homozygotes to CJD may reflect the protein-protein interactions between PrP^{C} and PrP^{Sc} which are required for conversion from PrP^{C} to PrP^{Sc} under the prion hypothesis. The differences between proteins in heterozygotes at codon 129 or 219 may hinder effective conversion thus preventing the onset of the disease. The presence of 50% of PrP with a different amino acid structure may be similar to the species barrier which is controlled by the number of amino acid differences between two species. Under the virino hypothesis, different strains of virino may have preferences for particular PrP alleles and thus homozygotes may be more susceptible than heterozygotes (Palmer and Collinge, 1992).

1.3.7 *In vitro* conversion

The prion theory predicts that PrP^{C} can be converted to the infectious, proteinase K resistant PrP^{Sc} by direct interaction: so far this has only been shown to occur in infected animals where the possibility of a viral component cannot be excluded. In order to show that the infectious agent is entirely proteinaceous and lacks nucleic acid PrP^{C} must be converted to PrP^{Sc} in a cell free system (to prevent viral replication) and this PrP^{Sc} should be infectious. Attempts at cell-free conversion has had some success as mixing PrP^{Sc} with labelled PrP^{C} under certain conditions then digesting with proteinase K has shown some labelled proteinase K resistant PrP (Caughey et al., 1995), implying that the interaction with the PrP^{Sc} has caused PrP to become proteinase resistant. However, the conversion process requires an excess of PrP^{Sc} and very little labelled PrP^{Sc} is produced. Although infectious PrP molecules have not been created in a cell-free system protein-protein interactions of PrP can confer specific levels of proteinase K resistance on PrP^{C}. The PrP^{Sc} proteins of different strains of transmissible mink encephalopathy (HY and DY) have different sensitivities to proteinase K digestion. These differences have been transmitted to PrP^{C} molecules in a cell free system. (Bessen et al., 1995).

The fact that the proteinase K resistant protein produced *in vitro* is in a solution of

infectious PrP^{Sc} makes it impossible to determine if the *in vitro* converted proteinase K resistant PrP is infectious. In order to discriminate (Hill et al., 1999) between *in vitro* converted recombinant PrP and the excess of PrP^{Sc} needed to initiate conversion, transgenic mice were used. Wild-type mice are resistant to hamster-adapted scrapie strain Sc237 but transgenic MH2M mice which contain part of the hamster protein with the murine N and C-terminals of the protein are susceptible to hamster-adapted scrapie and subsequent passages can infect wild-type mice. Hamster scrapie Sc237 PrP^{Sc} was used to *in vitro* convert recombinant MH2M protein made in a murine cell line. Proteinase K resistant MH2M PrP was produced but when injected into wild-type mice, it did not prove to be infectious but this may be due to the low amounts of PrP^{res} created. In wild-type mice *in vivo* produced MH2M PrP^{Sc} causes disease in 180 days but it is uncertain if the level of PrP^{res} produced is equivalent to the level used in the positive control experiments. Therefore the central tenet of the prion theory, that PrP^{c} can be converted to infectious PrP^{Sc} by protein interaction has not yet been demonstrated in a cell free system. This may be due to deficiencies in the cell-free system or that some other agent is required to catalyse conversion.

1.3.8 Species barrier

The species barrier, where transmissions from one species to another are less efficient than transmissions within species, is thought by the prion hypothesis to represent the difficulty a PrP^{Sc} protein from another species has in converting the PrP^{C} of the host due to structural differences between the two proteins Once conversion has been achieved, subsequent conversions are easier as they occur between PrP molecules of the same species which are often identical. Transgenic experiments to investigate the species barrier have been carried out in mice carrying copies of the hamster PrP gene in addition to murine PrP. When a transgenic mouse with 30-50 copies of the hamster PrP gene is infected with hamster scrapie strain Sc237 it develops the disease as quickly as a hamster, while mice with 4-8 copies of the hamster gene take a longer time but are still quicker than wild-type controls. The PrP^{Sc} proteins formed were from hamster PrP^{C} proteins rather than the endogenous mouse PrP^{C} proteins. This experiment shows that

the species barrier can be removed by expressing the PrP protein of the donor species in the recipient (Scott et al., 1989).

However, when these mice are infected with mouse-adapted scrapie the incubation period is longer than in a normal mouse and the PrP^{Sc} produced consists of murine PrP rather than hamster PrP, implying that the hamster PrP interfered with the interaction between the mouse PrP and mouse PrP^{Sc}. A different construct which contained the mouse PrP gene with the hamster sequence between codons 94 and 188 (MH2M) was made in attempt to make a PrP protein which could interact with both mouse and hamster PrP. The transgenic mice carrying this protein were equally susceptible to hamster or murine PrP^{Sc} and when they developed disease it could infect either hamsters or mice with similar incubation periods implying that this chimeric protein is unaffected by the species barrier between hamsters and mice.

However, introducing a human PrP gene into mice which were subsequently infected with human prions had no effect on susceptibility to inoculation to the human TSE. Although it is more difficult to infect a mouse with human TSEs than hamster TSE due to the greater sequence differences between rodents and primates than between rodents. When transgenic mice carrying human transgenes were crossed onto a PrP null background they became more susceptible to human TSE infection (Telling et al., 1995). A chimeric protein that contained human protein flanked by mouse protein (MHuM) was more susceptible to human TSEs than wild-type mice even on a wild-type background.

These results led to the development of the "protein X" theory which suggested that another protein was required to allow the conversion from PrP^{C} to PrP^{Sc} to take place by binding to specific regions in the C or N-terminals of the PrP protein. A mouse which expresses both murine and human PrP will only express murine protein X, which will preferentially bind to murine PrP allowing conversion of murine PrP by murine PrP^{Sc}. However, human PrP will not be converted by human PrP^{Sc} as human PrP has less affinity for protein X. On a null background human PrP can bind murine protein X and allow conversion to PrP^{Sc}. In the chimeric protein, protein X can interact with the murine

sequence, N-terminal or C-terminal while the human central region can interact with the human inoculum. However protein X has not yet been identified nor has it been shown that the molecule responsible for the effect is actually a protein, therefore further evidence is required to confirm the existence of protein X.

1.3.9 Prion and virino models of TSE disease

According to the prion theory sCJD occurs either by the spontaneous conversion of a PrP^{C} protein to PrP^{Sc} (Prusiner, 1996) or by a mutation of a PrP gene in a single cell to a familial CJD mutation which is then highly pre-disposed to assume the PrP^{Sc} structure. Both of these events produce a PrP^{Sc} protein that then propagates, either by the seeding or heterodimer interaction model resulting in disease. Alternatively, under the viral or virino hypothesis these individuals have become infected with a strain of virus/virino. However, there is as yet no evidence indicating that any of these hypotheses of the initiation of sCJD are correct nor whether host genetic susceptibility factors act by increasing the frequency of these rare events or by increasing the likelihood of disease developing once one of these events have occurred.

Familial TSEs are a more difficult disease to reconcile with the either the prion or virus/virino theory. Under the prion theory it is proposed that the conformation of the mutant protein makes it more likely to spontaneously convert to PrP^{Sc} which eventually leads to disease. Yet, the age of onset in familial TSEs is similar to sCJD despite a familial CJD patient producing 50% mutant protein from birth while a sporadic patient must await a spontaneous conversion event or a mutation. However, this effect is also seen in Alzheimer's disease where 10% of cases are linked to mutations in β-amyloid protein (APP), presenilin 1 and 2 (Das and Lal, 1997) while the other 90% of cases are apparently sporadic. Perhaps an age-dependent factor or many events unrelated to PrP are required to initiate disease which would explain the similar age-of-onset in sporadic and familial CJD. Under the virino hypothesis the mutation is not directly responsible for the disease but makes an individual orders of magnitude more susceptible to the agent, with GSS, FFI and familial CJD individuals susceptible to different virino strains.

The long incubation period of a virino infection would explain why the age of onset is similar in sporadic or familial CJD. Non-penetrance of familial TSE mutations may simply reflect individuals who are resistant to infection or have avoided infection.

Familial CJD and GSS mutated proteins are thought to be more likely to convert to a PrP^{Sc} conformation than the wild-type protein. Yet, sporadic disease is clinically and pathologically similar to CJD rather than GSS. Why are no sporadic GSS cases observed? Familial CJD is as common as GSS, as there are a similar number and type of mutations there is no reason why a somatic mutation to a GSS mutation should occur less frequently than a CJD mutation. Perhaps the spontaneous conversion of a wild-type protein to PrP^{Sc} is only possible to the conformation that results in CJD-like disease.

1.3.10 Yeast Prions

Evidence supporting the prion hypothesis has come from the field of fungal genetics. Two non-mendelian phenotypes in *Saccharomyces cerevisiae* whose mode of inheritance is uncertain are proposed to be transmitted by a prion-like mechanism. The [URE3] phenotype (Wickner, 1997) (the brackets indicate an apparent cytoplasmic mode of inheritance the capital letters that it acts dominantly) leads to a relaxation of repression of nitrogen catabolic enzymes. While the [PSI^+] (Patino et al., 1995) phenotype leads to skipping of translation termination codons. These phenotypes occur spontaneously at low frequencies (10^{-5}) but are then stably inherited although they can revert to the normal phenotype at low frequencies. The two phenotypes are related to two chromosomally encoded proteins, *Ure2*, which inhibits the transcription factor Gln3 and *Sup35* which is a transcription termination factor. Mutations in these genes can lead to phenotypes resembling [PSI^+] and [URE3], but such mutations result in Mendelian segregation of the phenotype in a recessive manner and the phenotype does not revert to wild type. Although Sup35p and Ure2p proteins are unrelated to each other, the N-terminal regions of both proteins are asparagine and glutamine rich. The rate of occurrence of the [PSI^+] phenotype can be increased by transient over-expression of *Sup35* or just the N-terminal region of the protein, the phenotype is then stably inherited,

even after the cessation of over-expression, conversely mutations in the N-terminal region of sup35 prevent the [PSI$^+$] phenotype from occurring.

Analysis of cell extracts from [PSI$^+$] and [PSI$^-$] cells (cells without the PSI phenotype) shows that the Sup35 protein in [PSI$^+$] cells forms aggregates which pellet when centrifuged while Sup35 in [PSI$^-$] cells do not. The aggregates were also more resistant to proteinase K digestion, but this was a general resistance to digestion rather than a resistant part of the protein as seen the 27-33Kda resistant core seen in PrP. A similar phenomenon is seen in [URE3] cells where the Ure2 protein is more proteinase resistant and form aggregates (Wickner, 1994). [PSI$^+$] cells can be cured of the phenotype (i.e. converted to wild type) by being grown in media containing guanidium chloride, which is a protein de-naturing agent or by transient over-expression of heat-shock protein 104 (Hsp104) Although cells lacking functional Hsp104 cannot become [PSI$^+$] in the first instance implying that a minimum level of Hsp104 is required to create the [PSI] phenotype but at higher levels it abolishes the [PSI$^+$] phenotype (Chernoff et al., 1995). The yeast prion phenotypes therefore appear to be related to a change in the structure of the proteins involved.

Attempts to prove the prion theory with mammalian PrP have not succeeded as recombinant PrP has not been shown to be infectious. The high propensity of the N-terminal and middle regions alone of Sup35 (NMSup35) (AA1-253) to form aggregates *in vitro* was used in experiments where recombinant NMSup35 protein was introduced to a cell through liposome fusion with a yeast cell spheroblast (Sparrer et al., 2000). A high proportion (1.5%) of the fused cells adopted the [PSI$^+$] phenotype (compared with the 10^{-5} background rate). This may indicate that the NMSup35 which forms fibrils *in vitro* catalyses fibril formation *in vivo* of the endogenous Sup35 proteins.

The existence of prions in fungi appears to be a form of reversible post-translational modification that allows the cell to adopt two phenotypes, although it has not been determined what advantage the [URE3] and [PSI] phenotypes confer on cells. This contrasts with mammalian PrP where aggregation of PrP results in cell death and precipitation of protein occurs extra-cellularly as well as intra-cellularly. Aggregation

of PrP does not appear to be reversible although this may be due to the fact that neurones are non-mitotic cells therefore any PrP^{Sc} remains in the cell and cannot be diluted by cell division even if the process of aggregation has ceased. The existence of apparent protein-protein interaction in yeast by a seeded nucleation model provides evidence that prions can occur but does not necessarily imply that TSEs are caused by prions as there is no functional or sequence similarities between PrP and yeast prions and infectious protease resistant PrP has not yet been produced *in vitro*. Therefore, more work is needed on mammalian PrP and yeast prions before stronger parallels with can be drawn between the two.

1.4 PrP Expression and TSEs

1.4.1 PrP expression patterns

The pattern of PrP expression is similar in all species examined. High levels of expression are found in the brain and CNS but PrP is also expressed a wide variety of other tissues. In adult mice the *Prnp* gene is primarily expressed in neuronal cells in the CNS and PNS and in astrocytes but it is also expressed in heart and lung and to a lesser extent in the spleen and liver (Oesch et al., 1985) (Caughey and et al., 1988). It is also expressed in resting T-lymphocytes where it may play a role in T-lymphocyte proliferation which is reduced in null mice (Mabbott et al., 1997). In sheep PrP has a similar expression pattern being highly expressed in neuronal tissue and also expressed in spleen, lymph node, lung, heart, skeletal muscle and intestine (Horiuchi et al., 1995,). In cattle PrP is expressed at high levels in brain, moderate levels in adrenal gland and skeletal muscle and at low levels in kidney and spleen, no expression was found in liver (Horiuchi et al., 1997). In rats (Saeki et al., 1996) PrP is expressed at high levels in brain and placenta and at moderate levels in lung, heart and testis, at low levels in spleen and kidney and no expression is detected in liver.

Levels of expression vary according to tissue and developmental stage. In embryonic mice (Manson et al., 1992) expression is not detected in 9.5 day embryos, but by day 13.5 it is detected in the CNS and other neuronal tissue and in the enamel forming cells of the tooth-buds. By day 16.5 the tooth-bud has the highest levels of expression and PrP is also expressed in the kidney. It is also expressed in extra-embryonic tissues from very early in development. At birth, levels of expression in all tissues increase until day 20 when adult levels of expression are achieved (Lazarini et al., 1991). Another study has however, found that the levels of PrP expression levels were increasing until six weeks of age (Miele, 1999).

1.4.2 PrP expression levels and TSEs

Experiments using transgenic mouse models have shown that the level of PrP expression play a role in disease. Transgenic 129/Ola mice hemizygous for functional *Prnp* (one functional copy of *Prnp*) infected with ME7 scrapie have an average incubation period of 293 days compared with 147 days in wild-type mice (Manson et al., 1994) (Manson, 1996) this effect was also seen with other scrapie strains (Table 1-4). (Sakaguchi et al., 1995) (Bueler et al., 1994) and in other hemizygous lines (Nagasaki and Zurich I).

Table 1-4 Effect of *Prnp* gene dosage on TSE incubation periods. Transgenic mice over-expressing or under-expressing PrP were infected intra-cerebrally with brain homogenate from various TSE strains. Edinburgh, Zurich I and Nagasaki lines are hemizygous mice carrying one functional copy of *Prnp* on the respective null background.

Mouse line	PrP levels (wild-type=1)	TSE Strain	Wild-type incubation period (mean +/- s.e.m)	Transgenic incubation period
Edinburgh	0.5	ME7	147+/-2	293+/-4
Edinburgh	0.5	301C	154+/-1	230+/-2
Edinburgh	0.5	301V	227+/-3	320+/-3
Zurich 1	0.5	RML	158+/-11	290+/-33
Nagasaki	0.5	CJD	138+/-12	259+/-27
Tg93L	2-4	Chandler	125+/-3	97+/-3
Tg94	8	Chandler	137+/2	76+/-2
Tg117	8	Chandler	130+/-5	79+/-2

This inverse relationship between expression levels and incubation period has also been

observed in transgenic mice which over-expresses PrP. I/lnJ transgenic mice which over-express PrP protein 2-4 fold have reduced incubation periods of 87 days when infected with the RML scrapie isolate compared with 125 days in wild-type animals. Another line which over-expresses PrP protein 8-fold has an even shorter incubation period of 76 days (Westaway et al., 1991). Any variation in PrP expression levels in humans may therefore affect incubation periods in human TSEs which could affect the number of individuals who develop disease if the incubation period is extended beyond the average human life-span.

1.4.3 PrP expression patterns and disease

While the levels of expression of PrP affect incubation periods in transgenic mice, the pattern of PrP expression in different tissues also affects susceptibility to TSE infection and the course of TSEs. Although in TSEs the main site of pathological damage and infectivity is the CNS, other organs can harbour infectivity. As most natural routes of TSE infection involve peripheral organs, the mechanisms by which infectivity enters the periphery, replicates and spreads through the body are important aspects of disease.

In TSE infected mice, infectivity can be detected in spleen implying that the infectious agent can replicate there as well as the CNS. In BSE infected cattle many different organs including the PNS and lymphoreticular system (LRS) cells contain infectivity, while in vCJD PrP^{Sc} has been found in the appendix and tonsils of patients. Upon intra-peritoneal (ip) infection with the 139a strain of scrapie, infectivity (as measured by bioassay in mice) spreads from the thoracic cord to the lumbar cord and brain as infectivity is detected in the spleen and as the splanchic nerve innervates the spleen and joins the spinal cord in the thoracic region, this suggests that infectivity replicates in the spleen and is then transported to the CNS via the sympathetic nervous system (Kimberlin and Walker, 1980).

In order to determine what role, if any, PrP expression plays in the spread of infectivity different strategies have been used to ectopically express PrP in particular tissues. One approach was to graft ectoderm tissue from an embryonic day 12.5 PrP over-expressing

transgenic murine embryo into the brains of adult PrP null mice. Upon intra-cerebral (ic) inoculation, vacuolation and proteinase K resistant PrP developed in the graft but the vacuolation did not spread to the surrounding null tissue. While some PrP^{Sc} was found in the areas surrounding the graft, it did not induce neurodegeneration or astrocytosis. This would suggest that extra-cellular deposits of PrP^{Sc} are not toxic to cells and may therefore be a by-product of infection (Brandner et al., 1996). Alternatively, extra-cellular PrP^{Sc} could be toxic only to PrP expressing cells by interacting with PrP^{C} bound to the cell surface. Infected grafts also see a breakdown of the blood-brain-barrier, which could be an important mechanism in spreading the agent to organs outside the CNS possibly via the circulatory system. However, neurograft mice infected via the ip, intra-ocular (io) or sub-cutaneous (sc) routes of infection do not develop vacuolation or PrP^{Sc} in the graft implying that PrP expression is needed in the periphery to allow infectivity to travel to the CNS via these routes. In contrast, wild-type animals are susceptible to infectivity by any of the four routes but have a shorter incubation period and are more susceptible to dilute doses of TSE when injected ic (Blattler et al., 1997) (Aguzzi et al., 1997).

A transgenic models where hamster PrP is expressed in wild-type mice under the control of the rat neurone specific enolase (NSE) promoter has shown that expression of hamster PrP solely in neurons is sufficient to make mice highly susceptible to hamster adapted scrapie strain 263K injected ic or ip implying that with this strain of scrapie expression of PrP in spleen is not necessary (Race et al., 1995) (Race et al., 2000). A different model (Raeber et al., 1997) used a transgenic mouse which expressed hamster PrP under the control of the glial fibrillary acidic protein promoter which resulted in expression in astrocytes. These mice were susceptible to hamster scrapie strain 263K injected inter-cerebrally but only if on a null background. On a wild-type background no disease occurred possibly due a species barrier effect, although passage of brains of both types of animals to hamsters showed that the wild type background mice did harbour infectivity. Therefore PrP expression in neurones or astrocytes is sufficient for disease to occur using strain 263K.

The presence of infectivity in spleens has led to experiments to determine the role of the

LRS in TSEs. Experiments with SCID mice have been carried out which show that while both SCID and wild-type mice are susceptible to ic infection with ME7 strain of scrapie, SCID mice do not develop disease if infected ip unless infected with very high doses. In contrast BSE agent does not cause disease in SCID mice suggesting that different strains have different requirements (Brown et al., 1997). To determine if PrP expression in the immune system is required for the transport of infectivity, PrP null mice who have a neurograft of wild-type tissue were irradiated to destroy their bone-marrow and then had their immune systems reconstituted with wild-type bone marrow. These mice do not develop TSE pathology in the graft when inoculated ip. Although wild-type mice irradiated and reconstituted with null or wild type bone-marrow do develop disease. This implies that expression of PrP in lymphocytes is not necessary for the spread of disease to the CNS, but that expression is required in an element of the immune system not destroyed by irradiation.

Severe combined immuno-deficiency (SCID) mice and wild-type mice are susceptible to disease upon ic infection, but SCID mice do not develop infectivity in their spleens. If the immune system is reconstituted with bone marrow from wild type or PrP null mice, ic challenge leads to the development of infectivity in the spleen, implying that the immune system needs to be functional but does not need to express PrP for infectivity to replicate in the spleen. However, SCID mice crossed with PrP null mice, to produce PrP null SCID mice, and whose immune system is then reconstituted with PrP wild-type bone marrow do not develop infectivity in the spleen when infected intra-cerebrally with scrapie strain ME7 (Brown et al., 1999). This implies that while PrP expression in lymphocytes is not necessary for infectivity to be transported to the spleen PrP expression is required in the spleen to allow the agent (PrP^{Sc} or a virino) to replicate.

One group of cells in the spleen which may be involved in the replication of the agent are follicular dendritic cells (FDC). FDCs express PrP and PrP^{Sc} accumulates in FDCs of CJD infected mice (Kitamoto et al., 1991). FDCs require B-cells to mature, therefore in SCID mice FDCs cannot mature unless B-cells are provided when the immune system is reconstituted by bone-marrow graft. In order to determine if PrP expression in FDCs is required for agent replication in the spleen, PrP null and wild-type mice were

γ–irradiated to destroy lymphocytes but not the long-lived FDCs. Bone marrow grafts from PrP null or wild type mice then reconstituted the immune system and the mice were infected ip or ic with mouse-adapted scrapie strain ME7. Regardless of the genotype of the graft, null mice remained resistant to infection. In wild-type mice the PrP status of the bone marrow did not affect ip or ic infectivity or spleen titre as all reconstituted wild-type mice developed disease and had high spleen titres. Peripheral expression of PrP is therefore necessary for the replication of infection in peripheral organs with FDCs being implicated in spleen. However as FDCs are only found in the spleen, other cells must be involved in the transmission of infectivity between the CNS and the spleen (Brown et al., 1999).

To further determine the role of the LRS in TSEs, three lines of mice have been made which express PrP under the control of different tissue specific promoters on a null background (Raeber et al., 1999). The IRF transgenic line contained a transgenic construct where PrP was expressed under the control of the interferon regulatory factor 1 promoter and the Ig heavy-chain enhancer expressed PrP in B and T lymphocytes and also in the spleen, although in the spleen PrP was expressed at 1000-fold higher levels than normal in one transgenic line and 100 times in another transgenic line. The transgenic construct was also "leaky" with PrP mRNA detected in the brain and PrP protein expression detected in FDCs. The alb transgenic line expressed PrP in the liver under the control of the albumin promoter, while the lck transgenic line expressed PrP in t-lymphocytes under the control of the lck promoter. Upon ip infection with the RML scrapie isolate, the lck and alb lines did not develop disease while the IRF line did, although with a prolonged incubation period relative to wild-type mice. Bioassay of spleens from IRF mice hemizygous for the transgene array showed that spleen was infectious at 14 days after inoculation while the brain did not became infectious until 365 days after inoculation. These experiments show that expression in T-lymphocytes alone or in hepatocytes does not replicate infectivity while the findings with the IRF promoter show that PrP expression in spleen and brain are sufficient to make null animals susceptible to infection although the delay in infectivity developing in the brain may be due to the low levels of expression in nervous tissue.

The mechanism by which the agent moves from the periphery to the CNS could be a physical transport system of the injected infectious molecules which move to the CNS, or it could be a domino effect where PrP^{Sc} converts the PrP^{C} in peripheral tissue and this conversion spreads to the CNS. The fact that lack of PrP expression in the PNS can prevent the passage of infectivity may suggest the latter theory is the mechanism. It could be that the immune system is necessary in the uptake of PrP^{Sc} (possibly at Peyer's patches in the gut or the spleen) and in replication but that infectivity reaches the CNS via the PNS. Although the replication in the spleen is not obligate depending on the strain of agent and amount of infectivity used.

These experiments show that any variation in an individual's PrP expression levels in particular organs could increase or decrease the probability of infectivity replicating in the periphery and reaching the CNS where disease develops. If a regulatory region mutation is discovered which affects susceptibility to diseases analysing the expression pattern of the mutant allele could implicate certain cell types involved in PrP transport to the CNS.

1.4.4 Expression level variation in other human diseases

The aim of this project is to identify regulatory regions of the human *PRNP* gene and identify mutations or polymorphisms within them which affect susceptibility to TSEs. No polymorphisms have been identified outside of the ORF which alter expression levels of the human *PRNP* gene. Although a G-A polymorphism has been identified at the 3' end of the human intron (intron 2 in other species) (Palmer et al., 1996). There is no known association between this mutation and disease, the effect of this polymorphism on expression levels *in vitro* has not been examined.

Mutations in transcription regulatory regions of other genes have been shown to either directly cause disease or increase susceptibility to disease. Mutations in regulatory regions which cause disease have been reported in β-Thalassemia, and X-linked Charcot-Marie-Tooth disease (Kazazian, 1990) (Ionasescu, 1998). However, these mutations result in a drastic reduction in expression of the gene in question resulting in

a null phenotype which is not a model for TSEs. A recent study of the promoter of the Apolipoprotein E (APOE) gene, which affects susceptibility to Alzheimer's disease (AD), illustrates how an increase in expression levels of a protein can increase the incidence of a sporadic disease.

There are many similarities between human TSEs and Alzheimer's disease (AD): both are neurological diseases with a late age of onset characterised by deposition of amyloid plaques in the brain. These plaques are composed of β-amyloid protein in AD and of PrP protein in TSE, although not all human TSEs have plaques. The majority of cases of both diseases are sporadic, but approximately 10% of cases of both CJD and AD are familial, liked to autosomal dominant mutations in PrP in TSEs and in AD mutations in any of three different proteins (β-amyloid protein (APP), presenilin 1 and 2) (Das and Lal, 1997).

Polymorphisms in the coding region of proteins have an effect on susceptibility to both sporadic TSE (the PrP protein) and sporadic AD (Apolipoprotein E) (Rubinsztein, 1997). Apolipoprotein E, which affects lipid uptake in the liver has three alleles, ε2, ε3 and ε4. In European populations 80% of alleles are ε3 with the ε2 and ε4 alleles found at a frequency of 10% each. Individuals heterozygous or homozygous for allele ε4 are eight times more likely to develop AD than non-carriers (Corder et al., 1993).

In order to determine if different levels of expression caused by promoter sequence variation affect susceptibility to late-onset Alzheimer's disease the regions shown to affect *APOE* expression (between –1017 and +406) have been sequenced in a case-control study. Six polymorphisms were identified in the transcription regulatory region, 5 Single Nucleotide Polymorphism point mutations in the *APOE* upstream region and one point mutation in the intronic region at +113 (*APOE* like PrP has two short untranslated exons). Two of these polymorphisms (-557, -456) were very rare but the other four (-491A-T, –427T-C, –219G-T and +113G-C) were quite common with the rarer allele being found at a frequency of 5% to 47%.

Due to the relatively short distance between the promoter and coding regions

considerable linkage disequilibrium could exist between the regulatory region polymorphisms and coding region polymorphisms which are known to affect susceptibility. Such linkage could mask a true association with disease or give an erroneous association. Analysing the frequency of each allele in individuals with each of the three genotypes at the other polymorphic positions shows that the –427C allele is closely linked to the ε2 allele in the coding region while the –219G allele is tightly linked to +113G. The +113G was also linked to the ε4 allele.

To determine if these mutations have any functional significance in regulating transcription, hepatoma and astrocytoma cell lines were transiently transfected with reporter gene constructs driven by the *APOE* gene upstream region. Constructs carrying the rare T allele at position –491 had 60% of the activity of the commoner A allele. A case-control study of the frequencies of alleles A and T in American and Spanish populations showed that the A allele was more common in affected individuals than controls in the Spanish group. Analysing genotypes showed that the AA genotype was more common in cases (80%) than controls (55%) but again this was only seen in the Spanish population as the American control group had a higher frequency of the A allele. The effect of the A allele was stronger if individuals carrying one or two copies of the ε4 allele were excluded.

A similar effect was seen in the –427 polymorphism where the genotype TC was more common in AD cases than in controls. Dividing the group into ε4 carriers and non-carriers showed that the TC genotype was had no effect on carriers but was three times more common among AD non-carriers than control non-carriers. The rare C allele also gave higher expression levels in astrocytoma cells then the T allele thus agreeing with the –491 result that increased *APOE* expression increases susceptibility to AD in individuals with no ε4 alleles.

Analysis of the haplotype which considers the effect of the -491–427 polymorphisms on the same allele shows that the TT (T at –491, T at –427) allele has lower transcriptional activity then AT while AC has higher levels. Therefore the effects of the two polymorphisms on expression are additive. Analysing the effect of the –491-427

haplotype on disease show that it has no effect on individuals who have one or two copies of the ε4 allele of *APOE* but there is a correlation between expression levels *in vitro* and susceptibility of ε4 non-carriers. Spanish individuals who have no ε4 alleles are at a 4.2-fold increased risk of developing AD if their genotype is AC/AT (high expressing and normal allele) compared with AT/AT individuals but a 4.3-fold lower risk if they are TT/AT (low expressing and normal allele) thus expression level affects susceptibility. In American individuals there is a similar effect but this it much weaker with AC/AT individuals being 1.3 more at risk than AT/AT while AT/TT individuals are 1.5 times less likely to develop AD.

This study shows that levels of expression can affect susceptibility to complex diseases such as AD. Thus while in sporadic AD the ε4 allele of *APOE* is an important susceptibility factor it is not the sole determinant and levels of expression of other alleles can affect an individual's susceptibility. The 16-fold difference in risk between AT/TT and AT/AC individuals in the Spanish population is caused by an expression difference of approximately 50%, when assayed in astrocytoma cells, which shows that the increase in risk is much larger than the rise in expression would suggest although there could be much larger expression variation in other tissues.

One problem with this study is that the effect of expression levels was much larger in Spanish individuals than American individuals. This difference was due to different allele frequencies in the two control groups as the affected groups had similar allele frequencies. The geographical difference in the allele frequencies of healthy individuals is possibly due to the different ethnic make-up of the two populations resulting in there being a more significant difference in the Spanish study. This shows the problem of ethnic background in detecting alleles with weak effects. Other studies have attempted to duplicate these findings in other Alzheimer's disease patients and not found any association in Chinese or American Alzheimer's case-control study (Chen et al., 1999) (Robeck et al., 1999), although they did replicate the linkage between the promoter alleles and the protein ε2 and 4 alleles.

The APOE promoter polymorphisms analysis has a lot of parallels with this study of the

PrP promoter, both diseases are complex but with genetic influences on susceptibility. The APOE study has shown that expression levels of a particular protein affects disease susceptibility. This study also showed that the effect of known susceptibility factors (protein polymorphisms) must be accounted for to detect the smaller effects of promoter polymorphisms. It also demonstrates that two or more promoter polymorphisms can act co-operatively in affecting expression levels.

1.5 Aims

The aim of this project is to characterise regulatory regions of the human *PRNP* gene and to determine if any polymorphisms affect an individual's susceptibility to sCJD or vCJD. An individual's susceptibility to sCJD and vCJD is affected by the codon 129 polymorphism in the amino acid sequence of PrP. However, this polymorphism is of little value in predicting which individual will develop sCJD as individuals carrying the most susceptible genotype (methionine homozygotes) are only 8-fold more likely to develop the disease than individuals of the most resistant genotype (heterozygotes) which is of little predictive value in a disease that affects only 1 in a million each year. Similarly, although all vCJD patients to date have been methionine homozygotes, this genotype is shared with 40% of the population. Therefore there may be other genetic or environmental factors which affect an individual's susceptibility to vCJD or sCJD.

Genetic susceptibility factors may be genes at different loci which may interact with PrP or variation in regulatory elements which affect the level of PrP expression. Differences in expression could be caused by mutations in the PrP promoter or mutations in the 5' or 3' untranslated region, which affect mRNA stability. Analysing the promoters of sCJD and vCJD patients might detect mutations which increase their expression levels and predispose them to develop the disease. As TSEs have very long incubation periods in humans, differences in expression which increase or decrease incubation periods, possibly extending them beyond the average human life-span, would be very important in disease.

Differences in expression levels of particular genes have been shown to directly cause some human genetic diseases (Ionasescu, 1998; Kazazian, 1990) and has been shown to influence susceptibility to Alzheimer's disease. Differences in expression levels have been shown to have a large effect on TSE incubation periods of transgenic mice which over or under-express PrP. Any variation in PrP expression levels could be confined to a particular tissue and still have an affect on susceptibility to CJD, as PrP gene

expression in peripheral organs is necessary for disease to occur through peripheral infection.

To determine if variation in PrP expression levels or patterns affects susceptibility to disease, the aim of this project is to identify sequences that control the expression of the human *PRNP* gene (Chapter 3). These regions will then be used to control expression of a reporter gene in transgenic mice (Chapter 4) to determine if the regulatory regions act in controlling expression in a tissue specific manner. In humans, the regulatory regions will then be sequenced in controls and CJD patients to determine if any polymorphisms exist which pre-dispose individuals to sCJD or vCJD (Chapter 5).

2 Chapter Two: Materials and Methods

2.1 General Materials and methods

All techniques described here are adapted from Sambrook et al. 1989 (Sambrook et al., 1989) unless otherwise stated

Bacterial strain used for cloning was XL-1 Blue

Strains used in lambda propagation were XL-1 Blue and XL-1 MRA

XL-1 Blue Genotype: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F' *proAB lacI^q ZΔM15 Tn 10 (Tet^r)*]

XL-1 MRA Genotype: *Δ(mcrA) 183 . Δ(mcrCB-hsdSMR-mrr) 173 endA1 supE44 thi-1 gyrA96 relA1 lac.*

2.2 DNA manipulation techniques

2.2.1 Electrophoresis

DNA was separated by electrophoresis in a 0.8-2.5% agarose in 1X TAE buffer gel and electrophoresed at 100V in 1XTAE buffer.

2.2.2 50X TAE

242g Tris, 57.1ml Glacial acetic acid, 100ml 0.5M EDTA pH 8, H_2O to 1 litre

2.2.3 Restriction endonuclease digests

All restriction enzymes used were from Promega. A typical digest composed:

1-2 μg of plasmid DNA
2 μl of 10X reaction buffer
5-10U of the restriction enzyme or enzymes
H_2O to 20 μl
The reaction was incubated at the appropriate temperature for 1-3 hours.

2.2.4 Ligation

Ligations were carried out using the Boehringer Mannheim Rapid ligation kit.

50ng of vector DNA
3X the molar amount of insert DNA
H_2O to 8 μl
2μl of DNA dilution buffer
10 μl of reaction buffer
1 μl of T4 DNA ligase

This was incubated at room temperature for 5 minutes for a "sticky-ended" ligation or 30 minutes for a "Blunt-end" ligation reaction. The reaction could then be used directly to transform *E.coli*.

2.2.5 Filling-in DNA Overhangs with Klenow fragment.

Klenow reactions were carried out using reagents in the Erase-a-base kit (Promega)
3μl of Klenow mix (5 units of Klenow fragment in 30μl of 20mM Tris.Cl pH8.0, 100mM $MgCl_2$) from the Erase-A-base kit (Promega) was added to
27μl of DNA in TE
Incubated for 3 minutes at 37°C
3μl of dNTP mix (0.125mM of dATP, dCTP, dGTP and dTTP) was added
Incubated for 5 minutes at 37°C.
Heat inactivated at 65°C for 10 minutes.

2.2.6 TE buffer

10mM Tris.HCl, 1mM EDTA pH 8.0

2.2.7 Making competent cells

One colony of *E.coli* strain Xl-1 Blue was added to 2mls of LB, and incubated overnight at 37^{O}C. The next morning 0.5ml of the overnight culture was added to 50mls of LB and grown at 37°C until the OD_{600} reaches 0.6. The cells were centrifuged at 3500g for ten minutes. The pellet of cells was resuspended in 25mls of ice-cold 100mM $CaCl_2$ then centrifuged again at 3500g for 10 minutes and the pellet resuspended in 3 mls of ice-cold 100mM $CaCl_2$ .The cells were then ready to be transformed.

2.2.8 Luria-Betani broth (LB)

10g bacto-tryptone
5g yeast extract
10g Sodium Chloride
H_2O to 1 litre
pH to 7.0
Autoclave at 121°C for 15 minutes.

2.2.9 X-Gal/Ampicillin plates

Autoclave 500mls of Agar broth at 121°C for 15 minutes, when it has cooled, add 250µl of 0.1M IPTG , 1ml 2% X-Gal, 500µl Ampicillin 100mg/ml. Mix and pour into 80mm petri dishes. To make an LB amp plate into an X-Gal amp plate spread 30µl of IPTG and 50µl of 2% X-Gal onto the surface of the plate. Allow to dry before use

2.2.10 Transforming *E.coli* using the Heat-shock method

100µl of competent cells were added to a 10µl ligation reaction (section 2.2.4). The reaction was placed on ice for 10 minutes then placed in a 42°C water-bath for 45

seconds. The cells were then placed on ice for a further two minutes then 500µl of LB broth was added and the cells incubated at 37°C with shaking for 1 hour. 200µl of cells were spread on an agar plate containing ampicillin, allowed to dry then incubated at 37°C overnight.

2.2.11 Small scale preparation of plasmid DNA using alkaline lysis

One bacterial colony was placed in 2.5ml of LB containing 10ng/ml of ampicillin and grown overnight shaking at 37°C. Two ml of culture were placed in an Eppendorf tube and centrifuged in a micro-centrifuge for 2 minutes at 13000g at room temperature. The pellet was resuspended in 100µl of DNA prep Solution 1. 200µl of DNA prep solution 2 was added and the cells vortexed briefly then placed on ice for five minutes. 150µl of DNA prep Solution 3 was then added and the cells centrifuged for five minutes at 13000g at room temperature. The supernatant was then transferred to a fresh Eppendorf tube and 2.5 volumes of 100% ethanol and 0.1 volumes of 3M Sodium Acetate added. The mixture was then place on ice for ten minutes then centrifuged at 13000g at room temperature for 10 minutes. The pellet was washed in 70% ethanol, air-dried then resuspended in 100µl of H_2O with 1µl of 10mg/ml RNase. The DNA was then stored at -20°C.

DNA Prep Solution 1

0.9% Glucose, 0.01M EDTA pH 8

DNA prep Solution 2

1% SDS, 0.2N NaOH

DNA Prep Solution 3

3M Potassium Acetate, 11.5% (v/v) Glacial acetic acid

2.2.12 Large scale preparation of plasmid DNA

200ml of LB containing 10ng/ml of ampicillin was inoculated with 100µl of an overnight culture or a single colony and incubated overnight at 37°C in a 1 litre flask. The culture was then centrifuged at 2500g for 10 minutes at 4°C. The pellet was resuspended in 9ml of ice-cold DNA prep solution I and split into two 50ml tubes and 0.5ml of freshly prepared lysozyme solution (10mg/ml in 10mM tris.Cl pH8) added to each tube. The tubes were mixed thoroughly then 9mls of solution II was added to each tube. The tubes were placed on ice for 5 minutes then 7mls of ice-cold solution III added to each tube, the tubes were mixed then placed on ice for 10 minutes. The tubes were then centrifuged at 2500g for 15 minutes at 4°C in a Beckmann GS-6R centrifuge. The supernatant was removed and 0.6 volumes of isopropanol added and left at room temperature for 10 minutes. The tubes were then centrifuged at 4000g for 15 minutes at 4°C the supernatant discarded and the DNA pellet washed in 10ml of 70% ethanol. The pellet was air-dried for 1 hour then resuspended in 4.8ml of H_2O. 4.5g of CsCl (Boehringer Mannheim) was then added and the DNA solution vortexed briefly then 20µl of a 10mg/ml solution of Ethidium Bromide added and vortexed again. The solution was then transferred to a Beckman quick seal tube and centrifuged at 80000g for 16-18 hours at 20°C in a Beckmann TL100 centrifuge using a TL100.4 rotor.

The lower band of EtBr containing the closed circular plasmid DNA was then removed with a hypodermic needle and transferred to a 15 ml tube. An equal volume of water saturated n-butanol was added to the tube which was then vortexed. The upper butanol layer which contained most of the ethidium bromide was removed and more water-saturated butanol added to the aqueous layer mixed and vortexed, the extraction procedure was repeated until no EtBr was visible. H_2O (4mls) and 10mls of 100% Ethanol was then added to the tube and the DNA was precipitated for 1 hour on ice. The tubes were then centrifuged at 4500g for 15 minutes in a Beckmann J2-21 centrifuge, washed with 5ml of 70% ethanol then air-dried for 10 minutes. The pellet was resuspended in 500µl of H_2O which was split into 4 aliquots and stored at -20°C.

2.2.13 Phenol/Chloroform extraction ethanol precipitation.

One volume of phenol and one volume of chloroform was added to the solution and centrifuged at 13000g for 15 minutes. The upper aqueous layer was taken and one volume of chloroform added to it. The mixture was then centrifuged at 13000g for 15 minutes at room temperature. The upper aqueous layer was taken and two volumes of 100% ethanol and 0.2 volumes of 3M Ammonium acetate pH 5.2 added. Incubated on ice for 10 minutes then centrifuged at 10000g for 15 minutes at room temperature. The pellet was washed in 70% ethanol, then centrifuged at 13000g for 2 minutes at room temperature, air-dried and the pellet resuspended in 0.5 ml TE.

2.3 Lambda techniques

2.3.1 Lambda strains

The lambda strain used was lambda dash (Sambrook et al., 1989) and the human genomic lambda library prepared from DNA prepared from the blood of an anonymous human donor (P.Estibeiro personal communications).

2.3.2 Growing lambda phage

E.coli strain XL-1 blue was grown to an OD_{600} of 0.5 and 100 µl mixed with 50 µl of a phage suspension and incubated at 37°C for 30 minutes. The mixture was then added to 2.5mls of melted 0.7% top agarose (agarose in LB) and poured onto a 90mm agar plate. The agar plates had been pre-warmed at 37°C to allow the top agarose to be spread on the plate before it set. For the initial screens of the Estibeiro genomic lambda library 12 large agar plates were used with 300µl of bacteria 0.5 µl of the library and 6.5 mls of top agarose for each plate this equals 3 genomic equivalents. When the agarose has set the plates were incubated overnight at 37°C

2.3.3 Lambda DNA preperation: Plate method (Promega)

Four to five 90mm agar plates were prepared as described in section 2.3.2 with top agarose containing enough lambda suspension sufficient to lyse all bacteria after over-night incubation. 3mls of SM buffer was added to each plate and the top agarose scraped of the agar and transferred to a COREX tube. The agarose was broken up with a spatula and the then left to stand at room temperature for 30 minutes to allow the phage to diffuse into the buffer. The tube was then centrifuged at 10000g for 10minutes at 4°C in a JA-21 rotor. The supernatant was taken and 0.3% volumes of chloroform added to kill any remaining bacteria. The reagents and protocol (from step IV) in the Promega wizard lambda prep DNA purification system were then used as follows 40µl of nuclease mixture added to 10mls of lambda lysate and incubated at 37^{0}C for 15 minutes 4mls of phage precipitate was added mixed and placed on ice for 30 minutes, the mixture was then centrifuged at 10 000g for 10 minutes. The supernatant was removed and the pellet resuspended in 500µl of phage buffer and transferred to a 1.5ml centrifuge tube and centrifuged at 10 000g for 10 seconds at room temperature to pellet any remaining insoluble matter. 1ml of purification resin was added to the tube and mixed by inverting the tube. The lysate was pipetted into a Luer-Lok 3ml syringe which was attached to a wizard minicolumn. The lysate was injected through the column and the flow-through discarded. The column was then washed by injecting 2mls of 80% isopropanol through the column. The column was then centrifuged for 2 minutes at 10 000g at room temperature to dry the column. Then 100µl of pre-heated H2O added to the column and the DNA eluted by centrifugation at 10 000g for 10 seconds at room temperature.

2.3.4 SM Buffer

2.9g NaCl, 1g $MgSO_4$ 25ml 1M Tris pH 7.5, 2.5ml 2%Gelatin, H_2O to 500ml. The solution was autoclaved at 121°C for 15 minutes.

2.3.5 TE buffer

10mM Tris.HCl 1mM EDTA pH 8.0

2.3.6 Lambda DNA Preparation:Qiagen midi method.

E.coli XL-1 Blue were grown to an optical density of 2.5 OD_{600} units in 10 mls LB supplemented with 0.01M $MgSO_4$ and 0.01% Maltose which is approximately equal to $2x10^{10}$ cells. The cells were centrifuged at 4000g for 10 minutes at room temperature then resuspended in 600µl of Sm buffer. Approximately 200µl of phage suspension was added to the cells and incubated at 37°C for 20 minutes the mixture was then added to 100mls of LB supplemented with 0.01M $MgSO_4$ and 0.01% Maltose. The cells were incubated over-night during which time the cells lysed and the culture became translucent. The Qiagen lambda midi kit protocol was then followed according to the manufacturer's instructions. The DNA was precipitated by centrifugation at 15,000g at room temperature.

2.3.7 Lambda DNA prep. Phage buffer

150mM NaCl, 40mM Tris-HCl pH7.4, 10mM $MgSO_4$

2.3.8 Lambda DNA prep. Phage precipitant

3% polyethylene glycol (PEG-8000) 3.3M NaCl

2.3.9 Lambda DNA prep. Nuclease mixture

0.25mg/ml RNase A, 0.25mg/ml DNase I, 150mM NaCl, 50% Glycerol

2.3.10 Lambda library Filter lift

To transfer phage from plaques on an agar plate to a membrane for hybridisation, an Amersham Hybond-N nylon membrane was placed on a petri dish containing the lambda plaques that were to be screened. The filter was orientated by stabbing through the filter and agarose 2-3 times in three places with a needle dipped in ink. Two filters were lifted from each petri dish. The first filter was left on the plate for 1 minute the second filter for two minutes. The filter was placed onto 3mm paper wetted with denaturing solution (1.5M NaCl 0.5 M NaOH) for 1 minute then transferred to 3mm paper wetted with neutralising solution (1.5M NaCl, 0.5 M Tris HCl pH 7.5) for 1 minute. Washed in 2X SSC then air-dried (DNA side up) at room temperature on paper towels for 10-20 minutes. The DNA was UV cross-linked to the filter using a UV cross-linker (Hybaid) at 0.60 J/cm^2 under saran-wrap. The filters were stored under saran-wrap at room temperature until ready to be hybridise

2.3.11 20X SSC

175.3g NaCl, 88.2g Sodium Citrate, 800 mls H_2O, pH to 7.0 make to 1litre with H_2O The solution was autoclaved at 121°C for 15 minutes.

2.3.12 De-naturing solution for plaque lifts using Hybond-N filters.

1.5M NaCl, 0.5 M NaOH

2.3.13 Neutralising solution for plaque lifts using Hybond-N filters.

1.5M NaCl, 0.5M Tris-Cl pH 7.5

2.3.14 Pre-Hybridisation buffer when using labelled oligonucleotides

5X SSC, 0.5% SDS, 1mm EDTA pH 8.0, 5X Denhart's solution.

2.3.15 Hybridisation buffer when using Labelled oligonucleotides.

6X SSC, 0.5% SDS, 100μg/ml Herring sperm DNA, 5X Denhart's

2.3.16 Denhart's Solution 50X

5g Ficoll, 5g Polyvinylpyrolidone, 5g Bovine serum albumin Fraction V, H_2O to 500ml. Filter-sterilise. Do not autoclave

2.3.17 Labeling reaction.

(Adapted from Pharmacia's T7 quickprime kit protocol)

10-20ng of DNA in TE buffer was denatured at 94^{0}C for 3 minutes then spun briefly in a benchtop centrifuge

10μl of quickprime reagent mix (aqueous solution of dATP, dGTP and dTTP with random 9-mer oligonucleotides)

5μl of (3000 Ci/mmol) α-32^P dCTP= 50mCi

1μl of T7 DNA polymerase

H_2O to a total of 50μl.

The reaction was mixed and centrifuged briefly then incubated at 37^{0}C for 20 minutes.

2.3.18 Separating Incorporated label from unincorporated label.

A sephadex column was made by adding 3mls of 10% (w/v in H_2O) G-50 sephadex (Amersham Pharmacia) to a 5ml pipette tip which had some glass wool blocking the bottom. The column was placed in a retort stand and 20 opened 1.5ml Eppendorf tubes

were placed in a rack underneath the column. A few mls of TE buffer was run through the column. When the TE had passed through the 50µl of the labelling reaction was added to the column then more TE was added. Twenty 6-drop fractions were placed into a scintillation counter rack counted with LKB 1219 Rackbeta scintillation counter. Two peaks appeared when the amount of radioactivity was plotted the first was the incorporated nucleotide which migrates quickly through the column the second peak is the unincorporated nucleotides. The 3-4 tubes containing the first peak were added together, denatured at 94^{0}C for 3 minutes, then placed on ice before being added to the hybridisation cylinder(s).

2.3.19 Hybridisation washing solutions PCR probes attached to Hybond-N filters.

First wash, 1X SSC, 0.1% SDS
Second wash 0.1X SSC, 0.1% SDS

2.3.20 Hybridisation buffer for PCR probe labelling

5X SSC, 0.5% SDS, 100µg/ml Herring sperm DNA, 5X Denharts

2.3.21 Hybridisation of lambda libraries.

The UV cross-linked filters were wetted in 5X SSC then placed in a hybridisation cylinder containing 50mls of hybridisation solution (0.5% SDS, 5X SSC, 5X Denharts solution, 100 µg/ml of herring sperm DNA) and incubated on a rotisserie at 65°C for 30 minutes.

