Molecular and Cellular Investigation of Alzheimer's Disease Associated Risk Loci: *BIN1* and *CD2AP*

A thesis submitted for the Degree of Doctor of Philosophy at
Cardiff University
School of Medicine

Anna Barrett
2017

Thesis summary

Genome wide association studies have identified genes associated with Late Onset Alzheimer's disease (AD). Pathway analyses have used this data to implicate a number of biological mechanisms in AD pathogenesis. Endocytosis has been implicated in AD and is a critical biological mechanism involved in the production of Aβ.

BIN1 and *CD2AP* are associated with AD and function in endocytosis. This thesis describes how depletion of BIN1 and CD2AP has contradicting effects on the processing of APP in a human brain cell line and affects endocytosis in different ways, suggesting multiple cellular trafficking mechanisms may be involved in Aβ production.

As most genetic variants associated with complex diseases are located in the non-coding region of the genome, they may contribute to disease susceptibility by disrupting gene regulation. Variants at the *BIN1* locus associated with AD are located approximately 30 Kb from the coding region, suggesting *BIN1* regulation may be a risk mechanism in AD.

BIN1 expression is influenced by *cis*-regulatory mechanisms in prefrontal cortex tissue. Differential allele expression implied that this *cis*-regulation was influenced by DNA variants. The variant responsible was not identified but there was suggestive evidence for an intronic variant associated with AD.

ChIP-Seq and DNase-Seq data identified chromatin modifications and transcription factor binding in immune cell types at the *BIN1* risk locus. Gene reporter assays showed that the *BIN1* locus was capable of functioning as a genetic enhancer. Furthermore, assays used to investigate DNA protein interactions showed that SPI1, an AD associated transcription factor with a critical immune function, bound to the *BIN1* locus. The *BIN1* index SNP had no effect on gene expression or protein binding, but may have a greater impact on additional disease relevant immune cell types.

Finally, gene-editing techniques were explored as a mechanism for generating isogenic cell models to investigate the effect of AD associated variants.

Declarations

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed …………………………………………..……… Date ………………….…………

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD

Signed …………………………………………..……… Date ………………….…………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed …………………………………………..……… Date ………………….…………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed …………………………………………..……… Date ………………….…………

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed .…………………………………………..……… Date ………………….…………

Acknowledgements

Firstly I would like to thank all my supervisors, Julie Williams, Lesley Jones, Matt Hill, Rhian Thomas and Mel Dunstan for giving me the opportunity to work with you all and on this project. Your advice and guidance throughout has been greatly appreciated.

A special mention would have to go to Matt and Rhian, who have adopted me along the way. The knowledge, passion and integrity you both have is inspiring. With your help, my PhD has become something I can really be proud of and I am so grateful.

I would like to especially thank Nick Bray and Carolina Toste, who selflessly gave up their time and offered to help me. You have both taught me so much and your input is truly valued.

I would like to thank all members of the AD lab team past and present; Taniesha, Jodie, Rhodri, Rachel, Helen, Adele, Gareth. You have been a joy to work with and have provided many hours of entertainment, especially in TC! I'd like to thank everyone else in the AD team, especially Sarah whose insane organisation skills have helped me on so many occasions. I'd also like to thank the rest of the lab staff based in HEB, who have helped me solve the many issues I've had over the years.

There is no way I would have got this far if it wasn't for my PhD buddy – Anna Burt. I am already struggling not having you opposite me every day, always happy to listen to me even when you may not have wanted to. Working with you has undoubtedly made me a better scientist and I have been so lucky to share this experience with you. I'm so glad I have had you to laugh, talk, complain, sing along to Heart and play the time tunnel with, so thank you so much for getting me this far. Once I'm over this PhD hangover, I will be glad we never invoked the pact…

So many thanks to my main bae Jordanna. Every now and then I have got a little bit terrified and then I see the look in your eyes. You have made me laugh on so many occasions where laughing was the last thing I wanted to do. We'll never be wrong together. I am so proud of you and you are going to ace your PhD, I will be waiting for you on the other side of your viva with the prosecco and chambourd! Forever's gonna start tonight!

I would also like to thank all the friends I have made during this PhD. You have all helped me enjoy my time in Cardiff and I know I have made some friends for life.

And finally, but by no means least – my family. Mum, Dad and Thomas, you have always been there for me and had faith in me even when I didn't have any in myself. You have done so much for me and as a small gesture to show my appreciation, this thesis is dedicated to you…enjoy!

"At the end of the day it's another day over"

Les Mis

Abbreviations

5mC: 5-methylcytosine

α-CTF: α-carboxyl terminal fragment

Aβ: Beta Amyloid

Ach: Acetylcholine

AChE: Acetylcholinesterase

AD: Alzheimer's Disease

APOE: Apolipoprotein

APP: Amyloid Precursor Protein

AP2: Adaptor Protein 2

ASE: Allele Specific Expression

β-CTF: β-C Terminal Fragment

BACE1: β-site APP cleaving enzyme 1

BAR: BIN-amphiphysin/Rvs

BCA: Bicinchoninic Acid

BIN1: Bridging Integrator 1

BP: base pair

BSA: Bovine Serum Albumin

CCP: Clathrin coated pits

CCV: Clathrin coated vesicles

CD2AP: CD2-Associated Protein

cDNA: Complementary DNA

CEU: Northern European Population from Utah

ChIP: Chromatin Immunoprecipitation

CLAP: Clathrin and AP-2

CME: Clathrin Mediated Endocytosis

CNS: Central Nervous System

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

CSF: Cerebrospinal fluid

Dextran-647: Alexa Fluor® 647 Dextran Conjugate

DHS: DNase I Hypersensitivity Site

DMEM: Dulbecco's Modified Eagle Medium

DMSO: Dimethyl sulfoxide

DNA: Deoxyribonucleic Acid

dNTP: Deoxyribonucleotide triphosphates

DS: Down Syndrome

dsDNA: double stranded DNA

DSB: Double strand break

EMSA: Electrophoretic Mobility Shift Assay

ER: Endoplasmic reticulum

ELISA: Enzyme-Linked Immunosorbent Assay

ESCRT: Endosomal Sorting Complex Required for Transport

FACS: Fluorescent Assisted Cell Sorting

gDNA: Genomic DNA

GFP: Green Fluorescent Protein

GWAS: Genome Wide Association Study

HDR: Homologous directed repair

HWE: Hardy Weinburg Equilibrium

IGAP: International Genomics of Alzheimer's Project

ILV: Intraluminal vesicles

LARII: Luciferase Assay Reagent II

LB: Luria-Bertani

LD: Linkage Disequilibrium

LOAD: Late onset Alzheimer's disease

MAF: Minor Allele Frequency

MRI: Magnetic Resonance Imaging

mRNA: messenger RNA

MVB: Multivesicular bodies

NFT: Neurofibrillary tangles

NGS: Next Generation Sequencing

NHEJ: Non-homology end joining

NT: nucleotide

NT siRNA: Non- targeting siRNA

NTC: Non-transfected control

ON: Over night

OR: Odds ratio

PAM: Protospacer Adjacent Motif

PBS: Phosphate Buffered Saline

PCC: Pearson's Correlation Coefficient

PCR: Polymerase Chain Reaction

PI: Phosphoinositide

PIP_2: Phosphatidylinositol 4, 5-bisphosphate

PS: Proline-Serine

PSEN1: Presenilin 1

PSEN2: Presenilin 2

RNA: Ribonucleic Acid

RNP: Ribonucleoprotein

RT: Room Temperature

sAPP: Soluble APP

SEM: Standard Error of the Mean

sgRNA: short guide RNA

SH3: Src homology 3

siRNA: short interfering RNA

SNP: Single Nucleotide Polymorphism

ssODN: single-stranded DNA oligonucleotides

TBE: Tris-Borate-EDTA

TF: Transcription factors

Tf-488: Alexa Fluor® 488 Transferrin Conjugate

TGN: Trans-golgi Network

TSS: Transcription start site

v/v: volume/volume

w/v: weight/volume

CONTENTS

1. Introduction

1.1 Alzheimer's Disease

Alzheimer's disease (AD) is the most common neurodegenerative disease and the main cause of dementia (7). AD is the fifth leading cause of death in the United States in individuals aged over 65. An estimated 5.5 million Americans suffer from AD, 5.3 million aged over 65. Should no intervention be developed, the number of AD patients is predicted to triple by 2050 (8). Due to an ageing population and the full time care dementia patients require in the later stages of disease, dementia is becoming an increasing economic burden. The health and social care costs of dementia are greater than that for cancer, stroke and chronic heart disease combined (9). In the United Kingdom over 520 000 people have AD and the cost of dementia is predicted to increase from £26 billion in 2014, to £55 billion in 2040 (10).

Current treatments aim to manage symptoms rather than prevent onset or halt disease progression. A greater understanding of AD pathogenesis would aid in the understanding of the biology of this disease and in the development of disease modifying therapies that could tackle this worldwide health issue.

1.1.1 History of Alzheimer's disease

Alois Alzheimer first described AD in 1906. Alzheimer examined a 51-year-old female who presented with severely impaired memory, aphasia, erratic behaviour, paranoia and auditory hallucinations. The post-mortem revealed the presence of extracellular neuritic plaques comprising of the amyloid beta peptide (Aβ), intracellular neurofibrillary tangles of tau protein (NFT) and

atherosclerotic changes in the brain (11, 12). These symptoms did not fit into a known disease at the time. The term "Alzheimer's disease" was first used by Kraepelin, often regarded as the father of modern psychiatry, in the 8th edition of his "Handbook of Psychiatry" in 1910 (13).

Little research on AD was carried out until the development of biochemical techniques in the 1970s, which led to an increased interest in the cholinergic system. In 1977, a cholinergic deficit in AD was confirmed (14, 15) and the discovery of selective cell death of cholinergic neurons in the nucleus basalis of Meynert provided further evidence for a cholinergic role in AD. This discovery linked the biochemical observations and the disease pathology for the first time (16). The identification of the cholinergic systems role in AD led to the first therapies to be developed: anti-cholinesterase inhibitors.

Work by Braak and Braak first identified the various stages in pathology of the disease in 1991 (17). Since then the development of more sophisticated neuroimaging techniques able to visualise neuronal atrophy *in vivo*, along with the use of PET ligand methods capable of imaging amyloid deposition and tau tangles in the brain, has allowed the study of AD pathology even in the early stages of the disease (18-20).

1.2 Clinical features of Alzheimer's disease

The average time course of AD dementia is 7-10 years and concludes with death. The symptom most commonly observed in the first instance is impairment of recent memory, however other cognitive deficits may also present, such as executive dysfunction and problem solving abilities. As the disease progresses to mild AD, which takes around 2-5 years, further symptoms become apparent such as language dysfunction, visuospatial difficulties, loss of insight and personality changes.

The moderate stage of disease (lasting between 2-4 years) displays more obvious problems with memory, affecting long-term memory. It also presents with increased cognitive problems resulting in the individual not being able to act independently in the community or perform routine tasks.

In the severe stages of AD, individuals require full time caregivers and are totally dependent upon them to complete all activities of daily living. In the most advanced stage of disease, individuals can become mute, non-ambulatory, unable to talk and cannot control bladder or bowel function (21).

1.2.1 Diagnosis

In order to conclusively diagnose AD, patients need to present with the clinical phenotype of progressive dementia with impairment of other cognitive domains (such as learning, language, reasoning, visuospatial abilities or behaviour), and the neuropathological hallmarks such as NFTs, amyloid plaques and neuronal cell death. As neuropathological investigations are difficult during life, the diagnosis of AD is predominantly influenced by clinical

evidence and takes a probabilistic approach, therefore only a diagnosis of probable AD is possible.

The criteria to describe the clinical diagnosis was first established in 1984 by the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (22). This report takes into account medical history, clinical examination, neuropsychological testing and laboratory assessments. Since 1984, the knowledge and understanding of AD has increased, leading to the need to revise these original criteria. The National Institute on Ageing – Alzheimer's Association (NIA-AA) have proposed revised guidelines that include the use of biomarkers to aid a diagnosis (23). A summary of these guidelines is described below.

The NIA-AA define a diagnosis for **dementia** as neuropsychiatric symptoms that:

- Interfere with an individuals ability to function at usual activities
- Have declined from previous levels of function and performance
- Are not explained by another psychiatric disorder
- Resulted in cognitive impairment that involves two of the following:
    o Impaired ability to acquire and remember new information
    o Impaired reasoning and handling of complex tasks/poor judgment
    o Impaired visual spatial abilities
    o Impaired language function
    o Changes in personality/behaviour

The NIA-AA have classified a clinical diagnosis of dementia as a result of AD (AD dementia) diagnosis as Probable or Possible AD. A diagnosis of **Probable AD dementia** is reached when:

- A patient meets the diagnosis for dementia previously described. In addition, the dementia must have the following characteristics:
    - Insidious onset
    - History of cognitive decline
- No evidence of cerebrovascular disease
- No evidence of features of Dementia with Lewy Bodies
- No evidence for Frontotemporal Dementia
- No evidence of progressive aphasia
- No evidence of a medical comorbidity that could have a substantial effect of cognition

An individual can gain a diagnosis of **Probable AD dementia with increased level of certainty** if there is additional evidence of a documented decline or a causative AD genetic mutation. Recent developments in the understanding of AD biomarkers can be used to strengthen a Probable AD diagnosis. Biomarkers include low cerebrospinal fluid (CSF) beta amyloid 42 (Aβ42) and positive PET amyloid imaging, which are indicative of Aβ deposition in the brain, and elevated CSF tau, decreased flurodeoxyglucose uptake on PET in temporoparietal cortex and evidence of atrophy identified via MRI, which are indicative of neuronal degeneration (24).

A **Possible AD dementia** diagnosis is received if an individual meets the clinical criteria of AD dementia but has:

- An atypical disease course
- Evidence of a comorbidity that could effect cognition

A diagnosis of Possible AD dementia with evidence of the AD pathological process would be given to patients who meet the clinical criteria for a non-AD

dementia but possess biomarker evidence for AD or who meet the neuropathological criteria for AD.

1.2.2 Current treatments

There are two categories of drugs currently available to treat the symptoms of AD: acetylcholinesterase (AChE) inhibitors and NDMA (N-methyl-D-aspartate) receptor antagonists.

Acetylcholine (ACh) is a neurotransmitter implicated in learning and memory (25). AChE is an enzyme, which breaks down ACh at the neuromuscular junction to terminate synaptic transmission in cholinergic neurons (26). AD patients have decreased acetylcholine synthesis and altered levels of acetylcholine receptors (27, 28). AChE inhibitors prevent the breakdown of acetylcholine in an attempt to restore/maintain levels in the brain (29). Three AChE inhibitors are commonly prescribed; rivastigmine (Exelon), galantamine (Razadyne, Reminyl) and donepezil (Aricept). AChE Inhibitors have been shown to temporarily improve symptoms and reduce their decline. Individuals taking AChE Inhibitors can have reduced anxiety and improvements in motivation, memory, concentration and ability to perform daily tasks (30).

Glutamate is an essential neurotransmitter in the central nervous system, primarily functioning in the hippocampal and neocortical regions of the brain. It is critical for cognition, learning and memory. NDMA receptors are enriched in the post-synaptic membrane and are a receptor for glutamate (31). Over stimulation by glutamate can increase intracellular calcium, which can impact on calcium homeostatic mechanisms and result in neuronal dysfunction and cell death (32). This process was first described in 1969 and was termed

excitotoxicity (33). An increased level of glutamate is seen in AD, which causes increased activation of the NDMA receptors resulting in excitotoxicity (34).

Memantine (Namenda) is a NDMA receptor antagonist preventing the effect of elevated glutamate levels that can lead to neuronal dysfunction. Memantine is prescribed to patients intolerant to AChE Inhibitors or with severe AD. Individuals taking Memantine can see a reduced decline of their symptoms, including disorientation and ability to perform daily tasks. There is some evidence suggesting Memantine may also alleviate additional symptoms such as delusions, aggression or agitation (30).

1.3 Pathology

1.3.1 Amyloid Plaques

Senile amyloid plaques observed in the brain of AD patients are extracellular deposition of Aβ peptides. Aβ typically consists of 40 or 42 amino acids residues and is cleaved from Amyloid Precursor Protein (APP). Aβ is predominantly generated at the neuronal membrane, although other cell types may contribute, and released into the extracellular space (35). Aβ40 is the most abundant Aβ species however Aβ42 is hydrophobic and more prone to aggregating and is consequently more abundant in amyloid plaques (36). Aβ42 was shown to be capable of causing the hyperphosphorylation of tau and neuritic dystrophy in cultured rat neurons (37) and the neurotoxic effect of the Aβ peptide in the hippocampus of healthy rats resulted in impaired memory (38). Aβ deposition appears to precede other AD-associated pathology (39, 40). Aβ plaques are always observed in AD and result in the generation of dystrophic neurites and glial activation (41, 42).

The progression of amyloid plaque deposition is fairly unpredictable, but three stages have been distinguished. In stage one, plaques are found in the basal portions of the frontal, temporal and occipital lobes. Stage two is characterised by all of the isocortical areas being affected, with mild plaque formation in the hippocampus and a lack of plaques observed in the primary sensory, motor and visual cortices. The third and final stage sees amyloid plaques in the primary isocortical areas and potentially in the cerebellum and subcortical nuclei (17). Amyloid plaque burden does not appear to correlate with dementia severity or duration (43-46).

There are two morphological categories of amyloid plaque, diffuse and dense-core. Dense-core amyloid plaques are predominantly found in the brains of AD patients whereas diffuse amyloid plaques have been observed in the brains of healthy individuals, therefore it is possible they may be a result of normal ageing. The neuropil is a dense network of cytoplasmic processes from nerve cells and glial cells. Dense-core amyloid plaques are associated with additional pathological events in the surrounding neuropil, such as increased neurite curvature, dystrophic neuritis, the recruitment and activation of astrocytes and microglia and synaptic and neuronal loss (47-54). Synaptic loss is an early event in the progression of AD (55).

Dense-core amyloid plaques comprise of a central mass of extracellular filaments that radiate outwards and come into contact with neuronal, astrocytic and microglial processes, forming neuritic dystrophies (54, 56). Plaques can have cytotoxic effects and reactive astrocytes and activated microglia are associated with dense-core amyloid plaques suggesting plaques potentially trigger the glial response (47, 50, 53). Furthermore, pathogenic soluble forms of Aβ have been implicated in inducing synaptic dysfunction and synapse loss, critical features of AD pathogenesis (57).

1.3.2 Neurofibrillary tangles

NFTs primarily consist of pairs of helical filaments consisting mainly of the hyperphosphorylated tau protein (58). Tau is a microtubule-associated protein located in the axon and plays a role in the stability of neuronal microtubules which function in axonal transport (59). If tau is abnormally modified, it may dissociate from the microtubule, causing these structures to collapse resulting in pathological lesions. Abnormally phosphorylated tau (P-tau) has been identified early in AD brains (60).

There are three distinguishable morphological stages of NFTs. Pre-NFTs are defined as having diffuse tau within the cytoplasm of otherwise normal looking neurons. Mature intraneuronal NFTs show filamentous aggregates of tau that displaces the nucleus and extraneuronal NFTs result from the death of a tangle bearing neuron (61, 62).

The pattern of NTF progression is predictable and has been described in six stages. NFTs initially appear in the transentorhinal region and entorhinal cortex in stage one. Stage two sees the appearance of NFTs in the CA1 region of the hippocampus. NFTs develop in the limbic structures such as the subiculum of the hippocampus formation during stage three and the amygdala, thalamus and claustrum in stage four. In stage five, NFTs develop in the isocortical areas and finally in the primary sensory, motor and visual areas in stage six (17, 63, 64).

NFT burden and distribution appears to correlate with the severity and duration of the dementia. Furthermore, the spreading of the NFTs between brain regions relates to the neuropsychological symptoms experienced (43, 45, 46). A correlation between NFT burden and astrocytosis and microgliosis has been observed independent of plaque burden, suggesting that NFT may also influence the glial response (46, 65).

1.3.3 Other features

AD displays symmetrical cortical atrophy predominantly affecting the medial temporal lobes whilst sparing the primary motor, sensory and visual cortices and is detectable via magnetic resonance imaging (MRI) during the early stages of disease (66). Neuronal and synaptic loss typically progresses in

parallel with NFTs, however as neuronal loss can exceed NFT burden, another mechanism of neuronal death may be involved (67-69).

Synapse loss has been observed to exceed neuronal loss, indicating that the loss of the synapses occurs prior to neuronal death. These surviving neurons have reduced connectivity with the surrounding neurons. Eventually synapse loss can lead to neuronal death and contribute to the cortical atrophy seen in AD (70-73).

Aβ can also deposit in vessel walls forming cerebral amyloid angiopathy, seen in approximately 80% of AD brains. Granuovacuolar degeneration of Hirano Bodies, intracellular aggregates of actin and associated proteins observed in neurons, has been detected in the brains of cognitively healthy elderly individuals, but are more frequent and severe in AD patients (74, 75).

1.4 APP processing and Aβ generation

Aβ is generated from the proteolytic cleavage of APP. Following APP synthesis in the endoplasmic reticulum (ER) APP is transported to the cell surface via secretory vesicles. α-secretase is a membrane bound endoprotease, thought to cleave APP primarily at the plasma membrane (76). α-secretase cleaves APP within the Aβ domain leaving a α-carboxyl terminal fragment (α-CTF) and releases the large soluble ectodomain sAPPα fragment. The ADAM protein family possesses α-secretase like activity, and ADAM10 and ADAM17 have been suggested to be active in the processing of APP (77, 78). The α-CTF fragment undergoes further cleavage by a γ-secretase to produce a p83 fragment, which is rapidly degraded and thought to have little function. Due to the α-secretase cleavage occurring in the Aβ domain, this prevents the production of the Aβ peptide and therefore this pathway is referred to as the non-amyloidogenic processing pathway (79).

Following transportation to the cell surface, APP can be re-internalised and enter the endosomal/lysosomal system, where it is thought amyloidogenic processing and Aβ generation occurs (80). Amyloidogenic processing occurs when APP is cleaved by a β-secretase, such as membrane bound BACE1 (β-site APP cleaving enzyme 1) (81). BACE1 is abundant in neurons and requires the acidic environment of the early/late golgi, early endosomes and endosomes for optimal activity and this is thought to be the rate-limiting factor in the generation of Aβ (82, 83).

β-secretase cleavage produces a soluble sAPPβ fragment and a membrane bound β-carboxyl terminal fragment (β-CTF) (81). γ-secretase cleaves β-CTF in the transmembrane protein domain, which mainly occurs in the trans-golgi

network (TGN) and early endosomes, with some evidence it may also occur in multivesicular bodies (MVB) (84, 85). The γ-secretase complex consists of 4 components: presenilin, Nicastrin, anterior pharynx defective 1 and presenilin enhancer 2 (86). γ-secretase cleavage generates Aβ and releases the APP intracellular domain (AICD). AICD can translocate to the nucleus, where it may regulate gene expression. Aβ40 or Aβ42 are generated depending on the γ-secretase cleavage site. Aβ peptides can assemble into amyloid fibrils. This is thought to occur via a nucleated polymerisation model, where an Aβ monomer acts as a nuclei from which fibrils grow (36). Amyloid fibrils are large and insoluble and form the amyloid plaques seen in AD brains. A summary of APP processing is shown in figure 1.1.
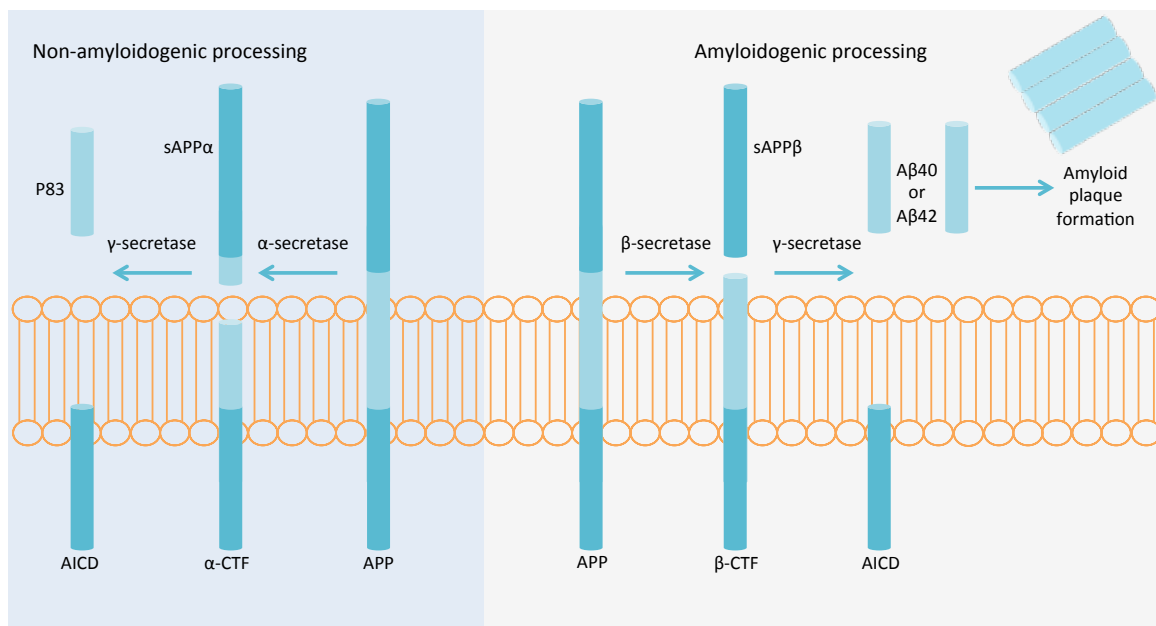
Figure 1.1. Non-amyloidogenic and Amyloidogenic APP processing. Non-amyloidogenic processing of APP occurs at the cell surface. Cleavage by α-secretase prevents the production of Aβ peptides. Amyloidogenic processing of APP occurs following APP internalisation. β-secretase cleaves APP in the endosomes to produce β-CTF and sAPPβ fragments. β-CTF is further cleaved by γ-secretase to generate Aβ peptides, which deposit to form amyloid plaques.

sAPPα has been found to have an important role in neuronal plasticity and survival and be protective against excitotoxicity (87, 88). Expression of sAPPα appears to alleviate some of the abnormalities seen in APP knockout mice, suggesting that a physiological function of APP may be mediated via sAPPα (89). In contrast to sAPPα, sAPPβ appears to have a detrimental effect and has been implicated in axonal pruning and neuronal cell death (90). Overexpression of β-CTF is thought to have a cytotoxic effect and has been linked to neuronal degeneration (91, 92).

1.5 Genetics of Alzheimer's Disease

Due to technological limitations, the genetics underlying AD was under investigated until the 1980s, when increasing numbers of autosomal dominant pedigrees were identified (93). The association between Down's syndrome (DS) (trisomy of chromosome 21) and AD was key in these initial genetic studies (94). Linkage studies and the discovery that amyloid plaques peptides were cleaved from APP, identified *APP* located on chromosome 21 as a candidate gene (95). Linkage studies also led to the identification of a novel gene on Chromosome 14 associated with early onset disease now known as presenilin 1 (*PSEN1*) (96). Studies in late onset familial and sporadic cases identified the association between gene dose of the apolipoprotein E (*APOE*) type 4 alleles and risk of AD in these cases (97, 98). Only with the recent development of genotyping chip technology were any further genetic associations identified.

1.5.1 Familial AD

Less than 5% of all AD cases occur in families in recognisable Mendelian inheritance patterns. Familial AD has an autosomal dominant pattern of inheritance and symptoms typically present before the age of 65. 10-15% of familial AD cases are estimated to have *APP* mutations, 70% of cases are estimated to have *PSEN1* mutations and 5% of cases are due to rare *PSEN2* mutations. The remaining familial AD patients are thought have *de novo* mutations (99, 100). Mutations in these genes generally result in increased production of Aβ42 and an earlier disease onset (101).

1.5.1.1 Evidence linking APP *to etiology of AD*

Investigations into the composition of amyloid plaques led to the identification of *APP* on chromosome 21q21 in 1987 (102). This discovery explained why individuals with DS develop the neuropathological features of AD in their fourth decade, as these individuals have three copies of *APP* (103). Individuals who have trisomy 21, but due to a translocation only have two copies of *APP*, have DS features but do not develop AD (104). Additionally, individuals with an *APP* duplication, but not trisomy 21, do not have DS but develop AD during their fifth decade (105). These findings all implicate overexpression of *APP* in the increased production of Aβ resulting in AD pathology.

Over 30 missense mutations have been identified in *APP* that result in familial AD, with most occurring within the secretase cleavage site of the transmembrane domain and adjacent Aβ domain and leading to increased Aβ production (106, 107). *APP* mutation carriers have an average age of onset of 45-55 (108). Missense mutations in the Aβ region of *APP* have also been identified which reduce β-secretase cleavage and therefore reduce the risk of developing AD (109).

1.5.1.2 Evidence linking Presenilin to etiology of AD

Presenilin genes encode for PSEN1 and PSEN2, which form a heterodimer that functions as the catalytic domain of the γ-secretase complex, the transmembrane cleaving enzyme responsible for the production of Aβ (110-112).

PSEN1

PSEN1 (Presenilin 1), located on chromosome 14q24.2, was identified as a pathogenic gene for Familial AD in 1995 (96) and encodes for the catalytic core of the γ-secretase complex (110). Missense mutations in *PSEN1* have complete penetrance and are the most common cause for familial AD, accounting for 18-50% of cases (113). The average age of onset of AD for individuals carrying a *PSEN1* mutation is 35-55 years, but can occur as early as mid-20s (108, 114). Missense mutations in *PSEN1* result in relatively increased levels of hydrophobic aggregate-prone Aβ42 peptides compared to Aβ40. This may be a result of either increased production of Aβ42 or decreased production of Aβ40 (115). Aβ42 deposition may be a preclinical event that occurs in *PSEN1* mutation carriers (116).

PSEN2

PSEN2 (Presenilin 2), located on chromosome 1q42.13, has high homology to *PSEN1* and was identified as a candidate gene in familial AD in 1995 (117). *PSEN2* mutations are relatively rare in the Caucasian population and result in a later and more variable age of onset between 40-70 years (108). *PSEN2* mutations also have lower penetrance than *PSEN1* mutations suggesting that environmental or other genetic factors may influence disease (118). *PSEN2* mutations have been associated with a relative increase in Aβ42, implying it also affects the cleaving activity of γ-secretase (115).

1.5.2 Sporadic AD

The vast majority of AD cases are sporadic and therefore do not run in families. Although sporadic AD can have an early onset, in the majority of cases onset occurs after the age of 65 and is referred to as Late Onset AD (LOAD). LOAD is a complex disease and has a heritability estimate of 60-80% (119). First degree relatives of an individual suffering from LOAD have a life time risk of 20-25%, in comparison to the general population which have a 10.4% risk (120), indicating genetics play an important role in this disease.

1.5.2.1 APOE

Apolipoprotein E (*APOE*), located on chromosome 19q13.32, has been implicated in both familial AD presenting with late onset and LOAD. *APOE* has three major allele variants, which can be distinguished by two SNPs (single nucleotide polymorphisms), rs429358 and rs7412. These allele variants are termed ε2, ε3 and ε4. The *APOE* ε4 allele was discovered as a risk gene for LOAD in 1993 (98). The dosage of the *APOE* ε4 allele is a major risk factor for LOAD, effecting both age of onset and cognitive decline (121, 122). Females with a single ε4 allele have a 4-fold increase in risk of developing LOAD, whereas men and women with two ε4 alleles show a 15-fold increase in risk (123-125). The ε2 allele is thought to be protective against AD risk and delay age on onset (126), however this allele has been linked to cardiovascular disease (127).

APOE is involved in lipid homeostasis. In the central nervous system APOE is primarily produced from astrocytes and transports cholesterol to neurons via APOE receptors (128, 129). There is evidence suggesting ε4 carriers may have diminished cholesterol transport in neurons and astrocytes (130). APOE is also capable of binding Aβ and functions in the clearance of Aβ aggregates and

soluble Aβ from the extracellular space. The ε3 allele has greater binding capacity resulting in less deposition, whereas the ε4 allele is thought to be less efficient in mediating clearance (131, 132). Furthermore, ε4 carriers have a higher plaque and tangle burden than other *APOE* polymorphisms (133).

1.5.2.2 Genome Wide Association Studies

Due to developments in sequencing technologies and the creation of chip based assay technology capable of genotyping millions of SNPs, Genome Wide Association Studies (GWAS) were made possible. GWAS are based on the hypothesis that common diseases are influenced by common genetic variation. GWAS use a case-control approach to determine allele frequency differences between healthy individuals and individuals suffering from a disease. This approach identifies genetic variation that is common in the population which may have a relatively small effect size in comparison to rare disorders, such as familial AD (134).

In October 2009, Harold *et al* published a GWAS of LOAD (135). A collaborative consortium was formed with parties based in Europe and the US termed Genetic and Environmental Risk in Alzheimer's Disease 1 Consortium (GERAD1). This allowed a combined collection of 19000 subjects to be studied. Following stringent quality control measures, 16000 subjects remained, making this study the largest GWAS to take place at that time. This study showed highly significant associations, $p=8.5 \times 10^{-10}$ and $p=1.3 \times 10^{-9}$, for *CLU* and *PICALM* loci respectively. This study also generated suggestive evidence for an association between *BIN1* and LOAD, however this did not reach genome-wide significance ($p=3.2 \times 10^{-6}$).

Simultaneously another LOAD GWAS was carried out by Lambert *et al* as part of the European AD Initiative Stage 1 (EADI1) (136). This study replicated the findings reported in Harold *et al*, and identified two genome wide significant markers at the *CLU* locus (p<9.3 x 10[-8]). Lambert *et al* also discovered another locus of interest that lies within a linkage disequilibrium block containing *CR1* (p=3.5 x 10[-9]).

In 2010, a three-stage analysis of GWAS data was performed studying 35000 participants, including 8371 AD cases. The aim was to identify additional loci associated with LOAD and replicate previous findings in an independent case-control population (137). This 3 stage analysis identified genome-wide significant associations between LOAD and the *APOE*, *CLU*, *PICALM* loci and two novel loci, one on chromosome 2 upstream of *BIN1* and another on chromosome 19 located near *BLOC1S3* and *MARK4*.

In 2011, Hollingworth *et al* aimed to identify more new susceptibility variants for AD by performing a three-stage association study based on the European population within the GERAD+ database and testing the loci with suggestive evidence for association in the Alzheimer's Disease Genetic Consortium (ADGC) GWAS (138). This analysis identified SNPs at the *BIN1*, *CR1*, *ABCA7* and *MS4A* loci with genome wide significance.

In a similar study The Alzheimer's Disease Genetic Consortium assembled a dataset for association analysis. The Stage 3 analyses strengthened the evidence for association with *CD33*, *EPHA1* and *CD2AP*. *CD2AP* gained genome-wide significance in this analysis with a p value of 8.6 x 10[-9] and the evidence of association for the *EPHA1* loci increased with a p-value of 6.0 x 10[-10] (139).

By 2011 ten LOAD susceptibility loci had been identified: *APOE*, *CR1*, *CLU*, *PICALM*, *BIN1*, *EPHA1*, *MS4A*, *CD33*, *CD2AP* and *ABCA7*. In 2013 the largest LOAD GWAS meta-analysis took place. The International Genomics of Alzheimer's Project (IGAP) conducted a two-stage meta-analysis using four GWAS samples with European ancestry (140). Stage 1 meta-analysis used data from ADGC, CHARGE, EADI and GERAD and used the European population reference haplotype data from the 1000 Genome Project to impute genotypes for up to 11863202 SNPs per data set. In addition to *APOE*, 14 genomic loci were identified to have a genome-wide significant association in stage 1. This included five newly associated loci; the genes closest to the most significantly associated SNPs were *HLA-DRB5-HLA-DRB1*, *PTK2B*, *SORL1*, *SLC24A4-RIN3* and *DSG2*. This study confirmed the previous findings of *ABCA7*, *BIN1*, *CD33*, *CLU*, *CR1*, *CD2AP*, *EPHA1*, *MS4A6A-MS4A4E* and *PICALM*. All SNPs found to have a p-value less that 1×10^{-3} were genotyped in stage 2. All loci identified in stage 1 reached genome-wide significance in the stage 2 replication, except for *CD33* and *DSG2*.

Seven new loci were identified in a combined analysis of stage 1 and stage 2 data. The strongest new association is from within an intron in *ZCWPW1*, however, this region of association covers around 10 genes therefore this hit is not necessarily the functional gene. Other novel loci identified includes *CELF1*, *NME8*, *FERMT2*, *CASS4*, *INPP5D* and *MEF2C*. This combined analysis also identified 13 loci with suggestive evidence of association for example SNPs downstream of *TREML2* and upstream of *TREM2*. This meta-analysis identified 11 new susceptibility loci and confirms the association of nine previously identified loci.

1.5.2.3 Post-GWAS studies

Following the success of GWAS, these data were used in different approaches to identify additional loci associated with LOAD. Two loci were identified via a gene wide burden analysis. This approach takes into account multiple variants at a single locus and therefore investigates more complex patterns of association. Using the IGAP data set, 2 novel loci were identified, *TP53INP1* (p=1.4x10^{-6}) and *IGHV1-67* (p=7.9 x 10^{-8}) (141).

When IGAP data and data from an independent sample were combined, *TRIP4* was identified as a LOAD risk locus (p=9.74 x 10^{-9}) (142). *PLD3* has been identified via burden analysis, suggesting multiple variants within this locus may affect LOAD susceptibility (143).

Despite the growing numbers of common variants associated with LOAD, it is estimated that 60% of the genetic component of LOAD is not accounted for by this common variation and *APOE* (144, 145). This missing heritability is more likely explained by variation with such a small effect size it is not currently detected in these studies, and rare variants with greater effect size, which are not represented in the current chip based technologies.

Next Generation Sequencing (NGS) technology, as oppose to SNP chip genotyping, has been used in an attempt to identify rare variants with a substantial effect size. Variants in exon 2 of *TREM2* were found to be overrepresented in LOAD patients and went on to be replicated in imputed GWAS data, independent samples and in an Icelandic population (146, 147). There are a number of additional loci with suggestive evidence of association, but, due to sample size and cost restrictions of NGS, are yet to reach significance: *UNC5C*, *ADAM10*, *ZNF628* and *AKAP9* (148-151).

In 2017 an approach for rare variant discovery used microarrays that targeted known exome variants. Secondary analysis used genotypes and imputed genotypes from independent samples. This three-stage case-control study used over 85000 subjects and identified three new genome wide significant nonsynonymous variants associated with AD. A new variant was found in *TREM2* (p=1.55 x 10^{-14}, OR=1.67), a protective variant was found *PLCG2* (p=5.38 x 10^{-10}, OR=0.68) and a risk variant was found in *ABI3* (p=4.56 x 10^{-10}, OR=1.43). These genes are all implicated in the innate immune response and are highly expressed in microglia, suggesting these functional pathways may contribute to AD susceptibility (152).

It is proposed that the additive or multiplicative effects of associated genes with one another could explain some of the missing heritability for LOAD. As the understanding of the genomic landscape progresses, the interaction between genes and variants and the contribution of this interaction to disease risk remains unclear.

Most models of AD risk do not consider the effects of multiple variants. GWAS assumes an additive model, where each SNP confers disease risk independent of each other. Models accounting for multiple variants are being developed. Polygenic risk scores (PGRS) use genomic profiles, which combine the effects of many associated genetic variants to predict risk of disease whereas multiplicative models of risk are able to investigate the relationship between genes (epistasis).

PGRS aggregate information across multiple variants to create a unique risk score that aims to predict risk based on a variant profile per individual. This approach can identify phenotypic associations that would not be detectable

using single loci with low effect sizes. These scores may increase the predictive power of genotyping and could be a used for personalised disease profile of an individual.

A PGRS model for AD that included APOE ε2, ε4, age, sex, the IGAP SNPs and SNPs with an AD association <0.05 was able to predict disease status with 78% accuracy (153). Although the low threshold for SNP inclusion added noise to the model, a higher power was achieved by including SNPs with even low level risks. The accuracy of this model was increased to 84% when using a pathologically confirmed cohort (154). It is thought that this model captures almost 90% of the SNP heritability that can actually be predictive for LOAD risk, even though this variation remains unknown (155). AD PGRS have been found to be associated with the risk of familial late-onset AD, accelerated progression from mild cognitive impairment to AD, cognitive scores and neuroimaging measures (156-162). PGRS will also have applications in identifying subjects at high risk of developing AD and aid in diagnosis.

Further explanation for the unexplained genetic variance for a particular trait, is gene-gene interactions (epistasis). Gene interactions are critical for biological processes such as gene regulation and signal transduction. Epistasis measures the interactive effects between one gene/variant and one or more other genes/variants. If a locus is examined as a single entity, the potential interactions are not taken into account and therefore the full contribution to disease risk may be missed.

A number of AD relevant gene-gene interactions were established before GWAS identified AD-associated genes. Interactions *IL6* and *IL10*, genes acting in the interleukin pathway, were identified (163) and interactions between *TF*

and *HFE*, implicated in the iron transport pathway, have been reported (164). Both molecular mechanisms have been implicated in LOAD.

Since GWAS, further epistatic interactions with AD-associated have been identified. Significant interactions were identified between GWAS identified LOAD genes *CLU-MS4A4E* and *CD33-MS4A4E* (165, 166). Furthermore, epistatic effects within known LOAD loci have identified associations with amyloid deposition (*BIN1-PICALM*) (167) and brain atrophy (168). Epistatic interactions may also contribute to the explanation of why some genetic effects are not replicated when examined in isolation.

1.5.3 AD genetics implicate disease mechanisms

1.5.3.1 Amyloid cascade hypothesis

Proteolytic cleavage of APP can result in the generation of Aβ peptides and therefore is fundamental in AD pathology. Evidence from the genetic studies of familial AD implicates APP proteolysis in the pathogenesis of AD. APP, PSEN1 and PSEN2 [discussed in chapter 1.5.1] are all vitally important in Aβ generation. Furthermore, the strongest risk factor for LOAD, *APOE*, has been implicated in Aβ clearance from the brain. Mutations in these genes can result in aggressive forms of familial AD as previously described. These observations led to the proposal of the Amyloid Cascade Hypothesis in 1992 (169, 170). The Amyloid Cascade Hypothesis suggested that the amyloid plaques seen in LOAD are the causative pathology and the subsequent NFTs, cell loss, vascular damage and dementia are a consequence of the plaque deposition. The hypothesis has since been updated to take in to account the importance of the soluble oligomers and is summarised in figure 1.2 (171).
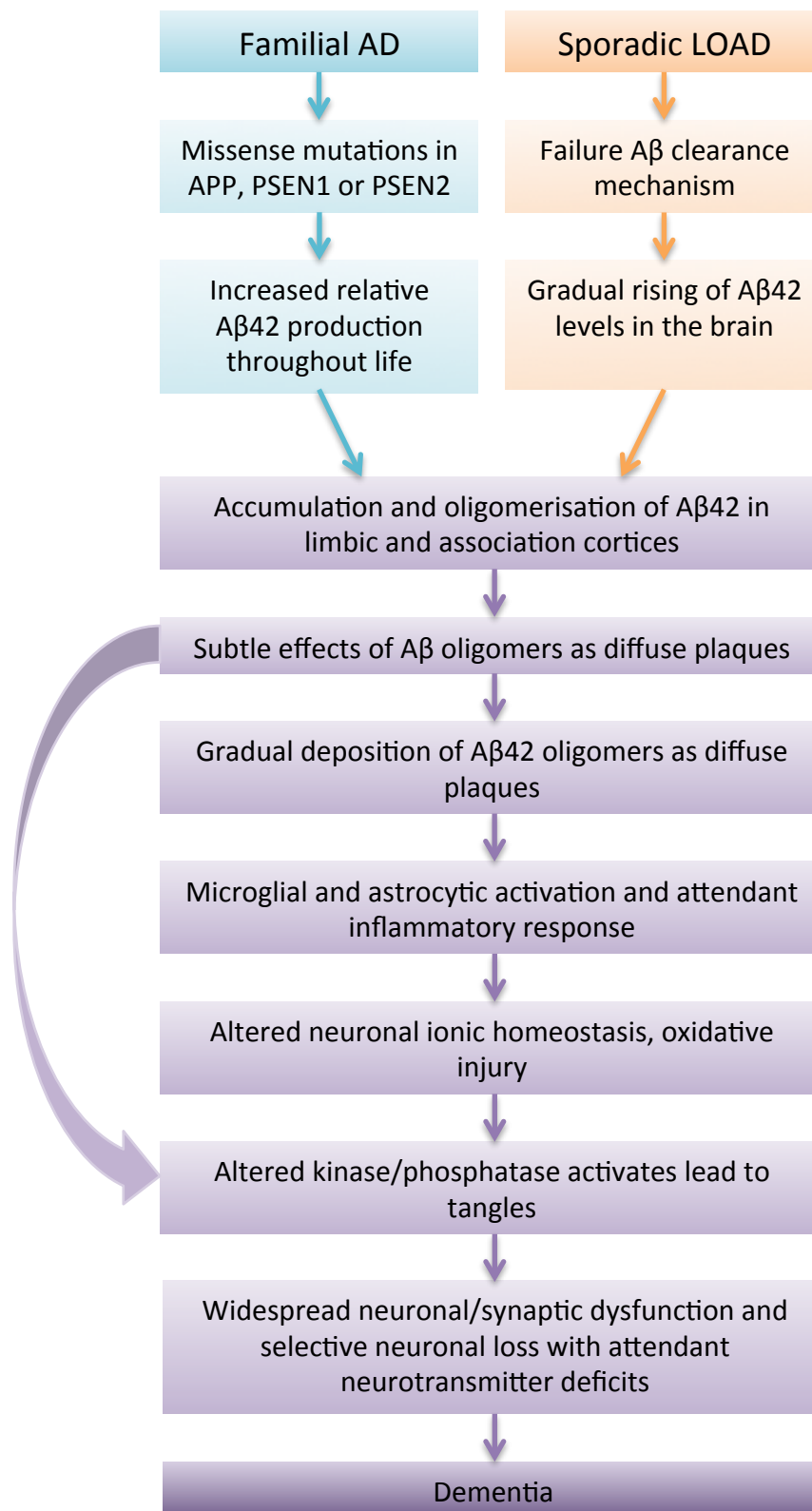
Figure 1.2. Amyloid Cascade Hypothesis. This hypothesis is the dominant model of AD pathogenesis. The diagram describes the progression of hypothesised pathogenic events resulting in AD. The arrow indicates that the Aβ oligomers may also have a direct effect on synapses and neurons in the brain, suggesting this cascade may not be truly linear. (Adapted from (2)

However, the amyloid cascade hypothesis remains controversial with evidence contradicting it. For example, Aβ plaques do not correlate with cognitive deficits as well as NFTs (172) and are often seen in the brains of cognitively normal people upon death (173, 174). Additionally, numerous therapies targeting Aβ have failed in treating this disease (175). These repeated failures are thought to be due to insufficient understanding of the drug mechanism, concerns over the drug-like properties of the compounds tested and whether patients in clinical trials are too advanced to benefit from therapeutic intervention (176, 177). Despite its controversies, the amyloid cascade hypothesis remains to be one of the leading suggested pathogenic mechanisms in AD and a vast amount of research surrounding it has been carried out.

Furthermore, the genetics of familial AD implicates a mechanistic link between Aβ and tau. Pathogenic mutations in genes involved in APP processing seen in familial AD lead to increased Aβ production and the development of AD, including tau pathology as observed in sporadic AD. However, mutations in *MAPT*, which encodes tau, does not cause familial AD, but can lead to other tauopathies (178), suggesting tau pathology alone is not sufficient to develop AD. It is therefore likely the Aβ generation initiates the disease process and drives tau pathology (179).

1.5.3.2 Pathway analysis

Pathway analysis is an experimental approach that aims to determine whether biological pathways are enriched in loci associated with a particular trait. A pathway analysis performed using the IGAP genetic association data used significantly associated loci with LOAD and loci that had failed to reach significance. This analysis identified a number of biological pathways enriched

in these loci: immune response, regulation of endocytosis, cholesterol transport, proteasome-ubiquitin activity, hematopoietic cell lineage, hemostasis, clathrin/AP2 adaptor complex and protein folding (180). This investigation identified cellular pathways that provide clues to the functional mechanisms involved in AD pathogenesis.

1.6 Alzheimer's Disease and Endocytosis

Dysregulation of endocytosis is associated with AD and other neurodegenerative disorders (181). Aberrant neuronal endocytosis is one of the earliest neuropathological changes in AD (182). Abnormal endosome morphology can be seen in DS decades before the onset of AD and precedes Aβ deposition (182-184). In an AD brain, the abnormal endosome morphology is a result of Aβ42 accumulation and the formation of Aβ oligomers within the MVB/late endosomes of neuronal processes and synaptic compartments (185). Endocytosis is critical for the amyloidogenic processing of APP, synaptic activity and neurotransmitter release, all of which are dysregulated in AD (21).

In addition to endosomal abnormalities, there is an increase in autophagic vacuoles, particularly in dystrophic neurites. This build up of autophagic vacuoles results in an accumulation of undigested or partially digested proteins throughout the brain, which is associated with neurodegeneration (186, 187). This build up of waste protein suggests that the efficient autophagic processes and lysosomal proteolysis may be defective in AD (188).

A number of genes significantly associated with LOAD are thought to have a role in endocytosis: *BIN1*, *PICALM*, *CD2AP*, *EPHA1* and *SORL1* (135, 138-140). The pathway analysis using GWAS data implicated endocytosis and clathrin/AP2 adaptor complex as functional pathways enriched with genes associated with LOAD, indicating these pathways must have a critical role in LOAD pathogenesis (180).

1.6.1 Clathrin Mediated Endocytosis and cellular trafficking

Endocytosis is vital for the function and survival of a cell. There are several pathways of endocytosis into the cell, including clathrin-dependent, caveolin-dependent and clathrin- and caveolin- independent internalisation. As APP is internalised via clathrin-mediated endocytosis (CME), this pathway has been most studied in relation to AD (189-191).

CME is the most prominent form of endocytosis that constitutively occurs in all mammalian cells. CME continuously transports nutrients, signaling receptors, low-density lipoprotein particles, cell adhesion molecules, ion channels and ion transporters via receptors into the cell (192, 193). CME is initiated by the formation of clathrin-coated pits (CCP) and clathrin-coated vesicles (CCV). Phosphatidylinositol 4, 5-bisphosphate (PIP_2) rich zones are generated in the plasma membrane. These PIP_2 rich zones contain high affinity receptors and their bound ligands. Adaptor Protein-2 (AP2) is recruited to the PIP_2 rich zones, which consequently recruits and binds triskelia, comprising of clathrin heavy chains and clathrin light chains (194). Triskelia assemble together to form a clathrin cage that functions to induce membrane curvature and the formation of CCPs (195-197). This membrane curvature invaginates the plasma membrane and once the CCP is sufficiently deep, dynamin is recruited to the neck of the CCP. Dynamin is a GTPase with a PIP_2 binding domain and a self-assembly domain (198, 199). Dynamin self-assembles and polymerises around the neck of the CCP, and upon GTP hydrolysis induces membrane scission at the neck of the CCP and a CCV pinches off from the plasma membrane (198, 200). The CCV then sheds its clathrin coat (201-203).

Once the clathrin coat has been lost, the vesicle fuses with the early endosome via a Rab5 containing protein complex (204). Rab proteins are small GTPases

thought to be key regulators of intracellular trafficking (205). Rab5 positive vesicles dock on to EEA1 and alsin, tethering proteins present on early endosomes (206, 207). Cargo in the early endosome is sorted and typically has three possible fates; be recycled back to the plasma membrane, be trafficked to the TGN or be transported to the lysosome for degradation (208).

The recycling of internalised molecules back to the plasma membrane is essential for plasma membrane homeostasis and can occur via a "fast" or "slow" method. Fast recycling is the direct transport of molecules from the early endosome back to the plasma membrane, regulated by Rab4 (209, 210). The slow recycling route involves the trafficking of cargo from the early endosome to the endocytic recycling compartment and then to the plasma membrane, a pathway that is regulated by Rab11 (211).

Retrograde trafficking is the process of cargo being transported from the early endosome to the TGN, which is mediated by a membrane sculpting/protein sorting complex known as retromer (212). The retromer complex comprises of a sorting nexin dimer, which binds to PIP_2 regions and drives membrane curvature, and a Vps26/Vps29/Vps35 trimer which sorts cargo for delivery to the TGN (213). Once at the TGN, the retromer derived vesicles dock and deliver their cargo (214). Retromer-mediated sorting controls the intracellular trafficking of a number of proteins, including Aβ, APP and BACE1 (215).

As cargo is transported out of the early endosome, the vesicle begins to mature into a MVB, which shares many similarities to a late endosome. Rab5, located in the early endosome, recruits effectors that activate Rab7, causing the maturation to Rab7 positive late endosomes (216). The maturation process begins with the inward budding of the membrane to create intraluminal vesicles (ILV). Non-ubiquitinated cargo is recycled back to the plasma

membrane and ILVs contain ubiquitinated cargo destined for degradation (217). The formation of the ILVs and the packaging of the ubiquitinated cargo require the action of the Endosomal Sorting Complex Required for Transport (ESCRT) (218, 219). The formation of the ILVs triggers the formation of MVB/late endosome (220). The late endosome will eventually fuse with a lysosome and its contents will be degraded. As well as transporting ILVs to the lysosome, MVB/late endosomes can also fuse with the plasma membrane to release ILVs in to the extracellular space (221). Released ILVs can contain many different types of molecules, including Aβ (222).

1.6.2 Trafficking of APP processing proteins

As the two pathways of APP processing are thought to occur at distinct cellular locations, the trafficking, cellular localisation and the co-localisation of APP and the secretases can directly impact Aβ generation. When CME is inhibited, APP processing is impacted. Typically it is thought increased delivery of APP to the cell surface or reduced internalisation will increase non-amyloidogenic processing, whereas increased internalisation or retention in the acidic organelles increases amyloidogenic processing (223).

Endocytosis is important for the cellular localisation of APP and BACE1. BACE1 is synthesised in the ER, transported to the golgi apparatus for post-translational modification, then mature BACE1 is trafficked to the cell surface or endosomes. Cell surface BACE1 can be internalised and recycled back to the cell surface, or trafficked into the endosomal lysosomal system. BACE1 internalisation is regulated by GTPase ADP ribosylation factor 6 (ARF6), which is associated with clathrin- and caveolin- independent endocytic routes (224, 225). As optimal BACE1 activity occurs within the acidic cellular compartments, such as the endosome, and APP cleavage by BACE 1 usually

occurs in the early endosome, aberrant trafficking of BACE1 could have a

subsequent effect on APP processing (224).

1.7 BIN1

1.7.1 Genetic association with LOAD

The first genetic evidence for Bridging Integrator 1 (*BIN1*) being associated with LOAD emerged from the 2009 genome wide association study (GWAS) performed by Harold *et al* (135). Although this study did not produce a genome-wide significant association between *BIN1* and LOAD, there was suggestive evidence with the most significant SNP reaching a p value of 3.2 x 10^{-6}. The association between *BIN1* and LOAD first reached genome-wide significance in 2010 in a study performed by Seshadri *et al* (137). This was a 3-stage meta-analysis identified rs744373 approximately 30 Kb upstream of *BIN1* with a meta-p value of 1.59 x 10^{-11}.

Hollingworth *et al* reported a significant association between rs744373 and LOAD in 2011 and reported an odds ratio (OR) of 1.17 (138). This provided independent support for the results reported by Seshadri *et al*. Naj *et al*, discovered another SNP in 2011, rs7561528, in the *BIN1* locus which reached genome-wide significance with a p-value of 5.2 x 10^{-14}, further establishing *BIN1* as an LOAD susceptibility locus (139). In 2013 the largest LOAD GWAS identified another SNP within the *BIN1* locus, rs6733839, located ~30 Kb upstream of *BIN1*, to have an overall meta-p value of 6.9 x 10^{-44} and an OR of 1.22 (140). A conditional analysis identified an additional independent association signal at rs7584040, indicating the presence of two functionally relevant loci at *BIN1* (Majounie et al, in prep). AD associated SNPs at the *BIN1* locus is shown in figure 1.3.

| ▮ rs7584040 | ▮ rs59335482 | ▮ rs6733839 | ▮ rs744373 |
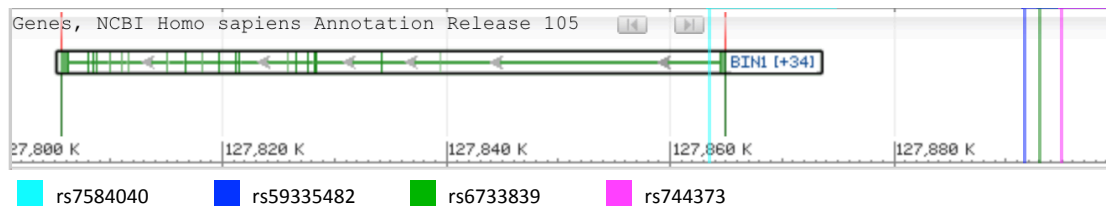
Figure 1.3. AD associated SNPs at the *BIN1* locus.

1.7.2 Gene Structure

BIN1 is located on chromosome 2q14.3 and is a member of the amphiphysin family (226). *BIN1* has a N-terminal BAR (BIN-amphiphysin/Rvs) domain, a PI (phosphoinositide) domain, a PS (Proline-Serine) domain, a CLAP (clathrin and AP-2) binding domain, a MYC-interacting domain and a C-terminal SH3 (Src homology 3) domain (226, 227). The N-BAR domain, present in all isoforms, is encoded by exons 1-10 and is capable of binding lipid membranes and inducing and sensing membrane curvature at sites such as endocytic pits of cytoplasmic endosomes (228, 229). The PI binding domain is encoded by exon 11 and is muscle specific (230-232). Exon 12 encodes the PS rich region present in all isoforms (233). Exons 13-16 encode the CLAP binding domain, which is only found in brain specific isoforms (230, 234, 235). This domain is capable of interacting with clathrin and AP2, vital components of CME (236, 237). The MYC-interacting domain is encoded by exon 17 and 18 (227) and the SH3 domain is encoded by exon 19 and 20, is present in all isoforms and binds proline rich motifs (238). *BIN1* is subject to extensive splicing, which generates 10 isoforms with diverse functions and tissue specific expression. The major isoforms differ due to the inclusion of exon 7, exon 11, brain specific exons 13-16 and exon 17 (239, 240). The protein domains, gene organisation and isoforms are illustrated in figure 1.4.
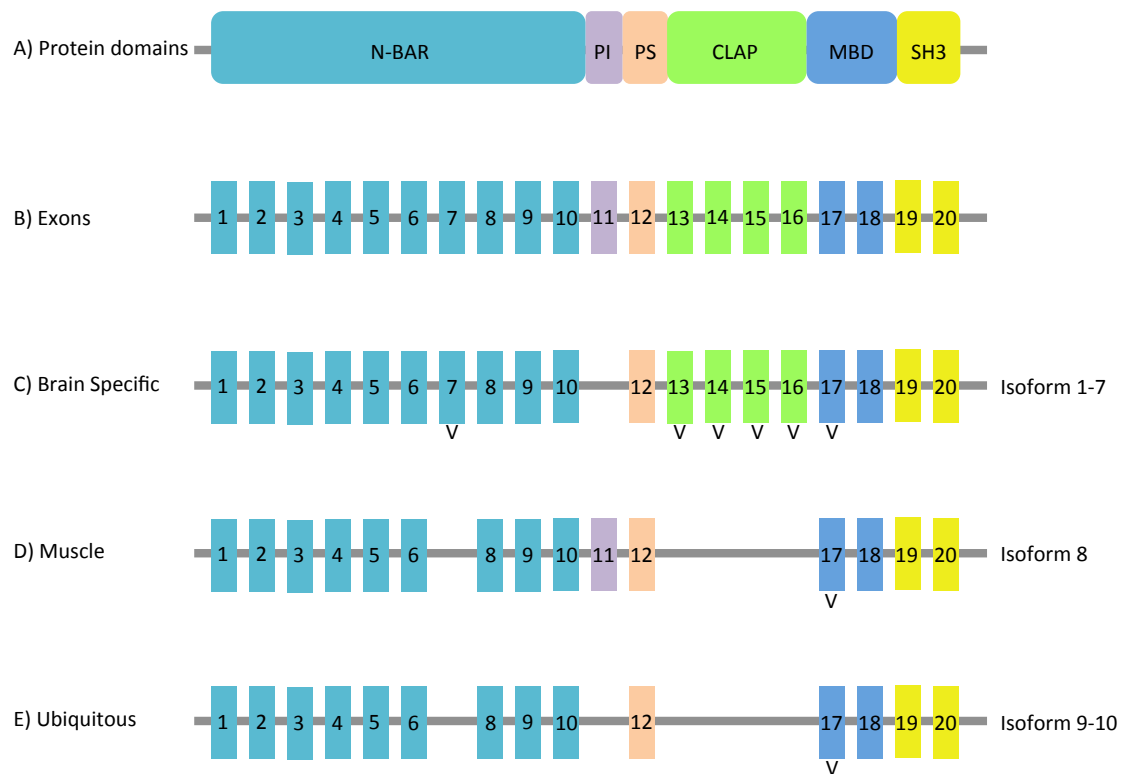
Figure 1.4. BIN1 protein domains, gene organisation and isoforms. A) Protein domains of BIN1, the coding exons are colour coded. B) Gene organisation of *BIN1*. C) Neuronal isoforms. Alternatively spliced exons are marked with a V. D) Muscle specific isoforms. E) Ubiquitous isoforms. Image adapted from (233).

1.7.3 BIN1 function in endocytosis and intracellular trafficking

The amphiphysin family of proteins is involved in endocytosis in both neuronal and non-neuronal cells and interacts with a number of proteins associated with CCPs via the SH3 domain (237, 240-245). It is thought that the amphipysins recruit effectors, such as clathrin, endophilin and dynamin, whilst simultaneously inducing membrane curvature at the neck of budding vesicles via the BAR domain (246, 247). This function requires the dimerisation of the BAR domain, which forms a positively charged concave surface which interacts with the negatively charged plasma membrane (229, 236, 248). BIN1 has been implicated in CME and intracellular trafficking (226, 249) and as the CLAP

domain is present only in brain isoforms, it is suggests that BIN1 may have a specific CME function in this tissue (236, 240, 244, 248). In order for endocytosis to be induced, BIN1 must be dephosphorylated by calcineurin (250). The BAR domain of BIN1 facilitates membrane curvature and the recruitment of dynamin by clustering PIP_2 in the plasma membrane. The accumulation of dynamin is ensured by the binding of the SH3 domain of BIN1 to the proline rich domain of dynamin (247, 251, 252). BIN1 has been characterised as a negative regulator of neuronal endocytosis, with downregulation resulting in an increase in early endosome rab5 activity (253).

Additional evidence implies BIN1 may also have essential functions in intracellular trafficking and endocytic recycling. Evidence suggests BIN1 may bind to the surface of intracellular vesicles and interact with SNX4, a member of the SNX protein family implicated in vesicular trafficking regulation. This interaction implies BIN1 may also have a role at the surface of newly formed endosomes and throughout intracellular regulation (254). RME1 is a protein essential for the successful transport of the endosome back to the membrane in order for cargo to be recycled. BIN1 orthologues have been shown to co-localise with RME1 on recycling endosomes and BIN1 knockdown in HeLa cells showed defective endosome recycling and abnormal endosome morphology, suggesting a further role for BIN1 in regulating endocytic recycling (249).

1.7.4 BIN1 and AD pathology

The endocytic function of BIN1 has led to investigations into BIN1 in relation to Aβ pathology, but the effect of BIN1 remains unclear. In human AD brains, BIN1 expression did not correlate with Aβ40/42 levels or neuritic plaques and BIN1 depletion had no effect on Aβ generation in neuronal-like cells (255, 256).

Knockdown and overexpression of the BIN1 neuronal isoform was found to have no effect of APP processing in neuroblastoma cells (256).

However, the number of BIN1 expressing neurons correlated with neuritic plaques in AD brain and BIN1 depletion resulted in an increase in Aβ secretion in HeLa cells expressing a familial AD mutation (257, 258). BIN1 was found to be crucial in the trafficking of BACE1 from the early endosomes and BIN1 depletion in neurons increased the co-localisation of APP and BACE1, resulting in increased intracellular Aβ accumulation within the axons (259). Additionally, 3 CpG sites within the *BIN1* locus were associated with Aβ load and the expression of three out of 14 isoforms investigated correlated with Aβ load and one isoform was negatively correlated with Aβ load (260).

1.7.5 BIN1 expression in AD patients

BIN1 is ubiquitously expressed, but the highest levels of expression are in the brain and muscle (231). Changes in *BIN1* expression in AD have been observed, however these observations are not consistent. *BIN1* mRNA transcript levels were shown to be increased in the frontal cortex of AD patients in comparison to healthy controls (261). When comparing BIN1 protein levels in AD brains to healthy brains, a modest increase in BIN1 expression was observed across all brains regions in AD brain, however further investigation identified that protein levels of neuronal isoform 1 decreased whereas ubiquitous isoform 9 increased, indicating a shift towards smaller isoforms in AD, which may be a result of neuronal loss (255). Additionally, increased levels of total BIN1 and the BIN1 neuronal isoform expression have been reported to be associated with a later age of onset and shorter disease duration (262) and an increase in *BIN1* mRNA levels was observed in the blood plasma of AD patients (263). Conversely, an 87% reduction of total BIN1 protein has been

observed in LOAD frontal cortex when compared to healthy individuals but remains unchanged in familial AD cases (256). Despite contradictory results, changes in *BIN1* expression appear to be associated with AD and are a pathogenic mechanism.

1.8 CD2AP

1.8.1.Genetic association with LOAD

CD2AP (CD2-associated protein) reached a genome wide significant association with LOAD in 2011 in two independent studies. These two studies combined their data to perform a combined analysis and identified rs9349407 to have a p value of 8.6×10^{-9} and an OR of 1.11 (138, 139). In the 2013 meta-analysis, rs10948363 was identified with a meta-p value of 5.2×10^{-11} and an OR of 1.1 (140). A conditional analysis identified an independent association signal at the *CD2AP* locus. Rs7745848, located upstream of *CD2AP* was identified with a p value of 1.33×10^{-5} (Majounie et al, in prep). AD associated SNPs at the *CD2AP* locus is shown in figure 1.5
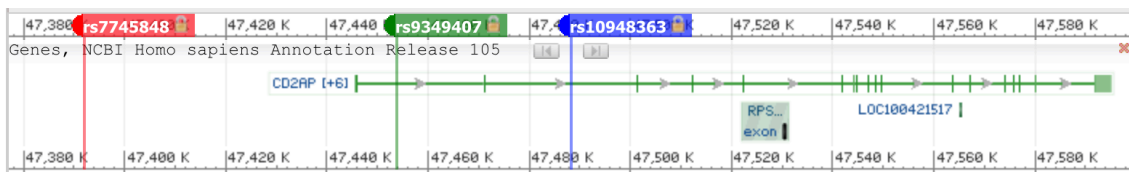
Figure 1.5. AD associated SNPs at the *CD2AP* locus.

1.8.2 Gene Structure

CD2AP is located at chromosome 6p12 and is comprised of 18 exons. Differential splicing of *CD2AP* results in eight different isoforms. CD2AP is a scaffolding molecule that regulates the actin cytoskeleton and is capable of binding actin via multiple actin binding sites, SH3 domains and Proline-rich regions (264). CD2AP has three N-terminal SH3 domains and a proline rich region, which are crucial in stabilising the contact between a T cell and antigen-presenting cell. The SH3 domains also interact with CD2 (a cell

adhesion molecule), ALIX (a cytoplasmic protein thought to be a regulator of the endo-lysosomal system) and Cbl (an adapter protein which acts in the protein tyrosine kinase signaling pathway) (265, 266). The SH3 domains are followed by a globular domain and a coiled-coil structure (264). The N terminus has an actin binding site that anchors CD2 at the site of cell contact (267, 268). The domain structure of CD2AP is illustrated in figure 1.6.

CD2AP protein domains
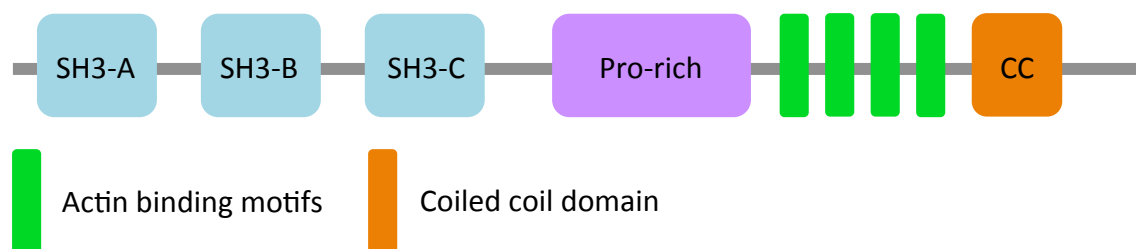
Figure 1.6. Protein domain structure of CD2AP. CD2AP has three N-terminal SH3 domains, a centrally located proline rich domain. There are multiple C-terminal actin binding domains and a coiled coil domains. Image adapted from (269).

1.8.3 CD2AP function

Numerous studies have identified variations in systemic immune responses in patients with AD, including the distribution and activation of lymphocytes (270, 271). CD2AP is a scaffolding protein that links actin to the immunological synapse (267) . CD2AP also binds to and facilitates the clustering of CD2, an adhesion molecule which binds ligands presented on antigen presenting cells and helps to create a tight, homogenous interaction between the membranes of T cells and antigen presenting cells (272). The binding of CD2AP is thought to aid in receptor patterning and cell polarisation in T cells (267).

The actin cytoskeleton plays a role in cell-cell adhesion (273) and CD2AP has been shown to regulate actin cytoskeletal dynamics via direct interactions with actin (274). Due to CD2AP interacting with actin and its role in cytoskeletal regulation, it has also been implicated in having a role in endocytosis (182, 268). The function of CD2AP in endocytosis is less understood than that of BIN1. CD2AP interacts with a number of components of endocytic machinery, such as rab4, AP2, cortactin and actin (247). CD2AP has been shown to form complexes with Cbl, an ubiquitin protein ligase, and endophilin, a BAR domain containing protein capable of inducing membrane curvature. It is proposed that Cbl mediated ubiquitination targets receptors for degradation. Interactions between Cbl and CD2AP recruit endophilin to induce membrane invagination at these sites. Following membrane invagination, cortactin is recruited to CD2AP. CD2AP links the endocytic complex to the cytoskeleton via its interaction with cortactin, which likely assists in the budding of the vesicle from the plasma membrane (268).

The co-expression of active rab4 or Cbl with CD2AP induces a significant enlargement of the early endosomes, indicating that a functional interaction between these proteins is essential for early endosome morphology (275). In addition to the early stages of endocytosis, CD2AP has also been implicated in exosome biogenesis and the degradative pathway (276, 277).

1.8.4 CD2AP and AD pathology

The *CD2AP* AD susceptibly locus is associated with increased plaque burden (278). In yeast, a functional homologue of CD2AP was identified as a suppressor of Aβ toxicity and CD2AP depletion in neuroblastoma cells transfected with wild-type APP resulted in a decrease in Aβ levels and a decrease in extracellular Aβ42/Aβ40 ratio (279, 280). In cortical neurons,

CD2AP is preferentially located in the dendrites and associates with early endosomes and APP. CD2AP depletion in cortical neurons led to an increase in Aβ42 in the dendrites but not the axons and this was due to inefficient degradation of membrane bound APP. Depletion of CD2AP prevented efficient sorting of APP for degradation and delivery to the lysosome. This resulted in APP accumulation in enlarged early endosomes, where amyloidogenic processing occurs. This inefficient sorting was proposed to be due to CD2AP having a critical role in MVB formation (259).

Breakdown of the blood brain barrier has been observed in AD patients and healthy aged individuals (281, 282). CD2AP deficient mice showed reduced blood brain barrier integrity, which could have detrimental effects on Aβ clearance and immune responses. This evidence indicates that CD2AP plays a key role in maintaining the blood brain barrier (283). It has also been observed that tau can have detrimental effects on blood brain barrier integrity and depletion of the fly orthologue of CD2AP enhanced tau toxicity in drosophila (284, 285). Therefore, loss of CD2AP function may also contribute to both amyloid and tangle pathology. CD2AP is ubiquitously expressed and expression does not associate with AD diagnosis or pathology (262, 264).

1.9 Gene Regulation

With only about 1.5% of the genome encoding for protein, a significant amount of the non-coding genome contains sequence elements that are capable of influencing the regulation of gene expression in diverse cellular environments (286, 287). Variation within these non-coding regulatory regions contribute to the phenotypic variation seen between individuals and can influence traits such as disease susceptibility. Protein coding genes typically have three types of regulatory DNA elements: a core promoter that is required for the recruitment of transcription apparatus and initiates transcription at the appropriate start site, a proximal promoter and distal DNA elements that can enhance or silence transcription (288, 289).

1.9.1 Eukaryotic transcription

The core eukaryotic promoter is located upstream of the protein coding region and is the site at which the transcriptional machinery assembles (288). This regulatory region binds Transcription factor II D, which subsequently coordinates the assembly of the transcription initiation complex and recruits RNA polymerase II (290, 291). The initiation of mRNA synthesis requires a number of multi-protein complexes called general transcription factors (292). Communication between RNA polymerase II, the initiation complex and the transcription factors (TF) is facilitated by the mediator complex (293-295).

The core promoter spans approximately 80 bp around the transcription start site (TSS). There are two classes of promoter in mammals; TATA box enriched promoters that initiate mRNA synthesis at a single TSS, and CpG-rich promoters that contain multiple TSS (296). Approximately 25% of mammalian

promoters contain a TATA box sequence motif that directs RNA polymerase II to the correct TSS. This can be the major TSS or several TSS located in a short nucleotide sequence (290). In CpG rich promoters, the TSS generally occur within or close to CpG islands, a genomic region enriched with CpG dinucleotides, which can result in diverse initiation patterns (297, 298). CpG poor promoters typically control genes whose expression is highly tissue specific, whereas CpG rich promoters are associated with multiple gene types including housekeeping, tissue specific and developmental genes (298). CpG islands are also associated with genes whose transcription is influenced by TFs that respond to upstream signals (299-301).

Promoters can be regulated by a number of DNA elements. The proximal promoter is generally located within 500 bp of the TSS and contains *cis*-elements capable of interacting with TFs. Enhancers, also capable of binding TFs, can work independently or synergistically with the core promoter to initiate transcription and can be located far from the gene which it influences. There are also B-recognition elements and initiator elements, which can interact with TFIID to elicit an effect on transcription initiation, and insulators, activators and repressors (291, 302).

1.9.2 Mechanisms of gene regulation

1.9.2.1 Trans-*regulation*

Trans-regulation is when the effects on gene expression are from a remote source, such as hormones or TFs that bind to DNA in a sequence specific manner to regulatory regions of the genome (303, 304). These are usually transcribed at distinct loci and are subject to their own regulation (305). *Trans*-regulatory factors equally influence transcription on both chromosomes.

1.9.2.2 Cis-*regulatory elements*

Cis-regulation is when the regulatory element is located on the same DNA molecule as the gene that it regulates and can therefore influence expression in an allele-specific manner. *Cis*-regulation variation is common in the human genome. *Cis*-acting elements, such as promoters, enhancers, silencers or locus control regions, can be influenced by sequence variation or epigenetic modification variation. Gene expression can also be influenced by DNA variation that can affect splicing or the stability of the gene product (306).

1.9.2.3 Enhancers

Enhancers are DNA elements that interact with *trans*-acting factors to enhance transcription. Enhancers contain sequence motifs, which bind TFs and can be located far from the TSS (307). Chromatin looping allows the enhancer to interact with the corresponding promoter and brings the enhancer-bound TF to the promoter (308). The bound TFs can then recruit co-factors, which can remodel chromatin structure, or co-activators, which determine the activity of the promoter and regulate RNA polymerase II activity (309). Enhancers have also been implicated in RNA polymerase II release (310).

It is presumed that enhancers act upon the closest promoter, however is has been shown that they can have an effect on the transcription at distant promoters and even on promoters located on a different chromosome (311-313). Most promoters associate with a single enhancer but around 25% associate with two or more enhancers and multiple enhancers and regulatory elements can have a combinatorial effect on gene regulation (314). Nucleosomes in the proximity of enhancers typically have modifications

associated with active chromatin and active enhancers typically do not contain nucleosomes so that the DNA is accessible to TFs (315).

Tissue specific expression is required for the unique development and function of each tissue. Enhancers are key regulatory elements controlling this process. Super-enhancers, which comprise of multiple enhancers within a certain region, are thought to be influential elements that control tissue specific transcription (316, 317). The establishment of super enhancers can be influenced by environmental factors, such as hormones during development, and each enhancer within a super enhancer binds tissue or lineage specific TFs (318-320). Genes associated with super enhancers are expressed at higher levels than those associated with single enhancer regions and are expressed in a more tissue specific manner, resulting in the necessary gene expression for the desired tissue type (316, 321-324).

1.9.2.4 Histone modification / chromatin structure

The organisation of genomic material greatly influences gene expression regulation (325). A nucleosome core is formed of 147 bp of DNA wrapped around a protein core consisting of histones H2A, H2B, H3 and H4 (326). Nucleosome cores are connected by linker DNA, ranging from 10-90 bp, forming a "beads on a string" nucleosomal array (327). Linker histones (H1 and H5) bind to the linker DNA at the site of DNA exit and entry to the nucleosome core, which organises the nucleosome array into a chromatin fiber (328, 329).

The interaction between DNA and histones greatly affects chromatin accessibility, which can affect protein binding to *cis*-elements necessary for transcription. DNA binding proteins, histone-modifying enzymes and proteins that bind to modified histones are capable of influencing chromatin structure

(325). Euchromatin is active chromatin where DNA is accessible for transcription, meaning that these regions are vulnerable to DNase I cleavage and DNase I hypersensitivity sites are hallmarks of active chromatin (315, 330). Heterochromatin is tightly packed chromatin and transcription is repressed in this state (331). Chromatin state depends upon the type of modifications on the core histones.

Histones have specific residues that can be modified via methylation, acetylation, phosphorylation or ubiquitination to influence transcription (331). Histone methylation and acetylation is performed by histone methyltransferases and histone acetyltransferases respectively (332, 333). Acetylation disrupts the interaction between the DNA and the histone causing the chromatin to become more accessible, and acetylated and methylated residues are capable of recruiting regulatory proteins to the chromatin (331).

Regulatory elements tend to have signature histone modifications. Active promoters are commonly marked by the di- and tri-methylation of lysine 4 on histone H3 (H3K4me2 and H3K4me3) and histone acetylation. Whereas, repressed promoters are typically marked by DNA methylation, H3K27me3 and H3K9me3, typical modifications in heterochromatin (325).

Enhancers are characterised by mono-, di- and tri-methylation of lysine 4 on histone H3 (H3K4me1, H3K4me2 and H3K4me3) (334-339). Active enhancers are enriched for acetylated H3K27 (H3K27ac), whereas inactive enhancers are enriched with H3K27me3 (340). The additional histone modification, H3K9me1, is associated with active enhancers, whereas H3K9me2 and H3K9me3 are associated with the repression of transcription from promoters and enhancers (325, 330, 337, 341). In genomic DNA (gDNA) CpG density and H3K4 tri-methylation are correlated, suggesting a mechanistic connection between the

two (297, 342). Increased levels of H3K4me3 modifications are associated with the TSS of many transcribed genes and CpGs at these locations are hypomethylated (325, 342).

1.9.2.5 DNA methylation

Methylation of cytosine residues within a CpG dinucleotide to form 5-methylcytosine (5mC) is a key epigenetic modification that influences gene expression (343). 5mC is produced when DNA methyltransferases add a methyl group to the 5-postion of the cytosine ring. 5mC is associated with gene silencing and plays a vital role in genomic imprinting, X-inactivation and lineage-specific gene expression regulation (344, 345).

In the mammalian genome, around 70-80% of CpG sites are methylated and highly methylated DNA sequences are located in satellite DNAs, repetitive elements, gene bodies and non-repetitive intergenic DNA (346). However, approximately 60% of CpG Islands, a region enriched with CpG dinucleotides, located in promoters are unmethylated (347, 348). Methylation at the CpG rich region of promoters are associated with gene repression, whereas low levels of methylation close to the TSS and methylation occurring in the gene body is associated with transcribed genes (349-351).

Methyl-CpG binding proteins (MeCPs) are able to interact with transcriptional repressors and specifically recognise methylated CpGs. Once bound, MeCPs recruit histone-modifying complexes to establish a repressive chromatin structure and stabilise patterns of gene expression (352-357). DNA methylation can also affect nucleosome positioning and can prevent the binding of TF to inhibit transcriptional activation (358-360).

1.9.3 Gene regulation in complex disease

Genetic variation can influence gene regulation in a number of ways. The first to be characterised were structural variants, such as large deletions or chromosomal translocations. These structural variants can separate regulatory elements from their target genes, resulting in lost or gain of target gene regulation, and are most often found in Mendelian disorders or cancers (361). In other instances, a gain of regulatory element function can be caused by copy number variants of the regulatory elements (362, 363).

In addition to large structural genomic changes, it has been shown that smaller genetic variants, such as SNPs, located within regulatory elements can modify disease risk. SNPs can impact on gene regulatory elements by disrupting the assembly of the transcriptional machinery. Changes in the binding of TFs are associated with SNPs in regulatory regions and can result in altered transcription levels (364). TF binding sites containing variants that alter protein binding are also enriched with disease-associated variants, suggesting TF binding may be a potential functional mechanism in disease risk (365).

Variation that affects enhancer function can have consequences on gene expression. SNPs disrupting the conserved enhancer for the *OCA2* gene, which is associated with eye colour, can result in a blue eye colour (366, 367) and mutations in an enhancer for *GATA1* can significantly affect TF binding and can result in erythroblastic leukemia and myelofibrosis in mice (368). Another well-characterised example involved SNPs associated with obesity located within introns of the *FTO* gene (369). It was found that this genomic region physically interacted with the promoter of another gene, *IRX3*, located 500 Kb from the associated variants. These variants showed association with *IRX3* expression and this investigation identified *IRX3* as having a causal role in obesity (370).

These examples demonstrate the effect variation in regulatory elements can have on biology and disease and how these effects can occur over larger genomic distances.

The majority of genetic variation associated with complex diseases is located outside of the protein-coding region and is thought to elicit its effect via the disruption of regulatory elements. Many risk loci across multiple phenotypes have been implicated in the regulation of transcriptional activity and are enriched with *cis*-regulatory elements (371). There are numerous examples of how specific sequence changes in DNA can change the interactions of proteins that complex with DNA to alter transcription. The characterisation of risk variants in relation to their impact on gene regulatory mechanisms on disease relevant tissues may provide insight into the functional mechanisms of disease, explain association signals observed and ultimately contribute to the knowledge of biological pathways that could be targeted by therapeutic interventions.