The labelled DNA probes were then added to the tubes and the cylinders were incubated at 65^{0}C overnight.

The Filters were washed in pre-heated (65^{0}C) 1X SSC, 0.1% SDS at 65^{0}C for 15 minutes, then washed in pre-heated (65^{0}C) 0.1X SSC, 0.1% SDS at 65^{0}C for 15 minutes. The damp filters were taped to 3mm paper with address labels dotted with radioactive

ink. Covered in saran wrap and exposed to XOMAT film overnight. The films were developed and any positive spots on the two duplicate filters compared. Where two spots overlapped perfectly, the radioactive ink marks on the address labels and the ink marks made with the needle when the filters were lifted from the plate were used to orientate the agar plate from which the filters were lifted. The plaque or group of plaques which aligned with the positive spot on the film was removed from the plate with a yellow Gilson tip, and placed into an Eppendorf tube which contained 1ml of Sm buffer and 1 drop of 100% chloroform was added to kill bacteria. The phage were allowed to diffuse into the buffer for 1 hour at room temperature then stored at 4°C. If necessary the phage in the buffer were then used to inoculate more bacteria at increasing dilutions in a subsequent round of screening until a single plaque corresponded to a positive spot.

2.4 PCR

2.4.1 PCR Oligonucleotides for exon 1 region

Oligonucleotides were designed using the GCG9 program Prime. Sequence positions are taken from genbank entry X83415 (Puckett et al., 1991) ".f" signifies that the oligo is on the forward strand ".r" signifies that the oligo is on the reverse strand

All oligonucleotides were manufactured by MWG (Milton Keynes UK).

HuPrPf	GGCACAGAGTGTGCGCCGGC	Position 21-40
HuPrP2f	AATTGGTCCCCGCCGCGACCT	Position 50-70
HuPrPr	TTACCTGCCTCGGTCGTGGAG	Position 210-188
HuPrP2r	TCGTGAGGAGAGGAGAAGCT	Position 197-178

Primers designed from the mammalian consensus sequences of PrP specific motifs (Westaway et al., 1994).

Mot3f	TAAAGATGATTTTTA
Mot2f	CCATTATGTAACG

2.4.2 PCR Oligonucleotides and conditions for the open reading frame

Oligonucleotides were designed using the GCG9 program Prime. Sequence positions are taken from genbank entry U29185 (Lee et al., 1998) ".f" signifies that the oligo is on the forward strand.

3'huprp CCTCCCTCATCCCACTATCAGGAAGATGAGGAA Position 26217-26185 Keorf.f ATGGCCAACCTTGGCTGCTGGATGCTGGTT Position 25503-25474

Reaction conditions	
GIBCO reagents	
Buffer (10X)	5.0µl
W-1	2.5µl
DNTPs	0.2mM each nucleotide
3'huprp primer	100pm
Keorf.f primer	100pm
DNA	approximately100ng
Taq	2.5U
H2O to	50.0µl
99°C	5 minutes
80°C	3 minutes
30 cycles of	
94°C	15 seconds
65°C	30 seconds
72°C	40 seconds
Then	
72°C	10 minutes
4°C	indefinite

2.4.3 PCR Oligonucleotides and conditions for upstream and intron probes for the lambda library

Oligonucleotides were designed using the GCG9 program Prime. Sequence positions are taken from genbank entry U29185 (Lee et al., 1998) ".f" signifies that the oligo is

on the forward strand ".r" signifies that the oligo is on the reverse strand.

Humprom1.f	CCTCCCTCTAAATCCTGTCCTTCTATC	Position 11953-11979
Humprom1.r	CATCACCTGTTCCTTCCTTCATTC	Position 12193-12170
Humprom2.f	GTAAAGAGACAAAGAGTGAGACAGGG	Position 11495-11520
Humprom2.r	GGAGAGCGAAGGATGAAGTGATAG	Position 11998-11975
Humprom3.f	TCTCACACAACACAGTGCAGTTTC	Position 13935-13958
Humprom3.r	CATACAAATCACATCCTACCCCTC	Position 14320-14297

Reaction conditions
GIBCO reagents

Buffer (10X)	5.0µl
W-1	2.5µl
DNTPs	0.2mM each nucleotide
Forward primer	100pm
Reverse primer	100pm
DNA	approximately100ng
Taq	2.5U
H2O to	50.0µl

94°C	3 minutes
35 cycles of	
94°C	45 seconds
56.6[a]/58.6[b]/55.4[c]°C	45seconds
72°C	45 seconds
Then	
72°C	10 minutes
4°C	indefinite

a=Humprom1.f and 1.r oligonucleotides, b=Humprom2.f and 2.r oligonucleotides, c=Humprom3.f and 3.r oligonucleotides

2.5 Manual Sequencing of plasmid DNA

2.5.1 Sequencing reaction

Adapted from Pharmacia T7 Sequenase 2.0 kit.

The DNA to be sequenced was denatured as follows

10 µl of sequencing denaturing solution (2mM EDTA , 2N NaOH)
5µg of plasmid DNA
H_2O to 100µl

Incubate at 37°C for 30 minutes

Sequencing denaturing solution
4µl of 500mM EDTA, 200µl of 10N NaOH 769 µl H_2O

Then 10µl of 3M NaAc and 300 µl of 100% ethanol was added to the mixture and the DNA precipitated at -70°C for 15 minutes. The mixture was then centrifuged at 13000g for 10 minutes in a benchtop centrifuge at room temperature, the pellet was washed in 70% ethanol, then resuspended in 7µl of H_2O. 2µl of sequenase buffer (200mM Tris.HCl pH 7.5, 100mM $MgCl_2$, 250mM NaCl) and 1µl of sequencing primer (1pmol/µl=1pmol) was added and the mixture was incubated at 37°C for 30 minutes to allow the primer to anneal. The reaction was then cooled on ice.

In a separate tube the sequencing reaction was set up as follows

1µl of DTT (0.1M),
2µl of sequenase dilution buffer (10mM Tris.hCl pH 7.5 5mM DTT),
0.5µl of dGTP,
2µl of H_2O,
0.2 µl of T7 polymerase (13U/µl),
0.4 µl of S^{35} dATP (10MCi/ml) was added.

This solution was added to the DNA/primer mix and incubated at room temperature for 3 minutes, then 3.5µl of this sequencing reaction was added to each of 4 tubes containing 2.5µl of dideoxy ATP, CTP,GTP or TTP. The tubes were incubated at 37°C for 5 minutes then 4µl of stop solution was then added to each tube. The samples were then stored at -20°C.

2.5.2 Sequencing denaturing solution

4µl of 500mM EDTA, 200µl of 10N NaOH 769 µl H_2O (2mM EDTA , 2N NaOH)

2.5.3 Running a sequencing gel.

Glass sequencing plates were washed with water and detergent then with industrial methylated spirits and one side of one plate was washed with dimethyldichlorosilane

(Pharmacia). The plates were separated by 0.5mm spacers and taped at the sides and bottom. The gel solution contained 6% (w/v) acrylamide diluted from the 40% acrylogel-5 mixture (BDH) 8M urea and 1X TBE. To 75mls of this solution 300µls of 10% w/v APS and 50µl of TEMED were added to catalyse polymerisation, mixed thoroughly and poured quickly between the plates the gel was allowed to set for two hours. When the gel had set it was put into a sequencing rig and 1X TBE added to the upper and lower reservoirs, the gel was run at 60 watts, 1800 volts and 100mA for 30-60 minutes to warm the gel. Sequencing reactions were heated to 80°C and 2.5µl loaded onto the gel. When the gel had finished running it was transferred to a sheet of 3mm paper then covered in Saran wrap (Dow Chemicals) and dried for 1 hour on a gel dryer at 80°C for two hours. When dried the Saran wrap was removed and the gel exposed to Biomax film (Kodak) overnight or longer if necessary.

2.5.4 Plasmid Sequencing Oligonucleotides

Designed using GCG9 program Prime. Sequence positions are taken from genbank entry U29185 (Lee et al., 1998) ".f" signifies the oligo is on the forward strand ".r" signifies the oligo is on the reverse strand.

1.f	GTGCTCATTCTAACTGGC	Position 11186-11203
3.f	TTGAGTGACAGTGATTGG	Position 11410-11427
5.f	TCAAGCAATCCTCCCATC	Position 11599-11616
10.f	ACACGGGTTTTTAAACGG	Position 12083-12100
17.f	TTCTCCTCTCCTCACGAC	Position 12742-12759
17a.f	ACGACCGAGGCAGGTAAAC	Position 12755-12773
19.f	AAAGCTCCCTTTACTGCG	Position 12997-13014
20.f	CGGAGCCACGTTAGGGAG	Position 13038-13055
22.f	GTGCTTGGGGGGATGGAG	Position 13298-13315
28b.f	TTTTAACCTCGCAATTCCACC	Position 13848-13868
28.f	TCACATCCATCCTCGCCATC	Position 13873-13892
29.f	GGCCGCGTTATTTTCTTG	Position 13985-14002
35.f	GGTCACAAAGATGGTTCTG	Position 14554-14572
36.f	TTCCTGATCCTCCTTTTG	Position 14608-14625
38.f	CCTGGAAAAGAGCCTAAATG	Position 14850-14874
47.f	TTATGTAAGGGAGCAGCC	Position 15704-15721
6.r	TCCTGTAGTCCCAGCTACTG	Position 11646-11627

6a.r	CAGAGTGAGCAACATAGTGAG	Position 11716-11696
8.r	AGTGAAATGGGCGTAAACCAT	Position 11913-11893
13.r	GAATGAGGAAGCAATTTCTG	Position 12337-12318
16.r	ACTGGCTCGGAGGACAGG	Position 12716-12699
17.r	GGGCGTTTACCTGCCTCG	Position 12777-12760
19.r	AACGGGCGGGGAGCGCGAGC	Position 13221-13202
24.r	TTCTCCAGGCGGTGGCAAC	Position 13469-13451
26.r	GAGGCTGGTTGTTCCTTG	Position 13695-13678
28.r	AATTCCTACAGCCACCTC	Position 13838-13821
30.r	CCTAAGGAAACGGGAAAC	Position 14069-14052
31.r	TTTATCCAATTCCCTGTTCC	Position 14147-14128
32.r	AATCACATCCTACCCCTC	Position 14314-14297
34.r	GATAACCCAAACATGCTC	Position 14225-14442
35.r	AACCATCTTTGTGACCCC	Position 14569-14552
36.r	CGAAATGCCCCCAATAAC	Position 14696-14679
37a.r	GCAGGTAAGTTCTCAGGAGTG	Position 14789-14769
37b.r	GTTTCTGGGTCACAGCTTC	Position 14766-14748
39.r	CACATATCACAGGCTCCA	Position 14973-14956
41.r	GCCCTTAGTGAGAGCAAT	Position 15141-15124
43.r	CAGGAGTCCTTAAAAACA	Position 15397-15380
45.r	ATCACAGCCTTTCCTGAC	Position 15575-15558
46.r	GGCTGCTCCCTTACATAATTC	Position 15721-15701
47.r	ACAGCCATGTTCAGTGTC	Position 15783-15800

Oligos designed from the pCAT3Basic plasmid (Genbank entryU57024)

49.r	CTTGATACTTACCTGCCC	Position 113-96
Rvprimer3	CTAGCAAAATAGGCTGTCCC	Position 3989-4008

2.6 Tissue culture

2.6.1 Cell lines

N_2a is a murine neuroblastoma cell line (ATCC CCL-131), SK-N-SH is a human Neuroblastoma cell line (ATCC HTB 11). Both cell-lines were grown in Dulbecco's modified Eagle medium (high glucose) with Glutmax I, 4500mg/l d-glucose and sodium pyruvate (Life technologies) supplemented with 10% foetal calf serum, 10U/ml penicillin and 10μg/ml streptomycin. The cell lines were grown at 37°C with 5% CO_2.

2.6.2 Passaging cells

The media was removed from the flask and washed with 10mls of phosphate buffered saline (PBS) without calcium or magnesium. For a 75ml flask 5mls of Trypsin (0.1X) in PBS was added and removed after one minute and the flask incubated at 37°C with 5% CO_2 for 5 minutes. The flask was then struck several times to dislodge the cells. 20mls of warmed media was then added to the cells and an aliquot was transferred to another flask and/or 6 well plates. N_2a cells were split at a dilution of 1/10. SK-N-SH cells were split at a dilution of 1 in 5.

2.6.3 Freezing and reviving cells

Cells were trypsinised in the same way as when the cells were being passaged but the cells were suspended in 5mls of MEM containing 8% DMSO (Sigma) and aliquoted into 3 cryo-tubes, wrapped in tissue paper and placed in a styrofoam box at -70°C overnight. The next day the tubes were transferred to liquid nitrogen. Frozen cells were revived by being placed in a 37°C water-bath to thaw the cells before being adding to 25mls of warm media. The next day the media was changed to remove the DMSO.

2.6.4 Calcium Phosphate Transient Transfection

Cells in six-well plates were grown to approximately 50% confluence in 2mls of media. For each well the transfection mix contained 2μg of the plasmid being tested, 1μg of control plasmid to normalise transfection efficiency (either pGL3-Control vector which encodes luciferase or pSV β-galactosidase control vector: both contain SV40 early promoter and enhancer sequences both are supplied by Promega) 12.4μl of 2M $CaCl_2$ and H_2O to 89.25 μl. To this 100μl of Hepes buffered saline (HBS) was added drop-wise while gently vortexing the tube. The tubes were left at room temperature for 30 minutes to allow precipitation to occur. The solution was added to the well and incubated at 37°C with 5% CO_2. After 24 hours the media was changed. After another 24 hours the cells

were harvested by removing the media and washing the cells in 3mls of PBS then 0.5mls of PBS was added and the cells scraped into an Eppendorf using a cell scraper. The cells were centrifuged at 10000rpm for 10 seconds in a benchtop centrifuge at room temperature and the supernatant removed. The pelleted cells can then be stored at -20°C.

A cell extract was prepared by two different methods depending on the control plasmid co-transfected with the plasmid being tested. When β-galactosidase was used, the pellet was resuspended in 100µl of 0.25M Tris.Cl pH 7.75, frozen in liquid nitrogen then melted at 37°C this process was repeated three times to lyse the cells. Alternatively when luciferase was used as a control plasmid, the cells were resuspended in 100µl 1X reporter lysis buffer (Promega) and left at room temperature for 15 minutes which lyses the cells. The lysed cells were spun at 13000rpm for five minutes in a benchtop centrifuge to pellet cellular debris. The supernatant was then taken and this cell extract was used in a CAT, luciferase or β-Galactosidase assay or stored at -20°C.

2.6.5 $CaCl_2$ 2M for transfections.

10.8g of $CaCl_2.6H_2O$ in 20ml H_2O, filter sterilize through a 22µm filter split into 1ml aliquot s store at -20°C.

2.6.6 Hepes Buffered saline (HBS)

1.6g NaCl, 0.074g KCl, 0.053g $Na_2HPO_4.12H_2O$, 0.2g Dextrose,1g HEPES, H_2O to 90ml pH to exactly 7.05 with 0.5N NaOH, H_2O to 100mls, Filter sterilize through a .22µm filter. Store in 5ml aliquots at -20°C .

2.6.7 β-Galactosidase assays

The Promega β-Galactosidase kit was used and its standard protocol followed. The 300µl reaction contained 50µl of cell extract 100µl of 1X reporter lysis buffer and 150µl of 2X reaction buffer. The mixture was vortexed then incubated at 37°C for 30 minutes 1M sodium carbonate (500µl) was added to stop the reaction. The OD_{420} was measured using a spectrophotometer (Beckmann Du650). The results of this assay allowed correction of transfection efficiencies so that a subsequent CAT assay using the PrP

promoter constructs reflects the strength of the promoter rather than any variation in transfection efficiency.

2.6.8 Luciferase assay

The Luciferase assay kit was used and the manufacturer's instructions followed (Promega). In a 96 well plate 20 µl of cell extract and 50 µl of luciferase assay reagent (Promega) were mixed together and immediately assayed for luciferase activity using a Wallac Victor 2 luminometer the plates were shaken 0.1mm at speed three for 1 second prior to reading.

2.6.9 Normalisation of luciferase and β-Galactosidase assays.

In a set of transfections 50µl of cell extract from the transfected cells with the lowest luciferase or β-galactosidase activity would be used in the CAT assay. Multiplying the activity of this extract by 50µl gives a constant number. Dividing this number by the activities of the other cell extracts gives the numbers of µl of those extracts which should be used.

Example

Cell extract	Luciferase activity (units)	Constant (units.µl)	Volume used (µl)
A	1567	78350	50
B	5623	78350	78350/5623=13.9
C	4690	78350	78350/4690=16.7

2.6.10 CAT assays

Chloramphenicol Acetyl Transferase activity was assayed by incubating up to 50µl of cell extract with 1µl of 0.025mCi/ml of C^{14} Chloramphenicol and 2µl of Acetyl Co-enzyme A 5mg/ml (Sigma) where less than 50µl of cell extract was used the volume was made up to 50µl with 0.25M Tris.Cl pH 7.75. As a positive control 0.2µl of CAT enzyme (Promega) was used. After incubation, 500µl of ethyl acetate was added and the mixture was vortexed for 30 seconds before being centrifuged for 3 minutes at 13000g then the upper organic phased was taken and dried in a speed vac for 40 minutes. The dried chloramphenicol was resuspended in 10µl of ethyl acetate before being spotted onto a TLC silica gel 60 plate (Merck). The plate was then placed in a chamber containing 50mls of 97% chloroform and 3% methanol. After 40 minutes the TLC plate was removed, air-dried for 5 minutes and exposed to film (Kodak X-OMAT) overnight or longer if necessary.

2.6.11 X-gal staining solution

0.2% X-gal from 2% stock in dimethyl formamide

2mM $MgCl_2$ (0.004g/10mls)

5mM $K_4Fe(CN)_6.H_2O$ (0.0211g/10mls)

5mM K_3 $Fe(CN)_6$ (0.01645g/10mls)

2.6.12 β-Galactosidase *in situ* cell staining

Cells in six well plates were washed twice with 2mls of PBS for five minutes on a gently shaking platform. The cells were then fixed with 2mls of a -20°C 1:1 mixture of methanol and acetone which was left on cells for 5 minutes, the cells were washed with 2mls of PBS for five minutes on a gently shaking platform. X-gal Staining solution (1

ml) was then added and the cells which were incubated overnight at 37°C

2.7 Transgenic techniques

2.7.1 Generation of transgenic mice.

Female C57B6/CBA F1 mice were super-ovulated and mated to males of the same genotype. The mice were killed and their oviducts dissected out in M2 media. Eggs were isolated in and transferred to a drop of M2 media containing hyluronidase (0.4mg/ml type IV-S Sigma H3884) for ten minutes to remove any cumulus cells attached to the eggs. The eggs were then washed in M2 media then M16 media then stored in a drop of M16 media under mineral oil in an incubator at 37°C at 5% CO_2 for one hour. Transgenic plasmid DNA were digested with restriction enzymes to remove the plasmid backbone and the fragments separated on a 1% agarose gel. The DNA was cut out and extracted from the agarose using the Qiaquick gel extraction kit (Qiagen). The DNA was eluted in PBS and diluted to a concentration of 1ng/μl before being microinjected into an egg pronucleus in M2 under mineral oil. The egg was injected until the pronucleus was seen to "swell" due to the injected DNA solution. 10-12 injected eggs were transferred into each oviduct of pseudo-pregnant (mated to vasectomised males the previous night) TO mice.

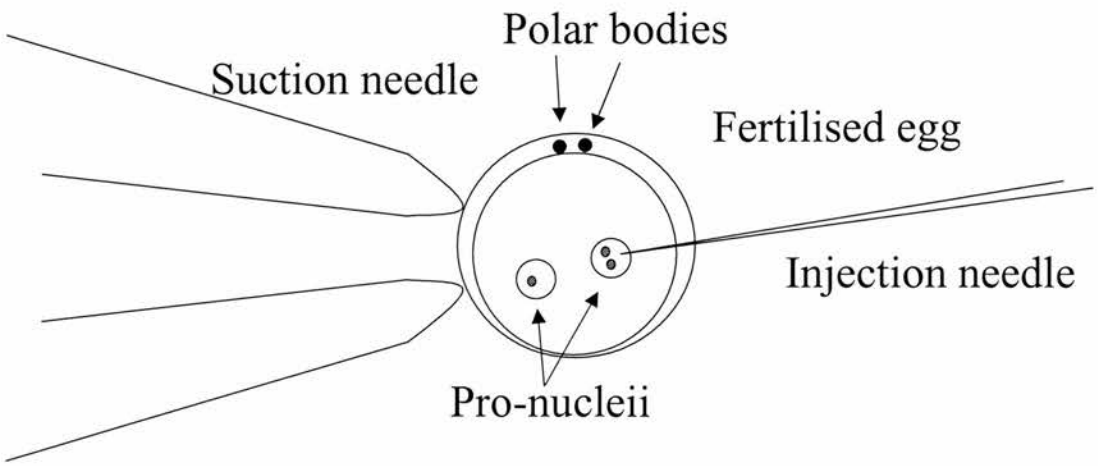


Figure 2-1 Microinjection of DNA into mouse eggs

M2 =1X Earls balanced salt solution (Gibco), 2mM $NaHCO_3$ (BDH), 0.0036g/100ml NaPyruvate (BDH), 20U/ml penicillin 20μg/ml streptomycin 0.4% BSA (Sigma A-9647), 20mM HEPES made to volume with water for injection.

M16 =1X Earls balanced salt solution (Gibco), 25mM $NaHCO_3$ (BDH), 0.0036g/100ml NaPyruvate (BDH). 20U/ml penicillin 20μg/ml streptomycin 0.4% BSA (Sigma A-9647) made to volume with water for injection

2.7.2 Analysis of embryos

Embryos were harvested in PBS the extra-embryonic membranes were removed before fixing and digested with proteinase K buffer (10mM Tris-HCl pH7.5, 1mM EDTA, 1% SDS 0.4mg/ml proteinase K (Boehringer)) overnight at 55°C. Then phenol/chloroform extracted and the DNA was ethanol precipitated and resuspended in 20μl of H_2O. The DNA was then amplified by PCR .

Conditions were

1X Gibco PCR buffer
1.5mM $MgCl_2$ 0.25mM each dNTP
0.1mg/ml BSA,
5ng oligo LacZ1 and LacZ2
0.05U/μl Gibco Taq
1μl of DNA
H_2O to 20 μl

PCR program
30 cycles of
95°C for 30 seconds
55°C for 60 seconds
72°C for 60 seconds.

Transgenic embryo diagnostic PCR oligonucleotides. GAPDH oligo positions were taken from genbank accession number NM008084. The sequence of the pLacf construct is unpublished.

LacZ1	gacaccagaccaactggtaatgg	Position 3354-3332
LacZ2	gcatcgagctgggtaataagcg	Position 2533-2554
G3PDH5	accacagtccatgccatcac	Position 566-585
G3PDH3	tccaccaccctgttgctgta	Position 1017-998

A control PCR reaction was carried out under identical conditions but using oligos G3PDH5 and G3PDH3 to amplify the murine GAPDH gene to control for variation in the quantity and quality of the DNA prep.

The embryo was fixed in 4% paraformaldehyde in PBS for an hour (E10.5 embryos) to 7 hours (E12.5 embryos). Then washed in PBS-Tween (0.1% tween-20) then stained in β-gal staining solution (.05M Na_2HPO_4 and .05M NaH_2PO_4, 2mM $MgCl_2$ 0.02% NP40, 0.01% sodium deoxycholate, 0.1% X-gal 0.24% DMF, 5mM $K_3Fe(CN)_6$, 5mM $K_4Fe(CN)_6$, 1mM EGTA pH 8 (ethylene-glycol-tetra-acetic acid)) overnight at 37°C.

2.8 Automated Sequencing

Genomic DNAs from sCJD, vCJD and control individuals were obtained from the CJD surveillance unit and were phenol extracted to remove any PrP which may still be in the DNA. The DNA was ethanol precipitated and resuspended in 20µl of H_2O. Two separate PCR reactions were carried out to amplify the upstream and intronic regulatory regions. The position of the oligonucleotides is taken from Genbank entry AF315723 (Full sequence in appendix).

The Qiagen Taq PCR core kit was used for both PCR reactions.

Upstream PCR 50µl reaction:

100pM of oligonucleotide up3.f
100pM of oligonucleotide up3.r
1X Q buffer,
1X PCR buffer which contains 15mM Mg^{2+}
5u Taq polymerase
200µM each dNTP

oligonucleotide up3.f TGCTTCCTCATTCCTGAGCC Position 1251-1270
oligonucleotide up3.r TTTACCTGCCTCGGTCGTG Position 1698-1680

Program
94°C for 3 minutes

60°C for 2 minutes
35 cycles of
72°C for 1minute
60 °C for 1minute
72 °C for 1 minute
then 72 °C for 10 minutes.

The PCR reaction product was purified using the Qiaquick gel extraction kit (Qiagen) and eluted in H_2O.

Intron PCR

Conditions 50µl reaction:

100pM of oligonucleotide Int1.f
100pM of oligonucleotide Int1.r
1X Q buffer,
1X PCR buffer Mg concentration 1.5mM,
5u Taq polymerase,
200µM each dNTP,

Oligonucleotide Int1.f TTCTCCTCTCCTCACGACCG Position 1668-1687
Oligonucleotide Int1.r TCTCTCCATCCCCCCAAGC Position 2246-2228

Program

94°C 3 minutes,
62°C 2 minutes,
35 cycles of
72°C 1minute,
62 °C 1minute,
72 °C 1 minute
then 72 °C for10 minutes.

The amplified product was purified using microcon filters

The cycle sequencing reaction was carried out using the "Big Dye" Sequencing Kit (Applied Biosystems). The sequencing oligonucelotides came from Microzone Ltd. The position of the oligonucleotides is taken from Genbank entry AF315723 (Full sequence in appendix).

Hup-F 5' CAT TCC TGA GCC TTT CAT TTT C Position 1259-1280
Hint-F 5' TCA CGA CCG AGG CAG GTA AAC Position 1679-1699

8 µl DNA,
8 µl kit PCR mix
3.2pM of sequencing oligonucleotide
H_2O to 20 µl

Program
98°C for 5 minutes,
then 25 cycles of
96°C for 30 seconds,
55°C for 15 seconds,
60 °C for 4 minutes.

To the 20µl sequencing reaction 80µl of 75% isopropanol was added, vortexed briefly and left at room temperature for 15 minutes then centrifuged at 13000g for 20 minutes at room temperature. The supernatant was removed and 250µl of 70% isopropanol added. The solution was vortexed briefly, the supernatant removed and the pellet dried by heating to 90°C and resuspended in 8µl of loading buffer. The sequencing reactions were then stored at -20°C. The sequencing gel was made using the Long ranger Singel packs (Flowgen Instruments ltd.) which were prepared according to the manufactures instructions and poured between two glass plates separated by a 0.2mm spacer. Before loading on a sequencing gel the samples were denatured by heating to 94°C for three minutes then stored on ice until loading. The gel was run at 2.99kV, 40mW at 51°C.

2.9 Statistical analysis

T-tests, ANOVA and χ^2 tests were carried out using Microsoft excel spread-sheet program. Fisher's exact tests were carried out using a web based program at www.physics.csbsju.edu/cgi-bin/uncgi/fisher.sh.

2.10 Solutions

2.10.1 EDTA 0.5M pH 8.0

Dissolve 186.1g of EDTA in 1L H2O, pH with NaOH

2.10.2 Magnesium Chloride 1M

Dissolve 203.3g of MgCl2 in 1 liter of H_2O.

2.10.3 Tris 1M pH 7.5/8

Dissolve 121.1g Tris in 800mls of H_2O. Adjust pH with HCl Make up to 1 litre with H_2O. The solution was autoclaved at 121°C for 15 minutes.

3 Chapter Three: *In vitro* characterisation of the human PrP promoter.

3.1 Introduction

3.1.1 Factors affecting susceptibility to CJD

Although the polymorphism at codon 129 is the only factor currently known to affect susceptibility to sCJD and vCJD it does not have a very strong effect on susceptibility to sCJD as most methionine homozygotes (the most susceptible genotype) do not develop sCJD. Similarly in vCJD although all the patients to date have been methionine homozygotes, this genotype is shared with 40% of the population therefore there may be other susceptibility factors which affect susceptibility to these diseases. Such factors could be environmental or genetic.

Genetic factors which affect susceptibility to CJD could be other genes whose proteins may interact with PrP, or variation in the level of PrP expression. Transgenic mice that over or under express PrP have shorter or longer TSE incubation periods than wild type mice (Manson et al., 1994) (Westaway et al., 1991), showing that expression levels can affect the course of disease. Similar expression level variation in humans, if it occurs, could affect the incubation period of sCJD or vCJD. The tissue specific pattern of PrP expression is also critical in disease, as expression in the periphery is needed to allow infectivity to replicate and spread to the CNS in mice infected experimentally by peripheral routes. In vCJD the exact mechanisms by which the BSE agent crosses the gut barrier following oral infection, replicates in the periphery, then spreads to the brain are unknown so if regulatory region polymorphisms that affect susceptibility exist, examining their affect on expression patterns could identify tissues that play a role in disease. Similarly, the tissues involved in the origin, replication and spread of the sCJD

agent are unknown. As the pathogenesis, clinical and pathological features of vCJD and sCJD are quite distinct, different tissues may play different roles in the development of the two diseases therefore polymorphisms that affect expression and alter susceptibility to one form of CJD need not necessarily affect the other form.

3.1.2 Regulation of gene expression.

In order to determine which sequences control expression of PrP in humans the sequences that regulate expression of *PRNP* need to be defined. The spatial and temporal control of gene expression is controlled by transcription factors that bind to particular DNA motifs (Tjian and Maniatis, 1994); the binding of transcription factors allows RNA polymerase II to initiate transcription. All genes have several transcription factor binding sites upstream of the start of transcription. This upstream region is often referred to as the promoter. The positioning of the transcription factor motifs in a promoter is critical as moving or reversing particular motifs can inactivate a promoter. Another class of regulatory elements are enhancers, which also contain transcription factor binding sites but the sites can be located thousands of bases upstream or downstream of the site of initiation of transcription and enhancer elements are unaffected by being reversed or moved (Fiering et al., 2000) (Blackwood and Kadonga, 1998).

The sequences upstream of all PrP genes that have been sequenced (mouse, mow, rat, human, hamster and sheep (Inoue et al., 1997) (Saeki et al., 1996) (Basler et al., 1986; Lee et al., 1998)) are similar in that they lacks a TATA box or initiator sequence and have several SP1 transcription factor binding sites near to the transcription start site. There are multiple transcription initiation sites in the ovine, murine, rat and hamster PrP genes, although not in the bovine PrP gene.

The initiation of transcription of the human gene, and hence the start of the first exon, has been identified by primer extension (Puckett et al., 1991). Examining the first exon of the ovine, murine and human PRNP genes shows that they are located within CpG islands (Lee et al., 1998), which is characteristic of genes expressed in several tissues. Normally in mammalian genomes, the cytosine residue in a cytosine guanine

dinucleotide is methylated to methyl cytosine, which is unstable and converts to a thymine. This mechanism results in the proportion of CpG dinucleotides being much lower than would be expected from the nucleotide frequencies. CpG islands are regions where the number of CpG dinucleotides is similar to what would be expected to occur by chance based on the frequency of each nucleotide. They occur where the cytosine in a cytosine-guanine dinucleotide is not methylated to the methyl-cytosine. As methylation of DNA inhibits transcription, the promoters of highly expressed genes are not methylated resulting in CpG islands near the promoters of genes. Analysis of the human *PRNP* sequence shows that in the 1kb around the first exon the GC content is 60.5% and the proportion of CpG dinucleotides is close to what would be expected from the nucleotide frequencies (91 expected 89 observed (98%)). This is a very high frequency as CpG islands are defined as regions over 200bp in length with a GC content over 50% which contains 60% or more of the expected number of CpG dinucleotides (Gardiner-Garden and Frommer, 1987). Therefore the presence of a CpG island surrounding the first exon suggests that the sequence either side of the exon may influence expression (Antequera and Bird, 1999).

Analysing the published upstream sequences of PrP genes in other species (Figure 3-1) shows that all species have two consensus binding sites for the transcription factor SP-1 (CCGCCC) upstream of the PrP gene transcription start site except for humans, who have only one SP1 site, and sheep, who have none although they do have two sites which align with the bovine SP1 sites and differ from the consensus by only one base (CCCCCC and CAGCCC). Rodents and humans have an AP-1 site (GGTGACTCA) and a reversed CCAAT box (ATTGG) in their upstream regions while ungulates have an AP-2 (CCCCGGGC) site (Kim et al., 1999) (Faisst and Meyer, 1992) .

SP-1, AP-1 and AP-2 are common transcription factors used to transcribe many genes while the CCAAT box is a motif found upstream of many genes and which several transcription factors such as CP1, CP2 and NF-1 can bind to (Chodosh et al., 1988). The human first exon (Puckett et al., 1991) is also larger than that of other species. Comparison of the first exons of published species shows that transcription of *PRNP* appears to start further upstream than in other species, this is confirmed by the presence

of two SP1 sites within the first exon (rather than immediately upstream as in other species) and a highly conserved region at the end of exon 1 in all species. All species have four conserved 10-15bp DNA motifs 150-300bp upstream of the transcription start site, although originally termed PrP specific motifs (Westaway et al., 1994), motif 4 is similar to the specific enhancer of the skeletal muscle gene (Hidaka et al., 1993) while motif 2 is identical to the binding site of E4BP4, a repressor protein related to the bZIP transcription factor family (Cowell et al., 1992).

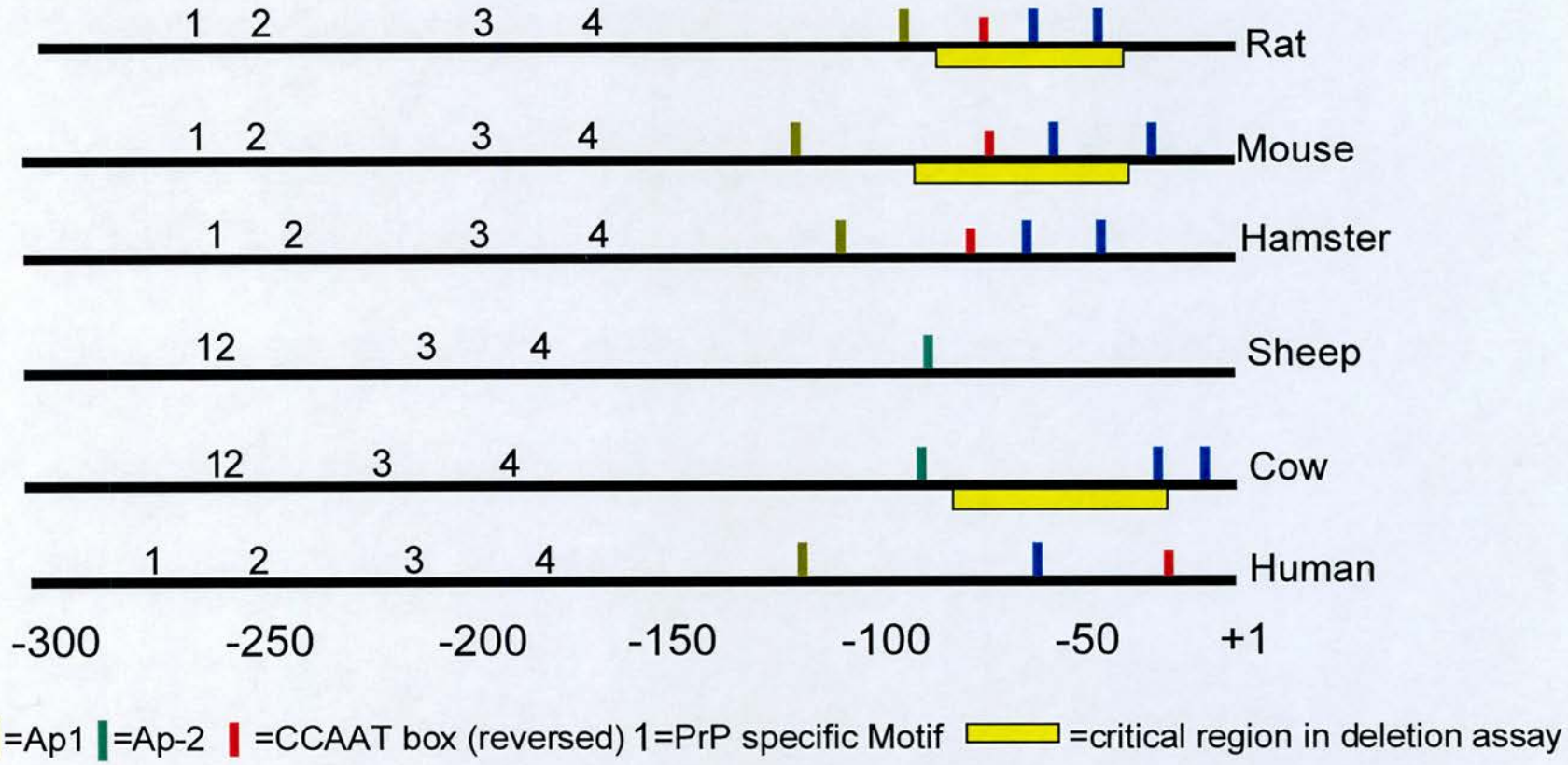
To identify sequences which control the expression of a gene, deletion analysis using a reporter gene is carried out. This is done by cloning a putative promoter region upstream of the reporter gene (such as chloramphenicol acetyl transferase (CAT), β-Galactosidase or luciferase) and transiently transfecting the construct into a cell line which normally expresses the gene whose promoter is being analysed. Once a region of DNA is shown to influence transcription, making further deletions in that region will show what effect a particular stretch of DNA has on transcription. A drop in reporter gene levels after deletion implies that the deleted region has a positive effect on transcription while a rise in reporter gene levels implies it has a negative effect.

Deletion analysis has been used to identify regulatory regions of the murine PrP gene in murine neuroblastoma cells (Baybutt and Manson, 1997), the rat PrP gene in rat pheochromocytoma and glioma cell lines (Saeki et al., 1996) and the bovine PrP gene in two bovine fibroblast cell-lines (Inoue et al., 1997). Analysis of these three promoters shows that they share a similar structure. The rat, murine and bovine PrP genes all have a promoter region between 30bp and 90bp upstream of the transcription start-site. Deletions upstream of this have no effect on transcription, except in the mouse where deletions have shown that the 300 base-pair region 360bp to 660bp upstream of the start site contains regions which have small positive and negative effect on transcription. Deleting the four PrP specific motifs has no effect on expression levels despite being ideal candidates for PrP specific transcription factor binding sites as they are well conserved in all PrP upstream regions. Perhaps they are required for transcription of PrP in particular conditions or at a developmental stage or a cell type different from those used in the reporter gene assays.

Deletions to -60 in cows and rats cause a large drop in activity while further deletions to -30 or -40 have a smaller additional effect. Further deletions which remove all upstream sequence to within 30 basepairs of the expression start site, have the same effect on transcription as deletions to –30 or –40. This may indicate that no transcription factors bind there or that this region is too small to bind sufficient transcription factors to allow transcription to occur. Each of the three promoters has two SP1 sites, with the deletion of the site furthest upstream leading to a drop in activity, but this may be due to both sites being required for transcription rather than the first site having a larger effect than the other site. Rat and mouse promoters also have an inverted CCAAT site, which also leads to a drop in activity when deleted. In deletion assays, the critical regions contain the AP1, CCAAT and two SP1 sites in the murine and rat PrP genes and an AP2 and two SP1 sites in the bovine gene (Figure 3-1).

The murine and bovine PrP genes also have regulatory regions within the first intron (this region was not examined in the rat PrP gene). In the bovine PrP gene (Inoue et al., 1997), deletions between +123 and +891 (the bovine exon 1 ends at +53 or +168 depending on alternative splicing see section 1.2.1) abolish all transcription activity even if the upstream promoter is intact. Conversely, removal of the upstream region abolishes transcription even if the intronic regulatory region is intact. Within this 768bp intronic regulatory region the only recognisable transcription factor binding site is an SP1 site. Whether this site is the crucial region requires further deletion analysis. The intronic region is not an enhancer element as cloning the reporter gene into the PrP first exon (effectively moving the intron region downstream of the reporter gene rather than upstream) abolished activity (Inoue et al., 1997).

Figure 3-1 Transcription factor binding sites upstream of the PrP first exon.



Analysis of the upstream regions of PrP genes in mouse rat cow sheep hamster and human show that the four PrP specific motifs are conserved. Within mammalian families there is also conservation of generic transcription factor binding sites such as SP1 AP1 and AP2 within rodent upstream sequences and an AP2 site in both ungulate sequences. Deletion assays of the Cow rat and mouse PrP genes show that the upstream region between approximately -90 and -30 contains regulatory elements.

In the murine PrP gene the first intron contains regions which have a positive and negative effect on transcription (Baybutt and Manson, 1997), the negative region is located within the first kilobase of the 2Kb intron while the positive region is located in the last 500bp of the intron. Unlike the bovine PrP gene, high levels of transcription can be achieved using only this second regulatory region while deletion of just the last 500bp of the intron abolishes almost all transcription even if the first promoter is left intact. Thus the murine intronic regulatory region is relatively stronger then the bovine intronic region which requires the first bovine promoter to allow transcription to occur.

One series of experiments which used a "half-genomic" construct which included the first exon and first intron but the second intron had been removed was used to drive expression of a PrP transgene which had deletions of the N-terminal (Shmerling et al., 1998). This construct gave a similar PrP expression pattern to the wild-type but lacked expression in Purkinje cells, which implies the existence of a Purkinje cell specific enhancer in the large second intron implying that there may be expression regulatory elements within the second intron (Fischer et al., 1996).

3.1.3 Plan

Based on studies of PrP regulatory regions in other species, the sequence upstream of the human first exon, the first exon itself and the intron region between the first and second exon-like untranslated exons may contain regulatory sequences. To determine if this is the case this region was cloned into the reporter gene construct pCAT3 Basic (Promega). This construct has no promoter but does contain the chloramphenicol acetyl transferase (CAT) gene that encodes a bacterial enzyme, which transfers acetyl groups from acetyl Co-A to chloramphenicol. The CAT gene was chosen as the reporter gene, rather than any of the other commonly used reporter genes such as luciferase or β-galactosidase, as the murine *Prnp* gene deletion study had been carried out in the same lab as this study using a similar CAT reporter system and therefore expertise in this technology was available.

The construct was transiently transfected into the human neuroblastoma cell line Sk-N-

Sh which expresses PrP, therefore if the region surrounding the first exon is a regulatory region of the *PRNP* gene, it would drive expression of the CAT gene when transiently transfected into the cells. The amount of CAT activity can be determined by incubating cell extracts with C^{14} labelled chloramphenicol and measuring the amount of acetylated chloramphenicol produced, as it migrates faster than un-acetylated chloramphenicol on a thin-layer chromatography plate. The level of expression is proportional to the amount of CAT activity, therefore by making a series of deletions within the proposed regulatory region the regions controlling expression can be defined.

3.2 Results

3.2.1 Cloning the putative regulatory region

To clone the putative regulatory region into the reporter construct, a human genomic lambda library (a kind gift from Dr. Peter Estibeiro, Centre for Genome Research, Edinburgh) was screened for the sequence surrounding the *PRNP* first exon. In order to screen the library a probe of human sequence from the first exon/upstream region was required. Attempts were made to make a probe used the Polymerase Chain Reaction (PCR) to amplify the first exon sequence published by Puckett and colleagues (Puckett et al., 1991)(Section 2.4.1). This published sequence was quite short (220bp) and very GC-rich (79.5%) (Puckett et al., 1991), therefore primer design was greatly constrained and the two forward primers (HuPrPf and HuPrP2f) and the two reverse primers (HuPrPr and HuPrP2r) that were designed, using less than optimal parameters, failed to amplify the region using PCR buffers and Taq polymerase from GIBCO and Qiagen and also failed to amplify using the Stratagene opti-prime PCR kit which contains 12 different PCR buffers of differing pH, $MgCl_2$ and KCl concentrations and PCR adjuncts formamide, DMSO, glycerol, BSA, $(NH_4)_2SO_4$ and Perfect Match™ reagent. The upstream motifs 2 and 3 which are highly conserved in the cattle sheep mice, rats and hamsters PrP gene upstream regions (Westaway et al., 1994) were used to design two short forward primers (Mot2f and Mot3f) in an attempt to amplify the sequence between the motifs and the first exon but these oligos also failed to amplify the region.

As the first exon could not be amplified by PCR, the open reading frame (ORF) of the *PRNP* gene was amplified by PCR as this exon has a larger published sequence (2.3Kb) with a lower GC-content (44%) than the first exon, allowing the design of optimal primers (Figure 3-2) (Puckett et al., 1991). This sequence would then be used as a probe to screen the lambda library. As the ORF is 12Kb downstream of the first exon and the average size of genomic DNA inserts in the lambda library is 15Kb it is unlikely that a single clone will contain both the ORF and the first exon, therefore once a lambda clone

containing the ORF was identified its 5' region would be used as a probe in a second round of screening to "walk" upstream to find a clone which contains the first exon.

3.2.2 Screening of the lambda library

The ORF region was amplified by PCR using oligonucleotides Keorf.f and 3'huprp (section 2.4.2) which gave a 743bp DNA fragment. This fragment was extracted from agarose using the Qiaquick gel extraction kit (Qiagen). The DNA was eluted in TE and used to make a P^{32} labelled DNA probe by the random primer method. This probe was then used to screen the human genomic lambda library.

The library was first diluted in 10-fold sequential dilutions from10^{-4} to 10^{-10} in SM buffer and plated out to determine the number of plaque forming units (pfu) per ml of the library. The 10^{-6} dilution contained 10 pfu therefore there are 10^{7} pfu per 100µl and 10^{8} pfu per ml. To screen a lambda library the total amount of insert DNA screened should be equal to three times the genome of the organism being screened. Therefore, to screen a human library 3 X $3X10^{9}$ base-pairs need to be screened. As the average DNA insert in lambda libraries is $1.5X10^{4}$ base-pairs, $6X10^{5}$ plaques need to be screened to cover the human genome, equal to 6µl of the lambda library.

In the first round screen twelve 150mm plates each containing 0.5 µl of library were screened and five plaques hybridised to the probe. Five further rounds of screening isolated three of these five initial positives; lambda clones 55, 79 and 80. DNA was prepared by lambda DNA plate prep and mapped. Restriction mapping of these lambda clones showed that they were identical to each other and the map agreed with the known restriction map of the PRNP ORF (Puckett et al., 1991), but showed that the ORF was located towards the 5' end of the clone and therefore there was little sequence 5' of it to use as a probe in a second screening (Figure 3-3) of the lambda library.

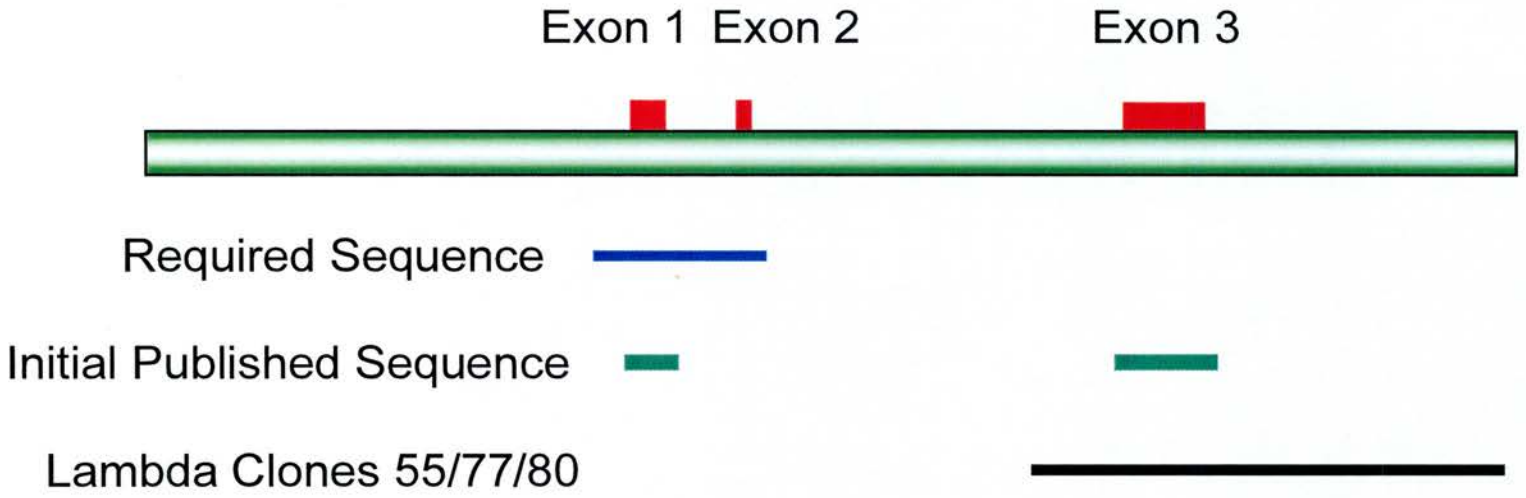


Figure 3-2 Initial attempts to clone the putative promoter region of *PRNP* In order to clone the putative regulatory region of *PRNP* (Blue line) the sequence available at the start of the study (green line) was used to amplify a probe by PCR. PCR on the exon1 region failed to amplify a product but a probe was made from the ORF in exon 3 and used to screen a lambda library and isolate lambda clones 55, 77 and 80 which were identical (Black line).

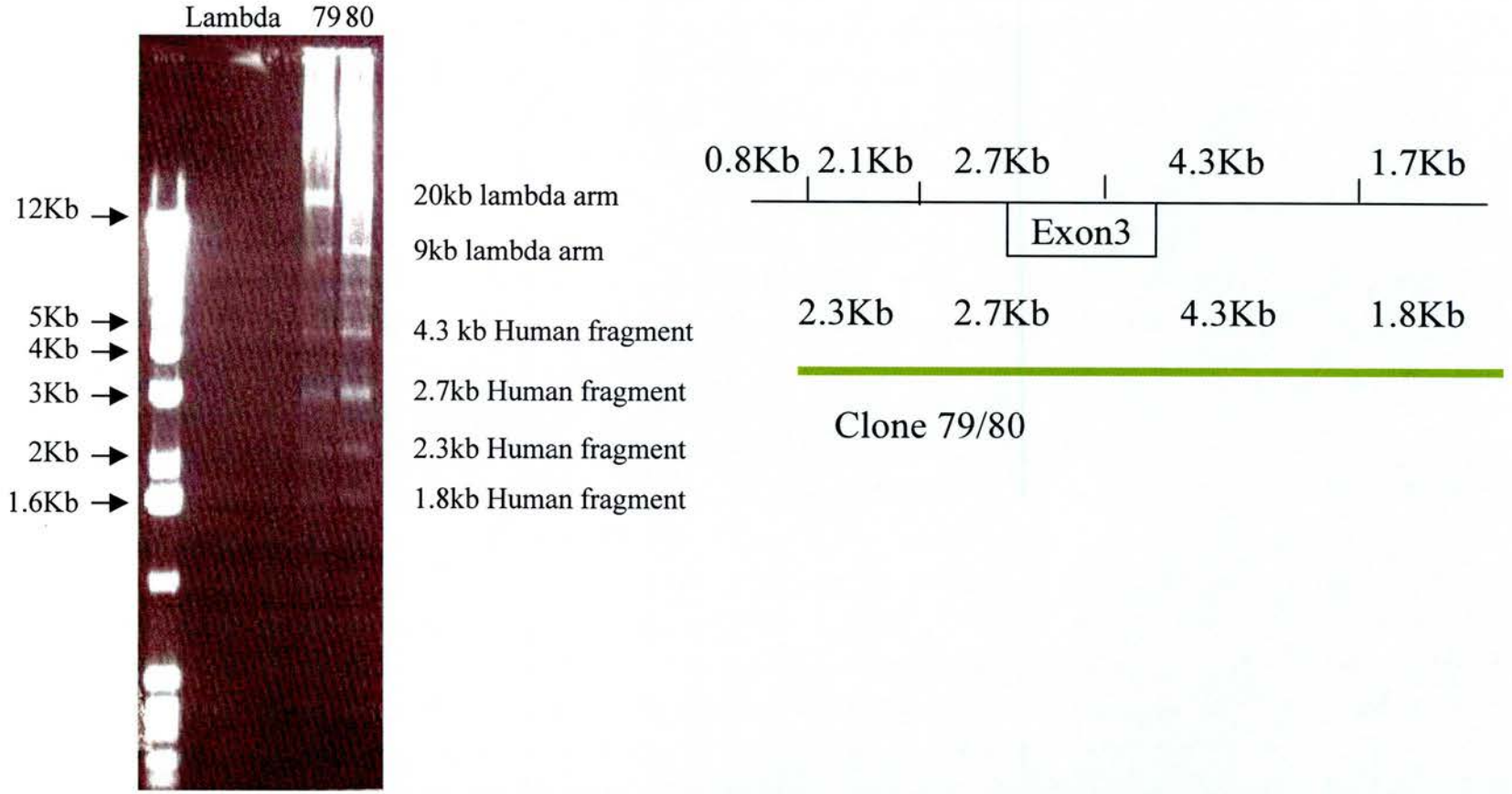


Figure 3-3 Mapping of ORF clones 79 and 80. *Xba*I digest of lambda clones 79 and 80 shows that they are identical. Comparing the restriction pattern with that of PrP gene (Lee et al.1997) shows that the lambda clones 79 and 80 only contain about 3Kb of sequence upstream of the third exon which contains the ORF of PrP.

However, at this time the sequence of the entire *PRNP* gene and 12 Kb of upstream sequence, was published (Lee et al., 1998) (Genbank Accession number U29185). Analysis of this sequence showed that there were discrepencies in the original sequence of the first exon which could have affected PCR efficiency. One of the two forward oligonucleotides had an error at the most 3' base which probably prevented annealing of the last base to the target strand and extension by the polymerase. The second forward oligonucleotide had two incorrect bases in the middle of the oligonucleotide which would reduce the efficiency of annealing to the target strand (Figure 3-4). The two reverse oligonucleotides were unaffected.

Figure 3-4 Discrepancies in the original sequence of the *PRNP* fist exon and flanking regions.

GAP program (GCG) alignment highlighting the differences between the human first exon sequences of Puckett *et al.*1991, which was used to design PCR primers, and Lee *et al.* 1998, which resequenced the same clone. First exon in **Bold**. The two forward oligonucleotides affected by the sequence changes are underlined. The sequences of the two reverse oligonucleotides were identical in the two sequences.

```
Puckett    1 GGCCGGCCGCCCGCCGCGGGGGCACAGAGTGTGCGCCGGCGCGCCGGCGA 50
             |||||||||||||||  |||||||||||||||||||||||  |||   |  |
Lee        1 GGCCGGCCGCCCGC..CGGGGGCACAGAGTGTGCGCCGGGCCGCGCGGCA 48

Puckett   51 ATTGGTCCCCGCCGCGACCTCCGCCCGCGAGCGCCGCCCGTTCCCTTCCC 100
             ||||||||||||  |||||||||||||||||||||||  ||||||||||
Lee       49 ATTGGTCCCCGCGCCGACCTCCGCCCGCGAGCGCCGCCGCTTCCCTTCCC 98

Puckett  101 CGCCCCGCGTCCCTCCCCGCCTCGGCCCCGCGCGTCGCCTCCAGTCGCTG 150
             ||||||||||||||||  |||||||||||||||||||||    ||   |
Lee       99 CGCCCCGCGTCCCTCCC..CCTCGGCCCCGCGCGTCGCCTGTCCTCCGAG 146

Puckett  151 CCAGTCGCTGACAGCCGCGGCGCCGCGAGCTTCTCCTCTCCTCACGACCG 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Lee      147 CCAGTCGCTGACAGCCGCGGCGCCGCGAGCTTCTCCTCTCCTCACGACCG 196

Puckett  201 AGGCAGGTAAACGCCCGGGG 220
             ||||||||||||||||||||
Lee      197 AGGCAGGTAAACGCCCGGGG 216
```

The publication of the full sequence of the human *PRNP* gene allowed the design of primers to amplify probes near the first exon which were not constrained by length or GC content. Three sets of PCR primers were designed using the prime program of GCG (Figure 3-5), two pairs of primers lay upstream of the first exon (Humprom1.f and

Humprom1.r creating a 241bp PCR product, Humprom2.f and Humprom2.r creating a 568bp PCR product) and one pair between the first exon and the second exon-like sequence (Humprom3.f and Humprom3.r creating a 381bp PCR product)(section 2.4.3). All three amplified at the recommended annealing temperature under standard PCR conditions (Figure 3-5). The PCR products were cloned into pGEMT and large scale plasmid DNA preparations made. These were digested with *SacII* and *NotI* to release the cloned sequence which were then radio-labelled using the random primer method and used to screen the lambda library. The first probe failed to hybridise to any plaques, the second hybridised to a very large proportion of plaques in the first round screen, further analysis of the sequence of this probe showed that it contained part of an Alu element of which there are approximately 300,000 in the human genome. This would account for the very high proportion of hybridisation. The third probe hybridised to two different clones 311 and 314 (Figure 3-6) which were isolated after three and four rounds of screening respectively.

Restriction mapping of Clone 311 (Figures 3-7 and 3-8) showed that the third probe (which is downstream of the first exon) is 14.7Kb long and located at the 3' end of 311 and the clone does not include the putative second untranslated exon. Clone 314 is 15Kb long and the first exon is located towards the 5' end of the clone. Mapping of this clone shows that it extends 1.5 Kb upstream of the first exon. As clone 314 contained the putative second exon and 1.5 Kb of sequence upstream of the first exon which contains regulatory elements in other species it was decided to sub-clone the upstream, first and second exons and the first intron from clone 314 into the pCAT3 Basic reporter plasmid.

Lambda clone 314 was digested with *EcoR*1 and *Bgl*II producing a 4.8Kb fragment, containing 1.5Kb of upstream sequence, the first exon and 3.1Kb of intron which included the cryptic second exon sequence. This fragment was sub-cloned into the *EcoR*1 and *BamH*1 sites in pBluescript (Stratagene) creating construct p4.8BS (the 4.8Kb fragment cloned into pBluescript)(Figure 3-9 and 3-11). Construct p4.8BS was then digested with *Kpn*1 and *Nhe*1 and sub-cloned into the *Kpn*1 and *Spe*1 sites of pCAT3 basic giving rise to construct p4.8CATBS (Figure 3-10 and 3-11).

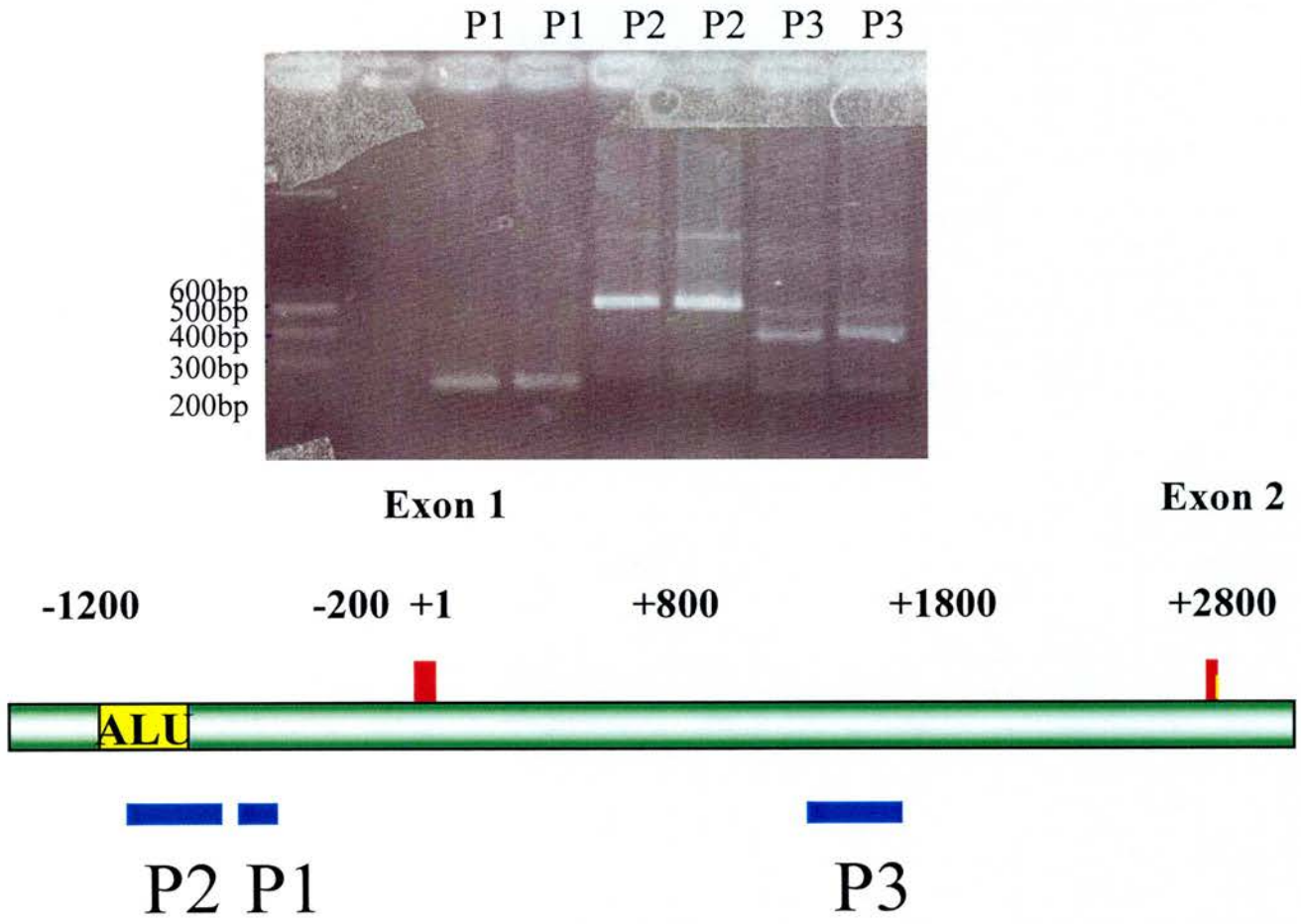


Figure 3-5 PCR probes for the putative promoter region of *PRNP* Using the genomic sequence of PRNP (Lee *et al.*1997) three probes were made by PCR and used to screen a human genomic lambda library. Human genomic DNA was amplified by PCR using primers from the upstream sequence of *PRNP* (P1 and P2) and intron of *PRNP* (P3). Expected sizes P1 241bp, P2 504bp, P3 386bp. After amplification the correct band was extracted from the agarose and cloned into pGEMT and subsequently used to screen the lambda library for a clone containing the sequence surrounding the *PRNP* first exon Probe P1 did not hybridise to any plaques while probe P2 hybridised to hundreds of plaques in the first round screen as it contained part of an Alu repetitive element of which there are 300,000 in the human genome. Probe P3 hybridised to two lambda clones 311 and 314.

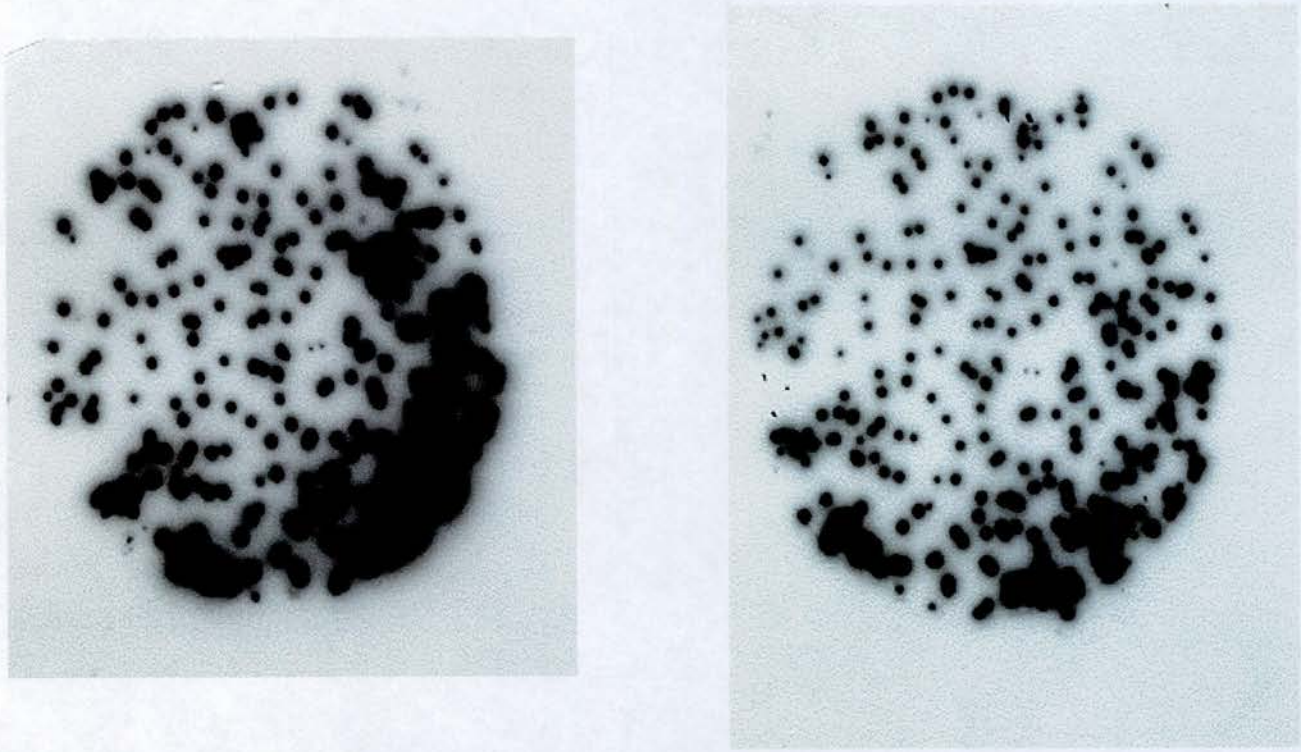


Figure 3-6 Final round screen of lambda library clone 314 using the P3 probe Hybridisation of human PRNP intron probe P3 with lambda plaques in the final round of a screen. The petri-dish was lifted in duplicate resulting in an identical pattern of plaques. The strength of hybridisation is greater in the first filter paper lifted from the petri-dish (left) than the second filter paper (right).

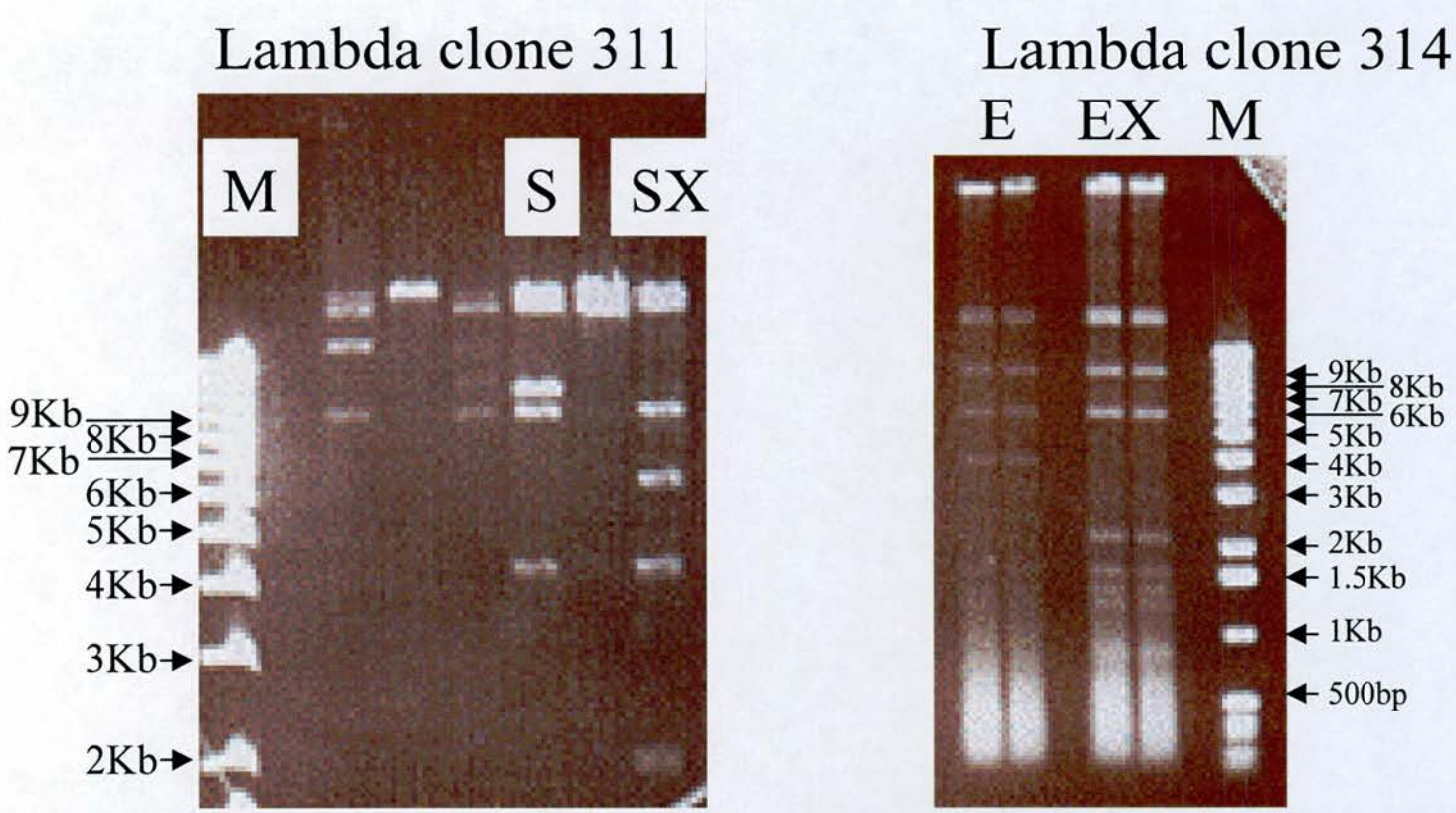


Figure 3-7 Restriction mapping of lambda clones 311 and 314 A)Digest of lambda clone 311 with *Sal*I (S) and *Sal*I-*Xba*I double digest (SX) the *Sal* I digest results in10 Kb and 4.5 Kb band but the double digest results in a 4.5 6.2 and a 2 Kb doublet. The 9Kb band is a lambda arm. This restriction pattern agrees with that of the published human sequence (Lee et al. 1997) B) 314 digest. 314 was digested with *EcoR*I and an *EcoR*I-*Xba*I. The digest pattern agrees with the published sequence The 10.2Kb band is a lambda arm. The clone contains a 5.6 1.6 1.5 and 1.4 *EcoR*I *EcoR*I fragments. *EcoR*I-*Xba* I fragments of 0.8 and 0 .9 and a 2.2 Kb *Xba*I-*Xba*I fragment.

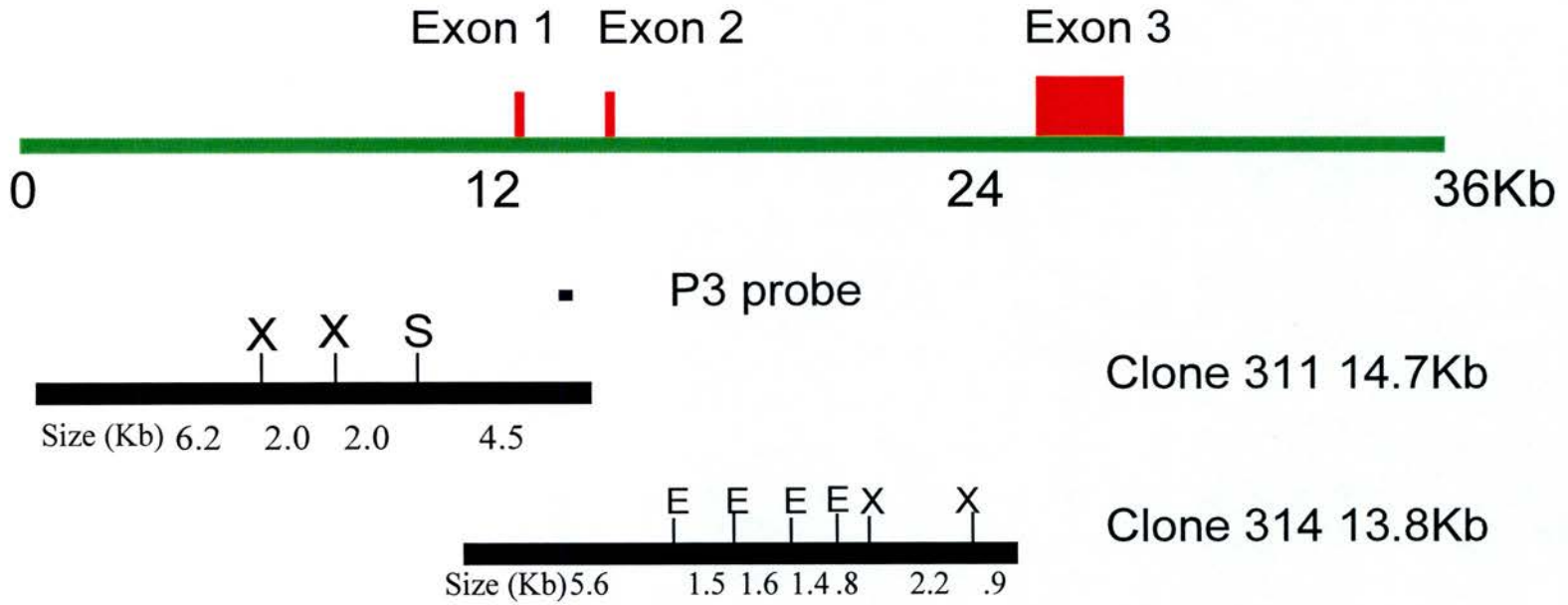
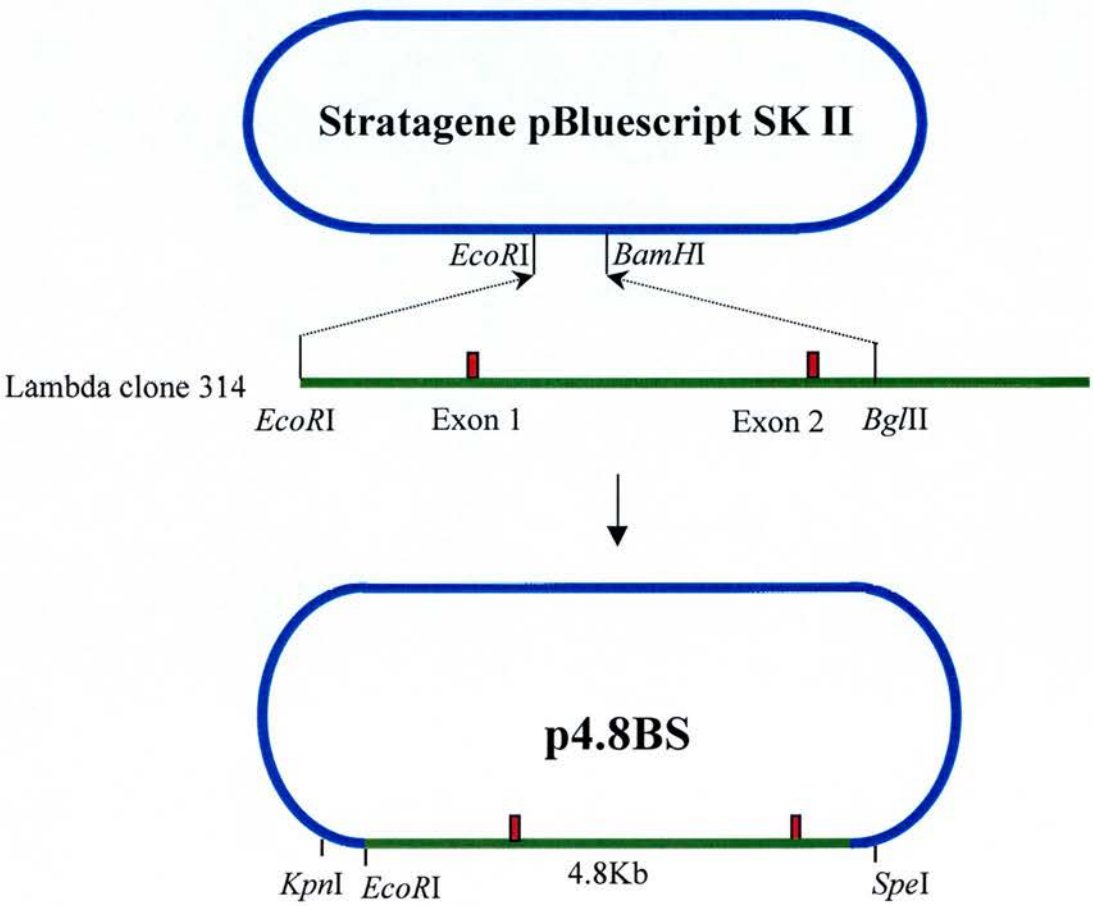


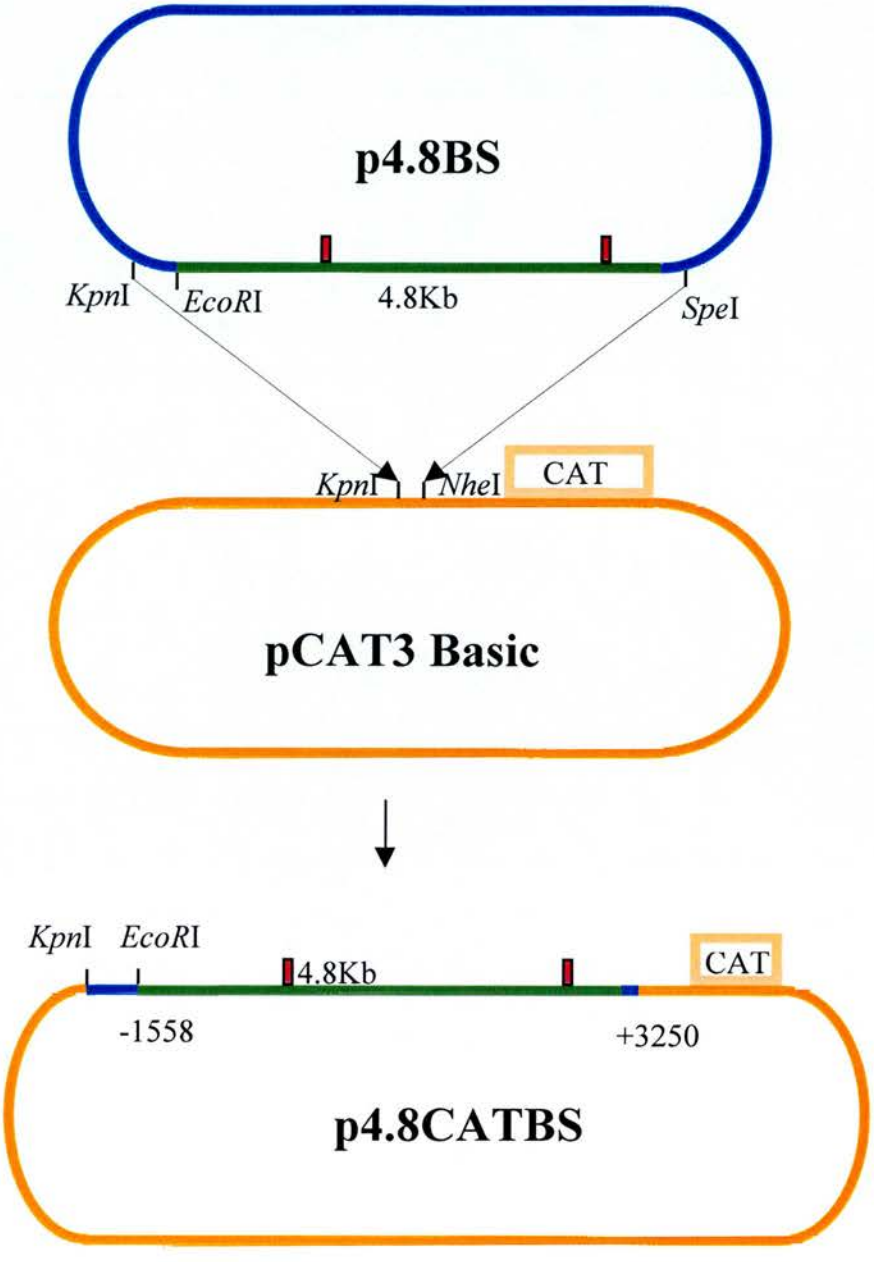
Figure 3-8 Map of lambda clones 311 and 314. Map of lambda clones 311 and 314, Clone 311 did not include the second exon whereas clone 314 did. Clone314 was therefore used to sub-clone the regulatory region into pBluescript despite clone 311 having a longer upstream sequence.E=EcoRI S=Sal I X=Xba I

Figure 3-9 Cloning putative promoter region from lambda clone 314 to pBluescript



Lambda clone 314 was digested with *EcoR*I and *Bgl*II giving a 4.8Kb fragment which contains upstream sequence the first exon and the putative second human untranslated exon. pBluescript was digested with with *EcoR*I and *BamH*1 and the 4.8Kb lambda fragment ligated into pBluescript (*Bgl*II and *BamH*I have compatible ends).

Figure 3-10 Sub-cloning the putative promoter region into pCAT3 Basic



p4.8BS was digested with *Kpn*I and *Spe*I, this 4.8Kb fragment was cloned into pCAT3Basic which had been digested with *Kpn*I and *Nhe*I, the two fragments were then ligated together the 5' overhang of *Spe*I is identical to the 5' over-hang of *Nhe*I. The resulting plasmid, p4.8CATBS, has an exonuclease 3 resistant restriction site(*Kpn*I) immediately upstream of an exonuclease 3 sensitive site (*EcoR*I) allowing unidirectional deletions to be made. This plasmid was the basis for all subsequent constructs.

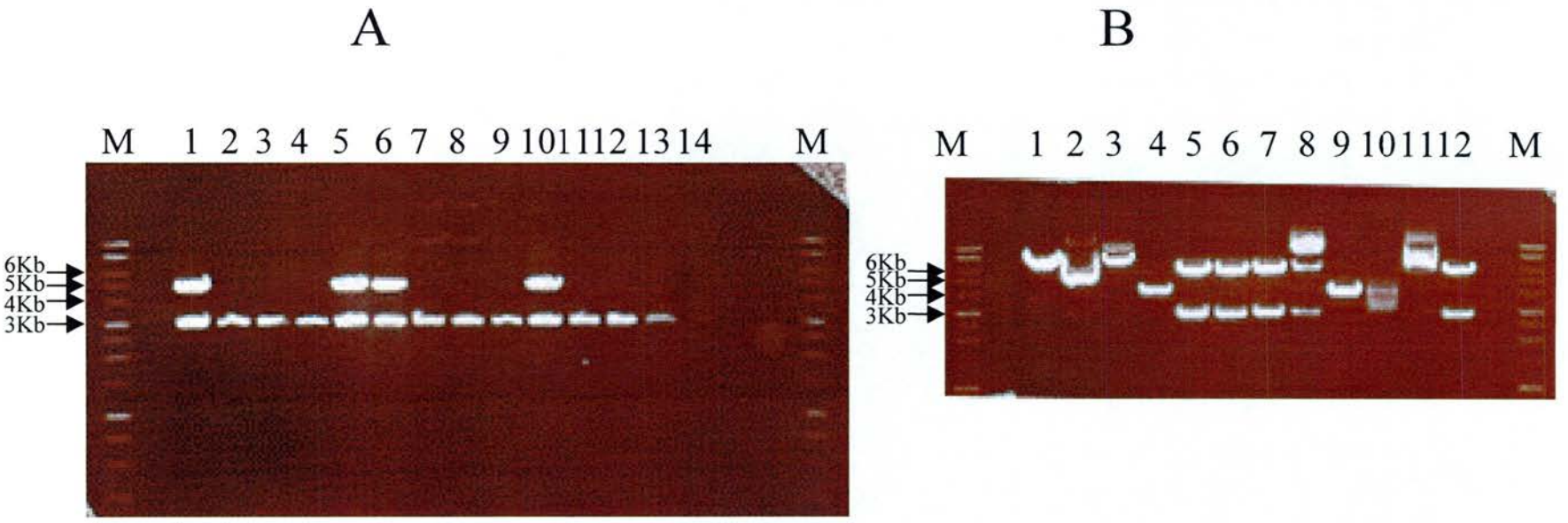


Figure 3-11 Digest analysis of cloning steps. A)Digest of putative p4.8BS transformant mini-preps. An *EcoR*I *Xba*I double digest cuts out the 4.8 Kb human fragment from the 3kb pBluescript plasmid. Lanes 1, 5, 6 and 10 contain the human sequence the other lanes contain the 3Kb pBluescript plasmid M=Size markers. B) Digest of putative p4.8CATBS transformant mini-preps with *Sal*I. p4.8CATBS cuts to give a 6.0Kb and a 2.8Kb fragment (Lanes 5,6,7,8 and 12). Lanes 4 and 9 contain the 4kb pCAT3 Basic plasmid.

The 4.8Kb fragment starts at position -1558 relative to the start of transcription and extends to position +3250 downstream of the transcription start site. This plasmid was the basis for all of the deletion constructs used in this Chapter. The pCAT3 Basic plasmid contains a chimeric intron between the poly-linker and the CAT ORF this should allow splicing to occur from the first exon or possibly the putative second exon to the CAT ORF exon.

3.2.3 Sequencing

The 4.8Kb fragment of lambda clone 314 cloned into pCAT3 basic was manually sequenced (Genbank entry AF315723 full sequence in appendix) and compared with the sequence published by Lee *et al.*, (Lee et al., 1998) (genbank number U29185) and the sequence released by the Sanger Centre as part of the human genome sequencing project (Genbank number dJ1068H6). There are seven sequence differences between the Lee sequence and the Sanger sequence. The lambda clone 314 sequence is identical to the Sanger sequence at these seven positions. This difference between the two publicly available sequences could be due to a lower sequencing error rate in the human genome project and/or the fact that Lee *et al.* used the Hela cell line S3 as the source of their clone and cell lines can accumulate mutations at a higher rate than entire organisms. There was one position in lambda clone 314 which differed from the Sanger and Lee sequences, a T to C transition at position +385 relative to the start of transcription, which creates a *BssH*II restriction site. Subsequent sequencing of a control population showed that this difference is a polymorphism (Chapter 5). Therefore the 4.8Kb of human sequence cloned from lambda clone 314 appears to be correct and has not undergone any mutations or rearrangements.

3.2.4 Identification of an upstream regulatory region

Construct p4.8CATBS, which contains sequence starting at position -1558 relative to the start of exon 1 and includes 4.8Kb of putative regulatory sequence, expressed the CAT reporter gene when transiently transfected into the human neuroblastoma cell-line

Sk-N-Sh. To determine which regions of the upstream sequence affected expression levels in this cell-line, a set of unidirectional deletions were created using the Erase-A-Base system (Promega). Plasmid -1558 (p4.8CATBS) was digested with *Kpn*I and *EcoR*I then incubated with Exonuclease III for varying periods of time. This allows the amount of DNA digested to be controlled as Exonuclease III digests DNA at a regular rate which is temperature-dependent (Figures 3-12 and 3-13). Exonuclease III digests 5' overhangs and blunt-ended DNA, but not certain 3' overhangs such as that generated by *Kpn*I. Thus a series of unidirectional deletions was produced. Incubating with S1 nuclease then removes the single stranded DNA. The resulting fragment was treated with Klenow fragment to fill in any over-hangs then ligated and transformed into *Escherichia coli*. These constructs were sequenced to determine the amount of digestion that occurred.

Analysing the rate of digestion of exonuclease III shows that it is influenced by GC content. The amount of DNA digested in 60-second intervals in the upstream region at 30°C declines with increasing GC content. As the first exon is found in a region of high GC content the rate of deletion slowed in the immediate upstream region creating several deletions in this region. Thus the –43 and –9 constructs were made by digesting for an extra minute which only removed 34 basepairs in contrast to the manufacturer's stated rate of 125bp/minute, this resulted in much closer deletions in the GC rich upstream region which is a regulatory region in other species. In contrast to the first set of Erase-A-Base deletions there was no correlation between the rate of digestion and GC content in the intron deletions although the digestion was carried out at a lower temperature (25°C) than the upstream deletions (30°C) (Figure 3-14).

Constructs -730, -364, -133, -43, -9 and +289 (the number indicates the start point of the human sequence relative to the start of exon 1) were made using the Erase-A-Base system and transiently transfected into Sk-N-Sh cells. Their level of CAT activity relative to construct -1558 is shown in Figure 3-15 and table 3-1. In each experiment construct –1558 was transfected in parallel to the other constructs being examined to allow the comparison of CAT expression levels. All constructs were co-transfected with either the pGL3-Control vector (Promega) which encodes luciferase or pSV β-

galactosidase (Promega) control vector both of which contain SV40 early promoter and enhancer sequences to control for variation in transfection efficiency. Early transfection experiments used the pSV β-galactosidase plasmid but when a luminometer capable of reading lucifcrase became available the pGL3-Control plasmid was used.

Removing the upstream sequence between -1558 and -43 had no effect on expression; statistical analysis showed no significant difference in CAT activity between the five constructs (One-way ANOVA p=0.51)(Figures 3-15 Table 3-1). This sequence includes the highly conserved region containing the four PrP specific motifs 150-300bp upstream of the transcription start site. A lack of transcriptional activity of this region has also been noted in reporter gene deletion studies in the mouse, cow and rat PrP genes (Baybutt and Manson, 1997) (Saeki et al., 1996) (Inoue et al., 1997). The functional significance of the four PrP specific motifs is therefore uncertain. Perhaps this conserved region regulates expression in different tissues or at particular stages of development. Deletion studies on mammalian PrP regulatory region have been carried out in only a limited number of cell types, (human and murine neuroblastoma cell-lines, two bovine fibroblast cell-lines and rat pheochromocytoma and glioma lines) so it is possible that using a wider variety of cell types would detect cell types where deletion of this region affects expression.

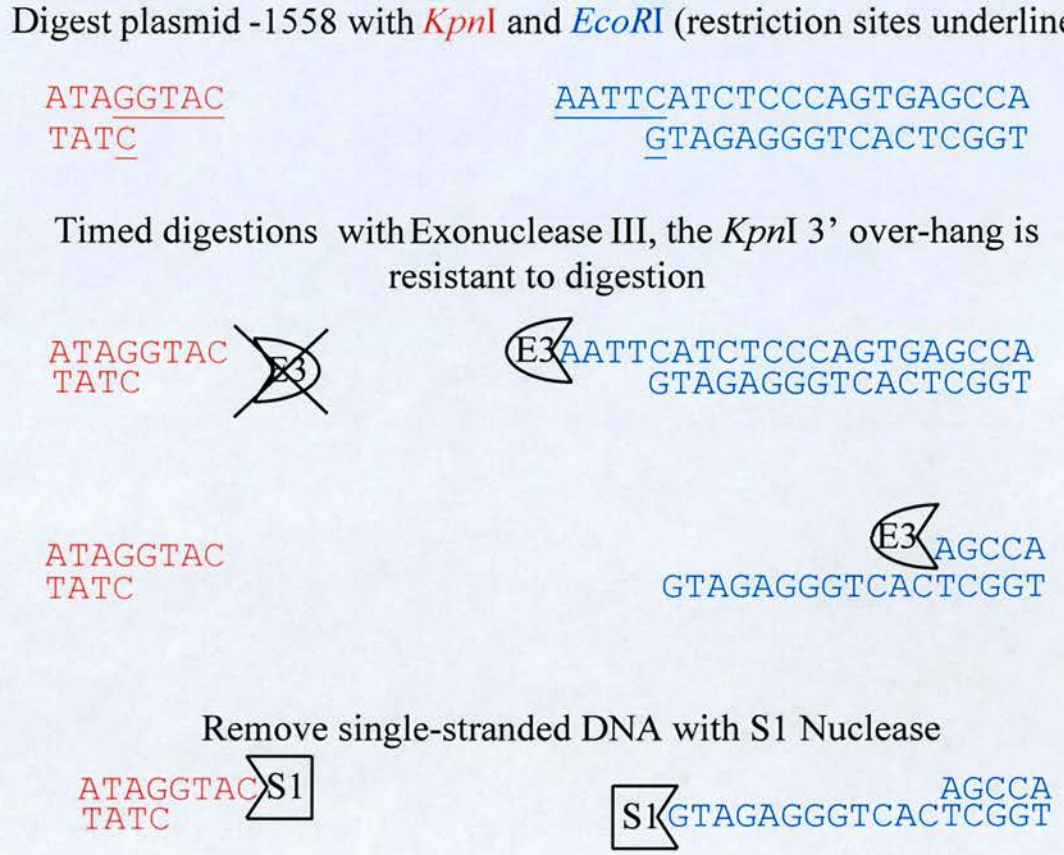
Deleting from –43 to –9 resulted in a significant drop in expression to 32% of construct –1558 ($p=10^{-5}$ t-test between –43 and –9 table 3-1). This implies that this region contains promoter activity. The location (relative to the start of transcription) of this upstream regulatory region in humans contrasts with that of rat, mouse and cow where removing sequence between –90 and –30 results in a drop in expression. However the human first exon is considerably larger than that of other species (Human=134bp Bovine = 53bp rat= 19-47bp). If the sequences of each species from the conserved motifs to the end of the first exon are aligned using the CLUSTALW program then the transcription factor binding sites, including the exonic SP1 sites in humans, and the critical regions in deletion assays align (Figure 3-16).

Therefore, the region between -43 and -9 in the human upstream region is approximately

equivalent to the -90 to -60 region in other species, implying that mammalian PrP promoters have a conserved structure. However, the alignment shows that the regulatory regions defined by deletion assay are not as well conserved as region further upstream which contains the PrP specific motifs. The only conserved transcription factor binding site in the regulatory regions is a reversed CCAAT motif in rodents and humans. This is therefore a good candidate for further studies, such as *in vitro* mutagenesis of the site and gel shift assays to determine if this sequence is bound by proteins in Sk-N-Sh cells and if the sites removal affects expression.

A further deletion which removes all upstream sequence, the first exon and part of the intron (construct +289) reduced expression to background levels ($p=0.96$ t-test between pCAT3basic and +289). This reduction may be due to the loss of transcription factor binding sites within the first exon which affect expression or that the first exon is required to initiate transcription or the first exon could be required to allow correct splicing to either the second exon or the reporter ORF exon. The most likely explanation is that there are transcription factor binding sites immediately downstream of the transcription initiation site with the most likely candidates being the SP1 sites. In the rodent and bovine analyses, removal of one of the SP1 sites resulted in a drop in expression. However, this deletion was not made in the human gene as disrupting the transcription initiation site could affect interpretation of such a construct.

Figure 3-12 Creation of deletions using the Erase-A-Base system



Erase-A-Base protocol plasmid p4.8CATBS was digested with *EcoR*I and *Kpn*I exonuclease III (E3). At one minute intervals the exonuclease digestion was ended and single-stranded DNA removed using S1 nuclease (S1). Any remaining over-hangs were filled in with Klenow fragment and the ends of each plasmid were ligated together

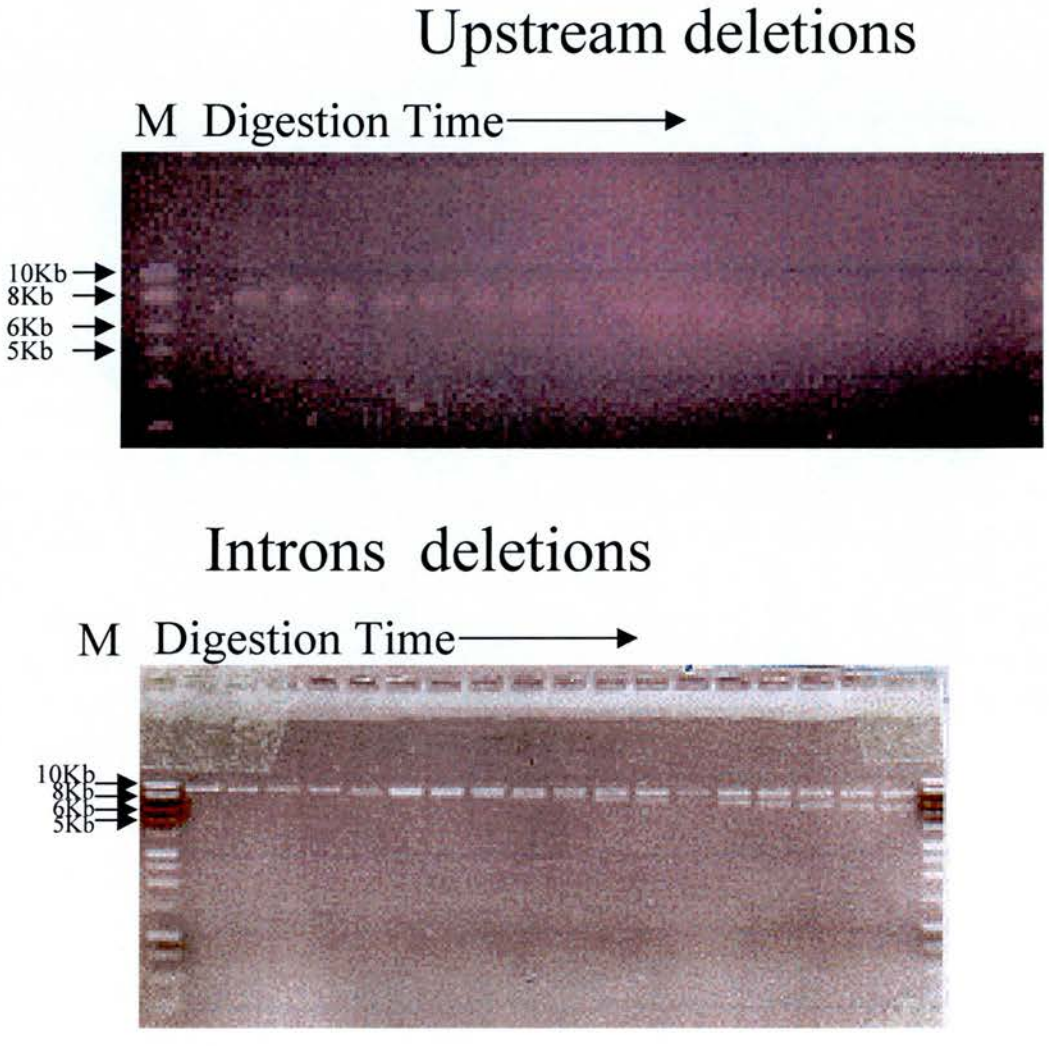


Figure 3-13 Creation of unidirectional deletions of construct -1558 Plasmid -1558 was digested with exonuclease 3 for 1 minute increments then treated with eS1 nuclease which removed the single-stranded DNA. An aliquot was then ran on a 0.8% agarose gel. On the left are the set of deletions used to make the upstream deletions, while the inton deletions are on the right. Due to an excess of ethidium bromide in the upstream deletion gel it is difficult to visualise the DNA bands .The intron deletions have two bands in the longer deletion times. The lower band may be due to partial digestion of the NsiI site by exonuclease III. But all clones recovered were mapped and had their junctions sequenced to ensure that only unidirectional fragments were used in subsequent transfection studies. M=Molecular mass marker.