1.10 Aims

Despite much research linking genes involved in endocytosis to AD, how they increase susceptibility for the disease remains unknown. This thesis aims to investigate the functional impact of BIN1 and CD2AP depletion on AD pathological processes and investigate the *BIN1* risk locus to determine how a non-coding region of the genome could affect the risk of developing AD.

BIN1 and CD2AP both play a role in endocytosis and have been associated with AD. In chapter 3, this thesis investigates how depletion of these genes, independently and in combination, affects amyloid processing and endocytosis with the aim of determining a disease relevant function in a human neuroglioma cell line.

GWAS have identified numerous genomic loci associated with LOAD, however how they influence risk mechanisms remains unknown. The *BIN1* risk locus is the second most significant locus associated with LOAD after *APOE*. The most significant variant is located approximately 30 Kb from the coding region of *BIN1* and how it confers AD risk remains unknown. As this locus is in an intergenic region, it is hypothesised that it is conferring risk for AD by disrupting *BIN1* regulation. Chapters 4 and 5 use a number of approaches to characterise the regulatory capacity of this locus and investigate the effect of AD associated genotypes.

2. Materials and Methods

2.1 Cell culture

2.1.1 Mammalian Cell Culture

2.1.1.1 H4 Neurogliomas

The human neuroglioma derived cell line known as H4 was established in 1971. H4s originated from a 37 year old Caucasian male's brain tumor tissue, diagnosed as neuroglioma with a glial origin (372). This cell line possesses an epithelial morphology.

The H4 neuroglioma cell line was cultured in Opti-MEM (Thermo Fisher Scientific, Newport, UK) supplemented with 4% fetal bovine serum (FBS) (Thermo Fisher Scientific) and incubated in 5% CO_2 at 37°C. The culture media was changed every 2-3 days. In order to passage this cell line, once approximately 80% confluent (or as desired) the media was removed and the cells were washed with Dulbecco's phosphate buffered saline (PBS) (Thermo Fisher Scientific). Cells were incubated in 0.25% Trypsin-EDTA (Thermo Fisher Scientific) for 5 minutes at 37°C. Once cells detached from the culture dish surface, the trypsin-EDTA was neutralised with the addition of normal FBS containing culture medium. Cells were resuspended in the desired amount and seeded into a new culture vessel. H4 cells were cryopreserved in FBS supplemented with 5% (volume/volume (v/v)) dimethyl sulfoxide (DMSO) and stored in liquid nitrogen.

2.1.1.2 Microglia SV40

The Immortalised Human Microglia - SV40 cell line was derived from Primary Human Microglia and is suitable for studies of human microglia in health and disease. This cell line was established from embryonic primary microglia cells in 1995 (373).

Flasks used to culture microglia SV40 cells were coated in collagen solution (Merck, Watford, UK) diluted in PBS to 5-10 µg/cm^2. The solution was incubated in a flask for 30 minutes, removed and the flask was washed with PBS. The microglia SV40 cell line was cultured in Prigrow III (Applied Biological Materials Inc., BC, Canada) supplemented with 10% FBS and incubated at 5% CO_2 at 37°C. The media was changed every 2-3 days. In order to passage this cell line, once approximately 80% confluent (or as desired) the media was removed and the cells were washed with PBS. Cells were then incubated in 0.25% Trypsin-EDTA for 5 minutes at 37°C. Once cells detached from the culture dish surface, the trypsin-EDTA was neutralised with the addition of normal FBS containing culture medium. Cells were resuspended in the desired amount of normal culture medium and seeded into a new collagen coated culture vessel. Microglia SV40 cells were cryopreserved in complete growth media supplemented with 10% (v/v) DMSO and stored in liquid nitrogen.

2.1.1.3 HEK293

HEK293 cells are a human embryonic kidney derived cell line with an epithelial morphology. This cell line was established in 1977 (374). HEK293 cells were cultured in DMEM (Dulbecco's Modified Eagle Medium) (Thermo Fisher Scientific) supplemented with 4% FBS. The media was changed every 2-3 days. In order to passage this cell line, once approximately 80% confluent (or as

desired) the media was removed and the cells were washed with PBS. Cells were then incubated in 0.25% Trypsin-EDTA for 5 minutes at 37°C. Once cells detached from the culture dish surface, the trypsin-EDTA was neutralised with the addition of normal FBS containing culture medium. Cells were resuspended in the desired amount and seeded into a new culture vessel. HEK293 cells were cryopreserved in complete growth media supplemented with 5% (v/v) DMSO and stored in liquid nitrogen.

2.1.1.4 THP-1

The THP-1 cell line is a human leukemic cell line cultured from a male with acute monocytic leukemia. The THP-1 cell line is monocytic in nature and was established in 1980 (375).

THP-1 cells are non-adherent and were grown in suspension and cultured in RPMI 1640 medium (Thermo Fisher Scientific) supplemented with 10% FBS and 2mM L-Glutamine (Thermo Fisher Scientific) and incubated at 5% CO_2 at 37°C. Culture medium was changed every 2-3 days. Cells were subcultured once the cell concentration exceeded 8×10^{-5} cells/mL. THP-1 cells were cryopreserved in complete growth media supplemented with 5% (v/v) DMSO and stored in liquid nitrogen.

2.1.1.4 BV2

BV2 cells are derived from immortalised murine neonatal microglial and are frequently used as a substitute for primary microglia (376). The BV2 cell line was established in 1992 (377).

BV2 cells grow both in suspension and adherently. BV2 cells were cultured in RPMI 1640 medium supplemented with 10% FBS and 2mM L-Glutamine and incubated at 5% CO_2 at 37°C. Culture medium was changed every 2-3 days. Once cells became confluent, cells in suspension were removed from the culture flask. Adherent cells were washed with PBS and then incubated in 0.25% Trypsin-EDTA for 5 minutes at 37°C. Once cells detached from the culture dish surface, the trypsin-EDTA was neutralised with the addition of normal culture medium. Cells were resuspended in the desired volume then combined with the cells in suspension. Both the cells in suspension and the adherent cells were seeded into a new culture flask in order to maintain the mixed population of cells. BV2 cells were cryopreserved in complete growth media supplemented with 5% (v/v) DMSO and stored in liquid nitrogen.

2.1.2 Bacterial Cell Culture

Stbl3™ *E.coli* strain (Thermo Fisher Scientific) and Subcloning Efficiency™ DH5α™ Competent cells (Thermo Fisher Scientific) were used for transformations and further culture.

Luria-Bertani (LB) broth was used in liquid culture of bacterium purchased from (Thermo Fischer Scientific). LB broth was supplemented with the necessary antibiotic for cell selection. Liquid cultures were incubated at 37°C in a shaking incubator, agitating the culture at 225 rpm. Cultures were cryopreserved in complete growth media supplemented with 5% (v/v) DMSO and stored at -80°C.

Agar broth (Thermo Fischer Scientific) was used for solid culture plates. The agar was supplemented with the appropriate antibiotics for colony selection. Plates are incubated at 37°C and stored at 4°C.

2.2 Cell Techniques

2.2.1 Oligofectamine™ Transfection of siRNA in H4 cell line

Transfection of siRNAs was performed using Oligofectamine™ Transfection Reagent (Thermo Fisher Scientific). Transfections were scaled for different tissue culture plates. The conditions and reagents required for transfections in different sized tissue culture plates are described in table 2.1.

Table 2.1. Transfection conditions for siRNA transfections using Oligofectamine Transfection Reagent scaled to different tissue culture plates.

Transfection Conditions (per well)		6 well plate	12 well plate	24 well plate	96 well plate
Number of cells		1.5×10^5	5×10^4	3×10^4	4800
Tube A (µL)	siRNA (X µM)	1	0.4	0.21	0.034
	Opti-MEM	179	71.6	37.76	6.03
Tube B (µL)	Oligofectamine Reagent	4	1.6	0.84	0.13
	Opti-MEM	16	6.4	3.38	0.54
Opti-MEM added to cells (µL)		800	320	168.78	26.94
Complex added to cells (µL)		200	80	42.19	6.73
12% FBS in Opti-MEM added to cells (µL)		500	200	105.49	16.8

Cells were plated at the desired density 24 hours prior to transfection and incubated at 37°C and 5% CO_2. On the day of transfection, siRNA (at the desired concentration) and Opti-MEM were combined in tube A. The oligofectamine transfection reagent was combined with Opti-MEM in tube B and incubated at room temperature (RT) for 10 minutes. Following incubation, tube A and B were combined and incubated for 20 minutes at RT. During this incubation, the transfection reagent forms stable complexes with the siRNA oligonucleotides, which can be efficiently transfected into eukaryotic cells.

The growth media was removed from the cells and the cells were washed once with Opti-MEM without serum. Cells were left in serum free media by adding the required amount of Opti-MEM to the wells as described in table 2.1. Following the 20-minute incubation, the complexes were added to cells. The cells were incubated for 4 hours at 37°C in 5% CO_2, then serum free media supplemented with 12% FBS was added to the wells (volume described in table 2.1). Cells were then cultured for 48 hours and used for further experimental analysis.

2.2.2 Plasmid Transfection using Lipofectamine™ 3000 Reagent

Plasmid transfections were performed using Lipofectamine™ 3000 Reagent (Thermo Fisher Scientific). Transfections were scaled for different tissue culture plates. The conditions and reagents required for transfections in different sized tissue culture plates are described in table 2.2.

Table 2.2. Transfection conditions for plasmid DNA transfections using Lipofectamine 3000 Reagent scaled to different tissue culture plates. X represents variable amounts of plasmid DNA. Y represents variable amounts of Lipofectamine™ 3000 Reagent.

Reagent (per well)		6 well plate	12 well plate	24 well plate
Tube A	Plasmid DNA	X ng	X ng	X ng
	P3000™ Reagent (μL)	5	4	1
	Opti-MEM (μL)	125	50	25
Tube B	Lipofectamine™ 3000 reagent (μL)	Y	Y	Y
	Opti-MEM (μL)	125	50	25
Volume of DNA-lipid complexes added to cells (μL)		250	100	50

Cells were plated at the desired density 24 hours prior to transfection. On the day of transfection, the desired amount of DNA, P3000 Reagent and Opti-MEM was combined into tube A. The desired volume of Lipofectamine Reagent and Opti-MEM was combined into tube B. The contents of each tube was combined at a 1:1 ratio and incubated at RT for 15 minutes. During this time the Lipid-DNA complexes form.

During DNA-lipid complex incubation, growth media was replaced with fresh media. The desired volume of DNA-lipid complexes was added directly to the cells and incubated for 48 hours. Following incubation, cells were used for further experimental analysis.

2.2.3 Cell lysis Protocol to generate protein lysates

10X Cell Lysis Buffer (Cell Signaling Technologies, Leiden, Netherlands) was used to lyse cells under non-denaturing conditions. The buffer consists of 20 mM Tris-HCl (pH 7.5), 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM beta-glycerophosophate, 1 mM Na_3VO_4 and 1 µg/mL leupeptin. This buffer was diluted to 1X and 1 mM of Phenylmethylsulfonyl fluoride was added prior to use.

In order to perform the cell lysis, the media was removed from the cells and the cells were washed twice with PBS to remove residual media. In a six well plate, 100 µL of 1X cell lysis buffer was added directly to cells and incubated on ice for five minutes. Cell scrapers were used in order to detach cells from the surface of the plate and the lysate was transferred to a microcentrifuge tube. The cell lysate was sonicated for five seconds and then centrifuged for 10 minutes at 14000 g at 4°C. The supernatant containing protein was removed and stored at -20°C.

2.3 Protein techniques

2.3.1 BCA (Bicinchoninic Acid) assay

A BCA assay is a method developed to measure the protein concentration of sample. In this instance, this assay was used to measure the protein concentration of lysates generated from transfected cells and nuclear protein extractions.

BCA standards were created by diluting 2 mg/mL Albumin in water (Thermo Fisher Scientific). The range of concentrations used as standards are shown in table 2.3.

Table 2.3. Dilution of 2mg/mL Albumin to create protein standards for BCA assay

Concentration of Albumin (mg/mL)
2
1.5
1
0.5
0.25
0.125
0.0625
0.0312

5 µL of each standard was added to a well of a clear flat bottomed 96 well plate. 5 µL of water was added to act as a blank for the standards. 5 µL of protein lysates were added to the plate. 5 µL of lysis buffer was added to act as blank for the protein lysates. Each well was duplicated to ensure accurate results.

The Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific) was used to determine protein concentration. Reagent B was added to Reagent A at a ratio of 1:50. 200 µL of the Reagent A and B mixture was added to the wells containing standards, blank and protein lysates, and pipette mixed. The plate was incubated at 37°C for 20 minutes. Following incubation, optical density was measured using a µQuant™ Microplate Spectrophotometer at 562nm wavelength.

Duplicate measurements were averaged. Standards and protein lysate readings were blanked using the appropriate blank measurements. In order to analyse the results, a graph of the protein standards concentration against their optical reading was plotted. A line of best fit was created and the y=mx+c equation was used to calculate the concentration of the protein lysate measured.

2.3.2 Western blotting

Western blotting is a semi-quantitative technique able to visualise a specific protein present in a lysate sample. In this instance, this technique was used to investigate BIN1 and CD2AP depletion and APP levels in transfected cells.

10% acrylamide gels were used in western blot experiments. These were made up of a resolving gel and a stacking gel for sample loading. The glass plates were assembled into the casting stand. To make the resolving gel for two 10% acrylamide gels of 0.75 mm width, 1.9 mL 1.5M tris (pH 8.8), 2.5 mL 30% acrylamide, 37.5 µL 20% Sodium Dodecyl Sulfate (SDS), 75 µL 10% APS, 6 µL TEMED (Tetramethylethylenediamine) and 3 mL H_2O was mixed and poured into the gel cast, leaving approximately 1.5 cm for the stacking gel. In order for the gel to set horizontally, isopropanol was added on top of the resolving gel and the gel was left to set for 10 minutes.

To create the stacking gel, 375 µL 1.5M Tris (pH 6.8), 750µl 30% acrylamide, 22.5 µL 20% SDS, 45 µL 10% APS, 4.5 µL TEMED and 3.15 mL H_2O was mixed. The isopropanol on the top of the resolving gel was discarded and residual liquid was removed with blotting paper. The stacking gel was added to the gel cast and the appropriate well comb was inserted. Once set, gels were stored for 2-3 days and in 1X PAGE [reagents to create 10X PAGE are described at

the end of this section]. The gel plates were slotted into the Mini-PROTEAN

Tetra Cell (Bio-Rad Laboratories Ltd, Watford, UK) and placed in a tank. If only

one gel was used, the plastic buffer dam was slotted into the tetra cell in

addition to the acrylamide gel. The tank was filled with 1x PAGE. 5 µL of Prism

Ultra Protein Ladder (Abcam, Bristol, UK) was loaded into one well in order to

determine protein size.

Lysate samples were prepared by combining one part 5X sample buffer

(0.125M Tris (pH 6.8), 4% SDS, 20% v/v Glycerol, 5% 2-Mercaptoethanol and

Bromophenol blue) with four parts protein lysate. Samples were heated at

95°C for 5 minutes. Following incubation, the lysate sample plus sample buffer

was loaded into the appropriate wells. Gels were electrophoresed at 150V for

the desired length of time.

Once the desired separation of the protein ladder was achieved, the protein

was transferred on to a nitrocellulose blotting membrane (GE Life Sciences,

Buckinghamshire, UK). 1L of transfer buffer (100 mL 10X PAGE, 200 mL

Methanol and 700 mL H_2O) was prepared. A piece of nitrocellulose blotting

membrane was cut to the appropriate size and soaked in transfer buffer.

The bottom half of the transfer sandwich was created using the gel holder

cassette, black foam, and blotting paper. One slide of the glass plate was

removed and the nitrocellulose blotting membrane was placed onto the gel. A

roller was used to ensure that no bubbles were present between the gel and

the nitrocellulose. The gel and nitrocellulose were peeled off the glass plate

and placed in the transfer sandwich, covered with blotting paper and foam,

and the clasp of the cassette was sealed. The transfer sandwich was placed in

an electrophoretic transfer unit and placed in transfer tank filled with transfer

buffer and an ice pack. If two gels were being transferred, 85V was supplied to

the transfer unit. If one gel was being transferred, 75V was supplied to the transfer unit. Electrophoretic transfer was carried out for 1 hour.

Following the transfer, blots were blocked by incubating in 5% milk (in PBS-T) for 1 hour. [Reagents for PBS-T are described at the end of this section]. Primary antibody was diluted in either 5% milk or 2% Bovine Serum Albumin (BSA) to the appropriate concentration depending on the antibody used. If the antibody was diluted in BSA, three 5-minute washes with PBS-T were performed following the blocking incubation. The blocking milk was removed and the primary antibody solution was added. The solution was agitated and incubated at 4°C over night (ON).

Following incubation, the primary antibody solution was removed and the blot was washed three times for 5 minutes in PBS-T. The appropriate secondary antibody was diluted in 5% milk at 1:15000 and added to the blots. The secondary antibody solution was agitated and incubated on the blot for 1 hour at RT then washed in PBS-T. Protein bands were detected with Pierce ™ ECL Western Blotting Substrate (Thermo Fisher Scientific). The substrate working solution was prepared by mixing equal parts of Detection Reagents 1 and 2. The substrate working solution was incubated on the blots for 5 minutes at RT. Excess substrate working solution was removed from the blot and the blot was placed in a film cassette. In a dark room, Amersham Hyperfilm ECL (GE Life Sciences) X ray films were exposed to the blot for the appropriate length of time. The X ray films were developed using the ECOMAX film processor (Protec, Oberstenfeld, Germany).

In order to remove the bound primary antibody from a blot, the blot was incubated in Restore™ PLUS Western Blot Stripping buffer (Thermo Fisher Scientific) at RT for 15 minutes with agitation. Following this, the blot was

washed three times for 5 minutes in PBS-T and then can be treated from the blocking step as described above.

Western blots were quantified via densitometry [described in section 2.5]. The Protein ladder was used to identify target proteins by molecular weight. An endogenous control was used in each blot using protein from a non-transfected cell population. Ideally a negative control would be performed using a cell line that does not express the protein of interest to demonstrate the specificity of the antibody. As BIN1, CD2AP, APP and GAPDH were expressed in cell lines available, a negative control was not possible. Targeting siRNA resulting in BIN1 and CD2AP depletion can be used to demonstrate the specificity of the BIN1 and CD2AP targeting antibodies.

Expression of the protein of interest was normalised to the expression of GAPDH, a housekeeping protein, to normalise for loading errors. Each normalised value was divided by the expression from a non-transfected control and expressed as a percentage of expression change compared to the control.

Reagents:

1.5M Tris pH 8.8 – 153.9 g Tris Base, 36.9g Tris HCl, 1L H_2O

1.5M tris pH 6.8 – 225g Tris HCl, 9g Tris Base, 1L H_2O

10X PAGE – 30.25g Tris Base, 144g Glycine, 10g SDS

PBS-T – 5 x PBS tablets (Sigma), 1 mL Tween, 1L H_2O

2.3.3 Enzyme-Linked Immunosorbent Assay (ELISA)

2.3.3.1 Human APP Quantification

Human APP protein was quantified using the Human APP Duoset ELISA following the manufacturers guidelines (R&D Systems, Abingdon, UK).

Plate Preparation

The mouse anti-human APP Capture Antibody was diluted to a working concentration of 4 µg/mL in PBS (Thermo Fisher Scientific). 100 µL of capture antibody was added to a 96-well Microplate and incubated ON at RT. Following incubation, the antibody was removed and each well was washed three times with Wash Buffer (0.05% Tween®20 in PBS). The wells were blocked by adding 300 µL of Reagent Diluent (1% BSA in PBS) and were incubated at RT for 1 hour. Following incubation, the Reagent diluent was removed and each well was washed three times with Wash Buffer.

Assay Procedure

Protein standards were created by diluting recombinant human APP with Reagent diluent. Standards between 0-20 ng/mL of human APP were generated. Samples from H4 cell lysates were diluted 1:15 in Reagent diluent. 100 µL of diluted samples, standards and appropriate blanks were added to the wells and incubated at RT for 2 hours. Following incubation, the samples, standards and blanks were removed and the wells were washed three times with wash buffer. 100 µL of biotinylated mouse anti-human APP Detection Antibody diluted in Reagent buffer (330 ng/mL) was added to the wells and incubated at RT for 2 hours. Following incubation, the detection antibody was removed and the wells were washed three times with wash buffer. 100 µL of substrate solution (1:1 mixture of Colour Reagent A (H_2O_2) and Colour Reagent B (Tetramethylbenzidine) was added to each well and incubated for 20 minutes at RT in the dark. Following incubation, 50 µL of Stop Solution (2 N H_2SO_4) was added to each well. The optical density of each well was determined on a µQuant™ Microplate Spectrophotometer using 450nm wavelength.

2.3.3.2 Quantification of APP related proteins

β-CTF, Aβ40, sAPPα, sAPPβ and BACE1 were quantified using ELISA kits from IBL International (Hamburg, Germany) and performed following the manufacturer's protocol. Wash buffer (40X) was diluted to 1X in deionised water. The Enzyme Conjugate (30X) was diluted 1:30 with Enzyme Conjugate diluent. The reconstituted protein standard provided was diluted in assay buffer to generate the appropriate protein standards. β-CTF protein standards ranged from 0.19-12 pmol/L. Aβ40 protein standards ranged from 1.56-100 pg/mL. sAPPα and sAPPβ protein standards ranged from 0.78-50 ng/mL. BACE1 protein standards ranged from 1.56-100 ng/mL. Test samples were diluted as desired with assay buffer.

100 μL of standards and diluted samples was added to the appropriate wells of the antibody coated microtiter plate. 100 μL of assay buffer was added to wells to act as a blank. For β-CTF, Aβ40, sAPPα and sAPPβ ELISAs the plate was incubated ON at 4°C. For the BACE1 ELISA, the plate was incubated for 1 hour at 37°C.

Following incubation, samples, standards and blanks were removed from the wells and the wells are washed thoroughly. Each well was filled with 1X wash buffer, incubated for 15 seconds then wash buffer was removed. This was repeated 8 times. Once all of the wash buffer was removed, 100 μL of 1X Enzyme Conjugate was added to each well and incubated for 1 hour at 4°C. Following incubation, the wells are washed 9 times using the same method as previously described. Once the wash buffer was removed, 100 μL of TMB Substrate Solution was added to the wells and the plate was incubated at RT for 30 minutes in the dark. Following incubation, 100 μL of TMB Stop Solution was added to the wells and mixed. The optical density of each well was

determined on a µQuant™ Microplate Spectrophotometer using 450nm wavelength.

2.3.3.3 Experimental Design

An endogenous control was used in each ELSIA using protein extracted from a non-transfected cell population. Ideally a negative control would be performed, using a cell line that does not express the protein of interest to demonstrate the specificity of the antibody. However, as APP, β-CTF, Aβ40, sAPPα, sAPPβ and BACE1 are expressed in cell lines available, a negative control was possible.

2.3.3.4 Calculating ELISA Results

The average optical density of the blanks was subtracted from the optical density of each sample and standard technical replicate. These blanked technical replicates were then averaged. Plotting the standard protein concentration against optical density generated a standard curve. A line of best fit was plotted to the standard curve.

The formula for the line of best fit (y=mx + c) was applied to calculate the protein concentrations of the samples from the optical density recorded. This value was then multiplied by the dilution factor and then divided by the total protein concentration of the sample lysate. This generates a target protein mass per mg of lysate protein (for example, APP ng/mg).

2.4 DNA techniques

2.4.1 DNA extraction

DNA was extracted from cell lines using QIAamp® DNA Mini Kit (Qiagen, Manchester, UK). Once cells reached confluency, cells were trypsinised and a maximum of 5×10^6 cells was added to a microcentrifuge tube. Cells were centrifuged for 5 minutes at 300 x g. The supernatant was removed and the cell pellet was resuspended in 200 µL of PBS. 20 µL of proteinase K was added to the resuspended cells. 200 µL of Qiagen Buffer AL was added to the cell suspension, mixed and incubated at 56°C for 10 minutes. 200 µL of 100% ethanol was added to the sample, mixed and then added to the QIAamp Mini spin column. The column was centrifuged for 1 minute at 6000 x g. The filtrate was discarded and 500 µL of Qiagen Buffer AW1 was added to the column. The column was centrifuged for 1 minute at 6000 x g. The filtrate was discarded and 500 µL of Qiagen Buffer AW2 was added to the column. The column was centrifuged for 3 minutes at 20000 x g. The filtrate was discarded and the column was centrifuged for a further minute. The column was placed into a clean microcentrifuge tube, 50 µL of distilled water was added to the column and incubated for 1 minute at RT. The DNA was eluted into the water and collected by centrifuging the column at 6000 x g for 1 minute.

2.4.2 Isolation of plasmids

Plasmids were isolated from transformed *E.coli* using QIAprep® Spin Miniprep Kit (Qiagen). Transformed cultures were incubated ON and collected in a microcentrifuge tube. The bacterial cells were harvested by centrifuging for 3 minutes at RT at 6800 x g. The supernatant was discarded and the cell pellet was resuspended in 250 µL of Qiagen buffer P1. Cells were lysed by the

addition of 250 µL of buffer P2. The tubes were inverted to mix and the lysis reaction was carried out for 5 minutes. To neutralise the reaction, 350 µL of Qiagen buffer N3 was added and mixed. The tube was then centrifuged for 10 minutes at 17900 x g. The supernatant was added to a QIAprep spin column and then centrifuged for 1 minute at 17900 x g. The filtrate was discarded, 500 µL of Qiagen buffer PB was added to the column and then centrifuged for 1 minute at 17900 x g. The filtrate was discarded, 750 µL of Qiagen buffer PE was added to the column then centrifuged for 1 minute at 17900 x g. The filtrate was discarded and the column was centrifuged for 1 minute at 17900 x g. The column was placed in a clean microcentrifuge tube. The DNA was eluted by adding of 50 µL of water to the column, incubated for 1 minute then centrifuged for 1 minute at 17900 x g.

2.4.3 Polymerase Chain Reaction (PCR)

The PCR is a technique that amplifies DNA located between two known locations. DNA primers are designed to complement DNA either side of the region of interest. Following the binding of these primers to their complementary sequence to form double stranded DNA, a thermostable Taq polymerase synthesises a complementary DNA strand using excess deoxyribonucleotide triphosphates (dNTPs) in a reaction mixture.

PCR consists of three stages, which are cycled to exponentially increase the product. The first stage is a denaturing step, which leads to the template DNA becoming single stranded. Next is the annealing stage where the PCR primers anneal to their complementary sequence within the template DNA, and finally the elongation step in which the new DNA strands are synthesised. A typical PCR will have around 30 cycles of these steps carried out at specific temperatures. All PCRs were carried out using C1000/S1000 Thermal Cyclers (Bio-Rad).

2.4.3.1 PCR Primer design

PCR primers were designed using the Primer3 web resources (http://bioinfo.ut.ee/primer3-0.4.0/). Primers were designed to be approximately 20bp in length and have an annealing temperature of around 60°C.

2.4.4 Gel Electrophoresis

Due to the negatively charged phosphate groups in DNA, DNA fragments can be separated electrophoretically. If a potential difference is applied to an agarose gel loaded with DNA, the DNA will migrate toward the anode and the speed of migration will be dependent on the size of the DNA fragment.

1-3% agarose gels were used depending on the DNA fragment size or the resolution required. A 1% agarose gel was made by dissolving 1 g of high-resolution agarose (Sigma) per 100 mL of 0.5X Tris-Borate-EDTA (TBE) buffer (Thermo Fisher Scientific). This was done by heating the solution until it becomes clear and no agarose was visible. 1 µL of ethidium bromide (10 mg/mL) was added per 100 mL of agarose solution in order to visualise the DNA. This solution was then poured into a gel cast with the appropriate comb to form the number of wells desired and left to set.

Once set, the gel was loaded into an electrophoresis tank and submerged in 0.5X TBE. PCR product to be ran on the gel was mixed with 6X loading dye (NEB, Hitchen, UK) and loaded into the formed gel. An appropriate DNA ladder (Thermo Fisher Scientific) was ran along side the PCR product in order to estimate the size of the fragment. Gels are typically electrophoresed at

approximately 100V until the loading dye has migrated ¾ down the length of the gel.

The DNA separation was visualised using a UV transilluminator (UVP) and photographed using Image Lab software (Bio-Rad).

2.4.5 Extracting DNA from an agarose gel

DNA was extracted from an agarose gel using QIAquick Gel Extraction Kit (Qiagen) and subsequently purified using QIAquick PCR purification Kit (Qiagen). All centrifugation steps were carried out for 1 minute at RT at 17900 x g. The DNA fragment was excised from the gel using a scalpel. Three volumes of Qiagen Buffer GQ was added to 1 volume of the gel (1mg ~ 100 µL). This sample was incubated at 50°C until the gel slice has dissolved. One gel volume of 100% isopropanol was added to the sample and mixed. The sample was applied to a QIAquickspin column and centrifuged. The filtrate was discarded and 500 µL of Qiagen buffer QG was added to the column. The filtrate was discarded and 750 µL of Qiagen buffer PE was added to the column and centrifuged. The filtrate was discarded and the column centrifuged again. The column was placed into a clean microcentrifuge tube and 50 µL of water was added to elute DNA. The column was incubated for 1 minute then centrifuged.

To purify the DNA, 5 volumes of Qiagen buffer PB was added to 1 volume of sample and mixed. The sample was added to a QIAquick spin column and centrifuged. The filtrate was discarded and 750 µL of Qiagen buffer PE was added to the column and centrifuged. The column was placed into a clean microcentrifuge tube and 50 µL of water was added to elute DNA. The column was incubated for 1 minute then centrifuged.

2.4.6 SNaPshot genotyping

Following PCR amplification, the PCR reaction was cleaned by adding 1 µL of Shrimp alkaline phosphatase (1000 units/mL) (NEB), 0.2 µL of exonuclease 1 (20 000 units/mL) (NEB) and 1.8 µL of water directly to the 12 µL of PCR product. This reaction was incubated at 37°C for 45 minutes, followed by incubation at 85°C for 15 minutes.

The SNaPshot reaction consists of 1 µL of extension primer (1 µM), 1 µL SNaPshot® Multiplex Reaction Mix (Thermo Fisher Scientific), 2 µL of cleaned PCR product and 6µL of water. This reaction was treated in the thermocycler with 30 cycles of 96°C for 10 seconds, 50°C for 5 seconds, and 60°C for 30 seconds. 8 µL of HiDi formamide (Thermo Fisher Scientific) was added following the SNaPshot reaction. SNaPshot results were obtained by analysing samples using the Applied Biosystem 3130 Genetic Analyser and the fluorescence peak height analysed using GeneMarker (Softgenetics).

2.5 Densitometry using ImageJ

Densitometry was used to quantify bands in Western Blots and Electrophoretic Mobility Shift Assays (EMSA). This technique was used to estimate protein expression and protein DNA binding.

Gel analysis required the image to be a grey-scale image and opened in ImageJ. The Rectangle Selections tool was used to draw a rectangle around the first lane. Analyse>gels>select first lane was selected. The rectangle was moved to the next lane and analyse>gels>select next lane was selected. This

was repeated for all lanes. Once each lane had been selected, a profile plot of each lane was drawn by selecting analyse>gels>plot lanes.

The profile plot represents the relative density of the contents of the rectangle over each lane. The straight-line selection tool was used to draw a line across the base of the peak to remove background noise. The Wand tool was used to highlight each peak of interest in the profile plot. When all of the peaks were highlighted, analyse>gels>label peaks was selected. The percentages were transferred to a spreadsheet for analysis.

2.6 Experimental Design and Statistical Analysis

In this PhD, biological replicates were defined as the same experiment performed on multiple samples and was used to test the variability between samples. In terms of tissue culture, cells originating from different passages are considered different samples. Technical replicates were defined as performing the same test multiple times on the same sample, to account for variation in the protocol itself.

The number of biological replicates is stated in the appropriate chapter. Graphically represented data displays the average value across biological replicates and error bars represent standard deviation.

Data was analysed using parametric statistical analysis, primarily either by Student's t-test or one way Analysis of Variance (ANOVA). The significance level was set as $p < 0.05$. These tests assume data is continuous, independent of one another and normally distributed. These factors were taken into account during experimental design.

Prior to analysis, data was tested for homogeneity of variance using a Levene's test. In the situation where data violated the assumption of homogeneity of variance ($p < 0.05$), data was log transformed and the Levene's test was repeated.

Following a significant ANOVA result, the post-hoc Tukey's HSD test, which performs selected pairwise comparisons, was performed to identify which group means significantly differed from one another. A p value < 0.05 was considered significant.

No statistical analysis was corrected for multiple testing in this thesis. This could increase the likelihood of encountering a type 1 error. All statistical analysis was performed using IBM SPSS statistics Version 23.

3. The effect of BIN1 and CD2AP depletion on APP processing

3.1 Introduction

3.1.1 APP Processing and Trafficking

Amyloid Precursor Protein (*APP*) has long been known as a risk gene for familial Alzheimer's disease (AD) and was first discovered in 1987 (102). As the understanding of familial AD grew, this led to the generation of the amyloid hypothesis, which states that amyloidogenic processing of APP is the key catalytic event in AD pathogenesis and that Aβ generation causes all other pathological symptoms (169). Since then, genome wide association studies have identified more genes associated with Late Onset Alzheimer's disease (LOAD) and a pathway analysis of the IGAP data identified an enrichment of genes involved in the regulation of endocytosis and the clathrin/AP2 complex (180), emphasising the importance of these functional pathways in AD.

APP trafficking is critical in the processing of APP. APP is synthesised in the endoplasmic reticulum and then transported through the golgi to the trans-golgi network (TGN) where post-translational modifications occur (378). From the TGN, APP is transported to the plasma membrane. APP can then be processed via the non-amyloidogenic or amyloidogenic pathways. Non-amyloidogenic processing of APP predominantly occurs at the plasma membrane due to the abundance of α-secretases at this location (76), whereas amyloidogenic processing of APP primarily occurs intracellularly, with a large amount of evidence suggesting that it occurs after APP has been internalised (379).

APP is internalised via clathrin-mediated endocytosis (CME) (189-191). The generation of Aβ within the endosomal compartments and TGN suggests that amyloidogenic processing of APP primarily occurs at these locations, where the acidic environment is optimal for BACE1 activity (190, 380, 381). Early endosomes are weakly acidic with a pH 6.8-6.1 (382), whereas the pH of the TGN averages around 5.9 (383). It is generally accepted that the early endosome is the main site of APP cleavage by BACE1 (189, 384-386). γ-secretase cleavage of β-CTF occurs in the transmembrane protein domain mainly in the TGN and early endosomes, with some evidence it may also occur in multivesicular bodies (MVB) (84, 85). γ-secretase cleavage generates Aβ and releases the intracellular domain of APP. From here, it is thought APP is trafficked to the late endosomes, then either to the lysosomes for degradation or recycled back to the plasma membrane (189, 385, 387-389).

APP has a YENPTY motif at the carboxyl terminus that is required for efficient internalisation into CCV and early endosomes (390). Studies using a modified form of APP, which lacks this motif, observed reduced internalisation and a reduction in Aβ generation (191, 391). Increasing APP trafficking to the cell surface or reducing it's internalisation reduced amyloidogenic processing (183). Furthermore, investigations which impaired APP trafficking to the cell surface or enhanced internalisation, increased amyloidogenic processing (392, 393). This evidence demonstrates the importance of APP endocytosis in the processing of APP. Dysregulation of endocytosis is associated with AD and other neurodegenerative disorders (181) with aberrant neuronal endocytosis being one of the earliest neuropathological changes in AD (182).

APP is produced in large quantities in neurons and is rapidly metabolised (394, 395). Neurons are polarised cells with complex endosomal networks in the soma, axons and dendrites (396, 397). Endogenous APP and BACE1 are

predominantly located in the soma and dendrites of neurons and it is thought that neuronal synaptic vesicles play an important role in APP uptake (224, 398). In hippocampal neurons, recycling microdomains such as dendritic spines and presynapses, were shown to be important sites of β-cleavage and that the recycling endosomes, distributed throughout the neuronal processes, are a major site of APP and BACE1 co-localisation. In axons, APP and BACE1 interact within golgi-derived vesicles as they are co-transported (399). This implies that the location of amyloidogenic processing of APP can vary across the cellular location of neurons and could potentially be influenced by different disease mechanisms.

3.1.2 BIN1 and CD2AP in AD

BIN1 is thought to have a critical function in CME by inducing membrane curvature and recruiting dynamin. [This function is further discussed in Chapter 1.7.3]. In rat neurons, BIN1 depletion was found to increase CME and over expression resulted in reduced CME (253). In 2013 the largest LOAD GWAS identified a SNP within the *BIN1* locus, located ~30 Kb upstream of *BIN1*, to have an overall meta-p value of 6.9×10^{-44} and an OR of 1.22 (140).

CD2AP has a role in cytoskeleton regulation and interacts with a number of components of the endocytic machinery (268). [CD2AP function is further described in chapter 1.8.3]. In the 2013 meta-analysis, a SNP located in the intron of *CD2AP* was identified with a meta-p value of 5.2×10^{-11} and an OR of 1.1 (140).

It is unknown how these variants, or functional variants tagged by these associations, infer an increased risk for LOAD. As they are located outside of the coding region it could be that they effect gene regulation and the risk

mechanism is a consequence of changes in BIN1 or CD2AP expression. BIN1 expression has been investigated in AD and in AD pathology, but the findings are inconsistent. [Further discussed in chapter 1.7.4 and 1.7.5]. Little research has been done into CD2AP expression in AD.

3.1.4 Aim

BIN1 and *CD2AP* are both significantly associated with AD and have been implicated in the regulation of endocytosis, a process thought to be crucial in the pathogenesis of AD (180). It could be hypothesised that BIN1 and CD2AP infer risk for AD by affecting the same cellular mechanisms. As common variants associated with complex disease are thought to have small effect sizes, this investigation took the approach of investigating crucial disease processes in cellular models with depleted BIN1, depleted CD2AP and depleted BIN1 and CD2AP in combination to investigate the additive functional effect of loss of these proteins.

As BIN1 and CD2AP function in endocytosis, this chapter aims to characterise how depletion of BIN1 and CD2AP, independently and in combination, affect both APP processing and CME. This investigation will use the human brain derived H4 neuroglioma cell line. This cell line is a model of a brain cell widely used in AD research and is easily genetically manipulated. H4s express endogenous levels of APP and were selected to investigate the effects of BIN1 and CD2AP on physiologically relevant levels of APP.

Protein depletion will be achieved by establishing siRNA mediated depletion cellular models. The effect of loss of BIN1 and CD2AP on APP processing was established by quantifying APP, APP metabolites and BACE1. As BIN1 and CD2AP have been implicated in endocytic processes and the internalisation of

APP via CME is critical in the processing of APP, the effect of protein depletion on CME will be investigated by quantifying uptake and subsequent localisation of transferrin, a commonly used functional measure of CME.

3.2 Methods

3.2.1 Optimisation of siRNA mediated BIN1 and CD2AP depleted cellular models

The primary criteria for the protein depleted cellular models were to achieve optimal protein depletion. A range of concentrations of targeting siRNA (12.5 nM, 25 nM, 50 nM) was investigated in order to determine their effect on protein expression. Following protein depletion, a number of additional factors were investigated. This included the effect of the siRNA transfection on cell viability and the expression of the non-targeted gene. The effect of total siRNA transfected on protein expression was also investigated. Conditions with minimal effect on cell viability and off-targets effects on protein expression were selected for functional assays.

3.2.1.1 siRNA transfection

H4 cells were plated at 1.5 x 10^5 cells per well in a 6 well plate 24 hours prior to transfection. [Cells were transfected as described in Chapter 2.2.1]. During optimisation the concentration of single targeting siRNAs used per well was 50 nM, 25 nM or 12.5 nM. The sequences of the siRNAs used are described in table 3.1. ON-TARGET*plus* Non-targeting control siRNA (NT siRNA) (Dharmacon, GE, UK) was used as a control siRNA that does not target any sequence in the human genome. Cells treated with NT siRNA reflect a baseline cellular response to siRNA transfection regardless of the target.

Table 3.1 Sequence of the siRNAs used in this investigation

BIN1 targeting siRNA (5'-3')	GGAGAUGAGCAAGCUCAACTT
CD2AP targeting siRNA (5'-3')	GAUACAUGCUACUCUCCAATT

Cells were co-transfected with BIN1 and CD2AP targeting siRNAs, siRNAs targeting BIN1 and CD2AP independently and with NT siRNA. The siRNA conditions initially investigated are described in table 3.2.

Table 3.2. siRNA concentrations investigated during optimisation of depletion models.

Transfected siRNA	Concentration of BIN1 targeting siRNA per well (nM)	Concentration of CD2AP targeting siRNA per well (nM)	Concentration of Non-targeting siRNA per well (nM)
BIN1 and CD2AP targeting	50	50	0
	50	25	0
	50	12.5	0
	25	50	0
	25	25	0
	25	12.5	0
	12.5	50	0
	12.5	25	0
	12.5	12.5	0
BIN1 targeting	50	0	0
	25	0	0
	12.5	0	0
CD2AP targeting	50	0	0
	25	0	0
	12.5	0	0
NT siRNA	0	0	100
	0	0	75
	0	0	62.5
	0	0	50
	0	0	37.5
	0	0	25
	0	0	12.5

3.2.1.2 Cell Viability Assay

The effect of siRNA concentration on cell viability was investigated. A cell viability assay was performed using CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega). Each transfection condition described in table 3.2 was scaled down to a 96-well plate format. 48 hours post transfection, 20 µL of CellTiter 96® AQueous One Solution Reagent was added to wells containing a total of 100 µL of culture medium. The CellTiter 96® AQueous One Solution Reagent contains a compound which is bioreduced by metabolically active cells to form a coloured formazan product soluble in cell culture medium. The plate was then incubated for 1.5 hours in the dark at 37°C and 5% CO_2. This time point had been previously optimised in the lab for this cell type. Following incubation, the absorbance of the medium at 490nm was recorded using a BioTek µQuant™ Microplate Spectrophotometer with Gen5™ software.

Three biological replicates were performed, with two technical replicates per biological replicate. Data were normalised to a non-transfected control (NTC). A one-sampled t-test was performed to determine whether cell viability of transfected cells was significantly different to the cell viability of NTC cells.

3.2.1.3 Protein Expression Quantification

Three biological replicates of each of the transfections described in table 3.2 were performed in 6 well plates. 48 hours post transfection, cells were lysed to generate protein lysates. [Cell lysis protocol is described in Chapter 2.2.3]. Protein concentration was determined via a BCA assay [described in Chapter 2.3.1]. The effect of the siRNAs on protein expression was determined via Western blotting protocol [described in Chapter 2.3.2]. 30 µg of cell lysate was used for Western Blotting.

The antibody targeting BIN1 was ab54764 (Abcam) and was diluted 1:1000 and incubated in 2.5% BSA ON (overnight) at 4°C. The antibody targeting CD2AP was sc-25272 (Santa Cruz Biotechnology, Inc, Heidelburg, Germany) and was diluted 1:1000 and incubated in 5% milk ON at 4°C. The housekeeping gene used to normalise protein expression was GAPDH and was targeted by ab8245 (Abcam). This detects GAPDH at 37 kDa. Ab8245 was diluted 1:25000 and incubated in 5% milk ON at 4°C.

HRP Horse Anti-Mouse IgG Antibody (Peroxidase) (Vector Laboratories Ltd, Peterborough) was used as a secondary antibody to detect the BIN1, CD2AP and GAPDH primary antibodies. The secondary antibody was used at a 1:15000 dilution in 5 % milk and incubated on the blots for 1 hour at room temperature (RT) in the dark.

Protein expression was quantified by performing densitometry on the blots using ImageJ software [described in chapter 2.5]. Protein loading errors were taken into the account as each sample was normalised to its own GAPDH

expression (400). Samples are then normalised to a NTC within the same biological replicate.

3.2.2 Quantification of APP and APP processing metabolites

Following depletion optimisation, cells were transfected using the conditions described in table 3.3. Following the transfection incubation, the growth media from cells was collected, the cells were lysed and the protein lysate collected. These were samples were used in the quantification of APP processing related proteins.

Table 3.3. Transfection conditions used to generate lysates following optimisation

Co-transfection	Concentration of BIN1 targeting siRNA per well (nM)	Concentration of CD2AP targeting siRNA per well (nM)	Concentration of NT targeting siRNA per well (nM)
25B 25C	25	25	0
12.5B 25C	12.5	25	0
25B	25	0	25
12.5B	12.5	0	37.5
25C	0	25	25
50NT	0	0	50

BACE1, APP and the APP processing metabolites Aβ40, β-CTF, sAPPα and sAPPβ were quantified in these siRNA-transfected cells via ELISA.

Quantification of intracellular APP, β-CTF and BACE1 required lysate dilution to be optimised prior to the ELISA being performed. Serial dilutions of lysates were performed and these samples underwent the ELISA protocol. Protein standards were used to generate a curve and the protein levels in lysate dilutions were determined. Following optimisation, dilutions were made so that the protein concentration fell within the exponential part of the protein standards graph and would therefore generate the most accurate result. For the APP ELISA, lysates were diluted 1:15. For the β-CTF ELISA, lysates were diluted 1:3. For the BACE1 ELISA lysates were diluted 1:2.For the quantification of extracellular Aβ40, sAPPα and sAPPβ, undiluted growth media was used in the ELISA.

For the BACE1 and β-CTF ELISA, only one technical replicate was performed due to the limited lysate availability. For the remaining ELISAs, two technical replicates were performed per sample. Two technical replicates were performed for every protein standard. A total of five biological replicates were used in each ELISA. [The full ELISA protocols are described in chapter 2.3.3].

To further demonstrate APP levels, APP expression was also quantified via Western blotting. The anti-APP antibody MAB348, clone 22c11 (EMD Millipore, MA, USA) was used at a 1:1000 dilution and incubated in 1% milk ON at 4°C. This antibody is predicted to recognise all three isoforms of APP at approximatley110kDa, 120kDa and 130kDa. HRP Horse Anti-Mouse IgG Antibody (Peroxidase) (Vector Laboratories Ltd, Peterborough) was used as a secondary antibody to detect the APP primary antibody. The secondary antibody was used at a 1:15000 dilution in 5 % milk and incubated on the blots for 1 hour at RT in the dark. Expression was normalised to GAPDH expression and quantified as previously described.

3.2.4 Investigating internalisation of Alexa-488 labeled transferrin as a measure of clathrin mediated endocytosis (CME)

Transferrin is internalised via CME is commonly used as a marker for CME (401). Dextran is a commonly used molecule to determine levels of fluid phase endocytosis (402). Flow cytometry was used to quantify fluorescently labeled transferrin and dextran molecules to investigate these uptake processes in BIN1 and CD2AP depleted cells. Four transfection conditions were used when investigating CME. These are described in table 3.4.

Table 3.4. Transfection conditions used to investigate CME.

Co-transfection	Concentration of BIN1 targeting siRNA per well (nM)	Concentration of CD2AP targeting siRNA per well (nM)	Concentration of NT targeting siRNA per well (nM)
25B 25C	25	25	0
25B	25	0	25
25C	0	25	25
50NT	0	0	50

3.2.4.1 Quantification of Alexa-488 labeled transferrin uptake via flow cytometry

48 hours post transfection, cells were washed twice in Opti-MEM at RT and then incubated in Opti-MEM containing 0.2% BSA (weight/volume) (w/v) at

37°C for 30 minutes. Cells were then washed at RT with Opti-MEM and incubated for 0, 5, 15 or 30 minutes in Opti-MEM containing 100nM Alexa Fluor® 488 Transferrin Conjugate (Thermo Fisher Scientific) (Tf-488) and 80mM Alexa Fluor® 647 Dextran Conjugate (Thermo Fisher Scientific) (Dextran-647). Following incubation, tissue culture plates were immediately placed on ice to prevent further uptake. Cells were washed with ice-cold 0.1M PBS and then incubated for 1 minute in ice-cold acid wash (0.2M acetic acid, 0.2M NaCl, pH 2). This acid wash removed any conjugate bound to the surface of the cell. The cells were then washed three times in RT PBS and trypsinised for 15 minutes (150 µL of 0.25% trypsin added to each well of a 12 well plate). Following incubation, equal volumes of trypsin inhibitor were added to the cells (75µg/mL DNase, 0.5mg/mL trypsin inhibitor in PBS containing 1% BSA (w/v)). The cells were then transferred to a microcentrifuge tube and centrifuged at 325 g for 3 minutes at 4°C. Cells were then washed in PBS containing 1% BSA (w/v) and DNase (10 mg/mL) and centrifuged at 325g for 3 minutes at 4°C. Cells were then resuspended in 200 µL of PBS containing 1% BSA (w/v) and 10 mg/mL DNase, filtered through a cell strainer (Sysmex Partec, Milton Keynes) and transferred to Fluorescent Assisted Cell Sorting (FACS) tubes (BD Falcon, NY, USA).

Cell-associated fluorescence was detected by flow cytometry on a BD FACSVerse™ analyser (BD Biosciences, Oxford, UK). Cell debris and aggregates were gated out and 10000 events counted. The data were analsyed using the geometric mean fluorescent intensity. Four biological replicates were performed.

3.2.4.2 Visualisation of Alexa-488 labeled transferrin internalisation via immunocytochemistry

Visualising Tf-488 uptake

To support the FACS data, Tf-488 uptake was visualised via immunocytochemistry. Cells were plated on to sterilised glass cover slips in a 24-well plate. Cells were transfected as described in table 3.4. The Tf-488 uptake assay was performed as described in Section 3.2.4.1, however excluding Dextran-647. Cells were incubated for 5 minutes or 30 minutes with the Tf-488. A negative control was performed which involved incubation for 5 minutes with Tf-488 on ice, which should inhibit uptake. Following the acid wash, cells were washed three times in PBS and then fixed by incubating in 3% Paraformaldehyde for 20 minutes at 4°C. Following fixing, cells were washed in PBS three times and then permeabilised by incubating cells for 5 minutes in PBS containing 0.2% Triton-X 100 (Sigma). Cells were then stained with DAPI (1 µg/mL) for 5 minutes, washed twice in PBS, once in distilled water and then mounted using Vectashield (Thermo Fisher Scientific) on to glass slides. Controls were performed which omitted the acid wash step to show that the acid wash sufficiently removed surface bound Tf-488. This therefore demonstrated that Tf-488 imaged following acid wash was intracellular (data shown in appendix figure 1.1).

Determining the location of Tf-488

Immunocytochemistry was used to investigate where in the cell Tf-488 was located following uptake. The Tf-488 uptake assay was performed as described

in Section 3.2.4.1, excluding Dextran-647. Cells were incubated for 5 minutes or 30 minutes with the Tf-488. Following the acid wash, cells exposed to Tf-488 for 5 minutes were washed three times in PBS and then fixed by incubating in 3% Paraformaldehyde for 20 minutes at 4°C. Cells exposed to Tf-488 for 30 minutes were washed three times in PBS, then fixed with 100% methanol at -20°C for 15 minutes. Following fixing, cells are washed in PBS three times and then permeabilised by incubating cells for 5 minutes in PBS containing 0.2% Triton-X 100 (Sigma). Cells were then incubated in blocking solution (5% BSA, 5% Goat Serum, 0.3% Triton-X) for 1 hour.

Cells exposed to Tf-488 for 5 minutes were incubated with an anti-EEA1 antibody (Abcam, ab2900) at a 1:800 dilution in antibody diluent (0.1% BSA, 0.3% Triton-X, PBS) ON at 4°C. Cells exposed to Tf-488 for 30 minutes were incubated with an anti-LAMP2 antibody (Developmental Studies Hybridoma Bank, H4B4) at a 1:200 dilution in antibody diluent ON at 4°C.

Following primary antibody incubation, cells were washed three times with antibody diluent. Cells probed with EEA1 were incubated in anti-rabbit 594 secondary antibody (Thermo Fisher Scientific) at a 1:1000 dilution in antibody diluent in the dark for two hours at RT. Cells probed with LAMP2 were incubated in anti-mouse 594 secondary antibody (Thermo Fisher Scientific) at a 1:1000 dilution in antibody diluent in the dark for two hours at RT. Negative controls were performed, which exposed cells only to the secondary antibody, to demonstrate that observed fluorescence was due to antibody binding (images shown in appendix figure 1.2).

Following secondary antibody incubation, cells were washed in antibody diluent three times and incubated in DAPI (1 µg/mL) for 5 minutes. Cells were

then washed three times in PBS and then mounted onto glass slides using

Vectashield (Thermo Fisher Scientific).

All cells were imaged using Leica DM6000B Upright Timelapse System at x 40

magnification. A total of three biological replicates were performed.

Co-localisation was quantified using the Coloc2 ImageJ plugin software.

Images were converted into a 16-bit images and the background was

subtracted. Co-localisation was measured and the Pearson Correlation

Coefficient (PCC) was used in subsequent statistical analyses.

3.3 Results

3.3.1 Protein depletion optimisation

3.3.1.1 Effect of siRNA transfection on cell viability

Initial investigations looked at the effect of siRNA transfection on cell viability by performing a cell proliferation assay. This was to identify any cytotoxic effects of siRNA transfection. A one-sample t-test was performed to determine whether any condition significantly effected cell viability compared to a non-transfected control. This revealed that no siRNA conditions reduced cell viability. Two conditions significantly increased cell metabolism and therefore cell viability. These conditions were the co-transfection of 25 nM BIN1 and 12.5 nM CD2AP targeting siRNAs (p=0.001) and 12.5 nM CD2AP targeting siRNA only (p=0.009). These conditions significantly increased absorbance and implied an increase in cell metabolism and therefore viability. There does not appear to be an siRNA concentration dependent affect as 12.5 nM CD2AP siRNA and 25 nM BIN1 siRNA were used in other transfection conditions and no effect on cell viability was observed. To avoid any affect on cell viability, these conditions were therefore excluded from functional assays. All other siRNA conditions showed no significant change in cell metabolism when compared to non-transfected control cells (figure 3.1).

	50	50	50	50	25	25	25	12.5	12.5	12.5	12.5	50	25	12.5	25	50	12.5	-	-	-	-	-	-	-
BIN1 siRNA	50	50	50	50	25	25	25	12.5	12.5	12.5	12.5	50	25	12.5	-	-	-	-	-	-	-	-	-	-
CD2AP siRNA	50	25	12.5	50	25	25	12.5	50	25	12.5	12.5	-	-	12.5	-	-	-	-	-	-	-	-	-	-
NT siRNA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100	75	62.5	50	37.5	25	12.5

siRNA transfection conditions (nM per well)
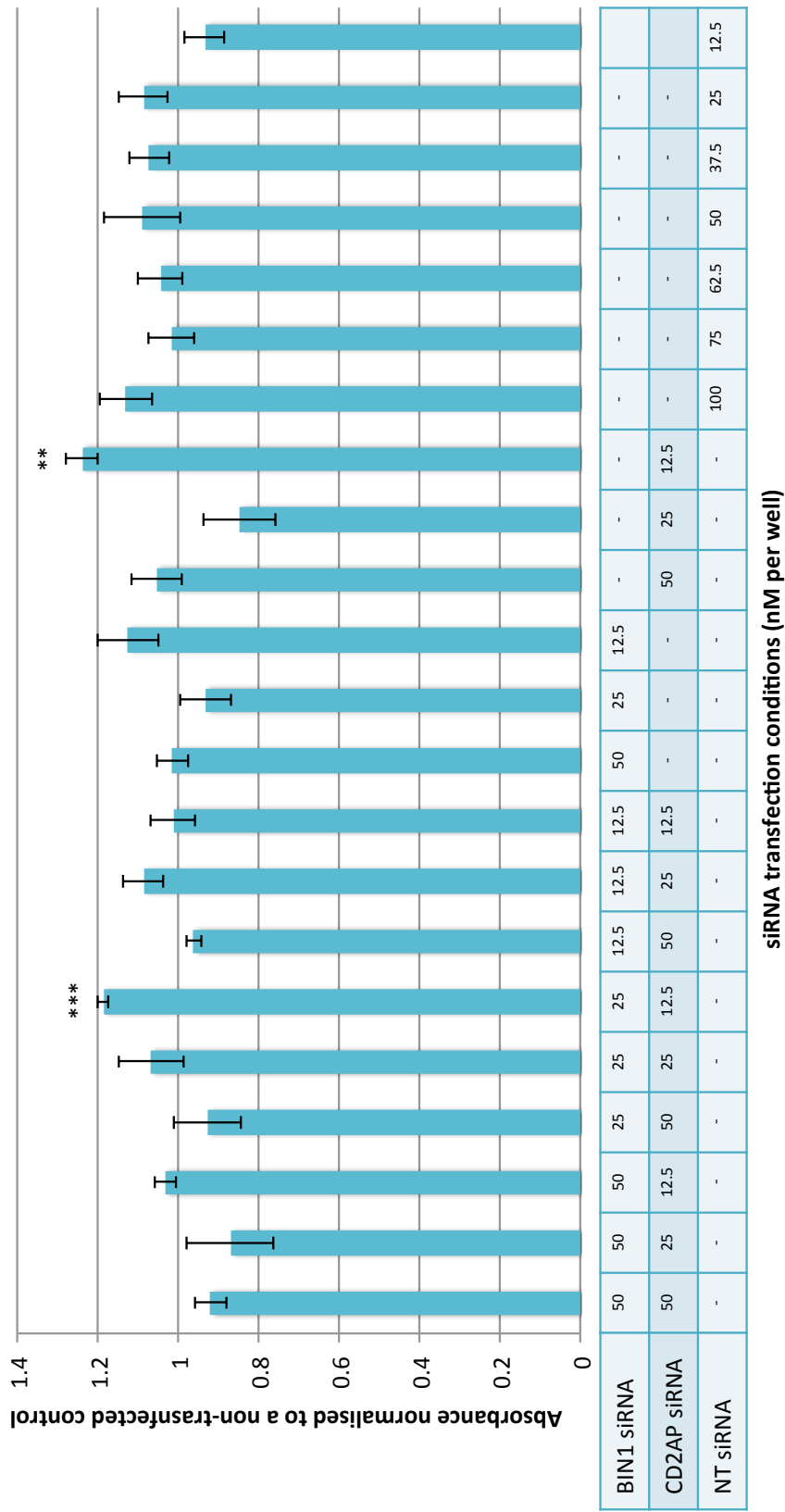
Figure 3.2 Cell proliferation assay to determine the effect of siRNA concentration on cell viability. Two technical replicates and three biological replicates were performed per siRNA condition. Data were normalised to a non-transfected control. The cell viability of two conditions was significantly different to that of the non-transfected control, 25nM BIN1 siRNA co-transfected with 12.5nM CD2AP siRNA, and 12.5nM CD2AP siRNA. In all other conditions cell viability was not significantly different to the non-transfected control. ** $p \leq 0.01$ *** $p \leq 0.001$. Error bars indicate standard error of the mean (SEM).

3.3.1.2 Effect of Non-targeting siRNA on BIN1 and CD2AP expression

To determine the effect of siRNA concentration on BIN1 and CD2AP expression, NT siRNA was transfected at different concentrations. The effect of NT siRNA concentration on BIN1 and CD2AP expression was investigated on concentrations ranging from 100 nM to 12.5 nM. Protein expression was investigated via Western blotting. Ab54764 detects BIN1 isoforms between 45-65 kDa and appears as three bands. Although, the content of each band has not been determined, the top band most likely contains isoform 5; the middle band most likely contains a combination of brain specific isoforms 3, 4, 6 and 7. The bottom band most likely contains ubiquitous isoforms 9 and 10 (403). CD2AP has a molecular mass of approximately 70 kDa and is detected as a single band (267).

A one-sampled t test revealed that no concentrations of NT siRNA had a significant effect on BIN1 or CD2AP expression levels. Statistical results are summarised in table 3.5. However, concentrations >50 nM had a variable effect on CD2AP expression therefore a total siRNA concentration above 50 nM was avoided in future experiments (figure 3.2).

Table 3.5. Summary of the one-tailed t-test showing no significant effect of NT

transfection

Concentration of NT siRNA (nM/well)	One-sampled t-test p value	
	BIN1	**CD2AP**
100	0.330	0.136
75	0.093	0.264
62.5	0.182	0.205
50	0.322	0.734
37.5	0.146	0.658
25	0.105	0.401
12.5	0.237	0.265

Figure 3.2. The effect of NT siRNA on BIN1 and CD2AP protein expression levels. A) Protein expression was quantified via Western Blotting. Protein expression was normalised to a non-transfected control and data is expressed as a percentage of normal protein expression. N=3. A one-sampled t-test revealed that NT siRNA does not significantly effect CD2AP expression at any concentration, however NT siRNA at a concentration >50nM appears to have a variable effect CD2AP expression. A one-sampled t-test revealed NT siRNA does not significantly effect BIN1 expression. Error bars indicate SEM. B) Representative Westerns demonstrating results described in A.

3.3.1.3 The effect of targeting siRNA on BIN1 and CD2AP expression

To determine the effect of BIN1 and CD2AP targeting siRNA at concentrations at 50 nM and below, targeting siRNAs were transfected alone and BIN1 and CD2AP expression investigated via Western blotting. 50, 25 and 12.5 nM of BIN1 targeting siRNA consistently produced good depletion of BIN1, with expression levels ranging from 2-12% of normal expression. However, BIN1 targeting siRNA appeared to have a concentration dependent effect on CD2AP expression, with the higher the BIN1 siRNA concentration the greater CD2AP expression. Due to the variability of the data, a one sampled t test determined this change in CD2AP expression was not significantly different to the NTC (50nM p=0.194, 25nM p=0.246, 12.5nM p=0.131). Despite not being significant, because of this observation, 50 nM concentrations of BIN1 siRNA were avoided in future investigations to minimise any off target effects.

siRNA targeting CD2AP showed almost complete depletion of CD2AP expression at all concentrations. A one-sampled t-test indicated CD2AP siRNA reduced BIN1 expression at a concentration of 50 nM and 12.5 nM (p=0.016 and p=0.002 respectively (figure 3.3). These concentrations were therefore avoided in further investigations in order to avoid any off target effects.

A

BIN1 siRNA	50	25	12.5	-	-	-	-
CD2AP siRNA	-	-	-	50	25	12.5	-

siRNA concentration (nM per well)

B

BIN1 siRNA	50	25	12.5	-	-	-	-
CD2AP siRNA	-	-	-	50	25	12.5	-

siRNA concentration (nM per well)
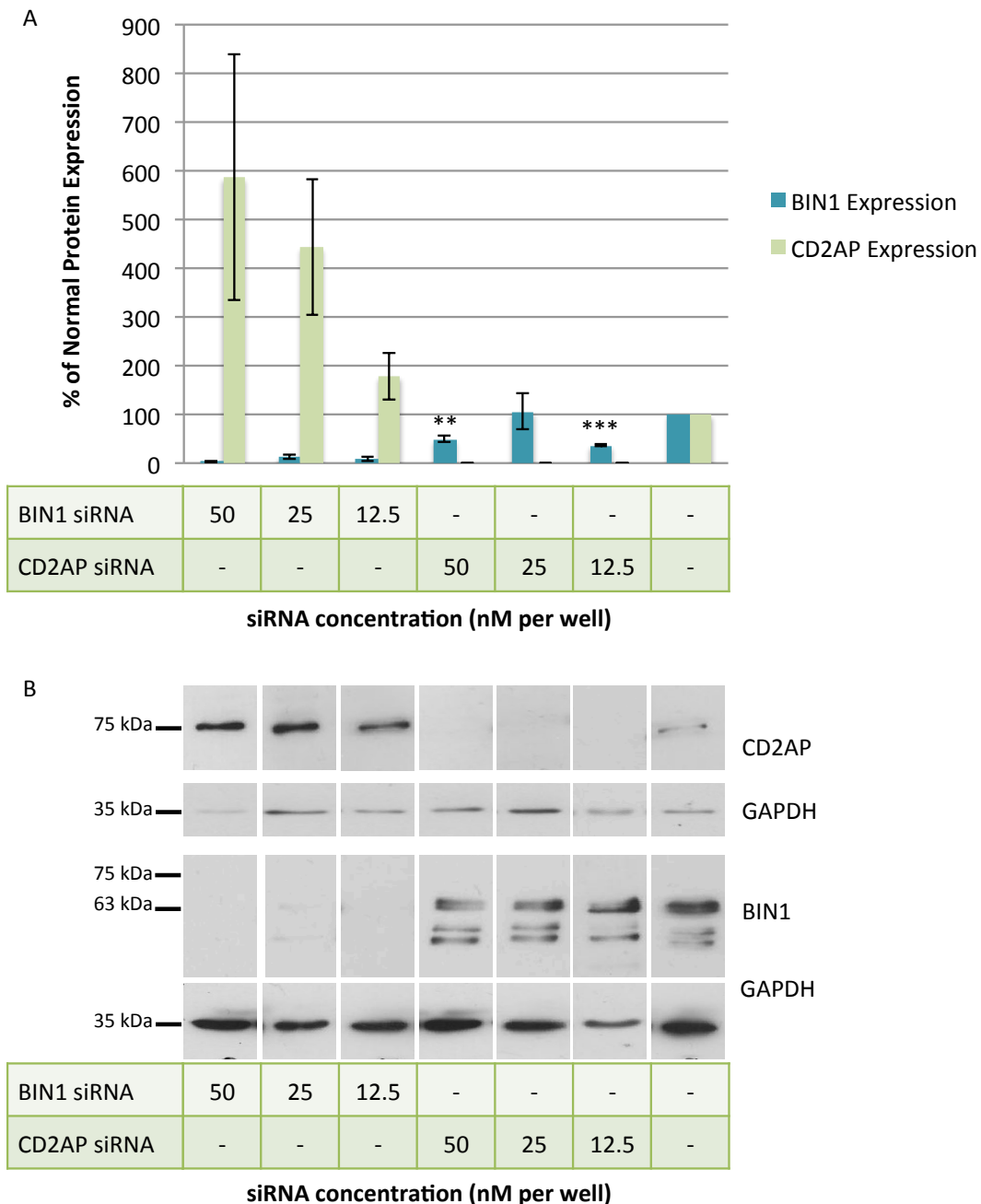
Figure 3.3. The effect of single targeting siRNAs on BIN1 and CD2AP expression. A) Protein expression was quantified via Western Blotting. Data is expressed as a percentage of normal protein expression. N=3. BIN1 siRNA appears to effect CD2AP expression is a concentration dependent manner, whilst reducing BIN1 expression to 2-12% of normal expression. A one-sampled t-test revealed CD2AP siRNA significantly reduces BIN1 expression at concentrations 50nM and 12.5nM. 25nM CD2AP targeting siRNA does not effect BIN1 expression. Complete CD2AP depletion was observed at all CD2AP siRNA concentrations. Error bars indicate SEM. B) Representative Westerns demonstrating results described in A. ** p≤0.01 ***p≤0.001.

3.3.1.4 Effect of combined BIN1 and CD2AP targeting siRNA on BIN1 and CD2AP expression

BIN1 and CD2AP siRNAs were transfected in combination in order to simultaneously deplete these two proteins. The effect of the combined BIN1 and CD2AP targeting siRNA on protein expression was quantified via Western Blotting. As a total siRNA concentration greater than 50 nM had a variable affect on CD2AP expression, combinations of BIN1 siRNA and CD2AP siRNA that totaled 50 nM or less were investigated.

This showed that a depletion of BIN1 of 10.9-35.4% of normal expression levels was achieved in all conditions. BIN1 siRNA appears to reduce all protein isoforms, except at 12.5nM BIN1 siRNA concentration where the depletion of each isoform is more variable. The greatest BIN1 depletion was observed with the 25 nM BIN1 and 25 nM CD2AP targeting siRNA transfection conditions.

In all siRNA transfection conditions, almost complete depletion of CD2AP was observed. This suggests that the potential concentration dependent effect of BIN1 siRNA on CD2AP expression is countered by using CD2AP targeting siRNA (figure 3.4).

A

| BIN1 siRNA | 25 | 25 | 12.5 | 12.5 | - |
| CD2AP siRNA | 25 | 12.5 | 25 | 12.5 | - |

siRNA transfection conditions (nM per well)

B

| BIN1 siRNA | 25 | 25 | 12.5 | 12.5 | - |
| CD2AP siRNA | 25 | 12.5 | 25 | 12.5 | - |

siRNA transfection conditions (nM per well)

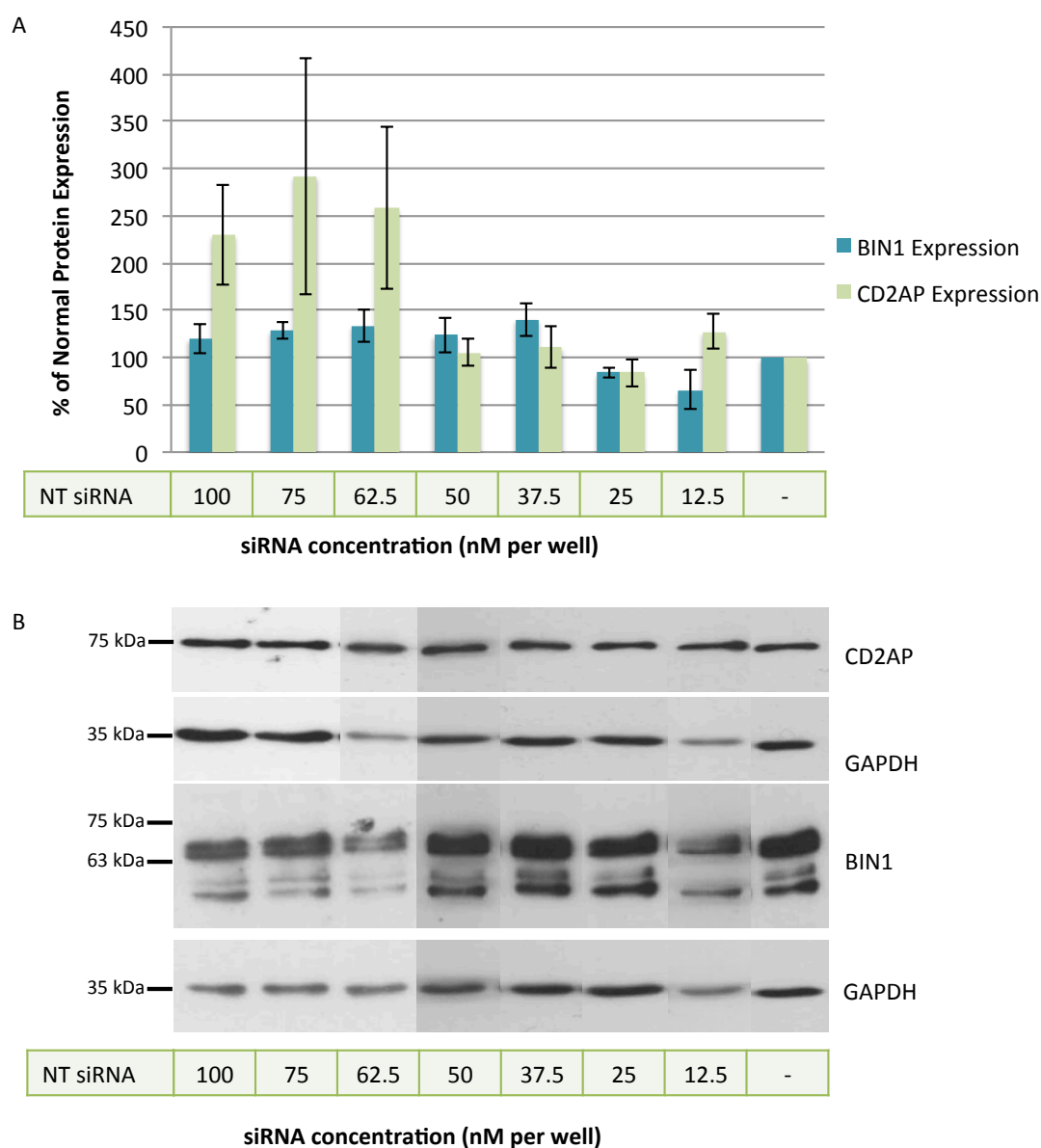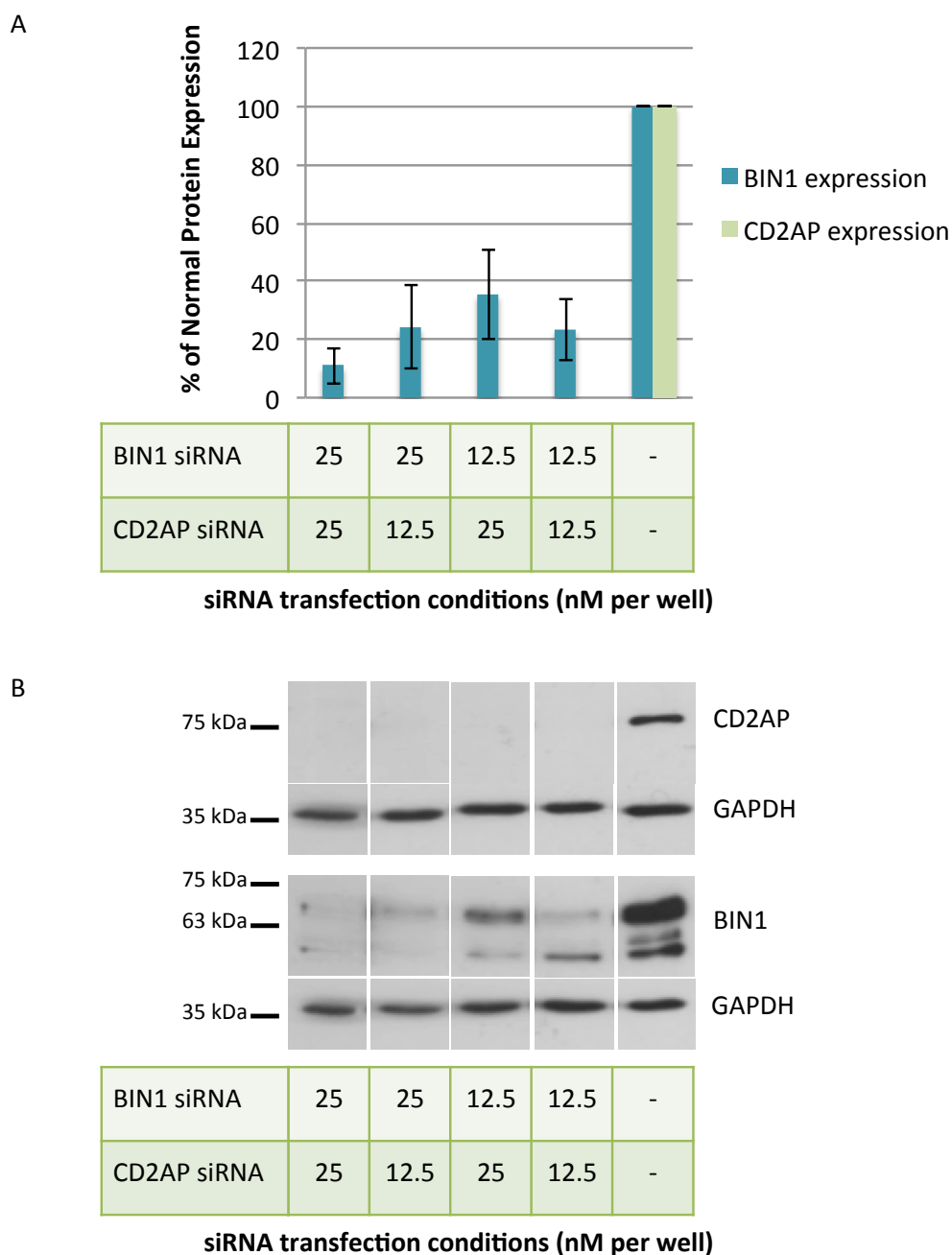**Figure 3.4 The effect of targeting siRNAs on BIN1 and CD2AP expression.** A) Protein expression was quantified via Western Blotting. Protein expression was normalised to a non-transfected control and data is expressed as a percentage of normal protein expression. N=3. Optimal BIN1 depletion was achieved when using 25nM BIN1 siRNA and 25nM CD2AP siRNA. This condition achieved 11% of the normal BIN1 expression and 0% of normal CD2AP expression. All conditions resulted in apparent complete depletion of CD2AP. Error bars indicate SEM. B) Representative Westerns demonstrating results described in A.

3.3.1.5 Optimised siRNA conditions

A concentration of NT siRNA greater than 50 nM appeared to have a variable affect on CD2AP protein levels (figure 3.2). Therefore the transfection conditions used for further work did not exceed a total siRNA concentration of 50 nM. As CD2AP siRNA at concentrations of 50 and 12.5 nM significantly reduced BIN1 expression (figure 3.3), these concentrations were avoided and 25 nM of CD2AP siRNA was used in further experiments.

25 nM BIN1 and 25 nM of CD2AP siRNA in combination produced the greatest depletion of BIN1 and CD2AP expression, reducing protein levels to 10.9% and 0% or normal protein expression levels respectively (figure 3.4). Therefore this transfection condition was selected as it produced the greatest BIN1 and CD2AP depletion and did not exceed 50 nM.

However, as BIN1 siRNA appeared to increase CD2AP expression in a concentration dependent manner (figure 3.3), a combination of 12.5 nM BIN1 and 25 nM CD2AP targeting siRNA was also used in an attempt to address any BIN1 siRNA concentration effect on CD2AP or any other affect not measured in this investigation.

BIN1 targeting siRNAs and CD2AP targeting siRNAs were transfected alone at the same concentrations used in the co-transfection to determine the effect of single protein knockdown at these concentrations. All siRNAs were transfected with NT siRNA to make up the final concentration of transfected siRNA to 50 nM to ensure every experiment is transfected with the same amount of siRNA. 50 nM of NT siRNA was used as a control. The transfection conditions used in the following experiments are summarised in table 3.6.

Table 3.6. Summary of the siRNA transfection conditions used to investigate the effect of BIN1 and CD2AP depletion in functional assays.

Transfection Condition	Concentration of BIN1 targeting siRNA (nM/well)	Concentration of BIN1 targeting siRNA (nM/well)	Concentration of BIN1 targeting siRNA (nM/well)
Depletion of BIN1 and CD2AP	25	25	0
Depletion of BIN1 and CD2AP	12.5	25	12.5
Depletion of BIN1	25	0	25
Depletion of BIN1	12.5	0	37.5
Depletion of CD2AP	0	25	25
NT siRNA control	0	0	50

3.3.1.6 Confirmation of BIN1 and CD2AP depletion

BIN1 and CD2AP depletion was confirmed prior to use in functional experiments. This was confirmed via western blotting and BIN1 and CD2AP depletion was quantified. When both BIN1 and CD2AP were targeted by siRNA (25 nM, 25 nM respectively) BIN1 expression was reduced by an average of 92.7% and CD2AP expression was reduced by and average of 98.8%. When both BIN1 and CD2AP were targeted by siRNA (12.5 nM, 25 nM respectively)

BIN1 expression was reduced by an average of 71.9% and CD2AP expression was reduced by and average of 96.8%. When BIN1 was targeted alone using 25 nM siRNA, BIN1 expression was reduced by an average of 73.4%, and when using 12.5 nM siRNA, BIN1 expression was reduced by an average of 68.6%. When targeting CD2AP alone with 25 nM of siRNA, CD2AP expression was reduced by an average of 87.6%. This data is shown in figure 3.5 and these depletion levels are assumed for the following experiments.

BIN1 siRNA	25	12.5	25	12.5	-
CD2AP siRNA	25	25	-	-	25
NT siRNA	-	12.5	25	37.5	25

siRNA concentration (nM per well)
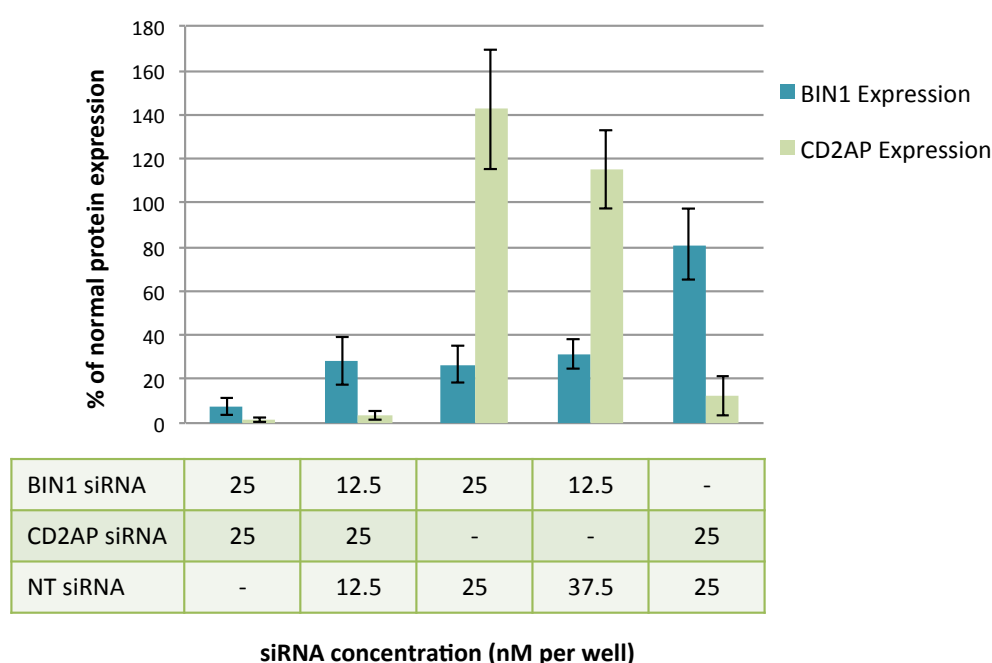
Figure 3.5 Levels of BIN1 and CD2AP in siRNA transfected cells used in downstream functional experiments. Protein expression was quantified via Western Blotting. Protein expression was normalised to a non-transfected control and data is expressed as a percentage on normal protein expression. This level of protein expression was assumed in the downstream functional assays. N=4. Error bars indicate SEM.

3.3.2 Quantifying proteins involved in APP processing

Following optimisation the effect of BIN1 and CD2AP depletion on APP processing was examined by quantifying various proteins involved in the APP

processing pathway. APP, β-CTF, Aβ40, sAPPα, sAPPβ and BACE1 were quantified in the optimised depletion models via ELISA.