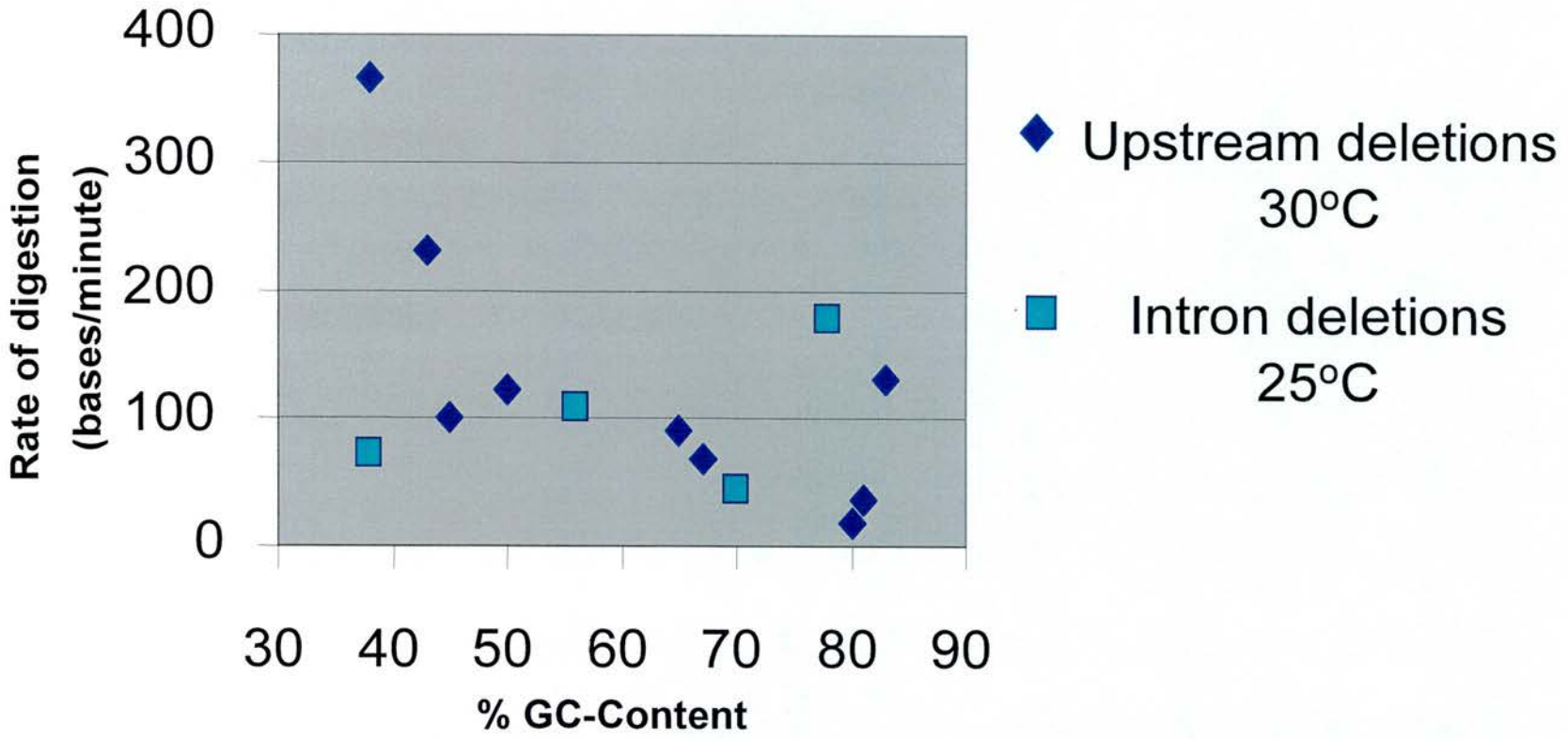
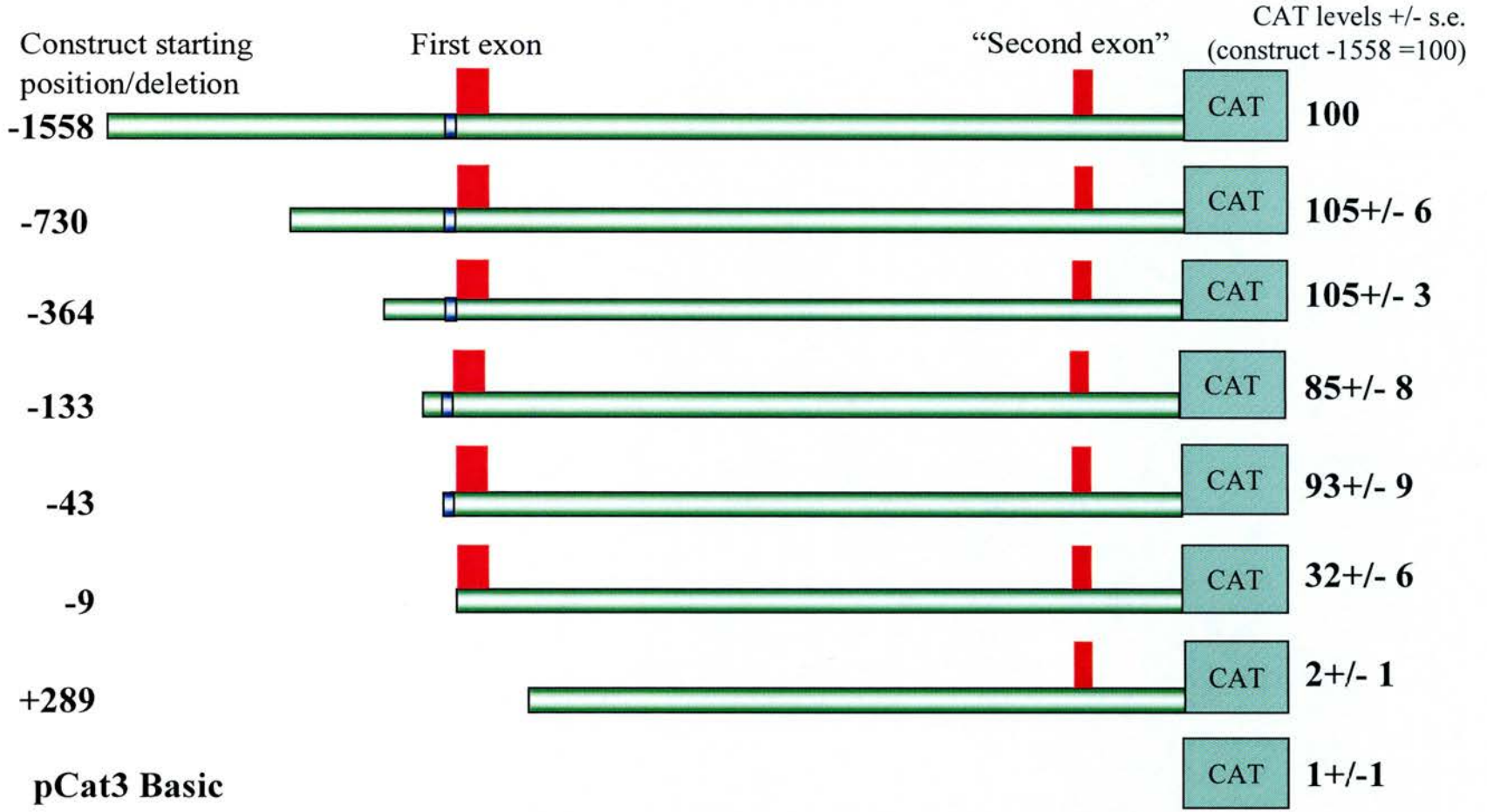


Figure 3-14 Effect of GC content on Exonuclease III Activity The Exonuclease III in the Erase-A-Base kit did not digest DNA evenly at 30°C as the rate of digestion decreased with increasing GC content (Dark Blue). This effect was not seen in a later series of erase-a-base deletions in the intron which were carried out at 25°C (light blue).

Figure 3-15 Cat expression levels of upstream deletion constructs



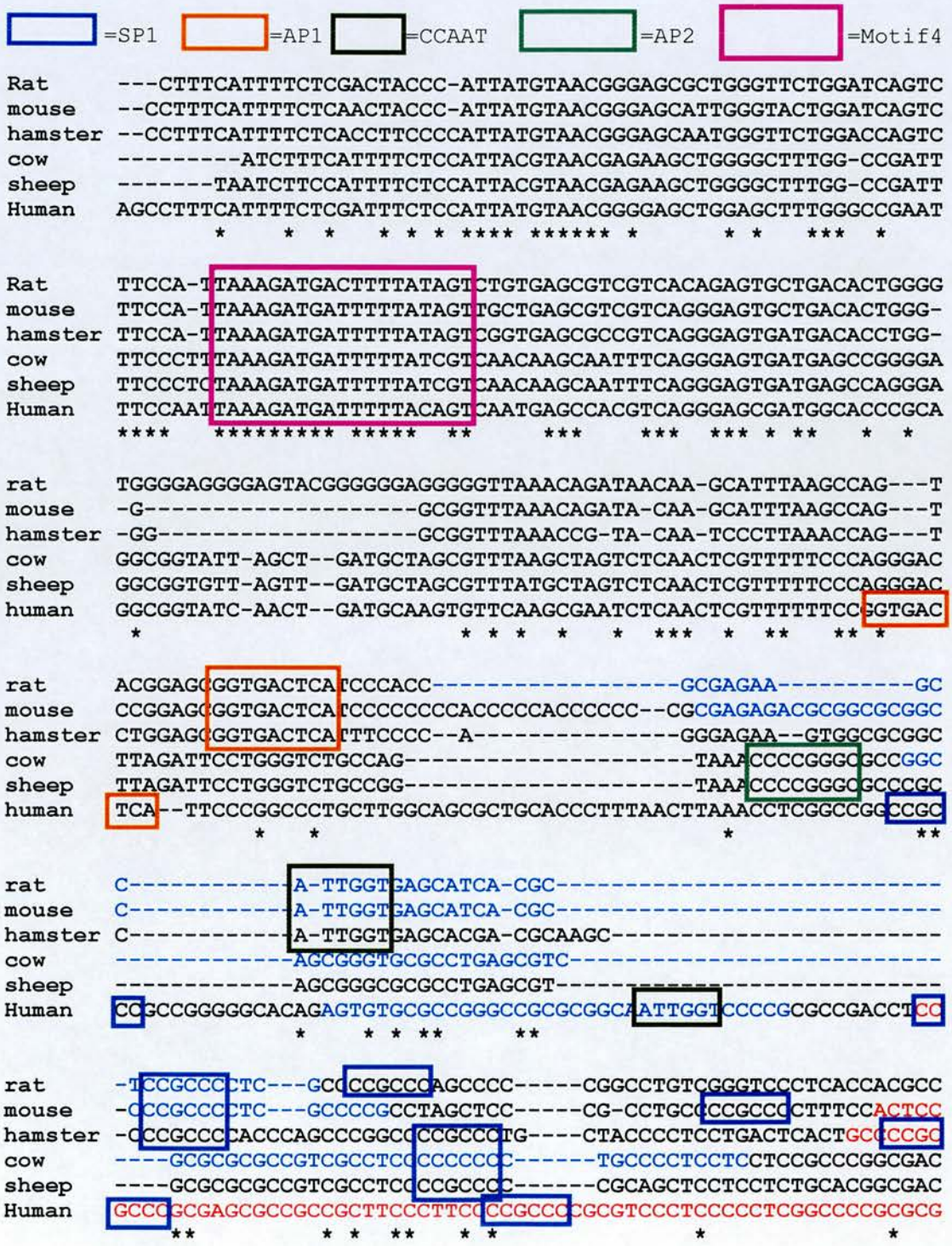
Constructs were transiently transfected into human Sk-N-SH neuroblastoma cells, the region between -43 and -9 (Blue Bar) have a large effect on expression while removal of the first exon abolishes expression to background levels.

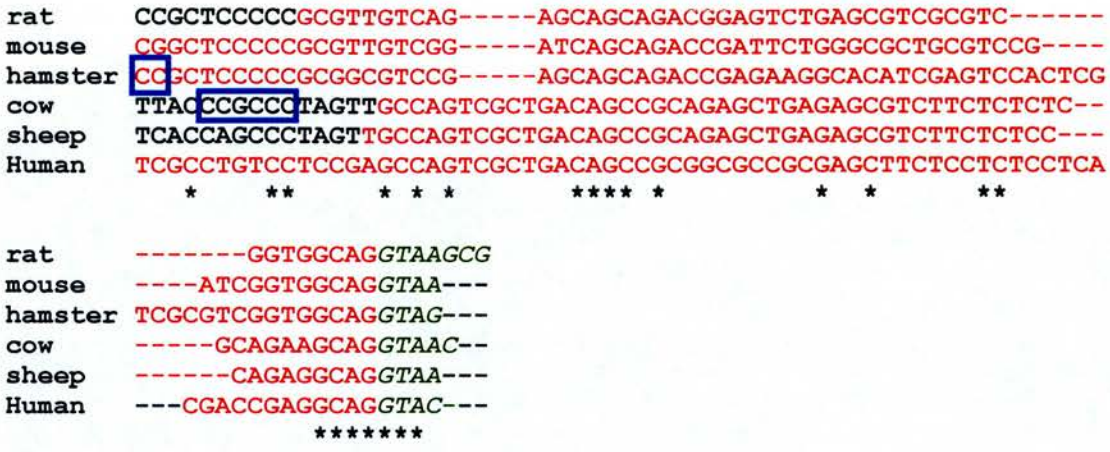
Table 3-1 CAT activity levels of *PRNP* regulatory regions constructs transfected into human Sk-N-Sh neuroblastoma cells. Constructs were transiently transfected in Sk-N-Sh cells between 2 and 30 times. Cells were co-transfected with either β-galactosidase or luciferase reporter genes under the control of an SV40 promoter to act as a control for variation in transfection efficiency. Construct –1558 was always transfected alongside the other construct and its expression level taken as 100. n=number of transfections

Construct	Mean CAT activity	Standard error	n
-1558	100	-	30
-730	105.5	5.5	4
-364	105.3	3.2	3
-133	84.8	8.4	9
-43	93.4	8.8	11
-9	32.3	6.1	12
+289	1.5	0.8	6
Δ+145:+3259	51.7	4.7	19
Δ+1925:+2106	103	0.1	2
Δ+1558:+2106	105.5	0.5	2
Δ+847:+2106	102.5	3.0	4
Δ+695:+2106	102.0	14.0	2
Δ+622:+2106	81.1	7.7	10
Δ+292:+2106	16.4	4.5	14
Δ+296:+1084	11.3	2.9	15
Δ+257:+2106	11.3	4.0	4
Δ+296:+380	56.6	9.0	9
Δ+2067:+3267	102.1	15.5	12
+292:+1082Rev.	50.3	5.5	22
+292:+1088@+2068	8.0	1.5	16
+292:+1088 Rev.@+2068	16.1	2.6	16
pCAT3 Basic	1.5	0.4	15

Figure 3-16 Alignment of mammalian PrP gene exon 1 and upstream sequences

CLUSTAL W alignment of human bovine murine, ovine and rat PrP upstream regions and exon 1. Key red=exon blue=regulatory region by deletion assay green=intron. *=conserved position in all species. Where exons have multiple start sites the most 5′ start site is used.




rat CCGCTCCCCCGCGTTGTCAG-----AGCAGCAGACGGAGTCTGAGCGTCGCGTC------
mouse CGGCTCCCCCGCGTTGTCGG-----ATCAGCAGACCGATTCTGGGCGCTGCGTCCG----
hamster CCGCTCCCCCGCGGCGTCCG-----AGCAGCAGACCGAGAAGGCACATCGAGTCCACTCG
cow TTACCCGCCCTAGTTGCCAGTCGCTGACAGCCGCAGAGCTGAGAGCGTCTTCTCTCTC--
sheep TCACCAGCCCTAGTTGCCAGTCGCTGACAGCCGCAGAGCTGAGAGCGTCTTCTCTCC---
Human TCGCCTGTCCTCCGAGCCAGTCGCTGACAGCCGCGGCGCCGCGAGCTTCTCCTCTCCTCA
rat -------GGTGGCAGGTAAGCG
mouse ----ATCGGTGGCAGGTAA---
hamster TCGCGTCGGTGGCAGGTAG---
cow -----GCAGAAGCAGGTAAC--
sheep ------CAGAGGCAGGTAA---
Human ---CGACCGAGGCAGGTAC---

3.2.5 Identification of an intronic regulatory region

The relatively high level of expression of construct –9 (32%) suggests that the region downstream of the transcription start site may contain sequences that regulate gene expression. To examine this possibility the intron and putative second human untranslated exon (Lee et al., 1998) were removed by digesting construct -1558 with *Xma*I and ligating the largest fragment to itself. The resulting construct (Δ+145:+3259) had only 52% of the expression levels of the full construct implying that the intron contains sequences that regulate transcription (Figure 3-17). To determine which sequences within the intron affected transcription a further set of unidirectional deletions was made using the Erase-A-Base system by digesting construct -1558 with *Nsi*I and *Csp45*I. (Figure 3-14) Digesting the plasmid back-bone with Exonuclease III removed the 5' overhang created by *Csp45*I, but not the 3'overhang created by the *Nsi*I site, thus creating a set of unidirectional deletions in the opposite direction to the upstream sequence of deletions (Constructs Δ+1925:+2106, Δ+1558:+2106, Δ+846:+2106, Δ+693:+2106 and Δ+254:+2106). The existence of an *Mlu*I site allowed the creation of construct Δ+292:+2106 by digesting construct –1558 with *Mlu*I and *Nsi*I, removing 5' and 3' over-hangs with S1 nuclease and ligating.

Transiently transfecting Sk-N-Sh cells with constructs which have deletions of the intron showed that removal of sequences between +622 and +2106 had no significant effect on expression (One-way ANOVA p=0.17 table 3-1). However a further deletion to +292 resulted in a drop in expression to 16% of the level of the full construct, implying that this region contains an intronic regulatory sequence (Yellow region Figure 3-17 and 3-18). To confirm that the sequence between +292 and +622 is the critical regulatory region of the intron, a deletion was made (Δ+292:+1082) by digesting construct –1558 with *Mlu*I and *Age*I, treating with Klenow fragment to fill in any remaining over-hangs and ligating the two ends of the plasmid together. This construct removed the critical region and gave equally low levels of expression as Δ+292:+2106 (t-test p=0.24),

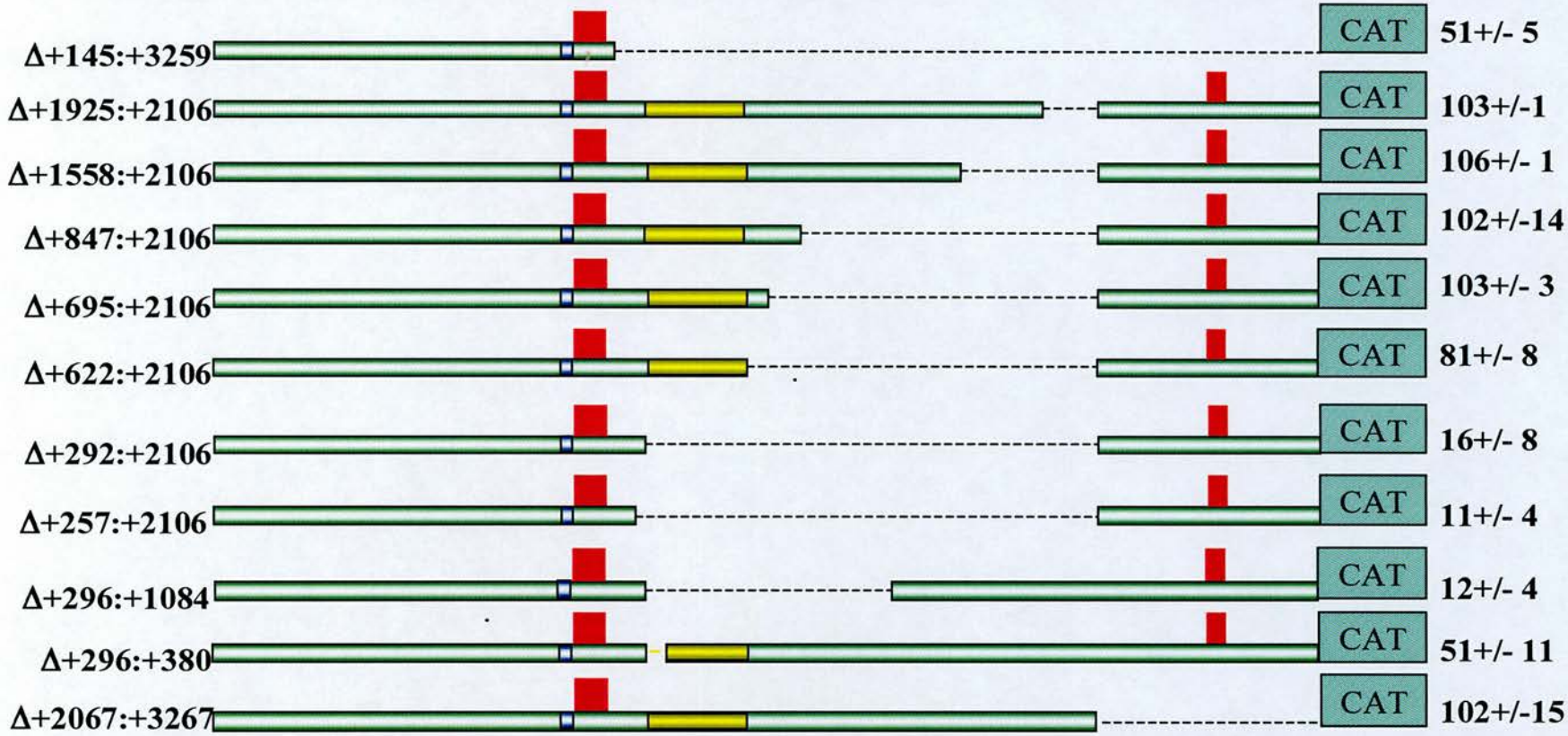
confirming that the drop in expression was due to removal of specific sequences rather than shortening of the intron. To determine if the 330bp intronic regulatory region can be further defined by deletions 84 bases of the 330bp critical region were removed (Δ+296:+380) by digesting construct -1558 with *Mlu*I and *BssH*II treating with Klenow fragment to fill in any remaining over-hangs and ligating the two ends of the plasmid together. Upon transient transfection this construct resulted in CAT activity of 52% of construct -1558 implying that this region within the intronic region may contain important regulatory regions.

One puzzling aspect of the intron deletion constructs is that expression levels are significantly higher when the whole intron is deleted (Δ+143:+3253; 52% of construct -1558) then when only the critical region is deleted. (Δ+292:+1082 and Δ+292:+2106 average expression level: 16%)(t-test $p=4X10^{-8}$ Table 3-1). This may be due to a repressor element near the second exon (between +2106 and +3253) or between +143 and +292. However, removing the region surrounding the second exon, by digesting construct -1558 with *Csp45*I and *Bgl*II (construct Δ+2067:+3253) filling in with Klenow fragment and ligating the two ends of the plasmid together, while leaving the intronic promoter region untouched had no effect on expression this implies that there is no repressor element in this region. It also implies that the second exon has no apparent function in transcriptional regulation in this experimental system, in contrasts with the murine deletion analysis where there was a second promoter region upstream of the second exon. The possibility remains that there is a repressor element between +143 and +292 i.e. between the first exon and the intronic regulatory region. However further deletion analysis within this region is needed to determine if this is correct.

To determine if the intronic region is an enhancer, i.e. can affect expression regardless of its position and orientation, the intronic region was reversed. This was achieved by digesting construct -1558 with *Mlu*I and *Age*I, filling in both fragments with Klenow fragment and religating the *Mlu*I-*Age*I fragment in the opposite orientation (Figures 3-19 and 3-20). This construct (+292:+1082R, R stands for reversed) both reversed and moved the critical region which should have no effect on the functioning of an enhancer,

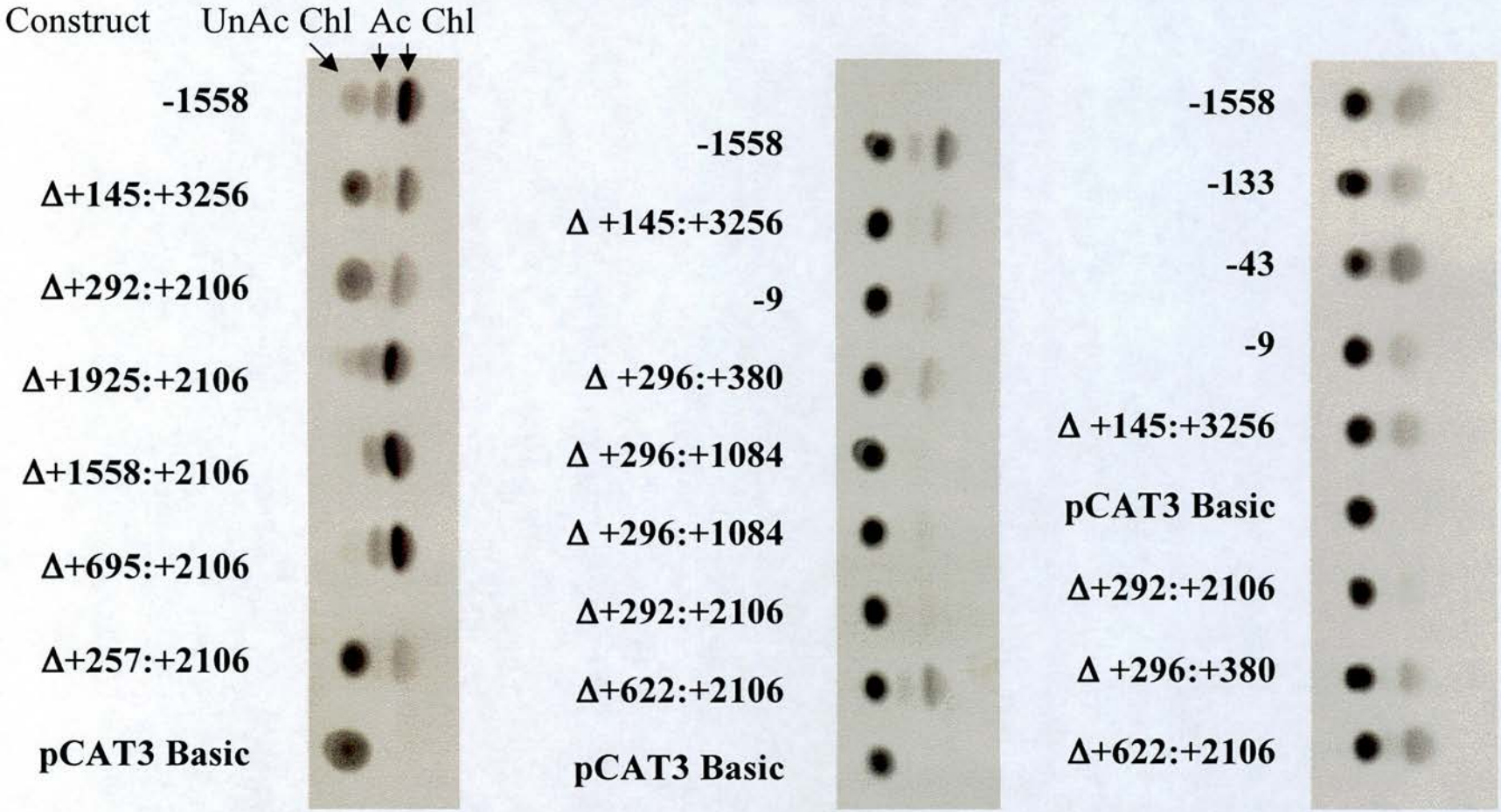
yet abolish the effect of any position dependent non-enhancer elements in this region. However, this construct had only 50% of CAT expression levels relative to construct -1558 implying that this region is not a true enhancer as being reversed and moved affects its expression level. To further determine the effect of position on the intronic element, the *Mlu*I-*Age*I fragment was ligated into the *Csp45*I site of construct Δ+292:+1082 in both the forward and reverse orientations (constructs +292:+1082@+2068 and +292:+1082rev @+2068). These two constructs drove expression at levels as low as when the intronic region was deleted. Therefore, the intronic region does not appear to contain an enhancer element. The intronic region must therefore act by regulating expression in a positional dependent manner, this may imply that it acts with the upstream regulatory region in driving expression from the first exon possibly in conjunction with regulatory elements within the exon.

Figure 3-17 The Effect of deletions within the intron on expression levels.



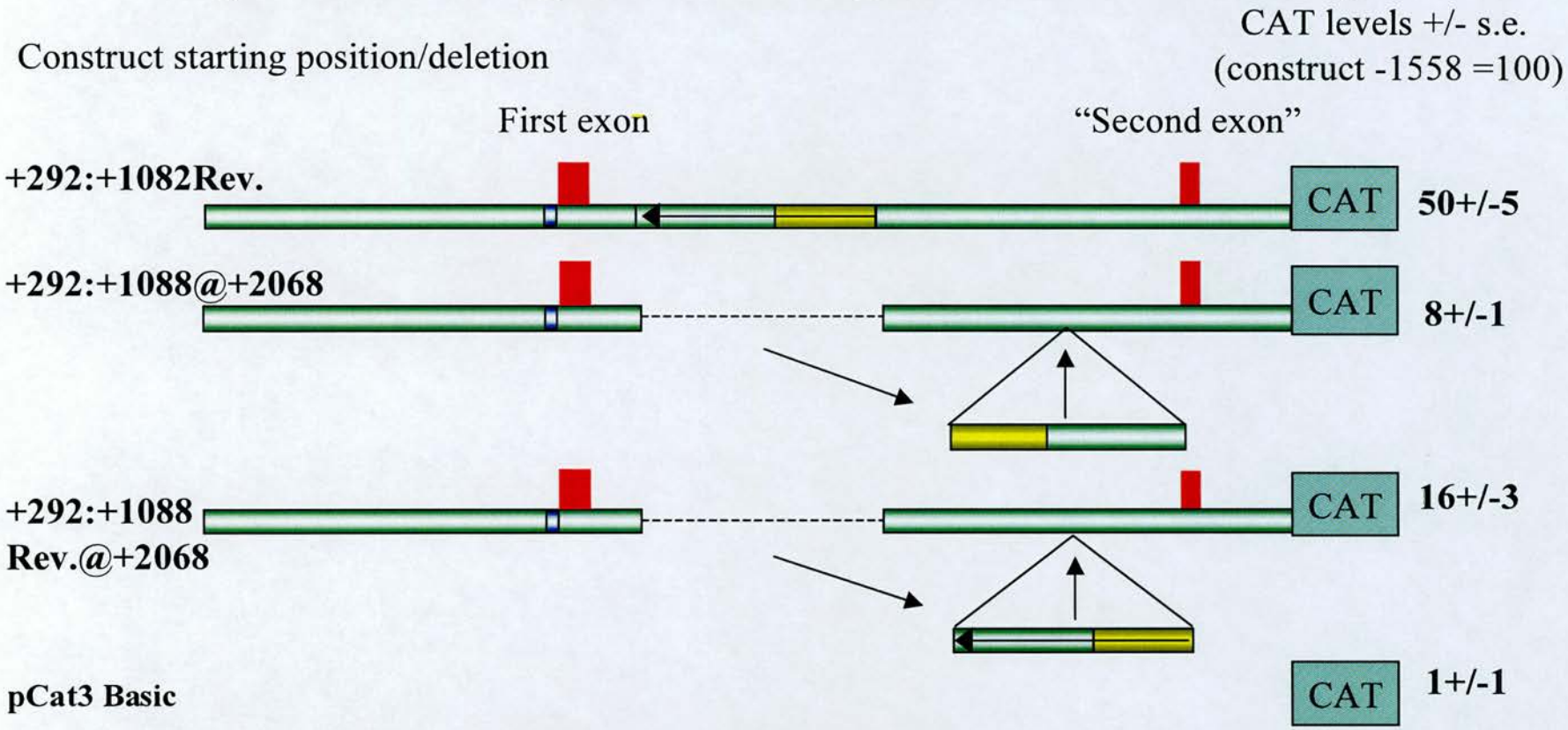
As construct Δ+145:+3259 has high levels of expression, a set of deletions within the intron was made, to determine which region affects expression. The drop in expression between Δ+622:+2106 and Δ+292:+2106 implied the 330bp region between +292 and +622 as the intronic regulatory region (yellow). This was confirmed by deletion Δ+296:+1084. A deletion which removed part of the region reduced expression by a large degree but not by as much as deleting the entire region implying that there may be two elements within the intronic region. The difference in expression levels when the intronic region is removed seen in constructs Δ+145:+3256 and Δ+292:+2106 is not due to a represser element near the second exon as seen by construct Δ+2067:+3267

Figure 3-18 Sample CAT assays of upstream and intron deletions



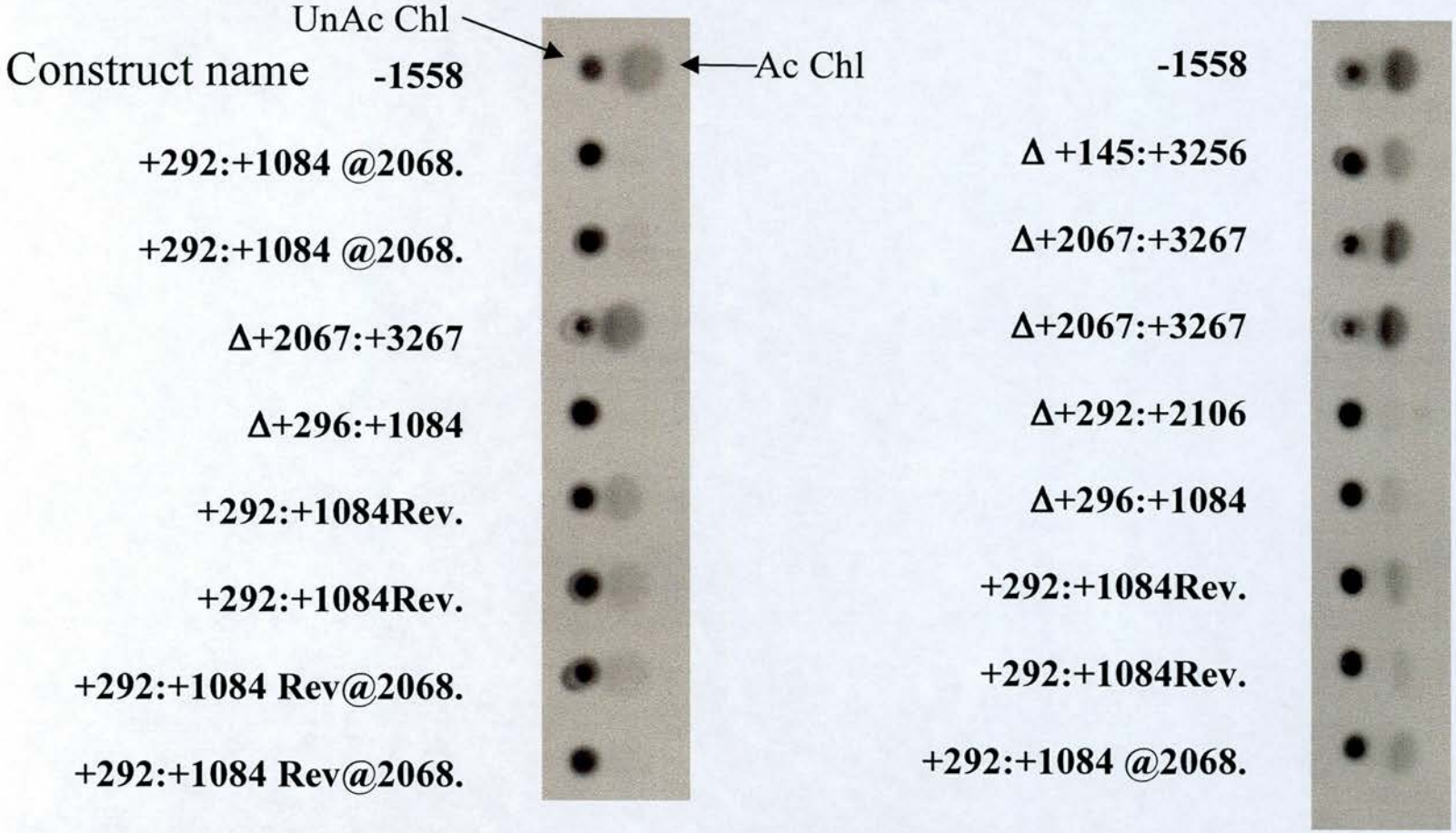
Sample CAT assays from constructs in figures 3-13 and 3-15 and UnAc-Chl=Unacetylated Chloramphenicol, AcChl= Acetylated Chloramphenicol, there are two bands of acetylated chloramphenicol as there are two acetylation sites on the chloramphenicol molecule. Cells were transiently transfected with the indicated plasmid then cell extracts were incubated with Chloramphenicol for 30 minutes, the chloramphenicol was then run on a silica TLC in Chloroform, the plates were exposed to film overnight

Figure 3-19 Testing the intronic region for enhancer activity



In order to determine if the intronic regulatory region is an enhancer construct -1558 was digested with *Mlu*I and *Age* I. This 700bp fragment had its overhangs filled in with Klenow fragment and ligated in the opposite direction (+292:+1082Rev). This alteration reduced expression but not by as much as removing the intron site. Inserting this fragment in either orientation in a site further downstream resulted in a reduction of expression implying that this region is not an enhancer.

Figure 3-20 Sample CAT assays of intron deletions and enhancer constructs.



Sample CAT assays from constructs in figures 3-13, 3-15 and 3-18. UnAc-Chl=Unacetylated Chloramphenicol, AcChl= Acetylared Chloramphenicol. Cells were transiently transfected with the indicated plasmid then cell extracts were incubated with Chloramphenicol for 30 minutes, the chloramphenicol was then run on a silica TLC in chloroform, the plates were exposed to film overnight

3.3 Discussion

Examining the upstream and intronic regulatory regions of the human *PRNP* gene defined in this deletion study shows that the upstream regulatory region is quite short at 34 bases (between –43 and –9) compared with that of other species. A reversed CCAAT site is the only consensus transcription factor binding site in this region. This site is also found in the critical upstream regions of the mouse and rat PrP genes, but not the bovine PrP gene. The human gene also has a regulatory region within the intron. This is also found in the first introns of the murine and bovine genes. The presence of regulatory elements in the intron which affect expression is not unprecedented as many genes have sequences within their introns which affect transcription, such as rat fatty acid synthase, human epidermal growth factor receptor, human myotonic dystrophy protein kinase gene, human fibroblast growth factor receptor, human smooth muscle (aortic type) alpha-actin gene, human noreprinephrine transporter gene and human tissue inhibitor of metalloproteinase1 (TIMP-1) (Storbeck et al., 1998) (Nakano et al., 1991) (McEwan and Ornitz, 1998) (Oskouian et al., 1997) (Gebhardt et al., 1999) (Kim et al., 1999) (Clark et al., 1997). Such elements are often found in introns between untranslated exons as is also seen in *PRNP*. The human intronic regulatory region of *PRNP* appears to be more similar to the bovine PrP gene intronic regulatory region then the murine intronic regulatory region, with a positive regulatory region in the 5' half of the intron. Unlike the bovine gene, which requires both upstream and intronic regions in order to drive expression of the reporter gene, in humans either region can function alone, although each drives expression at a significantly lower level than the full construct. The first exon is however essential for expression as its removal reduces expression to background levels.

The 330bp critical region of the human intron appears to contain at least two distinct elements as a small deletion (construct Δ+296:+380) within the intronic region reduced expression 49% but this is not as much as deleting the entire region (construct Δ+292:+1082) which reduces expression 88%. Analysing the sequence of the human

intronic region and the bovine intronic regulatory region identified by Inoue *et al* (Inoue et al., 1997) for transcription factor or enhancer consensus binding sites shows that the human intronic region contains a CCAAT/Enhancer binding protein beta (C/EBPβ) which is absent in the bovine intronic regulatory region. There is also an SP1 site in both intronic regulatory regions. The SP1 and C/EBP transcription factor binding sites may explain the expression levels seen in deletions Δ+292:+1082, Δ+296:+380 and +292:+1082R. In Δ+292:+1082 both the SP1 and C/EBP sites are removed, leading to a drop in expression while in Δ+296:+380 the C/EBP site was removed, but the SP1 site was left untouched. This resulted in levels of expression less than that of the full construct but greater than those where the intronic regulatory region was deleted. Reversal of the intronic regulatory region sequence (+292:+1082R) may have affected one of the two sites leading to the intermediate levels of expression seen. While moving both sites 1Kb downstream may prevent any transcription factors that bind to these sites from influencing expression (Constructs +292:+1082@+2068 and +292:+1082rev @+2068).

CCAAT enhancer binding proteins (C/EBP) are a family of leucine zipper transcription factors with six members α,β,χ,δ,ε and ζ. Despite its name, the consensus binding sequence for C/EBP β does not contain CCAAT; the consensus sequence is actually RNRTKNNGMAAK (R=A or G, N=A, C, G or T, K=G or T, M=A or C). Conversely the CCAAT reversed motif (ATTGG), which affects expression when deleted from the human and rodent upstream regions, can be bound to by several different transcription factors, the C/EBP family, CBF/NF-Y CTF and NF-I (Maity and Crombrugghe, 1998). Which protein actually binds to this region can only be determined by band-shift assays by adding competitive binding sequences. Another method of determining if the transcription factor binding sites found in the upstream and intronic regulatory regions are functional *in vivo* is to use *in vitro* mutagenesis to alter the transcription factor consensus binding sites, such as the two SP1 within the first exon the reversed CCAAT site in the upstream regulatory region and the C/EBP site and SP1 sites in the intronic regulatory region. Any alterations in expression levels when the mutated constructs are transiently transfected into Sk-N-Sh neuroblastoma cells would imply that these

transcription sites are important in regulating expression of *PRNP*.

The possible presence of a repressor region between the exon and the start of the intronic region suggests that the control of expression may be very complex. Rather than two distinct positive regulatory regions, one upstream and one in the intron, *PRNP* may be controlled by a single regulatory region stretching from –43 to +622, consisting of the positive elements upstream and between +292 and +622 in the intron, a possible negative element between +143 and +292 and potential regulatory elements within the exon (+1 to +134). Further deletion analysis is necessary in this region to test this hypothesis.

This study only examined 1.5Kb of upstream sequence and 3.1Kb of intron. It is therefore possible that there are other regions that regulate *PRNP* expression but were not included in this construct. One possible such element is the putative Purkinje cell specific element thought to be located at the 3' end of the murine second intron (Fischer et al., 1996). There may also be regulatory elements in the construct which have no effect on expression in neuroblastoma cells, but may regulate *PRNP* expression in other tissues; such elements would not be identified in this study. However, this study has identified the first regulatory regions of the *PRNP* gene. These regions can now be used to create transgenic constructs to see the effect of the regions on expression *in vivo* (Chapter 4) the regions can also be sequenced in vCJD patients, sCJD patients and controls to determine if any polymorphisms exist in these regions and if such polymorphisms have any effect on susceptibility to disease (Chapter 5).

4 Chapter Four: *In vivo* analysis of the human PrP promoter.

4.1 Introduction

4.1.1 Plan

The results of the transient transfection experiments identified two regions of the human *PRNP* gene which regulate expression. However, these regions have been only shown to control expression in human neuroblastoma cells. If the level of PrP expression influences disease and polymorphisms in these regions affect susceptibility to CJD then the effect each region has on expression in an entire organism needs to be examined. To determine if the upstream and intronic regulatory elements of the *PRNP* gene can regulate expression of a reporter gene in the PNS, CNS and other tissues where PrP is known to be expressed, transgenic mice were created which express the LacZ reporter gene under the control of the human *PRNP* upstream and intronic regulatory regions to determine if the regions identified as regulating expression *in vitro* control the expression levels, tissue specificity or temporal expression patterns *in vivo*. The upstream and intronic regions may have different effects on expression in different tissues if they contain tissue specific transcription factor sites.

Other studies have shown that human upstream promoter regions can replicate their expression patterns when introduced into mice, implying that mammalian transcription factors and promoters are interchangeable. Intronic regulatory regions of human genes have also been shown to affect expression patterns of a reporter gene in murine embryos. One example of this is the serotonin transporter gene. This gene has a variable number of tandem repeats (VNTR) polymorphism in the second intron, with alleles having 9, 10 or 12 copies of a 17bp repetitive element. Association studies of these alleles with

bipolar disorder (Manic-Depression) and other affective disorders have shown an association of these alleles with disease. Other studies however, have failed to replicate these findings; affective disorder association studies are notoriously difficult to replicate due to difficulties in diagnosing patients and the complex genetic aetiology of the diseases (Sher, 2000). In order to determine if the VNTR polymorphism has an effect on expression, *in vitro* and *in vivo* studies have been carried out. *In vitro* studies used a reporter gene construct where the 10 copy (Stin 2.10) and 12 copy (Stin2.12) alleles of the intronic polymorphism were inserted upstream of a LacZ gene driven by a minimal promoter and transiently transfected into murine embryonic stem cells have shown differences in expression between the two alleles (Finkerstrand et al., 1999). To examine the effect *in vivo* the constructs were injected into fertilised mouse eggs and the expression patterns of the resulting transgenic embryos were examined (MacKenzie and Quinn, 1999). While both alleles drove expression of the reporter gene in developing embryos the Stin2.12 allele construct expressed higher levels of the LacZ reporter gene in the rostral hind-brain of murine at day E10.5. Although there was no difference at E 12.5, this suggests that the two alleles have functional differences in embryonic development which could alter an individual's susceptibility to bipolar disorder in later life. This study also shows that relatively small differences in human regulatory regions can be detected in transgenic murine embryos.

To analyse the effect of the human *PRNP* gene regulatory regions *in vivo*, the upstream and intronic *PRNP* regulatory regions were cloned into a construct containing the β-galactosidase reporter gene. These constructs were then micro-injected into fertilised mouse eggs which were implanted into pseudo-pregnant mice. The 3.4Kb murine *Prnp* gene regulatory sequence identified by Baybutt and Manson (Baybutt and Manson, 1997) was used as a control to determine if expression of a reporter gene driven by the upstream and intronic regulatory regions of the murine *Prnp* gene can reproduce the known expression patterns of PrP. This was compared with the expression pattern of the construct containing the human upstream and intronic elements to determine if the differences between murine and human transcription factors affect the expression of a reporter gene controlled by the human *PRNP* regulatory regions in a murine

environment. Finally, constructs containing either the human upstream or intron regulatory region were injected to determine the effect of each of these elements individually. This was carried out to determine if the upstream and intronic elements act individually or co-operatively in particular tissues. Embryos and adult mice were analysed for the presence of the transgene by PCR of the extra-embryonic tissues or tail-snip DNA respectively. Temporal and tissue specific expression of PrP was examined by staining transgenic adults and embryos at different stages of development for β-galactosidase activity.

If this study shows that a particular regulatory region is identified as controlling expression in a particular tissue this can be used to predict the possible effect of any polymorphisms or mutations found in these regions in the sequencing study in Chapter 5. If these polymorphisms are associated with disease, then tissues whose expression is controlled by the regulatory region may be implicated in the initiation or progression of disease. Any regulatory region alleles pre-disposing individuals to disease would then be used to create new transgenic constructs so that a direct comparison of the two alleles can be made to determine in which tissue expression is altered.

Due to random integration of micro-injected constructs into the murine genome, inactivation of transgenes or ectopic activation by neighbouring enhancer elements is a potential problem in altering expression patterns. Therefore an expression pattern from a given transgenic construct needs to be observed in several embryos or transgenic lines. The micro-injection work and analysis of transgenic mice described in this Chapter was carried out jointly by myself and Dr. David Stott at the Department of Biology, Warwick University, UK; all the implantation of injected eggs was carried out by Dr.Stott.

4.2 Results

4.2.1 Creation of transgenic constructs.

The reporter gene plasmid used to make the transgenic constructs was placF which contains a LacZ reporter gene. It also contains sequences from the mouse protamine gene which provide an intron, 3'UTR and poly-adenylation signal; although coding sequences from the protamine gene are included they are not translated. The human regulatory region CAT assay deletion constructs -1558, -9, and Δ+143:+3259 from Chapter 3 were digested with *Sca*I and *Bst98*I to remove the pCAT3 basic plasmid backbone, filled in with Klenow fragment and ligated into the *SalI* site of the polylinker of placF, which had also been filled in with Klenow fragment. This resulted in constructs full-lacF, downstream-lacF and upstream-lacF (Figures 4-1, 4-2, 4-3, 4-5).

These constructs all contained the chimeric intron from the pCAT3 basic plasmid which serves to provide a splice acceptor site for the first exon, thereby avoiding splicing from the *PRNP* first exon to the protamine exon in LacF. A 3.5Kb *BamH*I-*Kpn*I murine fragment from construct PrP-7 (Baybutt and Manson, 1997) equivalent to the full human promoter was sub-cloned from the pCAT basic plasmid (Promega) into the *Sma*I site of the pLacF polylinker (Figure 4-4) by Dr. Herbert Baybutt (NPU). This construct was termed Murine-lacF. The exon-intron structure of the Murine-lacf construct differed from the human constructs as the murine sequence ended in the second exon and did not contain the chimeric intron (which is present in the human constructs) as it was originally cloned into the pCAT basic plasmid rather than the pCAT3 basic plasmid. In contrast the human placF constructs all contained a chimeric intron from pCAT3 basic and the human sequence ends in an intron thus contains a splice donor site.