3.3.2.1 BIN1 and CD2AP depletion had no effect on APP levels

Intracellular APP levels were quantified via ELISA in cellular lysates following siRNA transfection. APP levels were normalised to a NTC (treated only with media) and expressed as a percentage of this normal intracellular APP expression. A Levene's test revealed equal variances between each group (p=0.955). An ANOVA showed no transfection condition affected intracellular APP levels (p=0.913). This result indicated that the control siRNA did not have an effect on intracellular APP levels. Furthermore, reducing BIN1 and CD2AP levels had no significant effect on APP levels. This result was also demonstrated via Western Blotting (figure 3.6). Three bands were detected when probing for APP and with weights of approximately 122, 104 and 96 kDa. The contents of these bands have not been determined but most likely contain the three isoforms of APP. One isoform is predominantly expressed in neurons is 695 amino acids in length (404). The other two isoforms at 751 and 770 amino acids in length and are ubiquitously expressed (405). These bands also likely contain glycosylated APP and unglycosylated APP and products of APP processing [discussed in Chapter 1.4]. BIN1 and CD2AP depletion did not appear to affect the APP isoforms differently.

A

BIN1 siRNA	25	12.5	25	12.5	-	-
CD2AP siRNA	25	25	-	-	25	-
NT siRNA	-	12.5	25	37.5	25	50

siRNA concentration (nM per well)

B

BIN1 siRNA	25	12.5	25	12.5	-	-
CD2AP siRNA	25	25	-	-	25	-
NT siRNA	-	12.5	25	37.5	25	50

siRNA concentration (nM per well)

C
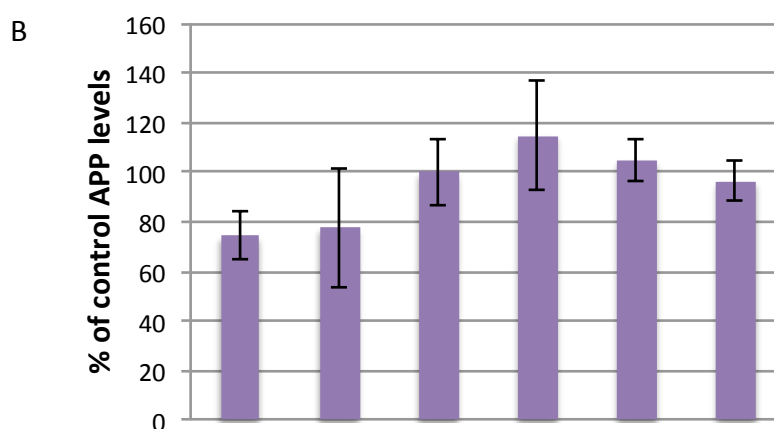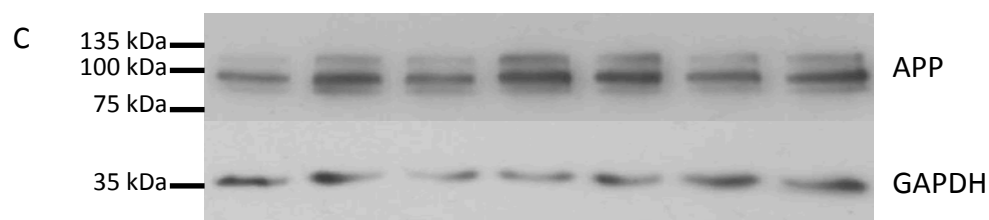
BIN1 siRNA	25	12.5	-	25	-	12.5	-
CD2AP siRNA	25	25	-	-	25	-	-
NT siRNA	-	12.5	50	25	25	37.5	-

siRNA concentration (nM per well)

Figure 3.6 The effect of BIN1 and CD2AP depletion of APP levels. A) Protein expression was quantified via ELISA. Protein expression was normalised to a non-transfected control and data is expressed as a percentage of control protein expression. N=5. BIN1 and CD2AP depletion was confirmed in lysates via Western Blotting. An ANOVA revealed no significant changes in APP expression were observed (p=0.913). B) Protein expression was quantified via Western blotting. Protein expression was normalised to a non-transfected control and data is expressed as a percentage of normal protein expression. N=3. Error bars indicate SEM. C) Representative Western demonstrating results described in B.

3.3.2.2 BIN1 and CD2AP depletion have opposing effects on β-CTF levels

β-CTF, a metabolite produced during the amyloidogenic processing of APP, was quantified in the cellular lysates produced following siRNA transfection. Data were expressed as a percentage change from normal β-CTF expression. A Levene's test revealed equal variances between the test groups (p=0.710). BIN1 and CD2AP depletion significantly affected β-CTF levels as detected by ANOVA (p≤0.001).

Using 12.5 nM siRNA to deplete BIN1 did not effect β-CTF levels when compared to the NT siRNA control (p=0.432), however a Tukey's HSD test showed a significant 30.6% increase in β-CTF levels when using 25 nM siRNA to deplete BIN1 expression when compared to the NT siRNA condition, suggesting an increase in β-secretase cleavage by of APP (p=0.029). Discrepancies in β-CTF levels between the two conditions used to deplete BIN1 may be a result of discrepancies in BIN1 depletion.

A significant 37.4% reduction in β-CTF levels was observed when using 25 nM siRNA to deplete CD2AP expression was compared to the NT siRNA control, suggesting a decrease in β-secretase cleavage by of APP (p=0.005).

Neither co-depletion condition had a significant effect of β-CTF levels when compared to NT siRNA conditions (p= 0.092 and p=0.177) or the CD2AP only depleted cells (p=794 and p=9.595. No significant difference in β-CTF levels was observed between the two BIN1 and CD2AP siRNA co-depletion conditions (p=0.999).

However, there was a significant difference in β-CTF levels between the two BIN1 and CD2AP depleted conditions and the two BIN1 only depleted conditions. 12.5 nM of BIN1 siRNA alone showed a significant increase when compared to both BIN1 and CD2AP depletion conditions (p=0.001 and p=0.003). 25 nM of BIN1 siRNA alone showed a significant increase when compared to both BIN1 and CD2AP depletion condition (p≤0.001 and p≤0.001). Data is shown in figure 3.7
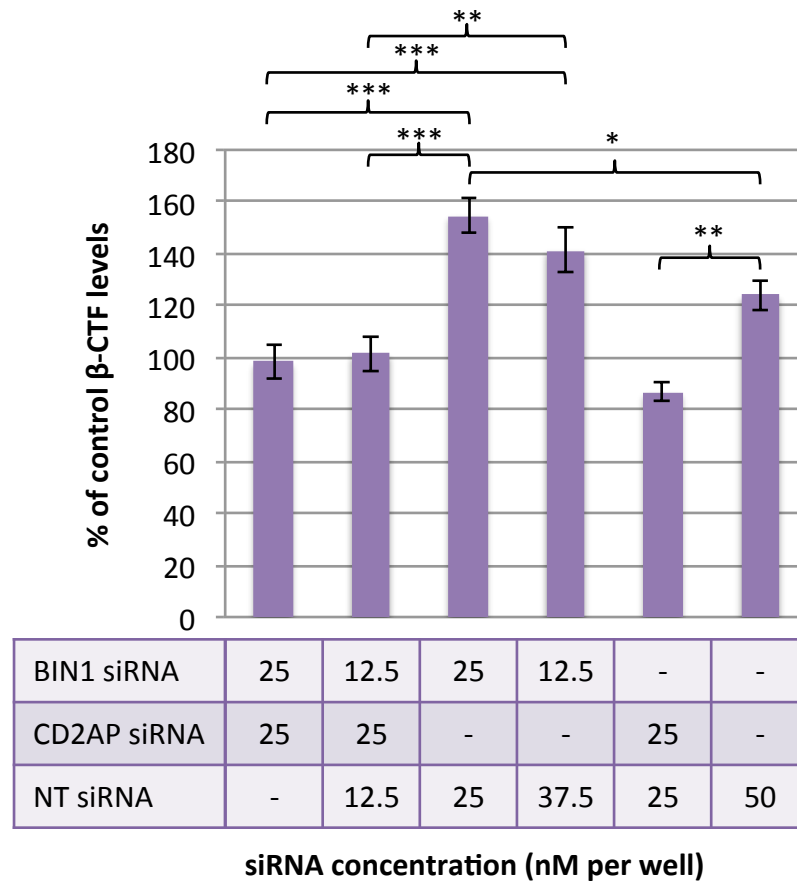
Figure 3.7. The effects of BIN1 and CD2AP depletion on β-CTF levels. β-CTF expression was quantified via ELISA. Protein expression was normalised to a non-transfected control and data are expressed as a percentage of control protein expression. N=5. BIN1 and CD2AP depletion was confirmed in lysates via Western Blotting. An ANOVA revealed significant changes in β-CTF levels were observed. When compared to the NT siRNA, Tukey's HSD test revealed 25nM BIN1 targeting siRNA significantly increased β-CTF levels (p=0.029) whereas 25nM CD2AP targeting siRNA significantly reduced β-CTF levels (p=0.005). All other conditions did not significantly affect β-CTF levels in comparison to cell treated with the NT siRNA. * p≤0.05 ** p≤0.01 ***p≤0.001. Error bars indicate SEM.

3.3.2.3 CD2AP depletion reduces Aβ40

Extracellular Aβ40 levels were quantified in the cell growth medium following siRNA transfection. A Levene's test showed no differences in variance between test groups (p=0.523). A one-way ANOVA revealed significant differences

109

between the test groups, suggesting BIN1 and CD2AP depletion significantly

affects extracellular Aβ40 levels (p=0.03). A Tukey's HSD revealed no

significant differences. However, a Dunnett's t-test was performed to compare

all groups to the NT siRNA control and this revealed a significant 43%

reduction in Aβ40 levels in the CD2AP depleted cells (p=0.023), suggesting a

reduction in γ-secretase cleavage of APP.

Neither BIN1 siRNA only conditions significantly affected Aβ40 levels (p=1 and

p=0.813) and neither co-transfection condition significantly affected Aβ40

levels (p=0.085 and 0.319). Data is shown in figure 3.8.

BIN1 siRNA	25	12.5	25	12.5	-	-
CD2AP siRNA	25	25	-	-	25	-
NT siRNA	-	12.5	25	37.5	25	50

siRNA concentration (nM per well)
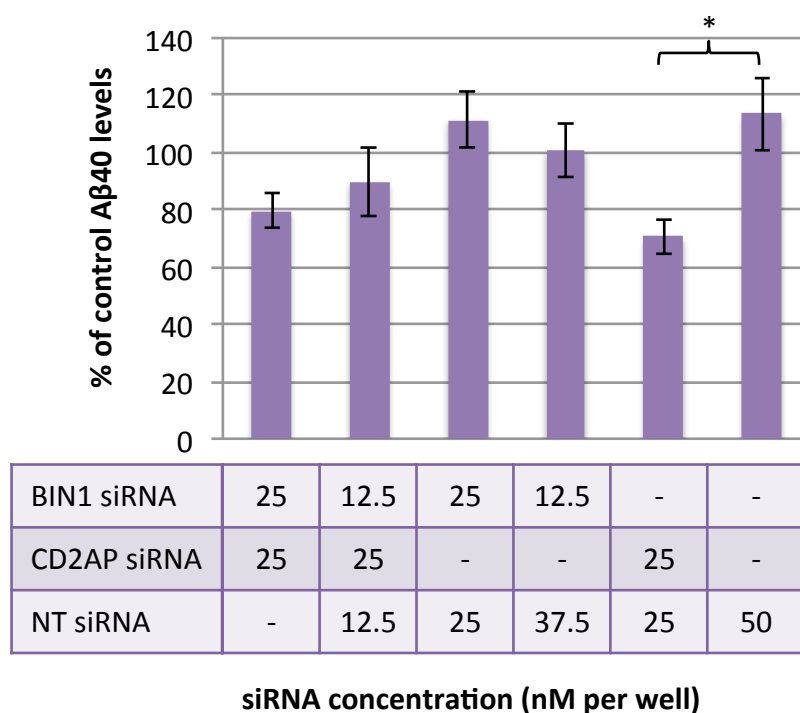
Figure 3.8. The effects of BIN1 and CD2AP depletion on extracellular Aβ40 levels. Aβ40 levels were quantified via ELISA. Data were expressed as a percentage of control Aβ40 levels. N=4. BIN1 and CD2AP depletion was confirmed in protein lysates. An ANOVA revealed that BIN1 and CD2AP depletion significantly effected Aβ40 levels (p=0.03). A Tukey's HSD test revealed no significant differences, but a Dunnett's t-test showed a significant reduction in Aβ40 levels in the CD2AP only depleted cells when compared to the NT control (p=0.023). * p≤0.05. Error bars indicate SEM.

3.3.2.4 BIN1 and CD2AP depletion has no effect of sAPPα levels

Extracellular sAPPα levels were quantified in the growth medium of cells following siRNA transfection. A Levene's test revealed no significant differences in the variances between the test groups (p=0.817). An ANOVA revealed there were no significant differences in sAPPα levels between the siRNA transfected cells (p=0.194). Data are shown in figure 3.9.

BIN1 siRNA	25	12.5	25	12.5	-	-
CD2AP siRNA	25	25	-	-	25	-
NT siRNA	-	12.5	25	37.5	25	50

siRNA concentration (nM per well)

Figure 3.9. The effects of BIN1 and CD2AP depletion on sAPPα levels. sAPPα levels were quantified via ELISA. Protein levels were normalised to a non-transfected control and data are expressed as a percentage of control sAPPα levels. N=5. BIN1 and CD2AP depletion was confirmed in protein lysates via Western Blotting. An ANOVA revealed no significant changes in sAPPα levels were observed (p=0.194). Error bars indicate SEM.

3.3.2.5 BIN1 and CD2AP depletion do not affect sAPP*β* levels

Extracellular sAPPβ levels were quantified in the growth medium of cells following siRNA transfection. A Levene's test revealed no significant differences in the variances between the test groups (p=0.338). An ANOVA revealed there were no significant differences in sAPPβ levels between the siRNA transfected cells (p=0.716). Data is shown in figure 3.10.

BIN1 siRNA	25	12.5	25	12.5	-	-
CD2AP siRNA	25	25	-	-	25	-
NT siRNA	-	12.5	25	37.5	25	50

siRNA concentration (nM per well)
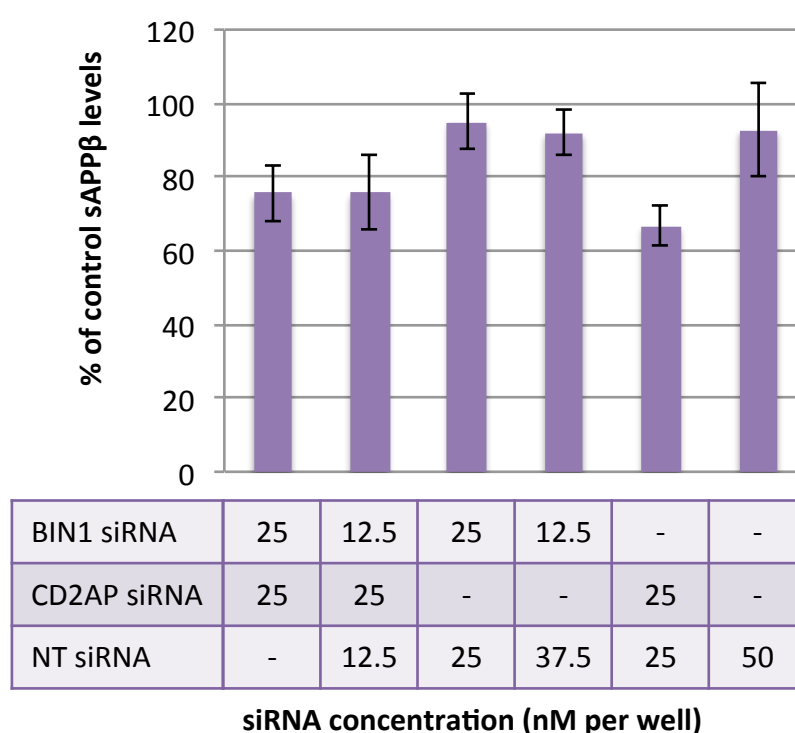
Figure 3.10. **The effects of BIN1 and CD2AP depletion on sAPPβ levels.** sAPPβ levels were quantified via ELISA. Protein levels were normalised to a non-transfected control. Data is expressed as a percentage of control sAPPβ levels. N=5. BIN1 and CD2AP depletion was confirmed in protein lysates via Western Blotting. An ANOVA revealed no significant changes in sAPPβ levels were observed (p=0.716). Error bars indicate SEM.

3.3.2.6 BIN1 and CD2AP depletion has no effect on sAPPβ:sAPPα ratio

Since no differences were observed in the absolute levels of either sAPPα or sAPPβ, extracellular sAPPα and sAPPβ levels were expressed as a ratio to determine whether there is any change in sAPP production relative to each other. A Levene's test showed there was no significant differences in variance between test groups (p=0.131). A one-way ANOVA revealed no significant changes in sAPPβ:sAPPα between any of the siRNA conditions (p=0.393). Data is shown in figure 3.11.

BIN1 siRNA	25	12.5	25	12.5	-	-
CD2AP siRNA	25	25	-	-	25	-
NT siRNA	-	12.5	25	37.5	25	50

siRNA concentration (nM per well)
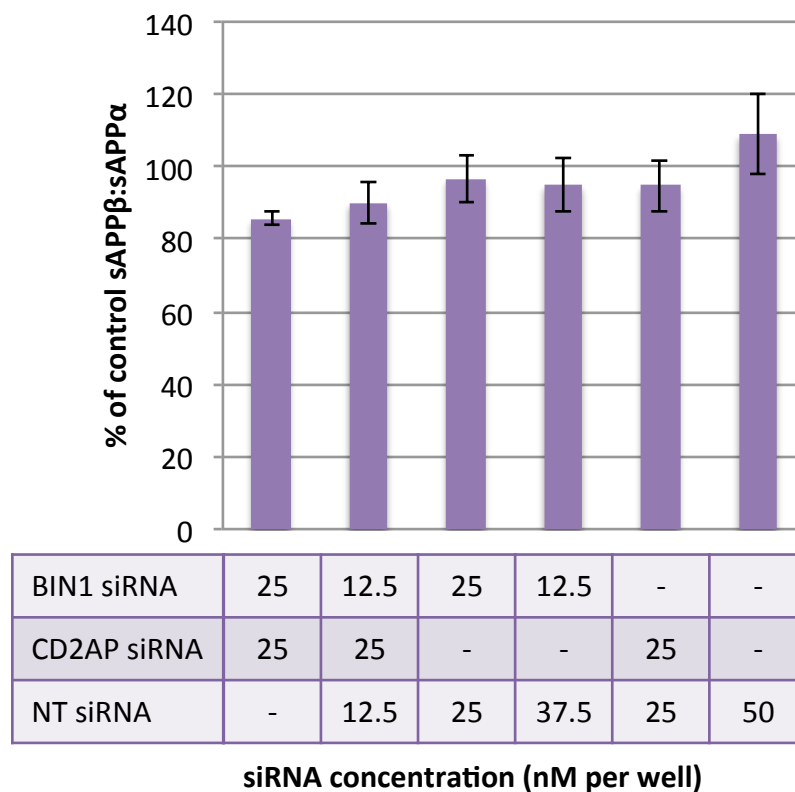
Figure 3.11. The effects of BIN1 and CD2AP depletion on sAPPβ:sAPPα ratio. sAPPα and sAPPβ levels were quantified via ELISA. Ratios were normalised to a non-transfected control and the data is expressed as a percentage of the control sAPPβ:sAPPα ratio. N=5. Protein depletion was confirmed in protein lysates via Western Blotting. An ANOVA revealed no significant changes in sAPPβ:sAPPα ratio were observed (p=0.393). Error bars indicate SEM.

3.3.2.7 CD2AP depletion reduces BACE1 levels

Due to no significant differences in APP processing being observed between the two BIN1 and CD2AP siRNA co-transfection conditions, only the 25 nM BIN1 and 25 nM CD2AP targeting siRNA condition was used when investigating BACE1 levels. A Levene's test showed significant differences in variance between the test groups (p=0.021). The data was then log transformed which created equal variances (p=0.082).

An ANOVA was performed which revealed a significant difference in BACE1 levels between the test groups (p≤0.001). BACE1 levels in BIN1 only depleted cells were not affected in comparison to NT siRNA control levels (p=0.808), but were significantly increased in comparison to the BIN1 and CD2AP depleted in combination and the CD2AP only depleted cells (p≤0.001 and p=0.005, respectively). CD2AP only depleted cells had significantly decreased levels of BACE1 in comparison to both BIN1 depleted cells and the NT siRNA control (p=0.005 and p=0.032).

However, BIN1 and CD2AP depletion resulted in a significant 83.1% reduction in BACE1 levels from control levels when compared to the NT siRNA control (p≤0.001) and BACE1 levels were significantly reduced when compared to BIN1 only depleted cells (p≤0.001). Furthermore, there was no significant difference between BIN1 and CD2AP depleted cells and CD2AP only depleted cells (adjusted p=0.156). This data is shown in figure 3.12.

BIN1 siRNA	25	25	-	-
CD2AP siRNA	25	-	25	-
NT siRNA	-	25	25	50

siRNA concentration (nM per well)

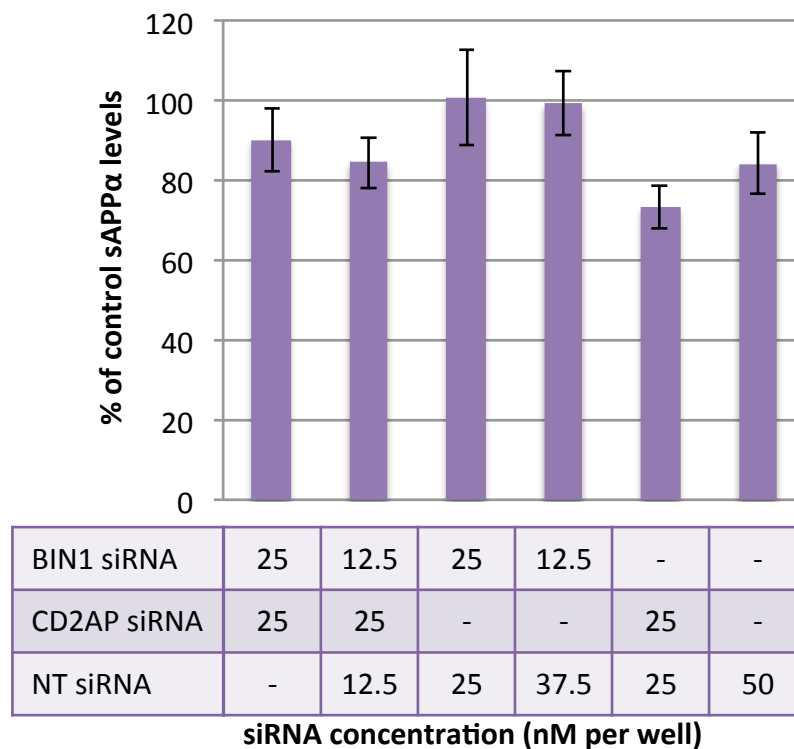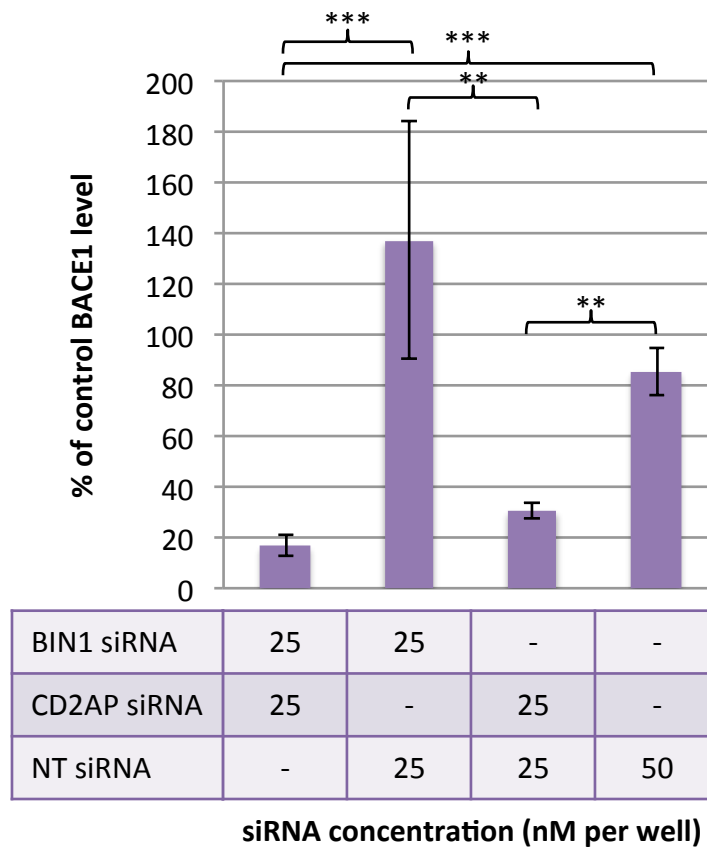**Figure 3.12. The effects of BIN1 and CD2AP depletion on BACE1 levels.** BACE1 levels were quantified via ELISA. Protein levels were normalised to a non-transfected control. Data are expressed as a percentage of the normal BACE1 levels. N=5. BIN1 and CD2AP depletion was confirmed in protein lysates via Western Blotting. Significant differences were detected between the BIN1 and CD2AP depleted cells and the BIN1 depleted cells and the NT control. CD2AP only depleted cells were significantly different to the BIN1 only depleted cells and the NT control. No difference between the NT control and BIN1 only depleted cells were observed. * $p \leq 0.05$ ** $p \leq 0.01$ *** $p \leq 0.001$. Error bars indicate SEM.

Table 3.7. Summary of BIN1 and CD2AP depletion of APP processing related proteins.

Protein	BIN1 depletion	CD2AP depletion	BIN1 and CD2AP depletion
APP	No effect	No effect	No effect
β-CTF	Increase	Decrease	No effect
Aβ40	No effect	Decrease	No effect
sAPPα	No effect	No effect	No effect
sAPPβ	No effect	No effect	No effect
sAPPβ:sAPPα	No effect	No effect	No effect
BACE1	No effect	Decrease	Decrease

3.3.3 Quantifying cellular uptake

3.3.3.1 Experimental conditions used to investigate cellular uptake.

In the previous experiments, BIN1 and CD2AP depletion together was achieved using two different combination of targeting siRNA, 25 nM BIN1 + 25 nM CD2AP targeting siRNA and 12.5 nM BIN1 + 25 nM CD2AP. This was done to avoid a potential off-target effect on the BIN1 siRNA on CD2AP expression observed during optimisation (figure 3.3). As no differences in APP or APP metabolites levels were detected between the two BIN1 and CD2AP depletion conditions, and for practicality reasons, only one co-depletion model was used in the following experiments. The siRNA conditions used in the following experiments, which focuses on investigating cellular uptake, are described in table 3.8.

Table 3.8. siRNA conditions used in the investigation of cellular uptake

Transfection Condition	Concentration of BIN1 targeting siRNA (nM/well)	Concentration of BIN1 targeting siRNA (nM/well)	Concentration of BIN1 targeting siRNA (nM/well)
Depletion of BIN1 and CD2AP	25	25	0
Depletion of BIN1	25	0	25
Depletion of CD2AP	0	25	25
NT siRNA control	0	0	50

3.3.3.2 CD2AP depletion increase CME

Transferrin is a common marker for functional CME and therefore Tf-488 uptake was quantified over time in each of the cellular depletion models and the geometric mean fluorescent intensity plotted (figure 3.13). Differences in Tf-488 levels were analysed by ANOVA then Tukey's HSD.

Figure 3.13. The uptake of Tf-488 over time by BIN1 and CD2AP depletion models. The geometric mean of fluorescent intensity is plotted against time to track the uptake of tf-488 over time. N=4. Error bars indicate SEM. Data was statistically analysed by an ANOVA then a Tukey's HSD. At 0 minutes, no statistical differences in Tf-488 levels were detected (p=0.693). At 5 minutes, there was no significant difference between NT and BIN1 depletion (p=0.884). A significant difference between NT and BIN1 and CD2AP co-depletion (p=0.021) and CD2AP only depleted cells (p=0.032). At 15 minutes, there was no significant difference between NT and BIN1 depletion (p=0.909), but there was a significant difference between NT and BIN1 and CD2AP co-depletion (p=0.001) and between NT and CD2AP only depleted cells (p=0.005). At 30 minutes, there was no significant difference between NT and BIN1 depletion (p=0.829), but there was significant difference between NT and BIN1 and CD2AP co-depletion (p=0.001) and between NT and CD2AP only depleted cells (p=0.005).

After 0 minutes of Tf-488 exposure, a Levene's test revealed no differences in variance between any of the siRNA conditions (p=0.916). An ANOVA revealed no changes in Tf-488 levels (p=0.693). This indicates that all conditions began with the same background level of Tf-488.

After 5 minutes of Tf-488 exposure, a Levene's test showed equal variances between the siRNA conditions (p=0.326). An ANOVA revealed a significant change in Tf-488 levels between the siRNA conditions (p=0.009). BIN1 depleted cells showed no differences in Tf-488 levels when compared to the NT siRNA control (p=0.884) but the CD2AP depleted cells showed at 2.03 fold increase in Tf-488 uptake when compared to the NT siRNA control (p=0.032). There was no significant difference between BIN1 only and CD2AP only depleted cells (p=0.110). BIN1 and CD2AP depleted cells showed a significant 2.1 fold increase in Tf-488 uptake when compared to the NT siRNA control (p=0.021), but showed no difference to CD2AP only depleted cells (p=0.996) or BIN1 only depleted cells (p=0.076).

After 15 minutes of Tf-488 exposure, a Levene's test showed equal variances between the siRNA conditions (p=0.483). An ANOVA revealed a significant change in Tf-488 levels between the siRNA conditions (p≤0.001). BIN1 depleted cells showed no differences in Tf-488 levels when compared to the NT siRNA control (p=0.909). However CD2AP depleted cells showed at 1.72 fold increase in Tf-488 levels when compared to the NT siRNA control (p=0.005) and CD2AP only depleted cells had a significant 1.55 fold increase when compared to BIN1 only depleted cells (p=0.015). The BIN1 and CD2AP depleted cells showed a significant 1.92 fold increase in Tf-488 uptake when compared to the NT siRNA control (p=0.001). The BIN1 and CD2AP depleted cells showed no difference to CD2AP only depleted cells (p=0.63), but showed

a significant 1.73 fold increase when compared to BIN1 only depleted cells (p=0.002).

After 30 minutes of Tf-488 exposure, a Levene's test showed equal variances between the siRNA conditions (p=0.116). An ANOVA revealed a significant change in Tf-488 levels between the siRNA conditions (p≤0.001). BIN1 depleted cells showed no differences when compared to the NT siRNA control (p=0.829). However, CD2AP depleted cells showed at 1.76 fold increase in Tf-488 uptake when compared to the NT siRNA control (p=0.005) and CD2AP only depleted cells had a significant 1.53 fold increase when compared to BIN1 only depleted cell (p=0.021). BIN1 and CD2AP depleted cells showed a significant 1.92 fold increase in Tf-488 uptake when compared to the NT siRNA control (p=0.001). The BIN1 and CD2AP depleted cells showed no difference to CD2AP only depleted cells (p=0.794), but showed a significant 1.67 fold increase compared to BIN1 only depleted cells (p=0.004). Data is shown in figure 3.13.

3.3.3.3 BIN1 and CD2AP depletion has no effect on fluid phase uptake

Dextran uptake is used as a fluid phase endocytosis marker, which is a non-specific process involving the uptake of extracellular fluid (406). Dextran-647 uptake was quantified in the depletion models over time and the geometric mean fluorescent intensity plotted (figure 3.14). This was an attempt to determine if CME is exclusively affected by BIN1 and CD2AP depletion and the changes in Tf-488 levels observed were not due to a change in continuous non-specific uptake. Differences in Dextran-647 levels were analysed by ANOVA.
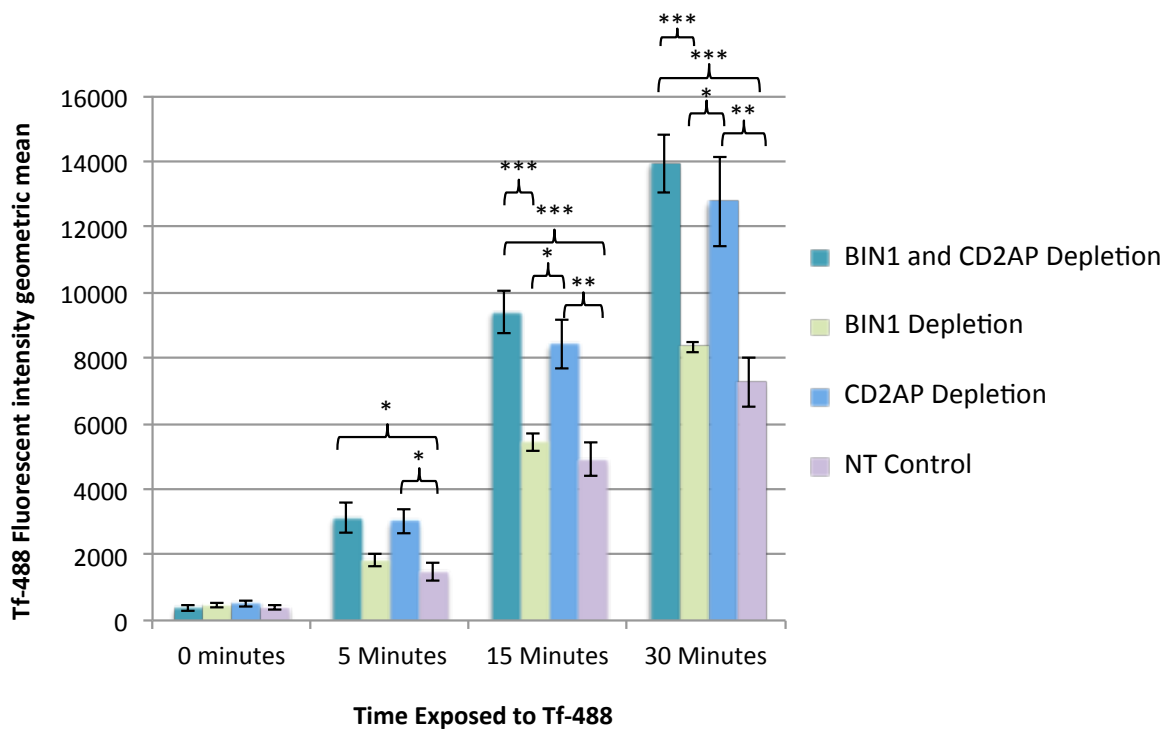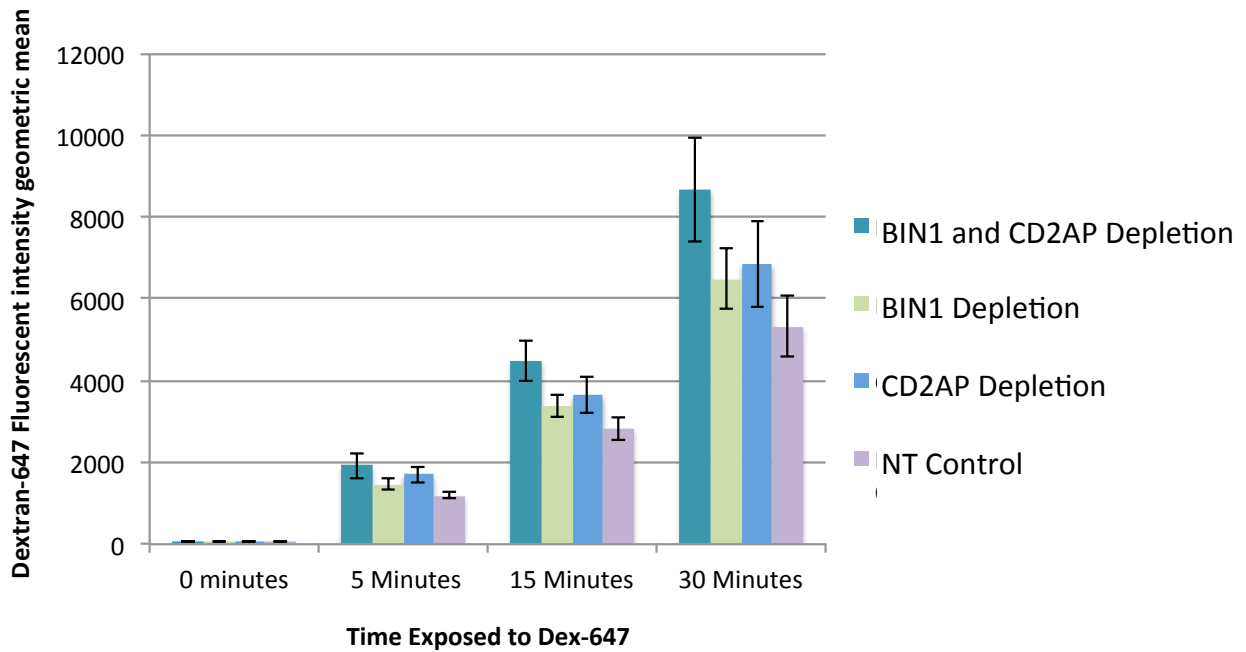
Figure 3.14. The uptake of Dextran-647 over time in depletion models. The geometric mean of fluorescent intensity is plotted against time to track the uptake of Dextran-647 over time. N=4. Error bars indicate SEM. ANOVA analysis revealed no significant changes in Dextran-647 uptake between all transfection conditions at all time points.

After 0 minute exposure to Dextran-647, a Levene's test showed equal variances between the siRNA conditions (p=0.28). An ANOVA revealed no significant changes in Dextran-647 uptake between the siRNA conditions (p=0.75). After 5 minutes exposure to Dextran-647, a Levene's test showed equal variances between the siRNA conditions (p=0.255). An ANOVA revealed no significant changes in Dextran-647 uptake between the siRNA conditions (p=0.121). After 15 minutes exposure to Dextran-647, a Levene's test showed equal variances between the siRNA conditions (p=0.575). An ANOVA revealed no significant changes in Dextran-647 uptake between the siRNA conditions (p=0.064). After 30 minutes exposure to Dextran-647, a Levene's test showed equal variances between the siRNA conditions (p=0.635). An ANOVA revealed no significant changes in Dextran-647 uptake between the siRNA conditions (p=0.166).

3.3.4 Visualising Tf-488 uptake via CME in BIN1 and CD2AP depleted cells

3.3.4.1 CD2AP depletion increases uptake of Tf-488 after 30 minutes

BIN1 and CD2AP depleted cells and NT siRNA control cells were exposed to Tf-488 for either 5 minutes, 30 minutes or on ice for 5 minutes. CME should be inhibited when cells are on ice therefore no Tf-488 uptake should occur. The transfected cells were then stained with DAPI and imaged at 40x magnification. All cells were imaged at the same exposure in order to image Tf-488 uptake over time.

When the BIN1 and CD2AP depleted cells and NT siRNA control cells were incubated on ice and exposed to Tf-488 for 5 minutes, no Tf-488 uptake was observed. This is consistent with CME inhibition and therefore no uptake occurring. Representative images are shown in figure 3.15.

BIN1 and CD2AP depleted cells and NT siRNA controls were exposed to Tf-488 at 37°C, where CME is not inhibited. After 5 minutes of Tf-488 exposure at 37°C, trace levels of Tf-488 can be observed in the BIN1 and CD2AP depleted cells (indicated by arrow in figure 3.16 A). This is consistent with the FACS results, which described an increase Tf-488 levels after 5-minute exposure (Figure 3.13). Tf-488 could not be detected in any other transfection condition. Images are shown in figure 3.16.

5 minutes on ice

A   BIN1 and CD2AP depletion

B   BIN1 depletion

C   CD2AP depletion
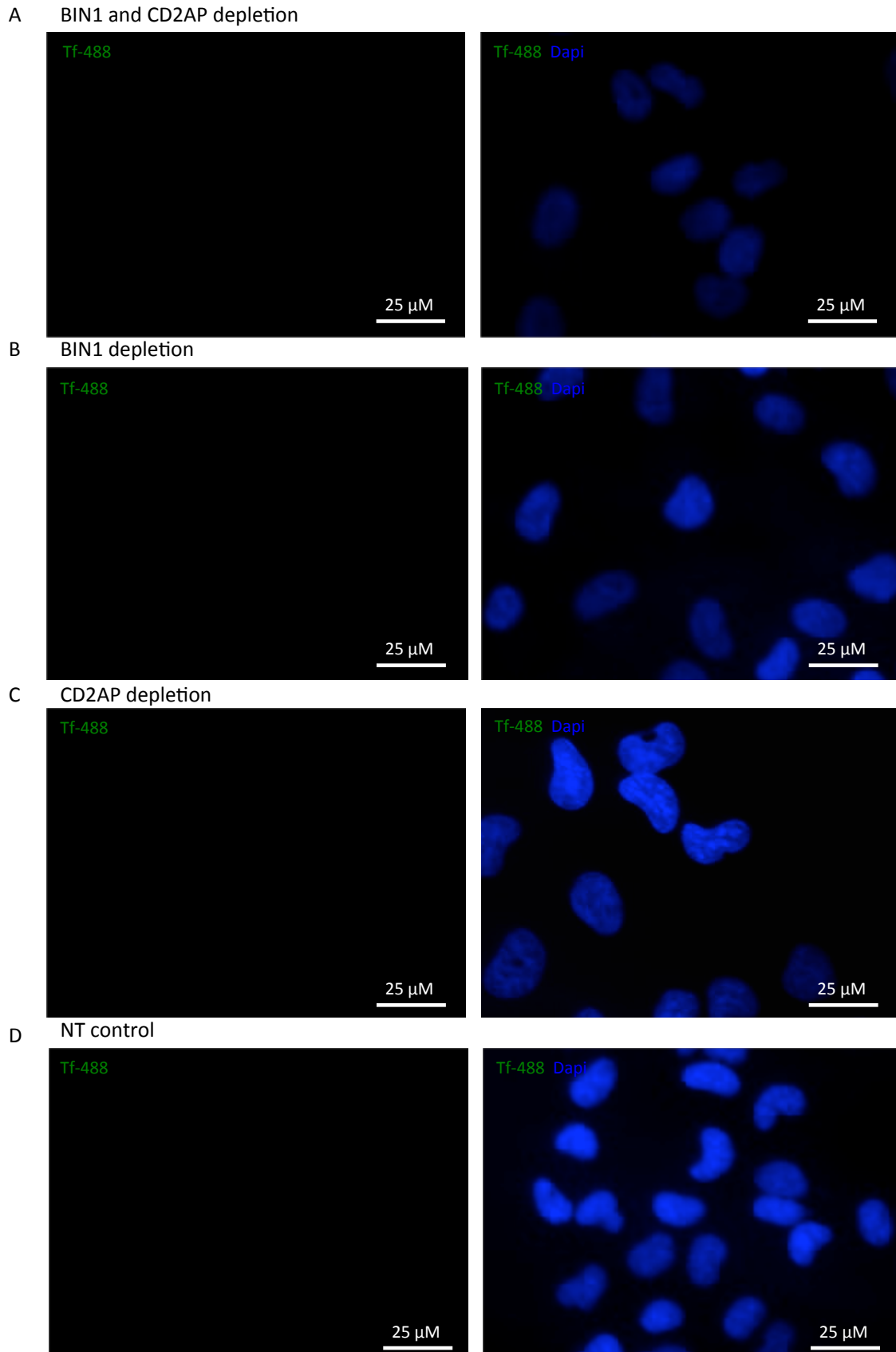
D   NT control

Figure 3.15. Tf-488 uptake in cells incubated on ice for 5 minutes. A) BIN1 and CD2AP depleted cells. B) BIN1 depleted cells. C) CD2AP depleted cells. D) NT control cells. In all conditions no Tf-488 can be detected. N=3. BIN1 and CD2AP depletion was confirmed via Western blotting. Left column shows Tf-488 only. Right column shows Tf-488 and DAPI.   123

5 minutes

A   BIN1 and CD2AP depletion

B   BIN1 depletion

C   CD2AP depletion

D   NT control

Figure 3.16. Tf-488 localisation after 5-minute exposure. A) BIN1 and CD2AP depleted cells. B) BIN1 depleted cells. C) CD2AP depleted cells. D) NT control cells. N=3. BIN1 and CD2AP depletion was confirmed via Western blotting. A trace amount of Tf-488 is detected in BIN1 and CD2AP depleted cells (indicated by arrow). No Tf-488 is detected in other cell populations. Left column shows Tf-488 only. Right column shows Tf-488 and DAPI.

After 30 minutes of Tf-488 exposure at 37°C, Tf-488 is visible in BIN1 and CD2AP depleted cells and the NT siRNA control. In the NT siRNA control, there appears to be diffuse detection of Tf-488 in the perinuclear region (figure 3.17 D). This diffuse localisation made it difficult to focus on during imaging. In contrast, the BIN1 depleted cells Tf-488 appears more punctate but again, located in the perinuclear region (figure 3.17 B). CD2AP depleted cells appear to also have punctate perinuclear localisation of Tf-488, in contrast to the NT siRNA control, and Tf-488 appears to be present in a slightly greater quantity than in the BIN1 depleted and NT siRNA control cells (figure 3.17 C).

BIN1 and CD2AP depleted cells appear to have more punctate Tf-488 localisation in comparison to the NT control. Tf-488 again appears to be located in the perinuclear region and appears to be present in slightly greater quantities than in the NT siRNA control or the BIN1 only depleted cells (figure 3.17 A). All images taken after 30 minutes are shown in figure 3.17.

The increase in Tf-488 uptake observed in the BIN1 and CD2AP depleted cells and the CD2AP only depleted cells in comparison to the BIN1 only depleted cells and NT siRNA control, is consistent with the FACS data at 30 minutes shown in figure 3.13.

30 minutes

A   BIN1 and CD2AP depletion

B   BIN1 depletion

C   CD2AP depletion
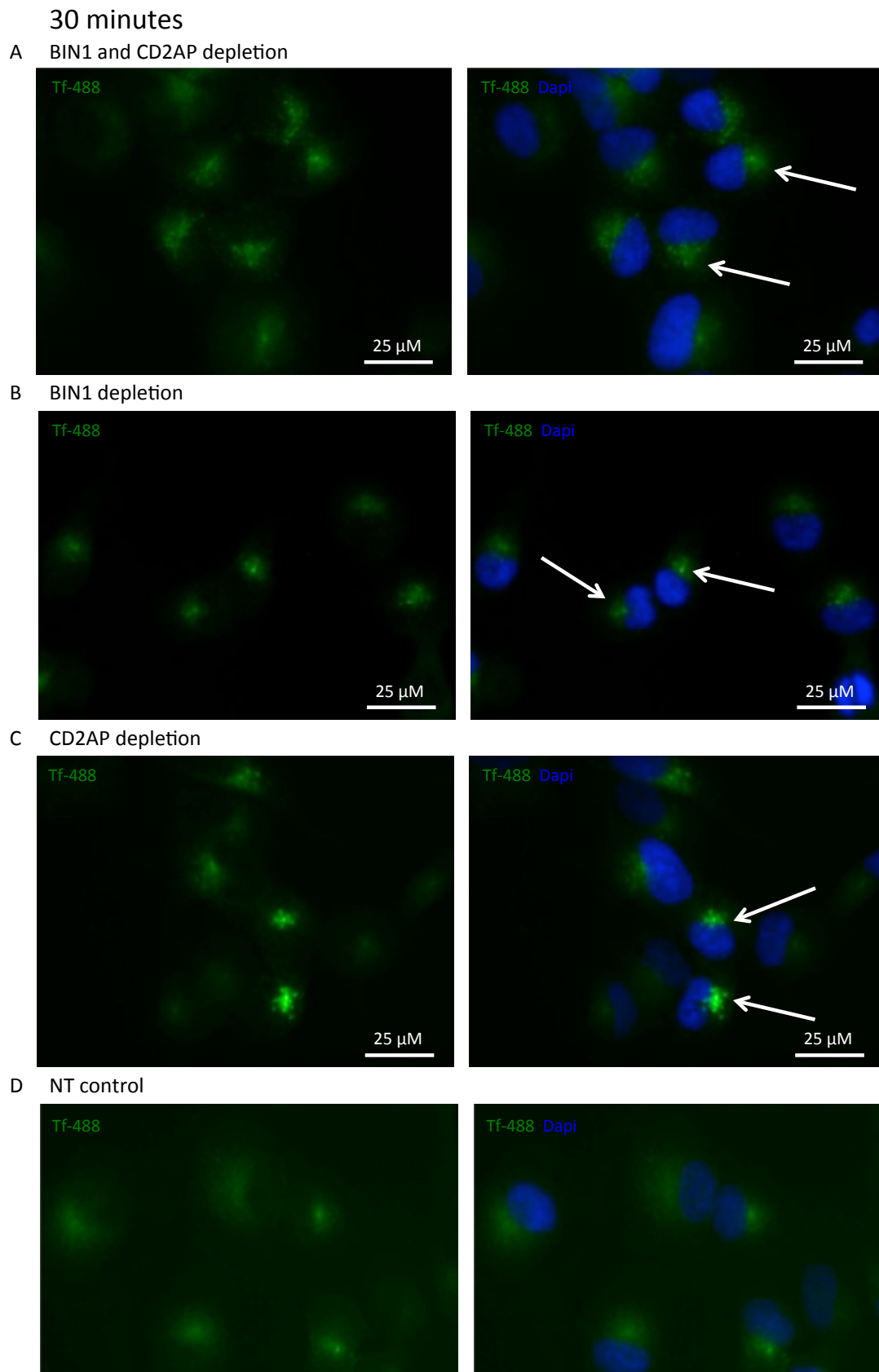
D   NT control

Figure 3.13. Tf-488 localisation after 30 minutes of exposure. A) BIN1 and CD2AP depleted cells. B) BIN1 depleted cells. C) CD2AP depleted cells. D) NT control cells. N=3. BIN1 and CD2AP depletion was confirmed via Western blotting. Diffuse perinuclear localisation of Tf-488 was detected in the NT control. Arrows indicate areas with increased punctate staining. Increased levels of Tf-488 are observed in A and C in comparison to B and D. Left column shows Tf-488 only. Right column shows Tf-488 and DAPI.

3.3.4.2 Tf-488 is localised to the early endosome after 5 minutes

All transfected cells were exposed to Tf-488 for 5 minutes. Cells were then probed with EEA1, an early endosome marker, and then stained with DAPI. Cells were imaged at 40x magnification and at optimal exposure for imaging Tf-488 after 5 minutes, which was consistent between biological replicates.

Tf-488 and EEA1 appears punctate and located in the perinuclear region. All conditions show some level of co-localisation between Tf-488 and EEA1, however complete localisation in not observed. This suggests some Tf-488 is located in the early endosome. EEA1 and Tf-488 co-localisation was quantified and a PCC value was generated. A Levene's test revealed equal variances between transfected populations (p=0.348). A one-way ANOVA revealed there were no significant differences in co-localisation between any of the transfected cell populations (p=0.648). Representative images are shown in figure 3.18.

5 minutes at optimal exposure

A   BIN1 and CD2AP depletion
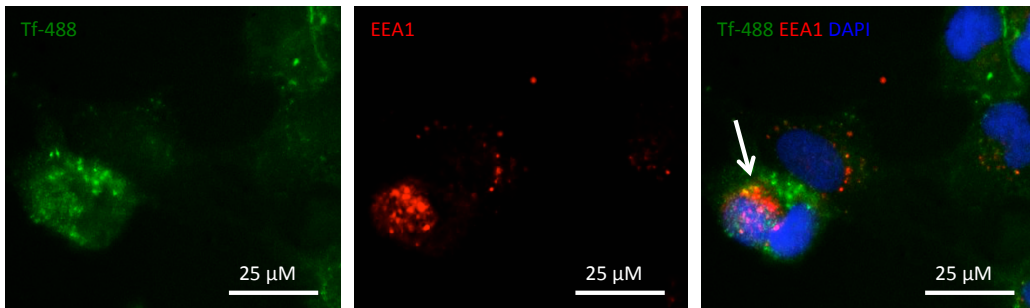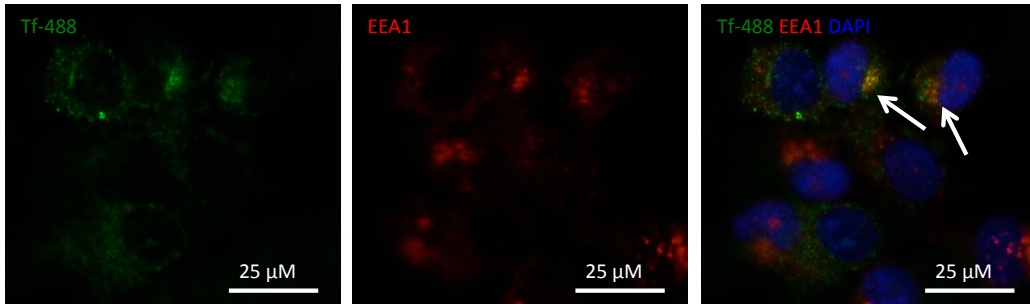
B   BIN1 depletion
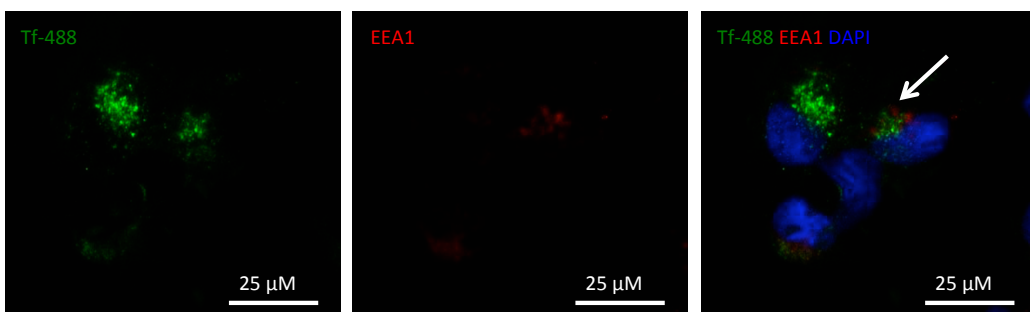
C   CD2AP depletion
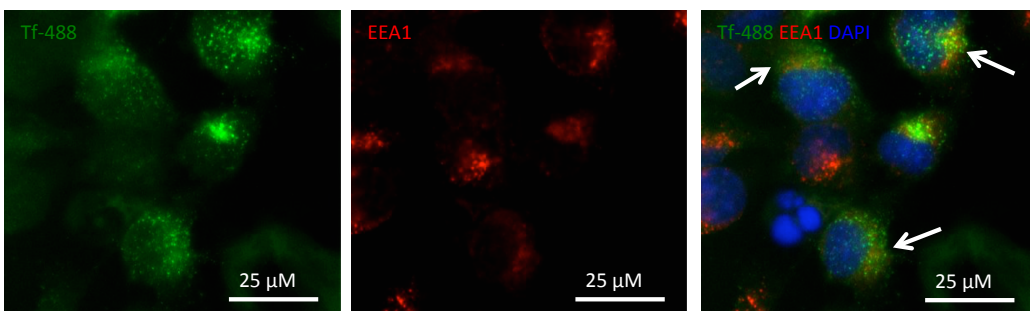
D   NT control

Figure 3.18. Tf-488 and EEA1 localisation after 5 minute exposure to Tf-488. A) BIN1 and CD2AP depleted cells. B) BIN1 depleted cells. C) CD2AP depleted cells. D) NT control cells. N=3. BIN1 and CD2AP depletion was confirmed via Western blotting. Left column shows Tf-488 only. Centre column shows EEA1 staining only. Right column shows Tf-488, EEA1 staining and DAPI. The right column shows some co-localisation between Tf-488 and EEA1 (indicated by arrows) suggesting Tf-488 is trafficked to the early endosome.

3.3.4.3 Tf-488 is localised to the lysosome after 30 minutes

All transfected cells were exposed to Tf-488 for 30 minutes. Cells were then probed with LAMP2, a lysosomal marker, and then stained with DAPI. Cells were imaged at 40x magnification and at optimal exposure for imaging Tf-488 after 30 minutes, which was consistent between biological replicates.

Consistent with previous observations shown in figure 3.17, BIN1 and CD2AP depleted cells and CD2AP only depleted cells appear to have greater levels of Tf-488 than the BIN1 only depleted cells and NT control (figure 3.19). Tf-488 appears punctate and in the perinuclear region in the BIN1 and CD2AP depleted cells, BIN1 only depleted cells and CD2AP only depleted cells. Punctate localisation of Tf-488 in the NT control is not clear, but remains in the perinuclear region.

All conditions show some level of co-localisation between Tf-488 and LAMP2, indicating Tf-488 is being trafficked to the lysosome (indicated by arrows in figure 3.19). Co-localisation between TF-488 and LAMP2 was quantified generating a PCC value for each transfected condition. A Levene's test revealed equal variances between the transfected conditions (p=0.105). A one-way ANOVA revealed significant differences between Tf-488/LAMP2 co-localisation between the transfected conditions (p=0.029). A Tukey HSD test revealed there was increased co-localisation in CD2AP depleted cells in comparison to BIN1 depleted cells (p=0.027). This could be explained by increased internalisation of Tf-488, resulting in an increased levels being trafficking to the lysosome. No other conditions showed any significant differences. Representative images are shown in figure 3.19.

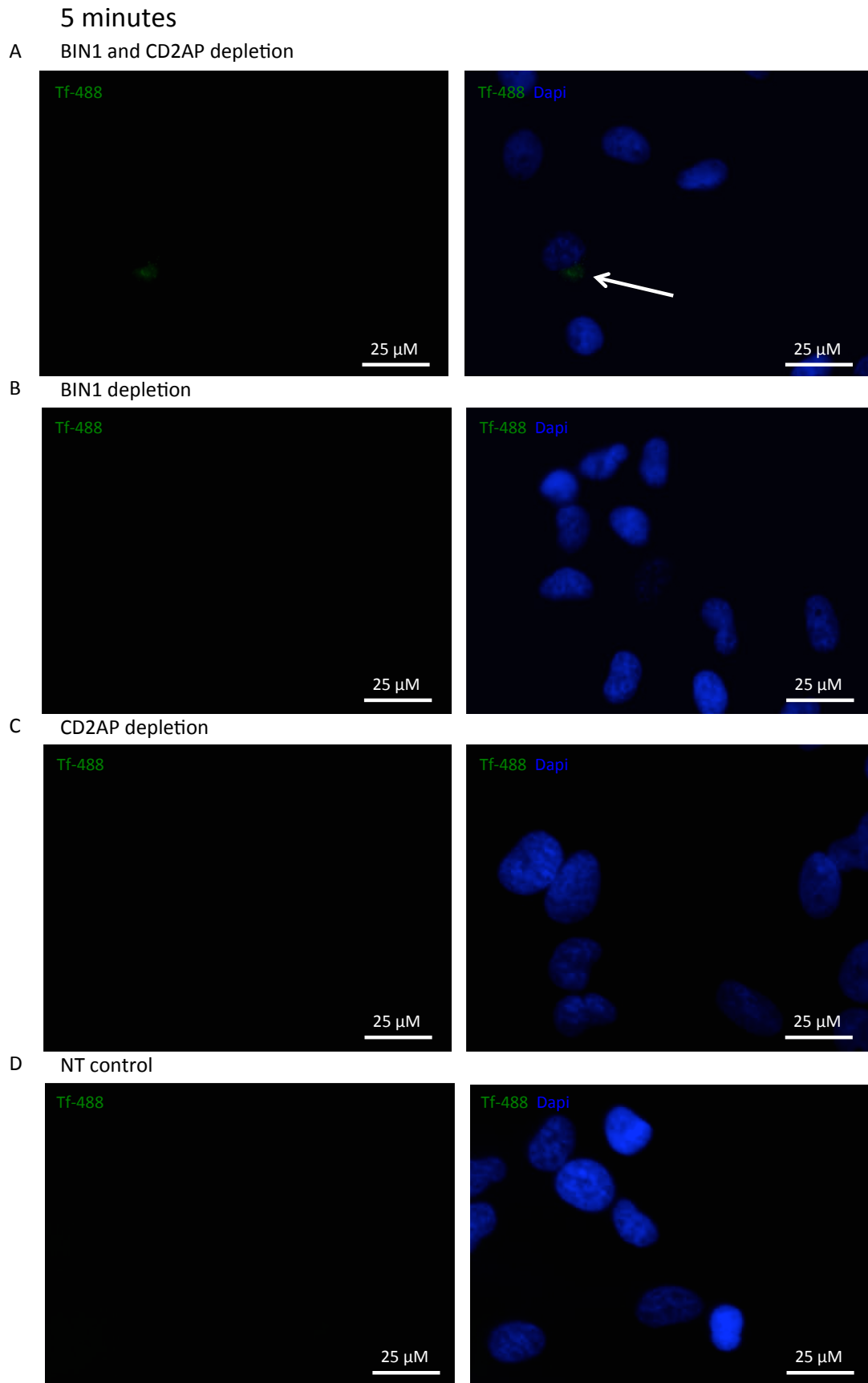## 30 minutes at optimal exposure
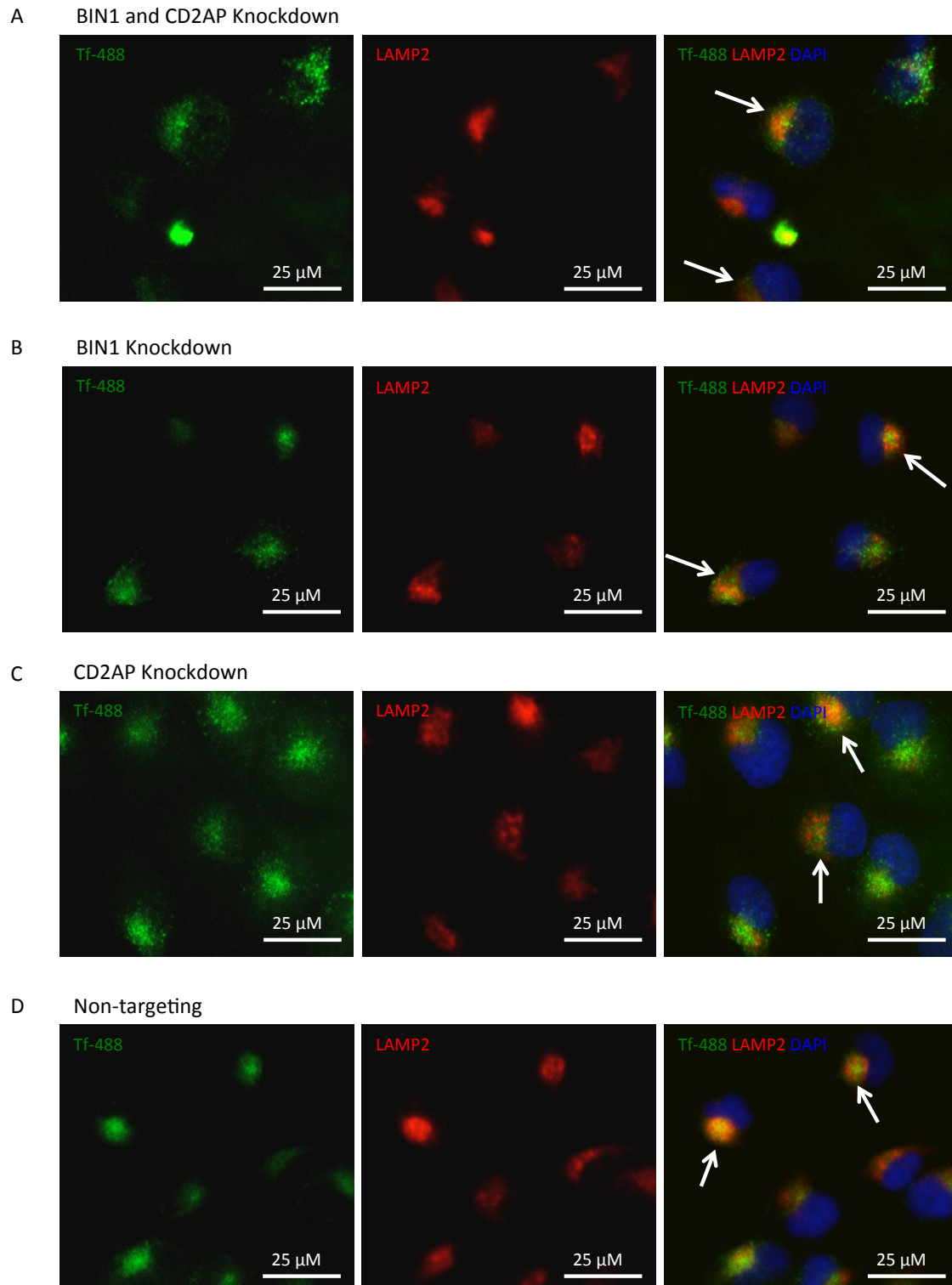
Figure 3.19. Tf-488 and LAMP2 localisation after 30 minute exposure to Tf-488. A) BIN1 and CD2AP depleted cells. B) BIN1 depleted cells. C) CD2AP depleted cells. D) NT control cells. N=3. BIN1 and CD2AP depletion was confirmed via Western blotting. Left column shows Tf-488 only. Centre column shows LAMP2 staining only. Right column shows Tf-488, LAMP2 staining and DAPI. The right column shows some co-localisation between Tf-488 and LAMP2 (indicated by arrows) suggesting Tf-488 is trafficked to the lysosome.

3.4 Discussion

BIN1 depletion resulted in an increase in β-CTF levels, suggesting a potential increase in β-secretase activity. CD2AP depletion resulted in a decrease in β-CTF and extracellular Aβ40 levels, and an increase in CME. When BIN1 and CD2AP are both depleted, APP and APP metabolite levels remain unchanged, however a decrease in BACE1 levels and an increase in CME was observed.

3.4.1 BIN1 depletion increases β-CTF levels

BIN1 depletion increased β-CTF levels, suggesting a potential increase in amyloidogenic processing. BIN1 depletion did not alter intracellular APP levels, this could suggest BIN1 depletion had no effect on APP generation and this is not the cause of the observed changes in β-CTF levels. However, as APP mRNA levels were not investigated the effect of BIN1 depletion on APP generation cannot be confirmed, as observations could also suggest an increase in APP production and an increase in APP clearance.

The increase in β-CTF levels in BIN1 depleted cells could indicate increased β-secretase cleavage. BIN1 depletion in mouse cortical neurons has been reported to increase β-cleavage and result in greater Aβ production and this coincided with increased cellular BACE1 levels (407). This increase in BACE1 levels was thought to be due to increased uptake and impaired cellular trafficking to the lysosomal pathway, resulting in enlarged BACE1 containing early endosomes and reduced BACE1 degradation. Furthermore, BIN1 depletion in primary cortical neurons increased intracellular Aβ in neuronal cell bodies, reduced BACE1 recycling back to the plasma membrane and increased BACE1 intracellular retention. BIN1 depletion also caused BACE1

accumulation in axonal early endosomes, the site where BACE1 and APP co-localise (259).

An apparent increase in BACE1 levels was observed here in BIN1 depleted cells, but this was not significant due to high variability. Although not significant, this data is tending towards being consistent with previous reports of increased BACE1 levels. The BACE1 levels measured were in the exponential region of the graph created by the protein standards provided, suggesting that lack of assay sensitivity is not the reason for the variability observed. An alternative explanation for the variation may be due to inconsistencies in BIN1 depletion. As BIN1 has been implicated in BACE1 recycling, to determine the effect of BIN1 depletion on BACE1 trafficking, further assays could investigate how depletion of BIN1 can affect cellular recycling in the H4 cells and the localisation of BACE1.

Extracellular Aβ40 levels remained unchanged in BIN1 depleted cells, suggesting no change in γ-secretase activity. No components of the γ-secretase complex were investigated in this study; therefore it cannot be determined whether levels of the γ-secretase complex have been altered. γ-secretase is located in multiple intracellular compartments (408). It is fairly well accepted that BACE1 is the rate-limiting factor in amyloidogenic processing, therefore one would expect an increase in β-CTF to result in an increase in Aβ40, should γ-secretase levels be unaltered. It in unknown why Aβ40 levels remain unchanged but this observation is consistent with another study, where PICALM, another AD associated gene with a critical function in endocytosis, was depleted in the H4 cell line (409). PICALM depletion resulted in an increase in β-CTF levels, but extracellular Aβ40 remained unchanged. Aβ40 levels observed in the H4 cell line were fairly low, so it could be possible

that the assay used to quantify Aβ40 was not sensitive enough to detect subtle variations.

Studies in other cell types have reported changes in Aβ levels following BIN1 depletion. An increase in extracellular Aβ40 and Aβ42 in BIN1 depleted mouse cortical neurons (407) and an increase in intracellular Aβ in BIN1 depleted neurons have been reported (259). As only extracellular Aβ40 was investigated, it could be that BIN1 depletion has an effect on Aβ secretion and a change in intracellular Aβ40 may have occurred. As BIN1 depletion has been reported to reduce BACE1 recycling to the plasma membrane (259), it could be possible that trafficking of other molecules to the plasma membrane is affected. BIN1 depletion could result in defects in Aβ40 trafficking or secretion, therefore a change in Aβ40 levels may not be reflected in the extracellular levels.

No significant changes in extracellular sAPPα, sAPPβ or the sAPPβ:sAPPα ratio were observed in BIN1 depleted cells. This observation contradicts the increase in β-CTF observed in this investigation, as sAPPβ is a product of β-secretase cleavage. Previous investigations have reported an increase in sAPPβ in BIN1 depleted neurons (407). As all extracellular metabolites were quantified in the culture medium, it could be that the dilution of these metabolites in the medium was so great that subtle changes could not be detected.

In this investigation, Tf-488 levels have been used as a proxy for CME as it has been previously been shown that transferrin is endocytosed via CME. It is assumed that increase in Tf-488 uptake would also be applicable to APP, as this is endocytosed through CME, although APP uptake was not directly investigated.

BIN1 depletion did not affect Tf-488 or Dextran-647 intracellular levels, indicating no change in CME or fluid phase uptake. This contradicts studies that reported an increase in Tf-488 uptake in BIN1 depleted rat neurons. This effect was hypothesised to be due to BIN1 sequestering dynamin at nascent vesicle necks, preventing efficient CME (253). The entire CLAP domain in BIN1 is critical for clathrin and AP2 interaction and CME and is only present in neuronal isoforms of BIN1 specifically expressed in the brain (233, 251). The presence of the neuronal isoform may differ in these cell types and could explain the differences in CME observed. Determining the BIN1 isoforms present in the H4s may help explain the effect of BIN1 depletion on CME. A reduction in BIN1 neuronal isoform 1 has been described in AD patients (255).

As no change in CME was detected in BIN1 depleted cells, this suggests that CME is not responsible for the increase in β-CTF levels. An increase in APP uptake via CME would explain the increase in β-CTF levels, as it would result in more APP coming into contact with BACE1 within the endosomes. It is possible that although Tf-488 and Dextran-647 levels are generally used to measure CME and fluid phase endocytosis, Tf-488 uptake may not be reflective of APP uptake and BIN1 may affect APP uptake via a different specific mechanism. For example, if BIN1 depletion affected the cell surface levels of LRP, the endocytic receptor of APP (410), this could affect APP uptake specifically but not overall CME levels.

BIN1 depletion may also affect additional intracellular pathways rather than CME. In neurons BIN1 functions in the intracellular trafficking of BACE1 and BIN1 depletion affects BACE1 recycling and degradation, but not cell surface levels or BACE1 endocytosis (259, 407). This is consistent with observations in this investigation as an apparent increase in cellular BACE1 levels was observed (although not significant) but no change in CME or fluid phase

uptake. In knockout BIN1 MEF cells, intracellular levels of transferrin were increased, but not due to increased internalisation (411) and BIN1 depleted HeLa cells showed impaired transferrin recycling rather than impaired uptake (249). Therefore, the changes in β-CTF and BACE1 levels observed in BIN1 depleted H4 cells, could be due to changes in other cellular trafficking functions and not CME.

BIN1 expression changes in AD remain unclear. Studies have reported a decrease in expression in total BIN1 expression and reduction in BIN1 neuronal isoform expression in AD brains (255, 256), which would be consistent with observations in this investigation as that BIN1 depletion may increase amyloidogenic processing of APP. However, there are multiple reports stating increases in BIN1 expression are associated with AD [further discussed in Chapter 1.7.5]. From the results of this investigation, a reduction in BIN1 expression would increase an individual's risk to AD. It is yet to be determined how AD associated variants within BIN1 affect expression levels, however rare deleterious coding mutations in BIN1 have been associated with AD in a Caribbean Hispanic population, suggesting loss of BIN1 function is a potential pathogenic mechanism (412).

3.4.2 CD2AP depletion decreases β-CTF and extracellular Aβ40 levels and increases CME

CD2AP depletion resulted in reduced β-CTF levels and reduced extracellular Aβ40 levels. CD2AP depletion did not alter intracellular APP levels, but as previously discussed, APP mRNA levels were not investigated so any effect on APP generation or clearance remains unknown.

Previous investigations in mouse neuroblastoma cells used epifluorescence microscopy to measure APP levels and found that CD2AP depletion delayed APP degradation. Depletion of CD2AP led to accumulation of APP in dendritic early endosomes, which are enlarged consistent with cargo accumulation. It was suggested that the organisation of APP into intraluminal vesicles (ILV) during multivesicular bodies (MVB) formation is impaired when CD2AP is depleted, preventing lysosomal degradation (259). Although APP levels remained unchanged when CD2AP was depleted in the H4 cell line, further investigation into APP degradation was not carried out.

CD2AP depletion resulted in a significant decrease in β-CTF levels, suggesting a decrease in β-secretase cleavage. BACE1 levels were reduced in CD2AP depleted cells, which would explain the apparent decrease in β-secretase activity observed. CD2AP depletion also resulted in a reduction in extracellular Aβ40. This is consistent with the reduction in β-CTF observed but could also suggests reduced γ-secretase activity. This is also consistent with previous findings which reported that CD2AP depletion in N2a mouse neuroblastoma cells decreased extracellular Aβ levels and Aβ40:42 ratio (280).

Extracellular sAPPα and sAPPβ levels were not affected in CD2AP depleted cells and there was no change in the sAPPβ:sAPPα ratio. As a decrease in β-CTF was observed, it would be expected to see a corresponding reduction in sAPPβ levels. Due to the dilution of the sAPP fragments in the culture medium, it is most likely the assay was not sensitive enough to detect changes in sAPPα or sAPPβ levels.

It has been previously hypothesised that CD2AP functions in ILV formation (259, 276). ILVs traffic cargo to the lysosome or it can be released into the extracellular space (221). If ILV formation is affected by CD2AP depletion, this

could impact on metabolite secretion via exosomes and explain why the sAPPβ levels to not reflect the β-CTF levels observed. To determine whether the extracellular metabolites reflect events within the cell, intracellular Aβ could be quantified. Discrepancies between intracellular and extracellular metabolites could be indicative of impaired secretion.

After 5 minutes of exposure to Tf-488, CD2AP depleted cells showed a significant 2.03 fold increase in Tf-488 internalisation in comparison to the NT siRNA control. As no change in fluid phase endocytosis was detected after 5 minutes, this suggests that CD2AP depletion specifically increases internalisation via CME, rather than non-selective uptake. Previous investigations have shown that increased internalisation of APP, increases amyloidogenic processing due to an increase in internalised APP being cleaved by β-secretase in the endosomes (392, 393). This is in contradiction to the observations in this investigation as a reduction in the products of amyloidogenic processing was observed when CD2AP is depleted, although this investigation only looked at global CME changes, not APP uptake specifically. At 15 minutes, a 1.72 fold increase in Tf-488 levels was observed in the CD2AP only depleted cells compared to the control, again suggesting increased internalisation.

After 30 minutes, CD2AP depleted cells showed a significant 1.76 fold increase in Tf-488 levels in comparison with the NT siRNA control. It has been previously reported that CD2AP depletion reduces APP trafficking to the lysosome and therefore impairs degradation of APP (259). Therefore, at this time point the increase in Tf-488 levels could reflect impaired recycling or degradation in addition to an increased internalisation. CD2AP depletion has been previously implicated in impaired APP degradation and reduced numbers of MVB (275, 277). APP sorting into MVB is critical for its degradation, so it has

been hypothesised that impaired MVB formation in CD2AP depleted cells may cause impaired lysosomal degradation of APP (259). CD2AP depletion could therefore affect multiple cellular trafficking pathways in addition to internalisation that could impact on APP processing.

There has been little investigation into CD2AP expression in AD and how common AD associated variants affect CD2AP expression remains unknown. Rare deleterious coding mutations have been identified in CD2AP to be associated with AD in Caucasian populations, suggesting loss of CD2AP function may be a mechanism which infers AD susceptibility (412). This would contradict some of the observations in CD2AP depleted cells in this investigation, which show a potential decrease in amyloidogenic processing of APP, but the underlying functional causes of the changes in APP processing observed remain unknown.

3.4.3 BIN1 and CD2AP depletion in combination do not have an effect on APP processing but reduce BACE1 levels and increase CME

BIN1 and CD2AP depletion in combination had minimal affect on APP processing. BIN1 and CD2AP depletion did not alter intracellular APP levels, however, as previously discussed mRNA levels were not investigated therefore it cannot be determined whether APP generation or degradation were affected.

BIN1 and CD2AP depletion did not have a significant affect on β-CTF levels when compared to the NT siRNA control, however β-CTF levels were significantly reduced when compared to BIN1 only depleted cells, which showed an increase in β-CTF. There was no significant difference between the

BIN1 and CD2AP depleted cells and the CD2AP only depleted cells, which showed a significant reduction in β-CTF when compared to the control levels.

A reduction in BACE1 levels was detected in BIN1 and CD2AP depleted cells compared to the NT siRNA control and BIN1 only depleted cells. There was no difference in BACE1 levels between BIN1 and CD2AP depleted in combination and CD2AP only depletion; it is therefore likely the loss of CD2AP is responsible for the reduction in BACE1. As β-CTF levels were not significantly reduced in the BIN1 and CD2AP depletion in combination cells, BIN1 depletion may rescue this phenotype via a mechanism independent of BACE1 levels.

When BIN1 and CD2AP are depleted in combination, extracellular Aβ40 levels were not significantly different to the NT siRNA control, which is contrary to the results from the CD2AP only depleted cells, in which extracellular Aβ40 levels were significantly reduced. BIN1 only depleted cells showed no change in Aβ40 levels. It could be possible that BIN1 depletion may rescue or counter the affect of CD2AP depletion on Aβ40 levels.

Extracellular sAPPα and sAPPβ were not significantly different to the NT siRNA control when BIN1 and CD2AP were depleted in combination. There was also no change observed in sAPPβ:sAPPα ratio. This was consistent with previous findings as no significant change in β-CTF levels were observed. As previously discussed, all extracellular metabolites were diluted in culture medium and the assays may not have been sensitive enough to detect a subtle changes in metabolites.

After 5 minutes of exposure to Tf-488, BIN1 and CD2AP depleted in combination resulted in a significant 2.1 fold increase in Tf-488 levels

compared to the NT siRNA control. At this early time point, this increase in Tf-488 levels is most likely due to an increase in internalisation via CME. At 15 minutes a 1.92 fold increase in Tf-488 levels is observed in the BIN1 and CD2AP depleted cells compared to the NT siRNA control, which again is probably due to increased internalisation. BIN1 and CD2AP depletion caused a significant 1.92 fold increase in Tf-488 after 30 minutes when compared to the NT siRNA control. However, at this time point the increase in Tf-488 could be due to increased internalisation and or potential defects in recycling or degradation. As BIN1 and CD2AP depletion in combination and CD2AP only depletion were not significantly different to each other at any time points, this suggests CD2AP depletion is likely to be responsible for the changes in CME.

As β-CTF and Aβ40 are significantly reduced in CD2AP only depleted cells, but not in BIN1 and CD2AP depleted in combination cells, this could suggest that BIN1 depletion rescues this phenotype. As CME is increased and there is no significant difference between the two depletion conditions, this suggests that if BIN1 depletion is rescuing β-CTF and Aβ40 levels, it is not via CME. In BIN1 and CD2AP depleted cells, BACE1 levels are significantly reduced compared to the BIN1 only depleted cells, indicating BIN1 depletion is not rescuing this phenotype by increasing BACE1 levels. BIN1 depletion could impact other cellular processes that either increase uptake of APP specifically or affect APP processing via mechanisms other than β-secretase activity.

3.4.4 Tf-488 is trafficked through the endosomal-lysosomal system

When cells were imaged for Tf-488 uptake after 30 minutes of exposure, as reflected in the FACS results, BIN1 and CD2AP depleted cells and CD2AP only depleted cells showed greater levels of Tf-488 located in the perinuclear area. BIN1 depleted cells show Tf-488 located in the perinuclear area but at lower

levels and NT siRNA control cells show Tf-488 diffused throughout the cytoplasm.

When investigating the localisation of transferrin following uptake, after five minutes of exposure there is some co-localisation of Tf-488 with EEA1, an early endosomal marker. Following uptake, cargo will be transported into EEA1 positive early endosomes, so these images are consistent with Tf-488 internalisation. No obvious change in Tf-488 co-localisation can be detected between BIN1 and CD2AP depletion, BIN1 only depletion, CD2AP only depletion and NT siRNA control.

After thirty minutes of Tf-488 exposure, some Tf-488 is co-localised with LAMP2, a lysosomal marker. In BIN1 and CD2AP depleted cells and CD2AP only depleted cells more Tf-488 was localised outside the lysosome when compared to the BIN1 only depleted cells and the NT control. This could be due to an increase in internalisation, as suggested by the FACS data, but at this point it would be difficult to conclude whether this is due to in overall increase in transferrin uptake, and therefore Tf-488 is located elsewhere in the endosomal-lysosomal system, or whether this is a result of impaired trafficking of cargo to the lysosome as previously reported in CD2AP depleted cells (259). The use of additional endosomal markers and investigations at different time points would help to determine the location of Tf-488 after 30 minutes and provide more information on cellular trafficking.

3.4.5 Concluding remarks

BIN1 and *CD2AP* are both significantly associated with LOAD. Variants at these loci are associated with an increased risk of AD, but how they increase risk for the disease remains unknown. BIN1 and CD2AP have both been implicated in

the same biological pathway of endocytosis, which is thought to be critical in the pathogenesis of AD. The cellular trafficking of APP is critical for it's processing, as its cellular location can determine which secretase is encountered. This investigation therefore investigated how loss of these two proetin impacted on APP processing and CME.

BIN1 depletion and CD2AP depletion appeared to have broadly opposing effects on β-CTF levels. BIN1 depletion increased β-CTF levels, which is potentially explained by an increase in BACE1 levels, whereas CD2AP depletion reduced β-CTF levels, which is potentially explained by the reduction in BACE1 levels observed. Depletion of BIN1 and CD2AP in combination showed minimal change in APP metabolites, but BACE1 levels remained significantly reduced in comparison to the NT siRNA control. This result was not different from the BACE1 levels observed in CD2AP only depleted cells, suggesting CD2AP depletion is responsible for the reduction in BACE1 levels. This also suggests that BIN1 depletion may counter the effects of CD2AP depletion on APP processing via a mechanism independent of BACE1.

BIN1 depletion did not significantly affect CME, suggesting increased internalisation is not responsible for the increase in β-CTF observed. BIN1 depletion likely results in increased β-CTF production via other cellular mechanisms, for example impaired BACE1 degradation or recycling. CD2AP depletion increased CME, which was unexpected, as increased internalisation has previously been implicated in increased β-CTF levels, contradicting observations in this investigation. This suggests that CD2AP depletion results in decreased β-CTF production independent of internalisation. CME in BIN1 and CD2AP depleted cells reflected that of CD2AP only depleted cells, suggesting loss of CD2AP is responsible for the increase in CME observed.

This data seems to suggest that CME is not responsible for the changes in APP

metabolites observed. It is likely that BIN1 and CD2AP have roles in more

complex cellular trafficking mechanisms, which may be affecting APP

processing.

4. Investigating *BIN1* allelic expression in relation to LOAD risk genotypes

4.1 Introduction

4.1.1 Allele Specific Expression

DNA variants associated with a complex trait can affect disease susceptibility via a number of mechanisms, one of which is by altering gene expression. Expression Quantitative Trait Loci (eQTL) are regions of the genome that contain DNA variants capable of influencing gene expression. Integrating GWAS and eQTL data can identify genes whose expression levels are associated with a complex trait (413). eQTLs identified in cerebellar tissue overlap with GWAS SNPs from a number of neurodegenerative diseases, including AD, indicating that changes in gene expression may be a mechanism influencing complex disease (414).

The underlying mechanism by which eQTLs are thought to act is in part mediated by DNA regulatory elements, such as those identified using DNase-seq. DNase I hypersensitivity sites (DHS) are regulatory elements used to map *cis*-regulatory elements throughout the genome (315). The distribution of non-coding genome wide significant associations in 207 diseases and 447 quantitative traits saw a 40% enrichment of GWAS SNPs in DHS, particularly in disease relevant cell types. Common variants associated with diseases are also enriched in recognition sequences of pathologically relevant transcription factors (371). Furthermore, when 11 complex traits were investigated, 79% of the heritability explained by imputed SNPs was due to their association with DHS, and a further enrichment of associated SNPs was observed in enhancer

associated DHS and in cell type specific DHS, providing further evidence that gene regulation is a likely mechanism by which genetic variants influence complex traits (415).

An Allele Specific Expression (ASE) Assay is a method used to measure *cis*-regulatory variation on gene expression. In the absence of heterozygous variants in the gDNA affecting *cis*-regulation, both alleles will be present in the mRNA in equal quantities. However, if an individual is heterozygous for a DNA variant that affects *cis*-regulation, each of the genomic alleles will be present in different quantities in the mRNA. An ASE assay uses a heterozygous transcribed SNP (present in the mRNA) to quantify the relative levels of the two alleles present in complementary DNA (cDNA), and therefore the mRNA. The ratio of the two alleles present in the mRNA can then be compared to the zygosity of the risk variant of interest to determine whether heterozygosity of this variant is associated with differential allele expression. Should the variant be associated with unequal allele expression, this could infer that this variant, or a tagged variant, is influencing the *cis*-regulation of this allele. Studying both alleles from the same cellular and tissue environment identifies genuine *cis*-acting effects by eliminating potential confounding effects from *trans*-acting elements (regulatory factors effecting both alleles equally). Such *trans*-acting elements that could affect mRNA levels in a particular cellular environment include tissue preparation, mRNA quality, environmental factors and *trans*-acting regulatory mechanisms (416-418). The principles of an ASE assay are illustrated in figure 4.1.
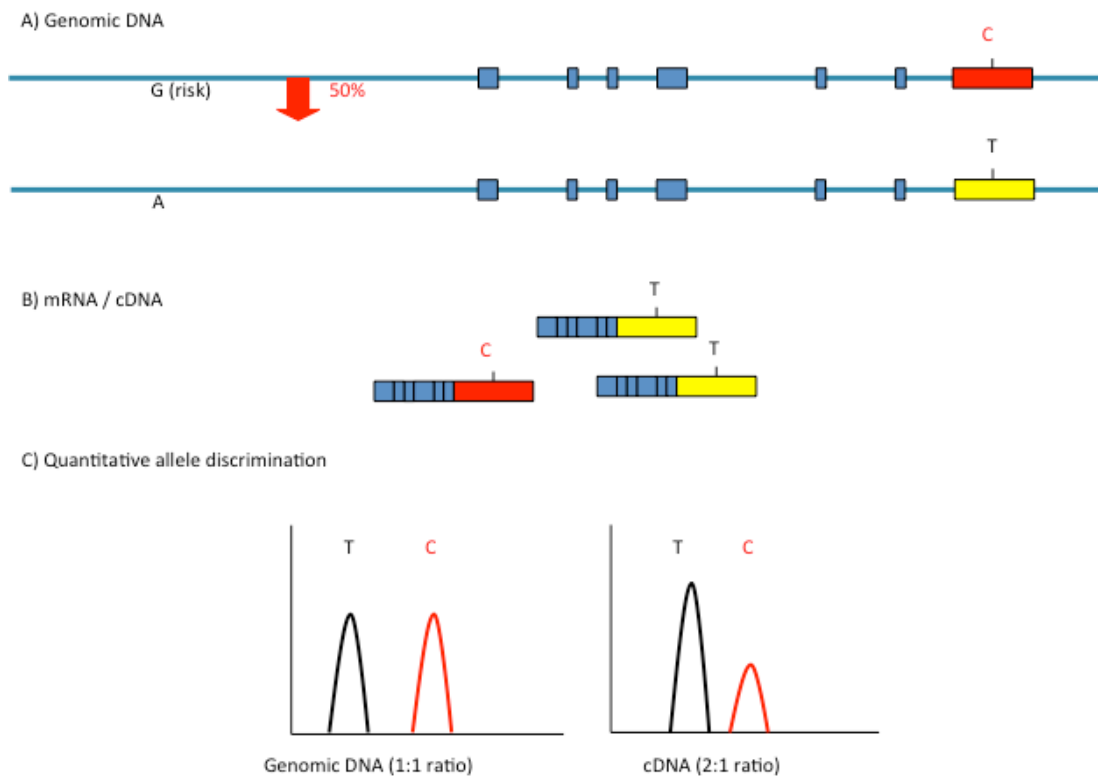
Figure 4.1. Principle of an Allele Specific Expression Assay. A) An individual is heterozygous for a *cis*-acting variant; each allele is present in the gDNA at a 1:1 ratio. In this example, the risk variant, G, reduces expression of this allele by 50%. This individual is also heterozygous for an exonic SNP (C/T). B) mRNA produced from each genomic copy can be distinguished by a heterozygous variants present in the coding DNA. mRNA is produced from the two alleles at a 2:1 ratio as half of the amount of mRNA is produced from the allele containing the risk variant. This allele expression ratio will also be present in cDNA. C) A quantitative technique that distinguishes between the two alleles can determine the allele ratios present in gDNA and cDNA (and therefore mRNA). The allele ratio observed in cDNA can be compared to the allele ratio observed in gDNA (assumed to be 1:1) and any deviation from the 1:1 ratio will indicate the presence if *cis*-acting variation. In this example, in cDNA the alleles will be present at a 2:1 ratio.

4.1.2 Investigating *cis*-regulation of *BIN1* in prefrontal cortex

Cis-regulation has been shown to occur in a high number of genes in the human brain (419). The prefrontal cortex plays an essential role in executive function, which involves a number of core cognitive components including

working memory (420). Neuronal and synapse loss and neurofibrillary tangles spread to all isocortical areas, including the prefrontal cortex, during the later stages of AD and isocortical regions become burdened with amyloid plaques during the early stages of AD pathology progression (17, 421).

Changes in *BIN1* expression have been described in the cortex tissue of AD patients (256, 261). Additionally, methylation profiles generated from prefrontal cortex tissue revealed suggestive evidence that differential methylation at the *BIN1* locus may affect LOAD susceptibility (422). As the prefrontal cortex is affected by AD neuropathology and regulatory and epigenetics changes at the *BIN1* locus have been identified to be associated with AD, it is an appropriate tissue to investigate *BIN1* allele expression and whether this is associated with a disease risk genotype.

4.1.3 SNPs of interest

There are a number of SNPs at the *BIN1* risk locus that have been identified to have an association with AD and have the potential to influence *cis*-regulatory factors.

4.1.3.1 Rs6733839: Most significant SNP

Rs6733839, located approximately 30 Kb upstream of *BIN1*, was identified as the most significant SNP at this locus in the largest meta-analysis of GWAS data for LOAD performed in 2013 ($p=6.9 \times 10^{-44}$). The minor allele with a T genotype has an odds ratio of 1.22 (140). Rs6733839 genotype has been associated with episodic memory and right inferior parietal atrophy, which is considered to be a predictive measure of AD prognosis (423-425).

4.1.3.2 Rs744373: First genome wide significant SNP identified at BIN1 locus

Rs744373 was the first SNP within the *BIN1* risk locus to reach genome-wide significance in 2010. Rs744373 was determined to have an odds ratio of 1.13 and is located just under 2 Kb away from rs6733839 (137). When investigating multilocus genotype patterns, rs744373 along with variants in *APOE* and *PICALM*, were associated with poorer episodic memory performance, even in individuals not suffering from AD (426). Furthermore, a healthy Chinese population homozygous for the rs744373 risk variant showed a poorer working memory performance, larger hippocampal volume and lower functional connectivity between the hippocampus and the dorsolateral prefrontal cortex (427). Rs744373 genotype is also correlated with tau levels, rate of cognitive decline and AD progression (428, 429).

4.1.3.3 Rs7584040: Conditional SNP

A conditional analysis utilising the genome wide complex trait analysis software was performed on stage 1 sample data from the IGAP consortium. This analysis investigated all variants within 500 Kb from rs6733839 and identified a secondary association signal at rs7584040, located within the first intron on *BIN1* (Majounie et al, in prep).

The r^2 values, representing the correlation coefficient between the AD associated variants are described in table 4.1.

Table 4.1. R^2 values between AD SNPs. Values were obtained from SNAP (SNP annotation and Proxy Search) Pairwise LD inquiry (accessed 21.8.2017) (430).

R^2 between AD associated SNPs	rs6733839	rs7584040
rs744373	0.58	0.131
rs7584040	0.056	-

4.1.3.4 Rs59335482: Functional Evidence

Rs59335482 is a three base pair indel located in a 6.7 Kb LD block containing rs6733839 and rs744373. Gene reporter assays showed that this variant is capable of influencing *BIN1* expression. The insertion allele was also found to be associated with increased *BIN1* mRNA expression in the frontal cortex of Alzheimer's disease brains (261).

4.1.4 Aim

This chapter aims to use an ASE assay to determine whether *BIN1* alleles are differentially expressed in prefrontal cortex tissue, which would indicate the presence of *cis*-regulatory factors. AD associated SNPs at the *BIN1* locus will be investigated to determine whether they explain any variation in *cis*-regulation observed.

4.2 Methods

4.2.1 Using publically available databases to functionally annotate variants of interest

There are a number of publically available databases that integrate genetic information with functional annotations. HaploReg v4.1 (accessed 18.5.17), hosted by the Broad institute, is a database that contains information on non-coding variants within haplotype blocks and compiles histone modification data (431). RegulomeDB v1.1 (accessed 18.5.17) gathers data from various sources to identify DNA features and describes regulatory elements found in the human genome (432). RegulomeDB Version 1.1 has a scoring system based on the functional confidence of a variant from 1 to 6. Lower scores indicate increasing evidence for the variant to be located in a functional region. Known eQTL variants are given a category 1 score, whereas variants lacking any functional annotation are labeled as category 6.

There are publically available databases that contain genotypic and gene expression data. These can be utilised to determine whether a specific variant is associated with a change in gene expression. Braineac (Brain eQTL Almanac) is a web-based resource that can be used to access the UK Brain Expression Consortium dataset. The UK Brain Expression Consortium dataset comprises of data from 134 human brains free from neurodegenerative disease and includes genotypes and gene expression data from multiple tissues across the human brain (5). This data can be stratified by genotype to determine whether a SNP acts as an eQTL (accessed 2.5.2017). GTEx (Genotype Tissue Expression Project) is resource that contains expression data from a number of tissues and researchers are able to use this information to investigate the relationship

between genetic variation and specific gene expression in human tissue (accessed 18.5.2017) (433).

Each of these databases was utilised to explore the potential functional and regulatory elements surrounding the SNPs of interest, particularly focusing on expression in brain tissue.

4.2.2 mRNA and gDNA samples

Studies were performed using gDNA and cDNA derived from post-mortem human brain tissue, which was obtained from the London Neurodegenerative Diseases Brain Bank by Dr Nicholas Bray and held at the MRC Centre for Neuropsychiatric Genomics and Genetics, Cardiff University, under a material transfer agreement. This collection is described in detail in (434). All subjects were free from psychiatric or neurological diagnosis at the time of death. In the present study, gDNA extracted from the brain of 116 unrelated adults was genotyped for the selected exonic SNP. cDNA synthesised from total RNA (by random priming) from the prefrontal cortex of 14 subjects identified as heterozygous for the *BIN1* exonic SNP was used for the ASE assay. Ethical approval for use of these samples to assess genetic effects on gene expression was provided to Dr Bray by The Joint South London and Maudsley and The Institute of Psychiatry NHS Research Ethics Committee (REF: PNM/12/13-102).

4.2.3 Heterozygous exonic SNP

A heterozygous exonic SNP is required to distinguish the mRNA copies produced from each chromosomal gene copy. *BIN1* is a highly conserved gene with very few exonic variants. SNPper sequence viewer (http://snpper.chip.org/bio/show-sequence/?TYPE=U&GENE=16170, accessed

17.11.15) identified SNPs transcribed in *BIN1*. Rs11554586 is located in the 5'UTR and is a SNP, a cytosine to a thymine. According to dbSNP, in the population sample "pilot_1_CEU_low_coverage_panel" attained from the 1000 Genomes Project, the C allele has a frequency of 0.817 and the T allele has a frequency of 0.183 (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?searchType=adhoc_search&type=rs&rs=rs11554586, accessed on 16.11.15) (435).

Applying the Hardy-Weinberg Equilibrium (HWE) principle predicts a frequency of heterozygotes of 0.299. Within a CEU population, similar to that of the sample population, approximately 30% of individuals will be heterozygote for rs11554586 and it is therefore likely heterozygotes will be present in the patient samples available for analysis.

4.2.4 SNaPshot Genotyping

SNaPshot genotyping uses the incorporation of a fluorescently tagged ddNTP to a primer located adjacent to the variant site. The fluorescence can be detected and will reflect the genotype of the variant. Primers were designed to amplify the region surrounding rs11554586, generating a product size of 165bp (table 4.2). This region was PCR amplified from the gDNA of 117 individuals and the negative controls (H_2O). The PCR reagents and thermocycler conditions are described in tables 4.3 and 4.4.

Table 4.2. Primer sequences used to amplify the 165 bp surrounding rs11554586

Forward Primer 5'-3'	CCTTTACTGCCCATCTCTGC
Reverse Primer 5'-3'	GTCAGTTGGCTCCGCTGT

Table 4.3. PCR reagents used to amplify rs11554586 region

Reagent	µL per Reaction
10X PCR Buffer (Qiagen)	1.2
Water	4.06
dNTPs (5mM)	0.96
Primers (5µM)	0.56 (x2)
DMSO (sigma)	0.6
HotStar Taq DNA polymerase (Qiagen)	0.06
DNA (5ng/µL)	4

Table 4.4. Thermocycling conditions used to amplify rs11554586 region

Step	Temperature (°C)	Time (Seconds)	Cycle
1	95	900	
2	95	40	
3	58	30	
4	72	40	Repeat steps 2-4 30 times
6	72	600	

These gDNA samples were genotyped for rs11554586 via SNaPshot genotyping [full method described in Chapter 2.4.6]. The SNaPshot genotyping reaction was performed using an extension primer targeting the antisense strand (5'-CAGGCCTCGCCCGGTGGCA-3').

14 rs11554586 heterozygous individuals had available prefrontal cortex cDNA for analysis. gDNA and cDNA from heterozygotes was amplified and genotyped for rs11554586 via the SNaPshot protocol plus water negative controls, following the protocol described above. SNaPshot genotyping

produces a fluorescent peak specific to the base incorporated on to the extension primer. The height of this peak is relative to the number of DNA copies that are incorporated with that specific base and therefore produces a quantitative measure that discriminates between the two alleles. cDNA genotyping was performed in duplicate and peak heights averaged for each sample. The peak heights of the G allele were divided by the peak heights for the A allele to create G:A allele ratio for each sample. gDNA of each sample was genotyped, the resulting G:A allele ratios were averaged between all gDNA samples. The average allele ratio from all gDNA samples was used to normalise the cDNA ratios to control for experimental inequalities in allele representation, as gDNA is assumed to have a 1:1 expression. Once normalised, any assay specific factors influencing the allelic representation will be corrected for and therefore the cDNA ratios will be representative of the expression of that specific allele. This can be analysed and any potential differential allele expression detected.

4.2.5 Genotyping of AD associated risk variants

Additional genotyping was performed on gDNA samples heterozygous for rs11554586 for which G:A ratios were available. A 163 bp region surrounding rs6733839 was amplified via PCR. The primer sequences, PCR reagents and thermocycling conditions are listed in tables 4.5, 4.6 and 4.7 respectively. The SNaPshot genotyping reaction was performed using the extension primer 5'-GTAAAAAGGGGAAAAGGGT-3'.

Table 4.5. Primers used to amplify region surrounding rs6733839

Forward Primer 5'-3'	TAGCCAGTGACTTACGCTGA
Reverse Primer 5'-3'	ACCTTCCCGTTCCATCCTGT

Table 4.6. PCR reaction used to amplify rs6733839

Reagent	µL per Reaction
10X PCR Buffer (Qiagen)	5
10mM dNTPs	1
10µM Forward Primer	0.5
10µM Reverse Primer	0.5
Water	13.75
HotStar Taq DNA Polymerase (Qiagen)	0.25
DNA (5 ng/µL)	4

Table 4.7. Thermocycler conditions to amplify rs6733839

Step	Temperature (°C)	Time (Seconds)	Cycles
1	95	900	
2	95	30	
3	60	30	
4	72	60	Repeat steps 2-4 30 times
5	72	500	

PCR primers and extension primers used in the SNaPshot genotyping of rs744373 are described in table 4.8. Primers used to genotype rs7584040 are described in table 4.8.

Table 4.8. PCR primers used to amplify region surrounding rs744373 and extension primers used for the genotyping of rs744373.

Forward PCR primer 5'-3'	CCTGGGAGACACTGGAGAAG
Reverse PCR primer 5'-3'	GCCTCCTGTCTTTCTGCAAG
Sense extension primer 5'-3'	ATCATGGGCAGCCTCTGAG
Antisense extension primer 5'-3'	GGGACAGGCAGGTCTGAGGC

PCR primers and extension primers used in the SNaPshot genotyping of rs7584040 are described in table 4.9.

Table 4.9. PCR primers used to amplify region surrounding rs7584040 and extension primers used for the genotyping of rs7584040.

Forward PCR primer 5'-3'	CAAACCTGGACTTGGCTGAG
Reverse PCR primer 5'-3'	CCTAAGGATGCAACCACGTG
Sense extension primer 5'-3'	AGGGAGGTATGGGGAAGCT
Antisense extension primer	GGGGGAGAGGGGGGCGCCTA

The PCR reagents and thermocycling conditions used to amplify rs744373 and rs7584040 are described in table 4.10 and 4.7 respectively.

Table 4.10. PCR reaction reagents used to amplify rs744373 and rs7584040

Reagent	µL per Reaction
10X PCR Buffer (Qiagen)	1.2
Water	4.66
dNTPs (2mM)	0.96
Forward Primer (5µM)	0.56
Reverse Primer (5µM)	0.56
HotStar Taq DNA polymerase (Qiagen)	0.06
DNA (5ng/µL)	4

Genotypes of rs59335482 were determined in samples heterozygous for rs11554586 via Sanger sequencing. DNA surrounding the rs59335482 was amplified using primers described in table 4.11. The PCR reaction and thermocycling conditions are described in table 4.12 and 4.13. The resulting PCR product is 152bp in length. The PCR product is ran on a 1% agarose gel, the appropriate sized band was excised and the DNA was extracted from the gel with the use of a QIAquick Gel Extraction Kit (Qiagen) and purified using a QIAquick PCR purification Kit. Purified DNA was sent to Genewiz® (Hope End, Takely, CM22 6TA) where samples were sequencing using their standard Sanger sequencing reaction conditions with the reverse sequencing primer (5'CAGGTGTGGTGGTTCGTA3').

Table 4.11. Primers used to amplify the region containing rs59335482

Forward sequencing primer 5'-3'	CCACCAAACCCAGCTAAT
Reverse Sequencing Primer 5'-3'	CAGGTGTGGTGGTTCGTA

Table 4.12. PCR reagents for rs59335482 sequencing

Reagents	Volume (µL)
Buffer	1.2
ddH$_2$O	4.66
dNTPs (2mM)	0.96
Primers (5pmol/µL)	0.56
Hot Star Taq	0.06
DNA	4

Table 4.13. Thermocycler conditions used to amplify the region surrounding rs59335482

Step	Temperature (°C)	Time (Seconds)	Cycle
1	95	600	
2	95	30	
3	58	60	
4	72	30	Repeat step 2-4 two times
5	95	30	
6	56	60	
7	72	30	Repeat step 5-7 two times
8	95	30	
9	54	60	
10	72	30	Repeat step 8-10 29 times
11	72	600	

4.3 Results

4.3.1 Functional databases indicate variants are located in potential regulatory regions in brain tissue

4.3.1.1 Rs6733839

The HaploReg v4.1 database provides evidence that rs6733839 may be within a regulatory region active in brain tissue. ChIP (Chromatin Imunnoprecipitation) -Seq data shows enhancer-associated modifications in this region in almost all brain tissue available in the database. Within the dorsolateral prefrontal cortex tissue, HaploReg v4.1 annotates this region to be associated with an active enhancer-flanking region and has H3K4me1 and H3K27ac, modifications indicative of enhancer activity.

RegulomeDB Version 1.1 annotates rs6733839 with a score of 5, suggesting that this variant lies within a TF binding or DNase peak. Using data from Pique-Regi et al, rs6733839 is annotated to lie in the proximity of MEF-2 and SPI1 binding motifs (436). ENCODE data suggests rs6733839 lies in a DNase Hypersensitive Site in the frontal cortex (437). ChromHMM regulatory regions, which utilises data from the Roadmap Epigenomics Project to annotate chromatin state (6, 438), describes the region to be a weak transcriptional chromatin state in inferior temporal lobe, anterior caudate, middle hippocampus, anugular gyrus, substantia nigra and dorsolateral prefrontal cortex tissue.

4.3.1.2 Rs744373

When querying rs744373 in the HaploReg v4.1 database, there is evidence of enhancer activity and promoter modifications in all adult brain tissues. Data from dorsolateral prefrontal cortex tissue shows evidence of enhancers and enhancer flanking regions within the proximity of rs744373 and H3K4me1 and H3K27ac modifications. Additionally, there is the presence of H3K9ac modifications, associated with active gene promoters.

RegulomeDB v1.1 annotates this variant with a score of 5 and describes an enhancer chromatin state in Angular Gyrus, Inferior temporal lobe, hippocampus middle, cingulate gyrus and dorsolateral prefrontal cortex (data from Roadmap Epigenomics Project).

4.3.1.3 Rs7584040

When querying rs7584040 in the HaploReg v4.1 database, there is evidence for this SNP being located near enhancers, promoters and a transcriptional start site in all adult brain tissues. Within dorsolateral prefrontal cortex tissue, there is evidence of rs7584040 being located near an active transcription start site, downstream promoter elements, H3K4me3 modifications associated with transcription start sites of active genes and H3K9ac modifications associated with active gene promoters. Additionally, there is evidence for H3K4me1 and H3K27ac modifications in this region associated with functional enhancer elements.

RegulomeDB v1.1 assigned rs7584040 a score of 4, indicative of minimal evidence of protein binding. ChIP-Seq data indicates USF1 TF binding in the SK-N-SH neuroblastoma cell line and annotations suggest this region is a DNase I hypersensitive site in the frontal cortex (data from ENCODE). ChromHMM regulatory regions describe the region as an active transcriptional

start site in angular gyrus, cingulate gyrus, substantia nigra, anterior caudate, hippocampus middle and dorsolateral prefrontal cortex tissues.

4.3.1.4 Rs59335482

Using Haploreg v4.1, rs59335482 is described in a genomic region marked with enhancer modifications, such as H3K27ac and H3K4me1, in brain tissues such as fetal thymus, hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe, angular gyrus and dorsolateral prefrontal cortex. Histone modifications that typically flank active promoters, such as H3K4me3 and H3K9ac, were identified in the proximity of rs59335482 in the hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe.

RegulomeDB v1.1 give rs59335482 a score of 5 and describes a weak transcriptional chromatin state in anterior caudate, hippocampus middle, inferior temporal lobe, angular gyrus and the substantia nigra. A quiescent chromatin region is described in the cingulate gyrus and the dorsolateral prefrontal cortex

4.3.1.5 BIN1 *expression in the brain*

When visualising gene expression using Braineac, *BIN1* is expressed throughout the brain but expression levels differ between brain regions. When this data is stratified by genotype of the SNPs of interest, no changes in gene expression in any brain tissue relative to genotype were observed; therefore within this database these SNPs are not functional eQTLs (figure 4.2). Rs59335482 is not within the Braineac database therefore this data could not be stratified by this genotype.
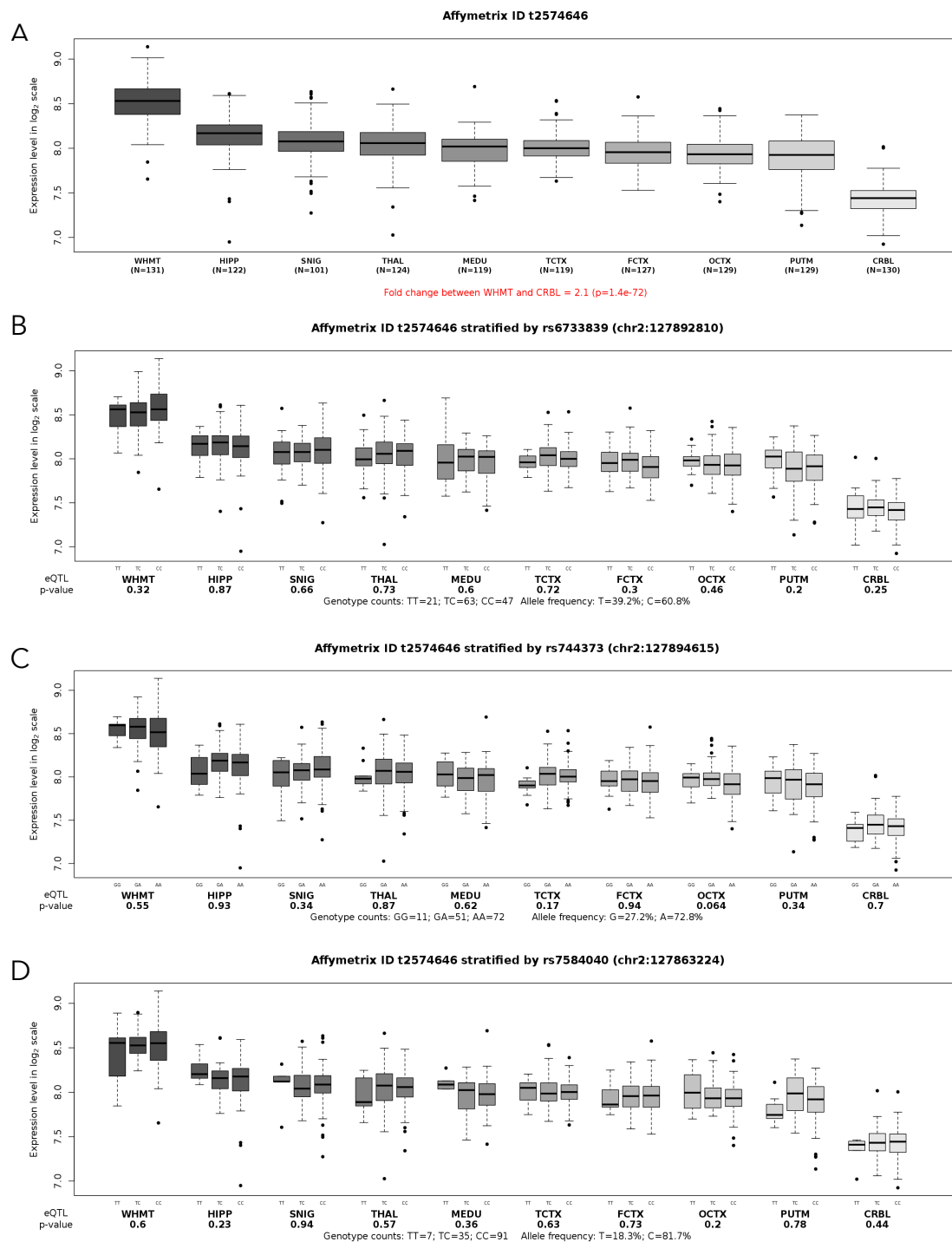
Figure 4.2. *BIN1* Expression throughout the brain. Data was obtained from www.braineac.org (accessed on 2.5.2017,(5)) which comprises of gene expression data in the following brain regions: Intralobular white matter, hippocampus, Substantia Nigra, Thalamus, Medulla, Temporal Cortex, Frontal Cortex, Occipital Cortex, Putamen and Cerebellar Cortex. A) *BIN1* transcript levels across 10 brain regions. B-D) *BIN1* transcript levels stratified by risk SNP genotype: rs6733839, rs744373 and rs7584040. In the Braineac database, none of these SNPs operate as an eQTL.

The GTEx database was used to determine whether the variants of interest were functional eQTLs for *BIN1* in frontal cortex tissue. Following Bonferroni correction for three tests (p=0.0167), the GTEx Analysis Release V6p (accessed 18.5.17) showed no evidence that rs6733839, rs744373 and rs7584040 were functional eQTLs in this tissue (p=0.082, p=0.046 and p=0.3 respectively) (figure 4.3).
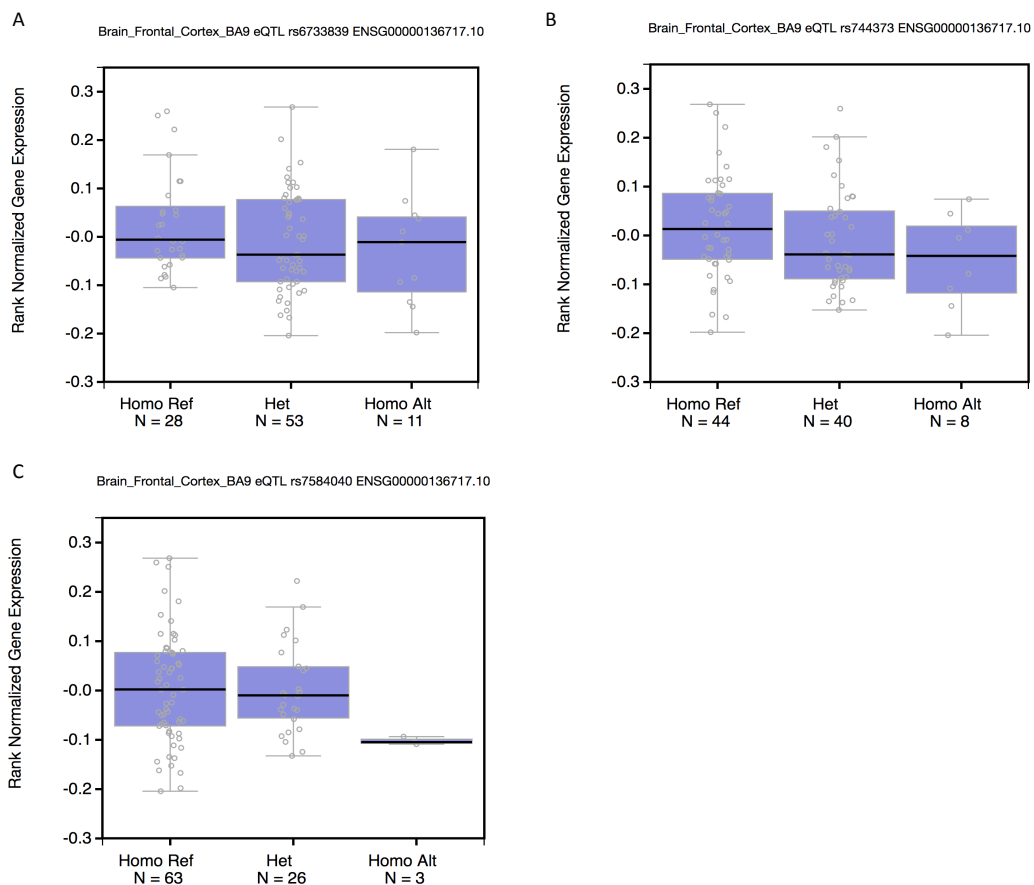
Figure 4.3. GTEx database was used to determine whether the variants of interest are functional eQTLs affecting *BIN1* expression in frontal cortex tissue. A) rs6733839 B) rs744373 C) rs7584040. These variants were determined not to be a functional eQTL for *BIN1* in prefrontal cortex tissue (p=0.082, p=0.046, p=0.3). Rs59335482 is not available in the GTEx database. (433).

When investigating the variants in additional tissues, there was suggestive evidence for rs6733839 as a functional eQTL in aorta tissue and whole blood (p=7.2 x 10 $^{-7}$ and p=1.5 x 10^{-5} respectively). There was suggestive evidence of rs744373 being an eQTL in aorta tissue (p=4.7 x 10^{-5}). rs7584040 was found to be a functional eQTL for *BIN1* in tissue from the tibial artery (p=2.2 x 10^{-15}), left ventricle of the heart (p=6.7 x 10^{-13}), esophagus muscularis (p=1.4 x 10^{-10}) and suggestive evidence in the aorta (p=7.3 x 10^{-7}), esophagus mucosa (p=6.5 x 10^{-7}), and pancreas (p=4.1 x 10^{-6}).

4.3.2 Heterozygosity of exonic SNP

PCR conditions were optimised using primers designed to amplify rs11554586 and confirmed via gel electrophoresis (figure 4.4).
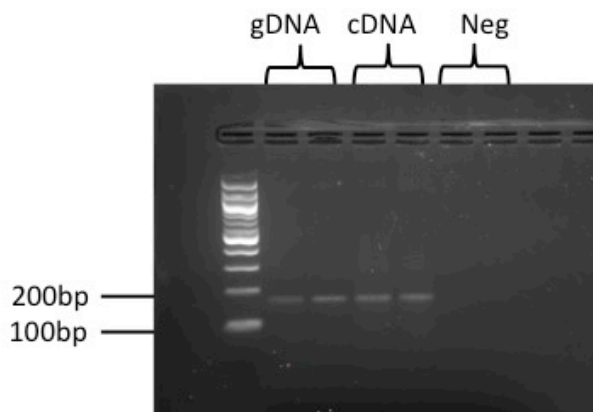
Figure 4.4. Amplification of 165 bp region surrounding rs11554586 in gDNA and cDNA. Water was used as a negative control. PCR products were electrophoresed on a 3% agarose gel.

These PCR conditions were used to amplify the rs11554586 region in gDNA and cDNA samples and subsequently used in SNaPshot genotyping of rs11554586. Genotypes were obtained and are summarised in table 4.14.

A X-squared test was performed to determine whether the observed genotype frequencies differ from the expected frequencies based on the observed allele frequencies. A X-Squared value of 0.6984 was obtained and a p value of

0.4033 indicated that the expected genotype frequencies were not significantly different from the observed genotype frequencies. The sample population is therefore in HWE and representative of the CEU population.

Table 4.14. Observed and Expected Genotypes. Expected genotype frequencies and expected genotype count was calculated based on the observed genotype count and allele frequencies.

Genotype	Observed Genotype Count	Expected Genotype frequency based on observed allele frequencies	Expected Genotype counts based on observed allele frequencies
GG homozygotes	87	0.7526	88
AA homozygotes	1	0.0176	2
GA heterozygotes	29	0.2299	27

4.3.3 Differential Expression of *BIN1* exonic variant

Following PCR amplification and SNaPshot genotyping of cDNA samples, the observed fluorescence peak height was recorded and normalised. Samples for which cDNA ratios showed a coefficient of variation >0.25 were excluded. Following successful SNaPshot genotyping and quality control checks, 14 samples had data. Figure 4.5 shows the G:A allele ratios of gDNA and cDNA. A Levene's test revealed that there were unequal variances between gDNA and cDNA ratios ($p \leq 0.001$) and therefore a two sampled t-test assuming unequal variances was performed to determine whether there was a significant

difference in allele ratios between gDNA and cDNA. This revealed that the G:A ratios of the cDNA is significantly different to that of gDNA (p=0.001) and thus deviates from a 1:1 allele expression ratio. This change from the 1:1 allele ratio indicates the presence of differential allelic expression and therefore indicates that *cis*-regulatory variation is operating on *BIN1* in adult prefrontal cortex.
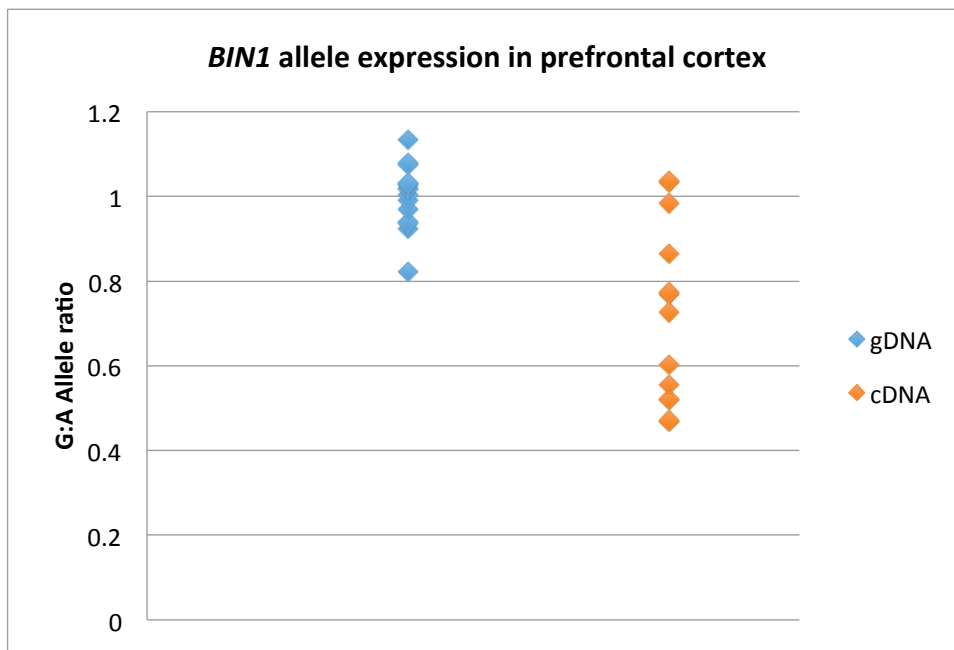
Figure 4.5. Allelic expression of *BIN1* at the expressed SNP rs11554586. 14 data points represent the experimental average G:A allele ratio for each individual. There is differential expression of the two alleles, as shown by the deviation from the 1:1 ratio in cDNA (p=0.001).

4.3.4 Suggestive evidence that rs7584040 may explain *cis*-regulation

The G:A allele ratios observed in the cDNA and the genotypes of rs6733839, rs744373, rs7584040 and rs59335482 are displayed in table 4.15.

Table 4.15. Allele ratios and zygosity of variants of interest

cDNA G:A Allele ratio	rs6733839	rs744373	rs7584040	rs59335482
0.773974147	CT	CT	CC	-/CCC
0.521357726	CT	CT	CC	-/CCC
1.036014658	CT	CT	CC	-/CCC
0.865119418	CT	CT	CC	-/CCC
1.032948655	CT	CT	CC	NA
0.471884365	CT	CT	CC	-/CCC
0.602978012	CT	CT	CT	-/CCC
0.519392134	CT	TT	CT	-/CCC
0.983747279	CT	CT	CC	-/CCC
0.769559491	TT	CT	CC	-/CCC
0.46898413	TT	CC	CT	CCC/CCC
1.100100889	CT	CT	CC	NA
0.726836047	CT	CT	CC	-/CCC
0.556687846	TT	CT	CC	-/CCC

A Levene's test for equality of variances was performed between homozygotes and heterozygotes for each variant and revealed equal variances. An independent samples t-test assuming equal variances was performed to compare allele ratios for homozygotes and heterozygotes for each variant. Should the variant account for the *cis*-regulation homozygotes should have an allele ratio of 1, which would be significantly different to the allele ratio of heterozygotes. These results are summarised in table 4.16. Due to amplification difficulties, some samples could not be sequenced to determine rs59335482 genotype. Of the samples successfully sequenced, only one

homozygote was identified and therefore statistical analysis could not be carried out for this variant.

Table 4.16. Summary of the statistical analysis performed to determine the differences in allele ratios between genotype. An independent samples t-test assuming equal variances was performed to compare homozygotes and heterozygotes for each SNP.

Variant	Levene's Test p value	Independent Samples T-Test p value
rs6733839	0.342	0.139
rs744373	0.079	0.093
rs7584040	0.070	0.062

4.4 Discussion

4.4.1 Functional databases indicate regulatory elements

Both Haploreg and Regulome DB annotated each variant as located within a region that had features indicative of being a regulatory element. In all brain tissues there was evidence of enhancer or promoter activity and histone modifications indicative of active regulatory elements were identified in the dorsolateral prefrontal cortex. This suggests that these variants are located in functionally active parts of the genome and therefore potentially could have an effect on *cis*-regulation.

Braineac and GTex did not associate any genotype with a change in *BIN1* expression levels, however these databases study total gene expression within these tissues. The approach to study allele specific expression is much more sensitive and this investigation studied a more specific brain region. GTex described rs7584040 as being an eQTL in tibial artery, left ventricle of the heart and esophagus muscularis tissue and described suggestive evidence for rs6733839 and rs744373 being eQTLs in non-disease relevant tissues. This demonstrates that rs7584040 is capable of influencing *cis*-regulation and rs6733839 and rs744373 potentially could.

4.4.2 Rs7584040 could potentially be a *cis*-acting variant

The allelic expression ratios observed indicate *BIN1* in prefrontal cortex tissue is influenced by *cis*-acting genetic variation as the allele ratios significantly differed from the 1:1 ratio (p=0.001). LOAD associated risk variants were

genotyped to determine whether changes in allele expression were correlated with zygosity of these variants.

Following statistical analysis, in this sample it appears that rs6733839 and rs744373 do not explain the *cis*-regulation observed as there was not a significant difference in allele ratios between heterozygotes and homozygotes. Despite not having sufficient data to determine whether rs59335482 genotype is statistically associated with a change in mRNA expression, it seems unlikely to be the *cis*-acting variant as the homozygote sample has an allele ratio of approximately 0.47, indicating the presence of differential *cis*-regulation not explained by rs59335482 genotype. However, as rs7584040 is approaching statistical significance (p=0.062) it is possible that this potentially affects *cis*-regulation in this tissue.

An analysis of differential methylation across the genome of a lymphoblastoid cell line found that rs7584040 lies within a 667 bp differentially methylated region which overlaps with a DNase I hypersensitivity site and a transcription factor binding region (439). Additionally, investigations have found that methylation levels of *BIN1* in the dorsolateral prefrontal cortex were significantly associated with increased LOAD risk and burden of AD pathology, suggesting disruption of DNA methylation may be associated with disease susceptibility (260, 422). Genetic variation has been shown to effect local methylation patterns in the cerebellum in a study of Bipolar Disorder (440). Therefore it could be possible for AD risk variants to elicit a functional impact by effecting allele specific methylation resulting in altered allele expression. Therefore this work provides suggestive evidence that rs7584040 may lie in a region that has an active role in gene regulation within the prefrontal cortex and could possibly effect *cis*-regulation by disrupting one of these mechanisms.

4.4.3 Sensitivity of assay was limited by sample size

The main limitation of this work was sample size. It has previously been calculated that 16 individuals heterozygous for the tagging SNP would be required to have 80% power to detect a *cis*-acting variant if this variant was found at a frequency of 0.05 in the population (418), but this sample size would not be sufficient to test the effect of such a variant. As the effect size of the functional variant is unknown, a power calculation cannot be used to determine the required sample size. This sample of 14 heterozygotes was sufficient to identify differential expression of the two alleles suggesting the presence of *cis*-acting variation. However, this sample size was not large enough to detect whether any of the AD associated risk SNPs were associated with this change in expression. In order to obtain interpretable data, a larger sample size should be used.

4.4.4 Future work

It is possible that further sequencing of the loci of interest could reveal additional variants that could affect *cis*-regulation. Genome wide sequencing of individuals that have an allele ratio significantly different to 1 could detect heterozygous variants that could be influencing *cis*-regulation. Although these variants may not currently be associated with AD, they could be tagged by a known SNP and contribute to the explanation of the *BIN1* index or conditional association signal. One issue with this approach again is sample size as it may be difficult to find a variant that overcomes a multiple testing penalty with such a limited available sample size.

To determine the directional change of expression, for example if the risk SNP increases allele expression, the variant would have to have a D'>0.8 with the

exonic SNP used to tag each allele. A pairwise linkage disequilibrium analysis using the SNAP version 2.2 online tool hosted by the Broad Institute, revealed a r^2 of 0.002 between rs11554586 and rs7584040, indicating rs11554586 would not be an appropriate tagging SNP (430). An alternative approach would be to perform long range PCR to determine which variants are located on the same allele in heterozygotes or statistical phasing, which would require a larger sample.

It is possible that the differential allele expression observed may be due to haplotype variation. Long-range allele-specific PCR could be used to determine the genotypes at multiple locations on each allele to distinguish haplotypes at this locus. Alternatively, if the D' is fairly high between variants of interest, haplotypes could be estimated via statistical phasing. Investigations could then be implemented to determine whether individuals who are heterozygous for a specific haplotype show differential allele expression. This approach would again require a greater sample size in order to determine significant changes. This approach has successfully identified risk haplotypes in complex diseases such as schizophrenia (441).

4.4.5 Concluding Remarks

Differential allele expression of *BIN1* was detected in prefrontal cortex tissue, indicating the presence of *cis*-acting regulatory variation. AD associated variants rs6733839, rs744373, rs7584040 and rs59335482 are located in genomic loci that have been annotated with evidence of regulatory DNA elements, suggesting they may elicit their effect by disrupting regulatory function. In prefrontal cortex tissue, none of these variants explain the differential allele expression observed, however rs7584040 is approaching statistical significance. As rs7584040 appears to be located in a regulatory

active genomic region, with evidence suggesting activity within the prefrontal

cortex, rs7584040, or a variant tagged by this SNP, could be responsible for

the change in allele expression. To validate this potential association, further

samples will need to be investigated for their allele expression levels and

rs7584040 genotype. Should this prove to be a significant association, studies

could go on to look at the methylation of this region or its interaction with

binding proteins in an attempt to understand the mechanism behind the

differential expression observed.

5. Characterising the Regulatory Capacity of the *BIN1* Risk Locus

5.1 Introduction

The majority of GWAS hits are located in non-coding regions of the genome. One of the ways DNA variants associated with a complex trait can affect disease susceptibility is via altering gene expression. Expression Quantitative Trait Loci (eQTLs) are regions of the genome that contain DNA variants capable of influencing gene expression. There are a number of different mechanisms by which eQTLs can affect gene regulation, specifically by altering mature mRNA levels. eQTL SNPs can affect epigenetic modification and transcription initiation by disrupting regulatory processes such as histone modification, transcription factor (TF) binding, enhancer activity (by altering chromatin conformation) and methylation (3). [Gene regulation and the affect of DNA variants are further described in Chapter 1.9].

5.1.1 Histone Modifications

DNA is bound by histones to form nucleosomes, which are the building blocks of eukaryotic chromatin. The binding of histones and other regulatory proteins to DNA results in a dynamic chromatin structure. Histones have distinct chemical modifications to the histone tail that can be read by specific proteins which can cause downstream events, such as transcriptional activation or gene repression (442). A nucleosome core consists of four histone proteins. Histone 3 (H3) receives the most extensive modifications and the biological significance of these has been most well characterised. H3 lysine 4 monomethylation

(H3K4me1) is associated with enhancer regions (334) and H3 lysine 27 acetylation (H3K27ac) is associated with increased activation of both enhancer and promoter regions (340). Histone modifications can exert their effect via influencing chromatin structure or by regulating the binding of effector molecules (331). eQTLs have been found to affect histone modifications. Furthermore, specific TFs have been identified which lead to histone modifications and variants located within TF binding sites have been shown to be correlated with changes in local histone modifications (443).

5.1.2 Transcription Factors

TF are key components in transcriptional regulation. Eukaryotic TFs are comprised of a DNA-binding domain and a transcription regulatory domain (444). TFs regulate the basal transcriptional apparatus by binding to specific gene promoters and inducing or repressing gene transcription by recruiting co-activators of the transcriptional machinery or aid in the remodeling of the chromatin structure (445).

Variants within the TF binding motif can directly affect TF binding, which could disrupt assembly of the transcription machinery or the recruitment of necessary co-factors. There are a number of examples where risk variants have created or destroyed a TF binding sites and have resulted in disease. For example, a risk variant associated with neuroblastoma formed an additional TF binding site resulting in overexpression of the *LMO1* oncogene (446). Furthermore, risk variants associated with colorectal cancer conferred disease risk by disrupting a TF binding site, altering the recruitment and binding affinity of TFs required for the transcriptional activation of the *FAS* gene, an inducer of the apoptotic signaling pathway (447).

In addition to changes in the binding motif, there are numerous other factors that influence TF binding, such as chromatin accessibility (448), co-factors that influence binding specificity (449), TF dimer interactions (450) and DNA sequences flanking the core TF binding site (451). The sequences flanking TF binding sites can have a pronounced affect on TF binding and this has been reflected in differential binding and gene expression observed *in vitro* (452). DNA conformation can contribute to differential binding to TF to various DNA sequences and it is thought that variation in the flanking region of a TF binding site could mediate an effect via DNA conformation (453).

5.1.3 Genetic Enhancers

Enhancers are DNA elements that interact with *trans*-acting factors to enhance transcription. Enhancers contain sequence motifs, which bind TFs (307). Variants can affect enhancer function by altering the binding affinity of regulatory proteins to the region or disrupt chromosomal conformations. Enhancers are thought to act via a looping mechanism where the enhancer and promoter physically interact allowing the exchange of transcriptional machinery and the stimulation of transcription (308, 454). This enhancer-promoter interaction is dependent on the 3 dimensional shape of the DNA, which could be affected by sequence variants. For example, a variant was identified that destabilised the enhancer-promoter loop with the *OCA2* gene, resulting in its downregulation (455).

eQTLs are also capable of influencing transcriptional elongation, co-transcriptional splicing, mRNA processing and modification and post transcriptional processing, such as mRNA degradation, polyadenylation and miRNA targeting (3). The ways by which eQTLs can influence gene expression is summarised in figure 5.1.
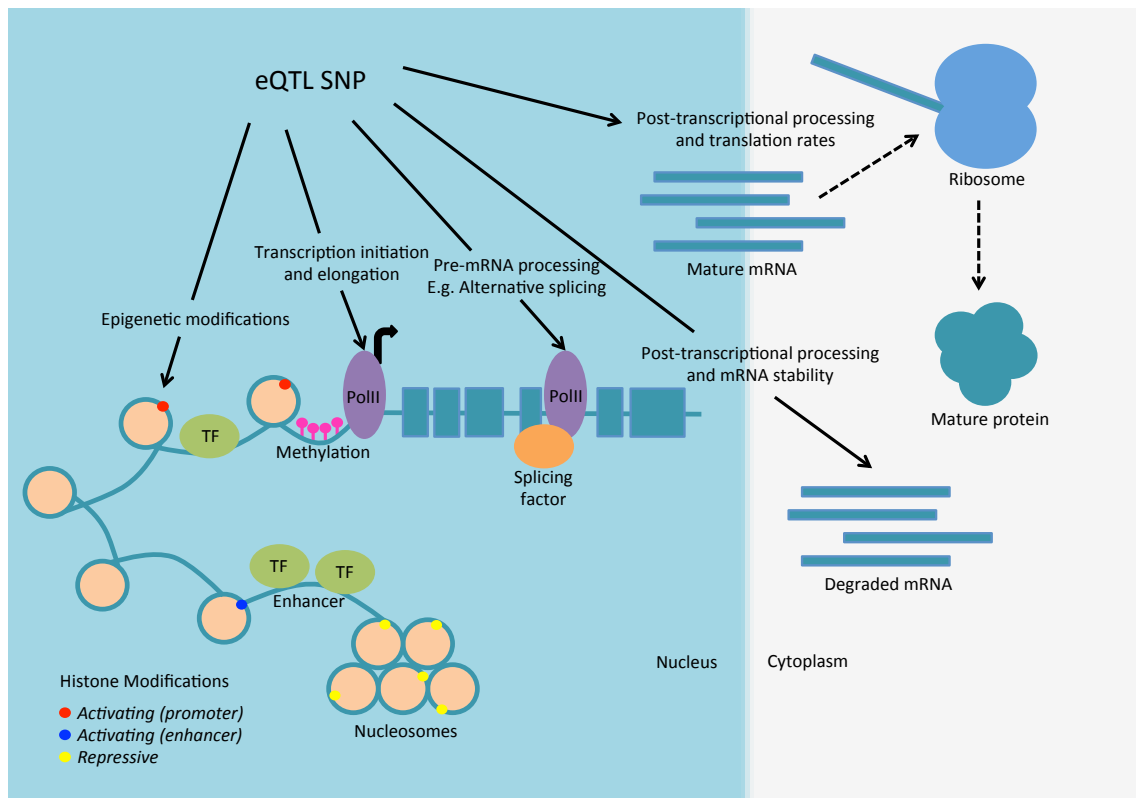
Figure 5.1. Regulatory mechanisms that can be affected by eQTL SNPs that can affect gene expression. eQTLs can affect epigenetic modifications and transcription initiation processes such as TF binding, histone modifications, enhancer activity and methylation. eQTLs can also affect transcriptional and co-transcriptional processes, such as splicing, mRNA processing and modification. eQTLs can also affect post-transcriptional mRNA processing, such as mRNA degradation, polyadenylation and targeting by miRNAs. Image adapted from (3).

5.1.4 Non-coding GWAS hits can identify variants that influence gene regulation

In vitro assays have been used to successfully investigate non-coding GWAS hits and have identified functional molecular mechanisms influenced by risk SNPs that affect gene regulation. For example, rs11603334 is in complete linkage disequilibrium (LD) with rs1552224, a SNP at the *ARAP1* locus significantly associated with type-2 diabetes (456). Gene reporter assays were used to show that the risk rs11603334 allele showed a two-fold increase in transcriptional activity in rat insulinoma b-cell lines. Further investigations found

that the rs1160334 risk allele disrupted PAX6 and PAX4 binding to the locus, both of which are transcriptional regulators (457). These data suggest rs11603334 is a functional variant regulating *ARAP1* expression and that changes in *ARAP1* expression may be the molecular consequence of the variants at this locus associated with Type 2 diabetes.

Variants located far from the gene can have an affect of gene regulation. A SNP located approximately 1.8 Kb upstream from the start codon of *ADRB2* which was associated with obesity proved to be functional by affecting the binding affinity of nuclear factors, resulting in decreased *ARDB2* expression (458). Furthermore, a pathway-based approach showed that phenotype-associated variants could be located up to 2 Mbs from the affected gene (459).

As the *BIN1* index SNP, plus LD partners ($r^2 > 0.8$), associated with LOAD is located approximately 30 Kb upstream of *BIN1,* it could be possible, as discussed in chapter 4, that this region could function as a regulatory DNA element genetic enhancer and the risk mechanism may have its effect by disrupting regulatory function.

5.1.1 Aims

This chapter aims to identify AD risk variants located in genomic sites of regulatory potential by utilising publically available databases containing ChIP-Seq and DNase-Seq data to investigate histone modification, local chromatin structure and protein binding. The regulatory activity of these elements will then be determined by implementing a gene reporter assay and an electrophoretic mobility shift assay (EMSA), both of which will examine the impact of rs6733839.

This investigation uses a number of disease relevant cell lines. The H4 cell line originated from a human neuroglioma and has been used in many AD cell biology investigations (409, 460-463). BIN1 depletion was investigated within this cell line and its affect on APP processing and endocytosis is discussed in chapter 3. Furthermore, chapter 4 describes regulatory features seen at the *BIN1* locus within brain tissues.

Monocytes are a type of white blood cell that is critical in innate immunity. Monocytes function in the immune response, in phagocytosis and can differentiate into macrophages and dendritic cells at the site of inflammation (464-466). Macrophages function in phagocytosis and in the innate and adaptive immune response by recruiting other functional immune cells. As the genetics of LOAD has implicated the immune response in AD pathology (180), such cell types would make an appropriate model. The human acute monocytic leukemia cell line, THP-1, resembles primary monocytes (375) and can be easily differentiated into a macrophage phenotype upon exposure to phorbol 12-myristate 13-acetate (467).

Microglia are the central nervous system's resident macrophage and a key component in neuroinflammation. Microglia survey the brain and are actively involved in maintaining the brains microenvironment by rapidly responding to pathogens or damage (468, 469). Microglia activation is a critical event in the neuroinflammatory response and microglia mediated chronic neuroinflammation has been linked to a number of neurological diseases, including AD (470). Due to this functional link to AD, two microglia cell lines were used to investigate *BIN1* regulation. This investigation used both the immortalised human microglia SV40 cell line and the immortalised murine

microglial BV2 cell line, which has been described as a valid substitute for

primary microglia (373, 376).

5.2 Methods

5.2.1 Interrogating databases to investigate protein binding, histone modifications and chromatin structure at the *BIN1* risk locus

HaploReg v4.1 annotates regulatory regions of the genome (431). The database uses data from The NIH Roadmap Epigenomics Mapping Consortium to annotate variants in a haplotype block with specific chromatin features (471). RegulomeDB v1.1 (accessed 18.5.17) compiles data from various sources to identify DNA features and describes regulatory elements found in the human genome (432). These tools were used to investigate the genomic features surrounding rs6733839 in blood and immune cells (http://archive.broadinstitute.org/mammals/haploreg/haploreg.php, accessed 2.8.2017).

The NIH Roadmap Epigenomics Mapping Consortium is a public resource of human epigenomic data (6). The resource contains data generated from next generation sequencing technologies that map DNA methylation, histone modification and chromatin accessibility across the genome in numerous tissues and cell types. In this investigation, this database was utilised to investigate H3K4me1 and H3K27ac modifications at the *BIN1* risk locus in brain tissue and monocyte cells and chromatin accessibility at this region within the monocyte cell type (http://www.roadmapepigenomics.org/ accessed 7.7.2017). Pott *et al* performed ChIP-Seq for the TF SPI1 on the THP-1 monocyte cell line (4). This data were accessed using Cistrome Dataset Browser, a comprehensive database of multi-resourced ChIP-Seq and DNase-Seq data, to determine SPI1 binding at the *BIN1* locus (http://cistrome.org/db/#/ accessed 7.7.2017)

5.2.2 Characterisation of Enhancer Activity using a Gene Reporter Assay

To investigate how the *BIN1* genomic region surrounding rs6733839 influences gene expression, this region was cloned upstream of the minimal promoter to act as a response element in the pGL4.23 plasmid, which encodes the luciferase reporter gene *luc2* (*Photinus pyralis*). Variants of the *BIN1* locus investigated differed in size, orientation and rs6733839 genotype in order to determine their effects of gene expression. Bioluminescence was quantified from the cells and acted as a proxy for *luc2* expression. The Dual-Luciferase® Reporter Assay System (Promega) was used in this gene reporter investigation.

5.2.2.1 Generating the plasmids

The Cistrome Dataset Browser was used to access ChIP-Seq data, which was used to investigate TF binding at the *BIN1* risk locus and determine potential enhancer regions for cloning. SPI1 binding regions in THP-1 cells showed TF binding within the region surrounding rs6733839. PCR primers were designed to amplify the genomic region surrounding rs6733839. Primer pair one was designed to amplify a region of 693bp, which spanned the entire SPI1 binding region identified by ChIP-Seq data. Primer pair two was designed to amplify a region of the genome where there was a dip in histone modifications identified using NIH Roadmap Epigenomics Mapping Consortium data, indicative of 5' end sequencing bias. Primer pair two amplified a region of 235bp surrounding rs6733839 to investigate a more localised effect of the variant. The PCR products were to be cloned into the KpnI cloning site in pGL4.23, therefore the KpnI recognition sequence was incorporated into the primers along with

additional bases required for restriction enzyme digests. The primer sequences are described in table 5.1.

Table 5.1. PCR primers used to amplify enhancer regions at *BIN1* risk locus

Primer Set 1	Forward Primer 5'-3'	CGATGGTACCAGAAACTGAGGCCAACTCCA
	Reverse Primer 5'-3'	CGATGGTACCACCCCTGTTTCCTCCTCTGT
Primer Set 2	Forward Primer 5'-3'	CGATGGTACCCCTGAGCCCCACTAAGATGA
	Reverse Primer 5'-3'	CGATGGTACCAGCATCGAGACTTCCCCTTC

DNA was amplified from sequenced CEU gDNA samples used in The International HapMap Project (472). The samples used were homozygous for either the non-risk or risk rs6733839 allele and shared common genotypes at all other SNPs amplified in these products. The PCR reagents and thermocycling conditions are described in table 5.2 and 5.3. PCR product size was confirmed via gel electrophoresis and purified using the QIAquick PCR purification kit (Qiagen).

Table 5.2. PCR reagents used in high fidelity PCR used to amplify *BIN1* enhancer regions

Reagent	Volume per Reaction (µL)
Water	32.5
5X Herculase II reaction Buffer	10
dNTPs (25mM)	0.5
DNA	4
10µM Forward Primer	1.25
10µM Reverse Primer	1.25
Herculase II fusion DNA polymerase	0.5

Table 5.3. Thermocycling conditions used to amplify *BIN1* enhancer regions

Step	Temperature (°C)	Time (Seconds)	Cycle
1	95	120	
2	95	20	
3	58	20	
4	72	30	Repeat Step 2-4 30 times
5	72	180	

Following amplification, the PCR product and pGL4.23 were digested separately by KpnI. The restriction enzyme digest reagents are described in table 5.4. The digestion reaction was incubated at 37°C for 1 hour. Once the pGL4.23 digestion had been incubated for 1 hour, 2.5 units of Shrimp Alkaline Phosphatase (NEB) was added to prevent religation of the plasmid, and this was incubated for a further 30 minutes at 37°C. Once all incubation steps were completed, the digestion products were purified using the QIAquick PCR purification kit (Qiagen) and the concentration of DNA determined using a NanoDrop™ 8000 Spectrophotometer (Thermo Scientific).

Table 5.4. Reagents for KpnI restriction enzyme digestion of PCR product and plasmid

Restriction Digest Reagent	Volume per Reaction (µL)
1 µg DNA	-
NEB Buffer 1.1	5
KpnI (NEB)	1
Total Volume	50

Each digested PCR insert was ligated into the digested pGL4.23 plasmid. The ligation reaction was carried out at a 12:1 insert to vector ratio. 100 ng of vector was combined with the appropriate amount of insert for a 12:1 ratio. A no-insert ligation negative control was performed. The reaction volume was then adjusted to 10 µL with the addition of water. 10 µL of 2X T7 Ligation Buffer and 1 µL of T7 ligase (NEB) was added and thoroughly mixed. This reaction was incubated at RT for 30 minutes.

The ligation products were transformed into Subcloning Efficiency™ DH5α™ Competent Cells (Thermo Fisher Scientific). 5 ng of ligation product was added to 50 µL DH5α™ Competent Cells and incubated on ice for 30 minutes. A Puc19 positive transformation control was performed. The cells underwent heat shock by being incubated at 42°C for 20 seconds then incubated on ice for 2 minutes. 950 µL of pre-warmed LB broth was added to the cells and then incubated in a shaking incubator at 225 rpm, at 37°C for 1 hour. 20 µL of each transformation was spread onto a LB plate containing 100 µg/mL ampicillin. The plates were incubated ON at 37°C.

A colony PCR was performed on clones present following incubation to determine the presence of the desired inserts. Primer set 1 and 2 described in table 5.1 were used. PCR reaction reagents and thermocycling conditions are described in table 5.5 and 5.6. Bacterial colonies were picked and tip placed in

PCR reagents. The tip was then placed into LB media containing ampicillin (100 µg/mL).

Table 5.5. Colony PCR reagents

Reagent	Volume per Reaction (µL)
PCR Buffer (10X) (Qiagen)	5
10mM dNTPs	1
10µM Forward Primer	0.5
10µM Reverse Primer	0.5
Water	17.75
HotStarTaq DNA Polymerase	0.25

Table 5.6. Colony PCR thermocycling conditions

Step	Temperature (°C)	Time (Seconds)	Cycle
1	95	900	
2	95	40	
3	58	30	
4	72	60	Repeat step 2-4 30 times
5	72	900	

PCR products were visualised via gel electrophoresis and colonies indicating the presence of the appropriate sized insert were expanded and cultured ON. Plasmids were isolated from these colonies using a QIAprep Spin Miniprep Kit (Qiagen). Isolated plasmids were sequenced to confirm the presence of the desired insert and determine the orientation of the insert (sequencing primers are described in table 5.7). 10 µL of plasmid was purified using the AMPure PCR purification system (Agencourt®) on the Biomek® NXP Laboratory Automation Workstation (Beckman Coulter). 5 µL of AMPure cleaned product

was combined with 5 µL of forward or reverse primer (5 pmol/µL). This was sequenced by GATC Biotech's LIGHTrun™ Sequencing service.

Table 5.7. Sequencing primer sequences located in pGL4.23 to sequence insert

Left Sequencing Primer (5'-3')	CGCTCTCCATCAAAACAAAA
Right Sequencing Primer (5'-3')	TCGAGCTTCCATTATATACCCTCT

5.2.2.2 Co-transfection of pGL4.23 and pGL4.73 into H4, Microglia-SV40, HEK293 and BV2

Following plasmid verification, plasmids were transfected into a number of cell lines. Microglia SV40, H4, BV2 and HEK293 cell lines were transfected using Lipofectamine 3000 [protocol described in Chapter 2.2.2]. Cells were plated into 24 well plates and incubated for 24 hours prior to transfection. All cells were co-transfected with pGL4.73 that encodes the luciferase reporter gene *hRluc* (*Renilla reniformis*) to act as a control for transfection efficiency. For each biological replicate, cells were independently transfected with pGL4.23 plasmids containing both sized *BIN1* inserts, in both orientations and with a risk or non-risk allele. Each transfection was repeated using plasmids that originated from independent transformed *E.coli* clones. Additional controls included transfecting cells with the same amount empty pGL4.23 containing no insert to determine baseline levels of expression from this plasmid and transfecting cells with pGL3-basic vector, which has no enhancer region, to distinguish background levels of *luc2* expression. Additional transfection parameters are described in table 5.8. Following transfection, cells were incubated at 37°C in 5% CO_2 for 48 hours. Non-transfected controls (NTC) were also performed.

Table 5.8. Transfection conditions for gene reporter assay in H4, microglia and BV2 cells

	Plating Density (cells/well)	Amount of *luc2* expressing plasmid (ng/well)	Amount of empty pGL4.23 (ng/well)	Amount of *pGL3-basic* (ng/well)	Amount of *hRluc* expressing plasmid (ng/well)
H4	3×10^4	600	600	600	20
Microglia	3×10^4	300	300	300	5
HEK293	8×10^4	300	300	300	1
BV2	2×10^4	1000	1000	1000	10

5.2.2.3 Optimising transfection of THP-1 cells

Optimisation of plasmid transfection using a GFP expressing plasmid was attempted in the THP-1 cell line to determine whether this cell line could be used for gene reporter assays. Transfection optimisation was performed on non-differentiated and differentiated THP-1 cells. THP-1 differentiation was performed by adding phorbol 12-myristate 13-acetate to a final concentration of 100nM to the culture medium and culturing the cells for 24 hours. The phorbol 12-myristate 13-acetate was then removed and replaced with normal culture media and transfected as normal. The transfection variables investigated are described in table 5.9. All transfection permutations were performed in 24 well plates and transfection efficiency was determined by visualising GFP expression using a fluorescent microscope. [The Lipofectamine 3000 transfection protocol is described in Chapter 2.2.2].

Lipofectamine™ 2000 reagent transfection requires cells to be plated out at their required density 24 hours prior to transfection. On the day of transfection,

the desired amount of DNA is made up to 25 µL with Opti-MEM into one tube. The desired volume of Lipofectamine 2000 Reagent is made up to 25 µL with Opti-MEM in another tube. The contents of each tube is then combined at a 1:1 ratio and incubated at RT from 5 minutes. During this time, Lipid-DNA complexes form. Following incubation, 50 µL of DNA-lipid complex is added to each well and incubated for 48 hours. A NTC was also performed.

Table 5.9. Transfection conditions investigated to optimise transfection in THP-1 cells

Transfection condition variable	Quantities investigated			
Number of cells/well	2×10^4	3×10^4	4×10^4	6×10^4
Amount of GFP (ng/well)	300	600	900	1000
Lipofectamine™ 2000 reagent (µL/well)	3	5	-	-
Lipofectamine™ 3000 reagent (µL/well)	1	1.5	2	2.5

5.2.2.4 Dual-luciferase Assay

Dual luciferase reporter (DLR) assays were performed using the Dual-Luciferase® Reporter Assay System (Promega UK Ltd, Southampton UK). Following the post transfection 48-hour incubation, cell lysates were produced using passive lysis buffer. 5X passive lysis buffer (supplied with DLR kit) was diluted 1:5 used ddH$_2$O. The transfection medium was removed from cells and cells were washed in PBS. 100 µL of 1X passive lysis buffer was added directly to cells. The 24-well plate containing passive lysis buffer was agitated and incubated at RT for 15 minutes.

Reconstituted Luciferase Assay Reagent II (LARII, supplied with kit) was thawed at RT. 1 volume of 50X Stop & Glo® Substrate was added to 50 volumes of Stop & Glo® Buffer. 20 µL of cell lysate was added to a luminometer plate

(Sigma). 100 μL of LARII was added to each well containing cell lysate. Luminescence was measured using a MicroLumatPlus LB96V Microplate Luminometer (EG&G Berthold). The measurement interval varied depending on cell type. Luminescence was quantified in the H4 and BV2 cells over a 10 second interval and a 1 second interval in the HEK293 cells. Each well was measured at least twice. Once luminescence from the LARII reaction had been measured, 100 μL of 1X Stop & Glo® Substrate was added and mixed. Luminescence generated from *hRluc* activity was then measured.

Quantified luminescence was blanked to NTC cell lysate readings. Each LARII reading is normalised to its own Stop & Glo® Substrate reading. Stop & Glo® Substrate reading measures *hRluc* expression and this can be used to normalise for transfection efficiency. Each normalised luminescence reading was then normalised to the pGL3-basic vector or empty pGL4.23 within its biological replicate. If the data were normalised to pGL3-basic expression, data were then normalised against the average luminescent reading from empty pGL4.23 vector across all biological replicates. Three biological replicates were performed in the H4 and BV2 cells. In the HEK293 cells, 10 biological replicates were performed using sense-orientated enhancers and 14 biological replicates were performed in antisense orientated enhancers. Data were expressed as a fold change of expression from the empty pGL4.23 vector. Data were analysed via Student's t-test or one-way ANOVA.

5.2.3 Electrophoretic Mobility Shift Assay

In order to investigate the effect of rs6733839 on protein-DNA interactions an electrophoretic mobility assay (EMSA) was performed. This technique is based on the notion that DNA-protein complexes will migrate slower than unbound DNA in a polyacrylamide gel and therefore changes in migration patterns due

to protein binding can be distinguished. An EMSA allows investigations into the cell types not used in the gene reporter assay and may yield more functionally relevant results.

5.2.3.1 Annealing Oligonucleotides

Biotinylated and non-biotinylated complementary oligonucleotides were designed to incorporate 50 bp around rs6733839 with both risk and non-risk alleles (Integrated DNA Technologies, Leuven, Belgium). Oligonucleotides sequences are described in table 5.10.

Table 5.10. Oligonucleotide sequences used in EMSA

Oligonucleotide	DNA sequence 5'-3'
Biotinylated Non-Risk Sense	/5Biosg/AAATCTCTGTTCTGCTTCTTAAAAA**C**ACC CTTTTCCCCTTTTTACTTTCAG-3'
Non-Risk Sense	AAATCTCTGTTCTGCTTCTTAAAAA**C**ACCCTTTTC CCCTTTTTACTTTCAG
Non-Risk Antisense	CTGAAAGTAAAAAGGGGAAAAGGGT**G**TTTTTAA GAAGCAGAACAGAGATTT
Biotinylated Risk Sense	/5Biosg/AAATCTCTGTTCTGCTTCTTAAAAA**T**ACC CTTTTCCCCTTTTTACTTTCAG
Risk Sense	AAATCTCTGTTCTGCTTCTTAAAAA**T**ACCCTTTTCC CCTTTTTACTTTCAG
Risk Antisense	CTGAAAGTAAAAAGGGGAAAAGGGT**A**TTTTTAAG AAGCAGAACAGAGATTT

Biotinylated oligonucleotides of the risk or non-risk alleles were annealed to non-biotinylated complementary oligonucleotides for visualisation. Non-biotinylated oligonucleotides of the risk or non-risk alleles were also

annealed to non-biotinylated complementary oligonucleotides for use as competitors.

To anneal oligonucleotides, complementary oligonucleotides were mixed together in a 1:1 ratio. Oligonucleotides were diluted to a final concentration of 1 pmol/µL in Tris buffer (10mM Tris, 1mM EDTA, 50mM NaCl (pH 8)). Oligonucleotides were annealed using a thermocycler and the following conditions: 95°C for 5 minutes, -1°C per 60 second cycle repeated 70 times. Independent annealing reactions were carried out per biological replicate.

5.2.3.2 Nuclear Protein Extraction

Nuclear protein was extracted from HEK293, BV2 and THP-1 cell lines using NE-PER Nuclear and Cytoplasmic Extraction Reagents (Thermo Fisher Scientific). HEK293 cells were harvested using trypsin and then centrifuged at 500 x g for 5 minutes. THP-1 cells are harvested by centrifugation at 500 x g for 5 minutes. BV2 cells in suspension were harvested by centrifugation, and adherent cells were trypsinised and then harvested via centrifugation. The supernatant was removed and cells were washed in PBS. 1-10 x 10^6 cells were transferred to a microcentrifuge tube and centrifuged for 3 minutes at 500 x g and the supernatant removed. Protease inhibitors were added to CER I reagent just prior to use. 100 µL of ice-cold CER I reagent was added per 1 x 10^6 cells, vortexed for 15 seconds and then incubated on ice for 10 minutes. 5.5 µL of ice-cold CER II reagent was added per 1 x 10^6, vortexed and incubated on ice for 1 minute. The sample was vortexed then centrifuged for 5 minutes at 4°C at 16000 x g.

The nuclear pellet was suspended in ice-cold NER (50 µL per 1 x 10^6) cells. The sample was vortexed for 15 seconds then incubated of ice for 40 minutes.

During this incubation, the sample was vortexed every 10 minutes. The sample was centrifuged for 10 minutes at 4°C at 16000 x g. The supernatant containing the nuclear extract is transferred to a pre-chilled tube and stored at -80°C. Protein concentrations were determined by BCA assay [method described in Chapter 2.3.1].

5.2.3.3 Protein Binding Reactions

Binding reactions and visualisation was performed using LightShift® Chemiluminescent EMSA Kit (Thermo Fisher Scientific). For HEK cells, 8 µg of nuclear extract was used per binding reaction. For THP-1 cells and BV2 cells, 10 µg of nuclear extract was used per binding reaction. To optimise the binding reaction, one of the optional reagents was added to each reaction (reagents listed in table 5.11). Binding reactions were performed with and without the competitor unlabeled oligonucleotides.

Table 5.11. Reagents for the optimisation of the binding reaction

Reagent	Final Amount
Water	Make up to 20 µL
10X Binding Buffer	2 µL
Poly (dI•dC) 1 µg/µL	1 µL
Optional: 50% Glycerol	1 µL
Optional: 1% NP-40	1 µL
Optional: 1M KCl	1 µL
Optional: 100mM MgCl$_2$	1 µL
Optional: 200mM EDTA	1 µL
Unlabeled Target DNA (1 pmol/µL)	0/4 µL
Protein Extract	8/10 µg
Biotin End-Labeled Target DNA	20 fmol

Reagents were added in the order listed in table 5.11. Prior to the Biotin End-Labeled Target DNA being added, the reaction is incubated at RT for 15 minutes. Once the Biotin End-Labeled Target DNA was added, the reaction was incubated at RT for a further 20 minutes. This optimisation was performed using both risk and non-risk allele oligonucleotides.

Following optimisation, the optimal binding reaction used THP-1 nuclear extract and EDTA. A free DNA control was performed where the binding reaction contained biotinylated DNA but no competitor non-labeled DNA and no nuclear protein. A nuclear protein:DNA binding reaction was performed which contained 10μg of nuclear protein and biotinylated DNA. A competitor reaction was performed which contained 10 μg of nuclear protein, biotinylated DNA and non-biotinylated DNA to show that the protein binding was specific. In addition, a supershift assay was performed which involved the nuclear protein-DNA binding reaction being carried out as normal, then following the protein binding incubation, 2 μL of SPI1 antibody (Santa Cruz Biotechnology-sc352) was added and incubated for 30 minutes at RT. If SPI1 is bound to the biotinylated oligonucleotides, the antibody will bind and therefore result in slower migration during electrophoresis. A supershift negative control was performed which used 2 μL of Human IgG Isotype Control antibody (Thermo Fisher Scientific). All reaction conditions were performed using risk and non-risk oligonucleotides and performed in three biological replicates. The binding reaction and supershift reactions are described in table 5.12.

Table 5.12. Summary of binding and Supershift reactions performed using THP-1 nuclear protein and EDTA reagent

Reagent	Free DNA control	Nuclear protein – DNA	Competitor	Supershift	Supershift negative control
Water	Make up to 20 µL	Make up to 20 µL	Make up to 20 µL	Make up to 22 µL	Make up to 22 µL
10X Binding Buffer	2 µL	2 µL	2 µL	2 µL	2 µL
Poly (dI•dC) 1 µg/µL	1 µL	1 µL	1 µL	1 µL	1 µL
200mM EDTA	1 µL	1 µL	1 µL	1 µL	1 µL
Unlabeled Target DNA (1 pmol/µL)	0 µL	0 µL	4 µL	0 µL	0 µL
Protein Extract	0 µg	10 µg	10 µg	10 µg	10 µg
Biotin End-Labeled Target DNA	20 fmol	20 fmol	20 fmol	20 fmol	20 fmol
SPI1 antibody	0 µL	0 µL	0 µL	2 µL	0 µL
Human IgG Isotype	0 µL	0 µL	0 µL	0 µL	2 µL

5.2.3.4 Visualisation of the Protein DNA interaction

A 5% Mini-PROTEAN® TBE precast DNA gel (Bio-Rad) was used to separate DNA oligonucleotides. The gel was placed into an electrophoresis unit filled with 0.5X TBE and the gel was pre-electrophoresed for 30 minutes by applying 100V. Once the binding reaction was complete, 5 µL of 5X loading buffer was added to the samples and then 20 µL loaded into the gel. 100V was applied and the gel was electrophoresed for 50 minutes.

The positively charged Biodyne B Nylon Membrane (Thermo Fisher Scientific) was soaked in 0.5X TBE for 10 minutes. The gel, nylon membrane and blotting paper were sandwiched and transferred to an electrophoretic transfer unit. 0.5X TBE cooled to 10°C was used as transfer buffer. 100V was applied for 20 minutes in order to transfer protein and DNA from the gel to the membrane.

Once the transfer was complete, the membrane was placed on a dry paper towel to allow the buffer to absorb into the membrane. The transferred DNA was cross-linked to the membrane by placing the membrane face down on a UV transilluminator for 15 minutes.

Blocking and Wash buffers were warmed to 37°C to ensure particulates were dissolved. The membrane was incubated with gentle shaking in blocking buffer for 15 minutes. The blocking buffer was then replaced with Stabilised Streptavidin-Horseradish Peroxidase Conjugate diluted 1:300 in blocking buffer and incubated for 15 minutes with gentle shaking. The conjugate/blocking solution was replaced with 1X wash buffer. This wash buffer was removed and replaced with more wash buffer and incubated for 5 minutes with gentle shaking. This wash step was repeated a further three times. The membrane was transferred to a new container and incubated in Substrate

Equilibration Buffer for 5 minutes with gentle shaking. The membrane was removed from the buffer and excess liquid drained. The Substrate Working Solution (Luminal/Enhancer Solution and Stable Peroxide Solution mixed at a 1:1 ratio) was added to the membrane and incubated for 5 minutes without shaking. The Substrate Working solution was removed from the membrane and excess buffer blotted from the membrane. The chemiluminescence was visualised using The OMEGA LUM™ G Imaging System (Aplegen).

Bands were quantified using ImageJ software [as described in Chapter 2.5]. To compare risk alleles to non-risk alleles, the density of the bands were expressed as a ratio and then a one-sampled t-test performed.

5.3 Results

5.3.1 Publically available data indicates *BIN1* risk locus may be regulatory active in immune cells

The HaploReg v4.1 database provides evidence that rs6733839 may be within a regulatory region active in Hematopoetic stem cells, B cells and CD14+ monocytes. ChIP-Seq data shows enhancer and promoter associated modifications in this region in these tissues, with some evidence of enhancer modifications in T cells (figure 5.2).