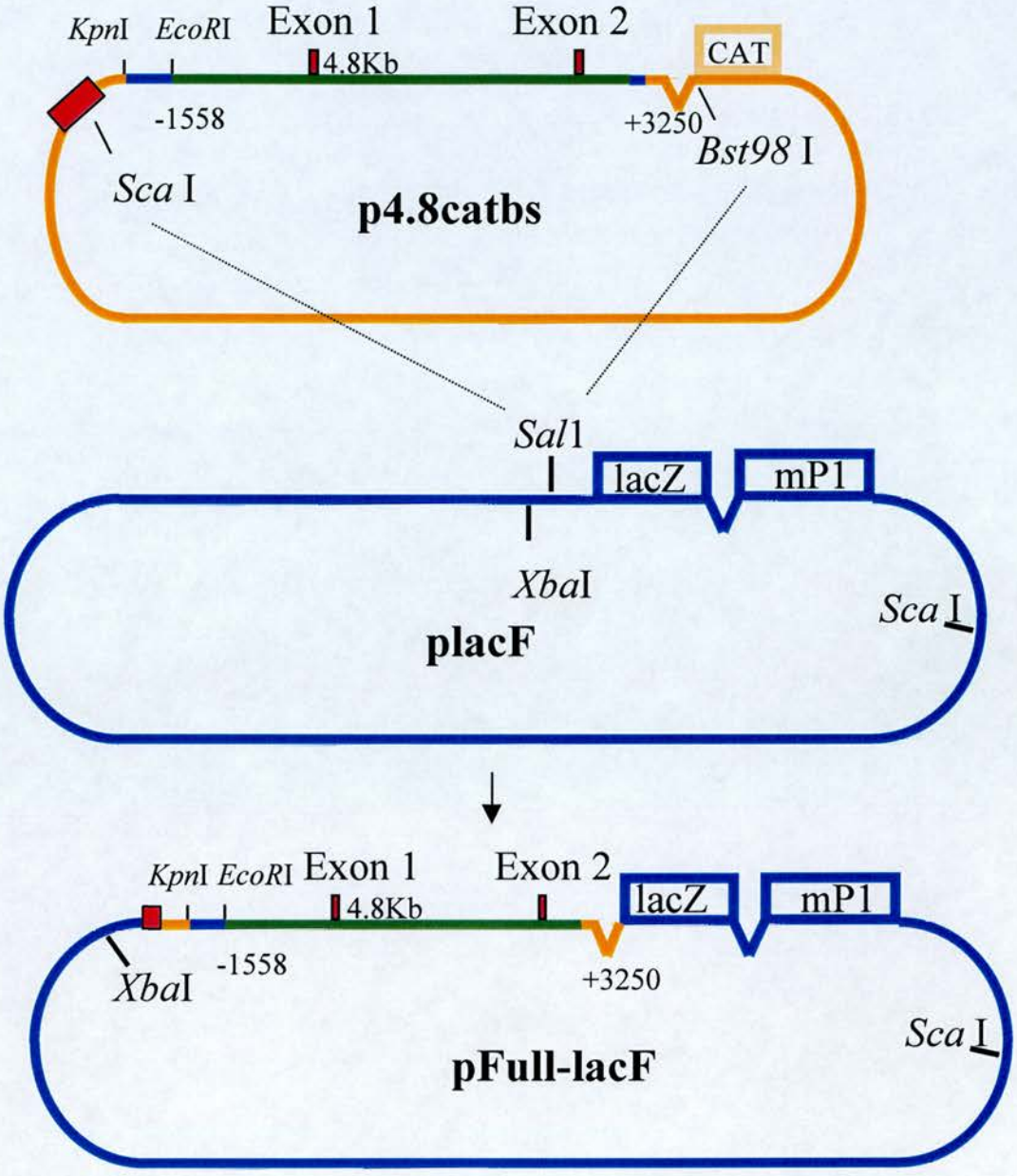


Figure 4-1 Cloning of pFull-LacF: The full human promoter construct p4.8BCAT (construct –1558 in deletion series) was digested with *Sca*I and *Bst98*I. This 5.2Kb *Sca*I-*Bst98*I fragment contains the human sequence plus flanking regions from pCAT3 basic. This fragment was cloned into the *Sal*I site of pLacF resulting in plasmid pFull-LacF. The entire plasmid was transiently transfected into N_2a cells but for the creation of transgenic mice the pFull-LacF plasmid backbone was removed by digesting with *Sca*I and *Xba*I.

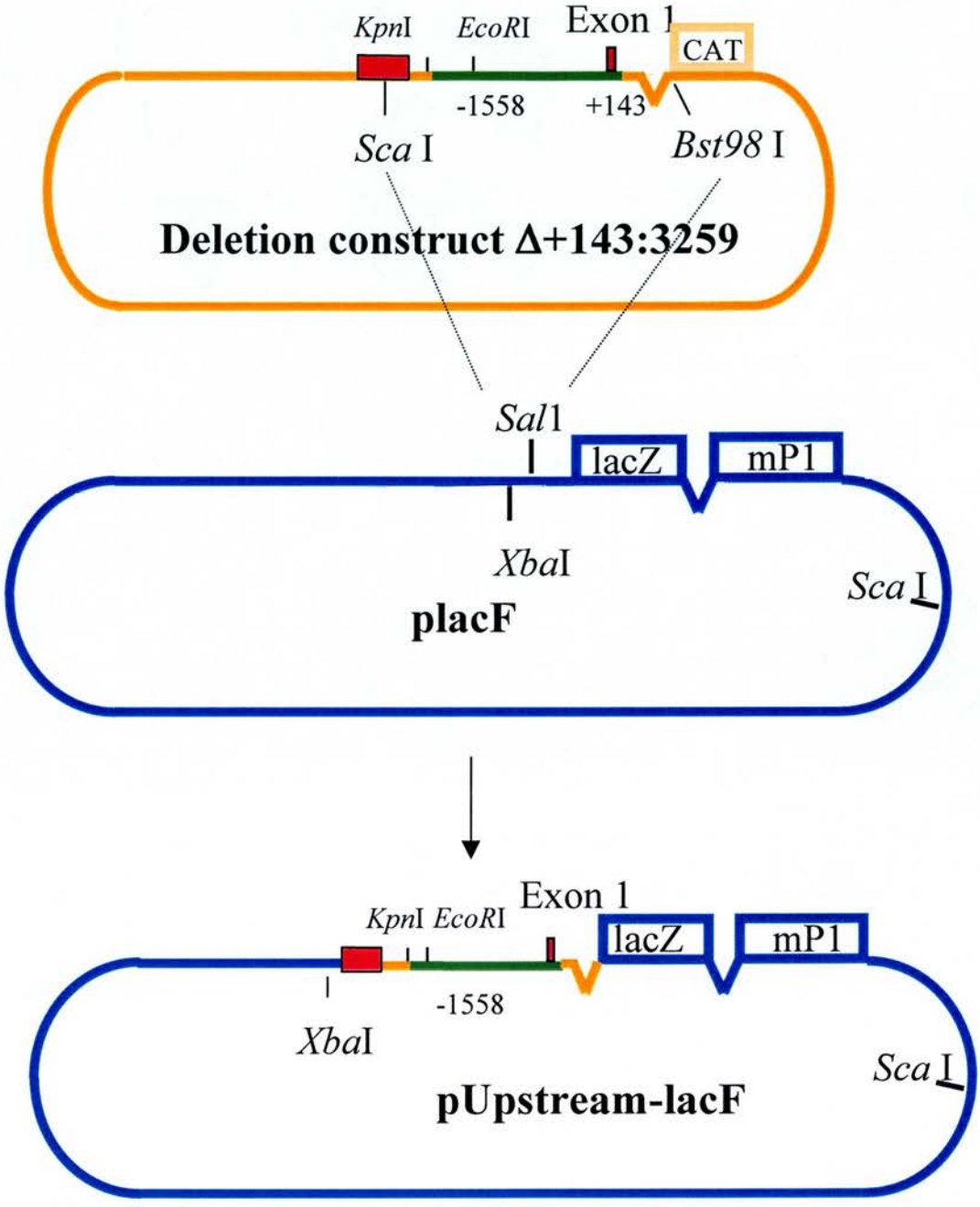


Figure 4-2 Cloning of pUpstream-LacF The human regulatory region construct Δ+143:3259 was digested with *Sca*I and *Bst98*I. This 2.1Kb *Sca*I-*Bst98*I fragment contains the human sequence plus flanking regions from pCAT3 basic. This fragment was cloned into the *Sal*I site of pLacF resulting in plasmid pUpstream-LacF. The entire plasmid was transiently transfected into N_2A cells but for the creation of transgenic mice the pLacF plasmid backbone was removed by digesting with *Sca*I and *Xba*I.

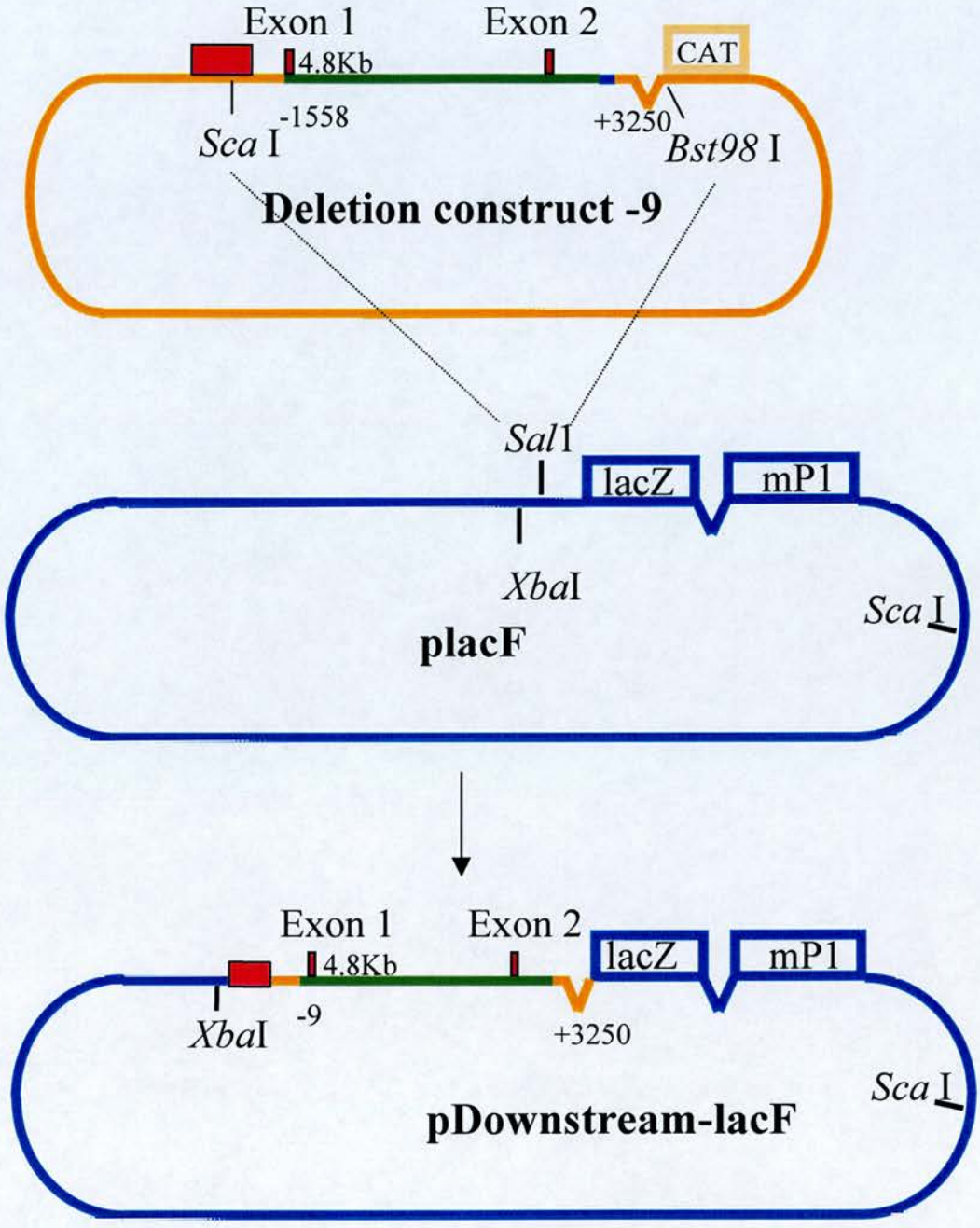


Figure 4-3 Cloning of pDownstream-LacF: The human regulatory region construct –9 was digested with *Sca*I and *Bst98*I. This 3.7Kb *ScaI-Bst98*I fragment contains the human sequence plus flanking regions from pCAT3 basic. This fragment was cloned into the *Sal*I site of pLacF resulting in plasmid pUpstream-LacF. The entire plasmid was transiently transfected into N_2A cells but for the creation of transgenic mice the pLacF plasmid backbone was removed by digesting with *Sca*I and *Xba*I.

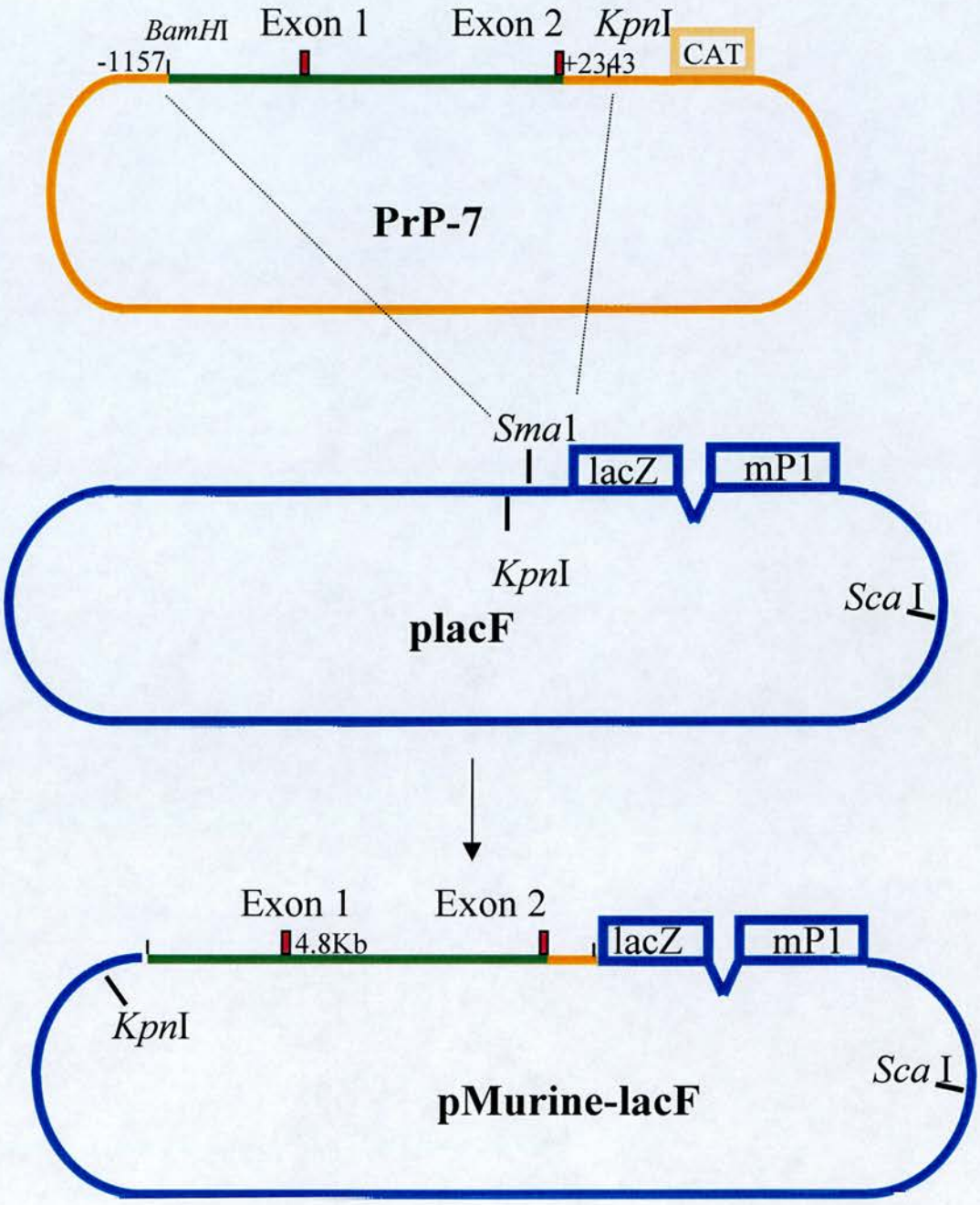
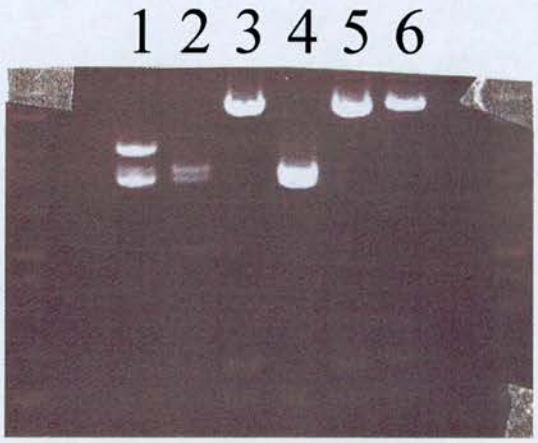
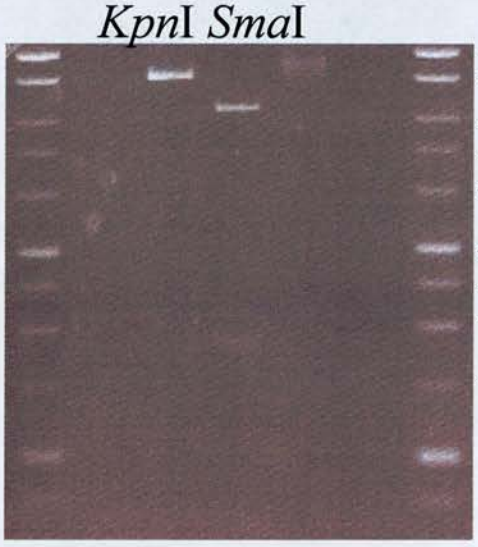


Figure 4-4 Cloning of pMurine-LacF: The murine regulatory region construct PrP-7 (Baybutt and Manson 1997) was digested with *BamH*I and *Kpn*I. *Bam*HI cuts upstream of the regulatory region while *BamH*I cuts within the murine second exon. This 3.4Kb *BamH*I-*Kpn*I fragment therefore contains the human sequence plus flanking regions from pCAT3 basic. This fragment was treated with Klenow fragment to fill in any over-hangs and cloned into the *Sma*I site of pLacF resulting in plasmid pFull-LacF. The entire plasmid was transiently transfected into N_2A cells but for the creation of transgenic mice the LacF plasmid backbone was removed by digesting with *Kpn*I and *Sca*I.

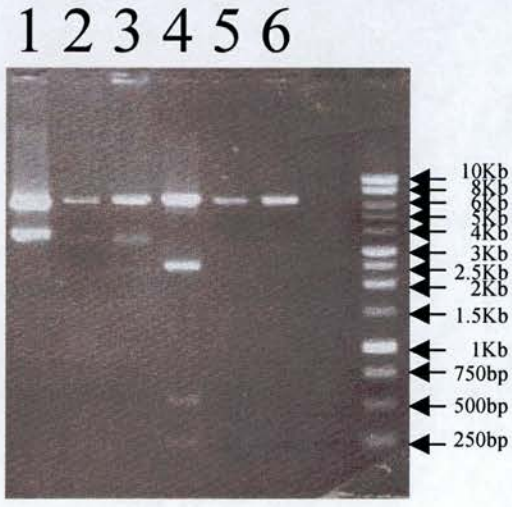
Figure 4-5 Restriction analysis of lacF cloning steps



Full-lacF: *EcoR*V digest of putative full-lacF clones. lanes 2 and 4 show 5.2Kb, 5.3Kb and 1Kb pattern of the full human sequence sub-cloned into placF in the forward orientation.



Upstream-lacF: *Kpn*I and *Sma*I digest of a putative clone of upstream-lacF. *Kpn*I digest gave a 8.3Kb and 0.1Kb band (0.1Kb band not visible). *Sma*I gave a 6.5Kb and 1.8Kb band



Downstream-lacF: *Bgl*II digest of three putative clone of downstream-lacF (lanes 1-3) show a 6.1Kb and 3.9Kb bands consistent with cloning of the downstream region of human DNA into placf in the forward orientation. *Sma*I digest of the same three clones(lanes 4-6) shows that only the first colony (lane 4) contains the 6.6Kb 2.5Kb, 0.5Kb and 0.25Kb bands

4.2.2 LacF constructs *in vitro*.

To determine if the human *PRNP* regulatory would function in mice and if the sub-cloning of the regulatory regions from the CAT constructs into lacF affected expression, the four lacF constructs were transiently transfected into murine N_2a neuroblastoma cells. As a control, a β-galactosidase gene under the control of an SV40 promoter was also transfected. Cells were transiently transfected using the same calcium phosphate protocol as the transient transfections in Chapter 3. The constructs were co-transfected with the pGL3 luciferase construct to normalise transfection efficiency. All four LacF constructs produced β-galactosidase activity in N_2a cells as measured by cell extract assay and by *in situ* staining (Table 4-1). As the murine regulatory regions had driven a CAT reporter gene in the same murine neuroblastoma cell-line (Baybutt and Manson, 1997), the sub-cloning of the regulatory region has not affected its efficiency or the ability of the LacZ reporter gene to produce the correct transcript. The full human construct had higher activity than the upstream or downstream constructs, which was similar to the results in the Sk-N-Sh cells with CAT reporter gene. However, the downstream construct expressed at higher levels than the upstream construct, a reversal of their relative strengths in Sk-N-Sh cells. The full human construct gave higher levels of expression than both the murine-lacF and an SV40 promoter control plasmid.

Table 4-1 β-galactosidase activity of lacF constructs transiently transfected into murine N_2a neuroblastoma cells.

n=number of transfections

Construct	Average β-galactosidase activity relative to Full-lacF.	n
SV40 LacZ	0.45	2
Murine-lacF	0.21	2
Full-lacF	1.00	2
Upstream-lacF	0.32	2
Downstream-lacF	0.69	2

The results from the transient transfection studies show that the sub-cloning of the human regulatory regions into the pLacf reporter construct did not disrupt the regulatory ability of the human regions. The relative differences between the full human construct and the upstream and downstream constructs may be due to different interactions with murine transcription factors. The low level of expression of the murine-lacF construct relative to the human construct could be due to differences in the exon intron structure between the two constructs as the murine construct does not contain any of the murine second intron and does not have a chimeric intron. Splicing therefore occurs from the first exon to the second exon which is fused to the LacF ORF containing exon. Whereas in the human constructs the first exon splices to the chimeric intron acceptor site, possibly including the second exon-like sequence, which includes the ORF.

4.2.3 Expression of lacF constructs *in vivo*

Having shown that the regulatory regions of the human *PRNP* gene can regulate expression of a reporter gene in murine cells, the human regulatory region lacF constructs were digested with *Xba*I and *Sca*I to remove the plasmid back-bone, while murine-lacf was digested with *Kpn*I and *Sca*I. The digested plasmids were run out on agarose and the fragment containing the regulatory regions and LacZ gene purified from the agarose using the Qiaquick gel extraction kit (Qiagen) eluted in PBS and micro-injected into the pro-nucleus of fertilised eggs. The results of the injections are shown in table 4-2.

Full-lacF: Two of five embryos recovered at E10.5 (Day 10.5 of development i.e. 10.5 days post coitus) carried the full-lacF transgene (as determined by PCR Figure 4-6). The normal pattern of PrP expression at E10.5 as previously reported (Miele, 1999) show that PrP is expressed in the telencephalon, mesencephalon and metencephalon the epithelia of the neural tube and the otic pits. However, no β-galactosidase staining was observed in any of the embryos (Table 4-2).

At developmental stage E12.5 eight embryos were recovered, including one which had not developed fully and appeared similar in size and level of development to an E10.5 embryo, four of the eight were transgenic (including the underdeveloped embryo) as determined by PCR. At this stage of development PrP expression has previously been detected (Manson et al., 1992) by *in situ* hybridisation in the developing brain and spinal cord and also in some non-neuronal tissues but no staining was observed after 24 hours incubation in X-gal staining solution. After staining for 48 hours weak β-galactosidase staining was seen in the tissue surrounding the 4th ventricle, and between the telencephalic ventricles (Figure 4-7). This staining was seen in all 7 fully developed E12.5 embryos however as only 3 of the 7 were transgenic, as determined by PCR, the staining may be due to endogenous β-galactosidase activity. Mammalian β-galactosidase is active at slightly acidic pH while bacterial β-galactosidase works best at neutral pH but is still active at slightly alkaline pH which inactivates mammalian β-galactosidase. The pH of the X-gal buffer was 7.0 this might therefore allow some endogenous β-galactosidase activity to be detected after prolonged staining. The fact that the staining was weak and there was no difference between transgenic and non-transgenic embryos suggests that the transgene was not functional in these embryos.

Upstream-lacF and Downstream-lacF: Two upstream LacF and one downstream LacF E12.5 embryo were produced. All three embryos were positive for the transgene by PCR but no staining was observed after 24 hours. After 48 hours of staining a similar pattern of staining to the background staining of the full-construct were observed. Again suggesting that there was no transgene expression.

Mouse-lacF: Three of six E12.5 embryos injected with the mouse construct contained the transgene but none stained for β-galactosidase. Although background X-Gal staining was again seen in both transgenic and non-transgenic embryos after 48 hours.

Expression in adults:

As the murine regulatory region construct failed to express the reporter gene in embryonic mice, some implanted eggs injected with the murine construct were allowed

to develop to term and expression in adult mice was examined. Seven transgenic adult mice were produced but no X-Gal staining was detected in paraformaldehyde fixed brain, heart, lung, kidney, spleen, muscle or liver in these animals. Therefore, these constructs do not appear to be able to express the reporter gene in embryonic or adult mice. The analysis of the adult mice was carried out by Dr.David Stott.

Table 4-2 Results of transgenic experiments

Construct	Eggs injected	Developmental Stage	Number of embryos/adults	Number of transgenics	β-Gal. Staining
Full-lacF	32	E10.5	5	3	0
Full-lacF	39	E12.5	8	4	0
Up-lacF	31	E12.5	2	2	0
Down-lacF	22	E12.5	1	1	0
Mouse-lacF	44	E12.5	6	3	0
Mouse-lacF	88	Adult	11	7	0

Figure 4-6 PCR analysis of Extra-embryonic tissues from 8 E12.5 embryos injected with the full-LacF construct

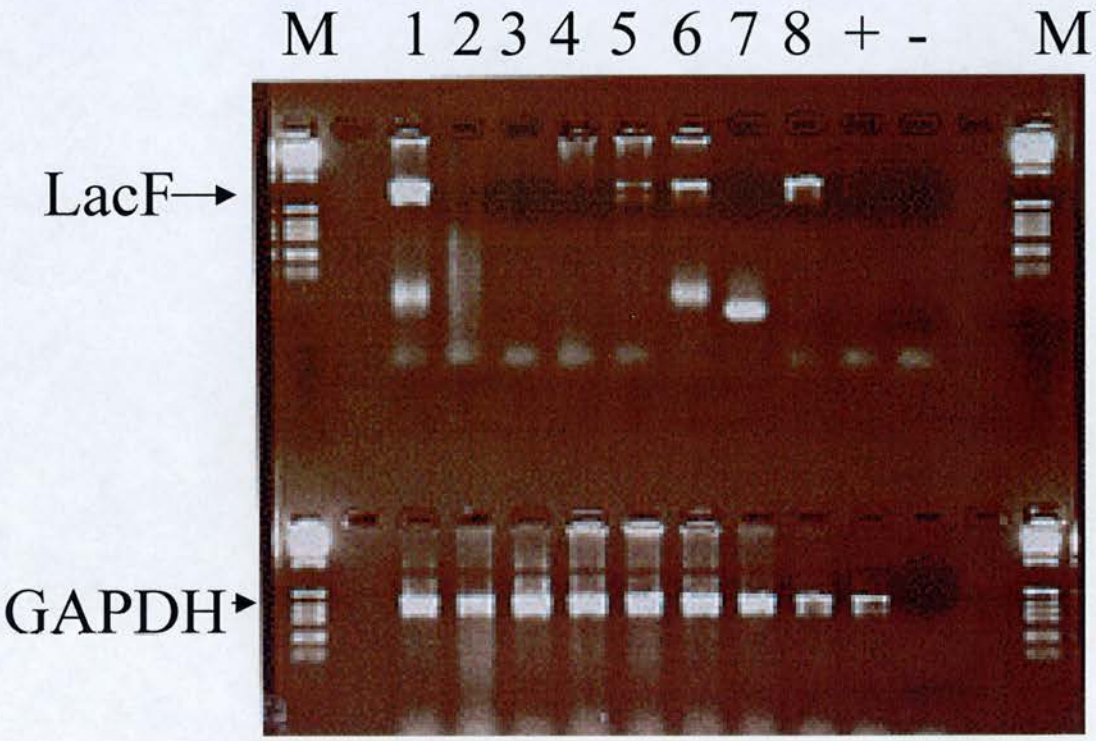


Figure 4-6 PCR analysis of Extra-embryonic tissues from E12.5 embryos injected with the full-LacF construct. The extra-embryonic membranes from the 8 embryos were used to prepare DNA which was used in PCR reactions to detect the lacZ open reading frame of the pLacf construct (Upper lanes) and GAPDH as a control (Lower lanes). Embryos 1, 5, 6 and 8 were positive. The positive control is DNA from a different line of transgenic mice which contain a pLacF construct, it gave a very weak signal in the Lacf PCR which is not visible in this picture but is just visible in the original photograph. The level of the GAPDH PCR product is also smaller in the control than most of the other embryos possibly due to a lower DNA concentration or degradation of the DNA

.

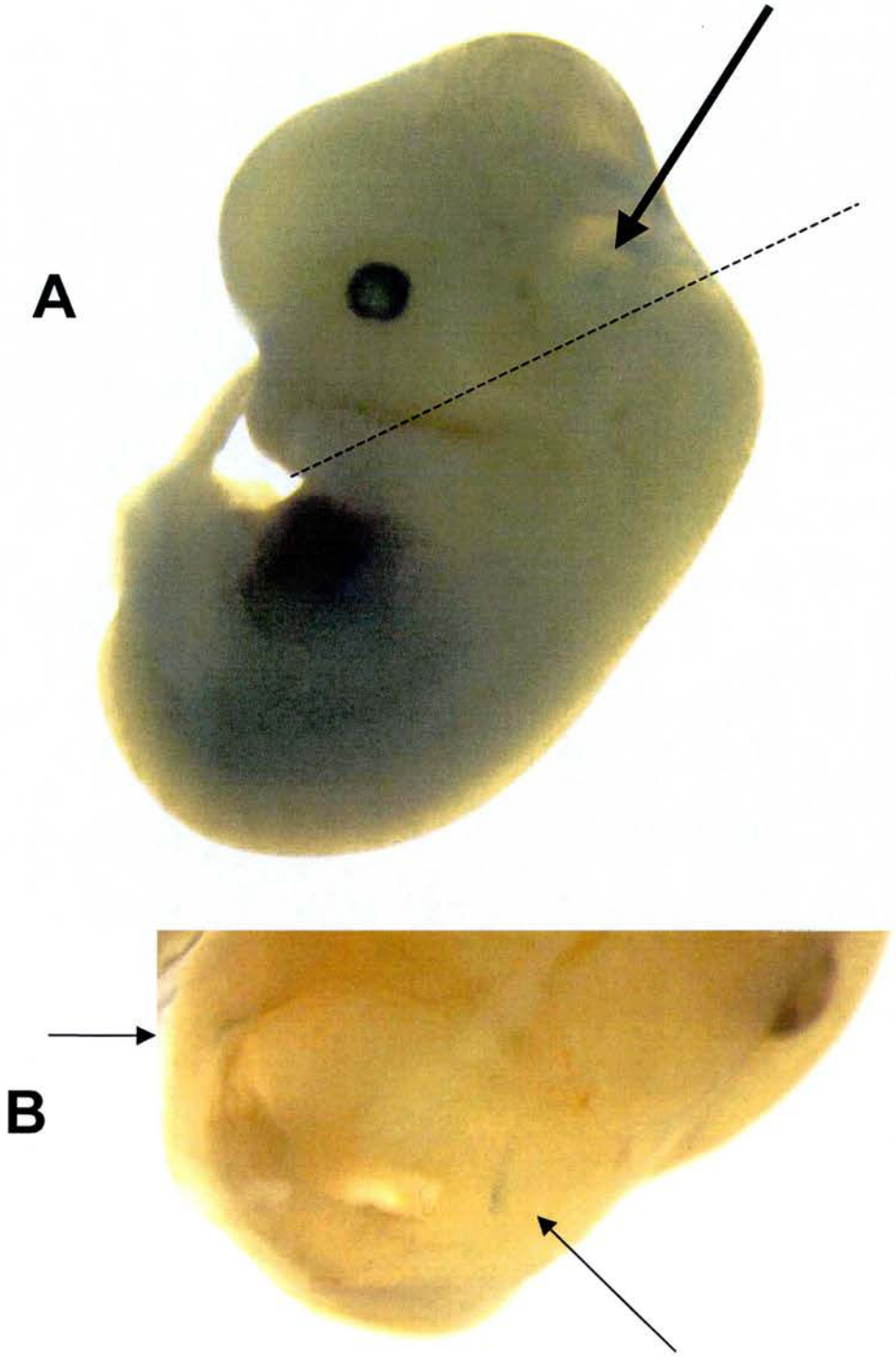


Figure 4-7 Background β-galactosidase in E12.5 embryos. Embryos which contained (A) or didn't contain (B) the transgene when analysed by PCR were stained for LacZ activity for 48hours. Weak staining was observed in the tissue surrounding the 4th ventricle (arrow in A). the non-transgenic embryo was dissected as shown by the dashed line and the β-galactosidase staining was seen in a thin ridge or tissue either side of the 4th ventricle (Arrows in B)

4.3 Discussion

4.3.1 Expression of PrP

PrP is expressed at high levels in the brain, spinal cord and PNS and at lower levels in many other non-neuronal tissues. The expression of PrP during murine development has been analysed by *in situ* hybridisation (Manson et al., 1992). At E9.5 no PrP was detected but at E13.5 and E16.5 abundant transcripts are found in the developing CNS, PNS and in non-neuronal tissues. Recent analysis (Miele, 1999) has detected PrP transcripts at E9 using RT-PCR. The level of PrP expression in the brain increases during development relative to a β-actin control until post-natal day 20 when it reaches its adult level (Lazarini et al., 1991). Attempts to determine the temporal and spatial pattern of PrP expression using wholemount *in situ* hybridisation using DIG labelled PrP riboprobes shows that PrP is undetectable at E8.5 but is detectable at E9.5 in the telencephalic, mesencephalic and metencephalic vesicles of the developing brain (Miele, 1999). Therefore a reporter gene should be detectable in the transgenic mice if the regulatory regions are sufficient to drive expression.

4.3.2 Possible reasons for the lack of expression of the constructs

Transgenic mouse embryos containing the full human promoter did not express the β-galactosidase transgene at E10.5 or E12.5, while the equivalent murine region similarly did not drive expression at E12.5 or in adult mice. Given this failure of the murine and full human constructs to express LacZ, it is unsurprising that the upstream-lacF and downstream-lacF constructs also failed to drive LacZ expression as they contain less regulatory sequence than the full human lacf construct.

Although in this study the human and murine sequences failed to promote expression

of a reporter gene in murine embryos or adult mice, a similar study which used 6.7Kb of bovine PrP sequence consisting of 4.3Kb of upstream sequence, the first exon the first intron and second exon, expressed a Green fluorescent protein (GFP) reporter gene in adult mice (Lemaire-Vieille et al., 2000). Implying that the upstream and intron regulatory regions of the bovine PrP gene can express a reporter gene although it is possible that the expression pattern from the bovine sequence could be due to vital regulatory elements in the 4.3 Kb of bovine upstream sequence which are missing from the 1.2 Kb of upstream sequence in the murine-lacF construct, or the 1.5Kb of upstream sequence in the human constructs. The bovine construct also had a different 3'UTR (SV40) than the lacF constructs which could also affect expression *in vivo*.

Another transgenic model has used the 3.5Kb of upstream and intronic murine sequence in construct Murine-lacF to drive expression of the murine PrP gene in transgenic mice made by micro-injection (G. Telling, personal communication). These mice show a similar pattern of PrP expression to that of normal PrP, suggesting that this region contains sufficient regulatory regions to reproduce the expression pattern of PrP in mice. One difference between the Murine-lacF construct used in this study and the PrP ORF construct driven by the upstream and intron region is that the PrP ORF construct has the 3' UTR of PrP rather than the mouse protamine gene 3' UTR that was used in the LacF constructs. As 3'UTRs play a role in mRNA stability its alteration could affect the levels of translation produced. This could account for the differences in expression between the two types of transgenic mouse.

There are several possible reasons for the failure of the human and murine lacF constructs to drive expression in murine embryos. There may be elements required for expression in the entire organism which have not been included in the murine and full human constructs. These elements may be further upstream or downstream i.e. within the second intron. Such a region in the second intron in mice immediately upstream of the third exon is believed to direct expression of PrP in Purkinje cells (Fischer et al., 1996). However, the bovine PrP regulatory region reporter construct expressed GFP in Purkinje cells in mice (despite lacking the second intron) suggesting that there is no Purkinje expression element in the second intron or that a Purkinje tissue specific

expression element is located in a different region of the bovine PrP gene and this was included in the construct.

Using larger constructs with more upstream elements or the intron might lead to expression which would identify further regulatory elements. Perhaps the neuroblastoma cells used in the testing of the constructs have unusually high levels of expression with a minimal promoter, and other regulatory elements either further upstream or within the second intron are required for expression in normal cells. It has been shown that sequences within the 3'UTR of the ovine PrP gene affect expression levels of a reporter gene in transient transfections in murine neuroblastoma cells so the human 3'UTR may also play a role in expression (Goldmann et al., 1999).

Depending on the location of their insertion into the genome, transgenes can be inactivated if they insert into euchromatic DNA which is transcriptionally inactive or if it inserts near a repressor element. It is therefore possible that the limited number of transgenic embryos and adults created all had the transgene silenced by its position of insertion. The production of further transgenic animals may result in the insertion of the transgene in a region of the genome which allows expression.

The lack of LacZ expression is unlikely to be due to a problem caused by the cloning of the regulatory regions from the CAT reporter constructs to the LacF constructs, such as interference with correct splicing or a mutation of the LacZ gene, as the constructs functioned in murine neuroblastoma cells. Another less likely scenario is that over-expression of the transgene could be toxic for the embryo resulting in selection *in utero* for embryos which do not express the transgene, although the LacZ protein is not toxic to mammals. Alternatively the insertion of large copy numbers of the transgene may have disrupted global gene regulation, perhaps by titrated out an essential transcription factor, which is incompatible with survival. Although it is difficult to then explain the existence of embryos that contained the transgene unless again their genomic position's chromatin structure prevented binding of transcription factors.

Expression may be occurring, but at a level below the limits of detection of this system although if levels of expression are below the level of detection of β-galactosidase then

expression would be so low that it is unlikely that it would represent faithful recapitulation of PrP expression. However the expression levels in transiently transfected murine neuroblastoma was high enough for β-galactosidase staining to be observed *in situ*. Although this could be an artefact of using that particular cell line.

4.3.3 Future work

As the most likely reason for the failure of reporter gene expression is due either to the low numbers of transgenic mice or absence of transcription elements, any future continuation of this work should initially make more transgenic embryos and adults to determine if the regulatory regions are indeed incapable of driving expression. If this fails to lead to expression, new *PRNP* reporter constructs could be made which contain more upstream sequence. Initially 4Kb of sequence should be used, as this amount of the bovine upstream region resulted in expression of a reporter gene in murine embryos. If the transgenic approach is unsuccessful, transient transfection studies in a range of murine and human cell types may determine the effect different regions have on levels of expression. The possibility remains that the mouse protamine 3' UTR may affect expression or mRNA stability *in vivo* but not in neuroblastoma cells, therefore using a different LacZ construct or a different reporter gene system such as luciferase or Green fluorescent protein may lead to a PrP-like expression pattern of the reporter gene.

If a future study shows that the upstream and/or intronic regulatory regions can drive expression of a reporter gene in mice, but not fully reproduce the tissue specific expression pattern of PrP, this regulatory region could be used to create constructs where the PrP protein is driven by the regulatory regions. These constructs could then be used to create transgenic mice by homologous recombination with an NPU PrP hemizygous embryonic cell line and used to create transgenic mice which could then be crossed with null mice to remove the endogenous PrP gene. Such a line of mice with limited PrP expression could be useful in experiments that examine the role of PrP expression in other tissues plays in disease, by showing which tissues require PrP expression to allow disease to develop by infection by different routes.

5 Chapter Five: Sequencing and analysis of *PRNP* regulatory regions.

5.1 Introduction

5.1.1 Association study of the *PRNP* regulatory region.

Work in transgenic mice has shown that the level of PrP expression affects the incubation of TSEs (section 1.4.2). Variation in human PrP expression levels may therefore affect susceptibility or incubation period to CJD. This Chapter describes an association study where the 34bp upstream regulatory region and 330bp intronic regulatory region of the *PRNP* gene, identified as regulating expression in Sk-N-Sh human neuroblastoma cells, were sequenced in sCJD, vCJD patients and in control individuals. If any polymorphisms in these regions modulate PrP expression levels or patterns and affect susceptibility to sCJD or vCJD they will be more frequent in CJD patients than controls. Polymorphisms or mutations could affect expression by increasing the binding of transcription factors or abolishing a repressor-binding site. Variation in expression could alter susceptibility to vCJD by affecting the uptake of infectivity in the gut, the replication of infectivity, or the transport of infectivity from the periphery to the CNS.

The age of onset of disease and the duration of disease was also examined to determine if regulatory regions polymorphisms affect disease through altered expression levels. In sCJD, the age of onset of disease could be lower in high expressing individuals according to the spontaneous misfolding hypothesis of sCJD as more PrP is produced the likelihood of mis-folding increases. In vCJD, variation in expression levels could shorten the incubation period by increasing the replication rate of the agent.

5.1.2 Association study theory

Association studies are used to detect susceptibility alleles in complex genetic diseases with multiple genetic and environmental influences, such as schizophrenia and bipolar disorder (Hodge, 1994). An association study is a non-parametric method of genetic analysis that does not specify any model of inheritance. Polymorphic loci close to candidate genes are analysed for any difference in allele frequency between affected and control individuals. If a polymorphic marker is tightly linked to a disease-causing mutation or a susceptibility allele then it will be more common in affected individuals than in controls. In this study, the polymorphism itself is proposed to affect disease. The effect of a susceptibility allele is described by the odds ratio, which measures the increased or decreased likelihood of someone carrying a particular allele developing the disease relative to someone who does not.

The ethnic origin of both affected and control individuals must be the same in an association study to ensure that any differences detected are due to disease rather than differing allele frequencies in different ethnic groups. Differences in allele frequencies in different ethnic groups is termed ethnic stratification and can lead to erroneous associations between alleles and disease(Hamer and Sirota, 2000). For example if the affected individuals in a study come from ethnic group A but the controls from groups A and B, then any allele found exclusively in group A would be more common in affected individuals than controls leading to an erroneous association. Therefore, care must be taken in selecting a control population to ensure no ethnic stratification exists.

In a case-control association study, a mutation that is found at the same frequency in affected individuals and control individuals is probably a polymorphism with no effect on susceptibility. If it is found in both groups but at a significantly higher level in affected individuals it is a susceptibility allele, while if found at significantly low levels in the affected population it may confer resistance to the disease. If it is only found in affected individuals, it may be directly causing the disease. It is unlikely that regulatory region mutations are causing vCJD, as epidemiological and transmission studies in mice suggest that BSE is the cause of vCJD. Regulatory region polymorphisms could

however affect susceptibility to vCJD. In sCJD, it is possible that some cases are familial but caused by a mutation outside of the ORF, familial CJD is diagnosed on the basis of a mutation in the ORF and then family history is examined for evidence of CJD-like diseases. A hypothetical mutation outside of the ORF that causes CJD could therefore be classified as sCJD if there was no strong family history of CJD-like neurological disease.

This study aims to detect polymorphisms that affect susceptibility to vCJD, sCJD, or unique mutations that may cause sCJD. In order to discriminate between a rare susceptibility allele and a disease causing mutation, a large number of control individuals must be sequenced to determine if the allele is found in the general population at low levels. The size of the control sample in this study was large enough to be 95% confident of detecting an allele present in the population at 1% (the definition of a polymorphism by convention) This is calculated by the formula:

Log(1-P)/Log(1-F)=n

Where P is the desired probability of finding a variant allele, F is the frequency of the variant allele in the control population, n is the number of alleles that need to be sequenced (n/2 individuals). Therefore, to be 95% certain of finding at least one copy of an allele this is present at 1% in the general population, 300 alleles (150 individuals) must be sequenced.

5.1.3 The study population.

The individuals sequenced in this study consisted of 22 vCJD patients, 29 sCJD and 149 control individuals. In this study sCJD and vCJD are considered as separate diseases as they have different aetiologies. Therefore, any alterations in expression may have different effects on susceptibility to the two diseases. The regulatory regions were also sequenced in control individuals to determine the presence and frequency of polymorphisms in the general population. The control population consisted of 28 individuals who were suspected of having some form of CJD but were later shown to

have a different neurological disease. As these individuals came from hospitals across the UK as did the CJD patients there is no danger of ethnic stratification. As there were only 28 such "Hospital" controls the rest of the control group came from random blood samples donated to the Blood Transfusion Service, 62 from Belfast, Northern Ireland and 59 from Edinburgh, Scotland. The *PRNP* ORF of all individuals in the study was sequenced in the CJD Surveillance Unit to ensure there were no mutations and to determine the status at codon 129.

There is a potential risk of ethnic stratification between the Edinburgh and Belfast bloods, which come from specific regions of the UK, and the Hospital controls and the vCJD and sCJD cases, which come from across the UK and there can be genetic variation between individuals from different regions of the UK (Mastana and Sokol, 1998). There is also the possibility that variation in susceptibility to CJD could have an ethnic basis. This is more than a theoretical risk as there is evidence of a difference in the incidence of vCJD in different regions of the UK. Analysis of the 51 vCJD cases in Great Britain up to 1999 has shown the incidence of vCJD in Northern Britain is significantly higher that of Southern Britain (CJD Surveillance unit 1999 Annual Report). In Northern Britain (Scotland, Yorkshire and Humberside, Northwest and the Northern region.) there were 2.57 cases per million people aged 16-54 in the 1991 census compared with 1.30 per million in Southern Britain (East and West Midlands, Southeast, Southwest, East Anglia and Wales). The reason for this difference is unknown, but it could reflect regional differences in genetic susceptibility to vCJD caused by differing ethnic composition of differing regions, such as the proportion of Celtic ancestry in inhabitants of different regions. Alternatively, differing regional dietary habits, such as increased consumption of sausages and meat pies in Northern Britain could affect exposure to BSE although both the total number of cases of BSE and its relative incidence were significantly higher in Southern Britain than in Northern Britain (Stevenson et al., 2000). The level of sCJD is lower in Northern Ireland than in the rest of the UK, but this is not statistically significant.

The three sub-groups of the control group will therefore have to be compared with each other to determine if there are any significant differences in polymorphism frequencies

between them. Such differences might have no effect on CJD susceptibility but simply reflect the Celtic origin of the Edinburgh and Belfast control groups that make up the majority of the controls. Alternatively any differences could reflect differing regional genetic susceptibility to vCJD and may account for the increased vCJD levels in Northern Britain. Comparing allele frequencies between the hospital controls and the Edinburgh and Belfast controls should determine whether any differences in polymorphic allele frequencies have a regional basis.

The regulatory region examined in this study is c.12Kb upstream of the open reading frame that contains the codon 129 Methionine/Valine polymorphism. This polymorphism affects susceptibility to sCJD with methionine homozygotes (MM) being more likely to develop CJD than valine homozygotes (VV) and heterozygotes (MV) are less likely to develop sCJD. In vCJD, all individuals to date have been methionine homozygotes. Due to the relatively short distance between the regulatory region and the ORF, significant linkage dis-equilibrium may occur between a promoter polymorphism and codon 129, which could give an erroneous association. For example, if a polymorphism occurred which has no effect on disease, but was linked to the methionine allele, then a higher frequency would be observed in vCJD patients (100% methionine allele frequency) than in a control population (c.60%). Similarly among sCJD patients, there is a higher frequency of methionine alleles than in the normal population. Such linkage disequilibrium would have a similar effect to ethnic stratification discussed previously, as the affected group has been effectively selected for one polymorphism while the control group has not. This could affect the frequencies of linked polymorphisms. Linkage disequilibrium can be estimated by examining the genotypes to determine if the alleles segregated independently. If alleles do not segregate independently, the frequency of regulatory region polymorphisms in control and affected individuals can be compared within individuals of genotypes MM, MV and VV to determine if any differences in allele frequency between controls and patients is due to linkage disequilibrium.

5.2 Results

5.2.1 Sequencing

Based on the results of the deletion assays (Chapter 3) the sequences surrounding the regulatory regions between –43 and –9 and between +292 and +622 were chosen to be sequenced as potential regulatory regions. In order to prevent bias in reading sequencing gels all individuals were coded before sequencing and the code broken when sequencing was completed.

Two PCR reactions were performed on genomic DNA (Figure 5-1 and 5-2). The first amplified a 448bp region between –310 and +139 using oligos Up3.f and Up3.r (Figure 5-1) this region included the upstream regulatory region between –43 and –9, the PCR product was excised from an agarose gel and purified using a Qiaquick gel extraction kit (Qiagen) (Figure 5-1). The second PCR reaction amplified the 577bp region between +109 and +684 using oligos Int1.f and Int1.r (Figure 5-1), this region includes the 330bp intronic regulatory region between +292 and +622. The PCR product was purified using Microcon filtration columns. The purified PCR products were sequenced using oligonucleotides UpSeq.f or IntSeq.f which overlapped the forward primer of each PCR reaction and fluorescent oligonucleotides and TAQ from the "Big Dye" Sequencing Kit (Applied Biosystems) in a cycle sequencing reaction. The fluorescent labelled sequence was then isopropanol precipitated, resuspended in loading buffer and run on an ABI377 automated sequencer. No PCR product was obtained from some individuals while sequence data was obtained in one region but not another in other individuals (Table 5-1). Of the five different groups the efficiency of the PCR reactions was poorest in Belfast controls, with just over half of the individuals (32/62) being sequenced in both upstream and downstream regions and almost a quarter (15/62) of the DNA samples failed to amplify. This reduced efficiency of the PCR reactions could be due to poor quality DNA from the Belfast controls although the ORF of all Belfast controls was sequenced successfully in the CJD Surveillance Unit.

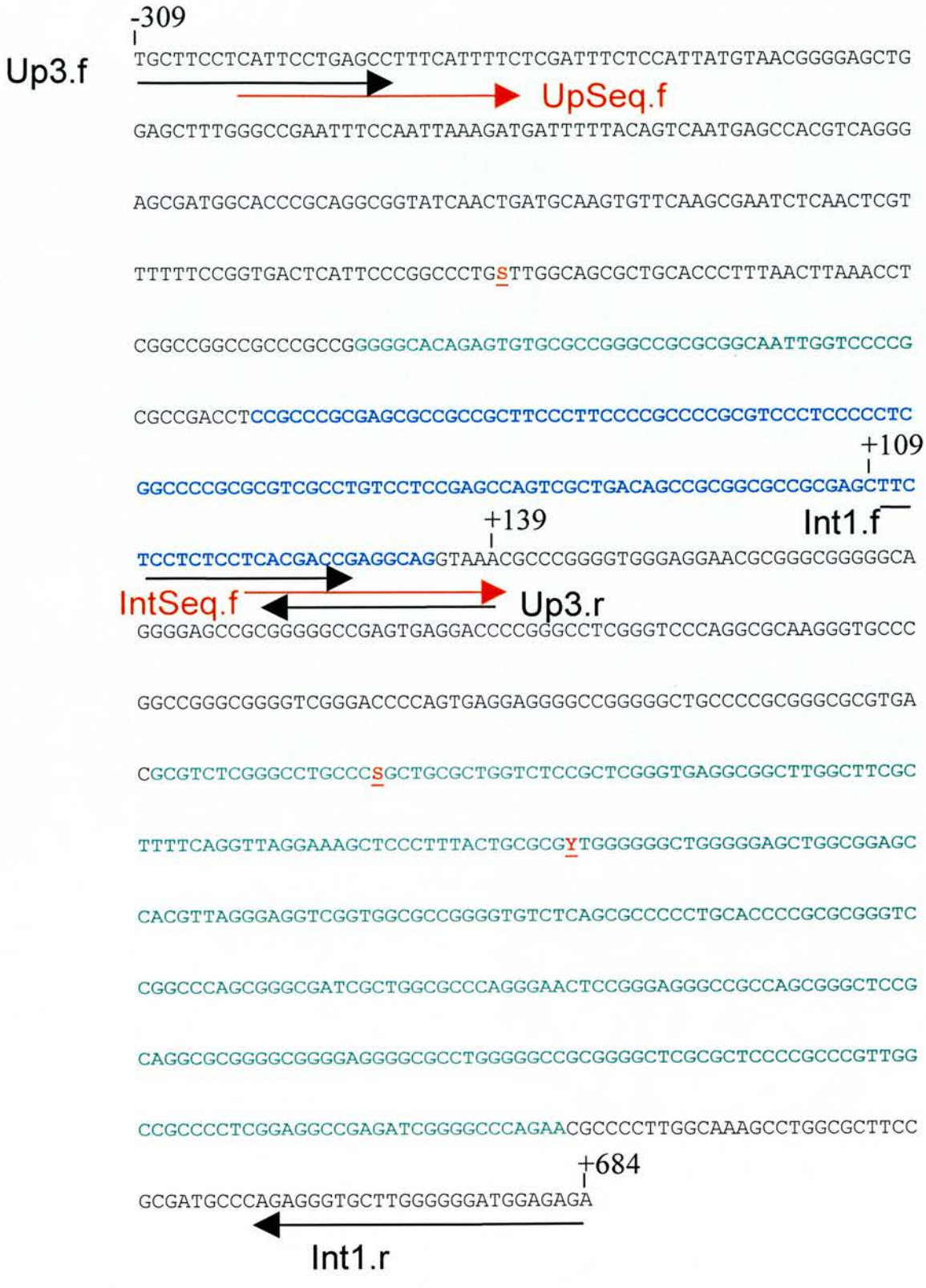


Figure 5-1 PCR primers and sequencing oligonucleotides. First exon in blue, regulatory regions defined by deletions in green, Polymorphisms in red underlined: S=G or C. Y=C or T PCR oligos =black arrows. Sequencing oligos = red arrows. See chapter 2 for details of the PCR reactions.

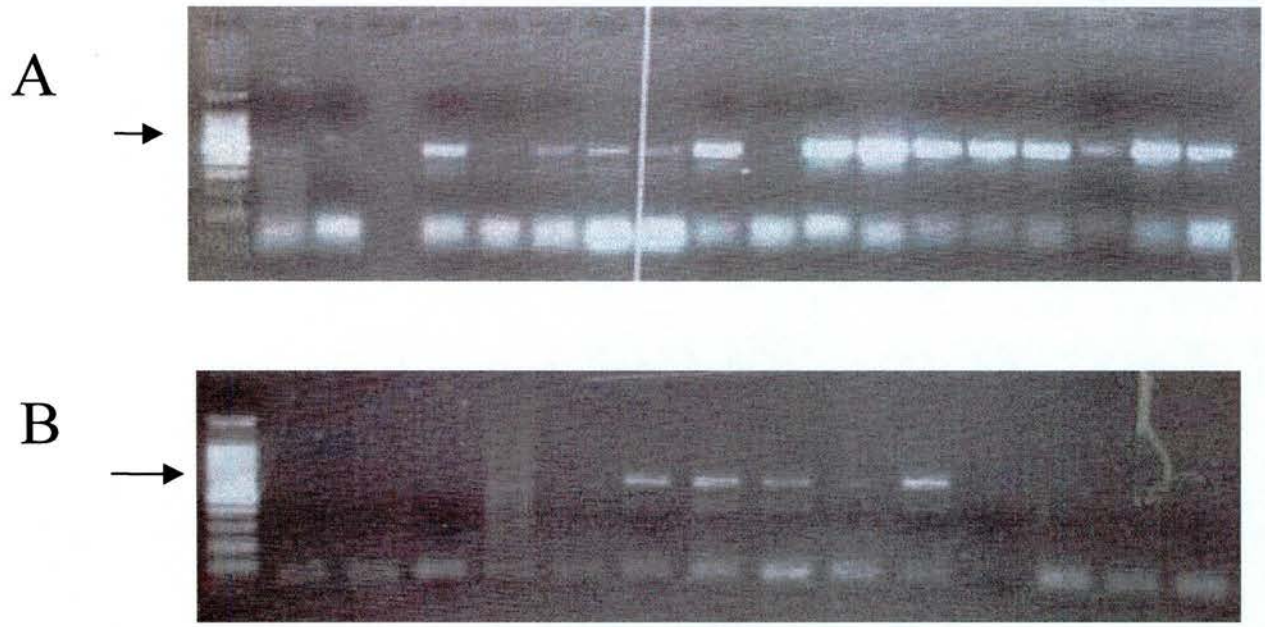


Figure 5-2 PCR amplification of genomic DNA for automated sequencing. Sample PCR amplification of the two PCR reactions which were sequenced. A=448bp Upstream sequence B= 577bp Downstream sequence.

Table 5-1 Efficiency of sequencing

	Sporadic CJD	vCJD	Hospital	Edinburgh	Belfast
Fully sequenced	25	16	23	45	32
Partially sequenced	3	3	1	10	15
Not sequenced	1	3	4	4	15
Total	29	22	28	59	62

5.2.2 Polymorphisms

Three single nucleotide polymorphisms (SNP) were identified (Figure 5-3): in the upstream region a C to G transversion occurred at position –101(Figure 5-4), in the intronic regulatory region a G to C transversion occurred at position +310 (Figure 5-5) and a T to C transition at position +385 (Figure 5-6)(Full data in appendix).

Differences in allele and genotype frequencies of the different polymorphisms, between controls and CJD patients were tested for significance using χ^2 tests or Fisher's exact test. These tests analyse differences in the distribution of numbers in a contingency table from what that expected assuming there is no relationship between the different categories in the table. Chi-squared (χ^2) tests are used on most tables, but this test become inaccurate when the expected values in a table become small and the more computationally demanding Fisher's exact test is then used. The criteria for using Fisher's exact test rather than a χ^2 test is when one of the expected values of a contingency table with four expected values is less than five or two expected values in a larger table are less than five. (Agresti, 1992).

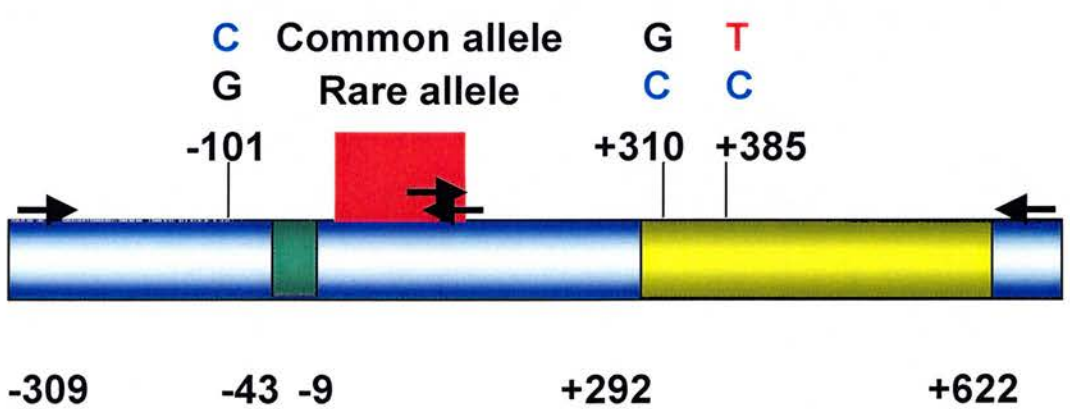
Initially the three sub-groups that that make up the control population were analysed for possible ethnic stratification, which could distort the association study. Statistical analysis using 3X2 contingency table χ^2 tests (Table 5-2) showed no significant

differences between the allele frequencies of the three control groups at codon 129 (p=0.31), position –101 (p=0.36), position +310 (p=0.65) or position +385 (p=0.93). Therefore the three control sub-groups show no evidence of ethnic stratification and can be combined into a single control group.

Table 5-3 Number and frequencies of the four *PRNP* polymorphisms in the three control groups.

Polymorphism	Hospital		Edinburgh		Belfast	
Allele (Common/rare)	Common	Rare	Common	Rare	Common	Rare
Codon 129(M/V)	34(0.63)	20(0.37)	85(0.72)	33(0.28)	79(0.64)	45(0.36)
-101(C/G)	44(0.91)	4(0.09)	82(0.84)	16(0.16)	76(0.88)	10(0.12)
+310(G/C)	42(0.92)	4(0.08)	97(0.95)	5(0.05)	68(0.94)	4(0.06)
+385(T/C)	43(0.93)	3(0.07)	95(0.95	5(0.05)	70(0.95)	4(0.05)

Figure 5-3 Polymorphisms of the *PRNP* regulatory region.



Three Single Nucleotide Polymorphisms were identified. In the control population at position –101, 5' of the upstream regulatory region (green) identified by deletion assay, a G to C transversion occurred. In the intronic region (yellow) two polymorphisms were identified, a G to C transversion at position +310 and a C to T transition at +385

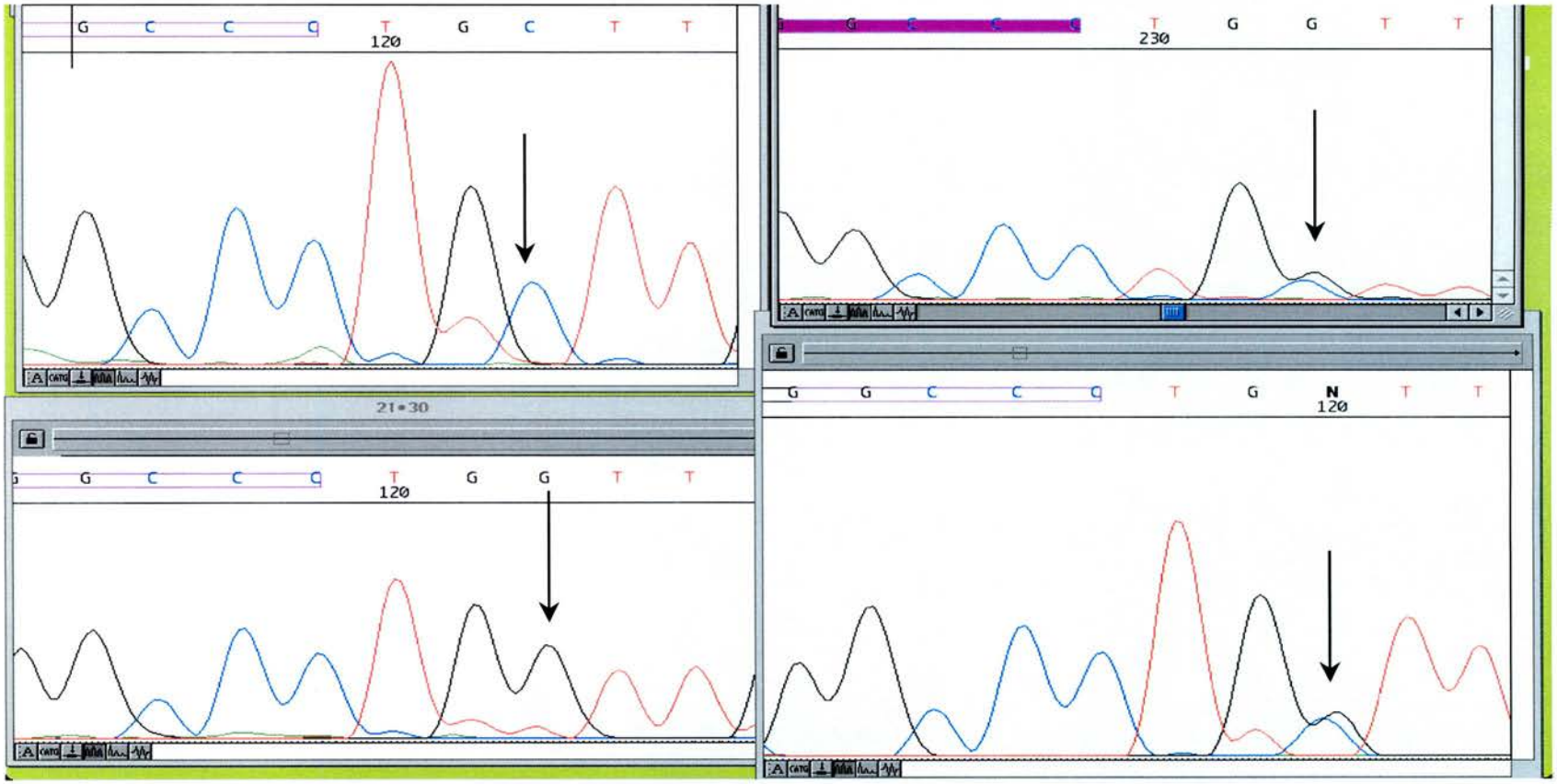


Fig 5-4 Polymorphism C-101G: ABI 377 trace of the C to G transition (Arrow). Upper left panel wild-type CC homozygote. Lower left panel GG homozygote. Right panels heterozygous (CG) individuals

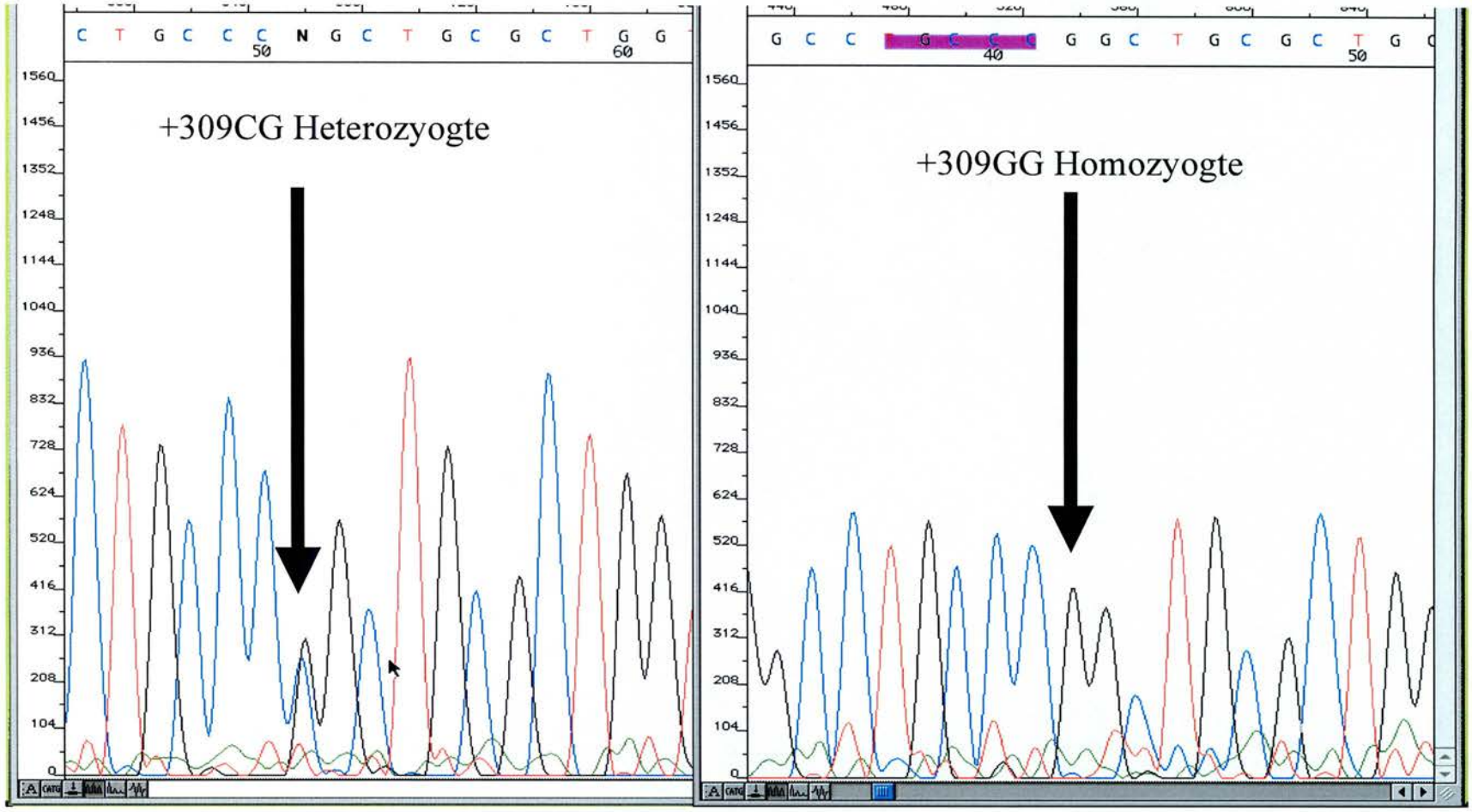


Fig 5-5 Polymorphism G+309C: ABI 377 trace of the G to C transition (Arrow). Left panel wild-type CG homozygote. Right panel GG homozygote. There were no CC homozygotes.

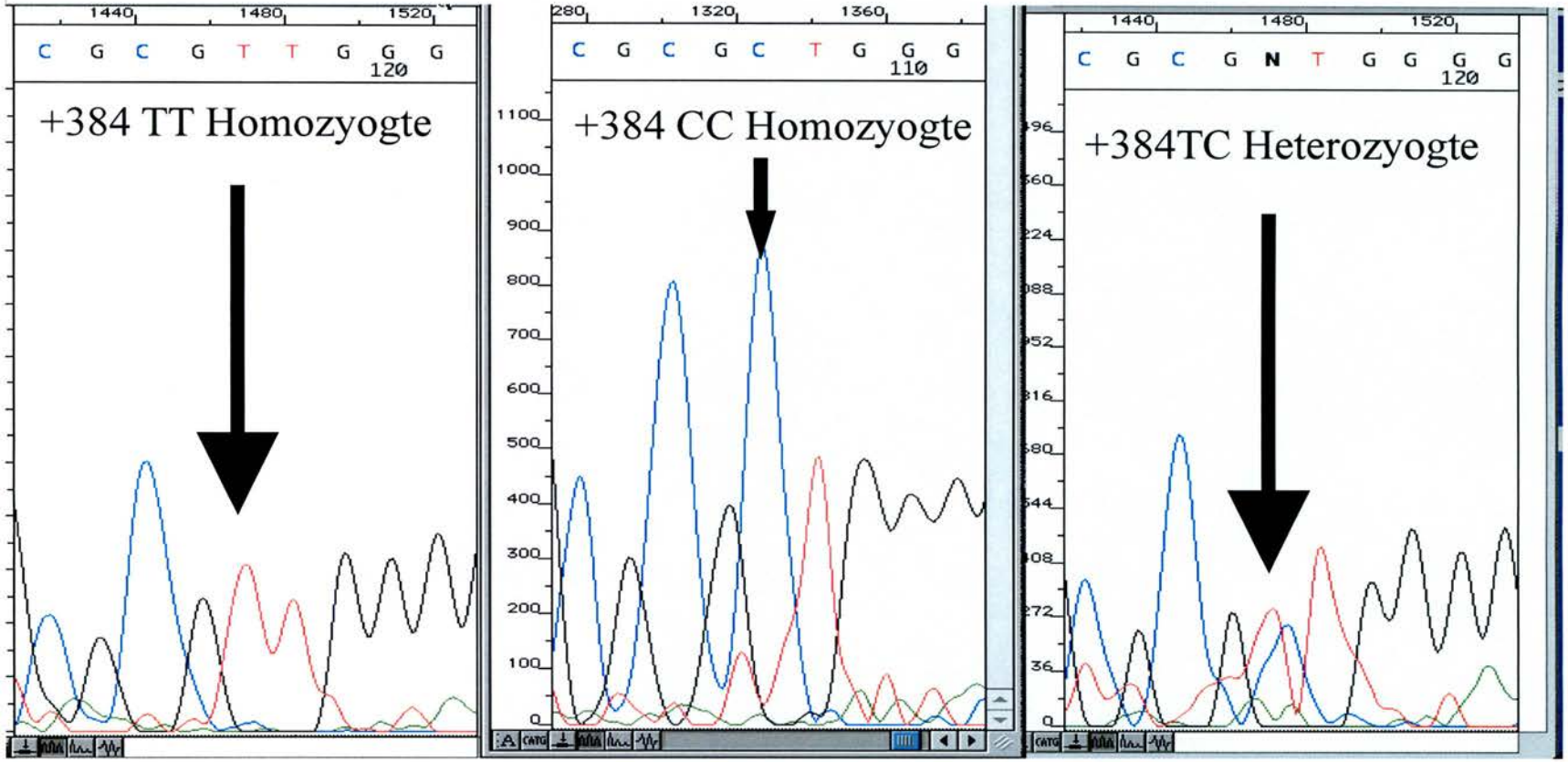


Fig 5-6 Polymorphism T+384C: ABI 377 trace of the T to C transition (Arrow). Left panel wild-type TT homozygote. Centre panel CC homozygote. Right panel a heterozygous (CT) individual.

5.2.3 Analysis of polymorphisms and disease

At position –101, 162 individuals were sequenced: 27 sCJD patients, 19 vCJD patients and 116 controls (49 from Edinburgh, 43 from Belfast and 24 hospital controls). Table 5-3 shows the genotypes and allele frequencies in affected individuals and controls. Analysing the frequency of G alleles between vCJD patients and the combined controls showed no significant difference (p=0.63 χ^2–test). There was a higher allele frequency of the rare G allele in sCJD patients than in controls, but this did not reach statistical significance (p=0.08 χ^2–test). To determine the effect of a polymorphism on an individual's susceptibility to disease, the genotypes of control and affected individuals were compared rather than the allele frequency. There was no significant difference in the distribution of genotypes between vCJD and controls (p=0.56 Fisher's exact test) or between sCJD and the control population (p=0.13 Fisher's exact test). Comparing the effect of carrying at least one –101G allele (i.e. treating CG heterozygotes and GG homozygotes the same) showed that there was a larger proportion of carriers in sCJD patients than in controls (10 of 27 sCJD, 28 of 116 controls). This is not a significant difference (Fisher exact test p=0.131), but there is a trend towards G allele carriers being more likely to develop sCJD with an odds ratio of 1.85.

The G to C polymorphism at +310 was sequenced in 26 sCJD, 16 vCJD and 110 controls (23 hospital controls, 51 Edinburgh and 36 Belfast controls). Despite the high frequency of the C allele in sCJD patients and its absence in vCJD patients (Table 5-4) neither of these differences were statistically significant (χ^2–test p=0.11 and 0.131 respectively). Analysing the difference in the number of carriers versus non-carriers in patients and controls was also non-significant in sCJD and vCJD (Fisher's exact test p=0.122 and 0.155 respectively).

The T to C polymorphism at +385 was sequenced in 26 sCJD 16 vCJD and 110 controls (23 hospital controls, 50 Edinburgh controls and 37 Belfast controls.). There was no significant difference in allele frequency between the control and sCJD groups (p=0.86) (Table 5-5). The frequency of the C allele in vCJD patients (0.156) was almost 3 times

that of the control group (0.054) and this difference just reaches statistical significance (p=0.049 Fisher's exact test). The effect of carrying a +385C allele on susceptibility to vCJD was examined by comparing the proportion of vCJD patients who were carriers with control (4 of 16 vCJD, 11 of 110 controls).This gives an odds ratio of 3.0 but the numbers are not large enough to reach significance (p=0.099).

Preliminary analysis of the three polymorphisms therefore shows that all are found in both disease and control groups suggesting that the polymorphisms do not act as the sole determinants of disease. The +310C allele is not found in vCJD patients possibly suggesting that it may decrease an individual's susceptibility to vCJD but the absence of this allele is not significant.

Table 5-3 C-101G allele. Genotype and allele numbers and frequencies in sCJD, vCJD and controls.

The numbers of alleles and genotypes in sCJD and vCJD patients were compared with those of the control population and deviations from what would be expected from random distribution analysed using Fisher's exact test of a chi-squared test.

	sCJD	vCJD	All Controls
CC	17(0.63)	13(0.68)	88(0.76)
CG	8(0.30)	6(0.32)	26(0.22)
GG	2(0.07)	0	2(0.02)
Fisher's exact test	p=0.13	p=0.56	
C	42(0.78)	32(0.84)	202(0.87)
G	12(0.22)	6(0.16)	30(0.13)
χ^2 test	p=0.08	p=0.63	

Table 5-4 G+310C allele. Genotype and allele numbers and frequencies in sCJD and vCJD and controls.

	sCJD	vCJD	All Controls
GG	20(0.77)	16	97(0.88)
CG	6(0.23)	0	13(0.12)
CC	0	0	0
Fisher's exact test	p=0.12	p=0.16	
G	46(0.88)	32	207(0.94)
C	6(0.12)	0	13(0.06)
χ^2 test	p=0.11	p=0.13	

Table 5-5 T+385C Genotype and allele numbers and frequencies in affecteds and controls.

	sCJD	vCJD	All Controls
TT	23(0.88)	12(0.75)	99(0.90)
TC	3(0.12)	3(0.19)	10(0.09)
CC	0	1(0.06)	1(0.01)
Fisher's exact test	p=0.83	p=0.13	
T	47(0.94)	27(0.84)	208(0.95)
C	3(0.06)	5(0.16)	12(0.05)
Fisher's exact test	p=0.86	p=0.049	

5.2.4 Linkage with codon 129

To determine if linkage between the regulatory region polymorphisms and codon 129 exists which would affect the frequencies of the polymorphisms in sCJD, vCJD and the control population, each polymorphism was examined for linkage disequilibrium.

C-101G: In order to determine if the -101G allele is in linkage dis-equilibrium with

either methionine or valine at codon 129 in the ORF, the –101/codon 129 haplotype was inferred from the genotypes of the control population. CJD patients were excluded from this analysis as they have been selected for methionine homozygosity and may also have been selected for particular regulatory region alleles and this could lead to an erroneous assumption of linkage to codon 129 methionine. The haplotype can be inferred in all individuals except those heterozygous at both positions (Table 5-6). A Fisher's exact test for independent assortment of haplotypes shows that the distribution is not random (p=0.005) and all 20 G alleles whose linkage can be inferred are linked to methionine at codon 129.

This method of analysis ignores double heterozygotes, as their haplotype cannot be deduced. However, the data from double heterozygotes can be used to estimate the frequency of the four haplotypes. An Estimation-Maximisation (EM) algorithm has been devised (Weir, 1990) which uses reiteration to find the haplotype frequencies which best fit the data. Firstly, an estimate is made of one haplotype frequency (in this example the frequency of –101C codon 129 methionine haplotype=MC) which allows the other three haplotype frequencies to be calculated.

Step 1. Estimate of MC haplotype frequency = p' MC (p' is an estimated frequency)

The other haplotypes frequencies can then be calculated from the allele frequencies of codon 129 Met and –101C in table 5-6

Step 2. p' MG= pM - p'MC (pM is the known frequency of the M allele.)

p' VC= pC - p'MC

p'VG=1- pM- pC + p'MC

The haplotype frequencies can then be used to calculate the proportion of double heterozygotes who have the two common alleles *in cis* (on the same chromosome i.e. MC/VG) rather than *in trans* (on different chromosomes i.e. MG/VC) assuming random mating.

Step 3 Frequency of in cis double heterozygotes = 2 X p'MC X p'VG

Frequency of in trans double heterozygotes = 2 X p'MG X p'VC

Proportion of double hets who are in cis = (2 X p'MC X p'VG)/((2 X p'MC X p'VG)+ (2 X p'MC X p'VG))=K'

A new estimate of the original estimated haplotype (p''MC) can then be calculated using the frequencies of MC haplotypes, which can be calculated from Table 5-6, plus the estimate of the number of double heterozygotes carrying that haplotype. The frequency of MC haplotypes is equal to twice the number of MMCC individuals (2 MC chromosomes) plus the number of MVCC, MMCG and in cis MVCG (double heterozygotes) who each have one MC haplotype, divided by the total number of haplotypes.

Step 4 p''MC = ((2 X pMMCC) +pMVCC +pMMCG+(K' X pMVCG))/total number of haplotypes.

pMMCC is the frequency of individuals of genotype MM at codon 129 and CC at -101, which can be calculated from Table 5-6.

The new estimate of the frequency of MC haplotypes p''MC is then used in step 2 to calculate the other haplotype frequencies and a new *in cis* homozygote proportion K'' in step 3 and a new estimate of MC frequency (p'''MC) in step 4. After several cycles, the estimates of MC frequency converge until eventually the new estimate equals the previous estimate. This gives the haplotype frequencies which are most likely to give the genotype combinations seen in the original table. The choice of initial MC haplotype does not affect the final frequency, as the same final haplotype frequencies are found using several different initial estimates.

Applying this method to –101 and codon 129 gives estimated haplotype frequencies which support linkage disequilibrium between –101G and 129C (MC haplotype frequency=0.586, MG=0.129, VC=0.284, VG=0.0000021). A larger study may find individuals who have VG alleles but this study indicates that this haplotype is absent or

extremely rare. This could reflect a relatively recent origin of the –101G mutation which occurred on a codon 129 methionine allele and not enough time has occurred to allow recombination. Examination of polymorphic markers upstream and downstream of *PRNP* could determine if all –101G alleles share a common haplotype, indicating a single mutation.

To control for the linkage of –101G to codon 129M the study population was sub-divided by codon 129 genotype. As all vCJD individuals and 22 of 29 sCJD patients are methionine homozygotes, they were compared with the 61 MM controls. Despite the high proportion of MM sCJD patients that carry at least one G allele compared with MM controls (resulting in a relative risk of 2.4), this difference did not reach statistical significance (p=0.095) (table 5-7). There were no MV and VV sCJD patients who carried -101G but the numbers were too small for this absence to be significant.

Table 5-6 Linkage disequilibrium between Codon 129 and –101

All individuals sequenced at –101G and codon 129 were classed according to their genotype combination. The genotypes were inferred in all individuals except the 10 double heterozygotes e.g. the 43 MM CC individuals have 86 M-C chromosomes, the 34 MV CC individuals have 34 M-C and 34 V-C chromosomes etc. All the chromosomes deduced in this manner were counted and random assortment tested in a chi-squared test.

n=116	Codon 129 Genotype		
-101 Genotype	MM	MV	VV
CC	43	34	11
CG	16	10	0
GG	2	0	0
Haplotypes	M	V	Total
C	136	56	192
G	20	0	20
Total	156	56	212

Chi-squared test p=0.005 Linkage disequilibrium between -101G and 129M

Table 5-7 Effect of –101G on sCJD and vCJD susceptibility in codon 129 methionine homozygotes

	vCJD	sCJD	Controls
-101G carriers	6	10	18
Non-carriers	13	10	43
Odds ratio and p value relative to controls.	OR=1.1 p=0.86	OR=2.4 p=0.095	

G+310C: In order to determine if the +310C allele is in linkage disequilibrium with the codon 129 polymorphism the haplotype was inferred from the genotypes of the control population. The C allele was found linked to both methionine and valine and although there was a higher frequency on V alleles there was no significant linkage disequilibrium (p=0.369) (Table 5-8). The EM Algorithm estimates that the haplotype frequencies are not what would be expected under independent assortment and show partial linkage disequilibrium between codon 129 valine and +310C (Table 5-9). If these haplotype frequencies are applied to the 220 chromosomes in table 5-8 the difference reaches significance (p=0.02). Therefore partial linkage disequilibrium may exist and has to be accounted for when analysing the effect of the polymorphism on disease.

Table 5-8 Linkage disequilibrium between Codon 129MV and +310 GC

n=110	Codon 129 Genotype		
+310 Genotype	MM	MV	VV
GG	53	36	8
CG	4	6	3
CC	0	0	0
Haplotypes	M	V	Total
G	146	55	201
C	4	3	7
Total	150	58	208

Fisher's exact test. p=0.304

Table 5-9 Estimated +310/codon 129 haplotype frequencies compared with the expected values under linkage equilibrium.

Haplotype	Expected	Algorithm Estimate	Difference
MG	0.667	0.686	+0.019
MC	0.042	0.023	-0.019
VG	0.274	0.255	-0.019
VC	0.017	0.036	+0.019

To control for linkage disequilibrium, the presence of +310C alleles was examined in controls, sCJD and vCJD patients of all three codon 129 genotypes (Table 5-10). Analysing MM individuals show no significant difference between the controls and either vCJD or sCJD. The numbers of sCJD patients of MV and VV genotypes are small and although statistical significance is not reached, there is a trend towards VV +310C carriers being more susceptible to sCJD and MV +310C carriers being extremely susceptible to sCJD.

Table 5-10 The effect of the +310C allele on sCJD and vCJD susceptibility by codon genotype 129.

Codon 129	MM			MV		VV	
Status	vCJD	sCJD	Controls	sCJD	Controls	sCJD	Controls
Carriers	0	2	4	2	6	2	3
Non-carriers	16	17	53	1	36	2	8
OR, p value	OR=0 p=0.62	OR=0.64 p=0.62		OR=12.0 p=0.08		OR=2.7 p=0.41	

T+385C: Analysing control haplotypes for linkage of the +385C allele to codon 129 shows that all of the 7 +385C alleles for which linkage can be determined in the control population are linked to Methionine alleles at codon 129 (Table 5-11). This does not reach significance (p=0.093) but the EM algorithm estimates the VC haplotype frequency at 3.6 X 10^{-7} which supports linkage disequilibrium between +385C and codon 129M. Although the rarity of the +385C allele means that a larger sample is

needed to confirm that no VC haplotypes exist.

Table 5-11 Linkage disequilibrium between Codon 129 and +385

n=110	Codon 129 Genotype		
+385 Genotype	MM	MV	VV
TT	51	37	11
TC	5	5	0
CC	1	0	0

Haplotypes	M	V	Total
T	144	59	203
C	7	0	7
Total	151	59	210

Fisher's exact test. p=0.095 +385C approaching linkage disequilibrium to 129 Met.

To control for linkage disequilibrium, the presence of +385C alleles was examined in controls, sCJD and vCJD of all three codon 129 genotypes (Table 5-12). In sCJD there were no significant differences between MM sCJD and controls, and while no MV or VV sCJD patients carried the allele this is not a significant difference. Analysing the numbers of vCJD patients carrying at least one copy of the +385C allele with MM controls shows that there is no statistically significant difference, although there is a trend towards it increasing susceptibility to vCJD (Odds ratio=2.8).

Table 5-12 Effect of +385C allele on sCJD and vCJD susceptibility by codon 129 genotype.

Codon 129	MM			MV		VV	
Status	vCJD	sCJD	Controls	sCJD	Controls	sCJD	Controls
Carriers	4	3	6	0	5	0	0
Non-carriers	12	16	51	3	37	4	11
OR and p value	OR=2.8 p=0.218	OR=1.6 p=0.54		OR=0 p=0.52		NA	

Linkage disequilibrium analysis of the three polymorphisms therefore shows that the -101G allele and the +385C allele are in linkage disequilibrium with codon 129M. The +310C allele is found linked to both V and M at codon 129 but is predominantly linked to V. Re-analysing the results taking linkage disequilibrium into consideration shows that –101G increases the risk of MM individuals to sCJD, +310C increases the risk of sCJD in codon 129 heterozygotes and that the +385C allele is more common in vCJD patients than MM controls. However none of these results reach statistical significance as all three alleles are relatively rare and do not have very large effects on disease susceptibility to reach highly significant results in a case-control study of this size.

As the three regulatory region polymorphisms are very close to each other, there may be considerable linkage disequilibrium between them as recombination will occur very infrequently. Analysing each pair of the three regulatory polymorphism shows that the haplotype distribution of –101G and +310C (Table 5-13) is not significantly different from random distribution even if the single double heterozygote is considered as a CG/GC or as a CC/GG. Similarly, despite both being linked to codon 129M, the -101G and +385C alleles do not appear to be linked to each other (Table 5-14). However, the two individuals heterozygous at positions –101 and +385 are also heterozygous at codon 129. This means that either –101G and +385C are linked together on the M allele of these individuals or that one of the polymorphisms is linked to V at codon 129. These two alternatives can only be resolved by cloning and sequencing both alleles from both individuals. There was no linkage disequilibrium between +310 and +385, all the haplotypes could be deduced as there were no double heterozygous individuals (Table 5-15)

Table 5-13 linkage disequilibrium between -101 and +310 in control population

n=99	-101 Genotype		
+310 Genotype	CC	CG	GG
GG	65	21	2
GC	10	1	0
CC	0	0	0
Haplotypes	C	G	Total
G	161	25	186
C	10	0	10
Total	171	25	196

Fisher's exact test. p=0.367 No linkage disequilibrium

Table 5-14 Linkage disequilibrium between –101CG and +385TC

n=100	-101 Genotype		
+385 Genotype	CC	CG	GG
TT	68	20	2
TC	8	2	0
CC	0	0	0
Haplotypes	C	G	Total
T	164	24	188
C	8	0	8
Total	172	24	196

Fisher's exact test. p=0.599 No linkage disequilibrium.

Table 5-15 Linkage disequilibrium between +310 and +385

n=109	+310 Genotype		
+385 Genotype	GG	CG	CC
TT	85	13	0
TC	10	0	0
CC	1	0	0
Haplotypes	G	C	Total
T	193	13	206
C	12	0	12
Total	205	13	218

Fisher's exact test. p=0.369 No linkage disequilibrium.

5.2.5 Could more than one regulatory region polymorphism have an effect?

The previous analysis of each polymorphism implicitly assumed that the other two polymorphisms had no effect on susceptibility, as the number of carriers of a given allele was compared with the number of non-carriers in patients and controls. This method of analysis excludes the possibility that two or possibly all three polymorphisms could each have an effect on susceptibility. To determine if more than one allele has an effect on susceptibility, the haplotype (the status of each regulatory region polymorphic site on each chromosome) of each individual was examined to determine if the regulatory region polymorphisms may act co-operatively. This can only be done in the 141 individuals who have been sequenced at all three polymorphic positions, (100 controls 25 sCJD and 16 vCJD). The "wild-type" haplotype CGT (-101:+310:+385) is the most frequent the others being **G**GT, C**C**T and CG**C** (rare allele in bold) each having one of the rare alleles. Two individuals are heterozygous at positions -101 and +310 and two heterozygous at positions -101 and +385. Due to the fact that a PCR product was

sequenced it is impossible to determine if these individuals have one rare allele on each chromosome or two rare alleles on one chromosome and a wild-type haplotype on the other i.e. a genotype of **G**GT/C**C**T or CGT/**GC**T for the two individuals heterozygous at the first two polymorphic sites.

Analysing the 141 individuals sequenced at all three positions (Table 5-16) showed that individuals who carried any of the rare alleles -101G, +310C or +385C are significantly more likely to develop sCJD than those who did not (odds ratio=2.93 p=0.02 χ^2-test). As only one sporadic and three control individuals carry rare alleles at two loci this increased risk is not due to an additive effect of the rare alleles. It may be due to two or more alleles each independently increasing susceptibility to sCJD. In this model, the effect of a polymorphism can only be analysed by comparing carriers to individuals who are homozygous for the common allele at all three positions. This is because comparing carriers of an allele to individuals who lack that allele but may be carrying susceptibility alleles at another locus could mask the effect of both alleles. To analyse the effect of each position, the 141 individuals sequenced at all three positions were sub-divided by any rare allele they possessed and by codon 129 genotype as this affects the frequency of all three rare alleles (Tables 5-17, 5-18, 5-19).

Table 5-16 Effect of carrying at least one rare allele on vCJD and sCJD susceptibility.

	vCJD	sCJD	Control
Carry at least one rare allele	9	17	42
Common allele at all three	7	8	58
Odds ratio and probability	OR=1.78	OR=2.93 p=0.02	

Table 5-17 The effect of –101G on susceptibility to vCJD and sCJD.

Individuals homozygous for –101C but carrying other regulatory region polymorphisms have been excluded, leaving only –101G carriers and wild-type individuals. The vCJD and sCJD patients were divided by codon 129 genotype and compared with controls of the same genotype. Differences in the proportions of carriers and non-carriers were used to calculate the odds ratio and the probability of such different proportions occurring by chance. a=includes two homozygotes, b=includes one individual heterozygous for +310C, c=includes two individuals heterozygous for +385C. NA=not applicable as no patients or controls carry a –101G allele.

	vCJD	sCJD				Controls		
Codon 129	MM	All	MM	MV	VV	All	M	MV	VV
-101G carrier	5	9	9[ab]	0	0	24	14[a]	10[bc]	0
Wild type	7	8	5	1	2	58	30	21	7
Total	12	17	14	1	2	82	44	31	7
Odds ratio	1.5	2.7	3.86	0	NA				
Probability	0.5	0.06	0.03	0.5	1.0				

Table 5-18 The effect of the +310C allele on susceptibility to vCJD and sCJD.

Individual of genotype +310GG carrying other regulatory region polymorphisms have been excluded, b=includes one individual heterozygous for –101G

	vCJD	sCJD				Controls			
Codon 129	MM	All	MM	MV	VV	All	MM	MV	VV
+310C carrier	0	6	2[b]	2	2	11	3	5[b]	3
Wild type	7	8	5	1	2	58	30	21	7
Total	7	14	7	3	4	69	33	26	10
Odds ratio	0	3.95	4.00	8.40	2.33				
Probability	0.41	0.02	0.16	0.07	0.48				

Table 5-19 The effect of +385C on susceptibility to vCJD and sCJD.

Individuals of genotype +385TT carrying other regulatory region polymorphisms have been excluded, a=Includes one homozygote, c=includes two individuals heterozygous for –101G. NA=not applicable as no patients or controls carry a +385C allele.

	vCJD	sCJD				Controls			
Codon 129	MM	All	MM	MV	VV	All	M	MV	VV
+385C carrier	4[a]	3	3	0	0	10	5	5[c]	0
Wild type	7	8	5	1	2	58	30	21	7
Total	11	11	8	1	2	68	35	26	7
Odds ratio	3.43	2.18	3.60	0	NA				
Probability	0.11	0.30	0.13	0.63	0.99				

Comparing the number of carriers of -101G and wild-type individuals in vCJD and MM controls showed no significant differences. Similarly, there were no significant differences when the effect of +310C was considered in vCJD patients. There was a trend towards +385C carriers having an increased risk of vCJD but this did not reach statistical significance.

Comparing sCJD and controls showed an excess of carriers of -101G relative to wild-type controls which approached statistical significance. When the linkage of -101G to codon 129 was considered and MM homozygotes were examined the effect of -101G was statistically significant and had a higher relative risk. No MV or VV sCJD individuals carried any -101G alleles, but this was not significant. The number of +310C carriers was significantly higher in sCJD patients than controls when compared with the number of wild-type individuals. Sub-dividing by codon 129 genotype showed no statistically significant differences between sCJD patients and controls. However, this lack of significance is due to the reduction in numbers and is not caused by linkage to codon 129, as the odds ratio was unchanged in MM individuals, greatly increased in MV individuals and decreased in VV individuals. The +385C allele was found in more sCJD patients than controls but this was not significant. Stratifying by codon 129 showed that there was an increased relative risk in MM sCJD (RR=3.43) but this difference was not

significant (p=0.1).

To summarise +310C and –101G independently increase an individual's susceptibility to sCJD. The +310C allele significantly increases susceptibility in all individuals but the effect was greatest in codon 129 heterozygotes while -101G increases susceptibility in methionine homozygotes. The +385C allele did not reach statistical significance as a susceptibility factor although there is a trend towards it being a susceptibility factor to both sCJD and vCJD in methionine homozygotes. One possible problem in this method of analysis is that the wild-type individuals used to compare with carriers are the same for each allele, therefore any sampling error of the proportions of healthy or CJD wild-type individuals will affect all the results.

5.2.6 Effect of polymorphisms on age of onset and duration of disease

In order to determine if the polymorphisms affect clinical parameters of disease the age of onset of the disease and the duration of disease was examined in vCJD and sCJD patients. Alterations in expression patterns or levels could influence an individual's susceptibility to infection or progression of disease once the disease has become established.

The age of onset of disease and duration of disease of the 16 vCJD patients sequenced at all three regulatory region polymorphic loci are shown in Figure 5-7. Neither -101G nor +385C had any significant effect on disease although the individual with the longest duration of disease was homozygous for +385C. No vCJD patient carries the +310C allele.

The age of onset of disease and duration of disease of the 25 sCJD patients sequenced at all three regulatory region polymorphic loci are plotted in Figure 5-8. Analysing the age of onset of disease showed a trend towards +385C carriers having a younger age of onset than non-carriers (58.3 years and 66.4 years respectively) but this is not significant (t-test p=0.19). Previous studies have that 18% of VV sCJD cases occur in people under

the age of 50 compared with 9% of MV cases and 4% of MM cases (Alperovitch et al., 1999). To determine if the regulatory region polymorphisms have an effect on disease when divided by codon 129 genotype the population was split into the 18 MM, 4 VV and 3MV individuals. Among MM homozygotes +385C carriers had a younger age of onset (58.3 years) than non-carriers (67.4) but this was not significant (t-test p=0.16). There was no effect seen in +310C or –101G carriers.

Three sCJD patients were atypical in being quite young and had very long duration of illness (The three individuals on the extreme left of figure 5-8). Two of these individuals were the only sCJD patients who were heterozygous at +310 and codon 129. This may suggest an interaction between the two polymorphisms, as the other +310C carriers are unremarkable as is the other codon 129 heterozygote. The other atypical individual was the youngest sCJD patient who was a methionine homozygote and homozygous for the common allele at the regulatory region polymorphisms. The patient was suspected of being a vCJD case on clinical grounds but neuropathological examination showed the features of sCJD. The re-classification of this individual removed a potential significant association between +385C and a younger age-of-onset in MM sCJD patients as if that individual is excluded the average age-of-onset of MM CJD patients who do not carry +385C rises to 69.4 years and the p value when the ages are compared to the +385C carriers is 0.016. However this is based on a very small number of carriers (3) and a larger study is required to determine if +385C affects the age-of-onset of sCJD.

Despite there being no statistically significant effect on clinical parameters in this study of sCJD and vCJD cases, the relationship between the three regulatory region polymorphisms and codon 129 are so complex with a large number of possible interactions that a very large study of 200 sCJD patients will be needed to detect any significant effect. Other phenotypic differences among sCJD patients that could be examined to determine if they are be influenced by the regulatory region polymorphisms, include neuropathology, PrP^{res} isotypes and specific clinical features such as the presence or absence of abnormal EEG patterns. These could also be examined in a larger study.