Description	Chromatin states (Core 15-state model)	Chromatin states (25-state model using 12 imputed marks)	H3K4me1	H3K4me3	H3K27ac	H3K9ac	DNase
Primary mononuclear cells from peripheral blood		19_DNase	H3K4me1_Enh			H3K9ac_Pro	
Primary T cells from peripheral blood							
Primary T cells effector/memory enriched from peripheral blood							
Primary T cells from cord blood		17_EnhW2					
Primary T regulatory cells from peripheral blood		19_DNase					
Primary T helper cells from peripheral blood							
Primary T helper naive cells from peripheral blood							
Primary T helper cells PMA-I stimulated							
Primary T helper 17 cells PMA-I stimulated							
Primary T helper memory cells from peripheral blood 1							
Primary T helper memory cells from peripheral blood 2							
Primary T CD8+ memory cells from peripheral blood							
Primary T helper naive cells from peripheral blood							
Primary T CD8+ naive cells from peripheral blood							
Primary monocytes from peripheral blood	7_Enh	14_EnhA2	H3K4me1_Enh		H3K27ac_Enh		
Primary B cells from cord blood		17_EnhW2	H3K4me1_Enh				
Primary hematopoietic stem cells	7_Enh	17_EnhW2					
Primary hematopoietic stem cells G-CSF-mobilized Male	7_Enh	17_EnhW2	H3K4me1_Enh				
Primary hematopoietic stem cells G-CSF-mobilized Female	7_Enh	17_EnhW2	H3K4me1_Enh				
Primary hematopoietic stem cells short term culture	7_Enh	16_EnhW1	H3K4me1_Enh				
Primary B cells from peripheral blood							
Primary Natural Killer cells from peripheral blood			H3K4me1_Enh				DNase
Primary neutrophils from peripheral blood		17_EnhW2					
Monocytes-CD14+ RO01746 Primary Cells	7_Enh	13_EnhA1	H3K4me1_Enh	H3K4me3_Pro	H3K27ac_Enh	H3K9ac_Pro	DNase

Figure 5.2. HaploReg v4.1 output for rs6733839 region in blood and immune cell types.

HaploReg v4.1 provides evidence for enhancer features in hematopoietic stem cells, B-cells and primary mononuclear cells from peripheral blood. There is also evidence of enhancer and promoter features in CD14+ Primary monocytes cells. 7_Enh indicates enhancers. 13_EnhA1 and 14_EnhA2 indicates active enhancers. 16_EnhW1 and 17_EnhW2 indicate weak enhancers. H3K4me1_Enh and H3K27ac_Enh indicate histones modifications associated with enhancers. H3K4me3_Pro and H3K9ac_Pro indicate histone modifications associated with promoters. DNase indicates DNase Hypersensitivity. Black indicates that the assay was not performed in this tissue.

RegulomeDB Version 1.1 has a scoring system based on the functional confidence of a variant from 1 to 6. Lower scores indicate increasing evidence for the variant to be located in a functional region. Known eQTL variants are

given a category 1 score, whereas variants lacking any functional annotation are labeled as category 6. RegulomeDB Version 1.1 annotates rs6733839 with a score of 5, suggesting this variant lies within a TF binding or DNase peak, but does not disrupt the site of binding (432). Using data from Pique-Regi *et al*, rs6733839 is annotated to lie in the proximity of MEF-2 and SPI1 binding motifs (436). ENCODE data suggests rs6733839 lies in a DNase Hypersensitive Site in monocyte CD14+ monocyte cells and B-Lymphocytes (437). ChromHMM regulatory regions, which utilises data from the Roadmap Epigenomics Project to annotate chromatin state (6, 438), describes the region as an enhancer in primary hematopoietic stem cells and primary CD14+ monocytes.

ChIP-Seq data from Roadmap Epigenomics Project in CD14+ monocyte cells shows the region surrounding rs6733839 has evidence of H3K4me1 and H3K27ac indicative of enhancer activity. DNase-Seq data showed that in CD14+ monocyte cells, this region is sensitive to DNase and is therefore a site of open chromatin, again consistent with this region being an active regulatory element (figure 5.3). Furthermore, ChIP-Seq data from human macrophages also indicates H3K4me1 and H3K27ac modifications at the BIN1 locus and DNase-Seq in macrophages also indicated this region is sensitive to DNase and is therefore a site of open chromatin (1) (figure 5.4).
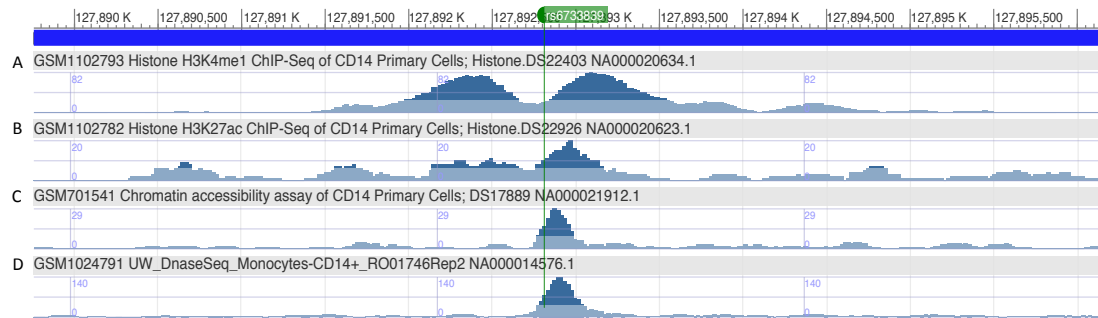
Figure 5.3. ChIP-Seq data from Roadmap for CD14+ monocytes. Image shows approximately 6 Kb surrounding rs6733839. X-axis is genomic location. Y-axis is normalised read counts, i.e., the greater the peak height, the more commonly this modification is found at this location. The green marker indicates the location of rs6733839. A and B show the presence of histone modifications that are indicative of enhancer activity surrounding rs6733839 (H3K4me1 and H3K27ac). C and D show area of open chromatin and DNase hypersensitivity surrounding rs6733839. All modifications are suggestive of regulatory activity within CD14+ monocytes at this locus. Data from (6).
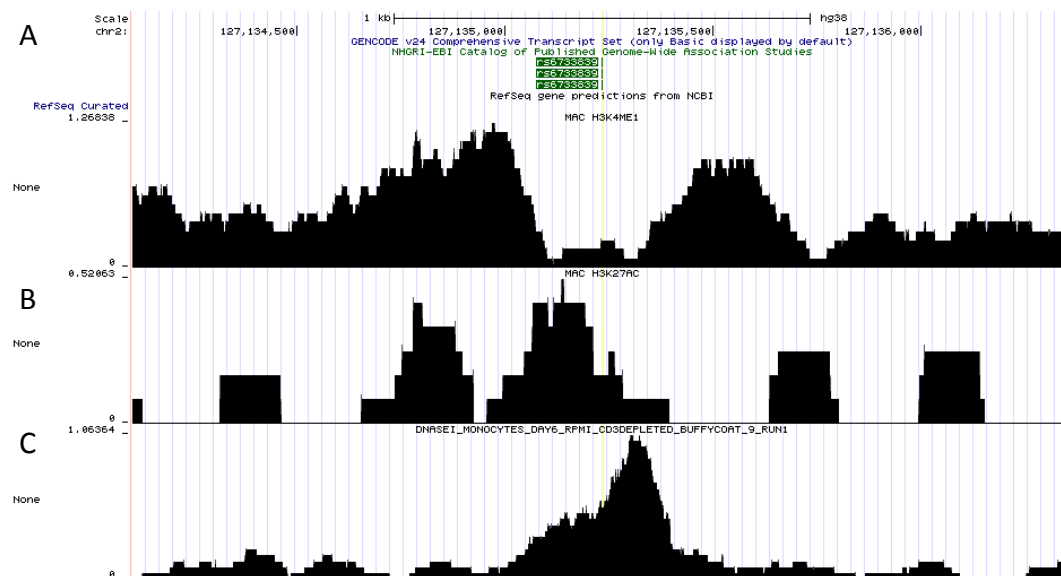
Figure 5.4. ChIP-Seq data from human macrophages. Image shows approximately 2 Kb surrounding rs6733839. X-axis is genomic location. Y-axis is normalised read counts, i.e., the greater the peak height, the more commonly this modification is found at this location. The green marker indicates the location of rs6733839. A and B show the presence of histone modifications that are indicative of enhancer activity surrounding rs6733839 (H3K4me1 and H3K27ac). C indicates area of DNase hypersensitivity surrounding rs6733839. All modifications are suggestive of regulatory activity within human macrophages at this locus. Data from (1).

These tools were then used to investigate regulatory modifications at the *BIN1*

locus in brain tissue. In contrast to the monocyte data, there were fewer

histone modifications indicative of enhancer activity across a number of

different brain tissues (figure 5.5). The histone and open chromatin profiles

surrounding the rs6733839 locus suggests that this region might have

regulatory functions in immune cell types but not brain tissue.
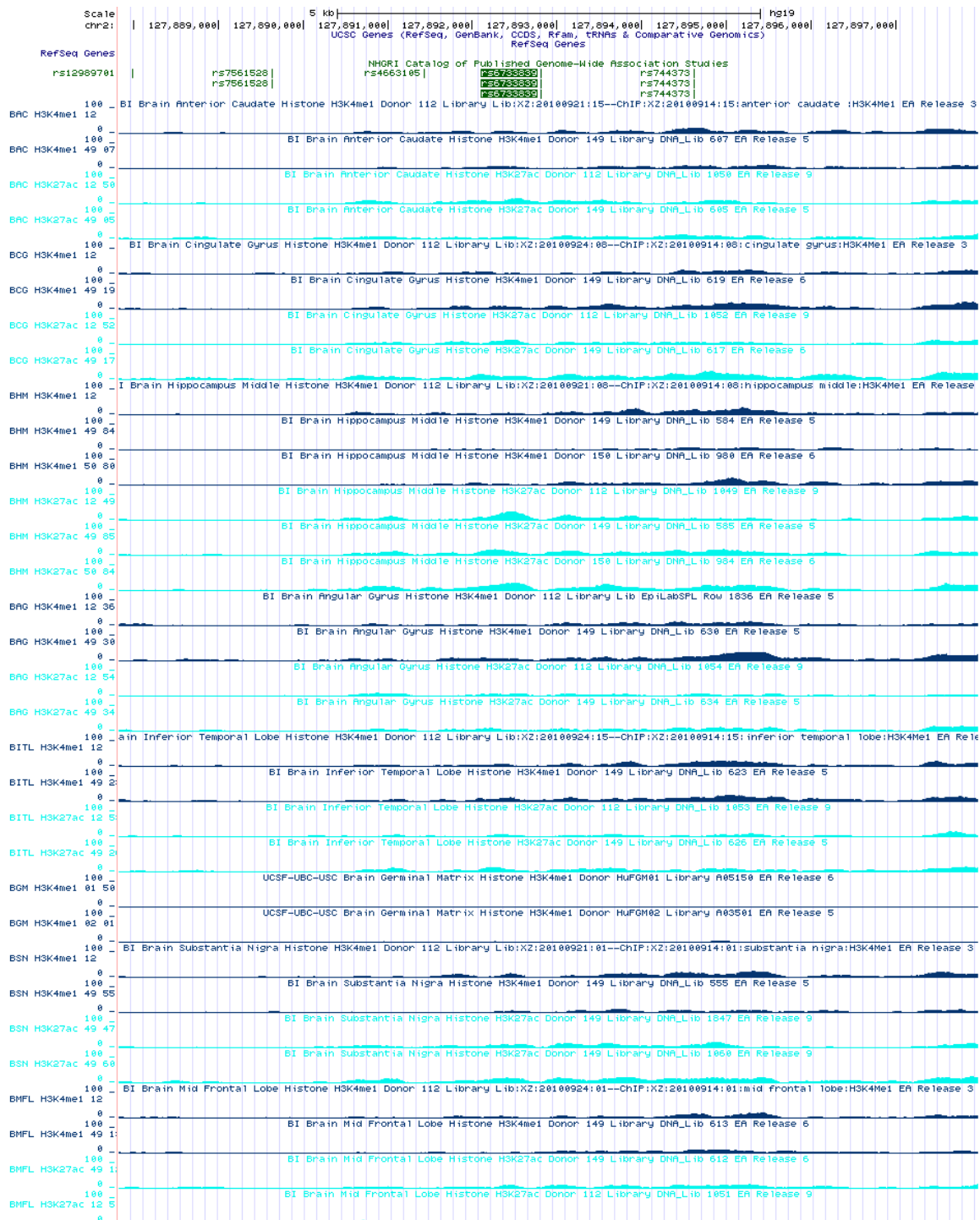
Figure 5.5. ChIP-Seq data from NIH Roadmap Epigenomics Mapping Consortium across a number of brain regions. . X-axis is genomic location. Y-axis is normalised read counts, i.e., the greater the peak height, the more commonly this modification is found at this location. Approximately 8 Kb is displayed with the location of rs6733839 indicated by the highlighted green marker. H3K4me1 and H3K27ac modifications are displayed across a number of brain regions: Anterior caudate, cingulate gyrus, hippocampus middle, angular gyrus, inferior temporal lobe, brain germinal matrix, substantia nigra and mid frontal lobe. Data indicates low levels of histone modification in these regions, suggesting this region is less regulatory active in brain regions than in monocytes. Data from (6).

Further to the histone modification data, Cistrome DB was used to investigate TF binding at the *BIN1* locus. ChIP-Seq performed in the THP-1 human monocytic cell line and from human macrophages identified SPI1 binding at the *BIN1* risk locus (figure 5.6 and figure 5.7 respectively).
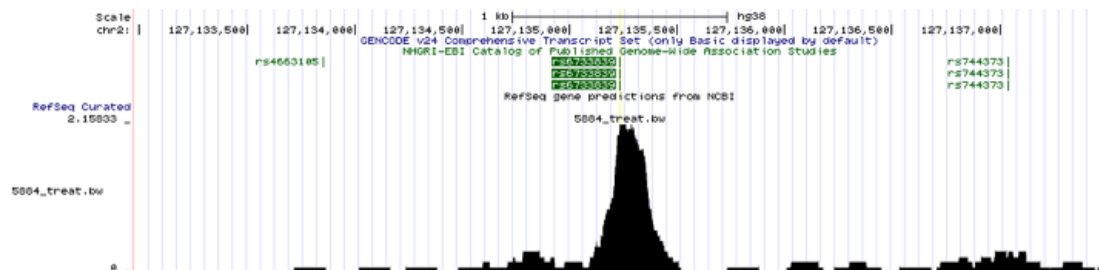
Figure 5.6. ChIP-Seq data indicate SPI1 binding at the *BIN1* locus in THP-1 cell line. . X-axis is genomic location. Y-axis is normalised read counts, i.e., the greater the peak height, the more commonly this modification is found at this location. The location of rs6733839 is indicated by the green highlighted marker. Data are taken from (4).
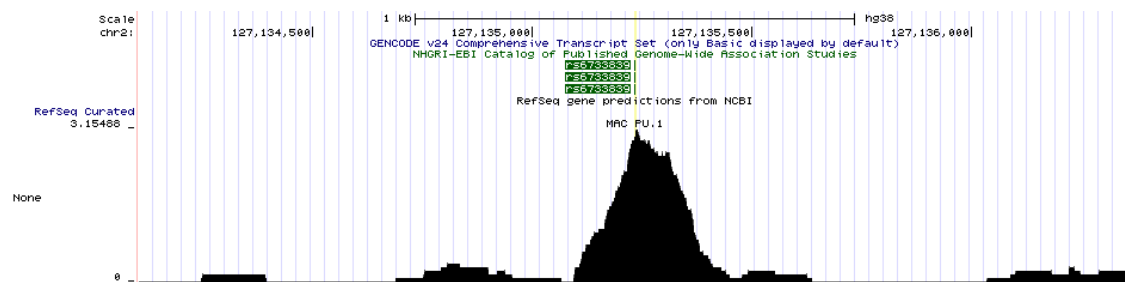
Figure 5.7. ChIP-Seq data indicate SPI1 binding at the *BIN1* locus in human macrophages. X-axis is genomic location. Y-axis is normalised read counts, i.e., the greater the peak height, the more commonly this modification is found at this location. The location of rs6733839 is indicated by the green highlighted marker. Data are taken from (473).

5.3.2 Gene Reporter Assay

As ChIP-Seq data indicates the presence of enhancer modifications and TF binding at the *BIN1* risk locus, a gene reporter assay was designed to characterise the enhancer function of this locus in multiple cell lines.

5.3.2.1 Amplification of potential enhancer regions of interest

Primers were designed to amplify potential enhancer regions to clone into the luciferase expressing plasmid pGL4.23. Primer set 1 was designed to amplify the entire potential SPI1 binding region shown in figure 5.6. An additional primer set was designed to amplify a more localised region surrounding rs6733839 that spanned the dip in the histone modifications (figure 5.3). A high-fidelity PCR was optimised to amplify these regions and the PCR products were visualised via gel electrophoresis (figure 5.8).
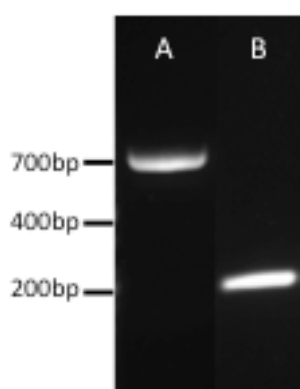
Figure 5.8. High fidelity PCR amplifies potential enhancer regions for insertion into pGL4.23 plasmid. Lane A shows the 693 bp fragment that encompasses the SPI1 binding region. Lane B shows the localised 235bp fragment surrounding rs6733839.

5.3.2.2 Validation enhancer cloning into pGL4.23

Following successful amplification of samples containing the appropriate genotype, inserts were ligated into pGL4.23 and transformed into competent *E.coli*. A colony PCR was performed on the transformed clones. The colony PCR identified clones that likely contained the insert present as appropriately sized PCR products were observed (figure 5.9). Clones that indicated the presence of the desired insert were further cultured and plasmid DNA was isolated. The plasmid DNA was purified and sent for Sanger sequencing which confirmed the presence of the desired insert, its orientation in relation to the endogenous promoter and its rs6733839 genotype. A representative chromatogram is shown in figure 5.10. No other DNA changes were observed. [Full insert sequences are shown in Appendix 2].
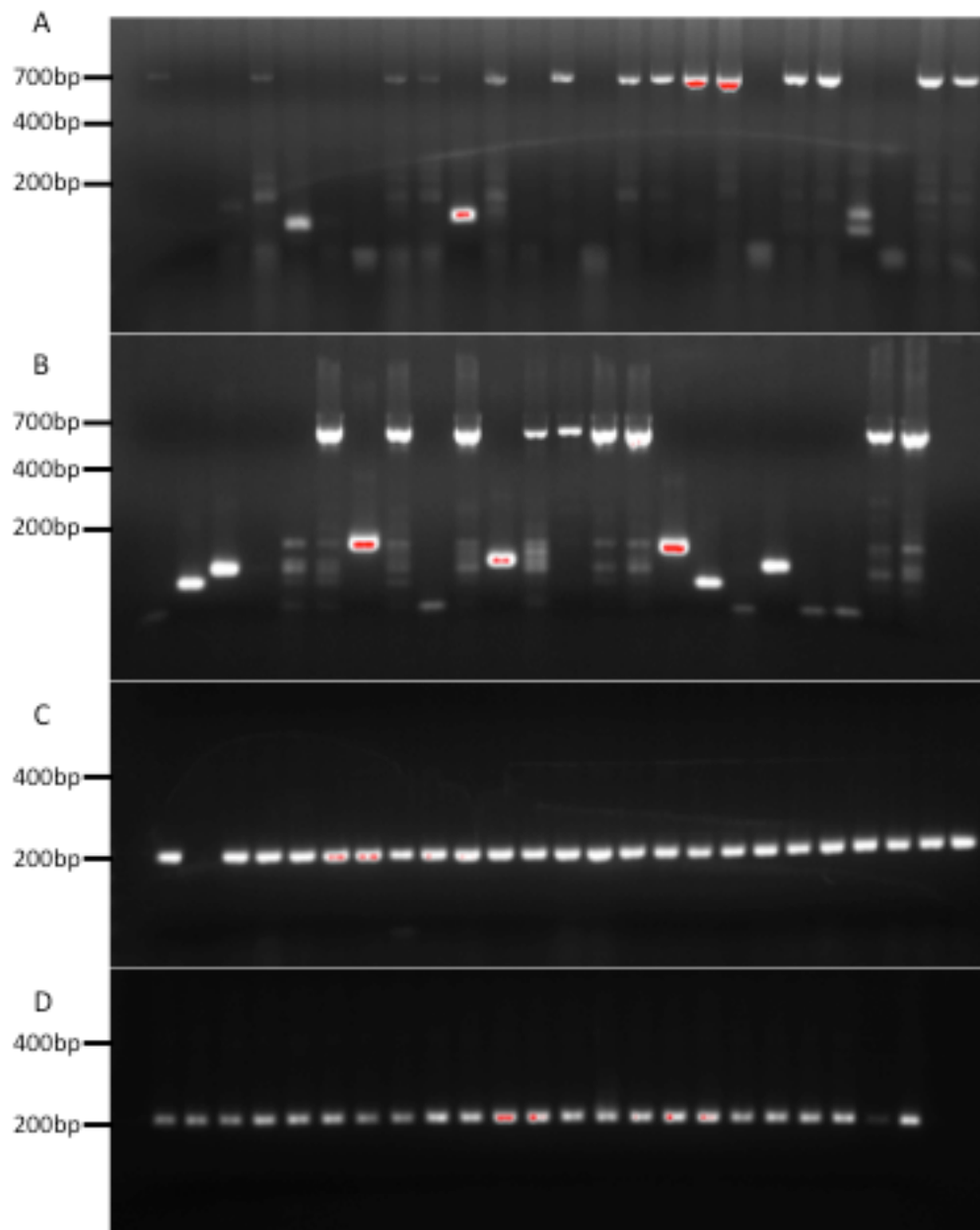
Figure 5.9. Colony PCR indicating transformed clones that contained cloned enhancer region. Each lane contains PCR products amplified from a different clone. A) PCR products from *E.coli* transformed with plasmid containing the 693 bp fragment with the non-risk allele. B) PCR products from *E.coli* transformed with plasmid containing the 693 bp fragment with the risk allele. C) PCR products from *E.coli* transformed with plasmid containing the 235 bp fragment with the non-risk allele. D) PCR products from *E.coli* transformed with plasmid containing the 235 bp fragment with the risk allele.
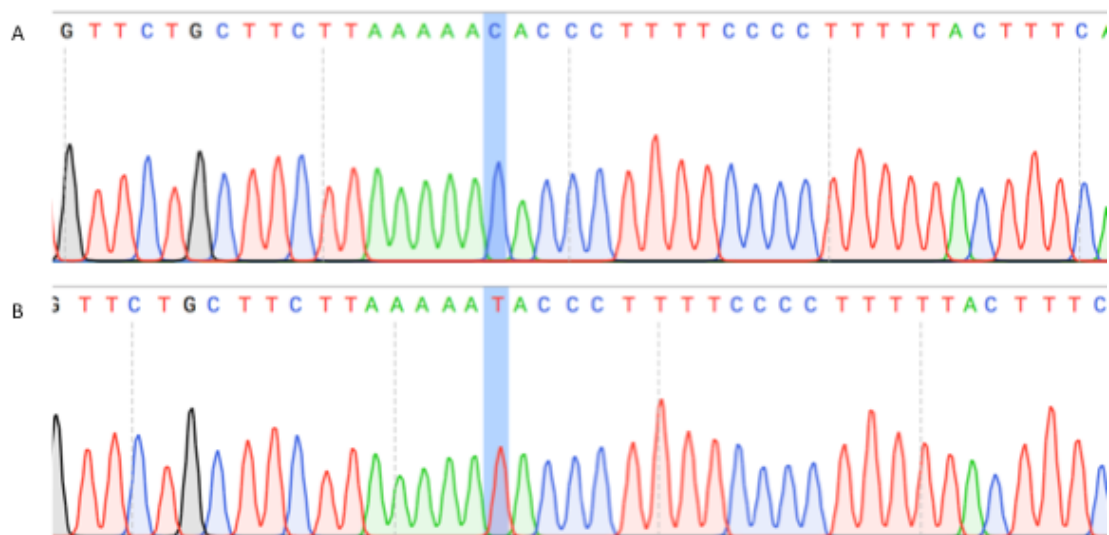
Figure 5.10. Representative chromatograms from pGL4.23 plasmids containing *BIN1* risk locus. A) Shows an insert homozygous for the non-risk rs6733839 allele in the antisense orientation (highlighted in blue). B) Shows a plasmid homozygous for the risk rs6733839 allele orientation (highlighted in blue).

5.3.2.3 Luciferase activity in H4 cell line indicates potential orientation-specific enhancer function

Plasmids were transfected in the H4 cell line [as described in Chapter 2.2.2]. The data were normalised to expression from the empty pGL4.23 for each biological replicate. Three biological replicates were performed using clone 1. Replication of the results was attempted using a second independent clone, which had originated from a separate transformed *E.coli* colony. Results are shown in figure 5.11. A Levene's test revealed the data to have unequal variances. A one-way ANOVA revealed a significant difference in luciferase activity between the test groups. A Dunnett's T3 test was performed to make pairwise comparisons.
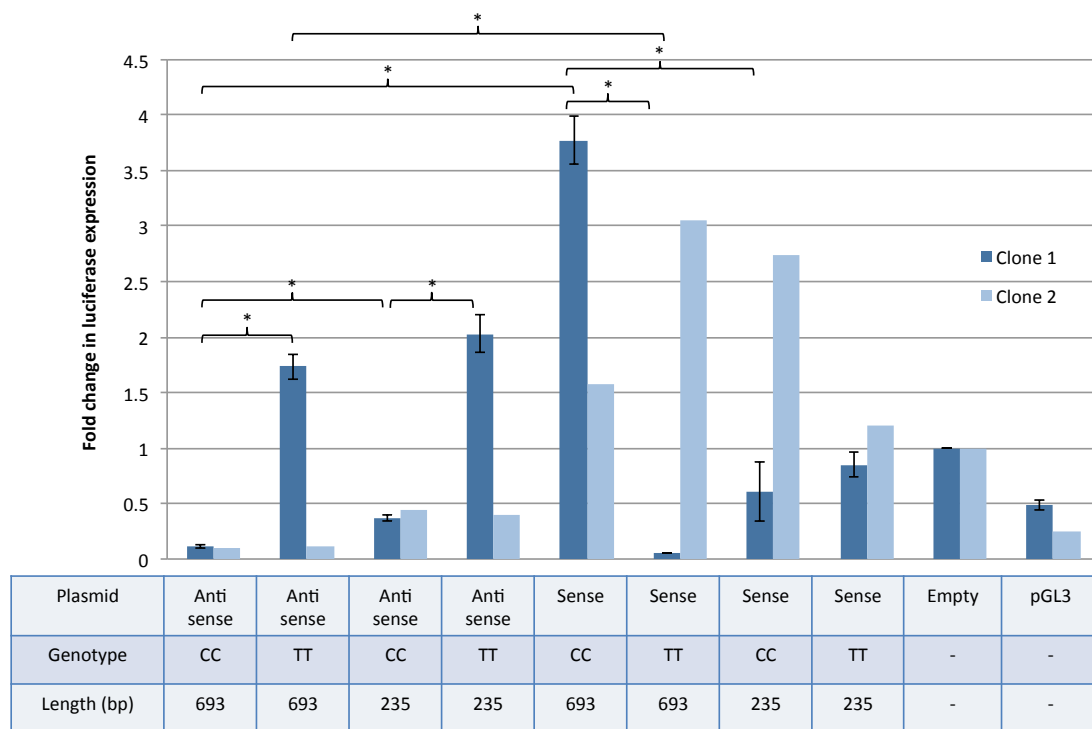
Figure 5.11. Luciferase expression in H4 cell line expressed as a fold change in comparison to the empty pGL4.23 plasmid. Three biological replicates were performed using clone one. One biological replicate was performed using clone 2. *p≤0.05. Error bars indicate SEM.

There are significant differences between orientations observed. There is a significant difference between the orientations of the 639 bp risk allele (p=0.026) and the non-risk allele (p=0.019). No other orientation specific differences were observed.

Enhancer size appeared to have a significant effect on the non-risk allele. A significant difference between the 639 bp and 235 bp antisense orientated non-risk allele was observed (p=0.021) and a significant difference between the 639 bp and 235 bp sense orientated non-risk allele was observed (p=0.010).

Finally, some allele specific changes in luciferase activity were detected. In the 693 bp and the 235 bp antisense orientated enhancer a significant allele

specific difference in luciferase was observed (p=0.025 and p=0.049 respectively). In the 693 bp sense orientated enhancer a significant allele specific difference in luciferase was observed (p=0.019). Despite these results, the inconsistencies between clones suggest further validation is required.

The most consistent relationship between allele and enhancer size with expected biology is enhancers in the antisense orientation, therefore this orientation was further investigated in other cell types. Three independent clones originating from separate transformed *E.coli* colonies were preliminarily screened and two independent clones were used in future experiments. As only three clones were screened, this introduces an element of selection bias. Therefore, further validation of results would require replication in a greater number of independent clones.

5.3.2.4 SV40 Microglia cell line was incompatible with this luciferase reporter assay

Microglia were transfected with antisense oriented enhancer plasmids and appropriate controls. Three biological replicates were performed. Luminescence was quantified in 20 µL of cell lysate over 10 seconds. Luminescence originating from the firefly luciferase reached the maximum possible reading on the plate reader used. This result was true for all cellular populations, including the NTC. Luminescence originating from the Renilla luciferase was measured at variable levels in transfected cells, which is expected as this expression is used to normalise for transfection efficiency. Very low levels were detected in the NTC, consistent with background luminescence. This result suggests that microglia appear to express a protein that reacts with the LARII reagent to produce luminescence (figure 5.12).

Plasmid	Anti sense	Anti sense	Anti sense	Anti sense	Empty	pGL3	NTC
Genotype	CC	TT	CC	TT			
Length (bp)	693	693	235	235			
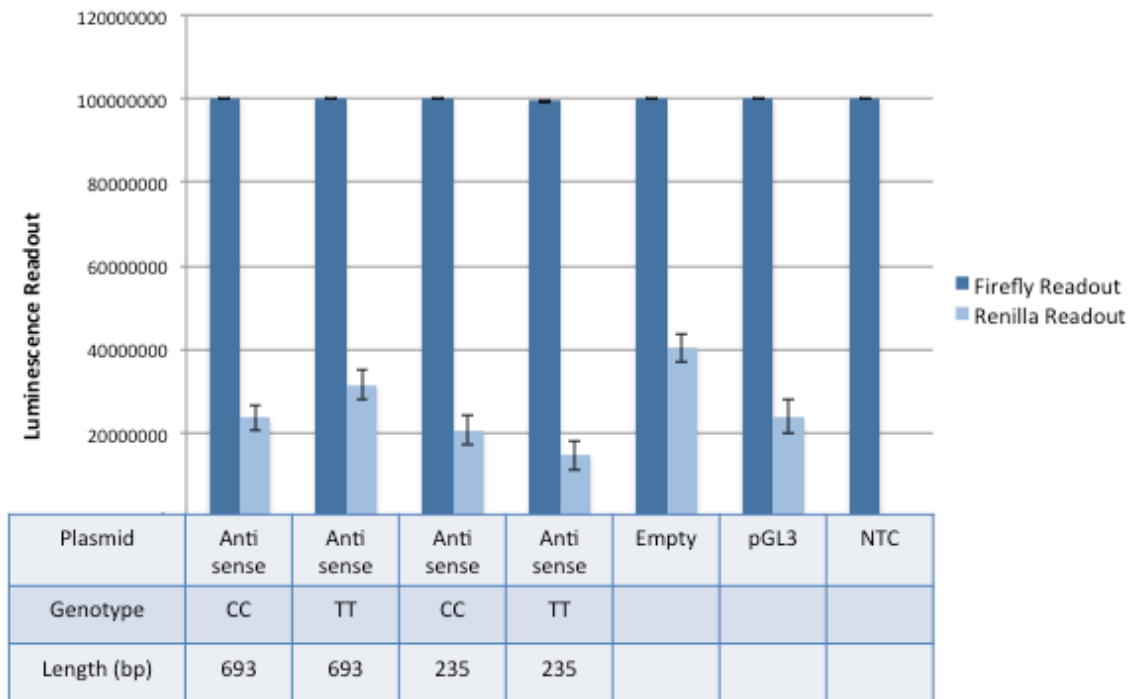
Figure 5.12. Luminescence readouts from pGL4.23 and pGL4.73 in the microglia cell line.
Data are from three biological replicates. High levels of luminescence detected from firefly luciferase were detected in all cellular populations, suggesting something present in the microglia cells interacts with LARII to produce luminesces. Luminescence from Renilla luciferase is measured at expected levels. Error bars indicate SEM.

To determine if changes in luciferase expression could be detected despite the high levels of background luminescence a reduced volume of lysate was measured (10 µL) and luminescence was quantified over 1 second measurement interval. Despite reduced levels of luminescence being quantified, high levels were detected in the NTC (figure 5.13).
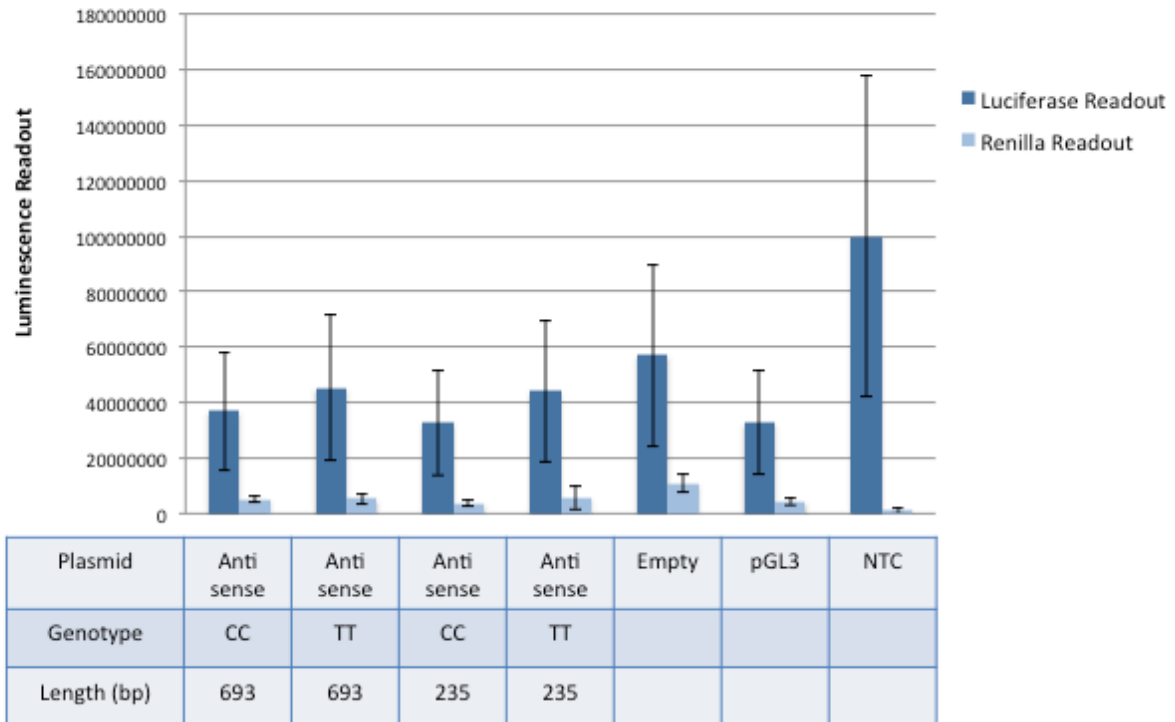
Figure 5.13. Luminescence readout from the microglia with reduced lysate and measurement interval. High levels of luminescence produced from the firefly luciferase are recorded in all cellular populations and measurements are variable. Reliable data were not yielded when using reduced lysate volume and measurement intervals. Error bars indicate SEM.

With the reduced cell lysate volume and reduced measurement interval the luminescence readout from transfected cells was variable and high levels of luminescence was detected from the NTC. In a troubleshooting attempt to determine what was causing the high levels of background luminescence, a luciferase assay was performed on microglia media, H4 media, collagen coating, NTC microglia and H4s and cells transfected with pGL4.23. The luminescence readouts from the various conditions investigated are shown in figure 5.14. This revealed that it was not the media or collagen coating that was causing the high levels of background luminescence. There is nothing in the literature that suggests that this microglia cell line expresses *luc2* but they

appear to produce a substrate that interacts with the LARII enzyme to produce luminescence. This indicates that the microglia cell line is an inappropriate cell line to use in this assay.

Collagen coating	Yes	No	No	Yes	No	No	Yes	No	No
Media	Prigrow	Prigrow	Opti-MEM	Prigrow	Prigrow	Opti-MEM	Prigrow	Prigrow	Opti-MEM
Cell Type	No	No	No	MG	MG	H4	MG	MG	H4
Transfection	No	No	No	No	No	No	pGL4.23	pGL4.23	pGL4.23
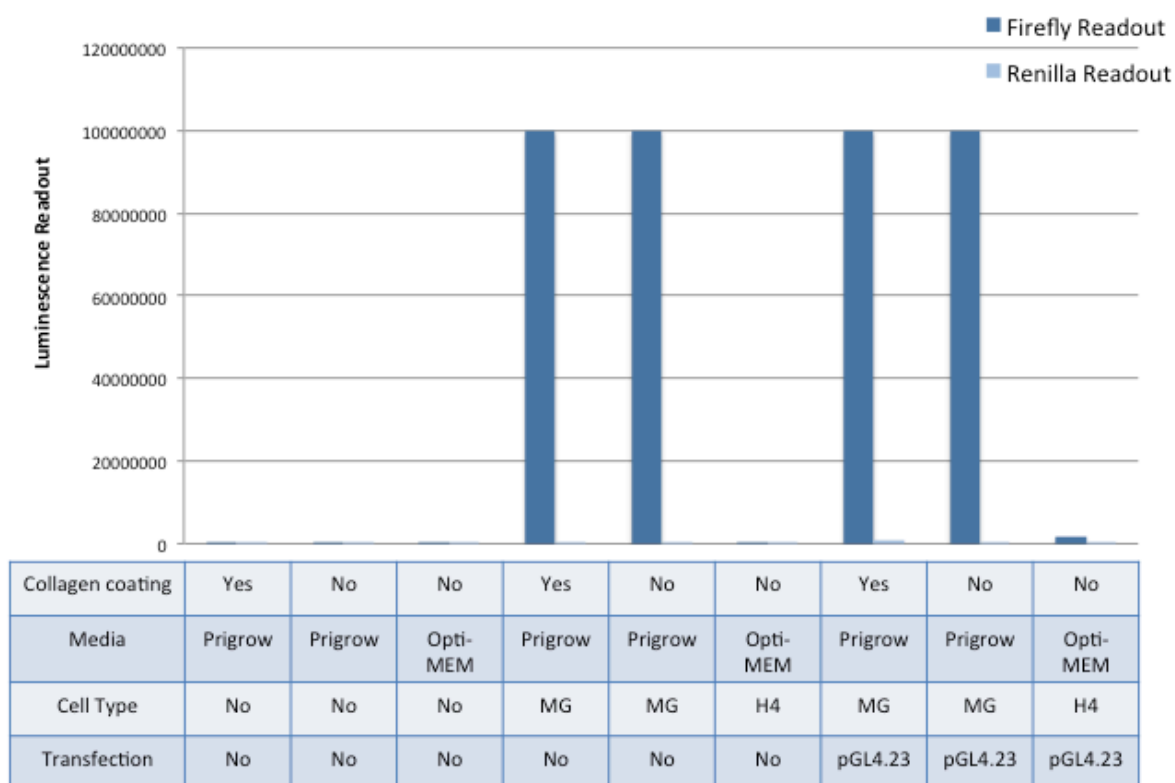
Figure 5.14. Luminescence readout from microglia troubleshooting investigation. Low levels of luminescence were detected in all cell culture medium and in H4 cells. The high levels of luminescence are detected in transfected and no-transfected microglia, suggesting microglia endogenously express something that interacts with LARII to produce luminescence. This makes this cell line incompatible with this experimental design. Error bars indicate SEM.

5.3.2.5 HEK293 cell line show orientation specific enhancer activity

HEK293 cells were transfected with two independent clones that had been previously screened. Luminescence data were normalised to the expression from the pGL3-basic vector for each biological replicate. The data were then further normalised to the average expression from the empty pGL4.23 vector across all biological replicates. For enhancers in the sense orientation, 10

biological replicates were performed. For enhancers in the antisense orientation, 14 biological replicates were performed.

Initially a Student's t-test was performed to determine the effect of orientation on luciferase activity. There was a significant difference in average fold change of luciferase activity between all antisense orientated enhancers and sense orientated enhancers (p≤0.001). The sense-orientated enhancers showed minimal luciferase activity, less so than the empty control vector, and were therefore excluded from further analysis (figure 5.15). It is possible that this lack of activity is due to this region having silencer functions in the sense orientation resulting in repression of gene expression.

Plasmid	Sense	Sense	Sense	Sense	Empty
Genotype	CC	TT	CC	TT	-
Length (bp)	693	693	235	235	-
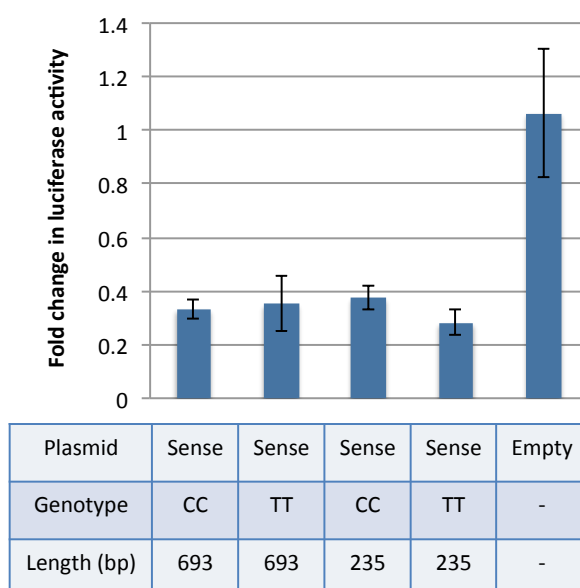
Figure 5.15. Fold change in luciferase activity from sense orientated enhancers in HEK293 cells. Sense orientated enhancers reduced luciferase activity in comparison to the empty control vector and were therefore excluded from further analysis. N=10. Error bars indicate SEM.

A one-way ANOVA revealed a significant difference in luciferase activity between the antisense orientated enhancers (p≤0.001). A Tukey's HSD test performed multiple comparisons between the groups. A significant decrease

in luciferase activity was observed in the 235 bp enhancer when compared to the 693 bp enhancer in the non-risk allele (p=0.015) and the risk allele (p=0.013). No allele specific effect on luciferase activity was observed. Data are shown in figure 5.16.

Plasmid	Anti sense	Anti sense	Anti sense	Anti sense	Empty
Genotype	CC	TT	CC	TT	-
Length (bp)	693	693	235	235	-
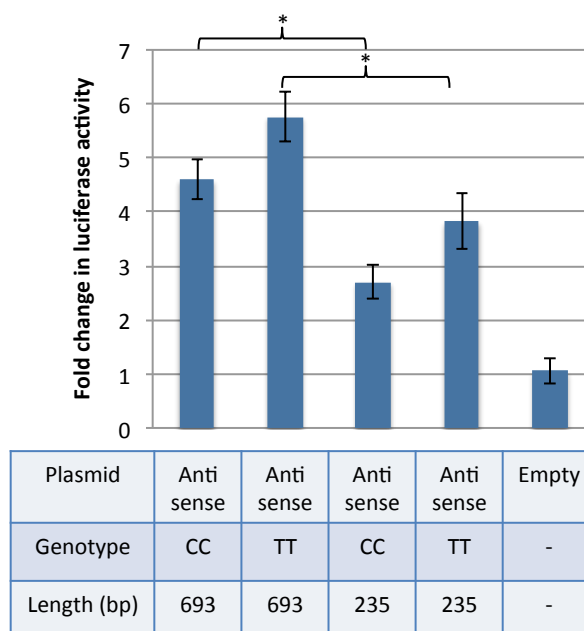
Figure 5.16. Fold change in luciferase expression from antisense orientated enhancers. Significant differences were observed between 639 bp and 235 bp enhancers. *p≤0.05. No allele specific effect was observed. N=14. Error bars indicate SEM.

5.3.2.6 Preliminary data in the BV2 cells

In an attempt to generate gene reporter data from a microglial cell type, the assay was performed in the mouse microglia cell line BV2 using enhancers in the antisense orientation. Despite low transfection efficiency, six biological replicates were performed using two independent clones and data were normalised to expression from the empty pGL4.23 vector per biological replicate. A one-way ANOVA revealed no significant differences between the enhancers (p=0.631). Despite no significant changes being observed, the expression patterns could be reflective of real biological mechanisms and, if

efficient transfection was achieved, may yield significant results. Data are shown in figure 5.17.

Plasmid	Anti sense	Anti sense	Anti sense	Anti sense	Empty	pGL3
Genotype	CC	TT	CC	TT		
Length (bp)	693	693	235	235		
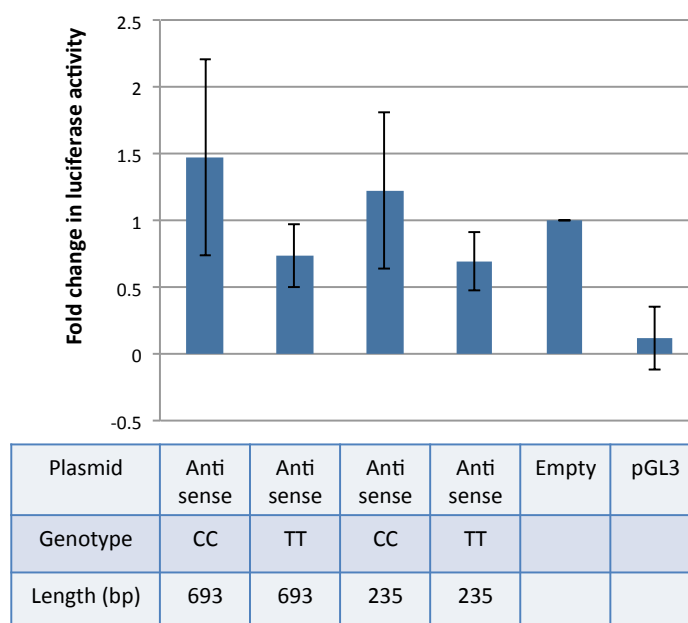
Figure 5.17. Luciferase expression in the BV2 cell line. Data are expressed as a fold change in luciferase expression in comparison to the empty pGL4.23 plasmid Data suggests a potential risk allele affect, but efficient transfection was not achieved. N=6. Error bars indicate SEM.

5.3.2.7 Efficient transfection of the THP-1 cells could not be achieved

Transfection of the GFP plasmid was performed in the THP-1 cells. Transfection efficiency was estimated by visualising GFP expression using fluorescent microscopy. Transfection was performed on undifferentiated and phorbol 12-myristate 13-acetate differentiated cells, but despite altering a number of variables, efficient transfection was not achieved and therefore the gene reporter assay was not performed in these cells.

5.3.3 EMSA showed that nuclear proteins interact with the *BIN1* risk locus

5.3.3.1 EMSA optimisation in HEK293, THP-1 and BV2 cells

The EMSA was optimised in a number of cell lines to determine the optimal conditions for DNA:protein interactions. The optimisations were performed in the HEK293, THP-1 and BV2 nuclear extract. Optimisations involved screening a number of reagents to identify optimal binding conditions and this was performed on both the risk and non-risk allele. Blots were overexposed to image faint bands present due to protein binding (figure 5.18). This could not detect the binding of nuclear proteins to the *BIN1* risk locus in the HEK293 cell line. Nuclear proteins bound specifically to the *BIN1* locus in the THP-1 and BV2 cells, but binding was more prominent in the THP-1 cells. The optimisation performed in the THP-1 cells showed minimal differences between conditions, but the clearest bands were observed in the binding reaction containing EDTA, therefore this condition was used in future investigations.
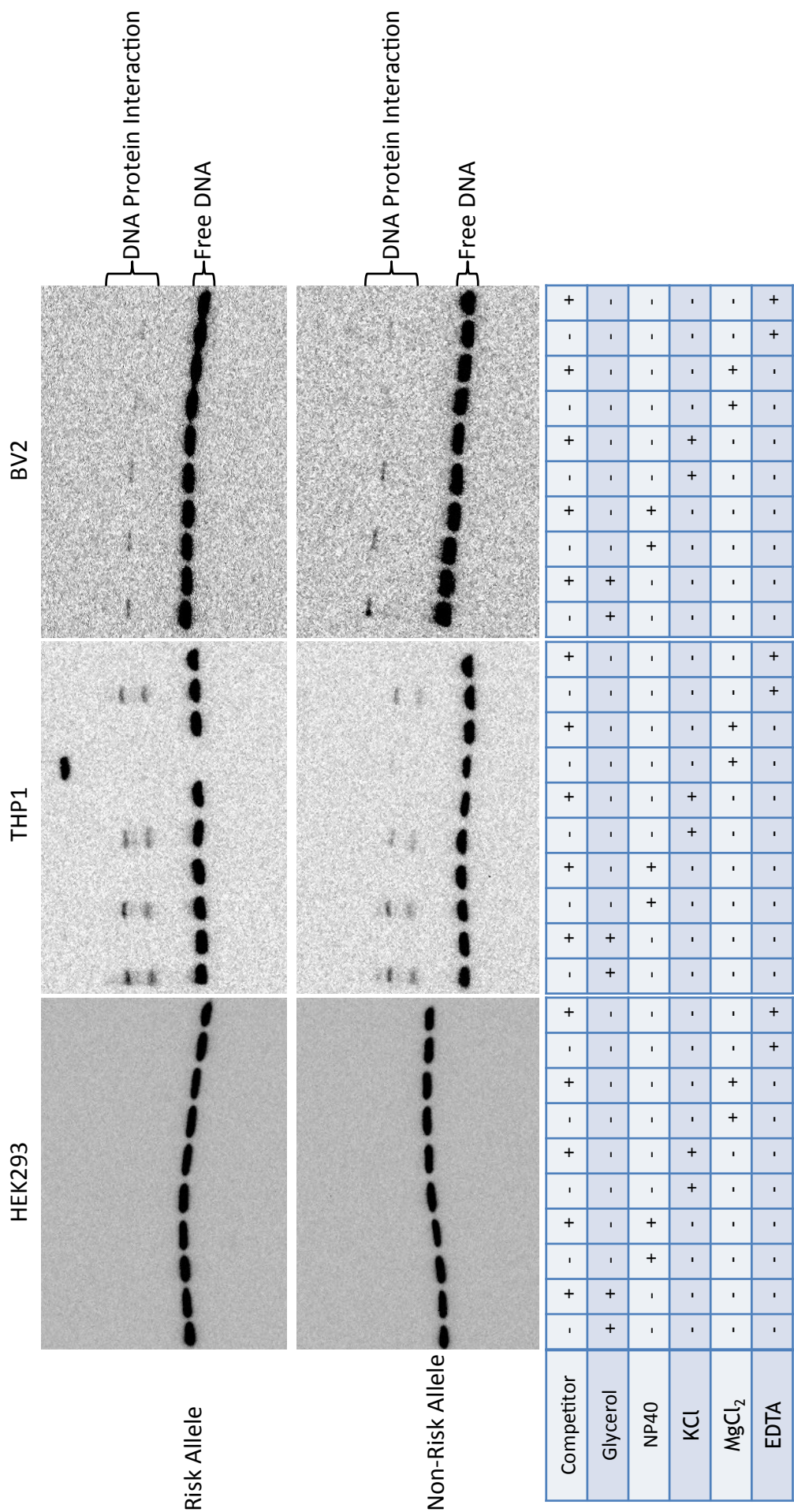
Figure 5.18. EMSA optimisation in HEK293, THP-1 and BV2 cells. Image processed in ImageJ. No DNA:protein interactions were detected in the HEK293 cell line. Minimal DNA:protein interaction was detected in the BV2 cell line. The most prominent DNA:protein interaction was detected in the THP-1 cells, with the clearest bands appearing in the EDTA containing binding reaction.

5.3.3.2 EMSA indicates SPI1 binds to the BIN1 *risk locus in THP-1 cells*

The optimised EMSA was performed in three biological replicates using the THP-1 cells. A supershift assay was also performed using a SPI1 antibody. Images were taken at low exposure to get unsaturated images of the free DNA and at high exposure to get clear images of the shifted bands. The two bands suggestive of protein:DNA interaction were quantified using ImageJ software. Shifted bands were normalised to free DNA to take into account loading errors. [Full method of quantifying bands is described in Chapter 2.5]. Data were then expressed as a risk:non-risk protein-binding ratio. The lack of bands in the lanes containing the competitor DNA indicates the binding is specific to the *BIN1* risk locus. A one-sampled t-test showed no significant difference in protein binding for either band in the non-supershifted reactions in relation to genotype (p=0.564 (top band) and p=0.432 (bottom band)).

In the supershifted reaction, a supershift has occurred, indicating SPI1 or an SPI1 containing complex is bound to the *BIN1* risk locus. The SPI1 bound bands were quantified using ImageJ. The average risk:non-risk protein-binding ratio was 0.655. This was revealed not to be significantly different from 1 in a one-sampled t-test (p=0.069), but is showing a potential trend. There is still an unsupershifted band present following SPI1 supershifting, suggesting that proteins other than SPI1 also bind this region. Data are shown in figure 5.19.

Allele	Risk			Non-Risk			Risk		Non-Risk	
Nuclear protein	-	+	+	-	+	+	+	+	+	+
Competitor	-	-	+	-	-	+	-	-	-	-
SPI1 antibody	-	-	-	-	-	-	+	-	+	-
IgG only	-	-	-	-	-	-	-	+	-	+
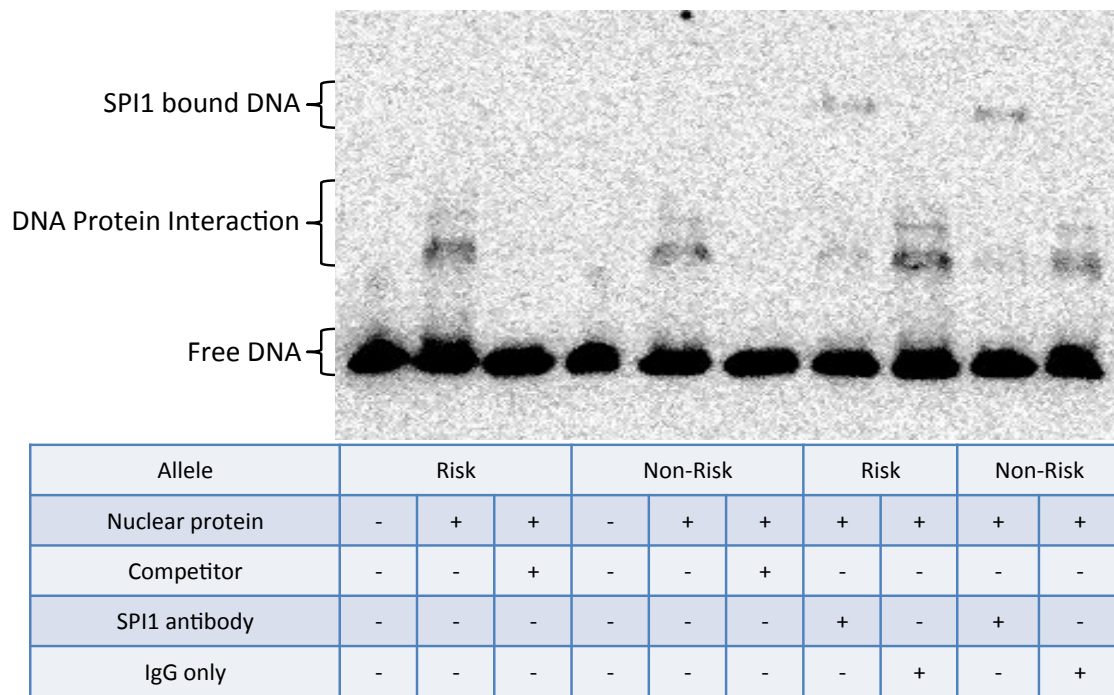
Figure 5.19. Representative EMSA and supershift assay performed in the THP-1 cell line.

Images were processed using ImageJ. Three biological replicates were performed. Shifting indicates proteins are interacting with the rs6733839 plus the surrounding sequence. The presence of supershifting when using the SPI1 antibody indicates that SPI1 binds to, or is present in a protein complex which binds to the *BIN1* risk locus. No significant differences in DNA: protein binding between risk and non-risk alleles was detected.

5.4 Discussion

ChIP-Seq data indicates modifications indicative of active regulatory mechanisms are present at the *BIN1* risk locus within monocytes and immune cells. Data from ChIP-Seq was used to design a gene reporter assay. This revealed that the *BIN1* risk locus is capable of acting as an enhancer and that the entire TF binding region is required for optimal function. The risk allele did not have a significant affect on enhancer function. DNA:protein binding assays revealed that nuclear proteins bind to the *BIN1* risk locus in THP-1 cells and one of these proteins is SPI1.

5.4.1 ChIP-Seq data imply regulatory function in monocytes and macrophages

Data from HaploReg and RegulomeDB suggest rs6733839 lies within a genomic region that has regulatory potential in blood and immune cell types. Data from HaploReg and RegulomeDB describing rs6733839 in neuronal and brain tissue has been previously discussed in chapter 4 and also indicates regulatory elements are present at this locus within these tissues.

ChIP-Seq data from The Roadmap Epigenomics Project denotes H3K4me1 and H3K27ac histone modifications and an open chromatin structure surrounding rs6733839, suggesting active enhancer activity, within monocytes and macrophages (figure 5.3 and 5.4). Furthermore, SPI1 TF binding was shown in the monocytic cell line THP-1 and macrophages further implicating a regulatory function at this genomic location (figure 5.6 and 5.7). This is in contrast to brain tissues, which showed minimal evidence of regulatory function at this locus (figure 5.5).

Monocytes are circulating leukocytes that are central in the innate immune system. Monocytes are incompletely differentiated phagocytic cells that give rise to a heterogeneous mononuclear phagocyte lineage in response to stress signals by expressing a variety of cell surface molecules, determining functional response (474, 475). When monocytes leave the blood and enter the affected tissue, they differentiate into macrophages, a specialised cell that detects, phagocytoses and destroys pathogens (466, 476). Microglia are the CNS resident macrophage and survey the brain for pathogens (54).

Numerous studies have linked neuroinflammation and AD pathology (477). Aβ deposition in the brain has been shown to induce inflammation, monocyte recruitment, infiltration and activation in the brain vasculature (478, 479). In post mortem brains of AD patients, activated microglia surround the Aβ plaques (480). Furthermore, the presence of Aβ can prime microglia cells, making them susceptible to secondary stimulus or promotes their activation (481). In AD, Aβ sustains chronic activation of primed microglia, leading to the production of inflammatory cytokines and chemokines, ultimately leading to neurodegeneration and neuronal loss (482).

Vascular patrolling monocytes have also been shown to act to eliminate soluble Aβ and microglia are implicated in Aβ clearance (483, 484). In addition, there is evidence that systemic and local chronic inflammation may contribute to neurodegeneration (485, 486).

Recent investigations are now beginning to link genetic mechanisms to immune cells and function. AD GWAS variants are enriched in enhancer orthologues with an immune function, implicating immune processes in AD.

This enrichment was stronger in distal enhancers over promoter regions, implicating distal enhancers in AD risk. One of the loci investigated was *BIN1* and was shown to be an active enhancer in BV2 and neuroblastoma cells (487). Furthermore, evidence suggests that variants exhibiting functional marks for chromatin accessibility, such as histone modifications and DNase I hypersensitivity, in immune cells significantly contribute to the heritability of AD (488).

5.4.2 *BIN1* risk allele can act as an regulatory element

In the H4 cells, the *BIN1* risk allele had a variable affect on luciferase activity. In some cases (antisense non-risk 693 bp, antisense non-risk 235 bp and sense risk 693 bp) expression in relation to the empty vector was significantly reduced and similar levels observed to that of the promoterless pGL3-basic vector, indicating the regulatory element has detrimental effects on expression and may abolish promoter function or act a genetic silencer.

In the sense orientation, no obvious pattern in changes of gene expression were observed and there are inconsistencies between independent clones, indicating further validation of these results is required. In the antisense direction, again there are inconsistencies between clones, but there does appear to be a pattern of increased activity in the risk allele (figure 5.11). Although this result has not been replicated, this would be consistent with observations of increased *BIN1* mRNA in the frontal cortex of AD patients (261).

The HEK293 cell line, although not disease relevant was used in this investigation due to its wide use, reliability during transfection and neuronal

phenotype. The efficient transfection and replication in independent clones generated consistent data that could reflect genuine biology. Firstly, a Student's t-test revealed the alleles were acting in an orientation specific manner, the sense orientation shows approximately a third of the expression of the empty vector, whereas the antisense orientation shows between an average 2.7 and 5.7 fold increase in activity (figure 5.15 and 5.16).

It has long been thought that enhancer's function independent of orientation. A recent study found that sequence orientation could impact chromosome architecture and effect proper enhancer-promoter interactions. It was found that orientation of the CTCF recognition sequence determined the formation of specific DNA looping interactions between promoters in mammalian cells (489). CTCF is a zinc finger DNA binding protein which functions as an insulator binding protein with a key role in genome looping (490, 491). Therefore it could be possible that the *BIN1* allele also functions in an orientation specific manner.

Secondly, a significant increase in activity was observed when the entire TF-binding region identified by ChIP-Seq was present in comparison to the localised region in the antisense orientation. In the non-risk allele containing enhancers, an average 1.7 fold increase was seen in the 693 bp fragment compared to the 235 bp fragment (p=0.015). For the risk allele containing enhancers, an average 1.5 fold increase in activity was seen in the 693bp fragment compared to the 235 bp fragment (p=0.013). This suggests that this entire TF region is required for optimal regulatory activity, rather than the localised region surrounding rs6733839.

The action of TFs can be influenced by a number of local features, such as the binding of additional TF, histone modification and chromatin remodeling. It may be that for optimal transcription regulated by TFs the entire 693 bp is required to create the appropriate environmental conditions.

The final gene reporter observation suggests the risk allele has no significant effect on enhancer function. The 693bp fragment showed an average of 25.3% increase in activity with the risk allele in comparison to the non-risk allele, however this results was not significant (p=0.235). The 235 bp fragment showed an average 41.4% increase in activity with the risk allele, but this was again not significant (p=0.263). Despite not being significant, this consistent observation of an increase in activity with the risk allele suggests it may have the potential to influence gene expression and this effect could be more prominent in a more disease relevant cell line. Alternatively, it could be that rs6733839 is not a functional variant at this locus, but is tagging an unknown functional variant that influences regulatory elements at this region of the genome.

In the BV2 cell line, luciferase activity was very variable due to poor transfection efficiency and few biological replicates. However this data does show a pattern of reduced enhancer activity with the risk allele and this pattern is consistent between enhancer sizes. Although replication is needed, this preliminary data may be reflective of genuine biology and a directional change in rs6733839 risk allele could demonstrate how the effect of risk SNPs may be different in different cell types.

As only a small difference in luciferase activity was observed between alleles, this may also be due to limitations of the gene reporter assay. Small changes in

reporter activity can be due to a number of confounding factors such as plasmid DNA concentration or quality (492). Another factor affecting the accuracy of the assay is the limited availability of regulatory proteins within the cell. During transfection, many copies of the reporter gene can enter the cell. However, as TFs are generally present at low concentrations in the cell, few of these gene copies will encounter all of the necessary proteins required for proper function of the control region (493). This will result in reporters expressing at different levels depending on their encounter with the required proteins (494, 495). In addition, the enhancer activity observed in the reporter assay may not be entirely reflective of endogenous enhancer activity as endogenous chromatin conformations cannot be formed and this may play a vital role, specifically in long-range enhancers.

Ideally this experiment would have been performed in all potential disease relevant cell lines available. Despite efficient transfection in the SV40 microglia cells, as demonstrated by the Renilla readouts, the SV40 microglia line appeared to produce a substrate that reacted with the LARII enzyme to produce luminescence, making this cell line incompatible with this gene reporter system (figure 5.12). The firefly luciferase gene *luc2*, is most commonly used in reporter assays due to its sensitivity and convenience, but alternative reporter genes are available. *Cypridina* luciferase and *Gaussia* luciferase can be used as reporter genes and originate from different organisms and therefore avoid the use of the LARII enzyme and may not produce background luminescence.

Efficient transfection of the THP-1 cells was not achieved using lipofection-based technologies. Efficient transfection of the THP-1 cells has been described using nucleofection, which uses electroporation as a delivery

method (496). As this approach was outside the scope of this project, BV2 cells were used as an alternative, but nucleofection could be utilised in the future.

5.4.3 DNA surrounding rs6733839 binds SPI1 in THP-1 cells

The shift in the EMSA indicated that THP-1 nuclear proteins bind to the DNA sequence containing rs6733839. Protein binding at this region suggests a regulatory function, which is consistent with the ChIP-Seq and gene reporter data (although the gene reporter assay was not performed in this cell line). Quantification of the bands indicated that rs6733839 did not affect overall protein binding (figure 5.19).

The additional shift observed when the SPI1 antibody is present indicates that either SPI1 binds to this locus or SPI1 is present in a protein complex that binds to this locus. This is consistent with the ChIP-Seq data from Pott *et al* (figure 5.6). SPI1, also known as PU.1, is a TF, which acts as a master regulator of myeloid and lymphoid development and is highly expressed in monocytes (497). The SPI1 locus was found to be associated with LOAD in a gene-wide study in 2014 (141) and evidence has since been found that reduced SPI1 expression may decrease AD risk (498). When quantifying the bands, a non-significant 35% reduction of SPI1 bound protein was observed with the risk allele compared to the non-risk allele (p=0.069). Although this result is approaching significance, due to the faintness of the shifted bands, the imaging required over exposure, reducing the accuracy of the quantification process. Therefore to have confidence in this result, further validation would be required (figure 5.19).

Following SPI1 mediated supershifting, a faint band remains, indicating the presence of proteins other than SPI1 bound to the BIN1 risk locus. Further investigation would focus on identifying this protein or proteins. Initial work would focus on MEF-2 TF, which has been shown to have a binding motif close to rs6733839.

Although this EMSA was performed in THP-1 cells, a ChIP-Seq study in the BV2 cell line identified over 5000 SPI1 targeted protein coding genes, one of which was *BIN1*. A pathway analysis of these SPI1 targeted genes identified diverse functional pathways essential for normal monocyte/macrophage function, including endocytosis (499). This finding links *BIN1*, SPI1 and endocytosis to monocyte and macrophage function, which is consistent with observations in this thesis. Therefore all these factors may function in a common biological mechanism which when dysregulated could increase risk for AD.

5.4.4 Future Work

Ideally, the ChIP-Seq, gene reporter and EMSA assays would be performed across of all of the cell lines in order to create a more comprehensive characterisation of the regulatory activity of this locus across multiple cell types. Further assay optimisation would be required for this to be achieved.

As ChIP-Seq data suggests rs6733839 is located within a regulatory genetic element and EMSA data suggests it is located in a SPI1 binding region in the THP-1 cell line, this could justify the use of CRISPR to generate isogenic cellular models with risk or non-risk rs6733839 allele. TF action and regulatory mechanisms can change depending on the state of the cell. A CRISPR model would allow the characterisation of the regulatory changes at the *BIN1* locus

across cell states, such as monocyte differentiation or microglia activation. This would help to determine whether rs6733839 has a significant effect when cells are in a particular state.

Furthermore, CRISPR-mediated epigenome editing could be implemented to investigate the functional significance of the *BIN1* risk locus in the regulation of *BIN1*, or potentially other genes, across multiple cell types and cell states. This technique is capable of targeting specific DNA regulatory elements with proteins that induce transcriptional silencing, such as the KRAB protein that triggers the formation of heterochromatin, or transcriptional activation, such as the HAT domain that acetylates histones (500, 501).

Master regulators, such as SPI1, function in combination with other TFs to create cell type specific regulatory regions (502), suggesting other TFs may have an effect at this locus. MEF-2 is a family of transcriptional regulators, which are encoded by four mammalian genes MEF-2-A, -B, -C and –D. MEF-2C was identified to be significantly associated with LOAD in the 2013 meta-analysis (p=3.2 x 10^{-8}) (140). MEF-2C is a TF that has been implicated in learning and memory, synaptic plasticity and control of synapse number (503-505). Furthermore, MEF-2C has been identified to function in combination with other regulators to create a distinct chromatin landscape in microglia (506).

RegulomeDB V 1.1 describes rs6733839 as in the vicinity of a MEF-2 binding motif as well as SPI1 (436). It could be possible that MEF-2C also binds to this locus, and rs6733839 may affect its binding or function. Therefore, repeating the EMSA across the cell lines using a MEF-2C antibody may show a cell type or allele specific effect on MEF-2C function at the *BIN1* locus.

5.4.5 Concluding remarks

Data from ChIP-Seq studies, gene reporter assays and the EMSA all indicate that the *BIN1* locus is implicated in gene regulatory mechanisms, unfortunately all these techniques could not be replicated in all of the available cell lines of interest. Despite this, this chapter provides evidence for further work at this locus. From these results, rs6733839 does not appear to have a significant impact on gene regulation. Future work would focus on determining whether the regulatory action of this locus differs between cell type and functional state and whether rs6733839 elicits its effect in a particular cell environment. SPI1 binds to the *BIN1* locus and may play a role in *BIN1* regulation. Other AD-associated TFs may also be of interest and their function at this locus could be determined.

6. General Discussion

This thesis investigated GWAS findings in a cellular and molecular context in order to shed light on potential risk mechanisms for AD. A number of approaches have been used which have investigated the impact of global BIN1 and CD2AP depletion on disease relevant cellular processes as well as *BIN1* expression and regulation in relation to specific AD associated DNA variants.

6.1 BIN1 and CD2AP have opposing effects of the processing of APP

Chapter 3 aimed to investigate the impact of BIN1 and CD2AP depletion on APP processing and clathrin mediated endocytosis (CME). Many of the AD associated genes function in the APP processing pathway in some manner. Understanding how these genes interact together and affect Aβ generation will expand the understanding of AD pathogenesis and allude to risk mechanisms.

As BIN1 and CD2AP have both been implicated in endocytosis, they were both individually depleted to ascertain loss of protein function on disease processes. As both proteins are thought to function in the same pathway, depletion of both proteins could produce a greater effect or counter the effect of each other. Therefore BIN1 and CD2AP were depleted in combination to determine the shared or additive effect on biological processes if loss of function was seen in both genes simultaneously. Data suggested that BIN1 and CD2AP depletion had broadly opposing affects on APP processing, which appeared to be independent of CME function.

6.1.1 BIN1 depletion indicates alternative risk mechanisms to CME

BIN1 depletion resulted in an increase in β-CTF generation in the H4 neuroglioma cells. BACE1 levels showed an apparent increase however data was inconsistent at the point of measuring and therefore were not significantly increased. BIN1 depletion in mouse neuroblastoma N2a cells transiently expressing human wild type APP was shown to increase Aβ generation. Although overall BACE1 levels were not measured, defective BACE1 recycling back to the plasma membrane was observed causing an increase in APP and BACE1 convergence in the early endosome, which was thought to contribute to the increased Aβ generation (259). In the N2a cell line, BIN1 depletion was also found to impair BACE1 trafficking from the early endosome to the lysosome, reducing BACE1 degradation and increasing cellular BACE1 levels (quantified via western blotting) (407). The consistent increase in β-CTF observed in BIN1 depleted H4 cells could potentially be explained by an increase in BACE1 activity or cellular BACE1 levels, however BACE1 was not consistently increased. Although the overall levels of BACE1 did not significantly change, the localisation of BACE1 in BIN1 depleted H4s may contribute to the increase in β-CTF levels. The cellular location of BACE1 was not determined during this investigation therefore any effect on BACE1 trafficking remains unknown.

Contradicting data from an investigation that studied gene expression in AD inferior temporal cortex samples showed a positive correlation between BIN1 expression and β-secretase activity, which was quantified with the use of β-secretase-specific fluorogenic substrate peptides conjugated to fluorescent reporter molecules (507). This observation contradicts investigations carried out *in vitro*, this could be potentially explained by the different cell type specific functions and isoforms of BIN1. As BIN1 is thought to interact directly

with BACE1 via the BAR domain (407) present in all BIN1 isoforms, BIN1 could affect BACE1 levels differently in multiple cell types. Investigations using brain tissue do not resolve different cell types and therefore the results may mask cell type specific effects. To understand how BIN1 may be affecting BACE1 in the H4 cell line, the BIN1 isoforms present and BIN1:BACE1 interactions would need to be determined.

BIN1 depletion had no effect on CME in the H4 cells. In neurons, BIN1 depletion resulted in increased CME, as detected by increased transferrin uptake. However, neurons contain the neuron specific isoform 1 of BIN1. The CLAP domain is thought to be important for BIN1 function in CME (233). The exons encoding CLAP domain of BIN1 is encoded by alternatively spliced exons exclusive in the seven brain specific isoforms. Isoform 1 contains all exons encoding the CLAP domain, which was proved to be essential for CME, as overexpression of isoforms lacking the CLAP domain had no effect on CME (253). As BIN1 depletion did not affect CME, it may be that these cells do not have CLAP domain containing BIN1 isoforms. Determining the BIN1 isoforms present in the H4 cell line may help explain the effect on CME observed.

Furthermore, neurons express amphiphysin 1 (AMPH1), a neuronal paralog of BIN1. AMPH1 is predominantly found in the neuronal synapse and is crucial for CME as it plays a major role in the fission of the clathrin coated pit to form a vesicle (508). AMPH1 is also implicated in endocytic recycling (249). It could be that AMPH1 may compensate or exacerbate the functional effects of BIN1 depletion and could contribute to the discrepancies in the functional impact of BIN1 depletion observed between neurons and other cell types. A reduction in AMPH1 has been observed in AD patients and reduced levels have been observed in brain regions with aggregated tau protein (509), suggesting AMPH1 may also be involved in AD pathobiology. The H4 cell line is unlikely to

express AMPH1, therefore it may be that BIN1 depletion in neurons affects AMPH1 which impacts on CME, whereas this affect is not observed in other cell types that do not express AMPH1.

Techniques such as CRISPR screening technologies can now be used to determine all genes involved in a particular biological mechanism. Loss of function CRISPR screening has identified genes essential for cell proliferation, drug resistance and viral infection (510-513). This approach could be used to determine which genes are essential for CME, or other disease processes, in disease relevant cells. This could determine cell type specific effects of AD associated genes or implicate new genes in disease mechanisms not currently investigated in AD.

BIN1 likely affects multiple risk mechanisms for AD in a vast amount of different tissues and cellular processes due to the numerous BIN1 isoforms. BIN1 isoforms differ in their tissue distribution and cellular functions (227, 230, 231, 514, 515). Not only is the BAR domain, capable of inducing membrane curvature, present in all isoforms, but so is the SH3 domain, which is responsible for the interaction with a number proteins involved in endocytosis (233). Exon 7 is an alternatively spliced exon encoding the BAR domain that allows interaction with dynamin 2 (516, 517). Dynamin 2 is ubiquitously expressed and involved in endocytosis, vesicle recycling, membrane fusion and microtubule stability (518). It is therefore possible that the different BIN1 isoforms have the potential to bind different proteins involved in endocytosis and could affect endocytic processes via multiple mechanisms.

In addition to BIN1 function in CME and APP processing, BIN1 has also been implicated in tau pathology. The level of the smaller BIN1 isoform 9, which has elevated expression in AD brains, correlated with the number of neurofibrillary

tangles (NFTs). Isoform 9 is predominantly expressed in astrocytes and this correlation may be due to reactive astrocytosis and be a consequence of neuronal loss (255). If NFT levels increase, this could result in increased neuronal death and could contribute to the shift in BIN1 isoforms levels. Furthermore, if BIN1 depletion results in increased Aβ production causing neuronal loss, the change in astrocyte predominant BIN1 isoforms could also be a consequence of this mechanism.

BIN1 has been shown to physically interact with tau, a correlation between BIN1 expression and tau-mediated neurotoxicity has been observed and BIN1 knockdown reduces tau neurotoxicity (261). Five CpG sites within the BIN1 locus are associated with tau tangle density, potentially implicating epigenetic regulation of BIN1 with AD pathology (260). BIN1 appears to link the microtubule skeleton to the cellular membrane via the tubular membrane structures it forms, which may influence the formation of NFT. BIN1 has also been shown to interact with CLIP170, a microtubule associated protein involved in microtubule stability (519).

Endocytic processes have been implicated in the cellular uptake and subsequent release of pathogenic tau contributing to tau propagation (520-523). Neuronal BIN1 isoforms can inhibit tau propagation whereas the loss of BIN1 function promotes tau propagation in rat neurons. It is proposed that loss of BIN1 increases uptake and tau aggregates in the endosomes ultimately resulting in tau pathology (253). Although this thesis links BIN1 function to Aβ generation, this evidence suggests that it could impact on tau pathology, again suggesting that BIN1 could affect AD risk via multiple cellular mechanisms.

6.1.2 CD2AP depletion implicates BACE1

CD2AP depletion reduced β-CTF and Aβ40 production, which was consistent with other reports in neuroblastoma cells expressing human APP, which showed CD2AP depletion reduced intracellular and secreted Aβ levels (280). CD2AP depletion resulted in decreased BACE1, which most likely contributes to the observed β-CTF and Aβ40 levels. Very little research has been done on the interaction of CD2AP and BACE1 and this may highlight a novel cellular mechanism, which may reduce amyloidogenic processing of CD2AP.

In this investigation, CD2AP depletion did not affect cellular levels of APP but appeared to reduce amyloidogenic processing in the H4 cell line. Contradictory results in neurons show an increase in APP and Aβ production was observed when CD2AP was depleted and this was thought to be due to a decrease in APP degradation (259). CD2AP depletion resulted in increased transferrin levels, suggesting increased internalisation via CME. As increased internalisation has been associated with increased Aβ generation (524), this suggests the mechanism by which CD2AP depletion reduces β-CTF and Aβ40 production is independent of internalisation.

Furthermore, a recent study has discovered a reduction in CD2AP expression in the peripheral blood lymphocytes of LOAD patients in a Chinese Han population, suggesting CD2AP may have a systemic involvement in LOAD (525). CD2AP has long been implicated in the immune response as it is associated with T-lymphocyte marker CD2, which stimulates T cell activation and the formation of the immunological synapse (526). As the immune system is becoming increasing implicated in AD, CD2AP likely contributes to disease risk by having a cell type specific effect on disease processes, including immune cells and neurons. Understanding the effect of CD2AP in different cell

types will give insight into cell type specific functions and how CD2AP may confer risk.

6.1.3 Is BACE1 responsible for changes in β-CTF?

As inhibition of CME is associated with an decrease in Aβ generation (524), this suggests that the effect observed on APP processing in BIN1 and CD2AP depleted cells is independent of CME. As BACE1 levels reflect β-CTF generation, this would suggest BACE1 is responsible.

Levels of BACE1 have been shown to be elevated in AD (527). In a hAPP transgenic mice harbouring the human AD Swedish and Indiana mutations, whose pathology is comparable to human AD, BACE1 accumulated in late endosomes at the presynaptic terminal *in vivo* (528). Recent work has shown that over expression of SNAPIN reduced BACE1 levels in the late endosomes of hAPP mice (529). SNAPIN is exclusively located on synaptic vesicles and interacts with SNARE complex, which has a crucial role in mediated vesicle fusion (530). As this implicates vesicle fusion in BACE1 regulation and BIN1 and CD2AP interact with the plasma membrane, they could affect BACE1 levels via vesicle fusion rather than internalisation, similar to SNAPIN but in other cell types. Therapies targeting BACE1 have been shown to reduce BACE1 levels, reduce amyloid levels and rescue memory loss in APP transgenic rats, indicating BACE1 activity is an active pathological mechanism (531).

6.1.4 BIN1 and CD2AP depletion suggest additional BACE1-independent risk mechanisms involved

Depletion of BIN1 appears to counter the effects of CD2AP depletion on APP processing as no significant changes in APP, β-CTF, Aβ40, sAPPα or sAPPβ

were observed when both proteins are depleted. When BIN1 and CD2AP are both depleted, BACE1 levels are low as seen in the CD2AP only depleted cells. This suggests BIN1 may be rescuing the effects of CD2AP depletion on APP processing via a BACE1 independent mechanism. Previous work has suggested that BIN1 depletion results in increase amyloidogenic processing of APP due to increased BACE1 accumulation (259, 407) whereas the data in this thesis suggests an additional mechanism is involved.

A large-scale siRNA screen attempted to map the regulatory landscape of APP processing. This identified a number of biological pathways that were involved in the regulation of APP processing. This included the "AD pathway" which included APP, γ-secretase complex, β-secretases and additional enzymes known to cleave APP. Additional significant pathways included "Notch receptor processing and trafficking", "membrane protein ectodomain proteolysis" and" Presenilin action in Notch and Wnt signaling". Pathways were also associated with gene transcription, mRNA splicing, protein translation (532). Approximately half of the current genes associated with AD functioned in at least one of these pathways, suggesting these genes could influence APP processing via numerous different mechanisms.

Understanding how AD associated genes affect APP processing is vital in understanding disease risk and pathology. A number of AD associated genes are being investigated for their effect on APP processing. For example, loss of ABCA7 function increases β-secretase activity and Aβ production (533, 534). PICALM depletion reduces β-secretase activity (409). Clearly the function of AD associated genes in APP processing is diverse and understanding how these genes work and interact together will be beneficial to understanding AD biology.

6.2 DNA variants can have cell type specific effects on gene regulation

6.2 1 DNA variants influence the *cis*-regulation of *BIN1* in prefrontal cortex tissue

Although further work remains, investigating the effects of BIN1 depletion chapter 3 establishes a functional role of BIN1 in amyloidogenic processing in brain cells. If reduced expression of BIN1 results in increased amyloidogenic processing, it could be hypothesised that aberrant regulation of *BIN1* in the brain may be a risk mechanism involved in AD. Although *BIN1* levels in AD brains have been investigated and a change in expression has been reported, the direction of this change remains contentious (255, 256, 261). Chapter 4 investigated *cis*-regulation of *BIN1* in prefrontal cortex tissue to determine whether AD associated variants could influence *BIN1* regulation.

Differential allele expression of *BIN1* was observed in prefrontal cortex tissue, implying that *BIN1* is under the influence of *cis*-acting regulatory variation in this tissue. It also implies there is a DNA variant capable of influencing this *cis*-regulation, resulting in different levels of allele expression when heterozygous. AD variants were investigated to determine whether they were likely to be the functional variant. Rs6733839, rs744373 and rs59335482 were not associated with differential allele expression. Rs7584040 was not significantly associated with differential allele expression, but was approaching significance (p=0.062). It could be possible that this intronic variant may reach significance with a greater sample size. Rs7584040 was identified in a conditional analysis (Majounie et al, in prep); therefore if rs7584040 was a functional variant that

affects the *cis*-regulation of *BIN1*, this could be the molecular explanation for this secondary association signal.