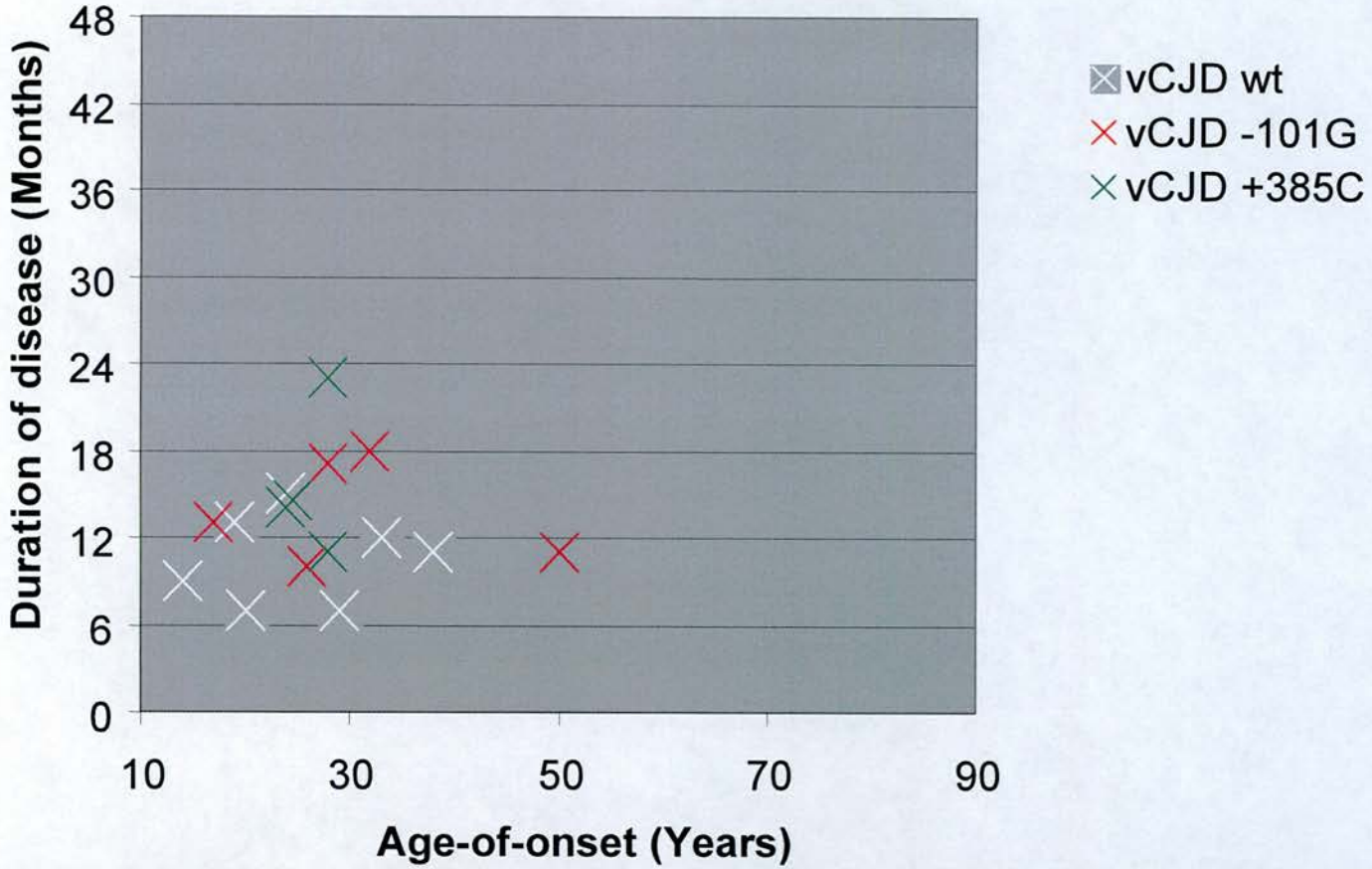


Figure 5-7 Effect of codon 129 and regulatory region polymorphisms on the age-of-onset and duration of vCJD.

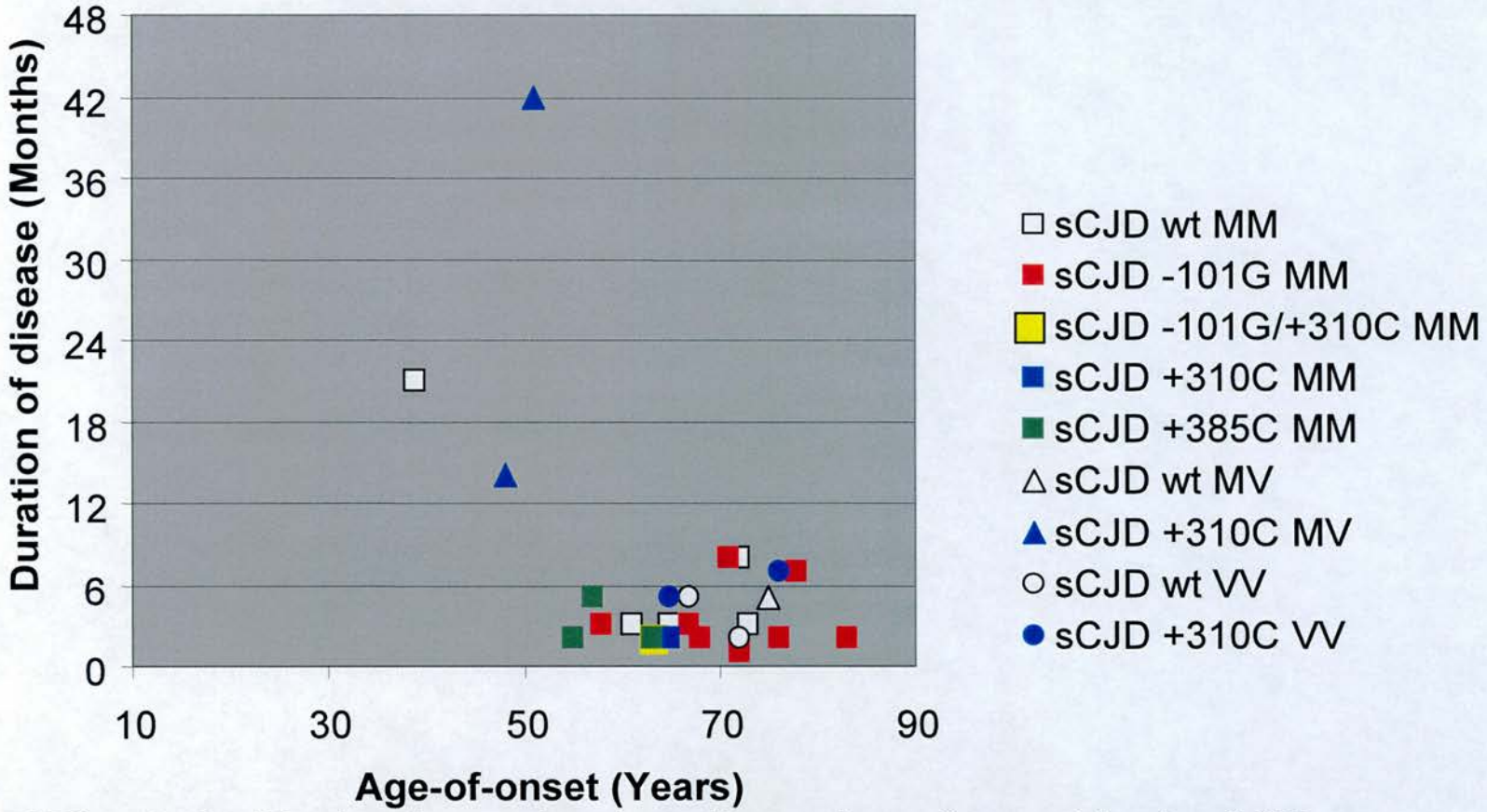


Figure 5-8 Effect of codon 129 and regulatory region polymorphisms on the age-of-onset and duration of sCJD .

5.3 Discussion

5.3.1 The effect of Polymorphisms on disease.

Three polymorphisms were identified within or adjacent to the regulatory regions of the *PRNP* as defined by deletion assay. Analysis of the presence of polymorphisms in sCJD, vCJD and control individuals has shown that carrying either of two of the polymorphisms (-101G and +310C) increases an individual's susceptibility to sCJD but not to vCJD, while there is a trend towards carriers of the +385C polymorphism being more susceptible to vCJD.

This study has identified methionine homozygous -101G carriers as a sub-group within methionine homozygotes that is more susceptible to sCJD. Previous studies have shown that codon 129 methionine homozygotes are more likely to develop sCJD than valine homozygotes while heterozygotes are least likely to develop sCJD. This results in 70% of sCJD cases occurring in the 40% of the Caucasian population who are MM (Alperovitch et al., 1999). The difference in susceptibility between MM and VV homozygotes is not necessarily due to the differing structures of the codon 129 M and V PrP proteins as valine homozygotes are more susceptible than methionine homozygotes to iatrogenic CJD. This study has now shown that MM -101G carriers (11% of the total Caucasian population) account for half of MM sCJD cases (35% of total sCJD cases). In contrast, MM non -101G carriers (29% of the total population, 35% of sCJD cases) are no more susceptible to sCJD than VV individuals (11% of the population, 16% of sCJD cases)(Figure 5-9). Therefore, the effect of codon 129 in affecting susceptibility to sCJD is influenced by linked regulatory polymorphisms. Perhaps homozygosity at codon 129 is the major susceptibility determinant of sCJD while -101G plays a minor role. The effect of -101G may not be sufficient to alter the susceptibility of heterozygotes and as -101G appears to be linked to methionine, its effect is only apparent in methionine homozygotes. This model would explain why methionine homozygotes are more susceptible to sCJD than valine homozygotes.

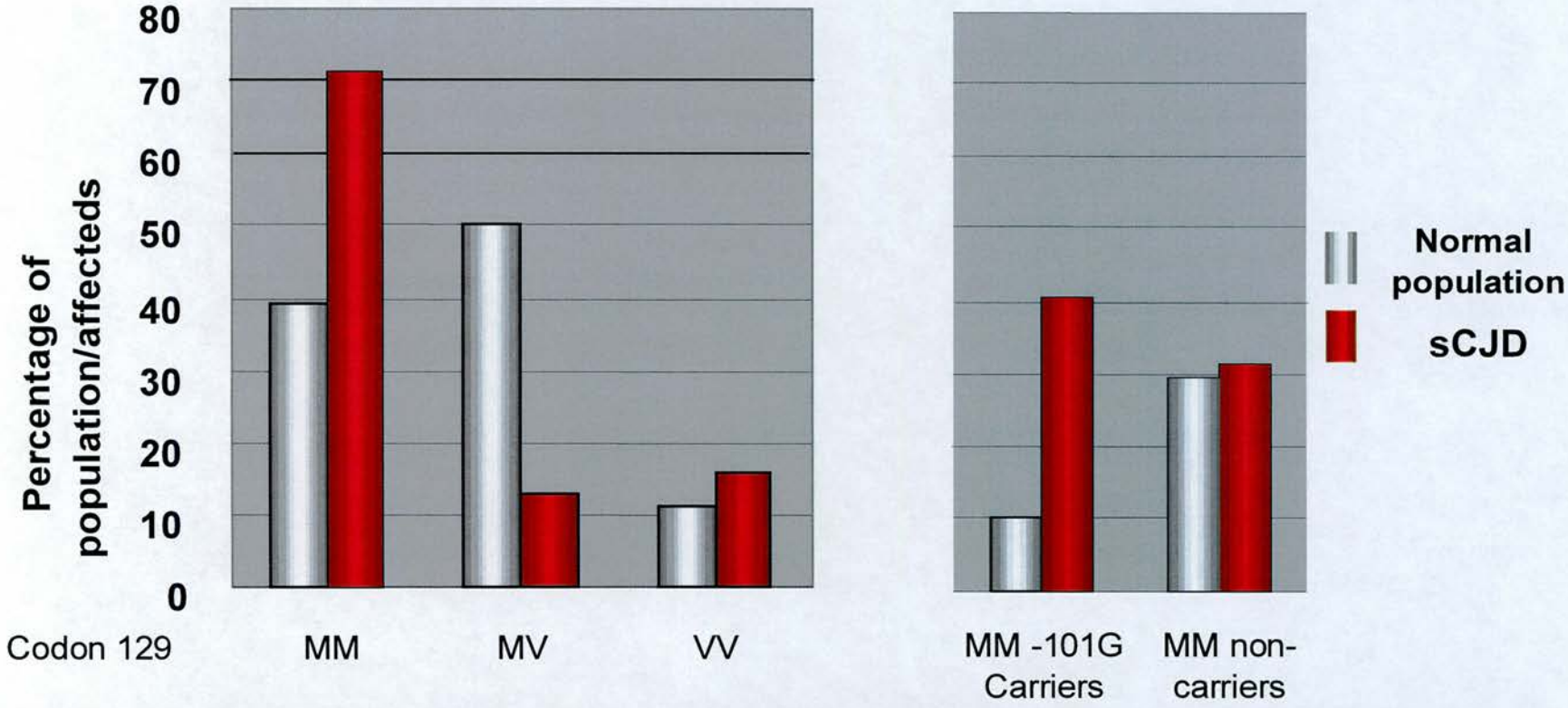


Figure 5-9 Interaction between codon 129 and -101G in affecting susceptibility to sCJD. Homozygosity at codon 129 increases susceptibility of MM homozygous individuals to sCJD with 70% of cases occurring in the 40% of the population who are methionine homozygous (left panel). This study has shown that sub-dividing methionine homozygotes by the presence or absence of the -101G allele shows that carriers are far more susceptible than non-carriers (ratio of red columns to white columns) with the 10% of the population that are MM -101G accounting for over 40%of disease. Although MM non-carriers are still more susceptible than MV individuals suggesting that -101G is a secondary factor in susceptibility to sCJD

The +310C allele increases susceptibility to sCJD when the effects of other polymorphisms are excluded. Its effect is strongest in MV individuals, although this is based on a very small sample. The +310C allele is absent in vCJD patients, but this can be accounted for by it being predominantly linked to valine at codon 129 as all vCJD patients are methionine homozygotes.

There is a trend towards +385C MM carriers being more susceptible to both sCJD and vCJD, but neither reaches significance. sCJD patients heterozygous for +385C also have a younger age of onset compared with other sCJD patients, but this is not significant although as there are only three such patients a larger sample size is needed to further examine the effect of this polymorphism on the age at onset of sCJD.

The different effects of the polymorphisms on susceptibility to sCJD and vCJD may reflect different effects variation in PrP expression plays in the two diseases. While vCJD is probably contracted by the oral route, the aetiology of sCJD is unknown; it may occur spontaneously by somatic mutation of the PrP ORF, or spontaneous mis-folding of a PrP protein to the pathogenic PrP^{Sc} conformation or be caused by an as yet unrecognised infectious agent. The results presented here suggest that two regulatory loci, -101 and +310 affect susceptibility to sCJD. These polymorphisms may achieve this by altering PrP expression patterns which may alter the efficiency with which infectivity replicates in an individual, resulting in differing susceptibilities to sCJD in the population. It is possible that increased expression in a particular tissue could increase susceptibility to spontaneous sCJD but have no effect on the probability of a vCJD infection developing into disease.

Differences in expression levels could also affect the age at disease onset or duration of disease (onset of clinical signs to death). However, carrying a rare allele at -101, +310 or +385 had no significant effect on the age distribution or length of disease in sCJD or vCJD patients in this study. Given the large number of potentially interacting factors the number of sCJD and vCJD patients in this study is too low to definitively rule out an effect of the regulatory alleles on the incubation period or age at disease onset of either

form of CJD. If this lack of association is confirmed in a larger study this would suggest that although susceptibility to sCJD is affected by the polymorphisms the incubation period is not. This could be reconciled with a model where the transition between the first step of sCJD (infection, somatic mutation or spontaneous conversion to PrP^{Sc}) and the start of the incubation period is controlled by PrP expression levels or patterns. This transition stage could be influenced by the polymorphisms, perhaps by localised over-expression in a tissue involved in the transport or replication of infectivity, but once this critical point has been passed the incubation periods are similar.

5.3.2 Polymorphisms and transcription factor binding sites.

The regulatory region polymorphisms may affect susceptibility to disease by altering binding of transcription factors. To determine if either of the alleles at any of the three polymorphisms altered a transcription factor consensus binding sites, the sequences were analysed using transcription factor databases (The GCG9 file tfsites.dat and the TFSITES search site at http://pdap1.trc.rwcp.or.jp/research/db/TFSEARCH.html).

At the -101 polymorphism neither the G nor C allele created or destroyed any known transcription factor binding consensus sites. The -101 polymorphism lies upstream of the transcription regulatory region identified by transient transfection studies in neuroblastoma cells (-43 to -9). It is possible that this region affects expression in other tissues which would not be detected using the neuroblastoma transient transfection system, or that the –101G polymorphism creates a new transcription factor binding site thus extending the upstream regulatory region.

Sequencing of the lambda clone sequence used in the transfection assays showed that its haplotype was CGC i.e. it contains the +385C polymorphism. The presence of the +385C allele creates a *BSSH*II restriction site which was used to create deletion Δ+296:+384. This relatively small deletion had a very large effect, reducing expression by 50% indicating that this region, which contains +310, is a good candidate region for

mutations which affect *PRNP* expression. As the wild-type allele lacks *the BssH*II restriction site this deletion cannot be replicated in the more common CGT haplotype.

The G to C polymorphism at +310 gives rise to a GC factor (GCF) consensus site on the complementary strand. GC**G**GGGC (GCCC**C**GC on coding strand). This factor has been shown to repress expression of EGFR (Epidermal growth factor receptor) and IGF-1R (insulin like growth factor 1 receptor), while mutation of the binding site in the adenylphosphoribosyltransferase gene (APRT) increases expression (She and Taylor, 1995) (Morrione et al., 1995) (Kitadai et al., 1993). The +310C polymorphism could therefore reduce expression of PrP. However, construct Δ+292:+382, which removed this putative GCF site, has only 50% of CAT expression levels relative to the full (-1558) construct. This result does not support the existence of a repressor element in this region. The GCF sequence may have no functional significance in *PRNP*, or that GCF acts as a positive transcription factor in *PRNP* in contrast to other genes. Alternatively, an uncharacterised transcription factor could bind to this sequence. Gel mobility assays are needed to determine if this sequence binds transcription factors *in vivo*. The polymorphism also removes a CpG site but creates one (C**G**G to C**C**G) whether this subtle change would affect methylation of cytosines which inhibits transcription is unknown.

The T to C polymorphism at +385 was also present in lambda clone 314 from which all the deletion constructs were made. This polymorphism creates the *BssH*II (GCGCG**C**) restriction site which was used in the creation of construct Δ+292:+382, which identified the 90bp region whose removal reduced expression by 50%. Analysing transcription factor binding consensus sites shows that the polymorphism creates a second AP-2 site (which overlaps an existing site), a P3A transcription factor site, a metal response element TGCRCNC consensus binding site (Langmade et al., 2000) and a modulator binding factor consensus site TGCRCRC (Palla et al., 1994). Clearly not all of these factors can bind to this site *in vivo* but binding of one of these transcription factors or a different factor in the polymorphism could result in higher expression than in wild-type alleles. As all of these sites are destroyed in construct Δ+292:+382 they are candidates for expression modulating factors. Gel-binding assays are needed to

determine if the transcription factors actually bind to their consensus sites created by the polymorphisms which could implicate a role *in vivo* which would correlate the increased risk of CJD of carriers of this polymorphism.

5.3.3 Future work

As the three polymorphisms found in this study appear to influence susceptibility to sCJD their effect *in vitro* and *in vivo* should be tested. *In vitro* mutagenesis can be used to firstly convert the construct used in the deletion assay (haplotype CG**C**) to the wild-type CGT haplotype. A second round of *in vitro* mutagenesis on that construct will then create **G**GT and C**C**T constructs which can be transfected into a range of cell types to determine the effects of the various polymorphisms on expression. To determine if any of the polymorphisms affect transcription factor binding, gel mobility assays can be performed using cell extracts from different cell lines and oligonucleotides encoding the three common and rare regulatory region alleles. The creation of a metal response element by the +385C polymorphism may affect the expression of PrP in response to metals. If permanently transfected cells lines were made with this construct they could be used to determine if expression is affected by copper levels or oxidative stress which have been proposed as possible functions of PrP (Brown et al., 1999).

The effect of the alleles *in vivo* could be examined by creating transgenic mouse models that contain reporter genes driven by the different alleles. These models could show in which tissues PrP expression is affected by the alleles and may therefore play a role in increasing susceptibility to disease. This might implicate particular tissues in disease pathogenesis which have not previously been recognised as important. It is also possible that the apparent effects of these alleles could be due to other polymorphic sites outside of the region sequenced which are in linkage disequilibrium with those identified in this study. This could be resolved by sequencing larger regions or carrying out deletion assays on a larger sequence.

As the number, frequency, effect and linkage disequilibrium of any regulatory region polymorphisms was unknown at the start of this project, this study was designed to

detect any rare polymorphism which had a very large effect on susceptibility to sCJD or vCJD. This study was not designed to achieve very high probability values for several polymorphisms each with relatively weak effects on disease as now appears to be the case. Now that the -101G and +310C alleles have shown an association with sCJD, and the +385C alleles may affect vCJD, further larger studies can be designed to confirm the effect of these polymorphisms on CJD. To examine the possible effect of +385 on vCJD, a study which sequences 50 vCJD and 100 codon 129 MM controls would reach a significance of 0.022 if the proportions found in this study were maintained. A future study of -101G which examines 50 codon 129 MM sCJD and 100 controls would reach a significance of 0.0001 if the same proportions of carriers and non-carriers found in this study is replicated. To determine the effect of +310C sequencing 20 MV sCJD and 40 MV controls would result in a significant result (p=0.002).

The polymorphisms may also have an effect in familial TSEs by increasing expression of a mutant PrP allele or by influencing disease if it is carried on the non-mutant allele. This could account for the reduced penetrance and clinical heterogeneity seen in some familial TSEs. Even if the regulatory region polymorphisms do not affect familial disease, they may serve as markers to facilitate haplotyping of mutated PrP genes to determine the relationships between different kindreds with a similar mutation. It would be interesting to determine if -101G and +310C are present in the Japanese population and have a similar effect on susceptibility to sCJD.

This result shows that in addition to the codon 129 polymorphism, two polymorphisms within or adjacent to regulatory regions of the *PRNP* gene affect sCJD susceptibility. These polymorphisms could act by increasing the overall levels of PrP expression or expression levels in particular tissues thereby increasing interactions between PrP molecules leading to a lower age at onset of disease or a greater susceptibility to disease. Further *in vivo*, *in vitro* and sequencing studies are needed to confirm the effect of these polymorphisms on susceptibility to variant and sporadic CJD.

6. Summary and Conclusions

This project is the first characterisation of regulatory regions of the human *PRNP* gene and the first to identify genetic loci outside of the PrP ORF which affect susceptibility to sCJD or vCJD. Based on studies of regulatory regions in other species, 4.8 Kb of sequence, containing 1.5Kb of upstream sequence, the first exon and 3.1Kb of intron, was cloned from a human genomic lambda library into a CAT reporter gene. Transient transfections in human neuroblastoma cells of this construct and of nested deletions of this construct, identified two regulatory regions, one between –43 and –9 bases upstream of the transcription initiation site and one within the intron between +292 and +622. The consensus transcription factor binding sites found in the upstream sequence are similar to that seen in other species while intronic regulatory regions have been found in the first introns of the murine and bovine PrP genes.

Having identified two regulatory regions in neuroblastoma cells the effect of the two regulatory regions *in vivo* was examined by making three constructs where expression of a LacZ reporter gene was driven by wither the 4.8Kb human fragment, the upstream and exon sequence alone or the exon and intron sequence. These constructs were designed to determine if the 4.8Kb region contained sufficient regulatory elements to express the reporter gene with the same tissue specificity as PrP and to determine the effects of the upstream and intronic regulatory region independently. Another construct which contained 3.4Kb of murine sequence containing upstream sequence, first exon and intron sequence of the *Prnp* gene was used as a control. Although all four constructs were successfully injected into fertilised eggs giving rise to transgenic embryos and adults, no LacZ expression was detected although faint background staining was observed in all embryos after staining for 48 hours. The failure of the constructs to express may be due to lack of sufficient regulatory elements to drive expression in entire organisms as all four constructs expressed LacZ when transiently transfected into murine neuroblastoma cells.

The upstream and intronic regions were sequenced in sCJD, vCJD and control individuals to determine if any polymorphisms or mutations exist in the regulatory regions which affect susceptibility or directly cause either form of CJD. Three polymorphisms were identified; a C to G transversion at position -101 relative to the start of exon one was found at an allele frequency of 13% in control individuals, a G to C transversion at +310 found at a frequency of 6% in controls and a T to C transition at position +385 found at a frequency of 5% in controls. While the +310 and +385 polymorphisms lay within the 330bp intronic region which affects expression, the -101G allele lay outside the upstream regulatory region, but was detected in sequencing as this position is adjacent to the upstream regulatory region. It is possible that sequence changes at this position could affect the binding of transcription factors and therefore the activity of the upstream regulatory region.

Analysis of the frequency of these polymorphisms in sCJD patients and controls showed that the -101G and +310C alleles each independently significantly increase an individual's risk of developing sporadic CJD. Both alleles were analysed for linkage disequilibrium to the polymorphic codon 129 methionine/valine polymorphism, which is known to affect sCJD and vCJD susceptibility. The -101G allele is linked to the methionine allele and effect of the -101G allele only reached significance in codon 129 methionine homozygotes. Similarly, the +310C allele is in linkage disequilibrium with valine at codon 129 and although there was a trend towards codon 129 heterozygotes who carry the +310C allele being more susceptible to sCJD than non-carriers, the number of sCJD cases was too small to reach statistical significance when sub-divided by codon 129 genotype. In vCJD, no allele reached statistical significance, although there was a non-significant trend for carriers of the +385C being almost three-fold more likely to develop vCJD. Due to the large number of potential interactions between the regulatory region and codon 129, a larger study is needed to confirm the associations found in this study.

How these polymorphisms may influence susceptibility to sCJD is not yet known. The most likely explanation is that the polymorphic alleles affect PrP expression levels. Reducing PrP expression levels results in a lengthening of TSE incubation periods in

transgenic mice and it is therefore difficult to envisage how a decrease in expression levels would increase susceptibility to human TSEs. The effect of the rare alleles is therefore more likely to be an increase rather than a decrease in expression levels. If the -101G and +310C alleles do increase PrP protein levels this could be due to increased transcription levels or, in the case of the +310C allele, an increase in mRNA stability which could result in an increase in PrP protein levels. Analysis of the transcriptional activity of these alleles *in vivo* and *in vitro* is necessary to determine if the polymorphisms actually affect expression levels.

Relatively small changes in expression levels could conceivably affect the course of disease (if the expression increase occurs in a tissue critical in the development of disease) therefore, the relative levels of expression driven by different alleles will have to be examined in a wide variety of human cell lines and tissues and in transgenic mice expressing reporter genes to detect differences between different alleles. It is possible that the apparent effect of the polymorphic alleles of increasing susceptibility to disease may not be caused by those alleles but by linked polymorphisms or mutations located outside of the region analysed in this study. *In vitro* testing of the effect of the polymorphisms on expression could rule this possibility out if a difference in expression levels can be detected between the different alleles.

The aetiology of sCJD is unknown. Previous work has identified heterozygosity at codon 129 as a protective factor and that methionine homozygotes are more likely to develop disease than valine homozygotes. This study has shown that codon 129 methionine homozygotes who carry -101G are more likely to develop sCJD than those who do not. This could be reconciled with a model where difficulties in interactions between the methionine and valine proteins in heterozygotes hinders the disease process making them less likely to develop sCJD. While an increase in expression levels caused by -101G leads to a greater likelihood of sCJD developing by increasing the rate of replication of infectivity perhaps in the spleen, or efficiency of transport of the agent, if the difference in expression levels is seen in the peripheral nervous systems. The interactions between the codon 129 proteins appears to be a more important factor than the levels of expression as codon 129 heterozygotes who carry -101G are at no more risk

than those who do not carry the allele. Codon 129 may therefore be the main genetic factor in determining susceptibility to sCJD with -101G as a secondary factor. But its linkage to methionine means that only methionine homozygotes are at increased risk. A similar model can be devised for the +310C allele, which has shown a trend towards increasing the susceptibility of codon 129 heterozygotes to sCJD. Again differences at codon 129 in the protein reduces the likelihood of disease occurring, but the +310C allele may increase expression of one allele perhaps making an individual effectively homozygous at codon 129 in a particular tissue if the majority of protein produced in that tissue is either valine or methionine at codon 129. This could then increase the likelihood of an individual developing sCJD. These two models are speculative and further epidemiological and molecular analysis is required to determine if the polymorphisms act in this manner. But these results have extended the spectrum of susceptibility to human TSEs beyond the open reading frame of the *PRNP* gene.

Susceptibility to iatrogenic CJD could also be influenced by the regulatory region polymorphisms and an examination of iCJD cases could show if the polymorphisms affect susceptibility and the incubation period of iCJD. The interaction of regulatory region polymorphisms with codon 129 could show a different effect in iCJD as, in contrast to sCJD, valine homozygotes are more likely to develop iCJD than methionine homozygotes.

Other forms of TSE such as BSE in cattle or CWD in deer could also be influenced by variation in PrP expression levels or patterns. The best candidate for regulatory region analysis may be scrapie in sheep as it has the most polymorphic PrP ORF suggesting that scrapie has exerted considerable selective pressure in sheep. The identification of regulatory regions could identify factors which cause scrapie and possibly lead to the selection of highly resistant animals to breed resistance into affected flocks. It is also possible that the apparent effect of some of the ovine protein polymorphisms could be due to linkage disequilibrium with regulatory region polymorphisms.

The existence of regulatory region polymorphisms in the immediate upstream and intronic regulatory regions that affect susceptibility to disease opens the possibility that

other regulatory regions of the *PRNP* gene which lie outside of the region examined in this study may also contain polymorphisms which alter expression and susceptibility to CJD. The discovery of such regions could identify factors which determine susceptibility to sCJD and/or vCJD which might allow a better estimate of the scale of the vCJD epidemic by determining the variation in susceptibility to vCJD in the population. A better understanding of sCJD and vCJD susceptibility factors would also allow the identification of uniquely susceptible individuals who could be candidates for any TSE therapies which may be developed.

Appendix

Sequencing data

Sequence of lambda clone 314

Genbank entry AF315723: Exons in red;

```
   1 gaattcatct cccagtgagc catctctgag gaaaggaggt aaggctctga atgcgatgtc
  61 aacctaagca ttacataaca ctgataggtt gtaaaatggc ctcgctgcct agtgctcatt
 121 ctaactggcc acaatccagt cctccctgtt gcactgagtc agctccctcc aaaggcggcg
 181 ctttccaggc ctcctggttg ccctttccca gcagaccttt caagtgctca cccacactca
 241 cataaacatg gcccaggcac tgtttacagc agctctcctc tgtgcattct cctggcttct
 301 cctgtctctc gccatttgta cccctccccc gcaacttgag tgacagtgat tggtcctgag
 361 atgcaaatat ttgatgcaca tatgtttaca atgcagcctc agcttttttt ttttttttg
 421 taaagagaca aagagtgaga cagggtcttg gcctatagcc ctgtggccca ggctggagtg
 481 cagtggcaca attaaagctc actgcagcct ctacctcctg ggctcaagca atcctcccat
 541 ctcagcctcc ccagtagctg ggactacagg actgtgccac cgttcccagc aatttttttt
 601 atttttgta gaaatggggt ctcactatgt tgctcactct ggtctcaaac tcctgagctc
 661 aagcaatctt cctgccttgg cctctgaaag tgctgggatt acaggcctga gccactgcac
 721 ctggctgtca acattcttaa atctctttcc ttacatgctt cctaaacctc tcacccaaaa
 781 ctaggagact agatgtccta ttttccccag ggcatgcctg gtttacgccc atttcacttt
 841 aaaagtgccc aatttgggta ataatttata agatcccccc ccctctaaat cctgtccttc
 901 tatcacttca tccttcgctc tcctttaaaa tgagacagtt gtcagcagga atcctgcgca
 961 agaacacacc accctgtttc atagaagata tctcaggtaa tgtgcaaaca cgggttttta
1021 aacggagcgc atttttctca tttgttaata tcaccaccta aatcatctct tgcctaaaac
1081 aaggagtaga aagtgaatga aggaaggaac aggtgatggt cagtgtcctt tctacgcctc
1141 aaaatttaag agtttatgtg aaaattcata aatattaatc tcaatccagg ttaagcaaaa
1201 ttttttgctc tcctctttag aaatttctgg ttgccaaagt tccagaaatt gcttcctcat
1261 tcctgagcct ttcattttct cgatttctcc attatgtaac ggggagctgg agctttgggc
1321 cgaatttcca attaaagatg atttttacag tcaatgagcc acgtcaggga gcgatggcac
1381 ccgcaggcgg tatcaactga tgcaagtgtt caagcgaatc tcaactcgtt ttttccggtg
1441 actcattccc ggccctgctt ggcagcgctg caccctttaa cttaaacctc ggccggccgc
1501 ccgccggggg cacagagtgt gcgccgggcc gcgcggcaat tggtccccgc gccgacctcc
1561 gcccgcgagc gccgccgctt cccttccccg ccccgcgtcc ctccccctcg gccccgcgcg
1621 tcgcctgtcc tccgagccag tcgctgacag ccgcggcgcc gcgagcttct cctctcctca
1681 cgaccgaggc aggtaaacgc ccggggtggg aggaacgcgg gcgggggcag gggagccgcg
1741 ggggccgagt gaggaccccg ggcctcgggt cccaggcgca agggtgcccg gccgggcggg
1801 gtcgggaccc cagtgaggag gggccggggg ctgccccgcg ggcgcgtgac gcgtctcggg
1861 cctgcccggc tgcgctggtc tccgctcggg tgaggcggct tggcttcgct tttcaggtta
1921 ggaaagctcc ctttactgcg cgctgggggg ctgggggagc tggcggagcc ccgttaggga
1981 ggtcggtggc gccggggtgt ctcagcgccc cctgcacccc gcgcgggtcc ggcccagcgg
2041 gcgatcgctg gcgcccaggg aactccggga gggccgccag cgggctccgc agggcgcggg
2101 gcggggaggg gcgcctgggg gccgcggggc tcgcgctccc cgcccgttgg ccgcccctcg
2161 gaggccgaga tcggggccca gaacgccccc tggcaaggcc tggcgcttcc gcgatgccca
2221 gagggtgctt gggggggatgg agagaggggc gcccgccggg ggagttccgg gagcctcggt
2281 gcctcccgcc gcagctgcag cgttcctccc gggaggcggc ccagcccttc atcctcgccg
2341 cctgagcttc tccgaggggg gctgcagcct tgcggccgtt gccaccgcct ggagaagcgg
2401 cccacgcgga ctgacgggcg ggggcggggc ctcgggcctc ggcgggggcg gggtccgggg
2461 aggccccacc ctctgttctc caggggcggg gagagaggag ctgcaggtct gcggcctggc
2521 cccaggtgcg atggcggacc ccagcttggc cagtcacatt cctcccagtc cccctggagg
2581 gagaacgctg gccatggggg gctccaagga acaaccagcc tcggatgacg acccttgggt
2641 caccggtctc cccacctgtg cggcaggcgc cttcacgttt cattattaaa caatggggag
2701 aaatccatgt ttactgtcct ttttaaggaa ttttttgctc ttctctttga ggtggctgta
2761 ggaaatagat tttttttta acctcgcaat tccaccacgg tcacatccat cctcgccatc
2821 gcagagccac agctctccgt ttttgtttcc tagcctccag attctcacac aacacagtgc
2881 agtttcactg ctgtaatgat gaggatcttc atggccgcgt tattttcttg ttctgagagc
```

```
2941 atcacggttt aattagcagt tccccatatg atttgaagtg tttcccgttt ccttagggaa
3001 aactcctggt agaataggat taaggatttt tacaaatata attatcaaaa acataggaac
3061 agggaattgg ataaatatgt taaacttctg gaaaaatcaa caacgctctt agatttgtag
3121 aagaaaggaa aaaatcacca gtggaaagga gcaattttac ttacacaaac acagagaagg
3181 tcttacagtg aaaaaaagct aaccagtaag gggaaaagca ggcagagggg taggatgtga
3241 tttgtatgtt atttatatct aacacaagtc ttccacaccg aaaggaaaat attaagatta
3301 taatagataa atggcaaaat gatgagtcat ttacacaata aaatgcaaat tagagcatgt
3361 ttgggttatc attttacatc tattaaaata accaaaataa ttaatagtaa cagcaaccct
3421 tgctggaagg ttgcccaaaa cttggcattt tcaagtgtct ggggaggtgg cagggctttg
3481 gggtcacaaa gatggttctg cagtcaattt tgtgaccttg gacaggctac ctaatttcct
3541 gatcctcctt ttgtccattc atagaatgga ggaaatgata gctactttct gcgtctgtat
3601 gtatgagtta ttgggggcat ttcgaaccag tgacaaacat tttgttaagc aatctggtga
3661 tgcattaaga agctggaagc tgtgacccag aaaccccact cctgagaact tacctgcaat
3721 ggaagaaaca aacaaacaaa aacaggcatg tattcctagc agaatgatct aaaattagaa
3781 cacctggaaa agagcctaaa tgtataacac cagggcagta gctaagaaaa ttatgacaca
3841 ttaactgaaa tgaacattat gtaaccacta aaaatcatga ttttggagcc tgtgatatgt
3901 ggggaaaaac tgacaagtaa aaaagtgggt tattaactgc acctgcttac tctaacgtga
3961 acgcatatgt gaaaaatctg aaaggaaaag cacagaaaat ggacgttttc attgaaattg
4021 tcggtgatct taattttcct ttggtgaata tattgctctc actaagggcg tttaaaaaat
4081 agttcacagg ttttaatttt ttagatgaaa tggacccaca gttttctgta agagaaagga
4141 gagattgtta tatttgctac ttagaataaa agatttttag ccaacgttgt ttccttttc
4201 aaatattttt ccatttttt agttgattaa tgatttagta atttgtgtat tgggtttttt
4261 taagaatcag ttcttagatt catttatcaa ttctagtttt ttgttgttgt ttttaaggac
4321 tcctgaatat ttttcaaaac tgaacaattt cagccatgtc tgagctttcc gtcttcctgg
4381 aggcacaaat ctagtttagc tgaaccacaa cagattgtac atatcctgca gaacctctgt
4441 ggtcttagga aggttgaaag tcaccaaatg tcacagaaaa tgagggtcag gaaaggctgt
4501 gatcacaagc tcttgatcca gaggccactc gggtgcatct gtgtgactga cacctcattg
4561 catcatcatt tcttgtaaaa tcagcttaat tttgcagaga aatggccaca gatttttaga
4621 attatgctga attatgtaag ggagcagcca ttttaaaaat aatgtaatca aataacaata
4681 ataacataac agaagtttct gcaccacact gacactgaac atggctgtca accaactagc
4741 caaaccaagg tgaaagattc tctgaacttc agttttaact tacttcgtcc attctccctt
4801 ttcagatc
```

Results of sequencing CJD patients and controls. Number=randomised i.d. number for each individual. –101,+310, +385 =Genotype at each of the regulatory region polymorphisms. Common allele in Capitals.Rare allele in small font. *NS*=not sequenced. Class=Origin of control (Hospital, Edinburgh or Belfast) or form of CJD (sporadic or variant). Codon 129 =Genotype at codon 129 in the ORF (ORFs sequenced at the CJD unit Edinburgh)

Number	-101	+310	+385	Class	Codon129
1	CC	GG	TT	Edinburgh	MM
2	CC	*NS*	*NS*	Edinburgh	MM
3	*NS*	GG	TT	Edinburgh	MM
4	*NS*	*NS*	*NS*	Hospital	MV
5	CC	GG	TT	Hospital	MV
6	*NS*	*NS*	*NS*	Edinburgh	MV
7	CC	GG	TT	Edinburgh	MM
8	*NS*	*NS*	*NS*	Edinburgh	MV
9	*NS*	GG	cc	Edinburgh	MM
10	CC	GG	TT	Edinburgh	MV
11	CC	GG	TT	Edinburgh	MV
12	*NS*	*NS*	*NS*	Variant	MM
13	CC	GG	Tc	Variant	MM
14	CC	GG	TT	Edinburgh	MM
15	CC	Gc	TT	Hospital	MM
16	CC	GG	TT	Sporadic	MM
17	Cg	GG	TT	Sporadic	MM
18	CC	Gc	TT	Belfast	VV
19	Cg	GG	TT	Sporadic	MM
20	CC	GG	TT	Edinburgh	MM
21	CC	GG	Tc	Edinburgh	MM
22	Cg	GG	TT	Edinburgh	MM
23	CC	GG	Tc	Sporadic	MM
24	CC	GG	TT	Edinburgh	VV
25	CC	GG	Tc	Belfast	MM
26	CC	GG	TT	Edinburgh	VV
27	CC	GG	TT	Belfast	MM
28	CC	GG	TT	Sporadic	MM
29	CC	GG	Tc	Sporadic	MM
30	gg	GG	TT	Sporadic	MM
31	*NS*	Gc	TT	Edinburgh	MM

32	*NS*	*NS*	*NS*	Edinburgh	MV
33	Cg	GG	TT	Edinburgh	MM
34	CC	GG	TT	Edinburgh	MV
35	CC	GG	TT	Edinburgh	MV
36	*NS*	GG	TT	Edinburgh	MV
37	CC	*NS*	*NS*	Edinburgh	MV
38	Cg	GG	TT	Edinburgh	MM
39	Cg	GG	TT	Variant	MM
40	Cg	GG	TT	Belfast	MM
41	CC	GG	TT	Edinburgh	MV
42	Cg	GG	TT	Sporadic	MM
43	CC	GG	Tc	Hospital	MM
44	*NS*	GG	TT	Sporadic	MM
45	CC	GG	cc	Variant	MM
46	CC	GG	TT	Edinburgh	MM
47	CC	GG	TT	Belfast	MM
48	CC	GG	TT	Hospital	MM
49	Cg	GG	TT	Edinburgh	MM
50	CC	GG	TT	Edinburgh	MM
51	*NS*	GG	TT	Edinburgh	MM
52	CC	Gc	TT	Edinburgh	MV
53	*NS*	*NS*	*NS*	Edinburgh	VV
54	Cg	*NS*	*NS*	Hospital	MM
55	CC	GG	TT	Belfast	MM
56	CC	GG	TT	Edinburgh	MM
57	CC	GG	TT	Hospital	MM
58	Cg	GG	TT	Belfast	MM
59	CC	GG	TT	Hospital	MM
60	CC	Gc	TT	Sporadic	MV
61	CC	GG	TT	Edinburgh	MM
62	CC	GG	TT	Hospital	MM
63	CC	GG	TT	Variant	MM
64	CC	GG	Tc	Edinburgh	MM
65	CC	GG	Tc	Variant	MM
66	Cg	GG	TT	Edinburgh	MV
67	CC	Gc	TT	Sporadic	MM
68	Cg	GG	TT	Variant	MM
69	CC	GG	TT	Belfast	MM
70	gg	GG	TT	Edinburgh	MM
71	CC	*NS*	*NS*	Sporadic	MM
72	Cg	GG	TT	Edinburgh	MV
73	Cg	GG	TT	Belfast	MV

74	CC	GG	TT	Edinburgh	MM
75	CC	Gc	TT	Hospital	MV
76	CC	GG	TT	Belfast	MV
77	CC	GG	TT	Sporadic	MM
78	Cg	GG	TT	Edinburgh	MM
79	CC	GG	TT	Sporadic	VV
80	CC	GG	TT	Edinburgh	VV
81	CC	GG	TT	Sporadic	MV
82	Cg	Gc	TT	Edinburgh	MV
83	CC	GG	TT	Hospital	MV
84	CC	*NS*	*NS*	Edinburgh	MV
85	CC	GG	Tc	Edinburgh	MV
86	CC	GG	TT	Edinburgh	MV
87	CC	GG	TT	Edinburgh	MM
88	*NS*	*NS*	*NS*	Variant	MM
89	*NS*	*NS*	*NS*	Sporadic	MM
90	CC	GG	Tc	Belfast	MV
91	Cg	*NS*	*NS*	Edinburgh	MM
92	CC	Gc	TT	Edinburgh	VV
93	Cg	GG	TT	Variant	MM
94	Cg	GG	TT	Edinburgh	MM
95	CC	GG	TT	Variant	MM
96	CC	GG	Tc	Sporadic	MM
97	CC	GG	TT	Hospital	MV
98	CC	GG	TT	Edinburgh	MM
99	CC	*NS*	*NS*	Variant	MM
100	CC	Gc	TT	Edinburgh	MM
101	CC	Gc	TT	Sporadic	MV
102	CC	GG	TT	Sporadic	VV
103	CC	Gc	TT	Sporadic	VV
104	Cg	Gc	TT	Sporadic	MM
105	CC	*NS*	*NS*	Belfast	MM
106	Cg	GG	TT	Edinburgh	MM
107	Cg	GG	TT	Edinburgh	MV
108	CC	*NS*	*NS*	Belfast	MM
109	gg	GG	TT	Edinburgh	MM
110	CC	GG	TT	Variant	MM
111	CC	GG	TT	Edinburgh	VV
112	CC	GG	TT	Edinburgh	MM
113	CC	GG	TT	Sporadic	MM
114	Cg	*NS*	*NS*	Sporadic	MM
115	Cg	GG	TT	Belfast	MV

116	Cg	GG	TT	Sporadic	MM
117	CC	*NS*	*NS*	Belfast	MV
118	*NS*	GG	*NS*	Edinburgh	MM
119	CC	GG	TT	Belfast	MV
120	*NS*	*NS*	*NS*	Variant	MM
121	CC	GG	TT	Edinburgh	MV
122	CC	GG	TT	Edinburgh	MM
123	CC	GG	TT	Edinburgh	MV
124	CC	GG	TT	Belfast	MM
125	CC	GG	TT	Hospital	MV
126	CC	GG	TT	Edinburgh	MV
127	CC	*NS*	*NS*	Variant	MM
128	Cg	GG	TT	Sporadic	MM
129	Cg	*NS*	*NS*	Variant	MM
130	CC	*NS*	*NS*	Belfast	MV
131	*NS*	*NS*	*NS*	Belfast	MM
132	*NS*	*NS*	*NS*	Belfast	MV
133	CC	*NS*	TT	Belfast	MM
134	CC	GG	TT	Variant	MM
135	*NS*	GG	TT	Belfast	MV
136	CC	GG	TT	Hospital	MM
137	Cg	GG	TT	Variant	MM
138	CC	GG	TT	Belfast	MV
139	CC	Gc	TT	Belfast	MV
140	CC	GG	TT	Hospital	VV
141	CC	GG	TT	Belfast	MV
142	CC	GG	TT	Hospital	VV
143	CC	Gc	TT	Belfast	MV
144	CC	GG	TT	Hospital	MM
145	Cg	GG	TT	Belfast	MM
146	Cg	GG	TT	Hospital	MV
147	CC	GG	TT	Belfast	MV
148	gg	GG	TT	Sporadic	MM
149	*NS*	Gc	TT	Belfast	MV
150	CC	GG	TT	Variant	MM
151	CC	GG	TT	Variant	MM
152	CC	GG	TT	Variant	MM
153	CC	GG	Tc	Belfast	MM
154	Cg	GG	TT	Belfast	MM
155	CC	Gc	TT	Hospital	MM
156	*NS*	*NS*	*NS*	Belfast	VV
157	*NS*	*NS*	*NS*	Hospital	*NS*

158	CC	GG	TT	Belfast	MM
159	Cg	GG	TT	Sporadic	MM
160	Cg	GG	TT	Variant	MM
161	*NS*	*NS*	*NS*	Belfast	MV
162	CC	GG	TT	Belfast	MM
163	CC	*NS*	*NS*	Belfast	MM
164	CC	GG	TT	Sporadic	MM
165	CC	GG	TT	Belfast	VV
166	CC	GG	TT	Belfast	MM
167	CC	*NS*	*NS*	Belfast	MV
168	CC	Gc	TT	Sporadic	VV
169	CC	GG	TT	Belfast	MV
170	CC	GG	Tc	Variant	MM
171	*NS*	*NS*	*NS*	Belfast	VV
172	CC	GG	Tc	Hospital	MV
173	Cg	GG	TT	Belfast	MM
174	CC	Gc	TT	Hospital	VV
175	Cg	GG	Tc	Hospital	MV
176	CC	GG	TT	Hospital	MM
177	Cg	*NS*	*NS*	Belfast	MM
178	Cg	GG	TT	Hospital	MV
179	CC	GG	TT	Belfast	MM
180	CC	GG	TT	Belfast	MV
181	Cg	GG	Tc	Belfast	MV
182	*NS*	*NS*	*NS*	Hospital	VV
183	*NS*	*NS*	*NS*	Belfast	VV
184	*NS*	*NS*	*NS*	Belfast	MV
185	CC	*NS*	*NS*	Belfast	VV
186	*NS*	*NS*	*NS*	Belfast	MV
187	*NS*	*NS*	*NS*	Belfast	VV
188	*NS*	*NS*	*NS*	Belfast	MV
189	*NS*	GG	TT	Belfast	VV
190	*NS*	*NS*	*NS*	Hospital	VV
191	CC	*NS*	*NS*	Belfast	MV
192	Cg	*NS*	*NS*	Belfast	MM
193	*NS*	*NS*	*NS*	Belfast	MV
194	CC	GG	TT	Belfast	MV
195	*NS*	*NS*	*NS*	Belfast	MV
196	CC	GG	TT	Hospital	MM
197	*NS*	*NS*	*NS*	Belfast	MV
198	*NS*	*NS*	*NS*	Belfast	MM
199	*NS*	*NS*	*NS*	Belfast	MM

200	*NS*	GG	TT	Belfast	MV

References

Agresti, A. (1992). A Survey of Exact Inference for Contegency Tables. Statistical Science *7*, 131-153.

Aguzzi, A., Bitiftler, T., Klein, M., Brandner, S., Raeber, A., Flechsig, E., and Weissmann, C. (1997). Neurotoxicity and neuroinvasiveness of prions. Brain Pathology *7*, 1137-1138.

Alper, T. (1985). Scrapie agent unlike viruses in size and susceptibility to inactivation by ionizing or ultraviolet-radiation. Nature *317*, 750.

Alperovitch, A., Zerr, I., Pocchiari, M., Mitrova, E., de Pedro Cuesta, J., Hegyi, I., Collins, S., Kretzschmar, H., van Duijn, C., and Will, R. G. (1999). Codon 129 prion protein genotype and sporadic Creutzfeldt-Jakob disease. The Lancet *353*, 1673-1674.

Antequera, F., and Bird, A. (1999). CpG islands as genomic footprints of promoters that are associated with replication origins. Current Biology *9*, R661-667.

Barbanti, P., Fabbrini, G., Salvatore, M., Petraroli, R., Cardone, F., Maras, B., Equestre, M., Macchi, G., Lenzi, G. L., and Pocchiari, M. (1996). Polymorphism At Codon-129 or Codon-219 Of Prnp and Clinical Heterogeneity In a Previously Unreported Family With Gerstmann- Straussler-Scheinker Disease (Prp-P102l Mutation). Neurology *47*, 734-741.

Basler, K., Oesch, B., Scott, M., Westaway, D., Walchli, M., Groth, D., McKinlay, M., Prusiner, S., and Weissmann, C. (1986). Scrapie and Cellular PrP Isoforms are encoded by the same chromosomal gene. Cell *46*, 417-428.

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

Belay, E. D. (1999). Transmissible spongiform encephalopathies in humans. Annual Review of Microbiology *53*, 283-314.

Bernoulli, C., Siegfried, J., Baumgartner, G., Regli, F., and Rabinowicz, T. (1977). Danger of accidental person-to-person transmission of CJD by surgery. Lancet *1*, 478-479.

Bessen, R. A., Kocisko, D. A., Raymond, G. J., Nandan, S., Lansbury, P. T., and Caughey, B. (1995). Non-genetic propagation of strain-specific properties of scrapie prion protein. Nature *375*, 698-700.

Blackwood, E., and Kadonga, J. (1998). Going the distance: A current view of enhancer action. Science *281*, 60-63.

Blattler, T., Brandner, S., Raeber, A., Klein<M., Voigtlander, T., Weissmann, C., and Aguzzi, A. (1997). PrP-expressing tissue required for transfer of scrapie infectivity from spleen to brain. Nature *389*, 69-73.

Brandner, S., Isenmann, S., Raeber, A., Fischer, M., Sailer, A., Kobayashi, Y., Marino, S., Weissmann, C., and Aguzzi, A. (1996). Normal Host Prion Protein Necessary For Scrapie-Induced Neurotoxicity. Nature *379*, 339-343.

Brown, D. R., and Besinger, A. (1998). Prion protein expression and superoxide dismutase activity. Biochemical Journal *334*, 423-429.

Brown, D. R., Hafiz, F., Glassmith, L. L., Wong, B.-S., Jones, I. M., Clive, C., and Haswell, S. J. (2000).

Consequences of manganese replacement of copper for prion protein function and proteinase resistance. EMBO *19*, 1180-1186.

Brown, D. R., Iordanova, N. K., Wong, B. S., Venien-Bryan, C., Hafiz, F., Glasssmith, L. L., Sy, M. S., Gambetti, P., Jones, I. M., Clive, C., and Haswell, S. J. (2000). Functional and structural differences between the prion protein from two alleles prnp(a) and prnp(b) of mouse. 267 *European Journal of Biochemistry*, 2452-2459.

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

Brown, D. R., Wong, B.-S., Hafiz, F., Clive, C., Haswell, S. J., and Jones, I. M. (1999). Normal prion protein has an activity like that of superoxide dismutase. Biochemical Journal *344*, 1-5.

Brown, K. L., Stewart, K., Bruce, M. E., and Fraser, H. (1997). Severely combined immunodeficient (SCID) mice resist infection with bovine spongiform encephalopathy. Journal Of General Virology *78*, 2707- 2710.

Brown, K. L., Stewart, K., Ritchie, D. L., Mabbott, N. A., Williams, A., Fraser, H., Morrison, W. I., and Bruce, M. E. (1999). Scrapie replication in lymphoid tissues depends on prion protein-expressing follicular dendritic cells. Nature Medicine *5*, 1308-1312.

Brown, K. L., Stewart, K., Ritchie, D. L., Mabbott, N. A., Williams, A., Fraser, H., Morrison, W. I., and Bruce, M. E. (1999). Scrapie replication in lymphoid tissues depends on prion protein-expressing follicular dendritic cells. Nature Medicine *5*, 1308-1312.

Brown, P., and Bradley, R. (1998). 1755 and all that: a historical primer of transmissible spongiform encephalopathy. British medical Journal *317*, 1688-1692.

Brown, P., Gibbs, C. J., Rodgersjohnson, P., Asher, D. M., Sulima, M. P., Bacote, A., Goldfarb, L. G., and Gajdusek, D. C. (1994). Human spongiform encephalopathy - the national-institutes-of-health series of 300 cases of experimentally transmitted disease. Annals of Neurology *35*, 513-529.

Brown, P., Preece, M., Brandel, J. P., Sato, T., McShane, L., Zerr, I., Fletcher, A., Will, R. G., Pocchiari, M., Cashman, N. R., D'Aignaux, J. H., Cervenakova, L., Fradkin, J., Schonberger, L. B., and Collins, S. J. (2000). Iatrogenic Creutzfeldt-Jakob disease at the millennium. Neurology *55*, 1075-1081.

Bruce, M., Chree, A., McConnell, I., Foster, J., Pearson, and Fraser, J. (1994). Transmission of BSE and scrapie to mice:strain variation and the species barrier. Phil. Trans. R. Soc. Lond. B *343*, 405-411.

Bruce, M. E. (1993). Scrapie strain variation and mutation. British Medical Bulletin *49*, 822-838.

Bruce, M. E., Will, R. G., Ironside, J. W., McConnell, I., Drummond, D., Suttie, A., McCardle, L., Chree, A., Hope, J., Birkett, C., Cousens, S., Fraser, H., and Bostock, C. J. (1997). Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent. Nature *389*, 498-501.

Brugere-Picoux, J. (1996). Clinical aspects of scrapie in sheep. In Transmissible Subacute Spongiform Encephalopathies: Prion Diseases, L. Court, ed.: Elsevier), pp. 29-36.

Bueler, H., Aguzzi, A., Sailer, A., Greiner, R. A., Autenried, P., Aguet, M., and Weissmann, C. (1993). Mice devoid of PrP are resistant to scrapie. Cell *73*, 1339-1347.

Bueler, H., Raeber, A., Sailer, A., Fischer, M., Aguzzi, A., and Weissmann, C. (1994). High prion and prpsc levels but delayed-onset of disease in scrapie-inoculated mice heterozygous for a disrupted prp

gene. Molecular Medicine *1*, 19-30.

Bugiani, O., Giaccone, G., Piccardo, P., Morbin, M., Tagliavini, F., and Ghetti, B. (2000). Neuropathology of Gerstmann-Straussler-Scheinker disease. Microsc Res Tech *50*, 10-15.

Cadore, J. L. (1996). Feline spongiform encephalopathy. Point Veterinaire *28*, 49-51.

Caughey, B., and et al. (1988). Detection of prion protein mRNA in normal and scrapie-infected tissues and cell lines. Journal of General Virology *69*, 711-716.

Caughey, B., Kocisko, D. A., Raymond, G. J., and Lansbury, P. T. (1995). Aggregates Of Scrapie-Associated Prion Protein Induce the Cell- Free Conversion Of Protease-Sensitive Prion Protein to the Protease-Resistant State. Chemistry & Biology *2*, 807-817.

Caughey, B. W., Dong, A., Bhat, K. S., Ernst, D., Hayes, S. F., and Caughey, W. S. (1991). Secondary structure-analysis of the scrapie-associated protein prp 27-30 in water by infrared-spectroscopy. Biochemistry *30*, 7672-7680.

Cervenakova, L., Goldfarb, L. G., Garruto, R., Lee, H. S., Gajdusek, D. C., and Brown, P. (1998). Phenotype-genotype studies in kuru: Implications for new variant Creutzfeldt-Jakob disease. Proceedings of the National Academy of Sciences of the United States of America *95*, 13239-13241.

Chapman, J., Benisrael, J., Goldhammer, Y., and Korczyn, A. D. (1994). The risk of developing creutzfeldt-jakob-disease in subjects with the prnp gene codon-200 point mutation. Neurology *44*, 1683-1686.

Chen, I., Baum, L., Ng, H., Chan, L., Sastre, I., Artiga, M., Valdivieso, F., Bullido, M., Chiu, H., and Pang, C. (1999). Apolipoprotein E promoter and alpha2-macroglobulin polymorphisms are not genetically associated with Chinese late onset Alzheimer's disease. Neuroscience letters *269*, 173-177.

Chernoff, Y. O., Lindquist, S. L., Ono, B., Ingevechtomov, S. G., and Liebman, S. W. (1995). Role of the chaperone protein hsp104 in propagation of the yeast prion-like factor [psi(+)]. Science *268*, 880-884.

Chiesa, R., Piccardo, P., Ghetti, B., and Harris, D. A. (1998). Neurological illness in transgenic mice expressing a prion protein with an insertional mutation. Neuron *21*, 1339-1351.

Chodosh, L., Baldwin, A., Carthew, R., and Sharp, P. (1988). Human CCAAT-Binding proteins have heterologous subunits. Cell *53*, 11-24.

Clark, I., Rowan, A., Edwards, D., Bech-Hansen, T., Manns, D., Bahr, M., and Cawston, T. (1997). Transcriptional activity of the human tissue inhibitor of mettaloproteinase1 (TIMP-1) gene in fibroblasts involves elements in the promoter, exon 1 and intron 1. Biochemical Journal *324*, 611-617.

Collinge, Whittington, Sidle, Smith, Palmer, and Clarke (1994). Prion protein is necessary for normal synaptic function. Nature *370*, 295-297.

Collinge, J., Brown, J., Hardy, J., Mullan, M., Rossor, M. N., Baker, H., Crow, T. J., Lofthouse, R., Poulter, M., Ridley, R., and al., e. (1992). Inherited prion disease with 144 base pair gene insertion. Clinical and pathological features. Brain *115*, 687-710.

Collinge, J., Palmer, M. S., and Dryden, A. J. (1991). Genetic predisposition to iatrogenic Creutzfeldt-Jakob disease. Lancet *337*, 1441-1442.

Collinge, J., Sidle, K., Meads, J., Ironside, J., and Hill, A. (1996). Molecular Analysis Of Prion Strain Variation and the Etiology Of New Variant CJD. Nature *383*, 685-690.

Come, J. H., Fraser, P. E., and Lansbury, P. T. (1993). A kinetic-model for amyloid formation in the prion diseases - importance of seeding. Proc. Natl. Acad. Sci., USA *90*, 5959-5963.

Corder, E., Saunders, A., Strittmatter, W., Schmechel, D., Gaskell, P., Small, G., Roses, A., Haines, J., and Pericak-Vance, M. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of AD in late onset families. Science *261*, 921-923.

Cowell, I., Skinner, A., and Hurst, H. (1992). Transcriptional repression by a novel member of the bZIP family of transcription factors. Mol Cell Biol *12*, 3070-3077.

Cuille, J., and Chelle, P. L. (1936). La tremblente du mouton est bien inoculble. CR. Sciences Acad. sci.paris *206*, 78-79.

dAignaux, J. H., Alperovitch, A., and Maccario, J. (1998). A statistical model to identify the contaminated lots implicated in iatrogenic transmission of Creutzfeldt-Jakob disease among French human growth hormone recipients. American Journal Of Epidemiology *147*, 597- 604.

Das, H., and Lal, H. (1997). Genes implicated in the pathogenesis of Alzheimer's disease. Front Biosci *2*, 253-259.

Deslys, J., Jaegly, A., Huillard, J., Mouthon, F., Billette, T., and Dormont, D. (1998). Genotype at codon 129 and susceptibility to Creutzfeldt-Jakob Disease. The Lancet *351*, 1251.

Dickinson, A. G., and Outram, G. W. (1988). Genetic aspects of unconventional virus infections: the basis of the virino hypothesis. Ciba Foundation Symposium *135*, 63-83.

Diringer (1996). Small virus-like structure in fractions from scrapie hamster brain. Lancet *343*, 894-895.

Doh-ura, K., Kitamoto, T., Sakaki, Y., and Tateishi, J. (1991). CJD discrepency. Nature *353*, 801-802.

Faisst, S., and Meyer, S. (1992). Compilation of vertebrate-encoded transcription factors. Nucleic Acids Res *20*, 3-26.

Fiering, S., Whitelaw, E., and Martin, D. (2000). To be or not to be active: the stochastic nature of enhancer action. Bioessays *22*, 381-387.

Finckh, U., Muller-Thomsen, T., Mann, U., Eggers, C., Marksteiner, J., Meins, W., Binetti, G., Alberici, A., Hock, C., Nitsch, R. M., and Gal, A. (2000). High prevalence of pathogenic mutations in patients with early-onset dementia detected by sequence analyses of four different genes. American Journal of Human Genetic *66*, 110-117.

Fink, J. K., Peacock, M. L., Warren, J. T., Roses, A. D., and Prusiner, S. B. (1994). Detecting prion protein gene-mutations by denaturing gradient gel-electrophoresis. Human Mutation *4*, 42-50.

Finkerstrand, C. E., Lovejoy, E. A., and Quinn, J. P. (1999). An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. FEBS Letters *458*, 171-174.

Fischer, M., Rulicke, T., Raeber, A., Sailer, A., Moser, M., Oesch, B., Brandner, S., Aguzzi, A., and Weissmann, C. (1996). Prion protein(PrP) with amino-proximal deletions restoring susceptibility of PrP knockout mice to scrapie. EMBO *15*, 1255-1264.

Foster, J. D., McKelvey, W. A., Mylne, M. J., Williams, A., Hunter, N., Hope, J., and Fraser, H. (1992). Studies on maternal transmission of scrapie in sheep by embryo transfer. Vet. Rec. *130*, 341-343.

Gabriel, J. M., Oesch, B., Kretzschmar, H., Scott, M., and Prusiner, S. B. (1992). Molecular-cloning of a candidate chicken prion protein. Proc. Natl. Acad. Sci., USA *89*, 9097-9101.

Gajdusek, D. (1996). Kuru in childhood:implications for the problem of whether BSE affects humans. In Transmissible Subacute Spongiform Encephalopathies: Prion Diseases, L. Court, ed.: Elsevier), pp. 15-28.

Gajdusek, D., and Zigas, V. (1959). Clinical, pathological and epidemiological study of an acute progressive degenerative disease of the central nervous system among natives of the eastern highlands of New Guinea. Am. J. Med. *26*, 442-469.

Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. J. Mol. Biol *196*, 261-282.

Gebhardt, F., Zanker, K., and Brandt, B. (1999). Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron I. The Journal of Biological Chemistry *274*, 13176-13180.

Ghani, A. C., Ferguson, N. M., Donnelly, C. A., and M., A. R. (2000). Predicted vCJD mortality in Great Britain: modelling the latest data puts a ceiling on the likely number of vCJD cases. Nature *406*, 583-584.

Ghetti, B., Dlouhy, S. R., Giaccone, G., Prelli, F., Piccardo, P., Farlow, M. R., Frangione, B., Bugiani, O., and Tagliavini, F. (1994). Gerstmann-Straussler-Scheinker disease (GSS). Neurobiology of Aging *15 S1*, 151.

Goldfarb, and al., e. (1992). FFI and familial CJD:disease phenotype determined by a DNA polymorphism. Science *258*, 806-808.

Goldfarb, L. G., Brown, P., Mitrova, E., Cervenakova, L., Goldin, L., Korczyn, A. D., Chapman, J., Galvez, S., Cartier, L., and Rubenstein, R. (1991). Creutzfeldt-Jacob disease associated with the PRNP codon 200Lys mutation: an analysis of 45 families. European. Journal. of. Epidemiology. *7*, 477-486.

Goldmann, W., O'Neill, G., Cheung, F., Charleson, F., Ford, P., and Hunter, N. (1999). PrP (prion) gene expression in sheep may be modulated by alternative polyadenylation of its messenger RNA. Journal of General Virology *80*, 2275-2283.

Hainfellner, J. A., Liberski, P. P., Guiroy, D. C., Cervenakova, L., Brown, P., Gajdusek, D. C., and Budka, H. (1997). Pathology and immunocytochemistry of a Kuru Brain. Brain Pathology *7*, 547-553.

Hamer, D., and Sirota, L. (2000). Beware the chopsticks gene. Mol Psychiatry *5*, 11-13.

Hidaka, K., Yamamoto, I., Arai, Y., and Mukai, T. (1993). The MEF-3 motif is required for MEF-2 mediated skeletal muscle-specific induction of the Rat adolase A gene. Molecular and Cell Biology *13*, 6469-6478.

Hill, A. F., Antoniou, M., and Collinge, J. (1999). Protease-resistant prion protein produced *in vitro* lacks detectable infectivity. Journal of General Virology *80*, 11-14.

Hill, A. F., Desbruslais, M., Joiner, S., Sidle, K. C. L., Gowland, I., Collinge, J., Doey, L. J., and Lantos, P. (1997). The same prion strain causes vCJD and BSE. Nature *389*, 448-450.

Hodge, S. (1994). What association analysis can and cannot tell us about the genetics of complex disease. American Journal of Medical genetics *54*, 318-323.

Horiuchi, M., Ishiguro, N., Nagasawa, H., Toyoda, Y., and Shinagawa, M. (1997). Alternative usage of exon 1 of bovine PrP mRNA. Biochemical and Biophysical Research Communications *233*, 650-654.

Horiuchi, M., Yamazaki, N., Ikeda, T., Ishiguro, N., and Shinagawa, M. (1995,). A cellular-form of prion protein (prpc) exists in many nonneuronal tissues of sheep. J. Gen. Virol. *76, No.Pt10,*, 2583-2587.

Hornshaw, M. P., McDermott, J. R., Candy, J. M., and Lakey, J. H. (1995). Copper-binding to the n-terminal tandem repeat region of mammalian and avian prion protein - structural studies using synthetic peptides. Biochemical and Biophysical Research Communications *214*, 993-999.

Hsiao, K. K., Groth, D., Scott, M., Yang, S. L., Serban, H., Raff, D., Foster, D., Torchia, M., Dearmond, S. J., and Prusiner, S. B. (1994). Serial transmission in rodents of neurodegeneration from transgenic mice expressing mutant prion protein. Proc. Natl. Acad. Sci., USA *91*, 9126-9130.

Hsiao, K. K., Scott, M., Foster, D., Groth, D. F., DeArmond, S. J., and Prusiner, S. B. (1990). Spontaneous neurodegeneration in transgenic mice with mutant prion protein. Science *250*, 1587-1590.

Hunter, N. (1997). PrP genetics in sheep and the implications for scrapie and BSE. Trends in Genetics *5*, 331-334.

Hunter, N., Goldmann, W., Smith, G., and Hope, J. (1994). Frequencies of PrP gene variants in healthy cattle and cattle with BSE in Scotland. Veterinary *135*, 400-403.

Hunter, N., Manson, J., Charleston, and Hope, J. (1994). Comparison of expression patterns of PrP mRNA in the developing sheep and mouse. Ann. NY Acad. Sc *724*, 353.

Ikeda, T., Horiuchi, M., Ishiguro, N., Muramatsu, Y., Kaiuwe, G. D., and Shinagawa, M. (1995). Amino-acid polymorphisms of prp with reference to onset of scrapie in suffolk and corriedale sheep in japan. J. Gen. Virol. *76*, 2577-2581.

Inoue, S., Tanaka, M., Horiuchi, M., Ishiguro, N., and Shinigawa, M. (1997). Characterisation of a bovine prion protein gene: The expression requires interaction between the promoter and intron. Journalk of Veterinary Medical Science *59*, 175-183.

Ionasescu, V. (1998). X-linked Charcot-Marie-tooth disease and connexin 32. Cell. Biol. int. *22*, 807-813.

Ironside, J. W. (1998). New-variant Creutzfeldt-Jakob disease. Neuropathology *18*, 131-138.

Kaneko, K., Ball, H. L., Wille, H., Zhang, H., Groth, D., Torchia, M., Tremblay, P., Safar, J., Prusiner, S. B., DeArmond, S. J., Baldwin, M. A., and Cohen, F. E. (2000). A synthetic peptide initiates Gerstmann-Straussler-Scheinker (GSS) disease in transgenic mice. Journal of Molecular Biology *295*, 997-1007.

Kazazian, H. (1990). The Thalassermia syndromes: Molecular Basis and prenatal diagnosis in1990. Seminars in Haemotology *27*, 209-228.

Kim, C., Kim, H., Cubells, J., and Kim, K. (1999). A previously undescribed intron and extensive 5' Upstream sequence but not Phox2a-mediated transactivation, are neccessary for high level cell type-specific expression of the human noreprinephrine transporter gene. Journal ofBiological Chemistry *274*, 6507-6518.

Kimberlin, R. H., and Walker, C. A. (1980). Pathogenesis of mouse scrapie: evidence for neural spread of infection to the CNS. J. Gen. Virol. *51*, 183-187.

Kirkwood, J. K., and Cunningham, A. A. (1994). Epidemiologic observations on spongiform

encephalopathies in captive wild animals in the british-isles. Vet. Rec. *135*, 296-303.

Kitadai, Y., Yamazaki, H., Ysui, W., Kyo, E., Yokozaki, H., Kajiyama, G., Johnson, A. C., Pastan, I., and Tahara, E. (1993). GC factor represses transcription of several growth factor/receptor genes and causes growth inhibiiton of human gastric carcinoma cell lines. Cell Growth and Differentiation *4*, 291-296.

Kitamoto, T., Muramoto, T., Mohri, S., Doh-Ura, K., and Tateishi, J. (1991). Abnormal isoform of prion protein accumulates in follicular dendritic cells in mice with Creutzfeldt-Jakob disease. J. Virol. *65*, 6292-6295.

Kitamoto, T., and Tateishi, J. (1994). Human prion diseases with variant prion proteins. Phil. Trans. R. Soc. Lond. B *343*, 391-398.

Langmade, S. J., Ravindra, R., Daniels, P. J., and Andrews, G. K. (2000). The transcription factor MTF-1 mediates metal-regulation of the mouse ZnT1 gene. Journal of Biological Chemistry *275*, 9377-9384.

Lantos, P., Bhatia, K., Doey, L., Al-Sarraj, S., Doshi, R., and Collinge, J. (1997). Is the neuropathology of new variantt CJD and kuru similar ? The Lancet *350*, 187-188.

Lasmezas, C. I., Deslys, J., Robain, O., Jaegly, A., Beringue, V., Peyrin, J., Fournier, J., Hauw, J., Rossier, J., and Dormont, D. (1997). Transmission of the BSE Agent to Mice in the Absence of Detectable Abnormal Prion Protein. Science *275*, 402-405.

Lazarini, F., Deslys, J. P., and Dormont, D. (1991). Regulation of the glial fibrillary acidic protein, beta-actin and prion protein messenger-rnas during brain-development in mouse. Molecular Brain Research *10*, 343-346.

Lee, H. S., Sambuughin, N., Cervenakova, L., Chapman, J., Pocchiari, M., Litvak, S., Qi, H. Y., Budka, H., Delser, T., Furukawa, H., Brown, P., Gajdusek, D. C., Long, J. C., Korczyn, A. D., and Goldfarb, L. G. (1999). Ancestral origins and worldwide distribution of the PRNP 200K mutation causing familial Creutzfeldt-Jakob disease. American Journal of Human Genetics *64*, 1063-1070.

Lee, I. Y., Westaway, D., Smit, A. F. A., Wang, K., Seto, J., Chen, L., Acharya, C., Ankener, M., Baskin, D., Cooper, C., Yao, H., Prusiner, S. B., and Hood, L. E. (1998). Complete genomic sequence and analysis of the prion protein gene region from three mammalian species. Genome Research *8*, 1022-1037.

Lemaire-Vieille, C., Schulze, T., Podevin-Dimster, V., Follet, J., Bailly, Y., Blanquet-Grossard, F., Decavel, J. P., Heinen, E., and Cesbron, J. Y. (2000). Epithelial and endothelial expression of the green fluorescent protein reporter gene under the control of bovine prion protein (PrP) gene regulatory sequences in transgenic mice. Proceedings of the National Academy of Sciences *97*, 5422-5427.

Li, A., Sakaguchi, S., Atarashi, R., Roy, B. C., Nakaoke, R., Arima, K., Okimura, N., Kopacek, J., and Shigematsu, K. (2000). Identification of a novel gene encoding a PrP-like protein expressed as chimeric transcripts fused to PrP exon 1/2 in ataxic mouse line with a disrupted PrP gene. Cellular and Molecular Neurobiology *20*, 553-567.

Li, G., and Bolton, D. C. (1997). A novel hanster prion protein mRNA contains an extra exon: increased expression in scrapie. Brain research *751*, 265-274.

Lledo, P. M., Tremblay, P., DeArmond, S. J., Prusiner, S. B., and Nicoll, R. A. (1996). Mice deficient for prion protein exhibit normal neural exciteability and synaptic transmission in the hippocampus. PNAS *93*, 2403-2407.

Mabbott, N., K. Brown, J. Manson, and Bruce, M. (1997). T-lymphocyte activation and the cellular form of the prion protein. Immunology *92*, 161-165.

MacKenzie, A., and Quinn, J. (1999). A serotonin transporter gene intron 2 polymorphic region correlated with affective disorders, has allele dependent differential enhancer-like properties in the mouse embryo. PNAS *96*, 15251-15255.

Maity, S., and Crombrugghe, B. (1998). Role of the CCAAT-bindingprotein CBF/NF-Y in transcription. TIBS *23*, 174-178.

Mallucci, G. R., Campbell, T. A., Dickinson, A., Beck, J., Holt, M., Plant, G., Pauw, K. W. d., Hakin, R. N., Clarke, C. E., Howell, S., Davies-Jones, G. A. B., Lawden, M., Smith, C. M. L., Ince, P., Ironside, J. W., Bridges, L. R., Dean, A., Weeks, I., and Collinge, J. (1999). Inherited prion disease with an alanine to valine mutation at codon 117 in the prion protein gene. Brain *122*, 1823-1837.

Manson, J. (1996). Prnp gene dosage, allelic specificity and gene regulation in the transmissible spongiform encephalopathies. Transmissible Subacute Spongiform Encephalopathies: Prion Diseases, 239-245.

Manson, J., Clarke, A., Hooper, Aitchison, L., McConnell, I., and Hope, J. (1994). 129/OLA mice carrying a null mutation in PrP that abolishes mRNA production are developmentally normal. Molecular Neurobiology *8*, 121-127.

Manson, J., Clarke, A., McBride, P., McConnell, I., and Hope, J. (1994). PrP gene dosage determines the timing but not the final intensity or distribution of lesions in scrapie pathology. Neurodegeneration *3*, 331-340.

Manson, J., West, J., Thompson, V., McBride, P., Kaufman, M., and Hope, J. (1992). The prion protein gene: a role in mouse embryogenesis? Development *115*, 117-122.

Manson, J. C., Hope, J., Clatke, A. R., Johnston, A., Black, C., and Macleod, N. (1995). PrP Gene dosage and long term potentiation. Neurodegeneration *4*, 113-115.

Manson, J. C., Jamieson, E., Baybutt, H., Tuzi, N. L., Barron, R., McConnell, I., Somerville, R., Ironside, J., Will, R., Sy, M. S., Melton, D. W., Hope, J., and Bostock, C. (1999). A single amino acid alteration (101L) introduced into murine PrP dramatically alters incubation time of transmissible spongiform encephalopathy. EMBO Journal *18*, 6855-6864.

Manuelidis , E., and Manuelidis, L. (1993). A transmissible CJD-like agent is prevalent in the human population. PNAS *90*, 7724-7728.

Marsh, R. (1992). Transmissible mink encephalopathy. Prion diseases of humans and animals, 300-307.

Mastana, S. S., and Sokol, R. J. (1998). Genetic variation in the east Midlands. Ann Hum Biol *25*, 43-68.

McEwan, D., and Ornitz, D. (1998). Regulation of the fibroblast Growth Factor receptor 3 Promoter and Intron I Enhancer by Sp1 family Transcription factors. The Journal of Biological Chemistry *273*, 5349-5357.

McKenzie, D., Bartz, J. C., and Marsh, R. F. (1996). Transmissible Mink Encephalopathy. Seminars In Virology *7*, 201-206.

McLean, C. A., Storey, E., Gardner, R. J. M., Tannenberg, A. E. G., Cervenakova, L., and Brown, P. (1997). The D178N (cis-129M) "fatal familial insomnia" mutation associated with diverse clinicopathologic phenotypes in an Australian kindred. Neurology *49*, 552-558.

Mead, S., Beck, J., Dickinson, A., Fisher, E. M. C., and Collinge, J. (2000). Examination of the human

prion protein-like gene Doppel for genetic susceptibility to sporadic and variant Creutzfeldt-Jakob disease. Neuroscience Letters *290*, 117-120.

Miele, G. (1999). Analysis of gene expression in Prnp+/+ and Prnp -/- (Edinburgh: University of Edinburgh).

Millson, G., Hunter, G., and Kimberlin, R. (1976). The physico-chemical nature of the scrapie agent. In Slow virus diseases of man and animals., R. Kimberlin, ed. (Amsterdam: North Holland), pp. 244-266.

Ministry of Agriculture, f. a. f. (1996). Bovine Spongiform Encephalopathy in Great Britain. A progress report. (London.

Moore, R. C., Lee, I. Y., Silverman, G. L., Harrison, P. M., Strome, R., Heinrich, C., Karunaratne, A., Pasternak, S. H., Chishti, M. A., Liang, Y., Mastrangelo, P., Wang, K., Smit , A. F. A., Katamine, 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. Journal of Molecular Biology *292*, 797-817.

Moore, R. C., Redhead, N. J., Selfridge, J., Hope, J., Manson, J. C., and Melton, D. W. (1995). Double replacement gene targeting for the production of a series of mouse strains with different prion protein gene alterations. Bio Technology *13*, 999-1004.

Morrione, A., deAnglis, T., and Baserga, R. (1995). The GC factor regulates the expression of the insulin-like growth factor-1 receptor. Cell Proliferation *28*, 659-671.

Murdoch, G. H., Sklaviadis, T., Manuelidis, E. E., and Manuelidis, L. (1990). Potential retroviral RNAs in Creutzfeldt-Jakob disease. J. Virol. *64*, 1477-1486.

Nakano, Y., Nishihara, T., Sasayama, S., Miwa, T., Kamada, S., and Kakunaga, T. (1991). Transcriptional regulatory elements in the 5' upstream and first intron regions of the human smooth muscle (aortic type) alpha-actin-encoding gene. Gene *99*, 285-289.

Nishida, N., Tremblay, P., Sugimoto, T., Shigematsu, K., Shirabe, S., Petromilli, C., Erpel, S. P., Nakaoke, R., Atarashi, R., Houtani, T., Torchia, M., Sakaguchi, S., DeArmond, S. J., Prusiner, S. B., and Katamine, S. (1999). A mouse prion protein transgene rescues mice deficient for the prion protein gene from Purkinje cell degeneration and demyelination. Laboratory Investigation *79*, 689-697.

Nitrini, R., Rosemberg, S., PassosBueno, M. R., daSilva, L. S. T., Iughetti, P., Papadopoulos, M., Carrilho, P. M., Caramelli, P., Albrecht, S., Zatz, M., and LeBlanc, A. (1997). Familial spongiform encephalopathy associated with a novel prion protein gene mutation. Annals Of Neurology *42*, 138-146.

Nixon, R., Mastrianni, J., Roberts, J., Prusiner, S., and Dearmond, S. (1996). Sporadic Occurrence Of Fatal Thalamic Insomnia Without a Prion Protein Gene Mutation. Journal Of Neuropathology and Experimental Neurology *55*, 225.

Oesch, B., Westaway, D., Walchli, M., McKinley, M. P., Kent, S. B., Aebersold, R., Barry, R. A., Tempst, P., Teplow, D. B., and Hood, L. E. (1985). A cellular gene encodes scrapie PrP 27-30 protein. Cell *40*, 735-746.

Oskouian, B., Rangan, V., and S., S. (1997). Regulatory elements in the first intron of the rat fatty acid synthase gene. Biochem. J. *324*, 113-121.

Owen, F., Poulter, M., Collinge, J., and Crow, T. J. (1990). Codon-129 changes in the prion protein gene in caucasians. American Journal of Human Genetics *46*, 1215-1216.

Owen, F., Poulter, M., Shah, T., Collinge, J., Lofthouse, R., Baker, H., Ridley, R., McVey, J., and Crow, T. J. (1990). An in-frame insertion in the prion protein gene in familial Creutzfeldt-Jakob disease. Brain Research. Molecular Brain Research *7*, 273-276.

Palla, F., Bonura, C., Anello, L., Di Gaetano, L., and Spinelli, G. (1994). Modulator factor-binding sequence of the sea urchin early histone H2A promoter acts as an enhancer element. Proc Natl Acad Sci U S A *91*, 12332-6.

Palmer, M., Dryden, A., Hughes, J., and Collinge, J. (1991). Homozygous prion protein genotype predisposes to sporadic CJD. nature *352*, 340-342.

Palmer, M. S., and Collinge, J. (1992). Human prion diseases. Current Opinion in Neurology and Neurosurgery *5*, 895-901.

Palmer, M. S., Mahal, S. P., Campbell, T. A., Hill, A. F., Sidle, K. C., Laplanche, J. L., and Collinge, J. (1993). Deletions in the prion protein gene are not associated with CJD. Human Molecular Genetics *2*, 541-544.

Palmer, M. S., Vanleeven, R. H., Mahal, S. P., Campbell, T. A., Humphreys, C. B., and Collinge, J. (1996). Sequence Variation Of Intron Of Prion Protein Gene, Crucial For Complete Diagnostic Strategies. Human Mutation *7*, 280-281.

Pan, K. M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R. J., Cohen, F. E., and al., e. (1993). Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. Proc. Natl. Acad. Sci., USA *90*, 10962-10966.

Parchi, P., Capellari, S., Chen, S. G., Petersen, R. B., Gambetti, P., Kopp, N., Brown, P., Kitamoto, T., Tateishi, J., Giese, A., and Kretzschmar, H. (1997). Typing prion isoforms. Nature *386*, 232-233.

Parchi, P., Capellari, S., Chin, S., Schwarz, H., Schecter, N., Butts, J., Hudkins, P., Burns, D., Powers, J., and Gambetti , P. (1999). A subtype of sporadic prion disease mimicking fatal familial insomnia. Neurology *52*, 1757-1763.

Parchi, P., Giese, A., Capellari, S., Brown, P., SchulzSchaeffer, W., Windl, O., Zerr, I., Budka, H., Kopp, N., Piccardo, P., Poser, S., Rojiani, A., Streichemberger, N., Julien, J., Vital, C., Ghetti, B., Gambetti, P., and Kretzschmar, H. (1999). Classification of sporadic Creutzfeldt-Jakob disease based on molecular and phenotypic analysis of 300 subjects. Annals of Neurology *46*, 224-233.

Parchi, P., Peterson, R., and Gambetti, P. (1996). New topics in familial prion diseases. Seminars in Virology *7*, 181-187.

Patino, M. M., Chernoff, Y. O., Liebman, S. W., and Lindquist, S. L. (1995). Analysis Of the Interaction Of Hsp104 and Yeast Prion Sup35 In Saccharomyces-Cerevisiae. Molecular Biology Of the Cell *6*, 35.

Pauly, P. C., and Harris, D. A. (1998). Copper stimulates endocytosis of the prion protein. Journal of Biological Chemistry *273*, 33107-33110.

Petraroli, R., and Pocchiari, M. (1996). Codon 219 polymorphism of PRNP in healthy caucasians and CJD patients. Am. J. Hum. Genet. *58*, 888-889.

Piccardo, P., Dlouhy, S. R., Lievens, P. M. J., Young, K., Thomas, D. P., Nochlin, D., Dickson, D. W., Vinters, H. V., Zimmerman, T. R., Mackenzie, I. R. A., Kish, S. J., Ang, L. C., DeCarli, C., Pocchiari, M., Brown, P., Gibbs, C. J., Gajdusek, D. C., Bugiani, O., Ironside, J., Tagliavini, F., and Ghetti, B. (1998). Phenotypic variability of Gerstmann-Straussler-Scheinker disease is associated with prion protein heterogeneity. Journal of Neuropathology and Experimental Neurology *57*, 979-988.

Pruisner, S. (1991). The molecular biology of prion diseases. Science *252*, 1515-1522.

Prusiner, S. B. (1996). Prion Biology and Diseases - Laughing Cannibals, Mad Cows, and Scientific Heresy. Medicinal Research Reviews *16*, 487-505.

Prusiner, S. B., McKinley, M. P., Groth, D. F., Bowman, K. A., Mock, N. I., Cochran, S. P., and Masiarz, F. R. (1981). Scrapie agent contains a hydrophobic protein. Proc. Natl. Acad. Sci., USA *78*, 6675-6679.

Puckett, C., Concannon, P., Casey, C., and Hood, L. (1991). Genomic structure of the human prion protein gene. American Journal of Human Genetics *49*, 320-329.

Race, R., Oldstone, M., and Chesebro, B. (2000). Entry versus blockade of brain infection following oral or intraperitoneal scrapie administration: role of prion protein expression in peripheral nerves and spleen. Virology *74*, 828-33.

Race, R. E., Priola, S. A., Bessen, R. A., Ernst, D., Dockter, J., Rall, G. F., Mucke, L., Chesebro, B., and Oldstone, M. B. A. (1995). Neuron-specific expression of a hamster prion protein minigene in transgenic mice induces susceptibility to hamster scrapie agent. Neuron *15*, 1183-1191.

Raeber, A. J., Race, R. E., Brandner, S., Priola, S. A., Sailer, A., Bessen, R. A., Mucke, L., Manson, J., Aguzzi, A., Oldstone, M. B. A., Weissmann, C., and Chesebro, B. (1997). Astrocyte-specific expression of hamster prion protein (PrP) renders PrP knockout mice susceptible to hamster scrapie. Embo Journal *16*, 6057-6065.

Raeber, A. J., Sailer, A., Hegyi, I., Klein, M. A., Rulicke, T., Fischer, M., Brandner, S., Aguzzi, A., and Weissmann, C. (1999). Ectopic expression of prion protein (PrP) in T lymphocytes or hepatocytes of PrP knockout mice is insufficient to sustain prion replication. Proceedings of the National Academy of Sciences of the United States of America *96*, 3987-3992.

Robeck, G., Cheung, B., Growdon, W., Deng, A., Ahuthota, P., Locascia, J., Greenberg, S., and Hyman, B. (1999). Lack of independent associations of apolipoprotein E promoter and intron 1 polymorphisms with Alzheimer's disease. Neuroscience Letters *272*, 155-158.

Rossi, G., Macchi, G., Porro, M., Giaccone, G., Bugiani, M., Scarpini, E., Scarlato, G., Molini, G. E., Sasanelli, F., Bugiani, O., and Tagliavini, F. (1998). Fatal familial insomnia - Genetic, neuropathologic, and biochemical study of a patient from a new Italian kindred. Neurology *50*, 688-692.

Rubinsztein, D. C. (1997). A review of the contribution of apo E genotypes to Alzheimer's disease risk and meta-analysis of apo E associations with other neuro-psychiatric disorders. Nutrition Metabolism and Cardiovascular Diseases *7*, 267-275.

Saeki, K., Matsumoto, Y., Hirota, Y., Matsumoto, Y., and Onodera, T. (1996). 3-Exon Structure Of the Gene Encoding the Rat Prion Protein and Its Expression In Tissues. Virus Genes *12*, 15-20.

Saeki, K., Matsumoto, Y., Matsumoto, Y., and T.Onodera (1996). Identification of a promoter region in the rat prion protein gene. Biochem. Biophys. Res. Comm *219*, 47-52.

Sakaguchi, Katamine, and Nishida (1996). Loss of cerebellar Perkinje cells in aged mice homozygous for a disrupted PrP gene. Nature *380*, 528-530.

Sakaguchi, S., Katamine, S., Shigematsu, K., Nakatani, A., Moriuchi, R., Nishida, N., Kurokawa, K., Nakaoke, R., Sato, H., Jishage, K., Kuno, J., Noda, T., and Miyamoto, T. (1995). Accumulation of proteinase k-resistant prion protein (prp) is restricted by the expression level of normal prp in mice inoculated with a mouse-adapted strain of the Creutzfeldt- Jakob-disease agent. J. Virol. *69*, 7586-7592.

Salvatore, and et al. (1996). Codon 200 mutation in a new family of Chilean origin with CJD. J. Neurol Neurosurg psychiatry *61*, 111-123.

Salvatore, M., Genuardi, M., Petraroli, R., Masullo, C., Dalessandro, M., and Pocchiari, M. (1994). Polymorphisms of the prion protein gene in italian patients with creutzfeldt-jakob-disease. Human Genetics *94*, 375-379.

Samaia, Mari, Vallada, and et al. (1997). A prion-linked psychiatric disorder. Nature *390*, 241.

Sambrook, J., Frisch, E. F., and Maniatis, T. (1989). Molecular cloning. A laboratory manual, 2nd edition (Cold Spring harbour , New York, USA: Cold Spring Harbour Laboratory Press).

Schatzl, Wopnfer, Gilsch, and al., e. (1997). Is codon 129 or prion protein polymorphic in human beings but not in animals. Nature *349*, 1603-1604.

Schatzl, H. M., Dacosta, M., Taylor, L., Cohen, F. E., and Prusiner, S. B. (1995). Prion protein gene variation among primates. Journal of Molecular Biology *245*, 362-374.

Schlapfer, J., Saitbekova, N., Gaillard, C., and Dolf, G. (1999). A new allelic variant in the bovine prion protein gene (PRNP) coding region. Animal Genetics *30*, 386-387.

Scott, M., Foster, D., Mirenda, C., Serban, D., Coufal, F., Walchli, M., Torchia, M., Groth, D., Carlson, G., and DeArmond, S. J. (1989). Transgenic mice expressing hamster prion protein produce species-specific scrapie infectivity and amyloid plaques. Cell *59*, 847-857.

She, B. R., and Taylor, M. W. (1995). Identification of the cis-elements requiredfor transcriptional control of the Chinese hamster ovary APRT gene. Gene *159*, 175-180.

Sher, L. (2000). Psychiatric diagnoses and inconsistent results of association studies in behavioral genetics. 2000 Feb;54(2):207-9 *54*, 207-209.

Shibuya, S., Higuchi, J., Shin, R. W., Tateishi, J., and Kitamoto, T. (1998). Codon 219 lys allele of PRNP is not found in sporadic Creutzfeldt-Jakob disease. Annals of Neurology *43*, 826-828.

Shmerling, D., Hegyi, I., Fischer, M., Blattler, T., Brandner, S., Gotz, J., Rulicke, T., Flechsig, E., Cozzio, A., vonMering, C., Hangartner, C., Aguzzi, A., and Weissmann, C. (1998). Expression of amino-terminally truncated PrP in the mouse leading to ataxia and specific cerebellar lesions. Cell *93*, 203-214.

Sigurdson, C. J., Williams, E. S., Miller, M. W., Spraker, T. R., Orourke, K. I., and Hoover, E. A. (1999). Oral transmission and early lymphoid tropism of chronic wasting disease PrPres in mule deer fawns (Odocoileus hemionus). Journal of General Virology *80*, 2757- 2764.

Simonic, T., Duga, S., Strumbo, B., Asselta, R., Ceciliani, F., and Ronchi, S. (2000). cDNA cloning of turtle prion protein. FEBS Letters *469*, 33-38.

Somerville, R. A. (1999). Host and transmissible spongiform encephalopathy agent strain control glycosylation of PrP. Journal of General Virology *80*, 1865-1872.

Somerville, R. A., Chong, A., Mulqueen, O. U., Birkett, C. R., Wood, S. C. E. R., and Hope, J. (1997). Biochemical typing of scrapie strains. Nature *386*, 564.

Sparrer, H. E., Santoso, A., Szoka, F. C., and Weissman, J. S. (2000). Evidence for the prion hypothesis: induction of the yeast [PSI+] factor by in vitro-converted Sup35 protein. Science *289*, 595-599.

Sternbach, G., Dibble, C., and Varon, J. (1997). From CJD to the Mad Cow Epidemic. Journal of Emergency Medicine *15*, 701-705.

Stevenson, M., Wilesmith, J., Ryan, J., morris, R., Lockhart, J., Lin, D., and Jackson, R. (2000). Temporal aspects of the epidemic of BSE in Great Britain: indivdual animal-associated risk factors for the disease. Veterinary Record *147*, 349-354.

Storbeck, C., Sabourin, L., Waring, J., and Korneluk, R. (1998). Definition of regulatory sequence elements in the promoter region and the first intron of the myotonic dystrophy protein kinase gene. The Journal of Biological Chemistry *273*, 9139-9147.

Telling, G. C., Scott, M., Mastrianni, J., Gabizon, R., Torchia, M., Cohen, F. E., Dearmond, S. J., and Prusiner, S. B. (1995). Prion propagation in mice expressing human and chimeric prP transgenes implicates the interaction of cellular PrP with another protein. cell *83*, 79-90.

Tjian, R., and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. Cell *77*, 5-8.

Tobler, I., Gauss, Deboer, and al., e. (1996). Altered circadian rhythms and sleep in mice devoid of prion protein. Nature *380*, 639-642.

vanDuijn, C. M., DelasnerieLaupretre, N., Masullo, C., Zerr, I., deSilva, R., Wientjens, D. P. W. M., Brandel, J. P., Weber, T., Bonavita, V., Zeidler, M., Alperovitch, A., Poser, S., Granieri, E., Hofman, A., and Will, R. G. (1998). Case-control study of risk factors of Creutzfeldt-Jakob disease in Europe during 1993-95. Lancet *351*, 1081-1085.

Weir, B. S. (1990). Genetic Data Analysis (Sunderland, Massachusetts: Sinauer Associates).

Weissmann, C., and Aguzzi, A. (1999). PrP's Double causes Trouble. Science *286*, 914-915.

Wells, G. A. H., Hawkins, S. A. C., Hadlow, W. J., and Spenser, Y. I. (1992). The discovery of bovine spongiform encephalopathy and observations on vacuolar changes. Prion diseases of humans and animals, 256-274.

Wells, G. A. H., Scott, A. C., Johnson, C. T., Gunning, R. F., Hancock, R. D., Jeffrey, M., Dawson, M., and Bradley, R. (1987). A novel progressive spongiform encephalopathy in cattle. Vet. Rec. *121*, 419-420.

Westaway, D., Cooper, C., Turner, S., Dacosta, M., Carlson, G. A., and Prusiner, S. B. (1994). Structure and polymorphism of the mouse prion protein gene. Proc. Natl. Acad. Sci., USA *91*, 6418-6422.

Westaway, D., Dearmond, S. J., Cayetanocanlas, J., Groth, D., Foster, D., Yang, S. L., Torchia, M., Carlson, G. A., and Prusiner, S. B. (1994). Degeneration of skeletal-muscle, peripheral-nerves, and the central-nervous-system in transgenic mice overexpressing wild-type prion prot ins. Cell *76*, 117-129.

Westaway, D., Goodman, P. A., Mirenda, C. A., McKinley, M. P., Carlson, G. A., and Prusiner, S. B. (1987). Distinct prion proteins in short and long scrapie incubation period mice. Cell *51*, 651-662.

Westaway, D., Mirenda, C. A., Foster, D., Zebarjadian, Y., Scott, M., Torchia, M., Yang, S. L., Serban, H., DeArmond, S. J., Ebeling, C., and al., e. (1991). Paradoxical shortening of scrapie incubation times by expression of prion protein transgenes derived from long incubation period mice. Neuron *7*, 59-68.

Wickner, R. B. (1997). A new prion controls fungal cell fusion incompatibility. Proceedings Of the National Academy Of Sciences Of the United States Of America *94*, 10012-10014.

Wickner, R. B. (1994). [ure3] as an altered ure2 protein - evidence for a prion analog in saccharomyces-cerevisiae. Science *264*, 566-569.

Will, R. G., Ironside, J. W., Zeidler, M., Cousens, S. N., Estibeiro, K., Alperovitch, A., Poser, S., Pocchiari, M., Hofman, A., and Smith, P. G. (1996). A New Variant Of Creutzfeldt-Jakob-Disease In the UK. Lancet *347*, 921-925.

Will, R. G., and Matthews, W. B. (1982). Evidence for case-to-case transmission of CJD. J. neurol. Neurosurg. Psychiatry *45*, 235-238.

Williams, E., and Young, S. (1993). Neuropathology of chronic wasting disease of mule deer (Odocoileus hemionus) and elk (Cervus elaphus nelsoni). Vet Pathol *30*, 36-45.

Windl, O., Giese, A., Schulz-Schaeffer, W., Zerr, I., Skworc, K., Arendt , S., Oberdieck, C., Bodemer, M., Poser, S., and Kretzschmar, H. A. (1999). Molecular genetics of human prion diseases in Germany. Human Genetics *105*, 244-252.

Wood, J. L. N., McGill, I. S., Done, S. H., and Bradley, R. (1997). Neuropathology of scrapie: A study of the distribution patterns of brain lesions in 222 cases of natural scrapie in sheep, 1982-1991. Veterinary Record *140*, 167-174.

Wuthrich, K., Riek, R., Wider, G., Garcia, F. L., Liu, A., Zahn, R., Billeter, M., Hornemann, S., and Glockshuber, R. (1999). Structural biology of prion proteins. Transfusion Clinique Et Biologique *6*, 31.

Zahn, R., Liu, A. Z., Luhrs, T., Riek, R., Schroetter, C., von Garcia, F. L., Billeter, M., Calzolai, L., Wider, G., and Wuthrich, K. (2000). NMR solution structure of the human prion protein. Proceedings of the National Academy of Sciences *97*, 145-150.

Zeidler, M., Stewart, G., Cousens, S., Estibeiro, K., and Will, R. (1997). Codon 129 genotype and new variant CJD. The Lancet *350*, 668.

Zimmermann, K., Turecek, P. L., and Schwarz, H. P. (1999). Genotyping of the prion protein gene at codon 129. Acta Neuropathologica *97*, 355-358.