Introns can include functional elements, such as intron splice enhancers and silencers that regulate alternative splicing (535, 536), trans-splicing elements (537) and other regulatory elements (538). Intronic polymorphisms can effect splicing by being located within a splice enhancer, branch point or by activating a cryptic splice site (539, 540).

A number of functional intronic SNPs have been found to contribute to susceptibility of complex disease. Two intronic SNPs located in *CD244*, a gene associated with susceptibility to rheumatoid arthritis, have been shown to increase transcriptional activity (541). An intronic variant within *FGFR2*, a gene involved in breast cancer, alters the binding affinity of TF Oct-1/Runx2, leading to increased *FGFR2* expression (542). Furthermore, intronic variants located in *GSK3B*, associated with increased risk for Parkinson's disease, implicated splicing and the risk allele was associated with increased levels of *GSK3N* transcripts lacking exons 9 and 11 (543).

As intronic variants have been shown to influence gene regulation via a number of mechanisms, rs7584040 could influence such regulatory elements in brain tissue. *BIN1* expression in the brain of AD patients remains unclear, but studies in the prefrontal cortex of AD patients revealed BIN1 levels were increased in neurons, but decreased in the neuropil areas (257). This suggests that other cell types present or the axons and dendrites in the neuropil area have different BIN1 expression to neuronal cell bodies, indicating potential cell type specific expression of BIN1. This thesis demonstrates that reduction of BIN1 expression in non-neuronal brain cells increases β-CTF production. As this study suggests a decrease in BIN1 expression in the neuropil areas of AD

prefrontal cortex, this could suggest that non-neuronal brain cells contribute to the production of Aβ, causing neuronal death in the prefrontal cortex. This thesis used an allele specific expression approach to investigate *BIN1* expression in brain tissue; therefore *BIN1* expression in specific cell types was not resolved. Further work could investigate the *cis*-regulation of *BIN1* in specific brain cell types, which may yield more significant results.

6.2.2 rs6733839 risk locus may be functional in immune cell types

Rs6733839 was not found to be the functional variant causing the differential allele expression of *BIN1* observed in prefrontal cortex tissue. However, chapter 5 identified chromatin modifications at the genomic region containing the index SNP, rs6733839, indicative of regulatory elements within monocytes using publically available ChIP-Seq and DNase-Seq data. This justified the *in vitro* study of this locus to investigate its ability to act as an enhancer in a number of disease relevant cell lines. A gene reporter assay in HEK293 cells determined that this locus is capable of acting as an enhancer in an orientation specific manner however the risk allele appears to have no significant affect on enhancer activity. As this locus was shown to have a regulatory function, this further implicates aberrant *BIN1* regulation as a disease risk mechanism. Although this enhancer is likely to function in immune cell types, the gene reporter assay revealed enhancer function in the HEK293 cells. This is possibly due to limitations of the assay, as it does not model the cell type specific open or closed chromatin structure. In the reporter plasmids the enhancer is likely to have an open structure when transfected.

Investigations into DNA:protein binding revealed that this locus binds transcription factor SPI1 in the THP-1 monocyte cell line, further implicating this locus in gene regulation within monocytic cell lines. Although rs6733839

genotype did not significantly affect protein binding in the THP-1 cells, it could be possible that it may have an affect in other immune cell types or specific cell states. These observations therefore provide additional evidence that the rs6733839 risk locus may confer its risk for AD by affecting gene regulatory factors in an immune cell type, however the utilised cell lines act only as a model and the results may differ *in vivo* where other cell types and tissue types are present.

SPI1 is specifically expressed in cells of hematopoietic linage and is crucial for the development of the myeloid and lymphoid lineages, critical for immune function (544-547). Activation of the innate inflammatory response is observed in AD brains and mosaic BIN1 knockout mice show an increase in incidence of inflammation (548, 549). BIN1 functions in phagocytosis in macrophages and regulates the expression of IDO1. IDO1 is an enzyme involved in tryptophan catabolism whose upregulation is required for the defense against parasites and pathogens (550-553). IDO1 expression has also been localised to plaques and tangles in the brain (554). *BIN1* mRNA levels in peripheral blood mononuclear cells and plasma BIN1 expression have been observed to be significantly increased in AD patients (263). Furthermore, a recent study which integrated epigenomic and transcriptomic annotations with heritability data, showed a strong enrichment for LOAD heritability in functional DNA elements related to innate immunity and indicated that monocyte functional elements are particularly relevant (555). If SPI1 binding to the *BIN1* locus in immune cell types is affected by sequence variation, this could affect the efficient regulation of BIN1 expression. In addition to changes in APP processing, this could also have downstream consequences for phagocytic, IDO1 and immune function in myeloid and lymphoid cells, which could potentially act as another risk mechanism for AD.

As shown in this thesis, BIN1 expression can influence APP processing. Non-neuronal cells process APP and generate Aβ in a similar manner to neuronal cells however non-amyloidogenic processing is the more prominent pathway and appears to be γ-secretase dependent (556). A less common splice variant of APP, Leukocyte derived-APP (L-APP), lacks exon 15 and is expressed in leucocytes, microglia and astrocytes (557). L-APP is processed via similar mechanisms to other APP isoforms to produce Aβ and sAPP and therefore may potentially be used to provide peripheral models of APP processing in the brain (558). BIN1 dysregulation is such cell types could therefore potentially affect APP processing outside of the brain.

ENCODE consortium data suggests that only 27% of distal regulatory elements interact with the closest promoter and that the average number of gene targets is 2.5 (313). Thus it could be possible that the *BIN1* risk locus regulates genes other than *BIN1*. Chromatin conformation capture techniques could be implemented to determine the genomic locations this locus interacts with and whether rs6733839 affects this interaction.

This investigation used a number of techniques to characterise the regulatory capacity of the *BIN1* locus, each with their own limitations. Genome wide epigenetic assays, such as ChIP-Seq and DNase-Seq can give insight into the regulatory state of a region, but not the effects of a particular variant, whereas plasmid based gene reporter assays and EMSAs can measure the effect of individual alleles, but are low throughput and can only investigate a single locus at a time. Massive parallel reporter assays have been recently developed to address these limitations. This assay is based upon gene reporter assays but uses unique barcodes in the 3' UTR of the reporter to differentiate expression of individual oligonucleotides and can be used to carry out large scale, sensitive and direct testing of potential regulatory variants. This large scale

approach can be implemented to identify causal variants that influence gene expression across cell types and populations and could contribute to a greater insight into how non-coding polymorphisms can mediate gene expression (559).

6.2.3 Cell type specific nature of disease SNPs

Expression quantitative trail loci (eQTLs) have been shown to function in a cell type specific manner (560, 561). eQTLs are enriched in GWAS data and disease SNPs have been shown to have very cell type specific effects (413). For example, rs612529 associated with inflammatory skin diseases, was shown to have no effect on VSTM1 (V-Set And Transmembrane Domain-Containing Protein 1) expression in B cells but a significant effect in monocytes and this was due to the allele specific binding of SPI1 to the gene promoter. Furthermore, this variant had an allele specific effect on DNA methylation in monocytes, but had no effect on methylation in neuropils, demonstrating how variants can have multiple cell type specific effects on gene regulation (562).

Chapters 4 and 5 suggest rs6733839 and rs7584040 could potentially affect *BIN1* regulation, but the effects of these SNPs are apparent in different cell types. rs7584040 may have a regulatory function in prefrontal cortex, a tissue affected by AD pathology in which changes in *BIN1* expression in AD patients have been observed. rs6733839 may effect *BIN1* regulation in a monocyte cell line as histone modifications and TF binding have been identified in this cell type in this region. This suggests that multiple SNPs may have multiple effects in multiple cell types to confer risk for AD that explain the association identified at the *BIN1* locus. This demonstrates the complexity of this disease and how pleiotropic the genetic associations with AD potentially are.

6.3 Gene editing technologies can be used to investigate the effect of variants in endogenous cellular environments

There are a number of biotechnologies that have the aim of direct genome editing. Such technologies could be utilised to create isogenic models that could provide a platform to elucidate the effects of specific variants in a cellular context.

CRISPR (Clustered regularly interspaced short palindromic repeats)- Cas (CRISPR-associated protein) is an adaptive multistep defense mechanism found in bacteria and archaea. The CRISPR-Cas defense mechanism is able to integrate short foreign DNA sequences, termed spacers, into the host genome at specific locations within the CRISPRs. CRISPR-Cas9 has become the method of choice for gene editing over previous technologies due to its high specificity, efficiency, ease and cost effectiveness.

Characterisation of functional disease associated variants will inevitably exploit gene-editing technologies, which will become commonplace in life science research. Appendix 3 explores this developing technology and how it could be used in the investigation of Alzheimer's disease genetics.

6.4 Why understanding genetic variants is necessary

GWAS have identified many disease associated genetic variants, but relatively few causal variants and the mechanisms by which they infer disease risk. As the majority of GWAS hits are located outside the coding region of genes, it is becoming increasingly apparent that gene regulation is likely the underlying biology of most GWAS associations. The underlying biology and how genetic variants affect gene regulation is poorly understood and this is made

increasingly challenging as the effects on gene regulation can be influenced by development, cell type and environmental factors.

Common diseases are highly polygenic and in reality there are likely many variants that have a small effect on gene expression and the effect of these variants will only be observed in specific cell types. As discussed in this thesis, BIN1 appears to function in many disease relevant biological pathways and *BIN1* expression appears to be influenced by different variants in different tissues.

Understanding genetic associations with disease can not only reveal disease biology, but can point to novel drug targets and help in the design of clinical trials. Understanding the complex mechanisms underlying GWAS associations will allow for the identification of biological pathways that could ultimately be targeted by therapeutics, elucidate the genetic basis of complex disease and open up opportunities for personalised medicine.

6.5 Concluding Remarks

BIN1 and *CD2AP* are both associated with LOAD and have been previously implicated in endocytosis. This thesis found that although they both function in the same biological process, depletion of these two genes has different effects on cellular disease mechanisms. BIN1 depletion results in an increase in β-CTF production independent to any detectable change in CME, whereas CD2AP depletion appears to reduce β-CTF and Aβ production by reducing BACE1 levels. CD2AP depletion increased CME, suggesting the change in APP processing was independent of APP internalisation. Depleting BIN1 and CD2AP in combination provided suggestive evidence that BIN1 can influence APP processing by an unknown BACE1-independent mechanism, as BACE1

levels remained reduced and no change in β-CTF and Aβ levels were observed. Additionally, it was found that CD2AP increased CME regardless of BIN1 depletion.

As BIN1 expression was shown to influence disease processes in brain cells, *BIN1* expression in the prefrontal cortex was investigated. This showed that *BIN1* expression can be influenced by DNA variants but the functional variant is yet to be determined. It appears that rs6733839 is not the functional variant in prefrontal cortex tissue, however ChIP-Seq and DNase-Seq data suggests rs6733839 is located within a regulatory element within monocyte cell lines. This result is supported by gene reporter data and DNA:protein interaction data suggesting SPI1, a transcription factor crucial for the development of the myeloid and lymphoid lineages associated with AD, binds to the *BIN1* locus. The effect of the risk allele remains to be determined. This thesis demonstrates that AD associated genes can effect disease processes via a number of cellular mechanisms and that DNA variants can potentially influence AD risk mechanisms in a cell type specific manner.

Bibliography

1.   Saeed S, Quintin J, Kerstens HH, Rao NA, Aghajanirefah A, Matarese F, et al. Epigenetic programming of monocyte-to-macrophage differentiation and trained innate immunity. Science. 2014;345(6204):1251086.

2.   Selkoe DJ, Hardy J. The amyloid hypothesis of Alzheimer's disease at 25years. Embo Molecular Medicine. 2016;8(6):595-608.

3.   Pai AA, Pritchard JK, Gilad Y. The Genetic and Mechanistic Basis for Variation in Gene Regulation. Plos Genetics. 2015;11(1).

4.   Pott S, Kamrani NK, Bourque G, Pettersson S, Liu ET. PPARG Binding Landscapes in Macrophages Suggest a Genome-Wide Contribution of PU.1 to Divergent PPARG Binding in Human and Mouse. Plos One. 2012;7(10).

5.   Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R, et al. Genetic variability in the regulation of gene expression in ten regions of the human brain. Nature Neuroscience. 2014;17(10):1418-28.

6.   Bernstein BE, Stamatoyannopoulos JA, Costello JF, Ren B, Milosavljevic A, Meissner A, et al. The NIH Roadmap Epigenomics Mapping Consortium. Nature Biotechnology. 2010;28(10):1045-8.

7.   Barker WW, Luis CA, Kashuba A, Luis M, Harwood DG, Loewenstein D, et al. Relative frequencies of Alzheimer disease, Lewy body, vascular and frontotemporal dementia, and hippocampal sclerosis in the state of Florida Brain Bank. Alzheimer Disease & Associated Disorders. 2002;16(4):203-12.

8.   Hebert PL, McBean AM, O'Connor H, Frank B, Good C, Maciejewski ML. Time until incident dementia among Medicare beneficiaries using centrally acting or non-centrally acting ACE inhibitors. Pharmacoepidemiology and Drug Safety. 2013;22(6):641-8.

9.   Luengo-Fernandez R, Leal J, Gray A. UK research spend in 2008 and 2012: comparing stroke, cancer, coronary heart disease and dementia. Bmj Open. 2015;5(4).

10.   Prince M, Knapp M, Guerchet M, McCrone P, Prina M, Comas-Herrera A, et al. Dementia UK: Update Second Edition report
produced by King's College London and the London School of Economics for the Alzheimer's Society. Alzheimer's Society2014.

11.   Alzheimer A. Uber eine eigenartige Erkrankung der Hirnrinde. Allgemeine Zeitschrift fuer Psychiatrie. 1907(64):146-8.

12.   Alzheimer A, Stelzmann RA, Schnitzlein HN, Murtagh FR. An English translation of Alzheimer's 1907 paper, "Uber eine eigenartige Erkankung der Hirnrinde". Clinical anatomy (New York, NY). 1995;8(6):429-31.

13.   Kraepelin E. Handbook of psychiatry, 8th edition. Kraepelin E, editor. Leipzig, Barth 1910.

14.   Bowen DM, Smith CB, White P, Goodhardt MJ, Spillane JA, Flack RHA, et al. Chemical pathology of organic dementias .1. Validity of biochemical measurements on human postmortem brain specimens. Brain. 1977;100(SEP):397-426.

15.   Bowen DM, Smith CB, White P, Flack RHA, Carrasco LH, Gedye JL, et al. Chemical pathology of organic dementias .2. Quantitative estimation of cellular changes in postmortem brainS. Brain. 1977;100(SEP):427-53.

16.   Whitehouse PJ, Price DL, Struble RG, Clark AW, Coyle JT, Delong MR. Alzheimers-disease and senile dementia - loss of neurons in the basal forebrain. Science. 1982;215(4537):1237-9.

17.   Braak H, Braak E. Neuropathological staging of alzheimer-related changes. Acta Neuropathologica. 1991;82(4):239-59.

18.	Sarikaya I. PET imaging in neurology: Alzheimer's and Parkinson's diseases. Nuclear Medicine Communications. 2015;36(8):775-81.

19.	Okamura N, Harada R, Furumoto S, Arai H, Yanai K, Kudo Y. Tau PET Imaging in Alzheimer's Disease. Current Neurology and Neuroscience Reports. 2014;14(11).

20.	Colliot O, Hamelin L, Sarazin M. Magnetic resonance imaging for diagnosis of early Alzheimer's disease. Revue Neurologique. 2013;169(10):724-8.

21.	Holtzman DM, Morris JC, Goate AM. Alzheimer's Disease: The Challenge of the Second Century. Science Translational Medicine. 2011;3(77).

22.	McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical-diagnosis of alzheimers-disease - report of the nincds-adrda work group under the auspices of department-of-health-and-human-services task-force on alzheimers-disease. Neurology. 1984;34(7):939-44.

23.	McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers & Dementia. 2011;7(3):263-9.

24.	Hampel H, Buerger K, Teipel SJ, Bokde ALW, Zetterberg H, Blennow K. Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. Alzheimers & Dementia. 2008;4(1):38-48.

25.	Hasselmo ME. The role of acetylcholine in learning and memory. Current Opinion in Neurobiology. 2006;16(6):710-5.

26.	Augustinsson KB. Comparative aspects of purification and properties of cholinesterases. Bulletin of the World Health Organization. 1971;44(1-3):81-+.

27.	Slotkin TA, Seidler FJ, Crain BJ, Bell JM, Bissette G, Nemeroff CB. Regulatory changes in presynaptic cholinergic function assessed in rapid autopsy material from patients with alzheimer-disease - implications for etiology and therapy. Proceedings of the National Academy of Sciences of the United States of America. 1990;87(7):2452-5.

28.	Xu YQ, Yan JQ, Zhou P, Li JJ, Gao HM, Xia Y, et al. Neurotransmitter receptors and cognitive dysfunction in Alzheimer's disease and Parkinson's disease. Progress in Neurobiology. 2012;97(1):1-13.

29.	Giacobini E. Cholinesterase inhibitors: new roles and therapeutic alternatives. Pharmacological Research. 2004;50(4):433-40.

30.	NICE. Donepezil, galantamine, rivastigmine and memantine for the treatment of Alzheimer's disease. Technology Appraisal Guidance (TA217); 2011.

31.	Olney JW. Excitotoxic amino-acids and neuropsychiatric disorders. Annual Review of Pharmacology and Toxicology. 1990;30:47-71.

32.	Lau A, Tymianski M. Glutamate receptors, neurotoxicity and neurodegeneration. Pflugers Archiv-European Journal of Physiology. 2010;460(2):525-42.

33.	Olney JW. Brain lesions obesity and other disturbances in mice treated with monosodium glutamate. Science. 1969;164(3880):719-&.

34.	Revett TJ, Baker GB, Jhamandas J, Kar S. Glutamate system, amyloid beta peptides and tau protein: functional interrelationships and relevance to Alzheimer disease pathology. Journal of Psychiatry & Neuroscience. 2013;38(1):6-23.

35.	Simons M, Keller P, De Strooper B, Beyreuther K, Dotti CG, Simons K. Cholesterol depletion inhibits the generation of beta-amyloid in hippocampal neurons. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(11):6460-4.

36.	Jarrett JT, Berger EP, Lansbury PT. The carboxy terminus of the beta-amyloid protein is critical for the seeding of amyloid formation - implications for the pathogenesis of alzheimers-disease. Biochemistry. 1993;32(18):4693-7.

37.	Jin M, Shepardson N, Yang T, Chen G, Walsh D, Selkoe DJ. Soluble amyloid beta-protein dimers isolated from Alzheimer cortex directly induce Tau hyperphosphorylation and

neuritic degeneration. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(14):5819-24.

38.   Sharma S, Verma S, Kapoor M, Saini A, Nehru B. Alzheimer's disease like pathology induced six weeks after aggregated amyloid-beta injection in rats: increased oxidative stress and impaired long-term memory with anxiety-like behavior. Neurological Research. 2016;38(9):838-50.

39.   Palmqvist S, Zetterberg H, Mattsson N, Johansson P, Minthon L, Blennow K, et al. Detailed comparison of amyloid PET and CSF biomarkers for identifying early Alzheimer disease. Neurology. 2015;85(14):1240-9.

40.   Oddo S, Caccamo A, Kitazawa M, Tseng BP, LaFerla FM. Amyloid deposition precedes tangle formation in a triple transgenic model of Alzheimer's disease. Neurobiology of Aging. 2003;24(8):1063-70.

41.   Dickson DW. The pathogenesis of senile plaques. Journal of Neuropathology and Experimental Neurology. 1997;56(4):321-39.

42.   Akiyama H, Barger S, Barnum S, Bradt B, Bauer J, Cole GM, et al. Inflammation and Alzheimer's disease. Neurobiology of Aging. 2000;21(3):383-421.

43.   Arriagada PV, Marzloff K, Hyman BT. Distribution of alzheimer-type pathological-changes in nondemented elderly individuals matches the pattern in alzheimers-disease. Neurology. 1992;42(9):1681-8.

44.   Hyman BT, Marzloff K, Arriagada PV. Lack of accumulation of senile plaques or amyloid burden in aging, alzheimers-disease and down-syndrome. Journal of Neuropathology and Experimental Neurology. 1993;52(3):264-.

45.   Giannakopoulos P, Herrmann FR, Bussiere T, Bouras C, Kovari E, Perl DP, et al. Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. Neurology. 2003;60(9):1495-500.

46.   Ingelsson M, Fukumoto H, Newell KL, Growdon JH, Hedley-Whyte ET, Frosch MP, et al. Early A beta accumulation and progressive synaptic loss, gliosis, and tangle formation in AD brain. Neurology. 2004;62(6):925-31.

47.   Itagaki S, McGeer PL, Akiyama H, Zhu S, Selkoe D. Relationship of microglia and astrocytes to amyloid deposits of alzheimer-disease. Journal of Neuroimmunology. 1989;24(3):173-82.

48.   Masliah E, Terry RD, Mallory M, Alford M, Hansen LA. Diffuse plaques do not accentuate synapse loss in alzheimers-disease. American Journal of Pathology. 1990;137(6):1293-7.

49.   Masliah E, Honer WG, Mallory M, Voigt M, Kushner P, Hansen L, et al. Topographical distribution of synaptic-associated proteins in the neuritic plaques of alzheimers-disease hippocampus. Acta Neuropathologica. 1994;87(2):135-42.

50.   Pike CJ, Cummings BJ, Cotman CW. Early association of reactive astrocytes with senile plaques in alzheimers-disease. Experimental Neurology. 1995;132(2):172-9.

51.   Knowles RB, Wyart C, Buldyrev SV, Cruz L, Urbanc B, Hasselmo ME, et al. Plaque-induced neurite abnormalities: Implications for disruption of neural networks in Alzheimer's disease. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(9):5274-9.

52.   Urbanc B, Cruz L, Le R, Sanders J, Ashe KH, Duff K, et al. Neurotoxic effects of thioflavin S-positive amyloid deposits in transgenic mice and Alzheimer's disease. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(22):13990-5.

53.   Vehmas AK, Kawas CH, Stewart WF, Troncoso JC. Immune reactive cells in senile plaques and cognitive decline in Alzheimer's disease. Neurobiology of Aging. 2003;24(2):321-31.

54.   Matsuda S, Yasukawa T, Homma Y, Ito Y, Niikura T, Hiraki T, et al. c-Jun N-terminal kinase (JNK)-interacting protein-1b/islet-brain-1 scaffolds Alzheimer's amyloid precursor protein with JNK. Journal of Neuroscience. 2001;21(17):6597-607.

55.   deToledo-Morrell L, Stoub TR, Bulgakova M, Wilson RS, Bennett DA, Leurgans S, et al. MRI-derived entorhinal volume is a good predictor of conversion from MCI to AD. Neurobiology of Aging. 2004;25(9):1197-203.

56.   Fiala JC. Mechanisms of amyloid plaque pathogenesis. Acta Neuropathologica. 2007;114(6):551-71.

57.   Spires-Jones TL, Hyman BT. The Intersection of Amyloid Beta and Tau at Synapses in Alzheimer's Disease. Neuron. 2014;82(4):756-71.

58.   Wisniewski HM, Narang HK, Terry RD. NEUROFIBRILLARY TANGLES OF PAIRED HELICAL FILAMENTS. Journal of the Neurological Sciences. 1976;27(2):173-81.

59.   Kadavath H, Hofele RV, Biernat J, Kumar S, Tepper K, Urlaub H, et al. Tau stabilizes microtubules by binding at the interface between tubulin heterodimers. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(24):7501-6.

60.   Nakamura K, Greenwood A, Binder L, Bigio EH, Denial S, Nicholson L, et al. Proline Isomer-Specific Antibodies Reveal the Early Pathogenic Tau Conformation in Alzheimer's Disease. Cell. 2012;149(1):232-44.

61.   Braak E, Braak H, Mandelkow EM. A sequence of cytoskeleton changes related to the formation of neurofibrillary tangles and neuropil threads. Acta Neuropathologica. 1994;87(6):554-67.

62.   Augustinack JC, Sanders JL, Tsai LH, Hyman BT. Colocalization and fluorescence resonance energy transfer between cdk5 and AT8 suggests a close association in pre-neurofibrillary tangles and neurofibrillary tangles. Journal of Neuropathology and Experimental Neurology. 2002;61(6):557-64.

63.   Arnold SE, Hyman BT, Flory J, Damasio AR, Van Hoesen GW. The Topographical and Neuroanatomical Distribution of Neurofibrillary Tangles and Neuritic Plaques in the Cerebral Cortex of Patients with Alzheimer's Disease. Cerebral Cortex. 1991;1(1):103-16.

64.   Braak H, Rub U, Schultz C, Del Tredici K. Vulnerability of cortical neurons to Alzheimer's and Parkinson's diseases. Journal of Alzheimers Disease. 2006;9:35-44.

65.   Serrano-Pozo A, Mielke ML, Gomez-Isla T, Betensky RA, Growdon JH, Frosch MP, et al. Reactive Glia not only Associates with Plaques but also Parallels Tangles in Alzheimer's Disease. American Journal of Pathology. 2011;179(3):1373-84.

66.   Dickerson BC, Bakkour A, Salat DH, Feczko E, Pacheco J, Greve DN, et al. The Cortical Signature of Alzheimer's Disease: Regionally Specific Cortical Thinning Relates to Symptom Severity in Very Mild to Mild AD Dementia and is Detectable in Asymptomatic Amyloid-Positive Individuals. Cerebral Cortex. 2009;19(3):497-510.

67.   Spires-Jones TL, De Calignon A, Matsui T, Zehr C, Pitstick R, Wu HY, et al. In vivo imaging reveals dissociation between caspase activation and acute neuronal death in tangle-bearing neurons. Journal of Neuroscience. 2008;28(4):862-7.

68.   Yoshiyama Y, Higuchi M, Zhang B, Huang SM, Iwata N, Saido TC, et al. Synapse loss and microglial activation precede tangles in a P301S tauopathy mouse model. Neuron. 2007;53(3):337-51.

69.   de Calignon A, Fox LM, Pitstick R, Carlson GA, Bacskai BJ, Spires-Jones TL, et al. Caspase activation precedes and leads to tangles. Nature. 2010;464(7292):1201-U123.

70.   Dekosky ST, Scheff SW. Synapse loss in frontal-cortex biopsies in alzheimers-disease - correlation with cognitive severity. Annals of Neurology. 1990;27(5):457-64.

71.   Scheff SW, Price DA. Synapse loss in the temporal-lobe in alzheimers-disease. Annals of Neurology. 1993;33(2):190-9.

72.   Scheff SW, Price DA, Schmitt FA, DeKosky ST, Mufson EJ. Synaptic alterations in CA1 in mild Alzheimer disease and mild cognitive impairment. Neurology. 2007;68(18):1501-8.

73.   Terry RD, Masliah E, Salmon DP, Butters N, Deteresa R, Hill R, et al. Physical basis of cognitive alterations in alzheimers-disease - synapse loss is the major correlate of cognitive impairment. Annals of Neurology. 1991;30(4):572-80.

74.   Ball MJ. Topographic distribution of neurofibrillary tangles, granulovacuoles and hirano bodies in hippocampus of aged and demented patients. Journal of Neuropathology and Experimental Neurology. 1978;37(5):586-.

75.   Xu M, Shibayama H, Kobayashi H, Yamada K, Ishihara R, Zhao P, et al. Granulovacuolar degeneration in the hippocampal cortex of aging and demented patients - a quantitative study. Acta Neuropathologica. 1992;85(1):1-9.

76.   Sisodia SS. Beta-amyloid precursor protein cleavage by a membrane-bound protease. Proceedings of the National Academy of Sciences of the United States of America. 1992;89(13):6075-9.

77.   Buxbaum JD, Liu KN, Luo YX, Slack JL, Stocking KL, Peschon JJ, et al. Evidence that tumor necrosis factor alpha converting enzyme is involved in regulated alpha-secretase cleavage of the Alzheimer amyloid protein precursor. Journal of Biological Chemistry. 1998;273(43):27765-7.

78.   Kuhn PH, Wang H, Dislich B, Colombo A, Zeitschel U, Ellwart JW, et al. ADAM10 is the physiologically relevant, constitutive alpha-secretase of the amyloid precursor protein in primary neurons. Embo Journal. 2010;29(17):3020-32.

79.   Allinson TMJ, Parkin ET, Turner AJ, Hooper NM. ADAMs family members as amyloid precursor protein alpha-secretases. Journal of Neuroscience Research. 2003;74(3):342-52.

80.   Haass C, Hung AY, Schlossmacher MG, Oltersdorf T, Teplow DB, Selkoe DJ. Normal cellular processing of the beta-amyloid precursor protein results in the secretion of the amyloid-beta peptide and related molecules. Alzheimers Disease: Amyloid Precusor Proteins, Signal Transduction, and Neuronal Transplantation. 1993;695:109-16.

81.   Vassar R, Bennett BD, Babu-Khan S, Kahn S, Mendiaz EA, Denis P, et al. beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. Science. 1999;286(5440):735-41.

82.   Walter J, Fluhrer R, Hartung B, Willem M, Kaether C, Capell A, et al. Phosphorylation regulates intracellular trafficking of beta-secretase. Journal of Biological Chemistry. 2001;276(18):14634-41.

83.   Vassar R, Kovacs DM, Yan R, Wong PC. The beta-Secretase Enzyme BACE in Health and Alzheimer's Disease: Regulation, Cell Biology, Function, and Therapeutic Potential. Journal of Neuroscience. 2009;29(41):12787-94.

84.   Tarassishin L, Yin YI, Bassit L, Li YM. Processing of Notch and amyloid precursor protein by gamma-secretase is spatially distinct. Proceedings of the National Academy of Sciences of the United States of America. 2004;101(49):17050-5.

85.   Takahashi RH, Milner TA, Li F, Nam EE, Edgar MA, Yamaguchi H, et al. Intraneuronal Alzheimer A beta 42 accumulates in multivesicular bodies and is associated with synaptic pathology. American Journal of Pathology. 2002;161(5):1869-79.

86.   Kimberly WT, LaVoie MJ, Ostaszewski BL, Ye WJ, Wolfe MS, Selkoe DJ. gamma-secretase is a membrane protein complex comprised of presenilin, nicastrin, aph-1, and pen-2. Proceedings of the National Academy of Sciences of the United States of America. 2003;100(11):6382-7.

87.   Furukawa K, Sopher BL, Rydel RE, Begley JG, Pham DG, Martin GM, et al. Increased activity-regulating and neuroprotective efficacy of alpha-secretase-derived secreted amyloid precursor protein conferred by a C-terminal heparin-binding domain. Journal of Neurochemistry. 1996;67(5):1882-96.

88.   Mattson MP. Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. Physiological Reviews. 1997;77(4):1081-132.

89.   Ring S, Weyer SW, Kilian SB, Waldron E, Pietrzik CU, Filippov MA, et al. The secreted beta-amyloid precursor protein ectodomain APPs alpha is sufficient to rescue the anatomical, behavioral, and electrophysiological abnormalities of APP-deficient mice. Journal of Neuroscience. 2007;27(29):7817-26.

90.   Nikolaev A, McLaughlin T, O'Leary DDM, Tessier-Lavigne M. APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. Nature. 2009;457(7232):981-U1.

91.   Yankner BA, Dawes LR, Fisher S, Villakomaroff L, Ostergranite ML, Neve RL. Neurotoxicity of a fragment of the amyloid precursor associated with alzheimers-disease. Science. 1989;245(4916):417-20.

92.   OsterGranite ML, McPhie DL, Greenan J, Neve RL. Age-dependent neuronal and synaptic degeneration in mice transgenic for the C terminus of the amyloid precursor protein. Journal of Neuroscience. 1996;16(21):6732-41.

93.   Bird TD, Sumi SM, Nemens EJ, Nochlin D, Schellenberg G, Lampe TH, et al. Phenotypic heterogeneity in familial alzheimers-disease - a study of 24 kindreds. Annals of Neurology. 1989;25(1):12-25.

94.   Olson MI, Shaw CM. PRESENILE DEMENTIA AND ALZHEIMERS DISEASE IN MONGOLISM. Brain. 1969;92:147-&.

95.   Stgeorgehyslop PH, Tanzi RE, Polinsky RJ, Haines JL, Nee L, Watkins PC, et al. The genetic-defect causing familial alzheimers-disease maps on chromosome-21. Science. 1987;235(4791):885-90.

96.   Sherrington R, Rogaev EI, Liang Y, Rogaeva EA, Levesque G, Ikeda M, et al. Cloning of a gene bearing missense mutations in early-onset familial alzheimers-disease. Nature. 1995;375(6534):754-60.

97.   Strittmatter WJ, Saunders AM, Schmechel D, Pericakvance M, Enghild J, Salvesen GS, et al. Apolipoprotein-e - high-avidity binding to beta-amyloid and increased frequency of type-4 allele in late-onset familial alzheimer-disease. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(5):1977-81.

98.   Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein-e type-4 allele and the risk of alzheimers-disease in late-onset families. Science. 1993;261(5123):921-3.

99.   Campion D, Dumanchin C, Hannequin D, Dubois B, Belliard S, Puel M, et al. Early-onset autosomal dominant Alzheimer disease: Prevalence, genetic heterogeneity, and mutation spectrum. American Journal of Human Genetics. 1999;65(3):664-70.

100.   Janssen JC, Beck JA, Campbell TA, Dickinson A, Fox NC, Harvey RJ, et al. Early onset familial Alzheimer's disease - Mutation frequency in 31 families. Neurology. 2003;60(2):235-9.

101.   Wu LY, Rosa-Neto P, Hsiung GYR, Sadovnick AD, Masellis M, Black SE, et al. Early-Onset Familial Alzheimer's Disease (EOFAD). Canadian Journal of Neurological Sciences. 2012;39(4):436-45.

102.   Kang J, Lemaire HG, Unterbeck A, Salbaum JM, Masters CL, Grzeschik KH, et al. The precursor of alzheimers-disease amyloid-a4 protein resembles a cell-surface receptor. Nature. 1987;325(6106):733-6.

103.   McCarron M, McCallion P, Reilly E, Mulryan N. A prospective 14-year longitudinal follow-up of dementia in persons with Down syndrome. Journal of Intellectual Disability Research. 2014;58(1):61-70.

104.   Prasher VP, Farrer MJ, Kessling AM, Fisher EMC, West RJ, Barber PC, et al. Molecular mapping of Alzheimer-type dementia in Down's syndrome. Annals of Neurology. 1998;43(3):380-3.

105.   Rovelet-Lecrux A, Hannequin D, Raux G, Le Meur N, Laquerriere A, Vital A, et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. Nature Genetics. 2006;38(1):24-6.

106.   Goate A, Chartierharlin MC, Mullan M, Brown J, Crawford F, Fidani L, et al. Segregation of a missense mutation in the amyloid precursor protein gene with familial alzheimers-disease. Nature. 1991;349(6311):704-6.

107.   Kwok JBJ, Li QX, Hallupp M, Whyte S, Ames D, Beyreuther K, et al. Novel Leu723Pro amyloid precursor protein mutation increases amyloid beta 42(43) peptide levels and induces apoptosis. Annals of Neurology. 2000;47(2):249-53.

108.   Holmes C. Genotype and phenotype in Alzheimer's disease. British Journal of Psychiatry. 2002;180:131-4.

109.	Jonsson T, Atwal JK, Steinberg S, Snaedal J, Jonsson PV, Bjornsson S, et al. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. Nature. 2012;488(7409):96-9.

110.	De Strooper B, Saftig P, Craessaerts K, Vanderstichele H, Guhde G, Annaert W, et al. Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. Nature. 1998;391(6665):387-90.

111.	Wolfe MS, Xia WM, Ostaszewski BL, Diehl TS, Kimberly WT, Selkoe DJ. Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. Nature. 1999;398(6727):513-7.

112.	Laudon H, Hansson EM, Melen K, Bergman A, Farmery MR, Winblad B, et al. A nine-transmembrane domain topology for presenilin 1. Journal of Biological Chemistry. 2005;280(42):35352-60.

113.	Theuns J, Del-Favero J, Dermaut B, van Duijn CM, Backhovens H, Van den Broeck M, et al. Genetic variability in the regulatory region of presenilin 1 associated with risk for Alzheimer's disease and variable expression. Human Molecular Genetics. 2000;9(3):325-31.

114.	Wisniewski T, Dowjat WK, Buxbaum JD, Khorkova O, Efthimiopoulos S, Kulczycki J, et al. A novel Polish presenilin-1 mutation (P117L) is associated with familial Alzheimer's disease and leads to death as early as the age of 28 years. Neuroreport. 1998;9(2):217-21.

115.	Scheuner D, Eckman C, Jensen M, Song X, Citron M, Suzuki N, et al. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased in vivo by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. Nature Medicine. 1996;2(8):864-70.

116.	Lippa CF, Nee LE, Mori H, St George-Hyslop P. A beta-42 deposition precedes other changes in PS-1 Alzheimer's disease. Lancet. 1998;352(9134):1117-8.

117.	Levylahad E, Wasco W, Poorkaj P, Romano DM, Oshima J, Pettingell WH, et al. Candidate gene for the chromosome-1 familial alzheimers-disease locus. Science. 1995;269(5226):973-7.

118.	Bird TD, LevyLahad E, Poorkaj P, Sharma V, Nemens E, Lahad A, et al. Wide range in age of onset for chromosome 1-related familial Alzheimer's disease. Annals of Neurology. 1996;40(6):932-6.

119.	Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, et al. Role of genes and environments for explaining Alzheimer disease. Archives of General Psychiatry. 2006;63(2):168-74.

120.	Silverman JM, Li G, Zaccario ML, Smith CJ, Schmeidler J, Mohs RC, et al. Patterns of risk in first-degree relatives of patients with alzheimers-disease. Archives of General Psychiatry. 1994;51(7):577-86.

121.	Blacker D, Haines JL, Rodes L, Terwedow H, Go RCP, Harrell LE, et al. ApoE-4 and age at onset of Alzheimer's disease: The NIMH genetics initiative. Neurology. 1997;48(1):139-47.

122.	Martins CAR, Oulhaj A, de Jager CA, Williams JH. APOE alleles predict the rate of cognitive decline in Alzheimer disease - A nonlinear model. Neurology. 2005;65(12):1888-93.

123.	Payami H, Montee KR, Kaye JA, Bird TD, Yu CE, Schellenberg GD. Alzheimers-disease, apolipoprotein e4, and gender. Jama-Journal of the American Medical Association. 1994;271(17):1316-7.

124.	Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease - A meta-analysis. Jama-Journal of the American Medical Association. 1997;278(16):1349-56.

125.	Altmann A, Tian L, Henderson VW, Greicius MD, Alzheimer's Dis Neuroimaging I. Sex Modifies the APOE-Related Risk of Developing Alzheimer Disease. Annals of Neurology. 2014;75(4):563-73.

126.	Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC, et al. Protective effect of apolipoprotein-e type-2 allele for late-onset alzheimer-disease. Nature Genetics. 1994;7(2):180-4.

127. Mahley RW, Huang YD, Rall SC. Pathogenesis of type III hyperlipoproteinemia (dysbetalipoproteinemia): questions, quandaries, and paradoxes. Journal of Lipid Research. 1999;40(11):1933-49.

128. Mahley RW, Rall SC. Apolipoprotein E: Far more than a lipid transport protein. Annual Review of Genomics and Human Genetics. 2000;1:507-37.

129. Bu GJ. Apolipoprotein E and its receptors in Alzheimer's disease: pathways, pathogenesis and therapy. Nature Reviews Neuroscience. 2009;10(5):333-44.

130. Rapp A, Gmeiner B, Huttinger M. Implication of apoE isoforms in cholesterol metabolism by primary rat hippocampal neurons and astrocytes. Biochimie. 2006;88(5):473-83.

131. Deane R, Sagare A, Hamm K, Parisi M, Lane S, Finn MB, et al. apoE isoform-specific disruption of amyloid beta peptide clearance from mouse brain. Journal of Clinical Investigation. 2008;118(12):4002-13.

132. Ladu MJ, Pederson TM, Frail DE, Reardon CA, Getz GS, Falduto MT. Purification of apolipoprotein-e attenuates isoform-specific binding to beta-amyloid. Journal of Biological Chemistry. 1995;270(16):9039-42.

133. Nagy Z, Esiri MM, Jobst KA, Johnston C, Litchfield S, Sim E, et al. Influence of the apolipoprotein-e genotype on amyloid deposition and neurofibrillary tangle formation in alzheimers-disease. Neuroscience. 1995;69(3):757-61.

134. Bush WS, Moore JH. Chapter 11: Genome-Wide Association Studies. Plos Computational Biology. 2012;8(12).

135. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nature Genetics. 2009;41(10):1088-U61.

136. Lambert JC, Heath S, Even G, Campion D, Sleegers K, Hiltunen M, et al. Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. Nature Genetics. 2009;41(10):1094-U68.

137. Seshadri S, Fitzpatrick AL, Ikram MA, DeStefano AL, Gudnason V, Boada M, et al. Genome-wide Analysis of Genetic Loci Associated With Alzheimer Disease. Jama-Journal of the American Medical Association. 2010;303(18):1832-40.

138. Hollingworth P, Harold D, Sims R, Gerrish A, Lambert J-C, Carrasquillo MM, et al. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. Nature Genetics. 2011;43(5):429-+.

139. Naj AC, Jun G, Beecham GW, Wang L-S, Vardarajan BN, Buros J, et al. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. Nature Genetics. 2011;43(5):436-+.

140. Lambert J-C, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nature Genetics. 2013;45(12):1452-U206.

141. Escott-Price V, Bellenguez C, Wang LS, Choi SH, Harold D, Jones L, et al. Gene-Wide Analysis Detects Two New Susceptibility Genes for Alzheimer's Disease. Plos One. 2014;9(6).

142. Ruiz A, Heilmann S, Becker T, Hernandez I, Wagner H, Thelen M, et al. Follow-up of loci from the International Genomics of Alzheimer's Disease Project identifies TRIP4 as a novel susceptibility gene. Translational Psychiatry. 2014;4.

143. Cruchaga C, Karch CM, Jin SC, Benitez BA, Cai YF, Guerreiro R, et al. Rare coding variants in the phospholipase D3 gene confer risk for Alzheimer's disease. Nature. 2014;505(7484):550-+.

144. So HC, Gui AHS, Cherny SS, Sham PC. Evaluating the Heritability Explained by Known Susceptibility Variants: A Survey of Ten Complex Diseases. Genetic Epidemiology. 2011;35(5):310-7.

145. Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Alzheimers Dis Genetics C. Alzheimer's Disease: Analyzing the Missing Heritability. Plos One. 2013;8(11).

146. Guerreiro R, Wojtas A, Bras J, Carrasquillo M, Rogaeva E, Majounie E, et al. TREM2 Variants in Alzheimer's Disease. New England Journal of Medicine. 2013;368(2):117-27.

147.   Jonsson T, Stefansson H, Steinberg S, Jonsdottir I, Jonsson PV, Snaedal J, et al. Variant of TREM2 Associated with the Risk of Alzheimer's Disease. New England Journal of Medicine. 2013;368(2):107-16.
148.   Wetzel-Smith MK, Hunkapiller J, Bhangale TR, Srinivasan K, Maloney JA, Atwal JK, et al. A rare mutation in UNC5C predisposes to late-onset Alzheimer's disease and increases neuronal cell death. Nature Medicine. 2014;20(12):1452-7.
149.   Kim M, Suh J, Romano D, Truong MH, Mullin K, Hooli B, et al. Potential late-onset Alzheimer's disease-associated mutations in the ADAM10 gene attenuate alpha-secretase activity. Human Molecular Genetics. 2009;18(20):3987-96.
150.   Kim JH, Song P, Lim H, Lee JH, Park SA, Alzheimer's Dis Neuroimaging I. Gene-Based Rare Allele Analysis Identified a Risk Gene of Alzheimer's Disease. Plos One. 2014;9(10).
151.   Logue MW, Schu M, Vardarajan BN, Farrell J, Bennett DA, Buxbaum JD, et al. Two rare AKAP9 variants are associated with Alzheimer's disease in African Americans. Alzheimers & Dementia. 2014;10(6):609-18.
152.   Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, et al. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. Nat Genet. 2017.
153.   Escott-Price V, Sims R, Bannister C, Harold D, Vronskaya M, Majounie E, et al. Common polygenic variation enhances risk prediction for Alzheimer's disease. Brain. 2015;138:3673-84.
154.   Escott-Price V, Myers AJ, Huentelman M, Hardy J. Polygenic risk score analysis of pathologically confirmed Alzheimer disease. Ann Neurol. 2017;82(2):311-4.
155.   Escott-Price V, Shoai M, Pither R, Williams J, Hardy J. Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. Neurobiology of Aging. 2017;49.
156.   Tosto G, Bird TD, Tsuang D, Bennett DA, Boeve BF, Cruchaga C, et al. Polygenic risk scores in familial Alzheimer disease. Neurology. 2017;88(12):1180-6.
157.   Rodriguez-Rodriguez E, Sanchez-Juan P, Vazquez-Higuera JL, Mateo I, Pozueta A, Berciano J, et al. Genetic risk score predicting accelerated progression from mild cognitive impairment to Alzheimer's disease. Journal of Neural Transmission. 2013;120(5):807-12.
158.   Verhaaren BFJ, Vernooij MW, Koudstaal PJ, Uitterlinden AG, van Duijn CM, Hofman A, et al. Alzheimer's Disease Genes and Cognition in the Nondemented General Population. Biological Psychiatry. 2013;73(5):429-34.
159.   Chauhan G, Adams HHH, Bis JC, Weinstein G, Yu L, Töglhofer AM, et al. Association of Alzheimer's disease GWAS loci with MRI markers of brain aging. Neurobiol Aging. 2015;36(4):1765.e7-.e16.
160.   Foley SF, Tansey KE, Caseras X, Lancaster T, Bracht T, Parker G, et al. Multimodal Brain Imaging Reveals Structural Differences in Alzheimer's Disease Polygenic Risk Carriers: A Study in Healthy Young Adults. Biol Psychiatry. 2017;81(2):154-61.
161.   Lupton MK, Strike L, Hansell NK, Wen W, Mather KA, Armstrong NJ, et al. The effect of increased genetic risk for Alzheimer's disease on hippocampal and amygdala volume. Neurobiol Aging. 2016;40:68-77.
162.   Mormino EC, Sperling RA, Holmes AJ, Buckner RL, De Jager PL, Smoller JW, et al. Polygenic risk of Alzheimer disease is associated with early- and late-life processes. Neurology. 2016;87(5):481-8.
163.   Infante J, Rodriguez-Rodriguez E, Mateo I, Llorca J, Luis Vazquez-Higuera J, Berciano J, et al. Gene-gene interaction between heme oxygenase-1 and liver X receptor-beta and Alzheimer's disease risk. Neurobiology of Aging. 2010;31(4):710-4.
164.   Robson KJH, Lehmann DJ, Wimhurst VLC, Livesey KJ, Combrinck M, Merryweather-Clarke AT, et al. Synergy between the C2 allele of transferrin and the C282Y allele of the haemochromatosis gene (HFE) as risk factors for developing Alzheimer's disease. Journal of Medical Genetics. 2004;41(4):261-5.

165.   Ebbert MTW, Ridge PG, Wilson AR, Sharp AR, Bailey M, Norton MC, et al. Population-based Analysis of Alzheimer's Disease Risk Alleles Implicates Genetic Interactions. Biological Psychiatry. 2014;75(9):732-7.

166.   Ebbert MTW, Boehme KL, Wadsworth ME, Staley LA, Mukherjee S, Crane PK, et al. Interaction between variants in CLU and MS4A4E modulates Alzheimer's disease risk. Alzheimers & Dementia. 2016;12(2):121-9.

167.   Hohman TJ, Koran ME, Thornton-Wells T, Alzheimer's Neuroimaging I. Epistatic Genetic Effects among Alzheimer's Candidate Genes. Plos One. 2013;8(11).

168.   Koran MEI, Hohman TJ, Meda SA, Thornton-Wells TA, Alzheimers Dis Neuroimaging I. Genetic Interactions within Inositol-Related Pathways are Associated with Longitudinal Changes in Ventricle Size. Journal of Alzheimers Disease. 2014;38(1):145-54.

169.   Hardy JA, Higgins GA. Alzheimers-disease - the amyloid cascade hypothesis. Science. 1992;256(5054):184-5.

170.   Selkoe DJ. The molecular pathology of alzheimers-disease. Neuron. 1991;6(4):487-98.

171.   Hardy J. Alzheimer's disease: The amyloid cascade hypothesis: An update and reappraisal. Journal of Alzheimers Disease. 2006;9:151-3.

172.   Nelson PT, Alafuzoff I, Bigio EH, Bouras C, Braak H, Cairns NJ, et al. Correlation of Alzheimer Disease Neuropathologic Changes With Cognitive Status: A Review of the Literature. Journal of Neuropathology and Experimental Neurology. 2012;71(5):362-81.

173.   Bennett DA, Schneider JA, Arvanitakis Z, Kelly JF, Aggarwal NT, Shah RC, et al. Neuropathology of older persons without cognitive impairment from two community-based studies. Neurology. 2006;66(12):1837-44.

174.   Aizenstein HJ, Nebes RD, Saxton JA, Price JC, Mathis CA, Tsopelas ND, et al. Frequent Amyloid Deposition Without Significant Cognitive Impairment Among the Elderly. Archives of Neurology. 2008;65(11):1509-17.

175.   Mullane K, Williams M. Alzheimer's therapeutics: Continued clinical failures question the validity of the amyloid hypothesis-but what lies beyond? Biochemical Pharmacology. 2013;85(3):289-305.

176.   Cummings J. What Can Be Inferred from the Interruption of the Semagacestat Trial for Treatment of Alzheimer's Disease? Biological Psychiatry. 2010;68(10):876-8.

177.   Selkoe DJ. Resolving controversies on the path to Alzheimer's therapeutics. Nature Medicine. 2011;17(9):1060-5.

178.   Hutton M, Lendon CL, Rizzu P, Baker M, Froelich S, Houlden H, et al. Association of missense and 5 '-splice-site mutations in tau with the inherited dementia FTDP-17. Nature. 1998;393(6686):702-5.

179.   Spires-Jones TL, Attems J, Thal DR. Interactions of pathological proteins in neurodegenerative diseases. Acta Neuropathologica. 2017;134(2):187-205.

180.   Jones L, Lambert JC, Wang LS, Choi SH, Harold D, Vedernikov A, et al. Convergent genetic and expression data implicate immunity in Alzheimer's disease. Alzheimers & Dementia. 2015;11(6):658-71.

181.   Nixon RA. Endosome function and dysfunction in Alzheimer's disease and other neurodegenerative diseases. Neurobiology of Aging. 2005;26(3):373-82.

182.   Cataldo AM, Peterhoff CM, Troncosco JC, Gomez-Isla T, Hyman BT, Nixon RA. Endocytic pathway abnormalities precede amyloid beta deposition in sporadic Alzheimer's disease and Down syndrome - Differential effects of APOE genotype and presenilin mutations. American Journal of Pathology. 2000;157(1):277-86.

183.   Cataldo AM, Barnett JL, Pieroni C, Nixon RA. Increased neuronal endocytosis and protease delivery to early endosomes in sporadic Alzheimer's disease: Neuropathologic evidence for a mechanism of increased beta-amyloidogenesis. Journal of Neuroscience. 1997;17(16):6142-51.

184.   Cataldo AM, Petanceska S, Terio NB, Peterhoff CM, Durham R, Mercken M, et al. A beta localization in abnormal endosomes: association with earliest A beta elevations in AD and Down syndrome. Neurobiology of Aging. 2004;25(10):1263-72.

185.   Takahashi RH, Almeida CG, Kearney PF, Yu FM, Lin MT, Milner TA, et al. Oligomerization of Alzheimer's beta-amyloid within processes and synapses of cultured neurons and brain. Journal of Neuroscience. 2004;24(14):3592-9.

186.   Nixon RA. Autophagy in neurodegenerative disease: friend, foe or turncoat? Trends in Neurosciences. 2006;29(9):528-35.

187.   Nixon RA, Yang DS, Lee JH. Neurodegenerative lysosomal disorders - A continuum from development to late age. Autophagy. 2008;4(5):590-9.

188.   Boland B, Kumar A, Lee S, Platt FM, Wegiel J, Yu WH, et al. Autophagy induction and autophagosome clearance in neurons: Relationship to autophagic pathology in Alzheimer's disease. Journal of Neuroscience. 2008;28(27):6926-37.

189.   Koo EH, Squazzo SL, Selkoe DJ, Koo CH. Trafficking of cell-surface amyloid beta-protein precursor .1. Secretion, endocytosis and recycling as detected by labeled monoclonal antibody. Journal of Cell Science. 1996;109:991-8.

190.   Koo EH, Squazzo SL. Evidence that production and release of amyloid beta-protein involves the endocytic pathway. Journal of Biological Chemistry. 1994;269(26):17386-9.

191.   Perez RG, Soriano S, Hayes JD, Ostaszewski B, Xia WM, Selkoe DJ, et al. Mutagenesis identifies new signals for beta-amyloid precursor protein endocytosis, turnover, and the generation of secreted fragments, including A beta 42. Journal of Biological Chemistry. 1999;274(27):18851-6.

192.   Schmid SL. Clathrin-coated vesicle formation and protein sorting: An integrated process. Annual Review of Biochemistry. 1997;66:511-48.

193.   Brodsky FM, Chen CY, Knuehl C, Towler MC, Wakeham DE. Biological basket weaving: Formation and function of clathrin-coated vesicles. Annual Review of Cell and Developmental Biology. 2001;17:517-68.

194.   Collins BM, McCoy AJ, Kent HM, Evans PR, Owen DJ. Molecular architecture and functional model of the endocytic AP2 complex. Cell. 2002;109(4):523-35.

195.   Kirchhausen T, Harrison SC. Protein organization in clathrin trimers. Cell. 1981;23(3):755-61.

196.   Ungewickell E, Branton D. ASSEMBLY UNITS OF CLATHRIN COATS. Nature. 1981;289(5796):420-2.

197.   Cocucci E, Aguet F, Boulant S, Kirchhausen T. The First Five Seconds in the Life of a Clathrin-Coated Pit. Cell. 2012;150(3):495-507.

198.   Hinshaw JE, Schmid SL. Dynamin self-assembles into rings suggesting a mechanism for coated vesicle budding. Nature. 1995;374(6518):190-2.

199.   Zheng J, Cahill SM, Lemmon MA, Fushman D, Schlessinger J, Cowburn D. Identification of the binding site for acidic phospholipids on the PH domain of dynamin: Implications for stimulation of GTPase activity. Journal of Molecular Biology. 1996;255(1):14-21.

200.   Bashkirov PV, Akimov SA, Evseev AI, Schmid SL, Zimmerberg J, Frolov VA. GTPase Cycle of Dynamin Is Coupled to Membrane Squeeze and Release, Leading to Spontaneous Fission. Cell. 2008;135(7):1276-86.

201.   Lee DW, Wu XF, Eisenberg E, Greene LE. Recruitment dynamics of GAK and auxilin to clathrin-coated pits during endocytosis. Journal of Cell Science. 2006;119(17):3502-12.

202.   Xing Y, Bocking T, Wolf M, Grigorieff N, Kirchhausen T, Harrison SC. Structure of clathrin coat with bound Hsc70 and auxilin: mechanism of Hsc70-facilitated disassembly. Embo Journal. 2010;29(3):655-65.

203.   Nandez R, Balkin DM, Messa M, Liang L, Paradise S, Czapla H, et al. A role of ocrl in clathrin-coated pit dynamics and uncoating revealed by studies of lowe syndrome cells. Elife. 2014;3.

204.   Gorvel JP, Chavrier P, Zerial M, Gruenberg J. Rab5 controls early endosome fusion invitro. Cell. 1991;64(5):915-25.

205.   Hutagalung AH, Novick PJ. Role of Rab GTPases in Membrane Traffic and Cell Physiology. Physiological Reviews. 2011;91(1):119-49.

206.   Simonsen A, Lippe R, Christoforidis S, Gaullier JM, Brech A, Callaghan J, et al. EEA1 links PI(3)K function to Rab5 regulation of endosome fusion. Nature. 1998;394(6692):494-8.

207.   Christoforidis S, McBride HM, Burgoyne RD, Zerial M. The Rab5 effector EEA1 is a core component of endosome docking. Nature. 1999;397(6720):621-5.

208.   Scott CC, Vacca F, Gruenberg J. Endosome maturation, transport and functions. Seminars in Cell & Developmental Biology. 2014;31:2-10.

209.   Vandersluijs P, Hull M, Webster P, Male P, Goud B, Mellman I. The small gtp-binding protein rab4 controls an early sorting event on the endocytic pathway. Cell. 1992;70(5):729-40.

210.   Choudhury A, Sharma DK, Marks DL, Pagano RE. Elevated endosomal cholesterol levels in Niemann-Pick cells inhibit Rab4 and perturb membrane recycling. Molecular Biology of the Cell. 2004;15(10):4500-11.

211.   Ren MD, Xu GX, Zeng JB, De Lemos-Chiarandini C, Adesnik M, Sabatini DD. Hydrolysis of GTP on rab11 is required for the direct delivery of transferrin from the pericentriolar recycling compartment to the cell surface but not from sorting endosomes. Proceedings of the National Academy of Sciences of the United States of America. 1998;95(11):6187-92.

212.   McGough IJ, Cullen PJ. Recent Advances in Retromer Biology. Traffic. 2011;12(8):963-71.

213.   Bonifacino JS, Hurley JH. Retromer. Current Opinion in Cell Biology. 2008;20(4):427-36.

214.   Bonifacino JS, Hierro A. Transport according to GARP: receiving retrograde cargo at the trans-Golgi network. Trends in Cell Biology. 2011;21(3):159-67.

215.   Rajendran L, Annaert W. Membrane Trafficking Pathways in Alzheimer's Disease. Traffic. 2012;13(6):759-70.

216.   Rink J, Ghigo E, Kalaidzidis Y, Zerial M. Rab conversion as a mechanism of progression from early to late endosomes. Cell. 2005;122(5):735-49.

217.   Hurley JH. ESCRT complexes and the biogenesis of multivesicular bodies. Current Opinion in Cell Biology. 2008;20(1):4-11.

218.   Williams RL, Urbe S. The emerging shape of the ESCRT machinery. Nature Reviews Molecular Cell Biology. 2007;8(5):355-68.

219.   Woodman PG, Futter CE. Multivesicular bodies: co-ordinated progression to maturity. Current Opinion in Cell Biology. 2008;20(4):408-14.

220.   Russell MRG, Nickerson DP, Odorizzi G. Molecular mechanisms of late endosome morphology, identity and sorting. Current Opinion in Cell Biology. 2006;18(4):422-8.

221.   Ludwig AK, Giebel B. Exosomes: Small vesicles participating in intercellular communication. International Journal of Biochemistry & Cell Biology. 2012;44(1):11-5.

222.   Mathivanan S, Simpson RJ. ExoCarta: A compendium of exosomal proteins and RNA. Proteomics. 2009;9(21):4997-5000.

223.   Chyung JH, Selkoe DJ. Inhibition of receptor-mediated endocytosis demonstrates generation of amyloid beta-protein at the cell surface. Journal of Biological Chemistry. 2003;278(51):51035-43.

224.   Sannerud R, Declerck I, Peric A, Raemaekers T, Menendez G, Zhou LJ, et al. ADP ribosylation factor 6 (ARF6) controls amyloid precursor protein (APP) processing by mediating the endosomal sorting of BACE1. Proceedings of the National Academy of Sciences of the United States of America. 2011;108(34):E559-E68.

225.   Donaldson JG, Porat-Shliom N, Cohen LA. Clathrin-independent endocytosis: A unique platform for cell signaling and PM remodeling. Cellular Signalling. 2009;21(1):1-6.

226.   Ren G, Vajjhala P, Lee JS, Winsor B, Munn AL. The BAR domain proteins: Molding membranes in fission, fusion, and phagy. Microbiology and Molecular Biology Reviews. 2006;70(1):37-+.

227.   Sakamuro D, Elliott KJ, WechslerReya R, Prendergast GC. BIN1 is a novel MYC-interacting protein with features of a tumour suppressor. Nature Genetics. 1996;14(1):69-77.

228.	Frost A, Unger VM, De Camilli P. The BAR Domain Superfamily: Membrane-Molding Macromolecules. Cell. 2009;137(2):191-6.
229.	Peter BJ, Kent HM, Mills IG, Vallis Y, Butler PJG, Evans PR, et al. BAR domains as sensors of membrane curvature: The amphiphysin BAR structure. Science. 2004;303(5657):495-9.
230.	Butler MH, David C, Ochoa GC, Freyberg Z, Daniell L, Grabs D, et al. Amphiphysin II (SH3P9; BIN1), a member of the amphiphysin/Rvs family, is concentrated in the cortical cytomatrix of axon initial segments and nodes of Ranvier in brain and around T tubules in skeletal muscle. Journal of Cell Biology. 1997;137(6):1355-67.
231.	WechslerReya R, Sakamuro D, Zhang J, Duhadaway J, Prendergast GC. Structural analysis of the human BIN1 gene - Evidence for tissue-specific transcriptional regulation and alternate RNA splicing. Journal of Biological Chemistry. 1997;272(50):31453-8.
232.	Lee EY, Marcucci M, Daniell L, Pypaert M, Weisz OA, Ochoa GC, et al. Amphiphysin 2 (Bin1) and T-tubule biogenesis in muscle. Science. 2002;297(5584):1193-6.
233.	Prokic I, Cowling BS, Laporte J. Amphiphysin 2 (BIN1) in physiology and diseases. Journal of Molecular Medicine-Jmm. 2014;92(5):453-63.
234.	Tsutsui K, Maeda Y, Seki S, Tokunaga A. cDNA cloning of a novel amphiphysin isoform and tissue-specific expression of its multiple splice variants. Biochemical and Biophysical Research Communications. 1997;236(1):178-83.
235.	Ramjaun AR, Micheva KD, Bouchelet I, McPherson PS. Identification and characterization of a nerve terminal-enriched amphiphysin isoform. Journal of Biological Chemistry. 1997;272(26):16700-6.
236.	Ramjaun AR, McPherson PS. Multiple amphiphysin II splice variants display differential clathrin binding: Identification of two distinct clathrin-binding sites. Journal of Neurochemistry. 1998;70(6):2369-76.
237.	McMahon HT, Wigge P, Smith C. Clathrin interacts specifically with amphiphysin and is displaced by dynamin. Febs Letters. 1997;413(2):319-22.
238.	Yu HT, Chen JK, Feng SB, Dalgarno DC, Brauer AW, Schreiber SL. Structural basis for the binding of proline-rich peptides to sh3 domains. Cell. 1994;76(5):933-45.
239.	Pineda-Lucena A, Ho CSW, Mao DYL, Sheng Y, Laister RC, Muhandiram R, et al. A structure-based model of the c-Myc/Bin1 protein interaction shows alternative splicing of Bin1 and c-Myc phosphorylation are key binding determinants. Journal of Molecular Biology. 2005;351(1):182-94.
240.	Wigge P, Kohler K, Vallis Y, Doyle CA, Owen D, Hunt SP, et al. Amphiphysin heterodimers: Potential role in clathrin-mediated endocytosis. Molecular Biology of the Cell. 1997;8(10):2003-15.
241.	Bauerfeind R, Takei K, De Camilli P. Amphiphysin I is associated with coated endocytic intermediates and undergoes stimulation-dependent dephosphorylation in nerve terminals. Journal of Biological Chemistry. 1997;272(49):30984-92.
242.	David C, McPherson PS, Mundigl O, DeCamilli P. A role of amphiphysin in synaptic vesicle endocytosis suggested by its binding to dynamin in nerve terminals. Proceedings of the National Academy of Sciences of the United States of America. 1996;93(1):331-5.
243.	Shupliakov O, Low P, Grabs D, Gad H, Chen H, David C, et al. Synaptic vesicle endocytosis impaired by disruption of dynamin-SH3 domain interactions. Science. 1997;276(5310):259-63.
244.	Micheva KD, Ramjaun AR, Kay BK, McPherson PS. SH3 domain-dependent interactions of endophilin with amphiphysin. Febs Letters. 1997;414(2):308-12.
245.	Ringstad N, Nemoto Y, DeCamilli P. The SH3p4/Sh3p8/SH3p13 protein family: Binding partners for synaptojanin and dynamin via a Grb2-like Src homology 3 domain. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(16):8569-74.
246.	Rao YJ, Haucke V. Membrane shaping by the Bin/amphiphysin/Rvs (BAR) domain protein superfamily. Cellular and Molecular Life Sciences. 2011;68(24):3983-93.

247.   Taylor MJ, Perrais D, Merrifield CJ. A High Precision Survey of the Molecular Dynamics of Mammalian Clathrin-Mediated Endocytosis. Plos Biology. 2011;9(3).

248.   Slepnev VI, Ochoa GC, Butler MH, Grabs D, De Camilli P. Role of phosphorylation in regulation of the assembly of endocytic coat complexes. Science. 1998;281(5378):821-4.

249.   Pant S, Sharma M, Patel K, Caplan S, Carr CM, Grant BD. AMPH-1/Amphiphysin/Bin1 functions with RME-1/Ehd1 in endocytic recycling. Nature Cell Biology. 2009;11(12):1399-U38.

250.   Marks B, McMahon HT. Calcium triggers calcineurin-dependent synaptic vesicle recycling in mammalian nerve terminals. Current Biology. 1998;8(13):740-9.

251.   Wigge P, McMahon HT. The amphiphysin family of proteins and their role in endocytosis at the synapse. Trends in Neurosciences. 1998;21(8):339-44.

252.   Picas L, Viaud J, Schauer K, Vanni S, Hnia K, Fraisier V, et al. BIN1/M-Amphiphysin2 induces clustering of phosphoinositides to recruit its downstream partner dynamin. Nature Communications. 2014;5.

253.   Calafate S, Flavin W, Verstreken P, Moechars D. Loss of Bin1 Promotes the Propagation of Tau Pathology. Cell Reports. 2016;17(4):931-40.

254.   Leprince C, Le Scolan E, Meunier B, Fraisier V, Brandon N, De Gunzburg J, et al. Sorting nexin 4 and amphiphysin 2, a new partnership between endocytosis and intracellular trafficking. Journal of Cell Science. 2003;116(10):1937-48.

255.   Holler CJ, Davis PR, Beckett TL, Platt TL, Webb RL, Head E, et al. Bridging Integrator 1 (BIN1) Protein Expression Increases in the Alzheimer's Disease Brain and Correlates with Neurofibrillary Tangle Pathology. Journal of Alzheimers Disease. 2014;42(4):1221-7.

256.   Glennon EBC, Whitehouse IJ, Miners JS, Kehoe PG, Love S, Kellett KAB, et al. BIN1 Is Decreased in Sporadic but Not Familial Alzheimer's Disease or in Aging. Plos One. 2013;8(10).

257.   Adams SL, Tilton K, Kozubek JA, Seshadri S, Delalle I. Subcellular Changes in Bridging Integrator 1 Protein Expression in the Cerebral Cortex During the Progression of Alzheimer Disease Pathology. Journal of Neuropathology and Experimental Neurology. 2016;75(8):779-90.

258.   Bali J, Gheinani AH, Zurbriggen S, Rajendran L. Role of genes linked to sporadic Alzheimer's disease risk in the production of beta-amyloid peptides. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(38):15307-11.

259.   Ubelmann F, Burrinha T, Salavessa L, Gomes R, Ferreira C, Moreno N, et al. Bin1 and CD2AP polarise the endocytic generation of beta-amyloid. Embo Reports. 2017;18(1):102-22.

260.   Yu L, Chibnik LB, Srivastava GP, Pochet N, Yang JY, Xu JS, et al. Association of Brain DNA Methylation in SORL1, ABCA7, HLA-DRB5, SLC24A4, and BIN1 With Pathological Diagnosis of Alzheimer Disease. Jama Neurology. 2015;72(1):15-24.

261.   Chapuis J, Hansmannel F, Gistelinck M, Mounier A, Van Cauwenberghe C, Kolen KV, et al. Increased expression of BIN1 mediates Alzheimer genetic risk by modulating tau pathology. Molecular Psychiatry. 2013;18(11):1225-34.

262.   Karch CM, Jeng AT, Nowotny P, Cady J, Cruchaga C, Goate AM. Expression of Novel Alzheimer's Disease Risk Genes in Control and Alzheimer's Disease Brains. Plos One. 2012;7(11).

263.   Sun L, Tan MS, Hu N, Yu JT, Tan L. Exploring the Value of Plasma BIN1 as a Potential Biomarker for Alzheimer's Disease. Journal of Alzheimers Disease. 2013;37(2):291-5.

264.   Kirsch KH, Georgescu MM, Ishimaru S, Hanfusa H. CMS: An adapter molecule involved in cytoskeletal rearrangements. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(11):6211-6.

265.   Kirsch KH, Georgescu MM, Shishido T, Langdon WY, Birge RB, Hanafusa H. The adapter type protein CMS/CD2AP binds to the proto-oncogenic protein c-Cbl through a tyrosine phosphorylation-regulated Src homology 3 domain interaction. Journal of Biological Chemistry. 2001;276(7):4957-63.

266.   Roldan JLO, Blackledge M, van Nuland NAJ, Azuaga AI. Solution structure, dynamics and thermodynamics of the three SH3 domains of CD2AP. Journal of Biomolecular Nmr. 2011;50(2):103-17.

267.   Dustin ML, Olszowy MW, Holdorf AD, Li J, Bromley S, Desai N, et al. A novel adaptor protein orchestrates receptor patterning and cytoskeletal polarity in T-cell contacts. Cell. 1998;94(5):667-77.

268.   Lynch DK, Winata SC, Lyons RJ, Hughes WE, Lehrbach GM, Wasinger V, et al. A cortactin-CD2-associated protein (CD2AP) complex provides a novel link between epidermal growth factor receptor endocytosis and the actin cytoskeleton. Journal of Biological Chemistry. 2003;278(24):21805-13.

269.   Dikic I. CIN85/CMS family of adaptor molecules. Febs Letters. 2002;529(1):110-5.

270.   Speciale L, Calabrese E, Saresella M, Tinelli C, Mariani C, Sanvito L, et al. Lymphocyte subset patterns and cytokine production in Alzheimer's disease patients. Neurobiology of Aging. 2007;28(8):1163-9.

271.   Pellicano M, Larbi A, Goldeck D, Colonna-Romano G, Buffa S, Bulati M, et al. Immune profiling of Alzheimer patients. Journal of Neuroimmunology. 2012;242(1-2):52-9.

272.   Shaw AS, Dustin ML. Making the T cell receptor go the distance: A topological view of T cell activation. Immunity. 1997;6(4):361-9.

273.   Baum B, Georgiou M. Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. Journal of Cell Biology. 2011;192(6):907-17.

274.   Tang VW, Brieher WM. FSGS3/CD2AP is a barbed-end capping protein that stabilizes actin and strengthens adherens junctions. Journal of Cell Biology. 2013;203(5):815-33.

275.   Cormont M, Meton I, Mari M, Monzo P, Keslair F, Gaskin C, et al. CD2AP/CMS regulates endosome morphology and traffic to the degradative pathway through its interaction with Rab4 and c-Cbl. Traffic. 2003;4(2):97-112.

276.   Kwon SH, Oh S, Nacke M, Mostov KE, Lipschutz JH. Adaptor Protein CD2AP and L-type Lectin LMAN2 Regulate Exosome Cargo Protein Trafficking through the Golgi Complex. Journal of Biological Chemistry. 2016;291(49):25462-75.

277.   Gauthier NC, Monzo P, Gonzalez T, Doye A, Oldani A, Gounon P, et al. Early endosomes associated with dynamic F-actin structures are required for late trafficking of H-pylori VacA toxin. Journal of Cell Biology. 2007;177(2):343-54.

278.   Shulman JM, Chen KW, Keenan BT, Chibnik LB, Fleisher A, Thiyyagura P, et al. Genetic Susceptibility for Alzheimer Disease Neuritic Plaque Pathology. Jama Neurology. 2013;70(9):1150-7.

279.   Treusch S, Hamamichi S, Goodman JL, Matlack KES, Chung CY, Baru V, et al. Functional Links Between A beta Toxicity, Endocytic Trafficking, and Alzheimer's Disease Risk Factors in Yeast. Science. 2011;334(6060):1241-5.

280.   Liao F, Jiang H, Srivatsan S, Xiao Q, Lefton KB, Yamada K, et al. Effects of CD2-associated protein deficiency on amyloid-beta in neuroblastoma cells and in an APP transgenic mouse model. Molecular Neurodegeneration. 2015;10.

281.   Wisniewski HM, Kozlowski PB. Evidence for blood-brain-barrier changes in senile dementia of the alzheimer type (SDAT). Annals of the New York Academy of Sciences. 1982;396(OCT):119-29.

282.   Montagne A, Barnes SR, Sweeney MD, Halliday MR, Sagare AP, Zhao Z, et al. Blood-Brain Barrier Breakdown in the Aging Human Hippocampus. Neuron. 2015;85(2):296-302.

283.   Cochran JN, Rush T, Buckingham SC, Roberson ED. The Alzheimer's disease risk factor CD2AP maintains blood-brain barrier integrity. Human Molecular Genetics. 2015;24(23):6667-74.

284.   Blair LJ, Frauen HD, Zhang B, Nordhues BA, Bijan S, Lin YC, et al. Tau depletion prevents progressive blood-brain barrier damage in a mouse model of tauopathy. Acta Neuropathologica Communications. 2015;3.

285.   Shulman JM, Imboywa S, Giagtzoglou N, Powers MP, Hu YH, Devenport D, et al. Functional screening in Drosophila identifies Alzheimer's disease susceptibility genes and implicates Tau-mediated mechanisms. Human Molecular Genetics. 2014;23(4):870-7.

286.   Lander ES, Int Human Genome Sequencing C, Linton LM, Birren B, Nusbaum C, Zody MC, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860-921.
287.   Heintzman ND, Ren B. Finding distal regulatory elements in the human genome. Current Opinion in Genetics & Development. 2009;19(6):541-9.
288.   Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome. Annual Review of Genomics and Human Genetics. 2006;7:29-59.
289.   Lemon B, Tjian R. Orchestrated response: a symphony of transcription factors for gene control. Genes & Development. 2000;14(20):2551-69.
290.   Juven-Gershon T, Kadonaga JT. Regulation of gene expression via the core promoter and the basal transcriptional machinery. Developmental Biology. 2010;339(2):225-9.
291.   Smale ST, Kadonaga JT. The RNA polymerase II core promoter. Annual Review of Biochemistry. 2003;72:449-79.
292.   Sikorski TW, Buratowski S. The basal initiation machinery: beyond the general transcription factors. Current Opinion in Cell Biology. 2009;21(3):344-51.
293.   Conaway JW, Florens L, Sato S, Tomomori-Sato C, Parmely TJ, Yao T, et al. The mammalian Mediator complex. Febs Letters. 2005;579(4):904-8.
294.   Malik S, Roeder RG. Dynamic regulation of pol II transcription by the mammalian Mediator complex. Trends in Biochemical Sciences. 2005;30(5):256-63.
295.   Levine M, Tjian R. Transcription regulation and animal diversity. Nature. 2003;424(6945):147-51.
296.   Carninci P, Sandelin A, Lenhard B, Katayama S, Shimokawa K, Ponjavic J, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nature Genetics. 2006;38(6):626-35.
297.   Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr ARW, James KD, Turner DJ, et al. Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. Plos Genetics. 2010;6(9).
298.   Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes & Development. 2011;25(10):1010-22.
299.   Bina M, Wyss P, Ren WH, Szpankowski W, Thomas E, Randhawa R, et al. Exploring the characteristics of sequence elements in proximal promoters of human genes. Genomics. 2004;84(6):929-40.
300.   Pares-Matos EI, Milligan JS, Bina M. Exploring transcription factor binding properties of several non-coding DNA sequence elements in the human NF-IL6 gene. Journal of Molecular Biology. 2006;357(3):732-47.
301.   Bina M, Crowely E. Sequence patterns defining the 5 ' boundary of human genes. Biopolymers. 2001;59(5):347-55.
302.   Juven-Gershon T, Hsu JY, Theisen JWM, Kadonaga JT. The RNA polymerase II core promoter - the gateway to transcription. Current Opinion in Cell Biology. 2008;20(3):253-9.
303.   Ing NH. Steroid hormones regulate gene expression posttranscriptionally by altering the stabilities of messenger RNAs. Biology of Reproduction. 2005;72(6):1290-6.
304.   Spitz F, Furlong EEM. Transcription factors: from enhancer binding to developmental control. Nature Reviews Genetics. 2012;13(9):613-26.
305.   Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. Interchromosomal associations between alternatively expressed loci. Nature. 2005;435(7042):637-45.
306.   Lo HS, Wang ZN, Hu Y, Yang HH, Gere S, Buetow KH, et al. Allelic variation in gene expression is common in the human genome. Genome Research. 2003;13(8):1855-62.
307.   He X, Duque T, Sinha S. Evolutionary Origins of Transcription Factor Binding Site Clusters. Molecular Biology and Evolution. 2012;29(3):1059-70.
308.   Nolis IK, McKay DJ, Mantouvalou E, Lomvardas S, Merika M, Thanos D. Transcription factors mediate long-range enhancer-promoter interactions. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(48):20222-7.

309.   Clapier CR, Cairns BR. The Biology of Chromatin Remodeling Complexes. Annual Review of Biochemistry. 2009;78:273-304.

310.   Ong CT, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. Nature Reviews Genetics. 2011;12(4):283-93.

311.   Geyer PK, Green MM, Gorces VG. Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome - the molecular-basis of transvection in drosophila. Embo Journal. 1990;9(7):2247-56.

312.   Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. Interchromosornal interactions and olfactory receptor choice. Cell. 2006;126(2):403-13.

313.   Sanyal A, Lajoie BR, Jain G, Dekker J. The long-range interaction landscape of gene promoters. Nature. 2012;489(7414):109-U27.

314.   Chepelev I, Wei G, Wangsa D, Tang QS, Zhao KJ. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. Cell Research. 2012;22(3):490-503.

315.   Gross DS, Garrard WT. NUCLEASE HYPERSENSITIVE SITES IN CHROMATIN. Annual Review of Biochemistry. 1988;57:159-97.

316.   Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, et al. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. Cell. 2013;153(2):307-19.

317.   Siersbaek R, Rabiee A, Nielsen R, Sidoli S, Traynor S, Loft A, et al. Transcription Factor Cooperativity in Early Adipogenic Hotspots and Super-Enhancers. Cell Reports. 2014;7(5):1443-55.

318.   Adam RC, Yang H, Rockowitz S, Larsen SB, Nikolova M, Oristian DS, et al. Pioneer factors govern super-enhancer dynamics in stem cell plasticity and lineage choice. Nature. 2015;521(7552):366-+.

319.   Shin HY, Willi M, Yoo KH, Zeng XK, Wang CC, Metser G, et al. Hierarchy within the mammary STAT5-driven Wap super-enhancer. Nature Genetics. 2016;48(8):904-+.

320.   Gosselin D, Link VM, Romanoski CE, Fonseca GJ, Eichenfield DZ, Spann NJ, et al. Environment Drives Selection and Function of Enhancers Controlling Tissue-Specific Macrophage Identities. Cell. 2014;159(6):1327-40.

321.   Huang JL, Liu X, Li D, Shao Z, Cao H, Zhang YY, et al. Dynamic Control of Enhancer Repertoires Drives Lineage and Stage-Specific Transcription during Hematopoiesis. Developmental Cell. 2016;36(1):9-23.

322.   Liu CF, Lefebvre V. The transcription factors SOX9 and SOX5/SOX6 cooperate genome-wide through super-enhancers to drive chondrogenesis. Nucleic Acids Research. 2015;43(17):8183-203.

323.   Ohba S, He XJ, Hojo H, McMahon AP. Distinct Transcriptional Programs Underlie Sox9 Regulation of the Mammalian Chondrocyte. Cell Reports. 2015;12(2):229-43.

324.   Vahedi G, Kanno Y, Furumoto Y, Jiang K, Parker SCJ, Erdos MR, et al. Super-enhancers delineate disease-associated regulatory nodes in T cells. Nature. 2015;520(7548):558-+.

325.   Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nature Reviews Genetics. 2011;12(1):7-18.

326.   Richmond TJ, Davey CA. The structure of DNA in the nucleosome core. Nature. 2003;423(6936):145-50.

327.   Li GH, Reinberg D. Chromatin higher-order structures and gene regulation. Current Opinion in Genetics & Development. 2011;21(2):175-86.

328.   Thoma F, Koller T, Klug A. Involvement of histone-h1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. Journal of Cell Biology. 1979;83(2):403-27.

329.   Widom J, Klug A. Structure of the 300a chromatin filament - x-ray-diffraction from oriented samples. Cell. 1985;43(1):207-13.

330.   Myers RM, Stamatoyannopoulos J, Snyder M, Dunham I, Hardison RC, Bernstein BE, et al. A User's Guide to the Encyclopedia of DNA Elements (ENCODE). Plos Biology. 2011;9(4).

331.   Kouzarides T. Chromatin modifications and their function. Cell. 2007;128(4):693-705.

332.   Ruthenburg AJ, Allis CD, Wysocka J. Methylation of lysine 4 on histone H3: Intricacy of writing and reading a single epigenetic mark. Molecular Cell. 2007;25(1):15-30.

333.   Anamika K, Krebs AR, Thompson J, Poch O, Devys D, Tora L. Lessons from genome-wide studies: an integrated definition of the coactivator function of histone acetyl transferases. Epigenetics & Chromatin. 2010;3.

334.   Heintzman ND, Stuart RK, Hon G, Fu YT, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nature Genetics. 2007;39(3):311-8.

335.   Koch CM, Andrews RM, Flicek P, Dillon SC, Karaoz U, Clelland GK, et al. The landscape of histone modifications across 1% of the human genome in five human cell lines. Genome Research. 2007;17(6):691-707.

336.   Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, et al. Methylation of histone H3 Lys 4 in coding regions of active genes. Proceedings of the National Academy of Sciences of the United States of America. 2002;99(13):8695-700.

337.   Barski A, Cuddapah S, Cui KR, Roh TY, Schones DE, Wang ZB, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007;129(4):823-37.

338.   Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. Embo Journal. 2011;30(20):4198-210.

339.   Wang ZB, Zang CZ, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. Nature Genetics. 2008;40(7):897-903.

340.   Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(50):21931-6.

341.   Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC, et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature. 2001;410(6824):120-4.

342.   Mikkelsen TS, Ku MC, Jaffe DB, Issac B, Lieberman E, Giannoukos G, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature. 2007;448(7153):553-U2.

343.   Holliday R, Pugh JE. Dna modification mechanisms and gene activity during development. Science. 1975;187(4173):226-32.

344.   Li E. Chromatin modification and epigenetic reprogramming in mammalian development. Nature Reviews Genetics. 2002;3(9):662-73.

345.   Bernstein BE, Meissner A, Lander ES. The mammalian epigenome. Cell. 2007;128(4):669-81.

346.   Li E, Zhang Y. DNA Methylation in Mammals. Cold Spring Harbor Perspectives in Biology. 2014;6(5).

347.   Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. A fraction of the mouse genome that is derived from islands of nonmethylated, cpg-rich dna. Cell. 1985;40(1):91-9.

348.   Bird AP. CPG-RICH ISLANDS AND THE FUNCTION OF DNA METHYLATION. Nature. 1986;321(6067):209-13.

349.   Meissner A, Mikkelsen TS, Gu HC, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. Nature. 2008;454(7205):766-U91.

350.   Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315-22.

351.   Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, et al. Dynamic DNA methylation across diverse human cell lines and tissues. Genome Research. 2013;23(3):555-67.

352.   Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. Trends in Biochemical Sciences. 2006;31(2):89-97.

353.   Fuks F, Hurd PJ, Wolf D, Nan XS, Bird AP, Kouzarides T. The Methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. Journal of Biological Chemistry. 2003;278(6):4035-40.

354.   Jones PL, Veenstra GJC, Wade PA, Vermaak D, Kass SU, Landsberger N, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics. 1998;19(2):187-91.

355.   Nan XS, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature. 1998;393(6683):386-9.

356.   Wade PA, Gegonne A, Jones PL, Ballestar E, Aubry F, Wolffe AP. Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nature Genetics. 1999;23(1):62-6.

357.   Kokura K, Kaul SC, Wadhwa R, Nomura T, Khan MM, Shinagawa T, et al. The ski protein family is required for MeCP2-mediated transcriptional repression. Journal of Biological Chemistry. 2001;276(36):34115-21.

358.   Campanero MR, Armstrong MI, Flemington EK. CPG methylation as a mechanism for the regulation of E2F activity. Proceedings of the National Academy of Sciences of the United States of America. 2000;97(12):6481-6.

359.   Iguchiariga SMM, Schaffner W. CpG methylation of the camp-responsive enhancer promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. Genes & Development. 1989;3(5):612-9.

360.   Chodavarapu RK, Feng SH, Bernatavichute YV, Chen PY, Stroud H, Yu YC, et al. Relationship between nucleosome positioning and DNA methylation. Nature. 2010;466(7304):388-92.

361.   Kleinjan DJ, van Heyningen V. Position effect in human genetic disease. Human Molecular Genetics. 1998;7(10):1611-8.

362.   Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. Science. 2007;315(5813):848-53.

363.   Gamazon ER, Cox NJ, Davis LK. Structural Architecture of SNP Effects on Complex Traits. American Journal of Human Genetics. 2014;95(5):477-89.

364.   Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A, Waszak SM, et al. Variation in Transcription Factor Binding Among Humans. Science. 2010;328(5975):232-5.

365.   Reddy TE, Gertz J, Pauli F, Kucera KS, Varley KE, Newberry KM, et al. Effects of sequence variation on differential allelic transcription factor occupancy and gene expression. Genome Research. 2012;22(5):860-9.

366.   Sturm RA, Duffy DL, Zhao ZZ, Leite FPN, Stark MS, Hayward NK, et al. A single SNP in an evolutionary conserved region within intron 86 of the HERC2 gene determines human blue-brown eye color. American Journal of Human Genetics. 2008;82(2):424-31.

367.   Duffy DL, Montgomery GW, Chen W, Zhao ZZ, Le L, James MR, et al. A three-single-nucleotide polymorphism haplotype in intron 1 of OCA2 explains most human eye-color variation. American Journal of Human Genetics. 2007;80(2):241-52.

368.   Ohneda K, Ohmori S, Ishijima Y, Nakano M, Yamamoto M. Characterization of a Functional ZBP-89 Binding Site That Mediates Gata1 Gene Expression during Hematopoietic Development. Journal of Biological Chemistry. 2009;284(44):30187-99.

369.   Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. Science. 2007;316(5826):889-94.

370.   Smemo S, Tena JJ, Kim KH, Gamazon ER, Sakabe NJ, Gomez-Marin C, et al. Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature. 2014;507(7492):371-+.
371.   Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. Science. 2012;337(6099):1190-5.
372.   Arnstein P, Taylor DON, Nelsonre.Wa, Huebner RJ, Lennette EH. Propagation of human tumors in antithymocyte serum-treated mice. Journal of the National Cancer Institute. 1974;52(1):71-84.
373.   Janabi N, Peudenier S, Heron B, Ng KH, Tardieu M. Establishment of human microglial cell-lines after transfection of primary cultures of embryonic microglial cells with the sv40 large t-antigen. Neuroscience Letters. 1995;195(2):105-8.
374.   Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a human cell line transformed by dna from human adenovirus type-5. Journal of General Virology. 1977;36(JUL):59-72.
375.   Tsuchiya S, Yamabe M, Yamaguchi Y, Kobayashi Y, Konno T, Tada K. Establishment and characterization of a human acute monocytic leukemia-cell line (THP-1). International Journal of Cancer. 1980;26(2):171-6.
376.   Henn A, Lund S, Hedtjarn M, Schrattenholz A, Porzgen P, Leist M. The Suitability of BV2 Cells as Alternative Model System for Primary Microglia Cultures or for Animal Experiments Examining Brain Inflammation. Altex-Alternativen Zu Tierexperimenten. 2009;26(2):83-94.
377.   Bocchini V, Mazzolla R, Barluzzi R, Blasi E, Sick P, Kettenmann H. An immortalized cell-line expresses properties of activated microglial cells. Journal of Neuroscience Research. 1992;31(4):616-21.
378.   Walter J, Haass C. Posttranslational modifications of amyloid precursor protein : ectodomain phosphorylation and sulfation. Methods in molecular medicine. 2000;32:149-68.
379.   Nordstedt C, Caporaso GL, Thyberg J, Gandy SE, Greengard P. Identification of the alzheimer beta/a4 amyloid precursor protein in clathrin-coated vesicles purified from pc12 cells. Journal of Biological Chemistry. 1993;268(1):608-12.
380.   Huse JT, Doms RW. Neurotoxic traffic: Uncovering the mechanics of amyloid production in Alzheimer's disease. Traffic. 2001;2(2):75-81.
381.   Tienari PJ, Ida N, Ikonen E, Simons M, Weidemann A, Multhaup G, et al. Intracellular and secreted Alzheimer beta-amyloid species are generated by distinct mechanisms in cultured hippocampal neurons. Proceedings of the National Academy of Sciences of the United States of America. 1997;94(8):4125-30.
382.   Maxfield FR, Yamashiro DJ. Endosome acidification and the pathways of receptor-mediated endocytosis. Advances in experimental medicine and biology. 1987;225:189-98.
383.   Demaurex N, Furuya W, D'Souza S, Bonifacino JS, Grinstein S. Mechanism of acidification of the trans-Golgi network (TGN) - In situ measurements of pH using retrieval of TGN38 and furin from the cell surface. Journal of Biological Chemistry. 1998;273(4):2044-51.
384.   Rajendran L, Honsho M, Zahn TR, Keller P, Geiger KD, Verkade P, et al. Alzheimer's disease beta-amyloid peptides are released in association with exosomes. Proceedings of the National Academy of Sciences of the United States of America. 2006;103(30):11172-7.
385.   Kinoshita A, Fukumoto H, Shah T, Whelan CM, Irizarry MC, Hyman BT. Demonstration by FRET of BACE interaction with the amyloid precursor protein at the cell surface and in early endosomes. Journal of Cell Science. 2003;116(16):3339-46.
386.   Small SA, Gandy S. Sorting through the cell biology of Alzheimer's disease: Intracellular pathways to pathogenesis. Neuron. 2006;52(1):15-31.
387.   Lorenzen A, Samosh J, Vandewark K, Anborgh PH, Seah C, Magalhaes AC, et al. Rapid and Direct Transport of Cell Surface APP to the Lysosome defines a novel selective pathway. Molecular Brain. 2010;3.

388.   Chia PZC, Toh WH, Sharples R, Gasnereau I, Hill AF, Gleeson PA. Intracellular Itinerary of Internalised beta-Secretase, BACE1, and Its Potential Impact on beta-Amyloid Peptide Biogenesis. Traffic. 2013;14(9):997-1013.

389.   Vetrivel KS, Thinakaran G. Amyloidogenic processing of beta-amyloid precursor protein in intracellular compartments. Neurology. 2006;66:S69-S73.

390.   Lai A, Sisodia SS, Trowbridge IS. Characterization of sorting signals in the beta-amyloid precursor protein cytoplasmic domain. Journal of Biological Chemistry. 1995;270(8):3565-73.

391.   Selkoe DJ, Yamazaki T, Citron M, Podlisny MB, Koo EH, Teplow DB, et al. The role of APP processing and trafficking pathways in the formation of amyloid beta-protein. Neurobiology of Alzheimer's Disease. 1996;777:57-64.

392.   Haass C, Koo EH, Mellon A, Hung AY, Selkoe DJ. Targeting of cell-surface beta-amyloid precursor protein to lysosomes - alternative processing into amyloid-bearing fragments. Nature. 1992;357(6378):500-3.

393.   Lee EB, Zhang B, Liu KN, Greenbaum EA, Doms RW, Trojanowski JQ, et al. BACE overexpression alters the subcellular processing of APP and inhibits A beta deposition in vivo. Journal of Cell Biology. 2005;168(2):291-302.

394.   Arai H, Lee VMY, Messinger ML, Greenberg BD, Lowery DE, Trojanowski JQ. Expression patterns of beta-amyloid precursor protein (beta-app) in neural and nonneural human tissues from alzheimers-disease and control subjects. Annals of Neurology. 1991;30(5):686-93.

395.   Lee JY, Retamal C, Cuitino L, Caruano-Yzermans A, Shin JE, van Kerkhof P, et al. Adaptor protein sorting nexin 17 regulates amyloid precursor protein trafficking and processing in the early endosomes. Journal of Biological Chemistry. 2008;283(17):11501-8.

396.   Schmidt MR, Haucke V. Recycling endosomes in neuronal membrane traffic. Biology of the Cell. 2007;99(6):333-42.

397.   Yap CC, Winckler B. Harnessing the Power of the Endosome to Regulate Neural Development. Neuron. 2012;74(3):440-51.

398.   Groemer TW, Thiel CS, Holt M, Riedel D, Hua Y, Hueve J, et al. Amyloid Precursor Protein Is Trafficked and Secreted via Synaptic Vesicles. Plos One. 2011;6(4).

399.   Das U, Wang L, Ganguly A, Saikia JM, Wagner SL, Koo EH, et al. Visualizing APP and BACE-1 approximation in neurons yields insight into the amyloidogenic pathway. Nature Neuroscience. 2016;19(1):55-+.

400.   Wu YH, Wu M, He GW, Zhang X, Li WG, Gao Y, et al. Glyceraldehyde-3-phosphate dehydrogenase: A universal internal control for Western blots in prokaryotic and eukaryotic cells. Analytical Biochemistry. 2012;423(1):15-22.

401.   Goldstein JL, Brown MS, Anderson RGW, Russell DW, Schneider WJ. Receptor-mediated endocytosis - concepts emerging from the ldl receptor system. Annual Review of Cell Biology. 1985;1:1-39.

402.   Wang JTH, Teasdale RD, Liebl D. Macropinosome quantitation assay. MethodsX. 2014;1:36-41.

403.   Zhou Y, Hayashi I, Wong J, Tugusheva K, Renger JJ, Zerbinatti C. Intracellular Clusterin Interacts with Brain Isoforms of the Bridging Integrator 1 and with the Microtubule-Associated Protein Tau in Alzheimer's Disease. Plos One. 2014;9(7).

404.   deSilva HAR, Jen A, Wickenden C, Jen LS, Wilkinson SL, Patel AJ. Cell-specific expression of beta-amyloid precursor protein isoform mRNAs and proteins in neurons and astrocytes. Molecular Brain Research. 1997;47(1-2):147-56.

405.   Zhang YW, Thompson R, Zhang H, Xu HX. APP processing in Alzheimer's disease. Molecular Brain. 2011;4.

406.   Swanson J. Fluorescent labeling of endocytic compartments. Methods in Cell Biology. 1989;29:137-51.

407.   Miyagawa T, Ebinuma I, Morohashi Y, Hori Y, Chang MY, Hattori H, et al. BIN1 regulates BACE1 intracellular trafficking and amyloid-beta production. Human Molecular Genetics. 2016;25(14):2948-58.

408.   Thinakaran G, Koo EH. Amyloid Precursor Protein Trafficking, Processing, and Function. Journal of Biological Chemistry. 2008;283(44):29615-9.

409.   Thomas RS, Henson A, Gerrish A, Jones L, Williams J, Kidd EJ. Decreasing the expression of PICALM reduces endocytosis and the activity of beta-secretase: implications for Alzheimer's disease. Bmc Neuroscience. 2016;17.

410.   Pietrzik CU, Busse T, Merriam DE, Weggen S, Koo EH. The cytoplasmic domain of the LDL receptor-related protein regulates multiple steps in APP processing. Embo Journal. 2002;21(21):5691-700.

411.   Muller AJ, Baker JF, DuHadaway JB, Ge K, Farmer G, Donover PS, et al. Targeted disruption of the murine Bin1/Amphiphysin II gene does not disable endocytosis but results in embryonic cardiomyopathy with aberrant myofibril formation. Molecular and Cellular Biology. 2003;23(12):4295-306.

412.   Vardarajan BN, Ghani M, Kahn A, Sheikh S, Sato C, Barral S, et al. Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Annals of Neurology. 2015;78(3):487-98.

413.   Zhu ZH, Zhang FT, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nature Genetics. 2016;48(5):481-+.

414.   Zou FG, Chai HS, Younkin CS, Allen M, Crook J, Pankratz VS, et al. Brain Expression Genome-Wide Association Study (eGWAS) Identifies Human Disease-Associated Variants. Plos Genetics. 2012;8(6).

415.   Gusev A, Lee SH, Trynka G, Finucane H, Vilhjalmsson BJ, Xu H, et al. Partitioning Heritability of Regulatory and Cell-Type-Specific Variants across 11 Common Diseases. American Journal of Human Genetics. 2014;95(5):535-52.

416.   Pant PVK, Tao H, Beilharz EJ, Ballinger DG, Cox DR, Frazer KA. Analysis of allelic differential expression in human white blood cells. Genome Research. 2006;16(3):331-9.

417.   Pollard KS, Serre D, Wang X, Tao H, Grundberg E, Hudson TJ, et al. A genome-wide approach to identifying novel-imprinted genes. Human Genetics. 2008;122(6):625-34.

418.   Bray NJ, O'Donovan MC. Investigating cis-acting regulatory variation using assays of relative allelic expression. Psychiatric Genetics. 2006;16(4):173-7.

419.   Bray NJ, Buckland PR, Owen MJ, O'Donovan MC. Cis-acting variation in the expression of a high proportion of genes in human brain. Human Genetics. 2003;113(2):149-53.

420.   Principles of Frontal Lobe Function. 2013.

421.   Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological Alterations in Alzheimer Disease. Cold Spring Harbor Perspectives in Medicine. 2011;1(1).

422.   De Jager PL, Srivastava G, Lunnon K, Burgess J, Schalkwyk LC, Yu L, et al. Alzheimer's disease: early alterations in brain DNA methylation at ANK1, BIN1, RHBDF2 and other loci. Nature Neuroscience. 2014;17(9):1156-63.

423.   Greenbaum L, Ravona-Springer R, Lubitz I, Schmeidler J, Cooper I, Sano M, et al. Potential contribution of the Alzheimer's disease risk locus BIN1 to episodic memory performance in cognitively normal Type 2 diabetes elderly. European Neuropsychopharmacology. 2016;26(4):787-95.

424.   Li JQ, Wang HF, Zhu XC, Sun FR, Tan MS, Tan CC, et al. GWAS-Linked Loci and Neuroimaging Measures in Alzheimer's Disease. Molecular Neurobiology. 2017;54(1):146-53.

425.   Fennema-Notestine C, Hagler DJ, McEvoy LK, Fleisher AS, Wu EH, Karow DS, et al. Structural MRI Biomarkers for Preclinical and Mild Alzheimer's Disease. Human Brain Mapping. 2009;30(10):3238-53.

426.   Barral S, Bird T, Goate A, Farlow MR, Diaz-Arrastia R, Bennett DA, et al. Genotype patterns at PICALM, CR1, BIN1, CLU, and APOE genes are associated with episodic memory. Neurology. 2012;78(19):1464-71.

427.   Zhang XL, Yu JT, Li J, Wang C, Tan L, Liu B, et al. Bridging Integrator 1 (BIN1) Genotype Effects on Working Memory, Hippocampal Volume, and Functional Connectivity in Young Healthy Individuals. Neuropsychopharmacology. 2015;40(7):1794-803.
428.   Wang HF, Wan Y, Hao XK, Cao L, Zhu XC, Jiang T, et al. Bridging Integrator 1 (BIN1) Genotypes Mediate Alzheimer's Disease Risk by Altering Neuronal Degeneration. Journal of Alzheimers Disease. 2016;52(1):179-90.
429.   Schmidt C, Wolff M, von Ahsen N, Zerr I. Alzheimer's Disease: Genetic Polymorphisms and Rate of Decline. Dementia and Geriatric Cognitive Disorders. 2012;33(2-3):84-9.
430.   Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PIW. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics. 2008;24(24):2938-9.
431.   Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Research. 2012;40(D1):D930-D4.
432.   Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. Genome Research. 2012;22(9):1790-7.
433.   Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nature Genetics. 2013;45(6):580-5.
434.   Duarte RRR, Troakes C, Nolan M, Srivastava DP, Murray RM, Bray NJ. Genome-wide significant schizophrenia risk variation on chromosome 10q24 is associated with altered cis-regulation of BORCS7, AS3MT, and NT5C2 in the human brain. American Journal of Medical Genetics Part B-Neuropsychiatric Genetics. 2016;171(6):806-14.
435.   Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Research. 2001;29(1):308-11.
436.   Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, Pritchard JK. Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Research. 2011;21(3):447-55.
437.   Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis C, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature. 2012;489(7414):57-74.
438.   Ernst J, Kellis M. ChromHMM: automating chromatin-state discovery and characterization. Nature Methods. 2012;9(3):215-6.
439.   Kuleshov V, Xie D, Chen R, Pushkarev D, Ma ZH, Blauwkamp T, et al. Whole-genome haplotyping using long reads and statistical methods. Nature Biotechnology. 2014;32(3):261-6.
440.   Gamazon ER, Badner JA, Cheng L, Zhang C, Zhang D, Cox NJ, et al. Enrichment of cis-regulatory gene expression SNPs and methylation quantitative trait loci among bipolar disorder susceptibility variants. Molecular Psychiatry. 2013;18(3):340-6.
441.   Bray NJ, Preece A, Williams NM, Moskvina V, Buckland PR, Owen MJ, et al. Haplotypes at the dystrobrevin binding protein 1 (DTNBP1) gene locus mediate risk for schizophrenia through reduced DTNBP1 expression. Human Molecular Genetics. 2005;14(14):1947-54.
442.   Jenuwein T, Allis CD. Translating the histone code. Science. 2001;293(5532):1074-80.
443.   McVicker G, van de Geijn B, Degner JF, Cain CE, Banovich NE, Raj A, et al. Identification of Genetic Variants That Affect Histone Modifications in Human Cells. Science. 2013;342(6159):747-9.
444.   Mitchell PJ, Tjian R. TRANSCRIPTIONAL REGULATION IN MAMMALIAN-CELLS BY SEQUENCE-SPECIFIC DNA-BINDING PROTEINS. Science. 1989;245(4916):371-8.
445.   Naar AM, Lemon BD, Tjian R. Transcriptional coactivator complexes. Annual Review of Biochemistry. 2001;70:475-501.
446.   Oldridge DA, Wood AC, Weichert-Leahey N, Crimmins I, Sussman R, Winter C, et al. Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. Nature. 2015;528(7582):418-+.

447.   Wang SZ, Wu SS, Meng QT, Li XB, Zhang JC, Chen R, et al. FAS rs2234767 and rs1800682 polymorphisms jointly contributed to risk of colorectal cancer by affecting SP1/STAT1 complex recruitment to chromatin. Scientific Reports. 2016;6.

448.   Guertin MJ, Martins AL, Siepel A, Lis JT. Accurate Prediction of Inducible Transcription Factor Binding Intensities In Vivo. Plos Genetics. 2012;8(3).

449.   Siggers T, Duyzend MH, Reddy J, Khan S, Bulyk ML. Non-DNA-binding cofactors enhance DNA-binding specificity of a transcriptional regulatory complex. Molecular Systems Biology. 2011;7.

450.   Jolma A, Yan J, Whitington T, Toivonen J, Nitta KR, Rastas P, et al. DNA-Binding Specificities of Human Transcription Factors. Cell. 2013;152(1-2):327-39.

451.   Gordan R, Shen N, Dror I, Zhou T, Horton J, Rohs R, et al. Genomic Regions Flanking E-Box Binding Sites Influence DNA Binding Specificity of bHLH Transcription Factors through DNA Shape. Cell Reports. 2013;3(4):1093-104.

452.   Levo M, Zalckvar E, Sharon E, Machado ACD, Kalma Y, Lotam-Pompan M, et al. Unraveling determinants of transcription factor binding outside the core binding site. Genome Research. 2015;25(7):1018-29.

453.   Yang L, Zhou TY, Dror I, Mathelier A, Wasserman WW, Gordan R, et al. TFBSshape: a motif database for DNA shape features of transcription factor binding sites. Nucleic Acids Research. 2014;42(D1):D148-D55.

454.   Matharu N, Ahituv N. Minor Loops in Major Folds: Enhancer-Promoter Looping, Chromatin Restructuring, and Their Association with Transcriptional Regulation and Disease. Plos Genetics. 2015;11(12).

455.   Visser M, Kayser M, Palstra RJ. HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. Genome Research. 2012;22(3):446-55.

456.   Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, Welch RP, et al. Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. Nature Genetics. 2010;42(7):579-U155.

457.   Kulzer JR, Stitzel ML, Morken MA, Huyghe JR, Fuchsberger C, Kuusisto J, et al. A Common Functional Regulatory Variant at a Type 2 Diabetes Locus Upregulates ARAP1 Expression in the Pancreatic Beta Cell. American Journal of Human Genetics. 2014;94(2):186-97.

458.   Naka I, Hikami K, Nakayama K, Koga M, Nishida N, Kimura R, et al. A functional SNP upstream of the beta-2 adrenergic receptor gene (ADRB2) is associated with obesity in Oceanic populations. International Journal of Obesity. 2013;37(9):1204-10.

459.   Brodie A, Azaria JR, Ofran Y. How far from the SNP may the causative genes be? Nucleic Acids Research. 2016;44(13):6046-54.

460.   Zhang C, Browne A, DiVito JR, Stevenson JA, Romano D, Dong YL, et al. Amyloid-beta Production Via Cleavage of Amyloid-beta Protein Precursor is Modulated by Cell Density. Journal of Alzheimers Disease. 2010;22(2):683-94.

461.   Abisambra JF, Fiorelli T, Padmanabhan J, Neame P, Wefes I, Potter H. LDLR Expression and Localization Are Altered in Mouse and Human Cell Culture Models of Alzheimer's Disease. Plos One. 2010;5(1).

462.   Kim TW, Pettingell WH, Jung YK, Kovacs DM, Tanzi RE. Alternative cleavage of Alzheimer-associated presenilins during apoptosis by a caspase-3 family protease. Science. 1997;277(5324):373-6.

463.   Natunen T, Takalo M, Kemppainen S, Leskela S, Marttinen M, Kurkinen KMA, et al. Relationship between ubiquilin-1 and BACE1 in human Alzheimer's disease and APdE9 transgenic mouse brain and cell-based models. Neurobiology of Disease. 2016;85:187-205.

464.   Auffray C, Fogg D, Garfa M, Elain G, Join-Lambert O, Kayal S, et al. Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. Science. 2007;317(5838):666-70.

465. Auffray C, Sieweke MH, Geissmann F. Blood Monocytes: Development, Heterogeneity, and Relationship with Dendritic Cells. Annual Review of Immunology. 2009;27:669-92.

466. Jakubzick C, Gautier EL, Gibbings SL, Sojka DK, Schlitzer A, Johnson TE, et al. Minimal Differentiation of Classical Monocytes as They Survey Steady-State Tissues and Transport Antigen to Lymph Nodes. Immunity. 2013;39(3):599-610.

467. Tsuchiya S, Kobayashi Y, Goto Y, Okumura H, Nakae S, Konno T, et al. Induction of maturation in cultured human monocytic leukemia-cells by a phorbol diester. Cancer Research. 1982;42(4):1530-6.

468. Hanisch U-K, Kettenmann H. Microglia: active sensor and versatile effector cells in the normal and pathologic brain. Nature Neuroscience. 2007;10(11):1387-94.

469. Block ML, Zecca L, Hong J-S. Microglia-mediated neurotoxicity: uncovering the molecular mechanisms. Nature Reviews Neuroscience. 2007;8(1):57-69.

470. Glass CK, Saijo K, Winner B, Marchetto MC, Gage FH. Mechanisms Underlying Inflammation in Neurodegeneration. Cell. 2010;140(6):918-34.

471. Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, Heravi-Moussavi A, et al. Integrative analysis of 111 reference human epigenomes. Nature. 2015;518(7539):317-30.

472. Gibbs RA, Belmont JW, Hardenbol P, Willis TD, Yu FL, Yang HM, et al. The International HapMap Project. Nature. 2003;426(6968):789-96.

473. Pham TH, Benner C, Lichtinger M, Schwarzfischer L, Hu YH, Andreesen R, et al. Dynamic epigenetic enhancer signatures reveal key transcription factors associated with monocytic differentiation states. Blood. 2012;119(24):E161-E71.

474. Geissmann F. Development of monocytes, macrophages, and dendritic cells (vol 327, pg 656, 2010). Science. 2010;330(6009):1318-.

475. Ginhoux F, Jung S. Monocytes and macrophages: developmental pathways and tissue homeostasis. Nature Reviews Immunology. 2014;14(6):392-404.

476. Das A, Sinha M, Datta S, Abas M, Chaffee S, Sen CK, et al. Monocyte and Macrophage Plasticity in Tissue Repair and Regeneration. American Journal of Pathology. 2015;185(10):2596-606.

477. Wyss-Coray T, Rogers J. Inflammation in Alzheimer Disease-A Brief Review of the Basic Science and Clinical Literature. Cold Spring Harbor Perspectives in Medicine. 2012;2(1).

478. Bornemann KD, Wiederhold KH, Pauli C, Ermini F, Stalder M, Schnell L, et al. A beta-Induce inflammatory processes in microglia cells of APP23 transgenic mice. American Journal of Pathology. 2001;158(1):63-73.

479. Gonzalez-Velasquez FJ, Moss MA. Soluble aggregates of the amyloid-beta protein activate endothelial monolayers for adhesion and subsequent transmigration of monocyte cells. Journal of Neurochemistry. 2008;104(2):500-13.

480. Akiyama H. Inflammation in Alzheimer's disease. Brain Pathology. 2000;10(4):707-8.

481. vanMuiswinkel FL, Veerhuis R, Eikelenboom P. Amyloid beta protein primes cultured rat microglial cells for an enhanced phorbol 12-myristate 13-acetate-induced respiratory burst activity. Journal of Neurochemistry. 1996;66(6):2468-76.

482. Perry VH, Nicoll JAR, Holmes C. Microglia in neurodegenerative disease. Nature Reviews Neurology. 2010;6(4):193-201.

483. Michaud JP, Bellavance MA, Prefontaine P, Rivest S. Real-Time In Vivo Imaging Reveals the Ability of Monocytes to Clear Vascular Amyloid Beta. Cell Reports. 2013;5(3):646-53.

484. Cho MH, Cho K, Kang HJ, Jeon EY, Kim HS, Kwon HJ, et al. Autophagy in microglia degrades extracellular beta-amyloid fibrils and regulates the NLRP3 inflammasome. Autophagy. 2014;10(10):1761-75.

485. Zhang RZ, Miller RG, Madison C, Jin X, Honrada R, Harris W, et al. Systemic immune system alterations in early stages of Alzheimer's disease. Journal of Neuroimmunology. 2013;256(1-2):38-42.

486. Cunningham C. Microglia and neurodegeneration: The role of systemic inflammation. Glia. 2013;61(1):71-90.

487.   Gjoneska E, Pfenning AR, Mathys H, Quon G, Kundaje A, Tsai LH, et al. Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease. Nature. 2015;518(7539):365-9.

488.   Gagliano SA, Pouget JG, Hardy J, Knight J, Barnes MR, Ryten M, et al. Genomics implicates adaptive and innate immunity in Alzheimer's and Parkinson's diseases. Annals of Clinical and Translational Neurology. 2016;3(12):924-33.

489.   Guo Y, Xu Q, Canzio D, Shou J, Li J, Gorkin DU, et al. CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. Cell. 2015;162(4):900-10.

490.   Lobanenkov VV, Nicolas RH, Adler VV, Paterson H, Klenova EM, Polotskaja AV, et al. A novel sequence-specific dna-binding protein which interacts with 3 regularly spaced direct repeats of the ccctc-motif in the 5'-flanking sequence of the chicken c-myc gene. Oncogene. 1990;5(12):1743-53.

491.   Ong C-T, Corces VG. CTCF: an architectural protein bridging genome topology and function. Nature Reviews Genetics. 2014;15(4):234-46.

492.   Butash KA, Natarajan P, Young A, Fox DK. Reexamination of the effect of endotoxin on cell proliferation and transfection efficiency. Biotechniques. 2000;29(3):610-+.

493.   Carey M, Smale ST. Transcriptional regulation in eukaryotes: Concepts, strategies, and techniques. Transcriptional regulation in eukaryotes: Concepts, strategies, and techniques. 2000:i-xxix, 1-640.

494.   Smith CL, Hager GL. Transcriptional regulation of mammalian genes in vivo - A tale of two templates. Journal of Biological Chemistry. 1997;272(44):27493-6.

495.   Mercola M, Goverman J, Mirell C, Calame K. Immunoglobulin heavy-chain enhancer requires one or more tissue-specific factors. Science. 1985;227(4684):266-70.

496.   Schnoor M, Buers I, Sietmann A, Brodde MF, Hofnagel O, Robenek H, et al. Efficient non-viral transfection of THP-1 cells. Journal of Immunological Methods. 2009;344(2):109-15.

497.   Turkistany SA, DeKoter RP. The Transcription Factor PU.1 is a Critical Regulator of Cellular Communication in the Immune System. Archivum Immunologiae Et Therapiae Experimentalis. 2011;59(6):431-40.

498.   Huang K-L, Marcora E, Pimenova AA, Di Narzo AF, Kapoor M, Jin SC, et al. A common haplotype lowers PU.1 expression in myeloid cells and delays onset of Alzheimer's disease. Nature neuroscience. 2017.

499.   Satoh J-I, Asahina N, Kitano S, Kino Y. A Comprehensive Profile of ChIP-Seq-Based PU.1/Spi1 Target Genes in Microglia. Gene regulation and systems biology. 2014;8:127-39.

500.   Hilton IB, D'Ippolito AM, Vockley CM, Thakore PI, Crawford GE, Reddy TE, et al. Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. Nature Biotechnology. 2015;33(5):510-U225.

501.   Thakore PI, D'Ippolito AM, Song LY, Safi A, Shivakumar NK, Kabadi AM, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nature Methods. 2015;12(12):1143-+.

502.   Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. Molecular Cell. 2010;38(4):576-89.

503.   Barbosa AC, Kim M-S, Ertunc M, Adachi M, Nelson ED, McAnally J, et al. MEF2C, a transcription factor that facilitates learning and memory by negative regulation of synapse numbers and function. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(27):9391-6.

504.   Li H, Radford JC, Ragusa MJ, Shea KL, McKercher SR, Zaremba JD, et al. Transcription factor MEF2C influences neural stem/progenitor cell differentiation and maturation in vivo. Proceedings of the National Academy of Sciences of the United States of America. 2008;105(27):9397-402.

505.   Martin JF, Schwarz JJ, Olson EN. Myocyte enhancer factor (mef) 2c - a tissue-restricted member of the mef-2 family of transcription factors. Proceedings of the National Academy of Sciences of the United States of America. 1993;90(11):5282-6.

506.   Lavin Y, Winter D, Blecher-Gonen R, David E, Keren-Shaul H, Merad M, et al. Tissue-Resident Macrophage Enhancer Landscapes Are Shaped by the Local Microenvironment. Cell. 2014;159(6):1312-26.

507.   Martiskainen H, Viswanathan J, Nykanen NP, Kurki M, Helisalmi S, Natunen T, et al. Transcriptomics and mechanistic elucidation of Alzheimer's disease risk genes in the brain and in vitro models. Neurobiology of Aging. 2015;36(2).

508.   Takei K, Slepnev VI, Haucke V, De Camilli P. Functional partnership between amphiphysin and dynamin in clathrin-mediated endocytosis. Nature Cell Biology. 1999;1(1):33-9.

509.   De Jesus-Cortes HJ, Nogueras-Ortiz CJ, Gearing M, Arnold SE, Vega IE. Amphiphysin-1 protein level changes associated with tau-mediated neurodegeneration. Neuroreport. 2012;23(16):942-6.

510.   Blomen VA, Majek P, Jae LT, Bigenzahn JW, Nieuwenhuis J, Staring J, et al. Gene essentiality and synthetic lethality in haploid human cells. Science. 2015;350(6264):1092-6.

511.   Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, et al. Identification and characterization of essential genes in the human genome. Science. 2015;350(6264):1096-101.

512.   Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, et al. Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science. 2014;343(6166):84-7.

513.   Marceau CD, Puschnik AS, Majzoub K, Ooi YS, Brewer SM, Fuchs G, et al. Genetic dissection of Flaviviridae host factors through genome-scale CRISPR screens. Nature. 2016;535(7610):159-+.

514.   Leprince C, Romero F, Cussac D, Vayssiere B, Berger R, Tavitian A, et al. A new member of the amphiphysin family connecting endocytosis and signal transduction pathways. Journal of Biological Chemistry. 1997;272(24):15101-5.

515.   DuHadaway JB, Lynch FJ, Brisbay S, Bueso-Ramos C, Troncoso P, McDonnell T, et al. Immunohistochemical analysis of Bin1/Amphiphysin II in human tissues: Diverse sites of nuclear expression and losses in prostate cancer. Journal of Cellular Biochemistry. 2003;88(3):635-42.

516.   Nicot AS, Toussaint A, Tosch V, Kretz C, Wallgren-Pettersson C, Iwarsson E, et al. Mutations in amphiphysin 2 (BIN1) disrupt interaction with dynamin 2 and cause autosomal recessive centronuclear myopathy. Nature Genetics. 2007;39(9):1134-9.

517.   Ellis JD, Barrios-Rodiles M, Colak R, Irimia M, Kim T, Calarco JA, et al. Tissue-Specific Alternative Splicing Remodels Protein-Protein Interaction Networks. Molecular Cell. 2012;46(6):884-92.

518.   González-Jamett AM, Momboisse F, Haro-Acuña V, Bevilacqua JA, Caviedes P, Cárdenas AM. Dynamin-2 function and dysfunction along the secretory pathway. Front Endocrinol (Lausanne). 2013;4:126.

519.   Meunier B, Quaranta M, Daviet L, Hatzoglou A, Leprince C. The membrane-tubulating potential of amphiphysin 2/BIN1 is dependent on the microtubule-binding cytoplasmic linker protein 170 (CLIP-170). European Journal of Cell Biology. 2009;88(2):91-102.

520.   Frost B, Jacks RL, Diamond MI. Propagation of Tau Misfolding from the Outside to the Inside of a Cell. Journal of Biological Chemistry. 2009;284(19):12845-52.

521.   Holmes BB, Devos SL, Kfoury N, Li M, Jacks R, Yanamandra K, et al. Heparan sulfate proteoglycans mediate internalization and propagation of specific proteopathic seeds. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(33):E3138-E47.

522.   Santa-Maria I, Varghese M, Ksiezak-Reding H, Dzhun A, Wang J, Pasinetti GM. Paired Helical Filaments from Alzheimer Disease Brain Induce Intracellular Accumulation of Tau Protein in Aggresomes. Journal of Biological Chemistry. 2012;287(24):20522-33.

523.   Wu JW, Herman M, Liu L, Simoes S, Acker CM, Figueroa H, et al. Small Misfolded Tau Species Are Internalized via Bulk Endocytosis and Anterogradely and Retrogradely Transported in Neurons. Journal of Biological Chemistry. 2013;288(3):1856-70.

524.   Cirrito JR, Kang JE, Lee J, Stewart FR, Verges DK, Silverio LM, et al. Endocytosis is required for synaptic activity-dependent release of amyloid-beta in vivo. Neuron. 2008;58(1):42-51.

525.   Tao QQ, Liu ZJ, Sun YM, Li HL, Yang P, Liu DS, et al. Decreased gene expression of CD2AP in Chinese patients with sporadic Alzheimer's disease. Neurobiology of Aging. 2017;56.

526.   Rabin EM, Gordon K, Knoppers MH, Luther MA, Neidhardt EA, Flynn JF, et al. Inhibition of t-cell activation and adhesion functions by soluble cd2 protein. Cellular Immunology. 1993;149(1):24-38.

527.   Yang LB, Lindholm K, Yan RQ, Citron M, Xia WM, Yang XL, et al. Elevated beta-secretase expression and enzymatic activity detected in sporadic Alzheimer disease. Nature Medicine. 2003;9(1):3-4.

528.   Ye X, Cai Q. Snapin-Mediated BACE1 Retrograde Transport Is Essential for Its Degradation in Lysosomes and Regulation of APP Processing in Neurons. Cell Reports. 2014;6(1):24-31.

529.   Ye X, Feng TC, Tammineni P, Chang Q, Jeong YY, Margolis DJ, et al. Regulation of Synaptic Amyloid-beta Generation through BACE1 Retrograde Transport in a Mouse Model of Alzheimer's Disease. Journal of Neuroscience. 2017;37(10):2639-55.

530.   Ilardi JM, Mochida S, Sheng ZH. Snapin: a SNARE-associated protein implicated in synaptic transmission. Nature Neuroscience. 1999;2(2):119-24.

531.   Wilson EN, Do Carmo S, Iulita MF, Hall H, Ducatenzeiler A, Marks AR, et al. BACE1 inhibition by microdose lithium formulation NP03 rescues memory loss and early stage amyloid neuropathology. Transl Psychiatry. 2017;7(8):e1190.

532.   Camargo LM, Zhang XD, Loerch P, Caceres RM, Marine SD, Uva P, et al. Pathway-Based Analysis of Genome-Wide siRNA Screens Reveals the Regulatory Landscape of App Processing. Plos One. 2015;10(2).

533.   Sakae N, Liu CC, Shinohara M, Frisch-Daiello J, Ma L, Yamazaki Y, et al. ABCA7 Deficiency Accelerates Amyloid-beta Generation and Alzheimer's Neuronal Pathology. Journal of Neuroscience. 2016;36(13):3848-59.

534.   Satoh K, Abe-Dohmae S, Yokoyama S, St George-Hyslop P, Fraser PE. ATP-binding Cassette Transporter A7 (ABCA7) Loss of Function Alters Alzheimer Amyloid Processing. Journal of Biological Chemistry. 2015;290(40):24152-65.

535.   Wang X, Wang KJ, Radovich M, Wang Y, Wang GH, Feng WX, et al. Genome-wide prediction of cis-acting RNA elements regulating tissue-specific pre-mRNA alternative splicing. Bmc Genomics. 2009;10.

536.   Tress ML, Martelli PL, Frankish A, Reeves GA, Wesselink JJ, Yeats C, et al. The implications of alternative splicing in the ENCODE protein complement. Proceedings of the National Academy of Sciences of the United States of America. 2007;104(13):5495-500.

537.   Gingeras TR. Implications of chimaeric non-co-linear transcripts. Nature. 2009;461(7261):206-11.

538.   Solis AS, Shariat N, Patton JG. Splicing fidelity, enhancers, and disease. Frontiers in Bioscience-Landmark. 2008;13:1926-42.

539.   Kwan T, Benovoy D, Dias C, Gurd S, Provencher C, Beaulieu P, et al. Genome-wide analysis of transcript isoform variation in humans. Nature Genetics. 2008;40(2):225-31.

540.   Coulombe-Huntington J, Lam KCL, Dias C, Majewski J. Fine-Scale Variation and Genetic Determinants of Alternative Splicing across Individuals. Plos Genetics. 2009;5(12).

541.   Suzuki A, Yamada R, Kochi Y, Sawada T, Okada Y, Matsuda K, et al. Functional SNPs in CD244 increase the risk of rheumatoid arthritis in a Japanese population. Nature Genetics. 2008;40(10):1224-9.

542.   Meyer KB, Maia AT, O'Reilly M, Teschendorff AE, Chin SF, Caldas C, et al. Allele-specific up-regulation of FGFR2 increases susceptibility to breast cancer. Plos Biology. 2008;6(5):1098-103.

543.   Kwok JBJ, Hallupp M, Loy CT, Chan DKY, Woo J, Mellick GD, et al. GSK3B polymorphisms alter transcription and splicing in Parkinson's disease. Annals of Neurology. 2005;58(6):829-39.

544.   Klemsz MJ, McKercher SR, Celada A, Vanbeveren C, Maki RA. The macrophage and b-cell specific transcription factor pu.1 is related to the ets oncogene. Cell. 1990;61(1):113-24.

545.   DeKoter RP, Singh H. Regulation of B lymphocyte and macrophage development by graded expression of PU.1. Science. 2000;288(5470):1439-41.

546.   McIvor Z, Hein S, Fiegler H, Schroeder T, Stocking C, Just U, et al. Transient expression of PU.1 commits multipotent progenitors to a myeloid fate whereas continued expression favors macrophage over granulocyte differentiation. Experimental Hematology. 2003;31(1):39-47.

547.   Scott EW, Simon MC, Anastasi J, Singh H. Requirement of transcription factor pu.1 in the development of multiple hematopoietic lineages. Science. 1994;265(5178):1573-7.

548.   Salminen A, Ojala J, Kauppinen A, Kaarniranta K, Suuronen T. Inflammation in Alzheimer's disease: Amyloid-beta oligomers trigger innate immunity defence via pattern recognition receptors. Progress in Neurobiology. 2009;87(3):181-94.

549.   Chang MY, Boulden J, Katz JB, Wang LW, Meyer TJ, Soler AP, et al. Bin1 ablation increases susceptibility to cancer during aging, particularly lung cancer. Cancer Research. 2007;67(16):7605-12.

550.   Gold ES, Simmons RM, Petersen TW, Campbell LA, Kuo CC, Aderem A. Amphiphysin IIm is required for survival of Chlamydia pneumoniae in macrophages. Journal of Experimental Medicine. 2004;200(5):581-6.

551.   Muller A, Heseler K, Schmidt SK, Spekker K, MacKenzie CR, Daubener W. The missing link between indoleamine 2,3-dioxygenase mediated antibacterial and immunoregulatory effects. Journal of Cellular and Molecular Medicine. 2009;13(6):1125-35.

552.   Muller AJ, DuHadaway JB, Donover PS, Sutanto-Ward E, Prendergast GC. Inhibition of indoleamine 2,3-dioxygenase, an immunoregulatory target of the cancer suppression gene Bin1, potentiates cancer chemotherapy. Nature Medicine. 2005;11(3):312-9.

553.   MacKenzie CR, Heseler K, Muller A, Daubener W. Role of indoleamine 2,3-dioxygenase in antimicrobial defence and immuno-regulation: Tryptophan depletion versus production of toxic kynurenines. Current Drug Metabolism. 2007;8(3):237-44.

554.   Bonda DJ, Mailankot M, Stone JG, Garrett MR, Staniszewska M, Castellani RJ, et al. Indoleamine 2,3-dioxygenase and 3-hydroxy-kynurenine modifications are found in the neuropathology of Alzheimer's disease. Redox Report. 2010;15(4):161-8.

555.   Lu Q, Powles RL, Abdallah S, Ou D, Wang Q, Hu Y, et al. Systematic tissue-specific functional annotation of the human genome highlights immune-related DNA elements for late-onset Alzheimer's disease. PLoS Genet. 2017;13(7):e1006933.

556.   Veeraraghavalu K, Zhang C, Zhang XQ, Tanzi RE, Sisodia SS. Age-Dependent, Non-Cell-Autonomous Deposition of Amyloid from Synthesis of beta-Amyloid by Cells Other Than Excitatory Neurons. Journal of Neuroscience. 2014;34(10):3668-73.

557.   Konig G, Monning U, Czech C, Prior R, Banati R, Schreitergasser U, et al. Identification and differential expression of a novel alternative splice isoform of the beta-a4 amyloid precursor protein (app) messenger-rna in leukocytes and brain microglial cells. Journal of Biological Chemistry. 1992;267(15):10804-9.

558.   Li QX, Fuller SJ, Beyreuther K, Masters CL. The amyloid precursor protein of Alzheimer disease in human brain and blood. Journal of Leukocyte Biology. 1999;66(4):567-74.

559.   Tewhey R, Kotliar D, Park DS, Liu B, Winnicki S, Reilly SK, et al. Direct Identification of Hundreds of Expression-Modulating Variants using a Multiplexed Reporter Assay. Cell. 2016;165(6):1519-29.

560.   Gregersen PK. Cell type-specific eQTLs in the human immune system. Nature Genetics. 2012;44(5):478-80.

561.   Brown CD, Mangravite LM, Engelhardt BE. Integrative Modeling of eQTLs and Cis-Regulatory Elements Suggests Mechanisms Underlying Cell Type Specificity of eQTLs. Plos Genetics. 2013;9(8).

562.   Kumar D, Puan KJ, Andiappan AK, Lee B, Westerlaken GHA, Haase D, et al. A functional SNP associated with atopic dermatitis controls cell type-specific methylation of the VSTM1 gene locus. Genome Medicine. 2017;9.

563.   Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, et al. Targeted Genome Editing Across Species Using ZFNs and TALENs. Science. 2011;333(6040):307-.

564.   Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67-+.

565.   Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315(5819):1709-12.

566.   Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science. 2013;339(6121):819-23.

567.   Gasiunas G, Barrangou R, Horvath P, Siksnys V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proceedings of the National Academy of Sciences of the United States of America. 2012;109(39):E2579-E86.

568.   Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012;337(6096):816-21.

569.   Chen F, Pruett-Miller SM, Huang Y, Gjoka M, Duda K, Taunton J, et al. High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases. Nature Methods. 2011;8(9):753-U96.

570.   Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nature Protocols. 2013;8(11):2281-308.

571.   Kinoshita A, Whelan CM, Berezovska O, Hyman BT. The gamma secretase-generated carboxyl-terminal domain of the amyloid precursor protein induces apoptosis via Tip60 in H4 cells. Journal of Biological Chemistry. 2002;277(32):28530-6.

572.   Asai M, Iwata N, Yoshikawa A, Aizaki Y, Ishiura S, Saido TC, et al. Berberine alters the processing of Alzheimer's amyloid precursor protein to decrease A beta secretion. Biochemical and Biophysical Research Communications. 2007;352(2):498-502.

573.   Shin J, Yu SB, Yu UY, Jo SA, Ahn JH. Swedish mutation within amyloid precursor protein modulates global gene expression towards the pathogenesis of Alzheimer's disease. Bmb Reports. 2010;43(10):704-9.

574.   Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nature Biotechnology. 2013;31(9):827-+.

575.   Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ. A rapid and general assay for monitoring endogenous gene modification. Methods Mol Biol. 2010;649:247-56.

576.   Qiu P, Shandilya H, D'Alessio JM, O'Connor K, Durocher J, Gerard GF. Mutation detection using Surveyor (TM) nuclease. Biotechniques. 2004;36(4):702-+.

577.   Toyama EQ, Herzig S, Courchet J, Lewis TL, Loson OC, Hellberg K, et al. AMP-activated protein kinase mediates mitochondrial fission in response to energy stress. Science. 2016;351(6270):275-81.

578.   Rahmanto YS, Jung JG, Wu RC, Kobayashi Y, Heaphy CM, Meeker AK, et al. Inactivating ARID1A Tumor Suppressor Enhances TERT Transcription and Maintains Telomere Length in Cancer Cells. Journal of Biological Chemistry. 2016;291(18):9690-9.

579.   Wills QF, Mellado-Gomez E, Nolan R, Warner D, Sharma E, Broxholme J, et al. The nature and nurture of cell heterogeneity: accounting for macrophage gene-environment interactions with single-cell RNA-Seq. Bmc Genomics. 2017;18.

580.   Vouillot L, Thélie A, Pollet N. Comparison of T7E1 and surveyor mismatch cleavage assays to detect mutations triggered by engineered nucleases. G3 (Bethesda). 2015;5(3):407-15.

581.   Yeung A, Hattangadi D, Blakesley L, Nicolas E. Enzymatic mutation detection technologies. Biotechniques. 2005;38(5):749-58.

582.   Inui M, Miyado M, Igarashi M, Tamano M, Kubo A, Yamashita S, et al. Rapid generation of mouse models with defined point mutations by the CRISPR/Cas9 system. Scientific Reports. 2014;4.

583.   Bialk P, Rivera-Torres N, Strouse B, Kmiec EB. Regulation of Gene Editing Activity Directed by Single-Stranded Oligonucleotides and CRISPR/Cas9 Systems. Plos One. 2015;10(6).

584.   Paquet D, Kwart D, Chen A, Sproul A, Jacob S, Teo S, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. Nature. 2016;533(7601):125-+.

585.   Hou ZG, Zhang Y, Propson NE, Howden SE, Chu LF, Sontheimer EJ, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proceedings of the National Academy of Sciences of the United States of America. 2013;110(39):15644-9.

586.   Muller M, Lee CM, Gasiunas G, Davis TH, Cradick TJ, Siksnys V, et al. Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. Molecular Therapy. 2016;24(3):636-44.

587.   Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, et al. In vivo genome editing using Staphylococcus aureus Cas9. Nature. 2015;520(7546):186-U98.

588.   Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng ZL, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. Nature. 2015;523(7561):481-U249.

589.   Guilinger JP, Thompson DB, Liu DR. Fusion of catalytically inactive Cas9 to Fokl nuclease improves the specificity of genome modification. Nature Biotechnology. 2014;32(6):577-+.

590.   Slaymaker IM, Gao LY, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science. 2016;351(6268):84-8.

591.   Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Joung JK. High-Fidelity CRISPR-Cas9 Nucleases with No Detectable Genome-Wide Off-Target Effects. Molecular Therapy. 2016;24:S288-S.

592.   Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nature Biotechnology. 2016;34(2):184-+.

593.   Dang Y, Jia GX, Choi J, Ma HM, Anaya E, Ye CT, et al. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. Genome Biology. 2015;16.

594.   Nielsen S, Yuzenkova Y, Zenkin N. Mechanism of Eukaryotic RNA Polymerase III Transcription Termination. Science. 2013;340(6140):1577-80.

595.   Cho SW, Kim S, Kim Y, Kweon J, Kim HS, Bae S, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Research. 2014;24(1):132-41.

596.   Fu YF, Reyon D, Joung JK. Targeted Genome Editing in Human Cells Using CRISPR/Cas Nucleases and Truncated Guide RNAs. Use of Crispr/Cas9, Zfns, and Talens in Generating Site-Specific Genome Alterations. 2014;546:21-45.

597.   Hendel A, Bak RO, Clark JT, Kennedy AB, Ryan DE, Roy S, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. Nature Biotechnology. 2015;33(9):985-U232.

598.   Richardson CD, Ray GJ, DeWitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nature Biotechnology. 2016;34(3):339-+.

599.   Chu VT, Weber T, Wefers B, Wurst W, Sander S, Rajewsky K, et al. Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nature Biotechnology. 2015;33(5):543-U160.

600.   Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nature Biotechnology. 2015;33(5):538-U260.

601.   Singh P, Schimenti JC, Bolcun-Filas E. A Mouse Geneticist's Practical Guide to CRISPR Applications. Genetics. 2015;199(1):1-U402.

602.   Srivastava M, Nambiar M, Sharma S, Karki SS, Goldsmith G, Hegde M, et al. An Inhibitor of Nonhomologous End-Joining Abrogates Double-Strand Break Repair and Impedes Cancer Progression. Cell. 2012;151(7):1474-87.

603.   Heyer WD, Ehmsen KT, Liu J. Regulation of Homologous Recombination in Eukaryotes. Annual Review of Genetics, Vol 44. 2010;44:113-39.

604.   Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. Nature Biotechnology. 2015;33(1):73-80.

605.   Kim S, Kim D, Cho SW, Kim J, Kim JS. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Research. 2014;24(6):1012-9.

Appendix 1. Immunocytochemistry controls
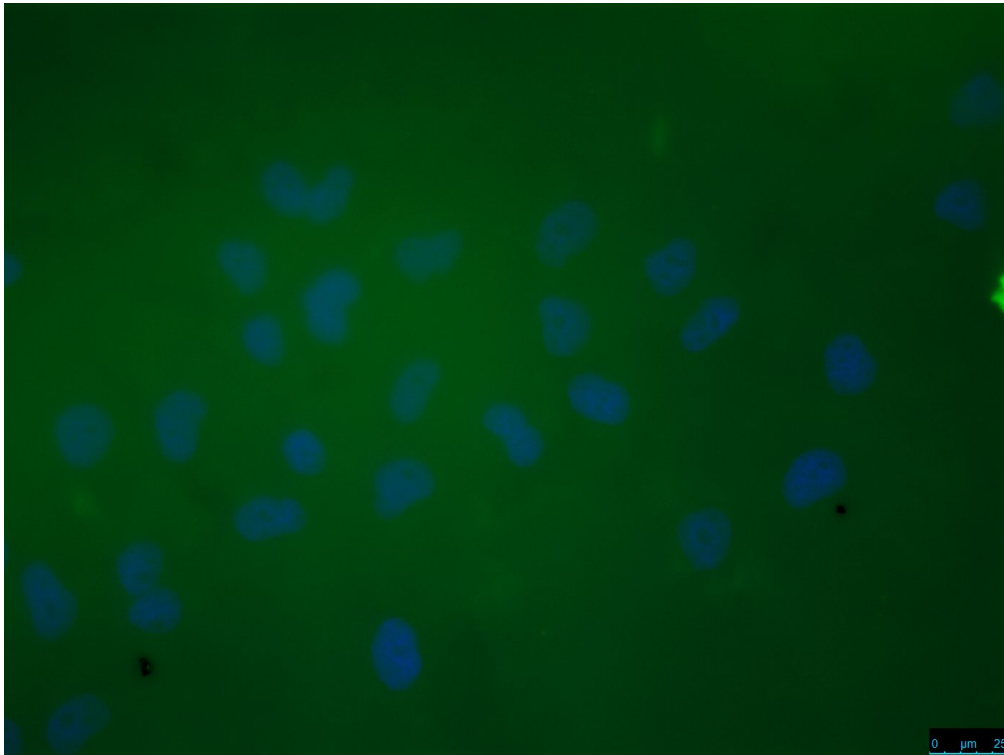
Appendix Figure 1.1 Representative image of Tf-488 uptake assay without acid wash. The presence of green background is surface bound Tf-488. This is not observed in cells that have undergone the acid wash step, demonstrating that the acid wash step is sufficient to remove surface bound Tf-488. N=3.
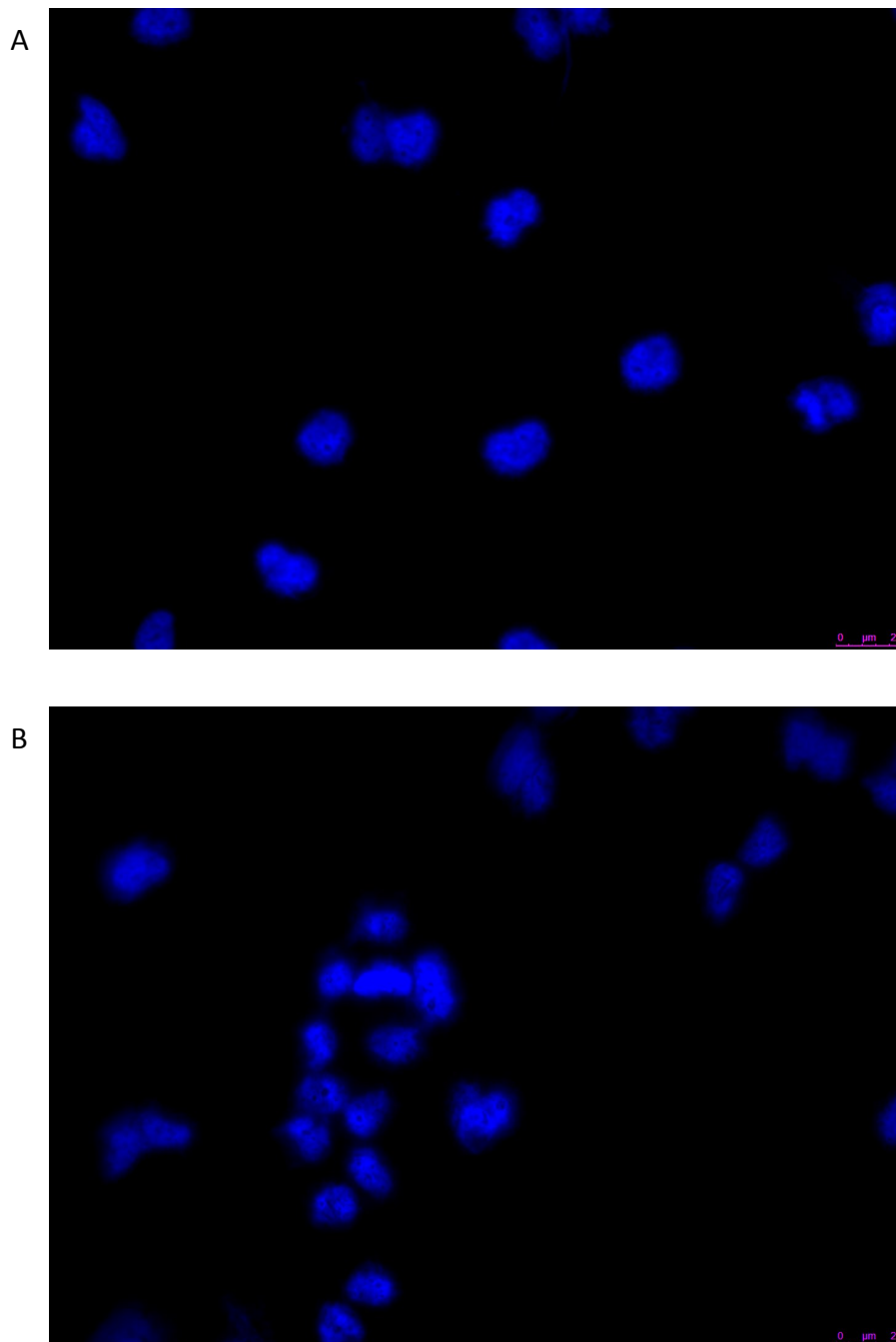
Appendix Figure 1.2 Representative secondary antibody only control images. A) H4 cells probed only with Mouse 594 secondary antibody and stained with DAPI. Lack of 594 staining indicates the secondary antibody binds specifically to the LAMP2 primary antibody. N=3. B) H4 cells probed only with Rabbit 594 secondary antibody and stained with DAPI. Lack of 594 indicates that the secondary antibody binding is specific to the EEA1 primary antibody. N=3.

Appendix 2. Complete DNA sequence of insert cloned into pGL4.23

Insert sequence is shown in red using sequencing primers described in table 5.7

Sense Orientation – 235 bp insert – risk

CCCCACCTTGCAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTGGCCTA
ACTGGCCGGTACCACCCCTGTTTCCTCCTCTGTAAGATGGGGTGAGATTCACCTACCCCACA
GGGCTGGTGTGAAGGTTAGGGTAGATTATGCTTGAAAAGCACTGTGTACAACAGTCACCTA
GAGTGCCTGCTCAAACGCAGGCAGCTGTCATTCTTATCATCTGTCGTGAGTGTCTGCTCTGG
GCCTGGCCCAGTGGCAGGTGTGGCCTGTGAGGACTCTCCCAAGGCAGGGGCAGTCCCTGCA
AGCCGGAGGCCGGAGGATGAGGGGGCCTTGTAGTCAGCATCGAGACTTCCCCTTCGCCTCTC
CCTCCACCTTCCCGTTCCATCCTGTTTCTAGTCTTCTCTGGTTCCTTCTTCTGAAAGTAAAAAG
GGGAAAAGGGTATTTTTAAGAAGCAGAACAGAGATTTCTTTCCTTTTTAAAAACTCCTATTT
AAAGTCCCAGTTTCCAGAAAAGTCAGCGTAAGTCACTGGCTATGCATAGGGAATAAAAGAA
GTCATCTTAGTGGGGCTCAGGGGCGGGGGGTGGGGGGTCCCTGCACTCAGAAGCCTCTGCC
CCTCCAAGTGTGTCTGCTCGGAGGGTGCTCTAGGGGATGAGGCGCTGCAGTTGGGGGGGCAA
TTAGGTGAGTGCAGGGCCAGGAGACCCCTCGATTCCATCTGCACCCGTGGCAGGCTGTTCCC
TTGCTGCTGGAGTTGGCCTCAGTTTCTGGTACCTGAGCTCGCTAGCCTCGAGGATATCAAGA
TCTGGCCTCGGCGGCCAAGCTTAGACACTAGAGGGTATATAATGGAAGCTCGACTTCCAGCT
TGGCAATCCGGTACTGTTGGTAAAGCCACCATGGAAGATGCCAAAACATTAAGAAGGGCCC
AGCGCCATTCTACCACTCGAAGACGGGACCGCCGGGCGAGCAGCTGGCACCAAAGGCATGA
AGCGCTACGGCTCTGGTGCTCCGGGCACCATTCGCCC

Antisense Orientation – 235 bp insert – risk

AAAAATAGCATTAGGCTGTCCCCAGTGCAGTGCAGGTGCCAGAACATTTCTCTGGCCTAACT
GGCCGGTACCCCTGAGCCCCACTAAGATGACTTCTTTTATTCCCTATGCATAGCCAGTGACTT
ACGCTGACTTTTCTGGAAACTGGGACTTTAAATAGGAGTTTTTAAAAAGGAAAGAAATCTCT
GTTCTGCTTCTTAAAAATACCCTTTTCCCCTTTTTACTTTCAGAAGAAGGAACCAGAGAAGAC
TAGAAACAGGATGGAACGGGAAGGTGGAGGGAGAGGCGAAGGGGAAGTCTCGATGCTGGT
ACCTGAGCTCGCTAGCCTCGAGGATATCAAGATCTGGCCTCGGCGGCCAAGCTTAGACACTA
GAGGGTATATAATGGAAGCTCGACTTCCAGCTTGGCAATCCGGTACTGTTGGTAAAGCCACC
ATGGAAGATGCCAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGA
CCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCGC
CTTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGATGAGCGTTC
GGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCAG
CGAGAATAGCTTGCAGTTCTTCATGCCCGTGTTGGGTGCCCTGTTCATCGGTGTGGCTGTGGC
CCCAGCTAACGACATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCC
ACCGTCGTATTCGTGAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTAC
CGATCATACAAAAGATCATCATCATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGCAT
GTACACCTTCGTGACTTCCCATTTGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGAGA
GCTTCGACCGGGACAAAACCATCGCCCTGATCATGAACAGTAATGGCAGTACCGGATTGCC
CAAGGGCGTATCCCTACCGCACCGCACCGCTTGGGTCCGATTCAATCATGCCCGCGACCCCA
TCTTCCGGAACAAAATAATCCCCGAAACCGCTTTCCTCAAGGGGGGGGCCATTTCACCACGG
CTTCTGCAAGGTTCACCACGCTGGGGCTACTTTGATTGGCGGGCTTTCGGGGCCATGGCCAA
GTCACCCCTTCCAAGGAAGAACGATTCTTTGCCAACTTTGCAAAACATAAAAAATCAACCAT
GCCCGCGGGGGGCCCACACCATTAAAATTTCTTCCCCTAGAAGACTCTCCCCATCAAAGTGT
TGACATTTGGACTATTTGCGAATAATCCCACATCAGGGGATGCTCCCCTCTCCAAATAATAT
GGTGAACGCTTGCGGCGAATACTGTCACCTTAGCGGCCCGTCCACTCGGGTACCCTTGCCAC
AGAATAAGCCAAGGTCTGTTCTATCTCACCGCAAGATGGTGAAAGAATGGTAATGACAGTG
AGGTGAAGGAAGAGTGCCCCTCACTATATAAGAGGGTAGACGATTGTAGCACCAGAACTAT
CAACAGTGGTGGCAAAACGTGGG

Sense Orientation – 235 bp insert – non risk

GAAAATAGCAATAGGCTGTCCCAGTTGCAGTGCAGGTGCCAGAACATTTCTCTGGCCTAACT
GGCCGGTACCAGCATCGAGACTTCCCCTTCGCCTCTCCCTCCACCTTCCCGTTCCATCCTGTT
TCTAGTCTTCTCTGGTTCCTTCTTCTGAAAGTAAAAAGGGGAAAAGGGTGTTTTTAAGAAGC
AGAACAGAGATTTCTTTCCTTTTTAAAAACTCCTATTTAAAGTCCCAGTTTCCAGAAAAGTCA
GCGTAAGTCACTGGCTATGCATAGGGAATAAAAGAAGTCATCTTAGTGGGGCTCAGGGGTA
CCTGAGCTCGCTAGCCTCGAGGATATCAAGATCTGGCCTCGGCGGCCAAGCTTAGACACTAG
AGGGTATATAATGGAAGCTCGACTTCCAGCTTGGCAATCCGGTACTGTTGGTAAAGCCACCA
TGGAAGATGCCAAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGAC
CGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCGCC
TTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGATGAGCGTTCG
GCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCAGC
GAGAATAGCTTGCAGTTCTTCATGCCCGTGTTGGGTGCCCTGTTCATCGGTGTGGCTGTGGCC
CCAGCTAACGACATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGCCCA
CCGTCGTATTCGTGAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCTACC
GATCATACAAAAGATCATCATCATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGCATG
TACACCTTCGTGACTTCCCATTTGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGAGAG
CTTCGACCGGGACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTGCCC
AAGGGCGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATTCAGTCATGCCCGCGACCCCAT
CTTCGGCAACCAGATCATCCCCGACACCGCTATCCTCAGCGTGGTGCCATTTCACCACGGCT
TCGGCATGTTCACCACGCTGGGCTACTTGATCTGCGGCTTTCGGGTCGTGCTCATGTACCGCT
TCGAGGAGGAGCTATTCTTGGCCAGCTTGCAAGACTTAAAAATTCAATCTTGCCTGTGGGTG
CCACACTATTTAACTTCTTCGCTAAGAAGCCTCTCATCGCCAAGTCCAACTAAGGACATTTG
CCAAAAATACCCCAGCGGGGGGGGCCCGCCTCAACAAGGAGTAAGTGAAGCCGGGGCCAAG
CCTTTCCCTTCCAACGGTTCCCACGGGGTTCGGGTTGGAAAAAACCGCCCCGCTTTTTGTTCC
CCCCAAAGGGGACAAAACCTCGTGGCAAAAAGAAGGGGCGGCTCCTTTCCAGAAATGGGGT
GTATTTATGCCCTCAGAAAGATCGGGGGTCACACGACGGGCCTAGAATCGCTTGTCTCCGTA
AATATAGTTGCCGTAAATCGAGGATGCTATAGCTTTTCGTAGGAACTGGAGCGCGCTAAGTT
GGGGCTCTGGACTCTCAATAGGT

Antisense Orientation – 235 bp insert – non risk

AAAAAAGCAATAGGCTGTCCCCCAGTGCAGTGCAGGTGCCAGAACATTTCTCTGGCCTAACT
GGCCGGTACCCCTGAGCCCCACTAAGATGACTTCTTTTATTCCCTATGCATAGCCAGTGACTT
ACGCTGACTTTTCTGGAAACTGGGACTTTAAATAGGAGTTTTTAAAAAGGAAAGAAATCTCT
GTTCTGCTTCTTAAAAACACCCTTTTCCCCTTTTTACTTTCAGAAGAAGGAACCAGAGAAGA
CTAGAAACAGGATGGAACGGGAAGGTGGAGGGAGAGGCGAAGGGGAAGTCTCGATGCTGG
TACCTGAGCTCGCTAGCCTCGAGGATATCAAGATCTGGCCTCGGCGGCCAAGCTTAGACACT
AGAGGGTATATAATGGAAGCTCGACTTCCAGCTTGGCAATCCGGTACTGTTGGTAAAGCCAC
CATGGAAGATGCCAAAAACATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGG
ACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGTGCCCGGCACCATCG
CCTTTACCGACGCACATATCGAGGTGGACATTACCTACGCCGAGTACTTCGAGATGAGCGTT
CGGCTGGCAGAAGCTATGAAGCGCTATGGGCTGAATACAAACCATCGGATCGTGGTGTGCA
GCGAGAATAGCTTGCAGTTCTTCATGCCCGTGTTGGGTGCCCTGTTCATCGGTGTGGCTGTG
GCCCCAGCTAACGACATCTACAACGAGCGCGAGCTGCTGAACAGCATGGGCATCAGCCAGC
CCACCGTCGTATTCGTGAGCAAGAAAGGGCTGCAAAAGATCCTCAACGTGCAAAAGAAGCT
ACCGATCATACAAAAGATCATCATCATGGATAGCAAGACCGACTACCAGGGCTTCCAAAGC
ATGTACACCTTCGTGACTTCCCATTTGCCACCCGGCTTCAACGAGTACGACTTCGTGCCCGA
GAGCTTCGACCGGGACAAAACCATCGCCCTGATCATGAACAGTAGTGGCAGTACCGGATTG
CCCAAGGGCGTAGCCCTACCGCACCGCACCGCTTGTGTCCGATTCAGTCATGCCCGCGACCC
CATCTTCGGGAACCAGATCATCCCCGACACCGCTATCCTCAGCGGTGGTGCCATTTCACCCC
GGCTTCGGCATGTTCACCACGCTGGGCTACTTGATCTGCGGCTTTCAGGGCAGGCTCATGGA
CCGCTTCCAAGGAAGGACTATTTCTTGCACACCTTGCAAAACTTAAAAATTCAACTTGCCGG
GCGGGGGCGCAACCATTTATCTTCTTTCTCATAGAACAGCTCCAGTGACAAATTGCAACCCA
GCTAATTTTGTCAAAAATACGCGCACGGGGAGTGCTCCTTCTCTTAAAAGGATAATGTTGGA
CGCTTGGGCTTAATGGTAGTTCGTTGCAACAGTGTCCCACCAGGGGACAT

Sense Orientation – 693 bp insert – risk

ACAAACTAGCAATTAGGCTGTCCCCAGTGCAGTGCAGGTGCCAGAACATTTCTCTGGCCTAA
CTGGCCGGTACCACCCCTGTTTCCTCCTCTGTAAGATGGGGTGAGATTCACCTACCCCACAG
GGCTGGTGTGAAGGTTAGGGTAGATTATGCTTGAAAAGCACTGTGTACAACAGTCACCTAG
AGTGCCTGCTCAAACGCAGGCAGCTGTCATTCTTATCATCTGTCGTGAGTGTCTGCTCTGGGC
CTGGCCCAGTGGCAGGTGTGGCCTGTGAGGACTCTCCCAAGGCAGGGGCAGTCCCTGCAAG
CCGGAGGCCGGAGGATGAGGGGGCCTTGTAGTCAGCATCGAGACTTCCCCTTCGCCTCTCCC
TCCACCTTCCCGTTCCATCCTGTTTCTAGTCTTCTCTGGTTCCTTCTTCTGAAAGTAAAAAGG
GGAAAAGGGTATTTTTAAGAAGCAGAACAGAGATTTCTTTCCTTTTTAAAAACTCCTATTTA
AAGTCCCAGTTTCCAGAAAAGTCAGCGTAAGTCACTGGCTATGCATAGGGAATAAAAGAAG
TCATCTTAGTGGGGCTCAGGGGCGGGGGGTGGGGGGTCCCTGCACTCAGAAGCCTCTGCCCC
TCCAAGTGTGTCTGCTCGGAGGGTGCTCTAGGGGATGAGGCGCTGCAGTTGGGGGGCAATT
AGGTGAGTGCAGGGCCAGGAGACCCCTCGATTCCATCTGCACCCGTGGCAGGCTGTTCCCTT
GCTGCTGGAGTTGGCCTCAGTTTCTGGTACCTGAGCTCGCTAGCCTCGAGGATATCAAGATC
TGGCCTCGGCGGCCAAGCTTAGACACTAGAGGGTATATAATGGAAGCTCGACTTCCAGCTTG
GCAATCCGGTACTGTTGGTAAAGCCACCATGGAAGATGCCAAAAACATTAAGAAGGGCCCA
GCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAAGC
GCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACATTACC
TACGCCAAGTACTTCGAGATGAACGTTCGGCTGGCAGAATCTATTAATCGCTATGGGCTTAA
TACAAACCATCGGTATCGTTGTGTTGCTTCGGATATTAGCTTGGGATTCTTCTTG

Antisense Orientation – 693 bp insert – risk:

AAAAACTAGCAATTAGGCTGTCCCCAGTGCAGTGCAGGTGCCAGAACATTTCTCTGGCCTAA
CTGGCCGGTACCAGAAACTGAGGCCAACTCCAGCAGCAAGGGAACAGCCTGCCACGGGTGC
AGATGGAATCGAGGGGTCTCCTGGCCCTGCACTCACCTAATTGCCCCCCAACTGCAGCGCCT
CATCCCCTAGAGCACCCTCCGAGCAGACACACTTGGAGGGGCAGAGGCTTCTGAGTGCAGG
GACCCCCACCCCCCGCCCCTGAGCCCCACTAAGATGACTTCTTTTATTCCCTATGCATAGCC
AGTGACTTACGCTGACTTTTCTGGAAACTGGGACTTTAAATAGGAGTTTTTAAAAAGGAAAG
AAATCTCTGTTCTGCTTCTTAAAAATACCCTTTTCCCCTTTTTACTTTCAGAAGAAGGAACCA
GAGAAGACTAGAAACAGGATGGAACGGGAAGGTGGAGGGAGAGGCGAAGGGGAAGTCTCG
ATGCTGACTACAAGGCCCCCTCATCCTCCGGCCTCCGGCTTGCAGGGACTGCCCCTGCCTTG
GGAGAGTCCTCACAGGCCACACCTGCCACTGGGCCAGGCCCAGAGCAGACACTCACGACAG
ATGATAAGAATGACAGCTGCCTGCGTTTGAGCAGGCACTCTAGGTGACTGTTGTACACAGTG
CTTTTCAAGCATAATCTACCCTAACCTTCACACCAGCCCTGTGGGGTAGGTGAATCTCACCC
CATCTTACAGAGGAGGAAACAGGGGTGGTACCTGAGCTCGCTAGCCTCGAGGATATCAAGA
TCTGGCCTCGGCGGCCAAGCTTAGACACTAGAGGGTATATAATGGAAGCTCGACTTCCAGCT
TGGCAATCCGGTACTGTTGGTAAAGCCACCATGGAAGATGCCAAAAACATTAAGAAGGGCC
CAGCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAA
GCGCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACATTA
CCTACGCCGAGTACTTCGAGATGAGCGTTCGGCTGGCAGAAGCTATGAAGCGCTATGGGCT
GAATACAAACCATCCGGATTCGTGGTGTGCACCGAAAATAAGCTTGCAGTTCTTCATGCCCG
TGTTGGGTGCCCTGTTCATCGGGGTGGCTTGGGCCCCAGCTAACGACATCTACAACGAAGCC
GAACTGCTGAACAGCATGGGCATCAGCCAGCCCACCGTCTTATTCTTGAACAAAAAGGGC
TGCAAAAAATCCTCACCGGGCAAAAAAGCTTCCCGTCTTACCAAAGATCGCTTCTTGGGAAA
CCAAACCGACACACGGAGGTTCAAAAAATTGGGACCTTTGGGGATTTCCATTTTGCCCCCCG
GTTTCGAGAAATCCAAATTTGCGGCCCAAAATTCTCTACGTGATCATACCTCTCTCTATATAA
GG

Sense Orientation – 693 bp insert – non risk:

GGGGCGTTTTTTTTTGCTGCACTGCAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAAC
ATTTCTCTGGCCTAACTGGCCGGTACCACCCCTGTTTCCTCCTCTGTAAGATGGGGTGAGATT
CACCTACCCCACAGGGCTGGTGTGAAGGTTAGGGTAGATTATGCTTGAAAAGCACTGTGTAC
AACAGTCACCTAGAGTGCCTGCTCAAACGCAGGCAGCTGTCATTCTTATCATCTGTCGTGAG
TGTCTGCTCTGGGCCTGGCCCAGTGGCAGGTGTGGCCTGTGAGGACTCTCCCAAGGCAGGGG
CAGTCCCTGCAAGCCGGAGGCCGGAGGATGAGGGGGCCTTGTAGTCAGCATCGAGACTTCC
CCTTCGCCTCTCCCTCCACCTTCCCGTTCCATCCTGTTTCTAGTCTTCTCTGGTTCCTTCTTCTG
AAAGTAAAAAGGGGAAAAGGGTGTTTTTAAGAAGCAGAACAGAGATTTCTTTCCTTTTTAA
AAACTCCTATTTAAAGTCCCAGTTTCCAGAAAAGTCAGCGTAAGTCACTGGCTATGCATAGG
GAATAAAAGAAGTCATCTTAGTGGGGCTCAGGGGCGGGGGGTGGGGGGTCCCTGCACTCAG

AAGCCTCTGCCCCTCCAAGTGTGTCTGCTCGGAGGGTGCTCTAGGGGATGAGGCGCTGCAGT
TGGGGGGCAATTAGGTGAGTGCAGGGCCAGGAGACCCCTCGATTCCATCTGCACCCGTGGC
AGGCTGTTCCCTTGCTGCTGGAGTTGGCCTCAGTTTCTGGTACCTGAGCTCGCTAGCCTCGAG
GATATCAAGATCTGGCCTCGGCGGCCAAGCTTAGACACTAGAGGGTATATAATGGAAGCTC
GACTTTCCAGCTTGGCAATCCGGTACTGTTGGTAAAAGCCACCATGGAAGATGCCAAAAAC
ATTAAGAAGGGCCCAGCGCCATTCTACCCACTCAAGACGGGACCGCCGGCGAGCAACTGCA
CAAAGCCATGAAGCGCTTACGCCCTGGTAGCCCG

Antisense Orientation – 693 bp insert – non risk:

ACAAACAGCAATTAGGCTGTCCCCAGTTGCAGTGCAGGTGCCAGAACATTTCTCTGGCCTAA
CTGGCCGGTACCAGAAACTGAGGCCAACTCCAGCAGCAAGGGAACAGCCTGCCACGGGTGC
AGATGGAATCGAGGGGTCTCCTGGCCCTGCACTCACCTAATTGCCCCCCAACTGCAGCGCCT
CATCCCCTAGAGCACCCTCCGAGCAGACACACTTGGAGGGGCAGAGGCTTCTGAGTGCAGG
GACCCCCCACCCCCCGCCCCTGAGCCCCACTAAGATGACTTCTTTTATTCCCTATGCATAGCC
AGTGACTTACGCTGACTTTTCTGGAAACTGGGACTTTAAATAGGAGTTTTTAAAAAGGAAAG
AAATCTCTGTTCTGCTTCTTAAAAACACCCTTTTCCCCTTTTTACTTTCAGAAGAAGGAACCA
GAGAAGACTAGAAACAGGATGGAACGGGAAGGTGGAGGGAGAGGCGAAGGGGAAGTCTCG
ATGCTGACTACAAGGCCCCCTCATCCTCCGGCCTCCGGCTTGCAGGGACTGCCCCTGCCTTG
GGAGAGTCCTCACAGGCCACACCTGCCACTGGGCCAGGCCCAGAGCAGACACTCACGACAG
ATGATAAGAATGACAGCTGCCTGCGTTTGAGCAGGCACTCTAGGTGACTGTTGTACACAGTG
CTTTTCAAGCATAATCTACCCTAACCTTCACACCAGCCCTGTGGGGTAGGTGAATCTCACCC
CATCTTACAGAGGAGGAAACAGGGGTGGTACCTGAGCTCGCTAGCCTCGAGGATATCAAGA
TCTGGCCTCGGCGGCCAAGCTTAGACACTAGAGGGTATATAATGGAAGCTCGACTTCCAGCT
TGGCAATCCGGTACTGTTGGTAAAGCCACCATGGAAGATGCCAAAAACATTAAGAAGGGCC
CAGCGCCATTCTACCCACTCGAAGACGGGACCGCCGGCGAGCAGCTGCACAAAGCCATGAA
GCGCTACGCCCTGGTGCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACATTA
CCTACGCCGAGTACTTCGAGATGAGCGTTCGGCTGGCAAAAGCTATGAAGCGCTATGGGCT
GAATACAAACCATCCGGATCGTGGTGTGCACCGAAAATAACCTTGCAATTCTTCATGCCCGG
GGTGGGGGGCCCTGTTCATCAGTGTGGGCTGGGGCCCCAGCTAACGAATTCTACAACGAGTT
CGAACTTGTTGACATCAGGGGGCATCACCCATTCCCACCGTCGTATTCCTGAAACAAGAAAT
GGCTCTAAAAAATCCTCAGGGTGCGAAGAAACCTCCCGATCCTACAAAGAATCTCCCTCTGG
GAATGCGAAACCTGTAACAAAGGTCTTCCGAGGGATGTACCCTTTGGTGAATTTCGATTGTG
CACCCGTCGTCAACAAAAGGCAATTTTGGGCCGGAATTATCTGCCGGACGAAAAACCCTCCT
CGGATGGGTTAAACGTAGGGGCATACCTAATTTCCTTGTTGTAGGG

Appendix 3. Creation of Cellular Model containing rs59335482 risk allele using CRISPR/Cas9n to be used in functional analysis

1. Introduction

1.1 Rs59335482, an insertion identified to have a functional impact on gene expression

The IGAP GWAS detected a significant association signal on chromosome 2 at position 127892810 (p=6.9 x 10^{-44}). This association signal lies approximately 30kb upstream of *BIN1* (140). Chapuis *et al*, fine mapped the *BIN1* risk locus by performing an imputation of genotypes in the European Alzheimer's Disease Initiative (EADI) cohort. Two SNPs were found to be associated with LOAD, rs4663105 and rs6733839, the latter being previously identified by Lambert *et al*. These SNPs were investigated for their ability to alter transcription *in vitro* by the use of gene reporter assays and neither SNP showed evidence for altering transcriptional activity in SKNSH-SY5Y and HEK cells (261).

Linkage disequilibrium (LD) analysis revealed that rs4663105 and rs6733839 were located within a 6.7 Kb LD block, which also contained rs744373, the first SNP at the *BIN1* locus shown to be genome wide significant (137). This indicates that a functional risk variant may be located within this LD block. Chapuis *et al* went on to sequence this LD block and identified eight polymorphisms.

These eight polymorphisms were imputed into the EADI1 GWAS data set. This found that one polymorphism was associated with an increased risk for LOAD,

rs59335482, an insertion of three cytosines. This association was replicated in the GERAD1 GWAS data and in an independent Flanders-Belgian population. This showed rs59335482 is significantly associated with an increased risk of LOAD with a genome-wide significant meta-analysed odds ratio of 1.20 with a p-value of 3.8×10^{-11}. Rs59335482 has a MAF of 0.27 and is located just under 26 Kb from the start of *BIN1* and approximately 1.3 Kb downstream of the association signal.

Following the identification of rs59335482, Chapuis *et al* performed a gene reporter assay to reveal whether this polymorphism influenced transcriptional activity. An increase in luciferase activity was observed in both neuroblastoma SH-SY5Y (+101%) and HEK293 (+33%) cells. The insertion allele was also found to be associated with increased *BIN1* mRNA expression in the frontal cortex of AD brains. Rs59335482 was imputed in 98 HapMap 3 individuals for whom *BIN1* RNA levels had been measured in lymphoblastoid cell lines. This again showed an association between the rs59335482 insertion allele with an increase in *BIN1* mRNA levels (261).

1.2 Imputation of rs59335482 in the GERAD dataset

In 2009, a GWAS of LOAD was performed involving over 16,000 individuals (135). This analysis identified two loci that reached the threshold for genome-wide significant association ($p \le 9.4 \times 10^{-8}$). Despite not reaching the genome-wide significance threshold, this stage 1 analysis identified two variants within *BIN1* that showed suggestive evidence of association at $p < 10^{-5}$.

In order to determine the association between rs59335482 and LOAD, this variant was imputed in a subset of the GERAD dataset consisting of 3332 cases and 9832 controls. Full details on samples and quality control criteria have

been described elsewhere (135). The dataset was imputed with SHAPEIT and IMPUTE2 software using 1000 genomes data (2012 release) as a reference panel. SNPs with info score quality estimates of less than 0.9 were excluded from the analysis. Association tests were performed under an additive model, using logistic regression, using PLINK. The analysis was adjusted for principal components and geographical region (performed by Nandini Badarinarayan). Data is shown in appendix table 3.1.

Appendix Table 3.1. PLINK output from imputation of rs59335482 in GERAD data set.
Imputation reveals that rs59335482 has a more significant association with LOAD in the dataset than the SNPs identified with suggestive associations.

Field	PLINK Output
Chromosome	2
SNP	rs59335482
Physical position (bp)	12789146
First Allele	TGGG
Second allele	T
Test	ADD
Number of non-missing genotypes	13164
Odds Ratio (OR)	1.175
Standard Error	0.03261
Lower bound of 95% CI for OR	1.102
Upper bound of 95% CI for OR	1.252
Coefficient t-statistic	4.94
p	7.82×10^{-7}

This analysis shows that rs59335482 is more significantly associated (p = 7.82 x 10^{-7}) with LOAD in this subset than the SNPs previously identified to have suggestive evidence of association. This further suggests that rs59335482

287

could be the functionally relevant variant that accounts for the GWAS association signal.

1.3 rs59335482 is associated with regulatory elements within brain tissue

HaploReg, hosted by the Broad institute, is a database that contains information on non-coding variants within haplotype blocks (431). This can be utilised to explore the potential functions of SNPs at disease-associated loci. Using LD data from the 1000 genomes project, SNPs and indels in LD can be identified along with diverse functional information.

HaploReg can be used to search for variants in high LD with the variant of interest. By specifying a minimum r^2 value, variants with an equal or greater correlation with the variant of interest than the specified r^2 will be identified. Submitting rs59335482 as a query SNP into HaploReg v4.1 with an r^2 specified to be greater than 0.8, identified no other variants in high LD with rs59335482. Functional data from the HaploReg database revealed that rs59335482 was associated with regulatory regions within a number of brain regions (Output is shown in appendix figure 3.1).

Group	Mnemonic	Description	Chromatin states (Core 15-state model)	Chromatin states (25-state model using 12 imputed marks)	H3K4me1	H3K4me3	H3K27ac	H3K9ac
Thymus	THYM	Thymus						
Thymus	THYM.FET	Fetal Thymus			H3K4me1_Enh			
Brain	BRN.HIPP.MID	Brain Hippocampus Middle		17_EnhW2	H3K4me1_Enh	H3K4me3_Pro	H3K27ac_Enh	
Brain	BRN.SUB.NIG	Brain Substantia Nigra		17_EnhW2			H3K27ac_Enh	H3K9ac_Pro
Brain	BRN.ANT.CAUD	Brain Anterior Caudate		18_EnhAc	H3K4me1_Enh		H3K27ac_Enh	H3K9ac_Pro
Brain	BRN.CING.GYR	Brain Cingulate Gyrus		18_EnhAc	H3K4me1_Enh		H3K27ac_Enh	H3K9ac_Pro
Brain	BRN.INF.TMP	Brain Inferior Temporal Lobe		17_EnhW2	H3K4me1_Enh		H3K27ac_Enh	H3K9ac_Pro
Brain	BRN.ANG.GYR	Brain Angular Gyrus					H3K27ac_Enh	
Brain	BRN.DL.PRFRNTL.CRTX	Brain_Dorsolateral_Prefrontal_Cortex			H3K4me1_Enh		H3K27ac_Enh	

Appendix figure 3.1. Haploreg detailed view of rs59335482. Regulatory chromatin states from DNase and histone ChIP-Seq data from Roadmap Epigenomics indicates rs59335482 is located within a region that appears to have active regulatory elements within various brain tissues. Database accessed on 14.11.2015.

288

Rs59335482 is located in a genomic region marked with enhancer modifications, such as H3K27ac and H3K4me1, in brain tissues such as fetal thymus, hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe, angular gyrus and dorsolateral prefrontal cortex. Histone modifications that typically flank active promoters, such as H3K4me3 and H3K9ac, were identified in the proximity of rs59335482 in the hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe. This data provides evidence that this variant may be located in a functional regulatory region capable of influencing gene transcription and that this region may be active within brain tissue.

RegulomeDB annotates rs59335482 with a score of 5, indicating potential transcription factor binding or DNase peak. A weak transcriptional chromatin state is described in the hippocampus, inferior temporal lobe and substantia nigra, whereas a quiescent chromatin state is described in the cingulate gyrus and dorsolateral prefrontal cortex (432).

1.4 Studying the effects of rs59335482 in a cellular context

Generating isogenic models, which differ only by rs59335482 genotype, would provide a platform to elucidate the effect of rs59335482 in a cellular context. As rs59335482 has previously been associated with changes in gene expression, primary investigations would focus of quantifying *BIN1* expression. mRNA levels could be quantified via quantitative PCR and protein levels quantified initially by Western Blot and finer changes may be observed via ELISA.

Should rs59335482 elicit an effect, this could have downstream implications in disease relevant cellular processes such as the amyloid processing pathway or

endocytic function. This model would provide a tool to study the effect on such processes in relation to this genotype.

1.5 Gene-editing as a way to create cellular isogenic models

There are a number of biotechnologies that have the aim of direct genome editing. Zinc finger (ZFN) and transcription activator-like effector nuclease (TALEN) strategies link endonuclease catalytic domains to DNA binding proteins to induce double strand breaks (DSB) at specific genomic loci (563). The latest development in gene editing technology is CRISPR-Cas9 which utilises a RNA guided nuclease to direct cleavage activity to specific loci (564).

CRISPR (Clustered regularly interspaced short palindromic repeats)- Cas (CRISPR-associated protein) is an adaptive multistep defense mechanism found in bacteria and archaea. The CRISPR-Cas defense mechanism is able to integrate short foreign DNA sequences, termed spacers, into the host genome at specific locations within the CRISPRs. Spacers are transcribed into non-coding RNAs, which with the help of specific Cas protein complexes are able to target foreign genetic material with the same or similar genetic sequence (viruses or plasmids), resulting in the degradation of invading material. This system can be manipulated to target specific DNA sequences of choice and induce DNA strand breaks at this location. (565)

CRISPR-Cas9 has become the method of choice for gene editing over previous technologies due to its high specificity, efficiency, ease and cost effectiveness. CRISPR-Cas9 is able to target almost anywhere in the genome as target selection is only limited by the Protospacer Adjacent Motif (PAM) sequence, which on average occurs every 8 bp (566). Targeting different loci can be achieved by altering the 20 bp protospacer of the guide RNA and leaving the

protein component unchanged, making this a more flexible tool than previous technologies, such as TALENs and Zinc fingers.

The Cas9 nuclease induces a DSB by cutting both DNA strands at a targeted location. Cas9 requires the action two conserved domains in order to cleave both DNA strands and create a DSB. The HNH and RuvC domains encode the nuclease active sites in Cas9. The HNH domain cleaves the sense strand complementary to the targeting guide RNA, whereas the RuvC domain cleaves the antisense strand (567). Mutating key catalytic residues within these domains can create a Cas9 nickase (Cas9n) enzyme, capable of cleaving only one DNA strand. The D10A Cas9 variant has an alanine substitution that inactivates the RuvC domain, leaving only the HNH domain functional to cut the sgRNA complementary strand (568). A double nickase approach uses two Cas9n enzymes simultaneously targeting opposite DNA strands to create a DSB. Following a DSB, endogenous repair mechanisms work to repair the break. By providing a DNA repair template with the desired modification, mechanisms that repair DNA via homology driven repair use this DNA template which can result in the inclusion of your genetic modification into the endogenous genome (569).

Ran *et al* investigated a number of variables in order to optimise the double nickase approach. SgRNA pairs that created a 5' overhangs and overlapped with less than 8 bp were consistently able to induce indels at multiple genomic loci at levels comparable to the wildtype Cas9. A greater than 100-fold increase in specificity was observed when using the nickase approach compared to Cas9. The nickase approach was found to induce homology driven repair at comparable levels to Cas9 (570). Due to these observations a double nickase approach was taken when trying to knock in rs59335482.

1.6 Aims

This chapter aims to genetically modify the endogenous genome of the H4 neuroglioma cell line, via a double CRISPR nickase approach, to contain the insertion allele of rs59335482. This isogenic model would be available to be used in downstream studies to characterise the variant.

H4 cells are derived from human neuroglioma cells and have been commonly used in AD studies (571-573). The H4 cell line was first described in 1974 and originated from a 37 year old Caucasian male diagnosed with a neuroglioma (372). Furthermore, as histone modifications typical of regulatory elements were identified within the proximity of rs59335482 in a number of brain tissues, it is possible that these regions are conserved in this brain derived cell line and therefore rs59335482 may be located in a regulatory active region of this genome.

2. Methods

2.1 Sequencing of wild type H4 neurogliomas

A 152 bp region containing rs59335482 was sequenced in the wildtype H4 neuroglioma cell line in order to determine the genotype prior to genetic modification. DNA was extracted from unmodified H4 neuroglioma cell line via the QIAamp DNA mini kit (Qiagen) [Method described in Chapter 2.4.1].

The region surrounding rs599335482 was amplified via PCR using the primer sequences described in appendix table 3.2. Appendix table 3.3 shows the PCR conditions and the thermocycler conditions are shown in appendix table 3.4.

Appendix Table 3.2. Primers used to amplify the region containing rs59335482

Forward sequencing primer 5'-3'	CCACCAAACCCAGCTAAT
Reverse Sequencing Primer 5'-3'	CAGGTGTGGTGGTTCGTA

Appendix Table 3.3. PCR reagents for rs59335482 sequencing

Reagents	Volume (µL)
Buffer	1.2
ddH$_2$O	4.66
dNTPs (2mM)	0.96
Primers (5pmol/µL)	0.56
Hot Star Taq	0.06
DNA	4

Appendix Table 3.4. Thermocycler conditions used to amplify the region surrounding rs59335482

Step	Temperature (°C)	Time (Seconds)	Cycle
1	95	600	
2	95	30	
3	58	60	
4	72	30	Repeat step 2-4 two times
5	95	30	
6	56	60	
7	72	30	Repeat step 5-7 two times
8	95	30	
9	54	60	
10	72	30	Repeat step 8-10 29 times
11	72	600	

PCR product size was confirmed via gel electrophoresis. 10 µL of PCR product was purified using the AMPure PCR purification system (Agencourt®) using Biomek® NXP Laboratory Automation Workstation (Beckman Coulter). 5 µL of AMPure cleaned product was combined with 5µl of forward or reverse primer (5 pmol/µL). This was sequenced by GATC Biotech's LIGHTrun™ Sequencing service.

2.2 Creation of CRISPR/Cas9n targeting plasmid

A double nickase gene editing approach, which predominantly follows the protocol by Ran *et al*, 2013 was used in this study (570). The CRISPR plasmid used was pSpCas9n(BB)-2A-Puro (PX462) purchased from Addgene.org (Addgene plasmid # 48141). This plasmid is 9200 bp in length and contains the

selectable marker of puromycin resistance and induces bacterial ampicillin resistance. This plasmid encodes the Cas9n gene, containing the D10A mutation.

2.2.1 Verifying PX462

PX462 was isolated from *E.coli* using the QIAprep Spin Miniprep Kit (Qiagen) [described in Chapter 2.4.2]. The restriction enzymes described in appendix table 3.5 were used to validate the plasmid obtained.

Appendix Table 3.5. Restriction enzymes used in the validation of PX462

Restriction Enzyme	DNA fragment sized produced from digestion of PX462 (bp)
Not1	9175
SacII	2414 + 6761
EcoRI	669 + 8506

Restriction enzyme digest reagents are listed in appendix table 3.6. Note that different buffers were compatible with each enzyme (appendix table 3.7). The reaction was incubated at 37°C for two hours. The digestion products were run on a 1% agarose gel to confirm the size of the DNA fragments.

Appendix Table 3.6. Restriction enzyme reaction reagents

Reagent	Volume µL
DNA	6
Buffer	2
Restriction Enzyme	1
Water	11
Total	20

Appendix Table 3.7. Compatible buffers used with each restriction enzyme

Enzyme	Buffer
EcoRI (NEB, cat. no. R0101S)	EcoRI buffer (NEB, cat. no. B7200S)
NotI (NEB, cat.no. R0189S)	Buffer 3.1 (NEB, cat. no. B7200S)
SacII (NEB, cat.no. R0157S)	CutSmart buffer (NEB, cat. no. B7200S)

2.2.2 Designing sgRNAs to target rs59335482

Short guide (sg) RNA sequences were designed using a web based guide RNA design tool, found at crispr.mit.edu, developed by the Zhang lab (574). Sequences were selected based on the nickase analysis output. High quality pairs that had a distance between targeted DNA sequences of 0-20bp (optimal offset) and GC content greater than 30% were selected. [A full description of potential sgRNA sequences is provided in appendix table 4.1].

Cloning oligonucleotides were designed in order to create the duplexes required for cloning. Sense and antisense oligonucleotides of each target sequence containing BpiI compatible sticky ends allowing for ligation into PX462 were created. An appended guanine base was incorporated on the sense strand and a cytosine on the antisense strand as the U6 RNA polymerase

prefers a guanine nucleotide as the first base of its transcript (575). [The final cloning oligonucleotide designs are shown in appendix table 4.2].

Single stranded DNA oligonucleotides (ssODN) were designed to act as a repair template resulting in the incorporation of the rs59335482 insertion (appendix table 3.8). SsODNs have a homologous sequence to that surrounding rs59335482 and therefore can be used as a template to drive homologous directed repair (HDR). All oligonucleotides were ordered from Eurofins Genomics (Ebersberg, Germany).

Appendix Table 3.8. ssODN repair template sequences incorporating rs59335482 (shown in bold).

Sense ssODN (5'-3')	CACCAAACCCAGCTAATTTTTTTTATTATTTTTTGTAGAG ATG**GGG**GGGGTCTCACTAAGCTGTCCAGGCTGGTCTT GAACTCCTGGCC
Antisense ssODN (5'-3')	GGCCAGGAGTTCAAGACCAGCCTGGACAGCTTAGTGA GACCCC**CCC**CATCTCTACAAAAAATAATAAAAAAAATTA GCTGGGTTTGGTG

2.2.3 Cloning sgRNA sequences into PX462

Annealing and Phosphorylating Cloning Oligonucleotides

The sense and antisense sgRNA oligonucleotides were annealed forming a DNA duplex with overhanging sticky ends for cloning into PX462. Oligonucleotides synthesised by conventional solid-phase phosphoramidite methods have a 5' hydroxyl group. In order for successful ligation into the PX462 plasmid, the oligonucleotides must have a 5' phosphate group in order to ligate to the 3' unmodified end of digested PX462.

The annealing and phosphorylation of the oligonucleotides was performed in one reaction and the reagents are described in appendix table 3.9. The reaction was incubated with the following temperature parameters: 37°C for 30 minutes; 95°C for 5 minutes to allow sufficient denaturation of the oligonucleotides; 95°C decreasing 5°C per minute until it reaches 25°C allowing for accurate annealing. Once the reaction was completed, the phosphorylated and annealed oligonucleotides were diluted 1:200 in nuclease free water.

Appendix Table 3.9. Reagents used in the annealing and phosphorylation of the sgRNA cloning oligonucleotides.

Reagent	Amount per reaction (μL)
sgRNA sense (100μM)	1
sgRNA antisense (100μM)	1
T4 ligation buffer, 10X (NEB, cat. no. B0202S)	1
T4 PNK (NEB, cat. no. M0201S)	1
ddH$_2$0	6
Total	10

Cloning of sgRNA duplexes into PX462

The digestion of PX462 by BpiI to create the cloning site (illustrated in appendix figure 3.2) and the ligation of the sgRNA duplexes into this cloning site was performed in one reaction. The duplex sgRNAs were ligated into PX462. A no-insert, PX462 only negative control ligation reaction was carried out. This digestion and subsequent ligation reaction was incubated at 37°C for 5 minutes, 21°C for 5 minutes, which was cycled 6 times (reagents are listed in appendix table 3.10).

Appendix figure 3.2. Plasmid map of PX462 showing the recognition sites of restriction enzyme BpiI. Digestion of BpiI will result in overhanging ends, compatible with ligation with the sgRNA duplexes.

Appendix Table 3.10. Ligation reaction reagents required for successful ligation of sgRNA duplex into PX462.

Reagent	Amount per reaction (µL)
PX462 (100 ng)	x
Diluted oligo duplex	2
Tango Buffer, 10X (Thermo Fisher Scientific)	2
DTT, 10 mM (Thermo Fisher Scientific)	1
ATP, 10mM (NEB)	1
FastDigest BpiI (Thermo Fisher Scientific)	1
T7 Ligase (Enzymatics, MA, US)	0.5
ddH$_2$0	To 20
Total	20

Once completed, the product was treated with PlasmidSafe™ DNase (Epicentre Biotechnologies, MI, USA) to digest any residual linearised DNA (reagents described in appendix table 3.11). The PlasmidSafe™ DNase reaction was incubated at 37°C for 30 minutes allowing for the enzyme to degrade any contaminating chromosomal dsDNA into deoxynucleotides. This

was followed by an incubation at 70°C for 30 minutes to deactivate the enzyme.

Appendix Table 3.11. PlasmidSafe™ DNase purification reaction reagents

Reagent	Amount per reaction (µL)
Digestion and Ligation Reaction Product	11
PlasmidSafe Buffer (Epicentre Biotechnologies)	1.5
PlasmidSafe ATP-Dependent DNase (Epicentre Biotechnologies)	1
ATP, 10mM (NEB)	1.5
Total	15

Transformation of PX462sgRNA into Chemically Competent Stbl3 E.coli

50 µL of One Shot® Stbl3TM chemically competent cells (Thermo Fisher Scientific) was thawed on ice for each transformation. 2 µL of plasmid DNA from the PlasmidSafe DNA purification reaction was added into the vial of cells and gently mixed. pUC19 control DNA (10 pg/µL) was used as a positive control for transformation. A negative control was performed using the no-insert PX462 negative ligation reaction. The vials were incubated on ice for 30 minutes. The cells underwent heat-shock by incubating the vials at 42°C for 45 seconds and then placed on ice for two minutes. 250 µL of pre-warmed S.O.C media (Thermo Fisher Scientific) was added to each vial. For the pUC19 control, the transformation mix was diluted 1:10 with LB media. The vials were incubated horizontally at 37°C for one hour in a shaking incubator. 25-100 µL of each transformation reaction was spread onto a pre-warmed selective LB agar plate (containing ampicillin 100 µg/mL) and incubated ON at 37°C. Plasmids that have successfully incorporated the duplex DNA are termed PX462sgRNA.

2.2.4 Confirming the creation of PX462sgRNAs

Plasmids were isolated from transformed minicultures by miniprep (QIAprep Spin Miniprep Kit, Qiagen). An EcoRI restriction enzyme digest was performed on plasmids isolated from transformed E.coli to confirm the presence of PX462sgRNA.

Following the successful ligation of the sgRNAs into PX462, the plasmid will be a length of 9178 bp. EcoRI digests are expected to produce DNA fragments of lengths of 669 bp and 8509 bp. The DNA products of this digest were run on a 1% agarose gel for analysis.

Purification and Sanger Sequencing of PX462sgRNA plasmids

Sanger sequencing was performed on plasmids thought to contain the sgRNA duplexes to confirm the presence of the sgRNA sequence within PX462 at the desired location. Sequencing primer sequences are shown in appendix table 3.12. Their location in the plasmid in relation to the sgRNA incorporation site is shown in appendix figure 3.3.

10 μL of each plasmid samples was purified for sequencing using AMPure PCR purification system (Agencourt®) using Biomek® NXP Laboratory Automation Workstation (Beckman Coulter). 5 μL of AMPure cleaned product was combined with 5 μL of forward or reverse primer (5 pmol/μL). This was sequenced by GATC Biotech's LIGHTrun™ Sequencing service.

Appendix Table 3.12. Sequencing Primers for Sanger sequencing of PX462-sgRNA

Forward Sequencing PX462-sgRNA Primer (5'-3')	GAGGGCCTATTTCCCATGATTCC
Reverse Sequencing PX462-sgRNA Primer (5'-3')	TTTGTCTGCAGAATTGGCGC

<div style="border:1px solid">

gagggcctatttcccatgattccttcatatttgcatatacgatacaaggctgtta
gagagataattggaattaatttgactgtaaacacaaagatattagtacaaaatac
gtgacgtagaaagtaataatttcttgggtagtttgcagttttaaaattatgtttt
aaaatggactatcatatgcttaccgtaacttgaaagtatttcgatttcttggctt
tatatatcttgtggaaaggacgaaacaccNNNNNNNNNNNNNNNNNNNNNNNgtttt
agagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaaaagt
ggcaccgagtcggtgctttttgttttagagctagaaatagcaagttaaaataag
gctagtccgttttttagcgcgtgcgccaattctgcagacaaatggctctagaggta

</div>

Appendix figure 3.3. 5'-3' sequence of PX462 spanning the cloning site for the sgRNA follwing BpiI digestion and sucessful incorporation of sgRNA sequences. Red indicates the sequence and location of the forward primer. Blue indicates the sequence complimentary to the reverse primer and relative location. N indicate the loccation of sgRNA incorporation, this sequence will differ for each of the different sgRNA designs.

2.3 PX462 transfection into H4 neuroglioma cell line

2.3.1 Transfection Optimisation

In order to gain the optimum transfection efficiency, initial experiments transfecting PX462 without cloning modifications and a GFP expressing plasmid were performed in order to optimise a number of experimental factors. GFP expression and puromycin resistance were used as markers for transfection efficiency.

The experimental factors to be optimised were the plating density of the cells, the amount of Lipofectamine 3000 reagent (Thermo Fisher Scientific) used, the concentration of plasmid DNA transfected and the optimal concentration of puromycin needed for efficient selection. [Lipofectamine 3000 transfection method is described in Chapter 2.2.2]

### 2.3.2 Co-Transfection of PX462-sgRNAs and ssODN

Once the transfection reaction was optimised, H4 cells were transfected in 12 well plates with three technical replicates per transfection. Cells were seeded at a density of 2.5×10^5 cells/well and incubated for 24 hours prior to transfection.

Co-transfections were performed with the left and right targeting plasmids and either sense or antisense repair template. Transfections were performed using 3 µL of Lipofectamine 3000 reagent per well following the protocol stated in Chapter 2.2.2. GFP plasmid was used as a positive control and as a visual measure of transfection efficiency. Cells were transfected with a total 1 µg of plasmid DNA (both left and right targeting plasmids) and 1 µL of either sense or antisense ssODN repair template (10 µM). In some instances, the PX462sgRNA plasmid transfected had been further treated with Plasmid-Safe™ ATP-dependent DNase as previously described. Negative controls were performed by following the Lipofectamine 3000 reagent protocol but without using DNA. The combinations of plasmids and guides co-transfected are shown in appendix table 3.13.

Appendix Table 3.13. Combinations of targeting plasmids and repair template used in CRISPR transfections.

Transfection Name	Right targeting plasmid (500 ng/well)	Left targeting plasmid (500 ng/well)	ssODN (10 nM/well)
Sg1(PD) sense	Sg1 (plasmid safe treated)	Sg1 (plasmid safe treated)	sense
Sg2 sense	Sg2	Sg2	sense
Sg2 antisense	Sg2	Sg2	antisense
Sg2 (PD) antisense	Sg2 (plasmid safe treated)	Sg2 (plasmid safe treated)	antisense

A schematic of the Cas9n targeting the DNA and generating double strand breaks using nickase pairs sg1 and sg2 are illustrated in appendix figure 3.4.
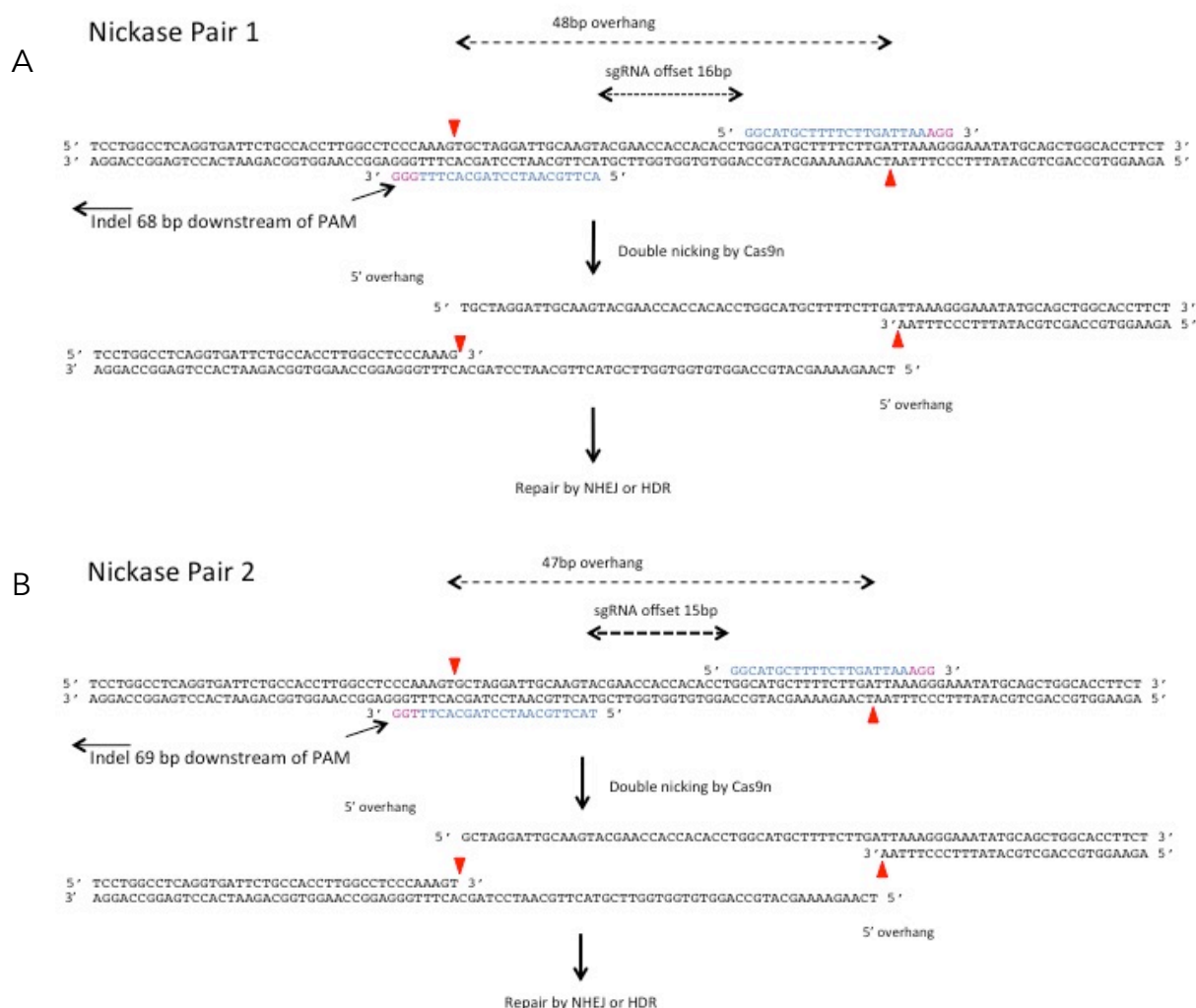
Appendix figure 3.4. Schematic diagram of DSB generation via a nickase approach with two different targeting pairs. A. Nickase pair 1 is used in transfection sg1(PD) sense. Diagram shows the sequence of the guide RNAs, where their complementary sequence is and where the nicking of the DNA strand will occur. B) Nickase pair 2 is used in transfections sg2 sense, sg2 antisense and sg2(PD) antisense. The guide sequences are shown binding to their target sequence and the location of the DNA nick is shown. Red triangles indicate Cas9n cut sites.

Puromycin was used to select for cells expressing PX462. Puromycin was added to all wells 24 hours post-transfection at a final concentration of 4 μg/mL. The transfection media was replaced with puromycin containing media 48 hours post-transfection. Cells were incubated in puromycin for at total of 72 hours.

2.3.3 Recovery of Remaining Cells

Following 72 hours of puromycin selection, cells were grown in Opti-MEM® Reduced Serum Media supplemented with 10% FBS and cultured at 37°C with 5% CO_2.

Cells not transfected with PX462sgRNA did not survive the selection procedure, as demonstrated by the negative controls. Once cells resistant to puromycin began to reach confluency, they were trypsinised and this mixed population of cells was expanded. DNA was extracted from these mixed populations of cells using Qiagen QIAamp DNA mini kit.

2.3.4 Single cell clonal populations

Post sg1(PD) sense transfection, cell populations that survived puromycin selection were plated out in an attempt to gain cellular colonies originating from single cells, where each resulting cell population has a common genome sequence.

Cells were counted and underwent serial dilutions in culture medium in order to create a cellular concentration of 5 cells/mL. 100 μL of this solution was seeded into 96 well plates and incubated at 37°C with 5% CO_2.

Four days post seeding, cell colonies became visible using light microscopy and it was determined whether wells contained single colonies. Wells containing multiple colonies or no colonies were discarded. Wells containing single colonies were expanded. Following expansion, DNA was extracted (using Qiagen QIAamp DNA mini kit) and cells stored in liquid nitrogen. 56 single cell colonies were established from transfection sg1(PD)sense.

2.4 Confirmation of Gene Editing

A number of techniques were implemented in order to determine whether the genetic modification had successfully been incorporated.

2.4.1 Surveyor Assay

Transgenomic® (USA) produce the Surveyor Mutation Detection Kit for standard gel electrophoresis, which can be used in order to locate polymorphisms within heteroduplex DNA. It utilises the Surveyor nuclease that specifically cuts at sites of base substitution, indel mismatches or similar variation resulting in cleavage of both DNA strands 3' of the mismatch (576). The subsequent DNA fragments can then be analysed via gel electrophoresis. This assay was used to identify the presence of the insertion in the DNA isolated from the mixed cellular populations that had undergone puromycin selection.

Primers were designed to amplify the DNA region surrounding rs59335482 producing a 500bp amplicon. Should the Surveyor nuclease cut in the location of the desired modification, DNA fragments of 150 bp and 350 bp is size would be produced. Primer sequences are shown in appendix table 3.14.

Appendix Table 3.14. Primers used to amplify SNP region for use in Surveyor assay

Left Primer 5'-3'	CACCTCCCCTAAAGAGCCTG
Right Primer 5'-3'	GGAAGAGAAGGTGCCAGCT

The PCR was carried out using the high fidelity Herculase polymerase (Agilent Technologies, Inc, USA). The PCR reagents are shown in appendix table 3.15 and the thermocycler conditions are described in appendix table 3.4.

Appendix Table 3.15. PCR reagents used in PCR for Surveyor Assay

Reagent	Volume per Reaction (µL)
Herculase Buffer	10
dNTP mix (25 mM)	0.5
Primers (10 µM)	2.5
Herculase Polymerase	0.5
DMSO	1
DNA	≤100 ng
Total Volume	50

This PCR protocol was used to amplify DNA extracted from cells, which had undergone the CRISPR transfection and puromycin selection (sample DNA), in addition to DNA obtained from wild type H4 cells (reference DNA). PCR reactions were performed in triplicate and then combined and purified using Qiagen PCR purification kit in order to obtain a DNA concentration >25 ng/µL.

DNA from cell populations that had undergone transfections using 3 different plasmids combinations was used for this assay. The plasmid pair and repair template used in each transfection are listed in appendix table 3.16.

Appendix Table 3.16. Transfections used Surveyor Assay

Transfection	Plasmid Pair	Repair Template
1	Sg1 (PD)	Sense
2	Sg2	Sense
3	Sg2	Antisense

Each transfection was performed in triplicate; therefore three cellular populations arose from each transfection following the puromycin selection. Each cellular population was analysed individually with the Surveyor assay.

The PCR product DNA concentrations were normalised to approximately 20 ng/µL. 9 µL of reference DNA PCR product, 9 µL of sample DNA PCR product and 2 µL of PCR buffer were combined. These samples were hybridised to each other with the intention to form heteroduplex DNA. Reference DNA was hybridised alone in order to act as a reference control. The tubes were then placed into a thermocycler and underwent the programme described in appendix table 3.17 in order to hybridise the PCR products.

Appendix Table 3.17. Thermocycler programme used for hybridisation step

Temperature (°C)	Time (seconds)
95	600
95 to 85	(-2.0°C/s)
85 to 75	(-0.3°C/s)
75°C	60
75 to 65	(-0.3°C/s)
65	60
65 to 55	(-0.3°C/s)
55	60
55 to 45	(-0.3°C/s)
45	60
45 to 35	(-0.3°C/s)
35	60
35 to 25	(-0.3°C/s)
25	60

Following hybridisation, each sample was digested with the Surveyor nuclease. The reaction components were added in the following order: 20 µL of annealed duplex, 2.5 µL MgCl, 0.5 µL ddH$_2$O, 1 µL of Surveyor Enhancer S and 1 µL of Surveyor nuclease S. The reaction was gently pipette mixed then incubated at 42°C for 30 minutes. Upon completion, 2 µL of Stop solution was added and mixed. The reaction products were ran on a 1% gel and visualised using a UV transilluminator.

A control experiment was performed alongside the test experiment which used plasmids provided with the Surveyor Mutation Detection Kit. These plasmids contain inserts that differ by one base pair. The Control G plasmid contains a guanine where the Control C plasmid contains a cytosine. PCR amplification

310

using the primers provided produces a fragment of 633bp. A negative control was performed where Control C PCR product was hybridised to another Control C PCR product. Control G and Control C were hybridised together. Digestion of this product with the Surveyor nuclease should yield products of 217bp and 416bp in length. This experiment was performed following the manufacturers instructions.

### 2.4.2 Sanger Sequencing of rs59335482

DNA surrounding the modification site was amplified using primers described in appendix table 3.2. The PCR reaction and thermocycling conditions have been previously described in appendix table 3.3 and 4. The resulting PCR product is 152bp in length. The PCR product is ran on a 1% agarose gel, the appropriate sized band was excised and the DNA is extracted from the gel with the use of a QIAquick Gel Extraction Kit (Qiagen) and purified using a QIAquick PCR purification Kit [methods described in Chapter 2.4].

Purified DNA was sent to Genewiz® (Takely, UK) where samples were sequencing using their standard sanger sequencing reaction conditions with the reverse sequencing primer (5'CAGGTGTGGTGGTTCGTA3'). Genewiz® Sanger sequencing was performed on DNA from mixed populations of cell following puromycin selection and on single cell populations originated from the Sg1(PD) sense transfection.

3 Results

## 3.1 H4 cell line does not contain the insertion

Sanger sequencing revealed the wildtype H4 cells line did not contain the 3 base pair insertion and therefore the attempt to induce this variant could proceed (appendix figure 3.5).
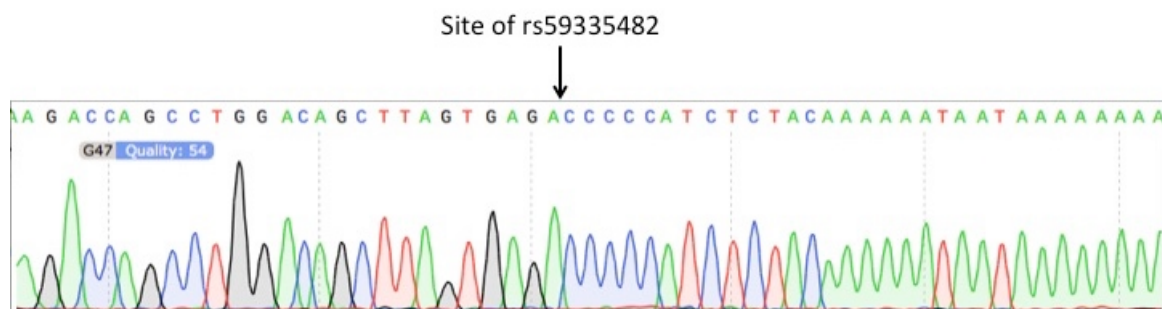

Appendix figure 3.5. Chromatogram showing the DNA sequence of the wildtype H4 cell line. Sequencing indicates the absence of the rs59335482. **insertion.**

3.2 PX462sgRNAs were successfully generated

Restriction enzyme digests using enzymes EcoRI, SacII and NotI verified the presence of PX462 (appendix figure 3.6). Following cloning of the guide RNAs EcoRI digestion indicates the presence of PX462 in transformed *E.coli* (appendix figure 3.7).
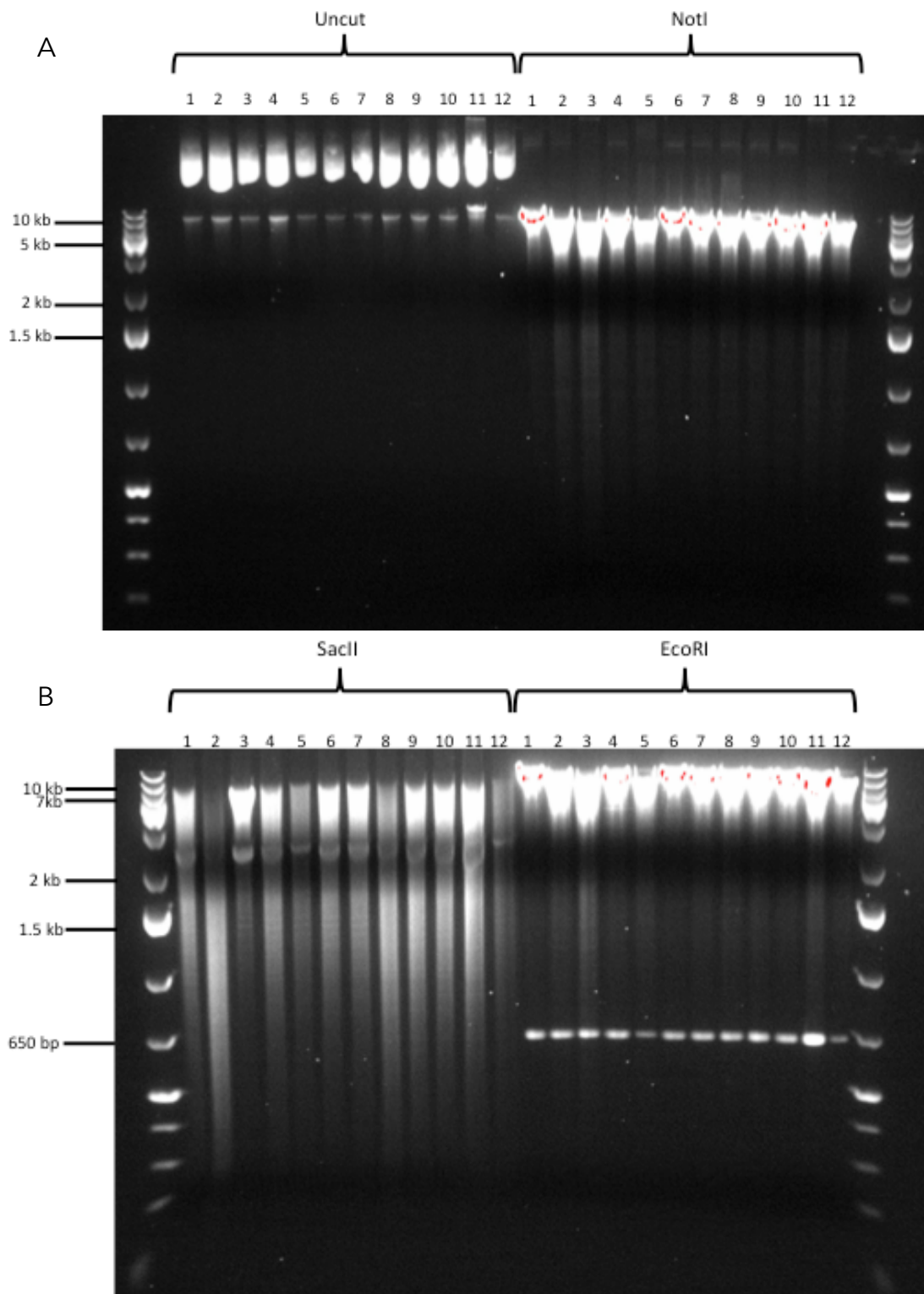
Appendix figure 3.6. Agarose gel containing the restriction enzyme digestion products. A. Uncut PX462 is run alongside PX462 digested with Not1. Not1 linearises PX462, creating a fragment of 9175 bp. B. PX462 digested with SacII created fragments of length 2414 bp and 6761 bp. PX462 digested by EcoRI produced fragments sized 669 bp and 8506 bp. These gel images show the correct size DNA fragment assuming successful digestion of PX462, therefore verifying the presence of PX462.
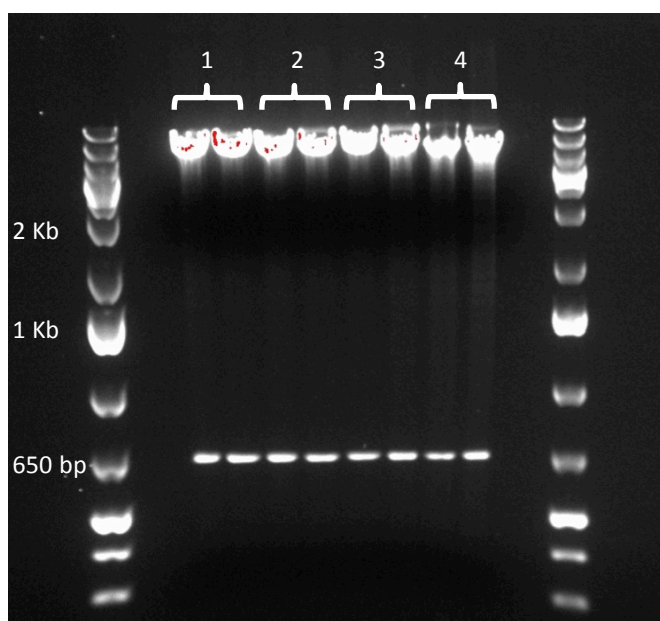
Appendix figure 3.7. Image of 1% agarose gel of PX462-sgRNA minipreps digested with EcoRI restriction enzyme. Lanes marked 1: two separate minipreps containing PX462sgR1. Lanes marked 2: two separate minipreps containing PX462sgL1. Lanes marked 3: two separate minipreps containing PX462sgR2. Lanes marked 2: two separate minipreps containing PX462sgL2. Expected DNA fragment sizes are 669 bp and 8509 bp. The gel image shows fragment sizes of approximately these sizes.

Sanger sequencing of isolated plasmids led to the confirmation of the successful incorporation sgRNA sequences into the PX462 plasmid. Six sequence-verified sgRNA pairs were available for use in further experiments. A list of PX462-sgRNA plasmids that have been verified with sequencing results is shown in appendix table 3.16. A chromatogram of PX462sgRNA-1R (shown in appendix figure 3.8) confirms the successful cloning of this sgRNA into PX462.

Appendix Table 3.16. Sequenced verified PX462-sgRNA plasmids

| PX462-sgRNA plasmids |
| --- |
| Sg1R |
| Sg1L |
| Sg2R |
| Sg2L |
| Sg4R |
| Sg4L |
| Sg6R |
| Sg6L |
| Sg7R |
| Sg7L |
| Sg8R |
| Sg8L |
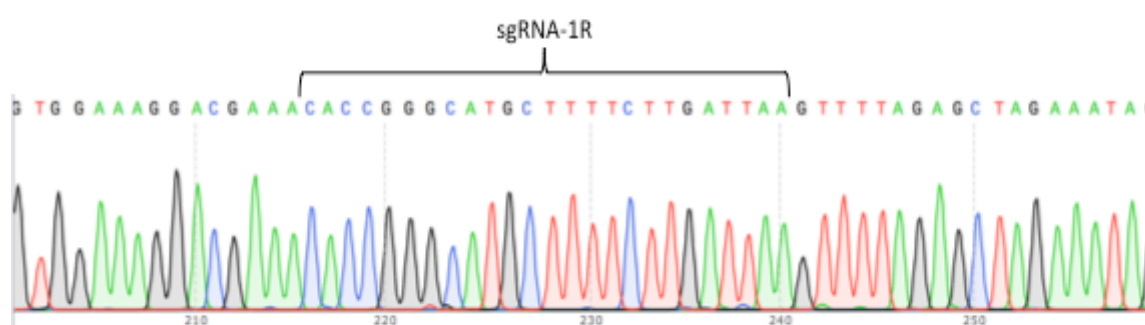
Appendix figure 3.8. Representative Chromatogram obtained from Sanger sequencing the PX462 plasmid, with the Nickase1-right designed sgRNA cloned into the Bpi1 cloning sites. Primers were designed to sequence the sense strand with the expected sequence CACCGGGCATGCTTTTCTTGATTAA . The presence of this sequence in the chromatogram confirms the successful cloning of the sgRNA into PX462. Complementary sequencing of the reverse strand was also performed to further verify the presence of the sgRNA sequence.

3.3 Co-transfection of PX462sgRNA and ssODN into H4 neuronal cell line

3.3.1 Transfection Optimisation

During optimisation, optimal transfection efficiency of approximately 50% was observed when using a cell density of 2.5×10^5 cells per well, 1 µg GFP plasmid, 3 µL of Lipofectamine 3000 reagent and 4 µL P3000 enhancer reagent per well of a 12 well plate therefore these transfection conditions were used to transfect the targeting plasmids (appendix figure 3.9).

A) GFP expression 24 hours post transfection
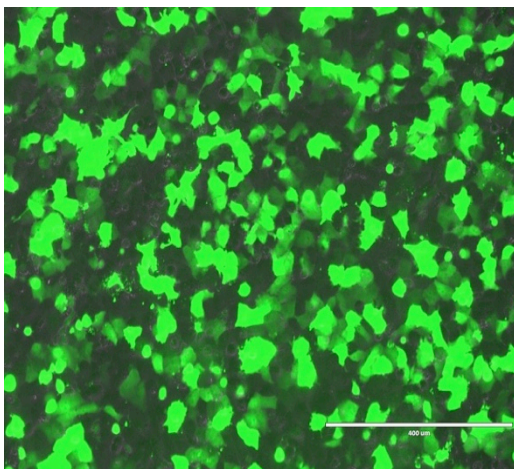
B) GFP expression 48 hours post transfection
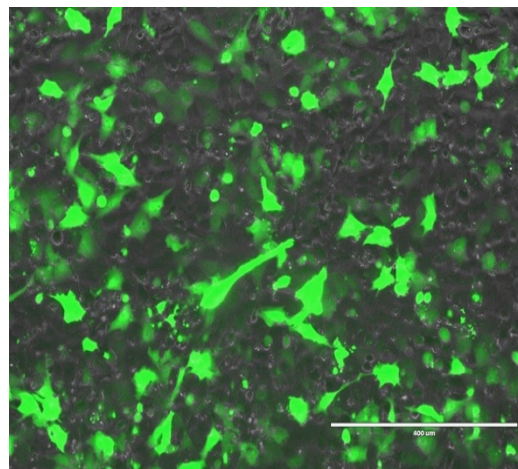
Appendix figure 3.9. Transfection of GFP plasmid. Overlay images taken at a visible light wavelength and 488 wavelength. A) Shows approximately 50% of cells expressing the GFP plasmid 24 hours after transfection. B) Showing expression of GFP in approximately 20% of cells 48 hours post transfection.

3.3.2 Transfection of PX462sgRNA

Targeting plasmids were co-transfected. Cells that survived puromycin selection were expanded. DNA extracted from these cells was screened for the presence of rs59335482 insertion allele. Cellular dilutions were carried out on the sg1(PD) sense population and single cell colonies obtained were expanded. DNA was extracted from these single cell populations was sequenced to screen for the presence of rs59335482 insertion allele.

3.4 Surveyor Assay

3.4.1 Amplification of BIN1 region surrounding rs59335482

Non-modified wild type H4 DNA and DNA extracted from cells that had successfully survived puromycin selection following transfection with the CRISPR plasmids was successfully amplified via PCR and used in downstream hybridisation experiments (appendix figure 3.10 and 11, respectively).
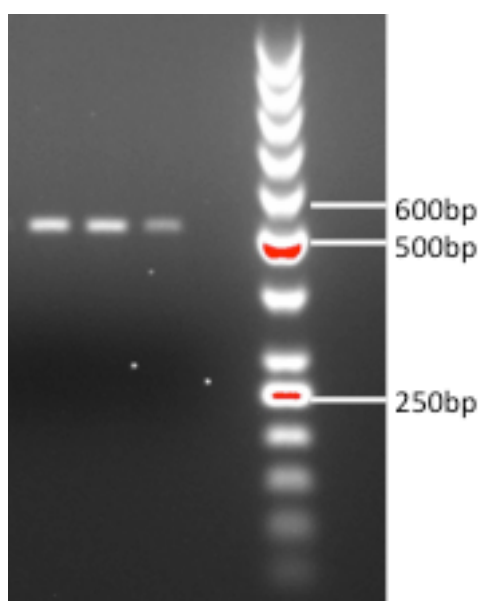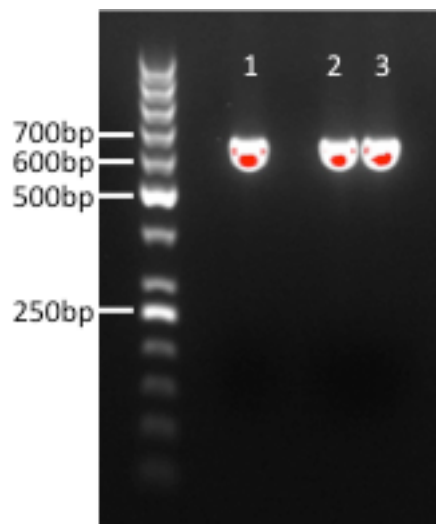

Appendix figure 3.10. Amplification of wild type rs59335482 region. Gel electrophoresis using 1% agarose performed using the PCR product from wild type H4s. The PCR product produced is approximately 500bp and was therefore used in downstream hybridisation experiments. PCR performed in triplicate and each product was run individually.
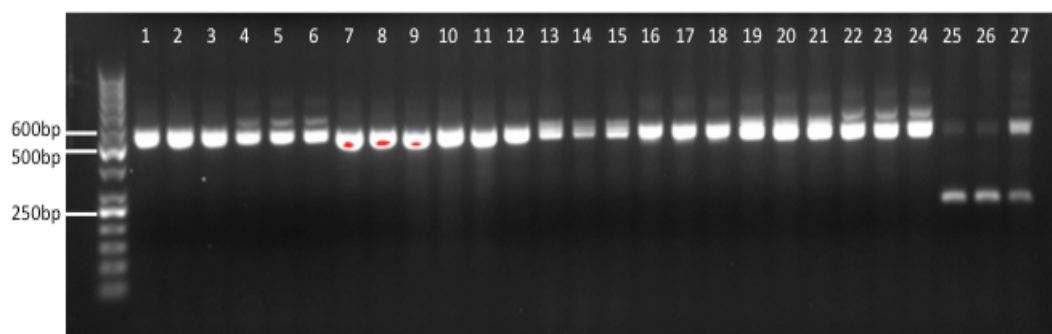
Appendix figure 3.11. The PCR product amplified from transfected cells following amplification of the region surrounding rs59335482. PCR products were electrophoresed on a 1% agarose gel. The expected fragment size is 500bp in length. The majority of samples show the expected fragment size. Some samples shows PCR products of differing lengths, potentially indicative of gene modification or errors during PCR. Lane 1 – 3: Triplicate PCR from transfection 1 replicate 1. Lane 4 – 6: Triplicate PCR from transfection 1 replicate 2. Lane 7 – 9: Triplicate PCR from transfection 1 replicate 3. Lane 10 – 12: Triplicate PCR from transfection 2 replicate 1. Lane 13 – 15: Triplicate PCR from transfection 2 replicate 2. Lane 16 – 18: Triplicate PCR from transfection 2 replicate 3. Lane 19 – 21: Triplicate PCR from transfection 3 replicate 1. Lane 22 – 24: Triplicate PCR from transfection 3 replicate 2. Lane 25 – 27: Triplicate PCR from transfection 3 replicate 3.

3.4.2 Amplification of control plasmids

PCR amplified the variable region of the control plasmids. 2 μL of each PCR product were analysed via gel electrophoresis that confirmed successful amplification of the desired region with the expected fragment size of 633bp (appendix figure 3.12).

318

Appendix figure 3.12. Amplification of region of control plasmids containing SNP. 1% agarose gel electrophoresis showing the PCR products amplified from the Surveyor assay control plasmids. Expected fragment size is 633bp. Lane 1: PCR product using Control G plasmid as template. Lane 2 and 3: PCR product using Control C plasmid as template

3.4.3 Surveyor nuclease digestion

Following the successful amplification of the appropriate regions, hybridisation was performed followed by the Surveyor nuclease digest. The products of the digest were analysed using gel electrophoresis and imaged using a UV transilluminator (appendix figure 3.13). The digestion products of the control plasmids produced the expected size products indicating the Surveyor nuclease successfully cleaved at the location of mismatches. The presence of multiple bands in the test and reference samples is potentially indicative of numerous SNPs being present in the wild type DNA. Lane 9 has a band at around 250bp that is absent from the reference control sample. This may indicate the presence of a variant not present in the wild type DNA sequence. Lane 11 and 12 contain the positive and negative control respectively.
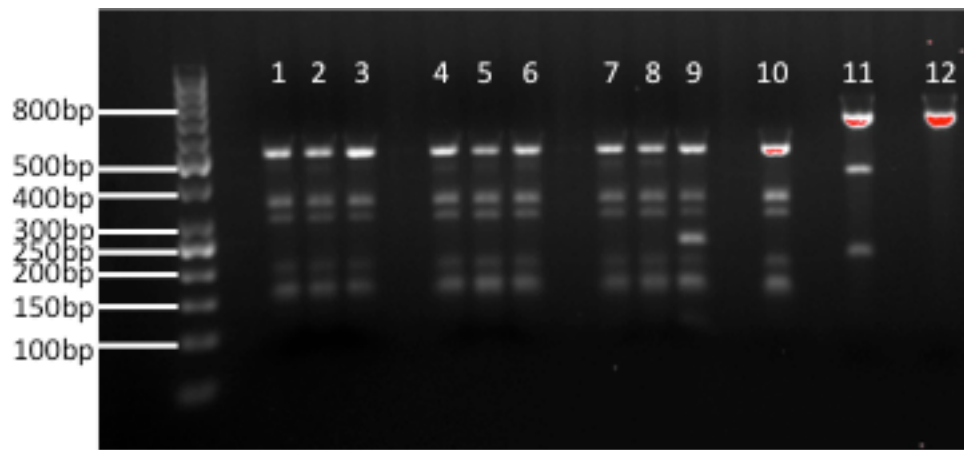
Appendix figure 3.13. Products of the Surveyor nuclease digest. Cleavage products of the Surveyor digest were separated on a 1% agarose gel. Lanes 1 – 3: Digestion products of each triplicate of transfection 1 hybridised to wild type DNA. Lanes 4 – 6: Digestion products of each triplicate of transfection 2 hybridised to wild type DNA. Lanes 7 – 9: Digestion products of each triplicate of transfection 3 hybridised to wild type DNA. Lane 10: Reference control - Contains the digestion product of wild type PCR product annealed another wild type PCR product. Lane 11: Positive Control – The digestion products of Control C PCR product annealed to Control G PCR product. This shows the expected fragment size of 633 bp (undigested product) and 217bp and 416bp (digested products). Lane 12: Negative control - The digestion products of Control C PCR product annealed to Control C PCR product. This shows the expected fragment size of 633 bp (undigested product).

3.5 Sanger Sequencing of rs59335482 region in mixed populations and single cell colonies indicate no incorporation of insertion

DNA extracted from a mixed population of transfected cells was sequenced in order to screen for the presence of rs59335482 insertion allele. Sequencing data revealed that there was no indication of the rs59335482 insertion allele present in the population (appendix figure 3.14).
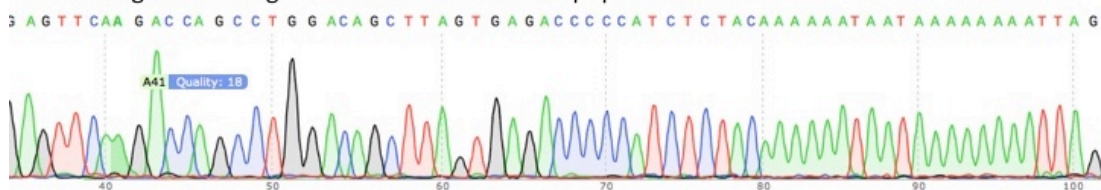
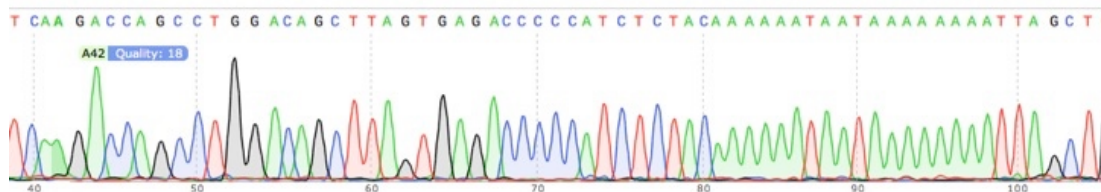Chromatogram from sg1(PD) sense mixed cellular populations
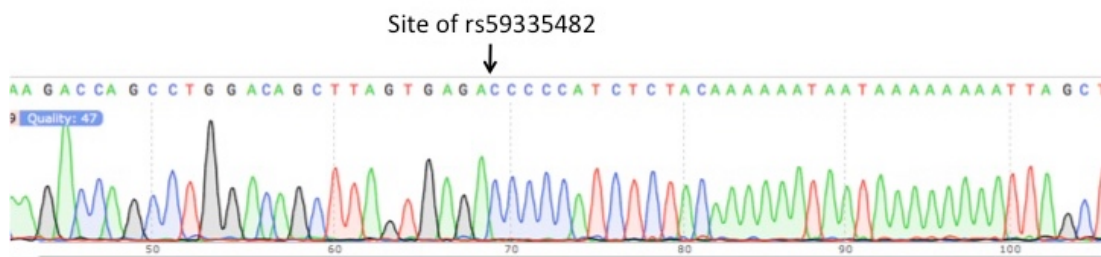
Chromatogram from sg2 sense mixed cellular populations
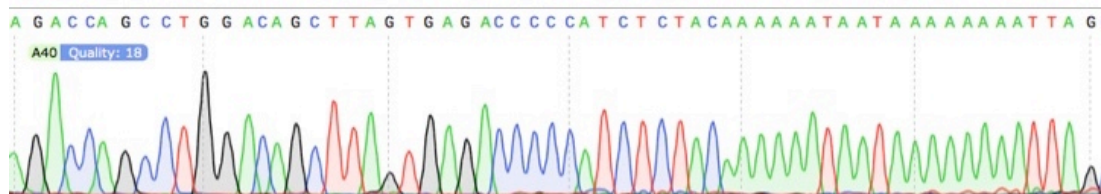
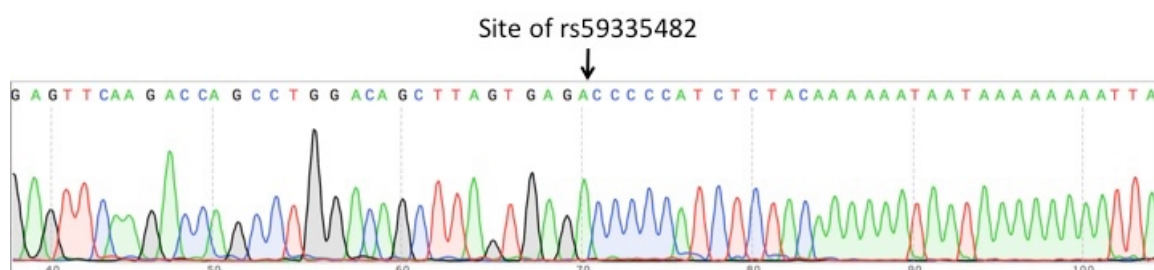Chromatogram from sg2 antisense mixed cellular populations

Chromatogram from sg2(PD) antisense mixed cellular populations

Appendix figure 3.14. Sanger sequencing of DNA extracted from mixed cellular populations following puromycin selection indicated that the rs59335482 insertion had not been incorporated. Chromatograms from four transfected cell populations sg1(PD) sense, sg2 sense, sg2 antisense and sg2 (PD) antisense showing the DNA sequence surrounding rs59335482.

In order to further confirm this as a representative result, the sequencing was repeated on single cell colonies generated from the cell populations transfected with sg1(PD) sense. This confirmed that the sequence observed in the mixed cellular populations was consistently observed in the single cell populations, indicating that the rs59335482 insertion allele was not present in any of the 56 single cell colonies sequenced (appendix figure 3.15).

Appendix figure 3.15. Representative chromatogram from the single cell colonies produced from the sg1(PD) sense transfected population. All chromatograms from 56 single cell colonies showed the same results, rs59335482 insertion had not been incorporated into these cells.

4 Discussion

This chapter aimed to establish isogenic cell models to investigate the effect of rs59335482 genotype on *BIN1*. Unfortunately, genetic modification was not achieved. The results, potential reasons as to why the modification was unsuccessful and what future directions this project could take are discussed below.

4.1 PX462 with targeting guide RNAs can be used to genetically modify the endogenous genome of cell lines

Despite not having success in knocking in the rs59335482 insertion, there have been numerous publications successfully using PX462 in CRISPR-Cas9n mediated indel generation. PX462 was used to successfully create a double knockout cell line by disrupting the two catalytic subunits of AMPK, and the authors observed a knock out modification frequency of 12.5% (577). Rahmanto *et al*, transfected PX462 using Lipofectamine 3000 into hEM3 cells and observed a deletion and an insertion on each allele of ARID1A and confirmed protein knockdown (578). PX462 was also transfected into HUES-2 stem cells via a Neon® Transfection system to knockout SAMHD-1 (579).

These studies have all used PX462 to establish gene knockout models via gene editing and therefore have validated the ability of PX462 to successfully cleave targeted DNA. However, there are currently no publications detailing the use of PX462 to knock in specific mutations.

4.2 Surveyor assay indicates the presence of additional SNPs or hybridisation errors

Mutations were searched for using an enzyme mismatch cleavage method, which utilises enzymes able to cleave heterogeneous DNA. The Surveyor nuclease cleaves DNA 3' of a mismatch or indel with high specificity (576). The Surveyor assay was used to determine whether variants were present within the CRISPR treated cells, which are absent from the wild type cell line. Multiple digestion products indicate that there were multiple cleavage sites present in the wildtype H4 DNA (appendix figure 3.13). There are a number of possible explanations as to why this may have occurred.

Vouillot *et al*, 2015 has reported optimal Surveyor activity is when heteroduplexes make up 60% of the DNA population (580). When the heteroduplex population was 5-30% they could not reliably evaluate the enzymes activity due to high background and non-specific cleavage products. If the presence of the mutation is low in the mixed cell population DNA, the Surveyor assay may not have been sensitive enough to detect it and what is being observed is non-specific cleavage products. The Surveyor enzyme also has a known 5'-exonulcease activity (581). This activity may result in DNA degradation causing further background. In combination, non-specific cleavage and exonuclease activity could be an explanation for the presence of multiple bands if differential exonuclease activity acted on cleaved products generating DNA strand length heterogeneity. Reducing the amount of Surveyor enzyme incubation time may prevent non-specific cleavage in future assays. However, as the multiple bands are consistent between different hybridisation reactions, it is unlikely that this is the sole explanation.
Due to the repetitive nature of the region of DNA, the hybridisation process may have misaligned resulting in the formation of mismatches in the hybrid

DNA. Diluting the DNA concentration to ≤50ng/μL, repeating the hybridisation step, or adding 1x PCR reaction buffer to products prior to annealing can attempt resolving sub-optimal hybridisation. However, again as the cleavage products are consistent between all samples and the reference control, this is unlikely.

If the hybridisation was accurate, the presence of multiple bands in the reference control indicates that there are sequence differences within the wild type DNA, suggesting heterozygote genotype. When studying this 500bp sequence in UCSC, it identifies 5 common SNPs found in the dbSNP build 144 (minor allele frequency (MAF) obtained from dbSNP): rs56193035 (MAF=0.11), rs61308109 (MAF=0.12), rs57109420 (MAF=0.11), rs4663105 (MAF=0.48), rs560590441 (MAF=0.31), rs59253801 (MAF=0.11). As these alternative SNPs are common in the population, it is possible that the H4 genome is heterozygous for a number of these SNPs, which has resulted in the Surveyor nuclease cleaving at a number of sites. Rs4663105, rs560590441 and rs59253801 are located 38, 1 and 46bp away from rs59335482 respectively and have therefore been Sanger sequenced during this investigation. All variants were found to be homozygous and not responsible for the cleavage patterns observed.

Further genotyping could be performed to determine the zygosity of the remaining SNPs in the H4 genome. Once the genotypes of the remaining SNPs are known, alternative primers could be designed in order to avoid heterozygous SNPs in the PCR product used in the Surveyor assay. Any cleavage of this product would be due to variants not in the endogenous genome.

In lane 9 of appendix figure 3.13, an additional band is present with a size of 250bp. As this band is absent from the reference sample and the remaining transfected cell populations, it may suggest the presence of a variant not found in the endogenous genome. This DNA sample is from the population of cells transfected with sg2 targeting plasmid and an antisense repair template (replicate 3), which during PCR amplification showed unexpected PCR products (appendix figure 3.11). The presence of the PCR product around 300bp indicates that a deletion event may have occurred during the gene-editing attempt.

When the 110 bp region surrounding rs59335482 in the sg2 antisense cell population was sequenced, no changes in DNA sequence were observed. The targeted sequences in this transfection are located 124bp and 75bp away from rs59335482, therefore it is possible an off target event may have occurred and generated a genetic modification elsewhere in this region of the genome. It may be of interest to carry out a more comprehensive investigation of this region within this cell population to determine if a modification has occurred and where. In the first instance, the 300bp fragment could be excised from the gel, sequenced and compared to the endogenous genome to elucidate any modifications.

4.3 Lack of Modification – CRISPR limitations

Following the inconclusive evidence from the Surveyor assay, mixed and single cell population sequencing revealed no modification at rs59335482.

4.3.1 Distance dependence for editing with CRISPR

It has been observed that the success of gene editing is reduced with increased distance from the cut site and is true for both Cas9 and Cas9n (582,

583). Paquet *et al* 2016 investigated the effect of distance between the intended point mutations and the cleavage site and observed an inverse relationship between rate of mutation incorporation and distance from the cleavage site. This observation was consistent between genetic loci and cell types. Therefore a distance effect may effect mutation incorporation via HDR (584).

Paquet's estimates for mutation incorporation are minimal when the cut to mutation distance approaches 45 bp. As plasmids were preferentially transfected based on the ranking in the output from the design tool, the plasmids used in this gene editing attempt cut the DNA no closer than 74 bp away from the mutation site. Based on Paquet's analysis, it could be assumed that cutting at this distance was unlikely to result in a modified genotype and could be one of the possible explanations why no gene editing was observed. Due to the repetitive nature of the sequence surrounding rs59335482 and the potential for off targeting effects, guides targeting closer to the mutation site were scored less highly than the more distant guides. Further attempts could utilise these alternative guides which may result in a higher modification efficiency, such as nickase pair seven whose right and left guide sequences cut 17 and 38bp from the mutation site respectively.

4.3.2 Reduced nickase activity

Bialk *et al*, 2015 showed that the overall level of gene editing is between 10-15 fold lower when using single a nickase enzyme in comparison with the CRISPR/Cas9 complex (583). When using a combination of nickase enzymes to create a double strand break efficient gene editing was observed, although levels remain reduced in comparison to Cas9, hence reduced Cas9n activity may be a factor in the lack of gene editing observed.

4.4 Developments and optimisation of gene editing approaches

The field of gene editing is fast evolving as researchers aim to optimise protocols and approaches in order to increase efficiency and reduce off target effects.

4.4.1 Cas9 variants that improve specificity

Variations from commonly used SpCas9 enzyme are available and possess different properties making them more amenable to gene editing. There are a number of Cas9 orthologues, which utilise a longer targeting sequence (585-587). Although a more complex PAM sequence means fewer available targets sequences, it also reduces the number of potential off target sites for any given guide.

In addition to the nickase modification, other point mutations within Cas9 have been shown to alter its function. The D1135E and VRER mutations increase SpCas9 specificity for NGG and NGCG PAMs respectively (588). Point mutations in the RuvC and HNH catalytic domains result in an inactive nuclease termed dCas9. dCas9 can be attached to various effector domains, which can be used in a number of applications. For example, one can use two dCas9 enzymes fused to a Fok1 nuclease, targeting opposite DNA strands. Only when the two fusion enzymes are bound within close proximity can the Fok1 monomers assemble to form the catalytically active Fok1 nuclease dimer cleave the dsDNA. This creates a dimerisation-dependent system improving specificity (589).

eSpCas9 has three codon substitutions within the non-targeted DNA strand groove (590). These mutations weaken non-target strand binding by the protein, causing in a reduction in genome wide off-target events whilst maintaining on target efficiency. SpCas9-HF contains four alanine substitutions within the region that interacts with the phosphate backbone of the targeted DNA, which significantly decreases genome wide off-target effects without compromising efficiency (591).

4.4.2 Optimising guide RNA design

Initial guide RNA design tools focused on designing guides that specifically target the genomic DNA with minimal homology to elsewhere in the genome. The tool used in this work, MIT CRISPR design tool, uses a specificity score described in Hsu *et al*, 2013, which takes into account mismatches in off-target sites and the effect of mismatches on cleavage efficiency, however there is evidence that off-targets sites could contain small indels or alternative PAM sequences and may still be targeted by Cas9 (574, 587). Despite this, the MIT CRISPR design tool remains to be one of the most comprehensive design tools for designing guides for a nickase approach. An improved design tool is now available is Benchling which designs guides based on Hsu *et al* specificity score and a cutting frequency determination score which more accurately predicts off target activity (592). As Cas9 specificity may vary between cell types, method of delivery and dosage, it is difficult to collate experimental data for all parameters and therefore the majority of design tools do not take them all into account.

Optimal guide length has been investigated and it was found that extending the sgRNA duplex by five nucleotides and mutating the fourth base in the continuous sequence of T bases, which follows the guide sequence and signals

for RNA polymerase III to pause, can significantly increase knockout efficiency (593, 594). It was also found that the addition of two guanine nucleotides at the 5' end of the guide termed "GGX20", demonstrated increased specificity with variable on-target efficiency, potentially due to changes in RNA structure or stability (595).

Shortening the guide to 17-18 nt was shown to increase the sensitivity of Cas9 to mismatches and more precise gene editing was observed. This may be due to the shorter guide being less tolerant to mismatches or possessing higher affinity for their target sites. When truncated guides were used in a Cas9 nickase approach, it was found to further reduce off-target events without compromising on-target efficiency (596).

Chemical modification of synthesised guides has been shown to have an effect on Cas9 specificity and enhance genome-editing efficiency. 2'-O-methyl, 2'-O-methyl 3'phosphorothioate or 2'-O-methyl 3'thioPACE modifications at the 5' and 3' end improves editing efficiency in human primary T cells and CD34+ hematopoietic stem and progenitor cells (597).

4.4.3 DNA donors optimisation

Alternative donor DNA designs have been investigated after it was found that following cleavage by Cas9 the 3' end of the non-target strand is preferentially released. An asymmetric DNA donor complementary to the non-target strand, which spanned the Cas9 cut site with 36bp on the PAM-distal side, and 91bp on the PAM proximal side was found to increase HDR frequency to approximately 57%. This effect was shown to be consistent between cell types and genomic loci. An optimised asymmetric donor resulted in a five-fold increase in HDR when using Cas9 nickase (598).

4.4.4 Inhibiting Non-homology end joining

Approaches aiming to reduce the number of off-target events and increase on-target insertions have explored the effects of inhibiting non-homology end joining (NHEJ). A number of studies have shown that inhibiting NHEJ repair can increase the number of on-target insertions created via HDR.

Chu *et al*, suppressed NHEJ by knocking down expression of vital components for successful NHEJ and using the small molecule SCR7 that inhibits DNA ligase IV function (599). They also investigated the effect of these approaches in combination with co-expressing Ad4 proteins which target DNA ligase IV for ubiquitination and proteasomal degradation.

By suppressing NHEJ, a fivefold increase in HDR in cell lines was observed, which was further increased with the co-expression of Ad4 proteins (599). This effect has been observed in a numerous cell lines, mice zygotes and at multiple loci (600, 601). This approach however can be cytotoxic or potentially mutagenic and therefore not necessarily appropriate for extensive genome engineering (602).

4.4.5 Synchronising the cell cycle

HDR occurs during the S and G2 phase as this is when sister chromatids are available to act as repair templates whereas NHEJ occurs throughout the cell cycle (603). Bialk *et al*, 2015 showed that gene-editing frequencies are increased if cells are held in the G1/S border for 24 hours and the released 4 hours before the introduction of the CRISPR/Cas9 complexes (583).

4.4.6 Cas9 delivery method and dosage

Lentiviral vectors are capable of producing stable gene expression once delivered into cells. Lentiviral-mediated delivery of Cas9 and guide RNA has been described to induce indels at on-target sites at up to 100% efficiency, however the prolonged expression of Cas9 can increase off-target events (512). Limiting Cas9 expression can improve specificity. This was observed by Hsu *et al* who, when reducing the dosage of Cas9 by five-fold, observed a seven-fold increase in specificity ratio and two-fold reduction in on-target efficiency (574).

Delivering recombinant Cas9 protein complexed with guide RNA into cultured human cells can bypass some of the limitations associated with plasmid transfection. Zuris *et al*, 2015 found that transient cationic lipid-mediated delivery of the Cas9 nuclease protein complexed with guide RNA forming a ribonucleoprotein (RNP) was highly efficient and improved the specificity of guide ten-fold in comparison to plasmid transfection (604). This increased specificity is thought to be due to a burst in Cas9 exposure when delivering RNP followed by rapid degradation of the protein resulting in reduced off-target effects (605). Zuris *et al* also used this approach to deliver Cas9 nickases and observed similar results to that seen when using plasmid transfection.

4.5 Concluding remarks

As no gene editing was observed at the desired location, optimising gene editing efficiency would be the primary objective. Utilising some of the developments discussed in this section could be beneficial, such as synchronising the cell cycle and inhibiting NHEJ may increase the likelihood of the rs59335482 insertion being incorporated. As the proximity of the guide

RNAs to the mutation has been shown to be important, using more proximal guide RNAs would be desirable and more complex design tool may aid in this. Following targeting optimisation, using an asymmetric repair template should increase the likelihood of mutation incorporation. Finally in an effort to reduce off target effects, the transient expression of the Cas9n enzymes is crucial and this could be addressed by delivering these complexes via a RNP rather than a plasmid whose expression may be more long term. Creating isogeneic rs59335482 models would aid in the characterisation of this variant and how it may influence disease mechanisms within a model brain cell.

Appendix 4. CRISPR oligo design

Appendix Table 4.1. sgRNA target sequences selected from crispr.mit.edu Nickase analysis output

Nickase Pair	Sense				Antisense				sgRNA offset
	Sequence 5'-3'	Quality	Off Targets	Genic off targets	Sequence 5'-3'	Quality	Off Targets	Genic off targets	
1	GGCATGCTTTTCTTGATTAAAGG	High	290	19	ACTTGCAATCCTAGCACTTTGGG	High	228	6	16
2	GGCATGCTTTTCTTGATTAAAGG	High	290	19	TACTTGCAATCCTAGCACTTTGG	High	180	12	15
3	CAAGTACGAACCACCACACCTGG	High	130	6	TGGCAGAATCACCTGAGGCCAGG	High	278	32	27
4	CAAGTACGAACCACCACACCTGG	High	130	6	CAAGGTGGCAGAATCACCTGAGG	High	267	34	22
5	TCCTGAGTAGCTGAGAACACAG	High	279	17	TAGGAGGATTGCTTGAGCCTGGG	High	188	12	8
6	CTCAGGTGATTCTGCCACCTTGG	High	229	21	CCAGGAGTTCAAGACCAGCCTGG	Medium	316	58	2
7	GGGGTCTCACTAAGCTGTCCAGG	High	194	14	AAAAAATTAGCTGGGTTTGGTGG	Medium	501	19	21
8	TCTCACTAAGCTGTCCAGGCTGG	High	213	24	AAAAAATTAGCTGGGTTTGGTGG	Medium	501	19	17
9	GGCATGCTTTTCTTGATTAAAGG	High	290	19	TGCAATCCTAGCACTTTGGGAGG	Medium	238	9	19

Appendix Table 4.2. Cloning Oligonucleotide designs for sgRNA target sites, incorporating BpiI compatible sticky ends and additional guanine base.

Name	Target 5'-3'	Oligonucleotide Sense Strand 5'-3'	Oligonucleotide Antisense Strand 5'-3'	Cut site to mutation distance (bp)
Nickase1-right	GGCATGCTTTTCTTGATTAAAGG	CACCGGGCATGCTTTTCTTGATTAA	AAACTTAATCAAGAAAAGCATGCCC	124
Nickase1-left	ACTTGCAATCCTAGCACTTTGGG	CACCGACTTGCAATCCTAGCACTTT	AAACAAAGTGCTAGGATTGCAAGTC	74
Nickase2-right	GGCATGCTTTTCTTGATTAAAGG	CACCGGGCATGCTTTTCTTGATTAA	AAACTTAATCAAGAAAAGCATGCCC	124
Nickase2-left	TACTTGCAATCCTAGCACTTTGG	CACCGTACTTGCAATCCTAGCACTT	AAACAAGTGCTAGGATTGCAAGTAC	75
Nickase3-right	CAAGTACGAACCACCACACCTGG	CACCGCAAGTACGAACCACCACACC	AAACGGTGTGGTGGTTCGTACTTGC	103
Nickase3-left	TGGCAGAATCACCTGAGGCCAGG	CACCGTGGCAGAATCACCTGAGGCC	AAACGGCCTCAGGTGATTCTGCCAC	42
Nickase4-right	CAAGTACGAACCACCACACCTGG	CACCGCAAGTACGAACCACCACACC	AAACGGTGTGGTGGTTCGTACTTGC	103
Nickase4-left	CAAGGTGGCAGAATCACCTGAGG	CACCGCAAGGTGGCAGAATCACCTG	AAACCAGGTGATTCTGCCACCTTGC	47
Nickase5-right	TCCTGAGTAGCTGAGAACACAGG	CACCGTCCTGAGTAGCTGAGAACAC	AAACGTGTTCTCAGCTACTCAGGAC	56
Nickase5-	TAGGAGGATTGCTTGAGCCTGGG	CACCGTAGGAGGATTGCTTGAGCCT	AAACAGGCTCAAGCAATCCTCCTAC	98

335

left				
Nickase6-right	CTCAGGTGATTCTGCCACCTTGG	CACCGCTCAGGTGATTCTGCCACCT	AAACAGGTGGCAGAATCACCTGAGC	59
Nickase6-left	CCAGGAGTTCAAGACCAGCCTGG	CACCGCCAGGAGTTCAAGACCAGCC	AAACGGCTGGTCTTGAACTCCTGGC	24
Nickase7-right	GGGGTCTCACTAAGCTGTCCAGG	CACCGGGGGTCTCACTAAGCTGTCC	AAACGGACAGCTTAGTGAGACCCCC	17
Nickase7-left	AAAAAATTAGCTGGGTTTGGTGG	CACCGAAAAAATTAGCTGGGTTTGG	AAACCCAAACCCAGCTAATTTTTTC	38
Nickase8-right	TCTCACTAAGCTGTCCAGGCTGG	CACCGTCTCACTAAGCTGTCCAGGC	AAACGCCTGGACAGCTTAGTGAGAC	24
Nickase8-left	AAAAAATTAGCTGGGTTTGGTGG	CACCGAAAAAATTAGCTGGGTTTGG	AAACCCAAACCCAGCTAATTTTTTC	38
Nickase9-right	GGCATGCTTTTCTTGATTAAAGG	CACCGGGCATGCTTTTCTTGATTAA	AAACTTAATCAAGAAAAGCATGCCC	124
Nickase9-left	TGCAATCCTAGCACTTTGGGAGG	CACCGTGCAATCCTAGCACTTTGGG	AAACCCCAAAGTGCTAGGATTGCAC	71